This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSUSC.2019.2904867, IEEE Transactions on Sustainable Computing

Two Level Colocation Demand Response with Renewable Energy

Huiting Xu[1], Xi Jin[2], Fanxin Kong[3], Qingxu Deng[1]
[1]School of Computer Science and Engineering, Northeastern University, China
[2]Shenyang Institute of Automation, Chinese Academy of Sciences, China
[3]Department of Computer and Information Science, University of Pennsylvania, United States
htingxu@gmail.com, jinxi@sia.cn, fanxink@seas.upenn.edu, dengqx@mail.neu.edu.cn

***Abstract*—Demand response is considered as a valuable functionality of the power grid and its potential impacts continue expanding with grid modernization. Colocation data centers (simply called colocation) are recognized as a notably promising resource for demand response due to their high power demand and remarkable potential in demand management. A major challenge of colocation demand response is the split incentive, that is, colocation operators desire demand response for financial compensation while tenants may not embrace demand response due to lack of incentives. Another key challenge is caused by renewable energy co-located with data centers. Demand response mechanisms overlooking the uncertainty of renewable would cause much inefficiency in terms of energy saving and economic aspects. Existing work considers the two challenges separately in the context of data centers. By contrast, this work jointly addresses them and specially studies mechanism design for colocation data centers in presence of co-located renewable.**

We propose a hierarchical demand response scheme, which is based on a new two-level market mechanism that results in a win-win situation for both parties, i.e., tenants who choose to reduce power demand obtain financial rewards from the operator, while the operator receives financial compensation from the electric power company due to its tenants' demand reduction. At each demand response period, the colocation operator solicits bids (amount of energy reduction) from tenants and tenants who choose to participate responds to the operator with their bids. The proposed mechanism provably converges to a unique equilibrium solution, and at the equilibrium, neither the operator or tenants can improve their individual economic performance by changing their own strategies. Further, we present a stochastic optimization based algorithm, which uses predictions of the co-located renewable to determine the colocation operator's best strategy. At the equilibrium, the algorithm has a provable economic performance guarantee in terms of the prediction error. We finally evaluate the designed mechanism via detailed simulations and the results show the efficacy and validate the theoretical analysis for the mechanism.

***Index Terms*—Demand response, data center, colocation, renewable, Stackelberg game, stochastic optimization**

I. Introduction

Demand response programs offer financial compensation to motivate consumers to adapt their power demand according to power supply conditions. For example, consumers may reduce their power loads in response to direct requests or price-peak warning signals from electric power companies. A key vision of the future smart grid is transiting from the paradigm of supply-follow-demand to the one of demand-follow-supply. Demand response is identified as the crucial functionality to help realize this vision [1].

This work studies a particularly promising resource for adopting demand response, which is colocation data centers (simply called colocation afterwards). On one hand, colocation is a favorable candidate for a demand response because it is critical data center segment requiring tremendous energy consumption. For example, colocation (e.g., those of Equinix) in the US consumed more than 14.5 billion kWh in 2011, which was about 4.8 times of the energy usage by large-scale cloud data centers (e.g., those of Google) [2]. Further, colocation has significant potential to allow flexible power demand management. The potential roots from the fact that server utilization varies a lot overtime. Actually, the average utilization is very low, only around 10-15%, and idle servers can take up to 60% of the peak power [2], [3]. There would be much power saving if adapting idle servers' operation status accordingly, for example, turning them off. One the other hand, it would be beneficial for colocation to participate into a demand response program. As to colocation, the energy cost takes up to 40% of the total cost of ownership [3], [4]. Adjusting its power consumption accordingly will bring significant financial compensation and thus offset the energy cost of colocation.

There are two key challenges to unleash the potential of colocation demand response. The first one is the split incentive, which is that colocation operators desire demand response while tenants hesitate about demand response. The operators' desire results from the benefit of financial compensation received by coloction if participating demand response. The tenants' hesitation stems from that existing pricing mechanisms used by colocation provide no incentives and thus can hardly motivate tenants to actively adapt their power usage [3], [5]. This challenge roots from the fact that colocation operators have little control over tenants' servers. Colocation operators (e.g., Equinix) mainly provide facilities support such as network, power and cooling, and rent out physical space to tenants for housing their own servers. Each tenant individually operates their own servers. Note that the operator of an owner-operated data center (e.g., those of Google), where servers and facilities belong to the single owner, has full control over them. This difference eliminates the feasibility of applying existing work such as [6]–[12], which study demand response for owner-operated data centers, to the case of colocation demand response.

The second challenge is induced by the co-located renewable energy, i.e., renewable co-located with a data center [13]. On one hand, colocation operators are increasingly integrating renewable into their data centers' power supply given the need of sustainable development and the financial benefit due to the lower cost of renewable (in the long run). For example, Equinix has increased the percentage of their renewable energy from 30% to 40% at the end of 2015 [14]. Further, thirty percent of the power supply for overall colocation would come from renewable sources by 2025 [15]. On the other hand, renewable such as wind and solar power is highly intermittent and variable. Their effective power output fluctuates and varies a lot over time due to the dependence on the external conditions (e.g., solar irradiation and wind speed) [16]–[18]. Both energy saving and economic performance would be much compromised if a demand response design overlooks this uncertainty of renewable.

Existing work addresses the above two challenges separately in the context of data centers. For example, research efforts including [3], [5], [19]–[22] are focused on the split incentive for colocation data centers and propose incentive mechanisms (i.e., pay or reward tenants) to motivate tenants to adapt their power consumptions. Existing work such as [23]–[25] and those surveyed in [13] is confined to renewable integration for owner-operated data centers and presents planning, scheduling, and routing approaches to accommodate the renewable's uncertainty. Differently, this work jointly addresses the two challenges and studies mechanism design for colocation data centers in presence of co-located renewable.

We propose to utilize hierarchical demand response, where tenants interact with colocation operators while colocation operators respond to electric power companies. Designing such a market mechanism that well addresses the above challenges, however, is a non-trivial task because of several difficulties as follows. First, both colocation operators and tenants may want to maximize their own benefits. Thus, a suitable market design should be distributed and allows both parties to carry out independent decision-making. Second, from an algorithmic perspective, a well-designed mechanism should have robust performance bound in terms of the impact of renewable uncertainty on the market.

To address these difficulties, we design a two-level market mechanism for colocation demand response (Fig. 1). Tenants who reduce their power demand can obtain financial rewards from colocation operators, while colocation operators receive financial compensation from the electric power company due to tenants' demand reduction. The market mechanism has a hierarchical decision-making structure. The colocation operator leads the market by deciding the amount of demand reduction that best responds to the requests from the electric power company. Tenants choose to follow the operator's actions by independently submitting to the operator with bids that best satisfy their own economic gains.

To demonstrate the feasibility and benefits of the proposed market mechanism, we study and analyze it using the framework of Stackelberg games [26]. Through rigorous game-theoretic analysis, we prove that 1) the market level among tenants converges to a unique Nash equilibrium where no tenant can improve its economic performance by changing only its own bid; and 2) the whole market converges to a unique Stackelberg equilibrium where both the colocation operator and tenants achieve the best economic gain in presence of the Nash equilibrium. Another key contribution is to demonstrate that the impact of renewable uncertainty on the Stackelberg equilibrium is bounded. For tenants, we show that the uncertainty does not affect the Nash equilibrium. For the colocation operator, to obtain good economic performance on average, we present a stochastic optimization based algorithm given the estimation of the likelihood of renewable. The algorithm has a provable performance guarantee in terms of the prediction errors, which is independent on any specific distribution of the prediction error. To be specific, our main contributions are summarized as follows.

- We newly study hierarchical demand response for colocation data centers in presence of co-located renewable, and specially, propose a distributed two-level market mechanism that creates a win-win situation for colocation operators and tenants. We further present a stochastic optimization based algorithm to optimize economic performance for the colocation at the equilibrium given the prediction of colocated renewable.
- Using theoretical analysis, we demonstrate that 1) the proposed market mechanism provably converges to the equilibrium solutions; 2) the renewable's uncertainty has provable bounded impact on the equilibriums.
- Through extensive simulations, we show the efficacy of the mechanism as well as validate the theoretical results under various settings with regard to both economic and algorithmic performance.

The rest of the paper is organized as follows. Section II provides the system model. Section III proposes the hierarchical demand response mechanism and analyzes its performance. Section IV presents a stochastic algorithm for incorporating renewable and bounds its performance. Section V evaluates the proposed mechanism. Section VI concludes the paper and points out the future work.

II. The System Model

Fig. 1 illustrates the schematic of hierarchical demand response, where level 1 is that the colocation operator interacts with its tenants while level 2 is that the operator responds to the electric power company. The scenario we consider in this paper is that an electric power company requests power demand reduction from a colocation during a time period (e.g., one hour). First, the electric power company initializes the demand response period by notifying the colocation operator with the demand reduction request. Second, the operator solicits bids from its tenants after receiving the demand reduction request after receiving the request. To appeal tenants to participate in reducing their power consumption, the

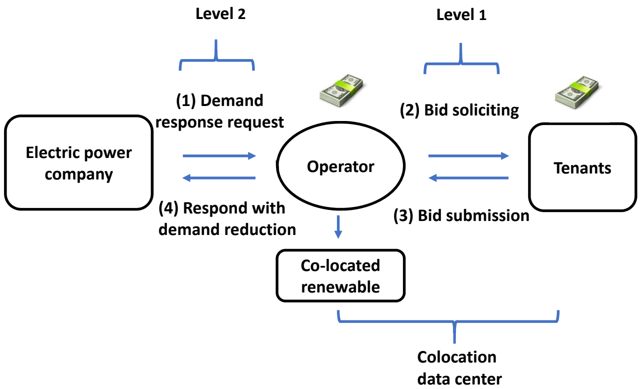


Fig. 1. Two-level market design for colocation demand response. Level 2 is that the colocation operator responds to the electric power company. Level 1 is that the operator interacts with its tenants.

operator provides a reward for each tenant who contributes to energy reduction. Tenants decide whether to decrease their power according to their own conditions (e.g., workload or timing requirements). Third, the operator decides how much power reduction is responded to the demand response signal. Finally, according to the operator's response, the electric power company offers financial compensation to the operator. In the following, we consider one demand response period and the proposed method can be repeatedly applied to periods [5], [27].

