Industrial and
Systems Engineering

Electric Storage with Stacked Services: Control and Valuation

Boris Defourny[1] and Somayeh Moazeni[2]

[1]Department of Industrial and Systems Engineering, Lehigh University, USA

[2]School of Business, Stevens Institute of Technology, USA

ISE Technical Report 18T-008

LEHIGH UNIVERSITY

Electric Storage with Stacked Services: Control and Valuation

Boris Defourny[a,*], Somayeh Moazeni[b]

[a]Department of Industrial and Systems Engineering, Lehigh University, 200 W Packer Ave, Bethlehem, PA 18015, USA
[b] School of Business, Stevens Institute of Technology, 1 Castle Point Terrace, Hoboken, NJ 07030, USA

ABSTRACT

This paper considers the management of electric storage resources performing two functions: frequency regulation and electricity price arbitrage. Frequency regulation is a grid service that amounts to letting the grid operator control the device within agreed-upon capacity bounds in exchange for payment at the regulation market price. Price arbitrage means using storage capacity to make transactions at the energy market price. This paper is concerned with the situation where the two functions are considered simultaneously and storage capacity is shared between the two functions dynamically. The real-time regulation characteristics relevant to the capacity allocation problem are described using point processes. The problem of optimizing operations is reduced to a continuous-time stochastic dynamic program. From structural properties of the problem, an exact solution method is derived, that relies on a separation of the state space into discrete and continuous parts. Numerical simulations are used to investigate the value of the storage asset for various market access assumptions. The results are expected to contribute to open questions in electricity market design.

Keywords: Markov decision processes, storage systems, real-time control, electricity markets

1. Introduction

Electric storage is instrumental in the further deployment of electric generation from renewables. Therefore, while most states in the U.S. already have Renewable Portfolio Standards (Barbose 2017), some states have set, or initiated the process of setting, capacity targets for electric storage.

* Corresponding author. Tel.: +16107584250.

E-mail addresses: defourny@lehigh.edu (B. Defourny), smoazeni@stevens.edu (S. Moazeni)

California, with assembly bill AB-2514, and Oregon, with house bill HB-2193, were precursor on this front. More recently, Massachussets has set a target of 200 MWh by 2020, pursuant to Chapter 188 of the Acts of 2016. New Jersey's AB-3723 has set a target of 600 MW by 2021, and 2000 MW by 2030. New York's AB-6571 has set the end of 2018 as the date to determine targets. The regulatory framework usually encourages further technical-economic analysis and allows for the adjustment of targets as implementation details are worked out. Oregon's public utility commission, for instance, in docket UM-1751, has helped to structure storage project selection criteria for utilities, using stakeholders input. Identified criteria include cost-effectiveness, location and benefits to the grid, diversity in ownership and technology, and experience acquisition. See Carnegie et al. (2013) and Akhil et al. (2015) on characteristics of pumped hydro, battery, compressed air, and flywheel. Each can provide a variety of services. This is valuable since in competitive markets, favorable conditions may not hold for an extended period of time. *Service stacking* refers to the provision of multiple services by storage resources, and is expected to increase their value (Forrester et al. 2017).

This paper considers the management of electric storage resources able to perform two functions: providing frequency regulation services, and performing intertemporal electricity price arbitrage. *Regulation* consists in following control signals from a balancing authority in order to correct second-by-second imbalances between power production and consumption. In this paper we approximate the statistical features of the exogenous real-time control signal relevant to our problem. *Energy arbitrage* consists in using storage capacity to gain from variations in electricity spot prices. Spot prices typically vary five-minute by five-minute in today's markets. We study the joint management of the two functions, establishing connections with the revenue management literature, and proposing approximations to model real-time control engineering aspects.

Concretely, the problem we consider is as follows. A state-of-the-world process, described as a continuous-time Markov chain, acts to jointly determine the energy price and the regulation price. Energy transactions take place in the energy market. We use a generic electric storage model: the energy storage capacity is bounded, conversion losses are incurred during energy transfers to and from storage. A continuous revenue stream can also be received in exchange for following an exogenous regulation control signal. By so doing, the storage resource can provide a regulation service paid for at the regulation price. The characteristics of the control signal that we retain are that (i) the deviations from an initial stored energy level remain bounded, and (ii) the deviations cancel out after a random duration. The bounds on exogenous storage level excursions is a decision, that determines the regulation service being provided.

A special characteristic of our problem is that opportunities of energy transactions and opportunities of providing regulation service only happen at random times. In this paper, the random times are described as point processes, with intensity specific to the operation type. This framework follows Moazeni and Defourny (2018) who use random times to describe uncertainty in market access opportunities. The framework is here extended to multiple markets (in fact, two). Most importantly, the problem and solution techniques are quite different.

The interplay between the endogenous and exogenous control of the stored energy level, uncertainty in energy and regulation prices, and uncertainty in times for decisions, lead to a dynamic capacity allocation problem. The main contributions of the paper can be summarized as follows. (i) We reduce the real-time control problem to a continuous-time Markov decision process. We use assumptions that allow us to analyze the problem and study structural properties. (ii) We contribute to exact solution methods for continuous-time Markov decision processes having discrete and continuous state variables (in our problem only one of the state variables is continuous). Our approach works by showing that restricting the state space to a specific finite subset of states can be done without loss of optimality. Subsequently the value function can be extended to the continuous state by solving an auxiliary optimization problem. (iii) We compare variants in the market model. We include the special case where operations are allowed to take place at any time. This is the limit case when the point process that describes market access becomes a continuum. In this case we assume that the execution of decisions is subject to a lead time, assumed to be needed by the system operator for oversight over operations.

While our problem belongs to the energy domain, we believe that our study of a mixed endogenous and exogenous control of capacity, application of stochastic modeling methodologies, and contributions to solution approaches, can be found relevant to other contexts. Below we review related work and our contributions in more detail.

1.1. Literature Review

We furnish pointers to representative related work in three areas: electric energy storage management, inventory control and revenue management, and Markov decision processes.

Electric energy storage management. The management of energy storage resources has first been investigated in the context of a single function. *Back-up power* is a function required to implement *uninterruptible power supply.* See Bekiarov and Emadi (2002). *Frequency regulation* with batteries has long been implemented in islanded systems. In Miller et al. (1996), battery storage is used as a substitute to fast-ramping diesel generators and its value comes from the fuel savings. Regulation leveraging the unique characteristics of electric storage resources is now envisioned at the level of the transmission grid. See PJM staff (2017b) for the most recent description of issues at stake, and Xiao et al. (2014) on performance pricing. See Pradhan and Patil (2016) and Xu et al. (2018) on battery operations in the regulation market. *Energy arbitrage* with batteries has been considered in several studies. See for instance van de Ven et al. (2013) and Bhattacharya et al.

(2018) for results on threshold policies, Moazeni et al. (2015) for risk-sensitive energy arbitrage, Hu and Defourny (2017) for energy arbitrage aware of battery deterioration. Natural gas storage furnishes related storage problems involving gas markets. See Thompson et al. (2009) and Lai et al. (2010). *Capacity firming* is concerned with the regulation of a stochastic energy source via energy storage. Work on this function includes *storage theory* which was motivated by hydro-storage applications and focuses of the stochastic aspects of the problem. See Moran (1954), Gani (1957), Harrison and Resnick (1976) and Prabhu (1998). Work on capacity firming for wind generation includes Kim and Powell (2011) for the study of optimal commitments, Qi et al. (2015) for the interplay with transmission planning, Dvorkin et al. (2017) for storage placement and sizing.

Providing multiple functions with a single storage asset is a practice that has been called *service stacking* in the electricity industry (Forrester et al. 2017). Service stacking is expected to increase revenue from storage assets and improve resilience to changes in market conditions. A real issue is that rules evolve rapidly in the regulation market. For instance, a change in January 2017 affected PJM's energy-neutral regulation signal and modified the characteristics of the regulation signals (FERC EL17-64-000, PJM staff 2017b).

Although work has been done on service stacking, the stochastic nature of the problem has rarely been entirely retained, and few structural optimality results have been reported. Donadee (2013) develop a dynamic programming model and finds significant value in the co-optimization of energy arbitrage and regulation, but uses perfect foresight of the price of electricity. Xi et al. (2015) consider energy arbitrage, up and down regulation markets, and back-up power services, and find that most of the value comes from regulation and to a lesser extent from energy arbitrage, but resort to rolling-horizon approximations. Berrada et al. (2016) consider regulation as well as energy arbitrage in a day-ahead and a spot market, and find that the day-ahead market is not utilized, but their study is done with perfect foresight of all prices. Staffell and Rustomji (2016) consider regulation and energy arbitrage in UK markets, but resort to ad-hoc heuristics to determine operations. Yu and Foggo (2017) develop stochastic models of regulation and energy markets, but optimize storage operations using perfect foresight.

In our paper, the ability to operate is intermittent and stochastic. This framework relates to the *transactive energy markets* envisioned in the white papers by Rahimi and Ipakchi (2016) and Cazalet et al. (2016), which aim at facilitating peer-to-peer energy transactions over the distribution network. We use point processes to describe opportunities of transactions. The utilization of point processes is standard in the study of transactive systems. For instance, Cont et al. (2010) study limit-order books and the interplay between arrival rates and bid-ask spreads.

Inventory control and revenue management. Within the field of operations research, our paper relates to the inventory control and revenue management literature. In our work, we describe the participation to the regulation market as if blocks of storage capacity were rented out to another party for some random time, at a certain price per unit of time and capacity. This is predicated on the fact that the regulation signal is energy-neutral after a random time, see Borsche et al. (2013) and PJM staff (2017b). During the time a block of capacity is rented out, the block is not available for supporting energy transactions. Given the analogy between participation in the regulation market and rental of storage capacity, our problem has similarities with the literature on capacity rationing in the presence of multiple customer classes, see Gans and Savin (2007), Papier and Thonemann (2010), Chong et al. (2017). However, in our problem the arrival processes for the different classes of "customers" are statistically dependent, because the "class" depends on the state-of-the-world process. Our problem also differs in the mixed control mode for the capacity.

The value of service stacking relates to the value of resource flexibility. The value of resource flexibility has been investigated in various ways. For instance, Chod and Rudi (2005) study a single resource system used to satisfy two price-sensitive demand classes. Their model is a two-stage optimization problem. Under their assumptions, the authors find that flexibility is most valuable under high demand variability and negative demand correlation. Akcay et al. (2010) discuss the value of flexible resources by formulating a stochastic dynamic program in discrete time over a finite horizon. However, the authors employ deterministic approximations. Al-Gwaiz et al. (2016) study the competition between flexible and inflexible electricity suppliers in a market with variations from renewables. They assume that the system operator curtails the production from renewables to reduce the market power of the flexible producers. Thus in that work a higher-level decision maker controls the variability of demand.

Markov decision processes. We reduce our problem to a continuous-time Markov decision process (Puterman 2005). The problem admits optimality conditions that can be expressed using uniformization techniques, by which the continuous-time problem can be mapped to a discrete-event problem (Lippman 1975, Puterman 2005, Gans and Savin 2007). However, our problem has a continuous state variable that makes the optimality condition system infinite-dimensional. We obtain results allowing us to circumvent this issue and to construct the exact value function on the entire state space. The solution is provably optimal, in contrast to generic discretization methods (Whitt 1978) or sampling procedures (Rust 1997) which can be used heuristically to produce approximate solutions. Thus in this paper we do not face a tradeoff between accuracy and computational cost. This allows us to carry out sensitivity analysis studies numerically.

1.2. Organization

The remainder of the paper is organized as follows. Section 2 provides background on regulation and presents a simulation study to clarify the impact of modeling approximations used in the sequel.

Section 3 presents the storage valuation model, formulates it as a continuous-time Markov decision process, and describes the optimality conditions satisfied by the value function. Section 4 studies the structure and sensitivity properties of the value function. Section 5 describes the solution method. Section 6 studies the value of the joint operations model over an optimal static allocation of the storage capacity. Section 7 presents the model with scheduled transactions and derive the corresponding optimality conditions and properties for the value function. Section 8 presents our numerical work and the comparisons between markets with random permissions and markets with scheduled transactions, and Section 9 concludes.

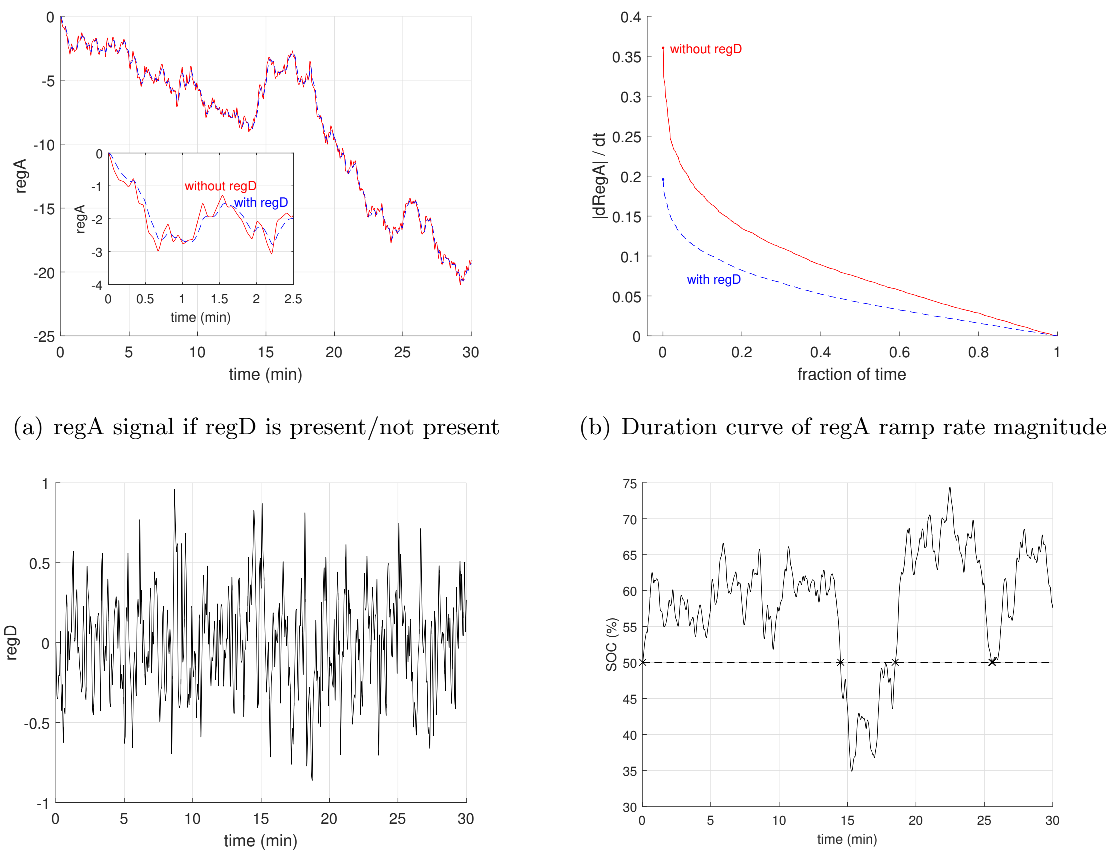
2. Statistical Features of the Regulation Signal

This section provides background on secondary frequency regulation and presents our analysis of issues at stake. For concreteness we focus on PJM's real-time control. Based on description by PJM staff (2017b), there are two regulation signals, called regA and regD. The regD signal is energy-neutral in the sense that the cumulated net power injection of a device that obeys this regulation signal will cancel out after some time. This is needed for devices with energy capacity constraints. RegD is offered to open the regulation market to electric storage resources, and is the regulation signal considered in this paper. Electric generators typically follow the regA signal, which varies more slowly and is not energy-neutral.

Based on our examination, historical regulation signal data cannot be very predictive, because the gains of the controller that produces the regulation signals are subject to modifications by the system operator. Therefore, for this work, we reconstruct a controller based on information in PJM staff (2017b), allowing us to generate the regulation signals for a variety of gains under a same sample path for the so-called "area control error" to correct. The details of the controller are in Appendix A. Note that when the resources cannot follow the regulation signals, the residual error is fed back to the input to the regulation controller. There is a mechanism in the controller that anticipates this effect given the resource constraints.

A first point to illustrate is the usefulness of the energy-neutral regulation signal. For the same sample path of real-time errors to correct, Figure 1(a) depicts 30 minutes of regA signal (in MW) in two cases: (i) when regA is the only signal produced, and (ii) when regD contributes to the regulation effort. The controller RegA is not energy-neutral, as seen from the drift in the signal. The relief in ramping effort is clarified in the inset. In Figure 1(b) we report the empirical duration curve of the derivative of the regA signal (fraction of time the ramp rate magnitude exceeds a value), estimated by finite differences over a period of 24 hours. RegD provides significant relief to RegA in terms of the ramp rate that is needed to follow RegA.

Figure 1 Regulation signals produced by the regA-regD controller considered in this section.



(a) regA signal if regD is present/not present

(b) Duration curve of regA ramp rate magnitude

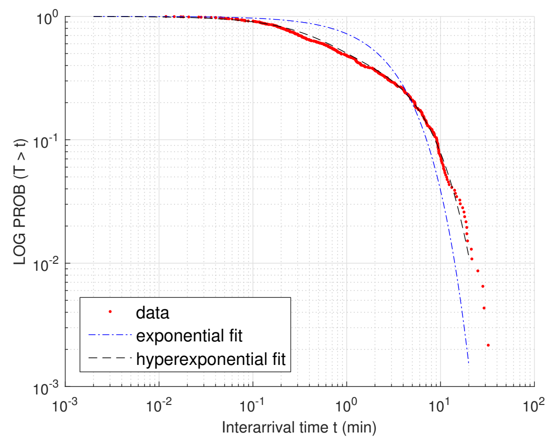
(c) regD signal

(d) State of charge (SoC), following regD

Figure 1(c) depicts the RegD signal. Figure 1(d) integrates this signal to describe the state of charge of a perfectly efficient storage device that would follow regD. Please note that it is possible to refine the controller to incorporate storage inefficiencies when following regD. Assuming the reference storage level is 50% of a given capacity, the figure also marks the return times to the reference level. We say that RegD attains energy neutrality at those times.

We let the energy neutrality times be the points of an arrival process, and set out to identify the distribution for the interarrival times. Since later on we consider continuous-time Markov decision processes, for convenience we restrict our search to phase-type distributions. See Neuts (1995) Ch. 2. We use the fitting algorithms from Feldmann and Whitt (1998), related to Prony's method, to fit a mixture of exponentials, also known as hyperexponential distribution. We find that the exponential distribution provides a reasonable approximation, and that a very good fit is already obtained with a mixture of two exponentials, as shown in Figure 2. Here the empirical mean of the interarrival time is $1/\hat{\lambda} = 3.0783$ minutes, used to determine the exponential fit $\text{Prob}(T > t) \simeq$

Figure 2 Complementary cumulative distribution function of the energy neutrality return time (in log scale).



$e^{-\hat{\lambda}t}$. The hyperexponential fit is $\text{Prob}(T > t) \simeq \hat{p}_1 e^{-\hat{\lambda}_1 t} + \hat{p}_2 e^{-\hat{\lambda}_2 t}$ with $\hat{p}_1 = 0.5397$, $\hat{p}_2 = 0.4603$, $\hat{\lambda}_1 = 0.1931$, $\hat{\lambda}_2 = 2.0721$.

We conclude this section with the following remarks. First, this study is presented to aid understanding of real-time regulation and bridge the gap with the interarrival time model used in the sequel. As with queueing systems, assuming exponentially distributed interarrival times is not fully accurate but is convenient. Second, the interarrival time distribution does depend on the regulation controller. The controller is designed by the system operator. We have presented results with gains adjusted to keep return times below 30 minutes with high probability (see FERC EL18-87-000, §16). This can be checked from Figure 2. From other experiments with the controller, we find that the mean return time can be reduced, but there is a tradedoff with the regA ramp rate reduction in the presence of RegD. Third, PJM's RegA-RegD control is chosen to illustrate an approach applicable to any controller. In the sequel, we treat the mean return time as a problem parameter characterizing the regulation market, and later we provide sensitivity results.

3. Model Description

We first give a compact overview of the model along with a summary of the notation. We then describe the components of the system in more detail.

The price for energy and the price of regulation are jointly determined by a continuous-time Markov chain referred to as the *background process*, see §3.1. For reference we reproduce the notation defined in §3.1: $m \in \{1, \ldots, M\}$ is the state of the background process, Q with elements q_{mn} is the transition rate matrix, p_m is the energy price when being in state m, and ϱ_m is the regulation price when being in state m.

The storage operations are of three types: charge, discharge, and provision of regulation services. Charge and discharge operations happen through the purchase and sale of electricity in the energy market. The electric energy storage capacity is K, the charging efficiency is $\eta^{\mathrm{c}} \in (0,1]$, and the discharging efficiency is $\eta^{\mathrm{d}} \in (0,1]$. See §3.2. For reference, a is the change of stored energy in absolute value, $C(a/\eta^c, m)$ is the cost from purchasing a quantity a/η^c when being in state m and $D(\eta^d a, m)$ is the payoff from selling a quantity $\eta^d a$ when being in state m.

