Event-Based EV Charging Scheduling in a Microgrid of Buildings

Qilong Huang, Member, IEEE, Li Yang, Member, IEEE, Chen Hou, Member, IEEE, Zhiyong Zeng, Member, IEEE, and Yaowen Qi, Member, IEEE

Abstract—With the popularization of electric vehicles (EVs), EV charging demand is becoming an important load in the building. Considering the mobility of EVs from building to building and their uncertain charging demand, it is of great practical interest to control the EV charging process in a microgrid of buildings to optimize the total operation cost while ensuring transmission security between the microgrid and the main grid. We consider this important problem in this article and make the following contributions. First, we formulate this problem as a Markov decision process (MDP) to capture the uncertain supply and EV charging demand in the microgrid of buildings. Besides reducing the total operation cost of buildings, the model also considers the power exchange limitation to ensure transmission security. Second, this model is reformulated within an event-based optimization (EBO) framework to mitigate the impact of large state and action space. A randomized parametric event-based control policy is proposed for the microgrid controller to determine the charge ratio for each building. Then, a selecting-to-charging principle is implemented for each building controller to determine the charge control of each EV in the building. Third, a constrained gradient-based policy optimization method is developed to find the optimal event-based control policy. Furthermore, an adjustment mechanism is also introduced to ensure transmission security during optimization. Numerical experiments are conducted to analyze the structure and the performance of the event-based control policy for EV charging in a microgrid of three buildings.

Index Terms—Building energy management, discrete event dynamic system, electric vehicle (EV), event-based optimization (EBO), Markov decision process (MDP).

I. INTRODUCTION

Due to reduced emission and dependency on fossil fuel, electric vehicles (EVs) have attracted more and more attention in recent years. To achieve carbon peak and neutrality targets, many countries have made great efforts to encourage EV popularization. For example, nearly 10% of global car sales were electric in 2021 which was four times compared to 2019. The total number of EVs was about 16.5 million which was triple compared to 2018 [1]. Although the EV popularization helps to mitigate the fossil-fuel crisis and environmental pollution, it poses a new challenge to the operation of microgrid if charging control is not available for increasing EVs [2].

As the building is the primary infrastructure that interacts with EVs, the impact of EV charging on the building is twofold. On the one hand, the charging profile of EVs will affect the energy operation of the building. The building energy operator needs to procure extra power to meet the demand for EV charging and achieve load balance. If inappropriate charging control policy implemented, the increasing operation cost of the building will ultimately be paid by the drivers. On the other hand, if only considering the charging control of EVs parked in the corresponding building, the charging actions of multibuildings may be homogeneous. This brings a new load peak to the grid, which can cause the aggregated load to exceed the contractual capacity. When this happens, it may congest the distribution feeders and transformers which results in voltage fluctuations in the microgrid [3].

In this context, it is of great practical interest to implement EV charging scheduling within a microgrid of buildings for economic operation and transmission security of multibuildings. This problem is nontrivial to solve due to the following difficulties.

First, there are uncertainties in the supply and demand side of the buildings. The uncertainty in the supply side of the buildings results from the uncertain generation of distributed renewable energy (DRE) [4]. The uncertainty in the demand side of the buildings originates from the uncertain charging demand of EVs. The arrival time, the parking time, and the required charging energy are all uncertain before the EV begins to park and charge in the building [5]. Second, the large number of EVs in the buildings leads to the large state space and action space. Currently, there are usually dozens to hundreds of charging piles in the buildings [6]. If all these charging piles are occupied, the charging states and actions that increase exponentially with the number of EVs will be high dimension, making it difficult to find an optimal charging control policy for this problem. Third, there is a limit to the aggregated load capacity in the microgrid of buildings. As aforementioned, a new load peak may appear if all the buildings implement a homogeneous charging control policy. Therefore, the power exchange of each building should also...
be investigated during scheduling to prevent an increase in peak load. This further increases the difficulty of solving this problem.

Based on the discussions above, we study the EV charging scheduling problem in a microgrid of buildings in this article. Compared with the published literature, the main contributions of this article are as follows.

1) The EV charging scheduling in a microgrid of buildings is formulated as a Markov decision process (MDP) to capture the uncertain DRE and EV charging demand in the microgrid. Besides reducing the total operation cost of the microgrid, the EV location transitions among buildings and the power exchange limitation of buildings are also incorporated into the model to depict the coupled charging relationship and ensure transmission security, respectively.

2) To mitigate the cures of dimensionality caused by the large number of EVs in the microgrid, the proposed model is reformulated within an event-based optimization (EBO) framework. A randomized parameteric event-based control policy is proposed for the microgrid controller to determine the charge ratio for each building. Then, a selecting-to-charging principle is implemented for each building controller to determine the charge control of each EV in the building.

3) A constrained gradient-based policy optimization method is developed to find the optimal event-based control policy. Furthermore, an adjustment mechanism is also introduced to ensure transmission security during optimization. Numerical experiments are conducted to analyze the structure and the performance of the event-based control policy for EV charging in a microgrid of three buildings.

The rest of this article is organized as follows. We briefly review the related literature in Section II, formulate the problem in Section III, present the solution methodology in Section IV, discuss the numerical results in Section V, and briefly conclude in Section VI.

II. LITERATURE REVIEW

In recent years, EV charging attracts much attention due to the popularization of EVs and their impact on the grid [7], [8]. Based on our survey, the control method and application scenarios for EV charging are the most active research areas. As this article mainly focuses on the EV charging control in the scenario of a microgrid of buildings, the detailed literature review for these areas is analyzed and introduced below.

The existing control approaches for EV charging can be mainly divided into two categories: the determined control approach and the stochastic control approach. In the determined charging control approach, the EV charging process is usually formulated as a determined programming model. In other words, it assumes the future EV charging demand is known a priori, such as arrival time, parking duration and required charging energy. Therefore, many traditional methods can be applied to search for an optimized charging control policy, such as linear/quadratic programming method [9], mixed-integer programming [10], heuristic approach [11], model predictive control (MPC) [12], etc. These methods enjoy high optimization efficiency due to the assumption of perfect information of uncertainty. However, this assumption is hard to obtain and these uncertainties cannot be underestimated in practice. For example, if the charging control policy is derived by assuming perfect information of EV parking events, this policy may be suboptimal or even infeasible when these EVs are absent due to the prediction error. Therefore, many researchers study the stochastic control method, such as scenario-based optimization [13], robust optimization [14], reinforcement learning [15], simulation-based policy improvement [16], etc. The first two methods usually try to convert the stochastic charging control problem into a determined control problem. These two methods should be carried out periodically and there is no experience accumulation among the control policies obtained at each time. The latter two methods usually formulate the problem as a MDP and can be considered as a state-based method. Due to the well-known curse of dimensionality, these methods suffer the large state space and action space considering the large number of EVs in a microgrid of buildings.

