Benefits of Storage Control for Wind Power Producers in Power Markets

Mainak Chowdhury, *Student Member, IEEE,* Milind Rao, *Student Member, IEEE,* Yue Zhao, *Member, IEEE,* Tara Javidi, *Senior Member, IEEE* Andrea Goldsmith, *Fellow, IEEE,*

***Abstract*—We consider a wind power producer (WPP) participating in a dynamically evolving two settlement power market. We study the utility of energy storage for a WPP in maximizing its expected profit. With random wind and price processes, the optimal forward contract and storage charging/discharging decisions are formulated as solutions of an infinite horizon stochastic optimal control problem. For the asymptotically small storage capacity regime, we precisely characterize the maximum profit increase brought by utilizing energy storage. We prove that, in this regime, an optimal policy uses storage to compensate for power delivery shortfall/surplus in real time, *without* changing the forward contracts from the optimal ones in the absence of energy storage. This policy also serves as an approximately optimal policy for the case of relatively small storage capacity. We also design a policy based on model predictive control (MPC) that is approximately optimal for general storage capacities. We numerically evaluate the developed policies for wind and price processes with representative statistics from real world data. It is observed that, as expected, the simple small storage approximation policy performs closely to the optimum when storage is relatively small, while the more complex stochastic MPC policy performs better for larger storage capacities.**

I. Introduction

Wind energy contributed to 17% of load serving by renewable energies in the U.S. in 2014 [1]. Globally, the installed wind power capacity has increased eight-fold in the past decade, and continues to grow thanks to the decreasing capital costs of harnessing wind power. A typical approach for integrating wind energy into the electricity grid is to let wind power producers (WPPs) participate in conventional multi-settlement power markets, where power is sold in multiple forward (e.g., day-ahead, hour-ahead, and 15-minutes ahead) markets. A summary describing participation of WPPs for multi-settlement markets in several Independent System Operators (ISOs) in the U.S. can be found in [2]. Such markets raise significant challenges for WPPs, particularly in the day-ahead market, where the vast majority of contracts of power delivery are made. In particular, as wind power generation is non-dispatchable, committing to forward power contracts raises fundamental challenges to WPPs, as future wind power generation is inherently uncertain due to difficulties of wind forecast [3]. Hence, a WPP has the risk of running a shortfall in delivering a forward power contract when its actual generation is insufficient. Making up for such a shortfall by fast-responding generation resources (e.g., buying power in the real time market) is usually costly. As wind power generation itself has very low variable cost, the cost due to compensating for the risk of shortfall constitutes a WPP's most significant operating cost.

There are a variety of approaches for the WPPs to reduce the risk of delivery shortfall. Methods that exploit statistical characteristics of wind generation such as improving wind power forecast [3] and aggregating diverse wind sources [4] directly reduce the uncertainty in the WPPs' knowledge of future wind power generation. Fast-ramping fuel-based backup generators can also compensate for the uncertainty of future wind power, albeit with a high operating cost. Recently, energy storage has emerged as an increasingly viable commercial technology due to its decreasing cost [5]; the ability of energy storage to shift energy across time provides great flexibility for WPPs to charge/discharge energy to compensate for any generation shortfall/surplus to meet forward contracts. Nonetheless, as energy storage is currently still expensive for large scale deployments, assessing its value in reducing operating cost and increasing profit is of primary interest to WPPs.

Without storage, the problem of a single WPP participating in a two-settlement (day-ahead and real time) market has been studied in [6], [7], [8] where the optimal (i.e. expected profit maximizing) forward contracts based on the statistics of wind generation were developed. With storage, different variants of a stochastic control formulation have been considered in [9], [10], [11], [12], [13], [14] for setting either the day ahead contracts or the charging policy or both. Optimal storage charging policy is studied in the absence of a power market in [9] where the total variation of the power output of a WPP is minimized. The value of providing co-located energy storage to a WPP has been studied in [10] and [11] for real time market participation only, where dynamic programming (DP) solutions and online algorithms were developed, respectively. Focusing on a single day of wind generation and storage operation, the day-ahead market is considered in assessing the role of storage co-located with wind power in [12], where the optimal day-ahead contracts and storage operation were solved. A dynamically evolving two-settlement market has been considered in [13] and [14] in which optimal operation of co-located storage and wind power are studied. There, the model is limited to

This research was supported in part by a 3Com Corporation Stanford Graduate Fellowship, and in part by the NSF under CPS Synergy grant 1330081.

M. Chowdhury, M. Rao, and A. Goldsmith are with the Dept. of Electrical Engineering, Stanford University, Stanford, CA, 94305, USA, (e-mail: mainakch,milind,andreag@stanford.edu).

Y. Zhao is with the Dept. of Electrical and Computer Engineering, Stony Brook University, Stony Brook, NY, 11794 USA, (e-mail: yue.zhao.2@stonybrook.edu).

T. Javidi is with the Dept. of Electrical and Computer Engineering, UC San Diego, La Jolla, CA, 92093, USA, (email:tara@ece.ucsd.edu).

one in which the forward market trades power *one time slot ahead*. As will be clarified next, this corresponds to the case of $D = 1$ in this paper. We note that, when hourly day-ahead and hourly real-time markets are considered, D can be as large as 24, which significantly complicates the study of the problem. Another work that considers storage and WPPs competing (as opposed to cooperating) with each other in a power market is [15], where the Nash equilibria in this competitive setting are derived. We further note that the use of energy storage has also been extensively studied from a system operator's perspective on unit commitment and optimal power flow problems (see, e.g., [16], [17] among others).

In this work we consider a WPP participating in a dynamically evolving conventional two-settlement (day-ahead and real time) market, while exploiting co-located energy storage (i.e., a battery). We consider that forward power contracts are made some time slots ahead of delivery, depending on the forward market structure. Wind power and prices are modeled as random processes with general statistics. We study the maximum utility of energy storage in the WPP's operation involving delivery of day ahead forward power contracts while limiting costs due to real time market interactions.

Since the computation of a policy minimizing an infinite horizon expected discounted cost is in general intractable due to a policy living in a multi-dimensional continuous space, we first focus on small storage asymptotes. We show that, for a small storage, an optimal policy involves employing storage to reduce the real time market interaction via appropriate charging (when there is excess energy) and discharging (when there is a deficit), *without* changing the forward contracts from the optimal ones in the absence of energy storage. We also show that the discounted infinite horizon profit is concave and increasing. Thus, calculating the precise incremental benefit of a small storage provides an upper bound that linearly increases with the battery capacity on the utility of a battery.

Next, for general battery capacity, we propose a method to obtain a convex quadratic approximation of the value function, namely the expected infinite horizon discounted cost. Based on this approximation, we develop a stochastic model predictive control (MPC) policy. We numerically evaluate the value function associated with the small battery approximation policy and the MPC policy. We observe that the small battery approximation policy performs very well when battery capacity is small, while the stochastic MPC performs better for larger storage.

The remainder of the paper is organized as follows. In Section II, we describe our system model, our objective in the optimal policy computation and the underlying assumptions made in the analysis. The optimal policy in the small battery asymptotic regime is characterized in Section III. We then describe a stochastic MPC policy based on quadratic approximations of the value function in Section IV. Numerical results are presented in Section V. Conclusions are drawn in Section VI.

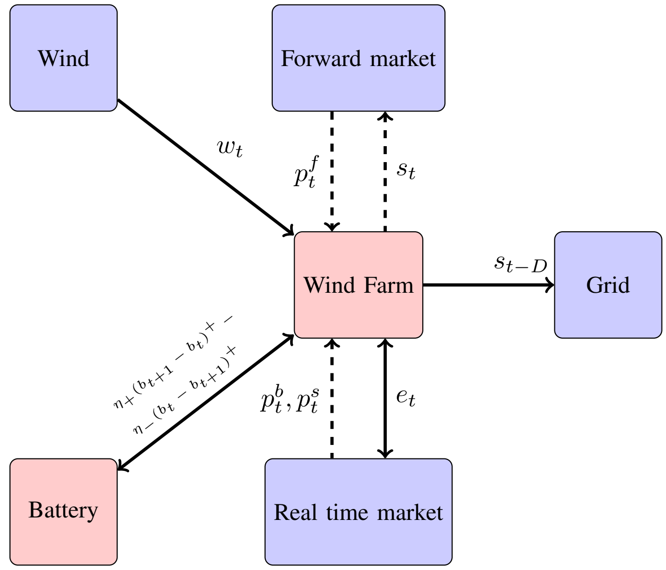


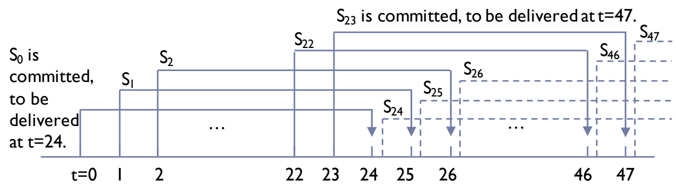
Fig. 1. System model of a wind farm participating in a two-settlement market. Red nodes represent the wind farm and the blue nodes represent external systems. Arrows with the labels represent flows between nodes. Solid arrows represent directions of energy flows (bidirectional arrows indicate energy may flow both ways). Dotted lines represent price and contract information flows.

