2024 IEEE 63rd Conference on Decision and Control (CDC)
December 16-19, 2024. MiCo, Milan, Italy

Real-time Pricing Mechanism for V2G Using Distributed Bilevel Optimization

Jiawei You, Xinliang Dai, Yuning Jiang*, Haoyu Yin, Yuanming Shi, and Colin N. Jones

Abstract— **This paper explores the impact of the burgeoning electric vehicle (EV) presence on distribution grid operations, highlighting the challenges they present to conventional pricing strategies due to their dual role as power consumers and suppliers, coupled with their energy storage capabilities. We propose an advanced real-time pricing model for the electricity market, employing a novel distributed bilevel optimization framework. This framework distinguishes between the distribution system operator (DSO) at the upper level and the EVs at the lower level, each aiming to optimize profit margins. The optimization includes power flow constraints at the upper level to ensure efficient operation within safe system limits, while model predictive control (MPC) is used to optimize lower-level EV responses. Additionally, we provide a rigorous convergence analysis of the proposed bilevel optimization method. Detailed convergence studies and simulation results demonstrate the effectiveness and superiority of the proposed algorithm.**

I. Introduction

The recent rise in urban electric vehicle (EV) adoption impacts the distribution grid due to their charging and discharging activities. With vehicle-to-grid (V2G) and grid-to-vehicle (G2V) capabilities, EVs contribute to both energy consumption and supply, boosting grid resilience and efficiency [5], [12], [25]. Since EVs are often parked for long periods, fast-charging technology reduces charging times, encouraging participation in V2G programs that support grid stability and efficiency. Electricity pricing influences EV charging behavior, prompting owners to charge during low-rate periods and discharge during peak times [13], [18], [26]. This interaction is crucial for a sustainable energy future.

The variability of EV power challenges traditional electricity pricing models, which rely on historical data and fixed intervals, often struggling with the market's dynamic nature [1], [17]. Intelligent pricing offers a more adaptive solution, promoting efficient energy use through strategies like critical peak [22], time-of-use [27], and real-time pricing [21]. Real-time pricing, which adjusts rates based on immediate user data, is particularly effective in managing EV integration into the grid. Using a Stackelberg game framework, where the distribution system operator (DSO) leads and consumers follow [23], this approach aligns with shifting market demands and supports a responsive and sustainable energy ecosystem.

Bilevel optimization, with its nested structure and interdependent variables, poses significant challenges. Various algorithms have been developed to tackle this [10], [11], [15], [19], often using first-order optimality conditions from the lower-level problem as constraints in the upper level [7]. While this simplifies the bilevel problem to a single level, it struggles with high-dimensional variables due to excessive equality constraints. To address this, gradient-based techniques like Approximate Implicit Differentiation (AID) [16] and Iterative Differentiation (ITD) [6] have been introduced, offering more efficient solutions for complex bilevel problems.

The energy storage characteristics of EVs require dynamic battery state analysis, adding uncertainty to the electricity market and differentiating them from traditional grid loads. DSOs must incorporate EV behavior predictions into pricing strategies. Current EV modeling methods include model-free approaches like Reinforcement Learning (RL) [8], [9], [24], which handles sequential optimization in V2G but faces convergence challenges, and model-based approaches like Model Predictive Control (MPC) [4], [14], [20], which excels in real-time and multi-scale optimization. Combining MPC with bilevel optimization could further enhance power market stability and efficiency.

Recent research on EV integration and electricity pricing often prioritizes grid revenue, overlooking load, physical models, and long-term uncertainties in EV charging, resulting in unrealistic grid solutions. Additionally, many studies neglect convergence analysis when assessing algorithms, potentially compromising market stability. In this paper, a novel bilevel formulation for the V2G pricing problem will be proposed in Section II. We employ an MPC-based strategy for real-time EV charging and discharging decisions at the lower level, while at the upper level, we account for the physical structure of the distribution network, integrating nonconvex optimal power flow with a price sequence as decision variables. A significant contribution of our work, detailed in Section III, is the development of a distributed bilevel optimization algorithm with comprehensive convergence guarantees. We propose a real-time pricing mechanism for V2G operations, employing a receding horizon approach. Our methods are validated through a case study using the IEEE 15-bus benchmark with 80 EVs, demonstrating the algorithms' effectiveness in closed-loop scenarios in Sec-

* corresponding author: yuning.jiang@ieee.org

The work of JY, YJ, CNJ was supported by the Swiss National Science Foundation (SNSF) under the NCCR Automation project, grant agreement 51NF40_180545. The work of YS was supported in part by the National Nature Science Foundation of China under Grant 62271318 and the Shanghai Rising-Star Program under Grant No. 22QA1406100.

JY and YS are with the School of Information Science and Technology, ShanghaiTech University, China, and JY is also a Visiting Student at EPFL.

Xinliang Dai is with the Institute for Automation and Applied Informatics, Karlsruhe Institute of Technology, Germany.

Yuning Jiang and Colin N. Jones are with the Automatic Control Laboratory, EPFL, Switzerland.

Haoyu Yin is with the Department of Electrical and Systems Engineering, Washington University in St. Louis, USA.

tion IV. Section V concludes the paper with outlooks.

Notations: Throughout the paper, we use boldface lower case and upper case letters to represent vectors and matrices, respectively. Moreover, we use notation $\mathbf{I}_n$ to denote the identity matrix in $\mathbb{R}^{n\times n}$, and $\mathbb{Z}_{z_1}^{z_2} = \{z \in \mathbb{Z} | z_1 \le z \le z_2\}$ denotes integer ranges from z_1 to z_2. Besides, notation $\mathrm{vec}(\mathbf{A})$ denotes the vector obtained by stacking all columns of $\mathbf{A}$ into one long vector. And notation $\|\cdot\|_{\mathrm{F}}$ defines the Frobenius norm for a given matrix. For given two matrices $\mathbf{A}$ and $\mathbf{B}$, we denote the inner product $\langle \mathbf{A}, \mathbf{B}\rangle = \mathrm{Tr}(\mathbf{A}^\top \mathbf{B})$. For a scale sequence $x_1, ..., x_n$, we use notation $\mathbf{x} = [x_1, ..., x_n]^\top \in \mathbb{R}^n$.

II. Problem Formulation

In this section, we introduce the system model for the radial distribution grid and the battery model for electric vehicles (EVs). Subsequently, we encapsulate the entire design problem of the real-time pricing mechanism as a bilevel optimization problem. The lower-level problem involves model predictive control (MPC) for managing EV charging/discharging at charging stations, and the upper-level problem handles optimal power flow (OPF) within the distribution grid.

