Optimal Energy Management Strategies for Electric Vehicles: Advanced Control and Learning-Based Perspectives

by

Qian Zhang
B.Sc., Harbin Institute of Technology, 2014
M.Sc., Harbin Institute of Technology, 2016

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Mechanical Engineering

Optimal Energy Management Strategies for Electric Vehicles: Advanced Control and Learning-Based Perspectives

by

Qian Zhang
B.Sc., Harbin Institute of Technology, 2014
M.Sc., Harbin Institute of Technology, 2016

Supervisory Committee

Dr. Yang Shi, Co-Supervisor
(Department of Mechanical Engineering)

Dr. Kui Wu, Co-Supervisor
(Department of Computer Sciences)

Dr. Curran Crawford, Departmental Member
(Department of Mechanical Engineering)

ABSTRACT

Motivated by the goal of transition to a zero-carbon-emission-based economy for climate change mitigation, electrification opportunities are more promising in the transportation sector. Electric Vehicles (EVs) are at the forefront of the energy transition at an expanded rapid pace in the transportation sector. To enable and enhance the energy efficiency, advanced control and optimization will play an important role in EV systems and infrastructure.

However, there are also some difficulties and limitations subject to the imperfection of management and control for EVs. Overall, to further the widespread adoption of EVs, the dissertation mainly includes two parts: 1) Power management for Plug-in Hybrid Electric Vehicles (PHEVs); 2) Charging control for Plug-in Electric Vehicles (PEVs). Chapter 2 deals with the power management and route planning problems for PHEVs, which aims to properly design the control algorithm to find the route that leads to the minimum energy consumption. Chapter 3 pays attention to the high workloads of the PEV in the electric power grids, which concentrates on studying a control algorithm leading to possible reductions in both computation and communication. Chapter 4 focuses on the charging control for PEVs, which explores how to improve the PEV charging efficiency while satisfying safety concerns. Chapter 5 modifies the results in Chapter 4 by taking battery capacity degradation into the optimization problem.

This dissertation proceeds with Chapter 1 by reviewing the state-of-the-art control methods for PEVs and PHEVs. Chapter 2 studies a novel control scheme of route planning with power management for PHEVs. By considering the power management of PHEVs, we aim to find the route that leads to the minimum energy consumption. The scheme adopts a two-loop structure to achieve the control objective. Specifically, in the outer loop, the minimum energy consumption route is obtained by minimizing the difference between the value function of current round and the best value from all previous rounds. In the inner loop, the energy consumption index with respect to PHEV power management for each feasible route is trained with Reinforcement Learning (RL). Under the RL framework, a nonlinear approximator structure, which consists of an actor approximator and a critic approximator, is built to approximate control actions and energy consumption. In addition, the convergence of value function for PHEV power management in the inner loop and asymptotical stability of the closed-loop system are rigorously guaranteed.

Chapter 3 investigates the self-triggered Model Predictive Control (MPC) with Integral Sliding Mode (ISM) method of a networked nonlinear continuous-time system subject to

state and input constraints with additive disturbances and uncertainties. Compared with the standard MPC strategy, the proposed control scheme is designed for PEV charging to reduce the high communication loads caused by a large-scale population of vehicles under centralized charging control architecture. In the proposed scheme, the constrained optimization problem is solved aperiodically to generate control signals and the next execution time, leading to possible reductions in both computation and communication. The motivation of using ISM approach is to reject matched uncertainties. A self-triggered condition that involves a comparison between the cost function values with different execution periods is derived. Besides, the robust MPC with ISM control strategy is rigorously studied depending on the self-triggered scheme.

Chapter 4 proposes a charging control algorithm for the valley-filling problem, while it meets individual charging requirements. We study a decentralized framework of PEV charging problem with a coordination task. An iterative learning-based model predictive charging control algorithm is developed to achieve the valley-filling performance. The design of the decentralized MPC meets individual charging requirements. The iterative learning method approximates the electricity price function and the system state sampled safe set to improve the accuracy of optimization problem calculations. The decentralized problem, in which the individual PEV minimizes its own charging cost, is formulated based on the sum of all power loads.

Chapter 5 studies a modified charging control algorithm based on the previous charging control algorithm in Chapter 4. We propose a charging control algorithm for PEVs using a decentralized MPC framework supplemented by the iterative learning method. By considering the battery aging of PEVs, we aim to find the optimal charging rate that leads to valley-filling performance. The scheme adopts the iterative learning-based method to solve the optimal control problem with the battery aging model. Specifically, the sampled safe set and price function are updated accordingly as the iteration number increases. The battery aging model involves the cost function to approach the real charging scenario. In addition, the recursive feasibility of the proposed optimal control problem for PEV charging with battery aging and asymptotical stability of the closed-loop system are rigorously studied.

Finally, in Chapter 6, the conclusions of the dissertation and some avenues for future potential research are presented.

Table of Contents

List of Tables

List of Figures

ACKNOWLEDGEMENTS

Throughout the writing of this dissertation, I have received a great deal of support and assistance.

I would first like to thank my supervisors, Professor Yang Shi and Professor Kui Wu whose expertise were invaluable in formulating the research questions and methodology. Their insightful feedback pushed me to sharpen my thinking and brought my work to a higher level. They are very patient, passionate, and always ready to provide insightful thoughts and suggestions on academic problems as well as career development. What I have learnt from them grew me not only as a researcher but also as a person. The experience of working together with them benefited and will continue to benefit me for the rest of my life. I also really appreciate some precious opportunities that were selflessly provided by Professor Yang Shi and Professor Kui Wu.

I would also like to thank my committee, Dr. Curran Crawford, for his valuable guidance throughout my studies. He has been continuously supporting and helping me. He provided me with the tools and all of the opportunities that I needed to choose the right direction to further my research. I also would like to express my appreciations to the External Examinar, Dr. Ahmad W. Al-Dabbagh, for his constructive comments in improving my dissertation.

I would like to acknowledge my experiences from my internship in the Applied Control and Information Processing Lab (ACIPL) at the University of Victoria (UVic) for their wonderful collaboration. Past senior students or current professors, Dr. Binxian Mu, Dr. Mingxi Liu, Dr. Yuanye Chen, Dr. Chao Shen, will be my role models. I would particularly like to single out my colleague at UVic, Jicheng Chen. I would like to thank him for his patient support during the preparation of my dissertation.

In addition, I could not have completed this dissertation without the support of my colleagues and friends, Dr. Changxin Liu, Dr. Qi Sun, Dr. Kunwu Zhang, Xinxin Shang, Dr. Songlin Zhuang, Yaning Guo, Henglai Wei, Tianyu Tan, Haoqiang Ji, Qi Zhou, Chen Ma, Huaiyuan Sheng, Zhuo Li, Xiang Sheng, Zhang Zhang, Tianxiang Lu, Chonghan Ma, Binyan Xu, Yufan Dai and Yue Song, who provided stimulating discussions as well as happy distractions to rest my mind outside of my research. I want to thank you all for the happy time we have spent on this small island.

Last but not least, I would like to thank my parents and all my families for their wise counsel and sympathetic ear. They are always there for me. I love them all with all my heart.

Victoria, B.C., Canada
September 2021

To my beloved parents

Acronyms

BEV Battery Electric Vehicle

DCFC Direct Current Fast Charger

DP Dynamic Programming

EM Electric Motor

EV Electric Vehicle

FEV Fuel-cell Electric Vehicle

G2V Grid-to-Vehicle

HEV Hybrid Electric Vehicle

ICE Internal Combustion Engine

ICEV Internal Combustion Engine Vehicle

ISM Integral Sliding Mode

MDP Markov Decision Processes

MPC Model Predictive Control

NCS Networked Control System

NN Neural Network

PEV Plug-in Electric Vehicle

PHEV Plug-in Hybrid Electric Vehicle

REV Range-extended Electric Vehicle

RHC Receding Horizon Control

RL Reinforcement Learning

SoC State of Charge

SoH State of Health

V2G Vehicle-to-Grid

Chapter 1

Introduction

Electric mobility is at the forefront of the energy transition in the transportation sector and is expanding at a rapid pace. Maintaining the same market value as in 2019, in 2020, the global Electric Vehicle (EV) fleet exceeded the 10-million mark, with almost 3-million new-vehicle sales [3, 4]. Despite this, the EV fleet only accounted for only 2.6% of global new-vehicle sales and about 1% of worldwide vehicle stock during 2019 [5]. Nevertheless, there are clear signs that the pace of adoption will continue to grow, with forecasts indicating that 57% of new-vehicle sales will be EVs by 2040 [6].

The advantages of EVs for global promotion include their benefits in local and global environmental protection, economic growth and energy security. EVs have zero tailpipe emissions and high energy efficiency, which are helpful for local air pollution reduction, especially in urban areas. To reduce the negative impacts of the transportation systems on global climate change to a greater extent, the electrification of the transportation systems and the application of new energy sources can be carried out simultaneously. The development of the EVs industry can stimulate economic growth, such as the EVs production supply chain, the charging infrastructure, and the business of operating the EVs. Since Battery Electric Vehicles (BEVs) and Hybrid Electric Vehicles (HEVs) utilize electric energy or environmentally friendly new energy sources, gasoline consumption can be reduced, thereby improving energy security.

According to the types of energy sources, energy carriers and the propulsion devices, the classification of vehicles can be depicted in Figure 1.1, including Internal Combustion Engine Vehicles (ICEVs), HEVs, BEVs, Fuel-cell Electric Vehicles (FEVs), Range-extended Electric Vehicles (REVs), and Plug-in Hybrid Electric Vehicles (PHEVs).

The impact of EVs on environmental protection, economic growth, energy security, or the power grids' operation can be identified and studied with mathematical models. A

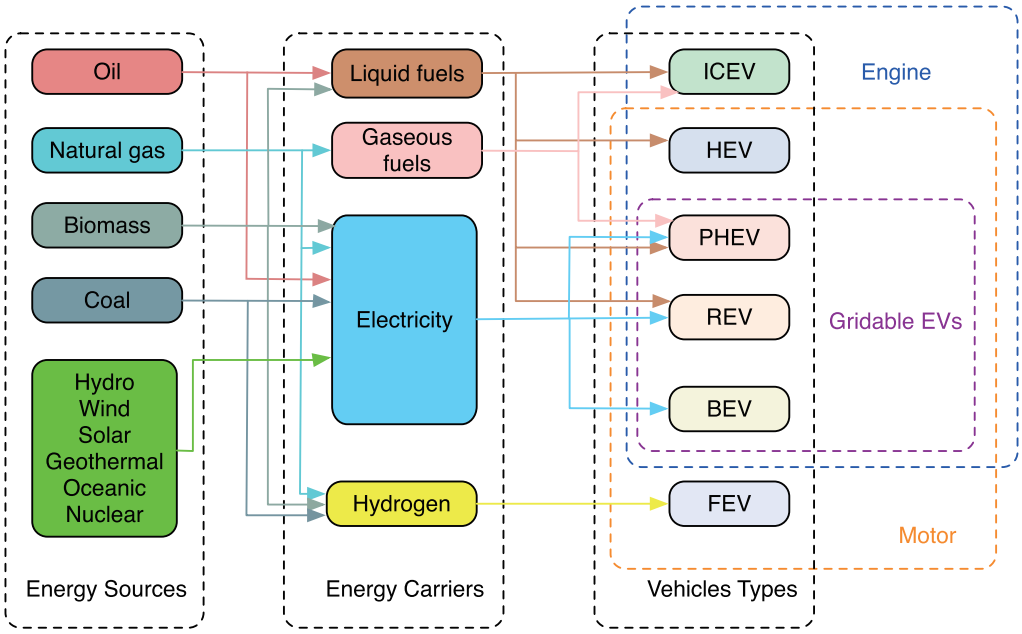


Figure 1.1: Classification of vehicles based on energy sources and carriers.

few existing results descriptively analyze the real-world EVs using the data collected from projects and experiments. These descriptive analyses are helpful for qualitative characterization of the use and charging behavior of EVs, but they are essentially a description of the status quo. The existing analysis is restricted in the context of the rapid development of transportation electrification. It requires tools that have sufficient flexibility to provide a basis for making decisions in rapidly changing circumstances.

In addition to the benefits mentioned above, the large-scale deployment of EVs has brought both challenges and opportunities to the operation of the power grids. On the one hand, the safety of the power grids will be disrupted by unmanaged EVs loads, especially at the distribution level where the limitation of the power grids' capacity is reached. On the other hand, the flexibility of charging requirements is often achieved by implementing smart charging strategies that can minimize the cost of upgrading the power grids' capacity. Besides, suppose that EV batteries connected to the power grids on a large scale are wisely managed. In that case, they will not enlarge the loads of power grids, and they will relieve the loads by utilizing them as energy storage devices.

To reduce CO_2 emissions significantly through the large-scale deployment of EVs, one of the crucial methods is to use new energy sources to charge the batteries of vehicles. Considering the intermittent nature of the new energy sources, it is necessary to develop

mechanisms as a part of smart grids to ensure the smooth integration of new energy sources into the energy systems [7, 8]. EVs can help by storing energy when there is a surplus and feeding it back to the power grids when there is a demand [9].

Indeed, EVs have the ability to store energy when used for transportation. On the one hand, vehicles are only driven for a small percentage of the day (around 1 hour - 1.5 hours in the U.S.). A large portion of the vehicles stays unused in parking lots (90% in the U.S.) [10]. Because EVs are equipped with large battery packs, they can be used as storage devices when parked, i.e., as a part of Vehicle-to-Grid (V2G) schemes [9, 11]. Thus, EVs dramatically increase the storage capacity of the power grids. Ref. [11] shows that replacing a quarter of the vehicles as EVs in the U.S. would double the current storage capacity of the power grids. On the other hand, considering that EVs are charged on demand, many EVs are charged every day. Consequently, the charging power loads of EVs may overload the power grids. It is highly desirable to design an EV charging management strategy (Grid-to-Vehicle (G2V)), taking into account the constraints of the distribution systems of power grids where EVs need to be charged. In addition, the EV route selection system is expected to maximize the generated energy when braking or driving downhill. In this way, EVs may reduce the frequency of charging, thereby expanding their cruising range, reducing the cost of their owners, and minimizing the peak value of EVs on the power grids.

In this chapter, a literature review on PHEVs and Plug-in Electric Vehicles (PEVs) is presented. Firstly, the modeling and power management of PHEVs are introduced. Secondly, the model and power management of PHEVs are first introduced. Then, the existing methodologies and research results on charging control of PEVs are reviewed. In addition, two problems: Power management method for PHEV and charging control for PEV are discussed in this dissertation. Finally, the motivation and research objectives of the dissertation are stated.

1.1 Literature Review on PHEVs and PEVs

This section elaborates the literature review on PHEVs and PEVs, including the background and the state-of-the-art techniques. The two essential problems of PHEVs and PEVs, i.e. power management and charging control, are analyzed and reviewed in detail.

1.1.1 Models of PHEVs

A. Powertrain Configurations

The objective of designing a controller for HEVs is to reduce the energy consumption based on power management. According to the powertrain configuration, HEVs can be separated into three categories: Series HEVs, Parallel HEVs and Combination HEVs [12]. The main difference among the configurations is whether there is a mechanical connection between the Internal Combustion Engine (ICE) and the drive train part.

- **Series hybrid electric vehicles**

  A typical series HEV has seven parts, including the ICE, generator, battery, rectifier, capacitors, converters and Electric Motor (EM). Figure 1.2 shows the power flow of the series HEV. The series HEV is characterized by the absence of a mechanical connection between the ICE and drive train part. While the vehicle is in motion, the ICE has two modes: 1) On-mode, i.e., the HEV is powered by fuel; 2) off-mode, i.e., the HEV is powered by the battery. In addition, the capacitor is an energy storage device that is located between the generator and the converter, being used to extend the cycle life and improve the battery efficiency [12]. When the output power of the generator is higher than the power requested by EM, the remaining power is used to charge capacitors and the battery. When the output power of the generator is less than the power requested by EM, the required energy is supplied from capacitors first.

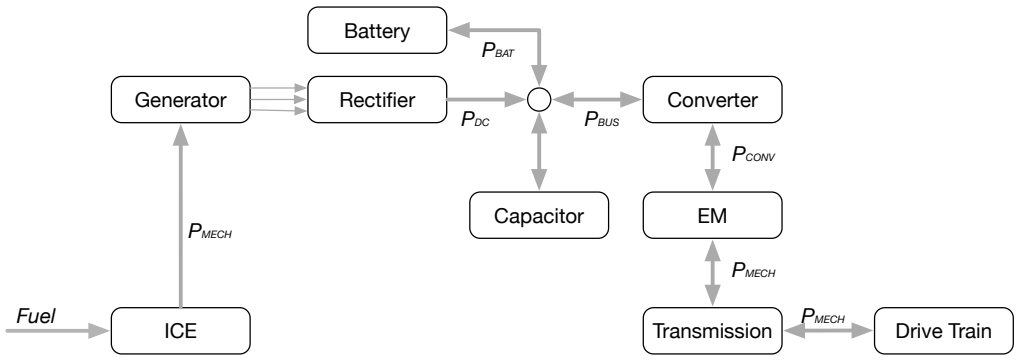


Figure 1.2: The power flow of a series HEV powertrain.

- **Parallel hybrid electric vehicles**

  A parallel HEV generally only has four parts, including the ICE, battery, converters and EM. As shown in Figure 1.3, the outputs of mechanical and electrical power are

connected with a parallel structure to drive the vehicle transmission. The distinct feature of the parallel HEV's control strategy is that the ICE is always turned on and operated at the maximum efficiency point. When the output power of ICE is less than the power requested from the vehicle transmission, the EM is turned on. Then, EM and ICE supply power to the transmission at the same time. When the output power of ICE is higher than the power requested by the vehicle transmission, the battery can be charged by the remaining power.

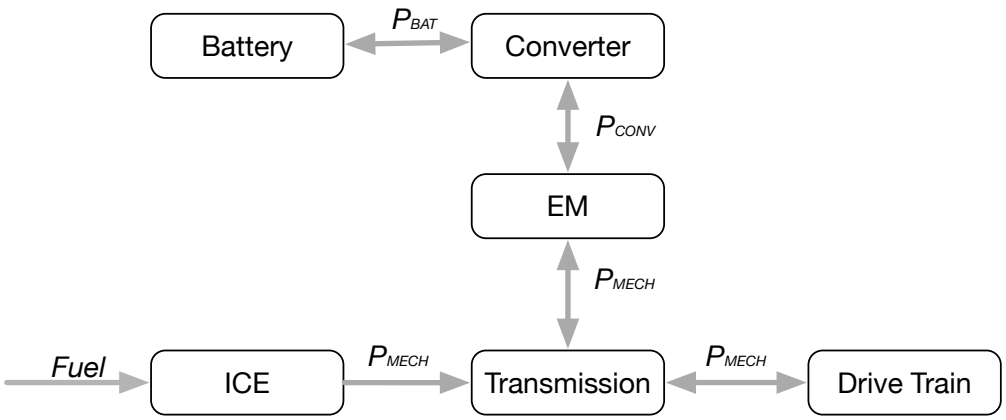


Figure 1.3: The power flow of a parallel HEV powertrain.

- **Parallel-series hybrid electric vehicles**

  As shown in Figure 1.4, the essential architecture is a combination of parallel and series hybrid configurations into a single package [13]. The parallel-series HEV has properties from configurations of both series HEVs and parallel HEVs. The complexity of the series HEV and parallel HEV in terms of the configuration is the main problem in practice.

B. Models of PHEVs

Generally speaking, the modeling of a PHEV can be characterized by five sub-systems [1]: (1) Vehicle dynamics, (2) Engine fuel consumption model, (3) Electric consumption model of the EM, (4) Transmission model, and (5) Battery model.

- **Vehicle Dynamics**

  The PHEV dynamics mainly refers to PHEV longitudinal dynamics. The output

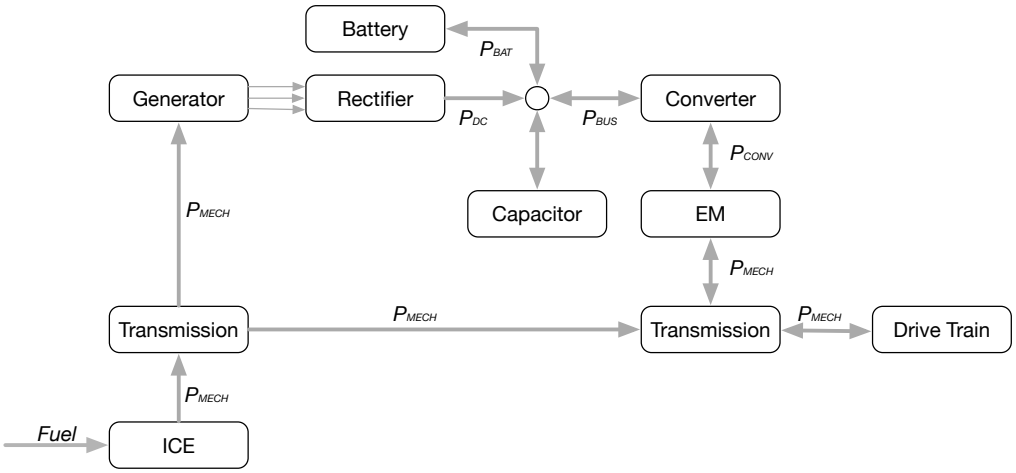


Figure 1.4: The combination of parallel and series hybrid configurations.

torque T_w can be obtained by the following equation:

$$\begin{aligned} T_w &= \eta_t \cdot i_{AMT} \cdot i_f (T_e + T_m) + T_B \\ &= [Mgf_r \cos\theta + \frac{1}{2} C_D \rho_d A_f v^2 + Mg \sin\theta + \sigma M \frac{dv}{dt}] \cdot r_w, \end{aligned} \tag{1.1}$$

where η_t represents the transmission efficiency. T_e, T_m and T_B are the torque of engine, the torque of EM, and the torque of braking, respectively. i_{AMT} and i_f denote the gear ratio of the automatic mechanical transmission (AMT) and the differential gear ratio, respectively. The parameters, f_r, g and θ, describe the operation condition of the vehicle, and denote the rolling resistance coefficient, gravity acceleration, road slope angle, respectively. $M = M_o + M_p$ denotes the equivalent vehicle mass including the vehicle mass M_o and the passenger load M_p. The parameters of the vehicle, r_w, C_D, ρ_d and A_f, represent the wheel radius, air drag coefficient, air density and frontal area, respectively.

To calculate vehicle speed, according to the kinematic constraints of series-parallel hybrid powertrain, the wheel speed ω_w can be calculated by the engine speed ω_e and EM speed ω_m as follows:

$$\omega_w = \frac{\omega_e}{i_{AMT} \cdot i_f} = \frac{\omega_m}{i_{AMT} \cdot i_f}. \tag{1.2}$$

- **Fuel Consumption Model of Engine**

Fuel economy is the most important assessment index of the power management

strategy for PHEVs since the engine is the main energy source in PHEV. For fossil fuel engines, their fuel consumption rate is a nonlinear function of the speed and power of the engine, which can be obtained by searching the consumption contours of the engine. Nevertheless, the fuel consumption contours of the engine depend on the type of fuel and the type of engine. As such, we model the fuel consumption based on a 1.5 litres gasoline engine, shown in Figure 1.5. The engine is modeled by the advanced vehicle simulator ADVISOR [14]. Indeed, some other engine fuel consumption models can also be applied.

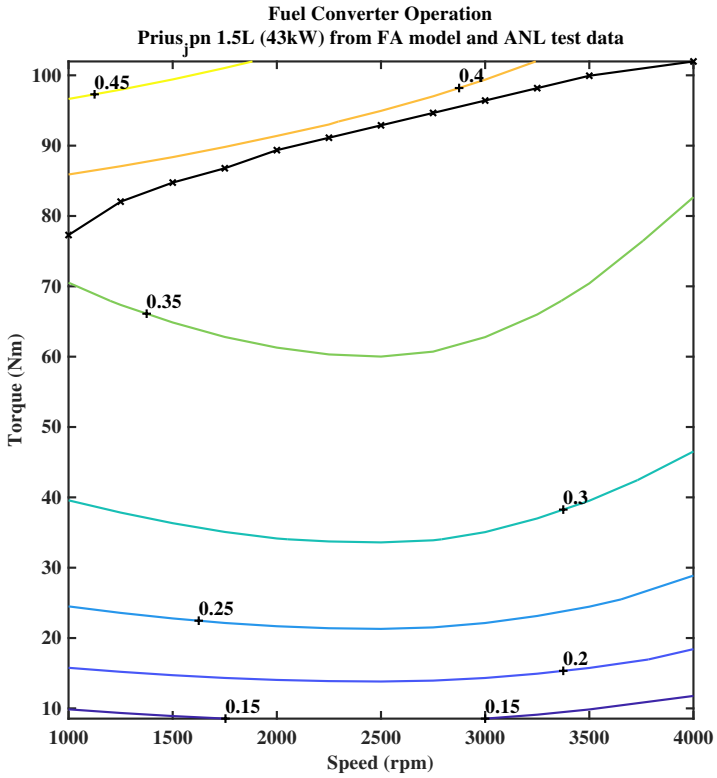


Figure 1.5: Fuel efficiency map of ICE.

- **Electric Consumption Model of EM**

  EM works under two operating modes, the driving mode and braking mode. The output power of EM can be described as follows:

$$P_m = \begin{cases} T_m \omega_m / \eta_m, & \text{Driving} \\ T_m \omega_m \eta_m, & \text{Braking} \end{cases} \tag{1.3}$$

The above model shows that the EM output power P_m depends on the EM torque T_m, the EM speed ω_m and the EM efficiency η_m. The EM output power P_m is a nonlinear function of the EM speed ω_m and the EM torque T_m. η_m can be obtained by searching the EM efficiency map that depends on the type of EM. As an example, Figure 1.6 shows the EM efficiency of a 30 kW permanent magnet motor, modeled by the advanced vehicle simulator ADVISOR [14].

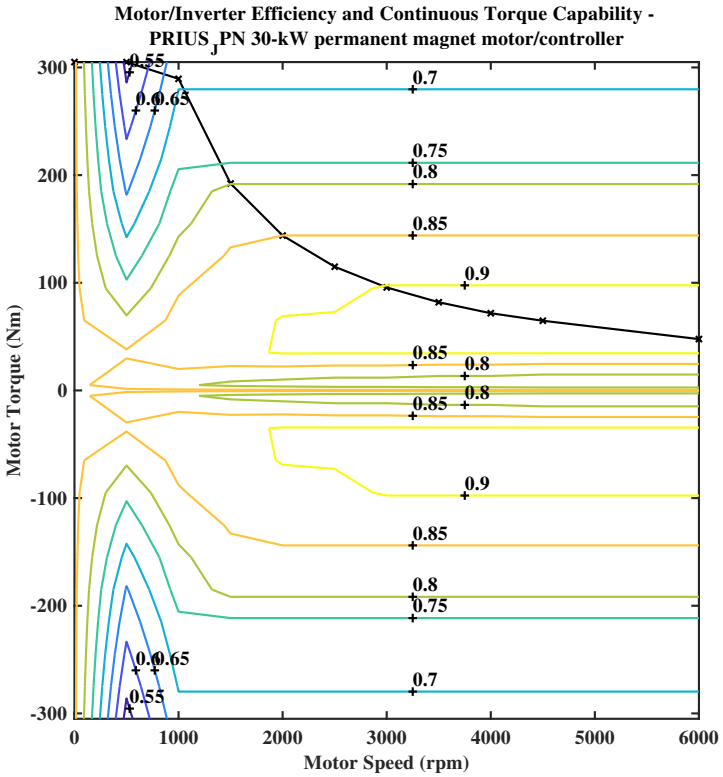


Figure 1.6: Efficiency map of EM.

- **Transmission Model**

  AMT can improve both the adaptability of driving conditions and the fuel economy for PHEVs through the dynamic coupling with engine and EM. Hence, its automatic gear-shifting model is needed for the power management of PHEVs. The detailed description of the gear-shifting strategy can be concluded by a realizable gear-shifting logic for better adaptability to the changes in the driving cycle and the road conditions [15]. The gear-shifting strategy includes the dynamic gear-shifting strategy and the economic gear-shifting strategy. Figure 1.7 shows the difference and relationship between the dynamic and economic gear-shifting schedules. The curve with circles

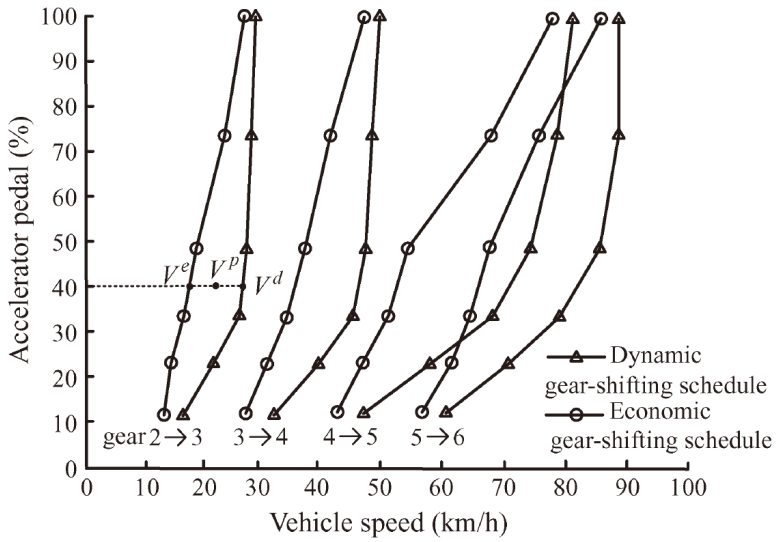


Figure 1.7: AMT gear-shifting map in two different strategies.

represents the economic gear-shifting strategy, which indicates upshifting to enlarge the fuel economy. The curve with triangles represents the dynamic gear-shifting strategy that keeps the low gear to enlarge the usage rate of the driving torque [16]. According to the accelerator pedal shown in Figure 1.7, the balanced gear-shifting point V^p can be calculated by the point V^d on the dynamic gear-shifting curve and the point V^e on the economic gear-shifting curve. Thus, the vehicle speed at the balanced point V^p can be calculated by the vehicle speed at the dynamic gear-shifting point V^d and the economic gear-shifting point V^e according to the distribution factor ϵ, as shown in the following:

$$V^p = \epsilon V^d + (1-\epsilon)V^e \tag{1.4}$$

The relationship between two torques before and after gear-shifting is $T_\omega^n = T_\omega^{n'}$, where the superscripts n and n' mean gear n and gear n', respectively. Hence, *in the dynamic gear-shifting strategy*, the speed V^d can be calculated by (1.2) and (1.5):

$$\begin{cases} T_\omega^n = \eta_t \cdot T_e(\omega_e^n) \cdot i_{AMT}^n \cdot i_f \\ T_\omega^{n'} = \eta_t \cdot T_e(\omega_e^{n'}) \cdot i_{AMT}^{n'} \cdot i_f, \end{cases} \tag{1.5}$$

where ω_e^n and $\omega_e^{n'}$ represent the rotation speed at gear n and n', respectively. $T_e(\omega_e^n)$ and $T_e(\omega_e^{n'})$ represent the engine torque working on n and n' gear modes with the

speed ω_e^n and $\omega_e^{n'}$, respectively, which can be obtained by searching the outer characteristic curves of the engine. i_{AMT}^n and $i_{AMT}^{n'}$ are the gear ratio of the AMT at gear n and n', respectively.

Similarly, *in the economy gear-shifting strategy*, the wheel torques before and after gear-shifting and the least fuel consumption torque of the engine with the current speed are equal, which can be obtained by searching the universal characteristic curves of the engine. The speed V^e can be calculated by (1.2) and (1.5) as well.

As a result, the gear-shifting strategy of six-speed AMT can be described as follows:

$$gb(k+1) = \begin{cases} 1, & gb(k)+s(k) < 1 \\ 6, & gb(k)+s(k) > 6 \\ gb(k)+s(k), & \text{otherwise} \end{cases} \tag{1.6}$$

where $gb(k)$ represents the gear number. $s(k)$ is the gear-shifting signal, whose value can be chosen from -1, 0 and $+1$. The three numbers represent the case of downshifting, sustain and upshifting, respectively [17].

- **Battery Model**

  State of Charge (SoC) is the main index to describe the battery capacity. Since battery SoC is always influenced by the temperature, it is difficult to perfectly model the transient behaviour of battery SoC. In this dissertation, we adopt a simplified battery SoC model.

  According to Kirchhoff's voltage law, the voltage of the battery in the equivalent circuit equation can be described as follows:

$$U_{batt} = U_{oc} - RI_{batt}, \tag{1.7}$$

  where U_{oc}, R, U_{batt} and I_{batt} are the open-circuit voltage, the internal resistance, the terminal voltage, and the internal current of battery, respectively.

  Then, the charge rate of the SoC can be calculated as follows:

$$SoC(k+1) = SoC(k) - \frac{I_{batt}\Delta t}{Q_{max}}, \tag{1.8}$$

  where Δt denotes the sampling time between the sampling time instant k and $k+1$, Q_{max} is the maximum value of battery capacity. The power of battery P_{batt} can be

described as follows:

$$P_{batt} = U_{oc} I_{batt} - I_{batt}^2 R_{int}. \tag{1.9}$$

Solving (1.8) and (1.9) for the rate of change of SoC, we have

$$SoC(k+1) = SoC(k) - \frac{U_{oc} - \sqrt{U_{oc}^2 - 4P_{batt}R_{int}}}{2Q_{max}R_{int}} \tag{1.10}$$

Note that the solutions of (1.9) for I_{batt} may not be unique. However, only the one with a larger value is admissible for negative power demands and capable of maximizing the efficiency for nonnegative power demands. Note that negative power demands only occur when SoC approaches the minimum values, and the battery has to recover with the help of the fuel engine.

Due to the difference between two operation modes of battery (i.e., charging and discharging), the internal resistance R_{int} can be described as:

$$R_{int} = \begin{cases} R_{dischg}, & \text{discharging} \\ R_{chg}, & \text{charging} \end{cases} \tag{1.11}$$

where R_{dischg} and R_{chg} represent the internal resistances for the battery being on discharging or charging mode, respectively.

1.1.2 Charging Control Algorithms of PEVs

Increased societal awareness of environmental issues associated with vehicular emissions has spurred the development of straightforward solutions for transportation. In this respect, electrified vehicles are emerging as the defining trend of transportation. EVs are more environmentally friendly, compared with the vehicles using fossil gas. Recently, a dramatic increase in the adoption of EVs has additionally been attributed to decreasing battery costs and cheaper electricity prices compared to escalating fuel prices. Despite the benefits to the environment, the growing electricity demand, due to the adoption of EVs, shows the negative impacts on the power grids, reliability and economical operation of power grids, such as voltage deviations, transformers and lines saturations, increase of electrical losses, etc. [18, 19]. For example, charging a single EV will potentially double the energy consumption of an average household [20]. In cases where millions of EVs simultaneously charge across the power grids, new peak load events will arise.

If the vehicle is in motion without a flexible conventional load, uncontrolled charging may cause different problems at network levels, such as the larger peak loads of power grids. In addition to the aforementioned literature, charging coordination strategies have been studied in a variety of aspects including control architecture, subjects and technologies [21]. Some main categories and corresponding representative papers are shown in Table 1.1.

Table 1.1: Classification and representative papers for charging control problems.

	Technology	Decentralized	Hierarchical	Centralized
For EV Users:	Game Theory	[22–25]	[26, 27]	
	ADMM	[28]	[29]	
	Optimization	[19, 30]	[31–33]	[34, 35]
	Programming	[36]	[32]	[37]
	Lagrangian Decomposition	[38]	[39]	
For Aggregator:	Game Theory	[40–44]	[27]	
	ADMM	[45]		
	Optimization	[19, 46–49]	[32]	
	Programming	[50]	[32, 51]	[52–54]
	Lagrangian Decomposition	[38]	[39, 55]	

Centralized Charging Control Architecture

Figure 1.8 shows the main architecture of centralized charging control, in which each EV directly connects to a central aggregator. The central aggregator needs to collect all the charging information of EVs, solve an optimization problem to schedule the charging process of each EV, and send the optimized control command back to all EVs. Obviously, the advantage of the centralized control schemes is that the central aggregator will generate the optimal control command based on the complete information of the entire system with all EVs. Briefly, the control methods can be classified into two perspectives: EV users and aggregators.

From EV users' perspectives, the increasing number of PHEVs can have a significant impact on the electric utility if charging control techniques are not properly designed. The developed system in [37] can dramatically mitigate the impact of PHEVs on the utility power grids and reduce the charging cost. Due to the increasing number of PHEVs, it

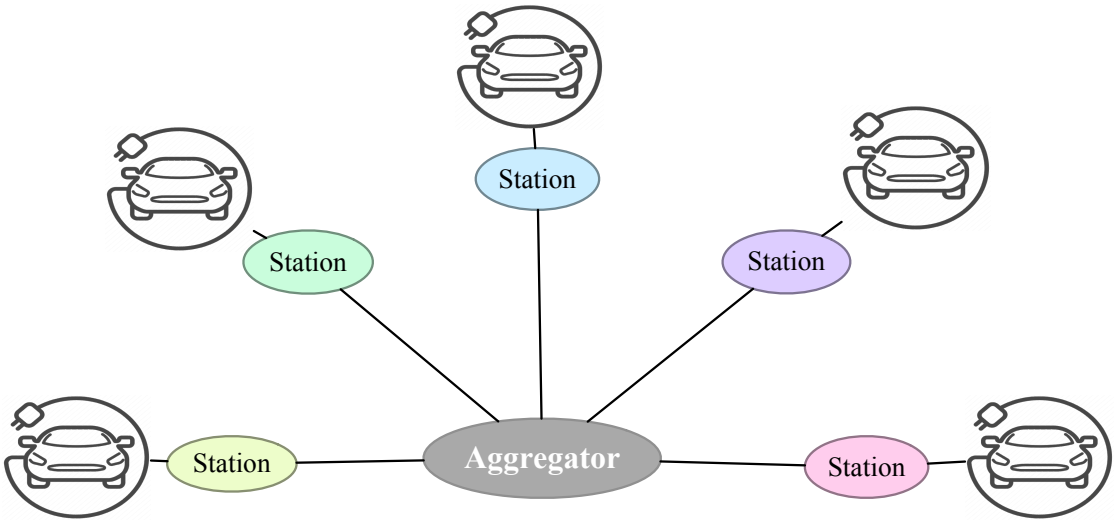


Figure 1.8: Centralized charging control architecture.

is critical to develop efficient charging coordination mechanisms that reduce the cost and impact of vehicle integration to power grids. In [35], Tang *et al.* provide a Model Predictive Control (MPC)-based algorithm for the optimal vehicle charging schedule, where the non-causal information about future vehicle arrivals is unknown in advance. However, its statistical information can be estimated based on the data-driven algorithm. In [52, 53], a two-step modeling framework is developed to merge and clean the useful EV charging data. Zhang *et al.* propose a data-driven model that captures the non-homogeneity and periodicity of the residential EVs charging behaviour [54].

Another class of charging control methods is designed for aggregators. In most of the proposed approaches, the aggregator in the closed-loop system can control the EV charging in a direct or indirect manner and serve as an interface with other entities, such as the transmission system operator or energy service providers [56]. To minimize the energy consumption based on EV drivers' preferences and technical limitations, Di Giorgio *et al.* present an event-driven MPC framework for managing charging operations of EVs in power grids by seeking the tracking of a reference load profile defined by the grid cooperator [34]. The proposed control approach allows random EV users to participate in demand-side management programs, which will play a crucial role in improving the stability and efficiency of smart charging control methods for power grids.

Since the charging rates of all EVs are decided by the central aggregator, it may not be suitable for each EV. What's more, if any mistake or failure happens at the central aggregator, there is a high possibility of failure in the entire system due to the centralized

structure. The centralized schemes also suffer from the high computational complexity of the optimal control problem and heavy communication loads. For instance, the large-scale population of vehicles and the growing number of control variables and constraints increase the size of the optimization problem. Besides, the redundant communication loads, bandwidth limitations, and expansive supporting infrastructure obstruct the broad adoption of centralized charging control approaches.

Decentralized Charging Control Architecture

To design the EV charging control algorithm, most existing results consider a centralized approach, which can not handle the impact of the increasing number of EVs. A distinct feature of the decentralized charging control method is that each EV will compute its own control signal locally and share the signal with other EVs or an aggregator to achieve the control objectives. Figure 1.9 depicts two different architectures of decentralized charging control, with and without the aggregator. Compared with centralized architecture, the aggregator plays a different role in a decentralized structure. In the decentralized structure, the aggregator is unnecessary and does not need to solve the optimal control problem. The decentralized architecture, illustrated in Figure 1.9 (a), reduces the communication loads by utilizing an indirect aggregator, which does not need to receive the complete information of the entire system (shown using dash lines). The control inputs will be broadcasted by the indirect aggregator to all the EVs. Another kind of decentralized architecture without using the indirect aggregator is shown in Figure 1.9 (b). The advantage of the decentralized architecture is that each EV solves the optimal control problem locally, which reduces the possibility of the whole system failures.

Decentralized charging control schemes without the aggregator are more efficient for EV users [40, 41]. Most of the existing methods for EV charging assume that the PEVs have the same power demands and battery conditions for charging. To avoid these strong assumptions, the work in [30] presents a decentralized EV charging approach to reduce customers' costs while increasing the energy efficiency of power grids. Another challenge from the charging problem is to minimize the cost of electric consumption. Economic MPC is very well suited for controlling smart energy systems since the electricity price and demand forecasts are easily integrated into the MPC optimization problem. In [57], the economic MPC minimizes the cost of electricity consumption for a single EV. To reduce the computation capability, in [58], the authors propose a decentralized algorithm to optimally schedule EV charging, which only requires each EV to solve its optimal control

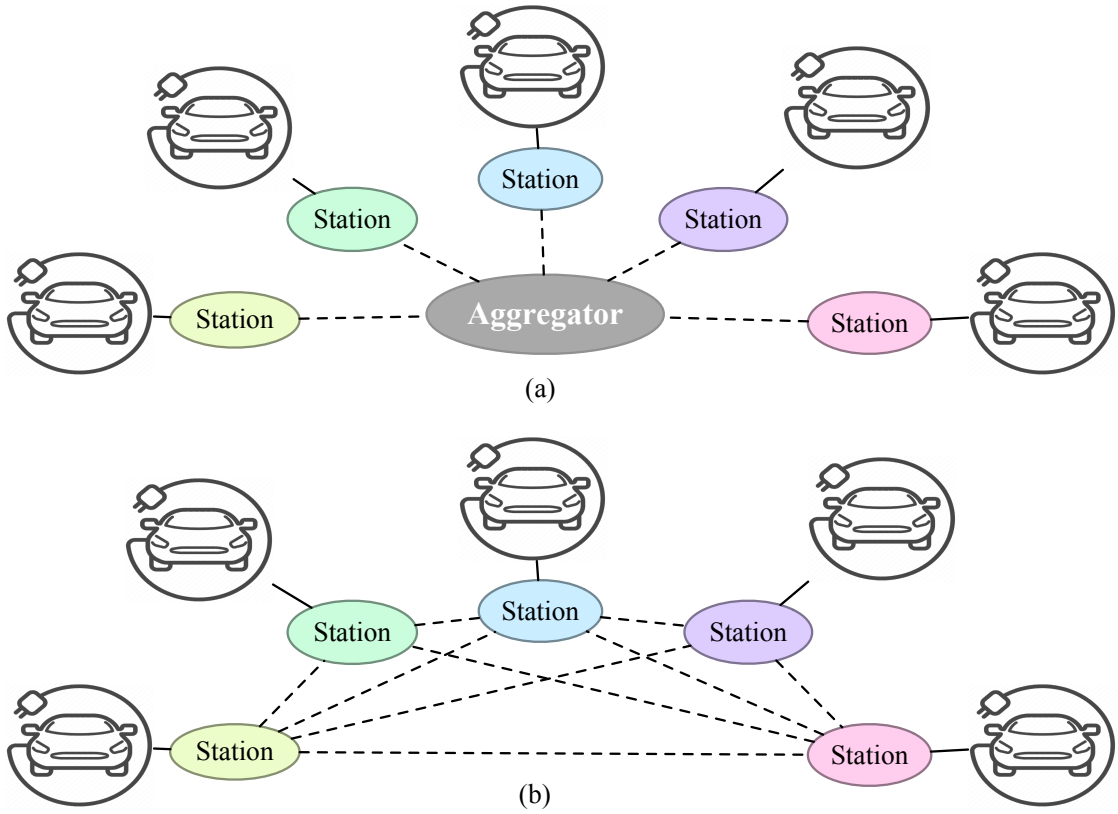


Figure 1.9: (a) Decentralized charging control architecture with an aggregator; (b) Decentralized charging control architecture without an aggregator.The dash lines represent the connection among stations.

problem locally.

The research on decentralized EV charging control methods for the aggregator shows the advanced use of communication technology and metering infrastructure that can increase the flexibility of power loads in power grids [59, 60]. Ref. [61] develops a decentralized event-driven EV charging control scheme for handling power demand variations caused by EVs' random arrivals and departures. To maximize the efficiency of EV charging, the authors in [49] develop a decentralized energy management system for regulating the energy flow among the photovoltaic system based on the MPC algorithm. The decentralized optimal controller can be embedded into local chargers [46, 47]. Due to the negative impacts on the distribution system of power grids caused by uncontrolled or improperly controlled EV charging, the control scheme for satisfying heterogeneous individual charging requirements and the distribution system constraints should be investigated [62]. The drawback of battery charging in traditional electric vehicles is the use of plug-in charg-

ing devices. To simultaneously transfer the information, including the battery status or emergency messages between the power grids and vehicles, the work in [63] proposes a decentralized wireless battery charging method.

Although the decentralized structure, shown in Figure 1.9 (b), enhances the autonomy of each EV, it generates a significant communication overhead, especially with the large-scale population of EVs.

Hierarchical Charging Control Architecture

Due to the aforementioned pros and cons of the centralized and decentralized charging control architectures, there is an increasing interest in the hierarchical architecture to reduce computational and communicational overheads. Furthermore, the algorithms based on the hierarchical charging control architecture retain the complete information of all EVs. Figure 1.10 depicts two typical structures of hierarchical charging control, with and without the aggregator, respectively. The architecture, illustrated in Figure 1.10 (a), shows EVs connect to the sub-aggregators under the decentralized structure, where all the sub-aggregators link to a central aggregator. Compared with a fully centralized structure, it reduces the communication loads by using sub-aggregators. The sub-aggregators gather certain information from EVs and transfer the scheduling information with the central aggregator until the equilibrium is obtained. Figure 1.10 (b) depicts another case of hierarchical charging control approaches, where the sub-aggregators act as a central aggregator for a group of EVs while sharing their scheduling information with each other under decentralized architecture. The advantages of hierarchical charging control schemes are three-fold: 1) reducing the communication and computation overheads; 2) decreasing the possibility of entire system failures; 3) enhancing the automation of each EV.

To meet plug-and-play charging requests of EV users in the distribution system of power grids, [33] presents a hierarchical control scheme based on MPC for tracking periodic references. To avoid the peak load of power grids, a price-based hierarchical scheme, using the architecture shown in Figure 1.10 (a), is proposed in [32] to minimize the EV charging cost with sub-aggregator supply constraints. A penalty price cost is surcharged for the load excursions.

