Meeting Inelastic Demand in Systems with Storage and Renewable Sources

Soongeol Kwon, Yunjian Xu, and Natarajan Gautam

***Abstract*—We consider a system where inelastic demand for electric power is met from three sources: the grid, in-house renewables such as solar panels, and an in-house energy storage device. In our setting, energy demand, renewable power supply, and cost for grid power are all time-varying and stochastic. Further, there are limits and inefficiency associated with charging and discharging the energy storage device. We formulate the storage operation problem as a dynamic program with parameters estimated from real-world demand, supply and cost data. As the dynamic program is computationally intensive for large-scale problems, we explore algorithms based on approximate dynamic programming (ADP), and apply them to a test data set. Using the real-world test data, we numerically compare the performance of two ADP-based algorithms against Lyapunov optimization based algorithms that require no statistical knowledge. Our results ascertain the value of storage and the value of installing a renewable source.**

I. Introduction

Renewable generation capacity is expanding rapidly to potentially reduce carbon dioxide emissions and dependence on fossil fuels. Being a source of non-dispatchable generation, renewable energy introduces variability into the energy portfolio, and further amplifies the difficulty of matching demand with supply in real time. Energy storage is an environmentally friendly candidate that can provide flexibility to the system and mitigate the impact of volatile renewable generation.

The focus of this paper is on the operation of electric storages operated by the electricity consumers who own distributed renewable generation and face time-varying and stochastic electricity prices. Our motivation stems from the potential of electricity consumers to own and use storage devices (e.g., major consumers like data centers [1] and individual consumers who own PHEVs [2]), and from a recent study that shows consumer ownership of storage can be socially beneficial [3]. We also note that there is a growing trend for residential consumers and data centers to own distributed renewable generation [4], [5].

In this paper, we consider a *consumer* of electricity with *inelastic demand*, i.e., in each time period the consumer has to consume a certain (time-varying and possibly random) amount energy that is independent of the price of electricity. Part of the demand can be met by a renewable energy source (such as photo-voltaic (PV) solar panels) that is situated locally and owned by the consumer. Note that renewable power *supply* is time-varying and stochastic. Remaining demand (if any), beyond what the renewable source can supply, is satisfied either from the *grid* or by an in-house *Energy Storage Device* (ESD) or both. Like power demand and renewable supply, *price* for power from the grid is also time-varying and stochastic.

In the last few years this problem has received a lot of attention. There exists a substantial literature on the operation of energy storage owned by renewable generators or system operators. The joint scheduling of variable wind generation and energy storage systems is studied in order to maximize the joint profit of wind farms and energy storage systems, through a two-stage stochastic programming formulation [6], and a model predictive control (MPC)-based approach [7]. The authors of [8] derive an upper bound on the marginal value of storage (at small installed capacities) for a transmission-constrained power network. A few recent works study the optimal operation of energy storage devices with an objective of minimizing the mismatch between the available renewable generation and system load [9], [10], [11].

Another well studied application of energy storage is the use of storages to arbitrage [12], [2]. A few recent works conduct a dynamic programming approach to derive the arbitrage value of electric storage, in the presence of dynamic pricing [13], [14]. Different from the setting in the present paper, this aforementioned literature assumes that the operator of energy storages (e.g., an arbitrager) has zero demand for electricity and puts no value on its own electricity consumption.

There have been recent studies on the operation of consumer-owned ESDs. The authors of [15] study the day-ahead scheduling of energy storage by analyzing a noncooperative game among consumers. There is a growing literature that applies Lyapunov optimization based on-line algorithms on the operation of consumer-owned ESDs [16], [17], [18]. These on-line algorithms are shown to be asymptotically optimal, as the storage capacity increases to infinity.[1] Unlike the Markov decision process (MDP), which is computationally complex and requires substantial statistical information of the system dynamics, Lyapunov optimization based algorithms use simple linear programs to make storage operation decisions based only on the current system state (e.g., the current storage

Soongeol Kwon and Natarajan Gautam are with the Department of Industrial and Systems Engineering, Texas A&M University (email: soongeol@email.tamu.edu, gautam@tamu.edu). Yunjian Xu (Corresponding author) is with the Engineering Systems and Design Pillar, Singapore University of Technology and Design, Singapore (email: yunjian_xu@sutd.edu.sg).

This work is partially supported by the AFOSR under Contract No. FA9550-13-1-0008, and the MIT-SUTD International Design Center (IDC) Grant IDG21400103. Thanks to Gareth Jones for collecting household power consumption data. A preliminary version of this paper appeared in the Proceedings of the 5th IEEE International Conference on Smart Grid Communications (SmartGridComm 2014).

[1]It is worth noting that these Lyapunov optimization based on-line algorithms may fail to achieve asymptotic optimality if the storage efficiency is less than 1, i.e., if the storage has non-negligible self-discharging [19].

level). Our numerical results show that the algorithm proposed in [17] performs well when the storage capacity is significantly larger than the maximum charging/discharging rates, i.e., when it takes many hours to fully charge and discharge the storage.

Closely related to the present paper, a few recent papers establish structural properties on optimal storage operation policies in a variety of MDP settings that incorporate time-varying (and/or stochastic) cost and demand [20], [21], [22]. The main results of these theoretic works are the existence of an optimal policy that can be characterized by (time-varying and possibly state dependent) operational thresholds. The computation of these thresholds usually becomes intractable for practical settings with stochastic renewable generation and varying electricity prices, for example, in our MDP setting where time is explicitly incorporated into the system state, and each time period lasts for only 5 minutes (288 time periods per day).

We formulate the storage operation problem as an MDP with periodic cycles. The parameters of the MDP are trained using a set of real data on electricity prices, solar generation, consumer demand. The *main contribution* of this paper is to implement and numerically compare approaches ranging from Markovian models, to hybrid methods based on statistics and optimization, to those that are based on Lyapunov optimization and require no historical information, under a variety of parameter settings on the storage size, the level of solar generation, as well as the maximum charging/discharging rates.

We introduce and test two approximate dynamic programming (ADP) based heuristic policies, which usually yield the best performance among all tested heuristics. The first ADP-based policy, which is referred to as One-step Lookahead Algorithm (OLA), chooses an action that minimizes the expected cost for the current and the next state. While OLA always treats the next stage as the terminal stage, the second ADP-based heuristic policy, which is referred to as One-step Roll-out algorithm (ORA), approximates the cost-to-go at every possible next system state by solving a deterministic (certainty-equivalent) optimization problem with future system stochasticity taking the expected value.

The insight we obtain from numerical experiments sheds some light on the effectiveness of different types of heuristic policies and the value of storage under various parameter settings. We summarize our key findings in the following.

1) Algorithms based on Lyapunov optimization (e.g., the one proposed in [17]) require minimum (almost negligible) computational efforts, and perform reasonably well when the storage capacity is significantly larger than the maximum charging/discharging rates. For fast-charging storage devices that can be fully charged within 2 hours, on the other hand, ADP-based algorithms (i.e., ORA and OLA) significantly outperform the one proposed in [17].
2) The value of storage (VoS, which is measured as the net benefit obtained by the consumer if she operates the storage according to an ADP algorithm) is much higher under 5-minute real-time pricing than that under hourly pricing, due the higher variability in cost in the former case. VoS increases sharply with the storage capacity only when the maximum charging/discharging rates grow in proportion to the storage capacity. In other words, the value of storage does not increase appreciably with increase in storage size, if the maximum charging/discharging rates remain fixed.

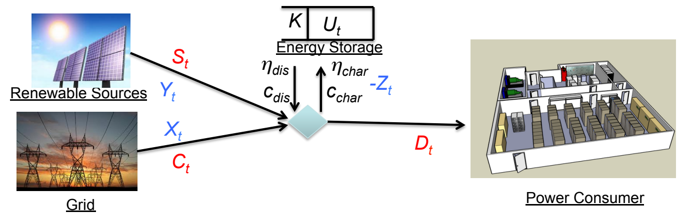


Fig. 1. Schematic representation of scenario and notation

The rest of the paper is organized as follows. We describe our problem in Section II. In Section III we develop a probabilistic model and suggest ADP-based approaches to solve it. We consider other heuristic algorithms in Section IV. We compare and discuss the performance of these algorithms in Section V by obtaining parameters using a real training data set and testing them. Finally, we make some brief concluding remarks in Section VI.

