Design and Optimal Operation of a Virtual Power Plant with Bidirectional Electric Vehicle Chargers

by

Saidur Rahman

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Computing Science

University of Alberta

Abstract

Virtual power plants (VPPs) can enhance reliability and efficiency of power systems with a high share of renewables. However, their adoption largely depends on their profitability, which is difficult to maximize due to the heterogeneity of their components, different sources of uncertainty and potential profit streams. In this thesis, we study a VPP that aggregates a fleet of electric vehicles (EVs), EV chargers with vehicle-to-grid (V2G) support, and possibly renewable energy systems, such as solar panels. This VPP generates profit by trading energy in day-ahead and imbalance electricity markets. In the first part of this thesis, we assume that the VPP owns and operates the EVs in addition to bidirectional chargers. We propose two profit-maximizing operating strategies for this VPP. Both strategies solve a two-stage stochastic optimization problem. In the first stage, energy bids are placed by solving a sequence of linear programs, each formulated for a specific forecast scenario. In the second stage, given the day-ahead commitments and real-time measurements, the decisions with respect to charging or discharging EVs are made sequentially for every hour, and adjustments to the day-ahead commitments are settled in the imbalance market. The two strategies differ in how they solve the sequential decision-making problem in the second stage. But, they both foresee the effect of their current (dis)charge decisions on the feasibility of fulfilling the EV charging demands using a one-step lookahead technique. The first strategy employs a heuristic algorithm to find a feasible charging schedule for every EV that is connected to a charger. The second one utilizes a soft actor-critic

reinforcement learning method with a differentiable projection layer that enforces constraint satisfaction. We empirically evaluate the proposed operating strategies using real market prices, solar traces, and EV charging sessions obtained from a network of chargers in the Netherlands, and analyze how the uptake of V2G could affect the profitability of this VPP.

In the second part of this thesis, we study this VPP under a more realistic assumption that the EVs are independently-owned, hence the VPP does not own or operate them. As a result, EV owners must be incentivized to participate in the VPP, and the VPP itself must remain profitable. We use contract theory to design optimal, incentive-compatible contracts between the VPP and EV owners, where each contract defines the maximum amount of energy that can be discharged from the battery in a fixed period of time, and the compensation the owner receives in return. We then propose a scheduling algorithm for the optimal operation of such a VPP that participates in a two-stage electricity market. This algorithm aims to maximize the VPP's profit, while (a) respecting the contracts that are accepted and currently valid, and (b) fulfilling the charging demand of each EV before it disconnects from the charger. We show that this algorithm increases the profitability of this VPP and allows EVs to offset the cost of charging their battery by enhancing grid reliability.

Preface

This thesis is the original work by Saidur Rahman. The results that are presented in Chapter 4 of this thesis are based on a conference publication which is co-authored by the author of this thesis: Saidur Rahman, Linda Punt, Omid Ardakanian, Yashar Ghiassi-Farrokhfal, Xiaoqi Tan, *On Efficient Operation of a V2G-Enabled Virtual Power Plant*, Accepted in the 9th ACM International Conference on Systems for Energy-Efficient Built Environments (ACM BuildSys), 2022.

As the first author, Saidur Rahman was responsible for developing the methodology, designing the experiments of the proposed methodology, conducting simulations, and reviewing the literature. The second author, Linda Punt, helped to formulate the problem and in pre-processing the datasets of Electric Vehicle (EV) charging and the multi-stage electricity markets. The third author and my thesis supervisor, Dr. Omid Ardakanian, reviewed the problem formulation, helped in the formulation of the optimization problems in the proposed methodology, edited the manuscripts, and provided excellent supervision. The fourth author, Dr. Yashar Ghiassi-Farrokhfal, provided knowledge on the electricity market and VPP operation in the Netherlands and edited the manuscript. Finally, the fifth author, Dr. Xiaoqi Tan, provided ideas about incentive mechanism design and edited the manuscripts.

Acknowledgements

I am grateful to my awesome thesis supervisor Dr. Omid Ardakanian for providing me the opportunity to do research with him. During the course of the work relationship that I have had with him, he has been instrumental in shaping me to become a better researcher and a more productive individual.

I would also like to thank Linda Punt, Dr. Yashar Ghiassi, Dr. Xiaoqi Tan, and the members of the Sustainable Computing Research Group (SCRG) for their valuable ideas and suggestions in helping me to improve my thesis.

I am immensely indebted to my parents for prioritizing academic excellence from my early childhood and ingraining the importance of education within me from an early age. Their constant unconditional support and encouragement helped me tremendously throughout my entire life. I dedicate all of my life's achievements to them.

I would also like to thank my friends, especially Zahid, Rafi, Navid, Saurin, Jessica, Najifa, Alvina, Tahsin, and Mahi who made my life in Canada super enjoyable and helped me maintain a work-life balance.

Finally, I would like to thank the Future Energy Systems (FES) at the University of Alberta and the Canada First Research Excellence fund for providing grants for my research.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In recent years, power grids have undergone significant changes owing to the integration of highly accurate sensor technology combined with advanced control systems, which allows optimal management of loads whose demands can be shaped (elastic loads) such as Electric Vehicles (EVs), thermostatically controlled loads, and storage systems. A Smart Grid (SG) is: "an electricity network that can intelligently integrate the actions of all users connected to it (generators, consumers, and those that do both) in order to efficiently deliver sustainable, economic, and secure electricity supply" [54]. Sensors installed in the SG measure voltage, current, and other physical quantities at regular intervals, producing time-series data that are then used by the SG for certain actions. These actions make it possible to operate the SG in a reliable and stable manner at lower costs and with a greater energy efficiency [29], [54], [100].

Due to the intermittency of renewable generation (e.g., solar) and the stochasticity of EV mobility, it is crucial for the SG to incorporate adaptive control strategies to prevent issues pertaining to grid stability and reliability. The available real-time sensor measurements (communicated to the SG using a reliable communication network) can be used to develop control strategies to optimize power flow, and also to develop event detection algorithms to identify and prevent unprecedented events. Towards this end, various techniques have been presented in the existing literature [9].

Figure 1.1 shows the operation of an SG with energy producers and various

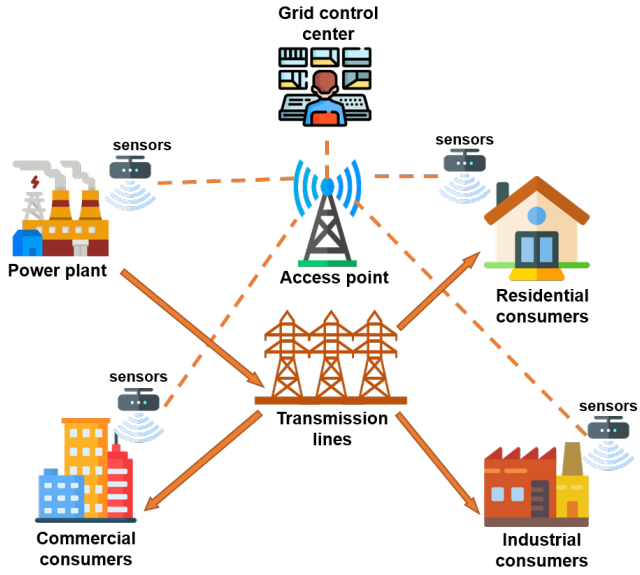


Figure 1.1: An illustration of how SG operates. Dashed lines show bidirectional digital communication between the grid control center and the consumers. Solid arrows show the power flow direction from the power plant to the consumers.

energy consumers that are typically part of the SG [8]. Sensors installed at the consumer side send measurements of energy usage to aggregation points via the communication network (dotted lines). These measurements are then sent to the grid control center. The grid operator can use these measurements to control the power plant operation or elastic loads. The electricity produced at the power plant is transmitted to the consumers via transmission lines. The solid orange arrows show the direction of power flow.

1.1 Integration of Distributed Energy Resources

Future power systems will heavily rely on distributed energy resources (DERs) as they provide energy at a lower cost than the electric grid, and enable greater resilience during adverse grid events. These resources, which generate, store, or reshape energy profiles, can be classified into five types: distributed generation units (*e.g.*, solar systems), battery storage, electric vehicle charging stations, grid-interactive appliances, power-to-heat resources (*e.g.*, heat pumps and thermal storage) [50]. Massive DER growth is expected in the next several

years. The cumulative DER capacity in the United States will reach 387GW by 2025 [113]. In Europe, DERs will provide 100GW of demand response, and in Australia, they will supply 30% to 45% of the total electricity demand by 2050 [34]. With the growing adoption of DERs, the concept of a virtual power plant (VPP) has become increasingly popular [4].

1.2 VPP Concept and Implementations

A VPP aggregates and orchestrates disparate DERs through sensing, communication, and control technology, to provide various services to the grid and increase the value of the DERs. For example, Tesla [109] and Swell Energy [105], in partnership with local utility companies, have implemented VPPs to support the grid by aggregating energy storage and solar systems in residential buildings. We study two novel VPP implementations in this thesis as shown diagrammatically in Figure 1.2. In Chapter 4, the VPP consists of solar systems, a fleet of EVs (owned by the VPP), and charging stations with vehicle-to-grid (V2G) support. This VPP combines renewable generation with mobile energy storage that can be charged and discharged subject to various constraints, such as fulfilling the EV charging demand by some deadline. Such a combination has been shown to enhance the mutual benefits of solar generation and flexibility of the EV fleet [71], [78]. While this type of VPP has been studied in the past (see [71] for example), the uncertainty in solar production and EV mobility together with a large number of decision variables have hindered the progress toward an optimal operating strategy. The limited number of studies, while insightful, exclude some important aspects, such as V2G [30], [117] make strong assumptions and simplifications, are not scalable, or are limited to specific network configurations, hence not conclusive [75], [78]. There are a few pilot projects of this type of VPP that are still in early stages (*e.g.,* the project in Utrecht [114] will combine 2,000 solar panels, 250 bidirectional chargers, and a car-sharing fleet). In Chapter 5, the VPP that we study consists of a fleet of EVs that are under independent ownership, and charging stations with vehicle-to-grid (V2G) support. This realistic setting

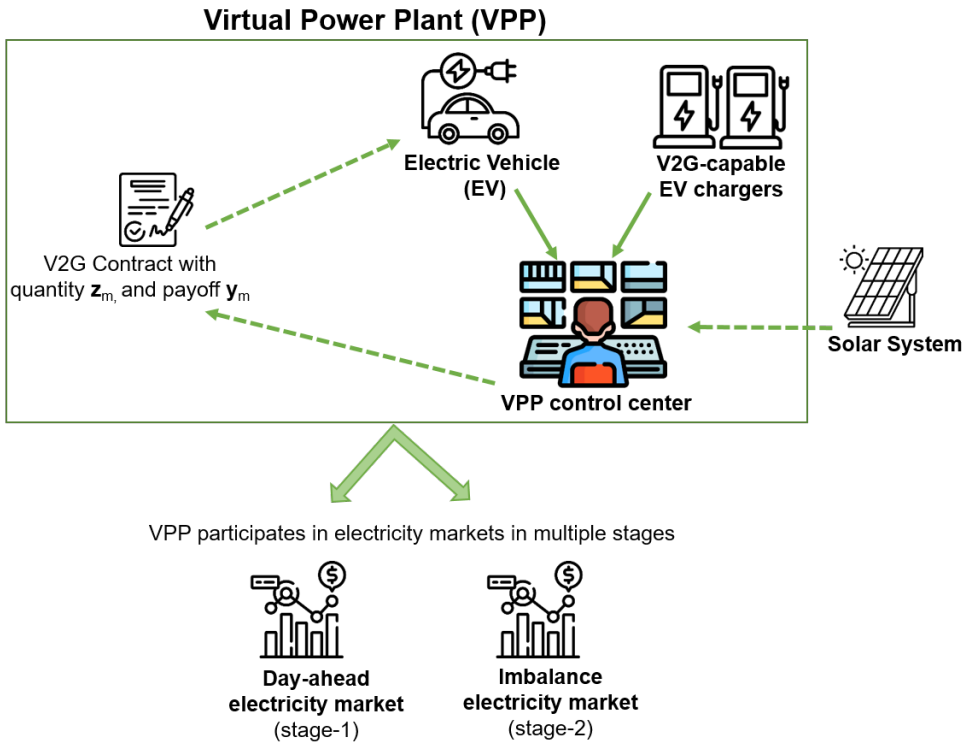


Figure 1.2: Communication between the various entities that make up the VPP. Dashed lines show entities that are optional, and solid lines represent the entities that are always part of the VPP that is studied in this thesis

has an additional challenge of offering reasonable monetary incentives to motivate the EV owners to participate in V2G since EV batteries would have a decreased life span due to having to go through the additional charge-discharge cycles by participating in V2G.

The VPPs can trade their aggregate energy in different stages of electricity markets, such as the wholesale market (day-ahead and intra-day), ancillary service, and capacity markets [50]. Participation in the wholesale electricity market, in particular, has attracted more attention [82] due to its simple form of bidding, higher predictability, manageable types of (energy) commitments, and the fact that it is by far the largest electricity market today. The participation of VPPs in this market could reduce wholesale prices considerably and cut end users' electricity bills. For example, Tesla's VPP in South Australia [101] can reduce the annual electricity bill of a typical customer by 30% by trading in the wholesale market.

Inspired by this, we consider a setting in which the VPP trades (buys/sells) energy in the day-ahead (DA) market. Due to the uncertainty in DERs such as stochastic EV mobility patterns, the amount of energy provided by the VPP in real time may not match its DA commitment. In this case, real-time deviations from the DA commitments are adjusted by the system operator, and the resulting financial cost or profit in the imbalance (IM) market is transferred to the VPP operator. Figure 1.2 illustrates the VPP's participation in these two markets. In the DA market, the (price-taker) VPP operator places an *energy bid* for every hour of the next day according to its predictions of the available solar energy, DA and IM market prices, and EV charging demands. On the operation day (*i.e.*, the next day), deviations from the DA bids are adjusted in the IM market according to imbalance prices. The price volatility in these markets further complicates the design of an optimal operating strategy for this VPP. The VPP operator can take advantage of V2G to shape the EV charging demand to increase its profit. Since the VPP operator owns and operates the EV fleet in the VPP setting that we study in Chapter 4, thus it does not collect payment from the EVs for charging their battery, nor does it compensate them for participating in V2G. Its profit solely depends on the amount of energy traded in the two markets in each hour. An example of this VPP is a car-sharing or taxi company that owns a fleet of EVs in addition to bidirectional chargers and solar systems installed across the city (as depicted in Figure 1.2). The VPP setting that we study in Chapter 5 assumes that the EVs are owned independently and thus the EV owners need to be appropriately incentivized to participate in V2G. The optimal operation of both of our VPP settings are a complex, two-stage stochastic optimization problem due to time-varying constraints and high uncertainty that can be attributed to intermittent solar generation, mobility and energy demand of EVs, and electricity market prices.

We develop practical operating strategies for this VPP that involve (a) bidding in the DA market according to the average solution of a sequence of linear programs defined for different *forecast scenarios*, *i.e.*, realizations of random variables; (b) charging or discharging EVs the next day via an online

algorithm to maximize the VPP's profit, while ensuring the demand of every EV can be met before it disconnects from the charger.

It is to be noted that all the computation takes place in the VPP control centre and we assume that the VPP control centre is equipped with servers that meet the stringent performance requirements of the VPP.

1.3 Incentive Mechanism for V2G Participation

Vehicle-to-Grid allows discharging of energy from the EV battery and injecting it into the grid. V2G can play a crucial role in providing ancillary services to the VPP. However, when EVs are owned by independent owners, they need to be incentivized (usually through monetary incentives) to motivate them to participate in V2G. This is primarily due to the accelerated EV battery degradation associated with frequent charge/discharge cycles when EVs participate in V2G. Therefore it is crucial to provide incentives to EV users (that offset some of the battery degradation cost) for V2G participation. The authors of [51], [115] studied VPPs for the scenario where a VPP integrates renewable generation and EV chargers with V2G capability. However, an important requirement that was not considered in previous work is ensuring the fulfillment of the EV charging demands. In this thesis, we overcome this limitation by presenting a scheduling algorithm and offering contracts that guarantee the feasibility of EV charging/discharging (i.e. ensuring fulfillment of EV owners' charge requirements by their charging deadline) despite the discharging of EV batteries as a result of V2G participation via the offered contracts. In addition, it is commonly assumed that energy storage systems are owned by the VPP operator. However, this restricting assumption implies that the VPP can only operate the resources it owns. As a result, the size (capacity) of this VPP would be smaller as it only operates the resources that it owns and therefore, it cannot be scaled due to the increased complexity of controlling all the VPP-owned entities. Relaxing this assumption raises the question pertaining to the incentive of third-party EV owners in V2G participation. Towards this end,

we seek an appropriate method for motivating EV owners to participate in EV discharge, by providing them monetary incentives, so that the VPP operator[1] can mitigate the surplus/deficit in energy with respect to to the day-ahead bids. Each of the EV owners is *self-interested* and therefore, they will attempt to claim a large incentive amount for the discharge power they allow. This problem is made even more difficult due to the information asymmetry that exists since the VPP operator does not have accurate knowledge about the discharge energy preference of the EV owners when participating in V2G. In this study, the V2G-discharge energy preference of an EV owner (assumed to be their private information) is unavailable to the VPP operator which necessitates designing a mechanism in which EV owners will be rewarded in accordance with their EV discharge amount. The said mechanism must also ensure that EV owners cannot *game* the system by being untruthful about their incentive threshold for V2G participation. Towards this end, we use Contract Theory [19].

Providing incentives to EV users for participation in V2G has been previously studied in the research works in [39], [53], [123]. However, this thesis is the first study that finds the efficacy of using contract theory for incentivizing EV users for V2G in the context of a VPP. In addition, the methodologies used for application of contract theory for V2G in [39], [53], [123] have several shortcomings that we tackle in our work. For example, the designed contracts do not ensure a feasible EV charge/discharge strategy (i.e. ensuring EV owners leave the charging station having their EV battery charged to their desired SoC level), when EV owners accept the proposed contract. In addition to that, simulation studies (based on real-life datasets) that show the efficacy of the offered contract are missing. Moreover, any scheduling mechanism for optimal utilization of the discharge energy from the EV owners under the uncertainties of future solar power, electricity market price, and EV mobility has not been provided. Finally, there is no discussion regarding the duration of the contract which can affect the discharge of energy as well as the incentive amount in the contract.

[1]We will use the terms *VPP operator* and *VPP* interchangeably

1.4 Research Questions

This thesis aims to answer the following research questions:

1. How to overcome the inherent challenges of operating a VPP that participates in 2-stage electricity markets?
2. What algorithms can be used to solve this stochastic decision-making problem under uncertainty and risk?
3. What is the profitability of this VPP under the proposed operating strategy?
4. How do different rates of V2G participation by EV owners affect the VPP profitability?
5. How to incentivize independent EV owners to participate in VPPs?
6. What are the factors that influence EV owners' willingness to participate in V2G?
7. What is the profitability of VPPs after accounting for the payoff provided to the EV owners?

Research Question 1 is answered in Chapter 3, where we present the mathematical model and the assumptions that we made to overcome the inherent challenges of operating the VPP participating in a 2-stage electricity market. We answer Research Questions 2 to 4 in Chapter 4, where we present our methodology to optimize the VPP operation and examine the VPP profitability using the two proposed profit-maximizing algorithms. To answer Research Questions 5-7, in Chapter 5 of this thesis, we use contract theory to incentivize independent EV owners, analyze the offered contracts, and finally, we analyze the resulting VPP profit.

1.5 Outline

Chapter 2 presents the related work that studies different VPP types and their trading platforms. We explain the methods that have been used to solve problems involving decision-making under uncertainty, such as model predictive control, stochastic optimization, and reinforcement learning. We also explain different ways of designing and offering incentives and their application to provide ancillary services or control DERs.

In Chapter 3, the problem formulation is presented, the assumptions are stated, and the 2-stages of VPP operation are discussed. This chapter then provides a detailed discussion about the working mechanisms of the proposed strategies for solving the stage-1 and stage-2 problems.

In Chapter 4, we explain the operating strategies that a VPP operator can utilize to maximize its profit by participating in the 2-stage market (explained in Chapter 3), under the uncertainty (of solar generation and EV mobility), risk (due to price volatility), and the requirement of fully charging the EVs before their departure from the charging station. Then, we evaluate the efficacy of the proposed strategies as the number of EVs that participate in V2G increases. We show that V2G can increase the VPP's profit by more than 42% compared to the case where V2G is not supported at all. Finally, we explain the data sources, as well as the baselines considered for evaluation.

In Chapter 5, we relax the assumption that the EV fleet is owned by the VPP operator. Hence, EV owners must be incentivized to participate in V2G while they are connected to a charger. We explain our contract theoretic approach to provide monetary incentives to EV owners via novel V2G contracts, thereby allowing them to enter into an agreement with the VPP operator so that their EVs can be discharged for a pre-specified duration of time in exchange for some monetary benefit. We also prove that a computationally efficient version of the contract optimization problem can be solved instead to obtain the contracts more efficiently. Then the choice of parameters in experiments is discussed, and a sensitivity analysis to different factors is performed. In addition to that, a heuristic algorithm that schedules the charging

of V2G-participating EVs for optimal VPP profit is presented. Chapter 5 also discusses the efficacy of the proposed contract theoretic approach for incentivizing EV owners for efficient VPP operation in terms of the VPP profit. The comparison is done with respect to a baseline for different levels of V2G participation.

Chapter 6 concludes this thesis, where the limitations of this work are outlined and the scope for improvement/future extension is discussed.

Chapter 2

Related Work

In this chapter, we first provide details of the different types of VPPs that have been presented in the literature along with their trading platforms and the respective goals they were designed to achieve. Secondly, we review the related work on optimizing the operation of VPPs under the inherent uncertainties of the DERs and electricity market prices. In particular, we review Robust Optimization (RO), Model Predictive Control (MPC), Stochastic Optimization (SO), and Reinforcement Learning (RL)-based approaches. Thirdly, we provide an overview of the approaches used to provide incentives to DER owners/operators, namely Mechanism Design (MD) and Contract Theory (CT). Finally, we conclude this chapter by highlighting the limitations of existing CT approaches that have been used to incentivize the stakeholders for V2G participation.

Reference	**Hydro plant**	**Solar**	**Wind farm**	**Battery**
[92], [74], [68], [38], [56], [88], [99], [107]	×	✓	✓	✓
[58]	×	✓	✓	×
[2]	✓	✓	×	✓
[110], [49], [79], [13]	×	×	✓	✓
[16]	✓	×	×	✓
[66], [93], [15]	×	✓	×	×
[122]	✓	✓	✓	×

Table 2.1: Taxonomy of related work that considered a VPP.