There are many ways that tenants can use to reduce energy, for example, lowering the frequencies of server's CPU [28], or transferring work loads to other data center [11]. The best and easy way to implement energy reduction is to turn off idle severs [5]. The tenants who respond and save energy will get a financial benefit from the operator. This financial benefit is limited so tenants have to compete with one another to maximise their own benefits. Both the operator and the tenants make their strategies due to their own financial benefits. In the rest of this section, we present how to model these behaviors of the operator and tenants. Note that although these following modeling approaches seem restricted, it is widely used in the power market literature, e.g. [29]–[32].

Table I lists notations used in this paper.

A. Colocation Model

We assume that there are N tenants in the colocation. We use q_i to denote the quantity of electricity that tenant $i \in [1, N]$ is willing to decrease. If tenant i decides not to respond to the operator, q_i is set to 0. The total energy reduction by all the tenants is thus

$$E = \sum_{i=1}^{N} q_i. \tag{1}$$

Each tenant i incurs a cost $C_i(q_i)$, when he/she reduces an amount of q_i energy. We assume that function $C_i(\cdot)$ is continuous, increasing, strictly convex, and with $C_i(0) = 0$. This cost includes switching cost, delay cost, and management cost [5], [33]. We uniformly model the total cost as convex

TABLE I
NOTATIONS USED IN THIS PAPER.

Notation	Description
N	total number of tenants
q_i	quantity of energy that tenant i saved
E	total energy reduction of tenants
$C_i(\cdot)$	tenant's cost function
p	financial reward price from the operator
b_i	the bid submitted by tenant i
$\mathbf{b}$	vector of tenants' bids
B	sum of bids
B_{-i}	sum of bids except tenant i
u_i	utility of tenant i
W_i	reward from operator
$K(\cdot)$	financial compensation from electric power company
$C_o(\cdot)$	cost function of the operator
ω	actual renewable generation
$\tilde{\omega}$	predicted renewable generation
ε	the prediction error of renewable generation
$\mathbb{E}[\varepsilon]$	mean of ε
$\mathbb{V}[\varepsilon]$	variance of ε
M	maximum amount of renewable generation
η	competitive ratio

function, because that can capture many common servers [33], [34]. The rationale of the convexity of $C_i()$ is as follows. For example, consider modeling all servers of a tenant as a M/M/n queue. The service delay has a convex relation with the number of active servers [35]. If we consider the delay as a cost of turning-off servers (thus energy reduction), the cost has a convex relation with energy reduction. In practice, $C_i()$ can be also empirically measured by observing the system. According to principle of diminishing returns, as the unite amount of energy tenants save increase, the related cost will increase more and more fast [36], [37]. In reality, when tenants turn off the server, the assigning jobs to each working server increase. At this time, if tenants try to turn off more server, the delay cost will increase more than the first time he/she turns off the server. We further assume the truthfulness of tenants, i.e., each tenant reveals a true $C_i()$ that provides reasonable cost for his/her power reduction.

For the operator side, we consider a market mechanism for the saving energy allocation, based on supply function bidding [38]. The operator offers financial reward to tenants to incentive them to reduce energy. He/she chooses a reward price p to clear the market (p can be dollars). This price p can stimulate the enthusiasm of tenants saving energy. It's quite easy to understand that the higher price can appeal more tenants to join in. We consider all tenants are price-taking customers in colocation. We assume that the relation between q_i and p is

$$q_i = b_i \times p, \tag{2}$$

each tenant submits a bid to operator as the response for reward price. We use the linear bid Eqn. (2) just for the ease of writing. Other functions $f(p)$ that are continuous, increasing, strictly convex and with $f(0) = 0$ can also apply here. Here b_i is the bid submitted by tenant i. We assume that each tenant's "supply" function is parameterized by a single parameter b_i

and it reflects the saving capability of tenant i. The total energy reduction by all tenants is

$$E = \sum_{i=1}^{N} b_i \times p. \tag{3}$$

We further use the following denotations to denote the sum bid (B) of all tenants and the sum bid (B_{-i}) except tenant i.

$$B = \sum b_i, (i \in N),\ B_{-i} = B - b_i, (i \in N). \tag{4}$$

B. Utility of Tenants

The main concern of tenants is the utility for energy saving. This reward is based on the reward offered by operator and the cost incurred by energy reduction. We use W_i to denote the reward provided by the operator and u_i to denote the utility of tenant i.

$$u_i = W_i - C_i(q_i),\ W_i = p \times q_i, \tag{5}$$

W_i is decided by the reward price and energy reduction from tenant i. If tenants want to participate into the reward competition, they need to decide their bids (b_i) and submit them to the operator.

C. Utility of Colocation Operator

The main concern of operator is potential of energy saving in colocation during the whole period. After collecting bids from all tenants, the colocation operator estimates the total amount of energy reduction and further estimates the utility for demand response. The utility is according to two key factors. One is the financial compensation from the electric power company due to responding to the demand response signal. The other one is the total cost for paying tenants who participate in the energy saving. Here we ignore the renewable energy and consider it in section IV. We use $K(\cdot)$ to denote the financial compensation from the electric power company and $C_o(\cdot)$ to denote the cost for the operator. Then the operator's utility is defined as

$$R(E) = K(E) - C_o(E), \tag{6}$$

we assume the function $K(\cdot)$ is a continuously differentiable, non-decreasing and concave function. Similarly, this kind of utility modeling is widely used when designing schemes to network economics [39] [40] [41]. The operator's cost $C_o(E)$ is the sum reward offered to tenants, i.e., the market clearing price times the total power reduction from the tenants as shown in Eqn. (39).

III. The Hierarchical Market

From the above system model, colocation is divided into a two-level structure. The first level is the operator responses to the electricity company and also interacts with tenants. The second level is tenants decide and submit their bids. Both the operator and tenants seek financial benefits without any cooperation. Hence, this structure can be seen as a hierarchical noncooperative problem and can be analyzed as a Stackelberg game [26]. In this section, we mainly analyze the equilibrium status between both sides in colocation.

In a Stackelberg game, each player is rational and aims to maximize their own utility. There are two different types of players in such a game: leader(s) and followers. Leaders make the first move and decide their best strategy. Then the followers determine their best strategies accordingly. The best strategy means the strategy can maximize their own utility. Both leaders and followers determine their best strategies based on the other players' responses. The Stackelberg equilibrium is usually regarded as the solution to this game [26].

In this work, the colocation operator is the only leader and those tenants who want to participate into the energy saving are the followers. The operator first decides a reward price and communicates it to the tenants. After receiving the price, those tenants calculate their energy reduction and decide their bids. Then they submit their bids to the colocation operator. The operator collects the bids and then adjusts the reward price for the second time if the total rewards for tenants is beyond his/her budget. The tenants follow the operator's strategy and change their strategies. It is easy see that this is an iteration process. We prove that the iteration converges to an equilibrium. At the equilibrium, neither the operator nor the tenants can change their own strategies to improve their utility.

A. Nash Equilibrium among Tenants

Tenants compete for rewards offered by operator. We analyze the equilibrium among tenants. First, by manipulating Eqn.(1), (2), (3) and Eqn.(5), we have the utility of tenant i is

$$u_i(b_i, p) = b_i \times E^2 \div (B)^2 - C_i(b_i \times E \div B), \tag{7}$$

The derivative of Eqn. (7) is

$$\frac{\partial u_i}{\partial b_i} = \frac{E^2}{B^2}[\frac{B_{-i} - b_i}{B} - \frac{B_{-i}}{E} \times C_i'(b_i \times \frac{E}{B})]. \tag{8}$$

Using Eqn.(4), Eqn. (8) is equivalent to

$$\frac{\partial u_i}{\partial b_i} = \frac{E^2}{(B_{-i} + b_i)^2}[\frac{B_{-i} - b_i}{B_{-i} + b_i} - \frac{B_{-i}}{E} \times C_i'(b_i \times \frac{E}{B_{-i} + b_i})]. \tag{9}$$

Because $C_i(\cdot)$ is continuous, increasing, strictly convex and with $C_i(0) = 0$, we have

$$\frac{B_{-i}}{E} \times C_i'(b_i \times \frac{E}{B_{-i} + b_i}) > 0 \tag{10}$$

Let $l = b_i \times \frac{E}{B_{-i}+b_i}$, then we have

$$\frac{\partial l}{\partial b_i} = \frac{E \times B_{-i}}{(B_{-i} + b_i)^2} > 0 \tag{11}$$