II. Problem Description

Before describing our model, we state a few key features of our problem setup. We consider inelastic demand, which must always be met instantaneously in real time and cannot be either postponed or cut back using incentives such as prices.[2] The ESD has a finite energy storage capacity. Price is non-negative and exogenous (not affected by the consumer's decisions). There are inefficiencies in charging and discharging, but no leakages (i.e., self-discharging) in the ESD. This assumption of 100% storage efficiency (no self-discharging) is reasonable since many popular types of modern batteries (e.g., Lead acid, Sodium Sulphur (NaS), Lithium ion, and Vanadium redox batteries) have negligible self-discharge ($0-5\%$ per month) [24], and further, the effective planning horizon for storage operation is usually no more than a week.

We now describe some notations used in this work (pictorially described in Fig. 1). We consider a discrete-time model where time periods are indexed by $t = 0, 1, \ldots$. The amount of energy in the ESD at the beginning of period t is denoted by U_t (in kWh). The *stochastic* uncontrollable variables are: D_t, the *demand* for energy in period t (in kWh); S_t, the energy *supply* from renewable source in period t (in kWh); C_t, the *cost* in period t for a unit of energy from grid (in \$/kWh). Table I summarizes the acronyms used in this paper.

There are constraints and inefficiencies in the ESD charging and discharging processes. A maximum of K kWh of energy can be stored in the ESD at any time. The ESD can be discharged and charged at a maximum rate of c_{dis} and c_{char} (in kW) respectively. Also, the ESD discharging and charging efficiencies are $\eta_{dis} \leq 1$ and $\eta_{char} \leq 1$ respectively (which we explain next). If ρ kWh of energy is used to charge the ESD

[2] We note that electricity consumption usually exhibits inelasticity in the short term [23], and that the setting of inelastic demand is used in many related works, e.g., [9], [20].

TABLE I
SUMMARY OF ACRONYMS

Acronym	Description
ESD	Energy Storage Device
MDP	Markov Decision Process
OLA	One-step Look-ahead Algorithm
ORA	One-step Roll-out Algorithm
TBA	Threshold-Based Approximation algorithm
NOA	Naive Opportunistic Algorithm
HWR	The on-line algorithm proposed in Huang, Walrand and Ramchandran [17]

in period t, then the increase in ESD level $U_{t+1} - U_t$ is $\rho\eta_{char}$ kWh. Likewise if ρ kWh of energy is needed from the ESD, then the increase in ESD level $U_{t+1} - U_t$ is $-\rho/\eta_{dis}$ kWh. Next we describe the *decision variables* under the control of the consumer. Let X_t, Y_t and Z_t be the energy drawn (in kWh) from the grid, renewable source and ESD respectively at period t. While $X_t \geq 0$ and $Y_t \geq 0$ for all t, Z_t can be positive or negative.

During every period t, given demand D_t, renewable supply S_t, cost C_t and ESD charge level U_t, we need to *determine* the supply from grid X_t, the draw from renewable source Y_t, and contribution from the ESD Z_t so that the long-run expected total cost is minimized subject to satisfying demand, staying within ESD capacities and other constraints such as dynamics and non-negativity. As in [20], [21], we are interested in minimizing the total expected discounted cost. This sequential decision making problem can be formulated mathematically as follows

$$\text{Minimize}_{(X_t,Y_t,Z_t)} \lim_{T\to\infty} \textstyle\sum_{t=0}^{T} \beta^t \, \mathbb{E}[C_t X_t] \tag{1}$$

$$\text{Subject to the following} \quad \forall t \in \{0,1,2,\ldots\}, \tag{2}$$

$$X_t + Y_t + Z_t \geq D_t, \tag{3}$$

$$0 \leq Y_t \leq S_t, \quad \psi Z_t \leq c_{dis}, \tag{4}$$

$$-\psi \min\{Z_t, 0\} \leq c_{char}, \tag{5}$$

$$U_{t+1} - U_t = -\max\{Z_t, 0\}/\eta_{dis} - \eta_{char} \min\{Z_t, 0\}, \tag{6}$$

$$0 \leq U_t \leq K, \quad X_t \geq 0, \tag{7}$$

where $\beta \in (0,1)$ is a discount factor, and ψ is a constant for time-unit conversion, i.e. number of time units per hour (viz. since c_{char} is in kW and Z_t in kWh, if length of period t is 1 second then $\psi = 3600$). In constraint (3) we have implicitly assumed free disposal of renewable generation.[3] For all $t \geq 0$, C_t, D_t, and S_t are modeled as discrete random variables, and all constraints in the above optimization problem must hold for every trajectory of realized demand, renewable supply, and cost.

Remark 1: We can reduce the above problem to a 1-dimensional control in X_t or Z_t by realizing that we can let $Y_t = S_t$ and $Z_t = D_t$ - X_t - Y_t, subject to the charging/discharging rate constraints [6], [16], [21]. However, for ease of presentation we will use all variables, not just X_t or Z_t.

Remark 2: It is shown in [16], [20], [22] that the optimal policy can be characterized by two thresholds. Given the system state at period t, namely t, D_t, S_t, C_t, and U_t, the optimal policy does the following: (i) if U_t lies between the two thresholds, do not charge or discharge the storage; (ii) if U_t lies below the lower threshold, greedily charge the storage up to this threshold; (iii) if U_t lies above the higher threshold, then discharge the storage to fulfil the demand. The threshold structure of optimal policies will provide some guidance on the design of heuristic policies (proposed in Sections III-B and IV). We finally note that when $\eta_{char} = \eta_{dis} = 1$ (no charging/discharging inefficiency), there exists a simpler optimal policy with a single threshold [14].

III. MDP: Probabilistic Model with Cycles

In this section, we first introduce the way we fit real data into an MDP model, and then discuss approaches to solve it. Analyzing the data described in [25], [4] and the NREL labs, it is evident that demand, solar PV supply and cost are time-varying and stochastic. However, it is also not unreasonable to assume that there are daily or weekly effects. In other words, there is a deterministic variability as well as stochastic variability. To model such a phenomenon we consider what we call *probabilistic model with cycles*.

Definition 1: An uncontrolled process $\{V_t\}_{t=1}^{\infty}$ is *cyclic* with cycle length N if the joint probabilistic distribution of $\{V_{\tau+\ell N}\}_{\tau=0}^{N-1}$ is identical for all $\ell \in \{1,2,\ldots\}$, where N is the number of periods in a cycle. Each cycle lasts for T hours, and therefore has $N = \psi T$ periods. $\square$

Based on the above definition, we assume that $\{D_t\}_{t=1}^{\infty}$, $\{S_t\}_{t=1}^{\infty}$ and $\{C_t\}_{t=1}^{\infty}$ are *cyclic* with cycle length T (one cycle typically is the equivalent of one day). Further, we write down for all $t \in \{0,1,\ldots\}$, with $n = (t \mod N)$,

$$D_t = d_n W_t^d + \delta_n, \quad S_t = s_n W_t^s, \quad C_t = c_n W_t^c + \gamma_n,$$

where $\{d_1, d_2, \ldots, d_N\}$, $\{\delta_1, \delta_2, \ldots, \delta_N\}$, $\{s_1, s_2, \ldots, s_N\}$, $\{c_1, c_2, \ldots, c_N\}$ and $\{\gamma_1, \gamma_2, \ldots, \gamma_N\}$, are sets of deterministic constants while $\{W_t^d\}_{t=1}^{\infty}$, $\{W_t^s\}_{t=1}^{\infty}$, and $\{W_t^c\}_{t=1}^{\infty}$, are stationary and independent *discrete time Markov chains* on state spaces $\mathcal{S}^d$, $\mathcal{S}^s$, and $\mathcal{S}^c$ and transition probability matrices $\mathcal{P}^d$, $\mathcal{P}^s$, and $\mathcal{P}^c$ respectively.

One can think of $\{s_1, s_2, \ldots, s_N\}$ as the power supplied by PV panels on a perfectly sunny day while $\mathcal{S}^s$ is a continuous set of values between 0 and 1. The *demand* and *cost* terms do not have such a nice interpretation and one would have to model them carefully based on data.

Note that since D_t, S_t and C_t are continuous, so are W_t^d, W_t^s and W_t^c.[4] However, to model as an MDP, we need state spaces $\mathcal{S}^d$, $\mathcal{S}^s$, and $\mathcal{S}^c$ to be discrete. We discretize W_t^d, W_t^s and W_t^c using discrete random variables $\tilde{W}_t^d$, $\tilde{W}_t^s$ and $\tilde{W}_t^c$ each of which take $M+1$ different values. For example the aforementioned W_t^s would be mapped from a statespace $\mathcal{S}^s =$

[3] It is optimal to use the renewable generation first to meet the demand, and then to charge the residual renewable generation to the ESD. If there is not enough storage capacity to absorb the residual renewable generation, then it is possible that $Y_t + Z_t > D_t$ (note that $X_t = 0$ and $Z_t \leq 0$ in this case).