Most existing works mainly consider the EV charging control in the scenario of the parking lot or the charging station [9], [10], [11], [13], [14], [15]. Due to the long parking time of EVs in the buildings, the interaction between EV charging and building energy operation has also gained attention in recent years. Many works consider the joint scheduling of EVs and building energy management in a single building. The real-time coordination between EV charging and distributed photovoltaic generation in a building is studied in [17], [18], and [19]. In [20], two demand response control algorithms are proposed to coordinate EV charging with controllable loads for the power balance of the building. In [21], an long short-term memory (LSTM)-based energy management scheme is proposed for EV charging in the commercial-building. However, the above works neglect the interaction between EV charging and energy operation of multibuildings in the same microgrid. To fill this gap, works related to the multibuildings are emerging over the last few years. In [12] and [22], a decentralized EV charging scheduling method and a distributed policy improvement method are proposed considering a microgrid of buildings with distributed wind power, respectively. However, these works neglect the transmission security on the aggregated power of buildings. The EV aggregation technique has also been studied to speed up the optimization for multibuildings [23], [24]. In [25], it jointly considers the energy scheduling of smart buildings and the charging requirements while preserving privacy. In general, the existing works neglect the location transitions of EVs within multibuildings which bring in coupled charging relationship. Furthermore, few works consider the homogeneous charging actions in the microgrid of buildings which may incur transmission capacity violation. This is critical for the microgrid of buildings to be grid-friendly.

For convenience, the comparisons between our work and related surveys are summarized and provided in Table I. The comparisons include whether the uncertainties in the
TABLE I

COMPARISON BETWEEN OUR WORK AND RELATED SURVEYS

| Reference | Uncertainties in supply and demand | Security on aggregated power | Location transitions of EVs | Interaction with multi-buildings | Optimization speedup |
|-----------|------------------------------------|------------------------------|-----------------------------|----------------------------------|--------------------|
| [9]       | ✓                                  | ✓                            | ✓                           | ✓                                | ✓                  |
| [10]      | ✓                                  | ✓                            | ✓                           | ✓                                | ✓                  |
| [11]      | ✓                                  | ✓                            | ✓                           | ✓                                | ✓                  |
| [12]      | ✓                                  | ✓                            | ✓                           | ✓                                | ✓                  |
| [13]      | ✓                                  | ✓                            | ✓                           | ✓                                | ✓                  |
| [14]      | ✓                                  | ✓                            | ✓                           | ✓                                | ✓                  |
| [15]      | ✓                                  | ✓                            | ✓                           | ✓                                | ✓                  |
| [16]      | ✓                                  | ✓                            | ✓                           | ✓                                | ✓                  |
| [17]      | ✓                                  | ✓                            | ✓                           | ✓                                | ✓                  |
| [18]      | ✓                                  | ✓                            | ✓                           | ✓                                | ✓                  |
| [19]      | ✓                                  | ✓                            | ✓                           | ✓                                | ✓                  |
| [20]      | ✓                                  | ✓                            | ✓                           | ✓                                | ✓                  |
| [21]      | ✓                                  | ✓                            | ✓                           | ✓                                | ✓                  |
| [22]      | ✓                                  | ✓                            | ✓                           | ✓                                | ✓                  |
| [23]      | ✓                                  | ✓                            | ✓                           | ✓                                | ✓                  |
| [24]      | ✓                                  | ✓                            | ✓                           | ✓                                | ✓                  |
| [25]      | ✓                                  | ✓                            | ✓                           | ✓                                | ✓                  |

Our Work

DRE and EV charging, the security on the aggregated power, the location transitions of EVs, the interaction with multi-buildings, and the optimization speedup are considered in the article. It can be observed that our work mainly studies an effective EV charging scheduling method for the microgrid of buildings considering the uncertainty, security, and EV location transition which is barely considered in the existing literature. Furthermore, an EBO method will be introduced to mitigate the curse of dimensionality caused by large state and action space during EV scheduling in a microgrid of buildings.

III. PROBLEM FORMULATION

A. System Description

We consider a microgrid of buildings as depicted in Fig. 1. In the microgrid, each building is equipped with DRE, hydrogen energy storage (HES), and charging piles. The buildings are required to provide charging service and keep load balance which help to improve the utilization of DRE and reduce the operation cost of the buildings. Typically, when the output of the DRE and HES cannot meet the demand for EV charging and buildings, the microgrid operation controller will procure power from the grid. The building can also sell power to the grid if the output of DRE is large. To reduce the EV charging impact and ensure transmission security, the microgrid operation controller should regulate EV charging behaviors in each building based on the supply and demand status. It is noted that this article mainly studies the EV charging scheduling in the microgrid of buildings. The EV drivers and the microgrid operation controller will benefit more if EV discharging is utilized. On the one hand, the charging cost can be reduced for EV drivers by discharging to earn profit. On the other hand, the microgrid of buildings can utilize the relatively cheap discharging service to achieve power balance compared with expensive electricity from the grid. However, the incorporation of EV discharging will make the model more complicated and cause significant differences in the control policy for EVs. This can be a future research topic.

In the following, we will first formulate this multistage stochastic decision problem as a MDP to capture the uncertain renewable energy and EV charging demand in each building. To simplify the discussions, the following assumptions are made. The first assumption is made as EVs are charged at a specific power level when connected to the charging piles. The second assumption is made as the generation cost of DRE is much cheaper than the procure cost from the grid, thus assuming to be zero [26]. The third assumption is made as the time-of-use price is usually used for residential buildings and office buildings in practice [27]. The fourth assumption is made as the prediction accuracy of the building load is relatively acceptable for engineering purpose [28]. Our model and method can also be extended to the case when the electricity price or the building load is uncertain by state augmentation.

1) The charge level of each EV is fixed.
2) The operation cost of DRE is omitted.
3) The electricity price from the grid is deterministic and known \textit{a priori}.
4) The load demand of the building is deterministic and known \textit{a priori}.
B. System Model

We consider this EV scheduling problem in a microgrid of buildings over the discretized horizon $t = 1, 2, \ldots, T$ where $t$ denotes the decision epoch and $\Delta T$ denotes the decision interval. There are $K$ buildings and $N$ EVs in the microgrid. The MDP model of the proposed problem is shown below.

1) System States: The system state at stage $t$ is defined as $S_t = [s^1_t, s^2_t, \ldots, s^K_t]$ where $k = 1, 2, \ldots, K$ and $s^k_t$ denotes the state of the $k$th building. For each $s^k_t$, it is defined as $s^k_t = [r^k_t, \lambda^k_t, n^k_{0,m}, n^k_{1,m}]$ where $r^k_t$ denotes the output of DRE in the building, $\lambda^k_t$ denotes the state of charge (SOC) of HES, $n^k_{0,m}$ denotes the number of EVs which must be charged at stage $t$, and $n^k_{1,m}$ denotes the number of EVs which can be charged at stage $t$.

2) Actions: The control action at stage $t$ is defined as $A_t = [a^1_t, a^2_t, \ldots, a^K_t]$ where $a^k_t$ denotes the specific action for the $k$th building. Each building should decide whether to provide charge service for the connected EVs at stage $t$. Therefore, there is $a^k_t = [z^1_{t,1}, z^1_{t,2}, \ldots, z^1_{t,N}]$ where $z^1_{t,i} \in \{0, 1\}$, $i = 1, 2, \ldots, N$. When $z^1_{t,i} = 1$, it means the $k$th building should provide charge service for the $i$th EV at stage $t$ if it is parked, otherwise $z^1_{t,i} = 0$. Obviously, there is $\sum_{k=1}^K z^1_{t,i} \leq 1$.

3) System Dynamics: As the energy status of EV and HES are both time-coupled, their system dynamics should be considered when action $A_t$ is decided for the current state $S_t$.