II. Problem Description

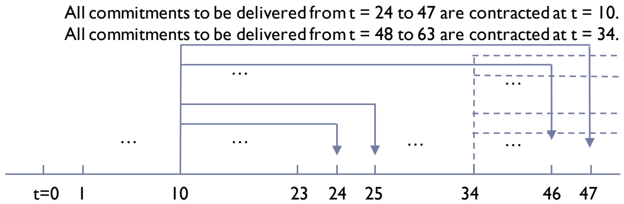
A. System model

We consider the infinite horizon problem of a WPP with co-located energy storage participating in a two settlement power market. The system diagram is depicted in Fig. 1. At each time instant t, the WPP performs the following two actions:

- **Forward market interaction**: The WPP promises to deliver a certain amount of electricity some number of hours, denoted by D_t, into the future. The WPP's revenue from this forward market interaction depends on the contract price and the amount of electricity promised. More generally, multiple forward contracts to be delivered at different future times can also be formed. Figures 2(a) and 2(b) depict two models of the forward market's timelines. In Figure 2(a), at every hour t, a forward contract in the amount of s_t is formed by a WPP, to be delivered $D_t = 24$ hours later. In Figure 2(b), at the 10^{th} hour of each day, a total of 24 forward contracts are formed to be delivered in each of the 24 hours in the next day, respectively.
  We employ the 1^{st} model, namely, $D_t = D$ for some fixed D, in deriving our main results. Generalizations of our results to other models is discussed in Section III-C. At each hour t, the contracted amount D hours earlier for the current hour t, denoted by s_{t-D}, needs to be delivered by the WPP. This necessarily entails the WPP's interactions with the real time market and its storage operation, as described next.
- **Real time market interaction and storage operation**: The WPP fulfils its earlier commitment made to be delivered at the current time. Due to the uncertainties in wind power generation, the WPP may either fall short of



(a)



(b)

Fig. 2. Examples of the timelines of forward market interactions: a) At every hour t, a forward contract in the amount of s_t is formed to be delivered $D = 24$ hours later, and b) In each day, at the 10^{th} hour, 24 forward contracts are formed to be delivered in each of the 24 hours of the next day.

its commitment, or it may have an excess left over after meeting all the forward contracts. In the former case, the WPP has to either buy power from the real time market or discharge power from its energy storage. In the latter case, the WPP sells power to the real time market or charges power to its energy storage.

The objective of the WPP is to minimize the infinite horizon expected discounted cost by choosing its actions appropriately. The notation associated with our model is summarized in Table I. Next, we present a more detailed description of the technical assumptions in our model in Section II-B.

B. Technical assumptions and model details

Our model makes the following assumptions:

- The price process (p_t^f, p_t^b, p_t^s) (where p_t^f is the forward contract price, p_t^b is the real time buying price, and p_t^s is the real time selling price) and wind process w_t are bounded and Lebesgue measurable. The wind farm is a price-taker, and actions of the wind farm do not impact the joint exogenous price and wind processes. The price and wind processes can be correlated.
- The stored energy in the battery is b_t. At each time instant, the WPP decides the level b_{t+1} to charge the battery to at the next time instant. Let the charging efficiency be denoted by η_+ and discharging efficiency by η_-. Thus, the energy consumed (and an negative amount means energy is extracted) for changing the battery level is given by $\eta_+(b_{t+1}-b_t)^+ - \eta_-(b_t - b_{t+1})^+$, where $(\cdot)^+$ denotes $\max(0,\cdot)$. Due to losses, $\eta_- \leq 1 \leq \eta_+$. Note that in batteries where the efficiency is 100% (lithium ion batteries are very close to ideal [18]), $\eta_+ = \eta_- = 1$, and the energy consumed/extracted for changing the battery level is $b_{t+1} - b_t$. At each time instant, we have the constraint that $b_{t+1} \in \mathcal{B}_t$ where $\mathcal{B}_t$ is a set dependent on current battery level b_t. This allows us to incorporate the following constraints in our model:
  - *Capacity constraints* The capacity of the battery is B. If we have an ideal battery, $\mathcal{B}_t = \mathcal{B} = [0, B]$. If we do not wish to completely charge or discharge the battery to prevent excessive wear, we can limit $\mathcal{B} = [\epsilon, B - \epsilon]$. This reduces the effective capacity to $B - 2\epsilon$.
  - *Ramping constraints* If the battery has power limits while charging or discharging, there is a ramping constraint $R \cdot B$ which is the maximum change in battery levels in a time instant. Ramping can be modeled as $\mathcal{B}_t = \mathcal{B} \cap [b_t - RB, b_t + RB]$.
- In addition to setting the battery level at time t, the WPP further makes a decision regarding sales and purchase of energy in the forward market. In particular, the WPP contracts s_t in the forward market to be delivered D time units later. The WPP receives contract price p_t^f for each unit contracted. The maximum contract is bounded by a sufficiently large constant S, i.e., $s_t \in [0, S]$.
- As the earlier contract s_{t-D} needs to be delivered from the realized wind and stored energy, the available energy e_t at time t, after delivering the contract and charging/discharging the battery to the appropriate level, is

$$e_t = w_t - \eta_+(b_{t+1} - b_t)^+ + \eta_-(b_t - b_{t+1})^+ - s_{t-D}. \quad (1)$$

- The WPP sells its excess $(e_t)^+$, if any, at the real-time market selling price p_t^s.
- The WPP meets its deficit, if any, by purchasing $(-e_t)^+$ units of energy at the real-time buying price p_t^b.

TABLE I
SUMMARY OF NOTATIONS USED

External processes	
p_t^f	Forward market price
p_t^b	Real time buying price
p_t^s	Real time selling price
w_t	Wind power
z_t	$(w_t, p_t^f, p_t^b, p_t^s)$
$\mathcal{F}_t$	All external realizations till time t: $\{z_s\}_{s\leq t}$
Battery parameters	
B	Battery capacity
b_t	Battery level at the beginning of time t
$\mathcal{B}_t$	Domain for battery level at time $t+1$
$\mathcal{B}$	If $\mathcal{B}_t$ is the same for all t we use this shorthand notation
R	Ramping constraint
η_+	Charging efficiency
η_-	Discharging efficiency
Other wind farm parameters	
D	The time after which the contract has to be delivered
s_t	The contract formed at time t, to be delivered at time $t+D$
s_{t-D}^{t-1}	Vector which contains all the power contracts to be offered from time t to $t+D-1$
e_t	The excess energy left over after meeting contract and charging/discharging battery
x_t	All prices (current and past time) and past actions; we call this the state of the system.
$\pi_{B,t}(\cdot)$	Mapping from state to action at time t when battery capacity is B. Action is in $\mathbf{R}^2$ and specifies the contract decision at time t and the battery level at $t+1$.
g	Stage cost (time dependence implicit in g)
g_f	Cost of interaction with the forward market
g_r	Cost of interaction with the real time market
$f(\cdot,\cdot,\cdot)$	Dynamics of the state evolution

The state of the WPP x_t in general consists of *the history of contracts to be delivered*, prices and wind energy realized, and the energy storage level. When the price and wind statistics are Markovian (to which our results are not restricted), the state only needs to include the current prices and wind as opposed to the history of prices and wind, i.e.,

$$x_t \triangleq (p_t^f, p_t^s, p_t^b, w_t, s_{t-D}^{t-1}, b_t, t).$$

The state space is $\mathcal{X}$. For brevity, we denote the price and wind processes by $z_t = (w_t, p_t^f, p_t^b, p_t^s)$ and the natural filtration [19] associated with the external stochastic process by $\{\mathcal{F}_t\}_{t\geq 0}$. Roughly speaking, the natural filtration satisfies the following properties:

- $\mathcal{F}_t \subseteq \mathcal{F}_{t+s}$ for $s \geq 0, t \geq 0$.
- Conditioning on $\mathcal{F}_t$ means that all realizations of the external stochastic process until time t, i.e., $\{z_s\}_{s\leq t}$, are known.

The action given a state x_t at time t consists of specifying the contract level s_t and the battery level b_{t+1} for time $t+1$. Thus, action $\pi_{B,t}(x_t) = (s_t, b_{t+1})$ and the space of all actions $\mathcal{A}$ is assumed to satisfy $\mathcal{A} = \{(s,b) | s \in [0,S], b \in [0,B]\}$.

We define the stage cost $g(x_t, \pi_B(x_t))$ as

$$g(x_t, \pi_B(x_t)) = g_f(p_t^f, s_t) + g_r(p_t^b, p_t^s, e_t) + \mathbf{I}_{\mathcal{B}_t}(b_{t+1}), \quad (2)$$

where g_f is the forward market cost component, g_r is the real-time market cost component and the third term limits the feasible range of b_{t+1}. $\mathbf{I}_{\mathcal{B}_t}(b_{t+1})$ denotes the standard indicator function which evaluates to zero if $b_{t+1} \in \mathcal{B}_t$ and is infinity otherwise. We consider

$$\begin{aligned} g_f(p_t^f, s_t) &= -p_t^f s_t, \\ g_r(p_t^b, p_t^s, e_t) &= p_t^b(-e_t)^+ - p_t^s(e_t)^+. \end{aligned} \quad (3)$$

Note that $g_r(\cdot)$ captures the risk (the greater the $-e_t$ is, the more the deficit and purchase from the real time market), whereas g_f captures the profits from the forward contract. Note that g, g_f and g_r are convex functions in their individual arguments. It can also be seen that the dynamics of the state is linear in the action and external processes, i.e. with a linear function f,

$$x_{t+1} = f(x_t, \pi_{B,t}(x_t), z_{t+1}). \quad (4)$$

We list some additional technical assumptions below.

- **Assumption 1**: The prices are such that $\beta^D p_t^s < \min(p_t^f) < \max(p_t^f) < \beta^D p_t^b$ with probability 1 for all $t \geq 0$. This is a sufficient condition to ensure that an agent with only a battery does not make a profit and also eliminates *infinite arbitrage* opportunities.
- **Assumption 2**: We consider contract policies in a (sufficiently large) neighborhood of the optimal "batteryless" contract policy denoted by $\mathcal{P}_L$. Specifically, we consider policies which are Lipschitz in B for all B in the following sense: there exists a finite (could be large) L so that for all considered policies $\pi_{B,t} = (s_{B,t}, b_{B,t})$ with battery of capacity B at time t,

  $$|s_{B,t}(x_t) - s_{0,t}^*(\tilde{x}_t)| \leq LB,$$

  where $s_{0,t}^*$ is the *optimal batteryless contract policy*, and $\tilde{x}_t$ equals x_t at all coordinates except for the battery capacity where it is zero. Note that $s_{0,t}^*$ can be computed based on just the external statistics by solving a news vendor problem [6]. From the constraint on the energy capacity, we also have $|b_{B,t}(x_t)| \leq B$.

Throughout the text, we use $o(B)$ to refer to any function $z(B)$ such that

$$\lim_{B\to 0} \frac{z(B)}{B} = 0,$$

and $\Theta(B)$ to refer to any function $z(B)$ such that

$$\lim_{B\to 0} \frac{z(B)}{B} = c$$

for some $c \neq 0$.

