Lossy Energy Storage to Cut Power Costs

Mohammad Rezaeian and Lachlan L. H. Andrew
Faculty of Information Technology, Monash University, Australia
mohammad.rezaeian@gmail.com, lachlan.andrew@monash.edu

***Abstract*—Under dynamic pricing, organizations can reduce their electricity cost by varying their load with price. When the end load is inflexible, storage can match the energy draw from the grid to the price. In this paper we present a general model for power cost management by shifting demand for high power away from high electricity price periods through imperfect energy storages. Using Lyapunov optimization, we present a policy that achieves a solution with bounded suboptimality. We demonstrate that the sampling rate must increase as either the battery capacity decreases or the charge rate increases.**

I. Introduction

Smart meters will allow dynamic electricity prices, to reduce peak power demand. In this context, users will seek to minimze their electricity bills by using local storage, such as electric vehicle (EV) batteries or uninterruptible power supplies (UPSs), to smooth their demand.

The optimal pattern of charging and discharging depends on future values both of the price of electricity bought from the grid and of the customer's demand. Since these values are not known, it is of interest to study charging policies guaranteed to work well with minimal assumptions about the future. Urgaonkar et al. [1] use Lyapunov optimization to derive a policy whose performance is at most a constant worse than the optimal charging given perfect information, but which makes no stochastic assumptions about inputs, needing only upper and lower bounds. Our paper extends [1] in several ways.

First, in Section II, we consider a more accurate model of storage. Storage devices lose energy both during (dis)charging and, gradually, simply when holding charge. We model both effects. Most work on storage management ignores this effect, although [2] has considered a general characterization of optimal policy with imperfect storages when perfect information is available. In addition, our results apply to a more general model of the wear-and-tear costs of charging and discharging. This allows us in Section III to derive an algorithm with performance guarantees analogous to those of [1].

Second, in Section IV we consider the impact of the rate of updating the charge/discharge rate. We show that updating the rate more frequently results in bounds that can be orders of magnitude tighter than those of [1]. The improvement is greatest when the storage capacity is small relative to its peak power rating, which is the case for UPSs.

When the control action is updated frequently, temporal correlation in the inputs becomes important. We provide a brief discussion of these issues at the end of the paper.

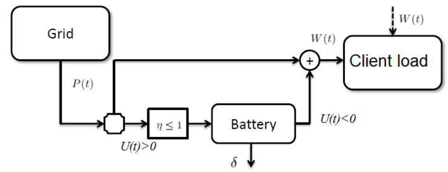

Figure 1. System model

II. System Model

We consider a discrete time model, with time slots of duration T indexed by k. Our task is an infinite horizon average cost control problem of a dynamical system whose state at time k is the battery charge, denoted $Y(k)$. Control is subject to the stability criterion that the state be bounded by $Y_{\min} \leq Y(k) \leq Y_{\max}$, for all k.

There are two exogenous stochastic inputs. The first is the (inelastic) power consumption of the customer at continuous time t, denoted $\bar{W}(t)$. For a given slot duration T, the discrete time power consumption is $W(k) = \bar{W}(kT)$, so that energy $TW(k)$ is consumed in slot k. The other input is the time varying electricity price, denoted by $\bar{S}(t) = S(kT)$. We assume that $\bar{W}(\cdot)$ and $\bar{S}(\cdot)$ are independent and stationary. Initially, we assume that the sampling time T is large enough that samples $W(k)$, respectively $S(k)$, are i.i.d..

A. System dynamics

The decision variable (the controlled input to the system) is the power that is bought from the grid at any given time, denoted by $P(k)$. The battery receives power $U(k) = P(k) - W(k)$, of which some is stored and some is lost.

We consider two types of energy losses: self discharging at a fixed power δ whenever the battery is non-empty, and the inefficiency η of charging and discharging. If $Y_{\min} > 0$, then the battery is never empty, and the state of charge of the battery, $Y(k)$, evolves as

$$Y(k+1) = Y(k) + Tf_\eta(P(k) - W(k)) - T\delta \qquad (1)$$

where

$$f_\eta(x) = \begin{cases} \eta x, & x > 0 \\ x, & x \leq 0 \end{cases}$$

for $0 < \eta \leq 1$ which is the efficiency of charging system. We assume $Y_{\min} \leq Y(0) \leq Y_{\max}$ and we wish to keep $Y(k)$ in this range for all time. In this model, for every unit of energy that is stored, only η can be recovered and at every time slot the battery consumes energy δ as self discharge. The controlled system with these inefficiencies is depicted in Figure 1.

