Peak-Aware Online Economic Dispatching for Microgrids

Ying Zhang, Mohammad H. Hajiesmaili, Sinan Cai, Minghua Chen, *Senior Member, IEEE,* and Qi Zhu, *Member, IEEE,*

***Abstract*—By employing local renewable energy sources and power generation units while connected to the central grid, microgrid can usher in great benefits in terms of cost efficiency, power reliability, and environmental awareness. Economic dispatching is a central problem in microgrid operation, which aims at effectively scheduling various energy sources to minimize the operating cost while satisfying the electricity demand. Designing intelligent economic dispatching strategies for microgrids, however, is drastically different from that for conventional central grids, due to two unique challenges. First, the demand and renewable generation uncertainty emphasizes the need for online algorithms. Second, the widely-adopted peak-based pricing scheme brings out the need for new peak-aware strategy design. In this paper, we tackle these critical challenges and devise peak-aware online economic dispatching algorithms. We prove that our deterministic and randomized algorithms achieve the best possible competitive ratios $2-\beta$ and $e/(e-1+\beta)$ in the fast responding generator scenario, where $\beta \in [0,1]$ is the ratio between the minimum grid spot price and the local-generation price. By extensive empirical evaluations using real-world traces, we show that our online algorithms achieve near offline-optimal performance. In a representative scenario, our algorithm achieves 17.5% and 9.24% cost reduction as compared to the case without local generation units and the case using peak-oblivious algorithms, respectively.**

***Index Terms*—Microgrids, Online Algorithm, Peak-Aware Scheduling, Economic Dispatching.**

NOTATIONS

This section lists the main notations used in this paper.

$e(t)$	The net electricity demand at time t in KWh.
$u(t)$	The amount of energy generated by local generators at time t in KWh.
$v(t)$	The amount of energy purchased from electricity grid at time t in KWh.
$p_e(t)$	The spot price of the electricity from grid at time t in \$/KWh.
$p_e^{\min}$	Minimum spot price of the electricity from grid, $\min_t p_e(t)$.
p_g	The unit cost of the electricity by local generators in \$/KWh.
β	Ratio between $p_e^{\min}$ and p_g.
p_m	The peak demand price of the electricity from grid in \$/KWh.
R^{u}	The maximum ramping up rate of local generator.
R^{d}	The maximum ramping down rate of local generator.
C	Local generator capacity.
T	Number of time slots in one charging period.
$\mathcal{T}$	Set of time slots in one charing period, $\{1,2,\cdots,T\}$.
$\mathbb{Z}^+$	The set of nonnegative integer numbers.

The work presented in this paper was supported in part by National Basic Research Program of China (Project No. 2013CB336700) and the University Grants Committee of the Hong Kong Special Administrative Region, China (Theme-based Research Scheme Project No. T23-407/13-N and General Research Fund No. 14201014).

Ying Zhang, Mohammad H. Hajiesmaili and Minghua Chen are with Department of Information Engineering, the Chinese University of Hong Kong. Sinan Cai is with School of Electrical Engineering, Xi'an Jiao Tong University. Qi Zhu is with Department of Electrical and Computer Engineering, University of California, Riverside.

I. INTRODUCTION

MICROGRID represents a promising paradigm of future electric power systems that autonomously coordinate distributed renewable energy source (*e.g.*, solar PVs), local generation unit (*e.g.*, gas generators), and the external grid to satisfy time-varying energy demand of a local community. As compared to traditional grids, microgrid has recognized advantages in cost efficiency, environmental awareness, and power reliability. Consequently, worldwide installed microgrid capacity has witnessed a phenomenon growth, reaching 866 MW in 2014, and is expected to reach 4,100 MW by 2020 [1].

Energy generation scheduling in microgrid determines the power output level of local energy sources and power to be procured from external grid, with the goal of minimizing the total cost over a pre-determined billing cycle. The scheduling plan should meet the time-varying energy demand and respect physical constraints of the generation units. Such problems have been studied extensively in the power system literature for traditional grids. Two main variants are unit commitment [2] and economic dispatching [3] problems. The unit commitment problem typically optimizes the start-up and the shut-down schedule of power generation units, whereas the economic dispatching problem optimally schedules the output levels given the on/off status as the input parameters. In this paper, we focus on economic dispatching problem in microgrid scenarios.

At first glance, economic dispatching in microgrid may appear to be a small-scale version of the classical urban-wide economic dispatching problem. However, the following two unprecedented challenges make the problem fundamentally different, thereby the previous solutions inapplicable.

$\triangleright$ **Demand and renewable generation uncertainty in microgrid.** Classical scheduling strategies for main grid rely on accurate prediction of future demand and dispatchable central generation [3]. However, without aggregation effect, the small-scale demand of microgrid is highly uncertain.

Meanwhile, the penetration of uncontrollable and intermittent renewable sources introduce uncertainty into generation scheduling. These observations motivate us to investigate new *online* scheduling strategies that do not rely on accurate prediction of future demand and renewable generation [4], [5].

▷ **Peak-based charging model of the external grid.** The real-world pricing scheme for consumers with large loads (such as universities or data centers) adopts a hybrid time-of-use and *peak-based* charging model where the electricity bill consists of both the total energy usage and the peak demand drawn over the billing cycle. The motivation is to encourage large customers to smooth their demand, thereby the utility provider can reduce its planned capacity obligations. The peak price is often more than 100 times higher than the maximum (on-peak) spot price, *e.g.*, 118 times for PG&E [6], and 227 times for Duke Energy Kentucky [7] [1]. Consequently, the contribution of peak charge in the electricity bill for a typical costumer can be considerable, *e.g.*, from 20% to 80% for several Google data centers [8]. These observations suggest that economic dispatching strategies with peak cost taken into account (referred to as peak-aware economic dispatching) may substantially reduce the total operating costs for microgrids as compared to economic dispatching strategies oblivious to peak cost (referred to as peak-oblivious economic dispatching). This is indeed the case as verified by our real-world trace-driven evaluation in Sec. IV.

Most of the previous researches on microgrid economic dispatching, that we are aware of and review in Sec. V, either adopt a peak-oblivious cost model, wherein the costumer bill is computed by total energy usage following a time-of-use pricing scheme, or rely on an accurate prediction of demand or renewable generation. In this paper, we tackle the peak-aware economic dispatching problem for microgrids by designing competitive online algorithms that do not rely on prediction of future input. Our main contributions are summarized as follows:

▷ We identify and formulate the peak-aware economic dispatching problem of minimizing the operating cost for microgrids under the hybrid time-of-use and peak-based pricing scheme in Sec. II. Notably, two aforementioned challenges change the structure of the problem fundamentally and call for online algorithm design.

▷ In Sec. III, we focus on "fast-responding" generator scenario, where the *ramping* constraints (*i.e.,* the maximum change in output level over successive steps) of local generators are ignored. We follow a *divide-and-conquer* approach and decompose the problem into multiple sub-problems, solve the sub-problems by their "rent-or-buy" nature, and then combine the solutions to obtain a solution for the original problem. We then demonstrate that the competitive ratios of our algorithms are $(2-\beta)$ and $e/(e-1+\beta)$ for deterministic and randomized versions respectively, where $\beta \in [0,1]$ is the ratio between the minimum grid spot price and the generator price. We prove that the ratios are the best possible. As such, these results characterize the fundamental *price of uncertainty* for the problem. The results in this part will help to solve the problem in "slow-responding" generator scenario [9].

▷ In Sec. IV, by extensive evaluations using real-world traces, we show that our online algorithms can achieve satisfactory empirical performance. Furthermore, our *peak-aware* online algorithms achieve near offline-optimal performance, and outperform the *peak-oblivious* designs under various settings. The substantial cost reduction shows the benefit and necessity of designing peak-aware strategies for economic dispatching in microgrids.

Some preliminary results in this paper were presented at ACM e-Energy 2015 [10] and all proofs are in the appendices.

II. Problem Formulation

In the microgrid economic dispatching problem, the objective is to orchestrate various energy sources to minimize the operating cost while satisfying the electricity demand.

We consider one billing cycle, which is a finite time horizon set $\mathcal{T} = \{1, \ldots, T\}$ with T discrete time slots of uniform length. In practice, the duration of one cycle is usually one month and the length of each time slot is 15 minutes [6]. In this paper, we quantize the electricity supply and demand to take only nonnegative integer values; this will simplify our presentation later on. Note that the quantization step can be arbitrarily small to achieve arbitrary granularity level.

Net electricity demand. Let $e(t)$ be the net electricity demand in time slot t, *i.e.*, the total electricity demand subtracted by the renewable generation. Note that since the renewable energy generation is in general very difficult to predict, we do not assume any specific stochastic model of $e(t)$, the pattern of which can be arbitrary.

