Mean-Conditional Value-at-Risk Optimal Energy Storage Operation in the Presence of Transaction Costs

Somayeh Moazeni, *Member, IEEE*, Warren B. Powell, *Member, IEEE*, and
Amir H. Hajimiragha, *Senior Member, IEEE*

Abstract—This paper addresses the formulation and solution of an optimal energy storage management problem under risk consideration and transaction costs of trading energy with the power grid. The price evolves as a stochastic process, capable of correctly explaining the seasonality effects as well as the tail fatness and spikiness in its distribution. Transaction costs capture the price impact of the storage operation on the electricity spot price. A risk analysis of an optimal risk neutral deterministic policy as well as the simple myopic policy indicates that the realized operational cost may notably differ from the expected cost by a considerable probability. This difference suggests that we need to consider risk. Using the downside risk measure of conditional value-at-risk, an optimal risk averse conversion and transmission strategy, among the grid, the renewable power generation source, and an energy storage is proposed to fully satisfy the electricity demand and minimize the expected operational cost as well as the risk. Our numerical study using data from NYISO demonstrates the impacts of risk consideration and the transaction cost parameters on the optimal strategy structure, its expected cost, and its risk.

Index Terms—Conditional value-at-risk (CVaR) optimization, energy storage, risk.

NOMENCLATURE

x_t^{IJ} — Amount of energy transferred from unit I to unit J at time step t [MWh].

β — Confidence level of the downside risk measure [-].

$\omega_{\mathbb{E}}, \omega_{\rho}$ — Risk aversion parameters [-].

α_t^{GI} — Transaction costs of exchanging energy from grid to unit I [$/MWh].

α_t^{IG} — Transaction costs of exchanging energy from unit I to grid [$/MWh].

$\gamma_{\Delta t}$ — Rate of energy loss over a time interval of length Δt [-].

η^{C} — Charging efficiency [-].

η^{D} — Discharging efficiency [-].

S^{cap} — Maximum capacity of the storage device [MWh].

ΔS^{C} — Maximum charging rate [-].

ΔS^{D} — Maximum discharging rate [-].

$S^{\max}$ — Maximum acceptable charge level [-].

$S^{\min}$ — Minimum acceptable charge level [-].

M — Number of Monte Carlo simulations.

$\tilde{P}_t$ — Energy price over the tth time step [$/MWh].

E_t — Energy output from the wind turbine over the tth time step [MWh].

D_t — Energy demand over the tth time step [MWh].

T — Planning horizon.

T_N — Number of time steps.

I. INTRODUCTION

ENERGY storage systems are becoming indispensable components of existing and future electricity grids by improving the grid economy, reliability, and stability [1], [2]. Different types of energy storage technologies with various response times can serve power system requirements in different time scales [3]. Thus, over a short time span (i.e., seconds to minutes), fast-response energy storage systems can offset the short-term fluctuations in demand and intermittent renewable power generation, and thereby improve the dynamic performance of the power systems. Over a long time span (i.e., hours to days), energy storage systems can potentially offer significant economic advantages by shifting the energy through absorbing surplus or inexpensive energy during off-peak hours, and then discharging it during on-peak hours, when electricity prices are typically high. In view of these considerations, grid-scale storage has drawn the attention of utilities throughout the world as a means to address many challenges they are dealing with, especially at present, when the penetration of intermittent and inflexible renewable power sources is on the rise. Grid-scale energy storage can provide a range of appealing benefits for the utilities, which include but are not limited to load leveling and peak shaving, smoothing power fluctuations, relieving transmission congestion, and deferring network upgrades [4]. However, an effective use of storage technology in energy and power systems is impossible without efficient operation

management, especially when their application over a long time period is considered; this is the main theme of the present study.

Variations of the storage operation management problem have been previously studied in the literature. Three control policies and their corresponding (expected) gained revenues are analyzed in [5], assuming probabilistic models for load and wind in the presence of the local network voltage perturbations. The sizing and management of a hydro pumping storage in an island power system with renewable penetration and deterministic generation cost are studied in [6]. The goal is to minimize the (expected) operation and annualized installation costs over all scenarios of wind, hydropower production, and load. Maximization of the (expected) market profit of a wind farm and a hydro pumped storage owner over a finite horizon to comply with his commitments in the market is discussed in [7]. The problem is formulated as a two-stage stochastic programming problem. An optimal bidding strategy in a day-ahead market, to maximize the (expected) discounted rewards from the bidding process for the owner of a hybrid system of renewable power generation and energy storage is investigated in [8]. This problem is modeled as a continuous-state Markov decision process and solved using approximate dynamic programming. A similar problem of maximizing the total discounted revenue of a storage owner when trading with the grid over a finite horizon and in the absence of any demand is studied in [9], where a simulation-based numerical method is proposed. An optimal energy commitment policy for the owner of a wind farm and a storage to maximize the (expected) revenue in an infinite horizon is investigated in [10], where both electricity price and wind power supply evolve as stochastic processes. In [11], a setting that involves conventional generation with a real time ex-ante price, renewable generation with a zero marginal cost, and an energy storage is considered. The goal is to fully serve elastic demand as well as to minimize the (average) cost of used conventional generation and investment in storage. The authors formulate the problem as a discrete time (average) cost stochastic dynamic programming problem over an infinite horizon and establish the existence of an optimal stationary policy. The optimal policy suggests to store when there is excess generation (over demand) and to extract energy from the storage device if there is excess demand (over generation). A virtual power plant consisting of an intermittent source, a storage facility, and a dispatchable power plant is considered in [12]. The virtual power plant sells and purchases electricity in both the day-ahead and the balancing markets in order to maximize the expected profit. The authors in [13] consider a weekly self-scheduling of a virtual power plant composed of intermittent renewable sources, storage system and a conventional power plant. The virtual power plant seeks to maximize its expected overall profit, while fulfilling its long-term bilateral contracts.

The vast majority of the literature on operation management of energy storage devices, including the aforementioned studies, minimize expected costs over the planning horizon, which implies that the controller is risk neutral. This ignores the reality that energy prices are highly volatile, exhibit pervasive spikes of extraordinary magnitude, and follow a fat-tailed distribution, e.g., see [14]. For example, the analysis in [15] on the daily average price data from April 1, 1998 to March 31, 2002 for the PJM Western Hub reports a skewness of 8.0 and kurtosis of 80.8, which clearly confirm the fat tail behavior in the electricity price distribution. The spikes may be due to sudden changes in the slope of the supply curve at the grid level. Indeed, price risk is the most significant risk exposure in energy systems [16]. Therefore, merely optimizing the expected cost (or revenue) is not sufficient and realistic risk considerations must be incorporated when looking for an efficient system operation strategy.

Here, we study the finite time horizon operation management of an energy system including a storage device, a time-variant renewable power generation source, and a load unit. The system owner is also responsible for satisfying the demand, with the capacity to trade power with the grid. The goal is to minimize the total operational cost and price risk exposure of the system. This model, although simple, is a building block in many advanced electric power systems.

The storage device is characterized by storage capacity limitations, conversion and dissipation losses, and maximum rates of conversion. Such a model encompasses various types of storage technologies and explains the essential storage behavior that arises in practice. Since wind power is one of the fastest-growing renewable energy sources (e.g., [17], [18]), this paper considers it for the energy generation. The time-varying load and electricity spot prices are reasonably explained by seasonality effects. The electricity price is assumed to evolve based on a mean reverting jump diffusion stochastic process, which is capable of addressing the main characteristics of energy prices. We calibrate the model to data for the New York Independent System Operator (NYISO).

First, we demonstrate that the realized operational cost of an optimal risk neutral deterministic policy as well as the myopic policy can differ significantly from the expected cost, with a considerable probability. For example, in almost 20% of the cases, the actual realized operational cost of the optimal risk neutral operation policy deviates from its expected cost by around 18% (see the second row of Table I-Policy$_E$). In 10% of the cases, the actual realized operational cost of this policy deviates from its expected cost by around 29% (see the third row of Table I-Policy$_E$). The deviation from expected costs becomes more significant as the confidence level increases. This clearly indicates the relevance of risk consideration in the context of storage operation management.

Second, we suggest incorporating a downside risk measure, such as the conditional value-at-risk (CVaR) in the objective function to properly model aversion to the (price) risk. The CVaR risk measure enjoys nice properties such as convexity [19] and has been widely used to model risk aversion in various industries, such as finance (e.g., [20], [21]) or electricity markets (e.g., [22], [23]). An efficient method to find a mean-CVaR optimal (deterministic) system operation policy is then developed. The approach relies upon a continuously differentiable smoothing function to approximate the non-differentiability from CVaR definition and an augmented Lagrangian method.

We then analyze the structure of the computed risk averse system operation strategy and compare it with the optimal risk neutral strategy. The optimal risk neutral strategy has a dual-threshold pattern; it keeps the storage level at its max-

imum when the gradient of the expected cost is positive and remains at the minimum charging status when the expected cost is decreasing over periods (see Fig. 5). However, the optimal risk averse strategy suggests interior levels and more frequent charging-discharging conversions. The frequency of changes in the charge level increases, as the CVaR confidence level approaches one (see the plots in Fig. 6). Furthermore, while the incurred expected operational cost of the optimal risk averse policy is only about 1% more than the expected cost of the optimal risk neutral policy, by adopting the optimal risk neutral operation policy the risk to contend with could be up to 13% more than what the risk averse strategy would have yielded (see Table VII). For example, the expected shortfall at 95% level of the optimal risk neutral policy is 9.99% more than that of the optimal risk averse storage operation policy.

