Incentivizing Energy Reduction for Emergency Demand Response in Multi-Tenant Mixed-Use Buildings

Nguyen H. Tran*, Chuan Pham*, Minh N. H. Nguyen*, Shaolei Ren‡, Choong Seon Hong*
*Department of Computer Science and Engineering, Kyung Hee University
‡Department of Electrical & Computer Engineering, University of California at Riverside

Abstract*—Emergency demand response, which is the last line of defense to avoid cascading failures during emergency events, has witnessed numerous crucial participants, including buildings and datacenters. However, even though the majority of datacenters are physically located in mixed-use buildings (MUBs), the existing studies on emergency demand response are non-coordinated approaches that separately focus on either buildings or datacenters, hence ignoring that both datacenters and non-datacenter (e.g., office) operations share the same MUB facilities (e.g., electricity supply). Furthermore, even when all MUB tenants (i.e., offices and datacenters) are jointly considered, tenants will incur different costs to shed energy for emergency demand response, thereby raising an issue of mis-aligned incentive for their participation. To overcome this *non-coordinated energy shedding and mis-aligned incentives*, we propose two incentive mechanisms in MUBs, such that the total incurred cost is minimized for energy shedding. The first mechanism, namely MECH-NA, is designed for non-strategic MUB tenants. In MECH-NA, the MUB operator provides a mechanism package including reward rate and a commitment profile with deviation penalty, based on which the MUB tenants will shed energy to maximize the reward and minimize their energy-shedding and deviation costs. We also design a distributed algorithm to implement MECH-NA that can achieve the minimum MUB cost. The second mechanism, namely MECH-SA, is a VCG-Kelly-based mechanism tailored to handle strategic MUB tenants. In MECH-SA, the operator announces both reward and energy shedding rules, based on which the tenants strategically participate in an bidding game. For this game, we not only show that there exists an efficient Nash equilibrium at which the total MUB cost is achieved, but also design a distributed algorithm to implement MECH-SA. Simulation results show that both MECH-NA and MECH-SA can obtain the optimal MUB cost, which outperforms partially or non-coordinated approaches.

I. Introduction

The power grid is becoming increasingly fragile with more generation volatility due to the aging infrastructure, frequent extreme weather, and/or wide adoption of renewables. While a generation-side solution (e.g., deploying utility-scale energy storage and reserves) is under development for smoothing the supply, it is generally capital-intensive. Hence, load-side solutions, also called demand response, have increased rapidly and become an effective technique to reshape consumer power demand via market-based approaches for grid stability in the face of time-varying electricity generation. In particular, emergency demand response (EDR) is a type of preventive demand response service that is called upon when power generation is anticipated to experience a shortfall. EDR has become one of the most widely-adopted demand response programs, representing 87% of demand reduction capabilities across all reliability regions [1]. EDR protects the power grid as the last line of defense against cascading blackouts by coordinating multiple energy consumers to shed their loads during emergency events. In the traditional sense, EDR occurs during natural disasters and/or extreme weathers. Nonetheless, as renewables are being increasingly incorporated into the grid and create more generation "shortfalls", EDR is expected to become more pervasive and important in stabilizing the grid and balancing supply and demand.

Ideal participants in the EDR programs are large yet flexible energy consumers, which include buildings and datacenters (DCs) [2]. In the U.S., buildings consume approximately 40% of the total generated electricity [3], while DCs have large yet flexible power demands [4]. However, *mixed-use buildings* (MUBs) housing both DC operations and non-DC operations (e.g., office spaces) have been largely overlooked in EDR programs. In fact, according to a report by Green Grid [5], "the majority of DCs are located within mixed-use facilities." A recent study [6] showed that large dedicated DCs (e.g., Google) only account for 4% of the total DC energy consumption, whereas the remaining 96% is used by other types of DCs (e.g., scientific computing cluster, colocation DCs, server rooms) that are mostly located in MUBs.

At present, buildings and DCs participate in EDR by means of their backup (diesel) generators which are economically costly and ecologically unfriendly. For example, as the most significant source of air pollution in California, diesel generation produces more than 50 times the amount of NOx and other toxic emissions as a typical power plant, and its usage is highly restricted [7]. Consequently, many have proposed to explore green alternatives for EDR. One approach proposed to reduce electricity consumption by cutting non-critical energy usage (e.g. heating, ventilation and air conditioning, or HVAC, switching off unused lighting) instead of backup diesel generation [8]. However, the flexibility of HVAC loads is often limited to a relatively small scope due to human comfort constraints. More importantly, the existing studies are not applicable to MUBs with DC operation since DCs are identified as "miscellaneous loads"; thus, the flexible electricity consumption of a DC is under-explored. A study based on real-world measurement states that DC operation can account for 50% or more of the total energy consumption of an MUB [9].

Similarly, the EDR for DCs has been extensively researched [4], [10], [11]. These works take advantage of widely available IT control knobs (e.g., turning on/off servers and workload migration). Further, a field study by Lawrence Berkeley National Laboratory has demonstrated that DCs can reduce energy consumption for EDR by 10-25%, without significantly impacting their normal operations [12]. Nonetheless, the existing research is targeted toward both owner-operated and multi-tenant DCs, where all the spaces and supporting infrastructure (e.g., cooling) are directly associated with DC operator [5]. Hence, despite sharing the main electrical power line in MUBs, the existing studies on demand response for buildings and DCs have been isolated from each other. This results in *non-coordinated* energy management, which leads to inefficient EDR for MUBs.

In this paper, we consider a multi-tenant MUB that houses a mix of both office tenants and DC tenants. Even when careful designers attempt to coordinate all tenants of the multi-tenant MUB, challenges escalate. Since electrical utilities cannot directly sub-meter each MUB tenant's energy usage, the EDR of MUB is directly managed by the MUB operator. However, while the operator only provides facilities (e.g., power, cooling), individual tenants that manage their own loads (e.g., DC servers, office HVACs) have individual costs and objectives to shed energy for EDR, inducing a *mis-aligned incentive* issue that further complicates the coordinated approach.

To overcome these challenges, we aim at approaches that not only coordinate both building offices and DCs, but also align their incentives to enable EDR in MUBs, such that the total incurred loss (e.g., latency performance degradation for DCs, thermal discomfort for office spaces) is minimized for energy shedding. During EDR, the MUB operator has multiple options: shedding DC energy (e.g., turning off unused servers, scaling down CPU frequency), shedding non-DC energy (e.g., increasing temperature set-point), and turning on on-site backup generation (usually, diesel generator). A *coordinated and incentive-aligned* approach is required because all three options mentioned have limitations and drawbacks: shedding DC energy can possibly degrade application performance, tuning HVAC temperature set-point for non-DC space results in human discomfort, while diesel generation contaminates the environment. Thus, the MUB operator must carefully control these three energy reduction knobs to minimize the overall negative impact while still satisfying the total energy reduction requirement for EDR.

Toward this end, we formulate an MUB cost minimization problem and propose two incentive-aligned mechanisms in which the MUB operator can coordinately incentivize the MUB tenants (i.e., non-DC, DC, and backup generator) for EDR such that the optimal HVAC temperature for non-DC space, server provisioning in DC, and usage of diesel generator can be obtained. We summarize our key contributions as follows:

- We propose the first mechanism, namely MECH-NA, for non-strategic MUB tenants. In this mechanism, the MUB operator supplies a package including reward rate and a commitment profile with a deviation penalty. Based on this package, the MUB tenants will shed energy to maximize the reward and minimize both energy-shedding and deviation costs. To enable the practicality of our approach, we also design a scalable and distributed algorithm that can achieve a MECH-NA equilibrium that is also the solution to the MUB cost minimization problem.
- We next propose the second mechanism, namely MECH-SA, for strategic tenants. MECH-SA is a VCG-Kelly-based mechanism that is tailored for EDR context in order to address the strategic behaviors of MUB tenants. In this mechanism, the operator first announces reward and energy shedding allocation rules resulted from a carefully chosen surrogate function. Then, the tenants strategically participate in an EDR bidding game driven by these allocation rules. For this game, we not only show that there exists an efficient Nash equilibrium at which the total MUB cost is achieved, but also design a scalable and distributed algorithm to implement MECH-SA.
- In order to validate the efficiency of the two proposed mechanisms, we conduct extensive simulations and show that both MECH-NA and MECH-SA can obtain the optimal MUB cost, which outperforms partially or non-coordinated approaches.

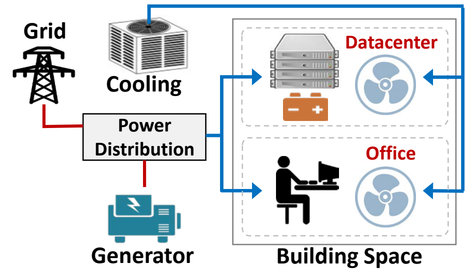


Fig. 1: A multi-tenant MUB architecture.

II. Background and Related Work

A. Background

Mixed-Use Building. In general, an MUB refers to a building with a combination of multiple distinct uses, such as residential, office, lab, and industrial, among others. In this paper, we explicitly focus on MUB that includes DC operation and a significant space for *non-DC* functions, i.e. offices. In an MUB with DC operation, the IT load usually accounts for 50% or more of a building's overall energy consumption [13]. Unless otherwise stated, we assume that the default function for the large non-DC space in an MUB is for offices[1]. Henceforth, "MUB" mentioned in this work refers to MUB with DC operations, which share the building space and main electrical power supply with the office operation. Servers often have dedicated backup power infrastructure (e.g., uninterrupted power supply, or UPS) for emergency. In some cases, a cooling system (e.g., chiller and cooling tower) may be shared between the DC and the office [5]. Note that DCs in an MUB are very diverse, ranging from state-of-the-art

[1]Dedicated DC, like Google, also has office space, but it is negligible compared to the DC part.

commercial DCs to scientific computing cluster and to small-/medium-size server rooms.

In general, there are two types of MUBs: (1) owner-operated, where both DC and non-DC parts are owned and managed by one entity/company (e.g., enterprise/campus building that has both DC and other usage); and (2) multi-tenant, where DC and non-DC parts belong to multiple different entities/companies, i.e., "tenants", which lease part of the building. In this paper, we focus on EDR by a multi-tenant MUB, which is more challenging than owner-operated MUBs due to the MUB operator's lack of control over tenants' loads. Fig. 1 illustrates a multi-tenant MUB architecture.

Emergency Demand Response. The EDR requires a mandatory power demand reduction response (with a significant penalty for non-compliance) for the participants, who are mostly paid for their availability for shedding loads even when no emergent signal is triggered [14]. At present, such programs are employed by many Independent System Operators, e.g., New England, where the customers' contracts can be established three years in advanced [15]. In particular, if there are some reliability issues in the grid (e.g., forecast capacity shortages or extreme weather event), the load serving entity (LSE) will trigger a signal to the customers from at least 10 minutes to one day in advance, and customers must comply with the notified reduction volume. The problem is that, at present, customers often participate in EDR using backup diesel generators, which is the least favorable choice both economically and ecologically. Hence, in this work, we present alternatives for the MUB operator for EDR by extracting the load reduction flexibilities enabled by recent advances in energy management, while only using the on-site diesel generation as a last resort.

B. Related Work

Energy management, both for DC and for building, has received much attention in the past.

For optimizing DC energy efficiency, many resource management approaches have been proposed and even implemented in real systems. These approaches include, but are not limited to, dynamically turning on/off servers for "energy proportionality" [16], [17], geographic load balancing to exploit location diversities for cost saving [18], [19], carbon footprint minimization [20], as well as brown energy reduction [21], [22], and cooling-aware scheduling [23], [24]. These studies, however, focus on energy management for dedicated DCs (e.g., Google), while isolating the significant amount of non-DC loads that co-exist in MUBs.

