Model Predictive Control for Industrial Energy Flexibility: Applications to Beverage Production Plants and Bidirectional Electric Vehicle Charging Hubs

vom Fachbereich Maschinenbau und Verfahrenstechnik

der Rheinland-Pfälzischen Technischen Universität Kaiserslautern-Landau

zur Verleihung des akademischen Grades

Doktor-Ingenieur (Dr.-Ing.)

genehmigte Dissertation

von

Herrn Behzad Heydaryan Manesh

unter der Leitung von

Prof. Dr.-Ing. Naim Bajcinca

Datum der mündlichen Prüfung: 14. Juli 2026

Dekan:	Prof. Dr. rer. nat. Roland Ulber
Berichterstatter:	Prof. Dr.-Ing. Naim Bajcinca
	Prof. Dr. Veit Hagenmeyer
	Prof. Dr.-Ing. Stefan Götz

D-386

Kurzfassung:

Diese Dissertation konzentriert sich auf die Entwicklung neuartiger Algorithmen zur Integration industrieller und gewerblicher Verbraucher in Programme zur Bereitstellung von Regelenergie für den Regelenergiemarkt in intelligenten Stromnetzen. Es werden generische und anpassungsfähige Energiemanagementsysteme vorgeschlagen, die es Großverbrauchern ermöglichen, nicht nur ihren Energieverbrauch zu optimieren, sondern auch am Demand Side Management (DSM) teilzunehmen, um ihre Leistungs- und Energieflexibilität dem Regelenergiemarkt sowie dem Spotmarkt anzubieten. Die vorgeschlagenen Steuerungs- und Managementstrukturen basieren auf modellprädiktiver Regelung (Model Predictive Control, MPC), um den Energieverbrauch entsprechend der Tarife zu optimieren, die Vergütung bei der Bereitstellung von Systemdienstleistungen bzw. bei Energiehandelsaktivitäten sicherzustellen und die Vorgaben der Verbraucher zu erfüllen. Diese entgegengesetzten Zielsetzungen führen zu einem Zielkonflikt zwischen der Erfüllung der Anforderungen des Stromnetzes und der Berücksichtigung der Präferenzen der Verbraucher im Entscheidungsprozess. Mit dem Hauptziel, Netzungleichgewichte auszugleichen, die zu Frequenzabweichungen führen, berücksichtigt die Arbeit sowohl die primären als auch die sekundären Anforderungen des Stromnetzes. Darüber hinaus wurden auch weitere Demand-Response-Strategien wie Lastverschiebung (Load Shifting) untersucht. Anhand von zwei zentralen Anwendungsfällen – Getränkeproduktionsanlagen und Ladeinfrastrukturen für Elektrofahrzeuge – wurde die Wirksamkeit der vorgeschlagenen Regelalgorithmen untersucht. Dabei kamen klassische Optimierungsmethoden sowie zugrunde liegende Solver wie IPOPT und Gurobi zum Einsatz.

Preface

The evolution of energy systems in the context of growing demand and the increasing integration of renewable sources presents both opportunities and challenges for the power grid. As the modern power systems shift toward sustainable and flexible energy solutions, the ability to adapt and optimize energy consumption in real time has become a critical aspect of industrial and commercial energy management. This thesis explores innovative methodologies to manage energy consumption across multiple domains, such as bidirectional charging hubs and beverage production plants, through optimal control strategies. The research presented here primarily focuses on developing energy management frameworks utilizing model predictive control (MPC) to optimize energy consumption and enhance the interaction between various stakeholders in energy markets. The various MPC approaches presented in this research provide practical solutions for balancing flexibility with operational constraints, enabling mainly industrial and commercial consumers to participate in demand response (DR) programs and contribute to grid stability through the provision of ancillary services. Throughout my academic journey, I have been deeply intrigued by the intersection of energy management systems, control and optimization techniques, and real-world applications. This thesis represents the culmination of several years of research, investigation, and collaboration with leading experts in the field of control and power systems. It aims to contribute to the growing body of knowledge on DR and energy flexibility, providing a framework that can be adapted to various prosumers, from manufacturing to electromobility. Hereby, I would like to express my sincere gratitude to my supervisor, Prof. Dr.-Ing. Naim Bajcinca, for his guidance and support throughout this pivotal period of my academic career. His insights and feedback have been invaluable in shaping the work presented in this thesis. I am also especially grateful to Dr. Mohammad Al Khatib, whose collaboration, kindness, and unfailing support were a source of strength and motivation throughout my PhD.

Abstract

This thesis proposes an economically optimal generic energy management framework to enable industrial, commercial, and residential producers/consumers (prosumers) to offer their power and energy flexibility in different energy markets, with a particular focus on mid-term and real-time energy markets and two main use-cases: beverage production plants and bidirectional electric vehicle (EV) charging hubs. The proposed framework utilizes the theory of model predictive control (MPC) in a hierarchical structure with different prediction horizons, which enables long-term and mid-term energy planning as well as real-time operational control. The optimal control framework is developed to enable the utilization of prosumers' flexibility in the "control energy market" and "spot market" through flexibility aggregators. The change in prosumers' pattern of power or energy consumption is within a specific range of services, where the loads (the demand side of the power grid) play a vital role by contributing to demand response (DR) programs. In this regard, first, the lowest layer of the proposed hierarchical energy management system has been synthesized for the two main applications. Then, the core energy management systems of the use cases at this layer have been extended to a unified and more generic hierarchical framework with the capability to be adapted to the two use cases to enable the prosumers for DR service provision in different time scales. However, this generic structure can be adapted to any prosumer and is not necessarily limited to the use cases explained in this thesis. In the use case of beverage production plants at the lower layer, a centralized MPC for economically optimal energy consumption and DR control of the plants has been suggested to assign the processing speed of the units inside the lines using a mixed-integer nonlinear program (MINLP). Also, the dynamical models of the processing units in the production lines have been considered for design purposes. In the use case of the bidirectional EV charging hub, the lower layer has been proposed in two different structures to optimally manage energy transactions between the charging hub (equipped with an energy buffer), the grid, and the EVs while considering DR and drivers' preferences. Mixed integer optimization frameworks have been proposed for decision-making on charging or discharging the EVs. Finally, the generic hierarchical structure has been established and adapted to the use cases while exploiting mixed-integer optimization frameworks accordingly. The effectiveness of the proposed controllers for each use case has been investigated through various simulation scenarios. The solvers, such as IPOPT and Gurobi, have been exploited as the underlying tools for solving the MPC problems.

Contents

Preface i

Abstract ii

1 Introduction 1
1.1 Background and Motivation . . . 1
1.2 Problem Statement . . . 3
1.3 Research Objectives and Questions . . . 4
1.4 Methodology Overview . . . 5
1.5 Conceptual Overview and Novelty . . . 6
1.6 Thesis Structure . . . 9

2 State of the Art and Preliminaries 10
2.1 Literature Review . . . 10
2.1.1 Demand Response in Industrial Processes . . . 10
2.1.2 Energy Management for Bidirectional Charging of EVs . . . 13
2.1.3 Research Gaps and Opportunities . . . 16
2.2 Preliminaries . . . 16
2.2.1 Fundamentals of Demand Response . . . 16
Incentive-based DR . . . 17
Time/Price-based DR . . . 17
2.2.2 Model Predictive Control . . . 19
2.2.3 Energy Market . . . 23
2.3 Conclusion . . . 27

3 Demand Response Control of Beverage Production Plants 28
3.1 Introduction to Methodology . . . 28
3.2 System Modeling for Beverage Production Plants . . . 28
3.2.1 Dynamic Distributed Parameter Model . . . 30
Continuous Machines and Conveyor Belts . . . 31
Binary Machines . . . 34
3.2.2 Network Delay Model . . . 34
3.2.3 Model Identification . . . 36
Nominal Power Consumption of Binary Machines . . . 36
Model Identification via Linear Least Squares . . . 36
Identification Procedure . . . 37
Plots of Fitted Curves . . . 38
3.3 Demand Response Objectives and Problem Formulation . . . 40
3.4 Control Strategy Design with MPC . . . 41
3.5 Simulation Scenarios . . . 44
3.5.1 Scenario I: One Production Line . . . 45
3.5.2 Scenario II: Three Production Lines . . . 46
3.6 Conclusion . . . 49

4 Optimal Control Strategy for Bidirectional EV Charging Hubs **50**
4.1 Introduction to Methodology . . . 50
4.2 System Modeling and Energy Market . . . 50
4.2.1 Battery Abstract Model . . . 51
4.2.2 Spot Energy Market . . . 53
4.2.3 Multi-asset optimization for energy trading . . . 54
4.3 Optimal Control Strategy . . . 55
4.3.1 Single-layer Controller (C_2) . . . 57
4.3.2 Double-layer Controller (C_3) . . . 59
The Upper Layer . . . 60
The Lower Layer . . . 61
4.3.3 Impact of Battery Power Constraints on the MPCs' Formulation 63
4.4 Simulation Results . . . 65
4.4.1 Scenario I: One-layer MPC (C_2) . . . 66
4.4.2 Scenario II: Double-layer MPC (C_3) . . . 68
4.4.3 Scenario III: C_3 with Binary Battery Model . . . 69
4.4.4 Scenario IV: C_3 with Quadratic State-dependent Power Constraints . . . 70
4.5 Conclusion . . . 71

5 Hierarchical Energy Management for Ancillary Services **73**
5.1 Introduction to Methodology . . . 73
5.2 Generic Scheme . . . 73
5.3 Use-Case of Beverage Production Plants . . . 76
5.3.1 Design of C_2 . . . 77
5.3.2 Design of C_3 . . . 82
5.4 Simulation Results of Use-Case I: Production Line with a Battery . . . 84
5.5 Use-Case of Bidirectional EV Charging Hubs . . . 90
5.5.1 The High-level Controller (C_2) . . . 92
5.5.2 The Low-level Controller (C_3) . . . 95
5.6 Simulation Results of Use-Case II: Bidirectional Charging Hub . . . 97
5.7 Conclusion . . . 98

6 Conclusion and Future Work **100**
6.1 Overview of Research Objectives and Methodology . . . 100
6.2 Summary of Key Findings . . . 101
6.3 Implications for Industry and Energy Markets . . . 102
6.4 Future Research Directions . . . 104
6.5 Closing Remarks . . . 106

References **107**

List of Author's Publications **117**

List of Figures

1.1 Conceptual design of EMS for various prosumers. 7

2.1 Different types of DR in mid-term. 18
2.2 MPC Scheme. 20
2.3 Current configuration of the energy market along with the physical grid. 23
2.4 Power imbalance observed by TSO over a specific period. 25
2.5 Hierarchical regulation of the grid's frequency. 26
2.6 Future configuration of the energy market along with the physical grid. 26

3.1 A beverage production line. 29
3.2 Binary machines of the line. 30
3.3 The schematic picture of a filling machine, involving a feeding and delivery conveyor belt. 31
3.4 Internal control loop of the PMSM. 31
3.5 Bottle density function (β function) at time t (continuous) and $t + \Delta t$ (dashed). 33
3.6 Time forward and backward transformation. 34
3.7 A component of the line in NDM. 34
3.8 Measured data and identified power curve for the filler machine. . . . 38
3.9 Measured data and identified power curve for the labeling machine. . 39
3.10 Measured data and identified power curve for the shrink packer. . . . 39
3.11 Measured data and identified power curve for conveyor belt. 39
3.12 Measured data and identified power curve for the palletizer. 40
3.13 The proposed power-based MPC architecture. 41
3.14 An example of a conjunction in a production line. 42
3.15 The bottle loss during a demand service. 43
3.16 Power consumption characteristics in terms of production rate. 44
3.17 Power consumption, bottle production rate, the speed of the blow molding machine, and the filler of one line for primary service. 46
3.18 Three beverage production lines. 47
3.19 Power consumption, bottle production rate, the speed of the blow molding machine, and the filler of three lines for primary service. . . . 48

4.1 A schematic diagram of the charging hub structure. 51
4.2 DC fast charging characteristics of some EV manufacturers [111]. This figure illustrates the amount of power required for fast-charging EVs from different car brands in terms of SOC. Some have almost constant power during charging mode, while others change considerably when SOC increases. 52
4.3 Energy transaction directions. u_c^+ and u_g^+ is the energy sold to all the EVs and the grid, and u_c^- and u_g^- is the energy bought from all the EVs and the grid. 54

4.4 Charging power characteristic and its bound. This plot shows how the upper bound on the charging powers changes in terms of SOC. . . 56
4.5 Restructure the charging hub for MPC in the upper layer. 59
4.6 The proposed C_3 control structure for the hub. 60
4.7 The comparison of a single-layer architecture and a double-layer architecture of the controller C_3 in terms of average computation time of the solver. The number of EVs starts from 5 and increases up to 700. A single-layer C_3 cannot provide a real-time solution for 675 EVs or more. The sample time is chosen 1 second. 62
4.8 Power profile in binary MPC. 64
4.9 EVs' charging power limits (upper subplot) and energy tariffs (lower subplot). 66
4.10 Simulation results of scenario I. The charging hub's economic benefit is greater when it participates in DR, as shown in the third subplot. . . 67
4.11 Simulation results of scenario II. The charging hub's economic benefit is greater when it participates in DR, as shown in the third subplot. . . 68
4.12 Simulation results of Scenario III. The energy tariffs change three times, leading to four main charging/discharging phases. 69
4.13 Simulation results of Scenario IV. The power of charging/discharging changes quadratically with the SOC of the EVs. 70

5.1 Generic hierarchical architecture for energy management of prosumers. 75
5.2 Production plant with a buffer of bottles. 76
5.3 Hierarchical architecture for energy management of a beverage production plant with battery. 77
5.4 Temporal interpretation of target production constraint in Eq. (5.22). . 81
5.5 Power profiles and production rate of the line for scenario I, sub-case 1. The power profiles in the upper subplot are zoomed in Fig. 5.6 for the third primary demand. 85
5.6 Power profiles during the third primary demand in scenario I, sub-case 1. 86
5.7 Bottle Production, bottles in the buffer, and binary signal in scenario I, sub-case 1. 87
5.8 Power profiles and production rate of the line for scenario I, sub-case 2. 88
5.9 Bottle Production, bottles in the buffer, and binary signal in scenario I, sub-case 2. 88
5.10 Power profiles and production rate of the line for scenario I, sub-case 3. 89
5.11 Bottle production, bottles in the buffer, and binary signal in scenario I, sub-case 3. 90
5.12 Hierarchical control structure for EV fleet. 92
5.13 The sequence of energy packets (from C_2) and the first corresponding power profile (from C_3). 93
5.14 Architecture of the charging hub. 94
5.15 Reference power profile (P_{ref}) and real power consumption (P). The red area shows the amount of energy that should be shifted to the off-peak time, and the green area shows the energy that has been shifted. . 94
5.16 Energy profile and economic benefit of the charging hub when there is no demand and when there is a demand for load shifting in Scenario II. The control strategy shifts the consumption to the off-peak time. The economic benefit of the charging hub is greater when the demand is satisfied. 98

5.17 Power profiles of the charging hub when there is no demand and when there is a demand in Scenario II. The SOC curves of some random EVs in the charging hub are also demonstrated in the lower subplot. The EVs with less parking time have little to no participation in DR. 98

List of Tables

2.1 **Overview of the Reviewed Literature.** 15

3.1 **Simulation parameters of the production line for scenario I.** 45
3.2 **Simulation parameters of the production line for scenario II.** 47

4.1 **Simulation parameters of the EVs and corresponding demands.** . . . 66
4.2 **Comparison of features of the two main proposed controllers.** 71

5.1 **Simulation parameters of the production line for use-case I.** 85
5.2 **Simulation parameters of the EV charging hub for use-case II.** 97

6.1 **Overview of the proposed controllers, their objectives, and application scenarios** 103

Chapter 1

Introduction

1.1 Background and Motivation

The increasing complexity of modern power grids, driven by growing electricity demand and the integration of volatile energy resources, necessitates the deployment of advanced energy management systems (EMS). These systems are essential for optimizing energy utilization, minimizing operational costs, and providing ancillary services. EMS solutions enable real-time monitoring, control, and decision-making to balance supply and demand efficiently, ultimately improving grid resilience and economic performance. Grid stability and energy optimization have been managed through supply-side management (SSM) strategies, which focus on expanding electricity generation through the construction of new plants to match consumption demands [1], [2]. Power plants adjust their production levels based on forecasted consumption patterns and grid conditions. However, with the increasing penetration of intermittent renewable energy sources like solar and wind energy, supply-side adjustments alone are often insufficient to maintain grid balance. The volatility of these energy sources introduces fluctuations that can lead to grid instability, frequency deviations, and increased reliance on backup fossil fuel-based generation, which contradicts decarbonization goals [3], [4].

To address these challenges, demand-side management (DSM) has emerged as a complementary strategy that focuses on adjusting electricity consumption, rather than solely relying on supply-side interventions. DSM enables consumers—ranging from industrial facilities and commercial buildings to residential households—to adapt their energy usage patterns in response to price signals, grid constraints, or fluctuations in renewable energy availability. In fact, DSM refers to a set of strategies and programs that aim to optimize electricity consumption patterns by encouraging consumers to modify their energy usage [5]. By exploiting flexible loads, DSM enhances the grid's ability to integrate renewable energy while minimizing the need for expensive and polluting backup power sources. Unlike traditional grid management, which relies on adjusting power generation, DSM enables a more flexible and decentralized approach by using the adaptability of consumers, particularly in industrial and commercial sectors [6]. Beyond integrating renewable energy, DSM offers multiple benefits, which are outlined below.

One of the main advantages of applying DSM is ensuring grid stability. As mentioned above, the integration of volatile sources into the power grid leads to increased variability and unpredictability in power generation, posing challenges to grid stability. DSM addresses these challenges by enabling consumers to adjust their power consumption in response to grid conditions, thereby aiding in balancing supply and demand. This dynamic adjustment is crucial for maintaining grid frequency and preventing outages. DSM also offers consumers the opportunity to manage and reduce their energy expenses by modifying consumption patterns in response to

pricing signals. By participating in programs that offer time-of-use (TOU) tariffs or real-time pricing (RTP), consumers can shift their energy-intensive activities to periods when electricity prices are lower [7]. This not only leads to direct cost savings but also reduces peak demand, which can lower overall energy prices in the market. Additionally, DSM can decrease the need for investments in new generation and transmission infrastructure, further reducing costs for both utilities and consumers. New research indicates that effective DSM can lead to significant economic benefits by optimizing energy usage and reducing acquisition costs [8], [9], [10].

Demand response (DR) is one of the main methods in DSM that refers to the voluntary or incentivized modification of electricity consumption patterns by loads (consumers) in response to external signals (typically from utility control centers or system operators) aimed at maintaining grid stability or reducing peak demand [11]. In fact, DSM involves both short-term and long-term strategic measures to improve energy efficiency and reduce costs. DR, as one of the key strategies in DSM, is concerned explicitly with dynamic and short-term adjustments to electricity use in response to grid requests, price signals, or incentives. The research presented in this thesis is primarily situated within the scope of DR, particularly in the context of beverage production plants and bidirectional electric vehicle (EV) charging hubs. DR has gained increasing attention as the need for grid flexibility has become more pronounced due to factors such as the surge in overall demand in industrial sectors and the aging of grid infrastructure[12].

DR programs are broadly categorized into time/price-based DR and incentive-based DR, each offering distinct mechanisms for adjusting electricity consumption in response to grid conditions [13]. Time/price-based DR relies on dynamic electricity pricing structures, where consumers voluntarily respond to changes in prices and change their consumption accordingly. This approach includes TOU pricing, where electricity rates vary between peak and off-peak hours; critical peak pricing (CPP), which imposes significantly higher prices during periods of extreme grid stress; and RTP, where prices fluctuate based on wholesale market conditions. These mechanisms encourage consumers to shift their electricity consumption to periods of lower cost, thereby reducing peak demand and alleviating grid congestion. In industrial settings and large-scale facilities, time/price-based DR plays a crucial role in optimizing operational schedules to minimize electricity costs while ensuring uninterrupted service. On the other hand, incentive-based DR programs offer direct compensation or incentives to consumers for providing flexibility in power or energy consumption. These programs typically involve agreements where consumers allow grid operators or aggregators to control or request reductions in power or energy consumption directly.

Incentive-based DR encompasses various mechanisms that encourage consumers to adjust their energy usage in exchange for financial benefits. These include direct load control (DLC), where utilities remotely manage specific loads; interruptible/curtailable load (ICL), where large consumers reduce consumption upon request for incentives; demand bidding, where industrial or commercial users offer consumption reductions in electricity markets; and ancillary services participation (ASP), where consumers support grid stability through services like frequency regulation. This approach plays a crucial role in maintaining grid reliability, especially with the integration of smart grid technologies that can enable automated and efficient DR participation.

1.2 Problem Statement

Flexible industrial and commercial consumers in the power grid are among the main targets for integrating into the control energy market, where they can systematically offer their flexibility for financial revenue. This service provision requires optimal decision-making by the load operators to manage power consumption for active DR participation while considering load constraints. Considering this necessity for an energy management system, we have targeted the following problems to address in the thesis.

Problem 1: Beverage production plants, which consume considerable energy and operate under specific production schedules, present unique challenges when integrating DR strategies. A significant barrier to effective DR implementation in such plants is the lack of autonomous control systems capable of adjusting production processes in real-time according to DR signals. These plants are characterized by complex binary (on-off) machinery, such as blow molding machines, and continuous machinery, including fillers, labeling machines, and conveyor belts, which must operate with precise processing speeds to meet quality standards and customer demands. These constraints complicate shifting energy consumption or reducing power without risking production schedules. Current DR strategies in industrial plants often rely on manual interventions or rigid scheduling, which are insufficient for responding to dynamic grid signals or fluctuating electricity prices. In fact, many of these plants still lack an intelligent control framework that can autonomously and efficiently modulate energy usage without violating the production constraints. Without such a system, plant managers must rely on costly human oversight, making large-scale participation in DR programs both inefficient and economically unfeasible. The main challenge lies in developing advanced autonomous control algorithms, such as model predictive control (MPC), that can balance power or energy flexibility with production needs. These systems would enable the plant to adjust operations autonomously, ensuring minimal impact on production schedules while optimizing energy consumption in response to changing grid conditions.

Problem 2: Managing power and energy flexibility in bidirectional EV charging hubs also presents several unique challenges. The primary issue is integrating vehicle-to-grid (V2G) technology, which requires coordination between the EVs, the charging hubs, and the grid. Unlike traditional charging hubs, V2G-enabled hubs must reverse the flow of energy from vehicles and energy buffers back to the grid when necessary. This dual flow of energy requires careful coordination to ensure that EV charging is optimized alongside the grid stability goals of DR programs. Another challenge lies in optimizing the energy exchange between the EV fleet, stationary energy buffers, and the grid. The energy storage capacity of EVs, along with their availability (i.e., when they are plugged in and ready for energy exchange), introduces a challenge in the power management process. Each EV's battery state of charge (SOC), charging preferences, and user schedules must be carefully accounted for to avoid potential disruptions. Similar to beverage production plants, an autonomous control system capable of managing the entire process is lacking. In EV charging hubs, coordinating the optimization of charging, discharging, and grid participation requires advanced control algorithms that not only manage energy transactions according to energy tariffs in the market but also integrate the hubs in DR.

1.3 Research Objectives and Questions

As mentioned earlier, we aim to develop advanced control strategies that optimize the energy management of industrial and commercial prosumers willing to offer their flexibility into the control energy market in the context of DR programs and system service provision. More precisely, the goal is to design advanced autonomous control systems for managing the power and energy consumption of active loads in DR, such that, using time/price-based or incentive-based DR strategies, these loads participate in the control energy market by offering their flexibility in an economically optimal manner. The control structure must be designed to capture not only the grid's requests with different time scales, but also the dynamics and operational constraints of the active loads. This objective has been pursued for two main applications: beverage production plants and bidirectional EV charging hubs.

Participation of beverage production plants and bidirectional EV charging hubs in DR programs enhances both grid stability and energy market integration. Pursuing our goal for these use cases has been approached as follows: first, to develop an MPC framework for beverage production plants that enable them to contribute to DR by adjusting processing speed of the units while maintaining production plans, and second, to integrate optimized asset allocation within the context of MPC for bidirectional EV charging hubs, allowing the hub operator to make optimal energy transactions based on market prices while simultaneously supporting DR initiatives. To achieve these overarching objectives, the thesis focuses mainly on the following specific items:

1. Developing an economic MPC framework for beverage production plants:
    - To design a control strategy that adjusts the processing speed of machines in response to DR signals while ensuring that production schedules are met.
    - To establish a trade-off between flexibility revenues and production losses by incorporating flexibility contracts into the optimization problem.
    - To simulate and validate the proposed control strategy using identified models and production constraints.
2. Integrating multi-asset optimization with MPC for energy transactions in bidirectional EV charging hubs:
    - To formulate an optimization framework that enables a bidirectional EV charging hub operator to maximize economic benefits while ensuring flexibility provision for grid services and meeting EV user preferences.
    - To ensure computational tractability by applying relaxation techniques to the optimization problem.
    - To demonstrate the effectiveness of the proposed approach for energy trading and DR participation through simulation studies.
3. Developing a generic MPC-based hierarchical EMS to deal with consumption planning and real-time control of a prosumer:
    - To develop a hierarchical structure in three main layers according to energy market requirements and different DR time scales

- To adapt the proposed hierarchical structure to the use case of a beverage production plant equipped with an energy storage unit
- To adapt the proposed hierarchical structure to the use case of EV charging hub
- To evaluate the performance of the adapted configurations for each use case through simulating various scenarios

To address the above objectives, we investigate the following key research questions:

1. How can the operational flexibility of beverage production plants be exploited for participating in DR without violating production constraints?
2. What are the trade-offs between production losses and economic incentives when optimizing power flexibility in beverage production plants?
3. How to design an MPC framework to make optimal energy transactions for bidirectional EV charging hubs?
4. How can a generic hierarchical MPC structure efficiently coordinate energy trading and power distribution while ensuring real-time implementation for both targeted systems (beverage plants and EV charging hubs)?
5. What are the economic and technical implications of participating in DR events for both use cases when developing the generic hierarchical MPC?

1.4 Methodology Overview

In the first part of the thesis, the focus is on developing MPC for beverage production plants, enabling them to optimize energy consumption in response to energy market signals while maintaining production constraints. Beverage plants cannot supply energy back to the grid; therefore, their flexibility lies in adjusting processing speeds to reduce or shift power consumption strategically. The methodology for these plants involves:

- Developing dynamic models that capture the relationship between processing speed, power consumption, and production throughput derived from real-world measurements of beverage processing machines.
- Designing an MPC that optimally adjusts machine processing speeds based on energy price fluctuations and DR signals, ensuring cost-efficient operation.
- Introducing power flexibility contracts, where beverage plants trade power flexibility with the grid by adjusting consumption profiles in exchange for financial revenues.

The second part focuses on bidirectional EV charging hubs, where energy flexibility is used for both energy trading and DR participation. Unlike beverage plants, EV hubs possess an additional degree of freedom—they can buy and sell energy using V2G technology, making energy transactions a crucial aspect of their operation. This requires a control structure that combines multi-asset optimization with real-time power management. Therefore, the methodology for these hubs includes:

- Modeling a multi-asset optimization that determines energy transactions in spot and balancing markets based on energy prices, market conditions, and system constraints. The goal is to maximize economic profitability while ensuring grid-supporting flexibility.
- Designing an MPC-based power management strategy optimizes charging and discharging profiles for individual EVs, ensuring grid constraints are met while respecting vehicle availability and user preferences.
- Using some relaxation techniques that ensure computational efficiency and enable real-time operation of the proposed approach.

In the third part, we seek a generic hierarchical solution for a prosumer that integrates the proposed design from the previous parts and can be adapted to the prosumer's specifications according to its operational constraints. As a result, this part contains:

- A hierarchical MPC approach for beverage production plants, where a higher-level controller sets optimal processing rate (or equivalently energy packets), and a lower-level controller refines machine-level adjustments to meet operational requirements such as target bottle production.
- A hierarchical MPC approach for EV charging hubs, where a higher-level controller sets optimal energy packets, and a lower-level controller refines power profiles to meet operational requirements such as target SOC.

To evaluate the effectiveness of the proposed methodologies, simulation studies have been conducted. The validation process includes:

- Implementing the MPC algorithm for beverage production plants in MATLAB using CasADi toolbox and IPOPT for solving the nonlinear optimization problems.
- Implementing the MPC algorithm and multi-asset optimization framework for EV charging hubs in MATLAB using YALMIP toolbox and Gurobi to solve mixed-integer programs (MIPs).
- Assessing the impact of the proposed strategies on economic benefits, energy flexibility provision, and system preferences, to ensure feasibility in real-world applications.

1.5 Conceptual Overview and Novelty

The controllers that are developed for the two main use-cases align in a conceptual design of EMS for prosumers at the demand side of the power grid. The conceptual structure of the thesis is shown in Fig. 1.1, where the main hierarchical EMS for prosumers is depicted. As illustrated in this figure, we have proposed a three-layer EMS for managing power and energy consumption of flexible prosumers (industrial, commercial, and residential prosumers) that copes with the energy markets in different time scales.

The prosumer is shown with "plant" in the figure, as usually used for the system we aim to control in the context of control theory. In this generic structure, a top-down approach for managing the prosumers as well as the requests originating

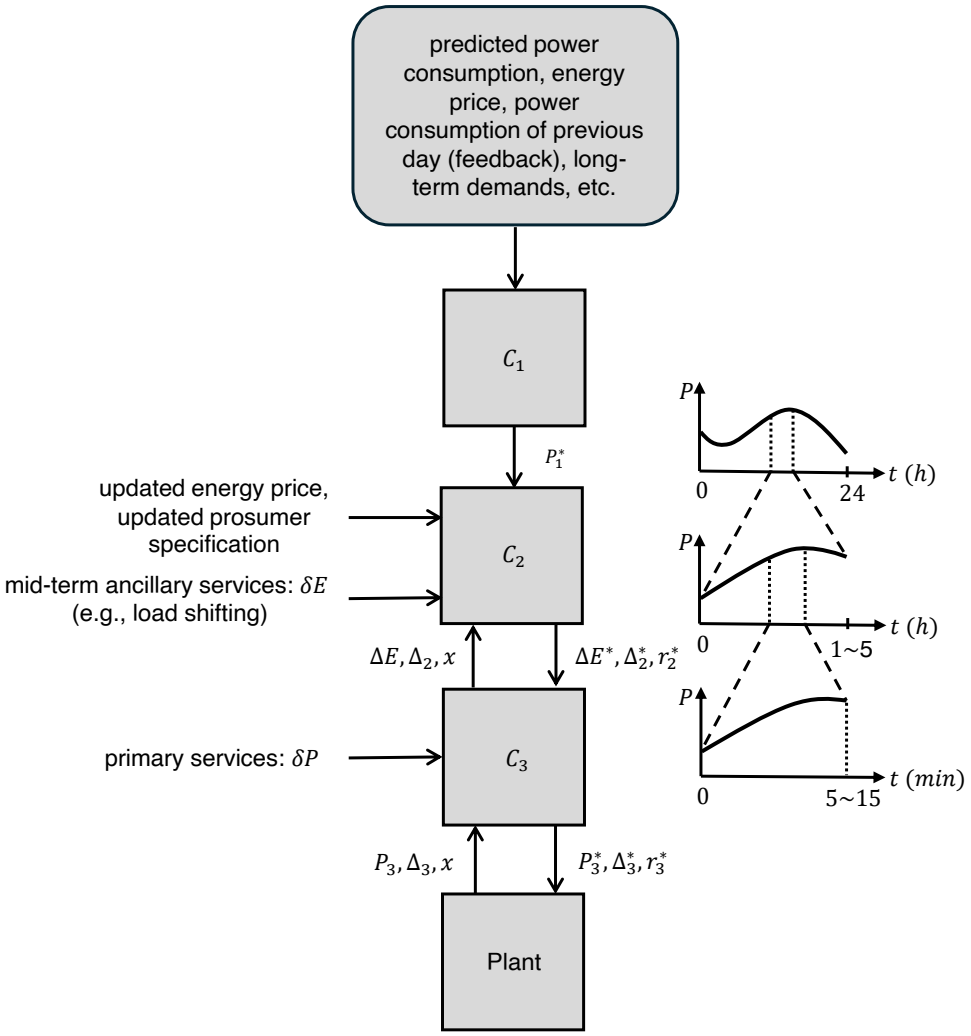


FIGURE 1.1: Conceptual design of EMS for various prosumers.

from different energy markets is proposed. This structure has been explained in detail in Chapter 5; however, a short description is provided here to create the conceptual overview of the thesis and its novelty. Starting from the top, the C_1 controller receives the required information to optimize the day-ahead power production of the prosumer and purchase it from the day-ahead market. This optimized power profile (P_1^*) as well as the updated energy tariffs, updated prosumer specifications, mid-term ancillary services, and feedback from C_3 controller are received by C_2 controller to modify optimally the day-ahead power profile coming from C_1 controller and find the optimal energy packets (ΔE^*) and other reference signals of the plant (Δ_2^*, r_2^*). This decision-making is accomplished in a shorter time interval, while the C_2 controller deals with the spot energy market with a more dynamic environment than the day-ahead market. Then, the controller at the lowest level (C_3) seeks the optimal power adjustments (P_3^*) with the highest changing rate (smallest sample time) while having the optimal decision from the C_2 controller, the primary demands from the grid (δp), and feedback from the plant (x).

We pursue the main concept illustrated in Fig. 1.1 in the thesis as follows: First, we develop the C_3 controller for a beverage production line to set up a real-time MPC solution for this use-case. Then, we establish two MPC frameworks, one at

the level of C_2 and one at the level of C_3 for EV charging hubs as the second use-case. Having the control frameworks established for these use-cases, we develop a generic hierarchical solution with a focus on C_2 and C_3 controllers for any flexible prosumers. Finally, these C_2 and C_3 controllers are adapted to the two main use-cases. The chapters are sorted according to this step-by-step conceptual design, and the structure of the thesis is explained in more detail in Section 1.6. The highest layer (C_1) is intentionally treated at a conceptual level and serves to illustrate how long-term forecasts, uncertainty-aware decision-making, and advanced methods such as stochastic MPC or emerging learning-based approaches can be integrated into the proposed framework. Although the explicit synthesis of a C_1 controller is not addressed in this thesis, its interface and role are clearly defined, ensuring compatibility with the developed C_2–C_3 control layers.

Novelty: The primary innovation of the thesis lies in the systematic synthesis of market-consistent MPC formulations for multi-time-scale power and energy management of flexible industrial, commercial, and residential loads and prosumers. The proposed structure enables the consistent translation of heterogeneous market signals and ancillary service requirements into physically feasible control actions, ensuring that fast power adjustments for balancing services remain compatible with previously committed energy schedules and underlying system dynamics. At the core, the contribution of this thesis lies in the market-oriented synthesis and dynamic instantiation of MPCs at the individual layers, and in demonstrating how these controllers can be coherently combined into a generic framework for multi-time-scale market participation.

Specifically the novelties can be discriminated in three sections. First, rather than introducing a purely conceptual hierarchical architecture, the thesis develops and validates concrete MPC formulations at various control layers and application domains, which subsequently motivates a generic, unified hierarchical framework. At the lowest control layer (C_3), the thesis develops an MPC formulation for primary DR in beverage production plants (Chapter 3), explicitly incorporating processing machines' dynamics, operational constraints, and real-time ancillary service requirements in small time scales. On the one hand, typically grid suppliers never take into account the process dynamics and constraints in the DR requirements. On the other hand, the production enterprises cannot neglect the dynamic effects corresponding to power spikes and for that matter feasibility constraints. Therefore, the first innovation provides a systematic methodology of integrating DRs in running production schemes, while also considering production requirements. In this sense, proper control actions can avoid the related overshoots that disturbs the controlled production. To this end, the controller C_3 exploits industrial production flexibility for real-time balancing services while preserving the feasibility of production constraints.

As a second contribution, in Chapter 4, we extend this philosophy to flexibility management of emerging V2G technology featuring a large flexibility potential in the future. Hereby, we introduce not only the C_3 controller with a fine sample time but also the C_2 control layer which operates in mid-term time scales. These controllers (i.e. C_2 and C_3 individually) implement multi-asset energy management, enabling coordinated participation in spot and balancing markets, respectively, while accounting for battery dynamics, charging constraints, and aggregated EV flexibility.

The insights gained from these application-specific MPC designs form finally the basis for the generic hierarchical control framework introduced in Chapter 5. In this chapter, the thesis combines the previously developed control concepts into a unified, modular architecture for flexible loads/prosumers, in which energy- and

power-oriented decision-making are explicitly separated across time scales and consistently coupled through shared state information and dynamic constraints. Although the resulting MPC formulations in Chapter 5 are not direct replicas of those in Chapters 3 and 4, they are systematically informed by the modeling choices, cost-function structures, and constraint formulations established therein.

1.6 Thesis Structure

The thesis is organized into six chapters, each focusing on different aspects of the research related to the optimization of beverage production systems and bidirectional EV charging hubs for participation in DR programs and energy market integration. It is crucial to note that the development of the proposed generic hierarchical MPC structure requires separate design approaches and foundations for each use case. Therefore, the subsequent chapters are structured as follows:

- **Chapter 2: Literature Review & Theoretical Preliminaries**

  This chapter reviews the existing literature on industrial DR, MPC for industrial DR, energy flexibility in industrial processes, and bidirectional EV charging, as well as the application of multi-asset optimization for the energy sector. It covers both theoretical preliminaries and literature review, identifying key challenges and research gaps that we address.

- **Chapter 3: Demand Response Control of Beverage Production Plants**

  The third chapter presents the development of an MPC strategy for beverage production plants. It outlines the dynamic models that link processing speed with power consumption and production throughput and explains how the MPC is designed to enable these plants to participate in DR events while not violating production constraints (see [14], [15], [16]). It mainly targets designing C_3 in Fig. 1.1 for beverage production lines.

- **Chapter 4: Optimal Control Strategy for Bidirectional EV Charging Hubs**

  This chapter focuses on bidirectional EV charging hubs, where the energy flexibility of the system is used for both DR participation and energy trading. This chapter describes two main MPC frameworks where energy transactions are determined based on multi-asset optimization, and the controller manages the charging and discharging of individual EVs (see [17], [18]). It mainly targets designing C_2 and C_3 in Fig. 1.1 for EV charging hubs.

- **Chapter 5: Hierarchical Energy Management for Ancillary Services**

  A generic hierarchical MPC framework (Fig. 1.1) aimed at managing energy and power consumption of flexible prosumers (applicable to both use-cases) is explained in this chapter. It describes a structure that enables long-term planning of energy transactions and load shifting, while assigning precise and high-resolution power profiles for prosumers (see [19]).

- **Chapter 6: Conclusion and Future Work**

  The final chapter summarizes the research findings, discusses the contributions to the field of DR and energy market integration, and outlines potential areas for future research.

Chapter 2

State of the Art and Preliminaries

A thorough understanding of existing research results and fundamental principles is essential for developing effective energy management strategies. This chapter is divided into two main sections. Section 2.1 reviews the literature on DR in industrial processes and bidirectional EV charging hubs, highlighting relevant methodologies and identifying research gaps. Section 2.2 presents the required preliminaries on MPC and the concept of the energy market, providing the foundation for the proposed methodologies in subsequent chapters. By integrating both the literature review and theoretical foundations, this chapter aims to establish the basis for the methodologies developed in Chapters 3 to 5.

2.1 Literature Review

As mentioned in Chapter 1, our goal is to focus on enabling industrial processes to participate in DR programs and bidirectional EV charging hubs to trade energy and contribute to DR events. This section provides a comprehensive review of existing research results, focusing on two key areas: (1) DR in industrial processes, particularly in beverage production plants, and (2) energy management strategies for bidirectional EV charging hubs, emphasizing their role in grid interaction and energy trading. By examining prior work, we identify existing methodologies, their limitations, and the need for advanced optimization frameworks that enhance both economic efficiency and grid stability.

2.1.1 Demand Response in Industrial Processes

Industrial facilities, particularly energy-intensive manufacturing plants, offer significant flexibility potential due to their large and controllable loads. An energy management system for industrial facilities based on the state task network and mixed integer linear programming (MILP) is proposed in [20], where the processing tasks are divided into nonschedulable and schedulable tasks, and it uses distributed energy resources to implement DR. Reference [21] presents the optimal scheduling models for industrial power load in smart grids as well as the solving methods of the corresponding optimization problems. Research by [22] highlights the barriers to DR adoption in industrial and commercial sectors, identifying concerns such as product quality, production disruptions, and revenue uncertainty. The study emphasizes the need to bridge knowledge gaps and adapt policies to overcome these obstacles, thereby promoting more involvement in DR programs. Also, a survey by [23] explores the role of aggregators in facilitating industrial DR, emphasizing their contribution to raising awareness and supporting implementation within industrial companies. An agent-based structure to integrate the flexibility of industrial and

residential DR, involving a central provider coordinating flexible plans for both industrial and residential aggregators, is presented in [24]. The approach, tested on the Danish sector of the Nordic electricity market, demonstrates how integrated flexibility can help stabilize power systems against the volatility of renewable energy. A study by [25] presents an optimization framework for cooperative industrial DR, where the industrial process incentivizes customers to adjust their schedules to better align with utility costs. The work demonstrates that cooperation leads to greater cost savings, especially when the industrial process has limited product storage, and the customer has high flexibility, using a distributed algorithm to solve the problem with minimal information sharing. The role of industrial loads in DR, specifically cement plants, is addressed in [26], highlighting their ability to adjust power consumption, but discrete power changes limit them. This reference proposes methods to enhance the granularity of power adjustments, enabling industrial loads to provide valuable ancillary services, such as regulation and load following, with the support of onsite energy storage systems.

The study in [27] presents an intelligent energy management framework with DR capabilities for industrial facilities, integrating components like industrial processes, thermostatically controlled loads, and renewable generation. The framework, optimized using MILP under TOU pricing, effectively reduces electricity costs and peak demand while meeting operational constraints, as demonstrated in a case study of tire manufacturing. A model-free DR scheme for industrial facilities, using a Markov decision process (MDP) and an actor-critic-based deep reinforcement learning algorithm, is proposed in [28] to optimize energy consumption. The proposed approach successfully reduced energy costs without compromising production, as demonstrated in a real-world industrial application. Also, a real-time price-based DR algorithm for industrial facilities is suggested in [29], utilizing a data-driven approach with long short-term memory recurrent neural networks (LSTM RNN) to forecast price uncertainties. The energy management is formulated as an MILP and successfully applied to a steel powder manufacturing process, demonstrating cost reduction and effective energy shifting while satisfying operational constraints. The research in [30] presents a dynamic simulation tool for industrial customers to assess the cost-benefit analysis of implementing DR strategies, aiming to mitigate the impact of intermittency of renewable energy sources. The tool evaluates industrial participation in reserve energy markets, considering technical restrictions, extra costs, and process-specific DR, with its application to a German paper factory.

A chapter by [31] provides an overview of DR technologies in the U.S., reviewing specific DR resources across industrial sectors. It characterizes attributes such as response frequency, response time, and load curtailment magnitude, and discusses enabling controls and communications technologies for DR participation. A cooperative DR scheme for load management in smart grids, proposed by [32], is formulated as a constrained optimization problem to generate a Pareto-optimal response strategy, reducing electricity costs for consumers. The scheme, applied to industrial refrigerated warehouses, is shown to outperform noncooperative DR strategies by reducing costs, electricity prices, and total energy consumption. Article [33] examines the barriers and motivations for industrial consumers participating in DR, identifying financial benefits as the primary motivation while also highlighting the opportunity to communicate non-financial benefits like preventing blackouts and managing renewable generation. The research in [34] reviews industrial DR opportunities in energy-intensive industries, highlighting flexibility potentials in cement manufacturing, aluminum smelting, and oil refinery industries. The paper discusses power and joint power-heat flexibility, noting significant energy cost reductions of

up to 34.2% and 20.7% for aluminum smelting and cement plants, respectively, and emphasizing the challenges and opportunities of industrial demand management in power systems.

DSM and DR are examined in [35], focusing on shifting and reducing electricity consumption for residential, commercial, and industrial customers. For industrial processes that face constraints such as production and inventory requirements, a fuzzy/expert-based system is proposed to rank loads and workstations dynamically for demand reduction. The system evaluates whether and by how much the plant can participate in a utility-initiated DR event while adhering to local operational constraints. Reference [36] proposes a two-level decision-making tree for assessing the participation of industrial processes in DR. The first level classifies processes using an analytic hierarchy process (AHP), while the second level estimates participation based on specific factors. A desirability coefficient helps system operators evaluate the attractiveness of incentive-based programs. Also, a data-driven framework for sustainable, intelligent manufacturing based on DR is presented in [37], aimed at energy-intensive industries in the Industry 4.0 context. Advanced technologies, such as cloud computing, IoT, and big data analytics, are utilized in this context to implement cleaner production strategies. The framework incorporates multi-level demand response models at the machine, shop floor, and factory levels to minimize energy costs.

