Data Center Decarbonization: A Hierarchical Scheme for Joint Online Electricity and Carbon Trading

Kekun Gao, Yuejun Yan, Yixuan Liu, Endong Liu, Zhaoyang Wang, Pengcheng You

***Abstract*—This paper studies the joint electricity and carbon trading strategy for a data center, aiming to reduce costs while integrating local renewable generation under uncertainty. To facilitate efficient and scalable decision-making, we propose a two-layer hierarchy where the lower layer focuses on the operation of all electrical equipment within the data center, while the upper layer manages the procurement of electricity and carbon product. At the lower layer, instead of device-level scheduling in real time, we exploit the inherent flexibility in demand, such as thermostatically controlled loads and flexible computing tasks, and aggregate them into virtual batteries. By this means, the upper-layer decision only needs to consider these virtual batteries, the size of which is generally small and independent of the data center scale. We further propose an online joint optimization algorithm based on Lyapunov optimization for purchasing electricity from the grid and carbon products from the carbon market at a manageable cost, despite uncertain and dynamic electricity prices, carbon product prices, renewable energy availability and virtual battery specifications. This approach fully leverages variations in price, carbon intensity and local renewable generation over time, as well as the inherent flexibility of data centers, to minimize costs while meeting carbon constraints. In particular, our algorithm under mild conditions can achieve bounded loss compared with the offline optimal cost, while strictly satisfying virtual battery operational constraints. Extensive simulation studies validate the theoretical analysis and illustrate the tradeoff between optimality and conservativeness.**

***Index Terms*—Lyapunov optimization, Online optimization, Flexibility aggregation, Virtual battery, Data center, Electricity market, Carbon trading.**

I. Introduction

Data centers, serving as the crucial infrastructure of the digital economy, are facing rapidly growing energy demands due to their expanding scale. According to the International Energy Agency, global data centers consumed between 240 and 340 TWh of electricity in 2022, accounting for about 1-1.3% of the world's total electricity demand [1]. The huge energy consumption results in high electricity bills and massive carbon emissions [2]. This highlights the core challenge for data centers to achieve sustainable development – decarbonization in an economically efficient manner.

This work was supported by NSFC through grants 72201007, 723B1001, 72431001, T2121002, 72131001, as well as CCF-Alibaba Innovative Research Fund. *(Corresponding author: Pengcheng You.)*

K. Gao, Y. Liu, and P. You are with the Department of Industrial Engineering and Management, College of Engineering, Peking University, Beijing, China (e-mail: pcyou@pku.edu.cn).

Y. Yan and Z. Wang are with Alibaba Group, Hangzhou, China.

E. Liu is with the School of Electrical and Mechanical Engineering, Heze University, Heze, China.

The cap-and-trade mechanism is a crucial approach towards such a goal, where data centers are required to cap their carbon emissions but can offset some excess emissions by purchasing carbon products (e.g. carbon allowances and Certified Emission Reductions (CERs)) in markets [3]–[7]. The associated expense of purchasing these products, referred to as "carbon cost", is a significant factor that data centers must account for. Additionally, electricity cost are also a critical factor, and the two are closely coupled. When electricity prices are low, data centers may increase electricity purchase to reduce total electricity cost, but this can raise carbon emissions if the grid's carbon emission intensity and carbon products price are high, leading to higher carbon costs. Conversely, low carbon emission intensity reduces carbon costs but may coincides with higher electricity prices. This interdependence between electricity and carbon costs creates a coupled relationship, where the decisions to minimize one type of cost inadvertently lead to increases in the other, ultimately driving up overall costs.

To reduce the huge electricity and carbon costs for data centers, an appealing option is to leverage the variations in market prices, grid's carbon emission intensity, renewable generations, and schedule the flexible devices accordingly. On the one hand, both electricity prices and carbon product prices exhibit real-time volatility, while the grid's carbon emission intensity also varies temporally [8]. Moreover, renewable energy sources integrated into data centers are characterized by high intermittency and volatility [9], [10]. On the other hand, a large portion of data center devices is temporally flexible, such as delay-tolerant computing loads [11], adjustable thermostatically controlling loads [12] and batteries. Therefore, by shifting these demands to periods with lower prices and "greener" electricity, data centers can effectively reduce both electricity and carbon costs.

A. Related Work

Currently, while most literature focuses on the electricity cost of data centers, joint management of electricity usage and carbon emissions is drawing increasing attention. Since carbon emissions result from electricity consumption, minimizing electricity costs and carbon costs are sometimes two conflicting goals. While low electricity prices prompt more electricity purchase, high grid carbon intensity or high carbon product prices at the same time may necessitate the opposite decision. To address this issue, some research has developed

models for carbon-aware data centers and proposed trade-off strategies to balance the reduction of electricity costs with the mitigation of carbon emissions [13]–[15]. They mainly utilize the temporal and spatial variability of electricity prices and carbon emission intensity to optimize workload scheduling. However, these research didn't account for the cap-and-trade mechanism, which will bring additional long-term carbon neutrality constraints and make the problem more complicated across timescales. Other studies, such as [16], [17], incorporate this mechanism to meet carbon neutrality constraints. However, they assume a fixed carbon emission budget throughout the budgeting period and only conduct a one-shot trade. In contrast, our work jointly and dynamically decides the purchase of electricity and carbon products, taking into account the time-varying carbon emission intensity and fluctuating prices in both electricity and carbon markets.

However, implementing these methods in hyperscale data centers with massive diverse flexible devices brings significant computational challenges due to handling diverse tasks and devices with heterogeneous resource and cooling demands. Therefore, incorporating device-level models into system-level cost optimization would dramatically increase computational complexity. To address this challenge, researchers have focused on aggregation methods that represent all flexible loads as a whole. This approach simplifies system-level decision-making, which only needs to consider the aggregate flexibility model. Currently, various aggregate flexibility models have been proposed, such as virtual battery approximation [18], convex polytopic approximation [19], [20] and many other approaches [21], [22]. Among these models, the virtual battery is particularly popular for its simple and compact form. It effectively captures the aggregate flexibility of various loads, such as temperature control systems, energy storage, pool pumps, electric vehicles, and deferrable tasks [18], [23]–[25]. However, existing aggregation methods mainly address offline scenarios, relying on complete information or perfect predictions of future uncertainties. In practice, future market prices, renewable generation, and energy demands are highly variable and hard to predict.

To address this problem, one stream of work utilizes model predictive control (MPC) [21], [22], [26], [27]. However, MPC-based methods require accurate statistical information of random parameters and predictions, and still exhibit high computational complexity. An alternative online optimization technique is Lyapunov optimization, which is prediction-free. Lyapunov optimization, originally developed for dynamic control of stochastic queuing systems [28], has been widely used for time-average stochastic optimization in various applications. It has been applied to manage different types of loads, such as energy storage [29]–[31], HVAC systems [32], and flexible workloads like data center batch tasks or EV charging, focusing on metrics like system cost, service delay, and operational constraints through the use of virtual queues to ensure these constraints are met [33]–[36]. However, almost all studies applied Lyapunov optimization to specific types of flexible loads, with tailored virtual queues. Data centers inherently possess massive heterogeneous devices and loads, requiring a more adaptable framework to handle this diversity. Thus, we extend these works by integrating a virtual battery into the Lyapunov framework, allowing for the aggregation of diverse flexible loads. However, unlike real batteries [29]–[31], virtual batteries have uncertain, time-varying specifications due to the inherent randomness of demands, and their dynamics involve non-negligible self-discharging. Consequently, existing algorithms urgently need improvement to effectively address these challenges.

B. Main Contributions

In this paper, we study the cost minimization problem of data centers with renewable generation, operating under uncertainties and subject to long-term carbon constraints. We propose a hierarchical scheme for joint online electricity and carbon trading, effectively reducing trading costs and enabling the data center to obtain a time-averaged cost close to the offline optimum. The main contributions of this work are summarized as follows:

- We propose a two-layer hierarchy, where the upper layer focuses on the coupling of electricity and carbon trading, and the lower layer aggregates flexible loads into dynamic virtual batteries. This structure enables efficient and scalable online decision-making by leveraging the flexibility of diverse data center devices while integrating various uncertain and time-varying factors such as market prices, carbon emission rates, renewable generation, and virtual battery specifications.
- We develop an online joint electricity and carbon trading algorithm for a data center that novelly incorporates a dynamic virtual battery into the Lyapunov optimization framework, and we enable the algorithm to adapt to the dynamic changes of the virtual battery, thereby ensuring its wide applicability to various types of loads.
- We conduct extensive simulation experiments based on real-life traces, revealing several interesting phenomena. For instance, our proposed algorithm achieves a 26.62% reduction in total cost compared to the online algorithm without management. We show that the proposed algorithm can reduce the average total cost to near-optimal levels, while ensuring feasibility at all times. Furthermore, we demonstrate the impact of the weight factor V and other important parameters, providing deeper insights into our algorithm.

The rest of this paper is organized as follows: Section II describes the system model and problem formulation. Section III proposes an online algorithm based on Lyapunov optimization. Section IV shows the numerical results, and Section V concludes the paper.

II. System Model

Consider the system operation of a data center in a discrete-time horizon $t \in \{0, 1, 2, ...\}$, which uses electricity power from a utility grid and local renewable generation to meet its energy demand and simultaneously purchases carbon allowance and CERs to neutralize the carbon emissions generated from electricity usage. We divide the overall demand into two categories: one is *controllable*, e.g., TCLs of cooling

systems and flexible computing tasks of IT equipment, and the other is *non-controllable* including necessary power to support basic services, uninterruptible computing tasks with priority, etc. We propose a two-layer hierarchy to aggregate controllable demand into (lower-layer) virtual batteries such that the flexibility can be exploited in a computationally efficient way to manage (upper-layer) electricity and carbon trading. A schematic diagram of the system model is given in Figure 1. We develop the detailed model for each component of the system below.

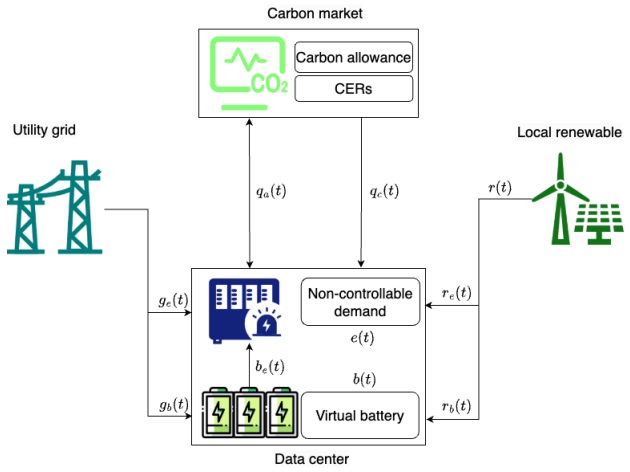

Fig. 1: Electricity and carbon flows in a data center with virtual batteries.

