Doctoral Thesis

Demand response methods for ancillary services and renewable energy integration in electric power systems

Author(s):
Koch, Stephan

Publication Date:
2012

Permanent Link:
https://doi.org/10.3929/ethz-a-009756530

Rights / License:
In Copyright - Non-Commercial Use Permitted
Demand Response Methods for Ancillary Services and Renewable Energy Integration in Electric Power Systems

A dissertation submitted to
ETH ZURICH

for the degree of
Doctor of Sciences

presented by
STEPHAN KOCH
Dipl.-Ing., University of Stuttgart
born December 30, 1980
citizen of Germany

accepted on the recommendation of
Prof. Dr. Göran Andersson, examiner
Prof. Dr. Ian A. Hiskens, co-examiner

2012
Preface

This thesis was written during my time as a Ph.D. student at the Power Systems Laboratory (PSL) at ETH Zurich from October 2007 to June 2012. It is closely related to the work that I did within the research project “Local Load Management”.

First, I would like to express my sincere gratitude to Professor Göran Andersson for the opportunity to pursue my Ph.D. studies at the Power Systems Laboratory. Through his support, guidance, and the freedom he gave me to follow my own ideas in research, he contributed greatly to this work. I have appreciated his positive spirit and open-minded attitude towards new research ideas, collaboration efforts, and projects throughout the duration of my Ph.D. studies.

I also would like to thank Professor Ian Hiskens from University of Michigan for his kind willingness to co-examine this thesis and for the inspiring discussions that we had at various conferences. I highly appreciate his passion for this field of research.

I am thankful to the members of the project “Local Load Management” for our well-working collaboration, particularly Professor Martin Wiederkehr, Dominik Meier, and Professor Rolf Gutzwiller from University of Applied Sciences North-Western Switzerland, Dominic Lendi from Landis+Gyr, Dr. Matthias Zwicky from Alpiq, and Alexander Küster from Swissgrid. Financial support and constructive feedback from Swisselectric Research, the latter especially from Dr. Michael Paulus and Dr. Martin Kauert, is gratefully acknowledged.

I would like to thank Professor Duncan Callaway for the opportunity to visit his group Energy Modeling and Control (EMAC) at University of California, Berkeley, for a period of five months in 2010. Working with him motivated me to follow new research directions and to drill down into the details of the modeling and control approaches that I worked with. I really enjoyed working with the other graduate students at the lab, especially Johanna Mathieu, Taylor Keep, and Froy Sifuentes.
Several Master’s students completed a semester thesis or Master’s thesis under my supervision, in particular Vanco Janev, Fernando Soto Bárcenas, Emil Iggland, Ifigeneia Stefanidou, Maria Zerva, Matthias Bucher, Chakrit Bhamornsiri, Aryestis Vlahakis, Fernando de Samaniego Steta, Philip Jonas, Martin Pfeiffer, Diego Adolf, Ganbayar Puntsagdash, Philipp Fortenbacher, and Samuel Pfaffen (in chronological order). I would like to acknowledge their motivation and hard work, part of which contributed to the research presented in this thesis.

There are numerous colleagues and friends that I worked with during my time at the Power Systems Laboratory. In particular, I would like to thank Marek Zima, Matthias Galus, Spyros Chatzivasileiadis, Maria Vrakopoulou, Osvaldo Rodríguez Villalón, Frauke Oldewurtel, Theodor Borsche, Evangelos Vrettos, and especially Andreas Ulbig and Kai Heussen for the inspiring discussions and fruitful collaboration. I thank Turhan Demiray for providing a MATLAB-based dynamic power system simulation environment which was very useful for my research, and for the entertaining coffee breaks which made all my programming challenges a lot easier to deal with. In the same sense, the legendary lunch breaks with Thilo Krause often enriched my days with the latest linguistic twists, and the hiking trips, ski weekends, and long nights at BQM with many PSL colleagues helped me to keep the balance between work and leisure. All in all, I would like to thank the entire group for our great time. I really enjoyed our nice and friendly atmosphere and the fun that we had inside and outside of ETL G-floor.

I would like to thank Sarah for the love, support, and understanding that she gave me while I was working towards my Ph.D. degree. I also thank my friends for the encouragement and support, especially Francesco for philosophical discussions and a never-ending series of running gags.

I am deeply grateful to my family for their tremendous amount of support, especially to my mother who greatly supported my education and always had an open ear for any difficulties, to my brother with whom I exchanged quite a number of very stress-relieving e-mails, especially when I was actually writing the thesis, to my grandmother for the continuous supply of cake and cookies, and to my father who surely would have liked to be part of this intense period of my life.

Stephan Koch

June 2012
Contents

Abstract ......................................................... xiii
Kurzfassung ...................................................... xv
List of Acronyms ................................................ xix
List of Figures .................................................... xxiii
List of Tables ..................................................... xxvi

1 Introduction .................................................. 1
  1.1 Background and Motivation .............................. 1
    1.1.1 History of Electric Power Systems ............... 1
    1.1.2 Low-Carbon Energy Scenarios and Policy ....... 2
    1.1.3 Intermittent Renewable Energy .................. 4
    1.1.4 Distributed Generation ............................. 5
    1.1.5 Liberalization of the Electricity Sector ....... 5
    1.1.6 SmartGrids ......................................... 6
    1.1.7 Demand Response Potential ...................... 11
    1.1.8 Objectives of this Work .......................... 13
  1.2 Contributions ........................................... 14
  1.3 Outline of the Thesis .................................. 15
  1.4 List of Publications ................................... 17

2 Modeling and Coordination of TCLs ...................... 19
  2.1 Introduction and Literature Review .................. 19
  2.2 Generic TCL Model ..................................... 21
2.2.1 Normalized Expression of the TCL State . . . . 22
2.2.2 Dynamic TCL Models . . . . . . . . . . . . . . . 22
2.2.3 Thermal and Electric Energy Content . . . . . . . 24
2.2.4 Analytical Solution of the Model Equations . . 24
2.2.5 Uncoordinated Steady-State Properties . . . . . 25
2.3 Communication Infrastructure . . . . . . . . . . . . . 28
2.4 TCL Coordination Based on “Willingness to Switch” . . 29
  2.4.1 The Coordination Strategy . . . . . . . . . . . . . 29
  2.4.2 Computation of Accepted Switching Prices . . . 31
  2.4.3 Refinement of the Error Signal . . . . . . . . . . . 34
2.5 Aggregate Population Model . . . . . . . . . . . . . . 35
2.6 Simulation of the Coordination . . . . . . . . . . . . . 37
2.7 Implications of the TCL Coordination . . . . . . . . . . 39
  2.7.1 Impact of the Control on the TCLs . . . . . . . . . 39
  2.7.2 Choice of the Optimization Step Size . . . . . . . 41
  2.7.3 Energy Constraint Violation Behavior . . . . . . . 41
  2.7.4 Adaptation of the TCL Switching Boundaries . . 42
2.8 Concluding Remarks . . . . . . . . . . . . . . . . . . . 45

3 Modeling and Coordination of Electric Water Heaters 47
  3.1 Introduction and Motivation . . . . . . . . . . . . . . . 47
  3.2 Literature Review . . . . . . . . . . . . . . . . . . . . 49
  3.3 Single Water Heater Model . . . . . . . . . . . . . . . 51
    3.3.1 Physical Effects . . . . . . . . . . . . . . . . . . . 52
    3.3.2 Heuristic Circular Mass Flow Term . . . . . . . . 55
    3.3.3 Merging the Model Pieces . . . . . . . . . . . . . 56
    3.3.4 Hysteretic Thermostat Controller . . . . . . . . . 59
    3.3.5 Modeling of Water Draws . . . . . . . . . . . . . 60
    3.3.6 Simulation of a Single Water Heater . . . . . . . . 63
    3.3.7 Remark on Model Validation . . . . . . . . . . . 64
  3.4 Water Heater Population Model . . . . . . . . . . . . . 66
3.4.1 Modeling of Population Parameters .... 66
3.4.2 Simulation of a Water Heater Population .... 67

3.5 Control Strategies ................. 71
3.5.1 Overview .................. 71
3.5.2 Gamma Control Strategy .......... 71
3.5.3 Switching Control Strategy .......... 74
3.5.4 Comparison of the Simulation Results .... 78

3.6 Concluding Remarks ................ 80

4 Probabilistic Modeling and Control of TCLs 81
4.1 Introduction and Literature Review .... 81
4.2 Modeling Approach ................ 83
4.2.1 Individual TCL Model .......... 83
4.2.2 TCL Population Model .......... 84
4.3 Information Transfer ................ 88
4.3.1 General Considerations .......... 88
4.3.2 Implemented Approach .......... 88
4.3.3 Controllability and Observability Properties .. 91
4.4 Control Approach ................ 91
4.4.1 Initial Considerations .......... 91
4.4.2 Control Problem Formulation .... 92
4.5 Case Study .................... 94
4.5.1 System Setup and Scenarios .... 94
4.5.2 Numerical Results .......... 96
4.6 Concluding Remarks ................ 98

5 System-Level Unit Models 99
5.1 Introduction and Literature Review .... 99
5.2 Power Nodes Modeling Framework .... 103
5.2.1 Modeling Domains .......... 103
5.2.2 Model of a Single Power Node .... 105
5.2.3 Characterization of Unit Properties .... 107
5.2.4 System-Level Performance Indicators . . . . . . . 109
5.3 Development of Unit Models . . . . . . . . . . . . . . 111
  5.3.1 Generation Units . . . . . . . . . . . . . . . . . . 111
  5.3.2 Storage Units . . . . . . . . . . . . . . . . . . . . 115
  5.3.3 Load Units . . . . . . . . . . . . . . . . . . . . . 117
5.4 Concluding Remarks . . . . . . . . . . . . . . . . . . . 122

6 Dispatch Strategies 123
  6.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . 123
  6.2 Framework for Multi-Stage Operation . . . . . . . . . 124
    6.2.1 Decomposition of Power Node Equation . . . . . 126
    6.2.2 Constraint Coordination and Reserves . . . . . . 127
  6.3 Model Predictive Control . . . . . . . . . . . . . . . . 129
  6.4 Compact Portfolio Notation . . . . . . . . . . . . . . . 130
    6.4.1 Element-Wise Vector Notation . . . . . . . . . . 130
    6.4.2 Matrix Notation . . . . . . . . . . . . . . . . . . 131
    6.4.3 Discretization . . . . . . . . . . . . . . . . . . . 132
    6.4.4 Grid Model Integration . . . . . . . . . . . . . . 133
    6.4.5 Endogenous Cost Function . . . . . . . . . . . . 134
    6.4.6 Definition of Auxiliary Power Nodes . . . . . . 135
  6.5 Use Cases and Control Problems . . . . . . . . . . . . 136
    6.5.1 Least-Cost Economic Dispatch . . . . . . . . . . 137
    6.5.2 Market-Based VPP Operation . . . . . . . . . . . 140
    6.5.3 Balancing of Schedule Deviations . . . . . . . . 141
    6.5.4 Provision of Frequency Control Reserves . . . . 147
    6.5.5 Peak Shaving . . . . . . . . . . . . . . . . . . . . 152
    6.5.6 Capacity Firming of Intermittent Generation . . 154
    6.5.7 Residual Load Ramp-Rate Reduction . . . . . . . 155
  6.6 Simulation Environment . . . . . . . . . . . . . . . . . 156
  6.7 Benchmark Power Node Portfolios . . . . . . . . . . . 157
  6.8 Simulation Examples . . . . . . . . . . . . . . . . . . . 158
## Contents

6.8.1 Day-Ahead Dispatch .................................. 158
6.8.2 Intra-Day Update .................................... 160
6.8.3 Market-Based Operation of VPPs ..................... 162
6.8.4 Pre-MTU Balancing ................................. 162
6.8.5 Primary Frequency Control .......................... 164
6.8.6 Secondary Frequency Control .......................... 166
6.8.7 Tertiary Frequency Control ........................... 168
6.8.8 Peak Shaving ........................................ 168
6.8.9 Capacity Firming of Intermittent Generation ........ 171
6.8.10 Residual Load Ramp-Rate Reduction ............... 173

6.9 Concluding Remarks .................................. 176

7 Economic Evaluation of Frequency Control Provision by Flexible Unit Portfolios 177

7.1 Introduction and Literature Review ..................... 177
7.2 Aggregators in Electricity Markets .................... 179
  7.2.1 The Role of Aggregators ............................ 179
  7.2.2 Distribution Grid Constraints ..................... 180
  7.2.3 Unit Monitoring Challenges ....................... 181
7.3 Modeling of Revenue Potential ......................... 182
  7.3.1 Regulatory Basis for Revenue Calculation .......... 182
  7.3.2 Net Operating Profit ................................ 184
7.4 Simulation Study .................................... 186
  7.4.1 Simulation Scenarios ............................... 187
  7.4.2 Numerical Results ................................ 188
7.5 Profit Sharing Methodology ......................... 197
  7.5.1 Business Value Model ............................. 198
  7.5.2 Actors and Activities ............................. 198
  7.5.3 Exchanges ...................................... 200
  7.5.4 Cash Flow Consolidation ......................... 203
  7.5.5 Application Example ............................. 205
7.6 Concluding Remarks .................................. 208
8 Customer-Level Under-Frequency Load Shedding

8.1 Introduction and Motivation

8.1.1 Conventional UFLS Approaches

8.1.2 Motivation for CL-UFLS

8.2 Household Infrastructure

8.3 System Design Principles

8.3.1 Non-Transparency of the Customer

8.3.2 Centralized Off-Line, Decentralized On-Line

8.3.3 Category-Wise Appliance Clustering

8.3.4 Consideration of the “Value of Lost Load”

8.3.5 Robustness Towards Varying System Inertia

8.3.6 Emulation of Ramp-Wise Load Shedding

8.4 Load Shedding Problem Formulation

8.4.1 Power System Components

8.4.2 Load Ratios

8.4.3 Distributed Generation Model

8.5 Load Shedding System Layout and Flow of Information

8.5.1 Conventional Load Shedding

8.5.2 Customer-Level Load Shedding

8.5.3 Remaining Load at Bus $n$

8.6 Frequency Threshold Computation

8.6.1 Conventional Load Shedding

8.6.2 Customer-Level Load Shedding

8.7 Case Study

8.7.1 Power System Model

8.7.2 Load Shedding System Parameterization

8.7.3 Simulation Scenarios

8.7.4 Numerical Results

8.8 Concluding Remarks

9 Conclusions and Outlook

9.1 Appliance-Level Coordination
9.1.1 Lessons Learned .......................... 244
9.1.2 Suggestions for Future Research ............ 245

9.2 System-Level Dispatch Strategies ............... 246
9.2.1 Lessons Learned .......................... 246
9.2.2 Suggestions for Future Research ............ 246

9.3 Profitability of Flexible Unit Portfolios .......... 247
9.3.1 Lessons Learned .......................... 247
9.3.2 Suggestions for Future Research ............ 248

9.4 Emergency Control by CL-UFLS ................. 248
9.4.1 Lessons Learned .......................... 248
9.4.2 Suggestions for Future Research ............ 249

Bibliography .................................. 250
Abstract

This Ph.D. thesis deals with the utilization of flexible demand-side resources in the electric power system for the provision of ancillary services, emergency control functionalities, and the operational flexibility that is crucial for the transition to a power system based on Renewable Energy Sources (RES).

A major part of the work is focused on modeling and control of large populations of Thermostatically Controlled Loads (TCLs) for heating and cooling applications. The principal objective of this research effort is to achieve a reliable setpoint tracking of the TCLs’ aggregate power consumption without compromising the users’ comfort. Tracking is achieved by using coordination algorithms on the device level which send switching signals to selected TCLs, or broadcast information that causes autonomous switching reactions by TCLs. The coordination algorithms used are 1) rule-based and 2) based on Model Predictive Control (MPC). The aggregate State of Charge (SOC) of the TCL population can be defined as a measure of internal TCL temperatures with respect to upper and lower temperature limits, indicating how much flexibility the TCLs exhibit without becoming too hot or too cold. This enables the integration of the aggregate SOC in dispatch algorithms.

The TCL populations coordinated by setpoint tracking algorithms are integrated into a supervisory dispatch framework which combines controllable loads, storage devices, and generation units in so-called Virtual Power Plants (VPPs). These flexible portfolios of physically diverse units are mathematically represented by a novel modeling approach called Power Nodes Modeling Framework (PNMF). Various control services for power systems, e.g., least-cost dispatch, ramping reduction, and frequency control, are formulated and simulated using the PNMF. We use MPC for dispatching the portfolio using predictions of load and intermittent renewable energy as an input.
Based on the PNMF, we furthermore investigate economic aspects of using controllable loads and storage units for ancillary service provision in power systems. A primary concern is the revenue that a flexible portfolio of coordinated units can achieve in control reserves markets at certain price levels. To approach this question, we conduct parameter variations over a set of time simulations of two benchmark portfolios. This yields a simulation-based estimation of the control potential of a certain portfolio composition. Together with assumptions based on the analysis of historical price data, an assessment of the economic potential of control reserve provision by VPPs is enabled. Furthermore, we use the modeling methodology $e^3$ value to propose a profit sharing between the market actors participating in the reserve provision scheme.

Finally, this work also elaborates on the issue of Under-Frequency Load Shedding (UFLS) in large-scale disturbance situations. Nowadays, under-frequency relays are present in the distribution substations of most interconnected grids. These relays trigger the disconnection of distribution feeders in the case of a decaying system frequency. We investigate in this thesis by means of transient system simulations how this mechanism could be replaced by a completely decentralized Customer-Level Under-Frequency Load Shedding (CL-UFLS) system which would leave the distribution feeders connected and shed low-priority load preferably. We propose a system design and a methodology for assigning frequency thresholds to the individual appliance classes in close coordination with the co-existing Conventional Under-Frequency Load Shedding (C-UFLS) system. We conduct time simulations to show the effectiveness of CL-UFLS and also study the effects of Distributed Generation (DG) on the frequency dynamics during load shedding events.
Kurzfassung

Diese Doktorarbeit beschäftigt sich mit der Nutzung von flexiblen Verbrauchern im elektrischen Energiesystem für die Bereitstellung von Systemdienstleistungen, Notfall-Regelungsfunktionen und betrieblicher Flexibilität für die Systemtransformation zu einem auf erneuerbaren Energien basierenden Stromnetz.


Die von Sollwertfolge-Algorithmen koordinierten Gerätepopulationen werden in ein übergeordnetes System zur Einsatzplanung (Dispatch) eingebunden, das steuerbare Lasten, Energiespeicher und Erzeugungseinheiten in sogenannten „Virtuellen Kraftwerken“, Virtual Power Plants (VPPs), zusammenfasst. Diese flexiblen Portfolios aus physikalisch unterschiedlichen Einheiten werden mathematisch durch ein neuartiges Modellierungskonzept, genannt „Power Nodes Modeling Framework (PNMF)“, repräsentiert. Verschiedene Regeldienstleistungen für...


<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC</td>
<td>Alternating Current</td>
</tr>
<tr>
<td>ACE</td>
<td>Area Control Error</td>
</tr>
<tr>
<td>AMI</td>
<td>Advanced Metering Infrastructure</td>
</tr>
<tr>
<td>AVR</td>
<td>Automatic Voltage Regulator</td>
</tr>
<tr>
<td>BESS</td>
<td>Battery Energy Storage System</td>
</tr>
<tr>
<td>CCS</td>
<td>Carbon Capture and Storage</td>
</tr>
<tr>
<td>CE</td>
<td>Continental Europe</td>
</tr>
<tr>
<td>CFD</td>
<td>Computational Fluid Dynamics</td>
</tr>
<tr>
<td>CHF</td>
<td>Swiss Franc</td>
</tr>
<tr>
<td>CHP</td>
<td>Combined Heat and Power</td>
</tr>
<tr>
<td>CLPU</td>
<td>Cold Load Pick-Up</td>
</tr>
<tr>
<td>CL-UFLS</td>
<td>Customer-Level Under-Frequency Load Shedding</td>
</tr>
<tr>
<td>COI</td>
<td>Center of Inertia</td>
</tr>
<tr>
<td>C-UFLS</td>
<td>Conventional Under-Frequency Load Shedding</td>
</tr>
<tr>
<td>DC</td>
<td>Direct Current</td>
</tr>
<tr>
<td>DG</td>
<td>Distributed Generation</td>
</tr>
<tr>
<td>DLR</td>
<td>Dynamic Line Rating</td>
</tr>
<tr>
<td>DR</td>
<td>Demand Response</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>DSM</td>
<td>Demand Side Management</td>
</tr>
<tr>
<td>DSO</td>
<td>Distribution System Operator</td>
</tr>
<tr>
<td>ENTSO-E</td>
<td>European Network of Transmission System Operators for Electricity</td>
</tr>
<tr>
<td>EPEX</td>
<td>European Power Exchange</td>
</tr>
<tr>
<td>ETP</td>
<td>European Technology Platform</td>
</tr>
<tr>
<td>EUR</td>
<td>Euro</td>
</tr>
<tr>
<td>EV</td>
<td>Electric Vehicle</td>
</tr>
<tr>
<td>EWH</td>
<td>Electric Water Heater</td>
</tr>
<tr>
<td>FACTS</td>
<td>Flexible Alternating Current Transmission System</td>
</tr>
<tr>
<td>FAPER</td>
<td>Frequency Adaptive Power-Energy Re-scheduler</td>
</tr>
<tr>
<td>GHG</td>
<td>Greenhouse Gas</td>
</tr>
<tr>
<td>HV</td>
<td>High Voltage</td>
</tr>
<tr>
<td>ICT</td>
<td>Information and Communication Technology</td>
</tr>
<tr>
<td>IEEE</td>
<td>Institute of Electrical and Electronics Engineers</td>
</tr>
<tr>
<td>LFC</td>
<td>Load Frequency Control</td>
</tr>
<tr>
<td>LM</td>
<td>Load Management</td>
</tr>
<tr>
<td>LMA</td>
<td>Load Manager Appliance</td>
</tr>
<tr>
<td>LMH</td>
<td>Load Manager Household</td>
</tr>
<tr>
<td>LPV</td>
<td>Linear Parameter-Varying</td>
</tr>
<tr>
<td>LTI</td>
<td>Linear Time-Invariant</td>
</tr>
<tr>
<td>LV</td>
<td>Low Voltage</td>
</tr>
<tr>
<td>MG</td>
<td>MicroGrid</td>
</tr>
<tr>
<td>MPC</td>
<td>Model Predictive Control</td>
</tr>
<tr>
<td>MTU</td>
<td>Market Time Unit</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>MV</td>
<td>Medium Voltage</td>
</tr>
<tr>
<td>O&amp;M</td>
<td>Operation &amp; Maintenance</td>
</tr>
<tr>
<td>OPF</td>
<td>Optimal Power Flow</td>
</tr>
<tr>
<td>PDE</td>
<td>Partial Differential Equation</td>
</tr>
<tr>
<td>PHEV</td>
<td>Plug-In Hybrid Electric Vehicle</td>
</tr>
<tr>
<td>PI</td>
<td>Proportional-Integral</td>
</tr>
<tr>
<td>PLC</td>
<td>Powerline Communication</td>
</tr>
<tr>
<td>PLL</td>
<td>Phase-Locked Loop</td>
</tr>
<tr>
<td>PMU</td>
<td>Phasor Measurement Unit</td>
</tr>
<tr>
<td>PNMF</td>
<td>Power Nodes Modeling Framework</td>
</tr>
<tr>
<td>PV</td>
<td>Photovoltaic</td>
</tr>
<tr>
<td>RES</td>
<td>Renewable Energy Sources</td>
</tr>
<tr>
<td>RMSE</td>
<td>Root Mean Square Error</td>
</tr>
<tr>
<td>SCADA</td>
<td>Supervisory Control and Data Acquisition</td>
</tr>
<tr>
<td>SOC</td>
<td>State of Charge</td>
</tr>
<tr>
<td>TCL</td>
<td>Thermostatically Controlled Load</td>
</tr>
<tr>
<td>TCP/IP</td>
<td>Transmission Control Protocol/Internet Protocol</td>
</tr>
<tr>
<td>TSO</td>
<td>Transmission System Operator</td>
</tr>
<tr>
<td>UCPTE</td>
<td>Union pour la Coordination de la Production et du Transport de l’Électricité</td>
</tr>
<tr>
<td>UCTE</td>
<td>Union for the Co-ordination of Transmission of Electricity</td>
</tr>
<tr>
<td>UFLS</td>
<td>Under-Frequency Load Shedding</td>
</tr>
<tr>
<td>VOLL</td>
<td>Value of Lost Load</td>
</tr>
<tr>
<td>VPP</td>
<td>Virtual Power Plant</td>
</tr>
<tr>
<td>WAMS</td>
<td>Wide-Area Monitoring System</td>
</tr>
</tbody>
</table>
List of Figures

2.1 Uncoordinated operation of a TCL population . . . . . . 27
2.2 Communication infrastructure in a “smart” household . 28
2.3 Calculation tasks and interfaces between the units . . . 31
2.4 Steps S1 – S6 of the coordination algorithm . . . . . . 33
2.5 Coordinated population of 600 TCLs . . . . . . . . . . 38
2.6 Impact of coordination on the operating cycle of TCL $i$ 40
2.7 Violation of the population energy constraints . . . . . 42
2.8 Adaptation of the switching boundaries . . . . . . . . . 44

3.1 Schematic representation of an EWH . . . . . . . . . . 52
3.2 Energy and mass balance for a regular disk . . . . . . . 53
3.3 Energy and mass balance for a mixing zone disk . . . 54
3.4 Energy and mass balance for top and bottom disks . . 54
3.5 Heuristic circular water mass flow . . . . . . . . . . . . 56
3.6 Water draw probability profile . . . . . . . . . . . . . . . 61
3.7 Exemplary water draw scenario . . . . . . . . . . . . . . 63
3.8 Simulation results of single unit: unimpeded operation . 65
3.9 Simulation results of single unit: blocked during day . 65
3.10 Simulation results of population: unimpeded operation . 69
3.11 Simulation results of population: blocked during day . 70
3.12 EWH population controlled by Gamma Control . . . . 74
3.13 EWH population controlled by Switching Control . . . 78
3.14 Comparison of tracking performance and user comfort . 79
4.1 State bin transition model . . . . . . . . . . . . . . . . 85
4.2 Unforced system evolution and eigenvalue comparison . 86
4.3 Control structure for TCL population . . . . . . . . . . 92
4.4 Simulation result for full state information . . . . . . . 97
5.1 Power system modeling domains . . . . . . . . . . . . . . 104
5.2 Notation for a single power node . . . . . . . . . . . . . 106
5.3 EWH as lumped energy storage . . . . . . . . . . . . . . 120
5.4 Water draw probability profile . . . . . . . . . . . . . . 121
6.1 Principle of receding horizon optimization . . . . . . . . 129
6.2 Power node setup for least-cost dispatch . . . . . . . . . 138
6.3 VPP operation with market exchanges . . . . . . . . . . 140
6.4 Swiss energy balancing regime . . . . . . . . . . . . . . 142
6.5 Self-balancing of a power node portfolio . . . . . . . . . 144
6.6 Balancing of external deviations by power node portfolio 144
6.7 Exemplary imbalance trajectory (peak load of 10 MW) . 147
6.8 Temporal structure of frequency control reserves . . . . 148
6.9 Frequency control with power node portfolio . . . . . . 149
6.10 Peak shaving, capacity firming, and ramp-rate reduction 153
6.11 Structure of the simulation environment . . . . . . . . . 156
6.12 Simulation example: day-ahead dispatch . . . . . . . . . 161
6.13 Simulation example: market-based VPP operation . . . . 163
6.14 Simulation example: primary control (Portfolio A) . . 165
6.15 Simulation example: primary control (Portfolio B) . . 165
6.16 Simulation example: secondary control (Portfolio A) . . 167
6.17 Simulation example: secondary control (Portfolio B) . . 167
6.18 Simulation example: peak shaving (base case) . . . . . . 169
6.19 Simulation example: peak shaving (application) . . . . . 169
6.20 Simulation example: peak shaving (comparison) . . . . . 170
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.21</td>
<td>Simulation example: capacity firming (base case)</td>
<td>171</td>
</tr>
<tr>
<td>6.22</td>
<td>Simulation example: capacity firming (application)</td>
<td>172</td>
</tr>
<tr>
<td>6.23</td>
<td>Simulation example: capacity firming (comparison)</td>
<td>172</td>
</tr>
<tr>
<td>6.24</td>
<td>Simulation example: ramp-rate reduction (base case)</td>
<td>174</td>
</tr>
<tr>
<td>6.25</td>
<td>Simulation example: ramp-rate reduction (application)</td>
<td>174</td>
</tr>
<tr>
<td>6.26</td>
<td>Simulation example: ramp-rate reduction (comparison)</td>
<td>175</td>
</tr>
<tr>
<td>7.1</td>
<td>Vertically integrated versus liberalized system structure</td>
<td>180</td>
</tr>
<tr>
<td>7.2</td>
<td>Coordination between aggregator and DSO</td>
<td>181</td>
</tr>
<tr>
<td>7.3</td>
<td>Swiss primary and secondary reserve prices, year 2011</td>
<td>183</td>
</tr>
<tr>
<td>7.4</td>
<td>Opportunity cost incurred by the generator</td>
<td>185</td>
</tr>
<tr>
<td>7.5</td>
<td>Number of water heaters over parameter ranges</td>
<td>187</td>
</tr>
<tr>
<td>7.6</td>
<td>Simulation result: Primary Control with Portfolio A</td>
<td>190</td>
</tr>
<tr>
<td>7.7</td>
<td>Simulation result: Primary Control with Portfolio B</td>
<td>192</td>
</tr>
<tr>
<td>7.8</td>
<td>Simulation result: Secondary Control with Portfolio A</td>
<td>194</td>
</tr>
<tr>
<td>7.9</td>
<td>Simulation result: Secondary Control with Portfolio B</td>
<td>196</td>
</tr>
<tr>
<td>7.10</td>
<td>$e^3$ value elements</td>
<td>197</td>
</tr>
<tr>
<td>7.11</td>
<td>Business value model for the aggregator</td>
<td>199</td>
</tr>
<tr>
<td>7.12</td>
<td>Cash in-/out-flows in the four selected scenarios</td>
<td>206</td>
</tr>
<tr>
<td>8.1</td>
<td>Ramp-wise load shedding illustration</td>
<td>216</td>
</tr>
<tr>
<td>8.2</td>
<td>Feeders and load classes at an arbitrary bus $n$</td>
<td>218</td>
</tr>
<tr>
<td>8.3</td>
<td>Flow of information between the entities</td>
<td>221</td>
</tr>
<tr>
<td>8.4</td>
<td>Finite state machine of load shedding mechanism</td>
<td>222</td>
</tr>
<tr>
<td>8.5</td>
<td>Topology of the 118-bus system</td>
<td>227</td>
</tr>
<tr>
<td>8.6</td>
<td>Distribution of load on the distribution feeders</td>
<td>229</td>
</tr>
<tr>
<td>8.7</td>
<td>Legend of the load classes for load shedding plots</td>
<td>231</td>
</tr>
<tr>
<td>8.8</td>
<td>Static analysis – C-UFLS only</td>
<td>232</td>
</tr>
<tr>
<td>8.9</td>
<td>Static analysis – CL-UFLS with unchanged C-UFLS</td>
<td>233</td>
</tr>
<tr>
<td>8.10</td>
<td>Static analysis – CL-UFLS with adapted C-UFLS</td>
<td>233</td>
</tr>
<tr>
<td>8.11</td>
<td>Simulation of scenario 1): C-UFLS, no DG</td>
<td>237</td>
</tr>
<tr>
<td>8.12</td>
<td>Simulation of scenario 2): C-UFLS, 50% DG</td>
<td>238</td>
</tr>
<tr>
<td>8.13</td>
<td>Simulation of scenario 3): C-UFLS + CL-UFLS, 50% DG</td>
<td>239</td>
</tr>
<tr>
<td>8.14</td>
<td>Comparison of the simulation results</td>
<td>241</td>
</tr>
</tbody>
</table>
# List of Tables

2.1 Notation for Chapter 2 ........................................ 21

3.1 Notation for Chapter 3 ........................................ 50
3.2 Assumed stochastic parameters for water draw scenarios 63
3.3 Parameters for individual water heater simulation .... 64
3.4 Parameters for creating the water heater population 67
3.5 Overview of water heater control strategies ........... 72
3.6 Water heater sub-populations .............................. 75

4.1 Notation for Chapter 4 ....................................... 83
4.2 Simulation and TCL parameters ........................... 85
4.3 Options for control and information transfer ........... 89
4.4 Options for controller parameterization ............... 95
4.5 Performance comparison ................................... 96

5.1 Notation for Chapter 5 ....................................... 102
5.2 Unit properties determined by power node constraints 108

6.1 Notation for Chapter 6 ....................................... 124
6.2 Directionality of control reserve provision ............ 127
6.3 Parameters used for imbalance creation ............... 147
6.4 Power node types for benchmark portfolios .......... 158
6.5 Benchmark portfolio parameters and constraints .... 159
LIST OF TABLES

6.6 Benchmark portfolio cost parameters ....................... 160
6.7 Day-ahead dispatch – balance terms ....................... 160
6.8 Market-based VPP operation – balance terms ............... 162
6.9 Primary control (Portfolio A) – Δ balance terms ........... 164
6.10 Primary control (Portfolio B) – Δ balance terms ........... 164
6.11 Secondary control (Portfolio A) – Δ balance terms ......... 166
6.12 Secondary control (Portfolio B) – Δ balance terms ......... 166
6.13 Peak shaving – balance terms ................................ 170
6.14 Capacity firming – balance terms ............................ 173
6.15 Ramp-rate reduction – balance terms ....................... 175

7.1 Notation for Chapter 7 ...................................... 178
7.2 Parameter variations for frequency control assessment ........ 187
7.3 Assumptions for economic parameters ....................... 188
7.4 Yearly financial results of the four selected scenarios ...... 206
7.5 Cash in-/out-flows in the four selected scenarios .......... 206

8.1 Notation for Chapter 8 ...................................... 210
8.2 Conventional load shedding stage plan ....................... 229
8.3 Assumed customer categories in 118-bus system ............. 230
8.4 Load classes, VOLL assumptions and freq. thresholds ....... 231
8.5 Numerical comparison of the simulation results ............. 241
Chapter 1

Introduction

This first chapter will give an overview of the recent developments in the electric power industry that motivate the present work. After a brief look at historical facts, we will focus on the tremendous changes that are currently occurring in energy policy, energy and power research, and power system operation. We will touch upon climate policy, renewable energies, power market liberalization, and the paradigm of “SmartGrids”, specifically the active control of the demand side, in order to set the scene for the chapters to follow.

1.1 Background and Motivation

1.1.1 History of Electric Power Systems

The development of electric power systems dates back to the end of the 19th century when the basic ideas of power supply infrastructures were conceived. The first stages of the electrification of society were initiated in parallel in the United States of America and Europe through three interrelated factors: major technological breakthroughs on the side of electricity generation such as the invention of the dynamo, novel electrified household appliances and industrial machines, as well as innovative business models [1]. Small-scale isolated power systems emerged when the first thermal power plants were constructed on industrial sites or in small neighborhoods. After an initial struggle between Alternating Current (AC) and Direct Current (DC) grids, AC prevailed as the standard for electric power supply. The individual small distribution grids were interconnected and gradually extended to larger regional and later interregional grids as transmission technology progressed and voltage
levels increased. The main argument for interconnection was the increased reliability resulting from the fact that a power deficit due to a generator outage could be compensated by other sources in the grid. These developments evolved in parallel in various parts of the world.

While in the USA the transmission systems were predominantly confined within the borders of the country, the geographical situation in Europe called for international cooperation in order to extend the power system to a large-scale, continent-wide grid. In 1951, the Union pour la Coordination de la Production et du Transport de l’Électricité (UCPTE), which was later called Union for the Co-ordination of Transmission of Electricity (UCTE), was formed in order to facilitate the development of the European interconnected system [2]. In the second half of the 20th century, major grid extensions were realized under the supervision of UCTE. Rules and regulations for Transmission System Operators (TSOs) were standardized in order to allow a coordinated operation. The predominant paradigm of that era was the efficient bulk power transmission over long distances, along with centralized power generation largely reliant on fossil fuels, hydro, and nuclear power. Vertical power flows between voltage levels were usually unidirectional as hardly any power generation was taking place in the lower voltage levels. Generation units were scheduled to follow the uncontrolled demand.

The historically grown structures of electric power systems are currently facing radical changes which were initiated around the beginning of the 21st century. The necessity to cut back on Greenhouse Gas (GHG) emissions has instilled a paradigm shift in energy policy targets and legislation, initiating a large-scale transformation towards a low-carbon (or even zero-carbon) energy system [3]. Renewable Energy Sources (RES), often intermittent and decentralized, are increasingly utilized for power production. In parallel, the electricity sector was liberalized and is now subject to (individually different degrees of) free-market competition in all European countries and many others around the world. These developments are increasing the complexity of planning and operating interconnected power systems.

1.1.2 Low-Carbon Energy Scenarios and Policy

The growing concern about the impacts of carbon dioxide emissions on the world’s climate systems were underlined by new research results in the first decade of the 21st century. According to these, a failure to
cap CO₂ levels in the atmosphere below certain critical levels within the next decades is likely to lead to severe consequences for natural climate systems, fauna and flora, and human civilization [4]. The costs of implementing an effective climate protection regime in order to achieve the transition towards a low-carbon energy system were estimated to be lower than dealing with the consequences of non-action [5]. In the context of political action against climate change, an emissions trading system has been put into place in Europe [6].

Energy system research has addressed these findings by focusing research efforts on the transition towards a more sustainable energy system, primarily by means of simulation studies and scenario analyses. The envisioned pathways to a low-carbon energy future are usually dependent upon assumption sets in the scenario simulations which are often highly uncertain and may be influenced by subjective reasoning and political preferences. Hence, a large number of completely different energy scenarios have emerged, describing a multitude of future energy systems with diverse ideas on possible generation portfolios. A comparative assessment of selected scenarios can be found in [7]. Some of the recently published scenarios rely on an expansion of nuclear power [8], fossil fuels with Carbon Capture and Storage (CCS) [9], or natural gas [10]. Other scenarios, such as [11, 12, 13, 14], favor high shares of RES for the decarbonization of the energy production. Global potentials of renewable energies have been shown to be sufficient and technically accessible for covering the entire global energy demand [15, 16]. Arguments for renewable energy utilization are steep industrial learning curves and large cost reduction potentials [17] and the absence of critical risks as in the case of, e.g., nuclear energy. Problems are associated with intermittency, low energy density, and the (presently) high generation cost which calls for financial incentive schemes. In general, the scientific and public discussion on advantages and disadvantages of different generation technologies as well as energy policy frameworks and incentive schemes can be highly controversial [18, 19, 20, 21].

The share that RES will account for in the world’s energy mix of 2050 and beyond is hard to predict. However, a continued expansion at substantial growth rates can be assumed. Many countries have implemented ambitious targets for RES expansion [22, 23] and a number of incentive schemes such as feed-in tariffs or quota systems [24]. These schemes are known to interact with the European emission trading system not always in the desired way [25], but there are strong arguments
for letting both instruments co-exist and complement each other [26]. These aspects of recent energy policy create a dynamic environment for the expansion of RES utilizations with high growth rates [27].

1.1.3 Intermittent Renewable Energy

The principal primary sources of renewable energy are the sun (solar radiation, ambient heat, precipitation, melting of snow, ocean currents, wind, waves, and biomass), gravitation of the moon (tidal energy), and radioactive fission processes in the earth (geothermal energy) [28]. All of these energy forms can be used for electricity generation. In general, the more concentrated forms of renewable energy, such as hydro power and biomass, are easier to harvest than the more dispersed sources such as wind and solar energy. However, the technically usable potentials under the constraints of environmental protection are far larger for wind and solar energy. Consequently, substantial percentages of future large RES contributions will have to come from these sources.

The basic challenge of wind and solar in-feeds in the power system lies in their intermittency on a variety of timescales [29, 30]: wind energy exhibits a strong dependency on the season, the time of day, and prevalent weather conditions. Hour-to-hour variations with strong autocorrelation can be observed, as well as uncorrelated ones in the range of seconds and minutes with a lower amplitude. Solar photovoltaic (PV) power generation exhibits dominant daily generation patterns, which also depend on the season, and short-term effects induced by clouds and other weather phenomena. In solar-thermal power production, the daily generation pattern can be influenced by using a heat storage [31].

The impacts of electricity generation by intermittent RES on power systems are manifold: the temporal variability of the RES generation poses a challenge to the dispatch of conventional power plants since the variation of the residual load (load minus intermittent generation) can trigger frequent start-up and shut-down events [32]. Large active-power ramps induced by intermittent generation can be higher than the usual load ramps and can also coincide with load changes in the opposite direction, as reported in [33] for the case of wind power. Increasing amounts of required active power reserves can also be observed as wind power penetration increases [34]. Another important factor is the increased need for long-distance power transmission since the best RES sites are often far away from the load centers.
1.1. Background and Motivation

In general, we can state that the flexibility of a power system is of critical importance for its hosting capacity for renewable energies [35]. On power markets, a lack of flexibility can lead to negative spot market prices since it may be more economical for a power plant to produce for a negative revenue than to incur shut-down or ramping costs [36].

Furthermore, there are short-time-scale impacts on power system stability induced by the increase in renewable energy. In the case of wind energy, doubly-fed induction generators have little inertia compared to synchronous machines, and inverter-connected generation units do not provide any inertia to the grid by default. Fault-ride-through capability, especially of wind power [37], is important for protecting the power system in the case of disturbances. Solar PV power generation can also influence the dynamic behavior of the interconnected network [38].

1.1.4 Distributed Generation

Along with the increasing penetration of renewable energies, the amount of decentralized production of electric energy in distribution grids, mostly referred to as Distributed Generation (DG), is on the rise. Distribution grids were originally not designed to accommodate large amounts of generation units, hence they are still lacking the same degree of automation and monitoring which is common in transmission systems. If only small amounts of DG are present, they can be installed according to a “fit and forget” paradigm and do not have to be actively managed. With DG becoming a major new paradigm of electric energy supply [39], the high density of generation in many distribution grids can cause novel issues that require active measures for integration [40]. Especially in the case of PV power generation, which is predominantly decentralized and also intermittent, a high penetration in a distribution grid may require measures to avoid over-voltages and component overloads [41].

1.1.5 Liberalization of the Electricity Sector

In many countries, the electricity markets were liberalized in the last 20 years. As with other vital infrastructures such as telecommunication networks or water supply, legislation and regulation was continuously adapted to the experiences during the liberalization process. Although the initial objectives of market liberalization, such as increased efficiency, lower prices, and inherent incentives to protect the environment, were not always attained right after the market openings [42], large
Chapter 1. Introduction

Gains in experience may remedy the negative aspects of liberalization over time [43].

As a result of the market openings, the transmission grids are increasingly utilized for energy trading, which increases the necessity to transport power over large distances. This can lead to a power system operation closer to its security margins [44]. Congestions between countries are usually accounted for by explicit or implicit capacity auctions [45], while intra-area congestions are usually handled by the responsible TSO, e.g., by re-dispatch and counter-trading.

The national jurisdiction on the implementation of energy market liberalization has led to a high diversity of energy market regimes. Especially in the field of ancillary service markets, the concrete market design, product structures, and the used nomenclature may vary significantly between countries [46].

We now take a brief look at the market opening process in Switzerland, as we base some of our analyses in later chapters on the Swiss situation. The liberalization of the Swiss electricity sector was designed as a two-stage process. As described in [47], the first part of the liberalization took place on January 1, 2009, when the wholesale electricity market as well as the ancillary service markets were opened. Final customers with an annual consumption above 100 MWh were enabled to opt out of the regulated tariffs and to freely choose their electricity supplier. A second stage of liberalization, scheduled for 2014, shall open the market also for small customers with a consumption below 100 MWh per year. Active power control reserves, which are of particular interest in this work, are tendered in weekly or daily auction processes which are administered by the TSO Swissgrid [48].

1.1.6 SmartGrids

The paradigm of “SmartGrids” emerged as a way to deal with the increasing complexity in the electric power industry as a consequence of RES expansion, energy efficiency requirements, and market liberalization. Although the term lacks a formal definition accepted by all stakeholders, the common understanding of “SmartGrids” is a combination of power engineering and Information and Communication Technology (ICT) for the implementation of advanced techniques for monitoring, control, and optimization on all grid levels [49]. Motivations for the implementation of SmartGrid features can be energy conservation,
enhanced controllability and stability of the power system, novel revenue opportunities in energy or ancillary service markets, or economic benefits due to more efficient grid operation. According to the Smart-Grids European Technology Platform (ETP) [50], a SmartGrid is:

“an electricity network that can intelligently integrate the actions of all users connected to it – generators, consumers and those that do both – in order to efficiently deliver sustainable, economic and secure electricity supplies. A smart grid employs innovative products and services together with intelligent monitoring, control, communication, and self-healing technologies in order to:

- Better facilitate the connection and operation of generators of all sizes and technologies
- Allow consumers to play a part in optimising the operation of the system
- Provide consumers with greater information and options for choice of supply
- Significantly reduce the environmental impact of the whole electricity supply system
- Maintain or even improve the existing high levels of system reliability, quality and security of supply
- Maintain and improve the existing services efficiently
- Foster market integration towards a European integrated market.”

Some key features of SmartGrids are elaborated below.

Information and Communication Technology

The key enabling technology for SmartGrids is ICT integrated with the power system infrastructure. This includes advanced control and monitoring equipment in transmission and distribution systems which allows operators to obtain a high degree of knowledge about the system’s operational state as well as to take action in real time. Utility business processes, such as the billing of customers, are facilitated by the utilization of Advanced Metering Infrastructure (AMI). Furthermore, the possibility to interact with the final customer, e.g., through dynamic pricing regimes, enhanced energy services, or direct load control, is established by ICT [51]. “SmartHome” technologies, originally designed for user comfort enhancement and energy efficiency increase, can be integrated with the grid ICT to allow a joint optimization according to local performance requirements and grid-wide objectives [52].
Chapter 1. Introduction

SmartGrid Features on Different Grid Levels
Many transmission-level power system features now attributed to the SmartGrid paradigm have been existing for quite some time since sophisticated monitoring and control is a long-standing requirement for transmission systems. This includes advanced Supervisory Control and Data Acquisition (SCADA) systems [53], Wide-Area Monitoring Systems (WAMS) [54], Dynamic Line Rating (DLR) [55], Phasor Measurement Units (PMUs) [56], as well as enhanced power flow control by Flexible Alternating Current Transmission Systems (FACTS) [57].

Conversely, real-time monitoring and control has not been necessary in traditional distribution system operation. For economic reasons, traditional distribution systems were designed conservatively with respect to peak load requirements, eliminating the need for active management in real time. As outlined above, this will have to change with increasing penetration of DG. Furthermore, new sources of demand, such as Electric Vehicles (EVs) or Plug-In Hybrid Electric Vehicles (PHEVs) [58], require a more sophisticated management of the distribution grid.

Energy Storage
In a power system largely based on intermittent renewable energy, the presence of energy storage is a key factor for efficient RES integration. Although energy storage is already common in today’s legacy power systems (mainly realized in the form of pumped-hydro storage plants), it is mostly utilized for day/night arbitrage trading, which in the aggregate renders the load curve flatter, allowing for a more efficient infrastructure utilization. Another use case is the provision of ancillary services, which takes advantage of the particular suitability of hydro power plants for flexible operation and fast ramping. However, storage is not a principal requirement for power system operation if sufficient dispatchable generation capacity is available for peak-load coverage.

The presence of controllable and uncontrollable generation and demand on all grid levels is complemented by the idea of “ubiquitous energy storage” [59], which advocates the use of storage on the premises of the final customer, in community-scale units in the distribution grid, as well as in the form of centralized storage facilities on the transmission level.

Operating a large variety of storage technologies on all grid levels requires a more sophisticated methodology for dispatch and optimization. As energy constraints can be far more dominant in decentralized
small-scale storage devices than in hydro power plants, the State of Charge (SOC) of an individual storage device, or a group of storage devices, has to be taken into account when charging or discharging decisions are made. The goal of the storage application is a determining factor for the control strategy. Based on this strategy, functional requirements for storage can be defined [60] which have to be compared with the physical properties of a certain storage technology in order to assess its suitability for a certain application.

**Load as an Active Control Resource**

Unlike in traditional power system operation, the active management of the demand side, which is often termed Demand Response (DR), Demand Side Management (DSM) or Load Management (LM), is an integral component of an envisioned SmartGrid infrastructure. In what follows, we adopt the term Demand Response for all activities related to active load control and optimization.

The main drivers for DR are 1) the need to compensate for the additional uncertainty brought about by intermittent renewable energy, 2) the possibility to react locally to problems caused by DG in the distribution grid, and 3) the so far unexploited flexibility potential available on the load side. Major barriers are 1) high cost of ICT infrastructure, 2) privacy concerns on the end-user side, and 3) technical and organizational complexity associated with coordinated operation of thousands, possibly millions, of interconnected small loads.

A large DR potential lies in heating and cooling loads due to their thermal inertia and corresponding inherent “slack” in their power consumption [61]. These loads are often controlled by hysteretic thermostats. We refer to these loads as Thermostatically Controlled Loads (TCLs).

The main use cases for load control are more traditional applications such as peak shaving and valley filling [62, 63] as well as fairly new use cases such as balancing services to market players [64] and the provision of ancillary services [65, 66, 67]. While the first two applications can be accomplished by an a-priori scheduling of demand, the latter two require interactions with the load on short time scales.

The methodologies to manage large populations of dispersed loads vary with the considered application. In order to influence the shape of the load curve on the time scale of hours, price signals may be transmitted to the final customers to incentivize a desired load shifting reaction
In this case, the reaction on the customers’ side may be both manual or automatic. A measurement of the power system frequency may be utilized for short-term reactions of TCLs, primarily targeted at replacing conventional primary control reserves [70]. Approaches based on multi-agent systems [71] and negotiation methodologies [72] are also found in the literature. In this work, we focus on methodologies for aggregation and direct load control which means that individual appliances are directly controlled by an electricity utility or a third-party entity called “aggregator”. Numerous works in this area were developed in parallel to this thesis, e.g., [73, 74, 75, 76, 77, 78].

As the integration of DR into power system control is the main focus of this work, a more detailed literature review is given in the subsequent chapters. In general, the active control of the demand side can complement or substitute the utilization of flexible generation and energy storage in many cases, although the time scales are often different: while many storage systems can store energy for days and weeks without major losses, and generators are only limited by their power constraints, DR control actions are usually limited to the scale of hours and below.

**Aggregators, Virtual Power Plants, and MicroGrids**

In order to face the challenge of integrating a multitude of dispersed units into power system operation, new technical and economic concepts were developed during the last years. The paradigm of “aggregation”, meaning a coordinated management of large numbers of small units, has gained much importance in the scientific discourse [79]. Novel entities called aggregators, which specialize in communicating with and co-optimizing these large unit groups [80], are envisioned to play an increasingly large role in power systems. Aggregation can be performed both from a market-driven and grid-driven perspective.

The market-driven approach is based on the fact that energy markets have been tailored for the participation of large centralized generators. Minimum bid sizes as well as the practical procedure of submitting bids to a power exchange or ancillary service market make it hard for small units operating on the scale of kilowatts to participate in power markets on their own. The concept of Virtual Power Plants (VPPs) [81] was created to combine a large number of these small units and to make them appear to the power market as a single entity. The individual units can be dispersed over various distribution grids and even control areas. Conversely, aggregation based on network topology is motivated
by the fact that most problems related to DG appear on the Medium Voltage (MV) and Low Voltage (LV) level. On the LV level, the idea to control individual parts of the grid as a single entity has led to the concept of MicroGrids (MGs) [82]. On the MV level, aggregated grid parts are often referred to as “cells” [83]. Both MGs/cells and VPPs are suitable for the integration of controllable loads, which establishes their relevance for the focus of this thesis.

Measures to Enhance System Security

Maintaining the current high reliability of the electric power supply is of utmost importance for the functioning of our societies. Prolonged power outages can have devastating effects on the economy, social life, public health, safety, and security in large geographical areas [84]. However, this reliability is expected to become increasingly jeopardized by effects of the significant transformation that the power system is facing. Blackouts and disturbances such as in the US and Italy in 2003 and in central Europe in 2006 [85] have demonstrated the vulnerability of modern electricity grids. The analysis of the latter event [86] has shown that the uncontrolled nature of DG disconnection and reconnection (in this case mostly caused by wind turbines without fault-ride-through capability) has made the mitigation of the disturbance more difficult. This underlines the necessity for enhancing the controllability of both generation and load and the application of new system operation strategies in normal operation and emergency situations. An example of such strategies is the so-called “controlled islanding” of parts of the grid in the case of a grid-wide disturbance and the adoption of novel Under-Frequency Load Shedding (UFLS) schemes. The latter may include an active shedding of individual loads on the customer level instead of entire distribution feeders. This requires communication infrastructure that reaches down to the final customer and a sophisticated methodology for designing the desired shedding response. Through such a system, active loads can also contribute to an increased system security.

1.1.7 Demand Response Potential

Having stressed the significance of DR for “SmartGrid” functionalities both in normal operation and in the case of large-scale disturbances, the question of the available potential of flexible loads for the various considered applications arises. Literature values for DR potentials are highly diverse due to their dependency upon the considered country,
the time of day, the season, and the way that the influence on the load is implemented. There is no standardized methodology to calculate DR potentials. Many known publications focus on a certain amount of power that can be blocked for a given period of time. Others take into account the possibility of a controlled load increase. The possible duration of the load modification is often stated as a measure for the available energy that can be shifted over time.

Some examples for DR potential assessments are mentioned in the following. Reference [87] summarizes a number of publications on DR potentials in Europe. As mentioned here, the European Network of Transmission System Operators for Electricity (ENTSO-E) System Adequacy Forecast 2009 – 2020 [88] states that only about 14 GW (or 1.6%) of the projected ENTSO-E Continental Europe (CE) peak load of 873 GW in 2020 can be influenced by a controlled load reduction (referred to as Load Management). A more optimistic view on DR is given in [89] with a peak reduction potential of 28 – 72 GW in Europe. However, it is important to note that these figures relate to a bird’s eye view on a diverse set of countries. According to [90], DR potential estimations mentioned in system adequacy studies are rather low in general, but most of these findings relate to interruptible load programs. Consequently, the significance of these potential estimations for the DR methods presented in this thesis can be considered relatively low.

In [91], a detailed investigation of DR potentials in Germany is given, stating a year-round flexible load potential of about 3 GW in the industry, 3 GW in the commercial sector, and 4 GW in the residential sector, which sums up to about 13% of the yearly maximum load (usually 75 – 80 GW). Additionally, seasonal potentials of 5 GW of air conditioning load and up to 18 GW of shiftable electric heating load are mentioned. Reference [92] also examines the load shifting potentials of the German power system. Flexible shares of the total load were estimated and mathematically modeled, yielding a total flexible potential of 11 – 13 GW in the residential and 7 – 10 GW in the commercial sector. These potentials were compared with other similar works. By using genetic algorithms, the flexible demand shares, modeled for a benchmark region with 500,000 inhabitants, were shifted in time in order to flatten the (residual) load curve in an optimal way. In the 2020 scenario, the achievable peak load reduction was quantified as 8.6 – 11% in the summer and as 9.5 – 14% in winter, while the minimum load increased substantially by 25 – 40%, depending on the scenario.
In summary, it can be noted that the significance of DR potential figures are strongly dependent upon the considered application. While for peak load reduction the overall power of the shiftable load during peak times is of crucial importance, other DR applications may provide substantial benefits for power system operation even with small numerical shares equal to a few percent of the total load. Considering that the total generation capacity designated as frequency control reserves is usually not more than 10 – 15% of the system’s peak load, a flexible load share of 3% can already make a significant contribution. For sophisticated DR applications such as balancing and frequency control, the smooth controllability of the resource and its usable flexible energy content appears to be more important than the sheer number of gigawatts.

This thesis is focused rather on how to exploit the available DR potentials instead of how to quantify them. We adopt in this work the perspective of treating flexible thermal load as a lossy energy storage that cannot feed energy back to the grid. In [93], an elaboration on this paradigm was presented with the intention to stress the dynamic and time-dependent nature of thermal load control potentials. In particular, this perspective captures the important fact that previous control actions have an effect on the currently available DR potential in real-time operation. The generalization of flexible thermal load within the Power Nodes Modeling Framework (PNMF), presented in Chapters 5 and 6, is formulated in accordance with this observation.

### 1.1.8 Objectives of this Work

Based on the developments described above, the objectives of the present work are:

- to develop methods for the control of flexible loads in the grid,
- to integrate controllable loads with storage devices and generation units in order to provide control services to the grid,
- to facilitate renewable energy integration by tailoring control and dispatch methods for dealing with intermittent generation,
- to explore economic aspects of demand-side integration in power systems, and
- to contribute to the aim of increasing the power system resilience by investigating novel UFLS methodologies.
1.2 Contributions

The main contributions of this thesis are:

1. Dynamic models of TCLs available from literature are extended for the usage in DR methods. A novel model of an Electric Water Heater (EWH) including thermal stratification is derived from physical principles and tailored for utility-scale DR simulations.

2. Several novel algorithms for aggregate setpoint tracking control with TCLs are presented. This includes rule-based coordination schemes and a Model Predictive Control (MPC) algorithm for the usage with a newly developed probabilistic state bin-transition model describing the behavior of the TCL population as internal model of the MPC controller.

3. A generalized framework for the system-level representation of generators, loads, and storage units in power system operation is presented. The approach, termed Power Nodes Modeling Framework (PNMF), allows the unified modeling and joint active-power dispatch of a diverse unit portfolio exhibiting different degrees of controllability. In the context of this thesis, the framework is applied mainly to the utilization of controllable loads for power system control purposes such as active-power balancing and control reserve provision.

4. Supervisory dispatch strategies based on receding-horizon optimization for the utilization of control flexibility of loads, generators, and storage devices for power system control are derived. The optimization problems including cost functions and constraint sets are formulated in the PNMF nomenclature.

5. Economic considerations on the introduction of sophisticated DR in power systems are presented. The profitability of ancillary service provision by flexible unit portfolios is investigated by parameter variations over sets of time simulations. A business value model based on the $e^3$ value methodology is presented and a methodology for sharing ancillary service profits is proposed.

