Peak Power Demand Analysis by Using Battery Buffers for Monotonic Controllers

Technical Report: KIT-MTA-2013-0003, Updated: August 15, 2013

Waqaas Munawar and Jian-Jia Chen
Department of Informatics, Karlsruhe Institute of Technology, Germany
{munawar, jian-jia.chen}@kit.edu

***Abstract*—Demand of electricity varies on hourly basis whereas the production is quite inelastic. This results in heavy fluctuation in the price of the electrical energy. Consequently many power companies have started offering time based tariffs. Data centers and industrial consumers of electricity are penalized for the peak power demand of their loads. To shave the peak power demand, a battery buffer can be adopted. The battery is charged during low load and is discharged during peak loads. One essential question is to analyze the reduction of the peak power demand by adopting battery buffers. The power loads are modeled in this paper by adopting the concept of arrival curves in Network Calculus. We analyze monotonic controllers, which have these two properties: (1) comparing one given trace of power loads and two initial battery statuses, if we start from higher battery status, the resulting battery status in the future will not become lower; (2) to increase the power demand at time slot t, the power loads released before t should be as close as possible to t. We present a simple and effective monotonic controller and also provide analyses for the peak power demand to the power grid. Our analysis mechanism can help determine the appropriate battery size for a given load arrival curve to reduce the peak.**

I. Introduction

Demand of the electricity varies on hourly basis. The trends shaping the overall demand curve include the "daily-component" due to office-home cycle of general population, "yearly-component" due to seasonal periodicity and other factors which are more random in nature. Electric power production, on the other hand, is quite inelastic. Due to the randomness of demand, the supply must be provisioned for the worst-case power loads which might only occur very seldom. This results in an over-provisioned supply network and generation setup. Consequently, additional costs are incurred which are equally shared by all the consumers irrespective of their contribution in the peak loading. To make the demand more uniform, the power generation companies have introduced time varying tariffs for bigger consumers such as data centers. The actual bill not only depends on how much power was consumed but also how quickly it was consumed. If the demands are $(d_1, d_2, ..., d_n)$, then the total bill is of the form $c_1 \cdot \sum_i d_i + c_2 \cdot \max_i\{d_i\}$. Peaks are often measured in the time scale of 15 mins to 1 hour. The peak penalty is also referred to as demand charge and it can be more than 300 times the regular cost of energy [5]. Consequently, the demand charge often exceeds the amount of all other charges combined together [5].

Three main approaches exist to mitigate the problem. I.e., rescheduling/shedding the loads [4], spatial redistribution of the loads [18] and employing a store of energy to minimize the peaks [2], [17], [20]. These three are mutually orthogonal approaches and can be employed independent of each other. In this paper we focus on the last one i.e. employing a storage system to shave the peaks.

The basic architecture by employing a storage system to shave the peaks is shown in Figure 1, in which the power grid is the service provider for the electricity, the local reservoir stores the spared energy for reducing the peak, and the controller decides when and how to charge (or analogously, store) and when and how to discharge (or analogously, consume).

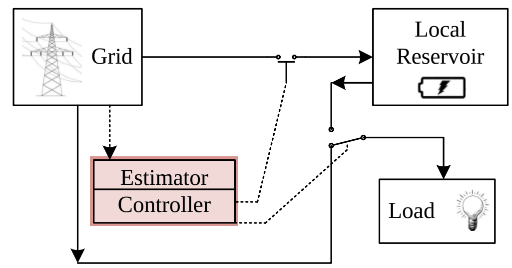

Fig. 1. Fundamental architecture of a peak shaving system.

Many technologies exist for storing the electrical charge [15], [19]. But battery electrical storage system (BESS) is the most promising due to the unique benefits that it offers. E.g., it does not depend on any geological features like compressed air and pumped hydro storage, it offers good cycle efficiency and, with maturing battery technology, it has become quite cost effective in recent years [6]. Throughout this paper we will be focusing on BESS, however the results are equally valid for other storage systems as well.

The peak shaving problem can be generalized as a minimax problem applicable to any item that can be stored in time of lower demand and then dispensed during the periods of higher demand. Economic lot sizing is one example. The is a commonly found general problem. But for the convenience of presentation we will be using the terminology of batteries and grid from here onwards.