The paper [38] presents a modified DR approach based on the interactive TOU model, enabling industrial customers and electricity suppliers to optimize their performance. The proposed model identifies peak and off-peak hours by analyzing the region's electrical load and industrial production/sales profiles. A multi-agent system for industrial companies to offer load flexibility in DR markets is described in [39], which helps DR aggregators integrate this system into virtual power plants, thereby addressing the gap between research and real-world implementation. The system is designed based on German DR market requirements to demonstrate its potential for future real-world adoption.

Some papers have targeted DR for industrial production lines. The research in [40] presents DR-Mgmt, a framework for optimizing industrial DR in asynchronous production lines. The goal is to balance production rates and power consumption during DR events. DR-Mgmt models machine activities using a temporal deterministic finite station machine approach and applies a heuristic local search to find near-optimal schedules. Also, the study in [41] suggests a production and maintenance scheduling model that integrates TOU electricity prices, balancing production capability, maintenance, and DR participation. An aggregate cost model is also formulated to optimize cost-effectiveness, and case studies show up to 19% cost savings compared to minimizing electricity costs alone and 14% savings over maintenance-focused strategies. The reference [42] presents a linear piece-wise approximation (LPWA) model for energy use in aluminum, steel, and cement production. Integrated into a microgrid's day-ahead energy management, the model enables industrial participation in energy and ancillary service markets through DR. The method presented in [43] optimizes batch and continuous task scheduling in copper production using an MILP based on the resource-task network (RTN) approach. Results show that optimal scheduling can reduce annual electricity costs by up to 14.2% without reducing production. Electrolytic refining offers the highest cost-saving potential, with off-gas handling and air separation also contributing. Claiming that the rise of fast-charging material handling equipment (MHE) complicates energy management in energy-intensive manufacturing, [44] proposes an integrated DR scheduling model that considers machine-MHE interactions, production demands,

battery charging, and dynamic electricity pricing. Case studies show the model achieves a 15.1% reduction in energy costs compared to conventional scheduling.

In [45], the role of DR in enhancing flexibility within the Norwegian energy system was evaluated using a stochastic TIMES model. The study highlighted that space heating in residential buildings provides the most cost-efficient DR potential, and that DR contributes to smoothing electricity prices, optimizing hydropower reservoir use, and increasing cross-border electricity trade. By exploiting DR, Norway can support European decarbonization goals through flexible electricity exports and reduced system costs. Also, the authors in [46] reviewed the growing role of artificial intelligence (AI) and optimization algorithms in enabling effective DSM. The study highlighted the complexity of DSM tasks and emphasized AI's potential to support real-time decision-making, dynamic pricing, consumer targeting, and device control. It also discussed challenges and opportunities in developing and deploying DR optimization strategies, particularly in the context of developing countries, and underscored the importance of selecting appropriate algorithms to minimize energy consumption costs.

2.1.2 Energy Management for Bidirectional Charging of EVs

Bidirectional EV charging hubs using V2G technology can actively participate in energy trading and DR by strategically controlling energy transactions. This capability allows them to support grid stability, mitigate peak demand, and generate revenue through energy arbitrage and ancillary services. Effective energy management in bidirectional charging hubs requires optimization strategies that balance multiple objectives, including cost minimization, DR service provision, and user convenience.

The study in [47] proposes a cost-minimizing dynamic programming (DP) approach for plug-in hybrid EV (PHEV) owners. Due to DP's computational complexity, a state-independent charging/discharging policy is introduced and proven optimal using stochastic inventory theory. A backward iteration algorithm ensures practical implementation, and from the grid operator's perspective, an optimal PHEV charging scheme flattens the load profile. Also, [48] proposes a control algorithm to balance electricity supply and demand using EV batteries as dynamic storage. A multi-objective optimization of charging and discharging is conducted using a genetic algorithm to maximize system efficiency and optimize bidirectional energy flows between homes, EVs, and the grid. The research in [49] proposes a voltage stability design integrating PV, wind turbines, EV storage, and the grid to mitigate voltage instabilities. A power management strategy coordinates multiple energy sources, using V2G technology for stability and sustainability. The results demonstrate effective synchronization under varying load and renewable conditions. The impacts of bidirectional EVs on power quality, voltage profile, and frequency stability have been investigated in [50], where a detailed model measures harmonic distortion in grid current, while an optimal frequency support strategy enhances grid stability in V2G mode. The results show that proper planning of EVs improves voltage and frequency stability while maintaining power quality within IEEE standards. Reference [51] explores the role of EVs in addressing energy crises and reducing climate change. It examines the current status and challenges of EV infrastructure, charging power levels, and topologies, comparing unidirectional and bidirectional charging systems. The study highlights the benefits and drawbacks of EVs, offering solutions to mitigate negative impacts and emphasizing their positive contributions.

The method proposed in [52] explores solutions to optimize the integration of EVs into the power grid by using slow charging stations and smart energy management systems to minimize grid strain. It also investigates utilizing EV batteries for V2G, vehicle-to-home (V2H), and other services to enhance grid stability and power quality. The reference [53] proposes electricity load scheduling algorithms for a power system with an aggregator and multiple customers with EVs, focusing on joint load scheduling for appliances and energy trading. It introduces both collaborative and non-collaborative approaches, with the collaborative approach maximizing social welfare and the non-collaborative approach addressing fairness using a tiered billing scheme, considering uncertainty in load demands. The article [54] proposes an energy management strategy for home photovoltaic (HPV) systems integrated with EV batteries to mitigate HPV intermittency and manage power quality issues caused by EV charging. The strategy efficiently charges the EV batteries either through the HPV or the grid, supports V2G for grid stability, and enhances the self-consumption of the HPV system while reducing peak power demand. Besides, an EMS for a DC distribution system integrating plug-in EVs with an AC grid through bidirectional converters is proposed in [55]. The EMS enables plug-in EV owners to select between fast and optimal energy exchange options while managing power flow to prevent grid issues, taking into account factors such as SOC, trip plans, and optimizing charging costs or revenues. The study in [56] presents a user-centric algorithm for bidirectional smart charging of EVs, considering user preferences, peer-to-peer (P2P) energy trading, and ancillary services to the grid. The model improves scheduling by incorporating user input, optimizing battery SOC, and enhancing EV participation in energy transactions and ASP.

Optimized energy management strategies have been suggested in [57] for EVs and battery swapping stations in a smart microgrid, considering both grid-connected and islanded modes. A price-incentive model coordinates EV and battery swapping stations' charging to minimize costs and maximize profits, while a fuzzy control method adjusts service prices based on battery SOC. [58] studies an EMS for smart microgrids that optimally allocates EVs, distributed generations, and responsive loads. It introduces a price-sensitive, priority-based smart charging approach, considers EVs for ancillary services like reactive power provision, and integrates automatically controlled switches for hourly reconfigurations, with numerical analysis validating its effectiveness. The chapter presented in [59] explores the impact of plug-in EV charging on power distribution systems and grid reliability. It highlights the challenges posed by renewable energy fluctuations and proposes energy management strategies to optimize charging costs and balance the grid. The method in [60] proposes a scheduling algorithm for EV charging stations to manage congestion and optimize power dispatch using chance-constrained programming. The method integrates TOU pricing and priority-based scheduling, showing a considerable cost reduction compared to uncoordinated charging.

The work presented in [61] is about a hierarchical MPC framework for energy management in smart grids that integrates EVs for load-frequency control, economic operation, and balancing of renewable energy. The proposed two-layer HiMPC particularly targets different time scales for demands and coordination between short-term and long-term control layers. It uses aggregators to manage large EV fleets, and the main focus is on optimization-based control. Also, [62] investigates realistic V2G strategies to balance EV mobility needs with grid costs, using real charging data and future time-varying electricity prices. The reviewed papers are listed in the Table. 2.1 based on five tags defined, including: application domain, control/optimization method, market participation level, flexibility sources, and time scale.

TABLE 2.1: **Overview of the Reviewed Literature.**

Application domain:
• Industrial: [20], [21], [24], [25], [26], [27], [29], [30], [35], [36], [37], [40], [41], [42], [43], [44], [45]
• Commercial: [47], [48], [49], [50], [51], [52], [53], [54], [55], [56], [57], [58], [59], [60], [61], [62]
Control/optimization method:
• Classical: [20], [27], [29], [42], [47], [48], [55], [56], [58], [60], [61]
• Machine learning/AI-based/heuristic: [28], [29], [35], [36], [37], [40], [41], [44], [46], [57]
• Multi-agent/game theory/survey based: [21], [22], [23], [24], [25], [31], [32], [33], [34], [39], [53]
Market participation level:
• Centralized wholesale markets: [20], [24], [26], [30], [42], [47], [50], [54], [58], [61]
• Decentralized & retail markets: [27], [29], [32], [36], [38], [53], [56], [60]
• Conceptual frameworks: [21], [22], [33], [35], [41], [44]
Flexibility sources:
• Task scheduling/storage units: [20], [25], [26], [27], [29], [37], [40], [41], [43], [47], [48], [49], [50], [51], [52], [53], [54], [55], [56], [57], [58], [59], [60], [61]
• Consumer's flexible load/facility: [24], [25], [27], [32], [52], [57], [58]
• Industrial equipment: [26], [30], [35], [37], [44]
Time scale:
• Operational: [20], [27], [28], [29], [37], [42], [46], [48], [50], [60], [61]
• Strategic planning: [31], [34], [38], [41], [43], [45], [62]
• Short-term/event-based: [21], [23], [32], [33], [35], [39], [40], [44], [47], [49], [51], [52], [53], [54], [55], [56], [57], [58], [59]

2.1.3 Research Gaps and Opportunities

Despite extensive research on DR, significant gaps remain in integrating automatic DR-based control methods, such as MPC, within industrial production environments, particularly in beverage production plants. Existing approaches predominantly rely on heuristic or rule-based methods, offering limited adaptability to dynamic market signals, system dynamics, and production variability ([28], [29], [35], [36], [37], [40]). Moreover, while the potential of industrial DR is recognized, there is insufficient exploration of how predictive algorithms can systematically trade off production efficiency and flexibility revenue across different energy markets ([33], [34]). Similarly, in the context of bidirectional EV charging hubs, systematically designed hierarchical MPC frameworks have rarely been applied for energy management [62]. They neglect the integration of financial multi-asset optimization techniques that could enhance market participation and financial revenue streams.

Current models typically address physical charging coordination without simultaneously optimizing the economic value of flexibility, and do not consider generic hierarchical, multi-time-scale control architectures that would enable coordinated participation in day-ahead, intraday, and balancing markets ([57], [58], [59], [60]). Hierarchical and multi-level control strategies are widely used in industrial control processes or energy systems ([63], [64], [65], [66], [67], [68], [69]). However, most existing approaches employ time-scale separation for computational or organizational reasons, without explicitly aligning control layers with heterogeneous market products and their contractual requirements. As a result, the interaction between mid-term energy market commitments (e.g., balancing and spot market) and short-term load services is often handled heuristically. Specifically, hierarchically organized multi-time-scale load management in coherence with market participation has not been addressed systematically in the literature.

2.2 Preliminaries

In this section, the theoretical bases relevant to this thesis are presented. Key concepts, methodologies, and frameworks that underpin the proposed energy management strategies are discussed to provide a foundation for the subsequent developments.

2.2.1 Fundamentals of Demand Response

As mentioned before, DR refers to the intentional modification of electricity consumption patterns by end-users in response to external signals, such as RTP, TOU rates, or requests for load reduction, to support grid stability and economic efficiency [70], [71]. It plays a critical role in enhancing the reliability of power systems, facilitating the integration of variable renewable energy sources, and reducing overall operational costs [72]. DR programs are typically classified into price-based programs, where users respond to fluctuating electricity prices, and incentive-based programs, where users receive financial rewards for providing predefined load reductions or flexibility services [73]. In other words, the key distinction between programs in these two categories is that incentive-based programs offer customers payments for delivering a specified load reduction within a set timeframe, whereas in price-based DR programs, consumers voluntarily reduce their load in response to economic signals. By actively managing consumption, DR reduces reliance on

costly and carbon-intensive peaking power plants, creating new opportunities for end-users to engage in energy markets and generate additional revenue streams.

Incentive-based DR

- Direct load control (DLC): The goal of DLC programs is to involve many small consumers, such as residential users. These programs enable utility control centers (also referred to as utilities) to directly control specific appliances at the consumer's premises, with common examples including air conditioners (ACs), lighting, water heaters, and pool pumps. These programs usually specify the number and duration of interruptions to ensure that end-user comfort is not significantly affected. Participants are compensated through discounts or benefits on their electricity bills and may also receive additional payments upon request. Managed directly by the utilities, these programs typically do not provide prior notice to the end-users before an interruption. DLC events can be triggered by either economic conditions or concerns about reliability [74].

- Interruptible/curtailable load (ICL): These load programs target medium and large consumers. Participants receive incentives to shut down specific loads or reduce their energy usage in response to signals from the utility. As with DLC programs, contracts typically define the maximum number and duration of such requests. These programs are mandatory, meaning customers may incur penalties if they fail to respond to a DR event. While utilities may request load reductions during reliability events, curtailed loads can also be traded in the market [75].

- Demand side bidding, and ancillary services: Demand-side bidding enables consumers to actively participate in the electricity market by submitting offers to reduce their load. Large customers typically participate directly, often utilizing advanced load management tools and strategies, while smaller consumers usually participate indirectly through third-party aggregators or load-serving entities (LSEs) [76]. The demand side may also participate in ancillary service markets, providing various system services on different time scales (regulation, spinning reserve, etc.).

Time/Price-based DR

- Time-of-use (TOU) tariffs: Electricity end-users billed at flat rates are typically unaware of fluctuations in electricity costs. Flat pricing reflects the average cost of electricity supply and can remain unchanged for several years. TOU pricing, by contrast, aims to more accurately represent the changing costs of electricity across different times of the day or seasons [75]. TOU pricing is a stepped-rate structure designed to reflect average market conditions based on the time of day when electricity is consumed, without accounting for daily volatility in supply costs. A typical TOU scheme includes a peak rate, an off-peak rate, and possibly a shoulder-peak rate, with periods determined by the utilities.

- Critical peak pricing (CPP): TOU tariffs reflect the long-term electricity supply costs associated with consumption during specific periods of the day. To account for short-term costs during critical periods for the power system, CPP

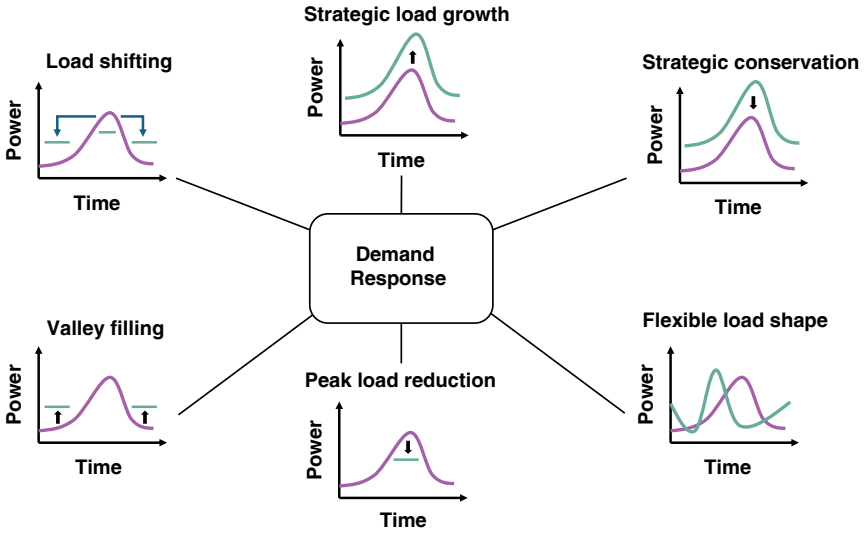


FIGURE 2.1: Different types of DR in mid-term.

can be introduced. CPP adds a time-independent high rate over existing TOU or flat rates, triggered by system conditions such as reserve shortages or extreme weather events that cause unexpected demand fluctuations. Contracts typically define the maximum number of critical days per year and the number of periods when the CPP rate may apply. Utilities usually announce a CPP event on short notice, ranging from several minutes to a few hours before the higher rate is activated. Two variants of CPP exist: extreme day pricing (EDP) and extreme day CPP. EDP imposes higher electricity prices for an entire 24-hour "extreme day," while extreme day CPP applies peak and off-peak rates, similar to standard CPP programs, but only during extreme days, with a flat rate used on all other days [70].

- Real-time pricing (RTP): RTP is a pricing model where the energy price is updated on short notice, usually hourly. With RTP, customers are directly exposed to fluctuations in the wholesale power market costs or changes in locational or zonal marginal prices. Currently, two notable RTP programs engage residential end-users in the U.S.: one by PJM and the other by the Midcontinent ISO (MISO). Both programs provide day-ahead market prices one day before actual power delivery, but their pricing methods differ. In the first program, end-users are charged based on real-time prices averaged over five-minute intervals within each hour, with the final prices determined at the end of the hour. In the second program, consumers are charged directly according to the day-ahead prices [13].

Industrial consumers account for a significant portion of the DR potential due to the scale and flexibility of their operations [77]. In manufacturing environments such as beverage production plants, DR is realized by rescheduling energy-intensive processes, reducing non-essential load consumption, or shifting production to off-peak periods, all while ensuring product quality and production targets are maintained [78]. Fig. 2.1 demonstrates different types of DR in the mid-term during a day. However, exploiting flexibility in industrial processes requires a careful balance between operational constraints, process interdependencies, and energy market participation

[79]. The complexity of industrial systems often necessitates using advanced automatic control strategies, such as MPC or optimization-based scheduling, to manage the trade-offs between production efficiency and flexibility revenue systematically.

2.2.2 Model Predictive Control

MPC is an advanced control strategy that relies on predicting future system behavior to compute control actions. Unlike conventional feedback controllers that respond reactively, MPC operates by solving a finite-horizon optimization problem at each sampling instant. This problem predicts the system's future behavior using a dynamic model, computes a sequence of control actions that minimize a predefined cost function, and satisfies system constraints such as input, state, and output limits. Typically, the first control input of the optimized sequence is applied to the system, after which the optimization is repeated at the next time step using updated measurements—a strategy known as receding horizon control. However, this is not the only option, and the designer can apply more than one computed optimal control input. This predictive capability allows MPC to anticipate future disturbances and correct trajectories proactively, making it exceptionally suitable for systems with complex dynamics, multivariable interactions, and tight constraints. The development of MPC dates back to the late 1970s in the process industries under the name dynamic matrix control (DMC), and it has since evolved into a mature control methodology with strong theoretical guarantees under certain conditions [80], [81].

The essence of the MPC framework lies in formulating and solving a constrained optimization problem that dynamically guides the evolution of control inputs and system states within a predefined horizon. The general formulation involves minimizing a cost function, often quadratic, over a prediction horizon N_p, subject to system dynamics and operational constraints. Formally, this can be expressed as:

$$\begin{aligned} \min_{\mathbf{u}_k} \quad & \sum_{k=0}^{N_p-1} \ell(\mathbf{x}_k, \mathbf{u}_k) + \ell_f(\mathbf{x}_{N_p}) \\ \text{subject to:} \quad & \mathbf{x}_{k+1} = f(\mathbf{x}_k, \mathbf{u}_k), \quad k = 0, \cdots, N_p - 1 \\ & g(x_k, u_k) \geq 0, \\ & \mathbf{x}_k \in \mathcal{X}, \quad \mathbf{u}_k \in \mathcal{U}, \\ & \mathbf{x}_{N_p} \in \mathcal{X}_f \end{aligned} \tag{2.1}$$

Here, ℓ represents the stage cost function, typically penalizing deviations from reference values, while ℓ_f denotes a terminal cost to enhance stability. Also, the functions f and g are nonlinear, and N_p is the prediction horizon. The system dynamics are given by either linear or nonlinear equations, depending on the application, and the constraints $\mathcal{X}$ and $\mathcal{U}$ define permissible sets for states ($\mathbf{x}_k$) and control inputs ($\mathbf{u}_k$). One of the most powerful aspects of MPC is its inherent capacity to incorporate hard constraints directly into the control design. This feature is particularly valuable in systems with actuator limitations or safety-critical boundaries. Advanced variants such as robust MPC and stochastic MPC further extend the formulation to handle uncertainties and probabilistic guarantees, offering flexibility in practical deployment. Figure 2.2 shows how MPC performs in time for tracking the setpoint.

Due to its flexibility and optimization-based nature, MPC has found widespread application in areas requiring sophisticated control strategies, such as chemical processes, robotics, transportation, and increasingly, energy systems. In modern energy

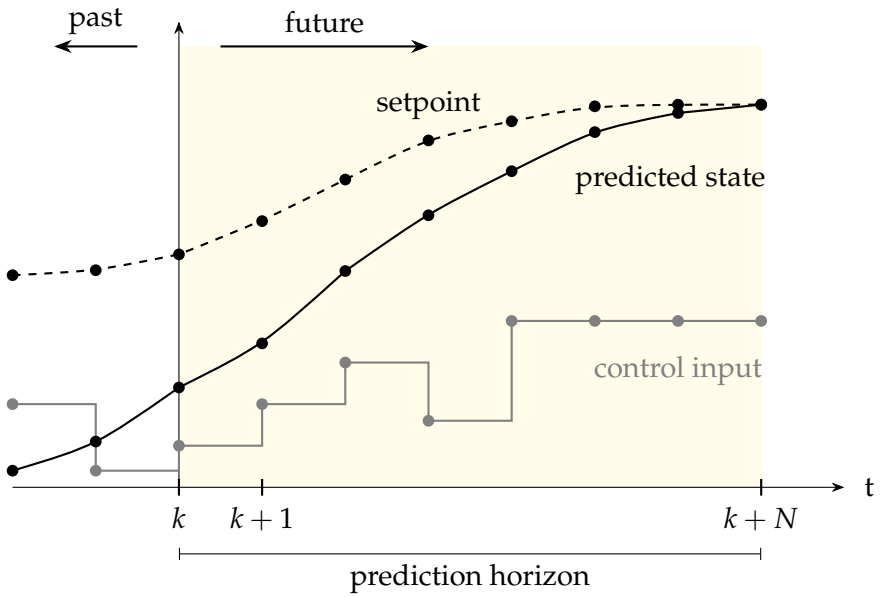


FIGURE 2.2: MPC Scheme.

systems, particularly those integrating renewable energy sources and distributed energy resources (DERs), MPC is well-suited to handle challenges arising from intermittency, forecast uncertainty, and the need for real-time operational efficiency. For instance, MPC can optimize the charging and discharging cycles of battery energy storage systems, schedule loads and generation in microgrids, and manage the grid interaction of EVs in bidirectional charging hubs. The ability of MPC to coordinate multiple energy assets while respecting constraints such as grid limits, emission targets, or market prices makes it a strong candidate for next-generation energy management systems. Moreover, extensions of MPC into economic and hierarchical frameworks [61], [82] allow for the simultaneous consideration of technical performance and economic objectives, particularly relevant in systems participating in DR programs or energy trading markets.

Despite the broad applicability of standard MPC, specific systems exhibit behaviors that purely continuous models cannot adequately capture. In many practical control problems, especially in modern energy systems, the dynamics include both continuous variables (e.g., power levels, temperatures) and discrete decisions (e.g., switching a device on or off, selecting an operating mode). These systems are commonly referred to as "hybrid systems", and addressing their control challenges has given rise to an essential extension of MPC known as hybrid MPC [83]. Hybrid systems are characterized by the interplay between continuous dynamics governed by differential or difference equations and discrete logic transitions. Such systems frequently arise in applications involving mode switching, logical constraints, or scheduling decisions. In the context of energy systems, hybrid behavior can be observed when coordinating energy assets with binary operational modes (e.g., turning a generator on or off, enabling or disabling EV charging, or activating backup storage), as well as in industrial processes with production stages or operating modes that require logical coordination [84]. Hybrid MPC extends the conventional MPC formulation by incorporating discrete variables directly into the optimization problem. This is typically achieved by formulating the problem as a mixed-integer optimization, where some decision variables are constrained to take on integer values—often binary—to model the on/off or discrete-mode behavior. A typical hybrid

MPC formulation can be expressed as:

$$\begin{aligned} \min_{\mathbf{u}_k, \mathbf{d}_k} \quad & \sum_{k=0}^{N_p-1} \ell(\mathbf{x}_k, \mathbf{u}_k, \mathbf{d}_k) + \ell_f(\mathbf{x}_{N_p}) \\ \text{subject to:} \quad & \mathbf{x}_{k+1} = f(\mathbf{x}_k, \mathbf{u}_k, \mathbf{d}_k), \qquad k = 0, \cdots, N_p - 1 \\ & g(\mathbf{x}_k, \mathbf{u}_k) \geq 0, \\ & \mathbf{x}_k \in \mathcal{X}, \quad \mathbf{u}_k \in \mathcal{U}, \quad \mathbf{d}_k \in \mathbb{Z}^m \\ & h(\mathbf{x}_k, \mathbf{u}_k, \mathbf{d}_k) \geq 0 \end{aligned} \tag{2.2}$$

Here, $\mathbf{d}_k$ denotes the vector of discrete decision variables at time k, which belongs to the set of integer numbers $\mathbb{Z}^m$, and the optimization problem may include logical constraints ($h(\mathbf{x}_k, \mathbf{u}_k, \mathbf{d}_k) \geq 0$), such as mutually exclusive modes or conditional activation rules. The combination of continuous and discrete control actions allows hybrid MPC to model and optimize hybrid behaviors while maintaining predictive control capabilities. However, the inclusion of integer variables significantly increases the computational complexity of the control problem, often making real-time implementation more challenging. Efficient mixed-integer programming (MIP) solvers and problem-specific relaxations are usually required to make online computation feasible [85]. Despite these challenges, the enhanced modeling fidelity offered by hybrid MPC enables it to deliver superior performance in systems where discrete decisions are essential.

The mixed logical dynamical (MLD) framework presents a generic problem formulation for hybrid MPC [86]. This hybrid dynamic model of the process integrates differential/difference equations with logical rules and constraints into a unified set of mixed-integer linear equations. This representation transforms the hybrid MPC problem into an MIP problem, which can be efficiently solved using modern solvers. The MLD systems provide a unified modeling framework for hybrid systems that combine continuous dynamics with logical decisions, switching behaviors, and constraints. The MLD formalism presents systems dynamics using real, binary, and integer variables within a mixed-integer linear system. This allows for the systematic representation of systems with on/off actuators, logical rules, and state-dependent constraints. To design MPC for MLD systems, the prediction model is given in the form of [83]:

$$\mathbf{x}_{k+1} = A\mathbf{x}_k + B_1\mathbf{u}_k + B_2\Delta_k + B_3\mathbf{z}_k \tag{2.3a}$$

$$\mathbf{y}_k = C\mathbf{x}_k + D_1\mathbf{u}_k + D_2\Delta_k + D_3\mathbf{z}_k \tag{2.3b}$$

where $\mathbf{x}_k$ is the state vector, $\mathbf{u}_k$ is the continuous input, Δ_k represents binary variables (e.g., on/off modes), and $\mathbf{z}_k$ are auxiliary continuous variables used to linearize logical conditions. The MPC problem is then formulated as a finite-horizon optimization problem, typically minimizing a cost function of the form:

$$\min_{\mathbf{u}_k, \Delta_k, \mathbf{z}_k} \quad \sum_{k=0}^{N_p-1} \ell(\mathbf{x}_k, \mathbf{u}_k, \Delta_k, \mathbf{z}_k) + \ell_f(\mathbf{x}_{N_p}) \tag{2.4}$$

subject to the MLD system dynamics, constraints on inputs and states, and logical conditions. The resulting optimization problem is a mixed-integer quadratic program (MIQP) or MILP, depending on the cost function.

While hybrid MPC primarily addresses systems with both discrete and continuous dynamics, additional complexity arises when the model includes nonlinearities, particularly those that include discrete variables. One of the main nonlinearities that often emerges in the problem formulation of hybrid systems is bilinearity. These usually occur in applications involving switching control or mode-dependent actuation, where variables interact multiplicatively. Such bilinearities introduce non-convexity that computationally complicates the optimization. To mitigate this problem, there are relaxation techniques to break down the bilinear constraints into some relaxed linear constraints, which can significantly reduce the complexity. The relaxation techniques in this regard include McCormick envelopes [87], [88], Reformulation-Linearization [89], or piecewise linear approximations [90]. The type of relaxation that we apply in the thesis is mainly based on the relaxation presented in [83] and [91]. Let us assume that a bilinearity in the optimization problem settings is in the form of $\delta \cdot x$, where δ is a binary variable and x is a continuous variable. We define:

$$z := \delta x \tag{2.5}$$

Also, we consider an upper and lower bound on the continuous variable x as follows:

$$\underline{x} \leq x \tag{2.6a}$$
$$x \leq \bar{x} \tag{2.6b}$$

where $\underline{x}$ and $\bar{x}$ are lower and upper bounds. The bilinear term $\delta \cdot x$ can then be represented by including the following linear inequalities in the optimization:

$$z \leq \delta \bar{x} \tag{2.7a}$$
$$z \geq \delta \underline{x} \tag{2.7b}$$
$$z \leq x - (1 - \delta)\underline{x} \tag{2.7c}$$
$$z \geq x - (1 - \delta)\bar{x} \tag{2.7d}$$

which means the new variable z can be used instead of the bilinear term δx if the above linear constraints are added to the problem formulation.

Hierarchical MPC extends the traditional MPC framework by decomposing the control problem into multiple layers, each responsible for different decision-making scopes or time scales. This structure is particularly advantageous when centralized MPC becomes computationally expensive or impractical due to system complexity or organizational constraints. In a typical hierarchical MPC architecture, the higher-level controller operates on a longer prediction horizon, focusing on strategic or economic objectives, and generates references or constraints for the lower-level controllers [82]. These subordinate controllers then solve local MPC problems over shorter horizons to track the references while reacting to fast dynamics and disturbances. Such a decomposition enables coordination between global objectives and local feasibility, improving scalability and responsiveness. In energy systems, hierarchical MPC is particularly effective when integrating distributed energy resources, such as photovoltaic units, batteries, and flexible loads. For example, in a multi-energy management framework, a supervisory MPC might optimize day-ahead energy procurement and load allocation, while lower-level controllers manage real-time operation of individual assets like EV chargers or HVAC units. The hierarchical structure also supports modularity, allowing different components to

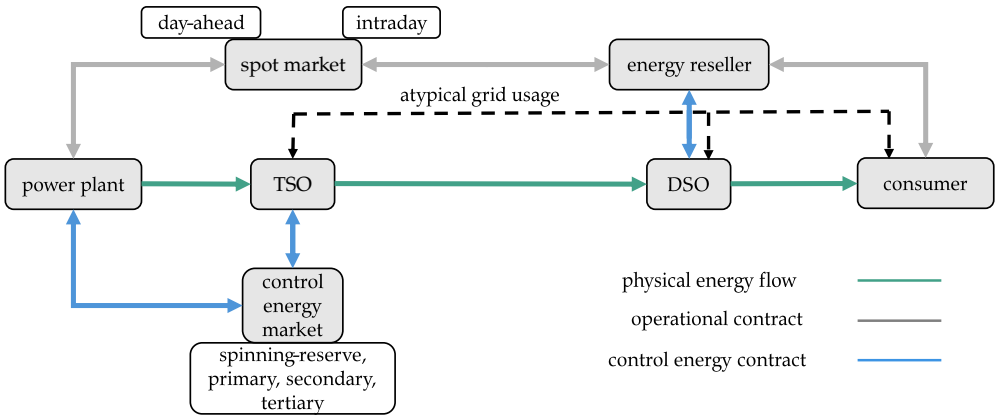


FIGURE 2.3: Current configuration of the energy market along with the physical grid.

be developed and operated independently while maintaining overall system coherence [92], [93].

2.2.3 Energy Market

The energy market is a complex network in which electricity is bought, sold, and balanced through various interconnected layers, including long-term contracts, day-ahead and intraday spot trading, as well as real-time balancing mechanisms [94], [95], [96]. Producers generate electricity, which is traded by energy resellers in the spot market and delivered to consumers through the transmission and distribution networks. To ensure that supply and demand are matched, grid operators rely on control energy markets—such as primary, secondary, and tertiary reserves—and increasingly on flexible, distributed resources aggregated from smaller consumers and producers. We elaborate more on the energy market mechanism in cooperation with the power grid to explain the current situation, followed by the improved structure for the future.

As depicted in Fig. 2.3, the physical energy is transmitted from the power plant (or any other generation units, including renewable generation) to the local distribution networks. The operator responsible for transmission is called the transmission system operator (TSO), and the operator responsible for distributing the energy is called the distribution system operator (DSO) [97]. The transmission of energy is carried out with high voltages (several hundred kilovolts), while the level of voltage for distribution is in the order of several hundred volts to several kilovolts (highly dependent on the area). The DSO distributes energy among endpoint consumers. At the signal level, there are two contracts for operational purposes (gray color) and the control energy market (blue color). Energy is traded through the spot market, which is the cornerstone of electricity trading, enabling real-time and near-term transactions to balance supply and demand at transparent prices. The two key components of the spot market are the day-ahead market and the intraday market. In the day-ahead market, an auction for electricity delivery 24 hours in advance is conducted, where generators and consumers submit their bids. In the intraday market, the trading is continuously conducted for the same day until minutes before delivery [98]. The energy tariffs in this market can suddenly fluctuate due to several factors affecting both generation and consumption, such as weather conditions.

The energy reseller (also called an electricity retailer or supplier) plays a purely commercial role in the power grid ecosystem, acting as an intermediary between wholesale electricity markets and end consumers. It buys electricity from wholesale markets (day-ahead, intraday) and offers the energy to the consumers through various pricing strategies. The pricing strategies include fixed rates, variable rates, dynamic pricing, and green energy plans. In the context of fixed-term pricing contracts, the reseller locks in a price per kWh for a fixed term (e.g., 12–24 months). In fact, the price remains constant regardless of wholesale market volatility. For the variable pricing strategy, the price per kWh fluctuates monthly/quarterly based on wholesale costs. In dynamic pricing, prices change hourly based on real-time wholesale markets (e.g., day-ahead prices), and consumers can observe the price 24 hours in advance. It allows shifting some high-energy-consuming tasks to the off-peak times. Ultimately, green energy plans prioritize the use of electricity generated from renewable sources, including solar, wind, geothermal, and hydropower. These plans aim to reduce reliance on fossil fuels and minimize environmental impact. Some resellers offer 100% renewable electricity options. The energy reseller, above all, uses pricing strategies to offer consumers, buying or selling energy in the spot market by bidding on the demands or possible supply of active consumers. The spot market determines the real-time value of electricity based on bids from suppliers and demand. Trading is performed by holding an auction for electricity delivery 24 hours in advance in the day-ahead market or by conducting a continuous energy transaction for adjustments on the same day in the intraday market.

The control energy market (also called the balancing market or reserve market) is a critical mechanism operated by TSO to ensure real-time stability of the power grid. It procures and dispatches reserves to correct imbalances between electricity supply and demand, maintaining frequency (e.g., 50 Hz in Europe, 60 Hz in the U.S.) and preventing blackouts. Electricity demand and generation must match instantaneously. When imbalances occur (e.g., a power plant fails or demand spikes), the grid's frequency deviates, and the control energy market compensates by activating reserves. The frequency changes due to imbalances are as follows:

- Frequency drops: when the demand is more than the supply
- Frequency rises: when the supply is more than the demand

Figure 2.4 illustrates the imbalances seen by TSO in the French grid over an 8-hour period [99]. As depicted in this figure, on one hand, when the imbalance is negative, it means that generation is less than production, and an "up-regulation" is required to balance the network. On the other hand, when the imbalance is positive, it means the generation exceeds consumption, and a "down-regulation" is mandatory.

In the current control energy market, TSO buys and activates reserves to balance the grid when demand is greater than generation. The reserves are provided by the supply side (generators) to restore the frequency of the grid, as shown in Fig. 2.3. In the event of a frequency deviation, the different layers of reserves are activated sequentially based on their response times: spinning and primary reserves react within seconds, followed by secondary reserves over several minutes, and finally tertiary reserves for sustained imbalances. More precisely, spinning reserves, which are backup electrical generation capacity, react very fast when there is an imbalance between generation and consumption. The reaction time is in the order of 1-10 seconds. The primary reserves include any generation units that can contribute to the frequency containment within 30 seconds. The secondary reserves activate within 30 seconds to 5 minutes, during which TSOs automatically dispatch pre-contracted

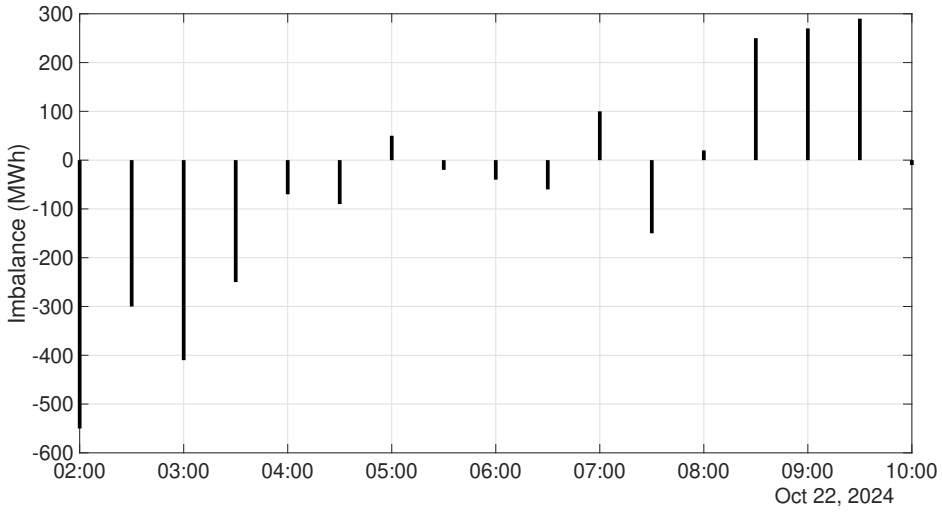


FIGURE 2.4: Power imbalance observed by TSO over a specific period.

reserves, such as gas plants. In 15 minutes to several hours, TSOs manually activate reserves (e.g., standby generators) to fully recover the frequency to the nominal value. For example, a diesel generator is started to cover prolonged shortages. A scheme of the control loops for frequency control of the power grid is depicted in Fig. 2.5.

This diagram illustrates the multi-level frequency control structure in a power system, which ensures grid stability following frequency deviations. The control mechanism starts with primary control, which is automatically and immediately activated to limit frequency deviations within a few seconds of an imbalance. Its task is to stabilize the system frequency by providing an initial power response, but not to restore the frequency to its nominal value. Once activated, it provides a "free reserve" to be taken over by the secondary control layer. The primary control is essential for the system's inertia-based response and buys time for the slower controls to engage. Secondary control then takes over if responsible and acts within tens of seconds to several minutes to restore the frequency to its nominal value. It corrects the residual imbalance and also restores the primary reserve. Once the secondary reserve is provided, if a more sustained or larger correction is required, the tertiary control is activated on a longer-term basis, either to support the secondary control or to replace it. With primary, secondary, and tertiary actions, the average frequency over long intervals can deviate slightly from the nominal value, causing a cumulative time error (drift of "synchronous time" relative to coordinated universal time (UTC)). For example, a persistent $+0.002$ Hz bias causes a clock to run fast by approximately 1.2 seconds per 10 hours. If unchecked, these minor errors (just a few millihertz over hours) accumulate into seconds of error per day. Time control (also called time error correction) is the slow-acting process that eliminates this drift. Its goal is to maintain the long-term mean system frequency exactly at its nominal value, ensuring that synchronous time remains in sync with UTC [100].

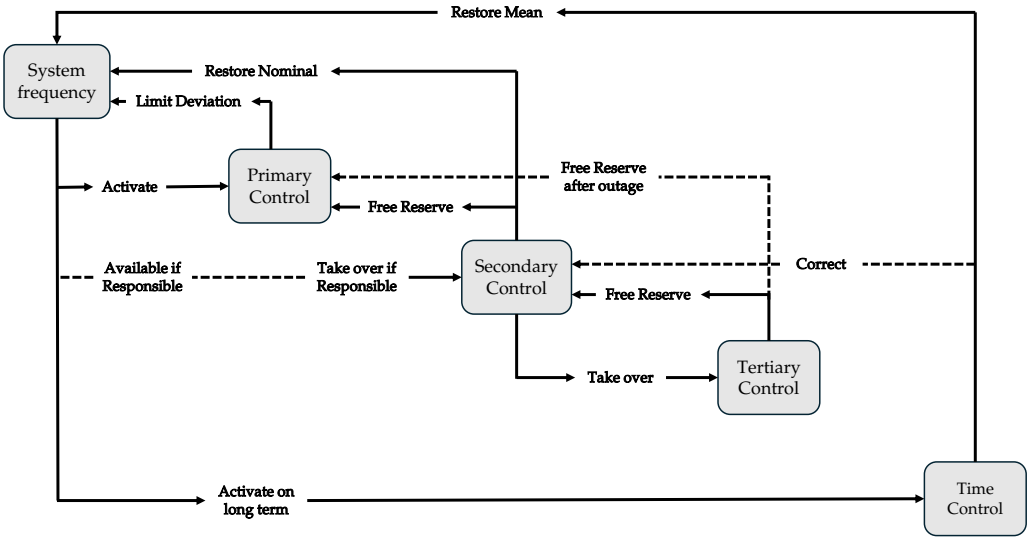


FIGURE 2.5: Hierarchical regulation of the grid's frequency.

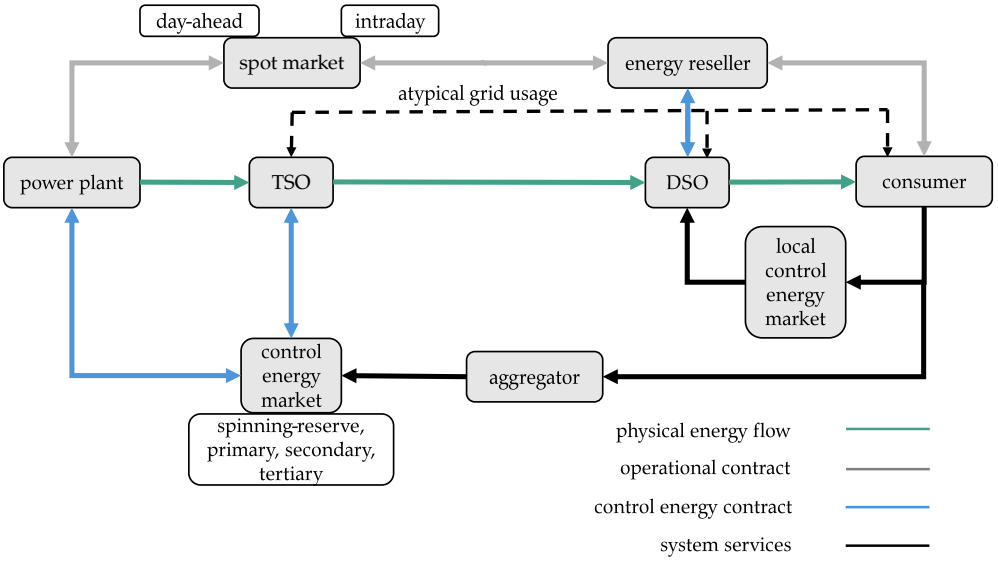


FIGURE 2.6: Future configuration of the energy market along with the physical grid.

As explained above about the current structure of the energy market and also sketched in Fig. 2.3, TSO takes action against frequency deviation through the offered bidding of the supply side (generation units) to overcome frequency drifting arising from power imbalance in the grid. It is crucial to note that current research in the domain of power grids also focuses on exploiting the flexibility provided by the demand side, as described in Chapter 1 and this chapter. To utilize this flexibility, the configuration depicted in Fig. 2.6 provides the opportunity to integrate active consumers into the control energy market. The flexibility aggregator aggregates the total flexibility that consumers provide and bids it in the control energy market, allowing each consumer to respond to grid imbalances and peak-time stress on time. This is highlighted in black in the figure. The loads can deliver the primary, secondary, and tertiary services if they receive an incentivizing auction from the TSO. Also,

the consumers can participate in other grid services, such as load shifting or load shedding, using incentive-based or time/price-based DR mechanisms to reduce the stress on the grid during peak times and optimize their energy costs. Therefore, the flexibility aggregator provides these system services to the DSO through a contract between itself and the energy reseller, as well as between the energy reseller and the consumer.

Atypical grid usage (shown in dashed lines in Fig. 2.3 and 2.6) refers to non-standard connections or operational scenarios that deviate from traditional electricity distribution practices. These exceptions often arise due to unique consumer needs, regulatory frameworks, or technological advancements. In this usage, large consumers bypass the DSO to connect directly to the TSO. Therefore, the TSO can directly send commands to consumers (typically those with high consumption) to enhance grid stability and resilience, as well as facilitate decentralized generation and peer-to-peer trading. This is in accordance with the DLC method introduced in 2.2.1, where some loads can be controlled directly by the grid operator.

2.3 Conclusion

This chapter reviewed the theoretical provisions and relevant literature on DR, MPC, and energy market integration. While existing approaches offer valuable insights, key gaps remain in multi-level energy management across different large consumers. These findings highlight the need for a generic control framework, which is developed in the subsequent chapters.