[4] These random parameters W_t^d, W_t^s and W_t^c are assumed to be not correlated although D_t, S_t and C_t might themselves be correlated due to the correlation in their deterministic components.

$[0,1]$ to $\tilde{\mathcal{S}}^s = [0,\ 1/M,\ 2/M,\ \ldots, 1]$. Thus $\tilde{W}_t^d$, $\tilde{W}_t^s$ and $\tilde{W}_t^c$ would each be $M+1$ state discrete time Markov chains with transition probability matrices $\mathcal{P}^d$, $\mathcal{P}^s$, and $\mathcal{P}^c$ respectively. By some abuse of notations, for the rest of this paper we let D_t, S_t, and C_t denote the corresponding discretized values of demand, renewable generation, and cost at stage t.

In Section V, we will use training data to estimate d_n, δ_n, s_n, c_n and γ_n for all $n \in \{1, \ldots, N\}$ as well as $\mathcal{P}^d$, $\mathcal{P}^s$, and $\mathcal{P}^c$. However, for the rest of this section we take a probabilistic approach assuming all the aforementioned parameters are known and formulate the system as an MDP.

We denote the **system state** at time t as a 5-tuple

$$\mathbf{x} = \{\{t/N\}+1, W_t^d, W_t^s, W_t^c, U_t\}, \tag{8}$$

(where $\{t/N\}$ denotes ($t \mod N$)) with state space $\mathcal{S}$ given by the cartesian product

$$\{1, 2, \ldots, N\} \times \tilde{\mathcal{S}}^d \times \tilde{\mathcal{S}}^s \times \tilde{\mathcal{S}}^c \times \tilde{\mathcal{S}}^u,$$

where $\tilde{\mathcal{S}}^u$ is the discrete set of values between 0 and K that U_t can take. We note that the constructed MDP is *stationary*, because the time dependency and correlations of consumer demand and renewable generation are incorporated by including in the system state a periodic Markov chain that describes time evolution.

Let the **action** at time t denote the amount of power to be supplied from the grid, i.e. X_t, with action space $\mathcal{A}(\mathbf{x})$ corresponding to the set of all possible real numbers that lie in the following interval

$$\mathcal{A}(\mathbf{x}) = \Big[\left(D_{\{t/N\}} - S_{\{t/N\}} - \min\{\eta_{dis} U_{\{t/N\}}, c_{dis}/\psi\}\right)^+, \left(D_{\{t/N\}} - S_{\{t/N\}} + \min\left\{\tfrac{K - U_{\{t/N\}}}{\eta_{char}}, c_{char}/\psi\right\}\right)^+ \Big], \tag{9}$$

where $(\cdot)^+ = \max\{\cdot, 0\}$. Here, the lower bound of the action space is the amount of energy needed from the grid to fulfil the demand, when the storage is greedily discharged for consumption, and the upper bound is the amount of energy needed to meet the demand and to greedily charge the storage.

As noted in Remark 1, the energy drawn from renewable source and ESD at period t is determined by X_t, i.e., $Y_t = S_t$ and $Z_t = D_t$ - X_t - Y_t.

Given any $\mathbf{x} \in \mathcal{S}$, $X_t \in \mathcal{A}(\mathbf{x})$, and $\mathbf{y} \in \mathcal{S}$, we can compute the **transition probability** $P_{\mathbf{xy}}(X_t)$ using appropriate Kronecker products of $\mathcal{P}^d$, $\mathcal{P}^s$, $\mathcal{P}^c$ and other matrices of zeros and ones (which are not explained due to space constraints). The next-stage storage level is given by

$$U_{t+1} = U_t - \max\{Z_t, 0\}/\eta_{dis} - \eta_{char} \min\{Z_t, 0\}, \tag{10}$$

where $Z_t = D_t - S_t - X_t$.

The **stage cost** at time t is the product of the corresponding power cost in state i times $X_t \in \mathcal{A}(\mathbf{x})$, $C_t X_t$.

By incorporating the time element into the state of the dynamic program, we have indeed formulated a "stationary" MDP (the quotes are because the state transition is stationary from one cycle of N values to the next cycle, but not within a cycle). For a given stationary policy which maps every possible system state $\mathbf{x}$ to a point in the action space $\mathcal{A}(\mathbf{x})$, the long-run discounted total cost corresponds to the objective function of our optimization problem defined in Section II (and also results in a feasible solution).

A. Exact MDP Solution

Note that the above MDP has a finite state-space and a finite action-space. There are many methods to obtain the optimal action $a \in \mathcal{A}(\mathbf{x})$ at state $\mathbf{x}$ for all $\mathbf{x} \in \mathcal{S}$. We consider a linear program (LP) based method. The following LP solves the optimal cost-to-go at each state $\mathbf{x}$, $\{J^*_{\mathbf{x}}\}_{\mathbf{x}\in\mathcal{S}}$ [26]:

$$\begin{aligned} &\mathrm{Max}_{\{J_{\mathbf{x}}\}} \quad \textstyle\sum_{\mathbf{x}\in\mathcal{S}} c_{\mathbf{x}} J_{\mathbf{x}} \\ &\text{s.t.} \quad g_{\mathbf{x}}(a) + \beta \sum_{\mathbf{y}\in\mathcal{S}} P_{\mathbf{xy}}(a) J_{\mathbf{y}} \geq J_{\mathbf{x}}, \;\; \forall \mathbf{x} \in \mathcal{S},\ \forall a \in \mathcal{A}(\mathbf{x}), \end{aligned} \tag{11}$$

where $\{c_{\mathbf{x}}\}_{\mathbf{x}\in\mathcal{S}}$ is a given vector with positive components, and $g_{\mathbf{x}}(a)$ is the stage cost at state $\mathbf{x}$ under action a. The optimal action to be taken at each state $\mathbf{x}$ can then be obtained by the following Bellman equation:

$$a^* \in \arg\min_{a\in\mathcal{A}(\mathbf{x})} \left\{ g_{\mathbf{x}}(a) + \beta \textstyle\sum_{\mathbf{y}\in\mathcal{S}} P_{\mathbf{xy}}(a) J^*_{\mathbf{y}} \right\}. \tag{12}$$

In this manner it is possible to determine the optimal action in each state (in theory).

In summary, the MDP algorithm works as follows: given the demand, renewable generation, cost, and storage level at stage t, we first obtain the discretized values D_t, S_t, C_t and U_t. Using those we compute the optimal action Z_t prescribed by the MDP. Then we obtain the actions $Y_t = S_t$ and $X_t = \max(D_t - S_t - Z_t, 0)$.

We note that an optimal solution to the LP formulated in (11) must exist. However in practice, one could encounter difficulties known as the *curse of dimensionality*. The exact MDP can be solved (especially by packages such as MATLAB) only when the action space is small, $\eta_{char} = \eta_{dis} = 1$, and C_t, D_t and S_t belong to a small discrete set for all $t \in \{1, \ldots, N\}$. In the next subsection, we adopt a common procedure usually referred to as **Approximate Dynamic Programming (ADP)** to deal with the curse of dimensionality.

B. Approximate MDP Solution

In this section, we introduce two simple ADP-based algorithms that will be tested against other heuristics as well as the exact MDP solution in Section V. An effective way to reduce the computation required by a dynamic program is to truncate the time horizon and at each stage make a decision based on look-ahead of a small number of stages [27], [28]. In particular, we will focus on the simplest ADP algorithms that look only a single stage ahead. Numerical results in Section V demonstrate that even these simplest ADP algorithms usually (significantly) outperforms Lyapunov optimization based algorithms. It is worth noting that, however, even these simplest ADP algorithms require much more computation than Lyapunov optimization based algorithms.

According to the simplest ADP-based policy, which is referred to as **One-Step Look-ahead algorithm** (OLA) in this paper, given the current state we determine the best action so that the expected cost for this state and the next state is minimized. Formally, given the current system state $\mathbf{x}$, the algorithm chooses an action $X_t \in \mathcal{A}_t$ that minimizes the following cost:

$$C_t X_t + \sum_{\mathbf{y}} \beta P_{\mathbf{xy}}(X_t) \cdot C_{t+1} \cdot (D_{t+1} - S_{t+1} - \eta_{dis} U_{t+1})^+ , \tag{13}$$

where $(\cdot)^+ = \max\{\cdot, 0\}$, and $\mathbf{y}$ is the system state at time $t+1$ that includes the parameters C_{t+1}, D_{t+1}, S_{t+1} and U_{t+1}. In this myopic version of one-step look-ahead policy, the stage $t+1$ is treated as the terminal stage, and therefore the policy fully discharges the storage to fulfill the demand at stage $t+1$. This would result in a stage cost at time $t+1$ of $C_{t+1}\left(D_{t+1} - S_{t+1} - \eta_{dis} U_{t+1}\right)^+$.

