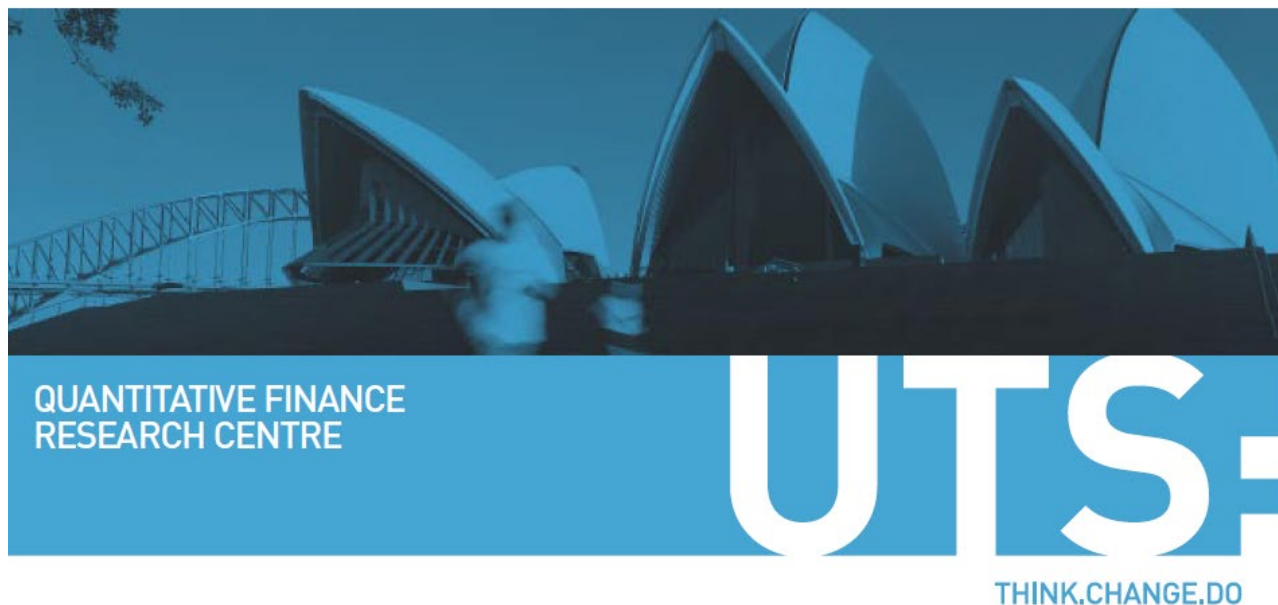


QUANTITATIVE FINANCE
RESEARCH CENTRE



UNIVERSITY OF
TECHNOLOGY SYDNEY



QUANTITATIVE FINANCE RESEARCH CENTRE

Research Paper 422

January 2021

A Computational Approach to Sequential Decision Optimization in Energy Storage and Trading

Paolo Falbo, Juri Hinz, Piyachat Leelasilapasart and Cristian Pelizzari

ISSN 1441-8010

www.qfrc.uts.edu.au

A computational approach to sequential decision optimization in energy storage and trading

Paolo Falbo^b, Juri Hinz^a, Piyachat Leelasilapasart^a, Cristian Pelizzari^b

^a*School of Mathematical and Physical Sciences - University of Technology Sydney*

^b*Department of Economics and Management - University of Brescia*

Abstract

Due to recent technical progress, using battery energy storages are becoming a viable option in the power sector. Their optimal operational management focuses on load shift and shaving of price spikes. However, this requires optimally responding to electricity demand, intermittent generation, and volatile electricity prices. More importantly, such optimization must take into account the so-called deep discharge costs, which have a significant impact on battery lifespan. We present a solution to a class of stochastic optimal control problems associated with these applications. Our numerical techniques are based on efficient algorithms which deliver a guaranteed accuracy.

Keywords: Approximate dynamic programming, energy trading, optimal control, power sector

1. Introduction

In many countries, Battery Energy Storage Systems (BESS) are becoming popular due to advantages in managing power dispatch, interconnection, and demand. Their growing acceptance is due to their ability to smooth out the intermittent and unreliable nature of Renewable Energy Sources (RES). Although RES penetration has shown an increasing trend with no expected end in the near future, their unreliable energy supply (Breton and Moe, 2009; Dinger, 2011) makes it difficult to incorporate RES into a modern electricity grid. However, in some niche applications a variety of BESS are already installed, where they provide operational efficiency and reduce costs by exploiting synergies between storage and renewables (Barton and Infield, 2004; Black and Strbac, 2007; Teleke et al., 2010; Kim and Powell, 2011). Nonetheless, the implementation of BESS is still not straightforward from economical and technical perspectives because of BESS costs and their sensitivity to deep discharge. In this domain,

Email addresses: `paolo.falbo@unibs.it` (Paolo Falbo), `Juri.Hinz@uts.edu.au` (Juri Hinz), `Piyachat.Leelasilapasart-1@student.uts.edu.au` (Piyachat Leelasilapasart), `cristian.pelizzari@unibs.it` (Cristian Pelizzari)

there is a large body of literature with diverse contributions to economic and technological aspects of BESS management. This literature has been reviewed in the excellent paper of Weitzel and Glock (2018). Furthermore, we emphasize Yang et al. (2014) and Kempener and Borden (2015), investigating the role of batteries with respect to RES, and Lu et al. (2014), on the optimal use of BESS for the so-called peak load shaving.

Let us describe an abstract, but typical framework for BESS application. The traditional electricity market players satisfy consumers' energy demand by purchasing electricity in advance, usually taking positions in advance at the *long-term market*. This market typically represented by the so-called day-ahead market for hourly delivery on the next day. However, depending on the situation long-term market may stand for any energy delivery agreements purchased prior to the delivery period. On the contrary, the imbalances during the delivery period must be compensated, as they occur, at a *short-term market*. Such real-time energy balancing can either be achieved through complex over-the-counter trading or, more realistically by participating in real-time energy auction, or by transferring supply from or to electricity grid at the so-called real time grid prices.

Figure 1 provides a simplified illustration of this optimal problem. However, in presence of storage and renewable generation facilities, the problem changes. Within this framework, the agent now os required to simultaneously take positions and setting optimum energy storage levels as shown in Figure 2. The decision optimization problem becomes significantly more complex due to the uncertainty stemming from the future battery capacity levels, electricity prices, and output of renewable energy.

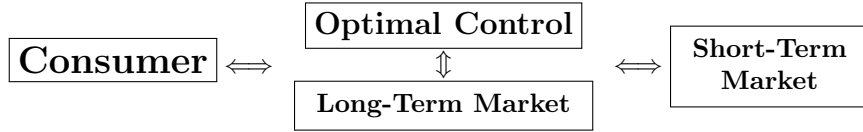


Figure 1: Traditional energy dispatch.

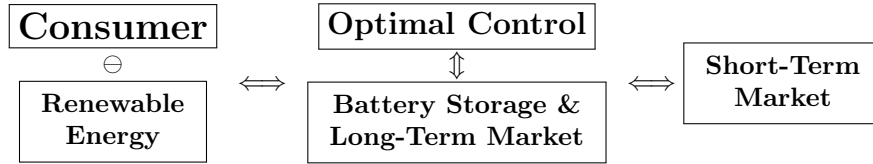


Figure 2: Energy dispatch in the presence of renewable energy and battery storage.

Typically, renewable energy sources such as wind and solar are notoriously

intermittent and unreliable. The potential of energy storage devices to address the highly erratic nature of renewable energy generation Breton and Moe (2009); Dinçer (2011) and energy demand has been discussed extensively in the literature (see Beaudin et al. (2010); Daz-Gonzalez et al. (2012); Evans et al. (2012); Kempener and Borden (2015); Yang et al. (2014)). Their incorporation into a modern energy grid will encourage more environmentally friendly policies which will also have significant impact on investor attitudes towards firms Ramiah et al. (2013); Chan and Walter (2014); Renneboog et al. (2008). The authors of Lu et al. (2014) studied the possible usage of battery storage systems to defer costly modifications to the energy grid by addressing peak loads in the power grid. An extensive recent review of available energy storage technologies has been given by Luo et al. (2015).

While there exist numerous types of energy storage systems, Beaudin et al. (2010) found that no single storage system consistently outperforms all the others for all types of renewable energy sources and applications. So for the sake of simplicity, this paper will assume that the energy retailer pictured in Figure 2 uses a battery device for storing energy. However, our methods and results can easily be extended to other types of storage technologies or even to the use of multiple types of storage devices. From a real options analysis point of view, the incorporation of energy storage devices into energy grid also poses interesting investment questions. The work by Bakke et al. (2016); Bradbury et al. (2014); Locatelli et al. (2016) examines the profitability of investing in energy storage devices. However, Schachter and Mancarella (2016) question the suitability of the current real options approach, stating that the risk neutrality assumption may not be appropriate for risk averse investors. The introduction of batteries also gives rise to important optimal stochastic control problems. The optimal dynamic storage and discharge of energy from battery devices has been examined in Dokuchaev (2016); Kim and Powell (2011); Oudalov et al. (2007); Teleke et al. (2010).