Regulation is performed by dedicating blocks of storage capacity to the following of a regulation signal. The regulation signal causes variations in the state of charge that are guaranteed to remain within the dedicated capacity and to cancel out after a random duration. The current state of charge determines the maximal number of blocks that can be assigned to regulation. See §3.3. For reference, u is the decision on the blocks assigned to regulation, and $R(u,m)$ is the equivalent lump-sum reward when being in state m and taking decision u, see §3.4.

Operations in the energy and regulation markets can only be initiated at specific random times. Permissions for charging arrive as a Poisson process of intensity λ^{c}. Permissions for discharging arrive as a Poisson process of intensity λ^{d}. Regulation service requests arrive as a Poisson process of intensity λ^{r}. The arrival processes are statistically independent. We denote the corresponding counting processes by $\{N_t^{\mathrm{c}}\}_{t\geq 0}$, $\{N_t^{\mathrm{d}}\}_{t\geq 0}$, and $\{N_t^{\mathrm{r}}\}_{t\geq 0}$.

The decision problem is formulated as a continuous-time Markov decision process over an infinite horizon, using a discounted expected return objective. See §3.4. For reference, $\gamma > 0$ is the discount rate and $V(k,\ell,m)$ is the value function where k is the stored energy for operations in the energy market, ℓ is the number of blocks in use for regulation, and m is the background process state.

3.1. Random Environment

The random environment is modeled as a continuous-time Markov chain. See Neuts (1995) Ch. 6. We refer to this state-of-the-world process as the *background process*. The states of the background process are labeled from 1 to M. We denote by m a generic element of the background process state space, and by Q the state transition rate matrix with elements q_{mn}. We assume that the process is ergodic. This ensures in particular that a stationary state distribution is uniquely defined. Let p_m be the energy price and let ϱ_m be the regulation price when being in background state m.

The background process can be formed by combining several continuous-time Markov chains. In this paper we assume the energy and regulation prices follow independent continuous-time Markov chains. Suppose the energy price process has n_1 states and let $Q^{\mathrm{en}} \in \mathbb{R}^{n_1 \times n_1}$ be its transition rate matrix. Similarly suppose the regulation price process has n_2 states and let $Q^{\mathrm{r}} \in \mathbb{R}^{n_2 \times n_2}$ be its transition rate matrix. Then the transition rate matrix $Q \in \mathbb{R}^{n_1 n_2 \times n_1 n_2}$ of the joint process over the product state space can be defined as

$$Q = Q^{\mathrm{r}} \otimes I_{n_1} + I_{n_2} \otimes Q^{\mathrm{en}} = Q^{\mathrm{r}} \oplus Q^{\mathrm{en}}, \tag{1}$$

where I_n is the identity matrix of dimension n, $\otimes$ denotes the Kronecker product, and $\oplus$ denotes the Kronecker sum, see Buchholz (1994). Assuming (i,j) refers to the i-th energy price state and j-th regulation price state, our state labeling convention maps $m=1$ to $(1,1)$, $m=2$ to $(2,1)$, $\ldots$, $m=n_1n_2$ to (n_1,n_2).

3.2. Operations in the Energy Market

Let $C(s,m)$ denote the expected cost of purchasing an energy quantity $s \geq 0$ when the background process is in state m. Let $D(s,m)$ denote the expected payoff from selling an energy quantity $s \geq 0$ when the background process is in state m. We assume that transactions are paid at the energy market price and executed instantaneously. In this case the functions C, D are linear in the quantity s and we have

$$C(s,m) = p_m s, \qquad\qquad D(s,m) = p_m s. \tag{2}$$

C and D are nonnegative in s and null at $s=0$, with $C(s,m) \geq D(s,m)$ for all $s \geq 0$.

Let a denote the variation in stored energy in absolute value. Due to storage inefficiencies, charging a corresponds to purchasing a quantity $s = a/\eta^c$ while discharging a corresponds to selling a quantity $s = a\eta^d$. Therefore, adding a to storage incurs a cost $C(a/\eta^{\mathrm{c}}, m)$ while using a from storage generates a payoff $D(a\eta^{\mathrm{d}}, m)$.

3.3. Operations in the Regulation Market

Regulation service events arrive as a Poisson process $\{N_t^r\}_{t\geq 0}$ with intensity λ^{r}. We consider two variants to describe the regulation service. Both accumulate rewards continuously in time, but later we show how these payments can be converted to equivalent expected lump-sum rewards for the purpose of optimization. This leads to an equivalent lump-sum regulation reward function

$$R(u,m) = r_m u \tag{3}$$

where the feasible set for the decision u and the state-dependent coefficient r_m will be defined later.

3.3.1. Rental Model. The first variant is a *rental model*. Each arrival is a request for one block of storage capacity. We assume the block size is 1, by rescaling energy units and prices if necessary. The service duration is not specified in the request but is known to be exponentially distributed with mean $1/\mu$. The regulation service durations and request interarrival times are mutually independent. At each request arrival, the decision is to accept $(u=1)$ or reject $(u=0)$ the request. If a request is accepted, a new block of capacity is reserved for the random duration of service, and the block is no longer available for supporting energy arbitrage operations. When the service ends, the block is released. During service time, a revenue of ϱ_m per unit of time is paid in exchange for the right of using the capacity block. We recall that ϱ_m denotes the regulation price when the background process is in state m. Here we assume that ϱ_m is set at the time of the request arrival and held for the entire service duration. The decision $u=0$ is the only feasible decision when the free storage capacity is inferior to the block size.

3.3.2. Mid-Charge Model. The second variant more closely reflects the nature of the regulation signal, but complicates the statement of results in the paper. We call it the *mid-charge model*. The regulation event arrival times are the times where storage resources are believed to hit a mid-charge level by the system operator, assuming they follow the regulation signal. The system operator controller generates the regulation signal in such a way that excursions from the assumed mid-charge level are bounded up and down by a same amount. This guarantees that the regulation signal can be followed by an energy-limited resource proportionally to some assigned capacity, reserved when the signal is picked up at arrival time. Interarrival times between assumed mid-charge level are independent and exponentially distributed with mean $1/\lambda^{\mathrm{r}}$, whence the arrival process $\{N_t^{\mathrm{r}}\}_{t\geq 0}$. At arrival times, the decision u is the change, positive or negative, in the number of capacity blocks committed to regulation until the next arrival. We count the number of blocks reserved for allowing deviations above the assumed mid-charge level, knowing that an equal number of blocks is also reserved for deviations below. The counted blocks are paid at the ongoing regulation price, which evolves during service.

3.4. Continuous-Time Markov Decision Process Formulation

The background process sojourn times, interarrival times and regulation service durations are exponentially distributed. Therefore, the control problem can be studied as a continuous-time Markov decision process (semi-Markov decision process with exponentially distributed sojourn times). The objective is the expected discounted total return over an infinite horizon. The expected discounted return, to be maximized, is the expected discounted revenues from discharging operations and regulation service provision, minus the expected discounted costs due to charging operations. Let $\gamma>0$ denote the discount rate in continuous-time.

The *system state* is made of the *background process state* (§ 3.1) and the *storage resource state*. The storage resource state is described by the quantity of stored energy k destined to price arbitrage in the energy market, and the number ℓ of capacity blocks in use to provide regulation services.

In the *rental model* (§ 3.3), the energy storage capacity constraint implies that $k+\ell\leq K$ needs to be satisfied by all feasible (k,ℓ). For the ease of exposition, K is assumed to be a multiple of the block size. Thus, the system state space is

$$\mathcal{S} := \{(k,\ell,m)\in[0,K]\times\{0,1,\ldots,K\}\times\{1,\ldots,M\} \mid k+\ell\leq K\}. \tag{4a}$$

State-dependent feasibility sets $\mathcal{A}^{\mathrm{c}}(k,\ell)$, $\mathcal{A}^{\mathrm{d}}(k,\ell)$, $\mathcal{A}^{\mathrm{r}}(k,\ell)$ for the charging decision, discharging decision, and regulation decision are defined for all states $(k,\ell,m)\in\mathcal{S}$ as follows, based on the condition that $(k+a,\ell,m)$, $(k-a,\ell,m)$, $(k,\ell+u,m)$ for $u\in\{0,1\}$ must remain in $\mathcal{S}$:

$$\mathcal{A}^{\mathrm{c}}(k,\ell)=[0,K-(k+\ell)], \qquad \mathcal{A}^{\mathrm{d}}(k,\ell)=[0,k], \qquad \mathcal{A}^{\mathrm{r}}(k,\ell)=\begin{cases}\{0,1\} & \text{if } k+\ell+1\le K,\\ \{0\} & \text{if } k+\ell+1>K.\end{cases} \tag{5a}$$

In the *mid-charge model* (§ 3.3), we introduce the additional constraint $k-\ell\ge 0$ to express the convention that reserving ℓ blocks for regulation means we should be able to immediately follow a signal instructing us to either charge or discharge up to the quantity ℓ assigned to regulation. In this case the state space becomes, using $\lfloor\cdot\rfloor$ for the floor function,

$$\mathcal{S}:=\{(k,\ell,m)\in[0,K]\times\{0,1,\ldots,\lfloor K/2\rfloor\}\times\{1,\ldots,M\}\mid k+\ell\le K,\ k-\ell\ge 0\}. \tag{4b}$$

Based on the condition that $(k+a,\ell,m)$, $(k-a,\ell,m)$ and $(k,\ell+u,m)$ for u integer must remain in $\mathcal{S}$, the feasibility sets defined for each $(k,\ell,m)\in\mathcal{S}$ from (4b) become

$$\mathcal{A}^{\mathrm{c}}(k,\ell)=[0,K-(k+\ell)],\ \mathcal{A}^{\mathrm{d}}(k,\ell)=[0,k-\ell],\ \mathcal{A}^{\mathrm{r}}(k,\ell)=\{-\ell,\ldots,-\ell+\min\{\lfloor k\rfloor,\lfloor K-k\rfloor\}\}\,. \tag{5b}$$

Since the decisions are made at permission times only, it is sufficient to consider the embedded discrete-time process defined by the epochs of the permission arrivals. See Puterman (2005) Chapter 11. In the sequel, we directly work with the value function $V(k,\ell,m)$, which gives the optimal value of the objective when starting from state (k,ℓ,m). The explicit objective is established in Appendix B. The key step is the conversion of the continuous-time discounted revenue from regulation into an equivalent lump-sum expected regulation reward function $R(u,m)=r_m u$. In the *rental model*, the reward coefficient r_m is

$$r_m=\frac{\varrho_m}{\mu+\gamma}. \tag{6a}$$

In the *mid-charge model*, the reward coefficient r_m is

$$r_m=e_m^{\top}(\gamma I-Q)^{-1}\varrho\,, \tag{6b}$$

where e_m is the m-th unit column vector in $\mathbb{R}^M$, $I\in\mathbb{R}^{M\times M}$ is the identity matrix, and $\varrho\in\mathbb{R}^M$ is the column vector with elements ϱ_m.

Gans and Savin (2007) have discussed equivalent lump-sum rewards of the form (6a) and explained that making μ state-dependent can be done at the cost of having to introduce state variables ℓ_m for each m. Rewards of the form (6b) are based on a different accounting scheme: instead of taking the expectation over a random duration, it is assumed that the continuous regulation reward is received forever. A reimbursement is then made when service ends, valued as the negative part of $R(u,n)$ when being in state n. Under (6b), λ^{r} can be made state-dependent without having to augment the state space.

3.5. Value Function and Optimality Conditions

The optimal control problem admits optimality conditions that can be expressed using uniformization techniques (Lippman 1975). To do this, define first the aggregate rate of events

$$\Lambda := \lambda^{\mathrm{d}} + \lambda^{\mathrm{c}} + \lambda^{\mathrm{r}} + \mu K + \bar{q} + \gamma, \qquad \bar{q} := \textstyle\sum_{m=1}^{M} \bar{q}_m, \qquad \bar{q}_m := \textstyle\sum_{n=1,\ n\neq m}^{M} q_{mn}. \tag{7}$$

By convention, μ is understood to be set to 0 for the mid-charge model which has no such parameter.

We assume that the system evolves with exponentially distributed inter-event times of mean $1/\Lambda$. However, when being in state (k,ℓ,m), the probability that the next event is a "null-event", that sends the system back into the same state, is $\nu_{\ell,m}/\Lambda$, where

$$\nu_{\ell,m} := \bar{q} - \bar{q}_m + \mu(K-\ell). \tag{8}$$

The value function associated with the control problem is defined as follows: for all $(k,\ell,m)\in\mathcal{S}$,

$$\begin{aligned} V(k,\ell,m) = {} & \frac{\lambda^{\mathrm{c}}}{\Lambda}\mathcal{H}^{\mathrm{c}}_{k\ell m}[V] + \frac{\lambda^{\mathrm{d}}}{\Lambda}\mathcal{H}^{\mathrm{d}}_{k\ell m}[V] + \frac{\lambda^{\mathrm{r}}}{\Lambda}\mathcal{H}^{\mathrm{r}}_{k\ell m}[V] \\ & + \frac{\mu\ell}{\Lambda}V(k,\ell-1,m) + \frac{1}{\Lambda}\Big(\sum_{\substack{n=1\\ n\neq m}}^{M} q_{mn} V(k,\ell,n)\Big) + \frac{\nu_{\ell,m}}{\Lambda}V(k,\ell,m), \end{aligned} \tag{9}$$

where the operators $\mathcal{H}^{c}_{k\ell m}[\cdot]$, $\mathcal{H}^{d}_{k\ell m}[\cdot]$, $\mathcal{H}^{r}_{k\ell m}[\cdot]$ are given by

$$\mathcal{H}^{\mathrm{c}}_{k\ell m}[V] = \max_{a\in\mathcal{A}^{\mathrm{c}}(k,\ell)} \left\{V(k+a,\ell,m) - C(a/\eta^{c},m)\right\}, \tag{10}$$

$$\mathcal{H}^{\mathrm{d}}_{k\ell m}[V] = \max_{a\in\mathcal{A}^{\mathrm{d}}(k,\ell)} \left\{V(k-a,\ell,m) + D(\eta^{d}a,m)\right\}, \tag{11}$$

$$\mathcal{H}^{\mathrm{r}}_{k\ell m}[V] = \max_{u\in\mathcal{A}^{\mathrm{r}}(k,\ell)} \left\{V(k,\ell+u,m) + R(u,m)\right\}. \tag{12}$$

The operator $\mathcal{H}^{c}_{k\ell m}[\cdot]$ gives the value function at the next state, based on the optimal amount a to charge, should a charge permission arrive while being in state (k,ℓ,m). The operator $\mathcal{H}^{d}_{k\ell m}[\cdot]$ gives the value function at the next state, based on the optimal amount a to discharge, should a discharge permission arrive. The operator $\mathcal{H}^{r}_{k\ell m}[\cdot]$ gives the value function at the next state, based on the optimal regulation decision u, should a regulation event arrive. The quantity $\mu\ell V(k,\ell-1,m)$, which is always null in the mid-charge model, is null at $\ell=0$ in the rental model, with $V(k,-1,m)$ left undefined. The system of equations (9), which needs to hold for all $(k,\ell,m)\in\mathcal{S}$, is referred to as $V=\mathcal{T}V$, adopting the operator-view of dynamic programming (Denardo 1967). The system is infinite-dimensional because the state variable k is continuous.

In our exposition we assume the arrival processes are Poisson processes. An extension to Markov-modulated arrival processes is possible by replacing the arrival rates $\lambda^{\mathrm{c}},\lambda^{\mathrm{d}},\lambda^{\mathrm{r}}$ by state-dependent rates $\lambda^{\mathrm{c}}_m,\lambda^{\mathrm{d}}_m,\lambda^{\mathrm{r}}_m$. The details are in Appendix C. For instance, one can set $\lambda^{\mathrm{c}}_m=\lambda^{\mathrm{d}}_m=0$ in some states m to forbid energy transactions in states associated with grid congestion.

4. Structural Properties of the Value Function

This section discusses the structure of the value function $V(k,\ell,m)$ in (9). This helps to differentiate our problem from previously studied storage problems. The proofs are in Appendix H.

Proposition 1 says that receiving in expectation more opportunities to act is never detrimental.

Proposition 1. *For each fixed state* $(k,\ell,m) \in \mathcal{S}$, *the value function* $V(k,\ell,m)$ *is nondecreasing in the arrival rates* $\lambda^c, \lambda^d, \lambda^r$.

Proposition 2 states a negative result for the *rental model*: having a greater expected regulation service duration is not necessarily beneficial or detrimental. The proof works by counterexamples.

Proposition 2. *With the rental model, there exists problems where the value function* $V(k,\ell,m)$ *is nonincreasing in the mean regulation service time* $1/\mu$, *and problems where* $V(k,\ell,m)$ *is nondecreasing in* $1/\mu$.

Proposition 3 states a negative result: having more energy in storage is not necessarily beneficial. This is due to the opportunity cost of having capacity occupied by stored energy, while this energy cannot be disposed of without the arrival of a discharge permission. One could of course increase the discharge permission rate to relax this constraint.

Proposition 3. *For each fixed* ℓ *and* m, *the value function* $V(k,\ell,m)$ *is not necessarily nondecreasing in the stored energy* k.

Proposition 4 states that having more capacity dedicated to regulation is detrimental to the value function. Note however that lump-sum payments will be received upfront when the capacity dedicated to regulation is increased. The level of commitment ℓ is viewed as a liability by the value function because the release of this commitment is driven by an exogenous process.

Proposition 4. *For each fixed* k *and* m, *the value function* $V(k,\ell,m)$ *is nonincreasing in* ℓ.

Proposition 5 concerns pure energy-arbitrage problems, when no capacity can be assigned to regulation. It establishes the optimality of "bang-bang" policies (subordinated to permission arrival times) for such problems. It also shows that the value function for such problems is affine in the stored energy. The proof relies on the property that the reward functions C, D are linear in a.

Proposition 5. *Consider the special case where operations can only happen in the energy market (set* $\ell = 0$ *and* $\lambda^r = 0$*). (1) Suppose that initially,* $k = 0$ *or* $k = K$. *Then without loss of optimality the decisions can be restricted to* $a \in \{-K, 0, K\}$, *and the stored energy state can be restricted to* $k \in \{0, K\}$. *(2) Furthermore, allowing now* $k \in [0, K]$ *and* $a \in [-K, K]$, *for each fixed* m *the value function is affine in* k.

Proposition 6 establishes that the value function as a function of k is piecewise linear with discontinuities. We refer to Figure 5 in Section 8 for visualizing these results. Proposition 6 relies on the property that the functions C and D are linear in the quantity a.

Proposition 6. *For each fixed ℓ, m, the value function $V(k,\ell,m)$ as a function of k is piecewise linear with discontinuities between the pieces. In the rental model, the function has right-discontinuities at $k = \{0,1,\ldots,K-1\} \cap \{k \in [0,K] : k+\ell \leq K\}$, and is left-continuous with right limits. In the mid-charge model, the discontinuities are (i) left-discontinuities at $k = \{1,\ldots,\lfloor K/2 \rfloor\} \cap \{k \in [0,K/2] : k-\ell \geq 0\}$, and (ii) right-discontinuities at $k = \{\lceil K/2 \rceil,\ldots,K-1\} \cap \{k \in [K/2,K] : k+\ell \leq K\}$. The function is right-continuous with left limits on $k \in [\ell, K/2]$ and left-continuous with right limits on $k \in [K/2, K-\ell]$.*

Proposition 7 establishes that the slope of the linear pieces described in Proposition 6 are a function of the background process state m only.

Proposition 7. *The directional derivatives $\partial_k^- V(k,\ell,m) := \lim_{\epsilon \downarrow 0^+}[V(k,\ell,m) - V(k-\epsilon,\ell,m)]/\epsilon$, defined where V is left-continuous, and the directional derivatives $\partial_k^+ V(k,\ell,m) := \lim_{\epsilon \downarrow 0^+}[V(k+\epsilon,\ell,m) - V(k,\ell,m)]/\epsilon$, defined where V is right-continuous, are independent of k and ℓ, that is, they only depend on the state m of the background process.*

5. Solution Method

In principle, the infinite-dimensional system of equations (9) can be solved approximately, by discretizing k and solving the resulting finite-dimensional approximate problem. Using results from Section 4 however, we can calculate the optimal value function exactly. Our approach works by formulating and solving two independent optimization problems, and then combining the results to define the value function on the entire state space. We refer to Section 8 for numerical results.