A natural way to improve OLA is to replace the myopic stage cost at state $\mathbf{y}$ by some approximated cost-to-go at this state. Formally, given the current system state $\mathbf{x}$, a **One-step Roll-out algorithm (ORA)** chooses an action $X_t \in \mathcal{A}_t$ that minimizes the following cost:

$$C_t X_t + \sum\nolimits_{\mathbf{y} \in \mathcal{S}} \beta P_{\mathbf{xy}}(X_t) \cdot \tilde{J}_{\mathbf{y}}, \tag{14}$$

where $\tilde{J}_{\mathbf{y}}$ is an approximation of the cost-to-go at system state $\mathbf{y}$, $J_{\mathbf{y}}$. We note that if the approximation is exact, i.e., if $\tilde{J}_{\mathbf{y}} = J_{\mathbf{y}}$, then the above Bellman recursion must yield the optimal action at the current system state $\mathbf{x}$.

The OLA simply treats stage $t+1$ as the terminal stage and let $\tilde{J}_{\mathbf{y}}$ be the stage cost at time $t+1$ in state $\mathbf{y}$. The ORA, on the other hand, solves a certainty equivalent optimization problem to approximate the cost-to-go at possible next-stage system states, where all random variables take the expected value, given that the system state at time $t+1$ is $\mathbf{y}$:

$$\begin{aligned}
\tilde{J}_{\mathbf{y}} = \ & \text{Minimize}_{\{\bar{X}_\tau, \bar{Y}_\tau, \bar{Z}_\tau, \bar{U}_\tau\}} \sum_{\tau=t+1}^{t+N} \beta^{\tau-t-1} \mathbb{E}[C_\tau \mid \mathbf{y}] \cdot \bar{X}_\tau \\
& \text{Subject to } \ \forall \tau \in \{t+1, t+2, \ldots, t+N\}, \\
& \bar{X}_\tau + \bar{Y}_\tau + \bar{Z}_\tau \geq \mathbb{E}[D_\tau \mid \mathbf{y}], \\
& 0 \leq \bar{Y}_\tau \leq \mathbb{E}[S_\tau \mid \mathbf{y}], \quad \psi \bar{Z}_\tau \leq c_{dis}, \\
& -\psi \min\{\bar{Z}_\tau, 0\} \leq c_{char}, \\
& \bar{U}_{\tau+1} - \bar{U}_\tau = -\max\{\bar{Z}_\tau, 0\}/\eta_{dis} - \eta_{char} \min\{\bar{Z}_\tau, 0\}, \\
& 0 \leq \bar{U}_\tau \leq K, \quad \bar{X}_\tau \geq 0,
\end{aligned} \tag{15}$$

where $\bar{U}_{t+1}$ is determined by the system state $\mathbf{y}$, $\mathbb{E}[\cdot\,|\mathbf{y}]$ denotes conditional expectation, and the minimization is taken over the variables $\bar{X}_\tau$, $\bar{Y}_\tau$, and $\bar{Z}_\tau$.

Certainty equivalent control is a simple and intuitive way to make sequential decisions under uncertainty. It is shown to be optimal for Linear-Quadratic-Gaussian (LQG) problems [29]. Certainty equivalence is the logic underlying the current practice of unit commitment (which is a deterministic optimization problem with random demand and renewable generation taking the expected value), and is proposed for the economic/environmental dispatch of power systems with intermittent renewable generation [30]. The certainty equivalent approximation results in significant computational savings by avoiding computing the exact cost-to-go at stage $t+1$, and on the other hand, makes the ORA suboptimal.

Since the system state space is continuous (due to the continuous storage level), given the current system state $\mathbf{x}$, it is impossible to evaluate the cost-to-go of every possible state $\mathbf{y}$ at stage $t+1$ (by solving the certainty equivalence optimization problem in (15)). In our simulation, we therefore discretize the space of demand, renewable supply, and cost at stage $t+1$, and restrict our attention to a finite set of actions X_t. Formally, given the current system state $\mathbf{x}$, ORA chooses an action X_t from a finite set of actions $\bar{\mathcal{A}}(\mathbf{x}) \in \mathcal{A}(\mathbf{x})$, in order to minimize

$$C_t X_t + \sum\nolimits_{\mathbf{y} \in \bar{\mathcal{S}}(\mathbf{x})} \frac{\beta P_{\mathbf{xy}}(X_t)}{\sum_{\mathbf{y} \in \bar{\mathcal{S}}(\mathbf{x})} P_{\mathbf{xy}}(X_t)} \tilde{J}_{\mathbf{y}}, \tag{16}$$

where $\bar{\mathcal{S}}(\mathbf{x})$ is a finite subset of $\mathcal{S}$ that includes the discretized states of demand, renewable supply, and cost, as well as a finite set of storage levels U_{t+1} resulting from the finite set of actions in $\bar{\mathcal{A}}(\mathbf{x})$ (according to Eq. (10)). In our simulation, the discretized states (of demand, renewable supply, and cost) are uniformly distributed in a compact set that is estimated from real data. The set of actions $\bar{\mathcal{A}}(\mathbf{x})$ explored by ORA at state $\mathbf{x}$ includes the following: (i) charge the battery only if there is surplus in renewable generation, i.e., $X_t = (D_t - S_t)^+$, (ii) greedily discharge the battery to meet the demand, i.e., $Z_t = \min\left\{c_{dis}/\psi, \eta_{dis} U_t, (D_t - S_t)^+\right\}$, (iii) greedily charge the battery, i.e., $Z_t = -\min\left\{c_{char}/\psi, (K - U_t)^+/\eta_{char}\right\}$, and several additional actions that are uniformly distributed between actions (ii) and (iii).

It is worth noting that unlike MDP, ORA does not discretize the storage level *in priori*. Under ORA, the (finite) set of next-stage storage levels in $\bar{\mathcal{S}}(\mathbf{x})$ is determined by the set of actions $\bar{\mathcal{A}}(\mathbf{x})$ as well as the current states U_t, D_t, and S_t.

IV. Other Heuristics

As a one-step rollout policy on top of certainty equivalent control, ORA uses the Markovian structure of our model (via the state transition probability from stage t to $t+1$). To assess how our key algorithm ORA performs with real data, we compare it against other certainty equivalence based algorithms that use neither the discrete framework nor the Markovian structure. For that, in this section we leverage upon existing approaches to develop two heuristic policies: TBA (threshold-based approximation) algorithm and NOA (naive opportunistic algorithm).

For both the heuristic TBA and NOA algorithms, we first solve the following (unconditioned) certainty equivalence (UCE) problem to obtain the variables $\hat{X}_\tau$, $\hat{Y}_\tau$, $\hat{Z}_\tau$, and $\hat{U}_\tau$, for $\tau = 1, 2, \ldots, N$.

$$
\begin{aligned}
\text{UCE: Minimize}_{\{\hat{X}_\tau,\hat{Y}_\tau,\hat{Z}_\tau,\hat{U}_\tau\}} & \sum_{\tau=1}^{N} \beta^{\tau-1}\,\mathbb{E}[C_\tau]\hat{X}_\tau, \\
\text{Subject to } & \forall \tau \in \{1,\ldots,N\}, \\
& \hat{X}_\tau + \hat{Y}_\tau + \hat{Z}_\tau \geq \mathbb{E}[D_\tau], \\
& 0 \leq \hat{Y}_\tau \leq \mathbb{E}[S_\tau], \\
& \psi \hat{Z}_\tau \leq c_{dis}, \\
& -\psi \min\{\hat{Z}_\tau, 0\} \leq c_{char}, \\
& \hat{U}_{\tau+1} - \hat{U}_\tau = -\max\{\hat{Z}_\tau,0\}/\eta_{dis} - \eta_{char}\min\{\hat{Z}_\tau,0\}, \\
& 0 \leq \hat{U}_\tau \leq K, \\
& \hat{X}_\tau \geq 0,
\end{aligned}
\tag{17}
$$

where $\hat{U}_{N+1} = \hat{U}_1$.

Heuristic TBA: Given the state at time t, namely, ($\{t/N\}$, D_t, S_t, C_t, U_t), we determine Z_t so that at time $t+1$, U_{t+1} is as close to $\hat{U}_{\{t/N\}+1}$ by appropriately charging or discharging. The goal is to reach threshold level $\hat{U}_{\{t/N\}+1}$ in the next time. Thus TBA is as follows (with $n = \{t/N\}$):

if $U_t < \hat{U}_{n+1}$, $Z_t = -\min\left\{\dfrac{\hat{U}_{n+1} - U_t}{\eta_{char}}, c_{char}/\psi\right\}$

else if $U_t = \hat{U}_{n+1}$, let $Z_t = 0$,

else $Z_t = \min\left\{\eta_{dis}(U_t - \hat{U}_{n+1}), c_{dis}/\psi\right\}$. Then, $X_t = \max\{D_t - S_t - Z_t, 0\}$ and $Y_t = S_t$.