The integration of electric energy storage systems yields challenging problems of optimal stochastic control type. Namely, while the traditional power generation considers as a sequence of independent decisions, the opportunity of storing electricity obviously intertwines actions made at different times: One realizes that storage facilities can be charged during the base load low-price hours, while discharging takes place during peak load high-price hours. However, a sound quantitative solution of such a problem requires sophisticated mathematical techniques, i.e. dynamic programming (Löhdorf and Minner, 2010; AzRISE, 2010), or calculus of variations (Flatley et al., 2016). Apart from rare cases where a dynamic programming problem can be solved analytically (Pham, 2009; Bäuerle and Rieder, 2011), these problems are usually based on appropriate discretization techniques and are addressed in terms of approximate dynamic programming methods (Powell, 2007). Although a huge variety of computational methods have been developed in this area, typical real-world problems are usually too complex for existing solution techniques. Usually, the so-called curse of dimensionality emerges especially when the number of decision factors (state dimension) is high. The present contribution extends the existing

state of the art in several aspects.

The first aspect is concerned with modeling of the complex financial and economic environment which electricity retailers are routinely confronted with. To this purpose, a realistic model of the interaction between energy trading and BESS management is suggested. As mentioned above, we consider energy trading on two time scales: the long-term (realized as day-ahead trading, futures or forward contracts) and short-term (devised to adjust to unexpected changes in the electricity demand in front of delivery, usually realized by the so-called intra-day trading, or by system balancing procedures). This structure is typical for all energy markets, with differences in price dynamics, liquidity, and spreads. The distinction between the short-term and the long-term energy trading is of primary importance to correctly evaluate the economic performance of a BESS, since the real-time re-balancing is essential in electricity business. The present work extends that of Hinz and Yee (2018a), who address BESS management within an over-simplistic setting. Namely, Hinz and Yee (2018a) merely models a battery installation as a passive buffer, which is only used to settle an energy imbalance against the electricity grid. In difference to this, the present work makes the realistic assumption that the battery level is controlled jointly with energy trading. In the present approach, the costs of deep discharge play a central role in the sequential optimization decision, which notably extends the current state of the art in this area.

The second aspect is methodological: we attempt solving a *class* of battery management problems rather than a specific problem. More precisely, we present here a computational methodology, whose routines implement stochastic switching algorithms in modern scientific language Julia¹ and make the entire code publicly available. Our approach realizes a highly customizable solution. Namely, the entire source code of our computations is available on Github and comprises several blocks which serve place holders and can be tailored to specific situations. For instance, the state space model for electricity prices (including seasonal and mean-reverting components) are not attempted to describe any typical electricity price pattern, but to serve a proxy which can be replaced, modified, and adjusted. Such flexibility is ensured by our general assumptions of linear state dynamics which encompass any ARMA model combined with appropriate seasonal and trend components. The same considerations apply to the modeling of deep discharge costs. Since there are diverse approaches to assess (in economic terms) how the battery life is affected by visiting a deep discharge state, we set up a simple generic function which penalizes any deep discharge leaving it to the user to describe the effects more specifically by reflecting a particular battery chemistry.

The third aspect is about a novel computation technique. In comparison to existing schemes (for instance Löhndorf and Minner (2010)), we optimize energy storage management using a combination of primal and dual schemes.

¹The commented source code of our computations is available in the GitHub development platform at the web page <https://github.com/LeePiyachat/Duality>.

More specifically, we apply a sub-gradient method introduced by Hinz (2014), Hinz and Yap (2016), and Hinz and Yee (2018a) to obtain, in a first step, an approximate solution of our stochastic optimal control problems whose numerical quality is examined, in the second step, using a dual method (based on pathwise dynamic programming, see Hinz and Yap, 2016; Hinz et al., 2018). That is, we provide simultaneous optimization of battery control and long-term trading in presence of the uncertainty resulting from RES generation, demand, and electricity price with *guaranteed precision*, which contrasts this work from all existing contributions, to the best of authors knowledge.

The paper is organized as follows: Section 2 details the model settings. Section 3 briefly reviews the solution technique adopted. Section 4 applies the solution technique to the present decision problem. Section 5 includes an illustrative case study, while Section 6 concludes.

2. Model Settings

In the following, we present an abstract but generic model for an electricity market, where an energy retailer has the obligation to meet the demand of its consumers using a combination of energy from renewable generation, contractual position from long-term market (for instance day-ahead trading) and short-term market (real-time trading, balancing market) as well as battery storage. Since in reality, decisions are made and revised periodically, we propose a dynamic optimization of discrete-time decision-making. The present a framework encompasses all important features of real-world energy trading, although it does not reflect a specific market, in order to illustrate our methodology in the most general setting. Following this approach, our algorithmic solution can be adjusted to serve numerous applications within a wide range of specific market situations.

We assume a given finite-time horizon $\{0, \dots, T\} \in \mathbb{N}$ and agree that the unit time corresponds to the energy delivery interval which can measure hours, days, or weeks, depending on the particular application. At any time $t = 0, \dots, T - 1$, an energy retailer has the obligation to satisfy, within $[t, t + 1]$, the unknown electricity demand of its customers while the retailer's renewable energy sources produce a random electricity amount. To manage the resulting energy shortage or surplus, the retailer must trade electricity optimally in advance (at time t) and decide to charge or discharge the battery. Such a retailer's revenue optimization is a typical sequential decision problem under uncertainty. That is, at any time $t = 0, \dots, T - 1$, an action must be chosen (that action will encompass both, energy trading and BESS management), which accrues some revenues and incurs some costs. However, this action also influences the transition to the subsequent states (next battery levels), which change all future revenues, costs, and decisions. The problems of this type are naturally formulated and solved in terms of the so-called Markov Decision Theory. In the subsequent paragraphs, we formulate our joint energy trading-BESS management optimization problem within this standard framework, which we trim to accordingly to make use of a novel method which relies on specific assumptions.

Let D_{t+1} denote the residual electricity demand (which stands for shortfall if $D_{t+1} < 0$ and excess if $D_{t+1} > 0$) for the delivery period $]t, t+1]$ beginning at time t . We assume that the random variable D_{t+1} is observed when the corresponding delivery period ends. Assume that, at each time $t = 0, \dots, T-1$, the producer can take a position F_t in a long-term market to ensure energy supply (if $F_t > 0$) or outflow (if $F_t < 0$) within $[t, t+1]$. Further consider the variable B_t , which stands for the decision to use the battery energy amount $|B_t|$, where $B_t > 0$ and $B_t < 0$ represent charging and discharging actions, respectively. Here, we agree that the BESS management actions must be decided before the start of the corresponding delivery period, in our model at time t . With these assumptions, the energy to be balanced through the short-term market given by

$$F_t - D_{t+1} - B_t \quad (2.1)$$

To introduce the costs/revenues introduce the random variables

$$\pi_t, \quad \underline{\Pi}_{t+1}, \quad \bar{\Pi}_{t+1}, \quad t = 0, \dots, T, \quad (2.2)$$

which stand for prices of electricity delivered within the time interval $]t, t+1]$. Thereby, we assume that π_t is the long-term market price (which is observed and paid at t), whereas $\underline{\Pi}_{t+1}$ and $\bar{\Pi}_{t+1}$ represent short term market prices where $\underline{\Pi}_{t+1}$ applies for procurement and $\bar{\Pi}_{t+1}$ for purchase of energy. Note that both short-term market prices $\underline{\Pi}_{t+1}$, $\bar{\Pi}_{t+1}$ are not observable at the time t become known at $t+1$ at the end of period $]t, t+1]$.

The relationships among prices of electricity delivered on different time scales and traded at different market places has been the topic of a lively discussion since the beginning of deregulation in electricity markets. For an overview of a detailed equilibrium analysis on this topic, we refer the interested reader to Hinz (2003) and the literature cited therein. For instance, it turns out that the connection between long-term (day-ahead) and the short-term (balancing auction, real-time) prices heavily depends on risk aversion, variable production costs, production capacities, and delivery commitments of all market participants.

Finally let us take into account operational costs associated C_t associated with battery management. Obviously, C_t must be modeled by a random variable which depends on the current battery level and on the energy amount delivered/absorbed by the battery within $]t, t+1]$. Let us postpone the specification of these costs and finalize a one-period revenue from BESS management as

$$R_t = F_t \pi_t - (F_t - B_t - D_{t+1})^- \bar{\Pi}_{t+1} + (F_t - B_t - D_{t+1})^+ \underline{\Pi}_{t+1} - C_t$$

for $t = 0, 1, \dots, T-1$. In order to maximize the expectation of the total revenue

$$R = \sum_{t=0}^{T-1} R_t \quad (2.3)$$

a dynamic stochastic control problem must be solved, since at any time t the decision to charge/discharge battery changes the situation at $t + 1$ and has a profound impact on next-period planning. Typically, such sequential decision problems has no closed-form solutions and is computationally challenging. The point here is that at any decision time, the situation is determined by several factors, such as energy prices, battery state and most importantly, energy demand.