Our contributions in this paper are as follows: To analyze the peak power demand by employing a storage system, we first characterize the loads by adopting the concept of arrival curves in Network Calculus [9]. That is, the power loads for the given time interval lengths are upper bounded by the given arrival curve. Moreover, we consider monotonic controllers, in which the following two properties hold: (1) comparing one given trace of power loads and two initial battery statuses, if we start from higher battery status, the resulting battery status in the future will not become lower; (2) to increase the power demand at time slot t, the power loads released before t should be as close as possible to t.

In Section IV, we provide a simple controller that charges/discharges the storage system based on a given threshold and prove that such a controller is a monotonic controller. We present a methodology in Section V to analyze the (worst-case) peak power demand to the power grid based on the arrival curve for monotonic controllers. The presented analysis technique is applied on the traces of real loads from New York region [16] and Indiana University campus [5]. The evaluations are presented in Section VI. We show some applications of the presented analysis technique in Section VII.

II. Related Work

The peak shaving problem in utility networks, also known as demand response, is gaining importance due to increased focus on smart grid technologies. There exist some results in the literature in this direction, such as [1]–[3], [10], [11], [13], [17], [20]. The work in this direction can be divided into two categories on the basis of two pricing models.

In the first model, the electricity provider controls the electrical load of the consumer and turns it off during the peak load times. The maximum duration for which the load can be turned off is agreed via contract. This approach is adopted in [1], [11].

In the second pricing model, the demand response is encouraged through controlling the pricing for the peak. This has emerged to be the most common approach toward encouraging a smoother demand. There have been works in this direction such as [3], [13], [17].

In both of these cases, the peak demand can be decreased by introducing a battery in the system. To determine the most appropriate battery size, and to measure the effectiveness of a control algorithm, an important milestone is to be able to quantify the peak load that can occur in presence of that battery and that control algorithm.

In this paper we present a methodology for quantifying the peaks. Fundamental difference between our work and the previous works such as [1], [11] is that we consider the load as inelastic which must be fulfilled at the time of demand. This assumption is more in line with contemporary market practices.

Among the works following the second pricing model, [3] offers online algorithms for shaving peaks and present worst-case competitive ratio analysis for these. This work aims to minimize the peak through specific algorithms whereas we aim to quantify them for any given monotonic algorithm. They also do not consider the practical aspects of inefficiencies in batteries. More recently, [13] tackle the same problem using approach of optimal control in context of data centers. Their focus, however, has been on decreasing the peaks on the average whereas, for billing, only the highest peak is considered.

III. System Model and Problem Definition

In this section we formalize the system model and present the problem definition. Figure 2 shows the abstract architecture for the peak shaving system considered in this paper.

A. Load Model

We discretize the time into time slots indexed through i. For time slot i, the plant requests ℓ_i amount of energy as a non-elastic load i.e., it is *non-postponable* and *non-reducible*. It is assumed that the connection to the grid is capable of fulfilling the worst-case peak demand at any instant. We assume that the loads are unknown until the beginning of time slot i. The loads could come irregularly, periodically, sporadically, or with jitters. To capture the worst-case behavior of the inelastic loads, we adopt the arrival curves in Network Calculus [9]. That is, the cumulative load R in time interval $[s, t)$ is upper-bounded by the upper arrival curve α, in which

$$R(s,t) = \sum_{i=s}^{t-1} \ell_i \le \alpha(t-1-s) \quad \forall s \le t. \tag{1}$$

If this condition holds then we can say that the cumulative load R complies to an arrival curve α. Intuitively, α can be seen as representative of worst-case cumulative electrical load. This curve can be derived from the historical loading curves using min-plus-deconvolution function. Throughout the analysis, we assume that arrival curve α is given.

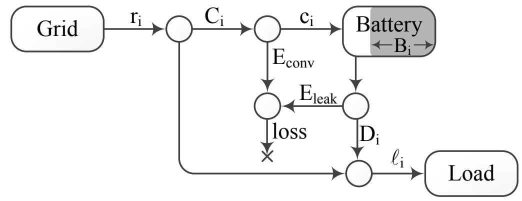

Fig. 2. Abstract Model of the system.

B. Battery Model

There is a battery in the system, with B_{max} as the maximum battery capacity. At time slot i, the battery level is B_i to denote the amount of available energy in the battery. At time slot i, the controller may decide to charge the battery with C_i amount of energy.