A. OPF-based DSO Decision-Making

Let us consider a radial distribution grid that connects multiple charging stations distributedly, where EVs engage in charging or discharging activities. The distribution system can be defined by the tuple $(\mathcal{N}, \mathcal{E}, \mathcal{G}, \mathcal{B})$, where $\mathcal{N} = \{0, 1, \cdots, N\}$ represents the bus set, with bus 0 assumed to be a substation with a fixed voltage. $\mathcal{E} \subset \mathcal{N} \times \mathcal{N}$ is the transmission line set with $|\mathcal{E}| = N$, $\mathcal{G} \subset \mathcal{N}$ is the generator set, and $\mathcal{B} \subset \mathcal{N}$ is the charge station set with $|\mathcal{B}| = M$.

To describe the power balance of the distribution grid, the relaxed DistFlow model is utilized as follows:

$$p_n = \sum_{k\in\mathcal{C}_n} P_k - P_n + r_n \ell_n, \tag{1a}$$

$$q_n = \sum_{k\in\mathcal{C}_n} Q_k - Q_n + x_n \ell_n, \tag{1b}$$

$$v_{\pi_n} = v_n + 2r_n P_n + 2x_n Q_n - (r_n^2 + x_n^2)\ell_n, \tag{1c}$$

$$\ell_n \ge (P_n^2 + Q_n^2)/v_{\pi_n} \tag{1d}$$

for all $n \in \mathcal{N}^0 := \mathcal{N}\backslash\{0\}$, where π_n is the parent bus of bus n and $\mathcal{C}_n$ is the children set of bus n. ℓ_n, P_n, and Q_n denote the squared current magnitude, the active and reactive power through the transmission line $(\pi_n, n) \in \mathcal{E}$, connecting bus n from its parent bus π_n, and r_n and x_n denote the resistance and reactance of the corresponding transmission line. Furthermore, p_n and q_n represent active power and reactive power injections at bus n given by

$$\begin{cases} \forall n \in \mathcal{G}\cap\mathcal{B}, & p_n = p_n^g - p_n^d - p_n^c,\ \ q_n = q_n^g - q_n^d, \\ \forall n \in \mathcal{G}\backslash\mathcal{B}, & p_n = q_n^g - p_n^d,\ \ q_n = q_n^g - q_n^d, \\ \forall n \in \mathcal{B}\backslash\mathcal{G}, & p_n = -p_n^c - p_n^d,\ \ q_n = -q_n^d, \\ \forall n \in \mathcal{N}\backslash(\mathcal{G}\cup\mathcal{B}), & p_n = -p_n^d,\ \ q_n = -q_n^d \end{cases} \tag{1e}$$

with (p_n^d, q_n^d) the load power demand at bus $n \in \mathcal{N}$, (p_n^g, q_n^g) the power generation at bus $n \in \mathcal{G}$, and the aggregative charging demand p_n^c at bus $n \in \mathcal{B}$.

The primary objective of the distribution system operator (DSO) is to operate the distribution grid within safe system limits while maximizing revenue by regulating electricity prices. Therefore, the OPF is employed to seek a deterministic solution:

$$\min_{\mathbf{p}^g, \mathbf{q}^g, \boldsymbol{\alpha}} \quad \Delta t \cdot \left(\sum_{n\in\mathcal{G}} c_n \cdot p_n^g - \sum_{n\in\mathcal{B}} \alpha_n \cdot p_n^c \right) \tag{2a}$$

$$\text{subject to } (\mathbf{p}^g, \mathbf{q}^g, \boldsymbol{\alpha}) \in \Omega \quad \text{with } \Omega := \tag{2b}$$

$$\left\{ \begin{pmatrix} \mathbf{p}^g \\ \mathbf{q}^g \\ \boldsymbol{\alpha} \end{pmatrix} \middle| \begin{array}{c} \exists\, \mathbf{P}, \mathbf{Q}, \mathbf{v}, \boldsymbol{\ell}, \mathbf{p}, \mathbf{q} \text{ such that (1) holds} \\ \sum_{k\in\mathcal{C}_0} P_k = p_0,\ \sum_{k\in\mathcal{C}_0} Q_k = q_0, \\ p_n^g \in [\underline{p}_n^g, \overline{p}_n^g],\ q_n^g \in [\underline{q}_n^g, \overline{q}_n^g], \\ v_n \in [\underline{v}_n, \overline{v}_n],\ \ell_n \in [0, \overline{\ell}_n],\ \alpha_n \le [\underline{\alpha}_n, \overline{\alpha}_n] \end{array} \right\}$$

with given (p_n^d, q_n^d) for all $n \in \mathcal{N}$, p_n^c for all $n \in \mathbb{B}$ and the price bound $(\underline{\alpha}_n, \overline{\alpha}_n)$ for all $n \in \mathcal{B}$.

The objective function (2a) is calculated as the difference between generation cost and the income from selling electricity to EVs. Here, coefficient c_n defines the price of generation power, and decision variable α_n defines the price of electricity for EV charge/discharge. The convex set Ω encompasses all power flow equations (1), as well as system limits on power generations $(\mathbf{p}^g, \mathbf{q}^g)$, squared voltage magnitude $\mathbf{v}$ and transmission lines in terms of squared current $\boldsymbol{\ell}$. Note that, in the present paper, we set $c_n, \underline{\alpha}_n > 0$ for all bus n to ensure the conic relaxation (1d) is exact.

Remark 1 (Exact Relaxation [2]) *Given a radial grid, the conic relaxation of the OPF problem in the branch flow model is exact, if the cost function is convexity and strict increase in the line losses.*

B. EV Charging Control Using MPC

Each EV connected to the distributed grid influences the demand side of the system. The primary impact of EVs on the grid comes from their batteries' charging or discharging behavior. In the present paper, we consider the linear battery model, and the battery dynamic model of EV i is given by

$$\mathrm{SoC}_i(k+1) = \beta_i \cdot \mathrm{SoC}_i(k) + \Delta t \cdot u_i(k) \tag{3}$$

with state of charge $\mathrm{SoC}_i(k)$ and charge/discharge rate $u_i(k)$ at time instant t_k, and Δ the charge/discharge interval. Here, constants $\beta_i \in (0, 1]$ represent efficiencies for self-discharge conversion, respectively. Given the recursive structure of the individual system equation (3), the future $\mathrm{SoC}_i(k)$, $k = 1, ..., N$, are determined only by inputs $\mathbf{u}_i(k-1) = [u_i(0), ..., u_i(k-1)]^\top$ and the initial state $\mathrm{SoC}_i(0)$, in particular

$$\mathrm{SoC}_i(k) = \beta_i^k \cdot \mathrm{SoC}_i(0) + \Delta t \cdot \sum_{j=0}^{k-1} \beta^{k-j-1} \cdot u_i(j). \tag{4}$$