Heuristic NOA: Given the state at time t, namely, ($\{t/N\}$, D_t, S_t, C_t, U_t), we adopt a naive (but intuitive) strategy – if C_t is cheap, charge the ESD as much as possible; and if D_t is much higher than S_t, discharge as much as possible; otherwise do what the certainty equivalence model suggests. For that we use $\mathbb{E}[C_N]$ as the grand cost (computed over entire cycle N), and $Var[C_N]$ the corresponding grand variance; also ϕ_c and ϕ are parameters to be tuned.[5] It leads to the following NOA (with $n = \{t/N\}$):

if $C_t < \mathbb{E}[C_N] - \phi_c\sqrt{Var[C_N]}$, then

$Z_t = -\min\{(K - U_t)/\eta_{char}, c_{char}/\psi\}$ otherwise,

if $D_t - S_t > \mathbb{E}[D_n] - \mathbb{E}[S_n] + \phi\sqrt{Var[D_n] + Var[S_n]}$,

$Z_t = \min\left\{D_t - S_t - \hat{X}_n, U_t\eta_{dis}, c_{dis}/\psi\right\}$

else

(i.e. $D_t - S_t < \mathbb{E}[D_n] - \mathbb{E}[S_n] + \phi\sqrt{Var[D_n] + Var[S_n]}$)

$Z_t = \min\left\{\max(S_t - D_t, \hat{Z}_n), c_{char}/\psi, (K - U_t)/\eta_{char}\right\}$.

Then $X_t = \max\{D_t - S_t - Z_t, 0\}$ and $Y_t = S_t$.

Heuristic HWR: Before ending this section, we revisit the algorithm proposed in Huang, Walrand and Ramchandran [17], which is referred to as **HWR** in this paper. The algorithm will be numerically tested in the next section. It is a remarkable online algorithm that is based on Lyapunov optimization. The algorithm proposed in [17] does not use any historical information and makes (myopic) decisions based only on current state information (such as D_t, S_t, C_t and U_t).

[5]We note that both parameters ϕ_c and ϕ are non-negative and bounded from the above. Since NOA is computationally simple, one can use the training data to test the performance of NOA under a finite set of feasible parameters ϕ_c and ϕ.

We now briefly outline the algorithm HWR. At each stage $t = 0, 1, \ldots$, the algorithm solves the following LP:

$$
\begin{aligned}
\text{Min}_{(\alpha_t^H,\beta_t^H,\gamma_t^H,\delta_t^H)} \quad & H_t^c\beta_t^H - H_t^s\gamma_t^H + H_t^r\delta_t^H \\
\text{Subject to} \quad & \beta_t^H + \delta_t^H \leq c_{char}, \\
& \gamma_t^H \leq c_{dis}, \\
& \delta_t^H \leq \max(S_t - D_t, 0), \\
& \alpha_t^H + \gamma_t^H = \max(D_t - S_t, 0), \\
& \alpha_t^H \geq 0,\ \ \beta_t^H \geq 0,\ \ \gamma_t^H \geq 0,\ \ \delta_t^H \geq 0,
\end{aligned}
\tag{18}
$$

where $H_t^c = \eta_{char}(U_t - \theta) + C_t/\epsilon$, $H_t^s = (U_t - \theta)/\eta_{dis} + C_t/\epsilon$, and $H_t^r = (U_t - \theta)/\eta_{dis}$, with $\theta = K - \eta_{char}c_{char}/\psi$, and

$\epsilon = \sup_{u\geq 0} C_u / \left[\eta_{char}(\theta - \min(\sup_{u\geq 0} D_u, c_{dis})/(\eta_{dis}\psi))\right]$.

The parameters (H_t^c, H_t^s, H_t^r) are designed to approximate the total operational cost. The relation between the decision variables of HWR and those used in this paper is

$$X_t = \alpha_t^H + \beta_t^H, \quad Z_t = \gamma_t^H - \beta_t^H - \delta_t^H.$$

For $t+1$ the algorithm updates the storage level as follows:

$$\psi U_{t+1} = \psi U_t - \gamma_t^H/\eta_{dis} + \eta_{char}(\beta_t^H + \delta_t^H).$$

It is shown in [17] that HWR achieves asymptotic optimality as the capacity of ESD K grows to infinity (when θ is big and ϵ is small). In the next section we will numerically compare the performance of HWR against the MDP and other heuristic policies in a setting with finite storage capacity.

V. Numerical Experimentation and Results

In Section V-A, we compare the performance and computational time of heuristic algorithms against MDP, under different sizes of discretized state spaces (for demand, solar generation, and cost). Our numerical results show that compared to MDP, the ORA algorithm requires much less computational time and only slightly increases the total cost (by less than 2%). Motivated by the excellent performance of ORA, in Section V-B we benchmark the performance of OLA, HWR, TBA and NOA against ORA under a variety of parameter settings on storage capacity, charging/discharging rates, and average solar generation. In Section V-C, we numerically explore the value of storage and solar generation under the same set of parameter settings considered in Section V-B.

Before representing the numerical results, we would like to briefly discuss the data we obtained and the way we train our algorithms. Our main purpose was to get a representative sample that adequately captures the deterministic and stochastic variability over time. All algorithms are implemented in Matlab R2014a on an Intel Core i7-3740 2.70GHz PC with 16GB memory.

For 5-minute granularity we obtained 26 days of demand, solar generation and cost data in a single month. In that spirit we collected demand data from households (http://www.doc.ic.ac.uk/~dk3810/data/), solar PV supply data from NREL (http://www.nrel.gov/midc/) and 5-minute electricity prices from New England ISO (http://iso-ne.org/). Note that for 5-minute granularity we have $\psi = 12$. We used 16 days of collected data to train the model, i.e., estimate/fit parameters in the MDP model (described in Section III) and the NOA algorithm (described in Section IV). We then use real data in the 10 remaining days (from the original 26 days) for testing. The length of the truncated horizons used in our simulation is long enough to evaluate the steady-state performance of the tested heuristic policies, with a discount factor $\beta = 0.99$.

For 1-hour granularity we obtained 5 years of 3 months' data of demand, solar generation and cost in June, July, and August of 2010-2014. We collected demand and (day-ahead hourly) price data from PJM (http://www.pjm.com/markets-and-operations/energy.aspx). For solar generation, we collected measurements of solar irradiance under the coverage of PJM, for the same 15-month period (http://www.nrel.gov/midc/bsc/). For 1-hour granularity we have $\psi = 1$. We used the first 12 months of collected data to train the model, and use real data in the 3 remaining months for testing.

To estimate d_t and δ_t for any $t \in [1, N]$, we use $D(1,t)$, $D(2,t)$, …, $D(16,t)$, the realized demands in 16 days, to compute $\delta_t = \min_i[D(i,t)]$ and $d_t = \max_i[D(i,t)] - \delta_t$. Likewise for c_t and γ_t. In case of supply s_t, the minimum value is zero. Then for the DTMCs $\{\tilde{W}_t^d\}_{t=1}^{\infty}$, $\{\tilde{W}_t^s\}_{t=1}^{\infty}$, and $\{\tilde{W}_t^c\}_{t=1}^{\infty}$, we first select the number of states $M+1$. The state space of these three DTMCs is a set of discrete values 0, $1/(M-1)$, $2/(M-1)$, …, 1. Then we estimate the elements of $\mathcal{P}^d$, $\mathcal{P}^s$, and $\mathcal{P}^c$ as the respective frequency of transition based on the 16 days' data (for the 5-min case) and the 12 months' data (for the 1-hour case). For all the 1-hour test cases, we use the weekdays' and the weekends' data (of demand and cost) to train two different transition matrices for weekdays and weekends, respectively. The objective is to capture the weekly fluctuation of demand and electricity prices through the constructed MDP model. For all the 1-hour test cases, each algorithm (MDP, ORA, OLA, HWR, TBA, NOA) makes storage operation decisions based on the corresponding transition matrix (of demand and cost) in weekdays and weekends.

