Robust Optimal Control of Electric Vehicles Charging for Stochastic and Differentially Private Demand

Tong Wu[1], Ravi Nikhil[1], Anna Scaglione[1], Sean Peisert[1], and Daniel Arnold[1]

[1]Affiliation not available

March 14, 2024

Robust Optimal Control of Electric Vehicles Charging for Stochastic and Differentially Private Demand

Tong Wu, *Member, IEEE*, Nikhil Ravi, *Student Member, IEEE*, Anna Scaglione, *Fellow, IEEE*, Sean Peisert, *Senior Member, IEEE*, Daniel Arnold, *Member, IEEE*,

***Abstract*—This paper presents a comprehensive stochastic optimization model that seamlessly integrates aggregate electric vehicle (EV) charging demand response with power grid system operations, leveraging the inherent flexibility of EV charging. Our main novel contribution is tackling the problem of uncertainty in the demand characteristics. In our stochastic model, we capture not only unknown user charging patterns but also the effect of a pseudo-randomized mechanism applied to provide differential privacy (DP) guarantees to users whose charging patterns are not disclosed. From a control perspective, the intrinsic randomness of the users charging needs, compounded with randomness introduced by the DP mechanism can easily result in infeasible solutions. To overcome this challenge, we adopt a robust optimal control strategy that encompasses the intersection of potential sampling scenario-based constraints. In addition, to manage the high-dimension of the control action space, we approximate the intersections of the feasible regions with a reduced set of polyhedron constraints. In conclusion, our case studies based on IEEE standard systems demonstrate that the proposed algorithm effectively addresses robustness, scalability, and differential privacy for EV users by dynamically adapting to control the demand response for renewable energy integration while consistently ensuring the privacy of EV drivers.**

***Index Terms*—EV Charging Control, Differential Privacy, Robust Control, Convex Approximation.**

I. Introduction

A. Background and Motivation

The adoption of electric vehicles (EVs) into our transportation system is a critical step toward achieving carbon-neutral transportation and significantly reducing greenhouse gas emissions that contribute to climate change [2]. Central to this goal is the alignment of EV charging practices with the generation of power from renewable distributed energy resources (DERs). However, renewable DERs are inherently non-dispatchable [3], which limits flexibility in power production. Nonetheless, EV charging, particularly at scale, offers a considerable degree of spatiotemporal flexibility in terms of both charging locations and time [4]. Exploiting this flexibility effectively requires addressing two main challenges: the uncertainty in EV charging demand and the need to protect EV users' privacy [5].

The preliminary version has been invited to IEEE Smartgridcomm 2023 [1]. Tong Wu, Nikhil Ravi, and Anna Scaglione are with the Department of Electrical and Computer Engineering, Cornell University, 10044 USA (e-mail: {tw385, as337}@cornell.edu). Daniel Arnold and Sean Peisert are with Lawrence Berkeley National Laboratory. This research was supported in part by the NSF under Grant NSF ECCS # 2210012, in part by the Director, Cybersecurity, Energy Security, and Emergency Response (CESER), Privacy-Preserving, Collective Cyberattack Defense of DERs (SHIELDERs) of the U.S. Department of Energy, under contract DE-AC02-05CH11231 and in part by the CESER, Risk Management and Tools and Technologies (RMT) program of the U.S. Department of Energy via Mitigation via Analytics for Grid-Inverter Cybersecurity (MAGIC) project under contract DE-AC02-05CH11231. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect those of the sponsors of this work.

To mitigate privacy concerns, this paper advocates the adoption of differential privacy (DP) [6], a methodology that provides theoretical bounds on the potential for information leakage when disseminating aggregate data. This approach ensures the visibility of broad trends without compromising individual data points, by introducing pseudo-random noise into the actual data values (typically following Laplacian or Gaussian distributions). Even with minimal alterations to individual records – insufficient to change crucial statistics – this method of randomization restricts the ability to infer specifics about any individual [7]. However, the challenge does not end there. The intrinsic uncertainty of future EV charging demand, when combined with the introduction of a DP mechanism, increases the risk of making infeasible decisions. Tackling this dual challenge of managing demand uncertainty while ensuring user privacy through DP is the primary focus of this paper.

B. Related Work

In this section, we review related work in three areas: the integration of EVs with demand response (DR) mechanisms, the incorporation of DP in EV charging strategies, and the application of robust optimization techniques in EV charging control.

a) Control of EVs with Demand Response: The domain of EV charging integrated with DR mechanisms is mainly classified into two research categories. The first explores the collective potential of EVs within the wholesale market, highlighting the flexibility, aggregate energy demand, and ramping capabilities of DR device groups [8, 9]. The second category [10, 11] investigates the synergy between real-time pricing and DR, focusing on households equipped with appliances (such as air conditioners, refrigerators, etc), EVs, and batteries. These studies propose a utility-maximization framework for DR that exploits the benefits of appliance use with power consumption, encouraging households to optimize power usage within certain thresholds. The dynamic nature of pricing synchronizes individual and collective gains, steering demand responses toward optimized system efficiency.

However, leveraging EVs' flexibility at a large scale introduces complexities. Traditional control methods, which have relied on predictive data such as expected EV arrivals, fall short in real-time EV management [12]. Deep reinforcement learning (DRL) emerges as a promising alternative, enabling decision-making in uncertain environments through neural networks [13]. By integrating Vehicle-to-Grid (V2G) technology with DRL, there is a potential to enhance both EV charging and power production, thereby supporting carbon-neutral transportation systems.

b) Differential Privacy in EV Charging: The exploration of DP in EV charging is nascent, with key studies focusing

on modifying the charging rate with DP noise and exploring distributed constrained optimization for EV charging under DP [14]. Recent advancements include an adaptive DP federated learning framework for EV charging infrastructures [15], aiming to detect anomalous traffic and enhance the privacy provisioning mechanism. However, these methods often neglect broader privacy issues related to user behavior patterns, including charging times, deadlines, capacity, and initial states of EVs. Exposure of this data could reveal detailed information about users' lifestyles and preferences. Moreover, it is important to recognize that many charging stations function at a fixed charging ratio, possibly diminishing the need for these existing DP approaches that fail to protect behavior patterns. Furthermore, the challenge of integrating existing DP approaches with distribution networks remains unaddressed, raising concerns about the impact of adding Laplacian noise on adhering to power flow constraints [16].

c) Robust Optimization: Incorporating DP into optimization problems introduces uncertainties that may challenge the applicability of traditional optimization methods. Robust optimization methodologies address this by categorizing uncertainties into adversarial robustness and scenario-based security constraints. Adversarial robustness involves an iterative minimax optimization process, aiming to identify resilient dispatch strategies against stochastic perturbations [17, 18] However, the identification of optimal attack policies and the potential of overly conservative strategies pose significant challenges. On the other hand, scenario-based security constraints, prioritize the feasibility of optimal actions across a wide range of scenarios, enhancing adaptability to real-world conditions. Despite their promise, these methods face computational challenges and the complexity of managing extensive scenario analyses [19–21].

C. Contributions and Organization

To navigate the challenges presented by current approaches, this paper introduces novel strategies for safe control of EV charging that incorporates DP. These strategies enhance the privacy of EV charging data while leveraging the intrinsic flexibility of aggregated EV charging models. Additionally, we apply constrained reinforcement learning to tackle the complexities of dynamic stochastic programming. The core contributions of this paper are as follows:

- An integrated control model is developed to harmonize aggregate EV charging DR with power grid system operations, utilizing the flexibility of EV charging to facilitate renewable energy integration.
- A DP mechanism is designed to protect the privacy of information related to EV charging events such as arrivals. This mechanism ensures ϵ-DP, with minimal impact on the optimality of the system's operation.
- To address the randomness introduced by the DP mechanism, we propose a method that incorporates the intersection of sampling scenario-based constraints, encapsulating a broad spectrum of possible sampling scenarios. The intersections of these sampling scenarios can encompass real-world unseen scenarios, thereby ensuring feasibility.
- Recognizing the high dimensionality of the control action space resulting from numerous potential scenarios, we introduce a technique that approximates this space with convex polyhedron constraints. This simplification facilitates the use of DRL methods by reducing the complexity of the action space.

The structure of the paper is as follows: Section II discusses aggregated EV charging models. Section III introduces the concept of DP, detailing our specific mechanism designed to safeguard the privacy of EV charging data. Section IV presents the V2G model, which is central to our approach for controlling EV charging and discusses our robust control methodology tailored for DP-enhanced EV models, aiming to address the challenges of privacy and system efficiency simultaneously. In Section V, we review the constrained reinforcement learning method, laying the groundwork for our control strategies. Section VI showcases the application of the primal-dual constrained reinforcement learning method within a case study, illustrating the implementation and benefits of our approach. The paper concludes with Section VII.

II. Aggregate EV Charging Model

In this section, we introduce how the charging flexibility of a collection of EVs can be depicted using a linear model, building on the concepts from [22, 23]. In the context of DR, the models for EV charging reveal considerable flexibility, presenting an array of energy consumption patterns.

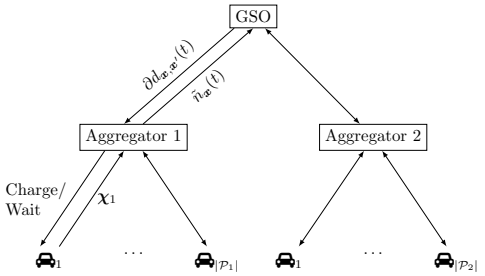


Figure 1. The three-tier system architecture. In this architecture, the users share their state with their aggregator, who then passes on differentially private arrival counts to the GSO. The GSO performs robust control optimization and communicates the optimal control actions back to the aggregators.

