DISS. ETH NO. 23554

Control of Residential and Commercial Loads for Power System Ancillary Services

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH

(Dr. sc. ETH Zurich)

presented by

EVANGELOS VRETTOS

Dipl.-Ing. National Technical University of Athens

born on 29 March 1987

citizen of

Greece

accepted on the recommendation of

Prof. Dr. Göran Andersson, examiner

Prof. Dr. John Lygeros, co-examiner

Prof. Dr. Steven H. Low, co-examiner

2016

To my family, especially to my grandfather.

Abstract

Due to environmental concerns related to burning fossil fuels and nuclear waste disposal, the energy mix has been changing during the last decades with the integration of Renewable Energy Sources (RES). However, RES inject fluctuating electric power into the grid because they depend on the availability of primary natural resources such as wind and solar power. These fluctuating power injections create frequency and voltage deviations from their nominal values, which have to be mitigated to maintain a secure and reliable power supply. Traditionally, the power system operators control the frequency and voltage using active and reactive power reserves offered by conventional generators. Therefore, a large integration of RES will challenge the traditional power system operation because the need for reserves increases and at the same time the portion of controllable generation resources decreases.

Although the traditional operation paradigm in power systems is to dispatch controllable generators to follow a variable electricity demand, it is also conceptually possible to control some portion of the demand to follow a variable generation power infeed. This idea was already proposed in the 1980's, and since then industrial loads participate in programs to support power system operation in many countries. However, there is a lot of potential for reserve provision from residential and commercial loads that remains widely untapped due to challenges related to control complexity, implementation costs, and regulatory aspects.

The main goal of this thesis is to develop methods to enable provision of power system reserves from residential and commercial loads, and verify their suitability for practical implementation in simulation and experimental studies. In addition, this thesis investigates using load control to reduce the electricity cost of individual customers in a way that is

beneficial for the power system. We consider thermal loads, such as refrigerators and Electric Water Heaters (EWHs) in residential buildings, as well as the Heating, Ventilation and Air-Conditioning (HVAC) systems of commercial buildings, because temporary interruptions in their operation are not noticed by the users due to thermal inertia.

The first part of this thesis presents predictive and rule-based controllers to minimize the electricity cost of buildings by shifting the energy consumption to the low-price intervals or consuming the energy produced by rooftop Photovoltaics (PV) within the building premises. The second part of this thesis develops several methods to provide power system Ancillary Services (AS), namely Primary Frequency Control (PFC), Secondary Frequency Control (SFC) and voltage regulation.

More specifically, we propose a decentralized stochastic control method to allow a large aggregation of refrigerators to provide PFC without real-time communication. In addition, we present centralized control algorithms to allow an aggregation of EWHs to track a SFC signal with different levels of information feedback. These algorithms are then extended to account for the constraints of Distribution Networks (DNs) and provide frequency reserves while regulating the DN voltage at the same time. Furthermore, we develop a state estimation method to enable SFC without the need for real-time communication between the central controller and the loads.

A main outcome of this thesis is the development of a hierarchical controller to allow aggregations of commercial buildings to provide SFC reserves reliably and accurately, while trying to maximize energy efficiency. The controller's performance is demonstrated in simulations with models of Swiss office buildings, and its technical feasibility is verified in frequency regulation experiments at a commercial building test facility at the Lawrence Berkeley National Laboratory (LBNL).

The results of this thesis show that with proper control design the flexibility of residential and commercial thermal loads can be used to provide reserves to the power system. Moreover, dynamic frequency studies with a two-area power system model show that a large integration of heterogeneous thermal loads in frequency control will help to reduce frequency deviations and enhance frequency stability.

Kurzfassung

Aufgrund von zunehmenden ökologischen Bedenken bezüglich der Nutzung von fossilen Brennstoffen und der Entsorgung von nuklearen Abfällen wird elektrische Energie in den letzten Jahrzenten vermehrt aus erneuerbaren Energiequellen (Renewable Energy Sources, RES) gewonnen. Da RES von natürlichen Ressourcen wie Wind und Sonneneinstrahlung abhängen, fluktuiert die elektrische Einspeisung ins Netz. Diese Fluktuationen verursachen wiederum Spannungs- und Frequenzabweichungen, welche entsprechend kontrolliert werden müssen um eine sichere und zuverlässige Energieversorgung zu garantieren. Frequenz und Spannung werden traditionellerweise durch Wirk- und Blindleistungsreserven, bereitgestellt durch konventionelle Kraftwerke, kontrolliert. Eine grossflächige Integration von RES ist deshalb eine Herausforderung für die herkömmliche Betriebsweise, da zum einen die Reserveanforderungen steigen und andererseits der Anteil an herkömmlichen regelbaren Generatoren sinkt.

Obwohl in diesem traditionellen Betriebsparadigma von Energiesystemen die Generatoren dem variablen Bedarf der Lastseite folgen, ist es konzeptionell auch möglich einen bestimmten Teil der Last so zu kontrollieren, dass sie einer variablen Energieeinspeisung folgt. Diese Idee wurde in den 1980er Jahren bereits vorgeschlagen, und seitdem unterstützen in vielen Ländern industrielle Lasten den Betrieb des Energiesystems. Trotzdem gibt es Potenzial für die Bereitstellung von Reserven aus häuslichen und kommerziellen Lasten, welche diesbezüglich bis heute weitgehend unerschlossen bleiben. Die Hauptgründe sind die Komplexität der Regelung eines solchen Systems, die Kosten einer Implementation und regulatorische Aspekte.

Das Hauptziel dieser Doktorarbeit ist die Entwicklung von Methoden,

welche die Kleinverbraucher und kommerzielle Lasten für Regelreserven nutzbar machen, sowie die Beurteilung der Umsetzbarkeit in Simulationen und Experimenten. Zudem untersucht diese Arbeit, wie mittels Laststeuerung die Elektrizitätskosten von einzelnen Kunden reduziert werden können und gleichzeitig Vorteile für den Systembetrieb bringen. Es werden thermische Lasten, wie Kühlschränke und elektrische Boiler (Electric Water Heaters, EWHs) in Wohnhäusern, als auch die Heiz-, Lüftungs- und Klimaanlagen in kommerziellen Gebäuden berücksichtigt, da kurzzeitige Betriebsunterbrüche aufgrund der thermischen Trägheit von Benutzern kaum bemerkt werden.

Im ersten Teil werden prädiktive und regelbasierte Regelungen zur Minimierung der Elektrizitätskosten von Gebäuden präsentiert. Diese verschieben den Energieverbrauch in Zeiträume mit niedrigeren Preisen oder konsumieren die Energie von Photovoltaik-Anlagen, welche in der Nähe der Gebäude vorhanden sind. Im zweiten Teil werden verschiedene Methoden entwickelt um Systemdienstleistungen anzubieten, vor allem Primärregelreserven (Primary Frequency Control, PFC), Sekundärregelreserven (Secondary Frequency Control, SFC) und Spannungshaltung.

Insbesondere wird eine dezentralisierte, stochastische Regelmethode vorgeschlagen. Diese ermöglicht die Aggregation von vielen Kühlschränken um PFC zu erbringen, ohne dass dabei Kommunikation in Echtzeit nötig ist. Zusätzlich werden zentralisierte Regelalgorithmen präsentiert, welche es erlauben einem SFC Signal mit einer Aggregation von EWH zu folgen. Dabei werden verschiedene Signale benutzt. Die Algorithmen werden anschliessend so erweitert, dass sie in Verteilnetzen eingesetzt werden können und neben der Frequenz auch die Spannung regulieren können. Des Weiteren entwickeln wir eine Methode zur Zustandsschätzung, welche es erlaubt SFC anzubieten, ohne dass eine Echtzeitkommunikation zwischen einem zentralen Regler und den Lasten besteht.

Ein Hauptresultat der Dissertation ist der entwickelte hierarchische Regler, der eine SFC-Bereitstellung aus einer Aggregation von kommerziellen Gebäuden ermöglicht. Dieser Regler garantiert eine zuverlässige und genaue Bereitstellung und maximiert gleichzeitig die Energieeffizienz. Die Leistungsfähigkeit wird mittels Simulationen von Schweizer Bürogebäuden durchgeführt und die technische Machbarkeit wurde in Experimenten zur Frequenzregelung in einer Testumgebung für kommerzielle Gebäude des Lawrence Berkeley National Laboratory (LBNL)

untersucht.

Die Resultate dieser Dissertation zeigen auf, dass Wohnhäuser als auch kommerzielle Gebäude Reserven mit einem geeigneten Reglerdesign bereitstellen können. Zudem zeigen dynamische Frequenzstudien in einem Energiesystem mit zwei Regionen, dass eine grossflächige Integration von heterogenen thermischen Lasten im Prozess der Frequenzregelung zu einer Reduktion von Frequenzabweichungen führt und die Frequenzstabilität erhöht.

Preface

This thesis summarizes most of my research activities during the last 4.5 years as a PhD student at the Power Systems Laboratory (PSL) of ETH Zurich. There are many people to thank for their help and support during this time.

First and foremost, I would like to express my deepest gratitude to my advisor Prof. Göran Andersson for giving me the opportunity to pursue my PhD at PSL. He has been an exemplary advisor who allowed me a lot of freedom in my research, and was there to support and advise me in the hard moments.

Special thanks go to Prof. John Lygeros from the Automatic Control Laboratory of ETH Zurich for being effectively my second advisor for a significant part of the PhD studies, and for accepting to be a co-examiner of this thesis. I benefited a lot from the inspiring discussions we had during the meetings of the HeatReserves project.

Another special thanks goes to Prof. Steven Low from the California Institute of Technology (Caltech) for being a co-examiner of this thesis. I appreciate the time he spent to read the thesis and his valuable comments that helped me improve the quality of the manuscript.

Many of the results of this thesis are the outcome of the two research projects I was involved in: the "SmartGrid-Polysun: Design Tool for Local Load Management" project and the HeatReserves project. I acknowledge the financial support from the Swiss Federal Office of Energy (BfE), swisselectric research, and Nano-Tera.ch for these two projects. I appreciate the stimulating discussions I had during the project meetings with Dr. Michael Moser, Dr. Martin Kauert, Dr. Andreas Witzig, Roland Kurmann, Dr. Stephan Koch, Dr. Theodor Borsche, Xiaojing

versity of California, Berkeley, for approving the experiments and for his unlimited support and positive attitude.

The students I supervised during their master theses or semester projects helped me investigate different ideas and contributed to some of the results included in this thesis. I would like to thank KuanLin Lai, Stavros Karagiannopoulos, Hanmin Cai, Vasileios Poulios, Adrian Martinez Gomez, Fengtian Zhu, Gregory Ledva, Charalampos Ziras, Jonas Lehr, Thierry Zufferey, Christina Tzanetopoulou, Philipp Lütolf, Antoine Gaillard, Felix Moritz, and Carsten Heinrich.

The offices G23 and G29 at PSL were exciting places to work due to the company of my officemates Dr. Maria Vrakopoulou, Dr. Stephan Koch, Dr. Theodor Borsche, Dr. Spyros Chatzivasileiadis, Dr. Tobias Haring, Dr. Petros Aristidou, Dr. Osvaldo Rodríguez Villalón, and Dr. Marc Scherer. I would like to thank you all for the interesting discussions we had and the relaxing atmosphere we maintained. Of course, I would like to thank all members of the group for the amazing time we spent together not only in PSL, but also outside of it during our excursions and other activities.

Last but not least, I owe a big thanks to my family for their guidance and love, for their support whenever I was grasping new opportunities, as well as for bearing the fact that I am away from home for the last 5 years. My friends in Switzerland and in Greece were an endless source of energy and happiness, and I would like to thank them for being companions in my life. Last but not least, I am deeply grateful for the love, support, and patience of Alina during my PhD studies.

<div align="right">

Evangelos Vrettos

November 2016

</div>

Contents

3   PV Self-Consumption Maximization in Buildings   **49**

   3.1   Introduction   49

   3.2   Building Modeling in Polysun   52

   3.3   Rule-Based Control Design   54

   3.4   Potential for PV Self-Consumption   60

   3.5   Conclusion   73

II   Demand Response for Ancillary Services   **75**

4   Demand-Side Ancillary Services in Power Systems   **77**

   4.1   Introduction   77

   4.2   Frequency Control in Power Systems   78

   4.3   Demand-Side Primary Frequency Control   83

   4.4   Demand-Side Secondary Frequency
Control   85

5   Primary Frequency Control with Refrigerators   **91**

   5.1   Introduction   91

   5.2   Modeling and Basic Control Design   95

   5.3   Refrigerator Startup Dynamics   98

   5.4   Compressor Lockout Constraints   101

   5.5   Improved Probabilistic Switching   103

   5.6   Robustness to Excessive Locking   104

   5.7   Modification of Thermostat Limits   107

   5.8   Robustness to Biased Frequency Deviations   110

   5.9   Considering Thermostat Resolution   115

   5.10 Accounting for Door Openings   118

   5.11 Parameters and Performance Metrics   120

   5.12 Benchmarking and Aggregation Size   122

   5.13 Sensitivity Analysis   125

   5.14 Performance with Biased Frequency Deviations   129

List of Acronyms

ACE Area Control Error

AGC Automatic Generation Control

AHU Air Handling Unit

AS Ancillary Service

BAS Building Automation System

CDF Cumulative Distribution Function

COP Coefficient of Performance

CWS Central Working Station

DA Daily-Asymmetric Reserves at each Building

DN Distribution Network

DR Demand Response

DSA Daily-Symmetric Reserves in Aggregate

DSB Daily-Symmetric Reserves at each Building

ENTSO-E European Network of Transmission System Operators for Electricity

EWH Electric Water Heater

FIT Feed-In Tariff

FLEXLAB Facility for Low Energy eXperiments

HA	Hourly-Asymmetric Reserves at each Building
HP	Heat Pump
HSA	Hourly-Symmetric Reserves in Aggregate
HSB	Hourly-Symmetric Reserves at each Building
HVAC	Heating, Ventilation and Air-Conditioning
IRA	Integrated Room Automation
KiBaM	Kinetic Battery Model
LBNL	Lawrence Berkeley National Laboratory
LP	Linear Program
LV	Low-Voltage
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MHSE	Moving Horizon State Estimation
MILP	Mixed-Integer Linear Program
MLD	Mixed Logical Dynamical
MPC	Model Predictive Control
MPPT	Maximum Power Point Tracker
MV	Medium-Voltage
NPV	Net Present Value
ODE	Ordinary Differential Equation
OPF	Optimal Power Flow
PC	Power Constraints
PCC	Point of Common Coupling
PDE	Partial Differential Equation
PDF	Probability Distribution Function

PEC	Power and Energy Constraints
PFC	Primary Frequency Control
PI	Proportional-Integral
PID	Proportional-Integral-Derivative
PJM	Pennsylvania, Jersey, and Maryland Power Market
PSL	Power Systems Laboratory
PV	Photovoltaics
RBC	Rule-Based Control
RC	Resistance-Capacitance
RES	Renewable Energy Source
RMSE	Root Mean Squared Error
RPC	Robust Problem with Power Constraints
RPEC	Robust Problem with Power and Energy Constraints
SAT	Supply Air Temperature
SC	Slab Cooling
SFC	Secondary Frequency Control
SLP	Sequential Linear Programming
SoC	State of Charge
SPC	Stochastic Problem with Power Constraints
SPEC	Stochastic Problem with Power and Energy Constraints
TABS	Thermally Activated Building Systems
TCL	Thermostatically Controlled Load
TSO	Transmission System Operator
UL	Uncontrollable Load
VAV	Variable Air Volume
VFD	Variable Frequency Drive

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background and Motivation

The power grid operation and planning is currently undergoing large changes due to increasing shares of Renewable Energy Sources (RES) such as wind and solar power. Although they are clean sources of energy, the power output of RES is typically variable and not fully controllable. The RES variability brings new challenges to the traditional transmission system operation because the generation and demand of electric power must be balanced at all times.

Power system frequency reflects the instantaneous balance between generation and demand of electric power. As long as the generation exactly meets the demand, the frequency is at its nominal value, e.g., 50 Hz in Europe and 60 Hz in North America. On the other hand, if the generation becomes lower than the demand, the frequency drops and vice versa. However, the grid frequency must remain close to its nominal value for reliability reasons. For this purpose, the Transmission System Operators (TSOs) rely on primary, secondary and tertiary frequency control reserves in the form of Ancillary Services (AS) to stabilize frequency after a disturbance and restore it to its nominal value.

It is generally accepted that larger shares of fluctuating RES will increase the amount and magnitude of frequency deviations and consequently the need for frequency control reserves [5]. Although reserves come traditionally from conventional generators, as RES shares increase

less generators will be in the energy mix and therefore additional resources are needed to provide frequency reserves.

Conceptually, loads can also provide frequency control by reducing their consumption when the frequency is low, and by increasing it when the frequency is high. The idea of controlling loads to provide power system AS (including frequency control) was already proposed in the 1980's [6]. Although mostly large industrial loads provided such services in the past, there is a growing interest in exploiting the flexibility of residential and commercial loads as well. If properly aggregated, loads can provide ancillary services more efficiently and at a lower cost compared with generators [7].

Apart from frequency deviations at the transmission system level, RES may create overvoltages and/or overloading of cables and transformers in the local Distribution Networks (DNs), where they are often connected. Load control can, in principle, help to mitigate these problems locally by increasing consumption to absorb the RES power surplus or by reducing consumption to compensate for RES power deficits.

During the intervals when power from RES is abundant, wholesale market electricity prices drop due to the low marginal costs of RES. If this is eventually reflected on the retail market, it can lead to time-varying end-customer electricity prices [8]. In this case, the customers will have an incentive to optimize their load profile by shifting the consumption to the low-price intervals in order to minimize electricity costs.

From the above discussion it becomes clear that new roles and opportunities arise for controllable loads in power systems. Controlling loads to minimize the cost of consumers or provide AS to the power system is commonly referred to as Demand Response (DR). Thermal loads are appropriate for DR because their consumption can be shifted in time without user discomfort due to thermal inertia. Examples of thermal loads are small residential Thermostatically Controlled Loads (TCLs) such as refrigerators, air conditioners, as well as Electric Water Heaters (EWHs), and large Heating, Ventilation and Air-Conditioning (HVAC) systems of commercial buildings.

TCLs are typically controlled by a hysteresis controller based on a temperature setpoint and a deadband. For heating loads, whenever the temperature falls below the lower deadband limit, the appliance turns on and keeps heating until the temperature reaches the higher deadband limit. At this point, the appliance turns off and a new cycle begins. To preserve the user comfort, the temperature must be kept within the

deadband. However, the actual on/off state of the appliance at a particular instance is not important, and this is why the consumption can be shifted in time.

The goal of this thesis is to develop new methods to enable efficient DR programs with thermal loads. In the first part of the thesis, we focus on customer-oriented DR applications aiming to minimize electricity costs by shifting consumption to low-price intervals or maximizing the absorption of locally generated RES energy, for example from rooftop Photovoltaics (PV). In the second part of the thesis, we consider provision of frequency control AS from aggregations of thermal loads, similarly to conventional generators.

The next section introduces the terminology adopted in this thesis for power system frequency control, Section 1.3 summarizes the contributions of this thesis, Section 1.4 presents its organization, whereas Section 1.5 provides a list of the publications published or submitted during the course of the thesis.

1.2 Frequency Control Terminology

Typically, a TSO controls the frequency in three steps, namely primary, secondary and tertiary control. In this thesis, we investigate the provision of Primary Frequency Control (PFC) and Secondary Frequency Control (SFC) from loads. PFC is sometimes referred to as governor control action, frequency response, droop control, or frequency containment reserve (FCR), and the name PFC is used throughout this thesis.

SFC is also known as automatic generation control (AGC), load frequency control (LFC), automatic frequency restoration reserve (FRR), or regulation service (mainly in the USA). In this thesis, we adopt the names SFC and frequency regulation, which we use interchangeably.

1.3 Contributions

The main contributions of this PhD thesis are the following:

- A Model Predictive Control (MPC) scheme to minimize costs in buildings under dynamic day-ahead and real-time electricity prices.

- Rule-Based Control (RBC) algorithms to maximize PV self-consumption in buildings, and the algorithms' integration in the commercial software tool *Polysun*.

- A decentralized stochastic control scheme to provide PFC reserves from refrigerators without real-time communication.

- RBC algorithms to provide SFC reserves from EWHs with various levels of information feedback to the central controller.

- A hierarchical control scheme to provide SFC and voltage regulation at the DN level simultaneously from the same load aggregation.

- A method to estimate the states of individual loads from aggregate power measurements while providing SFC.

- A hierarchical control framework to allow an aggregation of commercial buildings to participate in SFC reserve markets.

- Experimental demonstration of SFC reserve provision from a commercial building using historical and real-time signals from the Pennsylvania, Jersey, and Maryland Power Market (PJM).

- Dynamic frequency simulations using a two-area power system model to evaluate the effect of large shares of loads in PFC and SFC.

1.4   Thesis Organization

The thesis is organized into two parts: Part I focuses on customer-oriented DR applications and includes two chapters, whereas Part II focuses on provision of AS from load aggregations and includes eight chapters. The thesis is divided into the following chapters:

Part I

- **Chapter 2** develops a residential building model with a Heat Pump (HP), an EWH, a battery and PV panels for price-based DR applications. Using the building model, we design an MPC framework for building climate control. We use the MPC controller to evaluate the potential for electricity cost minimization by load shifting, as well as the building's response to day-ahead and real-time dynamic electricity prices.

- **Chapter 3** introduces simple and practical RBC algorithms to maximize the PV self-consumption in buildings using the flexibility of thermal loads and batteries. We integrate the RBC algorithms and the battery model of Chapter 2 in the commercial software tool Polysun, and perform building simulations to evaluate the financial benefits of PV self-consumption and to identify the optimal battery size.

Part II

- **Chapter 4** provides an overview of frequency control in power systems, with an emphasis on PFC and SFC, and discusses relevant control approaches for TCLs and commercial buildings from the literature.

- **Chapter 5** presents a decentralized stochastic control scheme to provide PFC reserves from aggregations of residential refrigerators without real-time communication. The controller has a high performance, avoids load synchronization, addresses a number of practical limitations, and is simple and cheap to implement. Extensive simulations show the controller's robustness and its suitability for PFC reserve provision.

- **Chapter 6** presents a modeling approach for populations of EWHs subject to random water draw events. The State of Charge (SoC) concept for EWHs is introduced and used within four RBC algorithms to control the aggregate consumption of EWHs, such that a SFC signal is tracked. Simulations with realistic EWH data and historical SFC signals highlight the value of state information in the control loop.

- **Chapter 7** proposes a hierarchical control algorithm to provide SFC reserves from aggregations of TCLs without stress on DNs with large penetration of RES. The algorithm relies on AC Optimal Power Flow (OPF) and integer optimization, and is tractable due to the hierarchical structure. The algorithm's performance is demonstrated in a simulation study with a benchmark DN.

- **Chapter 8** develops a Moving Horizon State Estimation (MHSE) method to estimate the states of individual TCLs in SFC schemes. The estimator works with limited communication, namely aggregate power measurements from substations and infrequent TCL

measurements from smart meters. Detailed simulations demonstrate the estimator's performance under different levels of process and measurement noise, as well as different types of SFC signals.

- **Chapter 9** introduces a hierarchical control framework to provide SFC from aggregations of commercial buildings. The framework is comprised of a day-ahead reserve scheduler, a building climate MPC, and a feedback controller to track the SFC signal. Different formulations are proposed leveraging techniques from robust and stochastic optimization. The potential for SFC reserves from buildings is estimated in extensive simulations.

- **Chapter 10** presents results from an experimental demonstration of SFC reserve provision by a commercial building. Using the control framework of Chapter 9 as a basis, we performed experiments over a two-week period using historical and real-time SFC signals from PJM. The results are very promising and show that SFC reserves can be indeed provided from commercial buildings at a high quality.

- **Chapter 11** combines the controllers developed in Chapters 5, 6 and 9 for PFC and SFC into a general control scheme. This scheme is used to perform dynamic frequency simulations with a two-area power system model for different shares of loads in frequency control. The results show that, with proper control design, loads can indeed replace conventional generators in PFC and SFC.

Finally, **Chapter 12** summarizes the key findings of this thesis and suggests directions for future work.

1.5   List of Publications

The work presented in this thesis has been reported in the following publications:

Journal Papers

1. **E. Vrettos**, E. Can Kara, J. MacDonald, G. Andersson, and D. Callaway, *Experimental demonstration of frequency regulation by commercial buildings – Part I: Modeling and hierarchical control*

design, IEEE Transactions on Smart Grid (Early Access Article), 2016.

2. **E. Vrettos**, E. Can Kara, J. MacDonald, G. Andersson, and D. Callaway, *Experimental demonstration of frequency regulation by commercial buildings – Part II: Results and performance evaluation*, IEEE Transactions on Smart Grid (Early Access Article), 2016.

3. **E. Vrettos**, C. Ziras, and G. Andersson, *Fast and reliable primary frequency reserves from refrigerators with decentralized stochastic control*, IEEE Transactions on Power Systems (Early Access Article), 2016.

4. **E. Vrettos**, F. Oldewurtel, and G. Andersson, *Robust energy-constrained frequency reserves from aggregations of commercial buildings*, IEEE Transactions on Power Systems, vol. 31, no. 6, pp. 4272-4285, 2016.

5. **E. Vrettos** and G. Andersson, *Scheduling and provision of secondary frequency reserves by aggregations of commercial buildings*, IEEE Transactions on Sustainable Energy, vol. 7, no. 2, pp. 850-864, 2016.

Conference Papers

1. C. Ziras, **E. Vrettos**, and G. Andersson, *Primary frequency control with refrigerators under startup dynamics and lockout constraints*, in IEEE PES General Meeting, Denver, USA, July 2015.

2. X. Zhang, **E. Vrettos**, M. Kamgarpour, G. Andersson, and J. Lygeros, *Stochastic frequency reserve provision by chance-constrained control of commercial buildings*, in European Control Conference (ECC), Linz, Austria, July 2015.

3. **E. Vrettos**, C. Ziras, and G. Andersson, *Integrating large shares of heterogeneous thermal loads in power system frequency control*, in IEEE PowerTech Conference, Eindhoven, the Netherlands, June 2015.

4. **E. Vrettos**, F. Oldewurtel, F. Zhu, and G. Andersson, *Robust provision of frequency reserves by office building aggregations*, in World Congress of the International Federation of Automatic Control (IFAC), Cape Town, South Africa, August 2014.

5. **E. Vrettos**, J. L. Mathieu, and G. Andersson, *Control of thermostatic loads using moving horizon estimation of individual load states*, in Power Systems Computation Conference (PSCC), Wroclaw, Poland, August 2014.

6. **E. Vrettos**, J. L. Mathieu, and G. Andersson, *Demand response with moving horizon estimation of individual thermostatic load states from aggregate power measurements*, in American Control Conference (ACC), Portland, USA, June 2014.

7. **E. Vrettos** and G. Andersson, *Combined load frequency control and active distribution network management with thermostatically controlled loads*, in IEEE International Conference on Smart Grid Communications, Vancouver, Canada, October 2013.

8. **E. Vrettos**, K. Lai, F. Oldewurtel, and G. Andersson, *Predictive control of buildings for demand response with dynamic day-ahead and real-time prices*, in European Control Conference (ECC), Zurich, Switzerland, October 2013.