Chapter 3

Demand Response Control of Beverage Production Plants

3.1 Introduction to Methodology

This chapter presents the methodological framework developed to enable beverage production plants to participate actively in DR programs and energy markets. Referring to Fig. 1.1, it is vital to mention that the aim in this chapter is to design the C_3 controller for beverage production plants. These production lines, which consist of energy-intensive machines operating in sequential stages, offer flexibility that can be utilized to shift or curtail energy consumption in response to external grid signals. However, utilizing this flexibility requires a control strategy that balances energy market incentives with production constraints, such as scheduled production quantity. The proposed methodology employs an MPC framework that integrates real-time operational decisions in response to market-based signals. In the control design procedure, each machine's working mode is modeled as either binary or continuous, and a flexibility contract that defines acceptable production losses in exchange for power reductions is incorporated. This approach is particularly suitable for beverage plants where certain stages have limited but exploitable temporal flexibility. The methodology prioritizes computational tractability, realistic modeling, and market alignment, ensuring it is both implementable and effective. It is important to mention that in the following, we use "components" and "units" interchangeably, both implying the set of all machines and belts in a production line.

3.2 System Modeling for Beverage Production Plants

A beverage production line consists of several processing machines and conveyor belts. In this subsection, we provide a detailed description of modeling the components of a line with different dynamic fidelity. Therefore, two primary approaches for modeling a beverage production line have been proposed. First, a distributed dynamic parametric model (DDPM) is developed to capture precisely the speed dynamics of the line's components, including the effects of permanent magnet synchronous motors (PMSMs) used in each unit. Next, a generic modeling framework is introduced to represent production line components abstractly, called the network delayed model (NDM). However, the focus is on DDPM for developing the proposed controller, and NDM remains only as a generic approach for modeling the line. Both models serve as the foundation for constructing a beverage production line, which can be extended to multiple lines to demonstrate scalability.

A beverage production line includes both binary and continuous machines. We assume that binary machines operate exclusively in two modes: normal mode and

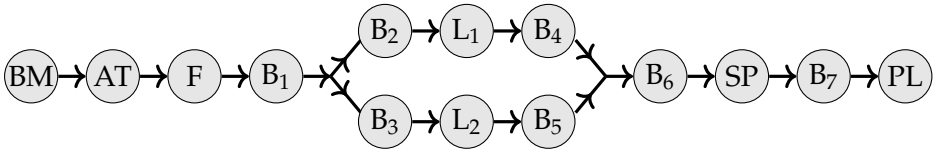
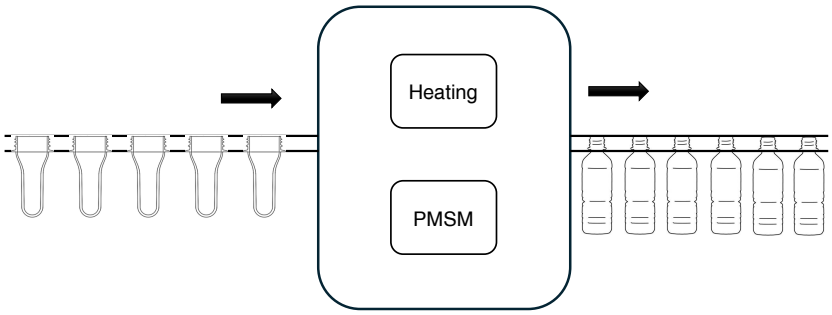


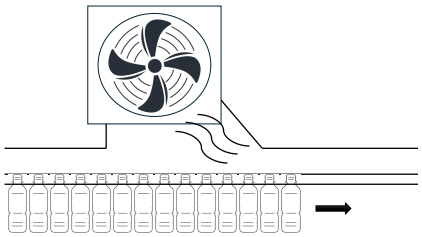
FIGURE 3.1: A beverage production line.

standby mode. In standby mode, the binary machine stops processing while other internal processes, such as heating, continue. In normal mode, the machine works in a full operational state and consumes its nominal power. Continuous machines provide full-range speed control continuously, and the power consumption changes accordingly. We consider the first two processing machines in a beverage production line to be the blow molding machine and the air transport system. These machines generally have several discrete modes of operation; however, we assume they have only two modes of operation and consider them as binary machines.

A blow molding machine is a special manufacturing device utilized to produce hollow plastic parts, particularly bottles and containers. It produces plastic bottles by first melting plastic pellets and forming them into a preform. This softened plastic is then clamped into a mold, where high-pressure air inflates it to take the shape of the mold cavity. After cooling, the solidified bottle is ejected, completing the cycle. The process is highly efficient, enabling mass production of uniform bottles with precise shapes and thicknesses. In normal mode, the blow molding machine processes formable preforms, subjecting them to intense heating inside the machine while operating at its nominal rotation speed, which is higher than in standby mode. In standby mode, the machine rotates at a reduced speed while maintaining a specific temperature to preserve the formability of the preforms. It receives only two reference signals for rotation, which indirectly influence the machine's internal heating adaptation. For simplicity, we do not explicitly model the internal heating control loop; instead, it is accounted for through the machine's power consumption. The machine's outflow rate depends on its operating mode. In standby mode, no preforms are processed, resulting in zero outflow. In normal mode, the machine produces empty bottles at a specific rate corresponding to the nominal speed of its PMSMs. The following machine is an air transport system that transfers empty bottles to the filler. The bottles move through this system, propelled by airflow generated by large fans driven by PMSMs. Each fan directs its circulated air through a pipe to a segment where the empty bottles are hung. The fans' speed indirectly controls the movement speed of the bottles in the air transport. Adjustable air pressure ensures controlled movement, reducing friction and minimizing bottle damage. This method enables high-speed, continuous transport, making it ideal for large-scale beverage production lines. Similar to the blow molding machine, the air transport operates in two distinct modes: normal mode and standby mode. In normal mode, the bottles are pushed forward at a nominal speed, whereas in standby mode, they come to a complete stop. The movement or stoppage of bottles depends on the filler's capacity to accept incoming bottles, ensuring smooth operation and preventing jams along the entire air transport system. Fig. 3.1 illustrates a beverage production line that includes several processing units and conveyor belts. In this figure, "BM" represents a blow molding machine, "AT" denotes air transport, "F" indicates filler, "B" represents a conveyor belt, "L" stands for labeling machine, "SP" denotes a shrink packer, and "PL" represents a palletizer. Also, Fig. 3.2 shows an image of both the blow molding machine and air transport in a beverage production line.



(A) Blow molding machine



(B) Air transport

FIGURE 3.2: Binary machines of the line.

A filling machine or filler is the unit of the production line that receives the empty bottles and fills them with the target liquid. It also has a capper that puts a cap on the bottles when they are filled enough. It usually fills the empty bottles while rotating them by PMSMs integrated inside the machine. Fig. 3.3 shows a typical filler in a beverage production line and how it processes empty bottles. Moreover, a conveyor belt is the unit that moves the bottles from one machine to the next. It usually has a long length, and its linear movement is the result of converting the rotation of PMSMs into linear movement. A labelling machine is a processing unit that receives the bottles from the filler and labels them with the required information. This is accomplished by the rotating PMSMs inside the units. Also, a shrink packer is a unit that sorts and wraps the bottles coming from the labeling machine so that small packs of bottles (e.g, 6 or 12 bottles in a pack) are prepared for processing by the palletizer. The palletizer rounds up the small packs of bottles and wraps a specific number of packs to fit them onto one pallet. It then accumulates these pallets to be prepared for transferring to the final customers. The processing speed of the conveyor belts, labeling machines, shrink packer, and palletizer can be controlled continuously.

3.2.1 Dynamic Distributed Parameter Model

Considering the type of machines (binary and continuous), this subsection elaborates on a dynamic model of the components of a beverage production line.

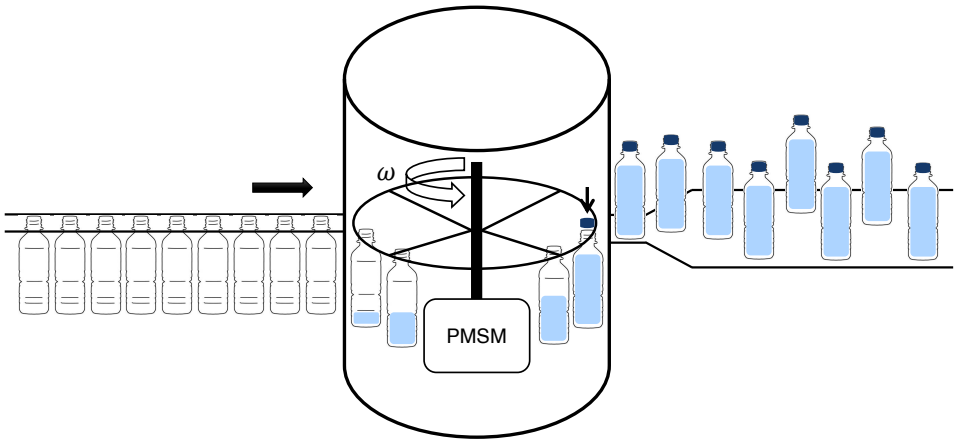


FIGURE 3.3: The schematic picture of a filling machine, involving a feeding and delivery conveyor belt.

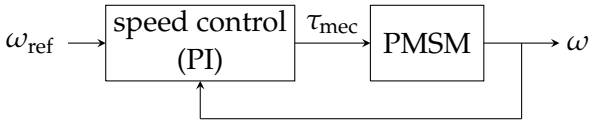


FIGURE 3.4: Internal control loop of the PMSM.

Continuous Machines and Conveyor Belts

These units process the bottles, exploiting the rotation of PMSMs that apply the rotational torque whose speeds are determined by the following dynamical equation in discrete time t:

$$\omega(t+1) = \omega(t) + \frac{T_s}{J}(-d\omega(t) - M + \tau_{\text{mec}}(t)), \quad t \in \mathbb{N} \tag{3.1}$$

where J, d, and M are the inertia, damping, and friction terms, respectively. The speed is shown by ω, τ_{mec} is the mechanical torque, and T_s is the sampling time. The torque τ_{mec} is generated by a PI controller designed to regulate the PMSM speed, ensuring it matches the reference speed ω_{ref}:

$$\tau_{\text{mec}}(t) = k_1(\omega_{\text{ref}}(t) - \omega(t)) + k_2 T_s \sum_{i=1}^{t}(\omega_{\text{ref}}(i) - \omega(i)) \tag{3.2}$$

k_1 and k_2 are proportional and integral coefficients of the PI controller, respectively (Fig. 3.4). We disregard electrical power losses in the machines and define the power consumption of a component using the following general form:

$$p(t) = d\omega^2(t) + M\omega(t) + \alpha n(t) + p_0 \tag{3.3}$$

where α, n, and p_0 are the power-bottle coefficient, the number of bottles in the component, and the power consumption at zero speed, respectively. In the case of the processing machines, such as the filler, α is zero; however, it is not zero for conveyor belts. Additionally, p_0 is 0 for conveyor belts, whereas it could be non-zero for other components. The control-oriented model of the components can then be

written as:

$$\begin{aligned}\mathbf{x}(t+1) &= \mathbf{x}(t) + T_s f(\mathbf{x}(t), \mathbf{u}(t)) \\ &= \mathbf{x}(t) + T_s \begin{bmatrix} \omega(t) \\ \frac{1}{J}(-d\omega(t) - M + \tau_{\text{mec}}(t)) \\ \omega_{\text{ref}}(t) \\ p(t) \\ i(t) - b(t) \end{bmatrix}\end{aligned} \tag{3.4}$$

where the state vector is $\mathbf{x} = [\phi \quad \omega \quad \phi_{ref} \quad e \quad n]^T$ and $\mathbf{u} = [i \quad \omega_{\text{ref}}]^T$ for the first component, such as blow molding machine, while $\mathbf{u} = \omega_{\text{ref}}$ for the rest, is the control input. The states ϕ and ϕ_{ref} are integrals of ω and ω_{ref}, respectively, and e is the consumed energy by the component. i and b represent the inflow of bottles into the component and the outflow of bottles from the component, respectively. The bottles move from the start to the end of the belt with a significant delay. In contrast, delays in units such as fillers and labeling machines are minimal, allowing us to assume that the inflow and outflow rates of these processing machines are equal:

$$i(t) = b(t) = \gamma\omega \tag{3.5}$$

γ is the bottle density inside a processing machine.

The primary challenge in modeling the conveyor belt lies in accurately characterizing the motion of the bottles. We address this using the transport equation (TE), which is inherently a partial differential equation (PDE) of the form:

$$\frac{\partial\beta(\tau,z)}{\partial t} + v\frac{\partial\beta(\tau,z)}{\partial z} = 0, \quad 0 \le z \le \bar{z}, \quad 0 \le \tau \tag{3.6}$$

where the initial and the boundary conditions are:

$$\begin{cases} \beta(0,z) = f_0(z) \\ \beta(\tau,0) = g(\tau) \end{cases} \tag{3.7}$$

where $f_0(z)$ is the initial distribution and $g(\tau)$ forms the boundary condition. In Eq. (3.6), $\beta(\tau,z)$ denotes the bottle density function on the belt at time τ and position z, assuming the beginning of the belt as the zero coordinate. In fact, to model bottle movement, we introduce the concept of bottle density, where the number of bottles per unit area on the conveyor belt varies with both time and position. This density function allows a straightforward calculation of the total bottle count. For simplicity, the width of the conveyor belt is normalized to unity. The "bottle density function" or "β-function" is defined as a two-variable function that specifies the number of bottles per unit length at each point on the conveyor belt at a given time (Fig. 3.5). The solution of Eq. (3.6) at a constant speed reads:

$$\beta(\tau,z) = \begin{cases} f_0(z - v\tau) & \tau \le \frac{z}{v} \\ g(\tau - \frac{z}{v}) & \tau > \frac{z}{v} \end{cases} \tag{3.8}$$

Since we aim to develop control algorithms for machine speed regulation, the assumption of a constant velocity v becomes invalid. Consequently, the solution must be adapted to accommodate variable speeds. Therefore, we introduce a new variable

θ as follows:

$$\theta = F(\tau) = \int_0^{\tau} v(s)ds, \tag{3.9}$$

yielding:

$$\frac{\partial \tilde{\beta}(\theta, z)}{\partial \theta} + \frac{\partial \tilde{\beta}(\theta, z)}{\partial z} = 0, \quad 0 \leq z \leq \bar{z}, \quad 0 \leq \theta \tag{3.10}$$

where:

$$\tilde{\beta} = \beta \circ F^{-1}, \; \tilde{g} = g \circ F^{-1} \tag{3.11}$$

The symbol "$\circ$" denotes the composition of two functions. In the new virtual time domain θ, the PDE is similar to the PDE in Eq. (3.6) with constant speed $v = 1$. Hence, the solution of this new PDE is ([15], [101]):

$$\tilde{\beta}(\theta, z) = \begin{cases} f_0(z - \theta) & \text{if} \quad \theta \leq z \\ \tilde{g}(\theta - z) & \text{if} \quad z \leq \theta \end{cases} \tag{3.12}$$

Now, the rate of change of the bottles on the belt in θ domain can be described as:

$$\frac{dn_b}{d\theta} = i_b(\theta) - o_b(\theta) = \tilde{\beta}(\theta, 0) - \tilde{\beta}(\theta, \bar{z}) \tag{3.13}$$

where i_b and o_b are the inflow and outflow of the belt. Moreover, the rate of change in the number of bottles on the belt is:

$$\frac{dn_b}{d\tau} = v(\tau)\frac{dn_b}{d\theta} = v(\tau)\left(\tilde{\beta}(\theta, 0) - \tilde{\beta}(\theta, \bar{z})\right) \tag{3.14}$$

Since $F(0) = 0$:

$$\tilde{\beta}(0, z) = \beta(0, z) = f_0(z) \tag{3.15}$$

For back-transformation in the time domain, we pose:

$$\tilde{g}(\theta - z) = g(\tau - \tau_z) \tag{3.16}$$

where τ_z denotes a backward shift in t domain. Using $F^{-1}(\theta - z) = \tau - \tau_z$ results in:

$$\tau_z = F^{-1}(\theta) - F^{-1}(\theta - z) \tag{3.17}$$

Eq. (3.15), (3.16), and (3.17) provide the unique solution to the transport equation in τ domain ($\beta(\tau, z)$) with variable speed v, assuming that we have the solution in θ domain ($\tilde{\beta}(\theta, z)$). Fig. 3.6 depicts the transformation between τ and θ domains. This analysis is valid only if $v(\tau) > 0, \quad \forall \tau > 0$. As a conclusive point, it is vital to

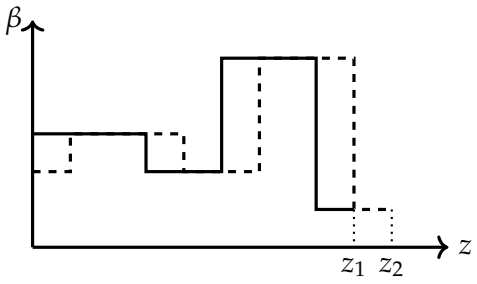


FIGURE 3.5: Bottle density function (β function) at time t (continuous) and $t + \Delta t$ (dashed).

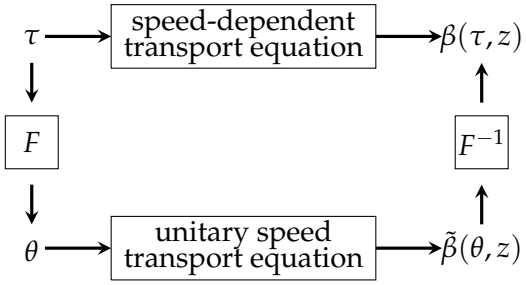


FIGURE 3.6: Time forward and backward transformation.

mention that the assumption in Eq. (3.5) does not hold for a conveyor belt, meaning the inflow rate is not equal to the outflow rate due to large delays for this unit, and its outflow rate is defined as follows:

$$b(\tau) = \beta(\tau, \bar{z}) \cdot v(\tau) \tag{3.18}$$

Binary Machines

These machines have a similar power consumption to continuous machines as described in Eq. (3.3) with $\alpha = 0$. However, the speed of these components can only take two values: nominal speed and standby speed. Therefore, we add a binary variable (δ_j) to model the speed variation of binary machines as follows:

$$\omega_{\text{ref},j} = \delta_j \cdot \omega_{\text{nom},j} + (1 - \delta_j) \cdot \omega_{\text{standby},j} \quad j \in \{\text{BM}, \text{AT}\} \tag{3.19}$$

where $\omega_{\text{nom},j}$ and $\omega_{\text{standby},j}$ are the nominal rotation speed and the speed at standby mode, respectively. The notations "BM" and "AT" are used for blow molding machine and air transport, respectively. In the following, we also assume that these two units work and stop together, enforcing $\delta_{\text{BM}} = \delta_{\text{AT}}$. The standby speed $\omega_{\text{standby},j}$ could generally get non-zero values, but in this thesis, for the sake of simplicity, we assume it is zero, and all power consumption of the binary machines is included in p_0. Also, Eq. (3.2) still holds for these machines.

3.2.2 Network Delay Model

A low-fidelity generic model where the sojourn time of bottles in the components varies over time is the network delay model (NDM), where each component is represented as a block with i) an inflow bottle rate (i), ii) an outflow bottle rate (b), and iii) a processing speed (ω). As illustrated in Fig. 3.7, incoming bottles experience a delay before exiting the component, which depends on the processing speed.

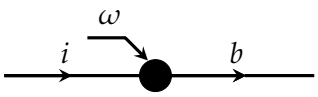


FIGURE 3.7: A component of the line in NDM.

Both processing machines and conveyor belts follow the same structural model, but with different parameter values. The dynamics of these machines are captured

by the following delayed difference equation (DDE):

$$\begin{bmatrix} n(t+1) \\ \phi(t+1) \end{bmatrix} = \begin{bmatrix} n(t) \\ \phi(t) \end{bmatrix} + T_s \cdot \begin{bmatrix} \rho(t) - b(t) \\ \omega(t) \end{bmatrix} \tag{3.20}$$

where:

$$b(t) = \omega(t)\gamma(t - \tau(t)), \tag{3.21a}$$
$$\rho(t) = \min(i(t), \omega(t)\overline{\gamma}), \tag{3.21b}$$
$$\gamma(t) = \frac{\rho(t)}{\omega(t)}, \tag{3.21c}$$
$$\phi(t - \tau(t)) = \phi(t) - \ell, \tag{3.21d}$$
$$0 \leq n(t) \leq c \tag{3.21e}$$
$$p(t) = h(\omega(t), n(t)) \tag{3.21f}$$

ρ denotes the bottle uptake rate and h is a nonlinear function of ω and n. The maximum value of bottle density is $\overline{\gamma}$, ℓ is the length a bottle moves from the beginning to the end of the production unit, c is the bottle capacity of the component, and τ is the time-varying delay of a unit to process a bottle. This model can be also written in a state-space form assuming $\mathbf{x}(t) = [n(t) \;\; \phi(t)]^\top$ as the state-vector, $\omega(t)$ as the control input, $i(t)$ as an internal input, and $b(t)$ as the output, with a given initial condition $\mathbf{x}(0) = [n_0 \;\; \phi_0]^\top$.

The proposed DDPM and NDM offer two levels of fidelity to address the behavior of machines, emphasizing the relationship between processing speed and power consumption of the components. The primary difference between the models lies in considering speed dynamics and bottle distribution on the conveyor belts. The DDPM provides a nonlinear second-order model of the machines, including the internal PI control loop for tracking the reference speed ω_{ref}. Additionally, it contains information about the bottle density function (β function) of conveyor belts for consideration in control purposes. On the other hand, the NDM model does not contain information about the distribution of bottles on conveyor belts. Moreover, it neglects the PI internal loop for following ω_{ref} and the time it takes for ω to match exactly ω_{ref}.

Although NDM is a model at an abstract level that is highly suitable for the analysis of interconnected production lines and heterogeneous processing units, we focus on DDPM in this thesis and retain NDM as a conceptual and analytical modeling framework that can be used to analyze various industrial production lines [102], [103]. It is also important to mention that since NDM is a generic model for any processing unit in a production line (not only beverage plants), a complete model of the unit that integrates the speed dynamics can be easily driven by adding an ordinary differential equation (ODE) to Eq. (3.20) and Eq. (3.21) as follows:

$$n(t+1) = n(t) + T_s\,(\rho(t) - b(t)), \tag{3.22a}$$
$$\phi(t+1) = \phi(t) + T_s\,\omega(t), \tag{3.22b}$$
$$\omega(t+1) = \omega(t) + T_s\,g(\omega(t), n(t), \omega_{\text{ref}}(t)), \tag{3.22c}$$
$$\rho(t) = \min(i(t), \omega(t)\overline{\gamma}), \tag{3.22d}$$
$$\gamma(t) = \frac{\rho(t)}{\omega(t)}, \tag{3.22e}$$

$$\phi(t - \tau(t)) = \phi(t) - \ell, \tag{3.22f}$$
$$b(t) = \omega(t)\gamma(t - \tau(t)), \tag{3.22g}$$
$$p(t) = h(\omega(t), n(t)) \tag{3.22h}$$

where g is a nonlinear function of its input arguments.

3.2.3 Model Identification

In order to design an effective predictive control system for the beverage production line, mathematical models that capture the dynamic behavior and power consumption patterns of the individual machines and conveyor belts involved must be established. We present the modeling and parameter identification approach adopted for the production units in Fig. 3.1, including both binary and continuously operating machines. The reason is that the units identified in this line are used to validate the effectiveness of our proposed control algorithm in two main simulation scenarios. The data for model identification are collected during a period of three years of extensive measurements in cooperation with a beverage enterprise (Brandenburger Urstromquelle GmbH) within the scope of the project Desprima [104]. The power consumption of the processing units is recorded at different processing speeds, for each machine and for the conveyor belts of a single line production.

Nominal Power Consumption of Binary Machines

The blow molding machine and air transport unit operate in a binary mode: either entirely on (nominal state) or off (standby state). Discrete levels, therefore, characterize their energy consumption patterns. The nominal and standby power consumption values, obtained from technical datasheets and validated through experimental measurement, are listed below:

- Blow molding machine:
  - Nominal consumption: 130 kW
  - Standby consumption: 0 kW
- Air transport:
  - Nominal consumption: 11.3 kW
  - Standby consumption: 0 kW

These values are treated as constants in the control model and serve as the base power levels when the machines are on.

Model Identification via Linear Least Squares

Unlike the binary machines, other components such as the filler, conveyor belts, labeling machine, shrink packer, and palletizer exhibit variable power consumption depending on their processing speed. A linear least squares parameter identification method has been implemented to model their behavior accurately. To capture the power variations in terms of processing speed (assuming the number of bottles is fixed on the conveyor belts), a second-order nonlinear model has been adopted for each unit, defined as:

$$p(t) = d\omega^2(t) + M\omega(t) + \tilde{p}_0 \tag{3.23}$$

where d, M, and $\tilde{p}_0$ are unit-specific parameters.

Identification Procedure

The identification process has been carried out systematically to ensure a reliable estimation of model parameters for each unit. The goal is to obtain accurate machine-specific models capturing the nonlinear relationship between power consumption and operational speed. The identification procedure involved the following key steps:

- **Data acquisition:** Operational data has been collected from the production line under various working conditions. For each machine, the input variable $\omega(t)$, representing either conveyor speed or another speed-equivalent metric (e.g., cycle frequency), has been recorded simultaneously with the corresponding electrical power consumption $p(t)$. The data sampling interval has been synchronized with the production control system (e.g., one-second resolution) to ensure time alignment. The data acquisition has been carefully performed to capture both transient and steady-state behaviors by including start-up, ramp-up, and nominal load periods.

- **Preprocessing:** Before model fitting, the raw data have been filtered to remove noise and transient spikes unrelated to mechanical activity, such as signal jitter or measurement artifacts. This step involved applying a low-pass filter and segmenting the dataset to focus on periods of active machine operation. Outliers due to sensor glitches or manual interventions are also removed. The filtered dataset is then normalized when necessary to facilitate numerical stability during optimization.

- **Model structure selection:** Based on physical insights and prior literature, a second-order nonlinear model of the form Eq. (3.23) was selected to represent the power consumption of the machines. This model structure captures both viscous (linear) and inertial (quadratic) effects of rotational movement, as well as a static baseline load $\tilde{p}_0$ associated with idle mechanisms or internal auxiliary systems.

- **Least squares parameter estimation:** The model parameters d, M, and $\tilde{p}_0$ are estimated by minimizing the mean squared error between the measured and modeled power values. The model is rewritten in a linear regression form as:

$$p(t_i) = [\omega^2(t_i)\ \omega(t_i)\ 1] \begin{bmatrix} d \\ M \\ \tilde{p}_0 \end{bmatrix}, \quad \text{for} \quad i = 1, \cdots, N \tag{3.24}$$

  where N is the number of data points. Defining the design matrix $A_{N\times 3}$ and the output vector $\mathbf{p}_{N\times 1}$ as:

$$A = \begin{bmatrix} \omega^2(t_1) & \omega(t_1) & 1 \\ \omega^2(t_2) & \omega(t_2) & 1 \\ \vdots & \vdots & \vdots \\ \omega^2(t_N) & \omega(t_N) & 1 \end{bmatrix}, \quad \mathbf{p} = \begin{bmatrix} p(t_1) \\ p(t_2) \\ \vdots \\ p(t_N) \end{bmatrix} \tag{3.25}$$

  The parameter vector $\theta = [d\ M\ \tilde{p}_0]^T$ is then obtained using the closed-form solution to the linear least squares problem:

$$\boldsymbol{\theta} = (A^\top A)^{-1} A^\top \mathbf{p} \tag{3.26}$$

This approach ensures computational efficiency and robustness, and it avoids the need for iterative solvers typically used in nonlinear optimization.

- **Model evaluation and validation:** After fitting, the identified model is validated on a separate portion of the dataset not used during identification. Model performance is evaluated using multiple criteria, including:
  - Root mean square error (RMSE)
  - Visual inspection of the fitted curve vs. actual data
- **Final model selection:** For each machine, the final model is selected based on the trade-off between accuracy and simplicity. The quadratic model is sufficient for all machines analyzed; no higher-order terms or piecewise approximations are necessary. The identified parameters are subsequently used in the control-oriented model of the plant to simulate power behavior under different operating schedules.

These evaluations ensured that the model not only fits the training data well but also generalizes to unseen operating conditions. In cases where the residuals indicated systematic errors or poor fit in specific speed ranges, the model is retrained with revised preprocessing parameters.

Plots of Fitted Curves

The comparison between the behavior of the identified models and measurements of each unit in the production line (Fig. 3.1) has been depicted in Fig. 3.8 to 3.12. Each plot consists of the following:

- The raw measurement points $(\omega(t_i), p(t_i))$, shown as a dashed curve
- The fitted model $\hat{p}(\omega) = d \cdot \omega^2 + M \cdot \omega + p_0$, shown as a continuous curve

These plots confirm that the proposed second-order polynomial model structure is sufficiently expressive to capture the units' power-speed relationships. In particular, the identified curves follow the nonlinear trends observed in the data while remaining smooth.

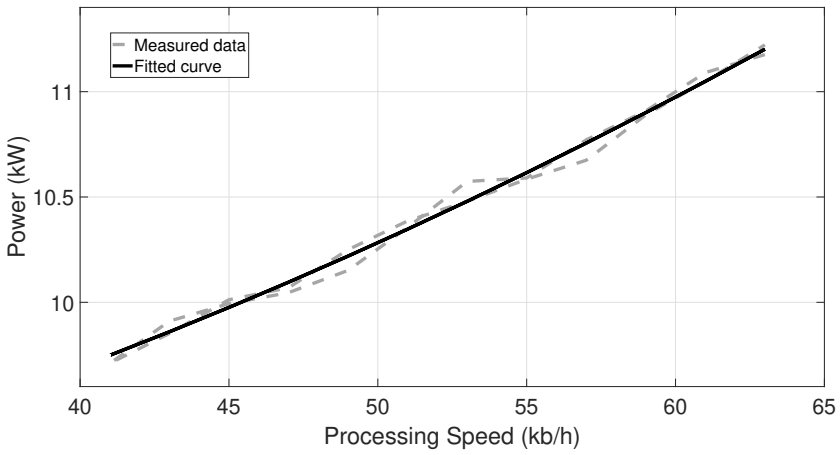


FIGURE 3.8: Measured data and identified power curve for the filler machine.

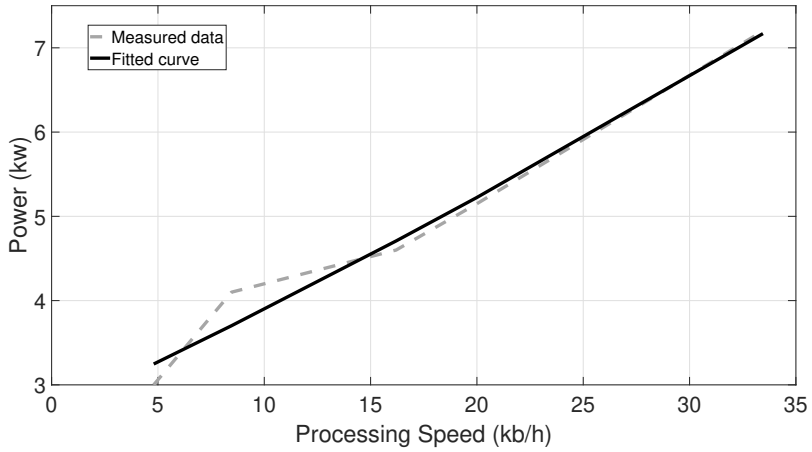


FIGURE 3.9: Measured data and identified power curve for the labeling machine.

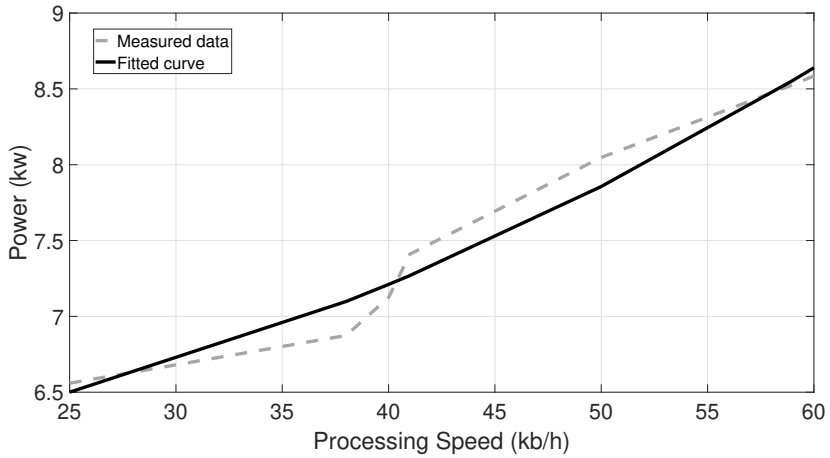


FIGURE 3.10: Measured data and identified power curve for the shrink packer.

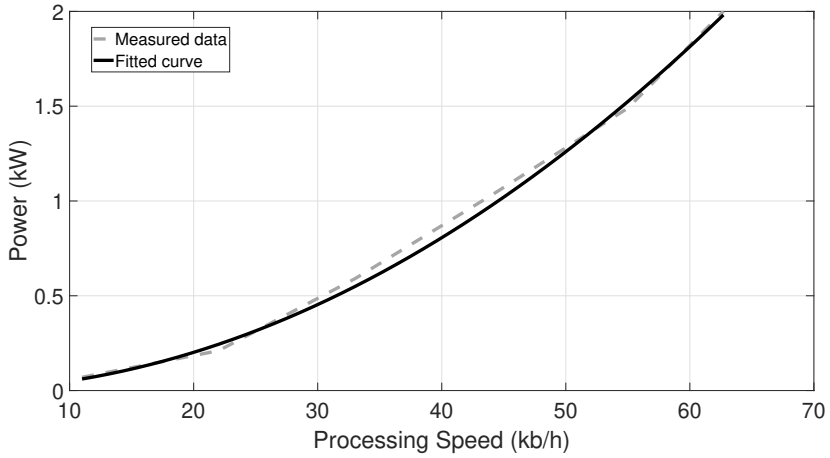


FIGURE 3.11: Measured data and identified power curve for conveyor belt.

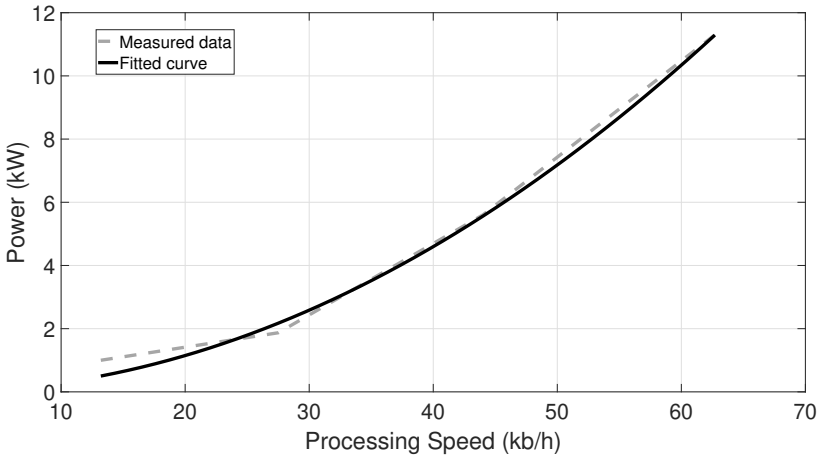


FIGURE 3.12: Measured data and identified power curve for the palletizer.

3.3 Demand Response Objectives and Problem Formulation

Beverage production plants can offer power and energy flexibility to electrical grids, generating revenue at the expense of bottle production losses. This revenue is determined through a flexibility contract between the beverage company and the energy reseller. Meanwhile, a separate contract between the company and its product customers sets a time constraint for producing a specific number of bottles within a given period. Flexibility is provided either by reducing power over short intervals in response to grid requests for frequency stabilization or by decreasing energy consumption over longer intervals to address energy supply and demand imbalances. This chapter focuses on frequency stabilization requests, referred to as “primary demands,” which fall under the broader category mentioned earlier as ancillary demands or ancillary services. Our proposed algorithm is designed to maximize the revenue of beverage processing plants, ensuring optimal operation of production lines under both normal conditions and DR scenarios while consistently meeting contractual bottle production requirements. The beverage company has two contracts: one with the grid and another with the production customer. In the contract with the grid, known as the "e-contract", it is determined how the company responds to the grid’s requests for changing power or energy consumption and how these services are provided by the company and are rewarded. Additionally, in the production contract, also known as a "p-contract", it is specified how much product will be provided to the customer, when, and at what price.

A primary demand should be satisfied in a short time, usually not later than 30 seconds. Therefore, the proposed MPC requires a short horizon to capture rapid changes in demand from the grid, as well as the transient response of the plant. The proposed DDPM is a high-fidelity model that we utilize to deal with primary demands. When a primary demand is sent from the grid to the line’s operator, the power consumption of the lines is reduced to assist the grid’s frequency stability. Therefore, the cost function of the optimization problem should be formulated in terms of the power consumption of the line to handle the demands. The following subsections establish the MPC problem formulation and the controller synthesis for primary demands.

3.4 Control Strategy Design with MPC

Fig. 3.13 illustrates the structure of the proposed MPC for primary services, where the controller processes three main reference signals to determine the optimal component speeds Ω_{ref} along the production lines and the on/off signal Δ for the binary machines. We begin by defining the cost function to formulate the optimization problem. By subsequently incorporating the relevant constraints, we construct the complete formulation of the problem. From an economic perspective, it is optimal for the beverage production lines to operate at lower power consumption levels, implying that one of the primary objectives of the optimization should be to minimize power usage and thereby reduce energy costs. For the sake of simplicity, in the rest of the chapter, we indicate time t with a subscript for the time-varying variables. Therefore, we consider the total power consumption of the lines with $\Sigma_l p_{l,t}$ with the corresponding power tariff $e_{1,t}$ in the cost function, where $p_{l,t}$ is:

$$p_{l,t} = \sum_{j=1}^{n_l} p_{j,l,t} \tag{3.27}$$

$p_{j,l,t}$ stands for the j-th unit of line number l at time t and n_l is the total number of units in line l. Normally, the components in the lines are assumed to run with their nominal speed and, consequently, the nominal power consumption of the lines will be $\Sigma_l p_{\text{nom},l}$ where:

$$p_{\text{nom},l} = \sum_{j=1}^{n_l} p_{\text{nom},j,l} \tag{3.28}$$

$p_{\text{nom},j,l}$ is the nominal power of the j-th unit of line l. Participation in primary DR programs involves deviating from this power consumption at time t, as indicated by Δp_t. The grid punishes any deviation from nominal power consumption $\Sigma_l p_{\text{nom},l}$, while it also leads to a reward from the energy market due to primary service provision. Thus, the terms $(\Sigma_l p_{l,t} - \Sigma_l p_{\text{nom},l} - \Delta p_t)^2$ and $-(\Delta p_t)^2$ are taken into account in the cost function, where the former is meant for power tracking and the latter implies a benefit regardless of the sign of Δp_t. It is assumed that the deviation from $\Sigma_l p_{\text{nom},l}$ is punished by the tariff $e_{2,t}$, and the reward of primary service is $e_{3,t}$.

The bottle production rate of line l at time t, denoted by $b_{l,t}$, directly influences the company's profit and must, therefore, be incorporated into the cost function. This production rate for line l is defined as follows:

$$b_{l,t} = b_{pl,l,t} \tag{3.29}$$

where $b_{pl,l,t}$ is the production rate of the palletizer which is placed at the end of line l,

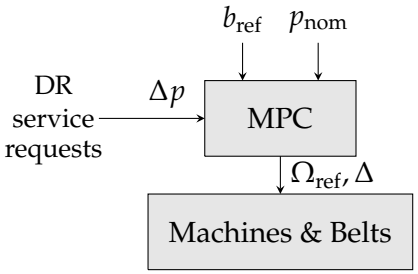


FIGURE 3.13: The proposed power-based MPC architecture.

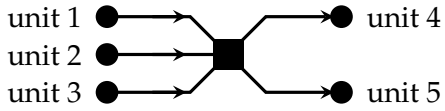


FIGURE 3.14: An example of a conjunction in a production line.

measured in bottles per second or kilo bottles per hour (kb/h). To reflect contractual obligations, we introduce a reference production setpoint $b_{\text{ref},l,t}$ for line l derived from the customer agreement to target the required bottle production. Therefore, the term $\Sigma_l(b_{l,t} - b_{\text{ref},l,t})^2$ is added to the cost function to penalize deviations from the planned production on each line, with a penalty weighted by the tariff $e_{4,l,t}$. At the same time, maximizing bottle production is economically beneficial as it increases the company's revenue. To capture this, we include an incentive term $-b_{l,t}$, scaled by the tariff $e_{5,l,t}$, encouraging higher output. The resulting total cost function is given by:

$$\begin{aligned} J_p = &\sum_{t=1}^{N} \Big(e_{1,t} \sum_{l=1}^{n_L} p_{l,t} + e_{2,t} (\sum_{l=1}^{n_L} p_{l,t} - \sum_{l=1}^{n_L} p_{\text{nom},l,t} - \Delta p_t)^2 \\ &- e_{3,t} \cdot \Delta p_t^2 + \sum_{l=1}^{n_L} e_{4,l,t} (b_{l,t} - b_{\text{ref},l,t})^2 - \sum_{l=1}^{n_L} e_{5,l,t} b_{l,t} \Big) + \sum_{l=1}^{n_L} e_{6,l} \mathcal{E}_{l,p} \end{aligned} \tag{3.30}$$

where N is the prediction horizon, n_L is the total number of lines, and the term $e_{6,l}\mathcal{E}_{l,p}$ will be discussed when explaining the constraints. The constraint set of the optimization problem encompasses not only the system dynamics and the component speed limitations but also incorporates the line topology and flexibility requirements. To ensure safe operation, the rotational speeds of all components are assumed to remain within predefined bounds. Additionally, the specific topology of each beverage line is reflected through a series of constraints that govern the transfer of bottles between consecutive units. Specifically, at a conjunction point inside line l, the summation of outgoing bottle flow rate from preceding units connected to that point must be equal to the incoming bottle flow rate of succeeding units:

$$\sum_{\kappa=1}^{p_{j,l}} b_{\kappa,t} = \sum_{\nu=1}^{s_{j,l}} i_{\nu,t}, \quad j = \{1, \cdots, c_l\} \tag{3.31}$$

where $p_{j,l}$ is the number of proceeding units at the j-th conjunction point in line l, $s_{j,l}$ is the number of succeeding units at the j-th conjunction point in line l, and c_l is the total number of conjunction points in line l. A sketch of a junction is provided in Fig. 3.14.

As the grid requests a power reduction, the MPC responds by lowering the component speeds, decreasing the bottle production rate. Consequently, the total number of bottles produced while providing primary services may fall short of the planned quantity. However, this reduction is only acceptable within a limit predefined by the contract; exceeding this threshold would compromise the ability to

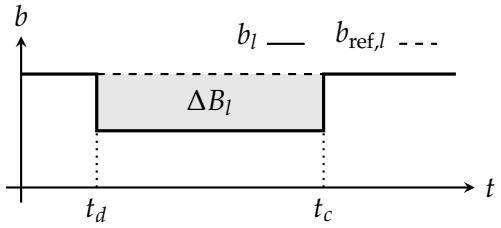


FIGURE 3.15: The bottle loss during a demand service.

recover production later through load shifting and risk violating the customer agreement. To prevent such scenarios, a flexibility constraint is included in the optimization problem, ensuring that the loss in bottle production remains within a contractually defined threshold:

$$n_{\text{loss},l} + \sum_{t=1}^{N} (b_{\text{ref},l,t} - b_{l,t}) - \mathcal{E}_{l,p} \leq \Delta B_l, \quad l \in \{1, \cdots, n_L\} \tag{3.32}$$

$n_{\text{loss},l}$ is the number of lost bottles at the beginning of the horizon, $\mathcal{E}_{l,p} \in \mathbb{R}_{0+}$ is a slack variable, and ΔB_l is the bottle flexibility the company provides for shifting the production to the future, all for the l-th line.

The non-negative slack variable $\mathcal{E}_{l,p}$, associated with a penalty factor $e_{6,l}$, is introduced to prevent infeasibility in the optimization problem, particularly at the onset of primary service provision. This term is added to the cost function in Eq. (3.30) and is meant to be minimized to reduce deviation from the scheduled production. The soft constraint in Eq. (3.32) ensures that, during primary service participation, the production rate $b_{l,t}$ remains close to the reference $b_{\text{ref},l,t}$ as the number of lost bottles $n_{\text{loss},l}$ approaches the threshold ΔB_l (see Fig. 3.15).