Local generation. There are local generators deployed in the microgrid with total generation capacity C, *i.e.*, they can jointly satisfy at most C amount of electricity demand for each time slot. [2]. We consider a practical setting where the generator's incremental power output in two consecutive slots is limited by the *ramping-up* and *ramping-down* constraints R^{u} and R^{d}, respectively. Most microgrids today employ small-capacity generators that are powered by gas turbines or diesel engines. These generators are "fast-responding" in the sense that they have large ramping-up/-down rates. Meanwhile, there are also "slow-responding" generators with small ramping-up/-down rates. We denote p_g as the cost of generating unit electricity using local generators.

Electricity from the external grid. The microgrid can also obtain electricity supply from the external grid for unbalanced electricity demand in an on-demand manner. We denote the spot price at time t from the external grid as $p_e(t)$. We assume that $p_e(t) \geq p_e^{\min} \geq 0$ [3]. we do not assume any stochastic model of $p_e(t)$. For ease of discussion later, we define $\beta \triangleq$

[1] In practice, the unit of peak price is \$/KW while the unit of spot price is \$/KWh. This estimation is obtained by assuming the peak demand lasts one hour.

[2] Conventionally a generator's capacity is measured in KW, we consider a discrete time setting, under which the generator's capacity is computed by its actual capacity in KW multiplied by the length of slot in hour. For example, for a generator with actual capacity $C = 10$KW and the length of slot is 0.25 hour, its capacity is 2.5KWh.

[3] We remark that the electricity spot price can sometime be negative in practice [11]. We restrict our attention to the case with $p_e(t) \geq 0$ in this study and leave the general case with negative price to future work.

$p_e^{\min}/p_g$ as the ratio between the minimum grid price and the unit cost of local generation.

Cost model. The microgrid operating cost in $\mathcal{T}$ includes the expense of purchasing electricity from the external grid and that of local generation. Let $v(t)$ be the amount of electricity purchased from the external grid and $u(t)$ be the amount of electricity generated locally.

The cost of grid electricity consists of volume charge and peak charge. The volume charge is simply the sum of volume cost in all the time slots, *i.e.*, $\sum_t p_e(t)v(t)$. In practice, the peak charge is based on the maximum single-slot power and the peak price unit is \$/KW [6], which is different from the spot price unit \$/KWh. Let the peak price in \$/KW be $\tilde{p}_m$ and the length of one time slot be δ (*e.g.*, 0.25 hour), we convert the peak price to \$/KWh as $p_m = \tilde{p}_m/\delta$. Consequently, the peak charge is $p_m \max_t v(t)$, *i.e.*, the peak demand over the billing cycle (in KWh) multiplied by p_m (in \$/KWh). This method is similar to the one used in [8]. We remark that p_m is usually more than 100 times larger than $p_e(t)$ [6].

For local generation, the cost of a generator to generate θ amount of electricity is commonly modeled as a quadratic function [2], *i.e.*, say, $a\theta^2 + b\theta + c$. The coefficient a is usually orders of magnitude smaller than b (*e.g.*, for a typical oil generator with capacity 15MW, $a = 0.007, b = 48.5$) [4]. Consequently, for small-capacity generators employed in microgrids, the quadratic term $a\theta^2$ is usually much smaller than the linear term $b\theta$ and is negligible. In this paper, we consider the homogeneous local generators and denote p_g as the unit generation cost. The total local generation cost is simply $\sum_t p_g u(t)$.

Putting together all the components, the microgird total operating cost over a billing cycle is given by

$$\mathsf{Cost}(\boldsymbol{u},\boldsymbol{v}) = \underbrace{\sum_{t\in\mathcal{T}} p_e(t)v(t) + p_m \max_{t\in\mathcal{T}} v(t)}_{\text{by external grid}} + \underbrace{\sum_{t\in\mathcal{T}} p_g u(t)}_{\text{by local generators}}. \tag{1}$$

Existing microgrid generation scheduling schemes [4], [5] did not consider the peak charge term $p_m \max_t v(t)$; we refer to these schemes as **Peak-Oblivious**. In this paper, we consider the **Peak-Aware Economic Dispatching** (**PAED**) problem as follows,

$$\begin{aligned}
\textbf{PAED} \quad \min_{\boldsymbol{u},\boldsymbol{v}} \quad & \mathsf{Cost}(\boldsymbol{u},\boldsymbol{v}) \\
\text{s.t.} \quad & u(t)+v(t) \geq e(t), \quad t\in\mathcal{T}, && \text{(2a)} \\
& u(t) \leq C, \quad t\in\mathcal{T}, && \text{(2b)} \\
& u(t+1)-u(t) \leq R^{\mathsf{u}}, \quad t\in\mathcal{T}, && \text{(2c)} \\
& u(t)-u(t+1) \leq R^{\mathsf{d}}, \quad t\in\mathcal{T}, && \text{(2d)} \\
\text{var.} \quad & u(t), v(t) \in \mathbb{Z}^{+}, \quad t\in\mathcal{T}.
\end{aligned}$$

The constraint in (2a) ensures that the electricity demand is satisfied. The constraint in (2b) is due to the generator capacity limitation. The constraints in (2c)-(2d) reflect the ramping up/down constraints respectively.

[4]This can be further verified by more examples from http://pscal.ece.gatech.edu/archive/testsys/generators.html.

In the offline setting where the net demand in the entire time horizon, *i.e.*, $e(t)$ for all t in $\mathcal{T}$, is given (by for example accurate prediction), problem **PAED** can be solved easily using dynamic programming. If we consider continuous supply, the optimization problem is convex. However, the net demand $e(t)$ in microgrid is hard to predict as it inherits substantial uncertainty. This motivates the need of online strategies that do not rely on net demand prediction to operate [5].

In this paper, we use competitive ratio (**CR**) as the metric to evaluate how good an online algorithm is. For an online algorithm $\mathcal{A}$, its competitive ratio is defined as the maximum ratio between the cost it incurs and the offline optimal cost over all inputs, *i.e.*,

$$\mathbf{CR}(\mathcal{A}) \triangleq \max_{\text{all inputs}} \frac{\text{Cost incurred by } \mathcal{A}}{\text{Offline optimal cost}}.$$

Clearly we have $\mathbf{CR} \geq 1$. It is desired to design online algorithms with small competitive ratios, since it guarantees that, for any input, the cost of the online algorithm is close to the offline optimal.

III. Fast-Responding Generator Case

In this section, we relax the ramping constraints (2c)-(2d) and consider the fast-responding generator scenario. Most generators employed in microgrids can ramp up/down very fast. For example, a diesel-based engine can ramp up/down 40% of its capacity per minute [12]. Considering the time scale of each slot (*e.g.*, 15 minutes), those generators can be thought as having no ramping constraints. That is, $R^{\mathsf{u}} = R^{\mathsf{d}} = C$. We note that even though we relax the ramping constraints, the relaxed problem, denoted as **FS-PAED**, still covers many practical scenarios in microgrids [5]. Moreover, the results in this section serves a building block for designing online algorithm for the original problem **PAED** with ramping constraints, which is presented in our technical report [9].

In the following, we focus on the scenario where the unit cost of local generators is always higher than that of external grid, i.e., $p_e(t) \leq p_g$. If $p_e(t) > p_g$, it is always optimal to use the local generator as much as possible ($u(t) = \max\{e(t), C\}$) for both online and offline algorithms. The demands in such time slots will incur equal costs for the online and offline algorithms and decrease the ratio between them. Thus ignoring such demands will not change the competitive analysis of the online algorithms. As a result, we can have $p_e^{\min} \leq p_g$ and $\beta \leq 1$.

We will first consider a special version of problem **FS-PAED**, named as **FS-PAED**^{k}, where the net demand only takes value 0 or 1. We design optimal online algorithms for problem **FS-PAED**^{k} and then extend the algorithms to solve the general problem **FS-PAED**.

A. Problem *FS-PAED*^{k} and An Optimal Offline Solution

We now consider a special version of problem **FS-PAED** as follows:

$$\begin{aligned}
\textbf{FS-PAED}^k: \quad \min \quad & \mathsf{Cost}(\boldsymbol{u}^k,\boldsymbol{v}^k) \\
\text{s.t.} \quad & u^k(t)+v^k(t) \geq e^k(t), \quad t\in\mathcal{T}, \\
\text{var.} \quad & u^k(t), v^k(t) \in \{0,1\}, \quad t\in\mathcal{T},
\end{aligned}$$

where $e^k(t)$ only takes value 0 or 1. To keep the problem interesting, we assume the capacity C to be larger than 1; thus the capacity constraint is inactive and removed.

Note that problem **FS-PAED**^{k} can be solved by dynamic programming, which however does not seem to bring significant insights for developing online algorithms. As such, in what follows, we study the offline optimal solution from another angle to reveal a useful structure, which we exploit to design efficient online algorithms.

Under the setting, the unit cost of local generation is more expensive than the spot price of the external grid, *i.e.*, $p_e(t) \leq p_g$. However, the expensive local generation can be leveraged to cut off the peak demand satisfied by the external grid and thus the prohibited peak charge from the external grid. Thus, the key in solving problem **FS-PAED**^{k} lies in balancing between the cost of using the expensive local generation and the peak charge of using the external grid. It turns out the optimal offline solution, as shown in Lemma 1, is developed by comparing the accumulated deficit of using the local generation and the unit peak charge.