Reference	VPP design goal
[47],[106],[97],[119],[73],[65], [72],[127],[94],[24],[81],[68]	Energy trading in electricity markets for profit
[92],[84],[65],[94],[24],[21],[87],[112]	Improving forecast of energy production/consumption
[76],[24],[105],[81],[68]	Providing energy backup and enhancing stability
[77],[119],[72],[83],[80],[81],[18]	Frequency regulation and capacity management
[119],[84],[65],[72]	Scheduling and balancing the VPP's loads

Table 2.2: VPP design goals in the literature.

2.1 Different VPP Types and their Trading Platforms

Extensive research has been conducted on a VPP that incorporates various combinations of DERs as well as some form of energy storage ([41], [64], [79]) as shown in Table 2.1. Bagchi et al. [13] quantify the additional gain of adding a stationary energy storage system to a VPP. The VPPs that combine wind and solar power with battery are considered in [107]. A combination of a conventional power plant, wind turbines, and energy storage is studied in [79]. Utkarsha et al. consider a VPP that aggregates prosumers in combination with energy storage [6]. Thermal generation [41], conventional power plants and wind turbines [79], grid-interactive appliances and HVAC [64] have also been combined with energy storage to implement a VPP. A VPP that integrates EVs without V2G has been the subject of many studies too (see for instance [30], [55], [117]). Jin et al. [55] explore the problem of EV charging with an external battery used for energy storage. Derakhshandeh et al. [30] consider a VPP that combines heat and power (CHP), solar PV, and EVs. Yao et al. [117] study VPPs that consist of different types of DERs, including EVs. Fewer studies consider a VPP that integrates renewable generation and EV chargers with V2G capability [51], [115]; these are the closest work to ours. We study the same type of VPP in a more practical setting and propose efficient operating strategies (for bidding and smart charging) that honor day-ahead commitments and guarantee the fulfillment of the EV charging demands, an important requirement, especially for V2G, that was not considered in previous work.

VPPs are capable of providing various prevalent challenges of the exist-

ing grid including, but not limited to, improving grid resiliency, and optimal DER utilization. Table 2.2, presents some of the existing VPP literature that has been designed with various goals. However, the VPP operator is typically assumed to be a profit seeker, with a few exceptions such as [13] where the VPP aims to become an energy-independent entity. Most related work considers the wholesale electricity market (typically the DA market) as the primary trading platform for the VPP operator. However, a simplified version of the wholesale market with a single stage and exogenous hourly prices is commonly considered [30], [41], [115]. Only a small number of papers envisage a two-stage model of trading in the electricity markets by accounting for the bidding in the DA market and the energy adjustments made in the IM and/or reserve market [13], [51], [64], [117]. We also adopt the two-stage trading model where the VPP operator trades energy in the DA market and settles its adjustments in the IM market. Despite this similarity, both of the VPP settings that we study have a unique configuration. The first VPP setting (studied in Chapter 4) includes EVs (owned by the VPP), chargers with V2G support, and solar systems. The second VPP setting (studied in Chapter 5) included independently-owned EVs and chargers with V2G support. The interactions among these DERs Besides, we quantify the additional gain that V2G provides in this type of VPP.

2.2 Decision Making Under Uncertainty

The optimal control of DERs in a VPP can be viewed as a decision-making problem under uncertainty and risk. This is primarily due to the intermittency of renewable generation, the uncertainty of EV mobility and charging demands, and the volatility of market prices. Thus, a wide range of techniques, from stochastic dynamic programming to robust optimization, can be applied to optimize the VPP operation.

2.2.1 Robust Optimization (RO)

Robust Optimization (RO) is one of the prevalent approaches used to take decisions under uncertainty, for the control of DERs and VPPs, as found in [70], [95]. The RO methods incorporate the uncertainty in the objective and constraint functions by taking expectation with respect to the control variable or considering the supremum of multiple functions. For example, a distributionally robust chance-constrained model is proposed in [125] to control several HVAC systems to absorb as much solar generation as possible. In another work, a chance-constrained energy management model is proposed in [95] to optimally control renewable generation and battery energy storage systems in a microgrid. Another variant of RO is used in [95] for the control of DERs. Several papers cast the control of DERs as a robust optimization problem [95], [125]. For example, a robust optimization problem is solved in [70] to develop a residential energy management system that controls PV systems, batteries, EVs, and thermostatically controlled loads. Another variant of RO is used in [95] for the control of DERs.

2.2.2 Model Predictive Control (MPC)

Model Predictive Control (MPC) is another approach that has been used to control DERs in an online fashion. In this approach, a model is utilized to predict the system dynamics and changes in the environment. This predictive model yields point estimates for future time slots, which are then plugged into a finite-horizon optimization problem. The solution to this optimization problem is the optimal control for the next time slot. MPC has been used for voltage stability, transmission line thermal control and management of energy by [10], [11], [46]. The authors of [12] have used MPC in order to develop a multi-level VPP control system that coordinates DERs (participating in the VPP) during the event of power mismatch or similar adverse event in the VPP. Vasirani et al. [115] adopt MPC to decide on the operation of a VPP that integrates wind turbines and EVs (with V2G) in an intra-day market. Distributed MPC is used in [14] to coordinate renewable generation in one

control area with storage in another area. To lower the computational cost of MPC, a neural network is trained in [60] to approximate the control policy of an MPC. This neural network is then used in an online fashion to control a solar-plus-battery system. In addition to having high computational costs, a major drawback of the MPC-based approaches is the need for an accurate predictive model. In our problem, the MPC-based approach performs poorly, because market prices, and in particular imbalance prices, are highly variable and depend on various factors that cannot be accurately modeled. Moreover, mispredictions could result in a violation of the charging constraints.

2.2.3 Stochastic Optimization (SO)

Stochastic Dynamic Programming (SDP) and Monte Carlo (MC) methods are two of the techniques that are used to solve SO problems, such as DER control under uncertainty. In SDP, similar to Dynamic Programming, the problem is solved by modeling it using Bellman Equations. The solution to these equations can then be used to compute *optimal* policies (operational strategies) via Bellman's principle of optimality [17]. The SDP method for DER control can be found in [67]. In MC methods, samples from a large number of different scenarios of a system are taken, and probabilistic models for the defined problem are then used to solve for values of the random variables so that eventually, an approximate solution to the problem can be found. The MC method for DER control can be found in [25]. In this thesis, one of our online solution approaches entails using RL, which is also one kind of SO and will be explained in the next paragraph.

2.2.4 Reinforcement Learning (RL)

Different types of Reinforcement Learning (RL) have been used in recent years to solve control tasks in the energy domain that have continuous and high dimensional action spaces [104]. RL algorithms can be broadly classified into two major types: model-based RL algorithms and model-free RL algorithms. A model-based RL algorithm requires an accurate model of the environment with which the RL agent interacts, in order to effectively find the *optimal* policy.

However, in most real-world applications, such as the VPP settings considered in this thesis, it is very difficult to accurately model the environment. Hence, model-free RL methods are advantageous for VPP and DER control because they do not require a separately trained model of the system dynamics. Instead, the RL agent continuously interacts with the environment to learn an optimal policy that governs the operation of DERs. For example, the authors of [86] utilize model-free RL for controlling battery operations and model-free RL control strategies are designed for EV charging in [7], [96]. Model-free RL algorithms can be further classified into action-value methods and policy gradient methods. In the case of action-value methods, the RL agent learns the actions' *values* and then selects an action based on its estimates of the action values. In contrast, the RL agent in policy-gradient methods aims to learn a *parameterized policy* and does not select the action based on the estimated action values. Actor-critic RL algorithms are one of the policy gradient methods that have been used to control the charging of EVs and stationary batteries [7], [86].

All traditional RL algorithms, however, suffer from one major issue during training and when deployed in the real world: due to the agent's inherent need to balance exploration and exploitation during the learning process for finding the optimal policy, they may take actions that violate operational or physical constraints [40]. In the context of controlling DER, there are usually several *hard constraints* that must be satisfied at all times. It is critical to ensure that they are satisfied in the deployment phase, and during training, if it takes place in a real environment rather than a simulator.

Safe RL

Action clipping or mapping the agent's action outside the RL loop using simple bounds can result in safe actions. However, the aforementioned techniques do not ensure the optimality of the resulting action, which in turn, does not ensure the optimal performance of the RL agent. Moreover, since the agent does not *learn* these hard constraints, action mapping would be required whenever the agent is deployed in any environment. There has been growing literature on

the adoption of various safe reinforcement learning techniques. For example, [40], [45], [52] have focused on techniques to learn constraint-satisfying policies to ensure safe actions. Various other techniques have been explored in the safe-RL literature. For example, the authors of [3] utilized a conditional-gradient technique, and authors of [37], and [102] have tweaked the reward function by adding the cost objective's weighted copy to it. Policy optimization using an initial safety set is another approach adopted by [116], [103] and [111]. A safety layer was added to ensure safe actions by [26]. The aforementioned studies in their proposed methods cannot provide any guarantee as to whether their obtained safe action is optimal in nature. Moreover, the RL agent does not *learn* the hard constraints of the system which implies that the agent's actions would require mapping to the safe region even in the deployment phase.

Safe RL Using a Differentiable Projection Layer in a Neural Network

To address the shortcoming of safe-RL techniques, various techniques have emerged in the safe-RL literature [48], [98] that enforce bounds on the agent's actions to satisfy the hard constraints. Bingqing et al. [23] show that a deep reinforcement learning agent trained with a differentiable projection layer embedded in the neural network can safely control inverters and building energy systems. To our knowledge, the application of safe-RL techniques to smart EV charging has not been explored in the literature yet. In this context, the charging deadlines and operational constraints of the battery can be viewed as hard constraints.

2.3 Incentive Mechanisms for DERs

The intermittent nature of renewable power generation calls for the active participation of end users, who are *rational* and self-interested agents. Therefore, incentive mechanisms are needed to attract them to participate in V2G to improve grid operation. A desirable incentive mechanism could maximize either the utility of the VPP operator or a social choice function, thereby increasing the utility of end users while optimizing the VPP's resource utilization and

profitability. We now expand on some of the existing approaches, namely MD and CT, that have been used to design such incentives.

2.3.1 Mechanism Design (MD)

Mechanism Design (MD) is defined as the allocation design of available resources under the assumption that the relevant information of the mechanism's participants is dispersed in the economy [89]. Since the complete information is unavailable to the allocation designer, therefore MD approaches aim to find an allocation technique such that it is in every user's best self-interest to reveal their private information in the mechanism [91]. MD has been used mostly for demand side management and demand response in the context of a smart grid [90].

The authors of [118] design a vehicle-to-aggregator interaction game where the aggregator is modeled as the coordinator and the EV batteries are modeled as independent players who provide frequency regulation service to the power grid by taking charging/discharging decisions. However, a major limitation of this work results from the assumption of a homogeneous setting where different preferences of EVs are not considered for charging/discharging decisions. In [120], a two-level reverse auction is implemented with a group-bidding mechanism where EV owners are incentivized for V2G participation via a feedback-based price scheme. In [57], EV charging and discharging are coordinated via a distributed mechanism that comprises day-ahead scheduling for the aggregator, and distributed coordination and distributed dispatch algorithms for the V2G services from the EVs. References [126] and [121] also use MD for incentivizing EV owners to participate in V2G. However, none of the incentive mechanisms presented in the aforementioned research works [57], [118], [120], [121], [126] discussed how these mechanisms are enforced in practice. That is, there is no penalty in place in the event that the EV agents participating in the mechanism deviate from their charging or discharging decision.

2.3.2 Contract Theory (CT)

Contract Theory (CT) is a principle from microeconomics in which rational and self-interested agents enter into an agreement with the contract provider and the agents receive a payoff (i.e. incentive) for their contribution. The agent contribution and payoff are both stipulated in the contract and any deviation from the stipulated contribution will result in a penalty being imposed on the agent. The agents (who are offered the contract) can be classified into different *types* according to the characteristics that influence their contract acceptance. For example, when providing contracts to EV owners for V2G participation, the EV owner types can depend on factors such as their EV laxity (i.e., maximum duration of time by which charging can be delayed while satisfying EV's charging demand by the deadline) and the EV owners' perceived battery degradation cost. However, owing to the self-centered and rational nature of the participating agents, they may disguise their true type to maximize their profits. Using the revelation principle [28], CT can help in designing contracts that incentivize the participating agents based on their true type under information asymmetry [19]. Each contract specifies the payoff to the participating agents in return for their service and the contracts vary across different types of participating agents. In Chapter 5, we discuss how incentive compatible contracts can be defined.

In [22], the authors consider a 5G network, where the Mobile Virtual Network Operator (MVNO) and Infrastructure Providers (InPs) enter into a contract theoretic agreement. Here, InPs are the employees and the MVNO is the employer. The InPs are provided incentives to provide services to the MVNO's customers. Similarly, in [27], the authors consider a 5G network where a Base Station (BS) wants to use data cached in Augmented Reality (AR) devices. Thus, the BS and AR devices enter into a contract theoretic agreement. Here, AR devices are the employees and the BS is the employer. The AR devices are provided incentives to provide caching service to the BS. In [44], contract theory is used for relay selection and incentive mechanisms in multi-carrier wireless systems under asymmetric information. The source

node is ill-informed of potential relay nodes and their private information such as channel conditions on the relay-destination link. They solve this problem by introducing a principal-agent model for the source and relays. In [61], the authors consider a vehicular network where a Road Side Unit (RSU) equipped with a Multi-access Edge Computing (MEC) server wants to offload its tasks to a set of neighboring EVs. Here, EVs are the employees and the RSU is the employer. The EVs are provided incentives to provide task computation service to the RSU. In [124], the authors consider the case where a Base Station (BS) needs to offload some of its traffic via Device-to-Device (D2D) communication to some User Equipment (UE). Here, UE devices are the employees and the BS is the employer. The UE devices are provided incentives to provide task computation service to the BS.

2.3.3 Contract Theory Applications for V2G

References [39], [53], [123] use contract theory to provide incentives to EV users for V2G participation. In particular, the authors in [123] use contract theory to incentivize EVs in a cloudlet-based Vehicle-to-Vehicle (V2V) setting where an external broker agent offers contracts to discharging EVs and resells the discharged energy to the EVs in demand of energy. However, the contracts designed in [123] may be in infeasible at certain times and therefore, a separate algorithm is used to re-design the infeasible contracts, creating a computational overhead. In [53], a game theoretic approach (using a non-cooperative Stackelberg game) was used to find the optimal pricing for discharging EVs and their discharging strategy and a contract theory-based approach was presented in order to incentivize EV users to participate in V2G. In [39], the authors present an algorithm that learns the optimal unit price (to incentivize owners to participate in V2G) using past interactions of the EV owners with the aggregator for V2G participation using a modified version of the Upper Confidence Bound (UCB) algorithm [104]. However, the authors do not incorporate the EV owners' perceived battery degradation cost in their algorithm design which can lead to sub-optimal performance of the algorithm. In Chapter 5, we attempt to overcome all of the aforementioned limitations in the

existing contract theory literature for V2G participation.

Chapter 3

Virtual Power Plant Model

In this chapter, we design a 2-staged Virtual Power Plan (VPP) operational method that enables it to overcome the inherent challenges of operating a VPP that participates in a 2-stage electricity market.

The VPP setting that we study consists of a fleet of EVs (all of which are owned by the VPP operator), solar systems (only in Chapter 4), and bidirectional chargers distributed across the city. The EVs visit charging stations, each containing multiple chargers, at random times to replenish their battery. They stay there for a certain amount of time before they start their next trip. This determines their charging deadline. We assume bidirectional chargers and solar inverters do not cause overloading or voltage violation problems, and consequently, ignore the grid constraints when optimizing the VPP operation. Moreover, we assume that the VPP operator is not certain about the arrival time, charging deadline, and energy demand of an EV before it arrives at a charging station. However, once an EV arrives, it communicates its energy demand and departure time (or deadline) to the VPP. It is also assumed that the VPP operator allows a subset of the EV fleet to participate in V2G, for example, they can be the EVs whose battery has a high remaining cycle life.

Because of the DERs that comprise the VPP, its demand and supply are both variable and flexible. The VPP operator trades (*i.e.*, buys or sells) energy in the wholesale day-ahead (DA) market, which is a pool-based energy market. The players in this market, place separate energy-price bids for every hour of

the next day, and each hour has an independent auction. DA markets are typically cleared based on the uniform pricing mechanism, hence the same (clearing) price applies to any market player (seller or buyer) whose bid is accepted. We further assume that the VPP operator is a price taker, which is reasonable given the size of a typical DA market today. Since the marginal cost of supplying power by the VPP is much lower than conventional generators that typically govern the clearing prices, the VPP's energy-price bid reduces to an energy (quantity) bid as the corresponding price does not affect the price it receives. This is a practical assumption for battery operators and aggregators in electricity markets [20], [59]. With this assumption, we can treat the DA prices as exogenous random variables, denoted by $\mathbf{P}^{DA} = [P_0^{DA}, \cdots, P_{23}^{DA}]$, and postulate that the VPP operator only bids for quantity (and not for price). At the time of operation, the DA market players might deviate from their DA commitments. Any such deviation must be financially settled through the imbalance (IM) market. The IM market prices reflect the additional costs incurred to serve the unexpected demand beyond wholesale commitments. As such, these prices are treated as exogenous random variables too. We assume the IM market adopts the single-pricing (aka one-pricing) model [36] – the most prevalent pricing scheme in the imbalance market today. In this model, selling and buying prices in each hour t are identical, denoted by P_t^{IM}. The vector $\mathbf{P}^{IM} = [P_0^{IM}, \cdots, P_{23}^{IM}]$ represents IM prices for one day.

In the DA bidding stage, the VPP operator does not deterministically know the market prices, available solar energy every hour of the next day, and the EV arrival and demand patterns on the next day. It must submit a vector of energy bids $\mathbf{X} = [x_0, \cdots, x_{23}]$ to the market, where x_t is positive when the VPP commits to sell energy in hour t of the next day and is negative when the VPP commits to buying energy in that hour. On the operation day, the VPP must schedule the charge and discharge of EVs that arrive at the charging stations. We denote this schedule by $\mathbf{Y}^n = [y_0^n, \cdots, y_{23}^n]$ with n being the index into the set of EVs that arrive the next day ($n \in \mathcal{N} = \{1, 2, ..., N\}$). Hence, every element y_t^n represents the amount of energy stored in (positive sign) or withdrawn from (negative sign) the battery of the n^{th} EV in hour t.

When EV n is not in a charging station, the respective elements of $\mathbf{Y}^n$ are set to zero.[1] Note that for a given vector $\mathbf{X}$ and a set of vectors $\mathbf{Y}^n$ ($\forall n \in \mathcal{N}$), the amount of energy that must be traded in the IM market, denoted by $\mathbf{Z} = [z_0, \cdots, z_{23}]$, can be computed from the energy balance equation, which states that the VPP's demand and supply must be equal in each hour. The VPP operator will buy enough energy from the two markets to satisfy the energy demand of all EVs by their deadlines. Additionally, the operator can sell energy (solar production and/or energy discharged from the batteries) in the markets. These decisions must be made so as to maximize the expected profit of the VPP operator. Energy bids are submitted to the DA market all at once on the day before the operation day, whereas (dis)charging and trading decisions in the IM market are made for every hour in an online fashion.

3.1 Notation

Let $\mathbf{E}^{solar} = [E_0^{solar}, \cdots, E_{23}^{solar}]$ be the available solar energy in every hour of the operation day, E_{max}^{solar} be the total peak generation capacity of the solar systems, $\mathcal{N}_t$ be the set of EVs that are connected to a charger owned by the VPP in hour t of the operation day, and $\mathcal{N} = \cup_{t \in \mathcal{T}} \{N_t\}$ be the set of all EVs that visit the charging stations on the operation day. Let $\mathcal{N}^D \subseteq \mathcal{N}$ denote the set of EVs that participate in V2G; hence, the operator can discharge their battery as long as it is possible to meet their charging demand by the deadline. We denote the EVs in $\mathcal{N}^D$ that are connected to chargers in hour t of the operation day by $\mathcal{N}_t^D \subseteq \mathcal{N}_t$. For the EV indexed by n, we denote the energy capacity of its battery by b_n, its arrival time by t_s^n, and the length of its charging session by τ^n. Hence, its departure time will be $t_e^n = t_s^n + \tau^n$. The maximum charge and discharge power supported by a charger are denoted by $\alpha_c, \alpha_d > 0$, respectively. We assume these maximum rates are the same for all chargers. Since the length of each time slot is one hour, we use the same notation (α_c, α_d) to represent the maximum amount of energy that can be

[1] Without loss of generality, we assume EV arrivals and departures occur in the beginning of a 1-hour time slot as it is the timescale of trading in the IM market.

stored or withdrawn from a battery in a single time slot. The state-of-charge (SoC) of the battery of EV n at time $t \in \mathcal{T}$ is denoted by SoC_t^n, which is bounded by δ_{min} and δ_{max}. The charge and discharge efficiencies are denoted by η_c and η_d, respectively. Given the energy demand of the EV and its SoC upon arrival, denoted by $\underline{SoC}^n$, we calculate the target SoC, which is denoted by $\overline{SoC}^n$.

Let $\mathbf{P}^{DA}, \mathbf{P}^{IM}, \mathbf{E}^{solar}, \mathbf{t}_s, \mathbf{t}_e, \underline{\mathbf{SoC}}, \overline{\mathbf{SoC}}$ be random vectors that collect the random variables and Ω is the set of possible realizations of these random variables.