Refs [26, 64] study the control schemes to minimize the energy cost by utilizing a set of sub-aggregators under another hierarchical architecture illustrated in Figure 1.10 (b). A non-cooperative game is designed in [26] to calculate the optimal charge profiles using the best response strategies for each sub-game among sub-aggregators. A mobility-aware

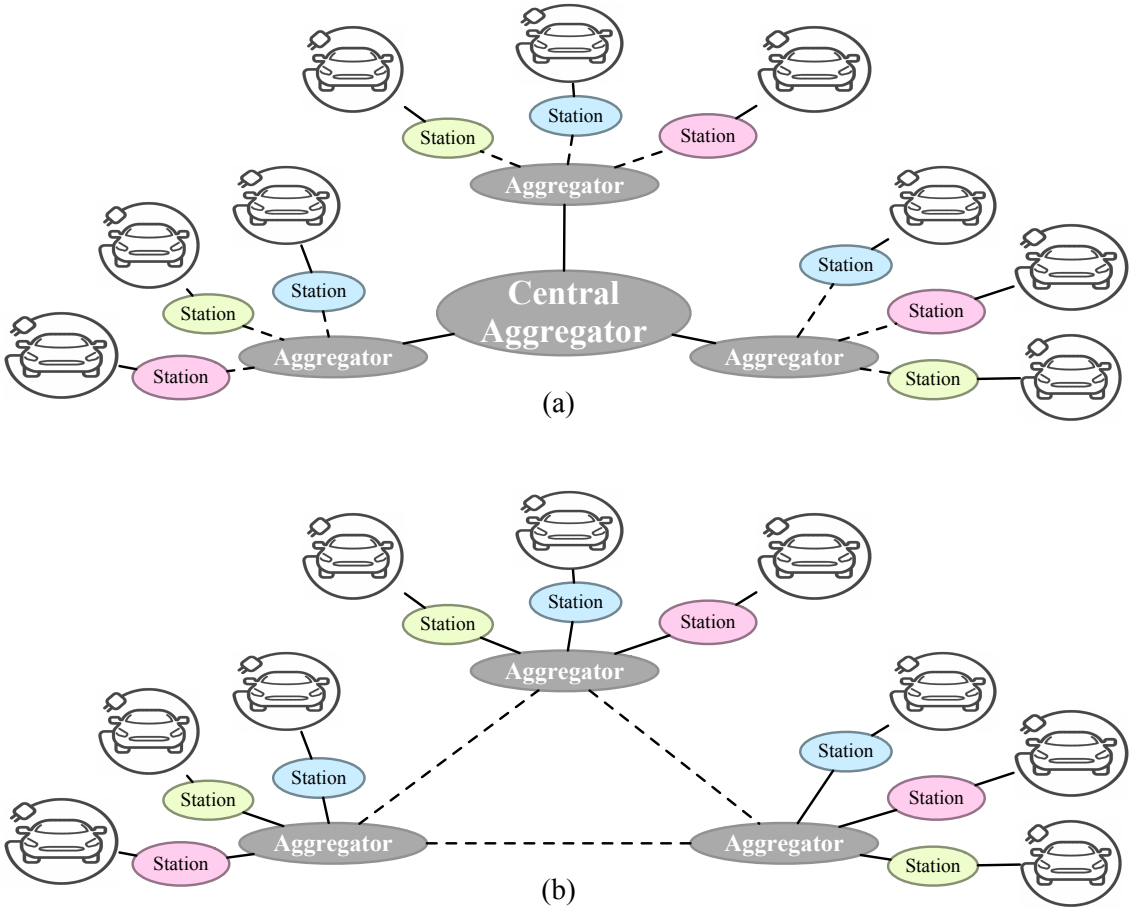


Figure 1.10: (a) Hierarchical charging control architecture with central aggregators: Centralized coordination of aggregators and decentralized charging scheduling of EVs; (b) Hierarchical charging control architecture without the central aggregator: Decentralized control architecture among aggregators, while centralized control structure among vehicles.

framework is proposed in [64] allowing the EVs to charge the battery at any location, based on the route information, average speed, charging conditions, and location of the charging stations. Since each sub-aggregator controls a set of charging stations to achieve collaboration among them, the proposed framework can be seen as a dynamic version of the hierarchical architecture in Figure 1.10 (b) with the EVs connected to multiple sub-aggregators.

From aggregators' perspectives, one of the objectives of designing a control scheme under a hierarchical architecture is to achieve a balance between minimizing the associated energy cost of EV users and maximizing the revenue of power grids' operators. Refs. [51,

65] design a control scheme to minimize the electricity supply cost under a hierarchical architecture in Figure 1.10 (a). The designed algorithm determines the charging schedule through three steps comprised of a set of sub-aggregators which manage a set of EVs. Besides, a hierarchical scheme based on the Stackelberg game is presented in [27] to make the decision depending on the electricity price that optimizes the EV charging cost and the revenue of selling energy under the architecture shown in Figure 1.10 (a). A trade-off between the benefit from EV charging and associated cost is captured, based on a given amount of energy requested by a set of EVs, to maximize the revenue of the power grids' operator. According to the proposed algorithm, the electricity price finally converges to an optimal Stackelberg equilibrium which is the optimal electricity price. Ref. [66] utilizes a multi-objective optimization-based method, while [48, 67] design a two-level control scheme for sub-aggregators (upper-level) and for each EV (lower-level), respectively. The proposed algorithm in [55] includes transformer capacity constraints in the cost function of the optimization problem by applying Lagrangian relaxation.

In conclusion, compared with traditional fossil fuel vehicles, EVs have potential benefits to the economy and environment due to their energy-efficient, cost-efficient and less greenhouse gas emissions [68]. The main challenge of the charging control for EVs is how to mitigate the negative impacts on the power grids, especially the uncontrolled recharging overlapping with other power loads, which could cause the security issues for power grids. This subsection surveys the charging control algorithms in the literature, which are categorized by three main groups based on the control architecture, with a focus on EV users and aggregator perspectives. From the EVs user's point of view, a proper recharging-and-discharging protocol shows more energy costs reductions and the opportunities of generating profits. From the aggregator or power grids point of view, the coordination of the EV charging helps to alleviate the negative impacts and enhance the power grids' capacity.

1.2 Literature Review on Energy Management Problems for EVs

This section is devoted to the literature review of the advanced control methods, e.g., Reinforcement Learning (RL) and MPC supplemented by the iterative learning method, and their applications to the power management of PHEVs and the charging control of PEVs.

1.2.1 Reinforcement Learning for HEV power management

RL provides a mathematical framework based on the optimal control problem and dynamic programming [69]. In an RL problem, there are typically four elements, including the controller, the environment, actions and rewards. In analogy to conventional control systems, the system state space is usually the observation space of the environment. The control objective of RL is to maximize the reward value by mapping the situations to actions. The typical decision-making method of RL is based on Markov Decision Processes (MDP).

The control policy, generated by RL, may not show which action to be taken, but it would discover which action generates the most rewards by trial and error. Another problem is that the actions may affect the immediate reward and the next state through all subsequent rewards.

Learning-based methods can be classified into two groups: Supervised learning and unsupervised learning. Supervised learning is a kind of learning algorithm that is learning from labeled training examples. Examples are used to identify different situations to specific categories. The main objective of supervised learning is that, the generated responses guarantee the actions that are correct in the situations. In supervised learning, the one that provides the labeled training examples information is called an external supervisor. Contrastingly, the main idea of unsupervised learning is to find the structure by using the collections of unlabeled data. Comparing the training process of supervised and unsupervised learning, RL can be treated as unsupervised learning since it does not depend on the labeled examples.

Learning-based power management of HEVs can be generally classified into two categories. One is offline supervised learning using Neural Networks (NNs) [70], the other is online RL [71–73]. Obviously, the learning-based power management for HEVs is easy to consider more complex scenarios since it does not require the accurate system model. In [74], the proposed approach predicts the power demand at the next step based on the Markov property. The Bellman optimality is derived to generate the control signals in order to minimize the expected energy cost and converge to the optimal value.

In addition, learning-based power management of HEVs can solve another type of urgent issues that the road types and traffic congestion greatly limit the reduction of energy consumptions [75]. There are two main categories of solutions using learning-based schemes. One is to improve the powertrain by training based on regular routes. The control algorithms are designed for a specific route with known traffic conditions. The control strategy for different driving behaviors can be customized after training utilizing learning-

based methods [76]. Typically, these control methods are usually used in hybrid electric buses [77, 78]. The other is to generate the control policy based on the predicted road types and traffic congestion levels. In [79, 80], Murphey *et al.* propose an NN control framework that can generate the prediction of roads types, traffic congestion levels and driving trends. Besides, the NN can be modified to improve the training process. Chen *et al.* propose an online controller for power management in [81], where the controller consists of two NNs considering the trip length and three different trip durations.

1.2.2 Model Predictive Control for EV Energy Problems

The feature of a networked control system is that the control and feedback signals are transmitted among sensors, controllers, and actuators in the form of information packages through a communication network. Along with advantages provided by Networked Control System (NCS), for instance, reducing the complexity and overall cost of the whole control systems, several challenges also lead to an increasing number of relevant research topics, such as fault-tolerant control strategies, time-delays, packet losses, resource constraints; see [82] for an example.

The idea of MPC or Receding Horizon Control (RHC) is came up in the 1960s [83]. The first result about the comprehensive exposition of generalized predictive control was presented in [84, 85]. The name of MPC comes from controlling an explicit model of the plant, using the prediction of the future output behavior. In MPC, the controller minimizes the tracking error, between the desired reference and the predicted output, over the future horizon, subject to state or input constraints. MPC has been proved to be a feasible control method for the plant with hard constraints on the input and state [86]. MPC solves optimal control problems by moving predictive windows to generate recurrent online solutions, which guarantees transient behaviors of closed-loop systems [86, 87]. The basic idea of MPC is depicted in Figure 1.11. In the MPC strategy, the controller solves an online finite horizon open-loop optimal control problem to obtain the optimal control sequence at each sampling time. The first control action is applied to the plant, where the initial state is the current state. Based on the information of next sampling time, the controller solves the optimal control problem again and repeats the procedure. At each sampling time, the desired feedback characteristics are enabled by the receding horizon mechanism when the plant's state and output are measured by sensors and sent to the controller.

In the 1990s, several works described and demonstrated the stability and the feasibility of the online optimization for linear system models [88, 89]. Many practical problems using

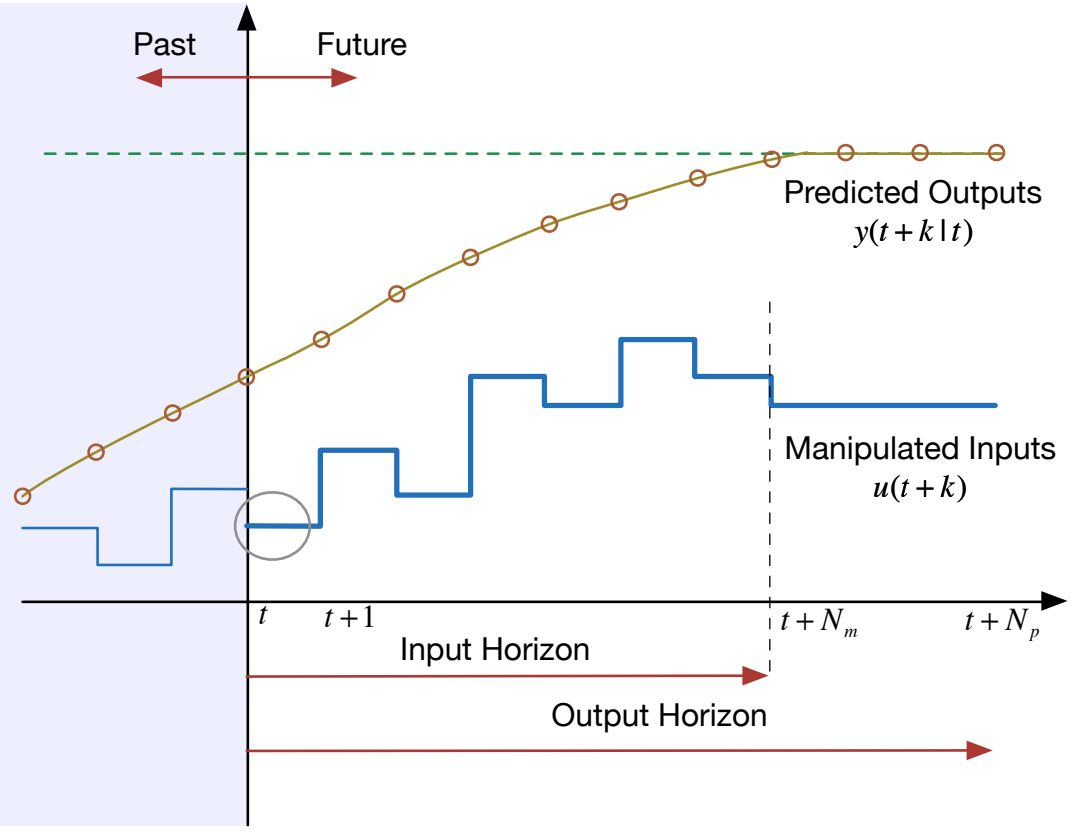


Figure 1.11: The receding horizon control strategy.

MPC strategies still necessitate being further investigated, such as the high complexity of models, the efficiency of the online optimization computation scheme and the reduction of communication loads for large-scale systems. Significantly, multiple goals need to be achieved at the same time with multiple dimensions to solve practical problems using MPC strategies. Also, it is always impossible to accurately obtain the mathematical models for the overall systems. We can take the charging control for PEVs as an example. The high-level objectives are to minimize energy usage and to meet individual requirements. Energy minimization further includes lower-level purposes, such as reducing the price cost, mitigating the negative impacts of battery capacity degradation.

Learning-based methods provide new opportunities to improve MPC schemes in industrial application scenarios [90]. Learning-based MPC schemes do not require an accurate model of the entire practical system. They approximate this knowledge during a training process. For example, when designing a charging control strategy, RL can be used to forecast the price or the power load [91, 92]. Also, iterative learning control is an effective

method for improving the transient response of uncertain dynamics that operate repetitively [93]. Typically, the iterative learning method is designed for repetitive operating systems or periodical processes [94].

The earliest work of iterative learning control combined with MPC was published in 1999 for chemical batch processes [95]. Iterative learning-based MPC has its abilities to reject disturbances and handle constraints [96, 97]. Based on the existing results, two sufficient conditions need to be guaranteed when the designed control algorithm utilizes the iterative learning-based MPC framework. One is the convergence to the desired equilibrium point, and the other is the constraint satisfaction at all iterations. Besides, MPC with iterative learning method can be utilized for energy problems. Comparing with traditional MPC method, MPC with iterative learning method is based on the stored measurement, input and error data of all previous trials. Despite that the periodicity of charging control problem is not as exact as in the control of industrial machine, such as autonomous racing cars, the charging condition of power grids change from day to day, which is an entire trial under iterative learning framework. The optimal control strategy for power management of PHEVs can be generated by iterative learning-based MPC, when the powertrain of PHEVs is treated as a repetitively operated dynamic system [98].

1.3 Motivations and Organization of Ph.D. Dissertation

In the previous sections, the powertrain configurations and models of PHEVs are briefly reviewed, followed by the development of charging control algorithms for PEVs. Though a few existing results for controlling and scheduling electric energy for EVs, many urgent issues in power management and charging control require to be investigated. In this dissertation, optimal energy management problems, including power management for PHEVs and charging control for PEVs, are discussed using advanced control and learning-based methods. A visual organization of this dissertation is shown in Figure 1.12.

Chapter 1 provides the literature review on power management for PHEVs and charging control for PEVs. In Section 1.1, the powertrain configurations and models of PHEVs are firstly reviewed. Then, the existing charging control algorithms for PEVs are summarized in the following subsection. The existing results on the problems considered in this dissertation are discussed in Section 1.2, including RL for power management and MPC supplemented by the iterative learning method. Finally, the motivation and organization of the Ph.D. dissertation are given in Section 1.3.

According to the daily utilization of PHEVs, a route leading to the minimum energy

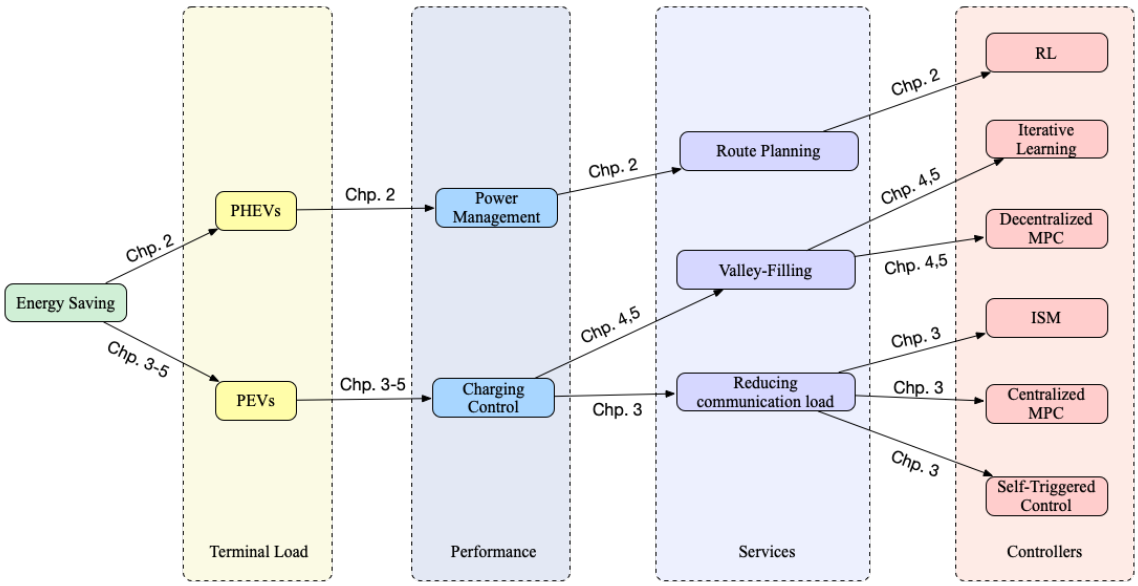


Figure 1.12: Organization of the dissertation.

consumption considered the power management is helpful for PHEVs' broad adoption. Route planning is usually separated from the PHEV power management in the existing literature. Most of the existing results related to route planning are to find routes through charging stations when the vehicle is in motion with low battery capacity. It would be more efficient for PHEV users to provide an option in the navigation system that leads to the lowest energy consumption that is the same as minimum time or distance used in the navigation system now. However, the current navigation system only provides the option of routes to the destination for minimum time or distances. An option of routes for lowest energy consumptions is also desired, especially for PHEVs. To address the problem, **Chapter 2** studies a control scheme of route planning with power management for PHEVs based on a two-loop structure using RL. Using Toyota Prius as an example and planning its route in a city with various navigation ranges, we demonstrate the effectiveness of the proposed control scheme by simulation studies.

Another limitation of impeding the PEVs' broad adoption is the long charging time and negative influence for power grids caused by PEV charging. Charging a large-scale population of PEVs at the same time yields high workloads of power grids, impacting efficiency and safety. The main difficulties using the centralized architecture for charging a large-scale population of PEVs come from the complexity of the optimal control problem and the high communication loads. In addition, there are redundant communication loads hindering the broad adoption of centralized charging control methods.

To tackle the difficulties, **Chapter 3** investigates a self-triggered MPC scheme integrated with Integral Sliding Mode (ISM) for networked uncertain nonlinear continuous-time system subject to state and input constraints. In the proposed scheme, both computation and communication loads can be reduced by aperiodically solving the optimization problem to generate the control signal and the next sampling instant. The matched uncertainty is suppressed by using the ISM method.

Next, **Chapter 4** proposes a decentralized charging control algorithm for the valley-filling problem to reduce the communication loads while meeting the individual charging requirements. An iterative learning-based decentralized model predictive charging control algorithm is developed for achieving the valley-filling performance, and the decentralized framework is utilized to meet the individual charging requirements.

Aside from the impacts of PEV charging on the power systems, it is also important to consider the battery's longevity. The drivers of PEVs have their own charging habits and driving behaviors. Battery charging and discharging in PEVs cause battery aging problem, further resulting in severe battery capacity degradation and safety issues. Thus, it is desired to consider battery aging during driving or charging PEVs. Considering the battery aging of PEVs, **Chapter 5** develops a charging control algorithm based on the previous algorithm presented in **Chapter 4**. Using the battery aging model enables updating the optimal charging rate at each iteration that minimizes the cost of PEVs leading to valley-filling performance.

Chapter 6 concludes the dissertation and provides some promising directions for future research.

Chapter 2

Route Planning and Power Management for PHEVs with Reinforcement Learning

2.1 Introduction

Continuous endeavor has been devoted to the sustainable development of modern transportation systems. Gasoline-powered vehicles, having extensively served the world for over a century, have caused serious concerns on air pollutants and greenhouse gas emissions. According to the U.S. Energy Information Administration (EIA), the emission of CO_2 from the U.S. gasoline and diesel motor vehicles in 2016 was $1,540$ million metric tons, which counted for 30% of total CO_2 emissions in the U.S. that year [99]. Modern society is in dire need of new types of vehicles that are more environmentally friendly. Answering this urgent call, various electric vehicles have been developed, including EVs and PHEVs. EVs are subject to the problems of short driving range and long charging time. To avoid these problems, PHEVs operate with both battery and gasoline. Nevertheless, PHEVs generally occur a higher energy cost than EVs. As such, it is much needed to further reduce the energy cost of PHEVs.

Power management of PHEVs handles the switch between different operating modes and real-time torque distribution of EMs and gasoline engines. Normally, a PHEV has five possible operating modes: Motor only, engine only, power-assist (motor and engine), recharging (engine charges the battery), and regenerative braking. Roughly speaking, existing solutions to PHEV power management are classified into four categories: rule-based,

optimization-based, prediction-based, and learning-based schemes.

The *rule-based* controller [100–102] determines the operating modes of PHEVs and torque distribution by the logic threshold, fuzzy logic control, or neural networks control. In practice, rule-based schemes may be impeded by several limitations, e.g., values of logic threshold parameters need to match the vehicle status and configurations [103], and a large database of experience and experiments is usually required [104].

In the *optimization-based* method, the torque distribution is obtained by solving a constrained optimization problem subject to vehicle dynamics. The objective function is defined by the energy consumption, and prior and prediction driving cycles are utilized to distribute the power between different power sources [105, 106]. Existing solutions include dynamic programming [107, 108], Pontryagin's minimum principle [109, 110], and MPC [111–113]. Generally speaking, the methods in this category normally have high computational overhead and high complexity in changing operating conditions [114].

In the *prediction-based* schemes, the global road information is needed. Methods in this category provide the torque distribution with known route and global traffic information [115, 116]. The main difficulty of this method is that the information needs to be obtained in advance from an onboard global positioning system (GPS), geographic information system (GIS), or advanced traffic flow modeling techniques.

Among the above three categories, the optimization method is online, while the other two are offline. The main challenge in the optimization method is on the modeling of vehicles and their driving conditions. *Learning-based* methods are thus proposed to tackle this challenge.

Learning-based methods [79], in particular RL methods [117–119], provide new opportunities to improve traditional and most commonly-used methods in industry. Different from previous approaches, learning-based methods do not require complete *a priori* information of driving cycles and vehicle models for power management. They learn this knowledge during a training process. In the RL framework, there are typically four elements: the agent, the environment, actions, and rewards. In the power management problem, the objective of RL agent is to minimize PHEVs' overall fuel consumption. Q-learning is one of the basic RL algorithms [69]. One of the main advantages of Q-learning is model-free. Training the power management policy requires discretization of the generally continuous state and control actions. The discrete representation allows a straightforward solution to the integration in power management problem for PHEVs.

All the existing solutions decouple path planning from PHEV power management. Path planning, which searches for a suitable path to a given destination, is usually considered

as a separate, high-level problem and thus beyond the scope of PHEV power management. While this mindset has the root for engineering convenience, we believe that path planning and PHEV power management are interwoven, and the two problems need to be tackled in an integrated framework. Suppose that there are two paths of the same length from a source to a destination in a city. One path is flat but has several traffic lights; the other has several upper-down slopes but no traffic lights. The shortest path-based path planning would treat the two paths equally. Path planning considering traffic conditions might select the second path because of no traffic lights [120, 121]. However, the first path may be the most energy-saving path if the slopes on the second path lead to more frequent operating mode switches and a higher energy cost. Clearly, better decisions on path planning rely on the power management of PHEVs. Motivated by this observation, we propose a hierarchical control scheme that unifies path planning and power management for PHEVs energy saving. Note that our solution differs significantly from trip-oriented predictive control [122], where a path is assumed in advance, and the energy management policy is derived for the given path without addressing any ad hoc road condition changes along the trip.

Briefly speaking, our solution adopts inner and outer control loops, with the former for the learning process by considering driving conditions and the latter for energy minimization. The inner loop focuses on generating the control policy by using the model-free RL algorithm for the power management problem of PHEVs. To guarantee the solvability of uncertain system dynamics, the nonlinear approximator structure is utilized to generate the control policy for a PHEV with an actor approximator and a critic approximator. The outer loop is designed for obtaining the minimum energy consumption route by minimizing the energy consumption overall candidate rounds. Figure 2.1 shows the system configuration of the inner-outer loop system in this work.

The main contributions of this chapter are as follows.

- An RL-based route planning algorithm, including inner and outer loops, is proposed for PHEVs by taking into account the road information. In the inner loop, the control policy is generated under the model-free RL framework with road conditions for power management problems. In the outer loop, the minimum energy route is obtained by minimizing the difference between the value function of the current round and the best value from all previous rounds.

- The model-free nonlinear approximator structure for PHEVs is proposed. Under the proposed structure, the convergence of the value function for power management is studied. Also, the closed-loop properties of the resulting PHEV power management

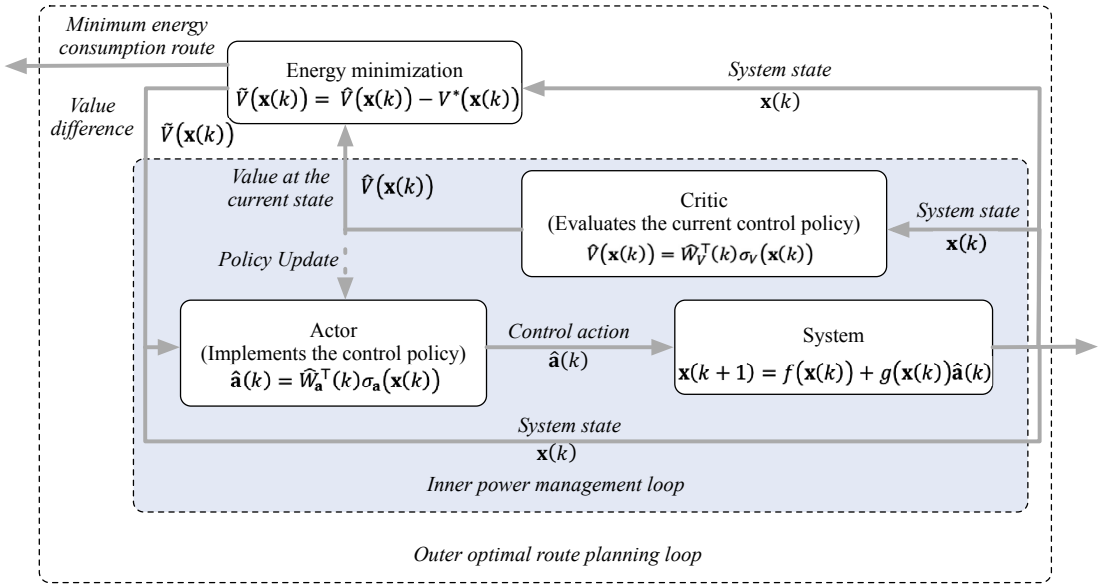


Figure 2.1: The framework of the inner-outer loop system

system are analyzed. In particular, the asymptotic stability of the closed-loop system is rigorously studied.

The rest of this chapter is organized as follows. In Section 2.2, we formally formulate the control problem and provide the algorithm design. Section 2.3 presents the main theoretical results and proofs, including the convergence of the value function for power management in the inner loop and asymptotic stability of the closed-loop system. Section 2.4 illustrates a simulation example, using Prius as the vehicle model, to verify the proposed algorithm and theoretical results. We conclude the chapter in Section 2.5.

The notations in this chapter are as follows. The symbol $\mathbb{R}$ denotes the real space. For a matrix X, $X^\top$ and X^{-1} denote the transpose and inverse of X (if X is invertible), respectively. Given a vector $\mathbf{x}$ and a matrix P, $\|\mathbf{x}\|_P = \sqrt{\mathbf{x}^\top P \mathbf{x}}$ is the P-weighted norm. For a given matrix P, $P > 0$ means that P is positive definite. $\bar{\lambda}(P)$ and $\underline{\lambda}(P)$ denote the maximum and minimum eigenvalues of P, respectively. $\mathrm{tr}(P)$ denotes the trace of P.

2.2 Problem Formulation and Algorithm Design

We first introduce our inner-outer loop framework, shown in Figure 2.1. The inner loop is for power management, and the outer loop is to find the minimum energy consumption route in the road networks. In Section 2.2.1, the constrained optimization problem

in the inner loop is solved to generate the control policy with a model-free RL algorithm. Section 2.2.2 studies the route planning problem by minimizing the energy consumption among the values from all rounds.

2.2.1 Inner Loop: Power Management Using Reinforcement Learning

To obtain a power management policy with RL, the agent learns from rewards for each action. The rewards represent interactions between the agent and the environment. The agent takes a control action based on the current state. The evaluated value for each state and action is updated by the reward. The map between states and actions is the control policy.

Power management with geographical information is modeled by the following nonlinear input affine system

$$\mathbf{x}(k+1) = f(\mathbf{x}(k)) + g(\mathbf{x}(k))\mathbf{a}(k). \tag{2.1}$$

Here, $\mathbf{x}(k)$ denotes the system state vector that consists of position $p(k)$, power demand at the wheel $P_{whl}(k)$, vehicle speed $v(k)$, and SoC of battery $SoC(k)$, i.e.,

$$\mathbf{x}(k) = [p(k) \quad P_{whl}(k) \quad v(k) \quad SoC(k)]^\top \in \mathbb{X} \subset \mathbb{R}^4,$$

where $\mathbb{X}$ denotes the system state space. $\mathbf{a}(k)$ is the action vector, including the desired output torque from engine $T_{e,des}(k)$, gear shift command of the transmission system $s(k)$, and direction $direc(k)$, i.e., $\mathbf{a}(k) = [T_{e,des}(k) \quad s(k) \quad direc(k)]^\top \in \mathbb{A}$, where $\mathbb{A}$ denotes the action space. k represents the duration of driving cycle. $s(k)$ is the gear-shifting signal, whose value is chosen from $\{-1, 0, +1\}$ representing downshifting, maintaining, and upshifting, respectively [17]. Functions, $f(\cdot)$ and $g(\cdot)$, do not have explicit forms and their meanings are explained in Remark 1.

Remark 1. *The utilization of the nonlinear input affine model for this work is mainly due to the following reasons. It is hard to obtain an accurate mathematical model for the overall system for power management. Also, the interactions between states and actions should be analyzed. The function $f(\cdot)$ describes the state updated rules, and the function $g(\cdot)$ represents the mapping between states and actions. Since generating the control policy for power management does not require complete information of overall system models, the*

nonlinear input affine system only involving related elements is a suitable system model.

The reward function, which describes the vehicle energy recuperation after action $\mathbf{a}(k)$ at state $\mathbf{x}(k)$, is given by

$$r(k) = -(Q_f(k) + Q_e(k))^2, \tag{2.2}$$

in which $Q_f(k)$ and $Q_e(k)$ denote fuel and electric consumption, respectively.

Rewards describe the interactions between the agent and the environment. In an RL problem, the evaluated value is defined as expected total future rewards. The control policy is trained to produce optimal actions, which maximizes the evaluated value for the state. The reward function involves two problems, i.e., power management and route planning, which will be introduced in detail in Section 2.2.2.

The value function satisfies the following equation

$$V(\mathbf{x}(k+1)) = \alpha[r(k+1) + \gamma V(\mathbf{x}(k+1)) - V(\mathbf{x}(k))] + V(\mathbf{x}(k)), \tag{2.3}$$

where $\mathbf{x}(k+1)$ is the next state, $\alpha \in (0,1)$ is the learning step size and $\gamma \in (0,1)$ is a discount factor. The relationship between the next and current states is discussed in Remark 3 in Section 2.2.2.

We need to model the constraints in power management. For this, assume that $\omega_e(k)$ denotes the engine speed, $T_e(k)$ represents the engine torque, $T_m(k)$ denotes the electric motor torque, $T_{req}(k)$ represents the requested torque of the vehicle. Let $\omega_{e,min}$ and $\omega_{e,max}$ denote the minimum and the maximum engine speeds, respectively. Similarly, we assume that $\omega_{m,min}$ and $\omega_{m,max}$ are the minimum and maximum electric motor speeds, respectively. Assume that $T_{e,min}(\omega_e(k))$ and $T_{e,max}(\omega_e(k))$ are the minimum and maximum engine torques with speed $\omega_e(k)$, respectively. Similarly, we assume that $T_{m,min}(\omega_m(k), SoC(k))$ and $T_{m,max}(\omega_m(k), SoC(k))$ are the minimum and maximum electric motor torques with speed $\omega_m(k)$ and $SoC(k)$, respectively. Let SoC_{min} and SoC_{max} represent the minimum and maximum values of SoC, respectively.

For safe and smooth operation of the engine, the electric motor and the battery, the following inequalities should be satisfied.

$$\omega_{e,min} \leq \omega_e(k) \leq \omega_{e,max} \tag{2.4a}$$

$$\omega_{m,min} \leq \omega_m(k) \leq \omega_{m,max} \tag{2.4b}$$

$$T_{e,min}(\omega_e(k)) \leq T_e(k) \leq T_{e,max}(\omega_e(k)) \tag{2.4c}$$

$$T_{m,min}(\omega_m(k), SoC(k)) \leq T_m(k) \leq T_{m,max}(\omega_m(k), SoC(k)) \tag{2.4d}$$

$$SoC_{min} \leq SoC(k) \leq SoC_{max}. \tag{2.4e}$$

Constrained by the working principles of the vehicle, the SoC, torques of engine and motor should satisfy the following equations, respectively.

$$T_m(k) + T_e(k) = T_{req}(k) \tag{2.5a}$$

$$SoC(k+1) = SoC(k) - \frac{U_{oc} - \sqrt{U_{oc}^2 - 4P_m R_{int}}}{2Q_{max}R_{int}} \tag{2.5b}$$

$$v(k+1) = \frac{T_w}{Mr_w} - g(f_r \cos\theta + \sin\theta) - \frac{1}{2M} C_D \rho_d A_f v(k)^2. \tag{2.5c}$$

After the constraints are formulated, the action $\mathbf{a}(k)$ is calculated by solving the following constrained optimization problem.

Problem 1 (Optimal Control Problem)**.**

$$\mathbf{a}^*(k) = \arg\max_{\mathbf{a}(k)} \{r(k+1) + V(\mathbf{x}(k+1))\} \tag{2.6}$$

subject to *constraints* (2.4a)-(2.4e) *and* (2.5a)-(2.5c).

The problem of RL-based power management is formulated with the constrained optimization problem. The actions and control policy will be produced by solving the optimal control problem. The following subsection studies the path planning problem with the goal of finding the minimum energy consumption route in a road network.

2.2.2 Outer Loop: Route Planning Algorithm

In this subsection, path planning is studied using a Q-learning algorithm to find the minimum energy consumption route. Using Q-learning algorithm, the next state $\mathbf{x}(k+1)$ depends on current state $\mathbf{x}(k)$. The action $\mathbf{a}(k)$ is uncertain before the agent arrives at the next state. Then the Q-value is utilized to evaluate the control policy μ. Also, at the current state $\mathbf{x}(k)$, the next state $\mathbf{x}(k+1)$ of the agent is uncertain after taking action $\mathbf{a}(k)$. For this reason, all the possibilities require to be considered. After the agent tries most of the actions in each state, the map between a given state and a specific action is constructed. At the end of the algorithm, the agent finds the optimal results by taking the highest value of action at each state.

Assume that a city map (e.g., Figure 2.2) with roads and geographic information is given, including the location of each road intersection, the vehicle speed limited by each road segment, length and slope of each road segment.

A route linking the source and the destination is decomposed into a series of road segments connected to each other. In each road segment, the distance, time, fuel consumption and electricity recuperation are calculated with the road information, such as the length l, road slope angle θ and vehicle speed v, respectively. The changes in road slope and vehicle speed affect the fuel consumption and electricity recuperation of the battery. For instance, the road may change from downhill to uphill; speed may change due to deceleration to stop signs, traffic lights or traffic jams and acceleration from stop signs or traffic lights.

The SoC of the battery is the core dynamic state in the control system. We assume that the battery must recover with the help of the engine when the SoC is below the threshold. $SoC(k)$ and $SoC_{ref}(k)$ are the current and reference SoC, respectively. ζ is the coefficient for SoC, whose value is determined by the empirical model introduced in Section 2.4.1.

The main difference between the proposed route planning algorithm and the shortest path algorithm is that the reward function (2.2) in the proposed method uses fuel consumption and battery recuperation, while the shortest path algorithm only considers the distance. The reward function (2.2) in the Q-learning algorithm is designed to evaluate the road for the route planning problem, including the fuel consumption $Q_f(k)$ and electricity recuperation $Q_e(k)$. The reward is the same as the one in the previous section. The fuel consumption $Q_f(k)$ and electricity recuperation $Q_e(k)$ in the k-th step are calculated based on the road slope angle $\theta(k)$, length $l(k)$, vehicle speed $v(k)$, $SoC(k)$ and $SoC_{ref}(k)$ as follows.

$$Q_f(k) = \phi(\theta(k), v(k), l(k)) \tag{2.7a}$$

$$Q_e(k) = \zeta(SoC(k) - SoC_{ref}(k)). \tag{2.7b}$$

To this end, a series of road segments is obtained by solving the following optimization problem.

Problem 2. *Given the values of the current round and the best value from all previous rounds, find a series of road segments to minimize the fuel consumption.*

Remark 2. *To reduce the computational overhead, we can define a circular boundary area that covers the start point and the destination, as shown in Figure 2.2. When the state approaches the boundary, we can add a large penalty so that the reward of the current state drops sharply.*

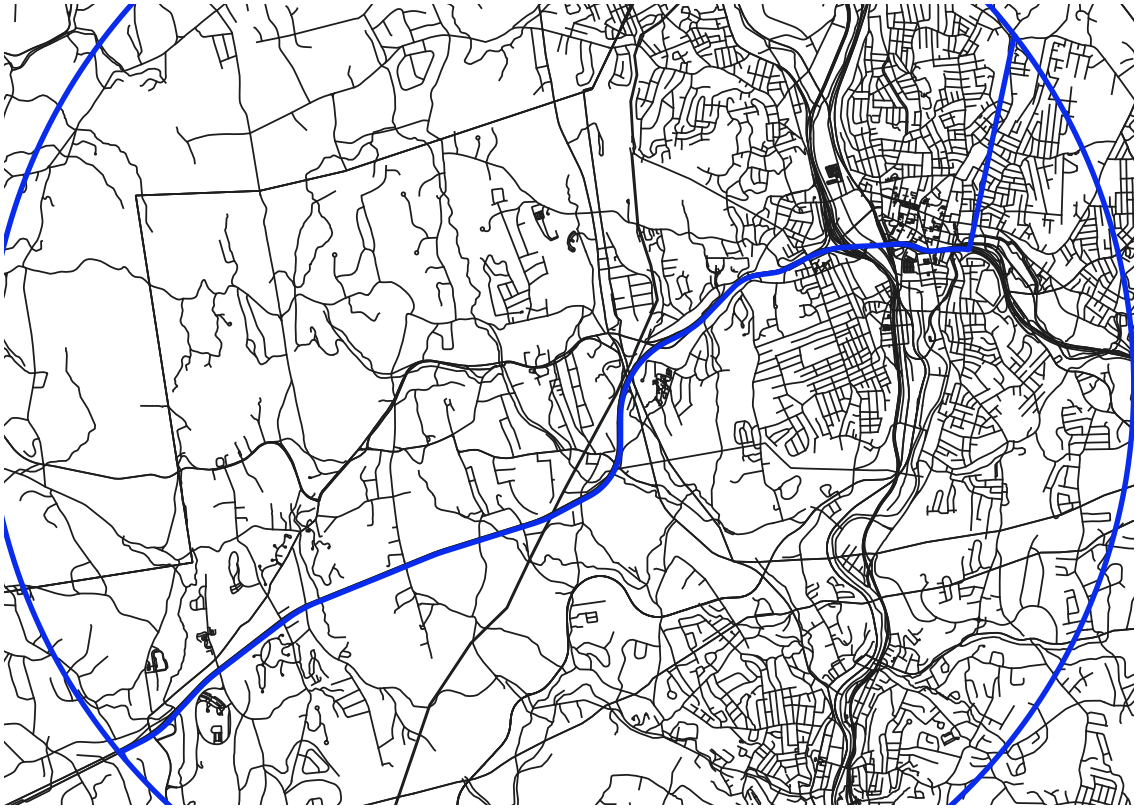


Figure 2.2: The blue circle represents the boundary area to reduce the computational overhead.

With reward function defined in (2.2), we design the following RL based route planning algorithm.

Remark 3. *The inner loop of the algorithm generates the updated action based on the estimation of the next state. The next state is estimated from the current state and possible action. The position* $p(k+1)$ *is obtained based on* $direc(k)$ *of the action* $\mathbf{a}(k)$ *at the current state. We assume that the power demand* $P_{whl}(k+1)$ *of the next state is post-action parameters, which are the same as the pair observed in the current state* $P_{whl}(k)$*. The* $SoC(k+1)$ *and velocity* $v(k+1)$ *are derived based on the empirical model introduced in the previous section. The outer loop of the algorithm records the value function from the inner loop and returns when the value function converges.*

In this chapter, we use the Q-learning algorithm to learn the Q-table. The Q-table is the value of the maximum expected future reward for choosing the optimal strategy in a given state. Before the learning algorithm begins, the values of the Q-table are randomized. As we gradually use the algorithm to learn, the Q-table will be updated with increasingly better

Algorithm 1: Route planning for estimating control policy μ

while $\tilde{V}(\mathbf{x}(k))$ *is greater than a given small threshold.* **do** /* outer loop*/
  while $\mathbf{x}(k)$ *is not the target* **do** /* inner loop*/
    Initialize $V(\mathbf{x}(k))$ arbitrarily at $k = 0$;
    Observe $\mathbf{x}(k)$;
    Choose $\mathbf{a}(k)$ at $\mathbf{x}(k)$ using the proposed policy;
    Take action $\mathbf{a}(k)$, observe $r(k)$ and $\mathbf{x}(k+1)$;
    Update the value at $\mathbf{x}(k+1)$ by (2.3);
    Set $k = k + 1$;
  end
  /* end of inner loop*/;
  Calculate $\tilde{V}(\mathbf{x}(k))$, which is the difference between $V(\mathbf{x}(k))$ of current round and the best value of $V^*(\mathbf{x}(k))$ from all previous rounds;
end
/* end of outer loop */

estimates by continuously applying Bellman's equation updates.

However, the main limitation of applying the Q-learning algorithm to this work is that the complicated road networks cause computational overheads and a huge Q-table. Among these issues, the actor-critic method is incorporated into the inner power management loop. Generally speaking, the critic estimates the value function, while the actor updates the action in the direction suggested by the critic. The actor-critic method has two independent parallel NNs. one is designed for optimizing actions. The other one is used to update the value function. We take the present state and enter this as input to the actor and the critic at each time. The critic computes the value of the action performed in that state, and the actor updates the weights with this Q-value. After updating the parameters, the actor generates the next state, and the critic uses this state to update its value. In particular, the power management problem and route planning are incorporated in the proposed algorithm, so we still need to build the Q-table. The Q-table is used to find the minimized energy consumption route in the outer loop.

2.3 System Analysis

In this section, the closed-loop properties of the proposed PHEV power management system are analyzed. First, the Hamilton-Jacobi-Bellman (HJB) equation for a nonlinear discrete-time system is derived. In addition, the convergence of the inner loop power man-

agement policy is discussed. In the end, the asymptotic stability of the closed-loop system is proved.

2.3.1 Hamilton-Jacobi-Bellman Equation

The value function (2.3) of the Q-learning algorithm is the main evaluation index of the control policy μ with given actions and system states. To prove the convergence of value function and stability of the closed-loop system with Q-learning control policy, the existence of the Hamiltonian function $H(\mathbf{x}(k), V(\mathbf{x}(k)), \mathbf{a}(k))$ for nonlinear dynamics (2.1) needs to be proved. First, we introduce a well-known result [123].

Lemma 1 ([123])**.** *Assume that $\mathbf{a}(k) \in \mathbb{A}$ is an admissible action arbitrarily selected for the nonlinear dynamics (2.1). If there exists a positive, uniformly convex, and continuously differentiable value function $V(\mathbf{x}(k))$ satisfying the following equations*

$$\begin{aligned} &\|\mathbf{x}(k)\|_Q^2 + \|\mathbf{a}(k)\|_R^2 + (\mathbf{x}(k+1) - \mathbf{x}(k))^\top \times \nabla V(\mathbf{x}(k))(\mathbf{x}(k+1) - \mathbf{x}(k)) \\ =&\|\mathbf{x}(k)\|_Q^2 + \|\mathbf{a}(k)\|_R^2 + \Delta\mathbf{x}^\top(k)\nabla V(\mathbf{x}(k))\Delta\mathbf{x}(k) \\ =&0, \end{aligned} \tag{2.8}$$

where Q and R are two positive definite matrices. $\Delta\mathbf{x}(k) = f(\mathbf{x}(k)) + g(\mathbf{x}(k))\mathbf{a}(k) - \mathbf{x}(k) \in \mathbb{X}$, and $\nabla V(\mathbf{x}(k))$ is the gradient vector of $V(\mathbf{x}(k))$, a suitable Hamiltonian function for the nonlinear system exists

$$\begin{aligned} H(\mathbf{x}(k), V(\mathbf{x}(k)), \mathbf{a}(k)) =&(\mathbf{x}(k+1) - \mathbf{x}(k))^\top \nabla V(\mathbf{x}(k))(\mathbf{x}(k+1) - \mathbf{x}(k)) \\ &+ \|\mathbf{x}(k)\|_Q^2 + \|\mathbf{a}(k)\|_R^2 \\ =&\|\mathbf{x}(k)\|_Q^2 + \|\mathbf{a}(k)\|_R^2 + \Delta\mathbf{x}^\top(k)\nabla V(\mathbf{x}(k))\Delta\mathbf{x}(k). \end{aligned} \tag{2.9}$$

The Hamiltonian function (2.9) is a nonlinear function of the state, value function and actions. Thus, by differentiating (2.9) with respect to actions $\mathbf{a}(k)$, the updated action is given by

$$\begin{aligned} \mathbf{a}^*(k) =& -\left(g^\top(\mathbf{x}(k))\nabla V(\mathbf{x}(k))g(\mathbf{x}(k)) + R\right)^{-1} \times g^\top(\mathbf{x}(k))\left(\nabla V(\mathbf{x}(k))f(\mathbf{x}(k))\mathbf{x}(k)\right) \\ =& -\frac{1}{2}R^{-1}g^\top(\mathbf{x}(k))\nabla V(\mathbf{x}(k+1)). \end{aligned} \tag{2.10}$$

Thus, the Hamiltonian function (2.9) for the nonlinear dynamics is expressed as

$$\begin{aligned} H(\mathbf{x}(k), V(\mathbf{x}(k)), \mathbf{a}^*(k)) =& \|\mathbf{x}(k)\|_Q^2 + \|\mathbf{a}^*(k)\|_R^2 + \Delta\mathbf{x}^{*^\top}(k) \cdot \nabla V(\mathbf{x}(k)) \Delta\mathbf{x}^*(k) \\ =& 0, \end{aligned} \tag{2.11}$$

where $\Delta\mathbf{x}^*(k) = f(\mathbf{x}(k)) + g(\mathbf{x}(k))\mathbf{a}^*(k) - \mathbf{x}(k) \in \mathbb{X}$, and the optimal action $\mathbf{a}^*(k)$ is obtained from (2.10).