It is well-known that the demand fluctuations follow a complex seasonal patterns and are difficult to model, particularly by Markovian processes required for the state dynamics of any sequential decision problem. At this point, we suggest a significant simplification: It turns out that under generic assumptions, the demand modeling can be split off from the strategy optimization. We show that merely one-step prediction d_t (conditional expectation on the most recent information) of the energy demand D_{t+1} is relevant. That is, the problem addressed in this work separates into two separate steps:

- i) Establishing a time-series model for the dynamics of the energy demand which serves at any time t the conditional expectation d_t of the demand D_{t+1} which occurs within $]t, t + 1]$.
- ii) Solving stochastic dynamic control problem, where the demand prediction d_t is not a decision variable at time t because optimal long-term position is calculated as deviation from demand prediction.
- iii) Running the optimal policy, for which the demand prediction must be available to enter optimal long-term positions.

Note that the second step is disentangled in the sense that for ii), the dynamics of the demand prediction is irrelevant. On this account, we consider only step ii) in the reminder of this work. However, notice that for strategy implementation in step iii), the demand prediction must be available at any decision time. Still, since its dynamics is not relevant, the user can alter or replace the entire demand prediction model without re-calculating the sequential decision strategy.

Let us establish such an approach and make some assumptions to obtain a model which is treatable by our numerical methodology. For numerical reasons, we suppose that there is finite set \mathbf{A} of all possible actions. At any time, the controller choses an action via an (optimal) decision rule which will be determined later. The actions can be defined in an abstract way, or they can be identified by integers, (or indexed by integer vectors) and their meaning for control is established merely by some functions defined on actions \mathbf{A} . For instance, to model diverse choices of the trade volume in the long-term market a function $f : \mathbf{A} \rightarrow \mathbb{R}$ is used with the following interpretation: Given the prediction d_t of the demand D_t occurring within $]t, t + 1]$, for the action $a_t \in \mathbf{A}$ chosen at time t by the controller, the energy volume traded in advance (long-term market) is

$$F_t = f(a_t) + d_t. \quad (2.4)$$

To some extent, $|f(a_t)|$ is the energy (bought if $f(a_t) > 0$, sold if $f(a_t) < 0$) on the top of the predicted demand d_t and will be referred to as a *safety margin*. The function f must be chosen manually and represents both, the granularity and range for safety margins.

Similarly, the battery management variable B_t is also determined by the action a_t chosen at time t (immediately before the start of delivery period $]t, t+1]$). Here, we again use an appropriate function on \mathbf{A} , but this time the modeling is more complex as the energy absorbed/delivered by the battery must take into account the current battery level and the physical constraints. To detail this, we suggest to discretize battery levels by a finite set \mathbf{P} . Having chosen at time $t = 0, \dots, T-1$ action $a_t \in \mathbf{A}$, we suppose that the current battery level $p_t \in \mathbf{P}$ transforms to the next level $p_{t+1} = \ell(p_t, a_t) \in \mathbf{P}$ in terms of a pre-specified level change function $\ell : \mathbf{P} \times \mathbf{A} \rightarrow \mathbf{P}$ representing technical restrictions of the battery (total capacity, electrical power). For instance, $a_t \mapsto \ell(p_t, a_t)$ can have values above and below p_t at a range which represents one-period charge/discharge power restrictions if p_t is in one of the intermediate battery levels. However, if p_t is the highest (the lowest) level, then $a_t \mapsto \ell(p_t, a_t)$ take only values below (above) p_t . Specifying the function ℓ requires some details of battery technology used, in particular choosing the maximal charge/discharge intensity along with the highest and the lowest (admissible) battery levels.

Notice that with this convention, the energy amount transformed from/to the storage is given by $B_t = b(p_t, a_t)$, where

$$b(p_t, a_t) = \begin{cases} \ell(p_t, a_t) - p_t, & \text{if } \ell(p_t, a_t) - p_t > 0 \\ \kappa \cdot (\ell(p_t, a_t) - p_t), & \text{if } \ell(p_t, a_t) - p_t \leq 0 \end{cases}, \quad (2.5)$$

with the constant $\kappa \in [0, 1]$ standing for battery efficiency. With these assumptions, we express the energy imbalance (2.1) in terms of action $a_t \in \mathbf{A}$ and battery level $p_t \in \mathbf{P}$ using prediction error

$$\varepsilon_{t+1} = D_{t+1} - d_t, \quad t = 0, \dots, T-1$$

as

$$\begin{aligned} F_t - D_{t+1} - B_t &= d_t + f(a_t) - (d_t + \varepsilon_{t+1}) - b(p_t, a_t) \\ &= f(a_t) - b(p_t, a_t) - \varepsilon_{t+1}. \end{aligned} \quad (2.6)$$

That is, an action $a_t \in \mathbf{A}$ not only triggers transition in battery level from p_t to $p_{t+1} = \ell(p_t, a_t)$, but also determines the the energy amount which needs to be balanced against at the short-term market. Having defined the excess and the shortage of the imbalance (2.6) by

$$\begin{aligned} \overline{E}_{t+1}(a_t, p_t) &= (f(a_t) - b(p_t, a_t) - \varepsilon_{t+1})^- \\ \underline{E}_{t+1}(a_t, p_t) &= (f(a_t) - b(p_t, a_t) - \varepsilon_{t+1})^+ \end{aligned}$$

the profit/loss of balancing costs at time $t = 0, \dots, T-1$ are modeled by

$$-\overline{E}_{t+1}(a_t, p_t)\overline{\Pi}_{t+1} + \underline{E}_{t+1}(a_t, p_t)\underline{\Pi}_{t+1}. \quad (2.7)$$

Now, using (2.4), the financial position for action $a_t \in \mathbf{A}$ is

$$F_t \pi_t = d_t \pi_t + f(a_t) \pi_t. \quad (2.8)$$

Finally, let us model the storage costs by

$$C_t = c(p_t, a_t), \quad (2.9)$$

reflecting the dependence on action $a_t \in \mathbf{A}$ and battery level $p_t \in \mathbf{P}$. The details of the costs must be described by an appropriate function

$$c : \mathbf{P} \times \mathbf{A} \rightarrow \mathbb{R}$$

specified in accordance to battery technology.

With the assumptions (2.7), (2.8), and (2.9), the profit/loss associated with action $a_t \in \mathbf{A}$ depends on prices π_t , $\bar{\Pi}_{t+1}$, $\underline{\Pi}_{t+1}$, the demand prediction d_t , and battery level $p_t \in \mathbf{P}$ as

$$R_t = -d_t \pi_t - f(a_t) \pi_t - \bar{E}_{t+1}(a_t, p_t) \bar{\Pi}_{t+1} + \underline{E}_{t+1}(a_t, p_t) \underline{\Pi}_{t+1} - c(p_t, a_t).$$

Observe that the term $d_t \pi_t$ depends neither on the action a_t nor on the battery level p_t . Since this quantity can not be changed by the decision optimization, the action-dependent part of the cumulative reward is modeled in terms of

$$\begin{aligned} \tilde{R}_t &= R_t + d_t \pi_t \\ &= -f(a_t) \pi_t - \bar{E}_{t+1}(a_t, p_t) \bar{\Pi}_{t+1} + \underline{E}_{t+1}(a_t, p_t) \underline{\Pi}_{t+1} - c(p_t, a_t) \end{aligned} \quad (2.10)$$

In what follows we show how to obtain a strategy which simultaneously takes positions on long-term market and controls a battery level to maximize the expectation of the demand-adjusted total revenue

$$\tilde{R} = \sum_{t=0}^{T-1} \tilde{R}_t. \quad (2.11)$$

Remark: Note that (2.11) differs from (2.3) by a random variable $\sum_{t=0}^{T-1} d_t \Pi_t$ which does not depend on the control policy. On this account, any maximizing the expectation of \tilde{R} constitutes a solution to our problem. As mentioned above, such approach avoids tedious modelling of energy demand dynamics. However notice that the results can not be used directly: For strategy implementation, the one-step demand prediction must be available at any time, thus a demand model is required to run the policy.