Moreover, we consider the economic cost for each time stage

$$L_i(u_i(k), \alpha_n(k)) = \alpha_n(k) \cdot \Delta t \cdot u_i(k) + \sigma_i \cdot u_i(k)^2,$$

which is parametric over the electricity price $\alpha_n(k)$ at time instant t_k. In this scenario, we assume that vehicle $i \in \mathcal{M}_n$

is connected via the charge station at bus n. Therefore, the first term represents the cost of charging if $u_i(k) > 0$, and it is the income from selling electricity to the grid otherwise. The second term penalizes the charge/discharge efforts with $\sigma_i > 0$. We further consider state and input constraints

$$\mathrm{SoC}_i(k) \in [0, C_i], \quad u_i(k) \in [\underline{u}_i, \overline{u}_i], \tag{5}$$

where C_i is the battery capacity, $\underline{u}_i \leq 0$ and $\overline{u}_i \geq 0$ define the bounds of the charging and discharging rates, respectively. To guarantee the recursive feasibility of the MPC scheme, we propose to transform (5) into the cost by using a relaxed logarithmic barrier function $\tilde{b}(\mathrm{SoC}_i(k), u_i(k))$ with weighting parameter $\rho_i > 0$ and

$$\begin{aligned}\tilde{b}_i(\mathrm{SoC}_i(k+1), u_i(k)) := {} & w_{i,1} \cdot b(\mathrm{SoC}_i(k+1)) \\ & + w_{i,2} \cdot [b(C_i - \mathrm{SoC}_i(k+1)) + \ln(C_i)] \\ & + w_{i,3} \cdot [b(u_i(k) - \underline{u}_i(k)) + \ln(-\underline{u}_i(k))] \\ & + w_{i,4} \cdot [b(\overline{u}_i(k) - u_i(k)) + \ln(\overline{u}_i(k))]\end{aligned}$$

with weights $w_{i,k} > 0$, $k = 1, 2, 3, 4$, which are chosen to ensure $\nabla \tilde{b}_i(0, 0) = 0$. Here, function b denotes the relaxed logarithmic barrier function by

$$b(z) = \begin{cases} -\ln(z) & \text{if } z > \delta \\ \dfrac{1}{2}\left(\left(\dfrac{z - 2\delta}{\delta}\right)^2 - 1\right) - \ln(\delta) & \text{if } z \leq \delta \end{cases} \tag{6}$$

with relaxation parameter $\delta > 0$, which is globally twice continuously differentiable and convex. The function $b(\cdot)$ approximates the standard log-barrier function with $\delta \to 0$. More details about $b(\cdot)$ refers to [3, Chapter 3].

Now, we can summarize the relaxed MPC problem into the following unconstrained convex nonlinear program,

$$\forall\, i \in \mathcal{M}_n, \quad \mathbf{u}_i^\star(\boldsymbol{\alpha}_n) := \arg\min_{\mathbf{u}_i} \; V_i(\mathbf{u}_i; \boldsymbol{\alpha}_n) \tag{7}$$

with the objective

$$\begin{aligned}V_i(\mathbf{u}_i; \boldsymbol{\alpha}_n) := {} & \eta_K \cdot (\mathrm{SoC}_i(K) - \mathrm{SoC}_i^c)^2 \\ & + \sum_{k=0}^{K-1} L_i(u_i(k), \alpha_n(k)) + \eta_i \cdot \tilde{b}_i(\mathrm{SoC}_i(k+1), u_i(k))\end{aligned}$$

parametric over the price sequence $\boldsymbol{\alpha}_n = [\alpha_n(0), ..., \alpha_n(N-1)]^\top$, and K the length of prediction horizon. Here, we consider a terminal cost with SoC_i^c the expected SoC of EV i at the end of the prediction horizon, η_i for all $i \in \mathbb{Z}_0^K$ are positive weighting parameters, the solution map $\mathbf{u}_i^\star(\cdot)$ is a function of $\boldsymbol{\alpha}_n$. Here, V_i is also parametric over the initial state $\mathrm{SoC}_i(0)$ but for notation simplification, we ignore it in the arguments of V_i.

Remark 2 *By the definition of L_i, for any given α_n, we have the objective function $V_i(\cdot; \alpha_n)$ of (7) is strongly convex and twice continuously differentiable while its gradient is Lipschitz continuous, i.e., we have $\underline{B} \cdot \mathbf{I}_N \preceq \nabla_u^2 V_i(\mathbf{u}_i; \boldsymbol{\alpha}_n) \preceq \overline{B} \cdot \mathbf{I}_N$ with $\underline{B}, \overline{B} > 0$.*

C. Bilevel Optimization Based Pricing Formulation

In this section, we propose a bilevel optimization-based pricing formulation. For this purpose, we introduce $p_n^c(k)$ as decision variables at the time stage k. These variables are based on the charge/discharge rate of EV i, we consider additional constraints

$$\sum_{i \in \mathcal{M}_n} u_i(k) \leq p_n^c(k), \quad n \in \mathcal{B} \tag{8}$$

in the OPF problem (2), similar to the budget constraint in the resource allocation problem. Moreover, we bound p_n^c to constrain the injection at bus $n \in \mathcal{B}$ by $p_n^c \in [\underline{p}_n^c, \overline{p}_n^c]$.

Next, we rewrite the OPF problem into a compact form,

$$\min_{\mathbf{z}(k), \boldsymbol{\alpha}(k)} \; f(\mathbf{z}(k), \boldsymbol{\alpha}(k), \mathbf{u}(k)) \tag{9a}$$

$$\text{s.t.} \quad (\mathbf{z}(k), \boldsymbol{\alpha}(k)) \in \overline{\Omega}(k), \tag{9b}$$

with objective function

$$\begin{aligned}f(\mathbf{z}(k), \boldsymbol{\alpha}(k), \mathbf{u}(k)) := {} & \Delta t \cdot \sum_{n \in \mathcal{G}} c_n \cdot p_n^g - \\ & \sum_{n \in \mathcal{B}} \left(\alpha_n \cdot \Delta t \cdot p_n^c - \omega_n \cdot b\left(p_n^c(k) - \sum_{i \in \mathcal{M}_n} u_i(k) \right) \right),\end{aligned}$$

which is smooth and nonconvex, and weights $\omega_n > 0$. Here, $\mathbf{z}$ stacks the decision variables $(\mathbf{p}^g, \mathbf{q}^g, \mathbf{p}^c)$, $\mathbf{u}(k)$ stacks all $u_i(k)$, $i \in \mathcal{M} := \bigcup_{n \in \mathcal{B}} \mathcal{M}_n$, and we reformulate the constraint (8) into the objective by using the relaxed barrier function. The main motivation of this constraint reformulation is to transfer the coupling between the upper- and lower-level problems from the constraint to the objective. As a result, the later algorithm design only requires DSO and EVs to exchange the primal information but without dual information. Constraint set $\overline{\Omega} := \Omega \times [\underline{\mathbf{p}}^c, \overline{\mathbf{p}}^c]$ includes the bounds of $\mathbf{p}^c$ into Ω. This results in the following proposition.