From Eqn. (10) and Eqn. (11), we can see that $C_i'(b_i \times \frac{E}{B_{-i}+b_i})$ is increasing as b_i. The reason is that in Eqn.(9), the left part of the minus sign is no greater than one. There are two different cases for the right part of the minus sign. One is that it is no less than 1, and the other is that it is less than 1. For the first case, $\frac{\partial u_i^2}{\partial^2 b_i}$ is always less than zero. Thus utility

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSUSC.2019.2904867, IEEE Transactions on Sustainable Computing

u_i is maximized when $b_i = 0$ given B_{-i}. For the second case, the second derivative of u_i is

$$\frac{\partial u^2}{\partial^2 b_i} = -\frac{2E^2}{(B_{-i}+b_i)^3}[\frac{B_{-i}-b_i}{B_{-i}+b_i} - \frac{B_{-i}}{E}C_i'(\frac{b_i \times E}{B_{-i}+b_i})] + \frac{E^2}{(B_{-i}+b_i)^2}[-\frac{2B_{-i}}{(B_{-i}+b_i)^2} - \frac{B_{-i}}{E}C_i''(\frac{b_i \times E}{B_{-i}+b_i})\frac{B_{-i} \times E}{(B_{-i}+b_i)^2}]. \tag{12}$$

If $\frac{B_{-i}-b_i}{B_{-i}+b_i}$ is no less than $\frac{B_{-i}}{E}C_i'(\frac{b_i \times E}{B_{-i}+b_i})$, $\frac{\partial u^2}{\partial^2 b_i}$ is less than zero. That means u_i is concave and u_i is maximized when $\frac{\partial u}{\partial b_i}$ is equal to zero. That is

$$\frac{B_{-i}-b_i}{B_{-i}+b_i} - \frac{B_{-i}}{E}C_i'(\frac{b_i \times E}{B_{-i}+b_i}) = 0. \tag{13}$$

Thus the solution to Eqn. (13) maximizes utility u_i. If $\frac{B_{-i}-b_i}{B_{-i}+b_i}$ is less than $\frac{B_{-i}}{E}C_i'(\frac{b_i \times E}{B_{-i}+b_i})$, it is difficult to decide whether the second derivative of u_i is greater or equal to zero. But we know that the first derivative of u_i is less than zero, i.e., u_i increases as to b_i. Thus u_i is maximized when $b_i = 0$ given B_{-i} as mentioned above. Putting them together, we have

$$b_i^* = 0 \quad or \quad \frac{B_{-i}^* - b_i^*}{B_{-i}^* + b_i^*} - \frac{B_{-i}^*}{E}C_i'(\frac{b_i^* \times E}{B_{-i}^* + b_i^*}) = 0, \tag{14}$$

where b_i^* denotes the best bid (or strategy) for tenant i, and B_{-i}^* denotes the sum best bids except tenant i. Accordingly, we have the following theories.

Lemma 1: If b_i^* is the best strategy at a Nash Equilibrium game, then $b_i^* < B_{-i}^* = \sum b_j^* - b_i^* (i \in N, j \in N)$, and each tenant i reduce less than $\frac{E}{2}$ energy at the equilibrium.

Proof: When $b_i^* = 0$, this lemma holds. When Eqn. (13) is true, $\frac{B_{-i}^* - b_i^*}{B_{-i}^* + b_i^*}$ must be positive because the second term in Eqn. (13) is positive. So lemma 1 holds. ∎

Corollary 2: No Nash equilibrium exists when $|N| = 2$.

Proof: This corollary follows Lemma 1. ∎

Theorem 3: If $|N| > 2$, the game among tenants has a unique Nash Equilibrium.

Proof: To prove theorem 3, we first see the following convex optimization problem.

$$F_i(q_i) = [1 + \frac{q_i}{E - 2q_i}C_i(q_i)] - \int_0^{q_i}[\frac{E}{(E-2x_i)^2}C_i(x_i) + p]dx \tag{15}$$

The first derivative of Eqn. (15) is

$$F_i'(q_i) = (1 + \frac{q_i}{E - 2q_i})C_i'(q_i) - p \tag{16}$$

The second derivative of Eqn. (15) is

$$F_i''(q_i) = \frac{E}{(E-2q_i)^2}C_i'(q_i) + (1 + \frac{q_i}{E-2q_i})C_i''(q_i) > 0 \tag{17}$$

Thus we can see that $F_i(q_i)$ is strictly convex. When $|N| > 2$, the Nash equilibrium of the game among tenants solves the following convex optimization problem

$$\min_{0<q_i<\frac{E}{2}} \Sigma_{i=1}^N F_i(q_i), \tag{18}$$

$$s.t. \quad \Sigma_{i=1}^N q_i = E. \tag{19}$$

Because $F_i(q_i)(q_i \in [0, \frac{E}{2}])$ is strictly convex, the optimization problem Eqn. (18)(19) is a strictly convex problem and thus has a unique solution. According to the optimality condition [42], the unique solution q_i^* is determined by

$$F_i'(q_i^*)(q_i - q_i^*) \geq 0 \tag{20}$$

Further, the unique solution of the convex optimization problem Eqn. (18)(19) must satisfy the following conditions:

$$[p^* - (1 + \frac{q_i^*}{E - 2q_i^*})C_i'(q_i^*)](q_i - q_i^*) \leq 0, \tag{21}$$

$$\Sigma_{i=1}^N q_i^* = E, \tag{22}$$

$$p^* > 0. \tag{23}$$

With the Nash equilibrium in Eqn.(14), we have

$$[\frac{B_{-i}^* - b_i^*}{B_{-i}^* + b_i^*} - \frac{B_{-i}^*}{E}C_i'(\frac{b_i^* \times E}{B_{-i}^* + b_i^*})](b_i - b_i^*) \leq 0. \tag{24}$$

Then, using Eqn.(2) and Eqn.(3), we furthermore have

$$q_i^* = \frac{E \times b_i^*}{B_{-i}^* + b_i^*}, \tag{25}$$

$$B_{-i}^* = \frac{E \times b_i^*}{q_i*} - b_i^*. \tag{26}$$

Putting Eqn.(26) into Eqn.(24), we have the following deduction:

$$[\frac{\frac{E \cdot b_i^*}{q_i^*} - 2b_i^*}{\Sigma b_i^*} - \frac{\frac{E \cdot b_i^*}{q_i^*} - b_i^*}{E}C_i'(q_i^*)](b_i - b_i^*) \leq 0 \tag{27}$$

$$\Rightarrow b_i^*[\frac{E - 2q_i^*}{\Sigma b_i^* \cdot q_i^*} - \frac{E - q_i^*}{E \cdot q_i^*}C_i'(q_i^*)](b_i - b_i^*) \leq 0$$

$$\Rightarrow \frac{q_i^*}{b_i^*} \cdot b_i^*[\frac{E - 2q_i^*}{\Sigma b_i^* \cdot q_i^*} - \frac{E - q_i^*}{E \cdot q_i^*}C_i'(q_i^*)](b_i - b_i^*) \leq 0$$

$$\Rightarrow [\frac{E - 2q_i^*}{\Sigma b_i^*} - (1 - \frac{q_i^*}{E})C_i'(q_i^*)](b_i - b_i^*) \leq 0. \tag{28}$$

According to Lemma 1, $E - 2q_i^*$ is greater than zero. Eqn. (28) mutiplied by $\frac{E}{E-2q_i^*}$ can not change the inequality. Thus,

$$[\frac{E}{\Sigma b_i^*} - \frac{E - q_i^*}{E - 2q_i^*}C_i'(q_i^*)](b_i - b_i^*) \leq 0. \tag{29}$$

We furthermore insert Eqn.(3) into Eqn.(29), and thus have the following inequality

$$[p^* - (1 + \frac{q_i^*}{E - 2q_i^*})C_i'(q_i^*)](b_i - b_i^*) \leq 0, \tag{30}$$

$$\Rightarrow [p^* - (1 + \frac{q_i^*}{E - 2q_i^*})C_i'(q_i^*)](b_i \cdot q_i - q_i^*) \leq 0. \tag{31}$$

Because p^* is greater than zero and b_i is arbitrary, Eqn. (31) is equivalent to Eqn. (24). Moreover, Eqn.(31) is equivalent to

Eqn. (21). Thus the Nash equilibrium satisfies the optimality conditions Eqn. (21) to (23) when $|N| > 2$, and solves the optimization problem Eqn. (18)(19). Finally, this optimization problem has a unique optimum solution, hence the Nash equilibrium exists and it is unique. The theorem holds. ■

When tenants reach the unique Nsh equilibrium, no tenants can improve their utility by changing their own bids. If there are n tenants participating in the market and each one with $q_i^* > 0$, we can know that $n \geq 3$ according to Lemma 1. Since $F_i(q_i)$ is convex, there exists at least one point (tenant k, $k \in N$) that makes $F_k'(q_k^*)$ equal to zero. That is

$$(1 + \frac{q_k}{E - 2q_k})C_k'(q_k^*) - p^* = 0, \tag{32}$$

$$p^* = (1 + \frac{q_k}{E - 2q_k})C_k'(q_k^*). \tag{33}$$

We use $G_k(q_k)$ to denote the right part of above equation:

$$G_k(q_k) = (1 + \frac{q_k}{E - 2q_k})C_k'(q_k^*) > 0. \tag{34}$$

The first derivative of $G_k(q_k)$ is

$$G_k'(q_k) = \frac{E}{(E - 2q_k)^2}C_k'(q_k) + (1 + \frac{q_k}{E - 2q_k})C_k''(q_k) > 0. \tag{35}$$

Thus $G_k(q_k)$ increases as to q_k. When $|N| \geq 3$, we have

$$G_k(\frac{E}{N}) \leq G_k(\frac{E}{3}) \tag{36}$$

Because q_j^* is less or equal than $\frac{E}{N}$, we have

$$p^* \leq G(\frac{E}{N}) \leq G(\frac{E}{3}) = C_k'(\frac{E}{3}) \tag{37}$$

From the above inequality, we can see that when $|N|$ is greater than 3, there is at least one Nash equilibrium, and p^* is no greater than $C_k'(\frac{E}{3})$.