C. The objective

We seek to obtain a policy

$$\pi_{B,t}\text{: Domain}(x_t) \to \mathbb{R}^2$$

that maps the state x_t at each time instant to an action $\pi_{B,t}(x_t)$ that minimizes the infinite horizon stage cost. Note that if the wind and prices are Markov processes with memory M, the domain for the state variable is $(x_t) = \mathbb{R}^{D+5M+1}$. We consider a discount factor of $\beta < 1$. We formulate the following optimal stochastic control (or dynamic programming) problem:

$$\begin{aligned} V_t(x_t) \quad &\triangleq \quad \min_{\pi_{B,t}} \mathsf{E}\left[\sum_{s=t}^{\infty} \beta^{s-t} g(x_s, \pi_{B,s}(x_s)) \mid \mathcal{F}_t\right] \\ &\text{s.t. } x_{s+1} = f(x_s, \pi_{B,s}(x_s), z_{s+1}) \ \forall\ s \geq t. \end{aligned} \quad (5)$$

Before we describe our main results we introduce additional notation related to the objective function. Let

$$V_{\pi_{B,t},t}(x_t) \triangleq \mathsf{E}\left[\sum_{s=t}^{\infty} \beta^{s-t} g(x_s, \pi_{B,s}(x_s)) \mid \mathcal{F}_t\right]$$

be the value function obtained by following some fixed policy $\pi_{B,t}$ conditioned on the realization at time t, i.e., on $\mathcal{F}_t$. Let $V_{\pi_{B,t},t}^+$ be defined as

$$V_{\pi_{B,t},t}^+(s_{t-D}^{t-1}, b) \triangleq \mathsf{E}\left[V_{\pi_{B,t},t}(x_t) \mid \mathcal{F}_{t-1}\right], \quad (6)$$

where the expectation is taken *before the randomness at time t is realized*, and b is the initial battery level at time t.

III. Analysis of the optimal policy

In this section we derive some properties of the optimal policies.

A. Structural properties

The Bellman operator of the dynamic program formulated in (5) is

$$\mathcal{T}_t V_t(x_t) \triangleq \inf_{u_t \in \mathcal{A}} g(x_t, u_t) + \beta \mathsf{E}_{z_{t+1}}[V_{t+1}(f(x_t, u_t, z_{t+1})) \mid \mathcal{F}_t].$$

In the simple case where the external statistics are stationary or cyclo-stationary, the value function and optimal policy can be obtained by solving the fixed point equation, $V = \mathcal{T}V$. Since the stage cost is bounded and the Bellman operator is a contraction, the fixed point equation has a unique solution. We call a stationary optimal policy (not necessarily unique) corresponding to this optimal solution π_B^*. In case the statistics are not stationary, the optimal value function $V_t(x_t)$ and a corresponding policy $\pi_{B,t}^*$ are even harder to compute; however our asymptotic characterizations in the following sections hold even for such general statistics.

The evaluation of $\pi_{B,t}^*$ for a finite B is in general intractable both analytically and computationally. We can, however, consider asymptotic analysis to give us further insights. In the following, we list some observations about an infinite battery asymptote first, followed by the optimal small battery characterization in Section III-B.

We now describe an upper bound on the value function when the battery is large and the discount factor is close to 1, i.e., we would like to minimize the average cost instead of the discounted total cost. As the energy storage capacity $B \to \infty$, it is apparent that the optimal policy of the wind farm is to store all realized wind and contract what is stored when the price at the forward market is highest. i.e. $p_t^f = \max p^f$.

Lemma 1. *The average stage profit realized with infinite battery is upper bounded by,*

$$V_{avg}^{\infty} = \mathsf{E}[w] \times \max p_f.$$

where $\mathsf{E}[w] = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathsf{E}[w_t]$ *is the average wind energy realized over all time.*

We can realize this only if there are no ramping constraints or other inefficiencies. This is the optimal policy as any interaction with the real time market (shortfall or excess) is suboptimal due to the absence of (infinite) arbitrage opportunities.

B. Small battery asymptotic analysis

As it is difficult to compute the optimal policy $\pi_{B,t}^*$ in general, in this section we focus on the incremental value of storage when the battery capacity B is small. As we consider small battery capacities, ramping constraints of the battery can be neglected in the analysis. In essence, we perform a sensitivity analysis of $V_{\pi_{B,t}^*,0}^+$ (6) at $B = 0$. Assuming that both our initial contract history and the battery level is zero to start with, we set $s_{1-D}^0 = \mathbf{0} \in \mathbf{R}^D$, the initial battery $b = 0$ and seek to compute $\frac{\partial V_{\pi_{B,t}^*,0}^+(\mathbf{0},0)}{\partial B}$.

We now state our main theorem:

Theorem 1. *At time t, consider the following policy $\pi_{B,t}$:*

- *$s_{B,t}$ is set to be the optimal batteryless contract.*
- *The optimal storage operation policy is to charge fully or discharge fully depending only on the external wind and price processes. The exact specification is in Program 1.*

For a small enough B, the above policy achieves $V_{\pi_{B,t}^*,t}^+(s_{t-D}^{t-1}, 0)$ to within $o(B)$.

Program 1 Optimize the next stage battery level

- If $\eta_+ p_t^b \mathbb{I}(s_{t-D} > w_t) + \eta_+ p_t^s \mathbb{I}(s_{t-D} < w_t) < -\beta\alpha_t(s_{t+1-D}^t, 0)$, **set** $b_{t+1} = B$.
- If $\eta_- p_t^b \mathbb{I}(s_{t-D} > w_t) + \eta_- p_t^s \mathbb{I}(s_{t-D} < w_t) > -\beta\alpha_t(s_{t+1-D}^t, 0)$, **set** $b_{t+1} = 0$.
- Else: **set** $b_{t+1} = b_t$.

$\alpha_t(.) \triangleq \frac{\partial V_{\pi_{B,t}^*,t}^+(\cdot,b)}{\partial b}$ in Program 1 is the expected cost of storing an additional unit of charge in the battery at time t. It is non-positive, since one can always make an expected profit (and the cost is the negative of profit) out of storing an additional unit of charge in the battery. The indicator function $\mathbb{I}(\cdot)$ is 1 if the condition is satisfied and 0 otherwise.

As a typical case, if

$$p_t^b < -\alpha_t(\cdot) < p_t^s, \text{ and } \beta, \eta_+, \eta_- \approx 1, \tag{7}$$

Program 1 reduces to the specification that we should charge if we have a wind excess event ($s_{t-D} < w_t$), and discharge if we have a wind deficit event ($s_{t-D} > w_t$). This is because, as we show later, for statistics where the distribution of the external processes does not change significantly over time scales of length D units (in a sense made precise in Appendix D), α_t is close to the negative of the contract price D time units earlier, $-p_{t-D}^f$, and the above conditions (7) hold under Assumption 1. For general statistics and correlation structure $\alpha(\cdot)$ can be computed using a dynamic program, whose complexity depends on the mixing time of the external processes. It may be intractable in general. Note that $\alpha_t(\cdot)$ needs to be computed only within $\Theta(B)$, as discussed more after the description of an elaborate version of this program in Appendix B.

We observe, in particular, that the optimal policy in the low battery regime *does not involve changing the optimal forward contract from that with no storage*, but instead involves deciding when to charge or fully discharge the battery to reduce the risk of energy shortfall (measured by $g_r(\cdot)$). For specific cases where the external price and wind processes are cyclostationary, this corresponds to charging when there is excess energy, and discharging when there is a deficit.

We prove Theorem 1 in the appendices. We can thereafter compute $\frac{\partial V_{\pi_{B,t}^*,0}^+(\mathbf{0},0)}{\partial B}$ by policy evaluation. For a system with a finite number of states, policy evaluation or evaluating the value functions corresponding to particular states can be done by solving a linear system of equations. It may also be evaluated using Monte Carlo simulations of state trajectories. The latter may be the only tractable option for larger state spaces due to complexity. As shown later in Section V, for some special wind and price processes, the incremental value of a small amount of storage may be evaluated in closed form.

C. Extension to other forward market models

In this section we describe extensions of the small battery analysis to another day-ahead market model commonly used in practice. The main difference of this variant from the model analyzed earlier is as follows: a) Forward contract decisions are made at one specified time every D' time units, b) At each time of decision, D' forward contracts are determined for the consecutive D' time units starting from D time units later. An example with $D' = 24$ and $D = 14$ is the following: at 10am of each day, 24 forward contracts are made for the 24 hours starting from the beginning of the next day. In comparison, the model analyzed in earlier sections evenly spread the forward contract decisions over all hours. We observe that the same dynamic programming formulation in Section II applies in this market model, and the change to this market model affects the *state* of the DP. However, in spite of the differences from the model considered earlier, in the low battery limit, there are many similarities about the nature of the policies achieving the optimal profits in both market models.

In particular, as we show in Appendix F, Theorem 1 continues to hold under this market model. The reason is similar to that in the earlier market model: it can be shown that the batteryless optimal policy solves the news-vendor problem; by the first order necessary conditions for optimality, any deviation in the contract policy would not affect the profit up to first order terms. The charging policy is also exactly the same as in the earlier case, with a possibly different $\alpha_t(\cdot)$.

IV. Approximate dynamic programming approaches

The optimal policy described in Section III-B becomes suboptimal for batteries of large capacities. In this section, we propose an efficiently computable heuristic policy when relatively large batteries are used. We first compute a convex quadratic approximation of the value function by converting costs and constraints to quadratic functions. We then describe how model predictive control approaches can yield better results.