A. System Architecture

In this paper, we consider a three-tier infrastructure for the EV charge scheduling problem, which focuses on dynamic interactions among three key stakeholders within the system. Fig. 1 shows that this infrastructure includes the Grid Side Operator (GSO), aggregators, and the EV owners (or the users/consumers), each driven by distinct yet interrelated objectives:

1) The GSO serves as the orchestrator, aiming to minimize the power generation cost and ensure efficient load management. To achieve these goals, strategic dispatch control is utilized, focusing on harnessing renewable energy generation for the charging of EVs.
2) Aggregators, intermediaries between the GSO and EV owners, aim to allocate loads based on factors such as charging demand, user preferences, and locations.
3) EV owners prioritize the seamless availability of cost-effective fast charging. Significantly, our framework

uniquely addresses user privacy concerns, safeguarding sensitive information such as arrival times, charging requirements, and departure deadlines – a novel aspect exclusive to our approach.

In Fig. 1, the system is illustrated with two aggregators. Upon arrival, users share their attributes (defined in the next subsection) with their respective aggregators. Subsequently, each aggregator calculates the differentially private total number of EVs in state $\boldsymbol{x}$ (described in Section III) and forwards this information to the GSO. The GSO, in turn, conducts robust control optimization (formulated in Section IV) and communicates the resulting control actions back to the aggregators. Following this, each aggregator employs its scheduling strategy to determine which EVs should charge in that specific interval.

B. EV Charging Model

Before diving into the details of the slack-charging EV model, we clarify our notation. We denote discrete variables in italics, represented as $z(t)$, while vectors or matrices are in bold, like $\boldsymbol{z}$. Changes over time are shown as $\partial z(t) = z(t+1) - z(t)$. The symbols $u(t)$ and $\delta(t)$ stand for the unit step and Kronecker delta function, respectively. The latter is 1 when its input is zero and 0 otherwise. We think of t as an element of the set $\mathcal{T}$ with equally spaced intervals. With time indices set apart by δT, we have $t = t\delta T$. Lastly, the feasible set $\mathcal{L}$, as detailed in Eq. (7), encompasses different instances of aggregated EV demand at each time t, represented by $\ell_{ev}(t)$.

Each EV, indexed by p, is characterized by the following quintuple of attributes: $(t_p^a, X_p, E_p, t_p^d, \rho_p)$. Here, t_p^a denotes the time of arrival, X_p represents the initial energy of the EV's battery, E_p stands for the battery capacity, t_p^d signifies the battery's deadline and ρ_p indicates the set charge rate based on the EV battery. The quantization step is standardized to $\delta T = 1$. Representing these parameters can be complicated, often requiring advanced clustering for aggregate EV demand models [22]. Therefore, we aim to streamline this model, focusing on fewer variables while retaining the intricacies of EV charging requirements. Towards this end, we first define the following two attributes:

1) The charging time $\chi_{p,r}$ - the number of intervals when the EV is actively charging at the rate ρ_p. It can be expressed as:

$$\chi_{p,r} = \left\lfloor \frac{E_p - X_p}{\rho_p} \right\rfloor. \tag{1}$$

2) The slack time $\chi_{p,s}$ - the number of intervals the EV remains at the charging station but is not charging. It can be defined as

$$\chi_{p,s} = t_p^d - t_p^a - \chi_{p,r}. \tag{2}$$

These elements pave the way for simplifying an EV's attributes to a more manageable set: $(\chi_{p,r}, \chi_{p,s}, \rho_p)$. The elements $\chi_{p,r}$ and $\chi_{p,s}$ are in the sets $\{0, \ldots, N_r - 1\}$ and $\{0, \ldots, N_s - 1\}$, respectively. We further denoted a vectorized form of the two elements as:

$$\boldsymbol{\chi}_p = \begin{bmatrix} \chi_{p,r} & \chi_{p,s} \end{bmatrix}^\top \in \mathcal{U}_{rs}, \tag{3}$$

where $\mathcal{U}_{rs} \triangleq \mathbb{R}^{\{0,\ldots,N_r-1\}\times\{0,\ldots,N_s-1\}} \subset \mathbb{N}_+^2$. When an EV starts charging, its remaining charging and slack times will gradually go down.

C. System State

In our model, at any time $t \in \mathcal{T}$, an EV's state is described by $(x_r, x_s) \in \mathcal{U}_{rs}$, with x_r indicating the charging duration and x_s the remaining time an EV driver has to stop charging and await further decisions. For simplicity, in subsequent sections, we use $\boldsymbol{x} = (x_r, x_s) \in \mathcal{U}_{rs}$ to represent a two-dimensional EV status. Our objective is to manage the progression of $n_{\boldsymbol{x}}(t)$, the number of EVs in the state $\boldsymbol{x} = (x_r, x_s)$ at time t. We define a function to model the entry of an EV at time t_p^a as:

$$a_p(t) = u(t - t_p^a), \tag{4}$$

which is one for all time intervals following the arrival of EV p and zero beforehand. Thus, the arrival process that increases the count of new vehicles entering a specific state is given by:

$$a_{\boldsymbol{x}}(t) = \sum_{p\in\mathcal{P}} \delta\left(\chi_{p,r} - x_r\right) \delta\left(\chi_{p,s} - x_s\right) a_p(t), \tag{5}$$

where the process $a_{\boldsymbol{x}}(t)$ is typically modeled as a non-stationary Poisson process, and $\mathcal{P}$ represents the set of all EVs in the system.

Considering the time discretization, the number of arrivals between intervals t and $t+1$ follows a Poisson distribution:

$$\mathbb{P}(a_{\boldsymbol{x}}(t) = n) = \frac{\lambda_{\boldsymbol{x}}(t)}{n!} e^{-\lambda_{\boldsymbol{x}}(t)}, \tag{6}$$

with $\lambda_{\boldsymbol{x}}(t)$ varying according to charging hours.

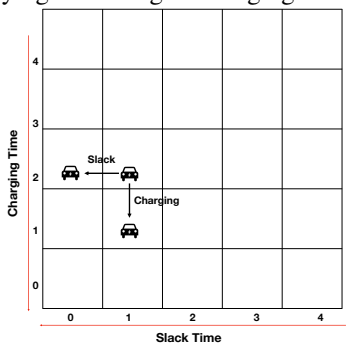


Figure 2. An illustration of the slack-charging EV model.

The number of EVs transitioning from state $\boldsymbol{x}$ to $\boldsymbol{x}'$ at time t is denoted by $\partial d_{\boldsymbol{x},\boldsymbol{x}'}(t)$. Thus, the expression $\big((\boldsymbol{x}' - \boldsymbol{x})\,\partial d_{\boldsymbol{x},\boldsymbol{x}'}(t)\big)_r$ indicates the EVs actively charging from $\boldsymbol{x}$ to $\boldsymbol{x}'$, and $\rho\big((\boldsymbol{x}' - \boldsymbol{x})\,\partial d_{\boldsymbol{x},\boldsymbol{x}'}(t)\big)_r$ represents the total charging power, reflecting the combined EV demands. Furthermore, the total number of EVs transitioning from state $\boldsymbol{x}$ equals the current EVs at that state. This is because the remaining time for each EV decreases, irrespective of whether they opt for charging or waiting. Thus, the combined feasible set for EVs at charging rate ρ is described by:

$$\begin{aligned}\mathcal{L} = \Big\{ \ell_{ev}(t) \mid \ell_{ev}(t) = \sum_{\forall \boldsymbol{x}} \sum_{\boldsymbol{x}'\in\mathcal{U}_{\boldsymbol{x}}} \rho\big((\boldsymbol{x}' - \boldsymbol{x})\,\partial d_{\boldsymbol{x},\boldsymbol{x}'}(t)\big)_r, \\ 0 \le \partial d_{\boldsymbol{x},\boldsymbol{x}'}(t) \le \overline{d}_{\boldsymbol{x},\boldsymbol{x}'}, \partial d_{\boldsymbol{x},\boldsymbol{x}'}(t) \in \mathbb{Z}^+, \\ \sum_{\boldsymbol{x}'\in\mathcal{U}_{\boldsymbol{x}}} \partial d_{\boldsymbol{x},\boldsymbol{x}'}(t) = n_{\boldsymbol{x}}(t), \forall t \in \mathcal{T} \Big\},\end{aligned} \tag{7}$$

where $(\cdot)_r$ denotes the value associated with the charging dimension and $\ell_{ev}(t)$ the aggregated EV demand at time t. The set $\mathcal{U}_{\boldsymbol{x}}$ is given by:

$$\mathcal{U}_{\boldsymbol{x}} = \left\{ \boldsymbol{x}' \mid \|\boldsymbol{x}' - \boldsymbol{x}\|_1 = \min\left(\|\boldsymbol{x}\|_1, 1\right), (\boldsymbol{x} - \boldsymbol{x}') \geq \mathbf{0} \right\}. \quad (8)$$

As depicted in Fig. 2, $\mathcal{U}_{\boldsymbol{x}}$ defines permissible state transitions, ensuring EVs can either charge or wait (slack) at any time, with transitions to states one unit apart (equality constraint), except near the origin, and maintaining movement within the state space's upper right quadrant (inequality constraints). These constraints require that movements should be restricted to one of the two axes – either charging or waiting.

