Modeling and Analysis of the Role of Fast-Response Energy Storage in the Smart Grid

Han-I Su and Abbas El Gamal
Department of Electrical Engineering
Stanford University
Stanford, CA 94305, USA
Email: hanisu@stanford.edu, abbas@ee.stanford.edu

Abstract—The large short time-scale variability of renewable energy resources presents significant challenges to the reliable operation of power systems. This variability can be mitigated by deploying fast-ramping generators. However, these generators are costly to operate and produce environmentally harmful emissions. Fast-response energy storage devices, such as batteries and flywheels, provide an environmentally friendly alternative, but are expensive and have limited capacity. To study the environmental benefits of storage, we introduce a slotted-time dynamic residual dc power flow model with the prediction error of the difference between the generation (including renewables) and the load as input and the fast-ramping generation and the storage (charging/discharging) operation as the control variables used to ensure that the demand is satisfied (as much as possible) in each time slot. We assume the input prediction error sequence to be i.i.d. zero-mean random variables. The optimal power flow problem is then formulated as an infinite horizon average-cost dynamic program with the cost function taken as a weighted sum of the average fast-ramping generation and the loss of load probability. We find the optimal policies at the two extremes of the cost function weights and propose a two-threshold policy for the general case. We also obtain refined analytical results under the assumption of Laplace distributed prediction error and corroborate this assumption using simulated wind power generation data from NREL.

I. INTRODUCTION

Electric power generation from fossil fuel contributes over 40% of the carbon emissions in the U.S. according to a 2011 EPA study [1]. These harmful emissions can be reduced by increasing the penetration of renewable generation from hydro, wind, and solar. However, the power generated from wind and solar is intermittent and uncertain, which presents significant challenges to power systems operation. The most commonly used approach to mitigating the variability of renewable generation is using medium and *fast-ramping* generators, such as combined-cycle combustion turbines and gas turbines (e.g., see [2]). As renewable energy penetration increases, however, more power from such generators is needed, which increases cost and offsets some of the environmental benefits of renewable energy [3]. More environmentally friendly means for mitigating renewable energy generation include exploiting geographic generation diversity [2], demand-response [4], and the use of energy storage [5], [6].

In particular, energy storage systems can be used for this purpose at two different time scales:

(i) Over a long time scale (days, hours), *bulk* energy storage systems, such as pumped hydroelectric storage and compressed air energy storage (CAES), can be "charged" by the excess renewable energy generation during off-peak hours and "discharged" during peak hours. In [7], a finite horizon dynamic dc optimal power flow problem with energy storage is formulated as a convex program. The round-trip storage efficiency is assumed to be perfect. For the single-bus case, under certain assumptions on the cost and the smoothness of the load profile, the optimal policy is to charge the storage at the beginning and discharge it at the end of the time period. In [8], the dynamic ac optimal power flow problem with energy storage is shown to be non-convex in general and sufficient conditions for strong duality are established.

(ii) Over a short time scale (minutes, seconds), *fast-response* energy storage systems, such as electric vehicle batteries and flywheels, can help smooth the output of renewable energy generators. In [9], the dynamic dc optimal power flow problem with energy storage is studied. The renewable generation is modeled as a sequence of discrete random variables. An approximate stochastic programming method is proposed and illustrated via numerical examples.

In this paper we focus on the short time scale use of storage. We consider a slotted-time model for demand and energy generation arising, for example, in real time (5–15 minute) dispatch. We view the generation as the sum of a predicted (forecasted) component and a prediction error component and similarly for the demand. We assume that the predicted (deterministic) component of the demand is met *perfectly* by the predicted component of the generation. Such balance in demand and generation can be achieved by solving a conventional (deterministic) static dc optimal power flow problem. We then consider a dynamic *residual* power flow problem with the difference between the generation and demand prediction errors as input and the fast-ramping generation and the storage (charging/discharging) operation as control sequences. To determine the policy for using fast-response energy storage and fast-ramping generation, we assume that the difference in prediction errors is a sequence of i.i.d. zero-mean random variables and formulate the dynamic

residual dc power flow problem as an infinite horizon average-cost dynamic program with the cost being a weighted sum of the average fast-ramping generation (which measures carbon emission cost) and loss of load probability. We find the optimal policies at the two extremes of the cost function weights and propose a two-threshold policy for the general case. At the one extreme, we find that storage can reduce the needed fast-ramping generation (relative to no storage) by a factor that approaches the round-trip storage inefficiency as storage capacity becomes large. At the other extreme, we find that the loss of load probability can be reduced to zero as storage capacity becomes large. Under the additional assumption of Laplace distributed prediction error, which is corroborated by simulated wind data from NREL [10], we also obtain closed form expressions for the cost function and the storage level and fast-ramping distributions in some special cases.

The next section introduces our residual power system model. In Section III, we present the problem formulation and summarize our results. The proofs of these results are given in [11]. In Section IV, we present numerical results using simulated Western wind dataset from NREL [10].

II. POWER SYSTEM MODELING

Consider an electric power system with conventional generators, renewable energy generators, loads, and energy storage devices. This power system may represent for example: (i) a transmission network with high renewable penetration, (ii) a distribution network with distributed renewable generators and energy storage devices, (iii) a microgrid not operated in the island mode where the power from the macrogrid acts as a fast-ramping generator [12], (iv) a wind farm with energy storage devices where the fast-ramping generation is bought from the electricity market, or (v) a stand-alone hybrid renewable energy system with battery storage [13], [14].

The power system consists of buses and connections between them. A bus may include a generator, a load, or an energy storage system directly connected to it. Due to the thermal constraint on each connection, there is a capacity limit on each connection. We will consider a dynamic dc power flow model (e.g., see [15]), which assumes that (i) the voltage magnitudes are constant, (ii) the angles of the complex bus voltages are small, (iii) the connections have no resistance, and (iv) the balance of the reactive power is ignored. Under these assumptions, the real power flowing through a connection can be approximated by a linear function of the phase angle difference of the two connected buses, where the slope of the linear function is the susceptance of the connection. Assuming that the phase angle of one of the bus be zero, then given the power flow injected into each bus such that the net injected power flow is zero, the remaining phase angles can be uniquely determined. Hence the dc power flow is also uniquely determined.