Likewise, for decades, researchers have developed numerous strategies to model, manage, and control the energy consumption for sustainability and/or reducing peak demand in different types of buildings [9]. Recent work includes energy management by taking into account both the energy demand and the thermal comfort of the occupants [25], [26]. Other common techniques for building energy management include lighting power reduction, global thermostat setpoint setback control, supply air temperature adjustment, pre-cooling, and use of a discharging energy storage device (e.g., battery) [8], [27]–[30]. However, the prior studies for (mixed-use) buildings treat DC loads as "miscellaneous" without judiciously exploring their high scheduling flexibilities.

As a promising load-side solution to stabilize the grid during generation shortfalls, extracting the load flexibilities of DCs [12], [31]–[33] and buildings [34] for EDR have received much attention recently. In particular, these studies have focused on achieving cost-effective and green EDR by reducing the usage of diesel generators. Nonetheless, the prior EDR work on DCs [12], [31]–[33] has only considered dedicated DCs, neglecting the non-DC loads that are physically colated with DCs but require distinctly different modeling and optimization approaches. While a preliminary study [34] has shown the effectiveness of coordinated energy management for cost-effective EDR by MUBs, it focuses on owner-operated MUBs, where the building manager has full control over all the control knobs. Thus, our work differs from and also is complementary to [34].

In this work, we focus on industrial/commercial MUBs with offices and datacenters, where HVAC and datacenter loads are the majority. Indeed, existing works on building energy management only consider HVAC loads, such as [8], [27]–[30]. Due to the large loads, the industrial/commercial MUBs are valid customers of EDR program. On the other hand, there are several works focusing on energy management of home appliances such as refrigerators, washing machine, electrical vehicle charging, etc., [35]–[37]. However, these works fall into a class of residential demand side management. Since the residential energy loads are usually small compared to industrial/commercial loads, the residential customers are not encouraged to participate in the EDR.

III. System Model

We consider a multi-tenant MUB composed of both office and DC tenants managed by one operator. We consider a one-period EDR, as in [11], [38]–[40], whose duration Δt is controlled by an LSE, e.g., 15 minutes or 1 hour.

A. HVAC Energy and User Comfort

We consider a set $\mathcal{N}_1$ of N_1 offices in the building. Each office has its own temperature setting knob, rather than being centrally controlled. Even though reducing the building office energy has attracted much attention, there exists a critical trade-off issue between building energy reduction and user comfort satisfaction, which dictates the performance of indoor environments. Generally, building energy is mainly contributed by: (a) HVAC system and (b) lighting and electrical equipment, which constitute 43% and 30% of building energy usage, respectively [3].

Energy reduction through HVAC control. Adopted from the energy-temperature correlation model in [41], the HVAC power of an office i, which depends on the difference between the mean outdoor temperature[2] T_i^{out} and indoor temperature

[2]We assume that all offices in the building have the same mean outdoor temperature. Furthermore, the temperature unit is Celsius.

T_i^{in} of office i controlled by an HVAC system, is given as follows:

$$q_{i,1}(T_i^{in}) = \frac{m_i}{M}\,|T_i^{in} - T_i^{out}|, \quad \forall i \in \mathcal{N}_1, \tag{1}$$

where m_i is the conductivity of the office i, and M is the power transformation that indicates the power efficiency of the HVAC system.

In practice, offices often put high priority on user comfort. Therefore, the indoor temperature is usually set to the comfort temperature $T^{cf}(T_i^{out})$, which is an affine function of the mean outdoor temperature [41]. Henceforth, for brevity, we will use T^{cf} without its argument. At this comfort temperature, the HVAC power consumption is $q_{i,1}(T^{cf})$. When the office tenants adjust the indoor temperature to T_i^{in} such that $q_{i,1}(T_i^{in}) \leq q_{i,1}(T^{cf})$ during the time slot Δt, the HVAC energy reduction is

$$\begin{aligned} e_i &:= \left(q_{i,1}(T^{cf}) - q_{i,1}(T_i^{in})\right)\Delta t \\ &= \kappa_i\,|T_i^{in} - T^{cf}|, \quad \forall i \in \mathcal{N}_1, \end{aligned} \tag{2}$$

where $\kappa_i := \frac{m_i}{M}\,\Delta t$.

Without loss of generality (w.l.o.g.), we consider the EDR time slot in summer (e.g., T_i^{out} = 32°C). Therefore, in order to have a non-negative reduced energy (i.e., $e_i \geq 0$) of the HVAC cooling system, we have

$$\bar{T}_i := T_i^{in} - T^{cf} \geq 0, \tag{3}$$

and we can rewrite (2) as an energy-temperature difference correlation as the following:

$$e_i = \kappa_i\,\bar{T}_i, \quad \forall i \in \mathcal{N}_1. \tag{4}$$

In this work, since HVAC is the prevailing energy consumption resource of buildings, we mainly consider the impact of an HVAC system on user comfort. However, our work can integrate other building consumption resources such as electrical lighting [42] or new applications of electric vehicle (EV) charging system [30]. By incorporating these loads into the model, we have additional decision variables and hence more scheduling freedom towards cost-efficient EDR.

User comfort model. Even though user comfort is an abstract concept and heavily depends on individual tastes, we consider a comfort model as in [41]. In this model, the user comfort consists of two components: (a) heat gain $G_i(t)$, which depends on the mean time t an office user spends in the room, and (b) heat loss $L_i(\bar{T}_i)$, which depends on the difference between indoor and comfort temperatures.

Since $\bar{T}_i \geq 0$ in our model, the heat loss model from [41] can be presented as follows:

$$L_i(\bar{T}_i) = \begin{cases} 3, & \bar{T}_i > R; \\ l\,\bar{T}_i, & 0 \leq \bar{T}_i \leq R, \end{cases} \tag{5}$$

where l and R are constants.

According to the AHSAH model [43], user comfort is the sum of heat gain and loss and is normalized in the range of $[-3, 3]$ encoding the feelings from cold to hot. Thus, we derive the user discomfort cost of an office i as follows:

$$C_i(\bar{T}_i) = \omega^{cf}\left(G_i(t) + L_i(\bar{T}_i)\right), \tag{6}$$

where ω^{cf} ($/heat loss discomfort) is the weight representing the unit discomfort cost of the office users due to the difference in indoor temperature and comfort temperature.

B. Datacenters

In our considered multi-tenant MUB, there is a set $\mathcal{N}_2$ of N_2 DC tenants, and each DC tenant $i \in \mathcal{N}_2$ manages S_i' homogeneous servers. A DC tenant with heterogeneous servers can be viewed as multiple virtual tenants, each having homogeneous servers.

Energy reduction of DCs. Even though DC tenants may use various control knobs (e.g., scaling down CPU frequencies, migrating loads to other places) for energy saving, the simple yet widely-studied approach that our study adopts as an example is turning off idle servers [11], [17], [40]. If tenant i has no intention to participate in EDR, all of its servers are active, and the workload will be evenly distributed to all servers to optimize performance [17]; hence, the total power consumption (i.e., cooling and IT) of this case is [11]

$$q_{i,2}(S_i') = S_i'\Big(p_{i,s} + p_{i,a}\frac{\lambda_i}{S_i'\mu_i}\Big)\,PUE, \tag{7}$$

where $p_{i,s}$ and $p_{i,a}$ are the static and active powers of each server, respectively, λ_i is the workload arrival rate, μ_i is a server's service rate measured in terms of the amount of workload processed per unit time, $\frac{\lambda_i}{S_i'\mu_i}$ is the server utilization with S_i' active servers, and PUE is the power usage effectiveness of a DC, which is measured by IT plus non-IT power consumption divided by IT power consumption. When participating in EDR by turning off s_i servers, the power consumption of DC tenant i is

$$q_{i,2}(s_i) = (S_i' - s_i)\left(p_{i,s} + p_{i,a}\frac{\lambda_i}{(S_i' - s_i)\mu_i}\right)PUE. \tag{8}$$

Therefore, the energy reduction by tenant i is

$$e_i := \left(q_{i,2}(S_i') - q_{i,2}(s_i)\right)\Delta t = \gamma_i s_i, \quad \forall i \in \mathcal{N}_2, \tag{9}$$

where $\gamma_i := p_{i,s}\,PUE\,\Delta t$ is a constant value of tenant i.

Turning servers off can have negative effects on tenant performance, inducing tenant costs. We rely on two typical costs that are widely used for tenants: the wear-and-tear cost and Service Level Agreement (SLA) cost [11], [17].

SLA cost. Since many Internet services hosted in DCs are sensitive to response/delay time, the SLA cost can be viewed proportionally to DC tenant mean response time. Using the M/M/1 queue, the mean response time of DC tenant i's workload is

$$d_i(s_i) := \frac{1}{\mu_i - \frac{\lambda_i}{S_i' - s_i}}. \tag{10}$$

We note that the queueing model has been widely used as a reasonable approximation for the actual service process [44], [45]. With this queueing model, we assume the tenant workload delay is negligible if it is below a predefined threshold d_0. Furthermore, $d_i(0) \leq d_0$, which means that DC tenants with full server capacity always guarantee the "best" SLA, i.e., no revenue loss for DC tenants due to service delay. On the other hand, DC tenant i has a "worst" SLA threshold D_i promised to its workload delay such that

$$d_i(s_i) \leq D_i, \tag{11}$$

which means

$$s_i + \frac{\lambda_i}{\mu_i - 1/D_i} \leq S'_i. \tag{12}$$

This can be interpreted to mean that DC tenant i needs at least $\frac{\lambda_i}{\mu_i-1/D_i}$ servers to guarantee a mean service delay of each workload less than D_i. Therefore, we assume that $0 \leq \frac{\lambda_i}{\mu_i-1/D_i} \leq S'_i$, and we have the following constraint of DC tenant i for turning off servers:

$$0 \leq s_i \leq S_i := S'_i - \frac{\lambda_i}{\mu_i - 1/D_i}. \tag{13}$$

The above constraint reflects the hard QoS constraint for a wide range of cloud applications such as multimedia, online gaming services [17], [19], [46]. When s_i increases, the workload distributed to the remaining active servers (i.e., $\frac{\lambda_i}{S'_i-s_i}$) increases due to the added migrating load, which leads to an increase in $d_i(s_i)$. With thousands of servers in a DC, we assume that s_i is a continuous variable [17].

Wear-and-tear cost. This cost occurs when tenants switch/toggle servers between active and idle states in every period and is linear with the number of turned-off servers [17]. Therefore, DC tenant i's total cost when turning off s_i servers is

$$C_i(s_i) = \omega^{wat}\, s_i + \omega^{sla}\, \lambda_i \max(d_i(s_i) - d_0, 0),\ \forall i \in \mathcal{N}_2, \tag{14}$$

where ω^{wat} (\$/server) and ω^{sla} (\$/delay) are the weights representing the unit cost of wear-and-tear and SLA, respectively.

C. Backup Generator

As HVAC and DCs together may not shed enough energy as required by the EDR, the MUB operator needs to resort to other control knobs, typically backup generators, to make up the energy reduction shortage. We assume that the MUB has a singleton set $\mathcal{N}_3$ of a generator.

Let $q_{i,3}$ and η'_i denote the power capacity and efficiency of the generator $i \in \mathcal{N}_3$, respectively, then its energy production during time slot Δt is

$$e_i = \eta_i\, q_{i,3}, \tag{15}$$

where $\eta_i := \eta'_i \Delta t$. Since a generator can control its power capacity to produce an amount of energy during a time slot, the backup generator cost is expressed as follows:

$$C_i(q_{i,3}) = \omega^{bg} \eta_i\, q_{i,3}, \quad i \in \mathcal{N}_3, \tag{16}$$

where ω^{bg} is the unit cost of the backup generation (e.g., diesel price).