Proposition 1 *Constraint set $\overline{\Omega}$ is convex and compact.*

It is obvious that $\overline{\Omega}$ is convex due to the relaxed DistFlow model (1). The compactness holds following the fact that we have box constraints for every entry of $\mathbf{z}$.

Remark 3 *We consider power demand p_n^d, bounds for $\mathbf{p}^g$, $\mathbf{q}^g$, $\mathbf{p}^c$ and $\boldsymbol{\alpha}$ are time-varying such that $\overline{\Omega}(k)$ is time-varying.*

Next, we take the action of EVs into account and construct the following bilevel optimization for real-time pricing in V2G systems,

$$\min_{(\mathbf{Z}, \boldsymbol{\mathcal{A}}) \in \Theta} F(\mathbf{Z}, \boldsymbol{\mathcal{A}}, \mathbf{U}^*) := \sum_{k=0}^{K-1} f(\mathbf{z}(k), \boldsymbol{\alpha}(k), \mathbf{u}^*(k)) \tag{10a}$$

$$\text{subject to } \; \mathbf{U}^* := \arg\min_{\mathbf{U}} \; V(\mathbf{U}; \boldsymbol{\mathcal{A}}) \tag{10b}$$

with $V(\mathbf{U}; \boldsymbol{\mathcal{A}}) := \sum_{i \in \mathcal{M}} V_i(\mathbf{u}_i; \boldsymbol{\alpha}_n)$, where $\mathbf{u}(k)$ and $\mathbf{u}_i^\top$ are the k-th column and i-th row of matrix $\mathbf{U}$, respectively, $\boldsymbol{\alpha}(k)$ and $\boldsymbol{\alpha}_n^\top$ are the k-th column and i-th row of matrix $\boldsymbol{\mathcal{A}}$, $\mathbf{Z} = [\mathbf{z}(0), ..., \mathbf{z}(N-1)]$ stacks $\mathbf{z}(k)$ for all $k \in \mathbb{Z}_0^{N-1}$, and constraint set Θ is given by $\Theta := \overline{\Omega}(0) \times \cdots \times \overline{\Omega}(N-1)$.

Dealing with (10) using the classical KKT-based approach presents significant challenges. This approach requires incorporating the KKT optimality conditions of the lower-level problems into the upper-level problem as additional nonconvex constraints. In particular, managing the resulting nonconvex optimization problem becomes difficult when a

large number of EVs are connected to the distribution grid. To this end, in the following section, we will propose a distributed bilevel optimization to solve (10), which can be easily scaled up.

III. Real-Time Pricing Mechanism for V2G

In this section, we introduce a bilevel optimization algorithm to deal with (10) in a distributed way, and based on this algorithm, a real-time pricing mechanism for V2G is then proposed in a receding horizon fashion.

A. Distributed Bilevel Optimization

The proposed algorithm has two main steps, the first step constructs a hypergradient-based approximate solution map for the lower-level problem and the second step then solves the upper-level problem by substituting the approximate solution into the problem. At κ-th iteration, each EV solves the lower-level MPC problem (10b) to optimal based on the current price $\boldsymbol{\alpha}_n^\kappa$ received from DSO in parallel for all $i \in \mathcal{M}$, i.e.,

$$\mathbf{u}_i^{\kappa+1} := \arg\min_{\mathbf{u}_i} \; V_i(\mathbf{u}_i; \boldsymbol{\alpha}_n^\kappa). \tag{11}$$

Based on the local solution $\mathbf{u}_i^{\kappa+1}$, we can construct the approximation of solution map $\mathbf{u}_i^\star(\cdot)$ by

$$\widetilde{\mathbf{u}}_i^\kappa(\boldsymbol{\alpha}_n) = \mathbf{u}_i^\star(\alpha_n^\kappa) + \frac{\mathrm{d}\mathbf{u}_i^\star(\boldsymbol{\alpha}_n^\kappa)}{\mathrm{d}\boldsymbol{\alpha}_n}(\boldsymbol{\alpha}_n - \boldsymbol{\alpha}_n^\kappa) \tag{12}$$

with $\mathbf{u}_i^\star(\alpha_n^\kappa) = \mathbf{u}_i^{\kappa+1}$ hyper-gradient $\frac{\mathrm{d}\mathbf{u}_i^\star(\boldsymbol{\alpha}_n^\kappa)}{\mathrm{d}\boldsymbol{\alpha}_n}$ given by

$$\frac{\mathrm{d}\mathbf{u}_i^\star(\boldsymbol{\alpha}_n^\kappa)}{\mathrm{d}\boldsymbol{\alpha}_n} = -\left[\frac{\partial^2 V_i(\mathbf{u}_i^{\kappa+1}; \boldsymbol{\alpha}_n^\kappa)}{\partial \mathbf{u}_i^2}\right]^{-1}\left[\frac{\partial^2 V_i(\mathbf{u}_i^{\kappa+1}; \boldsymbol{\alpha}_n^\kappa)}{\partial \mathbf{u}_i \partial \boldsymbol{\alpha}_n}\right].$$

Here, the Hessian matrix is always invertible as V_i is strongly convex and twice continuously differentiable over $\mathbf{u}_i$. Once the DSO receives $\widetilde{\mathbf{U}}^\kappa(\cdot)$ from all EVs, the upper-level problem is solved

$$(\mathbf{Z}^{\kappa+1}, \boldsymbol{\mathcal{A}}^{\kappa+1}) := \operatorname*{argmin}_{(\mathbf{Z},\boldsymbol{\mathcal{A}})\in\Theta} \; \widetilde{F}_\rho^\kappa(\mathbf{Z}, \boldsymbol{\mathcal{A}}, \widetilde{\mathbf{U}}^\kappa(\boldsymbol{\mathcal{A}})), \tag{13}$$

where the objective is given by the quadratic function $\widetilde{F}_\rho^\kappa(\mathbf{Z}, \boldsymbol{\mathcal{A}}, \widetilde{\mathbf{U}}^\kappa(\boldsymbol{\mathcal{A}})) :=$