In addition to risk consideration in the storage operation management problem, we introduce transaction costs into the model which include the cost of trading with the grid through an intermediate entity such as a utility or a power management company. The literature on the problem typically assumes that the decision maker is price taker and is a small player in a large market, i.e., the storage size is small or any demand shortfall can be made up by buying from the grid at the current spot price; see [6] or [9] for an explicit discussion. These assumptions ignore the storage operation's impact on the electricity spot prices. This concern is also briefly pointed out in [9, section 7]. The introduced transaction cost parameters in the definition of stage costs can explain effects of the storage operation and trading activities with the grid on the electricity spot prices. We study sensitivity of the structure of the optimal operation strategy, its cost, and its risk to these parameters.

The remainder of the paper is structured as follows. Section II presents the optimization model of the energy storage operation management problem studied in this paper. Risk analysis of the optimal risk neutral policy and myopic policy using the most popular downside risk measure is discussed in Section III. A computationally tractable approach to compute an optimal risk averse energy storage policy under CVaR is provided in Section IV. The results of our numerical experiment, policy analysis, and sensitivity analysis are presented in Section V. Concluding remarks are given in Section VI.

II. ENERGY STORAGE MANAGEMENT PROBLEM

Consider an energy system involving an electrical grid, a renewable power generation source (e.g., wind), an energy storage, and a demand, as shown in Fig. 1. The owner of the power generator and storage is also responsible to satisfy the demand for a finite planning horizon starting at time 0 and ending at time T. The time horizon $[0, T)$ is discretized into T_N intervals of length $\Delta t = T/T_N$. At each of the T_N time steps, the load is served by either the wind power, energy from the storage, or from the grid.

At the tth time step, the decision is on the following seven dimensional vector:

$$X_t \overset{\text{def}}{=} \left(x_t^{WD}, x_t^{GD}, x_t^{WG}, x_t^{WS}, x_t^{GS}, x_t^{SG}, x_t^{SD} \right) \quad (1)$$

for $t = 0, 1, \ldots, T_N - 1$. Here, x_t^{IJ} denotes the amount of energy transferred from unit I to unit J at time step t. The superscript W stands for power generator (e.g., wind), D for demand,
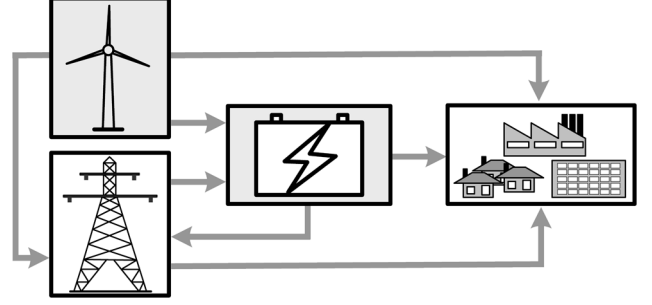

Fig. 1. Energy system diagram.

S for storage, and G for grid. All of these variables are *non-negative*.

Let S_t, referred to as the *storage level*, denote the fraction of the storage which is full at the tth time step. We assume S_0 is given and S_t evolves according to

$$S_{t+1} = (1 - \gamma_{\Delta t}) S_t + \frac{\eta^C \left(x_t^{GS} + x_t^{WS} \right) - \left(x_t^{SD} + x_t^{SG} \right)}{S^{\text{cap}}} \quad (2)$$

for $t = 0, 1, \ldots, T_N - 1$. Here, S^{cap} is the maximum capacity of the storage device. The constants $0 \leq \gamma_{\Delta t} \leq 1$ and $0 < \eta^C \leq 1$ represent the fraction of the storage charge level which is lost during the time interval of length Δt and the fraction of 1 MWh increase in the storage charge level as a result of 1-h charging at the 1 MW input power, respectively.

By implementing X_t at the tth time step, the system owner incurs the following cost (in dollar value) through exchanging energy with the grid:

$$C_t(\tilde{\mathbf{P}}_t, X_t) \overset{\text{def}}{=} \left(\tilde{P}_t^{GS} + \alpha_t^{GS} \right) x_t^{GS} + \left(\tilde{P}_t^{GD} + \alpha_t^{GD} \right) x_t^{GD}$$
$$- \eta^D \left(\tilde{P}_t^{SG} - \alpha_t^{SG} \right) x_t^{SG} - \left(\tilde{P}_t^{WG} - \alpha_t^{WG} \right) x_t^{WG} \quad (3)$$

where $\tilde{\mathbf{P}}_t = (\tilde{P}_t^{GS}, \tilde{P}_t^{GD}, \tilde{P}_t^{SG}, \tilde{P}_t^{WG})^\top$ refers to the electricity price, and $0 < \eta^D \leq 1$ indicates discharging efficiency. The cost function in (3) assumes that the grid pays for the amount of energy which is received, i.e., $\eta^D x_t^{SG}$, not the amount x_t^{SG} discharged from the storage.

The non-negative constants $\alpha_t^{GS}, \alpha_t^{GD}, \alpha_t^{SG}, \alpha_t^{WG}$ in (3) explain transaction costs of exchanging energy with the grid. Examples of such costs include maintenance costs of the energy storage technology (explained by α_t^{GS} and α_t^{SG}), or the maximum amount of power which can be absorbed by the grid (addressed through α_t^{SG} and α_t^{WG}). For example, when the grid cannot absorb energy anymore, we have $\alpha_t^{SG} = P_t^{SG}$ and $\alpha_t^{WG} = P_t^{WG}$; whence the system owner earns nothing by selling energy to the grid during that time step. In addition to these explicit costs, the system's operation and penetrating power to the grid may impact the real-time electricity price, especially if the storage size and amounts exchanged are significant. This price impact is captured through the transaction cost parameters. One may even consider impacts on the electricity price, and consequently the transaction cost parameters, to be functions of amount exchanged in the same or previous time intervals. However, in this paper, we assume that the transaction cost parameters are constants.

We refer to $C_t(\tilde{\mathbf{P}}_t, X_t)$ as the *stage cost*. A negative cost should be interpreted as revenue. The total system operational cost over the planning horizon $[0, T)$ then equals

$$\sum_{t=0}^{T_N-1} C_t(\tilde{\mathbf{P}}_t, X_t). \tag{4}$$

Describing the uncertain nature of electricity price $\tilde{\mathbf{P}}_t$ and time-varying loads and energy supplies require some modeling assumptions, which are explained next. In this paper, similar numerical values are used for prices, i.e., $\tilde{P}_t^{GS} = \tilde{P}_t^{GD} = \tilde{P}_t^{SG} = \tilde{P}_t^{WG} = \tilde{P}_t$; this is acceptable for most market situations.

[I] Energy Price ($\tilde{P}_t$): We assume a reduced form stochastic model accounting for intraday, weekly, and annual seasonality for the electricity spot prices. The *deterministic* seasonal component of $\tilde{P}_t$ is composed of an hour-of-day seasonal factor P_t^{hour}, a day-of-week seasonal term P_t^{day}, and a month-of-year seasonal part, P_t^{month}. These constants P_t^{hour}, P_t^{day}, and P_t^{month} are calculated as the average electricity price, respectively, over each of the hours of a day, over each of the days of a week, and over each of the months of a year. Hence, the energy price (in \$/MWh) over the tth time step is decomposed as

$$\tilde{P}_t = P_t^{\text{hour}} + P_t^{\text{day}} + P_t^{\text{month}} + \tilde{Y}_t^P. \tag{5}$$

This model explicitly incorporates price seasonality. The deseasonalized price $\tilde{Y}_t^P$ has the initial value $\tilde{Y}_0^P = \tilde{P}_0 - P_0^{\text{hour}} - P_0^{\text{day}} - P_0^{\text{month}}$.

Various reduced-form stochastic processes, as well as structural models, for $\tilde{Y}_t^P$ have been discussed in the literature. For a list of these processes and their characteristics, the reader is referred to [14]. Mean-reversion and the presence of spikes (or tail fatness) have been addressed as the main features of the electricity spot price distribution. Therefore, to capture these notable properties of energy prices, mean-reverting jump diffusion models have gained popularity, e.g., see [24], [25]. As a result, we also consider a mean-reverting jump diffusion model for $\tilde{Y}_t^P$ in this paper:

$$d\tilde{Y}_t^P = \lambda_P \left(\mu_P - \tilde{Y}_t^P \right) dt + \sigma_P d\tilde{W}_t + \tilde{J}_t d\tilde{q}_t \tag{6}$$

assuming that $\tilde{Y}_t^P$ evolves over continuous time. Here, λ_P is the speed of mean-reversion, and μ_P and σ_P denote the mean and volatility, respectively. In (6), $d\tilde{W}_t$ is the increment of the standard Brownian motion, $\tilde{J}_t$ indicates the proportional random jump size, and $d\tilde{q}_t$ is a Poisson process with intensity l.