9. **E. Vrettos**, A. Witzig, R. Kurmann, S. Koch, and G. Andersson, *Maximizing local PV utilization using small-scale batteries and flexible thermal loads*, in European Photovoltaic Solar Energy Conference (PVSEC), Paris, France, October 2013.

10. **E. Vrettos**, S. Koch, and G. Andersson, *Load frequency control by aggregations of thermally stratified electric water heaters*, in IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe), Berlin, Germany, October 2012.

Other Publications

1. **E. Vrettos**, G. Andersson, S. Koch, A. Witzig, *SmartGrid - Polysun: Design tool for local load management*, Final report of a project funded by the Swiss Federal Office of Energy (BFE), available online at www.bfe.admin.ch, September 2015.

2. **E. Vrettos**, X. Zhang, F. Oldewurtel, M. Kamgarpour, J. Lygeros, and G. Andersson, *Exploring the potential of buildings in the Swiss ancillary service market*, in Computational Optimization of Low-Energy Buildings Workshop (COLEB), Zurich, Switzerland, March 2014.

3. M. Kamgarpour, **E. Vrettos**, G. Andersson, and J. Lygeros, *Population of thermostatically controlled loads for the Swiss ancillary service market*, in Computational Optimization of Low-Energy Buildings Workshop (COLEB), Zurich, Switzerland, March 2014.

The following papers have been published in the course of the PhD studies, but their content is not included in the thesis:

Conference Papers

1. Q. Hu, F. Oldewurtel, M. Balandat, **E. Vrettos**, D. P. Zhou, C. J. Tomlin, *Model identification of commercial building HVAC systems during regular operation - Empirical results and challenges*, in American Control Conference (ACC), Boston, USA, July 2016.

2. V. Poulios, **E. Vrettos**, F. Kienzle, E. Kaffe, H. Luternauer, and G. Andersson, *Optimal placement and sizing of battery storage to increase the PV hosting capacity of low voltage grids*, in International ETG Congress, Bonn, Germany, November 2015.

3. G. S. Ledva, **E. Vrettos**, S. Mastellone, G. Andersson, and J. L. Mathieu, *Applying networked estimation and control algorithms to address communication bandwidth limitations and latencies in demand response*, in Hawaii International Conference on System Sciences (HICSS), Kauai, USA, January 2015.

4. S. Karagiannopoulos, **E. Vrettos**, P. C. López, M. Vrakopoulou, F. Oldewurtel, G. Andersson, and M. Zima, *On geographical allocation of primary frequency control reserves in large interconnected power systems*, in Power Systems Computation Conference (PSCC), Wroclaw, Poland, August 2014.

5. S. Karagiannopoulos, **E. Vrettos**, G. Andersson, and M. Zima, *Scheduling and real-time control of flexible loads and storage in electricity markets under uncertainty*, in European Energy Market Conference (EEM), Krakow, Poland, May 2014.

6. F. Oldewurtel, T. Borsche, M. Bucher, P. Fortenbacher, M. González Vayá, T. Haring, J. L. Mathieu, O. Mégel, **E. Vrettos**, and G. Andersson, *A framework for and assessment of demand response and energy storage in power systems*, in Bulk Power System Dynamics and Control Symposium - IX Optimization, Security and Control of the Emerging Power Grid (IREP), Rethymnon, Greece, August 2013.

7. **E. Vrettos**, F. Oldewurtel, M. Vasirani, and G. Andersson, *Centralized and decentralized balance group optimization in electricity markets with demand response*, in PowerTech Conference, Grenoble, France, June 2013.

The following papers have been submitted and are currently under review:

Journal Papers

1. C. Ziras, **E. Vrettos**, and S. You, *Controllability and stability of primary frequency control from thermostatic loads with delays*, Journal of Modern Power Systems and Clean Energy (in review).

2. J. Lehr, **E. Vrettos**, R. Jain, R. Rajagopal, and M. P. Everts, *Financial viability of residential PV/battery systems in California: A case-study*, Energy Journal (in review).

3. G. S. Ledva, **E. Vrettos**, S. Mastellone, G. Andersson, and J. L. Mathieu, *Investigating the impact of communication delays and plant-model mismatch in demand response*, IEEE Transactions on Power Systems (in review).

Part I

Customer-Oriented Demand Response Applications

Chapter 2

Price-Based Optimal Building Control

2.1 Introduction

This chapter concerns the optimal building operation in the presence of time-varying end-customer electricity tariffs. In principle, and with the appropriate price incentives, part of a building's consumption can be shifted in time due to thermal inertia. Optimal building control has gained a lot of attention in the literature, and some of the relevant previous works are summarized below.

Most of the early work investigated the use of building thermal mass for load shifting and peak shedding (see [9] for an overview of relevant simulation and experimental results up to 2003). For example, [10] investigated in a simulation study how the zone temperature setpoints can be optimally varied to reduce peak demand and take advantage of low nighttime electrical rates, whereas [11] experimentally demonstrated the potential for load shifting with optimal control. Predictive optimal control was compared against conventional control strategies (chiller-priority and storage-priority control) in [12] considering the minimization of either energy charges or peak demand charge, and in [13] considering real-time pricing.

Load shifting and peak load reduction have been addressed in more recent works as well. In [14], a simple optimization model that allows a

Table 2.1: Nomenclature of Chapter 2: symbols

Symbol	Unit	Description
A, k_{b}, q	-	PV model parameters
b_1	-	Battery capacity ratio parameter
b_2	1/h	Battery rate constant
C	J/K	Thermal capacity
C_{FL}	-	Numbers of battery cycles to failure
c_{w}	J/(Kg K)	Specific heat capacity of water
c_0	-	Parameter of HP model
c_1, c_2	K^{-1}	Parameters of HP model
$E, E_{\max}$	Wh	Total battery energy and battery capacity
E_1, E_2	Wh	Available and chemically bound battery energy
G	$\mathrm{W/m}^2$	Solar radiation
g	$\mathrm{m/s}^2$	Gravitational acceleration
I	A	Electric current
K_1, K_2	-	PV model parameters
k_{fz}	W/K	Heat exchange coeff. between T_{f} and T_{z}
k_{th}	1/s	Heat loss coefficient
k_{wf}	W/K	Heat exchange coeff. between $T_{\mathrm{w,r}}$ and T_{f}
m	kg	Mass
$\dot{m}_{\mathrm{w}}$	kg/s	Water mass flow rate
N_{PV}	-	Number of PV panels
n_{c}	-	Battery charging efficiency
n_{conv}	-	Battery converter efficiency
n_{d}	-	Battery discharging efficiency
P	W	Power
R	-	Range of a charge/discharge battery cycle
q_1	Ah	Available battery charge
q_2	Ah	Chemically bound battery charge
T	°C	Temperature
$T_{\min}^{\mathrm{c}}, T_{\max}^{\mathrm{c}}$	°C	RBC room cooling thresholds
$T_{\min}^{\mathrm{h}}, T_{\max}^{\mathrm{h}}$	°C	RBC room heating thresholds
$T_{\min}^{\mathrm{w}}, T_{\max}^{\mathrm{w}}$	°C	RBC water heating thresholds
U	V	Voltage
u	-	Building input
V	m/s	Vertical water velocity in the EWH
v	-	Building disturbance
x	-	Building state
$\Delta I_{\mathrm{p}}, \Delta U_{\mathrm{p}}$	-	PV model parameters
Δt	s	Discretization in time for EWH model
Δy	m	Discretization in space for EWH model
$\alpha_1 - \alpha_5$	-	Parameters of the battery lifetime model
$\varepsilon_{\mathrm{eff}}$	-	EWH turbulent mixing parameter
ρ	$\mathrm{kg/m}^3$	Density

consumer to adapt its demand profile in response to electricity prices was developed assuming all loads to be shiftable. Reference [15] presented a prototype that applies stochastic dynamic programming with weather forecasts to minimize the electricity cost by shifting heating and cooling loads as well as event-based loads (e.g., dishwashers).

Table 2.2: Nomenclature of Chapter 2: subscripts and superscripts

Subscript / superscript	Description
a	Ambient temperature
b	Battery-related variable
c	Battery charge power
cell	PV cell temperature
cw	Cold water
d	Battery discharge power
el	Electric power
f	Floor
k	Discrete time index
MPPT	Maximum power point tracking value
min/max	Minimum/maximum value of a variable
NOCT	Normal operating PV cell temperature
p	Peak value
s	Standard conditions (used in the PV model)
t	Continuous time index
th	Thermal power
w	Water (temperature)
wh	EWH-related variable
wr	Water return
ws	Water supply
z	Room temperature

In [16], the authors developed a stochastic Model Predictive Control (MPC) scheme based on a bilinear building model and weather predictions with the goal of increasing the energy efficiency. An MPC controller for building cooling systems equipped with thermal energy storage was proposed in [17] with the aim of achieving lower electricity costs. Reference [18] presented the practical implementation of an MPC controller in a building, in order to stabilize the power fluctuations due to high penetration of wind power. In [19], the minimization of electricity costs for residential buildings with elastic demand, storage, local generation and real-time pricing was formulated as a stochastic optimization problem and approximately solved using the so-called Lyapunov approach.

In this chapter, we develop a state-of-the-art model of a residential building and investigate its potential for Demand Response (DR). The building model includes a Heat Pump (HP) for space heating, Slab Cooling (SC) for space cooling, an Electric Water Heater (EWH), Photovoltaics (PV) and battery storage. We use the developed model within an MPC to estimate upper bounds on the potential for electricity cost reduction in a dynamic end-customer electricity price environment. In addition, we investigate the sensitivity of the building's response to day-

ahead and real-time price signals, and furthermore the effectiveness of employing real-time price control on top of day-ahead price profiles.

The remainder of this chapter is organized as follows. In Section 2.2, we introduce the building model including a detailed description of its components. In Section 2.3 we design the MPC controller, and in Section 2.4 we present simulation results from several case studies. Finally, Section 2.5 concludes this chapter, which is based on the work published in [20–22]. The chapter's nomenclature is introduced in Tables 2.1 and 2.2 (some of the EWH variables are summarized in Table 2.3).

2.2   Modeling

2.2.1   Building Thermal Model

We use a thermal model for buildings with Integrated Room Automation (IRA) systems (where heating, cooling, ventilation, blinds, and lighting are jointly controlled), which was developed in the course of the OptiControl project [23–25]. A single-zone model was obtained within this project by applying a common method to aggregate building zones in order to compute the building-wide energy use.

Figure 2.1a shows a schematic representation of the OptiControl building model. It includes 12 states that represent the temperatures in the room and in different layers in the floor, walls and ceiling. Different layers are used to model the heat transfer properties of different materials, as well as the influence of different heating and cooling actuators, e.g., the fact that floor heating affects the room temperature with some delay.

The model is a Resistance-Capacitance (RC), multiple-input-multiple-output bilinear model and it is validated against the well-known building simulation software TRNSYS [25]. The model's satisfactory accuracy and its relatively low complexity make it suitable for MPC. A model similar to the one considered here was used in an MPC implementation in a real building and was found to capture the building's thermal dynamics well [26]. Furthermore, this model is flexibly customizable and allows us to perform large-scale simulation studies with different representative building types.

(a) Schematic of the building thermal model with states the temperatures in the room, floor, walls and ceiling.

(b) Integration of an HP model into the building thermal model. The water supply and return temperatures are additional states.
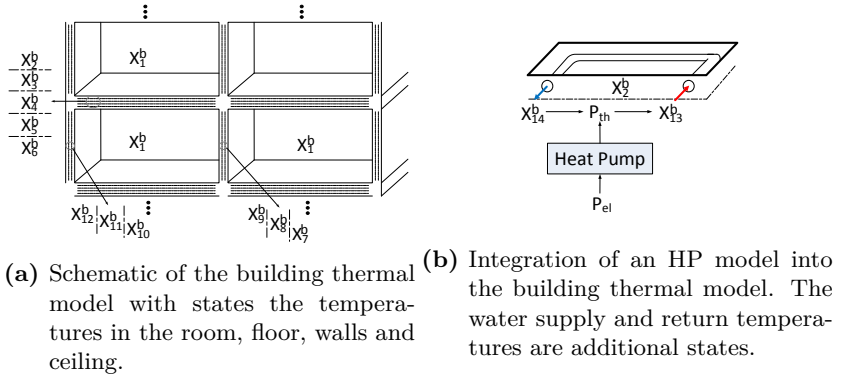
Figure 2.1: Extension of the OptiControl building model with a HP model.

The discrete-time bilinear model can be written in the form

$$x_{k+1} = Ax_k + B_u u_k + B_v v_k + \sum_{i=1}^{n_u} (B_{vu,i} v_k + B_{xu,i} x_k) u_{k,i} , \qquad (2.1)$$

where x_k, u_k, v_k are the state, input and disturbance vectors, respectively, and A, B_u, B_v and $B_{vu,i}, B_{xu,i}$ $i = 1, \ldots n_u$ are appropriate matrices that define the building dynamics. The n_x states include the temperatures of the room, walls, floor and ceiling (all measured in °C). The n_u IRA control inputs include heating and cooling power, ventilation, blind position, and lighting. The heating and cooling are represented in the thermal model as heat fluxes affecting the system states and their units are W/m², i.e., the heat fluxes are normalized by the floor area. The blind position is a number between 0 (fully closed) and 1 (fully open). The lighting is also normalized by the floor area and measured in W/m². The n_v disturbances include the ambient temperature in °C, the solar radiation in W/m², and the internal heat gains by the occupants and equipment in W/m².

Note that bilinearities exist in (2.1) between the control inputs u_k and the disturbances v_k, as well as between u_k and the states x_k. Three typical examples are: (i) bilinearity between the blind position and solar radiation; bilinearity between the blind position and room temperature; and (iii) bilinearity between the ventilation air flow rate and room temperature.

For optimization purposes, it is desirable to have a linear model because

any nonlinear model would result in a non-convex optimization problem. If the disturbances are fixed, for example according to their predicted values, the bilinearity between u_k and v_k vanishes and the system becomes time-varying. We handle the remaining bilinearity between x_k and u_k using a Sequential Linear Programming (SLP) technique, where the bilinear x_k terms are iteratively linearized around the most recently calculated x_k trajectory and the resulting convex Linear Program (LP) is solved until convergence, i.e., until the change in optimal solution between two iterations is smaller than a threshold. This SLP technique was shown to be a very good approximation of the original problem in [23, 24, 26] because the non-linearities are mild. Although the SLP is a heuristic approach and there are no guarantees for convergence to the globally optimal solution, practical experience has shown that convergence is usually achieved within a few iterations. Using the SLP technique, the building model can be written as the linear system

$$x_{k+1} = Ax_k + B_{u,k}u_k + B_v v_k \ , \tag{2.2}$$

where the input matrix $B_{u,k}$ is now time-varying.

In this chapter, we consider different building types that vary in terms of the actuators of the Heating, Ventilation and Air-Conditioning (HVAC) system, the insulation level of the building envelope, the window area fraction, and the magnitude of the internal heat gains. The interested reader is referred to [23, 24] for more information about the building models.

2.2.2   Heat Pump Model

In this section we integrate a steady-state HP model into the building thermal model of Section 2.2.1. The HP is modeled by the Coefficient of Performance (COP), i.e., the ratio of the output thermal power P_{th} to the input electric power P_{el}. In general, the COP is a function of the ambient temperature T_{a} and the water supply temperature T_{ws}. In this chapter, a linear function is assumed as in [27], which results in the following equations for COP and P_{th}

$$\text{COP} = c_0 + c_1 T_{\text{a}} + c_2 T_{\text{ws}} \tag{2.3}$$

$$P_{\text{th}} = (c_0 + c_1 T_{\text{a}} + c_2 T_{\text{ws}}) \cdot P_{\text{el}} \ . \tag{2.4}$$

To account for the interaction between the HP and the building, we model the thermal dynamics of the water heating loop by considering

two additional states: the water supply temperature T_{ws}, and the water return temperature T_{wr}. These dynamics can be modeled using the following state space model [27]

$$
\begin{bmatrix} \dot{T}_{\mathrm{ws}} \\ \dot{T}_{\mathrm{wr}} \\ \dot{T}_{\mathrm{f}} \end{bmatrix} = \begin{bmatrix} -\frac{\dot{m}_{\mathrm{w}} c_{\mathrm{w}}}{C_{\mathrm{ws}}} & \frac{\dot{m}_{\mathrm{w}} c_{\mathrm{w}}}{C_{\mathrm{ws}}} & 0 \\ \frac{\dot{m}_{\mathrm{w}} c_{\mathrm{w}}}{C_{\mathrm{wr}}} & -\frac{\dot{m}_{\mathrm{w}} c_{\mathrm{w}} + k_{\mathrm{wf}}}{C_{\mathrm{wr}}} & \frac{k_{\mathrm{wf}}}{C_{\mathrm{wr}}} \\ 0 & \frac{k_{\mathrm{wf}}}{C_{\mathrm{f}}} & -\frac{k_{\mathrm{wf}} + k_{\mathrm{fz}}}{C_{\mathrm{f}}} \end{bmatrix} \cdot \begin{bmatrix} T_{\mathrm{ws}} \\ T_{\mathrm{wr}} \\ T_{\mathrm{f}} \end{bmatrix} + \begin{bmatrix} \frac{1}{C_{\mathrm{ws}}} \\ 0 \\ 0 \end{bmatrix} \cdot P_{\mathrm{th}},
$$

$$(2.5)$$

where T_{f} denotes the floor temperature; $\dot{m}_{\mathrm{w}}$ is the mass flow rate of the water loop; c_{w} is the specific heat capacity of water; C_{ws} is the thermal capacity of supply water at temperature T_{ws}; C_{wr} is the thermal capacity of return water at temperature T_{wr}; C_{f} is the thermal capacity of floor at temperature T_{f}; k_{wf} is the heat exchange coefficient between T_{wr} and T_{f}; and k_{fz} is the heat exchange coefficient between T_{f} and the room temperature T_{z}. The HP parameters are given in Appendix B.

Since the input P_{el} depends on the state T_{ws}, the HP dynamics are nonlinear. To keep the resulting MPC problem convex, we linearize the dynamics by fixing T_{ws} in (2.3) to its steady state value. We estimate the steady state value of T_{ws} by taking the continuous-time building model, fixing the disturbance vector to its expected value, setting the derivatives to zero ($\dot{x} = 0$), and solving the resulting system of linear equalities. If this linearization approach is combined with a mixed cost function with a linear penalty for the electricity cost and a quadratic penalty for the HP power, then the resulting convex optimization problem will be a good approximation of the original one [27].

As depicted in Fig. 2.1b, we assume that the produced heat from the HP acts on the temperature of the first layer of the floor, i.e., T_{f} corresponds to $\mathrm{X}_2^{\mathrm{b}}$. We augment the state vector with T_{ws}, T_{wr} and integrate (2.5) in (2.2).

2.2.3 Electric Water Heater Model

Different types of EWH models have been proposed in the literature. Most of these models have only one state and assume a uniform temperature distribution in the water tank [28, 29]. Such models might not predict the EWH consumption accurately enough, because they do not capture the thermal stratification in the tank. For this reason, some researchers have considered thermal stratification. Reference [30] proposed a simplified stratified model considering only an upper hot water

layer and a lower cold water layer. A more detailed stratified model with six water layers was developed in [31], but without considering convective heat transfer phenomena.

In this section, we present a detailed thermally-stratified model for EWHs, which accounts for the heat transfer mechanisms inside the storage tank and demonstrates a good tradeoff between accuracy and complexity.

Governing Partial Differential Equation

If natural convection is neglected, the heat flow in the water tank is governed by the one-dimensional Partial Differential Equation (PDE)

$$\frac{\partial T}{\partial t} + V \frac{\partial T}{\partial y} = a\varepsilon_{\text{eff}} \frac{\partial^2 T}{\partial^2 y} - k_{\text{th}}(T - T_{\text{a}}) + Q(y, t) \; , \qquad (2.6)$$

where y denotes the position along the vertical axis of the tank, t denotes time, V is the vertical water velocity in the tank during a water draw, a is the thermal diffusivity, k_{th} is the heat loss coefficient, T_{a} is the ambient temperature, $Q(y, t)$ corresponds to the internal heat generation from the heater, and ε_{eff} accounts for the turbulent mixing at the tank inlet during a water draw [32].

If $\varepsilon_{\text{eff}} = 1$, the incoming water flow is laminar, whereas if $\varepsilon_{\text{eff}} \gg 1$, the flow is turbulent. Including ε_{eff} in (2.6) allows us to model the initial cooling of the whole tank right after a water draw occurs, which is observed in relevant experiments with EWHs.[1]

Numerical Solution Scheme

Equation (2.6) is a parabolic PDE with the additional convective term $V \frac{\partial T}{\partial y}$. A standard approach to numerically solve a parabolic PDE is a first-order upwind differencing scheme. However, such a scheme suffers from numerical diffusion problems that might lead to severe inaccuracies. For this reason, we use the second-order, three-level, finite difference Crank-Nicolson scheme proposed in [34] to numerically solve (2.6).

---

[1]Note that ε_{eff} is fixed to 1 in (2.6) if the simulation time step is greater than a few seconds due to numerical stability issues [33].
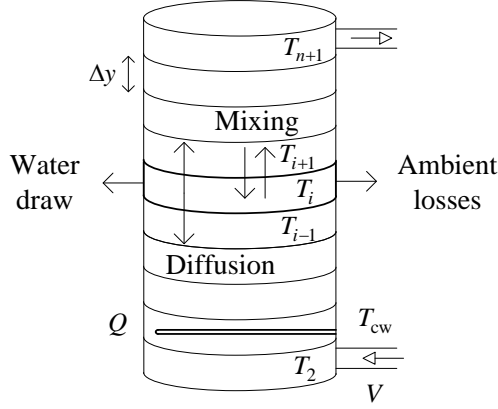
Figure 2.2: Schematic of the thermally-stratified EWH model.

The water tank is divided into a number of layers of the same volume along the tank's vertical axis, as shown in Fig. 2.2. The discretization in time is denoted by Δt, the discretization in space by Δy, the number of grid points in the y direction by n, and the temperature of the ith layer at time step k by T_i^k. Equation (2.6) can be written in the more general form

$$\frac{\partial T}{\partial t} = a\varepsilon_{\text{eff}}\frac{\partial^2 T}{\partial^2 y} + f\left(T, \frac{\partial T}{\partial y}\right) \tag{2.7}$$

$$f\left(T, \frac{\partial T}{\partial y}\right) = -V\frac{\partial T}{\partial y} - k_{\text{th}}(T - T_{\text{a}}) + Q(y,t) \ . \tag{2.8}$$

The Crank-Nicolson scheme is defined by the substitutions

$$\frac{\partial T}{\partial t} \rightarrow \frac{T_i^k - T_i^{k-1}}{\Delta t} \tag{2.9}$$

$$\frac{\partial^2 T}{\partial^2 y} \rightarrow \frac{1}{2}\left[\frac{T_{i-1}^k - 2T_i^k + T_{i+1}^k}{\Delta y^2} + \frac{T_{i-1}^{k-1} - 2T_i^{k-1} + T_{i+1}^{k-1}}{\Delta y^2}\right] \tag{2.10}$$

$$f\left(T, \frac{\partial T}{\partial y}\right) \rightarrow f\left[\frac{T_i^k + T_i^{k-1}}{2}, D\left(\frac{T_i^k + T_i^{k-1}}{2}\right)\right] \tag{2.11}$$

$$D\left(\frac{T_i^k + T_i^{k-1}}{2}\right) = \frac{\frac{T_{i+1}^k + T_{i+1}^{k-1}}{2} - \frac{T_{i-1}^k + T_{i-1}^{k-1}}{2}}{2\Delta y} \ . \tag{2.12}$$

Applying this scheme on (2.7) results in the following expressions for all tank layers apart from the top and bottom layers

$$-q_1 T_{i-1}^k + q_2 T_i^k + q_3 T_{i+1}^k =$$
$$q_1 T_{i-1}^{k-1} + q_4 T_i^{k-1} - q_3 T_{i+1}^{k-1} + k_{\text{th}} T_{\text{a}} + Q(x,t) \ , \quad (2.13)$$

$$\text{where:} \quad q_1 = \frac{e}{2} + d, \quad q_2 = \frac{1}{\Delta t} + e + \frac{k_{\text{th}}}{2}, \quad q_3 = d - \frac{e}{2} \ , \quad (2.14)$$

$$q_4 = \frac{1}{\Delta t} - e - \frac{k_{\text{th}}}{2}, \quad d = \frac{V}{4\Delta y}, \quad e = \frac{a\varepsilon_{\text{eff}}}{\Delta y^2} \ . \quad (2.15)$$

We consider two artificial layers to represent the incoming water temperature ($T_1 = T_{\text{cw}}$) and the ambient temperature ($T_{n+2} = T_{\text{a}}$), in order to facilitate the problem formulation in matrix form. Thus, the temperature of the bottom layer is denoted by T_2, and the temperature of the top layer by T_{n+1}. We use the following Ordinary Differential Equations (ODEs) as boundary conditions for the top and bottom layer

$$\frac{dT_2}{dt} = \frac{-a\varepsilon_{\text{eff}}}{\Delta y^2}(T_2 - T_3) + \frac{V}{\Delta y}(T_{\text{cw}} - T_2) - k'_{\text{th}}(T_2 - T_{\text{a}}) \ , \quad (2.16)$$

$$\frac{dT_{n+1}}{dt} = \frac{a\varepsilon_{\text{eff}}}{\Delta y^2}(T_n - T_{n+1}) + \frac{V}{\Delta y}(T_n - T_{n+1}) - k'_{\text{th}}(T_{n+1} - T_{\text{a}}). \quad (2.17)$$

By applying an explicit discretization scheme on these ODEs, we obtain

$$\frac{1}{\Delta t}T_2^k - 4dT_{\text{cw}} - k'_{\text{th}}T_{\text{a}} = q_5 T_2^{k-1} + e T_3^{k-1} \quad (2.18)$$

$$\frac{1}{\Delta t}T_{n+1}^k - k'_{\text{th}}T_{\text{a}} = q_5 T_{n+1}^{k-1} + q_6 T_n^{k-1} \quad (2.19)$$

$$\text{where:} \quad q_5 = \frac{1}{\Delta t} - e - 4d - k'_{\text{th}}, \quad q_6 = e + 4d \ . \quad (2.20)$$

In (2.16)–(2.20), k'_{th} is the heat loss coefficient of the top and bottom layers ($k'_{\text{th}} > k_{\text{th}}$ due to larger heat loss area). Equations (2.13),(2.18),(2.19) can be expressed in the matrix form

$$\mathbf{Z}_1 \mathbf{T}_{k+1} = \mathbf{Z}_2 \mathbf{T}_k + \mathbf{Z}_3 u_k \ , \quad (2.21)$$

where $\mathbf{T}_k \in \mathbb{R}^{n+2}$ is the temperature vector, $u_k \in \{0,1\}$ is a binary

on/off variable, and $\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3$ are defined as

$$\mathbf{Z}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ -4d & 1/\Delta t & 0 & 0 & 0 & \dots & 0 & 0 & 0 & -k'_{\text{th}} \\ 0 & -q_1 & q_2 & q_3 & 0 & \dots & 0 & 0 & 0 & -k_{\text{th}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \dots & & \vdots & & \\ 0 & 0 & 0 & 0 & 0 & \dots & -q_1 & q_2 & q_3 & -k_{\text{th}} \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1/\Delta t & -k'_{\text{th}} \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 1 \end{bmatrix} \tag{2.22}$$

$$\mathbf{Z}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & q_5 & e & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & q_1 & q_4 & -q_3 & 0 & \dots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \dots & & \vdots & \vdots & \\ 0 & 0 & 0 & 0 & 0 & \dots & q_1 & q_4 & -q_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & q_6 & q_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 1 \end{bmatrix} \tag{2.23}$$

$$\mathbf{Z}_3 = \begin{bmatrix} 0 & \dots & 0 & \frac{\eta P_{\text{el}}}{m_i c_{\text{w}}} & 0 & \dots & 0 \end{bmatrix}, \tag{2.24}$$

where P_{el} is the nominal electric power of the heating element, η is the electric efficiency, m_i is the water mass of the ith layer and c_{w} is the specific heat capacity of water. Although (2.24) implies that the heating zone has only one layer, larger zones can also be considered.