We consider a *slotted-time* model for the dynamics of the power system, where time is divided into slots each of length τ hours and power is constant over each time slot. In the numerical results in Section IV we assume $\tau = 1/6$ (10

minutes). Note that the constant power assumption implicitly implies that the balance of power at time scales shorter than τ can be achieved via regulation services. We now introduce the assumptions and models for the generators, loads, energy storage systems, and the single-bus residual power system.

Generation

We assume three classes of power generators;

(i) slow to moderate ramping generators, which include base load generators (coal-fired, hydro, and nuclear power plants), intermediate generators (combined-cycle combustion turbine), and peaking generators (gas turbines),

(ii) renewable generators (wind and solar) whose output is intermittent, and

(iii) fast-ramping generators (gas turbines) that are dedicated to compensating for the short time-scale variation in renewable generation.

The output from the slow to moderate ramping generators is assumed to be deterministic and is considered part of the predicted (forecasted) power flow.

Renewable generators: We model the power flow from the renewable generators as the sum of a predicted component and a prediction error component. The sum of this predicted component and the power from the slow to moderate ramping generators constitutes the total predicted generation power flow, which we assume to perfectly balance the predicted load. We model the prediction error component of renewable generation as a sequence of i.i.d. zero-mean random variables Δ_i, $i = 1, 2, \ldots$ with finite variance $E[\Delta^2] = \sigma^2$. We will also assume in some cases that the prediction error is a sequence of Laplace(λ) random variables, that is, the probability density function (pdf) of Δ_i is $f_\Delta(\delta) = (\lambda/2)e^{-\lambda|\delta|}$ for $\delta \in \mathbb{R}$ (and its standard deviation is $\sigma = \sqrt{2}/\lambda$). This assumption is akin to the exponential arrival and service time assumptions in queueing theory and Gaussian noise assumption in communication theory—all make their respective problems more tractable. In Section IV, we corroborate our Laplace assumption using simulated wind generation dataset from NREL.

Fast-ramping generators: These generators are used to compensate for the prediction error component of renewable generation. We assume that the total *generation power capacity* (the sum of the power ratings) of the fast-ramping generators is $G_{\max}$ MW, and denote the sequence of power supplied by these generators as G_i, $i = 1, 2, \ldots$, where $0 \leq G_i \leq G_{\max}$.

Load

As for the power flow for renewable energy, we model the load profile as the sum of a predicted component and a prediction error component. Since the load is less variable than the renewable generation under high penetration, we ignore the prediction error of the load profile.

Energy storage devices

We assume two types of energy storage devices, bulk energy storage and fast-response energy storage. The operation of the

bulk energy storage is scheduled according to the predicted component of the load and the energy generation (e.g., see [7], [8]), and is not considered further in this paper. The operation of the fast-response energy storage can follow the actual load and renewable generation. Let C_i and D_i, $i = 1, 2, \ldots$, be the *charging* and *discharging* power sequences of the fast-response energy storage, respectively. The storage is characterized by the following parameters.

- *Energy storage capacity* $\tau S_{\max}$ MW-h: This is the maximum amount of energy that can be stored. We refer to $S_{\max}$ as the *power storage capacity*. Generally the energy storage devices cannot be completely discharged, and there is also a limit on the minimum energy level. Here we use the minimum level as a reference and thus the lower limit is zero.
- *Rated storage output power* $D_{\max}$ MW: This is the maximum output (discharging) power, i.e., $0 \leq D_i \leq D_{\max}$ for all $i = 1, 2, \ldots$.
- *Rated storage power conversion* $C_{\max}$ MW: This is the maximum input (charging) power, i.e., $0 \leq C_i \leq C_{\max}$ for all $i = 1, 2, \ldots$.
- *Charging, discharging, and round-trip efficiencies*: The *charging efficiency* $\alpha_c \in (0,1)$ is the ratio of the charged power to the input power. The *discharging efficiency* $\alpha_d \in (0,1)$ is the ratio of the output power to the discharged power. The *round-trip storage efficiency* is $\alpha = \alpha_c \alpha_d$.
- *Storage efficiency*: This is the fraction of retained power over a time slot. We assume that the storage efficiency is very high and set it to one throughout.

Suppose that the *stored power* of the fast-response energy storage devices at the beginning of time slot $i = 1, 2, \ldots$ is S_i MW. Then the dynamics of fast-response energy storage can be represented as $S_{i+1} = S_i + \alpha_c C_i - (1/\alpha_d)D_i$ for $i = 1, 2, \ldots$ with constraints $0 \leq S_i \leq S_{\max}$, $0 \leq C_i \leq C_{\max}$, and $0 \leq D_i \leq D_{\max}$, where $S_1 = 0$.

Single-bus residual power system

We assume that the total predicted power flow (the sum of the power flow of the slow to moderated ramping generators, the predicted component of renewable generators, and the charging and discharging of the bulk energy storage systems) and the (predicted) load are perfectly balanced. Hence from this point onward, we consider only the *residual* dc power flow model consisting of the power flow of the fast-ramping generators, the prediction error component of the renewable generators, and the charging and discharging of the fast-response energy storage devices. Since the magnitude of the residual dc power flow is small relative to the bulk power flow, we assume that each connection in the residual power system has unlimited capacity. Hence, if the residual power system is still connected, a feasible dc power flow always exists when the net residual power flow injected into the system is zero. If the network is not connected because one or more connections is at capacity (due to the bulk power flow), the residual power system breaks up into several connected components, and our model and analysis applies to each component separately.
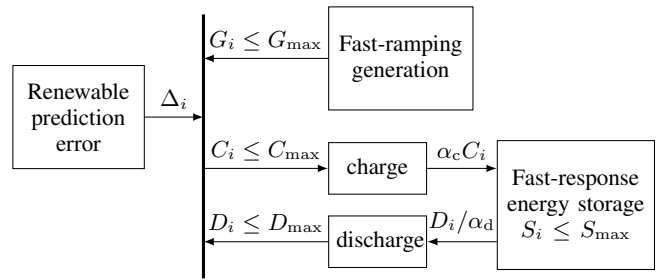

Fig. 1. A single-bus power system representation of the dynamic residual dc power flow model.