It is important to note that constraint in Eq. (3.32) also guides the total power consumption of the lines, $\Sigma_l p_l$, to optimally return from the reduced power trajectory $\Sigma_l p_{\text{nom},l} + \Delta p_t$ back toward the nominal power $\Sigma_l p_{\text{nom},l}$ as the number of lost bottles approaches its limit. Based on the constraints described above, the complete MPC problem formulation is as follows:

$$\min_{\Delta, \Omega_{\text{ref}}} \quad J_p \tag{3.33a}$$

subject to:

$$\mathbf{x}_{j,l,t+1} = \mathbf{x}_{j,l,t} + T_s \cdot f_{j,l,t}(\mathbf{x}_{j,l,t}, \mathbf{u}_{j,l,t}), \tag{3.33b}$$

$$j = \{1, \cdots, n_l\}, \quad l = \{1, \cdots, n_L\}, \quad t = \{1, \cdots, N\}, \tag{3.33c}$$

$$p_{j,l,t} = d_{j,l}\omega_{j,l,t}^2 + M_{j,l}\omega_{j,l,t} + \alpha_{j,l} n_{j,l,t} + p_{0,j,l}, \tag{3.33d}$$

$$\omega_{\text{ref},m,l} = \delta_l \cdot \omega_{\text{nom},m,l} + (1 - \delta_l) \cdot \omega_{\text{standby},m,l}, \quad m \in \{\text{BM}, \text{AT}\} \tag{3.33e}$$

$$i_{\text{bm},l} = \delta_l \cdot i_{\text{nom},l}, \tag{3.33f}$$

Eq. (3.27), (3.33g)

Eq. (3.28), (3.33h)

Eq. (3.29), (3.33i)

line topology (Eq. (3.31)), (3.33j)

$$\omega_{\min,j,l} \leq \omega_{j,l,t} \leq \omega_{\max,j,l}, \tag{3.33k}$$

Eq. (3.32) (3.33l)

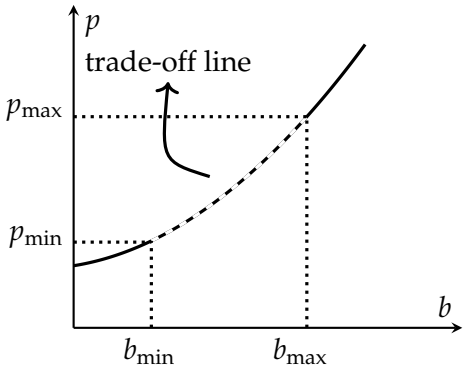


FIGURE 3.16: Power consumption characteristics in terms of production rate.

where $\Delta = [\delta_1, \cdots, \delta_{n_L}]^T$, $\Omega_{ref} = [\omega_{ref,1,1}, \cdots, \omega_{ref,n_1,1}, \cdots, \omega_{ref,1,n_L}, \cdots, \omega_{ref,n_{n_L},n_L}]^T$. Also, $i_{nom,l}$ is the nominal inflow rate of line l, $x_{j,l,t}$, $f_{j,l,t}$, $p_{j,l,t}$ are the state vector, nonlinear dynamics vector, and the power of the j-th unit in line l at time t, respectively, based on Eq.(3.3) and (3.4). Eq. (3.33) presents the complete formulation of the mixed-integer nonlinear optimization problem that must be solved to achieve an optimal power-based control strategy.

In the proposed MPC algorithm, the ancillary demands tend to decrease the speed of the machines and belts, while production constraints limit this reduction. The controller must decide how to balance the trade-off between the benefits of demand provision and bottle production. For a single component (machine or belt), let us assume that the power consumption of the component is $p = p_2\omega^2 + p_1\omega + p_0$ and the production rate is $b = \gamma\omega$. These equations are compatible with equations mentioned in Section 3.2. Therefore, the power consumption reduction of the component could be interpreted as $p_2\omega^2 + p_1\omega + p_0 \leq p_{nom}$. Also, the minimum production rate for bottles could be written as $b \geq b_{min}$. These two constraints and the speed limits (Eq. (3.33k)) form the trade-off line, as illustrated in Fig. 3.16, where the controller must find the optimal working point of the component on this line.

The cost function introduced in Eq. (3.30) assigns a reward for production to each line, resulting in a priority order among the lines. Also, each line is penalized differently for deviating from its reference bottle production rate. In contrast, the price of energy (e_1) is the same for all production lines.

Participation in DR is offered collaboratively by the lines, each of which makes its own contribution to the primary demands. The second term in the cost function of the proposed MPC $(\sum_{l=1}^{n_L} p_{l,t} - \sum_{l=1}^{n_L} p_{nom,l,t} - \Delta p_t)^2$ implies that all the lines participates cooperatively in DR such that the overall optimal service provision is fulfilled.

3.5 Simulation Scenarios

The proposed control strategy is validated through a simulation scenario conducted in MATLAB©. The scenario addresses the primary service for assisting the grid's frequency stability. Therefore, a nonlinear MPC is implemented using CasADi© — a MATLAB© toolbox for solving optimization problems, as introduced in [105]. All the simulations are performed with an 8-core CPU in a system with 32 GB RAM. In both scenarios, the prediction horizon is $N = 60$ and the control horizon is 1 sample. Also, the mixed-integer nonlinear optimization (3.33) is solved, once assuming the

binary section is stopped, and once assuming it works in nominal mode. In each case, the computation time is less than 0.5 seconds.

3.5.1 Scenario I: One Production Line

In this scenario, we assume that the controller developed in Section 3.4 is applied to one production line with the topology illustrated in Fig. 3.1. The MPC algorithm optimally adjusts the line inflow and component speeds to satisfy the demand requirements. Two primary demand events are introduced at $t = 150\,\text{s}$ and $t = 250\,\text{s}$, targeting a production line with a nominal power consumption of 200 kW. The first demand corresponds to a power reduction of $\Delta p = 40\,\text{kW}$, and the second to $\Delta p = 140\,\text{kW}$. The simulation parameters are mentioned in the Table. 3.1. Fig. 3.17 presents the line's power consumption, bottle production rate, and the speed profiles of the blow molding machine and the filler.

TABLE 3.1: **Simulation parameters of the production line for scenario I.**

Parameter	Value (unit)	Parameter	Value (unit)
T_s	0.5 s	k_1 shrink packer	3960 kg · m^2/s
d_f	6214.3 kg · m^2/s	k_2 shrink packer	3168 kg · m/s
d_b	6098 kg · m^2/s	k_1 palletizer	360,000 kg · m^2/s
d_l	1734.2 kg · m^2/s	k_2 palletizer	108,000 kg · m/s
d_{sp}	8317 kg · m^2/s	α belt	0.01 W
d_{pl}	34757.7 kg · m^2/s	p_0 filler	8.3672 kW
M_f	1390.2 kg · m^2/s^2	p_0 labeling machine	2.83 kW
M_b	1 kg · m^2/s^2	p_0 shrink packer	6 kW
M_l	6053.9 kg · m^2/s^2	p_0 palletizer	0.15 kW
M_{sp}	296.8 kg · m^2/s^2	γ_f	30.56
M_{pl}	1 kg · m^2/s^2	γ_l	15.28
k_1 filler	960 kg · m^2/s	e_1	0.02 c.u.*/W
k_2 filler	480 kg · m/s	e_2	0.5 c.u./W^2
k_1 belt	3600 kg · m^2/s	e_3	10 c.u./W^2
k_2 belt	2880 kg · m/s	e_4	0.01 c.u. s^2
k_1 labeling machine	135 kg · m^2/s	e_5	0.05 c.u. s
k_2 labeling machine	108 kg · m/s	p_{nom}	200.6 kW

* c.u. stands for "currency unit".

The first demand can be met by reducing the power consumption of non-binary machines and belts; therefore, the binary machines, which collectively consume 141.3 kW (over 70% of the total line consumption), do not participate in this DR event. The whole consumption decreases from 200.2 kW to 167.1 kW, leading to 3.31 € participation reward. The amount of reward per each kW is considered 0.1 €, which is within a realistic and acceptable range according to the primary control market [106]. As shown in the lower left subplot, the speed of the blow molder (ω_{bm}) remains unchanged during the first demand period, while the speed of the filler (ω_f) is significantly reduced. Additionally, the palletizer at the end of the line reduces its processing speed, resulting in a noticeable decrease in the final bottle production rate, as indicated by the blue curve labeled b in the upper right subplot. During the second demand interval, the burden of the demand is on the binary machines rather than the non-binary machines and belts. Therefore, the speed of the

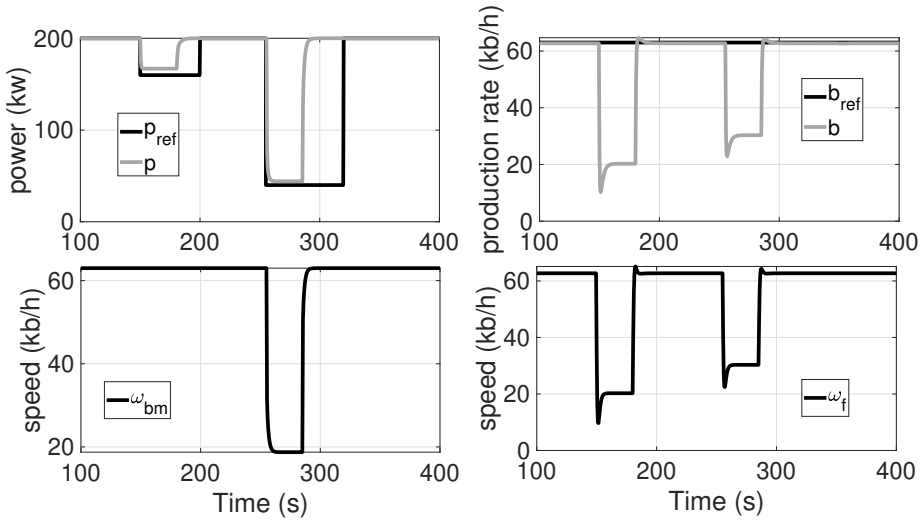


FIGURE 3.17: Power consumption, bottle production rate, the speed of the blow molding machine, and the filler of one line for primary service.

filler is reduced less than the first demand. This also holds for the palletizer, resulting in a less significant reduction in production rate during the second phase of demand. The whole consumption decreases from 200.2 kW to 44.2 kW, leading to 15.6 € participation reward. The amount of reward per each kW is considered 0.1 € . In both demands, the power consumption of the line does not precisely track the reference power because following the reference power is not the only goal the controller seeks. It is also important to make a trade-off between reducing power consumption and increasing bottle production. Also, the power curve (illustrated by gray color in the upper left subplot), regains its nominal value due to the flexibility contract introduced in Eq. (3.32). This profile returns to its nominal value before the demand event is completed.

3.5.2 Scenario II: Three Production Lines

To investigate the performance of the proposed MPC for more production lines, we have considered the line configuration in Fig. 3.18 to simulate the controller. In this configuration, two lines are similar to the line in Fig. 3.1, and one line has only one labeling machine. The parameters of the third line, as well as the parameters that are different from Scenario I, are presented in Table 3.2. The first two lines produce bottles with lower priority than the third line, which involves only one labeling machine. The mixed-integer nonlinear optimization is solved once, assuming the binary section is stopped, and once assuming it works in nominal mode. In each case, the computation time is less than 0.5 seconds. The simulation results are shown in Fig. 3.19, which includes four subplots. The top left subplot illustrates the power profiles of the lines and the reference power profile. Also, the total power consumption of the lines is plotted in gray. There are two intervals of primary demand, each requesting power reduction. The first demand is $\Delta p = 116.26$ kW and the second demand is $\Delta p = 468.52$ kW. The total power is reduced each time the demand is received. However, it does not fully track the new setpoint for power consumption, while the production rate has a significant impact on participation in demand services. Also, ΔB_l for each line is specific and limited, not allowing the lines to stay

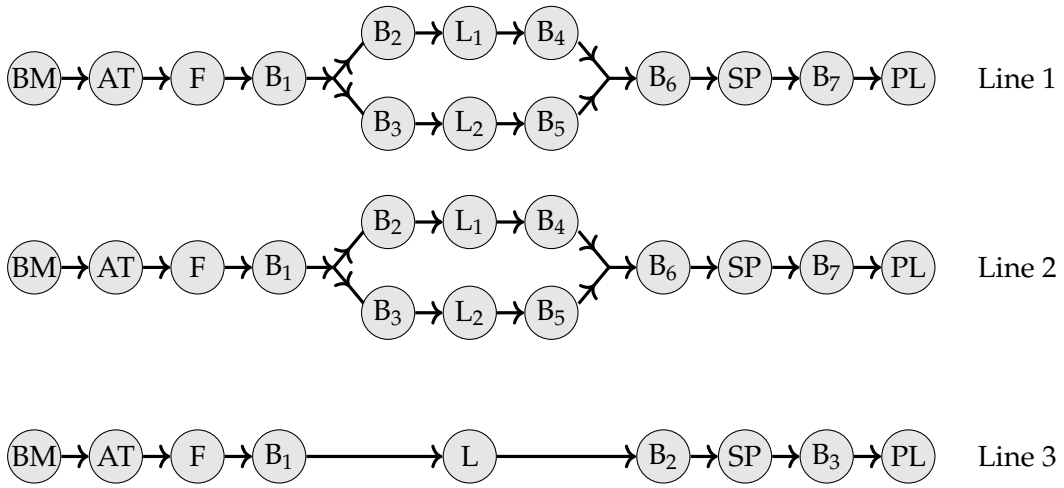


FIGURE 3.18: Three beverage production lines.

entirely in demand mode. Therefore, the total power consumption recovers to its nominal value of 585.94 kW before the setpoint (p_{ref}) returns to its nominal value. The power profile of the two identical lines (p_1 and p_2) is similar, and these two lines react accordingly to the demands. Considering 0.1 € per kW as the reward of the primary demands, the total reward for the first demand is 9.14 € and for the second demand is 36.75 €.

In both phases of the demand, the participation of the third line is considerably less than that of the others. The primary reason for this behavior is that the production rate of the third line is prioritized over the others. This usually indicates the importance of the production due to the contract between the production factory and the customer. The upper right subplot shows the production rate of the lines. The rate of production has been reduced twice due to the demands. Since the production rate of the third line has priority, the reduction in its production rate (b_3) is less than that of the others. The production rate of line 1 (b_1) and line 2 (b_2) is identical. In the first demand, only continuous machines of the three lines participate; however, the second demand is covered by the binary machines of lines 1 and 2.

Therefore, the reduction in b_1 and b_2 during the first interval is more than during the second interval. The entire portion of the second demand dedicated to line

TABLE 3.2: **Simulation parameters of the production line for scenario II.**

Parameter	Value (unit)
d_b belt 1 & 2 of line 3	18294 kg $\cdot$ m^2/s
M_b belt 1 & 2 of line 3	3.1 kg $\cdot$ m^2/s^2
d_l line 3	3468.4 kg $\cdot$ m^2/s
M_l line 3	12107.8 kg $\cdot$ m^2/s^2
p_0 labeling machine of line 3	3.11 kW
e_4 line 1 & 2	0.02 c.u.* s^2
e_5 line 1 & 2	0.06 c.u. s
e_4 line 3	0.07 c.u. s^2
e_5 line 3	0.19 c.u. s

* c.u. stands for "currency unit".

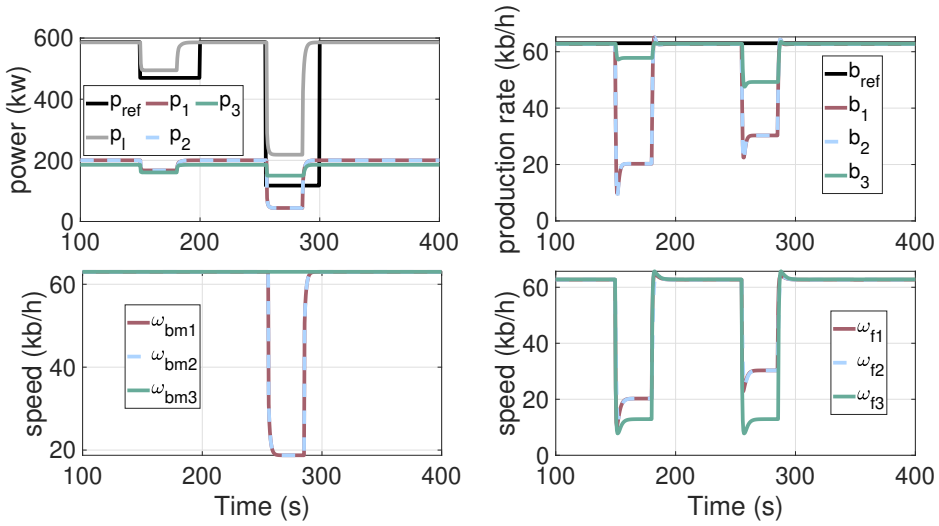


FIGURE 3.19: Power consumption, bottle production rate, the speed of the blow molding machine, and the filler of three lines for primary service.

3 is taken by its continuous machines, resulting in a greater reduction compared to the first demand. The lower left subplot depicts the processing speed of the blow molding machines of the three lines. These machines do not participate in the first demand, while the blow molding machines of lines 1 and 2 participate in the second demand. The reason is that the demand in the second instance is greater than the first, and the binary machines contribute due to their significant power consumption. However, the rate of the third blow molding machine has remained unchanged to avoid bottle loss at the in-feed of the filler. If this bottle loss is not avoided, it will propagate along the whole line, resulting in a significant loss in the final production. Finally, the lower right subplot illustrates the processing speed of the fillers. All the fillers participate in both demands with different processing rates. The third filler reduces its speed more than the others, implying that the demands are taken mainly through continuous machines for line 3. During the second demand, since it is primarily captured by the blow molding machines of lines 1 and 2, the precipitation of their continuous machines (including fillers) is less compared to the first demand.

The resulting optimization problem is, in general, computationally challenging. This complexity is a direct consequence of accurately representing the physical and operational characteristics of the production system and cannot be avoided without significantly compromising model fidelity. For a limited number of production lines, the resulting computation time remains below the chosen sampling interval, ensuring real-time feasibility. As the number of production lines increases, the computational burden grows accordingly. In such cases, scalability can be addressed through parallel computing strategies, solver selection, and the use of appropriate computational hardware. While a detailed scalability analysis is beyond the scope of this thesis, the presented results demonstrate that for the second scenario involving three production lines, the computation time of the proposed MPC remains below the sampling time and is sufficiently fast for our target real-time application.

3.6 Conclusion

This chapter presented a comprehensive framework for implementing DR strategies in beverage production plants through the application of MPC. The methodology is tailored to exploit the flexibility of production machines, particularly those with binary operational modes, to participate in power-based DR programs. The chapter introduced a multi-objective optimization approach, balancing production targets with DR incentives, and detailed the simulation environment, including machine modeling, demand profiles, and control scenarios. Simulation results demonstrated the effectiveness of the proposed algorithms in achieving energy cost savings and contributing to grid stability while minimizing production losses. The chapter concluded by highlighting the potential of industrial processes to act as active participants in the energy market through intelligent control. Since the proposed algorithm is a generic strategy for dealing with the problem of DR control across multiple production lines, it can also consider the problem of prioritizing production lines with stricter constraints. The research of this chapter has been conducted in the context of the Desprima project, funded by the German Federal Ministry for Economic Affairs and Energy (BMWE), from 2019 to 2023, under Grant 03ET1672G.

Chapter 4

Optimal Control Strategy for Bidirectional EV Charging Hubs

4.1 Introduction to Methodology

This chapter shifts the focus to the mobility sector, in which EV charging hubs equipped with V2G technology offer significant potential for DR participation and energy market integration. Referring to Fig. 1.1, it is vital to mention that the aim in this chapter is to design the C_2 and C_3 controllers "separately" for EV charging hubs, apply them "individually" to the plant, and validate their performance. It presents two advanced control algorithms designed to trade both power and energy flexibility from EV charging hubs. The C_2 controller has been designed in a single-layer structure while C_3 has been designed in a double-layer structure to reduce computational time. The proposed methods aim to achieve a dual objective: enhancing the economic profitability of the charging infrastructure and enabling the provision of ancillary services to the power grid. By integrating multi-asset energy optimization with MPC, the two algorithms address energy transactions and grid-oriented flexibility across two different time scales. The first approach (C_2) introduces a single-layer MPC based on a MILP formulation to optimize energy flows among EVs, stationary buffers, and the grid. The second approach (C_3) proposes a double-layer MPC structure that aggregates EVs into a virtual battery for high-level energy planning while distributing power among individual EVs using a static MILP at the lower layer. Both approaches use relaxation techniques to reduce computational burden, transforming inherently nonlinear problems into tractable MILP formulations.

4.2 System Modeling and Energy Market

The charging hub is equipped not only with EV chargers but also with stationary energy buffers that facilitate optimal energy management. While this work considers a single energy buffer for modeling and simulation purposes, the proposed control strategies are scalable and can be extended to configurations involving multiple buffers. A general schematic of the charging hub, including the layout of its storage systems, is illustrated in Fig. 4.1. It is worth noting that the number of EVs connected to the hub is not limited to those depicted in the figure. The inclusion of an energy buffer—typically possessing a larger capacity than individual EV batteries—serves multiple strategic purposes. First, it enables the hub operator to exploit time-varying electricity prices by storing energy during low-price periods (e.g., off-peak hours) and discharging during high-price intervals, thereby maximizing economic returns through energy arbitrage. Second, in the context of DR programs, the buffer enhances operational flexibility by allowing the hub to quickly adjust its

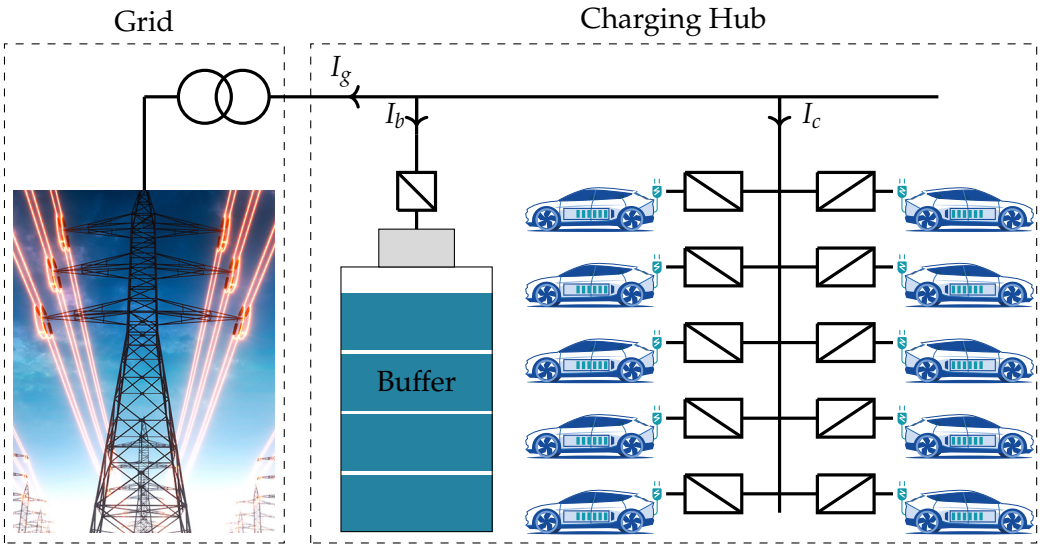


FIGURE 4.1: A schematic diagram of the charging hub structure.

power consumption in response to grid signals, independent of the availability or cooperation of EVs. Finally, the buffer can act as an intermediary supply source for EV charging when drawing power directly from the grid is economically or operationally suboptimal.

The charging hub operator is assumed to be the decision-making entity of an EMS that is responsible for the coordinated operation of the involved energy assets, including the grid supply, hub energy buffers, and connected EVs. Therefore, from the perspective of the operator, the charging hub is a medium of physically coupled flexible energy resources and the task of the operator consists in managing the energy flow pathways and corresponding transactions (i.e. purchases and sales). The latter i.e. individual bidirectional energy flow pathways are weighted by internal and external energy prices corresponding to the assets of the hub (local buffers and EVs) and energy markets (e.g. spot market), respectively. Lastly, the resulting transactions sum up to the cost function of an optimization problem which in this thesis is used to model the profit of the operator.

4.2.1 Battery Abstract Model

A simplified battery model based on Coulomb counting, as commonly used in the literature [107], [108], is adopted in this study. This model estimates the state of charge by accounting for the Coulombic efficiency and the capacity of the battery [107]:

$$\dot{x}(t) = \frac{\eta^{'}}{C} I(t) = \frac{\eta^{'}}{CV} P(t) = \frac{\eta}{C} P(t), \ t \geq 0 \tag{4.1}$$

x is the SOC of the battery, η is the Coulombic efficiency, C is the capacity of the battery, I is the injected current into the battery, V is the voltage of the battery, P is the injected power into the battery, and t is continuous time. It is evident that when $P > 0$, the battery is in charging mode, whereas for $P < 0$, it is discharging. This abstract battery model is applied to both the EVs and the stationary energy buffer within the charging hub. While real batteries exhibit complex behaviors—including nonlinear

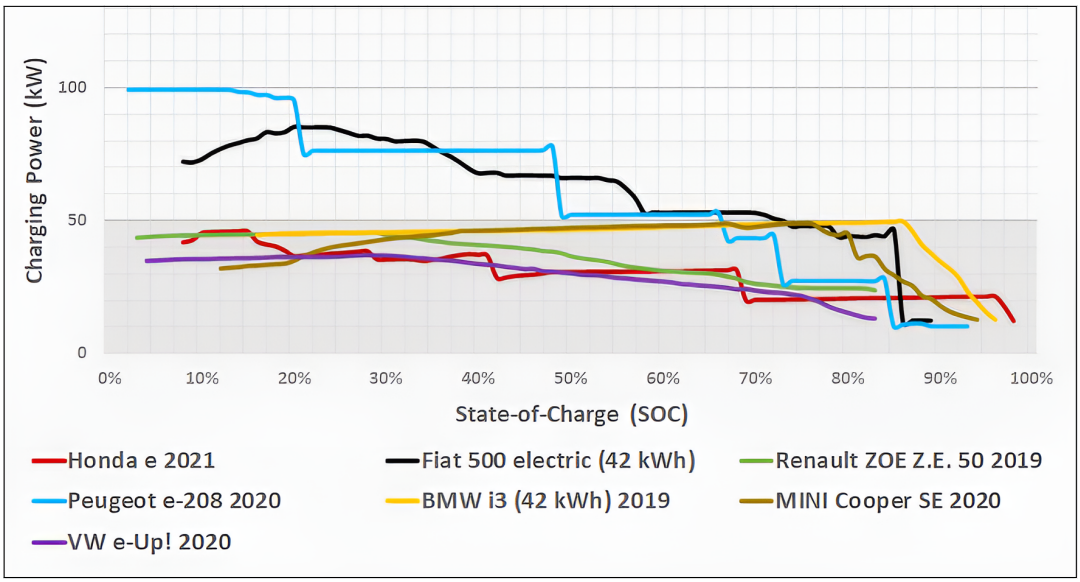


FIGURE 4.2: DC fast charging characteristics of some EV manufacturers [111]. This figure illustrates the amount of power required for fast-charging EVs from different car brands in terms of SOC. Some have almost constant power during charging mode, while others change considerably when SOC increases.

dynamics, temperature dependence, aging effects, and variable efficiency—such details are intentionally omitted in abstract models. This simplification enhances computational efficiency and facilitates high-level system analysis, aligning with one of the primary objectives of this study. In large-scale charging hubs with numerous EVs, the use of detailed battery models is often impractical and unnecessary; abstract models sufficiently capture the essential dynamics required for energy management and grid interaction without the added complexity of modeling individual battery characteristics.

The power applied to a battery (or equivalently, the charging/discharging current I) is subject to strict limitations due to safety concerns, power losses, and variations in internal resistance, which manifest as heat and can lead to elevated temperatures. These power limits are not constant and typically vary with the battery's SOC [109], [110]. To ensure efficient and safe operation with minimal energy loss and thermal stress, EV manufacturers usually provide recommended DC fast-charging profiles tailored to each vehicle model [111]. An example of such profiles is illustrated in Fig. 4.2.

To include the battery model in Eq. 4.1 into the control design, it is essential to differentiate between the charging and discharging powers of the storage systems. Thus, the power applied to the j-th EV in the charging hub at time t can be expressed as follows:

$$P_j(t) = \delta_{1j}(t)P_{1j}(t) + \delta_{2j}(t)P_{2j}(t), \tag{4.2a}$$

$$\delta_{1j}(t) + \delta_{2j}(t) \leq 1, \quad j \in \{1, \cdots, n\} \tag{4.2b}$$

where δ_{1j} and δ_{2j} are binary variables, $P_{1j} > 0$ is the charging power, $P_{2j} < 0$ is the discharging power, and n is the total number of EVs. Eq. (4.2) describes the three modes of the EVs where it can charge ($\delta_{1j} = 1$, $\delta_{2j} = 0$), discharge ($\delta_{1j} = 0$, $\delta_{2j} = 1$), or be in standby mode (neither charging nor discharging when $\delta_{1j} = \delta_{2j} = 0$). The battery can be charged or discharged by setting δ_{1j} or δ_{2j} to 1, respectively. This

equation can also be used for the buffer if the subscript notation j is replaced with b as follows:

$$P_b(t) = \delta_{1b}(t)P_{1b}(t) + \delta_{2b}(t)P_{2b}(t), \tag{4.3a}$$

$$\delta_{1b}(t) + \delta_{2b}(t) \leq 1 \tag{4.3b}$$

where δ_{1b}, and δ_{2b} are binary variables, and $P_{1b} > 0$, $P_{2b} < 0$. The primary motivation for modeling the batteries using Eq. (4.2) in the control design process is the differentiation between energy tariffs for purchasing energy from and selling energy to the storage components. For design purposes, the power injected into or drawn from the grid is described similarly:

$$P_g(t) = \delta_{1g}(t)P_{1g}(t) + \delta_{2g}(t)P_{2g}(t), \tag{4.4a}$$

$$\delta_{1g}(t) + \delta_{2g}(t) \leq 1 \tag{4.4b}$$

where $P_{1g} > 0$, $P_{2g} < 0$. The equation above enables the controller to differentiate between selling energy to and buying energy from the grid.

4.2.2 Spot Energy Market

As mentioned previously, the energy market serves as a platform for electricity trading, allowing producers, consumers, and intermediaries to exchange energy resources in response to supply and demand dynamics. This market is a complex network involving various participants, including resellers, who buy and sell energy based on market principles [112]. The main objective of the energy market is to guarantee a reliable energy supply at the lowest possible cost [94], [113]. This involves balancing supply and demand in real time while planning for future energy needs. The energy market is shaped by various factors, including regulations, technological advancements, and environmental fluctuations, which can affect energy prices and availability. The intraday market is a segment of the electricity market where energy transactions occur throughout the day, enabling participants to buy and sell energy closer to the time of actual delivery. This market allows for adjustments on short notice—often in hourly or even 15-minute intervals—helping energy suppliers, utilities, and large consumers better match supply and demand as conditions evolve in real time. As the intraday market operates continuously, participants can adjust their strategies based on the most up-to-date information, which is particularly valuable for managing variable renewable resources [114]. In fact, participants in the intraday market can optimally respond to unexpected changes in the grid, such as fluctuations in demand or supply disruptions originating from volatile renewable energy producers, and benefit from the reward of ancillary service provision.

The balancing market, also called the real-time market, operates to ensure that electricity supply and demand are balanced with a high time resolution (on the order of seconds). This market becomes crucial when there are discrepancies between forecasted demand and actual consumption or between scheduled generation and actual power production. Participants in the balancing market, such as grid operators and power consumers, adjust their power production or consumption to maintain grid stability [115]. Balancing market prices are typically higher than those in the intraday market due to the immediate need to secure additional resources or reduce power consumption. It plays a vital role in preserving the reliability and integrity of the power system, particularly during unforeseen events or sudden fluctuations in energy supply and demand. In this context, the charging hub is considered a consumer that trades energy in the market and participates in DR programs from

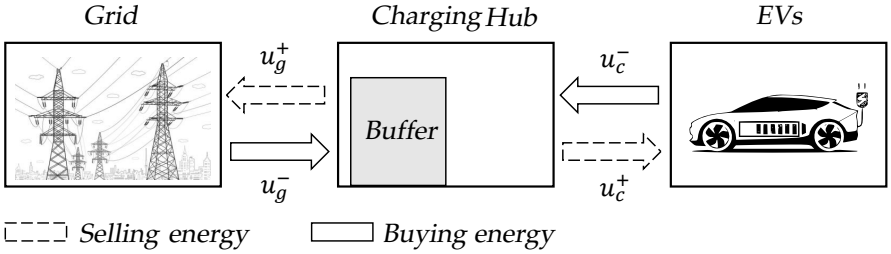


FIGURE 4.3: Energy transaction directions. u_c^+ and u_g^+ is the energy sold to all the EVs and the grid, and u_c^- and u_g^- is the energy bought from all the EVs and the grid.

intraday and balancing markets. As a result, the hub can adopt an optimal strategy to trade energy with the grid by exploiting the capacities of the EVs and the buffer while fulfilling grid demands (Fig. 4.3).

4.2.3 Multi-asset optimization for energy trading

The problem of allocating energy across multiple assets can be formulated as a deterministic energy optimization and control task: the charging hub coordinates coupled energy resources and energy flows—grid exchange, the stationary buffer, and the connected EVs—under operational constraints and time-varying tariffs. The energy stored in the EV fleet and in the buffer represents dispatchable energy states that can be scheduled over time. The proposed controllers regulate charging/discharging and grid exchange to maximize the operator's economic objective while respecting availability, power limits, and required SOC. Bidirectional V2G operation enables redistribution of energy between the storage systems and the grid, which is exploited for energy trading. Because buying and selling prices differ and vary over time, the controllers continuously adapt the energy dispatch and grid transactions according to the energy tariffs. The cumulative economic value of the charging hub at time $t = 0$ can be expressed as:

$$X_0 = c_0 + c_{b_0} u_{b_0} \tag{4.5}$$

where c_0 is the initial wealth of the charging hub, u_{b_0} is the initial energy stored in the buffer, and c_{b_0} is the price of the energy stored in the buffer at $t = 0$. The amount of energy stored in the charging hub or sold to the grid is represented by u_{1m}, where $m \in \{g, b, 1, \cdots, n\}$. The symbol g stands for the grid, and b stands for the buffer. Also, the amount of energy bought from EVs, taken from the buffer, and purchased from the grid is denoted by u_{2m}. Therefore, the wealth of the charging hub at time $t > 0$ is expressed as:

$$X_t = c_0 + c_{b_0} u_{b_0} + \sum_{\sigma=1}^{2} \left(u_{\sigma g}(t) G_\sigma(t) + u_{\sigma b}(t) B_\sigma(t) + \sum_{j=1}^{n} u_{\sigma j}(t) C_\sigma(t) \right) \tag{4.6}$$

where G_1 is the price of energy that the charging hub sells to the grid (u_{1g}), G_2 is the price of energy that the charging hub buys from the grid (u_{2g}), B_1 is the price of the energy stored in the buffer (u_{1b}), B_2 is the price of the energy taken from the buffer (u_{2b}), C_1 is the price of energy that the charging hub sells to the EVs (u_{1j}), and C_2 is the price. The energy prices of the buffer include grid fees, taxes, battery degradation costs, and other costs that the hub bears for storage. When the energy

of the buffer is sold to the grid or EVs, the charging hub profit margin must be also considered. The prices of the grid are obviously taken from the energy market. The prices determine the direction of energy flow during a trading phase or DR program.

4.3 Optimal Control Strategy

This section addresses the controller problem formulation for an EV charging hub, which includes several EVs and a buffer. The control systems must provide a real-time solution for energy transactions. The MPCs find the optimal flow of energy while considering operational constraints such as DR and terminal SOC.

Cost Function: The economic benefit of the charging hub within the horizon of the controller is driven from Eq. (4.6). The amount of energy at time $k \cdot T_s$ is calculated from the charging power (P_{1m}) and discharging power (P_{2m}) in discrete time as follows:

$$u_{1m}(k) = \kappa \cdot T_s \sum_{i=1}^{k} \delta_{1m}(i) P_{1m}(i), \tag{4.7a}$$

$$u_{2m}(k) = \kappa \cdot T_s \sum_{i=1}^{k} \delta_{2m}(i) P_{2m}(i) \tag{4.7b}$$

where T_s is the sampling time. It is assumed that during charging or discharging, the voltage for each component is constant and κ is the conversion factor from Joule to Megawatt. The cost function of the MPCs can now be written as:

$$J_N = \kappa \cdot T_s \cdot \sum_{k=1}^{N} \sum_{m} [\delta_{1m}(k) P_{1m}(k) W_{1m}(k) + \delta_{2m}(k) P_{2m}(k) W_{2m}(k)] \tag{4.8}$$

where $m \in \{g, b, 1, \cdots, n\}$, $W_{1g} = G_1$, $W_{2g} = G_2$, $W_{1b} = B_1$, $W_{2b} = B_2$, $W_{1j} = C_1$, $W_{2j} = C_2$, N is the prediction horizon. This cost function accounts for both the revenues and losses incurred by the charging hub during energy transactions. Specifically, the hub's wealth increases through energy sales and decreases with energy purchases. Consequently, the cost function aggregates all gains and losses to enable the MPCs to determine the optimal direction of energy flow. The objective function represents the allocation of energy among these assets over a finite prediction horizon. Electricity prices are treated as known or forecasted inputs rather than stochastic variables, resulting in a deterministic "Multi-asset Optimization" problem. While this formulation does not represent stochasticity arising from price variations, the goal is to jointly optimize multiple assets with different characteristics and constraints.

Common Constraints: Referring to Eq. (4.1) to Eq. (4.4), the dynamic models of the batteries are incorporated as equality constraints in the optimization problem. The SOC of the buffer and the EVs serve as state variables, while the injected power into these storage systems represents the control input. The grid power P_g is determined based on the energy conservation law once the other inputs have been computed.

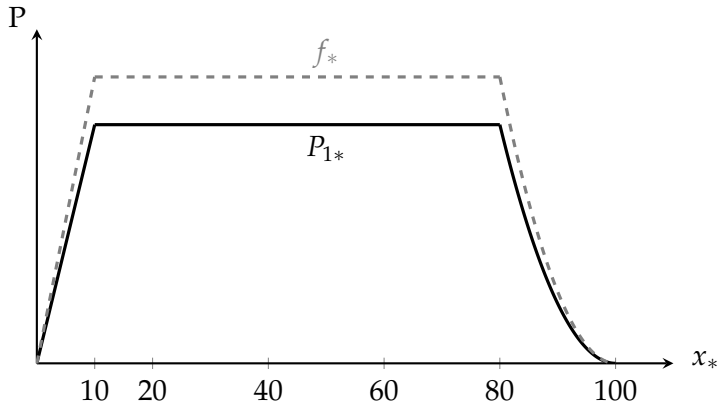


FIGURE 4.4: Charging power characteristic and its bound. This plot shows how the upper bound on the charging powers changes in terms of SOC.

$$\begin{cases} x_b(k+1) = x_b(k) + \dfrac{T_s \eta_b}{C_b} \left(\delta_{1b}(k) P_{1b}(k) + \delta_{2b}(k) P_{2b}(k)\right), & \text{(4.9a)} \\ \delta_{1b}(k) + \delta_{2b}(k) \leq 1, & \text{(4.9b)} \\ x_j(k+1) = x_j(k) + \dfrac{T_s \eta_j}{C_j} \left(\delta_{1j}(k) P_{1j}(k) + \delta_{2j}(k) P_{2j}(k)\right), & \text{(4.9c)} \\ \delta_{1j}(k) + \delta_{2j}(k) \leq 1, \quad k = 1, \cdots, N, \quad j = 1, \cdots, n & \text{(4.9d)} \end{cases}$$

As previously mentioned, the charging and discharging powers of the batteries are subject to limitations due to factors such as performance, safety, and longevity. Therefore, these power bounds must be explicitly considered in the controller synthesis:

$$\begin{cases} 0 \leq P_{1b}(k) \leq f_b(x_b(k)), & \text{(4.10a)} \\ P_{b,\min} \leq P_{2b}(k) \leq 0, & \text{(4.10b)} \\ 0 \leq P_{1j}(k) \leq f_j(x_j(k)), & \text{(4.10c)} \\ P_{j,\min} \leq P_{2j}(k) \leq 0 & \text{(4.10d)} \end{cases}$$

$P_{b,\min}$ and $P_{j,\min}$ are minimum discharging powers of the buffer and the j-th EV, respectively. The charging power of the buffer and the EVs is bounded by a nonlinear profile dependent on the "charging characteristic" of the batteries. The charging powers of the buffer and the EVs are constrained by a nonlinear profile that depends on the batteries' charging characteristics. This characteristic, typically provided by the battery manufacturer, defines how to charge the battery efficiently and is expressed as a function of the SOC. Fig. 4.4 illustrates a representative charging characteristic, along with the corresponding nonlinear functions f_* used in the formulation. Another constraint in the optimization problem arises from the physical architecture of the charging hub. Based on the law of conservation of energy, assuming no energy losses within the hub, the following constraint must be imposed:

$$P_g(k) + P_b(k) + \underbrace{\sum_{j=1}^{n} P_j(k)}_{P_c(k)} = 0 \tag{4.11}$$

where the positive direction of the storage systems and the grid is towards these players (Fig. 4.1).

4.3.1 Single-layer Controller (C_2)

The constraints discussed so far apply to both the single-layer MPC (C_2) and double-layer MPC (C_3) frameworks. This subsection focuses on the additional constraints and the specific problem formulation of the C_2 controller. Within the charging hub, each EV must reach its desired SOC within a predefined time window, as specified by the EV driver. This duration must be at least as long as the minimum time required to achieve the target SOC under continuous charging. Since the charging and discharging actions of all EVs are planned over the prediction horizon, a terminal constraint is included in the design to ensure that the desired SOC is ultimately reached:

$$x_j(k) \geq x_{T,j}, \quad \text{if } t_i + kT_s \geq t_{f,j} \tag{4.12}$$

$x_{T,j}$ is the desired terminal SOC for the j-th EV. Also, t_i is the beginning time of the prediction horizon, and $t_{f,j}$ is the final time that the j-th EV must be charged. To provide ancillary services and participate in DR programs, the energy consumption of the charging hub must be constrained according to set-point requests from the grid operator. The targeted demands correspond to mid-term demands from the intraday market and are defined as energy consumption reduction. This modeling approach is justified because the controller operates with prediction horizons spanning several minutes, enabling it to respond to mid-term demands effectively. Specifically, these demands are formulated to limit the total energy consumption of the charging hub over a defined time interval:

$$\kappa \cdot T_s \sum_{k=1}^{N} \left(p_b(k) + \sum_{j=1}^{n} p_j(k)\right) \leq e_{\max} \tag{4.13}$$

p_b is the power consumption of the buffer, p_j is the power consumption of the j-th EV, and $e_{\max}$ is the maximum allowed energy consumption. The controller is developed based on a moving-horizon strategy where it computes the optimal control sequences from $k = 1$ to $k = N$ by solving the optimization problem and applies only the first control at $k = 1$. Then, the whole computation is repeated at the next sampling time when the feedback is received from the storage systems. The whole optimization problem of the proposed MPC for the charging hub is expressed as follows:

$$(\text{OCP } 1): \begin{cases} \max\limits_{\Delta_k, P_{cs,k}} & J_N \\ \text{s.t.:} & \text{Eq. (4.4), (4.9)} - \text{(4.13)}, \\ & 0.1 \leq x_k \leq 0.9, \quad k = 1, \cdots, N, \end{cases} \tag{4.14}$$

where:

$$\begin{cases} \Delta_k = [\delta_{1b}(k)\ \delta_{2b}(k)\ \delta_{11}(k)\ \delta_{21}(k) \cdots \delta_{1n}(k)\ \delta_{2n}(k)]^T, \\ P_{cs,k} = [P_{1b}(k)\ P_{2b}(k)\ P_{11}(k)\ P_{21}(k) \cdots P_{1n}(k)\ P_{2n}(k)]^T, \\ x_k = [x_b(k)\ x_1(k) \cdots x_n(k)]^T \end{cases} \tag{4.15}$$

The last inequality in Eq. (4.14) assures safe and reliable battery usage. The whole optimization problem is a mixed-integer bilinear program (MIBP) with $4(n+1)N$ decision variables. However, it is possible to replace the bilinear terms in Eq. (4.9) with linear terms to simplify the optimization problem. By introducing auxiliary

variables:

$$\begin{cases} z_{1m} := \delta_{1m} P_{1m}, \\ z_{2m} := \delta_{2m} P_{2m} \end{cases} \tag{4.16}$$

the following constraints can be added to the problem to remove bilinear terms [83]:

$$\begin{cases} z_{1m} \leq \delta_{1m} \bar{P}_{1m}, \\ z_{1m} \geq 0, \\ z_{1m} \leq P_{1m}, \\ z_{1m} \geq P_{1m} - (1 - \delta_{1m}) \bar{P}_{1m}, \end{cases} \tag{4.17}$$

$$\begin{cases} z_{2m} \leq 0, \\ z_{2m} \geq \delta_{2m} \underline{P}_{2m}, \\ z_{2m} \leq P_{2m} - (1 - \delta_{2m}) \underline{P}_{2m}, \\ z_{2m} \geq P_{2m} \end{cases} \tag{4.18}$$

where $\bar{P}_{1m} > 0$, the upper bound of the charging power, and $\underline{P}_{2m} < 0$, the lower bound of the discharging power, are defined as follows:

$$\begin{cases} \bar{P}_{1j} = f_j(x_j), \\ \underline{P}_{2j} = P_{j,\min}, \\ \bar{P}_{1b} = f_b(x_b), \\ \underline{P}_{2b} = P_{b,\min}, \\ \bar{P}_{1g} = f_b(x_b) + \sum_{j=1}^{n} f_j(x_j), \\ \underline{P}_{2g} = P_{b,\min} + \sum_{j=1}^{n} P_{j,\min} \end{cases} \tag{4.19}$$

Eq. (4.8) and Eq. (4.9) can be reformulated in terms of auxiliary variables z_*:

$$J_N^z = \kappa \cdot T_s \sum_{k=1}^{N} \sum_{m} [z_{1m}(k) W_{1m}(k) + z_{2m}(k) W_{2m}(k)] \tag{4.20}$$

$$\begin{cases} x_b(k+1) = x_b(k) + \dfrac{T_s \eta_b}{C_b} \left(z_{1b}(k) + z_{2b}(k) \right), & \text{(4.21a)} \\ \delta_{1b}(k) + \delta_{2b}(k) \leq 1, & \text{(4.21b)} \\ x_j(k+1) = x_j(k) + \dfrac{T_s \eta_j}{C_j} \left(z_{1j}(k) + z_{2j}(k) \right), & \text{(4.21c)} \\ \delta_{1j}(k) + \delta_{2j}(k) \leq 1, \quad k = 1, \cdots, N & \text{(4.21d)} \end{cases}$$

Accordingly, the optimal control problem (OCP 1) can be formulated in terms of x_*, z_*, and P_* as an MILP:

$$(\text{OCP 2}) : \begin{cases} \max\limits_{\Delta_k, P_{cs,k}, Z_{cs,k}} & J_N^z \\ \text{s.t.:} & \text{Eq. (4.11)} - \text{(4.13)}, \text{(4.17)}, \text{(4.18)}, \text{(4.19)}, \text{(4.21)} \\ & 0.1 \leq x_k \leq 0.9, \quad k = 1, \cdots, N \end{cases} \tag{4.22}$$

where $Z_{cs,k} = [z_{1b}(k)\ z_{2b}(k)\ z_{11}(k)\ z_{21}(k) \cdots z_{1n}(k)\ z_{2n}(k)]^T$. Since the first sample of control sequences ($P_{cs,1}$) is applied to the system and the calculation is repeated in

the next time step, the "economic benefit" of the charging hub from the transactions is J_1^z. This equation concludes the single-layer optimal control (C_2) synthesis, and the double-layer MPC (C_3)is explained in the next subsection.