6. A possible system design of a Customer-Level Under-Frequency Load Shedding (CL-UFLS) mechanism is outlined and its impacts on the power system in emergency situations is illustrated by dynamic power system simulations.
1.3 Outline of the Thesis

This thesis is structured as follows:

**Chapter 2: Modeling and Coordination of TCLs.** In this chapter, a generic model of TCLs tailored for the integration into DR methods is presented. After introducing a generalized framework for modeling flexible cooling and heating appliances, a rule-based control strategy is presented which coordinates the group in such a way that it can track a setpoint trajectory with its aggregate power. Time simulations are conducted and the implications of the coordination approach for the individual participating units are discussed.

**Chapter 3: Modeling and Coordination of Electric Water Heaters.** A dynamic model of an electric water heater, which describes energy exchanges and the thermal stratification within the storage tank, is presented. The dynamic behavior of the model is investigated by open-loop simulations using an individual unit and a large unit population. Two control strategies for populations of electric water heaters represented by the developed model are explained and evaluated by simulating the controlled system.

**Chapter 4: Probabilistic Modeling and Control of TCLs.** A probabilistic model of TCL populations based on a Markov-chain modeling technique is developed. It yields a Linear Time-Invariant (LTI) state-space representation of a TCL population by describing the “probability mass” of appliance states within the temperature dead-band. A predictive controller is used to enable the group of TCLs to track a setpoint trajectory similar to the approaches presented in Chapters 2 and 3. Time simulations are conducted to evaluate the tracking performance and the impact on the individual TCLs.

**Chapter 5: System-Level Unit Models.** The novel modeling approach called PNMF is introduced. It consists of an abstract system-level representation of units (generators, loads, and storage devices) connected to the power system which enables the unified description of a diverse portfolio of units. After explaining the theoretical concept, parameter sets for the representation of different unit types are introduced and specific models for a number of units are presented.
Chapter 6: Dispatch Strategies. Based on the system-level unit models, several dispatch strategies for individual or aggregated units in power systems (generators, loads, and storage devices) are presented. Among the use cases are least-cost dispatch, control reserve provision, and ramp-rate reduction for the residual load.

Chapter 7: Economic Evaluation of Frequency Control Provision by Flexible Unit Portfolios. Flexible unit portfolios consisting of generators, loads, and storage devices can be used to generate revenues on ancillary service markets. The profitability of ancillary service provision is largely dependent on the achievable market price levels, the cost incurred for the service provision, as well as the opportunity cost incurred for not producing energy with generation assets. In order to assess the profitability, time simulations of a number of benchmark portfolios with varying parameter sets are conducted for the provision of primary and secondary control reserves. Using the modeling methodology $e^3$value for business value models, a profit sharing method for the ancillary service operating profits is proposed.

Chapter 8: Customer-Level Under-Frequency Load Shedding. In this chapter, a novel approach to Under-Frequency Load Shedding is presented and illustrated with simulations. It takes advantage of an envisioned pervasive communication infrastructure that reaches down to the final customer level. The approach consists of a centralized assignment of frequency thresholds to customer device classes (such as refrigerators, washing machines, and lights) and a decentralized shedding action triggered by the local detection of an under-frequency event. The performance of the novel approach is compared with a Conventional Under-Frequency Load Shedding (C-UFLS) system by conducting dynamic power system simulations using a benchmark power system.

Chapter 9: Conclusions and Outlook. Some key conclusions on the research results presented in this thesis are stated and suggestions for further work are given.
1.4 List of Publications

Cited here are only the publications directly related to this thesis.

Journal Papers


Conference Papers


Chapter 2

Modeling and Coordination of TCLs

In this chapter, we present dynamic models of heating and cooling appliances which are controlled by hysteretic thermostats. A generic dynamic model describing these Thermostatically Controlled Loads (TCLs) is introduced and analyzed with respect to dynamic properties of large groups of these loads. After that, a control strategy based on sending ON/OFF switching impulses to the TCL population is introduced which can impose a time-varying setpoint on the TCL population while respecting switching boundaries of the individual TCLs. Time-domain simulations are used for demonstration.

2.1 Introduction and Literature Review

Thermostatically Controlled Loads (TCLs) have been a long-standing matter of interest to power system researchers. These loads represent a significant percentage of the overall demand in many power systems since almost all small-scale heating and cooling loads, such as refrigerators, Electric Water Heaters (EWHs), air conditioning units, and heat pumps, are controlled by hysteretic thermostats.

TCLs exhibit unique dynamic characteristics. Due to their thermal inertia, they allow a temporary curtailment of the power supply without comfort loss, which however implies that they recover the unserved energy later [94]. The energy recovery leads to the important issue called Cold Load Pick-Up (CLPU) which denotes the load peak after the extended outage of an entire TCL population caused by synchronization.

19
of the TCLs’ duty cycles (loss of temporal diversity). CLPU has been studied by numerous researchers in the last decades, e.g., [95, 96].

Common models of TCLs are first-order differential equations with a switching control input, formulated traditionally in continuous time, which can be used to derive a probabilistic model based on Partial Differential Equations (PDEs) of a TCL population’s dynamics inside a temperature dead-band assuming that the population is homogeneous [97, 98]. These works have been extended by other research endeavors, e.g., [99, 100, 101]. Control strategies based on a discretization of the TCL duty cycles have also been proposed [102]. These approaches are further elaborated in Chapter 4. TCLs can also be controlled by means of voltage variations since they show different features depending on the supply voltage magnitude [103].

Due to the challenges associated with renewable energy generation such as increased balancing requirements, the inherently flexible TCLs are now widely considered for delivering system services like balancing and frequency control. Due to the thermal inertia and synchronizing behavior that leads to the (otherwise unwanted) CLPU effect, TCLs can also be utilized for controlled power increases instead of curtailment only. In order to be able to provide sophisticated control services, the accurate tracking of a desired power consumption trajectory is necessary. This goes beyond the requirements of peak shaving applications as, e.g., mentioned in [94]. Recent approaches to controlling TCL populations include rule-based techniques for sending switching impulses to individual TCLs, e.g., [104], although the controllability is not sufficient for accurate trajectory tracking.

In this chapter, we present a generic model of TCLs based on first-order dynamics between two switching thresholds. It is conceptually similar to the other approaches that utilize first-order models but it is formulated in a normalized way that allows a unified description of heating and cooling loads and an easier representation of the electric energy that can be shifted in time by manipulating aggregate power. Based on this model, a rule-based control strategy is developed and aggregate properties of the coordinated TCL population are analyzed.

The content of this chapter is based on [105] and [106]. The chapter proceeds as follows: Section 2.2 presents a generic TCL model based on a normalized internal temperature. Section 2.3 briefly discusses the communication infrastructure required for implementing the coordina-
2.2. Generic TCL Model

In this section, a thermal modeling framework for TCLs is developed. We derive a unified representation of heating and cooling appliances of different nature. The temperature dynamics in the TCL are represented by a first-order differential equation. By using an internal hysteresis switching controller which turns the TCLs ON or OFF, the temperature is kept between two switching thresholds.
2.2.1 Normalized Expression of the TCL State

As in the other publications on TCL modeling and control mentioned above, the dynamic state variable used here is the measured internal temperature $T$ [°C]. In order to derive a unified description of the device state independent of the temperature level and device type, a description of the internal thermal energy relative to the ambient temperature $T_{\text{amb}}$ [°C] can be used. This yields for cooling and heating devices:

$$E_{\text{th,cool}} = m \bar{c} (T_{\text{amb}} - T), \quad (2.1)$$

$$E_{\text{th,heat}} = m \bar{c} (T - T_{\text{amb}}), \quad (2.2)$$

where $m$ [kg] represents the mass contained in the device and $\bar{c}$ [J/kg K] the average specific heat capacity of the contents. The switching threshold temperatures of the TCL, $T_{\text{min}}$ and $T_{\text{max}}$ [°C], can be transformed similarly. The energy content $E_{\text{th}}$ is normalized to an interval of $[0, 1]$:

$$E_{\text{rel,cool}} = \frac{T_{\text{max}} - T}{T_{\text{max}} - T_{\text{min}}}, \quad (2.3)$$

$$E_{\text{rel,heat}} = \frac{T - T_{\text{min}}}{T_{\text{max}} - T_{\text{min}}}. \quad (2.4)$$

The thermal energy of the ambiance is equal to zero in absolute terms (reference level) and usually negative in relative terms. The latter expression is obtained by substituting $T = T_{\text{amb}}$ in (2.3) and (2.4).

2.2.2 Dynamic TCL Models

Now, the differential equations describing the evolution of the TCL state are introduced. Considered are refrigerators, freezers, and EWHs. The dynamic behavior is represented in terms of the relative thermal energy, which can be easily transformed back to the internal temperature if needed.

**Refrigerator and Freezer** The relative thermal energy content in cooling appliances evolves according to

$$\frac{dE_{\text{th}}^{\text{rel}}}{dt} = -\left(\frac{1}{\tau} + \frac{1}{\tau_{\text{open}}}\right) (E_{\text{th}}^{\text{rel}} - E_{\text{th, amb}}^{\text{rel}}) + \frac{k}{\tau} u, \quad (2.5)$$

where the thermal losses depend on the difference between the ambient and the inside energy level. The binary switching input is represented
2.2. Generic TCL Model

by \( u \in \{0, 1\} \). The time constant \( \tau [s] \), the amplification factor \( k [-] \), and the initial condition of the differential equation are determined by

\[
\tau = \frac{m \bar{\alpha}}{A} ,
\]
\[
k = \frac{\varepsilon_{th} P_{\text{el}}^{\text{rated}}}{A \bar{\alpha} (T_{\text{max}} - T_{\text{min}})} ,
\]
\[
E_{\text{th},0}^{\text{rel}} = E_{\text{th}}^{\text{rel}}(t = t_0) ,
\]

with the average heat transfer coefficient \( \bar{\alpha} [\frac{W}{m^2 K}] \) and the hull surface \( A [m^2] \). The factor \( \varepsilon_{th} [-] \) is the coefficient of performance of the cooling aggregate (including the efficiency of the compressor) and \( P_{\text{el}}^{\text{rated}} [W] \) is the rated power consumption of the TCL.\(^1\) Stochastic user interactions (door openings) are modeled by the disturbance input \( d \) (combined with the heuristic time constant \( \tau_{\text{open}} [s] \) during door-openings) and occasional variations of the time constant \( \tau \) which is linearly dependent on the mass.

**Electric Water Heater** For the water heater, the governing differential equation looks similar:

\[
\frac{dE_{\text{th}}^{\text{rel}}}{dt} = -\left(\frac{1}{\tau} + \frac{\dot{m}_{\text{demand}}}{m}\right) (E_{\text{th}}^{\text{rel}} - E_{\text{th}, \text{amb}}^{\text{rel}}) + \frac{k}{\tau} u ,
\]

where \( \tau, k, \) and \( E_{\text{th},0}^{\text{rel}} \) are the same as defined in (2.6) – (2.8) and \( \varepsilon_{th} = \eta_{th} \) is the thermal efficiency of the electric heating element. The variable \( \dot{m}_{\text{demand}} [\frac{kg}{s}] \) represents the mass flow of the water which is drawn from the tank and instantly replaced by fresh water assumed to enter the storage tank at ambient temperature.

Note that this is a simplified, single-state representation of an EWH which does not capture the dynamic behavior to the full extent. A more realistic modeling approach will be presented in Chapter 3.

For all TCLs modeled in the way presented above, a hysteresis switching controller acts on the input variable \( u \) and alters its value when a switching threshold is passed:

\[
u = \begin{cases} 
1 & \text{if } E_{\text{th}}^{\text{rel}} \leq 0 \\
0 & \text{if } E_{\text{th}}^{\text{rel}} > 1 
\end{cases}
\]

\(^1\)For simplicity, the rated power is regarded as constant here, although it actually depends on the supply voltage. In the cooling appliances, active power transients during the “on” phase due to thermodynamic effects are also neglected. If desired, corresponding duty-cycle and voltage dependent factors can be included in factor \( k \).
2.2.3 Thermal and Electric Energy Content

The relative thermal energy content, which evolves in the interval \([0,1]\) during normal operation, represents a different span of thermal energy for each device between the upper and lower switching boundaries \(T_{\text{max}}\) and \(T_{\min}\). This net energy is described by

\[
E_{\text{th}}^{\text{net}} = m \tilde{c} (T_{\text{max}} - T_{\min}) . \tag{2.11}
\]

As the relation between thermal and electric input power is given by

\[
P_{\text{th}}^{\text{rated}} = \varepsilon_{\text{th}} P_{\text{el}}^{\text{rated}}, \tag{2.12}
\]

the same can be stated for the thermal and electric energy contents. Taking into account (2.7) and (2.11), it is easy to show that the net electric energy span between the two switching boundaries is

\[
E_{\text{el}}^{\text{net}} = \frac{T}{k} P_{\text{el}}^{\text{rated}} . \tag{2.13}
\]

Furthermore, it can be shown using the results from (2.1) – (2.4) that the electric and thermal relative energy contents are the same:

\[
E_{\text{el}}^{\text{rel}} = E_{\text{th}}^{\text{rel}} . \tag{2.14}
\]

As this work is focused on the electric side of the TCLs, \(E_{\text{el}}^{\text{rel}}\) will be used as the state variable in what follows.

2.2.4 Analytical Solution of the Model Equations

The differential equations (2.5) and (2.9) are identical when user interactions are neglected, i.e., \(d = \dot{m}_{\text{demand}} = 0\). In this form, they can be solved analytically for a constant input \(u \in \{0,1\}\) and constant parameters. This yields the expression

\[
E_{\text{el}}^{\text{rel}} = E_{\text{el,amb}}^{\text{rel}} + (E_{\text{el,0}}^{\text{rel}} - E_{\text{el,amb}}^{\text{rel}}) e^{-\frac{1}{\tau} (t-t_0)} + k (1 - e^{-\frac{1}{\tau} (t-t_0)}) u . \tag{2.15}
\]

When this is reformulated with respect to time, the dwell time of the TCL in its current switching state can be expressed as

\[
\Delta t_{\text{dwell}} = -\tau \ln \frac{E_{\text{el,end}}^{\text{rel}} - E_{\text{el,amb}}^{\text{rel}} - k u}{E_{\text{el,0}}^{\text{rel}} - E_{\text{el,amb}}^{\text{rel}} - k u} . \tag{2.16}
\]
with $E_{\text{rel,end}}$ being the switching boundary that the TCL will run into if left in the current state. As this is the upper relative energy boundary when the TCL is switched ON and the lower boundary when it is switched OFF, the relation $E_{\text{rel,end}} = u$ holds.

Based on (2.16), the duration $\Delta t_{\text{cycle}}$ of the full autonomous TCL operating cycle can be derived, which is composed of the full dwell times $\Delta t_{\text{dwell,on}}$ and $\Delta t_{\text{dwell,off}}$ for both the ON and the OFF phase. The initial condition $E_{\text{rel,0}}$ is set to 0 and 1, respectively. This yields

$$\Delta t_{\text{cycle}} = \Delta t_{\text{dwell,on}} + \Delta t_{\text{dwell,off}}$$

(2.17)

$$\Delta t_{\text{dwell,on}} = -\tau \ln \frac{1 - E_{\text{rel,amb}} - k}{0 - E_{\text{rel,amb}} - k}$$

(2.18)

$$\Delta t_{\text{dwell,off}} = -\tau \ln \frac{0 - E_{\text{rel,amb}}}{1 - E_{\text{rel,amb}}}$$

(2.19)

With this information, the percentage of time that a TCL is switched on during its normal operation, $p_{\text{on}}$, can be expressed as

$$p_{\text{on}} = \frac{\Delta t_{\text{dwell,on}}}{\Delta t_{\text{cycle}}}$$

(2.20)

Alternatively, the daily energy consumption of the TCL $W_{\text{el}}^{\text{daily}}$ [kWh], which is easily obtainable from measurements, can be considered relative to the rated power in order to obtain $p_{\text{on}}$:

$$p_{\text{on}} = \frac{1,000}{24 \text{ h}} \frac{W_{\text{el}}^{\text{daily}}}{P_{\text{rated}}}$$

(2.21)

### 2.2.5 Uncoordinated Steady-State Properties

A group of TCLs in steady-state uncoordinated operation can be characterized by certain stochastic properties. The goal here is to derive the steady-state mean value and standard deviation of the overall power consumption of the TCL population depending on certain parameters. Especially the former will be useful when developing the coordination approach in Section 2.4. Furthermore, the electric energy stored in the TCL population will be discussed.
In a deterministic consideration, the current overall power consumption of a group of \( n \) TCLs is equal to

\[
P_{\text{el}}^{\text{total}} = \sum_{i=1}^{n} P_{\text{el},i}^{\text{rated}} u_i .
\]  

(2.22)

However, for the purposes of this section, the power consumption behavior of a TCL \( i \) being part of an uncoordinated group of \( n \) TCLs can be expressed as a Bernoulli-distributed stochastic variable \( X_i \in \{0, 1\} \) [107] which is scaled by the rated power consumption \( P_{\text{el},i}^{\text{rated}} \) [W] of the device:

\[
P_{\text{el,unc},i} = P_{\text{el},i}^{\text{rated}} X_i .
\]  

(2.23)

Note that this representation does not include any information about the operating cycle length, which is not of interest here. Conversely, a relevant quantity is \( p_{\text{on}} \) from (2.20) resp. (2.21), which can be regarded as the probability that the TCL is switched on, i.e., \( p[X_i = 1] = p_{\text{on},i} \).

It is then possible in accordance with [108] to derive the expected value \( \mathbb{E}[P_{\text{el,unc}}] \) and the standard deviation \( \sigma[P_{\text{el,unc}}] \) of a weighted sum of Bernoulli-distributed variables representing the switching behavior of \( n \) TCLs. As the TCLs are operated independently of each other, the covariances between them can be assumed equal to zero. This yields

\[
\mathbb{E}[P_{\text{el,unc}}] = \mathbb{E}\left[ \sum_{i=1}^{n} P_{\text{el},i}^{\text{rated}} X_i \right] = \sum_{i=1}^{n} P_{\text{el},i}^{\text{rated}} p_{\text{on},i} ,
\]  

(2.24)

\[
\sigma[P_{\text{el,unc}}] = \sqrt{\text{Var}\left[ \sum_{i=1}^{n} P_{\text{el},i}^{\text{rated}} X_i \right]} = \sqrt{\sum_{i=1}^{n} (P_{\text{el},i}^{\text{rated}})^2 p_{\text{on},i} (1 - p_{\text{on},i})} .
\]  

(2.25)

A slightly different reasoning can be applied in order to obtain the average electric energy content of the population. It is easy to see that the total electric energy content can be described by the weighted sum

\[
E_{\text{el}}^{\text{total}} = \sum_{i=1}^{n} E_{\text{el},i} = \sum_{i=1}^{n} E_{\text{el},i}^{\text{net}} E_{\text{el},i}^{\text{rel}} .
\]  

(2.26)

Consequently, the overall electric storage capacity \( E_{\text{el,max}}^{\text{total}} \) is obtained
Figure 2.1: Uncoordinated operation of a TCL population consisting of 100 refrigerators and 100 freezers, oscillating around a mean power value of about 6 kW while keeping the total energy content of the storage in the middle of the allowed range by setting $E_{\text{rel},i} = 1$ for $i = 1, \ldots, n$. The average relative electric energy for a specific TCL with known parameters can now be calculated using the analytical term from (2.15). As the dynamics of $E_{\text{el}}^{\text{rel}}$ evolving in the interval [0,1] as depicted in Figure 2.1 are quite similar to straight lines, the average value can be approximated by 0.5. Thus, for the average value of the entire population in uncoordinated operation holds:

$$E_{\text{el,av,unc}} = \frac{1}{n} \sum_{i=1}^{n} E_{\text{el},i}^{\text{net}} E_{\text{el,av,unc},i} \approx 0.5 E_{\text{el,max}}^{\text{total}}. \quad (2.27)$$

The evolution of an uncoordinated TCL population (100 simulated refrigerators and 100 freezers with diverse parameter sets chosen in realistic ranges) can be seen in Figure 2.1. Figure 2.1-a) depicts the relative thermal energy content of the individual units, Figure 2.1-b shows the aggregate energy level of the population, and Figure 2.1-c) shows the aggregate power consumption including the expected value and standard deviation (calculated both analytically and numerically). The proximity of the analytical and numerical values confirms the applicability of the considerations outlined in this section.
Chapter 2. Modeling and Coordination of TCLs

2.3 Communication Infrastructure

The coordination approach for TCLs, which is the main contribution of this chapter, requires the presence of a communication system based on Information and Communication Technology (ICT) that establishes the link between the central coordination entity and the TCLs on the household level. Since measurements from the individual TCLs are needed, a two-way communication approach is proposed. Within the household, the system is composed of two kinds of units: one central Load Manager Household (LMH) device and several Load Manager Appliance (LMA) units which are installed in the individual appliances, as depicted in Figure 2.2. Note that we depict both TCLs (denoted by underlined descriptions) as well as a number of appliances that are not TCLs. The latter do not take part in the coordination approach outlined in this chapter; instead, they are used in the novel Under-Frequency Load Shedding (UFLS) methodology presented in Chapter 8. The in-house link between the LMAs and the LMH can be realized with Powerline Communication (PLC) according to Konnex PL-132 [109]. For the communication between the household and the control center, two alternatives are being considered: a Low Voltage (LV) network PLC to the
nearest transformer station combined with a subsequent transmission over a proprietary utility communication channel or a Transmission Control Protocol/Internet Protocol (TCP/IP) link over a permanent Internet connection installed in the household. A determining factor for the choice of the communication channel is the minimum achievable latency and the sampling time of the system. In comparison with the TCL time constants, these have to be small in order to enable a reasonable performance of the coordination approach.

2.4 TCL Coordination Based on “Willingness to Switch”

In this section, the coordination method, which allows to influence the power consumption behavior of the TCL population over time, is presented. It is based on a two-way information exchange between the population and the control center. A general requirement is that the TCLs must be able to toggle their switching state according to information from the control center.

2.4.1 The Coordination Strategy

The proposed strategy for coordinating the duty cycles of the TCLs in the population consists of a local and a global computation module which exchange information through a defined interface. This interface is constituted by a momentary price offer from the TCL to the control center. The offered price would be paid to the TCL if it was forced to toggle its current ON/OFF state in that very instant, instead of later by its own switching controller. The rationale behind this is the ability to penalize the duration between an enforced switching instant and the autonomous switching instant. This shall be minimized to keep the impact of the control scheme on the TCLs as small as possible.

To calculate the switching price of a single TCL, the expression

\[ c_{sw}(t) = \lambda_{sw}(t) P_{el}^{rated} = \lambda_P \frac{\Delta t_{dwell}(t)}{\Delta t_{full}^{dwell}(t)} P_{el}^{rated} \]  

(2.28)

is proposed, where \( c_{sw}(t) \) [cent] is the current switching price offer, \( \lambda_P \) [\( \frac{\text{cent}}{\text{W}} \)] is a fixed price for a change in power consumption of the population (which can be specific for a certain TCL type),
\( \Delta t_{\text{dwell}}(t) \) [s] is the estimated dwell time that the TCL will still remain in its current \( ON/OFF \) state before reaching a switching boundary, and \( \Delta t_{\text{full, dwell}}(t) \) [s] is the full autonomous dwell time in the current state that the TCL would have in uninfluenced operation.

Having the current switching offers of all participating TCLs at its disposal, the algorithm running in the control center can “accept” certain offers at a specific instant, sending a switching impulse to the appropriate TCL and paying the switching cost to the customer.\(^2\) Thus, two different computation tasks have to be solved: a calculation of a momentary switching price curve on the device level and a time-step-wise global selection of devices to be switched compulsorily, taking into account their price offers.

In the discussions about “smart” electricity infrastructure, privacy and data protection concerns have emerged which need to be addressed accordingly. Under the requirement that an individual household participating in the proposed Demand Response (DR) scheme should not be transparent to the operating company in terms of size, number, and state of installed TCLs, an individual appliance addressing should be avoided. This is achieved by replacing the individual switching impulses by accepted (“clearing”) prices, which the TCLs react to autonomously. Any TCL offering a lower price than the accepted one is switched by a local decision. As there are two groups of TCLs, one currently \( ON \) that can be switched \( OFF \), and another group currently \( OFF \) that can be switched \( ON \), two different accepted prices must be transmitted. This avoids the need for the individualized collection of data. Furthermore, communication requirements are also far lower in a “broadcast” scheme than with individual addressing.

Based on these considerations, the information interfaces between the participating units in the coordination scheme can be defined. This was elaborated in more detail in [106] and is depicted in Figure 2.3. As shown on the right side, the LMA takes over the function of the internal switching controller of the TCL, i.e., toggling the \( ON/OFF \) state \( u \) based on the measured temperature \( T \) and the desired switching thresholds \( T_{\text{min}} \) and \( T_{\text{max}} \). The LMA forwards the information via the in-house communication link to the LMH in short time intervals, where the operating cycle prediction (estimation of dwell time \( \Delta t_{\text{dwell}} \)) and the

\(^2\)Note that other compensation schemes for customer participation can be used and the switching cost can be regarded as a virtual quantity.
2.4. TCL Coordination Based on “Willingness to Switch”

calculation of the relative switching price $\lambda_{sw}$ take place. The calculated price functions reach up to the next (predicted) autonomous switching instant and are transmitted to the central coordination. Note that this shall be done only when a TCL executes a switching action or when the prediction changes significantly due to user interaction in order to reduce the need for external communication.

2.4.2 Computation of Accepted Switching Prices

The main contribution of this chapter is the design of the coordination algorithm running in the control center. It is executed in discrete time steps $t_k$ with a fixed time step size $h_{opt} = t_{k+1} - t_k$.

Note that the TCLs can in principle switch their ON/OFF state asynchronously within the time span $h_{opt}$, which may lead to significant deviations from the power setpoint and cannot be influenced until the next optimization time step $t_{k+1}$. This is avoided in the following way: all TCLs that are going to switch in the upcoming time interval ($\Delta t_{dwell}(t_k) < h_{opt}$) are switched compulsorily and “slightly prematurely” in the current optimization step $t_k$.

The algorithm executed in each time step $t_k$ thus has to aggregate all the relevant information and calculate the accepted relative switching prices (clearing prices) $\lambda_{on}^*$ and $\lambda_{off}^*$ $\left[\text{cent/W}\right]$. The price determination resembles a “pay-as-bid” auction where all the TCLs are paid according to their submitted bids (if accepted), not according to the clearing price. It is
somewhat similar to a “priority list” methodology for unit commitment as outlined in [110]. After assessing the need for a power reduction or increase in the current time step, the relative switching price offers $\lambda_{sw}$ are grouped according to $ON$ and $OFF$ TCLs and sorted in ascending order. Thus, a “merit order” over the cumulated sum of the sorted rated powers is constructed for both $ON$ and $OFF$ TCLs. This allows to determine clearing prices that lead to a certain power change.

The algorithm consists of six steps (S1 – S6), which are described below in detail. A graphical illustration of each step is found in Figure 2.4.

S1: Assemble the information: build the population vectors of $ON/OFF$ states $u$ (and thus, obtain the current power consumption $P_{\text{el, current}}^\text{total}$), dwell times $\Delta t_{dwell}$, and relative switching prices $\lambda_{sw}$. Split $\lambda_{sw}$ according to $u$ to distinguish whether the device could be switched $ON$ or $OFF$:

$$\lambda_{on} = \lambda_{sw}(u = 0), \quad \lambda_{off} = \lambda_{sw}(u = 1).$$  

This step incorporates information that is sent from the TCLs to the central coordination in the period between $t_{k-1}$ and $t_k$.

S2: Predict the switchings that will take place in the upcoming time step: $\Delta t_{dwell} < h_{opt}$ (element-wise).

S3: Using this result, calculate the minimum accepted relative switching prices that have to be paid in order to prevent asynchronous switching:

$$\lambda_{on, min}^* = \max \lambda_{on}(\Delta t_{dwell} < h_{opt}), \quad (2.30)$$
$$\lambda_{off, min}^* = \max \lambda_{off}(\Delta t_{dwell} < h_{opt}). \quad (2.31)$$

and the corresponding consumption change $P_{\text{el, sync}}^\text{total}$. Note that this calculation is easy to visualize using the price “merit orders” mentioned above, as shown in Figure 2.4. However, the calculation can be done without this using (2.30) and (2.31).

S4: Compare the current consumption with the power setpoint $P_{\text{el, set}}^\text{total}$, taking into account the switchings caused by the prevention of asynchronous behavior in order to determine the required power to be switched additionally:

$$P_{\text{el, req}}^\text{total} = P_{\text{el, set}}^\text{total} - P_{\text{el, current}}^\text{total} - P_{\text{el, sync}}^\text{total}. \quad (2.32)$$
2.4. TCL Coordination Based on “Willingness to Switch”

Figure 2.4: Steps S1 – S6 of the coordination algorithm
S5: Construct the merit orders for positive and negative compulsory switchings by sorting the price offers and building the cumulative power change curves that can be achieved (“avd”) by a certain clearing price: \( \lambda^*_{\text{on}}(P_{\text{total,el,avd}}) \) and \( \lambda^*_{\text{off}}(P_{\text{total,el,avd}}) \). Afterwards, approximate the required power change by determining accepted switching prices for ON and OFF switching actions (while observing the minimum values computed in step S3 with \( P_{\text{total,el,avd}} = P_{\text{total,el,on,avd}}(\lambda^*_{\text{on}}) - P_{\text{total,el,off,avd}}(\lambda^*_{\text{off}}) \) (2.33)

by “walking up the merit order” in the required direction such that \( P_{\text{total,el,avd}} \approx P_{\text{total,el,req}} \).

S6: Broadcast the two clearing prices to all devices which then toggle their state if their own bid is lower than the relevant clearing price and log the paid sums \( c_{sw}(t_k) = \lambda_{sw}(t_k) P_{\text{rated,el}} \) locally in the household.

2.4.3 Refinement of the Error Signal

In the algorithm outlined in the previous section, it is proposed that the achieved power change should be approximately equal to the required power change denoted by \( P_{\text{total,el,avd}} \approx P_{\text{total,el,req}} \). This requirement still contains some flexibility, which shall be discussed now. Note that the tracking accuracy that is achieved by the proposed error signal refinement is most relevant for small populations of a couple of hundred TCLs of larger size since the “granularity” imposed by the discrete number of installed TCLs will be small for a large number of small TCLs.

It is clear that a strict equality between the two quantities cannot be demanded as the number of TCLs is finite and the change in consumption can only be made in discrete steps. Thus, the error signal

\[ P_{\text{total,el,err}} = P_{\text{total,el,avd}} - P_{\text{total,el,req}} \] (2.34)

can most likely not be made exactly equal to zero. However, the properties of the error signal can be actively influenced by the way the switching algorithm deals with the “last TCL”. This term denotes the TCL the switching of which would cause an overshoot of the achieved power change beyond the required power change. There are several possibilities for deciding whether or not to switch this “last TCL”:
2.5 Aggregate Population Model

1. Always switch the “last TCL” so there is always a small overshoot of the achieved beyond the required power change.

2. Never switch the “last TCL” so there is always a small shortage of the achieved compared with the required power change.

3. Switch the “last TCL” if it reduces the absolute of the error, $|P_{el, err}^{total}|$, in that time step, otherwise do not switch it.

4. Switch the “last TCL” depending on the historical error signal values in order to ensure that the error signal is close to zero-mean over time.

Here, the last option is considered to be the most useful. This is due to the fact that a zero-mean error signal ensures that the integral of the power consumption, i.e., the consumed energy, is equal to the energy that should have been consumed according to the power setpoint. The accumulated energy difference over all elapsed time steps $j$ up to the time step $k$ can be described by the term

$$\Delta W_{el}^{total} = \sum_{j=1}^{k} \left( P_{el,j}^{total} - P_{el, set,j}^{total} \right) h_{opt}. \quad (2.35)$$

For keeping this value close to zero, the following rule for switching the “last TCL” in time step $j$ is proposed:

- if $P_{el, req}^{total} < 0$ (consumption has to be decreased),
  - if $\Delta W_{el}^{total} \leq 0$: do not switch “last TCL”,
  - if $\Delta W_{el}^{total} > 0$: switch “last TCL”,

- if $P_{el, req}^{total} > 0$ (consumption has to be increased),
  - if $\Delta W_{el}^{total} \leq 0$: switch “last TCL”,
  - if $\Delta W_{el}^{total} > 0$: do not switch “last TCL”.

2.5 Aggregate Population Model

Now an approximation for the coordinated TCL population as a whole will be derived. This aggregate representation enables the population
to be operated as a single entity showing quasi-continuous behavior, represented by a first-order dynamical system. This formulation is particularly useful for the development of power system control applications using coordinated TCL populations.

First, the dynamic behavior of the State of Charge (SOC) of the TCL population depending on the consumed electric power will be analyzed, which is essential to assess the ability of the population to perform a certain control action without reaching the storage limits. For this purpose, a differential equation of the form

$$\frac{dE_{\text{el}}^{\text{total}}}{dt} = f(E_{\text{el}}^{\text{total}}, P_{\text{el}}^{\text{total}})$$ (2.36)

is sought. Note that the stochastic user interactions with the TCLs are neglected at this stage. The total power consumption of a population was already shown in (2.22), whereas the electric energy content was introduced in (2.26), and the total storage capacity was also derived in that context. The evolution of the SOC over time can be seen when all the participating TCLs, represented by the parameterized equations (2.5) and (2.9), are combined with (2.26). For \( n \) TCLs, this leads to

$$\frac{dE_{\text{el}}^{\text{total}}}{dt} = \sum_{i=1}^{n} \left( \frac{-1}{\tau_i} (E_{\text{el},i} - E_{\text{el,amb},i}) + P_{\text{el},i}^{\text{rated}} u_i \right)$$ . (2.37)

Note that due to the presence of a weighted sum in (2.37) it is not easily possible to derive a closed-form differential equation which allows a direct solution in time. However, similar aggregation problems can be found in standard power system control theory. One prominent example is the calculation of the Center of Inertia (COI) frequency in an interconnected power system comprising a number of generators with individual frequency dynamics. This problem has been treated extensively in power system literature, see, e.g., [111].

For describing the coordinated TCL populations, a similar approach is adopted. As the coordination algorithm treats all TCLs equally (i.e., independently of their time constant, operating cycle time, etc.), the current state \( E_{\text{el},i}^{\text{rel}} \) of a TCL \( i \in \{1, \ldots, n\} \) at a specific time can be assumed to be independent of the TCL parameters. This means that the state of a TCL with a large time constant is with the same probability in a certain range as the state of a TCL with a small time constant. Due to this effect, the coordinated population shows a uniform behavior which can be adequately approximated using just one time constant.
2.6 Simulation of the Coordination

It is supposed that the differential equation below shall represent the dynamic behavior of the TCL population:

\[
d\frac{E_{\text{el}}^{\text{total}}}{dt} = -\frac{1}{\tau} (E_{\text{el}}^{\text{total}} - E_{\text{el,amb}}^{\text{total}}) + P_{\text{el}}^{\text{total}}, \quad (2.38)
\]

for which the parameters have to be determined. Strictly speaking, the aggregate time constant of the population \(\tau\) can be defined as

\[
\frac{1}{\tau} = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{1}{\tau_i} E_{\text{el},i}^{\text{total}} - \frac{1}{\tau_i} E_{\text{el,amb},i}^{\text{total}} \right). \quad (2.39)
\]

This, however, is not a useful representation as it includes the TCL states which change over time. In practice, this time constant can be obtained by considering the steady-state solution of (2.38), which yields

\[
\tau = \frac{E_{\text{el,ss}}^{\text{total}} - E_{\text{el,amb}}^{\text{total}}}{P_{\text{el,ss}}^{\text{total}}}, \quad (2.40)
\]

where the stochastically derived values from (2.24) and (2.27) can be inserted as the steady-state quantities. The initial condition to (2.38) is given by substituting \(E_{\text{el},i}^{\text{rel}} = E_{\text{el},0,i}^{\text{rel}}\) for \(i = 1, \ldots, n\) in (2.26). This completes the aggregate first-order representation of the population.

2.6 Simulation of the Coordination

We evaluate the proposed control strategy with a TCL population of 300 refrigerators, 200 freezers, and 100 water heaters (total installed power of 385.4 kW, total electric storage capacity between the temperature bounds of 27.77 kWh), simulated over six hours. A variable simulation step size and zero-crossing detection is used. The optimization step size is equal to \(h_{\text{opt}} = 0.01\) h. The group is created with statistically distributed parameter sets and it is diverse in terms of time constants (minimum 1.8 h, maximum 43.8 h), rated power (minimum 100 W, maximum 4 kW), and operating cycle times (minimum 4.8 min, max 3.2 h). As depicted in Figure 2.5, a power consumption setpoint trajectory is imposed on the population. The setpoint is equal to the expected population power consumption value (46.84 kW) in the beginning, which is then decreased by 40% for half an hour from \(t = 1\) h and increased in the same way from \(t = 2\) h. From \(t = 4\) h, a sinusoid is imposed in order to show that arbitrary curve shapes can be tracked.
Figure 2.5-a) visualizes the individual relative energy contents of arbitrarily selected TCLs within the population, demonstrating the effect of the “earlier-than-normal” switchings on individual TCLs. Figure 2.5-b) shows the aggregate electric energy storage content between the upper and lower bounds, Figure 2.5-c) shows the aggregate population power consumption and its setpoint, and Figure 2.5-d) shows the power control error, here expressed in kW. Note that the spikes in this graph at $t = \{1 \text{ h}, 1.5 \text{ h}, 2 \text{ h}, 2.5 \text{ h}, 5 \text{ h}\}$ are caused by the step-wise setpoint changes in combination with the delay of one time step that the algorithm needs to bring the actual power to the vicinity of the new setpoint. In general, it can be seen that the control algorithm keeps the overall power consumption close to the setpoint, and that the approximation of the aggregate TCL population behavior developed in Section 2.5 is valid. Note that the initial transient visible in Figure 2.5-b) from $t = 0 \text{ h}$ to $t = 1 \text{ h}$ is due to the fact that the population is initialized slightly above its steady-state SOC of $0.5 \cdot E_{el}^{\text{total,max}}$. 
2.7 Implications of the TCL Coordination

In this section, we discuss some further properties of coordinated TCL populations that can be derived from the nature of the coordination algorithm presented in the previous sections.

2.7.1 Impact of the Control on the TCLs

As seen from the time-domain simulation in the previous section, the coordination algorithm has a shortening effect on the TCL duty cycles. In order to quantify this effect, two situations have to be distinguished. First, the central optimization algorithm will compulsorily switch all TCLs that would switch by themselves in the upcoming optimization time step. This serves to prevent TCLs from switching asynchronously between the optimization time steps, which would cause a change in active power consumption that cannot be influenced by the coordination. These compulsory switchings lead to a maximum operating cycle shortening equal to the value $h_{\text{opt}}$ per switching action. For instance, if $h_{\text{opt}}$ assumes a value around 30 s for devices with an operating cycle of 10 – 30 min, the cycle shortening caused by the need to maintain synchronism is relatively insignificant, at least unless the population is operated for long times close to the energy level boundaries. The second cause of operating cycle shortening are actual control actions on the current active power consumption of the whole group of TCLs. This can be quantified as described in this section.

The steady-state solution of (2.38) yields a relation between the steady-state power consumption and SOC of the group. Imposing a different power consumption will “charge” or “discharge” the aggregated thermal storage constituted by the TCLs. When the aggregate electric energy level of the storage is around 50%, the TCLs are allowed to run their usual duty cycles comprising the whole temperature range. Note that the switchings to maintain synchronism and those to keep the power consumption accurately on the setpoint are disregarded as their impact on the TCL duty cycles is relatively small. If the SOC shall be higher than the mean value, “earlier than normal” ON switchings are triggered for TCLs that are currently OFF. Vice versa, a lower SOC is achieved when the coordination forces TCLs that are currently ON to switch themselves OFF “earlier than normal”. By simulation, the relation between the aggregate SOC and the thermal energy range in which an
individual TCL $i$ is operated is found to be approximately linear as depicted in Figure 2.6.

As a chattering of the TCLs around their lower or upper switching boundary shall be avoided, the following operational constraints are imposed on the total relative electric energy level:

$$0.1 \leq E_{\text{total,rel}}^{\text{el}} \leq 0.9 \quad .$$  \hfill (2.41)

Referring again to Figure 2.6, this establishes worst-case operation regions for the TCLs where the highest operating cycle shortening is caused. For any TCL $i$ within the coordinated population, they are defined by the intervals $E_{\text{el},i}^{\text{rel}} \in [0, 0.2]$ and $E_{\text{el},i}^{\text{rel}} \in [0.8, 1]$, which is 20% of the normal range. If the evolution of $E_{\text{el},i}^{\text{rel}}$ is close to a straight line, the operating cycle time $\Delta t_{\text{cycle},i}$ will also be shortened to 20% of the autonomous full cycle time $\Delta t_{\text{cycle},i}^{\text{full}}$ in the worst case, i.e.,

$$\Delta t_{\text{cycle},\text{min},i} = 0.2 \cdot \Delta t_{\text{cycle},i}^{\text{full}} \quad .$$  \hfill (2.42)

If the TCL has a more prominently curved “first-order” (or higher order) $E_{\text{el},i}^{\text{rel}}$ evolution, this is still a reasonable approximation as the different shortening factors for the ON and OFF phases are at least partly compensated due to the summation in (2.17). Note that normally the TCL
group will be operated around an **SOC** of 50% in order to keep the overall switching cost (and thus the operating cycle shortening) low.

In all cases, the quantification of the operating cycle shortening is an average property derived from the nature of the coordination algorithm. It is not necessarily an exact value for an individual **TCL**. However, these considerations enable an estimation of the requirements on the communication infrastructure as the length of the **TCL** duty cycles has an impact on the amount of information transmitted between the different units in the coordination scheme.

### 2.7.2 Choice of the Optimization Step Size

Having analyzed the operating cycle impact of the coordination on the **TCLs**, the question of how to choose the optimization step size $h_{opt}$ can be tackled. With regard to the coordination algorithm presented in Section 2.4, the following requirement can be stated: as the algorithm has to be aware of an upcoming switching instant in order to maintain the synchronism of the TCL switching, the optimization step size should not be larger than the minimum full dwell time of a **TCL** in an **ON** or **OFF** state. Considering (2.42), this can be expressed for a population of $n$ **TCLs** as

$$h_{opt} < 0.2 \Delta t_{dwell,\text{min}} = 0.2 \min_{i \in \{1, \ldots, n\}} (\Delta t_{dwell,\text{on},i}, \Delta t_{dwell,\text{off},i}) .$$  \hspace{1cm} (2.43)

### 2.7.3 Energy Constraint Violation Behavior

Given the population coordination strategy and the proposed constraints on the population energy level as outlined in (2.41), it is interesting to regard the dynamic behavior of the population if these constraints are violated. For this purpose, a setpoint trajectory that brings the energy level gradually to a low value is imposed on a simulated **TCL** population of 500 refrigerators.

As shown in Figure 2.7, the operating cycle shortening gets more and more severe as the minimum energy level is approached. From about $t = 0.85$ h, the constraint (2.41) for normal population operation is already violated. The coordination is still working, although with a severe impact on the **TCLs**. Conversely, the algorithm is not able to follow the setpoint anymore from $t = 0.9$ h. The **TCLs** start chattering around
Figure 2.7: Violation of the population energy level constraints from \( t = 0.8 \) h to \( t = 1 \) h caused by an unsuitable power setpoint their switching boundary and there is a substantial deviation from the power consumption setpoint. This situation is only alleviated when the power setpoint is increased again at \( t = 1 \) h. It can be concluded that an operation of the population outside the permissible energy boundaries is both useless and harmful for the TCLs, and thus should be avoided.

### 2.7.4 Adaptation of the TCL Switching Boundaries

In a real operational situation where a TCL population is used for power system control such as the provision of secondary active power reserves, it is possible that the amount of energy storage capacity will be exceeded by the power setpoint trajectory imposed on the population. As discussed in the previous section, this has to be avoided. One possible way of mitigation is, e.g., the coupling of the TCL population with a generation unit that can adapt its production in order to keep the TCL population energy level within tolerable bounds. In the case of a biased (non-zero-mean) control signal that is imposed on the population for a longer time span, the necessity for such an SOC reset mechanism is obvious as even a large storage will eventually be full or empty.
2.7. Implications of the TCL Coordination

Conversely, an easy-to-implement option for the mitigation of short-term storage overflow or depletion situations is the adaptation of the switching boundaries of the TCLs in the population during runtime. This can be implemented in the form of a parallel shifting of the lower and upper switching temperature bounds, applied uniformly to all TCLs in order to constitute a non-discriminatory measure. The shifted switching bounds are

\[
T_{\text{max,shifted}} = T_{\text{max}} + \Delta T_{\text{shift}}, \quad (2.44)
\]

\[
T_{\text{min,shifted}} = T_{\text{min}} + \Delta T_{\text{shift}}, \quad (2.45)
\]

where \(\Delta T_{\text{shift}}\) shall be defined in relation to the individual temperature range of a TCL. To this end, (2.44) and (2.45) are inserted in the relative energy equations (2.3) and (2.4). The same is done for \(T = T_{\text{amb}}\) in order to transform the relative ambient energy level. This yields

\[
E_{\text{rel,shifted}}^{\text{el}} = E_{\text{rel}}^{\text{el}} + E_{\text{rel, bias}}^{\text{el}}, \quad (2.46)
\]

\[
E_{\text{rel,amb,shifted}}^{\text{el}} = E_{\text{rel,amb}}^{\text{el}} + E_{\text{rel, bias}}^{\text{el}}. \quad (2.47)
\]

Note that an upward shift of the temperature bounds \((\Delta T_{\text{shift}} > 0)\) increases the instantaneous relative energy of a cooling appliance and decreases the instantaneous relative energy of a heating appliance, and vice versa. Consequently, the bias factors are defined in opposite ways for the two appliance types:

\[
E_{\text{rel,cool, bias}}^{\text{el}} = \frac{\Delta T_{\text{shift}}}{T_{\text{max}} - T_{\text{min}}}, \quad (2.48)
\]

\[
E_{\text{rel,heat, bias}}^{\text{el}} = -\frac{\Delta T_{\text{shift}}}{T_{\text{max}} - T_{\text{min}}}. \quad (2.49)
\]

Defined in this way, the quantity \(E_{\text{rel, bias}}^{\text{el}}\) can now be used to alter the temperature switching bounds of all TCLs in the population. This can be done both manually or automatically during runtime.

In what follows, an exemplary implementation is considered. The temperature bands of the TCLs are shifted in the following way:

- if \(E_{\text{rel}}^{\text{total,el}} \leq 0.2\), set \(E_{\text{rel, bias}}^{\text{el}} = \min(0.2 - E_{\text{rel}}^{\text{total,el}}, 0.3)\), and
- if \(E_{\text{rel}}^{\text{total,el}} \geq 0.8\), set \(E_{\text{rel, bias}}^{\text{el}} = \max(0.8 - E_{\text{rel}}^{\text{total,el}}, -0.3)\).
Figure 2.8: Adaptation of the switching boundaries from $t = 0.75$ h in order to keep the TCLs within their allowed $E_{\text{rel}}^{\text{th}}$ range.

In order to smoothen the switching boundary adaptation, a first-order time delay with a time constant of $50 \cdot h_{\text{opt}}$ is used. As depicted in Figure 2.8, the boundary adaptation takes place from about $t = 0.8$ h and ends at about $t = 1.8$ h. The relative energy shift is purposely not imposed on the approximate model in order to show that the power setpoint trajectory would actually cause a negative energy level of the population according to its dynamic behavior.

The conclusion from the above simulation is that a shift of temperature boundaries, limited to a reasonable extent in consideration of user comfort, is an effective measure to keep the TCL population in an allowed dead-band range. Although it causes a temporary violation of the desired TCL temperature ranges, it is a useful tool to prevent the population from falling out of the coordinated operation.
2.8 Concluding Remarks

This chapter demonstrated the ability of large groups of TCLs, coordinated by a rule-based coordination strategy, to act like a distributed energy storage. The key features of the approach are the tight control that it imposes on the aggregate power consumption of the group and the analytical description of the SOC dynamics. The coordination algorithm is suitable for the application in higher-level control strategies that determine the power setpoint trajectory of the population according to a certain control goal. These can be strategies for providing primary or secondary frequency control, balancing schemes for integrating in-feeds from fluctuating Renewable Energy Sources (RES), or any other power system control function related to active power.
Chapter 3

Modeling and Coordination of Electric Water Heaters

In this chapter, we present a dynamic model of an Electric Water Heater (EWH) which describes the temperature dynamics including the thermal stratification in the storage tank. Based on this model, rule-based aggregate power control strategies (similar to the ones in Chapter 2) for populations of EWHs are designed and tested in simulations.

3.1 Introduction and Motivation

In many countries, Electric Water Heaters (EWHs) represent a significant share of the thermal load in power systems. Especially in countries with traditionally low electricity generation cost, the usage of electricity for water heating has been a widely adopted and economically attractive option in the last few decades. From an energy systems point of view, however, electric water heating has a number of downsides. If the used electricity is generated by fossil-fuel-fired thermal power plants, the overall efficiency of the water heating process chain is low compared to using a calorific gas-fired heater and a large carbon footprint is generated. Furthermore, the exergetic efficiency [112] of heating water to about 60°C (low exergy) with electricity (exergy = energy) is unfavorable. For these reasons, the installation of electric-only water heaters has been banned in various countries, e.g., in Switzerland [113].

In spite of this controversial assessment of electric water heating, we put a novel effort into modeling and managing the demand of EWHs in this work. This is due to the fact that the role of thermal loads may
become entirely different in an energy system that is largely based on renewable energies. Energy penetrations of wind power above about 25%, or of solar photovoltaic (PV) power generation above about 10 – 15%, are already likely to trigger in-feed situations in which intermittent production alone exceeds demand. Along with the so-called “must-run” generation, the superposition of wind and solar is likely to produce large quantities of power in the nearer future that either will have to be curtailed or consumed purposely by actively managed loads or storage devices. If EWHs can contribute to utilizing available energy that would otherwise be wasted, the energetic and exergetic efficiency can be considered of secondary importance. It may turn out that legislation concerning actively managed thermal loads will be reconsidered as more intermittent renewable energies penetrate the power systems around the globe, provided that suitable infrastructure and methodologies for active management of heating and cooling loads exist. Furthermore, resistive water heaters may also be replaced by heat pump water heaters, which improves the energetic (and exergetic) efficiency.

Having presented the motivation for this work, we now look at the requirements for accurately representing the behavior of EWHs in the power system. EWHs are usually controlled by a hysteretic thermostat which switches the heater ON and OFF depending on a measured temperature. In general, the internal temperature of the storage tank can be described by a single state if the assumption of a well-mixed tank holds. Real cylindrical storage tanks installed in an upright (vertical) position, however, usually exhibit a strong temperature stratification, meaning that the temperature at the top of the tank is much higher than at the bottom. This effect is caused by the demanded hot water being drawn from the top of the tank and the supplied cold water being injected at the bottom, and by density differences between cold and hot water which imply different buoyancy values.

Temperature stratification has a strong influence on the dynamic energy consumption behavior of EWHs. If only one heating element with an attached thermostat is present, it is usually located in the lower third of the tank. This means that a stratified tank subject to water draws will demand electric energy much sooner than a well-mixed tank since the measured temperature will fall below the thermostat switching threshold as soon as cold water rises above the position of the thermostat. This suggests that a well-mixed tank should not be assumed when the tank to be modeled is actually stratified since a single-state model
capturing the average tank temperature will exhibit a different power consumption behavior. Furthermore, the fact that hot water may still be available to the user at the top of the tank, even if the average tank temperature is already quite low, suggests that average temperature is not a good proxy for the degree of end use functionality that a stratified tank provides. Since all Demand Response (DR) methodologies have to balance end-use performance with the aggregate control flexibility, an accurate representation of the end-use functionality is necessary.

The temperature dynamics of hot water storage tanks are highly complex and represents a research domain of its own. Three-dimensional models of the water mass, as, e.g., used in Computational Fluid Dynamics (CFD), are required to capture the details of heat and mass transfers. This kind of model, however, is designed to study effects within a single water tank, e.g., the influence of the tank geometry on stratification and efficiency. Representing an aggregation of several thousand water heaters for DR applications requires a strong simplification of the model. This is possible since the individual properties of a single tank are not crucial for the behavior of the entire group.

### 3.2 Literature Review

Water heater models including stratification have been studied predominantly by researchers in thermal engineering and in solar-thermal energy. These studies are usually focused on increasing the efficiency of a single water heater. Reference [114] presents a detailed model of the thermodynamics of a solar water tank. Reference [115] describes a laboratory experiment with a real EWH in which the water temperature is measured in various locations along the tank height during water draws and heating periods. Reference [116] presents dynamic models, discretized in one spatial dimension, for various physical effects such as heat loss to the ambiance and heat exchanges inside the tank. Reference [117] studies the effect of inlet velocities. A comprehensive description of various discretization methods applied to the so-called “heat equation” which describes the heat diffusion in a medium is presented in [118].

Utilizing EWHs for DR has already been studied decades ago, long before large Renewable Energy Sources (RES) shares and the transition to a “SmartGrid” could be anticipated. In these times, EWH modeling was considered relevant due to the management of the Cold Load
Table 3.1: Notation for Chapter 3

<table>
<thead>
<tr>
<th>Var.</th>
<th>Unit</th>
<th>Meaning</th>
<th>Var.</th>
<th>Unit</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>[-]</td>
<td>Dynamic matrix</td>
<td>P</td>
<td>[W]</td>
<td>Power</td>
</tr>
<tr>
<td>A</td>
<td>[m²]</td>
<td>Area</td>
<td>p</td>
<td>[-]</td>
<td>Probability</td>
</tr>
<tr>
<td>B</td>
<td>[-]</td>
<td>Input matrix</td>
<td>s</td>
<td>[-]</td>
<td>Switching signal</td>
</tr>
<tr>
<td>b</td>
<td>[-]</td>
<td>Blocking signal</td>
<td>T</td>
<td>[°C]</td>
<td>Temperature</td>
</tr>
<tr>
<td>C</td>
<td>[-]</td>
<td>State meas. matrix</td>
<td>t</td>
<td>[s]</td>
<td>Time</td>
</tr>
<tr>
<td>D</td>
<td>[-]</td>
<td>Input meas. matrix</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>[m]</td>
<td>Disk width</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EXP</td>
<td>[-]</td>
<td>Exponential distr.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>[kWh]</td>
<td>Energy</td>
<td>v</td>
<td>[m/s]</td>
<td>Water velocity</td>
</tr>
<tr>
<td>h</td>
<td>[m]</td>
<td>Tank height</td>
<td>v</td>
<td>[-]</td>
<td>Internal sw. state</td>
</tr>
<tr>
<td>k</td>
<td>[-]</td>
<td>Time step</td>
<td>γ</td>
<td>[-]</td>
<td>Control signal</td>
</tr>
<tr>
<td>m</td>
<td>[kg]</td>
<td>Mass</td>
<td>η</td>
<td>[-]</td>
<td>Efficiency</td>
</tr>
<tr>
<td>ṁ</td>
<td>[kg/s]</td>
<td>Mass flow</td>
<td>λ</td>
<td>[-]</td>
<td>Expected value</td>
</tr>
<tr>
<td>N</td>
<td>[-]</td>
<td>Water heater set</td>
<td>μ</td>
<td>[-]</td>
<td>Mean</td>
</tr>
<tr>
<td>N</td>
<td>[-]</td>
<td>Normal distribution</td>
<td>ρ</td>
<td>[J/kg]</td>
<td>Density</td>
</tr>
<tr>
<td>n</td>
<td>[-]</td>
<td>Number</td>
<td>σ</td>
<td>[-]</td>
<td>Standard deviation</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subscript</th>
<th>Meaning</th>
<th>Subscript</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Initial</td>
<td>internal off</td>
<td>Off by internal controller</td>
</tr>
<tr>
<td>active off</td>
<td>Switched off actively</td>
<td>internal on</td>
<td>On by internal controller</td>
</tr>
<tr>
<td>active on</td>
<td>Switched on actively</td>
<td>j</td>
<td>Water heater index</td>
</tr>
<tr>
<td>agg</td>
<td>Aggregate</td>
<td>l</td>
<td>Heating element index</td>
</tr>
<tr>
<td>amb</td>
<td>Ambient</td>
<td>loss</td>
<td>Heat loss</td>
</tr>
<tr>
<td>bottom, b</td>
<td>Tank bottom</td>
<td>mix</td>
<td>Mixing zone</td>
</tr>
<tr>
<td>cat</td>
<td>Category</td>
<td>norm</td>
<td>Normalized</td>
</tr>
<tr>
<td>circflow, cf</td>
<td>Circular flow</td>
<td>off</td>
<td>Switched off</td>
</tr>
<tr>
<td>daily</td>
<td>Daily</td>
<td>on</td>
<td>Switched on</td>
</tr>
<tr>
<td>dead</td>
<td>Dead-band</td>
<td>on cand</td>
<td>Units for switch-on</td>
</tr>
<tr>
<td>diff</td>
<td>Heat diffusion</td>
<td>out</td>
<td>Water outlet</td>
</tr>
<tr>
<td>disk</td>
<td>Water layer</td>
<td>plug</td>
<td>Plug flow</td>
</tr>
<tr>
<td>draw</td>
<td>Water draw</td>
<td>sample</td>
<td>Sampling time</td>
</tr>
<tr>
<td>duration</td>
<td>Draw duration</td>
<td>set</td>
<td>Setpoint</td>
</tr>
<tr>
<td>e</td>
<td>Heating element</td>
<td>sim</td>
<td>Simulation</td>
</tr>
<tr>
<td>effective</td>
<td>Effective</td>
<td>th</td>
<td>Thermal</td>
</tr>
<tr>
<td>el</td>
<td>Electric</td>
<td>too cold</td>
<td>Below dead-band</td>
</tr>
<tr>
<td>hourly</td>
<td>Hourly values</td>
<td>too hot</td>
<td>Above dead-band</td>
</tr>
<tr>
<td>i</td>
<td>Disk index</td>
<td>top, t</td>
<td>Top of the tank</td>
</tr>
<tr>
<td>in dead-band</td>
<td>Within temp. dead-band</td>
<td>wall, w</td>
<td>Tank wall</td>
</tr>
<tr>
<td>inflow</td>
<td>Water inflow</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Pick-Up (CLPU) effect, as well as for peak shaving. Many known approaches utilize a one-state dynamic model with the assumption of an ideally mixed tank, which makes the problem a generic Thermostatically Controlled Load (TCL) aggregation and control problem, as treated in Chapters 2 and 4, so the references of these chapters are also relevant.
3.3. Single Water Heater Model

for the discussion here. If needed, specific properties of EWHs can be added to the model, as shown in [119]. The control strategies presented in the last part of this chapter are based on [120].