A. Virtual Battery

In this subsection, we use TCLs and flexible computing tasks, the two most common loads in the data center, as examples to show how virtual batteries can be aggregated. In particular, a virtual battery is specified by a tuple of parameters $(b_{char}(t), b_{dis}(t), b_{min}(t), b_{max}(t), \alpha)$, where $b_{char}(t)$/$b_{dis}(t)$ is the charge/discharge rate limit, $b_{min}(t)$/$b_{max}(t)$ is the lower/upper bound of battery capacity and α is the dissipation rate. Given such a specification, a virtual battery is modeled by two variables - the battery State-of-Charge (SoC) $b(t)$ and the energy injection/withdrawal $u(t)$ - subject to the following:

$$-b_{dis}(t) \le u(t) \le b_{char}(t), \quad \forall t, \tag{1a}$$

$$b(t+1) = \alpha b(t) + u(t), \quad \forall t, \tag{1b}$$

$$b_{min}(t) \le b(t) \le b_{max}(t), \quad \forall t. \tag{1c}$$

TCL: The role of a TCL is to maintain the required room temperatures for IT equipment. Such a requirement is typically captured by a set of thermodynamics equations and operational constraints:

$$\theta(t+1) = \alpha\theta(t) + (1-\alpha)\left(\theta_a(t) + cr(t) - bp(t)\right), \quad \forall t, \tag{2a}$$

$$\theta_r - \Delta \le \theta(t) \le \theta_r + \Delta, \quad \forall t, \tag{2b}$$

$$0 \le p(t) \le p_m, \quad \forall t. \tag{2c}$$

Here the variables are the room temperature $\theta(t)$ and the operating power $p(t)$ of the TCL. The external inputs include the operating power of IT devices in the room $r(t)$ and the ambient temperature $\theta_a(t)$. The parameters include the room temperature setpoint θ_r with a tolerance gap Δ, the maximum TCL power p_m, the coefficient c that reflects the temperature rise per unit of IT power, and the coefficient b that reflects the temperature drop per unit of the TCL power [18], [37].

Note that there exists a nominal value $p_o(t)$ for $p(t)$ that can set the immediate room temperature to the setpoint θ_r, given by

$$p_o(t) = \frac{\theta_a(t) + cr(t) - \theta_r}{b}.$$

By defining

$$b(t) := \frac{\theta_r - \theta(t)}{(1-\alpha)b}, \quad u(t) := p(t) - p_o(t),$$

we obtain

$$-p_o(t) \le u(t) \le p_m - p_o(t), \quad \forall t, \tag{3a}$$

$$b(t+1) = \alpha b(t) + u(t), \quad \forall t, \tag{3b}$$

$$-\frac{\Delta}{(1-\alpha)b} \le b(t) \le \frac{\Delta}{(1-\alpha)b}, \quad \forall t. \tag{3c}$$

It is obvious from (3) that the flexibility of a TCL can be represented as a virtual battery, parameterized by

$$\left(p_m - p_o(t),\ p_o(t),\ -\frac{\Delta}{(1-\alpha)b},\ \frac{\Delta}{(1-\alpha)b},\ \alpha\right).$$

Delay-Tolerant Computing Task: A lot of SQL or machine learning tasks are flexible with the processing schedule, as long as they are completed before a specified deadline. Consider a set of such tasks, with each task j parameterized by $(a^j, d^j, \bar{l}^j, e^j)$, denoting the arrival time, the deadline, the power consumption corresponding to the maximum processing speed, and the total energy required to complete the task, respectively. Let $l^j(t)$ be the allocated power to task j at time t. Then the task can be characterized as

$$0 \le l^j(t) \le \bar{l}^j, \quad \forall t \in [a^j, d^j), \tag{4a}$$

$$l^j(t) = 0, \quad \forall t \notin [a^j, d^j), \tag{4b}$$

$$\sum_t l^j(t) = e^j. \tag{4c}$$

The aggregate flexibility of this set of tasks can also be approximated as a virtual battery [18], [23]. Define in this case

$$b(t) := \sum_{\tau<t}\sum_j l^j(\tau), \quad u(t) := \sum_j l^j(t).$$

Obviously, $u(t)$ cannot exceed the sum of the maximum power consumption of all active tasks, i.e.,

$$0 \le u(t) \le \sum_{j:a^j \le t < d^j} \bar{l}^j. \tag{5}$$

Meanwhile, to guarantee that all active tasks at time t can be completed by their deadlines, the SoC $b(t)$ has to achieve a minimum level given by

$$b(t) \ge \sum_{j:d^j \le t} e^j + \sum_{j:a^j \le t < d^j} \max\left\{e^j - \left(d^j - t\right)\bar{l}^j,\ 0\right\},$$

where we have accounted for the energy delivered for all finished tasks up until time t. Note that the lower bound implies that each active task j will be allocated the maximum processing power from now on and can receive at most $(d^j - t)\bar{l}^j$ amount of energy. Similarly, we can derive an upper bound for the SoC $b(t)$ as

$$b(t) \leq \sum_{j:d^j \leq t} e^j + \sum_{j:a^j \leq t < d^j} \min\left\{e^j, \left(t - a^j\right)\bar{l}^j\right\}.$$

From above, the parameters of this aggregated virtual battery are available with a dissipation rate $\alpha = 1$.

Remark 1: As shown from the two virtual battery examples, the parameterization is likely time-varying and uncertain, depending on random ambient temperatures and computing task arrivals, etc. Therefore, managing such dynamic virtual batteries from flexibility aggregation is more challenging than operating real batteries with static specifications.

Note that multiple batteries can be readily aggregated into one [18]. Without loss of generality, we consider only one virtual battery in this work for ease of presentation. The analysis and results also generalize to the concurrent management of multiple virtual batteries. Specifically, in our setting, electricity from both local renewable generation and the utility grid can be used to charge the virtual battery.

We use $r_b(t)$ and $g_b(t)$ to denote respectively the electricity from the two sources. Besides, we define $b_e(t)$ to represent the energy withdrawal from the virtual battery to serve the non-controllable demand in the data center. Therefore, the characterization (1) of the virtual battery can be explicitly written as

$$0 \leq g_b(t) + r_b(t) \leq b_{char}(t), \quad \forall t, \tag{6a}$$

$$g_b(t) \geq 0, \quad r_b(t) \geq 0, \quad \forall t, \tag{6b}$$

$$0 \leq b_e(t) \leq b_{dis}(t), \quad \forall t, \tag{6c}$$

$$b(t+1) = b(t) + g_b(t) + r_b(t) - b_e(t) - (1-\alpha)b(t), \quad \forall t, \tag{6d}$$

$$b_{min}(t) \leq b(t) \leq b_{max}(t), \quad \forall t, \tag{6e}$$

where $b(0)$ is the given initial SoC. Furthermore, simultaneous charge and discharge are not allowed:

$$g_b(t) + r_b(t) > 0 \perp b_e(t) > 0, \quad \forall t. \tag{7}$$

B. Long-Term Carbon Constraint

The data center draws electricity from a utility grid with a time-varying carbon emission rate $\gamma(t) \geq 0$. The electricity that directly meets the non-controllable energy demand is denoted by $g_e(t)$, while $g_b(t)$ represents the electricity purchased for charging in the virtual battery. Thus, the carbon emissions from electricity purchase are given by $\gamma(t)(g_e(t) + g_b(t))$.

We consider a cap-and-trade system where the data center must ensure its cumulative carbon emissions over a given compliance period do not exceed a predetermined carbon allowance. Any excess emissions beyond this allowance require the purchase of additional carbon allowances or CERs from carbon markets to compensate for the shortfall. A compliance period, e.g., one year, is typically way longer than a trading period, e.g., 5 minutes, i.e., a compliance period could consist of over a hundred thousand trading periods. Therefore, instead of enforcing the explicit long-term carbon allowance limit, we propose a long-term time-average carbon constraint as follows:

$$\begin{aligned}&\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\{\gamma(t)[g_e(t) + g_b(t)]\} \\ \leq & A + \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\{q_a(t) + q_c(t)\},\end{aligned} \tag{8}$$

where A denotes a long-term time-averaged carbon allowance, and $q_a(t)$ and $q_c(t)$ denote the quantity of carbon allowance and CERs purchased from carbon markets, respectively. Note that an efficacious A could be well estimated by averaging the predetermined requirement over the compliance period. Since the compliance period is sufficiently long, (8) closely approximate the required cap on long-term carbon emissions.

The traded quantity is assumed bounded subject to market rules, and for simplicity, we assume that CERs are traded only once and can only be purchased, i.e.,

$$-\bar{q}_a \leq q_a(t) \leq \bar{q}_a, \quad \forall t, \tag{9}$$

$$0 \leq q_c(t) \leq \bar{q}_c, \quad \forall t, \tag{10}$$

where $\bar{q}_a$ and $\bar{q}_c$ are the maximum carbon allowance and the maximum CERs allowed to trade, respectively. Besides, some countries impose policies that restrict the carbon emissions offset by CERs below a specified proportion, denoted as η, of the total emissions [38], [39]. We account for such a restriction using also a long-term time-averaged constraint:

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\{q_c(t)\} \leq \eta \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\{\gamma(t)[g_b(t) + g_e(t)]\}. \tag{11}$$

C. Problem Formulation

Suppose the data center is equipped with local renewable generators, such as roof-top solar panels, which can provide an $r(t)$ amount of renewable energy at time t. $r(t)$ is random and intermittent in real time, and is difficult to forecast accurately. However, it shall always be upper bounded by inverters' capacity r_{max}, i.e.,

$$0 \leq r(t) \leq r_{max}.$$

The electricity generated from renewables is divided into three parts: $r_e(t)$ is directly supplied to the data center to meet its (non-controllable) demand; $r_b(t)$ can be stored in the virtual battery whenever necessary; the rest is just abandoned. This is captured by

$$0 \leq r_e(t) + r_b(t) \leq r(t), \quad \forall t, \tag{12a}$$

$$r_e(t) \geq 0, \quad \forall t. \tag{12b}$$

Denote the non-controllable demand of the data center at time t as $e(t)$. In total, it is satisfied by a combination of electricity supply from the local renewable generation, the virtual battery, and the utility grid, i.e.,

$$e(t) = r_e(t) + b_e(t) + g_e(t), \quad \forall t. \tag{13}$$

Note that selling electricity back to the utility grid is temporarily not allowed and the electricity purchase has an upper limit, i.e.,

$$g_e(t) \geq 0, \quad \forall t, \tag{14}$$

$$g_e(t) + g_b(t) \leq g_{max}, \quad \forall t. \tag{15}$$

We assume implicitly that the electricity available from the market is always sufficient to meet the data center demand and charge the virtual battery, i.e.,

$$g_{max} \geq \max_t\{e(t) + b_{char}(t)\}.$$

Let $p_e(t) \geq 0$, $p_a(t) \geq 0$, and $p_c(t) \geq 0$ denote the time-varying unit price of electricity, carbon allowance, and CER, respectively. Then the total energy and carbon cost for purchasing electricity and offsetting carbon emission is given by

$$c(t) = p_e(t)[g_e(t) + g_b(t)] + p_a(t)q_a(t) + p_c(t)q_c(t). \tag{16}$$

We formulate the long-term offline trading problem of the data center to minimize the time-average expected energy and carbon cost while satisfying all the operational constraints under uncertainty:

Offline Problem

$$\min \quad \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\{c(t)\} \tag{17a}$$

$$\text{s.t.} \quad (6) - (15) \tag{17b}$$

Note that the offline problem (17) is an offline stochastic optimization problem. However, in practice it is impossible to directly solve the offline problem (17) without the knowledge of uncertainty. We thus seek for an online solution in the next section.