B. The Stackelberg Equilibrium

This subsection analyzes the two-level game between the operator and tenants. When the tenants submit their a profile of bids (**b**), the operator tries to decide the best strategy by maximizing the utility function Eqn. (6). We assume that

$$K(E) = a \times \log(E + 1) \tag{38}$$

Here a is an adjustment coefficient. E does not involve the renewable generation and we will discuss it later. When $E = 0$, $K(E)$ is 0 and as E increases $K(E)$ increases. More commonly as $K(E)$ increases faster, $K(E)$ increases more and more slowly. The cost of the operator is the reward offered to the tenants. This cost $C_o(\cdot)$ is

$$C_o(E) = E \times p = E \times \frac{E}{B}. \tag{39}$$

Thus, the operator's utility is

$$R(E, \mathbf{b}) = a\log(E + 1) - \frac{E^2}{B}. \tag{40}$$

Based on Eqn. (40), we have the following definition.

Definition 1 (Best Response for the Operator): The best response of the operator is the best strategy for the operator with the tenants' bids **b**.

$$E(\mathbf{b}) := \arg\max R(\mathbf{b}). \tag{41}$$

Definition 2 (Stackelberg Equilibrium): A stackelberg equilibrium is a strategy E^* for the operator and a profile of strategies $\mathbf{b}^* = (b^*)_{\forall i}$ for tenants,

$$E^* = E(\mathbf{b}^*) \quad and \tag{42}$$

$$b_i^* = 0 \quad or \quad \frac{B_{-i}^* - b_i^*}{B_{-i}^* + b_i^*} - \frac{B_{-i}^*}{E}C_i'(\frac{b_i^* \times E}{B_{-i}^* + b_i^*}) = 0, \forall i. \tag{43}$$

The first and second derivatives of Eqn. (40) are

$$R'(E, \mathbf{b}) = \frac{a}{E + 1} - \frac{2E}{B},$$
$$R''(E, \mathbf{b}) = -\frac{a}{(E + 1)^2} - \frac{2}{B}. \tag{44}$$

We can see that the second derivative of Eqn. (40) is less than zero, and $R(E)$ is a concave function. Hence, $R(E)$ is maximized when $R'(E) = 0$. We thus have

$$\frac{a}{E^* + 1} - \frac{2E^*}{B} = 0, \; E^*(\mathbf{b}) = \sqrt{\frac{a \cdot B}{2} + \frac{1}{4}} - \frac{1}{2} \tag{45}$$

After tenants submit their bids, the operator can calculate the best amount of energy reduction by Eqn. (45) to maximize the utility. This best decision is dependent on tenant's bids and the financial compensation from the electric power company. Since the noncooperative game among tenants has a unique Nash Equilibrium, E^* is also unique according to Eqn. (45). We have the following major theorem.

Theorem 4: The two-level market (or the Stackelberg game) between the operator and tenants has a unique Stackelberg equilibrium $(E^*, \mathbf{b}^*)$.

Stackelberg game brings a win-win situation for colocation. Both operator and tenants receive financial benefits in the game. It also proves that operator successfully split the incentive of demand response. In Stackelberg equilibrium, both operator and tenants can not move forward to maximize their benefits. So a reward price provided by operator is accepted by tenants and a profile of bids offered by tenants is received by operator. Then both sides calculate their financial benefits according to the reward price and bids.

C. Iterative biding

The above subsections present the Stackelberg equilibrium for on the two-level market. This subsection provides the iterative process between the operator and tenants and shows how to converge to the equilibrium. The iteration is as follows.

(1) **Initialization**: the operator first communicates the message of energy reduction and the initial value $p(0)$. This value can be estimated from the historical data.

(2) $\mathbf{k^{th}}$**–iteration**: (i) Based on receiving the reward price $p(k)$ from the operator, each tenant i will decide their bids by

$$b_i(k) = [\frac{F_i^{'-1}(p(k))}{p(k)}]^+. \tag{46}$$

Here Eqn.(46) is based on the inverse function of Eqn.(2) and Eqn.(33). Then the tenants submit these bids to the colocation operator. (ii) After collecting all tenants' bids, the operator conducts updates accordingly. There are two steps. First, the operator calculates the best total amount of energy reduction by.

$$E(k) = \sqrt{\frac{a \cdot B_i(k)}{2} + \frac{1}{4}} - \frac{1}{2},\ B(k) = \sum_{i=1}^{N} b_i(k). \quad (47)$$

The second step is to decide the next round reward price based on the total amount of energy reduction.

$$p(k+1) = [p(k) - r(B(k) \cdot p(k) - E(k))]^{+}. \quad (48)$$

The scalar r is called the modulus of $p(\cdot)$, which is a constant in $[0, 1)$. In each iteration, both the operator and the tenants play their best response to the other's best strategies. One key question here is whether the iteration bidding will converge to a Stackelberg equilibrium. In the rest of this section, we prove that it surely converges to the equilibrium. To prove this is equivalent to show that the above iteration process is a contraction mapping [42]. The proof is as follows.

We rewrite the above iterative process as

$$p(k+1) = T(p(k)),\ k \in [0, N]. \quad (49)$$

where T is a mapping from a subset $Z(p(k) \in Z)$ of R into itself. Here, we assume that x is equal to $p(k)$ and y is equal to $p(k+n), (n \in N)$, and y is equal to $x + t, (t \in R)$. Thus,

$$T(x) = x - r(B \cdot x - E). \quad (50)$$

Because $E = \sqrt{\frac{a \cdot B}{2} + \frac{1}{4}} - \frac{1}{2}$, we have

$$T(x) = x - r[B \cdot x - (\sqrt{\frac{a \cdot B}{2} + \frac{1}{4}} - \frac{1}{2})]. \quad (51)$$

We assume that when $p(k+n) = y$, $\sum_{i=1}^{N} b_i^{'}(k+n) = B^{'}(k+n)$ and the energy reduction is $E^{'}$. $B^{'}(k+n)$ is written $B^{'}$.

$$\begin{aligned} T(y) &= T(x+t) = x + t - r[B^{'} \cdot (x+t) - E^{'}] \\ &= x + t - r[B^{'} \cdot (x+t) - \sqrt{\frac{a \cdot B'}{2} + \frac{1}{4}} + \frac{1}{2}]. \end{aligned} \quad (52)$$

Next we calculate the difference between $T(y)$ and $T(x)$ by comparing Eqn. (51) and Eqn. (52).

$$\begin{aligned} & T(y) - T(x) \\ &= t + r(B \cdot x - \sqrt{\frac{a \cdot B'}{2} + \frac{1}{4}} + \frac{1}{2} \\ &- B^{'} \cdot x - B^{'} \cdot t + \sqrt{\frac{a \cdot B'}{2} + \frac{1}{4}} - \frac{1}{2}) \end{aligned} \quad (53)$$

If Eqn. (49) is a contraction mapping, it must have have the following property:

$$\|T(y) - T(x)\| \le \alpha \|y - x\| = \alpha \|t\|, \quad (54)$$

where $\|\cdot\|$ is a norm and α is a constant in $[0, 1)$. We analyze the connection between the reward price (p) and the bid sum (B). When $|N| > 2$, the Nash equilibrium of the game among tenants solves the problem Eqn. (18). Eqn. (18) is a convex optimization problem and has a unique solution. Thus,

$$F_i^{'}(q_i) = (1 + \frac{q_i}{E - 2q_i})C_i^{'}(q_i) - p. \quad (55)$$

We assume that $F_i^{'}(q_i)$ is equal to M. If M is equal to zero, the game reaches the Nash equilibrium. Furthermore, we can obtain the unique solution when M is equal to zero. But before reaching the equilibrium, M is a non-zero real value. By manipulating Eqn.(55), we can have

$$(1 + \frac{q_i}{E - 2q_i})C_i^{'}(q_i) - p = M,\ q_i = \frac{E \cdot (p + M - C_i^{'}(q_i))}{2(M + p) - C_i^{'}(q_i)} \quad (56)$$

Based on Eqn. (2) and Eqn. (56), we have

$$b_i = \frac{E \cdot (p + M - C_i^{'}(q_i))}{[2(M + p) - C_i^{'}(q_i)] \cdot p}. \quad (57)$$

Now we can see that b_i deceases as to p, and B deceases as to p. We return to prove the contraction mapping of the iteration process. Based on Eqn. (53) and Eqn. (54), we have

$$\begin{aligned} \|T(y) - T(x)\| &= \|t + r(B \cdot x - \sqrt{\frac{a \cdot B}{2} + \frac{1}{4}} + \frac{1}{2} \\ &- B^{'} \cdot x - B^{'} \cdot t + \sqrt{\frac{a \cdot B'}{2} + \frac{1}{4}} - \frac{1}{2})\| \end{aligned} \quad (58)$$

Because Eqn.(54) is greater than zero, we have

$$C_i^{'}(q_i) - M < p < 2C_i^{'}(q_i) - M. \quad (59)$$

Because $y > x$, $B^{'}$ is greater than B. We thus have

$$\begin{aligned} & \|T(y) - T(x)\| \\ & \le \|t + r(B \cdot x - \sqrt{\frac{a \cdot B}{2} + \frac{1}{4}} \\ & - B^{'} \cdot x - B^{'} \cdot t + \sqrt{\frac{a \cdot B}{2} + \frac{1}{4}})\| \\ & = \|t + r(B \cdot x - B^{'} \cdot x - B^{'} \cdot t)\|. \end{aligned} \quad (60)$$

When r is small enough, $B \cdot x - B^{'} \cdot x - B^{'} < 0$. We have

$$\|T(y) - T(x)\| \le \|t\|.$$

Hence, there must be an α that meets the following inequality.

$$\|T(y) - T(x)\| \le \alpha \|t\| \quad \alpha \subseteq [0, 1]. \quad (61)$$

And this proves that Eqn.(49) is a contraction mapping.