For each EV, we use a tuple $(T^i_t, E^i_t, L^i_t)$ to represent its remaining parking time, remaining required charging energy and parking location. $L^i_t \in \{0, 1, 2, \ldots, K\}$ and $L^i_t = 0$ if the $i$th EV is on the road. Then, there is

$$T^i_{t+1} = \begin{cases} T^i_t - \Delta T, & \text{if } L^i_t > 0 \\ t^i_{t+1}, & \text{if } L^i_t \times (1 - L^i_t) > 0 \\ 0, & \text{if } L^i_{t+1} = 0 \end{cases}$$

$$E^i_{t+1} = \begin{cases} E^i_t - z^i_{t,i} P \Delta T \psi^c, & \text{if } L^i_t > 0 \\ \eta^i_{t+1}, & \text{if } L^i_t \times (1 - L^i_t) > 0 \\ 0, & \text{if } L^i_{t+1} = 0 \end{cases}$$

where $P$ is the charge power, $\psi^c$ denotes the charge efficiency, $z^i_{t,i} = \max(z^1_{t,i}, z^2_{t,i}, \ldots, z^N_{t,i})$, $t^i_{t+1}$ and $\eta^i_{t+1}$ are both nonnegative random variables which denote the stochastic characteristic of EV charging demand in the future. As the location transitions for EVs are time-variant, the location transition probability can be established as

$$P(L^i_{t+1} | L^i_t) = \begin{bmatrix} p_{11}(t) & p_{12}(t) & \cdots & p_{1K}(t) & p_{10}(t) \\ p_{21}(t) & p_{22}(t) & \cdots & p_{2K}(t) & p_{20}(t) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ p_{K1}(t) & p_{K2}(t) & \cdots & p_{KK}(t) & p_{K0}(t) \end{bmatrix}$$

where $p_{k1}(t)$ denotes the EV is parked in the $K$th building at stage $t$ and moves to the first building at stage $t + 1$, and the meaning of the remaining probabilities in (3) is similar.

According to [29], the system dynamics of HES in each building can be depicted as follows:

$$k^i_{t+1} = \begin{cases} \max\{k^i_t - h^i_t / \phi^H, 0\}, & \text{if } h^i_t \geq 0 \\ \min\{k^i_t - h^i_t \phi^P, \kappa^i\}, & \text{if } h^i_t < 0 \end{cases}$$

where $k^i_t$ is the stored hydrogen of HES in the $k$th building at stage $t$, $\kappa^i$ is the hydrogen storage capacity of HES, $\phi^H$ is the round-trip efficiency from hydrogen to power, $\phi^P$ is the round-trip efficiency from power to hydrogen, $h^i_t$ is the discharge power of HES if $h^i_t \geq 0$, otherwise is the charge power of HES. Considering $k^i_{t+1} = k^i_t \sigma_{Hi}$, where $k^i_{t+1}$ is the stored hydrogen energy and $\sigma_{Hi}$ is the lower heating value of hydrogen. (4) can be rewritten as follows by multiplying both sides with $\sigma_{Hi} / \kappa^i$:

$$h^i_{t+1} = \begin{cases} \max\{b^i_t - h^i_t / \psi^H, 0\}, & \text{if } h^i_t \geq 0 \\ \min\{b^i_t - h^i_t \psi^P / \kappa^i, 1\}, & \text{if } h^i_t \leq 0 \end{cases}$$

where $\kappa^i = \kappa^i \sigma_{Hi}$ denotes the energy capacity of HES, $\eta^H = \phi^H / \sigma_{Hi}$ denotes the discharge efficiency of HES and $\eta^P = \phi^P \sigma_{Hi}$ denotes the charge efficiency of HES.

For each HES in the building, we regulate that HES will discharge when the DRE cannot satisfy the building demand and will charge if DRE is sufficient to meet the building demand, i.e.,

$$h^i_t = \begin{cases} - \min\{r^i_t - l^i_t - \theta^i_t \eta^H, \eta^H \}, & \text{if } r^i_t \geq l^i_t + p^i_t \\ \min\{r^i_t + p^i_t - r^i_t \eta^H, \eta^H \}, & \text{if } r^i_t \leq l^i_t + p^i_t \end{cases}$$

where $l^i_k$ denotes the net demand in the $k$th building, $p^i_k = \sum_{i=1}^N z^i_{t,k} P$ is the total charge power in the $k$th building, $h^i_{t,c}$ and $h^i_{t,dc}$ satisfy

$$\begin{align*}
\eta^H \cdot h^i_{t,dc} &= \min\{h^i_{t,c}, h^i_{t,c}(1 - b^i_t) \kappa^i / \eta^P \} \\
\eta^H \cdot h^i_{t,c} &= \min\{h^i_{t,c}, (1 - b^i_t) \kappa^i / \eta^P \}
\end{align*}$$

in which $h^i_{t,c}$ is the maximum charge/discharge power of HES.

Based on the supply demand status in the building, each building can sell excess power or procure power from the grid. Therefore, the exchange power $g_{k,ij}^t$ between the building and the grid can be depicted as follows:

$$g_{k,ij}^t = \begin{cases} \max\{l^i_t + p^i_t - r^i_t - h^i_t, 0\}, & \text{if } h^i_t \geq 0 \\ - \max\{r^i_t + h^i_t - l^i_t - p^i_t, 0\}, & \text{if } h^i_t < 0. \end{cases}$$

4) Constraints: The action $A_t$ corresponding to the state $S_t$ should satisfy the following constraints:

$$\sum_{k=1}^{K} g_{k,ij}^t = G_t$$

$$\eta^H \cdot G_t \leq \overline{G}$$

$$0 \leq \bar{g}_k \leq \eta^H \cdot G_t$$

$$0 \leq E^i_t \leq E^c_{cap}$$

$$\eta^H \cdot l^i_t \leq P \cdot T^i_t$$

$$r^i_t + h^i_t + \bar{g}_k = l^i_t + p^i_t$$.
the battery capacity. Constraint (13) constrains the remaining required charging energy should not exceed the maximum energy that can be supplied during the remaining parking time. This constraint is used to satisfy the driver’s charging demand. Constraint (14) denotes the load balance in each building.

5) Objective Function: As the responsibility of the microgrid operator controller is to optimize the operation cost during scheduling, the following expected cumulative cost within a sliding window is chosen as the objective function considering the uncertain charging demand and output of DRE in the buildings, i.e.,

\[
J = \sum_{t=0}^{t_0 + T_w - 1} E^t \left[ \sum_{k=1}^{K} c_t(s^k_t, a^k_t) \right] \quad (15)
\]

where \( t_0 \) denotes the current decision stage, \( T_w \) denotes the sliding window, \( \pi \) denotes the EV scheduling policy, and \( c_t(s^k_t, a^k_t) \) denotes the one-step cost incurred by taking action \( a^k_t \) at state \( s^k_t \) for the \( k \)th building which is defined as follows:

\[
c_t(s^k_t, a^k_t) = \begin{cases} 
    o_t \left( l^k_t + p^k_t + h^k_{l,c} - r^k_t \right), & \text{if } p^k_t \leq r^k_t - l^k_t - h^k_{l,c} \\
    o_t \left( l^k_t + p^k_t - h^k_{l,dc} - r^k_t \right), & \text{if } p^k_t \geq r^k_t - l^k_t + h^k_{l,dc} \\
    0, & \text{else}
\end{cases}
\]

(16)

where \( o_t \) denotes the electricity price. It is noted that as the building load is inelastic and the building can sell/procure power in case of excess/insufficient generation, the cost of load shedding is not considered in (16).

