Optimal Peak-Minimizing Online Algorithms for Large-Load Users with Energy Storage

Yanfang Mo, Qiulin Lin, Minghua Chen, and S. Joe Qin

***Abstract*—The peak-demand charge motivates large-load customers to flatten their demand curves, while their self-owned renewable generations aggravate demand fluctuations. Thus, it is attractive to utilize energy storage for shaping real-time loads and reducing electricity bills. In this paper, we propose the first peak-aware competitive online algorithm for leveraging stored energy (e.g., in fuel cells) to minimize peak-demand charges. Our algorithm decides the discharging quantity slot by slot to maintain the optimal worst-case performance guarantee (namely, competitive ratio) among all deterministic online algorithms. Interestingly, we show that the best competitive ratio can be computed by solving a linear number of linear-fractional problems. We can also extend our competitive algorithm and analysis to improve the average-case performance and consider short-term prediction.**

I. INTRODUCTION

Utilities are motivating customers to play a more active role in power systems by introducing new pricing schemes, e.g., Time-of-Use (TOU), Real-Time Pricing (RTP), and Peak-Time Rebates (PTR). Notably, for large-load customers, in addition to the volume charge associated with the amount of consumed energy, utilities introduce the peak-demand charge [1], which sets a punitive charge on the maximum power during the billing period. Typically, the demand charge rate is over 100 HK$/kW, while the volume charge prices are around 50 cents/kWh in Hong Kong. Thus, it is no surprise that the peak-demand charge often accounts for a large portion of a commercial or industrial customer's bill, e.g., up to 90% for DC fast charging stations [2]. The heavy peak-demand charge drives customers to flatten their demand curves.

To reduce electricity purchases, energy consumers have been installing solar roofs or wind turbines. However, such volatile renewable generations aggravate demand fluctuations and render it harder to tame the peak-demand charge.

Energy storage has been a useful tool for peak shaving in power systems and is attracting large-load consumers' attention to reducing electricity bills. We herein focus on discharging the stored energy in an online fashion for peak-demand charge minimization. We consider discharging only from two aspects. First, recharging in an on-peak period is neither economic nor friendly to power systems. The volume charge rates are higher by TOU pricing, while recharging may increase the accumulated peak demand in a community microgrid. Second, it is inconvenient to recharge certain storage systems during use, such as fuel cells and pumped storage. Also, frequent charging and discharging behaviors may shorten the battery life. Last but not least, studying the storage discharging can shed light on the more complicated case regarding the charging/discharging joint optimization.

The work presented in this paper was supported in part by a Start-up Grant (Project No. 9380118) from City University of Hong Kong. Q. Lin was with The Chinese University of Hong Kong during this work.

Y. Mo, Q. Lin, M. Chen and S. J. Qin are with the School of Data Science, City University of Hong Kong, Kowloon, Hong Kong. E-mail: {yanfang.mo, qiulin.lin, minghua.chen, joe.qin}@cityu.edu.hk.

Designing an online discharging algorithm for reducing the peak-demand charge is challenging due to the non-cumulative nature of peak usage and the volatility of demands. Substantial existing results apply stochastic optimization or robust optimization. However, these methods may suffer from inaccurate estimations and the inefficiency of considering the worst absolute performance. Instead, we herein exploit the online algorithm design with competitive analysis, which relies little on future predictions and features the worst-case online-to-offline performance, referred to as Competitive Ratio (CR).

We propose the first peak-aware competitive online storage discharging algorithm with the optimal competitive relative performance guarantee regarding the peak-demand charge. Our algorithm adopts an intuitive and innovative CR-Pursuit framework [3], [4] and decides the discharging quantity slot by slot to maintain the optimal CR among all online algorithms. A unique technical contribution is that we obtain the best CR by solving a linear number of linear-fractional programs. We also extend the original algorithm to an adaptive one that improves average-case performance by exploiting the real-time information. Moreover, we adapt our algorithm to incorporate look-ahead information and obtain a better performance guarantee. Please refer to [5] for more technical details.