$$\langle \nabla_{\mathbf{Z}} F(\mathbf{Z}^\kappa, \boldsymbol{\mathcal{A}}^\kappa, \widetilde{\mathbf{U}}^\kappa(\boldsymbol{\mathcal{A}}^\kappa)), \mathbf{Z} - \mathbf{Z}^\kappa\rangle + \frac{\rho}{2}\|\mathbf{Z} - \mathbf{Z}^\kappa\|_{\mathrm{F}}^2$$
$$+\langle \nabla_{\boldsymbol{\mathcal{A}}} F(\mathbf{Z}^\kappa, \boldsymbol{\mathcal{A}}^\kappa, \widetilde{\mathbf{U}}^\kappa(\boldsymbol{\mathcal{A}}^\kappa)), \boldsymbol{\mathcal{A}} - \boldsymbol{\mathcal{A}}^\kappa\rangle + \frac{\rho}{2}\|\boldsymbol{\mathcal{A}} - \boldsymbol{\mathcal{A}}^\kappa\|_{\mathrm{F}}^2.$$

with penalty parameter $\rho > 0$ and the hyper-gradient $\nabla_{\boldsymbol{\mathcal{A}}} F(\mathbf{Z}^\kappa, \boldsymbol{\mathcal{A}}^\kappa, \widetilde{\mathbf{U}}^\kappa(\boldsymbol{\mathcal{A}}^\kappa))$ is given by

$$\partial_{\boldsymbol{\mathcal{A}}} F(\mathbf{Z}^\kappa, \boldsymbol{\mathcal{A}}^\kappa, \widetilde{\mathbf{U}}^\kappa)^\top + \left\langle \frac{\mathrm{d}\widetilde{\mathbf{U}}^\kappa(\boldsymbol{\mathcal{A}}^\kappa)}{\mathrm{d}\boldsymbol{\mathcal{A}}}, \partial_{\mathbf{U}} F(\mathbf{Z}^\kappa, \boldsymbol{\mathcal{A}}^\kappa, \widetilde{\mathbf{U}}^\kappa)^\top \right\rangle.$$

Lemma 1 *Solving Problem* (13) *is equivalent to one step of the projected gradient descent method with step size* $1/\rho$ *for solving the original problem*

$$\min_{(\mathbf{Z},\boldsymbol{\mathcal{A}})\in\Theta} \; \Phi(\mathbf{Z}, \boldsymbol{\mathcal{A}}) := F(\mathbf{Z}, \boldsymbol{\mathcal{A}}, \mathbf{U}^\star(\boldsymbol{\mathcal{A}})). \tag{14}$$

This follows the fact of $\mathrm{d}\widetilde{\mathbf{U}}^\kappa(\boldsymbol{\mathcal{A}}^\kappa)/\mathrm{d}\mathbf{A} = \mathrm{d}\mathbf{U}^\star(\boldsymbol{\mathcal{A}}^\kappa)/\mathrm{d}\mathbf{A}$ given by (12). To analyze the convergence of the proposed algorithm, we introduce the equivalent projected gradient step following Lemma 1 using the notation

$$\begin{bmatrix}\mathbf{Z}^{\kappa+1}\\ \boldsymbol{\mathcal{A}}^{\kappa+1}\end{bmatrix} = \mathrm{Proj}_\Theta\left(\begin{bmatrix}\mathbf{Z}^\kappa - 1/\rho \cdot \nabla_{\mathbf{Z}}\Phi(\mathbf{Z}^\kappa, \boldsymbol{\mathcal{A}}^\kappa)\\ \boldsymbol{\mathcal{A}}^\kappa - 1/\rho \cdot \nabla_{\boldsymbol{\mathcal{A}}}\Phi(\mathbf{Z}^\kappa, \boldsymbol{\mathcal{A}}^\kappa)\end{bmatrix}\right) \tag{15}$$

Now, we can summarize the convergence result as follows.

Theorem 1 *Let* ρ *chosen satisfying* $\rho > \tau/2$. *If one execute* $\overline{\kappa}$ *iterations of* (11) *and* (13) *to deal with bilevel optimization problem* (10)*, there exists a constant* $\mu > 0$ *such that the following inequality holds,*

$$\frac{1}{\overline{\kappa}}\sum_{\kappa=0}^{\overline{\kappa}-1}\left\|\widehat{\nabla}\Phi(\mathbf{Z}^\kappa, \boldsymbol{\mathcal{A}}^\kappa)\right\|^2 \le \frac{\Phi(\mathbf{Z}^0, \boldsymbol{\mathcal{A}}^0) - \Phi^*}{\mu \cdot \overline{\kappa}} \tag{16}$$

with $\Phi^* = \min_{(\mathbf{Z},\boldsymbol{\mathcal{A}})\in\Theta}\Phi(\mathbf{Z}, \boldsymbol{\mathcal{A}})$ *and* $\widehat{\nabla}\Phi(\mathbf{Z}^\kappa, \boldsymbol{\mathcal{A}}^\kappa) :=$

$$\rho \cdot \left(\mathrm{Proj}_\Theta\left(\begin{bmatrix}\mathbf{Z}^\kappa - 1/\rho \cdot \nabla_{\mathbf{Z}}\Phi(\mathbf{Z}^\kappa, \boldsymbol{\mathcal{A}}^\kappa)\\ \boldsymbol{\mathcal{A}}^\kappa - 1/\rho \cdot \nabla_{\boldsymbol{\mathcal{A}}}\Phi(\mathbf{Z}^\kappa, \boldsymbol{\mathcal{A}}^\kappa)\end{bmatrix}\right) - \begin{bmatrix}\mathbf{Z}^\kappa\\ \boldsymbol{\mathcal{A}}^\kappa\end{bmatrix}\right)$$

the generalized projected gradient at κ*-th iteration [11].*

The proof of Theorem 1 can be found in the appendix. This theorem implies that the iteration $(\mathbf{Z}^\kappa, \boldsymbol{\mathcal{A}}^\kappa)$ converges to a stationary point $(\mathbf{Z}^s, \boldsymbol{\mathcal{A}}^s)$ with

$$s := \arg\min_\kappa \; \left\|\widehat{\nabla}\Phi(\mathbf{Z}^\kappa, \boldsymbol{\mathcal{A}}^\kappa)\right\|^2$$

for the constrained bilevel optimization problem (10), whose generalized projected gradient norm sublinearly decays. Notice that we adopt the stationary point as the convergence criterion here due to the nonconvexity of the upper-level objective function Φ.

B. Receding Horizon Based Real-Time Pricing

Algorithm 1 outlines the proposed real-time pricing scheme. It is executed in a receding horizon fashion due to the decision-making of the lower-level MPC-based EV charging. One can see that the main idea of Algorithm 1 is approximately solving the bilevel pricing problem (10) at each time stage using one iteration of the proposed algorithm in Section III-A and then, shift the horizon with applying the first control input of the lower-level MPC solution.

Algorithm 1 Real-Time Pricing Scheme for V2G

Initialization:

- DSO chooses $\rho > 0$ such that $\rho > \tau/2$ and set $\mathrm{cl} = 0$;
- initializes a feasible guess $(\mathbf{Z}^{\mathrm{cl}}, \boldsymbol{\mathcal{A}}^{\mathrm{cl}}) \in \Theta^{\mathrm{cl}}$.