According to (2.10), the optimal action is calculated based on $f(\cdot)$ and $g(\cdot)$, whose accurate models are normally unknown. Hence we adopt an approximator structure to address this difficulty. In addition, the stability of the closed-loop system with the approximator structure and the approximate control policy is discussed in Section 2.3.3.

2.3.2 Convergence

The main objective of the proposed method is to find a series of road segments linking the source and the destination to minimize fuel consumption. Also, the agent will try most of the actions and states to generate the action sequence. In the following, we analyze the convergence of the actions.

To discuss the convergence of the actions, a good result in [124] Theorem 2.2 is recalled.

Lemma 2. *(Convergence [124]) Assume that the admissible action* $\mathbf{a}(k)$ *is obtained iteratively by solving* (2.8) *and updated using* (2.10)*, the sequence of solutions* $V(\mathbf{x}(k))$ *converges to the optimal HJB solution* $V^*(\mathbf{x}(k))$.

Proof. See Theorem 2.2 in [124]. □

The control actor and critic approximators are designed to approximate the optimal control input and Q-value, respectively. Since the approximation error is relatively small, we do not consider it in this work. Thus, the actual control signal and value function are defined as

$$\begin{aligned} \mathbf{a}(k) &= W_{\mathbf{a}}^\top \sigma_{\mathbf{a}}(\mathbf{x}(k)) = \mathbf{a}^*(k) \\ V(\mathbf{x}(k)) &= W_V^\top \sigma_V(\mathbf{x}(k)). \end{aligned}$$

In addition, the approximated control input and value function are represented as

$$\begin{aligned} \hat{\mathbf{a}}(k) &= \hat{W}_{\mathbf{a}}^\top(k) \sigma_{\mathbf{a}}(\mathbf{x}(k)) \\ \hat{V}(\mathbf{x}(k)) &= \hat{W}_V^\top(k) \sigma_V(\mathbf{x}(k)), \end{aligned}$$

where $\hat{W}_V(k)$ and $\hat{W}_{\mathbf{a}}(k)$ are the estimated weights of the optimal value function and action, respectively. $\sigma_V(\mathbf{x}(k))$ and $\sigma_{\mathbf{a}}(\mathbf{x}(k))$ denote the activation functions, which satisfy $\sigma_V(0) = 0$ and $\sigma_{\mathbf{a}}(0) = 0$ when $\|\mathbf{x}(k)\| = 0$. The following Bellman equation should be satisfied to guarantee the optimality.

$$r(k+1) + V(\mathbf{x}(k+1)) - V(\mathbf{x}(k)) = 0. \tag{2.12}$$

According to the definition of value function, the Bellman equation is rewritten as

$$\begin{aligned} e_B(k) =& r(k+1) + \hat{V}(\mathbf{x}(k+1)) - \hat{V}(\mathbf{x}(k)) \\ =& r(k+1) + \hat{W}_V^\top(k)\sigma_V(\mathbf{x}(k+1)) - \hat{W}_V^\top(k)\sigma_V(\mathbf{x}(k)). \end{aligned}$$

In addition, we can obtain an error related to the terminal constraint, which is defined as

$$e_T(k) = -\hat{W}_V^\top(k)\sigma_V(\hat{\mathbf{x}}(\infty)),$$

where $\hat{\mathbf{x}}(\infty)$ denotes the terminal state. Thus, the final error is obtained as

$$\begin{aligned} e_F(k) =& e_B(k) + e_T(k) \\ =& r(k+1) + \hat{W}_V^\top(k)\sigma_V(\mathbf{x}(k+1)) - \hat{W}_V^\top(k)\sigma_V(\mathbf{x}(k)) - \hat{W}_V^\top(k)\sigma_V(\hat{\mathbf{x}}(\infty)). \end{aligned} \tag{2.13}$$

With gradient descent, the updated estimation weight of value function is represented as

$$\hat{W}_V(k+1) = \hat{W}_V(k) - \alpha_V \frac{\tilde{\sigma_V}(\mathbf{x}(k))e_F(k)}{1 + \tilde{\sigma_V}^\top(\mathbf{x}(k))\tilde{\sigma_V}(\mathbf{x}(k))}, \tag{2.14}$$

where $\tilde{\sigma_V}(\mathbf{x}(k)) = \sigma_V(\mathbf{x}(k+1)) - \sigma_V(\mathbf{x}(k)) - \sigma_V(\hat{\mathbf{x}}(\infty))$, and the bound of $\tilde{\sigma_V}(\mathbf{x}(k))$ is $\tilde{\sigma}_{Vm} \leq \|\tilde{\sigma_V}(\mathbf{x}(k))\| \leq \tilde{\sigma}_{VM}$. In addition, the error between the estimation and actual weights is defined as

$$\tilde{W}_V(k) = W_V - \hat{W}_V(k)$$

According to the Bellman equation (2.12), we can obtain that

$$\begin{aligned} r(k+1) + V(\mathbf{x}(k+1)) - V(\mathbf{x}(k)) =& r(k+1) + W_V^\top\sigma_V(\mathbf{x}(k+1)) - W_V^\top\sigma_V(\mathbf{x}(k)) \\ =& r(k+1) + W_V^\top(\sigma_V(\mathbf{x}(k+1)) - \sigma_V(\mathbf{x}(k))) \\ =& 0. \end{aligned}$$

Thus, we can obtain $r(k+1) = -W_V^\top(\sigma_V(\mathbf{x}(k+1)) - \sigma_V(\mathbf{x}(k)))$, and rewrite (2.13) as

$$\begin{aligned} e_F(k) =& e_B(k) + e_T(k) \\ =& - W_V^\top(\sigma_V(\mathbf{x}(k+1)) - \sigma_V(\mathbf{x}(k))) + \hat{W}_V^\top(k)\sigma_V(\mathbf{x}(k+1)) \\ & - \hat{W}_V^\top(k)\sigma_V(\mathbf{x}(k)) - \hat{W}_V^\top(k)\sigma_V(\hat{\mathbf{x}}(\infty)) \\ =& - \tilde{W}_V^\top(k)(\sigma_V(\mathbf{x}(k+1)) - \sigma_V(\mathbf{x}(k))) - \hat{W}_V^\top(k)\sigma_V(\hat{\mathbf{x}}(\infty)). \end{aligned}$$

Then, the error of updated estimation weight (2.14) is described as

$$\begin{aligned} \tilde{W}_V(k+1) =& \tilde{W}_V(k) - \alpha_V \frac{\tilde{\sigma}_V(\mathbf{x}(k))\tilde{\sigma}_V^\top(\mathbf{x}(k))\tilde{W}_V(k)}{1 + \tilde{\sigma}_V^\top(\mathbf{x}(k))\tilde{\sigma}_V(\mathbf{x}(k))} \\ & - \alpha_V \frac{\tilde{\sigma}_V(\mathbf{x}(k))(\sigma_V^\top(\mathbf{x}(\infty)) - \sigma_V^\top(\hat{\mathbf{x}}(\infty)))W_V}{1 + \tilde{\sigma}_V^\top(\mathbf{x}(k))\tilde{\sigma}_V(\mathbf{x}(k))}. \end{aligned}$$

Next, we define the difference between the approximated control input and optimal control policy as

$$\tilde{\mathbf{a}}(k) = \hat{\mathbf{a}}(k) - \mathbf{a}^*(k) = \frac{1}{2}R^{-1}\hat{g}^\top(\mathbf{x}(k))\nabla\hat{V}(\mathbf{x}(k+1)) + \hat{W}_\mathbf{a}^\top(k)\sigma_\mathbf{a}(\mathbf{x}(k)).$$

Here, $\hat{g}(\cdot)$ is an estimated function from neural networks. With gradient descent, the updated estimation weight of action is represented as

$$\hat{W}_\mathbf{a}(k+1) = \hat{W}_\mathbf{a}(k) - \alpha_\mathbf{a}\frac{\sigma_\mathbf{a}(\mathbf{x}(k))\tilde{\mathbf{a}}^\top(k)}{1 + \sigma_\mathbf{a}^\top(\mathbf{x}(k))\sigma_\mathbf{a}(\mathbf{x}(k))}, \tag{2.15}$$

where $\alpha_\mathbf{a} > 0$ is a constant. Then, the error between the actual weight and estimation weight is defined as

$$\tilde{W}_\mathbf{a}(k) = W_\mathbf{a} - \hat{W}_\mathbf{a}(k)$$

Recall the optimal action, we have

$$\mathbf{a}^*(k) = W_\mathbf{a}^\top\sigma_\mathbf{a}(\mathbf{x}(k)) = -\frac{1}{2}R^{-1}g^\top(\mathbf{x}(k))\nabla\sigma_V^\top(\mathbf{x}(k+1))W_V.$$

Then we can obtain that

$$W_\mathbf{a}^\top\sigma_\mathbf{a}(\mathbf{x}(k)) + \frac{1}{2}R^{-1}g^\top(\mathbf{x}(k))\nabla\sigma_V^\top(\mathbf{x}(k+1))W_V = 0.$$

In addition, the error of control policy is rewritten as

$$\begin{aligned}\tilde{\mathbf{a}}(k) =& -\tilde{W}_{\mathbf{a}}^{\top}(k)\sigma_{\mathbf{a}}(\mathbf{x}(k)) - \frac{1}{2}R^{-1}g^{\top}(\mathbf{x}(k))\nabla\sigma_V^{\top}(\mathbf{x}(k+1))W_V \\ &+ \frac{1}{2}R^{-1}\hat{g}^{\top}(\mathbf{x}(k))\nabla\sigma_V^{\top}(\mathbf{x}(k+1))\hat{W}_V(k).\end{aligned} \tag{2.16}$$

Finally, substituting (2.16) into (2.15), we can rewrite the updated estimation weight of action as

$$\begin{aligned}\hat{W}_{\mathbf{a}}(k+1) =& \hat{W}_{\mathbf{a}}(k) - \alpha_{\mathbf{a}}\frac{\sigma_{\mathbf{a}}(\mathbf{x}(k))\tilde{\mathbf{a}}^{\top}(k)}{1+\sigma_{\mathbf{a}}^{\top}(\mathbf{x}(k))\sigma_{\mathbf{a}}(\mathbf{x}(k))} \\ =& \hat{W}_{\mathbf{a}}(k) - \alpha_{\mathbf{a}}\frac{\sigma_{\mathbf{a}}(\mathbf{x}(k))}{1+\sigma_{\mathbf{a}}^{\top}(\mathbf{x}(k))\sigma_{\mathbf{a}}(\mathbf{x}(k))} \times \Big(-\tilde{W}_{\mathbf{a}}^{\top}(k)\sigma_{\mathbf{a}}(\mathbf{x}(k)) \\ &- \frac{1}{2}R^{-1}g^{\top}(\mathbf{x}(k))\nabla\sigma_V^{\top}(\mathbf{x}(k+1))W_V \\ &+ \frac{1}{2}R^{-1}\hat{g}^{\top}(\mathbf{x}(k))\nabla\sigma_V^{\top}(\mathbf{x}(k+1))\tilde{W}_V(k)\Big).\end{aligned}$$

Then, the error of updated estimation weight of action is described as

$$\begin{aligned}\tilde{W}_{\mathbf{a}}(k+1) =& \tilde{W}_{\mathbf{a}}(k) + \alpha_{\mathbf{a}}\frac{\sigma_{\mathbf{a}}(\mathbf{x}(k))\tilde{\mathbf{a}}^{\top}(k)}{1+\sigma_{\mathbf{a}}^{\top}(\mathbf{x}(k))\sigma_{\mathbf{a}}(\mathbf{x}(k))} \\ =& \tilde{W}_{\mathbf{a}}(k) - \alpha_{\mathbf{a}}\frac{\sigma_{\mathbf{a}}(\mathbf{x}(k))}{1+\sigma_{\mathbf{a}}^{\top}(\mathbf{x}(k))\sigma_{\mathbf{a}}(\mathbf{x}(k))} \times \Big(\tilde{W}_{\mathbf{a}}^{\top}(k)\sigma_{\mathbf{a}}(\mathbf{x}(k)) \\ &+ \frac{1}{2}R^{-1}g^{\top}(\mathbf{x}(k))\nabla\sigma_V^{\top}(\mathbf{x}(k+1))W_V \\ &- \frac{1}{2}R^{-1}\hat{g}^{\top}(\mathbf{x}(k))\nabla\sigma_V^{\top}(\mathbf{x}(k+1))\tilde{W}_V(k)\Big).\end{aligned}$$

Remark 4. *The control policy generated by the RL algorithm has an effect on the current state and updated action. However, convergence is required for all the state-action pairs while they are updated recursively.*

2.3.3 Stability

After the convergence is proved, the stability of the closed-loop system is further investigated in this subsection. The optimal value function $V(\mathbf{x}(k))$ is taken as the Lyapunov function. Based on Theorem 1 in [125], the following theorem is derived.

Theorem 1. *Given* $\mathbf{a}(k) \in \mathbb{A}$ *and the initial state* $\mathbf{x}(0) \in \mathbb{X}$*, and the value function* $V(\mathbf{x}(k))$

satisfying Equation (2.8) with the condition $V(0) = 0$, the updated action (2.10) by using the Hamiltonian results is an admissible action for the nonlinear dynamics (3.1). The system is asymptotically stable using the control policy μ generated in Algorithm 1.

Proof. In order to prove the stability, a good result is recalled [125]. We start with the Lyapunov candidate.

$$L = V(\mathbf{x}(k)) + \frac{1}{2\alpha_V}\|\tilde{W}_V(k)\|^2 + \frac{1}{2\alpha_{\mathbf{a}}}\text{tr}(\tilde{W}_{\mathbf{a}}^\top(k)\tilde{W}_{\mathbf{a}}(k)), \tag{2.17}$$

where $\tilde{W}_V$ and $\tilde{W}_{\mathbf{a}}$ are the weight estimation errors of the dynamics and actions, respectively. $\alpha_V > 0$ and $\alpha_{\mathbf{a}} > 0$ are constants, which denote the convergence speeds of two approximators. Next, we analyze every term in (2.17). According to the system dynamics, we can obtain the difference of the value function with the action $\mathbf{a}(k+1)$.

$$\begin{aligned}
\Delta V(\mathbf{x}(k)) =& V(\mathbf{x}(k+1)) - V(\mathbf{x}(k)) \\
=& \nabla V^\top(\mathbf{x}(k))(f(\mathbf{x}(k)) + g(\mathbf{x}(k))\mathbf{a}(k) - (\mathbf{x}(k))) + \frac{1}{2}(f(\mathbf{x}(k)) + g(\mathbf{x}(k))\mathbf{a}(k) \\
& - (\mathbf{x}(k)))^\top \times \nabla^2 V(\mathbf{x}(k))(f(\mathbf{x}(k)) + g(\mathbf{x}(k))\mathbf{a}(k) - (\mathbf{x}(k))) \\
\leq& \nabla V^\top(\mathbf{x}(k))(f(\mathbf{x}(k)) + g(\mathbf{x}(k))\mathbf{a}(k+1) - \mathbf{x}(k)) \\
& + \frac{1}{2}(f(\mathbf{x}(k)) + g(\mathbf{x}(k))\mathbf{a}(k+1) \\
& - \mathbf{x}(k))^\top \times \nabla^2 V(\mathbf{x}(k))\left(f(\mathbf{x}(k)) + g(\mathbf{x}(k))\mathbf{a}(k+1) - \mathbf{x}(k)\right).
\end{aligned} \tag{2.18}$$

According to (2.8), rewrite (2.18) as

$$\begin{aligned}
\Delta V(\mathbf{x}(k)) =& -\frac{1}{2}\mathbf{a}^\top(k)(g^\top(\mathbf{x}(k))\nabla V(\mathbf{x}(k))g(\mathbf{x}(k)) + R)\mathbf{a}(k) - \|\mathbf{x}(k)\|_Q^2 \\
& + \left(\nabla V^\top(\mathbf{x}(k)) + (f(\mathbf{x}(k)) - \mathbf{x}(k))^\top \nabla^2 V(\mathbf{x}(k))\right) \\
& \times g(\mathbf{x}(k))(\mathbf{a}(k+1) - \mathbf{a}(k)) + \frac{1}{2}\mathbf{a}^\top(k+1)(g^\top(\mathbf{x}(k))\nabla V(\mathbf{x}(k))g(\mathbf{x}(k)))\mathbf{a}(k+1).
\end{aligned}$$

Since the updated action is obtained by (2.10), the above equation is rewritten as follows.

$$\begin{aligned}
\Delta V(\mathbf{x}(k)) =& -\|\mathbf{x}(k)\|_Q^2 - \|\mathbf{a}(k+1)\|_R^2 - \frac{1}{2}(\mathbf{a}(k+1) - \mathbf{a}(k))^\top \\
& \times (g^\top(\mathbf{x}(k))\nabla V(\mathbf{x}(k))g(\mathbf{x}(k)) + R)(\mathbf{a}(k+1) - \mathbf{a}(k)).
\end{aligned}$$

Since $g^\top(\mathbf{x}(k))\nabla V(\mathbf{x}(k))g(\mathbf{x}(k)) + R$, R and Q are positive definite, we can obtain

$$\Delta V(\mathbf{x}(k)) \leq -\|\mathbf{x}(k)\|_Q^2 \leq -\underline{\lambda}(Q)\|\mathbf{x}(k)\|^2 \tag{2.19}$$

Thus, the closed-loop system with the updated action $\mathbf{a}(k+1)$ is locally asymptotically stable. According to Lemma 2, due to the convergence of the value function $V(\mathbf{x}(k))$, the cost function of the system (2.1) with the updated action $\mathbf{a}(k+1)$ is finite. In addition, $\hat{V}(0) = 0$ is guaranteed. We define $L_{\mathbf{a}}(k) = \text{tr}(\tilde{W}_{\mathbf{a}}^\top(k)\tilde{W}_{\mathbf{a}}(k))$. The first difference is obtained as

$$\begin{aligned}\Delta L_{\mathbf{a}}(k) &= \text{tr}(\tilde{W}_{\mathbf{a}}^\top(k+1)\tilde{W}_{\mathbf{a}}(k+1)) - \text{tr}(\tilde{W}_{\mathbf{a}}^\top(k)\tilde{W}_{\mathbf{a}}(k)) \\ &= \frac{-2\alpha_{\mathbf{a}}\sigma_{\mathbf{a}}(\mathbf{x}(k))}{1+\sigma_{\mathbf{a}}^\top(\mathbf{x}(k))\sigma_{\mathbf{a}}(\mathbf{x}(k))}\text{tr}(\tilde{\mathbf{a}}(k)\tilde{\mathbf{a}}^\top(k)) + \frac{\alpha_{\mathbf{a}}\sigma_{\mathbf{a}}^\top(\mathbf{x}(k))\sigma_{\mathbf{a}}(\mathbf{x}(k))}{(1+\sigma_{\mathbf{a}}^\top(\mathbf{x}(k))\sigma_{\mathbf{a}}(\mathbf{x}(k))^2}\text{tr}(\tilde{\mathbf{a}}(k)\tilde{\mathbf{a}}^\top(k)).\end{aligned} \tag{2.20}$$

Considering the properties of trace operator, we substitute (2.16) and $\frac{\sigma_{\mathbf{a}}^\top(\mathbf{x}(k))\sigma_{\mathbf{a}}(\mathbf{x}(k))}{1+\sigma_{\mathbf{a}}^\top(\mathbf{x}(k))\sigma_{\mathbf{a}}(\mathbf{x}(k))} < 1$ into (2.20). By applying the upper bounds of matrices, the first difference is rewritten as

$$\begin{aligned}\Delta L_{\mathbf{a}}(k) \leq\ & \alpha_{\mathbf{a}}^2 R g_M \sigma_{VM} W_{VM}\left(\frac{\sigma_{\mathbf{a}}^\top(\mathbf{x}(k))\sigma_{\mathbf{a}}(\mathbf{x}(k))}{1+\sigma_{\mathbf{a}}^\top(\mathbf{x}(k))\sigma_{\mathbf{a}}(\mathbf{x}(k))}\right)^2 + \frac{\alpha_{\mathbf{a}}(g_M\sigma_{VM}RW_{VM})^2}{4\left(1+\sigma_{\mathbf{a}}^\top(\mathbf{x}(k))\sigma_{\mathbf{a}}(\mathbf{x}(k))\right)^2} \\ & - \frac{2\alpha_{\mathbf{a}}\sigma_{\mathbf{a}}^\top(\mathbf{x}(k))\sigma_{\mathbf{a}}(\mathbf{x}(k))}{1+\sigma_{\mathbf{a}}^\top(\mathbf{x}(k))\sigma_{\mathbf{a}}(\mathbf{x}(k))}\|\tilde{W}_{\mathbf{a}}(k)\|^2 + \alpha_{\mathbf{a}}^2\left(\frac{\sigma_{\mathbf{a}}^\top(\mathbf{x}(k))\sigma_{\mathbf{a}}(\mathbf{x}(k))}{1+\sigma_{\mathbf{a}}^\top(\mathbf{x}(k))\sigma_{\mathbf{a}}(\mathbf{x}(k))}\right)^2\|\tilde{W}_{\mathbf{a}}(k)\|^2 \\ & + \alpha_{\mathbf{a}}^2\left(\frac{\sigma_{\mathbf{a}}^\top(\mathbf{x}(k))\sigma_{\mathbf{a}}(\mathbf{x}(k))Rg_M\sigma_{VM}}{1+\sigma_{\mathbf{a}}^\top(\mathbf{x}(k))\sigma_{\mathbf{a}}(\mathbf{x}(k))}\right)^2\|\tilde{W}_V(k)\|^2 \\ & - \frac{\alpha_{\mathbf{a}}\sigma_{\mathbf{a}}^\top(\mathbf{x}(k))Rg_M\sigma_{VM}W_{VM}}{1+\sigma_{\mathbf{a}}^\top(\mathbf{x}(k))\sigma_{\mathbf{a}}(\mathbf{x}(k))}\|\tilde{W}_{\mathbf{a}}(k)\|^2,\end{aligned} \tag{2.21}$$

where g_M is the upper bound of function $g(\cdot)$. Since the bound of $\sigma_{\mathbf{a}}(\mathbf{x}(k))$ is $0 < \sigma_{\mathbf{a}m} \leq \|\sigma_{\mathbf{a}}\| \leq \sigma_{\mathbf{a}M}$, (2.21) is rewritten as

$$\begin{aligned}\Delta L_{\mathbf{a}}(k) \leq & \alpha_{\mathbf{a}}^2 R g_M \sigma_{VM} W_{VM}\left(\frac{\sigma_{\mathbf{a}M}^2}{1+\sigma_{\mathbf{a}m}^2}\right)^2 - \frac{2\alpha_{\mathbf{a}}\sigma_{\mathbf{a}M}^2}{1+\sigma_{\mathbf{a}m}^2}\|\tilde{W}_{\mathbf{a}}(k)\|^2 + \frac{\alpha_{\mathbf{a}}(g_M\sigma_{VM}RW_{VM})^2}{4\left(1+\sigma_{\mathbf{a}m}^2\right)^2} \\ & + \alpha_{\mathbf{a}}^2\left(\frac{\sigma_{\mathbf{a}M}^2}{1+\sigma_{\mathbf{a}m}^2}\right)^2\|\tilde{W}_{\mathbf{a}}(k)\|^2 + \alpha_{\mathbf{a}}^2\left(\frac{\sigma_{\mathbf{a}M}^2 Rg_M\sigma_{VM}}{1+\sigma_{\mathbf{a}m}^2}\right)^2\|\tilde{W}_V(k)\|^2 \\ & - \frac{\alpha_{\mathbf{a}}\sigma_{\mathbf{a}m}Rg_M\sigma_{VM}W_{VM}}{1+\sigma_{\mathbf{a}M}^2}\|\tilde{W}_{\mathbf{a}}(k)\|^2\end{aligned}$$

$$= -\alpha_{\mathbf{a}} \frac{2\sigma_{\mathbf{a}m}^2 + \sigma_{\mathbf{a}M} Rg_M \sigma_{VM} W_{VM}}{1+\sigma_{\mathbf{a}M}^2} \|\tilde{W}_{\mathbf{a}}(k)\|^2 \tag{2.22}$$

$$+ \alpha_{\mathbf{a}}^2 \left(\frac{\sigma_{\mathbf{a}M}^2 Rg_M \sigma_{VM}}{1+\sigma_{\mathbf{a}m}^2} \right)^2 \|\tilde{W}_V(k)\|^2$$

$$- \alpha_{\mathbf{a}}^2 \frac{\sigma_{\mathbf{a}m}^4}{(1+\sigma_{\mathbf{a}M}^2)^2} \|\tilde{W}_{\mathbf{a}}(k)\|^2. \tag{2.23}$$

Then, since we have $L_V(k) = \|\tilde{W}_V(k)\|^2$, we can obtain the first difference as follows

$$\begin{aligned}
\Delta L_V(k) =& \|\tilde{W}_V(k+1)\|^2 - \|\tilde{W}_V(k)\|^2 \\
=& \Big(\tilde{W}_V(k) - \alpha_V \frac{\tilde{\sigma_V}(\mathbf{x}(k)) \tilde{\sigma_V}^\top(\mathbf{x}(k)) \tilde{W}_V(k)}{1+\tilde{\sigma_V}^\top(\mathbf{x}(k))\tilde{\sigma_V}(\mathbf{x}(k))} \\
& - \alpha_V \frac{\tilde{\sigma_V}(\mathbf{x}(k))(\sigma_V^\top(\mathbf{x}(\infty)) - \sigma_V^\top(\hat{\mathbf{x}}(\infty)))\hat{W}_V(k)}{1+\tilde{\sigma_V}^\top(\mathbf{x}(k))\tilde{\sigma_V}(\mathbf{x}(k))} \Big)^\top \\
& \times \Big(\tilde{W}_V(k) - \alpha_V \frac{\tilde{\sigma_V}(\mathbf{x}(k)) \tilde{\sigma_V}^\top(\mathbf{x}(k)) \tilde{W}_V(k)}{1+\tilde{\sigma_V}^\top(\mathbf{x}(k))\tilde{\sigma_V}(\mathbf{x}(k))} \\
& - \alpha_V \frac{\tilde{\sigma_V}(\mathbf{x}(k))(\sigma_V^\top(\mathbf{x}(\infty)) - \sigma_V^\top(\hat{\mathbf{x}}(\infty)))\hat{W}_V(k)}{1+\tilde{\sigma_V}^\top(\mathbf{x}(k))\tilde{\sigma_V}(\mathbf{x}(k))} \Big) - \|\tilde{W}_V(k)\|^2 \\
=& -2\alpha_V \Big(\frac{\tilde{\sigma_V}^\top(\mathbf{x}(k))\tilde{\sigma_V}(\mathbf{x}(k))}{1+\tilde{\sigma_V}^\top(\mathbf{x}(k))\tilde{\sigma_V}(\mathbf{x}(k))} \|\tilde{W}_V(k)\|^2 \\
& - \frac{\tilde{W}_V^\top(k)\tilde{\sigma_V}(\mathbf{x}(k))(\sigma_V^\top(\mathbf{x}(\infty)) - \sigma_V^\top(\hat{\mathbf{x}}(\infty)))\hat{W}_V(k)}{1+\tilde{\sigma_V}^\top(\mathbf{x}(k))\tilde{\sigma_V}(\mathbf{x}(k))} \Big) \\
& + \alpha_V^2 \Big\| \frac{\tilde{\sigma_V}(\mathbf{x}(k))\tilde{\sigma_V}^\top(\mathbf{x}(k))}{1+\tilde{\sigma_V}^\top(\mathbf{x}(k))\tilde{\sigma_V}(\mathbf{x}(k))} \tilde{W}_V(k) \\
& - \frac{\tilde{\sigma_V}(\mathbf{x}(k))(\sigma_V^\top(\mathbf{x}(\infty)) - \sigma_V^\top(\hat{\mathbf{x}}(\infty)))\hat{W}_V(k)}{1+\tilde{\sigma_V}^\top(\mathbf{x}(k))\tilde{\sigma_V}(\mathbf{x}(k))} \Big\|^2.
\end{aligned}$$

Since we have $\frac{\tilde{\sigma}_V^\top(\mathbf{x}(k))\tilde{\sigma}_V(\mathbf{x}(k))}{1+\tilde{\sigma}_V^\top(\mathbf{x}(k))\tilde{\sigma}_V(\mathbf{x}(k))} < 1$ and $0 < \tilde{\sigma}_{Vm} \leq \|\tilde{\sigma}_V\| \leq \tilde{\sigma}_{VM}$, by applying Cauchy-Schwartz inequality, the first difference of Lyapunov function is rewritten as

$$\begin{aligned}
\Delta L_V &\leq -2\alpha_V \frac{\tilde{\sigma}_{VM}^2}{1+\tilde{\sigma}_{Vm}^2} \|\tilde{W}_V(k)\|^2 + \alpha_V^2 \|\tilde{W}_V(k)\|^2 \\
&= -\alpha_V \left(\frac{2\tilde{\sigma}_{VM}^2}{1+\tilde{\sigma}_{Vm}^2} - \alpha_V \right) \|\tilde{W}_V(k)\|^2.
\end{aligned} \tag{2.24}$$

Combining Equations (2.19) (2.23) and (2.24), we conclude the asymptotic stability.

□

Remark 5. *This section provides a basic analysis of the system with neural networks. Due to the complexity of the integrated framework, a multi-layer neural network is adopted to generate the control policy in this chapter. The discussion for the multi-layer neural network structure is an extension of the above result by applying the backpropagation (BP) algorithm [126]. We omit the detail of the BP algorithm since it can be found in exiting literature [126].*

2.4 Simulation Studies

In this section, the performance of the proposed method is demonstrated by an example with the Prius as the vehicle model. Our discourse begins with an explanation of vehicle models and simulation setup. In addition, the simulation results are explained, and the effectiveness of the proposed method is verified.

2.4.1 Toyota Prius Hybrid Vehicle Modeling

Although the complete and accurate mathematical model of the overall system is hard to obtain, the update rule of agent states should follow the vehicle dynamics and constraints of subsystems. In this section, the models of the agent are provided in the following. Generally speaking, a PHEV can be characterized by five sub-systems [1]: (1) Vehicle dynamics; (2) Engine fuel consumption model; (3) Electric consumption of the EM; (4) Transmission; and (5) Battery model. Specifically, the Toyota Prius is selected as the studied car. Since this car is equipped with a series-parallel hybrid system, it is hard to describe the model of its powertrain exactly and explicitly. In the following, we will briefly introduce the mathematical model of each subsystem. Figure 2.3 illustrates the schematic diagram of the five sub-systems.

Vehicle Dynamics

The vehicle dynamics for PHEVs mainly focus on PHEVs' longitudinal dynamics [1], which is illustrated in Figure 2.4.

The output torque T_w is obtained by the following equation

$$\begin{aligned} T_w &= \eta_t \cdot i_f (T_e + T_m) + T_B \\ &= [Mgf_r \cos\theta + \frac{1}{2} C_D \rho_d A_f v^2 + Mg \sin\theta + \delta M \frac{dv}{dt}] \cdot r_w, \end{aligned} \tag{2.25}$$

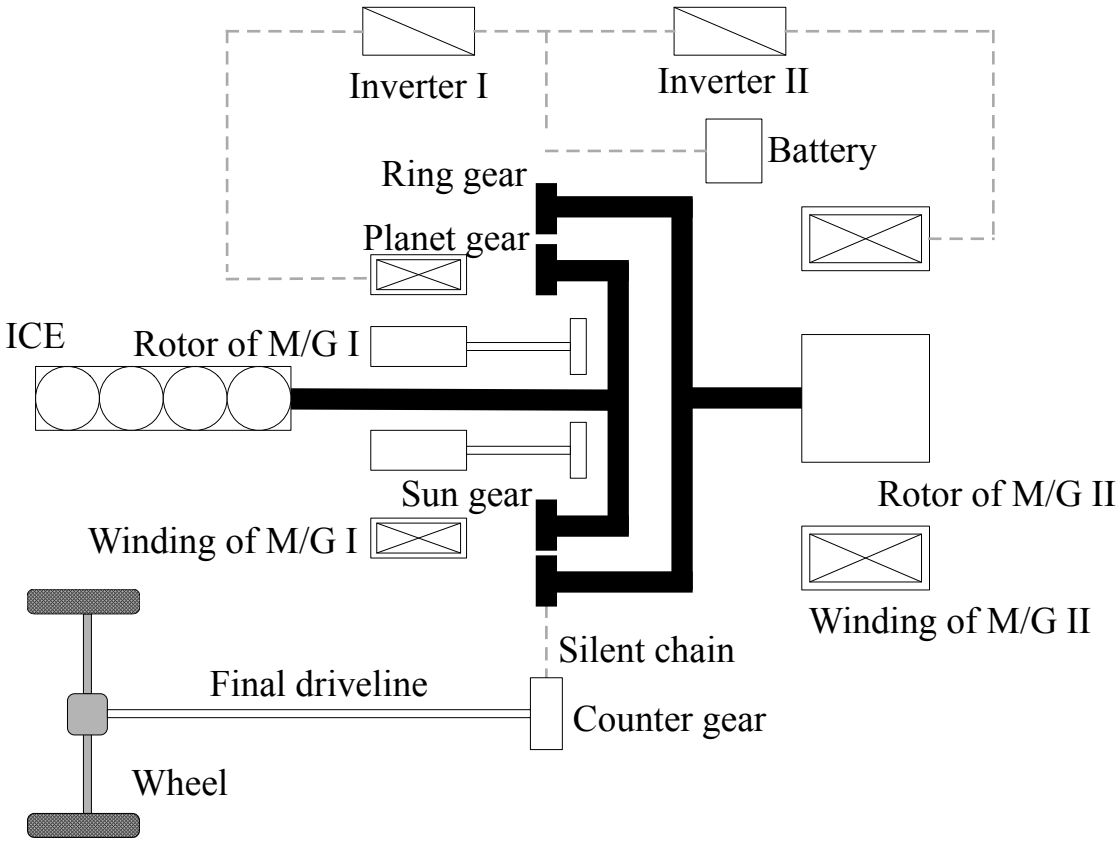


Figure 2.3: The schematic diagram of the Prius (series-parallel PHEVs) [2].

where η_t represents the transmission efficiency. T_e, T_m and T_B are the torque of engine, the torque of EM, and the torque of braking, respectively. i_f denotes the basic ratio of planetary gear in Prius, which is a fixed ratio and equals the ring radius divided by the sun gear radius.

The v denotes the vehicle speed. The parameters f_r, g and θ capture the operation condition of the vehicle and denote the rolling resistance coefficient, gravity acceleration, and road slope angle, respectively. $M = M_o + M_p$ denotes the equivalent vehicle mass including the vehicle mass (i.e., M_o) and the passenger load (i.e., M_p). The parameters of the vehicle, r_w, C_D, ρ_d and A_f, represent wheel radius, air drag coefficient, air density and frontal area, respectively.

Fuel Consumption Model of Engine

Fuel economy is the most important assessment index of the power management strategy for PHEVs. Using fossil fuel engines, the fuel consumption rate is a nonlinear function

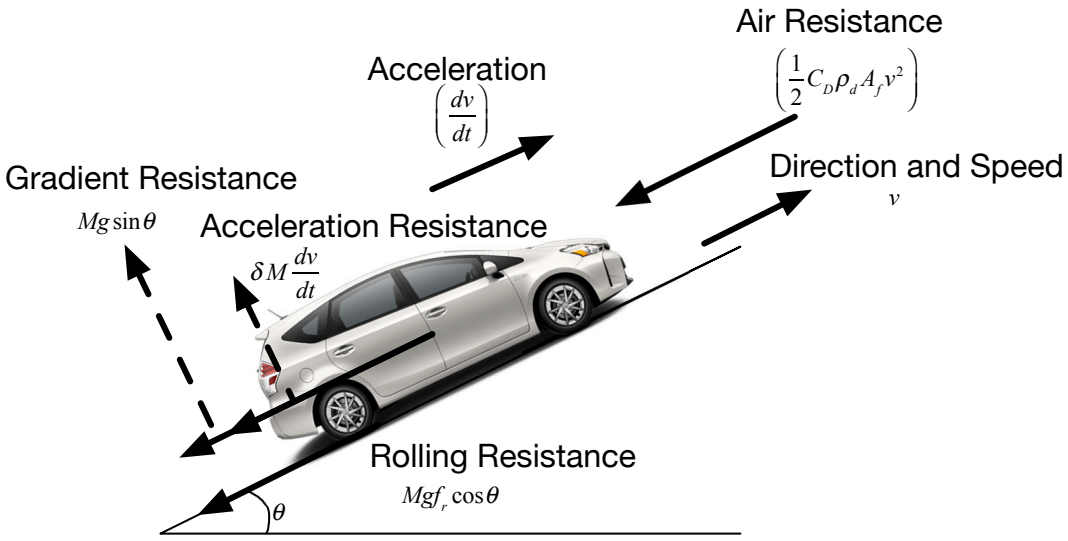


Figure 2.4: The driving model of the PHEVs.

of the speed and power of the engine. The fuel consumption rate is obtained by searching the consumption contours of the engine, which depends on the type of fuel and the type of engine.

The corresponding ICE efficiency is shown by the labeled number of each contour. As an example, it is a 1.5L S.I. Engine modeled by the advanced vehicle simulator ADVISOR [14], illustrated in Figure 2.5. The engine has a peak power of 43kW and a peak efficiency of 39%.

Electric Consumption Model of EM

EM works under two operating modes, the driving mode and the braking mode. The output power of EM is described as follows.

$$P_m = \begin{cases} T_m\omega_m/\eta_m, & \text{Driving} \\ T_m\omega_m\eta_m, & \text{Braking} \end{cases} \tag{2.26}$$

The above model shows that the EM output power P_m depends on the EM torque T_m, the EM speed ω_m, and the EM efficiency η_m. The efficiency η_m is obtained by searching the EM efficiency map that depends on the type of EM. As an example, the colourful contours in Figure 2.6 illustrate the EM efficiency of a 15 kW permanent magnet motor.

Toyota Prius employs an Electronically Controlled Continuously Variable Transmission

(ECVT) system, which plays a significant role in series-parallel PHEVs [2]. In PHEVs, the ECVT system is designed for ensuring the optimal operating modes of the engine and the motor related to the speed variations of the vehicles. In the simulation, the gear-shifting signal $s(k)$ is considered in the analysis of the energy consumption.

Battery Model

SoC is the main index to describe the battery operation. According to Kirchhoff's voltage law, considering the change rate of the SoC and the power of battery for the rate of change of SoC P_{batt}, SoC of next state is calculated as follows.

$$SoC(k+1) = SoC(k) - \frac{U_{oc} - \sqrt{U_{oc}^2 - 4P_{batt}R_{int}}}{2Q_{max}R_{int}}, \tag{2.27}$$

where U_{oc}, R_{int} and Q_{max} are the open-circuit voltage, the internal resistance and the maximum value of battery capacity, respectively.

Due to the difference between two operation modes of battery (i.e., charging and discharging), the internal resistance R_{int} can be described as

$$R_{int} = \begin{cases} R_{dischg}, & \text{Discharging} \\ R_{chg}, & \text{Charging} \end{cases} \tag{2.28}$$

where R_{dischg} and R_{chg} represent the internal resistances of the battery which indicate the battery is on discharging or charging mode, respectively. Values of the parameters related to vehicle dynamics are illustrated in Table 2.1.

2.4.2 Reinforcement Learning Simulations

In this subsection, the solution to the integrated problem using RL is proposed to find the minimal energy consumption route while taking power management into account. Using the RL algorithm, the environment is assumed to be a given city map with roads and geographic information, as shown in Figure 2.2. In the environment, the source and the destination are given, which are linked by a series of road segments. Between the two points, it is recommended to include various speed limits and road slope angles to demonstrate the effectiveness of the proposed method.

The action vector $\mathbf{a}(k)$ includes the desired output torque from engine $T_{e,des}(k)$, gear shift command of the transmission system $s(k)$ and the direction $direc(k)$. There are four

directions for the agent (vehicle) from the current position $p(k)$ to the next position $p(k+1)$ on the road map. The solution to the integrated problem using the RL algorithm is the instruction sequence with the form: [$p(A)$, $\mathbf{a}(1)$ (from A to 1), $p(1)$], [$p(1)$, $\mathbf{a}(2)$ (from 1 to 2), $p(2)$], $\cdots$, [$p(n)$, $\mathbf{a}(n)$ (from n to B), $p(B)$].

To control the computational overhead, we reduce the size of system states by only considering the road intersections in the simulation, i.e., actions are taken at the road intersections. In the training process, the velocity of the vehicle for each road segment is an average value which is calculated based on the length and height of each road segment. Since the velocity and slope angle are constants for each road segment, the required power for each road segment is computed based on (2.25). As a result, the power management process is simplified, and the computational load is reduced.

In the simulation, the training process is not affected by the battery performance. We use the SoC of the battery from 90% to 20% in this simulation. The SoC is calculated by the ratio between the current and total energy. We do not consider battery degradation, which unnecessarily complicates the simulation and may introduce too much noise in the simulation results.

The EM may work under two conditions: 1) providing the power to guarantee that the engine is working with high efficiency; 2) working as a generator to store the electric energy to the battery with the influence of inertia. The required power of EM is calculated by (2.26) with different working conditions. The SoC of next state is computed by (2.27) using $P_{batt} = P_m$.

The neural network is designed as a four-layer structure. The first layer includes the current state $\mathbf{x}(k)$, which is followed by two hidden layers with 10 neurons and 6 neurons, respectively. $\sigma_{\mathbf{a}}$ and σ_V are two ReLU activation functions with dropout rate 0.25. The last layer represents the Q-value for each direction of action $\mathbf{a}(k)$.

The main idea of power management for PHEVs is to avoid ICE operation points in the low-efficiency region. The efficiency contours of engine and EM are illustrated in Figure 2.5 and Figure 2.6, respectively. Also, Figure 2.5 and Figure 2.6 show that most corresponding operating points fall in the high efficiency region under different driving conditions and varying speeds.

Figure 2.7 shows the energy usage in different subsystems under power and regeneration modes. Each action is evaluated based on the required energy to travel on each road segment on the map (Figure 2.8). Based on the vehicle dynamics provided in Section 2.4.1, the required power is computed for the car to travel on the selected roads (using information regarding elevation, the road slope angle θ). The energy consumption is calculated by

Table 2.1: The values of vehicle dynamics for the simulation setup [1].

Parameters	Values
Calculated Mass M	$1332kg$
Coefficient of rolling resistance f_r	0.02
Air density ρ_d	$1.225kg/m^3$
Frontal area A_f	$2m^2$
Air drag coefficient C_D	0.5
Range of SoC	$90\% \sim 20\%$
α_V	0.9
$\alpha_{\mathbf{a}}$	0.0001
ϵ	0.9
U_{oc}	$144V$
R_{dischg}	0.03
R_{chg}	0.02
γ	0.9

multiplying the power with the travel time on the corresponding road segment. Figure 2.8 shows the route planning results, where the red route shows the most energy-efficient path and the green route is the shortest path which has a higher fuel consumption than the red route. The effectiveness of doing power management and route planning at the same time is verified.

2.5 Conclusion

In this chapter, we proposed a route planning algorithm using RL with the help of road geographical information. Our goal is to find the path that minimizes the energy usage of PHEVs. A two-loop structure is adopted to integrate both path planning and energy management. To be specific, the route of the minimum energy consumption is obtained in the outer loop, and the energy consumption index with respect to PHEV power management for each feasible route is trained by the RL framework in the inner-loop. To handle the power management for PHEVs using RL, the nonlinear approximator structure is built to approximate control actions and values by an actor approximator and a critic approximator, respectively. We also demonstrate the convergence of the value function for power management in the inner loop and asymptotic stability of the closed-loop system. Finally, simulation results show that the proposed algorithm provides an effective solution to the problem of routing PHEVs with the energy-efficient path.

Further investigations are needed in the future to reduce the computational overhead of

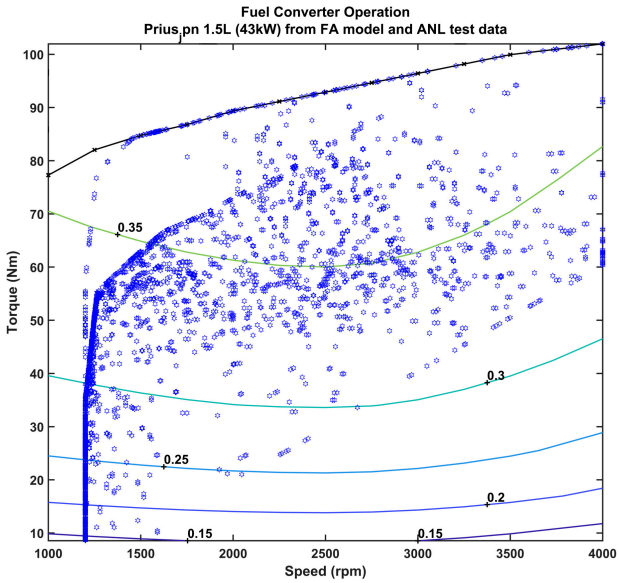


Figure 2.5: Fuel converter efficiency. The colourful contours with numbers illustrate the consumption contours of engine. The blue points represent the work points. The black contour shows the maximum value of fuel converter efficiency.

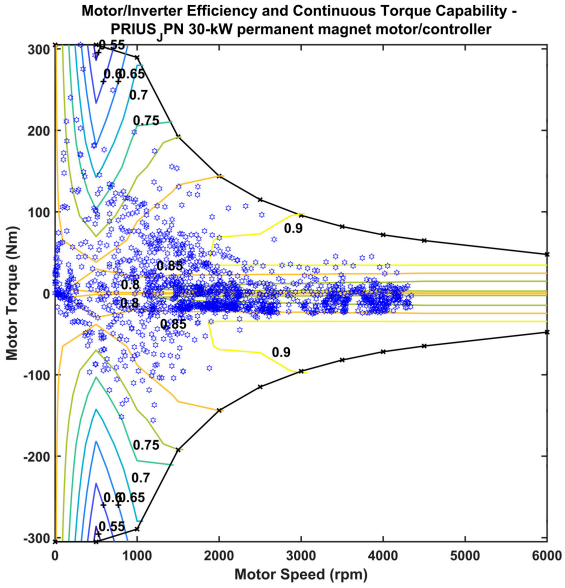


Figure 2.6: Motor controller efficiency. The colourful contours with numbers illustrate the consumption contours of electric motor. The blue points represent the work points. The two black contours show the maximum value of motor controller efficiency.

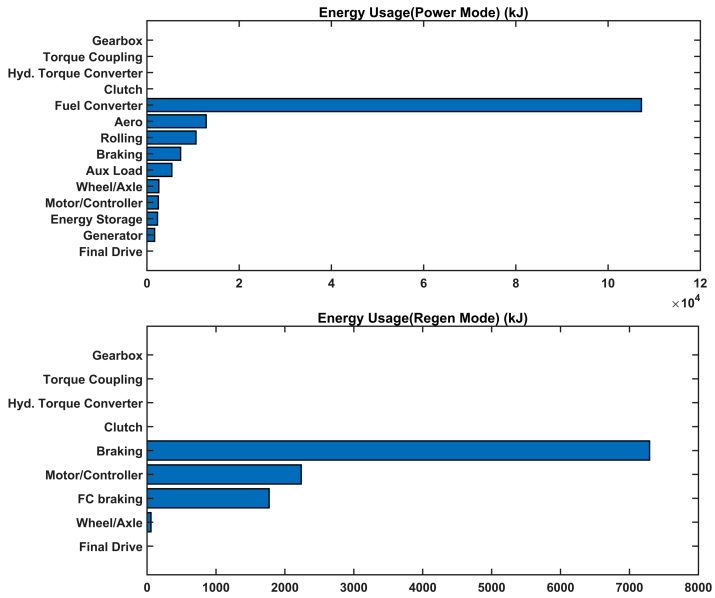


Figure 2.7: Energy usage under power and regen modes, respectively. The blue bars show the energy usages in different subsystems.