A. Quadratic approximation of value function

We seek to represent the value function by a quadratic approximation and use this to determine a policy to follow. The non-linearities in the control problem arise from the piecewise linear nature of the real time cost function and constraints (imposed by limited energy storage) on the charging levels and non-negative contracts in the forward market. As introduced in [20], we replace these with quadratic cost functions as follows:

$$g_f(p_t^f, s_t) = h_{f,1}\left(s_t + h_{f,2}\right)^2 \tag{8}$$

$$\begin{aligned} g_r(p_t^b, p_t^s, w_t + b_t - s_{t-D} - b_t^+) \quad (9) \\ = h_{r,1}\left(w_t + b_t - s_{t-D} - b_t^+ h_{r,2}\right)^2 \quad (10) \end{aligned}$$

$$\mathbf{I}_{[0,B]}(b_{t^+}) = \gamma\left(b_t^+ - \frac{B}{2}\right)^2, \tag{11}$$

where coefficients $h_{\cdot,\cdot}, \gamma$ are chosen as functions of expected external wind and price processes $\left(\mathsf{E}\left[p_t^f\right], \mathsf{E}\left[p_t^s\right], \mathsf{E}\left[p_t^b\right], \mathsf{E}\left[w_t\right], B\right)$ to approximate the piecewise linear costs and the energy storage constraint. This converts the problem to a linear-quadratic (LQ) control problem. The state can be written as $x_t = [w_t, b_t, s_{t-D}^{t-1}]$ and the action $u_t = [b_{t+1}, s_t]$. The state vector does not include the price, introducing another source of suboptimality. This is because policy recommendations from such an approach are the same irrespective of instantaneous prices. The stage cost and dynamics can now be written as

$$\begin{aligned} g_t^{\text{quad}}(x_t, u_t) &= \begin{bmatrix} x_t \\ u_t \\ 1 \end{bmatrix}^\mathsf{T} \begin{bmatrix} Q_t & q_t \\ q_t^\mathsf{T} & r_t \end{bmatrix} \begin{bmatrix} x_t \\ u_t \\ 1 \end{bmatrix} \\ x_{t+1} &= A_1 x_t + A_2 u_t + A_3 w_t, \end{aligned}$$

where $Q., q., A.$ are constants that can be straightforwardly computed. This problem has a convex quadratic stage cost and linear dynamics. The convex quadratic value function $V_{\text{quad}}(x_t)$ can be found by solving the Algebraic Ricatti Equations (ARE) resulting from the Bellman optimality equation. V_{quad} is a quadratic approximation of our original problem. This approach also delivers an *affine policy* as

$$u_t(x_t) = K_t x_t + k_t$$

for K_t, k_t determined through the ARE.

The quality of the approximation V_{quad} is much higher when the price process does not vary much. Another approach to obtaining a quadratic approximator of the value function which is not problem specific is presented in [21]. The authors use the $\mathcal{S}-$procedure to find a convex quadratic under-approximation to the value function using iterated Bellman equations.

B. Model predictive control

With any approximation to the value function, denoted by V_{app}, we can employ a model predictive control (MPC) approach to determine a policy by applying V_{app} as the *terminal cost*. In the LQ approximation to the problem described in Section IV-A, the variation in price is not taken into account while determining the policy. The proposed MPC approach refines the usage of the approximation by using the current price information. First, the expected value of the wind and price processes can be used for unseen future realizations in a certainty-equivalent MPC approach.

The M step certainty-equivalent MPC algorithm utilizes an $V_{\text{app}}(x_t)$ to obtain a policy as

$$\begin{aligned} u_t = \underset{s_t, b_{t+1}}{\text{argmin}} \min_{s_{t+1}^{t+M}, b_{t+2}^{t+M+1}} \sum_{j=0}^{M-1} \beta^j g(s_{t+j-D+1}^{t+j+1}, b_{t+j}^{t+j+1}, z'_{t+j}) \\ + \beta^M V_{\text{app}}(s_{t+M-D+1}^{t+M}, b_{t+M}, \mathsf{E}\left[z_{t+M}\right]). \end{aligned} \tag{12}$$

As defined earlier, $z_t = (p_t^f, p_t^s, p_t^b, w_t)$ represents the external stochastic processes. At each stage t, we see $z'_t = z_t$ and use $z'_{t+j} = \mathsf{E}\left[z_{t+j}\right] \ \forall j \in [M]$.

Furthermore, as opposed to the certainty-equivalent MPC approach, we also use a stochastic version of the MPC algorithm which is more robust as it samples the external wind and price processes from its probability distribution rather than simply using its expectation. We generate multiple such samples of length M for future realizations in a Monte Carlo approach. We choose the day ahead contract and the charging policy for the current time instant that minimizes the average costs across all such M-length samples and the appropriately discounted terminal function to approximate expected future cost. A policy minimizing this cost is selected. Further details explicitly characterizing the objective function to be minimized are presented in Appendix F. The MPC policy can be used in cases where the probability distribution for future realizations depends on past realizations of wind and prices as well. When V_{app} is a quadratic (e.g. the V_{quad} developed in Section IV-A), solving both certainty-equivalent and stochastic MPC is minimizing a quadratic function subject to linear constraints which can be done efficiently. As M gets larger, the terminal cost function V_{app} has a reduced impact because of the discount factor. Hence, for sufficiently large M, finding the MPC solution can be approximated by a linear program.

V. Numerical Results

In this section, we a) illustrate the optimal policy in the asymptotic regimes of energy storage, and b) extensively evaluate the asymptotically optimal and the proposed heuristic policies for all regimes of energy storage capacity.

A. Simulation setup

We first perform simulations employing a synthetic model for the external wind and price processes which are fit from PJM interconnection data from the year 2004-2005 [22]. We consider the case of a WPP in hourly day-ahead and hourly real-time markets for the case where horizon $D = 24$. For comparison, we also evaluate approximate $D = 4$ models with prices and wind averaged over 6 hour blocks. We fit Gaussian models to the prices and uniform distributions to the wind. The realizations of the wind and prices are independent across time and each other in this simulations but are cyclo-stationary. Furthermore, real world wind and price traces will be used for simulations toward the end of the section.

The code and data files for the simulation are available at [23]. We evaluate the averaged discounted profit over 16 realizations for two policies, and also evaluate upper bounds:

1) **Small battery approximate policy**: This policy sets contracts to the optimal batteryless contract s_t^*, and charges or discharges its battery if it sees an excess or deficit based on the realized wind and the contract it has to meet.
2) **Stochastic MPC**: We use the heuristic policy described in Section IV with 40 Monte Carlo samples for price and wind values with lookahead of $M = 40$ for $D = 4$, and 40 samples with lookahead $M = 48$ for $D = 24$. The terminal value function approximation used is computed from the LQ method described in Section IV-A.
3) **Clairvoyant bound**: Assuming complete knowledge of the *future* wind and price processes, an upper bound on the discounted profit can be computed using a linear program.
4) **Linear upper bound**: The precise incremental value of storage was computed in the small battery regime as shown in Section III-B. Now as the value function is convex in B, the profit is concave, and the computed derivative at 0 provides an upper bound as follows

$$\begin{aligned} & V_B^{\text{Linear}}(s^*_{-D+1}, b = 0) \\ & = V_{\pi^*_{0,t},t}(s^*_{-D+1}, 0) + \frac{\partial V_{\pi^*_{B,t},t}(s^*_{-D+1}, 0)}{\partial B}|_{B=0} B \\ & \geq V_{\pi^*_{B,t},t}(s^*_{-D+1}, 0). \end{aligned}$$

5) **Upper bound**: The upper bound we present in the figures is the minimum of the above two upper bounds.

B. Special case - Constant prices and stationary wind distribution

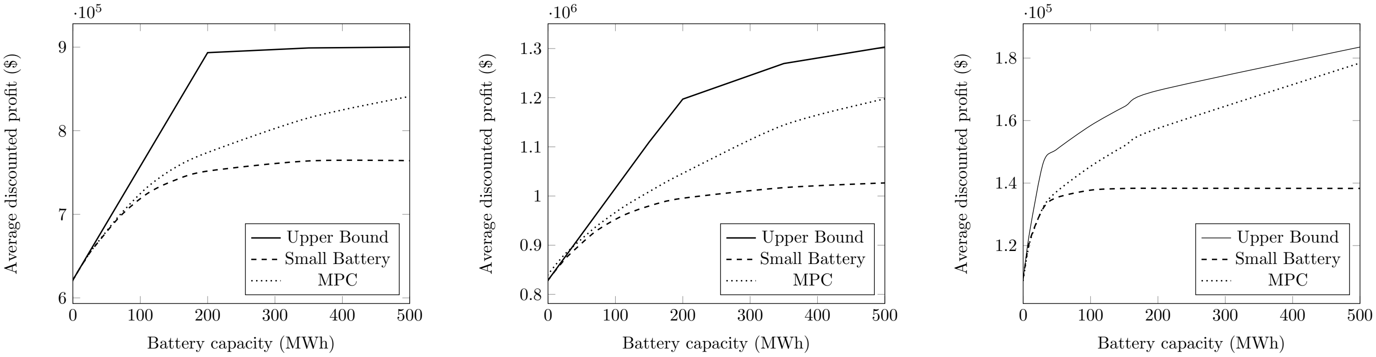
We start with the price process being constant and the wind distribution drawn from a stationary uniform distribution. The battery is assumed to be ideal. While this scenario is not realistic in practice, it allows us to compute the utility of the battery solely due to the variation in wind. In this scenario, we can evaluate the incremental value of energy storage in *closed-form*. For any finite horizon D, the optimal batteryless contract can be found by solving the following news-vendor problem,

$$\begin{aligned} s^* &= \min_s -p^f + \beta^D \mathsf{E}\left[p^b(s - w_{t+D})^+ - p^s(w_{t+D} - s)^+\right] \\ \Rightarrow s^* &= F_w^{-1}\left(\frac{p^f - \beta^D p^s}{\beta^D(p^b - p^s)}\right), \end{aligned} \tag{13}$$

where F_w is the cdf of the wind distribution. For a small battery and an optimal batteryless contract, the following holds:

$$\begin{aligned} \Delta V_B^+(B) &= \Pr(w < s^*)\left(-p^b B + \beta \Delta V_B^+(0)\right) && (14) \\ & \quad + \Pr(w > s^*)\left(\beta \Delta V_B^+(B)\right) \\ \Delta V_B^+(0) &= \Pr(w < s^*)\left(\beta \Delta V_B^+(0)\right) && (15) \\ & \quad + \Pr(w > s^*)\left(p^s B + \beta \Delta V_B^+(B)\right). \end{aligned}$$

The notations are explained as follows. We let $\Delta V_B^+(s^t_{t-D+1}, b)$ be the incremental cost of storage (nonpositive), i.e., the *difference* of a) the expected batteryless value function, from b) the expected value function at (s^t_{t-D+1}, b) from following the optimal policy under the small battery approximation. Notationally, if the first argument is omitted, the prior D contracted amounts are the optimal batteryless amounts s^*. In (14), the incremental value function with full battery is the sum of the cost when there is a deficit $(-p^b B + \beta \Delta V_B^+(0))$ and when there is an excess $(\beta \Delta V_B^+(B))$, weighted by the probability of deficit or excess while following the small battery optimal policy. Equation (15) can be similarly interpreted. The above simplification of stagecost is valid to $o(B)$ as described earlier. We can evaluate $\Pr(w > s^*)$ from the solution to the news-vendor problem described in Equation (13). If we follow the batteryless optimal contract and do not charge or discharge the battery



(a) (b) (c)

Fig. 3. Average discounted profit as a function of battery capacity for (a) constant price and stationary wind processes and $D = 4$, and (b) periodic and random price and wind processes with $D = 4$ modeled over 1 year and (c) periodic and random price and wind processes with $D = 24$ and 2 months modeled.