IV. Co-Located Renewable Energy

A colocation is usually co-located with renewable such as wind or solar power given the need of sustainable development and the lower cost of renewable [14]. This power generation provides further flexibility to optimize a colocation's power supply and thus maximize its economic performance. Meanwhile, the uncertainty of renewable causes difficulty in determining the optimal strategy for a colocation operator. In this section, we address this difficulty and study how the renewable generation impacts the operator's strategy in the Stackelberg game.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSUSC.2019.2904867, IEEE Transactions on Sustainable Computing

A. Model of Renewable Energy

At the beginning of a demand response period, the colocation operator predicts the amount of generated renewable (or renewable output) for the period, which is then used to make the best decision on E. We use a prediction error ε to capture the renewable's uncertainty, that is, $\omega = (1+\varepsilon)\widetilde{\omega}$, where ω is the actual renewable output and $\widetilde{\omega}$ is the predicted output. According to the standard assumptions in statistics, we use unbiased prediction, that is, the mean $\mathbb{E}[\varepsilon] = 0$ and $\mathbb{V}[\varepsilon] = \sigma^2$ [43] [44]. In addition, we assume the maximum amount of renewable generation is M. Then we have $\omega \leq M$ and $\widetilde{\omega} \leq M$.

B. Utility of Operator with Renewable Energy

In this part, we count in renewable energy to the utility function of operator side in Data center. The utility of operator is decided by two factors. One is the financial compensation from the electric power company due to responding to the demand response signal ($K(\cdot)$). The other one is the total cost for paying tenants who participate in the energy saving ($C_o(\cdot)$). If the operator decides not to turn on renewable facility, he/she only needs to set $\omega = 0$. Otherwise, $\omega \leq M$ (maximum amount of renewable generation). The operator's utility function with the actual renewable generation is

$$R(E^*, \omega) = K(E^* + \omega) - C_o(E^*), \tag{62}$$

The operation cost of renewable is negligible compared to the price of electricity from power grid [23]. Based on previous analysis (III section B), we can get the following results

$$E^* = \sqrt{a * B/2 + \frac{1}{4}} - \frac{1}{2} - [\omega]_0^M \tag{63}$$

$$\widetilde{E^*} = \sqrt{a * B/2 + \frac{1}{4}} - \frac{1}{2} - [\widetilde{\omega}]_0^M \tag{64}$$

Here E^* is the best solution with respect to actual renewable energy output and $\widetilde{E^*}$ is the optimal solution with the predicted renewable energy ($\omega = (1+\varepsilon)\widetilde{\omega}$). We only use $\widetilde{\omega}$ to estimate $\widetilde{E^*}$. When the operator counts the actual utility which he/she gains, the actual renewable generation ω should be used.

C. Performance Analysis

The performance of (Eqn. (63) and Eqn. (64)) depends on the accuracy of the predicted renewable power output. Because if ω is equal to $\widetilde{\omega}$, the best solution is the same as the actual results. This means our algorithm is efficient and effective. In this subsection, we adopt competitive analysis to evaluate and analyze our algorithm.

Competitive analysis is widely used to analyze online algorithms, in which the performance of an online algorithm is compared to the performance of an optimal offline algorithm that can view future information (here is the actual renewable output) in advance. An algorithm is competitive if its competitive ratio, i.e., the ratio between its performance and the offline algorithm's performance, is bounded [29] [45]. Our approach uses the on-line predicted renewable output to determine the operator's strategy. There is a need of understanding the economic performance gap between the strategy based on the predicted renewable and the actual optimal strategy. Using competitive analysis right fills the need and will answer questions such as how much is the gap and whether the gap is bounded. For example, if the gap or the competitive ratio is bounded by a linear relation with the prediction error, our approach will derive rather good performance. If bounded by an exponential relation, a small prediction error can cause large performance degradation.

The competitive ratio η is defined as the ratio of expectations of utilities calculated with actual renewable generation and with predicted renewable generation:

$$\eta = \frac{\mathbb{E}[R(E^*, \omega)]}{\mathbb{E}[R(\widetilde{E^*}, \omega)]} \tag{65}$$

Based on the assumption of renewable generation, we have a lemma here.

Lemma 5: If the prediction error of renewable generation is ε, and the expectation and variance are $\mathbb{E}[\varepsilon] = 0$ and $\mathbb{V}[\varepsilon] = \sigma^2$ respectively, we have $\mathbb{E}[\varepsilon^+] \leq \frac{\sigma}{2}$, $\mathbb{E}[\varepsilon^-] \geq -\frac{\sigma}{2}$, where $\varepsilon^+ = max(0, \varepsilon)$ and $\varepsilon^- = min(0, \varepsilon)$.

Lemma 5 gives the critical value of $\mathbb{E}[\varepsilon^+]$ and $\mathbb{E}[\varepsilon^-]$, and this help us analyse the boundary of competitive ratio. We do not repeat the proof of Lemma 5 in this paper which can be found in [43]. We have another Lemma as following

Lemma 6: For E^* given by Eqn.(63) and $\widetilde{E^*}$ given by Eqn.(64), we have $E^* = \widetilde{E^*} + \varphi$, and $\varphi = \varphi^+ + \varphi^-$, where $\varphi^+ = [-\widetilde{\omega} * \varepsilon^-]_0^{-\frac{M*\varepsilon^-}{1+\varepsilon^-}}$ and $\varphi^- = [-\widetilde{\omega} * \varepsilon^+]^0_{-\frac{M*\varepsilon^+}{1+\varepsilon^+}}$.

Proof: Based on $\omega = (1+\varepsilon)\widetilde{\omega}$ and Eqn.(63), we have the following equation

$$E^* = \sqrt{a * B/2 + \frac{1}{4}} - \frac{1}{2} - [(1+\varepsilon)\widetilde{\omega}]_0^M \tag{66}$$

We assume that $E^* = \widetilde{E^*} + \varphi$ and $\varphi = \varphi^+ + \varphi^-$, and we have

$$E^* = \sqrt{a * B/2 + \frac{1}{4}} - \frac{1}{2} - [\widetilde{\omega}]_0^M + \varphi \tag{67}$$

From Eqn.(66) and Eqn.(67), we have

$$\begin{aligned} [(1+\varepsilon)\widetilde{\omega}]_0^M &= [\widetilde{\omega}]_0^M - \varphi \\ \varphi &= [\widetilde{\omega}]_0^M - [(1+\varepsilon)\widetilde{\omega}]_0^M \end{aligned} \tag{68}$$

Next we analysis Eqn.(68) with ε^+ and ε^- separately. First, if $\varepsilon = \varepsilon^+$, we have

$$\begin{aligned} \varphi &= \widetilde{\omega} - (1+\varepsilon^+)\widetilde{\omega} \\ \varphi &= -\varepsilon^+ * \widetilde{\omega} \end{aligned} \tag{69}$$

Because $\omega \leq M$ and $\widetilde{\omega} \leq M$, we can get that

$$\widetilde{\omega} \leq \frac{M}{1+\varepsilon^+} \tag{70}$$

From Eqn.(69) and Eqn.(70), we can get

$$\varphi^- = [-\varepsilon^+ * \widetilde{\omega}]^0_{-\frac{M*\varepsilon^+}{1+\varepsilon^+}} \tag{71}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSUSC.2019.2904867, IEEE Transactions on Sustainable Computing

Second, if $\varepsilon = \varepsilon^-$, we have

$$\varphi = \widetilde{\omega} - (1+\varepsilon^-)\widetilde{\omega} \tag{72}$$
$$\varphi = -\varepsilon^- * \widetilde{\omega} \tag{73}$$

Also since $0 \le \omega \le M$ and $0 \le \widetilde{\omega} \le M$, we have

$$\widetilde{\omega} \le \frac{M}{1+\varepsilon^-} \tag{74}$$

From Eqn.(73) and Eqn.(74),we have the following

$$\varphi^+ = [-\varepsilon^- * \widetilde{\omega}]_0^{-\frac{M*\varepsilon^-}{1+\varepsilon^-}} \tag{75}$$

We now finish proof and Lemma 6 holds. ■

Based on Lemma 5 and Lemma 6, we introduce our main theorem of this section. This theorem bounds the competitive ratio. It also proves that if the prediction of renewable generation is accurate, the operator's utility is close to the optimal.