3.2 Two-Staged Approach to VPP Operation

The optimal VPP operation is the solution of the stochastic optimization in Problem (3.1). However, solving this problem via sample-average approximation (SAA) [62] is overly costly due to a large number of random variables that appear in the objective function and some of the constraints, leading to a huge optimization problem even for a small number of samples.

$$
\begin{aligned}
\underset{\mathbf{X},\ \mathbf{Y},\ \mathbf{Z}, AC, AD}{\text{maximize}} \quad & \mathbb{E}_{<\mathbf{P}^{DA}, \mathbf{P}^{IM}, \mathbf{E}^{solar}, \mathbf{t}_s, \mathbf{t}_e, \underline{\mathbf{SoC}}, \overline{\mathbf{SoC}}> \sim \Omega} \quad \mathbf{X}^\top \mathbf{P}^{DA} + \mathbf{Z}^\top \mathbf{P}^{IM} && \text{(3.1a)} \\
\text{subject to} \quad & -|\mathcal{N}_t| \cdot \alpha_c \leq x_t \leq E_{max}^{solar} + |\mathcal{N}_t^D| \cdot \alpha_d, \quad \forall t \in \mathcal{T} && \text{(3.1b)} \\
& x_t + z_t + y_t = E_t^{solar}, \quad \forall t \in \mathcal{T} && \text{(3.1c)} \\
& y_t = \sum_{n \in \mathcal{N}} y_t^n, \quad \forall t \in \mathcal{T} && \text{(3.1d)} \\
& y_t^n = AC_t^n + AD_t^n, \quad \forall n \in \mathcal{N}_t, \forall t \in \mathcal{T} && \text{(3.1e)} \\
& AD_t^n = 0, \quad \forall n \in \mathcal{N}_t \setminus \mathcal{N}_t^D, \forall t \in \mathcal{T} && \text{(3.1f)} \\
& -\alpha_d \leq AD_t^n \leq 0, \quad \forall n \in \mathcal{N}_t^D, \forall t \in \mathcal{T} && \text{(3.1g)} \\
& 0 \leq AC_t^n \leq \alpha_c, \quad \forall n \in \mathcal{N}_t, \forall t \in \mathcal{T} && \text{(3.1h)} \\
& SoC_{t+1}^n = SoC_t^n + \frac{AC_t^n \eta_c}{b_n} + \frac{AD_t^n}{\eta_d b_n}, \forall n \in \mathcal{N}_t, \forall t \in \mathcal{T} && \text{(3.1i)} \\
& \delta_{min} \leq SoC_t^n \leq \delta_{max}, \quad \forall n \in \mathcal{N}_t, \forall t \in \mathcal{T} && \text{(3.1j)} \\
& SoC_{t_s^n}^n = \underline{SoC}^n, \quad \forall n \in \mathcal{N}_t && \text{(3.1k)} \\
& SoC_{t_e^n}^n = \overline{SoC}^n, \quad \forall n \in \mathcal{N}_t && \text{(3.1l)}
\end{aligned}
$$

Recall that x_t is the DA energy bid placed for hour t of the next day and z_t is the amount of energy that would be traded in the IM market in that hour. The sign of z_t determines whether the VPP operator buys or sells in the IM

market: positive implies selling and negative implies buying. Constraint (3.1b) is an operational constraint that defines the DA hourly selling and buying bid caps for the VPP operator. We assume that the selling bid cap in any hour is the sum of the peak solar generation capacity and the maximum amount of energy that can be discharged from the connected EVs that participate in V2G in that hour. Similarly, the buying bid cap in any hour is the maximum amount of energy that can be stored in the connected EVs in that hour. Constraint (3.1c) is the energy balance equation and constraint (3.1d) expresses the total charging demand as the sum of the demands of individual chargers within charging stations. Constraint (3.1e) splits the contribution of each EV n to the total charging demand in time slot t into two parts: energy stored in its battery AC_t^n, and energy is withdrawn from its battery AD_t^n. This is necessary as charge and discharge efficiencies can be less than 1 in Constraint (3.1i). Constraints (3.1g)-(3.1h) set bounds for the amount of energy that can be stored or withdrawn from the battery of an EV. These bounds depend on the maximum charge and discharge rates supported by the chargers. As shown in Constraint (3.1f), for the EVs that do not participate in V2G, the amount of energy that can be withdrawn from their battery is set to zero. Constraint (3.1i) updates the SoC of each EV according to the amount of energy stored or withdrawn from its battery in the previous time slot and the respective inefficiency parameter. Finally, constraints (3.1j)-(3.1l) define bounds for the SoC of each EV and assign a value to it at arrival and departure time. Observe that all constraints are affine in Problem (3.1).

Remark 1. We do not need to introduce binary variables to ensure that AC_t^n and AD_t^n are not nonzero at the same time. This is because due to battery imperfections, a strategy that simultaneously charges and discharges the battery of one or multiple EVs will either increase the deficit energy that must be purchased from the IM market or decrease the surplus energy that could be sold in the IM market. Either way, assuming that IM prices are nonnegative, this strategy reduces the profit of the VPP and is therefore not optimal.

We decompose Problem (3.1) into two subproblems, namely stage 1 and

stage 2. The stage 1 problem determines the optimal bidding strategy in the DA market. This is not a sequential decision-making problem as hourly bids are submitted all at once. The stage 2 problem entails finding feasible schedules for charging or discharging the EVs as they arrive at the charging stations, while trading in the IM market to close the gap between day-ahead bids and the realized solar generation minus the total realized EV charging demand. This can be cast as a sequential decision-making problem. We formally define the two problems below:

3.2.1 Stage-1: Day-Ahead Energy Bidding

Given different forecasts for (hourly) DA and IM market prices, hourly solar production, EV arrivals, stay times, and energy demands for every hour t of the next day, where $t \in \mathcal{T} = \{0, 1, , ..., 23\}$, the goal is to compute day-ahead energy bids, *i.e.*, $\mathbf{X} = [x_0, \cdots, x_{23}]$, that maximize the expected profit of the VPP as a result of participating in both markets. This problem is solved in an offline fashion, typically in the beginning of the day before the operation day.

3.2.2 Stage-2: Real-Time Energy Trading and EV Charging

Given the energy bids submitted to the DA market the day before (*i.e.*, the stage 1 solution), the current price of the IM market and IM market price forecast for every hour until the end of the day, the deadline and unmet energy demand of the EVs that are currently in the charging stations, and forecast data for EV arrivals, stay (sojourn) times, and energy demands in future time slots, this problem concerns determining the amount of energy that must be charged/discharged in/from the EV batteries in the current time slot to maximize the expected profit of the VPP, such that the energy demand of every EV is guaranteed to be satisfied by their deadline. In other words, the stage 2 problem concerns determining y_t^n for every EV that is currently present in the charging stations while ensuring that it is possible to fulfill their demand before their departure. Once these values are fixed in hour t, the net difference

between demand and supply (z_t) will be traded in the IM market according to its current price.

Remark 2. In Chapter 5, we study a different VPP where we assume the EVs are under independent ownership, and the VPP comprises charging stations with vehicle-to-grid (V2G) support, but it does not own a solar farm. In this chapter, our primary focus is on the design of contracts rather than addressing the uncertainty caused by highly variable solar generation. This VPP also participates in the 2-stage electricity market in the same manner and hence we use the same problem formulation for both of the settings. We solve an additional challenge in the latter setting that pertains to motivating the EV owners to participate in V2G, by developing an appropriate incentive mechanism. Since the EV battery life would be reduced due to the additional charge-discharge cycles when participating in V2G, the EV owners need to be incentivized according to their perceived battery degradation cost. This is necessary to ensure their participation in the VPP.

3.3 Contracts for V2G Participants

We denote the V2G-participating EVs that have connected to the charger at hour t by $\mathcal{N}_t^D \subseteq \mathcal{N}_t$, where $\mathcal{N}_t$ denotes all the EVs (both V2G participating and non-V2G participating) that are connected to the charger at hour t. We represent the contracts offered to EV owners as energy-reward bundles, represented as $\{(g_m, w_m)\}_{m=1\cdots M}$, where each contract is designed for a specific type (denoted m). In the contract, w_m is the discharge energy that the EV owners contribute in V2G and g_m is the associated payoff to the EV owners accepting the contract. The hourly charging schedule for V2G participating EVs is denoted using the variable d_t and l_{V2G} represents the duration of the contract, which is the maximum duration of V2G participation of the EV owners.

Chapter 4

Operating Strategies

In this chapter, we show what algorithms can be used to solve the stochastic decision-making problem (under uncertainty and risk) that was introduced in the previous chapter. Specifically, we show that the stage 1 problem can be solved by an approximation technique that involves solving a sequence of deterministic LPs formulated for different forecast scenarios (*i.e.*, different hourly market prices, solar production levels, EV arrival times, and charging demands). Given the solution to the stage 1 problem, *i.e.*, the DA energy bids, we propose two algorithms to solve the stage 2 problem. Both utilize observations up to the current time slot of the operation day and make real-time decisions for (dis)charging EVs, and subsequently trading in the IM market. The decisions are made while ensuring that the charging deadlines can be met. Finally, we evaluate the profitability of this VPP under the proposed operating strategy and show how different levels of V2G participation by EV owners affect VPP profitability.

4.1 Stage-1 Problem: Linear Programming using Wait-and-See

The stage-1 solution method entails placing energy bids in the DA market by solving a number of linear programs for different forecast scenarios and taking the average of the solutions. This is called the wait-and-see (WS) approach in the field of stochastic programming. The WS approach [69] is computationally efficient and to solving a stochastic program with a large number of decision

variables is tractable, it yields a good approximation to the solution of the original stochastic program and provides a practical approach to handling a large number of decision variables and scenarios.

The WS approach approximates the solution of this stochastic linear program[1] that was introduced in the previous chapter. In this approach, we consider a large number of forecast scenarios and formulate a deterministic linear program (written below) for each forecast scenario $\omega \in \Omega$:

$$\underset{\mathbf{X},\,\mathbf{Y},\,\mathbf{Z},AC,AD}{\text{maximize}} \quad \mathbf{X}^\top \mathbf{P}^{DA} + \mathbf{Z}^\top \mathbf{P}^{IM} \tag{4.1a}$$

$$\text{subject to} \quad -|\mathcal{N}_t| \cdot \alpha_c \leq x_t \leq E_{max}^{solar} + |\mathcal{N}_t^D| \cdot \alpha_d, \quad \forall t \in \mathcal{T} \tag{4.1b}$$

$$x_t + z_t + y_t = E_t^{solar}, \quad \forall t \in \mathcal{T} \tag{4.1c}$$

$$y_t = \sum_{n \in \mathcal{N}} y_t^n, \quad \forall t \in \mathcal{T} \tag{4.1d}$$

$$y_t^n = AC_t^n + AD_t^n, \quad \forall n \in \mathcal{N}_t, \forall t \in \mathcal{T} \tag{4.1e}$$

$$AD_t^n = 0, \quad \forall n \in \mathcal{N}_t \setminus \mathcal{N}_t^D, \forall t \in \mathcal{T} \tag{4.1f}$$

$$-\alpha_d \leq AD_t^n \leq 0, \quad \forall n \in \mathcal{N}_t^D, \forall t \in \mathcal{T} \tag{4.1g}$$

$$0 \leq AC_t^n \leq \alpha_c, \quad \forall n \in \mathcal{N}_t, \forall t \in \mathcal{T} \tag{4.1h}$$

$$SoC_{t+1}^n = SoC_t^n + \frac{AC_t^n \eta_c}{b_n} + \frac{AD_t^n}{\eta_d b_n}, \forall n \in \mathcal{N}_t, \forall t \in \mathcal{T} \tag{4.1i}$$

$$\delta_{min} \leq SoC_t^n \leq \delta_{max}, \quad \forall n \in \mathcal{N}_t, \forall t \in \mathcal{T} \tag{4.1j}$$

$$SoC_{t_s^n}^n = \underline{SoC}^n, \quad \forall n \in \mathcal{N}_t \tag{4.1k}$$

$$SoC_{t_e^n}^n = \overline{SoC}^n, \quad \forall n \in \mathcal{N}_t \tag{4.1l}$$

We solve these deterministic linear programs independently and take the average of the respective solutions to efficiently compute near-optimal DA bids. We discard the tentative smart charging and IM market trading schedules because they will be recalculated in an online fashion (in stage 2), using more accurate data.

To create different forecast scenarios, we add white Gaussian noise with standard deviation being 10% of the realized value to the realized value of each random variable, namely $\mathbf{P}^{DA}$, $\mathbf{P}^{IM}$, $\mathbf{E}^{solar}$, $\mathbf{t}_s$, $\mathbf{t}_e$, $\underline{\mathbf{SoC}}$, $\overline{\mathbf{SoC}}$. We have found empirically that considering 1,000 forecast scenarios $(\omega_1, \cdots, \omega_{1000})$ provides

[1] It can be proved that WS yields a bound on SAA [62] as it interchanges the order of summation and maximization.

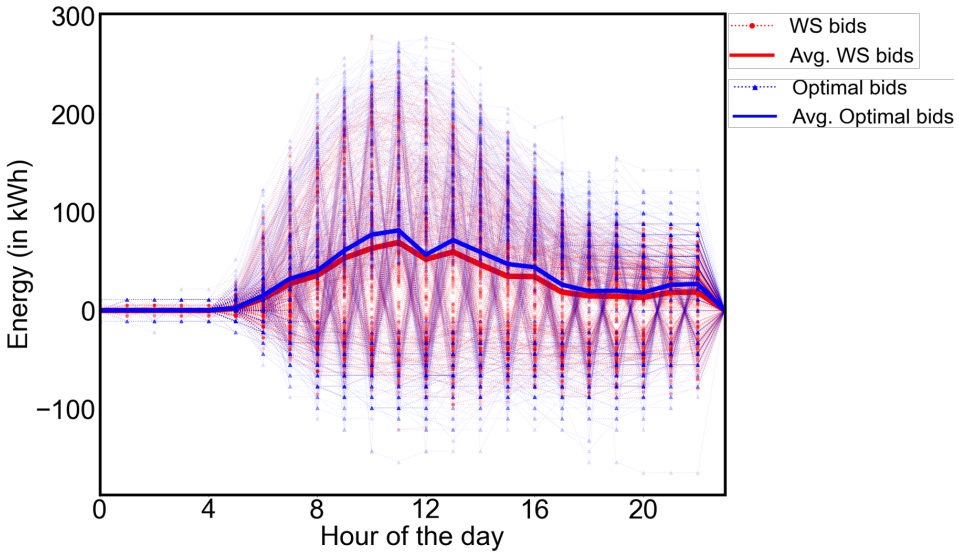


Figure 4.1: Comparison between the optimal DA bids (i.e. bids obtained with accurate knowledge of all random variables) and DA bids computed using the WS approach over one year.

a good-quality solution[2] and the total running time of solving 1,000 deterministic LPs is less than 15 minutes on an Intel core-i9 server with 128GB of memory. Once the deterministic LPs are solved, we take the average of the respective solutions and treat this as the energy bids that will be submitted to the DA market. Figure 4.1 compares the optimal energy bids (assuming perfect information) with the energy bids computed efficiently using the WS approach. It can be seen that the difference between the average hourly energy bids is generally insignificant.

Note that the objective function of Problem (4.1) is linear and its constraints are affine. We model this deterministic LP in CVXPY [31] and solve it using Gurobi [42]. Next, we propose two algorithms for solving the stage 2 problem given the DA energy bids, which is the solution of the stage 1 problem.

[2]Our experiment shows that the resulting DA bids do not vary noticeably if we consider more forecast scenarios. We omit the convergence analysis to save space.

4.2 Stage-2 Problem: Laxity-Lookahead Algorithm (LLA)

We propose two profit-maximizing strategies for the VPP considered in our work to solve the stage-2 problem. Both strategies place energy bids in the DA market by the WS approach. To hedge against the uncertainty of solar generation and satisfy the charging constraints, a decision-making problem is solved in an online fashion to (dis)charge the connected EVs and trade energy in the IM market. One strategy solves this problem using a heuristic algorithm, while the other adopts a policy learned via reinforcement learning. Nevertheless, they both perform a *laxity lookahead* to ensure that the problem remains feasible if they take a specific action at the present time.

4.2.1 Laxity LookAhead

The first algorithm we propose to solve the stage 2 problem is a heuristic algorithm, called Laxity-LookAhead (LLA). When we run LLA for hour t of the operation day, it computes the laxity of every EV that is currently in a charging station and uses this value to find out if the charging demand of this EV will be satisfied before its departure. This is essential for constraint enforcement.

The laxity of an EV is defined as the maximum amount of time we can delay its charging, while still being able to charge its battery to the desired SoC by the deadline. Specifically, the laxity of EV n, with departure time t_e^n, battery size b_n, and target SoC, $\overline{SoC}^n$, in time slot t is:

$$lax_t^n = t_e^n - t - \frac{(\overline{SoC}^n - SoC_t^n) \cdot b_n}{\alpha_c \eta_c}. \tag{4.2}$$

Note that the laxity of an EV can be calculated deterministically in any time slot after its arrival, because its deadline and energy demand are communicated to the VPP upon arrival. Nevertheless, the EV's laxity is unknown before it arrives. The basic idea of this algorithm is to identify all EVs that will have a *negative* laxity in the next time slot if they are not charged in the current time slot. These EVs must be charged at the maximum charge power supported by

Algorithm 1: LLA for EV charging

```
1  S_1 ← FindEVsWithNegativeLaxity(N_t) ;                    // lookahead
2  e_t ← ChargeEVs(S_1);
3  x_t ← x_t + e_t;
4  if x_t > E_t^solar then
5  |   S_2 ← FindEVsToDischarge(x_t − E_t^solar);
6  |   e_t ← DischargeEVs(S_2);
7  |   x_t ← x_t - e_t;
8  |   BuyFromImbalanceMarket(x_t − E_t^solar);
9  end
10 else if x_t ≤ E_t^solar then
11 |   S_3 ← FindEVsToCharge(E_t^solar − x_t);
12 |   e_t ← ChargeEVs(S_3);
13 |   x_t ← x_t + e_t;
14 |   SellToImbalanceMarket(E_t^solar − x_t);
15 end
```

the charger, otherwise the problem becomes infeasible. Next, depending on whether there is surplus solar energy in this time slot, other EVs are charged or discharged.

4.2.2 LLA Algorithm

The LLA algorithm utilizes three main functions (see Algorithm 1). In Line 1, the FindEVsWithNegativeLaxity($\mathcal{N}_t$) function returns the set of EVs that are presently connected to a charger and will have negative laxity in the next time slot if they are not charged in the current time slot. This set is denoted by $\mathcal{S}_1$ and is determined via a one-step laxity lookahead. Concretely, to calculate the laxity of an EV in the next time slot, we substitute t with $t+1$ in Equation (4.2) and let SoC^n_{t+1} be equal to SoC^n_t. The set $\mathcal{S}_1$ is then passed to the ChargeEVs($\mathcal{S}_1$) function (Line 2), which is responsible for charging these EVs at the maximum power supported by the charger. Once these EVs are charged, we add the energy delivered to these EVs to the day ahead commitment for this time slot, x_t, to update the amount of energy required in this time slot (Line 3). This value is then compared with the available solar energy in this time slot, E^{solar}_t. If the available solar energy is not enough to supply the demand (Line 4), we must discharge a subset of EVs or buy

the deficit from the IM market. Otherwise (Line 10), we can charge a subset of EVs or sell the surplus in the IM market. LLA uses a simple heuristic to specify the order in which we use smart charging and trade in the IM market in both cases.

In the case that $x_t > E_t^{solar}$, the FindEVsToDischarge($x_t - E_t^{solar}$) function gets the amount of deficit and returns the set of EVs, denoted by $\mathcal{S}_2$, that (a) participate in V2G, (b) have the highest laxity, and (c) their laxity will not become negative in the next time slot if we discharge them at the maximum power supported by the charger in this time slot by calling DischargeEVs($\mathcal{S}_2$). If there is not enough EVs in $\mathcal{S}_2$ to cover the deficit, we buy the remainder from the IM market according to the current market price (Line 8). This gives z_t. In the case that $x_t \leq E_t^{solar}$, the FindEVsToCharge($E_t^{solar} - x_t$) function gets the amount of surplus and returns the set of EVs, denoted by $\mathcal{S}_3$ not intersecting with $\mathcal{S}_1$, that have the lowest laxity. These EVs are charged at the maximum power in this time slot by calling ChargeEVs($\mathcal{S}_3$). If there is not enough EVs in $\mathcal{S}_3$ to absorb the surplus energy, we sell the remainder to the IM market according to the current market price (Line 14). This gives z_t.

Since LLA is an online algorithm, it does not assume the knowledge of the available solar energy and IM market prices in the next time slots of the day, and future EV arrival times, stay times, and energy demands. Moreover, it does not find the optimal EV charging strategy because regardless of future charging demands and market prices it always prioritizes (a) discharging EVs with highest laxity that will not have negative laxity in the next time slot over buying the deficit energy from the IM market; (b) charging EVs with lowest laxity over selling the surplus energy to the IM market.

4.3 Stage-2 Problem: Laxity Aware-Soft Actor Critic (LA-SAC)

The second algorithm we propose to solve the stage 2 problem is a model-free reinforcement learning algorithm that respects the EV charging deadlines.

This algorithm is designed based on SAC [43] and is called Laxity Aware-Soft Actor Critic (LA-SAC). It borrows the notion of laxity from LLA and incorporates one-step laxity lookahead in the differentiable projection layer embedded in the actor network. In this section, we describe the Reinforcement Learning (RL) framework, the algorithms used, the MDP formulation and we explain how the projection layer is used to ensure safe exploration and convergence to the optimal policy.

4.3.1 Reinforcement Learning

In RL, interactions between a decision-making agent and its surrounding environment are modelled as a Markov Decision Process (MDP) $(\mathcal{S}, \mathcal{A}, R, P, \gamma, \mu)$, where $\mathcal{S}$ is the set of states, $\mathcal{A}$ is the set of actions, R is the reward function, $P : \mathcal{S} \times \mathcal{A} \to \Delta(\mathcal{S})$ describes the next state distribution as a function of the current state and action, $\gamma \in [0, 1]$ is the discount factor, and $\mu : \Delta(\mathcal{S})$ is the distribution of starting states. Here, $\Delta(\mathcal{S})$ denotes the probability simplex over states. The interaction between the RL agent and the environment is as follows: at step t, the agent observes a state vector s_t and selects an action a_t according to a policy $\pi(\cdot|s_t)$, which is used to interact with the environment. The environment returns a reward r_{t+1} and the next state s_{t+1} to the agent. The goal of the RL agent is to learn a policy that maximizes its expected discounted cumulative reward (or return). We define the return at time t as $G_t \doteq \sum_{k=0}^{\infty} \gamma^k R_{t+k+1}$, and the value function $v^\pi(s)$ as the expected return when starting at s and following the policy π, given by: $v^\pi(s) \doteq \mathbb{E}_\pi \left[G_t \mid S_t = s \right]$. When the state space is continuous or large, the state value function is approximated as $v^\pi(s) \approx v_w(s)$, where the vector $w \in \mathbb{R}^d$ collects the parameters of this approximation. When deep neural networks are used for function approximation, we refer to it as *deep reinforcement learning*.

Actor-Critic Methods They are RL methods that learn the approximation to the policy in addition to the approximation to the value function. The *actor* refers to the policy and the *critic* refers to the value function learned by the agent. The main advantage of actor critic methods is their ability to tackle

continuous action spaces. In our work, we use the Soft-Actor Critic (SAC) algorithm [43], which is an off-policy, maximum-entropy policy gradient method, using a stochastic actor. We explain the SAC algorithm in the next Section.