We assume that the battery is non-ideal and may suffer from two types of inefficiencies. Firstly, it may suffer from the charging circuit inefficiency, called *conversion energy loss*. That is, out of the total energy C_i routed towards the battery, only a fraction of it, ϵ_{conv}, actually gets stored. The conversion efficiency of big batteries (e.g., Sodium-Sulfur) is normally 90% and 70% respectively [7], [12].

Secondly, we assume that the battery loses a fraction of its charge continuously with time like a leaky bucket, called *leakage energy loss*. We denote this fraction by ϵ_{leak}, where $0 \le \epsilon_{leak} < 1$.

C. Monotonic Controllers

In a system with a local energy buffer, the peak demand from the electrical grid does not only depend on the electrical load, but also the decisions of controller. That is, the decisions about when to charge and discharge the battery, naturally, has an influence on the peak demand. Hence, these must also be considered.

For time slot i, the controller has to decide how to charge/discharge the battery and how to request the electricity from the grid. That is, the controller decides

1) the request to the grid, r_i,
2) the charge to the battery, C_i, and the discharge from the battery D_i.

We assume that the control is causal and does not have the knowledge of future.

The request r_i to the grid is divided into two parts: (1) part of it, C_i, is used to charge the battery, and (2) the rest, $r_i - C_i$ is used to serve the demand ℓ_i. By the above definitions, a controller must ensure that $D_i + (r_i - C_i)$ is at least ℓ_i to meet the load demand ℓ_i.

Suppose that the controller decides to charge the battery with C_i at time slot i. The resulting battery charge at the beginning of the time slot is $B'_i = \max\{B_i + C_i(1-\epsilon_{conv}), B_{max}\}$. Due to the leakage current in the battery, the charge level in the battery at the beginning of the next time slot is $B_{i+1} = B'_i \cdot (1-\epsilon_{leak})$. The methods and techniques presented in this paper can trivially accommodate both of these inefficiencies. For ease of presentation we will only consider perfect batteries henceforth, i.e., $\epsilon_{conv} = \epsilon_{leak} = 0$.

We assume that at any time slot the controller can satisfy the load partially from the battery and partially by requesting the grid.

The fractional mixing is allowed. This is a technically feasible and fairly common assumption [14].

Suppose that the controller decides to discharge the battery with D_i at time slot i. The discharge must be at most B_i, i.e., $D_i \leq B_i$; otherwise, the battery is said with an underflow. Clearly, in a lossless system with $\epsilon_{leak} = 0$ and $\epsilon_{conv} = 0$, the total input to the system should be equal to total output for consecutive time slots in a time segment, i.e.

$$B_s + \sum_{i=s}^{t} r_i = B_{t+1} + \sum_{i=s}^{t} \ell_i \quad \forall s \leq t. \tag{2}$$

A controller is said to be feasible if all the demands can be satisfied. Grid has capability to satisfy worst-case load at any instant. Moreover, this paper analyzes the peak demand of *monotonic* controllers, in which it is a feasible controller and the battery charge status is *monotonic* for a given sequence of loads. Formally, a monotonic controller is defined to satisfy the following two properties in addition to the feasibility.

- M1 *First property of monotonicity.* A controller starting with a higher battery finishes up with a higher battery for a same loading sequence. Concretely, suppose that we are given a sequence of loads $\ell_s, \ell_{s+1}, \ldots, \ell_t$ for $s < t$ and two initial states of the battery charges, B_s^1 and B_s^2 with $B_s^1 \leq B_s^2$. By applying the controller on the sequence of loads, the battery charge level at the beginning of time slot $t+1$ is B_{t+1}^1 (B_{t+1}^2, respectively) when starting from B_s^1 (B_s^2, respectively) at time slot s. A controller is said to be *monotonic* if it guarantees that $B_{t+1}^1 \leq B_{t+1}^2$ under the assumption that $B_s^1 \leq B_s^2$.
- M2 *Second property of monotonicity.* A higher load near the point of measurement leads to higher request at the point of measurement. Concretely, suppose that we are given a sequence $\mathcal{L}$ of loads $\ell_s, \ell_{s+1}, \ldots, \ell_t$ for $s < t$. Consider another sequence $\mathcal{L}'$ of loads $\ell'_s, \ell'_{s+1}, \ldots, \ell'_t$ generated from $\mathcal{L}$ by shifting some load σ from time slot i to time slot $i+1$ for one $i \in \{s, s+1, t-1\}$. That is, $\mathcal{L}$ is identical with $\mathcal{L}'$ except that ℓ'_i is $\ell_i - \sigma$ and ℓ'_{i+1} is $\ell_{i+1} + \sigma$. The controller is said to be *monotonic* if the request to the grid at time t by considering $\mathcal{L}'$ is no less than the request to the grid at time t by considering $\mathcal{L}$.