3. A numerical solution technique

Sequential decision making is usually encompassed by discrete-time stochastic control and is addressed by the *Markov Decision Processes/Dynamic Programming*. This theory provides a variety of methods. However, approaching

analytical solutions may be cumbersome (Powell, 2007; Pham, 2009; Bäuerle and Rieder, 2011) and numerical approximations may often be far more practical. This work will utilize an implementation of fast and accurate algorithms (see Hinz and Yee, 2018b) to address specific control problems, assuming a finite-time horizon, a finite set of actions, convex reward functions, and a state process that follows linear dynamics. Although these assumptions are restrictive, they encompass a large class of practically important control problems and frequently yield approximate solutions with excellent precision and numerical performance. Let us briefly describe this approach. Suppose that state space $\mathbf{X} = \mathbf{P} \times \mathbf{Z}$ is the product of a finite set \mathbf{P} and an open convex set $\mathbf{Z} \subseteq \mathbb{R}^d$. Furthermore, assume that a finite set \mathbf{A} represents all possible actions. Given a finite-time horizon $\{0, 1, \dots, T\} \subset \mathbb{N}$, consider a *fully observable* controlled Markovian process $(X_t)_{t=0}^T := (P_t, Z_t)_{t=0}^T$ that consists of two parts. The discrete component $(P_t)_{t=0}^T$ describes the evolution of a finite-state controlled Markov chain, which takes values in a finite set \mathbf{P} . Further assume that, at any time $t = 0, \dots, T-1$, the controller chooses an action a from \mathbf{A} in order to cause the one-step transition from the mode $p \in \mathbf{P}$ to the mode $p' \in \mathbf{P}$ with probability $\alpha_{p,p'}(a)$, where $(\alpha_{p,p'}(a))_{p,p' \in \mathbf{P}}$ are pre-specified transition probability matrices for all $a \in \mathbf{A}$. Let us now turn to the evolution of the component $(Z_t)_{t=0}^T$ of the state process $(X_t)_{t=0}^T$. Here, we assume an uncontrolled evolution of such a component in the Euclidean space \mathbb{R}^d . The evolution is modeled by the difference equation

$$Z_{t+1} = W_{t+1}Z_t, \quad t = 0, \dots, T-1, \quad (3.1)$$

where $(W_t)_{t=1}^T$ are independent *disturbance matrices*. That is, the transition kernels \mathcal{K}_t^a governing the evolution of our controlled Markovian process $(P_t, Z_t)_{t=0}^T$ from time t to time $t+1$ are given, for each $a \in \mathbf{A}$, by

$$\mathcal{K}_t^a v(p, z) = \sum_{p' \in \mathbf{P}} \alpha_{p,p'}(a) \mathbb{E}(v(p', W_{t+1}z)), \quad p \in \mathbf{P}, z \in \mathbb{R}^d, t = 0, \dots, T-1,$$

which acts on each function $v : \mathbf{P} \times \mathbb{R}^d \rightarrow \mathbb{R}$ where the above expectations are well-defined. If the system is in the state (p, z) , the *rewards* of applying action $a \in \mathbf{A}$ at time $t = 0, \dots, T-1$ are expressed through $r_t(p, z, a)$. Having arrived at time $t = T$ in the state (p, z) , a final *scrap value* $r_T(p, z)$ is collected. Thereby, the reward functions $r_t : \mathbf{P} \times \mathbb{R}^d \times \mathbf{A} \rightarrow \mathbb{R}$, as well as the scrap function

$$r_T : \mathbf{P} \times \mathbb{R}^d \rightarrow \mathbb{R}, \quad (3.2)$$

are exogenously given for $t = 0, \dots, T-1$. At each time $t = 0, \dots, T-1$ the *decision rule* π_t is given by a mapping $\pi_t : \mathbf{P} \times \mathbb{R}^d \rightarrow \mathbf{A}$, prescribing at time t an action $\pi_t(p, z) \in \mathbf{A}$ for a given state $(p, z) \in \mathbf{P} \times \mathbb{R}^d$. Note that, at each time, the decision rule refers to the recent state of the system, representing a *feedback control*. A sequence $\pi = (\pi_t)_{t=0}^{T-1}$ of decision rules is called a *policy*. For each policy, $\pi = (\pi_t)_{t=0}^{T-1}$, the so-called policy value $v_0^\pi(p_0, z_0)$ is defined as the total expected reward

$$v_0^\pi(p_0, z_0) = \mathbb{E}^{(p_0, z_0), \pi} \left[\sum_{t=0}^{T-1} r_t(P_t, Z_t, \pi_t(P_t, Z_t)) + r_T(P_T, Z_T) \right].$$

In this formula, $\mathbb{E}^{(p_0, z_0), \pi}$ stands for the expectation with respect to the probability distribution of $(P_t, Z_t)_{t=0}^T$ defined by Markov transitions from (P_t, Z_t) to (P_{t+1}, Z_{t+1}) that are induced by the kernels $\mathcal{K}_t^{\pi_t(P_t, Z_t)}$ for $t = 0, \dots, T-1$, started at the initial point $(P_0, Z_0) = (p_0, z_0)$.

Now we turn to the optimization goal. A policy $\pi^* = (\pi_t^*)_{t=0}^{T-1}$ is called optimal if it maximizes the total expected reward over all policies $\pi \mapsto v_0^\pi(p, z)$. To obtain such a policy, one introduces, for $t = 0, \dots, T-1$, the so-called *Bellman operator*

$$\mathcal{T}_t v(p, z) = \max_{a \in \mathbf{A}} \left[r_t(p, z, a) + \sum_{p' \in \mathbf{P}} \alpha_{p, p'}(a) \mathbb{E}[v(p', W_{t+1} z)] \right], \quad (3.3)$$

for $(p, z) \in \mathbf{P} \times \mathbb{R}^d$, acting on all functions v where the stochastic kernel is well-defined. Consider the *Bellman recursion*, also referred to as backward induction:

$$v_T = r_T, \quad v_t = \mathcal{T}_t v_{t+1} \quad \text{for } t = T-1, \dots, 0. \quad (3.4)$$

Assuming that the reward functions are convex and globally Lipschitz and the disturbance matrices $(W_t)_{t=1}^T$ are integrable, there exists a solution $(v_t^*)_{t=0}^{T-1}$ to the Bellman recursion. The functions $(v_t^*)_{t=0}^{T-1}$ are called *value functions* and they determine an optimal policy $\pi^* = (\pi_t^*)_{t=0}^{T-1}$ via

$$\pi_t^*(p, z) = \arg \max_{a \in \mathbf{A}} \left[r_t(p, z, a) + \sum_{p' \in \mathbf{P}} \alpha_{p, p'}(a) \mathbb{E}[v_{t+1}^*(p', W_{t+1} z)] \right], \quad (3.5)$$

for $t = T-1, \dots, 0$ and $v_T^* = r_T$. In order to approximately solve the above Markov Decision problem, one needs to approximate the true value functions $(v_t^*)_{t=0}^{T-1}$ and the corresponding optimal policies $\pi^* = (\pi_t^*)_{t=0}^{T-1}$. Since the reward and scrap functions are convex in the continuous variable, the value functions are also convex and can be approximated by piecewise linear and convex functions. For this, introduce the so-called sub-gradient envelope $\mathcal{S}_{\mathbf{G}^m} f$ of a convex function $f : \mathbf{Z} \rightarrow \mathbb{R}$ on a grid $\mathbf{G}^m \subset \mathbf{Z}$ with m points:

$$(\mathcal{S}_{\mathbf{G}^m} f)(z) = \max_{g \in \mathbf{G}^m} (\nabla_g f)(z),$$

for $z \in \mathbf{Z}$, where $\nabla_g f$ is the tangent of f at grid point $g \in \mathbf{G}^m$. Using the sub-gradient envelope operator, define the double-modified Bellman operator as

$$\mathcal{T}_t^{m, n} v(p, z) = \mathcal{S}_{\mathbf{G}^m} \max_{a \in \mathbf{A}} \left(r_t(p, z, a) + \sum_{p' \in \mathbf{P}} \alpha_{p, p'}(a) \sum_{k=1}^n \nu_{t+1}^{(k)} v(p', W_{t+1}^{(k)} z) \right),$$

where the probability weights $(\nu_{t+1}^{(k)})_{k=1}^n$ correspond to the distribution sampling $(W_{t+1}^{(k)})_{k=1}^n$ of each disturbance matrix W_{t+1} . The corresponding backward induction

$$v_T^{m, n}(p, z) = \mathcal{S}_{\mathbf{G}^m} r_T(p, z), \quad (3.6)$$

$$v_t^{m, n}(p, z) = \mathcal{T}_t^{m, n} v_{t+1}^{m, n}(p, z), \quad t = T-1, \dots, 0, \quad (3.7)$$

yields the so-called double-modified value functions $(v_t^{m,n})_{t=0}^T$. Under appropriate assumptions regarding grid density and disturbance sampling, the double-modified value functions converge uniformly to the true value functions in (3.4) on compact sets (see Hinz, 2014). To gauge the quality of the above approximations, one determines two random variables whose expectations bound the true value function, i.e.

$$\mathbb{E}(v_0(p_0, z_0)) \leq v_0(p_0, z_0) \leq \mathbb{E}(\bar{v}_0(p_0, z_0)), \quad p_0 \in \mathbf{P}, z_0 \in \mathbf{Z}. \quad (3.8)$$

This technique is usually referred to as *pathwise stochastic control* and has gained increasing popularity over the recent decades. We refer the interested reader to Hinz and Yap (2016) and the literature cited therein for the technical details. Such a stochastic control exhibits a helpful *self-tuning* property. The closer the value function approximations resemble their true unknown counterparts, the tighter the bounds in (3.8) and the lower the standard errors of the bound estimates. The following sections provide an application of this technique to the above battery control problem.