D. EDR of MUB: Problem Formulation and Challenges

It is critical for the MUB to satisfy the EDR signals without causing much negative impact on the SLA performance of DC jobs or the user comfort in the building. Consequently, we consider the MUB's cost minimization problem for EDR as follows:

$$\mathbf{P_{mub}}: \quad \text{min.} \quad \sum_{i\in\mathcal{N}_1} C_i(\bar{T}_i) + \sum_{i\in\mathcal{N}_2} C_i(s_i) + \sum_{i\in\mathcal{N}_3} C_i(q_{i,3}) \tag{17}$$

$$\text{s.t.} \quad e_i = \kappa_i \bar{T}_i,\ \forall i \in \mathcal{N}_1, \tag{18}$$

$$e_i = \gamma_i s_i,\ \forall i \in \mathcal{N}_2, \tag{19}$$

$$e_i = \eta_i q_{i,3},\ \forall i \in \mathcal{N}_3, \tag{20}$$

$$\sum_{n\in\mathcal{N}_1} e_i + \sum_{i\in\mathcal{N}_2} e_i + \sum_{i\in\mathcal{N}_3} e_i = Q, \tag{21}$$

$$\text{var.} \quad 0 \leq \bar{T}_i \leq R, \quad \forall i \in \mathcal{N}_1, \tag{22}$$

$$0 \leq s_i \leq S_i, \quad \forall i \in \mathcal{N}_2, \tag{23}$$

$$0 \leq e_i \leq Q, \quad i \in \mathcal{N}_3. \tag{24}$$

In $\mathbf{P_{mub}}$, the objective is to minimize the MUB's total "cost" incurred for shedding energy for the EDR. The EDR is reflected in constraint (21) such that the MUB's response is equal to an an energy reduction target Q required by the LSE [40]. The MUB operator needs to solve a joint optimization problem by judicially optimizing server allocation, backup generator, and indoor temperature control such that the total MUB cost (which represents the overall negative impact of EDR) is minimized. In practice, these variables of $\mathbf{P_{mub}}$ are usually discrete; however, due to their large values, they can be approximated as continuous values and then rounded to the closest integers, similarly to existing works [39], [47]. Then, it is straightforward that $\mathbf{P_{mub}}$ is a convex problem that can be solved efficiently using the interior-point method. However, this centralized approach requires global access to all information of individual tenants, which may not be possible since all tenants are not willing to share their private data.

In practice, moreover, the MUB tenants are logically coupled (due to the shared total energy reduction required for EDR) but are physically self-managed in a separate manner. That is, they have their own controllers that are independent of each other. Thus, in order to utilize these independent controllers, a distributed implementation is essential. One may think about using the standard dual decomposition with (sub)gradient methods [48] for distributed optimization. However, since discomfort cost and backup generator cost are linear functions, the dual decomposition approach, which requires the cost function to be strictly convex, is not applicable [49]. Furthermore, shedding energy will increase tenant operational cost, raising the question of incentives for tenants' EDR participation.

To address those challenges, in the next two sections, we propose two incentive-aligned mechanisms that can coordinate two types of non-strategic and strategic MUB tenants to participate in the EDR such that the solutions of $\mathbf{P_{mub}}$ can be obtained.

IV. Incentive-aligned and Coordinated Mechanism for Non-strategic Tenants

In this section, we first design an incentive-aligned and coordinated mechanism for MUB non-strategic tenants. We then propose an algorithm that enables a practical implementation for this mechanism in a distributed and scalable manner.

A. Mechanism Design for Non-strategic Tenants

We denote an aggregated set of all MUB tenants by $\mathcal{N} = \mathcal{N}_1 \cup \mathcal{N}_2 \cup \mathcal{N}_3$. The mechanism for non-strategic tenants (MECH-NA) is described as follows:

MECH-NA:

1) **MUB operator** determines two components of the mechanism:
   (a) A *reward rate*, denoted by a vector $\boldsymbol{\delta} := \{\delta_i\}_{i\in\mathcal{N}}$, is the compensation price paid for every unit of energy shedding by tenants.
   (b) A *commitment*, characterized by a vector $\hat{\boldsymbol{e}} = \{\hat{e}_i\}_{i\in\mathcal{N}}$ such that
$$\mathbf{1}^T\hat{\boldsymbol{e}} = Q \tag{25}$$
   and a penalty $\rho(\boldsymbol{e} - \hat{\boldsymbol{e}})^2$ for commitment deviation where $\boldsymbol{e} = \{e_i\}_{i\in\mathcal{N}}$ is an energy shedding vector of MUB tenants.
2) **MUB tenants**, based on this mechanism components, decide their energy shedding as follows:
$$\max_{e_i\in[0,e_i^{max}]} \quad \delta_i e_i - c_i(e_i) - \frac{\rho}{2}(e_i - \hat{e}_i)^2, \tag{26}$$
   where
$$c_i(e_i) := \begin{cases} C_i(e_i/\kappa_i), & \forall i \in \mathcal{N}_1;\\ C_i(e_i/\gamma_i), & \forall i \in \mathcal{N}_2;\\ C_i(e_i/\eta_i), & \forall i \in \mathcal{N}_3, \end{cases} \tag{27}$$
   with its marginal cost
$$c_i'(e_i) = \begin{cases} \frac{\omega^{cf} l}{\kappa_i}, & \text{if } i \in \mathcal{N}_1;\\ \frac{\omega^{wat}}{\gamma_i} + \frac{\omega^{sla}}{\gamma_i}\frac{\lambda_i^2}{((S_i - e_i/\gamma_i)\mu_i - \lambda_i)^2}, & \text{if } i \in \mathcal{N}_2;\\ \frac{\omega^{bg}}{\eta_i}, & \text{if } i \in \mathcal{N}_3, \end{cases} \tag{28}$$
   and
$$e_i^{max} = \begin{cases} R/\kappa_i, & \forall i \in \mathcal{N}_1;\\ S_i/\gamma_i, & \forall i \in \mathcal{N}_2;\\ Q, & i \in \mathcal{N}_3. \end{cases} \tag{29}$$
3) **Mechanism equilibrium** is a vector $(\boldsymbol{\delta}^*, \hat{\boldsymbol{e}}^*, \boldsymbol{e}^*)$ satisfying (25) and (26) such that $\hat{\boldsymbol{e}}^* = \boldsymbol{e}^*$.

In MECH-NA, while the operator rewards tenants according to the commitment that satisfies the EDR signal, each tenant individually decides energy shedding to maximize its reward and minimize its costs, including its own cost and the commitment deviation cost. Even though the rationale behind the mechanism design is intuitive, its efficiency can be affected by many questions: Is there a mechanism equilibrium? If yes, then is it an optimal or sub-optimal solution to $\mathbf{P_{mub}}$? And how can this mechanism be implemented distributively?

To answer all the questions, we first propose a distributed algorithm for MECH-NA in Algorithm (Alg.) 1. Then we show that this algorithm can converge to an equilibrium of MECH-NA, which is an optimal solution to $\mathbf{P_{mub}}$.

Algorithm 1 Distributed Algorithm for MECH-NA

```
1: initialization: k = 0, set random positive (δ^(0), ê^(0));
2: repeat
3:     k ← k + 1;
4:     Tenant i decides e_i^(k) as in (30);
5:     The operator updates ê^(k) and δ^(k) by solving (31)
       and (32), respectively;
6: until ||δ^(k) − δ^(k−1)||_2 < ε.
```

B. Distributed Algorithm for MECH-NA

In MECH-NA, while MUB tenants rely on reward rate and commitment information to solve (26), the MUB operator must know the energy shedding of tenants to determine its reward rate and commitment. Thus, there is an interaction between MUB operator and tenants such that this mechanism cannot be implemented in only one round. Therefore, we propose the following iterative and distributed algorithm for the mechanism in Alg. 1. In principle, Alg. 1 enables the interaction in which the MUB operator and tenants exchange $(\hat{\boldsymbol{e}}, \boldsymbol{\delta})$ and $\boldsymbol{e}$, respectively, until they reach an agreement of $\hat{\boldsymbol{e}}^* = \boldsymbol{e}^*$.

Step 1: After receiving an EDR signal, the operator will announce the MECH-NA mechanism to all tenants.

Step 2: MUB tenants and operator interact sequentially as in line 4 and line 5 of Alg. 1, respectively. Specifically, MUB tenants solve the reward-minus-cost maximization problem (26) to obtain
$$e_i^{(k)} = \left[\hat{e}_i^{(k-1)} + \frac{\delta_i^{(k-1)}}{\rho} - \frac{c_i'(e_i^{(k)})}{\rho}\right]_{[0,e_i^{max}]} \tag{30}$$
where $[x]_{\mathcal{X}}$ denotes the projection of x onto the set $\mathcal{X}$. On the other hand, the operator will update the commitment $\hat{\boldsymbol{e}}^{(k)}$ as the solution to this problem
$$\begin{aligned} \text{min.} \quad & \hat{\boldsymbol{e}}^T\boldsymbol{\delta}^{(k-1)} + \frac{\rho}{2}\|\hat{\boldsymbol{e}} - \boldsymbol{e}^{(k)}\|_2^2 \\ \text{s.t.} \quad & \mathbf{1}^T\hat{\boldsymbol{e}} = Q, \\ \text{var.} \quad & \hat{\boldsymbol{e}} \geq 0 \end{aligned} \tag{31}$$
and update the reward rate as follows:
$$\boldsymbol{\delta}^{(k)} = \boldsymbol{\delta}^{(k-1)} + \rho(\hat{\boldsymbol{e}}^{(k)} - \boldsymbol{e}^{(k)}). \tag{32}$$

Step 3: If the interaction stops with the convergence condition in line 6, the MUB EDR proceeds: tenants shed their energy usage $\boldsymbol{e}^*$ and receive reward $\boldsymbol{\delta}^*$.

The design of Alg. 1 is based on basic principles of ADMM approach [49], which guarantees the convergence since $\mathbf{P_{mub}}$ is a convex problem.

Theorem 1. *Alg. 1 converges to a mechanism equilibrium, which is also a solution to* $\mathbf{P_{mub}}$.

Proof: Please see Appendix A. ■

Remarks:

1) The two-way communication exchanges between tenants and operator are composed of one-dimensional and two-dimensional information, respectively.
2) By submitting $e_i^{(k)}$, tenant i can keep its cost function $c_i(e_i)$ private, whereas the operator can hide the EDR signal Q and private information of other tenants from tenant i by only revealing the pair $(\delta_i^{(k)}, \hat{e}_i^{(k)})$. Furthermore, the reward rate is differentiated to tenants during iterations.
3) In each iteration, the operator needs to solve the strictly convex optimization (31) with an existing efficient algorithm such as the interior-point method. Otherwise, we can obtain the solution to this problem as

$$\hat{e}_i^{(k)} = \left[e_i^{(k)} - (\delta^{(k-1)} + \beta^{(k)})/\rho\right]_{[0,\infty]}, \tag{33}$$

where $\beta^{(k)}$ is the solution to the following equation

$$\sum\nolimits_{i\in\mathcal{N}} \left[e_i^{(k)} - (\delta^{(k-1)} + \beta^{(k)})/\rho\right]_{[0,\infty]} = Q, \tag{34}$$

which can be solved using a numerical method such as bisection search. In this case, the operator needs to calculate $\hat{e}_i^{(k)}$ for each tenant i. Thus, MECH-NA with Alg. 1 is scalable with complexity $O(N)$.

V. Incentive-aligned and Coordinated Mechanism for Strategic Tenants

In this section, we first tailor the VCG-Kelly mechanism to adapt to the MUB EDR context, based on which we next design an incentive-aligned and coordinated for MUB strategic tenants. Finally, we propose a scalable and distributed algorithm to enable practical implementation for this mechanism.