The system dynamics are governed by the state population tensor $n_{\boldsymbol{x}}(t)$ and transitions $d_{\boldsymbol{x},\boldsymbol{x}'}(t)$:

$$\begin{aligned} n_{\boldsymbol{x}}(t) &= a_{\boldsymbol{x}}(t) + \sum_{\boldsymbol{x}' \in \mathcal{U}_{\boldsymbol{x}}} \left[d_{\boldsymbol{x}',\boldsymbol{x}}(t) - d_{\boldsymbol{x},\boldsymbol{x}'}(t) \right], \forall \boldsymbol{x}, \\ \partial d_{\boldsymbol{x},\boldsymbol{x}'}(t) &= d_{\boldsymbol{x},\boldsymbol{x}'}(t) - d_{\boldsymbol{x},\boldsymbol{x}'}(t-1), \forall \boldsymbol{x}, \boldsymbol{x}' \in \mathcal{U}_{\boldsymbol{x}}, \end{aligned} \quad (9)$$

where $n_{\boldsymbol{x}}(t)$ quantifies the EVs present in the system at time t with state $\boldsymbol{x}$. Additionally, we introduce the vector $\boldsymbol{z}$ as $\boldsymbol{z} \triangleq (z_i)_{\forall i}$, where each z_i denotes one element of $\boldsymbol{z}$. In our framework, the term $\boldsymbol{d}^i_{\boldsymbol{x},\boldsymbol{x}'}(t)$ denotes the vector aggregating individual transitions $d_{\boldsymbol{x},\boldsymbol{x}'}$ for EVs located at bus $i \in \mathcal{N}$ across all pairs $\boldsymbol{x}, \boldsymbol{x}'$ within the set $\mathcal{U}_{\boldsymbol{x}}$, at time t. Similarly, $\boldsymbol{n}^i_{\boldsymbol{x}}(t)$ specifies the vector that enumerates all EVs at bus i, in state $\boldsymbol{x}$.

III. Differential Privacy of EV Arrivals

EV driver details can be discerned from the histogram of arrivals $a_{\boldsymbol{x}}(t)$, revealing the charging time $\chi_{p,r}$ and slack time $\chi_{p,s}$ of an EV p in a specific aggregation location. To safeguard user privacy, a privacy-preserving protocol should obscure specific user details using a pseudo-random mechanism. In this context, we adapt the DP framework to design a pseudo-randomized response mechanism, ensuring the privacy of each EV. The definition of DP follows:

Definition 1 (Differentially Private (DP) Mechanism) *For $\epsilon > 0$ and $\delta > 0$, a pseudo-randomized mechanism, $\mathbb{A}$, is said to be (ϵ, δ)-differentially private if, for all $\mathcal{Q} \subseteq \text{Range}(\mathbb{A})$ and for all adjacent[1] vectors $\boldsymbol{x}, \boldsymbol{x}' \in \text{Domain}(\mathbb{A})$ such that $\|\boldsymbol{x} - \boldsymbol{x}'\|_1 \leq 1$, the following holds:*

$$\mathbb{P}\left[\mathbb{A}(\boldsymbol{x}) \in \mathcal{Q}\right] \leq e^{\epsilon} \mathbb{P}\left[\mathbb{A}(\boldsymbol{x}') \in \mathcal{Q}\right] + \delta. \quad (10)$$

If $\delta = 0$, then $\mathbb{A}$ is said to be ϵ-differentially private (ϵ-DP).

DP ensures that for any two adjacent datasets $\boldsymbol{x}$ and $\boldsymbol{x}'$, the distributions of $\mathbb{A}(\boldsymbol{x})$ and $\mathbb{A}(\boldsymbol{x}')$ are very similar, with the total variation distance between them at most ϵ.

A. Randomized Response

Randomized Response (RR), initially developed by Warner [24], was later adapted as a DP framework [25] in the context of label DP to protect the privacy of the labels of the data points in datasets used for machine learning. Suppose the label of point p in a dataset is $\boldsymbol{\chi}_p$, then this simple DP mechanism, parameterized by $\varepsilon \geq 0$, answers the label query with a RR $\tilde{\boldsymbol{\chi}}_p \in \text{Range}(\mathbb{A})$ drawn from the following probability distribution:

$$\Pr[\tilde{\boldsymbol{\chi}}_p = \boldsymbol{x} \mid \boldsymbol{\chi}_p] = \begin{cases} \frac{e^{\varepsilon}}{e^{\varepsilon}+K-1}, & \text{for } \boldsymbol{x} = \boldsymbol{\chi}_p, \\ \frac{1}{e^{\varepsilon}+K-1}, & \|\boldsymbol{x} - \boldsymbol{\chi}_p\|_1 = 1, \\ 0, & \text{otherwise.} \end{cases} \quad (11)$$

where K is the number of labels such that $\|\boldsymbol{x} - \boldsymbol{\chi}_p\|_1 \leq 1$ and $\boldsymbol{x} \in \text{Domain}(\mathbb{A})$. Note, that in this case, the domain and range of the mechanism are equivalent.

Lemma 1 *EV pseudo-random mechanism by Eq. (11) is ϵ-DP.*

Proof: Consider any two adjacent inputs $\boldsymbol{\chi}$ and $\boldsymbol{\chi}'$ and any possible output $\tilde{\boldsymbol{x}}$ by Eq. (11). $\Pr[\tilde{\boldsymbol{\chi}}_p = \boldsymbol{x} \mid \boldsymbol{\chi}_p]$ is maximized when $\boldsymbol{x} = \boldsymbol{\chi}_p$, whereas $\Pr\left[\tilde{\boldsymbol{\chi}}_p = \boldsymbol{x} \mid \boldsymbol{\chi}'_p\right]$ is minimized when $\boldsymbol{x} \neq \boldsymbol{\chi}'_p$. This implies that

$$\frac{\Pr[\tilde{\boldsymbol{\chi}}_p = \boldsymbol{x} \mid \boldsymbol{\chi}_p]}{\Pr[\tilde{\boldsymbol{\chi}}_p = \boldsymbol{x} \mid \boldsymbol{\chi}'_p]} \leq \frac{\frac{e^{\varepsilon}}{e^{\varepsilon}+K-1}}{\frac{1}{e^{\varepsilon}+K-1}} = e^{\varepsilon}.$$

Thus, EV pseudo-random mechanism by Eq. (11) is ε-DP as desired.

B. Proposed DP Mechanism

To prevent the GSO or third-party analysts from infering an EV's initial state $\boldsymbol{\chi}_p$ from arrival histograms, it is crucial that an EV's initial state is obscured privately. As a result, this ensures the charging process remains concealed. The RR mechanism, which is a method responding to data queries, is applied by the aggregator to each EV's state it oversees before computing $\{a_{\boldsymbol{x}}(t)\}_{\boldsymbol{x} \in \mathcal{U}_{rs}}$. Each aggregator, being fully aware of EV states, uses the mechanism to produce differentially privatized arrival counts of EVs in state $\boldsymbol{x}$ at times $t \in \mathcal{T}$, $\tilde{a}_{\boldsymbol{x}}(t)$ given by:

$$\tilde{a}_{\boldsymbol{x}}(t) = \sum_{p \in \mathcal{P}} \delta\left(\tilde{\chi}_{p,1} - x_r\right) \delta\left(\tilde{\chi}_{p,2} - x_s\right) a_p(t), \quad (12)$$

where $\tilde{\boldsymbol{\chi}}_p$ is the output of the RR mechanism for EV p. Fig. 3, illustrates the mechanism, where the modified response has a 90% chance of being the true label of $(3, 1)$, and 2.5% chance of being one of its four neighbors at distance one.

In our mechanism, the parameter K depends on the true state $\boldsymbol{\chi}_p$. When the true state is away from the boundaries as shown in Fig. 3, K is equal to 5 (having 4 neighbors at a distance of one and itself). In scenarios where the state is at a corner, $K = 3$, and $K = 4$ when the state is at an edge.

The DP mechanism introduces randomness into EV arrival counts observed by the GSO. However, it is crucial to address the inherent uncertainty and ensure that the optimal actions derived from these counts remain feasible for the actual arriving EVs. In Section IV, we formulate the optimal power flow (OPF) with DP EVs as a scenario-constrained robust OPF. This reformulation considers a comprehensive set of sampling-based constraints, thereby guaranteeing that optimal solutions adhere to network constraints across all possible conditions.

[1] Adjacent vectors differ in only one record.

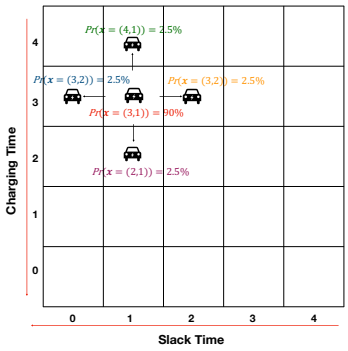


Figure 3. An illustration of the DP pseudo-random mechanism.

C. Parallel Decomposition and Privacy Cost

Even though we possess a matrix representing EV arrivals across all states, the DP mechanism is applied individually to each EV, where an EV's state is fed through the RR mechanism only once during its charging lifetime. Consequently, we employ the parallel composition of ε-DP to provide ε-DP for the entire system.

Theorem 1 ([26, Theorem 4]) *The Parallel Decomposition of ε-DP says that if $\mathbb{A}\left(\boldsymbol{\chi}_p\right)$ satisfies ϵ-differential privacy, then the mechanism which releases the results of all the disjoint datapoints $\left\{\mathbb{A}\left(\boldsymbol{\chi}_1\right), \ldots, \mathbb{A}\left(\boldsymbol{\chi}_{|\mathcal{P}|}\right)\right\}$ is ϵ-differentially private.*

This approach significantly outperforms sequential composition, where executing $|\mathcal{P}|$ times would suggest meeting $|\mathcal{P}|\epsilon$-DP. However, with parallel composition, we can assert that the overall privacy cost remains at ϵ.

IV. Robust Control of Differentially Private EV Models

Building on the foundation laid by the integration of DP to safeguard EV arrival patterns, we now turn our focus to the challenges introduced by the added randomness in the optimization process. This section delves into the development and implementation of a robust scenario-based control strategy. This approach is designed to mitigate the potential adverse effects of DP-induced randomness, ensuring the optimization algorithm remains resilient and feasible under varying conditions.