4.3.2 Soft Actor Critic (SAC)

The SAC algorithm [43] is a policy gradient algorithm which is more suitable for tackling continuous action space problems. In this case, a parameterized function (*i.e.*, a neural network) denoted π_θ, represents the policy of the RL agent. The policy parameter is updated towards the gradient direction of a performance function $J(\theta)$, as follows: $\theta \leftarrow \theta + \gamma \nabla_\theta J(\theta)$. The SAC algorithm encourages further exploration of the agent by incorporating an entropy measure in the reward function. The objective is to maximize both the entropy and expected return. The performance function is given by: $J(\theta) = \sum_{t=1}^{T} \mathbb{E}_{(s_t,a_t)\sim\rho_\pi}[r_t + \nu\mathcal{H}(\pi(.|s_t))]$.

Here, ρ_π denotes the marginal distribution for the state-action pairs (s_t, a_t) sampled from the policy π. The entropy measure and entropy importance are denoted by $\mathcal{H}$ and ν, respectively. We refer the readers to [43] for further details about SAC, including the policy and value functions along with their gradients. To enable the RL agent to do safe exploration and learn the hard constraints, we redefine the loss function as explained in the next section. The Adam optimizer is used with a learning rate of 0.0001, the discount factor is set to be 0.99, and the batch size is set to 72. We allow automatic entropy tuning, which automatically balances exploitation and exploration for the agent.

4.3.3 Markov Decision Process (MDP) formulation

State The state at time t, is a tuple s_t, where the first state variable is the moving average of solar generation for the particular time slot t in the past 3 days, denoted $\bar{E}_t^{solar}$. We use the moving average of the previous 3 days as one of the state variables to capture the diurnal pattern of solar generation and help the (model-free) RL agent to implicitly learn this temporal relationship. The second and third variables are the moving average of DA and IM market prices for the particular time slot t in the past 7 days, denoted $\bar{p}_t^{DA}$ and $\bar{p}_t^{IM}$

respectively. Although market prices change drastically over the course of the day, incorporating the average of previous-day market prices (in the same hour) could help the agent learn temporal patterns that might exist in the two markets. The rest of the state variables are the laxity and SoC level of each of the EVs present in the current time slot in a charging station, before the RL agent's (dis)charging action is implemented.

$$s_t = (\bar{E}_t^{solar}, \bar{p}_t^{DA}, \bar{p}_t^{IM}, \{lax_{t-1}^n\}_{n\in\mathcal{N}_t}, \{SoC_{t-1}^n\}_{n\in\mathcal{N}_t}) \tag{4.3}$$

Action The action space in continuous and multidimensional. Specifically, the action taken by the RL agent for every hour t forms a vector $a_t = [y_t^0, \cdots, y_t^{|\mathcal{N}_t|-1}]$ where y_t^n is the charge or discharge decision taken for the n^{th} EV that is present in one of the charging stations in this hour and satisfies the following conditions:

$$-\alpha_d \leq y_t^n \leq \alpha_c \qquad \forall n \in \mathcal{N}_t^D, \tag{4.4}$$

$$0 \leq y_t^n \leq \alpha_c \qquad \forall n \in \mathcal{N}_t \setminus \mathcal{N}_t^D. \tag{4.5}$$

Reward The reward obtained by the RL agent in hour t, denoted by r_t, is a scalar value calculated based on the profit that will be generated by taking actions that charge or discharge the EVs feasibly and trade the surplus or deficit in the IM market according to its current price. To help the agent in taking actions that result in the maximum return, we reward the agent as follows

$$r_t = z_t \cdot p_t^{IM} + \zeta \cdot \bar{lax}_t. \tag{4.6}$$

The first term is the immediate reward received by the agent. It depends on how much energy is traded in the IM market (*i.e.*, $z_t = E_t^{solar} - x_t - \mathbf{1}^\intercal a_t$) and the current market price. The second term is the average laxity of all the EVs that are present at the charging stations at time t, denoted as $\bar{lax}_t$, and a positive scaling factor ζ ($\zeta = 2$ in our experiments). As discussed earlier, the laxity of an EV characterizes charging flexibility. Thus, the second term is included to incentivize the agent to take actions that do not significantly lower this flexibility on average.

4.3.4 Safe Reinforcement Learning (safe-RL) using Differentiable Projection Layer

The RL algorithms could violate the hard constraints of the actuator or the environment during training (as they explore the space) and/or after deployment. Clipping or hand-crafted projection of the action taken by the RL agent to a *safe* region (aka the feasible set), can prevent the violation of hard constraints when the policy is used in practice. However, this does not guarantee that the resulting safe action is *optimal* too. Furthermore, the RL agent does not actually *learn* the hard constraints, hence the actions must be mapped to the safe region even after deployment.

To make the RL agent effectively *learn* the hard constraints without sacrificing optimality, we embed an optimization problem [5] that projects a point onto the safe region, in the neural network[3] used for approximating the policy function in the actor-critic method. The optimization problem (described below) is a differentiable layer within the actor-network, allowing the RL agent to learn the hard constraints through backpropagation. Consider the ℓ_2-norm projection $\mathcal{P}_{\mathcal{S}} : \mathbb{R}^n \to \mathcal{S}$ which maps a point in $\hat{a} \in \mathbb{R}^n$ to the point closest to it in a constraint set $\mathcal{S} \subseteq \mathbb{R}^n$, as shown below:

$$\mathcal{P}_{\mathcal{S}}(\hat{a}) = \underset{a \in \mathcal{S}}{\operatorname{argmin}} \|a - \hat{a}\|_2^2 \quad . \tag{4.7}$$

This is a convex optimization problem if $\mathcal{S}$ is a convex set. Thus, it can be solved using a standard solver, which constitutes the forward procedure in the neural network. The backward procedure is constructed using implicit function theorem [63] where the gradients of the solution variables of Equation (4.7) are calculated by differentiating the Karush-Kuhn Tucker (KKT) conditions at the solution. The respective gradients are then propagated back to the neural network. As a result, the agent is aware of the hard constraints in $\mathcal{S}$ through the parameter updates done when training the neural network.

Figure 4.2 illustrates how of an optimization problem embedded within the neural network as an implicit layer can help the neural network learn the hard

[3]The optimization problem can be implemented as a neural network layer using a Python library called `cvxpylayers`: `https://github.com/cvxgrp/cvxpylayers`.

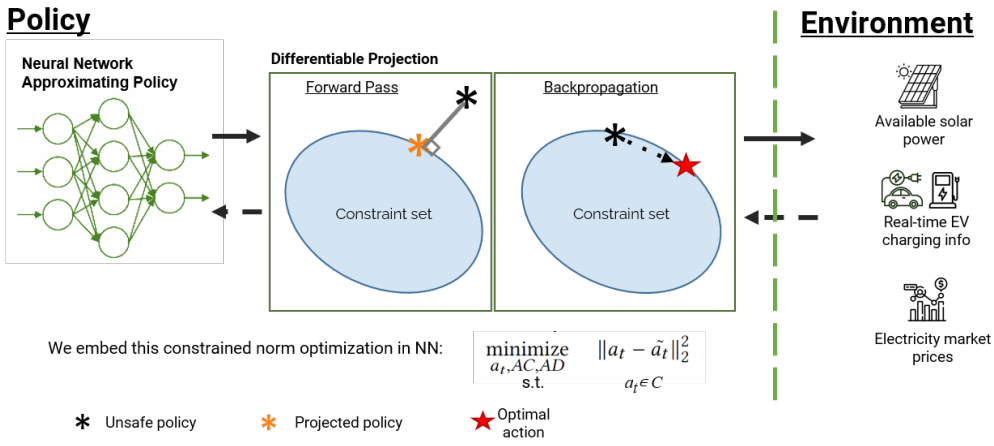


Figure 4.2: Illustration of how a neural network's unsafe actions are projected into the safe (constrained) region and how the neural network *implicitly* learns these hard constraints. Solid lines represent the forward pass and the dotted lines represent the backpropagation steps.

constraints of the optimization problem and thereby output *safe* actions.

4.3.5 LA-SAC Algorithm

The action is selected according to $\hat{\pi_\theta}$, which is the neural network that approximates the policy and is parameterized by θ. However, the output action of $\hat{\pi_\theta}$ does not guarantee that the physical constraints of the system are maintained which can result in issues like system voltage violation or damage to the EV battery, and are hence said to be *unsafe* actions. To get a *safe* policy, the output of this neural network is passed to a differentiable projection layer $\mathcal{P}$ that maps the action to the safe region by solving an optimization problem. Hence, we write the safe policy as $\pi_\theta(s_t) = \mathcal{P}_{\mathcal{S}}\big(\hat{\pi_\theta}(s_t)\big)$ where $\mathcal{S}$ is the safe region, *i.e.*, the feasible set of Problem (4.9) described below. The SAC agent is trained to minimize the following loss function:

$$\mathcal{L}(\theta, s_t) = -J(\theta) + \xi \|\pi_\theta(s_t) - \hat{\pi_\theta}(s_t)\|_2^2 \ , \tag{4.8}$$

where ξ is a non-negative hyper-parameter. We add $\mathcal{L}(\theta, s_t)$ to the policy function equation and to the function used for calculating the loss of automatic entropy tuning [43]. As a result, the hard constraints of our problem formula-

tion are learned by the neural network of the SAC agent via backpropagation.

The optimization problem that becomes the differentiable projection layer within the actor network of SAC is written below:

$$\underset{a_t, AC, AD}{\text{minimize}} \quad \|a_t - \tilde{a}_t\|_2^2 \tag{4.9a}$$

$$\text{subject to} \quad AD_t^n = 0, \qquad \forall n \in \mathcal{N}_t \setminus \mathcal{N}_t^D \tag{4.9b}$$

$$-\alpha_d \leq AD_t^n \leq 0, \qquad \forall n \in \mathcal{N}_t^D \tag{4.9c}$$

$$0 \leq AC_t^n \leq \alpha_c, \qquad \forall n \in \mathcal{N}_t \tag{4.9d}$$

$$y_t^n = AC_t^n + AD_t^n, \qquad \forall n \in \mathcal{N}_t \tag{4.9e}$$

$$SoC_{t+1}^n = SoC_t^n + \frac{AC_t^n \eta_c}{b_n} + \frac{AD_t^n}{\eta_d b_n}, \quad \forall n \in \mathcal{N}_t \tag{4.9f}$$

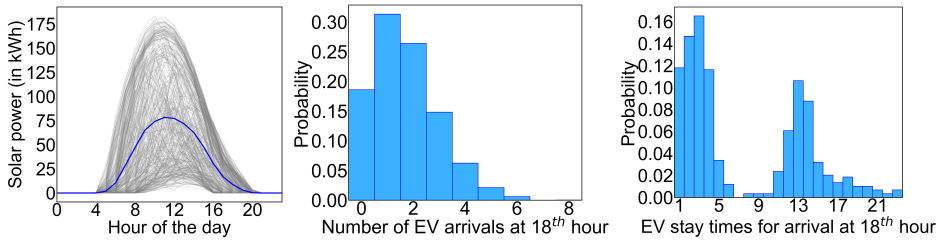
$$\delta_{min} \leq SoC_{t+1}^n \leq \delta_{max}, \qquad \forall n \in \mathcal{N}_t \tag{4.9g}$$

$$lax_{t+1}^n \geq 0, \qquad \forall n \in \mathcal{N}_t \tag{4.9h}$$

In this formulation, $\tilde{a}_t$ is the pre-projection action vector, *i.e.*, the set of actions taken by the RL agent concerning the connected EVs, before it is passed to the projection layer in the neural network. This optimization problem finds the post-projection action vector a_t that ensures the charging problem remains feasible for each EV and has the minimum Euclidean distance from the pre-projection action vector. The constraints define bounds for charge and discharge rates, and the SoC of batteries. The last constraint is to force the laxity of every EV to remain non-negative in the next hour if we implement a_t in this hour. Here lax_{t+1}^n can be defined by plugging in SoC_{t+1}^n in Equation 4.2, and replacing t with $t+1$. This ensures the problem remains feasible and all charging deadlines can be met.

4.4 Datasets

We combine four datasets that contain real solar traces, DA market prices, IM market prices, and EV charging sessions between January 1, 2020 and December 31, 2020, to create a test dataset that is used to evaluate the proposed VPP operating strategies and baselines (described in the next section). All these datasets pertain to the same region in Rotterdam, Netherlands. Specifically,



(a) (b) (c)

Figure 4.3: Empirical distribution of solar generation and EV mobility data: (a) daily solar generation where each gray curve represents the PV system output on a specific day and the blue curve represents the hourly mean solar generation over one year; (b) the probability mass function (pmf) of the number of EV arrivals in a specific hour of the day; (c) the conditional pmf of stay times for EVs arrived in that hour.

we pull hourly solar irradiance data via the Solcast API,[4] using the latitude and longitude of an arbitrary location in Rotterdam. This irradiance data is fed to the PVWatts model [32] to compute the power generated by a PV system located at these coordinates. The tilt angle of the panels is set to 51 degrees and their orientation angle to 270 degrees. The size of the PV system is defined according to the maximum EV charging demand, which is 200 kW-peak in our study. Figure 4.3a shows the daily solar production curves of this PV system. We obtain the hourly DA market price data for the Dutch market from the European Network of Transmission System Operators [35], and the IM market price data from the regional Transmission System Operator, called TENNET [108].

We use the dataset released by ElaadNL, a large charging infrastructure in the Netherlands [33], as our EV dataset. This dataset contains 10,000 charging sessions that occurred in several public EV chargers operated by EVnetNL.[5] Note that since the initial and target SoC levels are not reported in this dataset, we assume the target SoC of every EV is 1 and use the total energy charged into the battery in the respective charging session to calculate its initial SoC. We evaluate the proposed operating strategies (LLA and LA-SAC) and baselines on this test dataset.

We create a new dataset, separate from the test dataset described above,

[4]https://solcast.com/
[5]https://evnet.nl/

and use it only to train the LA-SAC agent in the second operating strategy. This is necessary because the amount of historical data (charging sessions) in the ElaadNL dataset is not enough to learn a near-optimal policy via reinforcement learning. To obtain a sufficiently large number of episodes, we synthesize realistic charging sessions. Specifically, we fit distributions to EV arrival times (depicted in Figure 4.3b), EV stay times (depicted in Figure 4.3c), and EV charging demands in the ElaadNL dataset. We assume that the number of arrivals in each hour of the day follows Poisson distribution with a parameter that depends on that particular hour. Since the stay time correlates with the arrival time, we fit a Gaussian mixture model, using Kernel Density Estimation (KDE), to the empirical distribution of stay times for EVs that arrived in a certain hour of the day. For the initial SoC levels, we use a truncated Gaussian distribution with mean 0.49 and standard deviation of 0.25, with the minimum and maximum SoC being 0.03 and 0.97 respectively; the two moments of the Gaussian distribution are defined according to the distribution of initial SoC in the test dataset. Furthermore, we assume that all EVs must be fully charged before they leave the charging station and set $\overline{SoC}^n$ accordingly. We set the size of all EV batteries to 80kWh (similar to Tesla Model 3), the maximum charge and discharge rates to 11kW, and the charge and discharge efficiencies to 0.98. We generate almost 3 years worth of EV charging sessions by sampling from these distributions. To create the training dataset, this data is combined with real market prices and solar traces that are collected from the same sources we used to create the test dataset, this time for a period that ends on December 31, 2019. This allows us to train the LA-SAC agent for up to 1,000 episodes, where each episode consists of 24 one-hour time steps, representing 1 day.

4.5 Baseline 1: Offline Deterministic (ORACLE)

This baseline solves Problem (4.1) using the actual values of the hourly market prices ($\mathbf{P}^{DA}, \mathbf{P}^{IM}$), solar generation ($\mathbf{E}^{solar}$), and EV mobility and energy de-

mand $(t_s^n, \tau^n, \underline{SoC}^n, \overline{SoC}^n)$. The solution would give the maximum profit the VPP can make by operating in the two-stage electricity market, *i.e.*, **X***, $\mathbf{Y}^{*n}$, and **Z***. Note that we do not need to solve the stage 2 problem separately because we obtain the optimal EV charging schedule as the LP is solved with perfect information. This baseline, which we call **ORACLE**, gives an upper bound on the VPP's profit. In practice, the VPP operates under significant uncertainty, hence it is impossible to generate this much profit.

4.6 Baseline 2: Current Practice in EV Charging (CHRG_ASAP)

The current practice in EV charging referred to as **CHRG_ASAP**, entails charging an EV at the maximum power as soon as it gets connected to a charger. This strategy minimizes the length of the charging session. It combines the solution of stage 1 and stage 2 problems when they are solved without taking advantage of V2G. More specifically, for solving stage 1 under this baseline, Problem (4.1) is modified by dropping Constraint (4.1g) and writing Constraint (4.1f) for all $n \in \mathcal{N}_t$, and then Problem (4.1) is solved using the aforementioned 2-staged approach.

In stage 2, EVs are charged at the maximum power when they arrive at a charging station. The main drawback of this strategy is the reduced flexibility, limiting the VPP options when it comes to addressing potential deviations from day-ahead commitments.

Figure 4.4 shows the EV charging schedule using the CHRG_ASAP baseline for an EV arriving with 20% SoC and requiring 100% SoC within its 8 hours of connection time to the charger, where the battery capacity of this EV is assumed to be 80kWh. It can be seen in the bar chart that the EV is idly connected to the charger for the last 2 hours of its charging duration. As a result, it is possible to *re-shape* the charging schedule of this EV leveraging this flexibility. The algorithms that we present in this thesis take advantage of the flexible nature of EV charging to maximize VPP profit.

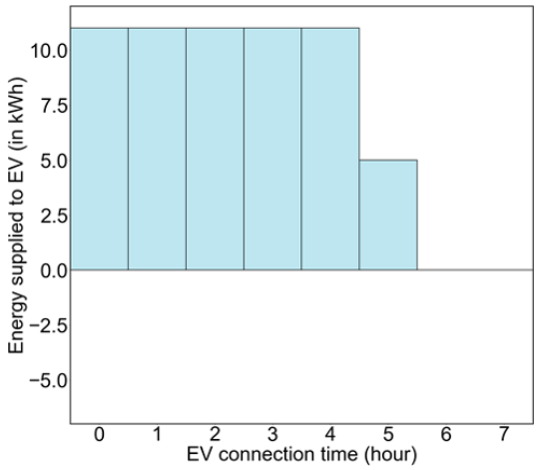


Figure 4.4: EV charging schedule with the CHRG_ASAP baseline. Negative values in the y-axis represent EV (dis)charge.

4.7 Annual VPP Profit

We now evaluate the performance of the proposed VPP operating strategies on the dataset described in Section 4.4. We investigate the profitability of the VPP in different scenarios considering five different V2G participation rates $(0\%, 25\%, 50\%, 75\%, 100\%)$, where the participation rate is defined as the ratio of EVs whose battery can be discharged to the total number of EVs that visit the charging stations in one day. Hence, 0% participation implies that *none* of the EVs that arrive at the charging stations can be discharged, and 100% participation implies that *all* EVs might be discharged as long as their charging demand can be met by their deadline. For a fixed participation rate, we randomly sample the required number of EVs to participate in V2G from the set of EVs that will visit the charging stations during the day.

Figure 4.5 compares the annual profit earned by the VPP when it adopts CHRG_ASAP, LLA, and LA-SAC.[6] We see that both LLA and LA-SAC greatly increase the VPP's profit for all V2G participation rates, while the profit earned under CHRG_ASAP is much lower and does not vary with the V2G participation rate because it does not take advantage of bidirectional charging. The increased profitability of the VPP when all EVs participate in V2G

[6]Recall that both strategies solve the stage 1 problem using a WS approach. But, for brevity, we just use the name of the the algorithm used in stage 2 to refer to each strategy.

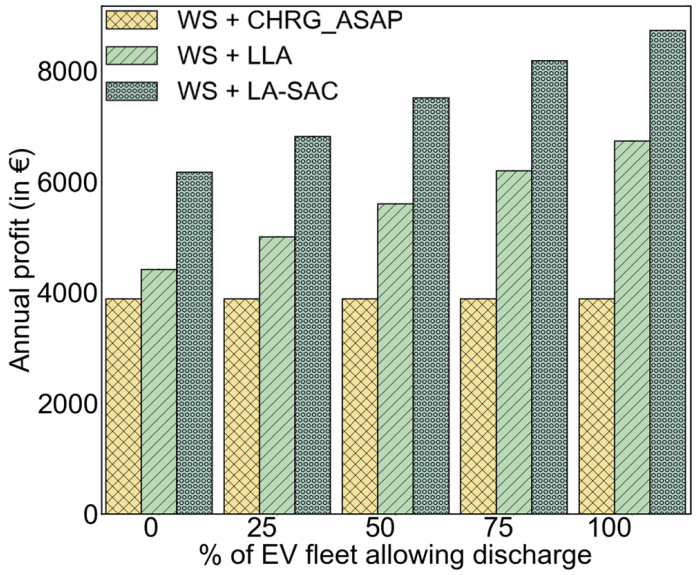


Figure 4.5: Annual profit earned by the VPP for various V2G participation rates.

compared to when V2G is not supported (52% and 42% increase under LLA and LA-SAC, respectively) highlights the importance of bidirectional charging in making this kind of VPP viable. Table 4.1 compares the annual profit earned by the VPP using each operating strategy as a percentage of the annual profit generated by ORACLE for the same V2G participation rate. The result indicates that for 100% V2G participation case, LA-SAC is the best-performing algorithm, achieving 51.4% of the profit that could be possibly earned if there was no uncertainty, followed by LLA which achieves 39.6% of the profit earned by ORACLE. The gap between the performance of the proposed operating strategies and ORACLE widens slightly as the V2G participation rate increases. We attribute this to the increased complexity of the problem when more storage capacity becomes available. It is important to note that LLA achieves around 40% of the profit that could be possibly earned if there was no uncertainty using a simple heuristic. However, should the VPP be able to afford the training cost of LA-SAC, its annual profit can be increased by up to €2,002 compared to when it adopts LLA.

	V2G participation (%)				
Algorithm	0	25	50	75	100
CHRG_ASAP	37.5%	32.4%	28.3%	25.2%	22.8%
LLA	42.7%	41.7%	40.8%	40.3%	39.6%
LA-SAC	59.6%	56.8%	54.8%	53.2%	51.4%

Table 4.1: The profit earned under different strategies as a percentage of the profit earned by ORACLE (assuming perfect information) for the same V2G participation rate.