A. Tailored VCG-Kelly Mechanism for MUB EDR

We assume that all MUB tenants are strategic and bid for a total amount of Q energy shedding to receive compensation rewards from the MUB operator. Each tenant i submits its bid θ_i to the operator, representing its aggressiveness for energy shedding. We denote the bid vector of all tenants by $\boldsymbol{\theta} = \{\theta_i\}_{i\in\mathcal{N}}$. We also denote the bid vector of all tenants excluding i by $\boldsymbol{\theta}_{\text{-}\boldsymbol{i}} = \{\theta_j\}_{j\in\mathcal{N}-i}$. The original VCG-Kelly mechanism [50] is presented in the context of resource buyer's utility maximization, which is in sharp contrast to our context on cost minimization of energy-shedding tenants. Therefore, we tailor the VCG-Kelly mechanism for the MUB EDR as follows.

MUB Operator. Lacking cost information of all tenants, the operator assumes that the cost of tenant i is represented by a surrogate function $V_i(x_i;\theta_i)$. With the input $\boldsymbol{\theta}$, the operator solves the following energy shedding allocation problem such that the (assumed) social cost is minimized:

$$\begin{aligned} \mathbf{P_{vcgk}}: \text{min.} \quad & \sum\nolimits_{i\in\mathcal{N}} V_i(x_i;\theta_i) & (35)\\ \text{s.t.} \quad & \sum\nolimits_{i\in\mathcal{N}} x_i = Q, & (36)\\ \text{var.} \quad & x_i \geq 0, \quad \forall i \in \mathcal{N}. & (37) \end{aligned}$$

Denote the solution of this problem as $\boldsymbol{x}^*(\boldsymbol{\theta}) = \{x_i^*(\boldsymbol{\theta})\}_{i\in\mathcal{N}}$. Then, the operator sets the *energy shedding rule*:

$$e_i(\theta_i, \boldsymbol{\theta}_{\text{-}\boldsymbol{i}}) = x_i^*(\boldsymbol{\theta}), \tag{38}$$

and the *reward rule*:

$$r_i(\theta_i, \boldsymbol{\theta}_{\text{-}\boldsymbol{i}}) = \sum_{j\neq i} V_j(x_j^*(\boldsymbol{\theta}_{\text{-}\boldsymbol{i}}), \theta_j) - V_j(x_j^*(\boldsymbol{\theta}), \theta_j), \tag{39}$$

where $x_j^*(\boldsymbol{\theta}_{\text{-}\boldsymbol{i}})$, $\forall j \in \mathcal{N} - i$, is the solution to the problem $\mathbf{P_{vcgk}}$ with input $\boldsymbol{\theta}_{\text{-}\boldsymbol{i}}$.

MUB Tenants. The payoff function of tenant i with bid θ_i is given as the following:

$$\Pi_i(\theta_i, \boldsymbol{\theta}_{\text{-}\boldsymbol{i}}) = r_i(\theta_i, \boldsymbol{\theta}_{\text{-}\boldsymbol{i}}) - c_i(e_i(\theta_i, \boldsymbol{\theta}_{\text{-}\boldsymbol{i}})), \tag{40}$$

where $c_i(e_i(\theta_i, \boldsymbol{\theta}_{\text{-}\boldsymbol{i}}))$, $\forall i \in \mathcal{N}$, is the tenant cost function defined as in (27).

Since the tenants strategically maximize their payoffs by adjusting energy shedding bids, there exists an energy shedding game $\mathbf{G}(\mathcal{N}, \{\Theta_i\}, \{\Pi_i\})$ defined as follows:

- *Players*: MUB tenants in the set $\mathcal{N}$;
- *Strategy*: $\theta_i \in \Theta_i$, $\forall i \in \mathcal{N}$;
- *Payoff function*: $\Pi_i(\theta_i, \boldsymbol{\theta}_{\text{-}\boldsymbol{i}})$, $\forall i \in \mathcal{N}$

where $\Theta_i = [0, \theta_i^{max}]$ is the strategy space of tenant i such that

$$\theta_i^{max} = \{\bar{\theta}_i : e_i(\bar{\theta}_i, \boldsymbol{\theta}_{\text{-}\boldsymbol{i}}) \leq e_i^{max}\}, \quad \forall i \in \mathcal{N}. \tag{41}$$

For this game, a bidding profile $\boldsymbol{\theta}^*$ is called a Nash Equilibrium (NE) if and only if

$$\theta_i^* = \arg\max_{\theta_i\in\Theta_i} \Pi_i(\theta_i, \boldsymbol{\theta}^*_{-\boldsymbol{i}}), \quad \forall i \in \mathcal{N}. \tag{42}$$

Definition 1. An NE $\boldsymbol{\theta}^*$ of the energy shedding game $\mathbf{G}(\mathcal{N}, \{\Theta_i\}, \{\Pi_i\})$ is called *efficient* if $e_i(\theta_i^*, \boldsymbol{\theta}_{\text{-}\boldsymbol{i}}{}^*)$, $\forall i \in \mathcal{N}$, is also an optimal solution to $\mathbf{P_{mub}}$.

For a single operator and multiple tenants, the Kelly mechanism [51] specifies both resource allocation and payment rules with one-dimensional bids such that sum of non-strategic tenants' valuations is maximized (full efficiency). However, the Kelly mechanism is not strategy-proof: the strategic tenants can reflect their bids untruthfully to gain more reward, resulting in compromised efficiency. On the other hand, the VCG mechanism [52] is both efficient and strategy-proof such that a truthful valuation report is a dominant strategy for all tenants. However, the buyers' valuation functions are infinite-dimensional, which prevents the VCG mechanism from having a scalable implementation, let alone privacy violation. Therefore, the marriage of these two mechanisms, the VCG-Kelly, combines all benefits of Kelly and VCG mechanisms: efficiency, one-dimensional bids, and strategy-proof property.

For the VCG-Kelly-based mechanism design, we see that the payoff of tenants (40) depends on the energy shedding (38) and reward allocation (39) rules, which in turn depend on the solution to $\mathbf{P_{vcgk}}$. Hence, the existence of an NE of $\mathbf{G}\left(\mathcal{N},\{\Theta_i\},\{\Pi_i\}\right)$ depends on the surrogate function. Therefore, the key challenges are: a) how to choose a surrogate function such that the NE of $\mathbf{G}\left(\mathcal{N},\{\Theta_i\},\{\Pi_i\}\right)$ is efficient[3] and b) designing a scalable and distributed implementation to obtain an efficient NE. We will address both issues in the next two sub-sections.

B. Mechanism Design for Strategic Tenants

By carefully choosing a surrogate function (to be disclosed later), we propose a strategy-proof mechanism for strategic tenants (MECH-SA) as follows.

MECH-SA:

1) **MUB operator** announces two components of the mechanism:
   (a) Energy reduction rule:
$$e_i(\theta_i,\boldsymbol{\theta}_{\text{-}\boldsymbol{i}}) = x_i^*(\boldsymbol{\theta}) = \begin{cases} \frac{\theta_i}{\theta_i+\theta_{\text{-}i}^{\Sigma}}Q, & \theta_i>0;\\ 0, & \theta_i=0.\end{cases} \quad (43)$$
   (b) Reward rule:
$$r_i(\theta_i,\boldsymbol{\theta}_{\text{-}\boldsymbol{i}}) = \frac{Q\theta_{\text{-}i}^{\Sigma}}{\alpha+1}\left[\left(\theta_{\text{-}i}^{\Sigma}\right)^{-(1+\alpha)} - \left(\theta_i+\theta_{\text{-}i}^{\Sigma}\right)^{-(1+\alpha)}\right] \quad (44)$$
   where $\theta_{\text{-}i}^{\Sigma} := \sum_{j\in\mathcal{N}-i}\theta_j$.
2) **MUB tenants**, based on these mechanism components, choose their bids to maximize payoff $\Pi_i(\theta_i,\boldsymbol{\theta}_{\text{-}\boldsymbol{i}}), \forall i\in\mathcal{N}$, inducing the game $\mathbf{G}\left(\mathcal{N},\{\Theta_i\},\{\Pi_i\}\right)$.
3) **Mechanism equilibrium** is an NE point of $\mathbf{G}\left(\mathcal{N},\{\Theta_i\},\{\Pi_i\}\right)$.

We next reveal our surrogate function and analyze the efficiency of MECH-SA in the following results.

Lemma 1. *By choosing the surrogate function*
$$V_i(x_i;\theta_i) = \frac{\theta_i Q}{1+\alpha}\left(\frac{\theta_i Q}{x_i}\right)^{-(1+\alpha)}, \forall i\in\mathcal{N}, \quad (45)$$
we obtain MECH-SA *as a VCG-Kelly-based mechanism.*

Proof: Please see Appendix B. ∎

Assumption 1. We assume that two following inequalities are satisfied:
$$c'_{i\in\mathcal{N}_3}(0) > \min_{i\in\mathcal{N}_1\cup\mathcal{N}_2} c'_i(0), \quad (46)$$
$$Q > \max_{i\in\mathcal{N}_1\cup\mathcal{N}_2} e_i^{max}. \quad (47)$$

Since $c_i(0)=0, \forall i\in\mathcal{N}$, the first assumption (46) prevents $\mathbf{P_{mub}}$ solutions from the most trivial case in which the MUB only relies on the backup generator for EDR, as in current practice. The second assumption (47) indicates that Q is sufficiently large such that no tenants (except the generator) can perform the EDR alone. Assumption 1 is essential for the following result.

[3]We note that the celebrated sum-weighted-log surrogate function of the Kelly mechanism cannot be applied to the MUB EDR.

Algorithm 2 Distributed Algorithm for MECH-SA

1: **initialization:** $k=0$, set a random value $\theta_{\text{-}i}^{\Sigma(0)}>0$;
2: **repeat**
3: $\quad k \leftarrow k+1$;
4: $\quad$ Tenant i updates $\theta_i^{(k)}$ as in (51), $\forall i$;
5: $\quad$ Operator updates $\theta_{\text{-}i}^{\Sigma(k)}$ as in (52), $\forall i$;
6: **until** $\left|Q-\sum_{i\in\mathcal{I}} e_i\big(\theta_i^{(k)},\boldsymbol{\theta}_{\text{-}\boldsymbol{i}}{}^{(k)}\big)\right|<\epsilon$.

Theorem 2. *With Assumption 1,* MECH-SA *has an efficient NE $\boldsymbol{\theta}^*$ with a set $\mathcal{M}$ of M tenants having positive bids. Furthermore, we have:*
$$M\geq 2, \quad (48)$$
$$c'_i(0) \geq \sum_{j\in\mathcal{N}-i}\theta_j^*, \quad \forall i\in\mathcal{N}:\theta_i^*=0, \quad (49)$$
$$\left(c'_{i\in\mathcal{N}_3}(0)\right)^{-1/\alpha} \leq \sum\nolimits_{i\in\mathcal{M}}\theta_i^* \leq \left(\min_{i\in\mathcal{N}_1\cup\mathcal{N}_2} c'_i(0)\right)^{-1/\alpha}. \quad (50)$$

Proof: Please see Appendix C. ∎

There are three implications from Theorem 2 corresponding to (48), (49), and (50), respectively. First, if an NE is efficient, then it has at least two tenants with positive energy shedding. Second, a tenant i will not participate in EDR if its base marginal cost is higher than the aggregated bids of others at the equilibrium. And third, the sum of positive bids is bounded within a range that depends on α and base marginal costs.