A. Problem Formulation: Stochastic Optimal Control

In this section, we formulate a multi-stage stochastic optimal control problem that addresses the integration of EVs with power grid operations, taking into account the complexities introduced by power flow constraints and adaptable EV charging. Our focus is to solve OPF with aggregated EVs to minimize fuel costs, underpinned by a formulation that seeks to balance economic dispatch within network constraints while managing EV and battery operations. The problem is formalized as:

$$\max_{\pi(\boldsymbol{s}(t-1))} \mathbb{E}_{\boldsymbol{\ell}_d(t^+)}\left[\sum_{t=\tau}^{\tau+T-1} r(\boldsymbol{s}(t), \boldsymbol{A}(t), \boldsymbol{\ell}_d(t))\right] \tag{13a}$$

$$\boldsymbol{s}(t) = f_t(\boldsymbol{s}(t), \boldsymbol{A}(t-1), \boldsymbol{\ell}_d(t)), \tag{13b}$$

$$\boldsymbol{A}(t) = \pi(\boldsymbol{s}(t-1)), \quad (\boldsymbol{s}(t), \boldsymbol{A}(t)) \in \Omega(t), \tag{13c}$$

where $\boldsymbol{s}(t)$ denotes the state vector, $\boldsymbol{A}(t)$ the control vector including all controllable devices, and $\boldsymbol{\ell}_d(t)$ the demands and renewable energies. The cost function is represented by $r(\boldsymbol{s}(t), \boldsymbol{A}(t), \boldsymbol{\ell}_d(t))$, with $\Omega(t)$ representing network and device bound constraints. The randomized control policy, $\pi(\boldsymbol{s}(t-1))$, guides the decision-making process to achieve network-constrained economic dispatch of power flows by controlling power generations, EVs, and batteries.

Let $\mathcal{S}$ denote the sampling space, encompassing various scenarios $s \in \mathcal{S}$ generated by applying the DP mechanism (designed in Section III) to the EV arrival matrix $a_{\boldsymbol{x}}, \forall \boldsymbol{x}$. Let $\mathcal{N}$ denote the set of nodes within the power system, $\mathcal{G}$ denote the set of nodes connected to power generators, and $\mathcal{B}$ denote the set of nodes connected to battery storage and EV charging substations. In particular, $\boldsymbol{s}(t) \triangleq [\boldsymbol{v}(t), \boldsymbol{soc}(t), (\boldsymbol{a}_{\boldsymbol{x}}^{s,i}(t))_{\forall s \in \mathcal{S}, \forall i \in \mathcal{N}}]^\top$, includes voltage phasors $\boldsymbol{v}(t)$, the state of charge vector $\boldsymbol{soc}(t)$ of all batteries in the system, and the EV arrival matrix $(\boldsymbol{a}_{\boldsymbol{x}}^{s,i}(t))_{\forall s \in \mathcal{S}, \forall i \in \mathcal{N}}$ for all sampling scenarios and buses. The control vector, $\boldsymbol{A}(t) = [\boldsymbol{g}^p(t); \boldsymbol{g}^q(t); \boldsymbol{p}_{ch}(t); \boldsymbol{p}_{dis}(t); (\partial \boldsymbol{d}_{\boldsymbol{x},\boldsymbol{x}'}^{s,i}(t))_{s \in \mathcal{S}, i \in \mathcal{N}}]^\top$, encompasses various elements. It includes both active and reactive generations, denoted as $\boldsymbol{g}^p(t) = (g_i^p(t))_{\forall i \in \mathcal{G}}$ and $\boldsymbol{g}^q(t) = (g_i^q(t))_{\forall i \in \mathcal{G}}$, respectively. Additionally, it includes the charge and discharge rates of batteries, which are represented as $\boldsymbol{p}_{dis}(t) = (p_{dis}^i(t))_{\forall i \in \mathcal{B}}$ and $\boldsymbol{p}_{ch}(t) = (p_{ch}^i(t))_{\forall i \in \mathcal{B}}$, respectively. Furthermore, the vector includes EV control actions, denoted by $\partial \boldsymbol{d}_{\boldsymbol{x},\boldsymbol{x}'}^{s,i}(t)$. All these components are dimensionally extensive due to the incorporation of multiple sampling scenarios, represented by $(\boldsymbol{a}_{\boldsymbol{x}}^{s,i}(t))_{\forall s \in \mathcal{S}, \forall i \in \mathcal{N}}$. Given the high dimensionality of both states and control actions, particularly in the context of RL, we face a vast search space that complicates the optimization process. To address this, Section IV-C presents a solution via polyhedron approximation, significantly simplifying the representation of constraints and actions.

1) Reward Function: In Eq. (13a), we define our reward function as $r(\boldsymbol{s}(t), \boldsymbol{A}(t), \boldsymbol{\ell}_d(t))$, or more compactly, $r(t)$. This function reflects our objectives, which include minimizing fuel costs, represented by $\varsigma_f(t)$ (equivalently, maximizing the negative of fuel costs), and smoothing the electric vehicle (EV) charging demand, denoted by $\varsigma_m(t)$. The formulation is as follows:

$$r(t) = -\varsigma_f(t) - \varsigma_m(t), \tag{14}$$

indicating that higher rewards are obtained for lower fuel costs and smoother EV demand profiles. To achieve the goal of reducing fuel costs, we delve into optimizing the operation of power grids through the integration of EVs, focusing on OPF with the two goals of reducing fuel costs and accommodating power flow constraints along with adaptable EV charging schedules. The primary aim is to explore the economic and operational benefits of aggregated EV participation in the power grid. To this end, we formulate the objective of minimizing fuel costs as follows:

$$\varsigma_f(t) = \sum_{i \in \mathcal{G}} (\upsilon_i (g_i^p(t))^2 + \varrho_i g_i^p(t) + \iota_i), \tag{15}$$

where υ_i, ϱ_i, and ι_i are positive coefficients, representing the quadratic and linear cost components of generation at

generator i at time t. To smooth EV charging demand, we incorporate DR mechanisms via EV charging management to enhance grid flexibility and reduce peak demand. This involves smoothing the charging demands of EVs to avoid sharp increases or decreases, thereby facilitating a more stable load profile across the grid. The objective is designed to measure the variability of EV charging demand:

$$\varsigma_m(t) = \sum_{i\in\mathcal{B}} \left\|\ell^i_{ev}(t) - \ell^i_{ev}(t-1)\right\|_2^2, \tag{16}$$

where $\ell^i_{ev}(t)$ denotes the charging demand of EVs at battery i and time t, aiming to minimize the variation in demand from one time step to the next.

2) Challenges in Robustness and Complexity: In Section III, we touched upon the unpredictability introduced by DP mechanisms, which can sometimes lead to infeasible control solutions. To effectively address this challenge, we incorporate scenario-based constraints to navigate the variability introduced by DP, covering a wide spectrum of potential sampling scenarios as elaborated in subsection IV-B.

After formulating the stochastic control problem, we encounter a second area of challenge: the complexity of a hybrid action space. This space comprises both integer variables for EV charging decisions and continuous variables for power generation and other control actions, presenting significant hurdles for conventional RL methods. To manage the hybrid action space, which comprises both integer and continuous variables, we draw on the Shapley-Folkman Lemma, as highlighted in [9], enabling the approximation of the aggregate feasible action space using continuous variables.

Furthermore, in subsection IV-C, to simplify the representation of the extensive set of constraints associated with all sampling scenarios, we employ a polyhedron approximation strategy. By integrating lower-order polyhedral constraints into our DRL framework, we significantly reduce the complexity of both actions and states. This simplification makes the learning process more tractable, enabling the DRL framework to more effectively address the intricacies of power grid operations with integrated EV charging strategies, achieving notable simplification and enhancing the overall robustness of our control strategies.

B. Robust OPF for DP EV Charging Control

Building on the stochastic control problem formulated in Section IV-A, we now focus on addressing the robust OPF for DP EV charging control. This shift is necessitated by the DP mechanism's inherent randomness, which, while ensuring privacy, introduces unpredictability that could lead to infeasible control solutions. Given this context, our objective expands to ensuring that optimal actions not only account for DP-induced variability but also remain feasible for the actual EV arrivals.

To this end, we cast the OPF problem within the framework of robust optimization, specifically tailored to accommodate a wide array of possible EV arrival matrices. This robust OPF formulation aims to guarantee that optimal solutions satisfy network constraints under all considered sampled arrival matrices, thus enhancing the reliability of the control solutions. The aggregate feasible set $\mathcal{L}_i$ at bus i in Eq. (7) across these scenarios is defined as $\mathcal{L}^s_i$ at bus i. The robust OPF model should consider the intersection of all the feasible EV sets for all sampled arrival matrices, i.e., $\ell_{i,ev}(t) \in \cap_{s\in\mathcal{S}}\mathcal{L}^s_i$ and $\boldsymbol{\ell}_{ev}(t) = (\ell_{ev}(t))_{i\in\mathcal{B}}$. Considering the intersections of all sampling scenario-based constraints, the robust OPF model can be expressed as

$$\max_{\boldsymbol{A}(t)} \quad \mathbb{E}_{\boldsymbol{\ell}_d(t)}\left[\sum_{t=\tau}^{\tau+T-1} r(t)\right] \tag{17a}$$

$$\mathbf{M}_b(\boldsymbol{p}_{dis}(t) - \boldsymbol{p}_{ch}(t) + \boldsymbol{\ell}_{ev}(t)) + \mathbf{M}_g\boldsymbol{g}^p(t) - \boldsymbol{\ell}^p_d(t) = \Re\{D(\boldsymbol{v}(t)(\boldsymbol{v}(t))^H\mathbf{Y}^H)\},\ \boldsymbol{\ell}_{ev}(t) \in \cap_{s\in\mathcal{S}}\mathcal{L}^s_i,\ \mathcal{L}^s_i \in (7)-(9) \tag{17b}$$

$$\mathbf{M}_g\boldsymbol{g}^q(t) - \boldsymbol{\ell}^q_d(t) = \Im\{D(\boldsymbol{v}(t)(\boldsymbol{v}(t))^H\mathbf{Y}^H)\}, \tag{17c}$$