***Lemma** 1:* An optimal offline solution of **FS-PAED**^{k}, denoted by $\left\{\left(\left(u^k(t)\right)^*, \left(v^k(t)\right)^*\right)\right\}_{\mathcal{T}}$, only takes value 0 and 1 and is given by $\left(u^k(t)\right)^* = e^k(k) - \left(v^k(t)\right)^*$ and

- if $\sigma > 1$, then $\left(v^k(t)\right)^* = e^k(t)$, for all t in $\mathcal{T}$,
- otherwise $\left(v^k(t)\right)^* = 0$, for all t in $\mathcal{T}$.

Here σ is a critical peak-demand threshold defined by

$$\sigma \triangleq \frac{1}{p_m}\left[\sum_{t\in\mathcal{T}} \left(p_g - p_e(t)\right) e^k(t)\right]. \tag{3}$$

Remark: The optimal solution constructed in Lemma 1 is computed given that the critical peak-demand threshold σ is determined. Meanwhile, σ can only be computed in the offline setting where the net demand in the entire horizon is given, and it turns out it is the sufficient statistics of the net demand for characterizing the ratio between the cost of an online algorithm and the offline optimal cost.

*B. Online Algorithms for Problem **FS-PAED**^{k}*

The challenge for the online algorithm comes from the fact that it cannot determine the value of critical peak-demand threshold σ ahead of time. This brings out a dilemma in online decision making: *to suffer deficit of local generator and bypass the peak charge or to pay for the peak and enjoy cheaper electricity from the grid*. The most *aggressive* strategy acquires electricity from the grid from the very beginning, while the most *conservative* strategy uses local generation to satisfy all the net demands in the entire horizon, to avoid the peak charge.

An important observation in online decision making for problem **FS-PAED**^{k} is that after purchasing electricity from the grid once, meaning the peak charge has already been paid (and will not be charged again during the current billing cycle), the microgrid should continue to use the cheap electricity from the grid until the end of the billing cycle. Then the key decision is to determine when to start to pay the peak-charge premium and buy electricity from the grid.

To pursue online algorithms with minimum competitive ratio, it turns out that it suffices to focus on online algorithms that switch from local generation to grid electricity procurement when the accumulated local generation deficit exceeds some threshold, say $s \cdot p_m$, where $s \in [0, \infty)$ is an algorithm-specific parameter.

For deterministic algorithms, these are the ones switching to grid electricity procurement at time τ that satisfies the following condition for the first time in the entire horizon:

$$\sum_{t=1}^{\tau} \left(p_g - p_e(t)\right) e^k(t) \geq s \cdot p_m.$$

The most aggressive strategy discussed above corresponds to $s = 0$, and the most conservative one corresponds to $s = \infty$. Randomized online algorithms can be then characterized by probability distributions of s.

1) An Optimal Deterministic Online Algorithm: The deterministic online algorithm we design is to set $s = 1$, which means that we will purchase electricity from the grid when the accumulated local generation deficit seen so far just equals the peak charge. We name this algorithm as Break-Even Economic Dispatching for problem **FS-PAED**^{k} (**BED-**k). We provide its performance guarantee in the following theorem.

***Theorem** 1:* The competitive ratio of **BED-**k is given by

$$\mathbf{CR}\left(\mathbf{BED}{-}k\right) = 2 - \beta,$$

and no other deterministic online algorithm can achieve a smaller competitive ratio. [5]

2) An Optimal Randomized Online Algorithm: We can design a randomized online by randomly picking a value of s and start to purchase electricity from the grid when $\sum_{\tau}(p_g - p_e(\tau)) \geq s \cdot p_m$. The core of the randomized algorithm design is by which distribution we generate s. It is easy to imagine that different probability distributions will lead to algorithms with different competitive ratios.

The probability distribution we choose is

$$f^*(s) = \begin{cases} \frac{e^s}{e-1+\beta}, & \text{when } s \in [0,1]; \\ \frac{\beta}{e-1+\beta}\delta(0), & \text{when } s = \infty; \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

We name this randomized algorithm as Randomized Economic Dispatching for problem **FS-PAED**^{k} (**RED-**k). Its competitive ratio can be computed by solving

$$\max_{\sigma} \frac{\text{Expected online cost}}{\text{Optimal offline cost}},$$

***Theorem** 2:* With the distribution given by $f^*(s)$ in (4), the competitive ratio of **RED-**k is given by

$$\mathbf{CR}\left(\mathbf{RED}{-}k\right) = \frac{e}{e-1+\beta},$$

and no other randomized online algorithm can achieve a smaller competitive ratio.

Remark: (i) In the deterministic online algorithm, setting $s = 1$ means that the microgrid will start to buy electricity

[5]Recall that $\beta \triangleq p_e^{\min}/p_g$ is the ratio between the minimum grid price and the unit cost of local generation

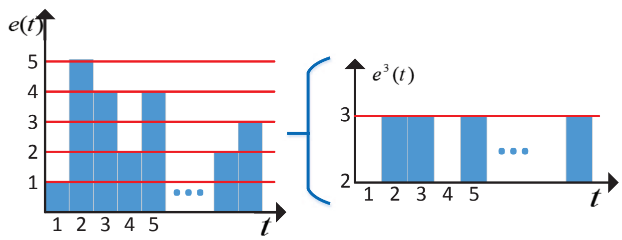

Fig. 1. An example of decomposing the demand into multiple layers and a microscopic view of layer 3

from the grid until the break-even condition is met. Similar to the ski rental problem [13], the break-even point turns out to be the best balance between being aggressive and conservative. (ii) The vigilant readers may notice that $f^*(s)$ is the same distribution that was adopted in solving the classic Bahncard problem [14], which is indeed similar to problem **FS-PAED**^{k} we study in this section. The basic version of **FS-PAED**^{k}, however, is different from Bahncard problem in the sense that the *discounted price* ($p_e(t)$ in this paper) is time varying.

C. From Problem *FS-PAED*^{k} to Problem *FS-PAED*

In this section, we design deterministic and randomized online algorithms for **FS-PAED** based on those of **FS-PAED**^{k}.

1) Net Demand Layering: Recall that $e(t)$ is assumed to take non-negative integer values. We divide the demand $e(t)$ into multiple layers such that the demand of each layer in each time slot is either 1 or 0, as shown in Fig. 1.

After layering, a bunch of sub-problems **FS-PAED**^{k} are obtained, which can be solved by the online algorithms **BED**-k or **RED**-k. However, unlike **FS-PAED**^{k}, the net demand of **FS-PAED** in some time slots can exceed the capacity of local generation, which makes it infeasible to ignore the whole picture while conquering each layer independently. For example, suppose the generation capacity is 4 for the case shown in Fig. 1. Even though the break-even points are not reached for all the layers in time slot 2, it is infeasible to set $u^k(2) = 1$ for all the layers (A capacity of 5 is needed to do so). Thus by taking into account the capacity constraint, we need to determine for which layers the demand should be satisfied by the grid while still keeping the algorithm competitive.

An obvious but critical observation is that the demands in the lower layers are denser than those in the upper layers. In addition, after being charged for the peak, we expect more demands to come to enjoy the cheap grid electricity. Consequently, it is always more economic to use the grid electricity to satisfy the denser demands, *i.e.*, the lower layers. In other words, in the proper algorithm design, the layers below $(e(t) - C)^+$ should always be satisfied by the grid. Meanwhile, for the layers above $(e(t) - C)^+$, if the demand is already satisfied by the grid, the online algorithm continues to acquire the electricity from the grid; otherwise, Algorithm **BED**-k or **RED**-k is applied with the same value s for all layers to obtain the sub-solutions. The solution is finally obtained by combining the sub-solutions. We summarize the resulting deterministic and randomized online algorithms, named as **BED** and **RED**, in Algorithm 1 and 2, respectively.