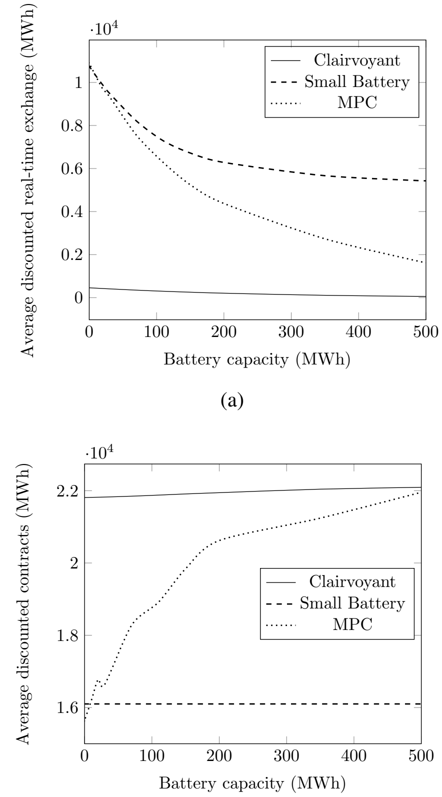
until the first time we have a non-zero contract to deliver ($s_{t-D} > 0$), the incremental value of battery evaluates to

$$\begin{aligned} \Delta V_B^+(s_{1-D}^* = 0, b = 0) = \quad & \Pr(w > s^*)\Delta V^+(B) \\ & +\Pr(w < s^*)\Delta V^+(0) \\ = \quad & -B\frac{(p^f - \beta^D p^s)(\beta^D p^b - p^f)}{\beta^D(p^b - p^s)(1-\beta)} \end{aligned} \tag{16}$$

This is confirmed in Fig. 3(a) where the incremental value of the battery using the small battery policy closely matches the asymptote obtained above (16). The approximation degrades when $B = 25MWh$. Beyond a capacity of $10MWh$, the stochastic MPC heuristic outperforms the small battery policy (albeit both are close to each other). To get a sense of the size of storage, the average wind energy generated at every time unit (over 6 hours) in this scenario is $200MWh$. In scenarios where the price and wind statistics are periodic, the asymptotes are found numerically as is done in Fig. 3(b) and Fig. 3(c).

C. Performance with general battery capacities

Fig. 3(a), 3(b), and 3(c) show the performance of the two policies for a wide range of energy storage capacities, in the stationary wind and constant price scenario and periodic external process scenarios with $D = 4$ and $D = 24$. The MPC heuristic is seen to perform quite well as it a) matches the asymptote (and the optimal small battery policy) at low battery levels, and b) approaches the clairvoyant bound at high energy levels. The small battery policy, which only employs the battery to minimize real-time interactions without changing the forward contracts from the optimal batteryless ones, is clearly suboptimal for values of energy storage beyond $100MWh$. In other words, when the battery capacity is higher than this, it is optimal to contract more than the optimal batteryless contract depending on the stored energy levels. In the extreme case with very high battery capacity, it is optimal to store energy and contract only when price is high. In Fig. 3(a), (finite) arbitrage across time cannot be done as prices are constant. Increasing discounted profits are seen in Fig. 3(b) and 3(c) as increased battery levels enable storing energy across periods until the forward market contract price is higher.



(a) (b)

Fig. 4. Periodically varying wind and fixed price with $D = 4$ and simulated for 1 year: (a) Average discounted magnitude of the real-time component of the stage cost (b) Average discounted contracts offered on the forward market.

In Fig. 4(a), the magnitude of real-time market interactions are plotted for the two policies and in the clairvoyant bound. It can be seen that the interactions, which are a proxy for risk since real-time market interactions are costly, are reduced more significantly for the MPC. Fig. 4(b) shows the average discounted contracts (sum of contracts with a discounting factor) made. The small battery policy does not contract more as capacity increases by design, and this is a source of its suboptimality. The MPC policy has its discounted contracts

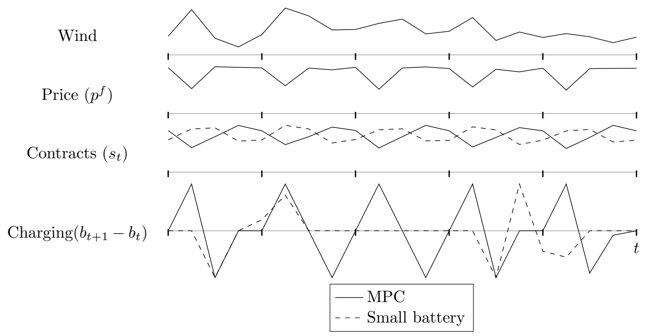


Fig. 5. Time trace for 21 time units with $D = 4$ with periodic wind and price statistics and capacity $30MWh$. All plots have been normalized.

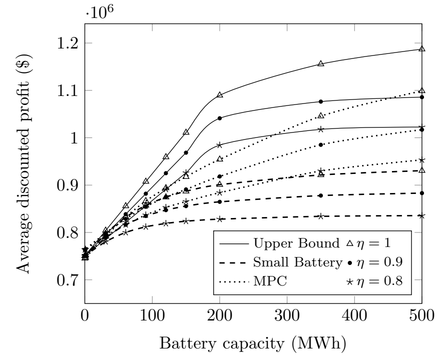
converging to that of the optimal clairvoyant policy as B increases; its increased contracts and reduced real-time market interactions contribute to good performance. In this scenario, energy storage allows for increased contracts, and at the same time reduced real time market interaction as it helps mitigate the randomness of wind energy generation. In the general case with varying prices, arbitrage brings additional benefit. It is interesting to note that MPC and the low complexity small battery heuristic have similar performance when $B < 20MWh$, although the policies have quite different computational requirements.

In Fig. 5, we can see the behavior of the MPC and small battery policies for a specific case with $D = 4$ and periodic wind and price statistics. The small battery policy contracts an amount based on the current forward contract price, statistics of the wind and price for each time instant; this can result in a higher contract for periods in which the contract price realization is low but the expected wind generation is high. The MPC policy actively takes into account arbitrage opportunities afforded by periodic price and wind variations and contracts more when the price is higher by charging and discharging the battery appropriately. This shows that the small battery policy brings gain through minimizing wind variations whereas the MPC policy can also arbitrage with prices.

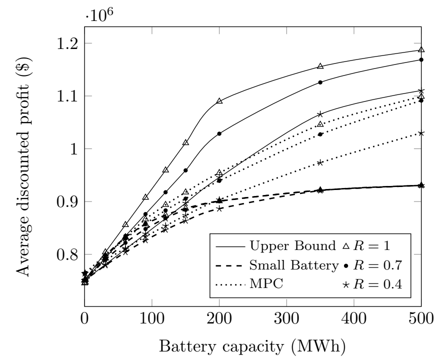
D. Impact of battery inefficiencies

Fig. 6(a) highlights the impact of battery charging and discharging inefficiencies in the periodically varying wind and price scenario for $D = 4$. Performance of both policies and the clairvoyant upper bound flatten out at degraded profit levels. Efficiency of the co-located energy storage is crucial in making a profit and justifying the expense of the battery.

Fig. 6(b) illustrates the impact of ramping factor R. A ramping factor $R > 0.7$ is not seen to degrade performance at higher battery levels, implying that large shifts in battery values are not common in the optimal policy for typical scenarios. The greatest impact of the ramping factor for the small battery policy is at lower battery levels while it affects MPC policy for all battery levels.



(a)



(b)

Fig. 6. Periodically varying wind and price statistics with $D = 4$ and simulated for 2 moths (a) Effect of battery inefficiency (b) Effect of the ramping constraint.

E. Performance on Real Data

Performance on real data with representative forecasting errors is shown in Fig. 7. In this simulation, PJM interconnection data for the years 2004 and 2005 were used to generate the wind and prices. The algorithms were run over 2 months which allowed for 12 realizations. Representative forecasting error distributions were assumed with the margin of error $(1\% - 10\%)$ increasing as the prediction horizon increased. In solving the policy decisions A uniform model was assumed for wind and Gaussian models for the prices. The policies are then evaluated using the real world wind and price traces.

As can be seen from the plot, the small battery policy performs quite well compared to the MPC policy for small storage capacities. The clairvoyant bound is also higher as real data can offer arbitrage conditions; this makes it profitable for a battery owner who does not generate wind power to operate on the grid. The performance of the small battery policy is not monotonically increasing as Assumption 1 on prices does not hold for real data. There are wide fluctuations in the prices and the forward contract price now can be larger than the buying price or lower than the real time selling price at a later time instant. This implies that storing energy in the battery may be sometimes not as profitable as selling it on the real time market.

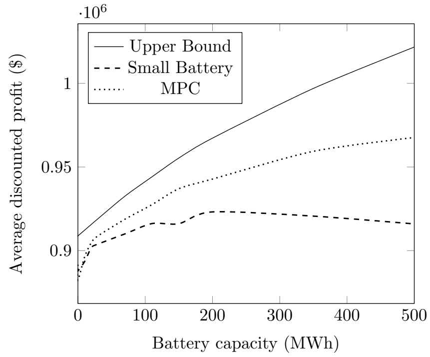


Fig. 7. Performance with $D = 4$ for 12 realizations of real data with representative errors over 2 months.