[II] Time-varying Energy Demand (D_t): Similar to electricity prices, loads are modeled through the hour-of-day, day-of-week, and month-of-year seasonal components, denoted respectively by D_t^{hour}, D_t^{day}, and D_t^{month}. Thus a reasonable approximation for the energy demand (in MWh) over the tth time step can be

$$D_t = D_t^{\text{hour}} + D_t^{\text{day}} + D_t^{\text{month}} + \mathbb{E}\left[\tilde{Y}_t^D \right] \tag{7}$$

where $\tilde{Y}_t^D$ captures stochastic fluctuations in load. Often in the literature, the deseasonalized load $\tilde{Y}_t^D$ is addressed by a linear autoregressive model, e.g., see [26], [27]

$$\tilde{Y}_t^D = \phi_D \tilde{Y}_{t-1}^D + \sigma_D \sqrt{\Delta t}\tilde{\epsilon}_t, \quad \tilde{\epsilon}_t \sim \mathcal{N}(0, 1) \tag{8}$$

with the initial value $Y_0^D = D_0 - D_0^{\text{hour}} - D_0^{\text{day}} - D_0^{\text{month}}$, and the given constants $|\phi_D| < 1$ and σ_D. Therefore, $\mathbb{E}[\tilde{Y}_t^D] = \phi_D^t Y_0^D$ and, consequently, we have

$$D_t = D_t^{\text{hour}} + D_t^{\text{day}} + D_t^{\text{month}} + \phi_D^t Y_0^D. \tag{9}$$

[III] Time-varying Energy Supply (E_t): The power delivered by a wind turbine is often represented by its power curve, which establishes a relation between the wind speed and the power. Cubic or approximate cubic power curves have been commonly used to represent the power curves of variable speed wind turbine generators, e.g., see [28], [29]. Here, we consider a power curve model in which the non-linear relationship between power and wind speed is represented by a cubic expression that saturates at the rated power output. Namely, we let the power output $p_{\tilde{W}_t}^{(1)}$ (in MW) be given by

$$p_{\tilde{W}_t}^{(1)} = \begin{cases} 0 & \text{if } \tilde{W}_t < 0 \text{ or } \tilde{W}_t > 25 \\ 10^{-6} \times \frac{1}{2} A\nu C_{\mathbf{p}} \tilde{W}_t^3 & \text{if } 0 \leq \tilde{W}_t < v_r \\ p_r & \text{if } v_r \leq \tilde{W}_t \leq 25 \end{cases} \tag{10}$$

where $\tilde{W}_t$ denotes the wind speed measured in m/s and v_r represents the minimal wind speed corresponding to the rated power. In (10), A captures the area in m^2 swept by the rotor blades of the turbine, p_r is the rated power in MW, $C_{\mathbf{p}}$ denotes the power coefficient parameter with the maximum theoretical value of $(16/27) \approx 0.59$, and $\nu \approx 1.3$ kg/m^3 is the density of air, e.g., see [30]. In the sequel, we make the approximation of disregarding the possibility of having wind speed beyond the cut-off speed.

We assume that we have 50 identical wind turbines, each of which has the rated power of $p_r = 4$ MW, with the power coefficient $C_{\mathbf{p}} = 0.5$ and $A = 50^2\pi$ (in m^2). Also, $v_r \approx 11.62$ (in m/s). Velocity of the wind is assumed to be equal to $\tilde{W}_t = (\tilde{Y}_t^E + \mu_E)^2$, with given $\tilde{W}_0$ and μ_E. Here, $\tilde{Y}_t^E$ evolves according to an AR(1) model (e.g., see [31]):

$$\tilde{Y}_t^E = \phi_E \tilde{Y}_{t-1}^E + \sigma_E \sqrt{\Delta t}\tilde{\epsilon}_t, \quad \tilde{\epsilon}_t \sim \mathcal{N}(0, 1) \tag{11}$$

with the deterministic initial value $\tilde{Y}_0^E = 0$ and the given $|\phi_E| < 1$ and σ_E. Using the fact that $\{\tilde{\epsilon}_i\}_{i=1}^t$ are independent standard normal random variables and the 6th moment formula of standard normals, we can compute $\mathbb{E}[\tilde{W}_t^3]$. Thus, the (total) wind energy output (measured in MWh) from the 50 wind turbines over the tth time step (one hour) can approximately be computed by

$$E_t = 50 \times \min \left\{ \frac{50^2\pi\nu}{2 \times 10^6} \left(\mu_W(t)^6 + 15\mu_W(t)^4\sigma_E^2 \left(\frac{1 - \phi_E^{2t}}{1 - \phi_E^2} \right) \right. \right.$$
$$\left. \left. + 45\mu_W(t)^2\sigma_E^4 \left(\frac{1 - \phi_E^{2t}}{1 - \phi_E^2} \right)^2 + 15\sigma_E^6 \left(\frac{1 - \phi_E^{2t}}{1 - \phi_E^2} \right)^3 \right), 4 \right\} \tag{12}$$

where $\mu_W(t) \stackrel{\text{def}}{=} \phi_E^t \tilde{Y}_0^E + \mu_E$.

It is worth mentioning that there are alternative approaches for modeling the wind energy, e.g., see [32].

A. Constraints and Admissible Energy Flows

Given the exogenous parameter values E_t, D_t, and the storage level S_t at time step t, the following constraints are imposed on the energy flows in (1):

$$x_t^{WD} = \min\{E_t, D_t\} \tag{13}$$
$$x_t^{WS} + x_t^{WD} + x_t^{WG} = E_t \tag{14}$$
$$x_t^{GD} + \eta^D x_t^{SD} + x_t^{WD} = D_t \tag{15}$$
$$x_t^{SD} + x_t^{SG} \leq \underline{S}_t S^{\mathrm{cap}} + \eta^C \left(x_t^{GS} + x_t^{WS} \right) \tag{16}$$
$$\eta^C (x_t^{GS} + x_t^{WS}) - (x_t^{SD} + x_t^{SG}) \leq \overline{S}_t S^{\mathrm{cap}} \tag{17}$$
$$x_t^{SD} + x_t^{SG} \leq \Delta S^D S^{\mathrm{cap}} \tag{18}$$
$$\eta^C (x_t^{GS} + x_t^{WS}) \leq \Delta S^C S^{\mathrm{cap}} \tag{19}$$

where the *maximum charging rate* $0 \leq \Delta S^C \leq 1$ and the *maximum discharging rate* $0 \leq \Delta S^D \leq 1$ are the maximum fraction of the storage which can be charged or discharged, respectively, over the tth time interval. Here, $\overline{S}_t = S^{\max} - (1 - \gamma_{\Delta t}) S_t$ and $\underline{S}_t = (1 - \gamma_{\Delta t}) S_t - S^{\min}$ represent the maximum charging and discharging potential in one time step. The non-negative constants $S^{\min}$ and $S^{\max}$ indicate the minimum and maximum acceptable charge levels, which are percentages of the maximum capacity S^{cap}. These constants, respectively, reflect the maximum depth of discharge due to cycle life considerations and storage degradation over time that precludes reaching to the maximum storage capacity. Constraint (15) implies that the demand is fully satisfied through the grid, the storage, or the wind power generator. Constraint (16) ensures $S_{t+1} \geq S^{\min}$ and constraint (17) guarantees that $S_{t+1} \leq S^{\max}$. Constraints (18) and (19) prevent the storage device from charging or discharging faster than what is allowed.

The model described above permits concurrent charge and discharge of the storage within the same time step. However, routing energy from the grid to the user via the storage might not be efficient, and some literature such as [33] circumvent it. Such a round-trip flow is, however, unlikely to happen at an optimal strategy of our model, as typically $\eta^C \eta^D < 1$ and transaction cost parameters have reasonable values.

The system (13)–(19) is always feasible, since $\bar{X}_t$, whose elements are defined as follows, satisfies the set of constraints:

$$\bar{x}_t^{WD} = F_t, \quad \bar{x}_t^{GD} = D_t - F_t, \quad \bar{x}_t^{SD} = 0,$$
$$\bar{x}_t^{WG} = E_t - F_t, \quad \bar{x}_t^{WS} = 0, \quad \bar{x}_t^{GS} = 0, \quad \bar{x}_t^{SG} = 0 \tag{20}$$

where $F_t \overset{\text{def}}{=} \min\{E_t, D_t\}$. This strategy is the unique feasible strategy to serve the demand D_t, in the absence of a storage device, i.e., $S^{\mathrm{cap}} = 0$ MWh.

From the equality constraints (13)–(15), the three decision variables x_t^{WD}, x_t^{WS}, and x_t^{SD} can be fully determined using the other four decision variables:

$$x_t \overset{\text{def}}{=} \left(x_t^{GS}, x_t^{GD}, x_t^{SG}, x_t^{WG} \right)^\top. \tag{21}$$

Therefore, a feasible x_t can be explained using constraints (16)–(19), stated in terms of x_t in (21), and the two additional inequality constraints

$$E_t - \min\{E_t, D_t\} \geq x_t^{WG} \tag{22}$$
$$D_t - \min\{E_t, D_t\} \geq x_t^{GD} \tag{23}$$

to ensure $x_t^{WS} \geq 0$ and $x_t^{SD} \geq 0$. These constraints together imply that at every time step t, $x_t^{GD} = 0$ or $x_t^{WG} = 0$.