Algorithm 1 BED: Optimal deterministic online algorithm for **FS-PAED**

Require: C,p_m,p_g,$p_e(t)$,$e(t)$,$\varsigma^0 = 0$
Ensure: $u(t)$,$v(t)$
1: **while** $\tau \in \mathcal{T}$ **do**
2: A threshold: $\varsigma^\tau = \max\{\varsigma^{\tau-1}, (e(\tau) - C)^+\}$.
3: For the layers below ς^τ, $v^k(\tau) = 1$, $u^k(\tau) = 0$
4: For the layers above ς^τ, run **BED**-k to obtain $u^k(\tau)$ and $v^k(\tau)$.
5: $u(\tau) = \sum_k u^k(\tau)$, $v(\tau) = \sum_k v^k(\tau)$
6: $\tau = \tau + 1$
7: **end while**

Algorithm 2 RED: Optimal randomized online algorithm for **FS-PAED**

Require: C,p_m,p_g,$p_e(t)$,$e(t)$,$\varsigma^0 = 0$
Ensure: $u(t)$,$v(t)$
1: **while** $\tau \in \mathcal{T}$ **do**
2: A threshold: $\varsigma^\tau = \max\{\varsigma^{\tau-1}, (e(\tau) - C)^+\}$.
3: For the layers below ς^τ, $v^k(\tau) = 1$, $u^k(\tau) = 0$
4: For the layers above ς^τ, run **RED**-k with the same randomized parameter s to obtain $u^k(\tau)$ and $v^k(\tau)$.
5: $u(\tau) = \sum_k u^k(\tau)$, $v(\tau) = \sum_k v^k(\tau)$
6: $\tau = \tau + 1$
7: **end while**

We show a toy example solved by **BED** to demonstrate idea of layering approach in Fig. 2. In the example, we consider a 9-slot horizon with demand in each slot being $1, 5, 3, 2, 4, 2, 1, 2, 3$, respectively. The local capacity is 4, $p_g = 5, p_e(t) = 2$, and $p_m = 8$. For the subproblem in each layer, the peak charge will be compensated by the cheaper grid electricity if the total demand in that layer is larger than 2. Given all inputs, the optimal offline solution uses only the grid for the subproblems with total demand larger than 2 (layer 1, 2 and 3) and uses only the local generators otherwise (layer 4 and 5). The optimal offline cost is calculated to be 79. For the online solution derived by **BED**, it uses the local generator for the first two nonzero demands because $\sum_{t=1}^{\tau}(p_g - p_e(t)) < p_m$ for $\tau = 1, 2$ and switch to the grid when extra demands come because $\sum_{t=1}^{\tau}(p_g - p_e(t)) > p_m$ for $\tau \geq 3$. It should also be noted that, in slot 2, the total demand exceeds the total capacity, thus we have to use the grid to satisfy that unit demand in the first layer. We use different colors to demonstrate by which source and for what reason each unit demand is satisfied in Fig. 2. By back-of-the-envelop calculation, the online cost is 94, and the ratio between the online cost and offline optimal cost is 1.19 for this particular example.

Even though the example is simple, it demonstrates two important and provable properties of **BED**: (i) For each layer, it will continue to use the grid after it uses it once, and (ii) when one layer uses the grid, all the layers below it use the grid too. The first property makes the solution and cost structure

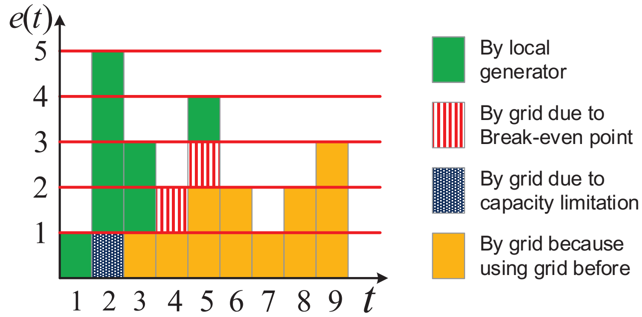

Fig. 2. Demonstration of **BED** with $C = 4$, different colors denoting different strategies of the algorithm.

similar to that of **BED**-k, while the second property makes the peak of $v(t)$ equal to the sum of the peaks of $v^k(t)$, *i.e.*, $\max_t \sum_k v^k(t) = \sum_k \max_t v^k(t)$. The two properties allow us to leverage the results in Sec. III-B to establish the competitive ratios of **BED** and **RED** in Theorem 3.

Theorem 3: The competitive ratios of **BED** and **RED** are given by

$$\mathbf{CR}\left(\mathbf{BED}\right) = 2 - \beta, \quad \text{and} \quad \mathbf{CR}\left(\mathbf{RED}\right) = \frac{e}{e - 1 + \beta}.$$

Further, no other deterministic and randomized online algorithm can achieve smaller competitive ratios.

In the next subsection, we discuss an intriguing impact of local generation capacity on the online algorithms' performance.

D. Critical Local Generation Capacity

The peak-aware economic dispatching aims at minimizing the sum of the peak charge (the term $p_m \max_{t\in\mathcal{T}} v(t)$ in (1)) and the volume charge (as the remaining part in (1)). The local generator provides the microgrid an option to use more expensive electricity (increase the volume charge) to reduce the peak (decrease the peak charge). An optimal solution is achieved with the best tradeoff between the two. Given an input, there is a threshold $\tilde{C}$, the demand below which should be satisfied by the grid and above which by the local generator. $\tilde{C}$ can be obtained by solving **FS-PAED** in an offline fashion without considering capacity constraint. It means that the optimal offline solution will not use the additional capacity even if it is larger than $\tilde{C}$.

We now discuss the impact of increasing local generation capacity C on the performance of offline and online algorithms. The offline algorithm will use full local capacity until C reaches $\tilde{C}$, and it will not use local capacity further beyond $\tilde{C}$. As such, one can expect that the operating cost of the offline algorithm is non-increasing as C increases. Meanwhile, the online algorithm, without knowing $\tilde{C}$ and with the tendency of reducing the peak with more expensive electricity, will try to exploit the whole capacity until it finds the break-even point, which turns out to be less economic and deviates more from the optimal solution. As a result, for the online algorithm, larger capacity may incur higher operating cost. We provide a concrete case-study by real world traces to confirm the above observation in Sec. IV.

Overall, we believe the above insights are important for microgrid operators to (a) determine the amount of local generation to invest in order to maximize the economic benefit, and (b) understand the importance of demand/generation prediction when performing peak-aware economic dispatching in microgrids.

IV. Experimental Results

We carry out numerical experiments using real-world traces to scrutinize the performance of our online algorithms under various practical settings. Our purpose is to investigate (i) the competitiveness of our online algorithms in comparison with the optimal offline one, (ii) the necessity of peak-awareness in economic dispatching of microgrids, and (iii) the performance of online algorithms under various parameter settings. More simulation results can be found in [9].

A. Experimental Setup

Electricity demand and renewable generation traces. We set the length of one billing circle as one month. We use the actual electricity demand of a college in San Francisco; its yearly demand is about 154GWh [15]. We inject renewable energy supply sources by a wind power trace of a nearby offshore wind station outside San Francisco with a total installed capacity of 12MW [16]. We then construct the net demand by subtracting the output level of the wind from the college electricity demand.

Energy source parameters. The electricity price $p_e(t)$ and peak price p_m are set based on the tariffs from PG&E [6] and $p_m = 17.56\$/\mathrm{KWh}$ while the electricity rate $p_e(t)$ varies from $0.056\$/\mathrm{KWh}$ to $0.232\$/\mathrm{KWh}$ for off-, mid-, and on-peak periods in different seasons. We set the unit cost of local generation p_g according to the monthly price of natural gas. Notably, the value of p_g could be less than $p_e(t)$ for some on-peak intervals. In such situations, generator plays its role not only by cutting off the peak but also by providing cheaper electricity as well. Finally, if not specified, the capacity of the local generator is set to be $C = 15\mathrm{MWh}$, which is around 60% of the peak net demand.

Cost benchmark. We use the cost incurred by only procuring electricity from the external grid, *i.e.*, $v(t) = e(t)$, as the benchmark. We demonstrate cost reduction to show the benefit of employing local generation units and the merits of our algorithms. The cost reduction originates from the cheaper electricity (in some on-peak intervals) and peak cut-off by local generators.

Comparison of algorithms. We compare our proposed peak-aware online economic dispatching algorithms **BED** and **RED** with (i) the optimal peak-aware offline solution (PA-OFFLINE) to evaluate the performance of the online algorithms, and (ii) the peak-oblivious online algorithms (PO-Online) in [5] to investigate the importance of peak-awareness.[6] We remark that both schemes in [4], [5] are peak-oblivious as they only consider volume charge but ignore peak charge.

[6]We remark that in [5], the joint unit commitment and economic dispatching problem in peak-oblivious manner is addressed and in this paper we compare the economic dispatching part with our algorithms.