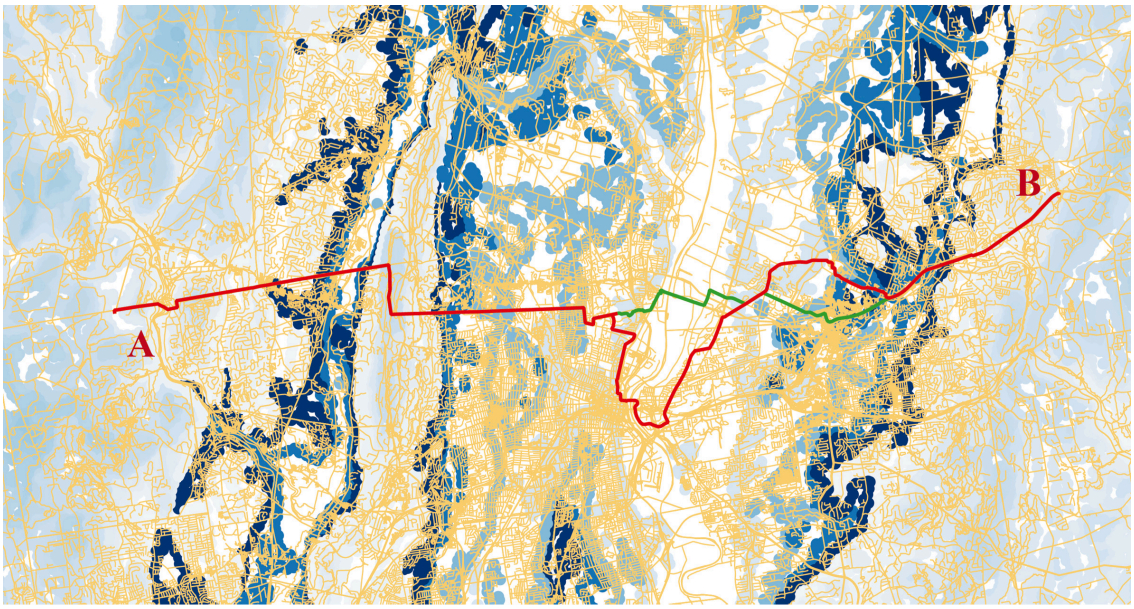


Figure 2.8: Route planning results. Background colors indicate the elevation ranging from 0m as white to 800m as deep blue. The yellow lines show the road networks. The red line illustrates the route produced by the proposed method with the fuel consumption of 134263 KJ, while the green line represents the shortest path with the fuel consumption of 215715 KJ. Two paths share the same source A directing eastward and ending at the destination B.

the RL algorithm. Searching for the optimal solution for the power management problem in a continuous state and the action spaces is also an interesting direction for future research.

Chapter 3

Uncertain PEV Charging Control: A Self-Triggered Robust MPC with ISM Method

3.1 Introduction

PEVs are strongly promoted and widely used owing to environmental protection and energy-saving considerations. In this context, the PEV charging strategy becomes an urgent research problem. In the following three chapters, we focus on the PEV charging control schemes. As discussed in Section 1.1.2, some practical issues need to be considered under different PEV charging architectures. For instance, the centralized PEV charging strategy will inevitably be subject to heavy communication loads because of the large-scale vehicle population. Also, there is a high possibility of failure in the entire system caused by uncertainties during the transmission. Among these issues, the self-triggered robust MPC with ISM strategy could be considered. This chapter aims to allow a trade-off between the performance and the reduction of communication for PEV charging under centralized architecture.

Event-triggered control balances the control performance and the communication resource usage [127]. The core idea of the event-triggered control scheme lies in that the sensor measurement and control signals are only exchanged between actuators, sensors and controllers when a particular event occurs [128–130]. Events are generally related to closed-loop system behaviors; for instance, the gap between the actual and predicted system state, due to the disturbance, violates a specific bound. The event-triggered condition

is based on the measurements which are monitored at each time instant. The system states and control inputs will be updated when the event-triggered condition is violated. There are several limitations when the event-triggered control applies to practical energy problems. For example, the uncertainty of PEVs remains high due to their mobility in transportation sectors. Event-triggered scheduling schemes to address the PEV charging problem are studied for handling power demand variations caused by PEVs' random arrivals and departures [61, 131].

Self-triggered control is a particular type of event-triggered control [132, 133], where the next execution time is precomputed at the current control execution time by using a self-triggered control scheme. Also, the next execution time depends on the stored measurements, predicted system states. The self-triggered control scheme naturally leads to a significant reduction of communication loads, thus making it a suitable candidate for the PEV charging control problem.

MPC has been proved to be an efficient control method for the plant with the input and state constraints [134]. Compared with other control methods, MPC provides recurrent online solutions for optimal control problems by moving predictive windows, which guarantees transient behaviors of closed-loop systems [86, 135]. From the practical point of view, systems are always influenced by uncertainties due to model mismatches or noises. Hence, researchers have proposed various approaches for robust MPC in the last decade [136, 137]. One of the challenges of designing a control algorithm for practical energy problems is dealing with model uncertainties and satisfying the stability of closed-loop systems [138, 139]. To tackle the problem, [140] studies a scheduling scheme of PEV charging under MPC framework, and shows a superior performance with robustness against uncertainties.

Due to the communication overheads of PEV charging control under centralized architecture, the self-triggered MPC strategy for PEV charging problems shows its efficiency for reducing the frequency of transmission [141]. In particular, a self-triggered MPC allows a trade-off between the closed-loop system performance and communication loads [142, 143]. Furthermore, a self-triggered MPC algorithm for nonlinear input-affine systems using adaptive sampling selection is investigated in [144], but the effects of disturbances and state constraints are not considered. However, under the centralized architecture, any noise or uncertainty happens during the transmission causing a high possibility of failure in the entire system since the charging rates of all PEVs is decided by the central aggregator. It is an urgent issue for PEV charging under the centralized architecture. A possible solution to this problem is using MPC integrated with other control methods.

Compared with different methods, ISM control strategy shows its efficiency for the systems with unmatched uncertainties [145], especially to address the battery charging problem of EVs [146].

Based on the aforementioned discussion, we propose a self-triggered MPC scheme with integrating ISM strategy to address the PEV charging problem with state and input constraints and uncertainties. The proposed controller solves an online finite horizon optimal control problem to obtain the optimal control sequence for the following time interval and the next execution time. The next execution time can be denoted as the time when the triggered condition is satisfied.

The remainder of this chapter is organized as follows. Section 3.2 formulate the problem of this work mathematically, followed by the objective of the proposed algorithm. In Section 3.3, the control strategies to achieve the desired objectives are discussed, including the setup of MPC framework, the self-triggered scheme and ISM component design. Section 3.4 contains a numerical example using nonlinear dynamics and an example using single PEV charging model to verify the theoretical results. The conclusion of this paper is given in Section 3.5.

3.2 Problem Formulation

Without loss of generality, we consider a nonlinear affine continuous-time system with the additive disturbance modeled by the following dynamics:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t)) + \mathbf{g}(\mathbf{x}(t))\mathbf{u}(t) + \omega(\mathbf{x}(t)), \quad t \geq 0, \quad \mathbf{x}(0) = \mathbf{x}_0, \tag{3.1}$$

where $\mathbf{x}(t) \in \mathbb{R}^n$ is the system state constrained by $\mathbf{x}(t) \in \mathbb{X} \subset \mathbb{R}^n$, $\mathbf{u}(t) \in \mathbb{R}^n$ denotes the control input constrained by $\mathbf{u}(t) \in \mathbb{U} \subset \mathbb{R}^n$, and $\omega(\mathbf{x}(t)) \in \mathbb{W} \subset \mathbb{R}^n$ is the disturbance. $\mathbb{U}$ and $\mathbb{W}$ are compact sets, containing the origin as an interior point. $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ is a Lipschitz continuous function that satisfies $\mathbf{f}(\mathbf{0}) = \mathbf{0}$, and the Lipschitz constant is $L \in (0, \infty)$. $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^n$ is a continuous function that satisfies $\mathbf{g}(\mathbf{0}) = \mathbf{0}$. Note that $\omega(\mathbf{x}(t))$ denotes the perturbation function with for the following expression

$$\omega(\mathbf{x}(t)) = \mathbf{g}(\mathbf{x}(t))\omega_m(t) + \mathbf{g}^{\perp}(\mathbf{x}(t))\omega_u(t), \tag{3.2}$$

where $\mathbf{g}^{\perp}(\mathbf{x}(t))$ is the orthogonal complement matrix of $\mathbf{g}(\mathbf{x}(t))$. $\omega_m(t)$ and $\omega_u(t)$ denote the matched and unmatched uncertainties, respectively. To guarantee the stability of the

closed-loop system, we assume that the function $\omega(\mathbf{x}(t))$ is bounded by a known positive scalar function $\rho(\mathbf{x}(t))$., i.e. $|\omega(\mathbf{x}(t))| \leq \rho(\mathbf{x}(t))$.

Remark 6. *The dynamics with the input-affine form* (3.1) *is required to obtain the explicit control law of the ISM controller and MPC controller, as clarified in Section 3.3.3 and Section 3.3.1, respectively.*

This chapter aims to design a control algorithm for PEV charging control under the centralized architecture to reduce communication loads and suppress matched uncertainties. We adopt an MPC framework for PEV charging control problem, supplemented by a self-triggered strategy to reduce the high communicational overheads and an ISM method to inhibit the matched uncertainties.

3.3 Control Strategies

In order to achieve the desired objectives, we propose a novel integration of MPC and ISM in this section. In the integrated MPC and ISM scheme, MPC is employed to fulfill the robust constraint satisfaction, and ISM is used to handle the uncertainties. The MPC controller will generate the control input for the nominal system with state and input constraints when receiving the new state measurements. The ISM controller is used to suppress the matched uncertainties. The compound control input is the sum of two control inputs from these two controllers. However, one challenge of implementing the compound controller lies in that the two controllers might be asynchronous. Under the self-triggered control scheme, the execution time of control inputs from the MPC controller can be calculated. The control input at the calculated execution time can be described as (3.3) and applied to the system (3.1).

$$\mathbf{u}(t) = \mathbf{u}^{MPC}(t) + \mathbf{u}^{ISM}(t), \tag{3.3}$$

where $\mathbf{u}^{MPC}(t)$ denotes the control input generated by the MPC controller and $\mathbf{u}^{ISM}(t)$ represents the control input from the ISM controller. The structure of the closed-loop system is shown in Figure 3.1.

3.3.1 MPC Problem Setup

For the system (3.1), take the sequence $\{t_k\}$, $k \in \mathbb{N}$, as the sampling time instants. At time t_k, let $\tilde{\mathbf{x}}(s; t_k)$ denote the predicted state trajectory satisfying the system dynamics

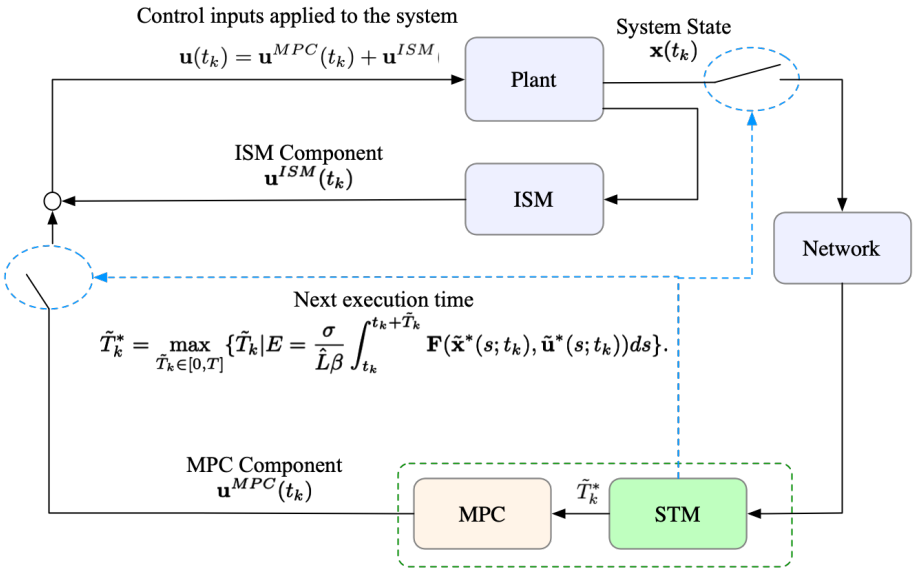


Figure 3.1: The structure of the closed-loop system.

(3.1), $s \in [t_k, t_k + T]$, where $T > 0$ is the prediction horizon. The predicted control input trajectory $\tilde{\mathbf{u}}(s; t_k)$ calculated by MPC is obtained by solving the finite horizon optimal control problem at time t_k corresponding to the predicted state trajectory $\tilde{\mathbf{x}}(s; t_k)$, $s \in [t_k, t_k + T]$. Let $\tilde{\mathbf{u}}^*(s; t_k)$, $s \in [t_k, t_k + T]$ denote the optimal control input trajectory that is the minimizer of the optimal control problem. Also, for every initial state, the optimal control input trajectory is unique. We define $\tilde{T}_k \in [0, T]$ as a particular time interval that is a finite prediction horizon. To design the control algorithm for PEV charging problem, $\tilde{T}_k$ can be chosen as an integer multiple of the sampling interval of power demand's baseline. At time $t_k + \tilde{T}_k$, the predicted state enters an invariant set $\Omega(\epsilon_T)$, which contains the origin as an interior point.

To design the optimal control problem, we define a function $\mathbf{F}(\tilde{\mathbf{x}}(s; t_k), \tilde{\mathbf{u}}(s; t_k)) = \|\tilde{\mathbf{x}}(s; t_k)\|_Q^2 + \|\tilde{\mathbf{u}}(s; t_k)\|_R^2$. The finite horizon cost function can be described as

$$\begin{aligned}\mathbf{F}_T(\tilde{\mathbf{x}}(s; t_k), \tilde{\mathbf{u}}(s; t_k), \tilde{T}_k) = &\int_{t_k}^{t_k+\tilde{T}_k} \frac{1}{\beta} \mathbf{F}(\tilde{\mathbf{x}}(s; t_k), \tilde{\mathbf{u}}(s; t_k)) \mathrm{d}s \\ &+ \int_{t_k+\tilde{T}_k}^{t_k+T} \mathbf{F}(\tilde{\mathbf{x}}(s; t_k), \tilde{\mathbf{u}}(s; t_k)) \mathrm{d}s,\end{aligned}$$

where $\beta \geq 1$ is a fixed constant, and the terminal cost function is defined by

$$\mathbf{V}_T(\tilde{\mathbf{x}}(t_k + T; t_k)) = \|\tilde{\mathbf{x}}(t_k + T; t_k)\|_P^2.$$

The cost function to be minimized in the optimal control problem is described in the following,

$$\mathbf{J}(\tilde{\mathbf{x}}(s;t_k),\tilde{\mathbf{u}}(s;t_k),\tilde{T}_k)=\mathbf{F}_T(\tilde{\mathbf{x}}(s;t_k),\tilde{\mathbf{u}}(s;t_k),\tilde{T}_k)+\mathbf{V}_T(\tilde{\mathbf{x}}((t_k+T;t_k)), \tag{3.4}$$

where Q, R and P are symmetric weighting matrices. Using the setup above, the optimal control problem can be defined as follows.

Problem 3. *(Optimal Control Problem)*

$$\begin{aligned}\tilde{\mathbf{u}}^*(s;t_k) &= \arg\min \mathbf{J}(\tilde{\mathbf{x}}(s;t_k),\tilde{u}(s;t_k),\tilde{T}_k)\\ s.t\quad & \dot{\tilde{x}}(s;t_k)=f(\tilde{x}(s;t_k))+g(\tilde{x}(s;t_k))\tilde{u}(s;t_k)),\quad \tilde{\mathbf{x}}(t_k;t_k)=\mathbf{x}(t_k)\\ & \tilde{\mathbf{u}}(s;t_k)\in\mathbb{U}, s\in[t_k,t_k+T]\\ & \tilde{x}(t_k+T;t_k)\in\Omega(\epsilon), s\in[t_k,t_k+T]\\ & \tilde{\mathbf{x}}(s;t_k)\in\mathbb{X}_{s-t_k}, s\in[t_k,t_k+T]\end{aligned} \tag{3.5}$$

where $\Omega(\epsilon)$ is the terminal set, and $\mathbb{X}_{s-t_k}$ is the tightened state constraint. $\mathbf{J}(\tilde{\mathbf{x}}(s;t_k),\tilde{u}(s;t_k),\tilde{T}_k)$ denotes the corresponding optimal value function and $\tilde{\mathbf{x}}^*(s;t_k)$ is the optimal state trajectory.

For the given constant ϵ, the terminal set $\Omega(\epsilon)$ is given by $\Omega(\epsilon)=\{\mathbf{x}(t)\in\mathbb{X}|\|\mathbf{x}(t)\|_P^2\leq\epsilon^2\}$. In order to satisfy the system state constraints, the tightened state constraints in [147, 148] can be recalled and applied here. The tightened state constraint is described as $\mathbb{X}_{s-t_k}=\mathbb{X}\sim\mathbb{T}_{s-t_k}$, where $\mathbb{T}_{s-t_k}=\{\mathbf{x}(t_k)\in\mathbb{R}^N|\|\mathbf{x}(t)\|_P\leq\rho\bar{\lambda}(\sqrt{P})(e^{L(s-t_k)}-1)/L\}$. We assume that $\Omega(\epsilon)\subseteqq\mathbb{X}_{s-t_k}$. The controller applies optimal control input $\tilde{\mathbf{u}}^*(s;t_k)$, $s\in[t_k,t_{k+1}]$ to the plant, where t_{k+1} denotes the next execution time obtained by the proposed self-triggered scheduler to be introduced in the following.

Assumption 1. *Given $\mathbf{f}(\mathbf{0})=\mathbf{0}$, the nonlinear system can be linearized as:*

$$\dot{\mathbf{x}}(t)=A\mathbf{x}(t)+B\mathbf{u}(t)+\omega(t), \tag{3.6}$$

where $A=\frac{\partial\mathbf{f}}{\partial\mathbf{x}}|_{(0,0)}$, $B=\frac{\partial\mathbf{g}}{\partial\mathbf{x}}|_{(0,0)}$, and the pair (A,B) is stabilizable.

Furthermore, the closed-loop system with a state-feedback control law $\dot{\hat{\mathbf{x}}}(t)=\hat{A}\hat{\mathbf{x}}(t)+\omega(t)$, where $\hat{A}=A+BK$ is asymptotically stable. In addition, a well-known result can be recalled in Lemma 1 of [141, 149].

Lemma 3. *Given two positive-definite matrices Q and R for the nominal system (3.1), with Assumption 1, there exists a state-feedback control law K, a symmetric matrix P and two constants ϵ_T and ξ such that:*

- *The inequalities $\dot{V}_L(\mathbf{x}) \leq -\frac{1}{2}\|\mathbf{x}\|^2_{Q^*}$ and $\dot{V}_L(\mathbf{x}) \leq -\xi V_L(\mathbf{x})$ hold for any $\mathbf{x}(t) \in \Omega(\epsilon_T)$ and $K\mathbf{x}(t) \in \mathbb{U}$, where $Q^* = Q + K^\top RK$ and $V_L(\mathbf{x}) = \frac{1}{2}\|\mathbf{x}\|^2_P$. The constant ξ is chosen to be $\frac{\lambda(Q^*)}{2\lambda(P)}$.*
- *The set $\Omega(\epsilon_T)$ is an invariant set, which is defined as $\Omega(\epsilon_T) = \{\mathbf{x} \in \mathbb{R}^n | V_L(\mathbf{x}) \leq \frac{1}{2}\epsilon_T^2\}$ for the nonlinear system under the local state-feedback control law with no disturbance $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t)) + \mathbf{g}(\mathbf{x}(t))K\mathbf{x}(t)$.*

Consequently, $\Omega(\epsilon_T)$ is the region of attraction for both the closed-loop system $\dot{\hat{\mathbf{x}}}(t) = \hat{A}\hat{\mathbf{x}}(t) + \omega(t)$ and the nonlinear system (3.1). To guarantee the recursive feasibility and stability of the closed-loop system, the following two inequalities are satisfied. There exist two κ_∞ - functions, α_1 and α_2, such that

$$\begin{aligned} \mathbf{F}(\mathbf{x}(t), \mathbf{u}(t)) &\geq \alpha_1(\|\mathbf{x}(t)\|), \\ \mathbf{V}_T(\mathbf{x}(t)) &\leq \alpha_2(\|\mathbf{x}(t)\|), \end{aligned} \tag{3.7}$$

for all $\mathbf{x}(t) \in \mathbb{X}$, $\mathbf{u}(t) \in \mathbb{U}$ and $\tilde{T}_k \in [0, T]$. Also, $\mathbf{F}(\mathbf{x}(t), \mathbf{u}(t))$ and $\mathbf{V}_T(\mathbf{x}(t))$ are Lipschitz continuous functions, where $\mathbf{F}(\mathbf{0}, \mathbf{0}) = \mathbf{0}$ and $\mathbf{V}_T(\mathbf{0}) = \mathbf{0}$. The Lipschitz constants are $L_1 \in (0, \infty)$ and $L_2 \in (0, \infty)$, respectively.

To show the relationship between the invariant set $\Omega(\epsilon_T)$ and the terminal constraint set $\Omega(\epsilon)$, we further summarize the following result based on Lemma 3 in [149].

Lemma 4. *For the given constants $\epsilon_T > \epsilon$ and local state-feedback control law $\tilde{\mathbf{u}}(s; t_k)$, satisfying $\tilde{\mathbf{u}}(s; t_k) \in \mathbb{U}$, $s \in [t_k, t_k + \tilde{T}_k]$, the following inequality holds for all $\tilde{\mathbf{x}}(s; t_k) \in \Omega(\epsilon_T)$*

$$\mathbf{J}^*(\tilde{\mathbf{x}}(t_k + \tilde{T}_k; t_k), \tilde{T}_k) - \mathbf{J}^*(\tilde{\mathbf{x}}(t_k; t_k), \tilde{T}_k) \leq -\int_{t_k}^{t_k + \tilde{T}_k} \frac{1}{\beta}\mathbf{F}(\tilde{\mathbf{x}}^*(s; t_k), \tilde{\mathbf{u}}^*(s; t_k))ds. \tag{3.8}$$

Proof. At time t_k, we have the optimal cost $\mathbf{J}^*(\tilde{\mathbf{x}}(s; t_k), \tilde{T}_k)$ by solving the optimal control problem, with the optimal control trajectory $\tilde{\mathbf{u}}^*(s; t_k)$ and the corresponding optimal state

trajectory $\tilde{\mathbf{x}}^*(s;t_k)$. The optimal cost $\mathbf{J}^*(\tilde{\mathbf{x}}(t_k;t_k),\tilde{T}_k)$ can be represented as

$$\begin{aligned}
&\mathbf{J}^*(\tilde{\mathbf{x}}(t_k;t_k),\tilde{T}_k) = \mathbf{F}_T(\tilde{\mathbf{x}}^*(s;t_k),\tilde{\mathbf{u}}^*(s;t_k),\tilde{T}_k) + \mathbf{V}_T(\tilde{\mathbf{x}}^*(t_k+T;t_k)) \\
=&\int_{t_k}^{t_k+\tilde{T}_k} \frac{1}{\beta}\mathbf{F}(\tilde{\mathbf{x}}^*(s;t_k),\tilde{\mathbf{u}}^*(s;t_k))\mathrm{d}s + \int_{t_k+\tilde{T}_k}^{t_k+T} \mathbf{F}(\tilde{\mathbf{x}}^*(s;t_k),\tilde{\mathbf{u}}^*(s;t_k))\mathrm{d}s \\
&+ \|\tilde{\mathbf{x}}^*(t_k+T;t_k)\|_P^2.
\end{aligned} \tag{3.9}$$

In addition, we define a predicted cost $\hat{\mathbf{J}}(\tilde{\mathbf{x}}(t_k+\tilde{T}_k),\tilde{T}_k)$ by applying the feasible control input trajectory $\tilde{\mathbf{u}}(s;t_k)$ for all $[t_k+T, t_k+\tilde{T}_k+T]$ in the following

$$\begin{aligned}
\hat{\mathbf{J}}(\tilde{\mathbf{x}}(t_k+\tilde{T}_k),\tilde{T}_k) &= \int_{t_k+\tilde{T}_k}^{t_k+T} \mathbf{F}(\tilde{\mathbf{x}}(s;t_k),\tilde{\mathbf{u}}(s;t_k))\mathrm{d}s + \|\tilde{\mathbf{x}}(t_k+\tilde{T}_k+T;t_k)\|_P^2 \\
&= \int_{t_k+\tilde{T}_k}^{t_k+T} \mathbf{F}(\tilde{\mathbf{x}}^*(s;t_k),\tilde{\mathbf{u}}^*(s;t_k))\mathrm{d}s \\
&\quad + \int_{t_k+T}^{t_k+\tilde{T}_k+T} \|\tilde{\mathbf{x}}(s;t_k)\|_{Q^*}^2\mathrm{d}s + \|\tilde{\mathbf{x}}(t_k+\tilde{T}_k+T;t_k)\|_P^2.
\end{aligned} \tag{3.10}$$

At time t_k, the optimal control problem is assumed to be feasible. Using Lemma 3, the following inequality holds

$$\|\tilde{\mathbf{x}}(t_k+\tilde{T}_k+T;t_k)\|_P^2 - \|\tilde{\mathbf{x}}^*(t_k+T;t_k)\|_P^2 \leq \int_{t_k+T}^{t_k+\tilde{T}_k+T} \|\tilde{\mathbf{x}}(s;t_k)\|_{Q^*}^2\mathrm{d}s. \tag{3.11}$$

Consequently, the predicted cost $\hat{\mathbf{J}}(\tilde{\mathbf{x}}(t_k+\tilde{T}_k),\tilde{T}_k)$ can be derived as

$$\hat{\mathbf{J}}(\tilde{\mathbf{x}}(t_k+\tilde{T}_k),\tilde{T}_k) \leq \int_{t_k+\tilde{T}_k}^{t_k+T} \mathbf{F}(\tilde{\mathbf{x}}^*(s;t_k),\tilde{\mathbf{u}}^*(s;t_k))\mathrm{d}s + \|\tilde{\mathbf{x}}^*(t_k+T;t_k)\|_P^2. \tag{3.12}$$

Considering the difference between the optimal costs at time t_k and $t_k+\tilde{T}_k$, the following inequalities hold

$$\begin{aligned}
&\mathbf{J}^*(\tilde{\mathbf{x}}(t_k+\tilde{T}_k;t_k),\tilde{T}_k) - \mathbf{J}^*(\tilde{\mathbf{x}}(t_k;t_k),\tilde{T}_k) \\
\leq&\hat{\mathbf{J}}(\tilde{\mathbf{x}}(t_k+\tilde{T}_k),\tilde{T}_k) - \mathbf{J}^*(\tilde{\mathbf{x}}(t_k;t_k),\tilde{T}_k) \\
\leq&\int_{t_k+\tilde{T}_k}^{t_k+T} \mathbf{F}(\tilde{\mathbf{x}}^*(s;t_k),\tilde{\mathbf{u}}^*(s;t_k))\mathrm{d}s + \|\tilde{\mathbf{x}}^*(t_k+T;t_k)\|_P^2 \\
&- \int_{t_k}^{t_k+\tilde{T}_k} \frac{1}{\beta}\mathbf{F}(\tilde{\mathbf{x}}^*(s;t_k),\tilde{\mathbf{u}}^*(s;t_k))\mathrm{d}s + \int_{t_k+\tilde{T}_k}^{t_k+T} \mathbf{F}(\tilde{\mathbf{x}}^*(s;t_k),\tilde{\mathbf{u}}^*(s;t_k))\mathrm{d}s
\end{aligned}$$

$$
\begin{aligned}
&+ \|\tilde{\mathbf{x}}^*(t_k + T; t_k)\|_P^2 \\
\leq &- \int_{t_k}^{t_k+\tilde{T}_k} \frac{1}{\beta}\mathbf{F}(\tilde{\mathbf{x}}^*(s; t_k), \tilde{\mathbf{u}}^*(s; t_k))\mathrm{d}s.
\end{aligned}
$$

The proof is complete. □

According to the definitions of $\Omega(\epsilon)$ and $\Omega(\epsilon_T)$, the terminal set $\Omega(\epsilon)$ is smaller than the set $\Omega(\epsilon_T)$ in Lemma 4, i.e. $\Omega(\epsilon) \subset \Omega(\epsilon_T)$.

Assumption 2. *Suppose that the closed-loop system satisfies input and state constraints, and it is asymptotically stable in $\Omega(\epsilon_T)$. The states need to satisfy the following properties.*

- *For each initial state $\mathbf{x}(t_0) \in \Omega(\epsilon_T)$, the corresponding state trajectory $\mathbf{x}(s; t_k)$ exists, where $s \in [t_k, t_k + T_k]$.*
- *For any $\epsilon > 0$, there is a $\delta > 0$ such that if the initial state $\mathbf{x}(t_0) \in \Omega(\epsilon_T)$ and $\|\mathbf{x}(t_0)\| \leq \delta$. Then the corresponding state trajectory $\mathbf{x}(s; t_k)$, where $s \in [t_k, t_k + T_k]$ satisfies $\|\mathbf{x}(s; t_k)\| \leq \epsilon$.*
- *For any initial state $\mathbf{x}(t_0) \in \Omega(\epsilon_T)$, the corresponding state trajectory $\mathbf{x}(s; t_k)$, where $s \in [t_k, t_k + T_k]$ satisfies $\lim_{t_k \to \infty} \mathbf{x}(s; t_k) = \mathbf{0}$.*

3.3.2 Self-Triggered Strategy

In this section, we propose a self-triggered control strategy for the nonlinear system (3.1). Let $\mathbf{x}(t_k)$ and $\mathbf{x}(t_k + \tilde{T}_k)$ be the actual state at time t_k and $t_k + \tilde{T}_k$, respectively. Assume that the actual state $\mathbf{x}(t_k)$ at time t_k follows the predicted state trajectory, i.e. $\mathbf{x}(t_k) = \tilde{\mathbf{x}}(t_k; t_k)$. $\mathbf{J}^*(\mathbf{x}(t_k), \tilde{T}_k)$ denotes the optimal cost that is obtained by solving the optimal control problem with the actual state $\mathbf{x}(t_k)$. The self-triggered strategy can be derived based on the optimal cost function $\mathbf{J}^*(\tilde{\mathbf{x}}(s; t_k), \tilde{T}_k)$, $s \in [t_k, t_k + \tilde{T}_k]$. The self-triggered condition is used to determine the next transmission time t_{k+1}, which is obtained by checking if the optimal cost function is decreased, i.e.,

$$
\mathbf{J}^*(\tilde{\mathbf{x}}(t_k + \tilde{T}_k; t_k), \tilde{T}_k) - \mathbf{J}^*(\mathbf{x}(t_k), \tilde{T}_k) < 0. \tag{3.13}
$$

According to Lemma 4, the following inequality holds

$$
\mathbf{J}^*(\tilde{\mathbf{x}}^*(t_k + \tilde{T}_k; t_k), \tilde{T}_k) - \mathbf{J}^*(\mathbf{x}(t_k), \tilde{T}_k) \leq - \int_{t_k}^{t_k+\tilde{T}_k} \frac{1}{\beta}\mathbf{F}(\tilde{\mathbf{x}}^*(s; t_k), \tilde{\mathbf{u}}^*(s; t_k))\mathrm{d}s. \tag{3.14}
$$

where $\mathbf{J}^*(\tilde{\mathbf{x}}^*(t_k+\tilde{T}_k;t_k),\tilde{T}_k)$ is the optimal cost by using the optimal state $\tilde{\mathbf{x}}^*(t_k+\tilde{T}_k;t_k)$ at time $t_k+\tilde{T}_k$. Hence, if $\mathbf{x}(t_k+\tilde{T}_k)=\tilde{\mathbf{x}}^*(t_k+\tilde{T}_k;t_k)$ holds, the Lyapunov candidate would be guaranteed to decrease. In addition, the following inequality can be obtained

$$\begin{aligned}\mathbf{J}^*(\tilde{\mathbf{x}}(t_k+\tilde{T}_k;t_k),\tilde{T}_k)-\mathbf{J}^*(\mathbf{x}(t_k),\tilde{T}_k)\leq&\mathbf{J}^*(\tilde{\mathbf{x}}(t_k+\tilde{T}_k;t_k),\tilde{T}_k)-\mathbf{J}^*(\tilde{\mathbf{x}}^*(t_k+\tilde{T}_k;t_k),\tilde{T}_k)\\&-\int_{t_k}^{t_k+\tilde{T}_k}\frac{1}{\beta}\mathbf{F}(\tilde{\mathbf{x}}^*(s;t_k),\tilde{\mathbf{u}}^*(s;t_k))\mathrm{d}s.\end{aligned}\tag{3.15}$$

$\mathbf{F}(\tilde{\mathbf{x}}(s;t_k),\tilde{\mathbf{u}}(s;t_k))$ is given when the optimal control problem is solved at time t_k. Furthermore, after a small time interval $\delta>0$, the actual state $\mathbf{x}(t_k+\delta)$ and the optimal state $\tilde{\mathbf{x}}^*(t_k+\delta;t_k)$ can be described as

$$\begin{aligned}\mathbf{x}(t_k+\delta)&=\mathbf{x}(t_k)+\int_{t_k}^{t_k+\delta}(\mathbf{f}(\tilde{\mathbf{x}}(s;t_k))+\mathbf{g}(\tilde{\mathbf{x}}(s;t_k))\tilde{\mathbf{u}}^*(s;t_k))\mathrm{d}s+\int_{t_k}^{t_k+\delta}\omega(s)\mathrm{d}s,\\\tilde{\mathbf{x}}^*(t_k+\delta;t_k)&=\mathbf{x}(t_k)+\int_{t_k}^{t_k+\delta}(\mathbf{f}(\tilde{\mathbf{x}}^*(s;t_k))+\mathbf{g}(\tilde{\mathbf{x}}^*(s;t_k))\tilde{\mathbf{u}}^*(s;t_k))\mathrm{d}s.\end{aligned}\tag{3.16}$$

Using the Lipschitz property of $f(\mathbf{x},\mathbf{u})$ in $\mathbf{x}$, we can obtain

$$E=\|\mathbf{x}(t_k+\delta)-\tilde{\mathbf{x}}^*(t_k+\delta;t_k)\|_P\leq L\int_{t_k}^{t_k+\delta}\|\mathbf{x}(s)-\tilde{\mathbf{x}}^*(s)\|_P\mathrm{d}s+\rho\bar{\lambda}(\sqrt{P})\delta.\tag{3.17}$$

By applying the Gronwall-Bellman inequality, the following inequality holds

$$E\leq\frac{\rho\bar{\lambda}(\sqrt{P})\delta}{L}(e^{L\delta}-1).\tag{3.18}$$

To derive a more detailed expression of (3.15), the following result based on Lemma 1 in [144] can be summarized.

Lemma 5. *Suppose the optimal cost $\mathbf{J}^*(\mathbf{x}(t),\tilde{T}_k)$ is Lipschitz continuous for all $\mathbf{x}(t)\in\mathbb{X}$. The Lipschitz constant can be obtained by*

$$\hat{L}=\frac{L_1\bar{\lambda}(\sqrt{Q})}{\beta L}(e^{L\tilde{T}_k}-1)+\frac{L_1\bar{\lambda}(\sqrt{Q})}{L}e^{L(T-\tilde{T}_k)}+\frac{L_2\bar{\lambda}(\sqrt{P})}{L}e^{L\tilde{T}_k}.\tag{3.19}$$

Proof. Suppose the current time is $t_k=0$. By solving the optimal control problem with two different initial states $\mathbf{x}(0)=\mathbf{x}_1(0)$ and $\mathbf{x}(0)=\mathbf{x}_2(0)$, the optimal state trajectories $\tilde{\mathbf{x}}_1^*(s;t_k)$, $\tilde{\mathbf{x}}_2^*(s;t_k)$, control sequences $\tilde{\mathbf{u}}_1^*(s;t_k)$, $\tilde{\mathbf{u}}_2^*(s;t_k)$ and the optimal cost values $\mathbf{J}^*(\tilde{\mathbf{x}}_1^*(s;t_k),\tilde{T}_k)$, $\mathbf{J}^*(\tilde{\mathbf{x}}_2^*(s;t_k),\tilde{T}_k)$ can be obtained, where $s\in[0,T]$. According to the

definition of the cost function, the optimal cost values with the optimal state trajectories are given by

$$\begin{aligned}\mathbf{J}^*(\tilde{\mathbf{x}}_1^*(s;t_k),\tilde{T}_k) &= \mathbf{F}_T(\tilde{\mathbf{x}}_1^*(s;t_k),\tilde{\mathbf{u}}_1^*(s;t_k),\tilde{T}_k)+\mathbf{V}_T(\tilde{\mathbf{x}}_1^*((t_k+T;t_k)),\\ \mathbf{J}^*(\tilde{\mathbf{x}}_2^*(s;t_k),\tilde{T}_k) &= \mathbf{F}_T(\tilde{\mathbf{x}}_2^*(s;t_k),\tilde{\mathbf{u}}_2^*(s;t_k),\tilde{T}_k)+\mathbf{V}_T(\tilde{\mathbf{x}}_2^*((t_k+T;t_k)).\end{aligned} \tag{3.20}$$

Given an alternative control input $\hat{\mathbf{u}}_1(s;t_k) = \tilde{\mathbf{u}}_2^*(s;t_k) \in \mathbb{U}$ with the initial state $\mathbf{x}_1(0)$. The corresponding state $\hat{\mathbf{x}}_1(s;t_k)$ and cost $\bar{\mathbf{J}}^*(\hat{\mathbf{x}}_1(s;t_k),\tilde{T}_k)$ are obtained by applying the alternative control input $\hat{\mathbf{u}}_1(s;t_k)$. Since $\mathbf{J}^*(\tilde{\mathbf{x}}_1^*(s;t_k),\tilde{T}_k)$ is the optimal cost value with the initial state $\mathbf{x}_1(0)$, $\mathbf{J}^*(\tilde{\mathbf{x}}_1^*(s;t_k),\tilde{T}_k) \leq \bar{\mathbf{J}}^*(\hat{\mathbf{x}}_1(s;t_k),\tilde{T}_k)$ holds. In addition, the following inequality holds using Lipschitz property of $\mathbf{F}(\mathbf{x}(t),\mathbf{u}(t))$ and $\mathbf{V}_T(\mathbf{x}(t))$ and Gronwall-Bellman inequality

$$\begin{aligned}&\mathbf{J}^*(\tilde{\mathbf{x}}_1^*(s;t_k),\tilde{T}_k)-\mathbf{J}^*(\tilde{\mathbf{x}}_2^*(s;t_k),\tilde{T}_k) \leq \bar{\mathbf{J}}^*(\hat{\mathbf{x}}_1(s;t_k),\tilde{T}_k)-\mathbf{J}^*(\tilde{\mathbf{x}}_2^*(s;t_k),\tilde{T}_k)\\ \leq &\int_{\tilde{T}_k}^{T} L_1\|\hat{\mathbf{x}}_1(s;t_k)-\tilde{\mathbf{x}}_2^*(s;t_k)\|_Q\mathrm{d}s + L_2\|\hat{\mathbf{x}}_1(s;t_k)-\tilde{\mathbf{x}}_2^*(s;t_k)\|_P\\ &+\int_0^{\tilde{T}_k}\frac{1}{\beta}L_1\|\hat{\mathbf{x}}_1(s;t_k)-\tilde{\mathbf{x}}_2^*(s;t_k)\|_Q\mathrm{d}s\\ \leq &\|\tilde{\mathbf{x}}_1^*(s;t_k)-\tilde{\mathbf{x}}_2^*(s;t_k)\|\left(\frac{L_1\rho\bar{\lambda}(\sqrt{Q})}{\beta L}(e^{L\tilde{T}_k}-1)+\frac{L_1\rho\bar{\lambda}(\sqrt{Q})}{L}e^{L(T-\tilde{T}_k)}\right.\\ &\left.+\frac{L_2\rho\bar{\lambda}(\sqrt{P})}{L}e^{L\tilde{T}_k}\right).\end{aligned} \tag{3.21}$$

Therefore, the proof is complete. □

From Lemma 5, Equation (3.15) can be rewritten as

$$\mathbf{J}^*(\tilde{\mathbf{x}}(t_k+\tilde{T}_k;t_k),\tilde{T}_k)-\mathbf{J}^*(\mathbf{x}(t_k),\tilde{T}_k) \leq \hat{L}E-\int_{t_k}^{t_k+\tilde{T}_k}\frac{1}{\beta}\mathbf{F}(\tilde{\mathbf{x}}^*(s;t_k),\tilde{\mathbf{u}}^*(s;t_k))\mathrm{d}s. \tag{3.22}$$

Since the optimal cost is guaranteed to decrease, we can obtain that

$$E < \frac{\sigma}{\hat{L}\beta}\int_{t_k}^{t_k+\tilde{T}_k}\mathbf{F}(\tilde{\mathbf{x}}^*(s;t_k),\tilde{\mathbf{u}}^*(s;t_k))\mathrm{d}s, \tag{3.23}$$

where $\sigma \in (0,1)$ is a constant. Furthermore, the next execution time can be described as

$t_{k+1} = t_k + \tilde{T}_k^*$, where

$$\tilde{T}_k^* = \max_{\tilde{T}_k \in [0,T]} \{\tilde{T}_k | E = \frac{\sigma}{\hat{L}\beta} \int_{t_k}^{t_k+\tilde{T}_k} \mathbf{F}(\tilde{\mathbf{x}}^*(s;t_k), \tilde{\mathbf{u}}^*(s;t_k))\mathrm{d}s\}. \tag{3.24}$$

3.3.3 ISM Component

To design the ISM component, a well-known result [150] can be recalled, which guarantees the system to evolve in the sliding mode. Given the nonlinear system dynamics (3.1), the nominal closed-loop system can be described as

$$\dot{\mathbf{x}}^{MPC}(t) = \mathbf{f}(\mathbf{x}(t)) + \mathbf{g}(\mathbf{x}(t))\mathbf{u}^{MPC}(t), \tag{3.25}$$

where $\mathbf{x}^{MPC}(t)$ denotes the system state trajectory of the nominal close-loop system under control input $\mathbf{u}^{MPC}(t) = \tilde{\mathbf{u}}^*(t)$. Recall the control input described as (3.3), we need to design the switching function $\mathbf{s}(t)$ as follows

$$\mathbf{s}(t) = \mathbf{s}_0(\mathbf{x}(t)) + \mathbf{z}(t), \tag{3.26}$$

where $\mathbf{s}(t) \in \mathbb{R}^n$. $\mathbf{s}_0(\mathbf{x}(t)) \in \mathbb{R}^n$ is the sliding variable, which is designed as the linear combination of system states. $\mathbf{z}(t) \in \mathbb{R}^n$ is the desired transient trajectory that can be utilized to induce the integral term. To derive the sliding mode equation, the derivative of $\mathbf{s}$ on the state trajectory needs to equal zero, i.e., $\dot{\mathbf{s}} = 0$. The equivalent control $\mathbf{u}_{eq}^{ISM}(t)$ needs to be designed and substituted into the motion equation. Then, based on the equivalent control $\mathbf{u}_{eq}^{ISM}(t)$, we can find the formulation of integral term $\mathbf{z}(t)$ to determine the formulation of $\mathbf{s}(t)$.

According to the $\mathbf{u}_{eq}^{ISM}(t)$, the derivative of desired transient trajectory $\mathbf{z}(t)$ on the state trajectory can be represented as

$$\dot{\mathbf{z}}(t) = -\frac{\partial \mathbf{s}_0(\mathbf{x}(t))}{\partial \mathbf{x}(t)}\big(\mathbf{f}(\mathbf{x}(t)) + \mathbf{g}(\mathbf{x}(t))\mathbf{u}^{MPC}(t)\big), \tag{3.27}$$

with respect to $\mathbf{z}(0) = -\mathbf{s}_0(\mathbf{x}(0))$. Thus, we have the motion equation of the system in sliding mode that can be described as

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t)) + \mathbf{g}(\mathbf{x}(t))\mathbf{u}^{MPC}(t), \tag{3.28}$$

which is of the same order as the nominal system dynamics (3.25). According to (3.27),

we assume that the matrix $\frac{\partial \mathbf{s}_0(\mathbf{x}(t))}{\partial \mathbf{x}(t)}\mathbf{g}(\mathbf{x}(t))$ is non-singular. Then, the discontinuous control component $\mathbf{u}^{ISM}(t)$ can be designed [150] as

$$\mathbf{u}^{ISM}(t) = -\mathbf{u}_M^{ISM}(\mathbf{x}(t))\text{sign}(\mathbf{s}(t)), \tag{3.29}$$

where $\mathbf{u}_M^{ISM}(\mathbf{x}(t)) > 0$ is a diagonal matrix and chosen to enforce the sliding mode on a manifold $\mathbb{S} = \{\mathbf{x} \in \mathbb{X} | \mathbf{s}(t) = 0\}$. To alleviate the chatting phenomenon due to the discontinuity of the control input, we modify the switching function (3.26) by letting $\mathbf{z}(0) = -\mathbf{s}_0(\mathbf{x}(0))$ and

$$\dot{\mathbf{z}}(t) = -\frac{\partial \mathbf{s}_0(\mathbf{x}(t))}{\partial \mathbf{x}}\{\mathbf{f}(\mathbf{x}(t)) + \mathbf{g}(\mathbf{x}(t))(\mathbf{u}^{MPC}(t) - \mathbf{u}_{eq}^{ISM}(t) - \mathbf{u}^{ISM}(t))\}. \tag{3.30}$$

Remark 7. *Due to the uncertain disturbance, the equivalent control is hard to compute directly. A well-known result [150] can be recalled. In [150], the equivalent value can be obtained by the average value, i.e.,* $\mathbf{u}_{eq}^{ISM}(t) = \mathbf{u}_{av}^{ISM}(t)$*, which can be computed based on the first-order linear filter. Thus, the average value can be defined as*

$$\mu\dot{\mathbf{u}}_{av}^{ISM}(t) + \mathbf{u}_{av}^{ISM}(t) = \mathbf{u}^{ISM}(t), \tag{3.31}$$

where μ *is a time constant and small enough to guarantee that the filter does not influence the switching actions.*

Only the unmatched uncertainties are suppressed by the proposed ISM method. The following lemma shows that the ISM controlled system is invariant concerning the matched uncertainty.