$$\underline{\boldsymbol{g}}^p \le \boldsymbol{g}^p(t) \le \overline{\boldsymbol{g}}^p, \quad \underline{\boldsymbol{g}}^q \le \boldsymbol{g}^q(t) \le \overline{\boldsymbol{g}}^q, \quad \underline{\boldsymbol{v}} \le |\boldsymbol{v}(t)| \le \overline{\boldsymbol{v}} \tag{17d}$$

$$0 \le \boldsymbol{p}_{ch}(t) \le \boldsymbol{P}^{ch}_{rated}, 0 \le \boldsymbol{p}_{dis}(t) \le \boldsymbol{P}^{dis}_{rated}, |\mathbf{K}\boldsymbol{v}(t)| \le \boldsymbol{sj}_{\max} \tag{17e}$$

$$\boldsymbol{soc}_{min} \le \boldsymbol{soc}(t) \le \boldsymbol{soc}_{max}, \forall t \in [\tau, \tau+T-1] \tag{17f}$$

$$\boldsymbol{soc}(t) = \boldsymbol{soc}(t-1) + \frac{\Delta t}{E_{cap}}\Big(\eta_{ch}\boldsymbol{p}_{ch}(t) - \frac{\boldsymbol{p}_{dis}(t)}{\eta_{dis}}\Big). \tag{17g}$$

where the active power demand vector is $\boldsymbol{\ell}^p_d(t) = [\ell^p_d(n,t)]_{\forall n\in\mathcal{N}}$, the reactive power demand vector is $\boldsymbol{\ell}^q_d(t) = [\ell^q_d(n,t)]_{\forall n\in\mathcal{N}}$, $\mathbf{Y}$ is the admittance matrix, and $\boldsymbol{v}(t) = [v(n,t)]_{\forall n\in\mathcal{N}}$ is the grid state in the AC power flow, i.e. $\boldsymbol{v}(t) = |\boldsymbol{v}(t)| \circ e^{\mathrm{j}\boldsymbol{\theta}(t)}, v(n,t) = |v(n,t)|e^{\mathrm{j}\theta(n,t)}$. The matrix $\mathbf{K} \triangleq \mathbf{YI}$, where $\mathbf{I} \in \mathbb{R}^{m\times n}$ is a network-directed graph incidence matrix, where m is the number of lines in the network. The complex power injection is given by $\boldsymbol{sj} = (\boldsymbol{v} \circ \boldsymbol{i}^*) = D(\boldsymbol{v}(\boldsymbol{i})^H)$, where $\boldsymbol{v} \in \mathcal{C}^N$ represents the voltage phasor vector, and $\boldsymbol{i} \in \mathcal{C}^N$ symbolizes the current phasor vector. $(\cdot)^*$ denotes the conjugate of the complex vector or matrix and $\circ$ denotes the Hadamard product (element-wise product). The function $D(\cdot)$ extracts the vector of diagonal elements from a matrix, while $(\boldsymbol{i})^H$ and $(\boldsymbol{i})^*$ represent the Hermitian and the conjugate of the vector $\boldsymbol{i}$, respectively. The current phasor vector can be further expressed as $\boldsymbol{i} = \mathbf{Y}\boldsymbol{v}$, thereby allowing $\boldsymbol{sj}$ to be rewritten as $D(\boldsymbol{v}\boldsymbol{v}^H\mathbf{Y}^H)$. Therefore, the active and reactive power injections are expressed as $\Re(\boldsymbol{sj}) = \Re\{D(\boldsymbol{v}\boldsymbol{v}^H\mathbf{Y}^H)\}$ and $\Im(\boldsymbol{sj}) = \Im\{D(\boldsymbol{v}\boldsymbol{v}^H\mathbf{Y}^H)\}$. $\boldsymbol{sj}_{\max}$ denotes the vector of long-term branch rating limits. Let $\mathbf{M}_g$ the matrix $\{0,1\}^{N\times G}$ that maps the generation vector $\boldsymbol{g}^p \in \mathbb{R}^{|\mathcal{G}|}$ to $\mathbb{R}^N$ as follows:

$$\begin{aligned}&[\mathbf{M}_g\boldsymbol{g}^p]_i = 0, [\mathbf{M}_g\boldsymbol{g}^q]_i = 0, \quad \forall i \in \mathcal{N}\setminus\mathcal{G}\\ &[\mathbf{M}_g\boldsymbol{g}^p]_i = g^p_j, [\mathbf{M}_g\boldsymbol{g}^q]_i = g^q_j, \quad \forall i \in \mathcal{G}, \quad \forall j \in [1,\cdots,G]\end{aligned} \tag{18}$$

and the matrix $\mathbf{M}_b \in \{0,1\}^{N\times B}$ functions as a mapping tool for the vectors of charging power $\boldsymbol{p}_{ch}(t)$, discharging power $\boldsymbol{p}_{dis}(t)$, and EV demands $\boldsymbol{\ell}_{ev}(t)$ across the entire network. It assigns zero to those buses that lack battery storage or EV charging substations. The variable soc denotes the state of charge (SoC) of the battery storage systems, with Eq. (17g) detailing the dynamic evolution of the SoC. The feasible set of Eq. (17) is denoted by Ω_t. The above optimization models are usually computationally intractable due to the complicating

constraints Eqs. (7) - (9), compounded by the addition of a large number of scenario-based constraints denoted by $s \in \mathcal{S}$.

C. Refined Polyhedron Constraints for Reduced Order

In this subsection, the intersections of feasible sets, i.e., Eqs. (7) - (9), is a polytope which means $\mathcal{L}$ can be replaced by convex combination of vertices $\boldsymbol{\nu} \in \mathcal{V}$. In discussing the simplification of EV constraints, it is crucial to understand that the feasible region of a problem constrained by linear equations forms a polytope. This insight enables us to map the EV feasible set into a lower-dimensional space, significantly reducing the complexity of dimensional control actions. Figure 4 depicts the process of low-dimensional reduction. For the intersection of a series of linear constraints, we can select four feasible vertices from the boundaries of this polytope, and the convex combination of these vertices is also feasible within the region of this intersection. Therefore, instead of controlling the high-dimensional actions $(\partial \boldsymbol{d}^{s,i}_{\boldsymbol{x},\boldsymbol{x}'}(t))_{s\in\mathcal{S},i\in\mathcal{N}}$, we can manage the convex combination scalars, significantly simplifying the control mechanism.

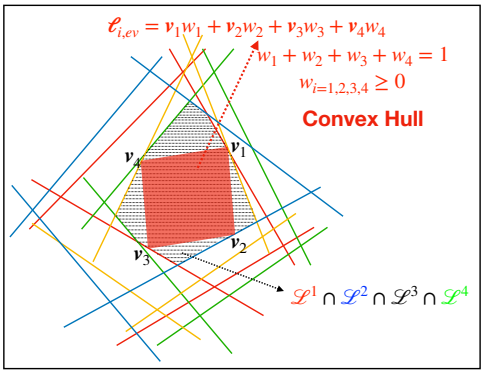


Figure 4. An illustration of the reduced Polyhedron approximation of intersections.

We denote t^+ by a time period, and t by a time point. Let $\boldsymbol{\ell}_{i,ev}(t^+) = [\ell_{i,ev}(\tau)]_{\tau=t}^{t+T-1}$ for a time period, and let $\boldsymbol{\ell}_{ev}(t^+) = [\boldsymbol{\ell}_{i,ev}(t^+)]_{i\in\mathcal{N}}$ for all buses. When we describe a convex combination of vertices, represented as $\boldsymbol{\nu}_\beta(t^+) \in \mathcal{V} = \text{Vert}(\cap_{s\in\mathcal{S}}\mathcal{L}^s)$, the expression is formalized as:

$$\boldsymbol{\ell}_{i,ev}(t^+) = \sum_{\beta=1}^{|\mathcal{V}|} \boldsymbol{\nu}_\beta(t^+) w_\beta(t), \quad \sum_{\beta=1}^{|\mathcal{V}|} w_\beta(t) = 1. \tag{19}$$

where $\boldsymbol{w}(t) \triangleq [w_\beta(t)]_{\forall\beta}$. Presuming the complexity of the convexified aggregate is inherently manageable— an assumption that holds for several tens of thousands of convexified load models — our goal is to pare down the complexity of $\cap_{s\in\mathcal{S}}\mathcal{L}^s$. With a set of price scenarios given as $\boldsymbol{\kappa}_\beta \in \mathbb{R}^\top, \beta \in \mathbb{R}$, our starting point is with $\mathcal{V} = \emptyset$. For each price scenario, we determine the optimal power profile:

$$\boldsymbol{\nu}_\beta = \arg\min \boldsymbol{\kappa}_\beta^\top \boldsymbol{\ell}_{i,ev}(t^+), \tag{20a}$$

$$\boldsymbol{\nu}_\beta \in \cap_{s\in\mathcal{S}}\mathcal{L}^s \tag{20b}$$

Following this, we incorporate the outcome, $\boldsymbol{\ell}_{i,ev}(t^+)$, into $\mathcal{V}$. Subsequently, we put forth another heuristic approach aimed at streamlining the aggregate region. This is achieved by excluding vertices that closely align with others, thereby reducing the facet count within the polytope. Therefore, $\boldsymbol{\ell}_{i,ev}(t^+) \in$ (17b) can replaced by $\boldsymbol{\ell}_{i,ev}(t^+) = \sum_{\beta=1}^{|\mathcal{V}|} \boldsymbol{\nu}_\beta(t^+) w_\beta(t)$, where $\boldsymbol{w}$ are the control variables. This significantly reduce the dimensional of $\partial d^s_{\boldsymbol{x},\boldsymbol{x}'}(t), \forall \boldsymbol{x}, \boldsymbol{x}' \in \mathcal{U}_{\boldsymbol{x}}, \forall s, \forall t = [\tau, \tau+t-1]$ to the number of $|\mathcal{V}|$. Eq. (17b) can be revised as

$$\begin{aligned} &\mathbf{M}_b(\boldsymbol{p}_{dis}(t) - \boldsymbol{p}_{ch}(t) + \boldsymbol{\ell}_{ev}(t)) + \mathbf{M}_g \boldsymbol{g}^p(t) - \boldsymbol{\ell}^p_d(t) = \\ &\Re\{D(\boldsymbol{v}(t)(\boldsymbol{v}(t))^H \mathbf{Y}^H)\}, \ \boldsymbol{\ell}_{ev}(t) = [\ell_{i,ev}(t)]_{i\in\mathcal{B}} \\ &\boldsymbol{\ell}_{i,ev}(t^+) = \sum_{\beta=1}^{|\mathcal{V}|} \boldsymbol{\nu}^i_\beta(t^+) w^i_\beta(t), \ \sum_{\beta=1}^{|\mathcal{V}|} w^i_\beta(t) = 1. \end{aligned} \tag{21}$$