4.3.2 Double-layer Controller (C_3)

The single-layer controller (C_2) involves $6(n+1)N$ independent decision variables, leading to a significant increase in computational burden as the number of EVs (n) grows. However, this is not a problem for C_2, while it has a large sampling time, and performing the calculations in its current structure is reasonable. The C_3 controller, on the other hand, must have a smaller sample time to capture primary demands, therefore, it cannot be simply designed as the C_2 controller is designed, and a different structure should be considered in its synthesis. One approach to mitigate this complexity is to aggregate all EVs into a single virtual EV, which collectively represents the total energy storage capacity of the individual EVs. As a result, there are $12N$ independent decision variables if the same approach of single-layer MPC is followed. Having the power of the virtual EV, a static optimization with $6n$ decision variables at a different layer of computation finds the optimal power of all EVs at each sample time. So, the total decision variables is $12N$+$6n$. Fig. 4.5 demonstrates the virtual structure of the charging hub with an accumulated EV (AEV). The capacity of the AEV is the sum of the capacity of all EVs:

$$C_{AEV} = C_1 + C_2 + \cdots + C_n \tag{4.23}$$

In turn, the SOC of the AEV (x_{AEV}) is calculated as follows:

$$x_{\text{AEV}}(t) - x_{\text{AEV}}(t_0) = \frac{\eta}{C_{AEV}} \int_{t_0}^{t} P_c(\tau)\, d\tau \tag{4.24}$$

In the virtual representation of the charging hub, only two energy storage systems are considered, simplifying the optimization problem. The core idea behind the double-layer controller is to compute the charging/discharging power P_c for the aggregated EV (AEV) in the upper layer, as depicted in Fig. 4.5, and then distribute

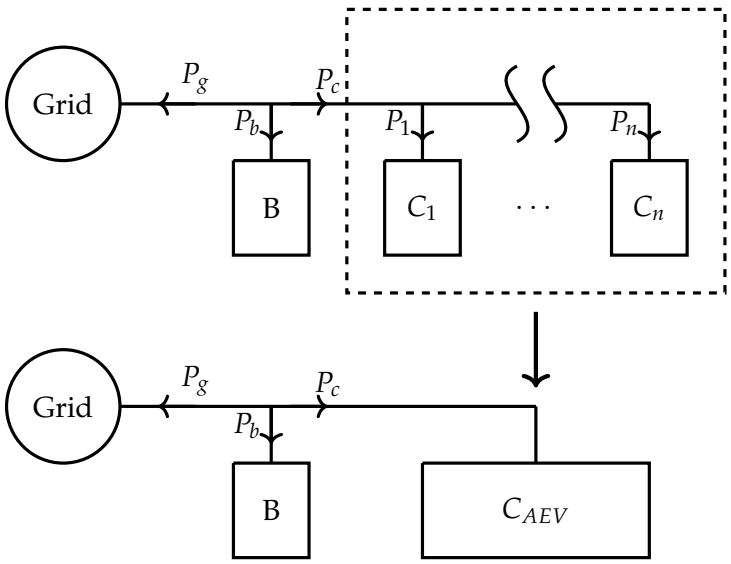


FIGURE 4.5: Restructure the charging hub for MPC in the upper layer.

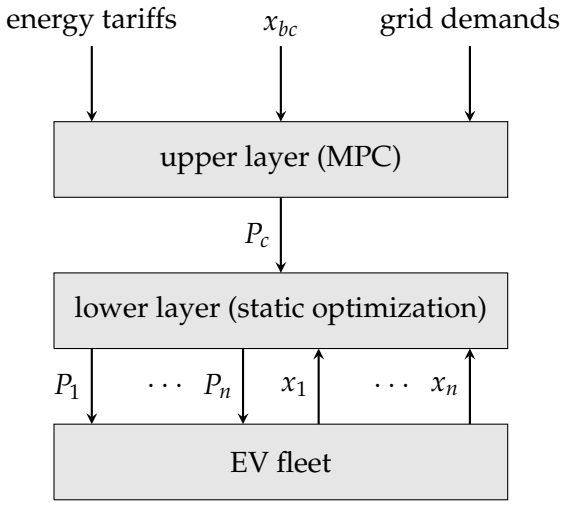


FIGURE 4.6: The proposed C_3 control structure for the hub.

this power among individual EVs in the lower layer. Specifically, the upper layer solves a reduced optimization problem, as defined in Eq. (4.14), involving fewer decision variables, and transmits the optimal P_c to the lower layer. The lower layer then determines the individual charging powers P_j for each EV by solving a separate static optimization problem. Fig. 4.6 illustrates the double-layer control architecture of C_3, where the state feedback from the charging hub to the upper layer is given by $x_{bc} = [x_b \; x_{\text{AEV}}]^T$.

The Upper Layer

The controller in the upper layer also seeks to maximize the economic benefit of the charging hub, considering energy tariffs and primary demand signals issued by the grid operator and priced in the balancing market. As previously noted, aggregating all EVs into an AEV significantly reduces the number of decision variables and alleviates the computational burden. Referring to Eq. (4.22), the optimization problem in the upper layer is formulated similarly, with $n = 1$. A key distinction is that the double-layer controller (C_3) operates at a higher time resolution and with a shorter prediction horizon than the single-layer MPC (C_2). Consequently, it can respond only to primary demands, which are defined in terms of instantaneous power reduction. As a result, the constraint in Eq. (4.13) in OCP 2 must be replaced with the following condition:

$$p_b(k) + p_c \leq p_{\max}(k), \; \forall k \in \{1, \cdots, N\} \tag{4.25}$$

where $p_{\max}$ is the maximum power consumption requested by the grid. The total number of decision variables in this layer is $12N$, significantly reducing computational complexity compared to the original formulation when n is large. The prediction horizon of C_2 spans several minutes, allowing the controller to enforce the terminal SOC constraint in Eq. (4.12) be satisfied for all EVs. In contrast, C_3 operates on a much shorter time scale—on the order of seconds—making it unsuitable for capturing the terminal SOC requirements of individual EVs. Therefore, the problem

formulation of the upper layer is:

$$(\text{OCP 3}): \begin{cases} \max\limits_{\Delta_k, P_{cs,k}, Z_{cs,k}} & J_N^z \\ \text{s.t.:} & \text{Eq. (4.11), (4.17), (4.18), (4.19), (4.21), (4.25),} \\ & 0.1 \leq x_k \leq 0.9, \quad k = 1, \cdots, N \end{cases} \tag{4.26}$$

The Lower Layer

The concept of aggregating all EVs into a single virtual EV simplifies the optimization problem in the upper layer, which is responsible for determining the optimal AEV power P_c across all time steps in the prediction horizon. Specifically, the upper layer computes $P_b(k)$ and $P_c(k)$ for $k = 1$ to $k = N$, and transmits the first sample $P_c(1)$ to the lower layer. The lower layer then allocates this power among individual EVs by determining P_j at $k = 1$, based on the charging priorities of the EVs. The corresponding optimization problem for this layer is formulated as follows:

$$(\text{OCP 4}): \begin{cases} \max\limits_{\tilde{\Delta}_1, P_{cc,1}, Z_{cc,1}} & \sum\limits_{j=1}^{n} \alpha_j x_j(2) \\ \text{s.t.:} & \sum\limits_{j=1}^{n} P_j(1) \leq P_c(1), \\ & x_j(2) = x_j(1) + \frac{T_s \eta_j}{C_j}(z_{1j}(1) + z_{2j}(1)), \\ & \delta_{1j}(1) + \delta_{2j}(1) \leq 1, \\ & \text{Eq. (4.17), (4.18) with } m = j, \\ & 0.1 \leq x_j(2) \leq 0.9, \end{cases} \tag{4.27}$$

where:

$$\begin{cases} \tilde{\Delta}_1 = [\delta_{11}(1)\ \delta_{21}(1) \cdots \delta_{1n}(1)\ \delta_{2n}(1)]^T, \\ P_{cc,1} = [P_{11}(1)\ P_{12}(1) \cdots P_{1n}(1)\ P_{2n}(1)]^T, \\ Z_{cc,1} = [z_{11}(1)\ z_{12}(1) \cdots z_{1n}(1)\ z_{2n}(1)]^T \end{cases} \tag{4.28}$$

The coefficients α_j are weighting factors selected by the designer to prioritize EVs that require faster charging. The lower layer optimization is solved at the current time step ($k = 1$) and includes $6n$ independent decision variables. By implementing the double-layer controller, the total number of decision variables is reduced to $12N + 6n$, significantly less than $6(n+1)N$ when n is large. However, a significant limitation of the C_2 controller is its inability to guarantee satisfaction of the terminal SOC constraint in Eq. (4.12), since the distribution of P_c is performed only for the current step without a prediction horizon. The lower-layer problem is therefore a static optimization, where the controller aims to allocate the optimal power to EVs with higher charging priorities. The economic benefit of the charging hub in this layer is defined analogously to the C_2 controller and corresponds to J_1^z.

The proposed double-layer MPC (C_3) is an efficient strategy to solve the optimal energy flow problem in the charging hub when the prediction horizon is short. In the case of primary demands, this architecture can be exploited to reduce computational complexity when the number of EVs is large. The planning for reaching a certain level of SOC for the EVs with time might deviate; however, this deviation for a short time (in the order of seconds) is compensable. Fig. 4.7 demonstrates a comparison between two different structures of the controller C_3. It shows the difference between a C_3 with a single-layer structure and a C_3 with a double-layer structure.

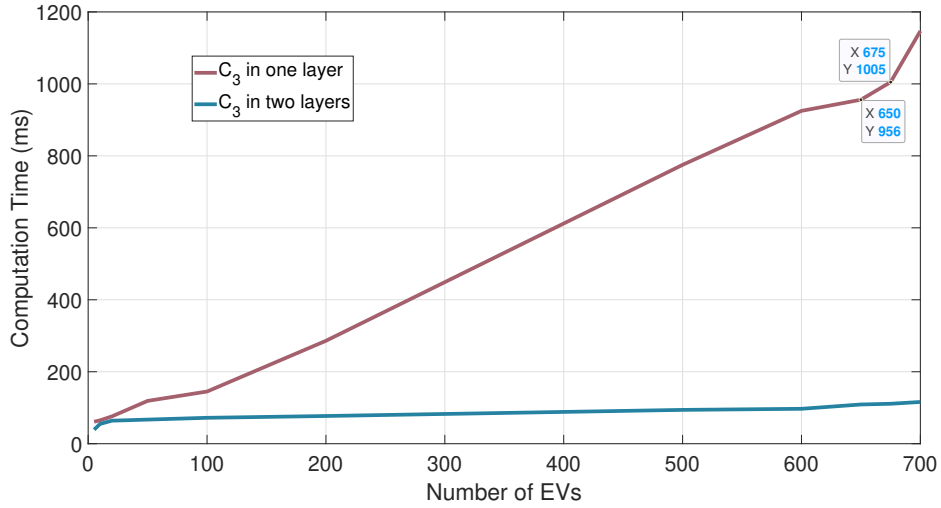


FIGURE 4.7: The comparison of a single-layer architecture and a double-layer architecture of the controller C_3 in terms of average computation time of the solver. The number of EVs starts from 5 and increases up to 700. A single-layer C_3 cannot provide a real-time solution for 675 EVs or more. The sample time is chosen 1 second.

If this controller is designed in one-layer similarly to the same approach that the C_2 controller is designed, by increasing the number of EVs and assuming the sample time of 1 second, it is not able to terminate the computation of the optimization problem on time (in less than 1 second). As shown in Fig. 4.7, the computation time of the proposed double-layer C_3 controller slightly increases with the number of EVs, while the computation time of a single-layer C_3 controller rises considerably with the number of EVs.

The charging hub can meet the demands originating from the grid using the flexibility provided by the EVs and the energy buffer. As mentioned in the design procedure, the C_3 controller targets primary demands to support the grid and stabilize the frequency when there is an imbalance between power generation and consumption in the grid. When the charging hub reduces its consumption or discharges the storage systems, it supports the grid to avoid frequency instability. The parametrization of primary demands for double-layer MPC is formulated using a model used to study synchronization phenomena in coupled oscillator systems, known as the Kuramoto model [116], [117]. If the power generation is reduced from P_g to $P_g - \Delta P_g$ due to a fault or any other factor that affects the power generation of the grid, this deficiency is sent to the consumers as a demand. Assuming that δp_{ch} is the amount of power that the charging hub reduces, then γ_{ch}, the contribution factor of the charging hub for primary demands, is $\frac{\delta p_{ch}}{\Delta P_g}$. In the case of load shifting that the proposed single-layer controller (C_2) has targeted, a similar strategy is adopted, and the contribution factor of the charging hub for load shifting (α_{ch}) is $\frac{\delta e_{ch}}{\Delta E_g}$. δe_{ch} is the amount of energy that the charging hub shifts to the off-peak time, and ΔE_g is the amount of energy required to be shifted by all the consumers.

4.3.3 Impact of Battery Power Constraints on the MPCs' Formulation

As mentioned earlier, EV manufacturers provide a charging and discharging regime for low-loss energy transfer between the EV and the grid. These regimes have various changes with SOC, which highly depend on the EV parameters and the manufacturer's strategy. The charging and discharging characteristics can be constant, linear, piecewise linear, or nonlinear. The charging or discharging optimal power (from the view of the manufacturer) is set based on the current value of SOC. In this regard, if we assume that the charging and discharging powers are constant, the control problem of the batteries is simplified to specifying the mode of the battery, ignoring the value of the charging or discharging power. This will simplify the problem settings of the MPC to a pure integer programming problem, which can be solved with less computational burden. Since the charging/discharging powers are known, we can write Eq. 4.9 as follows:

$$\begin{cases} x_b(k+1) = x_b(k) + \dfrac{T_s \eta_b}{C_b} \left(\delta_{1b}(k) P_{1b} + \delta_{2b}(k) P_{2b}\right), & \text{(4.29a)} \\ \delta_{1b}(k) + \delta_{2b}(k) \leq 1, & \text{(4.29b)} \\ x_j(k+1) = x_j(k) + \dfrac{T_s \eta_j}{C_j} \left(\delta_{1j}(k) P_{1j} + \delta_{2j}(k) P_{2j}\right), & \text{(4.29c)} \\ \delta_{1j}(k) + \delta_{2j}(k) \leq 1, \quad k = 1, \cdots, N, \quad j = 1, \cdots, n & \text{(4.29d)} \end{cases}$$

where P_{1b}, P_{2b}, P_{1j} and P_{2j} are constant. Therefore, the relaxation constraints are no longer required because the model does not contain a bilinear term. The only decision variables in this problem setting are δ_{1*}, δ_{2*}. The single-layer optimization problem or the double-layer optimization problem can be reformulated with the new assumption. For single-layer MPC, the optimization problem is:

$$\text{(OCP 5)} : \begin{cases} \max\limits_{\Delta_k} & J_N^{'} \\ \text{s.t.:} & \text{Eq. (4.11)} - \text{(4.13)}, \text{(4.29)}, \\ & 0.1 \leq x_k \leq 0.9, \quad k = 1, \cdots, N, \end{cases} \tag{4.30}$$

where:

$$J_N^{'} = \kappa \cdot T_s \sum_{k=1}^{N} \sum_{m} [P_{1m} \delta_{1m}(k) W_{1m}(k) + P_{2m} \delta_{2m}(k) W_{2m}(k)] \tag{4.31}$$

Δ_k is the vector of binary decision variables. A similar approach for double-layer MPC leads to the following problem formulation:

$$\text{(OCP 6 - Upper Layer)} : \begin{cases} \max\limits_{\Delta_k} & J_N^{'} \\ \text{s.t.:} & \text{Eq. (4.25)}, \text{(4.29)}, \\ & 0.1 \leq x_k \leq 0.9, \quad k = 1, \cdots, N \end{cases} \tag{4.32}$$

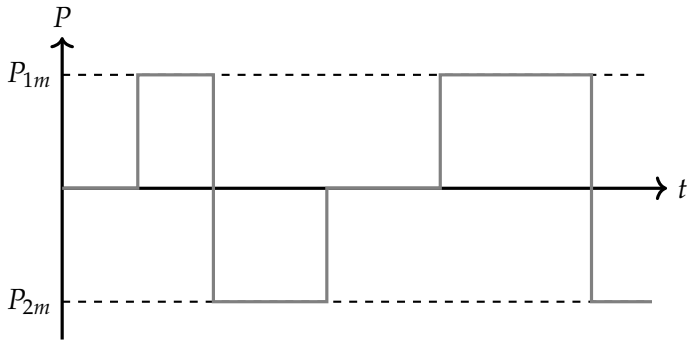


FIGURE 4.8: Power profile in binary MPC.

$$\text{(OCP 6 - Lower Layer)}: \begin{cases} \max\limits_{\delta_{1j},\delta_{2j}} & \sum\limits_{j=1}^{n} \alpha_j x_j(2) \\ \text{s.t.:} & \sum\limits_{j=1}^{n} \left(\delta_{1j}(1)P_{1j} + \delta_{2j}(1)P_{2j}\right) \leq P_c(1), \\ & x_j(2) = x_j(1) + \frac{T_s \eta_j}{C_j}(\delta_{1j}(1)P_{1j} + \delta_{2j}(1)P_{2j}), \\ & \delta_{1j}(1) + \delta_{2j}(1) \leq 1, \\ & 0.1 \leq x_j(2) \leq 0.9 \end{cases} \tag{4.33}$$

where P_{1j} and P_{2j} are constant. At both upper and lower layers, the only decision variables are binary Δ_k, δ_{1j}, and δ_{2j}. Fig. 4.8 depicts the three main levels of power that are sent to the battery by a binary MPC. The power profile fluctuates only between these levels, depending on the values of δ_{1*} and δ_{2*}.

Referring to the initial explanation in Subsection 4.3.3, the upper bound of the absolute value of the charging and discharging power of the EVs can be a second-order polynomial (quadratic) function of SOC. More precisely, a second-order polynomial function with maximum value at a specific value (possibly minimum) of SOC could be an approximation of some EV charging/discharging characteristics. Also, these second-order characteristics can be considered as upper bounds for the absolute value of charging/discharging powers. Therefore, we can write the charging/discharging powers as follows:

$$0 < P_{1m} \leq \alpha_{1m}(x_m - x_{1m})^2 + \beta_{1m} \tag{4.34a}$$

$$0 > P_{2m} \geq \alpha_{2m}(x_m - x_{2m})^2 + \beta_{2m} \tag{4.34b}$$

where α_{1m}, α_{2m}, β_{1m}, and β_{2m} are constant. x_{1m} and x_{2m} are the extremum points of the limiting curves. Regardless of whether we use relaxation techniques for the optimization problem, the Eq. 4.34 introduces a quadratic state-dependent constraint in our problem settings, replacing the Eq. 4.10. It is straightforward to use Eq. 4.34 to reformulate MPC again for single-layer and double-layer structures, while it is similar to the approach for binary MPC. The simulation results of this setup will be demonstrated in the next section.

4.4 Simulation Results

This section outlines the simulation setup used to evaluate the performance of the proposed control strategies. All simulations are carried out in MATLAB© R2022b, utilizing the YALMIP toolbox for optimization problem formulation and solution, with Gurobi© as the solver. The computations are executed on a machine equipped with an Intel Core i7 processor and 32 GB of RAM, ensuring efficient execution of the optimization routines. Key simulation settings include a prediction horizon of 100 minutes and a sampling time (T_s) of 5 minutes for the single-layer MPC (C_2), while the double-layer MPC (C_3) operates with a 30 seconds prediction horizon and a sampling time of 1 second. The sampling time and prediction horizon of the proposed C_3 controller are selected to reflect the requirements of fast ancillary services, in particular, primary DR. Such a temporal resolution is necessary to ensure accurate tracking of rapid power setpoint variations associated with primary demand activation, yielding a closer frequency nadir to nominal frequency [118], [119]. Additional simulation parameters are summarized in Table 4.1. The battery capacities of the EVs are selected based on realistic values from the current EV market [120]. Furthermore, the buffer capacity is set to 2.3 times the total capacity of all EVs to ensure sufficient flexibility not only to supply energy to the EVs but also to sell energy back to the grid when beneficial, simultaneously. In the case of the single-layer MPC, different departure times are assigned to each group of EVs, while in the double-layer MPC, weighting factors are introduced to observe the effect of prioritization. The control structures described in the previous sections are tested to confirm that they yield optimal solutions when operating in either energy trading mode or under DR conditions. The simulated charging hub comprises an energy buffer and five EV groups (each containing 10 EVs, totaling 50 EVs) capable of bidirectional charging across four simulation scenarios. It is important to note that the proposed control strategies are not restricted to this scale and can be extended to accommodate a larger number of EVs. The purpose of this section is to demonstrate the application of the methods introduced in this chapter. The charging limit profiles for each EV group, as well as the energy tariff profiles, are depicted in Fig. 4.9. The discharging limits are assumed to be the negative counterparts of the charging limits, although alternative profiles may also be defined. Specifically, EVs numbered 1 to 10 belong to Group 1, EVs 11 to 20 to Group 2, and so on, up to EV 50. The electricity and service prices considered in the simulations are inspired by representative values reported in European energy markets and are chosen to reflect typical magnitudes and relative price dynamics rather than exact market forecasts (see e.g., [121]). For time $t \leq 55$ minutes, the prices are chosen such that they result in charging the batteries, and for the rest of the time they are selected such that the batteries discharge to the grid. The benefit margin (considering the usage fee of the charging hub and degradation fee) has been considered to specify the prices for the EVs and the buffer. Looking to the lower subplot of 4.9 at $t \leq 55$ minutes, W_{2g} is less than W_{1b} and W_{1j}, denoting the fact that the energy is bought from the grid with a cheaper price and is sold to the EVs or stored in the buffer with a higher price. The difference comes from the benefit margin and usage fee of the installation of the charging hub. For $t \geq 55$ minutes, W_{2b} and W_{2j} are less than W_{1g}, implying that the charging hub operator buys energy from the EVs and sells it with the energy currently stored in the buffer to the grid with a higher price, considering usage fee, degradation fee, maintenance fee of the buffer and the benefit margin.

TABLE 4.1: **Simulation parameters of the EVs and corresponding demands.**

Parameter	Value (unit)	Parameter	Value (unit)
$\eta_b^{'}$	1	$t_{f,j}$ (Group 1)	90 (min)
$\eta_j^{'}$ (for all j)	1	$t_{f,j}$ (Group 2)	90 (min)
C_b	6944.4 (A.h)	$t_{f,j}$ (Group 3)	80 (min)
C_j (Group 1)	69.44 (A.h)	$t_{f,j}$ (Group 4)	100 (min)
C_j (Group 2)	52.08 (A.h)	$t_{f,j}$ (Group 5)	95 (min)
C_j (Group 3)	55.55 (A.h)	$p_{\max}$ (first demand)	-3100 (kW)
C_j (Group 4)	65.27 (A.h)	$p_{\max}$ (second demand)	-1890 (kW)
C_j (Group 5)	62.50 (A.h)	α_j (Group 1)	9
$\bar{P}_{1b}$	1750 (kW)	α_j (Group 2)	2
$\bar{P}_{1j}$ (Group 1)	35.0 (kW)	α_j (Group 3)	5
$\bar{P}_{1j}$ (Group 2)	30.8 (kW)	α_j (Group 4)	4
$\bar{P}_{1j}$ (Group 3)	31.2 (kW)	α_j (Group 5)	3
$\bar{P}_{1j}$ (Group 4)	28.0 (kW)	$\underline{P}_{2j}$ (Group 1)	-35.0 (kW)
$\bar{P}_{1j}$ (Group 5)	36.0 (kW)	$\underline{P}_{2j}$ (Group 2)	-30.8 (kW)
V	400 (V)	$\underline{P}_{2j}$ (Group 3)	-31.2 (kW)
N	20	$\underline{P}_{2j}$ (Group 4)	-28.0 (kW)
$\underline{P}_{2b}$	-1750 (kW)	$\underline{P}_{2j}$ (Group 5)	-36.0 (kW)

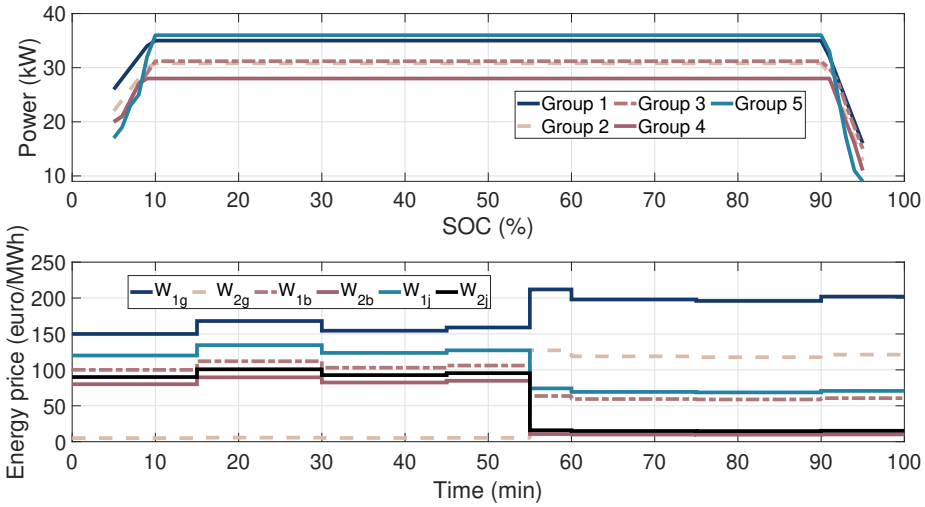


FIGURE 4.9: EVs' charging power limits (upper subplot) and energy tariffs (lower subplot).

4.4.1 Scenario I: One-layer MPC (C_2)

This scenario focuses on analyzing the performance of the C_2 controller when the charging hub simultaneously engages in energy trading and participates in DR programs. The simulation outcomes are illustrated in Fig. 4.10, where the SOC of selected EVs, the total energy stored in all storage systems, and the economic benefit of the charging hub are shown. The energy prices for buying from and selling to the grid, as well as the EV charging tariffs and the load-shifting demand signals, are

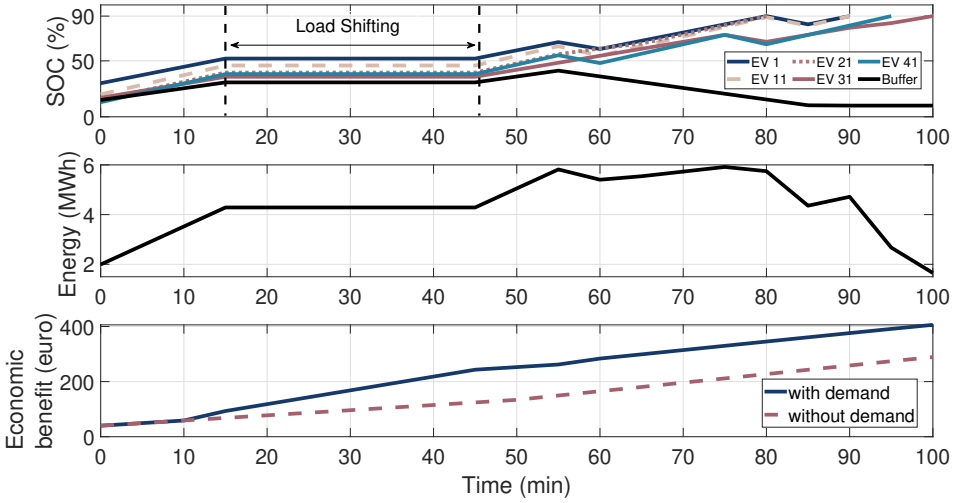


FIGURE 4.10: Simulation results of scenario I. The charging hub's economic benefit is greater when it participates in DR, as shown in the third subplot.

based on data from the daily German energy market [122]. The energy tariffs W_{1m} and W_{2m}, which change every 15 minutes in the spot market, incentivize the MPC to charge both the buffer and the EVs from $t = 0$ min to $t = 55$ min. In response to a grid-side demand to defer energy consumption from $t = 15$ min to $t = 45$ min, the MPC halts all battery charging during that interval. Charging resumes from $t = 45$ min to $t = 55$ min, and due to the energy price change at $t = 55$ min, some energy storage systems begin discharging, while some continue to charge. Nevertheless, the EVs must eventually reach their required SOCs as specified by user preferences. Consequently, the controller resumes charging the EVs up to 90% to satisfy the constraint in Eq. (4.12). As shown in Fig. 4.10, the EVs follow distinct charging patterns, with some requiring earlier charging than others. Assuming a grid-side demand of $\Delta E_g = 10$ MWh, the charging hub successfully shifts $\delta e_{ch} = 3.849$ MWh of energy to off-peak periods, yielding a flexibility contribution factor of $\alpha_{ch} = 0.3849$. The second subplot of Fig. 4.10 shows the energy stored in the storage systems (the buffer and EVs) while the third subplot illustrates the economic benefit of the charging hub for energy transactions and DR participation. In the first trading phase, the benefit increases until the demand for load shifting is issued. During the demand activation interval ($t = 15$ min to $t = 45$ min), the benefit grows faster due to the reward given by the grid for demand satisfaction. The amount of the reward is considered equal to 77.94 € for each MW per hour load shifting to reflect typical magnitudes as [106]. The increment in the benefit continues with a lower slope from $t = 45$ min to $t = 55$ min. The critical change of tariffs at $t = 55$ min results in discharging the EVs and the buffer, also the benefit rises with a slightly greater slope. After 5 minutes, all the EVs charge to satisfy the final SOC constraint; though it is not the optimal direction of energy flow, it is still beneficial for the charging hub. The buffer charges and discharges according to the pricing of energy, while contributing to ancillary service programs during the load-shifting. Since it does not have any SOC target, it discharges to the minimum value of SOC.

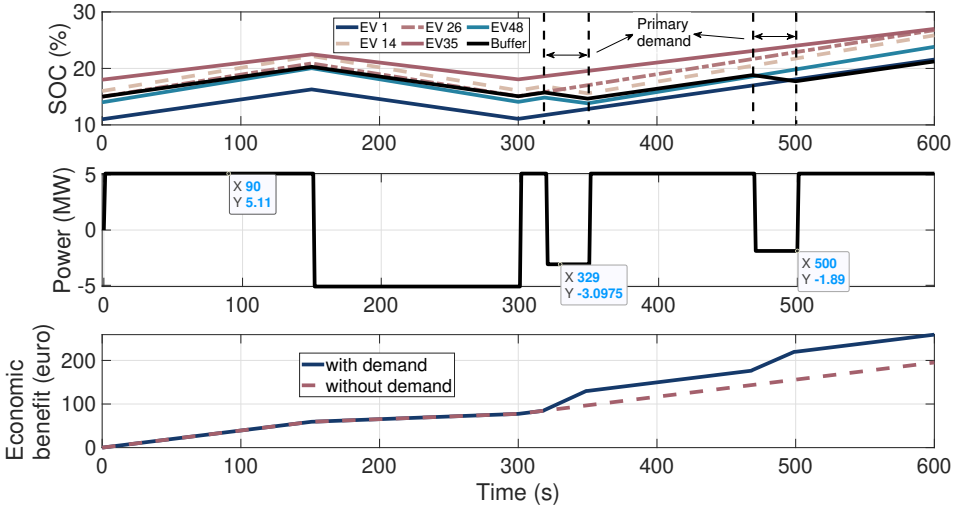


FIGURE 4.11: Simulation results of scenario II. The charging hub's economic benefit is greater when it participates in DR, as shown in the third subplot.

4.4.2 Scenario II: Double-layer MPC (C_3)

The objective of this simulation is to evaluate the behavior of the double-layer MPC when the charging hub receives primary requests during ongoing energy transactions. The prediction horizon is considered $N = 30$. The energy tariffs are modified twice throughout the simulation, resulting in three distinct operational phases, as illustrated in Fig. 4.11. In the first phase, from $t = 0\,\text{s}$, the controller begins charging to maximize the economic benefit of the charging hub. At $t = 150\,\text{s}$, the tariff changes prompt the controller to discharge the batteries until $t = 300\,\text{s}$, after which the new tariffs again incentivize charging. Two primary DR requests from the grid are issued between $t = 300\,\text{s}$ and the end of the simulation. The first DR event is active from $t = 321\,\text{s}$ to $t = 350\,\text{s}$, and the second from $t = 471\,\text{s}$ to $t = 500\,\text{s}$. During these intervals, the total power consumption is reduced to $-3.098\,\text{MW}$ and $-1.890\,\text{MW}$, respectively, corresponding to load-shifting contributions of $\delta p_{ch} = 8.208\,\text{MW}$ and $\delta p_{ch} = 7.000\,\text{MW}$, as shown in Fig. 4.11, subplot 2. Battery responses to these demands vary due to their assigned weighting coefficients α_j (Fig. 4.11, subplot 1). During the first DR event, EVs 1, 26, and 35 continue charging, while EVs 14 and 48, along with the energy buffer, discharge due to the lower priority (smaller α) of Groups 2 and 5. In contrast, during the second DR event, only the buffer discharges, while all EVs continue charging. The third subplot of Fig. 4.11 displays the economic benefit of the charging hub over time. The profit curve steadily increases, as the hub continually participates in either energy trading or DR. Notably, two distinct jumps in the curve reflect the financial rewards from DR participation. It is assumed that each DR event provides a reward of 40 € per MW of power reduction, per demand, given each event duration is 30 seconds. The grid's requested power reduction is $\Delta P_g = 80\,\text{MW}$ in each instance, and the charging hub's contribution factors are $\gamma_{ch} = 0.103$ and $\gamma_{ch} = 0.088$ for the first and second DR events, respectively.

It is always assumed that the EV drivers also benefit from the flexibility provided by their EVs for energy transactions and DR service provision if the EVs have been discharged during the parking time. Therefore, by the end of the charging period,

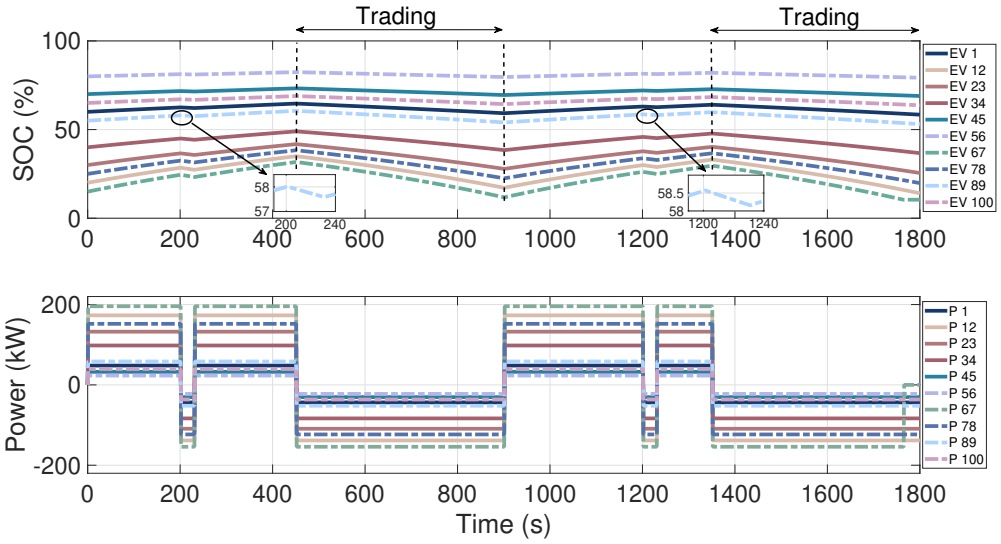


FIGURE 4.12: Simulation results of Scenario III. The energy tariffs change three times, leading to four main charging/discharging phases.

when the drivers want to depart, a portion of the revenue gained by the charging hub could be given to the EV drivers whose cars have been used for the services. This can be legally binding through a contract between the charging hub and the EV drivers.

4.4.3 Scenario III: C_3 with Binary Battery Model

In this scenario, the goal is to investigate the performance of the proposed binary double-layer MPC for different sets of energy tariffs, as well as under demand circumstances. The simulation results are demonstrated in Fig. 4.12 where the SOC and power profiles for 100 EVs are plotted. The simulation interval includes four main sub-intervals and two short periods for primary demands. In the first phase, all the EVs, including the buffer, are charging because of the energy tariffs. More precisely, the optimal direction of energy flow is from the grid to the batteries. During the first phase, a primary demand is sent to the charging hub, which results in discharging the EVs and the buffer. The demand lasts for 30 seconds (from $t = 200\,\text{s}$ to $230\,\text{s}$), and the batteries continue to charge after the demand is removed. The initial value of the SOC of EVs is different, and they have different levels of power for charging and discharging. In the second phase, the optimal direction of energy flow reverses, with energy flowing from the EVs and the buffer towards the grid. Therefore, the SOCs decrease in this phase, and the charging hub owner trades energy with the grid, leading to an economic benefit. In the third and fourth phases, which are from $t = 900\,\text{s}$ to $1350\,\text{s}$ and from $t = 1350\,\text{s}$ to $1800\,\text{s}$, respectively, the same pattern is observed. The primary demand between $t = 1200\,\text{s}$ and $t = 1230\,\text{s}$ enforces the batteries to discharge to provide their flexibility to the grid. The decision-making process for this binary controller is limited to determining the mode of the batteries (charging/discharging/zero power), resulting in a maximum of three power levels (including zero) for each EV, as shown in Fig. 4.12.

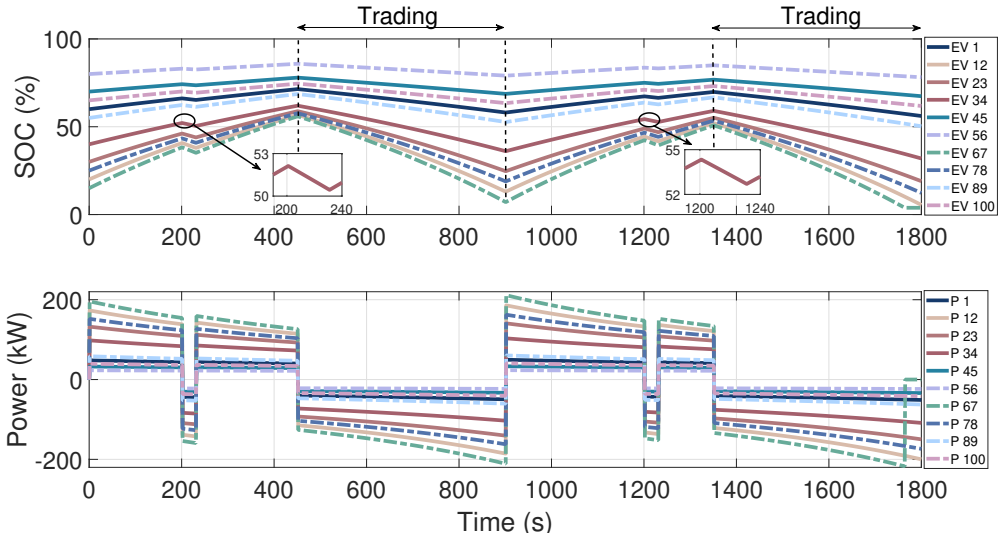


FIGURE 4.13: Simulation results of Scenario IV. The power of charging/discharging changes quadratically with the SOC of the EVs.

4.4.4 Scenario IV: C_3 with Quadratic State-dependent Power Constraints

It is also important to test the proposed double-layer MPC with quadratic constraints for charging and discharging powers. For this purpose, we have conducted a simulation for 100 EVs where the charging and discharging bounds are second-order functions similar to Eq. 4.34. The general form of these bounds is $\alpha_{1m}(x_m - 0.1)^2 + \beta_{1m}$ for charging and $\alpha_{2m}(x_m - 0.1)^2 + \beta_{2m}$ for discharging. We have chosen the same energy tariffs as in Scenario III to test the performance of the control algorithm. As a result, there are four main phases for the batteries, plus two short intervals with primary demands (Fig. 4.13). In the first phase, all EVs, including the energy buffer, are charging, as it is the optimal decision according to the energy tariffs. Specifically, the optimal energy flow is directed from the grid to the batteries. During this phase, a primary demand signal is sent to the charging hub, prompting the EVs and the buffer to discharge. This demand lasts for 30 seconds (from $t = 200$ to 230 s), after which the batteries resume charging. The EVs start with different initial SOC and possess varying charging and discharging power levels. In the second phase, the optimal energy flow reverses—energy is now supplied from the EVs and the buffer back to the grid. Consequently, SOC levels decrease during this period as the charging hub trades energy with the grid, generating economic gains. The third and fourth phases, spanning $t = 900$ s to 1350 s and $t = 1350$ s to 1800 s respectively, follow a similar pattern. A primary demand event between $t = 1200$ s and 1230 s triggers battery discharge, enabling the provision of flexibility services to the grid. As the lower subplot of Fig. 4.13 shows, the optimal charging/discharging power profiles are on their bound.

One of the main differences between the results of Scenarios III and IV is the profile of the SOCs' regimes. As shown in the upper subplots of Figures 4.12 and 4.13, the SOCs for Scenario III change linearly with time, while for Scenario IV they change nonlinearly with time. The reason is that in Scenario III, the SOCs are the integral of piece-wise constant functions, whereas in Scenario IV, the power profiles change nonlinearly with time. During the charging phases in Scenario IV, the charging power decreases for all EVs, resulting in a lower change in SOCs over

time. If there was no primary demand in Scenario III, since the intervals of charging and discharging are equal, the SOCs should have returned to their initial values. However, due to the demands, the final value of SOCs is less than the initial value. This reasoning does not necessarily hold for Scenario IV, because of nonlinear charging/discharging regimes.

4.5 Conclusion

Bidirectional charging enables the integration of the substantial battery capacities of EVs into grid services. This chapter has presented a case study of an EV charging hub enhanced with stationary energy storage systems and introduces a method that merges multi-asset optimization optimization with advanced control theory. The proposed algorithmic platforms serve two main objectives: facilitating the autonomous participation of future EV hubs and their connected vehicles in energy markets, and enabling the provision of ancillary services in both spot and balancing markets. A comparison of the two main control methods developed in this chapter has been explained in Table. 4.2.