Lemma 6. *For the system dynamics* (3.1) *with the additive uncertainties bounded by* ρ*, given the control law* (3.3) *with* (3.29)*, the unmatched uncertainties will not be amplified when the matched uncertainties are completely suppressed,*

Proof. Given the nonlinear system dynamics (3.1) and the designed sliding trajectory can be expressed as (3.27). Thus, the time derivation of the switching function can be determined as

$$\dot{\mathbf{s}}(t) = \frac{\partial \mathbf{s}_0(\mathbf{x}(t))}{\partial \mathbf{x}(t)}\dot{\mathbf{x}}(t) - \frac{\partial \mathbf{s}_0(\mathbf{x}(t))}{\partial \mathbf{x}(t)}\big(\mathbf{f}(\mathbf{x}(t)) + \mathbf{g}(\mathbf{x}(t))\mathbf{u}^{MPC}(t)\big). \tag{3.32}$$

Then, the equivalent control can be computed by using $\dot{\mathbf{s}}(t) = 0$. Thus, we can obtain

$$\begin{aligned}\dot{\mathbf{s}}(t) =& \frac{\partial \mathbf{s}_0(\mathbf{x}(t))}{\partial \mathbf{x}(t)}\big(\dot{\mathbf{x}}(t) - \mathbf{f}(\mathbf{x}(t)) - \mathbf{g}(\mathbf{x}(t))\mathbf{u}^{MPC}(t)\big) \\ =& \frac{\partial \mathbf{s}_0(\mathbf{x}(t))}{\partial \mathbf{x}(t)}\big(\mathbf{f}(\mathbf{x}(t)) + \mathbf{g}(\mathbf{x}(t))\big(\mathbf{u}^{MPC}(t) + \mathbf{u}_{eq}^{ISM}(t)\big) + \mathbf{g}(\mathbf{x}(t))\omega_m(t) \\ &+ \mathbf{g}^{\perp}(\mathbf{x}(t))\omega_u(t) - \mathbf{f}(\mathbf{x}(t) - \mathbf{g}(\mathbf{x}(t))\mathbf{u}^{MPC}(t)\big) \\ =& \frac{\partial \mathbf{s}_0(\mathbf{x}(t))}{\partial \mathbf{x}(t)}\big(\mathbf{g}(\mathbf{x}(t))\mathbf{u}_{eq}^{ISM}(t) + \mathbf{g}(\mathbf{x}(t))\omega_m(t) + \mathbf{g}^{\perp}(\mathbf{x}(t))\omega_u(t)\big) \\ =& 0.\end{aligned}$$

Thus, the equivalent control can be obtained as

$$\mathbf{u}_{eq}^{ISM}(t) = -\omega_m(t) - \Big(\frac{\partial \mathbf{s}_0(\mathbf{x}(t))}{\partial \mathbf{x}(t)}\mathbf{g}(\mathbf{x}(t))\Big)^{-1}\frac{\partial \mathbf{s}_0(\mathbf{x}(t))}{\partial \mathbf{x}(t)}\mathbf{g}^{\perp}(\mathbf{x}(t))\omega_u(t). \tag{3.33}$$

By applying the equivalent control into the system dynamics (3.1), the ISM controller can reject the matched uncertainty $\omega_m(t)$, and the equivalent system dynamics can be represented as

$$\begin{aligned}\dot{\mathbf{x}}(t) =& \mathbf{f}(\mathbf{x}(t)) + \mathbf{g}(\mathbf{x}(t))\mathbf{u}^{MPC}(t) \\ &+ \big(I - \mathbf{g}(\mathbf{x}(t))\Big(\frac{\partial \mathbf{s}_0(\mathbf{x}(t))}{\partial \mathbf{x}(t)}\mathbf{g}(\mathbf{x}(t))\Big)^{-1}\frac{\partial \mathbf{s}_0(\mathbf{x}(t))}{\partial \mathbf{x}(t)}\big)\mathbf{g}^{\perp}(\mathbf{x}(t))\omega_u(t).\end{aligned}$$

According to Propositions 2 and 3 in [151], the equivalent system dynamics can be replaced by

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t)) + \mathbf{g}(\mathbf{x}(t))\mathbf{u}^{MPC}(t) + \mathbf{g}^{\perp}(\mathbf{x}(t))\omega_u(t). \tag{3.34}$$

The equivalent system dynamics are invariant concerning the matched uncertainty, and the proof is complete. □

Remark 8. *The basic principle of sliding mode control is a decoupling control procedure that splits the original problem into lower dimensional independent problems. The invariant property of ISM controller means the siding motions that occur on the discontinuous lower dimensional sliding surfaces are always the same [152]. According to the definition of the decoupled control procedure in [152], we can define a non-singular transformation Σ between the sliding variable $\mathbf{s}(t)$ and a new vector $\hat{\mathbf{s}}(t)$, i.e. $\mathbf{s}(t) = \Sigma \cdot \hat{\mathbf{s}}(t)$ or $\hat{\mathbf{s}}(t) = \Sigma^{-1}\mathbf{s}(t)$. Moreover, according to (3.1), (3.26) and (3.27), if we consider the*

Lyapunov candidate as $V(t) = \frac{1}{2}\mathbf{s}(t)^\top \mathbf{s}(t)$*, we can obtain the time derivative of* V *as*

$$\begin{aligned}\dot{V}(t) =&\mathbf{s}(t)\dot{\mathbf{s}}(t) = (\Sigma^{-1}\hat{\mathbf{s}}(t))(\frac{d}{dt}(\mathbf{s}_0(\mathbf{x}(t)) + \mathbf{z}(t)))\\ =&(\Sigma^{-1}\hat{\mathbf{s}}(t))(\frac{\partial \mathbf{s}_0(\mathbf{x}(t))}{\partial \mathbf{x}}\dot{\mathbf{x}}(t) + \dot{\mathbf{z}}(t))\\ =&(\Sigma^{-1}\hat{\mathbf{s}}(t))(\frac{\partial \mathbf{s}_0(\mathbf{x}(t))}{\partial \mathbf{x}}(\mathbf{f}(\mathbf{x}(t)) + \mathbf{g}(\mathbf{x}(t))\mathbf{u}(t) + \omega(\mathbf{x}(t)) - \mathbf{f}(\mathbf{x}(t))\\ &+ \mathbf{g}(\mathbf{x}(t))\mathbf{u}^{MPC}(t)))\\ =&(\Sigma^{-1}\hat{\mathbf{s}}(t))(\frac{\partial \mathbf{s}_0(\mathbf{x}(t))}{\partial \mathbf{x}}\mathbf{g}(\mathbf{x}(t))\mathbf{u}^{ISM}(t) + \omega(\mathbf{x}(t))).\end{aligned} \tag{3.35}$$

Substitute (3.29) *into the above equation* (3.35)*, we can obtain*

$$\dot{V}(t) = (\Sigma^{-1}\hat{\mathbf{s}}(t))(\frac{\partial \mathbf{s}_0(\mathbf{x}(t))}{\partial \mathbf{x}}\mathbf{g}(\mathbf{x}(t))(-\mathbf{u}_M^{ISM} sign(\hat{\mathbf{s}}(t))) + \omega(\mathbf{x}(t))). \tag{3.36}$$

According to [152], we choose the transformation matrix candidate as $\Sigma = \big(\frac{\partial \mathbf{s}_0(\mathbf{x}(t))}{\partial \mathbf{x}}\mathbf{g}(\mathbf{x})\big)^\top$*. Thus, we can obtain that*

$$\dot{V}(t) = \hat{\mathbf{s}}^\top(t)\big(\mathbf{g}(\mathbf{x}(t))\omega(t) - \mathbf{u}_M^{ISM}(\mathbf{x}(t)) sign(\hat{\mathbf{s}}(t))\big). \tag{3.37}$$

Since the matrix $\mathbf{u}_M^{ISM}(\mathbf{x}(t))$ *enforces the sliding mode on the sliding manifold,* $\dot{V}(t) \leq 0$ *can be guaranteed.*

In summary, the developed control algorithm is presented in Algorithm 2. Algorithm 2 shows that the optimal control problem is solved when the self-triggered condition is violated. The control inputs and next execution time are obtained by solving the optimal control problem. According to the MPC strategy, the first element of the calculated control inputs applies to the PEV as the MPC component of the total control input. The remaining elements of the control inputs will be kept as the predicted MPC component. However, the ISM component of the total control input is calculated at every sampling time instant. When the self-triggered condition does not violate, the ISM component and the predicted MPC component at the current time instant are applied to the PEV.

Algorithm 2: Self-Triggered Robust MPC with ISM

Input: Current system state $\mathbf{x}(t_k)$
Output: Control input trajectory $\mathbf{u}(t_k)$; Next transmission time $t_k + \tilde{T}_k^*$

Set $t_k = 0$;
while do
  At time t_k, measure the current state $\mathbf{x}(t_k)$ of system (3.1);
  Solve the optimal control problem Problem 3 and obtain the next transmission time $t_k + \tilde{T}_k^*$ calculated by (3.24) and the optimal control trajectory $\tilde{u}^*(s; t_k)$ for this time interval;
  for $t_k < t_k + \tilde{T}_k^*$ **do**
    Obtain the control input component from ISM controller $u^{ISM}(t_k)$ by (3.29);
    Calculate the control input $u(t_k) = \tilde{u}^*(t_k; t_k) + u^{ISM}(t_k)$;
    Apply the control trajectory $u(t_k)$;
    Set $t_k = t_k + 1$
  end
  At time $t_k + \tilde{T}_k^*$, set $t_k = t_k + \tilde{T}_k^*$;
end

3.4 Simulation Results

3.4.1 Case A: Nonlinear Dynamics

We consider the following cart-damper-spring system [141].

$$\begin{cases} \dot{x}_1(t) = x_2(t), \\ \dot{x}_2(t) = -\dfrac{k}{M} e^{-x_1(t)} x_1(t) - \dfrac{h}{M} x_2(t) + \dfrac{u(t)}{M} + \dfrac{\omega(t)}{M}. \end{cases} \tag{3.38}$$

Here, the state is denoted by $\mathbf{x}(t) = [x_1(t) \quad x_2(t)]^\top$, consisting of the displacement of the cart $x_1(t)$, and its velocity $x_2(t)$. Also, the initial state satisfies the state constraint $[1, -0.3; 1, -0.3]$. $u(t)$ is the control input subject to the constraint $\|u\| \leq 1$. $\omega(t)$ denotes the additive disturbance, which is the white noise signal bounded by $\rho = \sup_{\omega(t)\in\mathbb{W}} \|\omega(t)\| = 4.0 \times 10^{-4}$ in the simulation. The signal $\eta_u(t)$ is the unmatched uncertain disturbance, which can be generated as the white noise signal bounded by $\sup_{\eta_u(t)\in\mathbb{W}} \|\eta_u(t)\| = 0.2$. $M = 1.25kg$ represents the overall mass of the system, $k = 0.9N/m$ is the constant factor of the nonlinear spring, and $h = 0.42N \cdot s/m$ is the damper factor. In the simulation, the weighting matrices are designed as $Q = [0.1, 0; 0, 0.1]$, $R = 0.1$ and $P = [0.5235, -0.1562; -0.1562, 0.2986]$. The initial value of the system state is given by $x(t) =$

$[1 \quad 0.2]^\top$. According to (3.26), the sliding surface [148] can be chosen as

$$\begin{aligned} s(t) =& x_1(t) + x_2(t) - x_1(0) - x_2(0) \\ & - \int_0^t x_2(\tau) - \frac{k}{M} e^{-x_1(\tau)} x_1(\tau) - -\frac{h}{M} x_2(\tau) + \frac{u(\tau)}{M} \mathrm{d}\tau. \end{aligned} \tag{3.39}$$

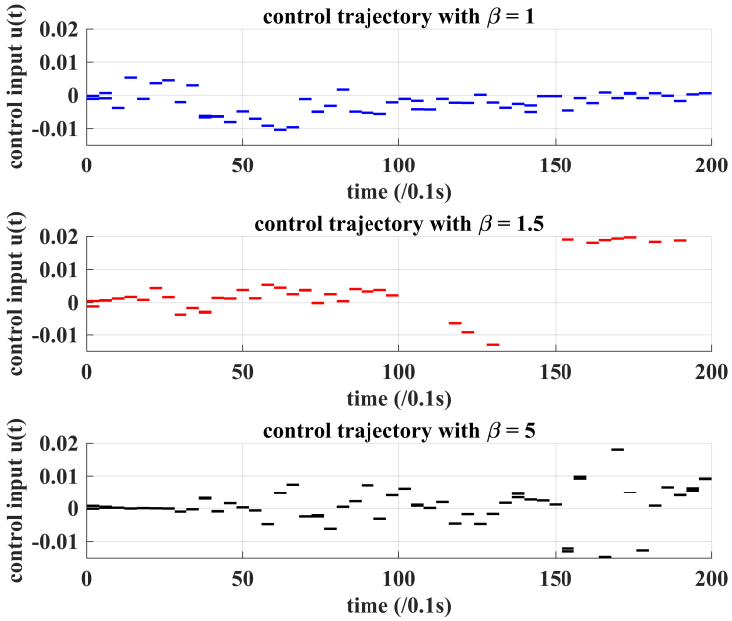


Figure 3.2: Comparison of control input u trajectory.

Figure 3.3 shows the state trajectories of the closed-loop system under Algorithm 2 with three different values of β. Figure 3.2 shows the resulting control input with $\beta = 1$, $\beta = 1.5$ and $\beta = 5$, respectively. From Figure 3.2, we can see that the proposed self-triggered MPC algorithm is feasible, and the control input satisfies the input constraints. The two polygons in Figure 3.3 represent the original state constraints and the terminal region, respectively. Figure 3.3 shows that the state trajectory starts from the initial point $[1 \quad 0.2]^\top$ and converges to the terminal region. In addition, the state constraints and robustness are satisfied. The triggering instants are plotted in Figure 3.4, from which we can see that the proposed self-triggered MPC algorithm can enlarge the sampling time interval to achieve the reduction of communication. Furthermore, the numerical comparison between the proposed method and periodic method is given in Table 3.1. In this table, we can see that with a larger β, the proposed self-triggered MPC scheme can save more on

communication resources and have a lower triggering frequency.

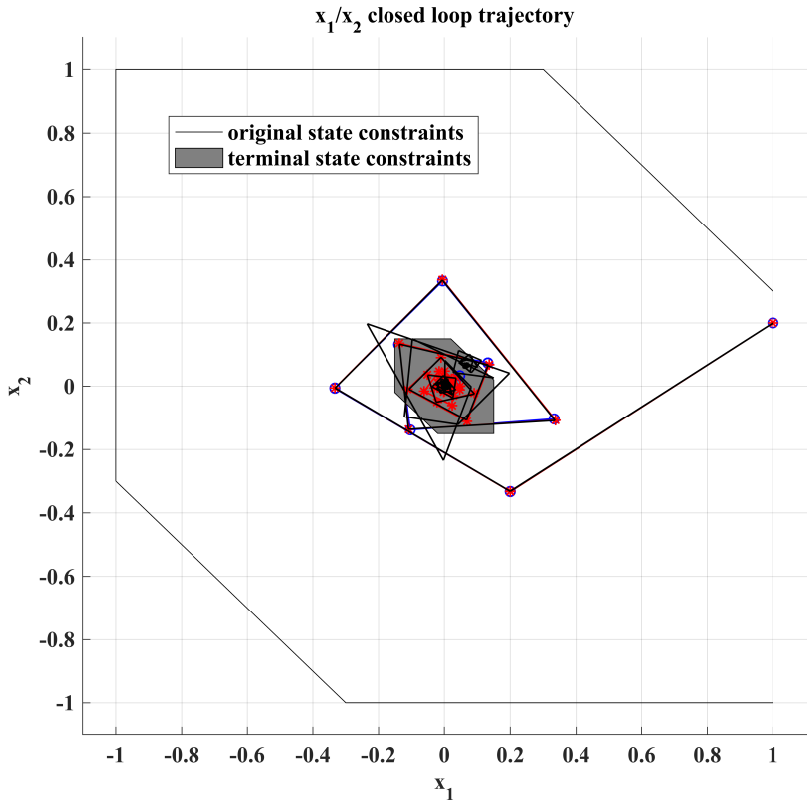


Figure 3.3: Convergence and bounds of state trajectories.

Table 3.1: The numerical comparison of the triggering frequency between the proposed method with different parameters and periodic method.

	Frequency	**Average number of triggering**
Periodic	1	0.1
$\beta = 1$	25.5%	0.39
$\beta = 1.5$	21%	0.47
$\beta = 5$	13.5%	0.74

3.4.2 Case B: PEV Charging

We apply the proposed charging control strategy to a single PEV to show the reduction of communication loads over time horizon $K = 19$h. For the PEV, the SoC of PEV at the current time instant $\tilde{k}$ is given by $SoC(\tilde{k}) \in [0, 1]$. Define $SoC_{ini} = 0.2$ and $SoC_{des} = 1.0$ as the initial SoC and desired SoC for the PEV, respectively. When the vehicle leaves, the

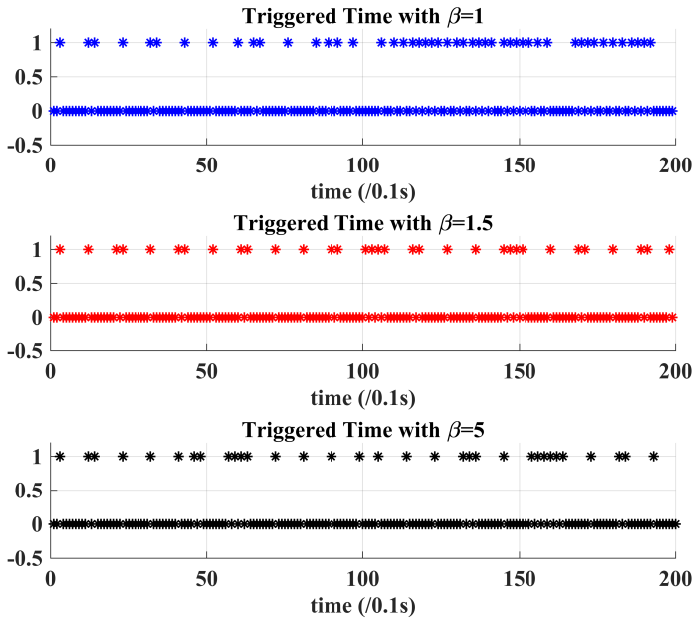


Figure 3.4: Comparison of triggering times with different value of β.

charged energy for the PEV $E_r = \bar{E}(SoC_{des} - SoC_{ini})$ is obtained. $\bar{E} = 20$ kWh is the maximum battery capacity of the PEV. Let $x(\tilde{k}) = E_r - SoC(\tilde{k})\bar{E}$ denote the energy of the PEV to be charged at time instant $\tilde{k}$. According to the charging dynamics, the following system model can be derived,

$$x(\tilde{k}+1) = x(\tilde{k}) + Bu(\tilde{k}) + \omega(t), \quad B = -\eta\Delta t\bar{P}. \tag{3.40}$$

where $u(\tilde{k}) \in \mathbb{R}$ is the charging rate, which is the control input signal to be designed and satisfies the control input constraint $0 \leq u(\tilde{k}) \leq 1$. $\eta = 0.85$, $\Delta t = 1$ and $\bar{P} = 6.6 \times 2$ kW represent the charging efficiency, sampling time period and maximum charging power, respectively. Also, we utilize the same setup of uncertainties as the Case A. $\omega(t)$ is the white noise signal bounded by $\rho = \sup_{\omega(t)\in\mathbb{W}} \|\omega(t)\| = 4.0 \times 10^{-4}$. The signal $\eta_u(t)$ is generated as the white noise signal bounded by $\sup_{\eta_u(t)\in\mathbb{W}} \|\eta_u(t)\| = 0.2$. The triggering instants are plotted in Figure 3.5, showing that the proposed self-triggered MPC algorithm can enlarge the sampling time interval to achieve the reduction of communication when the frequency of updating the baseline of the power load remains the same. The SoC trajectories are plotted in Figure 3.6, from which we can find that the enlarged sampling time interval improves the performance of PEV charging and meets the individual requirements.

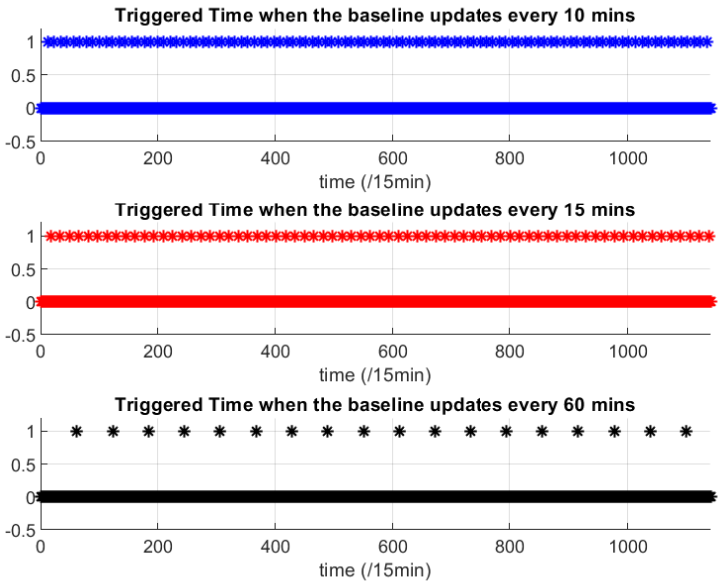


Figure 3.5: Comparison of triggering times with different updating frequency of the baseline of the power load.

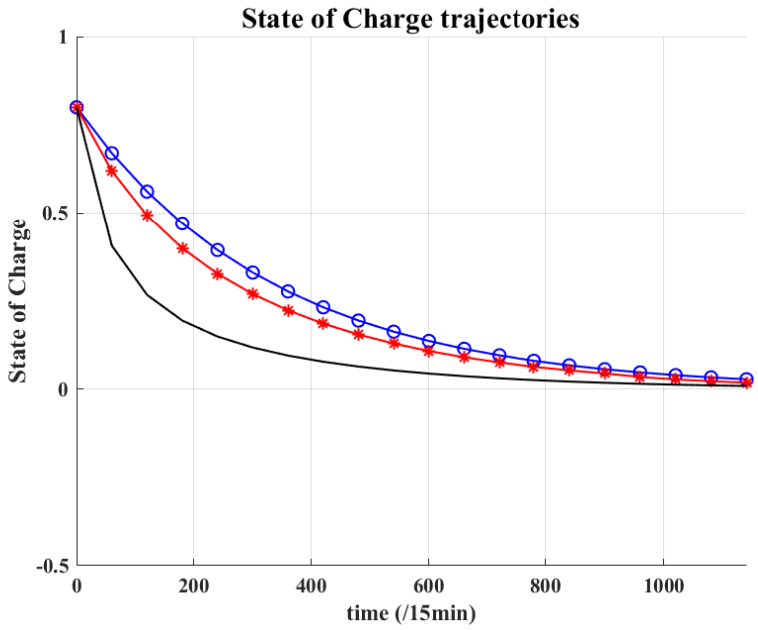


Figure 3.6: The SoC is influenced when the PEV charging is under different sampling time intervals. The blue line with circles, the red line with stars and the solid black line illustrate when the baseline updates every 10 minutes, 15 minutes and 60 minutes, respectively.

3.5 Conclusion

In this chapter, we proposed a self-triggered robust MPC with ISM strategy for nonlinear continuous-time systems subject to state and input constraints in the presence of additive disturbances. A self-triggering condition that involved comparing the cost function values with different execution periods was derived. Our proposed self-triggered control scheme was validated through simulation results. Also, the proposed approach allows a trade-off between the performance and the reduction of communication loads.

Chapter 4

Iterative Learning-Based Decentralized Model Predictive Charging Control for PEVs

4.1 Introduction

In 2018, the transport sector was reported to be the second-largest greenhouse gas emission source in Canada, and the emitted CO_2 reached 25% (185 megatons) of the total national emissions [153]. Due to the growing concerns on the high dependence on fossil fuels and environmental pollution, PEVs are becoming increasingly popular worldwide [154]. However, the massive integration of PEVs into the grid induces many potential problems such as voltage deviation, increased load variations, and the grid's power loss. It thus requires efficient strategies for load shifting, energy trading, and energy storage in the grid.

Smart charging of PEVs needs active control solutions to achieve some practical objectives, such as minimizing generation costs and shifting the power loads by optimization algorithms. In general, the solutions to the PEV charging control can be divided into two categories: Centralized and decentralized controls with different aggregator functions. In the centralized control algorithm, the aggregator is designed for directly managing the charging of all PEVs connected to the aggregator. The objective of charging algorithms is to minimize the deviation between the power demand required by an aggregator and PEVs [155] or to minimize generation costs [156]. However, there are several problems in centralized control algorithms caused by centralized data management schemes. A large amount of information must be transmitted and processed by the aggregator, producing

high communication loads. In Chapter 3, we have illustrated a solution of PEV charging control under the centralized architecture. However, in practice, the PEVs may be reluctant to share their private information with others. Also, the transmission of complete information introduces heavy communication loads, and optimization over a large population of PEVs yields high computational complexity. Hence, quite a few works propose decentralized control methods [46, 58] for large-scale PEVs, to preserve individual authority and relieve the computational burden.

In most of the proposed approaches, the aggregator in the closed-loop system can directly or indirectly control the charging of PEVs and serve as an interface with other entities such as the transmission system operators or energy service providers [56]. The decentralized charging architecture can be divided into two categories: With aggregators and without aggregators. Under the decentralized scheme without aggregators, the management of PEVs as a distributed resource can take advantage of the advanced communication technology and metering infrastructure to improve load flexibility [60]. However, each PEV keeps its private information while achieving charging control objective under the decentralized charging control scheme with the aggregator [40, 41].

Another challenge from the charging problem is to minimize the cost of energy consumption. One of the suitable solutions to control smart energy systems is economic MPC which integrates the electricity prices and forecasts the future vehicle arrival information. In [57], the economic MPC minimizes the cost of energy consumption for an individual PEV. With the increasing adoption of PEVs, efficient charging control schemes need to be developed to reduce the impact of vehicle integration to the power grids. In [35], Tang *et al.* provide an MPC-based algorithm for the optimal vehicle charging scheduling, based on the statistical information of the future vehicle arrivals. Processing the statistical information of the future vehicle arrivals naturally leads to adoption of data-driven approaches. In [52, 53], a two-step modeling framework is developed to merge and clean the useful information of PEV charging data. Zhang *et al.* propose a data-driven model that captures the non-homogeneity and the periodicity of the residential PEV charging behaviour [54]. To maximize the efficiency of PEV charging, [49] develops a decentralized energy management system for regulating the energy flow among the photovoltaic systems based on the MPC algorithm. However, in aforementioned results, there is a challenge in applying MPC for PEV charging when the price cost is not involved into the optimal control problem. The accurate price cost is important for designing the cost function under the MPC framework, but it is difficult to obtain in commercial applications. Iterative learning control can improve the transient response of uncertain dynamics operated repetitively or under the

periodical process [157]. Since the charging condition of power grids change from day to day can be taken as an entire trial under iterative learning framework, the price cost of PEV charging using an MPC method can be approximated iteratively.

Based on the above discussions, we tackle the PEV charging problem using a decentralized MPC framework, supplemented by an iterative learning method, to address the valley-filling control problem in the power grids subject to the individual charging constraints. Due to the complexity of the electricity price cost, this work applies an iterative learning method to approximate the price function and update system state constraints. The presented decentralized framework with the aggregator is designed to protect private information while meeting the individual charging requirements. Each PEV has its own charging scenarios, such as different charging power demands and charging periods. The proposed decentralized algorithm allows that each PEV solves its own optimal control problem to minimize the charging cost individually. Also, the proposed algorithm is designed for solving the valley-filling problem by managing all the charging rates and charging time of PEVs.

The remainder of this chapter is organized as follows. Section 4.2 formulates valley-filling problem under the decentralized charging architecture using an MPC framework, followed by the sampled safe set and control algorithm design under the iterative learning framework. Section 4.3 studies the Nash equilibrium for the valley-filling problem under a decentralized charging architecture. An example to verify the theoretical results is shown in Section 4.4. The conclusion of this chapter is given in Section 4.5.

4.2 Problem Formulation and Algorithm Design

In order to achieve valley-filling performance, the controller to be designed and implemented should: (1) fill the overnight power load valley by the charging power of PEVs; (2) minimize the electricity price and the deviation of power demands while each PEV is charged to its desired SoC; (3) guarantee the profiles of power loads under the limitation.

In the rest of this section, we first introduce the system dynamics based on the PEV charging model and the cost function in Section 4.2.1 using the optimization method. To improve the performance of iterations, the sampled safe set is studied in Section 4.2.2. The decentralized MPC scheme, supplemented by an iterative learning method, is proposed in Section 4.2.3.

4.2.1 Valley-Filling Using Optimization Method

We consider a decentralized charging control strategy for the valley-filling problem with PEVs' population size n over time horizon K. For the i-th individual PEV ($i \in \mathbb{N} := \{1, 2, \cdots, n\}$), the obtained SoC of PEV at the current time instant $\tilde{k}$ is given by $SoC_i(\tilde{k}) \in [0, 1]$. Define $SoC_{i,ini}$ and $SoC_{i,des}$ as the initial SoC and desired SoC for the i-th PEV, respectively. When the vehicle arrives, the charged energy for the i-th PEV $E_{i,r} = \bar{E}_i(SoC_{i,des} - SoC_{i,ini})$ would be known. $\bar{E}_i$ is the maximum battery capacity of the i-th PEV. Let $x_i(\tilde{k}) = E_{i,r} - SoC_i(\tilde{k})\bar{E}_i$ denote the energy of the i-th PEV to be charged at time instant $\tilde{k}$. According to the battery charging dynamics $SoC_i(\tilde{k}+1) = SoC_i(\tilde{k}) + \frac{\eta_i \Delta t \bar{P}_i}{\bar{E}_i} u_i(\tilde{k})$, the following system model can be derived,

$$x_i(\tilde{k}+1) = x_i(\tilde{k}) + B_i u_i(\tilde{k}), \quad B_i = -\eta_i \Delta t \bar{P}_i. \tag{4.1}$$

where $u_i(\tilde{k}) \in \mathbb{R}$ is the charging rate of the i-th PEV, and it is the control input signal to be designed and needs to satisfy the physical constraint $0 \leq u_i(\tilde{k}) \leq 1$ and $\sum_{\tilde{k}=k_i}^{k_i+K_i} u_i(\tilde{k}) = \gamma_i$, where $\gamma_i = -\frac{x_i(k_i)}{B_i}$. η_i, Δt and $\bar{P}_i$ represent the charging efficiency, sampling time period and maximum charging power, respectively.

The valley-filling charging control algorithm will start at time k and stop at time $k+K$. Similarly, k_i and $k_i + K_i$ are the start and end time of the i-th PEV charging process, respectively. The i-th PEV charging process does not exceed the valley-filling charging control time period, i.e., $\tilde{k} \in [k, k+K]$, $\bigcup_{i=1}^{n}[k_i, k_i + K_i] \subseteq [k, k+K]$. We assume that the control input signal is zero when the vehicle is not charging. Especially, we assume that the i-th PEV will not be charged outside the known charging time interval, i.e., $u_i(\tilde{k}) = 0$ when $\tilde{k} \in [k, k_i]$ or $\tilde{k} \in [k_i + K_i, k+K]$. As a result, the system state can be denoted as $x_i(\tilde{k}) = E_{i,r}$ when $\tilde{k} \in [k, k_i]$, and $x_i(\tilde{k}) = 0$ when $\tilde{k} \in [k_i + K_i, k+K]$. Furthermore, $x_i(k_i + K_i) = 0$ definitely guarantees that the i-th PEV is fully charged.

Based on the PEV charging dynamics presented above, we focus on the design of the decentralized charging controller, aiming to minimize the variance of the total power load profile by concerning the baseline of power demand. The calculated charging control input sequence for the i-th PEV at time $\tilde{k}$ will be calculated, as $\mathbf{u}_i(s; \tilde{k}) = [u_i(\tilde{k}; \tilde{k}), u_i(\tilde{k}+1; \tilde{k}), \cdots, u_i(k_i + K_i; \tilde{k})]$, $s \in [\tilde{k}, k_i + K_i]$. $u_i(\tilde{k}; \tilde{k})$ stands for the control input that will be applied to the i-th PEV at time $\tilde{k}$. $u_i(s; \tilde{k})$, $s \in [\tilde{k}+1, k_i + K_i]$, denote the predicted control inputs of the i-th PEV obtained at time $\tilde{k}$. The cost function of the optimization problem is

constructed as

$$J_i(\mathbf{u}_i(s;\tilde{k});\hat{U}_{-i}(\tilde{k})) = \sum_{s=\tilde{k}}^{k_i+K_i} p(u_i(s;\tilde{k});\hat{U}_{-i}(\tilde{k}))D_{\tilde{k}} + \delta(u_i(\tilde{k};\tilde{k}) - \bar{u}(\tilde{k}))^2, \tag{4.2}$$

where δ is an adjustable parameter. $\mathbf{u}_i(s;\tilde{k})$, $s \in [\tilde{k}, k_i + K_i]$ will be sent to the aggregator as the predicted control input to calculate the average value $\bar{u}(\tilde{k})$ used in the optimization problem of the next time instant, i.e., $\hat{u}_i(\tilde{k}) = u_i(\tilde{k};\tilde{k}-1)$. $\hat{U}_{-i}(\tilde{k})$ denotes the collection of predicted control inputs $\hat{u}_i(\tilde{k})$ of each PEV, at time $\tilde{k}$ except the i-th PEV, i.e., $\hat{U}_{-i}(\tilde{k}) = \{\hat{u}_{\tilde{i}}(\tilde{k})|\tilde{i} \in \mathbb{N}/i\}$. $\bar{u}(\tilde{k}) = \frac{1}{n}\sum_{i=1}^{n}\hat{u}_i(\tilde{k})$ denotes the average value of predicted control inputs, which will be calculated and broadcast by the aggregator. $p(u_i(s;\tilde{k});\hat{U}_{-i}(\tilde{k}))$ represents the electricity price function that is the derivative of the generation cost function. In this work, we formulate the generation cost function as $c(D_{\tilde{k}}) = \frac{1}{2}aD_{\tilde{k}}^2 + bD_{\tilde{k}}$, i.e., $p(u_i(s;\tilde{k});\hat{U}_{-i}(\tilde{k})) = aD_{\tilde{k}} + b$, where a and b are constants; $D_{\tilde{k}} = D(\tilde{k}) + \sum_{\tilde{i}\in\mathbb{N}/i}\hat{u}_{\tilde{i}}(\tilde{k}) + u_i(s;\tilde{k})$, where $D(\tilde{k})$ represents the power load baseline at time $\tilde{k}$. The profile of power load baseline comes from the power demand of appliances other than PEVs.

Remark 9. *The cost function consists of two terms. The first term denotes the electricity price function, representing the derivative of the generation cost function. The form of the price function has been utilized for the charging problem in [158]. The second term is designed for penalizing the deviation between the charging control input and the average control input to overcome the oscillating problem. All PEVs solve their optimization problems and update their control inputs, which could result in an oscillation between the current and the next charging interval. The oscillation will influence the convergence of the proposed algorithm negatively. As aforementioned, we can obtain that the cost function is convex with respect to the admissible charging control input.*

4.2.2 Sampled Safe Set of Charging Problem

The iterative learning control features the following properties: (1) The sampled safe set in the optimal control problem is updated with the iterations; (2) The robustness and accuracy are guaranteed as the iteration number increases; (3) The price function in the optimal control problem is also updated by solving the optimization problem iteratively to approximate the accurate price cost. Based on the stored measurements under the iterative learning framework, the sampled safe set increases the speed of solving the optimal control prob-

lem. In this subsection, we reformulate the cost function and design the sampled safe set based on the iterative learning method.

In the design phase, PEVs need to be fully charged with variant charging rates and time at every iteration. The sum of all power loads needs to be optimized under a decentralized architecture to satisfy the valley-filling performance. A suitable solution can be obtained after the convergence of the algorithm updated for price function and sampled safe set by the iterative learning method, which may actually result in a suboptimal charging control input for each PEV. A sampled safe set keeps all the possible solutions to each optimal control problem for each PEV. The set can be utilized to increase the speed to satisfy the valley-filling performance when solving the optimization problem.

The sampled safe set for each PEV is a collection of system states by applying the control input sequences. The set includes system state trajectories of successful iterations that are helpful to the recursive feasibility of the optimization problem. Inspired by [159], we define the sampled safe set for the charging problem in this work at the j-th iteration as

$$SS_i^j = \Big\{ \bigcup_{m \in M_i^j} \bigcup_{\tilde{k}=k_i}^{k_i+K_i} x_i^m(\tilde{k}) \Big\}, \tag{4.3}$$

where $M_i^j = \{m \in [0, j] | x_i^m(k_i + K_i) = 0\}$. M_i^j stands for the collection of indices associated with successful iterations m of the i-th PEV. $x_i^m(\tilde{k})$ is the system state at the m-th iteration of the i-th PEV at time $\tilde{k}$. SS_i^j is the collection of all state trajectories at iteration $m \in M_i^j$. Thus, when $m \leq j$, we have $M_i^m \subset M_i^j$. When $j = 0$, we assume that $SS_i^0 = \Big\{ \bigcup_{\tilde{k}=k_i}^{k_i+K_i} x_i^0(\tilde{k}) \Big\}$.

Based on the cost function (4.2), we define the iterative cost at the j-th iteration at time $\tilde{k}$. The cost function is rewritten as

$$J_i^j(\mathbf{u}_i^j(s;\tilde{k}); \hat{U}_{-i}(\tilde{k})) = \sum_{s=\tilde{k}}^{k_i+K_i} p^j(u_i^j(s;\tilde{k}); \hat{U}_{-i}(\tilde{k})) D_{\tilde{k}}^j + \delta(u_i^j(\tilde{k};\tilde{k}) - \bar{u}(\tilde{k}))^2, \tag{4.4}$$

where $\mathbf{u}_i^j(s;\tilde{k}) = [u_i^j(\tilde{k};\tilde{k}), u_i^j(\tilde{k}+1;\tilde{k}), \cdots, u_i^j(k_i+K_i;\tilde{k})]$, $s \in [\tilde{k}, k_i + K_i]$ is the calculated control input sequence at iteration j of the i-th PEV at time $\tilde{k}$. $u_i^j(s;\tilde{k})$ stands for the calculated control input of the i-th PEV at time $\tilde{k}$ for each sampling time instant at the j-th iteration. The price function is updated as

$$p^{j+1}(u_i^j(\tilde{k};\tilde{k}); \hat{U}_{-i}(\tilde{k})) = p^j(u_i^j(\tilde{k};\tilde{k}); \hat{U}_{-i}(\tilde{k})) + \lambda(c(D_{\tilde{k}}) - p^j(u_i^j(\tilde{k};\tilde{k}); \hat{U}_{-i}(\tilde{k}))), \tag{4.5}$$

where $\lambda \in [0,1]$ is the learning parameter.

4.2.3 Iterative Learning-Based Decentralized Model Predictive Control Algorithm Design

The following constrained optimal control problem is formulated using the model and the sampled safe set which are discussed in Section 4.2.1 and Section 4.2.2, respectively.

Problem 4. *(Optimal Control Problem (OCP))*

$$\min_{\mathbf{u}_i} J_i^j(\mathbf{u}_i^j(s;\tilde{k});\hat{U}_{-i}(\tilde{k})) = \sum_{s=\tilde{k}}^{k_i+K_i} \left(p^j(u_i^j(s;\tilde{k});\hat{U}_{-i}(\tilde{k}))D_{\tilde{k}}^j + \delta(u_i^j(s;\tilde{k}) - \bar{u}(\tilde{k}))^2 \right) \quad (4.6)$$

$$\textbf{\textit{s.t.}}\ x_i^j(\tilde{k}+1) = x_i^j(\tilde{k}) + B_i u_i^j(\tilde{k};\tilde{k}), \quad (4.7)$$

$$p^{j+1}(u_i^j(\tilde{k};\tilde{k});\hat{U}_{-i}(\tilde{k})) = p^j(u_i^j(\tilde{k};\tilde{k});\hat{U}_{-i}(\tilde{k})) + \lambda(c(D_{\tilde{k}}) - p^j(u_i^j(\tilde{k};\tilde{k});\hat{U}_{-i}(\tilde{k}))) \quad (4.8)$$

$$u_i^j(s;\tilde{k}) \in [0,1],\ x_i^j(\tilde{k}+1) \in SS_i^{j-1}, \quad (4.9)$$

$$x_i^j(k_i+K_i) = 0, \quad (4.10)$$

where $s \in [\tilde{k}, k_i+K_i]$, $B_i = -\eta_i \Delta t \bar{P}_i$ *and* $D_{\tilde{k}} = D(\tilde{k}) + \sum_{\tilde{i}\in\mathbb{N}/i} \hat{u}_{\tilde{i}}(\tilde{k}) + u_i(\tilde{k};\tilde{k})$.

Due to the coordination of local controllers to optimize the sum of all power loads, the optimal control input sequence $\mathbf{u}_i^*(s;\tilde{k}) = [u_i^*(\tilde{k};\tilde{k}), u_i^*(\tilde{k}+1;\tilde{k}), \cdots, u_i^*(k_i+K_i;\tilde{k})]$, $s \in [\tilde{k}, k_i+K_i]$, of the i-th PEV can be obtained after the convergence of the algorithm. $u_i^*(\tilde{k};\tilde{k})$ stands for the control input applied to the i-th PEV at time $\tilde{k}$ that obtains after the convergence of the algorithm. The optimal control input of the above optimal control problem can minimize the deviation of the sum of all power loads to realize the valley-filling performance and reduce the price cost. The decentralized control algorithm is presented in Algorithm 3.

4.3 Systems Analysis

In this section, the properties of the closed-loop system are analyzed. The sum of all power loads needs to be optimized under a decentralized architecture to satisfy the valley-filling performance with multiple local controllers. A suitable solution can be obtained after the convergence of the algorithm updated by the iterative learning method, which may actually

Algorithm 3: Iterative Learning-Based Decentralized Model Predictive Charging Control for PEVs

Input: Scheduling horizon $[k, k+K]$; Base load profile D; Number of PEVs n; Maximum iteration l_{max}; Tolerance ϵ_{stop}

Output: Charging control input for each PEV $u_i(\tilde{k})$, $i \in \mathbb{N}$

Set $\tilde{k} = k$;
while $\tilde{k} \neq k + K$ **do**
  The operator generates $D(\tilde{k})$;
  The operator broadcasts $D(\tilde{k})$ and $\bar{u}(\tilde{k})$ to all PEVs;
  for *Each PEV* i **do**
    if $\tilde{k} \notin [k_i, k_i + K_i]$ **then**
      $u_i(\tilde{k}; \tilde{k}) = 0$
    end
    else
      Set $j = 0$;
      Set $\epsilon = 0$;
      while $\epsilon \geq \epsilon_{stop}$ *and* $j \leq l_{max}$ **do**
        if *Problem 4 is feasible* **then**
          Compute $\mathbf{u}_i^*(s; \tilde{k})$;
          Update $p^{j+1}(u_i(\tilde{k}; \tilde{k}); \hat{U}_{-i}(\tilde{k}))$;
          Compute $\epsilon = p^{j+1}(u_i(\tilde{k}; \tilde{k}); \hat{U}_{-i}(\tilde{k})) - p^{j}(u_i(\tilde{k}; \tilde{k}); \hat{U}_{-i}(\tilde{k}))$;
          Update $SS_i^j = \left\{ \bigcup_{m \in M_i^j} \bigcup_{\tilde{k}=k_i}^{k_i+K_i} x_i^m(\tilde{k}) \right\}$
        end
        Set $j = j + 1$;
      end
    end
  end
  The operator collects $\mathbf{u}_i^*(s; \tilde{k})$ and updates $\bar{u}(\tilde{k}+1)$;
  for *Each PEV* i **do**
    Implement the charging control at time $\tilde{k}$;
    Update the state $x_i(\tilde{k}+1)$ according to (4.1);
  end
  Set $\tilde{k} = \tilde{k} + 1$;
end

result in a suboptimal charging control input for each PEV. The Nash equilibrium of the valley-filling problem needs to be considered.

Definition 1. *A collection of control inputs* $U^*(\tilde{k}) = \{u_i^*(\tilde{k}; \tilde{k}) | i \in \mathbb{N}\}$ *is Nash equilibrium if it satisfies that the cost of each PEV can be minimized by the corresponding individual*

control input $u_i^*(\tilde{k};\tilde{k})$*, i.e.,*

$$J_i(\mathbf{u}_i^*(s;\tilde{k});U_{-i}^*(\tilde{k})) \leq J_i(\mathbf{u}_i(s;\tilde{k});U_{-i}^*(\tilde{k})), \tag{4.11}$$

for all $\mathbf{u}_i^*(s;\tilde{k})$*,* $s \in [\tilde{k}, k_i + K_i]$*,* $i \in \mathbb{N}$*.* $J_i(\mathbf{u}_i^*(s;\tilde{k});U_{-i}^*(\tilde{k}))$ *stands for the optimal cost of* $\mathbf{u}_i^*(s;\tilde{k})$ *with* $U_{-i}^*(\tilde{k})$ *which denotes the collection of the optimal control inputs except the* i*-th PEV.*

Then, we need to show that the collection of optimal control inputs, obtained after the convergence of Algorithm 3, satisfies the above definition of the Nash equilibrium.

Lemma 7. *A collection of optimal control inputs* $U^*(\tilde{k}) = \{u_i^*(\tilde{k};\tilde{k})|i \in \mathbb{N}\}$ *for all PEVs, obtained after the convergence of Algorithm 3, is a Nash equilibrium, if* $u_i^*(\tilde{k};\tilde{k})$*,* $i \in \mathbb{N}$ *minimizes the corresponding cost function*

$$J_i(\mathbf{u}_i^*(s;\tilde{k});U_{-i}^*(\tilde{k})) = \sum_{s=\tilde{k}}^{k_i+K_i} p(u_i^*(s;\tilde{k});U_{-i}^*(\tilde{k}))D_{\tilde{k}} + \delta(u_i^*(s;\tilde{k}) - \bar{u}^*(\tilde{k}))^2,$$

with respect to the collection of the optimal control inputs except the i*-th PEV* $U_{-i}^*(\tilde{k})$ *and a calculated optimal average value* $\bar{u}^*(\tilde{k})$ *of all individual control inputs.*

Proof. Consider the collection of optimal control inputs $U^*(\tilde{k}) = \{u_i^*(\tilde{k};\tilde{k})|i \in \mathbb{N}\}$, where $u_i^*(\tilde{k};\tilde{k})$ is the optimal solution at the time instant $\tilde{k}$. Then, we have

$$\frac{1}{n}\big(u_i(\tilde{k};\tilde{k}) + \sum_{\tilde{i}\in\mathbb{N}/i} \hat{u}_{\tilde{i}}(\tilde{k})\big) \geq \frac{1}{n}\sum_{i\in\mathbb{N}} u_i^*(\tilde{k};\tilde{k}) = \bar{u}^*(\tilde{k}).$$

Consequently, we obtain that

$$\begin{aligned} J_i(\mathbf{u}_i^*(s;\tilde{k});U_{-i}^*(\tilde{k})) \leq & \sum_{s=\tilde{k}}^{k_i+K_i} p(u_i^*(s;\tilde{k});U_{-i}^*(\tilde{k}))D_{\tilde{k}} + \delta\Big(u_i(\tilde{k};\tilde{k}) - \frac{1}{n}\big(u_i(\tilde{k};\tilde{k}) \\ & + \sum_{\tilde{i}\in\mathbb{N}/i} \hat{u}_{\tilde{i}}(\tilde{k})\big)\Big)^2 \\ = & J_i(\mathbf{u}_i(s;\tilde{k});U_{-i}^*(\tilde{k})). \end{aligned}$$

Thus, we obtain that $u_i^*(\tilde{k};\tilde{k})$ for each PEV is the optimal solution of the Problem 4. According to the definition of Nash equilibrium, the collection of optimal control inputs $U^*(\tilde{k})$ is the Nash equilibrium of the proposed optimization problem. The proof is completed. □

If the Nash equilibrium exists, we need to prove that $u_i^*(\tilde{k};\tilde{k})$ is unique for the optimal control problem of the i-th individual PEV when it is subject to a fixed $\bar{u}(\tilde{k})$.

Lemma 8. *Given the average value $\bar{u}(\tilde{k})$ and control input $u_i(s;\tilde{k})$, the following $u_i^*(\tilde{k};\tilde{k})$ based on $D_{\tilde{k}}$,*

$$u_i^*(\tilde{k};\tilde{k}) = \frac{1}{2\delta}\max\{0, D_{\tilde{k}} - p(u_i(s;\tilde{k});\hat{U}_{-i}(\tilde{k})) + 2\delta\bar{u}(\tilde{k})\}, \tag{4.12}$$

is the unique optimal control input.