A significant benefit of this approach is its ability to concurrently execute constraint dimensional reductions across various charging stations. In contrast, the original optimization problems required the simultaneous consideration of all control variables associated with different charging stations. Thus, the optimization problem is formulated as Eq. (17), excluding Eq. (17b), combined with Eq. (21). This formulation is denoted as (**P1**). Through this dimensionality reduction, the dimensions of our states and actions are significantly reduced. Specifically, our states are given by $\boldsymbol{s}(t) = [\boldsymbol{v}(t), \boldsymbol{soc}(t), \boldsymbol{v}_\beta(t^+)]^\top$, and the control actions are represented as $\boldsymbol{A}(t) = [\boldsymbol{g}^p(t); \boldsymbol{g}^q(t); \boldsymbol{p}_{ch}(t); \boldsymbol{p}_{dis}(t); \boldsymbol{w}_\beta(t)]^\top$.

D. Disaggregation of Aggregated EV Demands

After determining the optimal w_i^*, we calculate the aggregate EV demand, represented as $\boldsymbol{\ell}_{i,ev} = \sum_{i=1}^{|\mathcal{V}|} \boldsymbol{\nu}_i w_i^*$. The term $\boldsymbol{\ell}_{i,ev}(t)$ indicates the aggregated EV demand at time t for bus i, and it corresponds to the first element of $\boldsymbol{\ell}_{i,ev}$. To achieve the disaggregation of the EV charging control, denoted as $\partial d^i_{\boldsymbol{x},\boldsymbol{x}'}(t), \forall \boldsymbol{x}, \forall \boldsymbol{x}' \in \mathcal{U}_{\boldsymbol{x}}$, one can solve a small-scale mapping problem individually for each i. More specifically, $\forall i, t$, the corresponding projection problem is independently resolved.

$$\begin{aligned} &\min \left\| \ell_{i,ev}(t) - \sum_{\forall \boldsymbol{x}} \sum_{\boldsymbol{x}'\in\mathcal{U}_{\boldsymbol{x}}} \left(\rho\,(\boldsymbol{x}' - \boldsymbol{x})\, \partial d^i_{\boldsymbol{x},\boldsymbol{x}'}(t)\right)_r \right\|_2^2 \\ &s.t. \ \ \partial d^i_{\boldsymbol{x},\boldsymbol{x}'}(t) \in (7) \end{aligned} \tag{22}$$

In this context, the disaggregation problem leverages current actual arrivals to make localized decisions. In this localized disaggregation approach, there is no interaction with centralized agents. Conversely, in the centralized OPF problem, the DP arrival matrix is transmitted to the central agent.

V. Constrained Reinforcement Learning

The task of directly resolving the optimization problem (**P1**) in is highly complex. This complexity arises due to the stochastic nature of the problem. To address the above challenge, in this section, we apply our CRL methodology in [27] for real-time predictive control of the OPF problem, incorporating aggregate EVs.

A. Constrained Twin Delayed Deep Deterministic Policy Gradient (TD3)

The TD3 method, which falls within an actor-critic framework, updates policy function parameters based on a critic or approximate value function [28]. The actor, π_ϕ, selects actions, while the critic, Q_ξ, evaluates these actions. Q-learning uses

temporal difference learning [29] to derive the value function from the Bellman equation [30].

In [27], a constrained TD3 approach is detailed. For our implementation, we adopt two target networks, specifically $Q_{\xi_1'}$ and $Q_{\xi_2'}$, and complement them with two critic networks, Q_{ξ_1} and Q_{ξ_2}. The methodology involves selecting the minimum value between the two value estimates, which promotes a consistent update process for the critic networks. Upon the suitable definition of the critic, we proceed to configure the actor network, introducing a constrained action space. Traditionally, the action network is trained to maximize the critic network's output. This is articulated as:

$$\phi \leftarrow \arg\max_{\phi} Q_{\xi_1}(\boldsymbol{s}(t-1), \pi_\phi(\boldsymbol{s}(t-1))). \tag{23}$$

where ϕ represents the parameters of the action network. We can employ either Q_{ξ_1} or Q_{ξ_2} to guide $\pi_\phi(\cdot)$ in updating ϕ. An action $\boldsymbol{A}(t^+) = \pi_\phi(\boldsymbol{s}(t-1))$ is deemed feasible if $\boldsymbol{A}(t)$ adheres to all its constraints, $\Omega(t)$. Consequently, the policy π_ϕ is derived by maximizing the critic network while ensuring compliance with $\zeta(t)$:

$$\max_{\phi} Q_{\xi_1}(\boldsymbol{s}(t-1), \pi_\phi(\boldsymbol{s}(t-1))) \quad s.t.\ \boldsymbol{A}(t) \in \Omega(t). \tag{24}$$

where $\boldsymbol{A}(t)$ is taken according to policy $\boldsymbol{A}(t) = \pi_\phi(\boldsymbol{s}(t-1))$.

B. Primal-Dual OPF Formulation

In this subsection, we aim to train the constrained policy function, denoted as $\pi_\phi(\cdot)$, for problem (**P1**). We introduce the normalized active power generation $\boldsymbol{g}^p(t)$ and the normalized reactive power generation $\boldsymbol{g}^q(t)$. Moreover, we define the normalized discharging power as $\boldsymbol{p}_{dis}(t)$ and the charging power as $\boldsymbol{p}_{ch}(t)$. Rather than directly controlling the high-dimensional variable $\partial \boldsymbol{d}_{\boldsymbol{x},\boldsymbol{x}'}$, our control actions are represented by w_β, which are convex combination weights. The actions $\boldsymbol{A}(t) = [\boldsymbol{a}_t, \cdots, \boldsymbol{a}_{t+T-1}]$ are expressed as follows:

$$\begin{aligned}
&\boldsymbol{a}(t) \triangleq [\hat{\boldsymbol{g}}^p(t), \hat{\boldsymbol{g}}^q(t), \hat{\boldsymbol{p}}_{ch}(t), \hat{\boldsymbol{p}}_{dis}(t), \boldsymbol{w}(t)]^\top, \\
&\boldsymbol{g}^p(t) \triangleq (1-\hat{\boldsymbol{g}}^p(t))\underline{\boldsymbol{g}^p} + \hat{\boldsymbol{g}}^p(t)\overline{\boldsymbol{g}^p}, \\
&\boldsymbol{g}^q(t) \triangleq (1-\hat{\boldsymbol{g}}^q(t))\underline{\boldsymbol{g}^q} + \hat{\boldsymbol{g}}^q(t)\overline{\boldsymbol{g}^q}, \\
&\boldsymbol{p}_{ch}(t) \triangleq \hat{\boldsymbol{p}}_{ch}(t)\overline{\boldsymbol{p}}_b, \quad \boldsymbol{p}_{dis}(t) \triangleq \hat{\boldsymbol{p}}_{dis}(t)\overline{\boldsymbol{p}}_b
\end{aligned} \tag{25}$$

where the power system environmental is controlled by the unnormalized power generation, unnormalized discharging and charging powers, and the unnormalized aggregated EV demand. The vector $\boldsymbol{w}(t)$ inherently lies within [0, 1] due to $\sum_\beta w_\beta(t) = 1$ and $w_\beta(t) \geq 0$. This can also be equivalently articulated through inequality constraints:

$$0 \leq \sum_{\beta=1}^{|\mathcal{V}|-1} w_\beta(t) \leq 1 \tag{26}$$

The last weight, $w_{|\mathcal{V}|}(t)$, is $w_{|\mathcal{V}|}(t) = 1 - \sum_{\beta=1}^{|\mathcal{V}|-1} w_\beta(t)$.