It is easy to show that matrix $\mathbf{Z}_1$ is nonsingular by construction, and thus invertible. Therefore, (2.21) can be rewritten in the form

$$\mathbf{T}_{k+1} = \mathbf{A}_k \mathbf{T}_k + \mathbf{B}_k u_k, \tag{2.25}$$

where $\mathbf{A}_k = \mathbf{Z}_1^{-1} \mathbf{Z}_2$ and $\mathbf{B}_k = \mathbf{Z}_1^{-1} \mathbf{Z}_3$.

Modeling Natural Convection

Natural convection is the combined effect of two opposite forces: buoyancy that accelerates warm water layers upwards, and viscous friction that opposes the fluid motion. To account for natural convection, the buoyant force at each water layer is calculated with

$$F_i^{\text{b}} = \min \left[0, \; g m_i \cdot \frac{\rho_{\text{ref}} - \rho_i}{\rho_{\text{ref}} + \rho_i} \right], \tag{2.26}$$
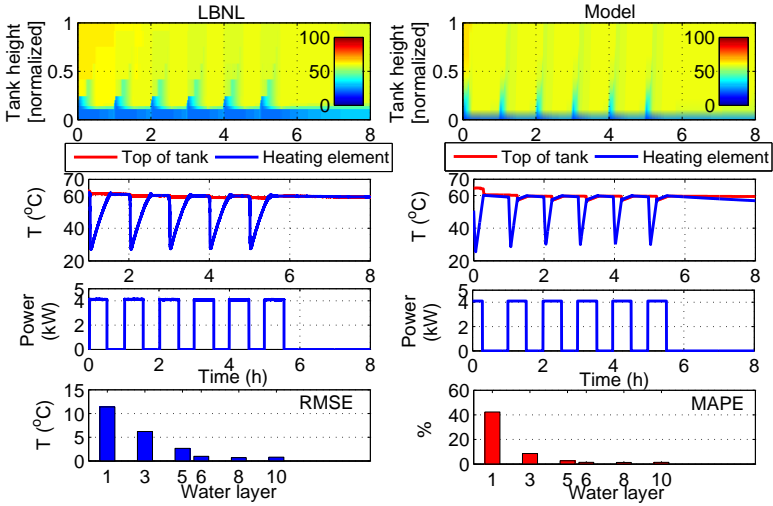
Figure 2.3: EWH model validation: comparison of simulation results (right) with experimental data (left). The model predicts the thermal stratification within the water storage tank sufficiently well.

where ρ_i is the density of the ith water layer, and ρ_{ref} is the minimum between ρ_{i+1} and the average density of the water layers above layer i. Using the natural convection velocity V_i, the viscous force at each layer is calculated with

$$F_i^{\text{v}} = -b \, V_i \; . \tag{2.27}$$

The parameter b depends both on the water viscosity and tank geometry. With both forces known, applying the Newton's law of motion gives us the displacement Δi of each water layer, which determines the degree of mixing with the upper layers according to

$$T_{j,k+1} = rT_{j+1,k} + (1-r)T_{j,k}, \qquad \text{for } j \in [i, i + \Delta i] \; , \tag{2.28}$$

where $r \in [0, 1]$ is the mixing ratio parameter. Both b and r can be identified from temperature measurements at different layers in the tank. Modeling buoyancy in such a way captures the initial cooling of the upper part of the tank, while the lower part is heated up, which is observed in relevant experiments. Equations (2.26), (2.27) and (2.28) are applied after (2.6) is numerically solved at each time step.

Model Tuning and Experimental Validation

The model was validated and tuned using measurements from ten different positions along the vertical axis of a cylindrical water tank from the Lawrence Berkeley National Laboratory (LBNL) [35]. The tank was subject to six water draws with a flow rate of 11.4 l/min, lasting for 3.5 minutes each, which occurred at the beginning of each hour. Six of the thermometers were placed in the center of six zones with equal volumes. The exact position of the other four thermometers was not specified, but a preliminary analysis suggested that they were clustered around the lowest thermometer.

Table 2.3: EWH Model Parameters

Parameter	Symbol	Value
Volume	-	190 l
Height	-	1.19 m
Power	P_{el}	4.1 kW
Efficiency	η	0.95
Heater position	-	0.24 m
Water thermal diffusivity	a	$0.1434 \cdot 10^{-6}$ m^2/s
Water specific heat capacity	c	4185.5 J/(kg K)
Tank heat loss coeff. 1	k_{th}	$6.3588 \cdot 10^{-7}$ 1/s
Tank heat loss coeff. 2	k'_{th}	$1.2382 \cdot 10^{-6}$ 1/s
Heating zone size	-	2
Friction parameter	b	1 kg/s
Mixing ratio	r	0.7
Discretization in time	Δt	10 s
Discretization in space	Δy	0.119 m
Number of layers	n	10
Inlet water temperature	T_{cw}	14.4°C
Ambient temperature	T_{a}	19.7°C

We give the fitted EWH parameters in Table 2.3, whereas Fig. 2.3 shows the model validation results. The plots on the left correspond to measurements, whereas the plots on the right correspond to model predictions. From top to bottom Fig. 2.3 shows: (i) the temperature distribution in the water tank, (ii) the temperature evolution at the top of the tank and at the heating element, (iii) the electric power consumption, and (iv) the Root Mean Squared Error (RMSE) and the Mean Absolute Percentage Error (MAPE) of the model predictions.

With the exception of the lowest water layer, the temperature errors are small, particularly for the upper tank layers (less than 3%). The increased model mismatch at the bottom of the tank is due to the uncertainty related with the lower thermometers' position. In addition, the heater operating intervals predicted by the model match the experimental results very well.

2.2.4   Battery Model

In this thesis, we consider lead-acid batteries as a storage technology due to their low investment cost, which makes them preferable for stationary applications. Since we are interested in the energy flows among the battery, the building loads, the PV and the grid, a model that describes the evolution of the stored energy is sufficient.

The Kinetic Battery Model (KiBaM) is appropriate to simulate lead-acid batteries [36].[2] This model considers the battery as a two-well system, where the first well contains the directly available charge q_1 and the second one contains the chemically bound charge q_2, which can be transformed to electricity only at a limited rate. KiBaM accounts for the capacity reduction at increased charge or discharge current, as well as the so-called battery recovery effect.

A schematic of the KiBaM model is shown in Fig. 2.4. The ratio between the widths of the two wells is called capacity ratio parameter and is denoted by b_1. The rate at which the bound charge becomes available is proportional to the difference in head of the two wells (h_1 and h_2) and is denoted by b_2. The battery can be described as the system of ODEs

$$\frac{dq_1}{dt} = -I - b_2' \cdot (h_1 - h_2) \tag{2.29}$$

$$\frac{dq_2}{dt} = b_2' \cdot (h_1 - h_2) \ , \tag{2.30}$$

where I is the charge/discharge current and $b_2' = b_2 b_1 (1 - b_1)$.

---

[2]The KiBaM equations for battery voltage hold for lead-acid batteries because of their flat discharge profile. These equations do not hold for lithium-ion batteries, which have a sloped discharge profile. However, if the goal is not to model the battery voltage during discharge, but instead to model the energy transfer and the battery lifetime, the two-well model of KiBaM is also applicable for lithium-ion batteries [37].
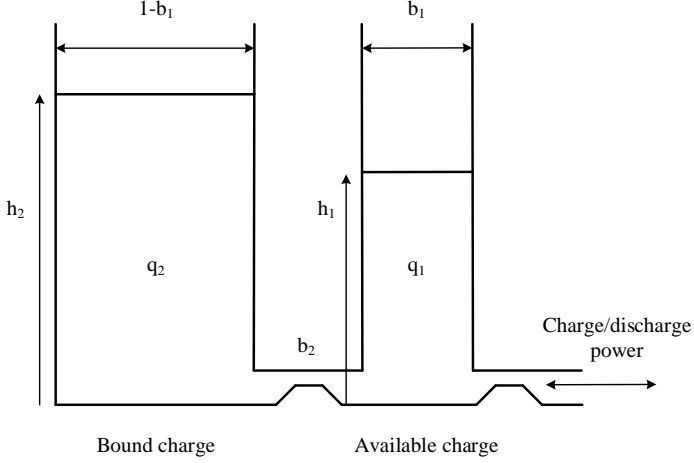
Figure 2.4: KiBaM models the battery as a two-well system, where the first well contains the available charge (q_1) and the second well contains the chemically bound charge (q_2).

Energy and Power Models

Since we are not interested in electrical quantities such as voltages and currents, we use the KiBaM variant proposed in [22]. Assuming that the battery terminal voltage is constant, the time evolution of the available energy E_k^1 and the chemically bound energy E_k^2 can be described by the discrete-time state space model

$$
\underbrace{\begin{bmatrix} E_{1,k+1} \\ E_{2,k+1} \end{bmatrix}}_{x_{k+1}^{\text{bat}}} = \underbrace{\begin{bmatrix} b_1 + b_3 b_4 & b_1(1-b_4) \\ b_3(1-b_4) & b_3 + b_1 b_4 \end{bmatrix}}_{A^{\text{bat}}} \cdot \underbrace{\begin{bmatrix} E_{1,k} \\ E_{2,k} \end{bmatrix}}_{x_k^{\text{bat}}} +
$$

$$
\underbrace{\begin{bmatrix} \frac{b_3 b_4 + b_1(1-b_2\Delta t)-1}{b_2} \cdot \frac{1}{n_\text{d}} & \frac{b_3 b_4 + b_1(1-b_2\Delta t)-1}{b_2} \cdot n_\text{c} \\ \frac{-b_3(b_2\Delta t-1+b_4)}{b_2} \cdot \frac{1}{n_\text{d}} & \frac{-b_3(b_2\Delta t-1+b_4)}{b_2} \cdot n_\text{c} \end{bmatrix}}_{B^{\text{bat}}} \cdot \underbrace{\begin{bmatrix} P_k^\text{d} \\ P_k^\text{c} \end{bmatrix}}_{u^{\text{bat}}}, \quad (2.31)
$$

where the total energy stored in the battery is equal to $E_k = E_{1,k} + E_{2,k}$, the subscript k denotes time, $b_3 = 1 - b_1$, $b_4 = e^{-b_2\Delta t}$, Δt is the discretization time, n_d is the discharging efficiency, n_c is the charging efficiency, P_k^d is the discharge power, and P_k^c is the charge power.

The battery round-trip efficiency is equal to $n_{\mathrm{rt}} = n_{\mathrm{c}} n_{\mathrm{d}}$. The DC/AC converter is simply modeled by its rated capacity and efficiency n_{conv}, which is assumed to be constant.

Apart from the energy model (2.31), KiBaM also models the maximum charge $P_{\mathrm{ch},k}^{\max}$ and discharge $P_{\mathrm{dis},k}^{\max}$ power as a function of the stored energy in the battery according to

$$
P_{\mathrm{ch},k}^{\max} = \frac{-b_2 \cdot b_1 \cdot E_{\max} + b_2 \cdot E_{1,k} \cdot e^{-b_2 \cdot \Delta t} + E_k \cdot b_2 \cdot b_1 \cdot (1 - e^{-b_2 \cdot \Delta t})}{1 - e^{-b_2 \cdot \Delta t} + b_1 \cdot (b_2 \cdot \Delta t - 1 + e^{-b_2 \cdot \Delta t})}
\tag{2.32}
$$

$$
P_{\mathrm{dis},k}^{\max} = \frac{b_2 \cdot E_{1,k} \cdot e^{-b_2 \cdot \Delta t} + E_k \cdot b_2 \cdot b_1 \cdot (1 - e^{-b_2 \cdot \Delta t})}{1 - e^{-b_2 \cdot \Delta t} + b_1 \cdot (b_2 \cdot \Delta t - 1 + e^{-b_2 \cdot \Delta t})} \; ,
\tag{2.33}
$$

where $E_{\max}$ is the nominal battery capacity. The SoC is defined as

$$
\mathrm{SOC}_k = \frac{E_k}{E_{\max}} \; .
\tag{2.34}
$$

Battery Lifetime Model

Battery lifetime estimation is essential to quantify the return on a battery investment. For this purpose, we apply the rainflow cycle counting method [22,38] that models the dependence of the lifetime on the range and number of charge/discharge cycles. The number of cycles to failure C_{FL} is a double exponential function of the charge/discharge cycle range R

$$
C_{\mathrm{FL}} = \alpha_1 + \alpha_2 \cdot e^{\alpha_3 \cdot R} + \alpha_4 \cdot e^{\alpha_5 \cdot R} \; .
\tag{2.35}
$$

The parameters α_1 to α_5 may be available from the battery manufacturer or obtained via a non-linear regression on lifetime data.

We consider twenty bins of the same width and allocate the battery cycles to them depending on their range. Let M_i denote the annual number of cycles with a range R_i, and $C_{\mathrm{FL},i}$ denote the respective cycles to failure. In each of these cycles, $1/C_{\mathrm{FL},i}$ of the entire battery lifetime is consumed, and so the cumulative annual damage is

$$
D = \sum_i \frac{M_i}{C_{\mathrm{FL},i}} \; .
\tag{2.36}
$$

For instance, if $D = 0.5$ at the end of an annual simulation, half of the battery lifetime has been consumed. In other words, the battery needs replacement every two years. Since the battery lifetime is drastically affected by deep discharge cycles, we use only part of the available capacity for daily cycling by introducing the constraint $\text{SOC} > \text{SOC}_{\min}$. The battery parameters are given in Appendix B.

2.2.5 Photovoltaic Model

We adopt a simple PV model for rooftop panels equipped with a Maximum Power Point Tracker (MPPT) from [39, 40]. The model parameters are based on the commercial product *BP 3160 photovoltaic module* and were verified with experimental data. The MPPT power of the array of PV panels is given by

$$P_{\text{PV}} = I_{\text{MPPT}} \cdot U_{\text{MPPT}} \cdot N_{\text{PV}} \ , \tag{2.37}$$

where N_{PV} is the number of panels. The variables I_{MPPT} and U_{MPPT} are the MPPT current and voltage computed with

$$I_{\text{MPPT}} = I_{\text{p}} \cdot \frac{G}{G_{\text{s}}} \cdot [1 + \Delta I_{\text{p}} \cdot (T_{\text{cell}} - T_{\text{s}})] \tag{2.38}$$

$$U_{\text{MPPT}} = U_{\text{p}} \cdot [1 + \Delta U_{\text{p}}(T_{\text{cell}} - T_{\text{s}})] + K_1 \cdot U_{\text{th}} \cdot \ln\left(\frac{G}{G_{\text{s}}}\right)$$
$$+ K_2 \cdot \left[U_{\text{th}} \cdot \ln\left(\frac{G}{G_{\text{s}}}\right)\right]^2 \tag{2.39}$$

$$U_{\text{th}} = \frac{A \cdot k_{\text{b}} \cdot T_{\text{cell}}}{q} \ , \tag{2.40}$$

where G is the solar radiation; I_{p} and U_{p} are the peak power current and voltage, respectively; ΔU_{p} and ΔI_{p} are the temperature effects on peak power voltage and current, respectively; and K_1, K_2 are constant coefficients. The variables G_{s}, T_{s}, A, k_{b} and q are the standard temperature, the standard radiation, the diode ideality factor, the Boltzmann constant and the charge of an electron, respectively. An approximate expression to calculate the cell temperature T_{cell} is [41]

$$T_{\text{cell}} = T_{\text{a}} + G \cdot (T_{\text{NOCT}} - 20) \ , \tag{2.41}$$

where T_{NOCT} denotes the normal operating cell temperature and T_{a} is the ambient temperature. The PV parameters are given in Appendix B.
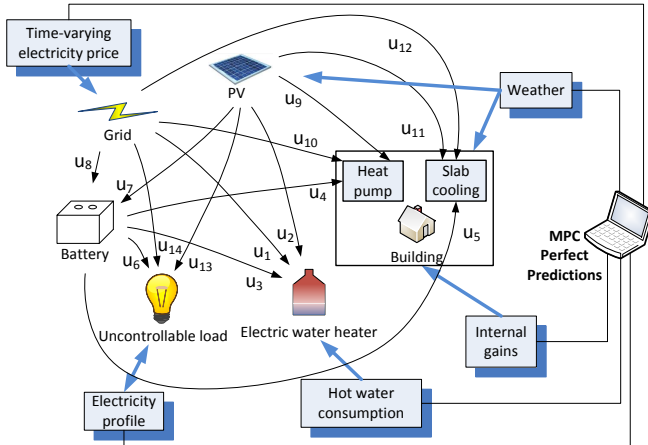
Figure 2.5: Schematic representation of the building system with all components and power flows between them.

2.2.6   Overall Building Model

An overview of the complete building model is given in Fig. 2.5. For optimization purposes one has to distinguish the power sources, because the power for the EWH, HP, SC, battery and Uncontrollable Load (UL) can come from the grid, from the PV, or from the battery. This is modeled by introducing separate control inputs (as shown in Fig. 2.5) and constraining them to sum up to the total power input for each component. The dynamics of the overall system are given by (2.42), where $x_k \in \mathbb{R}^{28}$, $u_k \in \mathbb{R}^{14}$ and $v_k \in \mathbb{R}^5$ are the overall system state, input and disturbance vectors, respectively

$$x_{k+1} = A_k x_k + B_u u_k + B_v v_k \ . \tag{2.42}$$

Observe that there is no arrow going from the PV to the grid in Fig. 2.5. This is because our controller prioritizes PV consumption by the building loads, instead of feeding the PV power to the grid. Of course, if there is excess PV power after supplying the building loads and charging the battery, this excess power will be fed to the grid. Prioritizing the PV self-consumption is an interesting control strategy in systems with low Feed-In Tariffs (FITs) for PV.

A typical Swiss-average residential building with an area of 100 m^2,
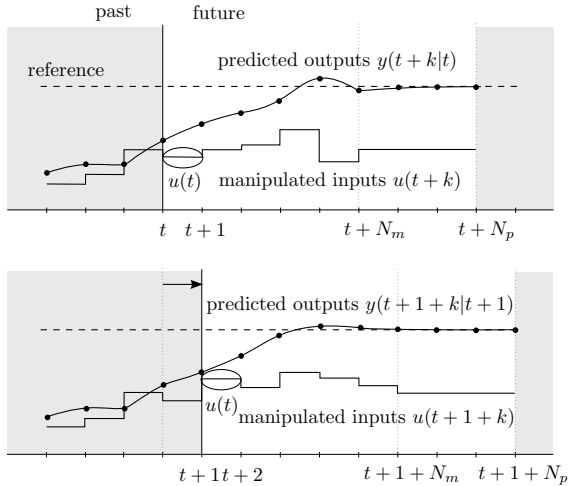
Figure 2.6: Schematic of receding horizon control in MPC [1].

heavy construction, low window area fraction, and south facade orientation is considered for the simulations of this chapter. The building is equipped with an HP with a nominal electric power of 2 kW, an SC installation with a nominal electric power of 1.2 kW and an EWH with a nominal power of 4.1 kW. The building may additionally include 12 PV panels with nominal power 160 W each, and a battery with a capacity of 5 kWh.

2.3 Optimal Building Control

2.3.1 Introduction to MPC

MPC is a predictive controller that uses a model of the system under control to account for the future evolution of the states, when deciding the control input at the current time step. MPC involves solving an optimization problem at each time step, and is executed in a receding horizon fashion. With reference to Fig 2.6, the MPC controller calculates the optimal control sequence across a prediction horizon, but it only applies the first input of the sequence. At the next time step the

$T_{\max}^{\mathrm{c}}$ ─────────────────────────

─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ $T_{\min}^{\mathrm{c}}$

─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ $T_{\max}^{\mathrm{h}}$

$T_{\min}^{\mathrm{h}}$ ─────────────────────────
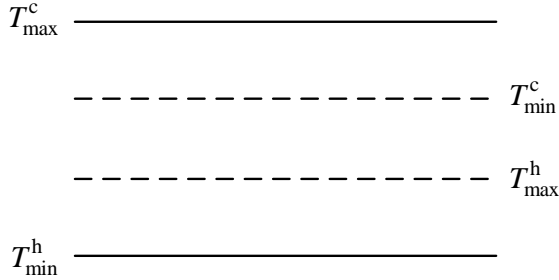
Figure 2.7: The temperature deadband logic considered for RBC. The solid lines correspond to the temperature thresholds to turn on heating or cooling. The dashed lines correspond to the temperature thresholds to turn off heating or cooling.

optimization problem is solved again and the process continues accordingly. The reader is referred to [1, 42] for an introduction to MPC.

2.3.2   Benchmark: Rule-Based Control

MPC is benchmarked against a Rule-Based Control (RBC) approach, which is a standard practice in building control. As the name indicates, RBC determines the control inputs based on rules of the type "if condition then action". Typically, the conditions in RBC correspond to crossings of temperature deadband thresholds.

We consider a temperature deadband with four thresholds for room temperature control, as shown in Fig. 2.7. The variables $T_{\min}^{\mathrm{h}}$ and $T_{\max}^{\mathrm{h}}$ are the lower and upper thresholds for heating operation, whereas $T_{\min}^{\mathrm{c}}$ and $T_{\max}^{\mathrm{c}}$ are the lower and upper thresholds for cooling operation. We consider the simple RBC design described by Algorithm 1.

The RBC parameters are selected as follows: $T_{\min}^{\mathrm{h}} = 21^{\circ}\mathrm{C}$, $T_{\max}^{\mathrm{h}} = 21.5^{\circ}\mathrm{C}$, $T_{\min}^{\mathrm{c}} = 22.5^{\circ}\mathrm{C}$, and $T_{\max}^{\mathrm{c}} = 23^{\circ}\mathrm{C}$. Therefore, two distinct and not overlapping temperature deadbands are considered for heating and cooling to avoid consecutive heating and cooling cycles and reduce unnecessary energy consumption. These thresholds result in an average room temperature of approximately $22^{\circ}\mathrm{C}$ during the whole year, which is equal to the middle of the deadband.

A simple RBC design (Algorithm 2) is used for water temperature control based on temperature measurements T_{w} at the 8th layer of the

---

Algorithm 1 RBC for room temperature control

1: Current heating on/off state ← Previous heating on/off state
2: Current cooling on/off state ← Previous cooling on/off state
3: **if** cooling is on and $T_z \leq T_{min}^c$ **then**
4:      Turn cooling off
5: **else if** cooling is off and $T_z \geq T_{max}^c$ **then**
6:      Turn cooling on
7: **end if**
8: **if** heating is on and $T_z \geq T_{max}^h$ **then**
9:      Turn heating off
10: **else if** heating is off and $T_z \leq T_{min}^h$ **then**
11:      Turn heating on
12: **end if**

Algorithm 2 RBC for water temperature control

1: **if** $T_w \leq T_{min}^w$ **then**
2:      Turn heater on
3: **else if** $T_w \geq T_{max}^w$ **then**
4:      Turn heater off
5: **else if** $T_{min}^w < T_w < T_{max}^w$ **then**
6:      Current heater state ← Previous heater state
7: **end if**

---

water tank (counting from the bottom). The deadband limits are fixed to $T_{min}^w = 50°C$ and $T_{max}^w = 60°C$, such that the average water values during the year are similar to the values obtained by the MPC controller.

2.3.3   MPC Problem Formulation

The goal of MPC is to minimize the electricity cost by using the building model (2.42) and predictions of future disturbances, i.e., weather conditions, occupancy, hot water consumption and electricity prices, while respecting system constraints. In this work perfect predictions of all disturbances are assumed, which provides us with an upper bound of the cost reduction with predictive control.

The following state constraints on room temperature, average water

tank temperature $\bar{x}_k^{\mathrm{wh}}$ and battery SoC are considered

$$21 \le x_{k,1}^{\mathrm{b}} \le 23 \ [^\circ\mathrm{C}] \tag{2.43}$$

$$55 \le \bar{x}_k^{\mathrm{wh}} \le 70 \ [^\circ\mathrm{C}] \tag{2.44}$$

$$0.2 \cdot E_{\max} \le E_{1,k} + E_{2,k} \le E_{\max} = 5 \ [\mathrm{kWh}] \ . \tag{2.45}$$

Referring to the numbering convention of Fig. 2.5, the input constraints can be written as

$$0 \le u_{k,i} \quad \forall\, i \in 1\ldots 14 \tag{2.46}$$

$$0 \le u_{k,1} + u_{k,2} + u_{k,3} \le P_{\mathrm{EWH}}^{\max} = 4.1 \tag{2.47}$$

$$0 \le u_{k,3} + u_{k,4} + u_{k,5} + u_{k,6} \le P_{\mathrm{d},k}^{\max} \tag{2.48}$$

$$0 \le u_{k,7} + u_{k,8} \le P_{\mathrm{c},k}^{\max} \tag{2.49}$$

$$0 \le u_{k,4} + u_{k,9} + u_{k,10} \le P_{\mathrm{HP}}^{\max} = 2 \tag{2.50}$$

$$0 \le u_{k,5} + u_{k,11} + u_{k,12} \le P_{\mathrm{SC}}^{\max} = 1.2 \tag{2.51}$$

$$0 \le u_{k,2} + u_{k,7} + u_{k,9} + u_{k,11} + u_{k,13} \le P_{\mathrm{PV},k} \tag{2.52}$$

$$P_{\mathrm{d},k}^{\max} = \xi_1^{\mathrm{T}} x_{k-1}^{\mathrm{bat}} \tag{2.53}$$

$$P_{\mathrm{c},k}^{\max} = \xi_2^{\mathrm{T}} x_{k-1}^{\mathrm{bat}} \tag{2.54}$$

$$P_{\mathrm{UL},k} = u_{k,6} + u_{k,13} + u_{k,14} \ , \tag{2.55}$$

where $P_{\mathrm{EWH}}^{\max}$ is the maximum EWH power, $P_{\mathrm{HP}}^{\max}$ is the maximum HP power, $P_{\mathrm{SC}}^{\max}$ is the maximum SC power, $P_{\mathrm{PV},k}$ and $P_{\mathrm{UL},k}$ are the available PV power and the uncontrollable load, respectively, both at time step k. All inputs are given in kW. In (2.53) and (2.54) $\xi_1, \xi_2 \in \mathbb{R}^2$ are constant vectors obtained directly from (2.32) and (2.33).