The remainder of this chapter is structured as follows: Section 3.3 presents the dynamic model of a single water heater based on physical effects within the water heater and some heuristic considerations. Apart from that, we present a modeling approach for user-induced water draws based on stochastic distributions and simulate an individual water heater in autonomous operation. Section 3.4 presents an approach to assemble a population of water heaters with statistical parameter distributions. Finally, Section 3.5 presents two control strategies that are able to impose a setpoint on the aggregate power consumption of a water heater population, and Section 3.6 concludes this chapter. Table 3.1 presents the notation used in this chapter.

3.3 Single Water Heater Model

In this work, we use a one-dimensional model of the temperature dynamics within the water tank. This is less computationally intense than a three-dimensional model but still allows to describe the dynamic behavior of the water heater sufficiently well. We consider a tank with two heating elements in the general case, as depicted in Figure 3.1, left.

We will utilize a spatial discretization along the height of the tank and subdivide the storage volume into a finite number of disks $N_{\text{disk}}$, as shown on the right side of Figure 3.1. We will formulate mass and energy balances for these individual disks in order to describe the dynamics within the tank. We distinguish between the following four disk types:

**Regular disk d):** Regular disks exchange heat and mass with their adjacent disks below and above and dissipate heat to the ambiance through the side wall of the tank.

**Mixing zone disk e):** A number of $N_{\text{mix}}$ disks around a heating element are assumed to be equally affected by the heating element (equal heat input once the heating element is turned on). This relates to the fact that the heating elements introduce a large amount of turbulence and mixing when they inject heat into the water. Otherwise, they function like regular disks.
Figure 3.1: Left: schematic representation of an EWH. Right: Division into finite disks. Labels: a) lower and upper heating elements, b) hot water outlet, c) cold water inlet, d) regular disks, e) mixing zone disk, f) top disk, g) bottom disk

**Top layer disk f):** The top layer disk exchanges heat and mass only with the disk below and dissipates heat to the ambiance through the top and the side wall of the tank.

**Bottom layer disk g):** This disk exchanges heat and mass with the disk above and dissipates heat through the bottom and the side wall of the tank. Furthermore, it receives the cold water from the inlet.

### 3.3.1 Physical Effects

We will explain in the following the physical effects governing the EWH dynamics, distinguished according to the disk type. The properties of water (density, heat conductivity, and specific heat capacity) are assumed to be constant and independent of temperature.

**Regular Disk**

Figure 3.2 presents the energy and mass balance for a regular disk. Following [116, 118], the rate of change in the temperature $\frac{K_s}{s}$ of the $i^{th}$ heating element due to heat diffusion is given by the following equation:

$$
\left( \frac{dT_i}{dt} \right)_{\text{diff}} = \frac{k}{\rho c d^2}(T_{i-1} - 2T_i + T_{i+1})
$$

(3.1)
for all disks $i = 1, \ldots, N_{\text{disk}}$, where $k \left[ \frac{\text{W}}{\text{m} \cdot \text{K}} \right]$ is the heat conductivity of the water, $T \left[ ^\circ\text{C} \right]$ is the disk temperature, $T_{i-1} \left[ ^\circ\text{C} \right]$ and $T_{i+1} \left[ ^\circ\text{C} \right]$ are the (lower and upper) adjacent disk temperatures, $\rho \left[ \frac{\text{kg}}{\text{m}^3} \right]$ is the water density, $c \left[ \frac{\text{J}}{\text{kg} \cdot \text{K}} \right]$ is the specific heat capacity of water, and $d$ [m] is the width of the disk in the direction of the tank height. The temperature change induced by the “plug flow” resulting from water draws is described by

$$
\left( \frac{dT_i}{dt} \right)_{\text{plug}} = \frac{\dot{m}}{m_i} (T_{i-1} - T_i),
$$

where $\dot{m} \left[ \frac{\text{kg}}{\text{s}} \right]$ is the mass flow rate of the drawn water and $m_i$ [kg] is the mass of the water in disk element $i$.

The rate of change in temperature due to heat losses to the environment is modeled by:

$$
\left( \frac{dT_i}{dt} \right)_{\text{loss}} = \frac{A_{\text{wall},i}}{m_i c} \left( T_i - T_{\text{amb}} \right),
$$

where $A_{\text{wall},i}$ [m$^2$] is the ring-shaped tank wall surface around disk $i$, $U \left[ \frac{\text{W}}{\text{m}^2 \cdot \text{K}} \right]$ is the heat transfer coefficient, and $T_{\text{amb}} \left[ ^\circ\text{C} \right]$ is the ambient temperature.

**Mixing Zone Disk**

Figure 3.3 presents a mixing zone disk adjacent to a heating element along with its associated energy balance terms. We model the heat
input into the mixing zone disk at heating element \( l \) by the term

\[
\left( \frac{dT_i}{dt} \right)_e = \frac{\eta_l P_{\text{el},l}^{\text{rated}}}{m_i c N_{\text{mix}}} u ,
\]

where \( \eta_l \) [-] and \( P_{\text{el},l}^{\text{rated}} \) [W] are the efficiency and the power rating of the heating element \( l \), respectively. \( N_{\text{mix}} \) is the number of disk elements that are considered to be in the heating zone of the element (all elements are assumed to get an equal share of the energy input from the element). The binary variable \( u \) represents the \( \text{ON} / \text{OFF} \) state of the heating element. All other terms are equal to those of the regular disk.

**Top Layer Disk**

Figure 3.4, left, shows the top layer disk including the term describing the energy dissipation to the ambiance:

\[
\left( \frac{dT_{\text{disk}}}{dt} \right)_{\text{loss, top}} = \frac{(A_{\text{wall},N_{\text{disk}}} + A_{\text{top}}) U}{m_{\text{N}_{\text{disk}}}} (T_{\text{disk}} - T_{\text{amb}}),
\]

**Figure 3.3: Energy and mass balance for a mixing zone disk**

**Figure 3.4: Energy and mass balance for a top layer disk (left) and a bottom layer disk (right)**
3.3. Single Water Heater Model

where $A_{\text{top}}$ is the surface area of the top of the tank. The diffusive heat exchange is limited to the adjacent disk below. The plug flow $\dot{m}_{\text{cf}} \left[ \frac{\text{kg}}{\text{s}} \right]$ gets ejected from the water tank from this disk.

**Bottom Layer Disk**

Figure 3.4, right, shows the bottom layer disk, also including the following term describing the heat loss to ambiance:

$$
\left( \frac{dT_1}{dt} \right)_{\text{loss,bottom}} = \frac{(A_{\text{wall,1}} + A_{\text{bottom}}) U m_1 c}{a_{w,b}} (T_1 - T_{\text{amb}}), \quad (3.6)
$$

where $A_{\text{bottom}}$ is the surface area of the tank bottom. The plug flow transports water from the inlet to this lowest disk of the tank. Diffusive heat exchange takes place with the adjacent disk above.

### 3.3.2 Heuristic Circular Mass Flow Term

Apart from the physical effects in the water heater, we introduce an additional heuristic mass flow that circulates within the water tank once a heating element is switched on. The rationale for this is the amount of turbulence and mixing that the heating element induces in the water column by natural convection. When relying on heat diffusion alone for distributing the heat injected by the heating elements within the tank, un-physical artifacts such as severe overheating of the mixing zone elements can be observed. A detailed modeling of density differences and buoyancy of hot water, however, would complicate the model substantially. For this reason, we utilize the heuristic but physically plausible circular mass flow term.

Figure 3.5 depicts the circular mass flow. The corresponding equation is

$$
\left( \frac{dT_i}{dt} \right)_{\text{circflow}} = \frac{\dot{m}_{\text{cf}}}{m_i} (T_{i-1} - 2T_i + T_{i+1}), \quad (3.7)
$$

where $\dot{m}_{\text{cf}} \left[ \frac{\text{kg}}{\text{s}} \right]$ is the magnitude of the heuristic mass flow. We consider this term for all disks above the lower edge of the mixing zone pertaining to a heating element. This is due to the upward nature of the convective effect emanating from the heating element. We can also describe the circular flow in terms of circular flow water velocity $v_{\text{cf}} \left[ \frac{\text{m}}{\text{s}} \right]$:

$$
\dot{m}_{\text{cf}} = \frac{m_i}{d} v_{\text{cf}}. \quad (3.8)
$$
We choose $v_{\text{cf}} = 0.05 \text{ m/s}$ for the simulations presented in this chapter, which yields a plausible mixing behavior without assuming the existence of excessively fast water movements within the water tank.

### 3.3.3 Merging the Model Pieces

Having defined the physical effects and a heuristic additional mass flow impacting the dynamic behavior of the water within the tank, we will merge the entire set of equations into a compact model. We use the following Linear Parameter-Varying (LPV) model:

$$\dot{x} = A(\dot{m}, u) x + B u,$$

(3.9)

where $A(\dot{m}, u)$ is the parameter-varying dynamic matrix depending on the mass flow $\dot{m}$ and the binary switching variable vector $u$, and $B$ is the input matrix. The state and input vectors are defined as follows:

$$x = [T_{\text{inflow}}, T_1, T_2, \ldots, T_{N_{\text{disk}}}, T_{\text{amb}}]^T,$$

(3.10)

$$u = [u_1, u_2]^T,$$

(3.11)

where $T_{\text{inflow}}$ [°C] denotes the temperature of the water inflow and $T_{\text{amb}}$ [°C] denotes the ambient temperature (usually room temperature of about 20°C in the case of EWHs). The temperatures $T_1, \ldots, T_{N_{\text{disk}}}$ are associated with the $N_{\text{disk}}$ discrete water layers within the tank, and $u_1$ and $u_2$ are switching variables for the two heating elements present in the EWH. Note that there might only be one heating element installed – in this case, the second element of $u$ is constrained to zero.

The dynamic matrix $A(\dot{m}, u)$ consists of the following components:

$$A(\dot{m}, u) = A_{\text{diff}} + A_{\text{plug}}(\dot{m}) + A_{\text{circflow}}(u) + A_{\text{loss}},$$

(3.12)
3.3. Single Water Heater Model

where $A_{\text{diff}}$ contains terms describing the heat diffusion inside the tank, $A_{\text{plug}}(\dot{m})$ contains the so-called "plug flow" model which describes the water mass flow moving upwards through the tank during water draws, and $A_{\text{loss}}$ contains the model for heat losses to the ambiance. The matrices are defined as follows:

$$A_{\text{diff}} = \begin{bmatrix}
0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & -\frac{k}{\rho c d^2} & \frac{k}{\rho c d^2} & 0 & \cdots & 0 & 0 \\
0 & \frac{k}{\rho c d^2} & -\frac{2k}{\rho c d^2} & \cdots & \ddots & \vdots & 0 \\
0 & 0 & \ddots & \ddots & \ddots & \ddots & 0 \\
\vdots & \vdots & \ddots & -\frac{2k}{\rho c d^2} & \frac{k}{\rho c d^2} & 0 & 0 \\
0 & 0 & \cdots & 0 & \frac{\dot{m}}{m_1} & -\frac{\dot{m}}{m_1} & 0 \\
0 & 0 & \cdots & 0 & 0 & \frac{\dot{m}}{m_2} & -\frac{\dot{m}}{m_2} \\
0 & 0 & \cdots & 0 & 0 & 0 & \ddots \\
0 & 0 & \cdots & 0 & 0 & 0 & 0 \\
\end{bmatrix}, \tag{3.13}$$

$$A_{\text{plug}}(\dot{m}) = \begin{bmatrix}
0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
\frac{\dot{m}}{m_1} & -\frac{\dot{m}}{m_1} & 0 & 0 & \cdots & 0 & 0 \\
0 & \frac{\dot{m}}{m_2} & -\frac{\dot{m}}{m_2} & \vdots & \vdots & \vdots \\
0 & 0 & \ddots & \ddots & \ddots & \ddots & 0 \\
\vdots & \vdots & \ddots & -\frac{2k}{\rho c d^2} & \frac{k}{\rho c d^2} & 0 & 0 \\
0 & 0 & \cdots & 0 & \frac{\dot{m}}{m_{N_{\text{disk}}}} & -\frac{\dot{m}}{m_{N_{\text{disk}}}} & 0 \\
0 & 0 & \cdots & 0 & 0 & 0 & \ddots \\
0 & 0 & \cdots & 0 & 0 & 0 & 0 \\
\end{bmatrix}, \tag{3.14}$$

$$A_{\text{loss}} = \begin{bmatrix}
0 & 0 & 0 & \cdots & \cdots & 0 & 0 & 0 \\
0 & -a_{b,w} & 0 & \cdots & \cdots & 0 & 0 & a_{b,w} \\
0 & 0 & -a_w & \cdots & \cdots & a_w & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \\
\vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \\
0 & 0 & \cdots & 0 & 0 & -a_w & 0 & a_w \\
0 & 0 & \cdots & 0 & 0 & 0 & -a_{t,w} & a_{t,w} \\
0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}, \tag{3.15}$$

where $a_{t,w}$ is the heat loss term for the top disk, $a_{b,w}$ for the bottom disk, and $a_w$ for all other disks. The heuristic circular flow term is expressed in the following way:
Chapter 3. Modeling and Coordination of EWHs

\[ A_{\text{circflow},1}(u_1) = \begin{bmatrix} 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & \frac{\dot{m}_{\text{cf},1} u_1}{m_i} & 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & \frac{\dot{m}_{\text{cf},1} u_1}{m_i} & 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & \frac{2 m_{\text{cf},1} u_1}{m_i} & \cdots & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & \frac{\dot{m}_{\text{cf},1} u_1}{m_i} & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 & \frac{\dot{m}_{\text{cf},1} u_1}{m_i} & 0 \end{bmatrix}, \] (3.16)

where the location of the terms within the matrix is determined by the location of the corresponding heating element and its mixing zone. If two heating elements are present, we define one circular flow matrix for each of them. The control input which represents the effect of the heating elements is associated with the mixing zone disks:

\[ B = \begin{bmatrix} 0 & \cdots & 0 & b_1 & \cdots & b_1 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 & b_2 & \cdots & b_2 & 0 & \cdots & 0 \end{bmatrix}^T, \] (3.17)

where \( b_1 \) and \( b_2 \) are the heat input terms pertaining to heating element 1 and 2 according to (3.4). The output equation defines which state and input variables are measured:

\[ y = C x + D u . \] (3.18)

We assume as an example that we measure the bottom and the top temperature of the tank and the power consumed by two heating elements. In this case, the matrices \( C \) and \( D \) look like:

\[ C = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}. \] (3.19)

In order to simulate the EWH dynamics, we discretize the model and perform a step-wise iteration of the resulting difference equation.
3.3. Single Water Heater Model

3.3.4 Hysteretic Thermostat Controller

A determining factor for the temporal evolution of the EWH’s power demand is the hysteretic thermostat, which is usually built into the heating element. Its main function is to turn the heating element ON and OFF depending on the measured water temperature. In the case of an autonomously operating EWH without any external control input, the following relation holds:

\[
    u_l(k) = \begin{cases} 
    0 & \text{if } T_{e,l}(k) \geq T_{e,l}^{\text{max}} \\
    1 & \text{if } T_{e,l}(k) \leq T_{e,l}^{\text{min}} \\
    u_l(k-1) & \text{if } T_{e,l}^{\text{min}} < T_{e,l}(k) < T_{e,l}^{\text{max}}
    \end{cases}, \tag{3.20}
\]

where \( u_l(k) \) is the binary variable denoting the power consumption of heating element \( l \) at time step \( k \), \( T_{e,l}(k) \) is the measured temperature at the heating element \( l \), and \( T_{e,l}^{\text{min}} \) and \( T_{e,l}^{\text{max}} \) are the lower and upper thresholds of the temperature dead-band.

In Section 3.5, we will deal with a control strategy that can block the power consumption of the EWH although its internal switching state is ON. In this case, it is important to distinguish between the internal switching state (unit “wants” to consume power or not) and its actual power consumption. Affirming the meaning of the variable \( u \) as the actual power consumption of the EWH, we introduce a new variable \( v \) for the internal switching state. Now the relation from (3.20) is formulated in \( v \). In unimpeded operation (EWH can consume power when the thermostat determines to do so), \( v = u \) holds by definition. A blocking mechanism which is able to cut the power supply to the EWH can be described by a binary variable \( b \), where \( b = 0 \) represents the unblocked state and \( b = 1 \) represents the blocked state. This yields the following relation between \( u \), \( v \), and \( b \):

\[
    u = v (1 - b) \tag{3.21}
\]

If two heating elements are present in the EWH, they usually operate in “interlocked” mode. This means that only one of the elements can be turned on at a time, and that one element is dominant over the other. Usually, the upper heating element, denoted with index \( l = 2 \), serves as an “emergency heating” in the case of a cold water tank and an immediate demand for hot water. Here, the upper element is defined as the dominant one. The interlock is then mathematically formulated
as follows:

\[
\begin{align*}
v_2(k) &= \begin{cases} 
0 & \text{if } T_{e,2}(k) \geq T_{e,2}^{\text{max}} \\
1 & \text{if } T_{e,2}(k) \leq T_{e,2}^{\text{min}} \\
v_2(k-1) & \text{if } T_{e,2}^{\text{min}} < T_{e,2}(k) < T_{e,2}^{\text{max}} 
\end{cases}, \quad (3.22) \\
v_1(k) &= \begin{cases} 
0 & \text{if } T_{e,1}(k) \geq T_{e,1}^{\text{max}} \lor v_2(k) = 1 \\
1 & \text{if } T_{e,1}(k) \leq T_{e,1}^{\text{min}} \land v_2(k) = 0 \\
v_1(k-1) & \text{otherwise}
\end{cases} \quad (3.23)
\end{align*}
\]

If the user specifies that only one of the heating elements is present, the algorithm simply operates this element according to its specified temperature limits.

The presented switching logic produces a so-called “sticky” thermostat state. This means that the thermostat remembers its state even if the actual ON/OFF state of the heater is altered during its transition to the opposite dead-band limit. For example, if a device has been turned ON \((v = 1)\) upon reaching the minimum temperature but is switched OFF (e.g., due to a power outage or an external control signal, i.e., \(u = 0\)) before it reaches the maximum temperature, the control signal \(v\) will remain in its original state (i.e., \(v = 1\)). It is assumed that the bimetallic strips used in thermo-mechanically controlled thermostats remember their state and the switching behavior produced by such devices might therefore be classified as sticky. For digitally controlled thermostats, however, the control signal may or may not be sticky depending on how the device is programmed [121]. Note that we will only consider EWHs with one heating element here and leave a detailed analysis of the implications of a second heating element to future research.

### 3.3.5 Modeling of Water Draws

In order to model random water draw events induced by customer behavior, we propose to generate a draw scenario for a single EWH by taking random values for draw starting time, draw duration, and flow rate from predefined probability distributions. For determining the draw times, a typical hourly water consumption profile (stating the probability of a water draw occurring) is used along with an exponential distribution for the time intervals between draws. The water draw duration and flow rate is drawn from another random distribution. The generated time series is then scaled to a previously determined total water consumption during the day. By this method, we are able to create
3.3. Single Water Heater Model

Figure 3.6: Profile describing the probability of a water draw occurring during the course of the day

a time series of mass flow rates $\dot{m}$ for the simulated time span. In what follows, we describe how we assemble the probability distributions and present a sample draw event time series.

**Probability Distributions**

For generating the water draw time series from a set of probability distributions, we proceed as follows:

1. For determining the probability of a draw event, we utilize a profile from [122] depicted in Figure 3.6 and further denoted by $p_{\text{draw}}(t)$. Furthermore, we determine the total number of water draws per day by drawing from a uniform distribution:

   $$n_{\text{daily}} \sim U(n_{\text{min}}^{\text{daily}}, n_{\text{max}}^{\text{daily}}), \quad (3.24)$$

   where $U(n_{\text{min}}, n_{\text{max}})$ denotes the uniform distribution between a minimum and a maximum value. We multiply the hourly draw probability with the total number of draws of the day to determine the hourly average water draws:

   $$n_{\text{avg hourly}}(t) = p_{\text{draw}}(t) \cdot n_{\text{daily}}. \quad (3.25)$$

2. We determine the vector of time intervals between water draws during the course of the day by

   $$\Delta t_{\text{draw}} \sim \mathcal{E}(\lambda'), \quad (3.26)$$

   where $\mathcal{E}(\lambda)$ is the exponential distribution with the expected value $\lambda$. By calculating a cumulative sum of $\Delta t_{\text{draw}}$, we assemble the vector $t_{\text{draw}}$ of time instants which denote the points in time when a draw takes place.
3. Now we need to assemble the draw duration vectors which are associated with the draw instants contained in \( t_{\text{draw}} \). We distinguish between two types of water draws: long draws that correspond to, e.g., a shower or bath, and shorter draws related to activities such as washing hands or cooking. We assume that a water draw is more likely to be a long one during times of the day when a lot of water is used. Therefore, we draw a random number from a uniform distribution \( U(0,1) \) and compare it with the value of the probability profile \( p_{\text{draw}}(t) \) scaled by a certain factor. When the random number is greater than the profile value, a short draw is assumed, and when it is below the profile value, a long draw is assumed. The profile is scaled such that, in total, a 50% chance for long and short water draws during the course of the day is obtained. The draw duration times are assembled in the vector \( \Delta t_{\text{duration}} \).

4. The flow rate will be determined by the total amount of hot water consumed during the day. In order to induce a variation of flow rates, we draw a normalized flow rate from a normal distribution for every draw event:

\[
\dot{m}_{\text{draw}} \sim \mathcal{N}(1, \sigma_{\dot{m}_{\text{draw}}}),
\]

where \( \mathcal{N}(\mu, \sigma) \) denotes the normal distribution with the mean value \( \mu \) and and the standard deviation \( \sigma \).

5. The overall resulting normalized mass flow time series \( \dot{m}_{\text{norm}}(k) \) is assembled (according to sampling instants \( k \) with the sampling time \( t_{\text{sample}} \)) with the information contained in the vectors \( t_{\text{draw}}, \dot{m}_{\text{draw}}, \) and \( \Delta t_{\text{duration}} \).

6. Finally, the draw time series is scaled such that the total daily consumption equals \( m_{\text{daily}} \).

**Numerical Example**

Having defined the methodology for creating a mass flow time series describing the water draws, we present an exemplary parameterization and a water draw time series. Table 3.2 displays the parameters used. Figure 3.7 depicts the exemplary water draw scenario.
Table 3.2: Assumed stochastic parameters for water draw scenarios

<table>
<thead>
<tr>
<th>Variable</th>
<th>Meaning</th>
<th>Distr.</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_{\text{daily}}$</td>
<td>Daily water usage</td>
<td>Normal</td>
<td>$\mu = 200$ l, $\sigma = 20$ l</td>
</tr>
<tr>
<td>$n_{\text{daily}}$</td>
<td>Number of draws per day</td>
<td>Uniform</td>
<td>$n_{\text{min}}^{\text{daily}} = 30$, $n_{\text{max}}^{\text{daily}} = 50$</td>
</tr>
<tr>
<td>$T_{\text{long}}^{\text{draw}}$</td>
<td>Long water draw duration</td>
<td>Normal</td>
<td>$\mu = 10$ min, $\sigma = 1$ min</td>
</tr>
<tr>
<td>$T_{\text{short}}^{\text{draw}}$</td>
<td>Short water draw duration</td>
<td>Normal</td>
<td>$\mu = 1$ min, $\sigma = 6$ s</td>
</tr>
<tr>
<td>$\dot{m}_{\text{draw}}$</td>
<td>Hot water flow rate</td>
<td>Normal</td>
<td>$\mu = 1$ l/min, $\sigma = 0.2$ l/min</td>
</tr>
</tbody>
</table>

Figure 3.7: Exemplary water draw scenario

3.3.6 Simulation of a Single Water Heater

We will now simulate the dynamic behavior of an individual EWH subject to water draws and periodic reheating triggered by the thermostat over the course of one day. We consider two scenarios: 1) unimpeded operation of the heating element during the entire day, 2) heating element blocked from 08:00 am to 08:00 pm for taking advantage of a low electricity tariff assumed to be active from 08:00 pm to 08:00 am.

The EWH and simulation parameters are summarized in Table 3.3. We utilize a water draw scenario created using the parameters shown in Table 3.2. The thermostat dead-band is chosen as $[50^\circ \text{C}, 60^\circ \text{C}]$.

Figure 3.8 shows the simulation of the EWH in unimpeded operation over 24 hours. Figure 3.8-a) depicts the water draws over the simulation time span, color-coded by the temperature of the drawn water. Figure 3.8-b) depicts the temperature along the tank height using the same color coding. Figure 3.8-c) shows the temperature at the heating element relative to the dead-band as well as the $ON/OFF$ state of the heating element. We see that the initial heating of the water
Table 3.3: Parameters for water heater simulation based on physical properties of water and assumptions on tank properties

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1.57</td>
<td>[m²]</td>
</tr>
<tr>
<td>Aₜₐₜₚ</td>
<td>0.3</td>
<td>[m²]</td>
</tr>
<tr>
<td>c</td>
<td>4,185.5</td>
<td>[kJ/K]</td>
</tr>
<tr>
<td>h</td>
<td>1</td>
<td>[m]</td>
</tr>
<tr>
<td>m</td>
<td>300</td>
<td>[kg]</td>
</tr>
<tr>
<td>vᶜᶠ</td>
<td>0.05</td>
<td>[m/s]</td>
</tr>
<tr>
<td>Nₘⁱₓ</td>
<td>4</td>
<td>[-]</td>
</tr>
<tr>
<td>Tₘᵦᵐ</td>
<td>20</td>
<td>[°C]</td>
</tr>
<tr>
<td>Tₑₜ₁ₙ</td>
<td>50</td>
<td>[°C]</td>
</tr>
<tr>
<td>tₛᵦₘᵉ</td>
<td>24</td>
<td>[h]</td>
</tr>
<tr>
<td>η</td>
<td>1</td>
<td>[-]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aₘᵦᵨₜᵫ</td>
<td>0.3</td>
<td>[m²]</td>
</tr>
<tr>
<td>Aₜₐᵦₛₙ,ᵢ</td>
<td>0.0097</td>
<td>[m²]</td>
</tr>
<tr>
<td>d</td>
<td>0.01</td>
<td>[m]</td>
</tr>
<tr>
<td>k</td>
<td>0.6</td>
<td>[W/mK]</td>
</tr>
<tr>
<td>mᵢ</td>
<td>3</td>
<td>[kg]</td>
</tr>
<tr>
<td>Nₜиск</td>
<td>100</td>
<td>[-]</td>
</tr>
<tr>
<td>Pₑᵦᵣₙₐₑ</td>
<td>4</td>
<td>[kW]</td>
</tr>
<tr>
<td>Tₑᵦₓₗₐₓ</td>
<td>60</td>
<td>[°C]</td>
</tr>
<tr>
<td>tₛᵦₐₘᵉ</td>
<td>10</td>
<td>[s]</td>
</tr>
<tr>
<td>U</td>
<td>0.5</td>
<td>[W/m²K]</td>
</tr>
<tr>
<td>ρ</td>
<td>1,000</td>
<td>[kg/m³]</td>
</tr>
</tbody>
</table>

(until $t = 4$ h) is relatively uniform. Then the heating element is switched OFF and the water draws gradually move the thermocline between cold and hot water upwards. At about $t = 11$ h, $t = 19$ h, and $t = 22$ h, the cold water reaches the thermostat and triggers the reheating of the tank.

Figure 3.9 shows the same simulation scenario with blocked heating element during the day. One can observe that the thermocline between hot and cold water moves upward about halfway before the water is reheated at $t = 20$ h. Note that the mixing effect brought about by the circular mass flow causes the temperatures above the element to equalize when the element is switched on. This implies that care should be taken when a heating cycle is started and then interrupted by external control since the user might not have sufficiently hot water anymore due to the averaging of the water temperatures.

### 3.3.7 Remark on Model Validation

The simulation outcome of the model was compared with a set of measurement data in [120]. The data set used was obtained and kindly made available by the Lawrence Berkeley National Laboratory. The experiment consists of a series of short strong draws in one-hour intervals applied to a 190-liter cylindrical water tank. The temperatures of the water within the tank were measured at different heights, thus giving insight into the stratification within the tank. Although an exact match of the temperature profiles over time could not be achieved, our EWH model yields physically plausible results with an accuracy that is deemed sufficient for simulations of large EWH populations. We omit further considerations for shortness and refer to [120] for details.
3.3. Single Water Heater Model

Figure 3.8: Simulation results of single unit: unimpeded operation

Figure 3.9: Simulation results of single unit: blocked during day
Chapter 3. Modeling and Coordination of EWHs

3.4 Water Heater Population Model

In this section, we will develop a representation of a population of EWH models by assembling statistical distributions for the governing parameters. The main purpose of this model is the investigation of the aggregate power consumption of a large unit population, both in autonomous operation and under the control of a central coordination entity. At the same time, the detailed representation of the water temperature layers within the tank enables us to evaluate the end-use performance of the units, i.e., the supply of hot water to the user.

3.4.1 Modeling of Population Parameters

We describe in the following a routine that creates a heterogeneous set of parameters for any desired number of EWHs. The devices are subject to variations of their tank volume, the power of their heating elements, and the total volume of water drawn in one day. We consider a certain amount of correlation between these parameters, which means, e.g., that a small EWH is more likely to have a small heating element and a small daily water draw volume.

We create a heterogeneous population of EWHs by the following methodology:

1. First, a set of tank sizes is defined by the water mass contained in the tank, yielding the tank category mass vector $m_{\text{cat}}$. A vector of percentages $p_{\text{cat}}$ is defined which represents the share that each EWH category has in the total population. Afterwards, the individual EWH contents $m_i$ are drawn from the vector $m_{\text{cat}}$ with a probability according to the percentages defined in $p_{\text{cat}}$.

2. A matrix $P_{\text{rated},\text{el,cat}}$ is defined, the columns of which contain possible power ratings of the heating elements according to the EWH category. For a water heater $j$ of a certain category (column), a random element (line) is drawn from the matrix based on a uniform discrete distribution to define the rated power $P_{\text{rated,el,j}}$ of this unit.

3. For each EWH category, an interval $[m_{\text{min,daily}}, m_{\text{max,daily}}]$ for the total water draw per day is defined. The actual daily draw volume $m_{\text{daily,j}}$ is determined by a uniform distribution between these bounds.
3.4. Water Heater Population Model

Table 3.4: Parameters for creating the water heater population

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Meaning</th>
<th>Value</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>( m_{\text{cat}} )</td>
<td>Water mass categories</td>
<td>[150, 200, 300, 400]</td>
<td>[kg]</td>
</tr>
<tr>
<td>( p_{\text{cat}} )</td>
<td>Category shares</td>
<td>[10%, 25%, 25%, 40%]</td>
<td>[-]</td>
</tr>
<tr>
<td>( P_{\text{rated el,cat}} )</td>
<td>Rated power per cat.</td>
<td>[2, 3, 4, 4]</td>
<td>[kW]</td>
</tr>
<tr>
<td>( m_{\text{min daily,cat}}, m_{\text{max daily,cat}} )</td>
<td>Daily draw bounds per cat.</td>
<td>[100, 130, 180, 200]</td>
<td>[kg]</td>
</tr>
<tr>
<td>( T_{\text{min set}}, T_{\text{max set}} )</td>
<td>Temp. setpoint bounds</td>
<td>[55, 65]</td>
<td>[°C]</td>
</tr>
<tr>
<td>( T_{\text{min dead}}, T_{\text{max dead}} )</td>
<td>Dead-band size bounds</td>
<td>[5, 15]</td>
<td>[K]</td>
</tr>
<tr>
<td>( U_{\text{min}}, U_{\text{max}} )</td>
<td>Heat loss coefficient bounds</td>
<td>[0.2, 1]</td>
<td>[\frac{\text{W}}{\text{m}^2 \text{K}}]</td>
</tr>
<tr>
<td>( x_0 )</td>
<td>Initial state vector</td>
<td>[10, 50, \ldots, 50, 20]</td>
<td>[°C]</td>
</tr>
</tbody>
</table>

4. The temperature setpoint \( T_{\text{set, j}} \) of water heater \( j \) and the width of the thermostat’s dead-band \( T_{\text{dead, j}} \) are drawn from uniform distributions in the intervals \([T_{\text{min set}}, T_{\text{max set}}]\) and \([T_{\text{min dead}}, T_{\text{max dead}}]\).

5. The thermal loss coefficient \( U \) is varied (based on a uniform distribution) in the interval \([U_{\text{min}}, U_{\text{max}}]\).

Table 3.4 shows an exemplary parameterization for creating the EWH population based on the described stochastic parameter variations.

3.4.2 Simulation of a Water Heater Population

Now we conduct a simulation of an EWH population consisting of 100 units in open loop (without a centralized control entity). Similar to the individual EWH simulation in Section 3.3.6, we simulate both 1) the unimpeded operation and 2) the “two-stage tariff” operation with blocking the heating elements from 08:00 am to 08:00 pm. We use the parameters given in Table 3.4.

Figure 3.10 shows the simulation of the unimpeded EWH population. Figure 3.10-a) shows the aggregated water draws in liters per minute of the overall population. The color indicates the average temperature of the drawn water, which is about 55°C. Figure 3.10-b) depicts the outlet temperatures of the EWHs. They remain above 50 degrees at all times. Figure 3.10-c) describes the relative thermal energy \( E_{\text{rel}} \) (i.e., the measured temperature relative to the dead-band) at the heating element. Note that this quantity stays between 0 and 1 when the temperature at the element is within its dead-band, and that water draws can move
it below 0). Figure 3.10-d) demonstrates the power consumption behavior of the population, which is relatively similar to the shape of the aggregated water draws.

In Figure 3.11, we demonstrate the effect of the power consumption blocking during the day. Figure 3.11-a) – Figure 3.11-d) show the same quantities as above. The differences in the results are as follows: In Figure 3.11-a), one can observe that the average outlet temperature is similar to the unimpeded operation, except for a strong drop just after 08:00 pm when the heating elements are released. This is due to the mixing effect of the heating elements, which eradicates the stratification and averages the layer temperatures. Due to the relatively low average temperature of the tanks induced by the water draws during the day, the outlet temperatures decrease just after the reheating starts. This is also observable in Figure 3.11-b). Figure 3.11-c) illustrates that the EWHs would actually consume power if they were not blocked, indicated by $E_{\text{rel}}^\text{th} \ll 0$ without the heating element being active. In Figure 3.11-d), one can observe the consumption being equal to zero between 08:00 am and 08:00 pm, and the subsequent large load peak starting at 08:00 pm. This effect is referred to as CLPU [95, 96].
Figure 3.10: Simulation results of population: unimpeded operation
Figure 3.11: Simulation results of population: blocked during day
3.5 Control Strategies

In this section, we present and evaluate two control strategies that coordinate the EWH population so as to track a time-varying setpoint with its aggregate power consumption. The model introduced in Section 3.4 is used as the “plant” to be controlled. This section is based on part of the results obtained in [120], where a variety of controllers is developed and tested.

After a short overview, we describe the two algorithms along with simulation results. Then we compare the tracking and end-use performance of the two controllers.

3.5.1 Overview

We consider the following two control strategies:

1. A control approach acting on the aggregated population by broadcasting percentages of the population to be blocked or released. No state information is obtained from the devices and a single signal is broadcast to all devices. Since the percentage to be switched ON or OFF is denoted by the variable $\gamma$, we refer to this approach as “Gamma Control”.

2. Direct temperature feedback from two temperature sensors located at the lower heating element ($T_{e,1}$) and at the hot water outlet ($T_{\text{out}}$) as well as information about the water heaters’ ON/OFF state ($u$) is known to the controller. It can transmit OFF as well as ON signals to individual devices of the population. This approach is referred to as “Switching Control”.

For a more detailed categorization and evaluation of control approaches, see [120]. Table 3.5 provides an overview of the different degrees of information feedback, possible ways to apply the control signals, and the prioritization of the external control signal versus the internal controller.

3.5.2 Gamma Control Strategy

This control strategy is characterized by a complete lack of information feedback from the individual units to the central controller. The controller has access to the following information:
Table 3.5: Overview of water heater control strategies

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Information Feedback</th>
<th>Control Signals</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gamma Control</td>
<td>Only aggregate power</td>
<td>Percentage of population to be switched ON or OFF</td>
<td>External control signal</td>
</tr>
<tr>
<td>Switching Control</td>
<td>ON/OFF state, outlet temperature, heating element temperature</td>
<td>Switching signals (toggling of ON/OFF state) to individual units</td>
<td>Internal controller</td>
</tr>
</tbody>
</table>

- the desired aggregate load in the current time step $P_{set}$,
- the measured aggregate load in the current time step $P_{agg}$, and
- the total installed power of the population $P_{inst}$.

**Strategy Description**

The control approach is based on the work presented in [121] and has already been investigated in [120]. It utilizes the control signal denoted as $\gamma$, which indicates the fraction of EWHs that shall be blocked or released (unblocked) in each time step. Since the load switching exhibits direct feed-through characteristics (i.e., the aggregate power can be adapted instantaneously by the controller without a dynamic response), we simply determine $\gamma$ in each time step by using the equation

$$ \gamma = \frac{P_{agg} - P_{set}}{P_{inst}}. \quad (3.28) $$

If the current actual power demand is larger than the desired value ($P_{agg} > P_{set}$), then $\gamma$ assumes a positive value in the range of $0 < \gamma \leq 1$, indicating that additional EWHs need to be blocked. If, on the other hand, the actual power demand is lower than the desired value ($P_{agg} < P_{set}$), $\gamma$ will be in the range $-1 \leq \gamma < 0$, indicating that EWHs that are blocked should be released.

An individual device is blocked or released based on a local decision by drawing a random number $n$ from a uniform distribution within the bounds $0 \leq n \leq 1$. A binary block signal $b$ is assigned locally to water heater $j$ according to the rule:

$$ b_j(k) = \begin{cases} 
1 & \text{if } b_j(k-1) = 0 \text{ and } \gamma > 0 \text{ and } n_j \leq \gamma \\
0 & \text{if } b_j(k-1) = 1 \text{ and } \gamma < 0 \text{ and } -n_j \geq \gamma \\
b_j(k-1) & \text{otherwise}
\end{cases} \quad (3.29) $$
where $j = 1, \ldots, N$ denotes the individual appliance, and $k$ and $k-1$ refer to the current and previous time-step. If the measured aggregate power consumption is higher than the setpoint (i.e., $\gamma > 0$), a device that draws a random number below $\gamma$ will assign a block signal to itself. In the opposite case, for $\gamma < 0$, a device that draws a random number below the magnitude of $\gamma$ will unblock itself. The usage of the uniform distribution ensures that the percentage of EWHs that change their blocking signal $b$ in a particular time-step is close to the percentage of EWHs that need to change their state. The decentralized draw of random numbers serves to facilitate the communication between the central coordination entity and the individual units. The block signal establishes the relation between $v_i$ and $u_i$ in accordance with Section 3.3.4.

**Simulation Results**

We simulate an EWH population of 1,000 units coordinated by the Gamma Control strategy. Figure 3.12 shows a time-domain simulation over a time span of one day. The parameterization of the population is the same as the one used in the open-loop simulations of Section 3.4. The following plots are shown:

Figure 3.12-a) depicts the total water draw of the population, color-coded by the mean temperature of the total amount of water drawn across the entire population.

In Figure 3.12-b), we show the outlet temperatures of 100 arbitrarily chosen units of the EWH population. Although most of the outlet temperatures remain above 50°C, some of them clearly attain unacceptable values around 20°C. This constitutes a comfort loss for some customers.

Figure 3.12-c) shows the relative thermal energy measured at the heating element, i.e., the temperature relative to its dead-band. It can be observed that some temperatures divert significantly from the dead-band interval [0,1] without a reaction from the internal controller. This is due to the fact that Gamma Control overrides the local thermostat.

The aggregate power consumption and its setpoint are shown in Figure 3.12-d). The setpoint is followed with a high accuracy.

The achieved results suggest that the Gamma Control strategy is suitable for setpoint tracking but has significant drawbacks in terms of user comfort. However, if a relatively benign control signal is imposed on the population (which regulates the population only slightly up and
Figure 3.12: EWH population controlled by Gamma Control

down compared to its natural power consumption), it will be easier to maintain customer comfort. This implies that the usable slack of power consumption will be smaller in this case than in what is presented here.

### 3.5.3 Switching Control Strategy

The Switching Control strategy is based on the work presented in [120], which in turn was built upon the ideas on rule-based TCL coordination described in Chapter 2. We utilize a toggling mechanism for the ON/OFF state of selected individual units in order to achieve the setpoint tracking. Since the water draws dominate the power consumption behavior, which makes it virtually impossible to predict the next
3.5. Control Strategies

Table 3.6: Water heater sub-populations

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{N}$</td>
<td>Set of all water heaters</td>
</tr>
<tr>
<td>$\mathcal{N}_{\text{on}}$</td>
<td>Set of all water heaters that are currently ON</td>
</tr>
<tr>
<td>$\mathcal{N}_{\text{off}}$</td>
<td>Set of all water heaters that are currently OFF</td>
</tr>
<tr>
<td>$\mathcal{N}_{\text{too hot}}$</td>
<td>Set of all water heaters for which $T_{e,1} \geq T_{e,1}^{\text{max}}$</td>
</tr>
<tr>
<td>$\mathcal{N}_{\text{too cold}}$</td>
<td>Set of all water heaters for which $T_{e,1} \leq T_{e,1}^{\text{min}}$</td>
</tr>
<tr>
<td>$\mathcal{N}_{\text{in dead-band}}$</td>
<td>$\mathcal{N} - (\mathcal{N}<em>{\text{too hot}} \cup \mathcal{N}</em>{\text{too cold}})$</td>
</tr>
</tbody>
</table>

autonomous switching instant (as in the algorithm in Chapter 2), we simplify the creation of the priority lists here.

Strategy Description

The setpoint tracking is achieved by selectively switching individual EWHs both ON and OFF based on information that is fed back to the controller. The central controller divides the EWHs into a number of different sub-populations based on the measured temperature at the heating element $T_{e,1}$, their desired temperature dead-band limits ($T_{e,1}^{\text{min}}$ and $T_{e,1}^{\text{max}}$), and their ON/OFF state (Table 3.6).

The algorithm is designed to respect internal OFF signals which occur when $T_{e,1} \geq T_{e,1}^{\text{max}}$, as well as ON signals when $T_{e,1} \leq T_{e,1}^{\text{min}}$. The following additional sub-populations are therefore identified:

1. $\mathcal{N}_{\text{internal on}} = \mathcal{N}_{\text{off}} \cap \mathcal{N}_{\text{too cold}}$ \hspace{1cm} (3.30)
2. $\mathcal{N}_{\text{internal off}} = \mathcal{N}_{\text{on}} \cap \mathcal{N}_{\text{too hot}}$ \hspace{1cm} (3.31)

These sub-populations (the composition of which can change in every time step) represent EWHs that are currently outside of their temperature dead-band, while their heating element has not reacted to this condition yet. This determines the set of EWHs that will switch autonomously in the upcoming time step. Using this information, the next step is to decide which EWHs should be switched. The following description outlines this process:

1. The difference between the desired aggregate power and the actual aggregate power of the population is computed:
   \[ \Delta P = P_{\text{set}} - P_{\text{agg}} \] \hspace{1cm} (3.32)

2. $\Delta P_{\text{effective}}$ is calculated according to
   \[ \Delta P_{\text{effective}} = \Delta P + \Delta P_{\text{internal off}} - \Delta P_{\text{internal on}} \] \hspace{1cm} (3.33)
If $\Delta P_{\text{effective}} < 0$ holds, the actual aggregate power, after considering internal switching actions, will be above $P_{\text{set}}$, and thus, switching OFF action is desired. Conversely, if $\Delta P_{\text{effective}} > 0$ holds, switching ON action is desired.

3. Candidates for OFF and ON switching actions are determined according to

$$N_{\text{off cand}} = N_{\text{on}} \cap N_{\text{in dead-band}} ,$$

$$N_{\text{on cand}} = N_{\text{off}} \cap N_{\text{in dead-band}} .$$

4. The final lists of EWHs to be switched ON or OFF ($N_{\text{active on}}$ and $N_{\text{active off}}$) are calculated as follows:

- For $\Delta P_{\text{effective}} < 0$: The EWHs in $N_{\text{off cand}}$ are ranked according to $(T_{e,1} - T_{e,1}^{\text{min}})$. A large difference means the EWH is relatively hot and therefore it has a high blocking priority. The final list of EWHs that will actually receive an OFF signal is determined by finding $n$ such that

$$\left| \sum_{j=0}^{n} P_{\text{rated}_{el,1,j}} + \Delta P_{\text{effective}} \right| \to \text{min}$$

is minimized (where $P_{\text{rated}_{el,1,j}}$ is the power of the lower heating element of the $j$th heater in the ranked list of block candidates). The result is a list of EWHs that will receive an OFF signal, $N_{\text{active off}}$.

- For $\Delta P_{\text{effective}} > 0$: The EWHs in $N_{\text{on cand}}$ are ranked according to $(T_{e,1}^{\text{max}} - T_{e,1})$. A large difference means the EWH is relatively cold and therefore it has a high switching priority. The final list of EWHs that will actually receive a release signal is determined by finding $n$ such that

$$\left| \sum_{j=0}^{n} P_{\text{rated}_{el,1,j}} - \Delta P_{\text{effective}} \right| \to \text{min}$$

is minimized (where $P_{\text{rated}_{el,1,j}}$ is the power of the lower heating element of the $j$th heater in the ranked list of release candidates). The result is a list of EWHs that will receive an ON signal, $N_{\text{active on}}$. 
3.5. Control Strategies

The controller then cycles through all devices \((j = 1, \ldots, |N|)\) and assigns a switch signal \(s\) according to:

\[
s_j(k) = \begin{cases} 
1 & \text{if } j \in N_{\text{active on}} \lor j \in N_{\text{internal on}} \\
0 & \text{if } j \in N_{\text{active off}} \lor j \in N_{\text{internal off}} \\
s_{j}(k-1) & \text{otherwise}
\end{cases}
\]  

(3.36)

Finally, the actual \(ON/OFF\) state of each EWH is determined by:

\[
u_j = s_j
\]  

(3.37)

Simulation Results

We simulate the Switching Control strategy in the same way as the Gamma Control strategy with a sampling time of 2 min. Figure 3.13 shows a time-domain simulation over a time span of one day. The parameterization of the population is the same as before. The following plots are shown:

Figure 3.13-a) shows the aggregate water draw profile, color-coded by the average outlet temperature. It can be seen that the average water temperature rises towards the end of the day when the power setpoint is larger.

Figure 3.13-b) presents the temperatures at the heating elements of 100 randomly selected units. All temperatures stay in a quite narrow corridor with a width of about 15 K. The lowest temperature attained by the plotted population subset is about 45°C.

Figure 3.13-c) shows the relative thermal energy measured at the heating element, i.e., the temperature relative to its dead-band. The downwards spikes are caused by strong water draws that suddenly subject the thermostat to cold water.

The aggregate power consumption of the population and its setpoint are shown in Figure 3.13-d). The aggregate power slope during the first hour is caused by the low initial temperature (selected to 50°C for illustration purposes) which leads to a CLPU effect [95, 96]. Since the controller is unable to override the local thermostat when the unit is outside of the dead-band, this effect is unavoidable for this initial condition. Later on, the signal is followed with a high accuracy.
3.5.4 Comparison of the Simulation Results

For a direct comparison of the two approaches, we present the aggregate power time series (as an indication of control quality) and outlet water temperatures (indicative of user comfort) together.

Figure 3.14-a) shows a comparison of the setpoint tracking performance and Figure 3.14-b) shows the control error as a percentage of the current setpoint value. In Figures 3.14-c) and 3.14-d), we depict a comparison of the EWH outlet temperatures.
It is evident that the Gamma Control approach introduced in Section 3.5.2 is able to follow the setpoint with an accuracy similar to the Switching Control approach from Section 3.5.3. Differences exist in the handling of EWHs outside the dead-band: while the Gamma Controller is able to override the local thermostat controller (which leads to the depicted comfort loss), the Switching Controller honors state constraints.

We conclude that Gamma Control is more suitable for smaller control actions where the deviation from autonomous operation are relatively small. One such application would be the tracking of a Load Frequency Control (LFC) signal around a baseline trajectory that is relatively close to the natural (uninfluenced) power consumption behavior of the EWH population.
The Switching Control approach is able to maintain user comfort by allowing individual EWH addressing while at the same time increasing the control accuracy when the population remains within the dead-band. The drawback of this is the substantial effort for communication and metering equipment.

3.6 Concluding Remarks

In this chapter, we presented a dynamic model of an EWH based on physical energy and mass balances within the water tank. The model dynamics can be shown to match experimental data relatively closely. A statistical water draw model was used to simulate the user-induced draws of hot water from an EWH and a population model was generated using stochastic parameter variations. We also showed by the application of rule-based control strategies for EWH populations that accurate setpoint tracking can be achieved. This profoundly extends the insights gained in Chapter 2 since we used a more detailed model of the internal dynamics and a model for user interactions instead of a simple one-state TCL model.

Future work may consist of a more detailed investigation of the internal water dynamics, particularly with respect to the dynamic behavior of the water temperatures when a heating element is activated. Furthermore, novel control strategies can be developed, e.g., with the aim of reducing the communication effort needed for accurate setpoint tracking and maintaining the end-use performance of the EWHs at the same time. When considering EWHs with a second heating element, which provides a quick re-heating possibility for the upper water layers, further opportunities to enhance the available power consumption “slack” without jeopardizing end-use functionality should be investigated.
Chapter 4

Probabilistic Modeling and Control of TCLs

In this chapter, we present a probabilistic modeling and control approach for Thermostatically Controlled Loads (TCLs). Based on a Markov-chain modeling technique, we derive a closed-form Linear Time-Invariant (LTI) model of a TCL population that can be used for simulation studies and control design. A Model Predictive Control (MPC) approach is used to track a time-varying setpoint in the aggregate power consumption. The control action is applied to the “plant” of individual TCLs represented by a first-order model for each unit. Numerical simulations are presented and tracking performance in different scenarios is assessed.

4.1 Introduction and Literature Review

A general overview about Thermostatically Controlled Load (TCL) modeling and control literature has already been presented in Chapter 2. Here we focus on research that is based on modeling the states of individual appliances within the temperature dead-band in an aggregate and probabilistic way. The common ground of these approaches is the representation of the set of TCL states by a continuous “probability mass”, the dynamics of which can be described in time and space by a Partial Differential Equation (PDE). The foundation for this work was laid by [97] and extended by various researchers, e.g., [98]. These works were published before a substantial share of intermittent renewable energy in power systems could be anticipated and thus were focused on the reduction of the fairly well-predictable peak load.
As outlined before, renewed interest in the management of aggregated TCL populations was triggered by the challenges associated with renewable energy expansion. The new idea of balancing and ancillary service provision by controllable loads established novel requirements on control methodologies. While statistical modeling approaches such as [97] were originally not designed to track rapidly changing setpoint trajectories, they possess the advantage of representing an entire TCL population in a closed form. In comparison with rule-based control approaches as presented in Chapters 2 and 3, this is a promising way to reduce the need for information exchange between the units and a central controller. The present chapter extends a line of research that is motivated by providing accurate setpoint tracking in aggregate power consumption while minimizing communication requirements. Known previous approaches include [75], in which aggregate TCL power consumption is controlled by changing the thermostat setpoints and insight into the system dynamics is gained by solving the PDE model from [97]. Reference [77] extends this work towards a discrete “state bin” modeling framework similar to [102], and the authors develop a bilinear controller to manipulate aggregate power consumption, again by temperature setpoint control. Since the authors of [97] had to assume parameter homogeneity for the model development, the approaches which build directly on their PDE model do so as well.

The main objective of our approach is to derive a framework that can incorporate various control design techniques and heterogeneous TCL parameters, and that also allows the use of state and parameter estimation. We develop a Linear Time-Invariant (LTI) model of a TCL population based on a discretization of the temperature space into “state bins” and a consideration of the aggregated probability mass, similar to [77, 102]. In a manner similar to [77], we are interested in understanding whether system identification and observer design methods can be used. We formulate the control problem in terms of an Model Predictive Control (MPC) framework that is highly flexible with respect to managing control objectives and constraints. Finally, anticipating that it will be difficult to measure in real-time the state of every TCL being controlled, we examine the influence that incomplete TCL state information has on controller performance.

The contents of this chapter are largely based on the paper [123]. Together with [124], this provided the basis for [125]. The chapter proceeds as follows: In Section 4.2 we present the aggregate probabilistic model
4.2. Modeling Approach

Table 4.1: Notation for Chapter 4

<table>
<thead>
<tr>
<th>Var. Unit</th>
<th>Meaning</th>
<th>Var. Unit</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>A [-]</td>
<td>Dynamic matrix</td>
<td>Q [-]</td>
<td>State penalty matrix</td>
</tr>
<tr>
<td>a [-]</td>
<td>TCL parameter</td>
<td>q [-]</td>
<td>Scalar penalty factor</td>
</tr>
<tr>
<td>B [-]</td>
<td>Control input matrix</td>
<td>R [-]</td>
<td>Input penalty matrix</td>
</tr>
<tr>
<td>C [-]</td>
<td>Output matrix</td>
<td>R [°C/kW]</td>
<td>Thermal resistance</td>
</tr>
<tr>
<td>C [kWh]</td>
<td>Thermal capacitance</td>
<td>u [-]</td>
<td>Control input vector</td>
</tr>
<tr>
<td>h [h]</td>
<td>Time step size</td>
<td>v [-]</td>
<td>Measurement noise vector</td>
</tr>
<tr>
<td>I [-]</td>
<td>Identity matrix</td>
<td>x [-]</td>
<td>State bin vector</td>
</tr>
<tr>
<td>J [-]</td>
<td>Cost function</td>
<td>y [-]</td>
<td>Output vector</td>
</tr>
<tr>
<td>m [-]</td>
<td>ON/OFF variable</td>
<td>δ [°C]</td>
<td>Temp. dead-band width</td>
</tr>
<tr>
<td>N [-]</td>
<td>Number</td>
<td>θ [°C]</td>
<td>Temperature</td>
</tr>
<tr>
<td>P [kW]</td>
<td>Electric power</td>
<td>ω [°C]</td>
<td>Noise process</td>
</tr>
<tr>
<td>p [-]</td>
<td>Probability</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subscript Meaning</th>
<th>Subscript Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>a Ambient</td>
<td>on ON switching state</td>
</tr>
<tr>
<td>bin State bin</td>
<td>P Power</td>
</tr>
<tr>
<td>g Gain</td>
<td>pred Prediction steps</td>
</tr>
<tr>
<td>i TCL index</td>
<td>pref Preferential</td>
</tr>
<tr>
<td>j Bin index</td>
<td>rated TCL rated power</td>
</tr>
<tr>
<td>k Time step index</td>
<td>rel Relative</td>
</tr>
<tr>
<td>j Prediction step index</td>
<td>set Setpoint</td>
</tr>
<tr>
<td>max Maximum</td>
<td>total Whole population</td>
</tr>
<tr>
<td>min Minimum</td>
<td>track Tracking</td>
</tr>
<tr>
<td>off OFF switching state</td>
<td></td>
</tr>
</tbody>
</table>

Based on Markov chains. In Section 4.3, we discuss the possibilities for transferring information between the central controller and the TCLs. The control approach is described in Section 4.4. Section 4.5 presents simulation results and Section 4.6 concludes this chapter. The used notation is summarized in Table 4.1.

4.2 Modeling Approach

4.2.1 Individual TCL Model

The individual TCL model is the same as the one used in [75]. The parameter $a$, which governs the thermal characteristics of each TCL $i$, is defined as:

$$a_i = e^{-h/(C_i R_i)}$$  \hspace{1cm} (4.1)

where $C_i$ [kWh] and $R_i$ [°C/kW] are the thermal capacitance and resistance of TCL $i$. Furthermore, $h$ [h] is the time step size.

The difference equation describing the temperature dynamics of TCL $i$ is:

$$\theta_{i,k+1} = a_i \theta_{i,k} + (1 - a_i)(\theta_{a,i} - m_{i,k} \theta_{g,i}) + \omega_{i,k}$$  \hspace{1cm} (4.2)
where \( k \) is the time step index, \( \theta_{i,k} \) is the internal temperature of the TCL, \( \theta_{a,i} \) is the ambient temperature, \( m_{i,k} \) is a discrete variable equal to 1 when the TCL is ON and 0 when it is OFF, and \( \omega_{i,k} \) is a noise process. The \( ON \) temperature gain \( \theta_{g,i} \) is:

\[
\theta_{g,i} = \begin{cases} 
R_i P_{\text{rated},i} & \text{for cooling devices} \\
-R_i P_{\text{rated},i} & \text{for heating devices}
\end{cases},
\]

(4.3)

where \( P_{\text{rated},i} \) is the TCL’s rated power. \( \theta_{g,i} \) is positive for cooling TCLs and negative for heating TCLs.

### 4.2.2 TCL Population Model

To simulate the behavior of a population of TCLs, we could aggregate thousands of single TCL models using (4.2); however, this would be computationally intensive and the aggregate system would not be in a form amenable to many control techniques. Instead, in this approach we will work with a discrete LTI system in state space form:

\[
x_{k+1} = Ax_k + Bu_k,
\]

(4.4)

\[
y_k = Cx_k
\]

(4.5)

which allows us to use a wide range of system analysis tools and advanced control techniques.