5.1. Restriction to a Finite Subset of States

Consider the restriction of $k \in [0,K]$ to the finite set $\{0,1,\ldots,K\}$. We refer to the discretized state space as $S = \mathcal{S} \cap \mathbb{N}^3$. The storage decision a is restricted to integer values so that S is closed under state transitions. For each $(k,\ell,m) \in S$, we define the discrete feasible sets $A^{\mathrm{c}}(k,\ell) = \mathcal{A}^{\mathrm{c}}(k,\ell) \cap \mathbb{N}$, $A^{\mathrm{d}}(k,\ell) = \mathcal{A}^{\mathrm{d}}(k,\ell) \cap \mathbb{N}$, and $A^{\mathrm{r}}(k,\ell) = \mathcal{A}^{\mathrm{r}}(k,\ell)$. We have $A^{\mathrm{d}}(k,\ell) = \{0,1,\ldots,k\}$ with the rental model, $A^{\mathrm{d}}(k,\ell) = \{0,1,\ldots,k-\ell\}$ with the mid-charge model, and $A^{\mathrm{c}}(k,\ell) = \{0,1,\ldots,K-(k+\ell)\}$ with both models. The set $A^{\mathrm{r}}(k,\ell)$ is expressed as in (5a) or (5b) except that rounding is now superfluous. We define the operators $H^{\mathrm{c}}_{k\ell m}[\cdot]$, $H^{\mathrm{d}}_{k\ell m}[\cdot]$, $H^{\mathrm{r}}_{k\ell m}[\cdot]$ from the corresponding operators $\mathcal{H}^{\mathrm{c}}_{k\ell m}[\cdot]$, $\mathcal{H}^{\mathrm{d}}_{k\ell m}[\cdot]$, $\mathcal{H}^{\mathrm{r}}_{k\ell m}[\cdot]$ with $\mathcal{A}^{\mathrm{c}}(k,\ell)$, $\mathcal{A}^{\mathrm{d}}(k,\ell)$, $\mathcal{A}^{\mathrm{r}}(k,\ell)$ replaced by $A^{\mathrm{c}}(k,\ell)$, $A^{\mathrm{d}}(k,\ell)$, $A^{\mathrm{r}}(k,\ell)$.

5.2. Optimality of a Value Function Calculated on the Finite Subset of States

Let v denote the value function satisfying (9) for all $(k,\ell,m) \in S$, with the operators $\mathcal{H}^{\mathrm{c}}_{k\ell m}[\cdot]$, $\mathcal{H}^{\mathrm{d}}_{k\ell m}[\cdot]$, $\mathcal{H}^{\mathrm{r}}_{k\ell m}[\cdot]$ replaced by $H^{\mathrm{c}}_{k\ell m}[\cdot]$, $H^{\mathrm{d}}_{k\ell m}[\cdot]$, $H^{\mathrm{r}}_{k\ell m}[\cdot]$. The discrete problem is finite-dimensional and can thus be solved with any dynamic programming method. It is convenient to use the linear optimization approach to dynamic programming, see Kallenberg (1994).

Let $v_{k\ell m}$ for $(k,\ell,m) \in S$ denote the optimization variable corresponding to v at (k,ℓ,m). Let $h^{c}_{k\ell m}$, $h^{d}_{k\ell m}$, $h^{r}_{k\ell m}$ be the optimization variables corresponding to the value of the operators $H^{c}_{k\ell m}$, $H^{d}_{k\ell m}$, $H^{r}_{k\ell m}$ applied to v. Let $L(k)$ denote the set of feasible values of ℓ given k. We have $L(k) = \{0,1,\dots,K-k\}$ for the rental model, and $L(k) = \{0,1,\dots,\min\{k,K-k\}\}$ for the mid-charge model. Then v can be found by solving the linear optimization problem

$$\begin{aligned}
\text{minimize} \quad & \textstyle\sum_{k=0}^{K} \sum_{\ell \in L(k)} \sum_{m=1}^{M} v_{k\ell m} && (13)\\
\text{subject to} \quad & (\Lambda - \nu_{\ell,m}) v_{k\ell m} = \lambda^{\mathrm{d}} h^{\mathrm{d}}_{k\ell m} + \lambda^{\mathrm{c}} h^{\mathrm{c}}_{k\ell m} + \lambda^{\mathrm{r}} h^{r}_{k\ell m} + \mu \ell v_{k,\ell-1,m} + \textstyle\sum_{\substack{n=1\\ n\neq m}}^{M} q_{mn} v_{k\ell n}, \\
& h^{\mathrm{c}}_{k\ell m} \geq v_{k+a,\ell,m} - C(a/\eta^{c}, m) \quad \text{for } a \in A^{\mathrm{c}}(k,\ell), \\
& h^{\mathrm{d}}_{k\ell m} \geq v_{k-a,\ell,m} + D(\eta^{d} a, m) \quad \text{for } a \in A^{\mathrm{d}}(k,\ell), \\
& h^{\mathrm{r}}_{k\ell m} \geq v_{k,\ell+u,m} + R(u, m) \quad \text{for } u \in A^{\mathrm{r}}(k,\ell), \qquad \forall\, (k,\ell,m) \in S.
\end{aligned}$$

The solution to the discrete problem actually gives the exact value function at the discrete states:

Proposition 8. *The optimal value function V such that $V = \mathcal{T}V$ over $(k,\ell,m) \in \mathcal{S}$ satisfies $V(k,\ell,m) = v(k,\ell,m) = v_{k\ell m}$ over $(k,\ell,m) \in S$.*

5.3. Extension to the Entire State Space

We assume that there is an interest in evaluating the value function V at noninteger values of k, otherwise the solution from (13) suffices.

Consider then the subproblem of operating a storage device of capacity 1 in the energy market only, restricted to the stored energy states $k \in \{0,1\}$. Let $w(k,m)$ denote the optimal value function for this problem. Adapting (7)-(9) by setting $K = 1$, $\ell = 0$, $\lambda^{\mathrm{r}} = 0$, $\mu = 0$, $k \in \{0,1\}$ leads to the optimality conditions for w, expressed using an aggregate rate of events $\Lambda_2 := \lambda^{\mathrm{c}} + \lambda^{\mathrm{d}} + \bar{q} + \gamma$ and null-event rates $\nu_m := \bar{q} - \bar{q}_m$ devoid of the regulation-related events,

$$\begin{aligned}
w(0,m) &= \tfrac{\lambda^{\mathrm{c}}}{\Lambda_2} H^{\mathrm{c}}_{0m}[w] + \tfrac{\lambda^{\mathrm{d}}}{\Lambda_2} w(0,m) + \tfrac{1}{\Lambda_2}\Big(\textstyle\sum_{\substack{n=1\\ n\neq m}}^{M} q_{mn} w(0,n)\Big) + \tfrac{\nu_m}{\Lambda_2} w(0,m), && (14)\\
w(1,m) &= \tfrac{\lambda^{\mathrm{c}}}{\Lambda_2} w(1,m) + \tfrac{\lambda^{\mathrm{d}}}{\Lambda_2} H^{\mathrm{d}}_{1m}[w] + \tfrac{1}{\Lambda_2}\Big(\textstyle\sum_{\substack{n=1\\ n\neq m}}^{M} q_{mn} w(1,n)\Big) + \tfrac{\nu_m}{\Lambda_2} w(1,m), \\
H^{\mathrm{c}}_{0m}[w] &= \max\{w(0,m),\, w(1,m) - C(1/\eta^{c}, m)\}, \\
H^{\mathrm{d}}_{1m}[w] &= \max\{w(1,m),\, w(0,m) + D(\eta^{d}, m)\}.
\end{aligned}$$

Let w_{0m} and w_{1m} be the optimization variables corresponding to $w(0,m)$ and $w(1,m)$. Define the auxiliary optimization variables $h^{\mathrm{c}}_{0m} = H^{\mathrm{c}}_{0m}[w]$ and $h^{\mathrm{d}}_{1m} = H^{\mathrm{d}}_{0m}[w]$. We can then find w by solving

$$\begin{aligned} \text{minimize} \quad & \textstyle\sum_{m=1}^{M} (w_{0m} + w_{1m}) \\ \text{subject to} \quad & (\lambda^{\mathrm{c}} + \bar{q}_m + \gamma) w_{0m} = \lambda^{\mathrm{c}} h^{\mathrm{c}}_{0m} + \textstyle\sum_{\substack{n=1 \\ n\neq m}}^{M} q_{mn} w_{0n} \\ & (\lambda^{\mathrm{d}} + \bar{q}_m + \gamma) w_{1m} = \lambda^{\mathrm{d}} h^{\mathrm{d}}_{1m} + \textstyle\sum_{\substack{n=1 \\ n\neq m}}^{M} q_{mn} w_{1n} \\ & h^{\mathrm{c}}_{0m} \geq w_{0m}, \qquad h^{\mathrm{c}}_{0m} \geq w_{1m} - C(1/\eta^c, m) \\ & h^{\mathrm{d}}_{1m} \geq w_{1m}, \qquad h^{\mathrm{d}}_{1m} \geq w_{0m} + D(\eta^d, m). \end{aligned} \tag{15}$$

The following result combines the value functions v and w to produce the optimal value function V.

PROPOSITION 9. *For the rental model, the value function V which satisfies* (9) *is given by*

$$V(k,\ell,m) = v_{\lceil k \rceil \ell m} + (w_{1m} - w_{0m})(k - \lceil k \rceil) \qquad \textit{for } k \in [0, K-\ell], \tag{16a}$$

where $v_{k\ell m}$ for all k,ℓ,m with k integer are obtained by solving the linear optimization problem (13)*, and w_{0m}, w_{1m} for all m are obtained by solving the linear optimization problem* (15)*.*

For the mid-charge model, the value function V which satisfies (9) *is given by*

$$V(k,\ell,m) = \begin{cases} v_{\lfloor k \rfloor \ell m} + (w_{1m} - w_{0m})(k - \lfloor k \rfloor) & \textit{for} \quad k \in [\ell, K/2], \\ v_{\lceil k \rceil \ell m} + (w_{1m} - w_{0m})(k - \lceil k \rceil) & \textit{for} \quad k \in [K/2, K-\ell], \end{cases} \tag{16b}$$

where $v_{k\ell m}$ is obtained from (13) *with the mid-charge model.*

Problem (15) and Proposition 9 are stated for a general transition rate matrix Q. If the energy and regulation price processes are independent, the size of (15) can be further reduced by formulating the problem on the states of the energy market only, using the transition rates of Q^{en}.

6. Value of Stacking Services

In this section, we define the value of stacking services as the value of the optimal policy operating jointly in the two markets, minus the value of the best static allocation policy. A static allocation policy sets in advance the storage capacity dedicated to the energy market and the storage capacity dedicated to the regulation market, and then operates the dedicated capacity optimally in each market. Thus, the static allocation divides up the storage capacity in two distinct resources, while with stacked services the optimal policy provides two functions with the same resource. Here we present structural results for static allocation, and refer to Section 8 for numerical results.

6.1. Value in the Energy Market as a Function of the Assigned Capacity

Let x be the storage capacity statically assigned to the energy market. Let $V_x^{\mathrm{en}}(k,m)$ denote the optimal value of operating in the energy market with assigned capacity x, starting from state (k,m). V_x^{en} can be computed by solving (9) with $K=x$, $\lambda^{\mathrm{r}}=0$, $\ell=0$, and then setting $V_x^{\mathrm{en}}(k,m)=V(k,0,m)$. Actually, V_x^{en} is linear in x, due to the linearity of C and D in the decision a:

Proposition 10. *The value of operating in the energy market with assigned storage capacity x is linear in x, that is, $V_x^{en}=xV_1^{en}$.*

Therefore, we only need to solve for V_1^{en}. Due the optimality of bang-bang policies established in Proposition 5, this problem was in fact already encountered and can be solved via the formulation (15). Hence we have

$$V_x^{\mathrm{en}}(0,m)=xw_{0m},\qquad V_x^{\mathrm{en}}(x,m)=xw_{1m} \tag{17}$$

where w_{0m}, w_{1m} are computed by solving (15).

Special Case. It is insightful to establish V_x^{en} in closed-form for a simple case. We assume the regulation and energy prices are independent processes, and the energy price is modeled using only two states, representing a low price and a high price state. We assume that m is initialized on its steady-state distribution (as determined by Q) and calculate $\bar{V}_x^{\mathrm{reg}}=\sum_{m=1}^{M}q_mV_x^{\mathrm{reg}}(0,m)$ where q_m (not to be confused with $\bar{q}_m$) is the steady-state probability of state m.

Proposition 11. *Suppose $Q^{en}=\begin{bmatrix}-q_{12} & q_{12}\\ q_{21} & -q_{21}\end{bmatrix}$ with energy prices $p_1<p_2$. Then we have*

$$\bar{V}_x^{en}=x\frac{\lambda^c}{\gamma}\left(\frac{q_{21}}{q_{12}+q_{21}}\right)\left[\frac{\lambda^d q_{12}D(\eta^d,2)-[\lambda^d(q_{12}+\gamma)+\gamma(q_{12}+q_{21}+\gamma)]C(1/\eta^c,1)}{(\lambda^c+\gamma)(q_{21}+\gamma+\lambda^d)+(\lambda^d+\gamma)q_{12}}\right]^+. \tag{18}$$

To interpret (18), we define the quantities

$$\Delta^{\mathrm{en}}:=D(\eta^{\mathrm{d}},2)-C(1/\eta^{\mathrm{c}},1),\qquad\qquad T^{\mathrm{en}}:=1/q_{12}+1/q_{21}. \tag{19}$$

Δ^{en} is the "efficiency-adjusted" energy price spread, and T^{en} is the expected duration of a low-price, high-price cycle. In steady-state, if operations were unrestricted, we could expect an average revenue stream $\Delta^{\mathrm{en}}/T^{\mathrm{en}}$ and an expected discounted cumulated revenue $\int_0^\infty(\Delta^{\mathrm{en}}/T^{\mathrm{en}})e^{-\gamma t}dt=\gamma^{-1}\Delta^{\mathrm{en}}/T^{\mathrm{en}}$, assuming charge-discharge operations are profitable ($\Delta^{\mathrm{en}}>0$). When permissions to charge and discharge arrive as Poisson processes, not every low-price, high-price cycle can be captured. An upper bound on the fraction of cycles that can be captured is found to depend on the ratios $\lambda^{\mathrm{c}}/q_{12}$ and $\lambda^{\mathrm{d}}/q_{21}$. This is expressed in the following bound.

Proposition 12. *Regarding* (18) *we have* $\bar{V}_x^{en}\le\frac{x}{\gamma}\left[\frac{\Delta^{en}}{T^{en}}\right]^+\min\left\{1,\left[\left(\frac{\lambda^c}{q_{12}}\right)^{-1}+\left(\frac{\lambda^d}{q_{21}}\right)^{-1}\right]^{-1}\right\}.$

This bound is simpler to interpret than (18). It says that the potential profit from price volatility can be captured if the charge and discharge permission rates dominate the price transition rates.

6.2. Value in the Regulation Market as a Function of the Assigned Capacity

Let y be the capacity statically assigned to the regulation market. Let $V_y^{\text{reg}}(\ell, m)$ denote the optimal value of operating in the regulation market with assigned capacity y, starting from state (ℓ, m). We set $\lambda^{\text{c}} = \lambda^{\text{d}} = 0$ in (7) and (9), so that V_y^{reg} can be computed by solving (9) with $K = \lfloor y \rfloor$, $\lambda^{\text{c}} = \lambda^{\text{d}} = 0$. The state k is set to k_0 and remains fixed at that value. To maximize the capacity available for regulation, with the rental model we set $k_0 = 0$, and with the mid-charge model we set $k_0 = \lfloor y/2 \rfloor$. The problem reduces to solving (13) with $K = \lfloor y \rfloor$, $\lambda^{\text{c}} = \lambda^{\text{d}} = 0$, $k = k_0$ and then setting $V_y^{\text{reg}}(\ell, m) = v_{k_0,\ell,m}$.

Proposition 13. *$V_y^{reg}(\ell, m)$ is nondecreasing in the assigned capacity y.*

6.2.1. Results for the Rental Model. We say that a function $f(\ell)$ of $\ell \in \mathbb{N}$ is concave in ℓ if $f(\ell)$ has nonincreasing differences, which means that $f(\ell+2) - f(\ell+1) \leq f(\ell+1) - f(\ell)$. We say that $f(\ell)$ is nonincreasing if $f(\ell+1) \leq f(\ell)$. We say that $f(\ell)$ is concave nonincreasing (CNI) if it is concave and nonincreasing.

Proposition 14. *Consider the rental model with $\lambda^c = \lambda^d = 0$ and $k = 0$. For each fixed m, the value function $V_y^{reg}(\ell, m)$ is CNI in ℓ.*

Proposition 14 relies on two lemmas from Lippman and Stidham (1977) routinely used in the literature, see e.g. Gans and Savin (2007). Our statements distinguish between concave and CNI.

Lemma 1. *Let $f(\ell)$ be concave over $\ell \in \{0, \ldots, K\}$. Then $g(\ell) = \ell f(\ell-1) + (K-\ell) f(\ell)$ is concave. If $f(\ell)$ is nonincreasing then $g(\ell)$ is nonincreasing. If $f(\ell)$ is CNI then $g(\ell)$ is CNI.*

Lemma 2. *Let $f(\ell)$ be concave over $\ell \in \mathbb{N}$. Then $h(\ell) = \max\{f(\ell), r + f(\ell+1)\}$ is concave. Furthermore, if $f(\ell)$ is CNI and $r \geq 0$, then $g(\ell)$ is CNI.*

Proposition 14 implies that the regulation acceptance policy has a threshold policy structure:

Proposition 15. *Consider the rental model with $\lambda^c = \lambda^d = 0$ and $k = 0$. For each state m, there exists a threshold $\ell_m^* \in \{0, \ldots, K\}$ such that at each regulation request arrival while being in state m, it is optimal to accept the request if $\ell < \ell_m^*$ and to reject the request if $\ell \geq \ell_m^*$.*

We conjecture that $V_y^{\text{reg}}(\ell, m)$ is concave in the integer capacity y, but we do not have a formal proof. What is simpler to establish is an upper bound on $V_y^{\text{reg}}(\ell, m)$. Recall that r is the column vector with elements r_m defined by (6a) and e_m is the m-th unit vector.

Proposition 16. *Under the rental model, $V_y^{reg}(\ell, m) \leq \dfrac{\lambda^r}{\gamma} e_m^\top (I - \gamma^{-1} Q)^{-1} r$.*

6.2.2. Result for the Mid-Charge Model. With the mid-charge model, it is optimal to immediately maximize the regulation commitment, since the expected payments from regulation are proportional to the committed capacity, and the regulation price is the ongoing price. The commitment is never modified, going forward ($u = 0$). If we start with the optimal commitment $\ell^* = \lfloor y/2 \rfloor$, using r_m in (6b) we have $V_y^{\text{reg}}(\lfloor y/2 \rfloor, m) = \lfloor y/2 \rfloor r_m = \lfloor y/2 \rfloor e_m^\top (\gamma I - Q)^{-1} \varrho$.

6.3. Optimal Static Allocation

The problem of finding the best static capacity allocation can be described as

$$\max \quad [V_x^{\text{en}}(k_0, m) + V_y^{\text{reg}}(\ell_0, m)] \quad \text{subject to} \quad x + y \leq K, \quad x \geq 0, \quad y \in \mathbb{N}, \tag{20}$$

where $k_0 = \ell_0 = 0$ for the rental model and $k_0 = \ell_0 = \lfloor y/2 \rfloor$ for the mid-charge model. Since V_x^{en} is increasing in x, we have $x = K - y$, thus the problem is one-dimensional and can be solved by enumeration over $y = 0, 1, \ldots, \lfloor K \rfloor$. The previous subsections have clarified how V_x^{en} and V_y^{reg} can each be computed based on (13). In our numerical work (Section 8.4), instead of using a single m we average the objective over the background process states using the steady-state probabilities.

7. Value of Random Permissions versus Scheduled Transactions

The market model with random permission arrivals can be compared to a market with scheduled transactions. In a market with scheduled transactions, decisions to operate in the energy market or assign storage capacity to regulation can take place at any time. However, operations need to be scheduled in advance. The rationale is that a utility company or system operator cannot allow arbitrary power transfers for security reasons, see e.g. Tian and Gross (1998). To model this situation, we assume that a fixed lead time $L > 0$ separates a decision from its actual implementation. Decisions are binding: they create a commitment to execute scheduled actions. Actually, informational delays are unavoidable even with existing "real-time" energy markets. For instance, the 5-minute locational marginal prices posted in real time are the prices of the last 5 minutes, not the prices of the next 5 minutes, since the exact demand of the next 5 minutes is unknown.

To derive a model with scheduled transactions from the model with random permissions, we formally set to infinity the rates of charging permission arrivals, discharging permission arrivals, and regulation request arrivals, so that storage decisions can take place continuously at any time.

We illustrate how the model should change in the case of the rental model. The state sojourn times of the background process, and the durations of the capacity reservations from previously accepted regulation service requests, are still exponentially distributed. As a result, the conditional probability of an event at time $t + \delta$ given the information at time t does not vary with the passage of time (that is, as t and $t + \delta$ roll forward), as long as the state does not change. Hence in the infinite horizon setting, the optimality of decisions conditional on being in a state does not change with the passage of time, and without loss of generality, the search for an optimal policy can be restricted to policies that make decisions at event arrival times only. The events to consider are reduced to the background process state changes and the end-of-service regulation events. At an event arrival time, say at time t, one can commit to an action executed at some time $t' \geq t + L$.