Theorem 7: If the variance of the prediction error of the renewable generation is bounded by σ^2, the utility given by Eqn.(62) has a competitive ratio(denoted by η), and we also have that $\eta = \frac{\mathbb{E}[R(E^*,\omega)]}{\mathbb{E}[R(\widetilde{E^*},\omega)]} \le 1 + \chi * \sigma$, where

$$\chi = \frac{\Big(K'(\widetilde{E^*}+\omega) - C_o'(\widetilde{E^*})\Big)\widetilde{\omega}}{2\mathbb{E}[K(\widetilde{E^*}+\omega) - C_o(\widetilde{E^*})]}. \tag{76}$$

Proof: Based on the definition of competitive ratio in Eqn.(65), the following deduction try to find the bound of the ratio.

$$\begin{aligned}
\eta &= \frac{\mathbb{E}[K(E^*+\omega) - C_o(E^*)]}{\mathbb{E}[K(\widetilde{E^*}+\omega) - C_o(\widetilde{E^*})]} \\
&= \frac{\mathbb{E}[K(\widetilde{E^*}+\varphi+\omega) - C_o(\widetilde{E^*}+\varphi)]}{\mathbb{E}[K(\widetilde{E^*}+\omega) - C_o(\widetilde{E^*})]} \\
&= \frac{\mathbb{E}[K(\widetilde{E^*}+\omega) - K(\widetilde{E^*}+\omega) + K(\widetilde{E^*}+\varphi+\omega)]}{\mathbb{E}[K(\widetilde{E^*}+\omega) - C_o(\widetilde{E^*})]} \\
&+ \frac{\mathbb{E}[C_o(\widetilde{E^*}) - C_o(\widetilde{E^*}) - C_o(\widetilde{E^*}+\varphi)]}{\mathbb{E}[K(\widetilde{E^*}+\omega) - C_o(\widetilde{E^*})]} \\
&= 1 + \frac{\mathbb{E}[K(\widetilde{E^*}+\varphi+\omega) - K(\widetilde{E^*}+\omega)]}{\mathbb{E}[K(\widetilde{E^*}+\omega) - C_o(\widetilde{E^*})]} \\
&- \frac{\mathbb{E}[C_o(\widetilde{E^*}+\varphi) - C_o(\widetilde{E^*})]}{\mathbb{E}[K(\widetilde{E^*}+\omega) - C_o(\widetilde{E^*})]}
\end{aligned} \tag{77}$$

Because $K(\cdot)$ is concave and $C_o(\cdot)$ is strictly convex, we introduce the properties of concave and convex function to boundary the competitive ratio. Then we have the following inequalities

$$\begin{aligned}
\eta &\le 1 + \frac{\mathbb{E}[K'(\widetilde{E^*}+\omega) * \varphi - C_o'(\widetilde{E^*}) * \varphi]}{\mathbb{E}[K(\widetilde{E^*}+\omega) - C_o(\widetilde{E^*})]} \\
&= 1 + \frac{\Big(K'(\widetilde{E^*}+\omega) - C_o'(\widetilde{E^*})\Big) * \mathbb{E}[\varphi]}{\mathbb{E}[K(\widetilde{E^*}+\omega) - C_o(\widetilde{E^*})]} \\
&\le 1 + \frac{\Big(K'(\widetilde{E^*}+\omega) - C_o'(\widetilde{E^*})\Big) * \mathbb{E}[\varphi^+]}{\mathbb{E}[K(\widetilde{E^*}+\omega) - C_o(\widetilde{E^*})]} \\
&= 1 + \frac{\Big(K'(\widetilde{E^*}+\omega) - C_o'(\widetilde{E^*})\Big) * \mathbb{E}[-\varepsilon^- * \widetilde{\omega}]}{\mathbb{E}[K(\widetilde{E^*}+\omega) - C_o(\widetilde{E^*})]} \\
&\le 1 + \frac{\Big(K'(\widetilde{E^*}+\omega) - C_o'(\widetilde{E^*})\Big)\widetilde{\omega}}{2\mathbb{E}[K(\widetilde{E^*}+\omega) - C_o(\widetilde{E^*})]}\sigma
\end{aligned} \tag{78}$$

Thus we have the following

$$\eta \le 1 + \chi * \sigma, \tag{79}$$

and Theorem 7 holds. ■

There are two important points from the above analysis. First, since χ in Eqn. (76) is constant, we can see that by Theorem 7, the competitive ratio has a linear relation with the standard deviation (σ) of the renewable prediction error. This result can be interpreted as that the economic performance of the colocation operator will be rather good if the renewable prediction is enough accurate. The competitive ratio decreases to 1, when the prediction error decreases to 0. (More description about this effect can be found in the Evaluation section.) Second, it is worth noting that the above analysis does not rely on any assumptions on the distribution of the prediction error other than zero mean and the bounded variance. This confirms that our analysis approach is more generalized compared to many existing analysis approaches.

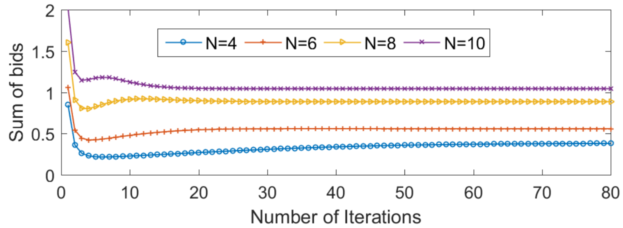
V. Evaluation

Till now we have proposed a two-level market mechanism for colocations, provided theoretical a Nash equilibrium and a Stackelberg equilibrium for the mechanism, and further analyzed the economic performance of colocation demand response with co-located renewable. In this section, we highlight the benefits of the market design using extensive simulations.

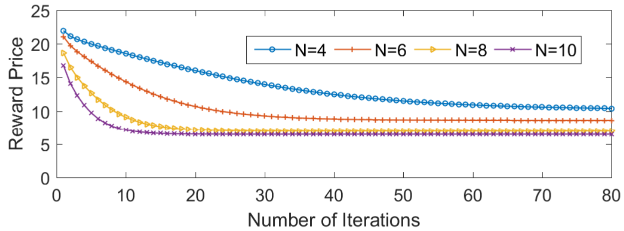
A. Simulation Setup

We consider a colocation with N tenants. The cost function of tenant i is set as $C_i(q_i)$= $\alpha_i * q_i + \beta_i * q_i^2$, where α_i is randomly generated in the range of [1, 6] and β_i is randomly drawn from [1, 5]. Note that the evaluation does not confine to a specific form of function, and other forms that satisfy the properties in Section II are also feasible. The compensation price (a in Eqn. (38)) is offered by the electric power company. The parameter N and a vary to allow sensitivity analysis in this evaluation. The step size r in the iterative biding (Section III-C) is set to 0.2. The total amount of energy reduction E is initialized as 50 when starting biding iterations.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSUSC.2019.2904867, IEEE Transactions on Sustainable Computing

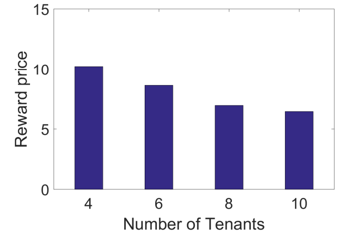


(a) Sum of all tenant's bids when the number of tenants (N) varies

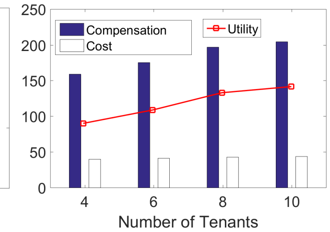


(b) Reward price when the number of tenants varies (N)

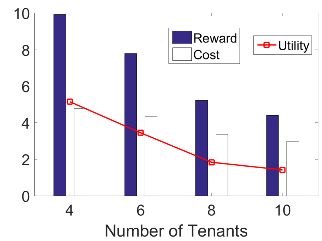
Fig. 2. Convergence of the iterative biding by showing tenant's sum bids and reward price.



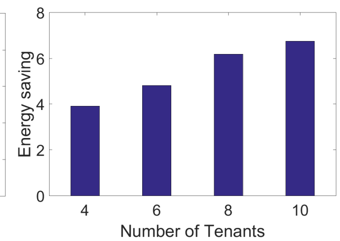
(a) Reward price



(b) The operator's compensation, cost, and utility



(c) Tenant 1's reward, cost, and utility



(d) Energy reduction

Fig. 3. Sensitivity analysis for tenants and the operator when the compensation price (a in Eqn. (38)) equals 100 and the number of tenants (N) varies.

The actual renewable generation ω and the prediction error (i.e., the standard deviation σ) also vary in the simulation.

B. Simulation Results

In this section, we present simulation results about mechanism convergence, sensitivity analysis for the market, and co-located renewable.

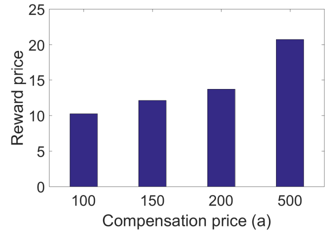
1) Results on the Mechanism Convergence: Fig. 2 demonstrates the convergence of the iterative biding. First, the convergence speed is rather fast, that is, values of parameters almost stay unchanged after tens of iterations. For example, for all four curves (with four different numbers of tenants) as shown in Fig. 2(b), the reward price nearly does not change after 70 iterations. A similar observation based on the sum bid of tenants can be also made from Fig. 2(a). Second, the iteration process converges faster when the number of tenants (N) is larger. For example, as shown in Fig. 2(b), the curve of $N = 10$ converges faster than other three curves of $N = 4, 6, 8$. The reason is as follows. When the number of tenants is larger, their sum bid also becomes more, as shown in Fig. 2(a). A larger sum bid, i.e., a larger B in Eqn. (48), the step between two consecutive iterations becomes larger, and thus the convergence speed grows.