Assume all TCLs in a population have the same temperature setpoint \( \theta_{\text{set}} \) and temperature dead-band width \( \delta \) (or normalized diverse dead-bands). Divide the dead-band into \( N_{\text{bin}}/2 \) temperature intervals. A TCL in a certain temperature interval can be either \( ON \) or \( OFF \). Divide each temperature interval into two state bins, one for TCLs that are \( ON \) and one for TCLs that are \( OFF \). This results in \( N_{\text{bin}} \) state bins. The state vector \( x \) contains the number of TCLs in each state bin, or, if normalized by the total number of TCLs, the fraction of TCLs in each state bin, which is equivalent to probability mass in the infinite system limit. In the remainder of the chapter, we will refer to \( x \) as a vector of probability mass. The \( A \)-matrix can be thought of as a (transposed) Markov transition matrix describing the probability of TCLs moving from one state bin to the next. Figure 4.1 shows how the state bins map to the temperature dead-band. In [123], the analytical derivation of the \( A \)-matrix is presented, which we compare with an identified \( A \)-matrix. We will discuss the structure of the matrices \( B \) and \( C \) and the vectors \( u \) and \( y \) in subsequent sections.
4.2. Modeling Approach

Figure 4.1: State bin transition model [123]

Table 4.2: Simulation and TCL parameters, adapted from [75]

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Meaning</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_{set}$</td>
<td>Temperature setpoint</td>
<td>20°C</td>
</tr>
<tr>
<td>$\delta$</td>
<td>Temperature dead-band width</td>
<td>0.5 K</td>
</tr>
<tr>
<td>$\theta_a$</td>
<td>Ambient temperature</td>
<td>32°C</td>
</tr>
<tr>
<td>$R$</td>
<td>Thermal resistance</td>
<td>$2 \text{ K kW}^{-1}$</td>
</tr>
<tr>
<td>$C$</td>
<td>Thermal capacitance</td>
<td>$8 - 12 \text{ kWh K}^{-1}$</td>
</tr>
<tr>
<td>$P_{rated}$</td>
<td>Rated power</td>
<td>14 kW</td>
</tr>
<tr>
<td>$h$</td>
<td>Time step</td>
<td>10 s</td>
</tr>
</tbody>
</table>

Comparison of Analytically-Derived $A$-matrix to Identified Matrix

We compare the analytically-derived $A$-matrix [123] with an identified $A$-matrix obtained by a Markov chain identification procedure. This technique can only be applied if full state information for the TCL population (or a subset thereof) is available to the identification algorithm.

For this comparison, we consider the case where the TCL parameter, $a$, is uniformly distributed in $[a_{\min}, a_{\max}]$:

$$p(a) = \begin{cases} 
1/(a_{\max} - a_{\min}) & \text{if } a_{\min} \leq a \leq a_{\max} \\
0 & \text{otherwise}
\end{cases} \quad (4.6)$$

The simulation and TCL parameters used in our analysis are listed in Table 4.2, and $a_{\min}$ and $a_{\max}$ are computed from these parameters.

Analytically-derived matrix: As described in [123], the analytically-derived $A$-matrix is calculated as follows: For each combination of possible starting and ending bins, we evaluate numerically the analytical expression for transition probability to generate the $A$-matrix. ON/OFF switching is represented in the model by the addition of state bins outside of the dead-band. In each time step, probability mass is
moved from starting bins to ending bins, without regard to the dead-band. After each time step, the probability mass ending up in a bin outside of the dead-band is switched \textit{ON} or \textit{OFF} by the TCL’s internal controller.

\textit{Markov-identified matrix:} We simulate 1,000 individual TCLs and use the MATLAB [126] function \texttt{hmmestimate} to identify the Markov Transition Matrix. We could also construct the matrix by counting the number of TCL transitions within the discretized dead-band over many time steps. Note that the Markov Transition Matrix has to be transposed in order to be equivalent to the $A$-matrix from (4.4).

\textit{Model comparison:} The dynamic behavior of the analytically-derived and Markov-identified models were compared with a simulated population of 1,000 individual TCLs. The top two plots in Figure 4.2 show the evolution of the unforced systems ($N_{\text{bin}} = 40$ for the analytical and identified $A$-matrices). The bottom three plots compare the eigenvalues of the analytically-derived and Markov-identified $A$-matrices, given...
different numbers of state bins, here simulated with a time step of 10 s. For the time-domain simulations, all TCLs are started in a single state bin, which leads to an oscillatory decay in aggregate power. The models produce similar results for the first 2 – 3 oscillations. After that, the disagreement between the three models increases. This implies that control actions that require long forecasts may not perform well, but that control actions with reasonably short forecast requirements (e.g., less than one hour) may not suffer from plant-model mismatch problems. If control actions are applied continuously to the system (“tight” control), this mismatch problem will likely be negligible.

The oscillatory decay is related to both TCL parameter heterogeneity and the number of state bins used in the state bin transition model. In the individual TCL model, homogeneous populations exhibit undamped oscillations, and these oscillations decay with increasing parameter heterogeneity as TCLs spread out over the temperature state space. In the bin transition model, there is a negative correlation between the number of bins and the damping of the oscillation: with increasing amount of bins, the oscillation becomes increasingly persistent. This finding is consistent with [77], where a large number of bins (200) was chosen in order to yield a non-decaying oscillation corresponding to a homogeneous system. Interestingly, the bin transition model shows a non-decaying behavior when large numbers of bins are used, even if the model has been constructed from a heterogeneous parameter set. Thus, for heterogeneous populations, the number of bins can be seen as a tuning variable which can be adapted to achieve a good matching of the decay behavior. However, as will be seen in Section 4.5, the number of bins does not have a large influence on a tightly controlled system.

The mismatch between the analytically-derived model and the simulated population of TCLs is due, at least in part, to the assumption made that the parameter distribution of TCLs is the same in each state bin (see [123]). For the simulated population of TCLs, the parameter distribution of TCLs in each state is different and depends on initial condition and simulation time.

The mismatch between the analytically-derived and the Markov-identified $A$-matrices can likely be attributed to the way $ON$/$OFF$ transitions are handled. In the analytical system we have added extra bins outside of the temperature dead-band. Conversely, the identified matrix has only one bin at each corner of the temperature dead-band, which takes
all probability mass that crosses the dead-band during each time step. We conclude that the analytically-derived and Markov-identified models are relatively consistent with the simulated population of TCLs, though in both cases some error results from the matrix construction methodology. We use the transposed Markov matrix in our subsequent simulations.

4.3 Information Transfer

4.3.1 General Considerations

In this section, we will not identify specific communications platforms or protocols. Instead, we will speak more abstractly about the types of information that might be exchanged. This work is based on a “direct load control” philosophy, meaning that a central controller collects information from the TCL population and transmits a control signal to which the TCLs react. Given the state bin transition model, there are many degrees of freedom for control and information transfer between the TCL population and the central controller including:

A) the nature of the control signal transmitted from the central controller to the TCLs,
B) the methodology for addressing the entire population or only a certain subset, and
C) the information that the TCLs transmit back to the controller.

In Table 4.3, we present a number of options for these three degrees of freedom. Clearly, our choices of options must correspond to capabilities of the actual system. In the following subsections, we explain the subset of options we have implemented. Future work will explore the other options.

4.3.2 Implemented Approach

Controller $\rightarrow$ TCL Population

We have implemented the approach A3/B3 according to Table 4.3. Specifically, we toggle ON/OFF TCLs in certain bins, which corresponds to moving TCLs from ON bins to OFF bins of equal temperature, or vice versa. The control input vector $u_{\text{bin}}$ only has half the
### 4.3. Information Transfer

Table 4.3: Options for control and information transfer

<table>
<thead>
<tr>
<th>Options</th>
<th>Feasible with State Bin Transition Model?</th>
<th>Implemented in this work?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Control signal from central controller to TCLs</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1. Shifting, expanding, or contracting the temp. dead-band [75, 77]</td>
<td>Yes, through definition of extra bins</td>
<td>No</td>
</tr>
<tr>
<td>A2. Blocking (switching OFF) TCLs at certain times</td>
<td>Yes, through definition of extra bins</td>
<td>No</td>
</tr>
<tr>
<td>A3. Toggling ON/OFF certain TCLs, as in [105]</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>B. Targeted TCLs</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B1. Selective addressing of individual TCLs</td>
<td>No, unless the current bin of each TCL is known</td>
<td>No</td>
</tr>
<tr>
<td>B2. Broadcast same control to all TCLs with uniform response</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>B3. Broadcast same control to all TCLs with differing responses</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>C. Information from TCLs to central controller</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1. Full state information of the entire population</td>
<td>Yes, through system ID</td>
<td>Yes</td>
</tr>
<tr>
<td>C2. Full state information of a subset of the population</td>
<td>Yes, through system ID and could include an observer</td>
<td>Yes</td>
</tr>
<tr>
<td>C3. No state information, only aggregate power measured</td>
<td>Yes, but may be difficult</td>
<td>No</td>
</tr>
<tr>
<td>C4. No state information, but TCL properties are known</td>
<td>Yes, through analytical derivation</td>
<td>Yes</td>
</tr>
</tbody>
</table>

number of independent elements as the bin state vector since for every bin that loses probability mass there will be another equal-temperature state bin that gains. Considering the sequence of bins in Figure 4.1, the control input $u_{bin}$ enters the system via the matrix $B$:

$$B = \begin{bmatrix}
-1 & 0 \\
0 & -1 \\
0 & 1 \\
\vdots \\
0 & \vdots \\
1 & 0
\end{bmatrix}.$$  \hspace{1cm} (4.7)

From the structure of $B$ it can be seen that no probability mass (or TCLs) can vanish from the state space through control actions. Based on our sign convention, positive elements of $u_{bin}$ correspond to turning TCLs ON and negative elements correspond to turning TCLs OFF. This property can be exploited by constraining $u_{bin}$ to only non-negative or non-positive values, if the system is designed such that the control can only work in one direction.
We target TCLs according to B3 as follows: If the controller chooses to switch a certain number of loads in a bin, it converts this number to a switch probability using available state information. Then, loads (using a local random number generator) switch with this probability, therefore the same control signal causes differing responses within the TCL population. This avoids the need to directly address individual TCLs, thus reducing the communication effort.

**TCL Population → Controller**

Knowing the measured temperatures of the entire population is the most convenient condition for control (option C1 in Table 4.3) since this means that the $A$-matrix can be identified using Markov Chain techniques, as outlined in Section 4.2.2. However, since this option is the most expensive to implement, we would like to reduce the need for measured state information. It is not beneficial to measure a subset of the states. Individual TCLs traverse all states, so measuring a subset of states would imply that all TCLs would have to be instrumented, which is equivalent to having full state information.\(^1\)

We propose instead to instrument a subset of TCLs (Option C2). At intervals, each TCL will transmit its internal temperature, with respect to its dead-band, to the controller. By aggregating these results, we obtain a noisy measurement of the state bin vector, which can be used directly or together with an observer (e.g., Kalman filter). The $A$-matrix can be identified using Markov Chain techniques on the noisy data.

Integrating measured information into the control system boils down to designing the output equation of the state bin transition model. The measured output $y = Cx + v$, where $v$ is measurement noise, does not include a direct feed-through term $Du$. This is because the controller acts on the TCLs, which is seen in the output one time step later. The matrix $C$ reflects whether only aggregate power is measured, or a measurement of the full state vector is included. In the former case:

$$C = P_{\text{rated}} N_{\text{TCL}} [0, \cdots, 0, 1, \cdots, 1], \quad (4.8)$$

where $N_{\text{TCL}}$ is the total number of TCLs, and $P_{\text{rated}}$ represents the overall installed power of the population. Since the system has only one scalar

\(^1\)A possible exception to this is TCLs reporting their switching actions to the controller, which yields information about the bins at the boundary of the dead-band.
output, the vector $v$ becomes a scalar as well: $v = v_p$, where $v_p$ is the aggregate power measurement noise. If the full state vector is included in the output, the matrix $C$ becomes:

$$C = \begin{bmatrix} I(N_{\text{bin}}) & 0, \cdots, 0, C_p, \cdots, C_p \end{bmatrix},$$

where $I(N_{\text{bin}})$ is the $N_{\text{bin}}$-dimensional identity matrix. The vector $v = [v_1, \ldots, v_{N_{\text{bin}}}, v_p]^T$ is measurement noise resulting from partial instrumentation of the population and other sources, with $v_j$ for $j = 1, \ldots, N_{\text{bin}}$ being the state measurement noise and $v_p$ being the aggregate power measurement noise.

### 4.3.3 Controllability and Observability Properties

The pair $[A, B]$ is not controllable: the controllability matrix is of rank $n - 1$. One degree of freedom is lost because the controller cannot drive all states to zero, since they represent the fraction of TCLs in each bin and must sum to one. However, if $C$ is defined as in (4.8), the system is output-controllable and aggregate power can be tracked. If $C$ is defined as in (4.9), the system is not output-controllable: the output controllability matrix is of rank $n - 2$. One degree of freedom is lost since the controller cannot drive all states to zero and another one is lost because aggregate power is dependent on the states. However, aggregate power can still be tracked. The pair $[A, C]$ is observable for both options of $C$.

### 4.4 Control Approach

#### 4.4.1 Initial Considerations

The state bin transition model is linear since the control input influences an absolute (albeit approximate) number of TCLs in a certain bin in each time step. This is a significant difference to the formulation in [77], in which the controller influences a percentage of TCLs in a bin, resulting in a nonlinear model. In our formulation, the system can theoretically be controlled to have less than zero or more than all of the TCLs in a bin ($x_j < 0$ or $x_j > 1$ for any $j = 1, \ldots, N_{\text{bin}}$), which is physically impossible. Consequently, we choose to use MPC since it can incorporate inequality constraints on states, not just on inputs.
MPC schemes are usually applied to control problems where certain variables must be controlled to (possibly varying) setpoints taking into account the (usually slow) system dynamics, such as in the process industry. In this case, special attention has to be given to the sampling time and prediction horizon in order to represent the dynamics properly within the controller. In our case, switching actions imposed on TCLs occur (in principle) instantly when the control is applied. Due to the state bin transition model formulation, this happens one step after the control is applied, but the whole range of possible outputs can be traversed in this one step. This means that a very short prediction horizon (minimum 2 steps for implementation reasons) is sufficient to capture the relevant input-output dynamics. In cases where the target trajectory is known several steps in advance, a longer prediction horizon may allow to further decrease the control impact on the TCL population.

### 4.4.2 Control Problem Formulation

Figure 4.3 illustrates the control structure consistent with the control and information transfer choices described in the previous section. The MPC outputs the control signal $u_{\text{bin}}$ (in terms of probability mass), which is converted to $u_{\text{bin,rel}}$ (in terms of switch probabilities) by using state information. Note that this preserves linearity in the plant / control input model. To formulate the control problem, a set of con-
4.4. Control Approach

Constraints, defined by the properties of probability mass evolving in the state space, is introduced:

\begin{align*}
0 & \leq x_j \leq 1 \quad \forall \ j = 1, \ldots, N_{\text{bin}} , \quad (4.10) \\
-1 & \leq u_{\text{bin},j} \leq 1 \quad \forall \ j = 1, \ldots, N_{\text{bin}}/2 . \quad (4.11)
\end{align*}

Besides these principal physical constraints, other constraints may be imposed based on the desired behavior of the controller.

A number of suitable cost function designs achieving certain design goals are discussed below. Here, we restrict ourselves to quadratic cost functions, as these usually yield smooth control behavior and are easy to solve by standard quadratic programming techniques. In the choice of penalty factors, there are several options, which include:

- penalization of the tracking error \( P_{\text{total}} - P_{\text{set}} \),
- penalization of the control input vector \( u_{\text{bin}} \), and
- penalization of the state vector deviation from a desired value \( x - x_{\text{set}} \).

In a general form, an appropriate cost function can be defined as follows:

\begin{equation}
J_k = \sum_{l=k}^{k+N_{\text{pred}}-1} (q_{\text{track}} (P_{\text{total}}(l) - P_{\text{set}}(l))^2 
\quad + u_{\text{bin}}^T(l)R_{\text{bin}}u_{\text{bin}}(l) + (x(l) - x_{\text{set}}(l))^TQ(x(l) - x_{\text{set}}(l))) . \quad (4.12)
\end{equation}

The following options are considered for influencing the control behavior in the desired way:

- \( R_{\text{bin}} = 0, Q = 0, q_{\text{track}} \neq 0 \) achieves the minimization of setpoint tracking error without any attention to the control action.
- To balance control actions against the tracking error, \( R_{\text{bin}} \) should be non-zero. In the most simple form, all main diagonal elements are chosen positive and equal, resulting in a uniform evolution of all elements of \( u_{\text{bin}} \). A reduction in switching can be achieved by prioritizing the switching from bins where an autonomous switching would happen soon anyway. This is achieved by altering the problem formulation in the following way: we introduce two new decision variable vectors \( u_{\text{bin},\text{on}} \) and \( u_{\text{bin,off}} \), which are tied to \( u_{\text{bin}} \) by means of an equality constraint:

\begin{equation}
u_{\text{bin,off}} = u_{\text{bin}} \quad . \quad (4.13)\end{equation}
We further establish the constraints $u_{\text{bin}, \text{on}} \geq 0$ and $u_{\text{bin}, \text{off}} \leq 0$ and add another term $J_{\text{pref}}$ into the cost function:

$$J_{\text{pref}} = \sum_{l=k}^{k+N_{\text{pred}}-1} (u_{\text{bin}, \text{off}}^T(l) R_{\text{bin}, \text{off}} u_{\text{bin}, \text{off}}(l)) + (u_{\text{bin}, \text{on}}^T(l) R_{\text{bin}, \text{on}} u_{\text{bin}, \text{on}}(l)) \quad (4.14)$$

The matrices $R_{\text{bin}, \text{on}}$ and $R_{\text{bin}, \text{off}}$ are diagonal matrices with a suitable dimension. $R_{\text{bin}, \text{on}}$ has a monotonically decreasing series of elements on the main diagonal and $R_{\text{bin}, \text{off}}$ has a monotonically increasing series. Depending on the desired penalization, the series can decrease/increase (piecewise) linearly, quadratically, or exponentially. Through this penalization strategy, a preferential switching in the bins close to the autonomous switching is achieved. The main diagonal of $R_{\text{bin}}$ from (4.12) is set to zero or to small equal values.

- Another alternative cost function design takes advantage of the flexibility provided by the bin transition model and results in a more complex controller behavior, although it is easy to formulate. The state vector $x$ can be penalized by the diagonal matrix $Q$ with respect to its deviation from a certain desired probability mass profile (e.g., the steady state profile) among the bins.

### 4.5 Case Study

#### 4.5.1 System Setup and Scenarios

To evaluate the performance of the control system, a simulated population of individual TCLs with process noise set to zero ($w_{i,k} = 0$) is controlled to track a highly variable setpoint. The test trajectory is composed of four different sections: 1) a series of steps resembling energy market or tertiary frequency control dispatch, 2) a sinusoid representing a smooth power adjustment trajectory, 3) some 5-minute power changes which is relevant for short-term ancillary service markets [127], and 4) a real Load Frequency Control (LFC) signal obtained from a European Transmission System Operator (TSO). The choice of these diverse setpoints demonstrates the versatility of the control approach.

The simulated model is parameterized according to the discussion in Section 4.2. Specifically, devices are heterogeneous in thermal capacitance ($C$) but homogeneous with respect to rated power ($P_{\text{rated}}$) and
thermal resistance ($R$). Therefore, the devices described by (4.2) are heterogeneous in the TCL parameter $a$ and homogeneous in the temperature gain $\theta_g$. A real population of TCLs would be heterogeneous in $\theta_g$, which would lead to additional diversity in dynamics during the ON phase. The study of the impact of rated power and thermal resistance heterogeneity on model outcomes can be addressed in future research.

Numerical parameters are shown in Table 4.2. Table 4.4 presents options for parameterizing the controller. Specifically, the columns represent three different choices for penalizing control inputs: “cst. $u_{bin}$” denotes a constant (equal) penalization of the elements in $u_{bin}$, “pref. $u_{bin}$” means that the bins where switching is imminent are preferred by the controller (according to (4.14)), and “state” means that the state is penalized with respect to the steady state of the bin transition model.

We define a set of simulation scenarios that provide insight into the control behavior for different parameterizations of the controller. One scenario will be shown graphically, while the others will only be considered in numerical performance comparisons. Twenty-seven scenarios can be explored by permutating the following parameterization options:

- variation of $N_{bin}$: 40, 60, 80 state bins,
- variation of cost function design:
  - equal penalty on $u_{bin}$ for all elements,
  - splitting of $u_{bin}$ in $u_{bin,on}$ and $u_{bin,off}$; high penalty on first half of $u_{bin,on}$, lower and linearly decreasing penalty for second half of $u_{bin,on}$; the same penalty vector flipped to penalize $u_{bin,off}$, or
  - penalizing the deviation from the steady state.
- variation of state information for controller: 100%, 30%, 10%.
4.5.2 Numerical Results

Numerical simulations were carried out using the scenario definitions above. For the MPC setup, the MATLAB toolbox YALMIP [128] was used. Figure 4.4 shows a simulation of the controlled system with full state vector information available to the controller. The parameters correspond to Case 8 in Table 4.5. The results show that the desired trajectory is tracked with good performance.

The State of Charge (SOC) of the aggregate system is defined as the center of gravity of device states with respect to the temperature dead-band. For cooling TCLs, all TCLs concentrated at the upper dead-band correspond to $SOC = 0$, at the lower dead-band to $SOC = 1$. The opposite relation holds for heating TCLs. Information on the aggregate SOC is relevant to power system operators, allowing them to dispatch the TCL population like an energy storage device.

Table 4.5 shows a numerical performance comparison of 12 of the 27 simulated scenarios. The parameter sets are presented in the upper section of the table, whereas the performance indicators are shown below. It can be seen that the relative Root Mean Square Error (RMSE), normalized by average aggregate power, of the setpoint tracking lies in the range of 0.49 – 2.07%, depending on the simulated case. These results suggest that tracking can be improved with a larger number of bins and/or a larger set of measured TCLs but the effect is relatively small. The preferential switching of TCLs from bins close to the autonomous switching point (pref. $u_{bin}$) yields the best tracking results.

Table 4.5: Performance comparison

<table>
<thead>
<tr>
<th>Case 1</th>
<th>Case 2</th>
<th>Case 3</th>
<th>Case 4</th>
<th>Case 5</th>
<th>Case 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{bin}$</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>60</td>
</tr>
<tr>
<td>Control penalty (Table 4.4)</td>
<td>$cst. \ u_{bin}$</td>
<td>$cst. \ u_{bin}$</td>
<td>$cst. \ u_{bin}$</td>
<td>pref. $u_{bin}$</td>
<td>state</td>
</tr>
<tr>
<td>Percentage measured [%]</td>
<td>100</td>
<td>30</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td><strong>Ctrl. error (rel. RMSE) [%]</strong></td>
<td>1.72</td>
<td>1.50</td>
<td>2.07</td>
<td>0.86</td>
<td>1.92</td>
</tr>
<tr>
<td><strong>Mean switching increase [%]</strong></td>
<td>173.65</td>
<td>221.80</td>
<td>284.54</td>
<td>116.33</td>
<td>211.48</td>
</tr>
<tr>
<td><strong>Min switching increase [%]</strong></td>
<td>70.59</td>
<td>100.00</td>
<td>111.76</td>
<td>56.25</td>
<td>80</td>
</tr>
<tr>
<td><strong>Max switching increase [%]</strong></td>
<td>325.00</td>
<td>450.00</td>
<td>527.27</td>
<td>233.33</td>
<td>416.67</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Case 7</th>
<th>Case 8</th>
<th>Case 9</th>
<th>Case 10</th>
<th>Case 11</th>
<th>Case 12</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{bin}$</td>
<td>60</td>
<td>60</td>
<td>80</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>Control penalty (Table 4.4)</td>
<td>pref. $u_{bin}$</td>
<td>state</td>
<td>$cst. \ u_{bin}$</td>
<td>pref. $u_{bin}$</td>
<td>state</td>
</tr>
<tr>
<td>Percentage measured [%]</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>10</td>
</tr>
<tr>
<td><strong>Ctrl. error (rel. RMSE) [%]</strong></td>
<td>0.61</td>
<td>0.96</td>
<td>0.78</td>
<td>0.49</td>
<td>0.79</td>
</tr>
<tr>
<td><strong>Mean switching increase [%]</strong></td>
<td>144.71</td>
<td>273.96</td>
<td>193.10</td>
<td>149.52</td>
<td>267.61</td>
</tr>
<tr>
<td><strong>Min switching increase [%]</strong></td>
<td>71.43</td>
<td>135.29</td>
<td>82.35</td>
<td>71.43</td>
<td>114.29</td>
</tr>
<tr>
<td><strong>Max switching increase [%]</strong></td>
<td>261.54</td>
<td>458.33</td>
<td>454.55</td>
<td>336.36</td>
<td>461.54</td>
</tr>
</tbody>
</table>
4.5. Case Study

Figure 4.4: Simulation result for full state information
Furthermore, we investigate how the control algorithm influences the \textit{ON/OFF} switching frequency of the TCL population. The number of \textit{ON/OFF} alterations (switchings) in each simulated case is compared with the number of switchings in the case where all TCLs operate autonomously. The average increase in switching frequency is roughly 100 – 330\%. This significant increase is caused by the shape of the desired trajectory which imposes large deviations of active power consumption with respect to the absence of control, thus forcing the population to sometimes operate far away from the mean SOC of 0.5. Again, the best results are achieved by the preferential switching method.

\section*{4.6 Concluding Remarks}

This chapter demonstrates the ability of a large number of thermostatically controlled loads to track a variable power signal. The bin transition modeling technique provides a model of the aggregate dynamic behavior of the TCL population which is accurate enough for use with model-based control techniques. Heterogeneity of the population was incorporated into the approach and tracking performance was demonstrated with a set of 1,000 TCLs.

Further research will address the reduction of communication requirements so as to increase the chance for cost-effective practical application of the technique. An issue is the quality of available state information to the controller, which makes the application of state estimation and filtering techniques promising for increasing tracking performance in communication-constrained environments.
Chapter 5

System-Level Unit Models

In this chapter, we present a framework for the system-level representation of distributed loads, generators, and storage units. This modeling approach, denoted as Power Nodes Modeling Framework (PNMF), consists of a unified description of a diverse portfolio of grid-connected units (individual or aggregated), which facilitates the development of operation strategies and evaluation of system performance. The models of a population of generic Thermostatically Controlled Loads (TCLs) shown in Chapter 2 as well as the Electric Water Heaters (EWHs) in Chapter 3 are represented by power node equations. Furthermore, individual energy storage devices and generation units are modeled using the PNMF nomenclature. We will use these models for control strategy development in Chapter 6 and for economic evaluation in Chapter 7.

5.1 Introduction and Literature Review

Power nodes are mathematical representations of power sources or sinks connected to the electricity system. They were first presented in [129, 130]. Note that the development of the basic modeling approach was a collaborative effort between the co-authors of these two publications, whereas the concept was later utilized for individually different purposes by each of the active researchers. As will be seen later, the work presented in this thesis is focused on the formulation of power system control services and the integration of Demand Response (DR) into dispatch strategies. Other applications of the concept independent of this thesis include conceptual considerations on power system control structures [131] and a structured assessment of power system flexibility [132].
The modeling concept builds upon the principle that any unit connected to the power system, either generation, load, or storage, requires the conversion of electric energy from and to different forms of energy. These forms can be called “supply forms” (e.g., primary energy such as natural gas) and “use forms” (e.g., end-use energy such as heat and light) of energy. The electric power balance in the grid can only be fulfilled due to the possibility to store at least some use- and supply energy forms and to accurately control the conversion processes, e.g., the furnace in a thermal power plant. To represent a grid-connected use- or supply process from a power system perspective, it is not necessary to consider the particular internal composition and physical properties of the process itself or of the intermediate energy conversion process. Unit properties relevant to the system level, such as storage capacity and ramping capability, can be represented in a unified way. We can thus model a unit connected to the power system as a lumped entity with characteristic parameters. This entity is called a “power node”.

Although the integration of energy storage and load flexibility into power systems has been discussed extensively in the literature, the Power Nodes Modeling Framework (PNMF) exhibits novel features. Some work that has been published in the same area is found to be methodologically similar but different in application focus, while other approaches pursue a similar goal but use a different set of tools. The multitude of ideas in the same area underlines the interest in a comprehensive modeling framework for power systems in the era of intermittent and distributed power generation, controllable loads, and energy storage. We will mention some of these concepts in the following.

Since the paradigm of Distributed Generation (DG) acquired more and more popularity, the challenge of coordinating large numbers of small units has shifted into the focus of many researchers. New concepts emerged to deal with the new complexity. The combination of DG and storage units into a single entity that can operate in power markets is referred to as Virtual Power Plant (VPP) [133]. VPPs are usually centrally controlled and operated using optimization algorithms. Their primary function is market participation but also technical control goals can be addressed. For this purpose, a distinction can be made between “technical” and “economic” VPPs [81].

The well-known Energy Hub concept [134] developed at ETH Zurich within the project “Vision of Future Energy Networks” [135] provides a modeling framework for multi-energy systems which can include thermal
and electric energy storage devices. One motivation for this work is the joint optimization of multiple-energy-carrier systems, such as interconnected networks for electricity, natural gas, and district heat [136, 137]. The reliability of the overall system can be modeled [138] and investment decisions under uncertainty of energy hubs can be made in a structured framework [139]. Energy storage can be incorporated but is not the main focus of the energy hub framework. A limiting factor for the application of energy hubs to large-scale systems is the computational complexity of optimizations using nonlinear models, although new ways to eliminate the nonlinearities emerged recently [140].

The key distinction between energy hubs and power nodes is that power nodes are focused on the electricity system and are tailored for the easy integration of a diverse unit portfolio into dispatch methodologies. Energy hubs are targeted at future energy systems where a large optimization potential lies in the interconnection of different energy carrier networks. Conversely, power nodes are designed for the utilization in today’s power system comprising novel elements such as controllable loads.

There are a number of approaches that emerged in parallel to the PNMF. In [141], the authors develop a modeling framework for energy storage in power systems following a rationale comparable to ours, while the way of integrating dispatch and control services is different. Reference [142] proposes a multi-stage optimization framework for the operation of storage units to manage energy imbalance, in which the storage units are represented in a nomenclature similar to the PNMF. Reference [143] proposes the application of convex optimization methods for the dispatch of storage devices in power systems, demonstrated by simulation of a price-based dispatch scenario with artificially generated price and demand signals. The usage of the PNMF for providing power system services as will be presented in Chapter 6 follows a similar rationale, albeit with a more detailed framework for the control services. Reference [144] presents control strategies for optimized feed-in (time shifting) and capacity firming of wind energy using a battery storage system. This work is similar to ours in that we are interested in the same kind of control services, but the storage modeling approach is quite different.

The issue of energy storage dispatch is a long-standing problem in the operation of hydro power plants, as explained in, e.g., [145]. As numerous other issues, such as water level and flow constraints, natural
inflows, seasonality, and the influence of the available head, need to be considered in hydro power modeling, a detailed review of hydro modeling and optimization approaches is beyond the scope of this work.

As we will see in Chapter 6, power nodes can be conveniently integrated into Model Predictive Control (MPC) strategies. The application of MPC to power system dispatch has been widely discussed due to its strong basis in optimization theory and its receding horizon capability which facilitates the execution in rolling time windows and enables the rejection of disturbances. For instance, [146] proposes the application of modified MPC problems to power system dispatch in the presence of load flexibility and intermittent renewable energy.
5.2. Power Nodes Modeling Framework

The remainder of this chapter is organized as follows: The PNMF is introduced in Section 5.2 and put into context with conventional power system models. Furthermore, the dynamic model of an individual power node is explained, parameter/constraint sets representing the characteristics of certain power system units are discussed, and balance terms for system performance evaluation are presented. Section 5.3 discusses the modeling of a number of different generators, storage devices, and loads as power nodes. Section 5.4 presents concluding remarks. The used notation is summarized in Table 5.1.

5.2 Power Nodes Modeling Framework

In this section, the PNMF is introduced. We will first explain the integration of power node models into an ordinary power system simulation, then introduce the model for an individual power node, characterize the properties of a number of standard units, and introduce a balance term framework for performance evaluation.

5.2.1 Modeling Domains

We define a “modeling domain” as an environment that contains a mathematical description of an underlying physical system. A power system simulation normally consists of grid component models and electrical representations of attached generators and loads. In order to integrate units modeled by the PNMF into a power system simulation, an additional modeling domain is created. As depicted in Figure 5.1, this power node domain interfaces the grid modeling domain with an external demand/supply domain containing information about energy use-and supply processes. These processes can be both externally driven, such as intermittent renewable energy supply, or controllable, such as the supply of fuel for dispatchable generators. While the power node and grid domains are considered integral parts of the electric energy system, the domain of demand/supply processes is considered external. In Figure 5.1, arrows indicate the energy (or power) flows that are taken into account. Empty arrowheads indicate energy that is exchanged with the environment, while solid arrowheads indicate energy flows into or across the modeling domains.

Interfaces between the modeling domains need to be defined in an unambiguous way in order to avoid model inconsistencies. An exchange
between domains is represented in terms of power in continuous time, which can be integrated in time to yield the exchanged energy. For example, the exchange between the power node domain and the grid domain is the active power fed into or consumed from the grid. If the power system is modeled including electro-mechanical dynamics, the system inertia provided by synchronous machines is part of the grid domain. Thus, the active power interface is equivalent to the mechanical power exerted by the prime mover. Grid losses are modeled inside the grid domain as well, while pre-grid losses, such as storage and conversion losses, are accounted for in the power node domain. This clear separation allows the PNMF to integrate with a number of different physical network representations common in power systems modeling.

In a power system model utilizing the PNMF, all grid-connected sources and sinks of electric power are represented by power nodes. This allows to account for all energy flows provided to or consumed from the grid as power node quantities, which enables an easy evaluation of the system’s performance. Real-world effects that cause supplied energy to be lost, or demanded energy to remain unserved, can be modeled in all kinds of units. For example, energy conversion implies conversion losses, power in-feed from wind turbines may be curtailed, and a load may get disconnected from the grid. In order to evaluate the performance of the overall system, it is necessary to keep track of these losses and to account for the value associated with them. For this purpose, balance terms as presented in Section 5.2.4 can be utilized.
5.2. Power Nodes Modeling Framework

5.2.2 Model of a Single Power Node

A single power node is structured as depicted in Figure 5.2. In comparison with Figure 5.1, the provided and demanded energies are lumped into an external process termed $\xi$ [MW], with $\xi < 0$ denoting use and $\xi > 0$ supply. The term $u_{\text{gen}} \geq 0$ [MW] describes a conversion corresponding to a power generation with efficiency $\eta_{\text{gen}}$ [-], while $u_{\text{load}} \geq 0$ [MW] describes a conversion corresponding to a consumption with efficiency $\eta_{\text{load}}$ [-]. The energy storage level is normalized to $0 \leq x_i \leq 1$ [-] with energy storage capacity $C \geq 0$ [MWh]. Figure 5.2 illustrates how the storage serves as a buffer between the external process $\xi$ and the two grid-related exchanges $u_{\text{gen}}$ and $u_{\text{load}}$. Internal energy losses associated with energy storage, e.g., physical, state-dependent losses, are modeled by the term $v_{\text{i}} \geq 0$ [MW], while enforced energy losses, e.g., curtailment/shedding of a supply/demand process, are denoted by the waste term $w_{\text{i}}$ [MW], where $w > 0$ denotes a loss of provided energy and $w < 0$ an unserved demand process.

**Generic Model**

The dynamics of a power node $i \in \mathcal{N} = \{1, \ldots, N\}$, which may be nonlinear in the general case, are described by:

\[
C_i \dot{x}_i = \eta_{\text{load},i} u_{\text{load},i} - \eta_{\text{gen},i}^{-1} u_{\text{gen},i} + \xi_i - w_{\text{i}} - v_{\text{i}}, \quad (5.1)
\]

s.t.

(a) $0 \leq x_i^{\min} \leq x_i \leq x_i^{\max} \leq 1$ ,
(b) $0 \leq u_{\text{gen},i}^{\min} \leq u_{\text{gen},i} \leq u_{\text{gen},i}^{\max}$ ,
(c) $0 \leq u_{\text{load},i}^{\min} \leq u_{\text{load},i} \leq u_{\text{load},i}^{\max}$ ,
(d) $0 \leq \xi_i \cdot w_{\text{i}}$ ,
(e) $0 \leq |\xi_i| - |w_i|$ ,
(f) $0 \leq v_{\text{i}}$ .

The power node variables $u_{\text{load}}, u_{\text{gen}}, \xi,$ and $w$ may in general be measured or not, controllable (by an algorithm or operator) or not, and driven by an external influence or not. These properties are dependent upon the physical process that is represented by the power node. Internal dependencies, e.g., state-dependent losses $v_i(x_i)$, can be modeled. Charge and discharge efficiencies may be non-constant in the general case, e.g., state-dependent: $\eta_{\text{load},i} = \eta_{\text{load},i}(x_i), \eta_{\text{gen},i} = \eta_{\text{gen},i}(x_i)$.

The constraints (a) – (f) denote a generic set of requirements on the variables. They are to express that (a) the State of Charge (SOC) is
Figure 5.2: Notation for a single power node [129]

normalized between zero and one and may also be subject to further constraints, (b, c) the grid variables are non-negative and bounded, (d) the supply/demand and the curtailment need to have the same sign, (e) the supply/demand curtailment cannot exceed the supply/demand itself, and (f) the storage losses are non-negative. Ramp-rate constraints, especially constraints on the derivatives $\dot{u}_{\text{gen},i}$ and $\dot{u}_{\text{load},i}$, can be included to represent the dynamic limitations of units, mostly generation assets, in a simplified way.

Apart from the constraints listed here, there may be additional bounds imposed on the variables, e.g., in order to define certain standard unit types with characteristic properties (cf. Section 5.2.3). A major motivation for this notation is to provide technology-independent categories that can be linked to the evaluation functions given in Section 5.2.4.

**Modeling a Power Node without Storage**

One of the main motivations for representing power system units as power nodes is the structured integration of energy storage into operational frameworks. However, power nodes are also useful to represent processes without energy storage properties such as intermittent renewable generation or conventional generation and load. A process without storage implies an algebraic coupling between the instantaneous quantities $\xi_i$, $w_i$, $u_{\text{gen},i}$, and $u_{\text{load},i}$; the term $v_i$ is naturally equal to zero since no storage is present. Equation (5.1) degenerates to

$$\xi_i - w_i = \eta_{\text{gen},i}^{-1} u_{\text{gen},i} - \eta_{\text{load},i} u_{\text{load},i} \quad .$$  \hspace{1cm} (5.2)
This equation holds for both externally driven and controllable power system units. A highly relevant case is a unit that is externally driven by a physical supply/demand process which cannot be influenced directly, but the possibility of curtailment exists. For a supply process, this means that the supplied energy is wasted, while for a demand process curtailment implies that the demand remains unserved. In both cases, the curtailment is brought about by the waste term \( w_i \). Examples are intermittent power generation \((0 \leq \xi_i = \xi_{\text{drv},i}(t), 0 \leq w_i \leq \xi_i)\) and classical load \((\xi_i = \xi_{\text{drv},i}(t) \leq 0, \xi_i \leq w_i \leq 0)\).

In the case of a fully controllable supply process such as a conventional generator, either the grid-related variables \( u_{\text{gen},i}, u_{\text{load},i} \), or the power exchange with the environment through \( \xi_i \) can be considered the controlled variables. The variable \( \xi_i \) then accounts, e.g., for primary energy supply.

**Affine Power Node Model**

Specializations and simplifications of the generic power node model have practical relevance for controller design and implementation. Here we present the example of an affine power node model which is suitable for describing a wide range of processes with state-dependent losses, such as heat storage units that dissipate energy to the ambiance due to a difference between the internal storage temperature and the ambient temperature. For this purpose, a linear dependency of \( v_i \) on the storage state \( x_i \) is assumed and the efficiencies are assumed constant in order to eliminate nonlinearities:

\[
C_i \dot{x}_i = \eta_{\text{load},i} u_{\text{load},i} - \eta_{\text{gen},i}^{-1} u_{\text{gen},i} + \xi_i - w_i - a_i (x_i - x_{ss,i})
\]  (5.3)

subject to suitable constraints. The steady-state storage level \( x_{ss,i} \) refers to the steady state of the differential equation in the absence of inputs, e.g., the thermal equilibrium of a heat storage with the ambiance, and \( a_i \) [MW] is a non-negative loss coefficient.

**5.2.3 Characterization of Unit Properties**

There are a limited number of practically relevant unit types. As discussed in Section 5.2.2, the various kinds of energy flows available in the generic power node model allow the modeling of a wide range of unit types. A certain practical unit type is thus classified by its characteristic subset of the possible modes of energy flow. A “unit” in the PNMF...
Table 5.2: Unit properties determined by power node constraints

<table>
<thead>
<tr>
<th>Variable(s)</th>
<th>Constraint(s)</th>
<th>Implication</th>
</tr>
</thead>
<tbody>
<tr>
<td>$u_{\text{gen},i}$, $u_{\text{load},i}$</td>
<td>$u_{\text{gen},i} = 0$</td>
<td>Load</td>
</tr>
<tr>
<td>$u_{\text{load},i}$</td>
<td>$u_{\text{load},i} = 0$</td>
<td>Generator</td>
</tr>
<tr>
<td>$u_{\text{gen},i} \cdot u_{\text{load},i} = 0$</td>
<td>One-conversion-unit storage</td>
<td></td>
</tr>
<tr>
<td>$- $</td>
<td>Two-conversion-unit storage</td>
<td></td>
</tr>
<tr>
<td>$C_i$</td>
<td>$C_i = 0$</td>
<td>Non-buffered unit</td>
</tr>
<tr>
<td>$C_i &gt; 0$</td>
<td>Buffered unit</td>
<td></td>
</tr>
<tr>
<td>$\xi_i$</td>
<td>$\xi_i = 0$</td>
<td>No external process</td>
</tr>
<tr>
<td>$\xi_i \geq 0$</td>
<td>Supply process</td>
<td></td>
</tr>
<tr>
<td>$\xi_i \leq 0$</td>
<td>Demand process</td>
<td></td>
</tr>
<tr>
<td>$\xi_i, w_i$</td>
<td>$\xi_i = \xi_{\text{drv},i}(t) \land w_i = 0$</td>
<td>Non-controllable</td>
</tr>
<tr>
<td>$\xi_i = \xi_{\text{drv},i}(t)$</td>
<td>Curtailable</td>
<td></td>
</tr>
<tr>
<td>$\xi_i$ arbitrary, $w_i = 0$</td>
<td>Controllable</td>
<td></td>
</tr>
<tr>
<td>$v_i$</td>
<td>$v_i = 0$</td>
<td>Lossless storage</td>
</tr>
<tr>
<td>$v_i \geq 0$</td>
<td>Lossy storage</td>
<td></td>
</tr>
<tr>
<td>$\dot{u}<em>{\text{gen},i}$, $\dot{u}</em>{\text{load},i}$</td>
<td>$\dot{u}<em>{\text{min}} \leq \dot{u}</em>{\text{gen},i} \leq \dot{u}_{\text{max}}$</td>
<td>Ramp-rate-constrained generation</td>
</tr>
<tr>
<td>$\dot{u}<em>{\text{min}} \leq \dot{u}</em>{\text{load},i} \leq \dot{u}_{\text{max}}$</td>
<td>Ramp-rate-constrained load</td>
<td></td>
</tr>
</tbody>
</table>

is an arbitrary generation, load, or storage device, or a group of devices which are coordinated and aggregated such that they behave like a single unit. The type distinction is established by a set of constraints on the variables used in (5.1), i.e., $u_{\text{load},i}$, $u_{\text{gen},i}$, $C_i$, $x_i$, $\xi_i$, $v_i$, and $w_i$. These constraints hold in addition to the principal constraints (a) – (f) in (5.1), providing a classification of units with different operational properties.

Table 5.2 establishes a set of basic properties defining the operational behavior of a unit modeled as a power node. The interpretation of the constraints is given in the following:

- **$u_{\text{load/gen},i}$**: A pure generation process would imply that $u_{\text{load},i} = 0$ at all times; a pure load cannot inject power, which is expressed by $u_{\text{gen},i} = 0$. In a bi-directional conversion system, both variables can assume non-zero values; both conversions can happen at the same time (e.g., in a pumped-hydro power plant with independent turbine and pump), or not (e.g., in an inverter-connected battery).

- **$C_i$**: The unit is modeled with ($C_i > 0$) or without energy storage capabilities ($C_i = 0$).

- **$\xi_i$**: This denotes a supply ($\xi_i > 0$) or demand process ($\xi_i < 0$). For a pure electricity storage (e.g., a battery), $\xi_i = 0$ holds.
• ξ_i, w_i: Constraints on ξ_i and w_i indicate the controllability of the power exchange with an external process. If ξ_i is driven by an external signal ξ_i = ξ_{drv,i}(t), e.g., induced by an intermittent supply, it may either be curtailable (w_i ≤ ξ_i, but no further constraint on w_i) or non-controllable (no curtailment possible: w_i = 0). If ξ_i is not externally driven, the unit is fully controllable. In this case, w_i = 0 can be assumed since the curtailment of a directly controllable process would be unnecessary.\(^1\)

• v_i: The storage is considered lossless if v_i = 0 and lossy if the term can assume non-zero values (v_i ≥ 0).

• ˙u_{load/gen,i}: The grid variables u_{load,i} and u_{gen,i} may be rate-constrained, which is reflected in continuous time by an upper and lower bound on their derivatives. This serves to model physical limitations on the rate of change of a power conversion process, e.g., due to thermal stress on power plant components.

Based on these properties, all unit types relevant for establishing the power balance in a power system can be classified and modeled in a simplified way within the PNMF. Some examples are given in Section 5.3. Additional constraints may be considered for specific applications.

### 5.2.4 System-Level Performance Indicators

In order to evaluate operation and control strategies acting on an electric grid interfaced with a set of power nodes, a number of performance indicators can be defined which allow to account for the energy flows over time in an aggregated form. For this purpose, balance terms depending on the instantaneous quantities (power/energy flow quantities) can be formulated that characterize the current operational state of the power system. These can be integrated over time to yield energy values which characterize the system performance over a certain time span. Below, a set of exemplary terms is presented, which can be extended to include also technology-dependent weighting terms for monetary cost or environmental impact. Examples for instantaneous balance terms indicating the current system state are:

\(^1\)Note that more detailed sets of constraints may be established for the power node variables in order to model particular units. It may be practical to allow a non-zero w_i even in the presence of a (partly) controllable ξ_i.
- Power supplied to grid: \( P_{\text{gen}}(t) = \sum_{i \in \mathcal{N}} u_{\text{gen},i}(t) \),

- Power consumed from grid: \( P_{\text{load}}(t) = \sum_{i \in \mathcal{N}} u_{\text{load},i}(t) \),

- Currently stored energy: \( E_{\text{stored}}(t) = \sum_{i \in \mathcal{N}} C_i x_i(t) \),

- Power supply available: \( \xi_{\text{supply}}(t) = \sum_{i \in \{i | \xi_i > 0\} \subset \mathcal{N}} \xi_i(t) \),

- Power demand: \( \xi_{\text{demand}}(t) = \sum_{i \in \{i | \xi_i < 0\} \subset \mathcal{N}} \xi_i(t) \),

- Power supply curtailed: \( w^+(t) = \sum_{i \in \{i | w_i(t) > 0\} \subset \mathcal{N}} w_i(t) \),

- Power demand not served: \( w^-(t) = \sum_{i \in \{i | w_i(t) < 0\} \subset \mathcal{N}} w_i(t) \),

- Power conversion loss: \( P_{\text{loss}}(t) = \sum_{i \in \mathcal{N}} \left( \frac{1 - \eta_{\text{gen},i}(t)}{\eta_{\text{gen},i}(t)} u_{\text{gen},i}(t) + (1 - \eta_{\text{load},i}(t)) u_{\text{load},i}(t) \right) \).

All of the above quantities can be restricted to certain unit types by placing restrictions on the index \( i \). For example, the consideration of all non-controllable non-buffered generation units would require a summation over the index \( i \in \{i | C_i = 0 \land \xi_i = \xi_{\text{drv},i}(t) \geq 0 \land w_i = 0\} \subset \mathcal{N} \).

Energy balance terms can be derived by time-integration over instantaneous balance terms in the time interval \([t_1, t_2]\), such as

- Electric energy supplied to grid: \( \int_{t_1}^{t_2} P_{\text{gen}}(t) \, dt \),

- Primary energy supplied: \( \int_{t_1}^{t_2} \xi_{\text{supply}}(t) \, dt \),

- Primary energy curtailed: \( \int_{t_1}^{t_2} w^+(t) \, dt \),

- Energy conversion losses: \( \int_{t_1}^{t_2} P_{\text{loss}}(t) \, dt \).
5.3 Development of Unit Models

We now make use of the framework presented above in order to derive a number of power node models for common power system units. These include generation units, energy storage devices, and both controllable and non-controllable loads.

5.3.1 Generation Units

In this work, generators are modeled as algebraic power nodes without storage. This implies that there is no dynamic behavior represented in the model, an exception being the hydro generator, which possesses a water reservoir. Nevertheless, for the dispatch and control services described in Chapter 6, a static model is suitable since dynamic limitations can be represented by a ramp-rate constraint. Note that in short-term control actions such as secondary frequency control, plant dynamics do play a role for the actual reaction to the control signal. This can be modeled outside of the power node domain by grid-side dynamic models. The power node model is then to be seen as a means for setpoint determination rather than a full representation of a generation unit. We will model controllable and non-controllable generators with the associated constraints and cost terms in the following.

Thermal Generation

Thermal generators rely on a chemically stored primary energy such as natural gas, biomass, or coal. The conversion process can be controlled such that the primary energy input and the electric energy output, coupled by the generator’s efficiency, can be freely determined within a certain range. Ramping constraints exist due to physical limitations of the combustion process and the turbine cycle. The power node equation reads

\[ u_{gen} = \eta_{gen} \xi \]  

subject to the constraints

\[ 0 \leq u_{gen}^{\text{min}} \leq u_{gen} \leq u_{gen}^{\text{max}}, \]
\[ \dot{u}_{gen}^{\text{min}} \leq \dot{u}_{gen} \leq \dot{u}_{gen}^{\text{max}}. \]  

In the simplified case of a constant generator efficiency, the fuel cost can be expressed as a linear term in the primary energy input \( \xi = \eta_{gen}^{-1} u_{gen} \).
The cost for Operation & Maintenance (O&M) of the generator are linear as well. The cost function $J_{\text{gen}} \left[ \text{EUR} \right]$ reads thus:

$$J_{\text{gen}} = \pi_{\text{fuel}} \eta_{\text{gen}}^{-1} u_{\text{gen}} + \pi_{\text{O&M}} u_{\text{gen}}$$ (5.7)

where $\pi_{\text{fuel}} \left[ \text{EUR/MWh} \right]$ is the fuel price and $c_{\text{O&M}} \left[ \text{EUR/MWh} \right]$ is the cost for O&M of the plant per produced unit of energy.

Non-constant generator efficiencies are usually modeled by a convex polynomial generation cost function. As described in [147], the heat rate of the generator, expressed in million British Thermal Units per megawatt-hour $\left[ \text{MBTU/MWh} \right]$, is inversely proportional to the generator’s efficiency and is often used in experimental identification of the energy input-output relation in a generator. By curve fitting techniques, the fuel cost depending on the power production can be expressed by the polynomial

$$J_{\text{gen}} = c_0 + c_1 P_{\text{p.u.}} + c_2 P_{\text{p.u.}}^2$$ (5.8)

where $c_0$, $c_1$, $c_2 \left[ \text{EUR} \right]$ are cost coefficients and $P_{\text{p.u.}}$ is the generator output power in per unit based on its maximum power output. In PNMF nomenclature, this can be expressed as

$$J_{\text{gen}} = c_0 + \frac{c_1}{u_{\text{gen}}^{\text{max}}} u_{\text{gen}} + \frac{c_2}{(u_{\text{gen}}^{\text{max}})^2} u_{\text{gen}}^2$$ (5.9)

O&M cost can be included in the same way as in (5.7). Start-up and shut-down costs are neglected in the context of this work since the unit commitment problem is not considered. As a side remark, we can state that the modeling of start-up and shut-down can be achieved by introducing a binary variable multiplied by $u_{\text{gen}}$ in the power node equation, the change of which can be penalized accordingly.

Depending on the type of power plant, significant additional cost can be incurred by ramping the power plant up and down for load-following, balancing, or frequency control. According to [148], the ramping cost can be approximated by a quadratic cost term, here expressed in a continuous-time framework:

$$J_{\text{gen,ramp}} = \pi_{\text{ramp}} \dot{u}_{\text{gen}}^2$$ (5.10)

where the derivative of the generated power is penalized quadratically, multiplied by the specific ramping cost $\pi_{\text{ramp}} \left[ \text{EUR/MW}^2/\text{h} \right]$. The ramping cost is dependent upon the physical properties of the generator and
represents the cost incurred through increased fuel usage and wear and tear of plant equipment.

The cost for Greenhouse Gas (GHG) emissions of the generator can be modeled in dependency on the primary energy input:

\[ J_{\text{gen, GHG}} = \pi_{\text{GHG}} \xi \]  

(5.11)

with the price of CO\(_2\)-equivalent emissions normalized to units of consumed primary energy \( \pi_{\text{GHG}} \) [EUR\(\text{MWh}^{-1}\)]. Depending on the considered legal framework for GHG emissions and the operational objectives of the power node portfolio, the GHG emission cost can be included in the generator cost function.

**Intermittent Generation**

Intermittent generation is defined by the fact that the power output depends on a source of primary energy that cannot be controlled. The availability of that energy source can exhibit characteristic hourly, daily, and seasonal variations, as well as short-term stochastic behavior. These patterns can be captured by time series (coming either from a simulation model or realized data) which enter the power node via the term \( \xi \). Depending on the conversion process, the power in-feed may be controlled to be less than the potentially available in-feed, which we will refer to as curtailment, denoted by \( w \). This is particularly relevant in the case of grid constraints in the vicinity of the in-feed as well as for system-level balancing at high intermittent Renewable Energy Sources (RES) penetration levels. In a wind farm, curtailment functions are usually implemented in the control systems and are accessible via remote control from a control center. Conversely, an aggregation of residential solar photovoltaic (PV) installations might not allow curtailment at all (modeled by the constraint \( w = 0 \)). The power node equation reads:

\[ u_{\text{gen}} = \eta_{\text{gen}}(\xi - w) \]  

(5.12)

subject to the constraints:

\[ 0 \leq u_{\text{gen}}^{\min} \leq u_{\text{gen}} \leq u_{\text{gen}}^{\max} \] ,  

(5.13)  
\[ 0 \leq w \leq \xi \] ,  

(5.14)  
\[ \xi = \xi_{\text{drv}}(t) \] .  

(5.15)

Since the time-varying primary energy input from wind and solar power, driven by the external process \( \xi_{\text{drv}}(t) \), is free of charge, it is not obvious
what the operating cost function to be used in an optimization problem must look like. The cost for O&M are similarly defined as for control-

able generation and the fuel cost is equal to zero. An ambiguity exists with respect to the cost of the curtailment through the term \( w \): On the one hand, it denotes the wasting of freely available energy, so no real cost is incurred by curtailing intermittent in-feeds – rather, the wasted energy eventually has to be supplied by controllable generation which incurs fuel cost. On the other hand, voluntary curtailment contracts or regulatory penalties of RES curtailment such as in Germany [149] may be in place, which means that curtailment implies real financial cost. In the cost function below, the second case is expressed. The first case can be modeled by omission of the second penalty term. The cost function \( J_{\text{gen}} \) reads:

\[
J_{\text{gen}} = \pi_{\text{O&M}} u_{\text{gen}} + \pi_{\text{curt}} w ,
\]

where \( \pi_{\text{curt}} \) is the curtailment cost. The cost for GHG emissions can be set equal to zero since any emissions incurred during the plant construction phase are not part of the dispatch optimization problem:

\[
J_{\text{gen,GHG}} = 0 .
\]

### Hydro-Electric Generation

The hydro power plant is modeled as a storage power node with an external energy input arising through natural water inflow into the reservoir. It may exhibit the possibility to spill stored water and is usually subject to evaporation losses. The power node equation reads:

\[
C \dot{x} = \eta_{\text{gen}}^{-1} u_{\text{gen}} + \xi - w - v
\]

subject to the constraints

\[
0 \leq u_{\text{gen}}^{\min} \leq u_{\text{gen}} \leq u_{\text{gen}}^{\max} ,
\]

\[
0 \leq x^{\min} \leq x \leq x^{\max} \leq 1 ,
\]

\[
\xi = \xi_{\text{drv}}(t) \geq 0 ,
\]

\[
v \geq 0 ,
\]

\[
w \geq 0 .
\]

The efficiency \( \eta_{\text{gen}} \) is dependent upon the head of the hydro power plant but can be assumed constant if the SOC variation is small on
the considered time scale. The term $\xi$ denotes the natural water inflow converted to a flow of potential energy and $v$ can contain a model for the evaporation losses from the reservoir. The shedding term $w$ can be used to model two different effects: the diversion of natural water in-flow such that it does not enter the reservoir as well as the spillage of water from the reservoir through a by-pass. In the first case, the additional constraint $w \leq \xi$ should be imposed since no more water inflow can be diverted than is available. In the second case, spillage can occur also at instants with no inflow into the reservoir.

The cost of operation of hydro power plants is influenced by the cost for O&M:

$$J_{\text{gen}} = \pi_{\text{O&M}} u_{\text{gen}}.$$  \hspace{1cm} (5.24)

The value of the water in the storage basin can be described by the opportunity cost that arises from using the water instead of keeping it in the storage for use at a later time. A discussion of the long-known concept of water value can already be found in [150] and is still referred to nowadays, e.g., in [151]. Further details are beyond the scope of this work.

In the least-cost dispatch strategy presented in Chapter 6, the value of using energy from a storage is implicitly taken into account by the optimizer since it tries to minimize overall operation cost. This, however, only holds for the time span within the optimization horizon. For large storage units that have discharge times beyond the hourly dispatch, a separation of the decision problem into long-, medium-, and short-term dispatch problems should be pursued.

### 5.3.2 Storage Units

Energy storage units are characterized by the ability to act as a load by storing electric energy and as a generation unit by feeding it back to the grid. Depending on the considered unit, the conversion efficiency may be significantly smaller than 1, and the storage can exhibit stand-by losses.