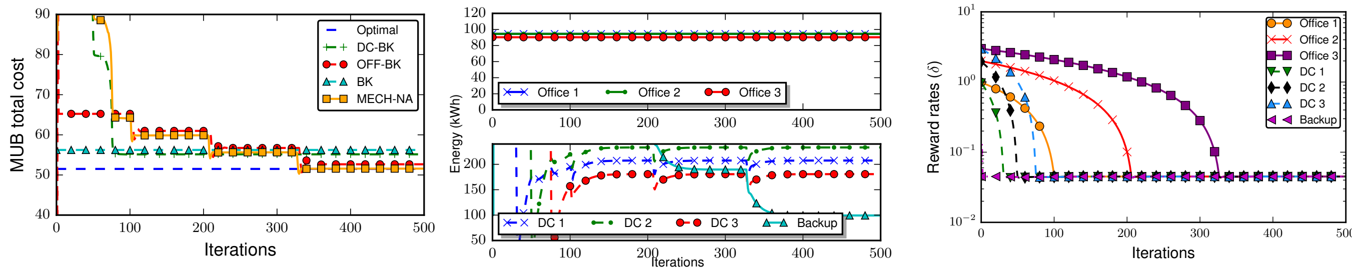
C. Distributed Algorithm for MECH-SA

We propose a distributed and scalable implementation for MECH-SA, which is presented in Alg. 2. The basic operations of this algorithm for the MUB EDR can be described as follows.

Step 1: After receiving an EDR signal, the operator will announce the MECH-SA rules (43) and (44) to all tenants.

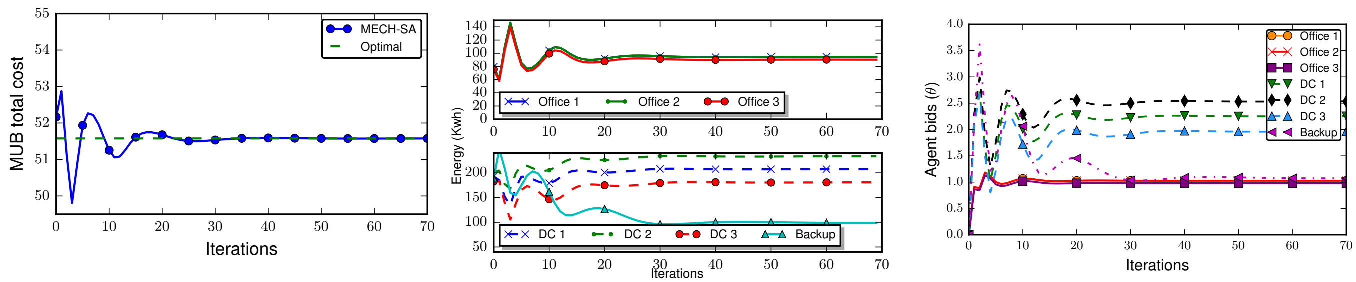
Step 2: MUB tenants and operator interact in line 4 and line 5 of Alg. 2, respectively, using simple communication exchanges: while each tenant strategically chooses its best response:
$$\theta_i^{(k)} = \left[\left(c'_i\Big(e_i\big(\theta_i^{(k)},\boldsymbol{\theta}_{\text{-}\boldsymbol{i}}{}^{(k-1)}\big)\Big)\right)^{-1/\alpha} - \theta_{\text{-}i}^{\Sigma(k-1)}\right]_{\Theta_i}, \quad (51)$$
the operator calculates $\theta_{\text{-}i}^{\Sigma}$ for each tenant i:
$$\theta_{\text{-}i}^{\Sigma(k)} = \left[\left(\delta^{(k)}\right)^{-1/\alpha} - \theta_i^{(k)}\right]_{[0,\infty]}, \forall i\in\mathcal{N}, \quad (52)$$
with
$$\delta^{(k)} = \left[\delta^{(k-1)} + \gamma^{(k)}\Big(Q-\sum\nolimits_{i\in\mathcal{N}} e_i\big(\theta_i^{(k)},\boldsymbol{\theta}_{\text{-}\boldsymbol{i}}{}^{(k-1)}\big)\Big)\right]_{[\epsilon,\infty]}. \quad (53)$$

Here ϵ is an arbitrarily small positive number to avoid undefined value in (52).



(a) MUB total cost convergence comparison. (b) Convergence of tenant energy shedding. (c) Convergence of reward rates.

Fig. 2: Convergence of MECH-NA in WeightSet 1, where all tenants have positive energy shedding.



(a) Convergence of MECH-SA to the optimal MUB cost. (b) Convergence of tenant energy shedding. (c) Convergence of tenant bids.

Fig. 3: Convergence of MECH-SA in WeightSet 1, where all tenants have positive energy shedding.

Step 3: If the interaction stops with the convergence condition in line 6, the EDR proceeds: tenant i reduces energy usage by an amount $e_i(\theta_i^*, \boldsymbol{\theta}_{\text{-}\boldsymbol{i}}{}^*)$ and receives its reward $R_i(\theta_i^*, \boldsymbol{\theta}_{\text{-}\boldsymbol{i}}{}^*)$.

The basic principles of Alg. 2 imitate the dual gradient method [53], which is widely used in both residential and commercial datacenter demand response, such as in [37] and [47], respectively.

Proposition 1. *Alg. 2 can converge to an efficient NE $\boldsymbol{\theta}^*$ with some conditions of the step-size rule and the curves of cost functions.*

Proof: Please see Appendix D. ■

Remarks:

1) The two-way exchanges between tenants and operator are one-dimensional bids.
2) With a single bid θ_i, tenant i can keep its cost function $c_i(e_i)$ private, whereas the operator can hide other individual bids from tenant i by revealing only the aggregated $\theta^{\Sigma}_{\text{-}i}$. Similar to MECH-NA, the reward of MECH-SA is differentiated to MUB tenants.
3) In each iteration, the operator needs to calculate $\theta^{\Sigma}_{\text{-}i}$ for each tenant i. Thus, Alg. 2 has the same complexity $O(N)$ as that of Alg. 1, which supports the scalability of MECH-SA.
4) Even though MECH-SA is strategy-proof and scalable with Alg. 2, it requires a stronger condition for NE existence (c.f. Assumption 1) and algorithm convergence (c.f. Proposition 1) than that of MECH-NA with Alg. 1.

VI. Simulation Results

In this section, we first describe the simulation settings and then present the results of the two proposed mechanisms to validate their efficacy.

A. Settings

For office tenants, the outdoor temperature is set to $T_i^{out} = 32^o$C for all case studies. We use the same model of [41] such that $T^{cf} = a + b \cdot T_i^{out}$, where a is set to 17.8, and b is set to 0.31. Thus, the comfort temperature T^{cf} is 27.7^{o}C. The heat gain $G_i(t)$ of each office is set in a range from 1.2 to 1.7, and its conductivity m_i is in a range from 200 to 450 J/s.C. Furthermore, we have $l = 2/7$ (discomfort/^{o}C), $R = 21/2^o$C, and $M = 0.05$. These values are conducted to emulate an approximate number of 100 occupants in an office [41].

For DC tenants, we set $S_i' = 2000$ servers, $\forall i$, and their corresponding service rates μ_i are increased from 2 to 6 jobs/s. We set $p_{i,a}$ and $p_{i,s}$ to 200 and 400 Watts, respectively, for all servers. The PUE is set to 1.5, which is typical for DCs [54]. The SLA delay threshold D_i is set to 1 second, which can be different depending the tenant's applications and services. In terms of DC tenant workload, we use the real trace of Facebook [55], from which we randomly collect 40 contiguous time slots. During considered time slots, workloads are normalized with respect to the total maximum capacity of tenants.

For the remaining parameters, unless stated otherwise, $Q = 1$ MWh for all cases. The backup generator efficiency is set to 1. For Alg. 1, we set $\rho = 10^{-5}$. For Alg. 2, the step size is set to $\gamma^{(k)} = 7 \cdot 10^{-5}/k$; moreover, we set $\alpha = 1.3$ for the surrogate function and initialize $\delta = 0.06$. Finally, $\epsilon = 10^{-4}$ in both algorithms.

B. Results

In order to evaluate the performance MECH-NA and MECH-SA, we compare the proposed mechanisms with the three following baselines:

1) BK: Similar to the current practice, this baseline only uses the backup generator for EDR, i.e., $e_{i\in\mathcal{N}_3} = Q$ and $e_i = 0$, $\forall i \in \mathcal{N}_1 \cup \mathcal{N}_2$.
2) DC-BK: Similar to the existing works [11], [47], [56], this baseline only relies on DCs and backup generator for EDR, i.e., $\bar{T}_i = 0$, $\forall i \in \mathcal{N}_1$, and constraint (21) is reduced to $\sum_{i\in\mathcal{N}_2} e_i + \sum_{i\in\mathcal{N}_3} e_i = Q$.
3) OFF-BK: As a partially coordinated approach for MUB EDR, this baseline only uses the HVAC control and the backup generator for EDR, i.e., $s_i = S_i'$, $\forall i \in \mathcal{N}_2$, and constraint (21) is reduced to $\sum_{n\in\mathcal{N}_1} e_i + \sum_{i\in\mathcal{N}_3} e_i = Q$.

Convergence and efficiency. We first illustrate the equilibrium convergence and optimality of MECH-NA with Alg. 1 and MECH-SA with Alg. 2 in Figs. 2 and 3, respectively. For ease of convergence observation, we first consider a small MUB system including three office tenants, three DC tenants and one backup generator. With a careful weight setting, namely WeightSet 1 that will be described later, these figures demonstrate an "ideal scenario," in which all tenants types have positive energy shedding.

Compared with other baselines, MECH-NA at equilibrium achieves the minimum cost, which is also the optimal MUB cost in Fig. 2a. Furthermore, the convergence of each tenant's energy shedding and the reward rates are shown in Figs. 2b and 2c, respectively. Similar to MECH-NA, MECH-SA at its equilibrium can obtain the optimal MUB cost and tenant energy shedding, as shown in Figs. 3a and 3b, respectively, whereas the corresponding convergent bids are shown in Fig. 3c.

Even though both proposed mechanisms, at their equilibria, can achieve the $\mathbf{P_{mub}}$ optimal objective, it can be observed in Figs. 2 and 3 that the convergence speeds of two mechanisms are different, where MECH-SA converges faster and smoother than MECH-NA. One of the reasons is the constant step-side of MECH-NA, whereas that of MECH-SA is fine-tuned diminishing.

Effect of weight parameters. We can see that the weight settings affect the $\mathbf{P_{mub}}$ solution and thus the equilibria of MECH-NA and MECH-SA. However, the unit-cost weights are individual parameters such that tenants have the freedom to decide the values at their own discretion. Hence, the operator cannot fully control the weights so that the $\mathbf{P_{mub}}$ solution can fall into an "ideal scenario." In order to evaluate the sensitivity of the proposed mechanisms to the weights, we classify the settings into the following cases.

1) WeightSet 1: As discussed above, this is an "ideal scenario" where all tenants have positive energy shedding, as in Figs. 2 and 3.
2) WeightSet 2: In this setting, the backup generator has sufficient high unit cost so that it is the last line for the EDR: the backup generator is only in use if all offices and DCs reach their maximum energy shedding. The effect of this setting to tenant energy shedding of MECH-NA and MECH-SA are shown in Figs. 4a and 4d, respectively.
3) WeightSet 3: In this setting, the office user comfort has sufficient high unit cost so that HVAC control is the last line for the EDR. The effects of this setting on tenant energy shedding of MECH-NA and MECH-SA are shown in Figs. 4b and 4e, respectively.
4) WeightSet 4: This setting gives DCs the first priority for EDR. The effects of this setting on tenant energy shedding of MECH-NA and MECH-SA are shown in Figs. 4c and 4f, respectively.