978-1-4799-3088-3/14/$31.00 ©2014 IEEE

Let $R_{\max}$ and $D_{\max}$ be the maximum physically possible charge and discharge rates for the battery, and P_{peak} be the maximum power that can be drawn from the grid. Numerical results (Fig. 3 of Section III-A2) suggest that Lyapunov optimization subject only to these constraints causes excessive swings in state of charge, and so we constrain the charge and discharge to be at most $R_m \le R_{\max}$ and $D_m \le D_{\max}$. The control actions $P(k)$ for any given state must satisfy

$$\begin{aligned} 0 \le P(k) &\le P_{\text{peak}} \\ W(k) - D_m + \delta \le P(k) &\le W(k) + \tfrac{R_m+\delta}{\eta} \\ W(k) + \delta - \tfrac{Y(k)-Y_{\min}}{T} \le P(k) &\le W(k) + \tfrac{Y_{\max}-Y(k)+T\delta}{T\eta}. \end{aligned} \tag{2}$$

We assume $P_{\text{peak}} \ge W_{\max} + \delta/\eta$, hence $P(k) = W(k) + \delta/\eta$ (no change in Y) is always a feasible action. Constraint (2) guarantees $Y_{\min} \le Y(k) \le Y_{\max}$ for all k. Note $R_m, D_m, R_{max}, D_{max}$ are powers, not energy per slot; this matters when we scale T.

B. Price model

The electricity price per kilowatt-hour (kWh) depends on global demand and availability of renewable energy. Utilities may also provide disincentives to high power draw to address their capacity for peak power; a natural approach is changing the price rate based on the power drawn. So in general the price is considered as a function of a stochastic process reflecting availability of renewable energy plus demand, and the total power drawn by the customer.

We measure the cost, denoted Z_1, per unit continuous time, to make things clearer when T changes. The cost in slot k is

$$TZ_1(k) = TP(k)\hat{C}(S(k), P(k)) + h(U(k)) \tag{3}$$

where $\hat{C}(S(k), P(k))$ is the user's electricity price when purchasing power $P(k)$ while the system state (including time-of-day, other users' loads and availability of renewables) is $S(k)$. Hence the electricity cost can be in general non-linear with purchased power and randomly time varying. We denote by $C_{\min}$ and $C_{\max}$ the minimum and maximum of price $\hat{C}(S, P)$ over all S and $0 \le P \le P_{\text{peak}}$. In (3), $h(.)$ is the cost due to reduction in battery life time due to charge/discharge, which is assumed to be positive everywhere except $h(0) = 0$.

The aim of energy management is to find the control actions $P(k)$ based on the state of system that minimizes the infinite horizon average cost under constraint (2), i.e., the solution to

$$\mathbf{P_1} : \text{Minimize} \lim_{k\to\infty} \frac{1}{k} \sum_{\tau=0}^{k-1} \mathsf{E}[Z_1(\tau)]$$

$$\text{Subject to (2) for all } k.$$

We denote the optimal value of this optimization by $\hat{\Phi}_1$. Note that this may be larger than the optimum achievable if (dis)charge rates $R_{\max}$ and $D_{\max}$ replace R_m and D_m.

The optimal policy for $\mathbf{P_1}$ needs to consider the fact that $P(k)$ influences the state of system after k. Finding the exact optimum requires dynamic programming which requires the statistics of $W(\cdot)$ and $S(\cdot)$ and is computationally prohibitive. Instead, we use the Lyapunov optimization to derive a computationally efficient policy of bounded suboptimality, in terms only of $\hat{C}$ and bounds on P and W.

Lyapunov optimization

Lyapunov optimization is a method for obtaining a greedy policy for on-line adjusting the controlled input to a system to minimize the average cost while imposing a form of stability. To minimize the average of a function f, at each time step it chooses the control to minimize the sum of the instantaneous value of f and the expected drift of a Lyapunov function, $E[dL/dt]$. The Lyapunov function $L(.)$ measures how far the state is from the desirable (stable) states. Since stability for our problem is the state $Y(k)$ remaining between $Y_{\min}$ and $Y_{\max}$, we use a quadratic Lyapunov function

$$L(Y(k)) = a_T(Y(k) - d_T)^2/2 \tag{4}$$

with $Y_{\min} < d_T < Y_{\max}$. The gain a_T, usually denoted $1/V$, typically enables the trade off between a soft stability constraint, such as low queue occupancy, and minimizing f. Since our setting has hard stability constraints on the battery state of charge, there is no tradeoff possible, and we follow [1] by setting a_T and d_T to yield the minimum optimality gap while enforcing stability for any sample path.