2) Impact of the Number of Tenants on the Market: Fig. 3 shows how the market varies with number of tenants (N). From Fig. 3(a), we can see that the reward price decreases as the number of tenants (N) increases. For example, the reward price when $N = 4$ is the largest among all the four settings. The reason here is two-fold. When the number of tenants is smaller, the operator needs each individual tenant to contribute more energy saving. Thus, the operator has to increases the reward price to motivate tenants to do so. Further, this result infers that it is not beneficial for a single tenant to fake and divide himself/herself as multiple smaller tenants.

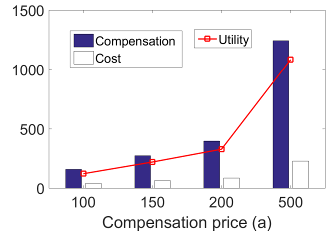
Fig. 3(b) illustrates how the operator's situation changes with the number of tenants. The blue bars represent the compensation received by the operator from the electric power company, and the white bars represent the cost of the operator, i.e., the total reward offered to tenants. The red curve is the operator's utility under four different settings. First, the compensation increases as the number of tenants grows, so does the utility. The reason is as follows. More tenants participating into the market results in more energy saving, demonstrated by Fig. 3(d). With more energy reduction, the operator can receive more compensation from the electric power company. Further, since the operator's cost, i.e., the total reward offered to tenants, increases slightly, as shown by white bars in Fig. 3(b). The operator's utility also becomes larger with an increased number of tenants. Furthermore, the operator's utility is always positive, and thus the operator is motivated and beneficial to participate into the market by offering reward to tenants.

Fig. 3(c) shows how a tenant's (here uses Tenant 1 as an example) situation changes with the number of tenants. The blue bars represent the reward received by Tenant 1, the white bars represent Tenant 1's cost caused by energy reduction, and the red curve is the tenant's utility. We can see that the tenant's reward and cost both decrease as the number of tenants increases. The major reason for this observation is that tenants participating into the market compete for the reward. Because of the competition, the reward price declines (as shown in From Fig. 3(a)) and energy reduction of individual tenant also goes down. Hence, each tenant's reward and cost decreases accordingly. Further, by the red curve, the tenant's utility declines as the number of tenants increases. An insight here is that the number of tenants willing to participate will saturate after the total number of tenants reaches a certain

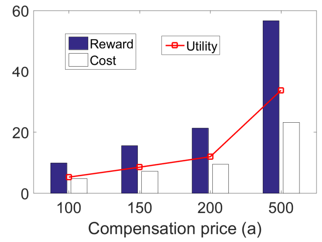
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSUSC.2019.2904867, IEEE Transactions on Sustainable Computing



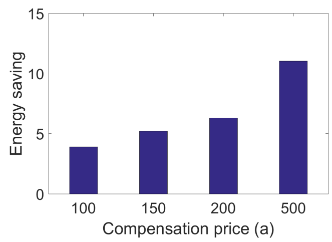
(a) Reward price



(b) The operator's compensation, cost, and utility



(c) Tenant 1's reward, cost, and utility



(d) Energy reduction

Fig. 4. Sensitivity analysis for tenants and the operator the number of tenants (N) equals 4 and the compensation price (a in Eqn. (38)) varies.

threshold (which is a very large number). At that time, it is not beneficial for new tenants to participate into the market anymore. By contrast, before that time, tenants always have positive utility, which motivates them to participate into the market. Furthermore, seeing both Fig. 3(b) and Fig. 3(c), we can conclude a win-win situation, i.e., with positive utility for both the operator and tenants.

Fig. 3(d) shows that the amount of energy reduction increases as the number of tenants grows. More tenants participating into the market results in higher competition on the reward among them. From Fig. 3(b), the total reward stays nearly unchanged with the number of tenants. So after splitting the total reward among more tenants, the amount of each tenant's energy reduction goes down. In spite of this, multiplied by a larger number of tenants, the total energy reduction still becomes more. The observation infers that more tenants participating into the market results in more benefits for the operator. This in turn follows the result in Fig. 3(b).

3) Impact of the Compensation on the Market: Fig. 4 depicts how the market varies with the compensation price (a in Eqn. (38)), which is offered by the electric power company to the operator. From Fig. 4(a), we can see that the reward price grows as the compensation price increases. A higher compensation price tends to motivate the operator to save more energy (and thus more compensation). To achieve this, the operator in turn needs to motivate tenants to reduce more energy, and thus increases the reward price. By Fig. 4(d), we can see that the higher reward price (caused by the increased compensation price) indeed leads to more energy reduction from tenants.

Fig. 4(b) illustrates how the operator's situation changes with the compensation price. The blue bars represent the compensation received by the operator, and the white bars

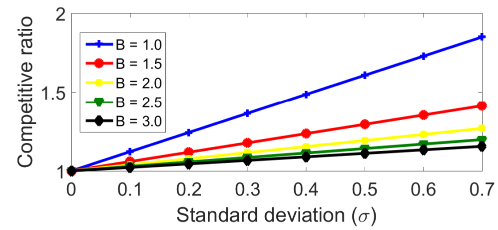


Fig. 5. The competitive ratio varies with the standard deviation and the sum bid B of all tenants at the equilibrium.

represent the cost of the operator. First, the compensation increases as its price increases, so does the operator's cost. As mentioned above, a higher compensation price results in more energy reduction, as shown in Fig. 4(d). Increased compensation price together with increased energy reduction causes the compensation increases very fast. Second, the red curve shows that the operator's utility also increases as the compensation price becomes higher. The reason is that the compensation increases much more than the operator' cost. Thus their difference, i.e., the utility, increases accordingly, instead of decreasing.

Fig. 4(c) shows how a tenant's (here uses Tenant 1 as an example) situation changes with the compensation price. The blue bars represent the reward received by the tenant, the white bars are Tenant 1's cost, and the red curve is the tenant's utility. Similarly, the tenant's reward, cost, and utility all go up as the compensation price becomes higher. This is also because of increased reward price and more energy reduction as mentioned above. By all the four figures, we can see that a higher compensation price more benefits not only the operator but also tenants participating the market.

4) Results about the Co-located Renewable: Fig. 5 demonstrates how the competitive ratio varies with standard deviation of renewable prediction error (σ) and the sum bid of tenants (B). First, we can see that when the standard deviation increases, the competitive ratio increases, i.e., the algorithm's performance declines. This observation well follows the intuition. Second, the competitive ratio linearly varies with the standard deviation. These two observations also validate the correctness of Theorem 7. Further, the proposed approach has rather good performance when the prediction is accurate enough. For example, when the standard deviation is less than 10%, the competitive ratio is less than 1.1. It is less than 1.2 when the standard deviation is less than 20%. Thus, the algorithm is well deployable in practice when the prediction error is acceptable. Third, as the sum bid (B) at the equilibrium increases, the competitive ratio decreases. For example, when $\sigma = 0.3$, the competitive ratio is about 1.35 as $B = 1$ and it is about 1.2 as $B = 1.5$. Fourth, the increasing rate of the competitive ratio decreases as the sum bid (B) increases. For example, the increasing rate (i.e., the slope) of $B = 1$ is larger than that of others such as $B_i = 2$. The reason for these two observations is that when more tenants join in the game or the demand response, the bids will increase. The operator thus

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSUSC.2019.2904867, IEEE Transactions on Sustainable Computing

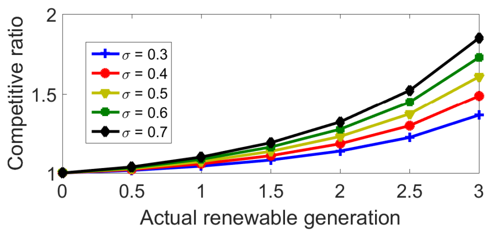


Fig. 6. The competitive ratio varies with the actual renewable generation.

receives more energy saving from tenants, and the impact of renewable for the compensation becomes relatively less. That is, more tenants offering more energy saving will offset the uncertainty of renewable. These observations partly confirm that demand response is key to accommodate renewable in our power systems.

Fig. 6 shows the competitive ratio varies with actual renewable generation (ω). From the figure, we can see that the competitive ratio increases as actual renewable output increases. This observation indicates that it does not encourage the operator to increase the renewable capacity aggressively but wisely. For example, the competitive ratio is less than 1.1 when actual renewable is less than 1, while it becomes over 1.8 when actual renewable is 3. Thus, we can conclude that it is better for the colocation operator to incorporate hybrid energy sources, for example, one part uses co-located renewable for a cost-saving purpose and the other part uses reliable sources such as the power grid to guarantee availability of the service. This insight also follows the key idea of GreenPlanning [23], that is, incorporating multiple energy sources to strike a balance between different goals. Further, this figure also shows that the competitive ratio increases as the renewable prediction error of renewable increases, which is also observed by Fig. 5. For example, the curve of $\sigma = 0.7$ is above all other curves of smaller σ.