The concern with delay in execution is the expected suboptimality of the executed decision at time t', compared to the decision adapted to the state at time t' that could be selected in the absence of lead time. Mathematically however, it suffices to replace the lump-sum reward associated to state m by the expected lump-sum reward at state n, multiplied by a discount $e^{-\gamma L}$, taking the expectation over n where n is the random state L units of time later when starting from state m.

A special consequence of lead times is that when L is sufficiently large, the background process will attain its stationary distribution from any starting state m. Consequently, the expected energy price in L units of time in the future has always the same value, and it becomes impossible to perform temporal arbitrage in the energy market. The policy that dedicates the entire capacity to regulation will thus furnish a lower bound on the optimal value of the problem and become tight for L large enough.

The details of the model with lead times are in Appendix D. We denote by $\widetilde{V}(k,\ell,m)$ the corresponding optimal value function. The two next propositions refine the discussion above.

Proposition 17. *$\widetilde{V}(k,\ell,m)$ is nonincreasing in L for any starting state (k,ℓ,m).*

Proposition 18. *There exists a critical lead time $L^* \geq 0$ beyond which optimal policies do not charge in the energy market and only operate in the regulation market. The critical L^* is nonincreasing in the discount rate γ and nondecreasing in the storage efficiency parameters η^c, η^d.*

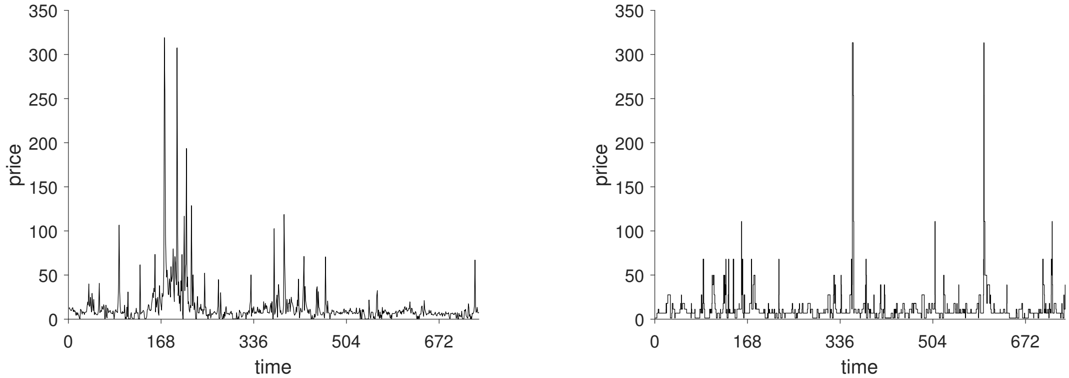
8. Numerical Results

This section presents our numerical work. We define our numerical setup in §8.1. We present the value function for the rental model in §8.2 and for the mid-charge model in §8.3. The value of stacking services versus static allocation is studied numerically in §8.4. The robustness of the policies with respect to the underlying stochastic model is investigated via backtesting experiments on out-of-sample empirical data in §8.5. The sensitivity to market parameters is studied numerically in §8.6. Additional details and results can be found in Appendices E, F, G.

8.1. Parameters for the Reference Case

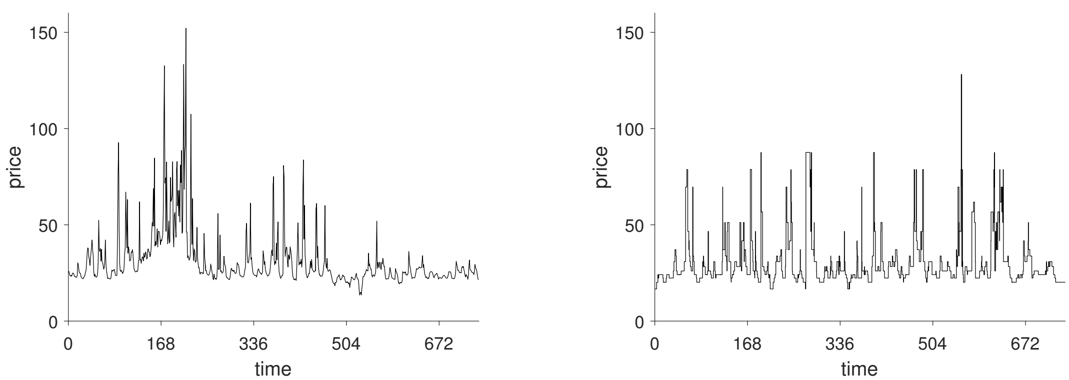
We consider a storage device with energy capacity $K = 5$ (divided up in blocks of 1 MWh) and efficiency $\eta^{\mathrm{c}} = \eta^{\mathrm{d}} = 0.95$. The round-trip efficiency $\eta^{\mathrm{c}}\eta^{\mathrm{d}} = 0.9025$ is typical for battery technology,

Figure 3 Sample path of the 12-state continuous-time Markov chain model for the regulation market price, calibrated using January 2017 hourly regulation price data (time in hours).



(a) Historical Data (January 2017) (b) One Simulated Sample Path

Figure 4 Sample path of the 313-state continuous-time Markov chain model for the energy market price, calibrated using the first quarter of 2017 hourly real-time energy price data (time in hours).

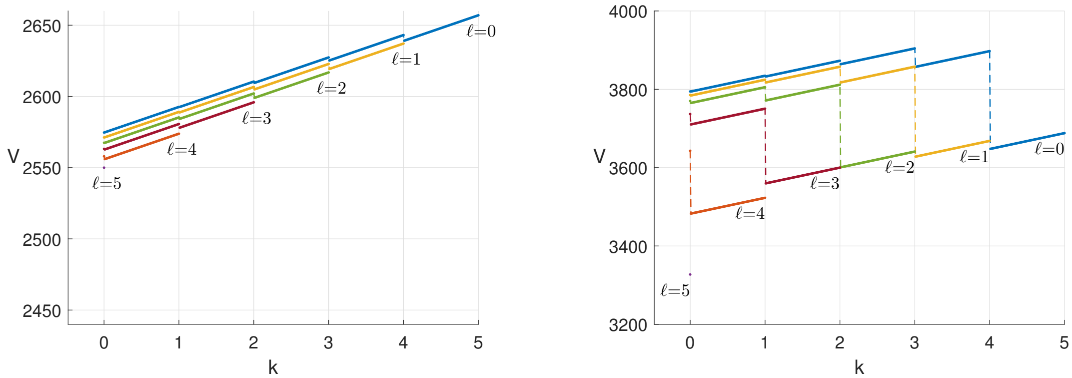


(a) Historical Data (January 2017) (b) One Simulated Sample Path

see Carnegie et al. (2013). The arrival rates of charge and discharge permissions are $\lambda^{\mathrm{c}} = \lambda^{\mathrm{d}} = 1.5$ (per hour). The reference case uses the rental model for regulation. The arrival rate of regulation service requests is $\lambda^{\mathrm{r}} = 0.5$ (per hour). The average duration of regulation service for one block is $1/\mu = 4$ (hours). The discount rate is $\gamma = 0.01$.

Historical PJM data from 2017 is used to fit the price processes and calibrate the matrix Q of the background process. The details on the data sets and the calibration method are in Appendix E. We have assigned to each state one of 24 time indices to model time-varying transition probabilities. In total, the background process that represents the combination of time indices, energy and regulation prices has $M = 3756$ states. For a comparison of January empirical data to a price sample path generated with the calibrated model, see Figure 3 for the regulation price and Figure 4 for the energy price.

Figure 5 Value function $V(k,\ell,m)$ for the rental model, as a function of the feasible states (k,ℓ).



(a) At state $m=1$ (lowest prices) (b) At state $m=3756$ (highest prices)

8.2. Value Function for the Rental Model

The reference case is solved via the 2-phase linear-optimization method developed in Section 5. Figure 5 presents the exact solution $V(k,\ell,m)$ as a function of k and ℓ for a fixed state m. We have also checked that refining the discretization of k prescribed in Section 5 increases the computational times without changing the optimal value function at integer k.

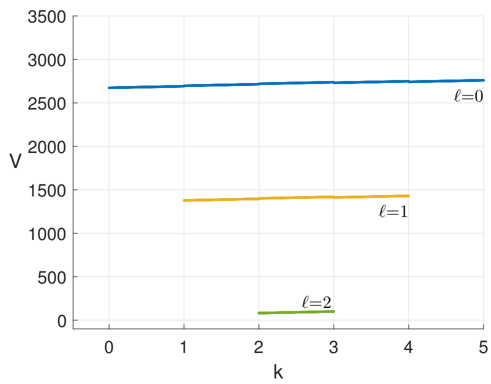
Figure 5(a) plots V at $m=1$ (lowest prices at energy time index 0) and Figure 5(b) plots V at $m=3756$ (highest prices at energy time index 23). The piecewise-linear discontinuous structure over k discussed in Propositions 6 and 7 is exhibited on these figures. We recall that the discontinuities at integer k are due to the sudden loss of a capacity block available for regulation services. Discontinuities are expected to be more pronounced when the regulation prices are higher. The slope of the affine parts between integer k are identical given the state m. Because Q has been constructed with the two markets evolving independently, the slopes actually depend on the energy market state only. Figure 5 also confirms visually Proposition 4 on V being nonincreasing in ℓ if all other variables are fixed.

Solution times are discussed in Appendix F where we solve problems of increasing size. An illustration of the behavior of the optimal policy derived from the value function V is provided in Appendix G for a particular sample path of the background process.

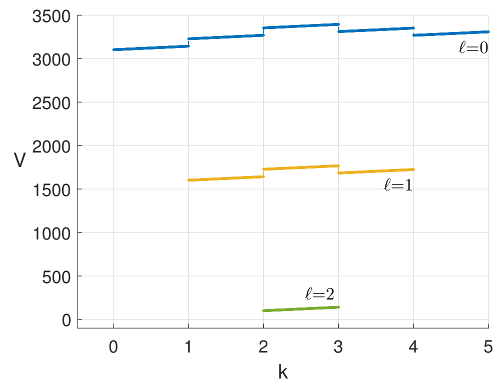
8.3. Value Function with the Mid-Charge Model

We solve a modification of the reference case that uses the mid-charge model for regulation. Figure 6 presents the exact solution $V(k,\ell,m)$ as a function of k and ℓ for a fixed state m. In particular, when $\ell=2$ blocks are reserved for up-regulation (and thus 2 blocks are also reserved for down-regulation), the energy level used for energy transactions can only vary over $k\in[2,3]$, reflecting the

Figure 6 Value function $V(k,\ell,m)$ for the mid-charge model, as a function of the feasible states (k,ℓ).



(a) At state $m=1$ (lowest prices)



(b) At state $m=3756$ (highest prices)

constraint that excursions from regulation could bring down the physical stored energy to $k-2=0$ or up to $k+2=5=K$. Compared to the rental model, there is a larger difference of value function as ℓ varies, that is explained by the magnitudes of the reward for accepting regulation and cost for ending regulation (recall from Section 3.3.2 that the accounting method to calculate the lump-sum payments uses an infinite duration of service.)

If the capacity K is set to 6 and the problem is solved again, the maximal possible value for ℓ is 3. The feasible space for k at $\ell=3$ is then reduced to a single point at $k=3$, and the top stair step described by V for $k\in[2,3]$ at fixed (ℓ,m) in Figure 6 is reduced to an isolated point at $k=3$.

8.4. Value of Dynamic Versus Static Allocation of Capacity

Let V^{static} denote the optimal value of (20), where in the objective m is averaged over the steady-state probabilities q_m of the background process. This value can be compared to $V^{\text{dyn}}=\sum_m q_m V(0,0,m)$, where V is the value function satisfying (9). The difference $V^{\text{dyn}}-V^{\text{static}}$ represents the improvement from optimizing in the class of dynamic allocation policies, instead of allocating capacity statically.

We first consider the reference case. We have $V^{\text{dyn}}=2639.2$, using $V(0,0,m)$ solved in §8.2. For the static allocation, we find that $V^{\text{stat}}=2535.6$, attained by allocating the full capacity to the regulation market (and thus never using the energy market). The relative improvement, defined as $(V^{\text{dyn}}-V^{\text{static}})/V^{\text{static}}$, is thus 4.1%.

Next, we discuss cases where the energy market is more attractive, relatively to the regulation market. For simplicity, we model the energy market using a 2-state continuous-time Markov chain, as in the special case of Section 6. The energy prices associated to the two states are p_1^{en} and p_2^{en}.

Table 1 reports results from a numerical study conducted by progressively increasing the efficiency-adjusted spread Δ^{en}, defined in (19) in terms of p_1^{en} and p_2^{en}. The energy spread is a good

Table 1 Value of stacking services, in various energy market contexts.

Instance			Optimal Static Allocation					With Stacked Services	
p_1^{en}	p_2^{en}	Δ^{en}	x	y	V_x^{en}	V_y^{reg}	V^{static}	V^{dyn}	Improvement
25	30	2.2	0	5	0.0	2535.6	2535.6	2535.6	0.0%
25	50	21.2	1	4	154.7	2396.9	2551.6	2738.2	7.3%
25	70	40.2	2	3	617.5	2126.6	2744.1	3134.8	14.2%
25	90	59.2	2	3	925.6	2126.6	3052.2	3664.3	20.1%

Storage parameters: $K=5$, $\eta^{\mathrm{c}}=\eta^{\mathrm{d}}=0.95$. Discount rate: $\gamma=0.01$.
Market parameters: $\lambda^{\mathrm{c}}=1.5$, $\lambda^{\mathrm{d}}=1.5$, $\lambda^{\mathrm{r}}=0.5$, $\mu=0.25$, $q_{12}^{\mathrm{en}}=0.1659$, $q_{21}^{\mathrm{en}}=0.3095$.

metric for the attractiveness of the energy market. The numerical results suggest that dynamic allocation policies do improve expected profits from operations, especially in the cases where the optimal static allocation would assign significant capacity to both markets.

8.5. Practical Implementation and Backtesting

Our numerical work uses processes calibrated from 2017 data. In this section we backtest the policy that uses the value function optimized on the calibrated model, by executing the policy on a real empirical scenario.

The value of a policy on a single scenario is only loosely related to the expected value of the policy, which is our optimality criterion. Furthermore, market conditions may change. Therefore, we repeat the backtesting experiment on several empirical scenarios of 30 days starting from the first day of a month. The 30-day time window is motivated by our choice of discount factor γ: we have $\int_0^{720} e^{-\gamma t}dt = 99.9253$ and $\int_{720}^{\infty} e^{-\gamma t}dt = 0.0747$, indicating that we neglect roughly 0.0747% of the total discounted return by truncating the return after 30 days.

The empirical scenarios come from the historical hourly real-time energy and regulation prices from PJM. We simulate a background state transition at each hour, even though the optimization model assumed stochastic sojourn times. The same rule used to associate time stamps and price levels to a background state m in order to estimate the background process transition matrix (see E) is used to convert observed hours and prices to a state m of the calibrated model. The decisions are obtained by solving (10), (11), (12) with the payoff functions evaluated at the observed prices and the value function V evaluated at m.

The results are reported in Table 2. The row "Dynamic policy" gives the empirical performance of the policy optimal on the 2017-calibrated process but simulated on the 2018 data. The row "Static policy" gives the empirical performance of the best static policy on the 2017-calibrated process; per the results of §8.4 the capacity is fully allocated to the regulation market. For all but one of the out-of-sample tests in 2018, the relative improvement of the dynamic policy over the static policy is significantly superior to the estimate stated in §8.4.

Table 2 Backtesting of policies optimized on 2017-data calibrated model: Empirical values.

	In-sample	Out-of-sample					
Scenario	Jan-2017	Jan-2018	Feb-2018	Mar-2018	Apr-2018	May-2018	Jun-2018
Static policy†	2122.6	23606.4	2470.7	2906.8	3605.6	2845.0	6053.1
Dynamic policy†	2145.8	28166.3	2325.0	3051.0	4117.9	3814.2	7452.2
Improvement	1.09%	19.32%	-5.90%	4.96%	14.21%	34.07%	23.11%

† discounted cumulated payoff over 30 days from 1st of the month, using discount $\gamma = 0.01$.

The return for January 2018 is abnormally high. Regulation and energy prices are in fact exceptionally high in the Jan-2018 scenario. This could have been caused by a disruption of the gas supply to industrial consumers, which includes fast-ramping gas-fueled power generators. See an article by EIA staff (2018). This episode illustrates limitations in the scope of the calibrated stochastic model, and the need for being aware of the normal market contexts in which it applies.

8.6. Random Opportunities to Act Versus Anytime Decisions with Lead-Time

We compare markets with random opportunities to act to markets with anytime decisions but subject to lead times. The comparisons are done given an ergodic background process, using its steady-state distribution as the initial state distribution. Different background processes are considered, formed by combining the regulation price model of the reference case (§8.1) to the simple parametric two-state energy price model described in §8.4.

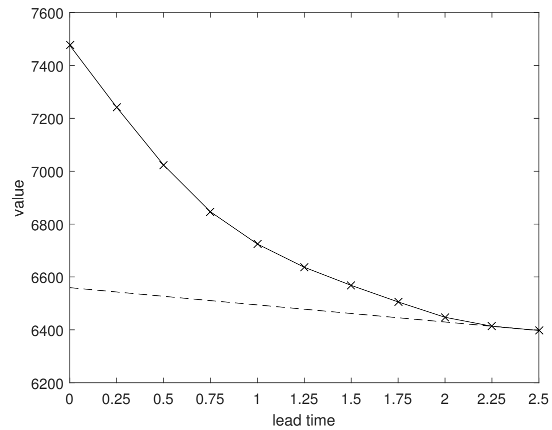
Let $V^{\lambda^c,\lambda^d,\lambda^r}$ denote the optimal value of the control problem in the market with random permissions when the arrival rates are λ^c, λ^d, λ^r, with initial storage state $k=0$, $\ell=0$. Let $\widetilde{V}^L$ denote the optimal value of the control problem in the market with lead time when the lead time is L, with initial storage state $k=0$, $\ell=0$.

8.6.1. Sensitivity to Lead Time. Proposition 17 states that $\widetilde{V}^L$ is nonincreasing in L. This is illustrated in Figure 7, which shows the sensitivity of the value of $\widetilde{V}^L$ with respect to the lead time L, using the 2-state energy market model with $p_1^{\mathrm{en}} = 25$ and $p_2^{\mathrm{en}} = 50$. We also display the value of the lower bound $\bar{V}^L$ calculated by assigning the entire capacity to regulation and accepting any incoming regulation request at any current price. The difference between $\widetilde{V}^L$ and $\bar{V}^L$ represents an upper bound on the value of being able to access both the energy and regulation markets, versus the regulation market only. The lower bound is ultimately attained, confirming the result in Proposition 18 that lead times ultimately erase the value of energy market access, irrespectively of the volatility of energy prices.

8.6.2. Iso-Value Curves in the Rate Parameter Space. For a given lead time L, define the set $\Upsilon(L)$ of arrival rates under which the value $V^{\lambda^c,\lambda^d,\lambda^r}$ is greater or equal to $\widetilde{V}^L$:

$$\Upsilon(L) = \{(\lambda^{\mathrm{c}}, \lambda^{\mathrm{d}}, \lambda^{\mathrm{r}}) \in \mathbb{R}^3_+ : V^{\lambda^c,\lambda^d,\lambda^r} \geq \widetilde{V}^L\}. \tag{21}$$

Figure 7 Variation of the optimal value $\widetilde{V}^L$ with the lead time L of the market with scheduled transactions. The lower bound $\underline{V}^L$ (dashed curve) is produced by dedicating the entire capacity to regulation.



Thus, a market with random permissions at rates $(\lambda^{\mathrm{c}}, \lambda^{\mathrm{d}}, \lambda^{\mathrm{r}}) \in \Upsilon(L)$ is at least as valuable as the market with scheduled transactions under lead time L, assuming the respective optimal policies are employed in each market.

From Proposition 1, $V^{\lambda^{\mathrm{c}},\lambda^{\mathrm{d}},\lambda^{\mathrm{r}}}$ is nondecreasing in the arrival rates. Along with $\widetilde{V}^L$ nonincreasing in L, this translates into the following monotonicity properties:

(i) $(\lambda_1^{\mathrm{c}}, \lambda_1^{\mathrm{d}}, \lambda_1^{\mathrm{r}}) \in \Upsilon(L) \quad \Rightarrow \quad (\lambda_2^{\mathrm{c}}, \lambda_2^{\mathrm{d}}, \lambda_2^{\mathrm{r}}) \in \Upsilon(L)$ for any $\lambda_2^{\mathrm{c}} \geq \lambda_1^{\mathrm{c}},\ \lambda_2^{\mathrm{d}} \geq \lambda_1^{\mathrm{d}},\ \lambda_2^{\mathrm{r}} \geq \lambda_1^{\mathrm{r}}$.

(ii) $L < L' \quad \Rightarrow \quad \Upsilon(L) \subseteq \Upsilon(L')$, that is, the set-valued mapping Υ is "expansive" in L.