III. The near optimal policy

The policy that we describe next requires subtracting an incentive α from the price. We call $\hat{C}(S, P) - \alpha$ a rebated price. We define $\psi(S)$ as the minimum value of $\alpha > 0$ such that, for all $0 \le P_1 \le P_2 \le P_{\text{peak}}$,

$$P_1(\hat{C}(S, P_1) - \alpha) \ge P_2(\hat{C}(S, P_2) - \alpha).$$

Accordingly, for any discount greater or equal to $\psi(S)$, the compound cost function $(\hat{C}(S, P(k)) - \psi(S))P(k)$ becomes decreasing with purchased power; i.e., $\hat{C}(S, .) - \psi(S)$ becomes a reward instead of price. If for all S, $\hat{C}(S, P)$ is continuous and increasing in P, then

$$\psi(S) = \hat{C}(S, P_{\text{peak}}) + P_{\text{peak}}\hat{C}'(S, P_{\text{peak}})$$

We denote by $\Omega_{\max}$ the minimum of $\psi(S)$ over all S (i.e., minus the maximum of price $-\psi(S)$ that if added to the price makes the compound price a reward), noting that for the case where $\hat{C}$ is only a function of S, $\Omega_{\max} = C_{\max}$, otherwise $\Omega_{\max} \ge C_{\max}$. We also denote $\Omega_{\min} = C_{\min}$ which is the maximum rebate that we can give and the compound cost function will remain increasing with $P(k)$.

Policy $\mathbf{P_1^g}$ generates $P(k)$ as the solution to

$$\mathbf{P_1^g} : \min_{P(k)} Z_1(k) + a_T(Y(k) - d_T)f_\eta(P(k) - W(k))$$

$$\text{Subject to:} \quad \begin{aligned} 0 \le \; & P(k) \le P_{\text{peak}} \\ W(k) + \delta - D_m \le \; & P(k) \le W(k) + \tfrac{R_m+\delta}{\eta}, \end{aligned} \tag{5}$$

where constants a_T and d_T are

$$a_T = \frac{\Omega_{\max} - \Omega_{\min}/\eta}{Y_{\max} - Y_{\min} - TR_m - T(D_m + \delta)}, \tag{6}$$

$$d_T = Y_{\min} + T(D_m + \delta) + \Omega_{\max}/a_T. \tag{7}$$

The above policy looks at load and price, $W(k), S(k)$ and the battery charge $Y(k)$ and solves $\mathbf{P_1^g}$ to find $P(k)$ for time k.

Note that $a_T(Y(t) - d_T)$ is a linear mapping from $Y(k) \in [Y_{\min} + T(D_m + \delta), Y_{\max} - TR_m]$ to a rebate in $[\Omega_{\min}, \Omega_{\max}]$. The policy adds this battery related rebate (or η times of it, if charging) to the electricity price $\hat{C}(S, P)$ and greedily minimizes the rebated cost. When the battery state of charge is low, the high rebate $\Omega_{\max}$ gives high incentive for charging the battery. This incentive linearly decreases as the battery state of charge increases.

We now show that the power management $P(k)$ obtained by this policy is feasible for $\mathbf{P_1}$ ($Y_{\min} \leq Y(k) \leq Y_{\max}$ for all k) and has time average cost that is at most a proven constant away from the time average cost of the best possible policy for $\mathbf{P_1}$.

Theorem 1. *For $\eta > \Omega_{\min}/\Omega_{\max}$ and a given $R_m \leq R_{max}, D_m \leq D_{max}$, the greedy algorithm $\mathbf{P_1^g}$ finds a near optimal power solution $P(t)$ which is feasible for $\mathbf{P_1}$ and its average cost is bounded by*

$$\lim_{t\to\infty} \frac{1}{t} \sum_{\tau=0}^{t-1} \mathsf{E}[Z_1(\tau)] \leq \hat{\Phi}_1 + a_T BT \tag{8}$$

where

$$B = \frac{\max\{R_m^2, D_m^2\}}{2}.$$

Note that $\Omega_{\max}$ is no less than the maximum electricity price, hence $\eta > \Omega_{\min}/\Omega_{\max}$ is not usually a restrictive conditions for applications with significant price fluctuations.

As we see in the proof of the theorem, the proof of performance bound only relies on the non-negativity of a_T. The value of a_T in (6) is the minimum value required to guarantee the feasibility of the solution (stability). By increasing this parameter the policy $\mathbf{P_1^g}$ still provides feasible solutions by emphasizing more on stability rather than cost reduction. In particular for $a_T \to \infty$, $\mathbf{P_1^g}$ enforces the solution of $P(k) = W(k)$ which is the same as disconnecting the energy storage. In this case the Lyapunov optimization theory does not give any bound on the performance of an optimal algorithm for using the energy storage.

The energy gap TR_m and $T(D_m+\delta)$ in the denominator of a_T are "guard intervals" from the limits of stability to make sure that in the safe region between $Y_{min} + T(D_m + \delta)$ to $Y_{max} - TR_m$ next state will not pass the stability region and if it passes this guard band, considering Lemma 2, the next state falls back into this safe region. This is the approach that is taken in the proof of feasibility part of Theorem 1 and finding the optimal values a_T, d_T in (6), (7).

A. Discussion of the policy $\mathbf{P_1^g}$

Policy $\mathbf{P_1^g}$ and Theorem 1 are extensions of results of [1] in which it is assumed

1) perfect storages ($\eta = 1, \delta = 0$)
2) $R_m = R_{max}, D_m = D_{max}$.
3) Wear and tear cost $h(U)$ is constant for $U \neq 0$.