Now, assuming a connected residual power system with unlimited connected capacities, we can aggregate the fast-ramping generators, the prediction error component of renewable generators, and the fast-response energy storage devices into a residual single-bus power system as depicted in Figure 1. We require the (negative) prediction error to be balanced as much as possible by the storage and fast-ramping generation. More specifically, when $\Delta_i \geq -G_{\max} - \min\{\alpha_d S_i, D_{\max}\}$, there is sufficient power capacity, and the balance constraint must be satisfied, that is, $G_i + D_i - C_i + \Delta_i \geq 0$. Note that if $G_i + D_i - C_i + \Delta_i > 0$, then there is excess generation that we assume to be curtailed. When $\Delta_i < -G_{\max} - \min\{\alpha_d S_i, D_{\max}\}$, then loss of load occurs. In this case, the fast-ramping generation is at the power capacity and the fast-response storage is discharged at the rated storage output power, i.e., $G_i = G_{\max}$, $C_i = 0$, and $D_i = \min\{\alpha_d S_i, D_{\max}\}$.

III. PROBLEM FORMULATION AND SUMMARY OF RESULTS

We wish to investigate the long-term role of fast-response energy storage in reducing the required fast-ramping generation and the loss of load probability under the assumptions introduced in the previous section, namely: (i) dc power flow, (ii) perfect balance of the total predicted power flow and the load, (iii) unlimited connection capacities of the residual power system, and (iv) i.i.d. zero-mean prediction error. Hence, we consider the single-bus residual power system depicted in Figure 1 and formulate the associated dynamic optimal power flow problem as an infinite horizon average-cost dynamic program (e.g. see [16]). For simplicity, we consider only the case where the rated storage output power and power conversion of fast-response energy storage are unconstrained, i.e., $\alpha_c C_{\max} = (1/\alpha_d)D_{\max} = S_{\max}$.

- *State space*: The stored power of the fast-response energy storage (stored power) s_i is the state at time $i = 1, 2, \ldots$, which is in the state space $[0, S_{\max}]$.
- *Control space*: The control of the fast-ramping generation and the charging and discharging of the fast-response energy storage at time i consists of deterministic non-negative mappings $g_i(s_i, \cdot)$, $c_i(s_i, \cdot)$, $d_i(s_i, \cdot)$ indexed by s_i. The control space $\mathcal{A}(s_i)$ is the set of these mappings

satisfying

$$0 \le g_i(s_i, \delta_i) \le G_{\max},$$
$$0 \le s_i + \alpha_c c_i(s_i, \delta_i) - (1/\alpha_d) d_i(s_i, \delta_i) \le S_{\max},$$
$$0 \le g_i(s_i, \delta_i) + d_i(s_i, \delta_i) - c_i(s_i, \delta_i)$$
$$\qquad + \max\{\delta_i, -G_{\max} - \alpha_d s_i\}$$

for all prediction error $\delta_i \in \mathbb{R}$. Note that the control space does not depend on the time index.

- *Dynamic system*: The stored power can be express as a dynamic system

$$S_{i+1} = S_i + \alpha_c c_i(S_i, \Delta_i) - \frac{1}{\alpha_d} d_i(S_i, \Delta_i)$$

for $i = 1, 2, \ldots$, where $S_1 = 0$, and the prediction error component of renewable generation (prediction error) Δ_i, $i = 1, 2, \ldots$ is a sequence of i.i.d. zero-mean random variables.

- *Policy*: A policy $\pi = \{\pi_i : i = 1, 2, \ldots\}$ is a sequence of controls, where $\pi(s_i) \in \mathcal{A}(s_i)$ for all $s_i \in [0, S_{\max}]$ and $i = 1, 2, \ldots$. Let $\mathcal{P}$ be the set of all policies. A policy is *stationary* if $\pi_i = \pi_j$ for all $i \ne j$.

- *Average cost*: For a policy $\pi \in \mathcal{P}$, we define two types of costs—the expected average fast-ramping generation (carbon emission cost) and the expected average loss of load probability—as follows

$$J_0(\pi) = \limsup_{n \to \infty} \mathrm{E}\left[\frac{1}{n} \sum_{i=1}^{n} g_i(S_i, \Delta_i) \right],$$

$$J_1(\pi) = \limsup_{n \to \infty} \mathrm{E}\left[\frac{1}{n} \sum_{i=1}^{n} F_\Delta(-G_{\max} - \alpha_d S_{i+1}) \right],$$

where F_Δ is the cumulative distribution function (cdf) of Δ_i and the expectations are over Δ_i, $i = 1, 2, \ldots$. We take the total average cost $J(\pi)$ to be a weighted sum of these two costs with weights $\rho_0, \rho_1 \ge 0$, i.e.,

$$J(\pi) = \rho_0 J_0(\pi) + \rho_1 J_1(\pi).$$

- *Minimum average cost* and *optimal policy*: The minimum average cost is $J^* = \inf_{\pi \in \mathcal{P}} J(\pi)$, and the policy π^* achieves the infimum is the optimal policy.

We first consider the two extreme cases $(\rho_0, \rho_1) = (1, 0)$ and $(\rho_0, \rho_1) = (0, 1)$, and find the optimal policy in each case. In Subsection III-C, we propose a stationary two-threshold policy for general (ρ_0, ρ_1) that includes these two optimal policies as special cases.

A. $(\rho_0, \rho_1) = (1, 0)$ Case

When $(\rho_0, \rho_1) = (1, 0)$, we wish to minimize the fast-ramping generation regardless of the loss of load probability. In the following theorem, we show that a simple greedy charging and discharging policy is optimal.