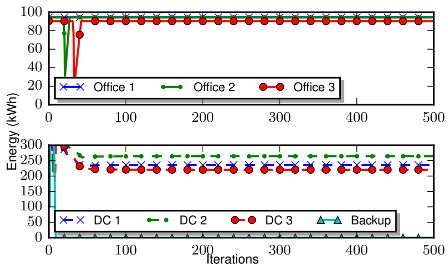
Despite various combinations of weight settings, Fig. 5 again shows that MECH-NA at its equilibrium achieves the minimum cost, which is also the optimal MUB cost of $\mathbf{P_{mub}}$, compared with other baselines. Intuitively, the weight settings affect to baseline performance: a) Fig. 5a shows that OFF-BK and DC-BK have higher costs than the others due to WeightSet 2, b) Fig. 5b shows that OFF-BK and OFF-DC have higher costs than the others due to WeightSet 3, and c) Fig. 5c shows that OFF-BK, without DCs for EDR, has the highest cost. Finally, Fig. 6 shows that MECH-SA at the equilibrium achieves the optimal MUB cost in all weight settings. Henceforth, since MECH-SA and MECH-NA have the same performance, we only present MECH-NA for all comparisons.

Effect of DC and non-DC workloads. Using a fixed weight setting and applying the Facebook trace for DC workloads in 40 time slots, we next evaluate the MUB cost performance of the proposed mechanisms in a new MUB setting with 10 offices and five DCs. We note that the dynamic DC workload also affects to the non-DC workload because the constraint (21) couples the EDR contributions of DCs and non-DC equipments. For example, when the DC has high workload and thus needs a large number of active servers, the EDR contribution of the DC is low, which forces the non-DC loads for EDR contribution to increase to satisfy the constraint (21). Fig. 7 shows that, even though the MUB cost varies with respect to DC and non-DC workload dynamics, MECH-NA can a) obtain the minimum MUB cost in all time slots and b) reduce the MUB total cost averaged over all time slots by up to 33%, 60%, and 73% compared with OFF-BK, DC-BK, and BK, respectively.

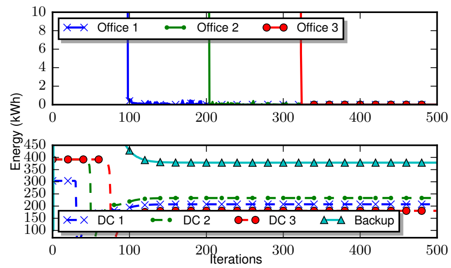
Effect of Q. Finally, using a fixed weight setting, we compare the proposed mechanism with other baselines by varying EDR signal Q. Fig. 8 shows that MECH-NA outperforms the partially and non-coordinated approaches. Especially, compared with current practice (BK), MECH-NA can reduce the MUB total cost up to 65.6%, 62.5%, and 59.4% when $Q = 1000, 1500$, and 2000 kWh, respectively. We see that when Q increases, the contribution of both DCs and offices to EDR decrease while that of backup generator increase, which can be explained by the increased DCs' wear-and-tear and SLA cost and office discomfort cost.

VII. Conclusions

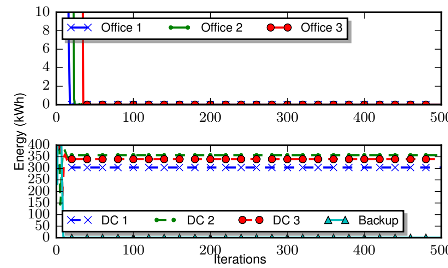
In this paper, we study MUB energy shedding for the EDR. In view that the existing studies on demand response by buildings and DCs have been largely isolated and resulted in



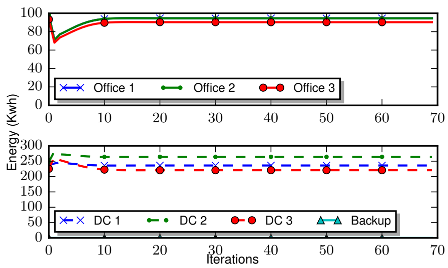
(a) MECH-NA, WeightSet 2



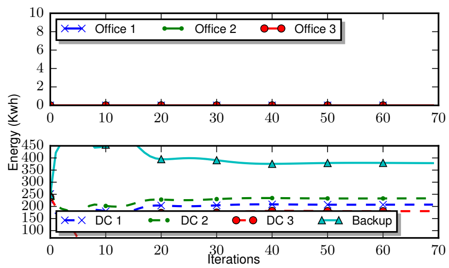
(b) MECH-NA, WeightSet 3



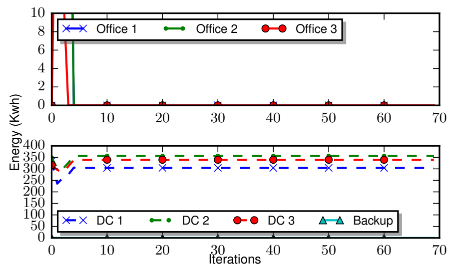
(c) MECH-NA, WeightSet 4



(d) MECH-SA, WeightSet 2

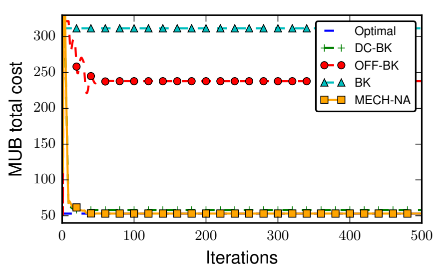


(e) MECH-SA, WeightSet 3

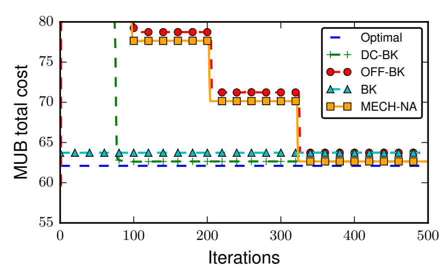


(f) MECH-SA, WeightSet 4

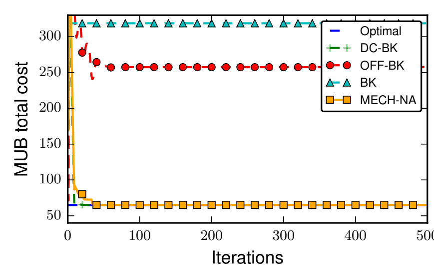
Fig. 4: Tenant energy shedding of MECH-NA and MECH-SA in WeightSets 2, 3, and 4.



(a) WeightSet 2



(b) WeightSet 3



(c) WeightSet 4

Fig. 5: MUB total cost of MECH-NA with WeightSets 2, 3 and 4.

non-coordinated energy management in MUBs, we propose two coordinated and incentive-aligned mechanisms to enable the EDR in MUBs, such that the total incurred loss (i.e., latency performance degradation for DCs, thermal discomfort for office, and diesel generation) is minimized for energy shedding during EDR. While the first mechanism is designed for non-strategic tenants, the second is designed to handle the strategic tenants. We also propose two scalable algorithms to implement the proposed mechanisms such that the mechanism equilibrium can be obtained in a distributed manner. We also conduct a case study to validate our proposed mechanisms, which achieve the optimal total incurred loss and thus outperform other partially or un-coordinated approaches.

Appendix A
Proof of Theorem 1

We first define

$$f(\boldsymbol{e}) = \sum\nolimits_{i\in\mathcal{N}} c_i(e_i). \tag{54}$$

Then $\mathbf{P_{mub}}$ can be presented in vector form as follows

$$\begin{aligned} \text{min.} \quad & f(\boldsymbol{e}) && (55)\\ \text{s.t.} \quad & \mathbf{1}^T\boldsymbol{e} = Q, && (56)\\ & 0 \le \boldsymbol{e} \le \boldsymbol{e}^{max}. && (57) \end{aligned}$$

Defining a set $\mathcal{E} = \{\boldsymbol{e} : 0 \le \boldsymbol{e}, \mathbf{1}^T\boldsymbol{e} = Q\}$ and a following indicator function

$$I_{\mathcal{E}}(\boldsymbol{e}) = \begin{cases} 0, & \boldsymbol{e} \in \mathcal{E}; \\ \infty, & \text{otherwise}, \end{cases} \tag{58}$$

we can rewrite $\mathbf{P_{mub}}$ as

$$\text{min.} \quad f(\boldsymbol{e}) + I_{\mathcal{E}}(\boldsymbol{e}) \tag{59}$$

By introducing auxiliary variable vector $\hat{e}$, we have an equivalent problem of $\mathbf{P_{mub}}$ as follows:

$$\begin{aligned} \text{min.} \quad & f(\boldsymbol{e}) + I_{\mathcal{E}}(\hat{\boldsymbol{e}}) && (60)\\ \text{s.t.} \quad & \boldsymbol{e} = \hat{\boldsymbol{e}}, && (61) \end{aligned}$$

with its augmented Lagrangian

$$L_\rho(\boldsymbol{e}, \hat{\boldsymbol{e}}, \boldsymbol{\delta}) = f(\boldsymbol{e}) + I_{\mathcal{E}}(\hat{\boldsymbol{e}}) + \boldsymbol{\delta}^T(\hat{\boldsymbol{e}} - \boldsymbol{e}) + \frac{\rho}{2}\|\hat{\boldsymbol{e}} - \boldsymbol{e}\|_2^2,$$

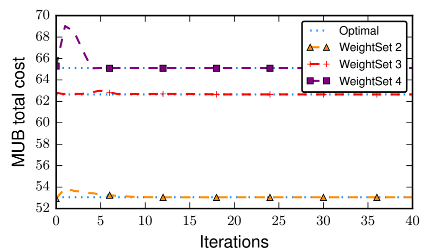


Fig. 6: MUB total cost of MECH-SA with WeightSets 2, 3, and 4.

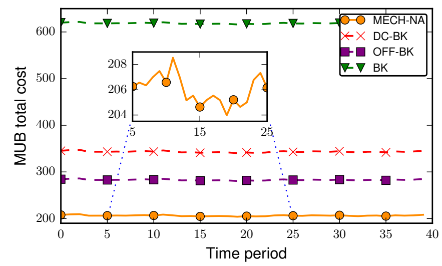


Fig. 7: MUB cost comparison with workload trace for DCs.

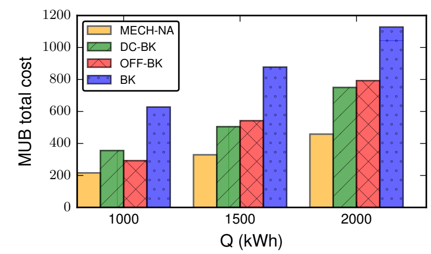


Fig. 8: MUB cost comparison with varying Q.

where ρ is a chosen parameter. Next, we obtain the following sequential decomposable updates:

$$\boldsymbol{e}^{(k)} = \arg\min_{0 \le \boldsymbol{e} \le \boldsymbol{e}^{max}} \left(f(\boldsymbol{e}) - \boldsymbol{e}^T \boldsymbol{\delta}^{(k-1)} + \frac{\rho}{2} \| \boldsymbol{e} - \hat{\boldsymbol{e}}^{(k-1)} \|_2^2 \right), \tag{62}$$

$$\hat{\boldsymbol{e}}^{(k)} = \arg\min \left(I_{\mathcal{E}}(\hat{\boldsymbol{e}}) + \hat{\boldsymbol{e}}^T \boldsymbol{\delta}^{(k-1)} + \frac{\rho}{2} \| \hat{\boldsymbol{e}} - \boldsymbol{e}^{(k)} \|_2^2 \right), \tag{63}$$

$$\boldsymbol{\delta}^{(k)} = \boldsymbol{\delta}^{(k-1)} + \rho(\hat{\boldsymbol{e}}^{(k)} - \boldsymbol{e}^{(k)}). \tag{64}$$

The first update (62) is the solution to the MUB tenants problems (26). Since all of these problems are convex, using the first-order condition, we can solve them to have the results (30) of Alg. 1. The second update (63) is the solution to the equivalent problem (31) of Alg. 1. Finally, because $\mathbf{P_{mub}}$ is a convex problem, the ADMM method can guarantee the convergence [49].