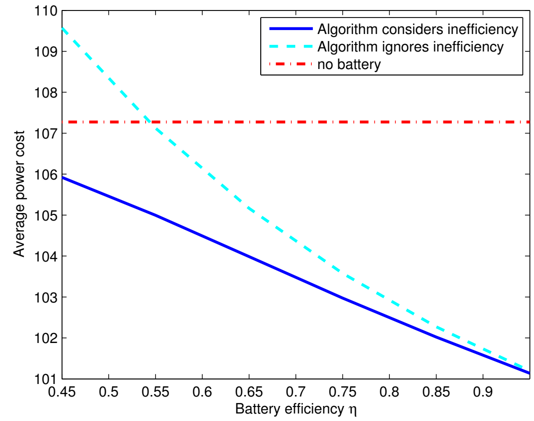

Figure 2. Average power cost reduction as compared to not using any storage for various battery efficiency when the inefficiency is considered and not considered in the algorithm

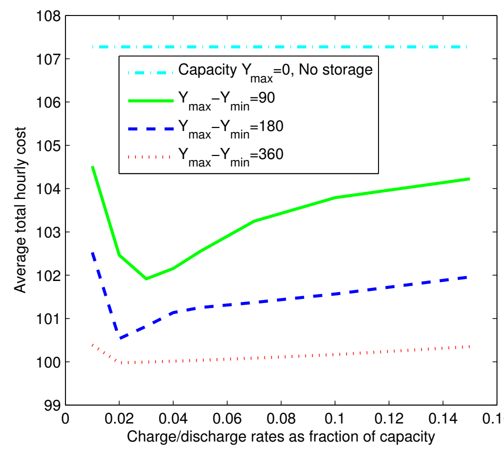

Figure 3. Cost as a function of rate of charge/discharge. The optimal rate is much less than the actual battery maximum charge/discharge rate.

Here we have considered the effect of imperfections and sampling time, with more general formulations towards improvement of the policy for practical purposes. Specialization to $\eta = 1, \delta = 0$ and $R_m = R_{max}, D_m = D_{max}$ recovers the result in [1].

1) Cost reduction benefit: Figure 2 compares the energy use of policy $\mathbf{P_1^g}$ with that of both using no storage and using the algorithm of [1] which assumes ideal storage. Here $W(\cdot)$ is the sum of a square wave of period $50T$ and a sinusoid of one third the amplitude and period $6T$, all scaled to $[100, 1000]$ kW; $\hat{C}(S, P) = (S + P)/200$; $S(.)$ has the same form as W, scaled to $[0.5, 3.5]\,(\text{kWh})^{-1}$; $Y_{\max} - Y_{\min} = 180$ kWh; $R_m = D_m = 30$ kW and $P_{\text{peak}} = 1$ MW. Note that for low efficiency energy storage devices (such as flywheels) it is necessary that the algorithm consider the inefficiency in order to obtain any benefit at all from using storage.

2) Charge rates: It is common for algorithms with performance bounds to perform on average slightly worse than those with weaker performance bounds. A related effect occurs here, with an interesting twist. The strongest bound on the performance occurs when the right hand side of (8) is minimized. Let us consider how that bound varies with the maximum charging rates we allow, R_m and D_m. The first term, $\hat{\Phi}_1$, is non-increasing in R_m and D_m, since it is the minimum value over a constraint set that is non-decreasing in R_m and D_m. However the second term, $a_T B n_c T$, increases rapidly for large R_m and D_m, and so the right hand side has a minimum for values of R_m and D_m that may be substantially smaller than the physical constraints $R_{\max}$ and $D_{\max}$. This is because policy $\mathbf{P_1^g}$ tends to respond to small changes in battery level by charging or discharging almost as fast as possible.

However, it is not possible to minimize the right hand side of (8) with respect to R_m and D_m, since $\hat{\Phi}_1$ is not known explicitly. The standard surrogate, used in [1], is to minimize the bound on the suboptimality with respect to $\hat{\Phi}_1(R_{\max}, D_{\max})$. Unless we can bound $\hat{\Phi}_1(R_m, D_m)$ in terms of $\hat{\Phi}_1(R_{\max}, D_{\max})$, the only way to obtain an absolute bound is to use the suboptimal values $R_m = R_{\max}$ and $D_m = D_{\max}$.

In practice to have the best performance of our policy, we need to tune rates R_m, D_m for overall minimum cost of the policy, or as a surrogate the bound $\Phi + a_T BT$. The rates R_m, D_m remain as the design parameters for the policy and for some applications the assumption $R_m = R_{max}, D_m = D_{max}$ in [1] is not a viable assumption.

Figure 3 shows the cost versus $R_m = D_m$ for three values of $Y_{\max}$ in the simulation mentioned before. The performance drop for high R_m, D_m is due to excessive feedback from battery in the linear mapping $a_T(Y(t) - d_T)$. High feedback gain makes battery state of charge have a huge cost implication in the objective function, overshadowing the actual electricity cost.