III. Online Algorithm

A. Problem Relaxation

The challenges of the offline problem (17) are twofold: First, not only are the energy demand $e(t)$, electricity prices $p_e(t)$, carbon allowance prices $p_a(t)$, CERs prices $p_c(t)$, and renewable generation $r(t)$ all random and difficult to predict accurately, but also the aggregated virtual battery parameters $(b_{char}(t), b_{dis}(t), b_{min}(t), b_{max}(t))$ are also dynamic and stochastic in nature. These parameters are not known in advance and will be sequentially revealed online. Second, we observe that not only are the decision variables $g_b(t)$, $g_e(t)$ and $b_e(t)$ temporally coupled due to the virtual battery's SoC constraints, but also $g_b(t)$, $g_e(t)$, $q_a(t)$ and $q_c(t)$ are interdependent due to the long-term carbon compliance constraints. Therefore, the convoluted decision-making problem imperatively calls for an efficacious online algorithm.

To address the above issues, we use Lyapunov optimization to design an online algorithm that achieves the optimal solution to the offline problem (17) asymptotically. Since the capacity constraint (6e) couples the entire time horizon, it makes attaining a good online algorithm difficult. Thus, we first relax the capacity constraint (6e) that is imposed at each time t into a time-average form consistent with (8) and (11). By summing up (6d) over time, taking expectation of both sides and then taking the average, we can obtain the following:

$$\begin{aligned}
&\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\{[g_b(t) + r_b(t) - b_e(t)]\} \\
&= \lim_{T\to\infty} \frac{1}{T} [\mathbb{E}\{b(T)\} - b(0) + \sum_{t=0}^{T-1} (1-\alpha)\mathbb{E}\{b(t)\}] \\
&= \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} (1-\alpha)\mathbb{E}\{b(t)\},
\end{aligned} \tag{18}$$

where the effect of $b(T) - b(0)$ diminishes as T goes to infinity in the last equality. Define $b_{max} := \max_t b_{max}(t)$, $b_{min} := \min_t b_{min}(t)$ and assume they exist and are finite. Given the capacity constraint (6e), any feasible charging/discharging decisions will also satisfy the following:

$$(1-\alpha)b_{min} \leq \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\{[g_b(t) + r_b(t) - b_e(t)]\} \leq (1-\alpha)b_{max}. \tag{19}$$

Note that the converse does not hold. Therefore, (19) is a relaxation of (6e).

From above, we propose the following relaxation problem to replace the offline problem:

Relaxed Offline Problem

$$\min \quad \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\{c(t)\} \tag{20a}$$

$$\text{s.t.} \quad (6a) - (6d), (7), (9) - (15), (19) \tag{20b}$$

This ensures that the optimal value of the relaxed offline problem (20) is always less than or equal to that of the offline problem (17). Let Y^{OPT} and Y^{REL} denote the optimal objective value of the offline problem (17) and the relaxed offline problem (20), respectively, where $Y^{REL} \leq Y^{OPT}$. Although we relax the constraint of the offline problem (17), the subsequent online algorithm is designed to account for this relaxation, ensuring that the original constraints are consistently satisfied.

B. Real-time Decision

In this subsection, we employ Lyapunov optimization to transform the relaxed offline problem into an online version that does not depend on the statistical properties of $e(t)$, $r(t)$, $p_e(t)$, $p_a(t)$, $p_c(t)$, $b_{min}(t)$, $b_{max}(t)$. This approach enables real-time decisions and obtains an approximate optimal solution to the original problem.

To handle long-term time-averaged constraints in the relaxed offline problem (20), we introduce virtual queues. These virtual queues decompose the long-term time-averaged constraints in the problem into each time t, allowing us to remove the long-term constraints and only focus on stabilizing the virtual queues at each time t. Here, stabilizing the queues means keeping their values bounded over time. In the following, we will explain why maintaining this stability ensures compliance with the associated long-term constraints. First, we define two virtual queues, $Q_a(t)$ and $Q_c(t)$, which are associated with the long-term carbon constraints (8) and (11), respectively.

$$Q_a(t+1) = max\{Q_a(t) - q_a(t) - q_c(t) - A + \gamma(t)[g_e(t) + g_b(t)], 0\}, \tag{21}$$

$$Q_c(t+1) = max\{Q_c(t) + q_c(t) - \eta\gamma(t)[g_b(t) + g_e(t)], 0\}. \quad (22)$$

Intuitively, the virtual queues $Q_a(t)$ and $Q_c(t)$ represent the historical accumulation of discrepancies: $Q_a(t)$ tracks the difference between carbon emissions and offset products, which increases with electricity purchases $\gamma(t)[g_e(t) + g_b(t)]$ and decreases with the allowance A and carbon products purchases $q_a(t)$ and $q_c(t)$. Meanwhile, $Q_c(t)$ reflects the backlog of surplus CERs, increasing with CERs purchases and decreasing with the offsetting of total carbon emissions. We further set the initial values $Q_a(0) = 0$ and $Q_c(0) = 0$. Specifically, from (21) and (22), we observe that:

$$Q_a(t+1) \geq Q_a(t) - q_a(t) - q_c(t) - A + \gamma(t)[g_e(t) + g_b(t)], \quad (23)$$

$$Q_c(t+1) \geq Q_c(t) + q_c(t) - \eta\gamma(t)[g_b(t) + g_e(t)]. \quad (24)$$

By taking the expectation of both sides and summing over $t \in \{0, 1, 2, ..., T-1\}$, we can get

$$\begin{aligned} &\mathbb{E}\{Q_a(T)\} - \mathbb{E}\{Q_a(0)\} \\ &\geq \sum_{t=0}^{T-1} \mathbb{E}\{\gamma(t)[g_e(t) + g_b(t)] - q_a(t) - q_c(t) - A\}, \end{aligned}$$

$$\mathbb{E}\{Q_c(T)\} - \mathbb{E}\{Q_c(0)\} \geq \sum_{t=0}^{T-1} \mathbb{E}\{q_c(t) - \eta\gamma(t)[g_e(t) + g_b(t)]\}.$$

Then, by dividing both sides by T and letting $T \to \infty$, if $Q_a(t)$ and $Q_c(t)$ are stable and finite, constraints (8) and (11) will be satisfied (we will prove the stability of the virtual queues $Q_a(t)$ and $Q_c(t)$ later).

Second, we define a virtual queue Q_v associated with the virtual battery SoC constraints. It differs slightly from the way we construct $Q_a(t)$ and $Q_c(t)$. We set $Q_v(t)$ as a constant shift from $b(t) - b_{min}(t)$:

$$Q_v(t) := b(t) - b_{min}(t) + Q_0, \quad \forall t \geq 1, \quad (25)$$

$$Q_v(0) := Q_0, \quad (26)$$

where Q_0 is a constant and will be specified later in Theorem 1. By stabilizing $Q_v(t)$ and carefully selecting Q_0, we effectively control the distance between $b(t)$ and its lower bound $b_{min}(t)$, ensuring that $b(t)$ evolves in alignment with $b_{min}(t)$. This approach helps us satisfy the constraint $b_{min}(t) \leq b(t) \leq b_{max}(t)$. It is particularly effective when $b_{min}(t)$ and $b_{max}(t)$ follow similar trends, which is common in virtual batteries from TCLs.

Combining (6d), (25) and (26), we obtain the evolution of $Q_v(t)$ as follows:

$$Q_v(t+1) = Q_v(t) + g_b(t) + r_b(t) - b_e(t) - (1-\alpha)b(t) - \varepsilon(t), \quad \forall t, \quad (27)$$

with $\varepsilon(t) := b_{min}(t+1) - b_{min}(t)$ for $t \geq 1$ and $\varepsilon(0) := b_{min}(1) - b(0)$.

After introducing the queues, we define $\Theta(t) = [Q_v(t), Q_a(t), Q_c(t)]$ as the state vector of all virtual queues. The Lyapunov function is defined as

$$L(\Theta(t)) = \frac{1}{2}[Q_v^2(t) + Q_a^2(t) + Q_c^2(t)], \quad (28)$$

which provides a scalar measurement of the sizes of the queues. The conditional one-slot Lyapunov drift is then given by the expected increment:

$$\Delta(\Theta(t)) = \mathbb{E}\{L(\Theta(t+1)) - L(\Theta(t)) | \Theta(t)\}. \quad (29)$$

It is shown in [28] that greedily minimizing the conditional one-slot Lyapunov drift at each time t ensures the stability of all queues, thereby guaranteeing the satisfaction of the corresponding long-term constraints. However, minimizing the Lyapunov drift only achieve the long term time-average constraints, but it may also result in high electricity and carbon costs. Therefore, we trade the expected cost (16) in one time slot off against the Lyapunov drift (29), as expressed by:

$$\Delta(\Theta(t)) + V\mathbb{E}\{c(t) \mid \Theta(t)\}, \quad (30)$$

where V is a non-negative weight parameter that balances the trade-off between the objective function and constraint feasibility. Instead of directly using the drift-plus-cost function, we take a linear upper bound of the Lyapunov drift to design our online algorithm, which facilitates the solution process and contributes to both the analysis of the solution and the performance guarantees. To derive this upper bound, we will first use the following lemma.