TABLE 4.2: **Comparison of features of the two main proposed controllers.**

Controller	Features
C_2	• Energy trading • Mid-term ancillary services • Target SOC at departure time • Higher sample time and prediction horizon
C_3	• Energy trading • Primary DR • No target SOC • Lower sample time and prediction horizon

Simulation results validate the effectiveness and feasibility of the proposed strategies within modern energy systems. These findings represent a significant step toward adaptive, autonomous energy management for EV charging hubs equipped with bidirectional charging, particularly in control energy markets. Importantly, the optimization schemes can be integrated into the operational and dispatch frameworks already used by distributed system operators. The hierarchical control architecture is designed for scalability and demonstrates low computational complexity, making it suitable for real-time implementation, even with a growing number of EVs. Looking forward, a promising research direction lies in embedding these control strategies into a unified hierarchical framework capable of operating across both short and long time scales, in alignment with IEEE PES standards for control energy. In this framework, the single-layer energy control algorithm would manage broader

time intervals—such as 15-minute sampling—for load shifting and mid-term market participation. Meanwhile, the double-layer power control algorithm would respond to primary grid demands in real time, with a temporal resolution of several seconds. Furthermore, these concepts can be extended beyond EV charging hubs to support sector coupling and be applied to other domains, such as industrial energy systems and large-scale residential networks. The research of this chapter has been conducted in the context of the DeRIVE project, funded by the German Federal Ministry of Transport (BMV), from 2022 to 2025, under Grant 01MV22014B.

Chapter 5

Hierarchical Energy Management for Ancillary Services

5.1 Introduction to Methodology

In this chapter, we propose a general three-level hierarchical MPC framework designed to manage energy consumption and storage systems in a flexible and optimized manner. However, the primary focus is on its second and third layer development. Therefore, the key innovation is separating decision-making into two coordinated layers: a higher-level strategic planner and a lower-level operational controller. This architecture enables long-horizon planning for energy transactions and load shifting, while maintaining high-resolution control over system dynamics. The flexibility and scalability of the proposed framework are demonstrated through two application scenarios:

- A bottle production plant, where machines with binary and continuous operation modes and integrated batteries provide process flexibility to support DR events without significantly impacting production output.
- A bidirectional EV charging hub, where EVs and an energy buffer participate in energy markets by shifting load and trading energy with the grid.

At the higher level, an optimization problem is solved at discrete intervals (e.g., every 15 minutes), assigning energy packets to the different assets based on price signals, system constraints, and grid conditions. The lower-level controller then ensures the delivery of these energy packets by solving a quadratic optimization problem with linear constraints over a finer time resolution, ensuring feasibility and compliance with operational limits. By decoupling long-term planning from real-time control, the framework provides a unified, robust, and efficient solution to managing diverse energy assets across multiple domains. Unlike traditional short-sighted control approaches, the hierarchical structure explicitly exploits energy flexibility over longer time horizons, allowing for dynamic and market-responsive energy management.

5.2 Generic Scheme

In current electricity markets, the day-ahead market serves as the main platform for scheduling electricity generation and consumption for the following day. Market participants submit bids based on forecasts, and the market clears by matching supply and demand, establishing prices and schedules for each hour. However, due to uncertainties such as renewable energy variability and demand fluctuations, discrepancies often arise between scheduled and actual electricity flows. To address

these discrepancies, the intraday market allows for adjustments closer to real-time, enabling participants to trade electricity in shorter intervals to balance supply and demand effectively. This sequential market structure enhances system flexibility and reliability, particularly in systems with high renewable energy penetration. Studies have proposed integrated models that optimize participation across the day-ahead market and the intraday market, demonstrating improved economic outcomes and system efficiency [123], [124], [125]. Furthermore, integrating continuous intraday trading into multi-market optimization frameworks has enhanced the operational flexibility of industrial energy systems, highlighting the importance of coordinated market participation strategies [126], [127].

Accordingly, this chapter proposes a generic hierarchical structure for contributive consumers to cope with energy markets at different time scales. The idea is to build a hierarchical energy management system to handle changes in different energy markets, from the slow variations (day-ahead market) to the fast changes (balancing market), with an appropriate performance target. However, in this thesis, we focus on developing control algorithms for the intraday market with variations of 15 minutes and the balancing market with variations of several seconds for frequency stability (primary demands). This generic management structure has been sketched in Fig. 5.1.

As shown in this figure, the proposed architecture has three main layers for decision making. Each layer corresponds to a time scale with its requirements and goals. Starting from the top, the controller C_1 is responsible for specifying the optimal power profile of the plant for the coming 24 hours. In fact, this controller buys the day-ahead power of the plant from the day-ahead market, considering various parameters and predictions. At this level of decision-making, it is important to have a good knowledge of the plant consumption on the next day, a reliable prediction of energy prices during this interval, what the power consumption of the plant was on the last day, and how far the plant is from completing its goal (e.g., bottle production, target SOC of the EVs, etc.) considering time constraints. The outcome of this controller is P_1^*, the day-ahead optimal power profile, which feeds into C_2 as its input. The C_2 controller is responsible for updating the power consumption of the plant in a shorter horizon, while handling mid-term ancillary services (δE, including load-shifting). In other words, this controller specifies the energy consumption (energy packets ΔE^*), setpoints of binary prosumers inside the plant (Δ_2^*), and setpoints of continuous prosumers (r_2^*) according to the feedback from C_3 (ΔE, Δ_2, x), updated energy prices, updated performance specifications of the plant, and mid-term ancillary services in a higher time resolution (approximately from 1 to 5 hours). It receives feedback from C_3 to modify its optimal decision for subsequent time steps, and updates the energy packets and setpoints of binary prosumers accordingly. This feedback includes the amount of energy consumed by the plant during one sample time of the C_2 controller (ΔE), the status of binary prosumers (Δ_2), and the states of the plant (x, e.g., SOC of the batteries or number of bottles produced by a production line). The C_3 controller is the lowest layer of the whole energy management system that directly manipulates the plant through the signals it sends to the actuators. This controller receives the optimal energy packets (ΔE^*), setpoints of binary prosumers (Δ_2^*) as well as setpoints of continuous prosumers (r_2^*), and finds the optimal power consumption of the plant (P_3^*), the setpoints of the binary prosumers (Δ_1^*), and the setpoints of the continuous prosumers (r_3^*) according to the primary services (δP). The controller has a shorter prediction horizon and, consequently, a shorter sample time (usually in the order of seconds), enabling the plant to react quickly to the

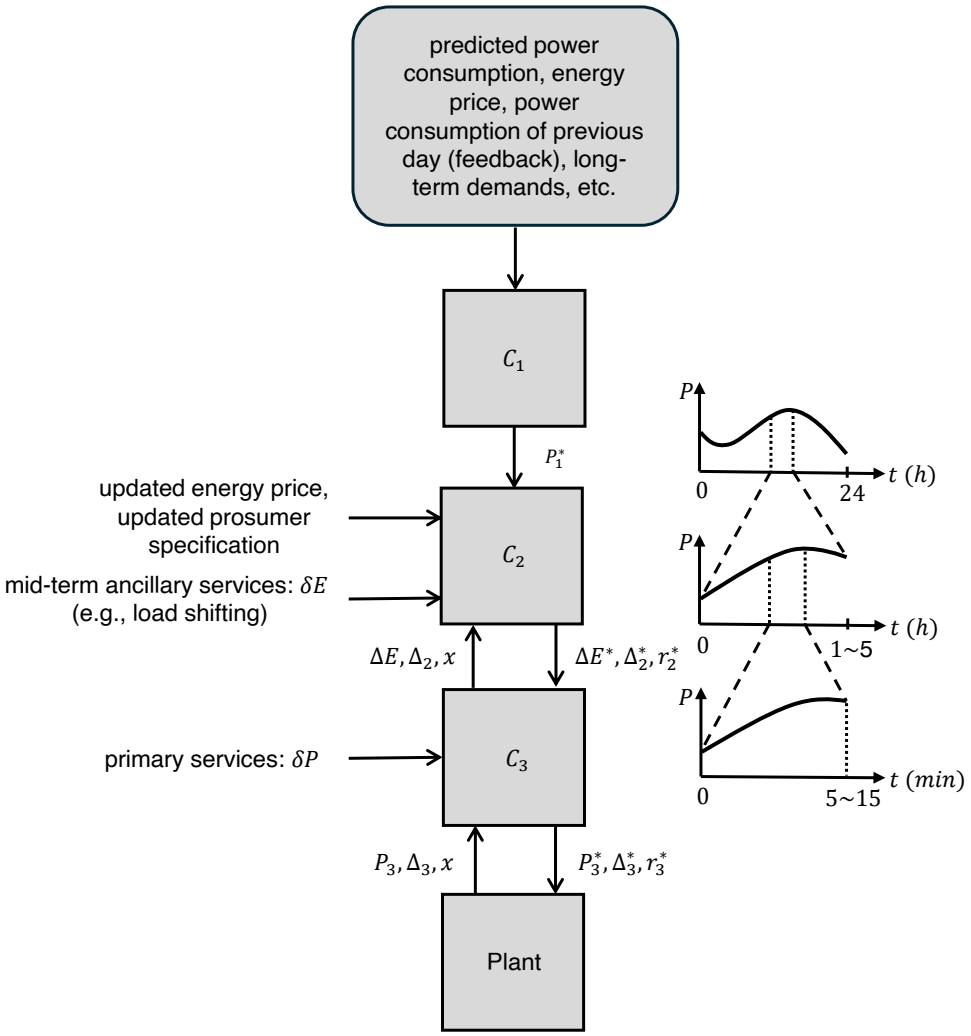


FIGURE 5.1: Generic hierarchical architecture for energy management of prosumers.

demands. The crucial point is that C_3 has a dynamic high-fidelity model of the prosumers, while such a model is not required at the higher levels of decision-making.

The proposed hierarchical control architecture is defined at an interface level to ensure modularity and applicability across different loads/prosumers and application domains. Rather than prescribing a fixed internal realization for each layer, the architecture specifies the time scale, decision scope, and exchanged quantities that define the interaction between layers. These interfaces consist of physically and economically interpretable signals such as power setpoints, energy tariffs, DR signals, user preferences, flexibility contracts, reference trajectories, and state-related information. The internal formulation of each layer, including the specific cost functions, constraints, and models, is application-dependent. The proposed approach allows individual layers to be modified, replaced, or extended without affecting the overall control framework. This interface-based abstraction is a deliberate design choice that supports scalability, reuse, and adaptation to heterogeneous plants while preserving a clear separation of responsibilities across time scales. It is important to mention

that we focus on developing the synthesis of C_2 and C_3 controllers, assuming that the day-ahead power profile P_1^* is available from C_1, while neglecting the method by which it is determined. However, we have some clues on how the C_1 controller can be designed. One way to synthesize this controller is to consider a stochastic model of energy prices and the plant's consumption and production, along with its time plan, to form the optimization problem. Moreover, there is room for utilizing machine learning and data-driven approaches to determine the optimal profile by collecting the necessary data, such as plant consumption, energy prices, and long-term demand, among other relevant factors. More explanation has been brought in Chapter 6 in this regard. Overall, the design of this controller is beyond the scope of this thesis.

5.3 Use-Case of Beverage Production Plants

Building upon the generic hierarchical control structure presented in Section 5.2, this section focuses on its customized design for a beverage production plant integrated with a battery storage system. The hierarchical controller is designed to use the flexibility offered by the battery to respond to DR signals—both primary and mid-term—without violating the production constraint. Moreover, we consider a buffer of empty bottles between the air transport machine and the filler to provide additional flexibility for production. The main reason for using this buffer is to run the continuous components of the line when the binary machines are off, thereby reducing the delay in satisfying bottle production goals. Also, when the continuous machines stop and the binary machines still work, the empty bottles can be stored in the buffer and processed later when the continuous machines resume processing. Considering this buffer in the line architecture, as shown in Fig. 5.2, provides the opportunity to control the binary units separated from the continuous section of the line for a limited time interval.

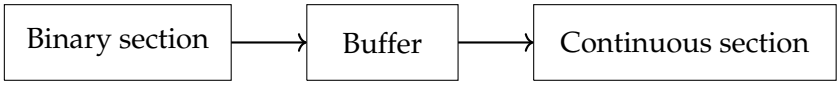


FIGURE 5.2: Production plant with a buffer of bottles.

As an example, if a demand necessitates that the binary machines stop, the continuous machines can continue production by taking bottles from the buffer. In this use case, the hierarchical energy management system finds the optimal processing speed of the continuous units, the mode of binary units, and the battery's power by C_2 for a long horizon, meeting secondary demands. It updates the optimal decisions by C_3 in a short horizon to capture primary demands. The proposed hierarchical controller is depicted in Fig. 5.3.

Starting from the top, the optimal day-ahead power profile P_{1d}^* is sent to the C_2 controller, considering factors such as predicted energy price and consumption of the plant. The controller is responsible for planning the process within a 1- to 5-hour horizon, as shown in Fig. 5.3. To specify the optimal working mode of the binary units, continuous units, and the battery, the controller requires additional information about the production line. C_2 must get the updated energy price prediction (μ_1), the price of bottles (μ_2), how many bottles (n_f) are mainly planned to be produced in the predefined time t_f, what the mid-term demand is asking for ($\delta E, \Delta P, T_d$), and how many bottles can be produced later (δn). The controller assigns the optimal processing speed for the continuous machines, the working mode of the binary units,

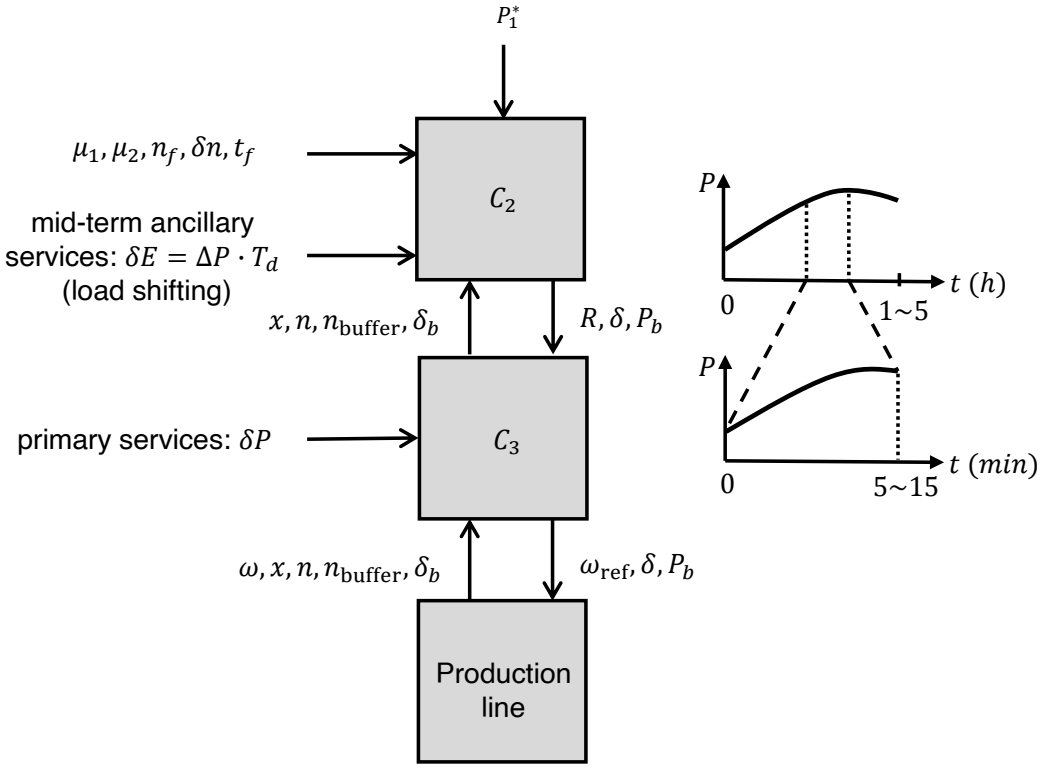


FIGURE 5.3: Hierarchical architecture for energy management of a beverage production plant with battery.

and the battery power. These are the setpoint for the controller C_3 and the C_2 controller receives the SOC of the battery x, the total number of produced bottles so far n, the number of bottles in the buffer n_{buffer}, and the current state of the binary units δ_b.

The C_3 controller not only gets the setpoints from C_2 but also the primary demand δp and calculates the reference speed of the continuous machines, the working mode of binary units, and the power of the battery. This controller, with a higher time resolution, monitors the line and battery and listens to the grid for primary services and reacts accordingly to modify the decision made by C_2. The accessibility of the battery is another important factor that both controllers, especially C_3, must consider for decision-making. If due to a low level of SOC or a failure, the battery cannot be used, both controllers should update their decision accordingly. The C_3 controller follows the decision made by the C_2 if there is no primary demand, no failure of the battery in a short time, and no disturbance to the units. In the design of C_2, it is assumed that all the continuous units are synchronous with processing speed R and no dynamics is considered for planning of the units. However, in the design of C_3, the dynamics of the machines are considered, and the continuous units can deviate from the synchronous processing speed R, issued by C_2, for a limited time.

5.3.1 Design of C_2

To synthesize the C_2 controller, we first need to formulate the energy consumption of the continuous units in terms of the number of bottles produced. Since the relationship between the power consumption of the components and their processing

speed is known, we can write:

$$P_i(t) = d_i\omega_i^2(t) + M_i\omega_i(t) + \alpha_i n_i(t) + p_{i,0} \tag{5.1}$$

This equation is similar to Eq. (3.3), where i is for the i-th unit. Since the rate of processing bottles by the continuous units is proportional to the processing speed of the components, one can write:

$$\omega_i = \lambda_i R \tag{5.2}$$

where R is the production rate at the end of the line, and λ_i is a constant. Therefore:

$$P_i(t) = d_i\lambda_i^2 R(t)^2 + M_i\lambda_i R(t) + \alpha_i n_i(t) + p_{i,0} \tag{5.3}$$

The total power of the continuous components is the summation of the power consumption of these units:

$$P_\sigma = \left(\sum_i d_i\lambda_i^2\right) R^2(t) + \left(\sum_i M_i\lambda_i\right) R(t) + \left(\sum_i \alpha_i n_i + P_{i,0}\right) \tag{5.4}$$

Assuming that the power consumption is constant during each sample time T_{s_2}, the consumed energy of the units is:

$$\Delta E_c = T_{s_2} \cdot P_\sigma = T_{s_2} A R^2(t) + T_{s_2} B R(t) + T_{s_2} P_0 \tag{5.5}$$

where A, B, and C are as follows:

$$\begin{cases} A = \sum_i d_i\lambda_i^2 \\ B = \sum_i M_i\lambda_i \\ P_0 = \sum_i \alpha_i n_i + P_{i,0} \end{cases} \tag{5.6}$$

This is a critical assumption at the planning level (C_2), which simplifies the synthesis procedure of the controller. The number of bottles produced by the line is then:

$$\Delta n(t) = T_{s_2} R(t) \tag{5.7}$$

assuming that the rate of production is constant during the interval. So, the energy consumption of the continuous units in terms of produced bottles is:

$$\Delta E_c = \frac{A}{T_{s_2}} (\Delta n)^2 + B(\Delta n) + T_{s_2} P_0 \tag{5.8}$$

We consider the bottle production rate R at each sample time T_{s_2} as the decision variable of the C_2 controller:

$$P_c(t) := P_\sigma(t) - P_0, \quad t = kT_{s_2}, \quad k = \{1, \cdots, N_2\} \tag{5.9}$$

By dropping t and showing discrete time instances with k, the power consumption of the total continuous units is as follows:

$$P_c(k) = AR^2(k) + BR(k) \tag{5.10}$$

This equation describes the power consumption of the continuous units in terms of the synchronous processing speed R and is the basic equation for developing the controller C_2. The goal of the planner C_2 is to minimize the energy cost of the plant

while maximizing its production. Additionally, this controller must meet the mid-term demands by utilizing the flexibility provided by the production line. By utilizing the buffer of bottles and the battery, the controller can manage the energy consumption of the line, ensuring that grid demands are satisfied with minimal impact on production schedules. The objective function of the MPC should be formulated such that the energy price is minimized while the bottle production is maximized. Thus, we consider the following cost function for designing C_2:

$$J_2 = \sum_{k=1}^{N_2} (\mu_1(k) T_{s_2} (P_c(k) + \delta(k) \cdot P_d) - \mu_2(k) \Delta n(k)) \tag{5.11}$$

where μ_1 is the price of energy, μ_2 is the bottle price, P_d is the power consumption of binary machines when they are on, δ is the binary variable for turning these machines on or off, and N_2 is the prediction horizon.

One of the main constraints at the C_2 level is the maximum processing speed of the continuous units. The units cannot process faster than a threshold imposed by their PMSMs. Hence, we consider the following constraint in the design of C_2:

$$0 \leq R(k) \leq R_{\max}, \quad k = 1, \cdots, N_2 \tag{5.12}$$

Also, the total number of bottles produced by the line (n) can be easily calculated as follows:

$$n(k+1) = n(k) + T_{s_2} R(k), \quad k = 1, \cdots, N_2 \tag{5.13}$$

For binary machines, the processing speed can only take two values, resulting in the following equations for bottle production and energy consumption:

$$n_d(k+1) = n_d(k) + T_{s_2} \cdot \delta(k) \cdot R_d \tag{5.14}$$

$$e_d(k+1) = e_d(k) + T_{s_2} \cdot \delta(k) \cdot P_d \tag{5.15}$$

where R_d and P_d are the bottle production rate and total power consumption of the binary machines, respectively, and are constant. As mentioned before, $\delta(k)$ is a binary variable determining whether to turn the binary machines on or off. Since the continuous machines are synchronized, the bottle production rate at the end of the line is assumed to be equal to the filler bottle intake rate. The number of bottles in the buffer between the air transport and the filler is then:

$$n_{\text{buffer}}(k+1) = n_{\text{buffer}}(k) + T_{s_2} \cdot (\delta(k) \cdot R_d - R(k)) \tag{5.16}$$

This holds if no other bottles are injected or taken in the middle of the line. This equation also implies that if the binary machines stop working and the continuous units continue to run, the number of bottles in the buffer decreases. On the other hand, if the binary units run and the processing speed of the continuous units reduces, the number of bottles in the buffer increases. If the binary units run and the continuous units work at full speed, there will be no change in the number of bottles in the buffer. It is also important to consider an upper and lower bound on the number of bottles in the buffer. We consider the lower bound because in the event of a failure in the binary units, the continuous part of the line should be able to continue production by drawing from the buffer of unprocessed bottles. Additionally, the buffer's capacity is limited and cannot obviously hold an infinite number of bottles.

As a result, another constraint in the optimization problem is:

$$n_{\min} \leq n_{\text{buffer}}(k) \leq n_{\max} \tag{5.17}$$

where $n_{\min}$ and $n_{\max}$ are the lower and upper bounds on the number of bottles in the buffer, respectively.

The battery, which can supply the continuous section of the line, is considered in this problem setting, where its accessibility and SOC amount are critical for decision-making at the level of C_2. There are several methods to write a mathematical model of a battery [128]; however, we use an abstract Coulomb counting model for planning [129]:

$$x(k+1) = x(k) - T_{s_2} \cdot \eta \cdot \delta_b(k) P_b(k) \tag{5.18}$$

where x is the SOC of the battery, η is its Coulombic efficiency, δ_b is a binary parameter that specifies the accessibility of the battery, and $P_b \geq 0$ is the power of the battery. It is assumed that the controller can only discharge the battery, and it is not accessible during charging. However, the accessibility of the battery is not only due to charging, but also a failure or sharing with other processes might affect its accessibility. Besides, the SOC cannot be less than a threshold due to safety and efficiency considerations:

$$x_{\min} \leq x(k), \quad k = 1, \cdots, N_2 \tag{5.19}$$

where $x_{\min}$ is the minimum SOC. Also, each battery can deliver power up to its nominal power, resulting in an upper bound on P_b in our problem setting:

$$0 \leq P_b(k) \leq P_{b,\max}(k), \quad k = 1, \cdots, N_2 \tag{5.20}$$

$P_{b,\max}(k)$ is the upper bound that depends on the battery's SOC.

As mentioned earlier, the controller C_2 is responsible for meeting the mid-term demands requested by the grid and should use the flexibility provided by the production line, the buffer of bottles, and the battery to participate in DR. The grid sends the demands parametrized in terms of power or energy (δE or ΔP in Fig. 5.3). One way to consider these demands in our problem setting is to add a hard constraint such that during the demand interval, the whole power consumption of the line is bounded to a limit set by the grid. This is the case for demands such as load shifting, which is our primary focus in this section:

$$0 \leq P_c(k) + \delta(k) \cdot P_d \leq P_{\text{ref}}(k) - \Delta P(k), \quad k = 1, \cdots, N_2 \tag{5.21}$$

P_{ref} is the reference power consumption of the line, which is usually equal to the nominal power consumption, and $\Delta P \geq 0$ is the amount of power the grid asks for reduction. It can be positive during the whole prediction horizon or for a shorter period.

The most important constraint of the problem is satisfying the production requirements. In fact, the main mission of the controller is to meet the production goals while utilizing flexibility to meet grid demands. The production contract with the customer, as discussed in Chapter 3, imposes a hard constraint on the processing speed of the line, requiring it to produce a minimum number of bottles within a specific timeframe. For example, the line is expected to produce 150,000 water bottles within 3 hours. The controller (planner) must manage the energy consumption of the whole line (binary and continuous machines) to satisfy this constraint eventually.

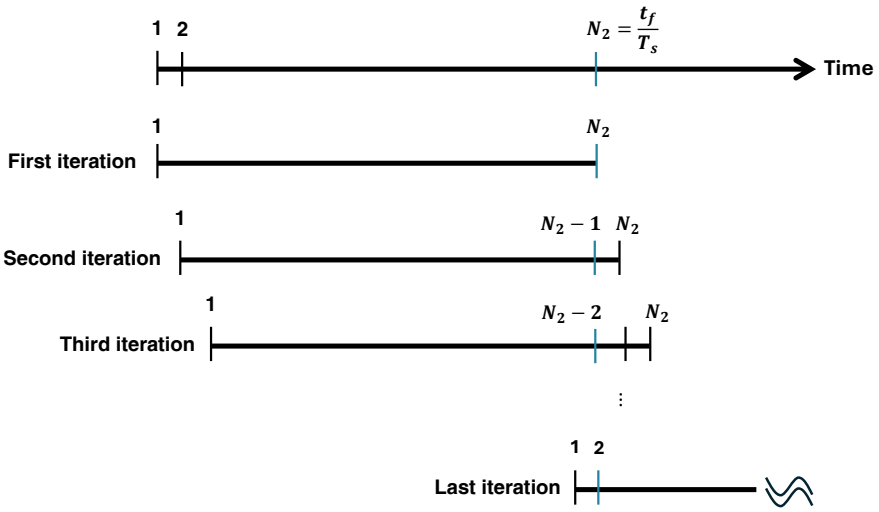


FIGURE 5.4: Temporal interpretation of target production constraint in Eq. (5.22).

Therefore, the following constraint must be considered in the optimization problem:

$$n(j) \geq n_f - \delta n, \quad j \in \{1, \cdots, N_2\}, \quad jT_{s_2} = t_f \tag{5.22}$$

n_f is the number of bottles targeted to be produced at time t_f, and δn is the number of bottles that can be shifted to the future to be produced. It is vital to mention that when MPC is running for planning, t_f is at the end of the horizon, meaning $j = N_2$. However, as time passes, the constraint must be satisfied before the end of the horizon ($j < N_2$). It means j will take values such as $N_2 - 1$, $N_2 - 2$, and so on. Fig. 5.4 illustrates the temporal interpretation of this constraint. Having formulated all the constraints and the objective function, one can write the following problem formulation for the C_2 controller:

$$\min_{R(k),\delta(k),P_b(k)} \quad J_2 = \sum_{k=1}^{N_2} \left(\mu_1(k) T_{s_2} (P_c(k) + \delta(k) \cdot P_d) - \mu_2(k) \Delta n(k)\right) \tag{5.23}$$

$$\text{s.t.:} \quad \begin{cases} P_c(k) = AR^2(k) + BR(k) - \delta_b(k) P_b(k), \\ n(k+1) = n(k) + T_{s_2} \cdot R(k), \\ 0 \leq R_{\min} \leq R(k) \leq R_{\max}, \\ x(k+1) = x(k) - T_{s_2} \cdot \eta \cdot \delta_b(k) P_b(k), \\ 0 \leq P_b(k) \leq P_{b,\max}(k), \\ x_{\min} \leq x(k), \quad k = 1, \cdots, N_2 + 1 \\ n_{\text{buffer}}(k+1) = n_{\text{buffer}}(k) + T_{s_2} \cdot (\delta(k) \cdot R_d - R(k)), \\ n_{\min} \leq n_{\text{buffer}}(k) \leq n_{\max}, \\ 0 \leq P_c(k) + \delta(k) \cdot P_d \leq P_{\text{ref}}(k) - \Delta P(k), \quad k = 1, \cdots, N_2 \\ n(j) \geq n_f - \delta n, \quad j \in \{1, \cdots, N_2\}, \quad jT_{s_2} = t_f. \end{cases}$$

The horizon of this controller is on the order of hours, typically ranging from 1 to 5 hours, resulting in a sample time T_{s_2} of 5 to 15 minutes.

5.3.2 Design of C_3

The planner (C_2) performs on a long time horizon and ignores the dynamics of the machines in the production line. This is the core idea for considering continuous units as synchronous processing machines that operate at the same speed R. Moreover, this controller captures the mid-term demands, which typically occur at a maximum frequency of every 15 minutes. Therefore, it is not necessary to consider high time resolution in the design procedure. The controller determines the processing speed of the continuous units, as well as the working mode of the binary units, thereby calculating the total power consumption of the line. However, we also want to target the primary demands: the accessibility of the battery at higher time resolutions and the dynamics of the machines. The more often the feedback from the processing machines and the battery is received, the faster the controller can modify the control signals accordingly.

In general, separating planning and higher time-resolution control into two distinct levels is essential for effectively managing system dynamics and computational complexity. The planning level operates on a coarser time scale and focuses on long-term decisions, such as scheduling, energy allocation, or strategic responses to forecasts. This layer can incorporate broader objectives, like economic optimization or resource coordination, without being overwhelmed by fine-grained dynamics. In contrast, the lower-level control operates at a higher time resolution to handle fast dynamics, uncertainties, and real-time adjustments, ensuring that the system remains stable and responsive to rapid changes. By decoupling these layers, we can exploit the strengths of each: the planning layer provides optimal guidance over extended horizons, while the control layer ensures precision and feasibility in execution. This hierarchical structure enhances scalability, robustness, and adaptability in complex dynamic systems.

At the lower layer (C_3), the goal is to follow the setpoints of processing speed, binary machine status, and battery power while capturing the primary demands and modifying the decision of C_2. The constraints are mostly the same as those we have considered for C_2, but with some exceptions. First, the dynamics of the processing machines must be considered at this level. Since the sampling time of the controller is in the order of several seconds, the transient response of the machines is now within the time frame of the controller. Therefore, we consider the dynamics of the machines in the problem formulation as follows:

$$\begin{cases} \omega_i(k+1) = \omega_i(k) + \dfrac{T_{s_3}}{J_i}(-d_i \cdot \omega_i(k) - M_i + \tau_{\text{mec},i}(k)) \\ \tau_{\text{mec},i}(k) = k_{1i}(\omega_{\text{ref},i}(k) - \omega_i(k)) + k_{2i} T_{s_3} \displaystyle\sum_{l=1}^{k} (\omega_{\text{ref},i}(l) - \omega_i(l)) \end{cases} \tag{5.24}$$

where T_{s_3} is the sampling time of the C_3 controller and the other variables are similarly defined as in Eq. (3.1) and Eq. (3.2). Now, the basic assumption is that the speed of the machines can be controlled independently, and they do not necessarily run synchronously. However, the deviation from the synchronous processing speed of the line cannot continue for long because it results in running machines with empty slots or extreme jamming on the conveyor belts. So, the following constraint allows

the machines to deviate from the synchronous speed R for a limited time:

$$\begin{aligned} &\gamma_i\omega_i(k) = \gamma_{i+1}\omega_{i+1}(k) + s_i(k), \quad i = 1, \cdots, m-1 \\ &-s_{\min,i} \leq T_{s_3} \sum_{k=1}^{N_3} s_i(k) \leq s_{\max,i}, \quad i = 1, \cdots, m-1 \end{aligned} \tag{5.25}$$

N_3 is the prediction horizon of C_3, s_i is a slack variable, and m is the total number of continuous units in the line, including conveyor belts. $s_{\min,i}$ and $s_{\max,i}$ are the bounds on the bottle mismatch between unit i and unit $i+1$. Furthermore, the bounds on the processing speed of the units must be considered in the design of the controller C_3:

$$\omega_{\min,i} \leq \omega_i(k) \leq \omega_{\max,i}, \quad k = 1, \cdots, N_3, \quad i = 1, \cdots, m \tag{5.26}$$

$\omega_{\min}$ and $\omega_{\max}$ are the minimum and maximum processing speeds of the i-th unit, respectively.

The constraint imposed by primary DR must also be considered in the synthesis of C_3. The grid requests always limit the maximum power consumption of the production line. Since the controller C_2 considers a constant power for the line within its sample time (P_{ref}), any request for reducing power consumption in C_3 is denoted by δp and is included in the optimization problem as follows:

$$\sum_{i=1}^{m} P_i(k) - \delta_b(k)P_b(k) + \delta(k)P_d \leq P_{\text{ref}} - \delta p(k) \tag{5.27}$$

Also, the number of bottles expected to be produced by the line within the horizon of C_3 can be calculated from the synchronous speed R issued by C_2. The controller C_3, however, can slightly deviate from this value to meet the primary demands. We incorporate the following constraint to track the production schedule with minimum deviation:

$$T_{s_3} \sum_{k=1}^{N_3} \gamma_m\omega_m(k) \geq T_{s_2}R - \epsilon, \quad \epsilon \geq 0 \tag{5.28}$$

ϵ is a slack variable that will be imported into the cost function appropriately. $\gamma_m\omega_m$ is the processing speed of the last continuous machine in the line (shrink packer).

The goal of the C_3 controller is to follow the setpoints sent by the C_2 controller (R, δ, and P_b) while responding to primary demands and handling fast dynamics and real-time adjustments. The definition of the cost function for C_3 can include the second-order terms for tracking the setpoints issued by the upper-layer controller and a reflection of the slack variable ϵ to have the minimum possible deviation from production goals in its horizon. The cost function that we consider for C_3 in this section is:

$$\begin{aligned} \min_{\omega_{\text{ref},i}(k), P_b(k), \delta(k), \epsilon, s_i(k)} \quad & J_3 = \epsilon + \alpha \sum_{i=1}^{m} \sum_{k=1}^{N_3} \left(\omega_i(k) - \lambda_i R\right)^2 + s_i^2(k)\big) + \\ & \sum_{k=1}^{N_3} (P_b(k) - P_{b,\text{ref}})^2 + \sum_{k=1}^{N_3} (\delta(k) - \delta_{\text{ref}})^2 \end{aligned} \tag{5.29}$$

where α is a positive coefficient, $P_{b,\text{ref}}$ is the power of the battery set by C_2, and δ_{ref} is the binary variable for binary units defined by C_2. Finally, the whole optimization problem for the C_3 controller can be written as:

$$\min_{\omega_{\text{ref},i}(k),P_b(k),\delta(k),\epsilon,s_i(k)} \quad J_3 = \epsilon + \alpha \sum_{i=1}^{m}\sum_{k=1}^{N_3}\left((\omega_i(k)-\lambda_i R)^2 + s_i^2(k)\right) + \tag{5.30}$$

$$\sum_{k=1}^{N_3}(P_b(k)-P_{b,\text{ref}})^2 + \sum_{k=1}^{N_3}(\delta(k)-\delta_{\text{ref}})^2$$

$$\text{s.t.:} \quad \begin{cases} P_i(k) = d_i\omega_i^2(k) + M_i\omega_i(k), \quad i=1,\cdots,m, \quad k=1,\cdots,N_3, \\ \omega_i(k+1) = \omega_i(k) + \dfrac{T_{s_3}}{J_i}(-d_i\cdot\omega_i(k) - M_i + \tau_{\text{mec},i}(k)), \\ \tau_{\text{mec},i}(k) = k_{1i}(\omega_{\text{ref},i}(k)-\omega_i(k)) + k_{2i}T_{s_3}\sum_{l=1}^{k}(\omega_{\text{ref},i}(l)-\omega_i(l)), \\ \omega_{\min,i} \le \omega_i(k) \le \omega_{\max,i}, \\ \sum_{i=1}^{m} P_i(k) - \delta_b(k)P_b(k) + \delta(k)P_d \le P_{\text{ref}}(k) - \delta p(k), \\ 0 \le P_b(k) \le P_{b,\max}(k), \\ T_{s_3}\sum_{k=1}^{N_3}\gamma_m\omega_m(k) \ge T_{s_2}R - \epsilon, \quad \epsilon \ge 0, \\ \gamma_i\omega_i(k) = \gamma_{i+1}\omega_{i+1}(k) + s_i(k), \quad i=1,\cdots,m-1 \\ -s_{\min,i} \le T_{s_3}\sum_{k=1}^{N_3} s_i(k) \le s_{\max,i}, \\ s_{\max,i}, s_{\min,i} \ge 0, \\ x(k+1) = x(k) - T_{s_3}\eta\cdot\delta_b(k)\cdot P_b(k), \\ x_{\min} \le x(k). \end{cases}$$

The optimization problem is an MIQP that can be solved numerically using most optimization tools and solvers.

5.4 Simulation Results of Use-Case I: Production Line with a Battery

We assume that the production line in Fig. 3.1 is now equipped with a battery that has a nominal power to supply the continuous machines in nominal mode. This means that the battery can fully supply continuous machines if needed. The simulation results are shown in Fig. 5.5 to 5.11. Also, the parameters used for simulating different sub-cases are given in Table. 5.1. Since the production line is the same as in Scenario I in Chapter 3, we skip the parameters of the dynamical model of the line, such as d_i or M_i. The results are for different settings of the optimization problems Eq. (5.23) and Eq. (5.30). This scenario includes four primary demands and one load shifting. It is also assumed that there is a buffer for preforms between the binary units and continuous units of the line. To investigate the proposed hierarchical controller, three main sub-cases have been defined to evaluate its performance.

In the first sub-case, we assume the line has flexibility for three primary demands but not for load-shifting, and the battery is available. In other words, the line with the nominal production rate of 17.5 b/s allows a maximum of 3×30×17.5 = 1,575 bottles to be produced later to offer primary services to the grid. However, it does

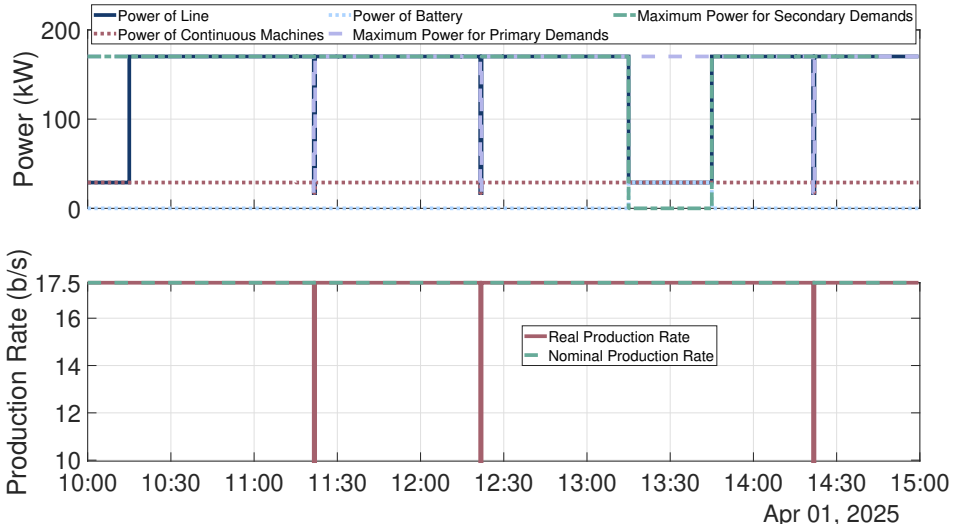


FIGURE 5.5: Power profiles and production rate of the line for scenario I, sub-case 1. The power profiles in the upper subplot are zoomed in Fig. 5.6 for the third primary demand.

not offer any flexibility for long-term power reduction, resulting in more bottle production loss during the simulation interval. The duration of the whole scenario is 5 hours on a specific day, starting from 10:00:00 to 15:00:00. There are two subplots in Fig. 5.5, the upper subplot demonstrates the power profiles of the line (including continuous and binary machines), as well as the upper bound of power consumption for ancillary services. The lower subplot displays the actual production rate of the line alongside its nominal (reference) production rate.

The continuous units start from the beginning to process the bottles they take from the buffer of preforms, leading to the nominal power consumption of this section of the line, which is 28.94 kW. The binary units begin processing with a 15-minute lag starting at $t =$ 10:15:00. Three primary demands are received at $t =$ 11:21:30, 12:21:30, and 14:21:30, each requesting a 90% reduction in the whole power reduction of the line. The whole power consumption of the production line is 170.24 kW, and each primary demand lasts for 30 seconds. The mid-term demand is also received at $t =$ 13:15:00 and lasts until $t =$ 13:45:00, requesting that the line be fully stopped. Since the flexibility for primary demands has been considered up to 3 requests, the

TABLE 5.1: **Simulation parameters of the production line for use-case I.**

Parameter	Value (unit)	Parameter	Value (unit)
μ_1	0.2 €/kWh	N_2	20
μ_2	0.22 €/b	t_f	5 h
A	0.0687 J · s	$\omega_{\min,i}$	0 kb/h
B	0.4514 J	$\omega_{\max,i}$	63 kb/h
$R_{\min}$	0 kb/h	$s_{\min,i}$	175 b
$R_{\max}$	63 kb/h	$s_{\max,i}$	175 b
$n_{\min}$	15,750 b	$x_{\min}$	0.1
$n_{\max}$	63,000 b	α	100

line fully participates in primary demands by turning off the binary units and reducing the processing speed of the continuous section. On the other side, since the line operator does not provide any flexibility for mid-term demands (δn in Eq. (5.23)) and the battery is available, the controller exploits the battery to supply the continuous section. Though the binary units are off during this period, the continuous section of the line takes the preforms from the buffer to keep the production alive. The lower subplot of Fig. 5.5 shows how the production rate of the line is affected by primary demands but not the mid-term demand. Additionally, a zoom of the power profiles for the third primary demand is depicted in Fig. 5.6. It shows more clearly how the power profiles change during a primary demand, which is usually a short-term demand. The demand reduces the power consumption of the continuous section to 17.02 kW and turns off the whole binary units.

Figure 5.7 illustrates the bottle variation in the buffer, the whole produces bottles, and the on-off signal for binary machines (δ in Eq. (5.23) and Eq. (5.30)) in three subplots, from top to bottom, respectively. At the beginning of the simulation, the buffer of preforms is full with 63,000 bottles. Since the continuous section starts to process 15 minutes before the binary section, the number of bottles in the buffer decreases. At t =10:15:00, the number of bottles remains constant in the buffer due to equal speed of incoming and uptake preforms. Also, there are two slight drops when the first two primary demands are received. The reason is that the binary units stop functioning during the demands, but the continuous section remains operational. When the mid-term demand is issued, the number of bottles in the buffer decreases considerably for the same reason. This drastic decline is followed by another decrease due to the third primary demand, resulting in reaching the lowest level of bottles in the buffer. We have considered an upper bound and a lower bound for the buffer. The buffer's capacity is limited, and a minimum number of bottles should be inside the buffer in case of an emergency stop of the binary section of the line.

The second subplot illustrates the number of produced bottles when there is flexibility for DR services and when there is no such flexibility, meaning the entire line operates in nominal mode for the entire interval. When there is a primary demand, the black profile (produced bottles without flexibility) exceeds the real profile of the produced bottles, though the difference is not notable. This originates from the fact

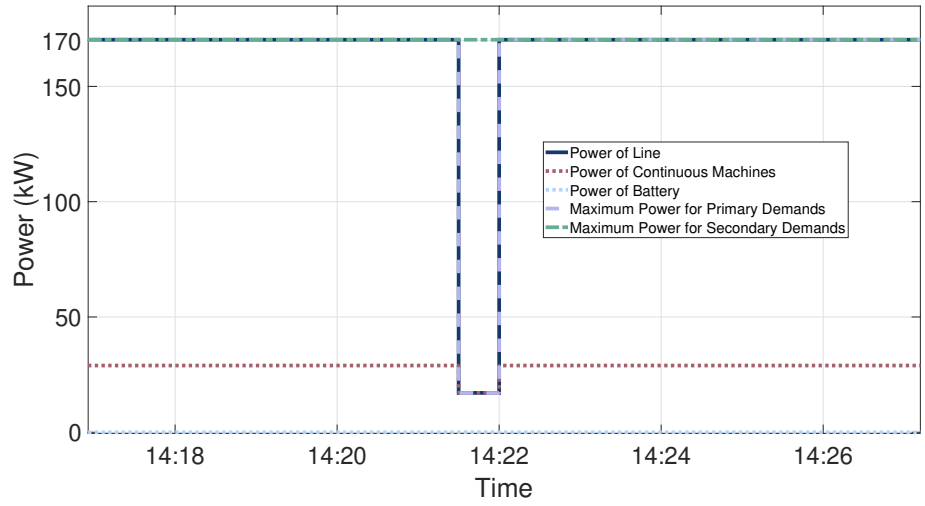


FIGURE 5.6: Power profiles during the third primary demand in scenario I, sub-case 1.