**Pumped-Hydro Storage**

The pumped-hydro storage power plant is an extension of the hydro-electric generator modeled above. Additionally, it contains a load term denoting the pumping function of the plant:

$$C \dot{x} = \eta_{\text{load}} u_{\text{load}} + \eta_{\text{gen}} u_{\text{gen}}^{-1} + \xi - w - v$$ \hspace{1cm} (5.25)
subject to the constraints

\begin{align}
0 \leq x_{\text{min}} & \leq x \leq x_{\text{max}} \leq 1, \quad (5.26) \\
0 \leq u_{\text{min}} & \leq u_{\text{gen}} \leq u_{\text{max}} \quad , \quad (5.27) \\
0 \leq u_{\text{min}} & \leq u_{\text{load}} \leq u_{\text{max}} \quad , \quad (5.28) \\
\xi = \xi_{\text{drv}}(t) & \geq 0 \quad , \quad (5.29) \\
v & \geq 0 \quad , \quad (5.30) \\
w & \geq 0 \quad . \quad (5.31)
\end{align}

The cost function of pumped-hydro power plants is based on the marginal operation cost attributed to operating the pump and turbine set, respectively. In order to avoid the curtailment of freely available (intermittent) energy although it could be stored, the operational cost should be attributed to the generation (discharge) side of the plant in system-level dispatch algorithms:

\begin{equation}
J_{PH} = \pi_{PH} u_{\text{gen}} \quad , \quad (5.32)
\end{equation}

where $\pi_{PH}$ [\text{EUR}/\text{MWh}] is the marginal cost of operating the storage. For ancillary service provision, however, it can be beneficial to attribute the cost equally to charging and discharging since control actions in both directions constitute a utilization of the flexible resource:

\begin{equation}
J_{PH} = \frac{\pi_{PH}}{2} u_{\text{gen}} + \frac{\pi_{PH}}{2} u_{\text{load}} \quad . \quad (5.33)
\end{equation}

**Battery Storage**

The Battery Energy Storage System (BESS) is modeled as a power node with in-feed and load term and a certain efficiency.

The charge and discharge management of a real BESS requires the detailed modeling of the physical battery properties including its internal parasitic losses depending on state variables such as the stored electric charge and its internal temperature. For this purpose, equivalent electrical representations of the electro-chemical interactions, such as in the third-order model presented in [152], are particularly useful. In a high-level dispatch strategy, however, specific physical properties such as the battery temperature should not be explicitly considered. These properties should rather be encapsulated in a local control system, yielding an aggregate value for the battery capacity (which may vary with the
5.3. Development of Unit Models

operating conditions) and the SOC. This is why a power node representation is useful for representing the battery towards the power system. The equation is

\[ C \dot{x} = \eta_{\text{load}} u_{\text{load}} - \eta_{\text{gen}}^{-1} u_{\text{gen}} - v \]  \hspace{1cm} (5.34)

subject to

\[ 0 \leq x_{\text{min}} \leq x \leq x_{\text{max}} \leq 1 , \]  \hspace{1cm} (5.35)

\[ 0 \leq u_{\text{gen}}^{\text{min}} \leq u_{\text{gen}} \leq u_{\text{gen}}^{\text{max}} , \]  \hspace{1cm} (5.36)

\[ 0 \leq u_{\text{load}}^{\text{min}} \leq u_{\text{load}} \leq u_{\text{load}}^{\text{max}} , \]  \hspace{1cm} (5.37)

\[ v \geq 0 . \]  \hspace{1cm} (5.38)

The term \( v \) may contain a physical model for the stand-by losses through self-discharge. Depending on the battery type and control strategy considered, it may be necessary to include a ramping constraint on \( u_{\text{gen}} \) and \( u_{\text{load}} \). As ramping capabilities of batteries are usually good compared to thermal generators [153], this will not be necessary in most cases.

The operation cost of batteries is caused by limited lifetime (expressed in number of charge/discharge cycles) and investment cost. The same argument as for the pumped-hydro plant applies to the battery in dispatch situations; the operation cost should be attributed to the generation side:

\[ J_{\text{bat}} = \pi_{\text{bat}} u_{\text{gen}} , \]  \hspace{1cm} (5.39)

where \( \pi_{\text{bat}} \) [EUR/MWh] is the marginal cost of operating the battery. For ancillary service provision, both directions can be included in the cost function in order to penalize the control resource utilization:

\[ J_{\text{bat}} = \frac{\pi_{\text{bat}}}{2} u_{\text{gen}} + \frac{\pi_{\text{bat}}}{2} u_{\text{load}} . \]  \hspace{1cm} (5.40)

5.3.3 Load Units

The load models presented below can represent a variety of use processes, which shall not be explicitly modeled. The two categories of load to be addressed are conventional (non-controllable, maybe scheddable) loads represented by static power nodes and loads with thermal inertia, e.g., aggregations of Thermostatically Controlled Loads (TCLs).

Conventional Load

A conventional load is modeled as a static power node with a load term coupled through an efficiency to the use process \( \xi \) which is driven
Chapter 5. System-Level Unit Models

externally. It may also be possible to shed power, which is denoted by the term \( w \). The power node equation is thus:

\[
    u_{\text{load}} = -\eta_{\text{load}}^{-1}(\xi - w) .
\]  

(5.41)

It is subject to the constraints

\[
0 \leq u_{\text{load}} \leq u_{\text{load}}^\text{max}, \quad \xi \leq w \leq 0, \quad \xi = \xi_{\text{drv}}(t) \leq 0 .
\]  

(5.42)-(5.44)

The cost function for a non-conventional load is equal to zero since penalizing an external use process is not meaningful. If shedding of the load is possible, it is usually to be penalized by a high value since it is a disruptive emergency action. The penalty in the cost function may correspond to the Value of Lost Load (VOLL) [154, 155, 156]:

\[
    J_{\text{loadshed}} = -\pi_{\text{VOLL}} w ,
\]  

(5.45)

where \( \pi_{\text{VOLL}} \text{ [EUR MWh]} \) equals the VOLL. There are also formulations of the VOLL which consider a dependency on the outage duration. This could be modeled by a piece-wise constant VOLL that changes with the number of periods in which \( w \) was unequal to zero, which would turn an associated optimization problem into a mixed-integer problem. We will not go into the details of this modeling discussion here.

**Generic Thermal Load Model**

A generic thermal load is modeled as a power node with associated storage and is described by a power node equation with a load term, a state-dependent thermal loss term, and an external demand term:

\[
    C\dot{x} = \eta_{\text{load}} u_{\text{load}} - a(x - x_{ss}) + \xi
\]  

(5.46)

subject to the constraints:

\[
0 \leq x_{\text{min}} \leq x \leq x_{\text{max}} \leq 1, \quad 0 \leq u_{\text{load}} \leq u_{\text{load}}^\text{max}, \quad \xi = \xi_{\text{drv}}(t) \leq 0 .
\]  

(5.47)-(5.49)

In Chapter 2, we introduced a first-order differential equation for the aggregated population of TCLs. Specifically, we used the following formulation in (2.38):

\[
    \frac{dE_{\text{el}}^{\text{total}}}{dt} = -\frac{1}{\tau} (E_{\text{el}}^{\text{total}} - E_{\text{el,amb}}^{\text{total}}) + P_{\text{el}}^{\text{total}} .
\]  

(5.50)
5.3. Development of Unit Models

This affine first-order model has the same structure as (5.46) and can easily be transformed to PNMF nomenclature by

\[
x = \frac{E_{\text{el}}^{\text{total}}}{C}, \quad x_{ss} = \frac{E_{\text{el,amb}}^{\text{total}}}{C}, \quad u_{\text{load}} = P_{\text{el}}^{\text{total}}, \quad a = C, \quad \eta_{\text{load}} = 1. \quad (5.51)
\]

The cost function of the thermal load shall represent the need to maintain the users’ comfort and control flexibility at the same time. This is mapped into a heuristic state penalty \( \pi_{\text{th}} [\text{EUR}_h] \) on the deviation from a defined SOC level:

\[
J_{\text{th}} = \pi_{\text{th}} (x - x_{\text{ref}})^2. \quad (5.52)
\]

The rationale of the quadratic penalization is that a larger deviation from the reference value should be penalized overproportionally.

**Electric Water Heater**

The modeling of Electric Water Heater (EWH) aggregations is an extension of the thermal load model presented above. In the following, a power node model which describes the dynamic behavior of an EWH’s energy storage content \( E_{\text{ewh}} \) is derived. The description here is largely based on the corresponding section in [157] where the modeling process is described in more detail. An illustration of a generic EWH and its energy in- and out-flows is shown in Figure 5.3.

Contrary to the aim of the modeling process in Chapter 3, we are now interested in representing the overall flexibility that an aggregation of EWHs provides for shifting load back and forth throughout the day. For this purpose, we use a single-state dynamic model that we aggregate over the entire population. A coordination algorithm similar to the ones presented in Chapter 3 will be required in order to enable the management of the EWH population as though it were a single entity.

**Thermal Energy Content** We start out with the modeling by stating that a single EWH draws a certain power \( P_{\text{load}} [\text{W}] \) from the grid. The energy extracted from the storage by drawing warm water is accounted for by the term \( \dot{E}_{\text{water,warm}} [\text{W}] \). Analogously, \( \dot{E}_{\text{water,cold}} [\text{W}] \) accounts for the energy contained in the in-flowing cold water. Due to a temperature gradient between the inside of the storage and the ambiance, heat losses occur. They are represented by the heat loss rate \( \dot{Q}_{\text{loss}} [\text{W}] \).
The evolution of the energy content $E_{\text{ewh}}$ [W] over time can thus be described by

$$\frac{dE_{\text{ewh}}}{dt} = P_{\text{load}} - \dot{Q}_{\text{loss}} + \dot{E}_{\text{water,cold}} - \dot{E}_{\text{water,warm}}. \quad (5.53)$$

By heating water up from the inlet temperature $T_{\text{water,cold}}$ [°C] to the outlet temperature $T_{\text{water,warm}}$ [°C], thermal energy is stored. We assume that these two temperatures are lower and upper bounds for the average water temperature in the tank. Consequently, the storage capacity of the water heater $C_{\text{ewh}}$ [Wh] can be calculated according to

$$C_{\text{ewh}} = c_{\text{water}}\rho_{\text{water}}V_{\text{ewh}}(T_{\text{water,warm}} - T_{\text{water,cold}}) \quad (5.54)$$

with the specific heat capacity $c_{\text{water}}$ (here expressed in [Wh/°C/kg]), the water density $\rho_{\text{water}}$ [kg/m$^3$], and the tank volume $V_{\text{ewh}}$ [m$^3$].

The stored energy inside the EWH can be derived from the average water temperature $T_{\text{ewh}}$ [°C], which serves as a dynamic state variable of the device. The temperature is mapped to the interval of [0,1] in relation to inlet and outlet temperatures similar to the normalization presented in Chapter 2. This mapping yields the SOC $x_{\text{ewh}} [-]$. The energy content $E_{\text{ewh}}$ [Wh] is equal to

$$E_{\text{ewh}} = C_{\text{ewh}}x_{\text{ewh}} = C_{\text{ewh}}\frac{T_{\text{ewh}} - T_{\text{water,cold}}}{T_{\text{water,warm}} - T_{\text{water,cold}}} \quad (5.55)$$

We stated in Chapter 3 that the average energy content of the EWH is not a good proxy for the user comfort since hot water will accumulate
5.3. Development of Unit Models

Figure 5.4: Profile describing the probability of a water draw occurring during the course of the day at the top of the tank. Here we take advantage of the assumption that the stratification is always present and that hot water is available to the customer if the SOC $x_{\text{ewh}}$ is above 0.2.

**Heat Losses** The heat losses to the ambiance are described by a heat conduction term:

$$\dot{Q}_{\text{loss}} = \lambda_{\text{ewh}} \frac{A_{\text{ewh}}}{d_{\text{ewh}}} (T_{\text{ewh}} - T_{\text{amb}}),$$

(5.56)

where $\lambda_{\text{ewh}}$ [W m K] denotes the heat conductivity through the tank insulation, $A_{\text{ewh}}$ [m$^2$] the tank surface, $d_{\text{ewh}}$ [m] the insulation thickness, and $T_{\text{amb}}$ [$^\circ$C] the ambient temperature.

**Water Draw Profile** We use the same warm water draw profile as in Chapter 3, taken from [122] and depicted again in Figure 5.4. As stated in [157], we derive from that profile the energy that is drawn from the population of EWHs over time:

$$WWD_{\text{agg,ewh}}(t) = p_{\text{ewh}} n_{\text{hh}} s_{\text{hh}} wwd_{p,\text{daily}} WWD_{\text{norm}}(t),$$

(5.57)

where $p_{\text{ewh}}$ denotes the share of households that own an EWH, $n_{\text{hh}}$ is the number of households, $s_{\text{hh}}$ is the average number of persons per household, $wwd_{p,\text{daily}}$ [m$^3$] is the daily hot water usage per person and $WWD_{\text{norm}}(t)$ is the norm water draw profile from Figure 5.4. Further details are given in [157].

**Power Node Model** The dynamic behavior of the energy $E_{\text{ewh}}$ in the EWH can be conveniently written as

$$C\dot{x} = u_{\text{load}} - a(x - x_{\text{ss}}) + \xi.$$

(5.58)
Based on the modeling presented above, the following identities hold:

\[
\begin{align*}
  u_{\text{load}} &= P_{\text{load}}, \\
  x &= \frac{E_{\text{ewh}}}{C_{\text{ewh}}}, \\
  a &= \lambda_{\text{ewh}} A_{\text{ewh}} (T_{\text{water,warm}} - T_{\text{water,cold}}) d_{\text{ewh}}, \\
  \xi &= c_{\text{water}} \rho_{\text{water}} (T_{\text{water,cold}} - T_{\text{water,warm}}) W W D_{\text{agg,ewh}}(t).
\end{align*}
\] (5.59, 5.60, 5.61, 5.62)

As in the case of the generic thermal load, we penalize a deviation from a desired SOC with a factor of \( \pi_{\text{ewh}} \) [EUR h]:

\[
J_{\text{ewh}} = \pi_{\text{ewh}} (x - x_{\text{ref}})^2.
\] (5.63)

### 5.4 Concluding Remarks

In this chapter, the PNMF was demonstrated to be a versatile framework for modeling units connected to the power system in a unified way. The parameterization of the power node equation and the associated constraint set allows to represent system-level unit properties of physical units, such as ramping constraints and energy storage capacity, within the power node model. A number of exemplary unit parameterizations with associated operational cost functions have been presented. In the upcoming chapter, we will utilize the PNMF and the presented unit models for the development of optimization-based control strategies in order to provide a number of different power system services with a flexible unit portfolio.

We have focused so far on the consistency and understandability of the framework and considered only linear and affine power node models for simplicity. Further research potential lies in the integration of nonlinearities into the approach, e.g., in the form of working-point-dependent efficiencies. However, the solution of optimization problems involving nonlinearities can become challenging, while the rationale of the control strategy for providing a certain power system service is not much affected by whether the underlying model is nonlinear or not. Consequently, we will restrict ourselves to linear/affine power node models also in the upcoming chapter and leave the extension to nonlinear models to future research.
Chapter 6

Dispatch Strategies

The power node models presented in Chapter 5 are now used to formulate dispatch strategies for a portfolio of units for power system control tasks or economic objectives. The portfolio can represent an entire electricity system or a Virtual Power Plant (VPP) operating in a larger system. For this purpose, we introduce a framework for multi-stage operation of the power node portfolio, present a number of Model Predictive Control (MPC) strategies for different use cases, and conduct simulations for demonstrating the developed approaches.

6.1 Introduction

Since we elaborated on the basics and the relevant literature for the dispatch of Virtual Power Plants (VPPs) in power systems already in Chapter 5, we do not formulate an extensive introduction here and rather stick to outlining the contents of this chapter. It is structured as follows: Section 6.2 presents a framework for multi-stage operation of the power node portfolio, which enables the combined implementation of day-ahead, intra-day, and real-time strategies. Section 6.3 discusses the usage of Model Predictive Control (MPC) for power system dispatch problems. In Section 6.4, we introduce a compact notation for power node portfolios which is used for the formulation of control problems for various use cases in Section 6.5. In Section 6.6, the simulation environment for simulating the dispatch problems is described. In Section 6.7, we present benchmark portfolios for the simulation of the use cases, which are used in the simulation examples presented in Section 6.8. Section 6.9 presents some concluding remarks. The used notation is summarized in Table 6.1.
### Table 6.1: Notation for Chapter 6

<table>
<thead>
<tr>
<th>Var.</th>
<th>Unit</th>
<th>Meaning</th>
<th>Var.</th>
<th>Unit</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>[-]</td>
<td>Dynamic matrix</td>
<td>R</td>
<td>EUR</td>
<td>Input penalty matrix</td>
</tr>
<tr>
<td>a</td>
<td>[MW]</td>
<td>Loss coefficient</td>
<td>r</td>
<td>EUR</td>
<td>Input penalty vector</td>
</tr>
<tr>
<td>B</td>
<td>[-]</td>
<td>Input matrix</td>
<td>δR</td>
<td>EUR</td>
<td>Input rate pen. matrix</td>
</tr>
<tr>
<td>B</td>
<td>[-]</td>
<td>Line susceptance (p.u.)</td>
<td>S</td>
<td>MW</td>
<td>Frequency droop</td>
</tr>
<tr>
<td>C</td>
<td>[MWh]</td>
<td>Storage capacity</td>
<td>S_B</td>
<td>[-]</td>
<td>Rated apparent power</td>
</tr>
<tr>
<td>f</td>
<td>[Hz]</td>
<td>System frequency</td>
<td>t</td>
<td>[h]</td>
<td>Time</td>
</tr>
<tr>
<td>H</td>
<td>[s]</td>
<td>Generator time constant</td>
<td>u</td>
<td>MW</td>
<td>Grid out-/in-feed</td>
</tr>
<tr>
<td>J</td>
<td>[EUR]</td>
<td>Cost function (discrete)</td>
<td>w</td>
<td>MW</td>
<td>Waste term</td>
</tr>
<tr>
<td>k</td>
<td>[-]</td>
<td>Simulation step index</td>
<td>x</td>
<td>[-]</td>
<td>State vector</td>
</tr>
<tr>
<td>l</td>
<td>[-]</td>
<td>Prediction step index</td>
<td>x</td>
<td>[-]</td>
<td>State of Charge</td>
</tr>
<tr>
<td>M</td>
<td>[-]</td>
<td>Set of buses</td>
<td>Y</td>
<td>[-]</td>
<td>LFC signal</td>
</tr>
<tr>
<td>M</td>
<td>[-]</td>
<td>Number of buses</td>
<td>α</td>
<td>[-]</td>
<td>Imbalance cost coeff.</td>
</tr>
<tr>
<td>N</td>
<td>[-]</td>
<td>Set of power nodes</td>
<td>Δ</td>
<td>[-]</td>
<td>Deviation</td>
</tr>
<tr>
<td>N</td>
<td>[-]</td>
<td>Number of power nodes</td>
<td>η</td>
<td>[-]</td>
<td>Efficiency</td>
</tr>
<tr>
<td>P</td>
<td>[MW]</td>
<td>Power</td>
<td>ξ</td>
<td>[MW]</td>
<td>External process</td>
</tr>
<tr>
<td>Q</td>
<td>[EUR]</td>
<td>State penalty matrix</td>
<td>π</td>
<td>EUR</td>
<td>Price</td>
</tr>
<tr>
<td>q</td>
<td>[EUR]</td>
<td>State penalty vector</td>
<td>ω</td>
<td>[rad]</td>
<td>Angular frequency</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subscript</th>
<th>Meaning</th>
<th>Subscript</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ask</td>
<td>Spot market ask price</td>
<td>load</td>
<td>Load, grid consumption</td>
</tr>
<tr>
<td>bid</td>
<td>Spot market bid price</td>
<td>m</td>
<td>Bus index</td>
</tr>
<tr>
<td>cap</td>
<td>Capacity</td>
<td>n</td>
<td>Bus index</td>
</tr>
<tr>
<td>d</td>
<td>Discretized</td>
<td>netinj</td>
<td>Net injection</td>
</tr>
<tr>
<td>db</td>
<td>Dead-band</td>
<td>offset</td>
<td>Constant offset</td>
</tr>
<tr>
<td>del</td>
<td>Delivery time</td>
<td>opt</td>
<td>Optimization horizon</td>
</tr>
<tr>
<td>dev</td>
<td>Deviation</td>
<td>peak</td>
<td>Peak load</td>
</tr>
<tr>
<td>drv</td>
<td>Externally driven</td>
<td>prov</td>
<td>Provided reserves</td>
</tr>
<tr>
<td>endo</td>
<td>Endogenous</td>
<td>p.u.</td>
<td>Per unit</td>
</tr>
<tr>
<td>exch</td>
<td>Exchange</td>
<td>ramp</td>
<td>Ramp-rate</td>
</tr>
<tr>
<td>gen</td>
<td>Generation, grid injection</td>
<td>R-export</td>
<td>Control energy export</td>
</tr>
<tr>
<td>i</td>
<td>Power node index</td>
<td>R-import</td>
<td>Control energy import</td>
</tr>
<tr>
<td>imb</td>
<td>Imbalance</td>
<td>s</td>
<td>Sampling</td>
</tr>
<tr>
<td>lead</td>
<td>Lead time</td>
<td>ss</td>
<td>Steady state</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Superscript</th>
<th>Meaning</th>
<th>Superscript</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ctrl</td>
<td>Control power node</td>
<td>rt</td>
<td>Real-time</td>
</tr>
<tr>
<td>dyn</td>
<td>Dynamic</td>
<td>sch</td>
<td>Day-ahead schedule</td>
</tr>
<tr>
<td>imb</td>
<td>Imbalance</td>
<td>sec</td>
<td>Secondary</td>
</tr>
<tr>
<td>long</td>
<td>Excess energy</td>
<td>short</td>
<td>Lack of energy</td>
</tr>
<tr>
<td>max</td>
<td>Maximum, upper bound</td>
<td>slack</td>
<td>Slack power node</td>
</tr>
<tr>
<td>min</td>
<td>Minimum, lower bound</td>
<td>spot</td>
<td>Spot market</td>
</tr>
<tr>
<td>month</td>
<td>Monthly fee</td>
<td>stat</td>
<td>Static</td>
</tr>
<tr>
<td>MTU</td>
<td>Market Time Unit</td>
<td>tert</td>
<td>Tertiary</td>
</tr>
<tr>
<td>neg</td>
<td>Negative</td>
<td>ts</td>
<td>Time series</td>
</tr>
<tr>
<td>pos</td>
<td>Positive</td>
<td>upd</td>
<td>Intra-day update</td>
</tr>
<tr>
<td>prim</td>
<td>Primary</td>
<td>*</td>
<td>Predicted value</td>
</tr>
<tr>
<td>ref</td>
<td>Reference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rel</td>
<td>Relative</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 6.2 Framework for Multi-Stage Operation

This section is based on [130]. Power node dispatch problems can be formulated either with respect to absolute power and energy quantities, or with respect to a deviation from a previously determined work-
6.2. Framework for Multi-Stage Operation

ing point. Whether absolute or relative coordinates are used is determined by the nature of the considered control application. This section presents a decomposition of the affine power node equation introduced in Section 5.2.2 for different planning and operation stages: day-ahead scheduling, intra-day schedule updates, and control service provision. The economic optimization of a power node portfolio is handled by the former two stages, while the latter stage is particularly relevant for balancing and ancillary services. Note that unit commitment is not addressed here, although it can be modeled by a mixed-integer formulation of a power node dispatch problem. Longer-term power system planning is not of interest for the present work, either.

The three stages have the following properties:

1. **Day-ahead dispatch**: an operating point schedule for the controllable variables; established once a day for the coming day on the basis of operation cost and predictions for uncertain variables,

2. **Intra-day rescheduling**: alteration of the operating point schedule; several updates a day, incorporating shorter-term predictions for uncertain variables, and

3. **Control service provision**: short-term relative changes to the operating point schedules that were determined in the previous two stages. This can be utilized to pursue short-term control goals such as balancing. In the case of security-relevant control reserves requiring guaranteed availability, a control band has to be reserved around the working point trajectory, imposing additional constraints on the day-ahead dispatch and intra-day rescheduling.

The degrees of freedom related to each of the decision and control problems shall be modeled separately. The actual power node model variables are therefore decomposed into three shares consisting of scheduled values (sch), schedule updates as deviations from the scheduled values (upd), thus formulating the real-time (rt) behavior as deviation from the planned baseline:

\[ \mathcal{N} = \mathcal{N}^{\text{sch}} + \Delta \mathcal{N}^{\text{upd}} + \Delta \mathcal{N}^{\text{rt}} \]  

(6.1)

with \( \mathcal{N} = \{u_{\text{gen}}, u_{\text{load}}, \xi, w\} \) being the instantaneous power node variables. The physical storage loss term \( v \) has to be dealt with separately.
6.2.1 Decomposition of Power Node Equation

The decomposition of the power node equation and constraints is based on the analogous decomposition of the storage state variable:

\[ x = x^{sch} + \Delta x^{upd} + \Delta x^{rt}, \]  
\[ \dot{x} = \dot{x}^{sch} + \Delta \dot{x}^{upd} + \Delta \dot{x}^{rt}. \]  

The goal is to formulate separate power node dynamics for each of the shares, such that in superposition they constitute the original power node equation. As condition for superposition, the differential equation has to be linear. This decomposition is thus not applicable for the general case (5.1) but it can be shown to hold for the affine case (5.3). If a coordinate translation \( \hat{x} = x - x_{ss} \) is applied to the affine model (5.3), the result is:

\[ C_i \dot{\hat{x}}_i = \eta_{load,i} u_{load,i} - \eta_{gen,i}^{-1} u_{gen,i} + \xi_i - w_i - a_i \hat{x}_i. \]  

The power node equation is thus linear in \( \hat{\cdot} \)-coordinates, enabling the application of the superposition principle.

For the decomposition of \( \hat{x} \), the offset \( x_{ss} \) can be associated with any of the shares of \( x \) in (6.2). We choose \( \hat{x}^{sch} = x^{sch} - x_{ss} \) and consequently \( \Delta \hat{x}^{upd} = \Delta x^{upd} \) and \( \Delta \hat{x}^{rt} = \Delta x^{rt} \). As a result, the original coordinates can be used for the three related power node formulations:

1. Power node equation for the scheduling problem:

\[ C_i \dot{x}^{sch}_i = \eta_{load,i} u^{sch}_{load,i} - \eta_{gen,i}^{-1} u^{sch}_{gen,i} + \xi^{sch}_i - w^{sch}_i - a_i (x^{sch}_i - x_{ss,i}). \]  

2. Schedule update equation, formulated as a deviation:

\[ C_i \Delta \dot{x}^{upd}_i = \eta_{load,i} \Delta u^{upd}_{load,i} - \eta_{gen,i}^{-1} \Delta u^{upd}_{gen,i} + \Delta \xi^{upd}_i - \Delta w^{upd}_i - a_i \Delta x^{upd}_i. \]  

3. Real-time balancing and control power node dynamics, formulated as the difference between the actual realization and the schedule: \( \Delta \mathbb{N}^{rt} = \mathbb{N}(t) - (\mathbb{N}^{sch} + \Delta \mathbb{N}^{upd}) \):

\[ C_i \Delta \dot{x}^{rt}_i = \eta_{load,i} \Delta u^{rt}_{load,i} - \eta_{gen,i}^{-1} \Delta u^{rt}_{gen,i} + \Delta \xi^{rt}_i - \Delta w^{rt}_i - a_i \Delta x^{rt}_i. \]
6.2. Constraint Coordination and Reserves

The power node constraints (5.1) (a) – (f) have been formulated as ‘physical’ limitations of the unit operation ranges. The multi-stage formulation requires a coordination of constraints between the stages that is compliant with those original power node constraints.

Resources for the real-time control of power systems, e.g., for Load Frequency Control (LFC) provision, are capacities reserved for activation when imbalances occur:

\[
\begin{align*}
- \Delta u_{\text{rt}, \text{neg}}^{\text{gen},i} & \leq \Delta u_{\text{rt}}^{\text{gen},i} \leq \Delta u_{\text{rt}, \text{pos}}^{\text{gen},i}, \\
- \Delta u_{\text{rt}, \text{pos}}^{\text{load},i} & \leq \Delta u_{\text{rt}}^{\text{load},i} \leq \Delta u_{\text{rt}, \text{neg}}^{\text{load},i},
\end{align*}
\]

(6.8) 
(6.9)

where \((\text{rt, pos})\) and \((\text{rt, neg})\) indicate the constraints associated with the provision of positive and negative control reserve, respectively.

Nowadays it is not common in power system operation to deliver control reserves through units with energy constraints relevant on the time scale of the reserve provision. Pumped-hydro power plants, which are naturally energy-constrained by their water reservoir, usually have sufficient storage capacity to securely deliver the contracted control reserves without risk of depletion or overflow of their storage. This can be vastly different in the case of reserve provision by controllable thermal loads, small-scale Combined Heat and Power (CHP) units, or Plug-In Hybrid Electric Vehicles (PHEVs), which have a significantly smaller capacity to store energy in proportion to their power capacity. In these cases, it may be necessary to also reserve a storage control band:

\[\begin{align*}
- \Delta x_{i}^{\text{rt, pos}} & \leq \Delta x_{i}^{\text{rt}} \leq \Delta x_{i}^{\text{rt, neg}},
\end{align*}\]

(6.10)

The nomenclature of \(\Delta x_{i}^{\text{rt, pos}}\) for the lower and \(\Delta x_{i}^{\text{rt, neg}}\) for the upper bound is due to positive and negative reserves being formulated from a grid perspective, whereas \(x\) is formulated from a power node perspective. The implications of reserve provision by energy-constrained generation and load units are summarized in Table 6.2.

| Table 6.2: Directionality of control reserve provision |
|-----------------------------------|--------------------|--------------------|
| **Generation** | **Positive Reserve** | **Negative Reserve** |
| \(\Delta u_{\text{rt}}^{\text{gen},i}\) | \(\rightarrow\) \(\Delta x_{i}^{\text{rt}}\) | \(\Delta u_{\text{rt}}^{\text{gen},i}\) | \(\rightarrow\) \(\Delta x_{i}^{\text{rt}}\) |
| \(\Delta u_{\text{rt}}^{\text{load},i}\) | \(\rightarrow\) \(\Delta x_{i}^{\text{rt}}\) | \(\Delta u_{\text{rt}}^{\text{load},i}\) | \(\rightarrow\) \(\Delta x_{i}^{\text{rt}}\) |
Control reserves are security-critical and are typically procured with considerable lead-time. This requirement of availability calls for the reservation of a control band in the power nodes represented by (5.1) to be taken into account in the day-ahead scheduling stage:

\[(a) \quad \Delta x_{i}^{rt, pos} \leq x_{i}^{sch} \leq 1 - \Delta x_{i}^{rt, neg},\]
\[(b) \quad 0 \leq u_{\text{min}, gen,i} + \Delta u_{\text{gen}, i}^{rt, neg} \leq u_{\text{sch}, gen,i}^{gen} \leq u_{\max, gen,i}^{gen} - \Delta u_{\text{rt}, pos}^{gen,i},\]
\[(c) \quad 0 \leq u_{\text{load}, i}^{\text{min}} + \Delta u_{\text{load}, i}^{rt, pos} \leq u_{\text{sch}, load,i}^{load} \leq u_{\max, load,i}^{load} - \Delta u_{\text{rt}, neg}^{load,i}.\]

For the schedule update, the above absolute constraints are then formulated relative to the pre-planned trajectory:

\[(a) \quad \Delta x_{i}^{rt, pos} - x_{i}^{sch} \leq \Delta x_{i}^{upd} \leq 1 - \Delta x_{i}^{rt, neg} - x_{i}^{sch},\]
\[(b) \quad u_{\text{min}, gen,i}^{gen} + u_{\text{rt}, neg}^{gen,i} - u_{\text{sch}, gen,i}^{gen} \leq u_{\text{upd}, gen,i}^{gen} \leq u_{\max, gen,i}^{gen} - u_{\text{rt}, pos}^{gen,i} - u_{\text{sch}, gen,i}^{gen},\]
\[(c) \quad u_{\text{load}, i}^{\text{min}} + u_{\text{rt}, pos}^{load,i} - u_{\text{sch}, load,i}^{load} \leq u_{\text{upd}, load,i}^{load} \leq u_{\max, load,i}^{load} - u_{\text{rt}, neg}^{load,i} - u_{\text{sch}, load,i}^{load}.\]

The constraints ensure that trajectories scheduled in one stage do not influence the feasibility of trajectories formulated in another stage with respect to the original power node constraints. All other constraints of (5.1) (d) – (f) can be transformed accordingly. Note that the nonlinear constraint (d) can be easily recast as a linear constraint since \(\xi\) is either always smaller or always greater than 0 for most processes. Additional ramping constraints, which are constraints on the time-derivative of \(u_{\text{load}}\) and \(u_{\text{gen}}\), would be formulated entirely analogous to the above constraints. The main parameters of the two stages for day-ahead scheduling and update are:

1. optimization frequency (e.g., every day, every second hour),
2. the sampling time and the length of the look-ahead horizon,
3. available predictions at the time of carrying out the optimization (e.g., 24-hour-ahead wind forecast with 15-minute time resolution), and
4. the time lag between the execution of the optimization and the realization, in liberalized settings given, e.g., by the gate closure time of the energy exchange.
6.3. Model Predictive Control

Traditionally, MPC is used for optimally controlling dynamical systems so that their controlled variables stay close to or follow an externally given setpoint while keeping the control inputs as low as possible. This is achieved by minimizing an objective function with costs attributed to both control error and control inputs. In order to enable a closed-loop-like reaction to disturbances and to plant-model mismatch, the optimization is carried out in a receding horizon fashion with periodic observation (or estimation) of state variables of the plant.

For the dispatch of a power node portfolio, MPC is used in a slightly different sense. Within the Power Nodes Modeling Framework (PNMF), the state of the “plant” is the set of dynamic State of Charge (SOC) variables \( x_i \) of the storage units, and the control inputs are the power quantities \( u_{\text{load},i}, u_{\text{gen},i}, \xi_i, \) and \( w_i \) of all units. Predictions for externally driven in-feeds and loads are included as time-varying equality constraints on the corresponding input variables. Figure 6.1 illustrates the principle of the optimization carried out in a receding horizon.

The goal of the optimization is the minimization of cost incurred by the power node portfolio, either for economically serving the connected loads or in order to fulfill externally given control tasks. In this regard, the goals of the MPC approach coincide with those of traditional power system dispatch. However, it explicitly includes loads and storage de-
vices, and inter-temporal constraints such as ramp-rate limitations and dynamic SOC variables are considered by the predictive optimization.

The use of MPC for power system dispatch has been addressed by other researchers. In [146], several variants of power system dispatch including wind and controllable load are presented. It is argued that the existence of inter-temporal constraints, such as ramping constraints or the presence of energy storage, requires the usage of look-ahead dispatch algorithms, e.g., based on MPC. As a principal limitation, the computational complexity of optimizing large numbers of decision variables over long time horizons is mentioned, and a methodology for overcoming these limitations is formulated. In contrast, [143] argues that through recent advances in convex optimization, large multi-period dispatch problems can be solved efficiently if formulated in a suitable way.

Note that the optimizations in this work are formulated deterministically based on available predictions. A structured integration of the uncertainty into the optimization, e.g., by stochastic Dynamic Programming [158], stochastic MPC [159], or Robust Optimization [160], will increase the realism of the results but is beyond the scope of this work.

6.4 Compact Portfolio Notation

In this section, we formulate model equations for an entire portfolio of power nodes. We will restrict ourselves to the affine formulation of the power node equation as given in (5.3) in Chapter 5. First, the continuous-time equations are formulated in a vector notation with an element-wise multiplication operator. Then the problem is recast in a matrix notation and transformed into a discrete-time dynamical model subject to equality and inequality constraints.

6.4.1 Element-Wise Vector Notation

The portfolio $\mathcal{N}$ of power nodes $i \in \mathcal{N} = \{1, \ldots, N\}$ is defined by the set of differential equations

$$
\mathbf{C} \otimes \dot{x} = \eta_{\text{load}} \otimes u_{\text{load}} - \eta_{\text{gen}}^{-1} \otimes u_{\text{gen}} + \xi - w - a \otimes (x - x_{\text{ss}}), \tag{6.11}
$$

where $\otimes$ represents the element-wise multiplication of two vectors of equal dimension. Note that (6.11) also contains the static power nodes.
6.4. Compact Portfolio Notation

Let \( i_{stat} \in \{i \mid i \in N \land C_i = 0\} \). For implementation, it is practical to separate the static power nodes and to include them in the optimization as a vectorized equality constraint. Equation (6.11) has the initial condition \( x_0 = x(t = t_0) \) and is subject to the following constraints:

\[
\begin{align*}
0 \leq x_{\text{min}} &\leq x \leq x_{\text{max}} \leq 1, \\
0 \leq u_{\text{min}} &\leq u_{\text{gen}} \leq u_{\text{max}} \leq 1, \\
0 \leq u_{\text{min}} &\leq u_{\text{load}} \leq u_{\text{max}} \leq 1, \\
\dot{u}_{\text{min}} &\leq \dot{u}_{\text{gen}} \leq \dot{u}_{\text{max}}, \\
\dot{u}_{\text{min}} &\leq \dot{u}_{\text{load}} \leq \dot{u}_{\text{max}}, \\
\xi_i = \xi_{\text{drv},i}(t) \quad \forall \quad i \in \{i \mid i \in N \land i \text{ externally driven}\}.
\end{align*}
\]

The additional constraints set forth in (5.1), \( 0 \leq \xi \cdot w, 0 \leq |\xi| - |w|, \) and \( 0 \leq v \) are omitted here since these requirements on the power nodes can be ensured implicitly by appropriate constraints on \( u_{\text{gen}}, u_{\text{load}}, \) and \( \xi \) for realistic units (cf. Sections 5.3.1 – 5.3.3).

6.4.2 Matrix Notation

In order to make the notation more compact, we define the state and input variable vectors for the dynamic state variables \( x_{\text{dyn}} \) with \( i_{\text{dyn}} \in \{i \mid i \in N \land C_i > 0\} = 1, \ldots, N_{\text{dyn}} \), and \( u_{\text{gen},i}, u_{\text{load},i}, \xi_i, \) and \( w_i \) with \( i \in N = 1, \ldots, N \):

\[
\begin{align*}
x &= [x_1, \ldots, x_{N_{\text{dyn}}}]^T, \\
u &= [u_{\text{gen},1}, \ldots, u_{\text{gen},N}, u_{\text{load},1}, \ldots, u_{\text{load},N}, \\
\xi_1, \ldots, \xi_N, w_1, \ldots, w_N]^T. 
\end{align*}
\]

For distinguishing between the static power nodes \( (C_i = 0) \) and the ones possessing inherent storage \( (C_i > 0) \), we introduce the vectors \( u_{\text{dyn}} \) and \( u_{\text{stat}} \):

\[
\begin{align*}
u_{\text{dyn}} &= [u_{\text{gen},1}^{\text{dyn}}, \ldots, u_{\text{gen},N_{\text{dyn}}}^{\text{dyn}}, u_{\text{load},1}^{\text{dyn}}, \ldots, u_{\text{load},N_{\text{dyn}}}^{\text{dyn}}, \\
\xi_1^{\text{dyn}}, \ldots, \xi_{N_{\text{dyn}}}^{\text{dyn}}, w_1^{\text{dyn}}, \ldots, w_{N_{\text{dyn}}}^{\text{dyn}}]^T, \\
u_{\text{stat}} &= [u_{\text{gen},1}^{\text{stat}}, \ldots, u_{\text{gen},N_{\text{stat}}}^{\text{stat}}, u_{\text{load},1}^{\text{stat}}, \ldots, u_{\text{load},N_{\text{stat}}}^{\text{stat}}, \\
\xi_1^{\text{stat}}, \ldots, \xi_{N_{\text{stat}}}^{\text{stat}}, w_1^{\text{stat}}, \ldots, w_{N_{\text{stat}}}^{\text{stat}}]^T.
\end{align*}
\]

Using this nomenclature, the dynamical system describing the power node portfolio can be formulated:

\[
\dot{x} = A(x - x_{ss}) + B_{\text{dyn}} u_{\text{dyn}}
\]
with the initial condition $\mathbf{x}_0 = \mathbf{x}(t = t_0)$ and subject to the constraints

$$B^\text{stat} \mathbf{u}^\text{stat} = 0$$ (6.18)

as well as the rest of the constraints from (6.12).

The matrices $\mathbf{A}$, $B^\text{stat}$, and $B^\text{dyn}$ contain the appropriate coefficients such that the power node system is accurately represented by (6.17) and (6.18).

### 6.4.3 Discretization

The dynamic model from (6.17) is now discretized (denoted by subscript $d$) for the use in MPC. We obtain the discrete-time linear system

$$\mathbf{x}(k + 1) = \mathbf{A}_d \mathbf{x}(k) + \mathbf{B}_d^\text{dyn} \mathbf{u}^\text{dyn}(k) + \mathbf{x}_\text{offset}$$ (6.19)

with the initial condition $\mathbf{x}_0 = \mathbf{x}(k = 0)$, the constant offset $\mathbf{x}_\text{offset}$ (which comes from the discretization of the term $\mathbf{A}\mathbf{x}_\text{ss}$), and subject to the constraints

$$0 \leq \mathbf{x}^\text{min} \leq \mathbf{x}(k) \leq \mathbf{x}^\text{max} \leq 1 ,$$ (6.20)

$$0 \leq \mathbf{u}^\text{min} \leq \mathbf{u}(k) \leq \mathbf{u}^\text{max} ,$$ (6.21)

$$\mathbf{B}_d^\text{stat} \mathbf{u}^\text{stat}(k) = 0 ,$$ (6.22)

$$\mathbf{u}_\text{drv}(k) = \mathbf{u}_\text{ts}^\text{drv}(k) ,$$ (6.23)

where the profile variables $\xi_i$ enter as time series in $\mathbf{u}_\text{ts}^\text{DRV}(k)$. The ramping constraint

$$\dot{\mathbf{u}}^\text{min} \leq \dot{\mathbf{u}} \leq \dot{\mathbf{u}}^\text{max}$$ (6.24)

is reformulated by using $\dot{\mathbf{u}} = \lim_{t \to 0} \frac{\mathbf{u}(t) - \mathbf{u}(t-h)}{h}$ to a discrete approximation:

$$h \dot{\mathbf{u}} \approx \mathbf{u}(k) - \mathbf{u}(k-1) =: \delta \mathbf{u}(k) ,$$ (6.25)

which yields

$$\delta \mathbf{u}^\text{min} \leq \delta \mathbf{u}(k) \leq \delta \mathbf{u}^\text{max}$$ (6.26)

with $\delta \mathbf{u}^\text{min} = h \dot{\mathbf{u}}^\text{min}$ and $\delta \mathbf{u}^\text{max} = h \dot{\mathbf{u}}^\text{max}$. This formulation of the power node system is amenable to the use in control frameworks. Up to now, the power nodes are still completely decoupled, i.e., they can be controlled independently of each other. An additional set of constraints is needed in order to formulate a meaningful power system control problem: the power balance in the grid, represented by power flow equations.
### 6.4. Compact Portfolio Notation

#### 6.4.4 Grid Model Integration

Consider a power grid composed of $M$ buses denoted by $m, n \in \mathcal{M} = \{1, \ldots, M\}$ and a set of $N$ power nodes $i \in \mathcal{N} = \{1, \ldots, N\}$, representing a number of single or aggregated units. The mapping can be formulated by index sets $\mathcal{N} \to \mathcal{M}$. The power node indices are divided into sets $\mathcal{N}_m \subseteq \mathcal{N}$ associated with one bus each; the following properties hold for $\mathcal{N}_m$: $\mathcal{N}_m \cap \mathcal{N}_n = \emptyset$ for $m \neq n$, and $\bigcup_{m \in \mathcal{M}} \mathcal{N}_m = \mathcal{N}$.

The net power injection to a grid node $m \in \mathcal{M}$ is thus:

$$P_{\text{netinj},m} = \sum_{i \in \mathcal{N}_m} u_{\text{gen},i} - \sum_{i \in \mathcal{N}_m} u_{\text{load},i} \quad . \quad (6.27)$$

In general, the power systems literature offers many options to model a power system, depending on the questions of relevance to the study. In principle, the power node domain can be interfaced with many grid model types, such as Direct Current (DC) or Alternating Current (AC) power flow, static or dynamic grid models. This is due to the clear separation from the electro-mechanical domain.\footnote{In most cases it is appropriate to model the power-exchange $u_{\text{gen/load}}$ as a power injection to the respective bus. In case of a dynamical grid model and the power node being a synchronous machine, the proper interface would be the mechanical power exerted on its shaft.}

To illustrate the approach, this section formulates a network represented by linear DC power flow equations. The DC network representation is used, e.g., in an active-power dispatch of a unit portfolio in a capacity-constrained transmission system. The DC power flow assumes small angle differences, a constant, flat voltage profile, and neglects the resistance of lines. While voltage angles are generally small, the critical assumptions are the flat voltage profile and the negligible resistance [161].

The power flow is governed by the following equations:

$$P_{\text{exch},m} = \sum_{\substack{n \in \mathcal{M} \setminus \{m\}}} B_{mn} (\delta_m - \delta_n) \quad , \quad (6.28)$$

$$0 = \sum_{m=1}^{M} (P_{\text{netinj},m} - P_{\text{exch},m}) \quad , \quad (6.29)$$

where $\delta_m$ is the voltage angle at bus $m$ and $B_{mn} = 1/X_{mn}$ is the inverse of the line reactance.
The line flows may be subject to capacity constraints $P^\text{min}_{mn} < 0$ and $P^\text{max}_{mn} > 0$:

$$P^\text{min}_{mn} \leq B_{mn}(\delta_m - \delta_n) \leq P^\text{max}_{mn}. \quad (6.30)$$

The system frequency can be described by an aggregate inertia model:

$$H \dot{\omega} = \sum_{m=1}^{M} P_{\text{netinj},m}, \quad (6.31)$$

where $H$ [s] is the aggregate inertia constant and $\omega$ [rad] is the angular frequency of the system.

The integration of the power nodes into AC power flow and AC Optimal Power Flow (OPF) methods is not so straightforward due to the non-linearity of the power flow equations. This problem is studied in [162].

### 6.4.5 Endogenous Cost Function

In order to formulate optimization strategies for the power node portfolio, a cost function which penalizes the decision variables of the problem according to certain control objectives has to be conceived. The penalty terms should – as much as possible – be based upon real monetary costs in order to yield an economically optimal operation.

The cost incurred by the operator of the power node portfolio depends upon two main components: the cost of operating the units themselves (endogenous cost), and the cost or revenue based on other factors such as energy sales, schedule imbalances, or ancillary service provision (exogenous cost). Note that revenues are denoted as negative costs. In this section, we formulate the endogenous cost. The exogenous cost and revenue terms will be defined in Section 6.5.

For the scheduling in absolute quantities, we consider the following cost function for an individual time step $k$:

$$J_{\text{endo}}(k) = (x(k) - x^\text{ref})^T Q (x(k) - x^\text{ref}) + q^T (x(k) - x^\text{ref})$$
$$+ (u(k) - u^\text{ref})^T R (u(k) - u^\text{ref}) + r^T (u(k) - u^\text{ref})$$
$$+ \delta u^T(k) \delta R \delta u(k), \quad (6.32)$$

where $x^\text{ref}$ and $u^\text{ref}$ are reference values for state and input variable vectors, $Q$, $R$, and $\delta R$ are quadratic penalty matrices of appropriate size, and $q$ and $r$ are penalty vectors. The individual terms in the cost
function are as follows: The first line penalizes a deviation of the state from a desired target value. Penalizing state deviation is only meaningful in cases when actual financial costs are incurred by the deviation, or when the state shall be kept in the vicinity of a certain level, e.g., in order to reduce the risk of a storage depletion or overflow. The second line penalizes all power quantities except for the physical loss term \( v \). This includes generator cost terms (linear and/or quadratic) for fuel and Operation & Maintenance (O&M) costs, and penalties for curtailments of load and generation (the latter is relevant when actual compensation payments have to be made, e.g., for Renewable Energy Sources (RES) curtailments). The last line represents the ramping cost incurred by working point changes. This is particularly relevant for thermal power plants where thermal stress is an important factor for unit lifetime.

When the optimization problem is formulated in relative quantities (as in the “schedule update” and “real-time control” stages as defined in Section 6.2), the variable substitutions \( x = x^{sch} + \Delta x \) and \( u = u^{sch} + \Delta u \) are done. As the scheduling stage is calculated first without any influence of update and real-time stages, \( x = x^{sch} \) and \( u = u^{sch} \) can be assumed for the schedule stage.

The endogenous cost function for the relative model formulation is given in the following:

\[
J^\text{rel\ endo}(k) = \Delta x^T(k)Q \Delta x(k) + (2(x^{sch}(k) - x^{ref})^TQ + q^T)\Delta x(k) + \Delta u^T(k)R \Delta u(k) + (2(u^{sch}(k) - u^{ref})^TR + r^T)\Delta u(k) + \Delta \delta u^T(k) \delta R \Delta \delta u(k) + 2 \delta u^{sch}(k) \delta R \Delta \delta u(k) .
\] (6.33)

For shortness of notation, \( x^{sch} + \Delta x^{upd} \) can be called the (updated) \( x^{sch} \) when the real-time part is considered. The same holds for \( u \).

### 6.4.6 Definition of Auxiliary Power Nodes

For the formulation of certain control objectives, it is useful to define two specific types of power nodes, which we will refer to as “Control Power Node” and “Slack Power Node”. The Control Power Node serves to model the effect of external control signals that influence the power node portfolio, while the Slack Power Node provides a representation of a power source or sink that is not part of the considered unit portfolio but still to be included in the control problem. Both Control and Slack Power Nodes do not possess inherent storage (\( C^{ctrl} = 0, C^{slack} = 0 \)).
Chapter 6. Dispatch Strategies

The equations defining a Control and a Slack Power Node are

\[ \xi^{\text{ctrl}} = u^{\text{ctrl}}_{\text{gen}} - u^{\text{ctrl}}_{\text{load}}, \quad (6.34) \]

\[ \xi^{\text{slack}} = (\eta^{\text{slack}}_{\text{gen}})^{-1} u^{\text{slack}}_{\text{gen}} - \eta^{\text{slack}}_{\text{load}} u^{\text{slack}}_{\text{load}}. \quad (6.35) \]

As the Slack Power Node represents external physical generation and consumption, the efficiencies are included in the equation. It can be subject to power and ramping constraints. The Control Power Node, however, is a purely virtual entity and thus possesses neither efficiencies nor constraints. It is driven by the external process

\[ \xi^{\text{ctrl}} = \xi^{\text{ctrl,ts}}_{\text{drv}}(k), \quad (6.36) \]

where \( \xi^{\text{ctrl,ts}}_{\text{drv}}(k) \) represents the time series of the external control signal.

As we will see in the remainder of this chapter, the Control Power Node is of most practical use in the relative formulation:

\[ \Delta \xi^{\text{ctrl}} = \Delta u^{\text{ctrl}}_{\text{gen}} - \Delta u^{\text{ctrl}}_{\text{load}}, \quad (6.37) \]

\[ \Delta \xi^{\text{ctrl}} = \Delta \xi^{\text{ctrl,ts}}_{\text{drv}}(k). \quad (6.38) \]

The auxiliary power nodes enter the power node portfolio as additional optimization variables subject to the equality constraints from (6.34) – (6.38). They are purposely not included in the system formulation in Section 6.4.2 since they will be used separately in the optimization problem. As the auxiliary power nodes are purely algebraic, they do not possess dynamics to be discretized.

6.5 Use Cases and Control Problems

The control of power node portfolios for economic objectives or the provision of power system control services can now be formulated using the power node decomposition mentioned above. We build upon the experience gained with the usage of MPC strategies for control services in [64, 163] before the development of the Power Node concept and generalize the ideas to a comprehensive use case taxonomy for flexible portfolios in power systems. The following cases are considered:

1. Least-Cost Economic Dispatch,
2. Market-Based VPP Operation,
3. Balancing of Schedule Deviations,
4. Provision of Frequency Control Reserves,
5. Capacity Firming of Intermittent Generation,
6. Peak Shaving, and
7. Residual Load Ramp-Rate Reduction.

All control cases will be formulated in an MPC strategy based on the predicted values of externally driven variables. For electric load forecasts, a wide range of established methods exists [164] and advances in the development of wind and solar power forecasts have been made [165, 166]. In this work, we restrict ourselves to using realized time series of predictions and actual values for the externally driven variables. We simulate the scenarios using a perfect forecast. This is foremost due to the lack of periodically updated forecast data, but also because we focus on the formulation of the optimization problem for the considered control case. Forecast errors can be incorporated by replacing the perfect-prediction time series with the forecast time series available at the respective instant in time.

6.5.1 Least-Cost Economic Dispatch

Background
The economic dispatch of a power system establishes a benchmark for operating the system at the lowest possible cost, in its most basic form without considering the effect of network losses and/or constraints. According to the traditional power system operation paradigm, generators are the only controllable resources in the system. Thus, standard economic dispatch problems are formulated in terms of minimizing generation cost for serving a given load demand. The textbook definition of economic dispatch, e.g., reported by [167], is as follows:

$$\min \sum_{\text{gen}_i} C_i(P_{\text{gen},i}) \quad (6.39)$$

subject to the constraints

$$P_{\text{min},i} \leq P_{\text{gen},i} \leq P_{\text{max},i}, \quad (6.40)$$
$$\sum_{\text{gen}_i} P_{\text{gen},i} = P_{\text{load}} \quad , \quad (6.41)$$

where $P_{\text{gen},i}$ is the power output of generator $i$, $P_{\text{min},i}$ and $P_{\text{max},i}$ are generator $i$’s output limits, $P_{\text{load}}$ is the total system demand, and $C_i(P_{\text{gen},i})$ is the marginal cost of energy production of generator $i$. 
Chapter 6. Dispatch Strategies

Figure 6.2: Power node setup for least-cost dispatch including grid

If the electricity network model (i.e., the power flow equations) is included in the problem, it becomes an OPF optimization. The power flow can be formulated as a full AC power flow or in approximate forms such as the frequently used DC approximation (see Section 6.4.4).

Figure 6.2 shows the power node setup for the least-cost dispatch including the grid constraints. The optimization takes into account the endogenous cost terms for all units represented by power nodes as well as all constraints on the individual power nodes. If the grid topology is considered as presented in Section 6.4.4, the power flow equations as well as line limits enter the problem as additional constraints (linear in case of DC power flow equations). These parameters, denoted by dashed arrows in Figure 6.2, are constant over different optimization periods. Denoted by solid arrows are the quantities that change from step to step: state information and predictions are passed from the power node portfolio to the optimizer, which then determines setpoints for the power nodes.

**Day-Ahead Dispatch**

The day-ahead optimization usually takes place in the afternoon of the day preceding the actual operation. One day in advance, the prediction of the occurring load and fluctuating RES in-feed is still rather inaccurate. For instance, the wind prediction Root Mean Square Er-
ror (RMSE) can be as large as 20 – 30% depending on the forecast method [168]. This highlights the importance of intra-day adjustments in the case of large RES penetrations.

**Cost Function** Since the objective of the dispatch is cost minimization and the power node portfolio comprises the entire system, we only need to consider the endogenous cost of the power node portfolio. The cost function in time step $k$ is thus:

$$J(k) = \sum_{l=k}^{k+N_{opt}-1} J_{\text{endo}}^*(l),$$

(6.42)

where $J_{\text{endo}}^*(l)$ is the cost function using the predicted values for $x$ and $u$, denoted by $x^*$ and $u^*$.

**Intra-Day Update**

Intra-day predictions are usually much more accurate and have a shorter forecast horizon, e.g., 4 or 12 hours. Due to computation times and market gate closure, there is also a time lag between prediction and execution of, e.g., one hour. This leads to a reduction in accuracy of the forecast since it is already one hour old when it is executed, and a current measurement of the real values cannot be employed directly for the upcoming period.

**Cost Function** Since the update is performed in relative coordinates, the endogenous cost function also needs to be considered in relative coordinates: $J_{\text{endo}}^{rel}$. Since the day-ahead dispatch usually takes into account larger time horizons, the determination of optimal storage levels based on the day-ahead dispatch is also relevant for the intra-day dispatch update. In order to tie the update to the day-ahead dispatch, the deviation of the storage level $\Delta x^*$ from the day-ahead schedule is penalized quadratically. This reflects the cost for deviations of the storage levels from their previously determined optimal trajectory as well as transaction costs due to reduced intra-day liquidity and exchange commissions. Note that also linear (absolute) penalties may be of practical relevance. The cost function for the dispatch update reads:

$$J(k) = \sum_{l=k}^{k+N_{opt}-1} \left( J_{\text{endo}}^{rel}(l) + \Delta x^* \Delta x^* C_{\text{dev}} \Delta x^*(l) \right),$$

(6.43)

where $C_{\text{dev}}$ is the deviation penalty matrix of appropriate size.
6.5.2 Market-Based VPP Operation

The dispatch presented in the previous section concerns the system-level cost minimization. This is a practically applicable problem in a vertically integrated utility structure and an approximation of the outcomes of merit-order dispatch in a perfectly competitive market environment. In contrast, the operator of a small VPP in a larger electricity market does not seek to minimize his cost of operation but to maximize his profit. This can involve utilizing cheap generators preferentially and managing the demand, but also trading energy on the market. In the present setting, only trading on the spot market is considered.

Assuming that the VPP operator is a price taker, day-ahead predictions of the spot market price can be used to schedule exchanges with the market. The spot price of the upcoming day can be predicted by various methods, e.g., auto-regressive time-series modeling techniques [169]. The VPP operator thus seeks to minimize a cost function, which can assume negative values to represent profits, by dispatching the units in an appropriate way.

The power node portfolio is defined as above in Section 6.4 with an endogenous cost function defined in Section 6.4.5. Additionally, a Slack
6.5. Use Cases and Control Problems

Power Node as in Section 6.4.6 is defined which models the import and export to and from the considered portfolio. It is described by

\[
(\eta_{\text{gen}}^{\text{spot}} - 1) u_{\text{gen}}^{\text{spot}} - u_{\text{load}}^{\text{spot}} = \xi_{\text{spot}}. \tag{6.44}
\]

The efficiency \( \eta_{\text{gen}}^{\text{spot}} \) models the average efficiency of the power generation outside the considered power node portfolio, which may be useful for an accounting of primary energy usage or Greenhouse Gas (GHG) emissions. Figure 6.3 shows the setup for this use case.

The cost function taking into account the import and export to and from the power node portfolio is as follows:

\[
J(k) = \sum_{l=k}^{k+N_{\text{opt}}-1} \left( \pi_{\text{bid}}^{\star}(l) u_{\text{load}}^{\text{spot} \ast}(l) - \pi_{\text{ask}}^{\star}(l) u_{\text{gen}}^{\text{spot} \ast}(l) \right) + \sum_{l=k}^{k+N_{\text{opt}}-1} J_{\text{endo}}^{\ast}(l),
\]

where \( \pi_{\text{bid}}^{\star}(l) \) is the predicted spot price for energy sales by the VPP and \( \pi_{\text{ask}}^{\star}(l) \) is the predicted spot price for purchases by the VPP in time period \( l \). This formulation allows the approximate modeling of an average bid-ask spread in the market. From an optimization perspective, the spread serves to penalize the trading actions in order to avoid optimization artifacts such as simultaneous sales and purchases.