Proof. Define a Lagrangian

$$L_i(u_i(\tilde{k};\tilde{k}), A) = J_i(\mathbf{u}_i(s;\tilde{k});\hat{U}_{-i}(\tilde{k})) + A(\gamma_i - \sum_{s=\tilde{k}}^{k_i+K_i}(u_i(s;\tilde{k}))), \tag{4.13}$$

where A is the Lagrange multiplier. Based on the cost function $J_i(\mathbf{u}_i(s;\tilde{k});\hat{U}_{-i}(\tilde{k}))$, we have two conditions, $\frac{\partial L_i}{\partial A} = 0$ and $\frac{\partial L_i}{\partial u_i(\tilde{k};\tilde{k})} \leq 0$, where $u_i(\tilde{k};\tilde{k}) \in [0,1]$. Obviously, the condition $\frac{\partial L_i}{\partial A} = 0$ recalls $\sum_{\tilde{k}=k_i}^{k_i+K_i} u_i(\tilde{k}) = \gamma_i$. Then, we have

$$J_i(\mathbf{u}_i(s;\tilde{k});\hat{U}_{-i}(\tilde{k})) = \sum_{s=\tilde{k}}^{k_i+K_i} p(u_i(s;\tilde{k});\hat{U}_{-i}(\tilde{k}))D_{\tilde{k}} + \delta(u_i(s;\tilde{k}) - \bar{u}(\tilde{k}))^2 \tag{4.14}$$

and

$$\frac{\partial L_i}{\partial u_i(\tilde{k};\tilde{k})} = p(u_i(s;\tilde{k});\hat{U}_{-i}(\tilde{k})) + 2\delta(u_i(\tilde{k};\tilde{k}) - \bar{u}(\tilde{k})) - A\begin{cases} = 0, & u_i(\tilde{k};\tilde{k}) > 0, \\ < 0, & u_i(\tilde{k};\tilde{k}) = 0. \end{cases} \tag{4.15}$$

Then, the control inputs can be represented based on $\bar{u}(\tilde{k})$ and Lagrange multiplier A.

$$u_i^*(\tilde{k};\tilde{k}) = \frac{1}{2\delta}\max\{0, A - p(u_i(s;\tilde{k});\hat{U}_{-i}(\tilde{k})) + 2\delta\bar{u}(\tilde{k})\}.$$

Besides, we use the unique value $D_{\tilde{k}}$ for each given $\bar{u}(\tilde{k})$, where $D_{\tilde{k}}$ ensures $\sum_{\tilde{k}=k_i}^{k_i+K_i} u_i(\tilde{k}) = \gamma_i$. Thus, we have $D_{\tilde{k}}$ and $u_i(\tilde{k};\tilde{k}) = u_i^*(\tilde{k};\tilde{k})$. Since $J_i(\mathbf{u}_i(s;\tilde{k});\hat{U}_{-i}(\tilde{k}))$ is convex, $u_i^*(\tilde{k};\tilde{k})$ is unique. The proof is completed. □

Remark 10. *The Lagrangian function consists of two terms. The first term denotes the cost function. The second term represents the constraints with a Lagrange multiplier. In this*

work, there is only one decision variable to minimize the cost function. Inspired by [46], we replace the constraints related to the charging dynamics (4.7) and the terminal state constraint (4.10) by $\sum_{\tilde{k}=k_i}^{k_i+K_i} u_i(\tilde{k}) = \gamma_i$ *where* $\gamma_i = -\frac{x_i(k_i)}{B_i}$. *The replacement is designed to show the KKT conditions [160]. We assume that* $u_i(\tilde{k}) = 0$ *holds when* $\tilde{k} \in [k, k_i]$ *or* $\tilde{k} \in [k_i + K_i, k + K]$. *According to the system model (4.1) for each vehicle, we list all the equations at each sampling time and sum the equations together. Since* B_i *is a constant, we can obtain that* $\sum_{\tilde{k}=k_i}^{k_i+K_i} u_i(\tilde{k}) = \gamma_i$ *if* $x_i(k_i + K_k) = 0$ *holds.*

Finally, based on the above discussions, we need to show that the Nash equilibrium is the solution to the valley-filling problem.

Theorem 2. *Given* $p(u_i(s;\tilde{k}); \hat{U}_{-i}(\tilde{k}))$ *and the collection of charging control inputs* $U^*(\tilde{k}) = \{u_i^*(\tilde{k};\tilde{k}) | i \in \mathbb{N}\}$ *for PEVs, obtained after the convergence of the algorithm, is a Nash equilibrium. Then the following valley-filling properties are satisfied,*

$$\begin{aligned} &\bar{u}^*(k_i) \geq \bar{u}^*(k_i + K_i), \text{ if } D(k_i) \leq D(k_i + K_i), \\ &D(k_i) + \bar{u}^*(k_i) \leq D(k_i + K_i) + \bar{u}^*(k_i + K_i), \\ &D(k_i) + \bar{u}^*(k_i) = P_i, \end{aligned} \tag{4.16}$$

where P_i *is a constant.*

Proof. For the i-th PEV $[k_i, k_i + K_i]$, the control input sequence is denoted by $\mathbf{u}_i = [u_i(k_i), u_i(k_i + 1), \cdots, u_i(k_i + K_i)]$. Each control input satisfies the control constraint and $u_i(k_i) + u_i(k_i + K_i) \leq \gamma_i$. We define $\alpha_i = \frac{u_i(k_i+K_i)-u_i(k_i)}{2}$, $\beta_i = \frac{u_i(k_i+K_i)+u_i(k_i)}{2}$, so we can obtain $u_i(k_i) = \beta_i - \alpha_i$, $u_i(k_i + K_i) = \beta_i + \alpha_i$. We define the minimum of the cost function as a Bellman equation. Then, we have

$$\begin{aligned} V_i = &\min \left(2\delta\alpha_i - \frac{1}{2}(\bar{u}(k_i + K_i) - \bar{u}(k_i)) \right) + g(\beta_i) \\ &+ \frac{1}{4\delta} \left(p(u_i(s; k_i + K_i); \hat{U}_{-i}(k_i + K_i)) - p(u_i(s; k_i); \hat{U}_{-i}(k_i)) \right)^2, \end{aligned}$$

where $g(\beta_i)$ represents a term that is not related to α_i. Then we obtain the optimal solution of α_i based on the optimal value of β_i as $\alpha_i^*(\beta_i^*) = \arg\min_{\alpha_i \in [-\beta_i, \beta_i]} (\alpha_i - \zeta)^2$, where $\zeta = \frac{1}{2}(\bar{u}(k_i + K_i) - \bar{u}(k_i)) - \frac{1}{4\delta} \left(p(u_i(s; k_i + K_i); \hat{U}_{-i}(k_i + K_i)) - p(u_i(s; k_i); \hat{U}_{-i}(k_i)) \right)$.

Besides, we can obtain that

$$\begin{cases} 0 < \alpha_i^* \leq \zeta, & \text{if } \zeta > 0, \\ \alpha_i^* = \zeta = 0, & \text{if } \zeta = 0, \\ \zeta \leq \alpha_i^* < 0, & \text{if } \zeta < 0. \end{cases} \tag{4.17}$$

To prove these properties, we can construct the contradiction. In the valley-filling problem considered in this work, the suitable start charging time should be in the valley of the baseline, if possible. Thus, if there exist k_i and K_i satisfying $D(k_i) < D(k_i+K_i)$, $\bar{u}^*(k_i) < \bar{u}^*(k_i + K_i)$ and $D(k_i) + \bar{u}^*(k_i) > D(k_i + K_i) + \bar{u}^*(k_i + K_i)$. Then, we have $\bar{u}^*(k_i) < \bar{u}^*(k_i + K_i)$ and $\zeta < \frac{1}{2}(\bar{u}^*(k_i + K_i) - \bar{u}^*(k_i))$.

Furthermore, we obtain that $\alpha_i^* < \frac{1}{2}(\bar{u}^*(k_i + K_i) - \bar{u}^*(k_i))$ due to (4.17). Then, since $p(u_i(s; \tilde{k}); \hat{U}_{-i}(\tilde{k}))$ is strictly increasing, we can obtain that

$$p(u_i(s; k_i + K_i); \hat{U}_{-i}(k_i + K_i)) > p(u_i(s; k_i); \hat{U}_{-i}(k_i)).$$

Thus, we have

$$u_i^*(k_i + K_i) - u_i^*(k_i) = 2\alpha_i^* < \bar{u}^*(k_i + K_i) - \bar{u}^*(k_i).$$

In addition, we have $\bar{u}^*(k_i + K_i) - \bar{u}^*(k_i) < \bar{u}(k_i + K_i) - \bar{u}(k_i)$. However, $U^*(\tilde{k})$ is the Nash equilibrium and satisfies $\bar{u}^*(\tilde{k}) = \bar{u}(\tilde{k})$. There is a contradiction. Thus, we have $\bar{u}^*(k_i) \geq \bar{u}^*(k_i+K_i)$ and $D(k_i)+\bar{u}^*(k_i) \leq D(k_i+K_i)+\bar{u}^*(k_i+K_i)$, if $D(k_i) \leq D(k_i+K_i)$. The conditions are satisfied. The proof is complete. □

4.4 Simulation Results

In this section, the performance of the proposed method is demonstrated by a numerical example. Values of the parameters related to vehicle dynamics are illustrated in Table 4.1. The information of PEVs, utilized in the charging control problem, is shown in Figure 4.1-4.4. Figure 4.1 and Figure 4.2 show that the number of PEVs arrival and departure during the charging time, respectively. Figure 4.3 represents the total number of vehicles that are connected to the grid during the charging time. Figure 4.4 illustrates the desired SoC of each vehicle for charging at the initial time instant. Thus, the control inputs at the sampling time $\tilde{k}$ can be calculated.

We use the residential power load profiles [161] as the baseline for the charging prob-

Table 4.1: The values of parameters related to vehicle dynamics and charging control problems.

Parameters	Values
$SoC_{i,des}$	0.8 - 1.0
$SoC_{i,ini}$	0.1 - 0.3
ϵ_{stop}	0.06
the i-th PEV battery capacity $\overline{E}_i$	20 kWh
maximum charging power $\overline{P}_i$	6.6×2 kW
charging efficiency η	85%
generation cost function of electric energy price $c(D_{\tilde{k}})$	$c(D_{\tilde{k}}) = 5.8 \times 10^{-7} D_{\tilde{k}}^2 + 0.06 D_{\tilde{k}}$
price function $p(D_{\tilde{k}})$	$p(D_{\tilde{k}}) = 11.6 \times 10^{-7} D_{\tilde{k}} + 0.06$
begin time of charging k	k is the time instant at 3 : 00 PM
time period under the proposed control algorithm K	$K = 18$h

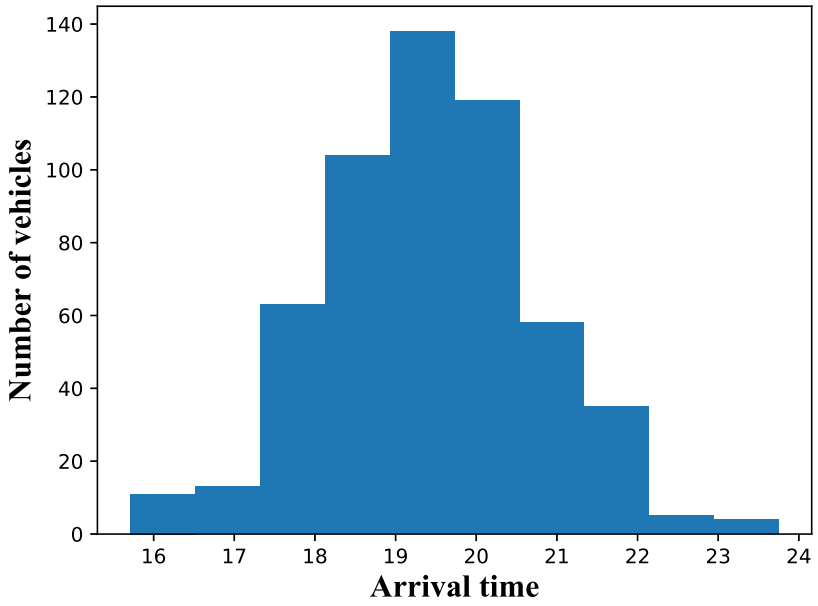


Figure 4.1: The number of PEVs plug-in to the power grids. The count starts at 3:00 PM and lasts for 9 hours.

lem. Since the charging controller of residential areas considered in this work is to fill the overnight valley, the simulation time (18 hours) is from 3:00 PM to the following day at 9:00 AM. Figure 4.5 shows the power load profiles of the charging problem. The red dashed line represents the limitation of the power grids' loads. The green dashed line illustrates the baseline of the power load, which demonstrates the profile without PEV charging

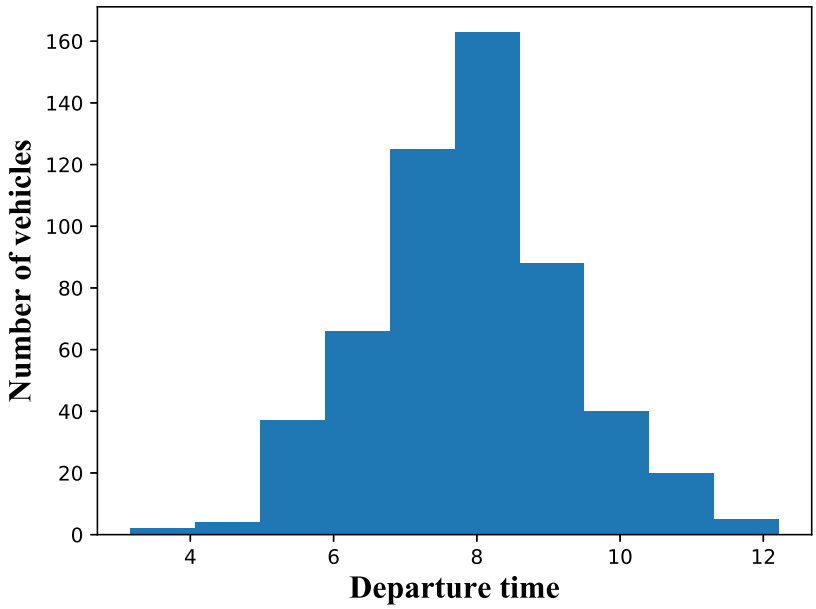


Figure 4.2: The number of PEVs plug-out of the power grids. The count starts at 3:00 AM in the following morning and lasts 9 hours.

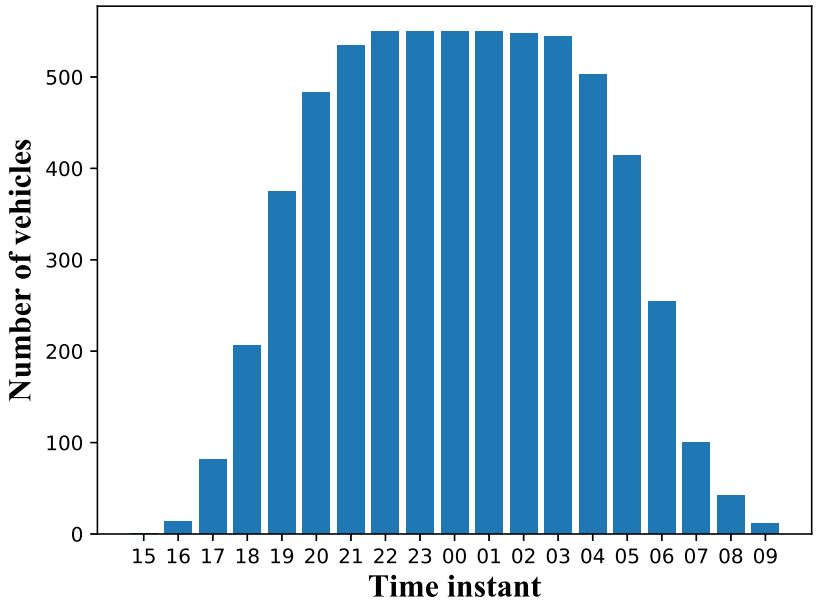


Figure 4.3: The number of PEVs can be used for charging control. The count for the simulation starts at 3:00 PM and lasts for 18 hours.

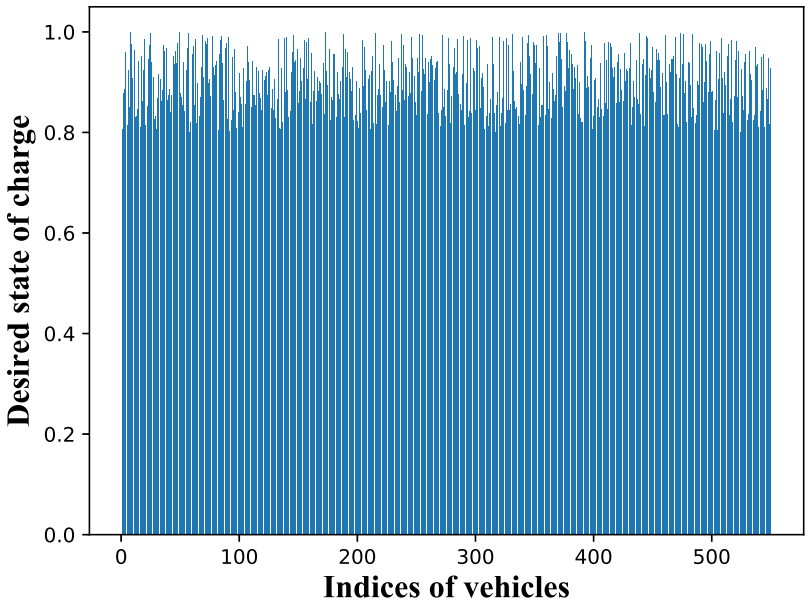


Figure 4.4: The desired state of charge of each PEV for charging control.

requirements. The blue dashed line represents the power load profile without achieving the PEV charging objective. The two yellow lines denote the PEVs charging power load profiles before (dashed line) and after (solid line) applying the charging control rates, respectively. One of the main objectives of the charging control problem is to avoid overload of the grid by managing the charging time of each PEV. The power load of PEV charging significantly impacts the grid, which causes safety problems for the grid, as shown using the orange line. As shown in Figure 4.5, the proposed method provides a possible solution to the valley-filling objective and can meet the charging requirements of each PEV.

4.5 Conclusion

In this work, we studied a decentralized model predictive charging control algorithm, supplemented by the iterative learning method for the PEV charging problem. To improve the safety of the power grids and meet individual charging requirements, the sampled safe set of charging was involved in the optimal control problem. The proposed framework generated the optimal control inputs with the updating price function iteratively to estimate the electricity price cost. There were several limitations of this work. For example, the battery needed to be replaced regularly since the battery aging problem appeared in any battery.

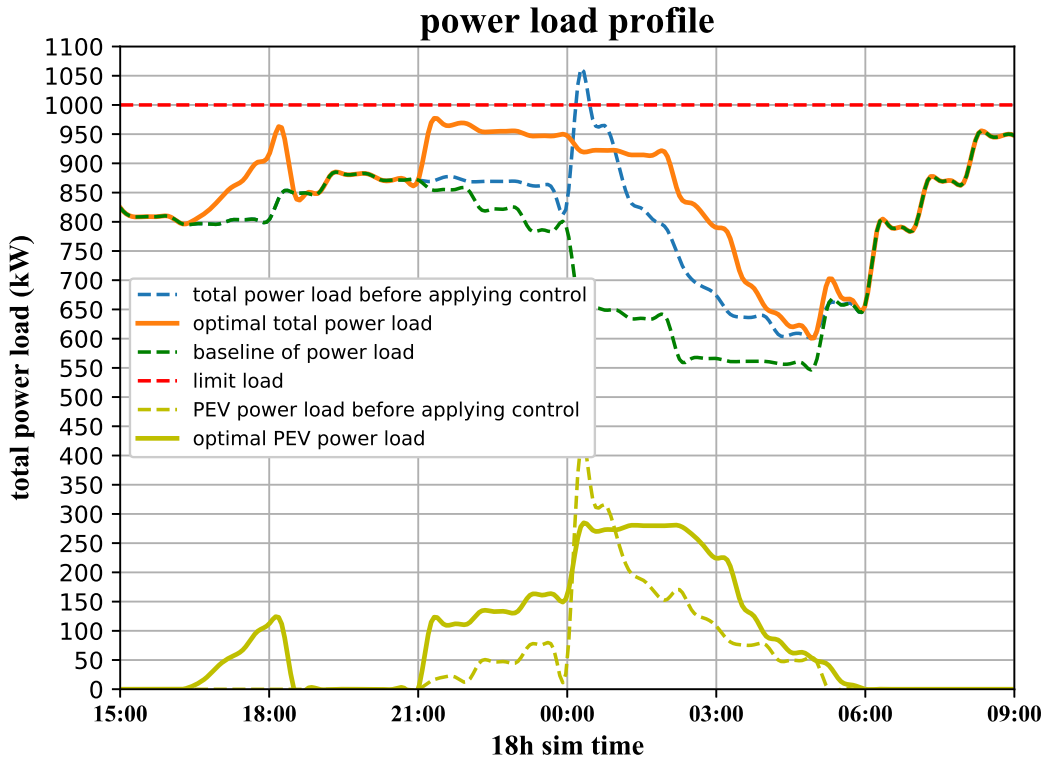


Figure 4.5: The power load profiles of the charging control

The cost of power spikes, voltage drop and battery aging model would be left for future work.

Chapter 5

Decentralized MPC for PEV Charging Control with Battery Aging: An Iterative Learning-Based Method

5.1 Introduction

In Chapter 4, we have investigated iterative learning-based decentralized model predictive charging control to address the valley-filling problem. Yet, there are some practical issues that need to be considered. As a simple example, batteries need to be replaced periodically due to battery aging problems. The capacity of a battery usually degrades gradually over time during charging and discharging. Among these issues, the battery aging factor should be incorporated into the design of the charging control scheme. Following the discussion in Chapter 4, we further improve the formulated optimal control problem by involving the battery capacity degradation function. In this chapter, we focus on the recursive feasibility of the modified optimal control problem, the stability of the closed-loop system and the convergence of the proposed algorithm, based on the same charging dynamics and sampled safe set as ones in Chapter 4.

Looking into control methods for the charging processes of PEVs in power grids [162, 163], one goal is to reduce negative effects on power systems. Aside from impacts of PEV charging on power systems, it is also important to consider the battery's longevity. The determinants of batteries' longevity is their charging range and lifetime. The demand for output and input power makes the battery discharge and charge, which causes battery aging problems, resulting in severe capacity degradation of batteries in PEVs [164]. For example,

a four-year 60 kWh battery is as effective as a 54 kWh battery. The maximum driving range of the PEV with a 60 kWh battery is around 420 km, while the one with a 54 kWh battery is only about 350 km [165].

Battery aging problem is always complex [166–168] since it is a result of ages, cycle counts, charging speeds, charging levels and temperatures [169]. All these effects may occur simultaneously, and some of them dominate the aging mechanism and lead to a faster battery capacity degradation [170]. Theoretically, a full charge and full discharge cycle maximizes the energy efficiency. However, at this point, batteries work under high pressures, and their capacities accelerate abate, which reduces the driving range and lifespan of batteries. Recently, the favored solution to battery charging problems is only utilizing 60% of the battery capacity. The battery is charged to 80% and stops discharging at 20%. As the capacity decreases, the battery management service will require a higher charge bound and a lower discharge one to approach the driving range until the driving range shrinks, obviously.

Briefly speaking, there are two solutions to the battery aging problem. The first one is applying a hybrid energy storage system for compensating the cost of battery aging. Ref. [171] proposes a topology and power management of the multi-mode hybrid energy storage system for EVs based on the rule-based control strategy and the power-balancing strategy. Besides, [172] provides a novel battery hybrid energy storage system to extend the driving range of EVs, in which the battery is not used to collect energy from the regenerative braking directly.

However, the vehicle needs to carry more battery cells using a hybrid energy storage system and applying a co-design control method. Another method is to integrate the battery capacity degradation cost into the cost function and constraints of the optimal control problem [173, 174]. In addition, [175] sets up an optimization-based smart charging scheme in a microgrid, which aims to minimize energy costs, battery aging effects and CO_2 emissions, and to maximize power grids' efficiency. To adequately combine with the perspectives of all stakeholders, multi-objective decision processes are required. Although a 28.1% reduction of battery aging is reported for some cases, the applied battery model neither considers calendar aging nor the batteries' thermal behavior. In [176], a smart charging application using an empirical battery aging model is set up, which considers electro-thermal influences. The case study for validation, however, neglects important factors from an EV user perspective, such as energy and battery aging costs. For PHEVs, the bidirectional charging can be considered to further reduce the operating cost of PHEVs considering dynamic energy tariffs and battery aging cost [177]. Also, Bai *et al.*, in [178], propose a two-level

optimization energy management strategy to suppress the battery aging for PHEVs, with one for the power distribution by a variable-threshold Dynamic Programming (DP) algorithm and the other for adaptive power allocation based on the low-pass filtering algorithm and power management module.

Another challenge from the PEV charging is how to minimize energy consumptions, electricity price costs, etc., while meeting individual requirements. MPC is a promising solution to this problem since it allows to integrate the energy consumptions and forecast the future charging scenarios. In [57], an economic MPC scheme is developed to minimize the cost of energy consumption for an individual PEV. In [35], Tang *et al.* present an MPC-based algorithm for the optimal vehicle charging scheduling, based on the prediction of the future vehicle arrivals.

Several approaches using the MPC framework show their efficiency for PEV charging problem, such as Lagrangian Decomposition [179, 180], Alternating Direction Method of Multipliers (ADMM) [181], Monte Carlo Markov chain [182]. Ref. [183] proposes a hierarchical control algorithm for regenerative braking problem considering battery aging using MPC and pneumatic pressure control methods. Ref. [184] presents an MPC method using the DP algorithm for a simple battery aging model in the formulations, which studies the effects of combined driving and V2G safe usage on the lifetime performance of relevant commercial Li-ion batteries. Due to the complexity of the electricity price cost and battery aging model, a possible solution to overcome the difficulty is to estimate and update the price function and battery capacity degradation, involved in the optimal control problem, under an iterative learning framework. Inspired by the work [159], the iterative learning method can be utilized to approximate the price and battery capacity degradation function under the MPC framework for PEV charging.

The main idea of the valley-filling performance is to use all possible PEVs charging inputs to compensate for lower power load conditions. In fact, the battery aging problem of PEVs limits the actual charging rate. However, based on existing results on the valley-filling problem, the battery aging effects are usually ignored because they are influenced by different operating conditions, such as charging level, usage frequency, operating temperature, life cycle, driving conditions, etc. Although this is for engineering convenience, an algorithm integrating battery degradation and price functions in the optimization problem is desired for practice. Motivated by this observation, we propose an iterative learning-based decentralized model predictive charging control algorithm to solve the valley-filling problem while considering battery capacity degradation. All PEVs' conditions are assumed to be the same even when the battery degradation is taken into account.

The control objective is to guarantee valley-filling performance while meeting individual charging requirements. Generally speaking, the proposed solution consists of two steps: (1) updating the battery capacity degradation cost and price function, and (2) optimizing the charging rates. The first step focuses on approximating the battery capacity degradation cost and price function that is then utilized to update the cost function and constraints of the optimization problem in the second step. The second step is to generate optimal charging inputs by solving the optimization problem individually under a decentralized MPC structure.

The main contributions of this work are as follows:

- A charging control algorithm is proposed for PEVs to address the valley-filling problem, where the battery capacity degradation is taken into account. The proposed algorithm, which includes the calculation of sampled safe set, price function and battery capacity degradation function, can generate the control inputs for the PEV charging problem. The sampled safe set, the price function and battery capacity degradation function are updated after solving the optimal control problem iteratively.

- This work is constructed based on a decentralized charging architecture within multiple local controllers using MPC method. The studies on the recursive feasibility of the proposed optimal control problem and the stability of the closed-loop system are investigated based on the Nash equilibrium.

The remainder of this chapter is organized as follows. Section 5.2 formulates the valley-filling problem using MPC framework, supplemented by the iterative learning method considering battery capacity degradation. Section 5.3 presents the main theoretical results, including the recursive feasibility of the proposed optimal control problem and the stability of the closed-loop system. Section 5.4 gives an example to verify the theoretical results. The conclusion of this chapter is given in Section 5.5.

5.2 Problem Formulation and Algorithm Design

The main objectives of integrating the battery aging characteristics into the charging control scheme design are four-fold: (1) to achieve the valley-filling performance; (2) to minimize the energy cost and the deviation of power demands; (3) to reduce the negative impacts caused by battery aging problem; (4) to meet the safety and security requirements of the power grids.

5.2.1 PEV Charging Control Using Optimization Method

We consider a decentralized charging control strategy for the valley-filling problem with the population size of PEVs n over time horizon K. For the i-th individual PEV ($i \in \mathbb{N} := \{1, 2, \cdots, n\}$), the obtained SoC of PEV at the time instant $\tilde{k}$ is given by $SoC_i(\tilde{k}) \in [0, 1]$. Define $SoC_{i,ini}$ and $SoC_{i,des}$ as the initial SoC and desired SoC for the i-th PEV, respectively. When the vehicle arrives, the charged energy for the i-th PEV $E_{i,r} = \bar{E}_i(SoC_{i,des} - SoC_{i,ini})$ can be obtained. $\bar{E}_i$ is the maximum battery capacity of the i-th PEV. Let $x_i(\tilde{k}) = E_{i,r} - SoC_i(\tilde{k})\bar{E}_i$ denote the energy of the i-th PEV to be charged at time instant $\tilde{k}$. According to the charging dynamics, the following system model can be derived,

$$x_i(\tilde{k}+1) = x_i(\tilde{k}) + B_i u_i(\tilde{k}), \quad B_i = -\eta_i \Delta t \bar{P}_i, \tag{5.1}$$

where $u_i(\tilde{k}) \in \mathbb{R}$ is the charging rate of the i-th PEV, and it is the control input signal to be designed and needs to satisfy the physical constraint $-1 \leq u_i(\tilde{k}) \leq 1$ and $\sum_{\tilde{k}=k_i}^{k_i+K_i} u_i(\tilde{k}) = \gamma_i$, where $\gamma_i = -\frac{x_i(k_i)}{B_i}$. η_i, Δt and $\bar{P}_i$ represent the charging efficiency, sampling time period and maximum charging power, respectively.

The valley-filling charging control algorithm will start at time k and stop at time $k+K$. Similarly, k_i and $k_i + K_i$ are the start and end time of the i-th PEV charging process, respectively. The i-th PEV charging process does not exceed the valley-filling charging control time period, i.e., $\tilde{k} \in [k, k+K]$, $\bigcup_{i=1}^{n}[k_i, k_i + K_i] \subseteq [k, k+K]$. We assume that the control input signal is zero when the vehicle is not charging. As a result, the system state can be denoted as $x_i(\tilde{k}) = E_{i,r}$ when $\tilde{k} \in [k, k_i]$, and $x_i(\tilde{k}) = 0$ when $\tilde{k} \in [k_i + K_i, k+K]$. Furthermore, $x_i(k_i + K_i) = 0$ definitely guarantees that the i-th PEV is fully charged.

Based on the PEV charging dynamics presented above, we focus on the design of the decentralized charging controller, aiming to minimize the variance of the total power load profile by concerning the baseline of power demand. The calculated charging control input sequence for the i-th PEV at time $\tilde{k}$ will be calculated, as $\mathbf{u}_i(s; \tilde{k}) = [u_i(\tilde{k}; \tilde{k}), u_i(\tilde{k}+1; \tilde{k}), \cdots, u_i(k_i + K_i; \tilde{k})]$, $s \in [\tilde{k}, k_i + K_i]$. $u_i(\tilde{k}; \tilde{k})$ stands for the control input that will be applied to the i-th PEV at time $\tilde{k}$. The cost function of the optimization problem is constructed as

$$\begin{aligned} J_i(\mathbf{u}_i(s; \tilde{k}); \hat{U}_{-i}(\tilde{k}); x_i(\tilde{k})) = \sum_{s=\tilde{k}}^{\tilde{k}+K_i} \Big(\sum_{i \in \mathbb{N}} \big(p(u_i(s; \tilde{k}); \hat{U}_{-i}(\tilde{k})) D_{\tilde{k}} + f_i(u_i(s; \tilde{k})) \big) \\ + \delta(u_i(\tilde{k}; \tilde{k}) - \bar{u}(\tilde{k}))^2 \Big) + x_i(\tilde{k})^2 + Q_i(\hat{x}_i(\tilde{k} + K_i | \tilde{k})), \end{aligned} \tag{5.2}$$

where δ is an adjustable parameter. $\mathbf{u}_i(s;\tilde{k})$, $s \in [\tilde{k}, k_i + K_i]$ will send to the aggregator as the predicted control input to calculate the average value for the optimization problem of the next time instant, i.e., $\hat{u}_i(\tilde{k}) = u_i(\tilde{k};\tilde{k}-1)$. In this work, we formulate $c(D_{\tilde{k}}) = \frac{1}{2}aD_{\tilde{k}}^2 + bD_{\tilde{k}}$ and $p(u_i(s;\tilde{k});\hat{U}_{-i}(\tilde{k})) = aD_{\tilde{k}} + b$, as well; $D_{\tilde{k}} = D(\tilde{k}) + \sum_{\tilde{i}\in\mathbb{N}/i} \hat{u}_{\tilde{i}}(\tilde{k}) + u_i(s;\tilde{k})$, where $D(\tilde{k})$ represents the power load baseline at time $\tilde{k}$. The battery degradation function is formulated by $f_i(u_i(s;\tilde{k})) = a_i u_i^2(s;\tilde{k}) + b_i u_i(s;\tilde{k}) + c_i$. a_i, b_i and c_i are constants, but they are different among each PEV.

5.2.2 Sampled Safe Set for Battery SoC

According to the sampled safe set defined in 4.2.2, we define the sampled safe set for the charging problem in this work at the j-th iteration as

$$SS_i^j = \Big\{ \bigcup_{m\in M_i^j} \bigcup_{\tilde{k}=k_i}^{k_i+K_i} \hat{x}_i^m(\tilde{k}) \Big\}, \tag{5.3}$$

where $M_i^j = \{m \in [0, j] \mid \hat{x}_i^m(k_i + K_i) = 0\}$. M_i^j stands for the collection of indices associated with successful iterations m of the i-th PEV. $\hat{x}_i^m(\tilde{k})$ is the predicted system state at iteration m of the i-th PEV at time $\tilde{k}$. SS_i^j is the collection of all state trajectories at iteration $m \in M_i^j$. Thus, when $m \leq j$, we have $M_i^m \subset M_i^j$. When $j = 0$, we assume that $SS_i^0 = \Big\{ \bigcup_{\tilde{k}=k_i}^{k_i+K_i} x_i^0(\tilde{k}) \Big\}$.

Based on the cost function (5.2), we define the iterative cost at iteration j at time $\tilde{k}$. The cost function is rewritten as

$$\begin{aligned} J_i^j(\mathbf{u}_i^j(s;\tilde{k});\hat{U}_{-i}(\tilde{k});x_i^j(\tilde{k})) = \sum_{s=\tilde{k}}^{\tilde{k}+K_i} \Big(\sum_{i\in\mathbb{N}} \big(p^j(u_i^j(s;\tilde{k});\hat{U}_{-i}(\tilde{k}))D_{\tilde{k}}^j + f_i^j(u_i^j(s;\tilde{k})) \big) \\ + \delta(u_i^j(\tilde{k};\tilde{k}) - \bar{u}(\tilde{k}))^2 \Big) + x_i^j(\tilde{k})^2 + Q_i^{j-1}(\hat{x}_i^{j-1}(\tilde{k}+K_i|\tilde{k})), \end{aligned} \tag{5.4}$$

where $\mathbf{u}_i^j(s;\tilde{k}) = \{u_i^j(\tilde{k};\tilde{k}), u_i^j(\tilde{k}+1;\tilde{k}), \cdots, u_i^j(k_i+K_i;\tilde{k})\}$, $s \in [\tilde{k}, k_i + K_i]$ is the calculated control input sequence at iteration j of the i-th PEV at time $\tilde{k}$. $u_i^j(s;\tilde{k})$ stands for the calculated control input of the i-th PEV at time $\tilde{k}$ for each sampling time instant at iteration j. The price function is updated as

$$p^{j+1}(u_i^j(\tilde{k};\tilde{k});\hat{U}_{-i}(\tilde{k})) = p^j(u_i^j(\tilde{k};\tilde{k});\hat{U}_{-i}(\tilde{k})) + \lambda(c(D_{\tilde{k}}) - p^j(u_i^j(\tilde{k};\tilde{k});\hat{U}_{-i}(\tilde{k}))), \tag{5.5}$$

where $\lambda \in [0,1]$ is the learning parameter.

5.2.3 Decentralized Model Predictive Charging Control Algorithm Design

The following constrained optimal control problem is formulated using the model and the sampled safe set which are discussed in Section 5.2.1 and Section 5.2.2, respectively.

Problem 5. *(Optimal Control Problem (OCP))*

$$\begin{aligned}
\min_{\mathbf{u}_i} \; J_i^j(\mathbf{u}_i^j(s;\tilde{k});\hat{U}_{-i}(\tilde{k});x_i^j(\tilde{k})) = &\sum_{s=\tilde{k}}^{\tilde{k}+K_i} \Big(\sum_{i\in\mathbb{N}} \big(p^j(u_i^j(s;\tilde{k});\hat{U}_{-i}(\tilde{k}))D_{\tilde{k}}^j + f_i^j(u_i^j(s;\tilde{k}))\big) \\
&+ \delta(u_i^j(\tilde{k};\tilde{k}) - \bar{u}(\tilde{k}))^2\Big) + x_i^j(\tilde{k})^2 \\
&+ Q_i^{j-1}(\hat{x}_i^{j-1}(\tilde{k}+K_i|\tilde{k})) \qquad (5.6)\\
\textbf{\textit{s.t.}}\; \hat{x}_i^j(\tilde{k}+1) = &\hat{x}_i^j(\tilde{k}) + B_i u_i^j(\tilde{k};\tilde{k}), \qquad (5.7)\\
p^{j+1}(u_i^j(\tilde{k};\tilde{k});\hat{U}_{-i}(\tilde{k})) = &p^j(u_i^j(\tilde{k};\tilde{k});\hat{U}_{-i}(\tilde{k})) + \lambda(c(D_{\tilde{k}}) - p^j(u_i^j(\tilde{k};\tilde{k});\hat{U}_{-i}(\tilde{k}))) \\
&\qquad (5.8)\\
u_i^j(s;\tilde{k}) \in &[-1,1],\; \tilde{x}_i^j(\tilde{k}+1) \in SS_i^{j-1}, \qquad (5.9)\\
x_i^j(k_i+K_i) = &0, \qquad (5.10)
\end{aligned}$$

where $s \in [\tilde{k}, k_i+K_i]$, $B_i = -\eta_i \Delta t \bar{P}_i$ *and* $D_{\tilde{k}} = D(\tilde{k}) + \sum_{\tilde{i}\in\mathbb{N}/i} \hat{u}_{\tilde{i}}(\tilde{k}) + u_i(\tilde{k};\tilde{k})$.

Due to the coordination of several local controllers to optimize the sum of all power loads, the optimal control input sequence $\mathbf{u}_i^*(s;\tilde{k}) = \{u_i^*(\tilde{k};\tilde{k}), u_i^*(\tilde{k}+1;\tilde{k}), \cdots, u_i^*(k_i+K_i;\tilde{k})\}$, $s \in [\tilde{k}, k_i+K_i]$, of the i-th PEV obtained after the convergence of the algorithm. $u_i^*(s;\tilde{k})$, $s \in [\tilde{k}, k_i+K_i]$ stands for the control input of the i-th PEV at time $\tilde{k}$ obtained after the convergence of the algorithm. $\hat{x}_i^{j-1}(\tilde{k}+K_i|\tilde{k})$ denotes the corresponding predicted terminal system state at $(j-1)$-th iteration. $Q_i^{j-1}(\hat{x}_i^{j-1}(\tilde{k}+K_i|\tilde{k}))$ is the terminal cost. The optimal control input of the above optimal control problem can minimize the deviation of the sum of all power loads to realize the valley-filling performance and reduce the price cost. The decentralized control algorithm is presented in Algorithm 4.

Algorithm 4: Iterative Learning-Based Decentralized MPC for the Valley-Filling Problem with Battery Aging

Input: Scheduling horizon $[k, k+K]$; Base load profile D; Number of PEVs n; Maximum iteration j_{max}; Tolerance ϵ_{stop}

Output: Charging control input for each PEV $u_i(\tilde{k})$, $i \in \mathbb{N}$

Set $\tilde{k} = k$;
while $\tilde{k} \neq k+K$ **do**
  The operator generates $D(\tilde{k})$;
  The operator broadcasts $D(\tilde{k})$ and $\bar{u}(\tilde{k})$ to all PEVs;
  for *Each PEV* i **do**
    if $\tilde{k} \notin [k_i, k_i + K_i]$ **then**
      $u_i(\tilde{k}; \tilde{k}) = 0$;
    else
      Set $j = 0$;
      Set $\epsilon = 0$;
      while $\epsilon \geq \epsilon_{stop}$ *and* $j \leq j_{max}$ **do**
        if *Problem 5 is feasible* **then**
          Compute $\mathbf{u}_i^*(s; \tilde{k})$;
          Update $p^{j+1}(u_i(\tilde{k}; \tilde{k}); \hat{U}_{-i}(\tilde{k}))$;
          Compute $\epsilon = p^{j+1}(u_i(\tilde{k}; \tilde{k}); \hat{U}_{-i}(\tilde{k})) - p^{j}(u_i(\tilde{k}; \tilde{k}); \hat{U}_{-i}(\tilde{k}))$;
          Update $SS_i^j = \left\{ \bigcup_{m \in M_i^j} \bigcup_{\tilde{k}=k_i}^{k_i+K_i} x_i^m(\tilde{k}) \right\}$;
        end
        Set $j = j + 1$;
      end
    end
  end
  The operator collects $\mathbf{u}_i^*(s; \tilde{k})$ and updates $\bar{u}(\tilde{k}+1)$;
  for *Each PEV* i **do**
    Implement the charging control at time $\tilde{k}$;
    Update the state $x_i(\tilde{k}+1)$ according to (5.1);
  end
  Set $\tilde{k} = \tilde{k} + 1$;
end

5.3 System Analysis

We discuss the existence and uniqueness of the Nash equilibrium in Section 4.3 for the valley-filling problem under iterative learning-based decentralized MPC framework. In this section, we will study the convergence of the proposed algorithm, the recursive feasibility

of the optimal control problem and the stability of the closed-loop systems.

5.3.1 Convergence

In this subsection, the optimal control inputs, generated by the proposed controller, and corresponding state converges to the steady-state trajectory. Thus, we need to show: (1) the cost function of j-th iteration is non-creasing as the index j increases; (2) the steady-state trajectory is a locally optimal solution to the approximation of the optimal control problem.

Theorem 3. *Given the charging system* (5.1) *with the proposed controller and the sampling safe set SS^j at the j-th iteration. Suppose that the assumptions hold. The optimal value of the cost function is non-increasing with the iteration index j.*

Proof. To simplify the representation of the equations, we use $J_i^j(x_i^j(k_i))$ as the cost function $J_i^j(\mathbf{u}_i^j(s;\tilde{k});\hat{U}_{-i}(\tilde{k});x_i^j(\tilde{k}))$ for short when we discuss the system states of the cost function. Firstly, we need to find the lower bound of the cost value of the j-th iteration.

$$
\begin{aligned}
&J_i^{j-1}(x_i^{j-1}(k_i))\\
=&\sum_{s=k_i}^{k_i+K_i}\Big(p^{j-1}(u_i^{*,j-1}(s;k_i);\hat{U}_{-i}(k_i))D_{k_i}^{*,j-1}+\delta(u_i^{*,j-1}(k_i)-\bar{u}^{j-1}(k_i))^2\\
&+f_i^{j-1}(u_i^{*,j-1}(s;k_i))\Big)+\sum_{s=k_i+K_i+1}^{k+K}\Big(p^{j-1}(u_i^{*,j-1}(s;k_i);\hat{U}_{-i}(k_i))D_{k_i}^{*,j-1}\\
&+\delta(u_i^{j-1}(k_i)-\bar{u}^{j-1}(k_i))^2+f_i^{j-1}(u_i^{*,j-1}(s;k_i))\Big)+x_i^{j-1}(k_i)^2\\
\geq&\sum_{s=k_i}^{k_i+K_i}p^{j-1}(u_i^{*,j-1}(s;k_i);\hat{U}_{-i}(k_i))D_{k_i}^{*,j-1}+\delta(u_i^{j-1}(k_i)-\bar{u}^{j-1}(k_i))^2+f_i^{j-1}(u_i^{*,j-1}(s;k_i))\\
&+x_i^{j-1}(k_i)^2+Q_i^{j-1}(\hat{x}_i^{*,j-1}(k_i+K_i|k_i))\\
\geq&\min_{\mathbf{u}_i}\Big[\sum_{s=k_i}^{k_i+K_i}p^{j}(u_i^{*,j}(s;k_i);\hat{U}_{-i}(k_i))D_{k_i}^{*,j}+\delta(u_i^{j}(k_i)-\bar{u}^{j}(k_i))^2+f_i^{j-1}(u_i^{*,j-1}(s;k_i))\\
&+x_i^{j}(k_i)^2+Q_i^{j-1}(\hat{x}_i^{*,j-1}(k_i+K_i|k_i))\Big]\\
=&J_i^j(x_i^j(k_i))
\end{aligned}
$$

where

$$
Q_i^{j-1}(\hat{x}_i^{*,j-1}(k_i+K_i|k_i))=\sum_{s=k_i+K_i+1}^{k+K}\Big(p^{j-1}(u_i^{*,j-1}(s;k_i);\hat{U}_{-i}(k_i))D_{k_i}^{*,j-1}
$$

$$+ f_i^{j-1}(u_i^{*,j-1}(s;k_i)) + \delta(u_i^{j-1}(k_i) - \bar{u}^{j-1}(k_i))^2\Big).$$

Then, we realize that $J_i^j(x_i^j(k_i))$ also is upper bounded the trajectory.

$$\begin{aligned}
&J_i^j(x_i^j(k_i)) \\
\geq& p^j(u_i^{*,j}(s;k_i);\hat{U}_{-i}(k_i))D_{k_i}^{*,j} + \delta(u_i^j(k_i) - \bar{u}^j(k_i))^2 + f_i^j(u_i^j(s;k_i)) + J_i^j(x_i^j(k_i+1)) \\
\geq& p^j(u_i^{*,j}(s;k_i);\hat{U}_{-i}(k_i))D_{k_i}^{*,j} + \delta(u_i^j(k_i) - \bar{u}^j(k_i))^2 + f_i^j(u_i^j(s;k_i)) \\
&+ p^j(u_i^{*,j}(s;k_i+1);\hat{U}_{-i}(k_i+1))D_{k_i+1}^{*,j} + \delta(u_i^j(k_i+1) - \bar{u}^j(k_i+1))^2 \\
&+ f_i^j(u_i^j(s;k_i+1)) + J_i^j(x_i^j(k_i+2)) \\
\geq& \lim_{\tilde{k}\to k+K}\Big[\sum_{s=k_i}^{k_i+K_i}\Big(p^j(u_i^{*,j}(s;\tilde{k});\hat{U}_{-i}(\tilde{k}))D_{\tilde{k}}^{*,j} + \delta(u_i^j(\tilde{k}) - \bar{u}^j(\tilde{k}))^2 + f_i^j(u_i^j(s;\tilde{k}))\Big) \\
&+ J_i^j(x_i^j(\tilde{k}))\Big]
\end{aligned}$$

Since the Nash equilibrium is asymptotically stable for the closed-loop system and the $p^j(u_i^j(s;\tilde{k});\hat{U}_{-i}(\tilde{k}))D_{\tilde{k}}^j + \delta(u_i^j(\tilde{k}) - \bar{u}^j(\tilde{k}))^2 + f_i^j(u_i^j(s;\tilde{k}))$ is continue, the following equation holds.

$$\lim_{\tilde{k}\to k+K} J_i^j(x_i^j(k_i)) = J_i^j(x_i^j(k_i+K_i)) = 0 \tag{5.11}$$

Then, we can obtain that

$$\begin{aligned}
J_i^j(x_i^j(k_i)) \geq& \sum_{s=k_i}^{k_i+K_i}\Big(p^j(u_i^{*,j}(s;\tilde{k});\hat{U}_{-i}(\tilde{k}))D_{\tilde{k}}^{*,j} + \delta(u_i^{*,j}(\tilde{k}) - \bar{u}^j(\tilde{k}))^2 + f_i^j(u_i^{*,j}(s;\tilde{k}))\Big) \\
&+ x_i^j(\tilde{k})^2 + Q_i^j(\hat{x}_i^{*,j}(k_i+K_i|k_i)) \\
=& J_i^{j+1}(x_i^{j+1}(k_i))
\end{aligned} \tag{5.12}$$

Consequently, the iteration cost is decreasing as the j increasing.

$$J_i^{j-1}(x_i^{j-1}(k_i)) \geq J_i^j(x_i^j(k_i)) \geq J_i^{j+1}(x_i^{j+1}(k_i)) \tag{5.13}$$

The proof is completed. □

Then, we will show that the optimal control inputs, generated by the proposed controller, and corresponding state converges to the steady-state trajectory. Due to the definition of the terminal set and k_i, K_i, k, K, the terminal set of solving the optimal control problem in $[k_i, K_i]$ is a subset of the one in $[k, K]$.