VI. Conclusions

We have studied the problem of a WPP participating in a dynamically evolving two settlement power market, with the help of a co-located energy storage. To maximize the expected profit of a WPP, the optimal forward contract and the storage operation (i.e. charging/discharging) decisions are formulated as an infinite horizon stochastic optimal control problem. We characterize analytically the optimal operations for the asymptotic regimes of small storage and show that the optimal operation for small storage involves only a reduction in the real time market interaction, without changing the forward contracts from the optimal ones in the absence of storage. For an intermediate storage capacity, we propose a stochastic model predictive control (MPC) policy based on a quadratic approximation of the value function. We numerically evaluate the simple asymptotically optimal policy and the MPC policy. We observe that, as expected, while the simple policy associated with the small battery approximation works well for small batteries, the more complex MPC works better for larger ones. The precise threshold on battery capacity where the different policies are best depends on the exact parameters of the system.

Acknowledgment

We thank Dr. Mahnoosh Alizadeh and the anonymous reviewers for helpful comments and discussions.

Appendix A

In this appendix, we state and prove Lemma 2 about some properties of $V^+_{\pi^*_B,t}$.

Lemma 2. *The following hold:*

- $V^+_{\pi^*_{B,t},t}$ *is differentiable in* B.
- $V^+_{\pi^*_{B,t},t}(s^{t-1}_{t-D}, b)$ *is differentiable in* b.

Proof. The finite first moment assumption on the wind together with the bounded price processes implies that the expected stage cost at each time is bounded.

- The batteryless optimal contract is bounded above by an absolute constant if the expected absolute first moment of the wind process is finite [6]. Also, if the discount factor $\beta < 1$, the series is absolutely summable. This implies that the expected stage cost is finite and the dominated convergence theorem can be used to interchange the expectation and the summation. We note that the incremental change in the value function due to the battery cannot be infinite due to the fact that the prices are bounded. This means that the value function is Lipschitz and absolutely continuous in B at $B = 0$. This implies differentiability of the value function almost everywhere with respect to the Lebesgue measure, in any interval that B may lie in, in particular, a neighborhood around 0.
- From the boundedness of prices it follows that $V^+_{\pi^*_{B,t},t}$ is Lipschitz in b. Thus $V^+_{\pi^*_{B,t},t}$ is differentiable in b (similar argument as in the last paragraph).

□

Appendix B

In this appendix, we prove Theorem 1. We first show that an optimal storage operation policy is dependent only on the external processes (in particular, it is independent of the contract policy). For any fixed storage operation policy dependent only on external wind and price statistics we show that an optimal contract policy is the batteryless optimal policy.

1) Optimal storage operation policy: From the differentiability of $V^+_{\pi^*_{B,t},t}$ in b established in Appendix A, we have the following expression:

$$V^+_{\pi^*_{B,t},t}(s^{t-1}_{t-D}, b) = V^+_{\pi^*_{B,t},t}(s^{t-1}_{t-D}, 0) + \alpha_t(s^{t-1}_{t-D})b + o(B).$$

$\alpha_t(\cdot)b$ in the above expression is the incremental cost (or the negative of the incremental value) of having an initial battery level of b. Note that $\alpha_t < 0$ can be computed based on the history of all realizations till (and including) time $t-1$ and the future statistics of all external wind and price processes. Note that, as shown in Appendix C, the $o(B)$ term follows from the Lipschitz continuity of α_t in its contract arguments. Hence even if each entry of s^{t-1}_{t-D} is specified within $\Theta(B)$ of the batteryless optimal contract, the above relation still holds. We now characterize an optimal battery charging policy within $\mathcal{P}_L$ (defined in Assumption 2). Note that by an optimal policy we mean any policy $\pi_{B,t}$ such that the corresponding $V^+_{\pi_{B,t},t}$ is different from the $V^+_{\pi^*_{B,t},t}$ by no more than $o(B)$. We have the following inequality.

$$\begin{aligned} V_{\pi^*_{B,t},t}(x_t) &\le h_t(s_t, b_{t+1}, x_t) \\ &= -p^f_t(s_t) + p^b_t(s_{t-D} - w_t + \\ &\quad \eta_+(b_{t+1} - b_t)^+ - \eta_-(b_t - b_{t+1})^+)^+ \\ &\quad - p^s_t(-s_{t-D} + w_t \\ &\quad - \eta_+(b_{t+1} - b_t)^+ + \eta_-(b_t - b_{t+1})^+)^+ \\ &\quad + \beta V^+_{\pi^*_{B,t},t}(s^t_{t+1-D}, b_{t+1}). \end{aligned}$$

Note that $h_t(\cdot,\cdot,\cdot)$ differs from $V_{\pi_{B,t},t}(x_t)$ in that the former corresponds to the value function for the optimal policy from all time instants $t+1$ onwards (with the only potentially suboptimal action being at time t), whereas in the latter, we follow the suboptimal policy $\pi_{B,t}$ for all time. Note also that x_{t+1} is specified by the choices of the contract and battery pair

s_t, b_{t+1} and by the external random processes. The (s_t, b_{t+1}) can be chosen either according to the policy $\pi^*_{B,t}$ (to get an equality) or can be chosen suboptimally, thereby increasing the one-stage cost.

We now look into the optimality conditions that the best one-step b_{t+1} needs to satisfy. We first note that $-b_t + b_{t+1}$ is bounded in absolute value by B. For a given s_{t-D}, b_t, w_t we then have the following:

$$\begin{aligned}&\frac{\partial h(s_t, b_{t+1}, x_t)}{\partial b_{t+1}}|_{b_{t+1}\to b_t+}\\&= \eta_+ p^b_t \mathbb{I}(s_{t-D} > w_t)\\&\quad +\eta_+ p^s_t \mathbb{I}(s_{t-D} < w_t) + \beta\alpha_t(s^t_{t+1-D}),\ \text{if } |s_{t-D} - w_t| > 2B,\end{aligned}$$

and

$$\begin{aligned}&\frac{\partial h(s_t, b_{t+1}, x_t)}{\partial b_{t+1}}|_{b_{t+1}\to b_t-}\\&= \eta_- p^b_t \mathbb{I}(s_{t-D} > w_t)\\&\quad +\eta_- p^s_t \mathbb{I}(s_{t-D} < w_t) + \beta\alpha_t(s^t_{t+1-D}),\ \text{if } |s_{t-D} - w_t| > 2B.\end{aligned}$$

The notation $\frac{\partial f}{\partial x}|_{x\to a+}$ refers to the right derivative of f with respect to x at $x = a$ and $\frac{\partial f}{\partial x}|_{x\to a-}$ refers to the corresponding left derivative. Note that we are interested in the value of $\mathsf{E}[h]$, and specifying the storage operation policy when w_t lies in $[s_{t-D} - 2B, s_{t-D} + 2B]$ is sufficient for us to know $\mathsf{E}[h]$ within $o(B)$. The following lemma makes this precise:

Lemma 3. *Policies differing only in the specification for $b_{t+1} \in [0, B]$ for the wind realization w_t lying in a set of measure $\Theta(B)$ yield $\mathsf{E}[h(s_t, b_{t+1}, x_t)]$ differing by at most $o(B)$.*

This follows from the fact that the difference in $\mathsf{E}[h(s_t, b_{t+1}, x_{t+1})]$ for two policies in $\mathcal{P}_L$ differing only in the behavior when $|s_{t-D} - w_t| \le 2B$ is of the form

$$\Pr(|w_t - s_{t-D}| \le B) \times \Theta(B) = o(B).$$

An implication of the above lemma is that the choice of storage operation policy in $|s_{t-D} - w_t| \le 2B$ does not matter and an optimal next stage storage operation policy is dependent only on external wind and price processes. An optimal next stage storage operation policy can be written as in Program 1, reproduced in Program 2.

$\alpha_t(.)\left(\triangleq \frac{\partial V^+_{\pi^*_{B,t},t}(\cdot,b)}{\partial b}\right)$ in Program 2 is the expected cost of storing an additional unit of charge in the battery and is non-positive, since one can always make an expected profit out of storing an additional unit of charge in the battery. Note that $\alpha_t(s^t_{t+1-D})$ is Lipschitz in the contract policy, i.e., it is known to be within $\Theta(B)$ (discussed in Appendix C) among all policies in $\mathcal{P}_L$. Strictly speaking one should have $\Theta(B)$ instead of 0 on the right hand side in Program 2 or Program 1, as done in Program 3. This is because the set of w_t for which $\frac{\partial h}{\partial b}$ is $\Theta(B)$ is also of measure $\Theta(B)$. By Lemma 3, all policies satisfying Program 3 yield $\mathsf{E}[h]$ that are the same up to first order terms in B.

Note that, if we start from a battery level $b_0 = 0$, the battery level is always either 0 or B. Note also that terms in s^t_{t+1-D} are known to within $\Theta(B)$. This is true for time $t < D$ because initial contracts are known, and for time $t \ge D$ by Assumption 1 by which we have that the contract policy does not deviate from the batteryless optimal contract policy by more than $\Theta(B)$. Thus $\alpha_t(\cdot)$ can be computed by knowing *just the external statistics*. In some special cases, this can even be obtained in closed form (Appendix E).

We now characterize the optimal contract policy by fixing a (potentially time dependent) storage operation policy dependent only on the external statistics of wind and prices.