We introduce dual variables $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ in relation to Eq. (17), alongside the augmented penalty parameters $\boldsymbol{\alpha}$:

$$\begin{aligned}
&\boldsymbol{\lambda} = [\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \boldsymbol{\lambda}_3]^\top, \boldsymbol{\mu} = [\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\mu}_3, \boldsymbol{\mu}_4]^\top, \\
&\boldsymbol{\alpha}_\lambda = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3]^\top, \boldsymbol{\alpha}_\mu = [\boldsymbol{\alpha}_5, \boldsymbol{\alpha}_6, \boldsymbol{\alpha}_7, \boldsymbol{\alpha}_8]^\top
\end{aligned} \tag{27}$$

Furthermore, we reformulate the equality and inequality constraints of Eq. (17) in a more compact manner:

$$\boldsymbol{\omega}_\lambda(t) = \begin{bmatrix}
\mathbf{M}_b(\boldsymbol{p}_{dis}(t) - \boldsymbol{p}_{ch}(t) + \boldsymbol{\ell}_{ev}(t)) + \mathbf{M}_g \boldsymbol{g}^p(t) - \boldsymbol{\ell}_d^p(t) \\
-\Re\{D(\boldsymbol{v}(t)(\boldsymbol{v}(t))^H \mathbf{Y}^H)\}, \\
\mathbf{M}_g \boldsymbol{g}^q(t) - \boldsymbol{\ell}_d^q(t) - \Im\{D(\boldsymbol{v}(t)(\boldsymbol{v}(t))^H \mathbf{Y}^H)\}, \\
\boldsymbol{soc}(t) - \boldsymbol{soc}(t-1) + \frac{\Delta t}{E_{cap}}\left(\eta_{ch}\boldsymbol{p}_{ch}(t) - \frac{\boldsymbol{p}_{dis}(t)}{\eta_{dis}}\right)
\end{bmatrix} \tag{28}$$

$$\boldsymbol{\omega}_\mu(t) = \begin{bmatrix}
|\mathbf{K}\boldsymbol{\vartheta}(t)| - \boldsymbol{s}_{\max} \\
\boldsymbol{soc}(t) - \boldsymbol{soc}_{max} \\
\boldsymbol{soc}_{min} - \boldsymbol{soc}(t) \\
\sum_{\beta=1}^{|\mathcal{V}|-1} w_\beta(t) \leq 1
\end{bmatrix}_+ \tag{29}$$

Algorithm 1: Constrained Reinforcement Learning for OPF with aggregate EVs

Initialize critic network Q_{ξ_1}, Q_{ξ_2}, and actor network π_ϕ with random parameters ξ_1, ξ_2 and ϕ;
Initialize target networks $\xi_1' \leftarrow \xi_1$, $\xi_2' \leftarrow \xi_2$, $\phi' \leftarrow \phi$;
Initialize replay buffer $\mathcal{B}$, and set primal and dual update periods pu and du;
for $t = 1 : T$ **do**
  Select action $\boldsymbol{A}(t) \sim \pi_\phi(\boldsymbol{s}(t-1))$;
  Observe reward $r(t)$ by Eq. (14);
  Obtain the new state $\boldsymbol{s}(t) = \textbf{env}(\boldsymbol{A}(t))$ by taking action $\boldsymbol{A}(t)$;
  Store transition tuple $(\boldsymbol{s}(t-1), \boldsymbol{A}(t), r(t), \boldsymbol{s}(t))$ in $\mathcal{B}$;
  Sample mini-batch of N transitions $\{(\boldsymbol{s}^{n-1}, \boldsymbol{A}^n, r^n, \boldsymbol{s}^n) | n = 1, \cdots, N\}$ from $\mathcal{B}$;
  $y \leftarrow r(t) + \gamma \min_{i=1,2} Q_{\xi_i'}(\boldsymbol{s}^n, \pi_\phi(\boldsymbol{s}^n))$;
  Update critics: $\xi_{i=1,2} \leftarrow \arg\min_{\xi_{i=1,2}} \frac{1}{N}\sum(y - Q_{\xi_{i=1,2}}(\boldsymbol{s}^{n-1}, \boldsymbol{A}^n))^2$;
  if t *mod* pu **then**
    Update ϕ by the deterministic policy gradient: $\phi \leftarrow \phi - \eta \nabla \mathcal{L}_\phi(\boldsymbol{s}^{n-1}, \boldsymbol{s}^n)$, where η is the learning rate and $\mathcal{L}_\phi$ is defined in Eq. (30);
    Update target networks by [27];
  if t *mod* du **then**
    Update the dual variables by Eq. (31).

With the definition of Eq. (25), the augmented Lagrangian is:

$$\begin{aligned}
\min_{\phi} \mathcal{L}_\phi = &- Q_{\xi_1'}(\boldsymbol{s}(t-1), \pi_\phi(\boldsymbol{s}(t-1))) + \Bigg(\boldsymbol{\lambda}^\top \boldsymbol{\omega}_\lambda(t) + \\
&\boldsymbol{\mu}^\top \boldsymbol{\omega}_\mu(t) + \left\| \begin{bmatrix} \mathrm{diag}(\boldsymbol{\alpha}_\lambda) & \mathbf{0} \\ \mathbf{0} & \mathrm{diag}(\boldsymbol{\alpha}_\mu) \end{bmatrix} \begin{bmatrix} \boldsymbol{\omega}_\lambda(t) \\ \boldsymbol{\omega}_\mu(t) \end{bmatrix} \right\|_2^2 \Bigg)
\end{aligned} \tag{30}$$

where $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ are the dual variable vectors, and $\boldsymbol{\alpha}_\lambda$ and $\boldsymbol{\alpha}_\mu$ are positive scalars that penalize the augmented terms.

The primal dual update requires minimizing the Lagrangian function and then maximizing the dual function With the definition of Eq. (25), the dual variables gradient update is:

$$\begin{aligned}
\boldsymbol{\lambda}^{k+1} &\leftarrow \boldsymbol{\lambda}^k + \mathrm{diag}(\boldsymbol{\alpha}_\lambda)\boldsymbol{\omega}_\lambda(t), \\
\boldsymbol{\mu}^{k+1} &\leftarrow \boldsymbol{\mu}^k + \mathrm{diag}(\boldsymbol{\alpha}_\mu)\boldsymbol{\omega}_\mu(t),
\end{aligned} \tag{31}$$

where $\boldsymbol{\lambda}^{k+1}$ and $\boldsymbol{\mu}^{k+1}$ are updated by batch samples. We conduct the primal-dual update alternatively to optimize ϕ of $\pi_\phi(\cdot)$ while enforcing the feasibility of both equality and

inequality constraints. The convergence proof is provided in our previous research [27]. The training process of the constrained reinforcement learning algorithm, described above, is summarized in Algorithm 1.

VI. Case Studies

In our experiments, we used the IEEE 14-bus system, equipped with three EV charging stations and battery systems at Buses 8 and 12, and complemented by three wind power generation units. We also evaluated the IEEE 30-bus system, which has four EV charging stations and battery systems located at Buses 6, 12, 20, and 25, alongside another four wind power generation units. For training the constrained DRL, we used PyTorch, integrating actual demand profiles from Texas. The wind power profiles were based on data from NREL Wind [31], while the EV arrival pattern was simulated using the model in [4]. Charging and slack durations were capped at 6 hours.

A. Experimental Setup

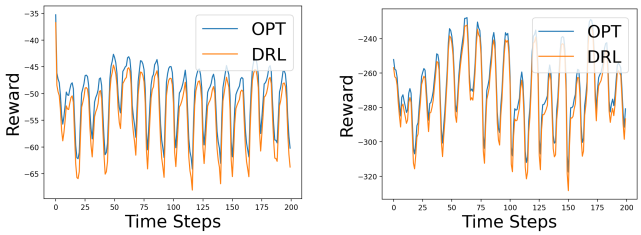
In the context of reinforcement learning, we have set the buffer size at 500, with a discount factor for the reward of 0.99. The rate at which the target network updates is 0.005, and policy updates with delay occur once every two iterations. Both network parameters are optimized using the Adam optimizer with a learning rate of 10^{-3}. Following each time step, the networks are trained on a mini-batch of 100 transitions that are uniformly sampled from the replay buffer. The number of sampled arrival matrices generated for the DP EV arrivals is 20. We select four vertices to approximate intersections within the reduced polyhedron. The EV arrival process is based on [4]. The actor network's architecture is given as follows. For both the IEEE 14-bus and IEEE 30-bus systems, the setup includes an initial cplx-STGCN layer for feature extraction, succeeded by a Complex-valued Neural Network (cplx-NN) layer, as mentioned in [32]. On the other hand, the critic network, which estimates the long-term discounted reward, features a straightforward design: a three-layer real-NN, each furnished with 256 neurons.

B. Testing Curves

After training, we test our policy function with unseen cases. The derived outcomes, illustrated in Fig. 5(a) using 200 sample visuals, emphasize our policy's consistent trend toward almost optimal results, with an optimality gap of only 3.06%. Likewise, we scale the proposed algorithm to the IEEE 30-bus systems. For the training process, the difference between our DRL method and standard optimization is a only 2.15%. For the testing process, a curated selection from 200 sample outcomes is depicted in Fig. 5(b). The result implies our policy's approximation for near-optimal decisions, even in the absence of future EV arrival data and renewable energy generations, with 2.01% optimality gap.

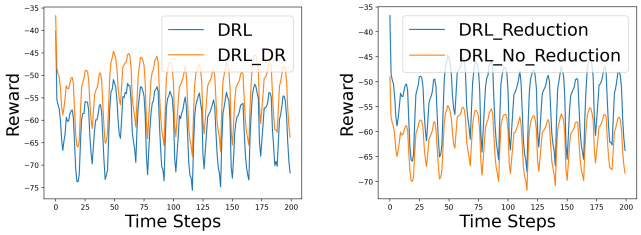
C. Demand Response of EV Controls

In this section, we underscore the advantages of employing demand response in EV control. By optimizing the use of wind power for EV charging, our primary objective is to substantially cut down on fuel costs. The performance of DRL without demand responses is depicted in Fig. 6(a). Clearly, the CRL integrated with demand responses yields higher rewards than its counterpart without them. Fig. 7 presents a comparative analysis between the projected aggregate EV demand incorporating demand response strategies and that without such strategies. The outcomes indicate that the policy function can effectively forecast actions: when wind power is abundant, it prioritizes charging more vehicles. In contrast, during low wind power periods, the policy directs a greater number of EVs towards a relaxed mode rather than charging. Our tests on average fuel cost revealed noteworthy findings. With the implementation of DRL alongside EV demand response controls, we registered a 13.28% drop in fuel expenses compared to setups devoid of DRL demand response. This significant reduction in average fuel cost is attributed to the control policy's adeptness at harnessing renewable energy sources for EV consumption, consequently reducing fuel costs.



(a) Testing curves of rewards (the 14-bus system). (b) Testing curves of rewards (the 30-bus system).

Figure 5. The testing curves for the CRL with EV demand response in the IEEE 14-bus system and the IEEE 30-bus system are illustrated in figures (a) and (b), accompanied by the optimal rewards calculated through the optimization method, labelled as "OPT".