3) Large storage: It was noted in [1] that Theorem 1 shows that if R_m and D_m are fixed then taking $Y_{\max} - Y_{\min}$ to infinitely makes $\mathbf{P_1^g}$ optimal, despite no future knowledge. This is because the optimal policy in the absence of battery constraints is the sum of the greedily minimizer $Z_1(k)$ and a constant to induce a "drift" towards charging. For large $Y_{\max} - Y_{\min}$, the scalar $(Y(k) - Y_{\min})/(Y_{\max} - Y_{\min})$ tends to a constant far from the boundaries 0 and 1 meaning that short-term fluctuations almost never cause the capacity constraints to affect the policy. That constant encodes the optimal charging drift.

4) Possible extension: Self-discharge is more accurately modeled as being proportional to the state of charge [3]. By incorporating the self discharging into load and indicating the combined one by $\hat{W}(t)$, because both $W(t)$ and $Y(t)$ are observed at any time t, we can devise a policy based on the stochastic input $\hat{W}(k)$ in $\mathbf{P_1^g}$ which will be a feasible policy for $\mathbf{P_1}$, but the performance bound is not guaranteed. Preliminary research shows that the design of a policy with performance bound in this case requires a different proof method or non-Lyapunov optimization methods due to the multiplicative factor $(1 - \delta)$ appearing in the state recursion $Y(t+1) = (1-\delta)Y(k) + P(k) - W(k)$ (for $\eta = 1$) which is not conventional for Lyapunov optimizations.

B. Discussions towards the proof of Theorem 1

Here we consider a randomized stationary control policy as a solution for a relaxed version of problem $\mathbf{P_1}$ which we call $\bar{\mathbf{P}}$. The randomized policy may not be feasible for $\mathbf{P_1}$ due to a relaxation of the constraints, but we use the existence of such a randomized policy in the proof of Theorem 1

Since $Z_1(k)$ depends on time only through variables $P(k), S(k)$ and $W(k)$, we can write a generic cost function as $Z(P, S, W) = P\hat{C}(S, P) + h(P - W)/T$ and derive an optimal randomized stationary policy. This policy is a set of conditional probabilities $q(.|S, W)$ for any value of S, W.

The **randomized stationary policy** generates $P(k) = P$ randomly with the probability $q(P|S(k), W(k))$ that solves the optimization

$$\bar{\mathbf{P}} : \min_{q(.|S,W)} \mathsf{E}[Z(P, S, W)]$$

Subject to
$$\begin{array}{rl} 0 \leq & P \leq P_{\text{peak}} \\ W + \delta - D_{\max} \leq & P \leq W + \frac{R_{\max}+\delta}{\eta} \\ \mathsf{E}[f_\eta(P - W)] & = \delta \\ q(P|S, W) \geq 0, & \sum_P q(P|S, W) = 1 \end{array} \tag{9}$$

where

$$\mathsf{E}[Z(P, S, W)] = \sum_{S,W} \sum_P Pr(S, W) q(P|S, W) Z(P, S, W)$$

$$\mathsf{E}[f_\eta(P - W)] = \sum_{S,W} \sum_P Pr(S, W) q(P|S, W) f_\eta(P - W).$$

Since $W(k)$ and $S(k)$ are independent and each one is i.i.d over time, we have fixed distributions for $W(k)$ and $S(k)$ independent of time and $Pr(S, W) = Pr(S)Pr(W)$. Knowing the two distributions $Pr(S), Pr(W)$, we can derive the optimal randomized policy $q(.|S, W)$. The problem is feasible as there exists at least one feasible point, i.e., $q(P = W + \delta/\eta|S, W) = 1, q(P \neq W + \delta/\eta|S, W) = 0$. Hence, there always exists an optimal stationary policy for $\bar{\mathbf{P}}$.

Now consider the following optimization which as compared to $\mathbf{P_1}$ does not have $Y(k)$ to carry history and create time dependency.

$$\bar{\mathbf{P}}_\mathbf{1} : \text{Minimize} \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathsf{E}[Z_1(k)]$$

Subject to (5) and $\overline{P - W}^\eta = \delta$, where

$$\overline{P - W}^\eta \equiv \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathsf{E}[f_\eta(P(k) - W(k))].$$

This problem only depends on $S(k)$ and $W(k)$ which have statistical description independent of time. Due to time independence, the time average of the expectation of the cost and f_η are the same as the expectation of the generic (time independent) cost and f_η in $\bar{\mathbf{P}}$. Therefore, this problem is equivalent to $\bar{\mathbf{P}}$ and they both have the same optimal value which we denote by $\bar{\Phi}$. Moreover, $\bar{\mathbf{P}}_\mathbf{1}$ is a relaxed version of