For MDP, we consider 13 discretized actions that are uniformly distributed in the continuous action space expressed in Eq. (9). Given the current system state, ORA considers a finite number of possible next system states resulting from a set of 7 actions,[6] and chooses the action that minimizes the approximated cost-to-go. For the data set we use, increasing the number of explored actions (beyond 7) leads to negligible improvement in the performance of ORA. The OLA algorithm takes into account the same set of 7 actions as the ORA algorithm. The other three heuristic polices (HWR, TBA, and NOA), on the other hand, do not discretize the action space. For all the numerical experiments we use $c_{char} = c_{dis}$ and $\eta = \eta_{char} = \eta_{dis}$.

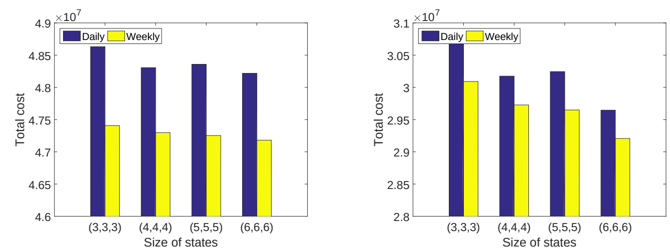
A. Benchmarking Heuristics Against MDP

In this subsection, we will 1) compare the performance of MDP and ORA under the daily model (where a single transition matrix is trained using real data) and the weekly model (where two different transition matrices are trained using weekday's and weekends' data, respectively), and 2) benchmark the performance and computational time of heuristic algorithms against MDP under different sizes of discretized states.

All numerical results presented in this subsection have a 1-hour interval. The two representative parameter settings considered in this subsection are:

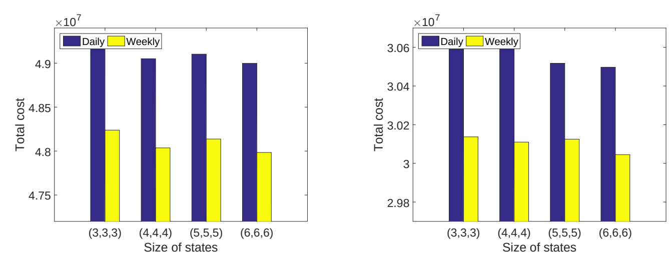
1) $\eta = 1$, $K = 600$ kWh, $c_{char} = c_{dis} = 300$ kW, $K/E[D_t] = 1.0572$, $E[S_t]/E[D_t] = 0.5357$;
2) $\eta = 0.85$, $K = 800$ kWh, $c_{char} = c_{dis} = 100$ kW, $K/E[D_t] = 2.17$, $E[S_t]/E[D_t] = 0.468$.

For MDP, we use 7 discretized storage levels under the first parameter setting, and 9 discretized storage levels under the second parameter setting. (Note that the other heuristic policies do not discretize storage level.)



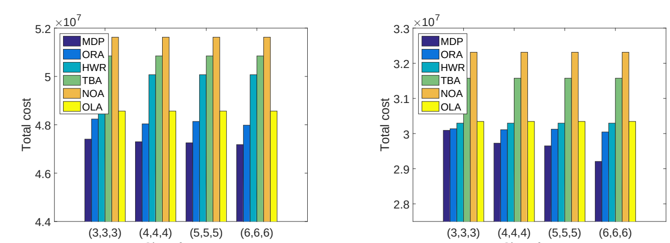
The first parameter setting The second parameter setting

Fig. 2. Total cost resulting from MDP under the daily and weekly models



The first parameter setting The second parameter setting

Fig. 3. Total cost resulting from ORA under the daily and weekly models



The first parameter setting The second parameter setting

Fig. 4. Cost comparison under the weekly model and various sizes of discretized states

In Fig. 2 and 3 we present the total discounted cost resulting from MDP and ORA under the daily model and weekly model. We note from these figures that the incorporation of weekly

[6] Note that the choice of these 7 actions depends on the current system state U_t, D_t and S_t (cf. the discussion at the end of Section III).

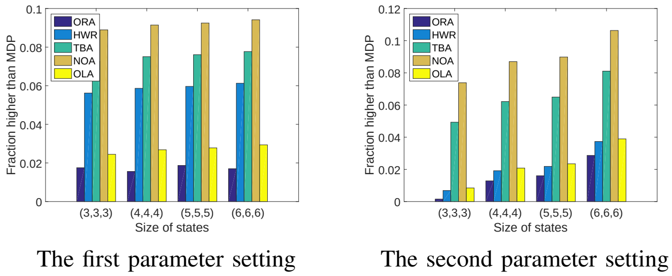


Fig. 5. Cost comparison under the weekly model and various sizes of discretized states

TABLE II
COMPUTATIONAL TIME UNDER THE FIRST PARAMETER SETTING (IN SECOND)

	(3,3,3)	(4,4,4)	(5,5,5)	(6,6,6)
MDP	25.72	345.48	4516.00	21880
ORA	14.46	22.57	45.3	63.08
HWR	2.023	2.023	2.023	2.023
TBA	0.001	0.001	0.001	0.001
NOA	0.0016	0.0016	0.0016	0.0016
OLA	0.069	0.07	0.071	0.073

fluctuation mildly improves the performance of MDP and ORA (by about $1-2.5\%$). We will therefore apply the weekly model for MDP and ORA in all our 1-hour tests throughout the section. Here we do not include the comparison for OLA because incorporating weekly fluctuation leads to negligible performance improvement for OLA.

In Fig. 4 and 5 we compare the total discounted cost resulting from all algorithms under different sizes of the discretized states (of demand, solar generation, and cost). We observe from Fig. 4 that while the performance of MDP is slightly improved as the number of discretized states increases, the performance of other heuristic policies are not sensitive to the size of state space. We observe from Fig. 5 that ORA achieves the best performance: the gap between ORA and MDP is almost always less than 2% (except in the $(6,6,6)$ case under the second parameter setting). OLA is the second best, and results in about $1-4\%$ more cost than MDP. HWR achieves similar performance as OLA under the second parameter setting, but leads to much higher cost than OLA under the first parameter setting. This is in correspondence with the observation in Section V-B: HWR performs well when the number of hours needed to fully charge the storage, K/c_{char}, is big. We have implemented the ORA and OLA algorithms with much larger state space (up to $(40,40,40)$); the performance of both algorithms remains almost the same as the size of state space increases from $(6,6,6)$ to $(40,40,40)$.

TABLE III
COMPUTATIONAL TIME UNDER THE SECOND PARAMETER SETTING (IN SECOND)

	(3,3,3)	(4,4,4)	(5,5,5)	(6,6,6)
MDP	57.76	1376.6	29671	117081
ORA	11.07	21.89	65.72	94.5
HWR	1.834	1.834	1.834	1.834
TBA	0.001	0.001	0.001	0.001
NOA	0.0014	0.0014	0.0014	0.0014
OLA	0.072	0.0908	0.0962	0.1123

In Table II and III we compare the computational time of all algorithms for the scheduling of one-week storage operation. While the size of discretized states significantly (mildly) increases the computational time of MDP (ORA, respectively), it has little influence on the computational time of the other heuristic algorithms. We note that the computational time of MDP is much higher under the second parameter setting, mainly because of the high value of K/c_{char} that leads to more discretized states of storage level. It is also worth noting that ORA takes much less computational time than MDP in all cases, and that somewhat surprisingly, OLA is faster than HWR.

In summary, ORA is comparable to MDP: with problem sizes of $(3,3,3)$ and $(4,4,4)$, it leads to $0.2-2\%$ more total cost than MDP, and is faster than MDP in the $(4,4,4)$ case. OLA leads to $1-4\%$ more total cost than MDP, and requires negligible computational time. HWR achieves similar performance as OLA under the second parameter setting with $K/c_{char}=8$, and leads to significantly higher cost than OLA under the first parameter setting with $K/c_{char}=2$.

B. Benchmarking Heuristics Against ORA

Here we compare the five heuristics ORA, OLA, HWR, TBA and NOA but without MDP. Motivated by the excellent performance of ORA (cf. Section V-A), we compare the other four algorithms against ORA in the next set of experiments. In addition we let storage charging/discharging efficiency $\eta=0.85$, and consider both 1-hour/5-min intervals (corresponding to $N=24$ and $N=288$, respectively). We vary the storage capacity K and scale the average PV supply $\mathbb{E}[S_t]$ based on the average demand $\mathbb{E}[D_t]$. While we will consider variations, we will mainly consider the baseline of: Hours of "average" demand in storage, i.e. $K/\mathbb{E}[D_t]$ as 2.17; Hours to fully charge/discharge, i.e. K/c_{char} as 8; Ratio of average PV supply to average demand, $\mathbb{E}[S_t]/\mathbb{E}[D_t]$, as 0.468.