Consider a prediction horizon N and define

$$\mathbf{x} := [x_0^\top, \ldots, x_N^\top]^\top \in \mathbb{R}^{28(N+1)} \tag{2.56}$$

$$\mathbf{u} := [u_0^\top, \ldots, u_{N-1}^\top]^\top \in \mathbb{R}^{14N} \tag{2.57}$$

$$\mathbf{v} := [v_0^\top, \ldots, v_{N-1}^\top]^\top \in \mathbb{R}^{5N} \ . \tag{2.58}$$

With this notation, the MPC problem can be formulated as

$$\min_{\mathbf{u}} \quad w_0 \mathbf{c}^\top \mathbf{u} + w_1 \mathbf{u}^\top \mathbf{Q}\mathbf{u} + w_2 (\mathbf{x} - \mathbf{T}_{\mathrm{ref}})^\top \mathbf{H}(\mathbf{x} - \mathbf{T}_{\mathrm{ref}}) \tag{2.59a}$$

$$\text{s.t.} \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \mathbf{S}\mathbf{u} \le \mathbf{s} \tag{2.59b}$$

$$\mathbf{G}\mathbf{x} \le \mathbf{g} \tag{2.59c}$$

$$\mathbf{F}\mathbf{x} = \mathbf{f} \tag{2.59d}$$

$$\mathbf{x} = \mathbf{A}x_0 + \mathbf{B_u}\mathbf{u} + \mathbf{B_v}\mathbf{v} \ , \tag{2.59e}$$

where $\mathbf{c} \in \mathbb{R}^{14N}$ denotes the electricity cost vector for the whole horizon, w_0, w_1, w_2 are weighting factors, and the matrices $\mathbf{A}, \mathbf{B_u}, \mathbf{B_v}, \mathbf{S}, \mathbf{G}, \mathbf{F}$ and vectors $\mathbf{s}, \mathbf{g}, \mathbf{f}$ are of appropriate sizes. Equations (2.59b) and (2.59c) are the input and state inequality constraints, respectively, (2.59b) is the equality constraints, and (2.59b) represents the building dynamics.

The objective function in (2.59a) consists of three terms: (a) a linear penalty for the electricity cost, (b) a quadratic term penalizing the HP power, and (c) a quadratic term penalizing the deviations of room temperature from a reference value. The entries of matrices $\mathbf{Q}$ and $\mathbf{H}$ are either 0 or 1 such that the quadratic penalties are applied only to the HP inputs (u_4, u_9 and u_{10} in Fig. 2.5), and room temperature deviations from the center of the comfort zone ($22°\mathrm{C}$), respectively.

The three terms of the objective function are given different weights w_0, w_1 and w_2. The quadratic penalty for the HP power is required to linearize the HP dynamics (see Section 2.2.2). By incorporating the penalty for the room temperature deviations, an average annual temperature close to $22°\mathrm{C}$ can be achieved, which is also the average annual temperature with RBC. This ensures that the energy and cost savings achieved by MPC are due to its predictive nature, rather than reduction of user comfort. We tuned the weighting factors in order to achieve a good trade-off among cost minimization, smooth HP operation and reference temperature tracking, and the values $w_0 = 1$, $w_1 = 0.2$ and $w_2 = 0.2$ were chosen, because they performed well in preliminary simulations. A larger w_1 value would penalize the HP power more and result in spreading its operation during the day, thus drastically reducing the potential for load shifting. Similarly, a larger w_2 value would penalize aggressively small temperature deviations from $22°\mathrm{C}$ and limit the potential for load shifting.

We perform the optimization with an hourly time step and a prediction horizon of 16 hours. Preliminary simulations showed that increasing the prediction horizon up to 16 hours reduces the cost drastically, whereas an horizon longer than 16 hours does not reduce the cost significantly. Each optimization problem is solved in 3 seconds on average, using CPLEX through a YALMIP interface [43] in a 4-core machine (2.83 GHz) with 8 GB RAM. Due to the short computation time, a larger optimization problem with shorter time steps would also be tractable.

Note that the power consumption of the EWH, HP and SC are assumed to be continuous variables in (2.59). We made this simplifying assump-

tion to speed up the solution of the optimization problem. In practice, continuous power modulation is possible with the so-called modulating HPs, whereas continuous power control of an EWH can be achieved with voltage control [44]. However, today most of small HPs are not modulating and most EWHs operate in an on/off mode without power modulation. The on/off mode of operation can be integrated into (2.59) by introducing binary variables, which will of course increase the computation time.

Recall that the disturbance predictions are assumed to be perfect. If the building model was also perfect, (2.59) could be applied in an open-loop fashion: at time step k the MPC problem is solved and the optimal control sequence is applied for all time steps within the prediction horizon N without resolving the MPC problem. However, we do not follow this approach because the building model is not perfect due to linearization of the nonlinear HP model.

2.4   Cost Savings and Price Sensitivity

2.4.1   Case Study Description

Building Configurations

Depending on the availability of PVs and batteries, we differentiate among the four building configurations of Table 2.4. Note that IRA is typically used in office buildings because it provides high comfort while being energy efficient. Therefore, the building models developed within the OptiControl project are targeted to office buildings. However, we use a modified model version in this chapter to model a residential building. We simplify the model by neglecting lighting and ventilation, and by fixing the blinds' position to 0.5. In addition, we augment the model's state vector by introducing the HP model of Section 2.2.2. As a result, the heating control input is the HP electric power consumption instead of the heat flux.

Price Signals

The building is controlled using price signals for energy purchases from the grid. Three different price signals are considered in this chapter as summarized in Table 2.5.

Table 2.4: Building configurations

Component	Case I	Case II	Case III	Case IV
PV	-	-	✓	✓
Battery	-	✓	-	✓

Table 2.5: Price signals

	Case A	Case B	Case C
Price signal	day-night	day-ahead dynamic	real-time

First, we consider a simple day-night tariff, which is a common practice today in many countries. More specifically, we use the tariff from *Elektrizitätswerk der Stadt Zürich* (EWZ), which has a day tariff of 0.185 CHF/kWh (from Monday to Saturday, 6h-22h), and a night tariff of 0.095 CHF/kWh (from Monday to Saturday 22h-6h); the night tariff is also used on Sundays and public holidays [8].

Second, we consider a day-ahead time-varying tariff that reflects the marginal electricity costs of the wholesale market. This tariff is based on the Swiss spot market prices of 2009, and can be used to shift demand towards time periods with high RES infeed, and consequently low marginal costs [8].

And third, we consider real-time pricing, namely instantaneous price changes occurring at delivery time. Real-time price signals can be used to control the consumption of an aggregation of buildings in order to track a power reference signal.

In order to prioritize the local PV utilization, we use a zero FIT for the PVs in the MPC problem formulation. However, to achieve reasonable estimates of the annual operational costs, we apply a FIT equal to half of the day tariff for the PVs when post-processing the economic results.

Uncontrollable Building Loads

The electricity consumption of ULs (lighting, televisions, ovens, etc.) is modeled using the data from [45]. The data include power measurements of different loads with a resolution of 1 second from a typical household for the week $21 - 28$ April 2011, which is repetitively used throughout the whole year. While this assumption is reasonable for most uncontrollable loads, this is generally not the case for lighting that

has a strong seasonal dependence. Seasonality effects on uncontrollable loads are neglected in this work due to lack of appropriate data, but they could be easily incorporated.

External Inputs

External inputs to the building include weather data, internal heat gains due to people and equipment, and hot water consumption. The weather data are comprised of the outside dry-bulb air temperature, the wet-bulb temperature and the solar radiation, and are taken from archived forecasts of the COSMO-7 numerical weather prediction model operated by MeteoSwiss. A residential building is considered in this work with an occupancy profile from 6 pm each day till 7 am next day. Typical values for internal gains are taken from the Swiss standard SIA [20]. A probability profile of water draws from [21] is used to simulate daily hot water consumption, which is taken as equal to 280 liters.

2.4.2   MPC Performance

Figures 2.8 and 2.9 depict the optimal building operation for Case IV-B during a typical week in winter and summer, respectively. The HP and SC power consumption is shifted towards low-price intervals, which indicates that the controller exploits the thermal inertia. Similarly, the battery charges from the grid when the price is low, and supplies the load when the price is high. Note that the battery charge power decreases as the SoC approaches 100% due to the charge transfer constraints of the KiBaM model in (2.32). In summer, most of the EWH energy is shifted to low-price intervals using the battery rather than the EWH. In contrast, in winter, the EWH additionally consumes power from the grid during low-price periods and stores this energy in its thermal mass. During the summer week, the PV production is high enough to cover both the controllable and uncontrollable loads, while the excess energy is stored in the battery.

The EWH dominates the electricity consumption in both weeks, whereas the shares of the HP and SC are significantly smaller. The need for cooling is low even during the summer week, because the average ambient temperature is around 20°C. Although the average ambient temperature during the winter week is approximately 1.5°C, the space heating needs are reduced due to the high internal gains and the high COP of
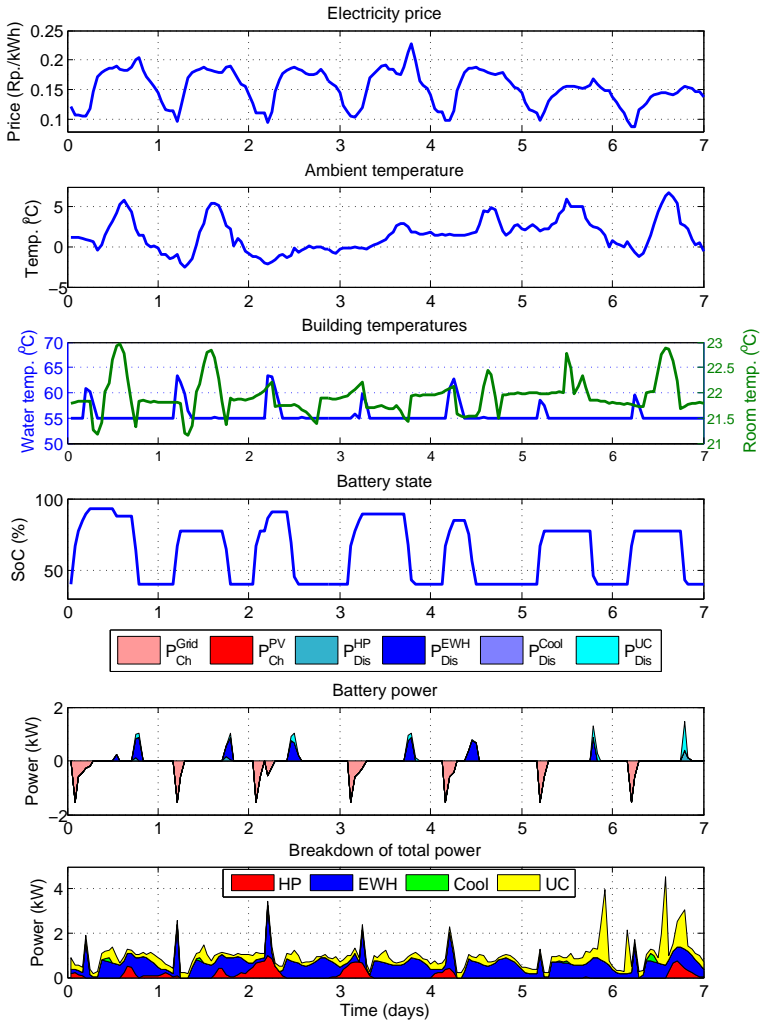
Figure 2.8: Optimal building operation during a typical winter week [Case IV-B]. (a) Electricity price; (b) Ambient temperature; (c) Room and water temperatures; (c) Battery SoC; (d) Battery charge-discharge profile; (e) Power of controllable loads and ULs.
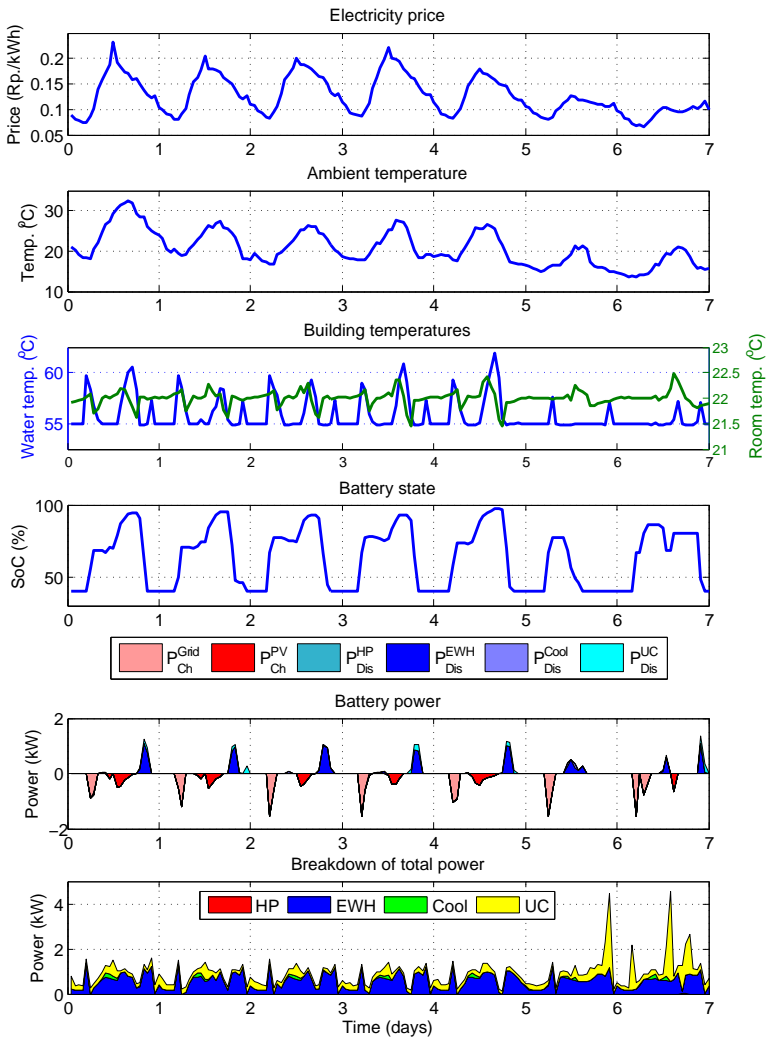
Figure 2.9: Optimal building operation during a typical summer week [Case IV-B]. (a) Electricity price; (b) Ambient temperature; (c) Room and water temperatures; (c) Battery SoC; (d) Battery charge-discharge profile; (e) Power of controllable loads and ULs.

Table 2.6: Room temperature comparison between RBC and MPC

Controller	Min Temp.	Max Temp.	Mean Temp.
RBC	20.55°C	23.00°C	21.97°C
MPC	21.08°C	23.00°C	21.95°C

Table 2.7: Water temperature comparison between RBC and MPC

Controller	Min Temp.	Max Temp.	Mean Temp.
RBC	50.19°C	59.97°C	55.57°C
MPC	54.88°C	62.11°C	55.47°C

the HP (the COP is 4.2 on average). On the other hand, the EWH consumes a lot of energy due to the high thermostat setpoints $(55 - 70°C)$ and the large hot water demand.

2.4.3 Cost Savings Potential

In this section, we assess the potential of MPC to minimize electricity costs in buildings compared with RBC. We perform annual simulations with the four building configurations of Table 2.4, and for the price cases A and B. Tables 2.8 and 2.9 compare RBC and MPC in terms of annual electricity costs and energy imports from the grid.[3]

As mentioned in Section 2.3.2, the temperature thresholds of RBC were selected with two criteria in mind: (a) excessive cycling is avoided; and (b) the average room and water temperatures during the year are roughly the same for RBC and MPC. Indeed, the resulting average temperatures are nearly identical for RBC and MPC according to Tables 2.6 and 2.7. Therefore, the comparison between MPC and RBC is fair, i.e., any cost reduction achieved by MPC can be attributed to the prediction of disturbances and prices rather than on compromising user comfort.

---

[3]Similar simulations were performed in [20], but the savings calculated in this thesis are generally smaller than those reported in [20]. There are two reasons for this. First, a seasonal setting was applied in [20] where heating was only allowed during winter months and cooling only during summer months. In contrast, this seasonal setting was not used in the simulations of this thesis. Second, in order to avoid consecutive heating and cooling cycles, a more advanced RBC that consists of four threshold temperatures (see Section 2.3.2) was implemented in this thesis but not in [20].

Table 2.8: Comparison between RBC and MPC for case A

Benchmark (RBC)				
	Electricity cost		Energy consumption	
	[CHF]	[%]	[kWh]	[%]
RBC: I-A, II-A	1349.90	-	8379.80	-
RBC: III-A, IV-A	1058.10	-	7088.90	-
Performance bound (MPC)				
	Electricity cost		Energy consumption	
	[CHF]	[%]	[kWh]	[%]
I-A	1180.96	−12.52	8182.04	−2.36
II-A	1138.90	−15.63	8468.37	+1.06
III-A	833.95	−21.18	6140.32	−13.38
IV-A	797.42	−24.64	6371.58	−10.12

Table 2.9: Comparison between RBC and MPC for case B

Benchmark (RBC)				
	Electricity cost		Energy consumption	
	[CHF]	[%]	[kWh]	[%]
RBC: I-B, II-B	1299.80	-	8379.80	-
RBC: III-B, IV-B	1021.60	-	7088.90	-
Performance bound (MPC)				
	Electricity cost		Energy consumption	
	[CHF]	[%]	[kWh]	[%]
I-B	1190.54	−8.41	8216.27	−1.95
II-B	1142.34	−12.11	8473.00	+1.11
III-B	874.33	−14.42	6160.60	−13.10
IV-B	835.95	−18.17	6364.87	−10.21

In case I-A, the MPC reduces the cost by approximately 12% and the energy consumption by approximately 2% compared with the RBC. By adding a battery to the system the additional savings are approximately 3%. However, the cost reduction comes with an increase in energy consumption due to battery efficiency losses (note that in case II-A, the MPC results in a consumption 1% higher than that of RBC, yet it still reduces the cost).

In a building with PVs the benefits of using MPC are pronounced. In case III.A energy imports from the grid and costs decrease by approx-
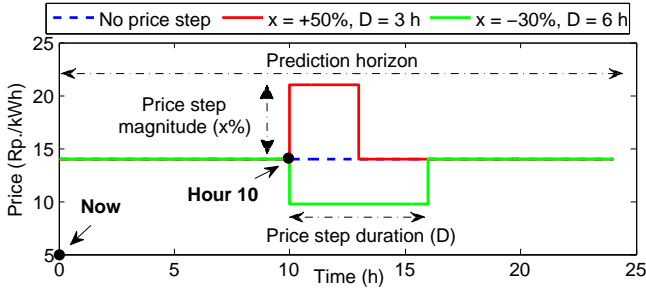
Figure 2.10: Explanation of the price step used in the sensitivity analysis for the day-ahead price signals.

imately 21% and 13%, respectively. Therefore, the building thermal inertia can be efficiently used to increase local PV utilization in buildings. If a battery is used with a PV installation, the battery achieves additional savings of around 3%. Qualitatively similar results are obtained for the day-ahead dynamic tariff (case B), but the savings are smaller in absolute terms.[4]

Based on these results, we carry out a simple economic evaluation of a battery installation for electricity cost reduction. Assuming a capital cost of 200 CHF/kWh for lead-acid batteries [46] and constant savings for every year equal to 48.2 CHF (which is the best case according to Table 2.9), the investment's payback period is approximately 20 years. Since the typical battery lifetime is $5 - 15$ years [46], using batteries for this application is not economically viable with the current electricity price levels, battery efficiencies and capital costs. According to our results, a battery price less than 145 CHF/kWh is necessary to render the investment in battery profitable.

2.4.4 Sensitivity to Day-Ahead Price Signals

We carry out a sensitivity analysis to investigate the building's response to day-ahead price signals in winter. For this purpose, we apply a constant price for every hour of the next day, with the exception of

---

[4]The values reported in this chapter are upper bounds of the potential savings for two main reasons. First, the disturbance predictions are assumed to be perfect in the MPC. And second, the gap between the MPC and the RBC can be, in principle, reduced by appropriately tuning the RBC, e.g., based on experience.
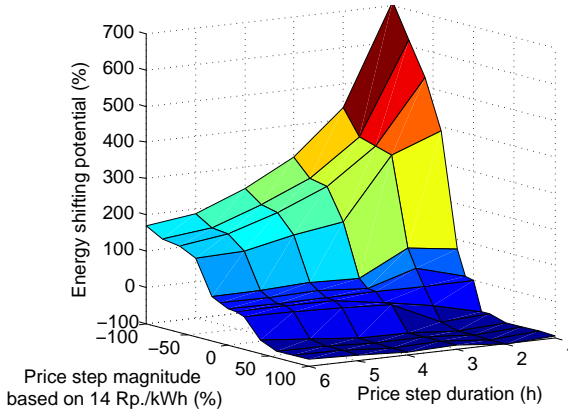
Figure 2.11: Energy shifting potential as a function of the magnitude and
duration of the day-ahead price steps (Case IV).

hour 10 when we apply a price step of variable magnitude and duration, as shown in Fig. 2.10. We choose hour 10 as the starting time of the price step because both the HP and the EWH normally operate during this hour. We present the results of the sensitivity analysis for Case IV in Fig. 2.11, where the energy shifting potential is defined as the consumption during the price step divided by the consumption without any price step.

Percentage-wise the building is more responsive to negative than positive price steps, in particular if the duration of the negative price step is short. Interestingly, most of the energy shifting potential can be achieved by a price step duration of up to 2 hours. For a -100% price step with a duration of 1 hour the energy consumption increases by a factor of 7, whereas for a $+100\%$ price step with a duration of 1 hour the energy consumption decreases by approximately 93%. This means that the building is very sensitive to price signals that are known day-ahead. For price steps with a magnitude less than 30%, the energy is shifted using only the thermal inertia of the building and EWH. For larger price steps, the battery is also used because the round-trip efficiency is around 74% (including converter losses).
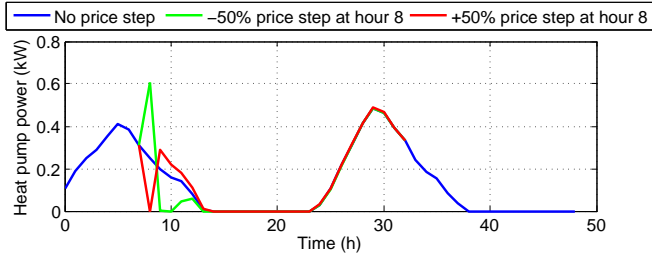
Figure 2.12: HP Response to -50% and $+50\%$ real-time price signals.
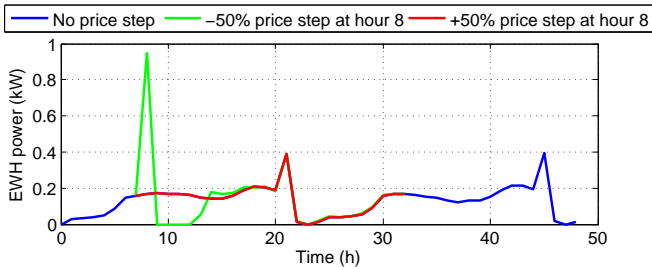
Figure 2.13: EWH Response to -50% and $+50\%$ real-time price signals.

2.4.5 Sensitivity to Real-Time Price Signals

A real-time price signal is a signal sent to the building at the time when the energy delivery takes place, i.e., it is not known a priori like in Section 2.4.4. In this section, we assume that a building uses MPC to optimize its operation based on a day-ahead price profile. However, an additional price signal may be superimposed on the day-ahead price profile during physical energy delivery.

Figure 2.12 and Fig. 2.13 depict the HP and EWH response, respectively, to a -50% and a $+50\%$ price signal occurring at hour 8. Both the HP and the EWH are able to respond to negative price signals by instantly increasing their consumption and shifting energy from the following hours. However, the behavior is different for positive price signals. The HP switches off during hour 8 and compensates by increasing its consumption in the following 5 hours, whereas the EWH cannot modify its consumption at all. Although these results are only for hour 8, a similar behavior was observed for any other hour when both the
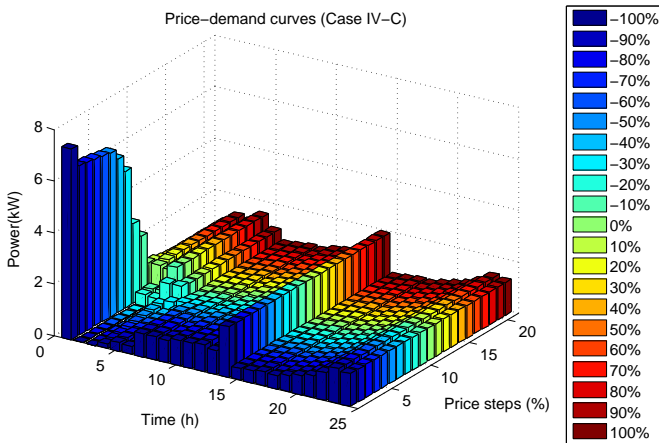
Figure 2.14: Building demand-price curves for real-time price signals at
hour 8 (Case IV-C). The curves show the sensitivity of de-
mand on the magnitude of the price signal.

HP and the EWH operate. Of course, positive price signals occurring
at times when the consumption is zero, as well as negative price signals
occurring at times when the consumption is maximum, will not affect
the building behavior.

There are two reasons why the the EWH cannot respond to positive
real-time price signals. First, the EWH dynamics are faster than the
HP dynamics and the EWH thermal inertia is lower. Second, recall that
the deviation of EWH temperature from the setpoint is not penalized in
the objective function (2.59a), as it is done for the room temperature.
Therefore, the EWH temperature remains close to the lower deadband
limit for most of the day in order to minimize energy consumption, and
so there is practically no slack for further power reduction. Note that
the HP potential for power reduction depends heavily on the weighting
factor w_2 in (2.59a): the better the tracking of 22°C (staying in the
middle of the comfort zone), the higher the potential.

By repeating the same analysis for several price steps in the range
$[-100\%, 100\%]$ and by aggregating the responses of all building com-
ponents, we derive the demand-price curves of Fig. 2.14. This figure
shows how much the building consumption of the following hours will
change if a real-time price signal is sent at hour 8. For the building

under consideration, real-time price signals modify consumption up to 6 hours in the future. Again, it is obvious that the effect of negative price signals on the consumption is much higher than that of positive price signals. Demand-price curves of different buildings can be used by load aggregators to incorporate the economic value of electricity for end-users in power markets, as in [47], or to minimize balancing energy costs due to RES infeed prediction errors.

2.5 Conclusion

In this chapter, we introduced a benchmark residential building model with an HP, an EWH, a battery and PV panels. We used the model within an MPC scheme to optimize the building's consumption under dynamic end-customer electricity prices. Simulations showed that the MPC outperforms a conventional RBC and results in cost reductions in the range $8 - 25\%$, depending on the building configuration and electricity price profile.

Furthermore, we investigated the response of a building controlled with an MPC to day-ahead and real-time step price signals. The results showed that the building consumption is much more sensitive to negative price signals (decreased price) than positive price signals (increased price).

Chapter 3

PV Self-Consumption Maximization in Buildings

3.1 Introduction

Large shares of Photovoltaics (PVs) have been introduced in the power system over the last years, and projections show that the worldwide installed PV capacity will continue to increase [48]. A large portion of the PV power worldwide is installed on the roofs of residential and commercial buildings [49]. Depending on financial incentives, locally produced PV energy can be either self-consumed in the building premises or fed into the grid. Maximizing self-consumption of PV energy is preferable from a technical point of view for the following reasons [50]: (a) fewer overvoltages occur, (b) violations of cable and transformer thermal limits are less likely, and (c) the losses in the distribution network are reduced.