For the purpose of visualization, we reduce the dimension of $\Upsilon(L)$ by imposing $\lambda^{\mathrm{c}} = \lambda^{\mathrm{d}}$ (symmetry between the charge and discharge permission arrival rates). We can then represent the boundary of $\Upsilon(L)$ in the plane $(\lambda^{\mathrm{c}}, \lambda^{\mathrm{r}})$. The boundary of $\Upsilon(L)$ in the plane $(\lambda^{\mathrm{c}}, \lambda^{\mathrm{r}})$ with $\lambda^d = \lambda^c$ is the curve

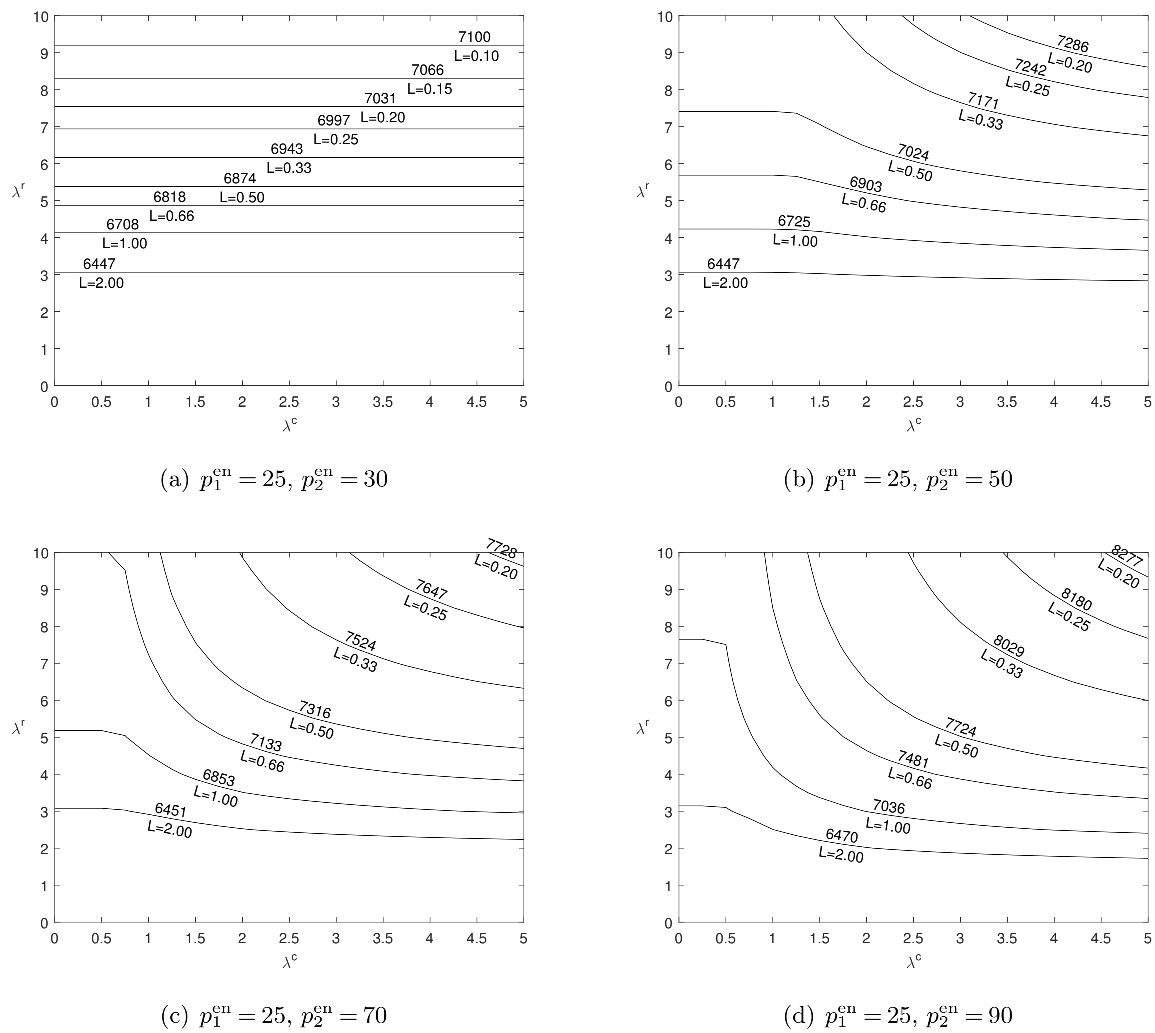
$$B(L) := \{(\lambda^c, \lambda^r) \in \mathbb{R}_+^2 : V^{\lambda^c,\lambda^c,\lambda^r} = \widetilde{V}^L\} = \mathrm{Proj}_{13}\left[\partial[\Upsilon(L) \cap \{(\lambda^{\mathrm{c}}, \lambda^{\mathrm{d}}, \lambda^{\mathrm{r}}) \in \mathbb{R}_+^3 : \lambda^{\mathrm{c}} = \lambda^{\mathrm{d}}\}]\right],$$

$$\mathrm{Proj}_{13}(A) := \{(x_1, x_3) \in \mathbb{R}^2 : (x_1, x_2, x_3) \in A\}. \tag{22}$$

Figure 8 depicts the "iso-value" curves $B(L)$ obtained with various lead times L from 0.1 to 2 (hours). The parameters of the energy market model that are used for the background process are given in the captions of subfigures 9(a) to 9(d); the parameter values are those also used in Table 1.

In Figure 9(a), the efficiency-adjusted energy price spread is insufficient to warrant operations in the energy market. Therefore, the capacity is entirely dedicated to regulation market operations, and the rates of charge and discharge permissions $\lambda^{\mathrm{c}} = \lambda^{\mathrm{d}}$ are irrelevant. It can be seen that as the lead-time L is decreased, the optimal value $\widetilde{V}^L$ of the control problem increases (the numerical values are indicated on the curves). This improvement is matched by a corresponding increase in the arrival rate λ^{r} of regulation requests in the market with random permissions.

Figure 8 Iso-value curves $B(L)$ plotted in the plane $(\lambda^{\mathrm{c}}, \lambda^{\mathrm{r}})$ for the values of $\widetilde{V}^L$ corresponding to various lead times L. The values of $\widetilde{V}^L$ and L are indicated on the corresponding curves. Subfigures refer to various energy market price spreads.



(a) $p_1^{\mathrm{en}} = 25$, $p_2^{\mathrm{en}} = 30$

(b) $p_1^{\mathrm{en}} = 25$, $p_2^{\mathrm{en}} = 50$

(c) $p_1^{\mathrm{en}} = 25$, $p_2^{\mathrm{en}} = 70$

(d) $p_1^{\mathrm{en}} = 25$, $p_2^{\mathrm{en}} = 90$

As we move from Figure 9(a) to Figure 9(d), the energy price spread is increased, making operations in the energy market more attractive. When the energy market is employed by the optimal policy, increasing λ^{c} offers an alternative to increasing λ^r for matching a decrease in L.

From this study, we conclude to the existence of arrival rates ensuring an equivalence, in terms of expected value of storage operations, between the markets with random opportunities to act, and the markets with continuous access but decisions subject to a lead time.

9. Conclusion

Motivated by the development of electricity markets accessible to electric storage, we have considered the management of electric storage with stacked services, and in particular, the joint utilization of the storage capacity for providing regulation services and for carrying out price arbitrage in the energy market. Even though regulation is driven by real-time control engineering considerations, approximations are possible that lead to problems that can be analyzed using established operations research methodologies. We have clarified the structure of the optimal value function for this type of problems, considering two variants for describing the constraints of regulation service commitments and the payment schemes.

The theoretical results of this paper are useful to derive a numerical solution algorithm. There is no generic method to solve Markov decision processes with continuous state variables exactly, but in this work we could exploit model properties to determine exact value functions. We use exact solutions to perform sensitivity analysis numerically, which could not have been done reliably otherwise. Our goal was to compare the value of the storage asset under different market conditions. Such discussions are expected to have ramifications in market design.

The value of service stacking relates to the value of being able to dynamically assign storage capacity to support multiple functions. We have confirmed that if operations in the energy market and the regulation market are both individually profitable to start with, then there is extra value in co-optimizing operations and dynamically sharing the capacity, as compared to allocating the capacity statically between the two functions. In our numerical tests, the extra value was higher when the optimal static allocation would have divided up the capacity in parts of comparable size. It was also high in our policy backtesting experiments.

We have modeled market access using point processes. The arrival rates should reflect the characteristics of the markets. We have provided tools to study how the storage asset value varies as a function of these rates. Recently, there have been discussions between battery storage investors and system operators regarding the magnitude and duration of the regulation signal excursions away from a reference level. Modifications of engineering aspects impacting the regulation market are likely (FERC EL18-87-000). We are optimistic that our technical framework has enough flexibility to be adapted to a wide range of future contexts and can be useful to provide insights. An aspect that we have emphasized in our numerical work is the equivalence, in terms of asset value and by finding suitable pairings of parameters, between markets where actions can only happen at random times, and markets where actions are permitted at any time but with informational delay. Both market types allow a system operator to exercise a specific form of supervisory control.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. 1610825 and 1610302.

References

Akcay Y, Balakrishnan A, Xu S (2010) Dynamic assignment of flexible service resources. *Production and Operations Management* 19(3):279–304.

Akhil A, Huff G, Currier A, Kaun B, Rastler D, Chen S, Cotter A, Bradshaw D, Gauntlett W (2015) DOE/EPRI electricity storage handbook in collaboration with NRECA. Technical Report SAND2015-1002, Sandia National Laboratories.

Al-Gwaiz M, Chao X, Wu O (2016) Understanding how generation flexibility and renewable energy affect power market competition. *Manufacturing & Service Operations Management* 19(1):114–131.

Barbose GL (2017) U.S. renewables portfolio standards: 2017 annual status report. Technical Report LBNL-2001031, Lawrence Berkeley National Laboratory.

Bekiarov S, Emadi A (2002) Uninterruptible power supplies: classification, operation, dynamics, and control. *17th Annual IEEE Applied Power Electronics Conference and Exposition (APEC)*, volume 1, 597–604.

Berrada A, Loudiyi K, Zorkani I (2016) Valuation of energy storage in energy and regulation markets. *Energy* 115, Part 1:1109–1118.

Bhattacharya A, Kharoufeh J, Zeng B (2018) Structured storage policies for energy distribution networks. *IISE Transactions* 50(8):683–698.

Borsche T, Ulbig A, Koller M, Andersson G (2013) Power and energy capacity requirements of storages providing frequency control reserves. *2013 IEEE Power & Energy Society General Meeting*, 1–5.

Brémaud P (1981) *Point Processes and Queues: Martingale Dynamics* (New York, NY: Springer).

Buchholz P (1994) A class of hierarchical queueing networks and their analysis. *Queueing Systems* 15:59–80.

Carnegie R, Gotham D, Nderitu D, Preckel P (2013) Utility scale energy storage systems: Benefits, applications, and technologies. Technical report, State Utility Forecasting Group, Purdue University, West Lafayette, IN.

Cazalet E, Martini PD, Price J, Woychik E, Caldwel J (2016) Transactive energy models. *NIST Transactive Energy Challenge: Business and Regulatory Models Working Group* 1–44.

Chod J, Rudi N (2005) Resource flexibility with responsive pricing. *Operations Research* 53(3):532–548.

Chong K, Henderson S, Lewis M (2017) Two-class routing with admission control and strict priorities. *Probability in the Engineering and Informational Sciences* 1–16.

Cont R, Stoikov S, Talreja R (2010) A stochastic model for order book dynamics. *Operations Research* 58(3):549–563.

Denardo E (1967) Contraction mappings in the theory underlying dynamic programming. *SIAM Review* 9(2):165–177.

Donadee J (2013) Optimal operation of energy storage for arbitrage and ancillary service capacity: The infinite horizon approach. *2013 North American Power Symposium*, 1–6.

Dvorkin Y, Fernandez-Blanco R, Kirschen D, Pandzic H, Watson J, Silva-Monroy C (2017) Ensuring profitability of energy storage. *IEEE Transactions on Power Systems* 32(1):611–623.

EIA staff (2018) January's cold weather affects electricity generation mix in Northeast, Mid-Atlantic. Today in Energy, Jan 23, 2018, U.S. Energy Information Administration, URL `www.eia.gov/todayinenergy/detail.php?id=34632`.

Feldmann A, Whitt W (1998) Fitting mixtures of exponentials to long-tail distributions to analyze network performance models. *Performance evaluation* 31(3-4):245–279.

FERC EL17-64-000 (2017) Complaint of Renewable Energy Systems Americas and Invenergy Storage Development LLC. URL `https://elibrary.ferc.gov/idmws/common/opennat.asp?fileID=14560432`.

FERC EL18-87-000 (2018) Order rejecting tariff revisions. URL `https://elibrary.ferc.gov/idmws/common/OpenNat.asp?fileID=14859266`.

Forrester S, Zaman A, Mathieu J, Johnson J (2017) Policy and market barriers to energy storage providing multiple services. *The Electricity Journal* 30:50–56.

Gani J (1957) Problems in the probability theory of storage systems. *Journal of the Royal Statistical Society, Series B* 19(2):181–206.

Gans N, Savin S (2007) Pricing and capacity rationing for rentals with uncertain durations. *Management Science* 53(3):390–407.

Güler O, den Hertog D, Roos C, Terlaky T (1993) Degeneracy in interior point methods for linear programming: a survey. *Annals of Operations Research* 46:107–138.

Harrison J, Resnick S (1976) The stationary distribution and first exit probabilities of a storage process with general release rule. *Mathematics of Operations Research* 1(4):347–358.

Hu Y, Defourny B (2017) Optimal price-threshold control for battery operation with aging phenomenon: a quasiconvex optimization approach. *Annals of Operations Research* 1–28, URL `http://dx.doi.org/10.1007/s10479-017-2505-4`.

Kallenberg L (1994) Survey of linear programming for standard and nonstandard Markovian control problems. Part I: Theory. *ZOR – Mathematical Methods of Operations Research* 40:1–42.

Kim JH, Powell WB (2011) Optimal energy commitments with storage and intermittent supply. *Operations Research* 59(6):1347–1360.

Lai G, Margot F, Secomandi N (2010) An approximate dynamic programming approach to benchmark practice-based heuristics for natural gas storage valuation. *Operations Research* 58(3):564–582.

Lippman S (1975) Applying a new device in the optimization of exponential queueing systems. *Operations Research* 23(4):687–710.

Lippman S, Stidham S (1977) Individual versus social optimization in exponential congestion systems. *Operations Research* 25(2):233–247.

Mannor S, Simester D, Sun P, Tsitsiklis JN (2007) Bias and variance approximation in value function estimates. *Management Science* 53(2):308–322.

Megiddo N (1991) On finding primal- and dual- optimal bases. *ORSA Journal on Computing* 3(1):63–65.

Miller N, Zrebiec R, Delmerico R, Hunt G, Achenbach H (1996) A VRLA battery energy storage system for Metlakatla, Alaska. *Proc. of the 11th Annual Battery Conference on Applications and Advances*, 241–248.

Moazeni S, Defourny B (2018) Optimal control of energy storage under random operation permissions. *IISE Transactions* 50(8):668–682.

Moazeni S, Powell WB, Hajimiragha AH (2015) Mean-conditional value-at-risk optimal energy storage operation in the presence of transaction costs. *IEEE Transactions on Power Systems* 30(3):1222–1232.

Moran P (1954) A probability theory of dams and storage systems. *Aust. J. Appl. Sci.* 5:116–124.

Neuts M (1995) *Matrix-geometric solutions in stochastic models: An algorithmic approach* (Dover), revised edition.

Papier F, Thonemann U (2010) Capacity rationing in stochastic rental systems with advance demand information. *Operations Research* 58(2):274–288.

PJM staff (2017a) Ancillary service market results. URL `www.pjm.com/-/media/markets-ops/ancillary/synchronized-service/historical-ancillary-service-data-2017.ashx`.

PJM staff (2017b) Implementation and rationale for PJM's conditional neutrality regulation signals. URL `www.pjm.com/~/media/committees-groups/task-forces/rmistf/postings/regulation-market-whitepaper.ashx`.

Prabhu N (1998) *Stochastic Storage Processes* (Springer), second edition.

Pradhan P, Patil R (2016) Optimal battery operation in frequency regulation markets considering its health factors. *2016 IEEE/PES Transmission and Distribution Conference and Exposition*, 1–5.

Puterman ML (2005) *Markov decision processes: discrete stochastic dynamic programming* (Wiley).

Qi W, Liang Y, Shen ZJM (2015) Joint planning of energy storage and transmission for wind energy generation. *Operations Research* 63(6):1280–1293.

Rahimi F, Ipakchi A (2016) Using a transactive energy framework. *IEEE Electrification Magazine* 23–29.

Rust J (1997) Using randomization to break the curse of dimensionality. *Econometrica* 65(3):487–516.

Staffell I, Rustomji M (2016) Maximising the value of electricity storage. *Journal of Energy Storage* 8:212–225.

Thompson M, Rasmussen H, Davison M (2009) Natural gas storage valuation and optimization: A real options application. *Naval Research Logistics* 56(30):226–238.

Tian Y, Gross G (1998) OASISNET: an OASIS network simulator. *IEEE Transactions on Power Systems* 13(4):1251–1258.

van de Ven P, Hegde N, Massoulié L, Salonidis T (2013) Optimal control of end-user energy storage. *IEEE Transactions on Smart Grid* 4(2):789–797.

Whitt W (1978) Approximations of dynamic programs, I. *Mathematics of Operations Research* 3(3):231–243.

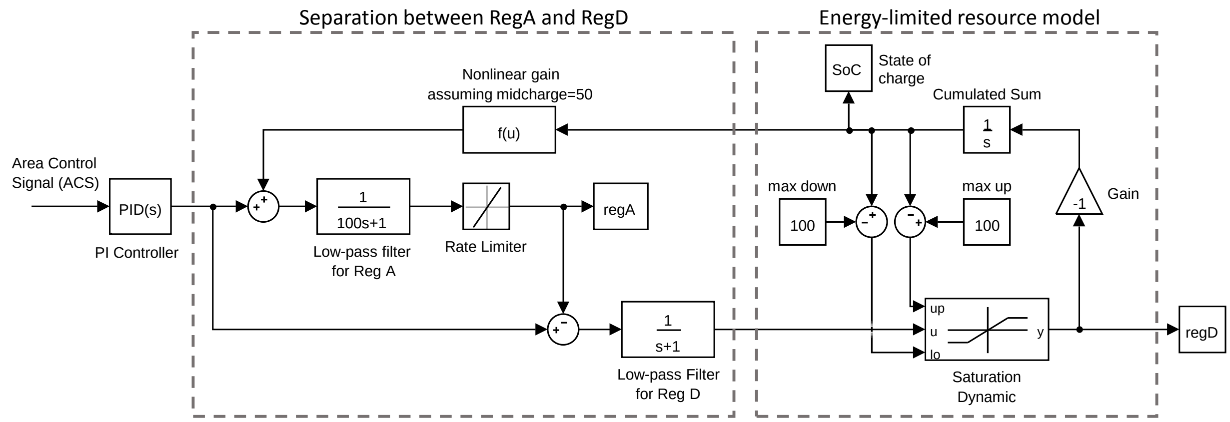
Xi X, Sioshansi R, Marano V (2015) A stochastic dynamic programming model for co-optimization of distributed energy storage. *Energy Systems* 1–31.

Xiao Y, Su Q, Bresler F, Carroll R, Schmitt J, Olaleye M (2014) Performance-based regulation model in PJM wholesale markets. *2014 IEEE PES General Meeting*, 1–5, URL http://dx.doi.org/10.1109/PESGM.2014.6939371.

Xu B, Shi Y, Kirschen DS, Zhang B (2018) Optimal battery participation in frequency regulation markets. *IEEE Transactions on Power Systems* 1–1.

Yu N, Foggo B (2017) Stochastic valuation of energy storage in wholesale power markets. *Energy Economics* 64:177–185.

Figure 9 Block-diagram of the regulation controller that generates the regulation signals regA and regD.



Appendix A: Architecture of the Real-Time Regulation Controller

The controller is derived from the partial description given in PJM staff (2017b). A control-theoretic analysis is beyond the scope of this paper, we only sketch the functioning to provide some intuition.

The block diagram is presented in Figure 9. Two regulation signals are produced: regA for ramp-limited resources and regD for energy-limited resources (see Figure 9). A proportional-integral (PI) controller is used to drive a balancing error to zero, up to some bias. The output signal of the PI controller is fed into a low-pass filter (1/(100s+1)) and a rate-limiter to produce regA. The output signal of the PI controller, net of the regA contribution, is passed through a saturation block to keep the state of charge within limits, and becomes the regD signal. The regA and regD signals are then sent to the resources, using a distributed network communication protocol. It is understood that the signals are scaled to the capacity of the resources providing regulation services. Imperfect compliance simply adds to the balancing error used as input to the PI controller.