The analysis technique presented in this paper is only applicable to monotonic controllers. In Section IV, we show a practical example of such controllers.

D. Problem Definition

The peak measurement problem is defined as follows: *Given a monotonic controller $\mathcal{C}$, which serves as interconnect between the electrical grid, a battery of size B_{max} and inelastic electrical loads whose demands comply to an (upper) arrival curve α, the objective is to find the maximum peak request among the unknown demands.*

IV. Constant Request Controller

The constant request controller (see Algorithm 1) always requests a constant amount R_{th} from the grid following the battery constraints. If the total request is more than load, the surplus is charged to the battery and if it is less, the deficiency is fulfilled from the battery. If battery and R_{th} combined together are still not enough to satisfy the load, the underflow occurs and the unsatisfied load is supported directly from the grid. Analogously, if requested energy, R_{th}, is more than free capacity in battery and the load combined together, then overflow occurs and the request from the grid is decreased to the sum of demand (ℓ_i) and free capacity of battery ($B_{max} - B_i$).

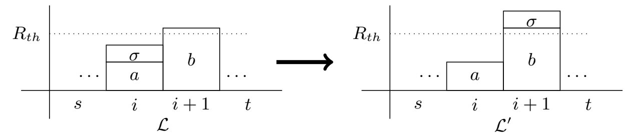

Fig. 3. Constant request controller - Proof of monotonicity (M2)

This controller covers the offline-optimal. If R_{th} is set to the average of the loads in the budgeting period, this controllers gives the optimal reduction in peak request provided the battery is of sufficient size, i.e., underflow and overflow never occurs. The pseudo-code of this algorithm is presented in 1.

It can be shown that this is a monotonic controller as follows.

Lemma 1: The constant request algorithm satisfies the property M1 for monotonic controllers.

Proof: We prove that the statement for M1, defined in Section III-C holds. Suppose $B_{t+1}^1 > B_{t+1}^2$. Then, in the previous time slot t, there are three possibilities.

1) $B_t^1 < B_t^2$. Considering the three possible values of request r_i, it is not possible to have battery status reverse in two consecutive time slots for the same loading sequence.
2) $B_t^1 = B_t^2$. If the battery status is same in slot t, then for identical load ℓ_t and identical request, it is not possible to have different battery status in slot $t+1$.
3) $B_t^1 > B_t^2$. Considering the three possible values at slot t for request r_t, it is the only possible status the batteries can have to end up in state $B_{t+1}^1 > B_{t+1}^2$ for slot $t+1$.

Applying the same argument backwards progressively to all slots until slot s would imply $B_s^1 > B_s^2$. This is a contradiction. ■

Lemma 2: The constant request algorithm satisfies the property M2 for monotonic controllers.

Proof: We prove that the statement for M2, defined in Section III-C holds. That is, $\mathcal{L}$ is identical with $\mathcal{L}'$ except that ℓ'_i is $\ell_i - \sigma$ and ℓ'_{i+1} is $\ell_{i+1} + \sigma$ where $\sigma \leq \ell_i$. Moreover, for the simplicity of the terminologies, we assume that $0 < \sigma \ll \min(B_{max}, a, b)$. Both the loading sequences $\mathcal{L}$ and $\mathcal{L}'$ consist of three parts $[s, i-1]$,$[i, i+1]$ and $[i+2, t]$. The first and the last are identical as shown in Figure 3.

Firstly, we consider the first interval $[s, i-1]$. Since the request for $\mathcal{C}$ depends on the load and the left over battery and both of these are exactly same for both $\mathcal{L}$ and $\mathcal{L}'$ in this interval hence the total request for interval $[s, i-1]$, $\sum_{j=s}^{i-1} r_j$ will also be same. I.e., the total output from the system and total input to the system for both $\mathcal{L}$ and $\mathcal{L}'$ in interval $[s, i-1]$. Hence, using Equation (2) and starting from same battery level, the battery level at slot i (B_i for $\mathcal{L}$ and B'_i for $\mathcal{L}'$) will also be the same.