6.5.3 Balancing of Schedule Deviations

For a utility that serves loads and/or owns generation assets, the task of energy balancing on distinctive time slices, which are usually called Market Time Units (MTUs), is an important aspect of everyday operation. Usually the market players are organized in balancing entities, which have different names depending on the country and are called Balance Groups in Switzerland. From a system perspective, the main goal of the balancing regime is to incentivize an accurate following of the exchange schedules submitted to the Transmission System Operator (TSO), day-ahead or intra-day, by the market participants. In order to achieve this, financial penalties are imposed for differences between actual and scheduled energy exchanges.

The cost incurred by market participants for the imbalances is strongly dependent upon the market design. Here, the Swiss balancing regime
Chapter 6. Dispatch Strategies

Figure 6.4: Swiss energy balancing regime

[170] is considered. It consists of a penalty for long and short energy imbalances on 15-minute time slices. The penalization is different depending on whether the entire control area is long or short, and thus, whether the balance group has a stabilizing or destabilizing effect on the overall system.

Figure 6.4 shows the scheme for penalty calculation. In case of destabilizing behavior, the balancing energy price is based on the price for positive and negative control energy ($\pi_{R\text{-import}}$ and $\pi_{R\text{-export}}$), while a stabilizing balance group is billed with the (more favorable) spot market price $\pi_{\text{spot}}$. The penalty factors $\alpha_1, \ldots, \alpha_4$ further define the imbalance penalty regime. In 2012, they were chosen as follows: $\alpha_1 = \alpha_4 = 1.3$, $\alpha_2 = \alpha_3 = 0.7$. This implies that, besides the control energy billing for destabilizing behavior, another 30% of penalty is incurred by the balance group for both excess and deficit of energy.

The penalty regime is mathematically modeled in a cost function by comparing the balancing energy price with the current spot price. The rationale behind this is that the need to consume or inject balancing energy could have been avoided by more appropriate scheduling of energy purchases and sales on the spot market. We formulate the imbalance cost function for the current MTU as follows:

$$J_{\text{imb}}^{\text{MTU}} = Cost_{\text{imb}}^{\text{MTU}} - Revenue_{\text{imb}}^{\text{MTU}}$$

with the cost and revenue terms

$$Cost_{\text{imb}}^{\text{MTU}} = \left( \pi_{R\text{-import}} \alpha_1 - \pi_{\text{spot}} \right) \left( 1 - m \right) E_{\text{imb}}^{\text{short}}$$
$$+ \pi_{\text{spot}} \left( \alpha_4 - 1 \right) m E_{\text{imb}}^{\text{short}},$$

$$Revenue_{\text{imb}}^{\text{MTU}} = \left( \pi_{R\text{-export}} \alpha_3 - \pi_{\text{spot}} \right) m E_{\text{imb}}^{\text{long}}$$
$$+ \pi_{\text{spot}} \left( \alpha_2 - 1 \right) \left( 1 - m \right) E_{\text{imb}}^{\text{long}},$$
where $m = 1$ holds when the control area is long and $m = 0$ holds when the control area is short. As it is not known in advance whether the control area is long or short, $m = 0.5$ is chosen to reflect a 50% chance that the balance group is stabilizing or destabilizing. $E_{\text{imb}}^{\text{long}}$ is the excess energy in the considered MTU and $E_{\text{imb}}^{\text{short}}$ is the energy deficit. By reformulating, we derive:

$$J_{\text{imb}}^{\text{MTU}} = \pi_{\text{imb}}^{\text{short}} E_{\text{imb}}^{\text{short}} - \pi_{\text{imb}}^{\text{long}} E_{\text{imb}}^{\text{long}}$$

(6.49)

with the short and long penalty coefficients (here already with $m = 0.5$ inserted):

$$\pi_{\text{imb}}^{\text{short}} = \frac{1}{2} \pi_{\text{R-import}} \alpha_1 + \pi_{\text{spot}} (0.5 \alpha_4 - 1)$$

(6.50)

$$\pi_{\text{imb}}^{\text{long}} = \frac{1}{2} \pi_{\text{R-export}} \alpha_3 + \pi_{\text{spot}} (0.5 \alpha_2 - 1)$$

(6.51)

Figures 6.5 and 6.6 show two power node setups for different balancing tasks. Figure 6.5 depicts the self-balancing of a power node portfolio that contains non-controllable or only partly controllable units which introduce an uncertainty. The imbalance arises within the power node portfolio and is also compensated there. The second case, depicted in Figure 6.6, shows the balancing of a deviation external to the portfolio. In this case, the imbalance is imposed through a time series from a Control Power Node as presented in Section 6.4.6 (in this case called Imbalance Power Node in order to avoid confusion). The remaining power nodes then act so as to compensate the externally imposed imbalance. It shall be noted that both formulations yield equivalent results, but one representation may be more convenient than the other depending on the considered system. We will only consider the “external imbalance” case here, i.e., the case depicted in Figure 6.6.

A further distinction is made between imposing only one control action at the beginning of an MTU (“Pre-MTU Balancing”) and adapting generation and consumption in smaller time intervals within the MTU (“Intra-MTU Balancing”). We will only formulate the “Pre-MTU Balancing” strategy in detail and comment on some aspects of the “Intra-MTU Balancing” strategy.

**Pre-MTU Balancing**

The first balancing setup consists of an adaptation of active power set-points in the very beginning of a Market Time Unit. A very-short-term prediction of the production and consumption in the MTU is needed,
Chapter 6. Dispatch Strategies

Figure 6.5: Self-balancing of a power node portfolio

Figure 6.6: Balancing of external deviations by power node portfolio
which can be delivered either by a simple assumption of persistence of a current measurement or more sophisticated very-short-term prediction methods for load \cite{171} and renewable energy in-feeds \cite{172}. Power setpoints are then adjusted accordingly for the entire duration of the MTU. This requires an execution of the optimization problem once per MTU and the entire problem may be formulated in a sampling time equal to the MTU time step (in the present case 15 min).

We impose the imbalance time series through the Imbalance Power Node given in $\Delta$-formulation:

$$\Delta u_{\text{imb \ load}}^\text{imb}(k) = - \Delta \xi_{\text{imb}}^\text{imb}(k) = \frac{E_{\text{dev}}(k)}{t_s} \quad ,$$

(6.52)

where $E_{\text{dev}}(k) \,[\text{MWh}]$ is the time series of deviations of the energy exchanges per MTU. A Slack Power Node serves to model the actual exchange of balancing power with the system. In order to enable a different penalization of consumption and in-feed of balancing power, separate variables are used:

$$\Delta u_{\text{slack \ gen}}^\text{slack} - \Delta u_{\text{slack \ load}}^\text{slack} = \Delta \xi_{\text{slack}}^\text{slack} \quad (6.53)$$

subject to the constraints

$$0 \leq \Delta u_{\text{slack \ gen}}^\text{slack} \quad , \quad 0 \leq \Delta u_{\text{slack \ load}}^\text{slack} \quad .$$

(6.54)

The power balance of the system holds for all $N$ power nodes of the portfolio plus the slack and imbalance power node ($N + 2$ power nodes in total). It is formulated in $\Delta$-quantities as well:

$$\sum_{i=1}^{N+2} (\Delta u_{\text{gen \, i}}^* (l) - \Delta u_{\text{load \, i}}^* (l)) = 0 \quad .$$

(6.55)

The cost function to be minimized by the balancing controller is created by comparing the cost incurred by balancing energy import or export with the actual spot market price at that moment ("mark-to-market"):

$$J(k) = \sum_{l=k}^{k+N_{\text{opt}}-1} J_{\text{endo}}^\text{rel \,}^* (l)$$

$$+ \sum_{l=k}^{k+N_{\text{opt}}-1} \left( \pi_{\text{short}}^* (l) - \pi_{\text{spot}}^* (l) \right) t_s u_{\text{gen \, i}}^\text{slack \ gen \,}^* (l)$$

(6.56)

$$- \sum_{l=k}^{k+N_{\text{opt}}-1} \left( \pi_{\text{long}}^* (l) - \pi_{\text{spot}}^* (l) \right) t_s u_{\text{load \, i}}^\text{slack \ load \,}^* (l) \quad .$$
**Intra-MTU Balancing**

A slightly different approach to the balancing problem is the compensation of arising deviations *within* the Market Time Unit. This requires a shorter sampling time than the length of the MTU. A predictive optimization can be used as well, but in this case the optimization horizon can be curtailed at the end of the MTU.

Although this strategy may be attractive for a Balance Group due to the possibility to minimize its balancing energy cost, we do not formulate it here in detail. The reason is a significant drawback of the strategy for the system level: since the optimization is based on energy values in a fixed time interval, the instantaneous power quantities can vary significantly from the average over the MTU. An optimal balancing of the *energy* exchanges by the balance group may therefore lead to short-term *power* imbalances that need to be compensated by frequency control structures in the control area, e.g., by the secondary frequency controller. This effect has been mentioned in [131, 173] and is clearly undesirable since cost savings on the side of the balance group lead to cost increases for the overall system. Thus, regulatory boundaries should be imposed on such strategies, e.g., by limiting the power deviation from the average of the MTU.

If this strategy were to be formulated in the present framework, the essential input for the predictive optimization would be a very-short-term forecast of consumption and production over the remaining duration of the current time slice. Along with the already realized values of these variables (that are assumed to be measured), this would enable the optimizer to compare the predicted total energy with its target and react such that the difference is minimized.

**Imbalance Trajectory Creation**

Now we briefly outline a simple approach for generating a test trajectory for balancing evaluation. We follow an approach similar to [174]. It consists of a synthesis of the stochastic power deviation time series from a normal distribution with zero mean and a certain standard deviation, described as a percentage of the system peak load. In order to introduce a correlation between consecutive time steps $k - 1$ and $k$ (which is an observable property of real imbalance time series), a *window* of a certain width [0,1] around the probability of $\Delta P_{\text{dev}}(k - 1)$ in the cumulative distribution function is used to impose a restriction on the current value $\Delta P_{\text{dev}}(k)$. The power deviation is sampled on a smaller time frame (one
Table 6.3: Parameters used for imbalance creation

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imbalance mean</td>
<td>0</td>
</tr>
<tr>
<td>Imbalance standard deviation</td>
<td>2% of peak load</td>
</tr>
<tr>
<td>Maximum power deviation</td>
<td>5% of peak load</td>
</tr>
<tr>
<td>Window width</td>
<td>0.3</td>
</tr>
<tr>
<td>Sampling time for power deviation</td>
<td>1 min</td>
</tr>
<tr>
<td>Sampling time for Market Time Unit</td>
<td>15 min</td>
</tr>
</tbody>
</table>

Figure 6.7: Exemplary imbalance trajectory (peak load of 10 MW)

minute or less) and the energy deviation is computed from the power time series. Figure 6.7 shows an exemplary synthesized time series and Table 6.3 shows the used parameters.

6.5.4 Provision of Frequency Control Reserves

Frequency control in power systems serves to maintain a stable system frequency in the presence of load variations, power plant outages, and intermittent generation. According to the classification used by the European Network of Transmission System Operators for Electricity (ENTSO-E) [175], frequency control comprises primary, secondary, and tertiary control reserves contracted by the TSO.

As illustrated in [176], the properties of control reserves can vary significantly from country to country, even within a common synchronous
area. Therefore, the common guidelines for the Continental Europe (CE) system (ENTSO-E CE, corresponding to the area of the former Union for the Co-ordination of Transmission of Electricity (UCTE)) set forth in [175] (Policy P1 and Appendix A1: LFC and Performance) are used here as a basis. All kinds of control reserves (primary, secondary, and tertiary) are tendered by the TSO in individually specified quantities. These are usually determined by probabilistic considerations [177] and are dependent upon the size and generation portfolio of the control area. They consist of reserve products for both positive (generation increase or load decrease) and negative (generation decrease or load increase) reserves. Figure 6.8 depicts roughly the relation between the activation and delivery times for the three kinds of control reserves.

Primary control reserves are activated by a decentralized proportional controller within the speed governors of generating units. In Continental Europe, the primary reserve has to be fully activated after 30 s and has to be sustained for a maximum of 15 min after a larger disturbance. However, it has to be available continuously and at all times. The droop must be set such that the reserve is fully activated at a frequency deviation of 200 mHz. Secondary control reserves are activated by a Proportional-Integral (PI) controller usually operated by the TSO. These reserves are mainly used to relieve primary control, to bring the system frequency back to its nominal value, and to ensure the maintenance of the scheduled tie-line exchanges with other control areas. The secondary control signal transmitted by the TSO to the providing units in its control zone is dependent on the Area Control Error (ACE), which should be controlled to zero. In principle, secondary reserve capacity can be offered asymmetrically (meaning, a different scaling of positive and negative reserves), although this is used in few countries.
Tertiary control is a manually activated reserve which is used to relieve the secondary control reserves. It has to be fully activated after 15 min.

The advantage of providing frequency control reserves by a portfolio of flexible units consists of the ability to take advantage of the different unit properties. The splitting of control signals among generators and storage devices has been discussed before. For instance, in [178] it is proposed to split the control signal by using frequency filters. This allows to distribute the high-frequency components to the storage and the low-frequency components to the generation units. The approach is implementable with relative ease and transparent with respect to its effects. However, it lacks the possibility to base the splitting directly on economic criteria, conduct the splitting optimally in that sense, and to react on-line to changing operating conditions. The present work explicitly addresses these issues.

**Control Signal Tracking**

Figure 6.9 depicts the setup for frequency control reserve provision using a portfolio of power nodes enhanced by a Control Power Node and a Slack Power Node. The tracking of the control signal is achieved by imposing the power balance constraint in $\Delta$-quantities on the $N + 2$
power nodes:

\[
\sum_{i=1}^{N+2} (\Delta u^*_{gen,i}(l) - \Delta u^*_{load,i}(l)) = 0 \tag{6.57}
\]

over the length of the prediction horizon. Note that we omit the generation cost of the generator in our optimization problem since the net production reduction or increase is either insignificant (primary control) or compensated by control energy payments. This is explained in detail in Chapter 7.

The control signal itself enters the problem via the Control Power Node, which is simply

\[
\Delta u^\text{ctrl}_\text{load} = -\Delta \xi^\text{ctrl}. \tag{6.58}
\]

Since imposing the control signal as an equality constraint onto the system may lead to infeasibility of the optimization, it is practical to allow a deviation from the control signal by introducing a virtual Slack Power Node that may cover some of the control action:

\[
\Delta u^\text{slack}_\text{gen} = \Delta \xi^\text{slack}. \tag{6.59}
\]

As a deviation from the control signal should only be a last resort to avoid infeasibility, we will penalize the slack contribution quadratically in the cost function with a penalty factor \(\pi^\text{dev}[\text{EUR/MW}^2]\) which should be chosen large in comparison with the other cost terms. The cost function for frequency control is then as follows:

\[
J(k) = \sum_{l=k}^{k+N_{\text{opt}}-1} \left( J^\text{rel}_\text{endo}(l) + \pi^\text{dev}(\Delta u^\text{slack}_\text{gen}(l))^2 \right). \tag{6.60}
\]

The presented setup allows us to track, and deviate from, control signals that are imposed on the portfolio. In what follows, we will discuss the specifics of using the setup with primary, secondary, and tertiary frequency control signals.

**Primary Frequency Control**

Primary frequency control is a proportional control which adapts the power production of a unit (usually a generation unit) based on the frequency deviation from its set value. The negative inverse of the proportionality factor between frequency deviation and power change is
6.5. Use Cases and Control Problems

referred to as droop [179], here denoted by $S \left[ \frac{\text{Hz}}{\text{MW}} \right]$. It can be expressed as

$$S = -\frac{\Delta f}{\Delta P} = -\frac{\Delta f_{\text{p.u.}}}{\Delta P_{\text{p.u.}}} \frac{f^{\text{ref}}}{S_B},$$  \hspace{1cm} (6.61)

where $\Delta f = f - f^{\text{ref}}$ is the system frequency deviation, $\Delta P = P - P^{\text{ref}}$ is the deviation from the power production setpoint, p.u. denotes per-unit values, and $S_B$ is the apparent power base of the system. In the ENTSO-E [88] CE, the droop has to be set such that the entire primary frequency reserve is activated at a frequency deviation of $\pm 200$ mHz. This means that for a control band of 1 MW, the system frequency deviation must be multiplied by 1 MW/0.2 Hz = 5 $\frac{\text{MW}}{\text{Hz}}$ to derive the control signal. Furthermore, a dead-band of $\pm 20$ mHz is imposed where no control action shall be taken, so the system frequency must be pre-processed accordingly. We will refer to the measured frequency deviation with the filtered-out deadband as $\Delta f_{\text{db}}(k)$. The control signal input to the power node portfolio as presented in (6.58), scaled by the offered primary control capacity $P_{\text{prim prov}}$, is thus

$$\Delta \xi_{\text{ctrl,ts}}(k) = P_{\text{prim prov}} \frac{200 \text{ mHz}}{Y} \Delta f_{\text{db}}^{*}(k).$$ \hspace{1cm} (6.62)

### Secondary Frequency Control

Secondary frequency control is a PI control loop on the control area level. It is responsible for keeping the system frequency at its nominal value and for maintaining the scheduled tie-line exchanges with other control areas on an integral basis. Due to the integrating part, there can only be one central controller per control area since several integrating controllers acting on the power balance of the system can lead to destabilizing effects. The control signal is derived from the ACE [179] and is issued in percent of the secondary control band:

$$\Delta \xi_{\text{ctrl,ts}}(k) = -P^{\text{sec prov}} Y(k)$$ \hspace{1cm} (6.63)

with $Y \in [-100\%, 100\%]$ being the secondary control signal in percent and $P^{\text{sec prov}}$ being the offered secondary frequency control capacity.

### Tertiary Frequency Control

In contrast to the previous two frequency control services, tertiary frequency control is a manually activated reserve. It is used by the TSO
to relieve secondary control reserves in order to avoid saturation of the contract reserve capacity. The reserve activation is defined by three parameters:

1. Activation Lead Time $\Delta t_{\text{lead}}$: lead time from calling the reserve until actual delivery,

2. Minimum Delivery Time $\Delta t_{\text{min}}^{\text{del}}$: minimum time span in which the delivery has to be sustained (the TSO cannot ask for an earlier deactivation), and

3. Maximum Delivery Time $\Delta t_{\text{max}}^{\text{del}}$: maximum time span in which the delivery has to be sustained (the TSO cannot ask for a longer delivery).

In Switzerland, both lead time and minimum delivery time are equal to 15 minutes [180]. While the maximum delivery time is theoretically infinite, in practice it is usually confined to one hour.

In the case of activation, a message containing the required power increase or decrease $\Delta P_{\text{tert}}$ is sent. The power node portfolio needs to deliver the requested power by a collective power change, so the input of the Control Power Node is equal to

$$\Delta \xi_{\text{ctrl},\text{ts}}(k) = -\Delta P_{\text{tert}}(k). \quad (6.64)$$

### 6.5.5 Peak Shaving

Another interesting application of a portfolio of controllable units in the distribution grid is the reduction of the distribution system’s peak load which needs to be served by imported power from the transmission level. Both grid operators and electricity companies serving final customers have an interest in a low peak demand since usually a peak power tariff has to be paid. For example, the following peak demand billing model holds for the Swiss electricity system:

$$C_{\text{peak}} = \frac{1}{12} \pi_{\text{peak}} \sum_{i=1}^{12} P_{\text{peak},i}^{\text{month*}}, \quad (6.65)$$

where $C_{\text{peak}}$ [CHF] is the overall annual cost for peak power, $\pi_{\text{peak}}$ [CHF/MW] is a price multiplier (equal to CHF 25,600 in 2011 according to [181]),
and $P_{\text{peak},i}^{\text{month}}$ [MW] is the maximum load drawn from the transmission grid during one month. On the level of the entire power system, peak shaving helps to ensure that peak demand can be met by the available generation assets with a certain safety margin (often referred to as generation adequacy).

As the balancing application, peak shaving can be performed by a power node portfolio either for the portfolio itself or as a service to the rest of the system. The first case is considered here. Figure 6.10 shows the corresponding setup, which is the same as for the following cases “Capacity Firming” and “Ramp-Rate Reduction”. We consider the following setup for the peak shaving application: the portfolio, consisting of loads, generators, and/or storage units, exchanges energy with its surroundings, e.g., an external supplier contracted through a bilateral agreement. This party is represented by the Slack Power Node:

$$u_{\text{slack}}^{\text{gen}} - u_{\text{slack}}^{\text{load}} = \xi^{\text{slack}}. \quad (6.66)$$

The power balance of production and consumption is formulated in absolute quantities:

$$\sum_{i=1}^{N+1} (u_{\text{gen},i}^*(l) - u_{\text{load},i}^*(l)) = 0. \quad (6.67)$$
By dispatching its flexible units accordingly, the power node portfolio has a direct influence on the residual load (which can also be negative in the general case) that is exerted on the external party, i.e., the Slack Power Node. For the peak shaving application, we restrict the residual load to be positive, i.e., \( u_{\text{slack}} = 0 \). The following cost function yields the desired behavior of peak demand reduction by control actions of the power node portfolio:

\[
J(k) = \sum_{l=k}^{k+N_{\text{opt}}-1} J^*_{\text{endo}}(l) + \frac{1}{12} \pi_{\text{peak}} \max_{l \in [k,k+N_{\text{opt}}-1]} \left( u_{\text{slack}}^* (l) \right) . \tag{6.68}
\]

Note that this cost function is only meaningful if used in a multi-period optimization for an entire month, or for a day-ahead optimization on the day with the expected maximum demand since the peak price is only relevant for the maximum demand per month.

### 6.5.6 Capacity Firming of Intermittent Generation

One of the challenges associated with integrating intermittent renewable energy is the lack of dependability of the installed capacity. While a dispatchable power plant such as natural gas or biomass plant is able to reliably contribute to serving the system’s peak load, contributions by wind and solar generation are subject to weather influences and uncertainty.

The so-called “capacity credit”, particularly relevant for wind energy, stipulates the percentage of the installed capacity that can be assumed to produce power at all times \([182]\), usually quantified to a few percent of the total installed capacity in the case of wind and photovoltaic (PV) energy. In practice, however, prolonged period of almost no intermittent RES in-feeds can be observed. The overall capacity credit of PV power generation is equal to zero since no production takes place at night. During day time, it is also highly dependent on cloud coverage, so that exact values are difficult to derive.

One way to deal with the capacity uncertainty is to try to increase the minimum of the net power in-feed of a certain group of units, which essentially leads to an increased firmly provided (dependable) capacity to the system. This involves an optimization problem that strives to maximize the minimum export from the unit portfolio. References \([144]\) and \([183]\) present similar methods to approach this problem.
In PNMF nomenclature, the net export from the portfolio can be represented by a Slack Power Node that absorbs the excess power:

\[ u_{\text{slack}}^{\text{load}} = -\xi_{\text{slack}}. \]  

(6.69)

The optimization problem is thus formulated with a cost function involving the endogenous cost of the portfolio and the negative minimum load of the Slack Power Node with the capacity premium price \( \pi_{\text{cap}} \text{[\text{EUR/MW}]} \):

\[
J(k) = \sum_{l=k}^{k+N_{\text{opt}}-1} J_{\text{endo}}^*(l) - \pi_{\text{cap}} \cdot \min_{l \in [k, k+N_{\text{opt}}-1]} u_{\text{slack}}^{\text{load}}(l). \]

(6.70)

### 6.5.7 Residual Load Ramp-Rate Reduction

In the presence of intermittent generation in the power system, the ramp-rate of the residual load, which has to be covered by the controllable generators, can exhibit large ramp-rates compared to the original load curve. This is mainly due to naturally occurring intermittency of wind and solar as well as an adverse coincidence of load ramps with intermittent generation ramps. Reference [33] stresses the importance of managing the residual load ramp rate.

A portfolio of units represented by power nodes can contribute to reducing the overall system’s ramp rate by using its flexibility of power generation and/or consumption. As in the case of peak shaving, this can be achieved by controlling for a ramp decrease in the Slack Power Node exchange:

\[
J(k) = \sum_{l=k}^{k+N_{\text{opt}}-1} \left( J_{\text{endo}}^*(l) + \pi_{\text{ramp}}^{\text{slack}} \frac{1}{t_s} (\delta u_{\text{load}}^{\text{slack}}(l))^2 \right), \]

(6.71)

where \( \pi_{\text{ramp}}^{\text{slack}} \text{[\text{EUR/MW}^2/h]} \) is a residual load ramping penalty, \( t_s \) is the sampling time, and \( \delta u_{\text{load}}^{\text{slack}}(l) = u_{\text{load}}^{\text{slack}}(l) - u_{\text{load}}^{\text{slack}}(l-1) \) is the slack power change from one time step to the other.
Figure 6.11: Structure of the simulation environment

6.6 Simulation Environment

Figure 6.11 depicts the structure of the simulation environment. The power node setup and simulation case is defined in the module “System Definition”. The module “Problem Builder” parses through the definition of the power node topology and builds the aggregated dynamical system with parameters and constraints according to Section 6.4.2.

The assembled dispatch problem is then passed to the “Simulator” module, where the optimization problem is pre-solved for variable input parameters. The MPC dispatch algorithm evaluates the solution in a receding horizon fashion and thus determines setpoints for the power nodes. In practice, these setpoints would then be sent to the respective units or to a coordination algorithm which is able to achieve setpoint tracking with distributed Thermostatically Controlled Load (TCL) populations. The results are saved and evaluated in an appropriate graphical and numerical way.
6.7 Benchmark Power Node Portfolios

We have created a number of benchmark portfolios that represent certain typical characteristics of modern power systems such as the presence of intermittent in-feeds and loads, storage devices, and controllable loads. The parameters of the modeled units were chosen in accordance with literature values of the respective technologies. Table 6.4 summarizes the used power nodes models with the relevant parameters and constraints. The benchmark portfolios are summarized in Tables 6.5 and Table 6.6. They are described in the following:

1. Large System Portfolio with RES: This portfolio represents an aggregated set of generation, load, and storage units in a larger area, e.g., in a country. The system has a conventional peak load of 10 GW and consists furthermore of a large aggregated water heater load, natural gas and biomass generation, wind and solar PV generation, a pumped-hydro energy storage, and aggregated stationary batteries. Import and export is not considered here, but can be modeled easily by adding a Slack Power Node.

2. VPP Portfolio with RES: This portfolio represents a VPP with a large renewable energy share, some conventional load, and a number of controllable assets that provide flexibility. The portfolio exhibits a conventional load peak of 30 MW and a total generation capacity (including intermittent generation) of 80 MW. A Slack Power Node models exchanges via a power exchange.

3. Control-VPP Portfolio A – Battery and Generator: For reserve provision, the combination of a generator and a battery is proposed. A Control Power Node imposes the control signal and the Slack Power Node accounts for deviations from that signal. The simplicity of the portfolio design ensures that the results of detailed parameter variations (such as presented in Chapter 7) remain understandable, which would not be the case for many units with many variable parameters. We provide a base case parameterization which will be varied later on in Chapter 7.

4. Control-VPP Portfolio B – Thermal Load and Generator: Here we use a thermal load and a conventional generator for reserve provision. The design rationale is the same as for Portfolio 3.
Table 6.4: Power node types for benchmark portfolios

<table>
<thead>
<tr>
<th>Power Node</th>
<th>Equation</th>
<th>Parameters</th>
<th>Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Generators</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural Gas</td>
<td>$\eta^{-1}<em>{gen}u</em>{gen} = \xi$</td>
<td>$\eta_{gen}$</td>
<td>$u_{gen}$, $u_{gen}^{-1}$</td>
</tr>
<tr>
<td>Biomass</td>
<td>$\eta^{-1}<em>{gen}u</em>{gen} = \xi$</td>
<td>$\eta_{gen}$</td>
<td>$u_{gen}$, $u_{gen}^{-1}$</td>
</tr>
<tr>
<td>Wind</td>
<td>$\eta^{-1}<em>{gen}u</em>{gen} = \xi - w$</td>
<td>$\eta_{gen}$</td>
<td>$\xi = \xi_{drv}(t)$</td>
</tr>
<tr>
<td>Solar PV</td>
<td>$\eta^{-1}<em>{gen}u</em>{gen} = \xi - w$</td>
<td>$\eta_{gen}$</td>
<td>$\xi = \xi_{drv}(t)$</td>
</tr>
<tr>
<td><strong>Loads</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water Heaters</td>
<td>$C \dot{x} = \eta_{load}u_{load}$</td>
<td>$\eta_{load}$, $C$, $a$, $x_{ss}$ $x$, $u_{load}$</td>
<td>$\xi = \xi_{drv}(t)$</td>
</tr>
<tr>
<td>Conventional Load</td>
<td>$\eta_{load}u_{load} = w - \xi$</td>
<td>$\eta_{load}$</td>
<td>$u_{load}$, $w$, $\xi = \xi_{drv}(t)$</td>
</tr>
<tr>
<td><strong>Storage Devices</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pumped-Hydro</td>
<td>$C \dot{x} = \eta_{load}u_{load}$</td>
<td>$\eta_{load}$, $\eta_{gen}$, $C$ $x$, $u_{load}$, $u_{load}$</td>
<td>$\xi = \xi_{drv}(t)$</td>
</tr>
<tr>
<td>Battery</td>
<td>$C \dot{x} = \eta_{load}u_{load} - \eta^{-1}<em>{gen}u</em>{gen}$</td>
<td>$\eta_{load}$, $\eta_{load}$, $C$ $x$, $u_{gen}$, $u_{load}$</td>
<td></td>
</tr>
<tr>
<td><strong>Auxiliary Units</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>$\eta_{load}u_{load} = \xi$</td>
<td>$\eta_{load}$</td>
<td>$\xi = \xi_{drv}(t)$</td>
</tr>
<tr>
<td>Slack</td>
<td>$u_{gen} = \xi$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6.8 Simulation Examples

In this section, simulation examples for each of the formulated dispatch strategies are presented using the benchmark systems defined above. We will show time simulations over the range of two days to two weeks. Longer simulations are conducted in Chapter 7. Aggregated properties of the simulations are presented by evaluating a number of balance terms such as presented in Section 5.2.4 over the simulation time span.

6.8.1 Day-Ahead Dispatch

The day-ahead dispatch is simulated with Portfolio 1. It serves to demonstrate how a day-ahead prediction for load and intermittent in-feeds is used to derive an optimal dispatch for the portfolio. Figure 6.12-a) shows the in-feeds into (positive) and out-feeds from the grid (negative), which have the same absolute values in every time step. The reaction of the flexible units to the intermittent in-feeds can be observed during the entire time span. Figure 6.12-b) shows the profile variables imposed on the system through the variables $\xi_i$. Curtailments of excess in-feeds from wind and solar energy take place during the first day only, as depicted in Figure 6.12-c). Load shedding is avoided during the entire time span. Figure 6.12-d) shows the storage levels of the pumped-hydro storage, the aggregated battery storage units, and the aggregated controllable loads. Table 6.7 shows the associated balance terms for the simulated time span.
Table 6.5: Benchmark portfolio parameters and constraints

<table>
<thead>
<tr>
<th>Power Nodes</th>
<th>Parameters</th>
<th>Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Large System Portfolio with RES</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural Gas</td>
<td>$\eta_{gen} = 0.45$</td>
<td>$1 \text{ GW} \leq u_{gen} \leq 5 \text{ GW}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$-4 \text{ GW/h} \leq \dot{u}_{gen} \leq 4 \text{ GW/h}$</td>
</tr>
<tr>
<td>Biomass</td>
<td>$\eta_{gen} = 0.4$</td>
<td>$0.2 \text{ GW} \leq u_{gen} \leq 2 \text{ GW}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$-1 \text{ GW/h} \leq \dot{u}_{gen} \leq 1 \text{ GW/h}$</td>
</tr>
<tr>
<td>Wind</td>
<td>$\eta_{gen} = 1$</td>
<td>$0 \text{ GW} \leq u_{gen} \leq 15 \text{ GW}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$0 \text{ GW} \leq w \leq 15 \text{ GW}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\xi = \xi_{drv}(t)$</td>
</tr>
<tr>
<td>Solar PV</td>
<td>$\eta_{gen} = 1$</td>
<td>$0 \text{ GW} \leq u_{gen} \leq 18 \text{ GW}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$0 \text{ GW} \leq w \leq 18 \text{ GW}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\xi = \xi_{drv}(t)$</td>
</tr>
<tr>
<td>Conventional Load</td>
<td>$\eta_{load} = 1$</td>
<td>$0 \text{ GW} \leq u_{load} \leq 10 \text{ GW}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$0 \text{ GW} \leq w \leq 10 \text{ GW}$</td>
</tr>
<tr>
<td>Water Heaters</td>
<td>$\eta_{load} = 1, C = 24 \text{ GWh}, a = 150 \text{ MW}, x_{ss} = 0.2$</td>
<td>$0 \text{ GW} \leq u_{load} \leq 4 \text{ GW}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$0.2 \leq x \leq 1$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\xi = \xi_{drv}(t)$</td>
</tr>
<tr>
<td>Pumped-Hydro</td>
<td>$\eta_{load} = 0.8, \eta_{gen} = 0.95, C = 40 \text{ GWh}$</td>
<td>$0 \text{ GW} \leq u_{load} \leq 5 \text{ GW}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$0 \leq x \leq 1$</td>
</tr>
<tr>
<td>Battery</td>
<td>$\eta_{load} = 0.88, \eta_{gen} = 0.88, C = 20 \text{ GWh}$</td>
<td>$0 \text{ GW} \leq u_{load} \leq 2 \text{ GW}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$0.2 \leq x \leq 1$</td>
</tr>
<tr>
<td><strong>2. VPP Portfolio with RES</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biomass</td>
<td>$\eta_{gen} = 0.4$</td>
<td>$2 \text{ MW} \leq u_{gen} \leq 20 \text{ MW}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$-10 \text{ MW/h} \leq \dot{u}_{gen} \leq 10 \text{ MW/h}$</td>
</tr>
<tr>
<td>Wind</td>
<td>$\eta_{gen} = 1$</td>
<td>$0 \text{ MW} \leq u_{gen} \leq 40 \text{ MW}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$0 \text{ MW} \leq w \leq 40 \text{ MW}$</td>
</tr>
<tr>
<td>Solar PV</td>
<td>$\eta_{gen} = 1$</td>
<td>$0 \text{ MW} \leq u_{gen} \leq 40 \text{ MW}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$0 \text{ MW} \leq w \leq 40 \text{ MW}$</td>
</tr>
<tr>
<td>Conventional Load</td>
<td>$\eta_{load} = 1$</td>
<td>$0 \text{ MW} \leq u_{load} \leq 30 \text{ MW}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$0 \text{ MW} \leq w \leq 30 \text{ MW}$</td>
</tr>
<tr>
<td>Pumped-Hydro</td>
<td>$\eta_{load} = 0.95, \eta_{gen} = 0.8, C = 400 \text{ MW}$</td>
<td>$0 \text{ MW} \leq u_{load} \leq 20 \text{ MW}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$0 \leq x \leq 1$</td>
</tr>
<tr>
<td>Battery</td>
<td>$\eta_{load} = 0.88, \eta_{gen} = 0.88, C = 100 \text{ MWh}$</td>
<td>$0 \text{ MW} \leq u_{load} \leq 20 \text{ MW}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$0.2 \leq x \leq 1$</td>
</tr>
<tr>
<td>Slack (Market)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td><strong>3. Control-VPP Portfolio A</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural Gas</td>
<td>$\eta_{gen} = 0.45$</td>
<td>$-4 \text{ MW} \leq \Delta u_{gen} \leq 4 \text{ MW}$</td>
</tr>
<tr>
<td>Battery</td>
<td>$\eta_{load} = 0.88, \eta_{gen} = 0.88, C = 60 \text{ MWh}$</td>
<td>$0.2 \leq x \leq 1$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$0 \text{ MW} \leq \Delta u_{gen} \leq 6 \text{ MW}$</td>
</tr>
<tr>
<td>Control</td>
<td>$\eta_{load} = 1$</td>
<td>$\Delta \xi = \Delta \xi_{ctrl}(t)$</td>
</tr>
<tr>
<td>Slack</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td><strong>4. Control-VPP Portfolio B</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural Gas</td>
<td>$\eta_{gen} = 0.45$</td>
<td>$-4 \text{ MW} \leq \Delta u_{gen} \leq 4 \text{ MW}$</td>
</tr>
<tr>
<td>Water Heaters</td>
<td>$\eta_{load} = 1, C = 48 \text{ MWh}, a = 0.75 \text{ MW}, x_{ss} = 0.2$</td>
<td>$0.2 \leq x \leq 1$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$-6 \text{ MW} \leq u_{load} \leq 6 \text{ MW}$</td>
</tr>
<tr>
<td>Control</td>
<td>$\eta_{load} = 1$</td>
<td>$\Delta \xi = \Delta \xi_{ctrl}(t)$</td>
</tr>
<tr>
<td>Slack</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>
Table 6.6: Benchmark portfolio cost parameters

<table>
<thead>
<tr>
<th>Power Nodes</th>
<th>Cost term</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Large System Portfolio with RES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biomass</td>
<td>$\pi_{\text{fuel}} + \pi_{\text{O&amp;M}}$</td>
<td>70 EUR/MWh</td>
</tr>
<tr>
<td></td>
<td>$\pi_{\text{ramp}}$</td>
<td>0.1 EUR/(MW^2/h)</td>
</tr>
<tr>
<td>Biomass</td>
<td>$\pi_{\text{fuel}} + \pi_{\text{O&amp;M}}$</td>
<td>146 EUR/MWh</td>
</tr>
<tr>
<td></td>
<td>$\pi_{\text{ramp}}$</td>
<td>0.1 EUR/(MW^2/h)</td>
</tr>
<tr>
<td>Wind</td>
<td>$\pi_{\text{O&amp;M}}$</td>
<td>10 EUR/MW</td>
</tr>
<tr>
<td>Solar PV</td>
<td>$\pi_{\text{O&amp;M}}$</td>
<td>3 EUR/MWh</td>
</tr>
<tr>
<td>Conventional Load</td>
<td>$\pi_{\text{curt}}$</td>
<td>-1,000 EUR/MWh</td>
</tr>
<tr>
<td>Water Heaters</td>
<td>$\pi_{\text{EWH}}$</td>
<td>100 EUR</td>
</tr>
<tr>
<td>Pumped-Hydro</td>
<td>$\pi_{\text{PHP}}$</td>
<td>5 EUR/MWh</td>
</tr>
<tr>
<td>Battery</td>
<td>$\pi_{\text{bat}}$</td>
<td>50 EUR/MWh</td>
</tr>
<tr>
<td>2. VPP Portfolio with RES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biomass</td>
<td>$\pi_{\text{fuel}} + \pi_{\text{O&amp;M}}$</td>
<td>146 EUR/MWh</td>
</tr>
<tr>
<td>Wind</td>
<td>$\pi_{\text{O&amp;M}}$</td>
<td>10 EUR/MWh</td>
</tr>
<tr>
<td>Solar PV</td>
<td>$\pi_{\text{O&amp;M}}$</td>
<td>3 EUR/MWh</td>
</tr>
<tr>
<td>Pumped-Hydro</td>
<td>$\pi_{\text{PHP}}$</td>
<td>5 EUR/MWh</td>
</tr>
<tr>
<td>Battery</td>
<td>$\pi_{\text{bat}}$</td>
<td>50 EUR/MWh</td>
</tr>
<tr>
<td>Water Heaters</td>
<td>$\pi_{\text{EWH}}$</td>
<td>100 EUR</td>
</tr>
<tr>
<td>3. Control-VPP Portfolio A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural Gas</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Battery</td>
<td>$\pi_{\text{bat}}$</td>
<td>50 EUR/MWh</td>
</tr>
<tr>
<td>Control</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Slack</td>
<td>$\pi_{\text{slack}}$</td>
<td>1,000 EUR/MWh</td>
</tr>
<tr>
<td>4. Control-VPP Portfolio B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural Gas</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Water Heaters</td>
<td>$\pi_{\text{EWH}}$</td>
<td>100 EUR</td>
</tr>
<tr>
<td>Control</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Slack</td>
<td>$\pi_{\text{slack}}$</td>
<td>1,000 EUR/MWh</td>
</tr>
</tbody>
</table>

Table 6.7: Day-ahead dispatch – balance terms

| Natural gas energy produced | 736,300.21 MWh |
| Biomass energy produced    | 134,690.18 MWh |
| Wind energy produced       | 688,879.30 MWh |
| Wind energy curtailed      | 5,699.48 MWh   |
| Solar PV energy produced   | 907,816.81 MWh |
| Solar PV energy curtailed  | 8,211.99 MWh   |
| Energy consumed by conventional load | 2,256,602.33 MWh |
| Energy not served to conventional load | 0.00 MWh |
| Energy consumed by water heaters | 161,745.37 MWh |
| Energy fed into batteries  | 11,282.67 MWh  |
| Energy taken from batteries| 14,017.29 MWh  |
| Energy lost in batteries   | 3,265.37 MWh   |
| Energy fed into pumped-hydro storage | 274,921.95 MWh |
| Energy taken from pumped-hydro storage | 222,848.53 MWh |
| Energy lost in pumped-hydro storage | 65,999.24 MWh |

6.8.2 Intra-Day Update

Since the intra-day update is only marginally different from the day-ahead dispatch, we omit it here for shortness. An exemplary simulation is presented in [130].
Figure 6.12: Simulation example: day-ahead dispatch
6.8.3 Market-Based Operation of VPPs

The market-based operation of VPPs is demonstrated by simulating Portfolio 2 over a time span of 15 days as depicted in Figure 6.13. The simulation shows how the optimization reacts to the varying price and selectively imports and exports electricity via the market. For simplicity and demonstration purposes, we perform the simulation with a perfect price prediction. Figure 6.13-a) depicts again the network in-feeds and out-feeds, Figure 6.13-b) shows the curtailment which is equal to zero at all times, Figure 6.13-c) shows the spot price time series, and Figure 6.13-d) depicts the SOC evolution. It can be observed that the battery is not utilized due to the high marginal operation cost compared to the revenue potential, but the pumped-hydro plant actively contributes to energy arbitrage utilizing the time-varying price. Table 6.8 summarizes the balance terms.

Table 6.8: Market-based VPP operation – balance terms

<table>
<thead>
<tr>
<th>Energy Source</th>
<th>Quantity</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biomass energy produced</td>
<td>672.50</td>
<td>MWh</td>
</tr>
<tr>
<td>Wind energy produced</td>
<td>1,852.21</td>
<td>MWh</td>
</tr>
<tr>
<td>Solar PV energy produced</td>
<td>2,035.62</td>
<td>MWh</td>
</tr>
<tr>
<td>Energy consumed by conventional load</td>
<td>6,769.81</td>
<td>MWh</td>
</tr>
<tr>
<td>Energy fed into batteries</td>
<td>0.00</td>
<td>MWh</td>
</tr>
<tr>
<td>Energy taken from batteries</td>
<td>0.00</td>
<td>MWh</td>
</tr>
<tr>
<td>Energy lost in batteries</td>
<td>0.00</td>
<td>MWh</td>
</tr>
<tr>
<td>Energy fed into pumped-hydro storage</td>
<td>3,292.09</td>
<td>MWh</td>
</tr>
<tr>
<td>Energy taken from pumped-hydro storage</td>
<td>2,558.39</td>
<td>MWh</td>
</tr>
<tr>
<td>Energy lost in pumped-hydro storage</td>
<td>778.08</td>
<td>MWh</td>
</tr>
<tr>
<td>Energy imported from spot market</td>
<td>5,408.68</td>
<td>MWh</td>
</tr>
<tr>
<td>Energy exported to spot market</td>
<td>2,465.50</td>
<td>MWh</td>
</tr>
</tbody>
</table>

6.8.4 Pre-MTU Balancing

We omit the simulation of Pre-MTU Balancing here for shortness and refer to [64], where a similar case related to wind forecast errors is presented.
6.8. Simulation Examples

Figure 6.13: Simulation example: market-based VPP operation
6.8.5 Primary Frequency Control

Primary frequency control is simulated using both Control-VPP Portfolio A and B. The control signal is scaled to 10 MW (provided that the full reserve has to be activated at a frequency deviation of ±200 mHz). It is applied to the portfolios assuming that they have a constant operating point (base case). Note that it is straight-forward to reserve the control bands in an economic dispatch (using the methodology outlined in Section 6.2.2) and to apply the control signal on top of a time-varying working point trajectory. We simulate the control scenario with a 60% storage share, i.e., the generator control band amounts to 40% of the total control band. The storage capacity is equivalent to 10 hours for the battery and equal to 8 hours for the water heaters.

Figure 6.14 shows a time simulation of primary control provision using Portfolio A over 48 hours. It can be seen that the control action is completely absorbed by the power plant since its control band is sufficiently large. In case of a larger frequency deviation, the battery will be used to cover part of the control action. In this way, the battery mainly serves as a backup to enable the control power plant to reduce its reserved control band size. For Portfolio B, depicted in Figure 6.15, the results are different due to the different cost structure (no cycling cost). The balance terms are summarized in Tables 6.9 and 6.10.

<table>
<thead>
<tr>
<th>Table 6.9: Primary control (Portfolio A) – Δ balance terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive regulation demanded by control signal: 2.78 MWh</td>
</tr>
<tr>
<td>Negative regulation demanded by control signal: 32.90 MWh</td>
</tr>
<tr>
<td>Positive contribution (generation increase) by generator: 2.78 MWh</td>
</tr>
<tr>
<td>Negative contribution (generation decrease) by generator: 32.90 MWh</td>
</tr>
<tr>
<td>Positive contribution (generation increase) by battery: 0.00 MWh</td>
</tr>
<tr>
<td>Negative contribution (load increase) by battery: 0.00 MWh</td>
</tr>
<tr>
<td>Positive contribution (generation increase) by slack: 0.00 MWh</td>
</tr>
<tr>
<td>Negative contribution (generation decrease) by slack: 0.00 MWh</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 6.10: Primary control (Portfolio B) – Δ balance terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive regulation demanded by control signal: 2.78 MWh</td>
</tr>
<tr>
<td>Negative regulation demanded by control signal: 32.90 MWh</td>
</tr>
<tr>
<td>Positive contribution (generation increase) by generator: 24.32 MWh</td>
</tr>
<tr>
<td>Negative contribution (generation decrease) by generator: 0.00 MWh</td>
</tr>
<tr>
<td>Positive contribution (load decrease) by water heaters: 1.51 MWh</td>
</tr>
<tr>
<td>Negative contribution (load increase) by water heaters: 55.94 MWh</td>
</tr>
<tr>
<td>Positive contribution (generation increase) by slack: 0.00 MWh</td>
</tr>
<tr>
<td>Negative contribution (generation decrease) by slack: 0.00 MWh</td>
</tr>
</tbody>
</table>
6.8. Simulation Examples

Figure 6.14: Simulation example: primary control (Portfolio A)

Figure 6.15: Simulation example: primary control (Portfolio B)
6.8.6 Secondary Frequency Control

Secondary frequency control is simulated in the same setting as primary frequency control above, also using both Portfolio A and B. Again, we scale the control signal to the reserved control band of $\pm 10$ MW.

Figure 6.16 presents the control simulation scenario using Portfolio A. It can be observed that the control action is flexibly split between the generator and the battery. While the generator takes over most of the smaller deviations from zero, the utilization of the battery is mainly restricted to moves to the outside of the generator control band. This behavior is due to the cycling cost of the battery. The SOC is rising slightly in the first half of the simulation.

Figure 6.17 presents the same scenario using Portfolio B. Since the water heaters do not exhibit any cycling cost, the splitting of the control signal is considerably different. The generator assumes a negative bias throughout the two simulated days. The water heater control contribution moves up and down around the relatively steady generator trajectory, covering moves in both directions. The SOC is maintained around 50% throughout the simulation.

The balance terms of the secondary frequency control simulation are shown in Tables 6.11 and 6.12.

<table>
<thead>
<tr>
<th>Table 6.11: Secondary control (Portfolio A) – Δ balance terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive regulation demanded by control signal</td>
</tr>
<tr>
<td>Negative regulation demanded by control signal</td>
</tr>
<tr>
<td>Positive contribution (generation increase) by generator</td>
</tr>
<tr>
<td>Negative contribution (generation decrease) by generator</td>
</tr>
<tr>
<td>Positive contribution (generation increase) by battery</td>
</tr>
<tr>
<td>Negative contribution (load increase) by battery</td>
</tr>
<tr>
<td>Positive contribution (generation increase) by slack</td>
</tr>
<tr>
<td>Negative contribution (generation decrease) by slack</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 6.12: Secondary control (Portfolio B) – Δ balance terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive regulation demanded by control signal</td>
</tr>
<tr>
<td>Negative regulation demanded by control signal</td>
</tr>
<tr>
<td>Positive contribution (generation increase) by generator</td>
</tr>
<tr>
<td>Negative contribution (generation decrease) by generator</td>
</tr>
<tr>
<td>Positive contribution (load decrease) by water heaters</td>
</tr>
<tr>
<td>Negative contribution (load increase) by water heaters</td>
</tr>
<tr>
<td>Positive contribution (generation increase) by slack</td>
</tr>
<tr>
<td>Negative contribution (generation decrease) by slack</td>
</tr>
</tbody>
</table>
6.8. Simulation Examples

Figure 6.16: Simulation example: secondary control (Portfolio A)

Figure 6.17: Simulation example: secondary control (Portfolio B)
6.8.7 Tertiary Frequency Control

Tertiary frequency control provision consists of following a rectangular or trapezoidal control signal. Since this is straight-forward to simulate similar to primary and secondary control, we omit it here for shortness.

6.8.8 Peak Shaving

In order to demonstrate the effect of peak shaving control actions, we compare the same simulation of a portfolio over the time span of two weeks with and without the peak penalty term according to Section 6.5.5 in the cost function of the optimization. We focus here on the day-ahead optimization case, where a total of four days is taken into account by the optimizer, but only the control action for the upcoming day is implemented. Note that the same methodology is usable for intra-day updates in a receding horizon fashion, provided that the prediction horizon is long enough to cover the next peak which is to be avoided.

Figure 6.18 depicts the base case in which relatively high load peaks have to be covered by importing energy into the portfolio through the Slack Power Node. Figure 6.19 shows the simulation with active peak shaving term where the import is avoided altogether during the first week by ramping up the biomass production and dispatching the storage units accordingly. In the second week, strong reduction of the peak load can be observed as well. In Figure 6.20, the difference between the previously shown two scenarios is illustrated. Table 6.13 shows a number of balance terms on the difference between the base case and the peak shaving case. It can be seen that the peak load to be covered by the slack has been reduced by over 70%.
Figure 6.18: Simulation example: peak shaving (base case)

Figure 6.19: Simulation example: peak shaving (application)
Figure 6.20: Simulation example: peak shaving (above: base case, below: peak shaving active)

Table 6.13: Peak shaving – balance terms

<table>
<thead>
<tr>
<th>Base case:</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Biomass energy produced</td>
<td>624.50 MWh</td>
<td></td>
</tr>
<tr>
<td>Wind energy produced</td>
<td>1,801.56 MWh</td>
<td></td>
</tr>
<tr>
<td>Solar PV energy produced</td>
<td>1,936.46 MWh</td>
<td></td>
</tr>
<tr>
<td>Energy consumed by conventional load</td>
<td>6,232.95 MWh</td>
<td></td>
</tr>
<tr>
<td>Energy not served to conventional load</td>
<td>0.00 MWh</td>
<td></td>
</tr>
<tr>
<td>Energy fed into batteries</td>
<td>50.77 MWh</td>
<td></td>
</tr>
<tr>
<td>Energy taken from batteries</td>
<td>0.03 MWh</td>
<td></td>
</tr>
<tr>
<td>Energy lost in batteries</td>
<td>6.10 MWh</td>
<td></td>
</tr>
<tr>
<td>Energy fed into pumped-hydro storage</td>
<td>337.73 MWh</td>
<td></td>
</tr>
<tr>
<td>Energy taken from pumped-hydro storage</td>
<td>210.61 MWh</td>
<td></td>
</tr>
<tr>
<td>Energy lost in pumped-hydro storage</td>
<td>74.06 MWh</td>
<td></td>
</tr>
<tr>
<td>Energy imported from external source</td>
<td>2,048.28 MWh</td>
<td></td>
</tr>
<tr>
<td>Energy exported to external sink</td>
<td>0.00 MWh</td>
<td></td>
</tr>
<tr>
<td>Maximum power imported from external source</td>
<td>18.45 MW</td>
<td></td>
</tr>
</tbody>
</table>

| Including peak shaving:                   |                                |                                |
| Biomass energy produced                   | 2,053.20 MWh                   |                                |
| Wind energy produced                      | 1,801.56 MWh                   |                                |
| Solar PV energy produced                  | 1,936.46 MWh                   |                                |
| Energy consumed by conventional load      | 6,232.95 MWh                   |                                |
| Energy not served to conventional load    | 0.00 MWh                       |                                |
| Energy fed into batteries                 | 245.54 MWh                     |                                |
| Energy taken from batteries               | 193.37 MWh                     |                                |
| Energy lost in batteries                  | 55.83 MWh                      |                                |
| Energy fed into pumped-hydro storage      | 142.96 MWh                     |                                |
| Energy taken from pumped-hydro storage    | 270.84 MWh                     |                                |
| Energy lost in pumped-hydro storage       | 51.54 MWh                      |                                |
| Energy imported from external source      | 366.01 MWh                     |                                |
| Energy exported to external sink          | 0.00 MWh                       |                                |
| Maximum power imported from external source| 5.22 MW                        |                                |
6.8. Simulation Examples

6.8.9 Capacity Firming of Intermittent Generation

Capacity firming represents the complementary case to peak shaving since we try to maximize the minimum in-feed instead of minimizing the maximum consumption. The same methodology as above is applied: a base case simulation is performed and compared with the same scenario simulated including the capacity firming term in the cost function.

Figure 6.21 presents the simulation of the base case, where the storage units stay inactive and the biomass generation is at its minimum. Figure 6.22 presents the simulation including the capacity firming control. It can be seen that the biomass generation is ramped up and the storage units are emptied in times of low intermittent in-feeds. Figure 6.23 presents the difference between the two cases. Table 6.14 summarizes the balance terms for the capacity firming, showing that the minimum export could be increased by almost 600%.

Figure 6.21: Simulation example: capacity firming (base case)
Figure 6.22: Simulation example: capacity firming (application)

Figure 6.23: Simulation example: capacity firming (above: base case, below: capacity firming active)
Table 6.14: Capacity firming – balance terms

<table>
<thead>
<tr>
<th>Base case:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Biomass energy produced</td>
<td>624.50 MWh</td>
</tr>
<tr>
<td>Wind energy produced</td>
<td>1,801.56 MWh</td>
</tr>
<tr>
<td>Solar PV energy produced</td>
<td>1,936.46 MWh</td>
</tr>
<tr>
<td>Energy fed into batteries</td>
<td>0.00 MWh</td>
</tr>
<tr>
<td>Energy taken from batteries</td>
<td>0.00 MWh</td>
</tr>
<tr>
<td>Energy lost in batteries</td>
<td>0.00 MWh</td>
</tr>
<tr>
<td>Energy fed into pumped-hydro storage</td>
<td>0.00 MWh</td>
</tr>
<tr>
<td>Energy taken from pumped-hydro storage</td>
<td>0.00 MWh</td>
</tr>
<tr>
<td>Energy lost in pumped-hydro storage</td>
<td>0.00 MWh</td>
</tr>
<tr>
<td>Energy imported from external source</td>
<td>0.00 MWh</td>
</tr>
<tr>
<td>Energy exported to external sink</td>
<td>4,362.53 MWh</td>
</tr>
<tr>
<td>Minimum power exported to external sink</td>
<td>2.32 MW</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>With capacity firming:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Biomass energy produced</td>
<td>1,861.44 MWh</td>
</tr>
<tr>
<td>Wind energy produced</td>
<td>1,801.56 MWh</td>
</tr>
<tr>
<td>Solar PV energy produced</td>
<td>1,936.46 MWh</td>
</tr>
<tr>
<td>Energy fed into batteries</td>
<td>208.10 MWh</td>
</tr>
<tr>
<td>Energy taken from batteries</td>
<td>168.62 MWh</td>
</tr>
<tr>
<td>Energy lost in batteries</td>
<td>47.97 MWh</td>
</tr>
<tr>
<td>Energy fed into pumped-hydro storage</td>
<td>28.77 MWh</td>
</tr>
<tr>
<td>Energy taken from pumped-hydro storage</td>
<td>192.62 MWh</td>
</tr>
<tr>
<td>Energy lost in pumped-hydro storage</td>
<td>25.72 MWh</td>
</tr>
<tr>
<td>Energy imported from external source</td>
<td>0.00 MWh</td>
</tr>
<tr>
<td>Energy exported to external sink</td>
<td>5,723.85 MWh</td>
</tr>
<tr>
<td>Minimum power exported to external sink</td>
<td>13.62 MW</td>
</tr>
</tbody>
</table>

6.8.10 Residual Load Ramp-Rate Reduction

The residual load ramping reduction is demonstrated in the same way as in the previous two cases. We show the simulation of the base case in Figure 6.24, the application of the ramp-rate reduction in Figure 6.25, and a comparison of the base case and the ramp-rate reduction case in Figure 6.26. Table 6.15 presents balance terms which demonstrate the effectiveness of the ramp-rate reduction method: the average ramp rate was reduced by over 90%.
Figure 6.24: Simulation example: ramp-rate reduction (base case)

Figure 6.25: Simulation example: ramp-rate reduction (application)
6.8. Simulation Examples

Figure 6.26: Simulation example: ramp-rate reduction (above: base case, below: ramp-rate reduction active)

Table 6.15: Ramp-rate reduction – balance terms

<table>
<thead>
<tr>
<th></th>
<th>Base case:</th>
<th>Including ramp-rate reduction:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biomass energy produced</td>
<td>624.50 MWh</td>
<td>1,651.97 MWh</td>
</tr>
<tr>
<td>Wind energy produced</td>
<td>1,080.94 MWh</td>
<td>1,080.94 MWh</td>
</tr>
<tr>
<td>Solar PV energy produced</td>
<td>1,161.88 MWh</td>
<td>1,161.88 MWh</td>
</tr>
<tr>
<td>Energy consumed by conventional load</td>
<td>6,232.95 MWh</td>
<td>6,232.95 MWh</td>
</tr>
<tr>
<td>Energy not served to conventional load</td>
<td>0.00 MWh</td>
<td>0.00 MWh</td>
</tr>
<tr>
<td>Energy fed into batteries</td>
<td>13.57 MWh</td>
<td>53.18 MWh</td>
</tr>
<tr>
<td>Energy taken from batteries</td>
<td>0.00 MWh</td>
<td>20.71 MWh</td>
</tr>
<tr>
<td>Energy lost in batteries</td>
<td>1.63 MWh</td>
<td>9.21 MWh</td>
</tr>
<tr>
<td>Energy fed into pumped-hydro storage</td>
<td>18.52 MWh</td>
<td>75.46 MWh</td>
</tr>
<tr>
<td>Energy taken from pumped-hydro storage</td>
<td>0.00 MWh</td>
<td>17.48 MWh</td>
</tr>
<tr>
<td>Energy lost in pumped-hydro storage</td>
<td>2.78 MWh</td>
<td>13.26 MWh</td>
</tr>
<tr>
<td>Energy imported from external source</td>
<td>3,397.72 MWh</td>
<td>3,397.72 MWh</td>
</tr>
<tr>
<td>Energy exported to external sink</td>
<td>0.00 MWh</td>
<td>0.00 MWh</td>
</tr>
<tr>
<td>Average ramp-rate of external source</td>
<td>1.52 MW/h</td>
<td>0.14 MW/h</td>
</tr>
</tbody>
</table>
6.9 Concluding Remarks

In this chapter, we have demonstrated how the PNMF can be used to formulate strategies for control service provision. The starting point was the decomposition of the power node equation into several temporal stages and the compact formulation of the power node portfolio. We then continued with the presentation of an MPC approach that can be adapted to different control objectives, which in turn can be pursued by operating a power node portfolio. We find that adaptations of constraints and the cost function, along with the introduction of the auxiliary power nodes (“Control Power Node” and “Slack Power Node”), allow to model a variety of different use cases of flexible unit portfolios in power systems.
Chapter 7

Economic Evaluation of Frequency Control Provision by Flexible Unit Portfolios

In this chapter, we present an economic evaluation of frequency control provision (primary and secondary control) by flexible unit portfolios, also referred to as Virtual Power Plants (VPPs), consisting of generators, energy storage units, and controllable thermal loads. This investigation relates directly to the Power Nodes Modeling Framework (PNMF) optimization strategies presented in Chapter 6. Time simulations over parameter ranges are undertaken so as to determine the control potential provided by two benchmark portfolios. Ancillary service market data from Switzerland are used to assess the value of using energy-storing units for frequency control instead of generators only. This yields a quantitative evaluation of the commercial potential of control services by coordinated flexible unit portfolios. Furthermore, we present an approach for sharing the profit among the various participating actors by taking advantage of the e³-value methodology for modeling value exchanges between actors in a market setting.