2) Optimal contract policy: We start with decomposing the expression $V_{\pi^*_{B,t}}(x_t)$ as the summation of two terms: one being the expected real time cost due to the initial contract and the second being the expected revenue from contracts from the time t onwards. We observe that $V_{\pi^*_{B,t}}$ can be written as

$$\begin{aligned}V_{\pi^*_{B,t}}(x_t) = \mathsf{E}\Big[&\sum_{v=t}^{t+D-1}\beta^{v-t} g_r(p^b_v, p^s_v, e_v) \mid \mathcal{F}_t\Big]\\&+\sum_{v=t}^{\infty}\beta^{v-t}\mathsf{E}\Big[-p^f_v s_v\\&+\beta^D \mathsf{E}\Big[g_r(p^b_{v+D}, p^s_{v+D}, e_{v+D}) \mid \mathcal{F}_v\Big]\Big]\end{aligned} \tag{17}$$

We focus on finding the optimal contract policy s_t. We observe that the inner expectation conditioned on $\mathcal{F}_v$ can be computed to $o(B)$ accuracy by using the storage operation policy which is just a function of the random variables defined in $\mathcal{F}_v$ (this is proved in Appendix D). In addition, for any L_1, L_2 such that $|L_1|, |L_2| < L$, we can consider the batteryless optimal policy at time t, s^*_t and have that for $\tilde{s}_t = s^*_t + L_1 B, \bar{s}_t = s^*_t + L_2 B$,

$$\begin{aligned}&\Big|\Big(-p^f_t\tilde{s}_t + \beta^D\mathsf{E}\Big[g_r(p^b_{t+D}, p^s_{t+D}, \tilde{e}_{t+D}) \mid \mathcal{F}_t\Big]\Big)\\&\quad -\Big(-p^f_t\bar{s}_t + \beta^D\mathsf{E}\Big[g_r(p^b_{t+D}, p^s_{t+D}, \bar{e}_{t+D}) \mid \mathcal{F}_t\Big]\Big)\Big|\\&\quad = o(B).\end{aligned}$$

Thus, given a fixed storage operation policy dependent *only* on the external statistics of the wind and the price processes, we observe that the contract choice does not affect the expected cost by more than $o(B)$ as B approaches 0. Intuitively, the increased profits from the contracts are canceled by the losses due to the increased expected real time market interaction if the contract is already at the one specified by the batteryless optimal policy. Hence an *optimal contract policy* in the asymptotically small battery regime is just the batteryless optimal contract policy and an *optimal storage operation policy* is to follow Program 1. Thus Theorem 1 is proved.

Appendix C

In this appendix, we prove that α_t is also locally Lipschitz around the contracts.

We note that the expected increase in the real time interaction due to a change in the contract by $\Theta(B)$ is also $\Theta(B)$ due to the fact that $\mathsf{E}_{w_{t+1}}[p^b_t(w_{t+1} - s_{t+1-D})^- - p^s_t(w_{t+1} - s_{t+1-D})^+]$ is Lipschitz in s_{t+1-D}. Since $b < B$, we have $b\Theta(B) = o(B)$, and

$$\alpha_t(s^{t-1}_{t+1-D} + \Theta(B))b = \alpha_t(s^{t-1}_{t+1-D})b + o(B)$$

Program 2 Optimize the next stage battery level

- If $\frac{\partial h(s_t,b_{t+1},x_t)}{\partial b_{t+1}}|_{b_{t+1}\to b_t+} = \eta_+ p_t^b \mathbb{I}(s_{t-D} > w_t) + \eta_+ p_t^s \mathbb{I}(s_{t-D} < w_t) + \beta\alpha_t(s_{t+1-D}^t) < 0$, **set** $b_{t+1} = B$.
- If $\frac{\partial h(s_t,b_{t+1},x_t)}{\partial b_{t+1}}|_{b_{t+1}\to b_t-} = \eta_- p_t^b \mathbb{I}(s_{t-D} > w_t) + \eta_- p_t^s \mathbb{I}(s_{t-D} < w_t) + \beta\alpha_t(s_{t+1-D}^t) > 0$, **set** $b_{t+1} = 0$.
- Else: **set** $b_{t+1} = b_t$.

Program 3 Optimize the next stage battery level (similar to Program 2)

- If $\frac{\partial h(s_t,b_{t+1},x_t)}{\partial b_{t+1}}|_{b_{t+1}\to b_t+} = \eta_+ p_t^b \mathbb{I}(s_{t-D} > w_t) + \eta_+ p_t^s \mathbb{I}(s_{t-D} < w_t) + \beta\alpha_t(s_{t+1-D}^t) \leq \mathbf{\Theta}(\mathbf{B})$, **set** $b_{t+1} = B$.
- If $\frac{\partial h(s_t,b_{t+1},x_t)}{\partial b_{t+1}}|_{b_{t+1}\to b_t-} = \eta_- p_t^b \mathbb{I}(s_{t-D} > w_t) + \eta_- p_t^s \mathbb{I}(s_{t-D} < w_t) + \beta\alpha_t(s_{t+1-D}^t) \geq \mathbf{\Theta}(\mathbf{B})$, **set** $b_{t+1} = 0$.
- Else: **set** $b_{t+1} = b_t$.

where $s_{t-D}^{t-1} + \Theta(B)$ is taken to refer to any contract history which is within a constant times B of each contract in s_{t-D}^{t-1}, i.e.,

$$s_{t-D}^{t-1} + \Theta(B) \triangleq \{z_{t-D}^{t-1} : |z_i - s_i| < \Theta(B) \text{ for } i \in \{t-D,\dots,t-1\}\}.$$

Appendix D

In this appendix we show that, given a storage operation policy, the expectation at a future time can be computed to within $o(B)$ and hence can be used to compute the optimal contract policy. The batteryless optimal s^* is such that

$$\begin{aligned}&-p_t^f s^* \\ &+\beta^D \mathsf{E}\Big[g_r(p_{t+D}^b, p_{t+D}^s, -s^* + w_{t+D} \\ &-\eta_+(b_{t+D+1}-b_{t+D})^+ + \eta_-(-b_{t+D+1}+b_{t+D})^+ \mid \mathcal{F}_t\Big] \\ &= 0\end{aligned}$$

is minimized at s^*. A first order necessary condition for this to hold is

$$\begin{aligned}&\frac{\partial\left(-p_t^f s^* + \beta^D \mathsf{E}\Big[g_r(p_{t+D}^b, p_{t+D}^s, e_{t+D}) \mid \mathcal{F}_t\Big]\right)}{\partial s^*} \\ &\quad = 0 \\ &\equiv -p_t^f + \beta^D \mathsf{E}\left[p_{t+D}^b(s^* - w_{t+D})\mathbb{I}(w_{t+D} < s^*) \mid \mathcal{F}_t\right] \\ &\quad + \beta^D \mathsf{E}\left[p_{t+D}^s(-s^* + w_{t+D})\mathbb{I}(w_{t+D} > s^*) \mid \mathcal{F}_t\right] = 0.\end{aligned}$$

With a fixed storage operation policy dependent only on the external wind and price statistics, and here we choose $s^*+\lambda B$ WLOG (as will be shown by the end of this section), the change in the cost is

$$\begin{aligned}&-p_t^f(\lambda B) \\ &+\beta^D \mathsf{E}\Big[g_r(p_{t+D}^b, p_{t+D}^s, w_{t+D}-s^*-\lambda B-\eta_+(b_{t+D+1}-b_{t+D})^+ \\ &-\eta_-(b_{t+D}-b_{t+D+1})^+) \mid \mathcal{F}_t\Big] \\ &= \lambda B\Big(-p_t^f \\ &\quad +\beta^D \mathsf{E}\left[p_{t+D}^b(s^*-w_{t+D})\mathbb{I}(w_{t+D} < s^*) \mid \mathcal{F}_t\right] \\ &\quad +\beta^D \mathsf{E}\left[p_{t+D}^s(w_{t+D}-s^*)\mathbb{I}(w_{t+D} > s^*) \mid \mathcal{F}_t\right]\Big) \\ &\quad +B\Big(\beta^D \mathsf{E}\Big[\frac{\eta_+(-b_{t+D}+b_{t+D+1})^+ - \eta_-(b_{t+D}-b_{t+D+1})^-}{B} \\ &\qquad \mid \mathcal{F}_t\Big]\Big)+o(B) \\ &= B\Big(\beta^D \mathsf{E}\Big[\frac{\eta_+(-b_{t+D}+b_{t+D+1})^+ - \eta_-(b_{t+D}-b_{t+D+1})^-}{B} \\ &\qquad \mid \mathcal{F}_t\Big]\Big)+o(B).\end{aligned}$$

The first order (in B) terms in the last expression is independent of λ and depends only on the storage operation policy and the external statistics.

Appendix E

We compute α_t in this appendix. We note that, by the results of the previous appendix, we can focus only on the real time market interaction at the next time instant, as every subsequent time will only have a $o(B)$ effect on the value function. We also restrict ourselves to a contract s_{t-D} which is within $\Theta(B)$ of the batteryless optimal contract s^*. Taking partial derivatives with respect to b_t, and using Assumption 1, we get that

$$\begin{aligned}\alpha_t = \beta\frac{\partial}{\partial b}(\mathbb{E}[&p_t^b(s^*+\eta_+(-b_t+b)^+ \\ &-\eta_-(b_t-b)^- - w_t)^+ \\ &-p_{t+1}^s(s^*+\eta_+(-b_t+b)^+ \\ &-\eta_-(b_t-b)^- - w_{t+1})^- \mid \mathcal{F}_t])+\Theta(B)\end{aligned}$$

If the wind and price statistics are such that conditioning on $\mathcal{F}_t$ is the same as conditioning on $\mathcal{F}_{t-D}$ (one special case where this condition holds is if the external processes are independent across time) then this expression evaluates (within $\Theta(B)$) to

$$-\beta p_{t-D}^f,$$

which by assumption 1 (and assuming $\eta_- = \eta - + = 1$), is between $p_{\tilde{t}}^s$ and $p_{\tilde{t}}^b$ for all $\tilde{t}$, w.p. 1. This means that optimal battery charging policy in this case (under the assumptions above) is simply to charge if there is an excess and discharge if there is a deficit.

Appendix F

In this appendix, the formulations of the M-step certainty equivalent Model Predictive Control (MPC) and the stochastic MPC model are presented.

The M step certainty-equivalent MPC algorithm utilizes an approximate value function $V_{\text{app}}(x_t)$ to obtain a policy as

$$u_t = \underset{s_t, b_{t+1}}{\text{argmin}} \min_{s_{t+1}^{t+M}, b_{t+2}^{t+M+1}} \sum_{j=0}^{M-1} \beta^j g(s_{t+j-D+1}^{t+j+1}, b_{t+j}^{t+j+1}, z'_{t+j}) + \beta^M V_{\text{app}}(s_{t+M-D+1}^{t+M}, b_{t+M}, z'_{t+M}). \tag{18}$$

As defined earlier, $z_t = (p_t^f, p_t^s, p_t^b, w_t)$ represents the external stochastic processes. At each stage t, we see $z'_t = z_t$ and use $z'_{t+j} = \mathsf{E}[z_{t+j}] \ \forall j \in [M]$.