Theorem 1: For $(\rho_0, \rho_1) = (1, 0)$, the optimal policy π^* is given in Table I, and it is stationary.

When the prediction error $\delta \ge 0$, the optimal policy charges the storage using the excess renewable generation as much as

$\pi_i^*(s)$	$g_i^*(s, \delta)$	$c_i^*(s, \delta)$	$d_i^*(s, \delta)$
$(1/\alpha_c)(S_{\max} - s) \le \delta$	0	$(1/\alpha_c)(S_{\max} - s)$	0
$0 \le \delta < (1/\alpha_c)(S_{\max} - s)$	0	δ	0
$-\alpha_d s \le \delta < 0$	0	0	$-\delta$
$-G_{\max} - \alpha_d s \le \delta < -\alpha_d s$	$-\delta - \alpha_d s$	0	$\alpha_d s$
$\delta < -G_{\max} - \alpha_d s$	$G_{\max}$	0	$\alpha_d s$

possible. When the prediction error $\delta < 0$, the storage is first discharged to compensate for as much of the renewable power deficit as possible. The fast-ramping generation is then used to compensate for the remaining renewable generation deficit (if any). Thus, the optimal policy never charges the storage using fast-ramping generation. This greedy policy is the result of the linearity of the cost function J_0 and the imperfect round-trip storage efficiency.

Now we consider the extreme case where the fast-ramping generation capacity $G_{\max}$ is unlimited. If there is no storage, that is, $S_{\max} = 0$, then the average cost is $\mathrm{E}[\Delta^+]$, where $x^+ = \max\{x, 0\}$. For unlimited power storage capacity $S_{\max}$, the minimum average cost is given in the following proposition.

Proposition 1: For $(\rho_0, \rho_1) = (1, 0)$ and unlimited $G_{\max}$ and $S_{\max}$, the minimum average cost is $J_0(\pi^*) = (1 - \alpha) \mathrm{E}[\Delta^+]$.

Comparing the average costs for no storage and unlimited power storage capacity, we show that storage can reduce the amount of needed fast-ramping generation (relative to no storage) by a factor no smaller than the round-trip storage inefficiency. This is not surprising because our i.i.d. zero-mean prediction error assumption implies that over the long term, the excess energy is roughly equal to the deficit. With infinite capacity and $\alpha = 1$, storage can compensate for almost all the variation in renewable generation. However, when $\alpha < 1$, it can compensate for at most this fraction of the variation and the rest needs to be compensated for by fast-ramping generation.

For the rest of this subsection, we assume that the prediction error Δ_i, $i = 1, 2, \ldots$ is a sequence of i.i.d. zero-mean Laplace(λ) random variables. First, we establish the following closed-form expression for the minimum average cost.

Proposition 2: For $(\rho_0, \rho_1) = (1, 0)$, the minimum average cost under the Laplace assumption is

$$J_0(\pi^*) = \frac{1 - e^{-\lambda G_{\max}}}{2\lambda} \left(\frac{1 - \alpha}{1 - \alpha e^{-(1/\alpha_c - \alpha_d)\lambda S_{\max}/2}} \right).$$

We consider the sizing of the storage using the above closed-from expression. The effectiveness of storage with capacity $S_{\max}$ in reducing the fast-ramping generation can be measured by the derivative

$$\left| \frac{\partial J_0(\pi^*)}{\partial S_{\max}} \right| = \frac{\alpha_d (1 - e^{-\lambda G_{\max}})(1 - \alpha)^2 e^{-(1/\alpha_c - \alpha_d)\lambda S_{\max}/2}}{4 \left(1 - \alpha e^{-(1/\alpha_c - \alpha_d)\lambda S_{\max}/2} \right)^2},$$

which decreases close to exponentially in $S_{\max}$. Hence only a small power storage capacity is sufficient to achieve most

of the reduction in fast-ramping generation. For example, for typical round-trip storage efficiency 60%–80% (e.g., see [17]), 80% of this reduction can be achieved with power storage capacity less than 4 standard deviations of the prediction error. *Stationary distribution of stored power*: Note that under the optimal policy in Theorem 1, the stored power sequence is a homogeneous Markov process. In the following, we find the stationary distribution for this Markov process under the Laplace assumption. Furthermore, using this stationary distribution, we can find the distribution of the fast-ramping generation.

Proposition 3: The cdf of the stationary distribution of the stored power under the optimal policy in Theorem 1 and the Laplace assumption is

$$F_S(s) = \frac{1 - 0.5(1+\alpha)e^{-(1/\alpha_{\mathrm{c}} - \alpha_{\mathrm{d}})\lambda s/2}}{1 - \alpha e^{-(1/\alpha_{\mathrm{c}} - \alpha_{\mathrm{d}})\lambda S_{\max}/2}}$$

for $0 \le s < S_{\max}$, $F_S(s) = 0$ for $s < 0$, and $F_S(s) = 1$ for $s \ge S_{\max}$. The corresponding distribution of the fast-ramping generation is

$$F_G(g) = 1 - \frac{1-\alpha}{2(1 - \alpha e^{-(1/\alpha_{\mathrm{c}} - \alpha_{\mathrm{d}})\lambda S_{\max}/2})} e^{-\lambda g}$$

for $0 \le g < G_{\max}$, $F_G(g) = 0$ for $g < 0$, and $F_G(g) = 1$ for $g \ge G_{\max}$.

Expected average loss of load probability: Using the stationary distribution of the stored power sequence in Proposition 3, we can readily find the following expression for the expected average loss of load probability

$$J_1(\pi^*) = \frac{1}{2} e^{-\lambda G_{\max}} \left(\frac{1-\alpha}{1 - \alpha e^{-(1/\alpha_{\mathrm{c}} - \alpha_{\mathrm{d}})\lambda S_{\max}/2}} \right).$$

For no storage, $J_1(\pi^*) = (1/2)e^{-\lambda G_{\max}}$. As $S_{\max}$ tends to ∞, $J_1(\pi^*) = ((1-\alpha)/2)e^{-\lambda G_{\max}}$. Thus, using the optimal policy for $(\rho_0, \rho_1) = (1, 0)$, storage can reduce the expected average loss of load probability by as much as the round-trip storage inefficiency.