Proposition 1: The one-slot conditional Lyapunov drift plus cost satisfies:

$$\begin{aligned} &\Delta(\Theta(t)) + V\mathbb{E}\{c(t)|\Theta(t)\} \\ \leq &B + \frac{1}{2}(1-\alpha)^2 b^2(t) + \varepsilon(1-\alpha)b(t) \\ &+ V\mathbb{E}\{p_e(t)[g_b(t) + g_e(t)] + p_a(t)q_a(t) + p_c(t)q_c(t)|\Theta(t)\} \\ &+ Q_v(t)\mathbb{E}\{g_b(t) + r_b(t) - b_e(t) - (1-\alpha)b(t) - \varepsilon(t)|\Theta(t)\} \\ &- (1-\alpha)\mathbb{E}\{b(t)[g_b(t) + r_b(t) - b_e(t)]|\Theta(t)\} \\ &+ Q_a(t)\mathbb{E}\{\gamma(t)[g_e(t) + g_b(t)] - q_a(t) - q_c(t) - A|\Theta(t)\} \\ &+ Q_c(t)\mathbb{E}\{q_c(t) - \eta\gamma(t)[g_b(t) + g_e(t)]|\Theta(t)\}, \end{aligned} \quad (31)$$

where $b_{char} := \max_t b_{char}(t)$ and $b_{dis} := \max_t b_{dis}(t)$ represent the maximum charge and discharge rates over time, respectively, $\varepsilon := \max_t |\varepsilon(t)|$ indicates the maximum variation in the lower bound of virtual battery capacity, and $\gamma_{max} := \max_t \gamma(t)$ is the maximum carbon emission rate from the utility grid.

The proof is provided in Appendix A.

Using this upper bound, we obtain the following optimization problem:

Online Problem

$$\begin{aligned} &\min \\ &[Q_v(t) - (1-\alpha)b(t) + Vp_e(t) + \gamma(t)Q_a(t) - \eta\gamma(t)Q_c(t)]g_b(t) \\ &- [Q_v(t) - (1-\alpha)b(t) + Vp_e(t) + \gamma(t)Q_a(t) - \eta\gamma(t)Q_c(t)]b_e(t) \\ &- [Vp_e(t) + \gamma(t)Q_a(t) - \eta\gamma(t)Q_c(t)]r_e(t) \\ &+ [Q_v(t) - (1-\alpha)b(t)]r_b(t) + [Vp_a(t) - Q_a(t)]q_a(t) \\ &+ [Vp_c(t) - Q_a(t) + Q_c(t)]q_c(t) \end{aligned} \quad (32)$$

$$\text{s.t. (6a) - (6d), (7), (9), (10), (12)- (15)}$$

The objective represents the weighted sum of the upper bound on the Lyapunov drift and the electricity and carbon costs. This

optimization problem at each t is a linear programming problem that depends only on the current system and queue states, allowing the online problem (32) to be solved without future information. We develop an online algorithm that decides $\{r_e(t), r_b(t), g_e(t), g_b(t), b_e(t), q_a(t), q_c(t)\}$ in real time based on observations of the current queue backlog $Q_a(t)$, $Q_c(t)$, $Q_v(t)$ and other system states such as electricity price $p_e(t)$, carbon allowance price $p_a(t)$, CERs price $p_c(t)$, available local renewable $r(t)$, SoC of virtual battery $b(t)$, maximum charging/discharging rate $b_{char}(t)/b_{dis}(t)$ and non-control demand $e(t)$ by solving the linear programming online problem (32). The online algorithm is summarized in Algorithm 1.

Algorithm 1 Online Joint Electricity and Carbon Purchase Decision-Making Algorithm

1: **Initialize** the queue backlog $Q_a(0) = 0$, $Q_c(0) = 0$ and $Q_v(0) = Q_0$.
2: **for** each time slot t **do**
3: Observe $\{Q_a(t), Q_c(t), Q_v(t), p_e(t), p_a(t), p_c(t), r(t), b(t), b_{char}(t)/b_{dis}(t)$ and $e(t)\}$.
4: Determine $\{r_e^*(t), r_b^*(t), g_e^*(t), g_b^*(t), b_e^*(t), q_a^*(t), q_c^*(t)\}$ by solving the online problem (32).
5: Implement the decision and update $b(t+1)$, $Q_a(t+1)$, $Q_c(t+1)$, $Q_v(t+1)$ following the dynamics (6d), (21), (22) , (27) respectively.
6: **end for**

C. Performance Analysis

Before summarizing the properties of the online algorithm, we first analyze the optimal solution of the online problem (32). In the following lemma, we prove that $Q_a(t)$ and $Q_c(t)$ have finite upper bounds $\bar{Q}_a$ and $\bar{Q}_c$ respectively. Based on that, we analyze the optimal charging and discharging decisions in the online problem (32).

Lemma 1: The optimal solution to the online problem (32) satisfies the following:

1) If $Q_v(t) \leq -V\bar{p}_e - \gamma_{max}\bar{Q}_a + (1-\alpha)b(t)$ holds, we have $b_e(t) = 0$ and $g_b(t) + r_b(t) = b_{char}(t)$.
2) If $Q_v(t) \geq \eta\gamma_{max}\bar{Q}_c + (1-\alpha)b(t)$ holds, we have $g_b(t) = 0$, $r_b(t) = 0$, and $b_e(t) = b_{dis}(t)$,

where $\bar{p}_e := \max_t p_e(t)$, $\bar{Q}_a$ and $\bar{Q}_c$ are the upper bounds of $Q_a(t)$ and $Q_c(t)$ respectively which will be given in the proof. See Appendix B for the proof.

Lemma 1 contributes more insights into Algorithm 1 as follows.

Theorem 1: If $r(t)$, $e(t)$, $p_e(t)$, $p_a(t)$, $p_c(t)$, $b_{char}(t)$, $b_{dis}(t)$, $\gamma(t)$, $\forall t$, are i.i.d., Algorithm 1 achieves a time-average electricity and carbon cost that exceeds the optimum Y^{OPT} by at most a constant B^*/V, i.e.,

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} \mathbb{E}\{c(t)\} \leq Y^{OPT} + B^*/V, \tag{33}$$

where

$$\begin{aligned} B^* := B &+ \frac{1}{2}(1-\alpha)^2 b_{max}^2 + \varepsilon(1-\alpha)b_{max} \\ &+ \max\{|\underline{Q_v}|, |\bar{Q_v}|\}[(1-\alpha)(b_{max}-b_{min}) + \varepsilon] \\ &+ (1-\alpha)^2 \max\{|b_{min}|, |b_{max}|\}^2, \end{aligned}$$

with

$$\underline{Q_v} := -V\bar{p}_e - b_{dis} - \varepsilon - (1-\alpha)(b_{max}-b_{min}) - \gamma_{max}\bar{Q}_a,$$

$$\bar{Q}_v := \delta - V\bar{p}_e - b_{dis} - \varepsilon - (1-\alpha)(b_{max}-b_{min}) - \gamma_{max}\bar{Q}_a$$

and $\delta := \min_t[b_{max}(t) - b_{min}(t)]$.

Moreover, given

$$Q_0 := -V\bar{p}_e - b_{dis} - \varepsilon - (1-\alpha)(b_{max}-b_{min}) - \gamma_{max}\bar{Q}_a,$$

if $0 < V \leq V_{max}$ holds with

$$\begin{aligned} V_{max} := \frac{1}{\bar{p}_e}[&\delta - b_{char} - b_{dis} - \gamma_{max}\bar{Q}_a - \eta\gamma_{max}\bar{Q}_c \\ &-2(1-\alpha)(b_{max}-b_{min}) - 2\varepsilon], \end{aligned} \tag{34}$$

it is guaranteed that

$$b_{min}(t) \leq b(t) \leq b_{max}(t), \quad \forall t,$$

holds.

The proof is given in Appendix C.

Theorem 1 implies a trade-off between optimality and conservativeness. An increased value of V reduces the optimality gap, thus enhancing the performance of the proposed online algorithm. However an excessively large V may lead to $Q_v(t)$,$Q_a(t)$ and $Q_c(t)$ that go beyond the feasible range and result in violating the SoC limits and carbon constraints. Conversely, a smaller V value enhances the stability of the queues but leads to a larger optimality gap. Thus, selecting a good value of V is, the key to the successful application of the proposed online algorithm.

Recall in Section III-B we show that the long-term carbon constraints (8) and (11) in the offline problem (17) can be satisfied if $Q_a(t)$ and $Q_c(t)$ are bounded. This condition has been proved in lemma 1. Therefore, we can conclude that the solution of Algorithm 1 satisfies all constraints of the offline problem (17).

IV. Numerical Results

In this section, we evaluate the proposed Lyapunov optimization-based online algorithm through simulation experiments. We first describe our experimental setup for the evaluation. Then we present the evaluation results and give a detailed analysis.

A. Experimental Setup

We consider a total time duration of 30 days, where each time t corresponds to 1 hour. For the electricity price, we leverage the historical day-ahead price from NY-ISO [40]. We approximate the hourly carbon allowance prices using daily carbon prices from the German carbon market in 2022. Similarly, we derive the CER prices from German market data. According to the Fudan Carbon Price Index [41], CER prices are generally lower than carbon allowance prices, although this varies over time, and we consider this characteristic in our CER approximation. We calculate the hourly carbon emission intensity trace for utility electricity using hourly data on electricity generation and CO_2 emissions from California ISO [42]. PV energy is considered as renewable energy

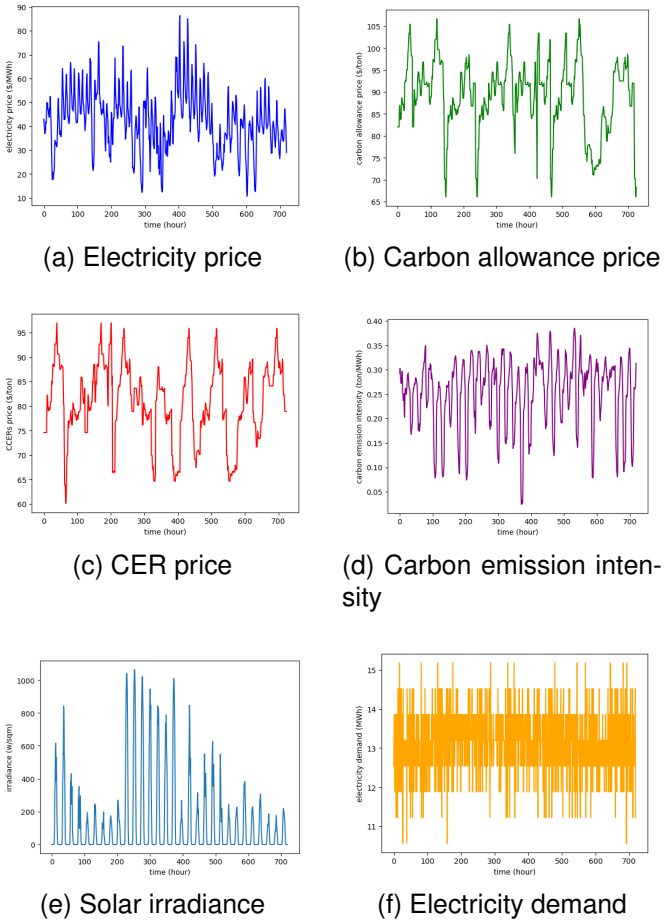

Fig. 2: Real-life traces of electricity price, carbon allowance price, CER price, carbon emission intensity, solar irradiance, and non-controllable electricity demand.