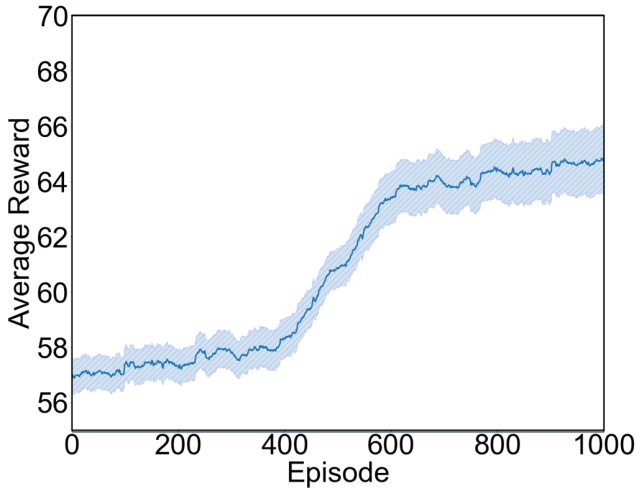
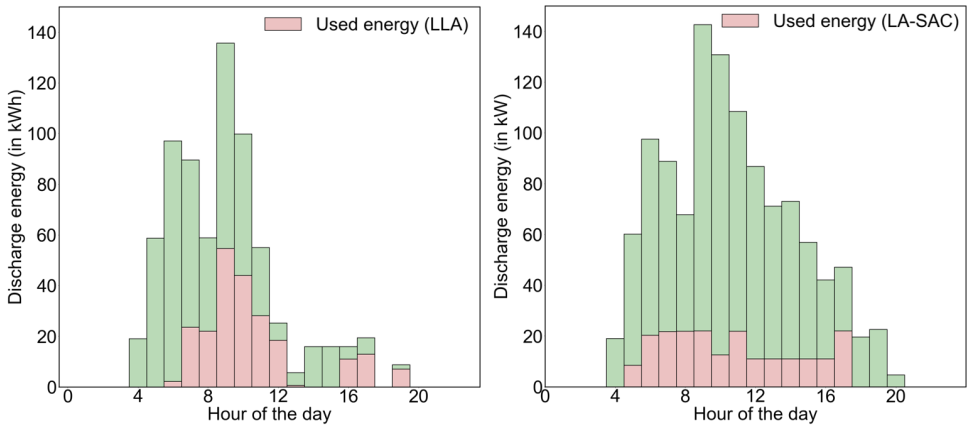


Figure 4.6: The reward obtained per episode by LA-SAC for 100% V2G participation.

4.8 Learning Curve of LA-SAC

We briefly discuss the number of episodes required to train a policy using the LA-SAC algorithm. We use 3 random seeds (independent trials) to train a policy using this algorithm. In each trial, the policy is trained for 1,000 episodes. Figure 4.6 depicts the learning curve of the RL agent assuming 100% participation in V2G. The solid curve corresponds to the mean over 3 trials and the shaded region shows one standard error from the mean. It can be seen that the RL agent demonstrates improved performance after around 600 episodes (days).

We witnessed stable performance when the learned policy (after 1,000 episodes) was evaluated on the test dataset.



(a) (b)

Figure 4.7: Comparing discharge energy usage of LLA and LA-SAC. The summation of the green and pink bars represents the total available discharge energy from the EVs and the pink bar represents the discharge energy used by the respective algorithm. (a) discharge energy usage by LLA; (b) discharge energy usage by LA-SAC

4.9 Comparison between LLA and LA-SAC

Figure 4.7 compares the discharge energy usage by the algorithms, namely LLA and LA-SAC, that are used to solve the second stage of the 2-stage optimization problem. As explained in the previous chapter, stage 2 requires EVs to be discharged in order to mitigate the deficit in solar energy in real time. It can be inferred upon comparing both bar charts in Figure 4.7 that LLA takes a less conservative approach in discharging EVs (refer to hours 8 to 12 in Figure 4.7) which reduces the available discharge energy in the following hours of the day. In contrast, LA-SAC discharges EVs more conservatively and this ensures that the discharge energy availability is spread throughout the entire day since EV mobility has an intermittent pattern, which results in less reliance in buying electricity from the imbalance market, thereby increasing VPP profit. LA-SAC takes the aforementioned approach in discharging EVs because the reward function of the RL agent (see Equation 4.6) is designed such that the agent receives a higher reward for maintaining higher average hourly laxity.

Chapter 5

Incentivizing EVs to Participate in the VPP

In this chapter, we explain how contract theory can be used to incentivize independent EV owners to participate in a VPP, and perform sensitivity analysis to the factors such as the perceived battery degradation cost and contract duration that influence EV owners' willingness to participate in V2G. Finally, we analyze the profitability of the VPP after accounting for the payoff provided to the EV owners.

5.1 Modeling the VPP and EV Owners

In contract theory, each contract specifies the contribution from the agent or agents who are offered the contract, and the corresponding payoff that they receive. Effective contract design requires a careful definition of the utility functions of the contract recipients and the contract provider. We use the principal-agent model to describe the interaction between the VPP and EV owners. In this model, the VPP (principal) has bargaining power, meaning that it defines the set of contracts that will be offered to each EV owner (agent). An EV owner can simply accept or decline an offer proposed by the In this section, we introduce the contract theory-specific terminologies and present the utility function of the VPP operator and EV owners.

5.1.1 EV Owner Type

Let N_t be the set of EVs that arrive at the charging station in hour t of a day and $\mathcal{N} = \cup_{t\in\mathcal{T}}\{N_t\}$ be the set of all EVs that arrive at the charging station on this day, i.e., in a 24 hour period ($\mathcal{T} = \{1, \cdots, 24\}$). Each EV owner $n \in \mathcal{N}$ has some private information that influences the amount of energy they are willing to provide to the VPP through V2G. We classify EV owners into different *types* based on their willingness to contribute to the VPP. Without loss of generality, we assume the type of an EV owner belongs to an interval, $[\underline{\theta}, \overline{\theta}]$. We quantize the type with a quantization factor M such that the collection of types forms a discrete set denoted $\Theta = \{\theta_1, \cdots, \theta_M\}$ where $\underline{\theta} \leq \theta_1 < \cdots < \theta_M \leq \overline{\theta}$. A greater value of m in θ_m indicates that the EV owner is more willing to have their battery discharged so as to participate in V2G. The VPP operator does not deterministically know the type of an EV owner since it is determined based on their private information. However, we assume that the VPP operator knows the probability distribution over different types, which implies that it knows an arbitrary EV owner belongs to a certain type θ_m with probability π_m (where $\sum_{m=1}^{M} \pi_m = 1$).

5.1.2 Contract Structure

To increase V2G participation, the VPP offers monetary incentives to EV owners for the maximum amount of energy that will be withdrawn from their battery over a time period of length l_{V2G}, starting when the contract is accepted. Recall that the VPP operator does not know the actual type of an EV owner. Our goal is to design *incentive compatible* contracts which are, *self-revealing* [89], thus allowing us to design the best contracts despite the information asymmetry that exists. The VPP operator will offer energy-reward bundle contracts $\{(g_m, w_m)\}_{m=1\cdots M}$ to EV owners, where each contract is designed for a specific type. For example, an EV owner of type θ_m should accept (g_m, w_m), with w_m being the amount of discharge energy that they will provide and g_m being the associated payoff, which is a strictly increasing function of θ_m.

5.1.3 Utility Functions

VPP's Utility

The VPP's utility when an EV of type θ_m participates in V2G can be defined as follows:

$$U_{VPP}(g_m, w_m) = u(w_m) - g_m \tag{5.1}$$

Here $u(w_m)$ represents the revenue that can be generated by discharging w_m from the battery of this EV, and g_m is the EV owner's payoff for their participation in V2G. Since the VPP is assumed to be risk-averse, we let $u(\cdot)$ be a concave function of w_m [85]. Thus, the expected utility of the VPP from the participation of an arbitrary EV in V2G would be:

$$U_{VPP}^{total} = \sum_{m=1}^{M} \pi_m \cdot \big(u(w_m) - g_m\big) \tag{5.2}$$

and its expected utility when $|\mathcal{N}|$ EVs participate in V2G would be $U_{VPP}^{total} \cdot |\mathcal{N}|$.

EV Owner's Utility

The utility of an EV of type θ_m that accepts contract (g_m, w_m) is:

$$U_{EV}(g_m, w_m; \theta_m) = g_m - \mathcal{C}(w_m; \theta_m), \tag{5.3}$$

where $\mathcal{C}(.)$ represents the cost incurred by an EV of type θ_m due to discharging w_m amount of energy from their battery (e.g., the battery degradation cost). Using the fixed per kWh degradation model proposed in [1], we write the cost function for type θ_m as:

$$\mathcal{C}(w_m; \theta_m) = \frac{w_m}{\theta_m} \cdot \gamma \cdot c \ , \tag{5.4}$$

where $\gamma = \frac{V_{batt}}{L}$, $c = \frac{1}{2 \cdot b_n \cdot DoD/100}$, and V_{batt} represents the battery price or its perceived value, L represents its nominal cycle life, DoD represents its nominal depth of discharge, and b_n represents its capacity. Note that $\mathcal{C}(.)$ is a concave function of w_m. The economic significance of this design is that the marginal utility of an increase in w_m is always higher for higher EV types. Thus, the

willingness to receive payment for an increase in w_m is always higher for higher types.

The outside utility of an EV owner, i.e., when they reject all contracts, is assumed to be zero.

5.2 Contract Design

Let $V_{EV}(\theta_{m'}, \theta_m) = U_{EV}(g_{m'}, w_{m'}; \theta_m)$ represent the utility of an EV owner whose true type is θ_m, but declares their type to be $\theta_{m'}$. The VPP operator aims to maximize their expected utility by designing and offering a specific contract for each type of EV owner, satisfying incentive compatibility and individual rationality requirements. We define these requirements below.

Individual Rationality (IR): The contract accepted by an EV owner should guarantee that their utility will be non-negative (i.e., higher than the outside utility). This can be written as:

$$U_{EV}(g_m, w_m; \theta_m) = V_{EV}(\theta_m, \theta_m) = g_m - \frac{w_m \cdot \gamma \cdot c}{\theta_m} \geq 0$$

Incentive Compatibility (IC): EV owners do not gain any advantage by lying about their true type. Consequently, they prefer a contract that is specifically designed for their type and declare their type truthfully. This can be written as:

$$V_{EV}(\theta_m, \theta_m) \geq V_{EV}(\theta_{m'}, \theta_m) \quad \forall \theta_m, \theta_{m'} \in \Theta \tag{5.5}$$

5.2.1 Finding Optimal Contracts

The optimal contracts are the solution of the following utility maximization problem subject to the constraints defined above:

$$\begin{aligned}
\underset{\{(g_m,w_m)\}}{\text{maximize}} \quad & \sum_{m=1}^{M} \pi_m \cdot (u(w_m) - g_m) \qquad (5.6)\\
\text{subject to} \quad & (\mathbf{OC1}) \quad g_m - \frac{w_m \cdot \gamma \cdot c}{\theta_m} \geq 0, \quad \forall \theta_m \in \Theta \\
& (\mathbf{OC2}) \quad g_m - \frac{w_m \cdot \gamma \cdot c}{\theta_m} \geq g_l - \frac{w_l \cdot \gamma \cdot c}{\theta_m}; \\
& \quad \forall m, l \in \Theta, m \neq l \\
& (\mathbf{OC3}) \quad 0 \leq w_1 \leq ... \leq w_M \\
& (\mathbf{OC4}) \quad 0 \leq g_1 \leq ... \leq g_M
\end{aligned}$$

The solution to this problem is a set of contracts, denoted $\{(g_m^*, w_m^*)\}_{m=1\cdots M}$, which are offered to the EVs, maximizing the expected utility of the VPP. Observe that there is a total of $M(M-1)$ constraints in the form of (**OC2**) in the above optimization problem, hence the number of constraints is not a linear function of M. As shown in Figure 5.1, the computational overhead of solving this optimization problem increases drastically with the number of EV owner types. We propose a more tractable version of this optimization problem by rewriting Constraint (**OC1**) and Constraint (**OC2**), and by dropping Constraint (**OC3**) so as to reduce the number of constraints as follows:

$$\begin{aligned}
\underset{\{g_m,w_m\}}{\text{maximize}} \quad & \sum_{m=1}^{M} \pi_m \cdot (u(w_m) - g_m) \qquad (5.7)\\
\text{subject to} \quad & (\mathbf{C1}) \quad g_1 - \frac{w_1 \cdot \gamma \cdot c}{\theta_1} = 0 \\
& (\mathbf{C2}) \quad g_m - \frac{w_m \cdot \gamma \cdot c}{\theta_m} = g_{m-1} - \frac{w_{m-1} \cdot \gamma \cdot c}{\theta_m} \\
& \quad \forall m \in \Theta \setminus \{1\} \\
& (\mathbf{C3}) \quad 0 \leq g_1 \leq ... \leq g_M
\end{aligned}$$

The resulting optimization Problem (5.7) can be solved faster than the original problem and the running time increases moderately as we consider more types (see Figure 5.1). In the next section, we prove that these two problems are equivalent.

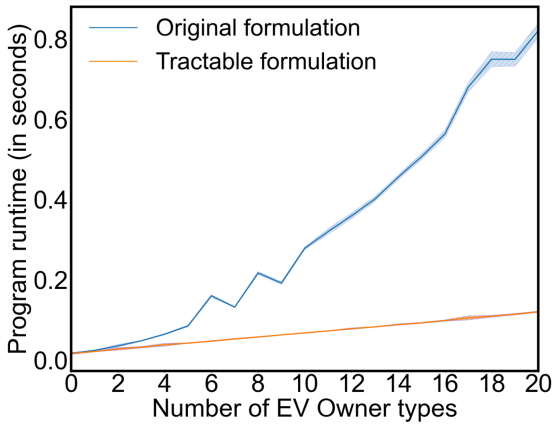


Figure 5.1: Run-time comparison between the original optimization problem with our proposed version which is more tractable. Each point shows the average running time over 5 trials and the shaded region shows three standard errors around the mean.

Note that we add an additional constraint to these optimization problems to ensure that the maximum discharge energy offered in a contract is upper bounded by the maximum amount of energy that can be possibly discharged from the battery within the contract duration: $w_m \leq \alpha_d \cdot l_{V2G}$, and this constraint is incorporated for practical reasons.

5.2.2 Proof of Equivalence

We borrow the following definitions from [19]:

- Downward Incentive Constraints (DICs):
  The IC constraints between *type* i and *type* j with $j \in \{1, ..., i-1\}$, which can be written as $V_{EV}(\theta_i, \theta_i) \geq V_{EV}(\theta_j, \theta_i)$

- Local Downward Incentive Constraints (LDICs):
  The IC constraints between *type* i and *type* $(i-1)$, which can be written as $V_{EV}(\theta_i, \theta_i) \geq V_{EV}(\theta_{i-1}, \theta_i)$

- Upward Incentive Constraints (UICs):
  The IC constraints between *type* i and *type* j, with $j \in \{i+1, ..., N\}$ which can be written as $V_{EV}(\theta_i, \theta_i) \geq V_{EV}(\theta_j, \theta_i)$

- Local Upward Incentive Constraints (LUICs):

The IC constraints between *type* i and *type* $(i+1)$, which can be written as $V_{EV}(\theta_i, \theta_i) \geq V_{EV}(\theta_{i+1}, \theta_i)$

Theorem 1 For the optimal solution of Problem (5.6), the individual rationality constraint for the lowest type is binding, that is, $V_{EV}(\theta_1, \theta_1) = 0$, and consequently, the individual rationality constraint for higher types also hold and hence can be ignored.

Proof: We want to prove the following relationship:

$$V_{EV}(\theta_m, \theta_m) \geq V_{EV}(\theta_1, \theta_1) = 0 \tag{5.8}$$

Step 1. Since the optimal solution is incentive compatible, we can write this inequality ($\forall \theta_m \in \Theta$):

$$V_{EV}(\theta_m, \theta_m) \geq V_{EV}(\theta_1, \theta_m) \tag{5.9}$$

Step 2. When the same contract (in this case (g_1, w_1)), is offered to EV owners of types θ_m and θ_1, the attained utility is higher for the higher type EV owner, and therefore, the following holds:

$$\begin{aligned} &V_{EV}(\theta_1, \theta_m) \geq V_{EV}(\theta_1, \theta_1) \\ &\Rightarrow g_1 - \frac{w_1 \cdot \gamma \cdot c}{\theta_m} \geq g_1 - \frac{w_1 \cdot \gamma \cdot c}{\theta_1} \end{aligned} \tag{5.10}$$

This clearly holds because: $\theta_1 < ... < \theta_m < ... < \theta_M$.

Step 3. We prove by contradiction that at the optimal solution, the IR condition for the lowest type (i.e. type θ_1) must be binding.

We can write the following if the IR condition for type θ_1 is not binding:

$$\begin{aligned} &V_{EV}(\theta_1, \theta_1) > 0 \\ &\Rightarrow g_1 - \frac{w_1 \cdot \gamma \cdot c}{\theta_1} > 0 \end{aligned}$$

Now consider (g_1', w_1) which is a new contract with $g_1' = -\epsilon + g_1$, where ϵ is a positive constant and $\epsilon \leq g_1 - \frac{w_1 . \gamma . c}{\theta_1}$

The following is the type θ_1 EV owner's utility when they are offered this contract:

$$\begin{aligned} &g_1' - \frac{w_1 \cdot \gamma \cdot c}{\theta_1} > -\epsilon + g_1 - \frac{w_1 \cdot \gamma \cdot c}{\theta_1} > 0 \\ &\Rightarrow -\epsilon + V_{EV}(\theta_1, \theta_1) > 0 \end{aligned}$$

Since the IR condition is satisfied with the new contract (g_1', w_1) and this contract increases the VPP profit, it is, therefore, a better contract than the optimal solution. This contradiction suggests that the IR constraint for type θ_1 must be binding so that no other contract better than the optimal contract can be found.

We have proved the following relationship:

$$V_{EV}(\theta_m, \theta_m) \geq V_{EV}(\theta_1, \theta_m) \geq V_{EV}(\theta_1, \theta_1) = 0$$

The first inequality follows directly from the assumption that the incentive compatibility constraint has been satisfied, the second inequality follows from the proof in step 2 and the last equality is proved in step 3. (PROVED)

As a result, we are able to remove Constraint (**OC1**) of Problem (5.6) and replace it with Constraint (**C1**) (which is shown on Problem (5.7)). This reduces the total number of constraints since Constraint (**OC1**) had to be incorporated for all EV owner types present in the set Θ, whereas Constraint (**C1**) is written for the lowest EV owner type only.

Theorem 2 The monotonicity condition, i.e. $0 \leq w_1 \leq ... \leq w_M$ can be derived from the IC constraints.

Proof: Let us consider two types of EV owners: θ_j and θ_k such that $\theta_k > \theta_j$. The IC constraints for these EV owner types are given by:

$$V(\theta_k, \theta_k) \geq V(\theta_j, \theta_k)$$

$$\Rightarrow g_k - \frac{w_k \cdot \gamma \cdot c}{\theta_k} \geq g_j - \frac{w_j \cdot \gamma \cdot c}{\theta_k} \tag{5.11}$$

Similarly,

$$V(\theta_j, \theta_j) \geq V(\theta_k, \theta_j)$$

$$\Rightarrow g_j - \frac{w_j \cdot \gamma \cdot c}{\theta_j} \geq g_k - \frac{w_k \cdot \gamma \cdot c}{\theta_j} \tag{5.12}$$

Adding up the 2 sides of the above 2 inequalities (i.e., Equations 5.11 and 5.12), we get:

$$\begin{aligned}
&g_k + g_j - \frac{w_k \cdot \gamma \cdot c}{\theta_k} - \frac{w_j \cdot \gamma \cdot c}{\theta_j} \geq \\
&g_j + g_k - \frac{w_j \cdot \gamma \cdot c}{\theta_k} - \frac{w_k \cdot \gamma \cdot c}{\theta_j} \\
&\Rightarrow -\frac{w_k \cdot \gamma \cdot c}{\theta_k} + \frac{w_j \cdot \gamma \cdot c}{\theta_k} \geq \frac{w_j \cdot \gamma \cdot c}{\theta_j} - \frac{w_k \cdot \gamma \cdot c}{\theta_j} \\
&\Rightarrow \frac{\gamma \cdot c \cdot (w_j - w_k)}{\theta_k} \geq \frac{\gamma \cdot c \cdot (w_j - w_k)}{\theta_j}
\end{aligned}$$

Reorganizing the above inequality yields:

$$\begin{aligned}
&\theta_j \cdot \gamma \cdot c \cdot (w_j - w_k) \geq \theta_k \cdot \gamma \cdot c \cdot (w_j - w_k) \\
&\Rightarrow \theta_k(w_k - w_j) \geq \theta_j(w_k - w_j)
\end{aligned}$$

Since $\theta_k > \theta_j$, we conclude that $w_k \geq w_j$. Because the discharge energy from any arbitrary EV owner type m, i.e. w_m, is assumed to be positive, we have:

$$0 \leq w_1 \leq ... \leq w_M \tag{5.13}$$

and thus, our proof is complete. (PROVED)

This theorem allows us to drop Constraint (**OC3**) to simplify Problem (5.6).

Theorem 3 For the optimal solution of Problem (5.6), we have:

1. All DICs are satisfied when LDIC is satisfied
2. All UICs are satisfied when LUIC is satisfied

Proof. Let us consider 3 types of EV owners: θ_{j-1}, θ_j and θ_{j+1}, where $\theta_{j-1} < \theta_j < \theta_{j+1}$. In Step 1, we prove that LDIC implies DIC.

Step 1. The IC constraint between type θ_j and θ_{j-1} constitutes an LDIC

as follows:

$$V_{EV}(\theta_j, \theta_j) \geq V_{EV}(\theta_{j-1}, \theta_j) \tag{5.14}$$

$$\Rightarrow g_j - \frac{w_j \cdot \gamma \cdot c}{\theta_j} \geq g_{j-1} - \frac{w_{j-1} \cdot \gamma \cdot c}{\theta_j}$$

$$\Rightarrow g_j - g_{j-1} \geq \frac{(w_j - w_{j-1}) \cdot \gamma \cdot c}{\theta_j} \tag{5.15}$$

According to the constraint: $\theta_1 < ... < \theta_m < ... < \theta_M$, we know that θ_{j+1} must be larger than θ_j, and as a result, the following inequality holds.

$$\frac{(w_j - w_{j-1}) \cdot \gamma \cdot c}{\theta_j} \geq \frac{(w_j - w_{j-1}) \cdot \gamma \cdot c}{\theta_{j+1}} \tag{5.16}$$

Since Equation 5.15 holds and Equation 5.16 holds, therefore the following also holds:

$$g_j - g_{j-1} \geq \frac{(w_j - w_{j-1}) \cdot \gamma \cdot c}{\theta_j} \geq \frac{(w_j - w_{j-1}) \cdot \gamma \cdot c}{\theta_{j+1}} \tag{5.17}$$

Now, from the above inequality, the following can be deduced:

$$g_j - g_{j-1} \geq \frac{(w_j - w_{j-1}) \cdot \gamma \cdot c}{\theta_{j+1}} \tag{5.18}$$

$$\Rightarrow g_j - \frac{w_j \cdot \gamma \cdot c}{\theta_{j+1}} \geq g_{j-1} - \frac{w_{j-1} \cdot \gamma \cdot c}{\theta_{j+1}}$$

$$\Rightarrow V_{EV}(\theta_j, \theta_{j+1}) \geq V_{EV}(\theta_{j-1}, \theta_{j+1}) \tag{5.19}$$

The following is the LDIC constraint for EV owner of type θ_{j+1}:

$$V_{EV}(\theta_{j+1}, \theta_{j+1}) \geq V_{EV}(\theta_j, \theta_{j+1}) \tag{5.20}$$

Now, since Equation 5.19 holds and Equation 5.20 holds, the following also holds:

$$V_{EV}(\theta_{j+1}, \theta_{j+1}) \geq V_{EV}(\theta_{j-1}, \theta_{j+1}) \tag{5.21}$$

Therefore, for each type θ_j, if the incentive constraint with respect to type θ_{j-1} holds, (i.e. the LDIC is satisfied), then all other downward incentive constraints are also satisfied. Thus, we are able to reduce the set of downward

incentive constraints to the set of LDICs.