Moreover, for the third interval, $[i+2, t]$, the loads in $\mathcal{L}$ and $\mathcal{L}'$ are the same from $i+2$ to t. Therefore, if the inequality $B_{i+2} \geq B'_{i+2}$ always holds, by Lemma 1, we can say $B_t \geq B'_t$.

Therefore, to prove the monotonicity for M2, we just have to prove the inequality $B_{i+2} \geq B'_{i+2}$. There are four cases, depending on the relation of R_{th}, ℓ_i, and ℓ_{i+1}:

Case 1 $R_{th} \geq \max(\ell_i, \ell_{i+1})$: Since the input to the system in both time slots is more than the output, hence the battery will be charged in both time slots. For $\mathcal{L}$, starting with B_i at i, the battery state at instance $i+2$ is given as

$$B_{i+2} = \min(B_{max}, B_i + 2R_{th} - \ell_i - \ell_{i+1}) \tag{3}$$

for $\mathcal{L}'$, at instances $i+1$ and $i+2$ the battery state will be

$$B'_{i+1} = \min(B_{max}, B_i + R_{th} - \ell_i + \sigma) \tag{4}$$
$$B'_{i+2} = \min(B_{max}, B_{i+1} + R_{th} - \ell_{i+1} - \sigma). \tag{5}$$

Substituting the value of B_{i+1} from Equation (4) in (5)

$$B'_{i+2} \leq \min(B_{max}, B_i + 2R_{th} - \ell_i - \ell_{i+1}). \tag{6}$$

Hence,

$$B'_{i+2} \leq B_{i+2} \tag{7}$$

The rest of the three cases can also be proved similarly. ■

Theorem 1: A controller $\mathcal{C}$, that in presence of battery of maximum size B_{max} and load ℓ_i requests r_i such that,

$$r_i = \begin{cases} B_{max} + \ell_i - B_i, \text{if } R_{th} - (\ell_i - B_i) > B_{max} \text{ (overflow)} \\ \ell_i - B_i, \text{if } R_{th} - (\ell_i - B_i) < 0 \text{ (underflow)} \\ R_{th}, \text{ otherwise} \end{cases}$$

where R_{th} is a constant, is a monotonic controller.

Proof: This is simply based on Lemmas 1 and 2. ■

V. Analysis Technique

In this section we present the core contribution of this paper, i.e. , the technique to analyze the peak electrical demand from grid given a load arrival curve and a monotonic control scheme to manage battery.

The core idea of the analysis technique is based on following three observations:

1) more the amount of charge in a battery, more it will be able to help in the reduction of peak demand from grid
2) releasing the load nearer to the point of measurement results in a more drained battery at the point of measurement than releasing the load further back in time, and
3) the worst-case load in any time interval $(t, t+\Delta)$ is given by the $[0, \Delta)$ portion of arrival curve.

On basis of these, it can be concluded that the peak demand from grid will occur when the maximum load is released as nearer to the point of measurement as possible. Repeating this process for all the points of interest gives the maximum grid demand at those points. A simple comparison among all these results reveal the maximum peak that may occur.

Lemma 3: Suppose that we are given a sequence of loads $\ell_s, \ell_{s+1}, \ldots, \ell_t$ for $s < t$ and two initial states of the battery charges, B_s^1 and B_s^2 with $B_s^1 \leq B_s^2$. By applying a monotonic controller on the sequence of loads, the request to the grid at time t by starting from B_s^1 is no more than the request to the grid at time t by starting from B_s^2.

Proof: Suppose a monotonic controller $\mathcal{C}$ in time slot t, requests r_t^1 when starting from B_s^1 and r_t^2 when starting from B_s^2. If $r_t^2 > r_t^1$, then $\mathcal{C}$ does not follow M1 (first property of monotonicity). This is a contradiction. ■

Algorithm 1: Constant request algorithm

Input: Load demand: ℓ_i,
Algorithm parameter: R_{th},
Battery parameters: B_{max}, B_i.
Output: Grid request: r_i,
Battery state: B_{i+1}.