$\mathbf{P}_1$ and hence $\bar{\Phi} \leq \hat{\Phi}$ because both have the same objective function and any feasible point of $\mathbf{P_1}$ is a feasible point for $\bar{\mathbf{P}}_1$. In fact any policy satisfying (the third constraint in) (2) would yield $Y_{\min} \leq Y(k) \leq Y_{\max}$ and according to (1),

$$\mathsf{E}[Y(k) - Y(0)] = -n\delta + \sum_{k=0}^{n-1} \mathsf{E}[f_\eta(P(k) - W(k))]$$

and hence,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathsf{E}[f_\eta(P(k) - W(k))] = \overline{P - W}^\eta = \delta$$

Therefore it is a feasible policy for $\bar{\mathbf{P}}_\mathbf{1}$. Obviously, any control policy (including the optimal policy) for $\bar{\mathbf{P}}_\mathbf{1}$ can be defined only based on the input variables $W(k)$ and $S(k)$ and not state $Y(k)$. Note that the optimization $\bar{\mathbf{P}}_\mathbf{1}$ replaces the hard third constraint of (2) with the soft constraint $\overline{P - W}^\eta = \delta$ which relaxes the optimization $\mathbf{P_1}$. The soft constraint only guarantees that in expectation the battery charge will remain at its initial value, but it may allow sample path $Y(k)$ exceeds its limits of $Y_{\min}$ and $Y_{\max}$ and hence the solution of $\bar{\mathbf{P}}_\mathbf{1}$ may not be feasible for $\mathbf{P_1}$.

We use the existence of a stationary randomized policy for the relaxed version and $\bar{\Phi} \leq \hat{\Phi}$ in the proof of performance bound of Theorem 1 which applies Lyapunov optimization approach. We also use the following lemma in the proof of feasibility part of Theorem 1.

Lemma 2. *For $\mathbf{P_1^g}$, the following propositions hold,*

$$\begin{array}{ll} Y(k) > d_T - \Omega_{\min}/(a_T\eta) & \Rightarrow Y(k+1) \leq Y(k) \\ Y(k) < d_T - \Omega_{\max}/a_T & \Rightarrow Y(k+1) \geq Y(k) \end{array} \quad (10)$$

Lemma 2 implies existence of a negative feedback for $Y(k)$ in Policy $\mathbf{P_1^g}$ (high Y stops charging, low Y stops discharging).

Proof of Theorem 1:

Performance bound: Define Lyapunov function $L(Y(k)) = a_T(Y(k) - d_T)^2/2$. From (1),

$$\begin{aligned} a_T\frac{(Y(k+1) - d_T)^2 - (Y(k) - d_T)^2}{2} \\ = a_T\frac{T^2(f_\eta(P(k) - W(k)) - \delta)^2}{2} \\ + a_T T(Y(k) - d_T)(f_\eta(P(k) - W(k)) - \delta) \quad (11) \end{aligned}$$

From constraint (5), the first term in the right hand side is at most $a_T BT^2$, hence for the greedy algorithm the Lyapunov drift is

$$\Delta L(Y(k)) \triangleq E[L(Y(k+1)) - L(Y(k))|Y(k)] \leq a_T BT^2 + a_T T(Y(k) - d_T)E[f_\eta(\tilde{P}(k) - W(k)) - \delta|Y(k)] \quad (12)$$

Adding T times the unit time average cost to both side we get

$$\Delta L(Y(k)) + TE[\tilde{Z}_1(k)|Y(k)] \leq a_T BT^2 + a_T T(Y(k) - d_T) \times E[f_\eta(\tilde{P}(k) - W(k)) - \delta|Y(k)] + TE[\tilde{Z}_1(k)|Y(k)] \quad (13)$$

where $\tilde{Z}_1(k)$, $\tilde{P}(k)$ is the value of $Z_1(k)$ and $P(k)$ obtained by the greedy algorithm at time k. Since the greedy algorithm for any given $Y(k)$ minimizes $a_T(Y(k) - d_T)f_\eta(\tilde{P}(k) - W(k)) + \tilde{Z}_1(k)$, it opportunistically minimizes the conditional expectations given any $Y(k)$, as compared to any other policy, including the optimal randomized stationary policy $q(.|S(k), W(k))$. Hence, we can write

$$\Delta L(Y(k)) + TE[\tilde{Z}_1(k)|Y(k)] \leq a_T BT^2 + a_T T(Y(k) - d_T) \times E[f_\eta(\bar{P}(k) - W(k)) - \delta|Y(k)] + TE[\bar{Z}_1(k)|Y(k)] \quad (14)$$

for any $Y(k)$ resulting from the greedy algorithm, where $\bar{Z}_1(k)$ and $\bar{P}(k)$ are the value of the Z_1 and P obtained by the optimal randomized stationary policy of $\bar{\mathbf{P}}$ given the states S and W at time slot k. Noting that for $\bar{\mathbf{P}}$, $\mathsf{E}[f_\eta(\bar{P}(k) - W(k))] = \delta$,

$$\Delta L(Y(k)) + TE[\tilde{Z}_1(k)|Y(k)] \leq a_T BT^2 + T\bar{\Phi}_1 \quad (15)$$