Appendix B
Proof of Lemma 1

With $V_i(x_i;\theta_i) = \frac{\theta_i Q}{1+\alpha} \left(\frac{\theta_i Q}{x_i} \right)^{-(1+\alpha)}, \forall i \in \mathcal{N}$, the Lagrangian of $\mathbf{P_{mub}}$ is

$$L(\{x_i\},\{\nu_i\},\zeta) = \sum_{i\in\mathcal{N}} \frac{\theta_i Q}{1+\alpha} \left(\frac{\theta_i Q}{x_i} \right)^{-(1+\alpha)} - \sum_{i\in\mathcal{N}} \nu_i x_i - \zeta \Big(\sum_{i\in\mathcal{N}} x_i - Q \Big). \tag{65}$$

The KKT condition [58] is

$$\begin{cases} \frac{\partial L}{\partial x_i} = (\theta_i Q)^{-\alpha} x_i^{-\alpha} - \nu_i - \zeta = 0, & \forall i \in \mathcal{N}, \\ \nu_i x_i = 0, & \forall i \in \mathcal{N}, \\ x_i \ge 0, \nu_i \ge 0, & \forall i \in \mathcal{N}, \\ \sum_i x_i = Q. \end{cases} \tag{66}$$

We see that, if $x_i > 0$, then $\nu_i = 0$; thus, we have

$$x_i = \theta_i Q \zeta^{1/\alpha} > 0. \tag{67}$$

Combining the above equation with the last equation of the KKT condition, we have $\zeta^{1/\alpha} = \frac{1}{\sum_{i\in\mathcal{N}} \theta_i}$, which is plugged into (67) to produce

$$x_i^*(\boldsymbol{\theta}) = \frac{\theta_i}{\sum_{i\in\mathcal{N}} \theta_i} Q. \tag{68}$$

Therefore, with this $x_i^*(\boldsymbol{\theta})$, the reward function (39) can be presented as follows:

$$\begin{aligned} r_i(\theta_i, \boldsymbol{\theta}_{\textbf{-}\boldsymbol{i}}) &= \sum_{j\neq i} V_j(x_j^*(\boldsymbol{\theta}_{\textbf{-}\boldsymbol{i}}), \theta_j) - V_j(x_j^*(\boldsymbol{\theta}), \theta_j) \\ &= \frac{Q\theta_{\text{-}i}^{\Sigma}}{\alpha+1} \left[\left(\theta_{\text{-}i}^{\Sigma}\right)^{-(1+\alpha)} - \left(\theta_i + \theta_{\text{-}i}^{\Sigma}\right)^{-(1+\alpha)} \right]. \end{aligned} \tag{69}$$

Appendix C
Proof of Theorem 2

The following lemma is an essential building block to prove Theorem 2.

Lemma 2. *If $\mathbf{P_{mub}}$ has a solution in which at least two tenants have positive energy shedding, then there exists a corresponding unique efficient NE with at least two positive bids of $\mathbf{G}\left(\mathcal{N}, \{\Theta_i\}, \{\Pi_i\}\right)$.*

Proof: We first find the optimality conditions for $\mathbf{P_{mub}}$. The Lagrangian of $\mathbf{P_{mub}}$ can be expressed as follows:

$$L(e_i, \nu_i, \zeta_i \delta) = \sum_{i\in\mathcal{N}} c_i(e_i) - \nu_i e_i + \zeta_i(e_i - e_i^{max}) + \delta \Big(\sum_{i\in\mathcal{N}} e_i - Q \Big), \tag{70}$$

then we have the KKT condition for $\mathbf{P_{mub}}$:

$$\begin{cases} c_i'(e_i) - \delta - \nu_i + \mu_i = 0, & \forall i \in \mathcal{N}, \\ \nu_i e_i = 0 \text{ and } \nu_i \ge 0, & \forall i \in \mathcal{N}, \\ \zeta_i(e_i - e_i^{max}) = 0 \text{ and } \zeta_i \ge 0, & \forall i \in \mathcal{N}, \\ \sum_{i\in\mathcal{N}} e_i = Q \text{ and } e_i \ge 0, & \forall i \in \mathcal{N}. \end{cases} \tag{71}$$

This condition specifies a pair of optimal primal-dual solutions $(\{e_i^*\}, \delta^*)$ such that:

$$\begin{cases} c_i'(e_i^*) = \delta^* > 0, & \text{if } 0 < e_i^* < e_i^{max}, \\ c_i'(e_i^{max}) \le \delta^*, & \text{if } e_i^* = e_i^{max}, \\ c_i'(0) \ge \delta^*, & \text{if } e_i^* = 0, \\ \sum_{i\in\mathcal{N}} e_i^* = Q. \end{cases} \tag{72}$$

We next find the efficient NE conditions for a bidding profile $\boldsymbol{\theta}^*$ of $\mathbf{G}\left(\mathcal{N}, \{\Theta_i\}, \{\Pi_i\}\right)$. We have

$$\frac{\partial u_i}{\partial \theta_i}(\theta_i^*, \boldsymbol{\theta}_{\textbf{-}\boldsymbol{i}}{}^*) = \frac{\theta_{\text{-}i}^{*\Sigma} D}{\left(\theta_i^* + \theta_{\text{-}i}^{*\Sigma}\right)^2} \Big(\left(\theta_i^* + \theta_{\text{-}i}^{*\Sigma}\right)^{-\alpha} - c_i'\big(e_i(\theta_i^*, \boldsymbol{\theta}_{\textbf{-}\boldsymbol{i}}{}^*)\big) \Big). \tag{73}$$

Defining

$$g\left(\sum_{i\in\mathcal{N}}\theta_i^*\right)=\left(\sum_{i\in\mathcal{N}}\theta_i^*\right)^{-\alpha}, \quad (74)$$

in order for a profile $\boldsymbol{\theta}^*$ to be an efficient NE according to (42), we use the following first-order condition:

$$\begin{cases} c_i'\big(e_i(\theta_i^*,\boldsymbol{\theta}_{\text{-}\boldsymbol{i}}^*)\big)=g(\theta_i^*+\theta_{\text{-}i}^{*\Sigma}), & \text{if } 0<\theta_i^*<\theta_i^{max},\forall i,\\ c_i'\big(e_i(\theta_i^{max},\boldsymbol{\theta}_{\text{-}\boldsymbol{i}}^*)\big)\le g(\theta_i^{max}+\theta_{\text{-}i}^{*\Sigma}), & \text{if } \theta_i^*=\theta_i^{max},\forall i,\\ c_i'\big(e_i(0,\boldsymbol{\theta}_{\text{-}\boldsymbol{i}}^*)\big)\ge g(0+\theta_{\text{-}i}^{*\Sigma}), & \text{if } \theta_i^*=0,\forall i,\\ \sum_{i\in\mathcal{N}}e_i(\theta_i^*,\boldsymbol{\theta}_{\text{-}\boldsymbol{i}}^*)=Q. \end{cases} \quad (75)$$

Comparing (72) and (75), we see that they have similar structures. We will show that, for a pair of $(\{e_i^*\},\delta^*)$, there exists a unique NE $\boldsymbol{\theta}^*$ such that $(\{e_i^*\},\delta^*)$ and $\boldsymbol{\theta}^*$ satisfy

$$g(\sum_{i\in\mathcal{N}}\theta_i^*)=\delta^*, \quad (76)$$

$$e_i(\theta_i^*,\boldsymbol{\theta}_{\text{-}\boldsymbol{i}}^*)=e_i^*,\forall i\in\mathcal{N}, \quad (77)$$

and vice versa. We first show the forward case. Since $g(.)$ is strictly decreasing and $\lim_{x\to 0}g(x)=\infty$, given δ^*, there exists a unique $\sum_{i\in\mathcal{N}}\theta_i^*=g^{-1}(\delta^*)>0$, where $g^{-1}(.)$ is the inverse of $g(.)$. Combining with (77), we obtain $\theta_i^*=e_i^*g^{-1}(\delta^*)/Q$, $\forall i$. For the reverse case, given an NE $\boldsymbol{\theta}^*$, we have a corresponding $(\{e_i^*\},\delta^*)$ according to (76) and (77).

Therefore, when $\mathbf{P_{mub}}$ has a solution $(\{e_i^*\},\delta^*)$ with at least two positive values of e_i, we see that there exists a corresponding $\boldsymbol{\theta}^*$ such that $\theta_{\text{-}i}^{\Sigma^*}>0$, $\forall i$. For a fixed $\boldsymbol{\theta}_{\text{-}\boldsymbol{i}}^*$, there exists a unique solution θ_i^* to (75), which is the solution to $\max_{\theta_i}u_i(\theta_i,\boldsymbol{\theta}_{\text{-}\boldsymbol{i}}^*)$, $\forall i$, because $g(\theta_i^*+\theta_{\text{-}i}^{*\Sigma})$ is strictly decreasing, and $c_i'\big(e_i(\theta_i^*,\boldsymbol{\theta}_{\text{-}\boldsymbol{i}}^*)\big)$ is strictly increasing with respect to θ_i^*. Therefore, (75) is a necessary and sufficient condition for this $\boldsymbol{\theta}^*$ to be a unique efficient NE corresponding to $(\{e_i^*\},\delta^*)$.

Finally, we show that, if $\boldsymbol{\theta}'$ has only one positive element, then it is not an NE. Assume an tenant i has $\theta_i'>0$ and $\theta_{-i}'=0$, then the reward to this tenant is zero. Therefore, this tenant will decrease its bid to 0. ■

Now we prove the main results of Theorem 2.

(a) First, with Assumption 1, we see that $\mathbf{P_{mub}}$ has at least two tenants with positive energy shedding; so, we can apply Lemma 2 to have $M\ge 2$.

(b) Second, we see that (49) is the result of the third inequality of (75).

(c) Finally, in the remaining part of the proof, we will show the correctness of inequality (50). From Lemma 2, for a pair $(\{e_i^*\},\delta^*)$ and the corresponding $\boldsymbol{\theta}^*$, we must have

$$g(\sum_{i\in\mathcal{N}}\theta_i^*)=\delta^*\le c_{i\in\mathcal{N}_3}'(0). \quad (78)$$

Otherwise, if $\delta^*>c_{i\in\mathcal{N}_3}'(0)$, then from the second line of (72) we have $e_i=Q$, $i\in\mathcal{N}_3$, so that $e_i=0$, $\forall i\in\mathcal{N}_1\cup\mathcal{N}_2$, contradicting Assumption 1. Since $g(.)$ is strictly decreasing and positive, (78) implies

$$\sum_{i\in\mathcal{M}}\theta_i^*\ge\left(c_{i\in\mathcal{N}_3}'(0)\right)^{-1/\alpha}. \quad (79)$$

On the other hand, from the first line of (72), we have

$$g(\sum_{i\in\mathcal{N}}\theta_i^*)=\delta^*=c_i'(e_i^*)\ge\min_{i\in\mathcal{N}_1\cup\mathcal{N}_2}c_i'(0) \quad (80)$$

since $c_i'(.)$ is positive non-decreasing, which implies

$$\sum_{i\in\mathcal{M}}\theta_i^*\le\left(\min_{i\in\mathcal{N}_1\cup\mathcal{N}_2}c_i'(0)\right)^{-1/\alpha}. \quad (81)$$

APPENDIX D
PROOF OF PROPOSITION 1

The best response (51) of tenant i chooses $\theta_i^{(k)}\in\Theta_i$ such that

$$c_i'\big(e_i(\theta_i^{(k)},\boldsymbol{\theta}_{\text{-}\boldsymbol{i}}^{k-1})\big)=(\theta_i^{(k)}+\theta_{\text{-}i}^{\Sigma(k-1)})^{-\alpha}=\delta^{(k-1)}, \quad (82)$$

where the second equation comes from (52), which is used by the operator to modulate the information $\theta_{\text{-}i}^{\Sigma}$ to all tenants. The update of $\delta^{(k)}$ is re-written here as

$$\delta^{(k)}=\left[\delta^{(k-1)}+\gamma^{(k)}\Big(Q-\sum_{i\in\mathcal{N}}e_i\big(\theta_i^{(k)},\boldsymbol{\theta}_{\text{-}\boldsymbol{i}}^{(k-1)}\big)\Big)\right]_{[\epsilon,\infty]}. \quad (83)$$

Both (82) and (83) imitate the dual gradient method [53] of $\mathbf{P_{mub}}$, whose convergence depends on step-sides rules $\gamma^{(k)}$ and/or the curve of cost function $c_i(e_i)$. By Lemma 2, when Alg. 2 converges to a $\mathbf{P_{mub}}$'s optimal primal-dual solution $(\{e_i^*\},\delta^*)$, it also corresponds to an efficient NE $\boldsymbol{\theta}^*$.