Based on this model, the microgrid operator controller should minimize \( J \) to find an optimal scheduling policy for each decision stage. However, due to the coupled constraint (10) and the large state space and action space of the problem, the exact optimal solution of the above model can rarely be derived [30]. In Section IV, we will explore an event-based approach with gradient search to approximately solve the problem.

IV. Solution Methodology

In this section, an event-based solution method with gradient search is developed to solve the proposed EV scheduling problem in a microgrid of buildings. The procedure for this method is as follows: First, the proposed model is reformulated within an EBO framework by introducing the event and event-based action. This can mitigate the impact of large state space. Second, based on this model, a randomized parametric event-based control policy is defined to determine the charge ratio for each building in the microgrid. Each building operator selects EVs to charge based on this charge ratio and a heuristic selection principle. This can mitigate the impact of large action space. Third, a constrained gradient-based policy optimization method is developed to find the optimal randomized parametric event-based policy based on the performance gradient. To ensure transmission security, an adjustment mechanism is also proposed to keep the feasibility of the policy during optimization.

A. Event Definition

Due to the well-known curse of dimensionality, the state space of the proposed model will increase exponentially with the number of buildings and EVs. Therefore, we propose an EBO framework to solve this large state space difficulty. In the EBO, the event is defined to depict the state transition and the decision is event-triggered [31]. Compared with the state-triggered method, the EBO can reduce the computation burden of the problem. When appropriately defining the event, EBO can be applied to solve MDPs with large state space [32]. Due to these advantages, EBO has been applied in various fields, such as heating, ventilation, and air conditioning (HVAC) control [32], [33], communication network [34], stock trading [35], etc.

In this article, our idea to use EBO comes from the fact that it may not need to describe the detailed EV charging state and output of the DRE which incurs the large state space. Instead, we can use an event to depict the elastic ratio of the EV charging, DRE and HES. The larger elastic ratio means the larger control margin during scheduling.

Based on this idea, we can first define the elastic ratio of EV charging in each building, i.e.,

\[
I_{t, EV} = \frac{\langle n^k_{t,c} - n^k_{t,m} \rangle P}{\langle n^k \rangle P} \quad (17)
\]

where \( \langle n^k \rangle \) denotes the number of charging piles in the \( k \)th building. Equation (17) describes the elastic degree of charging which can be delayed. Second, the elastic ratio of HES in each building can be defined as its SOC, i.e., \( I_{t, HES} = b^k_t \). The larger \( b^k_t \) means that the HES can provide a more flexible charge/discharge power for operation cost optimization. Third, the elastic ratio of DRE can be defined as follows:

\[
I_{t, DRE} = \frac{\bar{r}^k_t}{r^k} \quad (18)
\]

where \( \bar{r}^k_t \) denotes the generation capacity of the DRE in the \( k \)th building. Equation (18) describes the generation level of the DRE and the higher level indicates the larger saving potential of the operation cost.

Based on the elastic ratios introduced above, we can finally define the event as follows:

\[
E = \{ e^k_t | t = 1, 2, \ldots, T, k = 1, 2, \ldots, K \} \quad (19)
\]

where

\[
e^k_t = \{ s^k_{t-1} < s^k_t > (I_{t, EV} + I_{t, HES} + I_{t, DRE})/3 \in [0, 1] \} \quad (20)
\]

In (19), \( e^k_t \) denotes the triggered event in the \( k \)th building at stage \( t \). The value of \( e^k_t \) is within [0, 1]. If we divide this interval equally with discrete unit set as 0.1, the event space for each building is limited to 10 which is largely reduced compared with the large state space in the original MDP model.

B. Randomized Parametric Event-Based Control Policy

Another difficulty of the original MDP is the large action space. It is of great computation burden to compute the charge
control variables \( z^k_{t+i} \) for each EVs. Therefore, we propose a randomized parametric event-based control policy to mitigate the impact of large action space.

The charge control of each EV is implemented in two steps. First, the microgrid operator controller decides a parametric charge ratio \( a_t^k \) for each building which chosen as the event-based action, i.e., \( a_t^k = r_t^k \in [0, 1] \). In this way, the total charge power for each building can be described as follows:

\[
p_t^k = \left[n_t^{k,m} + a_t^k (n_t^{k,e} - n_t^{k,m}) \right] P.
\] (21)

As \( a_t^k \in [0, 1] \), the action space of the proposed problem can be largely restricted. Based on [33], the performance of randomized policy may be better than deterministic policy and easier to obtain. Therefore, we will find a randomized parametric event-based control policy \( \sigma \) for the proposed problem, i.e.,

\[
\sigma : \mathcal{E} \rightarrow \mathcal{P}(a_t^k)
\] (22)

where \( \mathcal{P} \) denotes a probability distribution over action space. Equation (22) means that the action \( a_t^k \) will be chosen based on probability distribution \( \mathcal{P} \) and observed event \( e_t^k \). When an optimal randomized parametric event-based control policy is obtained, the action with the highest probability can be selected for implementation in practice.

Second, after the microgrid operator controller allocates the charge ratio for each building, the charge controller in each building should decide which EV should be charged and keep the total number of charged EVs within \( n_t^{k,m} + a_t^k (n_t^{k,e} - n_t^{k,m}) \). Therefore, we use a modified least-laxity-longer-processing-time-first (mLLLP) principle to select EVs to charge, which is introduced in [36]. The mL LLP principle generates a complete order among EVs and selects EVs based on the remaining processing time \( E_t^i / P \) and the EV laxity \( T_t^i - E_t^i / P \).

C. Constrained Gradient-Based Policy Optimization for EBO

Due to the existence of (10), the proposed model is coupled among all the buildings in the microgrid. Therefore, we propose a constrained gradient-based policy optimization method to search the optimal randomized parametric event-based policy considering this coupled constraint.

First, we neglect (10) and the proposed model can be decoupled into a single building optimization problem. For each building, let \( \sigma \) and \( \nu \) denote two event-based policies. Then, the following state-based performance difference formula can be derived based on [37]:

\[
j_{t_0}^{k,\sigma}(s_{t_0}^k) - J_{t_0}^{k,\nu}(s_{t_0}^k) = j_{t_0}^{k,\sigma} P_{t_0}^{k,\sigma}(j_{t+1}^{k,\sigma} - j_{t+1}^{k,\nu})
+ j_{t_0}^{k,\nu} P_{t_0}^{k,\nu}(j_{t+1}^{k,\nu} - j_{t+1}^{k,\nu})
+ \sum_{t=t_0+T} j_{t}^{k,\sigma} \left[j_{t_0}^{k,\sigma} - j_{t_0}^{k,\nu} + (P_{t_0}^{k,\sigma} - P_{t_0}^{k,\nu}) j_{t_0}^{k,\nu} \right]
\] (23)

where \( j_{t_0}^{k,\sigma} \) and \( j_{t_0}^{k,\mu} \) denote the value function from stage \( t_0 \) to \( t_0 + T_0 - 1 \) corresponding to the event-based policy \( \sigma \) and \( \nu \) for the \( k \)-th building, \( \beta_t^{k,\sigma} \) is the state distribution at stage \( t \) corresponding to \( \sigma \), \( P_{t_0}^{k,\sigma} \) and \( P_{t_0}^{k,\nu} \) denote the transition probability at stage \( t \) corresponding to \( \sigma \) and \( \nu \). \( c_t^{k,\sigma} \) and \( c_t^{k,\nu} \) denote the one-step cost at stage \( t \) corresponding to \( \sigma \) and \( \nu \). It is noted that the initial distribution \( \beta_0^{k,\sigma} \) is independent with policy \( \sigma \) or \( \nu \). Using recursion and \( \beta_t^{k,\sigma} P_{t_0}^{k,\sigma} = \beta_0^{k,\sigma} \), the last equation can be obtained.