Online:

1) *Lower-level phase:* EVs connected to the distribution grid in parallel,
   a) based on the current price at the charge station and the current SoC, solve
   $$\mathbf{U}^{\mathrm{cl}} = \arg\min_{\mathbf{U}} \; \mathcal{V}(\mathbf{U}; \boldsymbol{\mathcal{A}}^{\mathrm{cl}}) := \sum_{i\in\mathcal{M}} V_i(\mathbf{u}_i; \boldsymbol{\alpha}_n^{\mathrm{cl}});$$
   b) construct the approximate solution map $\widetilde{\mathbf{U}}^{\mathrm{cl}}(\boldsymbol{\mathcal{A}})$ and send it to DSO;
   c) apply $u_i^*(0)$ the first optimal input to charge or discharge in one slot.
2) *Upper-level phase:* DSO deals with
   $$(\mathbf{Z}^{\mathrm{cl}+1}, \boldsymbol{\mathcal{A}}^{\mathrm{cl}+1}) := \operatorname*{argmin}_{(\mathbf{Z},\boldsymbol{\mathcal{A}})\in\Theta^{\mathrm{cl}}} \; \widetilde{F}_\rho^{\mathrm{cl}}(\mathbf{Z}, \boldsymbol{\mathcal{A}}, \widetilde{\mathbf{U}}^{\mathrm{cl}}(\boldsymbol{\mathcal{A}}))$$
   to update the price and set $\mathrm{cl} \leftarrow \mathrm{cl} + 1$.

In Algorithm 1, DSO needs to start with a feasible point, i.e., $(\mathbf{Z}^0, \boldsymbol{\mathcal{A}}^0) \in \Theta^0$. This could be easily achieved by first initializing a feasible price for a time window with length

$K > 0$ and then, dealing with the OPF problem with fixed prices and ignoring the EVs. During the closed-loop, EVs connected with the distribution grid first deal with the lower-level MPC problem (11) based on the current initial value of SoC and the current price sequence $\boldsymbol{\alpha}_n$ at all bus $n \in \mathcal{B}$. In practice, this unconstrained convex nonlinear program could be efficiently solved by using the Newton-type method. The solution is then used to construct the map $\widetilde{\mathbf{U}}^{\mathrm{cl}}(\cdot)$, where cl is the closed-loop index. At the same, the first control input $u_i^*(0)$ is applied to the process as a fixed charge/discharge rate for a time slot Δt. As the pricing mechanism is essentially a leader-follower problem, the DSO recomputes the shifted price by solving (13) after it receives the local solution map approximation. Here, the constraint set Θ^{cl} is adapted only such that it is time-varying.

IV. Numerical Case Study

In this section, we validate the proposed algorithms using benchmark data and conduct analyses of the algorithm's performance in the closed-loop setting.

A. Experiment Setting

The distribution grid model is built based on the IEEE 15-bus benchmark using `MATPOWER 7.1` toolbox, where a generator is situated at the root node of the power grid. At the upper level, generation coefficients are set to $c_n = 0.2\ \$/\mathrm{kW}$ for all $n \in \mathcal{G}$, and the electricity price is bounded as $\alpha \in [0.5\,\$/\mathrm{kW}\cdot\mathrm{h}, 1\ \$/\mathrm{kW}\cdot\mathrm{h}]$. In the lower-level MPC problem, the parameter of the logarithmic barrier function is $\delta = 0.0001$. The capacity of each EV is given by $C_i = 100\,\mathrm{kW}\cdot\mathrm{h}$. The initial SoC and expected SoC_i^e are randomly initialized. We conduct tests with different horizon lengths with $K \in \{3, 5, 10\}$ and different numbers of EVs $|\mathcal{M}| \in \{10, 30, 100\}$. For the open-loop configuration, we assume the initial value of the SoC for the EV remains constant in each iteration. Then we analyze the convergence of the algorithm. For the closed-loop configuration, the EV at the lower level adjusts its behavior based on electricity prices and modifies its SoC accordingly, implying that the initial state differs in each iteration. We utilize `CAsaDi` to solve the OPF problem as well as the MPC problem. Moreover, we execute our experiment with different prediction horizons, $K = 5$ and $K = 10$.

B. Closed-Loop Performance

In this section, we evaluate our model within a closed-loop framework under various horizon settings. Specifically, we operate under the assumption that across the entire distribution grid, all charging stations adhere to a uniform electricity pricing structure at each time interval. Our experiments simulate a practical plug-and-play scenario, accommodating the dynamic nature of electric vehicle (EV) interactions—where certain EVs are compelled to depart at predetermined times while others arrive at the charging stations at specific moments. Here, we consider 10 EVs connected to the IEEE 15-bus benchmark grid via 3 charge stations.

Figure 1 provides a comprehensive view of the collective power consumption by electric vehicles (EVs) at charging stations. Figure 2 and 3 show the operation costs for the DSO and the corresponding pricing strategies over the same timeframe, with two predictive horizons, respectively. These two plots illustrate how costs and pricing fluctuate in tandem with power usage patterns. This reflects a dynamic pricing model that adjusts in real-time to demand and supply changes, a crucial mechanism for maintaining grid balance and promoting energy consumption when it is most abundant and least costly. Additionally, the pattern of these fluctuations mirrors the dynamics of a Stackelberg game, where, at the point of dynamic equilibrium, neither the DSO nor the consumers reach an optimal point, highlighting the complex interplay between supply, demand, and pricing strategies. These fluctuations in cost and pricing harmonize with the observed trends in power consumption, showcasing a closed-loop pricing scheme that skillfully adapts to immediate demand-side dynamics.

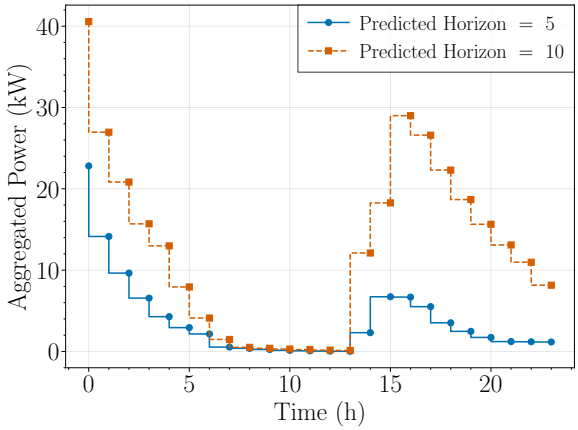

Fig. 1. Sum of aggregated power at charge stations

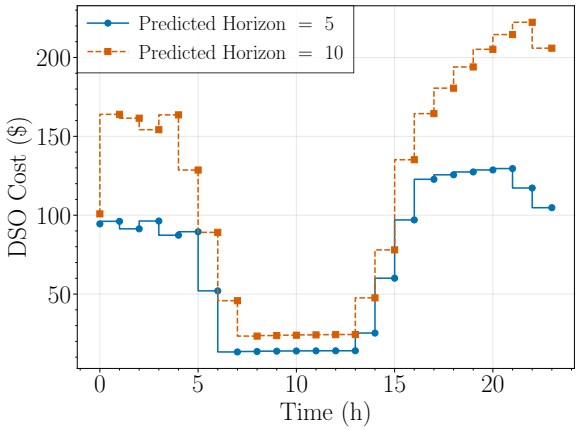

Fig. 2. Operation cost of distribution system

V. Conclusion

This study addresses the challenges of integrating EVs into the distribution grid, highlighting the limitations of conventional pricing mechanisms in accommodating EVs as both consumers and suppliers, as well as energy storage units. We propose an innovative real-time pricing strategy using distributed bilevel optimization, clearly defining the