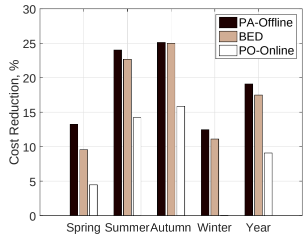

Fig. 3. Cost reduction for different seasons and the whole year

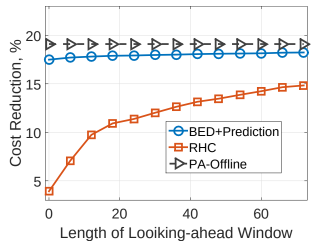

Fig. 4. Cost reduction with different length of looking-ahead window

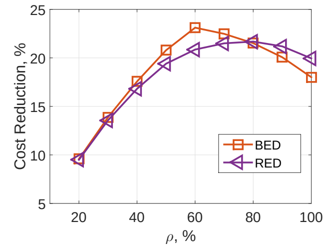

Fig. 7. Comparison of **RED** and **BED** under different local capacities

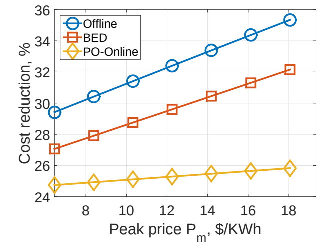

Fig. 8. Cost reduction with different peak price on a small microgrid testbed

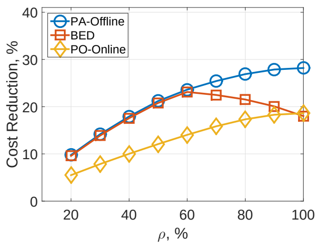

Fig. 5. Cost reduction vs different local capacities

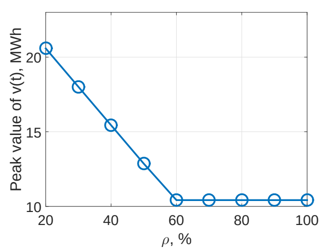

Fig. 6. $\max v(t)$ of PA-Offline vs ρ

B. Benefits of Employing Local Generators

Purpose. The purpose of this experiment is two-fold. First, compare the potential savings of microgrid in different seasons, in which the demand pattern, the wind output, and the cost parameters differ. Second, compare the cost reduction of peak-aware algorithms against peak-oblivious ones. The results are shown in Fig. 3.

Observations. The most notable observations from Fig. 3 are the following. First of all, the cost reduction varies over seasons and the most significant one occurs in the summer. This is because the gas price is lower and the grid electricity price is higher in the summer than those of the other seasons, thus employing local generators brings more benefit. Second, the performance of our proposed **BED** is superior than PO-Online algorithm. In particular, PO-Online cannot reduce the cost in the winter, but our algorithm **BED** can still achieve cost reduction. The reason is that, as $p_g > p_e(t)$ always holds in the winter, PO-Online algorithm always purchases cheaper electricity from the gird, which gives no cost reduction as compare to the benchmark strategy. In contrast, our **BED** algorithm reduces the cost by exploiting (the expensive) local generation to reduce the peak demand served by the external grid, and consequently our algorithm can save operating cost. On average, **BED** reduces the annual cost by 17.49%, while PO-Online reduces the cost only by 9.08%. *Third*, the performance of **BED** in practice is close to that of the offline optimal.

C. Benefit of Prediction

Purpose. For the online algorithm design, predicting the future is believed to be an effective mechanism to improve the performance [17] and many prediction algorithm are proposed [18], [19]. For the problem in this paper, the prediction will help us to notice the break-even point earlier and thus bring benefit. In this part, we evaluate how helpful prediction is by changing the looking-ahead window Δ from 0 to 24 time slots (two days). We also compare its performance with that of receding horizon control [20] algorithm (**RHC**) under different Δ and the results are shown in Fig 4.

Observation. As we can observe, the looking-ahead window will increase the performance of RHC significantly, but for our online algorithm **BED**, the cost reduction is not increased so much, only from 17.49% to 18.23%. This result indicates that predicting the near future contributes little to the overall performance. The underlying logic is as follows, the difference between the online solution with prediction and that without prediction only happens during the period (looking-ahead window) before the 'break-even point', the length of which is much smaller than that of the whole billing cycle. In other words, the prediction will not change the online solutions too much, thus the performance will almost remain the same. The good news is that, the performance of **RED** is close to optimal and the property of being insensitive to prediction will not make it less attractive compared with other design, like **RHC**.

*D. The Performance of **BED** under Different Local Generation Capacities*

Purpose. At first glance, one may imagine that larger local generator leads to larger design space and thus larger cost reduction is expected. However, as discussed in Sec. III-D, this is not the case for online algorithms that do not have the complete future knowledge of price and demand. We carry out an experiment to verify and elaborate the observation. For convenience, we define $\rho = C/\max e(t)$ as the ratio of local generation capacity over the peak net demand and change ρ from 20% to 100%. The result is shown in Fig. 5. The experiment in this part is carried out with the data from July (one billing cycle).

Observations. The results for PA-OFFLINE and PO-Online algorithms follow the intuition that more local capacity brings more cost reduction. For **BED**, however, we observe that the cost reduction increases when ρ increases from 20% to 60%, and decreases as ρ continues to increase from 60% to 100%. As we discussed in Sec. III-D, there exists a critical local generation capacity $\tilde{C}$ beyond which the peak charge and the overall cost will not decrease further. In Fig. 6, we report the peak grid demand $\max v(t)$ in one month versus ρ just for PA-OFFLINE algorithm. Results show that the peak value of $v(t)$ does not decrease as ρ increases from 60% to 100%, evincing that $\tilde{C}$ is about 60% of the maximum demand in this case. As discussed in Sec. III-D, $\tilde{C}$ can be computed by solving problem **FS-PAED** in an offline manner.

However, the online algorithm, without knowing $\tilde{C}$ and with the tendency of reducing the peak charge by using more expensive local generation, will try to exploit the entire local generation capacity until the cost-benefit break-even point is reached, which turns out to be less economic and deviate from the offline optimal. As a result, for the online algorithm, larger capacity may incur higher operating cost, as shown in Fig. 5.

This experiment, together with the discussions in Sec. III-D, show that it is important for the microgrid operator to set the local generation capacity right at $\tilde{C}$ to cope with online algorithms to achieve maximum cost reduction. A possible way to set $\tilde{C}$ is to use the historical data as the input to the offline algorithm and obtain the critical capacity.

*E. The Performance of **RED***

Purpose. In this part, we compare the empirical performance of the deterministic online algorithm **BED** and randomized online algorithm **RED** under different local capacities. The cost of **RED** is computed by running the algorithm 1000 times and taking the average.

Observations. Even though **RED** is better than **BED** in terms of competitive ratio, it is not always the case empirically because the competitive ratio only characterizes the performance in the worst case. As we can see, when ρ is less than 80%, **BED** outperforms **RED** while the other way around if ρ is larger than 80%. Furthermore, when ρ increases from 80% to 100%, the performance of **BED** degrades drastically, while the cost reduction of **RED** almost remains the same. This observation indicates that, to ensure that **BED** has good performance, we need to carefully determine the local capacity but additional local capacity will not harm **RED** much, which can be viewed as another advantage of **RED**.

F. Empirical Evaluations Using Traces from a Real-world Small-scale Microgrid

Purpose. In this simulation, we replace the previous trace with a new one, which is from a test-bed building at College of Engineering Center for Environmental Research and Technology of UC Riverside and spans three months from May to July. The building has 20 office rooms, 2 conference rooms, one large open area with cubicles, and 7 other miscellaneous rooms. The building HVAC system consists of 16 packaged rooftop units. In addition to its small scale, the building is connected to solar PV and several charging stations, both of which introduce additional demand uncertainties. As a result, the demand fluctuates more than the previous data set we use. The simulation result is shown in Fig 8.

Observations. On this new data set, the cost reduction is more significant (at least 40% for the offline case) than the previous results and will increase with larger peak price p_m. This result indicates that peak-aware scheduling is more beneficial with more fluctuating demand and larger peak prices.

V. Related Work

Microgrid is attracting substantial attention from both academic and industrial communities due to its economic and environmental benefits, evidenced by a number of real-world pilot microgrid projects [21].

With the penetration of renewable energy in microgrids, conventional economic dispatching approaches based on accurate demand prediction for power grid [3] are not applicable as the local demand is highly uncertain and is hard to predict accurately. Online convex optimization [4] and Lyapunov optimization [22] are popular approaches to design online algorithms in face of uncertainty of future demand. In recent years, competitive online algorithm design is advocated by researchers to design online algorithms with strong worst-case performance guarantee for power system operation. Examples include the microgrid unit commitment and economic dispatching algorithm under the volume charging model [5], EV charing algorithm [23], dynamic provisioning of data centers [24], *etc*.

The peak-based charging model has been considered in the cost minimization problem for data centers in [8], [25] and for content delivery network in [26]. In the microgrid scenario, distributed energy storage scheduling [27], demand response [28]–[30], HVAC controlling for buildings [31], and climate control for storage systems [32] are also studied with peak-charging taken into consideration, assuming prediction of future input with certain level accuracy. In contrast, we design competitive online algorithms that does not rely on prediction of future input and achieves strong worst-case performance guarantee. We summarize the main differences between existing literatures and our work in Table I.

The special case **FS-PAED**^{k} of the economic dispatching problem in this paper can be considered as a generalization of the classic Bahncard problem [14], in the sense that the 'discounted price' is time-varying. The Bahncard problem and its solutions have also found application in the instance acquisition problem of cloud computing [35].

VI. Discussion and Future Work

In this paper, we devised online economic dispatching algorithms for microgrids, with peak charging model taken into account. We developed both deterministic and randomized online algorithms with best possible competitive ratios following a divide-and-conquer approach. In addition to sound theoretical performance guarantees, the empirical evaluations based on real-world traces also corroborated our claim on the importance of peak-awareness in scheduling and the merit of our algorithms.