sources in this work. Hourly solar irradiance data of PV is taken from a photovoltaic power station in China. We adjust electricity demand data from a specific data center to reflect the proportions from long-term and spot markets, preserving the original demand trends and patterns. They are shown in Fig. 2 The maximum energy level $b_{max}(t)$ of the aggregated virtual battery varies between 5500 and 6000 kWh, while the minimum energy level $b_{min}(t)$ ranges from 1000 to 1500 kWh. The charging rate $b_{char}(t)$ and discharging rate $b_{dis}(t)$ can fluctuate arbitrarily between 400 kW and 500 kW. We assume that we know $\delta = 4012, b_{char} = 500, b_{dis} = 500$ in advance. Additionally, we assume the virtual battery's self-discharge factor $\alpha = 0.95$, the area of PV is 10000 m^2 and the allocated carbon allowance $A = 4000$.

B. Results

In this subsection, We describe three benchmark algorithms and compare them with the proposed online algorithm. Subsequently, we analyze the impact of the weight coefficient V and the parameter η on performance. Finally, we examine the feasibility of the algorithm in meeting the virtual battery's SoC constraints and the long-term carbon constraints.

Cost Performance: We compare the proposed online algorithm with three algorithms: (1) an online algorithm that ignores both the flexibility of data center loads and fluctuations in external prices, (2) an algorithm based on Lyapunov optimization that optimizes only electricity transactions without considering carbon trading optimization, (3) an offline optimal algorithm with perfect future information. As shown in Fig.3, our proposed algorithm achieves average cost that is close to optimal and much lower than the other two benchmarks even though we only consider current information. This comparison illustrates that by leveraging the inherent flexibility of data center loads, the algorithm optimizes trading decisions in response to real-time electricity and carbon product prices, carbon emissions, and local renewable generation, achieving results that are close to the optimal offline algorithm.

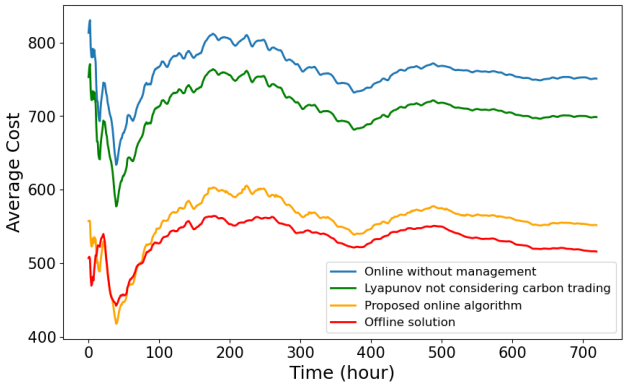

Fig. 3: Cost comparison across different algorithms.

TABLE I: COST COMPARISON AMONG ALGORITHMS AND THE PROPOSED ALGORITHM

Algorithm	**Total cost**	**Reduction**
Online without management	750.9	-
Lyapunov not considering carbon trading	699.5	6.85%
Proposed online algorithm	551.3	26.62%
Offline solution	515.5	31.34%

Table I summarizes the costs across different algorithms. Compared to the online algorithm without management, the proposed algorithm significantly reduces total cost from 750.9$ to 551.3$, achieving a substantial 26.62% reduction. In contrast, an algorithm based on Lyapunov optimization – focused solely on electricity transactions without carbon trading – yields a reduction of only 6.85%. The proposed algorithm's performance is the closest to that of offline optimization.

Additionally, we observe how the parameter V influences the cost reduction. It can be observed from Fig. 4 that the average cost decreases in V, which aligns with our analysis. Furthermore, we perform a sensitivity analysis on the parameter η, which represents a specified proportion of carbon emissions offset by CERs. As η increases, it indicates the potential for purchasing more CERs products for carbon offsetting. As shown in Fig. 5, with the rising η, the average cost for the data center decreases. This suggests that our algorithm opts to purchase more lower-priced CERs for carbon offsetting, thereby reducing costs.

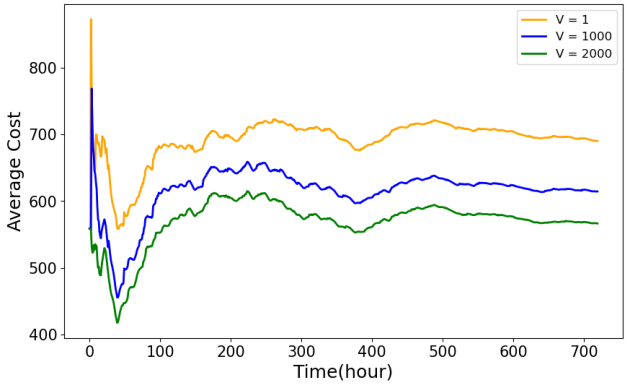

Fig. 4: Average cost with different V.

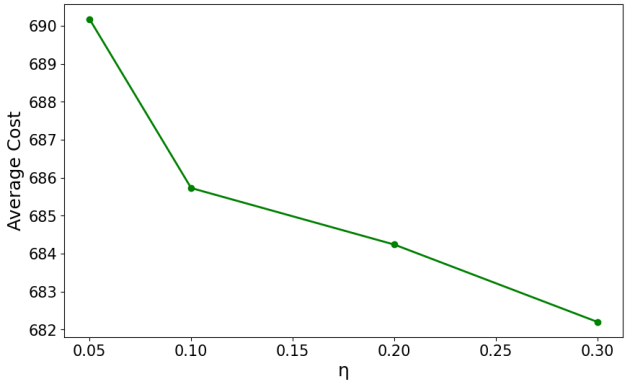

Fig. 5: Average cost with different η.

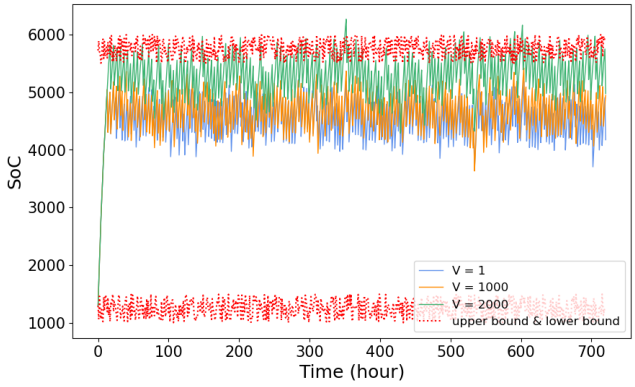

Fig. 6: Variation of SoC under different settings of V.

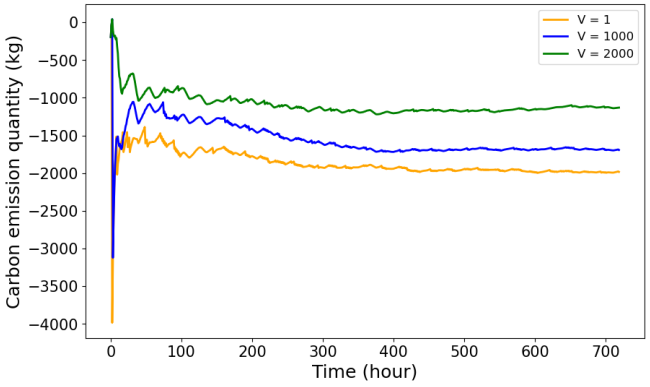

Fig. 7: Variation of net carbon emissions over time under different V.

Constraint satisfaction: Next, we check the feasibility of the algorithm by observing whether the virtual battery SoC always remains between its upper and lower bounds under our algorithm, and under what circumstances it may exceed these bounds. As we previously show in Theorem 1, if we set the weight parameter to V_{max}, the maximum value that can ensure feasibility, then the virtual battery SoC will remain within its bounds at all times. However, the calculation of V_{max} requires prior knowledge of some global parameters, such as δ, ϵ and b_{char}. In practice, lack of these parameters may lead to an inappropriate selection of V, resulting in the SoC of the virtual battery exceeding its bounds. As illustrated in Fig.6, when we set $V = 1$ and $V = 1000$, which means $V \leq V_{max}$, we observe that the SoC of the virtual battery remains within its bounds. However, when $V = 2000 > V_{max}$, it can be observed that the SoC of the virtual battery may exceed its limits. Therefore, by setting an appropriate V, we can achieve close-to-optimal cost reduction while strictly respecting all operational constraints. Moreover, as shown in Fig.7 , we also observe that the carbon constraint is consistently satisfied. This indicates that, under our algorithm, the data center continues to ensure carbon constraints are met while minimizing costs.

V. Conclusions

This work studies the problem of minimizing costs in joint electricity and carbon trading decisions for data center with massive flexible loads under uncertainties and carbon constraints. We propose a two-layer hierarchy: the lower layer aggregates the operation of all electrical equipment in the data center into a dynamic virtual battery, while the upper layer decides the procurement of electricity and carbon products to minimize the long-term time-average cost. Then an online algorithm based on Lyapunov optimization is further developed to optimize electricity and carbon product trading for data centers, which considers uncertainties in energy prices, carbon allowance and CER prices, grid carbon intensity, renewable energy availability, and virtual battery specifications. This algorithm balances the stability of virtual queues with immediate costs at each time t. As the parameter V increases, the data center can achieve a time-averaged cost close to the offline optimum. It also satisfies virtual battery constraints and ensures long-term carbon compliance during decarbonization. At last, simulation results validate the theoretical analysis and show the effectiveness of our algorithm in reducing the long-term electricity and carbon costs. The proposed online algorithm achieves a near-optimal offline cost compared to an algorithm focused solely on electricity transactions without carbon trading. Additionally, simulations show that as carbon emissions are offset by CERs, our algorithm increases the purchase of lower-priced CERs, further reducing costs.

References

[1] Data centres and data transmission networks. [Online] Available: https://www.iea.org/reports/data-centres-and-data-transmission-networks.