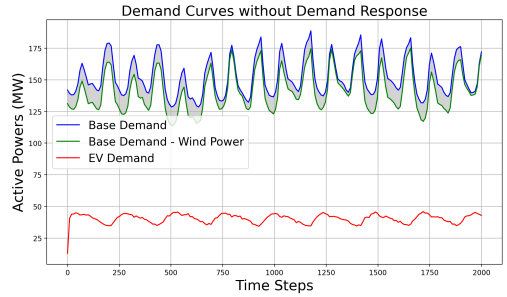


(a) Testing curves of rewards with demand responses. (b) Testing curves of rewards with dimensionality reduction.

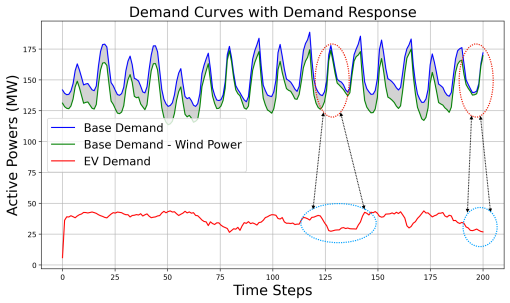
Figure 6. (a) Comparative testing trajectories of the CRL with and without EV demand response; "DRL_DR" denotes CRL incorporating EV demand response. (b) Testing trajectories contrasting the CRL with action space reduction against the CRL without reduction.

D. Action Space Dimensionality Reduction

It is also worth noting that the computational time for the original control actions, without dimensionality reduction, is significantly higher, taking around 605 mins, compared to around 118 mins when dimensionality reduction is applied. We further emphasize the significance of dimensionality reduction in reinforcement learning. Fig. 6(b) demonstrates that the training curves without dimensionality reduction fail to find the optimal actions that reduce the fuel costs. With this dimensionality reduction technique, we achieved an average 16.74% drop in fuel costs. This can be attributed to the fact that the particularly large action space, resulting from the scenarios, makes it challenging for the learning process to identify the optimal actions.



(a) Demand Curves without Demand Response.



(b) Demand Curves with Demand Response.

Figure 7. The active power demand curves encompass base demands, wind power contributions, and EV demands. The shaded areas represent wind power output. Red circles highlight periods of low wind power generation, which, in turn, result in reduced charging power for EVs, as indicated by the blue circles.

VII. Conclusion

In this work, we developed a control model that harmonizes EV charging DR seamlessly with power grid operations, leveraging the inherent flexibility of EV charging. We addressed privacy concerns by implementing DP mechanisms to safeguard the patterns of EV drivers' arrivals. However, the inherent randomness introduced by the DP mechanism posed challenges, occasionally leading to infeasible solutions. To overcome this, our robust control strategy considered a range of potential arrival scenarios, though this introduced complexity due to the high-dimensional nature of control actions. By devising and applying refined polyhedron constraints, we were able to approximate the intersections of feasible regions effectively. Through case studies using IEEE standard systems, we demonstrated our approach's efficacy in reducing generation fuel costs by optimizing the use of renewable resources for EV charging, while simultaneously preserving the privacy of EV drivers.

References

[1] T. Wu, A. Scaglione, A. P. Surani, D. Arnold, and S. Peisert, "Network-constrained reinforcement learning for optimal ev charging control," in *2023 IEEE SmartGridComm*.

[2] M. Tran, D. Banister, J. D. Bishop, and M. D. McCulloch, "Realizing the electric-vehicle revolution," *Nature climate change*, vol. 2, no. 5, pp. 328–333, 2012.

[3] A. Mohamed, V. Salehi, T. Ma, and O. Mohammed, "Real-time energy management algorithm for plug-in hybrid electric vehicle charging parks involving sustainable energy," *IEEE Trans. Sustain. Energy*, vol. 5, no. 2, pp. 577–586, 2013.

[4] M. Alizadeh, A. Scaglione, J. Davies, and K. S. Kurani, "A scalable stochastic model for the electricity demand of electric and plug-in hybrid vehicles," *IEEE Trans. Smart Grid*, vol. 5, no. 2, pp. 848–860, 2013.

[5] A. Unterweger, F. Knirsch, D. Engel, D. Musikhina, A. Alyousef, and H. de Meer, "An analysis of privacy preservation in electric vehicle charging," *Energy Informatics*, vol. 5, no. 1, pp. 1–27, 2022.

[6] C. Dwork, "Differential privacy," in *International colloquium on automata, languages, and programming*. Springer, 2006, pp. 1–12.

[7] C. Dwork, A. Roth *et al.*, "The algorithmic foundations of differential privacy," *Foundations and Trends® in Theoretical Computer Science*, vol. 9, no. 3–4, pp. 211–407, 2014.

[8] Y. Li *et al.*, "Coordinating flexible demand response and renewable uncertainties for scheduling of community integrated energy systems with an electric vehicle charging station: A bi-level approach," *IEEE Trans. Sustain. Energy*, vol. 12, no. 4, pp. 2321–2331, 2021.

[9] K. Hreinsson, A. Scaglione, M. Alizadeh, and Y. Chen, "New insights from the shapley-folkman lemma on dispatchable demand in energy markets," *IEEE Trans. Power Syst.*, vol. 36, no. 5, pp. 4028–4041, 2021.

[10] N. Li, L. Chen, and S. H. Low, "Optimal demand response based on utility maximization in power networks," in *IEEE PES general meeting*. IEEE, 2011, pp. 1–8.

[11] R. Yao, X. Lu, H. Zhou, and J. Lai, "A novel category-specific pricing strategy for demand response in microgrids," *IEEE Trans. Sustain. Energy*, vol. 13, no. 1, pp. 182–195, 2021.

[12] C. Wei, J. Xu, S. Liao, and Y. Sun, "Aggregation and scheduling models for electric vehicles in distribution networks considering power fluctuations and load rebound," *IEEE Trans. Sustain. Energy*, vol. 11, no. 4, pp. 2755–2764, 2020.

[13] 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, 2019.

[14] S. Han, U. Topcu, and G. J. Pappas, "Differentially private distributed constrained optimization," *IEEE Trans. Autom. Control*, vol. 62, no. 1, pp. 50–64, 2016.

[15] S. Islam, S. Badsha, S. Sengupta, I. Khalil, and M. Atiquzzaman, "An intelligent privacy preservation scheme for ev charging infrastructure," *IEEE Trans. Ind. Inform.*, vol. 19, no. 2, pp. 1238–1247, 2022.

[16] T. Ding, S. Zhu, J. He, C. Chen, and X. Guan, "Differentially private distributed optimization via state and direction perturbation in multiagent systems," *IEEE Trans. Autom. Control*, vol. 67, no. 2, pp. 722–737, 2021.

[17] P. Donti, A. Agarwal, N. V. Bedmutha, L. Pileggi, and J. Z. Kolter, "Adversarially robust learning for security-constrained optimal power flow," *NeurIPS*, vol. 34, pp. 28 677–28 689, 2021.

[18] A. Agarwal, P. L. Donti, J. Z. Kolter, and L. Pileggi, "Employing adversarial robustness techniques for large-scale stochastic optimal power flow," *Electric Power Systems Research*, vol. 212, p. 108497, 2022.

[19] D. Bertsimas, E. Litvinov, X. A. Sun, J. Zhao, and T. Zheng, "Adaptive robust optimization for the security constrained unit commitment problem," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 52–63, 2012.

[20] Z. Chen, L. Wu, and Y. Fu, "Real-time price-based demand response management for residential appliances via stochastic optimization and robust optimization," *IEEE Trans. Smart grid*, vol. 3, no. 4, pp. 1822–1831, 2012.

[21] K. Margellos, P. Goulart, and J. Lygeros, "On the road between robust optimization and the scenario approach for chance constrained optimization problems," *IEEE Trans Autom. Control*, vol. 59, no. 8, pp. 2258–2263, 2014.

[22] M. Alizadeh, A. Scaglione, A. Applebaum, G. Kesidis, and K. Levitt, "Reduced-order load models for large populations of flexible appliances," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 1758–1774, 2014.

[23] K. Hreinsson, A. Scaglione, and V. Vittal, "Aggregate load models for demand response: Exploring flexibility," in *2016 IEEE GlobalSIP*.

[24] S. L. Warner, "Randomized response: A survey technique for eliminating evasive answer bias," *Journal of the American Statistical Association*, vol. 60, no. 309, pp. 63–69, 1965.

[25] B. Ghazi, N. Golowich, R. Kumar, P. Manurangsi, and C. Zhang, "Deep learning with label differential privacy," *NeurIPS*, vol. 34, pp. 27 131–27 145, 2021.

[26] F. D. McSherry, "Privacy integrated queries: an extensible platform for privacy-preserving data analysis," in *Proc. ACM SIGMOD Int. Conf. Manag. Data.*, 2009, pp. 19–30.

[27] T. Wu, A. Scaglione, and D. Arnold, "Constrained reinforcement learning for predictive control in real-time stochastic dynamic optimal power flow," *IEEE Trans. Power Syst.*, 2023.

[28] S. Fujimoto, H. Hoof, and D. Meger, "Addressing function approximation error in actor-critic methods," in *ICML*, 2018, pp. 1587–1596.

[29] R. S. Sutton, "Learning to predict by the methods of temporal differences," *Machine learning*, vol. 3, no. 1, pp. 9–44, 1988.

[30] R. Bellman, "Dynamic programming," *Science*, vol. 153, no. 3731, pp. 34–37, 1966.

[31] C. Draxl, A. Clifton *et al.*, "The wind integration national dataset (wind) toolkit," *Applied Energy*, vol. 151, pp. 355–366, 2015.

[32] T. Wu, A. Scaglione, and D. Arnold, "Complex-value spatio-temporal graph convolutional neural networks and its applications to electric power systems ai," *IEEE Trans. Smart grid*, 2023.