Demand response and energy storage management can be used to effectively "shave the peak" of the demand so that the peak-charging can be reduced. Our work and algorithm are orthogonal to demand response and energy storage management, in the sense that we focus on orchestrating the local and external supply to further cut short the peak-charging. Thus one can apply demand response, energy storage management, and our algorithm to reduce the peak-charging by optimizing both the demand and supply.

TABLE I
Comparisons between most related literature on microgrid economic dispatching and peak-charging aware operations in power system.

Reference	Topic	Energy supply and storage components	Considering peak (demand) charge?	Relying on prediction?	With performance guarantee?
[32]	Climate control	Grid	Yes	Yes	No
[29]	Demand response	Grid and battery	Yes	Yes	No
[30]	Economic dispatching	Grid, local generator and battery	Yes	Yes	No
[31]	HVAC control	Grid	Yes	Yes	No
[33]	Economic dispatching	Grid, local generator and battery	No	Yes	No
[5]	Economic dispatching	Grid and local generator	No	No	Yes
[34]	Economic dispatching	Grid and battery	No	No	Yes
This work	Economic dispatching	Grid and local generator	Yes	No	Yes

An interesting future direction is to study the microgrid economic dispatching problem under accurate or noisy prediction of future demand and renewable generation, and characterize the averaged performances with respect to different stochastic patterns. Furthermore, it also deserves effort to extend the results in this paper to the scenarios with heterogeneous local generators and to the settings considering demand response and energy storage systems.

References

[1] "navigant research," internet: http://www.navigantresearch.com.

[2] S. A. Kazarlis, A. Bakirtzis, and V. Petridis, "A genetic algorithm solution to the unit commitment problem," *IEEE Trans. Power Syst.*, vol. 11, no. 1, pp. 83–92, 1996.

[3] Z.-L. Gaing, "Particle swarm optimization to solving the economic dispatch considering the generator constraints," *IEEE Trans. Power Syst.*, vol. 18, no. 3, pp. 1187–1195, 2003.

[4] B. Narayanaswamy, V. K. Garg, and T. Jayram, "Online optimization for the smart (micro) grid," in *Proc. ACM e-Energy*, 2012, pp. 1–19.

[5] L. Lu, J. Tu, C.-K. Chau, M. Chen, and X. Lin, "Online energy generation scheduling for microgrids with intermittent energy sources and co-generation," in *Proc. ACM SIGMETRICS*, 2013, pp. 53–66.

[6] "pacific gas and electric company, rate information," internet: http://www.pge.com/nots/rates/tariffs/rateinfo.shtml.

[7] "duke energy kentucky, electric tariff," internet: http://www.duke-energy.com/rates/kentucky/electric.asp.

[8] H. Xu and B. Li, "Reducing electricity demand charge for data centers with partial execution," in *Proc. ACM e-Energy*, 2014, pp. 51–61.

[9] Y. Zhang, M. H. Hajiesmaili, S. Cai, M. Chen, and Q. Zhu, "Peak-aware online economic dispatching for microgrids," *arXiv preprint arXiv:1508.03996*, 2015.

[10] Y. Zhang, M. H. Hajiesmaili, and M. Chen, "Peak-aware online economic dispatching for microgrids," in *Proc. ACM e-Energy*, 2015, pp. 27–36.

[11] E. Fanone, A. Gamba, and M. Prokopczuk, "The case of negative day-ahead electricity prices," *Energy Economics*, vol. 35, pp. 22–34, 2013.

[12] A. Vuorinen, *Planning of optimal power systems*. Ekoenergo Oy Espoo, Finland, 2007.

[13] A. R. Karlin, M. S. Manasse, L. Rudolph, and D. D. Sleator, "Competitive snoopy caching," *Algorithmica*, vol. 3, no. 1-4, pp. 79–119, 1988.

[14] R. Fleischer, "On the bahncard problem," *Theoretical Computer Science*, vol. 268, no. 1, pp. 161–174, 2001.

[15] "california commercial end-use survey," internet: http://capabilities.itron.com/CeusWeb.

[16] "national renewable energy laboratory," internet: http://wind.nrel.gov.

[17] T. Lu, M. Chen, and L. L. Andrew, "Simple and effective dynamic provisioning for power-proportional data centers," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 6, pp. 1161–1171, 2013.

[18] J. W. Taylor, L. M. de Menezes, and P. E. McSharry, "A comparison of univariate methods for forecasting electricity demand up to a day ahead," *International Journal of Forecasting*, vol. 22, no. 1, pp. 1–16, 2006.

[19] T. Zhu, A. Mishra, D. Irwin, N. Sharma, P. Shenoy, and D. Towsley, "The case for efficient renewable energy management in smart homes," in *Proc. ACM BuildSys*, 2011, pp. 67–72.

[20] J. Mattingley, Y. Wang, and S. Boyd, "Receding horizon control," *IEEE Control Systems Magazine*, vol. 31, no. 3, pp. 52–65, 2011.

[21] M. Barnes, "Real-world microgrid-an overview," in *IEEE International Conference on System of Systems Engineering*, 2007.

[22] Y. Huang, S. Mao, and R. Nelms, "Adaptive electricity scheduling in microgrids," in *Proc. IEEE INFOCOM*, 2013, pp. 1142–1150.

[23] S. Zhao, X. Lin, and M. Chen, "Peak-minimizing online EV charging." in *Proc. Allerton*, 2013, pp. 46–53.

[24] M. Lin, A. Wierman, L. L. Andrew, and E. Thereska, "Dynamic right-sizing for power-proportional data centers," *IEEE/ACM Trans. Netw.*, vol. 21, no. 5, pp. 1378–1391, 2013.

[25] C. Wang, B. Urgaonkar, Q. Wang, and G. Kesidis, "A hierarchical demand response framework for data center power cost optimization under real-world electricity pricing," in *Proc. IEEE MASCOTS*, 2014, pp. 305–314.

[26] M. Adler, R. K. Sitaraman, and H. Venkataramani, "Algorithms for optimizing the bandwidth cost of content delivery," *Computer Networks*, vol. 55, no. 18, pp. 4007–4020, 2011.

[27] A. Mishra, D. Irwin, P. Shenoy, and T. Zhu, "Scaling distributed energy storage for grid peak reduction," in *Proc. ACM e-Energy*, 2013, pp. 3–14.

[28] S. Tang, Q. Huang, X.-Y. Li, and D. Wu, "Smoothing the energy consumption: Peak demand reduction in smart grid," in *Proc. IEEE INFOCOM*, 2013, pp. 1133–1141.

[29] S. Paudyal, C. Canizares, K. Bhattacharya *et al.*, "Optimal operation of industrial energy hubs in smart grids," *IEEE Trans. Smart Grid*, vol. 6, no. 2, pp. 684–694, 2015.

[30] D. Zhang, N. Shah, and L. G. Papageorgiou, "Efficient energy consumption and operation management in a smart building with microgrid," *Energy Conversion and Management*, vol. 74, pp. 209–222, 2013.

[31] T. X. Nghiem and G. J. Pappas, "Receding-horizon supervisory control of green buildings," in *Proc. Amer. Control Conf.*, 2011, pp. 4416–4421.

[32] M. C. Bozchalui, C. Cañizares, K. Bhattacharya *et al.*, "Optimal operation of climate control systems of produce storage facilities in smart grids," *IEEE Trans. Smart Grid*, vol. 6, no. 1, pp. 351–359, 2015.

[33] M. Motevasel and A. R. Seifi, "Expert energy management of a micro-grid considering wind energy uncertainty," *Energy Conversion and Management*, vol. 83, pp. 58–72, 2014.

[34] C.-K. Chau, G. Zhang, and M. Chen, "Cost minimizing online algorithms for energy storage management with worst-case guarantee," *To appear in IEEE Trans. Smart Grid.*

[35] W. Wang, B. Li, and B. Liang, "To reserve or not to reserve: Optimal online multi-instance acquisition in iaas clouds," in *Proc. USENIX ICAC*, 2013.

[36] A. C.-C. Yao, "Probabilistic computations: Toward a unified measure of complexity," in *Proc. IEEE FOCS*, 1977, pp. 222–227.

Appendix

A. A Proposition to Prove Theorem 1 and 2

To facilitate our analysis, we provide Proposition 1 to characterizes the ratio between the online cost and the offline optimal cost with respect to different s and σ. Recall that different s and σ characterize different online algorithms and inputs.

Proposition 1: The ratio between the cost of a deterministic online algorithm with parameter s and the offline optimal cost,

denoted by $h(\mathcal{A}_s,\sigma)$, is given by:
when $\sigma \le 1$,

$$h(\mathcal{A}_s,\sigma) = \begin{cases} 1, & \text{if } s > \sigma, \\ 1 + \frac{1-\sigma+s}{\sigma}(1-\beta), & \text{otherwise;} \end{cases} \quad (5)$$

when $\sigma > 1$,

$$h(\mathcal{A}_s,\sigma) = \begin{cases} 1 + \frac{(\sigma-1)(1-\beta)}{(\sigma-1)\beta+1}, & \text{if } s > \sigma, \\ 1 + \frac{s(1-\beta)}{(\sigma-1)\beta+1}, & \text{otherwise.} \end{cases} \quad (6)$$

Proof: We denote the number of time slots with demand 1 by T, and the number of time slot using the local generator according to the online algorithm by T^s. The following relations are used in the derivation.

$$\sigma p_m = \sum_{t=1}^{T}(p_g - p_e(t)) \le T(p_g - p_e^{\min}) \quad (7a)$$

$$\sum_{t=1}^{T^s}(p_g - p_e(t)) \le s p_m \quad (7b)$$

We compute the ratios in the following different cases.