Next, we prove that LUIC implies UIC.

Step 2. The IC constraint between type θ_j and θ_{j+1} constitutes a LUIC, which is:

$$\begin{aligned} & V(\theta_j, \theta_j) \geq V(\theta_{j+1}, \theta_j) \\ & \Rightarrow g_j - \frac{w_j \cdot \gamma \cdot c}{\theta_j} \geq g_{j+1} - \frac{w_{j+1} \cdot \gamma \cdot c}{\theta_j} \\ & \Rightarrow g_{j+1} - g_j \leq \frac{(w_{j+1} - w_j) \cdot \gamma \cdot c}{\theta_j} \end{aligned} \tag{5.22}$$

According to the constraint: $\theta_1 < ... < \theta_m < ... < \theta_M$, we know that θ_j must be larger than θ_{j-1}, and as a result, the following inequality holds.

$$\frac{(w_{j+1} - w_j) \cdot \gamma \cdot c}{\theta_j} \leq \frac{(w_{j+1} - w_j) \cdot \gamma \cdot c}{\theta_{j-1}} \tag{5.23}$$

Since Equation 5.22 holds and Equation 5.23 holds, therefore the following also holds:

$$g_{j+1} - g_j \leq \frac{(w_{j+1} - w_j) \cdot \gamma \cdot c}{\theta_j} \leq \frac{(w_{j+1} - w_j) \cdot \gamma \cdot c}{\theta_{j-1}} \tag{5.24}$$

Now, from the above inequality, the following can be deduced:

$$\begin{aligned} & g_{j+1} - g_j \leq \frac{(w_{j+1} - w_j) \cdot \gamma \cdot c}{\theta_{j-1}} \\ & \Rightarrow g_{j+1} - \frac{w_{j+1} \cdot \gamma \cdot c}{\theta_{j-1}} \leq g_j - \frac{w_j \cdot \gamma \cdot c}{\theta_{j-1}} \\ & \Rightarrow V_{EV}(\theta_j, \theta_{j-1}) \geq V_{EV}(\theta_{j+1}, \theta_{j-1}) \end{aligned} \tag{5.25}$$

The following is the LUIC constraint for EV owner of type θ_{j-1}:

$$V_{EV}(\theta_{j-1}, \theta_{j-1}) \geq V_{EV}(\theta_j, \theta_{j-1}) \tag{5.26}$$

Now, since Equation 5.25 holds and Equation 5.26 holds, the following also holds:

$$V_{EV}(\theta_{j-1}, \theta_{j-1}) \geq V_{EV}(\theta_{j+1}, \theta_{j-1}) \tag{5.27}$$

Therefore, for each type θ_j, if the incentive constraint with respect to type θ_{j+1} holds, (i.e. the LUIC is satisfied), then all other upward incentive constraints (i.e IC constraints for θ_j relative to higher θ_j's) are also satisfied. This allows us to reduce the set of upward incentive constraints to the set of LUICs. (PROVED)

We prove in the next theorem (Theorem 4) that the LUIC constraints can be omitted. As a result, Constraint (**OC2**) in Problem (5.6) can be replaced with Constraint (**C2**) in Problem (5.7). This allows us to reduce the total number of constraints significantly, which is important to achieve faster runtime for the optimization problem.

Theorem 4 An EV owner will not accept a contract designed for a higher type than their own type, because this leads to a negative utility.

Proof. Let us consider 3 different EV owners belonging to types θ_{j-1}, θ_j, and θ_{j+1} respectively. Let the following be the optimal contracts that satisfy individual rationality constraints but not incentive compatibility constraints: (g^*_{j-1}, w^*_{j-1}), (g^*_j, w^*_j), and (g^*_{j+1}, w^*_{j+1}) respectively. More specifically, these contracts are designed by the VPP operator to pay each EV owner just enough to keep their utility zero while maximizing their profit. For example, the contract (g^*_j, w^*_j) is designed for EV owner type θ_j and it results in a utility for EV owner type θ_j that equals the minimum value of IR constraint. Therefore the following holds for EV owner of type θ_j:

$$g^*_j - \frac{w^*_j \cdot \gamma \cdot c}{\theta_j} = 0 \tag{5.28}$$

Similarly, the following 2 equations hold for EV owners of type θ_{j-1} and type θ_{j+1} respectively:

$$g^*_{j-1} - \frac{w^*_{j-1} \cdot \gamma \cdot c}{\theta_{j-1}} = 0 \tag{5.29}$$

$$g^*_{j+1} - \frac{w^*_{j+1} \cdot \gamma \cdot c}{\theta_{j+1}} = 0 \tag{5.30}$$

When the 3 contracts: (g^*_{j-1}, w^*_{j-1}), (g^*_j, w^*_j), and (g^*_{j+1}, w^*_{j+1}) are offered, the EV owner of type θ_j chooses the contract (g^*_{j-1}, w^*_{j-1}) instead of the contract designed for their *true* type, i.e. (g^*_j, w^*_j). This is because, when the EV owner of type θ_j chooses the contract (g^*_{j-1}, w^*_{j-1}), their utility is as follows:

$$g^*_{j-1} - \frac{w^*_{j-1} \cdot \gamma \cdot c}{\theta_j} > 0 \tag{5.31}$$

Equation 5.31 shows a positive utility for the EV owner of type θ_j and it is greater than the utility of this EV owner when choosing the contract designed for their own type (the utility given in Equation 5.28). This is because, $\theta_j > \theta_{j-1}$, and hence the second term of the left-hand side of the inequality has a lower value in Equation 5.31, when compared to the second term of the left-hand side of Equation 5.28. Therefore, LDIC constraints must be incorporated so that the EV owners do not deviate from the contracts designed by the VPP operator for their designated types.

However, the EV owner of type θ_j does not prefer to choose the contract designed by the VPP operator for type θ_{j+1}. This is because, when the EV owner of type θ_j chooses the contract (g^*_{j+1}, w^*_{j+1}), their utility will be as follows:

$$g^*_{j+1} - \frac{w^*_{j+1} \cdot \gamma \cdot c}{\theta_j} < 0 \tag{5.32}$$

Equation 5.32 shows a negative utility for the EV owner of type θ_j and it is less than the utility of this EV owner when choosing the contract designed for their own type (the utility given in Equation 5.28). This is because, $\theta_j < \theta_{j+1}$, and hence the second term of the left-hand side of the inequality has a greater value in Equation 5.32 when compared to the second term of the left-hand side

of Equation 5.28. Therefore, the LUIC constraints of type-θ_j EV owner can be omitted since an EV owner has no incentive to choose a contract designed for a higher type than their own. (PROVED)

The significance of proving Theorem 4 is that having proven Theorem 3, the Constraint (**OC2**) in Problem (5.6) can be replaced with Constraint (**C2**) in Problem (5.7).

Theorem 5 For the optimal solution of Problem (5.6), all the LDICs are binding.

Proof. We will prove this by contradiction. The LDIC for an EV owner of type θ_k that is not binding can be written as:

$$\begin{aligned} & V(\theta_k, \theta_k) > V(\theta_{k-1}, \theta_k) \\ & \Rightarrow g_k - \frac{w_k \cdot \gamma \cdot c}{\theta_k} > g_{k-1} - \frac{w_{k-1} \cdot \gamma \cdot c}{\theta_k} \end{aligned} \tag{5.33}$$

Let (g'_k, w_k) be a new contract where $g'_k = -\epsilon + g_k$, where ϵ is a positive constant and $\epsilon < g_k$.
When a type θ_k EV owner utility is offered the contract (g'_k, w_k), their utility would become:

$$g'_k - \frac{w_k \cdot \gamma \cdot c}{\theta_k} \tag{5.34}$$

The VPP operator can fix the value of ϵ, such that Equation 5.33 is still satisfied (i.e., it is set equal to the gap between the two sides of this inequality), which results in increased VPP profit, by offering the minimum payment to the EV owners in the new contract (g'_k, w_k).

Since the IC condition is satisfied with the new contract (g'_k, w_k) and this contract increases VPP profit, it is a contradiction. Therefore, the LDIC constraints must be binding. (PROVED)

Theorem 5 has been used to write Constraint (**C2**) as an equality in Problem (5.7).

5.3 Analysis of the Optimal Contracts

We now consider a case study where we set the number of types to a default value of $M = 3$ and examine the energy-reward bundles to find out how many unique contracts will be found as a result of solving the optimization problem.

We define $u(.)$, the first term in the utility function of the operator given in Equation (5.1), as $u(w_m) = \kappa \cdot log(w_m + 1)$ where κ is a hyper-parameter. This function is concave and increasing in its domain for positive values of κ. This hyper-parameter reflects the importance of V2G for the VPP. A large value of κ indicates that the VPP highly values the flexibility offered by EVs through V2G. This can be due to high volatility of market prices and renewable production. A small value of κ however suggests that the VPP does not heavily rely on V2G to optimize its operation. As κ gets smaller, the payoff of each contract has to go down to ensure the VPP's profit is maximized.

Recall that γ is the perceived value of the battery per charge/discharge cycle. Although it is possible to calculate γ according to the price and cycle life of lithium-ion batteries that are commonly used in EVs, EV owners may think that the value of their battery is higher than its actual price or its lifespan is shorter than the nominal value reported by the manufacturer. This affects their utility and the contract they would accept accordingly. Specifically, when γ is large, EV owners should receive higher payoff to participate in V2G. To investigate how the perceived value of the battery changes the value of the optimal contracts, we assign different values to γ.

In this case study, we assign a value to κ from $\{0.01, 0.05, 0.1, 0.5, 1.0\}$ and to γ from $\{0.01, 0.05, 0.1, 0.5, 1.0, 5.0, 10.0\}$. Notice that the price per cycle of typical EV batteries is around 0.1 €/kWh today.

Figure 5.2 shows the energy-reward bundles that are obtained for different values of κ and γ, when l_{V2G} (i.e., the contract duration) is set to 1 hour. Interestingly, regardless of the value of M, we found at most 3 unique contracts in these cases. From the set of empirical results shown in Figure 5.2, we choose $\kappa = 0.1$, and $\gamma = 1.0$ which results in offering 2 unique contracts for 2 EV owner types. Note that the aforementioned κ and γ value combination results

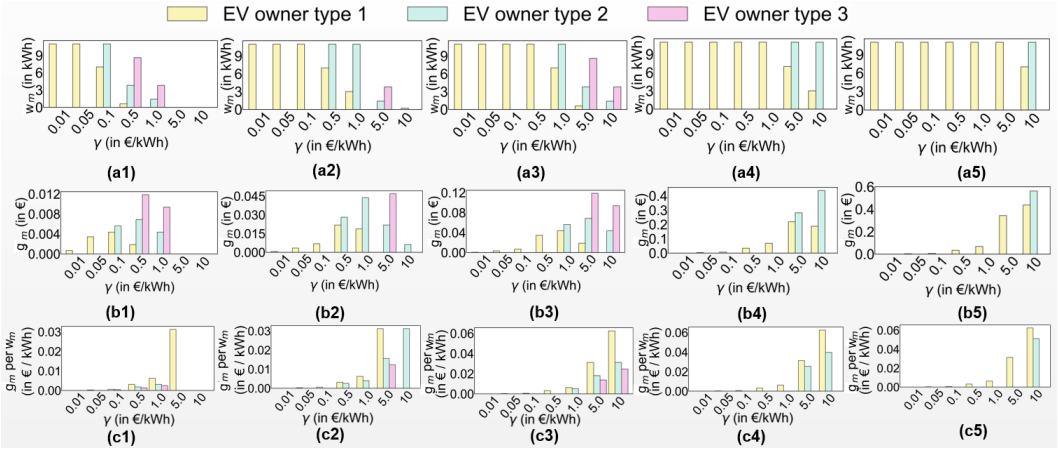


Figure 5.2: For $l_{V2G} = 1$. The κ values from the left to the right column are set as $0.01, 0.05, 0.10, 0.5, 1.0$ respectively. For various γ values (shown on the x-axis of each bar chart), the y-axis of bar charts (a1 to a5) in the first-row show EV discharge energy in offered contract, the y-axis of bar charts (b1 to b5) in the second row show payment offered to EV owners in offered contract, and the y-axis of bar-charts (c1 to c5) in the third row show the payment offered per unit of energy discharged, to EV owners in the offered contract. The number of bars represents the effective number of EV owner types for designing the optimal contract for the chosen parameter values.

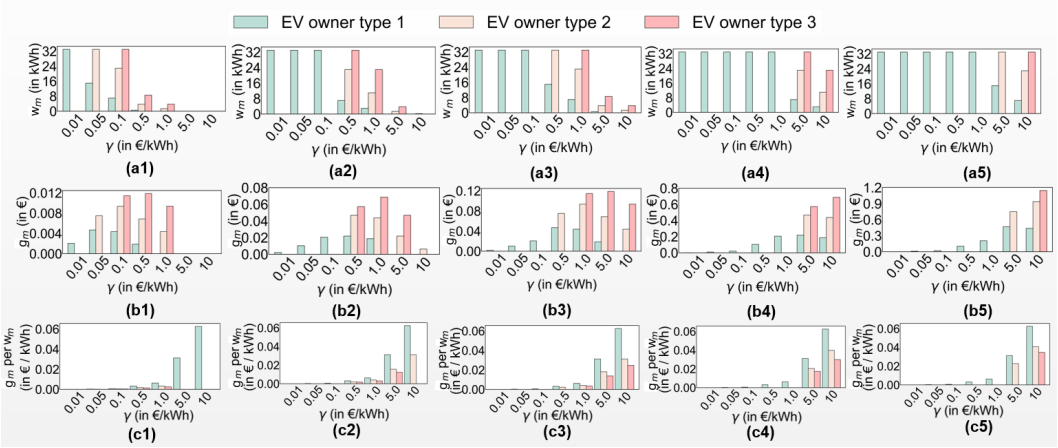


Figure 5.3: For $l_{V2G} = 3$. The κ values from the left to the right column are set as $0.01, 0.05, 0.10, 0.5, 1.0$ respectively. For various γ values (shown on the x-axis of each bar chart), the y-axis of bar charts (a1 to a5) in the first-row shows EV discharge energy in offered contract, the y-axis of bar charts (b1 to b5) in the second-row shows payment offered to EV owners in offered contract, and the y-axis of bar-charts (c1 to c5) in the third-row shows the payment offered per unit of energy discharged, to EV owners in the offered contract. The number of bars represents the effective number of EV owner types for designing the optimal contract for the chosen parameter values.

in the most number of EV types with the highest possible discharge energy, and results in contracts where each contract has a reasonably low payoff to respective EV owners compared to imbalance market price from offering the contracts. By setting the parameters to the aforementioned values, the payoff to EV owners (from (bar chart c3)) of the higher type is 0.0063 €/kWh and to EV owners of the lower type is 0.0051 €/kWh, which are respectively 12.6% and 10.2% of the average IM price of 0.05 €/kWh. As a result, the EV owners receive a reasonable payoff for their V2G contribution, and at the same time, the VPP operator benefits from discharging EVs by reducing the amount of electricity bought from the imbalance market. A similar sensitivity analysis to parameters is also performed for contracts having longer duration. Figure 5.3 shows the energy-reward bundles for the same setting except for l_{V2G} is set to 3 hours in this case. We found from Figure 5.3 that the best-fit values are: $\kappa = 0.05$, $\gamma = 0.5$. The effective number of EV owner types along with the offered monetary incentive and the respective discharge energy can be found in Table 5.1 for the different values of l_{V2G} that we considered in our case study.

5.4 Simulating VPP Operation with Contracts

We analyze the VPP operation given the designed contracts for the EV fleet through simulation. When an EV connects to a bidirectional charger, the VPP calculates its laxity and uses a probabilistic function that maps the laxity upon arrival to the type of an EV owner to estimate its type.[1]

Given the laxity of EV n, we first ignore a) the contracts that are not meaningful for this EV, either because of its short stay time or low energy content of its battery, b) the contracts that if executed the charging deadline of the respective EV cannot be satisfied. Let us define d^n_{max} as the maximum discharge energy available from EV n upon its arrival at the charging station.

[1]Laxity is one of the many factors that determine the type of an EV owner. The other factors, such as their perception of the battery degradation cost, are difficult to model. This is why we use a probabilistic function to determine the EV owner type given their laxity at arrival time.

It is calculated from the reported EV laxity upon arrival using this equation:

$$d^n_{max} = lax^n_t \cdot \psi_{lax} \cdot \alpha_d,$$
$$\psi_{lax} = \frac{\alpha_c \cdot \eta_c \cdot \eta_d}{\alpha_d + \alpha_c \cdot \eta_c \cdot \eta_d},$$

where ψ_{lax} denotes the portion of the laxity that can be used to discharge the battery at the maximum power so that it can be charged to the same level by the end of the laxity if charged at the maximum power.

Now, suppose the specified amount of discharge energy in contracts (g_1, w_1) to (g_m, w_m) is less than or equal to d^n_{max} (i.e., $g_1 \leq \cdots \leq g_m \leq d^n_{max}$) and the amount of discharge energy in contract (g_{m+1}, w_{m+1}) is greater than d^n_{max}. These are the only contracts that will be offered to EV n to ensure that the EV charging problem remains feasible. Accepting any of the remaining contracts, i.e. (g_{m+1}, w_{m+1}) up to (g_M, w_M), could make the charging problem infeasible, so they will not be offered to this EV. Additionally, the VPP will not offer any contract to the EVs that stay in the charging station less than the contract duration l_{V2G}, since no such contract can be executed. Similarly, the VPP will not offer contracts that have a higher amount of discharge energy than the energy stored in the EV battery upon arrival. This is because these contracts cannot be executed.

Next we explain the probabilistic mapping of the observed laxity of EV n to one of the EV types. As mentioned earllier, we assume the EV type is a function of laxity and some external factors. We use lax_{min} to denote the minimum laxity for which a discharge energy of w_m amount is feasible and the maximum possible EV laxity is found (as per our dataset) as lax_{max}. We want to map lax^n_t (i.e., the observed laxity of EV n) to θ_x, where θ_x is between θ_m and θ_M. We calculate θ_x by solving the equation below:

$$\frac{\theta_x - \theta_m}{lax^n_t - lax_{min}} = \frac{\theta_M - \theta_m}{lax_{max} - lax_{min}} \tag{5.35}$$

Finally, we find the EV owner type (denoted θ_i) by sampling (and rounding off the resulting value) from a Gaussian distribution (truncated between 0 and 1) whose mean is θ_x. The variance of this distribution is a fixed value and it is set as 0.1 in our simulations.

5.4.1 Designing a VPP Operating Strategy

Once the EV owner accepts the contract (g_i, w_i), the VPP operator needs to set a schedule to use up the discharge energy w_i within l_{V2G} time slots while ensuring that the VPP profit is maximized. We use $\bar{w}_i^n$ and l_{V2G}^n to denote respectively the remaining discharge energy (as specified in the contract) and the remaining contract duration for the n^{th} EV within the set $\mathcal{N}_t^D$. We use a heuristic EV scheduling algorithm (shown in Algorithm 2), which schedules the charge and discharge of EV batteries in an attempt to increase the VPP profit.

The laxity of EV n at time slot t can be calculated given the EV departure time, battery capacity, target SoC, the maximum charging rate supported by the charger, and the charge efficiency of the battery:

$$lax_t^n = t_e^n - t - \frac{(\overline{SoC}^n - SoC_t^n) \cdot b_n}{\alpha_c \eta_c} \tag{5.36}$$

Algorithm 2: VPP Operating Strategy

```
1  S_1 ← FindEVsWithNegativeLaxity(N_t \ N_t^D) ;        // lookahead
2  e_t ← ChargeEVs(S_1);
3  x_t ← x_t + e_t;
4  P̄^IM ← GetPriceForecasts(t, t + l_V2G);
5  d_t ← GetEVChargingSchedule(N_t^D, P̄^IM);
6  UpdateContractParameters(N_t^D);
7  x_t ← x_t + d_t;
8  if x_t > 0 then
9  |  BuyFromImbalanceMarket(x_t);
10 end
11 else if x_t < 0 then
12 |  SellToImbalanceMarket(x_t);
13 end
```

In Algorithm 2, which runs in the beginning of each time slot, the FindEVsWithNegativeLaxity() function in Line 1 uses the set of non-V2G participating EVs (that are present at the charging station) as input. Using laxity lookahead (described in Section 4.2.1), this function returns the set of EVs (denoted by $\mathcal{S}_1$) that will have a negative laxity in the next time slot and

therefore must be charged at charging power of α_c kW in the current time slot to keep their laxity non-negative. When calculating the laxity lookahead, the t variable is substituted with $t+1$ in Equation 5.36 and SoC^n_{t+1} is set equal to SoC^n_t. Then in Line 2, the set $\mathcal{S}_1$ is passed to the ChargeEVs() function that charges the EVs in $\mathcal{S}_1$ at α_c charging power. The output of the ChargeEVs() function, denoted e_t, is added to the day ahead commitment at time t (i.e., x_t) to update the total amount of energy required by the VPP in this time slot (Line 3). In Line 4, the GetPriceForecasts() function returns a vector of imabalance market price forecasts, starting from the current time step t up to time step $t + l_{V2G}$. In this thesis, the GetPriceForecasts() has a simple implementation and returns the expected values of the imbalance market prices for the next l_{V2G} time slots as the forecast values. In line 5, the GetEVChargingSchedule() function returns the charging profile of all the V2G-participating EVs that are present at the charging station in the current time slot by solving the optimization problem (given in Eqaution 5.37). The solution is the optimal charging/discharging schedule of each EV until the end of their stay time or until $t+l_{V2G}$, whichever happens earlier. Line 5 is executed for the V2G-participating EVs in every time slot to update their EV schedules using more accurate information of the stochastic variables (namely, EV mobility patterns and electricity market prices) for best results. The obtained EV schedule ensures that their laxity will stay non-negative at every time step of their charging duration. In Line 6, the UpdateContractParameters() updates the V2G-participating EVs' remaining contract duration according to this rule: $l^n_{V2G} = l^n_{V2G} - 1$, and their remaining discharge energy (as per the contract) according to this rule: $\bar{w}^n_i = \bar{w}^n_i - AD^{*n}$, where AD^{*n} is the solution of the optimization problem. This function also updates the set of V2G-participating EVs (i.e. $\mathcal{N}^D_t$) by adding the newly arrived V2G participating EVs to $\mathcal{N}^D_t$. The amount of energy required in this time slot is updated again by adding the energy from the V2G-participating EVs, denoted d_t, to the day ahead commitment for this time slot, x_t, in Line 7. Note that the day-ahead commitments are calculated using the 'Wait-and-See' method explained in Section 4.1. Now, if the value of x_t is positive (Line 8), then the

VPP purchases the required x_t amount of energy from the IM market (Line 9). In the case where x_t is negative (Line 11), there is some surplus energy available to the VPP even after charging the EVs, and hence the VPP sells the surplus in the IM market (Line 13).