Theorem 4. *Given the charging system (5.1) with the proposed controller and the sampled safe set SS^j at the jth iteration. Suppose that the assumptions hold. Also, assume that the optimal control inputs and the corresponding states converge to the steady-state trajectories $x_i^{j_{max}} = \lim_{j\to j_{max}} x_i^j$ and $u_i^{j_{max}} = \lim_{j\to j_{max}} u_i^j$ as the iteration j increases. If the terminal states in the sampled safe set, the optimal control inputs and the corresponding states are the optimizer of the optimal control problem (Problem 5).*

Proof. We assume that the optimal control inputs and the corresponding states converge to the steady-state trajectories as the iteration j increases. Thus, we can obtain that

$$\begin{aligned}
&J_i(x_i^{j_{max}}(\tilde{k}))\\
\geq& p^{j_{max}}(u_i^{j_{max}}(s;\tilde{k});\hat{U}_{-i}(\tilde{k}))D_{\tilde{k}}^{j_{max}} + \delta(u_i^{j_{max}}(\tilde{k}) - \bar{u}^{j_{max}}(\tilde{k}))^2 + f_i^{j_{max}}(u_i^{j_{max}}(s;\tilde{k}))\\
&+ J_i^{j_{max}}(x_i^{j_{max}}(\tilde{k}+1))\\
\geq& p^{j_{max}}(u_i^{j_{max}}(s;\tilde{k});\hat{U}_{-i}(\tilde{k}))D_{\tilde{k}}^{j_{max}} + \delta(u_i^{j_{max}}(\tilde{k}) - \bar{u}^{j_{max}}(\tilde{k}))^2 + f_i^{j_{max}}(u_i^{j_{max}}(s;\tilde{k}))\\
&+ p^{j_{max}}(u_i^{j_{max}}(s;\tilde{k}+1);\hat{U}_{-i}(\tilde{k}+1))D_{\tilde{k}+1}^{j_{max}} + \delta(u_i^{j_{max}}(\tilde{k}+1) - \bar{u}^{j_{max}}(\tilde{k}+1))^2\\
&+ f_i^{j_{max}}(u_i^{j_{max}}(s;\tilde{k}+1)) + J_i^{j_{max}}(x_i^{j_{max}}(\tilde{k}+2))\\
\geq& \sum_{h=0}^{K_i} p^{j_{max}}(u_i^{j_{max}}(s;\tilde{k}+h);\hat{U}_{-i}(\tilde{k}+h))D_{\tilde{k}+h}^{j_{max}} + f_i^{j_{max}}(u_i^{j_{max}}(s;\tilde{k}+h))\\
&+ \delta(u_i^{j_{max}}(\tilde{k}+h) - \bar{u}^{j_{max}}(\tilde{k}+h))^2 + J_i^{j_{max}}(x_i^{j_{max}}(\tilde{k}+K_i))
\end{aligned} \tag{5.14}$$

Besides, the terminal cost converges at a steady-state $J_i^{j_{max}}(x_i^{j_{max}}(\tilde{k}+K_i)) = x_i^{j_{max}}(k_i+K_i)^2$. According to Eq. (5.14), we have

$$\begin{aligned}
J_i^{j_{max}}(x_i^{j_{max}}(\tilde{k})) \geq& \sum_{h=0}^{K_i} p^{j_{max}}(u_i^{j_{max}}(s;\tilde{k}+h);\hat{U}_{-i}(\tilde{k}+h))D_{\tilde{k}+h}^{j_{max}}\\
&+ \delta(u_i^{j_{max}}(\tilde{k}+h) - \bar{u}^{j_{max}}(\tilde{k}+h))^2 + f_i^{j_{max}}(u_i^{j_{max}}(s;\tilde{k}+h)) + x_i^{j_{max}}(k_i+K_i)^2
\end{aligned} \tag{5.15}$$

Consequently, the cost associated with feasible input and state trajectories $x_i^{j_{max}}(\tilde{k}+K_i)$ and $u_i^{j_{max}}(\tilde{k}+K_i)$ are the lower bound of the optimal cost $J_i^{j_{max}}(x_i^{j_{max}}(\tilde{k}))$. Thus, $x_i^{j_{max}}$ and $u_i^{j_{max}}$ is the optimizer of the optimal control problem (Problem 5) as the iteration index j increases.

Then, we need to prove that $x_i^{j_{max}}(\tilde{k}+K_i+1)$ and $u_i^{j_{max}}(\tilde{k}+K_i+1)$ also are the optimal solution to the optimal control problem (Problem 5) with the initial state $x_i(\tilde{k}) = x_i^{j_{max}}(\tilde{k})$. The corresponding optimal cost is $J_i^{j_{max}}(x_i^{j_{max}}(\tilde{k}))$. Due to the time invariance of the optimal control problem, we begin with the initial condition $x_i(k_i) = x_i^{j_{max}}(k_i)$ at

time $\tilde{k} = k_i$.

By contradiction, we assume that the optimal solution to Problem 5 is $\tilde{x}_i^{j_{max}}(K_i+1)$ and $\tilde{u}_i^{j_{max}}(K_i+1)$ are different from $x_i^{j_{max}}(K_i+1)$ and $u_i^{j_{max}}(K_i+1)$.

If we have multiple feasible trajectories, for $\mu \in (0,1)$, $\hat{\mathbf{x}}_i^{j_{max}}(K_i) = \{x_i^{j_{max}}(k_i), \cdots, x_i^{j_{max}}(k_i), \mu x_i^{j_{max}}(\tilde{k}) + (1-\mu)x_i^{j_{max}}(\tilde{k}), x_i^{j_{max}}(\tilde{k}+1), \cdots, x_i^{j_{max}}(K_i+1)\}$ The corresponding inputs are feasible for the Problem 5.

Then we define that

$$J_i^{j_{max}}(x_i^{j_{max}}(K_i+1)) = \sum_{h=0}^{K_i} \Big(p^{j_{max}}(u_i^{j_{max}}(s;\tilde{k}+h);\hat{U}_{-i}(\tilde{k}+h)) D_{\tilde{k}+h}^{j_{max}} + f_i^{j_{max}}(u_i^{j_{max}}(s;\tilde{k}+h)) + \delta(u_i^{j_{max}}(\tilde{k}+h) - \bar{u}^{j_{max}}(\tilde{k}+h))^2 \Big) + x_i^{j_{max}}(K_i+1)^2 \tag{5.16}$$

Since $x_i^{j_{max}}(\tilde{k}+K_i)$ and $u_i^{j_{max}}(\tilde{k}+K_i)$ are the optimal solution to Problem 5, we obtain that

$$\begin{aligned} J_i^{j_{max}}(\hat{\mathbf{x}}_i^{j_{max}}(K_i)) &> J_i^{j_{max}}(x_i^{j_{max}}(K_i)) \\ J_i^{j_{max}}(\hat{\mathbf{x}}_i^{j_{max}}(K_i+1)) &> J_i^{j_{max}}(x_i^{j_{max}}(K_i+1)) \end{aligned} \tag{5.17}$$

Then we have

$$\begin{aligned} &\sum_{h=0}^{K_i} p^{j_{max}}(u_i^{j_{max}}(s;\tilde{k}+h);\hat{U}_{-i}(\tilde{k}+h)) D_{\tilde{k}+h}^{j_{max}} + \delta(\hat{\mathbf{u}}_i^{j_{max}}(\tilde{k}+h) - \bar{u}^{j_{max}}(\tilde{k}+h))^2 \\ &\quad + f_i^{j_{max}}(u_i^{j_{max}}(s;\tilde{k}+h)) + \hat{\mathbf{x}}_i^{j_{max}}(K_i)^2 \\ &> \sum_{h=0}^{K_i} p^{j_{max}}(u_i^{j_{max}}(s;\tilde{k}+h);\hat{U}_{-i}(\tilde{k}+h)) D_{\tilde{k}+h}^{j_{max}} + \delta(u_i^{j_{max}}(\tilde{k}+h) - \bar{u}^{j_{max}}(\tilde{k}+h))^2 \\ &\quad + f_i^{j_{max}}(u_i^{j_{max}}(s;\tilde{k}+h)) + x_i^{j_{max}}(K_i)^2 \end{aligned} \tag{5.18}$$

Thus, we have $\hat{\mathbf{u}}_i^{j_{max}}(\tilde{k}+h) = u_i^{j_{max}}(\tilde{k}+h)$ and $\hat{\mathbf{x}}_i^{j_{max}}(\tilde{k}) = x_i^{j_{max}}(\tilde{k})$. Then we can obtain

$$\begin{aligned} &\sum_{h=0}^{K_i+1} p^{j_{max}}(u_i^{j_{max}}(s;\tilde{k}+h);\hat{U}_{-i}(\tilde{k}+h)) D_{\tilde{k}+h}^{j_{max}} + \delta(\hat{\mathbf{u}}_i^{j_{max}}(\tilde{k}+h) - \bar{u}^{j_{max}}(\tilde{k}+h))^2 \\ &\quad + f_i^{j_{max}}(u_i^{j_{max}}(s;\tilde{k}+h)) + \hat{\mathbf{x}}_i^{j_{max}}(K_i+1)^2 \\ &> \sum_{h=0}^{K_i+1} p^{j_{max}}(u_i^{j_{max}}(s;\tilde{k}+h);\hat{U}_{-i}(\tilde{k}+h)) D_{\tilde{k}+h}^{j_{max}} + \delta(u_i^{j_{max}}(\tilde{k}+h) - \bar{u}^{j_{max}}(\tilde{k}+h))^2 \\ &\quad + f_i^{j_{max}}(u_i^{j_{max}}(s;\tilde{k}+h)) + x_i^{j_{max}}(K_i+1)^2 \end{aligned}$$

which implies

$$J_i^{j_{max}}(\hat{\mathbf{x}}_i^{j_{max}}(K_i+1)) > J_i^{j_{max}}(x_i^{j_{max}}(K_i+1)) \tag{5.19}$$

Since $\hat{\mathbf{u}}_i^{j_{max}}(\tilde{k}+h) = u_i^{j_{max}}(\tilde{k}+h)$ and $\hat{\mathbf{x}}_i^{j_{max}}(\tilde{k}) = x_i^{j_{max}}(\tilde{k})$, we have

$$\begin{aligned}
&p^{j_{max}}(u_i^{j_{max}}(s;\tilde{k});\hat{U}_{-i}(\tilde{k}))D_{\tilde{k}}^{j_{max}} + \delta(\hat{\mathbf{u}}_i^{j_{max}}(\tilde{k}) - \bar{u}^{j_{max}}(\tilde{k}))^2 + f_i^{j_{max}}(u_i^{j_{max}}(s;\tilde{k})) \\
&+ (\hat{\mathbf{x}}_i^{j_{max}}(K_i+1) - x(K_i+1))^2 + J_i^{j_{max}}(\hat{\mathbf{x}}_i^{j_{max}}(K_i+1)) \\
>&p^{j_{max}}(u_i^{j_{max}}(s;\tilde{k});\hat{U}_{-i}(\tilde{k}))D_{\tilde{k}}^{j_{max}} + \delta(u_i^{j_{max}}(\tilde{k}) - \bar{u}^{j_{max}}(\tilde{k}))^2 + f_i^{j_{max}}(u_i^{j_{max}}(s;\tilde{k})) \\
&+ (x_i^{j_{max}}(K_i+1) - x(K_i+1))^2 + J_i^{j_{max}}(x_i^{j_{max}}(K_i+1))
\end{aligned}$$

Thus, we have

$$J_i^{j_{max}}(\hat{\mathbf{x}}_i^{j_{max}}(K_i+1)) > J_i^{j_{max}}(x_i^{j_{max}}(K_i+1)) \tag{5.20}$$

Assume that $\bar{x}_i^{j_{max}}(K_i+1)$ is a convex combination of $x_i^{j_{max}}(K_i+1)$, i.e.,

$$\bar{x}_i^{j_{max}}(K_i+1) = \sum_{s=k_i+1}^{k_i+K_i} \frac{1}{k_i+K_i}\big(\frac{1}{2}\hat{x}_i^{j_{max}}(s) + \frac{1}{2}x_i^{j_{max}}(K_i+1)\big) \tag{5.21}$$

Also, $\bar{x}_i^{j_{max}}(K_i+1)$ can be expressed as based on $\tilde{x}_i^{j_{max}}(K_i+1)$ and $x_i^{j_{max}}(K_i+1)$.

$$\bar{x}_i^{j_{max}}(K_i+1) = \frac{\mu}{2K_i}\tilde{x}_i^{j_{max}}(K_i+1) + \frac{2K_i-\mu}{2K_i}x_i^{j_{max}}(K_i+1) \tag{5.22}$$

Then, we can obtain that

$$J_i^{j_{max}}(x_i^{j_{max}}(K_i+1)) < J_i^{j_{max}}(\bar{x}_i^{j_{max}}(K_i+1)) < J_i^{j_{max}}(\hat{x}_i^{j_{max}}(K_i+1)) \tag{5.23}$$

where the value of $J_i^{j_{max}}(\hat{x}_i^{j_{max}}(k_i))$ is maximum at $\tilde{k}$.

Thus, we have

$$J_i^{j_{max}}(\tilde{x}_i^{j_{max}}(K_i+1)) < J_i^{j_{max}}(\bar{x}_i^{j_{max}}(K_i+1)) < J_i^{j_{max}}(x_i^{j_{max}}(K_i+1)) \tag{5.24}$$

Consequently, we have a contradiction and $x_i^{j_{max}}(\tilde{k}+K_i)$ and $u_i^{j_{max}}(\tilde{k}+K_i)$ are the optimal solution to Problem 5. The proof is completed. □

5.3.2 Recursive Feasibility and Stability

As mentioned in Section 4.3, there is an optimal solution to the optimization control problem for driving the system state to the Nash equilibrium when the point is in the sampling safe set SS_i^j. The properties of SS_i^j and Q_i^j in Section 5.2.2 are provided for recursive feasibility and stability analysis.

Theorem 5. *(Recursive feasibility) For the charging system (5.1) controlled by the proposed control method (Problem 5), with the given sampling safe set of i-th PEV SS_i^j at j-th iteration in Section 5.2.2. Suppose the assumptions hold, the optimization control problem is feasible for all $\tilde{k} \in [k_i, k_i + K_i]$ at every iteration $j \geq 1$.*

Proof. We assume that the sampling safe set at iteration 0, SS_i^0 is not empty. According to the definition in Section 5.2.2, we have that $SS_i^0 \subset SS_i^{j-1}$ for all $j \geq 1$. Hence, the sampling safe set at $(j-1)$-th iteration SS_i^{j-1} is also nonempty. Particularly, we already know that one of the feasible trajectory $\mathbf{x}_{i,0} \in SS_i^0 \subset SS_i^{j-1}$, where $\mathbf{x}_{i,0}$ denotes the initial state trajectory of i-th PEV.

At time $\tilde{k} = k_i$ of j-th iteration, the predictive state trajectory

$$\hat{\mathbf{x}}_{i,0}^j(k_i|k_i) = \{\hat{x}_{i,0}(k_i|k_i), \cdots, \hat{x}_{i,0}(k_i + K_i|k_i)\} \in SS^{j-1}$$

and the related input sequence $\hat{\mathbf{u}}_i^0 = \{\hat{u}_{i,0}^0, \cdots, \hat{u}_{i,N-1}^0\}$, which satisfies the input and state constraints in the optimal control problem (Problem 5). Hence, $\hat{\mathbf{x}}_{i,0}^j(k_i)$ and $\hat{\mathbf{u}}_i^0$ are the feasible solution at time $\tilde{k} = k_i$ of the j-th iteration.

At time $\tilde{k}$ of j-th iteration, we suppose that the optimal control problem (Problem 5) is feasible, and $\mathbf{x}_i^{*,j}$ and $\mathbf{u}_i^{*,j}$ are the optimal state trajectory and input sequence, respectively. The terminal constraint enforces $\hat{x}_i^{*,j-1}(\tilde{k} + K_i|\tilde{k}) \in SS^{j-1}$ and $Q_i^{j-1}(\hat{x}_i^{*,j-1}(\tilde{k} + K_i|\tilde{k}))$. By the definition of $Q_i^{j-1}(\hat{x}_i^{*,j-1}(\tilde{k} + K_i|\tilde{k}))$, we have $\hat{x}_i^m(\tilde{k}) = \hat{x}_i^{*,j-1}(\tilde{k} + 1|\tilde{k})$. Due to the state update method, we have that

$$x_i^j(\tilde{k} + 1) = \hat{x}_i^{*,j-1}(\tilde{k} + 1|\tilde{k}). \tag{5.25}$$

Thus, we have the input sequence at time $\tilde{k}$ of the j-th iteration $\{u_i^{*,j}(\tilde{k}+1|\tilde{k}), \cdots, u_i^{*,j}(\tilde{k}+K_i|\tilde{k})\}$ and the related feasible state trajectory $\{x_i^{*,j}(\tilde{k}+1|\tilde{k}), \cdots, x_i^{*,j}(\tilde{k}+K_i+1|\tilde{k})\}$, which satisfies the input and state constraints in Problem 5. Hence, $\{u_i^{*,j}(\tilde{k}+1|\tilde{k}), \cdots, u_i^{*,j}(\tilde{k}+K_i|\tilde{k})\}$ and $\{x_i^{*,j}(\tilde{k}+1|\tilde{k}), \cdots, x_i^{*,j}(\tilde{k}+K_i+1|\tilde{k})\}$ is the feasible solution for the Problem 5 at time $\tilde{k}+1$. The proof is completed. □

Thus, the recursive feasibility of the optimal control problem is studied. Then, we need to show that the Nash equilibrium is asymptotically stable.

Theorem 6. *(Stability) For the charging system (5.1) controlled by the proposed control method (Problem 5), with the given sampling safe set of i-th PEV $SS_i^{j_{max}}$ at j_{max}-th iteration in Section 4.2.2. Suppose the assumptions hold, the Nash equilibrium is asymptotically stable for the closed-loop system.*

Proof. We need to show that the optimal cost $J_i^{j_{max}}(x_i^{j_{max}}(\tilde{k}))$ is a Lyapunov candidate for the Nash equilibrium. When $x_i^{j_{max}}(\tilde{k}) \in \mathbb{R}\backslash\{x_F\}$, we have that $J_i^{j_{max}}(x_i^{j_{max}}(\tilde{k})) > 0$. When $x_i^{j_{max}}(\tilde{k}) = x_F$, we have that $J_i^{j_{max}}(x_i^{j_{max}}(\tilde{k})) = 0$. With these assumptions, we need to show that $J_i^{j_{max}}(x_i^{j_{max}}(\tilde{k}))$ is non-increasing. Firstly, we know that

$$\hat{x}_i^{*,j_{max}-1}(\tilde{k}+1|\tilde{k}) = x_i^{j_{max}}(\tilde{k}+1). \tag{5.26}$$

Hence, we have that

$$J_i^{j_{max}}(\hat{x}_i^{*,j_{max}-1}(\tilde{k}+1|\tilde{k})) = J_i^{j_{max}}(x_i^{j_{max}}(\tilde{k}+1)). \tag{5.27}$$

Then, we need to show the the cost function is non-increasing with the iteration index j_{max}. Eq. (5.26) implies that

$$J_i^{j_{max}}(\hat{x}_i^{*,j_{max}}(\tilde{k}+1|\tilde{k})) = J_i^{j_{max}}(x_i^{j_{max}}(\tilde{k}+1)) \tag{5.28}$$

Based on the optimal inputs and the corresponding optimal trajectory, the optimal cost is obtained as follows.

$$\begin{aligned}
&J_i^{j_{max}}(x_i^{j_{max}}(\tilde{k})) \\
=&\min_{\mathbf{u}_i}\Big[\sum_{s=k_i}^{k_i+K_i}\Big(p^{j_{max}}(u_i^{j_{max}}(s;\tilde{k});\hat{U}_{-i}(\tilde{k}))D_{\tilde{k}}^{j_{max}} + f_i^{j_{max}}(u_i^{j_{max}}(s;\tilde{k})) \\
&+\delta(u_i^{j_{max}}(\tilde{k}) - \bar{u}^{j_{max}}(\tilde{k}))^2\Big) + x_i^{j_{max}}(\tilde{k})^2 + Q_i^{j_{max}-1}(\hat{x}_i^{j_{max}-1}(\tilde{k}+K_i|\tilde{k}))\Big] \\
=&\sum_{s=\tilde{k}+1}^{k_i+K_i}\Big(p^{j_{max}}(u_i^{j_{max}}(s;\tilde{k});\hat{U}_{-i}(\tilde{k}))D_{\tilde{k}}^{j_{max}} + f_i^{j_{max}}(u_i^{j_{max}}(s;\tilde{k})) + \delta(u_i^{j_{max}}(\tilde{k}) - \bar{u}^{j_{max}}(\tilde{k}))^2\Big) \\
&+\sum_{s=k_i}^{\tilde{k}}\Big(p^{j_{max}}(u_i^{j_{max}}(s;\tilde{k});\hat{U}_{-i}(\tilde{k}))D_{\tilde{k}}^{j_{max}} + \delta(u_i^{j_{max}}(\tilde{k}) - \bar{u}^{j_{max}}(\tilde{k}))^2
\end{aligned}$$

$$
\begin{aligned}
&+ f_i^{j_{max}}(u_i^{j_{max}}(s;\tilde{k}))\Big) + x_i^{j_{max}}(\tilde{k})^2 + Q_i^{j_{max}-1}(\hat{x}_i^{j_{max}-1}(\tilde{k}+K_i|\tilde{k}))\\
=&\sum_{s=k_i}^{\tilde{k}}\Big(p^{j_{max}}(u_i^{j_{max}}(s;\tilde{k});\hat{U}_{-i}(\tilde{k}))D_{\tilde{k}}^{j_{max}} + \delta(u_i^{j_{max}}(\tilde{k}) - \bar{u}^{j_{max}}(\tilde{k}))^2 + f_i^{j_{max}}(u_i^{j_{max}}(s;\tilde{k}))\Big)\\
&+ J_i^{j_{max}}(\hat{x}_i^{j_{max}-1}(\tilde{k}+1|\tilde{k}))\\
=&\sum_{s=\tilde{k}+1}^{k_i+K_i}\Big(p^{j_{max}}(u_i^{j_{max}}(s;\tilde{k});\hat{U}_{-i}(\tilde{k}))D_{\tilde{k}}^{j_{max}} + \delta(u_i^{j_{max}}(\tilde{k}) - \bar{u}^{j_{max}}(\tilde{k}))^2 + f_i^{j_{max}}(u_i^{j_{max}}(s;\tilde{k}))\Big)\\
&+\sum_{s=k_i}^{\tilde{k}}\Big(p^{j_{max}}(u_i^{j_{max}}(s;\tilde{k});\hat{U}_{-i}(\tilde{k}))D_{\tilde{k}}^{j_{max}} + \delta(u_i^{j_{max}}(\tilde{k}) - \bar{u}^{j_{max}}(\tilde{k}))^2 + f_i^{j_{max}}(u_i^{j_{max}}(s;\tilde{k}))\Big)\\
&+\sum_{s=k_i+K_i+1}^{k+K}\Big(p^{j_{max}}(u_i^{j_{max}}(s;\tilde{k});\hat{U}_{-i}(\tilde{k}))D_{\tilde{k}}^{j_{max}} + \delta(u_i^{j_{max}}(\tilde{k}) - \bar{u}^{j_{max}}(\tilde{k}))^2\\
&+ f_i^{j_{max}}(u_i^{j_{max}}(s;\tilde{k}))\Big)\\
\geq&\sum_{s=\tilde{k}+1}^{k_i+K_i}\Big(p^{j_{max}}(u_i^{j_{max}}(s;\tilde{k});\hat{U}_{-i}(\tilde{k}))D_{\tilde{k}}^{j_{max}} + \delta(u_i^{j_{max}}(\tilde{k}) - \bar{u}^{j_{max}}(\tilde{k}))^2 + f_i^{j_{max}}(u_i^{j_{max}}(s;\tilde{k}))\Big)\\
&+ p^{j_{max}}(u_i^{j_{max}}(\tilde{k};\tilde{k});\hat{U}_{-i}(\tilde{k}))D_{\tilde{k}}^{j_{max}} + \delta(u_i^{j_{max}}(\tilde{k}) - \bar{u}^{j_{max}}(\tilde{k}))^2 + f_i^{j_{max}}(u_i^{j_{max}}(s;\tilde{k}))\\
&+ (x_i^{j_{max}}(k_i+K_i) - x_i^{j_{max}}(\tilde{k}))^2 + p^{j_{max}}(u_i^{j_{max}}(k_i+K_i;\tilde{k});\hat{U}_{-i}(k_i+K_i))D_{k_i+K_i}^{j_{max}}\\
&+ \delta(u_i^{j_{max}}(k_i+K_i) - \bar{u}^{j_{max}}(k_i+K_i))^2 + f_i^{j_{max}}(u_i^{j_{max}}(s;\tilde{k}))\\
\geq& p^{j_{max}}(u_i^{j_{max}}(\tilde{k};\tilde{k});\hat{U}_{-i}(k_i))D_{\tilde{k}}^{j_{max}} + \delta(u_i^{j_{max}}(\tilde{k}) - \bar{u}^{j_{max}}(\tilde{k}))^2 + f_i^{j_{max}}(u_i^{j_{max}}(s;\tilde{k}))\\
&+ J_i^{j_{max}}(x_i^{j_{max}}(\tilde{k}+1))
\end{aligned}
\tag{5.29}
$$

Thus, we can conclude that the optimal cost is non-creasing.

$$
\begin{aligned}
J_i^{j_{max}}(x_i^{j_{max}}(\tilde{k}+1)) - J_i^{j_{max}}(x_i^{j_{max}}(\tilde{k})) \leq& - p^{j_{max}}(u_i^{j_{max}}(\tilde{k};\tilde{k});\hat{U}_{-i}(k_i))D_{\tilde{k}}^{j_{max}}\\
&+ \delta(u_i^{j_{max}}(\tilde{k}) - \bar{u}^{j_{max}}(\tilde{k}))^2 + f_i^{j_{max}}(u_i^{j_{max}}(s;\tilde{k}))\\
<&0
\end{aligned}
\tag{5.30}
$$

We have Eq. (5.30) holds and the cost function (5.2) is continue. The equilibrium is asymptotically stable. The proof is completed. □

5.4 Simulation Results

In this section, the performance of the proposed method is demonstrated by a numerical example. Our discourse begins with an explanation of the battery capacity degradation model involved in the optimal control problem and simulation setup. In addition, the simulation results are explained, and the effectiveness of the proposed method is verified.

The lithium-ion type of battery is widely used in the PEVs. In this section, we use the $LiFePO_4$ battery in Nissan Leaf as an example. Figure 5.1 shows the State of Health (SoH) degradation under different conditions and the approximate function of battery degradation using the data from [165]. The SoH is a kind of battery condition. The new batteries have 100% SoH which deteriorates over time. For instance, a 20 kWh battery under 90% SoH is performed like an 18 kWh battery. In this Chapter, we combine the SoH scenario with the SoC. If the battery under 90% SoH, it would stop charging at 90% SoC. In Figure 5.1, the dash lines illustrate different batteries capacity degradation in Nissan Leaf model 2019, 2015 and 2014, respectively. The lines with circles show the PEVs under varying use level, use over 20000 km per year (high use) and under 8000 km per year (low use). The two lines with stars illustrate charging with 120 Volt and 240 Volt, respectively. The lines with squares show if charging use Direct Current Fast Charger (DCFC), never use DCFC (Never Use) or frequently use DCFC (Use over three times per month). Finally, we calculate and approximate the average value of battery capacity degradation under different scenarios over time, shown as the blue line in Figure 5.1. Values of the parameters related to vehicle dynamics are illustrated in Table 5.1.

Table 5.1: The values of parameters related to vehicle dynamics and setup for charging control problems.

Parameters	**Values**
$SoC_{i,des}$	0.8 - 1.0
$SoC_{i,ini}$	0.1 - 0.3
ϵ_{stop}	0.06
the i-th PEV battery capacity $\overline{E}_i$	20 kWh
maximum charging power $\overline{P}_i$	6.6×2 kW
charging efficiency η	85%
generation cost function of electric energy price $c(D_{\tilde{k}})$	$c(D_{\tilde{k}}) = 5.8 \times 10^{-7} D_{\tilde{k}}^2 + 0.06 D_{\tilde{k}}$
price function $p(D_{\tilde{k}})$	$p(D_{\tilde{k}}) = 11.6 \times 10^{-7} D_{\tilde{k}} + 0.06$
begin time of charging k	k is the time instant at $15:00$
time period under the proposed control algorithm K	$K = 18$h

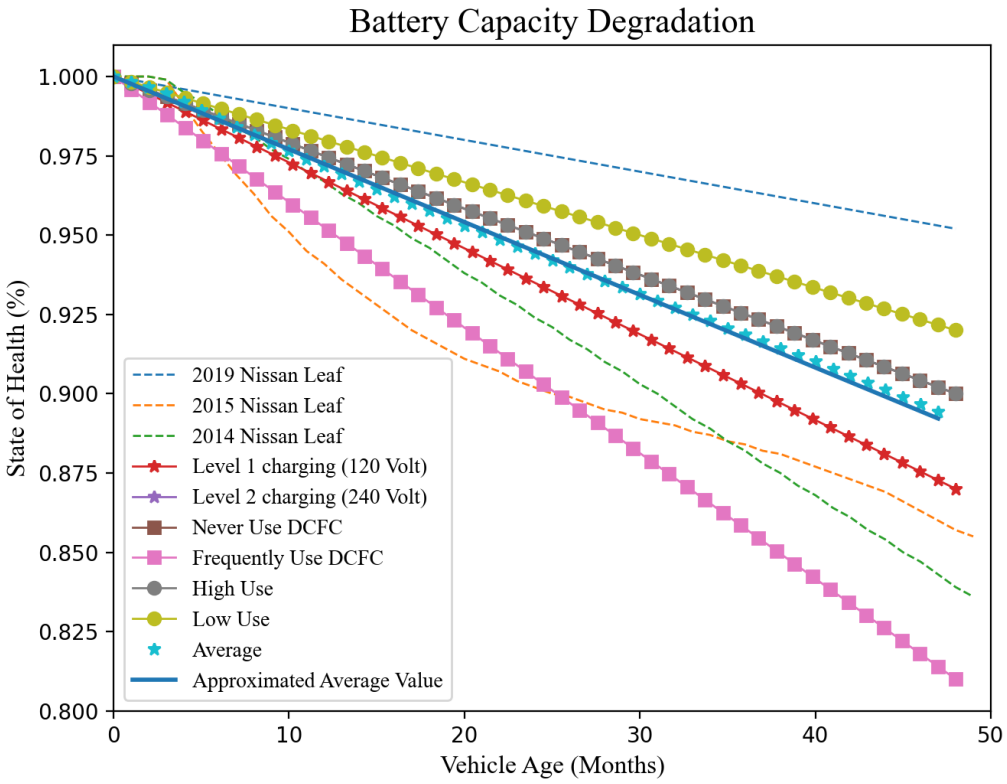
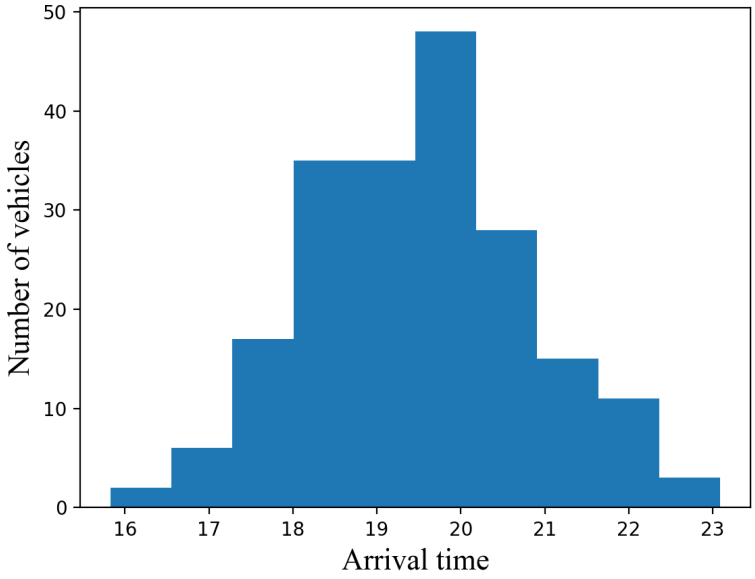


Figure 5.1: Battery capacity degradation under different conditions

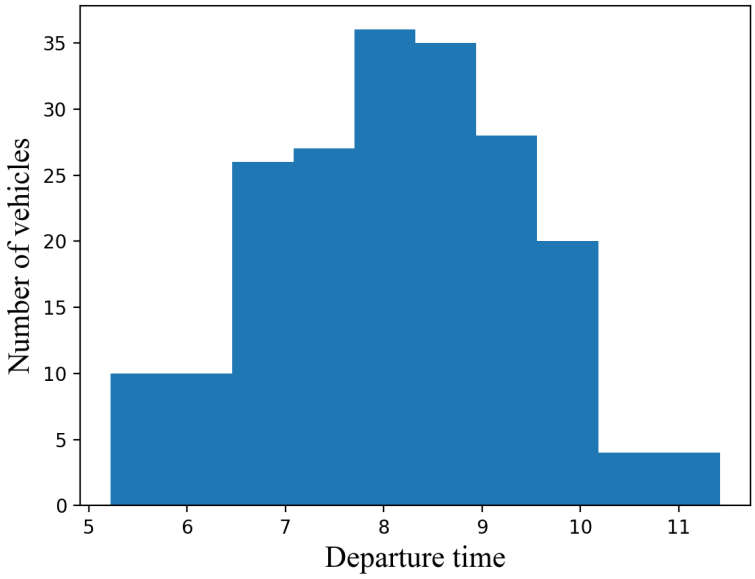
Figure 5.2(a) and Figure 5.2(b) show that the number of PEVs arrival and departure during the charging time, respectively. Figure 5.3(a) represents the total number of vehicles that are connected to the grid during the charging time. Figure 5.3(b) illustrates the desired SoC of each vehicle for charging at the initial time instant. Figure 5.4 represents the SoH of each vehicle for charging at the initial time instant. We assume that the SoH of each vehicle remains unchanged during the charging time period. Thus, the control inputs at the sampling time $\tilde{k}$ can be calculated.

We use the residential power load profiles [161] as the baseline for the charging problem. Since the charging controller of residential areas considered in this work is to fill the overnight valley, the simulation time (18 hours) is from 15:00 in the afternoon to the following day at 9:00.

Figure 5.5 and Figure 5.6 show the charge rate and SoC profiles during the simulation time, respectively. As shown, each PEV meets its own individual requirement. Figure 5.7 shows the power load profiles of the charging problem. The red dash line represents the



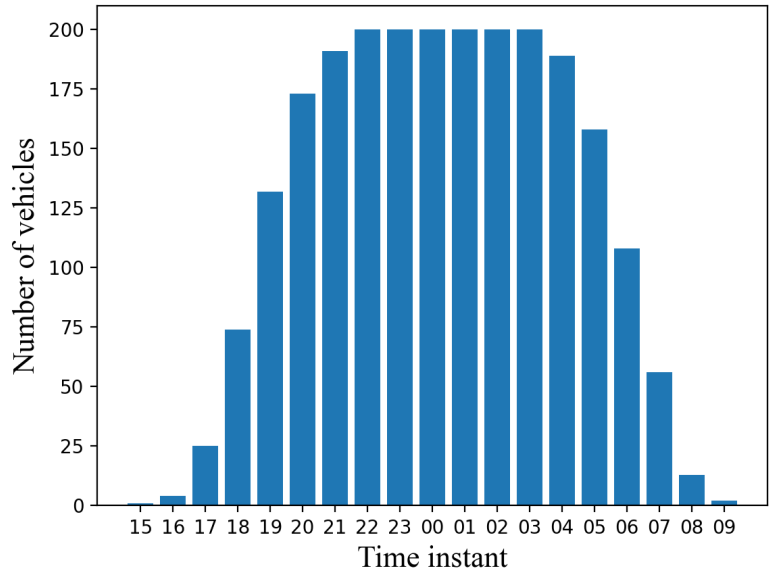
(a) The number of PEVs plug-in to the power grids. The count starts at 15:00 and lasts for 9 hours.



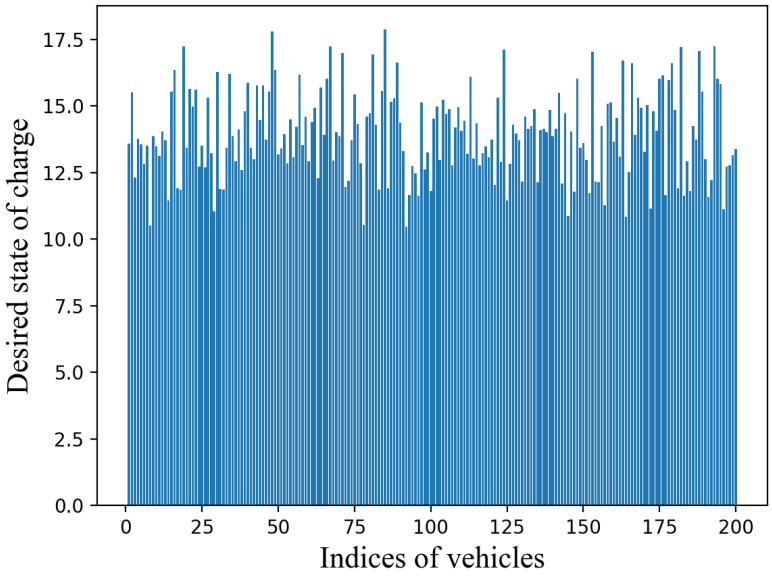
(b) The number of PEVs plug-out of the power grids. The count starts at 3:00 in the following morning and lasts 9 hours.

Figure 5.2: Arrival and departure time of PEVs for charging control

limitation of the power grids' loads. The green dash line illustrates the baseline of the power



(a) The number of PEVs can be used for charging control. The count for the simulation starts at 15:00 and lasts for 18 hours.



(b) The desired state of charge of each PEV for charging control.

Figure 5.3: Information of PEVs for charging control

load, which demonstrates the profile without PEV charging requirements. The blue dash line represents the power load profile without achieving the PEV charging objective. The

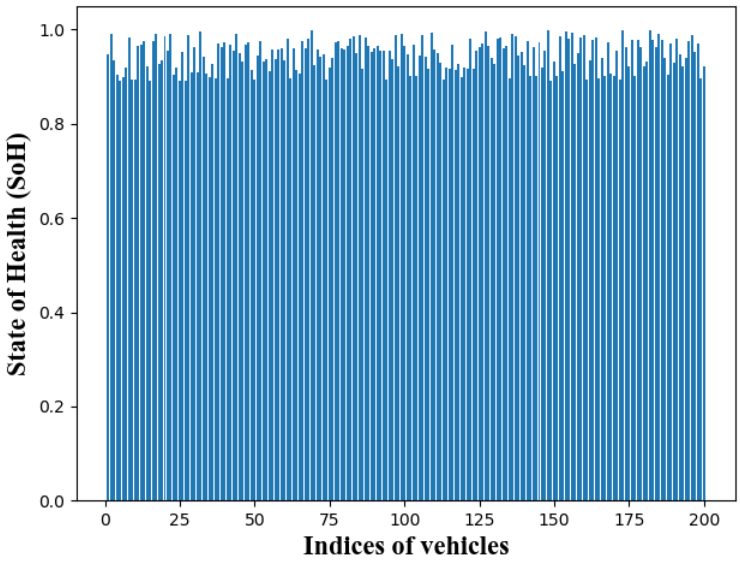


Figure 5.4: The State of Health (SoH) of each PEV for charging control.

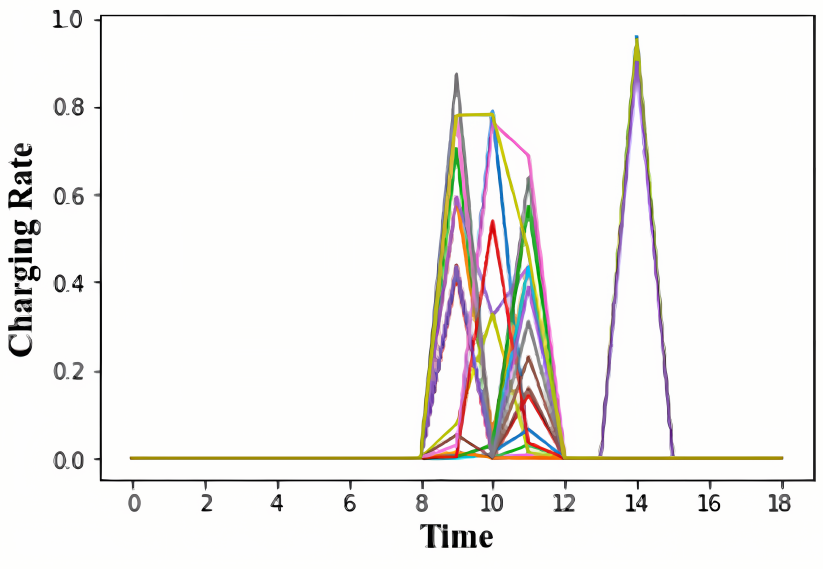


Figure 5.5: The charging rates calculate by solving the optimal control problem.

yellow lines denote the PEV charging power load profiles applying the charging control rates. One of the main objectives of the charging control problem is to avoid overload of the grid by managing the charging time of each PEV, which can be achieved by using the proposed method. The power load of PEV charging significantly impacts the grid, which causes safety problems for the grid, as shown using the orange line. As shown in Figure 5.7, the proposed method provides a possible solution to the valley-filling objective and can

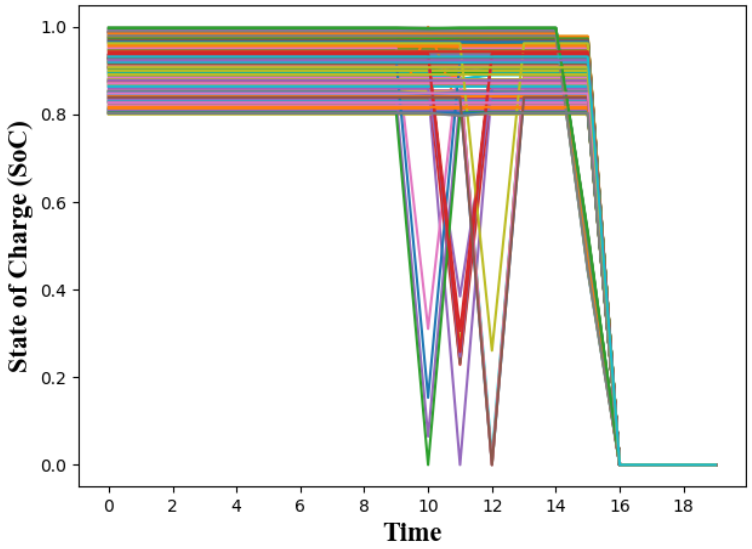


Figure 5.6: The State of Charge (SoC) by applying the charging control inputs.

meet the charging requirements of each PEV.

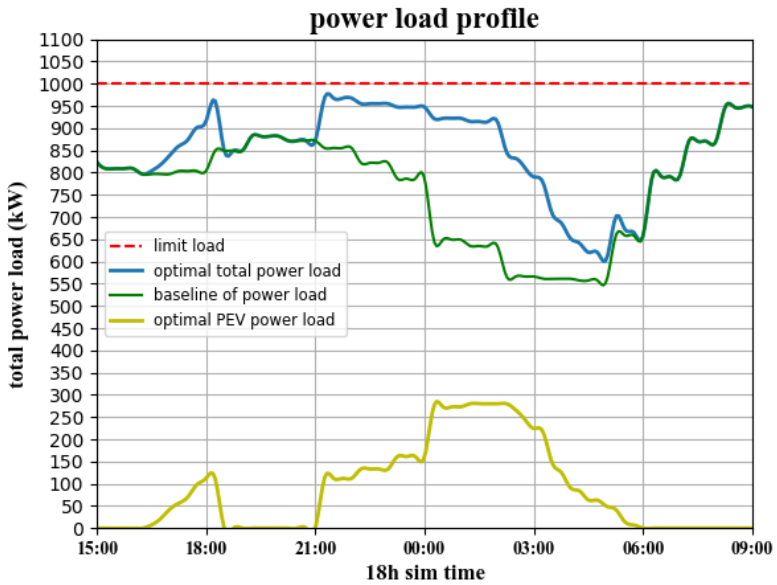


Figure 5.7: The power load profiles of the charging control

5.5 Conclusion

In this chapter, we studied a charging control algorithm using iterative learning-based decentralized MPC for PEVs with battery aging problems. Our goal was to find the optimal charging rate that minimizes the price cost of PEVs leading to valley-filling performance. The proposed scheme adopted the iterative learning-based method to improve the optimal control problem involving the battery capacity degradation model. Specifically, the sampled safe set and price function were updated accordingly as the iteration number increased. We also demonstrated the recursive feasibility of the proposed optimal control problem for PEV charging and asymptotical stability of the closed-loop system. Finally, simulation results showed that the proposed algorithm provided an effective solution to the problem of PEV charging with the battery aging model. Further investigations were needed in the future to reduce the computational overhead of the MPC algorithm.

Chapter 6

Conclusion and Future Work

This Ph.D. dissertation concerns two problems in the field of utilizing control strategies for EVs: Power management for PHEVs and charging control for PEVs plugged in the power grids. We have designed appropriate control methods for practical problems integrated with power management for PHEVs or PEVs charging, including route planning, communication loads reduction under centralized charging architecture, overnight valley-filling and battery aging under decentralized charging architecture. This chapter summarizes the main results of the dissertation, followed by a brief introduction of several potential topics for future research.

6.1 Conclusion

Chapter 2 studied a novel control scheme of route planning with power management for PHEVs. We found the route that leads to the minimum energy consumption, integrated the power management of PHEVs with route planning problems. The proposed scheme adopted a two-loop structure to achieve the control objective. In the inner loop, a nonlinear approximator structure was built to approximate control actions and energy consumption, which consisted of an actor approximator and a critic approximator under the RL framework. Finally, the optimized route was obtained by the minimum energy consumption converged in the outer loop.