[2] N. Jones, "How to stop data centres from gobbling up the world's electricity," *Nature*, vol. 561, no. 7722, pp. 163–166, 2018.

[3] Cap and trade basics. [Online] Available: https://www.c2es.org/content/cap-and-trade-basics/.

[4] S. Huang, Z. Wen, J. Chen, and N. Cui, "Optimal technology investment under emission trading policy," *Journal of Systems Science and Systems Engineering*, vol. 29, no. 2, pp. 143–162, 2020.

[5] M. Yang and X. Chen, "Green technology investment strategies under cap-and-trade policy," *IEEE Transactions on Engineering Management*, vol. 71, pp. 3867–3880, 2022.

[6] A. Entezaminia, A. Gharbi, and M. Ouhimmou, "A joint production and carbon trading policy for unreliable manufacturing systems under cap-and-trade regulation," *Journal of Cleaner Production*, vol. 293, p. 125973, 2021.

[7] S. Du, Y. Zhu, Y. Zhu, and W. Tang, "Allocation policy considering firm's time-varying emission reduction in a cap-and-trade system," *Annals of Operations Research*, vol. 290, pp. 543–565, 2020.

[8] P. X. Gao, A. R. Curtis, B. Wong, and S. Keshav, "It's not easy being green," *ACM SIGCOMM Computer Communication Review*, vol. 42, no. 4, pp. 211–222, 2012.

[9] Y. Zhou, D. K. Yau, P. You, and P. Cheng, "Optimal-cost scheduling of electrical vehicle charging under uncertainty," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 4547–4554, 2017.

[10] Z. Wang, H. Zhang, X. Cao, E. Liu, H. Li, and J. Zhang, "Modeling and detection scheme for zero-dynamics attack on wind power system," *IEEE Transactions on Smart Grid*, vol. 15, no. 1, pp. 934–943, 2024.

[11] W. A. Hanafy, Q. Liang, N. Bashir, D. Irwin, and P. Shenoy, "Carbonscaler: Leveraging cloud workload elasticity for optimizing carbon-efficiency," *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, vol. 7, no. 3, pp. 1–28, 2023.

[12] L. Cupelli, T. Schütz, P. Jahangiri, M. Fuchs, A. Monti, and D. Müller, "Data center control strategy for participation in demand response programs," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 11, pp. 5087–5099, 2018.

[13] Z. Zhou, F. Liu, R. Zou, J. Liu, H. Xu, and H. Jin, "Carbon-aware online control of geo-distributed cloud services," *IEEE Transactions on Parallel and Distributed Systems*, vol. 27, no. 9, pp. 2506–2519, 2015.

[14] H. Dou, Y. Qi, W. Wei, and H. Song, "Carbon-aware electricity cost minimization for sustainable data centers," *IEEE Transactions on Sustainable Computing*, vol. 2, no. 2, pp. 211–223, 2017.

[15] W.-T. Lin, G. Chen, and H. Li, "Carbon-aware load balance control of data centers with renewable generations," *IEEE Transactions on Cloud Computing*, vol. 11, no. 2, pp. 1111–1121, 2022.

[16] S. Ren and Y. He, "Coca: Online distributed resource management for cost minimization and carbon neutrality in data centers," in *ACM Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis (SC)*, 2013, pp. 1–12.

[17] H. He, H. Shen, and D. Liang, "Cost-minimizing online algorithm for internet green data centers on multi-source energy," *Concurrency and Computation: Practice and Experience*, vol. 31, no. 21, pp. 5044–5056, 2019.

[18] H. Hao, A. Somani, J. Lian, and T. E. Carroll, "Generalized aggregation and coordination of residential loads in a smart community," in *IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 2015, pp. 67–72.

[19] F. L. Müller, J. Szabó, O. Sundström, and J. Lygeros, "Aggregation and disaggregation of energetic flexibility from distributed energy resources," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 1205–1214, 2017.

[20] K. Mukhi and A. Abate, "An exact characterisation of flexibility in populations of electric vehicles," in *IEEE Conference on Decision and Control (CDC)*, 2023, pp. 6582–6587.

[21] T. Li, B. Sun, Y. Chen, Z. Ye, S. H. Low, and A. Wierman, "Learning-based predictive control via real-time aggregate flexibility," *IEEE Transactions on Smart Grid*, vol. 12, no. 6, pp. 4897–4913, 2021.

[22] M. Khatibi, S. Rahnama, P. Vogler-Finck, J. Dimon Bendtsen, and A. Afshari, "Towards designing an aggregator to activate the energy flexibility of multi-zone buildings using a hierarchical model-based scheme," *Applied Energy*, vol. 333, pp. 120 562–120 573, 2023.

[23] D. Madjidian, M. Roozbehani, and M. A. Dahleh, "Energy storage from aggregate deferrable demand: Fundamental trade-offs and scheduling policies," *IEEE Transactions on Power Systems*, vol. 33, no. 4, pp. 3573–3586, 2017.

[24] L. Zhao, W. Zhang, H. Hao, and K. Kalsi, "A geometric approach to aggregate flexibility modeling of thermostatically controlled loads," *IEEE Transactions on Power Systems*, vol. 32, no. 6, pp. 4721–4731, 2017.

[25] F. Al Taha, T. Vincent, and E. Bitar, "An efficient method for quantifying the aggregate flexibility of plug-in electric vehicle populations," *IEEE Transactions on Smart Grid*, 2024.

[26] E. Biyik and A. Kahraman, "A predictive control strategy for optimal management of peak load, thermal comfort, energy storage and renewables in multi-zone buildings," *Journal of building engineering*, vol. 25, pp. 100 826–100 837, 2019.

[27] B. Li, B. Wu, Y. Peng, and W. Cai, "Tube-based robust model predictive control of multi-zone demand-controlled ventilation systems for energy saving and indoor air quality," *Applied Energy*, vol. 307, pp. 118 297–118 308, 2022.

[28] M. Neely, *Stochastic network optimization with application to communication and queueing systems*. Springer Nature, 2022.

[29] T. Li and M. Dong, "Online control for energy storage management with renewable energy integration," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2013, pp. 5248–5252.

[30] T. Li and M. Dong, "Real-time energy storage management with renewable integration: Finite-time horizon approach," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 12, pp. 2524–2539, 2015.

[31] T. Li and M. Dong, "Residential energy storage management with bidirectional energy control," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 3596–3611, 2018.

[32] L. Yu, T. Jiang, and Y. Zou, "Online energy management for a sustainable smart home with an hvac load and random occupancy," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 1646–1659, 2019.

[33] G. Zhang, S. Zhang, W. Zhang, Z. Shen, and L. Wang, "Distributed energy management for multiple data centers with renewable resources and energy storages," *IEEE Transactions on Cloud Computing*, vol. 10, no. 4, pp. 2469–2480, 2020.

[34] L. Yu, T. Jiang, and Y. Zou, "Distributed real-time energy management in data center microgrids," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 3748–3762, 2016.

[35] T. Li and M. Dong, "Real-time residential-side joint energy storage management and load scheduling with renewable integration," *IEEE Transactions on Smart Grid*, vol. 9, no. 1, pp. 283–298, 2016.

[36] Y. Zhou, D. K. Yau, P. You, and P. Cheng, "Optimal-cost scheduling of electrical vehicle charging under uncertainty," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 4547–4554, 2017.

[37] Q. Fang, J. Wang, Q. Gong, and M. Song, "Thermal-aware energy management of an hpc data center via two-time-scale control," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 5, pp. 2260–2269, 2017.

[38] A. Y. Lo and R. Cong, "After cdm: Domestic carbon offsetting in china," *Journal of Cleaner Production*, vol. 141, pp. 1391–1399, 2017.

[39] L. Li, F. Ye, Y. Li, and C.-T. Chang, "How will the chinese certified emission reduction scheme save cost for the national carbon trading system?" *Journal of environmental management*, vol. 244, pp. 99–109, 2019.

[40] Energy market and operational data. [Online] Available: https://www.nyiso.com/energy-market-operational-data.

[41] Carbon price index of fudan. [Online] Available: https://rcsd.fudan.edu.cn/fdtjzs/zsjj.htm.

[42] California iso emissions. [Online] Available: https://www.caiso.com/todays-outlook/emissions.

[43] L. Huang and M. J. Neely, "Utility optimal scheduling in processing networks," *Performance Evaluation*, vol. 68, no. 11, pp. 1002–1021, 2011.

VI. Appendix

A. Proof of Lemma 1

The one time slot Lyapunov drift expression is as follows:

$$\begin{aligned} L(\Theta(t+1)) - L(\Theta(t)) =& \frac{1}{2}[Q_v^2(t+1) - Q_v^2(t) \\ &+ Q_a^2(t+1) - Q_a^2(t) \\ &+ Q_c^2(t+1) - Q_c^2(t)]. \end{aligned} \tag{35}$$

Note that taking square of both sides of (21), (22) and (27), respectively, yields

$$\begin{aligned} Q_a^2(t+1) \leq & Q_a^2(t) + \left[\gamma(t)\big(g_e(t) + g_b(t)\big) - q_a(t) - q_c(t) - A\right]^2 \\ & + 2Q_a(t)\left[\gamma(t)\big(g_e(t) + g_b(t)\big) - q_a(t) - q_c(t) - A\right], \end{aligned}$$

$$Q_c^2(t+1) \le Q_c^2(t) + \left[q_c(t) - \eta\gamma(t)\big(g_b(t)+g_e(t)\big)\right]^2 + 2Q_c(t)\left[q_c(t) - \eta\gamma(t)\big(g_b(t)+g_e(t)\big)\right],$$

$$\begin{aligned} Q_v^2(t+1) &= Q_v^2(t) + [g_b(t)+r_b(t)-b_e(t)-(1-\alpha)b(t)-\varepsilon(t)]^2 \\ &\quad + 2Q_v(t)\left[g_b(t)+r_b(t)-b_e(t)-(1-\alpha)b(t)-\varepsilon(t)\right] \\ &\le Q_v^2(t) + [g_b(t)+r_b(t)-b_e(t)-(1-\alpha)b(t)-\varepsilon(t)]^2 \\ &\quad + 2Q_v(t)\left[g_b(t)+r_b(t)-b_e(t)-(1-\alpha)b(t)-\varepsilon(t)\right]. \end{aligned}$$