The technical advantages of PV self-consumption are understood by power system regulators in some countries, which have adopted (or are discussing) regulatory measures to motivate local utilization of PV energy in buildings. For example, in Germany the Feed-In Tariff (FIT) for PVs has fallen below the residential customer electricity tariff [51]. This

Table 3.1: Nomenclature of Chapter 3: symbols

Symbol	Unit	Description
$b_1 - b_4, \gamma$	-	PV model parameters
C_{bat}	€/kWh	Battery capital cost
C_{el}	cents/kWh	Electricity cost
C_{FIT}	cents/kWh	PV feed-in tariff
CI	€/year	Annual cash inflow
CO	€/year	Annual cash outflow
d	%	Discount rate
E	Wh	Electric energy
G	W/m^2	Solar radiation
N	years	Investment horizon
P	W	Power
R	€/year	Annual cash flow
r	%/year	Annual electricity price increase
Q_{bat}	kWh	Battery energy capacity
S_{HP}	-	On/off state of the HP
T	°C	Temperature
t	-	Time index
η	-	PV efficiency
ξ	-	PV self-consumption ratio
α	-	PV excess energy parameter

Table 3.2: Nomenclature of Chapter 3: subscripts and superscripts

Subscript / superscript	Description
a	Ambient temperature
b	Building temperature
c	PV cell temperature
ch	Battery charge power
dis	Battery discharge power
el	Electric power
exp	Energy export
imp	Energy import
l	Uncontrollable load
min/max	Minimum/maximum value of a variable
n	Net power
s,i	The ith layer of the water storage tank
k	Year index
th	Thermal power

creates an interest for self-consumption of PV electricity even in the absence of additional regulatory measures, such as the previously existing bonus on self-consumed PV electricity in Germany [52]. However, further investigation is needed to verify whether this interest translates into a business case.

The financial advantages of PV self-consumption must be taken into account in the planning phase of rooftop PV installations. For this

purpose, algorithms to maximize local PV utilization should be incorporated into software for energy simulations in buildings. In this context, this chapter presents four control algorithms to maximize PV self-consumption with controllable thermal loads and batteries. To illustrate their practical applicability, the algorithms were integrated into Polysun, a commercial software for energy simulations in buildings [53, 54], which was extended with new features.

Optimal operation strategies for residential and commercial buildings with PV installations and battery storage have attracted the interest of many researchers. Some papers, for example [19, 20, 55, 56], investigated planning and operation strategies for residential buildings with flexible thermal loads and/or battery storage to reduce electricity costs, but without explicitly addressing the problem of PV self-consumption. Other papers, such as [57, 58], studied the potential of PV self-consumption increase using stationary lithium-ion batteries and/or electric vehicles. More relevant to this chapter is the work presented in [59], where control of event-based loads (e.g., washing machines and dishwashers) and storage are compared in terms of potential for PV self-consumption optimization in residential buildings.

The contribution of this chapter is threefold: (a) we develop four simple rule-based control algorithms for batteries and flexible thermal loads and integrate them in Polysun; (b) we evaluate the potential for PV self-consumption maximization and estimate the respective cost savings by performing annual energy simulations; and (c) we investigate the effect of uncertain parameters, such as the battery costs and FITs, on the optimal building configuration. In this chapter, the thermal load is a Heat Pump (HP) and the goal is to shift its consumption towards intervals with large PV production.

This chapter is based on the results published in [60] and is organized as follows. Section 3.2 starts with a brief overview of Polysun and then introduces the models used to simulate the PV, HP, hydronic system, and battery. Section 3.3 presents the developed control algorithms for PV self-consumption maximization. Finally, Section 3.4 discusses simulation results from a case study, which are used to estimate the PV self-consumption potential and the respective savings. The chapter's nomenclature is shown in Tables 3.1 and 3.2.

3.2   Building Modeling in Polysun

Polysun is a software tool developed by the Swiss company Vela Solaris AG for thermal simulations in buildings [53,54]. It includes a large number of building components and templates, and enables annual energy simulations with realistic weather and heat demand data. In addition to thermal simulations, Polysun also supports simulations of rooftop PV systems. In this section, we present the mathematical models used in Polysun to simulate the PV, HP and hydronic system.

3.2.1   Photovoltaic Model

PV modules are modeled using the general purpose model proposed in [61]. The model assumes that the PV is equipped with a Maximum Power Point Tracker (MPPT), and expresses the module efficiency η as a function of solar radiation G and cell temperature T_c

$$\eta(G, T_c = 25^oC) = b_1 + b_2 G + b_3 \ln G \tag{3.1}$$

$$\eta(G, T_c) = \eta(G, 25^oC)[1 + b_4(T_c - 25)] \ , \tag{3.2}$$

where b_1, b_2, b_3 are fitting parameters, and b_4 is the module temperature coefficient. The cell temperature can be calculated based on the ambient temperature T_a according to

$$T_c = T_a + \frac{\gamma G}{1000} \ , \tag{3.3}$$

where γ is a parameter related to the rear ventilation of the PV module. The necessary data to identify the four model parameters using linear fitting techniques are:

- Three values of PV module efficiency at different radiation conditions.

- The PV module efficiency at standard test conditions of 1000 W/m^2 radiation and 25°C cell temperature, as well as the efficiency at 1000 W/m^2 but at a different temperature.

Due to its simplicity, the model can be applied for a variety of PV cell technologies including crystalline silicon (cSi), amorphous silicon (a-Si) and copper-indium-diselenite (CIS) cells. The inverter efficiency is modeled based on measured values at 100%, 50% and 10% partial load. Efficiencies for other load levels are calculated with linear interpolation between these values.

3.2.2 Building and Hydronic System Model

The dynamic thermal response of a building is highly dependent on the topology of pipes and pumps, on the size of water storage tank, and on the controller's settings. An adequate representation of the hydronic system is necessary to accurately predict the seasonal performance factor of a HP [62], and reference [63] showed that Polysun algorithms are suitable for this purpose.

The numerical model of Polysun implements the plug-flow approach [64] in which fluid elements such as pipes, heat exchangers, and storage tanks realize the mass transport at every time step. In order to calculate the thermal energy demand, a single-node building model is used with an empirical heat transfer coefficient for the heating device (e.g., radiator or floor heating). Adequate modeling of the building envelope is critical, because it is a key thermal storage element in addition to the water storage tank. The hot water storage tank is divided into eleven water layers with different temperatures in order to capture the thermal stratification effects. Numerical weather data from the Meteonorm database [65] are used as external input to calculate the thermal energy demand of the building.

3.2.3 Heat Pump Model

Air-to-water HPs for heating and cooling are considered as flexible thermal loads. The HP model is based on the thermal and electric power consumption at different temperatures at the evaporator side (T_{evap}) and at the air side (T_{air}). These values are typically measured on predefined sampling points according to the test standards EN 255 or EN 14511 (e.g., the notation A2/W35 indicates measurement at an air intake temperature of 2°C and a heating water outlet temperature of 35°C). Additional sampling points can be used as input values in order to improve the accuracy and application range of the simulation model. A similar approach could be applied for water-water or brine-water HPs.

In the time-domain simulation algorithm, the operating point of the HP is evaluated at each calculation step. The electric and thermal HP power, P_{el} and P_{th} respectively, are evaluated by interpolation between the sampling points. These two quantities are related with the Coeffi-

cient of Performance (COP) according to

$$\text{COP} = \frac{P_{\text{th}}}{P_{\text{e}}} \ . \tag{3.4}$$

3.2.4 Battery Model

The battery model developed in Section 2.2.4 was integrated in Polysun and used for the simulations of this chapter.

3.3 Rule-Based Control Design

3.3.1 Problem Statement

We define the PV self-consumption ratio as the ratio between the PV energy consumed in the building premises $E_{\text{PV}\to\text{bld}}$, i.e., before the Point of Common Coupling (PCC), and the total PV energy yield E_{PV}. We assume the meter configuration of Fig. 3.1, but a single bi-directional meter could also be used instead of the two one-directional meters M1 and M2. Let us denote by E_{exp} the energy exported to the grid and recorded by meter M1, and by E_{imp} the energy imported from the grid and recorded by meter M2. With this notation, the PV self-consumption ratio can be defined as

$$\xi = \frac{E_{\text{PV}\to\text{bld}}}{E_{\text{PV}}} = \frac{E_{\text{PV}} - E_{\text{exp}}}{E_{\text{PV}}} \ . \tag{3.5}$$

To maximize ξ, the export to the grid E_{exp} should thus be minimized.

We develop four rule-based control algorithms for the HP and the battery in order to maximize ξ. In principle, the MPC approach of Chapter 2 could be used for the same purpose. However, the goal of this chapter is to model the thermal distribution system in detail (in addition to modeling the building envelope), and using an MPC with such a model would be computationally intensive. Instead, we attempt to investigate if simple rule-based control algorithms can increase the PV self-consumption ratio significantly.

The key idea of the algorithms is that PV energy should be utilized preferably within the building premises, instead of being fed to the grid. For this reason, the HP operation is shifted to intervals when the PV
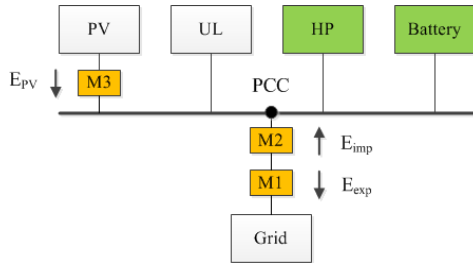
Figure 3.1: The layout of the system and the metering topology. The controllable components and electricity meters (M1-M3) are highlighted with dark color. The grid is shown only to explain the metering topology, but no grid constraints are considered in this chapter.

production is high. The PV energy is stored as heat in the hot water storage tank and can be utilized later depending on the building thermal energy demand. Similarly, any PV energy surpluses can be stored in the battery provided that it is not fully charged. During hours with low or zero PV production, the battery discharges to cover the load in order to reduce the electricity import from the grid.

Therefore, two storage options are available in the building: thermal storage and battery. Algorithm 1 (A1) assumes that only the thermal storage is present. Algorithm 2 (A2) considers the battery but does not use the thermal storage for PV self-consumption. Algorithms 3 and 4 consider both storage options, but employ different control priorities with respect to PV energy surpluses. Algorithm 3 (A3) prioritizes the HP, whereas Algorithm 4 (A4) stores energy first in the battery. The choice of the priority depends on various efficiencies, the thermal energy demand of the building, the electricity consumption profile, and it might exhibit seasonal patterns. Since it is generally difficult to identify a priori the appropriate priority, annual simulations can be used for this purpose.

3.3.2 Conventional Control

The proposed algorithms are compared with a conventional control algorithm (A0), where there is no battery and the HP is operated based on its internal controller. The operation principle of this controller is
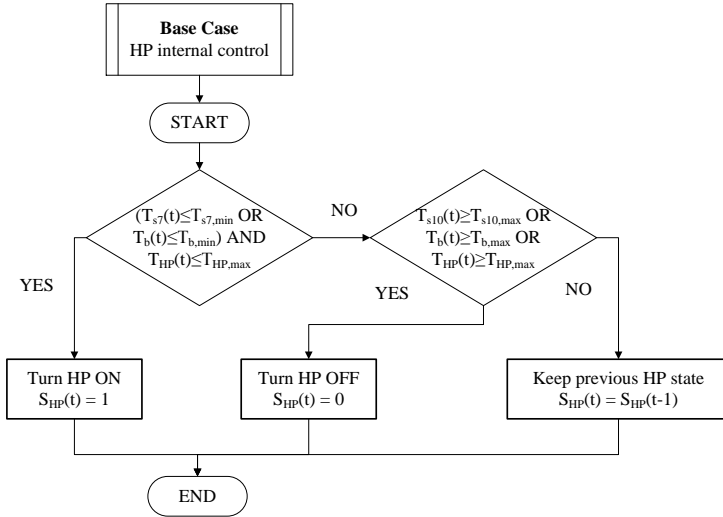
Figure 3.2: Flowchart of the conventional control algorithm without consideration of PV self-consumption (algorithm A0).

shown in Fig. 3.2. The water temperatures at the 7th and 10th layer of the storage tank ($T_{\text{s},7}$ and $T_{\text{s},10}$), as well as the building temperature T_{b}, are used as feedback for the controller, and the respective setpoints are denoted by $T_{\text{s},7,\min}$, $T_{\text{s},10,\max}$, $T_{\text{b},\min}$, and $T_{\text{b},\max}$. The internal temperature of the HP (T_{HP}) is also used as a feedback variable to allow switching off for security reasons whenever a limit $T_{\text{HP},\max}$ is exceeded. We denote the HP state by S_{HP}, and the current time step by t. According to Fig. 3.2, the HP state changes when any of the aforementioned setpoints is reached. In any other case, the HP state remains the same as in the previous time step. Note that the internal controller can also handle minimum operation time constraints; however, these are not shown in Fig. 3.2 for the sake of simplicity.

3.3.3   Heat Pump-Only Smart Control

The control logic of algorithm A1 for smart HP control is shown in Fig. 3.3. Let us denote the available PV power by P_{PV} and the total uncontrollable load by P_{l}. Similarly to algorithm A0, whenever any of the setpoints $T_{\text{s},7,\min}$, $T_{\text{s},10,\max}$, $T_{\text{b},\min}$ and $T_{\text{b},\max}$ is reached, the HP
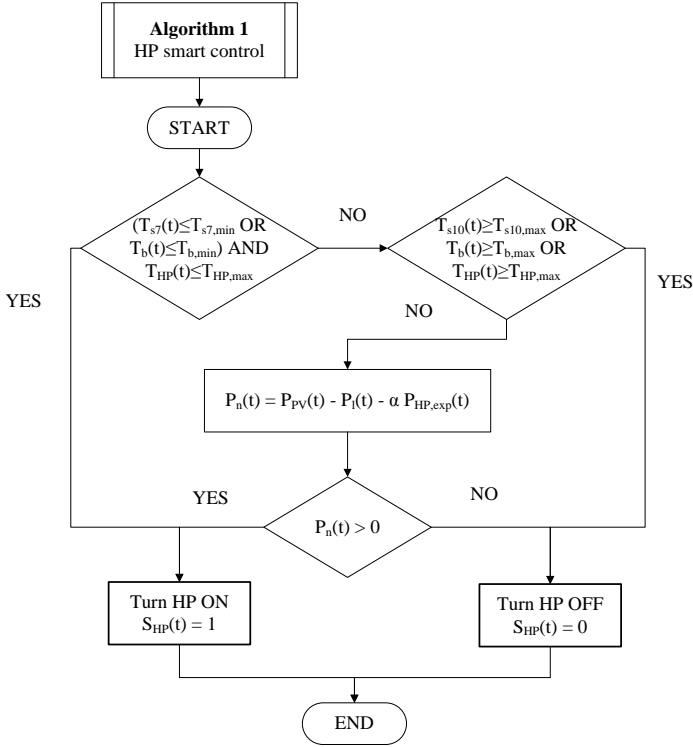
Figure 3.3: Flowchart of the HP-only smart control algorithm for PV self-consumption maximization (algorithm A1).

turns on or off based on its internal controller. In any other case, the HP state is determined based on the net power P_n, which represents the power surplus in the building. In Fig. 3.3, $P_{HP,exp}$ denotes the power that the HP is expected to consume if it turns on at the current time step. The parameter $\alpha \in [0, 1]$ determines which part of $P_{HP,exp}$ must be covered by P_{PV} to allow HP operation (higher α values will result in a less frequent HP operation).

3.3.4 Battery-Only Smart Control

Figure 3.4 presents the system operation under algorithm A2. The HP is not utilized for PV energy management in this case, but it operates
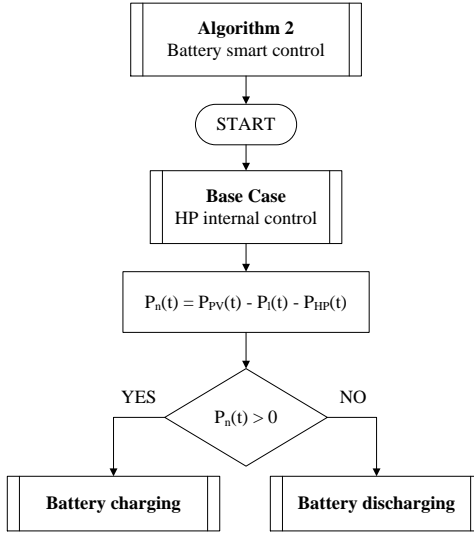
Figure 3.4: Flowchart of the battery-only smart control algorithm for PV
self-consumption maximization (algorithm A2). The battery
charging and discharging algorithms are shown in Fig. 3.5.

based on algorithm A0 and is treated as an uncontrollable load in the
calculation of P_n. The battery control during charging and discharging is shown in Fig. 3.5. If a power surplus occurs and the battery is
not fully charged, the charging mode is enabled. On the other hand, if
a power deficit occurs and the battery is not fully discharged, the discharging mode is enabled. If the discharging or charging limit is reached,
any additional power is imported from or fed to the grid, respectively.
Finally, the battery idles if the State of Charge (SoC) is 100% or below
SOC_{min}.

3.3.5   Heat Pump-Priority Smart Control

Algorithm 3 is described in Fig. 3.6. Since the HP has priority over the
battery, the first part of algorithm A3 is identical to algorithm A1. After
the HP state is fixed, the battery enters the charging or discharging
mode depending on the sign of P_n. The same logic as in algorithm A2
applies here for the control of battery charging and discharging.
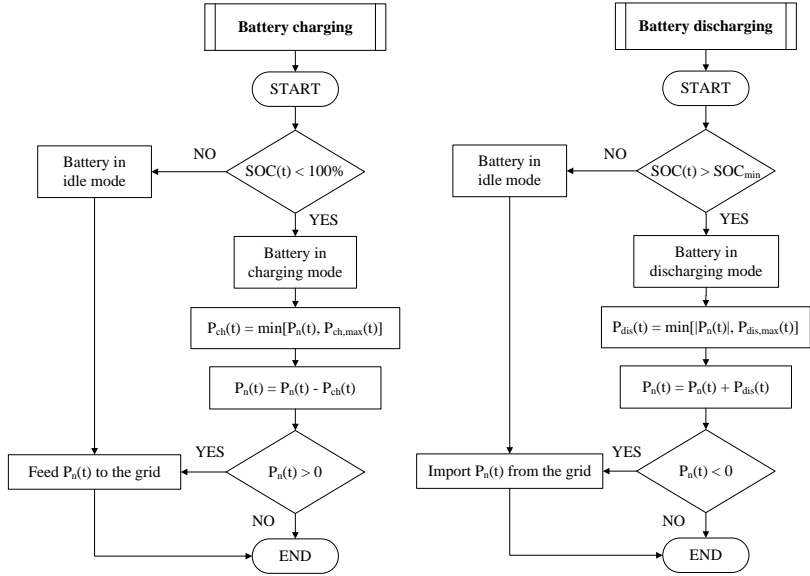
Figure 3.5: Flowchart of the battery charging and discharging algorithms, which are used within algorithms A2-A4.

3.3.6 Battery-Priority Smart Control

Algorithm 4 is the most involved of the developed algorithms and is shown in Fig. 3.7. The discharging phase is identical to algorithm A3 but the charging priority is reversed, i.e., in the presence of a power surplus the battery is charged first. If the charging power limit is reached or the battery is fully charged, the remaining power is consumed by the HP (if it is more than the expected power consumption of the HP; otherwise, it is fed to the grid).

3.3.7 Control Implementation into Polysun

We implement algorithms A0-A4 using logic statements and incorporate them into Polysun using a component called programmable controller. This component allows us to program logical statements in a high-level user-friendly language. In addition, specific operation intervals for the thermal loads can be defined using the same component.
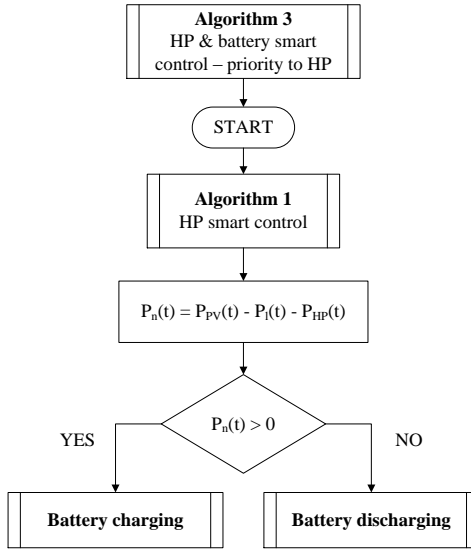
Figure 3.6: Flowchart of the smart control algorithm for PV self-consumption maximization that considers both the HP and the battery, and gives priority to the HP (algorithm A3). The battery charging and discharging algorithms are shown in Fig. 3.5.

3.4   Potential for PV Self-Consumption

3.4.1   Case Study Description

We show the algorithms' performance considering a single-family residential building (2 parents both working and 1 child) as a case study. We design the system in Polysun with the parameters and assumptions summarized below:

- The HP is rated at 10.1 kW thermal power and 3.3 kW electric power at W2/W35 operating conditions.

- The capacity of the hot water storage tank is equal to 500 liters.

- The rooftop PV system consists of 60 panels rated at 180 W each, which leads to a total installed power of 10.8 kWp.
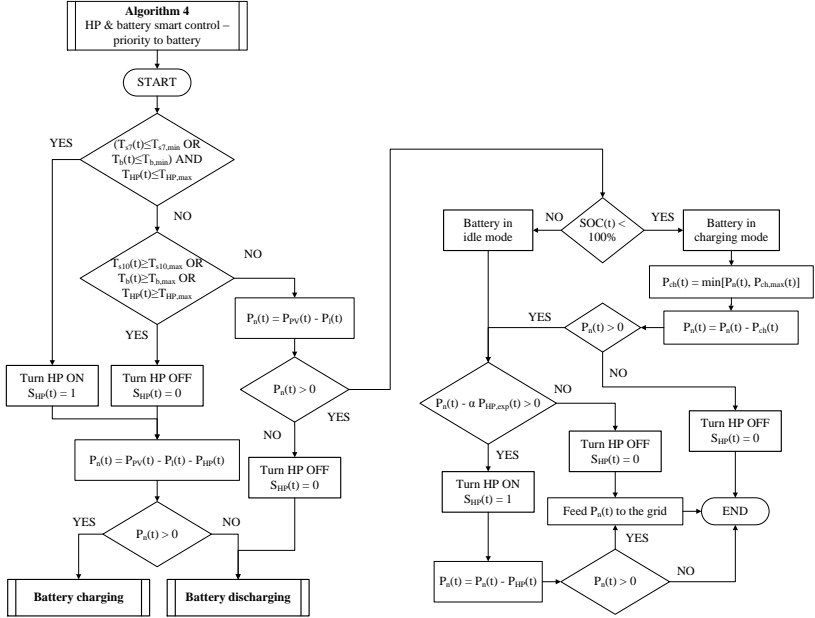
Figure 3.7: Flowchart of the smart control algorithm A4 for PV self-consumption maximization that considers both the HP and the battery, and gives priority to the battery. The battery charging and discharging algorithms are shown in Fig. 3.5.

- The battery is a Hoppecke 24 OpzS 3000 module with a nominal cell voltage of 2 V and a total capacity of 6 kWh. The values of the battery lifetime parameters are taken from [22] and are given in Appendix B.

- The hot water temperature setpoint is fixed to 50°C, and the average hot water demand is assumed equal to 200 liters/day.

- The electric Uncontrollable Load (UL) is modeled adopting a profile with an annual consumption of 3103 kWh. The profile was generated using a load profile generator, which simulates the occupants' behavior based on a desire model, and includes typical operation patterns for more than 100 electrical devices [66, 67]. After individual load curves are generated for each device using the desire model, the total UL profile is calculated by adding up the energy use of each device at each point in time.
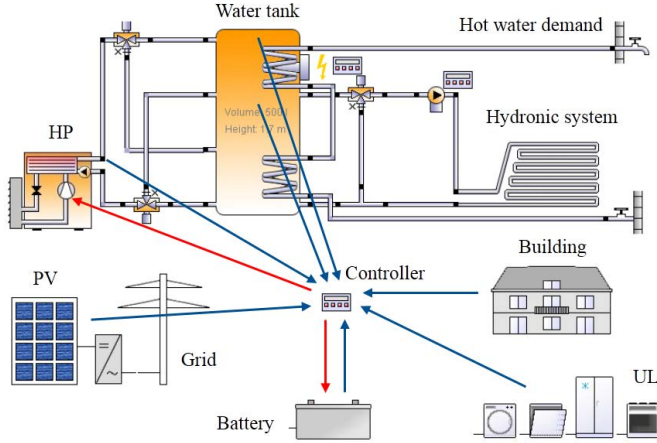
Figure 3.8: The building diagram in Polysun.

Table 3.3: Controller parameters for algorithms A0-A4

Parameter	Symbol	Value
Tank lower temperature setpoint	$T_{s,7,min}$	45°C
Tank upper temperature setpoint	$T_{s,10,max}$	70°C
Building lower temperature setpoint	$T_{b,min}$	21°C
Building upper temperature setpoint	$T_{b,max}$	25°C
HP maximum temperature	$T_{HP,max}$	65°C
Minimum battery charge	SOC_{min}	30%
PV excess energy parameter	α	1

The building diagram in Polysun with all components and the details of the hydronic system is shown in Fig. 3.8. The programmable controller is in the center of the diagram; its inputs are indicated with blue arrows, whereas its outputs with red arrows. Note that apart from the programmable controller there are two more components with the same icon, which are controllers for the internal electric heating element of the water tank and the position of the two-way mixing valves. We keep separate controllers for these devices, because they are not used for PV self-consumption maximization. The chosen control parameters for algorithms A0-A4 are shown in Table 3.3.
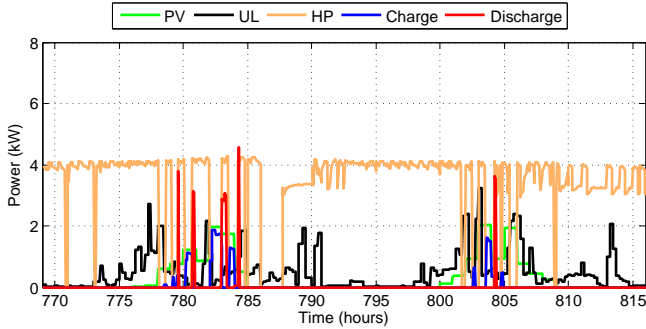
Figure 3.9: Production and consumption of system components during two typical days in winter (January).
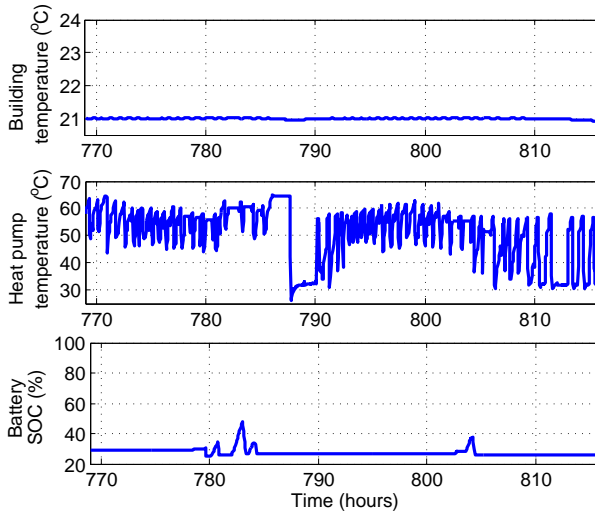
Figure 3.10: Building temperature, HP temperature, and battery SoC during two typical days in winter (January).