Based on (23), we can extend the performance difference formula for EBO, i.e.,

\[
j_{t_0}^{k,\sigma}(s_{t_0}^k) - J_{t_0}^{k,\nu}(s_{t_0}^k) = \sum_{t=t_0}^{T} j_t^{k,\sigma} P_t^{k,\sigma}(j_{t+1}^{k,\sigma} - j_{t+1}^{k,\nu})
+ \sum_{t=t_0}^{T} j_t^{k,\nu} P_t^{k,\nu}(j_{t+1}^{k,\nu} - j_{t+1}^{k,\nu})
+ \sum_{t=t_0}^{T} j_t^{k,\nu} (P_{t}^{k,\sigma} - P_{t}^{k,\nu}) j_{t_0}^{k,\nu}
\] (24)

where \( S \) denotes the state space. When policy \( \nu \) is close to \( \sigma \), the performance gradient at stage \( t_0 \) can be derived as follows after observing event \( e_{t_0}^k \):

\[
\frac{\partial j_{t_0}^{k,\sigma}(s_{t_0}^k)}{\partial \sigma_{t_0}} = \beta_{t_0}^{k,\sigma} \sum_{s_{t_0}^k} \beta^{k,\sigma}_{t_0} \left[j_{t_0}^{k,\sigma} - j_{t_0}^{k,\nu} + (P_{t_0}^{k,\sigma} - P_{t_0}^{k,\nu}) j_{t_0}^{k,\nu} \right]
\] (25)

where \( \sigma_{t_0} \) denotes the detailed charge control policy at stage \( t_0 \) for the \( k \)-th building.

It is noted that policy \( \sigma \) is a randomized policy which selects the action with specific probability. Suppose there are \( M \) actions for the \( k \)-th building at stage \( t_0 \) and each action is denoted as \( \alpha_{t_0}^{k,m} \). Let \( p_{t_0}^{k,m} \) denote the probability to choose \( \alpha_{t_0}^{k,m} \). Then, there is the following relationship. The proof is given in Appendix

\[\begin{align}
\frac{\partial P_{t_0}^{k,\sigma}}{\partial \sigma_{t_0}^{k,m}} &= \frac{M}{\sum_{i=1}^{M} p_{t_0}^{k,i} - p_{t_0}^{k,m}} P(s_{t_0+1}^{k} | s_{t_0}^{k}, \alpha_{t_0}^{k,m}) \\
&\quad + \sum_{i \neq m} \frac{-p_{t_0}^{k,i}}{\sum_{i=1}^{M} p_{t_0}^{k,i}} P(s_{t_0+1}^{k} | s_{t_0}^{k}, \alpha_{t_0}^{k,i}) \\
\frac{\partial j_{t_0}^{k,\sigma}}{\partial \sigma_{t_0}^{k,m}} &= \frac{M}{\sum_{i=1}^{M} p_{t_0}^{k,i} - p_{t_0}^{k,m}} c(s_{t_0}^{k}, \alpha_{t_0}^{k,m}) \\
&\quad + \sum_{i \neq m} \frac{-p_{t_0}^{k,i}}{\sum_{i=1}^{M} p_{t_0}^{k,i}} c(s_{t_0}^{k}, \alpha_{t_0}^{k,i}).
\end{align}\]
Substituting (26) and (27) into (25), the policy gradient can be finally obtained which is shown below

\[
\frac{\partial J_{t_b}^{k,\sigma}(e_k^t)}{\partial \sigma_{t_0}^{k,\alpha}} = \frac{1}{\Delta} \left\{ \sum_{i,m} p_{t_b}^{k,i} \sum_{\alpha_{t_0}^{k,\alpha}} \partial_{\sigma_{t_0}^{k,\alpha}} V(s_{t_0}^{k,\alpha}, \sigma_{t_0}^{k,\alpha}) \right\}
\]

denotes the incurred future total cost when taking action \( \alpha_{t_0}^{k,\alpha} \) for current state \( s_{t_0}^{k,\alpha} \) and there is

\[
\frac{\partial J_{t_b}^{k,\sigma}(e_k^t)}{\partial \sigma_{t_0}^{k,\alpha}} = \left( \frac{\partial J_{t_b}^{k,\sigma}(e_k^t)}{\partial p_{t_b}^{k,1}}, \frac{\partial J_{t_b}^{k,\sigma}(e_k^t)}{\partial p_{t_b}^{k,2}}, \ldots, \frac{\partial J_{t_b}^{k,\sigma}(e_k^t)}{\partial p_{t_b}^{k,M}} \right).
\]

Based on the policy gradient (28), the randomized parametric event-based control policy can be updated as follows during policy optimization:

\[
\sigma_{t_0,j+1} = \sigma_{t_0,j} - \delta_j \frac{\partial J_{t_b}^{k,\sigma}(e_k^t)}{\partial \sigma_{t_0}^{k,\alpha}}
\]

where \( \sigma_{t_0,j} \) denotes the updated event-based policy at \( j \)th iteration, \( \delta_j = 1/(1 + \xi j) \) denotes the update step at \( j \)th iteration and \( \xi \) denotes the decay factor.

Due to the uncertainties in the DRE generation and EV charging demand, it is impractical to analytically compute (28) under expectation. Therefore, the Monte Carlo simulation method is adopted to estimate (28). The estimation algorithm is summarized in Algorithm 1.

As mentioned before, the derived policy gradient neglects the coupled constraint (10). To satisfy this transmission security constraint, the following adjustment mechanism is proposed to ensure the feasibility of the policy \( \sigma \).

Adjustment Step I: If the total exchange power exceeds the upper bound of (10) by \( \Delta \), the total number of EVs to be charged should be reduced by \( \Delta/P \). Each building should reduce the number of charged EVs by \( \left( \frac{\partial J_{t_b}^{k,\sigma}}{\partial \sigma_{t_0}^{k,\alpha}} \Delta / \sum_{k=1}^{K} \frac{\partial J_{t_b}^{k,\sigma}}{\partial \sigma_{t_0}^{k,\alpha}} \right) / P \). The policy \( \sigma \) can be updated by solving the following quadratic programming problem:

\[
\begin{align*}
\min_{p_{t_b,j+1}^{k,i} \in [0,1]} \sum_{i=1}^{M} \left( p_{t_b,j+1}^{k,i} - p_{t_b,j}^{k,i} \right)^2 & \\
\text{s.t.} \sum_{i=1}^{M} p_{t_b,j+1}^{k,i} \sigma_{t_0}^{k,i} & = \sum_{i=1}^{M} p_{t_b,j}^{k,i} \sigma_{t_0}^{k,i} + \left( 1 - \frac{\partial J_{t_b}^{k,\sigma}}{\partial \sigma_{t_0}^{k,\alpha}} \Delta / P \right) \left( \frac{\partial J_{t_b}^{k,\sigma}}{\partial \sigma_{t_0}^{k,\alpha}} \right) / P (n_{t,c} - n_{t,m})
\end{align*}
\]