B. $(\rho_0, \rho_1) = (0, 1)$ Case

When $(\rho_0, \rho_1) = (0, 1)$, we wish to minimize the expected average loss of load probability. To minimize this probability, we would like to keep the stored power as high as possible. This turns out to be the optimal policy in this case.

Theorem 2: For $(\rho_0, \rho_1) = (0, 1)$, the optimal policy π^* minimizing the expected loss of load probability is given in Table II, and it is stationary.

Since there is no cost of fast-ramping generation, in some cases the optimal policy charges the storage with fast-ramping generation to minimize the loss of load probability. In the following, we show that the benefits of storage to the loss of load probability can be unbounded.

Proposition 4: Using the optimal policy for $(\rho_0, \rho_1) = (0, 1)$, if F_Δ and $G_{\max}$ satisfy $\limsup_{x \to \infty} x F_\Delta(-x) \le c$ for some constant $c \ge 0$ and

$$\mathrm{E}\left[\alpha_{\mathrm{c}}(G_{\max} + \Delta)^+ - (1/\alpha_{\mathrm{d}})(G_{\max} + \Delta)^- \right] > 0, \quad (1)$$

where $x^+ = \max\{x, 0\}$ and $x^- = \max\{-x, 0\}$, then the expected average loss of load probability tends to 0 as $S_{\max}$ tends to ∞.

Unlike the case of $(\rho_0, \rho_1) = (1, 0)$, we are not able to find closed form expressions for the optimal cost functions or the stationary distributions under the Laplace assumption. However, it can be verified that for $\alpha > e^{-\lambda G_{\max}}$, Laplace$(\lambda)$ satisfies the sufficient conditions in Proposition 4, and thus the expected average loss of load probability tends to 0 as $S_{\max}$ tends to ∞. When there is no fast-ramping generation, i.e., $G_{\max} = 0$, the sufficient condition in Proposition 4 given by (1) does not hold. However, the optimal policy reduces to a special case of the optimal policy in Theorem 1. Thus, storage can only reduce the expected loss of load probability by a factor no smaller than the round-trip storage inefficiency.

C. Two-threshold policy

For the general average cost with weights $\rho_0, \rho_1 \ge 0$, we propose the following stationary two-threshold policy.

Two-threshold Policy: Let $0 \le S_{\mathrm{c}} \le S_{\mathrm{d}} \le S_{\max}$. The two-threshold policy is parameterized by a "charging threshold" S_{c} and a "discharging threshold" S_{d} as given in Tables III.

When the prediction error $\delta \ge 0$, the storage is charged as much as possible using the excess in renewable generation. If the stored power after this charging is above S_{c}, then the fast-ramping generation is not used. However, if it is below S_{c}, then the storage is charged as close to S_{c} as possible using fast-ramping generation. When the prediction error $\delta < -G_{\max}$, the storage must be discharged to balance the prediction error. If there is still unbalanced prediction error after the storage is discharged to S_{d}, then fast-ramping generation is used such that the the stored power is as close to S_{d} as possible. When the prediction error $-G_{\max} \le \delta < 0$, the case where the stored power is either lower than S_{c} or higher than S_{d} is similar to the above cases. When the stored power is between S_{c} and S_{d}, only the fast-ramping generation is used to balance the prediction error, and the stored power is unchanged.

Note that the optimal policies for the two extreme cases in Theorem 1 and 2 are two-threshold policies with parameters $(0, 0)$ and $(S_{\max}, S_{\max})$, respectively.

TABLE II
OPTIMAL POLICY IN THEOREM 2 FOR $(\rho_0, \rho_1) = (0, 1)$.

$\pi_i^*(s)$	$g_i^*(s, \delta)$	$c_i^*(s, \delta)$	$d_i^*(s, \delta)$
$0 \le \delta - \frac{S_{\max} - s}{\alpha_{\mathrm{c}}}$	0	$\frac{S_{\max} - s}{\alpha_{\mathrm{c}}}$	0
$-G_{\max} \le \delta - \frac{S_{\max} - s}{\alpha_{\mathrm{c}}} < 0$	$\frac{S_{\max} - s}{\alpha_{\mathrm{c}}} - \delta$	$\frac{S_{\max} - s}{\alpha_{\mathrm{c}}}$	0
$0 \le \delta + G_{\max} < \frac{S_{\max} - s}{\alpha_{\mathrm{c}}}$	$G_{\max}$	$G_{\max} + \delta$	0
$-\alpha_{\mathrm{d}} s \le \delta + G_{\max} < 0$	$G_{\max}$	0	$-\delta - G_{\max}$
$\delta + G_{\max} < -\alpha_{\mathrm{d}} s$	$G_{\max}$	0	$\alpha_{\mathrm{d}} s$

TABLE III
TWO-THRESHOLD POLICY PARAMETERIZED BY (S_c, S_d).

(a) $0 \le s < S_c$

$\pi_i^*(s)$	$g_i^*(s,\delta)$	$c_i^*(s,\delta)$	$d_i^*(s,\delta)$
$(1/\alpha_c)(S_{max} - s) \le \delta$	0	$(1/\alpha_c)(S_{max} - s)$	0
$(1/\alpha_c)(S_c - s) \le \delta < (1/\alpha_c)(S_{max} - s)$	0	δ	0
$(1/\alpha_c)(S_c - s) - G_{max} \le \delta < (1/\alpha_c)(S_c - s)$	$(1/\alpha_c)(S_c - s) - \delta$	$(1/\alpha_c)(S_c - s)$	0
$-G_{max} \le \delta < (1/\alpha_c)(S_c - s) - G_{max}$	G_{max}	$G_{max} + \delta$	0
$-G_{max} - \alpha_d s \le \delta < -G_{max}$	G_{max}	0	$-G_{max} - \delta$
$\delta < -G_{max} - \alpha_d s$	G_{max}	0	$\alpha_d s$