II. PROBLEM FORMULATION

Consider an on-peak period of T time slots. Let c be the storage capacity and $\bar{\delta}$ be the ceiling discharging quantity in a time slot. Given the demand profile $\boldsymbol{d} \in \mathbb{R}^T$, we formulate the Peak-Aware Discharging (PAD) problem as follows:

$$\textbf{PAD:} \quad \min_{\boldsymbol{\delta} \in \mathbb{R}^T} \quad \max_{t \in [T]} (d_t - \delta_t)$$

$$\text{subject to} \quad \sum_{t=1}^{T} \delta_t \leq c; \quad \text{(Inventory Constraint)}$$

$$0 \leq \delta_t \leq \min\{\bar{\delta}, d_t\}, \text{ for all } t \in [T],$$

where $[T] \triangleq \{1, 2, \ldots, T\}$, δ_t is the discharging quantity of time slot t, and $\max_{t \in [T]} (d_t - \delta_t)$ is the peak usage. We can easily solve the offline PAD and obtain the optimal objective value $v(\boldsymbol{d})$ by linear programming. However, owing to the volatile real-time demands, we merely know in advance the lower and upper bounds of the demand in each time slot, namely $d_t \in [\underline{d}, \overline{d}]$. An online discharging algorithm $\mathfrak{A}$ has to determine δ_t without knowing $d_k, k > t$, while the prior knowledge includes the sequentially revealed demands and discharging quantities of past time slots together with the demand bounds. Let $v_{\mathfrak{A}}(\boldsymbol{d})$ be the ultimate peak usage under the algorithm $\mathfrak{A}$ and the demand profile $\boldsymbol{d}$. The CR of $\mathfrak{A}$ is $CR_{\mathfrak{A}} = \max_{\boldsymbol{d} \in \mathcal{D}} \frac{v_{\mathfrak{A}}(\boldsymbol{d})}{v(\boldsymbol{d})}$, where $\mathcal{D} = \{\boldsymbol{d} \in \mathbb{R}^T \mid \underline{d}\mathbf{1} \leq \boldsymbol{d} \leq \overline{d}\mathbf{1}\}$. Our objective is to design an online algorithm such that its CR is as small as possible. Despite the generality of PAD, to the best of our knowledge, no existing results have studied optimal competitive online algorithms for PAD. Without loss of generality, we assume that $c \leq T\underline{d}$ as we consider large-load customers.

III. Optimal Online Discharging Algorithm

Let π^* be the optimal competitive ratio among all online algorithms for PAD, given the parameters $T, c, \bar{\delta}, \underline{d}, \overline{d}$. Our optimal online discharging algorithm pCR-PAD(π^*) sequentially determines the discharging quantity of time slot t as follows:

$$\delta_t = [d_t - \pi^* v(\boldsymbol{d}^t)]^+,$$

where $[x]^+ \triangleq \max\{x, 0\}$, $\boldsymbol{d}^t = [d_1 \ d_2 \ \cdots \ d_t \ \underline{d} \ \cdots \ \underline{d}]'$, and $v(\boldsymbol{d}^t)$ is the optimal objective value of PAD under $\boldsymbol{d}^t$.

Theorem 1. *pCR-PAD(π^*) achieves the optimal competitive ratio among all online algorithms for PAD.*

The above theorem and the algorithm pCR-PAD(π^*) are concise and intuitive. We attain the optimal discharging algorithm by maintaining the online-to-offline ratio of peak usage under the reference input $\boldsymbol{d}^t$ to be no more than the best possible CR π^*, namely, $\max_{k \leq t}(d_k - \delta_k)/v(\boldsymbol{d}^t) \leq \pi^*$. A critical issue remains unsolved: how can we obtain the smallest CR π^*? While π^* varies along with the parameters $T, c, \bar{\delta}, \underline{d}, \overline{d}$, we can compute π^* according to the following theorem.

Theorem 2. *Given $T, c, \bar{\delta}, \underline{d}, \overline{d}$, it holds that*

$$\pi^* = \max_{\boldsymbol{d} \in \mathcal{D}, k \in [T], k > \lfloor c/\overline{d} \rfloor} \frac{\sum_{i \in [k]} d_i - c}{\sum_{i \in [k]} v(\boldsymbol{d}^i)}. \quad (1)$$

By Theorem 2, we next show how to obtain π^* via a linear number of linear-fractional programs. For each $k \in [T]$, the right part of (1) is equivalent to the following linear-fractional program, where we substitute a variable v_i for $v(\boldsymbol{d}^i)$:

$$\max_{\boldsymbol{d} \in \mathcal{D}} \frac{\sum_{i \in [k]} d_i - c}{\sum_{i \in [k]} v(\boldsymbol{d}^i)} := \max_{\boldsymbol{d} \in \mathcal{D}, v_i, \delta_{ij}} \frac{\sum_{i \in [k]} d_i - c}{\sum_{i \in [k]} v_i}$$

$$\begin{aligned} \text{subject to} \quad & \sum_{j=1}^{T} \delta_{ij} = c, \text{ for all } i \in [T]; \\ & 0 \leq \delta_{ij} \leq \bar{\delta}, \text{ for all } i, j \in [T]; \\ & d_j - \delta_{ij} \leq v_i, \text{ for all } 1 \leq j \leq i \leq T; \\ & \underline{d} - \delta_{ij} \leq v_i, \text{ for all } 1 \leq i < j \leq T. \end{aligned}$$