Given E_t, D_t, and S_t, the constraints (16)–(19) and (22)–(23) can be presented as a system of linear inequalities $A x_t \leq b_t$. Throughout this paper, we refer to the set of $x_t \in \mathbb{R}_+^4$ satisfying $A x_t \leq b_t$ by $\mathcal{X}_t$, which is bounded and convex.

B. Myopic and Optimal Risk Neutral Policies

At any time step t, a naive approach to adopt a system operation strategy is to simply minimize the tth stage cost and solve the following linear programming problem:

$$\min_{x_t \in \mathcal{X}_t} \quad C_t(\tilde{P}_t, x_t). \tag{24}$$

This policy, referred to as the myopic policy and denoted henceforth by Policy_M, discharges the storage as quickly as possible. While computing the myopic policy is straightforward, it entirely ignores the influence of decisions on the future state of the storage charge level, and hence the total operational cost. Hence, it is not a practically interesting strategy.

An alternative to the myopic policy, in which the total cost over the planning horizon is taken into account, is a solution of the following problem:

$$\min_{x_0 \in \mathcal{X}_0, \dots, x_{T_N-1} \in \mathcal{X}_{T_N-1}} \mathbb{E} \left[\sum_{t=0}^{T_N-1} C_t(\tilde{P}_t, x_t) \right] \tag{25}$$

which is referred to as an optimal risk neutral strategy and is denoted by Policy_E. When the transaction cost parameters are constants (i.e., do not depend on x_t), and parameters D_t and E_t can be approximated well by some deterministic (but time-varying) values, problem (25) is reduced to a linear programming problem.

Similar to the myopic policy, a solution for problem (25) can be computed easily either through a closed-form formula (for special cases) or using linear programming solution techniques and solvers.

Whether constraints are treated deterministically or probabilistically, Policy_E does not depend on the assumed stochastic dynamics model for $\tilde{Y}_t^P$ and, consequently, for the price $\tilde{P}_t$. However, the energy price, appearing in the objective function, is known to be highly volatile and spiky, e.g., see [14]. This fat tail behavior of the energy prices motivates the importance of using some downside risk (tail risk) measure, when looking for an optimal storage operation management strategy.

III. RISK ANALYSIS AND DOWNSIDE RISK MEASURES

The price risk is one of the most significant risks in the energy industry [16]. One of the most widely used risk measures is *value-at-risk* (VaR), e.g., see [16] or [34]. This downside risk measure is particularly appropriate for loss distributions with fat tail behavior such as that in the energy prices. For a given time horizon and confidence level β, the value-at-risk at the confidence level β is the smallest cost (loss in market value) over the time horizon that is exceeded with probability (no greater than) $1 - \beta$, e.g., see [19]. In other words, $\mathrm{VaR}_\beta(\mathrm{cost})$ is the level β-quantile of the loss distribution, i.e.,

$$\mathrm{VaR}_\beta(\mathrm{cost}) = \inf \{\ell \in \mathbb{R} : \Pr(\mathrm{cost} \leq \ell) \geq \beta\}. \tag{26}$$

Fig. 2 illustrates the probability distribution function of the total system operational cost, using the input data presented in Section V and Policy_E. Here we assumed that all of the transaction cost parameters are equal to 0. Table I reports the VaR values of the total costs corresponding to Policy_E and
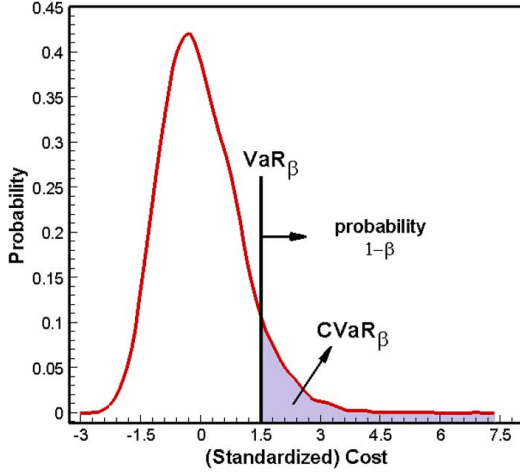
Fig. 2. Probability distribution function of operational cost.

TABLE I
VaR_β OF THE SYSTEM OPERATIONAL COST CORRESPONDING TO Policy_E AND Policy_M OVER 168 HOURS COMPUTED USING 20 000 MONTE CARLO PATHS. EXPECTED VALUES OF THE TOTAL COST CORRESPONDING TO Policy_E AND Policy_M EQUAL 15 458 318.42 *and* 15 530 033.15, RESPECTIVELY

Policy_E			
β [%]	VaR_β [$]	Absolute Difference [$]	Relative Difference [%]
75	17, 659, 736.25	2, 201, 417.83	14.24
80	18, 241, 611.01	2, 783, 292.59	18.01
85	19, 049, 807.51	3, 591, 489.10	23.23
90	20, 000, 727.86	4, 542, 409.45	29.38
95	21, 877, 231.73	6, 418, 913.31	41.52
99	25, 336, 499.63	9, 878, 181.21	63.90
99.90	30, 041, 183.60	14, 582, 865.19	94.34

Policy_M			
β [%]	VaR_β [$]	Absolute Difference [$]	Relative Difference [%]
75	17, 715, 420.01	2, 185, 386.86	14.07
80	18, 291, 168.95	2, 761, 135.81	17.78
85	19, 100, 810.54	3, 570, 777.39	22.99
90	20, 137, 750.95	4, 607, 717.80	29.67
95	21, 815, 224.10	6, 285, 190.96	40.47
99	25, 496, 654.67	9, 966, 621.52	64.18
99.90	30, 118, 873.21	14, 588, 840.06	93.94

Policy_M. As the results in this table indicate, while by minimizing the expected operational cost and adopting Policy_E, one expects to incur an operational cost of \$15 458 318.42, for example by looking at $\text{VaR}_{90\%}$, in 10% of cases the incurred cost will be 4 542 409.45 (around 29.38%) more than what is expected. Thus, to obtain an operation plan with less exposure to market risk and variations in the electricity spot price, it is important to include a downside risk measure in the objective function.

While VaR is still a firm-wide measure of risk in calculating regulatory and economic capital, the VaR function is non-convex, non-smooth, and has many local minima, see e.g., [19]. An attractive alternative to VaR is the coherent risk measure *conditional value-at-risk* (CVaR), also known as *average value-at-risk, mean excess loss*, or *mean shortfall*.

For a given time horizon $\bar{t}$ and confidence level β, CVaR_β is the conditional expectation of the loss above VaR_β for the time horizon $\bar{t}$. Thus, CVaR_β can be defined as

$$\text{CVaR}_\beta(\text{cost}) = \mathbb{E}\left[\text{cost} : \text{cost} \geq \text{VaR}_\beta(\text{cost})\right]. \quad (27)$$

Without referencing to VaR, a more direct way of defining CVaR is (see, e.g., [35])

$$\text{CVaR}_\beta(\text{cost}) = \min_\alpha \left(\alpha + \frac{1}{1-\beta}\mathbb{E}\left[[\text{cost} - \alpha]^+\right] \right) \quad (28)$$

where $[z]^+ = \max(z, 0)$. Minimizing CVaR typically leads to a strategy with a small VaR. Next, we discuss the CVaR risk consideration for the system operational cost of interest.

IV. OPTIMAL RISK AVERSE ENERGY STORAGE POLICY

An optimal risk averse (deterministic) policy can be derived by solving the following problem:

$$\min \quad \omega_\mathbb{E}\mathbb{E}\left[\sum_{t=0}^{T_N-1} C_t\left(\tilde{P}_t, x_t\right)\right] + \omega_\rho\rho\left[\sum_{t=0}^{T_N-1} C_t\left(\tilde{P}_t, x_t\right)\right],$$
$$\text{s.t.} \quad x_0 \in \mathcal{X}_0, \ldots, x_{T_N-1} \in \mathcal{X}_{T_N-1}. \quad (29)$$

Here, $\rho[\cdot]$ is a risk measure. The non-negative parameters $\omega_\mathbb{E}$ and ω_ρ, where $\omega_\mathbb{E} + \omega_\rho = 1$, capture the aversion of the decision maker to the price risk. We focus on the conditional value-at-risk as the risk measure. Using (28), a mean-CVaR optimal deterministic energy storage policy is determined by

$$\min_{x \in \mathbb{R}_+^{4T}} \quad \omega_\mathbb{E}\mathbb{E}[\tilde{c}^\top x] + \omega_\rho\text{CVaR}_\beta[\tilde{c}^\top x]$$
$$= \min_{\alpha \in \mathbb{R}, x \in \mathbb{R}_+^{4T}} \quad \omega_\mathbb{E}\mathbb{E}[\tilde{c}^\top x] + \omega_\rho\alpha + \frac{\omega_\rho}{1-\beta}\mathbb{E}\left[[\tilde{c}^\top x - \alpha]^+\right] \quad (30)$$

where $x = (x_0, \ldots, x_{T_N-1})^\top$ and $\tilde{c} = (\tilde{c}_0, \ldots, \tilde{c}_{T_N-1})^\top$. To solve problem (30), [35] suggest to replace the piecewise linear function $[z]^+$ with a set of linear constraints, as follows:

$$\min_{\alpha \in \mathbb{R}, x \in \prod_{t=0}^{T_N-1} \mathcal{X}_t} \quad \omega_\mathbb{E}\mathbb{E}[\tilde{c}^\top x] + \omega_\rho\alpha + \frac{\omega_\rho}{M(1-\beta)}\sum_{\ell=1}^{M} z_\ell$$
$$z_\ell - x^\top \tilde{c}^{(\ell)} + \alpha \geq 0, \ell = 1, \ldots, M,$$
$$z_\ell \geq 0, \ell = 1, \ldots, M. \quad (31)$$

Here, M denotes the number of Monte Carlo scenario paths. Here, the Monte Carlo simulation is used to approximate the mathematical expectations (for mean as well as the expected value in the definition of CVaR) in the objective function.