7.1 Introduction and Literature Review

In Chapter 6, we presented use cases for flexible unit portfolios comprising generators, controllable loads, and storage devices. A key application of these Virtual Power Plants (VPPs) in power systems is the provision of frequency control reserves. A crucial question for this use
Table 7.1: Notation for Chapter 7

<table>
<thead>
<tr>
<th>Var.</th>
<th>Unit</th>
<th>Meaning</th>
<th>Var.</th>
<th>Unit</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>EUR</td>
<td>Cost</td>
<td>α</td>
<td>[-]</td>
<td>Storage share</td>
</tr>
<tr>
<td>E</td>
<td>MWh</td>
<td>Energy</td>
<td>β</td>
<td>[-]</td>
<td>Admin fee coeff.</td>
</tr>
<tr>
<td>K</td>
<td>[-]</td>
<td>No. of time steps</td>
<td>γ</td>
<td>[-]</td>
<td>Control resp. fee coeff.</td>
</tr>
<tr>
<td>k</td>
<td>[-]</td>
<td>Time step</td>
<td>δ</td>
<td>[-]</td>
<td>Storage fee coeff.</td>
</tr>
<tr>
<td>M</td>
<td>[-]</td>
<td>Mapping</td>
<td>Δ</td>
<td>[-]</td>
<td>Change</td>
</tr>
<tr>
<td>n</td>
<td>[-]</td>
<td>Number</td>
<td>Π</td>
<td>EUR</td>
<td>Profit</td>
</tr>
<tr>
<td>R</td>
<td>EUR</td>
<td>Revenue</td>
<td>π</td>
<td>MWh</td>
<td>Price</td>
</tr>
<tr>
<td>T</td>
<td>[-]</td>
<td>Time span</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subscript</th>
<th>Meaning</th>
<th>Subscript</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>activ</td>
<td>Activities</td>
<td>MTU</td>
<td>Market Time Unit</td>
</tr>
<tr>
<td>actor</td>
<td>Market actor</td>
<td>net</td>
<td>Net value</td>
</tr>
<tr>
<td>bat</td>
<td>Battery</td>
<td>opp</td>
<td>Opportunity</td>
</tr>
<tr>
<td>capa</td>
<td>Capacity</td>
<td>prov</td>
<td>Provision</td>
</tr>
<tr>
<td>en</td>
<td>Energy</td>
<td>ramp</td>
<td>Ramping</td>
</tr>
<tr>
<td>exch</td>
<td>Exchanges</td>
<td>s</td>
<td>Sampling</td>
</tr>
<tr>
<td>fuel</td>
<td>Generator fuel</td>
<td>spot</td>
<td>Spot market</td>
</tr>
<tr>
<td>gen</td>
<td>Generator</td>
<td>strg</td>
<td>Storage</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Superscript</th>
<th>Meaning</th>
<th>Superscript</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>base</td>
<td>Base case value</td>
<td>feed-in</td>
<td>Feed-in of energy</td>
</tr>
<tr>
<td>bat</td>
<td>Battery</td>
<td>fixed</td>
<td>Fixed one-time fee</td>
</tr>
<tr>
<td>capped</td>
<td>Value with upper bound</td>
<td>floored</td>
<td>Value with lower bound</td>
</tr>
<tr>
<td>CL</td>
<td>Controllable loads</td>
<td>gen</td>
<td>Generation</td>
</tr>
<tr>
<td>cons</td>
<td>Consumption</td>
<td>max</td>
<td>Maximum</td>
</tr>
</tbody>
</table>

case is the revenue that can be achieved with a certain amount of control capacity provision on ancillary service markets. Previous work on the value of controllable demand has been done mainly for the cases of load shifting and the exploitation of price differences, such as in [184]. The value of energy storage devices for short-term dispatch and balancing was assessed in [185]. Frequency control ancillary service provision by Plug-In Hybrid Electric Vehicles (PHEVs) was described and economically evaluated in [186], demonstrating substantial revenue potentials.

In this chapter, we examine a market-based approach to ancillary service provision by flexible unit portfolios consisting of storage, controllable load, and generation units. In this setup, individual market players, called aggregators, control unit populations in a suitable way for delivering control reserves. The aggregators can both be separated from or integrated in established electricity utilities. The main point of investigation is the contribution that a certain type of unit can make to a control product. This requires a consideration of the inherent energy constraints that any kind of storage unit exhibits (acceptable internal temperature range in the case of thermal loads). This question is approached similarly to [163] by a time simulation setup using benchmark portfolios described in the Power Nodes Modeling Framework (PNMF) nomenclature. The results of the time-domain simulations are merged
with an analysis of historical price data, in this case of the Swiss frequency control reserves market, in order to estimate the financial revenue potential of a certain unit portfolio bidding into the market.

Some results of this chapter are based on [187]. It is structured as follows: Section 7.2 discusses properties of liberalized electricity markets and the role of aggregators. In Section 7.3, the modeling of the revenue from ancillary service provision by flexible unit portfolios is presented. Section 7.4 presents a case study using the benchmark portfolios 3 and 4 from Chapter 6. Section 7.5 presents the modeling of a method for profit sharing between the various actors. Section 7.6 presents concluding remarks. The used notation is summarized in Table 7.1.

### 7.2 Aggregators in Electricity Markets

In the “old world” of vertically integrated electricity utilities, the tasks of electricity generation, network operation, and supplying final customers were usually carried out by the same company. The simultaneous fulfillment of several roles enabled the company to perform a joint optimization of their generation and network assets. This is not possible in the same way in liberalized electricity systems. Depending on the regulatory framework, the Distribution System Operators (DSOs) may need to accommodate various electricity suppliers in their grid domain, while suppliers may be active in more than one distribution grid. Figure 7.1 illustrates the change from a vertically integrated utility structure to a liberalized electricity system. The exchange of energy between producers and suppliers is handled via an electricity market, and the generation companies have to bid control reserves and other grid services into ancillary service markets administered by the Transmission System Operator (TSO). The information exchange between grid operators and market players is usually restricted in order to guarantee non-discriminatory conditions. This has a number of implications for the introduction of “SmartGrid” functionalities in power systems, which we will briefly discuss in the following.

#### 7.2.1 The Role of Aggregators

Aggregators are relatively new entities in electricity systems which possess the ability to influence a number of grid-connected units via a suitable communication interface. The units are coordinated, usually
Figure 7.1: Vertically integrated versus liberalized system structure

by a centralized optimization, in order to fulfill a certain control goal as a group. Aggregators may operate in different parts of the electricity network and utilize the units in their portfolio for trading in electricity and ancillary service markets. Note that, depending on the market design, an aggregator can also be understood as an entity that coordinates the units in a certain area of the network in the sense of a MicroGrid (MG) [82], but we do not consider this setting here. We use the term aggregator synonymously to VPP operator.

Aggregators are market players by nature. Their aim is the commercially successful operation of their connected units, be it in the form of energy schedule optimization or in the form of power system control services. They can both be separated from or associated with an established electricity utility. The control actions that they impose on the connected power system units may influence the load flows as well as transformer and line loadings of one or several distribution systems. Thus, care has to be taken if relevant distribution grid constraints exist in the area of operation of an aggregator.

7.2.2 Distribution Grid Constraints

The existence of grid constraints is a self-evident fact for any distribution system. Although traditionally distribution grids were designed such that the peak load could be comfortably accommodated without
running the risk of transformer or line/cable overloading, coordinated control actions on large amounts of flexible demand may quickly drive the system towards the constraints. For example, negative tertiary control provided by electric water heaters, called by the TSO during the evening peak, is likely to increase the peak load to an unacceptable level. In the long run, structured information interfaces must be established between the DSOs and the aggregators that control the flexible demand in their area. Figure 7.2 outlines this situation. The interface can be established, e.g., in the form of time-varying power ranges that denote the permissible change in power consumption by the flexible load.

### 7.2.3 Unit Monitoring Challenges

Another complication of ancillary service provision by coordinated dispersed units is the need to prove to the TSO that the power delivery actually happened as requested. For instance, the relevant documents for the Swiss power system [188] stipulate that all units providing secondary control reserves have to deliver continuous on-line measurements to the TSO. In the case of generation units, this requires expensive instrumentation and telemetry. When a large number of dispersed units is delivering the same service, this degree of on-line monitoring may be difficult to achieve. Consequently, it may be necessary to adapt the regulatory framework of ancillary service provision depending on the nature of the control resources.
7.3 Modeling of Revenue Potential

The profitability of the application of load flexibility for power system control purposes depends, among other factors, on the design of the underlying electricity market, specifically on wholesale energy and ancillary service markets. While most of today’s wholesale markets are tailored for international electricity trading, ancillary service markets remain largely in the national domain. This implies that there is a large amount of diversity among the ancillary service market structures between countries. Reference [46] presents a comprehensive overview on different ancillary service market structures and nomenclatures. For instance, the reserve delivery times and specifications of control service quality are widely different. As an example, we use the Swiss market design as described in [180] as a basis for the evaluation of ancillary service provision. In other market settings, the used regulatory framework for revenue generation may be entirely different, so the methodology used here may have to be adapted in order to yield meaningful results.

The Swiss control reserves market is based on an auctioning system with a Pay-as-Bid pricing method [189] for primary, secondary, and tertiary control reserves. Bids can be entered online and are accepted or rejected by the TSO, who represents the only source of demand in these markets according to the calculated control reserve requirements for the control area. The auction result is not disclosed in detail, but the average price of the most expensive accepted 20 MW is published by Swissgrid [48].

The financial compensation for primary and secondary control is outlined below. Note that, due to the national nature of the ancillary service market, the capacity prices are issued in Swiss Franc (CHF). Since the energy price is based on the SwissIX spot market price of the European Power Exchange (EPEX) [190], it is issued in Euro (EUR).

7.3.1 Regulatory Basis for Revenue Calculation

**Revenue of Primary Control** Primary control is compensated by a capacity fee per MW of control band provided for a certain time span. Bids can only be placed symmetrically (equal control band for positive and negative deviation from the working point). Currently, the auctions take place once a week. Energy compensation is not used since system frequency deviation is relatively symmetrical and does not exhibit longer deviations from zero. The prices for primary control capacity in the year 2011 are depicted in Figure 7.3-a).
7.3. Modeling of Revenue Potential

Revenue of Secondary Control  In the Swiss ancillary service market, secondary frequency control is compensated by a capacity fee and an energy fee. While the capacity price per MW of control band provided for a certain time span is determined by the auction process, the energy price follows a fixed scheme: the control energy is averaged over a time slice of 15 minutes and paid for with an hourly spot market price including a bonus of ±20%. In case of a generation increase (or load decrease), the ancillary service provider receives the spot price +20%, in case of a generation decrease (or load increase) he pays the spot price −20%. In order to smoothen price spikes, the provider’s revenues are floored by the weekly base price and his costs are capped by the weekly base price. The prices for secondary control capacity in the year 2011 are presented in Figure 7.3-b). The energy prices based on the spot market prices are depicted in Figure 7.3-c).

Capacity Price Assumption  For the analysis of revenue potentials and the profitability of ancillary service provision by flexible unit portfolios, we consider the average price as depicted in Figure 7.3 for the primary and secondary control capacity. This is an idealistic assumption, since
in a Pay-as-Bid auction, perfect bidding (placement of the bid at the clearing price) is needed to achieve this result. The calculations consequently serve as a performance benchmark, not necessarily as a realistic prediction of achievable revenues. For reasons of comparability, we convert the prices from CHF to EUR. We assume a conversion factor of 1.25 CHF/EUR.

7.3.2 Net Operating Profit

Based on the market design presented above, we evaluate the maximum net profit that can be made (under the idealistic assumption of perfect bidding) by using a combined portfolio of power nodes to provide ancillary services. To calculate the operating profit $\Pi$ [EUR], we use the following formula:

$$\Pi = R_{\text{capa}} + R_{\text{en,net}} - C_{\text{strg}} - C_{\text{ramp}} - \Delta C_{\text{fuel}} - C_{\text{opp}},$$  

(7.1)

where $R_{\text{capa}}$ [EUR] is the capacity revenue of the ancillary service provision, $R_{\text{en,net}}$ [EUR] is the net control energy revenue, $C_{\text{strg}}$ [EUR] is the storage cycling cost, $C_{\text{ramp}}$ [EUR] is the ramping cost of the generator, $\Delta C_{\text{fuel}}$ [EUR] is the change in fuel cost through the ancillary service provision compared to a constant setpoint, and $C_{\text{opp}}$ [EUR] is the opportunity cost incurred by not using the generator control band for energy production. The individual revenue and cost terms are defined as follows:

$$R_{\text{capa}} = \pi_{\text{capa}} T_{\text{prov}} P_{\text{prov}},$$  

(7.2)

with the control reserve capacity price $\pi_{\text{capa}}$ [EUR/MW·h], the duration of reserve provision $T_{\text{prov}}$ [h], and the provided symmetrical power control band of width $P_{\text{prov}}$ [MW]. The following term describes the energy revenue (secondary control):

$$R_{\text{en,net}} = \sum_{k_{\text{MTU}}=1}^{K_{\text{MTU}}} 1.2 \pi_{\text{floored}} (k_{\text{MTU}}) E_{\text{feed-in}} (k_{\text{MTU}}) - \sum_{k_{\text{MTU}}=1}^{K_{\text{MTU}}} 0.8 \pi_{\text{capped}} (k_{\text{MTU}}) E_{\text{cons}} (k_{\text{MTU}}),$$  

(7.3)

with the capped and floored spot prices $\pi_{\text{spot}}^{\text{capped}}$ and $\pi_{\text{spot}}^{\text{floored}}$ and with the fed-in or consumed control energy (netted over market time units
7.3. Modeling of Revenue Potential

\( k_{\text{MTU}} = 1, \ldots, K_{\text{MTU}} \) of 15 minutes) \( E_{\text{net}}^{\text{feed-in}} \) and \( E_{\text{net}}^{\text{cons}} \). Describing the storage cycling cost for all time steps \( k = 1, \ldots, K \) is straight-forward:

\[
C_{\text{strg}} = \sum_{k=1}^{K} \left( \frac{1}{2} \pi_{\text{bat}} \Delta u_{\text{bat}}^{\text{gen}}(k) + \frac{1}{2} \pi_{\text{bat}} \Delta u_{\text{load}}^{\text{bat}}(k) \right) t_s ,
\]

where \( \pi_{\text{bat}} [\text{EUR/MWh}] \) is the marginal battery storage cost and \( t_s \) is the sampling time of the problem. The generator ramping cost is equal to

\[
C_{\text{ramp}} = \sum_{k=1}^{K} \pi_{\text{ramp}} t_s \left( \Delta u_{\text{gen}}^{\text{gen}}(k) - \Delta u_{\text{gen}}^{\text{gen}}(k - 1) \right)^2 ,
\]

where \( \pi_{\text{ramp}} [\text{EUR/MW}^2/\text{h}] \) is the marginal ramping cost. The change in fuel cost, depending on the fuel price \( \pi_{\text{fuel}} [\text{EUR/MWh}] \), is equal to

\[
\Delta C_{\text{fuel}} = \sum_{k=1}^{K} \pi_{\text{fuel}} t_s \Delta u_{\text{gen}}^{\text{gen}}(k) .
\]

Figure 7.4: Opportunity cost incurred by the generator due to the reservation of the control band
The opportunity cost incurred by the generator represents the financial loss associated with reserving the control band in the generator’s operating range. Let \( \alpha \in [0, 1] \) be the share of the control band that is covered by the energy-storing unit (battery, thermal load). As depicted in Figure 7.4, we consider the financial loss due to two effects: first, the reduction of the possible energy production range to the upside (by the positive control band) when the wholesale market price is higher than the fuel cost and the generator wants to produce at maximum power; and second, the forced production of energy (by the negative control band) when the market price is lower than the fuel cost and the generator wants to produce at minimum power. These two effects are mapped into an opportunity cost function which reads

\[
C_{\text{opp}} = \sum_{k=1}^{K} (1 - \alpha) t_s P_{\text{prov}} \cdot |\pi_{\text{spot}}(k) - \pi_{\text{fuel}}(k)|. \tag{7.7}
\]

For each time step, the energy portion \((1 - \alpha) t_s P_{\text{prov}}\) that either could have been produced for a profit (when \(\pi_{\text{spot}}(k) > \pi_{\text{fuel}}(k)\)) or that had to be produced for a loss (when \(\pi_{\text{spot}}(k) < \pi_{\text{fuel}}(k)\)) is summed up. Note that this kind of modeling constitutes a certain simplification of reality. To be exact, one would have to perform a unit commitment and economic dispatch of the power plant with and without the reserved control band, the difference of which constitutes the opportunity cost. In that way, start-up/shut-down and ramping can be considered. Since this would introduce an additional dependency on individual parameters of the power plant, we opt for the presented simplified approach. Note that another slight error is introduced in the ramping cost since it is quadratic and up- and down-ramping for working point changes would have to be superposed on the ramping caused by ancillary service provision. However, we do not deem this effect crucial for this study and consequently neglect it.

### 7.4 Simulation Study

In the following, we conduct a simulation study to evaluate the frequency control approach using the benchmark portfolios A (generator and battery) and B (generator and water heater population). We define simulation scenarios which shall enable an illustration and evaluation of the benefits arising from using energy storage for ancillary services.
7.4. Simulation Study

Table 7.2: Parameter variations for frequency control assessment

<table>
<thead>
<tr>
<th></th>
<th>Battery</th>
<th>Water heater</th>
</tr>
</thead>
<tbody>
<tr>
<td>Storage share ((\alpha) [%])</td>
<td>[10:10:90]</td>
<td>[10:10:90]</td>
</tr>
<tr>
<td>Energy cap. (C/\Delta u_{load}^{\text{max}}) [h]</td>
<td>[1, 2.5, 5, 7.5, 10]</td>
<td>[3:1:8]</td>
</tr>
</tbody>
</table>

Figure 7.5: Number of water heaters over parameter ranges

7.4.1 Simulation Scenarios

In order to quantify the benefits of energy storage for primary and secondary frequency control, we take the following approach: Real system frequency and secondary frequency control signal data (available in 10-second resolution over a time span of 30 days) are scaled to a benchmark control band of \(P_{prov} = \pm 10\, \text{MW}\). This is a realistic bid size in many ancillary service markets and is useful for easy comparison. We will simulate the system over a time span of 30 days with variations in power and energy capacity of the units. In particular, we are interested in the share of the control band that an energy-storing unit (storage device or thermal load) can securely account for, which provides a measure for the value of using the combination of units instead of the generator alone. The following parameters are varied: 1) the share \(\alpha\) of the control band that is covered by the energy-storing unit (defining its rated power), while the generator control band is scaled to \(1 - \alpha\), and 2) the storage capacity of the energy-storing unit, measured in hours of charging at rated power disregarding the efficiency.

The simulation scenario parameter sets are described in Table 7.2. The battery storage capacity is sized simply according to the desired duration of charging at maximum power. For the water heater population, a more detailed approach is necessary: since the water heater control band shall be symmetrical (up- and down-regulation of the power consumption equally possible), the autonomous energy demand of the water
heaters must be taken into account. In most cases, water heaters have a relatively low capacity factor (average power demand divided by rated power). The symmetry of the control band is achieved by selecting the working point such that the mean value of energy demand is consumed when the control signal is equal to zero. The daily draw profile is scaled such that 75% of the water contained in the tank is drawn during one day. These facts imply that for a larger storage capacity less water heaters are needed since the energy demand is higher. Figure 7.5 depicts the number of water heaters over the parameter ranges.

We present the most important economic parameters for the case study in Table 7.3. The cost function coefficients are defined in accordance with the portfolio definitions in Chapter 6. We carry out all simulations with a prediction horizon of $N = 2$ (time step of 10 s), which enables to honor inter-temporal and state constraints while acknowledging that control signals are hardly predictable over longer time horizons.

### 7.4.2 Numerical Results

We conduct the parameter variations described above in order to assess the value of the joint reserve provision. The following quantities are of interest for the performance evaluation:

a) deviation from the control signal (in the form of Root Mean Square Error (RMSE), normalized to the size of the control band),

b) mean storage State of Charge (SOC) level deviation from the working point,

c) mean storage power deviation from the working point,

d) mean generator power deviation from the working point,

e) cost incurred by the portfolio operation and revenue earned by the ancillary service provision,

f) net profit from the portfolio deviation,

g) net profit change (compared to ancillary service provision by generator only), and

h) net profit per MW of battery storage / per individual water heater.

<table>
<thead>
<tr>
<th>Table 7.3: Assumptions for economic parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spot market price and control signals</strong></td>
</tr>
<tr>
<td>Fuel price for generator</td>
</tr>
<tr>
<td>Capacity price Primary Control</td>
</tr>
<tr>
<td>Capacity price Secondary Control</td>
</tr>
</tbody>
</table>
Primary Control with Portfolio A

The provision of primary control reserves with Portfolio A (Battery + Generator) is depicted in Figure 7.6. The battery share is varied between 10 and 90% of the control band, and the battery capacity is varied between 1 and 10 h. The parts a) – h) of the figure show:

a) The control signal deviation is non-zero only for battery shares above 80%. Even then, the deviation stays small with a maximum of about 0.17%. This is due to the fact that the maximum frequency deviation rarely exceeds ±100 mHz, so only about half of the control band is utilized. The battery storage capacity does not have a strong impact on the control signal deviation, which also holds true for all other variables discussed below.

b) The mean SOC deviation from the scheduled working point is equal to zero below 70 – 80% since the storage is hardly used for the reserve provision. Above this storage share level, the battery storage operates mostly in its upper SOC region.

c) The mean storage power deviation becomes positive for a control band share above 80% as well, which corresponds to the increased mean SOC. Below 80%, it is equal to zero.

d) The mean generator power deviation is constantly equal to about −0.525 MW for power shares below 80%. This is due to the negative bias of the control signal, which is entirely accommodated by the generator. Above a level of 80%, the generator produces slightly more on average.

e) The totally incurred cost for the 30 days of operation (lower surface) ranges from about EUR 100,000 down to about zero. This is due to the decrease in opportunity cost as well as the negative bias of the control signal. The revenue (upper surface, from capacity bidding only since the control energy is not compensated in primary control) is constant at about EUR 220,000.

f) The net profit slopes up from about EUR 120,000 to almost EUR 220,000 along with a change from 10% to 90% of battery share.

g) The net profit change with respect to the base case (generator only) ranges from about 15 up to about 100%.

h) The net profit per MW of battery power is equal to about EUR 12,000 for the considered 30 days of operation for battery shares below 70%. Exceeding this share reduces the profit due to the increased cycling of the battery storage.
Figure 7.6: Simulation result: Primary Control with Portfolio A
Primary Control with Portfolio B

Figure 7.7 shows the aggregated simulation results of primary control provision with Portfolio B (Battery + Electric Water Heaters). As before, we vary the storage share between 10 and 90%. The energy capacity is varied from 3 – 8 hours for a symmetrical control band. The parts a) – h) of the figure show:

a) The control signal deviation starts at a storage share of about 50%. Below this level, the portfolio is able to fulfill the control task perfectly. The discrepancy to the results of Portfolio A can be explained by the fact that the water draw profile induces a variation of the SOC even if the control signal were equal to zero.

b) The average SOC deviation (from a base value of 0.6) is equal to 0.4 for small water heater shares. This is due to the fact that the optimizer chooses to move the SOC to that region in the beginning of the time simulation. The average SOC drops to about 0.1 for a share of about 60% and then rises again to almost 0.3 when the water heater share attains 90%. This behavior is relatively arbitrary and largely depends upon the decision variable penalization and control signal time series.

c) The average storage power deviation fluctuates between about 0.1 MW for storage shares of 10% and slightly above 0.8 MW for 90% and small energy capacity. This also depends on the decision variable penalties and control signal time series.

d) The average generator power deviation follows the same shape as the mean storage power deviation and fluctuates between about $-0.45$ MW and $+0.3$ MW.

e) The totally incurred cost is equal to about EUR 100,000 for a storage share of 10% and slopes down to about 15,000 for 90%. The revenue generated is constant over the entire parameter span and is equal to about EUR 220,000.

f) The net profit varies – consequently – from about EUR 120,000 to EUR 205,000 as the storage share increases.

g) The net profit change with respect to the base case (generator only) has a similar shape and varies between about 15% and 95%.

h) The net profit per water heater varies between about EUR 4 and EUR 15 for the considered 30 days of operation. The slightly counter-intuitive result of increasing returns for a larger storage capacity is due to the number of water heaters decreasing in this direction of the parameter variation (see Figure 7.5).
Chapter 7. Economic Evaluation of Frequency Control

Figure 7.7: Simulation result: Primary Control with Portfolio B
Secondary Control with Portfolio A

Figure 7.8 presents the provision of secondary control with Portfolio A (Battery + Generator). The battery power share and energy capacity are varied in the same way as before. We observe the following details in the simulation results:

a) The control signal deviation is only equal to zero for small battery shares. This is due to the fact that the energy contained in the secondary control signal quite easily exceeds the available storage capacity. Unlike primary control, a secondary control signal usually exhibits frequent attainment of positive and negative extremes, which implies the full activation of the available reserves.

b) The average SOC deviation ranges between about 0.1 and 0.4. The maximum values are attained on a ridge-like curve in the vicinity of the diagonal of the parameter space.

c) The mean storage power deviation is close to zero for small storage shares and attains increasingly higher values of up to 0.5 as the storage share increases to 90%. The storage utilization does not change significantly with the storage capacity.

d) The mean generator power deviation largely follows the same shape with a relatively constant offset. It varies between about −0.8 and −0.2, being slightly steeper than the surface in c) for large storage shares.

e) The totally incurred cost ranges between about EUR 90,000 to approximately EUR 45,000. Note that the slope of the surface in battery-share direction becomes positive again for battery shares above 60 – 70%. This is due to the increased storage cycling. The attained revenue is constant at about EUR 220,000.

f) The net profit consequently varies between a minimum of EUR 130,000 and a maximum of about EUR 175,000.

g) The net profit change with respect to the base case (generator only) varies between about 10% to almost 55%.

h) The net profit per MW of storage power is equal to almost EUR 12,000 for small storage shares and slopes down to about EUR 5,000 as the share increases. This indicates a decreasing marginal value of replacing generation capacity control band with energy storage since the actual utilization of the storage causes cycling cost and energy losses.
Figure 7.8: Simulation result: Secondary Control with Portfolio A
Secondary Control with Portfolio B

Figure 7.9 presents the provision of secondary control with Portfolio B (Generator + Electric Water Heaters). The following observations can be made in part a) – h) of the figure:

a) The control signal deviation remains close to zero for all water heater shares below 80 – 90%. The non-zero values for 8 hours of storage capacity and a storage share of about 70% can be regarded as a random event.

b) The average SOC deviation is approximately v-shaped in the direction of increasing storage shares. It is close to zero for small storage shares and slopes downwards to about −0.2 for a storage share of around 60%. Above that, the average SOC slopes up again to a maximum of about 0.2.

c) The mean storage power deviation has a similar shape, being close to zero for small storage shares, negative at about −0.3 MW for a storage share of about 60%, and positive at about 0.35 MW for a storage share of 90%.

d) The mean generator power deviation exhibits the same shape as c) with a relatively constant offset of about −0.8 MW.

e) The totally incurred cost in EUR slopes downward in the direction of the storage share, starting at EUR 95,000 down to slightly below zero for a storage share of 90%. The negative cost is due to the change in fuel cost induced by the reserve provision along with a lack of storage cycling cost.

f) The net profit consequently ranges from about EUR 125,000 to EUR 220,000 between small and large storage shares.

g) The net profit change with respect to the base case (generator only) ranges between 15% and about 85%.

h) The net profit per water heater ranges from about EUR 5 to EUR 14 for the 30 days of operation, depending on the parameter set. The explanation for the upward slope in the direction of the storage capacity is explained as for the case “Primary Control with Portfolio B”.
Figure 7.9: Simulation result: Secondary Control with Portfolio B
7.5 Profit Sharing Methodology

Having derived the attainable operating profit by two flexible benchmark portfolios, we proceed with the development of a methodology for sharing the profit between different players in the market system. For the representation of the considered market environment, we take advantage of a modeling methodology named $e^3value$ [191]. It serves to create so-called business value models which allow the description of exchange processes of services and financial compensations between various entities according to a certain market design. It has been successfully applied to the modeling of Distributed Generation (DG) business cases [192] and other business fields. In the context of the present work, modeling activities with the $e^3value$ methodology have been carried out in [193]. The purpose of this modeling is, in our case, the structured representation and illustration of the exchanges between the aggregator and its surroundings, and, in the case of privately owned thermal loads, the final customers.

We will only briefly introduce $e^3value$ and refer the interested reader to [191]. At its core, the methodology consists of a (graphic and mathematical) representation of actors, their various business activities, and the exchanges between them. This enables the calculation of the profitability for each of the actors. The methodology comprises the following principal elements which are also depicted in Figure 7.10:

**Actor:** An actor is an entity in the market that pursues certain goals by conducting business-related value activities.

**Value Activity:** A value activity is a set of business-related actions that generates a revenue or incurs a cost for a certain actor.

**Value Port:** A value port is a connector that enables a value exchange (i.e., exchange of certain goods/services/money) between different actors. These exchanges are associated with specific value activities.

![Figure 7.10: $e^3value$ elements](image-url)
Start Stimulus: A start stimulus is a principal need or desire of an actor for a certain good or service, which creates a demand and a chain of value exchanges.

End Stimulus: An end stimulus is the end of a value exchange chain.

AND and OR gates: These are routing elements that create logical relations between exchanges and enable the merging of exchange paths.

7.5.1 Business Value Model

The business value model of the ancillary service provision by an aggregator is depicted in Figure 7.11. We will introduce the various actors, activities, and exchanges in the following. Note that the present model only considers exchanges that are influenced by the ancillary service provision. A detailed analysis and parameterization of the base case (electricity system without the aggregator) as carried out in [193] is omitted here for shortness. For this reason, we also omit the modeling of the DSO since we do not consider the effects of the aggregator’s business on the distribution system in this study.

7.5.2 Actors and Activities

The following actors and activities are considered in the model:

Transmission System Operator (TSO): The TSO is responsible for controlling and operating the transmission grid (usually comprising the voltage levels of 220 kV and 380 kV in Europe). This includes monitoring and control of the current grid topology (position of breakers and switches within the grid) and the voltage in all parts of the transmission grid. The principal activity of interest for our analysis is the contracting of ancillary service providers by determining the required control reserve capacity, administrating the auction process for the reserve tendering, and also calling the reserves when needed. The ancillary services are paid for by grid users via grid usage fees levied by the TSO.

Aggregator (Agg): The Aggregator is a novel entity in the power system which operates between the TSO, Power Plants, Electricity Supplier, and Final Customer. It contracts both generation capacity and final customers in cooperation with the electricity suppliers in order to deliver ancillary services to the TSO.
Figure 7.11: Business value model for the aggregator
**Power Plant (PP):** The power plant produces and sells electricity via the wholesale electricity market. It also delivers ancillary services, in our case frequency control reserves, to the TSO. A bilateral contractual relation with the Aggregator serves to integrate a generator control band into the portfolio administrated by the Aggregator.

**Electricity Supplier (Supp):** The Electricity Supplier serves the final customer with electricity at a certain price. The contract modalities can be arbitrarily defined and usually consist of a certain base price, an energy price per kWh of used electricity, and in the case of large customers also a power capacity fee per kW of peak load. The electricity is usually bought on the electricity market or via bilateral contracts. Note that the electricity supply is not depicted in Figure 7.11 since it is not influenced by the Aggregator’s business model. The relevant activity for our analysis is Allowing Load Aggregation, which means that the Electricity Supplier consents and collaborates with the aggregator in contracting flexible loads of final customers. This can include the offering of a Demand Response (DR) tariff to the final customers, the financial benefits of which are enabled by the Aggregator’s business.

**Wholesale Market (WhMa):** The Wholesale Market is an electricity exchange where trading with other market actors takes place. As the other actors are not modeled, the Wholesale Market serves as a source of revenue generation.

**Equipment and ICT Supplier (ICT):** The Equipment and ICT Supplier delivers the necessary equipment and Information and Communication Technology (ICT) to the Aggregator.

**Final Customer (Cust):** The Final Customer contracts an Electricity Supplier for the supply with electric energy. The relevant activity is the response to a control signal sent by the aggregator, which is compensated by a response fee.

**Storage Owner (StrgOw):** The Storage Owner possesses and operates an energy storage device (in our case, a battery) that the Aggregator includes in his portfolio. The Storage Owner gets compensated by a fee for his services to the Aggregator.

**7.5.3 Exchanges**

Now the value exchanges between the actors are described and quantified. We refer to the net operating profit calculation presented in
Section 7.3.2 and make use of the simulation results from Section 7.4. A number of degrees of freedom exist, which call for assumptions and/or business model design decisions. We will outline one possible variant to design the value exchanges. In order to calculate the cash in- and out-flows of the actors, we will consolidate the exchanges for the different activities and actors. The following exchanges take place:

**Aggregator Capacity Fee (AggCapFee):** The capacity fee earned by the aggregator is equal to the overall capacity revenue earned by the portfolio described by (7.2):

\[
AggCapFee = R_{\text{capa}}.
\] (7.8)

**Aggregator Energy Fee (AggEnFee):** Similarly, the energy fee is equal to the total net energy revenue according to (7.4):

\[
AggEnFee = R_{\text{en,net}}.
\] (7.9)

**Power Plant Capacity Fee (PPCapFee):** We assume that, in the base case, the power plant offers 100% of its control capacity band (in our case ±10 MW) to the ancillary service market. In our considered business model, we assume that the generator does not bid directly into the ancillary service market anymore. The generator share of the control band \((1 - \alpha)\) is instead provided to the Aggregator, and the remaining share, \(\alpha\), is used for energy production for the Wholesale Market. This means that the power plant does not earn any capacity fee anymore. Compared to the base case \((\alpha = 0)\) where it used to earn \(R_{\text{capa}}\), the revenue is the negative of the original capacity fee:

\[
PPCapFee = -R_{\text{capa}}.
\] (7.10)

**Power Plant Energy Fee (PPEnFee):** The same is true for the energy fee:

\[
PPEnFee = -R_{\text{en,net}}.
\] (7.11)

**Control Service Compensation (CtrlServComp):** The control service compensation is an internal flow within the aggregator. It is equal to the revenue earned by the ancillary service provision:

\[
CtrlServComp = AggCapFee + AggEnFee.
\] (7.12)
Chapter 7. Economic Evaluation of Frequency Control

**Generation Coupling Fee (GenCoupFee):** The generation coupling fee shall compensate the provision of control capacity by the power plant. We dimension the fee such that the power plant does not incur any profit or loss with respect to bidding its capacity into the market. The main incentive for the power plant to participate in the proposed scheme is the reduction of its ramping activity which allows for a longer equipment life in thermal power plants. We use the aggregator’s control capacity and energy revenues as a basis and deduct the reduction in opportunity cost (with respect to the base case) due to using the control band share $\alpha$ for energy production. We also adjust for changes in fuel cost with respect to the base case. Note that further adjustments may have to be made for the changes in the power plant’s control energy revenue, which we neglect here for simplicity. Consequently, the generation coupling fee is defined as follows:

$$GenCoupFee = CtrlServComp - \alpha \cdot C^\text{base}_\text{opp} - \left(\Delta C^\text{fuel}_\text{base} - \Delta C^\text{fuel}_\text{base}\right).$$

(7.13)

**Administration Fee (AdminFee):** In order to incentivize the participation of the electricity supplier (only relevant for Portfolio B) and to cover costs incurred by the need to communicate with the customers, we propose the following administration fee as a percentage $\beta$ of the Aggregator’s revenue that can be attributed to the storage share $\alpha$:

$$\text{AdminFee} = \begin{cases} 
0 & \text{for Portfolio A} \\
\alpha \cdot \beta \cdot CtrlServComp & \text{for Portfolio B} 
\end{cases}.$$  

(7.14)

**Wholesale Electricity Fee (WhElecFee):** The wholesale electricity fee is defined as the storage control band share $\alpha$ times the opportunity cost incurred in the base case $C^\text{base}_\text{opp}$:

$$WhElecFee = \alpha \cdot C^\text{base}_\text{opp}.$$  

(7.15)

The rationale behind this is that a share $\alpha$ of the control band is now utilized for energy production instead of ancillary service provision, so the opportunity cost is reduced by this amount.

**Equipment and ICT Cost (EquipICTFee):** The equipment and ICT cost is incurred for the communication and control infrastructure in a centralized control center ($C^\text{fixed}_\text{ICT}$) and on the premises of the final
customers \( (n_{\text{CL}} C_{\text{ICT}}^\text{CL}, \text{where} n_{\text{CL}} \text{is the number of final customers}) \) or, respectively, on the premises of the storage owner \( (C_{\text{ICT}}^\text{bat}) \):

\[
\text{EquipICTFee} = \begin{cases} 
C_{\text{ICT}}^\text{fixed} + C_{\text{ICT}}^\text{bat} & \text{for Portfolio A} \\
C_{\text{ICT}}^\text{fixed} + n_{\text{CL}} C_{\text{ICT}}^\text{CL} & \text{for Portfolio B}
\end{cases}
\] (7.16)

**Control Response Fee (CtrlRespFee):** This fee is paid by the Aggregator to the final customers as a compensation for offering their controllable load capacity. We base the fee on the Aggregator’s revenue, multiplied by the storage share \( \alpha \) and another modifier \( \gamma \):

\[
\text{CtrlRespFee} = \begin{cases} 
0 & \text{for Portfolio A} \\
\alpha \cdot \gamma \cdot \text{CtrlServComp} & \text{for Portfolio B}
\end{cases}
\] (7.17)

**Storage Fee (StrgFee):** The storage fee compensates the operation of the energy storage device. We define it as the cycling cost plus the generator’s opportunity cost reduction \( \alpha C_{\text{base}}^\text{opp} \), modified by a factor \( \delta \):

\[
\text{StrgFee} = \begin{cases} 
C_{\text{strg}} + \alpha \delta C_{\text{base}}^\text{opp} & \text{for Portfolio A} \\
0 & \text{for Portfolio B}
\end{cases}
\] (7.18)

### 7.5.4 Cash Flow Consolidation

In accordance with Figure 7.11, we enumerate the actors, activities, and exchanges in the form of column vectors, which is practical for a mathematical representation of the exchanges and the attribution of cash flows to the activities and actors. The following vectors are defined:

\[
\text{actors}^\top = \left[ \text{TSO, Agg, PP, Supp, WhMa, ICT, Cust, StrgOw} \right], (7.19)
\]

\[
\text{activities} = \left[ \begin{array}{c} \text{CtrlResContr} \\
\text{CtrlServProv} \\
\text{Agg + Ctrl} \\
\text{AncServProv} \\
\text{EnergyProd} \\
\text{AllowingAgg} \\
\text{Trading} \\
\text{EquipICTSupp} \\
\text{RespToCtrl} \\
\text{StrgProv} \end{array} \right], \text{ exchanges} = \left[ \begin{array}{c} \text{AggCapFee} \\
\text{AggEnFee} \\
\text{PPCapFee} \\
\text{PPEnFee} \\
\text{CtrlServComp} \\
\text{GenCoupFee} \\
\text{AdminFee} \\
\text{WhElecFee} \\
\text{EquipICTFee} \\
\text{CtrlRespFee} \\
\text{StrgFee} \end{array} \right].
\]
In our case, the dimensions of the vectors are \( n_{\text{actor}} = 8 \), \( n_{\text{activ}} = 10 \), and \( n_{\text{exch}} = 11 \). The cash flow consolidation is performed by a linear mapping between actors, activities, and exchanges. We introduce the mapping matrices \( M_{\text{exch}} \) (\( n_{\text{activ}} \times n_{\text{exch}} \)) which maps the exchanges to the activities and \( M_{\text{act}} \) (\( n_{\text{actor}} \times n_{\text{activ}} \)) which maps the activities to the actors:

\[
\text{activities} = M_{\text{exch}} \cdot \text{exchanges}, \quad (7.20)
\]

\[
\text{actors} = M_{\text{act}} \cdot \text{activities}. \quad (7.21)
\]

In \( M_{\text{exch}} \), we denote an inflow of money into a value activity with +1 and the outflow out of a value activity with −1. Unrelated exchange/activity combinations get a 0. In \( M_{\text{act}} \), we assign a 1 to an activity pertaining to an actor and a 0 to unrelated activity/actor combinations. According to Figure 7.11, the matrices are defined as follows:

\[
M_{\text{exch}} =
\begin{bmatrix}
-1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & -1 & -1 & 0 & -1 & -1 \\
0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}, \quad (7.22)
\]

\[
M_{\text{act}} =
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}. \quad (7.23)
\]

Note that the columns of \( M_{\text{exch}} \) must sum to zero since every exchange makes a positive contribution to exactly one activity and a negative contribution to another. The columns of \( M_{\text{act}} \) must sum to one since every activity is associated with exactly one actor. We calculate the financial in- and out-flow \( \text{CashInOut} \) to and from each actor by applying the presented linear maps to the vector of financial exchanges \( \text{CashExch} \):

\[
\text{CashInOut} = M_{\text{act}} \cdot M_{\text{exch}} \cdot \text{CashExch}. \quad (7.24)
\]
7.5.5 Application Example

In order to demonstrate the application of the profit sharing methodology, we consider one example for each of the four simulated scenarios. For each portfolio, we take one data point with a relatively high net operating profit from the parameter scans described in Table 7.2. For both Portfolio A and B, we select a storage share of 60% ($\alpha = 0.6$) and a storage capacity of 5 h; in the case of water heaters, this corresponds to 8,955 units in the population, which means that one water heater contributes on average $\pm 670$ W of controllable power. We approximate the results for one year by multiplying the revenue earned during the 30 simulated days by 365/30. The following (exemplary) numerical values were selected for the exchange parameters: $\beta = 0.02$, $\gamma = 0.4$, $\delta = 0.775$, $C_{\text{ICT}}^{\text{fixed}} = 10,000$ EUR, $C_{\text{ICT}}^{\text{bat}} = 10,000$ EUR, and $C_{\text{ICT}}^{\text{CL}} = 10$ EUR.

Table 7.4 presents the financial results for the chosen scenarios as obtained from the simulations presented in Section 7.4. We will briefly discuss the results with respect to the differences between the scenarios.

- **Prim. A** (Primary control provision by Portfolio A): This scenario is characterized by the dominance of the generation unit in following the control signal, while the battery storage serves as a backup that is only used when needed. This is due to the control signal being small enough to be covered almost always by the generator and due to the cycling cost incurred by the storage. Changes in ramping cost, cycling cost, and fuel cost are small.

- **Prim. B** (Primary control provision by Portfolio B): The contribution of the water heaters is more significant here, since water heaters incur no cycling cost but are able to smooth the control signal, which in turn saves ramping cost for the generator.

- **Sec. A** (Secondary control provision by Portfolio A): Here, the battery storage is needed more often for control signal tracking, which incurs cycling cost and also drives the generator fuel cost up. Note that the energy revenue from secondary control is negative, which is due to the negative bias of the control signal and is compensated by fuel cost savings with respect to the case of not providing reserves at all (not presented in this table).

- **Sec. B** (Secondary control provision by Portfolio B): In this scenario, the control signal is smoothed by the water heaters, so less ramping cost is incurred. The fuel cost is driven by a small increase in thermal losses in the water heaters.
Table 7.4: Yearly financial results of the four selected scenarios

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_{\text{capa}}$</td>
<td>2,628,000</td>
<td>2,628,000</td>
<td>2,628,000</td>
<td>2,628,000</td>
</tr>
<tr>
<td>$R_{\text{en,net}}$</td>
<td>0</td>
<td>0</td>
<td>-68,985</td>
<td>-68,985</td>
</tr>
<tr>
<td>$C_{\text{strg}} - C_{\text{base strg}}$</td>
<td>30</td>
<td>0</td>
<td>100,980</td>
<td>0</td>
</tr>
<tr>
<td>$C_{\text{ramp}} - C_{\text{base ramp}}$</td>
<td>-23</td>
<td>-60,306</td>
<td>-4,000</td>
<td>-12,116</td>
</tr>
<tr>
<td>$\Delta C_{\text{fuel}} - \Delta C_{\text{base fuel}}$</td>
<td>-19</td>
<td>2,032</td>
<td>46,095</td>
<td>6,281</td>
</tr>
<tr>
<td>$C_{\text{opp}} - C_{\text{base opp}}$</td>
<td>-879,167</td>
<td>-879,167</td>
<td>-879,167</td>
<td>-879,167</td>
</tr>
</tbody>
</table>

| II - II<sub>base</sub> per water heater | 879,179 | 937,441 | 736,092 | 885,003 |
| II - II<sub>base</sub> per battery MW | -104.68 | -98.83 | 146,529 | 122,682 |

Table 7.5: Cash in-/out-flows in the four selected scenarios

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Transmission System Operator</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Aggregator</td>
<td>123,418</td>
<td>115,329</td>
<td>123,418</td>
<td>128,465</td>
</tr>
<tr>
<td>PowerPlant</td>
<td>46,095</td>
<td>2,032</td>
<td>46,095</td>
<td>6,281</td>
</tr>
<tr>
<td>Supplier</td>
<td>0</td>
<td>31,536</td>
<td>0</td>
<td>30,708</td>
</tr>
<tr>
<td>Wholesale Market</td>
<td>-879,167</td>
<td>-879,167</td>
<td>-879,167</td>
<td>-879,167</td>
</tr>
<tr>
<td>ICT Provider</td>
<td>20,000</td>
<td>99,550</td>
<td>20,000</td>
<td>99,550</td>
</tr>
<tr>
<td>Final Customer</td>
<td>0</td>
<td>630,720</td>
<td>0</td>
<td>614,164</td>
</tr>
<tr>
<td>Storage Owner</td>
<td>689,655</td>
<td>0</td>
<td>689,655</td>
<td>0</td>
</tr>
<tr>
<td>Sum</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 7.12: Cash in-/out-flows in the four selected scenarios
Table 7.5 presents the monetary exchanges between the actors in numerical form, which is also visualized graphically in Figure 7.12. It can be seen that, for our parameterization of the value exchanges, there is a win-win situation for the participating actors. The additional monetary inflow into the system arises due to the fact that the power plant can use its generation capacity for energy production and sales on the market, as opposed to reserving the control band for ancillary service provision. This is depicted by a negative bar for the actor “Energy Exchange”, which obviously is to be interpreted as the source of additional revenue, not a financial loss incurred by the energy exchange.

For Portfolio A, most of the additional revenue is going to the storage owner (about EUR 690,000). The economic viability of operating the battery for this kind of revenue largely depends on the cost structure of the technology and the considered business model (ownership of the storage, buying or renting, combined use cases, etc.). At the present cost of about US$ 1,500 per kW of installed power [194] (amounting to US$ 9 million for a 6 MW battery), the economic viability may be questionable, but prices are expected to decrease steeply in the future and multi-objective utilizations of the battery can be considered. Details on this question are beyond the scope of this thesis. Furthermore, the ICT supplier earns a flat fee of EUR 10,000 for the control center equipment and another EUR 10,000 for the on-site equipment. The aggregator generates a revenue of about EUR 123,000.

For Portfolio B, the additional revenue is distributed between the final customers (about EUR 631,000 for primary and EUR 614,000 for secondary control), the electricity supplier (about EUR 31,000), the ICT equipment supplier (about EUR 100,000), and the aggregator (about EUR 115,000 for primary and EUR 128,000 for secondary control).

The amount of money transferred to the final customers corresponds to EUR 70.43 (CHF 88.04 with our assumed exchange rate) for primary control and EUR 68.58 (CHF 85.73) for secondary control per year for a single water heater. In relation to the total cost of electricity of an exemplary household consuming 5,000 kWh per year at CHF 0.14/kWh (average price) with an assumed monthly base fee of CHF 10.00, this amounts to 10.5% – 10.7% of the electricity bill. Note that we assume an unchanged electricity cost for the water heater operation – which may imply that cost differences induced by time-variable tariffs have to be compensated by the electricity utility, which in turn requires a redistribution of the profits generated by the ancillary service provision.
If this financial compensation is sufficient for incentivizing customer participation will largely depend upon the customers’ motivation or reluctance to engage in a novel utilization of their domestic appliances which might have noticeable effects on the appliance behavior (e.g., by additional ON and OFF switchings during the day). The trust that the customers have in their electricity suppliers and in novel players in the electricity market is therefore a major factor for successful implementation. By means of a suitable marketing strategy (such as pointing out the advantages of a “SmartGrid” for the transition towards a clean energy supply), an adoption of the proposed scheme may be achieved.

7.6 Concluding Remarks

We discussed the integration of flexible portfolios consisting of controllable loads and storage devices in conjunction with generation units into frequency control structures. A business value model including a methodology for sharing the profits between different actors based on the $e^3value$ framework was created and parameterized for the control reserve provision by a flexible unit portfolio administered by an aggregator. The optimization-based control strategy derived in Chapter 6 proved to be useful for distributing the control actions in an economically optimal way on the participating units while honoring state constraints of storage devices. We find that the simulation-based sizing of the storage units for frequency control is a promising approach since it takes into account time series properties, such as the energy contained in the control signal, as well as the duration of deviations from zero.

The simulated scenarios show that the combination of generators and energy-storing units yields a substantial net operating revenue increase which may be used to amortize the investment in the joint portfolio. For a real deployment of the described control systems, further investigations of the storage level management over time, accurate modeling of battery degradation, and ancillary service bidding strategies are necessary. Furthermore, the simplifying assumptions in this work should be addressed in future research. For instance, the market prices of both ancillary service and wholesale markets are assumed to be unaffected by the additional ancillary service provision by the aggregator. While this is realistic for a small and unique aggregator in a large system, the widespread implementation of such systems will have an impact on the overall market liquidity and price structure. The opportunity cost incurred by the generator should also be subject to further investigations.
Chapter 8

Customer-Level Under-Frequency Load Shedding

In this chapter, a novel Customer-Level Under-Frequency Load Shedding (CL-UFLS) approach is introduced. After a brief review of Conventional Under-Frequency Load Shedding (C-UFLS) approaches from literature, the motivation for a new system design is discussed and the necessary household infrastructure for implementing such a scheme is outlined. After that, some design principles and a concrete proposal of the information flow and the load shedding frequency threshold design is presented. Dynamic power system simulations illustrate the effectiveness of the approach.

8.1 Introduction and Motivation

Under-Frequency Load Shedding (UFLS) is a long-known means to counteract a dangerous frequency decay in electric power systems and to restore the balance between load and generation in emergency situations. It serves as a “last line of defense” against a system-wide disturbance that could easily lead to a complete collapse of the power system. This underlines the importance of UFLS for system security. However, the downside of conventional under-frequency load shedding approaches (denoted as C-UFLS here) is the loss of power supply as well as Distributed Generation (DG) in the parts of the grid that are
affected by shedding actions. This motivates the Customer-Level Under-Frequency Load Shedding (CL-UFLS) approach which is presented in this chapter.

This work is largely based on [195] and [196]. A preceding study with a simplified dynamic model of a power system was presented in [197]. The chapter is structured as follows: In the remainder of Section 8.1, we present some literature on Conventional Under-Frequency Load Shedding (C-UFLS) and introduce the CL-UFLS ideas. Section 8.2 presents the considered infrastructure on the customers’ premises and Section 8.3 introduces some basic principles for designing the CL-UFLS system. Section 8.4 presents a mathematical formulation of the load shedding problem, focusing on the description of the remaining load per bus after a certain shedding action, both for C-UFLS and CL-UFLS. Section 8.5 presents a possible layout and organizational structure of the CL-UFLS system and Section 8.6 discusses the precomputation of frequency thresholds. In Section 8.7, a simulation example with detailed numerical parameterization is presented in order to demonstrate the impact of the CL-UFLS approach. Section 8.8 offers some concluding remarks. The used notation is summarized in Table 8.1.
8.1.1 Conventional UFLS Approaches

We start with a review of conventional UFLS literature. Although there are a number of central load shedding approaches in which the loads are tripped based on the on-line minimization of an objective function as in [198, 199], most systems in use nowadays utilize under-frequency relays in substations which trip certain distribution feeders or entire transformers if an under-frequency situation is detected. Various methods for the decentralized measurement of the instantaneous power system frequency are available (see, e.g., [200] for the commonly used Phase-Locked Loop (PLL) measurement scheme, and [201] for a brief discussion of other available techniques and an extended PLL-based methodology). These techniques are implemented in a wide range of commercially available under-frequency relays.

The C-UFLS schemes implemented in different geographical regions are diverse with respect to the handling of an under-frequency situation. They may include only a frequency measurement \( f \) and the step-wise shedding of load according to a given stage plan as outlined in [175]. A common extension, often termed “adaptive UFLS”, is the estimation of the frequency decay gradient \( df/dt \) in order to assess the lack of active power. See, e.g., [202, 203, 204] for a discussion of such approaches.

8.1.2 Motivation for CL-UFLS

In most of the load shedding schemes discussed above, the shedding action is implemented through frequency relays which trip entire distribution feeders in the case of a disturbance. The relays may be located on the lower-voltage side of the High Voltage (HV)/Medium Voltage (MV) transformers, or on subsequent MV/MV or MV/Low Voltage (LV) transformers in the distribution network. Thus, entire portions of the distribution network are de-energized by the trippings caused by the under-frequency relays. This means that the consumers on these feeders are not supplied with any electricity.

At the same time, all DG units such as smaller wind farms, decentralized Combined Heat and Power (CHP) plants, or photovoltaic (PV) generation units on the shed feeders are lost. In the presence of large amounts of DG, the intended load shedding schemes can thus lead to significant involuntary generation shedding. This problem was described in, e.g., [205]. Furthermore, the decision to trip a certain distribution feeder implies the shedding of a relatively large amount of power. This is due
to the load shedding relays being located on a relatively high level in the vertical structure of the power system. If the location of the shedding action was more oriented towards the LV side of the grid, i.e., closer to the customer level, this would allow a more fine-tuned load reduction and the avoidance of over-shedding. At the same time, a higher number of frequency relays would have to be installed, which makes the system more expensive and possibly cumbersome to maintain.

One of the solution strategies is the migration of the shedding mechanism to the customer level utilizing two-way communication channels to individual appliances. This would allow to keep the distribution feeders energized and possibly installed DG units connected while minimizing the impact on the customer. First approaches to this issue exist in the literature [206]. Apart from UFLS, the problem of under-voltage load shedding can also be regarded on the customer level [207].

One important distinction has to be made with respect to frequency-dependent load control of consumer appliances. In the literature, there are numerous examples of the consideration of frequency-responsive cooling and heating loads which shall contribute to frequency stability. The original idea goes back to [70], where the term Frequency Adaptive Power-Energy Re-scheduler (FAPER) for such appliances is introduced. This idea has been picked up in a lot of current research such as [208, 209]. The operational behavior of such schemes is significantly different from the case considered here, although both deal with a customer-level frequency measurement and decentralized reaction. In the FAPER-like approaches, the control contribution comes from a frequency dead-band which works in both directions and is small enough to make an impact also for normal-operation frequency deviations. This is not the case in this setup, which is tailored for the rejection of larger disturbances without automatic re-activation after the shedding.

8.2 Household Infrastructure

The CL-UFLS methodology utilizes a communication system which is usually referred to as “SmartHome” infrastructure [210]. The idea behind this is the connection of individual electric appliances on the consumers’ premises via in-house communication systems. This provides novel functionalities for the customers, such as comfort features related to home automation, energy efficiency monitoring, and visualization of domestic electricity consumption. An information gateway to the
outside of the household can enhance comfort functionalities, e.g., by remote control of home appliances. At the same time, sophisticated Demand Response (DR) approaches requiring two-way communication are enabled. Novel business cases may be created, e.g., the provision of control reserves by aggregated and coordinated groups of flexible household appliances [93]. CL-UFLS provides an additional benefit by increasing the security of supply while utilizing the same infrastructure.