Remark: In applications, sequential decision problems frequently appear in a slightly different formulation than given above. Usually, the costs of control depend on both, the recent and the next state. That is, instead of a previously introduced reward $r_t(P_t, Z_t, a)$ for taking action a in the situation (P_t, Z_t) , a modeling may naturally suggest $\tilde{r}_t(P_t, Z_t, P_{t+1}, Z_{t+1}, a)$ where the action a taken at time t in the situation (P_t, Z_t) but reward is observed and returned at $t + 1$ with a random outcome depending on the next-time situation (P_{t+1}, Z_{t+1}) . Fortunately, this context is seamlessly covered by the formal setting introduced above. It turns out that since the expectation of cumulative reward is being maximized, a pre-conditioning $\tilde{r}_t(P_t, Z_t, P_{t+1}, Z_{t+1}, a)$ on the information available at time t can be applied. That is, having determined the control rewards as being next-state dependent

$$\tilde{r}_t : \mathbf{P} \times \mathbb{R}^d \times \mathbf{P} \times \mathbb{R}^d \times \mathbf{A} \rightarrow \mathbb{R} \quad t = 0, \dots, T-1 \quad (3.9)$$

for each $p \in \mathbf{P}$, $z \in \mathbb{R}^d$ and $t = 0, \dots, T-1$ one averages them

$$r_t(p, z, a) = \sum_{p' \in \mathbf{P}} \alpha_{p,p'}(a) \mathbb{E}(\tilde{r}_t(p, z, p', W_{t+1}z, a)), \quad (3.10)$$

to obtain reward functions as introduced in the standard setting (3.2).

4. BESS as Stochastic Switching with Linear State Dynamics

Let us construct a model that fulfills all the assumptions of Section 2 such that the methodology presented in Section 3 becomes applicable. For this, we introduce the four-dimensional uncontrolled state evolution $(Z_t)_{t=0}^T$

$$Z_t = [1, \quad Z_t^{(2)}, \quad Z_t^{(3)}, \quad Z_t^{(4)}, \quad Z_t^{(5)}]^\top, \quad t = 0, \dots, T$$

which carries constant entry one its first component. This is a minor increase of state dimension, allowing to encompass a broad class of dynamics while fulfilling linear (3.1) restriction. Let us agree that the processes $(\pi_t)_{t=0}^T$, $(\underline{\Pi}_t)_{t=0}^T$, $(\overline{\Pi}_t)_{t=0}^T$, and $(\varepsilon_t)_{t=0}^T$, are functions on the components of $(Z_t)_{t=0}^T$ as follows:

$$\begin{aligned} \pi_t &= g^{(2)}(t, Z_t^{(2)}), & \underline{\Pi}_t &= g^{(3)}(t, Z_t^{(3)}), \\ \overline{\Pi}_t &= g^{(4)}(t, Z_t^{(4)}), & \varepsilon_t &= g^{(5)}(t, Z_t^{(5)}) \end{aligned} \quad (4.1)$$

where the deterministic affine-linear transformations $(g^{(i)}(t, \cdot))_{t=0}^T$ for $i = 2, \dots, 5$ appropriately describe trends and seasonal patterns.

For a numerical case study, let us address the above framework more specifically. First, we suggest modeling the long-term price component as a function of an auto-regressive process. Therefore, consider a sequence $(N_{t+1}^{(2)})_{t=0}^T$ of independent and standard normal random variables and introduce the auto-regressive state process $(Z_t^{(2)})_{t=0}^T$ such that

$$Z_{t+1}^{(2)} = \mu + \phi Z_t^{(2)} + \sigma N_{t+1}^{(2)}, \quad Z_0^{(2)} = z_0^{(2)} \in \mathbb{R}, \quad (4.2)$$

with parameters $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}_+$, and $\phi \in [0, 1]$. To embed the evolution (4.2) into state process $(Z_t)_{t=0}^T$ recall that the first component is equal to one for $t = 0, \dots, T$, which allows the desired linear dynamics

$$\begin{bmatrix} 1 \\ Z_{t+1}^{(2)} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \mu + \sigma N_{t+1}^{(2)} & \phi \end{bmatrix} \begin{bmatrix} 1 \\ Z_t^{(2)} \end{bmatrix} \quad t = 0, \dots, T. \quad (4.3)$$

Other components can be modeled similarly, as time dependent affine-linear functions of auto-regressions. For simplicity, we suggest independent identically distributed random variables

$$Z_t^{(3)} = N_t^{(3)}, \quad Z_t^{(4)} = N_t^{(4)}, \quad Z_t^{(5)} = N_t^{(5)}, \quad t = 0, \dots, T \quad (4.4)$$

which yields a linear state dynamics (3.1) with the following disturbance matrices

$$W_t = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \mu + \sigma N_t^{(2)} & \phi & 0 & 0 & 0 \\ 0 & 0 & N_t^{(3)} & 0 & 0 \\ 0 & 0 & 0 & N_t^{(4)} & 0 \\ 0 & 0 & 0 & 0 & N_t^{(5)} \end{bmatrix} \quad t = 0, \dots, T.$$

Here $(N_t = (N_{i=2}^{(i)})_{i=2}^5)_{t=0}^T$ is a sequence of independent standard normally distributed random variables. To describe the dynamics (4.1), the state variables must be scaled and shifted appropriately. The seasonality is reflected by functions

$$g^{(i)}(t, z) = u_t^{(i)} + s_t^{(i)} z^{(i)} \quad z^{(i)} \in \mathbb{R}, \quad t = 0, \dots, T, \quad (4.5)$$

with deterministic shift $u_t^{(i)} \in \mathbb{R}$ and scale $s_t^{(i)} \in]0, \infty[$ coefficients, $i = 2, \dots, 5$.

To describe the evolution of the controlled part $(P_t)_{t=0}^T$ of the state dynamics, we assume that the finite set \mathbf{P} includes battery levels, which are equidistantly spaced with step size $\Delta > 0$ between $p = \min \mathbf{P}$ and $\bar{p} = \max \mathbf{P}$ levels. Given the level change function $\ell : \mathbf{P} \times \mathbf{A} \rightarrow \mathbf{P}$, the transitions are non-random

$$\alpha_{p,p'}(a) = \begin{cases} 1 & \text{if } p' = \ell(p, a), \\ 0 & \text{else,} \end{cases} \quad p \in \mathbf{P}, a \in \mathbf{A}. \quad (4.6)$$

Having defined the state evolution $(Z_t, P_t)_{t=0}^T$ and the processes (4.1) via functions (4.5) on states, observe that the rewards (2.10) depend on the current and on the next state, as in (3.9)

$$-f(a)\pi_t - \bar{E}_{t+1}(a, p)\bar{\Pi}_{t+1} + \underline{E}_{t+1}(a, p)\underline{\Pi}_{t+1} - c(p, a) \quad (4.7)$$

since due to (4.1), π_t is a function of Z_t while $\bar{E}_{t+1}(a, p)$, $\bar{\Pi}_{t+1}$, $\underline{E}_{t+1}(a, p)$, $\underline{\Pi}_{t+1}$ are functions of Z_{t+1} . Using (3.10), we transform (4.7) to the standard form of the reward

$$-f(a)\pi_t - \bar{e}_t(a, p)\bar{\pi}_t + \underline{e}_t(a, p)\underline{\pi}_t - c(p, a). \quad (4.8)$$

In this equation, the expected surplus $\underline{e}_t(a, p)$ and shortage $\bar{e}_t(a, p)$ of the imbalance are obtained as integrals

$$\underline{e}_t(p, a) = \int_0^\infty x \mathcal{N}(f(a) - b(p, a) - u_{t+1}^{(5)}, (s_{t+1}^{(5)})^2)(dx) \quad (4.9)$$

and

$$\bar{e}_t(p, a) = \int_{-\infty}^0 (-x) \mathcal{N}(f(a) - b(p, a) - u_{t+1}^{(5)}, (s_{t+1}^{(5)})^2)(dx) \quad (4.10)$$

respectively, for all $p \in \mathbf{P}$ and $a \in \mathbf{A}$, where $\mathcal{N}(\xi, \varsigma^2)$ denotes the normal distribution with mean $\xi \in \mathbb{R}$ and variance $\varsigma^2 \in \mathbb{R}_+$. Furthermore $\underline{\pi}_t$ and $\bar{\pi}_t$ are the expectations of $\underline{\Pi}_{t+1}$ and $\bar{\Pi}_{t+1}$ at time t , given by

$$\underline{\pi}_t = u_t^{(3)} \quad \bar{\pi}_t = u_t^{(4)}, \quad t = 0, \dots, T. \quad (4.11)$$

With all ingredients now in place, we to define the reward functions in accordance to (2.10) by

$$r_t(p, (z^{(1)}, \dots, z^{(5)}), a) = -f(a)(u_t^{(2)} + s_t^{(2)}z^{(2)}) - \bar{e}_t(p, a)u_t^{(4)} + \underline{e}_t(p, a)u_t^{(3)} - c(p_t, a_t), \quad (4.12)$$

for all $a \in \mathbf{A}$, $p \in \mathbf{P}$, $(z_t^{(1)}, \dots, z_t^{(5)}) \in \mathbb{R}^5$, and $t = 0, \dots, T-1$.