This approach can quickly become computationally rather expensive and inefficient as the number of simulation paths increases; each new scenario path introduces a new decision variable z_ℓ and a new constraint to the optimization problem. However, to have an accurate estimation of the tail of the cost distribution, a large number of Monte Carlo paths is required.

Alternatively, using the smoothing technique in [36], the function $[z]^+$ can be approximated with a continuously differentiable piecewise quadratic function $\varrho_\epsilon(z)$ using a small resolution parameter $\epsilon > 0$:

$$\varrho_\epsilon(z) = \begin{cases} z & \text{if } z > \epsilon \\ \frac{z^2}{4\epsilon} + \frac{1}{2}z + \frac{1}{4}\epsilon & \text{if } -\epsilon \leq z \leq \epsilon \\ 0 & \text{if } z < -\epsilon. \end{cases} \quad (32)$$

Note that $\varrho_\epsilon(z) \geq 0$ for every z. It is shown in [36] that this smoothing approach can be significantly more efficient than

the linear programming formulation (31) for the CVaR minimization problem. The convergence of stationary points of the smoothed sample average approximation problem, with the increase of sample size, has been investigated in [37]. More precisely, [37] proves that accumulation points of the stationary points of the smoothed sample average approximation problem, to solve the ϱ_ϵ-smoothed problem, are almost surely weak stationary points of their counterparts in the true problem, as the sample size increases.

Applying function (32), the objective function in (30) is then reduced to the following convex and continuously differentiable piecewise quadratic function:

$$\omega_\mathbb{E}\mathbb{E}[\tilde{c}^\top x] + \omega_\rho \alpha + \frac{\omega_\rho}{1-\beta}\mathbb{E}\left[\varrho_\epsilon(\tilde{c}^\top x - \alpha)\right]$$

$$= \omega_\mathbb{E}\mathbb{E}[\tilde{c}]^\top x + \omega_\rho \alpha + \frac{\omega_\rho}{M(1-\beta)}\sum_{\ell=1}^{M}\varrho_\epsilon\left(x^\top \tilde{c}^{(\ell)} - \alpha\right). \quad (33)$$

We refer to a solution of problem (29) with an objective function as in (33) by Policy_C. In this formulation, the number of decision variables or the constraints do not depend on the number of Monte Carlo sample paths. This problem can be solved using an augmented Lagrangian method, e.g., see [38, Section 17.3]. The augmented Lagrangian, denoted by $\mathcal{L}_A(x, s, \alpha, \lambda; \mu)$, for this problem can be written as follows:

$$\omega_\mathbb{E}\mathbb{E}[\tilde{c}]^\top x + \omega_\rho \alpha + \frac{\omega_\rho}{M(1-\beta)}\sum_{\ell=1}^{M}\varrho_\epsilon\left(x^\top \tilde{c}^{(\ell)} - \alpha\right)$$

$$- \sum_{t=0}^{T_N-1}\lambda_t^\top(Ax_t - b_t + s_t) + \frac{\mu}{2}\sum_{t=0}^{T_N-1}\|Ax_t - b_t + s_t\|_2^2$$

$$(34)$$

where $x_t \geq 0$ and the slack variables $s_t \geq 0$. The vector λ_t includes the Lagrange multipliers corresponding to the constraints $Ax_t + s_t = b_t$. At every iteration k, the algorithm then fixes the penalty parameter $\mu^{(k)} > 0$, fixes the Lagrange multipliers at the current estimate $\lambda^{(k)}$, and performs minimization of the augmented Lagrangian with respect to $x, s,$ and α. The computed approximate minimizer is then used to update the penalty parameter and Lagrange multipliers. Such a minimizer can be calculated using an interior point method [39] for nonlinear minimization with bound constraints.

V. NUMERICAL ILLUSTRATION

We use 5-min prices (in $/MWh) and hourly demands (in MW) for NYISO from 12 am of January 1, 2007 (Monday) to 12 am of December 31, 2011 (Saturday). These data sets are publicly available at NYISO website. We replace the negative prices with 1/MWh, and take the average of 12 measurements from 5-min prices for each hour to derive hourly prices. The estimation process for the deterministic seasonal trend is then as follows: 1) For every hour of a day, P_t^{hour} is set as the mean of the price of that hour over all of the days in our data set. Then for every hour, the computed P_t^{hour} is subtracted from the hourly prices to derive hourly residual prices. 2) For each day in a week, P_t^{day} is computed as the mean of the average hourly residual prices per day over all of the weeks in the historical data set. Then for each hour this value is subtracted from the hourly residual prices to obtain updated hourly residual prices. 3) For each month of a year, P_t^{month} is the mean of the average updated hourly residual prices per month over the five years 2007 to 2011. This value is then subtracted from each updated

TABLE II
HOUR OF DAY SEASONALITY FACTORS

Hour	Time Interval	Demand D_t^{hour} [MW]	Price P_t^{hour} [$/MWh]
0	[12 am , 1 am)	5159.62	52.92
1	[1 am , 2 am)	4943.20	47.81
2	[2 am , 3 am)	4819.20	43.88
3	[3 am , 4 am)	4781.16	42.30
4	[4 am , 5 am)	4885.42	44.07
5	[5 am , 6 am)	5225.77	48.49
6	[6 am , 7 am)	5713.78	57.95
7	[7 am , 8 am)	6160.71	61.54
8	[8 am , 9 am)	6515.21	66.08
9	[9 am , 10 am)	6756.23	71.13
10	[10 am , 11 am)	6898.87	72.75
11	[11 am , 12 pm)	6973.70	73.00
12	[12 pm , 1 pm)	7006.82	75.05
13	[1 pm , 2 pm)	7015.28	77.30
14	[2 pm , 3 pm)	7017.51	80.85
15	[3 pm , 4 pm)	7029.42	81.30
16	[4 pm , 5 pm)	7036.35	85.72
17	[5 pm , 6 pm)	6974.70	88.22
18	[6 pm , 7 pm)	6881.89	82.30
19	[7 pm , 8 pm)	6773.92	80.02
20	[8 pm , 9 pm)	6595.16	77.30
21	[9 pm , 10 pm)	6309.24	68.12
22	[10 pm , 11 pm)	5918.03	61.52
23	[11 pm , 12 am)	5492.11	56.51

TABLE III
DAY OF WEEK SEASONALITY FACTORS

Day	Demand D_t^{day} [MW]	Price P_t^{day} [$/MWh]
Monday	174.19	2.43
Tuesday	266.76	2.49
Wednesday	263.84	3.42
Thursday	224.28	1.17
Friday	146.66	0.34
Saturday	-468.64	-3.65
Sunday	-592.24	-5.45

TABLE IV
MONTH OF YEAR SEASONALITY FACTORS

Month	Demand D_t^{month} [MW]	Price P_t^{month} [$/MWh]
January	-221.78	10.29
February	-253.70	6.04
March	-520.57	-2.71
April	-682.52	-0.97
May	-454.68	1.69
June	659.15	10.29
July	1454.98	13.99
August	1147.22	-0.79
September	297.15	-8.56
October	-524.26	-14.23
November	-627.55	-15.38
December	-307.61	0.23

hourly residual price to attain deseasonalized hourly prices. A similar process is carried out for loads. The estimated values for seasonal terms are presented in Tables II–IV.

Using the maximum likelihood method, we then obtain estimations for the parameters in $\tilde{Y}_t^P$ and $\tilde{Y}_t^D$. This is reported in Table V, with $l = 0.27$, and $\tilde{J}_t \sim \mathcal{N}(0.03, 0.41)$ for $\tilde{Y}_t^P$.

TABLE V
PARAMETERS FOR $\tilde{Y}_t^P$, $\tilde{Y}_t^D$, AND E_t

$\mu_P = 4.35$	$\lambda_P = 37.48$	$\sigma_P = 2.08$
$\phi_D = 0.97$	$\sigma_D = 138.08$	
$\phi_E = 0.95$	$\sigma_E = 0.9$	$\mu_E = 3$

TABLE VI
PARAMETERS FOR ENERGY STORAGE

$S^{\min} = 0.1$	$\Delta S^C = 0.2$	$\eta^C = 0.75$
$S^{\max} = 0.9$	$\Delta S^D = 0.25$	$\eta^D = 0.9$
$S_0 = S^{\min}$	$S^{\text{cap}} = 1,000$ [MWh]	$\gamma_{\Delta t} = 0$

We let $\tilde{Y}_0^P = -5.88$ and $\tilde{Y}_0^D = -63.63$, which are the de-seasonal demand and price at hour 12 am of January 1, 2007. In the first week of January 2007, the maximum load and minimum load had been 7003.60 MW and 4223.10 MW, respectively. We assume that the energy demand to be served by the system owner is 25% of the demand calibrated from historical data from NYISO. Demand over the time horizon of our study is illustrated in Fig. 3.