We first estimate the state transition matrices (of demand, solar generation, and cost) using training data. We tried various alternatives for size of the state space. We chose number of states in demand, renewable generation, and cost Markov chain to be (4,4,4). Incidentally, when we increased the number of states to (10,10,10), the results approximately remain the same.

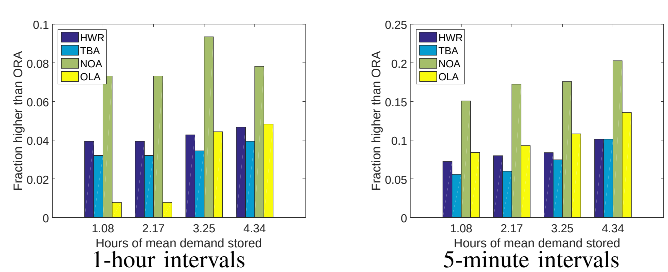
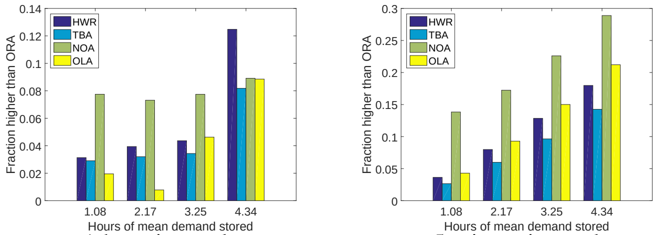


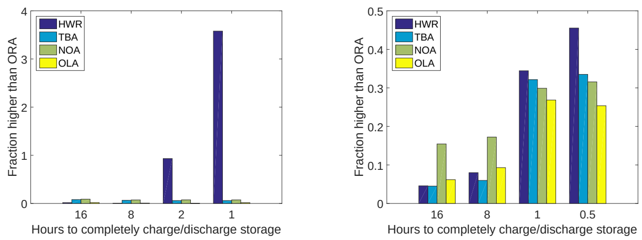
Fig. 6. Heuristics' performance over varying storage capacity (via $K/\mathbb{E}[D_t]$) keeping c_{char} constant

The simulation results are described in Fig. 6-9 where we compare the five heuristic algorithms with the y-axis denoting



1-hour intervals 5-minute intervals

Fig. 7. Heuristics' performance over varying storage capacity (via $K/\mathbb{E}[D_t]$) keeping $K/c_{char} = 8$



1-hour intervals 5-minute intervals

Fig. 8. Heuristics' performance over varying hours to completely charge/discharge storage (via K/c_{char}) keeping K fixed

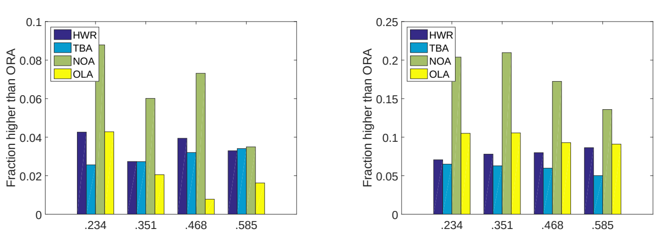
$(b-a)/a$ where a is the minimum total discounted cost (that is obtained by the ORA algorithm) and b is the corresponding heuristic's total discounted cost. While the left side displays correspond to the 1-hour case, the right side figures are based on 5-min real data.

For the testing we let $U_0 = K/2$, i.e., the initial storage level is 50% of storage capacity. For NOA, we selected tolerance parameters $\phi_c = \phi = 0.25$ by testing several options. Interestingly the 0.25 value is robust and the solutions do not change with much higher or lower values of ϕ. We observe from these four figures that for most all 1-hour cases, the ADP-based OLA algorithm performs slightly worse than ORA, and yields the minimum total discounted cost among the four heuristics. In many 5-minute cases, however, OLA algorithm performs worse than some other heuristics (e.g., TBA). This is intuitive since OLA always treats the next stage (the next five minutes in 5-minute case) as the terminal stage and completely ignores the system dynamics after the next stage.

In Fig. 6, we fix the maximum charging rate c_{char} and vary the storage capacity K. Note that the ratio K/c_{char} is 4, 8, 16, and 32 hours for the four cases, respectively. We observe from Fig. 6 that HWR performs reasonably well in all cases, and yields $2\% - 13\%$ more cost than ORA. Also, we observe that TBA achieves the minimum cost (among the four heuristics) in the 5-minute case. In Fig. 7, we fix the ratio $K/c_{char} = 8$ and vary the storage capacity K. This is a more practical setting since the maximum charging rate usually grows (nearly) proportionally to the storage capacity. The performance gap between ORA and the four heuristics increases with storage capacity.

In Fig. 8 we fix the storage capacity K and vary the capacity to charging rate ratio K/c_{char}. The parameter setting in our simulation is motivated by the development of fast-charging batteries [31].[7] We observe from Fig. 8 that the performance of HWR heavily depends on the ratio K/c_{char}: the performance gap between HWR and ORA is mild when this ratio is larger than 8 (i.e., it takes more than 8 hours to fully charge the storage); however, for fast-response storage devices with $K/c_{char} \leq 2$, both ORA and OLA significantly outperform HWR.

Finally, in Fig. 9, we vary the ratio of average solar supply to average demand ($\mathbb{E}[S_t]/\mathbb{E}[D_t]$) while fixing the other parameters. The parameter setting is motivated by the fast growing installment of solar panels on the consumer side. We note that as the solar penetration increases, ORA still outperforms the other four heuristics (including OLA), especially in the 5-minute case.

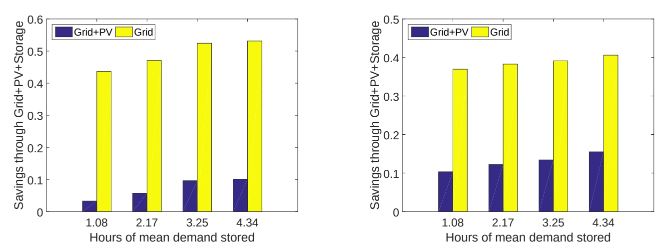


1-hour intervals 5-minute intervals

Fig. 9. Heuristics' performance over varying ratio of average solar supply to average demand (via $\mathbb{E}[S_t]/\mathbb{E}[D_t]$)

C. The Value of Storage and PV

There are costs to install a solar PV system and/or an ESD. A natural question to ask is whether the PV and/or ESD installation was worth it. For that we consider two parameters: value of storage and value of PV and storage. We use the same test data as the previous sub-section and policy based on ORA. In Fig. 10-13, the y-axis denotes $(a-b)/a$ where b is the total discounted cost using both PV and storage, while a in the left bars correspond to the (optimal) use of only PV,[8] and a in the right bars correspond to the use of neither PV nor storage.



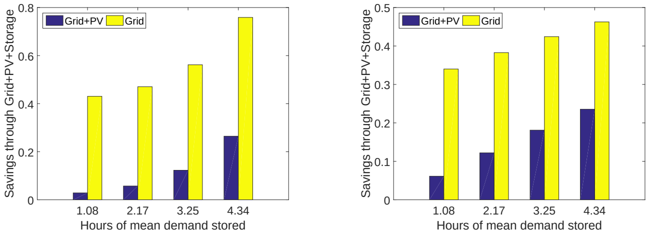
1-hour intervals 5-minute intervals

Fig. 10. Value of storage/PV over varying storage capacity (via $K/\mathbb{E}[D_t]$) keeping c_{char} constant

The left bars present the fractional cost savings due the operation of storage, and can be therefore viewed as an illustrator on the value of storage. Similarly, the right bars illustrates the value of storage and PV. We observe from Fig.

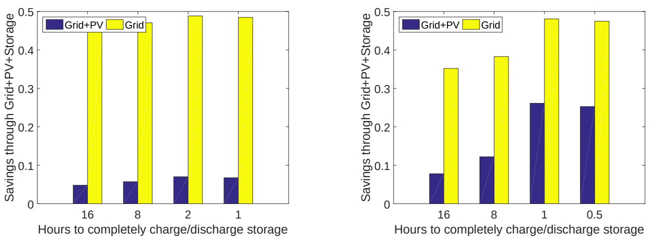
[7] For example, the lithium-ion titanate batteries are capable of recharging to 95% of full capacity within approximately ten minutes [31].

[8] We note that without storage, the optimal operation of PV is trivial: simply use as much solar generation as possible to fulfill the current demand.