Finally, let us introduce the last ingredient – the scrap function. Here, we assume that the entire electricity from the BESS can be sold at the long-term market at time T :

$$r_T(p_T, (z^{(1)}, \dots, z^{(5)})) = p\pi_T = p(u_T^{(2)} + s_T^{(2)}z^{(2)}), \quad (4.13)$$

for $p \in \mathbf{P}, (z^{(1)}, \dots, z_T^{(5)}) \in \mathbb{R}^5$. With these definitions, we have formalized the optimal management of battery energy storage system as a stochastic control problem and can address its numerical solution in the next section.

Remark: Note that the rewards (4.12) and scrap (4.13) functions depend only on only second component $z^{(2)}$ of the state variable $z = (z^{(1)}, \dots, z^{(5)})$. That is, modeling the state evolution can be reduced to first two components using linear dynamics as in (4.3).

5. A Numerical Illustration

Consider a BESS with a total capacity of $\chi \in \mathbb{R}_+$ MWh. We assume that the positions $\mathbf{P} \subset [0, \chi]$ represent a grid of all feasible battery levels ranging from the minimum level $\underline{p} = \min \mathbf{P}$ to the maximum level $\bar{p} = \max \mathbf{P}$. Such discretization of battery levels (which are continuous by their physical nature) is a tribute we have to pay to make our optimal switching approach applicable. However, since the numerical procedures we use are very efficient, the degree of discretization can be realized at a sufficiently fine granularity. Further, assume that the space of actions is the Cartesian product of two finite sub-spaces:

$$\mathbf{A} = \{1, 2, \dots, \bar{a}^{(1)}\} \times \{1, 2, \dots, \bar{a}^{(2)}\},$$

with the interpretation that, having taken action $a = (a^{(1)}, a^{(2)}) \in \mathbf{A}$, the retailer chooses a certain safety margin $f^{(1)}(a^{(1)}) \in \mathbb{R}$ through the first action component $a^{(1)}$ and determines, at the same time, a potential battery charge/discharge $f^{(2)}(a^{(2)})$ via the second action component $a^{(2)}$. In fact, we assume that both sets, that of the safety margins

$$\{f^{(1)}(a^{(1)}) : (a^{(1)}, a^{(2)}) \in \mathbf{A}\}$$

and that of the charge/discharge amounts

$$\{f^{(2)}(a^{(2)}) : (a^{(1)}, a^{(2)}) \in \mathbf{A}\},$$

are represented by discrete grids, ranging from their minimum values $\underline{f}^{(1)}$ and $\underline{f}^{(2)}$ to their maximum values $\bar{f}^{(1)}$ and $\bar{f}^{(2)}$, respectively. In this setting, the safety margin function is given by $f(a^{(1)}, a^{(2)}) = f^{(1)}(a^{(1)})$ and the BESS management variable is defined by

$$\ell(p, (a^{(1)}, a^{(2)})) = \arg \min_{p' \in \mathbf{P}} |p' - (p + f^{(2)}(a^{(2)}))|,$$

for all $p \in \mathbf{P}$ and $(a^{(1)}, a^{(2)}) \in \mathbf{A}$. With this variable, the loss due to battery inefficiency is described as in (2.5).

The state process $(Z_t)_{t=0}^T := (Z_t^{(1)}, Z_t^{(2)})_{t=0}^T$ is modeled as in Section 4 with parameters $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}_+$, and $\phi \in [0, 1]$. To describe the trend and the seasonality in the evolution of the long-term price, we assume that the price is given by

$$\pi_t = g^{(2)}(t, Z_t^{(2)}) = u_t^{(2)} + s_t^{(2)} Z_t^{(2)}, \quad t = 0, \dots, T,$$

T	μ	ϕ	σ	$Z_0^{(2)}$	τ	κ	$ \mathbf{P} $	$ \mathbf{A} $	
48	1	0.9	1	10	24	1	21	13×9	
p	\bar{p}	$\bar{a}^{(1)}$	$\bar{a}^{(2)}$	$\bar{f}^{(1)}$	$\bar{f}^{(2)}$	ζ	χ		
0	150	13	9	-10	10	-10	10	1	150

Table 1: Values of the parameters used in the numerical illustration.

with deterministic coefficients $u_t^{(2)} = -1 + \cos(2\pi t/\tau)$ and $s_t^{(2)} = 1 + \sin^2(2\pi t/\tau)$, where parameter $\tau > 0$ represents the period length. Figure 3 depicts the state process $(Z_t^{(2)})_{t=0}^T$ and the long-term electricity prices in Euros $(g^{(2)}(t, Z_t^{(2)}))_{t=0}^T$ for the parameters listed in Table 1. Further, we obtain the expected short-term prices as

$$\underline{\pi}_t = u_t^{(3)}, \quad \bar{\pi}_t = u_t^{(4)}, \quad t = 0, \dots, T$$

where we have supposed $u_t^{(3)} = 5$ and $u_t^{(4)} = 50$ for all $t = 0, \dots, T$. Finally, we

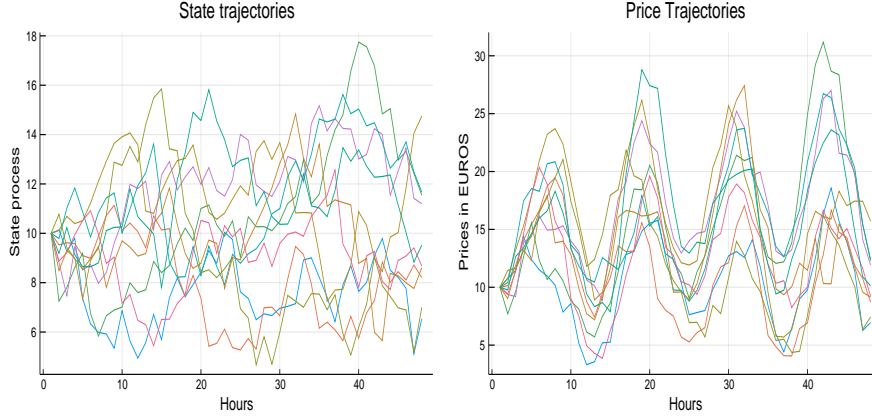


Figure 3: The left and right plots depict sample paths for the state process $(Z_t^{(2)})_{t=0}^T$ and the corresponding long-term electricity prices $(\pi_t)_{t=0}^T$, respectively.

suggest modeling deep discharge costs by the following function:

$$c(p, a) = \eta_1 (1 + \eta_2 p/\chi)^{-1}, \quad p \in \mathbf{P}, a \in \mathbf{A},$$

with parameters $\eta_1 \geq 0, \eta_2 > 0$. Note that this function depends on the ratio p/χ , which measures the depth of discharge. Such function increasingly penalizes the total reward as the battery level approaches zero. To examine the effect of this penalization, we compare optimal battery levels in Figure 4, which depicts level evolutions under the assumption that the battery starts at the lowest level at time $t = 0$. The bottom plot shows that, in the absence of deep discharge costs ($\eta_1 = 0$), low battery levels are reached routinely. On the

contrary, the upper plot shows that battery levels rarely fall below 30% of the total capacity.

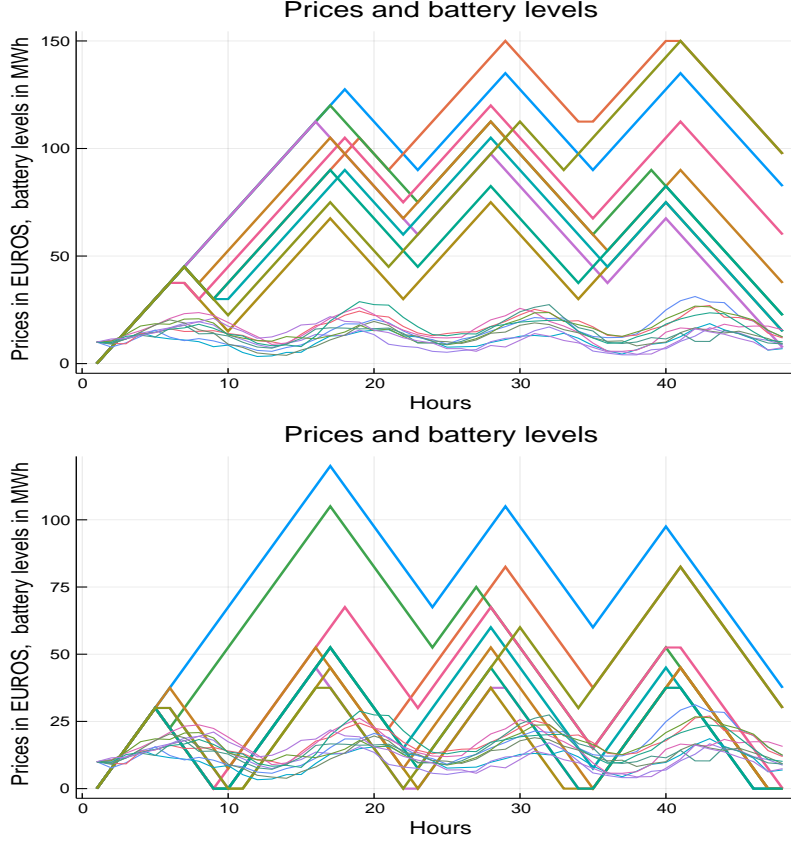


Figure 4: Forward electricity prices and battery levels based on the parameter values in Table 1, with $\eta_1 = 100$, $\eta_2 = 15$ (upper plot) and with $\eta_1 = 0$ (lower plot). The price trajectories, taken from the right panel of Figure 3, are depicted in the lower part of each panel.