3.4.2   Seasonal Operation Results

In this section, we present operation results from typical days in winter, spring, summer and autumn for algorithm A3 (control with priority to HP), which demonstrates the best performance over the year. We
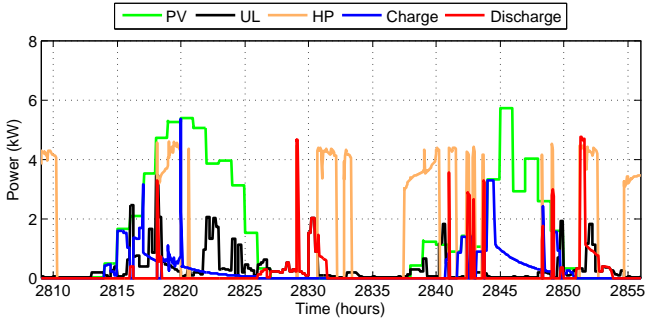
Figure 3.11: Production and consumption of system components during two typical days in spring (April).
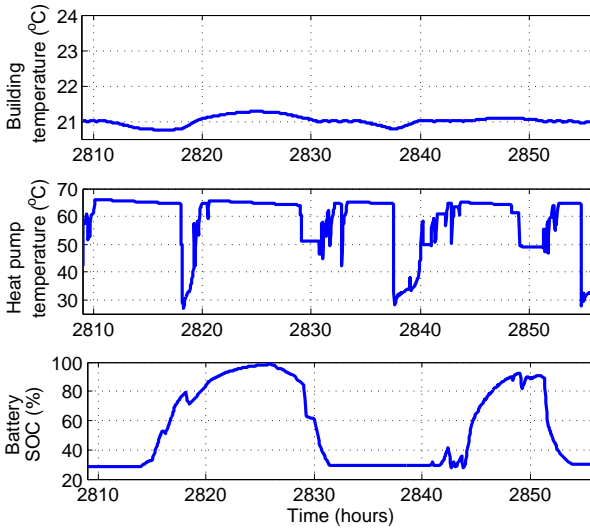
Figure 3.12: Building temperature, HP temperature, and battery SoC during two typical days in spring (April).

focus on PV production, UL and HP consumption, battery charge and discharge energy, building temperature, battery SoC, and HP internal temperature, and show simulation results in Figs. 3.9-3.16.

The seasonal patterns of HP operation can be clearly seen in the figures. The HP runs most of the time during a cold winter day, whereas it
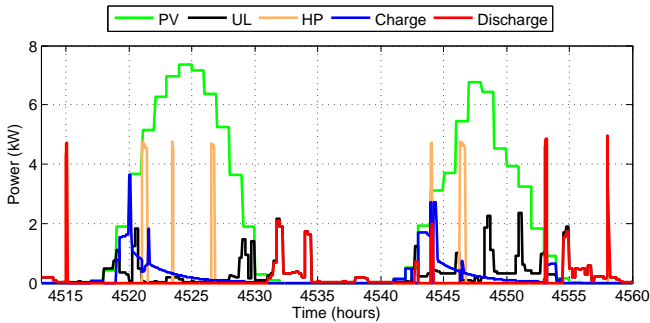
Figure 3.13: Production and consumption of system components during two typical days in summer (July).
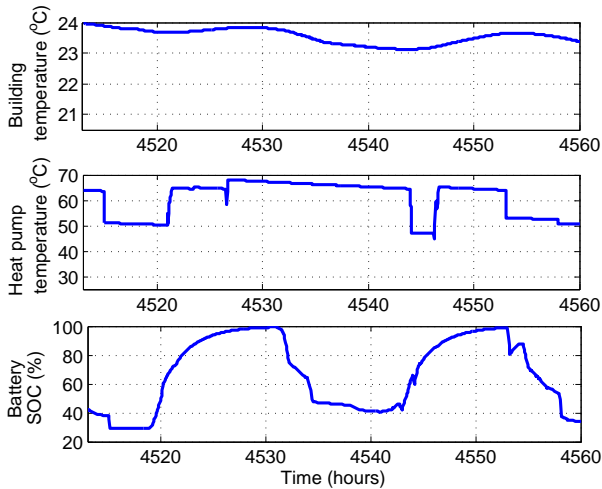
Figure 3.14: Building temperature, HP temperature, and battery SoC during two typical days in summer (July).

turns on only a few times during a summer day. In winter, when PV production is low, the HP operation is determined by the building heat demand, i.e., the HP consumes the minimum amount of energy that is required to keep the building temperature at 21°C. Thus, the potential for PV energy management via HP control is very limited.

In summer, the PV production is high; however, only a small part of
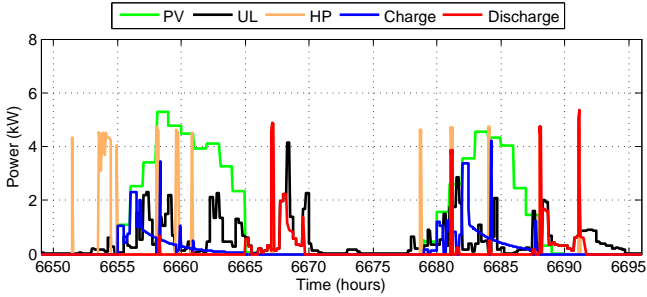
Figure 3.15: Production and consumption of system components during two
typical days in fall (October).
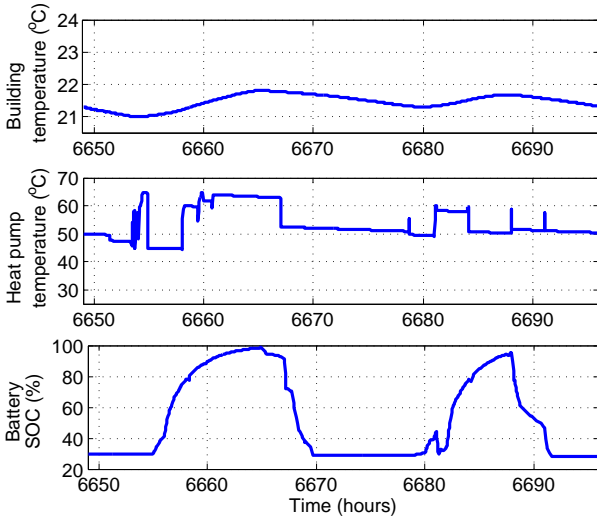
Figure 3.16: Building temperature, HP temperature, and battery SOC dur-
ing two typical days in fall (October).

the PV energy surplus is stored as thermal energy. This is because the
prolonged HP operation leads to overheating due to the low building
heat demand. This can be seen in Fig. 3.14, where the HP temperature
remains often at or above $T_{\mathrm{HP,max}}$. This is in contrast to Fig. 3.10,
where the HP temperature in winter is significantly lower on average.

Our simulations indicate that the potential for PV self-consumption

maximization via HP control is higher in spring, and to a lesser extent in autumn. For example, the HP can effectively absorb the excess PV power during the first day of Fig. 3.11, because high PV production coincides with increased heat demand and low HP temperatures (see Fig. 3.12).

If PV power surpluses exist during daytime, the battery exhibits the expected pattern: it charges during daytime and discharges in the evening and night hours to cover the load. Note that the charging cycles might be interrupted by smaller discharging cycles, if the HP turns on and the PV power is not sufficient. Due to low PV production in winter, the battery might stay at low SoC (even below SOC_{min}) for prolonged periods of time.

Recall that algorithm A3 prioritizes the HP over the battery, as shown in Fig. 3.15. During the interval $6655 - 6658$ hours, PV power is available but $P_{PV} < P_{HP} + P_{UL}$; therefore, the HP is off and the excess power charges the battery. Afterwards, the PV power increases and the UL demand decreases, so that the algorithm follows the predefined priority and turns on the HP.

3.4.3 Annual Economic Results

This section presents economic results based on annual simulations using Polysun. Note that in these simulations the battery capacity is not optimized, but this will be investigated in Section 3.4.4. Table 3.4 compares the algorithms A0-A4 with respect to the energy imports from the grid, the energy exports to the grid, the HP and total electricity consumption, the battery charging and discharging energy, the PV self-consumption ratio, and the resulting electricity costs.

Using only the HP and the building thermal inertia in algorithm A1 increases the PV self-consumption by approximately 1.5% compared with algorithm A0. On the other hand, the battery alone (algorithm A2) demonstrates a much higher potential, and leads to a PV self-consumption increase of roughly 15.5%. When the HP and the battery are controlled simultaneously, a marginally higher self-consumption ratio is achieved when the priority during charging is given to the HP. Algorithm A3 leads to the highest PV self-consumption ratio, which is equal to 36.46%.

Algorithm A3 gives the best results also in terms of savings for the building owner, which are approximately 85 € per year. Although increasing

Table 3.4: Annual simulation results for algorithms A0-A4

	A0	A1	A2	A3	A4
Import (MWh)	8.83	8.76	7.73	7.64	7.70
Export (MWh)	8.15	8.00	6.58	6.45	6.55
HP consum. (MWh)	7.73	7.80	7.76	7.80	7.77
Total consum. (MWh)	10.83	10.90	10.86	10.90	10.87
Bat. charge (MWh)	−	−	1.66	1.66	1.60
Bat. discharge (MWh)	−	−	1.22	1.22	1.18
Self-consumption (%)	19.71	21.14	35.19	36.46	35.49
Electricity bill (€)	1248.6	1249.2	1169.3	1163.0	1163.4

the PV self-consumption usually reduces the electricity bill, this is not the case for algorithm A1. The reason is that A1 operates the system in a less energy efficient way, increases the thermal losses, and eventually requires more electric energy imports for the HP.

Clearly, the battery alone achieves most of the potential for PV self-consumption maximization. HP control further reduces the annual cost only by 6 €. However, this reduction comes likely at virtually zero cost, since the marginal cost of integrating the HP in the already installed controller is expected to be negligible. Similarly to algorithm A1, shifting the HP demand in algorithm A3 increases the thermal losses and the annual HP consumption. Nevertheless, these additional losses are covered by battery discharge energy, instead of energy imports from the grid. This explains why algorithm A3 achieves a lower cost than algorithm A2, whereas algorithm A1 increases the cost compared with algorithm A0.

Intuitively, the potential for PV self-consumption maximization with smart HP control depends on the volume of the hot water storage tank. We investigate this dependence by running annual simulations with two larger tanks (1000 l and 1500 l) and compare the results with the base-case tank of 500 l. In these simulations we consider a system without battery and use the control algorithms A0 and A1. Furthermore, we keep the PV installed capacity and HP nominal power constant, because they are in practice determined by the available roof area and building thermal energy needs, respectively.

Table 3.5 summarizes the simulation results for the tanks of 1000 l and 1500 l with respect to PV self-consumption ratio, electricity costs,

Table 3.5: Effect of hot water tank volume on PV self-consumption

	1000 liters		1500 liters	
	A0	A1	A0	A1
HP consumption (MWh)	8.15	8.27	8.87	8.82
Self-consumption (%)	20.03	22.46	20.83	28.96
Electricity bill (€)	1356.4	1355.5	1537.9	1416.9

and HP electrical consumption. Results for a tank volume of 500 l are presented in Table 3.4. The system behaves in a similar way for water tank volumes of 500 l and 1000 l. However, with a tank of 1500 l, algorithm A1 clearly outperforms the base-case algorithm A0, in terms of both PV self-consumption and electricity costs. In this case, the building thermal inertia is high enough to allow efficient shifting of the HP consumption without increasing the energy losses. Note that increasing the tank volume leads to a higher HP consumption, because more energy is needed to heat up a larger water volume at the same temperature setpoint.

Our analysis indicates that in the presence of a large water tank, HP control is efficient for PV self-consumption maximization. If the tank volume is below a threshold, then the HP alone does not improve the local PV utilization significantly. In such cases, additional control actions might be necessary, for example, the pumps and mixing valves in Fig. 3.8 could be also controlled in parallel to the HP to improve performance. Doing so is likely to avoid high HP temperatures that prohibit HP operation when PV energy surpluses are available, as explained in Section 3.4.2.

3.4.4   Investment Evaluation

In this section, we perform annual simulations using algorithm A3 with six different battery capacities ($Q_{bat} = \{0.6, 1.6, 3, 6, 12, 18\}$ kWh), while keeping the storage tank volume constant at 500 liters. In this case, we also assess the battery investment over the PV installation's lifetime via parametric analysis with respect to the battery capacity, the battery capital cost, and the FITs.

The FIT for residential PV installations up to 40 kW was 13.54 cents/kWh in October 2013 [51]. Once the PV installation is completed, the building owner will receive the FIT for every kWh exported to the grid for 20

years [68]. For electricity costs we use the EON tariff for end-customers in Stuttgart, which is $C_{el,0} = 26.64$ cents/kWh (almost twice as much as the FIT). According to [52], the average residential electricity price in Germany increased from 17.14 cents/kWh in 2004 to 26.79 cents/kWh in 2013. This corresponds to an average annual increase of r = 4.5%, and it was considered for the analysis of this chapter.

The capital cost of lead-acid batteries has been decreasing recently as the technology is getting more mature. Reference [46] reported a capital cost range of $150 - 300$ €/kWh for lead-acid batteries in 2009. The battery type Hoppecke OpzS could be purchased in 2013 at a price range of $60 - 200$ €/kWh, depending on the battery capacity and the retailer. Based on this information, we consider the ten different battery costs $C_{bat} = \{60, 80, 100, 120, 140, 160, 180, 200, 220, 240\}$ €/kWh. The annual operation and maintenance cost is assumed equal to 2% of the battery capital cost [22]. The controller cost is hard to estimate, but it is expected to be significantly lower than the cost of the battery due to its simplicity, therefore it is neglected here. The investment is assessed considering a period of $N = 20$ years, which is a typical value for PV lifetime, assuming a discount rate of $d = 8\%$.

We use the annual simulation results along with the economic parameters to assess the battery investment using the Net Present Value (NPV) as a performance indicator. With the exception of the battery, we assume that the other building components exhibit the same behavior for each year of the investment period. This means that the occupant consumption patterns do not change, the PV module's performance does not degrade over time, and the HP, hydronic system, and storage tank operate reliably without maintenance or replacement costs. Since this assumption is not valid for the battery, we calculate the expected battery lifetime using the lifetime model of Section 2.2.4, as well as the number of replacements needed within N years. We discount the costs associated with the battery replacements using the discount rate d, and include them in the NPV calculation. However, the same battery efficiency is assumed throughout the whole investment period.

The FIT of the German renewable energy act (EEG) for residential PVs was decreasing in 2013, with a monthly degression dependent on the amount of new PV capacity connected to the network [52]. For this reason, it is hard to estimate the evolution of the FIT even in the near future. To investigate the effect of this uncertainty on investment decisions, we repeat our analysis for three different FITs, namely $C_{FIT} =$
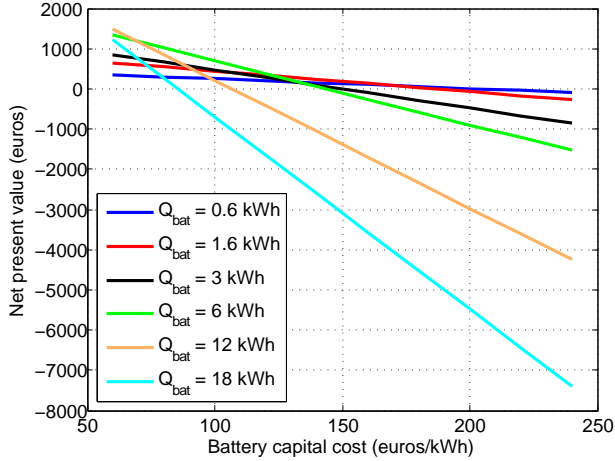
Figure 3.17: The dependence of investment NPV on battery capacity and capital cost for a FIT of 13.54 cents/kWh.

Table 3.6: Dependence of battery lifetime on battery capacity

Battery capacity (kWh)	0.6	1.6	3	6	12	18
Battery lifetime (years)	2.8	3.6	4.3	4.5	4.5	4.8

$\{13.54, 12.00, 10.63\}$ cents/kWh. The first value corresponds to the FIT in October 2013, whereas the other two tariffs correspond to expected tariffs in October 2014 assuming monthly degressions of 1%, and 2%, respectively.

Based on the above, the investment NPV can be calculated with

$$\text{NPV} = \sum_{k=0}^{N} \frac{R_k}{(1+d)^k} \tag{3.6}$$

$$R_k = CI_k - CO_k, \quad \forall k \geq 1 \tag{3.7}$$

$$CI_k = BC - |C_{\text{FIT}} \cdot E_{\exp,k} - C_{\text{el},k} \cdot E_{\text{imp},k}|, \quad \forall k \geq 1 \tag{3.8}$$

$$CO_k = 0.02 \cdot C_{\text{bat}} \cdot Q_{\text{bat}} + m \cdot C_{\text{bat}} \cdot Q_{\text{bat}}, \quad \forall k \geq 1 \tag{3.9}$$

$$C_{\text{el},k} = C_{\text{el},0} \cdot (1+r)^k, \quad \forall k \geq 1 \tag{3.10}$$

$$R_0 = C_{\text{bat}} \quad, \tag{3.11}$$

where R_k is the net cash flow in year k, CI_k and CO_k are the cash
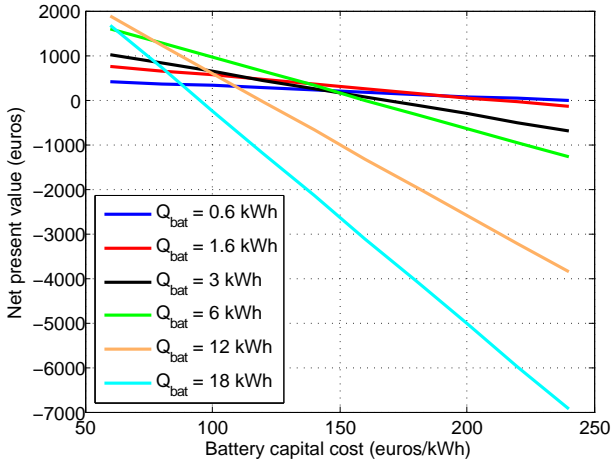
Figure 3.18: The dependence of investment NPV on battery capacity and capital cost for a FIT of 12.00 cents/kWh.

inflows and outflows, respectively, and BC is the annual electricity cost for the base case. The interpretation of (3.8) is that the cash inflows are the annual savings resulting from the battery investment. The cash outflows include the annual battery operation and maintenance cost, and the battery replacement cost. Note that the replacement cost occurs only when the battery reaches the end of its lifetime ($m = 1$). Equation (3.10) expresses the annual increase in electricity cost and (3.11) represents the fact that the net flow at the beginning of the investment is equal to the battery capital cost.

The calculated lifetimes for each Q_{bat} are reported in Table 3.6. In general, increasing the battery capacity leads to longer lifetimes because the battery operates at a higher SoC on average, and the number of deep discharge cycles is reduced.

The parametric analysis results are shown in Figs. 3.17-3.19, for a FIT of 13.54 cents/kWh, 12.00 cents/kWh, and 10.63 cents/kWh, respectively. Based on these results, a number of interesting observations can be made. If $C_{bat} = 60$ €/kWh, the capacity $Q_{bat} = 12$ kWh results in the highest NPV for all FITs (in excess of 2000 €). On the other hand, if $C_{bat} = 240$ €/kWh, the capacity $Q_{bat} = 0.6$ kWh performs the best but the investment is not viable. In all cases, small capacities lead to NPVs that are less sensitive to C_{bat}. On the contrary, large capacities
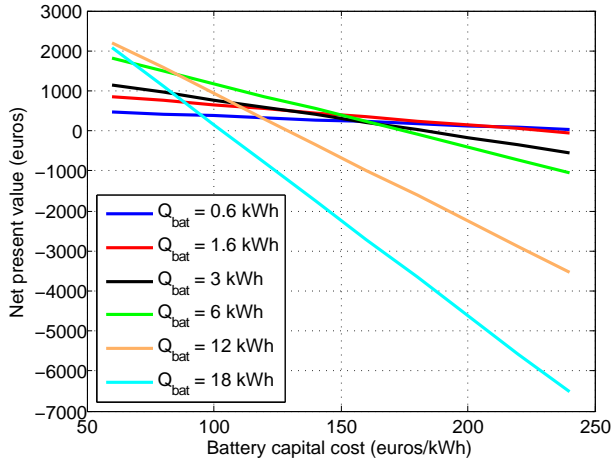
Figure 3.19: The dependence of investment NPV on battery capacity and capital cost for a FIT of 10.63 cents/kWh.

result in a higher NPV for low battery costs, but increasing the battery cost quickly leads to a negative NPV. Note that in Fig. 3.19, a 0.6 kWh capacity achieves a positive NPV for all costs in the considered range. In addition, the capacity $Q_{bat} = 18$ kWh is a dominated solution for all FITs, because there exists always another capacity that achieves a higher NPV for any C_{bat}.

Although the battery capital cost is known at the time of investment, the annual cash flows will be affected by the uncertainty related to battery costs due to the necessary replacements. Depending on the expected evolution of battery costs and the investor's attitude against risk, larger or smaller capacities can be selected based on Figs. 3.17-3.19.

As a final remark, a battery may further reduce the costs in the presence of time-varying retail electricity prices. In this case, a predictive control approach (as the one of Chapter 2) is more appropriate, but the control design becomes more demanding.

3.5 Conclusion

In this chapter, we investigated the potential for PV self-consumption maximization in buildings using batteries and flexible thermal loads.

For this purpose, we developed four rule-based control algorithms and integrated them in the commercial software Polysun. We illustrated the algorithms' performance considering a residential building with a HP and lead-acid batteries as a case study.

We found that controlling the HP alone has a small potential for PV self-consumption maximization, especially for small water tanks. In addition, the HP-only control might increase the annual electricity costs due to higher thermal losses. These results are qualitatively different to the ones of Chapter 2 and highlight the importance of using detailed load models. Engaging HPs in PV energy management applications might also require control of other building components, such as mechanical pumps and mixing valves.

On the other hand, batteries demonstrate a good potential for PV self-consumption maximization and cost minimization. Combining thermal with battery storage achieves the best results, because the synergies between them are efficiently exploited. Our results showed that installing batteries for local PV utilization is becoming an attractive investment due to the decreasing trends in battery costs and FITs.

Part II

Demand Response for Ancillary Services

Chapter 4

Demand-Side Ancillary Services in Power Systems

4.1 Introduction

With increasing shares of fluctuating Renewable Energy Sources (RES), more Ancillary Services (AS) are required in power systems [5]. AS are those services that are necessary to support the transmission of energy from the production resources to the loads in a reliable way. Although different Transmission Systems Operators (TSOs) might procure different types of AS, the following AS are generally common [69]:

- Frequency control (active power reserves)

- Voltage control (reactive power reserves)

- Black start

- Compensation of active power losses

Frequency control reserves are the focus of this thesis. Although frequency reserves come mainly from generators, flexible loads can in principle also provide reserves. The idea of using flexible loads to balance the power grid, which is known as Demand Response (DR), was already proposed in the 1980's [6]. Nowadays, DR has become a very active field of research firstly, because the need for AS is increasing and secondly, due to recent advancements in communication technologies. Loads with thermal storage are particularly suitable for AS, because their consumption can be shifted in time without impact on consumer comfort.

It is common in the literature to differentiate between (i) large aggregations of small residential Thermostatically Controlled Loads (TCLs) such as refrigerators, air conditioners, space and water heaters, and (ii) small aggregations of large Heating, Ventilation and Air-Conditioning (HVAC) systems of commercial buildings. Typically, TCLs are easier to control but the main challenge is to achieve coordinated control of thousands of devices without excessive communication. Commercial buildings are well-suited for reserve provision because they have a large thermal inertia, and they typically include a Building Automation System (BAS), i.e., some communication and control infrastructure is already in place. However, the challenge is that the HVAC systems are typically complex with many control variables and cascaded control loops.

In the second part of this thesis, we develop methods to allow aggregations of residential and commercial thermal loads to provide frequency reserves. The goal of this chapter is to recap relevant aspects of frequency control in power systems and summarize the existing literature.

4.2   Frequency Control in Power Systems

Typically, a TSO controls the frequency in three steps: primary, secondary and tertiary control. Primary control is a decentralized, proportional controller that stabilizes the frequency after a disturbance. Secondary control is a centralized controller that restores the frequency to its nominal value, and maintains the desired exchanges between neighboring control areas. Tertiary control releases secondary control in case of prolonged disturbances and is typically manually activated. This sequential activation of frequency control reserves is shown in Fig. 4.1.
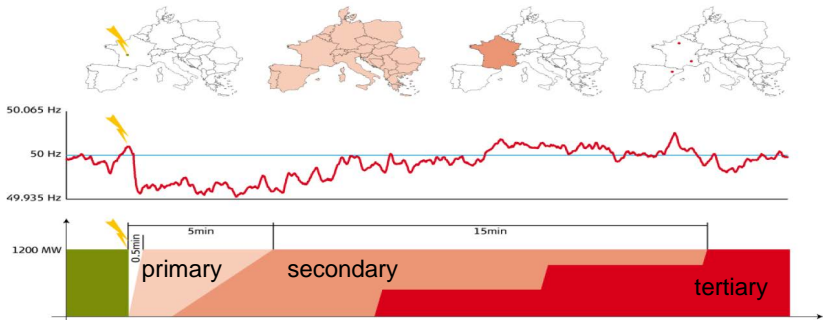
Figure 4.1: The sequential activation of primary, secondary and tertiary frequency control reserves (source: www.swissgrid.ch).

In this thesis, we focus on primary and secondary frequency control because they are automatic. In the following sections, we summarize some important aspects of scheduling and activation of Primary Frequency Control (PFC) and Secondary Frequency Control (SFC) in power systems [70].

4.2.1 Primary Frequency Control

In some countries provision of PFC reserves is mandatory for specific generators, whereas in other countries there exist organized markets for PFC. In the first case the reserve price is typically regulated, whereas in the second case the providers bid their PFC reserve capacity in pay-as-bid or marginal pricing auctions. Minimum resolution limitations of the bid PFC reserve capacity exist in some countries, and they are typically less than 1 MW. In addition, the PFC reserves are mostly symmetrical (i.e., the up- and down-reserve capacities must be equal) with a few exceptions, for example Denmark and the Netherlands.

An important aspect of PFC is that it must be activated very fast, for example, within 30 seconds in Switzerland (see Fig. 4.1). Furthermore, PFC is decentralized since each generator responds proportionally to the local frequency measurements according to its droop characteristic [71]. PFC is critical for power system stability, and in order to keep the control robust, it does not rely on communication. Due to the proportional control and the lack of communication, PFC results in a steady-state frequency error after a disturbance, and it is the task of

SFC to compensate for this error.

Theoretically, the PFC action is finalized after approximately 5 minutes because SFC has taken over (see Fig. 4.1). Therefore, provision of PFC reserves involves little energy delivery, and remuneration is typically awarded only for the PFC capacity (not for the energy). Typically, the TSOs have prequalification tests to assess the suitability of a reserve provider for PFC.

Different names are sometimes used to refer to PFC, for example, governor control action, droop control, frequency response, and frequency containment reserve (FCR). Throughout this thesis we use the name PFC.

4.2.2   Secondary Frequency Control

In most of Europe, SFC reserves are procured in a market setting, i.e., the generators bid their reserve capacity and price in weekly or daily auctions. The requirements of the auctions and SFC reserve products vary among countries. The minimum bid size is typically in the range 1 to 10 MW, e.g., 5 MW in Switzerland [72]. However, there exist some markets with low minimum bid sizes, for example, the Pennsylvania, Jersey, and Maryland Power Market (PJM) with a minimum bid size of 100 kW [73]. In many countries only symmetric reserves are allowed (equal up- and down-reserve capacities), whereas in other countries asymmetric reserves are also accepted.