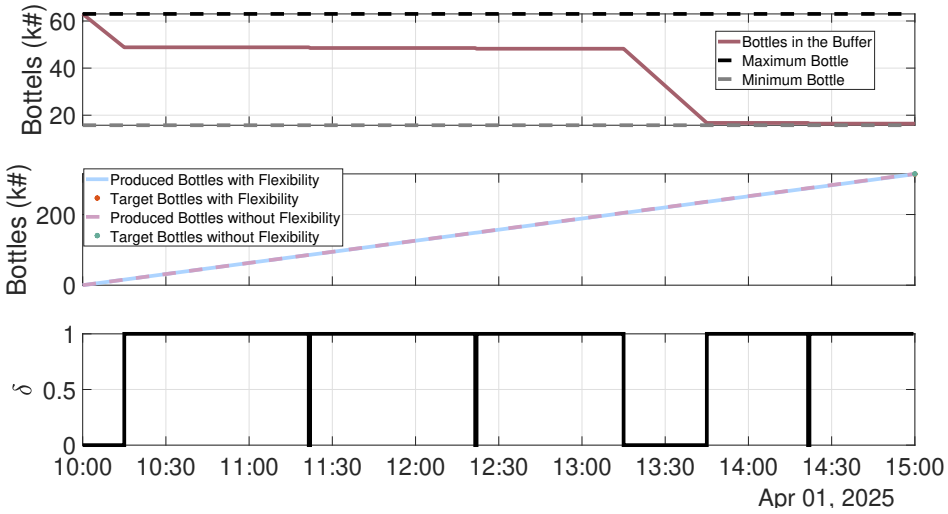


FIGURE 5.7: Bottle Production, bottles in the buffer, and binary signal in scenario I, sub-case 1.

that the line offers little flexibility (maximum 1,575 bottles). If the line operates at its nominal speed for the entire 5 hours, it produces 315,000 bottles. Assuming a maximum of 1,575 bottles for primary demand, the target number of bottles in both cases is almost the same, leading to the overlapping of the green and red points on the second subplot at the end. The third subplot of Fig. 5.7 illustrates the binary signal for making the binary units on or off. These units, as previously mentioned, start to work with a 15-minute delay and turn off during all demands. Therefore, the δ signal fluctuates accordingly, as expected.

The second sub-case involves providing flexibility for mid-term demands while making the battery unavailable for supplying the continuous section of the line. The simulation results are presented in Fig. 5.8 and 5.9. Similar to the previous sub-case, the continuous section starts from the beginning of the interval to process the preforms, while the binary section begins with a 15-minute delay (upper subplot of Fig. 5.8). During the primary demands that are received at the same time as the previous sub-case, the binary and continuous sections of the line participate in DR. During the mid-term demand, unlike the previous sub-case, the entire line is turned off because the battery is unavailable to supply the continuous units. However, there are two intervals (from $t =$ 11:30:00 to 11:45:00 and from $t =$ 14:30:00 to 14:45:00) in which the binary units are turned off. Since one of the motivations, as considered in the cost function of Eq. (5.23) and Eq. (5.30), is to reduce the energy consumption of the line, the binary units stop functioning in these two intervals, and the buffer supplies the continuous section. This is, of course, due to the sufficient number of bottles present in the buffer. The total flexibility considered this time is $17.5\times900\times2$ + $17.5\times30\times3$ = 33,075 bottles. This number is chosen such that the line can support a maximum of three primary demands each for 30 seconds and a maximum duration of 30 minutes, full stop, due to mid-term demands. The power consumption of the continuous part (shown with a red curve in the upper subplot of Fig. 5.8) falls only during mid-term demand. Examining the lower subplot of this figure reveals that the entire flexibility considered for mid-term demands is utilized within this interval. The production rate drops four times, each time an ancillary service is requested by the grid to be provided by the line.

The remaining plots of the second sub-case are sketched in Fig. 5.9. It contains

three main subplots, each highlighting the variation in key signals from the simulation. Starting from the top, the number of bottles in the buffer has three principal declines and two slight drops. The first decrease is due to the delay in turning on the binary units. As a result, the continuous part takes the unfilled bottles from the buffer. The other two main changes in the number of bottles are due to different incoming and outgoing rates of bottles. The binary units do not function while the continuous section works. The small drops during primary demands have the same reason; considering this fact, the processing speed of continuous sections is slower than the nominal. The second subplot indicates the number of produced bottles at each time as well as the nominal profile of bottle production when there is no demand. The number of bottles produced follows the ideal curve by the first three demands (2 primary and one mid-term), and the distance between these curves increases when the next mid-term demand is received and the ideal curve significantly dominates the blue curve.

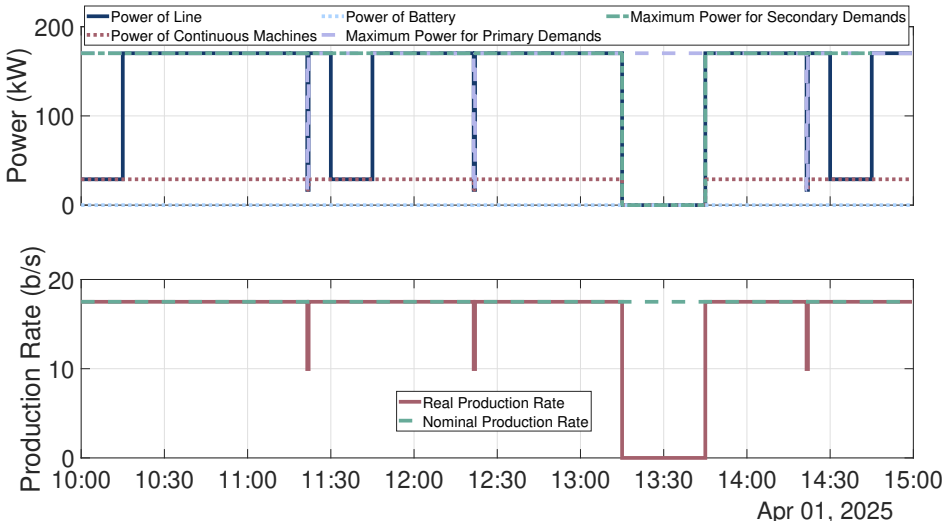


FIGURE 5.8: Power profiles and production rate of the line for scenario I, sub-case 2.

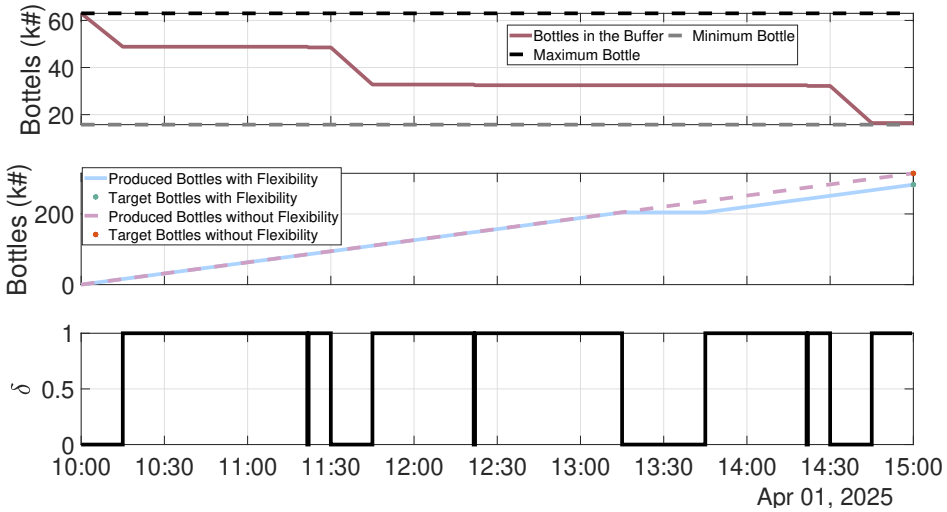


FIGURE 5.9: Bottle Production, bottles in the buffer, and binary signal in scenario I, sub-case 2.

However, the number of produced bottles always stays within the bounds and targets the minimum final value highlighted with a red point at the end (second subplot of Fig. 5.9). The third subplot shows the binary signal of the binary units. As expected, it demonstrates when the binary units are on or off. The availability of the battery and the flexibility together have been considered in the third sub-case. The simulation results are plotted in Fig. 5.10, and 5.11. As shown in Fig. 5.10, the continuous part functions from the beginning, getting different processing rates during the first 45 minutes, from $t =$ 10:00:00 to $t =$ 10:45:00. From $t =$ 10:45:00 on, the entire line works in full nominal mode. Similar to previous sub-cases, there are three primary demands and one mid-term demand. The controller uses the battery during the mid-term demand to supply the continuous section. The number of bottles planned for production is less than the ideal number. Therefore, the line is not required to operate in nominal mode for the whole duration. Consequently, the binary section does not function during the first 45 minutes, and the continuous section operates at a processing speed that is sometimes less than nominal. This strategy reduces the energy price while still achieving the target number of bottles. The lower subplot of Fig. 5.10 shows the production rate of the line. This rate has been reduced from 10:15:00 to 10:45:00, as well as during the primary demands. As depicted in Fig. 5.11, the number of bottles in the buffer has two primary declines. The first one is between 10:15:00 and 10:45:00, in which the binary section is turned off and the continuous section is on. The second one is during the mid-term demand. When the processing rate of the continuous section is less than that of the binary section, it results in a reduction in the number of bottles in the buffer. The second subplot illustrates how the number of produced bottles falls short of the ideal curve in this sub-case. The main gap occurs from 10:15:00 to 10:45:00, when the line's production rate is significantly reduced. This difference increases further when three primary demands are met, resulting in approximately 226 fewer bottles than the ideal case. Ultimately, the number of produced bottles is 297,996 (the red point in the subplot), exceeding the minimum targeted number of 281,925. The third subplot illustrates the binary signal for turning the binary units on and off, as expected.

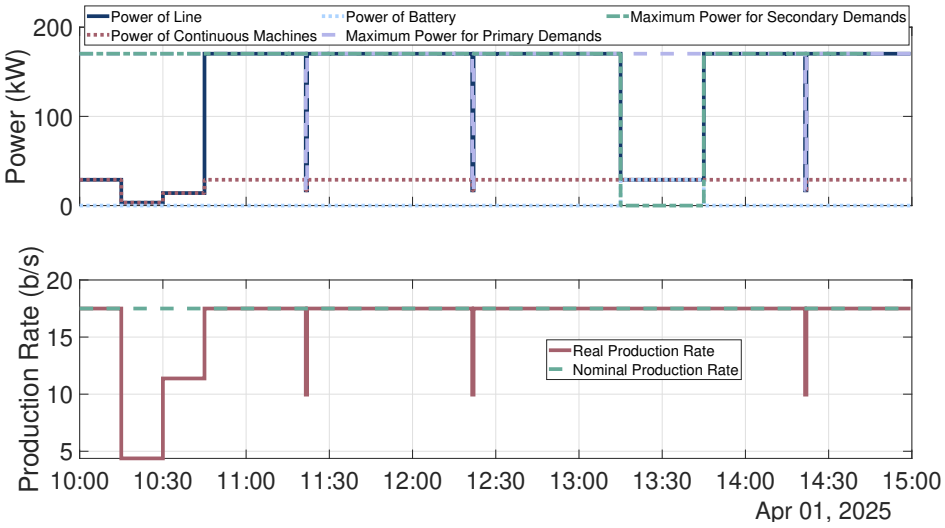


FIGURE 5.10: Power profiles and production rate of the line for scenario I, sub-case 3.

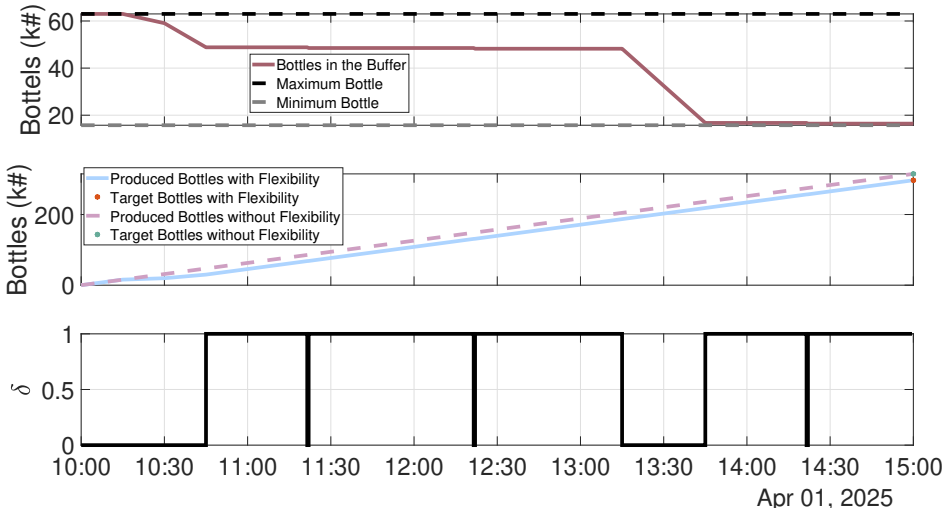


FIGURE 5.11: Bottle production, bottles in the buffer, and binary signal in scenario I, sub-case 3.

5.5 Use-Case of Bidirectional EV Charging Hubs

In this use-case, the focus is to adapt the proposed hierarchical structure to an EV charging hub. We assume the hub operates with a centralized control system that oversees energy prices, grid demand, and EV charging needs. The EV batteries and the energy buffer are modeled using a simplified battery representation. The model is defined by the SOC dynamics as follows:

$$x_i(k+1) = x_i(k) + T_{s_3} \cdot \eta_i P_i(k), \quad \forall i \in \{1, \cdots, n\} \cup \{b\} \tag{5.31}$$

Let $x_i(k)$ denote the SOC of the i-th battery in the charging hub, which includes both EVs and the energy buffer. The variable k represents the discrete time index, while T_{s_3} denotes the sampling interval. The parameter η_i is a scaling factor influenced by the coulombic efficiency, voltage, and capacity of battery i. The power associated with battery i at time k is represented by $P_i(k)$. The hub is assumed to serve n EVs and one buffer, labeled b. This battery model balances computational efficiency and realistic behavior, enabling tractable yet credible simulations.

The control strategy adopts a hierarchical structure with two decision-making layers, as illustrated in Fig. 5.12. Each control level employs a distinct battery model to derive optimal control actions. The lower-level controller (C_3) utilizes the battery representation given in Eq. (5.31), whereas the higher-level controller (C_2) adopts an energy-based battery model. Specifically, the design of C_2 necessitates a unified parametrization of all storage elements to ensure consistency with its control objectives. Accordingly, at this level, each battery is modeled as follows:

$$c_i = c_i(\Delta E_i, T_{a_i}, T_{d_i}), \quad \forall i \in \{1, \cdots, n\} \tag{5.32}$$

where c_i stands for the i-th EV, ΔE_i is the total energy required to charge the EV to the desired value of SOC, T_{a_i} is the arrival time of the EV, and T_{d_i} is its departure time. In the case of buffer, it is parametrized as:

$$c_b = c_b(\Delta E_b, T_{a_b}, T_{d_b}) \tag{5.33}$$

We assume that the buffer is available from the beginning of the time horizon ($T_{a_b} = 0$) and remains in the charging hub forever ($T_{d_b} = \infty$). However, due to its limited capacity, the time it takes to become fully charged is finite. For EVs, the departure time T_{d_i} also defines the moment by which the target SOC must be reached. The sampling time of the higher-level controller C_2 is denoted by T_{s_2}, and its task is to determine the optimal allocation of the energy increments ΔE_i over the interval $[T_{a_i}, T_{d_i}]$ for the EVs, and ΔE_b over $[T_{a_b}, T_{d_b}]$ for the buffer. The energy allocated at sampling instant k is denoted by $\Delta E_i(k)$ for EVs and $\Delta E_b(k)$ for the buffer. For controller design, it is essential to distinguish between positive and negative values of $\Delta E_i(k)$ and $\Delta E_b(k)$. In particular, energy purchase and energy sale must be separated to enable proper formulation of the multi-asset optimization problem for C_2. This distinction is necessary because different pricing applies to buying and selling energy involving the EVs, the buffer, and the grid. Consequently, the energy transactions at each time step k are structured accordingly:

$$\Delta E_i(k) = \delta_{1i}(k)\Delta E_{1i}(k) + \delta_{2i}(k)\Delta E_{2i}(k), \quad \forall i \in \{1, \cdots, n\} \cup \{b\} \tag{5.34}$$

where $\delta_{1i}(k)$ and $\delta_{2i}(k)$ are binary variables and:

$$\delta_{1i}(k) + \delta_{2i}(k) \leq 1, \tag{5.35a}$$
$$0 < \Delta E_{1i}(k) \leq \Delta E_{\max,i}(k), \tag{5.35b}$$
$$\Delta E_{\min,i}(k) \leq \Delta E_{2i}(k) < 0 \tag{5.35c}$$

$\Delta E_{\max,i}(k)$ and $\Delta E_{\min,i}(k)$ are the maximum and minimum energy that can be assigned to the i-th battery at time k. In designing the C_2 controller, the battery model must be considered, as the SOC is critical for ensuring feasible decision-making. Specifically, if a battery is at its minimum SOC, assigning a negative energy packet (i.e., setting $\delta_{2i}(k) = 1$) is infeasible. Similarly, if the battery is fully charged, there is no remaining capacity to accept additional energy (i.e., $\delta_{1i}(k) = 1$). The battery dynamics expressed in terms of energy packets are given by:

$$x_i(k+1) = x_i(k) + T_{s_2} \cdot \gamma_i \Delta E_i(k), \quad i \in \{1, \cdots, n\} \cup \{b\} \tag{5.36}$$

γ_i is constant and depends on the physical specifications of the battery, and T_{s_2} is the sampling time. The models presented in Eq. (5.31) and Eq. (5.36) are employed during the controller design to ensure that the battery SOC remains within its allowable bounds. A similar modeling strategy is employed to represent the energy packets exchanged with the power grid. Importantly, the price asymmetry between the energy purchased from the grid and the energy sold back to it is a key consideration in the control synthesis. Consequently, following the same approach used for the batteries in Eq. (5.34), the model for the grid energy packets is formulated as follows:

$$\Delta E_g(k) = \delta_{1g}(k)\Delta E_{1g}(k) + \delta_{2g}(k)\Delta E_{2g}(k) \tag{5.37}$$

where:

$$\delta_{1k}^{[g]} + \delta_{2k}^{[g]} \leq 1, \tag{5.38a}$$
$$0 < \Delta E_{1g}(k) \leq \Delta E_{\max,b}(k) + \sum_{i=1}^{n} \Delta E_{\max,i}(k), \tag{5.38b}$$
$$\Delta E_{\min,b}(k) + \sum_{i=1}^{n} \Delta E_{\min,i}(k) \leq \Delta E_{2g}(k) < 0 \tag{5.38c}$$

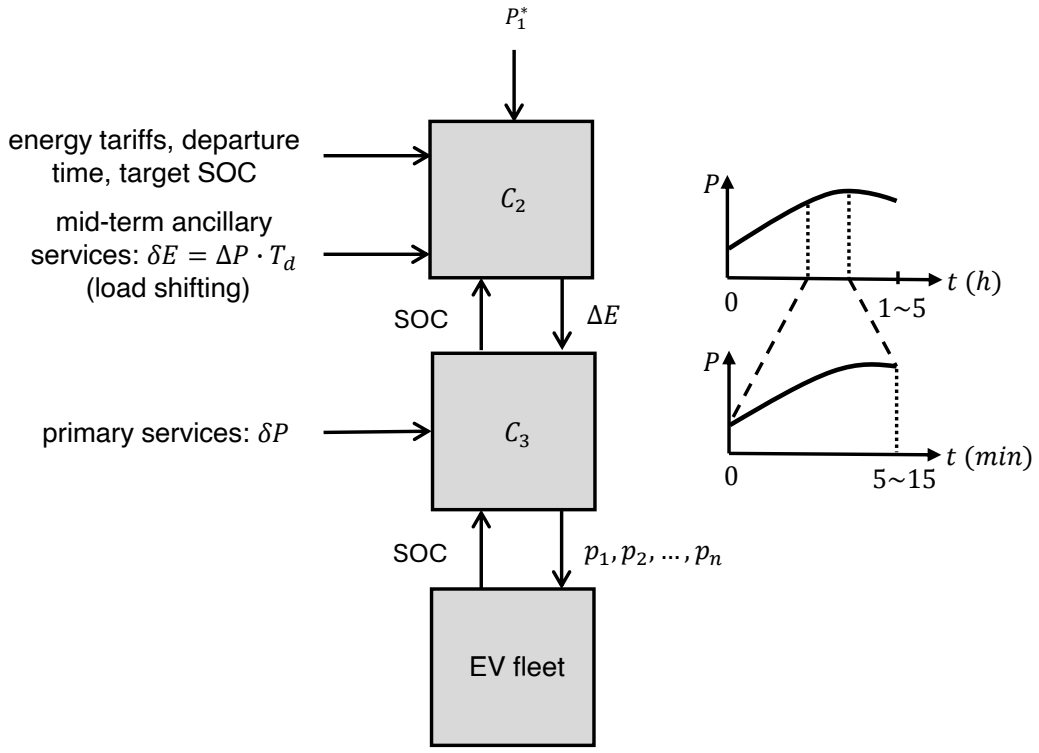


FIGURE 5.12: Hierarchical control structure for EV fleet.

The proposed control architecture and its temporal strategy are illustrated in Fig. 5.12 and Fig. 5.13, respectively, where decision-making is structured into two hierarchical levels. At the upper level, the controller C_2 is implemented as an MPC that receives the day-ahead power schedule and ancillary service requirements. It determines the optimal allocation of battery energy packets $\Delta E_i(k)$ over its prediction horizon. This controller employs the concept of multi-asset optimization to determine the optimal direction of energy flow, thereby maximizing the economic return of the charging hub. Additionally, C_2 must remain responsive to ancillary grid services, such as load shifting requests. At the lower level, the controller C_3 operates with finer temporal resolution and computes the optimal battery power profiles based on the energy packet allocations provided by C_2, as depicted in Fig. 5.13. Furthermore, C_3 strives to ensure that the resulting power profiles remain as close as possible to the preferred charging trajectories of the EVs.

5.5.1 The High-level Controller (C_2)

At the higher level of the proposed control architecture, the controller C_2 is primarily responsible for planning and making long-term decisions. As a predictive controller, it determines the optimal energy trading strategy, while simultaneously responding to grid-related service demands. The formulation of this controller is complete once its objective function and constraints are fully specified. Since the objective of C_2 is to schedule energy transactions and allocate energy packets among the batteries to maximize the overall economic benefit of the charging hub, all costs and revenues associated with energy trading must be incorporated into its objective function. The financial gain of the charging hub increases when energy is sold to the grid or connected EVs, and when energy is stored in the buffer. Conversely, purchasing energy

from the grid or EVs or reducing stored buffer energy leads to economic loss. The cost function includes both gain and loss terms and is formulated as follows:

$$J_2 = \sum_{k=1}^{N_2} \sum_{j=1}^{n+2} \left(W_{1j}(k)\delta_{1j}(k)\Delta E_{1j}(k) + W_{2j}(k)\delta_{2j}(k)\Delta E_{2j}(k) \right) \tag{5.39}$$

where N_2 is the prediction horizon, $W_{1j}(k)$ is the tariff for selling energy (storing in the case of buffer), and $W_{2j}(k)$ is the tariff for buying energy (discharging in the case of buffer). Also, the energy tariffs for EVs are considered equal: $W_{11}(k) = W_{12}(k) = \cdots = W_{1n}(k)$, $W_{21}(k) = W_{22}(k) = \cdots = W_{2n}(k)$. The subscript $n+1$ stands for the buffer (b) and $n+2$ for the grid (g). The controller determines the optimal allocation of energy packets based on the predicted energy prices over its prediction horizon of length $N_2 \cdot T_{s_2}$. It identifies the most advantageous trajectory for energy flow to maximize the economic return of the charging hub. It is important to note that this formulation assumes perfect foresight of energy prices and the arrival times of EVs within the prediction horizon of C_2.

We also assume the flow of energy between the grid and batteries is lossless. As shown in Fig. 5.14, the positive direction of energy for each player (grid, EVs, and buffer) is towards that player. Thus, the conservation of energy reads:

$$\Delta E_g(k) + \Delta E_c(k) + \Delta E_b(k) = 0, \quad k = 1, \cdots, N_2 \tag{5.40}$$

$\Delta E_c(k)$ is the summation of the energy of all EVs ($\sum_{i=1}^{n} \Delta E_i(k)$). Also, the sequence of energy packets assigned throughout the prediction horizon must ensure that the target SOC of each EV is met by the time of its departure from the charging hub. As a result, the drivers' charging requests must be incorporated into the problem formulation as explicit constraints to ensure the effectiveness of the control strategy. Accordingly, the allocation of energy packets across the horizon must supply the required energy for each EV:

$$\sum_{k=k_{a_i}}^{k_{d_i}} \Delta E_i(k) = \Delta E_i, \quad \forall i \in \{1, \cdots, n\} \tag{5.41}$$

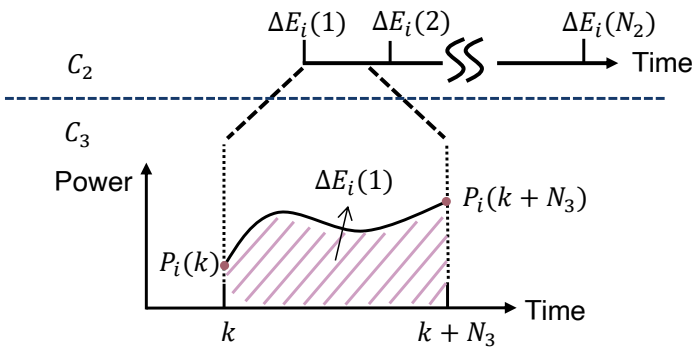


FIGURE 5.13: The sequence of energy packets (from C_2) and the first corresponding power profile (from C_3).

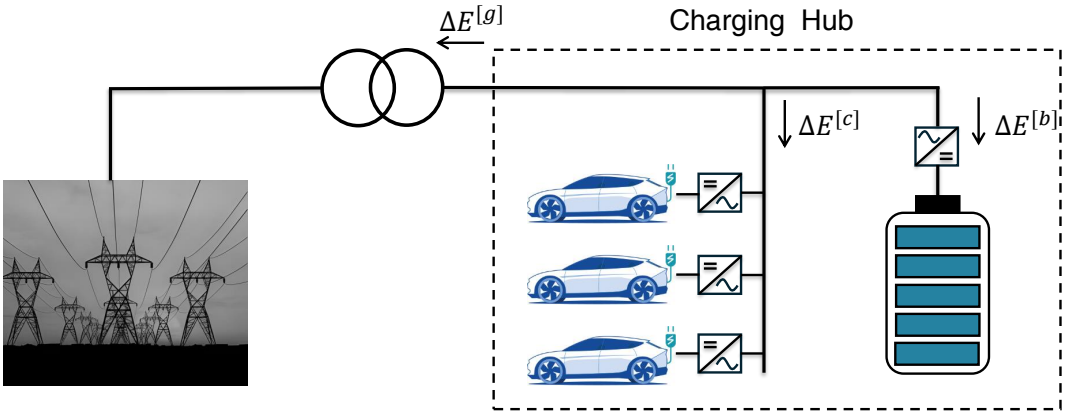


FIGURE 5.14: Architecture of the charging hub.

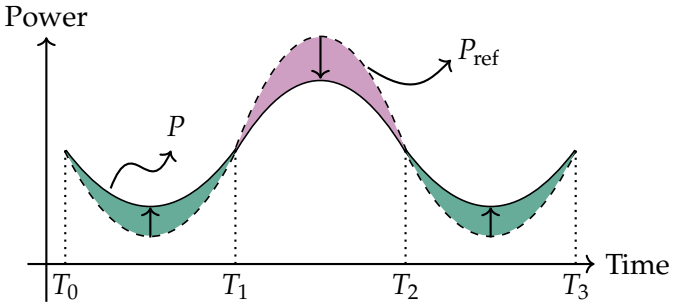


FIGURE 5.15: Reference power profile (P_{ref}) and real power consumption (P). The red area shows the amount of energy that should be shifted to the off-peak time, and the green area shows the energy that has been shifted.

k_{a_i} and k_{d_i} are the time that the i-th EV arrives and departs, respectively. The charging hub predicts its power consumption for the upcoming day and submits a corresponding power request to the grid based on this consumption profile. An example of this profile is depicted in Fig. 5.15, denoted as P_{ref}. If the charging hub's power consumption exceeds this requested profile, it will incur a penalty from the grid. As a result, the charging hub's power consumption must be constrained to this profile, with exceptions allowed when ancillary services are requested. For instance, in the case of a load-shifting demand from the grid during off-peak periods, the charging hub can exceed its usual consumption during peak times to compensate for the shortfall, as illustrated by the P profile in Fig. 5.15. This demand satisfaction is incorporated into the problem formulation as a constraint, as follows:

$$\sum_{k=k_1}^{k_2} \left(P_{\text{ref},k} \cdot T_{s_2} - \sum_{i=1}^{n+1} \Delta E_i(k) \right) \geq \delta E = \Delta P \cdot T_d \tag{5.42}$$

where $n+1$ stands for the buffer (b), k_1 is the time sample that the demand begins k_2 is the time sample that the demand terminates, δE is the amount of energy required to be shifted, and ΔP is the maximum power allowed during the demand interval,

and T_d is the duration of the demand. Eq. (5.42) concludes the synthesis of the high-level controller C_2. The whole optimization problem is then formulated as:

$$\max_{\delta_{1j}(k),\delta_{2j}(k),\Delta E_{1j}(k),\Delta E_{2j}(k)} J_2 \tag{5.43}$$

$$\text{s.t.} \quad \text{Eq. (5.34)} - \text{(5.38)}, \text{(5.40)}, \text{(5.41)}, \text{(5.42)}$$

which is an MIBP. To reduce problem complexity, a set of relaxation constraints can be added to transform the problem into an MILP. Suppose we define new auxiliary variables to capture bilinear terms:

$$\begin{cases} z_{1j}(k) := \delta_{1j}(k)\Delta E_{1j}(k), \\ z_{2j}(k) := \delta_{2j}(k)\Delta E_{2j}(k). \end{cases} \tag{5.44}$$

The optimization problem can be equivalently reformulated as an MILP if the following holds for the selling energy:

$$z_{1j}(k) \geq 0, \tag{5.45a}$$
$$z_{1j}(k) \leq \delta_{1j}(k) \cdot \Delta E_{\max,j}(k), \tag{5.45b}$$
$$z_{1j}(k) \geq \Delta E_{1j}(k) - (1 - \delta_{1j}(k))\Delta E_{\max,j}(k), \tag{5.45c}$$
$$z_{1j}(k) \leq \Delta E_{1j}(k), \tag{5.45d}$$

and the following holds for buying energy:

$$z_{2j}(k) \leq 0, \tag{5.46a}$$
$$z_{2j}(k) \geq \delta_{2j}(k) \cdot \Delta E_{\min,j}(k), \tag{5.46b}$$
$$z_{2j}(k) \leq \Delta E_{2j}(k) - (1 - \delta_{2j}(k))\Delta E_{\min,j}(k), \tag{5.46c}$$
$$z_{2j}(k) \geq \Delta E_{2j}(k). \tag{5.46d}$$

The relaxed version of the original optimization problem in the form of MILP is then formulated as follows:

$$\max_{v} J_2 = \sum_{k=1}^{N_2} \sum_{j=1}^{n+2} \left(W_{1j}(k) z_{1j}(k) + W_{2j}(k) z_{2j}(k) \right) \tag{5.47a}$$
$$\text{s.t.: } \Delta E_j(k) = z_{1j}(k) + z_{2j}(k) \tag{5.47b}$$
$$\text{Eq. (5.34)} - \text{(5.38)}, \text{(5.40)}, \text{(5.41)}, \text{(5.42)}, \text{(5.45)}, \text{(5.46)} \tag{5.47c}$$

where again, the superscript $n+1$ stands for the buffer (b) and $n+2$ stands for the grid (g). Also, the matrix of decision variables is $v = [v_j(k)]_{6(n+2)\times N_2}$ where $v_j(k) = [\delta_{1j}(k)\ \delta_{2j}(k)\ \Delta E_{1j}(k)\ \Delta E_{2j}(k)\ z_{1j}(k)\ z_{2j}(k)]^T$.

5.5.2 The Low-level Controller (C_3)

The higher-level controller is responsible for planning the long-term energy consumption of the charging hub and determining the optimal energy packets to allocate. However, the battery's actuation signal is expressed in terms of power (or current), which must be specified in smaller steps to charge or discharge the battery effectively. The low-level controller ensures that the optimal energy packets are achieved by the end of its horizon, while striving to maintain the charging (or discharging) profile as closely as possible to the recommended power profile for

the batteries. In essence, the high-level controller sets the desired energy level for the batteries, while the low-level controller determines the optimal charging (or discharging) trajectory to reach that level, as depicted in Fig. 5.13. To ensure that the required energy is successfully stored in the EVs, the following equality is imposed as a constraint for the C_3 controller:

$$T_{s_3} \cdot \sum_{k=1}^{N_3} P_i(k) = \Delta E_i(k), \quad \forall i \in \{1, \cdots, n\} \tag{5.48}$$

N_3 is the prediction horizon, and $P_i(k)$ is the power consumption of the i-th battery at time k. Moreover, the total power that the charging hub consumes must not violate the upper bound set by the grid. Otherwise, the charging hub will deteriorate the power balance in the grid and will be punished with a fine.

$$\sum_{i=1}^{n} P_i(k) + P_b(k) - s_k \leq P_{\max,k} - \delta p_k, \quad s_k \geq 0, \tag{5.49}$$

$P_b(k)$ denotes the power of the buffer, $P_{\max,k}$ is the maximum power allowed to consume, δp_k is the primary demand, and s_k is a slack variable, all at time k. Furthermore, there is a bound on the amount of power a battery can take or release, which needs to be taken into account in the lower-level controller design:

$$P_{\min,i}(k) \leq P_i(k) \leq P_{\max,i}(k), \quad \forall i \in \{1, \cdots, n\} \cup \{b\}, \tag{5.50}$$

$P_{\min,i}(k)$ and $P_{\max,i}(k)$ are minimum and maximum power of the i-th battery at time k. The dynamic model of the battery must be incorporated into the synthesis of the C_3 controller to ensure that the SOC of the batteries remains within the predefined operating range. Thus, Eq. (5.50) serves as one of the constraints in the optimization problem for the controller. The objective of the C_3 controller is to align the energy consumption of the batteries with the optimal energy packets, while generating a power profile that closely follows the recommended charging (or discharging) trajectory for the batteries. Therefore, the problem formulation for the MPC in the C_3 controller is as follows:

$$\min_{s_k, P_i(k)} J_3 = \sum_{k=1}^{N_3} \Big(\sum_{i=1}^{n+1} \lambda_i \cdot (P_i(k) - P_{\text{rec},i})^2 + s_k \Big) \tag{5.51a}$$

$$\text{s.t. Eq. (5.31), (5.48), (5.49), (5.50)} \tag{5.51b}$$

$$0.1 \leq x_i(k) \leq 0.9, \tag{5.51c}$$

where $P_{\text{rec},i}$ is the recommended charging (or discharging) power of the i-th battery at time k, which depends on its SOC. λ_i is a design parameter that should be tuned according to the performance specifications. For a safe and reliable operation of the batteries, we assume that their SOC is limited between 10% and 90% and have added it to the optimization problem. The optimization problem has a quadratic cost function with linear constraints and can be easily solved with most quadratic optimization toolboxes.

TABLE 5.2: **Simulation parameters of the EV charging hub for use-case II.**

Parameter	Value (unit)	Parameter	Value (unit)
$P_{\max}$	266.67 kW	n	100
T_d	1 h	N_2	20
λ_i for all i	0.1	N_3	45
T_{s_2}	15 min	T_{s_3}	20 s

5.6 Simulation Results of Use-Case II: Bidirectional Charging Hub

In this section, a simulation of the proposed control algorithm for load-shifting in a charging hub with 100 EVs and one buffer is conducted. The MILP has been solved using the YALMIP toolbox in MATLAB©, with Gurobi© as the underlying solver. The EVs have random arrival and departure times, as well as varying initial SOC values. The simulation begins at 10:00:00 and ends at 15:00:00. The number of EVs in the charging hub peaks between 12:00:00 and 13:00:00, resulting in increased energy demand. Fig. 5.16 illustrates the energy consumption and cumulative benefit profiles of the charging hub during this interval, with and without demand. As shown in green, when there is a demand for load-shifting, the load between 12:00:00 and 13:00:00 is shifted to off-peak hours. The total requested demand is $\delta E = 450\,\text{kWh}$, and this flexibility provision significantly alleviates the strain on the grid during the peak period. The second subplot demonstrates the economic profit of the charging hub over time. Considering the demand, since some of the total energy consumption is shifted to the interval before the demand, the energy cost is higher, and the benefit is lower. The demand begins at 12:00:00, the time at which the economic benefit curve with demand (green curve) approaches the benefit curve without demand (black curve). The green curve exceeds the black curve during the demand with a significant jump and stays above until the end of the time. This is due to not only an optimal strategy adopted by the hierarchical controller but also due to the reward of service provision.

The prediction horizon for the C_2 controller is 5 hours, with a sampling time of 15 minutes. The parameters of the simulation are listed in the Table. 5.2. Fig. 5.17 shows the power profile of the charging hub and the variation in the SOC of some of the EVs. It is assumed that the recommended charging (or discharging) power for the EVs ($P_{\text{rec},i}$) remains constant. The green power profile in the upper subplot illustrates reduced consumption during peak times when there is a demand, with increased consumption outside of this interval. The EVs that stay longer in the charging hub and can provide energy flexibility are charged during off-peak periods (before and after the demand). In contrast, EV 95, which arrives at 10:30:00 and departs at 14:15:00, offers flexibility during peak hours by allowing for a half-hour discharge. A similar behavior is observed with EV 23. EVs 11 and 49 charge mostly but pause their charging briefly during the demand window. These results confirm that the proposed control structure efficiently allocates energy packets based on the available flexibility to meet the demand.

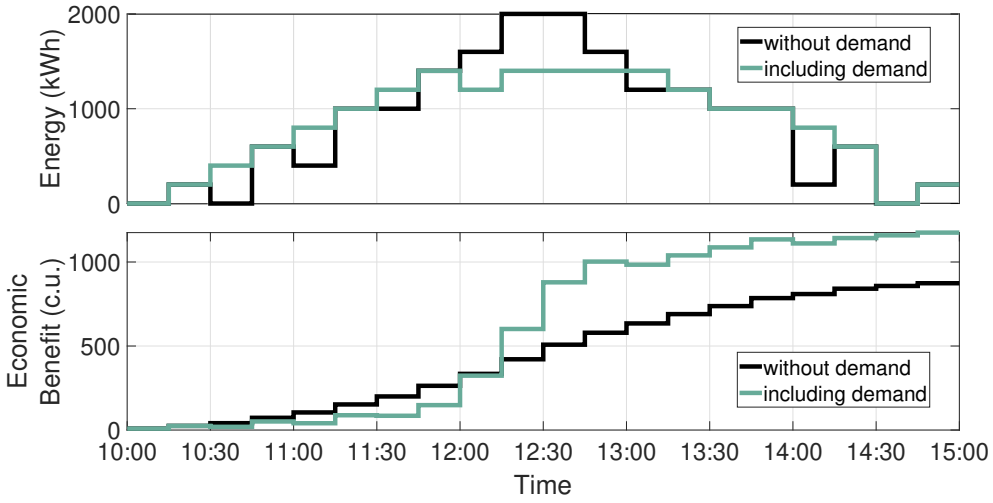


FIGURE 5.16: Energy profile and economic benefit of the charging hub when there is no demand and when there is a demand for load shifting in Scenario II. The control strategy shifts the consumption to the off-peak time. The economic benefit of the charging hub is greater when the demand is satisfied.

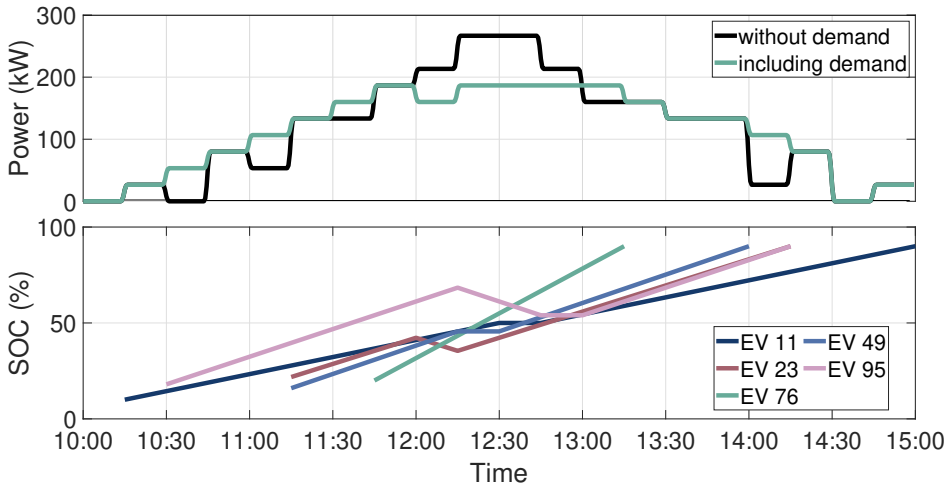


FIGURE 5.17: Power profiles of the charging hub when there is no demand and when there is a demand in Scenario II. The SOC curves of some random EVs in the charging hub are also demonstrated in the lower subplot. The EVs with less parking time have little to no participation in DR.

5.7 Conclusion

This chapter has presented a comprehensive hierarchical energy management framework tailored for the coordination of primary and mid-term energy demands in complex, dynamic environments. By introducing a modular and scalable structure composed of three control layers—strategic planning (C_1), mid-level operational adjustment (C_2), and low-level real-time control (C_3)—the proposed approach successfully bridges the temporal and functional gaps between long-term market planning

and immediate system needs. This tiered design empowers contributive consumers, such as industrial plants and EV charging stations, to participate in multiple energy markets while maintaining operational reliability and flexibility.

The central focus of this chapter has been the design and implementation of the C_2 and C_3 controllers. These two levels serve as the operational backbone of the framework, reacting to short-term market fluctuations and real-time service demands while preserving the broader strategic vision set by the higher-level C_1 controller. The C_2 controller manages intraday decisions by allocating energy packets and adjusting process setpoints based on mid-term demand signals and system feedback, thus enabling mid-term responsiveness to market volatility and internal system changes. Simultaneously, the C_3 controller ensures rapid compliance with primary services, such as frequency regulation or immediate load adjustments, through a fine-grained, real-time control horizon. A key contribution of this hierarchical architecture is its ability to decouple long-horizon planning from real-time operation, facilitating economic optimization and technical feasibility. While many traditional energy management strategies struggle to adapt across diverse time scales or suffer from excessive computational complexity, this framework ensures responsiveness without sacrificing planning integrity. It allows for the optimal deployment of different types of flexibility—both binary and continuous, from energy buffers and flexible consumers—to be achieved according to their characteristics and constraints. Furthermore, the chapter provided a generic scheme that situates the proposed control architecture within the broader context of energy market operations. By acknowledging the interactions between day-ahead planning, intraday optimization, and real-time balancing, this work contributes to the growing body of literature that advocates for cross-market integration and multi-temporal optimization. Importantly, it also underscores the role of industrial and mobile energy actors not merely as passive consumers but as active contributors to energy system stability. Although the day-ahead controller C_1 is not synthesized within the scope of this thesis, its theoretical outline has been discussed to complete the conceptual framework. Future work may pursue the development of C_1 using stochastic optimization or machine learning techniques that exploit historical and real-time data to construct robust and predictive planning models. This advancement would fully realize the vision of end-to-end hierarchical energy management—from forecasting and procurement to deployment and feedback. In conclusion, the hierarchical control framework developed in this chapter advances the field of intelligent energy management for flexible consumers. By combining predictive control methods, real-time responsiveness, and market integration, it offers a viable pathway for managing energy assets in a distributed, uncertain, and market-driven energy landscape. The proposed methodology is particularly relevant for industrial facilities and EV infrastructure seeking to enhance their economic performance and DR participation. The research of this chapter has been partly conducted in the context of the eCells project, funded by the German Federal Ministry for Economic Affairs and Energy (BMWE), from 2024 to 2027, under Grant 03El4088A.

Chapter 6

Conclusion and Future Work

6.1 Overview of Research Objectives and Methodology

This thesis has investigated the integration of flexible energy consumers into modern energy markets through the development of advanced control strategies. The central objective was to design model-based control frameworks that enable industrial processes and EVs charging infrastructure to respond dynamically to market signals, contribute to grid stability, and optimize energy costs—all without compromising operational performance. To achieve this goal, the research focused on modeling, controlling, and optimizing flexible energy assets under both continuous and binary operational constraints.