$r_i = R_{th}$
$B_{i+1} = r_i + B_i - \ell_i$
/* Applying battery constraints */
if $B_{i+1} > B_{max}$ **then**
  // Overflow
  $r_i = r_i - (B_{i+1} - B_{max})$
  $B_{i+1} = B_{max}$
else if $B_{i+1} < 0$ **then**
  // Underflow
  $r_i = r_i - B_{i+1}$
  $B_{i+1} = 0$
return r_i, B_{i+1}

Lemma 4: Suppose that we are given a sequence $\mathcal{L}$ of loads $\ell_s, \ell_{s+1}, \ldots, \ell_t$ for $s < t$. Consider another sequence $\mathcal{L}'$ of loads $\ell'_s, \ell'_{s+1}, \ldots, \ell'_t$ generated from $\mathcal{L}$ by shifting some load σ from time slot i to time slot $i+1$ for one $i \in \{s, s+1, t-1\}$. That is, $\mathcal{L}$ is identical with $\mathcal{L}'$ except that ℓ'_i is $\ell_i - \sigma$ and ℓ'_{i+1} is $\ell_{i+1} + \sigma$. For a monotonic controller, the request to the grid at time t by considering $\mathcal{L}'$ is no less than the request to the grid at time t by considering $\mathcal{L}$.

Proof: Suppose a monotonic controller $\mathcal{C}$ in time slot t, requests r_t for $\mathcal{L}$ and r'_t for $\mathcal{L}'$. If $r_t > r'_t$, then $\mathcal{C}$ does not follow M2 (second property of monotonicity). This is a contradiction. ■

Theorem 2: Under a given arrival curve α and a monotonic controller, the peak request the controller requests to the grid at time slot t can be found by evaluating a sequence $\mathcal{L}$ of loads $\ell_1, \ell_2, \ldots, \ell_t$ under the given initial battery charge B_1, in which

$$\sum_{j=i}^{t} \ell_j = \alpha(t-i). \qquad \forall i = 1, 2, \ldots, t \tag{8}$$

Proof: Lets suppose there exist a loading sequence $\mathcal{L}'$ that results in a higher request at time instant t than $\mathcal{L}$, starting with same B_1. Then from Lemma 4, $\mathcal{L}'$ must have an interval $[i, t]$ such that $\sum_{j=i}^{t} \ell'_j > \sum_{j=i}^{t} \ell_j$. This implies that $\mathcal{L}'$ does not follow the arrival curve α and hence is not a valid loading sequence. ■

A. Convergence

Ability to analyze a *bounded* interval does not insure that maximum possible peak, that might occur in future, has been discovered. In this section we present a method to determine a feasible cutoff point for search. The core idea is to determine the upper and the lower bounds of the peaks and the search can be terminated where these two meet. Starting with an empty battery and analyzing for all time slots of an interval gives the highest peak for that interval. Analogously, starting with a full battery and analyzing for the same interval, gives the minimum possible peak for that interval. Once these lower and the upper bounds meet, this can be used as a cutoff point. It can be shown that, starting from a full battery, a peak higher than the one at the point of convergence will never occur, however, convergence is not guaranteed for all monotonic controllers.

Theorem 3: Given are loading sequence $\mathcal{L}$ and an arrival curve α, such that they follow Equation (8). Also given is a battery of maximum capacity B_{max} and a monotonic controller $\mathcal{C}$. Lets say,

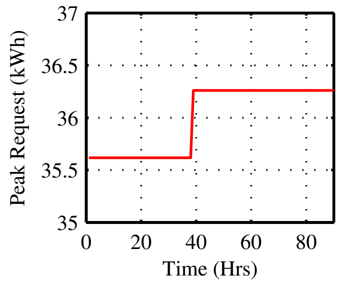

Fig. 4. Peak requests for constant request control algorithm

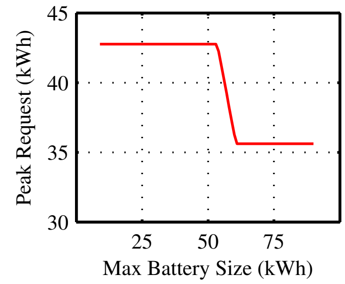

Fig. 5. Effect of battery size on the peak request

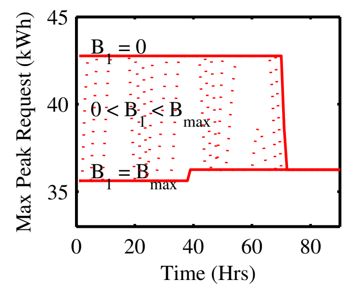

Fig. 6. Convergence interval

starting with an empty battery $B_1 = 0$, the peak requests of $\mathcal{C}$ for $\mathcal{L}$, are r_i^{emp} and starting with a full battery $B_1 = B_{max}$ they are r_i^{full}. Starting from first time slot, the first instant j, such that $r_j^{full} = r_j^{emp}$, defines the convergence interval as $[0, j]$. Starting from a full battery for $\mathcal{L}$ using $\mathcal{C}$, the highest possible peak is r_j^{full}.