The reserve is requested from the generators via a signal sent by the TSO, typically every $2-4$ seconds. There are two main activation rules: (a) the pro-rata activation, where the reserve request is proportional to the capacity; and (b) the merit-order activation, where the reserve is requested based on the short-term marginal costs of the generators. In this thesis, we treat the future reserve requests as uncertain, because they are unknown at the time when the reserve capacities are scheduled.

It is a standard practice today to have a single SFC signal in a control area, which is typically the output of a Proportional-Integral (PI) controller with the Area Control Error (ACE) as an input [71]. However, a TSO could alternatively provide a number of reserve products and each provider could choose the product to offer its reserve capacity. This can be achieved by decomposing the original signal into multiple SFC signals with different energy contents and ramping rates using filtering or optimization-based techniques [74–78]. Such an approach is already

applied in the PJM control area, where the SFC signal is split into a slow (RegA) and a fast (RegD) component [73].

Typically, the provision of SFC reserves is remunerated separately for the accepted reserve capacity in the auction (availability) and the requested reserve energy by the TSO (utilization). In some countries the reserve capacity auctions are pay-as-bid, whereas in others the auctions have a common clearing price. The remuneration of reserve energy is also country dependent; for example, in Switzerland the reserve energy remuneration is coupled with the energy price in the spot market. In PJM, the reserve provider is additionally remunerated based on the tracking performance of the SFC signal [79].

Different names are used to refer to SFC, for example, Automatic Generation Control (AGC), load frequency control (LFC), automatic frequency restoration reserve (FRR), or frequency regulation service (mainly in the USA). In this thesis we adopt the names SFC and frequency regulation, which we use interchangeably.

4.2.3   Specific Frequency Reserve Products

Many AS markets have specific frequency reserve products. In this section, we briefly describe some products commonly used in AS markets in the USA [69]. However, a complete review of frequency reserve products is outside the scope of this thesis.

The AS for frequency support can be divided in two groups. The first group contains AS for normal conditions, such as regulating reserve and load following. The second group consists of AS for contingency conditions, such as spinning and non-spinning reserve, and replacement reserve.

Regulating Reserve (Frequency Regulation)

As already mentioned, frequency regulation is the term commonly used in the USA to refer to SFC. Frequency regulation is automatically provided by resources that are online (i.e., connected to the system) and its goal is to correct for the minute-to-minute deviations between the total generation and the load, under normal conditions.

Load Following

This AS is similar to frequency regulation, but it has a slower ramp rate and fewer sign changes per unit of time. The goal of load following is to bridge the gap between the hourly energy markets and frequency regulation, also under normal conditions. Another difference with frequency regulation is that load following can be manually activated.

In systems with load following, the frequency regulation signal is usually uncorrelated and crosses zero often. In contrast, if there is only one SFC reserve product that combines frequency regulation and load following, then the SFC signal can be biased to the positive or negative direction for a significant amount of time.

Spinning Reserve

Spinning reserve is an AS supplied by generators that are online and not fully loaded. The resource starts providing the spinning reserve immediately after the contingency, the full reserve must be deployed within 10 minutes, and the resource must be able to sustain the reserve for two hours. Spinning reserve is typically activated automatically using an AGC signal.

Non-Spinning Reserve

In contrast to spinning reserve, the non-spinning reserve is provided by generators that are not necessarily online and synchronized to the grid. However, the spinning and non-spinning reserve have the same response requirements: the full reserve must be provided within 10 minutes after the contingency, and the resource must be able to sustain the reserve for two hours. Typically, non-spinning reserves are dispatched using an AGC signal.

Replacement (Supplemental) Reserve

This contingency reserve must be fully deployed within 30 or 60 minutes, and can last for two to four hours. Both online and offline resources can provide this reserve, and AGC is not necessary for activation. The replacement reserve is a type of tertiary control reserve.

4.3 Demand-Side Primary Frequency Control

Reference [6] proposed already in 1980 a simple frequency-responsive controller to adjust the temperature limits of TCLs based on frequency measurements to provide a governor-type action. Much of the early work on DR focused on modeling populations of TCLs based on first principles [80–83]. For example, [82] proposed a well-known model consisting of a set of Fokker-Planck equations that describe the Probability Distribution Function (PDF) of temperature in a TCL aggregation. In the rest of this section, we summarize the most relevant approaches from the literature to provide PFC with TCL aggregations.

4.3.1 Deterministic Decentralized Approaches

References [84, 85] developed deterministic approaches that rely on a proportional frequency-dependent temperature deadband of the thermostat to provide PFC. The authors found that load control reduces the frequency excursions after a sudden loss of generation, and smoothes the short-term frequency fluctuations. Reference [86] introduced a more generic decentralized approach for load participation in PFC, which is based not only on the frequency deviation but also on its evolution over time. A simple rule-based controller with delays was developed in [87] and used in an experimental demonstration to provide PFC with residential appliances using a commercially available smart meter for frequency measurements, a smart load controller, and smart sockets.

Such deterministic approaches provide an effective initial response to frequency deviations. For this reason, the European Network of Transmission System Operators for Electricity (ENTSO-E) published a demand connection code that suggests a proportional shift of the setpoint temperatures of thermal loads in response to frequency deviations. Despite their simplicity, deterministic approaches have an important limitation: they cause rebound effects and tend to synchronize the on/off cycles of individual devices, which might introduce non-decreasing oscillations in system frequency [88,89]. Furthermore, [90] showed a main disadvantage of the ENTSO-E demand connection code: it allows for various controller implementations that are compatible with the code, but result in very different power response signatures.

4.3.2   Randomized Decentralized Approaches

In order to de-synchronize the responses of TCLs while providing PFC, a number of randomized approaches were proposed in the literature. For example, a control approach with random on/off frequency thresholds and turn-on delay times was used in an experimental demonstration with 150 residential cloth dryers and 50 Electric Water Heaters (EWHs) in [91]. Reference [92] proposed a randomized load control scheme, where each load monitors the frequency over random time instances and responds according to a simple control policy based on frequency thresholds. Using this scheme the authors derived the evolution of frequency's mean value and variance over time, as well as estimates of the frequency recovery time and probability of frequency overshoot after a contingency.

Reference [88] proposed an interesting stochastic approach to control a population of refrigerators. Each device is modeled as a Markov-jump linear system with transition probabilities between the on and off states that depend indirectly on the frequency deviation. The authors derived closed-form expressions for the mean value and the variance of the average temperature within the population, and proved that the closed-loop system is asymptotically stable. However, the main limitation of this approach is that the control adjusts the properties of the temperature's PDF at steady state, and therefore results in slow responses that are unacceptable for PFC.

Another relevant work is [93], where a a decentralized stochastic controller was developed to allow a TCL aggregation to collectively track a desired aggregate power profile. The main innovation of [93] over previous works is that the average heating rate of the population itself is used as a control variable. The formulations of [93] result in control laws for individual appliances, namely temperature limit changes and switching rates, and allow an aggregator to estimate the available flexibility from the aggregation. However, the control approach of [93] is able to respond to demand reduction requests, but not to demand increase requests, and therefore it is not appropriate for PFC.

More related to our work is [94] where a decentralized stochastic controller based on probabilistic switching was proposed for refrigerators to provide PFC. This controller results in very fast responses that are suitable for PFC, but it has a number of limitations that will be discussed in Chapter 5.

4.3.3   Approaches based on Utility Maximization

A number of works have proposed methods to provide PFC while minimizing the cost of reserve provision or maximizing the social welfare. For example, [95] proposed a mixed integer linear optimization problem formulation to allocate trigger frequencies to a set of on/off devices such that they collectively provide PFC reserves at minimum cost. The authors of [96] designed a decentralized PFC strategy for loads whose power can be modulated within a certain range around their nominal value. In addition, they developed an optimization scheme to maximize the participation of loads in PFC, while compensating for the payback energy in an optimal way.

Reference [97] proposed a frequency support method for residential loads based on utility functions and bidding, which guarantees maximal social welfare. The authors showed that by adding a frequency-dependent price component to the market price, the contribution of each load to the PFC reserve can be computed in a decentralized way. A similar topic was investigated in [98], where the authors formulated a load control optimization problem, and developed a synchronous and an asynchronous algorithm to solve it in a decentralized way based on local frequency measurements. The follow-up work [99] extended the methods of [98] for a multi-machine power system and proved that the system under the proposed load-side PFC is asymptotically stable.

4.4   Demand-Side Secondary Frequency Control

4.4.1   SFC with Thermostatically Controlled Loads

Thermostat and On/Off Load Control Approaches

The early work on centralized control of TCL aggregations focused on provision of load following or frequency regulation. Reference [100] developed control strategies to provide such services via a thermostat setpoint manipulation, and showed that a simple linear model can describe the TCL aggregate transient dynamics. The authors of [101] developed a controller based on an input-output, second-order, linear time-invariant model to regulate the aggregate power of a TCL aggregation via broadcasts of thermostat setpoint changes. The aggregate TCL dynamics

under thermostat setpoint control were described by a transfer function in [102], and a linear quadratic regulator was used to control the aggregate power.

Other researchers developed control strategies based on direct on/off control of loads. A strategy that relies on two-way communication links between the aggregator and each TCL was proposed in [103]. Coordinated control using signal broadcasts is possible with probabilistic switching, where a fraction of the devices switches stochastically to provide the desired service. Reference [104] showed how bi-directional control (up- and down-regulation) can be achieved by broadcasting a switching fraction signal that only switches off devices. In [105], a statistical load modeling approach based on Markov Chains was introduced and a Model Predictive Control (MPC) scheme was used to control a population of TCLs with switching probability broadcasts.

A broadcast controller was used in [106] together with an optimization process to determine the switching fractions that result in good tracking of a given reference power trajectory. The control design of [106] relied on the evolution of the PDF of temperature extending the work of [82]. Reference [107] proposed a hierarchical framework for demand-side frequency control. The upper layer is centralized and computes optimal control gains for the participating loads depending on the bus where they are connected. The lower layer is decentralized, and each device computes switching probabilities based on the control gains.

While most of the works considered generic TCL models, some papers used more detailed device-specific models in direct load control studies. For example, [28] and [29] used "single point" EWH models assuming a uniform temperature profile within the water tank. In [30], an EWH stratified model was used, where only an upper hot water layer and a lower cold water layer were considered. In [31], the authors developed a six-layer stratified model (without considering convection) to simulate EWHs subject to off-peak schedules. In addition, a broadcast controller was proposed in [108] to allow the aggregate power of a population of air conditioners to track a power reference signal taking into account compressor lockout times.

State Estimation to Reduce Communication

The above works, as well as other relevant approaches that rely either on probabilistic switching or rule-based controllers [21,30], assumed access to TCL state measurements (temperature and on/off state). More

recent works proposed state estimation methods to reduce the needs for real-time state measurements and consequently the implementation costs. For example, [3] used a Kalman Filter to estimate the state of a Markov model for a TCL aggregation, whereas [109] used a similar model and a Moving Horizon State Estimation (MHSE) approach. Reference [110] developed a particle filter to estimate the states of a simple four-state aggregate system model, whereas [111] proposed a state estimation scheme based on Partial Differential Equations (PDEs).

4.4.2 SFC with Commercial Buildings

Most of the early work on commercial buildings focused on the development of models for DR applications [112, 113], and on using the building's thermal mass for load shifting and peak shedding to minimize the energy costs and demand charges [10, 11, 13]. Similar applications were considered more recently in [17, 114], where MPC approaches were proposed.

Early Work on Ancillary Services

Some works investigated the potential of commercial buildings to participate in AS markets. A retail store and an office building participated in a pilot program for non-spinning reserves in the California Independent System Operator's AS market using global temperature adjustments and the OpenADR protocol [115]. In [116], spinning reserve experiments with a duration of 15 minutes were performed curtailing the air conditioning load of a hotel. The results showed that the load response is fast and that the room temperature rise remains acceptable.

More recently, [117] used a detailed model of a Variable Air Volume (VAV) HVAC system to simulate the performance of setpoint adjustments in zone temperature, duct static pressure, supply air temperature, and chilled water temperature while providing spinning reserve. However, there is a limited amount of theoretical, simulation or experimental work on frequency regulation with commercial buildings.

Feasibility of Tracking SFC Signals

A few recent papers studied the feasibility of tracking SFC signals by controlling the power consumption of various HVAC system components. Reference [118] investigated the control of Heat Pumps (HPs)

via regulation of the refrigerant's flow rate with valves. A dynamic model was fitted based on step response tests on a real medium-size HP of a commercial building, and simulation results showed that the refrigerant flow rate control is appropriate for frequency regulation.

The suitability of static pressure and zone temperature setpoint adjustments for frequency regulation was investigated in [119] in a simulation study with a very detailed commercial building model. According to the results, static pressure setpoint changes provide a faster initial power response, but the control performance gradually degrades due to the compensating reactions of the dampers.

The authors of [120,121] investigated frequency regulation via fan power control in buildings with VAV systems. Simulations showed that up to 15% of a building's fan power can be successfully offered as reserves, if the frequency band of the regulation signal is $f \in [1/(10 \text{ min}), 1/(4 \text{ sec})]$. The follow-up work [122] included chiller control enlarging the frequency band to $1/(60 \text{ min})$.

SFC Reserve Capacity Estimation

A challenge in frequency regulation with commercial buildings is to determine the reserve capacity that can be offered reliably to the TSO. This is a very important topic because buildings are energy-constrained resources, and it is only marginally discussed in the literature. MPC was used in [123] to quantify the flexibility of a commercial building online and offer it to a utility. Reference [124] proposed a simulation-based method to estimate the reserve capacity, but without considering the time-coupling across different scheduling intervals. Recently, [125] developed a virtual battery model for a commercial building and used it to estimate the building's rate limits and energy capacity for frequency regulation.

Baseline Estimation and Building Aggregations

An important challenge for both the SFC performance and financial settlement is to estimate the baseline consumption of the building, i.e., the consumption without frequency regulation [126]. Baseline estimation was performed online in [127] using a low-pass filter. If MPC is used, as in [123], the baseline power is known ahead of time, which is advantageous because it facilitates the financial settlement.

Due to the minimum bid size limitations of several AS markets, aggregations of a few tenths or hundreds of commercial buildings are necessary to gain access to the market. Unfortunately, SFC from commercial building aggregations is not sufficiently studied in the literature.

An MPC scheme was designed in [128] to provide reserves from building HVAC systems in parallel to generators. However, approximate ramping rates were used in [128] to represent the building aggregation instead of realistic building models. Reference [129] investigated the real-time allocation of control actions within a building aggregation while providing frequency regulation, but without a priori reserve capacity scheduling. Different thermal load types were considered in [130], but the participation factors were empirically estimated instead of systematically determined.

Experimental Demonstrations

Although the above simulation works provide some evidence that frequency regulation can be provided by commercial buildings, experimental verification is necessary to build confidence for wide-spread implementation. To the best of our knowledge, there have been only a few experimental demonstrations and field tests so far. Perhaps the first demonstration project was reported in [131], where the feasibility of offering up- and down-regulation products with university campus buildings was investigated using global temperature adjustments and changing the ventilation power.

Reference [132] experimentally demonstrated that fans in commercial buildings can provide frequency regulation. The experiment was performed using archived regulation signals from PJM and the OpenADR protocol, whereas the frequency of the fan's Variable Frequency Drive (VFD) was used as a control variable in an open-loop fashion. The results showed that the delays introduced by OpenADR and the BAS are low enough to allow for a sufficient tracking accuracy.

Experimental results for frequency regulation via fan control in an auditorium of a university campus building were reported in [127]. Two control approaches were tested and shown to perform well: (i) superimposing a fan command on the output of the existing fan controller, and (ii) changing the air flow rate setpoint. Reference [123] reported results from a small set of experiments where the static duct pressure setpoint

was randomly varied within a range every minute, and the effects on fan speed and room temperatures were monitored.

A data-driven dynamic model for a variable speed HP was developed in [133, 134] and used in frequency regulation experiments in a lab-scale microgrid in [134]. The authors concluded that although using the supply water temperature setpoint as a control variable is sufficient to pass PJM's prequalification test, direct control of the compressor's frequency results in a much faster response.

Reference [135] developed a frequency regulation controller for a chiller consisting of a Proportional-Integral-Derivative (PID) loop with the chilled water supply temperature setpoint as a control variable, a high-pass filter of the regulation signal, and a baseline estimator. The authors demonstrated the effectiveness of the controller in an experiment with an office building using actual RegA and RegD signals from PJM [73]. The follow-up work [136] performed an additional set of experiments on another building and identified the BAS delay time, the chiller ramp-rate limits, the chiller minimum cooling power constraints, and the transient variations in the Coefficient of Performance (COP) as important issues for practical implementation. Relevant is also the work of [137], where experiments were performed to analyze the efficiency of fast DR actions in a commercial building.

Chapter 5

Primary Frequency Control with Refrigerators

5.1 Introduction

A literature review on Primary Frequency Control (PFC) methods for aggregations of Thermostatically Controlled Loads (TCLs) was provided in Chapter 4. References [88, 92–94] are particularly interesting because they designed decentralized stochastic controllers to control an aggregation of refrigerators without real-time communication and avoided synchronization.

We follow this line of research and develop a new decentralized controller for refrigerators[1] that relies on switching probabilities, which are computed directly based on the frequency deviation. This is in contrast to [92], where the randomization is with respect to the time instance when each load reacts to a given frequency deviation. Our approach is also fundamentally different to [88, 93], where the devices' switching

---

[1]Although we consider refrigerators, the proposed approach can be applied to other TCLs with compressors and similar dynamics.

Table 5.1: Nomenclature of Chapter 5: symbols

Symbol	Unit	Description
C	kJ/$^\circ$C	Refrigerator thermal capacitance
e_{r}	%	Instantaneous reserve error
$e_{\mathrm{b,mape}}$	%	Baseline MAPE
$e_{\mathrm{r,mape}}$	%	Reserve MAPE
$e_{\mathrm{p,rmse}}$	-	Switching probability RMSE
D	-	Duty cycle
D^{a}	-	Actual duty cycle (with PFC)
D^{n}	-	Nominal duty cycle (without PFC)
D^{d}	-	Desired duty cycle (with PFC)
D^{r}	-	PFC reserve capacity in terms of duty cycle
$E_{\mathrm{cl}}^{\mathrm{d}}$	kJ	Daily energy consumption without door openings
$E_{\mathrm{op}}^{\mathrm{d}}$	kJ	Daily energy consumption with door openings
F_{off}	-	CDF of the lock-off time in the aggregation
F_{on}	-	CDF of the lock-on time in the aggregation
K_{c}	-	Corrective temperature gain
$K_{\mathrm{r},t}$	$^\circ$C	Resetting factor for the thermostat limits
l_{off}	-	Lock-off event as a random variable
l_{on}	-	Lock-on event as a random variable
$L_{\mathrm{off}}^{\mathrm{st}}$	-	Steady-state fraction of refrigerators locked-off
$L_{\mathrm{on}}^{\mathrm{st}}$	-	Steady-state fraction of refrigerators locked-on
$L_{\mathrm{off},t}^{\mathrm{tr}}$	-	Transient fraction of refrigerators locked-off
$L_{\mathrm{on},t}^{\mathrm{tr}}$	-	Transient fraction of refrigerators locked-on
m	-	Refrigerator on/off state

rates are post-calculated based on the desired evolution of mean temperature and the average heating rate, respectively. The main advantage of the probabilistic switching approach is that it achieves a much faster PFC response compared with previous works.

A similar probabilistic switching approach for PFC with refrigerators was recently proposed in [94]. This approach relies on four main assumptions: (i) each refrigerator consumes a constant amount of power in the on state; (ii) each refrigerator can cycle on and off arbitrarily often; (iii) the frequency deviation signal is filtered in order to be zero-mean, i.e., it crosses zero very often; and (iv) the refrigerators are not exposed to door openings. Assumptions (i), (ii) and (iv) were also made in [93]. However, these four assumptions will likely not hold in a real-world implementation. The power consumption of a refrigerator is typically higher at the beginning of each on cycle, its compressor cannot turn on and off arbitrarily often, and its temperature is affected by door openings. Furthermore, in practice the frequency deviation signal can be significantly biased to the positive or negative direction.

In this chapter, we maintain the same basic probabilistic switching

Table 5.2: Nomenclature of Chapter 5: symbols (continued)

Symbol	Unit	Description
N_{ev}	s	Duration of a step frequency deviation event
N_{on}	s	Maximum lock-on time within the aggregation
N_{off}	s	Maximum lock-off time within the aggregation
N_r	-	Size of refrigerator aggregation
N_{rec}	s	Recovery period after a step frequency deviation
$N_{r,a}$	-	Number of refrigerators activated for PFC
N_s	s	Duration of startup dynamics
N_{sim}	s	Simulation period
N_δ	s	Half-period of the frequency deviation signal
P_{agg}	W	Actual aggregate electric power
P_b	W	Uncontrolled aggregate electric power
P_d	W	Desired aggregate electric power (with PFC)
P_{res}	W	Reserve capacity of the aggregation
P_n	W	Nominal refrigerator electric power
q	-	Switching probability of the benchmark controller
R	$^\circ$C/kW	Refrigerator's thermal resistance with closed door
R_{op}	$^\circ$C/kW	Refrigerator's thermal resistance with open door
r	-	Normalization factor for the resetting factor
S_{off}	-	Survival function of the lock-off time in the aggregation
S_{on}	-	Survival function of the lock-on time in the aggregation
S_u	-	Refrigerator startup dynamics profile
$s_{net,t}$	-	Average net switching rate in the aggregation
$s_{off,t}$	-	Average rate at which refrigerators switch off
$s_{on,t}$	-	Average rate at which refrigerators switch on

structure, which is recapped in Section 5.2, but enhance it with several additions to drop the aforementioned limiting assumptions. Specifically, we design a controller that can cope with loads with time-varying power profiles during the startup phase, as well as limitations on compressor cycling (the so-called lockout constraints). These are the topics of Sections 5.3, 5.4, and 5.5.

We show that the control performance is improved if the probabilistic switching is combined with a deterministic resetting of thermostat temperature limits. We investigate this in Section 5.7 and analytically calculate the necessary resetting factor.

In Section 5.8, we show the need for an additional control loop to satisfy temperature requirements and we provide analytical expressions for upper and lower bounds of the control gain of this loop. In addition, we present a simple method to estimate the population's mean temperature in a decentralized way.

A common limitation in approaches that rely on resetting of the thermostat's temperature limits is the thermostat resolution. We address this issue in Section 5.9 by adding another randomization layer in the con-

Table 5.3: Nomenclature of Chapter 5: symbols (continued)

Symbol	Unit	Description
T	$^\circ$C	Refrigerator temperature
T_{a}	$^\circ$C	Ambient temperature
$T_{\max}$	$^\circ$C	Higher deadband limit of thermostat
$T_{\min}$	$^\circ$C	Lower deadband limit of thermostat
$\bar{T}_{\mathrm{nom}}$	$^\circ$C	Nominal mean temperature (without PFC)
T_{set}	$^\circ$C	Thermostat setpoint $[T_{\mathrm{set}} = 0.5 \cdot (T_{\max} + T_{\min})]$
$\dot{T}_{\mathrm{d}}$	$^\circ$C/s	Average temperature decrease rate
$\dot{T}_{\mathrm{i}}$	$^\circ$C/s	Average temperature increase rate
$\bar{T}_{\mathrm{cl}}$	$^\circ$C	Average refrigerator temperature with closed door
$\bar{T}_{\mathrm{op}}$	$^\circ$C	Average refrigerator temperature with open door
t_{off}	s	Duration of refrigerator's off cycle
t_{off}^{l}	s	Refrigerator lock-off time
t_{on}	s	Duration of refrigerator's on cycle
t_{on}^{l}	s	Refrigerator lock-on time
u	-	Peak power factor for startup dynamics modeling
w	$^\circ$C/s	Noise term for refrigerator's external disturbances
x	-	Fraction of switched loads for PFC under startup dynamics
z	-	Thermostat resetting event as a Bernoulli random variable

troller, and we analytically show that the side-effect of the new control layer is a monotonic increase in the variance of refrigerator temperatures within the aggregation. This side-effect is mitigated by introducing a bound on the maximum deviation from the population's mean temperature. Another addition compared with [88, 94] is that we model the refrigerator's door opening events, and propose a simple way to explicitly account for them in the control design. This is the topic of Section 5.10.

In Section 5.11, we introduce the model parameters as well as the metrics and benchmarks used to assess the control performance. In Section 5.12, we analytically investigate the dependence of PFC performance on the size of refrigerator aggregation, and we backup our findings with simulations. The controller's sensitivity on important design and load parameters such as the reserve capacity, the peak power during the startup phase, the lockout time and the PFC activation deadband is investigated in detail in Section 5.13.

Sections 5.14, 5.15 and 5.16 present simulation results showing the controller's performance with biased frequency deviations, limited thermostat resolution, and door openings. Last, we discuss some practical issues for a real-world implementation in Section 5.17, whereas Section 5.18 concludes. This chapter is based on the work reported in [138, 139], and its nomenclature is summarized in Tables 5.1-5.6.

Table 5.4: Nomenclature of Chapter 5: Greek letters

Symbol	Unit	Description
α	1/s	Thermal parameter of continuous-time model
β	°C/kJ	Thermal parameter of continuous-time model
γ	°C/Hz	Auxiliary variable for PFC
ΔD	-	Duty cycle change
ΔD_t^{op}	-	Increase in duty cycle due to door openings
ζ	-	Auxiliary variable for thermostat resetting events
Δf	Hz	Frequency deviation
ΔT	°C	Width of thermostat's deadband
ΔT_{b}	°C	Bound for refrigerator's thermostat limits deviation
ΔT_{lim}	°C	Change in refrigerator's thermostat limits
ΔT_{res}	°C	Minimum resolution of refrigerator's thermostat
Δt	s	Discretization time step
δ	Hz	Magnitude of frequency deviation signal
ε	°C	Tolerance for temperature deviation during a frequency event
ϵ	°C	Tolerance for temperature deviation after a frequency event
η	-	Coefficient of performance
λ	-	Auxiliary variable equal to $1 - K_{\mathrm{c}}$
μ_{dur}	s	Mean duration of a door opening event
μ_{op}	-	Average number of door openings per day
ν	-	Parameter for modeling of startup dynamics
ξ	-	Normalized energy consumption increase due to door openings
ρ	-	Switching probability of the proposed controller
ϱ	-	Switching probability for resetting of thermostat limits
σ_{dur}	s	Standard deviation of the duration of a door opening event
σ_{op}	-	Standard deviation of the number of door openings per day
χ	-	Switching action of a refrigerator (Bernoulli random variable)

Table 5.5: Nomenclature of Chapter 5: probability operators

Operator	Description
$\mathbb{E}\left[\cdot\right]$	Expected value of a random variable
$\mathbb{N}$	Normal probability distribution
$\mathbb{P}\left[\cdot\right]$	Probability of a random variable
$\mathrm{SD}\left[\cdot\right]$	Standard deviation of a random variable
$\mathbb{U}$	Uniform probability distribution
$\mathrm{Var}\left[\cdot\right]$	Variance of a random variable

5.2 Modeling and Basic Control Design

5.2.1 Refrigerator Model

Consider a refrigerator i without freezer controlled by an on/off hysteresis controller. We use a standard first-order differential equation to model the lumped temperature $T_i(t)$ of the refrigerator including the

Table 5.6: Nomenclature of Chapter 5: subscripts and superscripts

Subscript/Superscript	Description
i	Refrigerator index within an aggregation
min/max	Minimum/maximum value of a variable or parameter
k, t	Time indexes
$(\bar{\cdot})$	Average value of a variable or parameter
$(\hat{\cdot})$	Estimated value of a variable or parameter
$[\cdot]_+$	Caps a variable to non-negative values

air and solid mass (compartment and content) [83, 88, 140]

$$\dot{T}_i(t) = \alpha_i\big[T_{a,i} - T_i(t)\big] + w_i(t) \qquad \text{when OFF} \qquad (5.1)$$

$$\dot{T}_i(t) = \alpha_i\big[T_{a,i} - T_i(t)\big] - \beta_i P_{n,i} + w_i(t) \qquad \text{when ON,} \qquad (5.2)$$

where $\alpha = 1/RC$, $\beta = \eta/C$, C is the thermal capacitance, R is the thermal resistance, T_a is the room temperature, η is the Coefficient of Performance (COP), and P_n is the nominal power. The variable $w(t)$ is a noise term that aggregates the effect of external disturbances, e.g., door openings, changes in food content, and variations of T_a. In this chapter, we first design the decentralized controller neglecting these external disturbances. Then, in Section 5.10 we show how the effect of door openings, which is the dominant external disturbance, can be integrated in the refrigerator model and control design.