Furthermore, in our numerical studies, to approximate E_t we assume $\tilde{Y}_0^E = 0$, and ϕ_E, σ_E, μ_E as in Table V. In addition, for the energy storage we use the values in Table VI. We then use the average of 50 path samples as E_t in our case study. The energy E_t over the time horizon 168 hours is illustrated in Fig. 4. Next, in Section V-A, different policies are compared. Sensitivity to the transaction cost parameters is investigated in Section V-B.

A. Policy Analysis

Fig. 5 illustrates storage levels corresponding to Policy_E and Policy_M, when all of the transaction cost parameters are equal to zero. The dashed line in this plot indicates the scaled expected price. Policy_M discharges the storage to the minimum level $S^{\min}S^{\text{cap}}$ as fast as possible (in this example, in the first time step) and keeps it at the minimum level for the rest of the planning horizon. In contrast, Policy_E increases the storage level toward the maximum level $S^{\max}S^{\text{cap}}$ when the gradient of the expected price is positive, but discharges it toward the minimum level $S^{\min}S^{\text{cap}}$ when the expected price is decreasing. The speed of switching between $S^{\min}$ and $S^{\max}$ depends on ΔS^C and ΔS^D. For example, when $\Delta S^C \geq 0.8$ and $S^D \geq 0$, Policy_E is a two-threshold policy with abrupt changes in S_t and moves with the gradient of the expected price, when all of the transaction cost parameters are equal to zero. For $\Delta S^C = 0.20$ and $\Delta S^D = 0.25$, as assumed in Table VI, switching between the maximum and minimum charge levels requires four time steps (hours). Note that the configurations of both Policy_E and Policy_M depend on neither the stochastic fluctuations of $\tilde{Y}_t^P$ nor the average of spikes in the set of generated scenario paths.

Fig. 6 depicts the storage levels of Policy_C for several confidence levels β, when $\omega_{\mathbb{E}} = 1/51$ and $\omega_\rho = 50/51$. Unlike Policy_E, Policy_C is not a two-threshold policy, even when all of the transaction cost parameters are zero; by implementing Policy_C, S_t may remain in other (interior) storage levels. However, this pattern depends highly on the chosen confidence level β. As the value of the confidence level increases, the storage profile deviates more from the two-threshold configuration.

The costs and risks of the three policies Policy_C, Policy_E, and Policy_M are presented in Table VII. Here, $\varepsilon_{\text{CVaR}_\beta}$ and
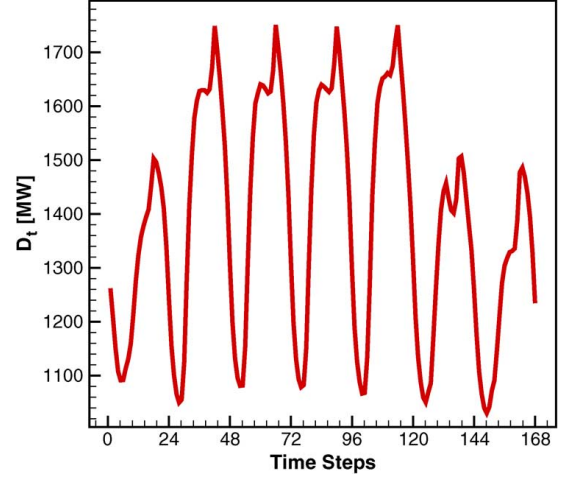

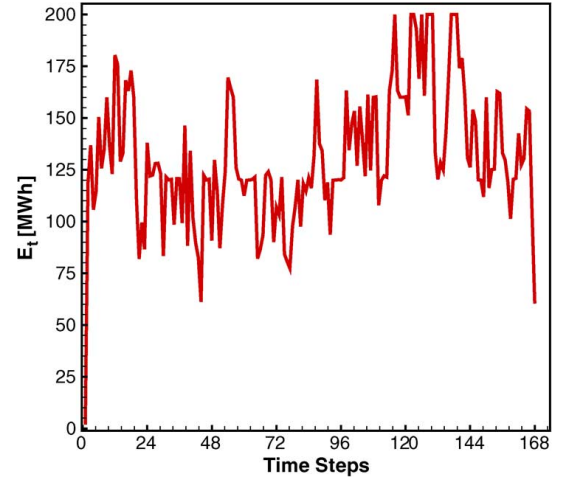Fig. 3. Load over the time horizon (168 hours).

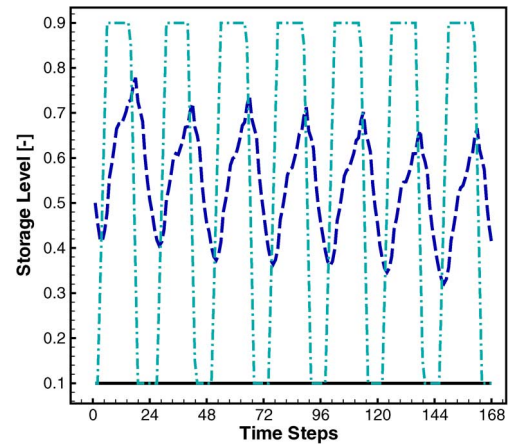Fig. 4. Wind energy (average of 50 simulated paths).

Fig. 5. Storage levels corresponding to Policy_E (dotdashed blue line) and Policy_M (solid black line). The dashed blue line indicates the expected price (scaled by $(1/2P_0)$).

$\varepsilon_{\text{mean}}$ denote the relative difference between the CVaR and expected costs of policies, respectively. As the table indicates, by implementing Policy_C the incurred operational cost is only about 1% more than the optimal expected cost. However, by
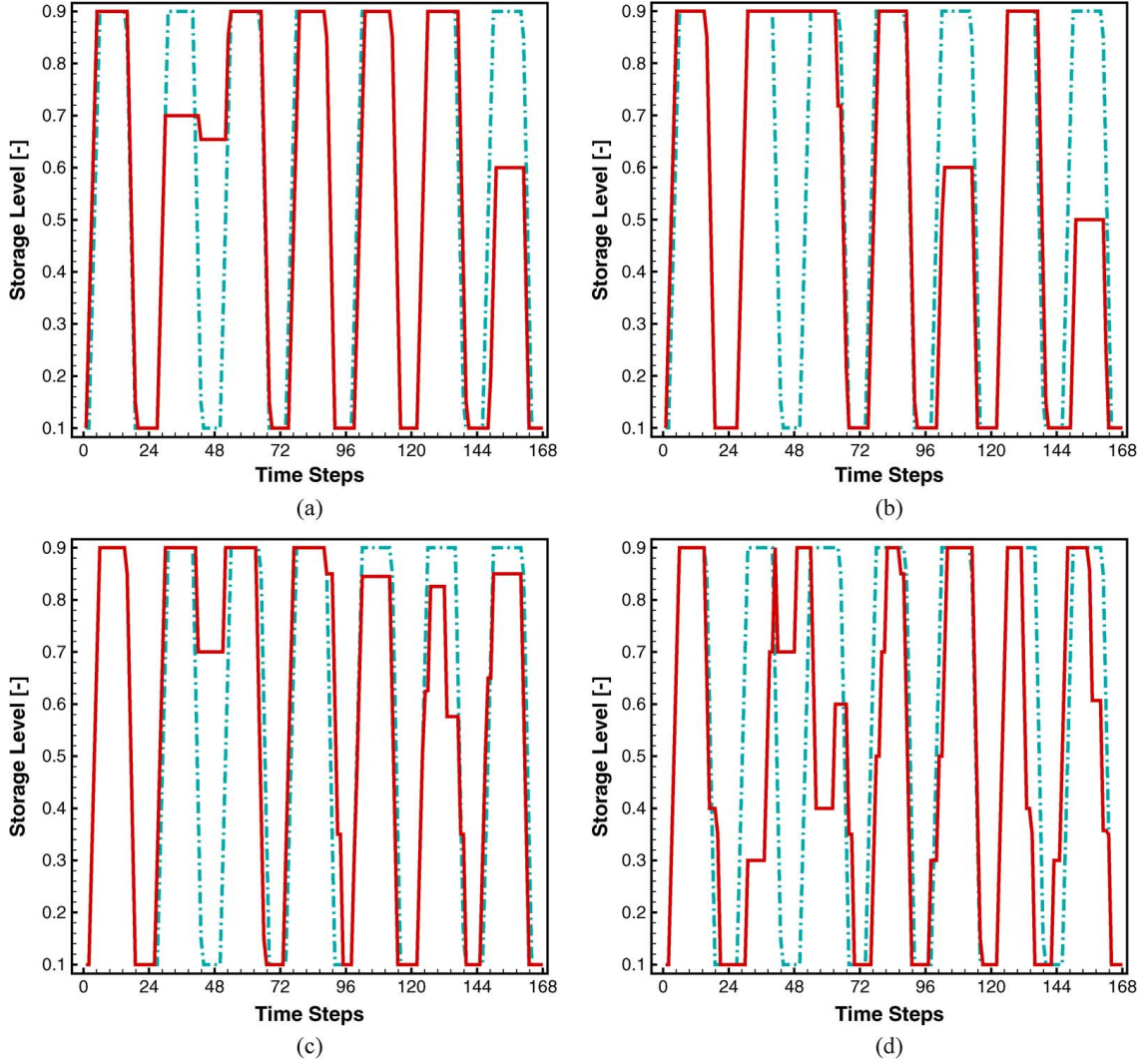
Fig. 6. Storage levels corresponding to Policy_C (solid red line) and Policy_E (dotdashed blue line). In these plots $\omega_{\mathbb{E}} = 1/51$ and $\omega_\rho = 50/51$ and the transaction cost parameters equal zero; (a) $\beta = 85\%$; (b)$\beta = 90\%$; (c) $\beta = 99.00\%$; (d) $\beta = 99.90\%$.