1-hour intervals 5-minute intervals

Fig. 11. Value of storage/PV over varying storage capacity (via $K/\mathbb{E}[D_t]$) keeping $K/c_{char} = 8$



1-hour intervals 5-minute intervals

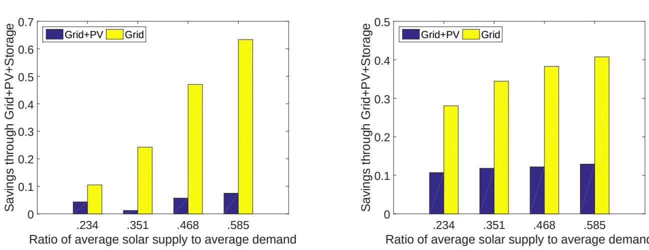
Fig. 12. Value of storage/PV over varying hours to completely charge/discharge storage (via K/c_{char}) keeping K fixed

10-13 that the value of storage is much higher in 5-minute cases, due to the higher variability in costs under 5-minute real-time pricing than that under hourly pricing.

As shown in Fig. 10, the values of storage and PV do not increase appreciably with increase in storage size (without increasing rates of charging/discharging). We observe from Fig. 11 that the value of storage increases sharply with the storage capacity K, when the maximum charging rate c_{char} grows in proportional to K. We observe from Fig. 12 that the values of storage and PV increase with the maximum charging/discharging rate of the storage. Fig. 13 shows that the values of storage and PV increase with average solar PV generation.

VI. Conclusion

Although deceptively easy to state, the problem of determining energy mix from the grid, renewable source and storage device is fairly complex to solve. We implemented six policies MDP, ORA, OLA, HWR, TBA and NOA, and compared their performance using real data of energy demand, renewable generation, and electricity prices.



1-hour intervals 5-minute intervals

Fig. 13. Value of storage/PV over varying ratio of average solar supply to average demand

The following were our findings.

1) ORA outperforms the other four heuristics in all cases, and at the same time, requires the most computing power among the five heuristics. For 1-hour cases, ORA results in $0.2-2\%$ more total cost than MDP, and OLA leads to $1-4\%$ more total cost than MDP. For many 5-minute cases, however, ORA significantly outperforms OLA; this is intuitive, since OLA always treats the next stage (the next five minutes in this case) as the terminal stage and completely ignores the system dynamics after the next five minutes. TBA performs well in some 5-minute cases. Tuning tolerance and coefficient parameters had virtually no effect on NOA.
2) HWR is an easily implementable algorithm that needs no training. Its performance to a large extent depends on the number of hours to fully charge storage (i.e. K/c_{char}). It achieves almost the same total discounted cost as ORA when K/c_{char} is large (e.g. > 8). On the other hand, the two ADP-based algorithms, ORA and OLA, significantly outperform HWR for the case with $K/c_{char} \leq 2$.
3) Under the one-step roll-out algorithm (ORA) and hourly pricing, value of storage is not too high with $K/c_{char} = 8$. Value of storage is much higher in 5-minute cases, because the cost is much more fluctuating under 5-minute real-time pricing than that under hourly pricing. Value of storage would improve greatly if the storage size increases along with speed of charging and discharging. As solar penetration goes higher, storage has more value.

References

[1] R. Urgaonkar, B. Urgaonkar, M. J. Neely, and A. Sivasubramaniam, "Optimal power cost management using stored energy in data centers," *Proc. of the ACM SIGMETRICS*, 2011.

[2] S. B. Peterson, J. F. Whitacre, and J. Apt, "The economics of using plug-in hybrid electric vehicle battery packs for grid storage," *Journal of Power Sources*, vol. 195, no. 8, pp. 2377–2384, 2010.

[3] R. Sioshansi, "Welfare impacts of electricity storage and the implications of ownership structure," *The Energy Journal*, vol. 31, pp. 173–198, 2010.

[4] Z. Liu, Y. Chen, C. Bash, A. Wierman, D. Gmach, Z. Wang, M. Marwah, and C. Hyse, "Renewable and cooling aware workload management for sustainable data centers," *ACM SIGMETRICS Performance Evaluation Review*, vol. 40, no. 1, 2012.

[5] D. Noll, C. Dawes, and V. Rai, "Solar community organizations and active peer effects in the adoption of residential pv," *Energy Policy*, vol. 67, no. 2, pp. 330–343, 2014.

[6] J. Garcia-Gonzalez, R. D. la Muela, L. Santos, and A. Gonzalez, "Stochastic joint optimization of wind generation and pumped-storage units in an electricity market," *IEEE Tran. Power Systems*, vol. 23, pp. 460–468, 2008.

[7] L. Xie, Y. Gu, A. Eskandari, and M. Ehsani, "Fast mpc-based coordination of wind power and battery energy storage systems," *Journal of Energy Engineering*, vol. 138, no. 2, pp. 43–53, 2012.

[8] S. Bose and E. Bitar, "Variability and the locational marginal value of energy storage," in *IEEE Conference on Decision and Control*, 2014.

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

[10] J. Qin and R. Rajagopal, "Dynamic programming solution to distributed storage operation and design," in *IEEE PES General Meeting*, 2013.

[11] N. Gast, D. C. Tomozei, and J. Y. L. Boudec, "Optimal generation and storage scheduling in the presence of renewable forecast uncertainties," *IEEE Tran. Smart Grid*, vol. 5, no. 3, pp. 1328–1339, 2014.

[12] F. Graves, T. Jenkins, and D. Murphy, "Opportunities for electricity storage in deregulating markets," *The Electricity Journal*, vol. 12, no. 8, pp. 46–56, 1999.

[13] P. Mokrian and M. Stephen, "A stochastic programming framework for the valuation of electricity storage," *Proc. of 26th USAEE/IAEE North American Conference*, 2006.

[14] J. Qin, R. Sevlian, D. Varodayan, and R. Rajagopal, "Optimal electric energy storage operation," *Proc. of 2012 PES General Meeting*, 2012.

[15] I. Atzeni, L. G. Ordonez, G. Scutari, D. P. Palomar, and J. R. Fonollosa, "Demand-side management via distributed energy generation and storage optimization," *IEEE Tran. Smart Grid*, vol. 4, no. 2, pp. 866–876, 2013.

[16] I. Koutsopoulos, V. Hatzi, and L. Tassiulas, "Optimal energy storage control policies for the smart power grid," *Proc. of IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 2011.

[17] L. Huang, J. Walrand, and K. Ramchandran, "Optimal demand response with energy storage management," in *IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 2012, pp. 61–66.

[18] S. Lakshminarayana, T. Q. Quek, and H. Poor, "Cooperation and storage tradeoffs in power grids with renewable energy resources," *IEEE J. on Selected Areas in Communications*, vol. 32, no. 7, pp. 1386–1397, 2014.

[19] J. Qin, Y. Chow, J. Yang, and R. Rajagopal, "Online modified greedy algorithm for storage control under uncertainty." [Online]. Available: http://arxiv.org/abs/1405.7789

[20] P. van de Ven, N. Hegde, L. Massoulié, and T. Salonidis, "Optimal control of end-user energy storage," *IEEE Tran. Smart Grid*, vol. 4, no. 2, pp. 789–797, 2013.

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

[22] Y. Xu and L. Tong, "On the operation and value of storage in consumer demand response," *Proc. of 53th IEEE Conference on Decision and Control (CDC)*, 2014.

[23] R. Wilson, "Architecture of power markets," *Econometrica*, vol. 40, no. 4, pp. 1299–1340, 2002.

[24] K. C. Divya and J. Østergaard, "Battery energy storage technology for power systems - an overview," *Electric Power Systems Research*, vol. 7, no. 4, pp. 511–520, 2009.

[25] C. Ren, D. Wang, B. Urgaonkar, and A. Sivasubramaniam, "Carbon-aware energy capacity planning for datacenters," in *IEEE 20th International Symposium on Modeling, Analysis & Simulation of Computer and Telecommunication Systems (MASCOTS)*, 2012, pp. 391–400.

[26] D. P. de Farias and B. V. Roy, "The linear programming approach to approximate dynamic programming," *Operations Research*, vol. 51, no. 6, pp. 850–865, 2003.

[27] S. M. Ross, *Introduction to Stochastic Dynamic Programming*. Academic Press, 1983.

[28] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Athena Scientific, 2005, vol. I.

[29] D. P. Joseph and T. J. Tou, "On linear control theory," *AIEE Transactions: Applications and Industry*, vol. 80, no. 4, pp. 193–196, 1961.

[30] L. Xie and M. D. Ilic, "Model predictive economic/environmental dispatch of power systems with intermittent resources," in *IEEE Power and Energy Society General Meeting*, 2009.

[31] M. Etezadi-Amoli, K. Choma, and J. Stefani, "Rapid-charge electric-vehicle stations," *IEEE Trans. Power Delivery*, vol. 25, no. 3, pp. 1883–1887, 2010.