Furthermore, let us illustrate the safety margins in Figure 5. Since there is no significant difference between both graphs, deep discharge costs seem to have a moderate impact on safety margins. In both graphs, we merely see a tendency to buy energy through higher safety margins when electricity prices are low.

Finally, we provide a brief discussion on the value function illustrated in Table 2 and in Figure 6. Each row of Table 2 corresponds to a discretized battery level. The columns “Lower interval” and “Upper interval” include the empirical confidence intervals for the lower and upper estimate of the value function, respectively. This calculation was based on the assumption that the initial electricity price, Π_0 , was equal to 10. The confidence bounds, obtained

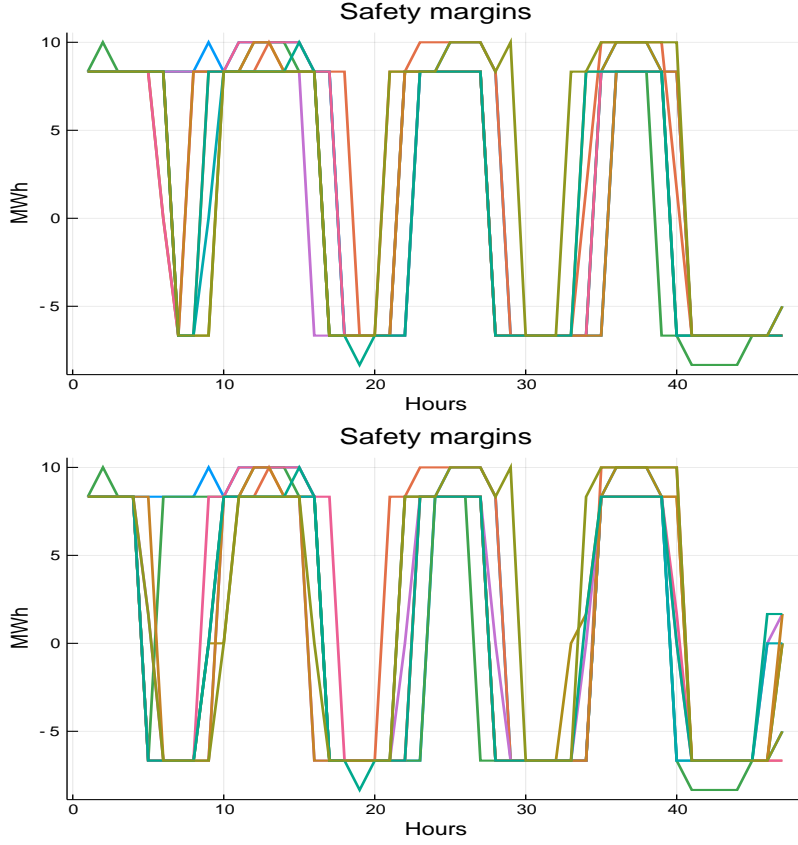


Figure 5: Safety margins based on the parameter values in Table 1, with $\eta_1 = 100$, $\eta_2 = 15$ (upper plot) and with $\eta_1 = 0$ (lower plot). Contrary to Figure 4, the price trajectories, taken from the right panel of Figure 3, are not reported here to avoid a confusing overlapping with the safety margin plots.

by the diagnostic methods in Hinz and Yap (2016) based on a pathwise dynamic approach, are tight, which certifies the high precision of our solution obtained in terms of the sub-gradient method described in Section 3. The approximate value functions delivered by the sub-gradient method are depicted in Figure 6 for a range of values of the initial state variable $Z_0^{(2)}$ (represented on the horizontal axis), with different curves standing for different initial battery levels p_0 . Here, we can notice that the value function is increasing in the initial battery level p_0 (the energy owned at the beginning yields a certain return) and it interacts with the initial state variable, $Z_0^{(2)}$, which is also, by construction, the initial electricity price Π_0 . Recall that a higher price at the beginning causes subsequent prices to be high on average (by the increasing trend of the autoregressive state process). Therefore, if the battery is well charged at time $t = 0$, the retailer can sell electricity (within the time horizon) and obtain a substantial profit; if, on the contrary, the initial level of the battery is low, the retailer must pay more for the initial charge.

Notice that the numerical results for the value function do not allow a direct interpretation in terms of a total revenue. The reason is that the rewards of our model (2.10) do not take into account the retailer's income from fixed delivery contracts (refer to the remark after (2.10)).

Level p (MWh)	Lower interval	Upper interval
0	[-484.8457, -484.7002]	[-484.845728, -484.7002]
8	[-261.5587, -261.4229]	[-261.558277, -261.4219]
15	[-82.4095, -82.2604]	[-82.409453, -82.2604]
23	[78.0581, 78.2014]	[78.058582, 78.2023]
30	[227.8284, 227.9784]	[227.828407, 227.9784]
38	[370.0146, 370.1593]	[370.015052, 370.1603]
45	[506.2079, 506.3552]	[506.207923, 506.3552]
53	[637.4470, 637.6007]	[637.447020, 637.6007]
60	[764.6185, 764.7616]	[764.620788, 764.7631]
68	[888.3149, 888.4576]	[888.314914, 888.4576]
75	[1008.8882, 1009.0220]	[1008.890685, 1009.0233]
83	[1126.5152, 1126.6511]	[1126.515153, 1126.6511]
90	[1241.4532, 1241.5858]	[1241.455821, 1241.5871]
98	[1353.6041, 1353.7311]	[1353.604121, 1353.7311]
105	[1463.2317, 1463.3635]	[1463.234482, 1463.3646]
113	[1570.0781, 1570.2042]	[1570.078093, 1570.2042]
120	[1674.4280, 1674.5548]	[1674.428008, 1674.5548]
128	[1775.6500, 1775.7734]	[1775.649952, 1775.7734]
135	[1868.0094, 1868.1358]	[1868.009369, 1868.1358]
143	[1947.9315, 1948.0580]	[1947.931502, 1948.0580]
150	[2023.2787, 2023.4052]	[2023.278738, 2023.4052]

Table 2: Solution diagnostics based on 50 trajectories (with $\eta_1 = 100$ and $\eta_2 = 15$) and $\Pi_0 = 10$.

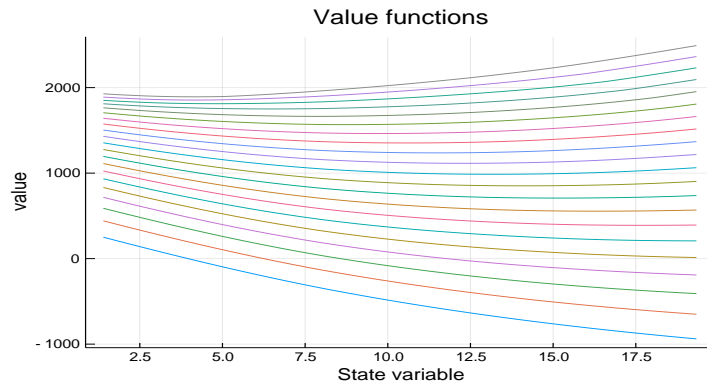


Figure 6: Value functions at $t = 0$ under optimal policies. Each curve corresponds to a battery level $p_0 \in \mathbf{P}$: the lowest curve is associated with the lowest battery level, 0, the second lowest curve is associated with the second lowest battery level, 8, and so on. Each curve is drawn with respect to a range of values for the state variable $Z_0^{(2)}$. The graphs are drawn with costs of deep discharge ($\eta_1 = 100$ and $\eta_2 = 15$).

Finally, we exemplarily elaborate on a typical economic application addressing a stylized investment and capacity allocation problem. In this context, one of the most important questions is to determine the optimal installed capacity and the type of the battery. Having assumed zero costs of deep discharge, Figure 7 depicts the value function, starting with an empty battery, in dependence on storage capacity. Let us refer to this value as the *initial storage value*. In line with intuition, a higher storage yields a higher value represented by a monotonically increasing concave curve. Since the initial investment in BESS is usually linear in the capacity put in place, this curve could be used to determine the optimal capacity by equating the value of the marginal storage to that of the marginal investment.

Our numerical experiments suggest, that dealing with twenty to fifty equidistant levels shall yield sufficiently precise results (i.e. the numerical outcomes do not change significantly if the granularity becomes finer). However, the discretization of actions (which yields only a finite number of safety margins and battery controls) is a delicate issue. Here, it may be advisable to compare numerical outcomes from several models. Still, our experiments suggest that the optimal strategies are of bang-bang type, meaning that they apply just few extremal (usually largest and smallest) safety margins and battery charging/discharging actions. For this reason, we believe that good results are achievable by small action spaces.