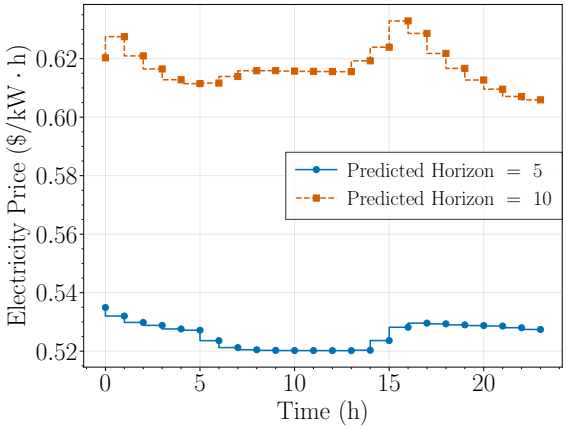

Fig. 3. Real-time electricity price

roles of the DSO and EVs, with both aiming to maximize profits. Our model incorporates power flow constraints at the DSO level for efficient and secure grid operation and uses MPC to precisely manage EV charging and discharging. The approach's effectiveness and robustness are demonstrated through detailed convergence analysis.

Appendix

As the objective $\Phi(\mathbf{Z}, \boldsymbol{\mathcal{A}})$ includes three terms, a linear term, a bilinear term, and a relaxed barrier term such that it is twice continuously differentiable and its gradient is Lipschitz, i.e., there exists a constant $\tau > 0$ such that

$$\begin{aligned}&\Phi(\mathbf{Z}^{\kappa+1}, \boldsymbol{\mathcal{A}}^{\kappa+1}) - \Phi(\mathbf{Z}^{\kappa}, \boldsymbol{\mathcal{A}}^{\kappa}) \qquad (17)\\ &\leq \langle \nabla_{\mathbf{Z}} \Phi(\mathbf{Z}^{\kappa}, \boldsymbol{\mathcal{A}}^{\kappa}), \mathbf{Z}^{\kappa+1} - \mathbf{Z}^{\kappa} \rangle + \frac{\tau}{2} \left\| \mathbf{Z}^{\kappa+1} - \mathbf{Z}^{\kappa} \right\|^2 + \\ &\langle \nabla_{\boldsymbol{\mathcal{A}}} \Phi(\mathbf{Z}^{\kappa}, \boldsymbol{\mathcal{A}}^{\kappa}), \boldsymbol{\mathcal{A}}^{\kappa+1} - \boldsymbol{\mathcal{A}}^{\kappa} \rangle + \frac{\tau}{2} \left\| \boldsymbol{\mathcal{A}}^{\kappa+1} - \boldsymbol{\mathcal{A}}^{\kappa} \right\|^2.\end{aligned}$$

Then, following Lemma 1, the nonexpansiveness of projection operation yields

$$\left\langle \begin{bmatrix} \mathbf{Z}^{\kappa} - 1/\rho \cdot \nabla_{\mathbf{Z}} \Phi(\mathbf{Z}^{\kappa}, \boldsymbol{\mathcal{A}}^{\kappa}) - \mathbf{Z}^{\kappa+1} \\ \boldsymbol{\mathcal{A}}^{\kappa} - 1/\rho \cdot \nabla_{\boldsymbol{\mathcal{A}}} \Phi(\mathbf{Z}^{\kappa}, \boldsymbol{\mathcal{A}}^{\kappa}) - \boldsymbol{\mathcal{A}}^{\kappa+1} \end{bmatrix}, \begin{bmatrix} \mathbf{Z}^{\kappa} - \mathbf{Z}^{\kappa+1} \\ \boldsymbol{\mathcal{A}}^{\kappa} - \boldsymbol{\mathcal{A}}^{\kappa+1} \end{bmatrix} \right\rangle \leq 0,$$

which can be rewritten as

$$\left\langle \begin{bmatrix} \nabla_{\mathbf{Z}} \Phi(\mathbf{Z}^{\kappa}, \boldsymbol{\mathcal{A}}^{\kappa}) \\ \nabla_{\boldsymbol{\mathcal{A}}} \Phi(\mathbf{Z}^{\kappa}, \boldsymbol{\mathcal{A}}^{\kappa}) \end{bmatrix}, \begin{bmatrix} \mathbf{Z}^{\kappa} - \mathbf{Z}^{\kappa+1} \\ \boldsymbol{\mathcal{A}}^{\kappa} - \boldsymbol{\mathcal{A}}^{\kappa+1} \end{bmatrix} \right\rangle \geq \rho \left\| \begin{bmatrix} \mathbf{Z}^{\kappa} - \mathbf{Z}^{\kappa+1} \\ \boldsymbol{\mathcal{A}}^{\kappa} - \boldsymbol{\mathcal{A}}^{\kappa+1} \end{bmatrix} \right\|^2.$$

Substituting this inequality and (15) into (17) yields

$$\Phi(\mathbf{Z}^{\kappa+1}, \boldsymbol{\mathcal{A}}^{\kappa+1}) - \Phi(\mathbf{Z}^{\kappa}, \boldsymbol{\mathcal{A}}^{\kappa}) \leq \left(\frac{\tau}{2 \cdot \rho^2} - \frac{1}{\rho} \right) \left\| \widehat{\nabla} \Phi(\mathbf{Z}^{\kappa}, \boldsymbol{\mathcal{A}}^{\kappa}) \right\|^2.$$

As ρ is chosen with $\rho > \tau/2$, we define $\mu = \frac{1}{\rho} - \frac{\tau}{2 \cdot \rho^2} > 0$ such that conducting the telescoping for the inequality above yields

$$\begin{aligned}\mu \cdot \sum_{\kappa=0}^{\overline{\kappa}-1} \left\| \widehat{\nabla} \Phi(\mathbf{Z}^{\kappa}, \boldsymbol{\mathcal{A}}^{\kappa}) \right\|^2 &\leq \Phi(\mathbf{Z}^{0}, \boldsymbol{\mathcal{A}}^{0}) - \Phi(\mathbf{Z}^{\overline{\kappa}}, \boldsymbol{\mathcal{A}}^{\overline{\kappa}}) \\ &\leq \Phi(\mathbf{Z}^{0}, \boldsymbol{\mathcal{A}}^{0}) - \Phi^{*}.\end{aligned}$$

Then, both side times $\overline{\kappa}/\mu$ concludes the proof.