The special element of this control loop is the internal model of the state of charge. The regD signal, changed of sign, is passed through an integrator ($1/s$) to model a state of charge (SoC) normalized between 0 and 100. Suppose that $u(t)$ denotes this virtual state of charge at time t. Consider the function $f(u) = k_f[\log(0.25) - \log(u/100) - \log(1-u/100)](1_{u<50} - 1_{u\geq 50})$ where $k_f > 0$ is a gain parameter. The function f satisfies the following properties. (i) At the mid-charge level $u = 50$, $f(50) = 0$. (ii) Below the mid-charge level ($u < 50$), $f(u) > 0$. (iii) Above the mid-charge level ($u > 50$), $f(u) < 0$. (iv) $\lim_{u\to 0} = \infty$ and $\lim_{u\to 100} = -\infty$.

In the control loop, the signal $f(u)$ is added to the input signal for RegA. This is done in order to help drive back the state of charge to the mid-charge level.

Appendix B: Explicit Objective

We introduce the following notation for the purpose of describing the objective of the continuous-time Markov decision process explicitly. Let $\tau_{c,i}$, $\tau_{d,j}$, $\tau_{r,l}$ be the arrival time of respectively the i-th charge permission, the j-th discharge permission and the l-th regulation service review event. Let (k_0, ℓ_0, m_0) be the state at time 0 and let $(k_{c,i}, \ell_{c,i}, m_{c,i})$, $(k_{d,j}, \ell_{d,j}, m_{d,j})$, $(k_{r,l}, \ell_{r,l}, m_{r,l})$ be the state at times $\tau_{c,i}$, $\tau_{d,j}$, $\tau_{r,l}$ respectively. Let $(p_{c,i}, \varrho_{c,i})$, $(p_{d,j}, \varrho_{d,j})$ $(p_{r,l}, \varrho_{r,l})$ be the energy and regulation prices at states $m_{c,i}$, $m_{d,j}$, $m_{r,l}$ respectively.

Let π refer to a Markov policy mapping states to decisions. Let $a_{c,i} = A^\pi_{\rm c}(k_{{\rm c},i}, \ell_{{\rm c},i}, m_{{\rm c},i})$ be the i-th charged amount, let $a_{d,j} = A^\pi_{\rm d}(k_{{\rm d},j}, \ell_{{\rm d},j}, m_{{\rm d},j})$ be the j-th discharged amount, and let $u_{r,l} = A^\pi_{\rm r}(k_{{\rm r},l}, \ell_{{\rm r},l}, m_{{\rm r},l})$ be the l-th regulation decision. In the *rental model* (see § 3.3), $u_{r,l}$ is 0-1 indicator that the l-th regulation request is rejected ($u_{{\rm r},l} = 0$) or accepted ($u_{{\rm r},l} = 1$). In the *mid-charge model*, $u_{r,l}$ is the change made at time $\tau_{r,l}$ in the regulation commitment. The decisions are said to be feasible if $a_{c,i} \in \mathcal{A}^{\rm c}(k_{{\rm c},i}, \ell_{{\rm c},i})$, $a_{d,j} \in \mathcal{A}^{\rm d}(k_{{\rm d},j}, \ell_{{\rm d},j})$, $u_{r,l} \in \mathcal{A}^{\rm r}(k_{{\rm r},l}, \ell_{{\rm r},l})$, where the sets are defined by (5a) or (5b). The class of all policies π mapping states to feasible decisions is denoted by Π.

For the *rental model*, let $T_{{\rm r},l}$ denote the random duration of the l-th regulation request, which is exponentially distributed with mean $1/\mu$. Define the random discounted payoff of the l-th regulation service request as $\zeta_{{\rm r},l} = \int_{t=0}^{T_{{\rm r},l}} e^{-\gamma t}\varrho_{{\rm r},l}dt$. Note that $\mathbb{E}[\zeta_{{\rm r},l}] = \varrho_{{\rm r},l}/(\mu+\gamma)$. The value function of the optimal control problem is then defined as

$$\begin{aligned} V(k,\ell,m) := \max_{\pi\in\Pi}\ \mathbb{E}\Big[&-\textstyle\sum_{i=1}^\infty e^{-\gamma\tau_{{\rm c},i}} C\left(a_{{\rm c},i}/\eta^{\rm c}, m_{{\rm c},i}\right) + \sum_{j=1}^\infty e^{-\gamma\tau_{{\rm d},j}} D\left(a_{{\rm d},j}\eta^{\rm d}, m_{{\rm d},j}\right) \\ &+ \textstyle\sum_{l=1}^\infty e^{-\gamma\tau_{{\rm r},l}} \zeta_{{\rm r},l} u_{{\rm r},l} \ \Big|\ (k_0,\ell_0,m_0) = (k,\ell,m)\Big]. \end{aligned}$$

Without loss of optimality we can take the expectation over the durations $T_{r,l}$. Then the equivalent lump-sum reward for accepting the l-th regulation request is $\mathbb{E}[\zeta_{r,l}]$. In general, the reward for accepting a regulation request while being in state m is calculated as $r_m = \varrho_m/(\mu+\gamma)$, as in (6a). Using $R(u,m) = r_m u$, the objective can then be written explicitly as

$$\begin{aligned} V(k,\ell,m) := \max_{\pi\in\Pi}\ \mathbb{E}\Big[&-\textstyle\sum_{i=1}^\infty e^{-\gamma\tau_{{\rm c},i}} C\left(a_{{\rm c},i}/\eta^{\rm c}, m_{{\rm c},i}\right) + \sum_{j=1}^\infty e^{-\gamma\tau_{{\rm d},j}} D\left(a_{{\rm d},j}\eta^{\rm d}, m_{{\rm d},j}\right) \\ &+ \textstyle\sum_{l=1}^\infty e^{-\gamma\tau_{{\rm r},l}} R(u_{{\rm r},l}, m_{{\rm r},l}) \ \Big|\ (k_0,\ell_0,m_0) = (k,\ell,m)\Big]. \end{aligned}$$

For the *mid-charge model*, let $\tau_{r,0} = 0$, $\ell_{r,0} = \ell_0$. Let (k_t, ℓ_t, m_t) be the state at time t. Note that $u_{r,l} = \ell_{r,l} - \ell_{r,l-1}$. The continuous reward from regulation is the quantity

$$\begin{aligned} \int_0^\infty e^{-\gamma t}\varrho_{m_t}\ell_t dt &= \textstyle\sum_{l=0}^\infty \int_{\tau_{r,l}}^{\tau_{r,l+1}} e^{-\gamma t}\varrho_{m_t}\ell_{r,l}dt \\ &= \textstyle\sum_{l=0}^\infty \left[\int_{\tau_{r,l}}^\infty e^{-\gamma t}\varrho_{m_t}\ell_{r,l}dt - \int_{\tau_{r,l+1}}^\infty e^{-\gamma t}\varrho_{m_t}\ell_{r,l}dt\right] \\ &= \textstyle\sum_{l=0}^\infty \left[\int_{\tau_{r,l}}^\infty e^{-\gamma t}\varrho_{m_t}\ell_{r,l}dt\right] - \sum_{l=1}^\infty \left[\int_{\tau_{r,l}}^\infty e^{-\gamma t}\varrho_{m_t}\ell_{r,l-1}dt\right] \\ &= \textstyle\int_0^\infty e^{-\gamma t}\varrho_{m_t}\ell_0 dt + \sum_{l=1}^\infty \left[\int_{\tau_{r,l}}^\infty e^{-\gamma t}\varrho_{m_t}u_{r,l}dt\right] \\ &= \textstyle\ell_0\left[\int_0^\infty e^{-\gamma t}\varrho_{m_t}dt\right] + \sum_{l=1}^\infty e^{-\gamma\tau_{r,l}}\left[\int_0^\infty e^{-\gamma t'}\varrho_{m_{t'+\tau_{r,l}}}dt'\right] u_{r,l}\,. \end{aligned}$$

In the last right-hand side, the quantities inside brackets are independent of the decision policy. Furthermore, the process $\varrho_{m_{t'+\tau_{r,l}}}$ for $t' \geq 0$ is equivalent to the process $\varrho_{m'_t}$ for $t \geq 0$ with m'_t distributed as the background process started at $m'_0 = m_{r,l}$. Therefore, without loss of optimality, we can replace the quantities within brackets by their expectation conditioned on the state at time $\tau_{r,l}$ for the purpose of optimization. Let e_m denote the m-th unit column vector in $\mathbb{R}^M$ and let $\varrho \in \mathbb{R}^M$ be the column vector with elements ϱ_m. Using $Prob(m_t = n \mid m_0 = m) = e_m^\top \exp(Qt)e_n$, we evaluate $r_m := \mathbb{E}\left[\int_0^\infty e^{-\gamma t}\varrho_{m_t}dt \mid m_0 = m\right] = \sum_{n=1}^M \int_0^\infty e^{-\gamma t}e_m^\top \exp(Qt)e_n\varrho_n dt = e_m^\top\left[\int_0^\infty \exp((Q-\gamma I)t)dt\right]\varrho$, which leads to $r_m = e_m^\top(\gamma I - Q)^{-1}\varrho$, as announced in (6b). Hence, using $R(u,m) = r_m u$, the objective can be written as with the *rental model*, up to the additional constant term $\mathbb{E}\left[\ell_0\int_0^\infty e^{-\gamma t}\rho_{m_t}dt\right] = R(\ell_0, m_0)$. To unify the formulations, we define $V(k,\ell,m)$ for the *mid-charge model* as the objective of the problem with the constant term $R(\ell_0, m_0)$ neglected. If $\ell_0 = 0$ then $R(\ell_0, m_0) = 0$ and there is no constant term being neglected.

Appendix C: Value Function with State-Dependent Arrival Rates

The arrival processes can be extended to Markov-Modulated Arrival Processes by using the device of the background process and assigning rates to each state of the background process. Conditionally on being in state m, permissions and requests arrive at rates λ^{c}_m, λ^{d}_m, and λ^{r}_m. The optimality conditions for the value function in the case of the rental model for instance can be adapted as follows. Let $\bar{\lambda}^{\mathrm{c}} := \sum_{m=1}^{M} \lambda^{\mathrm{c}}_m$, $\bar{\lambda}^{\mathrm{d}} := \sum_{m=1}^{M} \lambda^{\mathrm{d}}_m$, and $\bar{\lambda}^{\mathrm{r}} := \sum_{m=1}^{M} \lambda^{\mathrm{r}}_m$. Define the state-independent aggregate event rate A and the null-event probability $\bar{\nu}_{\ell,m}/A$ as

$$A := \bar{\lambda}^{\mathrm{d}} + \bar{\lambda}^{\mathrm{c}} + \bar{\lambda}^{\mathrm{r}} + \mu K + \bar{q} + \gamma,$$
$$\bar{\nu}_{\ell,m} := (\bar{\lambda}^{\mathrm{d}} - \lambda^{\mathrm{d}}_m) + (\bar{\lambda}^{\mathrm{c}} - \lambda^{\mathrm{c}}_m) + (\bar{\lambda}^{\mathrm{r}} - \lambda^{\mathrm{r}}_m) + \bar{q} - \bar{q}_m + \mu(K-\ell).$$

Conditionally on being in state (k,ℓ,m), λ^{c}_m/A is the probability that the next event is a charge permission, A^{d}_m/A is the probability that the next event is a discharge permission, λ^{r}_m/A is the probability that the next event is a regulation service request, q_{mn}/A is the probability that the next event is a price change from m to n, $\ell\mu/A$ is the probability that the next event is a regulation service release. Thus the value function of the optimal control problem satisfies

$$\begin{aligned} V(k,\ell,m) = &\frac{\lambda^{\mathrm{c}}_m}{A}\mathcal{H}^{\mathrm{c}}_{k\ell m}[V] + \frac{\lambda^{\mathrm{d}}_m}{A}\mathcal{H}^{\mathrm{d}}_{k\ell m}[V] + \frac{\lambda^{\mathrm{r}}_m}{A}\mathcal{H}^{\mathrm{r}}_{k\ell m}[V] \\ &+ \frac{\mu\ell}{A}V(k,\ell-1,m) + \frac{1}{A}\Big(\sum_{\substack{n=1\\ n\neq m}}^{M} q_{mn}V(k,\ell,n)\Big) + \frac{1}{A}\bar{\nu}_{\ell,m}V(k,\ell,m), \end{aligned}$$

where the operators $\mathcal{H}^{\mathrm{c}}_{k\ell m}$, $\mathcal{H}^{\mathrm{d}}_{k\ell m}$, $\mathcal{H}^{\mathrm{r}}_{k\ell m}$ are as in Equations (10), (11), (12).

Appendix D: Value Function with Anytime Decisions and Scheduled Transactions

When all types of decisions can be taken at any time, without loss of optimality we restrict our attention to decisions to be taken at event arrivals. Suppose the state is (k_-,ℓ_-,m_-) just before the event, and (k,ℓ,m) just after the event. If the event is a background state transition, we have $k=k_-$, $\ell=\ell_-$, $m\neq m_-$. If the event is a regulation release event, we have $k=k_-$, $\ell=[\ell_- - 1]^+$, $m=m_-$. Decisions are taken given the state (k,ℓ,m). Let (k',ℓ') denote the post-decision storage state, with the understanding that $[k'-k]^+$ is a charging decision, $[k-k']^+$ is a discharging decision, and $(\ell'-\ell)$ is a change of regulation commitment decision. The set of feasible post-decision storage states given (k,ℓ) is

$$\widetilde{\mathcal{A}}(k,\ell) := \{(k',\ell') \in [0,K]\times\{0,\ldots,K\} : \ell' \geq \ell,\ k'+\ell' \leq K\}.$$

Let n be the state when the decision is executed. Given the lead time L, we have $\mathrm{Prob}(m_{t+L}=n \mid m_t=m) = [\exp(QL)]_{mn}$. Let $R^0_n(k,\ell,k',\ell')$ be the lump-sum reward at execution time, given that (i) the storage state at the time the decision was taken was (k,ℓ), (ii) the post-decision storage state was (k',ℓ'), (iii) the state at execution time is n. Using the functions C and D, the lump-sum regulation reward r_n given by (6a), and the 0-1 indicator functions $\mathbf{1}_{k'<k}$ and $\mathbf{1}_{k'>k}$, the lump-sum reward at execution time is

$$R^0_n(k,\ell,k',\ell') = \mathbf{1}_{k'<k}R(\eta^d(k-k'),n) - \mathbf{1}_{k'>k}C((k'-k)/\eta^c,n) + (\ell'-\ell)r_n \text{ for } (k',\ell') \in \widetilde{\mathcal{A}}(k,\ell).$$

Let $\boldsymbol{R}^0(k,\ell,k',\ell')$ be the column vector with elements $R_n^0(k,\ell,k',\ell')$. The expected lump-sum reward at decision time when being in state (k,ℓ,m) and choosing the post-decision state (k',ℓ') can then be expressed as

$$\widetilde{R}(k,\ell,m,k',\ell') = e^{-\gamma L} e_m^\top [\exp(QL)] \boldsymbol{R}^0(k,\ell,k',\ell') \quad \text{for } (k',\ell') \in \widetilde{\mathcal{A}}(k,\ell).$$

Let $\widetilde{V}(k,\ell,m)$ denote the optimal value function of the modified control problem. The aggregated rate of events for the problem with discounted objective is defined as

$$\widetilde{\Lambda} = \mu K + \bar{q} + \gamma,$$

to be compared to (7). The state-dependent null-event rate exclusive of discount is

$$\widetilde{\nu}_{\ell m} = (\bar{q} - \bar{q}_m) + (K - \ell)\mu.$$

The optimality conditions for the value function are then expressed as

$$\widetilde{V}(k,\ell,m) = \frac{1}{\widetilde{\Lambda}} \left(\sum_{\substack{n=1 \\ n \neq m}}^{M} q_{mn} \widetilde{\mathcal{H}}_{k\ell n}[\widetilde{V}] \right) + \frac{1}{\widetilde{\Lambda}} \ell \mu \widetilde{\mathcal{H}}_{k,\ell-1,m}[\widetilde{V}] + \frac{1}{\widetilde{\Lambda}} \widetilde{\nu}_{\ell m} \widetilde{V}(k,\ell,m),$$
$$\widetilde{\mathcal{H}}_{k\ell m}[\widetilde{V}] := \max_{(k',\ell') \in \widetilde{\mathcal{A}}(k,\ell)} [\widetilde{V}(k',\ell',m) + \widetilde{R}(k,\ell,m,k',\ell')].$$

Appendix E: Background Process for the Reference Case

The background process for the reference case in Section 8 has its transition rate matrix set to $Q = Q^{\mathrm{r}} \oplus Q^{\mathrm{en}}$, following (1) and assuming the energy and regulation markets are statistically independent. This simplifies the calibration of the price models, which is not a central topic of the paper. One cannot exclude that the energy and regulation markets are statistically dependent, but there is a bias-variance tradeoff in the empirical estimation of Q from market data. With (1), there are $\mathcal{O}(n_1^2 + n_2^2)$ transition rates to estimate, compared to $\mathcal{O}((n_1 + n_2)^2)$ if we attempt to estimate Q directly. See Mannor et al. (2007) for a discussion of the bias-variance tradeoff in the context of Markov decision processes.

The matrix Q^{r} is the transition rate matrix of a 12-state continuous-time Markov chain that describes the regulation market price process. To each regulation market state is associated one of 12 regulation prices, determined by K-means clustering of January 2017 hourly regulation price data (see PJM staff (2017a)). The empirical mean sojourn time in a state and the empirical transition probabilities are used to calibrate the matrix Q^{r}. We give the matrix Q^{r} below, where the j-th row label gives the regulation price ϱ_j:

$$Q^{\mathrm{r}} = \begin{array}{r} \mathit{1.31} \\ \mathit{6.84} \\ \mathit{11.35} \\ \mathit{17.92} \\ \mathit{27.60} \\ \mathit{38.98} \\ \mathit{49.85} \\ \mathit{68.09} \\ \mathit{110.81} \\ \mathit{193.47} \\ \mathit{253.55} \\ \mathit{313.32} \end{array} \begin{bmatrix} -0.446 & 0.269 & 0.108 & 0.031 & 0.023 & 0 & 0.008 & 0.008 & 0 & 0 & 0 & 0 \\ 0.119 & -0.335 & 0.140 & 0.052 & 0.012 & 0.009 & 0 & 0.003 & 0 & 0 & 0 & 0 \\ 0.094 & 0.439 & -0.690 & 0.110 & 0.016 & 0.008 & 0.008 & 0 & 0.016 & 0 & 0 & 0 \\ 0.033 & 0.150 & 0.300 & -0.767 & 0.100 & 0.067 & 0.067 & 0.033 & 0.017 & 0 & 0 & 0 \\ 0.100 & 0.133 & 0.167 & 0.133 & -0.833 & 0.100 & 0.100 & 0.100 & 0 & 0 & 0 & 0 \\ 0 & 0.158 & 0.053 & 0.158 & 0.316 & -0.947 & 0.105 & 0.053 & 0.053 & 0 & 0 & 0.053 \\ 0.125 & 0 & 0.125 & 0.125 & 0.063 & 0.188 & -0.875 & 0.125 & 0.063 & 0 & 0 & 0.063 \\ 0 & 0.231 & 0.077 & 0.154 & 0 & 0.231 & 0.154 & -0.923 & 0 & 0.077 & 0 & 0 \\ 0 & 0.143 & 0 & 0 & 0.429 & 0 & 0.143 & 0.286 & -1.000 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1.000 & 0 & 0 & 0 & -1.000 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.000 & 0 & -1.000 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.500 & 0 & 0.500 & -1.000 \end{bmatrix}.$$

Figure 3(a) plots the regulation price data set. Figure 3(b) shows one sample path of the regulation price process generated using Q^{r}.

The matrix Q^{en} relative to the energy market is calibrated as follows. To each state is assigned one of 20 energy prices and one of 24 time indices. The time indices are labeled from 0 to 23, roughly corresponding to one hour of the day. At each transition of the continuous-time Markov chain, the time index is deterministically incremented by 1 modulo 24, while the next price is drawn randomly. We construct Q^{en} such that $Q^{\text{en}}_{ii} = -1$. This implies that state transitions happen hourly in expectation. We use real-time hourly locational marginal prices indexed by their hour stamps from PJM to estimate the transition rates. The price levels are determined by K-means clustering of the hourly price data. Using data for the first quarter of 2017, we obtain a 313-state continuous-time Markov chain for the energy market, after having eliminated the states corresponding to price-time pairs that were not observed. The 20 prices are 16.70, 20.19, 22.36, 24.13, 26.04, 28.32, 30.89, 33.82, 37.13, 41.69, 46.60, 51.23, 56.71, 61.81, 69.64, 78.73, 87.60, 101.18, 128.13, 154.99. For a comparison of the January empirical data to a sample path of the energy price generated with the calibrated model, see Figure 4. Overall, the background process has $M = 12 \cdot 313 = 3756$ states.