6. Conclusions

Battery Energy Storage Systems (BESS) have the potential to change the landscape of future energy generation and trading. However, the existing sys-

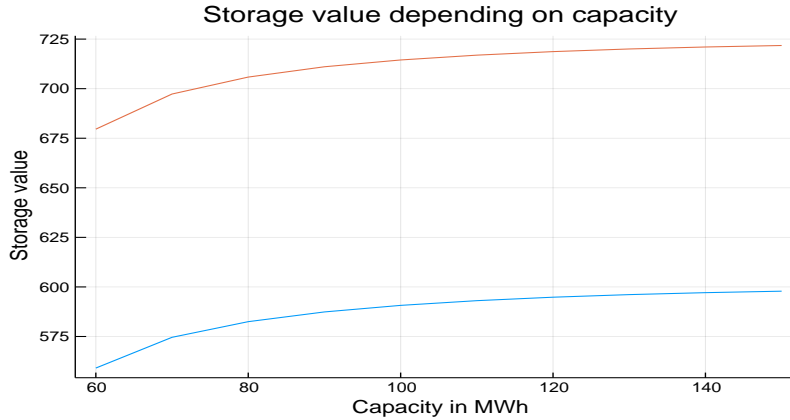


Figure 7: Initial value of an empty battery at $\Pi_0 = 10$ and for the efficiency parameter $\kappa = 0.97$ (lower line) and $\kappa = 1.0$ (upper line).

tems are costly and sensitive to diverse operational issues (such as deep discharge), therefore a thorough investigation of their optimal management is essential. To engage with this development, we investigate the problem of electricity storage management in presence of deep discharge costs within a well defined market structure. In particular, we advance a simultaneous optimization of BESS management and energy trading in presence of the uncertainty resulting from RES generation, demand, and electricity prices. We provide evidence that charging/discharging decisions anticipate nicely changes in electricity prices and avoid, at the same time, deep discharging. Despite the obvious mathematical complications of joining energy trading and BESS management in a stochastic framework, the optimization problem is successfully solved. To this purpose, we adopt for the first time a combination of primal and dual schemes that provide an approximately optimal solution with guaranteed accuracy. Our methodology is based on highly time performing computational schemes, whose routines are publicly available. Our approach is also characterized by flexibility, which allows the general problem to be specialized by adopting arbitrary functions in a more realistic investigation.

References

- AzRISE, August 2010. Study of compressed air energy storage with grid and photovoltaic energy generation. Report, AzRISE, Arizona Research Institute for Solar Energy.
URL <http://www.u.arizona.edu/~sreynold/caes.pdf>
- Bakke, I., Fleten, S., Hagfors, L., Hagspiel, V., Norheim, B., Wogrin, S., 2016. Investment in electric energy storage under uncertainty: a real options approach. *Computational Management Science* 13 (3), 483–500.

- Barton, J. P., Infield, D. G., 2004. Energy storage and its use with intermittent renewable energy. *IEEE Transactions on Energy Conversion* 19 (2), 441–448.
- Bäuerle, N., Rieder, U., 2011. *Markov Decision Processes with Applications to Finance*. Universitext. Springer, Berlin, Germany.
- Beaudin, M., Zareipour, H., Schellenberglabe, A., Rosehart, W., 2010. Energy storage for mitigating the variability of renewable electricity sources: An updated review. *Energy for Sustainable Development* 14 (4), 302–314.
- Black, M., Strbac, G., 2007. Value of bulk energy storage for managing wind power fluctuations. *IEEE Transactions on Energy Conversion* 22 (1), 197–205.
- Bradbury, K., Pratson, L., Patio-Echeverri, D., 2014. Economic viability of energy storage systems based on price arbitrage potential in real-time u.s. electricity markets. *Applied Energy* 114, 512 – 519.
- Breton, S.-P., Moe, G., 2009. Status, plans and technologies for offshore wind turbines in Europe and North America. *Renewable Energy* 34 (3), 646–654.
- Chan, P., Walter, T., 2014. Investment performance of “environmentally-friendly” firms and their initial public offers and seasoned equity offers. *Journal of Banking & Finance* 44, 177 – 188.
- Dinger, F., 2011. The analysis on photovoltaic electricity generation status, potential and policies of the leading countries in solar energy. *Renewable and Sustainable Energy Reviews* 15 (1), 713–720.
- Dokuchaev, N., 2016. Optimal energy storing and selling in continuous time stochastic multi-battery setting.
- Daz-Gonzalez, F., Sumper, A., Gomis-Bellmunt, O., Villaffila-Robles, R., 2012. A review of energy storage technologies for wind power applications. *Renewable and Sustainable Energy Reviews* 16 (4), 2154 – 2171.
- Evans, A., Strezov, V., Evans, T., 2012. Assessment of utility energy storage options for increased renewable energy penetration. *Renewable and Sustainable Energy Reviews* 16 (6), 4141 – 4147.
- Flatley, L. C., MacKay, R. S., Waterson, M., 2016. Optimal strategies for operating energy storage in an arbitrage or smoothing market. *Journal of Dynamics and Games* 3 (4), 371–398.
- Hinz, J., 2003. Modelling day-ahead electricity prices. *Applied Mathematical Finance* 10 (2), 149–161.
- Hinz, J., 2014. Optimal stochastic switching under convexity assumptions. *SIAM Journal on Control and Optimization* 52 (1), 164–188.

- Hinz, J., Tarnopolskaya, T., Yee, J., 2018. Efficient algorithms of pathwise dynamic programming for decision optimization in mining operations. *Annals of Operations Research*.
- Hinz, J., Yap, N., 2016. Algorithms for optimal control of stochastic switching systems. *Theory of Probability & Its Applications* 60 (4), 580–603.
- Hinz, J., Yee, J., 2018a. Optimal forward trading and battery control under renewable electricity generation. *Journal of Banking & Finance* 95, 244–254.
- Hinz, J., Yee, J., 2018b. rcss: R package for optimal convex stochastic switching. *The R Journal* 10 (2), 38–54.
- Kempener, R., Borden, E., January 2015. Battery storage for renewables: Market status and technology outlook. Report, IRENA, International Renewable Energy Agency.
URL <http://www.irena.org/publications/2015/Jan/Battery-Storage-for-Renewables-Market-Status-and-Technology-Outlook>
- Kim, J. H., Powell, W. B., 2011. Optimal energy commitments with storage and intermittent supply. *Operations Research* 59 (6), 1347–1360.
- Locatelli, G., Invernizzi, D., Mancini, M., 2016. Investment and risk appraisal in energy storage systems: A real options approach. *Energies* 104, 114 – 131.
- Löhndorf, N., Minner, S., 2010. Optimal day-ahead trading and storage of renewable energies - An approximate dynamic programming approach. *Energy Systems* 1 (1), 61–77.
- Lu, C., Xu, H., Pan, X., Song, J., 2014. Optimal sizing and control of battery energy storage system for peak load shaving. *Energies* 7 (12), 8396–8410.
- Luo, X., Wang, J., Dooner, M., Clarke, J., 2015. Overview of current development in electrical energy storage technologies and the application potential in power system operation. *Applied Energy* 137, 511 – 536.
- Oudalov, A., Chartouni, D., Ohler, C., 2007. Optimizing a battery energy storage system for primary frequency control. *IEEE Transactions on Power Systems* 22 (3), 1259–1266.
- Pham, H., 2009. Continuous-Time Stochastic Control and Optimization with Financial Applications. *Stochastic Modelling and Applied Probability*. Springer, Berlin, Germany.
- Powell, W. B., 2007. Approximate Dynamic Programming: Solving the Curses of Dimensionality. *Wiley Series in Probability and Statistics*. Wiley, Hoboken, NJ, USA.
- Ramiah, V., Martin, B., Moosa, I., 2013. How does the stock market react to the announcement of green policies? *Journal of Banking & Finance* 37 (5), 1747 – 1758.

- Renneboog, L., Horst, J., Zhang, C., 2008. Socially responsible investments: Institutional aspects, performance, and investor behavior. *Journal of Banking & Finance* 32 (9), 1723 – 1742.
- Schachter, J., Mancarella, P., 2016. A critical review of real options thinking for valuing investment flexibility in smart grids and low carbon energy systems. *Renewable and Sustainable Energy Reviews* 56, 261 – 271.
- Teleke, S., Baran, M. E., Bhattacharya, S., Huang, A. Q., 2010. Optimal control of battery energy storage for wind farm dispatching. *IEEE Transactions on Energy Conversion* 25 (3), 787–794.
- Weitzel, T., Glock, C. H., 2018. Energy management for stationary electric energy storage systems: A systematic literature review. *European Journal of Operational Research* 264 (2), 582–606.
- Yang, Y., Li, H., Aichhorn, A., Zheng, J., Greenleaf, M., 2014. Sizing strategy of distributed battery storage system with high penetration of photovoltaic for voltage regulation and peak load shaving. *IEEE Transactions on Smart Grid* 5 (2), 982–991.