Taking expectation with respect to $Y(\tau)$ from the greedy algorithm and summing over τ we can write

$$\sum_{k=0}^{n-1} E[\tilde{Z}_1(k)] \leq a_T BnT + n\bar{\Phi}_1 + \frac{L(Y(0)) - L(Y(k))}{T} \quad (16)$$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} E[\tilde{Z}_1(k)] \leq \bar{\Phi}_1 + a_T BT \leq \hat{\Phi}_1 + a_T BT \quad (17)$$

Feasibility:

To show the feasibility of Policy $\mathbf{P_1^g}$ for $\mathbf{P_1}$, we show that assuming $Y_{\min} \leq Y(k) \leq Y_{\max}$, for the choice of a_T, d_T in (6),(7), $Y_{\min} \leq Y(k+1) \leq Y_{\max}$.

To show the above conclusion, if $d_T - \Omega_{\min}/(a_T\eta) < Y(k) \leq Y_{\max}$ then from (10), $Y(k+1) \leq Y(k) \leq Y_{\max}$ and if $Y(k) \leq d_T - \Omega_{\min}/(a_T\eta)$ then (assuming $\Omega_{\min}/\eta < \Omega_{\max}$ and using the definition of a_T) $Y(k+1) \leq d_T - \Omega_{\min}/(a_T\eta) + TR_m \leq Y_{\max} + d_T - Y_{\min} - T(D_m + \delta) - \Omega_{\max}/a_T = Y_{\max}$, i.e., $Y(k) \leq Y_{\max} \Rightarrow Y(k+1) \leq Y_{\max}$.
On the other hand, for $Y_{\min} \leq Y(k)$ if $Y(k) < d_T - \Omega_{\max}/a_T$ then from (10), $Y_{\min} \leq Y(k) \leq Y(k+1)$ and if $Y(k) \geq d_T - \Omega_{\max}/a_T$ then $Y(k+1) \geq d_T - \Omega_{\max}/a_T - TD_m - T\delta = Y_{\min}$. Hence $Y(k) \geq Y_{\min} \Rightarrow Y(k+1) \geq Y_{\min}$. ■

Proof of Lemma 2:

For the first proposition we need to show that the solution $P^*(k)$ to $\mathbf{P_1^g}$ at time k satisfies $P^*(k) \leq W(k)$ whenever $a_T(Y(k) - d_T) > -\Omega_{min}/\eta$. This is because if $P^*(k) > W(k)$, the value of the objective in $\mathbf{P_1^g}$ (assuming $a_T(Y(k) - d_T) > -\Omega_{min}/\eta$) is

$$(\eta a_T(Y(k) - d_T) + V\hat{C}(S(k), P^*(k)))P^*(k) + h(P^*(k) - W(k)) > (\eta a_T(Y(k) - d_T) + V\hat{C}(S(k), W(k)))W(k)$$

hence $P(k) = W(k)$ is a better solution than $P^*(k) > W(k)$ which contradicts the optimality of $P^*(k)$.

For the second proposition, if $P^*(k) < W(k)$, then the value of the objective in $\mathbf{P_1^g}$ is

$$(a_T(Y(k) - d_T) + \hat{C}(S(k), P^*(k)))P^*(k) + h(P^*(k) - W(k)) > (a_T(Y(k) - d_T) + \hat{C}(S(k), P^*(k)))P^*(k) \quad (18)$$

which based on the definition Ω_{max} indicates that if $a_T(Y(k)-d_T) < -\Omega_{max}$, the RHS of (18) (under assumption $P^*(k) < W(k)$) is greater or equal to $(a_T(Y(k)-d_T) + \hat{C}(S(k), W(k)))W(k)$ hence $P(k) = W(k)$ is a better solution than $P^*(k) < W(k)$ which contradicts the optimality of $P^*(k)$. ■

IV. Sampling period and non-i.i.d. sources

Let us now examine the impact of the discrete time step, T, to obtain insight into how to choose a suitable value. Unlike most prior work on Lyapunov optimization, we have explicitly considered T so that the discrete time processes are embedded in real time. This avoids confusion between terms that represent *power* and those that represent *energy used per time slot*, which are numerically equal if the time slot is normalized to have duration 1.

First, consider the dependence on T of the gain a_T of the Lyapunov function. To ensure that $a_T > 0$, we require

$$T < (Y_{\max} - Y_{\min})/(R_m + D_m + \delta) \equiv T_{\max}. \qquad (19)$$

For such T, a_T is monotonically decreasing in T. If T is close to $T_{\max}$ then the factor a_T in the bound becomes extremely large, indicating that T should decrease as either $Y_{\max} - Y_{\min}$ decreases or the charge rates increase.