5.4.2 Scheduling EV Charging

The optimal solutions AC^* and AD^* obtained from the optimization Problem (5.37) constitute the charging schedule of all EVs such that the discharge energy of the EV owners' chosen contract is used to maximize the VPP profit. In the optimization problem formulation, $\bar{p}_t^{IM}$ in the objective function in Equation (5.37a) represents the predicted imbalance market price at time t. The set $\mathcal{T}^n$ contains time steps starting from the current time step t and upto t_e^n (i.e. the stay time of EV n at the charging station). Constraint (5.37b) ensures that discharging does not occur for EVs once their contract duration time has elapsed. The EV discharging amount for EV n at time t, (i.e., AD_t^n) must be bounded according to minimum value between the remaining discharge energy specified in the contract and the maximum discharging power supported by the EV charger which is shown in Constraint (5.37c). Constraint (5.37d) ensures that the total discharge amount of EV n at time t does not exceed its remaining contract-specified discharge energy amount, $\bar{w}_i^n$. The charging energy is bounded in Constraint (5.37e) using the maximum charger-supported charging rate. Constraint (5.37f) defines the EV schedule for EV n in time slot t as the summation of the energy that is charged in the EV battery AC_t^n, and energy discharged from its battery, AD_t^n. This is essential since charge and discharge efficiencies (shown in Constraint (5.37j)) could be less than 1. Constraints (5.37g) bounds the SoC level of the EV battery and the SoC level at arrival time and departure time are defined in Constraints (5.37h) and (5.37i) respectively. The SoC of EV batteries at time t are updated based on the charging and discharging power supplied to them at the previous time step and taking into account the battery charge/discharge efficiencies and it is defined in Constraint (5.37j). Finally, Constraint (5.37k) uses laxity lookahead to ensure positive laxity of the EV at all times during its charging/discharging

process. It is to be noted that all constraints in optimization Problem (5.37) are affine.

$$\underset{AC,AD}{\text{maximize}} \quad \sum_{t}^{t+l_{V2G}} (-x_t - d_t) \cdot \bar{p}_t^{IM} \tag{5.37a}$$

$$\text{subject to} \quad AD_t^n = 0, \quad \forall n \in \mathcal{N}_t^D \tag{5.37b}$$
$$\forall t \in \mathcal{T}^n \setminus \{t_s^n, ..., (t_s^n + (l_{V2G}^n - 1))\}$$
$$-\min\left(\bar{w}_i^n, \alpha_d\right) \leq AD_t^n \leq 0, \tag{5.37c}$$
$$\forall n \in \mathcal{N}_t^D, \forall t \in \mathcal{T}^n$$
$$\sum_{t}^{t+l_{V2G}^n} AD_t^n \leq \bar{w}_i^n, \quad \forall n \in \mathcal{N}_t^D, \forall t \in \mathcal{T}^n \tag{5.37d}$$
$$0 \leq AC_t^n \leq \alpha_c, \ \ \forall n \in \mathcal{N}_t^D, \forall t \in \mathcal{T}^n \tag{5.37e}$$
$$d_t^n = AC_t^n + AD_t^n, \quad \forall n \in \mathcal{N}_t^D, \forall t \in \mathcal{T}^n \tag{5.37f}$$
$$\delta_{min} \leq SoC_t^n \leq \delta_{max}, \quad \forall n \in \mathcal{N}_t^D, \forall t \in \mathcal{T}^n \tag{5.37g}$$
$$SoC_{t_s^n}^n = \underline{SoC}^n, \quad \forall n \in \mathcal{N}_t^D \tag{5.37h}$$
$$SoC_{t_e^n}^n = \overline{SoC}^n, \quad \forall n \in \mathcal{N}_t^D \tag{5.37i}$$
$$SoC_{t+1}^n = SoC_t^n + \frac{AC_t^n \eta_c}{b_n} + \frac{AD_t^n}{\eta_d b_n}, \tag{5.37j}$$
$$\forall n \in \mathcal{N}_t^D, \forall t \in \mathcal{T}^n$$
$$lax_{t+1}^n \geq 0, \quad \forall n \in \mathcal{N}_t^D, \forall t \in \mathcal{T}^n \tag{5.37k}$$

5.4.3 Dataset and Parameters Used for Experiments

Datasets containing real traces of EV charging sessions and electricity market prices (DA and IM) from the Rotterdam region in the Netherlands were used that began on January 1, 2020 and ended on December 31, 2020. These datasets were used for evaluating the proposed contract theoretic approach for V2G participation in addition to the scheduling algorithm with the baseline (described in the following section).

Regarding the electricity market prices, the DA market price data were used from the the European Network of Transmission System Operators [35], and the IM market prices from the regional Transmission System Operator, called TENNET [108]. For the EV data, the charging dataset released by ElaadNL (which is a large-scale charging infrastructure in the Netherlands [33]) was used. This dataset comprises of 10,000 charging events that occurred in

various EVnetNL[2]-operated public chargers. It is to be noted that the initial (i.e. arrival) and target/final SoC level of the EV batteries are not reported in this dataset and hence it is assumed for all EVs to have target SoC of 1. The initial SoC is calculated based on the total charging energy required by this EV during its charging session.

In our experiments, $\alpha_c, \alpha_d > 0$ are both taken to be 11 kW, δ_{min} and δ_{max} are assumed to be 0.03 and 0.97 respectively and η_c and η_d are both set to 0.98. Regarding the parameters pertaining to battery degradation cost, L is assumed to be $3,500$ full cycles, DoD is assumed to be 100%, and b_n is taken to be 80 kWh (which is the battery capacity of Tesla Model 3).

5.4.4 Annual Profit Comparison

The difference in profit earned between the proposed algorithm leveraging contract theory with the prevalent practice of charging EVs (dubbed CHRG_ASAP, which entails charging the EVs at the maximum charging rate at the moment that the EV is connected to the charger and assumes the absence of V2G technology) is shown in Figure 5.4. We incorporate the payment received by the VPP from EV owners after they have completed charging. The EV owners pay 0.13 €/kWh, which is the 95^{th} percentile of the buying price in the imbalance market. It can be seen that the VPP operator can achieve significant improvement in annual profit when utilizing V2G technology despite offering contracts to EV owners in comparison to the prevalent practice of EV charging without V2G. Furthermore, when comparing between the annual VPP profit for different l_{V2G} values of the contract, it can be deduced from Figure 5.4 that offering contracts where $l_{V2G} = 3$ results in a higher annual profit for the VPP operator. This can be explained using Table 5.1 which shows that the total discharge energy available to the VPP operator is greater when setting $l_{V2G} = 3$ compared to $l_{V2G} = 1$. We find that the annual VPP profit increases (in comparison to the profit earned when adopting CHRG_ASAP baseline) by 6.5% and 23.1% when contracts are offered to 25% and 100% of the EV fleet respectively when $l_{V2G} = 1$. The annual VPP profit increases (in comparison

[2]https://evnet.nl/

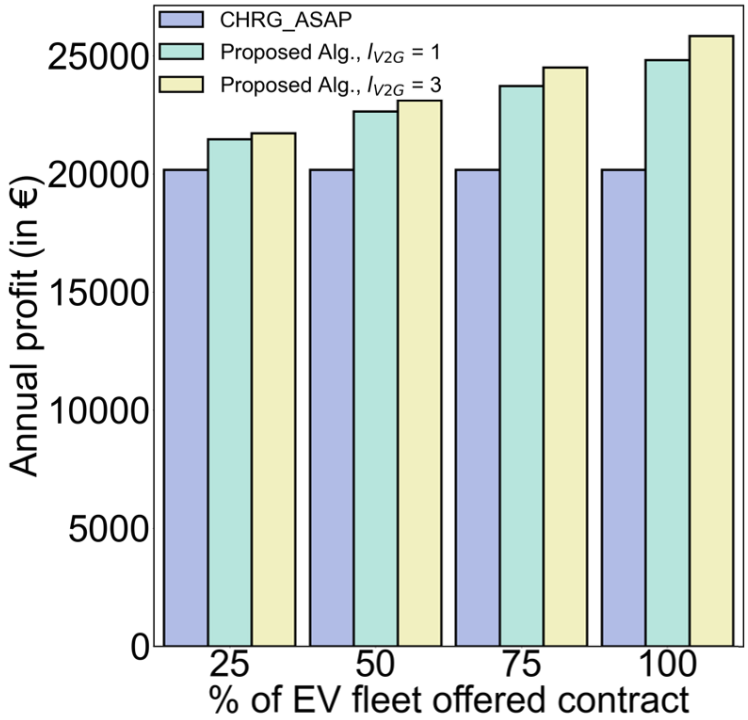


Figure 5.4: Comparison of annual VPP profit for various l_{V2G} values with the baseline.

	Contract (Incentive in €, Energy in kWh)		
l_{V2G} **(hours)**	**EV Type-1**	**EV Type-2**	**EV Type-3**
1	$0.044, 7$	$0.056, 11$	N/A
3	$0.021, 7$	$0.047, 23$	$0.057, 33$

Table 5.1: Contracts for l_{V2G} values

to the profit earned when adopting CHRG_ASAP baseline) by 7.7% and 28.1% when contracts are offered to 25% and 100% of the EV fleet respectively when $l_{V2G} = 3$.

Figure 5.5 shows the number of EVs in V2G participation over a period of 1 year of simulation. It shows a lower number of EVs in V2G participation when l_{V2G} is set to 3. The decrease in the number of V2G-participating EVs for $l_{V2G} = 3$ is because fewer EVs stay long enough at the charging station to participate in V2G when $l_{V2G} = 3$. As a result, the VPP operator filters out a greater number of EVs when $l_{V2G} = 3$ in comparison to when $l_{V2G} = 1$.

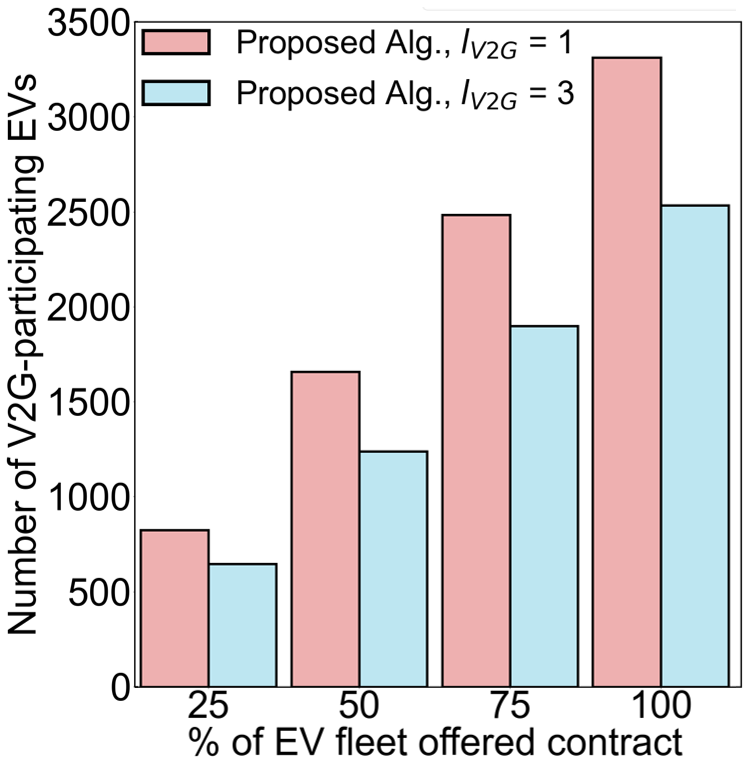


Figure 5.5: Comparison of the number of V2G-participating EVs for various l_{V2G} values during the 1 year period of simulation.

Chapter 6

Conclusion

Flexibility in the power grid is of utmost importance as it keeps to maintain reliability and stability during adverse grid events, which are expected to occur more frequently due to the increasing penetration of intermittent renewable energy sources in the grid and rising temperature. With the advent of bidirectional charging technology, power system operators can leverage the flexibility offered by EVs when they are connected to the grid to replenish their battery. This reduces the cost of providing flexibility.

In this thesis, we considered a VPP that aggregates and orchestrates DERs (with or without solar systems), an EV fleet, and bidirectional chargers. This VPP participates in multi-stage electricity markets. Given the uncertainty of EV mobility and solar generation, the risk involved in the multi-stage electricity market participation, and the challenge of fulfilling EV charging demand, this problem is complex and non-trivial. In the absence of efficient operating strategies, it is hard to create such VPPs to improve the reliability of the power grid.

The goal of this thesis is to find strategies that allow such a VPP to address the aforementioned uncertainties and risks, while being profitable. Towards this end, we proposed operational strategies for this VPP when participating in multi-stage electricity markets, and evaluated the profitability of this VPP under the proposed operational strategies, by carrying out simulations using real data (EV mobility, solar generation, and market prices).

We then extended our research to the scenario where a VPP comprises an

independently-owned EV fleet with V2G-enabled chargers, and participates in a two-stage electricity market. We designed several incentive-compatible contracts. By accepting a contract, the EV owners enter into an agreement with the VPP for V2G participation that ensures that the EV owners are provided sufficient incentives considering their battery degradation (due to V2G participation).

Throughout the thesis, we have answered all the research questions that were presented in the first chapter. We summarize them below:

- In Chapter 3, we showed how to overcome the inherent challenges of operating a VPP that participates in a two-stage electricity market by considering an emerging type of VPP that integrates a fleet of EVs with bidirectional chargers and solar systems. We highlighted that operating the DERs in a VPP is a challenging task owing to the large number of stochastic processes that govern demand, supply, and market prices.

- We presented algorithms to solve the stochastic decision-making problem which determines VPP operation. We proposed efficient and practical operating strategies for this VPP when participating in a two-stage electricity market in Chapter 4. Specificaly, we proposed one heuristic and one RL-based algorithm with a differentiable projection layer to maximize the profit of this VPP on the operation day, given the day-ahead commitments. This VPP places energy bids in the DA market according to the average solution of a sequence of deterministic linear programs solved for different realizations of random variables (i.e., forecast scenarios), and trades in the IM market to honor its day-ahead commitments and satisfy the EV charging demands.

- In Chapter 4, we investigated the profitability of this VPP under the proposed operating strategy. We carried out the evaluation of VPP profitability using real data pertaining to a specific region in the Netherlands and compared it with the offline optimal and current EV charging

baselines. In our results, we showed that the proposed operating strategies exhibit strong performance and outperform the prevalent practice of EV charging.

- We showed how varying levels of V2G participation could affect the VPP profit. Our results show that enabling V2G can substantially increase the profit of this kind of VPP due to increased flexibility available to the VPP.

In Chapter 5 of this thesis, we studied how to incentivize independent EV owners to participate in VPPs, and what factors influence EV owners' willingness to participate in V2G. We explored the profitability of VPPs after accounting for the payoff provided to the EV owners. Towards this end, we explored a related but structurally different VPP, in which the EV owners and the owner of the charging stations (*i.e.*, the aggregator) are different agents. In this case, the aggregator was required to design a billing mechanism to charge the EV owners for the service and another mechanism to incentivize them to participate in V2G (given the battery degradation cost), while making sure that the VPP remains profitable and charging deadlines are met.

- We showed how to incentivize independent EV owners to participate in VPPs by proposing a contract theoretic approach to address the surplus or deficit in the available energy faced by the VPP operator in real-time due to the inherent stochasticity of EV mobility and energy market prices.

- We analyzed the factors that influence EV owners' willingness to participate in V2G by presenting a sensitivity analysis of the parameters and contract duration when designing the contracts.

- We conducted simulations (based on real datasets) and showed a significant profit improvement in the proposed V2G-based scheme compared to the prevalent practice of V2G-disabled EV charging. Furthermore, a scheduling algorithm has been presented that optimizes the EV

charge/discharge schedule of V2G-participating EVs to maximize VPP profit.

We plan to address the following limitations of this thesis in future work:

- To calculate the laxity of an EV upon arrival at the charging station, the EV departure time was assumed to be always accurate. In future work, we plan to design incentive mechanisms such that it is in the EV owners' best interest to communicate accurate departure times to the VPP.
- We plan to consider grid constraints in the design of VPP operating strategies.
- The operational costs of the VPP are not taken into account during VPP profit calculation. We plan to revisit the profitability of the VPP after incorporating the operating costs.
- The efficacy of the control methodologies (developed and evaluated using Netherlands electricity market data) was not evaluated on electricity markets of other countries and regions. We intend to run experiments using data from other markets and jurisdictions.
- Only one policy gradient RL method, namely Soft Actor-Critic (SAC) was used for solving the stage-2 (being sequential decision-making) problem. So, future work lies on comparing the LA-SAC method against other policy gradient RL techniques such as Deep Deterministic Policy Gradient (DDPG) method, Trust Region Policy Optimization (TRPO), or Proximal Policy Optimization (PPO) by using differentiable projection within the aforementioned algorithms to solve the stage-2 problem.

We have developed all the operating strategies in this thesis to be generalizable to other VPP settings that consist of other kinds of DERs. Therefore, we believe that the algorithms and incentive mechanisms presented in this work would prove beneficial to the research community in providing insight into the design of more innovative, yet effective VPP control strategies, thereby

accelerating the adoption of VPPs. We envision a green, safe, and sustainable power grid in which VPPs provide significant flexibility at the minimum cost.

References

[1] K. Abdulla *et al.*, "Optimal operation of energy storage systems considering forecasts and battery degradation," *IEEE Transactions on Smart Grid*, vol. 9, no. 3, pp. 2086–2096, 2016.

[2] W. S. W. Abdullah, M. Osman, M. Z. A. Ab Kadir, R. Verayiah, N. F. Ab Aziz, and M. A. Rasheed, "Techno-economics analysis of battery energy storage system (bess) design for virtual power plant (vpp)–a case study in malaysia," *Journal of Energy Storage*, vol. 38, p. 102 568, 2021.

[3] J. Achiam *et al.*, "Constrained policy optimization," in *International Conference on Machine Learning*, PMLR, 2017, pp. 22–31.

[4] K. Adu-Kankam *et al.*, "Towards collaborative virtual power plants: Trends and convergence," *Sustainable Energy, Grids and Networks*, vol. 16, pp. 217–230, 2018.

[5] A. Agrawal *et al.*, "Differentiable convex optimization layers," *Advances in neural information processing systems*, vol. 32, 2019.

[6] U. Agwan *et al.*, "Pricing in prosumer aggregations using reinforcement learning," in *Proceedings of the Twelfth ACM International Conference on Future Energy Systems*, ACM, 2021, pp. 220–224.

[7] A. Al Zishan *et al.*, "Adaptive control of plug-in electric vehicle charging with reinforcement learning," in *Proceedings of the Eleventh ACM International Conference on Future Energy Systems*, ACM, 2020, pp. 116–120.

[8] A. Alabdulatif *et al.*, "Privacy-preserving cloud-based billing with lightweight homomorphic encryption for sensor-enabled smart grid infrastructure," *IET Wireless Sensor Systems*, vol. 7, no. 6, pp. 182–190, 2017. DOI: https://doi.org/10.1049/iet-wss.2017.0061. eprint: https://ietresearch.onlinelibrary.wiley.com/doi/pdf/10.1049/iet-wss.2017.0061. [Online]. Available: https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/iet-wss.2017.0061.

[9] I. Alotaibi, M. A. Abido, M. Khalid, and A. V. Savkin, "A comprehensive review of recent advances in smart grids: A sustainable future with renewable energy resources," *Energies*, vol. 13, no. 23, p. 6269, 2020.

[10] M. Amini and M. Almassalkhi, "Trading off robustness and performance in receding horizon control with uncertain energy resources," in *2018 Power systems computation conference (PSCC)*, IEEE, 2018, pp. 1–7.

[11] M. Amini and M. Almassalkhi, "Optimal corrective dispatch of uncertain virtual energy storage systems," *IEEE Transactions on Smart Grid*, vol. 11, no. 5, pp. 4155–4166, 2020.

[12] M. Amini, A. Khurram, A. Klem, M. Almassalkhi, and P. D. Hines, "A model-predictive control method for coordinating virtual power plants and packetized resources, with hardware-in-the-loop validation," in *2019 IEEE Power & Energy Society General Meeting (PESGM)*, IEEE, 2019, pp. 1–5.

[13] A. Bagchi *et al.*, "Adequacy assessment of generating systems incorporating storage integrated virtual power plants," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 3440–3451, 2019.

[14] K. Baker *et al.*, "Distributed MPC for efficient coordination of storage and renewable energy sources across control areas," *IEEE Transactions on Smart Grid*, vol. 7, no. 2, pp. 992–1001, 2016.

[15] B. Behi, A. Arefi, P. Jennings, A. Gorjy, and A. Pivrikas, "Advanced monitoring and control system for virtual power plants for enabling customer engagement and market participation," *Energies*, vol. 14, no. 4, p. 1113, 2021.

[16] B. Behi, A. Baniasadi, A. Arefi, A. Gorjy, P. Jennings, and A. Pivrikas, "Cost–benefit analysis of a virtual power plant including solar pv, flow battery, heat pump, and demand management: A western australian case study," *Energies*, vol. 13, no. 10, p. 2614, 2020.

[17] R. Bellman and R. E. Kalaba, *Dynamic programming and modern control theory*. Citeseer, 1965, vol. 81.

[18] S. Bianchi, A. De Filippo, S. Magnani, G. Mosaico, and F. Silvestro, "Virtus project: A scalable aggregation platform for the intelligent virtual management of distributed energy resources," *Energies*, vol. 14, no. 12, p. 3663, 2021.