While the idea of building automation is already omnipresent in today’s commercial and industrial facilities, home automation approaches have not yet managed to pass the threshold towards a mass implementation in consumer households. This is mainly due to high cost, lack of standardization, and cumbersome installation and maintenance of in-house communication and control systems. However, recent developments show significant advances towards commercialization and mass implementation of home automation enabling technologies, e.g., Powerline Communication (PLC) or wireless communication systems that can be integrated into home appliances, consume small amounts of power, and may be produced in large quantities at low cost [210, 211]. These developments appear to support the idea of a future wide-spread implementation of “SmartHome” technologies.

8.3 System Design Principles

Several degrees of freedom exist in the design of the CL-UFLS system. In a large power system with several million appliances spread throughout the system in individual households and commercial buildings, the flows of information have to be designed such that sensitive information is protected, the system is ensured to react as expected, clear advantages over more conventional approaches exist, and the operation is feasible from a practical point of view. These aspects are addressed by the following design principles which characterize the load shedding system developed in this work:

1) non-transparency of the customer,
2) centralized off-line computation, decentralized on-line disturbance reaction,
3) category-wise appliance clustering,
4) consideration of the Value of Lost Load (VOLL) or similar ranking criteria,
5) robustness towards varying system inertia, and
6) emulation of ramp-wise load shedding for smooth stabilization of
the frequency decay.

Each of these points is justified and discussed with respect to its implications in the following.

### 8.3.1 Non-Transparency of the Customer

The first and obvious requirement established for the CL-UFLS is that the customer should not lose privacy with respect to the possession and operation of appliances present on his or her premises. In fact, the adequate consideration of privacy concerns is recognized to be a major factor in the design of “SmartHome” and “SmartGrid” technologies [212]. This point is taken into account through the anonymous collection of aggregated data without the necessity to directly address and track individual households or even appliances.

### 8.3.2 Centralized Off-Line, Decentralized On-Line

As a centralized triggering of shedding actions on individual consumers’ premises may be too slow and unreliable for UFLS, an autonomous reaction of the consumer loads is considered. It is assumed here that the system frequency can be measured locally in residential households and commercial buildings by the in-house equipment and shedding actions according to pre-set frequency thresholds can be triggered. As accurate “rate of change of frequency” \((df/dt)\) measurements are quite costly, the CL-UFLS system shall not rely on this information. Furthermore, it is clear that a 100% system penetration with CL-UFLS capability may take a long time to realize. Thus, the system has to be designed such that it can coexist with existing feeder-wise load shedding schemes.

### 8.3.3 Category-Wise Appliance Clustering

In order to avoid a large number of decision variables for the setting of frequency thresholds, the appliances should be clustered into categories. Due to consumer-induced usage profiles, it appears practical to define the categories according to the function of the appliances. These categories account for specific portions of load during the day, which implies that the CL-UFLS should be adaptable according to the load situation.
Another advantage is that the value associated with operating these similar appliances at a specific time can be assumed approximately uniform. This implies a constant VOLL for the appliances within the same category.

### 8.3.4 Consideration of the “Value of Lost Load”

The VOLL is a well-established term for the cost of a certain electricity outage. It is usually given in monetary units per kWh of energy not supplied to a customer. Different methodologies exist for such a quantification, such as the assessment of costs through macro-economic modeling, surveys on the customers’ directly incurred costs or on the willingness to pay for the avoidance of an outage, or empirical studies on the effects of major blackouts. Considering additionally the regional diversity of consumer structures and the usage of electricity, it is close to impossible to state universally applicable numerical values for outage costs. This argument is supported by the high diversity of VOLL assessments in the international literature, which is summarized in several surveys combining these results, e.g., [154, 156]. It can be seen that even the relation between residential and industrial load can be significantly different, as can be seen, e.g., in [155] which states a relatively high value for the residential VOLL.

These findings suggest that, for the purpose of this work, an assumption has to be made concerning the VOLL of different customer categories and load classes. This can be adapted towards more reality-based values when a more detailed study for a specific country is conducted. The concept shall be designed such that a change of VOLL assessments can be integrated easily into the system.

### 8.3.5 Robustness Towards Varying System Inertia

The splitting of a power system during a disturbance, as well as a high penetration of wind generators and inverter-connected DG, can lead to a significant reduction of the electromechanical inertia of the power system. Naturally, a small inertia and the correspondingly higher volatility of the system frequency provides a challenge to any kind of UFLS system which cannot detect the under-frequency situation and perform the tripping arbitrarily fast. However, the load shedding system should be designed such that a variable system inertia within reasonable bounds is not a major problem that would impede its effectiveness.
Figure 8.1: Illustration of the effect of ramp-wise load shedding by assigning a random frequency threshold from a uniform distribution to individual units (power ratings: 2 – 6 kW for water heaters, 80 – 120 W for cooling appliances)

8.3.6 Emulation of Ramp-Wise Load Shedding

In the case of conventional under-frequency relays in substations, the reduction of system load has to be performed in a sequence of relatively large load steps. This is due to the amount of aggregated load behind a single frequency relay which can either be shed or not shed. However, when a frequency response mechanism is present on the appliance level, this can be exploited for a much smoother disturbance reaction. One way is the assignment of random frequency threshold values to all appliances within one class. In order to approximate a ramp-wise load reduction between two fixed frequency values, these random numbers should be drawn from a uniform distribution between those boundaries.

Figure 8.1 illustrates the effect of uniformly distributed random values, in this case between 49 and 48.7 Hz (considering a nominal power system frequency of 50 Hz). In the left column, the device category of typical water heaters is considered (equal shares of 2, 3, 4, 5, and 6 kW rated power). In the right column, typical cooling appliances are
shown (consisting of equal shares of 80, 90, 100, 110, and 120 W rated power). In both cases, the aggregation of 1,000, 10,000, and 100,000 units is considered. In the top plots, the shed power in per unit over the attained system frequency is illustrated. In the bottom plots, the deviation from an ideal ramp is shown, confirming the intuitive notion that a better approximation is achieved by higher numbers of appliances. The maximum deviation of about 2% for 1,000 appliances and much less than 1% for 10,000 appliances and more shows that it is justified to consider the shedding by uniform random values approximately ramp-wise. Comparing the right and the left plots, it is apparent that the absolute power rating of the appliances is of low relevance to this matter.

8.4 Load Shedding Problem Formulation

In this section, the load shedding problem under consideration is formulated mathematically. The main goal is to describe both conventional load shedding based on frequency relays in distribution feeders as well as the newly introduced CL-UFLS of individual appliances, while taking into account the distribution of the load on the network buses. For simplicity and transparency, all calculations related to load shedding are performed in MW instead of per unit. Note that we use 50 Hz as the nominal system frequency.

8.4.1 Power System Components

Considered is a power system consisting of several transmission lines which interconnect a number of buses with attached central generators and distribution feeders. These feeders can be considered as a model for “portions” of load that can be disconnected by a single frequency relay. That is, if the shedding in the underlying distribution system is realized by the tripping of entire MV/MV or MV/LV transformers, these can also be described as feeders since the distribution network structure is not modeled for the purposes of this work.

Apart from the feeder subdivision, the load at each bus can be subdivided into load classes according to the associated functions that the load fulfills (such as cooking, washing, freezing, etc.). In this formulation, the load classes constitute a categorization which is distinct from
the distribution feeder subdivision. As the load class is a virtual categorization while the feeder subdivision represents the physical structure of the system, the load classes can be considered to be distributed on the individual feeders as well. This is illustrated in Figure 8.2.

Equations (8.1) – (8.3) show the nomenclature for the buses \( n \), the load classes \( i_n \), and the feeders \( j_n \) at bus \( n \):

\[
\begin{align*}
\text{Bus: } n & \in \mathcal{N} = \{1, \ldots, N\}, \\
\text{Load class at bus } n: i_n & \in \mathcal{L}_n = \{1, \ldots, N_{L_n}\}, \\
\text{Distribution feeder at bus } n: j_n & \in \mathcal{F}_n = \{1, \ldots, N_{F_n}\}.
\end{align*}
\]

For simplicity, the load classes are defined independently of the buses, so the number of load classes \( N_{L_n} \) is equal for all \( n \) and they represent the same kind of load (e.g., residential cooling load). Note that this does not imply that all load classes have to be actually present at all buses.

### 8.4.2 Load Ratios

The load at bus \( n \), \( P_{L,n} \), can be modeled by commonly used aggregated load models for power system studies, e.g., as found in [213]. In principle, each feeder and load class could be represented by a different load model according to its characteristics, although this might only be needed for a detailed study of a specific system. In this work, we will
represent the load at one bus by the same model. As the load shedding system will act independently of the load characteristics, the static load $P_{L0,n}$ at bus $n$ is considered for the computations. The same holds for the load classes and the feeder loads, as stated below.

For the subdivision of $P_{L0,n}$ into load classes $P_{L0,i_n}$, it is practical to define the ratios $r_{L,i_n}$ between the individual load classes $i$ on bus $n$ and the total bus load. The load distribution onto the different feeders $j$ at bus $n$ is handled in the same way:

$$r_{L,i_n} = \frac{P_{L0,i_n}}{P_{L0,n}}, \quad (8.4)$$
$$r_{F,j_n} = \frac{P_{F0,j_n}}{P_{F0,n}}. \quad (8.5)$$

In order to render the notation more compact, these ratios can be written as ratio vectors for each bus $n$ in the system, the 1-norm of which has to be equal to 1 by definition:

$$r_{L,n} = [r_{L,1_n}, \ldots, r_{L,N_{L_n}}]^T, \quad (8.6)$$
$$r_{F,n} = [r_{F,1_n}, \ldots, r_{L,N_{F_n}}]^T. \quad (8.7)$$

In the case of the CL-UFLS, one further detail has to be considered: the total penetration of the CL-UFLS system into a certain appliance class may be much less than 100% (for instance, only half of the refrigerators in a country could be equipped with the household load shedding capability). For this reason, the penetration ratio $r_{pen,i_n} \in [0, 1]$ is defined, which indicates the ratio of appliances in a class equipped with CL-UFLS capability. This can also be gathered in the vector

$$r_{pen,n} = [r_{pen,1_n}, \ldots, r_{pen,N_{L_n}}]^T. \quad (8.8)$$

The load shedding system penetration is assumed to be uniform throughout the system, i.e., $r_{pen,n}$ is equal for all $n$.

### 8.4.3 Distributed Generation Model

DG units in the distribution network can have a significant impact on system behavior in the case of a UFLS event. For the purposes of this work, these are modeled as negative constant-power loads and are distributed on the distribution feeders introduced above. Further dynamics of the DG units are not considered, as they are assumed to be connected via further transformers which provide a certain amount of
decoupling through their inductance. Note that the focus is here on the interaction of DG with UFLS, and that autonomous disconnections of the DG through unit protection devices are not considered.

The power that is injected by DG into bus \( n \) is equal to \( P_{DG,n} \), which is a value that can be set according to the desired DG penetration. In the same way as in (8.5), the distribution on the feeders emanating from bus \( n \) is described by the ratio

\[
 r_{DG,jn} = \frac{P_{DG,jn}}{P_{DG,n}} ,
\]

or in vector notation for bus \( n \):

\[
 r_{DG,n} = [r_{DG,1n}, \ldots, r_{DG,N_{F_n}}]^T .
\]

### 8.5 Load Shedding System Layout and Flow of Information

Based on the design principles outlined above, a concrete implementation of the CL-UFLS system and the flow of information can be designed. The communication system needed for implementing the scheme under consideration may run on a variety of hardware platforms. It has to support the following characteristics:

1. Assignment of a random frequency threshold to each individual appliance based on a uniform probability distribution determined by the load shedding system administrator, e.g., grid control center. This requires the clustering of appliances according to their load class (e.g., stoves, refrigerators, lighting, etc.). Essentially, the vectors \( r_{L,n} \) and \( r_{F,n} \) as described in (8.6) and (8.7) have to be obtained from aggregated measurements or estimations.

2. Decentralized measurement of the power system frequency on the household level. This can be performed either in the individual appliances or in the customers’ switchboards.

3. Comparison of the currently measured power system frequency with the assigned threshold (including a delay for secure detection) and triggering of the disconnection of the appliance in case of a system frequency below the threshold value.
The conventional load shedding is assumed to stay in place and has to exhibit almost the same characteristics (with the exception that the thresholds are not derived from a probability distribution).

Figure 8.3 depicts the flow of information between the load shedding administrator and the units. The information obtained from the units and the subsequent assignment of the conventional load shedding thresholds $f_{thr,j_n}$ is visualized, as well as the customer-level thresholds $f^{cl,1}_{thr,i_n}$ and $f^{cl,2}_{thr,i_n}$, which are the boundaries of the uniform distribution from which the individual appliance thresholds are taken (cf. Section 8.3.6). The time delays $\Delta t_{fr, detect,j_n}$ and $\Delta t_{fr, detect,i_n}$ are further design parameters which are explained in the next sections.

### 8.5.1 Conventional Load Shedding

The conventional load shedding system is composed of frequency relays (denoted by superscript fr) assumed to be present in every distribution feeder of the entire system. An arbitrary relay on feeder $j_n$ at bus $n$ is characterized by three parameters: the threshold frequency $f_{thr,j_n}^{fr}$, the detection time delay $\Delta t_{fr, detect,j_n}$ and the tripping time delay $\Delta t_{fr,trip,j_n}$. In case an under-frequency situation is detected, the shedding command is issued after $\Delta t_{detect,j_n}^{fr}$ has passed and the measured frequency $f_n$ did not return above the threshold. Once the shedding command has been issued, the actual tripping is performed after $\Delta t_{trip,j_n}^{fr}$. The behavior of the frequency relay can be illustrated by the finite state machine shown in Figure 8.4.
We now define the remaining load at one feeder as the difference of the load before the shedding and the shed load:

\[ P_{F0,rem,j_n} = P_{F0,j_n} - P_{F0,shed,j_n} . \]  

Corresponding to the load class and feeder ratios defined above, the shedding ratio of the feeder \( j_n \) is defined as the relation of the currently shed load value to the value before the shedding:

\[ r_{fr,shed,j_n} = P_{L0,shed,j_n}/P_{L0,j_n} . \]  

As a single feeder is either shed or not shed, \( r_{shed,j_n} \in \{0, 1\} \) must hold. The following equations describe the logic of the shedding, taking into account the theoretical shedding (theo) which would be caused in the case of instantaneous shedding, the shedding if only detection time delay was considered (detect), and the true shedding including the tripping delay:

\[ r_{fr,theo,shed,j_n}(t) = \begin{cases} 0 & \text{for } f_n(t) > f_{thr,j_n} , \\ 1 & \text{for } f_n(t) \leq f_{thr,j_n} , \end{cases} \]  

\[ r_{fr,detect,shed,j_n}(t) = r_{fr,theo,shed,j_n} \max_{\tilde{t} \in [t - \Delta t_{detect,j_n}, t]} f_n(\tilde{t}) , \]  

\[ r_{fr,shed,j_n}(t) = r_{fr,detect,shed,j_n}(t - \Delta t_{trip,j_n}) . \]  

Having established the requirement that the load should not be reactivated automatically once the frequency returns above the threshold, the shedding ratio is restricted to be monotonically increasing, i.e., for the time instants \( t_1 \) and \( t_2 \) with \( t_1 \leq t_2 \) holds

\[ r_{fr,shed,j_n}(t_1) \leq r_{fr,shed,j_n}(t_2) . \]
The main design parameters for the load shedding system are thus the frequency thresholds and the detection time delays of the frequency relays. The tripping time delay represents the latency of the frequency relay and can only be influenced by alteration of the tripping mechanism.

### 8.5.2 Customer-Level Load Shedding

The shedding ratios of the CL-UFLS (denoted by superscript cl) are determined as follows: Based on the ramp-wise load shedding emulation outlined in Section 8.3.6, two frequency thresholds have to be given to each load class. The actual threshold of each appliance is then determined on the household level. In an analogous way to the conventional load shedding, we define the shed and remaining load caused by one load class \( i_n \) at bus \( n \), as well as the shedding ratio:

\[
P_{L0,\text{rem},i_n} = P_{L0,i_n} - P_{L0,\text{shed},i_n} \quad , \quad (8.17)
\]

\[
r_{\text{shed},i_n}^{\text{cl}} = P_{L0,\text{shed},i_n}/P_{L0,i_n} \quad . \quad (8.18)
\]

As the load shedding is performed ramp-wise, \( r_{\text{shed},i_n} \in [0,1] \) holds in contrast to the conventional load shedding. The following equations describe the logic of the shedding, taking into account the theoretical shedding (theo) which would be caused in the case of instantaneous shedding, the shedding if only detection time delay was considered (detect), and the true shedding including the tripping delay:

\[
r_{\text{shed},i_n}^{\text{cl, theo}}(t) = \begin{cases} 
0 & \text{for } f_n(t) > f_{\text{thr},i_n}^{\text{cl},1} \\
\max_{\tilde{t}\in[t-\Delta t_{\text{detect},i_n},t]} f_n(\tilde{t}) & \text{for } f_{\text{thr},i_n}^{\text{cl},1} < f_n(t) \leq f_{\text{thr},i_n}^{\text{cl},2} \\
r_{\text{pen},i_n}^{\text{cl}} & \text{for } f_n(t) \leq f_{\text{thr},i_n}^{\text{cl},2}
\end{cases} \quad , \quad (8.19)
\]

\[
r_{\text{shed},i_n}^{\text{cl, detect}}(t) = r_{\text{shed},i_n}^{\text{cl, theo}}(t - \Delta t_{\text{detect},i_n}) \quad , \quad (8.20)
\]

\[
r_{\text{shed},i_n}^{\text{cl}}(t) = r_{\text{shed},i_n}^{\text{cl, detect}}(t - \Delta t_{\text{trip},i_n}) \quad . \quad (8.21)
\]

As in the case of conventional load shedding, the shedding ratio is restricted to be monotonically increasing, i.e., for the time instants \( t_1 \) and \( t_2 \) with \( t_1 \leq t_2 \) holds

\[
r_{\text{shed},i_n}^{\text{cl}}(t_1) \leq r_{\text{shed},i_n}^{\text{cl}}(t_2) \quad . \quad (8.22)
\]
8.5.3 Remaining Load at Bus $n$

Any combination of the previously modeled C-UFLS system and the newly introduced CL-UFLS system has to take into account the fact that the same load can only be tripped once. Thus, the effects of both load shedding systems have to be concatenated. This requires one further assumption, namely the distribution of one individual load class at one bus onto the feeders at this bus. It is assumed here that a certain load class $P_{L,i_n}$ at bus $n$ is equally distributed onto the $N_{F_n}$ feeders at this bus, i.e.,

$$P_{LF,i_n,j_n} = P_{L,i_n}/N_{F_n} \quad \forall i_n \in \mathcal{L}_n, j_n \in \mathcal{F}_n \quad .$$  \hspace{1cm} (8.23)

Under this requirement, the concatenation of customer-level and conventional load shedding is performed in the following way: the remaining load at the bus due to a certain CL-UFLS ratios $r_{L,n}$ is calculated, which is then subject to a further reduction by the conventional load shedding ratio $r_{F,n}$. The remaining load after the CL-UFLS is described by

$$P_{clL0,rem,n} = r_{T,L,n} \left[1 - r_{clshed,n}\right] P_{L0,n} \quad .$$  \hspace{1cm} (8.24)

The effect of the conventional load shedding is independent of the class-dependent load reduction due to the requirement from (8.23). Thus, it can be assumed to act on the aggregated remaining load after the CL-UFLS:

$$P_{frL0,rem,n} = r_{T,F,n} \left[1 - r_{frshed,n}\right] P_{clL0,rem,n} \quad .$$  \hspace{1cm} (8.25)

Inserting (8.24) into (8.25), the remaining load at bus $n$ is equal to:

$$P_{L0,rem,n} = r_{T,F,n} \left[1 - r_{frshed,n}\right] r_{T,L,n} \left[1 - r_{clshed,n}\right] P_{L0,n} \quad .$$  \hspace{1cm} (8.26)

8.6 Frequency Threshold Computation

In this section, the frequency threshold computation will be presented both for the conventional and the customer level load shedding.

8.6.1 Conventional Load Shedding

In this work, we consider a load shedding stage plan as commonly used in the ENTSO-E Continental Europe (CE) region [175]. It will consist of four load shedding stages defined by the thresholds $f_{frstage1}, \ldots, f_{frstage4}$
8.6. Frequency Threshold Computation

and the detection time delays $\Delta t^{fr}_{\text{detect, stage1}, \ldots, \text{stage4}}$. The tripping time delays $\Delta t^{fr}_{\text{trip}}$ are assumed to be given by the frequency relays in use.

We take the following simple approach for parameterizing the conventional load shedding system: The stages are selected heuristically, roughly based on the recommendations set forth in [175]. The distribution of the frequency thresholds onto all feeders in the system is then performed in the following way:

1. select the current load shedding stage $k \in \{1, \ldots, 4\}$,

2. assign distribution feeders (here by random pre-selection, in reality often according to a non-discriminatory rotation scheme) to the $k^{th}$ load shedding stage (i.e., frequency threshold and detection time) until the shed power at stage $k$ corresponds to the stage plan requirement, and

3. progress to the next load shedding stage until a sufficient amount of feeders have been assigned to each stage.

8.6.2 Customer-Level Load Shedding

Having established the appliance categories for a given system, the associated frequency thresholds must be determined. As stated above, the threshold tuning should take into account the VOLL of load class $i_n$ at bus $n$, denoted by $VOLL_{i_n}$, or by the vector $VOLL_n$. As stated in Section 8.4.1, the load classes $i_n$ are defined identically for all $n$, and thus, $VOLL_n$ is identical for all $n$.

Optimization-based tuning methods can be used in order to find the frequency thresholds $f^{cl,1}_{\text{thr},i_n}$ and $f^{cl,2}_{\text{thr},i_n}$. In order to reduce the complexity of this work, we choose a simple method that considers the VOLL while ensuring a constant relation between shed load and system frequency for the entire frequency span of the CL-UFLS. The following algorithm is used:

1. Determine the frequency span of the CL-UFLS: it should not start above a certain minimum deviation of the frequency from its nominal value, e.g., $f^{\text{cl,1}}_{\text{thr, max},i_n} = 48.8$ Hz, in order to avoid activation in normal operation situations. We define furthermore that the
**CL-UFLS** should end before the first conventional load shedding stage with a certain margin, e.g., \( f_{\text{thr},\text{min},i_n}^{\text{cl,1}} = f_{\text{thr},\text{max},j_n}^{\text{fr}} + 0.2 \text{ Hz} \).

2. Calculate the aggregated shedding slope \( m_{\text{agg}} \) [\( \frac{\text{MW}}{\text{Hz}} \)]:

\[
m_{\text{agg}} = \frac{\Delta P_{L,\text{total}}^{\text{cl}}}{\Delta f_{\text{total}}^{\text{cl}}} \tag{8.27}
\]

with the total power \( \Delta P_{L,\text{total}}^{\text{cl}} = \sum_{n=1}^{N} r_{\text{pen},n}^{T} r_{L,n} P_{L,n} \) and the frequency span \( \Delta f_{\text{total}}^{\text{cl}} = f_{\text{thr},\text{max},i_n}^{\text{cl,1}} - f_{\text{thr},\text{min},i_n}^{\text{cl,2}} \).

3. Sort the appliance categories according to the vector \( \text{VOLL}_n \) in ascending order. This defines the shedding sequence of the individual categories with decaying frequency.

4. Calculate for each load class the frequency span such that the shedding slope of the load class is equal to aggregated shedding slope: \( m_i = m_{\text{agg}} \), where \( m_i \) is calculated for class \( i \) in an analogous way to \( m_{\text{agg}} \) in (8.27).

Numerical results of an exemplary application of this approach are given in Section 8.7.2. It is further determined that the detection time delays are set to identical values for all classes. The selected value depends on the transient behavior of the used power system frequency measurement system.

After having introduced the **CL-UFLS**, it can be seen that the amount of power shed by the conventional load shedding system is different. This is due to the fact that the shed feeders have a lighter loading since part of the load has already been tripped. Depending on the system penetration of the **CL-UFLS**, the conventional load shedding system settings can be adapted. This is illustrated in Section 8.7.2.

### 8.7 Case Study

In the following, we conduct a simulation study on the effectiveness of the **CL-UFLS** system. We will introduce the used power system model and the disturbance scenarios, present the load shedding system parameterization, and finally conduct numerical simulations in order to investigate the interaction between **C-UFLS**, **CL-UFLS**, and **DG** during a disturbance.
8.7. Case Study

8.7.1 Power System Model

As stated above, the performance of the proposed approach is evaluated on the IEEE 118-bus network which is considered complex enough to capture the difficulties that may arise in a real system. The data of the model are retrieved from a snapshot available at [214]. It includes 19 generators, 177 lines, 99 load buses, and 7 transmission level transformers. The system used in the simulations contains slight modifications so as to render the model in a more realistic configuration. Thus, in order to represent the different voltage levels, 19 more transformers are added to connect the MV generator buses (6.9 – 24 kV) with the HV transmission level buses (400 kV). Moreover, since there were no dynamic data available, typical values provided by [215] were used for the simulations. Note that a critical influence of the exact numerical values on the results is not to be expected. Figure 8.5 shows a single-line diagram of the network.

All generators are represented by the classical model and are equipped with primary frequency control and Automatic Voltage Regulator (AVR). Details on the controllers, as well as the dynamic data used in the simulations, can be found in [195]. The setpoint of the scheduled mechanical power \( P_{m,\text{set}} \) of each generator is determined according to the mentioned IEEE 118-bus system power flow case.

Figure 8.5: Topology of the 118-bus system [195]
The load model for the load at bus \( n \), \( P_{L,n} \), implies a voltage dependency of the load, while a frequency dependency is not considered:

\[
P_{L,n} = P_{L0,n}(V_T/V_0)^\alpha,
\]

\[
Q_{L,n} = Q_{L0,n}(V_T/V_0)^\beta.
\]

(8.28)

(8.29)

The exponents \( \alpha \) and \( \beta \) are set both equal to 2 representing constant impedance characteristics.

The model of the network is implemented using NEPLAN [216], a power system dynamic simulator which is suitable for transient stability analysis. In order to interact with the system during run-time using customized algorithms, the NEPLAN model was converted to a dynamic simulation environment for MATLAB [126] described in [217].

### 8.7.2 Load Shedding System Parameterization

Both load shedding systems are parameterized according to the methodologies stated in Section 8.6. Note that the detection time delays are set to a uniform value, which is adapted to the transient behavior of the instantaneous bus frequency measurement realized by a PLL system.

#### Conventional Load Shedding

For a specific load situation, the load at each bus of the system is known. For the purpose of parameterizing the load shedding system, distribution feeders have to be artificially created as they are not part of the available data. As already stated above, the feeders represent only portions of load to be disconnected in one piece since the distribution system topology is not modeled. Therefore, we take a simple approach which creates a reasonably realistic number of feeders and distribution of load:

First, the cumulated active power load in the whole system is compared with the cumulated active power limits of the generators, which yields the current load factor of the entire system. By proportional scaling, the load at each bus is adapted such that the system is at maximum loading. Based on this, the number of feeders at each bus is determined by assuming a maximum feeder capacity. We assume a maximum of 20 MW per feeder, and consequently end up with a varying number of feeders at different buses in the range between 1 and 20.
8.7. Case Study

Table 8.2: Conventional load shedding stage plan

<table>
<thead>
<tr>
<th>$f_{thr}$</th>
<th>$\Delta t_{detect}$</th>
<th>$\Delta t_{trip}$</th>
<th>$P_{F, shed}/P_{F}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>49.0 Hz</td>
<td>0.10 s</td>
<td>0.05 s</td>
<td>15.00%</td>
</tr>
<tr>
<td>48.7 Hz</td>
<td>0.10 s</td>
<td>0.05 s</td>
<td>15.00%</td>
</tr>
<tr>
<td>48.4 Hz</td>
<td>0.10 s</td>
<td>0.05 s</td>
<td>15.00%</td>
</tr>
<tr>
<td>48.1 Hz</td>
<td>0.10 s</td>
<td>0.05 s</td>
<td>15.00%</td>
</tr>
</tbody>
</table>

Figure 8.6: Distribution of load on the distribution feeders

The load at each bus of the unscaled system (original load situation) is distributed with some stochastic variability (scaling factors from a normal distribution with mean 1 and 20% standard deviation) on the feeders. The last feeder at one bus has to take the remainder of the load so that the original bus load is not changed. Figure 8.6 shows the load per feeder, clustered according to buses. Note that the feeder load can be slightly higher than the established maximum of 20 MW due to the stochastic load assignment, which is a case that can also temporarily occur in reality. The parameters for the conventional load shedding are based on the stage plan presented in Table 8.2.

Customer-Level Load Shedding

In order to parameterize the CL-UFLS, the load classes have to be created first. In this work, this is entirely assumption-based, however reasonably realistic. In a first step, the 118-bus system is divided into three regions, which are assumed to have different compositions of customer categories (residential, commercial, industrial). The assumed shares of each customer category in the regions are summarized in Table 8.3.

The customer categories are now subdivided into subcategories, the total set of which constitutes the load classes for the CL-UFLS. In Table 8.4, the 14 load classes created for the present system are summarized. In the column “Rel. Share”, the percentage is stated that the load class accounts for with respect to its customer category. This is actually only valid for one specific load situation and will consequently vary during the course of the day. Based on the calculations performed
Chapter 8. Customer-Level UFLS

Table 8.3: Assumed customer categories in 118-bus system

<table>
<thead>
<tr>
<th>Customer Category</th>
<th>Region 1</th>
<th>Region 2</th>
<th>Region 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residential</td>
<td>70.00%</td>
<td>20.00%</td>
<td>50.00%</td>
</tr>
<tr>
<td>Industrial</td>
<td>20.00%</td>
<td>70.00%</td>
<td>10.00%</td>
</tr>
<tr>
<td>Commercial</td>
<td>10.00%</td>
<td>10.00%</td>
<td>40.00%</td>
</tr>
</tbody>
</table>

Bus numbers 1 – 33; 34 – 74; 75 – 112; belonging to cat. 113 – 115; 116 118

here, a category-wise load profile can easily be used for the creation of time-dependent frequency threshold assignments. For details on the time-varying load shedding potential of residential customers with exemplary calculations for Germany, see [197].

Assumptions for the VOLL per customer category are shown in the column “VOLL(category)”. Note that the VOLL value is expressed in EUR/kWh of energy not supplied – which means that a certain load shedding action has to be combined with an assumed outage duration in order to yield a monetary value. We assume this to be 1 hour in our simulated scenarios. This may be exaggerated in the case of smaller-scale shedding actions. However, in the case of larger disturbances causing under-frequency situations in the entire system, the actual outage duration will probably be longer. For a direct comparability of the results, the VOLL value should be calculated based on estimated outage durations depending on the magnitude of the disturbance. This, however, is beyond the scope of this work.

As further presented in Table 8.4, a VOLL modifier is introduced in order to rank the customer comfort loss brought about by the disconnection of specific appliance types. With this, the modified VOLL(class) is derived. Next to this column, the assumptions about CL-UFLS penetration are summarized, which is a moderate 50% for selected appliance types in the present case. The frequency thresholds $f_{\text{cl,}1}^{\text{thr},i_n}$ and $f_{\text{cl,}2}^{\text{thr},i_n}$ are calculated based on the algorithm presented in Section 8.6. Finally, the load shedding time delays are stated (identical detection time related to measurement transients, shedding-system-related tripping time).

Now the load shedding system performance is analyzed by a static calculation of the remaining load according to system frequency. The main purpose of this calculation is to gain an understanding of the interaction between the customer-level and the conventional load shedding system, which influence each other in spite of the fact that they operate in different frequency ranges. Note that this influence does not occur in operational terms but rather due to the two shedding systems operating on the same loads.
### Table 8.4: Load classes, VOLL assumptions and frequency thresholds

<table>
<thead>
<tr>
<th>Category</th>
<th>Class</th>
<th>Rel. Share</th>
<th>VOLL(category)</th>
<th>Modifier</th>
<th>VOLL(class)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residential</td>
<td>Small Devices</td>
<td>23%</td>
<td>15 EUR/kWh</td>
<td>2</td>
<td>30 EUR/kWh</td>
</tr>
<tr>
<td>Residential</td>
<td>TV-HiFi</td>
<td>7%</td>
<td>15 EUR/kWh</td>
<td>2</td>
<td>30 EUR/kWh</td>
</tr>
<tr>
<td>Residential</td>
<td>Cooking</td>
<td>10%</td>
<td>15 EUR/kWh</td>
<td>1</td>
<td>15 EUR/kWh</td>
</tr>
<tr>
<td>Residential</td>
<td>Lighting</td>
<td>9%</td>
<td>15 EUR/kWh</td>
<td>2</td>
<td>30 EUR/kWh</td>
</tr>
<tr>
<td>Residential</td>
<td>Washing</td>
<td>12%</td>
<td>15 EUR/kWh</td>
<td>1</td>
<td>15 EUR/kWh</td>
</tr>
<tr>
<td>Residential</td>
<td>Heating</td>
<td>3%</td>
<td>15 EUR/kWh</td>
<td>0</td>
<td>0 EUR/kWh</td>
</tr>
<tr>
<td>Residential</td>
<td>Warm water</td>
<td>14%</td>
<td>15 EUR/kWh</td>
<td>0</td>
<td>0 EUR/kWh</td>
</tr>
<tr>
<td>Residential</td>
<td>Cooling</td>
<td>22%</td>
<td>15 EUR/kWh</td>
<td>0</td>
<td>0 EUR/kWh</td>
</tr>
<tr>
<td>Industrial</td>
<td>Interruptible</td>
<td>10%</td>
<td>25 EUR/kWh</td>
<td>0</td>
<td>0 EUR/kWh</td>
</tr>
<tr>
<td>Industrial</td>
<td>Non-interr.</td>
<td>90%</td>
<td>25 EUR/kWh</td>
<td>2</td>
<td>50 EUR/kWh</td>
</tr>
<tr>
<td>Commercial</td>
<td>Heating</td>
<td>20%</td>
<td>30 EUR/kWh</td>
<td>0</td>
<td>0 EUR/kWh</td>
</tr>
<tr>
<td>Commercial</td>
<td>Cooling</td>
<td>15%</td>
<td>30 EUR/kWh</td>
<td>0</td>
<td>0 EUR/kWh</td>
</tr>
<tr>
<td>Commercial</td>
<td>Lighting</td>
<td>15%</td>
<td>30 EUR/kWh</td>
<td>1</td>
<td>30 EUR/kWh</td>
</tr>
<tr>
<td>Commercial</td>
<td>Remaining</td>
<td>50%</td>
<td>30 EUR/kWh</td>
<td>3</td>
<td>90 EUR/kWh</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Category</th>
<th>Class</th>
<th>$r_{\text{pen},i_n}^{\text{cl}}$</th>
<th>$r_{\text{thr},i_n}^{\text{cl,1}}$</th>
<th>$r_{\text{thr},i_n}^{\text{cl,2}}$</th>
<th>$\Delta t_{\text{detect},i_n}^{\text{cl}}$</th>
<th>$\Delta t_{\text{trip},i_n}^{\text{cl}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residential</td>
<td>Small Devices</td>
<td>0%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Residential</td>
<td>TV-HiFi</td>
<td>0%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Residential</td>
<td>Cooking</td>
<td>0%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Residential</td>
<td>Lighting</td>
<td>0%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Residential</td>
<td>Washing</td>
<td>50% 49.1962</td>
<td>49.1000</td>
<td>0.1 s</td>
<td>0.1 s</td>
<td></td>
</tr>
<tr>
<td>Residential</td>
<td>Heating</td>
<td>50% 49.8000</td>
<td>49.7760</td>
<td>0.1 s</td>
<td>0.1 s</td>
<td></td>
</tr>
<tr>
<td>Residential</td>
<td>Warm water</td>
<td>50% 49.7760</td>
<td>49.6637</td>
<td>0.1 s</td>
<td>0.1 s</td>
<td></td>
</tr>
<tr>
<td>Residential</td>
<td>Cooling</td>
<td>50% 49.6637</td>
<td>49.4874</td>
<td>0.1 s</td>
<td>0.1 s</td>
<td></td>
</tr>
<tr>
<td>Industrial</td>
<td>Interruptible</td>
<td>50% 49.4874</td>
<td>49.4488</td>
<td>0.1 s</td>
<td>0.1 s</td>
<td></td>
</tr>
<tr>
<td>Industrial</td>
<td>Non-interr.</td>
<td>0%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Commercial</td>
<td>Heating</td>
<td>50% 49.4488</td>
<td>49.3044</td>
<td>0.1 s</td>
<td>0.1 s</td>
<td></td>
</tr>
<tr>
<td>Commercial</td>
<td>Cooling</td>
<td>50% 49.3044</td>
<td>49.1962</td>
<td>0.1 s</td>
<td>0.1 s</td>
<td></td>
</tr>
<tr>
<td>Commercial</td>
<td>Lighting</td>
<td>0%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Commercial</td>
<td>Remaining</td>
<td>0%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

**Figure 8.7:** Legend of the load classes for load shedding plots

Figures 8.8, 8.9, and 8.10 depict the static analysis of the two interacting load shedding systems over a range of possible system frequencies. In each of the figures, plot a) shows the remaining system load according to bus in MW, plot b) shows the shed load according to load class, and plot c) shows the cumulated VOLL according to class. For the plots referring to load classes, the legend presented in Figure 8.7 holds. It is important to note that in Figure 8.9 the conventional load shedding is impacted in two ways: First, the actual system load is less than what the conventional load shedding was designed for. In fact, the first shedding stage of the stage plan has already been fulfilled before even
reaching it. The second effect is that the shed load per stage is reduced, as part of the load on the shed feeders is already off. These effects can be desirable or not. If a performance of the system with CL-UFLS similar to the conventional system is desired, this can be achieved by adapting the frequency thresholds of the conventional load shedding as shown in Figure 8.10. In this case, the first load shedding stage is completely eliminated, and the others are adapted by some percent in order to meet the thresholds stipulated in the stage plan. The developed load shedding threshold parameter sets will be used in the simulations presented in the following.

8.7.3 Simulation Scenarios

A significant loss of active power is needed in the system in order to cause a frequency decay that triggers the UFLS systems. We choose a simultaneous tripping of the buses at bus 80 and 89. This disturbance affecting three generators at the same time might not be likely to occur in a real power system. However, cascading events including controlled or uncontrolled islanding of parts of the interconnected system can very well cause a lack of active power of this magnitude. We therefore create
Figure 8.9: Static analysis of the load shedding system – CL-UFLS with unchanged C-UFLS

Figure 8.10: Static analysis of the load shedding system – CL-UFLS with adapted C-UFLS
this benchmark situation only for ease of implementation as a proxy for more complex real disturbances involving a large lack of active power.

Furthermore, a 50% DG power penetration is considered, modeled as a negative load, Poisson-distributed in 0.2 MW portions on the system feeders. The power generation of the central generators is reduced proportionally. This can be justified if the considered DG is only temporarily available generation such as PV generation around noon, which is not reliable enough to make a substantial impact on unit commitment.

Three scenarios are simulated: 1) no DG penetration (and, consequently, no DG loss) and conventional load shedding, 2) 50% DG and conventional load shedding, and 3) 50% DG and 50% penetration of customer level load shedding with adaptation of the conventional load shedding thresholds as described in Section 8.7.2. We omit the simulation of CL-UFLS without any DG since it is identical to scenario 3). This is due to the ability of the CL-UFLS to completely avoid the activation of C-UFLS in the simulated scenario.

8.7.4 Numerical Results

Now we present numerical simulation results using the disturbance scenario defined above. We use a constant time step size of 10 ms in the simulations. Event-handling as described in [217] is performed for important system events outside of the normal step size. The generator outage is triggered 0.5 s into the simulation and the total simulation time is 10 s.

In all result plots (Figures 8.11, 8.12, and 8.13) the structure is as follows: in part a), the frequencies of the 17 remaining generators are depicted, in part b) the measured bus frequencies (by PLL measurement), in part c) the remaining load at the buses in the system, and in part d) the totally incurred VOLL. The latter is based on the (quite arbitrary) assumption that the outage duration is equal to one hour.

In the following, we explain the individual simulation results in detail. After that, we directly compare the results with each other and discuss the differences.
8.7. Case Study

Scenario 1) C-UFLS, no DG

Figure 8.11 shows the result of C-UFLS without any influence of DG. The following effects can be observed in part a) – d) of the figure:

a) The frequencies of the remaining synchronous generators decay together in an oscillatory manner. The minimum frequency, about 48.8 Hz, is attained approximately 2 seconds after the disturbance. After that, the frequency decay is halted by the load shedding action of the first stage, and the temporary excess of generation takes the frequency back to about 50 Hz.

b) The measured frequency follows a path that is similar to the actual generator frequencies. The main difference consists of the strong spikes right after the disturbance which is caused by the transient response of the PLL frequency measurement mechanism.

c) Most of the load shedding occurs around 1.7 – 2.0 seconds after the disturbance. This is caused by the measured frequencies hitting the first conventional load shedding stage. A first-stage load shedding already takes place at some buses right after the disturbance, which is an undesirable effect of the PLL artifact shown in b) and can be avoided through more sophisticated filtering.

d) The total VOLL amounts to slightly more than EUR 24 million caused by the activation of the first load shedding stage.

Scenario 2) C-UFLS, 50% DG

Figure 8.12 shows the result of C-UFLS with DG share of 50%.

a) The initial frequency trajectory follows the same path as in scenario 1). When the first load shedding stage kicks in at 49 Hz, the frequency decay can be stopped, but the frequency is not taken back to its original value. The reason for this is the lost DG in the parts of the grid that were affected by load shedding. In the present scenario, the effect is not terminal since the frequency decay could be stopped nonetheless. For a larger disturbance, however, the loss of DG can imply the secondary triggering of the second load shedding stage.

b) Again, the behavior of the measured bus frequencies except for the transient spikes after the disturbance.

c) The load shedding behavior is the same as for scenario 1).

d) The VOLL is also equal to the value attained in scenario 1).
Scenario 3) C-UFLS and CL-UFLS, 50% DG

The last simulation depicted in Figure 8.13 consists of both C-UFLS and CL-UFLS with a DG penetration of 50%. The results show:

a) Here the effect of the CL-UFLS can be observed, which arrests the frequency decay already above 49 Hz. Only one generator briefly attains values of below 49 Hz, but no conventional load shedding is triggered. The frequency is quickly stabilized, attaining a value of about 50.2 Hz at the end of the simulation.

b) The measured frequencies behave in a similar manner as before.

c) The first involuntary load shedding is avoided here due to the adaptation of the C-UFLS (lower activation frequency). Most of the load shedding takes place about 0.7 – 1.3 seconds after the disturbance.

d) The incurred VOLL is marginal with a value of EUR 0.1 million.
Figure 8.11: Simulation of scenario 1): C-UFLS, no DG
Chapter 8. Customer-Level UFLS

Figure 8.12: Simulation of scenario 2): C-UFLS, 50% DG
Figure 8.13: Simulation of scenario 3): C-UFLS + CL-UFLS, 50% DG
In order to sum up the previous results, we directly compare the aggregate values for frequency response, load shedding, DG shedding, and incurred VOLL. Figure 8.14, part a) – d), show the following quantities:

a) Here we show a comparison of the individual generator frequencies (thin lines) including the Center of Inertia (COI) frequency [111] for the three simulated scenarios. The direct comparison between the COI frequencies shows that the frequency decay is arrested in approximately the same point for scenarios 1) and 2), and considerably higher for scenario 3). This is due to the fact that CL-UFLS is activated at a higher frequency than C-UFLS. The differences in stabilization behavior are also clearly visible. While for scenario 1), the temporary excess of generation brings the frequency back to its nominal value without the need for primary control activation, this is not the case for scenario 2) due to the shed DG. For scenario 3) a slight overshoot above 50 Hz is created since more load is shed earlier than for the previous two scenarios. Note that this behavior is dependent on the magnitude of the disturbance.

b) This plot depicts the overall load shedding for the three scenarios. It can be seen that scenarios 1) and 2) have almost the same load shedding characteristics. Scenario 3), however, is characterized by more and earlier load shedding (since CL-UFLS starts to be active at a higher frequency).

c) This plot demonstrates how the DG shedding takes place in scenario 2). The curve shape is similar to the load shedding, which is plausible since the DG shedding is caused by tripping load shedding relays.

d) Finally, we compare the incurred VOLL for the three scenarios. It can be seen that scenarios 1) and 2) have highly similar progressions, while scenario 3) is not even visible on this scale. The fact that the VOLL of scenario 3) is negligible, while the total load shedding is actually higher than in the other two scenarios, is explained by the predominant shedding of low-priority load that can be disconnected at zero cost.

Table 8.5 shows a numerical comparison of the final simulation outcomes.
### Table 8.5: Numerical comparison of the simulation results

<table>
<thead>
<tr>
<th>Simulation results</th>
<th>Scenario 1)</th>
<th>Scenario 2)</th>
<th>Scenario 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>COI frequency nadir</td>
<td>48.88 Hz</td>
<td>48.86 Hz</td>
<td>49.18 Hz</td>
</tr>
<tr>
<td>COI frequency at $t=10$ s</td>
<td>49.96 Hz</td>
<td>49.16 Hz</td>
<td>50.22 Hz</td>
</tr>
<tr>
<td>Total load shed</td>
<td>558.92 MW</td>
<td>558.92 MW</td>
<td>641.09 MW</td>
</tr>
<tr>
<td>Total DG in-feed lost</td>
<td>0.00 MW</td>
<td>212.95 MW</td>
<td>0.00 MW</td>
</tr>
<tr>
<td>Total VOLL incurred</td>
<td>24.150 M EUR</td>
<td>24.150 M EUR</td>
<td>0.099 M EUR</td>
</tr>
</tbody>
</table>

![Individual generator and Center of Inertia frequencies](image)

**Figure 8.14: Comparison of the simulation results**

- **a)** Individual generator and Center of Inertia frequencies
- **b)** Overall shed load
- **c)** Overall shed DG power
- **d)** Value of Lost Load
8.8 Concluding Remarks

In this chapter, we presented a methodology for designing UFLS mechanisms on the customer level. The presented mathematical formulation of the shed and remaining load share at a certain distribution feeder leads enables a structured assessment of the interaction between C-UFLS and CL-UFLS systems. A methodology for assigning frequency thresholds to individual appliance classes has been proposed and evaluated by a static analysis of the shed load depending on the system frequency. Dynamic time simulations have shown the applicability of the approach to a 118-bus benchmark system. Further work may include the study of concrete system design for real transmission systems. Another point of interest is the assessment of distribution-level impacts of the sudden disconnections of load on a feeder that remains connected to the rest of the power system.

Opportunities for further work are, e.g., the development of an organizational concept for coordinating a multitude of distribution systems with different CL-UFLS penetrations, a more detailed investigation of the effects of DG on the load shedding schemes, and the study of daily load profiles and corresponding time-varying frequency threshold assignment for the appliance classes.
Chapter 9

Conclusions and Outlook

In this last chapter, we try to summarize and consolidate the work that was presented in this thesis. The four main topics of this work (coordination algorithms for Thermostatically Controlled Loads (TCLs), system-level dispatch strategies using the Power Nodes Modeling Framework (PNMF), the profitability of ancillary service provision by flexible unit portfolios, and Customer-Level Under-Frequency Load Shedding (CL-UFLS)) are discussed in separate sections. We will focus on the conclusions and experiences gained during the development process and also give suggestions for future research.

9.1 Appliance-Level Coordination

The first part of this thesis, Chapters 2, 3, and 4, was focused on appliance-level modeling and coordination strategies. In Chapter 2, a generic model for Thermostatically Controlled Loads (TCLs) was developed. We used the analysis of the dynamic behavior of large numbers of these models to develop a control strategy that imposes a setpoint on the TCL population. The implications of the coordination algorithm on the appliance level during normal operation of the power system were also discussed. In Chapter 3, this generic model was extended to a multi-state temperature layer model of a water heater tank. The main purpose of this modeling effort was to evaluate the effect of water draws on the setpoint tracking and how the control may affect user comfort. In Chapter 4, we used a Markov-chain-like modeling technique to describe the TCL population as a whole by a linear state-space model. The dynamic behavior and eigenvalues were analyzed and a control strategy based on Model Predictive Control (MPC) was developed.
9.1.1 Lessons Learned

We learned from our simulations in Chapter 2 that rule-based coordination algorithms are promising to enable a population of TCLs to track a time-varying power setpoint in the presence of fast and reliable communication equipment. Since this setpoint can be changed on short notice (i.e., just one time step ahead), the approach is particularly suitable for ancillary service provision. For practical implementation, one of the main issues appears to be the duty cycle estimation on the appliance level, which can be omitted if the coordination algorithm is simplified to a mere sorting of TCLs according to their position inside the temperature dead-band. Another issue is the need for synchronized transmission of the control commands and the communication of state information back to the coordination entity. If future commercially available TCLs will be equipped with a sophisticated communication interface by default, an approach like this will be implementable with relative ease. However, as outlined in [218], a market penetration of these modified appliances may take up to several decades. For a near-term implementation of sophisticated Demand Response (DR) methods, it seems advisable to pursue a reduction of the communication effort, which was a key motivation for the developments presented in Chapter 4.

As to the modeling of electric water heaters presented in Chapter 3, the stratification can be modeled with straight-forward energy and mass balances of over a finite number of layers. In this modeling, the trade-off between physical accuracy and the practical applicability to utility-scale systems of several tens of thousands of units has to be taken into account. We use a mixture of physical effects and heuristic terms in the model to derive a system which yields physically plausible simulation results. We find that the problem of bi-linearity between input and state can be approached by means of a Linear Parameter-Varying (LPV) system formulation. By simulating the system without any external control, one can observe the dominant influence of water draws on the electricity consumption behavior of the water heaters. If the population is controlled by a blocking strategy, one must consider that the upward control potential (load increase) can be limited if the population of water heaters is already too hot.

The probabilistic modeling presented in Chapter 4 appears to be a suitable way to reduce the communication effort contained in coordinating a TCL population. We find through open-loop simulations of the uncontrolled TCL population that the Markov-chain modeling technique in-
9.1. Appliance-Level Coordination

Introduces artifacts into the dynamics of the aggregate population model. These consist of the different oscillatory behavior of the population of individual TCLs and the aggregate population model, and of the dependency of the oscillation decay on the number of modeled state bins. However, simulations show that the accuracy of the aggregate population model dynamics is sufficient for calculating short-term control actions with a predictive model-based controller.

9.1.2 Suggestions for Future Research

For the rule-based coordination approach, we propose the following aspects to be investigated in future work: First, the application of the coordination to other specific TCL models such as detailed, multi-state models of heat pumps, can be explored. The approach can also be extended to the aggregation of more complex building energy systems. For practical application, it is also necessary to model the specific requirements of the TCL type, such as minimum ON/OFF times, and to consider these in the coordination algorithm. As to the control accuracy, further investigations on the effects of communication latency should be conducted.

The representation of electric water heaters can be extended by exploring alternatives to the presented modeling technique, such as a more sophisticated representation of the thermodynamics within the tank. Since all thermodynamic modeling approaches have to utilize a number of simplifying assumptions, there are certain degrees of freedom in the modeling decisions. For example, a more accurate representation of the convective propagation of hot water during the heating phase can be attempted, while it is clear that the model must stay simple enough to be suitable for utility-scale system simulations. Other control algorithms for setpoint tracking can also be explored, e.g., algorithms that allow a reduction of the communication effort.

The probabilistic Markov-chain modeling technique can be extended by introducing appliance representations that are more realistic than first-order dynamic temperature models. The utilization of state estimation and identification techniques was already explored in [125]. This can be extended further, e.g., by investigating the effect of communication latency on control performance.

For all of the presented approaches, a further piece of work may consist of the integrated simulation of a TCL population with a detailed re-
presentation of one or several distribution systems. By this means, local voltage variations and the handling of distribution-level grid constraints can be explored.

9.2 System-Level Dispatch Strategies

In Chapter 5, we shifted the focus from the appliance level up to the system level. By introducing the Power Nodes Modeling Framework (PNMF), we facilitate the modeling of individual and aggregated units in the power system for capturing relevant properties while abstracting from the physical details. Based on this notation, we formulated models for common power system units, such as controllable and uncontrollable load, intermittent and dispatchable generation, and energy storage devices.

Based on the PNMF, we developed a number of optimization-based dispatch strategies in Chapter 6 which can be applied to a portfolio of units in the power system represented by the PNMF. We presented a comprehensive framework for delivering control services to the power system by using predictive optimization techniques.

9.2.1 Lessons Learned

We conclude from our modeling efforts that indeed a consistent framework can be formulated that comprises unit properties such as storage capacity, constraints such as minimum/maximum power and ramp-rate, dispatchability properties such as total and partly existing freedom to control the energy output of a unit, and (both financial and figurative) cost terms associated with the decision variables. The formulated balance terms enable an easy performance evaluation of a time simulation result with respect to economic and environmental criteria.

As shown in Chapter 6, the formulation of control applications can be done in an equally consistent way, facilitated by the introduction of two auxiliary power nodes, the Control Power Node and the Slack Power Node. Different control goals in the power system can be pursued by suitable cost function and constraint design.

9.2.2 Suggestions for Future Research

Potentials for future research exist for both the modeling approach and the dispatch strategies. First, the power node models can be extended
by nonlinear components, e.g., non-constant efficiency terms. This will make the solution of optimization problems more difficult but it will increase the degree of realism of the mathematical representation of power system units. Furthermore, the interconnection of different power nodes outside of the power system domain (i.e., by introducing coupling terms between different power nodes) enables the modeling of multi-state energy storage systems, such as hydro power plant cascades.

As to the dispatch strategies, we only considered perfect forecasting for the time series prediction up to the prediction horizon. This is an idealistic assumption which serves to establish a performance benchmark for the presented methodologies. Since we focus on the conceptual clarity and consistence of the presented optimization approaches, we deem perfect predictions sufficient for the purposes of this thesis. For the goal of obtaining realistic quantitative results, the inclusion of prediction errors will be necessary.

9.3 Profitability of Flexible Unit Portfolios

In Chapter 7, we analyzed the revenue potentials of ancillary service provision by flexible unit portfolios. The methodology was based on the PNMF and consisted of a set of one-month time simulations with 10-second sampling and variations of the underlying parameters. These simulation results were then used as an input for a profit sharing methodology.

9.3.1 Lessons Learned

Based on the time simulation, we were able to develop an understanding of the share of the frequency control band that controllable loads and storage devices are able to cover. Due to the avoided opportunity cost which would normally arise from the larger generator control band, we found substantial revenue increases for the provision of ancillary services by combined portfolios.

A statistical assessment of deviations from the control signal can be made for longer simulation runs, which enables a sizing of the energy-storing unit with respect to the control capacity that it has to provide.
9.3.2 Suggestions for Future Research

Further research may be directed towards a more detailed modeling of power system units, especially with respect to their cost of operation in various operating conditions. Apart from that, the consideration of prediction errors in the control signal can further increase the realism.

We considered in this thesis the direct coupling of an energy-storing unit to a generator. In practice, short-term schedule changes or tertiary control reserves may be used to refill the storage in order to avoid the need for a fixed generator control band. This may affect the profitability of the reserve provision in a positive way since the opportunity cost may be drastically reduced. The investigation of this approach requires a more detailed consideration of refill lead times and rules for when and how much to refill the storage, which may be subject to future work.

9.4 Emergency Control by CL-UFLS

Chapter 8 was focused on the investigation of Customer-Level Under-Frequency Load Shedding (CL-UFLS) techniques which serve to stabilize a decaying grid frequency. We started out with a mathematical representation of load groups, clustered according to distribution feeder and load class, in order to model the selective disconnection of loads.

For the load classes, we used a stochastic assignment of frequency thresholds drawn from a normal distribution in order to enable a smooth, ramp-like CL-UFLS. Special attention was paid to the interaction of the CL-UFLS with the Conventional Under-Frequency Load Shedding (C-UFLS) system. Dynamic time-domain simulations of the IEEE 118-bus system [214] were conducted that demonstrate the effectiveness of the CL-UFLS system.

9.4.1 Lessons Learned

Preliminary investigations of the CL-UFLS system were undertaken in [197]. This work served as a basis for the modeling and simulation of Under-Frequency Load Shedding (UFLS) methods in this thesis. Note that the aggregate simulation of the power system frequency dynamics by a single swing equation provides a simplified representation of the load shedding since it does not include any geographical distribution of the system frequency and the load disconnection. For a more accurate
9.4. Emergency Control by CL-UFLS

representation of the dynamic behavior of a power system in an under-frequency situation, a dynamic power system model is needed. We used the IEEE 118-bus system since it is sufficiently large to be reasonably realistic and sufficiently small to be comprehensible and computationally tractable.

We find that the CL-UFLS system, if fast enough for a timely reaction to frequency disturbances, can make a substantial contribution to system security. It is certain to interact with any C-UFLS system, so care has to be taken in the design and coordination of frequency threshold assignments. Distributed Generation (DG) has an adverse impact on the effectiveness of conventional load shedding. This can be alleviated by a step-by-step replacement of the C-UFLS system by a CL-UFLS strategy.

9.4.2 Suggestions for Future Research

The results presented in this thesis may serve as a basis for a concrete investigation of introducing a CL-UFLS system in a real power system. This can be done by equipping realistic power system models, such as detailed transmission system models of entire countries, with the CL-UFLS models and observing the results in direct comparison with the established C-UFLS methodologies.

From a methodological standpoint, some improvements can be made to what we presented in this thesis. We made the simplifying assumption that the DG acts as a negative load in the power system. Thus, we used the same load model for the DG as for the conventional load. In order to increase the degree of realism of the simulation, a detailed modeling of the dynamic DG control systems (especially the response to frequency and voltage variations) should be conducted since this can have an impact on the dynamic behavior of the entire system.

Another aspect of the CL-UFLS is the influence that it has on the distribution grid. A sudden disconnection of large amounts of distribution load may lead to a voltage rise in the affected distribution feeders. This problem is a non-issue in C-UFLS since the shedding of the entire feeder will completely de-energize the affected part of the distribution grid. In order to avoid equipment failures due to these induced voltage variations (and consequently, secondary power outages), a detailed study of this effect is imperative before practical implementation of such a system.
Bibliography


[58] European Technology Platform on Smart Systems Integration. (2009, October) European Roadmap Electrification of Road


[96] N. Saker, M. Petit, and J. Coullon, “Demand side management of electrical water heaters and evaluation of the cold load pick-up


Decision and Information Sciences Division, Argonne National Laboratory, USA, Tech. Rep. ANL/DIS-10-1, November 2009.