References

[1] H. Allcott. Rethinking real-time electricity pricing. *Resource and Energy Economics*, 33(4):820–842, 2011.

[2] M. Farivar and S. H. Low. Branch flow model: Relaxations and convexification—part i. *IEEE Trans. Power Syst.*, 28(3):2554–2564, 2013.

[3] C. Feller. *Relaxed barrier function based model predictive control*. PhD thesis, University of Stuttgart, 2017.

[4] F. Garcia-Torres, D. G. Vilaplana, C. Bordons, P. Roncero-Sánchez, and M. A. Ridao. Optimal management of microgrids with external agents including battery/fuel cell electric vehicles. *IEEE Trans. on Smart Grid*, 10(4):4299–4308, 2019.

[5] H. R. Gholinejad, J. Adabi, and M. Marzband. Hierarchical energy management system for home-energy-hubs considering plug-in electric vehicles. *IEEE Trans Ind. Informat.*, 58(5):5582–5592, 2022.

[6] R. Grazzi, L. Franceschi, M. Pontil, and S. Salzo. On the iteration complexity of hypergradient computation. In *Proc. Int. Conf. Mach. Learn.*, pages 3748–3758. PMLR, 2020.

[7] P. Hansen, B. Jaumard, and G. Savard. New branch-and-bound rules for bilevel linear programming. *SIAM Journal on Scientific and Statistical Computing*, 13:273–, 09 1992.

[8] Q. Huang, Q.-S. Jia, Z. Qiu, X. Guan, and G. Deconinck. Matching ev charging load with uncertain wind: A simulation-based policy improvement approach. *IEEE Trans. on Smart Grid*, 6(3):1425–1433, 2015.

[9] H. Jahangir, S. S. Gougheri, B. Vatandoust, M. A. Golkar, A. Ahmadian, and A. Hajizadeh. Plug-in electric vehicle behavior modeling in energy market: A novel deep learning-based approach with clustering technique. *IEEE Trans. on Smart Grid*, 11(6):4738–4748, 2020.

[10] K. Ji, J. Yang, and Y. Liang. Bilevel optimization: Convergence analysis and enhanced design. In *Proc. Int. Conf. Mach. Learn.*, pages 4882–4892. PMLR, 2021.

[11] K. Ji and L. Ying. Network utility maximization with unknown utility functions: A distributed, data-driven bilevel optimization approach. *arXiv preprint arXiv:2301.01801*, 2023.

[12] W. Kempton and J. Tomić. Vehicle-to-grid power implementation: From stabilizing the grid to supporting large-scale renewable energy. *Journal of power sources*, 144(1):280–294, 2005.

[13] W. Kempton and J. Tomić. Vehicle-to-grid power implementation: From stabilizing the grid to supporting large-scale renewable energy. *J. Power Sources*, 144(1):280–294, 2005.

[14] P. Kou, D. liang Liang, L. Gao, and F. Gao. Stochastic coordination of plug-in electric vehicles and wind turbines in microgrid: A model predictive control approach. *IEEE Trans. on Smart Grid*, 7:1537–1551, 2016.

[15] R. Liu, X. Liu, S. Zeng, J. Zhang, and Y. Zhang. Value-function-based sequential minimization for bi-level optimization. *IEEE Trans. Pattern Anal. Mach. Intell.*, 45(12):15930–15948, 2023.

[16] J. Lorraine, P. Vicol, and D. Duvenaud. Optimizing millions of hyperparameters by implicit differentiation. In *Proc. Int. Conf. Artif. Intell. Statist.,*, pages 1540–1552. PMLR, 2020.

[17] T. Namerikawa, N. Okubo, R. Sato, Y. Okawa, and M. Ono. Real-time pricing mechanism for electricity market with built-in incentive for participation. *IEEE Trans. on Smart Grid*, 6(6):2714–2724, 2015.

[18] W. Shi and V. W. Wong. Real-time vehicle-to-grid control algorithm under price uncertainty. In *2011 IEEE Int. Conf. Smart Grid Commun.*, pages 261–266, 2011.

[19] Y. Shi, L. Lian, Y. Shi, Z. Wang, Y. Zhou, L. Fu, L. Bai, J. Zhang, and W. Zhang. Machine learning for large-scale optimization in 6g wireless networks. *IEEE Commun. Surveys Tuts.*, 25(4):2088–2132, 2023.

[20] M. A. Velasquez, J. Barreiro-Gomez, N. Quijano, A. I. Cadena, and M. Shahidehpour. Intra-hour microgrid economic dispatch based on model predictive control. *IEEE Trans. on Smart Grid*, 11(3):1968–1979, 2020.

[21] S. Wang, S. Bi, and Y. A. Zhang. Reinforcement learning for real-time pricing and scheduling control in ev charging stations. *IEEE Trans. on Industrial Informatics*, 17(2):849–859, 2021.

[22] Z. Wang and F. Li. Critical peak pricing tariff design for mass consumers in great britain. In *2011 IEEE Power and Energy Society General Meeting*, pages 1–6, 2011.

[23] D. Xie, M. Liu, L. Xu, and W. Lu. Multiplayer nash–stackelberg game analysis of electricity markets with the participation of a distribution company. *IEEE Syst. J*, 17(3):3658–3669, 2023.

[24] L. Yan, X. Chen, J. Zhou, Y. Chen, and J. Wen. Deep reinforcement learning for continuous electric vehicles charging control with dynamic user behaviors. *IEEE Trans. on Smart Grid*, 12(6):5124–5134, 2021.

[25] M. Yilmaz and P. T. Krein. Review of the impact of vehicle-to-grid technologies on distribution systems and utility interfaces. *IEEE Trans. Power Electron.*, 28(12):5673–5689, 2012.

[26] F. Zhang, Q. Yang, and D. An. Cddpg: a deep-reinforcement-learning-based approach for electric vehicle charging control. *IEEE IoT J.*, 8(5):3075–3087, 2020.

[27] Y. Zhou, F. Gao, S. Ci, Y. Yang, and Y. Xu. Time-of-use pricing in retail electricity market: Step tariff vs. usage-based schemes. In *2016 Int. Conf. Probabilistic Methods Applied to Power Syst.*, pages 1–5, 2016.