Note that \( \partial J_{t_b}^{k,\sigma}/\partial \sigma_{t_0}^{k,\alpha} \) can be considered as the marginal operation cost for the \( k \)th building. The proposed adjustment mechanism allocates the reduced or increased number of charged EVs for each building based on this marginal cost. When required to reduce the charge demand, the building with large marginal cost should largely reduce the number of charged EVs. On the contrary, when required to increase the charge demand, the building with small marginal cost should largely increase the number of charged EVs. The motivation of solving (32) and (33) is to minimize the probability difference between the adjacent policies \( \sigma_{t_0,j+1} \) and \( \sigma_{t_0,j} \) while reduce or increase the expected charge ratio to meet the reduced or increased requirement of charging demand for the \( k \)th building.
The idea of this adjustment mechanism is shown in Fig. 2. As there exists a large number of discrete variables and nonlinear constraints, the feasible policy space can be considered disconnected. When (10) is not violated, the policy update happens within a feasible policy set. When violated, the policy update should transfer to another feasible policy set based on this adjustment mechanism. In the end, the proposed constrained gradient-based policy optimization for the problem is summarized in Algorithm 2. Note that the gradient search method can ensure finding an improved policy at each iteration which can guarantee the global optimum in theory [38]. However, due to the intrinsic deficiencies of the gradient search method, the proposed method may encounter trapping into the local optimum if an inappropriate update step is selected. In the next section, we will analyze the optimality of the proposed method through numerical experiments.

V. NUMERICAL RESULTS

In this section, we evaluate the proposed method via simulations. The charge control policies for different types of buildings, the performance, and comparison of the proposed solution method will be analyzed in the following experiments.

A. Parameter Settings

We take building load data from [39] for experiments as shown in Table II. In the experiment, we consider there are three buildings in the microgrid. The first two are residential buildings and the rest is an office building. We take distributed wind generation data from [4] and set it as the predicted value. The actual output of DRE in each building is assumed to follow normal distribution with the predicted value as the mean and 10% of the predicted value as the standard deviation. Fig. 3 shows the realization of the actual output of DRE in three buildings. The time-of-use electricity price is shown in Table III.

![Fig. 2. Illustration of the adjustment mechanism.](image)

**Algorithm 2** Constrained Gradient-Based Policy Optimization

1: for $t_0 = 1, 2, \ldots, T$ do
2: Set $j \rightarrow 0$ and select the initial policy $\sigma_{0,j}$.
3: for $k=1,2,\ldots,K$ do
4: Compute the policy gradient based on Algorithm 1 when observing event $e_{k,0}$.
5: end for
6: Check whether constraint (10) is violated. If violated, update policy by adjustment mechanism and go to Step 3.
7: If $||(\partial J_{k,0}^{b,\sigma_j}/\partial \sigma_{0,j})||_2 \leq \epsilon$ or $||(\partial J_{k,0}^{b,\sigma_j}/\partial \sigma_{0,j}) - (\partial J_{k,0}^{b,\sigma_{j-1}}/\partial \sigma_{0,j})||_2 \leq \epsilon$, go to Step 1.
8: Update policy using (31) and go to Step 3.
9: end for

![Fig. 3. Realization of actual wind output in three buildings.](image)

**TABLE II**

| Time(h) | $k = 1$ | $k = 2$ | $k = 3$ |
|---------|---------|---------|---------|
| 1       | 139     | 180     | 367     |
| 2       | 103     | 120     | 353     |
| 3       | 144     | 180     | 333     |
| 4       | 127     | 147     | 333     |
| 5       | 151     | 180     | 433     |
| 6       | 150     | 180     | 387     |
| 7       | 67      | 84      | 520     |
| 8       | 202     | 240     | 567     |
| 9       | 216     | 264     | 820     |
| 10      | 151     | 191     | 1053    |
| 11      | 147     | 170     | 967     |
| 12      | 150     | 258     | 973     |

**TABLE III**

| Price      | Time       |
|------------|------------|
| 0.3515CNY/kWh | 23:00-6:00 |
| 0.8135CNY/kWh | 7:00-10:00 |
| 0.4883CNY/kWh | 11:00-18:00 |
| 0.8135CNY/kWh | 19:00-22:00 |
distribution of the trip distance and the electric drive efficiency as introduced in [4]. We take the parameters of HES from [41]. The detailed parameters are shown in Table IV.

In the experiment, we consider this problem on a daily basis with $T = 48$, $AT = 30$ min and $T_w = 12$. The event is evenly discretized by 0.1. The action selection probabilities of the initial policy $\sigma$ are all set as equal. Note that the parameters in the experiment are for illustration purposes, which should be adjusted in practice.

### B. Result Analysis

We firstly look into the optimized selection probability of each action at each decision stage for these three buildings. The results are shown in Fig. 4(a)–(c). Note that these probabilities are obtained after observing the occurred event at each decision stage which is shown in Fig. 5(a)–(c). In Fig. 4(a)–(c), the darker color indicates a higher selection probability for the specific action. It can be seen that the selection probabilities tend to be high near the departure time, such as 7:00 in building #1 and building #2 and 17:00 in building #3, while the selection probabilities tend to be low near the arrival time, such as 17:30 in building #1 and building #2 and 7:30 in building #3. The reason why the charging probability is small at arrival and becomes large at departure lies in two aspects. The first is that the parking deadline approaches and the charging demand should be satisfied. Another important reason is that the distributed wind power begins to increase during the time interval (2:30–7:30) and (14:30–17:00) for building #3. Therefore, it indicates that the randomized parametric event-based control policy can be reduced to a determined event-based control policy during these periods. Note that as there are few EVs in building #1 and building #2 during (8:00–16:00), there is no action whose selection probability is significantly higher than others. The same reason holds for building #3 during (24:00–6:00). Furthermore, as there is no EV parked in building #3 after 20:00 considering the time window $T_w$ set as 6 h, the policy gradient keeps zero and the selection probability for each action remains unchanged and equal. Comparing Fig. 4(a) with (b), it can be seen that their selection probability distributions are similar. This is caused by the similar departure time of EVs in building #1 and building #2. The main difference lies in (16:30–17:00) where the low selection probability of building #1 appears earlier than building #2. This is caused by the longer trip time between building #2 and building #3. Comparing Fig. 4(c) with (a) and (b), it can be seen that the selection probability distributions are almost complementary due to the coupled charging relationship between building #3 and the rest two buildings.