TABLE VII
EXPECTED VALUE AND CVaR_β OF THE OPERATIONAL COST

	Policy_C $\omega_E = 0,\ \omega_\rho = 1$	Policy_E	Policy_M
Mean	15, 483, 053.51	15, 458, 318.42	15, 530, 033.15
$\varepsilon_{\text{mean}}$	0.16%	–	0.46%
$\text{CVaR}_{85\%}$	21, 216, 461.07	21, 529, 402.82	21, 576, 631.39
$\varepsilon_{\text{CVaR}_{85\%}}$	–	1.47%	1.70%
Mean	15, 522, 149.73	15, 458, 318.42	15, 530, 033.15
$\varepsilon_{\text{mean}}$	0.41%	–	0.46%
$\text{CVaR}_{90\%}$	21, 521, 120.95	22, 546, 664.41	22, 593, 562.81
$\varepsilon_{\text{CVaR}_{90\%}}$	–	4.77%	4.98%
Mean	15, 591, 503.39	15, 458, 318.42	15, 530, 033.15
$\varepsilon_{\text{mean}}$	0.86%	–	0.46%
$\text{CVaR}_{95\%}$	23, 782, 833.07	26, 158, 851.56	26, 290, 658.08
$\varepsilon_{\text{CVaR}_{95\%}}$	–	9.99%	10.54%
Mean	15, 625, 009.39	15, 458, 318.42	15, 530, 033.15
$\varepsilon_{\text{mean}}$	1.07%	–	0.46%
$\text{CVaR}_{99.90\%}$	30, 946, 785.70	35, 127, 905.09	35, 288, 053.30
$\varepsilon_{\text{CVaR}_{99.90\%}}$	–	13.51%	14.02%

adopting Policy_E the risk to contend with could be up to 13% more than what Policy_C would have yielded.

B. Sensitivity to Transaction Cost Parameters

In this section, we consider nonzero transaction cost parameters and investigate impacts of changes in the transaction cost parameters on the strategies Policy_E and Policy_C.

Fig. 7 illustrates the storage levels associated with the optimal storage policies for three instances of these parameters, when $\beta = 90\%, \omega_{\mathbb{E}} = 1/51$, and $\omega_\rho = 50/51$. A comparison of these plots with the plot (b) in Fig. 6 clearly suggests that the transaction cost parameters directly influence the structure of the optimal storage policy and its storage level. In Fig. 7(a), the profile of the storage corresponding to Policy_E is exactly the same of the storage level of Policy_E in Fig. 6(b), which had been obtained under the assumption that all of the transaction cost parameters are equal to zero. However, even for this small change in the transaction cost parameters, Policy_C yields a different storage profile within the time steps 64–67 and 123–140, and 151–162. As α^{GS} increases, Policy_E differs from the risk-neutral policy in Fig. 6(b) over the time steps 31–43, in which the storage will not be charged as it would have been previously. It makes sense since now buying energy from the grid to store in the storage device for next time steps incurs additional cost due to the transaction cost $\alpha^{GS} > 0$, and would no
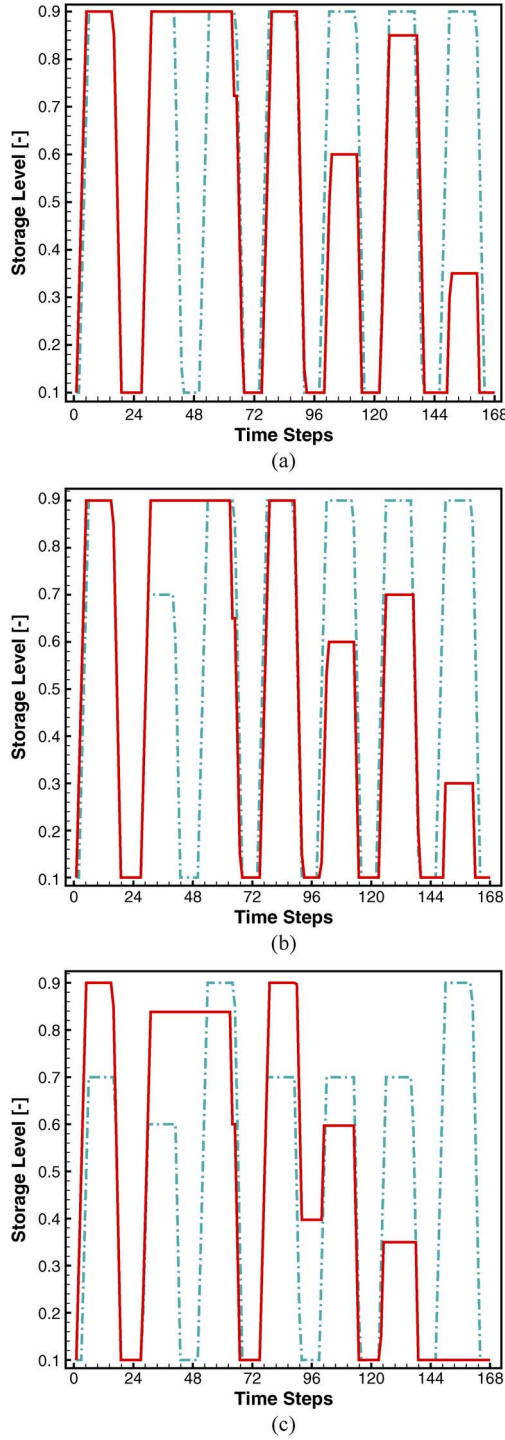
Fig. 7. Storage levels corresponding to Policy_C (solid red line) and Policy_E (dotdashed green line) with three non-zero values for α^{GS}. Other transaction cost parameters are equal to zero; (a) $\alpha^{GS} = 1$; (b) $\alpha^{GS} = 3$; (c) $\alpha^{GS} = 6$.

the CVaR risk measure with a positive risk aversion parameter ω_ρ. A thorough analysis on the impact of these parameters, when the objective function and constraints are linear, can be derived using the results from parametric linear programming (e.g., see [40]).

VI. CONCLUSIONS AND FUTURE WORK

An optimal energy storage management problem under the headings of risk consideration and transaction costs of exchanging energy with an electrical grid was studied in this paper. A renewable power generation source and energy storage are managed to fully satisfy an energy load with the capability to trade energy with the grid.

The fat tail behavior of energy prices and a risk analysis of the optimal risk neutral strategy were provided to motivate downside risk consideration in the problem. An efficient method to compute mean-CVaR optimal strategies was developed based on a smoothing piecewise quadratic function and an augmented Lagrangian method. Using the real world data from NYISO, we showed that the optimal risk aversion strategy can be quite different from the optimal risk neutral strategy. This difference becomes more prominent as the CVaR confidence level β increases. A sensitivity analysis of the optimal storage policies, their expected costs, and price risks was provided.

Several issues are left for future investigation. Extending this study for more general storage models, e.g., considering the impact of full charges or discharges on the storage life or constraints regarding safety concerns or time varying efficiency parameters, is an interesting direction for future work. Extending and applying the developed methodology and risk analysis for more realistic settings, e.g., other applications or other distributional model assumptions for the uncertain demands and particularly energy prices would also be a beneficial study. Furthermore, when reasonable values for loads and energy supplies are not available, the constraints determining admissible energy flows can be replaced with their corresponding chance constraints. Investigating the structure of the optimal risk averse and risk neutral policies with chance constraints to define admissible policies over the planning horizon would be insightful. Solution techniques such as [41] can then be adopted. Including capital cost and opportunity cost in the total operational cost within the risk averse operation management framework can be another research avenue.

For managing the operation of an energy storage, one may adopt a robust optimization approach with an appropriate uncertainty set, e.g., see [42] and [43] for the latest research on the theory of robust optimization, and [44] and [45] for applications of this approach for planning purposes in energy and power systems. It is interesting to compare the configuration and performance of the robust risk-averse energy storage operation management policy with those of the nominal risk-averse policy.

longer be beneficial. As this transaction cost increases, for example to $\alpha^{GS} = 6$, as in Fig. 7(c), the storage level further decreases during the time steps 3–16, 28–43, 75–91, 99–115, and 123–139. By increasing the transaction cost parameter α^{GS}, the storage level corresponding to Policy_C is further pushed toward the interior levels. This analysis suggests the importance of including the transaction cost parameters and their estimation in the energy storage management problem, particularly when a risk averse policy is sought and the objective function includes

REFERENCES

[1] P. F. Ribeiro, B. K. Johnson, M. L. Crow, A. Asroy, and Y. Liu, "Energy storage systems for advanced power applications," *Proc. IEEE*, vol. 89, no. 12, pp. 1744–1756, Dec. 2001.