Appendix F: Computational Times

The reference case of Section 8 is solved by the 2-phase linear-optimization method developed in Section 5. Please note that the value iteration algorithm is not competitive in terms of computational times for reaching the level of precision attained by the linear optimization approach, although of course value iteration can be used to reach approximate solutions. During development we used one iteration of value iteration to verify the correctness of our linear optimization solutions.

We have carried out computational tests on problem instances of increasing dimensions, obtained by fitting the energy price process on an increasing number of states. We use the barrier algorithm of a commercial solver to solve (13) with the cross-over phase disabled whenever possible. See Güler et al. (1993) for the discussion of the robustness of interior point methods to degeneracy. The cross-over phase converts the interior-point solution to a basic solution (Megiddo (1991)), and is enabled by default by commercial solvers. We found that its effect on computational times for our problem class is detrimental, see Table 3. The table reports the size of the background process model and compares the times to solve (13) with and without the cross-over phase. These tests did not detect any significant discrepancy among the two solutions (column Sup-Error).

Appendix G: Optimal Policy for the Reference Case

An illustration of the behavior of the optimal policy derived from the optimal value function V of the reference case of Section 8.2 is provided in Figure 10. One sample path of the background process over

Table 3 Times (seconds) for solving (13) with the barrier algorithm.

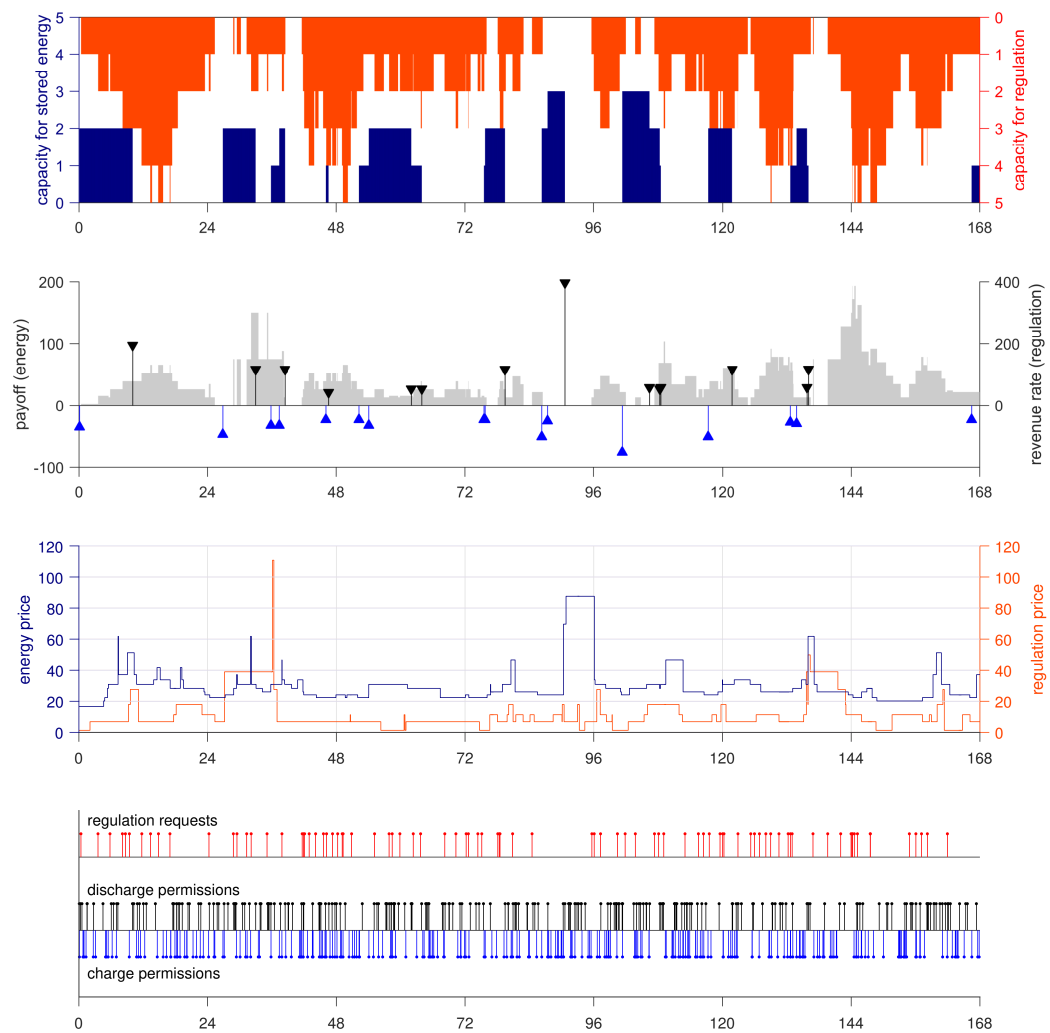
Number of background process states			Times, with Crossover		
(Regulation)	(Energy)	Total (M)	Enabled	Disabled	Sup-Error
12	2	24	0.8s	0.7s	$9.0 \cdot 10^{-5}$
12	20	240	147.7s	14.6s	$5.2 \cdot 10^{-5}$
12	64	768	5,000s	54.5s	$2.1 \cdot 10^{-4}$
12	313	3756	—	1455s	—

Sup-error = maximum solution discrepancy among all $v_{k\ell m}$ and $h_{k\ell m}$ between the two settings.

one week (168 hours) is generated, as well as the sample path for the permissions and requests times. At the permission or request times, an optimal action is selected following (10), (11) or (12) evaluated at V, depending on the nature of the event. When capacity blocks are reserved for regulation ($\ell_t > 0$), the release times are simulated as well.

During the simulation, we keep track of the stored energy level k_t and we record the amounts $C(a_t/\eta^{\mathrm{c}}, m_t) = a_t p_t/\eta^{\mathrm{c}}$ and $D(\eta^{\mathrm{d}} a_t, m_t) = a_t p_t \eta^{\mathrm{d}}$ paid at charging times and received at discharging times. We also keep track of each individual block of capacity reserved for regulation services, along with the revenue rate $\rho_{m_{t_i}}$ at which each ongoing reservation i was accepted at some time $t_i \leq t$, $i = 1, \ldots, \ell_t$. The total instantaneous revenue per hour from ongoing regulation services is obtained by summing up the revenue per hour from each block. At a capacity release event arrival, one of the reserved blocks is released, sampled uniformly among the reserved blocks. By so doing we sample the regulation revenue, rather than the equivalent expected lump-sum revenue used for optimizing V.

Figure 10 Sample path of the capacity utilization and payoffs, in response to prices and arrivals.



The first graph of Figure 10 shows the instantaneous allocation of the storage capacity, distinguishing between the capacity used for energy storage (in blue), the capacity used for regulation (in orange), and the free capacity (in white). The second graph shows the times and the magnitude of the undiscounted payments from energy market transactions (blue markers when buying, black markers when selling), along with the total undiscounted revenue rate from providing regulation services (in gray). The third graph shows the sample paths for the energy price (in blue) and the regulation price (in orange). The fourth graph shows the arrival times of the charging, discharging and regulation request events. For all graphs the horizontal axis is the time, in hours. From the first graph it is clear that the capacity utilization varies dynamically.

Appendix H: Proof of Propositions

PROPOSITION 1. *For each fixed state $(k,\ell,m)\in\mathcal{S}$, the value function $V(k,\ell,m)$ is nondecreasing in the arrival rates $\lambda^c,\lambda^d,\lambda^r$.*

Proof of Proposition 1. Without gain in optimality, we extend the policy space to randomized policies. Let V_1 denote the value function for the problem with rates $\lambda_1^{\mathrm{c}},\lambda_1^{\mathrm{d}},\lambda_1^{\mathrm{r}}$. Consider the problem with rates $\lambda_2^{\mathrm{c}},\lambda_2^{\mathrm{d}},\lambda_2^{\mathrm{r}}$ such that $\lambda_2^{\mathrm{c}}\geq\lambda_1^{\mathrm{c}}$, $\lambda_2^{\mathrm{d}}\geq\lambda_1^{\mathrm{d}}$, $\lambda_2^{\mathrm{r}}\geq\lambda_1^{\mathrm{r}}$, and the randomized policy that (i) at each charging permission arrival, chooses $a\in\arg\max_{a\in\mathcal{A}^c(k,\ell)}\{V_1(k+a,\ell,m)-C(a/\eta^{\mathrm{c}},m)\}$ with probability $p^{\mathrm{c}}:=\lambda_1^{\mathrm{c}}/\lambda_2^{\mathrm{c}}$, and chooses $a=0$ with probability $(1-p^c)$; (ii) at each discharging permission arrival, chooses $a\in\arg\max_{a\in\mathcal{A}^d(k,\ell)}\{V_1(k-a,\ell,m)+D(a\eta^{\mathrm{d}},m)\}$ with probability $p^{\mathrm{d}}:=\lambda_1^{\mathrm{d}}/\lambda_2^{\mathrm{d}}$, and chooses $a=0$ with probability $(1-p^{\mathrm{d}})$; (iii) at each regulation event arrival, chooses $u\in\arg\max_{u\in\mathcal{A}^r(k,\ell)}\{V_1(k,\ell+u,m)+R(u,m)\}$ with probability $p^{\mathrm{r}}:=\lambda_1^{\mathrm{r}}/\lambda_2^{\mathrm{r}}$, and chooses $u=0$ with probability $(1-p^{\mathrm{r}})$. Then everything happens as if the rates of the arrival processes were $\lambda_1^{\mathrm{c}},\lambda_1^{\mathrm{d}},\lambda_1^{\mathrm{r}}$ and the policy that attains V_1 was employed, showing that $V_2\geq V_1$, in the sense $V_2(k,\ell,m)\geq V_1(k,\ell,m)$ for all $(k,\ell,m)\in\mathcal{S}$. $\square$

PROPOSITION 2. *With the rental model, there exists problems where the value function $V(k,\ell,m)$ is nonincreasing in the mean regulation service time $1/\mu$, and problems where $V(k,\ell,m)$ is nondecreasing in $1/\mu$.*

Proof of Proposition 2. Suppose that the regulation price is 0 in all states m, while the energy prices vary. Then, $\ell>0$ causes a loss of capacity that could have been used to operate profitably in the energy market. In this case, since the capacity lost to ℓ is released at rate $\ell\mu$, the value function $V(k,\ell,m)$ is nonincreasing in $1/\mu$.

Now, consider a different case, with functions C and D linear in the quantity a, and energy prices such that $(\sup_{m_1}\{D(\eta^{\mathrm{d}},m_1)\}-\inf_{m_2}\{C(1/\eta^{\mathrm{c}},m_2)\})<0$. This is sufficient to guarantee that it is never optimal to operate in the energy market. Assume regulation prices are 0 except at some state m where $\varrho_m>0$. Then it is optimal to always accept regulation requests when being in state m, since those pay out a lump-sum reward $r_m=\varrho_m/(\mu+\gamma)$. Since r_m is decreasing in μ, $V(k,\ell,m)$ is this time nondecreasing in $1/\mu$. $\square$

PROPOSITION 3. *For each fixed ℓ and m, the value function $V(k,\ell,m)$ is not necessarily nondecreasing in the stored energy k.*

Proof of Proposition 3 With the functions C and D linear in the quantity a, suppose that $(\sup_{m_1}\{D(\eta^{\mathrm{d}},m_1)\}-\inf_{m_2}\{C(1/\eta^{\mathrm{c}},m_2)\})<0$. This is sufficient to guarantee that it is never optimal to operate in the energy market. Then, for states where $k > 0$, there exists at least one state m where it is optimal to fully discharge, and there is no state where it is optimal to charge. Suppose that λ^{d} is arbitrarily small. Then, having energy in storage that cannot be easily discharged occupies capacity for some time, that would otherwise be profitably used for performing regulation. In this case V is nonincreasing in k. $\square$

PROPOSITION 4. *For each fixed k and m, the value function $V(k,\ell,m)$ is nonincreasing in ℓ.*

Proof of Proposition 4. Recall that the expected reward r_m for increasing the regulation commitment of one unit is received as a lump sum at the regulation event time. Thus the only difference between state (k,ℓ,m) and state $(k,\ell+1,m)$ is that the free capacity in state $(k,\ell+1,m)$ to store energy or increase the regulation commitment is smaller. This implies that $V(k,\ell+1,m) \le V(k,\ell,m)$. $\square$

PROPOSITION 5. *Consider the special case where operations can only happen in the energy market (set $\ell = 0$ and $\lambda^r = 0$). (1) Suppose that initially, $k = 0$ or $k = K$. Then without loss of optimality the decisions can be restricted to $a \in \{-K, 0, K\}$, and the stored energy state can be restricted to $k \in \{0, K\}$. (2) Furthermore, allowing now $k \in [0,K]$ and $a \in [-K,K]$, for each fixed m the value function is affine in k.*

Proof of Proposition 5. Divide up the capacity K equally into N devices of capacity $K_i = K/N$, each device i being operated individually, subject to the constraint that $k_i \in \{0, K/N\}$ and $a_i \in \{-K/N, 0, K/N\}$. This represents a discretization of $[0,K]$ into $\{0, K/N, 2K/N, \dots, K\}$. As the reward structure is additive and the transition function is linear, there is no interaction among the devices, and all can employ the same optimal policy. Assuming they are all started from the same state k_i, equal either to 0 or to K/N, and since they are subject to the same background process and event arrivals, their actions are identical, and the aggregated action is $a = \sum_i a_i \in \{-K, 0, K\}$. This establishes the first part of the lemma.

For the second part, consider the value function at state $k = (j/N)K$, $j \in \{0, 1, \dots, N\}$, written $V(k,m)$. Starting from state $k = (j/N)K$ is equivalent to starting with j devices of capacity K/N fully charged and $N - j$ devices of capacity K/N fully discharged. Thus we have $V(k,m) = jV_1(K/N, m) + (N-j)V_1(0,m)$ in terms of the individual value functions $V_i(k_i, m)$ of each device i, showing that the value function V is linear in j. The results follow by letting N tend to ∞, invoking the limit theorems from Whitt (1978) and in particular Theorem 5.2 on the convergence of dynamic programs with discretized state spaces. $\square$

PROPOSITION 6. *For each fixed ℓ, m, the value function $V(k,\ell,m)$ as a function of k is piecewise linear with discontinuities between the pieces. In the rental model, the function has right-discontinuities at $k = \{0, 1, \dots, K-1\} \cap \{k \in [0,K] : k + \ell \le K\}$, and is left-continuous with right limits. In the mid-charge model, the discontinuities are (i) left-discontinuities at $k = \{1, \dots, \lfloor K/2 \rfloor\} \cap \{k \in [0, K/2] : k - \ell \ge 0\}$, and (ii) right-discontinuities at $k = \{\lceil K/2 \rceil, \dots, K-1\} \cap \{k \in [K/2, K] : k + \ell \le K\}$. The function is right-continuous with left limits on $k \in [\ell, K/2]$ and left-continuous with right limits on $k \in [K/2, K-\ell]$.*

Proof of Proposition 6. We prove the result for the rental model. Let $f(k,\ell)$ denote the unused capacity in state $(k,\ell,m) \in \mathcal{S}$ able to accommodate new regulation requests. We have $f(k,\ell) = K - \ell - \lceil k \rceil$. Indeed, note first that $f(k,\ell) \in \{0, 1, \dots, K\}$ since capacity for regulation is reserved by blocks. Let k be integer with $k + \ell \le K - 1$. Then $f(k,\ell) = K - \ell - k$ while $\lim_{\epsilon \downarrow 0^+} f(k+\epsilon, \ell) = f(k,\ell) - 1$. As the capacity available for new regulation requests drops at the points $\lim_{\epsilon\downarrow 0^+} k+\epsilon$, the value function will potentially drop at the same points.

Now, given $(k,\ell,m)\in\mathcal{S}$, decompose k as $k=k_1+k_2$ with $k_1=\lfloor k\rfloor$ and $k_2=k-k_1\in[0,1)$. Recall that C and D are linear in the quantity a. We can view k_2 as the charge level of an independent storage unit of capacity ($\lim_{\epsilon\downarrow 0^+} 1-\epsilon$) operated in the energy market only, since this capacity is too small to constitute a block of capacity for regulation. By Proposition 5 however, it is optimal to operate at the boundaries of this capacity, allowing us to conclude that without loss of optimality operations can be restricted to integer k. By the reasoning of Proposition 5, the value function will drop at the points ($\lim_{\epsilon\downarrow 0^+} k+\epsilon$), and in the interval $(k,k+1)$, the value function will be affine in k.

The proof for the mid-charge model follows from similar arguments. Recall that we assume that K is integer for convenience. Given ℓ, the capacity available for new regulation becomes $f(k,\ell)=\lfloor K/2\rfloor-\ell-(\lceil |k-K/2|-\delta_{1/2}\rceil)$ where $\delta_{1/2}:=(K/2-\lfloor K/2\rfloor)\in\{0,1/2\}$. The value function will drop at each downward jump of $f(k,\ell)$ and vary linearly otherwise. □

Proposition 7. *The directional derivatives $\partial_k^- V(k,\ell,m):=\lim_{\epsilon\downarrow 0^+}[V(k,\ell,m)-V(k-\epsilon,\ell,m)]/\epsilon$, defined where V is left-continuous, and the directional derivatives $\partial_k^+ V(k,\ell,m):=\lim_{\epsilon\downarrow 0^+}[V(k+\epsilon,\ell,m)-V(k,\ell,m)]/\epsilon$, defined where V is right-continuous, are independent of k and ℓ, that is, they only depend on the state m of the background process.*

Proof of Proposition 7. This is a continuation of the proof of Proposition 6. Since the device with charge level k_2 is operated independently of the device with charge level k_1 and operates in the energy market only, its value function $V_2(k_2,m)$ and optimal policy are independent of the state k_1 and regulation state ℓ. By Proposition 5, for each fixed m, $V_2(k_2,m)$ is affine in k_2. Since variations in the value function V at noninteger k can only be due to the $V_2(k_2,m)$ part, the result follows. □

Proposition 8. *The optimal value function V such that $V=\mathcal{T}V$ over $(k,\ell,m)\in\mathcal{S}$ satisfies $V(k,\ell,m)=v(k,\ell,m)=v_{k\ell m}$ over $(k,\ell,m)\in S$.*

Proof of Proposition 8. This is a direct consequence of the proof of Proposition 6, which shows that without loss of optimality, operations can be limited to the states where k is integer. □

Proposition 9. *For the rental model, the value function V which satisfies* (9) *is given by*

$$V(k,\ell,m)=v_{\lceil k\rceil \ell m}+(w_{1m}-w_{0m})(k-\lceil k\rceil)\qquad \text{for } k\in[0,K-\ell], \tag{16a}$$

where $v_{k\ell m}$ for all k,ℓ,m with k integer are obtained by solving the linear optimization problem (13)*, and w_{0m},w_{1m} for all m are obtained by solving the linear optimization problem* (15)*.*

For the mid-charge model, the value function V which satisfies (9) *is given by*

$$V(k,\ell,m)=\begin{cases} v_{\lfloor k\rfloor \ell m}+(w_{1m}-w_{0m})(k-\lfloor k\rfloor) & \text{for} \quad k\in[\ell,K/2],\\ v_{\lceil k\rceil \ell m}+(w_{1m}-w_{0m})(k-\lceil k\rceil) & \text{for} \quad k\in[K/2,K-\ell],\end{cases} \tag{16b}$$

where $v_{k\ell m}$ is obtained from (13) *with the mid-charge model.*

Proof of Proposition 9. This is a consequence of Proposition 7. The directional derivatives $\partial_k^- V(k,\ell,m)$ associated to each state m are identified as being $(w_{1m} - w_{0m})/(1-0)$, using the property that V is affine in k on the intervals $(k, k+1]$. Equation (16a) reflects the left-continuity of V in k established in Proposition 6. Equation (16b) is established similarly and reflects the needed adjustments from the portion of V that was shown to be right-continuous. □

PROPOSITION 10. *The value of operating in the energy market with assigned storage capacity x is linear in x, that is, $V_x^{en} = xV_1^{en}$.*

Proof of Proposition 10. Define the normalized state variables $\kappa = k/x$ and decisions $\alpha = a/x$. By linearity of C and D in the quantity, we have $C(s,m) = xC(s/x,m)$ and $D(s,m) = xD(s/x,m)$. Using the notation of Section B, let $\alpha_{c,i} = a_{c,i}/x$ and let $\alpha_{d,j} = a_{d,j}/x$. Then we have

$$\begin{aligned} V_x^{\text{en}} &= \max_\pi \mathbb{E}^\pi \left[\sum_{j=1}^\infty e^{-\gamma\tau_{\text{d},j}} D(a_{\text{d},j}\eta^{\text{d}}, m_{\text{d},j}) - \sum_{i=1}^\infty e^{-\gamma\tau_{\text{c},i}} C(a_{\text{c},i}/\eta^{\text{c}}, m_{\text{c},i}) \,\Big|\, k_0 = 0,\ m_0 \sim q \right] \\ &= \max_\pi \mathbb{E}^\pi \left[\sum_{j=1}^\infty e^{-\gamma\tau_{\text{d},j}} xD(\alpha_{\text{d},j}\eta^{\text{d}}, m_{\text{d},j}) - \sum_{i=1}^\infty e^{-\gamma\tau_{\text{c},i}} xC(\alpha_{\text{c},i}/\eta^{\text{c}}, m_{\text{c},i}) \,\Big|\, \kappa_0 = 0,\ m_0 \sim q \right] \\ &= xV_1^{\text{en}}. \qquad \square \end{aligned}$$