The upper limit on T is because the third and fourth terms in the denominator of a_T in (6) are there to ensure that the state of charge Y will not "overshoot" the feasible region within a single time slot. If the rate of charging or discharging is updated sufficiently often, this guard region is not required, and a_T can approach

$$a_{\min} = \frac{\Omega_{\max} - \Omega_{\min}/\eta}{Y_{\max} - Y_{\min}}. \qquad (20)$$

This phenomenon is limited to settings where hard constraints are imposed, which are not usually treated by Lyapunov optimization.

The bound in Theorem 1 on the optimality gap also contains a factor proportional to T. However, this optimality gap is for the case that $S(\cdot)$ and $W(\cdot)$ are i.i.d sequences. As $(T \downarrow 0)$ these sequences as samples of continuous time processes $\bar{S}(\cdot), \bar{W}(\cdot)$ will not satisfy the i.i.d assumption. Without considering the importance of this assumption in Theorem 1, the performance bound may give the illusion that there is no performance penalty for using an algorithm such as $\mathbf{P_1^g}$ that makes no use of knowledge of the future.

In order to estimate the performance with rapid updates (small T), we must consider dependence within $S(\cdot)$ and $W(\cdot)$. Two approaches, described in [4], are to compare $\mathbf{P_1^g}$ against the optimal solution assuming (S, W) is a finite-state Markov processes [5] or comparing $\mathbf{P_1^g}$ against an algorithm with bounded lookahead for arbitrary bounded (S, W) [6]. We will briefly describe the former.

Let us first recall the discrete time case, and then relate that to the underlying continuous time process. Model (S, W) as an ergodic finite state Markov chain, denoted M_T. Each state of M_T is a renewal state. Select an arbitrary state and denote it by 0. The number of slots between successive visits to 0 are i.i.d. distributed as a random variable R_T. Then $n_c = \mathbb{E}[R_T^2]/\mathbb{E}[R_T]$ is a measure of the convergence time of M_T. By (4.82) of [4],

$$\begin{aligned}\lim_{k\to\infty} \frac{1}{k}\sum_{\tau=0}^{k-1} \mathsf{E}[Z_1(\tau)] &\le \hat{\Phi}_1 + a_T B \frac{\mathbb{E}[(TR_T)^2]}{\mathbb{E}[TR_T]} \\ &= \hat{\Phi}_1 + a_T B T n_c.\end{aligned}$$

Now assume that (S, W) arises by sampling a continuous time Markov chain M for $(\bar{S}, \bar{W})$ periodically at intervals of T. Let the renewal periods be denoted as R, and let $T_c = \mathbb{E}[R^2]/\mathbb{E}[R]$. The remaining task is to relate Tn_c to T_c. Typically TR_T will be larger than R, because the process may enter and leave state 0 within a single sampling period. Reducing T can thus be expected to reduce Tn_c. However, if T is short compared with both the average sojourn time in state 0 and the excursion times away from state 0, then it is likely that all transitions both into and out of state 0 are observed, and the differences between R_T and R will consist primarily of quantization error. Consequently, the tightest bound achievable using this approach is

$$\lim_{k\to\infty} \frac{1}{k}\sum_{\tau=0}^{k-1} \mathsf{E}[Z_1(\tau)] \le \hat{\Phi}_1 + a_{\min} B T_c \qquad (21)$$

which is obtained when $T \ll T_{\max}$ and T is much less than the minimum of the sojourn times in state 0 and in the set of other states. Further reduction of T increases the computational load without reducing the performance bound.

A. Detailed models capturing slow variability

The bound (21) is potentially extremely loose. A detailed model of power price and demand must consider diurnal variations, causing T_c to be at least a day. Greater accuracy, and a much looser bound, results from modeling annual variation. However, the optimal solution depends only on recent history, becoming more limited as the battery capacity decreases [2]. This suggests that tighter bounds could be obtained using the techniques of Section 4.9.2 of [4].

Acknowledgements

LA is supported by ARC grants FT0991594 and DP130100156. While performing this work, the authors were at Swinburne University of Technology.

References

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

[2] R. N. Jesudasan, L. L. H. Andrew, and H. L. Vu, "Scheduling inefficient storage," in *Proc. IEEE PES General Meeting*, 21-25 Jul 2013.

[3] G. L. Plett, "Extended Kalman filtering for battery management systems of LiPB-based HEV battery packs: Part 3. state and parameter estimation," *Journal of Power sources*, vol. 134, no. 2, pp. 277–292, 2004.

[4] M. J. Neely, *Stochastic Network Optimization with Application to Communication and Queueing Systems.* Morgan & Claypool, 2010.

[5] L. Huang and M. J. Neely, "Max-weight achieves the exact $[O(1/V), O(V)]$ utility-delay tradeoff under markov dynamics," *arXiv preprint arXiv:1008.0200*, 2010.

[6] M. J. Neely, "Universal scheduling for networks with arbitrary traffic, channels, and mobility," in *Decision and Control (CDC), 2010 49th IEEE Conference on.* IEEE, 2010, pp. 1822–1829.