Another limitation of impeding the PEVs' broad adoption was the long charging time and negative influence for grids caused by PEV charging. A large-scale amount of PEVs yielded high workloads of the electrical power grids, impacting efficiency and safety. The influence of charging PEVs on the electrical power grids required further research. **Chap-**

ter 3 investigated the self-triggered MPC with ISM to address the PEV charging problem subject to state and input constraints with additive disturbances and uncertainties. The motivation of using the ISM approach was to reject matched uncertainties. The constrained optimization problem has been solved aperiodically to generate control signals and the next execution time in the proposed scheme utilizing the self-triggered method, leading to reductions in both computation and communication. Consequently, the proposed method reduced the communication loads due to a large-scale amount of vehicles connected to the grids. **Chapter 4** presented a PEV charging control algorithm to address the valley-filling problem of power grids, which satisfied the individual charging requirements. An iterative learning-based model predictive charging control algorithm has been developed. The design of the decentralized MPC met the individual charging requirements. Compared with the uncontrolled power of grids, the proposed method avoid overload of the grids.

The drivers of PEVs had their own charging habits and driving behaviors, which affected the lifespan of the batteries of PEVs. For safety concerns, battery aging has been a problem that needs to be considered during driving or charging PEVs. **Chapter 5** studied a modified charging control algorithm based on the previous charging control algorithm studied in **Chapter 4**. Our goal was to find the optimal charging rate that minimized the price cost of PEVs leading to valley-filling performance. The cost function of the optimal control problem involved battery capacity degradation. The proposed scheme adopted the iterative learning-based method to improve the optimal control problem involving the battery capacity degradation. Specifically, the sampled safe set and price function were updated accordingly as the iteration number increased. As a result, the proposed scheme addressed the charging problem while the battery capacity degradation was considered.

6.2 Future Work

This dissertation focuses on the design of controllers to solve the power management for PHEVs and PEV charging problem. Furthermore, due to the environmental protection issues, there are a lot of interesting problems that deserve further research. We briefly provide some future study topics in the following.

- **RL Based Power Management for PHEVs with Auxiliary Systems**
  Since the power management problem of PHEVs is designed for the main power chain in the propulsion systems [164], some auxiliary systems require the amount of power. For instance, the air-conditioning systems may cost up to 25% or more of the

total fuel consumption. In the existing results, the auxiliary power consumption or loads are often ignored or treated as a constant since they are difficult to be measured. Due to the simplified power loads of auxiliary systems, the prediction of SoC and the control policy may be inaccurate. It is helpful for power management of PHEVs to involve the power loads of auxiliary systems in obtaining accurate prediction of states and designing the control policy. However, it is hard to build the accurate model for the power management of PHEVs with auxiliary systems. To tackle this challenge, RL shows its efficiency for dealing with uncertain systems, especially for the power management of PHEVs.

- **Multi-Mode Power Management for PHEVs Using An Iterative Learning-Based MPC**

  Generally speaking, a PHEV has five possible operating modes: Motor only, engine only, power-assist (motor and engine), recharging (engine charges the battery), and regenerative braking. In recent research results on PHEVs, these operating modes of PHEVs are independent for designing the optimal control policy. Also, the driving range is still an obstacle to the broad adoption of PHEVs. A multi-objective optimal control problem, including reducing energy consumption and negative impacts caused by battery capacity degradation, shows its efficiency for practice [185]. It is necessary to further reduce energy consumption by designing a multi-mode control method that solves multi-objective optimal control problems. For example, when the SoC of the battery is at an urgent low point, two operating modes (power-assist and recharging) can be used simultaneously to extend the driving range of the PHEV. The advantage of applying MPC for power management in PHEVs is allowing the multi-variable controller with the specified constraints on system states and inputs. However, MPC relies on the available accurate system model. Iterative learning-based MPC can devise the models by measured data when certain system dynamics models are not available. An MPC control algorithm for power management of PHEVs, coupled with an iterative learning method, can be utilized to solve the optimal control problem for multi objectives when the accurate system model is not available.

Appendix A

Publications

The following is a list of publications during Ph.D. studies.

- Journal Papers

  J1 **Q. Zhang**, K. Wu, and Y. Shi, "Route Planning and Power Management for PHEVs with Reinforcement Learning," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 4751-4762, 2020.

  J2 **Q. Zhang**, Y. Shi, and K. Wu, "Iterative Learning-based Decentralized Model Predictive Charging Control for PEVs with Battery Aging Model," under preparation.

  J3 **Q. Zhang**, Y. Shi, and K. Wu, "Model Predictive Control Application for Electric Vehicles: A Survey," under preparation.

- Conference Papers

  C1 **Q. Zhang**, Y. Shi, and K. Wu, "Self-Triggered Robust MPC with ISM for Constrained Nonlinear Input-Affine Systems," in *Proceedings of the 28th IEEE International Symposium on Industrial Electronics (ISIE)*. Vancouver, Canada: IEEE, June 2019.

  C2 **Q. Zhang**, Y. Shi, and K. Wu, "Iterative Learning-Based Decentralized Model Predictive Charging Control for Plugin Electric Vehicles," in *Proceedings of the 4th IEEE International Conference on Industrial Cyber-Physical Systems (ICPS)*. Victoria, Canada: IEEE, May 2021. (Winning the ICPS 2021 Best Student Paper Award)

Bibliography

[1] Chao Yang, Liang Li, Sixiong You, Bingjie Yan, and Xian Du. Cloud computing-based energy optimization control framework for plug-in hybrid electric bus. *Energy*, 125:11–26, 2017.

[2] John M Miller. Hybrid electric vehicle propulsion system architectures of the e-CVT type. *IEEE Transactions on Power Electronics*, 21(3):756–767, 2006.

[3] Global EV outlook 2021. https://www.iea.org/reports/global-ev-outlook-2021. Accessed September 1, 2021.

[4] Global energy review 2021. https://www.iea.org/reports/global-energy-review-2021. Accessed September 13, 2021.

[5] Global EV outlook 2020. https://www.iea.org/reports/global-ev-outlook-2020. Accessed September 1, 2021.

[6] Addressing the barriers to EV adoption. https://www.geotab.com/white-paper/barriers-to-ev-adoption/. Accessed September 26, 2021.

[7] Hassan Farhangi. The path of the smart grid. *IEEE Power and Energy Magazine*, 8(1):18–28, 2009.

[8] Ahmed Yousuf Saber and Ganesh Kumar Venayagamoorthy. Plug-in vehicles and renewable energy sources for cost and emission reductions. *IEEE Transactions on Industrial Electronics*, 58(4):1229–1238, 2010.

[9] Willett Kempton and Jasna Tomić. Vehicle-to-grid power fundamentals: Calculating capacity and net revenue. *Journal of Power Sources*, 144(1):268–279, 2005.

[10] Jasna Tomić and Willett Kempton. Using fleets of electric-drive vehicles for grid support. *Journal of Power Sources*, 168(2):459–468, 2007.

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

[12] Kamil Çağatay Bayindir, Mehmet Ali Gözüküçük, and Ahmet Teke. A comprehensive overview of hybrid electric vehicle: Powertrain configurations, powertrain control techniques and electronic control units. *Energy Conversion and Management*, 52(2):1305–1313, 2011.

[13] Ali Emadi. *Handbook of Automotive Power Electronics and Motor Drives*. CRC press, 2017.

[14] National Renewable Energy Lab. Advisor 2003 documentation. http://bigladdersoftware.com/advisor/docs/advisor_doc.html. Accessed February 4, 2018.

[15] Liang Li, Chao Yang, Yahui Zhang, Lipeng Zhang, and Jian Song. Correctional DP-based energy management strategy of plug-in hybrid electric bus for city-bus route. *IEEE Transactions on Vehicular Technology*, 64(7):2792–2803, 2014.

[16] Hyeoun-Dong Lee, Seung-Ki Sul, Han-Sang Cho, and Jang-Moo Lee. Advanced gear-shifting and clutching strategy for a parallel-hybrid vehicle. *IEEE Industry Applications Magazine*, 6(6):26–32, 2000.

[17] Chan-Chiao Lin, Huei Peng, Jessy W Grizzle, and Jun-Mo Kang. Power management strategy for a parallel hybrid electric truck. *IEEE Transactions on Control Systems Technology*, 11(6):839–849, 2003.

[18] J García-Villalobos, I Zamora, JI San Martín, FJ Asensio, and V Aperribay. Plug-in electric vehicles in electric distribution networks: A review of smart charging approaches. *Renewable and Sustainable Energy Reviews*, 38:717–731, 2014.

[19] Evangelos L Karfopoulos and Nikos D Hatziargyriou. Distributed coordination of electric vehicles providing V2G services. *IEEE Transactions on Power Systems*, 31(1):329–338, 2015.

[20] Abouzar Ghavami, Koushik Kar, and Aparna Gupta. Decentralized charging of plug-in electric vehicles with distribution feeder overload control. *IEEE Transactions on Automatic Control*, 61(11):3527–3532, 2016.

[21] Nanduni I Nimalsiri, Chathurika P Mediwaththe, Elizabeth L Ratnam, Marnie Shaw, David B Smith, and Saman K Halgamuge. A survey of algorithms for distributed charging control of electric vehicles in smart grid. *IEEE Transactions on Intelligent Transportation Systems*, 21(11):4497–4515, 2019.

[22] Amir-Hamed Mohsenian-Rad, Vincent WS Wong, Juri Jatskevich, Robert Schober, and Alberto Leon-Garcia. Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid. *IEEE Transactions on Smart Grid*, 1(3):320–331, 2010.

[23] Mosaddek Hossain Kamal Tushar, Adel W Zeineddine, and Chadi Assi. Demand-side management by regulating charging and discharging of the EV, ESS, and utilizing renewable energy. *IEEE Transactions on Industrial Informatics*, 14(1):117–126, 2017.

[24] Hung Khanh Nguyen and Ju Bin Song. Optimal charging and discharging for multiple PHEVs with demand side management in vehicle-to-building. *Journal of Communications and Networks*, 14(6):662–671, 2012.

[25] Zhaoxi Liu, Qiuwei Wu, Shaojun Huang, Lingfeng Wang, Mohammad Shahidehpour, and Yusheng Xue. Optimal day-ahead charging scheduling of electric vehicles through an aggregative game model. *IEEE Transactions on Smart Grid*, 9(5):5173–5184, 2017.

[26] Chathurika P Mediwaththe and David B Smith. Game-theoretic electric vehicle charging management resilient to non-ideal user behavior. *IEEE Transactions on Intelligent Transportation Systems*, 19(11):3486–3495, 2018.

[27] Wayes Tushar, Walid Saad, H Vincent Poor, and David B Smith. Economics of electric vehicle charging: A game theoretic approach. *IEEE Transactions on Smart Grid*, 3(4):1767–1778, 2012.

[28] Jose Rivera, Christoph Goebel, and Hans-Arno Jacobsen. Distributed convex optimization for electric vehicle aggregators. *IEEE Transactions on Smart Grid*, 8(4):1852–1863, 2016.

[29] Chao-Kai Wen, Jung-Chieh Chen, Jen-Hao Teng, and Pangan Ting. Decentralized plug-in electric vehicle charging selection algorithm in power systems. *IEEE Transactions on Smart Grid*, 3(4):1779–1789, 2012.

[30] Milad Latifi, Amir Rastegarnia, Azam Khalili, and Saeid Sanei. Agent-based decentralized optimal charging strategy for plug-in electric vehicles. *IEEE Transactions on Industrial Electronics*, 66(5):3668–3680, 2018.

[31] Akshay Malhotra, Giulio Binetti, Ali Davoudi, and Ioannis D Schizas. Distributed power profile tracking for heterogeneous charging of electric vehicles. *IEEE Transactions on Smart Grid*, 8(5):2090–2099, 2016.

[32] Dhananjay M Anand, Rupert Tull de Salis, Yijie Cheng, James Moyne, and Dawn M Tilbury. A hierarchical incentive arbitration scheme for coordinated PEV charging stations. *IEEE Transactions on Smart Grid*, 6(4):1775–1784, 2015.

[33] Caroline Le Floch, Somil Bansal, Claire J Tomlin, Scott J Moura, and Melanie N Zeilinger. Plug-and-play model predictive control for load shaping and voltage control in smart grids. *IEEE Transactions on Smart Grid*, 10(3):2334–2344, 2017.

[34] Alessandro Di Giorgio, Francesco Liberati, and Silvia Canale. Electric vehicles charging control in a smart grid: A model predictive control approach. *Control Engineering Practice*, 22:147–162, 2014.

[35] Wanrong Tang and Ying Jun Angela Zhang. A model predictive control approach for low-complexity electric vehicle charging scheduling: Optimality and scalability. *IEEE Transactions on Power Systems*, 32(2):1050–1063, 2016.

[36] Changsun Ahn, Chiao-Ting Li, and Huei Peng. Optimal decentralized charging control algorithm for electrified vehicles connected to smart grid. *Journal of Power Sources*, 196(23):10369–10379, 2011.

[37] Tan Ma and Osama A Mohammed. Optimal charging of plug-in electric vehicles for a car-park infrastructure. *IEEE Transactions on Industry Applications*, 50(4):2323–2330, 2014.

[38] Omid Ardakanian, Catherine Rosenberg, and Srinivasan Keshav. Distributed control of electric vehicle charging. In *Proceedings of the Fourth International Conference on Future Energy Systems*, Berkeley, California, May 2013.

[39] Junjie Hu, Shi You, Morten Lind, and Jacob Østergaard. Coordinated charging of electric vehicles for congestion prevention in the distribution grid. *IEEE Transactions on Smart Grid*, 5(2):703–711, 2013.

[40] Zhongjing Ma, Duncan S Callaway, and Ian A Hiskens. Decentralized charging control of large populations of plug-in electric vehicles. *IEEE Transactions on Control Systems Technology*, 21(1):67–78, 2011.

[41] Zhongjing Ma, Duncan Callaway, and Ian Hiskens. Decentralized charging control for large populations of plug-in electric vehicles: Application of the nash certainty equivalence principle. In *Proceedings of IEEE International Conference on Control Applications*, Yokohama, Japan, Sept. 2010. IEEE.

[42] Chenye Wu, Hamed Mohsenian-Rad, and Jianwei Huang. Vehicle-to-aggregator interaction game. *IEEE Transactions on Smart Grid*, 3(1):434–442, 2011.

[43] Jun Tan and Lingfeng Wang. Enabling reliability-differentiated service in residential distribution networks with PHEVs: A hierarchical game approach. *IEEE Transactions on Smart Grid*, 7(2):684–694, 2015.

[44] Woongsup Lee, Lin Xiang, Robert Schober, and Vincent WS Wong. Electric vehicle charging stations with renewable power generators: A game theoretical analysis. *IEEE Transactions on Smart Grid*, 6(2):608–617, 2014.

[45] Liang Zhang, Vassilis Kekatos, and Georgios B Giannakis. Scalable electric vehicle charging protocols. *IEEE Transactions on Power Systems*, 32(2):1451–1462, 2016.

[46] Mingxi Liu, Phillippe K Phanivong, Yang Shi, and Duncan S Callaway. Decentralized charging control of electric vehicles in residential distribution networks. *IEEE Transactions on Control Systems Technology*, 27(1):266–281, 2017.

[47] Mingxi Liu, Yang Shi, and Huijun Gao. Aggregation and charging control of PHEVs in smart grid: A cyber–physical perspective. *Proceedings of the IEEE*, 104(5):1071–1085, 2016.

[48] Weifeng Yao, Junhua Zhao, Fushuan Wen, Yusheng Xue, and Gerard Ledwich. A hierarchical decomposition approach for coordinated dispatch of plug-in electric vehicles. *IEEE Transactions on Power Systems*, 28(3):2768–2778, 2013.

[49] Juan P Torreglosa, Pablo García-Triviño, Luis M Fernández-Ramirez, and Francisco Jurado. Decentralized energy management strategy based on predictive controllers for a medium voltage direct current photovoltaic electric vehicle charging station. *Energy Conversion and Management*, 108:1–13, 2016.

[50] Andres Ovalle, Julian Fernandez, Ahmad Hably, and Seddik Bacha. An electric vehicle load management application of the mixed strategist dynamics and the maximum entropy principle. *IEEE Transactions on Industrial Electronics*, 63(5):3060–3071, 2016.

[51] Zhiwei Xu, Zechun Hu, Yonghua Song, Wei Zhao, and Yongwang Zhang. Coordination of PEVs charging across multiple aggregators. *Applied Energy*, 136:582–589, 2014.

[52] Erotokritos Xydas, Charalampos Marmaras, Liana M Cipcigan, Nick Jenkins, Steve Carroll, and Myles Barker. A data-driven approach for characterising the charging demand of electric vehicles: A UK case study. *Applied Energy*, 162:763–771, 2016.

[53] Emre C Kara, Jason S Macdonald, Douglas Black, Mario Bérges, Gabriela Hug, and Sila Kiliccote. Estimating the benefits of electric vehicle smart charging at non-residential locations: A data-driven approach. *Applied Energy*, 155:515–525, 2015.

[54] Xiaochen Zhang and Santiago Grijalva. An advanced data driven model for residential electric vehicle charging demand. In *Proceedings of IEEE Power & Energy Society General Meeting*, Denver, CO, USA, Jul. 2015. IEEE.

[55] Wei Qi, Zhiwei Xu, Zuo-Jun Max Shen, Zechun Hu, and Yonghua Song. Hierarchical coordinated control of plug-in electric vehicles charging in multifamily dwellings. *IEEE Transactions on Smart Grid*, 5(3):1465–1474, 2014.

[56] João A Peças Lopes, Filipe Joel Soares, and Pedro M Rocha Almeida. Integration of electric vehicles in the electric power system. *Proceedings of the IEEE*, 99(1):168–183, 2010.

[57] Rasmus Halvgaard, Niels K Poulsen, Henrik Madsen, John B Jørgensen, Francesco Marra, and Daniel Esteban Morales Bondy. Electric vehicle charge planning using economic model predictive control. In *Proceedings of 2012 IEEE International Electric Vehicle Conference*, Greenville, SC, USA, Mar. 2012. IEEE.

[58] Lingwen Gan, Ufuk Topcu, and Steven H Low. Optimal decentralized protocol for electric vehicle charging. *IEEE Transactions on Power Systems*, 28(2):940–951, 2012.

[59] Matthias D Galus, Marina González Vayá, Thilo Krause, and Göran Andersson. The role of electric vehicles in smart grids. *Wiley Interdisciplinary Reviews: Energy and Environment*, 2(4):384–400, 2013.

[60] Hamed Mohsenian-Rad et al. Optimal charging of electric vehicles with uncertain departure times: A closed-form solution. *IEEE Transactions on Smart Grid*, 6(2):940–942, 2014.

[61] Mingxi Liu, Phillippe K Phanivong, and Duncan S Callaway. Electric vehicle charging control in residential distribution network: A decentralized event-driven realization. In *Proceedings of the 56th IEEE Conference on Decision and Control (CDC)*, Melbourne, VIC, Australia, Dec. 2017. IEEE.

[62] Mingxi Liu, Phillippe K Phanivong, and Duncan S Callaway. Customer-and network-aware decentralized EV charging control. In *Proceedings of Power Systems Computation Conference (PSCC)*, Dublin, Ireland, Jun. 2018. IEEE.

[63] Chih-Cheng Huang, Chun-Liang Lin, and Yuan-Kang Wu. Simultaneous wireless power/data transfer for electric vehicle charging. *IEEE Transactions on Industrial Electronics*, 64(1):682–690, 2016.

[64] Joy Chandra Mukherjee and Arobinda Gupta. Distributed charge scheduling of plug-in electric vehicles using inter-aggregator collaboration. *IEEE Transactions on Smart Grid*, 8(1):331–341, 2016.

[65] Stijn Vandael, Bert Claessens, Maarten Hommelberg, Tom Holvoet, and Geert Deconinck. A scalable three-step approach for demand side management of plug-in hybrid vehicles. *IEEE Transactions on Smart Grid*, 4(2):720–728, 2012.

[66] Xiaomin Xi and Ramteen Sioshansi. Using price-based signals to control plug-in electric vehicle fleet charging. *IEEE Transactions on Smart Grid*, 5(3):1451–1464, 2014.

[67] Chengcheng Shao, Xifan Wang, Xiuli Wang, Chao Du, and Biyang Wang. Hierarchical charge control of large populations of EVs. *IEEE Transactions on Smart Grid*, 7(2):1147–1155, 2015.

[68] Wenjing Shuai, Patrick Maillé, and Alexander Pelov. Charging electric vehicles in the smart city: A survey of economy-driven approaches. *IEEE Transactions on Intelligent Transportation Systems*, 17(8):2089–2106, 2016.

[69] Richard S Sutton and Andrew G Barto. *Reinforcement Learning: An Introduction*. MIT press, 2018.

[70] Qiuming Gong, Yaoyu Li, and Zhongren Peng. Power management of plug-in hybrid electric vehicles using neural network based trip modeling. In *Proceedings of American Control Conference (ACC)*, St. Louis, MO, USA, Jun. 2009. IEEE.

[71] Rui Xiong, Jiayi Cao, and Quanqing Yu. Reinforcement learning-based real-time power management for hybrid energy storage system in the plug-in hybrid electric vehicle. *Applied Energy*, 211:538–548, 2018.

[72] Jingda Wu, Hongwen He, Jiankun Peng, Yuecheng Li, and Zhanjiang Li. Continuous reinforcement learning of energy management with deep Q network for a power split hybrid electric bus. *Applied Energy*, 222:799–811, 2018.

[73] Yuan Zou, Teng Liu, Dexing Liu, and Fengchun Sun. Reinforcement learning-based real-time energy management for a hybrid tracked vehicle. *Applied Energy*, 171:372–382, 2016.

[74] Chan-Chiao Lin, Huei Peng, and JW Grizzle. A stochastic control strategy for hybrid electric vehicles. In *Proceedings of the American Control Conference (ACC)*, Boston, MA, USA, June. 2004. IEEE.

[75] Jungme Park, Zhihang Chen, Leonidas Kiliaris, Ming L Kuang, M Abul Masrur, Anthony M Phillips, and Yi Lu Murphey. Intelligent vehicle power control based on machine learning of optimal control parameters and prediction of road type and traffic congestion. *IEEE Transactions on Vehicular Technology*, 58(9):4741–4756, 2009.

[76] Craig KD Harold, Suraj Prakash, and Theo Hofman. Powertrain control for hybrid-electric vehicles using supervised machine learning. *Vehicles*, 2(2):267–286, 2020.

[77] Zheng Chen, Liang Li, Xiaosong Hu, Bingjie Yan, and Chao Yang. Temporal-difference learning-based stochastic energy management for plug-in hybrid electric buses. *IEEE Transactions on Intelligent Transportation Systems*, 20(6):2378–2388, 2018.

[78] Huachun Tan, Hailong Zhang, Jiankun Peng, Zhuxi Jiang, and Yuankai Wu. Energy management of hybrid electric bus based on deep reinforcement learning in continuous state and action space. *Energy Conversion and Management*, 195:548–560, 2019.

[79] Yi Lu Murphey, Jungme Park, Zhihang Chen, Ming L Kuang, M Abul Masrur, and Anthony M Phillips. Intelligent hybrid vehicle power control—part I: Machine learning of optimal vehicle power. *IEEE Transactions on Vehicular Technology*, 61(8):3519–3530, 2012.

[80] Yi Lu Murphey, Jungme Park, Leonidas Kiliaris, Ming L Kuang, M Abul Masrur, Anthony M Phillips, and Qing Wang. Intelligent hybrid vehicle power control—part II: Online intelligent energy management. *IEEE Transactions on Vehicular Technology*, 62(1):69–79, 2013.

[81] Zheng Chen, Chunting Chris Mi, Jun Xu, Xianzhi Gong, and Chenwen You. Energy management for a power-split plug-in hybrid electric vehicle based on dynamic programming and neural networks. *IEEE Transactions on Vehicular Technology*, 63(4):1567–1580, 2014.

[82] Joo P Hespanha, Payam Naghshtabrizi, and Yonggang Xu. A survey of recent results in networked control systems. *Proceedings of the IEEE*, 95(1):138–162, 2007.

[83] Hanna Michalska and David Q Mayne. Robust receding horizon control of constrained nonlinear systems. *IEEE Transactions on Automatic Control*, 38(11):1623–1633, 1993.

[84] David W Clarke, Coorous Mohtadi, and PS Tuffs. Generalized predictive control—part I. the basic algorithm. *Automatica*, 23(2):137–148, 1987.

[85] David W Clarke, Coorous Mohtadi, and PS Tuffs. Generalized predictive control—part II extensions and interpretations. *Automatica*, 23(2):149–160, 1987.

[86] David Q Mayne, James B Rawlings, Christopher V Rao, and Pierre OM Scokaert. Constrained model predictive control: Stability and optimality. *Automatica*, 36(6):789–814, 2000.

[87] Huiping Li and Yang Shi. *Robust Receding Horizon Control for Networked and Distributed Nonlinear Systems*. Springer, 2017.

[88] Robert R Bitmead, Michel Gevers, and Vincent Wertz. *Adaptive Optimal Control: The Thinking Man's GPC*. Prentice Hall of Australia, 1990.

[89] Eduardo Fernandez-Camacho and Carlos Bordons-Alba. *Model Predictive Control in The Process Industry*. Springer, 1995.

[90] Juncheng Zhu, Zhile Yang, Monjur Mourshed, Yuanjun Guo, Yimin Zhou, Yan Chang, Yanjie Wei, and Shengzhong Feng. Electric vehicle charging load forecasting: A comparative study of deep learning approaches. *Energies*, 12(14):2692, 2019.

[91] Adriana Chiş, Jarmo Lundén, and Visa Koivunen. Reinforcement learning-based plug-in electric vehicle charging with forecasted price. *IEEE Transactions on Vehicular Technology*, 66(5):3674–3684, 2016.

[92] Morteza Dabbaghjamanesh, Amirhossein Moeini, and Abdollah Kavousi-Fard. Reinforcement learning-based load forecasting of electric vehicle charging station using Q-learning technique. *IEEE Transactions on Industrial Informatics*, 17(6):4229–4237, 2020.

[93] Douglas A Bristow, Marina Tharayil, and Andrew G Alleyne. A survey of iterative learning control. *IEEE Control Systems Magazine*, 26(3):96–114, 2006.

[94] YangQuan Chen, Kevin L. Moore, and Hyo-Sung Ahn. *Iterative Learning Control*. Springer US, Boston, MA, 2012.

[95] Kwang S Lee, In-Shik Chin, Hyuk J Lee, and Jay H Lee. Model predictive control technique combined with iterative learning for batch processes. *AIChE Journal*, 45(10):2175–2187, 1999.

[96] Jingyi Lu, Zhixing Cao, and Furong Gao. Multipoint iterative learning model predictive control. *IEEE Transactions on Industrial Electronics*, 66(8):6230–6240, 2018.

[97] Youqing Wang, Donghua Zhou, and Furong Gao. Iterative learning model predictive control for multi-phase batch processes. *Journal of Process Control*, 18(6):543–557, 2008.

[98] Hong-Qiang Guo, Cong-Zhi Liu, Jia-Wang Yong, Xing-Qun Cheng, and Fahad Muhammad. Model predictive iterative learning control for energy management of plug-in hybrid electric vehicle. *IEEE Access*, 7:71323–71334, 2019.

[99] U.S. Energy Information Administration. How much carbon dioxide is produced from burning gasoline and diesel fuel? https://www.eia.gov/tools/faqs/faq.php?id=307&t=11. Accessed September 13, 2018.

[100] Xiangyu Wang, Liang Li, Kai He, and Congzhi Liu. Dual-loop self-learning fuzzy control for AMT gear engagement: Design and experiment. *IEEE Transactions on Fuzzy Systems*, 26(4):1813–1822, 2017.

[101] Xiangyu Wang, Liang Li, and Chao Yang. Hierarchical control of dry clutch for engine-start process in a parallel hybrid electric vehicle. *IEEE Transactions on Transportation Electrification*, 2(2):231–243, 2016.

[102] Jiankun Peng, Hongwen He, and Rui Xiong. Rule based energy management strategy for a series–parallel plug-in hybrid electric bus optimized by dynamic programming. *Applied Energy*, 185:1633–1643, 2017.

[103] Marco Sorrentino, Gianfranco Rizzo, and Ivan Arsie. Analysis of a rule-based control strategy for on-board energy management of series hybrid vehicles. *Control Engineering Practice*, 19(12):1433–1441, 2011.

[104] Jong-Seob Won and Reza Langari. Intelligent energy management agent for a parallel hybrid vehicle-part II: Torque distribution, charge sustenance strategies, and performance results. *IEEE Transactions on Vehicular Technology*, 54(3):935–953, 2005.

[105] Maciej Wieczorek and Mirosław Lewandowski. A mathematical representation of an energy management strategy for hybrid energy storage system in electric vehicle and real time optimization using a genetic algorithm. *Applied Energy*, 192:222–233, 2017.

[106] Qi Jiang, Florence Ossart, and Claude Marchand. Comparative study of real-time HEV energy management strategies. *IEEE Transactions on Vehicular Technology*, 66(12):10875–10888, 2017.

[107] Liangfei Xu, Fuyuan Yang, Jianqiu Li, Minggao Ouyang, and Jianfeng Hua. Real time optimal energy management strategy targeting at minimizing daily operation cost for a plug-in fuel cell city bus. *International Journal of Hydrogen Energy*, 37(20):15380–15392, 2012.

[108] Antonio Sciarretta, Michael Back, and Lino Guzzella. Optimal control of parallel hybrid electric vehicles. *IEEE Transactions on Control Systems Technology*, 12(3):352–363, 2004.

[109] Namwook Kim, Sukwon Cha, and Huei Peng. Optimal control of hybrid electric vehicles based on pontryagin's minimum principle. *IEEE Transactions on Control Systems Technology*, 19(5):1279–1287, 2010.

[110] Lorenzo Serrao, Simona Onori, Antonio Sciarretta, Yann Guezennec, and Giorgio Rizzoni. Optimal energy management of hybrid electric vehicles including battery aging. In *Proceedings of the 30th American Control Conference (ACC 2011)*, San Francisco, California, USA, Jun 2011. IEEE.

[111] Stefano Di Cairano, Daniele Bernardini, Alberto Bemporad, and Ilya V Kolmanovsky. Stochastic MPC with learning for driver-predictive vehicle control and its application to HEV energy management. *IEEE Transactions on Control Systems Technology*, 22(3):1018–1031, 2014.

[112] Fengjun Yan, Junmin Wang, and Kaisheng Huang. Hybrid electric vehicle model predictive control torque-split strategy incorporating engine transient characteristics. *IEEE Transactions on Vehicular Technology*, 61(6):2458–2467, 2012.

[113] Shuo Zhang, Rui Xiong, and Fengchun Sun. Model predictive control for power management in a plug-in hybrid electric vehicle with a hybrid energy storage system. *Applied Energy*, 185:1654–1662, 2017.

[114] Cong Hou, Minggao Ouyang, Liangfei Xu, and Hewu Wang. Approximate pontryagin's minimum principle applied to the energy management of plug-in hybrid electric vehicles. *Applied Energy*, 115:174–189, 2014.

[115] Qiuming Gong, Yaoyu Li, and Zhong-Ren Peng. Trip-based optimal power management of plug-in hybrid electric vehicles. *IEEE Transactions on Vehicular Technology*, 57(6):3393–3401, 2008.

[116] Yong Sun, Xiangyu Wang, Liang Li, Jialei Shi, and Quan An. Modelling and control for economy-oriented car-following problem of hybrid electric vehicle. *IET Intelligent Transport Systems*, 13(5):825–833, 2019.

[117] Xue Lin, Yanzhi Wang, Paul Bogdan, Naehyuck Chang, and Massoud Pedram. Reinforcement learning based power management for hybrid electric vehicles. In *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design (ICCAD 2014)*, San Jose, California, USA, Nov. 2014. IEEE/ACM.

[118] Chang Liu and Yi Lu Murphey. Optimal power management based on Q-learning and neuro-dynamic programming for plug-in hybrid electric vehicles. *IEEE Transactions on Neural Networks and Learning Systems*, 31(6):1942–1954, 2019.

[119] Teng Liu, Xiaosong Hu, Shengbo Eben Li, and Dongpu Cao. Reinforcement learning optimized look-ahead energy management of a parallel hybrid electric vehicle. *IEEE/ASME Transactions on Mechatronics*, 22(4):1497–1507, 2017.

[120] Jie Lin, Wei Yu, Xinyu Yang, Qingyu Yang, Xinwen Fu, and Wei Zhao. A real-time en-route route guidance decision scheme for transportation-based cyberphysical systems. *IEEE Transactions on Vehicular Technology*, 66(3):2551–2566, 2016.

[121] Md Abdus Samad Kamal, Tomohisa Hayakawa, and Jun-ichi Imura. Road-speed profile for enhanced perception of traffic conditions in a partially connected vehicle environment. *IEEE Transactions on Vehicular Technology*, 67(8):6824–6837, 2018.

[122] Hai Yu, Ming Kuang, and Ryan McGee. Trip-oriented energy management control strategy for plug-in hybrid electric vehicles. *IEEE Transactions on Control Systems Technology*, 22(4):1323–1336, 2014.

[123] Frank L Lewis, Draguna Vrabie, and Vassilis L Syrmos. *Optimal Control*. John Wiley & Sons, 2012.

[124] Zheng Chen and Sarangapani Jagannathan. Generalized Hamilton–Jacobi–Bellman formulation-based neural network control of affine nonlinear discrete-time systems. *IEEE Transactions on Neural Networks*, 19(1):90–106, 2008.

[125] Qiming Zhao, Hao Xu, and Sarangapani Jagannathan. Neural network-based finite-horizon optimal control of uncertain affine nonlinear discrete-time systems. *IEEE Transactions on Neural Networks and Learning Systems*, 26(3):486–499, 2014.

[126] Frauke Günther and Stefan Fritsch. neuralnet: Training of neural networks. *The R Journal*, 2(1):30–38, 2010.

[127] WPMH Heemels, Karl Henrik Johansson, and Paulo Tabuada. An introduction to event-triggered and self-triggered control. In *Proceedings of the 51st IEEE Conference on Decision and Control (CDC 2012)*, Maui, Hawaii, USA, Dec. 10-13, 2012.

[128] Joris Sijs, Mircea Lazar, and WPMH Heemels. On integration of event-based estimation and robust MPC in a feedback loop. In *Proceedings of the 13th ACM International Conference on Hybrid Systems: Computation and Control*, Stockholm Sweden, Apr. 2010.

[129] Florian D Brunner, WPMH Heemels, and Frank Allgöwer. Robust event-triggered MPC for constrained linear discrete-time systems with guaranteed average sampling rate. *IFAC-PapersOnLine*, 48(23):117–122, 2015.

[130] Shahab Heshmati-alamdari, Alina Eqtami, George C Karras, Dimos V Dimarogonas, and Kostas J Kyriakopoulos. A self-triggered position based visual servoing model predictive control scheme for underwater robotic vehicles. *Machines*, 8(2):33, 2020.

[131] Linni Jian, Yanchong Zheng, Xinping Xiao, and CC Chan. Optimal scheduling for vehicle-to-grid operation with stochastic connection of plug-in electric vehicles to smart grid. *Applied Energy*, 146:150–161, 2015.

[132] TMP Gommans and WPMH Heemels. Resource-aware MPC for constrained nonlinear systems: A self-triggered control approach. *Systems & Control Letters*, 79:59–67, 2015.

[133] Adolfo Anta and Paulo Tabuada. To sample or not to sample: Self-triggered control for nonlinear systems. *IEEE Transactions on Automatic Control*, 55(9):2030–2042, 2010.

[134] David Q Mayne, María M Seron, and SV Raković. Robust model predictive control of constrained linear systems with bounded disturbances. *Automatica*, 41(2):219–224, 2005.

[135] Jingyuan Zhan and Xiang Li. Self-triggered consensus of multi-agent systems via model predictive control. In *Proceedings of the 14th IFAC Symposium on Large Scale Complex Systems Theory and Applications (LSS 2016)*, Riverside, California, USA, May 28, 2016.

[136] Florian D Brunner, WPMH Heemels, and Frank Allgöwer. Numerical evaluation of a robust self-triggered MPC algorithm. *IFAC-PapersOnLine*, 49(22):151–156, 2016.

[137] Wilbur Langson, Ioannis Chryssochoos, SV Raković, and David Q Mayne. Robust model predictive control using tubes. *Automatica*, 40(1):125–133, 2004.

[138] Florian David Brunner, Maurice Heemels, and Frank Allgöwer. Robust self-triggered MPC for constrained linear systems: A tube-based approach. *Automatica*, 72:73–83, 2016.

[139] Daniel Limón, T Alamo, Francisco Salas, and Eduardo F Camacho. Input to state stability of min–max MPC controllers for nonlinear systems with bounded uncertainties. *Automatica*, 42(5):797–803, 2006.

[140] Di Wu, Nikitha Radhakrishnan, and Sen Huang. A hierarchical charging control of plug-in electric vehicles with simple flexibility model. *Applied Energy*, 253:113490, 2019.

[141] Huiping Li and Yang Shi. Event-triggered robust model predictive control of continuous-time nonlinear systems. *Automatica*, 50(5):1507–1513, 2014.

[142] Lars Grüne and Jürgen Pannek. *Nonlinear Model Predictive Control: Theory and Algorithms*. Springer, 2011.

[143] Alina Eqtami, Shahab Heshmati-alamdari, Dimos V Dimarogonas, and Kostas J Kyriakopoulos. Self-triggered model predictive control for nonholonomic systems. In *Proceedings of European Control Conference (ECC2013)*, Zurich, Switzerland, Jul. 17-19, 2013.

[144] Kazumune Hashimoto, Shuichi Adachi, and Dimos V Dimarogonas. Self-triggered model predictive control for nonlinear input-affine dynamical systems via adaptive control samples selection. *IEEE Transactions on Automatic Control*, 62(1):177–189, 2017.

[145] Matteo Rubagotti, Antonio Estrada, Fernando Castaños, Antonella Ferrara, and Leonid Fridman. Integral sliding mode control for nonlinear systems with matched and unmatched perturbations. *IEEE Transactions on Automatic Control*, 56(11):2699–2704, 2011.

[146] Abdellah Lassioui, Hassan El Fadil, Aziz Rachid, Zakariae El-Idrissi, Tasnime Bouanou, Fatima Zahrae Belhaj, and Fouad Giri. Modelling and sliding mode control of a wireless power transfer system for BEV charger. *International Journal of Modelling, Identification and Control*, 34(2):171–186, 2020.

[147] Changxin Liu, Jian Gao, Huiping Li, and Demin Xu. Aperiodic robust model predictive control for constrained continuous-time nonlinear systems: An event-triggered approach. *IEEE Transactions on Cybernetics*, 48(5):1397–1405, 2017.

[148] Gian Paolo Incremona, Antonella Ferrara, and Lalo Magni. Asynchronous networked MPC with ISM for uncertain nonlinear systems. *IEEE Transactions on Automatic Control*, 62(9):4305–4317, 2017.

[149] Hong Chen and Frank Allgöwer. A quasi-infinite horizon nonlinear model predictive control scheme with guaranteed stability. *Automatica*, 34(10):1205–1217, 1998.

[150] Vadim Utkin and Jingxin Shi. Integral sliding mode in systems operating under uncertainty conditions. In *Proceedings of the 35th IEEE Conference on Decision and Control (CDC1996)*, Kobe, Japan, Dec. 1996. IEEE.

[151] Fernando Castaños and Leonid Fridman. Analysis and design of integral sliding manifolds for systems with unmatched perturbations. *IEEE Transactions on Automatic Control*, 51(5):853–858, 2006.

[152] Vadim I Utkin. *Sliding Modes in Control and Optimization*. Springer Science & Business Media, 2013.

[153] Greenhouse gas emissions 2020. https://www.canada.ca/en/environment-climate-change/services/environmental-indicators/greenhouse-gas-emissions.html. Accessed March 13, 2021.

[154] Emmanouil S Rigas, Sarvapali D Ramchurn, and Nick Bassiliades. Managing electric vehicles in the smart grid using artificial intelligence: A survey. *IEEE Transactions on Intelligent Transportation Systems*, 16(4):1619–1635, 2014.

[155] Filipe Joel Soares, PM Rocha Almeida, and JA Pecas Lopes. Quasi-real-time management of electric vehicles charging. *Electric Power Systems Research*, 108:293–303, 2014.

[156] Keenan Valentine, William G Temple, and K Max Zhang. Intelligent electric vehicle charging: Rethinking the valley-fill. *Journal of Power Sources*, 196(24):10717–10726, 2011.

[157] Hyo-Sung Ahn, YangQuan Chen, and Kevin L Moore. Iterative learning control: Brief survey and categorization. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 37(6):1099–1121, 2007.

[158] Frank P Kelly, Aman K Maulloo, and David KH Tan. Rate control for communication networks: Shadow prices, proportional fairness and stability. *Journal of the Operational Research Society*, 49(3):237–252, 1998.

[159] Ugo Rosolia and Francesco Borrelli. Learning model predictive control for iterative tasks. A data-driven control framework. *IEEE Transactions on Automatic Control*, 63(7):1883–1896, 2017.

[160] Stephen Boyd, Stephen P Boyd, and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[161] SCE Load Profiles. https://www.sce.com/regulatory/load-profiles/2019-static-load-profiles. Accessed February 3, 2020.

[162] Shima Nazari, Francesco Borrelli, and Anna Stefanopoulou. Electric vehicles for smart buildings: A survey on applications, energy management methods, and battery degradation. *Proceedings of the IEEE*, 109(6):1128–1144, 2020.

[163] NBG Brinkel, WL Schram, TA AlSkaif, I Lampropoulos, and WGJHM Van Sark. Should we reinforce the grid? Cost and emission optimization of electric vehicle charging under different transformer limits. *Applied Energy*, 276:115285, 2020.

[164] Ali Ahmadian, Mahdi Sedghi, Ali Elkamel, Michael Fowler, and Masoud Aliakbar Golkar. Plug-in electric vehicle batteries degradation modeling for smart grid studies: Review, assessment and conceptual framework. *Renewable and Sustainable Energy Reviews*, 81:2609–2624, 2018.

[165] Electric vehicle battery health. https://www.geotab.com/blog/ev-battery-health/. Accessed September 26, 2021.

[166] Vincenzo Marano, Simona Onori, Yann Guezennec, Giorgio Rizzoni, and Nullo Madella. Lithium-ion batteries life estimation for plug-in hybrid electric vehicles. In *Proceedings of IEEE Vehicle Power and Propulsion Conference*, Dearborn, MI, USA, Sept. 2009. IEEE.

[167] Samuel Pelletier, Ola Jabali, Gilbert Laporte, and Marco Veneroni. Battery degradation and behaviour for electric vehicles: Review and numerical analyses of several models. *Transportation Research Part B: Methodological*, 103:158–187, 2017.

[168] Jens Vetter, Petr Novák, Markus Robert Wagner, Claudia Veit, K-C Möller, JO Besenhard, Martin Winter, Margret Wohlfahrt-Mehrens, Christoph Vogler, and Abderrezak Hammouche. Ageing mechanisms in lithium-ion batteries. *Journal of Power Sources*, 147(1-2):269–281, 2005.

[169] DP Abraham, EM Reynolds, PL Schultz, AN Jansen, and DW Dees. Temperature dependence of capacity and impedance data from fresh and aged high-power lithium-ion cells. *Journal of the Electrochemical Society*, 153(8):A1610, 2006.

[170] Anthony Barré, Benjamin Deguilhem, Sébastien Grolleau, Mathias Gérard, Frédéric Suard, and Delphine Riu. A review on lithium-ion battery ageing mechanisms and estimations for automotive applications. *Journal of Power Sources*, 241:680–689, 2013.

[171] Bin Wang, Jun Xu, Binggang Cao, and Xuan Zhou. A novel multimode hybrid energy storage system and its energy management strategy for electric vehicles. *Journal of Power Sources*, 281:432–443, 2015.

[172] Jian Cao and Ali Emadi. A new battery/ultracapacitor hybrid energy storage system for electric, hybrid, and plug-in hybrid electric vehicles. *IEEE Transactions on Power Electronics*, 27(1):122–132, 2011.

[173] Zhongjing Ma, Suli Zou, and Xiangdong Liu. A distributed charging coordination for large-scale plug-in electric vehicles considering battery degradation cost. *IEEE Transactions on Control Systems Technology*, 23(5):2044–2052, 2015.

[174] Karl Schwenk, Stefan Meisenbacher, Benjamin Briegel, Tim Harr, Veit Hagenmeyer, and Ralf Mikut. Integrating battery aging in the optimization for bidirectional charging of electric vehicles. *arXiv preprint arXiv:2009.12201*, 2020.

[175] Ridoy Das, Yue Wang, Ghanim Putrus, Richard Kotter, Mousa Marzband, Bert Herteleer, and Jos Warmerdam. Multi-objective techno-economic-environmental optimisation of electric vehicle for energy services. *Applied Energy*, 257:113965, 2020.

[176] Martin Petit, Eric Prada, and Valérie Sauvant-Moynot. Development of an empirical aging model for Li-ion batteries and application to assess the impact of Vehicle-to-Grid strategies on battery lifetime. *Applied Energy*, 172:398–407, 2016.

[177] Benedikt Lunz, Hannes Walz, and Dirk Uwe Sauer. Optimizing vehicle-to-grid charging strategies using genetic algorithms under the consideration of battery aging. In *Proceedings of IEEE Vehicle Power and Propulsion Conference*, Chicago, IL, USA, Sept. 2011. IEEE.

[178] Yunfei Bai, Hongwen He, Jianwei Li, Shuangqi Li, Ya-xiong Wang, and Qingqing Yang. Battery anti-aging control for a plug-in hybrid electric vehicle with a hierarchical optimization energy management strategy. *Journal of Cleaner Production*, 237:117841, 2019.

[179] Alessandro Di Giorgio, Andrea Di Maria, Francesco Liberati, Vincenzo Suraci, and Francesco Delli Priscoli. Lagrangian decomposition based multi agent model predictive control for electric vehicles charging integrating real time pricing. *arXiv preprint arXiv:1607.02895*, 2016.

[180] Alessandro Di Giorgio, Alessandro Giuseppi, Roberto Germanà, and Francesco Liberati. Decentralised model predictive control of electric vehicles charging. In *Proceedings of 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, Bari, Italy, Oct. 2019. IEEE.

[181] Roberto Germanà, Francesco Liberati, and Alessandro Di Giorgio. Decentralized model predictive control of plug-in electric vehicles charging based on the alternating direction method of multipliers. In *Proceedings of 2020 28th Mediterranean Conference on Control and Automation (MED)*, Saint-Raphaël, France, Sept. 2020. IEEE.

[182] Mehdi Rahmani-Andebili, Massimiliano Bonamente, and James A Miller. Charging management of plug-in electric vehicles in San Francisco applying Monte Carlo Markov chain and stochastic model predictive control and considering renewables

and drag force. *IET Generation, Transmission & Distribution*, 14(25):6179–6188, 2020.

[183] Jian Wu, Xiangyu Wang, Liang Li, Yongchang Du, et al. Hierarchical control strategy with battery aging consideration for hybrid electric vehicle regenerative braking control. *Energy*, 145:301–312, 2018.

[184] Scott B Peterson, Jay Apt, and JF Whitacre. Lithium-ion battery cell degradation resulting from realistic vehicle and vehicle-to-grid utilization. *Journal of Power Sources*, 195(8):2385–2392, 2010.

[185] Samuel Filgueira da Silva, Jony Javorski Eckert, Fabrício Leonardo Silva, Ludmila CA Silva, and Franco Giuseppe Dedini. Multi-objective optimization design and control of plug-in hybrid electric vehicle powertrain for minimization of energy consumption, exhaust emissions and battery degradation. *Energy Conversion and Management*, 234:113909, 2021.