(b) $S_c \le s \le S_d$

$\pi_i^*(s)$	$g_i^*(s,\delta)$	$c_i^*(s,\delta)$	$d_i^*(s,\delta)$
$(1/\alpha_c)(S_{max} - s) \le \delta$	0	$(1/\alpha_c)(S_{max} - s)$	0
$0 \le \delta < (1/\alpha_c)(S_{max} - s)$	0	δ	0
$-G_{max} \le \delta < 0$	$-\delta$	0	0
$-G_{max} - \alpha_d s \le \delta < -G_{max}$	G_{max}	0	$-G_{max} - \delta$
$\delta < -G_{max} - \alpha_d s$	G_{max}	0	$\alpha_d s$

(c) $S_d < s \le S_{max}$

$\pi_i^*(s)$	$g_i^*(s,\delta)$	$c_i^*(s,\delta)$	$d_i^*(s,\delta)$
$(1/\alpha_c)(S_{max} - s) \le \delta$	0	$(1/\alpha_c)(S_{max} - s)$	0
$0 \le \delta < (1/\alpha_c)(S_{max} - s)$	0	δ	0
$-\alpha_d(s - S_d) \le \delta < 0$	0	0	$-\delta$
$-\alpha_d(s - S_d) - G_{max} \le \delta < -\alpha_d(s - S_d)$	$-\delta - \alpha_d(s - S_d)$	0	$\alpha_d(s - S_d)$
$-G_{max} - \alpha_d s \le \delta < -\alpha_d(s - S_d) - G_{max}$	G_{max}	0	$-G_{max} - \delta$
$\delta < -G_{max} - \alpha_d s$	G_{max}	0	$\alpha_d s$

IV. NUMERICAL RESULTS

We compare the theoretical results in the previous section with numerical results using the simulated Western wind dataset from NREL [10], which is based on numerical weather prediction (NWP) models. This dataset recreates the potential wind power generation of more than 30 000 wind turbines in the western U.S. from 2004 to 2006, and the wind power data are sampled every 10 minutes. We select 50 offshore sites in California with highest power densities. Figure 2a depicts the total power output of these 50 sites in every 10 minutes over a two weeks period. Since this dataset does not include forecast data, we illustrate our results using the following simple 10-minute-ahead linear predictor with maximum likelihood (ML) estimator for the Laplace parameter. Other more sophisticated predictors can be readily used with the ML estimator to obtain more refined results. Let r_i, $i = 1, 2, \ldots, n$ be the simulated wind power generation sequence. We consider a 10-minute-ahead linear predictor of the wind power at time i based on the 6 samples of the generation in the past hour $\hat{r}_i = c_0 + \sum_{j=1}^{6} c_j r_{i-j}$. The coefficients $c_0, c_1, \ldots, c_6$ chosen minimize the squared error of the data of 2004 used as a training set. Let $\delta_i = r_i - \hat{r}_i$, $i = 1, 2, \ldots, n$ be the prediction error. Figures 2b and 2c depict the 10-minute-ahead prediction and the prediction error sequences for the two-week period, respectively. Based on the Laplace assumption, the ML estimate of parameter λ is $\lambda = n/\sum_{i=1}^{n} |\delta_i|$. Figure 3 plots the empirical pdf of the prediction error and the best fit Laplace pdf. The maximum absolute difference between the empirical cdf of the prediction error and the Laplace cdf is 0.018.

A. Results for $(\rho_0, \rho_1) = (1, 0)$

Figure 4 compares the minimum average costs in Proposition 2 and using the three-year simulated wind data for various values of power storage capacities and round-trip storage efficiencies $\alpha = 60\%$ and 80%. The maximum absolute difference between the theoretical and the simulated costs normalized by the theoretical cost is less than 6% and 8% for $\alpha = 60\%$ and $\alpha = 80\%$, respectively. Thus, the Laplace distribution appears to be an acceptable approximation of the simulated wind generation data from NREL.

Figure 5a compares the empirical pdf of the stored power of the simulated wind generation data to the stationary pdf under the Laplace distribution assumption in Proposition 3. The corresponding empirical pdf of the fast-ramping generation and its stationary pdf are shown in Figure 5b. Note again the simulation results corroborate well with the theory.

B. Results for general (ρ_0, ρ_1)

To demonstrate the two-threshold policy, we implemented a dynamic programming method by discretizing the state space and then running the value iteration [18]. For the values of G_{max}, S_{max}, ρ_0, and ρ_1 used in the following numerical examples, we find that the policy obtained from the value iteration is a discretized two-threshold policy.