Due to (6c), (6e), (9), (10) and (15), we have

$$\begin{aligned} &[g_b(t)+r_b(t)-b_e(t)-(1-\alpha)b(t)-\varepsilon(t)]^2 \\ =&[g_b(t)+r_b(t)-b_e(t)]^2 + [(1-\alpha)b(t)+\varepsilon(t)]^2 \\ &-2[g_b(t)+r_b(t)-b_e(t)][(1-\alpha)b(t)+\varepsilon(t)]] \\ \le& max\{b_{char}^2(t), b_{dis}^2(t)\} + (1-\alpha)^2 b^2(t) + \varepsilon^2(t) \\ &+2\varepsilon(t)(1-\alpha)b(t) - 2(1-\alpha)b(t)[g_b(t)+r_b(t)-b_e(t)] \\ &-2\varepsilon(t)[g_b(t)+r_b(t)-b_e(t)] \\ \le& max\{b_{char}^2, b_{dis}^2\} + (1-\alpha)^2 b^2(t) + \varepsilon^2 + 2\varepsilon(1-\alpha)b(t) \\ &-2(1-\alpha)b(t)[g_b(t)+r_b(t)-b_e(t)] + 2\varepsilon\max\{b_{dis}, b_{char}\}. \end{aligned}$$

Similarly, we can also obtain

$$\begin{aligned} &[\gamma(t)[g_e(t)+g_b(t)] - (q_a(t)+q_c(t)+A)]^2 \\ =&(\gamma(t)[g_e(t)+g_b(t)])^2 + (q_a(t)+q_c(t)+A)^2 \\ &-2\gamma(t)[g_e(t)+g_b(t)][q_a(t)+q_c(t)+A] \\ \le&(\gamma(t)[g_e(t)+g_b(t)])^2 + (q_a(t)+q_c(t)+A)^2 \\ \le&\gamma_{max}^2 g_{max}^2 + (\bar{q}_a+\bar{q}_c+A)^2, \end{aligned}$$

$$\begin{aligned} &[q_c(t) - \eta\gamma(t)(g_b(t)+g_e(t))]^2 \\ =&q_c^2(t) + [\eta\gamma(t)(g_b(t)+g_e(t))]^2 - 2q_c(t)[\eta\gamma(t)(g_b(t)+g_e(t))] \\ \le&q_c^2(t) + [\eta\gamma(t)(g_b(t)+g_e(t))]^2 \\ \le&\bar{q}_c^2 + \eta^2\gamma_{max}^2 g_{max}^2. \end{aligned}$$

Therefore taking the conditional expectation of (35) and adding $V\mathbb{E}\{c(t)|\Theta(t)\}$ results in the following bound on the one-step Lyapunov drift plus cost

$$\begin{aligned} &\Delta(\Theta(t)) + V\mathbb{E}\{c(t)|\Theta(t)\} \\ \le& B + \frac{1}{2}(1-\alpha)^2 b^2(t) + \varepsilon(1-\alpha)b(t) \\ &V\mathbb{E}\{p_e(t)[g_b(t)+g_e(t)] + p_a(t)q_a(t) + p_c(t)q_c(t)|\Theta(t)\} \\ &+ Q_v(t)\mathbb{E}\{g_b(t)+r_b(t)-b_e(t)-(1-\alpha)b(t)-\varepsilon(t)|\Theta(t)\} \\ &- (1-\alpha)\mathbb{E}\{b(t)[g_b(t)+r_b(t)-b_e(t)]|\Theta(t)\} \\ &+ Q_a(t)\mathbb{E}[\gamma(t)\{[g_e(t)+g_b(t)] - q_a(t) - q_c(t) - A|\Theta(t)\} \\ &+ Q_c(t)\mathbb{E}\{q_c(t) - \eta\gamma(t)[g_b(t)+g_e(t)]|\Theta(t)\}, \end{aligned} \tag{36}$$

where $B := \frac{1}{2}([max\{b_{char}^2, b_{dis}^2\} + \varepsilon^2 + 2\varepsilon\max\{b_{dis}, b_{char}\}] + [\gamma_{max}^2 g_{max}^2 + (\bar{q}_a+\bar{q}_c+A)^2] + [\bar{q}_c^2 + \eta^2\gamma_{max}^2 g_{max}^2])$.

B. Proof of Lemma 1

We first prove the existence of the upper bounds of $Q_a(t)$ and $Q_c(t)$. Define $\bar{p_a} := \max_t p_a(t)$ and $\underline{p_c} := \min_t p_c(t)$. We impose an assumption that $\bar{p}_a \ge \underline{p_c}$ as a rational premise. If this condition were not met, the procurement of CERs would become economically infeasible given the excessive pricing.

For the proof concerning the queue $Q_a(t)$, we assume that the carbon emissions from purchasing electricity from the grid are significant enough so that $\gamma_{\max} g_{\max} > A$. This assumption is reasonable because if the initial carbon allowance A were always sufficient, there would be no need for carbon emission control in data centers.Additionally, since excess carbon emissions are mainly offset by purchasing $q_a(t)$, we assume $\overline{q}_a > \gamma_{\max} g_{\max}$. Now we use mathematical induction method to prove that $Q_a(t) \le V\bar{p}_a + \gamma_{\max} g_{\max}$ for all time t. For $t = 0$, $Q_a(0) = 0 \le V\bar{p}_a + \gamma_{\max} g_{\max}$ holds. Suppose the above inequality holds for time t. If $Q_a(t) \le V\bar{p}_a$, then according to equation (21), we have $Q_a(t+1) \le V\bar{p}_a + \gamma_{\max} g_{\max}$. If $V\bar{p}_a < Q_a(t) \le V\bar{p}_a + \gamma_{max} g_{max}$, we can get $Vp_a(t) - Q_a(t) < 0$. To minimize the objective of **P3**, we strive to maximize the value of $q_a(t)$. So we have $Q_a(t+1) < Q_a(t) \le V\bar{p}_a + \gamma_{max} g_{max}$. Therefore, the upper bound $\bar{Q}_a := V\bar{p}_a + \gamma_{max} g_{max}$.

With respect to the queue $Q_c(t)$, we will prove that $Q_c(t) <= \bar{Q}_c := \bar{Q}_a - V\underline{p_c} + \bar{q}_c$. For $t = 0$, $Q_c(0) = 0 \le \bar{Q}_a - V\underline{p_c} + \bar{q}_c$ holds. Suppose the above inequality holds for time t. If $Q_c(t) \le \bar{Q}_a - V\underline{p_c}$, based on (10), we have $Q_c(t+1) \le \bar{Q}_a - V\underline{p_c} + \bar{q}_c$. If $\bar{Q}_a - V\underline{p_c} < Q_c(t) \le \bar{Q}_a - V\underline{p_c} + \bar{q}_c$, to minimize the objective of **P3**, we need to make the value of $q_c(t)$ as small as possible. Thus we can obtain $Q_c(t+1) \le Q_c(t) \le \bar{Q}_a - V\underline{p_c} + \bar{q}_c$.

Through the aforementioned analysis, we have the conclusion that $\bar{Q}_a := V\bar{p}_a + \gamma_{max} g_{max}$ and $\bar{Q}_c := \bar{Q}_a - V\underline{p_c} + \bar{q}_c$.

Next, we show the optimal charging and discharging solutions in some cases below:

1. If $Q_v(t) \le -V\bar{p}_e - \gamma_{max}\bar{Q}_a + (1-\alpha)b(t)$: We have $Q_v(t) - (1-\alpha)b(t) + Vp_e(t) + \gamma(t)Q_a(t) - \eta\gamma(t)Q_c(t) \le 0$ using $Q_c(t) \ge 0$. To minimize the objective of **P3**, we strive to maximize the values of $g_b(t)$, $r_b(t)$ and while minimizing $b_e(t)$ as much as possible. Thus in this case we can obtain:

$$\begin{cases} r_b(t) = \min\{r(t), b_{char}(t)\} \\ g_b(t) = b_{char}(t) - r_b(t) \\ b_e(t) = 0. \end{cases}$$

2. If $Q_v(t) \ge \eta\gamma_{max}\bar{Q}_c + (1-\alpha)b(t)$: We have $Q_v(t) + Vp_e(t) + \gamma(t)Q_a(t) - \eta\gamma(t)Q_c(t) - (1-\alpha)b(t) \ge 0$ using $p_e(t) \ge 0$ and $Q_a(t) \ge 0$. To minimize the objective of **P3**, we need to make the values of $g_b(t)$, $r_b(t)$ as small as possible, while maximizing $b_e(t)$ as much as possible. Thus we can obtain the following:

$$\begin{cases} r_b(t) = 0 \\ g_b(t) = 0 \\ b_e(t) = b_{dis}(t), \end{cases}$$

where we assume $e(t) \ge b_{dis}(t)$ without loss of generality.

C. Proof of Theorem 1

Assume that for any time t the following conditions hold: $b_{char}(t) \ge (1-\alpha)b(t) + \varepsilon$ and $b_{dis}(t) \ge \varepsilon - (1-\alpha)b(t)$. This assumption ensures that the charging or discharging rates at each time slot is sufficient to cover the variations in SoC limits and self-dissipation of the virtual battery. It's reasonable because the virtual battery's power rating must be sufficient to handle both self-dissipation and capacity variations. First,

we prove the feasibility of virtual battery operation when Q_0 and V are properly chosen. We use mathematical induction method to prove that for all time t

$$0 \leq b(t) - b_{min}(t) \leq \delta. \tag{37}$$

Note that $Q_v(t) = b(t) - b_{min}(t) + Q_0$ where $Q_0 := -V\bar{p}_e - b_{dis} - \varepsilon - (1-\alpha)(b_{max} - b_{min}) - \gamma_{max}\bar{Q}_a$. So (37) is equivalent to the following

$$\begin{aligned} -V\bar{p}_e - b_{dis} - \varepsilon - (1-\alpha)(b_{max} - b_{min}) - \gamma_{max}\bar{Q}_a \leq Q_v(t) \\ \leq \delta - V\bar{p}_e - b_{dis} - \varepsilon - (1-\alpha)(b_{max} - b_{min}) - \gamma_{max}\bar{Q}_a. \end{aligned} \tag{38}$$

For $t = 0$, $b(0)$ is the given initial SoC and satisfies $b_{min}(0) \leq b(t) \leq b_{max}(0)$. Suppose that (37) and (38) hold at time t. We prove through induction that they also holds at time $t+1$.