Proof: We generate loading sequences $\mathcal{L}$ of length q, and $\mathcal{L}'$ of length $q+1$ from α using Equation (8). $\mathcal{L}$ and $\mathcal{L}'$ differ only for load at slot 1. I.e., if load for first slot in $\mathcal{L}'$ is ℓ'_1, then $\mathcal{L}' = \{\ell'_1, \mathcal{L}\}$. Lets say we start with an empty battery ($B'_1 = 0$) for $\mathcal{L}'$, then, at the start of slot 2, either the battery will accumulate some charge or remain empty. Hence, $B'_2 \geq B'_1$ and $B'_1 = 0 = B_1$. The $[2, q+1]$ portion of $\mathcal{L}'$ is identical as $\mathcal{L}$. Using Lemma 3, we can therefore state that request, r_q at q for $\mathcal{L}$ will be no less than r'_{q+1} at $q+1$ for $\mathcal{L}'$. Hence, when starting with an empty battery the peak requests for loading sequences generated from Equation (8) will decrease monotonically.

Similarly, if we start with full battery for $\mathcal{L}'$, then $B'_2 \leq B_{max}$, whereas the starting point for $\mathcal{L}$ will be B_{max}. Hence, starting with full battery, using the Lemma 3 we can state that request, r_q at q for $\mathcal{L}$ will be no more than r'_{q+1} at $q+1$ for $\mathcal{L}'$. Clearly, when starting with a full battery the peak requests for loading sequences generated from Equation (8) will increase monotonically.

Considering the above two cases, trivially their intersection will result in the highest possible peak, when starting B_{max} for loading sequences generated from Equation (8). ■

B. Computational Complexity

In first step, we generate the worst-case trace from the given arrival curve. The worst-case trace is the loading sequence that follows Equation (8). This is one time operation and for trace size of n, its complexity is $\mathcal{O}(n)$. Once the worst case trace is obtained it can be used repeatedly to determine the peaks for each time slot till convergence occurs. If we converge in k steps then the complexity will be $\mathcal{O}(k^2)$.

VI. Evaluations

To evaluate our technique for Algorithm 1 presented in Section IV we use the electrical load and price traces for the month of October 2007 from the New York region [16]. We use the maximum battery size of 60kWh unless otherwise mentioned. R_{th} is set to slightly higher than the mean of the total monthly load. We present the evaluations for three important criteria i.e. comparison of peak power demand, effect of battery size, and convergence interval.

Peak power demand: Peak requests for the evaluated control algorithm are shown in Fig. 4. We start with $B_1 = B_{max}$. With time, the battery depletes and its ability to dampen the peaks reduces. Hence a step near the 40^{th} hour. However, later as the worst part of the trace passes out the system enters an equilibrium.

Effect of the battery size: The battery size has a positive effect on the peak. As shown in Fig. 5, increasing the battery size results in decreased peak request from the grid. However, increasing it beyond a certain point does not yield any more benefit.

Convergence interval: Fig. 6 shows the convergence interval for Algorithm 1. We tested with some intermediate battery values between 0 and B_{max} and we found that the peak request never exits the envelope determined by the lower and upper bound.

VII. Applications

The ability to estimate peak demand for different control algorithms and battery sizes, is the central component in design of a peak shaving system. In this section we present some of its applications.

A. Battery Cost Determination

The premier application of the presented peak estimation technique is the selection of appropriate battery size. Batteries constitute the major portion of costs of any peak shaving system [8]. We present the results of application of the our scheme in Fig. 7. Here we specifically simulate a case study of Indiana University campus [5]. The battery costs are amortized with an interest rate of 10%. To estimate the maximum savings we implemented an offline optimal algorithm. We see that a battery based system can reduce the demand charges by more than 500,000 USD.

B. Controller Design

Another application of the presented scheme is to determine the design parameters of the controller such as charging rate, threshold to be used, etc. Controller design parameters influence the effectiveness of the controller and an inappropriate selection can not only result in

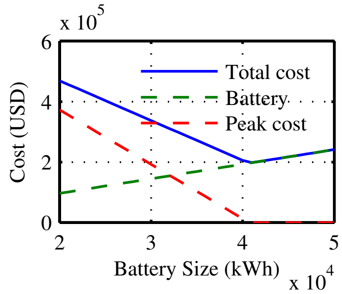

Fig. 7. Total cost and battery size

an ineffective controller, but can also cause an increase in the peak demand thereby causing an increase the electricity bill.