The total charge power and observed event at each stage are also shown in Fig. 5(a)–(c). Due to the stochastic charging demand of EVs and their distinct departure, the charging behavior occurs during (16:00–10:00) for building #1 and building #2 and during (10:00–18:00) for building #3. Furthermore, it can be found that the peak of the total charging power occurs later in building #2 than building #1 after 15:00. This is also because the trip time from building #3 to building #2 is longer than that from building #3 to building #1. Based on the delayed feature of the optimized selection probability in building #3, the peak occurrence of charging power is also delayed to 15:00 and gradually decreases due to the departure of EVs. It can be seen that the event index reaches a peak during (7:00–8:00) for these three buildings. This is because the distributed wind power generation reaches a peak during this period which means a high elastic ratio of DRE according to (18). After this period, the event index begins to decrease for these buildings due to the decreased elastic ratio of EV charging caused by EV departure according to (17). Note that the high event index around 15:00 is also caused by the high generation of distributed wind power. Comparing Fig. 5(c) with (a) and (b), it is seen that the event index of building #3 is lower than the event index of building #1 and building #2.

| Parameter | Setting | Parameter | Setting |
|-----------|---------|-----------|---------|
| $E_{em}$  | 36kWh   | $P$       | 3.6kW   |
| $\phi^p$  | 0.92    | $s_i$     | 166.63kWh |
| $h_{em}$  | 50kW    | $\eta^f$  | 0.82    |
| $\eta^f$  | 0.62    | $\xi$     | 0.1     |
| $\Omega$  | 5600kW  | $\xi$     | 0.1     |
on average. This indicates the relatively low elasticity of office building #3. The main reason is that the total charging power of building #3 is higher than other residential buildings. This causes the frequent usage of the HES and the reduction of HES elastic as shown in Fig. 6. Due to the low generation of distributed wind power and few parked EVs, the event index reaches its lowest during (1:00–5:00) and (19:00–24:00) for building #3.

The control strategy of the HES in each building is shown in Fig. 6(a)–(c). It can be found that the trend of the SOC for each HES is similar to the trend of the observed event in Fig. 5. This is because the SOC of HES is one of the main factors which influence the value of the observed event. The difference between the trend of the observed event and the SOC of HES is caused by the distributed wind power generation and EV charging elasticity. In Fig. 6(a)–(c), the decrease of SOC indicates the HES provides power for balancing building load and EV charging load and the increase of SOC indicates the excess generation of distributed wind power. It can be seen that the HES in all these buildings will provide power for EV charging at the beginning and begins to be charged during (5:00–8:00) due to the increase of the distributed wind power in this period. The difference between building #3 and the rest buildings mainly lies in the period after 9:00. The HES of residential buildings #1 and #2 will begin to provide power only after 17:00 as more and more EVs arrive at home. On the contrary, the HES in building #3 will begin to discharge at early 9:00. This is because the proposed method will obtain the information of large charging demand and large generation of distributed wind power during a later period (14:00–16:00) from the generated sample paths. This causes two discharging events to happen at 9:00 and 16:30 for the HES of building #3.

To analyze the performance of the optimized event-based control policy, we compare the derived event-based policy with the rule-based charging policy and the ideal charging policy. The rule-based charging policy will meet the EV charging demand as soon as possible once connected to the charging pile in the building. The ideal charging policy is derived by implementing the MPC method with precise information of the EV charging demand and wind power generation and the same length of the sliding window. Particularly, the optimal scheduling of HES is also considered as the control variable in MPC. The performance of the above three policies is shown in Table V. It can be seen that the rule-based charging policy achieves the highest operation cost as the EV charging control has no relationship with the building load and supply. On the contrary, the ideal charging achieves the lowest operation cost. However, this policy cannot be implemented in practice due to its requirement of seeing the future. Comparing the event-based policy with ideal charging, it can be found that the performance of our policy is closer to the ideal policy and better than the rule-based charging policy. The cost difference between event-based charging and ideal charging is only 154CNY which means the optimization gap is less than 1%. This indicates the proposed constrained gradient-based policy optimization can avoid trapping into local optimum and

| Table V | Performance Comparison of Charging Control Policies |
|---------|---------------------------------------------|
| Term    | Rule-based | Event-based | Ideal |
| Total Operation Cost | 28163.9CNY | 24965.9CNY | 24411.9CNY |
| Min. Exchange Power   | -2766.1kW  | -2887.1kW  | -2876.3kW  |
| Max. Exchange Power   | 5965.2kW   | 5589.1kW   | 5560.5kW   |
demonstrates the effectiveness of the proposed EV charging control method.

Fig. 7 shows the total exchange power between the microgrid and the main grid under the above three policies. It can be seen that the total exchange power of the microgrid reaches a peak around 19:00 due to the high building load in building #1 and building #2 at night. It can also be seen that the rule-based charging policy exceeds the maximum transmission power \( G = 5600 \text{ kW} \) while the peak of the latter two policies is below this upper bound. The maximum total exchange power of these policies is shown in Table V. This indicates that the main reason why event-based charging and ideal of these policies is shown in Table V. This indicates that that the minimum iteration number is 5 which happens at 7:30 and the maximum iteration number is 25.3 iterations on average to find the optimal event-based charging policy.

Therefore, Fig. 8 shows the total iteration number at each decision stage. It can be seen that the minimum iteration number is 5 which happens at 7:30 and the maximum iteration number is 25.3 iterations on average to find the optimal event-based charging policy.

VI. CONCLUSION

In this article, the EV charging scheduling problem in a microgrid of buildings is studied to optimize the total operation cost of the microgrid while ensuring its transmission security. The MDP formulation is introduced to capture the uncertain supply and EV charging demand in the buildings. To mitigate the large state and action space difficulties, we reformulate it within an EBO framework with searchable control policy space. A constrained gradient-based policy optimization approach is proposed to find an optimal randomized parametric control policy for EV charging. We analyze the structure of the control policy through numerical experiments and demonstrate the proposed method can reduce the total operation cost while ensuring transmission security for the microgrid of buildings.

APPENDIX A

PROOF OF (26) AND (27)

Proof: Based on the randomized control policy \( \sigma \), there is

\[
P(k_{i}^{m} | s_{k}^{m}) = \frac{p_{k}^{m}}{\sum_{i=1}^{M} p_{k_{i}^{m}}} \tag{34}
\]

\[
p_{k}^{m} = \sum_{j=1}^{M} P\left(a_{i}^{k_{i}^{m}} | s_{k}^{m}\right) P\left(s_{k}^{m} | l_{k_{i}^{m}, M}^{i}, a_{k_{i}^{m}}^{i}\right) \tag{35}
\]

For the selection probability \( P(a_{i}^{k_{i}^{m}} | s_{k}^{m}) \), there is

\[
\frac{\partial P(a_{i}^{k_{i}^{m}} | s_{k}^{m})}{\partial p_{k}^{m}} = \begin{cases} \frac{\sum_{i=1}^{M} p_{i}^{m} - p_{i}^{m}}{(\sum_{i=1}^{M} p_{i}^{m})^{2}}, & \text{if } i = m \\ \frac{-p_{i}^{m}}{(\sum_{i=1}^{M} p_{i}^{m})^{2}}, & \text{if } i \neq m \end{cases} \tag{36}
\]

As only \( P(a_{i}^{k_{i}^{m}} | s_{k}^{m}) \) depends on \( p_{k}^{m} \) in (35), (26) is obtained by taking derivative of (35) with \( p_{k}^{m} \) and substituting (36) into it.

Similarly, for the one-step cost \( c_{k_{i}^{m}, a_{i}^{k_{i}^{m}}} \), there is

\[
c_{k_{i}^{m}}(s_{k}^{m}) = \sum_{i=1}^{M} P(a_{i}^{k_{i}^{m}} | s_{k}^{m}) c(s_{k}^{m}, a_{i}^{k_{i}^{m}}). \tag{37}
\]
Taking derivative of (37) with respect to $p_{th}^{k,m}$ and substituting (36) into it, (27) can be obtained.