$\triangleright$ **Case 1:** $\sigma \le 1$. The optimal offline solution is always using the local generator and the cost is $\text{Cost}_{\text{off}} = Tp_g$.

Case 1.1: $s > \sigma$. In this case, the online algorithm will not turn to the grid before the input ends. Therefore, the online cost is exactly the same as the offline cost, thereby the ratio is 1.

Case 1.2: $s \le \sigma$. It turns out that there is a critical time slot T^s that for all $1 \le t \le T^s$, the online algorithm uses the local generator and for time slots $T^s < t \le T$, it turns to the grid, thereby we have $\text{Cost}_{\text{on}} = T^s p_g + \sum_{t=T^s+1}^{t=T} p_e(t) + p_m$. Hence, we get the following ratio:

$$\begin{aligned} R(s,z) &= \frac{T^s p_g + \sum_{t=T^s+1}^{t=T} p_e(t) + p_m}{Tp_g} \\ &= 1 + \frac{\sum_{t=1}^{T^s}(p_g - p_e(t)) - \sum_{t=1}^{t=T}(p_g - p_e(t)) + p_m}{Tp_g} \\ &\overset{(E_1)}{\le} 1 + (1-\sigma+s)\frac{p_m}{Tp_g}, \\ &\overset{(E_2)}{\le} 1 + \frac{1-\sigma+s}{\sigma}\frac{p_g - p_e^{\min}}{p_g} \\ &= 1 + \frac{1-z+s}{z}(1-\beta) \end{aligned}$$

Inequality (E_1) is due to (7a), (7b) and (E_2) is due to the fact that $zp_m = \sum(p_g - p_e(t)) \le T(p_g - p_e^{\min})$.

$\triangleright$ **Case 2:** $\sigma > 1$. The optimal offline solution is always acquiring the electricity from the grid and the cost is $\text{Cost}_{\text{off}} = \sum_{t=1}^{T} p_e(t) + p_m$.

Case 2.1: $s > \sigma$, In this case, the online algorithm always uses the local generator and thus the online cost is $\text{Cost}_{\text{on}} = Tp_g$. Hence, the ratio is as follows:

$$\begin{aligned} R(s,z) &= \frac{Tp_g}{\sum_{t=1}^{T} p_e(t) + p_m} \\ &= \frac{\sum_{t=1}^{T} p_e(t) + p_m + \sum_{t=1}^{T}(p_g - p_e(t)) - p_m}{\sum_{t=1}^{T} p_e(t) + p_m} \\ &\overset{(E_3)}{\le} 1 + \frac{(\sigma-1)p_m}{Tp_e^{\min} + p_m} \\ &= 1 + \frac{(\sigma-1)}{p_e^{\min}\frac{T}{p_m} + 1} \\ &\overset{(E_4)}{\le} 1 + \frac{(\sigma-1)}{p_e^{\min}\frac{\sigma}{p_g - p_e^{\min}} + 1} \\ &\overset{(E_5)}{=} 1 + \frac{(\sigma-1)(1-\beta)}{(\sigma-1)\beta+1} \end{aligned}$$

Inequality (E_3) is by $p_e(t) \ge p_e^{\min}$ and $\sigma p_m = \sum_{t=1}^{T}(p_g - p_e(t))$; (E_4) is by $\frac{T}{p_m} \ge \frac{\sigma}{p_g - p_e^{\min}}$, which comes from (7a); we have the last equality (E_5) by substituting $\beta = \frac{p_e^{\min}}{p_g}$.

Case 2.2: $s \le \sigma$. Like case 1.2 here we have $T \ge T^s$. Therefore, the online algorithm uses the local generator for the first T^s time slots and turns to the grid afterwards. In this case, the online cost is $\text{Cost}_{\text{on}} = T^s p_g + \sum_{t=T^s+1}^{t=T} p_e(t) + p_m$, and the ratio is

$$\begin{aligned} R(z,s) &= \frac{T^s p_g + \sum_{t=T^s+1}^{t=T} p_e(t) + p_m}{\sum_{t=1}^{T} p_e(t) + p_m} \\ &= 1 + \frac{\sum_{t=1}^{T^s}(p_g - p_e(t))}{\sum_{t=1}^{T} p_e(t) + p_m} \\ &\overset{(E_6)}{\le} 1 + \frac{sp_m}{Tp_e^{\min} + p_m} \\ &\overset{(E_7)}{\le} 1 + \frac{s}{p_e^{\min}\frac{\sigma}{p_g - p_e^{\min}} + 1} \\ &\overset{(E_8)}{=} 1 + \frac{s(1-\beta)}{(\sigma-1)\beta+1} \end{aligned}$$

Inequality (E_6) is by $p_e(t) \ge p_e^{\min}$ and (7b); (E_7) is by $\frac{T}{p_m} \ge \frac{\sigma}{p_g - p_e^{\min}}$, which comes from (7a); (E_8) is obtained by substituting $\beta = \frac{p_e^{\min}}{p_g}$.

The proof is completed. ■

B. Proof of Theorem 1

Proof: The best deterministic online algorithm with smallest **CR** can be obtained by solving

$$\min_s \max_\sigma h(\mathcal{A}_s,\sigma). \quad (8)$$

The problem is non-convex and thus challenging on the first sight. However, given a deterministic online algorithm $\mathcal{A}_s$, it turns out the worst cost ratio is obtained when $\sigma = s$, in which case the online algorithm pays for the peak-charge premium but there is no net demand to serve anymore. This can also be obtained by studying the property of $h(\mathcal{A}_s,\sigma)$. Thus we have

$$\max_\sigma h(\mathcal{A}_s,\sigma) = h(\mathcal{A}_s,s) = \begin{cases} 1 + \frac{1}{s}(1-\beta), & \text{if } s \le 1, \\ 1 + \frac{s(1-\beta)}{(s-1)\beta+1}, & \text{otherwise.} \end{cases}$$

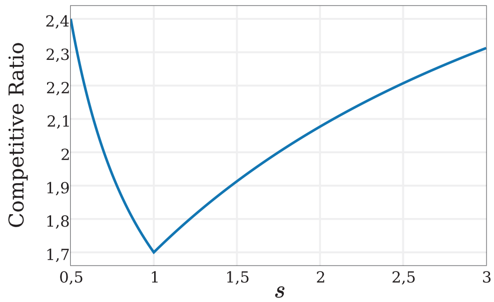

Fig. 9. Competitive ratio of $\mathcal{A}_s$ as a function of s, with $\beta = 0.3$.

Leveraging this observation, the problem in (8) can be solved easily by studying the extreme points of the two functions

To visualize how the competitive ratio varies as s changes, we plot the competitive ratio for different values of s in Fig. 9 for the case where $\beta = 0.3$. ■

C. Proof of Theorem 2

Proof: Recall that

$$f^*(s) = \begin{cases} \frac{e^s}{e-1+\beta}, & \text{when } s \in [0,1]; \\ \frac{\beta}{e-1+\beta}\delta(0), & \text{when } s = \infty; \\ 0, & \text{otherwise.} \end{cases}$$

When $\sigma \leq 1$,

$$\begin{aligned} \int_s h(s,\sigma) f^*(s) ds &= \int_0^\sigma (1 + \frac{1-\sigma+s}{\sigma}(1-\beta)) \frac{e^s}{e-1+\beta} ds \\ &+ \int_\sigma^1 \frac{e^s}{e-1+\beta} ds + \frac{\beta}{e-1+\beta} \\ &= \frac{\beta}{e-1+\beta} + \int_0^1 \frac{e^s}{e-1+\beta} ds \\ &+ \int_0^\sigma \frac{1-\sigma+s}{\sigma}(1-\beta) \frac{e^s}{e-1+\beta} ds \\ &= 1 + \int_0^\sigma \frac{1-\sigma+s}{\sigma}(1-\beta) \frac{e^s}{e-1+\beta} ds \\ &= 1 + \frac{1-\beta}{e-1+\beta} \\ &= \frac{e}{e-1+\beta} \end{aligned}$$

When $\sigma > 1$,

$$\begin{aligned} \int_s h(s,\sigma) f^*(s) ds &= \int_0^1 (1 + \frac{1-\sigma+s}{\sigma}(1-\beta)) \frac{e^s}{e-1+\beta} ds \\ &+ \int_1^\sigma \frac{e^s}{e-1+\beta} ds \\ &+ (1 + \frac{(\sigma-1)(1-\beta)}{(\sigma-1)\beta+1}) \frac{\beta}{e-1+\beta} \\ &= 1 + \frac{1}{e-1+\beta} \frac{1-\beta}{(\sigma-1)\beta+1} (1 + (\sigma-1)\beta) \\ &= 1 + \frac{1-\beta}{e-1+\beta} \\ &= \frac{e}{e-1+\beta} \end{aligned}$$

We observe that the value of $\int_s h(s,\sigma) f^*(s) ds$ has nothing to do with σ, then

$$\max_\sigma \int_s h(s,\sigma) f^*(s) ds = \frac{e}{e-1+\beta}.$$

Next we will provide Lemma 2 to prove that no other randomized online algorithm can achieve a smaller competitive ratio.