[19] P. Bolton and M. Dewatripont, *Contract theory*. MIT press, 2004.

[20] K. Bradbury *et al.*, "Economic viability of energy storage systems based on price arbitrage potential in real-time us electricity markets," *Applied Energy*, vol. 114, pp. 512–519, 2014.

[21] D. I. Candra, K. Hartmann, and M. Nelles, "Economic optimal implementation of virtual power plants in the german power market," *Energies*, vol. 11, no. 9, p. 2365, 2018.

[22] Z. Chang, D. Zhang, T. Hämäläinen, Z. Han, and T. Ristaniemi, "Incentive mechanism for resource allocation in wireless virtualized networks with multiple infrastructure providers," *IEEE Transactions on Mobile Computing*, vol. 19, no. 1, pp. 103–115, 2018.

[23] B. Chen *et al.*, "Enforcing policy feasibility constraints through differentiable projection for energy optimization," in *Proceedings of the Twelfth ACM International Conference on Future Energy Systems*, ACM, 2021, pp. 199–210.

[24] T. Cioara, M. Antal, V. T. Mihailescu, C. D. Antal, I. M. Anghel, and D. Mitrea, "Blockchain-based decentralized virtual power plants of small prosumers," *IEEE Access*, vol. 9, pp. 29 490–29 504, 2021.

[25] G. E. Constante-Flores and M. S. Illindala, "Data-driven probabilistic power flow analysis for a distribution system with renewable energy sources using monte carlo simulation," *IEEE Transactions on Industry Applications*, vol. 55, no. 1, pp. 174–181, 2019. DOI: 10.1109/TIA.2018.2867332.

[26] G. Dalal *et al.*, "Safe exploration in continuous action spaces," *arXiv preprint arXiv:1801.08757*, 2018.

[27] T. N. Dang, K. Kim, L. U. Khan, S. A. Kazmi, Z. Han, and C. S. Hong, "On-device computational caching-enabled augmented reality for 5g and beyond: A contract-theory-based incentive mechanism," *IEEE Internet of Things Journal*, vol. 8, no. 24, pp. 17 382–17 394, 2021.

[28] P. Dasgupta, P. Hammond, and E. Maskin, "The implementation of social choice rules: Some general results on incentive compatibility," *The Review of Economic Studies*, vol. 46, no. 2, pp. 185–216, 1979.

[29] A. Delle Femine, D. Gallo, C. Landi, A. Lo Schiavo, and M. Luiso, "Low power contactless voltage sensor for low voltage power systems," *Sensors*, vol. 19, no. 16, p. 3513, 2019.

[30] S. Y. Derakhshandeh *et al.*, "Coordination of generation scheduling with pevs charging in industrial microgrids," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 3451–3461, 2013.

[31] S. Diamond *et al.*, "CVXPY: A Python-embedded modeling language for convex optimization," *Journal of Machine Learning Research*, vol. 17, no. 83, pp. 1–5, 2016.

[32] A. P. Dobos, "Pvwatts version 5 manual," 2014. [Online]. Available: https://www.osti.gov/biblio/1158421.

[33] *Elaadnl open data*. [Online]. Available: https://platform.elaad.io/download-data/.

[34] Energy Networks Australia, *Electricty network transformation roadmap: Final report*, 2017. [Online]. Available: https://www.energynetworks.com.au/resources/reports/electricity-network-transformation-roadmap-final-report/.

[35] *Entso-e transparency platform*. [Online]. Available: https://transparency.entsoe.eu.

[36] EU, *Decisions of the agency for the cooperation of energy cooperation of energy regulators no 18-2020*, 2020. [Online]. Available: https://nordicbalancingmodel.net/roadmap-and-projects/single-price-model/.

[37] S. Fujimoto *et al.*, "Benchmarking batch deep reinforcement learning algorithms," *arXiv preprint arXiv:1910.01708*, 2019.

[38] T. Gabderakhmanova, J. Engelhardt, J. M. Zepter, *et al.*, "Demonstrations of dc microgrid and virtual power plant technologies on the danish island of bornholm," in *2020 55th International Universities Power Engineering Conference (UPEC)*, IEEE, 2020, pp. 1–6.

[39] Y. Gao *et al.*, "A contract-based approach for ancillary services in v2g networks: Optimality and learning," in *2013 Proceedings IEEE INFOCOM*, IEEE, 2013, pp. 1151–1159.

[40] J. Garcıa *et al.*, "A comprehensive survey on safe reinforcement learning," *Journal of Machine Learning Research*, vol. 16, no. 42, pp. 1437–1480, 2015.

[41] M. Giuntoli and D. Poli, "Optimized thermal and electrical scheduling of a large scale virtual power plant in the presence of energy storages," *IEEE Transactions on Smart Grid*, vol. 4, no. 2, pp. 942–955, 2013.

[42] Gurobi Optimization, LLC, *Gurobi Optimizer Reference Manual*, 2021. [Online]. Available: https://www.gurobi.com.

[43] T. Haarnoja *et al.*, "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor," in *International conference on machine learning*, PMLR, 2018, pp. 1861–1870.

[44] Z. Hasan and V. K. Bhargava, "Relay selection for ofdm wireless systems under asymmetric information: A contract-theory based approach," *IEEE Transactions on Wireless Communications*, vol. 12, no. 8, pp. 3824–3837, 2013.

[45] M. Hasanbeig, A. Abate, and D. Kroening, "Cautious reinforcement learning with logical constraints," in *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*, ser. AAMAS '20, Auckland, New Zealand: International Foundation for Autonomous Agents and Multiagent Systems, 2020, pp. 483–491, ISBN: 9781450375184.

[46] I. A. Hiskens and B. Gong, "Mpc-based load shedding for voltage stability enhancement," in *Proceedings of the 44th IEEE Conference on Decision and Control*, IEEE, 2005, pp. 4463–4468.

[47] R.-A. Hooshmand, S. M. Nosratabadi, and E. Gholipour, "Event-based scheduling of industrial technical virtual power plant considering wind and market prices stochastic behaviors-a case study in iran," *Journal of cleaner production*, vol. 172, pp. 1748–1764, 2018.

[48] N. Hunt *et al.*, "Verifiably safe exploration for end-to-end reinforcement learning," in *Proceedings of the 24th International Conference on Hybrid Systems: Computation and Control*, 2021, pp. 1–11.

[49] R. Iacobucci, B. McLellan, and T. Tezuka, "Costs and carbon emissions of shared autonomous electric vehicles in a virtual power plant and microgrid with renewable energy," *Energy Procedia*, vol. 156, pp. 401–405, 2019.

[50] IRENA, *Innovation landscape brief: Market integration of distributed energy resources*, 2019. [Online]. Available: `https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2019/Feb/IRENA_Market_integration_distributed_system_2019.pdf`.

[51] J. Iria *et al.*, "Optimal bidding strategy for an aggregator of prosumers in energy and secondary reserve markets," *Applied Energy*, vol. 238, pp. 1361–1372, 2019.

[52] N. Jansen, B. Könighofer, S. Junges, A. Serban, and R. Bloem, "Safe reinforcement learning using probabilistic shields," 2020.

[53] A. G. Jember *et al.*, "Game and contract theory-based energy transaction management for internet of electric vehicle," *IEEE Access*, vol. 8, pp. 203 478–203 487, 2020.

[54] N. Jenkins, C. Long, and J. Wu, "An overview of the smart grid in great britain," *Engineering*, vol. 1, no. 4, pp. 413–421, 2015, ISSN: 2095-8099. DOI: `https://doi.org/10.15302/J-ENG-2015112`. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S2095809916300224`.

[55] C. Jin *et al.*, "Optimizing electric vehicle charging with energy storage in the electricity market," *IEEE Transactions on Smart Grid*, vol. 4, no. 1, pp. 311–320, 2013.

[56] L. Ju, Z. Tan, J. Yuan, Q. Tan, H. Li, and F. Dong, "A bi-level stochastic scheduling optimization model for a virtual power plant connected to a wind–photovoltaic–energy storage system considering the uncertainty and demand response," *Applied energy*, vol. 171, pp. 184–199, 2016.

[57] E. L. Karfopoulos *et al.*, "Distributed coordination of electric vehicles providing v2g regulation services," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 2834–2846, 2015.

[58] M. J. Kasaei, M. Gandomkar, and J. Nikoukar, "Optimal management of renewable energy sources by virtual power plant," *Renewable energy*, vol. 114, pp. 1180–1188, 2017.

[59] M. Kazemi *et al.*, "Operation scheduling of battery storage systems in joint energy and ancillary services markets," *IEEE Transactions on Sustainable Energy*, vol. 8, no. 4, pp. 1726–1735, 2017.

[60] F. Kazhamiaka *et al.*, "Adaptive battery control with neural networks," in *Proceedings of the Tenth ACM International Conference on Future Energy Systems*, ACM, 2019, pp. 536–543.

[61] Kazmi *et al.*, "A novel contract theory-based incentive mechanism for cooperative task-offloading in electrical vehicular networks," *IEEE Transactions on Intelligent Transportation Systems*, 2021.

[62] S. Kim *et al.*, "A guide to sample average approximation," in *Handbook of Simulation Optimization*, Springer, 2015, ch. 8, pp. 207–243.

[63] S. G. Krantz and H. R. Parks, *The implicit function theorem: history, theory, and applications*. Springer Science & Business Media, 2002.

[64] Z. Liang *et al.*, "Risk-constrained optimal energy management for virtual power plants considering correlated demand response," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 1577–1587, 2019.

[65] Z. Liang, Q. Alsafasfeh, T. Jin, H. Pourbabak, and W. Su, "Risk-constrained optimal energy management for virtual power plants considering correlated demand response," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 1577–1587, 2017.

[66] Y. Liu, H. Xin, Z. Wang, and D. Gan, "Control of virtual power plant in microgrids: A coordinated approach based on photovoltaic systems and controllable loads," *IET Generation, Transmission & Distribution*, vol. 9, no. 10, pp. 921–928, 2015.

[67] R. Lu *et al.*, "Multi-stage stochastic programming to joint economic dispatch for energy and reserve with uncertain renewable energy," *IEEE Transactions on Sustainable Energy*, vol. 11, no. 3, pp. 1140–1151, 2020.

[68] M. Maanavi, A. Najafi, R. Godina, M. Mahmoudian, and E. MG Rodrigues, "Energy management of virtual power plant considering distributed generation sizing and pricing," *applied sciences*, vol. 9, no. 14, p. 2817, 2019.

[69] A. Madansky, "Inequalities for stochastic linear programming problems," *Management Science*, vol. 6, no. 2, pp. 197–204, 1960.

[70] F. Y. Melhem *et al.*, "Energy management in electrical smart grid environment using robust optimization algorithm," *IEEE Transactions on Industry Applications*, vol. 54, no. 3, pp. 2714–2726, 2018.

[71] F. Mwasilu *et al.*, "Electric vehicles and smart grid interaction: A review on vehicle to grid and renewable energy sources integration," *Renewable and Sustainable Energy Reviews*, vol. 34, pp. 501–516, 2014.

[72] P. Naina, H.-S. Rajamani, and K. Swarup, "Modeling and simulation of virtual power plant in energy management system applications," in *2017 7th International Conference on Power Systems (ICPS)*, IEEE, 2017, pp. 392–397.

[73] J. Naughton, M. Cantoni, and P. Mancarella, "A modelling framework for a virtual power plant with multiple energy vectors providing multiple services," in *2019 IEEE Milan PowerTech*, IEEE, 2019, pp. 1–6.

[74] J. Naughton, H. Wang, S. Riaz, M. Cantoni, and P. Mancarella, "Optimization of multi-energy virtual power plants for providing multiple market and local network services," *Electric Power Systems Research*, vol. 189, p. 106 775, 2020.

[75] H.-M. Neumann *et al.*, "The potential of photovoltaic carports to cover the energy demand of road passenger transport," *Progress in Photovoltaics: Research and Applications*, vol. 20, no. 6, pp. 639–649, 2012.

[76] H. Nguyen Duc and N. Nguyen Hong, "Optimal reserve and energy scheduling for a virtual power plant considering reserve activation probability," *Applied Sciences*, vol. 11, no. 20, p. 9717, 2021.

[77] F. Oest, M. Radtke, M. Blank-Babazadeh, S. Holly, and S. Lehnhoff, "Evaluation of communication infrastructures for distributed optimization of virtual power plant schedules," *Energies*, vol. 14, no. 5, p. 1226, 2021.

[78] Pacific Northwest National Library, *Utilizing electric vehicles to assist integration of large penetrations of distributed photovoltaic generation*, 2012.

[79] Pandžić *et al.*, "Offering model for a virtual power plant based on stochastic programming," *Applied Energy*, vol. 105, pp. 282–292, 2013.

[80] S.-W. Park and S.-Y. Son, "Interaction-based virtual power plant operation methodology for distribution system operator's voltage management," *Applied Energy*, vol. 271, p. 115 222, 2020.

[81] M. Pasetti, S. Rinaldi, and D. Manerba, "A virtual power plant architecture for the demand-side management of smart prosumers," *Applied Sciences*, vol. 8, no. 3, p. 432, 2018.

[82] B. S. Pedasingu *et al.*, "Bidding strategy for two-sided electricity markets: A reinforcement learning based framework," in *Proceedings of the 7th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*, ACM, 2020, pp. 110–119.

[83] C. Pop, M. Antal, T. Cioara, I. Anghel, I. Salomie, and M. Bertoncini, "A fog computing enabled virtual power plant model for delivery of frequency restoration reserve services," *Sensors*, vol. 19, no. 21, p. 4688, 2019.

[84] T. Popławski, S. Dudzik, P. Szelag, and J. Baran, "A case study of a virtual power plant (vpp) as a data acquisition tool for pv energy forecasting," *Energies*, vol. 14, no. 19, p. 6200, 2021.

[85] J. W. Pratt, "Risk aversion in the small and in the large," *Econometrica*, vol. 32, no. 1/2, pp. 122–136, 1964, ISSN: 00129682, 14680262. [Online]. Available: `http://www.jstor.org/stable/1913738` (visited on 08/30/2022).

[86] B. Qi *et al.*, "Energyboost: Learning-based control of home batteries," in *Proceedings of the Tenth ACM International Conference on Future Energy Systems*, ACM, 2019, pp. 239–250.

[87] G. Raveduto, V. Croce, M. Antal, C. Pop, I. Anghel, and T. Cioara, "Dynamic coalitions of prosumers in virtual power plants for energy trading and profit optimization," in *2020 IEEE 20th Mediterranean Electrotechnical Conference (MELECON)*, IEEE, 2020, pp. 541–546.

[88] E. Rodrigues, G. Osório, R. Godina, A. Bizuayehu, J. Lujano-Rojas, and J. Catalão, "Grid code reinforcements for deeper renewable generation in insular energy systems," *Renewable and Sustainable Energy Reviews*, vol. 53, pp. 163–177, 2016.

[89] B. Salanié, *The economics of contracts: a primer*. MIT press, 2005.

[90] P. Samadi, A.-H. Mohsenian-Rad, R. Schober, V. W. Wong, and J. Jatskevich, "Optimal real-time pricing algorithm based on utility maximization for smart grid," in *2010 First IEEE International Conference on Smart Grid Communications*, IEEE, 2010, pp. 415–420.

[91] P. Samadi, H. Mohsenian-Rad, R. Schober, and V. W. Wong, "Advanced demand side management for the future smart grid using mechanism design," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1170–1180, 2012.

[92] J. C. Sarmiento-Vintimilla, E. Torres, D. M. Larruskain, and M. J. Pérez-Molina, "Applications, operational architectures and development of virtual power plants as a strategy to facilitate the integration of distributed energy resources," *Energies*, vol. 15, no. 3, p. 775, 2022.

[93] H. Sharma and S. Mishra, "Techno-economic analysis of solar grid-based virtual power plant in indian power sector: A case study," *International Transactions on Electrical Energy Systems*, vol. 30, no. 1, e12177, 2020.

[94] F. Sheidaei and A. Ahmarinejad, "Multi-stage stochastic framework for energy management of virtual power plants considering electric vehicles and demand response programs," *International Journal of Electrical Power & Energy Systems*, vol. 120, p. 106 047, 2020.

[95] Z. Shi *et al.*, "Distributionally robust chance-constrained energy management for islanded microgrids," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 2234–2244, 2018.

[96] M. Shin, D.-H. Choi, and J. Kim, "Cooperative management for pv/ess-enabled electric vehicle charging stations: A multiagent deep reinforcement learning approach," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 5, pp. 3493–3503, 2019.

[97] T. Sikorski, M. Jasiński, E. Ropuszyńska-Surma, *et al.*, "A case study on distributed energy resources and energy-storage systems in a virtual power plant concept: Technical aspects," *Energies*, vol. 13, no. 12, p. 3086, 2020.

[98] T. D. Simão *et al.*, "AlwaysSafe: Reinforcement learning without safety constraint violations during training," in *Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems*, IFAAMAS, 2021, pp. 1226–1235.

[99] J. Soares, N. Borges, C. Lobo, and Z. Vale, "Vpp energy resources management considering emissions: The case of northern portugal 2020 to 2050," in *2015 IEEE Symposium Series on Computational Intelligence*, IEEE, 2015, pp. 1259–1266.

[100] E. Y. Song, G. J. FitzPatrick, and K. B. Lee, "Smart sensors and standard-based interoperability in smart grids," *IEEE sensors journal*, vol. 17, no. 23, pp. 7723–7730, 2017.

[101] *South australia's virtual power plant*. [Online]. Available: `https://www.energymining.sa.gov.au/growth_and_low_carbon/virtual_power_plant`.

[102] A. Stooke, J. Achiam, and P. Abbeel, "Responsive safety in reinforcement learning by pid lagrangian methods," in *International Conference on Machine Learning*, PMLR, 2020, pp. 9133–9143.

[103] Y. Sui *et al.*, "Safe exploration for optimization with gaussian processes," in *International Conference on Machine Learning*, PMLR, 2015, pp. 997–1005.

[104] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[105] *Swell Energy's Virtual Power Plants*. [Online]. Available: https://www.swellenergy.com/.

[106] S. I. Taheri, M. B. Salles, and E. C. Costa, "Optimal cost management of distributed generation units and microgrids for virtual power plant scheduling," *IEEE Access*, vol. 8, pp. 208 449–208 461, 2020.

[107] M. A. Tajeddini *et al.*, "Risk averse optimal operation of a virtual power plant using two stage stochastic programming," *Energy*, vol. 73, pp. 958–967, 2014.

[108] *Tennet export data*. [Online]. Available: https://www.tennet.org/english/operational_management/export_data.aspx.

[109] *Tesla Virtual Power Plant*. [Online]. Available: https://www.tesla.com/support/energy/tesla-virtual-power-plant-pge-2022.

[110] R. Torabi, Á. Gomes, and F. Morgado-Dias, "Energy transition on islands with the presence of electric vehicles: A case study for porto santo," *Energies*, vol. 14, no. 12, p. 3439, 2021.

[111] M. Turchetta, F. Berkenkamp, and A. Krause, "Safe exploration in finite markov decision processes with gaussian processes," *Advances in Neural Information Processing Systems*, vol. 29, pp. 4312–4320, 2016.

[112] Z. Ullah, G. Mokryani, F. Campean, and Y. F. Hu, "Comprehensive review of vpps planning, operation and scheduling considering the uncertainties related to renewable energy sources," *IET Energy Systems Integration*, vol. 1, no. 3, pp. 147–157, 2019.

[113] *United States distributed energy resources outlook*. [Online]. Available: https://www.woodmac.com/news/editorial/der-growth-united-states/.

[114] *Utrecht wants to be the first city to use its electric car fleet as a giant battery*. [Online]. Available: https://www.fastcompany.com/90705832/utrecht-wants-to-be-the-first-city-to-use-its-electric-car-fleet-as-a-giant-battery.

[115] M. Vasirani *et al.*, "An agent-based approach to virtual power plants of wind power generators and electric vehicles," *IEEE Transactions on Smart Grid*, vol. 4, no. 3, pp. 1314–1322, 2013.

[116] A. Wachi and Y. Sui, "Safe reinforcement learning in constrained markov decision processes," in *International Conference on Machine Learning*, PMLR, 2020, pp. 9797–9806.

[117] Y. Wang *et al.*, "Interactive dispatch modes and bidding strategy of multiple virtual power plants based on demand response and game theory," *IEEE Transactions on Smart Grid*, vol. 7, no. 1, pp. 510–519, 2016.

[118] C. Wu *et al.*, "Vehicle-to-aggregator interaction game," *IEEE Transactions on Smart Grid*, vol. 3, no. 1, pp. 434–442, 2011.

[119] A. G. Zamani, A. Zakariazadeh, S. Jadid, and A. Kazemi, "Stochastic operational scheduling of distributed energy resources in a large scale virtual power plant," *International Journal of Electrical Power & Energy Systems*, vol. 82, pp. 608–620, 2016.

[120] M. Zeng *et al.*, "An incentivized auction-based group-selling approach for demand response management in v2g systems," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 6, pp. 1554–1563, 2015.

[121] M. Zeng *et al.*, "Group bidding for guaranteed quality of energy in v2g smart grid networks," in *2015 IEEE International Conference on Communications (ICC)*, IEEE, 2015, pp. 5266–5271.

[122] G. Zhang, C. Jiang, and X. Wang, "Comprehensive review on structure and operation of virtual power plant in electrical system," *IET Generation, Transmission & Distribution*, vol. 13, no. 2, pp. 145–156, 2019.

[123] K. Zhang, Y. Mao, S. Leng, *et al.*, "Optimal energy exchange schemes in smart grid networks: A contract theoretic approach," in *2016 IEEE/CIC International Conference on Communications in China (ICCC)*, IEEE, 2016, pp. 1–6.

[124] Y. Zhang, L. Song, W. Saad, Z. Dawy, and Z. Han, "Contract-based incentive mechanisms for device-to-device communications in cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 10, pp. 2144–2155, 2015.

[125] Y. Zhang *et al.*, "Distributionally robust building load control to compensate fluctuations in solar power generation," in *2019 American Control Conference (ACC)*, 2019, pp. 5857–5863.

[126] W. Zhong *et al.*, "Efficient auction mechanisms for two-layer vehicle-to-grid energy trading in smart grid," in *2017 IEEE International Conference on Communications (ICC)*, IEEE, 2017, pp. 1–6.

[127] Y. Zhou, Z. Wei, G. Sun, K. W. Cheung, H. Zang, and S. Chen, "Four-level robust model for a virtual power plant in energy and reserve markets," *IET Generation, Transmission & Distribution*, vol. 13, no. 11, pp. 2036–2043, 2019.