VIII. Conclusion

In this paper we presented a methodology to calculate the peak demand for a given battery under monotonic controllers and electrical load quantification using arrival curves. We presented a simple and effective example of a monotonic control algorithm for such a system and evaluated it against actual trace of electricity consumption from New York region and Indiana University. The presented methodology can help easing the adoption of smart grids of future in following ways. It helps in choosing an appropriate battery size considering realistic load profiles. It helps in fine tuning the control algorithm parameters. It can also help ease the integration of renewable energy.

The possible future directions to this work include generalizing the technique to include non-monotonic controllers and elastic loads.

References

[1] H. Aalami, M. P. Moghaddam, and G. Yousefi. Demand response modeling considering interruptible/curtailable loads and capacity market programs. *Applied Energy*, 87(1):243–250, 2010.

[2] A. Bar-Noy, Y. Feng, M. Johnson, and O. Liu. When to reap and when to sow–lowering peak usage with realistic batteries. *Experimental Algorithms*, pages 194–207, 2008.

[3] A. Bar-Noy, M. Johnson, and O. Liu. Peak shaving through resource buffering. *Approximation and Online Algorithms*, pages 147–159, 2009.

[4] D. Callaway and I. Hiskens. Achieving controllability of electric loads. *Proceedings of the IEEE*, 99(1):184–199, 2011.

[5] I. U. Campus Administrative Training Series (CATS). Electricity bill break up for indiana university. http://www.indiana.edu/~cats/presentations/2011/Electricity2011.pdf.

[6] GVEA. GVEA Ni-Cd BESS project. http://www.gvea.com/energy/bess.

[7] H. Ibrahim, A. Ilinca, and J. Perron. Energy storage systemscharacteristics and comparisons. *Renewable and Sustainable Energy Reviews*, 12(5):1221–1250, 2008.

[8] L. Jones. Electric Transmission Texas(ETT) signs contract for largest utility-scale battery in the U.S. . ETT Press Release: Sep 1, 2009.

[9] J.-Y. Le Boudec and P. Thiran. *Network calculus: a theory of deterministic queuing systems for the internet*. Springer, 2001.

[10] J.-Y. Le Boudec and D.-C. Tomozei. A demand-response calculus with perfect batteries. *MMB&DFT'12*, pages 273–287.

[11] J.-Y. Le Boudec and D.-C. Tomozei. Demand response using service curves. In *ISGT Europe*, pages 1–8. IEEE, 2011.

[12] S. Lemofouet and A. Rufer. A hybrid energy storage system based on compressed air and supercapacitors with maximum efficiency point tracking (mept). *Industrial Electronics, IEEE Transactions on*, 53(4):1105–1115, 2006.

[13] N. Li, L. Chen, and S. H. Low. Optimal demand response based on utility maximization in power networks. In *Power and Energy Society General Meeting*, pages 1–8. IEEE, 2011.

[14] A. Mishra, D. Irwin, P. Shenoy, and T. Zhu. Scaling distributed energy storage for grid peak reduction. In *4th international conference on Future energy systems*, pages 3–14. ACM, 2013.

[15] T. Nanahara. Pumped hydro storage. *JOURNAL-JAPAN INSTITUTE OF ENERGY*, 81(9; ISSU 905):811–818, 2002.

[16] New York Independent System Operator, NYISO. http://nyiso.com/.

[17] A. Oudalov, R. Cherkaoui, and A. Beguin. Sizing and optimal operation of battery energy storage system for peak shaving application. In *Power Tech*. IEEE, 2007.

[18] A. Qureshi et al. Plugging into energy market diversity. In *Proc. 7th ACM Workshop on Hot Topics in Networks (Hotnets-VII)*, 2008.

[19] P. F. Ribeiro, B. K. Johnson, M. L. Crow, A. Arsoy, and Y. Liu. Energy storage systems for advanced power applications. *Proceedings of the IEEE*, 89(12):1744–1756, 2001.

[20] R. Urgaonkar, B. Urgaonkar, M. J. Neely, and A. Sivasubramaniam. Optimal power cost management using stored energy in data centers. SIGMETRICS '11, pages 221–232, New York, NY, USA. ACM.