***Lemma** 2:* For any randomized online algorithm $\mathcal{A}_f$ for problem **FS-PAED**^{k}, we have

$$\mathbf{CR}(\mathcal{A}_f) \geq \frac{e}{e-1+\beta}.$$

Proof: The idea is to choose a randomized input, denoted by $g(\sigma)$, and compute the ratio between the online cost and offline optimal cost by the best deterministic online algorithm for this input. Yao's Principle [36] says that the computed ratio is a lower bound for any randomized online algorithm. The particular distribution we use is given by

$$g^*(\sigma) = \begin{cases} \frac{e}{e-1+\beta}\sigma e^{-\sigma}, & \text{when } \sigma \in [0,1], \\ \frac{e}{e-1+\beta}[(\sigma-1)\beta+1]e^{-\sigma}, & \text{otherwise.} \end{cases} \quad (9)$$

When $s \leq 1$,

$$\begin{aligned} \int_\sigma h(s,\sigma) g^*(\sigma) d\sigma &= \int_0^s g^*(\sigma) d\sigma + \int_s^1 (1 + \frac{(1-\sigma)s}{z}) g^*(\sigma) d\sigma \\ &+ \int_1^{+\inf} (1 + \frac{s(1-\beta)}{(\sigma-1)\beta+1}) g^*(\sigma) d\sigma \\ &= 1 + \frac{e(1-\beta)}{e-1+\beta} \cdot \\ &\left[\int_s^1 (1-\sigma+s) e^{-\sigma} dz + \int_1^{+\inf} s e^{-\sigma} d\sigma\right] \\ &= 1 + \frac{e(1-\beta)}{e-1+\beta} e^{-1} \\ &= \frac{e}{e-1+\beta} \end{aligned}$$

When $s > 1$,

$$\begin{aligned} \int_\sigma h(s,\sigma) g^*(\sigma) d\sigma &= \int_0^1 g^*(\sigma) d\sigma \\ &+ \int_1^s (1 + \frac{(\sigma-1)(1-\beta)}{(\sigma-1)\beta+1}) g^*(\sigma) d\sigma \\ &+ \int_s^{+\inf} (1 + \frac{s(1-\beta)}{(\sigma-1)\beta+1}) g^*(\sigma) d\sigma \\ &= 1 + \frac{e(1-\beta)}{e-1+\beta} \cdot \\ &\left[\int_1^s (\sigma-1) e^{-\sigma} d\sigma + \int_s^{+\inf} s e^{-\sigma} d\sigma\right] \\ &= 1 + \frac{e(1-\beta)}{e-1+\beta} e^{-1} \\ &= \frac{e}{e-1+\beta} \end{aligned}$$

Similarly, the value of $\int_\sigma h(s,\sigma) g^*(\sigma) d\sigma$ has nothing to do with s, then

$$\min_s \left[\int_\sigma h(s,\sigma) g^*(\sigma) d\sigma\right] = \frac{e}{e-1+\beta}.$$

■

Then we can establish Lemma 2 and the proof for Theorem 2 is completed. ■

D. Proof of Theorem 3

Proof: Firstly, if some energy demands of the input exceed the capacity constraints, we can construct a new input by sequently removing the demand exceeding the capacity and the following demands in the same layer (like the first layer from time slot 2 in Fig 2). Then compared with the original input, the online cost and offline cost are reduced by the same amount, which will lead to a larger competitive ratio. Then we only need to focus on the input whose demand is always smaller than the capacity.

Furthermore, due to the two properties of the algorithm described in the paragraph before Theorem 3, we can have $\max_t \sum_k v^k(t) = \sum_k \max_t v^k(t)$, and $\begin{cases} u(t) = \sum_k u^k(t) \\ v(t) = \sum_k v^k(t) \end{cases}$. Then $\mathsf{Cost}(\boldsymbol{u}, \boldsymbol{v}) = \sum_k \mathsf{Cost}(\boldsymbol{u}^k, \boldsymbol{v}^k)$. This property still holds for the offline cost. We denote $\tilde{r}$ as the competitive ratio for each layer, meaning

$$\mathsf{Cost}(\boldsymbol{u}^k, \boldsymbol{v}^k) \leq \tilde{r}\mathrm{Cost}_{\mathrm{off}}^k, \forall k.$$

Then by summing the above inequality over k, we can have

$$\mathsf{Cost}(\boldsymbol{u}, \boldsymbol{v}) \leq \tilde{r}\mathrm{Cost}_{\mathrm{off}}.$$

For **BED**, $\tilde{r} = 2 - \beta$ and for **RED**, $\tilde{r} = \frac{e}{e-1+\beta}$ for the randomized case, which establish the upper bound of the competitive ratios.

Furthermore, note that **FS-PAED**^{k} is a special case of **FS-PAED**. Since we cannot obtain smaller competitive ratios for **FS-PAED**^{k}, we cannot obtain smaller competitive ratios for **FS-PAED**.

The proof for Theorem 3 is completed. ■

Ying Zhang received his B.Eng degree from the Department of Electronic Engineering at Shanghai Jiao Tong University in 2013, and is currenty pursuing the Ph.D. degree in Department of Information Engineering at the Chinese University of Hong Kong.

His research interests include energy systems (e.g., microgrids and energy-efficient data centers) and distributed optimization.

Mohammad H. Hajiesmaili received his B.S. degree from Department of Computer Engineering at Sharif University of Technology, Tehran, Iran in 2007. He received his M.Sc. and Ph.D. degrees from the Electrical and Computer Engineering Department at the University of Tehran, Iran, in 2009 and 2014, respectively. He was a researcher at the School of Computer Science, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran, from 2008 to 2013, and a research assistant in the Institute of Network Coding, the Chinese University of Hong Kong, from 2013 to 2014. He is currently a postdoctoral researcher at the Department of Information Engineering, the Chinese University of Hong Kong. His research interests include energy systems and multimedia networking.

Sinan Cai enrolled in Xian Jiao Tong University at School of Electrical Engineering in 2012, majoring Electrical Engineering And Automation. He completed internship in China Electric Power Research Institute in 2014 and in Zhejiang Xutang Power Generation CO.Ltd in 2015 respectively. His research interests focus on Power System Analysis, Optimization & Control.

Minghua Chen (S'04 M'06 SM' 13) received his B.Eng. and M.S. degrees from the Dept. of Electronic Engineering at Tsinghua University in 1999 and 2001, respectively. He received his Ph.D. degree from the Dept. of Electrical Engineering and Computer Sciences at University of California at Berkeley in 2006. He spent one year visiting Microsoft Research Redmond as a Postdoc Researcher. He joined the Dept. of Information Engineering, the Chinese University of Hong Kong in 2007, where he is currently an Associate Professor. He is also an Adjunct Associate Professor in Institute of Interdisciplinary Information Sciences, Tsinghua University. He received the Eli Jury award from UC Berkeley in 2007 (presented to a graduate student or recent alumnus for outstanding achievement in the area of Systems, Communications, Control, or Signal Processing) and The Chinese University of Hong Kong Young Researcher Award in 2013. He also received several best paper awards, including the IEEE ICME Best Paper Award in 2009, the IEEE Transactions on Multimedia Prize Paper Award in 2009, and the ACM Multimedia Best Paper Award in 2012. He is currently an Associate Editor of the IEEE/ACM Transactions on Networking. His recent research interests include energy systems (e.g., microgrids and energy-efficient data centers), distributed optimization, multimedia networking, wireless networking, network coding, and delay-constrained communication.

Qi Zhu (M12) received the B.E. degree in Computer Sciences from Tsinghua University, Beijing, China, in 2003, and the Ph.D. degree in Electrical Engineering and Computer Sciences from the University of California, Berkeley in 2008. He is an Assistant Professor with the Department of Electrical and Computer Engineering in the University of California, Riverside (UCR). Prior to joining UCR, he was a Research Scientist with the Strategic CAD Labs at Intel Corporation from 2008 to 2011. His current research interests include model-based design and software synthesis for cyber-physical systems, cyber-physical system security, energy-efficient buildings and infrastructures, and system-on-chip design.He received the National Science Foundation (NSF) CAREER award in 2016. He received best paper awards at the Design Automation Conference (DAC) 2006, DAC 2007, and International Conference on Cyber-Physical Systems (ICCPS) 2013.