PROPOSITION 11. *Suppose $Q^{en} = \begin{bmatrix} -q_{12} & q_{12} \\ q_{21} & -q_{21} \end{bmatrix}$ with energy prices $p_1 < p_2$. Then we have*

$$\bar{V}_x^{en} = x\frac{\lambda^c}{\gamma}\left(\frac{q_{21}}{q_{12}+q_{21}}\right)\left[\frac{\lambda^d q_{12} D(\eta^d,2) - [\lambda^d(q_{12}+\gamma) + \gamma(q_{12}+q_{21}+\gamma)]C(1/\eta^c,1)}{(\lambda^c+\gamma)(q_{21}+\gamma+\lambda^d) + (\lambda^d+\gamma)q_{12}}\right]^+. \tag{18}$$

Proof of Proposition 11. By Proposition 5, when operations are restricted to the energy market, bang-bang policies are optimal. Given that $p_1 < p_2$, only two policies need to be considered: (1) the trivial policy that does nothing and has value 0, and (2) the policy that fully charges at the first charging permission arriving while being in state 1, and fully discharges at the first discharging permission arriving while being in state 2. To calculate the value function of the nontrivial policy, the rate of aggregated events can be reduced to $\Lambda = \lambda^c + \lambda^d + q_{12} + q_{21} + \gamma$. Let v_{km} denote the value function at $k = 0,1$ and $m = 1,2$ for the problem with capacity $K = 1$. The value function satisfies the linear system

$$\begin{aligned} \Lambda v_{01} &= \lambda^{\text{d}} v_{01} + \lambda^{\text{c}}[v_{11} - C(1/\eta^{\text{c}},1)] + q_{12}v_{02} + q_{21}v_{01}, \\ \Lambda v_{02} &= \lambda^{\text{d}} v_{02} + \lambda^{\text{c}} v_{02} + q_{21}v_{01} + q_{12}v_{02}, \\ \Lambda v_{11} &= \lambda^{\text{d}} v_{11} + \lambda^{\text{c}} v_{11} + q_{12}v_{12} + q_{21}v_{11}, \\ \Lambda v_{12} &= \lambda^{\text{d}}[v_{02} + D(\eta^{\text{d}},2)] + \lambda^{\text{c}} v_{12} + q_{21}v_{11} + q_{12}v_{12}. \end{aligned}$$

In matrix form,

$$\begin{bmatrix} -(\lambda^{\text{c}}+q_{12}+\gamma) & q_{12} & \lambda^{\text{c}} & 0 \\ q_{21} & -(q_{21}+\gamma) & 0 & 0 \\ 0 & 0 & -(q_{12}+\gamma) & q_{12} \\ 0 & \lambda^{\text{d}} & q_{21} & -(\lambda^{\text{d}}+q_{21}+\gamma) \end{bmatrix} \begin{bmatrix} v_{01} \\ v_{02} \\ v_{11} \\ v_{12} \end{bmatrix} = \begin{bmatrix} \lambda^{\text{c}} C(1/\eta^{\text{c}},1) \\ 0 \\ 0 \\ -\lambda^{\text{d}} D(\eta^{\text{d}},2) \end{bmatrix}.$$

From this we get

$$v_{01} = \frac{\lambda^{\mathrm{c}}}{\gamma}\left(\frac{q_{21}+\gamma}{q_{12}+q_{21}+\gamma}\right)Z, \qquad v_{02} = \frac{\lambda^{\mathrm{c}}}{\gamma}\left(\frac{q_{21}}{q_{12}+q_{21}+\gamma}\right)Z,$$
$$Z := \frac{\lambda^{\mathrm{d}} q_{12} D(\eta^{\mathrm{d}},2) - [\lambda^{\mathrm{d}}(q_{12}+\gamma)+\gamma(q_{12}+q_{21}+\gamma)]C(1/\eta^{\mathrm{c}},1)}{(\lambda^{\mathrm{c}}+\gamma)(q_{21}+\gamma+\lambda^{\mathrm{d}})+(\lambda^{\mathrm{d}}+\gamma)q_{12}}.$$

Using the steady-state probabilities $q_1 = q_{21}/(q_{12}+q_{21})$ and $q_2 = q_{12}/(q_{12}+q_{21})$, and taking into account the value function of the two possible policies, we get

$$\bar{V}_1^{\mathrm{en}} = \max\{0, q_1 v_{01} + q_2 v_{02}\} = \max\left\{0, \frac{q_{21}v_{01}+q_{12}v_{02}}{q_{12}+q_{21}}\right\} = \max\left\{0, \frac{\lambda^{\mathrm{c}}}{\gamma}\left(\frac{q_{21}}{q_{12}+q_{21}}\right)Z\right\}.$$

□

Proposition 12. *Regarding* (18) *we have* $\bar{V}_x^{en} \le \frac{x}{\gamma}\left[\frac{\Delta^{en}}{T^{en}}\right]^+ \min\left\{1, \left[\left(\frac{\lambda^c}{q_{12}}\right)^{-1} + \left(\frac{\lambda^d}{q_{21}}\right)^{-1}\right]^{-1}\right\}.$

Proof of Proposition 12. Suppose $\Delta^{\mathrm{en}} := D(\eta^{\mathrm{d}},2) - C(1/\eta^{\mathrm{c}},2) > 0$. Set $T^{\mathrm{en}} = 1/q_{12} + 1/q_{21}$. Then we have

$$\begin{aligned}
\bar{V}_1^{\mathrm{en}} &= \frac{\lambda^{\mathrm{c}}}{\gamma}\left(\frac{q_{21}}{q_{12}+q_{21}}\right)\left[\frac{\lambda^{\mathrm{d}} q_{12} D(\eta^{\mathrm{d}},2) - [\lambda^{\mathrm{d}}(q_{12}+\gamma)+\gamma(q_{12}+q_{21}+\gamma)]C(1/\eta^{\mathrm{c}},1)}{(\lambda^{\mathrm{c}}+\gamma)(q_{21}+\gamma+\lambda^{\mathrm{d}})+(\lambda^{\mathrm{d}}+\gamma)q_{12}}\right] \\
&\le \frac{\lambda^{\mathrm{c}}}{\gamma}\left(\frac{q_{21}}{q_{12}+q_{21}}\right)\frac{\lambda^{\mathrm{d}} q_{12}[D(\eta^{\mathrm{d}},2) - C(1/\eta^{\mathrm{c}},1)]}{\lambda^{\mathrm{c}} q_{21} + \lambda^{\mathrm{d}} q_{12}} \\
&= \frac{1}{\gamma}\left(\frac{q_{12}q_{21}}{q_{12}+q_{21}}\right)\left(\frac{\lambda^{\mathrm{c}}\lambda^{\mathrm{d}}}{\lambda^{\mathrm{c}} q_{21} + \lambda^{\mathrm{d}} q_{12}}\right)[D(\eta^{\mathrm{d}},2) - C(1/\eta^{\mathrm{c}},2)] \\
&= \frac{1}{\gamma}\left(\frac{1}{q_{12}}+\frac{1}{q_{21}}\right)^{-1}\left(\frac{q_{12}}{\lambda^{\mathrm{c}}}+\frac{q_{21}}{\lambda^{\mathrm{d}}}\right)^{-1}[D(\eta^{\mathrm{d}},2) - C(1/\eta^{\mathrm{c}},2)] \\
&= \gamma^{-1}\frac{\Delta^{\mathrm{en}}}{T^{\mathrm{en}}}\left[\left(\frac{\lambda^{\mathrm{c}}}{q_{12}}\right)^{-1} + \left(\frac{\lambda^{\mathrm{d}}}{q_{21}}\right)^{-1}\right]^{-1}.
\end{aligned}$$

Under the same assumptions we also have

$$\begin{aligned}
\bar{V}_1^{\mathrm{en}} &= \frac{\lambda^{\mathrm{c}}}{\gamma}\left(\frac{q_{21}}{q_{12}+q_{21}}\right)\left[\frac{\lambda^{\mathrm{d}} q_{12} D(\eta^{\mathrm{d}},2) - [\lambda^{\mathrm{d}}(q_{12}+\gamma)+\gamma(q_{12}+q_{21}+\gamma)]C(1/\eta^{\mathrm{c}},1)}{(\lambda^{\mathrm{c}}+\gamma)(q_{21}+\gamma+\lambda^{\mathrm{d}})+(\lambda^{\mathrm{d}}+\gamma)q_{12}}\right] \\
&\le \frac{\lambda^{\mathrm{c}}}{\gamma}\left(\frac{q_{21}}{q_{12}+q_{21}}\right)\frac{\lambda^{\mathrm{d}} q_{12}[D(\eta^{\mathrm{d}},2) - C(1/\eta^{\mathrm{c}},1)]}{\lambda^{\mathrm{c}}\lambda^{\mathrm{d}}} = \gamma^{-1}\frac{\Delta^{\mathrm{en}}}{T^{\mathrm{en}}}.
\end{aligned}$$

□

Proposition 13. $V_y^{reg}(\ell, m)$ *is nondecreasing in the assigned capacity* y.

Proof of Proposition 13. Any policy feasible under capacity y is also feasible under capacity $y+1$. □

Lemma 1. *Let* $f(\ell)$ *be concave over* $\ell \in \{0,\dots,K\}$. *Then* $g(\ell) = \ell f(\ell-1) + (K-\ell)f(\ell)$ *is concave. If* $f(\ell)$ *is nonincreasing then* $g(\ell)$ *is nonincreasing. If* $f(\ell)$ *is CNI then* $g(\ell)$ *is CNI.*

Proof of Lemma 1. Let $f(\ell)$ be concave, which means that $f(\ell+2) - 2f(\ell+1) + f(\ell) \le 0$ for $\ell = 0,\dots,K-2$. Implicitly we assume $K \ge 2$. Then we have $g(\ell+2) - 2g(\ell+1) + g(\ell) = \ell(f(\ell+1) - 2f(\ell) + f(\ell-1)) + (K-\ell-2)(f(\ell+2) - 2f(\ell+1) + f(\ell)) \le 0$, showing that g is concave.

Let $f(\ell)$ be nonincreasing, that is, $f(\ell+1) - f(\ell) \le 0$ for $\ell = 0,\dots,K-1$. Implicitly we assume $K \ge 1$. Then we have $g(\ell+1) - g(\ell) = \ell(f(\ell) - f(\ell-1)) + (K-\ell-1)(f(\ell+1) - f(\ell)) \le 0$, showing that g is nonincreasing.

□

LEMMA 2. *Let $f(\ell)$ be concave over $\ell \in \mathbb{N}$. Then $h(\ell) = \max\{f(\ell), r + f(\ell+1)\}$ is concave. Furthermore, if $f(\ell)$ is CNI and $r \geq 0$, then $g(\ell)$ is CNI.*

Proof of Lemma 2. Let $h(\ell) = \max_{u \in \{0,1\}}\{(1-u)f(\ell) + u(r + f(\ell+1))\}$. Let u_ℓ denote the optimal solution for $h(\ell)$ with the highest index. Note that $u_\ell = 0$ implies $f(\ell) > r + f(\ell+1)$, and that $u_\ell = 1$ implies $f(\ell) \leq r + f(\ell+1)$.

We have $u_{\ell+1} \leq u_\ell$. Indeed, by contradiction suppose that $1 = u_{\ell+1} > u_\ell = 0$. This would imply $r + f(\ell+2) \geq f(\ell+1)$ and $f(\ell) > r + f(\ell+1)$, which in turns implies $f(\ell+2) - 2f(\ell+1) + f(\ell) > 0$, contradicting the concavity of f. Therefore, $u_{\ell+1} \leq u_\ell$.

Let $\delta = h(\ell+2) - 2h(\ell+1) + h(\ell)$. If $(u_{\ell+2}, u_{\ell+1}, u_\ell) = (0,0,0)$ then $\delta = f(\ell+2) - 2f(\ell+1) + f(\ell) \leq 0$. If $(u_{\ell+2}, u_{\ell+1}, u_\ell) = (0,0,1)$ then $\delta = f(\ell+2) - f(\ell+1) + r < 0$ due to $u_{\ell+1} = 0$. If $(u_{\ell+2}, u_{\ell+1}, u_\ell) = (0,1,1)$ then $\delta = f(\ell+1) - f(\ell+2) - r \leq 0$ due to $u_{\ell+1} = 1$. If $(u_{\ell+2}, u_{\ell+1}, u_\ell) = (1,1,1)$ then $\delta = f(\ell+3) - 2f(\ell+2) + f(\ell+1) \leq 0$. In all cases $\delta \leq 0$, showing h is concave in ℓ.

Let $\delta' = h(\ell+1) - h(\ell)$ and use now the additional assumptions $f(\ell+1) - f(\ell) \leq 0$, $r \geq 0$. If $(u_{\ell+1}, u_\ell) = (0,0)$ then $\delta' = f(\ell+1) - f(\ell) \leq 0$. If $(u_{\ell+1}, u_\ell) = (0,1)$ then $\delta' = -r \leq 0$. If $(u_{\ell+1}, u_\ell) = (1,1)$ then $\delta' = f(\ell+2) - f(\ell+1) \leq 0$. In all cases $\delta' \leq 0$, showing h is nonincreasing in ℓ. □

PROPOSITION 14. *Consider the rental model with $\lambda^c = \lambda^d = 0$ and $k = 0$. For each fixed m, the value function $V_y^{reg}(\ell, m)$ is CNI in ℓ.*

Proof of Proposition 14. Let $\mathcal{T}$ denote the Bellman operator with values $V_{t+1} = \mathcal{T} V_t$ defined by

$$V_{t+1}(0,\ell,m) = \frac{\lambda^{\mathrm{r}}}{\Lambda} H^{\mathrm{r}}_{0\ell m}[V_t] + \frac{\mu\ell}{\Lambda} V_t(0,\ell-1,m) + \frac{1}{\Lambda}\Big(\sum_{\substack{n=1\\ n\neq m}}^{M} q_{mn} V_t(0,\ell,n)\Big)$$
$$+ \frac{\mu(K-\ell)}{\Lambda} V_t(0,\ell,m) + \frac{\bar{q} - \bar{q}_m}{\Lambda} V_t(0,\ell,m).$$

The mapping is contractive by definition of Λ. Therefore, $V_t = \mathcal{T}^t V_0$ converges to the unique fixed point of $\mathcal{T}$ as $t \to \infty$. The fixed point satisfies $V = \mathcal{T} V$ and coincides with the optimal solution to (9) and (13) with $\lambda^{\mathrm{c}} = \lambda^{\mathrm{d}} = 0$ and $k = 0$.

Let $V_0 \equiv 0$. In particular V_0 is CNI in ℓ. Suppose that V_t is CNI in ℓ. Then $\ell V_t(0,\ell-1,m) + (K-\ell)V_t(0,\ell,m)$ is CNI in ℓ for each fixed m by Lemma 1, and $H^{\mathrm{r}}_{0\ell m}[V_t]$ is CNI in ℓ for each fixed m by Lemma 2. The set of CNI functions is closed under addition and positive scaling, thus $V_{t+1}(0,\ell,m)$ is CNI in ℓ for each fixed m. □

PROPOSITION 15. *Consider the rental model with $\lambda^c = \lambda^d = 0$ and $k = 0$. For each state m, there exists a threshold $\ell^*_m \in \{0, \ldots, K\}$ such that at each regulation request arrival while being in state m, it is optimal to accept the request if $\ell < \ell^*_m$ and to reject the request if $\ell \geq \ell^*_m$.*

Proof of Proposition 15. From Proposition 14, the value function $V(0,\ell,m)$ is CNI in ℓ for each m. For a fixed m, let $f(\ell) = H^{\mathrm{r}}_{0\ell m}[V] = \max_{u\in\{0,1\}}\{(1-u)V(0,\ell,m) + u(r_m + V(0,\ell+1,m))\}$. Let u_ℓ denote the optimal solution of higher index associated to the maximization. By the arguments of the proof of Lemma 2, the concavity of V in ℓ implies $u_{\ell+1} \leq u_\ell$. This is also true at $\ell = K$ where $u_K = 0$. Thus $u_0 \geq \cdots \geq u_K$ and the threshold policy structure holds for each m. □

Proposition 16. *Under the rental model,* $V_y^{reg}(\ell, m) \le \dfrac{\lambda^r}{\gamma} e_m^\top (I - \gamma^{-1}Q)^{-1} r.$

Proof of Proposition 16. We establish an upper bound on $V_y^{\text{reg}}(0,m)$ since $V_y^{\text{reg}}(0,m) \ge V_y^{\text{reg}}(\ell,m)$ by Proposition 14. As V_y^{reg} is nondecreasing in y, we let the capacity y tend to infinity. In this case, it is optimal to accept all incoming regulation requests (we assume $r_m \ge 0$ for all m), irrespectively of the current state ℓ. The probability of having the background process started from state m in state n at time t is $p_{nt} = [\exp(Qt)]_{mn}$ where $\exp(Qt) = \sum_{j=0}^{\infty} (Qt)^j / j!$. Thus if there is an arrival at time t, the expected lump-sum reward at time t is $\bar{r}(t) = \sum_{n=1}^{M} p_{nt} r_n = e_m^\top [\exp(Qt)] r$. By the Campbell-Hardy theorem, see Brémaud (1981), we have

$$\lim_{y\to\infty} V_y^{\text{reg}}(0,m) = \mathbb{E}\left[\int_0^\infty e^{-\gamma t} \bar{r}(t) dN_t^{\text{r}}\right] = \int_0^\infty e^{-\gamma t} \lambda^{\text{r}} \bar{r}(t) dt$$

and we can conclude that

$$\begin{aligned} \lim_{y\to\infty} V_y^{\text{reg}}(0,m) &= \lambda^{\text{r}} \int_0^\infty e^{-\gamma t} e_m^\top [\exp(Qt)] r dt = \lambda^{\text{r}} e_m^\top \left[\int_0^\infty \exp((Q-\gamma I)t) dt\right] r \\ &= \lambda^{\text{r}} e_m^\top [\gamma I - Q]^{-1} r = \frac{\lambda^{\text{r}}}{\gamma} e_m^\top [I - \gamma^{-1} Q]^{-1} r. \end{aligned}$$

□

Proposition 17. $\widetilde{V}(k,\ell,m)$ *is nonincreasing in* L *for any starting state* (k,ℓ,m).

Proof of Proposition 17. Let $\widetilde{V}^L$ and $\widetilde{V}^{L'}$ denote the value functions of the modified problem stated in D with lead time L and L' respectively, for some lead times L, L' such that $0 < L < L'$. We omit the starting state (k,ℓ,m) from the notation. Let $\pi^{L'}$ denote an optimal policy attaining $\widetilde{V}^{L'}$. For the problem with lead time L, we extend the set of admissible policies to those able to take decisions at any continuous time. Let $\bar{V}^L$ denote the corresponding optimal value. Then we have $\bar{V}^L \ge \widetilde{V}^L$, and actually $\bar{V}^L = \widetilde{V}^L$ because the policy space extension is made without gain in optimality. We then construct an admissible policy for the problem with lead time L as follows: at each event arrival, reset the time to $t=0$, wait during a period of duration $L'-L$, and make the decision for the action implemented at time $t+L=L'$ based on the state history during the period $[0, L'-L]$. Then, the value $\widetilde{V}^{L'}$ is attained by the policy that only observes the state at reset time 0 and makes the decision according to $\pi^{L'}$. We can then conclude that $\bar{V}^L \ge \widetilde{V}^{L'}$ with the use of an improved policy, namely, one based on the state at time $L'-L$ rather than on the state at older time 0. □

Proposition 18. *There exists a critical lead time* $L^* \ge 0$ *beyond which optimal policies do not charge in the energy market and only operate in the regulation market. The critical* L^* *is nonincreasing in the discount rate* γ *and nondecreasing in the storage efficiency parameters* η^c, η^d.

Proof of Proposition 18. Recall that the background process is assumed to be ergodic. The probabilities $\text{Prob}(m_{t+L} = n \mid m_t = m)$ tend to the steady-state probabilities $p_{\infty,n} := \lim_{L\to\infty} \text{Prob}(m_{t+L} = n | m_t = m) = \lim_{L\to\infty} [\exp(QL)]_{mn}$ which are independent of the starting state m. This implies that the expected price at which energy will be purchased or sold is independent of the current state, and that no expected value can be gained from buying and selling energy later on. Regulation services, however, will still be paid at the expected regulation price, and thus the optimal policy should allocate the entire capacity to regulation. The critical lead time is expected to be nonincreasing in the discount rate γ and nondecreasing in the storage efficiency parameters η^{c}, η^{d}, since increasing the discount rate and decreasing the efficiency can only reduce the expected value from round-trip charge-discharge cycles. □