Case 1. Suppose $-V\bar{p}_e - b_{dis} - \varepsilon - (1-\alpha)(b_{max} - b_{min}) - \gamma_{max}\bar{Q}_a \leq Q_v(t) \leq -V\bar{p}_e - \gamma_{max}\bar{Q}_a + (1-\alpha)b_{min}$. By **Lemma 2**, it can be inferred that $b_e(t) = 0$, $g_b(t) + r_b(t) = b_{char}(t)$, then

$$\begin{aligned} Q_v(t+1) =& Q_v(t) + b_{char}(t) - (1-\alpha)b(t) - \varepsilon(t) \\ \geq& Q_v(t) \\ \geq& -V\bar{p}_e - b_{dis} - \varepsilon - (1-\alpha)(b_{max} - b_{min}) - \gamma_{max}\bar{Q}_a, \end{aligned}$$

$$\begin{aligned} Q_v(t+1) \leq& Q_v(t) + b_{char} - (1-\alpha)b_{min} + \varepsilon \\ \leq& -V\bar{p}_e - \gamma_{max}\bar{Q}_a + b_{char} + \varepsilon \\ \leq& \delta - V\bar{p}_e - b_{dis} - \varepsilon - (1-\alpha)(b_{max} - b_{min}) - \gamma_{max}\bar{Q}_a. \end{aligned}$$

In the proof of the upper bound, we utilize $0 < V \leq V_{max}$ and $b(t) \geq b_{min}(t) \geq b_{min}$.

Case 2. Suppose $-V\bar{p}_e - \gamma_{max}\bar{Q}_a + (1-\alpha)b_{min} < Q_v(t) < \eta\gamma_{max}\bar{Q}_c + (1-\alpha)b_{max}$, then

$$\begin{aligned} Q_v(t+1) \geq& Q_v(t) - b_{dis}(t) - (1-\alpha)b(t) - \varepsilon(t) \\ \geq& Q_v(t) - b_{dis} - (1-\alpha)b_{max} - \varepsilon \\ >& -V\bar{p}_e - \gamma_{max}\bar{Q}_a - (1-\alpha)(b_{max} - b_{min}) - b_{dis} - \varepsilon, \end{aligned}$$

where we utilize $b(t) \leq b_{min}(t) + \delta \leq b_{max}(t) \leq b_{max}$ in the second inequality.

$$\begin{aligned} Q_v(t+1) \leq& Q_v(t) + b_{char}(t) - (1-\alpha)b(t) - \varepsilon(t) \\ \leq& Q_v(t) + b_{char} - (1-\alpha)b_{min} + \varepsilon \\ <& \eta\gamma_{max}\bar{Q}_c + (1-\alpha)(b_{max} - b_{min}) + b_{char} + \varepsilon \\ \leq& \delta - V\bar{p}_e - b_{dis} - \varepsilon - (1-\alpha)(b_{max} - b_{min}) - \gamma_{max}\bar{Q}_a, \end{aligned}$$

where we utilize $0 < V \leq V_{max}$ and $b(t) \geq b_{min}(t) \geq b_{min}$.

Case 3. Suppose $\eta\gamma_{max}\bar{Q}_c + (1-\alpha)b_{max} \leq Q_v(t) \leq \delta - V\bar{p}_e - b_{dis} - \varepsilon - (1-\alpha)(b_{max} - b_{min}) - \gamma_{max}\bar{Q}_a$. By **Lemma 2** we can infer $g_b(t) = 0$, $r_b(t) = 0$ and $b_e(t) = b_{dis}(t)$, then

$$\begin{aligned} & Q_v(t+1) \\ =& Q_v(t) - b_e(t) - (1-\alpha)b(t) - \varepsilon(t) \\ \geq& Q_v(t) - b_{dis} - (1-\alpha)b_{max} - \varepsilon \\ \geq& \eta\gamma_{max}\bar{Q}_c - b_{dis} - \varepsilon \\ \geq& -V\bar{p}_e - b_{dis} - \varepsilon - (1-\alpha)(b_{max} - b_{min}) - \gamma_{max}\bar{Q}_a, \end{aligned}$$

$$\begin{aligned} & Q_v(t+1) \\ =& Q_v(t) - b_e(t) - (1-\alpha)b(t) - \varepsilon(t) \\ \leq& Q_v(t) - b_{dis}(t) - (1-\alpha)b(t) + \varepsilon \\ \leq& Q_v(t) \\ \leq& \delta - V\bar{p}_e - b_{dis} - \varepsilon - (1-\alpha)(b_{max} - b_{min}) - \gamma_{max}\bar{Q}_a. \end{aligned}$$

In the proof of the lower bound, we utilize $0 < V \leq V_{max}$ and $b(t) \leq b_{max}(t) \leq b_{max}$.

Through the aforementioned induction, we have proved that (38) and its equivalent conclusion (37) hold at all time t. Using the definition of δ, (37) further indicates that $b_{min}(t) \leq b(t) \leq b_{max}(t)$ at all time t.

Next we prove (33). Before proving this part, we need to use the following lemma.

Lemma 2: If $r(t)$, $e(t)$, $p_e(t)$, $p_a(t)$, $p_c(t)$, $b_{char}(t)$, $b_{dis}(t)$, $\gamma(t)$ $\forall t$, are i.i.d over time t, then there exists a stationary, randomized policy that satisfies the constraints (6a),(6b),(6c), (6d),(7), (8), (9), (10),(11),(12),(13),(14),(15),(19) and provides the following guarantees:

$$(1-\alpha)b_{min} \leq \mathbb{E}\{g_b^*(t) + r_b^*(t) - b_e^*(t)\} \leq (1-\alpha)b_{max}, \tag{39}$$

$$\mathbb{E}\{(q_a^*(t) + q_c^*(t)) + A\} \geq \mathbb{E}\{\gamma(t)(g_b^*(t) + g_e^*(t))\}, \tag{40}$$

$$\mathbb{E}\{\eta\gamma(t)(g_b^*(t) + g_e^*(t))\} \geq \mathbb{E}\{q_c^*(t)\}, \tag{41}$$

$$\mathbb{E}\{p_e(t)[g_b^*(t) + g_e^*(t)] + p_a(t)q_a^*(t) + p_c(t)q_c^*(t)\} = Y^{REL}. \tag{42}$$

where the expectations above are with respect to the stationary distributions of $\{r(t),\ e(t),\ p_e(t),\ p_a(t),\ p_c(t),\ b_{char}(t),\ b_{dis}(t),\ \gamma(t)\}$ and the randomized control decisions.

This result has been proven in [43]. It is omitted here for brevity.

Our online algorithm is designed to minimize the upper bound of drift plus penalty indicated in (31) over all possible feasible control decisions. The feasible control policy includes the optimal, stationary, randomized policy given in **Lemma 3**. Therefore, we can obtain the following inequality:

$$\begin{aligned} & \Delta(\Theta(t)) + V\mathbb{E}\{c(t)|\Theta(t)\} \\ \leq& B + \frac{1}{2}(1-\alpha)b^2(t) + \varepsilon(1-\alpha)b(t) \\ & + V\mathbb{E}\{p_e(t)[g_b(t) + g_e(t)] + p_a(t)q_a(t) + p_c(t)q_c(t)|\Theta(t)\} \\ & + Q_v(t)\mathbb{E}\{g_b(t) + r_b(t) - b_e(t) - (1-\alpha)b(t) - \varepsilon(t)|\Theta(t)\} \\ & - (1-\alpha)\mathbb{E}\{b(t)[g_b(t) + r_b(t) - b_e(t)]|\Theta(t)\} \\ & + Q_a(t)\mathbb{E}\{\gamma(t)\{[g_e(t) + g_b(t)] - q_a(t) - q_c(t) - A|\Theta(t)\} \\ & + Q_c(t)\mathbb{E}\{q_c(t) - \eta\gamma(t)[g_b(t) + g_e(t)]|\Theta(t)\} \\ \leq& B + \frac{1}{2}(1-\alpha)b^2(t) + \varepsilon(1-\alpha)b(t) \\ & + V\mathbb{E}\{p_e(t)[g_b^*(t) + g_e^*(t)] + p_a(t)q_a^*(t) + p_c(t)q_c^*(t)\} \\ & + Q_v(t)\mathbb{E}\{g_b^*(t) + r_b^*(t) - b_e^*(t) - (1-\alpha)b(t) - \varepsilon(t)\} \\ & - (1-\alpha)\mathbb{E}\{b(t)[g_b^*(t) + r_b^*(t) - b_e^*(t)]\} \\ & + Q_a(t)\mathbb{E}[\gamma(t)\{[g_e^*(t) + g_b^*(t)] - q_a^*(t) - q_c^*(t) - A\} \\ & + Q_c(t)\mathbb{E}\{q_c^*(t) - \eta\gamma(t)[g_b^*(t) + g_e^*(t)]\} \\ \leq& B + \frac{1}{2}(1-\alpha)b_{max}^2 + \varepsilon(1-\alpha)b_{max} + VY^{REL} \\ & + \max\{|\underline{Q_v}|, |\bar{Q}_v|\}[(1-\alpha)(b_{max} - b_{min}) + \varepsilon] \\ & + (1-\alpha)^2\max\{|b_{min}|, |b_{max}|\}^2 \\ \leq& B^* + VY^{OPT}, \end{aligned} \tag{43}$$

where $\underline{Q_v} := -V\bar{p}_e - b_{dis} - \varepsilon - (1-\alpha)(b_{max} - b_{min}) - \gamma_{max}\bar{Q}_a$, $\bar{Q}_v := \overline{\delta} - V\bar{p}_e - b_{dis} - \varepsilon - (1-\alpha)(b_{max} - b_{min}) - \gamma_{max}\bar{Q}_a$. $\underline{Q_v}$

and $\bar{Q}_v$ are the lower and upper bound of $Q_v(t)$ respectively, as shown in (38). The constant

$$\begin{aligned} B^* := B + \frac{1}{2}(1-\alpha)^2 b_{max}^2 + \varepsilon(1-\alpha)b_{max} \\ + \max\{|\underline{Q_v}|, |\bar{Q}_v|\}[(1-\alpha)(b_{max}-b_{min})+\varepsilon] \\ + (1-\alpha)^2 \max\{|b_{min}|, |b_{max}|\}^2. \end{aligned}$$

The third inequality utilizes the properties of the policy given in **Lemma 3**. The last inequality utilizes $Y^{REL} \leq Y^{OPT}$.

Taking expectation over on both sides of (43) and summing over $t \in \{0, 1, 2, ..., T-1\}$, we can get

$$\begin{aligned} &\mathbb{E}\{L(\Theta(T)\} - \mathbb{E}\{L(\Theta(0)\} + V\sum_{t=0}^{T-1}\mathbb{E}\{c(t)\} \\ &\leq TB^* + TVY^{OPT}, \end{aligned} \tag{44}$$

where we have applied Law of Total Expectation. Rearranging the terms in the above equation and dividing both sides by $V \cdot T$. When $T \to \infty$, since $L(\Theta(T))$ and $L((\Theta(0))$ are finite, we can get:

$$\lim \frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\{c(t)\} \leq Y^{OPT} + \frac{B^*}{V}. \tag{45}$$