(a) Wind power
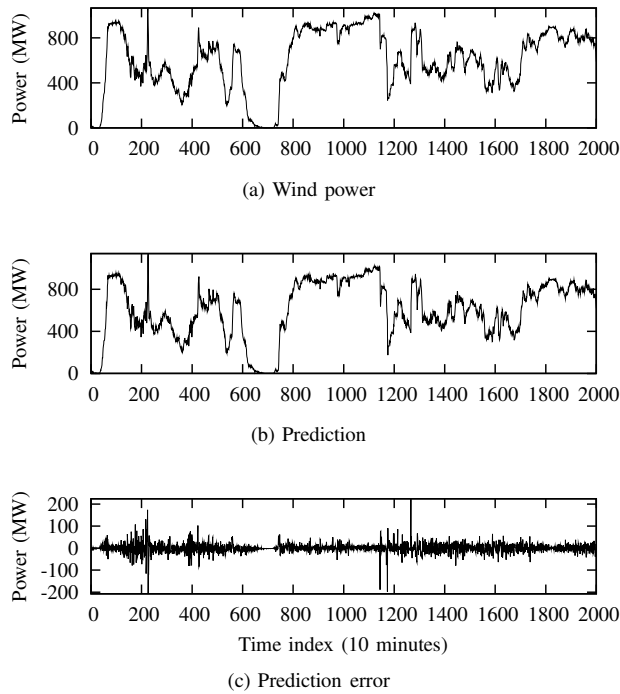

(b) Prediction

(c) Prediction error

Fig. 2. Figures (a), (b), and (c) show the wind power, 10-minute-ahead prediction, and prediction error sequences for two weeks, respectively. The average wind power is 560.26 MW, the mean absolute value of the prediction error is 13.99 MW, and the standard deviation of the prediction error is 20.88 MW.
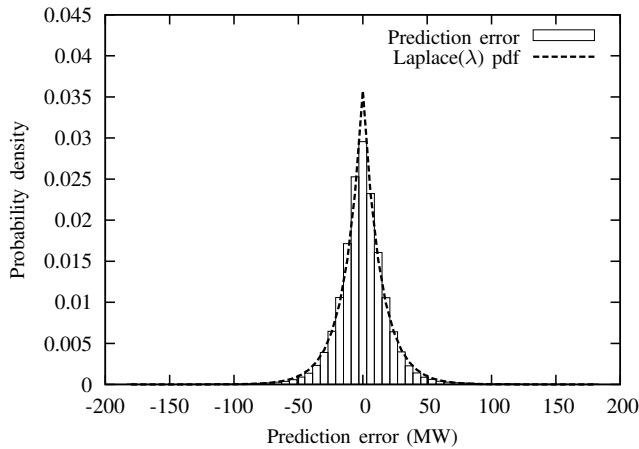

Fig. 3. The empirical pdf of the 10-minute-ahead wind power prediction error for three years versus the Laplace(λ) pdf with $1/\lambda = 13.99$.
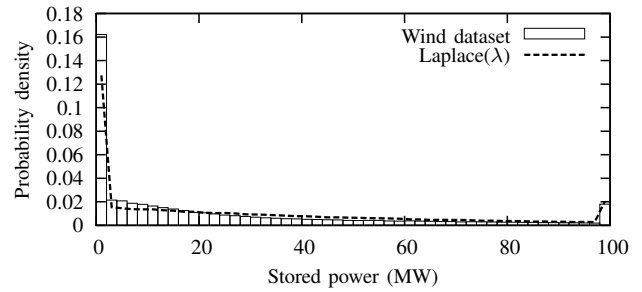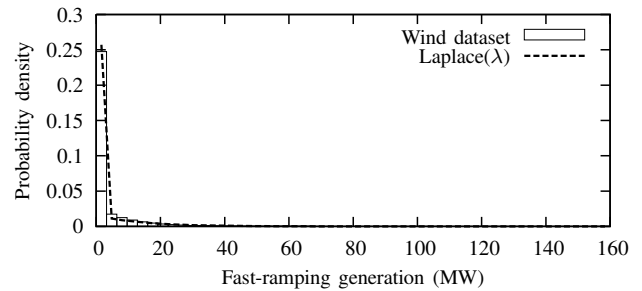

Fig. 4. The optimal expected average fast-ramping generation for the wind data versus power storage capacity for round-trip efficiencies $\alpha = 60\%$ and 80% and $G_{\max} = 160$ MW.

(a) Fast-response storage

(b) Fast-ramping generation

Fig. 5. Figures (a) and (b) show the pdfs of the stored power and the fast-ramping generation, respectively, for $\alpha = 60\%$, $S_{\max} = 100$ MW, and $G_{\max} = 160$ MW.

Tradeoff between J_0 and J_1: Figure 6 shows the tradeoff between the fast-ramping generation and the loss of load probability for no storage and for storage capacities $S_{\max} = 50$ MW and $S_{\max} = 100$ MW with fast-ramping generation capacity $G_{\max} = 160$ MW. Note that the results for the Laplace pdf corroborate very well with the simulated wind generation data. As shown in the figure the loss of load probability is improved by more than two orders of magnitude by using power storage capacity less than 5 standard deviations

of the prediction error.

Empirical distribution of stored power: The empirical pdf of the stored power for $(\rho_0, \rho_1) = (10^{-6}, 1)$ is shown in Figure 7. We expect the probability concentrates at 0, S_c, S_d, and $S_{\max}$. The empirical pdf in Figure 7 indeed has two peaks at S_d and $S_{\max}$. The probability that the stored power is below S_d is very small since the transition probability of the stored power from a value above S_d to below S_d is less than $F_\Delta(-G_{\max}) = 5.4 \cdot 10^{-6}$.

Tradeoff between $S_{\max}$ and $G_{\max}$: In Figure 8, we compare

the two ways of mitigating renewable energy variability; using fast-ramping generation and using fast-response storage. We fix the expected average fast-ramping generation at 3.6 MW (corresponding to 80% maximum reduction in the fast-ramping generation) and the loss of load probability at $2 \cdot 10^{-6}$ (corresponds to one loss of load event every 10 years). To achieve these goals with minimum power storage capacity, we need $G_{\max} = 170$ MW and $S_{\max} = 60$ MW. To reduce the fossil fuel generation and to achieve the same goals, we can replace 1 MW of $G_{\max}$ with 1.3 MW of $S_{\max}$.
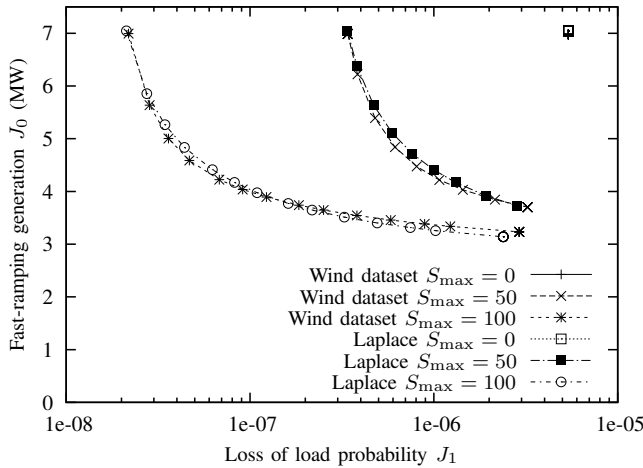
Fig. 6. The tradeoff between the fast-ramping generation and the loss of load probability for $G_{\max} = 160$ MW under two-threshold policies.
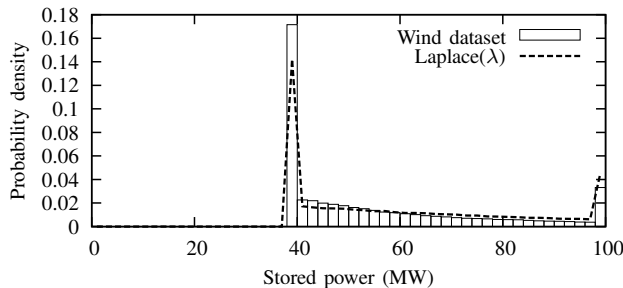

Fig. 7. The pdfs of the stored power under the two-threshold policy with $\alpha = 60\%$, $G_{\max} = 160$ MW, $S_{\max} = 100$ MW, $S_c = 12$ MW, and $S_d = 38$ MW.