REFERENCES

[1] K. Managan, "Demand repsonse: A market overview," 2014, http://enaxisconsulting.com.
[2] EnerNOC, "Ensuring U.S. grid security and reliability: U.S. EPA's proposed emergency backup generator rule," 2013, http://www.whitehouse.gov/sites/default/files/omb/assets/oira_2060/2060_12102012-2.pdf.
[3] F. Jazizadeh, A. Ghahramani, B. Becerik-Gerber, T. Kichkaylo, and M. Orosz, "Human-Building Interaction Framework for Personalized Thermal Comfort-Driven Systems in Office Buildings," *Journal of Computing in Civil Engineering*, vol. 28, no. 1, pp. 2–16, jan 2014.
[4] A. Wierman, Z. Liu, and H. Mohsenian-Rad, "Opportunities and challenges for data center demand response," in *IEEE IGCC*, Dallas, TX, jun 2014.
[5] The Green Grid, "Pue: A comprehensive examination of the metric," 2012.
[6] NRDC, "Scaling up energy efficiency across the data center industry: Evaluating key drivers and barriers," *Issue Paper*, Aug. 2014.
[7] Santa Babara County, "Air Pollution Control District," http://www.ourair.org/do-you-really-need-a-diesel-generator/.
[8] T. Wei, T. Kim, S. Park, Q. Zhu, S. X.-D. Tan, N. Chang, S. Ula, and M. Maasoumy, "Battery management and application for energy-efficient buildings," in *DAC*, 2014.
[9] Y. Agarwal, T. Weng, and R. K. Gupta, "The energy dashboard: Improving the visibility of energy consumption at a campus-wide scale," in *BuildSys*, 2009.
[10] B. Aksanli and T. S. Rosing, "Providing regulation services and managing data center peak power budgets," in *DATE*, 2014.
[11] S. Ren and M. A. Islam, "Colocation demand response: Why do I turn off my servers?" in *USENIX ICAC*, Philadelphia, PA, jun 2014, pp. 201–208.
[12] G. Ghatikar, V. Ganti, N. E. Matson, and M. A. Piette, "Demand response opportunities and enabling technologies for data centers: Findings from field studies," 2012.
[13] S. K. Aggarwal, L. M. Saini, and A. Kumar, "Electricity price forecasting in deregulated markets: A review and evaluation," *International Journal of Electrical Power & Energy Systems*, vol. 31, no. 1, pp. 13–22, 2009.
[14] EnerNOC, "Demand response: A multi-purpose resource for utilities and grid operations," 2009. [Online]. Available: http://www.enernoc.com/our-resources/white-papers/demand-response-a-multi-purpose-resource-for-utilities-and-grid-operators
[15] PJM, "Emergency demand response (load management) performance report-2012/2013."
[16] Q. Wu, "Making facebook's software infrastructure more energy efficient with autoscale," 2014.

[17] M. Lin, A. Wierman, L. L. H. Andrew, and E. Thereska, "Dynamic Right-sizing for Power-proportional Data Centers," in *Proceedings IEEE INFOCOM*, Shanghai, China, apr 2011.
[18] A. Qureshi, R. Weber, H. Balakrishnan, J. Guttag, and B. Maggs, "Cutting the electric bill for internet-scale systems," in *SIGCOMM*, 2009.
[19] 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 *Proceedings IEEE INFOCOM*, San Diego, CA, mar 2010.
[20] P. X. Gao, A. R. Curtis, B. Wong, and S. Keshav, "It's not easy being green," *SIGCOMM Comput. Commun. Rev.*, 2012.
[21] I. Goiri, K. Le, T. D. Nguyen, J. Guitart, J. Torres, and R. Bianchini, "Greenhadoop: leveraging green energy in data-processing frameworks," in *EuroSys*, 2012.
[22] N. Deng, C. Stewart, D. Gmach, M. Arlitt, and J. Kelley, "Adaptive green hosting," in *ICAC*, 2012.
[23] J. Moore, J. Chase, P. Ranganathan, and R. Sharma, "Making scheduling "cool": Temperature-aware workload placement in data centers," in *USENIX ATC*, 2005.
[24] Q. Tang, S. K. S. Gupta, and G. Varsamopoulos, "Thermal-aware task scheduling for data centers through minimizing heat recirculation," in *CLUSTER*, 2007.
[25] L. A. Hurtado, P. H. Nguyen, W. L. Kling, and W. Zeiler, "Building Energy Management Systems - Optimization of comfort and energy use," in *2013 48th International Universities' Power Engineering Conference (UPEC)*. IEEE, sep 2013, pp. 1–6. [Online]. Available: http://ieeexplore.ieee.org/articleDetails.jsp?arnumber=6714910
[26] P. H. Shaikh, N. B. M. Nor, P. Nallagownden, I. Elamvazuthi, and T. Ibrahim, "A review on optimized control systems for building energy and comfort management of smart sustainable buildings," *Renewable and Sustainable Energy Reviews*, vol. 34, pp. 409–429, jun 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032114001889
[27] D. Irwin, S. Barker, A. Mishra, P. Shenoy, A. Wu, and J. Albrecht, "Exploiting home automation protocols for load monitoring in smart buildings," in *BuildSys*, 2011.
[28] A. Mishra, D. Irwin, P. Shenoy, J. Kurose, and T. Zhu, "Greencharge: Managing renewableenergy in smart buildings," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 7, pp. 1281–1293, Jul. 2013.
[29] M. Maasoumy and A. Sangiovanni-Vincentelli, "Modeling and optimal control algorithm design for HVAC systems in energy efficient buildings," 2011, https://escholarship.org/uc/item/0475s253.
[30] T. Wei, Q. Zhu, and M. Maasoumy, "Co-scheduling of HVAC control, EV charging and battery usage for building energy efficiency," in *ICCAD*, 2014.
[31] L. Zhang, S. Ren, C. Wu, and Z. Li, "A truthful incentive mechanism for emergency demand response in colocation data centers," in *INFOCOM*, 2015.
[32] N. Chen, X. Ren, S. Ren, and A. Wierman, "Greening multi-tenant data center demand response," in *IFIP Performance*, 2015.
[33] Z. Liu, I. Liu, S. Low, and A. Wierman, "Pricing data center demand response," in *SIGMETRICS*, 2014.
[34] N. H. Tran, C. Pham, S. Ren, and C.-S. Hong, "Coordinated energy management for emergency demand response in mixed-use buildings," in *GreeNets Workshop*, 2015.
[35] A.-H. Mohsenian-Rad, V. W. S. Wong, S. Member, J. Jatskevich, R. Schober, and A. Leon-Garcia, "Autonomous Demand-Side Management Based on Game-Theoretic Energy Consumption Scheduling for the Future Smart Grid," *IEEE Transactions on Smart Grid*, vol. 1, no. 3, pp. 320–331, dec 2010.
[36] A.-H. Mohsenian-Rad and A. Leon-Garcia, "Optimal Residential Load Control With Price Prediction in Real-Time Electricity Pricing Environments," *IEEE Transactions on Smart Grid*, vol. 1, no. 2, pp. 120–133, sep 2010.
[37] N. Gatsis and G. B. Giannakis, "Residential Load Control: Distributed Scheduling and Convergence With Lost AMI Messages," *IEEE Transactions on Smart Grid*, vol. 3, no. 2, pp. 770–786, jun 2012.
[38] M. Ghamkhari and H. Mohsenian-Rad, "Profit maximization and power management of green data centers supporting multiple slas," in *IEEE ICNC*, 2013, pp. 465–469.
[39] Z. Liu, M. Lin, A. Wierman, S. H. Low, and L. L. H. Andrew, "Greening geographical load balancing," in *Proc. ACM SIGMETRICS*, San Jose, CA, 2011, pp. 233–244.
[40] L. Zhang, S. Ren, C. Wu, and Z. Li, "A truthful incentive mechanism for emergency demand response in colocation data centers," in *IEEE INFOCOM*, Hong Kong, China, 2015.
[41] A. H.-y. Lam, Y. Yuan, and D. Wang, "An occupant-participatory approach for thermal comfort enhancement and energy conservation in buildings," in *Proc. ACM e-Energy*, New York, New York, USA, jun 2014, pp. 133–143.
[42] R. Yang and L. Wang, "Development of multi-agent system for building energy and comfort management based on occupant behaviors," *Energy and Buildings*, vol. 56, pp. 1–7, 2013.
[43] ASHRAE, "ASHRAE standard 55-2010:Thermal Environmental Conditions for Human Occupancy," 2010.
[44] Z. Liu, M. Lin, A. Wierman, S. H. Low, and L. L. H. Andrew, "Geographical load balancing with renewables," *ACM SIGMETRICS Performance Evaluation Review*, vol. 39, no. 3, pp. 62–66, dec 2011.
[45] A. Gandhi, M. Harchol-Balter, R. Das, and C. Lefurgy, "Optimal power allocation in server farms," *ACM SIGMETRICS Performance Evaluation Review*, vol. 37, no. 1, pp. 157–168, 2009.
[46] P. Wang, L. Rao, X. Liu, and Y. Qi, "D-pro: Dynamic data center operations with demand-responsive electricity prices in smart grid," *IEEE Transactions on Smart Grid*, vol. 3, no. 4, pp. 1743–1754, December 2012.
[47] 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*, vol. 33, no. 12, pp. 2892–2905, dec 2015.
[48] D. P. Palomar and M. Chiang, "Alternative Distributed Algorithms for Network Utility Maximization: Framework and Applications," vol. 52, no. 12, pp. 2254–2269, dec 2007.
[49] S. Boyd, "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, jan 2010.
[50] B. Hajek, "VCG-Kelly Mechanisms for Allocation of Divisible Goods: Adapting VCG Mechanisms to One-Dimensional Signals," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 6, pp. 1237–1243, aug 2007.
[51] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan, "Rate Control for Communication Networks: Shadow Prices, Proportional Fairness and Stability," *The Journal of the Operational Research Society*, vol. 49, no. 3, pp. 237–252, 1998.
[52] V. Krishna, *Auction Theory*. Academic Press, 2009. [Online]. Available: https://books.google.com/books?id=qW1128ktG1gC&pgis=1
[53] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and distributed computation: numerical methods*. Prentice-Hall, Inc., jan 1989.
[54] N. Chen, X. Ren, S. Ren, and A. Wierman, "Greening multi-tenant data center demand response," in *IFIP WG 7.3 Performance*, Sydney, Australia, oct 2015.
[55] "Facebook Dashboard: PUE & WUE." [Online]. Available: https:/PrinevilleDataCenter/app_399244020173259
[56] N. Chen, X. Ren, S. Ren, and A. Wierman, "Greening multi-tenant data center demand response," *Performance Evaluation*, vol. 91, pp. 229–254, sep 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0166531615000620
[57] M. Lubin and I. Dunning, "Computing in Operations Research Using Julia," *INFORMS Journal on Computing*, vol. 27, no. 2, pp. 238–248, apr 2015.
[58] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, mar 2004.