[2] B. Dunn, H. Kamath, and J. M. Tarascon, "Electrical energy storage for the grid: A battery of choices," *Science*, vol. 334, no. 6058, pp. 928–935, 2011.

[3] H.-I. Su and A. E. Gamal, "Modeling and analysis of the role of fast-response energy storage in the smart grid," in *Proc. 9th Annual Allerton Conf. Communication, Control, Computing*, Urbana, IL, USA, 2011, pp. 719–726.

[4] J. Eyer and G. Corey, Energy Storage for the Electricity Grid: Benefits and Market Potential Assessment Guide, 2010, SANDIA Report.

[5] J. P. Barton and D. G. Infield, "Energy storage and its use with intermittent renewable energy," *IEEE Trans. Energy Convers.*, vol. 19, no. 2, pp. 441–448, Jun. 2004.

[6] P. D. Brown, J. A. P. Lopes, and M. A. Matos, "Optimization of pumped storage capacity in an isolated power system with large renewable penetration," *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 523–531, May 2008.

[7] J. G. Gonzalez, R. M. R. de la Muela, L. M. Santos, and A. M. Gonzalez, "Stochastic joint optimization of wind generation and pumped-storage units in an electricity market," *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 460–468, May 2008.

[8] N. Löhndorf and S. Minner, "Optimal day-ahead trading and storage of renewable energies: An approximate dynamic programming approach," *Energy Syst.*, vol. 1, pp. 61–77, 2010.

[9] R. Carmona and M. Ludkovski, "Valuation of energy storage: An optimal switching approach," *Quant. Finance*, vol. 10, no. 4, pp. 359–374, 2010.

[10] J. H. Kim and W. B. Powell, "Optimal energy commitments with storage and intermittent supply," *Oper. Res.*, vol. 59, no. 6, pp. 1347–1360, 2011.

[11] P. Harsha and M. Dahleh, "Optimal management and sizing of energy storage under dynamic pricing for the efficient integration of renewable energy," *IEEE Trans. Power Syst.*, to be published.

[12] H. Pandzic, J. M. Morales, A. J. Conejo, and I. Kuzle, "Offering model for a virtual power plant based on stochastic programming," *Appl. Energy*, vol. 105, pp. 282–292, 2013.

[13] H. Pandzic, I. Kuzle, and T. Capuder, "Virtual power plant mid-term dispatch optimization," *Appl. Energy*, vol. 101, pp. 134–141, 2013.

[14] A. Eydeland and K. Wolyniec, *Energy and Power Risk Management: New Developments in Modeling, Pricing, Hedging*, 1st ed. New York, NY, USA: Wiley, 2002.

[15] T. Kanamura and K. Ōhashi, "A structural model for electricity prices with spikes: Measurement of spike risk and optimal policies for hydropower plant operation," *Energy Econ.*, vol. 29, no. 5, pp. 1010–1032, 2007.

[16] M. Burger, B. Graeber, and G. Schindlmayr, *Managing Energy Risk: An Integrated View on Power and Other Energy Markets*, 1st ed. New York, NY, USA: Wiley, 2007.

[17] Z. Hameed, J. Vatn, and J. Heggset, "Challenges in the reliability and maintainability data collection for offshore wind turbines," *Renew. Energy*, vol. 36, pp. 2154–2165, 2011.

[18] U. S. Department of Energy Report, U. S. Wind Energy Production and Manufacturing Reaches Record Highs, 2013, pp. 1–3.

[19] P. Artzner, F. Delbaen, J. Eber, and D. Heath, "Coherent measures of risk," *Math. Finance*, vol. 9, pp. 203–228, 1999.

[20] G. C. Pflug and W. Romisch, *Modeling, Measuring and Managing Risk*. Singapore: World Scientific, August 2007.

[21] H. Follmer and A. Schied, *Stochastic Finance: An Introduction in Discrete Time*, 3rd ed. Berlin, Germany: De Gruyter, Jan. 2011.

[22] A. Downward, D. Young, and G. Zakeri, "Electricity contracting and policy choices under risk-aversion," *Oper. Res.*, pp. 1–43, 2012.

[23] M. Carrión, A. B. Philpott, A. J. Conejo, and J. M. Arroyo, "A stochastic programming approach to electric energy procurement for large consumers," *IEEE Trans. Power Syst.*, vol. 22, no. 2, pp. 744–754, May 2007.

[24] A. Cartea and M. Figueroa, "Pricing in electricity markets: A mean reverting jump diffusion model with seasonality," *Appl. Math. Finance*, vol. 12, no. 4, pp. 313–335, 2005.

[25] H. Geman and A. Roncoroni, "Understanding the fine structure of electricity prices," *J. Business*, vol. 79, no. 3, pp. 1225–1261, 2006.

[26] J. Hinman and E. Hickey, "Modeling and forecasting short-term electricity load using regression analysis," *Inst. Reg. Policy Studies*, pp. 1–51, 2009.

[27] L. J. Soares and M. C. Medeiros, "Modeling and forecasting short-term electricity load: A comparison of methods with an application to Brazilian data," *Int. J. Forecast.*, vol. 24, no. 4, pp. 630–644, 2008.

[28] T. Ackermann, *Wind Power in Power Systems*, 2nd ed. New York, NY, USA: Wiley, May 2012.

[29] C. Carrillo, A. F. Obando Montaño, J. Cidrás, and E. Díaz-Dorado, "Review of power curve modelling for wind turbines," *Renew. Sustain. Energy Rev.*, vol. 21, pp. 572–581, 2013.

[30] D. J. MacKay, *Sustainable Energy-Without the Hot Air*, 1 ed. Cambridge, U.K.: UIT Cambridge, 2009.

[31] D. C. Hill, D. McMillan, K. R. W. Bell, and D. Infield, "Application of auto-regressive models to U.K. wind speed data for power system impact studies," *IEEE Trans. Sustain. Energy*, vol. 3, no. 1, pp. 134–141, 2012.

[32] T. Burton, D. Sharp, N. Jenkins, and E. Bossanyi, *Wind Energy Handbook*. New York, NY, USA: Wiley, Jun. 2011.

[33] P. M. van de Ven, N. Hegde, L. Massoulie, and T. Salonidis, "Optimal control of end-user energy storage," *arXiv:1203.1891*, pp. 1–9, 2012.

[34] D. Duffie and J. Pan, "An overview of value at risk," *J. Deriv.*, vol. 4, no. 3, pp. 7–49, 1997.

[35] R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," *J. Risk*, vol. 2, no. 3, pp. 21–41, 2000.

[36] S. Alexander, T. F. Coleman, and Y. Li, "Minimizing VaR and CVaR for a portfolio of derivatives," *J. Bank. Finance*, vol. 30, no. 2, pp. 583–605, 2006.

[37] H. Xu and D. Zhang, "Smooth sample average approximation of stationary points in nonsmooth stochastic optimization and applications," *Math. Program.*, vol. 119, pp. 371–401, 2009.

[38] J. Nocedal and S. Wright, *Numerical Optimization*, 2nd ed. New York, NY, USA: Springer, Jul. 2006.

[39] T. F. Coleman and Y. Li, "An interior trust region approach for nonlinear minimization subject to bounds," *SIAM J. Optimiz.*, vol. 6, no. 2, pp. 418–445, 1996.

[40] A. Holder, "Parametric LP analysis," in *Wiley Encyclopedia of Operations Research and Management Science*, J. Cochran, Ed. New York, NY, USA: Wiley, 2010.

[41] A. Nemirovski and A. Shapiro, "Convex approximations of chance constrained programs," *SIAM J. Optimiz.*, vol. 17, no. 4, pp. 969–996, 2006.

[42] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Robust Optimization*. Princeton, NJ, USA: Princeton Univ. Press, 2009.

[43] D. Bertsimas, D. B. Brown, and C. Caramanis, "Theory and applications of robust optimization," *SIAM Rev.*, vol. 53, no. 3, pp. 464–501, 2011.

[44] F. Babonneau, J. P. Vial, and R. Apparigliato, "Robust optimization for environmental and energy planning," in *Handbook on Uncertainty and Environmental Decision Making*, ser. International Series in Operations Research and Management Science, J. A. Filar and A. Haurie, Eds. New York, NY, USA: Springer Verlag, 2010, pp. 79–126.

[45] A. H. Hajimiragha, C. A. Canizares, M. W. Fowler, S. Moazeni, and A. Elkamel, "A robust optimization approach for planning the transition to plug-in hybrid electric vehicles," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2264–2274, Nov. 2011.

Somayeh Moazeni (M'13) is a postdoctoral research associate in the Department of Operations Research and Financial Engineering in Princeton University, Princeton, NJ, USA.

Warren B. Powell (M'06) is a Professor in the Department of Operations Research and Financial Engineering at Princeton University, Princeton, NJ, USA, and the director of the Princeton laboratory for ENergy Systems Analysis (PENSA).

Amir H. Hajimiragha (S'09–M'10–SM'13) is a lead application engineer at the GE GridIQ Innovation Center, Markham, ON, Canada.