Adaptive pCR-PAD Algorithm: Next, we shall improve the average-case performance by absorbing real-time information while maintaining the optimal worst-case performance guarantee. Specifically, the adaptive pCR-PAD algorithm improves pCR-PAD(π^*) by maintaining a better ratio π_t^* in each time slot, namely,

$$\delta_t = [d_t - \pi_t^* \cdot v(\boldsymbol{d}^t)]^+, \text{ for all } t \in [T],$$

$$\text{where } \pi_t^* = \max_{d_i \in [\underline{d}, \overline{d}], t \leq k \leq T} \frac{\sum_{i \in \{t, \ldots, k\}} d_i - \left(c - \sum_{i=1}^{t-1} \delta_i\right)}{\sum_{i \in \{t, \ldots, k\}} v(\boldsymbol{d}^i)}.$$

We see that π_t^* is no more than π^* and non-increasing in t. Here is the simple intuition behind the adaptive pCR-PAD: if the observed data $d_k, k < t$ can never be the first $t-1$ elements of a worst-case demand profile regarding the online-to-offline ratio of peak usage, then we should update our knowledge on worst cases and maintain a better ratio π_t^*, instead of π_{t-1}^*.

pCR-PAD Algorithm with Look-Ahead Information: We extend our optimal discharging algorithm to the case where there is a look-ahead window of size $W \in [T]$. That is, we know $d_k, k \in [1, \min\{T, t + W\}]$ at time slot t. We denote by π^W the optimal CR in this case. Again, we can compute π^W by linear-fractional programs, as shown below.

Theorem 3. *Given $T, c, \bar{\delta}, \underline{d}, \overline{d}, W$, it holds that*

$$\pi^W = \max_{\boldsymbol{d} \in \mathcal{D}, \mathcal{I} \subseteq [T]} \frac{\sum_{i \in \mathcal{I}} d_i - c}{\sum_{i \in \mathcal{I}} v\left(\boldsymbol{d}^{\min\{T, i+W\}}\right)}.$$

IV. Simulation Results

By the real data from an electric vehicle charging station, we uniformly set $T = 20$, $\underline{d} = 442.8$, and $\overline{d} = 1020.9$ for all instances. In practice, we can estimate these parameters based on historical data and update them in real time. The capacity rate refers to the ratio between the storage capacity c and the average total consumption of sampled demand profiles. The peak usage rate means the ratio between the ultimate peak usage under each algorithm and the original peak demand. Fig. 1 illustrates that our adaptive pCR-PAD algorithm attains better peak reduction effects in both the average case (marked points in lines) and the worst 15% case (cap lines) under different storage capacities, in comparison with several threshold-based algorithms and Receding Horizon Control (RHC) algorithms of the same look-ahead window size $W = T/4$. On average, our adaptive algorithm improves cost saving on the peak-demand charge by about 19% as compared to alternatives.

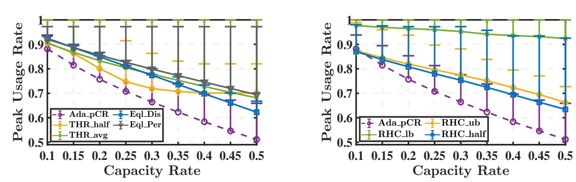

Fig. 1: Comparison with threshold-based and RHC algorithms.

References

[1] J. L. Neufeld, "Price discrimination and the adoption of the electricity demand charge," *J. Econ. Hist.*, pp. 693–709, 1987.

[2] G. Fitzgerald and C. Nelder, "Evgo fleet and tariff analysis: Phase 1: California," 2017.

[3] Q. Lin, H. Yi, J. Pang, M. Chen, A. Wierman, M. Honig, and Y. Xiao, "Competitive online optimization under inventory constraints," *Proc. SIGMETRICS*, vol. 3, pp. 1–28, 2019.

[4] H. Yi, Q. Lin, and M. Chen, "Balancing cost and dissatisfaction in online EV charging under real-time pricing," in *IEEE INFOCOM*, 2019, pp. 1801–1809.

[5] Y. Mo, Q. Lin, M. Chen, and S. J. Qin, "Online peak-demand minimization using energy storage," *ArXiv e-prints*, 2021, arXiv:2103.00005.