VI. Conclusions

Demand response is a crucial functionality of the power grid to realize the key vision of demand-follow-supply. This paper is focused on a notably promising demand response resource, colocation data centers, and specially studies mechanism design for colocation demand response in presence of co-located renewable. We propose a hierarchical demand response scheme that jointly addresses the split incentive between a colocation operator and its tenants, and the uncertainty of co-located renewable. The scheme is based on a two-level market mechanism, by which tenants receive rewards from the operator given their power reduction and the operator obtains financial compensation from the electric power company. The mechanism provably converges to a unique equilibrium solution, where neither the operator or tenants can improve their individual utility by changing their own strategies. Further, we present a stochastic optimization based algorithm to optimize economic performance for the colocation operator at the equilibrium. Performing competitive analysis demonstrates a robust guarantee of the operators's financial gain in terms of the renewable prediction error. Finally, we evaluate the designed scheme using extensive simulations. Results illustrate the win-win situation between a colocation operator and its tenants, and validate the linear relation between the operator's economic performance and the renewable prediction error. As to future work, we plan to study the truthfulness of the operator and tenants, e.g., how to ensure tenants provide a reasonable cost function for their power reduction, as well as extend the current market mechanism to a setting of geographically distributed data centers, which may have different demand response periods and renewable generation.

Acknowledgment

The authors would like to...

References

[1] U.S. Department of Energy, "The smart grid: An introduction," http://www.energy.gov/, [Online accessed March, 2015].

[2] NRDC, "Scaling up energy efficiency across the data center industry: Evaluating key drivers and barriers," 2014, issue Paper.

[3] M. A. Islam, H. Mahmud, S. Ren, and X. Wang, "Paying to save: Reducing cost of colocation data center via rewards," in *IEEE International Symposium on High Performance Computer Architecture (HPCA)*, 2015.

[4] A. Greenberg, J. Hamilton, D. A. Maltz, and P. Patel, "The cost of a cloud: research problems in data center networks," *ACM SIGCOMM computer communication review*.

[5] S. Ren and M. A. Islam, "Colocation demand response: Why do i turn off my servers?" in *International Conference on Autonomic Computing (ICAC)*. Usenix, 2014.

[6] X. Liu and F. Kong, "Datacenter power management in smart grids," *Foundations and Trends® in Electronic Design Automation*, 2015.

[7] A. Wierman, Z. Liu, I. Liu, and H. Mohsenian-Rad, "Opportunities and challenges for data center demand response," in *International Green Computing Conference*. IEEE, 2014.

[8] N. H. Tran, D. H. Tran, S. Ren, Z. Han, E.-N. Huh, and C. S. Hong, "How geo-distributed data centers do demand response: A game-theoretic approach," *IEEE Transactions on Smart Grid*, 2016.

[9] G. Ghatikar, "Demand response opportunities and enabling technologies for data centers: Findings from field studies," 2014.

[10] A. Rahman, X. Liu, and F. Kong, "A survey on geographic load balancing based data center power management in the smart grid environment," *IEEE Communications Surveys & Tutorials*, 2014.

[11] F. Kong, X. Lu, M. Xia, X. Liu, and H. Guan, "Distributed optimal datacenter bandwidth allocation for dynamic adaptive video streaming," in *Proceedings of the 23rd ACM international conference on Multimedia*. ACM, 2015.

[12] N. H. Tran, C. Pham, M. N. Nguyen, S. Ren, and C. S. Hong, "Incentivizing energy reduction for emergency demand response in multi-tenant mixed-use buildings," *IEEE Transactions on Smart Grid*, 2018.

[13] F. Kong and X. Liu, "A survey on green-energy-aware power management for datacenters," *ACM Computing Surveys (CSUR)*, 2015.

[14] Equinix, "Green data centers," http://www.equinix.com/.

[15] Steve Hassell, "Will the data center of the future be covered in solar panels powering 52 kw racks?" http://www.datacenterjournal.com/data-center-future-covered-solar-panels-powering-52-kw-racks/.

[16] F. Kong, C. Dong, X. Liu, and H. Zeng, "Blowing hard is not all we want: Quantity vs quality of wind power in the smart grid," in *INFOCOM*. IEEE, 2014.

[17] ——, "Quantity versus quality: Optimal harvesting wind power for the smart grid," *Proceedings of the IEEE*, 2014.

[18] S. Bahrami, M. H. Amini, M. Shafie-khah, and J. P. Catalao, "A decentralized renewable generation management and demand response in power distribution networks," *IEEE Transactions on Sustainable Energy*, 2018.

[19] N. H. Tran, C. T. Do, S. Ren, Z. Han, and C. S. Hong, "Incentive mechanisms for economic and emergency demand responses of colocation datacenters," *IEEE Journal on Selected Areas in Communications*, 2015.
[20] N. Chen, X. Ren, S. Ren, and A. Wierman, "Greening multi-tenant data center demand response," *Performance Evaluation*, vol. 91, pp. 229–254, 2015.
[21] N. H. Tran, T. Z. Oo, S. Ren, Z. Han, E.-N. Huh, and C. S. Hong, "Reward-to-reduce: An incentive mechanism for economic demand response of colocation datacenters," *IEEE Journal on Selected Areas in Communications*, 2016.
[22] Q. Sun, S. Ren, C. Wu, and Z. Li, "An online incentive mechanism for emergency demand response in geo-distributed colocation data centers," in *Proceedings of the Seventh International Conference on Future Energy Systems*. ACM, 2016.
[23] F. Kong and X. Liu, "Greenplanning: Optimal energy source selection and capacity planning for green datacenters," in *ACM/IEEE 7th International Conference on Cyber-Physical Systems (ICCPS)*. IEEE, 2016, pp. 1–10.
[24] A. Khosravi, L. L. Andrew, and R. Buyya, "Dynamic vm placement method for minimizing energy and carbon cost in geographically distributed cloud data centers," *IEEE Transactions on Sustainable Computing*, 2017.
[25] M. S. Hasan, Y. Kouki, T. Ledoux, and J.-L. Pazat, "Exploiting renewable sources: when green sla becomes a possible reality in cloud computing," *IEEE Transactions on Cloud Computing*, 2017.
[26] T. Basar and G. J. Olsder, *Dynamic noncooperative game theory*. Siam, 1999.
[27] Z. Liu, M. Lin, A. Wierman, S. H. Low, and L. L. Andrew, "Greening geographical load balancing," in *Proceedings of the ACM SIGMETRICS joint international conference on Measurement and modeling of computer systems*. ACM, 2011.
[28] F. Kong, W. Yi, and Q. Deng, "Energy-efficient scheduling of real-time tasks on cluster-based multicores," in *Design, Automation & Test in Europe Conference & Exhibition (DATE)*. IEEE, 2011.
[29] F. Kong and X. Liu, "Distributed deadline and renewable aware electric vehicle demand response in the smart grid," in *IEEE Real-Time Systems Symposium (RTSS)*, 2015.
[30] L. Chen, N. Li, S. H. Low, and J. C. Doyle, "Two market models for demand response in power networks," in *IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 2010.
[31] R. Johari and J. N. Tsitsiklis, "Parameterized supply function bidding: Equilibrium and efficiency," *Operations research*, 2011.
[32] F. Kong, X. Liu, and I. Lee, "Joint rate control and demand balancing for electric vehicle charging," in *IEEE/ACM Third International Conference on Internet-of-Things Design and Implementation (IoTDI)*. IEEE, 2018.
[33] M. Lin, A. Wierman, L. L. Andrew, and E. Thereska, "Dynamic right-sizing for power-proportional data centers," *IEEE/ACM Transactions on Networking (TON)*, 2013.
[34] A. Wierman, L. L. Andrew, and A. Tang, "Power-aware speed scaling in processor sharing systems," in *INFOCOM*, 2009.
[35] L. Rao, X. Liu, L. Xie, and W. Liu, "Minimizing electricity cost: optimization of distributed internet data centers in a multi-electricity-market environment," in *INFOCOM*. IEEE, 2010.
[36] Z. Liu, I. Liu, S. Low, and A. Wierman, "Pricing data center demand response," *ACM SIGMETRICS Performance Evaluation Review*, 2014.
[37] E. Le Sueur and G. Heiser, "Dynamic voltage and frequency scaling: The laws of diminishing returns," in *Proceedings of the international conference on Power aware computing and systems*, 2010.
[38] P. D. Klemperer and M. A. Meyer, "Supply function equilibria in oligopoly under uncertainty," *Econometrica: Journal of the Econometric Society*, pp. 1243–1277, 1989.
[39] A. Nagurney, *Network economics: A variational inequality approach*. Springer Science & Business Media, 2013, vol. 10.
[40] S. H. Low and D. E. Lapsley, "Optimization flow control. i. basic algorithm and convergence," *IEEE/ACM Transactions on networking*, vol. 7, no. 6, pp. 861–874, 1999.
[41] S. Shakkottai, R. Srikant *et al.*, "Network optimization and control," *Foundations and Trends® in Networking*, vol. 2, no. 3, pp. 271–379, 2008.
[42] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and distributed computation: numerical methods*. Prentice hall Englewood Cliffs, NJ, 1989.
[43] Z. Liu, A. Wierman, Y. Chen, B. Razon, and N. Chen, "Data center demand response: Avoiding the coincident peak via workload shifting and local generation," *Performance Evaluation*, vol. 70, no. 10, pp. 770–791, 2013.
[44] L. Gan, A. Wierman, U. Topcu, N. Chen, and S. H. Low, "Real-time deferrable load control: handling the uncertainties of renewable generation," in *Proceedings of the fourth international conference on Future energy systems*. ACM, 2013, pp. 113–124.
[45] A. Borodin and R. El-Yaniv, *Online computation and competitive analysis*. cambridge university press, 2005.