V. ACKNOWLEDGMENTS

REFERENCES

[1] "Inventory of U.S. greenhouse gas emissions and sinks: 1990 – 2009," U.S. Environmental Protection Agency, Apr. 2011. [Online]. Available: http://epa.gov/climatechange/emissions/usinventoryreport.html
[2] "Western wind and solar integration study," The National Renewable Energy Laboratory, May 2010. [Online]. Available: http://www.nrel.gov/wind/systemsintegration/wwsis.html
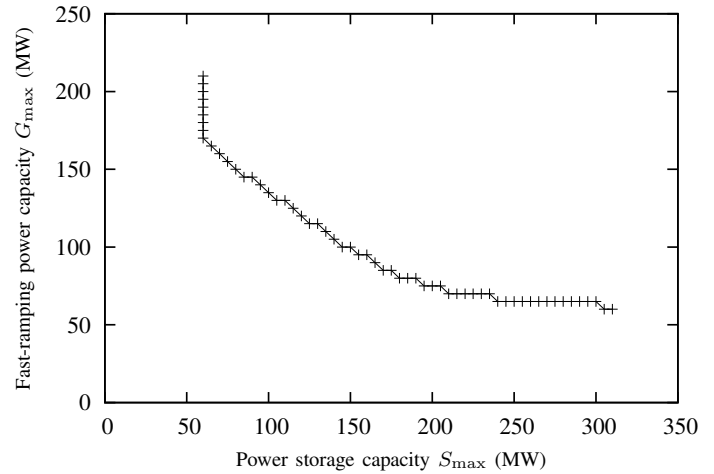
Fig. 8. The tradeoff between the fast-ramping generation power capacity and the power storage capacity for $J_0 \leq 3.6$ MW and $J_1 \leq 2 \cdot 10^{-6}$.

[3] E. K. Hart and M. Z. Jacobson, "A monte carlo approach to generator portfolio planning and carbon emissions assessments of systems with large penetrations of variable renewables," *Renewable Energy*, vol. 36, no. 8, pp. 2278–2286, 2011.
[4] B. Kirby and M. Milligan, "Utilizing load response for wind and solar integration and power system reliability," in *WindPower 2010*, Dallas, Texas, May 2010.
[5] P. Denholm, E. Ela, B. Kirby, and M. Milligan, "The role of energy storage with renewable electricity generation," National Renewable Energy Laboratory, Tech. Rep., Jan. 2010.
[6] E. Bitar, R. Rajagopal, P. Khargonekar, and K. Poolla, "The role of co-located storage for wind power producers in conventional electricity markets," in *2011 American Control Conference*, June 2011.
[7] K. Chandy, S. Low, U. Topcu, and H. Xu, "A simple optimal power flow model with energy storage," in *Decision and Control (CDC), 2010 49th IEEE Conference on*, Dec. 2010, pp. 1051 –1057.
[8] D. Gayme and U. Topcu, "Optimal power flow with distributed energy storage dynamics," in *American Control Conference (ACC), 2011*, 2011.
[9] H. Oh, "Optimal planning to include storage devices in power systems," *Power Systems, IEEE Transactions on*, 2011.
[10] C. Potter, D. Lew, J. McCaa, S. Cheng, S. Eichelberger, and E. Grimit, "Creating the dataset for the western wind and solar integration study," in *7th International Workshop on Large Scale Integration of Wind Power and on Transmission Networks for Offshore Wind Farms*, May 2008.
[11] H. Su and A. E. Gamal, "Modeling and analysis of the role of fast-response energy storage in the smart grid," 2011. [Online]. Available: http://arxiv.org/abs/1109.3841
[12] R. Lasseter, A. Akhil, C. Marnay, J. Stephens, J. Dagle, R. Guttromson, A. S. Meliopoulous, R. Yinger, , and J. Eto, "The certs microgrid concept," Apr. 2002. [Online]. Available: http://certs.lbl.gov/certs-der-micro.html
[13] M. Deshmukh and S. Deshmukh, "Modeling of hybrid renewable energy systems," *Renewable and Sustainable Energy Reviews*, vol. 12, no. 1, pp. 235–249, 2008.
[14] J. L. Bernal-Agustín and R. Dufo-López, "Simulation and optimization of stand-alone hybrid renewable energy systems," *Renewable and Sustainable Energy Reviews*, vol. 13, no. 8, pp. 2111–2118, 2009.
[15] R. Christie, B. Wollenberg, and I. Wangensteen, "Transmission management in the deregulated environment," *Proceedings of the IEEE*, vol. 88, no. 2, pp. 170–195, Feb. 2000.
[16] A. Arapostathis, V. S. Borkar, E. Fernández-Gaucherand, M. K. Ghosh, and S. I. Marcus, "Discrete-time controlled markov processes with average cost criterion: A survey," *SIAM Journal on Control and Optimization*, vol. 31, no. 2, pp. 282–344, 1993.
[17] R. Schainker, "Executive overview: energy storage options for a sustainable energy future," in *Power Engineering Society General Meeting, 2004. IEEE*, June 2004, pp. 2309–2314.
[18] D. Bertsekas, Ed., *Dynamic Programming and Optimal Control*, 3rd ed. Athena Scientific, 2007.