REFERENCES

[1] Global EV Outlook 2022, International Energy Agency, Paris, France, 2022.
[2] N. Kumar, B. Sivanesan, and P. L. So, “Impact of priority criteria on electric vehicle charging scheduling,” IEEE Trans. Transport. Electrific., vol. 1, no. 3, pp. 200–210, Oct. 2015.
[3] J.-H. Teng, S.-H. Liao, and C.-K. Wen, “Design of a fully decentralized controlled electric vehicle charger for mitigating charging impact on power grids,” IEEE Trans. Ind. Appl., vol. 53, no. 2, pp. 1497–1505, Mar. 2017.
[4] Q. Huang, Q.-S. Jia, and X. Guan, “A multi-timescale and bilevel coordination approach for matching uncertain wind supply with EV charging demand,” IEEE Trans. Autom. Sci. Eng., vol. 14, no. 2, pp. 694–704, Apr. 2016.
[5] D. Wu, H. Zeng, C. Lu, and B. Boulet, “Two-stage energy management for office buildings with workplace ev charging and renewable energy,” IEEE Trans. Transport. Electrific., vol. 3, no. 1, pp. 225–237, Mar. 2017.
[6] Z. Xiaodi, S. Yunliang, Y. Junwei, and D. Qing, “The method of charging piles planning in parking lot,” in Proc. IEEE Int. Conf. Power Syst. Technol. (POWERCON), Sep. 2016, pp. 1–5.
[7] A. S. Al-Oqaili et al., “Review on scheduling, clustering, and forecasting strategies for controlling electric vehicle charging: Challenges and recommendations,” IEEE Access, vol. 7, pp. 128353–128371, 2019.
[8] J. Kang, H. Kong, Z. Lin, and A. Dang, “Mapping the dynamics of electric vehicle charging demand within Beijing’s spatial structure,” Sustain. Cities Soc., vol. 76, Jan. 2022, Art. no. 103507.
[9] B. Sun, Z. Huang, X. Tan, and D. H. K. Tsang, “Optimal scheduling for electric vehicle charging with discrete charging levels in distribution grid,” IEEE Trans. Smart Grid, vol. 9, no. 2, pp. 624–634, Mar. 2018.
[10] A. M. Koufakis, E. S. Rigas, N. Bassiliades, and S. D. Ramchurn, “Offline and online electric vehicle charging scheduling with V2Y energy transfer,” IEEE Trans. Intell. Transp. Syst., vol. 21, no. 5, pp. 2128–2138, May 2020.
[11] M. F. Shaaban, S. Mohamed, M. Ismail, K. A. Qaraqe, and E. Serpentin, “Joint planning of smart ev charging stations and DGs in eco-friendly remote hybrid microgrids,” IEEE Trans. Smart Grid, vol. 10, no. 5, pp. 5819–5830, Sep. 2019.
[12] Y. Yang, Q.-S. Jia, X. Guan, Z. Zhang, Q. Gu, and G. Deconinck, “Decentralized EV-based charging optimization with building integrated wind energy,” IEEE Trans. Autom. Sci. Eng., vol. 16, no. 3, pp. 1002–1017, Jul. 2019.
[13] A. Fallah-Mehrjardi, M. H. Yaghmaeef, and A. Leon-Garcia, “Charge scheduling of electric vehicles in smart parking-lot under future demands uncertainty,” IEEE Trans. Smart Grid, vol. 11, no. 6, pp. 4949–4959, Nov. 2020.
[14] S. Zhao, X. Lin, and M. Chen, “Robust online algorithms for peak-minimizing EV charging under multistage uncertainty,” IEEE Trans. Autom. Control, vol. 62, no. 11, pp. 5739–5754, Nov. 2017.
[15] H. Li, Z. Wan, and H. He, “Constrained EV charging scheduling based on safe deep reinforcement learning,” IEEE Trans. Smart Grid, vol. 11, no. 3, pp. 2427–2439, May 2020.
[16] Q. Huang, Q.-S. Jia, and X. Guan, “Robust scheduling of EV charging load with uncertain wind power integration,” IEEE Trans. Smart Grid, vol. 9, no. 2, pp. 1043–1054, Mar. 2018.
[17] J. Liu, Q. Wu, M. Shahidehpour, C. Li, S. Huang, and W. Wei, “Transactive active-time electric vehicle charging management for commercial buildings with PV on-site generation,” IEEE Trans. Smart Grid, vol. 10, no. 5, pp. 4939–4950, Sep. 2019.
[18] B. Kandpal and A. Verma, “Demand peak reduction of smart buildings using feedback-based real-time scheduling of EVs,” IEEE Syst. J., early access, Oct. 19, 2021, doi: 10.1109/JSYST.2021.3117715.
[19] C. A. Roy, N. D. Bequette, F. Gei, M. Büscher, R. Sabelnien, and D. Driesen, “Electric vehicle charging in an office building microgrid with distributed energy resources,” IEEE Trans. Sustain. Energy, vol. 5, no. 4, pp. 1389–1396, Oct. 2014.
[20] J. Hu, Z. Zhang, J. Lu, J. Yu, and J. Cao, “Demand response control of smart buildings integrated with security interconnection,” IEEE Trans. Control Syst. Technol., vol. 21, no. 3, pp. 435–445, Jun. 2012.
[21] H. Zhou et al., “LSTM-based energy management for electric vehicle charging in commercial-building prosumers,” J. Mod. Power Syst. Clean Energy, vol. 9, no. 5, pp. 1205–1216, Sep. 2021.
Li Yang (Member, IEEE) received the Ph.D. degree in control science and engineering from the Nanjing University of Science and Technology, Nanjing, China, in 2009. She is currently a Professor with the School of Automation, Nanjing University of Science and Technology, where she has been the Dean of the School of Automation since 2022. Her research interests include network optimization and control, space-ground integrated networks, theory and method of wireless network protocol.

Chen Hou (Member, IEEE) received the M.S. and Ph.D. degrees in control science and engineering from Tsinghua University, Beijing, China, in 2015 and 2019, respectively. He was a Visiting Scholar with the University of Illinois at Urbana-Champaign (UIUC), Champaign, IL, USA, in 2018, under the supervision of Prof. Tamer Basar. He held the position of a Senior Engineer, a Chief Architect, and the WeTransport FT Manager with the Tencent Future Network Laboratory, Tencent Ltd., Beijing, from 2019 to 2021. He joined the China Agricultural University (CAU), Beijing, as an Associate Professor in 2021. He holds over 200 granted and issued invention patents. His research interests include the operation and control for cyber-physical systems (CPS), the Internet of Things (IoT), and edge computing, and their applications in industry and agriculture.

Dr. Hou received the Tsinghua University Excellent Ph.D. Dissertation Award for his M.S. and Ph.D. degrees.

Zhiyong Zeng (Member, IEEE) received the Ph.D. degree in electrical engineering from Zhejiang University, Hangzhou, China, in July 2016. From September 2016 to February 2021, he was a Post-Doctoral Associate with Duke University, Durham, NC, USA. He is currently a Faculty Member with the Nanjing University of Science and Technology, Nanjing, China. His research interests include modular pulse synthesizers for noninvasive brain magnetic stimulation, vector sequences optimizations of modular multilevel converters, and fault-tolerant control of grid-connected power converter.

Yaowen Qi is currently pursuing the Ph.D. degree with the School of Automation, Nanjing University of Science and Technology, Nanjing, China. His research interests include network routing planning and network reliability analysis.