In stochastic MPC, we use the distribution of the future realizations of wind and prices to generate multiple samples $z'(i), i \in [N]$, where N is the number of samples. We use $z'(i)_t = z_t$ and sample $z'(i)_{t+1}^{t+M} \overset{\text{iid}}{\sim} \Pr(z_{t+1}^{t+M})$ to generate N samples indexed by $i \in [N]$.

$$\begin{aligned} u_t = &\underset{s_t, b_{t+1}}{\text{argmin}} \min_{s_{t,i}^{t+M,i}, b_{t+1,i}^{t+M+1,i}} \\ &\frac{1}{N} \sum_{i=1}^{N} \sum_{j=0}^{M-1} \beta^j g(s_{t+j-D+1,i}^{t+j+1,i}, b_{t+j,i}^{t+j+1,i}, z'(i)_{t+j}) \\ &\quad + \beta^M V_{\text{app}}(s_{t+M-D+1,i}^{t+M,i}, b_{t+M,i}, \mathsf{E}[z(i)_{t+M}]) \\ &\text{s.t.} \quad s_{t,i} = s_t, \ b_{t+1,i} = b_t \qquad \forall i \end{aligned} \tag{19}$$

Appendix G

In this section we argue that a different market model as discussed in Section III-C would yield the same contract and storage operation policies as the market model assumed in the earlier appendices does. We start with the contract policy described in Appendix B-2, and make the following observation from the discussions after Equation (17): under a different market model as in Section III-C, any deviations in the contract policy which are $\Theta(B)$ from the batteryless optimal policy would affect the cost (when the contract is realized) only by a term whose magnitude is less than first order in the size of the battery (i.e., the deviations in the cost are $o(B)$).

The derivation of the storage operation policy follows a very similar derivation as in Appendix B-1.

References

[1] U.S. Energy Information Administration, "Annual energy outlook," Tech. Rep., 2014.

[2] J. Rogers and K. Porter, Utility Wind Integration Group, "Wind power and electricity markets," Tech. Rep., 2011.

[3] P. Pinson, "Wind energy: Forecasting challenges for its operational management," *Statistical Science*, vol. 28, no. 4, pp. 564–585, 2013.

[4] Y. Zhao, J. Qin, R. Rajagopal, A. Goldsmith, and H. V. Poor, "Wind aggregation via risky power markets," *IEEE Trans. Power Syst.*, vol. 30, no. 3, pp. 1571–1581, 2015.

[5] H. L. Ferreira, R. Garde, G. Fulli, W. Kling, and J. P. Lopes, "Characterisation of electrical energy storage technologies," *Energy*, vol. 53, pp. 288–298, 2013.

[6] E. Y. Bitar, R. Rajagopal, P. P. Khargonekar, K. Poolla, and P. Varaiya, "Bringing wind energy to market," *IEEE Trans. Power Syst.*, vol. 27, no. 3, pp. 1225–1235, 2012.

[7] M. Zugno, T. Jónsson, and P. Pinson, "Trading wind energy on the basis of probabilistic forecasts both of wind generation and of market quantities," *Wind Energy*, vol. 16, no. 6, pp. 909–926, 2013.

[8] J. M. Morales, A. J. Conejo, and J. Pérez-Ruiz, "Short-term trading for a wind power producer," *IEEE Trans. Power Syst.*, vol. 25, no. 1, pp. 554–564, 2010.

[9] H.-I. Su and A. El Gamal, "Modeling and analysis of the role of energy storage for renewable integration: Power balancing," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4109–4117, 2013.

[10] P. Harsha and M. Dahleh, "Optimal management and sizing of energy storage under dynamic pricing for the efficient integration of renewable energy," *IEEE Transactions on Power Systems*, vol. 30, no. 3, pp. 1164–1181, 2015.

[11] J. Qin, Y. Chow, J. Yang, and R. Rajagopal, "Online modified greedy algorithm for storage control under uncertainty," *IEEE Transactions on Power Systems*, vol. 31, no. 3, pp. 1729–1743, May 2016.

[12] E. Bitar, R. Rajagopal, P. Khargonekar, and K. Poolla, "The role of co-located storage for wind power producers in conventional electricity markets," in *Proc. American Control Conference (ACC)*, 2011, pp. 3886–3891.

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

[14] S. Kurandwad, C. Subramanian, A. Vasan, V. Sarangan, V. Chellaboina, A. Sivasubramaniam *et al.*, "Windy with a chance of profit: bid strategy and analysis for wind integration," in *Proc. 5th International Conf. on Future Energy Syst.* ACM, 2014, pp. 39–49.

[15] J. Taylor, D. S. Callaway, K. Poolla *et al.*, "Competitive energy storage in the presence of renewables," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 985–996, 2013.

[16] R. Jiang, J. Wang, and Y. Guan, "Robust unit commitment with wind power and pumped storage hydro," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 800–810, 2012.

[17] D. Gayme and U. Topcu, "Optimal power flow with large-scale storage integration," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 709–717, 2013.

[18] K. Divya and J. Østergaard, "Battery energy storage technology for power systems—an overview," *Electric Power Syst. Research*, vol. 79, no. 4, pp. 511–520, 2009.

[19] B. Øksendal, *Stochastic differential equations*. Springer, 2003.

[20] M. Rao, M. Chowdhury, Y. Zhao, T. Javidi, and A. Goldsmith, "Value of storage for wind power producers in forward power markets," in *Proc. American Control Conference (ACC)*, 2015, pp. 5686–5691.

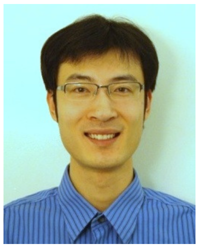
[21] Y. Wang, B. O'Donoghue, and S. Boyd, "Approximate dynamic programming via iterated bellman inequalities," *International Journal of Robust and Nonlinear Control*, 2014.

[22] National Renewable Energy Laboratory, "Eastern wind integration and transmission study," Tech. Rep., 2010.

[23] "Github code repository," https://github.com/milindmrao/WindFarm_MAT, accessed: 2016-03-30.

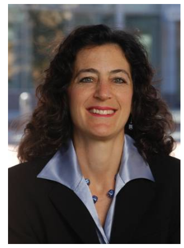
MAINAK CHOWDHURY is pursuing a Ph.D. degree in the Wireless Systems Laboratory at Stanford University. He received a Bachelor of Technology in Electrical Engineering from the Indian Institute of Technology, Kanpur, India (2011), and a Master of Science in Electrical Engineering from Stanford University (2013). His research interests include communication systems with a large number of transmitters or receivers and applications involving stochastic control.

MILIND RAO (S'14) received the B. Tech degree in Electrical Engineering from IIT Madras, India in 2013 where he graduated with the Siemens award for academic proficiency. He received the M.S. degree in Electrical Engineering from Stanford University in 2015 and is currently pursuing the Ph.D. degree in the Wireless Systems Laboratory at Stanford University. He is a recipient of the Stanford Graduate, DAAD and KVPY fellowships. His research interests are in stochastic control, optimization, signal processing, and wireless communications.



YUE ZHAO (S'06, M'11) is an assistant professor of electrical and computer engineering at Stony Brook University. He received the B.E. degree in electronic engineering from Tsinghua University, Beijing, China in 2006, and the M.S. and Ph.D. degrees in electrical engineering from the University of California, Los Angeles (UCLA), Los Angeles in 2007 and 2011, respectively. His current research interests include renewable energy integration, smart grid, optimization and game theory, and statistical signal processing.

Tara Javidi (S'96-M'02) studied electrical engineering at the Sharif University of Technology from 1992 to 1996. She received her MS degrees in Electrical Engineering: Systems, and Applied Mathematics: Stochastics, from the University of Michigan, Ann Arbor, MI. She received her PhD in EECS from University of Michigan, Ann Arbor in 2002. From 2002 to 2004, Tara was an assistant professor at the electrical engineering department, University of Washington, Seattle. She joined University of California, San Diego, in 2005, where she is currently an associate professor of electrical and computer engineering. She was a Barbour Scholar during 1999-2000 academic year and received an NSF CAREER Award in 2004. Her research interests are in communication networks, stochastic resource allocation, and wireless communications



ANDREA GOLDSMITH (S'90-M'93-SM'99-F'05) is the Stephen Harris professor in the School of Engineering and a professor of Electrical Engineering at Stanford University. She was previously on the faculty of Electrical Engineering at Caltech. Her research interests are in information theory and communication theory, and their application to wireless communications and related fields. She co-founded and served as Chief Scientist of Plume WiFi, and previously co-founded and served as CTO of Quantenna Communications, Inc. She has also held industry positions at Maxim Technologies, Memorylink Corporation, and AT&T Bell Laboratories. Dr. Goldsmith is a Fellow of the IEEE and of Stanford, and has received several awards for her work, including the IEEE ComSoc Edwin H. Armstrong Achievement Award as well as Technical Achievement Awards in Communications Theory and in Wireless Communications, the National Academy of Engineering Gilbreth Lecture Award, the IEEE ComSoc and Information Theory Society Joint Paper Award, the IEEE ComSoc Best Tutorial Paper Award, the Alfred P. Sloan Fellowship, the WICE Technical Achievement Award, and the Silicon Valley/San Jose Business Journal's Women of Influence Award. She is author of the book "Wireless Communications" and co-author of the books "MIMO Wireless Communications" and "Principles of Cognitive Radio," all published by Cambridge University Press, as well as an inventor on 28 patents. She received the B.S., M.S. and Ph.D. degrees in Electrical Engineering from U.C. Berkeley.

Dr. Goldsmith has served on the Steering Committee for the IEEE Transactions on Wireless Communications and as editor for the IEEE Transactions on Information Theory, the Journal on Foundations and Trends in Communications and Information Theory and in Networks, the IEEE Transactions on Communications, and the IEEE Wireless Communications Magazine. She participates actively in committees and conference organization for the IEEE Information Theory and Communications Societies and has served on the Board of Governors for both societies. She has also been a Distinguished Lecturer for both societies, served as President of the IEEE Information Theory Society in 2009, founded and chaired the student committee of the IEEE Information Theory society, and chaired the Emerging Technology Committee of the IEEE Communications Society. At Stanford she received the inaugural University Postdoc Mentoring Award, served as Chair of Stanford's Faculty Senate and for multiple terms as a Senator, and currently serves on its Budget Group, Committee on Research, and Task Force on Women and Leadership.