The thesis has focused on two main applications, each addressing a distinct but complementary application domain:

- The first part has concentrated on demand response control in beverage production lines, where production machines, some operating in binary mode, are modeled and controlled to shift or reduce energy consumption in response to energy market requirements. Generic MPC-based strategies are introduced to target not only short-term power flexibility (primary control) but also optimizing energy consumption over longer periods (secondary control), while considering production goals and associated flexibility costs.

- The second part has addressed the optimal control strategy for bidirectional EV charging hubs, where a hierarchical energy management system has been developed to facilitate load shifting, energy trading, and provision of ancillary services. A two-level optimization approach is proposed, with a strategic layer for energy packet allocation and a real-time layer for charging/discharging decisions. The architecture has integrated MPC techniques and multi-asset optimization to facilitate participation in multiple energy markets while preserving vehicle user requirements and grid constraints.

Across both domains, the common methodological thread is using MPC as the core decision-making tool. MPC is chosen for its ability to handle multivariable systems, incorporate predictions, and enforce constraints over finite horizons—qualities essential for managing flexibility in complex energy systems. In addition, the thesis has introduced hierarchical control structures that separate planning and operational decisions across different time scales, reflecting the layered structure of modern electricity markets.

Simulation-based validations are carried out using realistic scenarios, system parameters, and market profiles to demonstrate the feasibility and effectiveness of the proposed approaches. The results confirm that both industrial plants and EV hubs

can provide valuable flexibility services while maintaining their primary operational functions.

In summary, the research presented in this thesis contributes novel control frameworks for participatory energy systems, offering scalable and market-integrated solutions for demand-side flexibility. The methodologies developed here can support the broader transition toward decentralized and resilient energy systems, particularly in the context of increasing renewable energy integration and electrification.

6.2 Summary of Key Findings

The research has led to several key findings that advance the understanding and application of flexible energy management in both industrial and transportation contexts. These findings are summarized below, corresponding to the two primary case studies and the general methodological contributions.

1. Optimal Control of Beverage Production Lines

One of the main findings is that industrial processes, such as beverage production lines, which are traditionally viewed as rigid and inflexible, can be controlled to provide significant energy flexibility. The developed MPC-based strategies showed that:

- Machines with binary and continuous operation modes can be effectively modeled to participate in DR programs.
- A short-horizon control strategy targeting power reduction can enable these machines to contribute to primary frequency control without violating process constraints.
- Production losses due to flexibility provision can be quantified and controlled using performance indices, enabling plant operators to participate in energy markets with clear expectations.
- Both long and short-term service provision are possible via the proposed unified hierarchical structure for primary and secondary demands

These findings demonstrate that, with suitable control design, even tightly constrained processes can participate actively in the energy market.

2. Energy Management for EV Charging Hubs

The second major finding concerns the development and validation of generic control frameworks for bidirectional EV charging hubs. Key insights include:

- A two-level architecture can effectively manage energy packets and charging/discharging decisions across different time scales, aligning with intraday and balancing market signals.
- The system can dynamically allocate charging loads, allowing some EVs to temporarily adjust their charging patterns to support grid needs and compensate by shifting their charging to off-peak intervals.

- The integration of an energy buffer into the system enhances its responsiveness and flexibility, especially when providing ancillary services such as load shifting.
- Multi-asset optimization and MPC, when combined, offer a robust mechanism for coordinating energy transactions while respecting technical constraints and user requirements.

This work confirms that EV charging hubs are not merely passive loads, but can also actively offer flexibility to the grid, provided that advanced control mechanisms are in place.

3. Methodological Contributions

Beyond the application-specific insights, the thesis also provides general methodological contributions:

- The hierarchical control design, separating long-term planning and short-term execution, provides a scalable approach for systems operating under multiple time-scale market signals.
- The formulation of flexibility metrics and their incorporation into the optimization problem bridges the gap between physical system behavior and market-oriented decision-making.
- The control designs are computationally efficient and implementable in real-world systems, offering practical pathways toward deployment.

To enable a clear evaluation of the proposed approaches, it is important to explicitly state that the controllers developed in this thesis are not intended as mutually competing solutions, but rather as goal-oriented control strategies designed for different operational requirements, market products, and time scales. Each controller is formulated to address a specific control objective under a given set of system constraints, ranging from fast power tracking for demand response to energy-oriented optimization and multi-asset coordination. Consequently, the performance of each controller is assessed with respect to its intended purpose, rather than through a direct one-to-one comparison across all approaches. To clarify the design goals, requirements, and application scenarios of all proposed controllers, Table 6.1 summarizes their main characteristics and indicates which controller should be employed under which operational conditions.

6.3 Implications for Industry and Energy Markets

The proposed methodologies highlight a transformative opportunity for industrial plants and EV charging hubs to become active participants in modern energy systems. Traditionally treated as passive consumers, these entities can now provide valuable flexibility services to the grid without compromising their operational objectives. For industrial production lines, particularly in energy-intensive sectors, integrating MPC with flexibility quantification enables dynamic adaptation to energy market signals. This not only opens up new revenue streams through DR programs but also promotes process-aware energy contracting, allowing the industrial loads to make a balance between production performance and energy service contributions.

TABLE 6.1: **Overview of the proposed controllers, their objectives, and application scenarios**

Controller	System	Primary objective	Time scale	Market level	Recommended application
C_3	Beverage plant	Optimal power consumption and primary service provision	Seconds	Balancing market	Short-term grid support with minimal production disruption
C_2 (single-layer MPC)	EV charging hub	Energy trading and load-shifting	Minutes to hours	Spot/balancing market	Mid-term grid support with financial revenue
C_3 (double-layer MPC)	EV charging hub	Energy trading and primary service provision	Seconds	Balancing market	Short-term grid support with financial revenue
Hierarchical controller ($C_2 - C_3$)	Beverage plant + battery	Optimal power consumption and ancillary service provision	Seconds to several hours	Spot/balancing market	Short-term/mid-term grid support with financial revenue
Hierarchical controller ($C_2 - C_3$)	EV charging hub	Energy trading and ancillary service provision	Seconds to several hours	Spot/balancing market	Short-term/mid-term grid support with financial revenue

For EV charging infrastructure, especially those equipped with bidirectional capabilities, the proposed hierarchical control framework enables coordinated participation in both energy and ancillary markets. Charging hubs can respond in real-time to grid needs while maintaining user preferences, effectively acting as distributed energy storage systems. These findings are particularly relevant in the context of increasing renewable energy penetration, where load flexibility and fast-response capabilities are crucial. Furthermore, the thesis emphasizes the need for updated market designs and incentive mechanisms that accommodate such responsive consumers. By aligning control algorithms with market operations, this research supports the broader energy transition and encourages the adoption of flexible, intelligent systems across industrial and electromobility sectors.

6.4 Future Research Directions

While substantial advancements in integrating flexible industrial loads and EV charging hubs into energy markets are presented, several ways remain open for future exploration. These directions aim to extend the applicability, robustness, and practicality of the proposed methodologies.

First, future work could explore integrating stochastic and robust optimization techniques into the control frameworks to handle uncertainties more effectively. Variations in production schedules, electricity prices, and EV user behavior introduce significant uncertainty that could be better addressed through probabilistic or scenario-based modeling approaches. Embedding such methods into MPC could enhance the reliability of decision-making under real-world variability. The main idea is to consider the future price of energy as a random variable with a known probability distribution or multiple discrete scenarios. The energy prices are generated by probabilistic models or from historical market data. Then, MPC minimizes the expected cost across the probabilistic scenarios or optimizes a risk measure such as conditional CVaR. As an example, we might have 10 price trajectories for the next 24 hours, each with a probability. The MPC optimizes charging schedules that are optimal on average or that limit the risk of high costs. Moreover, EV arrival times can be modeled as random variables with a probability distribution, and the charging/discharging patterns are planned accordingly. The possible arrival-departure profiles within the horizon of MPC can be generated using, e.g., historical patterns and time-of-day statistics. The potential scenarios also include not just the time of arrival, but the initial SOC and departure time. MPC minimizes expected cost (or a risk measure like CVaR) across all these scenarios while meeting constraints for each. As real arrivals happen, the MPC re-optimizes in a receding horizon manner, progressively replacing forecasts with actual data.

Second, scaling the developed frameworks to deal with multiple interacting industrial sites or a network of EV charging hubs remains an open challenge. This would require decentralized or distributed control strategies that ensure scalability, privacy, and computational efficiency. Agent-based modeling or cooperative game theory could provide valuable tools in this context. In fact, for several production lines, the large and coupled optimization problem can be split into local MPCs (one per hub or one per line factory) that coordinate the shared/coupling variables. Each line or hub solves its local MPC and shares the coupling variables with other local MPCs recursively until the primary cost is optimized. Also, each hub or set of lines can be modeled as a self-interested agent, to coordinate behaviors via market mechanisms and ensure fair value sharing to keep everyone engaged.

Third, flexibility contracts and energy trade-offs could be further developed to include dynamic pricing and bilateral negotiations. Future research could investigate how digital platforms and smart contracts can support automated flexibility trading between consumers and grid operators, leading to more responsive and decentralized market structures. More precisely, the idea is to move from static, pre-defined DR agreements to dynamic, real-time pricing and bilateral negotiations between hubs/plants and the grid or other market participants. Each EV hub or production site evaluates flexibility offers (e.g., shiftable load, curtailment potential, V2G discharge) against dynamic price signals from the grid or energy market. Flexibility contracts can be parameterized by price vs. quantity curves or time-varying availability windows, allowing participants to adjust offers in response to market fluctuations. Bilateral negotiations allow hubs/plants to trade energy/flexibility directly with other consumers or aggregators, potentially using automated negotiation protocols with constraints on timing, capacity, and risk. For EV charging hubs, users may have different charging priorities or a tolerance for delayed charging. MPC integrates dynamic contracts into the cost function, influencing charging/discharging schedules in real time. For production lines, MPC considers shifting machine operations to periods of high flexibility prices, balancing production loss vs. market revenue.

Fourth, the human and behavioral dimensions of flexibility provision, particularly for EV users, deserve deeper investigation. Understanding user preferences, willingness to participate, and the impact of incentives on behavior will be crucial to implementing the proposed algorithms. Coupling technical control with behavioral economics could yield richer models of participation. It means incorporating EV user behavior and production personnel flexibility into MPC by modeling their preferences, constraints, and willingness to participate. For EV charging hubs, users may have different charging priorities, or tolerance for delayed charging. These qualitative measures can be represented as utility functions, participation probabilities, or soft constraints in the MPC. For production lines, operators' flexibility in scheduling shifts or tolerating minor production delays can be modeled as adjustable constraints or cost penalties reflecting labor agreements or comfort levels. This idea captures realistic participation limits and increases the acceptability of control strategies. Reward points can be integrated into MPC cost functions as additional terms representing expected behavioral response. Behavioral economics models (e.g., prospect theory, bounded rationality) can be utilized to predict how users respond to price signals, delays, or curtailment requests.

Fifth, the role of reactive power in supporting voltage control should be investigated, especially in distribution networks with high shares of renewable generation and flexible loads. While this thesis primarily focuses on active power management, the ability of industrial loads and EV charging hubs to provide reactive power support can significantly enhance voltage stability. Future research could explore how coordinated control strategies can be extended to include reactive power optimization within existing MPC frameworks. Integrating voltage regulation constraints and exploiting inverter capabilities at EV hubs or industrial converters could unlock new ancillary services and improve local voltage profiles. The proposed MPC frameworks can be extended such that they simultaneously optimize active and reactive power. The voltage constraints at key buses in the distribution network can be added to the optimization problem to maintain voltage within acceptable limits. Multi-objective MPC strategies or weighted cost functions can be developed to balance active power optimization (frequency stability, load shifting, energy cost reduction) and reactive power provision (voltage support, local network stability).

Sixth, the proposed algorithms for EV charging hubs, though, consider an abstract model for batteries to reflect the changes in SOC, but the state of health (SOH) is still not considered in the optimization problems. One direction for future research could be to consider the SOH of the batteries in the optimization problem to adopt control strategies that degrade the life of the battery less. First, SOH, as a dynamic variable that evolves according to battery usage (charge/discharge rates, depth of discharge, temperature effects), must be introduced. Then, SOH-dependent constraints or cost penalties in the MPC objective should be included. Higher degradation rates (fast charging, deep discharges) increase cost, and slower, controlled charging preserves battery life but may reduce revenue from flexibility. To quantify the impact of each control action on SOH battery aging models (empirical or physics-based) can be exploited. It is also possible to apply SOH-aware MPC across a fleet of EVs or a network of charging hubs to optimize aggregate flexibility while minimizing fleet-wide degradation. It prioritizes charging/discharging of EVs with higher remaining SOH.

Ultimately, practical validation through pilot projects or industry-academic collaborations will be crucial in bridging the gap between simulation-based results and real-world deployment. Such validation could test the robustness of the algorithms, uncover implementation bottlenecks, and offer critical feedback to refine the control strategies further. By addressing these future directions, research can move closer to realizing the full potential of intelligent demand-side participation in next-generation energy systems.

6.5 Closing Remarks

A comprehensive exploration of how industrial processes and EV charging infrastructures can be transformed from passive consumers into active and intelligent participants in modern energy systems has been targeted in this work. It bridges the gap between physical system operations and economic decision-making in energy markets. The transition toward a decarbonized and decentralized energy landscape will require the seamless integration of responsive demand-side resources. The methodologies developed in the thesis demonstrate that such integration is not only technically feasible but also economically beneficial when approached with the proper optimization tools and market-aligned strategies. While challenges remain in terms of energy price prediction, user preferences, and practical implementation, the results of the research provide a strong foundation upon which future developments can be built. Ultimately, the ability of industrial consumers and charging networks to interact intelligently with the grid will be a defining element of tomorrow's energy systems. The insights and frameworks offered here hope to inspire further research and innovation toward that goal.

Bibliography

[1] K. Karunanithi, S. Saravanan, B. Prabakar, S. Kannan, and C. Thangaraj, "Integration of demand and supply side management strategies in generation expansion planning," *Renewable and Sustainable Energy Reviews*, vol. 73, pp. 966–982, 2017.

[2] S. Gheouany, H. Ouadi, F. Giri, I. Jrhilifa, N. Mounir, et al., "Optimal supply-side and demand-side management strategies for energy efficiency in residential buildings using particle swarm optimization," *IFAC-PapersOnLine*, vol. 58, no. 13, pp. 188–193, 2024.

[3] S. D. Ahmed, F. S. Al-Ismail, M. Shafiullah, F. A. Al-Sulaiman, and I. M. El-Amin, "Grid integration challenges of wind energy: A review," *IEEE Access*, vol. 8, pp. 10 857–10 878, 2020.

[4] M. Shafiullah, S. D. Ahmed, and F. A. Al-Sulaiman, "Grid integration challenges and solution strategies for solar pv systems: A review," *IEEE Access*, vol. 10, pp. 52 233–52 257, 2022.

[5] G. Strbac, "Demand side management: Benefits and challenges," *Energy Policy*, vol. 36, no. 12, pp. 4419–4426, 2008.

[6] M. S. Bakare, A. Abdulkarim, M. Zeeshan, and A. N. Shuaibu, "A comprehensive overview on demand side energy management towards smart grids: Challenges, solutions, and future direction," *Energy Informatics*, vol. 6, no. 1, p. 4, 2023.

[7] A. R. Khan, A. Mahmood, A. Safdar, Z. A. Khan, and N. A. Khan, "Load forecasting, dynamic pricing and dsm in smart grid: A review," *Renewable and Sustainable Energy Reviews*, vol. 54, pp. 1311–1322, 2016.

[8] M. E. Honarmand, V. Hosseinnezhad, B. Hayes, M. Shafie-Khah, and P. Siano, "An overview of demand response: From its origins to the smart energy community," *IEEE Access*, vol. 9, pp. 96 851–96 876, 2021.

[9] H. Karimianfard, M. R. Salehizadeh, and P. Siano, "Economic profit enhancement of a demand response aggregator through investment of large-scale energy storage systems," *CSEE Journal of Power and Energy Systems*, vol. 8, no. 5, pp. 1468–1476, 2022.

[10] C. Dou, X. Zhou, T. Zhang, and S. Xu, "Economic optimization dispatching strategy of microgrid for promoting photoelectric consumption considering cogeneration and demand response," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 3, pp. 557–563, 2020.

[11] S. S. Siddiquee, B. Howard, K. Bruton, A. Brem, and D. T. O'Sullivan, "Progress in demand response and its industrial applications," *Frontiers in Energy Research*, vol. 9, p. 673 176, 2021.

[12] M. H. Shoreh, P. Siano, M. Shafie-khah, V. Loia, and J. P. Catalão, "A survey of industrial applications of demand response," *Electric Power Systems Research*, vol. 141, pp. 31–49, 2016.

[13] N. G. Paterakis, O. Erdinç, and J. P. Catalão, "An overview of demand response: Key-elements and international experience," *Renewable and Sustainable Energy Reviews*, vol. 69, pp. 871–891, 2017.

[14] B. Heydaryan and N. Bajcinca, "Economic model predictive control of industrial demand response," in *2022 XXVIII International Conference on Information, Communication and Automation Technologies (ICAT)*, IEEE, 2022, pp. 1–6.

[15] B. Heydaryan, J. Tousi, and N. Bajcinca, "Model predictive control of industrial demand response for production lines," in *2022 IEEE International Symposium on Advanced Control of Industrial Processes (AdCONIP)*, IEEE, 2022, pp. 151–157.

[16] B. Heydaryan, M. Al Khatib, and N. Bajcinca, "Demand response control of beverage production plants with binary units," *IEEE Access*, vol. 13, pp. 187 121–187 133, 2025.

[17] B. Heydaryan, M. Al Khatib, M. Hess, and N. Bajcinca, "Optimal energy transactions for bidirectional charging stations enabling grid ancillary services," in *2024 European Control Conference (ECC)*, IEEE, 2024, pp. 2003–2009.

[18] B. Heydaryan, M. Al Khatib, and N. Bajcinca, "Optimal control strategy for bidirectional ev charging hubs contributing to energy market," *Journal of Energy Storage*, vol. 119, p. 116 282, 2025.

[19] B. Heydaryan, M. Al Khatib, and N. Bajcinca, "Hierarchical energy management for load shifting in bidirectional charging hubs," in *2025 33rd Mediterranean Conference on Control and Automation (MED)*, IEEE, 2025, pp. 441–446.

[20] Y. M. Ding, S. H. Hong, and X. H. Li, "A demand response energy management scheme for industrial facilities in smart grid," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 4, pp. 2257–2269, 2014.

[21] H. Cui and K. Zhou, "Industrial power load scheduling considering demand response," *Journal of cleaner production*, vol. 204, pp. 447–460, 2018.

[22] L. Scharnhorst, D. Sloot, N. Lehmann, A. Ardone, and W. Fichtner, "Barriers to demand response in the commercial and industrial sectors–an empirical investigation," *Renewable and Sustainable Energy Reviews*, vol. 190, p. 114 067, 2024.

[23] J. Stede, K. Arnold, C. Dufter, G. Holtz, S. von Roon, and J. C. Richstein, "The role of aggregators in facilitating industrial demand response: Evidence from germany," *Energy Policy*, vol. 147, p. 111 893, 2020.

[24] H. Golmohamadi, R. Keypour, B. Bak-Jensen, and J. R. Pillai, "A multi-agent based optimization of residential and industrial demand response aggregators," *International Journal of Electrical Power & Energy Systems*, vol. 107, pp. 472–485, 2019.

[25] A. Allman and Q. Zhang, "Distributed cooperative industrial demand response," *Journal of Process Control*, vol. 86, pp. 81–93, 2020.

[26] X. Zhang, G. Hug, J. Z. Kolter, and I. Harjunkoski, "Demand response of ancillary service from industrial loads coordinated with energy storage," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 951–961, 2017.

[27] J. Wang, Y. Shi, and Y. Zhou, "Intelligent demand response for industrial energy management considering thermostatically controlled loads and evs," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 6, pp. 3432–3442, 2018.

[28] X. Huang, S. H. Hong, M. Yu, Y. Ding, and J. Jiang, "Demand response management for industrial facilities: A deep reinforcement learning approach," *IEEE Access*, vol. 7, pp. 82 194–82 205, 2019.

[29] R. Lu, R. Bai, Y. Huang, Y. Li, J. Jiang, and Y. Ding, "Data-driven real-time price-based demand response for industrial facilities energy management," *Applied Energy*, vol. 283, p. 116 291, 2021.

[30] J. Rodríguez-García, C. Alvarez-Bel, J.-F. Carbonell-Carretero, M. Alcázar-Ortega, and E. Peñalvo-López, "A novel tool for the evaluation and assessment of demand response activities in the industrial sector," *Energy*, vol. 113, pp. 1136–1146, 2016.

[31] S. Kiliccote, D. Olsen, M. D. Sohn, and M. A. Piette, "Characterization of demand response in the commercial, industrial, and residential sectors in the united states," *Advances in Energy Systems: The Large-scale Renewable Energy Integration Challenge*, pp. 425–443, 2019.

[32] K. Ma, G. Hu, and C. J. Spanos, "A cooperative demand response scheme using punishment mechanism and application to industrial refrigerated warehouses," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 6, pp. 1520–1531, 2015.

[33] N. Lashmar, B. Wade, L. Molyneaux, and P. Ashworth, "Motivations, barriers, and enablers for demand response programs: A commercial and industrial consumer perspective," *Energy Research & Social Science*, vol. 90, p. 102 667, 2022.

[34] H. Golmohamadi, "Demand-side management in industrial sector: A review of heavy industries," *Renewable and Sustainable Energy Reviews*, vol. 156, p. 111 963, 2022.

[35] S. Mohagheghi and N. Raji, "Dynamic demand response: A solution for improved energy efficiency for industrial customers," *IEEE Industry Applications Magazine*, vol. 21, no. 2, pp. 54–62, 2014.

[36] S. Dehghan-Dehnavi, M. Fotuhi-Firuzabad, M. Moeini-Aghtaie, P. Dehghanian, and F. Wang, "Decision-making tree analysis for industrial load classification in demand response programs," *IEEE Transactions on Industry Applications*, vol. 57, no. 1, pp. 26–35, 2020.

[37] S. Ma, Y. Zhang, Y. Liu, H. Yang, J. Lv, and S. Ren, "Data-driven sustainable intelligent manufacturing based on demand response for energy-intensive industries," *Journal of Cleaner Production*, vol. 274, p. 123 155, 2020.

[38] S. S. Kholerdi and A. Ghasemi-Marzbali, "Interactive time-of-use demand response for industrial electricity customers: A case study," *Utilities Policy*, vol. 70, p. 101 192, 2021.

[39] S. Woltmann and J. Kittel, "Development and implementation of multi-agent systems for demand response aggregators in an industrial context," *Applied Energy*, vol. 314, p. 118 841, 2022.

[40] A. A. Desta, H. Badis, and L. George, "Demand response scheduling in industrial asynchronous production lines constrained by available power and production rate," *Applied Energy*, vol. 230, pp. 1414–1424, 2018.

[41] F. Dababneh, L. Li, R. Shah, and C. Haefke, "Demand response-driven production and maintenance decision-making for cost-effective manufacturing," *Journal of Manufacturing Science and Engineering*, vol. 140, no. 6, p. 061 008, 2018.

[42] N. Gerami, A. Ghasemi, A. Lotfi, L. G. Kaigutha, and M. Marzband, "Energy consumption modeling of production process for industrial factories in a day ahead scheduling with demand response," *Sustainable Energy, Grids and Networks*, vol. 25, p. 100 420, 2021.

[43] F. T. Röben, D. Liu, M. A. Reuter, M. Dahmen, and A. Bardow, "The demand response potential in copper production," *Journal of Cleaner Production*, vol. 362, p. 132 221, 2022.

[44] L. Yun, L. Li, and S. Ma, "Demand response for manufacturing systems considering the implications of fast-charging battery powered material handling equipment," *Applied Energy*, vol. 310, p. 118 550, 2022.

[45] M. Ahang, P. M. S. Seljom, and A. Tomasgard, "Analysis of the impact of demand response on the norwegian energy system," *Energy Systems*, pp. 1–28, 2023.

[46] A. N. F. Ali, M. F. Sulaima, I. A. W. A. Razak, A. F. A. Kadir, and H. Mokhlis, "Artificial intelligence application in demand response: Advantages, issues, status, and challenges," *IEEE Access*, vol. 11, pp. 16 907–16 922, 2023.

[47] X. Wang and Q. Liang, "Energy management strategy for plug-in hybrid electric vehicles via bidirectional vehicle-to-grid," *IEEE Systems Journal*, vol. 11, no. 3, pp. 1789–1798, 2015.

[48] G. Merhy, A. Nait-Sidi-Moh, and N. Moubayed, "Control, regulation and optimization of bidirectional energy flows for electric vehicles' charging and discharging," *Sustainable Cities and Society*, vol. 57, p. 102 129, 2020.

[49] P. Siddhartha, T. Sujeeth, B. Shiva, and J. Ramprabhakar, "Integration of renewable energy sources with power management strategy for effective bidirectional vehicle to grid power transfer," *Procedia Computer Science*, vol. 218, pp. 9–23, 2023.

[50] A. Kazemtarghi, S. Dey, and A. Mallik, "Optimal utilization of bidirectional evs for grid frequency support in power systems," *IEEE Transactions on Power Delivery*, vol. 38, no. 2, pp. 998–1010, 2022.

[51] S. Habib, M. M. Khan, F. Abbas, and H. Tang, "Assessment of electric vehicles concerning impacts, charging infrastructure with unidirectional and bidirectional chargers, and power flow comparisons," *International Journal of Energy Research*, vol. 42, no. 11, pp. 3416–3441, 2018.

[52] M. Lukianov, I. Verbitsky, E. R. Cadaval, and R. Strzelecki, "An overview of bidirectional ev chargers: Empowering traction grid-powered chargers," *Power Systems Research and Operation: Selected Problems III*, pp. 191–230, 2023.

[53] B.-G. Kim, S. Ren, M. Van Der Schaar, and J.-W. Lee, "Bidirectional energy trading and residential load scheduling with electric vehicles in the smart grid," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 7, pp. 1219–1234, 2013.

[54] V. T. Tran, M. R. Islam, K. M. Muttaqi, and D. Sutanto, "An efficient energy management approach for a solar-powered ev battery charging facility to support distribution grids," *IEEE Transactions on Industry Applications*, vol. 55, no. 6, pp. 6517–6526, 2019.

[55] M. Tabari and A. Yazdani, "An energy management strategy for a dc distribution system for power system integration of plug-in electric vehicles," *IEEE Transactions on Smart Grid*, vol. 7, no. 2, pp. 659–668, 2015.

[56] A. Al-Obaidi, H. Khani, H. E. Farag, and M. Mohamed, "Bidirectional smart charging of electric vehicles considering user preferences, peer to peer energy trade, and provision of grid ancillary services," *International Journal of Electrical Power & Energy Systems*, vol. 124, p. 106 353, 2021.

[57] M. Zhang and J. Chen, "The energy management and optimized operation of electric vehicles based on microgrid," *IEEE Transactions on Power Delivery*, vol. 29, no. 3, pp. 1427–1435, 2014.

[58] M. Anand, S. Golshannavaz, W. Ongsakul, and A. Rajapakse, "Incorporating short-term topological variations in optimal energy management of mgs considering ancillary services by electric vehicles," *Energy*, vol. 112, pp. 241–253, 2016.

[59] A. Rajapandiyan and S. Thangavel, "Energy management in distribution system due to the integration of renewable energy powered ev charging infrastructures," in *Power Systems Operation with 100% Renewable Energy Sources*, Elsevier, 2024, pp. 281–297.

[60] D. Li, A. Zouma, J.-T. Liao, and H.-T. Yang, "An energy management strategy with renewable energy and energy storage system for a large electric vehicle charging station," *Etransportation*, vol. 6, p. 100 076, 2020.

[61] F. Kennel, D. Görges, and S. Liu, "Energy management for smart grids with electric vehicles based on hierarchical mpc," *IEEE Transactions on industrial informatics*, vol. 9, no. 3, pp. 1528–1537, 2012.

[62] S. Woo, L. Strobel, Y. Yuan, M. Pruckner, and T. E. Lipman, "Exploring bidirectional charging strategies for an electric vehicle population," *Applied Energy*, vol. 397, p. 126 361, 2025.

[63] O. Palizban and K. Kauhaniemi, "Hierarchical control structure in microgrids with distributed generation: Island and grid-connected mode," *Renewable and Sustainable Energy Reviews*, vol. 44, pp. 797–813, 2015.

[64] W. Jiang, C. Yang, Z. Liu, M. Liang, P. Li, and G. Zhou, "A hierarchical control structure for distributed energy storage system in dc micro-grid," *IEEE Access*, vol. 7, pp. 128 787–128 795, 2019.

[65] C. Wang, B. Wang, and F. You, "Demand response for residential buildings using hierarchical nonlinear model predictive control for plug-and-play," *Applied Energy*, vol. 369, p. 123 581, 2024.

[66] L. Zheng, B. Zhou, Y. Cao, S. W. Or, Y. Li, and K. W. Chan, "Hierarchical distributed multi-energy demand response for coordinated operation of building clusters," *Applied Energy*, vol. 308, p. 118 362, 2022.

[67] C. Moya, W. Zhang, J. Lian, and K. Kalsi, "A hierarchical framework for demand-side frequency control," in *2014 American Control Conference*, IEEE, 2014, pp. 52–57.

[68] P. Paudyal, P. Munankarmi, Z. Ni, and T. M. Hansen, "A hierarchical control framework with a novel bidding scheme for residential community energy optimization," *IEEE Transactions on Smart Grid*, vol. 11, no. 1, pp. 710–719, 2019.

[69] Y.-Q. Bao, Y. Li, Y.-Y. Hong, and B. Wang, "Design of a hybrid hierarchical demand response control scheme for the frequency control," *IET Generation, Transmission & Distribution*, vol. 9, no. 15, pp. 2303–2310, 2015.

[70] M. H. Albadi and E. F. El-Saadany, "A summary of demand response in electricity markets," *Electric power systems research*, vol. 78, no. 11, pp. 1989–1996, 2008.

[71] P. Palensky and D. Dietrich, "Demand side management: Demand response, intelligent energy systems, and smart loads," *IEEE transactions on industrial informatics*, vol. 7, no. 3, pp. 381–388, 2011.

[72] International Energy Agency, *Status of Power System Transformation 2019: Power System Flexibility*. Paris: OECD Publishing, 2019, Licence: CC BY 4.0, ISBN: 92-64-44482-3. DOI: 10.1787/7c49400a-en [Online]. Available: https://www.iea.org/reports/status-of-power-system-transformation-2019

[73] P. Pinson, H. Madsen, et al., "Benefits and challenges of electrical demand response: A critical review," *Renewable and Sustainable Energy Reviews*, vol. 39, pp. 686–699, 2014.

[74] K.-Y. Huang and Y.-C. Huang, "Integrating direct load control with interruptible load management to provide instantaneous reserves for ancillary services," *IEEE Transactions on power systems*, vol. 19, no. 3, pp. 1626–1634, 2004.

[75] Q. Zhang and J. Li, "Demand response in electricity markets: A review," in *2012 9th international conference on the European Energy market*, IEEE, 2012, pp. 1–8.

[76] C. Alvarez, A. Gabaldón, and A. Molina, "Assessment and simulation of the responsive demand potential in end-user facilities: Application to a university customer," *IEEE Transactions on Power Systems*, vol. 19, no. 2, pp. 1223–1231, 2004.

[77] P. Siano, "Demand response and smart grids—a survey," *Renewable and sustainable energy reviews*, vol. 30, pp. 461–478, 2014.

[78] J. Kirkerud, N. O. Nagel, and T. Bolkesjø, "The role of demand response in the future renewable northern european energy system," *Energy*, vol. 235, p. 121 336, 2021.

[79] A.-H. Mohsenian-Rad and A. Leon-Garcia, "Optimal residential load control with price prediction in real-time electricity pricing environments," *IEEE transactions on Smart Grid*, vol. 1, no. 2, pp. 120–133, 2010.

[80] E. F. Camacho, C. Bordons, E. F. Camacho, and C. Bordons, *Constrained model predictive control*. Springer, 2007.

[81] D. Q. Mayne, "Model predictive control: Recent developments and future promise," *Automatica*, vol. 50, no. 12, pp. 2967–2986, 2014.

[82] R. Scattolini and P. Colaneri, "Hierarchical model predictive control," in *2007 46th IEEE conference on decision and control*, IEEE, 2007, pp. 4803–4808.

[83] A. Bemporad and M. Morari, "Control of systems integrating logic, dynamics, and constraints," *Automatica*, vol. 35, no. 3, pp. 407–427, 1999.

[84] J. B. Rawlings, D. Q. Mayne, M. Diehl, et al., *Model predictive control: theory, computation, and design*. Nob Hill Publishing Madison, WI, 2017, vol. 2.

[85] F. Borrelli, A. Bemporad, and M. Morari, *Predictive control for linear and hybrid systems*. Cambridge University Press, 2017.

[86] S. Karaman, R. G. Sanfelice, and E. Frazzoli, "Optimal control of mixed logical dynamical systems with linear temporal logic specifications," in *2008 47th IEEE Conference on Decision and Control*, IEEE, 2008, pp. 2117–2122.

[87] J. K. Scott, M. D. Stuber, and P. I. Barton, "Generalized mccormick relaxations," *Journal of Global Optimization*, vol. 51, no. 4, pp. 569–606, 2011.

[88] A. Mitsos, B. Chachuat, and P. I. Barton, "Mccormick-based relaxations of algorithms," *SIAM Journal on Optimization*, vol. 20, no. 2, pp. 573–601, 2009.

[89] H. D. Sherali and W. P. Adams, *A reformulation-linearization technique for solving discrete and continuous nonconvex problems*. Springer Science & Business Media, 2013, vol. 31.

[90] A. Bärmann, R. Burlacu, L. Hager, and T. Kleinert, "On piecewise linear approximations of bilinear terms: Structural comparison of univariate and bivariate mixed-integer programming formulations," *Journal of Global Optimization*, vol. 85, no. 4, pp. 789–819, 2023.

[91] N. Bajcinca, M. V. Pedrosa, and V. Yfantis, "Mixed-integer model predictive control of hybrid impulsive linear systems," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 6775–6780, 2020.

[92] M. Ellis, H. Durand, and P. D. Christofides, "A tutorial review of economic model predictive control methods," *Journal of Process Control*, vol. 24, no. 8, pp. 1156–1178, 2014.

[93] A. Parisio, E. Rikos, and L. Glielmo, "A model predictive control approach to microgrid operation optimization," *IEEE Transactions on Control Systems Technology*, vol. 22, no. 5, pp. 1813–1827, 2014.

[94] S. Bjarghov et al., "Developments and challenges in local electricity markets: A comprehensive review," *IEEE Access*, vol. 9, pp. 58 910–58 943, 2021.

[95] B. Cheng, T. Levin, Z. Zhou, and A. Botterud, "Electricity markets and long-duration energy storage: A survey of grid services and revenue streams," *Current Sustainable/Renewable Energy Reports*, vol. 12, no. 1, pp. 1–12, 2025.

[96] L. Silva-Rodriguez, A. Sanjab, E. Fumagalli, A. Virag, and M. Gibescu, "Short term wholesale electricity market designs: A review of identified challenges and promising solutions," *Renewable and Sustainable Energy Reviews*, vol. 160, p. 112 228, 2022.

[97] D. Mayorga Gonzalez, J. Myrzik, and C. Rehtanz, "The smart power cell concept: Mastering tso–dso interactions for the secure and efficient operation of future power systems," *IET Generation, Transmission & Distribution*, vol. 14, no. 13, pp. 2407–2418, 2020.

[98] EPEX SPOT, *EPEX SPOT – European Power Exchange*, https://www.epexspot.com/en, Accessed: 2025-07-03, Paris, France: EPEX SPOT SE, 2008.

[99] Réseau de Transport d'Électricité (RTE), *RTE - France's Transmission System Operator*, https://www.rte-france.com/en/home, Accessed: 2024-10-22, 2024.

[100] North American Electric Reliability Corporation (NERC), "Reference document: Balancing and frequency control," NERC, Tech. Rep., 2023, Accessed: 2024-07-21. Accessed: Jul. 21, 2024. [Online]. Available: https://www.nerc.com/comm/RSTC_Reliability_Guidelines/Reference_Document_NERC_Balancing_and_Frequency_Control.pdf

[101] R. J. LeVeque and R. J. Leveque, *Numerical methods for conservation laws*. Springer, 1992, vol. 132.

[102] J. Tousi, M. Al Khatib, and N. Bajcinca, "Model predictive control of demand response for large scale production lines and networks," in *2022 XXVIII International Conference on Information, Communication and Automation Technologies (ICAT)*, IEEE, 2022, pp. 1–6.

[103] J. Tousi, B. Heydaryan, M. Al Khatib, and N. Bajcinca, "Model predictive control strategies for industrial demand response of beverage production lines," in *2023 European Control Conference (ECC)*, IEEE, 2023, pp. 1–6.

[104] "Wie die industrie zukünftig zur netzstabilität beitragen kann." Aktualisiert am: 14.11.2024, energieforschung.de, Accessed: Mar. 3, 2026. [Online]. Available: https://www.energieforschung.de/de/aktuelles/news/2022/wie-die-industrie-zukuenftig-zur-netzstabilitaet-beitragen-kann

[105] J. A. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl, "Casadi: A software framework for nonlinear optimization and optimal control," *Mathematical Programming Computation*, vol. 11, pp. 1–36, 2019.

[106] Bundesnetzagentur, *Description of the balancing process and the balancing markets in germany*, Accessed: 20-Feb-2025. [Online]. Available: https://www.regelleistung.net/en-us/

[107] S. Zhang, X. Guo, X. Dou, and X. Zhang, "A data-driven coulomb counting method for state of charge calibration and estimation of lithium-ion battery," *Sustainable Energy Technologies and Assessments*, vol. 40, p. 100 752, 2020.

[108] K. Zaghib, A. Mauger, and C. Julien, "Rechargeable lithium batteries, from fundamentals to applications; franco, aa, ed," *Woodhead Publ. Ltd: Cambridge, UK*, pp. 319–351, 2015.

[109] C. M. Vertgewall, M. Trageser, M. Kurth, and A. Ulbig, "Modeling probabilistic driving and charging profiles of commercial electric vehicles," *Electric Power Systems Research*, vol. 212, p. 108 538, 2022.

[110] J. Shi, T. Zeng, and S. Moura, "Electric fleet charging management considering battery degradation and nonlinear charging profile," *Energy*, vol. 283, p. 129 094, 2023.

[111] M. Kane, *Analysis: Honda e fast charging dissapoints*, https://insideevs.com/news/512213/honda-e-fast-charging-analysis, Accessed: 2024-07-05.

[112] E. Celebi and J. D. Fuller, "Time-of-use pricing in electricity markets under different market structures," *IEEE Transactions on Power Systems*, vol. 27, no. 3, pp. 1170–1181, 2012.

[113] D. Zhao, S. Coyle, A. Sakti, and A. Botterud, "Market mechanisms for low-carbon electricity investments: A game-theoretical analysis," *IEEE Transactions on Energy Markets, Policy and Regulation*, 2023.

[114] K. Neuhoff, N. Ritter, A. Salah-Abou-El-Enien, and P. Vassilopoulos, "Intraday markets for power: Discretizing the continuous trading?" Energy Policy Research Group - University of Cambridge, Research Report 1609, Mar. 2016.

[115] I. G. Marneris, C. G. Roumkos, and P. N. Biskas, "Towards balancing market integration: Conversion process for balancing energy offers of central-dispatch systems," *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 293–303, 2019.

[116] H. Taher, S. Olmi, and E. Schöll, "Enhancing power grid synchronization and stability through time-delayed feedback control," *Physical Review E*, vol. 100, no. 6, p. 062 306, 2019.

[117] G. Filatrella, A. H. Nielsen, and N. F. Pedersen, "Analysis of a power grid using a kuramoto-like model," *The European Physical Journal B*, vol. 61, pp. 485–491, 2008.

[118] X. Zhao, Y. Xue, and X.-P. Zhang, "Fast frequency support from wind turbine systems by arresting frequency nadir close to settling frequency," *IEEE Open Access Journal of Power and Energy*, vol. 7, pp. 191–202, 2020.

[119] I. M. Alotaibi, M. Abido, and M. Khalid, "Primary frequency regulation by demand side response," *Arabian Journal for Science and Engineering*, vol. 46, no. 10, pp. 9627–9637, 2021.

[120] Recharged Editorial. "Ev car battery capacity: Kwh, range & battery health explained." Accessed: 2026-01-13, Recharged. [Online]. Available: https://recharged.com/articles/ev-car-battery-capacity

[121] EPEX SPOT. "Market results," Accessed: Mar. 11, 2024. [Online]. Available: https://www.epexspot.com/en/market-results

[122] SMARD. "Electricity market data in germany." Accessed: 2025-02-15. [Online]. Available: https://www.smard.de/en

[123] B. Finnah, J. Gönsch, and F. Ziel, "Integrated day-ahead and intraday self-schedule bidding for energy storage systems using approximate dynamic programming," *European Journal of Operational Research*, vol. 301, no. 2, pp. 726–746, 2022.

[124] N. Nolzen, A. Ganter, N. Baumgärtner, F. J. Baader, L. Leenders, and A. Bardow, "Where to market flexibility? integrating continuous intraday trading into multi-market participation of industrial multi-energy systems," *Computers & Chemical Engineering*, p. 109 026, 2025.

[125] M. Kremer, R. Kiesel, and F. Paraschiv, "An econometric model for intraday electricity trading," *Philosophical Transactions of the Royal Society A*, vol. 379, no. 2202, p. 20 190 624, 2021.

[126] H. Golmohamadi and R. Keypour, "Stochastic optimization for retailers with distributed wind generation considering demand response," *Journal of Modern Power Systems and Clean Energy*, vol. 6, no. 4, pp. 733–748, 2018.

[127] J. Meese, B. Dahlmann, M. Zdrallek, and A. Voelschow, "Intraday redispatch-optimal scheduling of industrial processes at day-ahead and continuous intraday market," in *International ETG Congress 2017*, VDE, 2017, pp. 1–6.

[128] Y. Wang et al., "A comprehensive review of battery modeling and state estimation approaches for advanced battery management systems," *Renewable and Sustainable Energy Reviews*, vol. 131, p. 110 015, 2020.

[129] M. Behrendt, N. Bajcinca, F. Zenith, and U. Krewer, "Model predictive control of a hybrid fuel cell & battery power system," *IFAC Proceedings Volumes*, vol. 45, no. 15, pp. 131–136, 2012.

List of Author's Publications

[1] Behzad Heydaryan, Mohammad Al Khatib, and Naim Bajcinca, "Demand response control of beverage production plants with binary units". In: *IEEE Access* 13 (2025), pp. 187121-187133.

[2] Behzad Heydaryan, Mohammad Al Khatib, and Naim Bajcinca, "Hierarchical energy management for load shifting in bidirectional charging hubs," In: *33rd Mediterranean Conference on Control and Automation (MED)*, Tangier, 2025, pp. 441-446.

[3] Behzad Heydaryan, Mohammad Al Khatib, and Naim Bajcinca, "Optimal control strategy for bidirectional EV charging hubs contributing to energy market". *Journal of Energy Storage*, vol. 119, pp. 116282, 2025.

[4] Behzad Heydaryan, Mohammad Al Khatib, Markus Hess, and Naim Bajcinca, "Optimal energy transactions for bidirectional charging stations enabling grid ancillary services". In: *European Control Conference (ECC)*, Stockholm, 2024, pp. 2003-2009.

[5] Behzad Heydaryan, Mohammad Al Khatib, and Naim Bajcinca, "Model predictive control of industrial demand response for production lines". In: *IEEE International Symposium on Advanced Control of Industrial Processes (AdCONIP)*, Vancouver, 2022, pp. 151-157.

[6] Behzad Heydaryan, Javad Tousi, and Naim Bajcinca, "Economic model predictive control of industrial demand response," In: *IEEE XXVIII International Conference on Information, Communication and Automation Technologies (ICAT)*, Sarajevo, 2022, pp. 151-157.

Curriculum Vitae

Behzad Heydaryan Manesh

Research associate at Rheinland-Pfälzische Technische Universität

(RPTU)

Department of Mechanical and Process Engineering

Research & Education

05/2021 – 12/2026:	Research assistant in Energy Management System (EMS), RPTU • Development & implementation of control-based algorithms for research projects in the domain of energy management (electromobility and sector coupling)
05/2021 – 07/2026:	PhD candidate in control system at RPTU • Topic of the thesis: “Model Predictive Control for Industrial Energy Flexibility: Applications to Beverage Production Plants and Bidirectional Electric Vehicle Charging Hubs”
09/2017 – 02/2021:	Research assistant at Sharif University of Technology • Development of robust control algorithms for offshore wind farms
09/2014 – 01/2017:	Master program in control theory, Tarbiat Modarres University • Topic of thesis: Design & implementation of a robust bilateral teleoperation algorithm to control n-DoF nonlinear manipulators
09/2010 – 09/2014:	Bachelor program in electrical engineering, Isfahan University of Technology • Topic of project: Design an intelligent system for image processing and word seeking in “Scramble Words” game

Kaiserslautern, 16 September 2026