The model is discretized with a discretization time step Δt. Let $T_{\min,i}$ and $T_{\max,i}$ denote the deadband limits of the thermostat's hysteresis controller and $m_{i,t} \in \{0, 1\}$ denote the compressor's on/off state. Using the discrete-time model we can derive the duration of the on and off cycles

$$t_{on,i} = R_i C_i \ \ln\left[\frac{T_{\max,i} - T_{a,i} + \eta_i R_i P_{n,i}}{T_{\min,i} - T_{a,i} + \eta_i R_i P_{n,i}}\right] \qquad (5.3)$$

$$t_{off,i} = R_i C_i \ \ln\left[\frac{T_{\min,i} - T_{a,i}}{T_{\max,i} - T_{a,i}}\right], \qquad (5.4)$$

and the device's duty cycle $D_i = t_{on,i}/(t_{on,i} + t_{off,i})$.

5.2.2   Aggregation Model

Consider a population of N_r refrigerators providing PFC. For control purposes we are interested in the aggregate power of the population

$P_{\text{agg},t}$, which depends on the aggregate duty cycle, i.e., the proportion of loads that are in the on state

$$D_t = \frac{\sum_i m_t^i}{N_{\text{r}}} \ , \tag{5.5}$$

where $i \in [1, N_{\text{r}}]$ is the refrigerator index. If N_{r} is sufficiently large, then the aggregate power at any time will be approximately equal to

$$P_{\text{agg},t} \approx N_{\text{r}} D_t \bar{P}_{\text{n}} \ , \tag{5.6}$$

where $\bar{P}_{\text{n}}$ is the mean refrigerator power in the aggregation. Without PFC, and if w_t is sufficiently small, the duty cycle will be approximately constant and equal to a nominal value D^{n}.

5.2.3 PFC with Probabilistic Switching

PFC reserves can be provided by perturbing the population's duty cycle around D^{n} proportionally to the frequency deviation Δf_t with a gain D^{r} that corresponds to the PFC reserve capacity. Therefore, to activate the correct amount of reserve power after a frequency deviation Δf_t, the duty cycle should be controlled to the desired value

$$D_t^{\text{d}} = D^{\text{n}} + D^{\text{r}} \cdot \frac{\Delta f_t}{\Delta f_{\text{max}}} \ , \tag{5.7}$$

where Δf_{max} is the frequency deviation for full PFC reserve activation, e.g., $\Delta f_{\text{max}} = 0.2$ Hz in continental Europe.

Denote by $\Delta D_t = D_t^{\text{d}} - D_{t-1}^{\text{d}}$ the change in the desired duty cycle between two consecutive time steps. Note that ΔD_t depends on Δf_t and Δf_{t-1}, which can be measured locally by each refrigerator, and so the duty cycle change can be achieved in a decentralized way with probabilistic switching. Each device calculates the switching probability

$$q_t = \begin{cases} \frac{\Delta D_t}{1 - D_{t-1}^{\text{d}}}, & \text{if } \Delta D_t \geq 0 \\ -\frac{\Delta D_t}{D_{t-1}^{\text{d}}}, & \text{if } \Delta D_t < 0 \ , \end{cases} \tag{5.8}$$

and generates a random number uniformly distributed in $[0, 1]$. If the number is smaller than q_t, the device will switch to contribute to PFC. The first line of (5.8) is the probability for devices that are off to switch on, whereas the second line is the probability for devices that are on to

switch off. Due to the law of large numbers, the proportion of switched devices will be very close to q_t, and the population will collectively provide the required PFC reserve.

In the remainder of this chapter, we extend the basic probabilistic switching approach of this section to account for startup dynamics, lockout constraints, biased frequency deviations, limited thermostat resolution, and door openings.

5.2.4   Communication Aspects

Assuming that the refrigerators know the value of D^{r} and measure the frequency deviation locally, they can provide PFC in a decentralized way without real-time communication. However, some infrequent communication might be needed (or desired) for the proposed control scheme to work effectively.

For example, the value of D^{r} reflects the PFC reserve capacity, which is in some cases periodically determined by the reserve market. Therefore, whenever the reserve market is cleared (e.g., once a week), the updated value of D^{r} should be communicated to the refrigerators. As it will be shown in the following chapters, some of the controller parameters depend on the properties of the refrigerator population. Therefore, if these properties change over time, the respective controller parameters should be updated. A communication channel with low data transfer can be used to enable these types of infrequent communication.

In the above cases, the refrigerators provide PFC reserves in a semi-autonomous way: the real-time control is decentralized, while some of the controller parameters are updated infrequently through a communication link. We provide a more detailed discussion on the communication aspects in Section 5.17.

5.3   Refrigerator Startup Dynamics

Typically, a refrigerator's compressor consumes more power during the startup phase (approximately 30 seconds) due to a smaller COP, higher evaporator temperature, and higher motor current [141].[2]   Since this

---

[2]The higher motor current is the dominant factor for the first few seconds after the startup, whereas the smaller COP is the main factor afterwards.

additional power can be as high as 25% of the nominal power P_{n}, it is important to account for it in PFC to avoid large overshoots.

The startup dynamics of a refrigerator i are modeled using a peak power factor $u_i > 0$ and the duration of the startup phase $N_{\mathrm{s},i}$. We assume that the power becomes immediately equal to $(1 + u_i)P_{\mathrm{n},i}$ when the refrigerator switches on, and then it linearly decreases until it becomes equal to $P_{\mathrm{n},i}$ after $N_{\mathrm{s},i}$ seconds. Therefore, the refrigerator power in the on state can be expressed as

$$P_{i,t} = P_{\mathrm{n},i}[1 + S_{\mathrm{u},i}(t)] = P_{\mathrm{n},i}\left[1 + u_i\left[1 - \frac{t}{N_{\mathrm{s},i}}\right]_+\right], \qquad (5.9)$$

where $S_{\mathrm{u},i}$ is the startup power profile and $[\cdot]_+$ caps its argument to positive values, i.e., $S_{\mathrm{u},i}(t) = 0$ for $t > N_{\mathrm{s},i}$.

ΔD_t in (5.8) is the fraction of loads that need to switch at each time step, and it depends only on Δf_t and Δf_{t-1}. However, in the presence of startup dynamics, the fraction of loads that need to switch at the current time step depends additionally on the number of loads that switched due to previous reserve activations. Therefore, the fraction of loads to be switched is in general different from ΔD_t, and it is denoted by x_t in the rest of this chapter.

For control design we assume to know only the average values $\bar{P}_{\mathrm{n}} = \mathbb{E}\left[P_{\mathrm{n},i}\right]$, $\bar{u} = \mathbb{E}\left[u_i\right]$ and $\bar{N}_{\mathrm{s}} = \mathbb{E}\left[N_{\mathrm{s},i}\right]$ in the population. Based on the average values, we define the average startup duty cycle profile

$$\bar{S}_{\mathrm{u}}(t) = \bar{u} \cdot \left[1 - \frac{t}{\bar{N}_{\mathrm{s}}}\right]_+, \qquad (5.10)$$

and we use it to recursively calculate x_t with

$$x_t = \frac{1}{1 + \nu_t}\left[D_t^{\mathrm{d}} - D_{t-1}^{\mathrm{a}} - \sum_{k=t-N_{\mathrm{s}}+1}^{t-1} x_k \bar{S}_{\mathrm{u}}(t - k + 1)\right] \qquad (5.11)$$

$$D_t^{\mathrm{a}} = D_{t-1}^{\mathrm{a}} + x_t . \qquad (5.12)$$

D_t^{a} is the actual duty cycle of the population, whereas $\nu_t = \bar{S}_{\mathrm{u}}(1)$ holds if the value of the bracket at the right hand side of (5.11) is non-negative, and $\nu_t = 0$ holds if the value of the bracket is negative. This differentiation is necessary because there exist no dynamics at shutdown. The term $1/(1 + \nu_t)$ in (5.11) reduces the fraction of switched loads x_t as

much as the expected overshoot due to the current startup dynamics, whereas the summation term of (5.11) recursively accounts for the effect of startup dynamics of past control actions on the current x_t.

Using (5.11) and (5.12) in the controller, we estimate the effect of refrigerators' startup dynamics on aggregate power and modify the switching probability accordingly. Since the startup dynamics are accounted for using the mean values $\bar{P}_n$, $\bar{u}$ and $\bar{N}_s$, the control will not be perfect. However, it is possible to characterize its performance analytically as shown by Proposition 1.

Proposition 1. *Denote by $N_{s,min}$ and $N_{s,max}$ the minimum and maximum startup duration in the population of refrigerators. The estimated aggregate power due to refrigerator startup dynamics obtained using the average values $\bar{P}_n$, $\bar{u}$, $\bar{N}_s$ and (5.10) is an upper bound of the actual aggregate power up to $N_{s,min}$ seconds after the reserve activation, irrespective of the probability distribution of $N_{s,i}$.*

Furthermore, if the random variable $N_{s,i}$ follows the uniform distribution $N_{s,i} \sim [N_{s,min}, N_{s,max}]$, then the estimated aggregate power is an upper bound for up to t_{lim} seconds after the reserve activation, whereas it is a lower bound afterwards, where t_{lim} is given by

$$t_{lim} = \frac{N_{s,max}\left(N_{s,min} + N_{s,max}\right)}{3N_{s,max} - N_{s,min}} \;. \tag{5.13}$$

Proof. The proof is given in Appendix A.   ∎

Another effect of the startup dynamics is that they increase the refrigerator's average power, both with and without PFC. This happens because typically the COP is low during the startup phase, and so the additional electric power does not translate into more cooling. With reference to the left plot of Fig. 5.1, and using (5.9), it is easy to derive the following relationship between the parameters of startup dynamics and the mean power of an individual refrigerator

$$\bar{P}_i = \frac{\sum_{t=0}^{t_{\mathrm{on},i}+t_{\mathrm{off},i}} P_{i,t}\Delta t}{t_{\mathrm{on},i} + t_{\mathrm{off},i}} = D_i P_{\mathrm{n},i} + \frac{0.5 \cdot u_i P_{\mathrm{n},i} N_{\mathrm{s},i}}{t_{\mathrm{on},i} + t_{\mathrm{off},i}} \;. \tag{5.14}$$

The first term of (5.14) is the refrigerator's mean power in the absence of startup dynamics. The numerator of the second term is the area of the triangle shown in the left plot of Fig. 5.1. Using the mean values
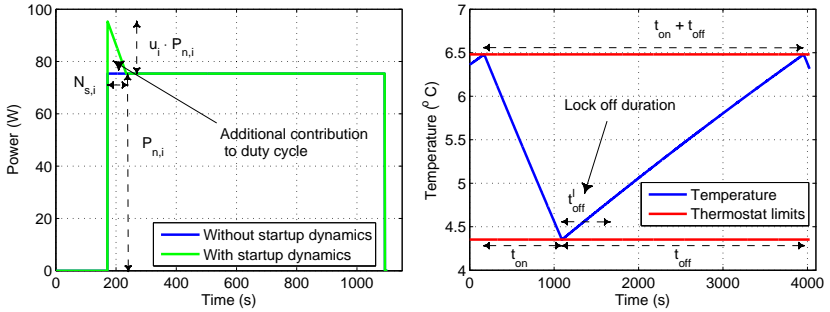
Figure 5.1: Left: Effect of startup dynamics on a refrigerator's power and duty cycle. Right: Evolution of the refrigerator's temperature during an on/off cycle.

of the parameters in (5.14), we can get a similar expression for the population's increased aggregate power that is used as a baseline for PFC evaluation purposes.

5.4 Compressor Lockout Constraints

Lockout times are usually employed to avoid the compressor's frequent switching that decreases efficiency and reduces lifetime [108]. Let us denote by t_{on}^l the lock-on time, i.e., the minimum duration the device must operate after switching on, and by t_{off}^l the lock-off time, i.e., the minimum duration the device must remain off after switching off. When a compressor is locked, it cannot react to frequency deviations; therefore, it is important to estimate the fraction of locked devices. In the rest of this section, we present a simple estimation method based on local frequency measurements to keep the PFC fully decentralized.

5.4.1 Steady-State Lockout due to Thermostatic Control Actions

Even without PFC (at steady state), at any time step t a fraction L_{on}^{st} of refrigerators is locked in the on state and a fraction L_{off}^{st} is locked in the off state due to thermostatic control actions. Considering the cycle of an individual refrigerator in the right plot of Fig. 5.1, the lock

event at the off state l_{off} can be modeled as the Bernoulli probability distribution

$$f_{\text{off}}^l(l_{\text{off}}) = \begin{cases} \frac{t_{\text{off}}^l}{t_{\text{on}}+t_{\text{off}}}, & \text{if } l_{\text{off}} = 1 \\ 1 - \frac{t_{\text{off}}^l}{t_{\text{on}}+t_{\text{off}}}, & \text{if } l_{\text{off}} = 0 \ . \end{cases} \quad (5.15)$$

Using the properties of Bernoulli distribution, the fraction of refrigerators that is locked at the off state is

$$L_{\text{off}}^{\text{st}} := \frac{1}{N_{\text{r}}} \cdot \sum_{i=1}^{N_{\text{r}}} l_{\text{off},i} = \mathbb{E}\left[l_{\text{off}}\right] = \frac{t_{\text{off}}^l}{t_{\text{on}} + t_{\text{off}}} \ . \quad (5.16)$$

Since t_{off}^l, t_{on} and t_{off} are random variables, $L_{\text{off}}^{\text{st}}$ is also a random variable with a so-called ratio distribution. Without knowledge of the probability distribution of t_{off}^l, t_{on}, t_{off}, the expectation of $L_{\text{off}}^{\text{st}}$ can be approximated with a first order Taylor expansion [142]

$$\mathbb{E}\left[L_{\text{off}}^{\text{st}}\right] = \mathbb{E}\left[\frac{t_{\text{off}}^l}{t_{\text{on}} + t_{\text{off}}}\right] \approx \frac{\mathbb{E}\left[t_{\text{off}}^l\right]}{\mathbb{E}\left[t_{\text{on}}\right] + \mathbb{E}\left[t_{\text{off}}\right]} \ . \quad (5.17)$$

A second order approximation can be also used if the variance of $t_{\text{on}}+t_{\text{off}}$ is known. Using a similar procedure, one can get the approximation for the steady-state lockout at the on state

$$\mathbb{E}\left[L_{\text{on}}^{\text{st}}\right] \approx \frac{\mathbb{E}\left[t_{\text{on}}^l\right]}{\mathbb{E}\left[t_{\text{on}}\right] + \mathbb{E}\left[t_{\text{off}}\right]} \ . \quad (5.18)$$

5.4.2 Transient Lockout due to PFC

When providing PFC the refrigerators are exposed to additional switching actions that increase the fraction of locked devices. Let us denote by $L_{\text{on},t}^{\text{tr}}$ and $L_{\text{off},t}^{\text{tr}}$ the time-varying fractions of the devices that lock at the on and off state due to PFC. Since $L_{\text{on},t}^{\text{tr}}$ and $L_{\text{off},t}^{\text{tr}}$ depend on the past reserve activations, they can be estimated recursively given sufficient statistical information. We assume that the Cumulative Distribution Functions (CDFs) of the lockout times are known and we denote them by $F_{\text{on}}(t_{\text{on}}^l)$ for lock-on time and $F_{\text{off}}(t_{\text{off}}^l)$ for lock-off time. We further define $N_{\text{on}} = \max\left[t_{\text{on}}^l\right]$ and $N_{\text{off}} = \max\left[t_{\text{off}}^l\right]$. The fractions $L_{\text{on},t}^{\text{tr}}$ and

$L_{\mathrm{off},t}^{\mathrm{tr}}$ can be computed using

$$L_{\mathrm{on},t}^{\mathrm{tr}} = \sum_{k=0}^{t-1} c_k x_k S_{\mathrm{on}}(t-k), \quad c_k = \begin{cases} 1, & \text{if } x_k \geq 0 \\ 0, & \text{if } x_k < 0 \end{cases} \tag{5.19}$$

$$L_{\mathrm{off},t}^{\mathrm{tr}} = \sum_{k=0}^{t-1} d_k x_k S_{\mathrm{off}}(t-k), \quad d_k = \begin{cases} 0, & \text{if } x_k \geq 0 \\ 1, & \text{if } x_k < 0 \end{cases}, \tag{5.20}$$

where $S_{\mathrm{on}} = 1 - F_{\mathrm{on}}$ and $S_{\mathrm{off}} = 1 - F_{\mathrm{off}}$ are the survival functions of the CDFs, $S_{\mathrm{on}}(t-k) = 0$ for $t-k > N_{\mathrm{on}}$, and $S_{\mathrm{off}}(t-k) = 0$ for $t-k > N_{\mathrm{off}}$. The total fraction of locked devices is obtained by adding the steady-state and transient contributions

$$L_{\mathrm{on},t} = \mathbb{E}\left[L_{\mathrm{on}}^{\mathrm{st}}\right] + L_{\mathrm{on},t}^{\mathrm{tr}}, \quad L_{\mathrm{off},t} = \mathbb{E}\left[L_{\mathrm{off}}^{\mathrm{st}}\right] + L_{\mathrm{off},t}^{\mathrm{tr}} . \tag{5.21}$$

5.5 Improved Probabilistic Switching

In Sections 5.3 and 5.4 we presented methods to account for the startup dynamics and lockout constraints when providing PFC with an aggregation of refrigerators. To integrate these methods in the probabilistic switching approach of Section 5.2.3, it suffices to substitute the switching probability calculation of (5.8) with

$$\rho_t = \begin{cases} \dfrac{x_t}{1 - D_{t-1}^{\mathrm{a}} - L_{\mathrm{off},t-1}}, & \text{if } x_t \geq 0 \\[2mm] -\dfrac{x_t}{D_{t-1}^{\mathrm{a}} - L_{\mathrm{on},t-1}}, & \text{if } x_t < 0 . \end{cases} \tag{5.22}$$

Observe that the total fraction of loads that need to switch (x_t) is normalized by the fraction of the loads that are available to respond. If $x_t \geq 0$, the loads that can respond are those that are off and unlocked, i.e., the normalization factor is $1 - D_{t-1}^{\mathrm{a}} - L_{\mathrm{off},t-1}$. On the other hand, if $x_t < 0$, the loads that can respond are those that are on and unlocked, i.e., the normalization factor is $D_{t-1}^{\mathrm{a}} - L_{\mathrm{on},t-1}$.

The proposed controller is easy to implement at the device level. The required hardware is shown in Fig. 5.2 and includes a frequency meter and a micro-controller to control the switch of the compressor. The micro-controller consists of a random number generator and a floating-point unit able to perform elementary mathematical operations (additions, subtractions, multiplications and divisions). In addition, a link to the thermostat is established in order to modify its temperature limits while providing PFC; this is the topic of Section 5.7.
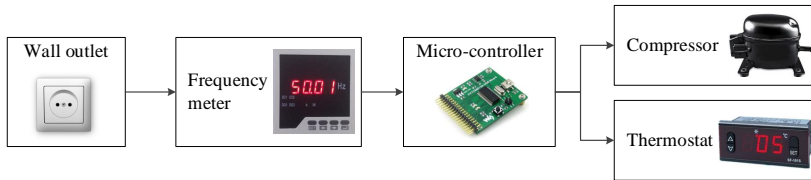
Figure 5.2: The required hardware at the device level to implement the proposed decentralized stochastic controller.

5.6　Robustness to Excessive Locking

The switching probability in (5.22) reflects the fraction of the population that needs to switch to provide PFC at each time step. The reserve will be provided reliably as long as there are enough refrigerators available to switch or, in other words, as long as $\rho_t \leq 1$. However, the controllability might be lost due to the lockout constraints. In the following sections, we identify the key parameters that contribute to the risk for loss of controllability, and propose countermeasures to build robustness to excessive refrigerator locking.

5.6.1　Risk for Loss of Controllability

In general, loss of controllability can occur in both directions (both when increasing and decreasing power). However, it is more likely to lose controllability in the positive PFC direction, i.e., while reducing the aggregate power, because the duty cycle is typically less than 0.5. In the rest of this section, we provide intuition on when controllability is lost without considering the startup dynamics for simplicity.

Consider a time step t when positive PFC is requested, i.e., $x_t < 0$. If startup dynamics are neglected, then $x_t = D_t^{\mathrm{d}} - D_{t-1}^{\mathrm{d}}$, and D_{t-1}^{a} is the same as D_{t-1}^{d}. Therefore, the lower branch of (5.22) can be written as

$$\rho_t = -\frac{D_t^{\mathrm{d}} - D_{t-1}^{\mathrm{d}}}{D_{t-1}^{\mathrm{d}} - L_{\mathrm{on},t-1}} \ . \tag{5.23}$$

The condition to lose controllability gives

$$\rho_t > 1 \quad \Leftrightarrow \quad D_t^{\mathrm{d}} - L_{\mathrm{on},t-1} < 0 \ . \tag{5.24}$$

For this to occur, D_t^{d} must take its minimum value, i.e., $D_t^{\mathrm{d}} = D^{\mathrm{n}} - D^{\mathrm{r}}$, which leads to[3]

$$D^{\mathrm{n}} - D^{\mathrm{r}} - L_{\mathrm{on},t-1} < 0 . \tag{5.25}$$

In addition, controllability is at risk when the fraction of devices locked in the on state ($L_{\mathrm{on},t-1}$) is high. Recall that $L_{\mathrm{on},t-1}$ consists of a constant steady-state part and a transient part that can be simplified in our case to

$$L_{\mathrm{on},t-1}^{\mathrm{tr}} = \sum_{k=0}^{t-1} c_k \cdot \left(D_k^{\mathrm{d}} - D_{k-1}^{\mathrm{d}} \right) \cdot S_{\mathrm{on}}(t-1-k) , \tag{5.26}$$

where c_k is defined in (5.19). Intuitively, $L_{\mathrm{on},t-1}^{\mathrm{tr}}$ is maximized when there exist frequent increases in duty cycle between two consecutive time steps k and $k+1$ within an interval $[0, t-1]$. For example, this is the case when the frequency oscillates for $k \in [0, t-1]$.

Consider a signal Δf_t that oscillates around zero with a magnitude δ and a period $2N_\delta$ seconds for $k \in [0, t-1]$, and an under-frequency event equal to $\Delta f_{\max}$ at time step t. Assuming that the oscillation is a triangle wave, the differential $df_k = \Delta f_k - \Delta f_{k-1}$ for each time step k of any increasing half-period of the wave is $df_k = \delta/N_\delta$. Consequently, $D_k^{\mathrm{d}} - D_{k-1}^{\mathrm{d}}$ is computed using (5.7) as

$$D_k^{\mathrm{d}} - D_{k-1}^{\mathrm{d}} = D^{\mathrm{r}} \cdot \frac{\Delta f_k - \Delta f_{k-1}}{\Delta f_{\max}} = \frac{D^{\mathrm{r}} \delta}{\Delta f_{\max} N_\delta} . \tag{5.27}$$

For the increasing half-periods of the wave $c_k = 1$ holds, whereas for the decreasing half-periods $c_k = 0$ holds. Using (5.27), equation (5.26) can be equivalently written as

$$L_{\mathrm{on},t-1}^{\mathrm{tr}} = \sum_{k=0}^{n_1} \sum_{j=n_2}^{n_3} \frac{D^{\mathrm{r}} \delta}{\Delta f_{\max} N_\delta} \cdot S_{\mathrm{on}}(t-1-j) , \tag{5.28}$$

---

[3]We can verify that the loss of controllability is more likely to occur in the positive PFC direction (while reducing the aggregate power) by comparing the condition (5.25) with the respective condition for the negative PFC direction, which can be obtained with similar arguments and is $1 - D^{\mathrm{n}} - D^{\mathrm{r}} - L_{\mathrm{off},t-1} < 0$. The risk for loss of controllability is higher to the positive direction if $D^{\mathrm{n}} - D^{\mathrm{r}} - L_{\mathrm{on},t-1} < 1 - D^{\mathrm{n}} - D^{\mathrm{r}} - L_{\mathrm{off},t-1}$, which is equivalent to $L_{\mathrm{off},t-1} - L_{\mathrm{on},t-1} < 1 - 2D^{\mathrm{n}}$. For typical values of the nominal duty cycle (around 0.25), typical lock-on times, and typical lock-off times, the last condition is satisfied. Note that the nominal duty cycle can be as high as 0.4 for parts of the day in the presence of door openings; in this case, the last condition might not hold, and therefore it is also likely to lose controllability to the negative PFC direction.

where $n_1 = (t-N_\delta-1)/2N_\delta$, $n_2 = 2N_\delta k$, and $n_3 = 2N_\delta k+N_\delta-1$. Combining (5.23), (5.21) and (5.28), the condition for loss of controllability for the considered signal Δf_t is

$$D^{\mathrm{n}} - D^{\mathrm{r}} - \mathbb{E}\left[L_{\mathrm{on}}^{\mathrm{st}}\right] - \frac{D^{\mathrm{r}}\delta}{\Delta f_{\max}N_\delta} \sum_{k=0}^{n_1} \sum_{j=n_2}^{n_3} S_{\mathrm{on}}(t-1-j) < 0. \quad (5.29)$$

From (5.29) one can see that controllability is at stake if (i) a large PFC reserve band is offered (if D^{r} is high), (ii) the magnitude of oscillation δ of Δf_t is high, (iii) the period of oscillation $2N_\delta$ of Δf_t is low, or (iv) the double summation is a large number, i.e., the maximum lock-on time N_{on} is high.[4]

Using the parameter values $D^{\mathrm{n}} = 0.25$, $D^{\mathrm{r}} = 0.12$, $\Delta f_{\max} = 0.2$ Hz, $\mathbb{E}\left[L_{\mathrm{on}}^{\mathrm{st}}\right] = 0.02$, and $N_{\mathrm{on}} = 80$ s in (5.29), we can calculate the critical magnitude of oscillation to lose controllability for different periods of oscillation.[5] For example, if $N_\delta = 1$ s the critical magnitude is $\delta = 6$ mHz, if $N_\delta = 10$ s we get $\delta = 57$ mHz, whereas if $N_\delta = 20$ s the critical magnitude is $\delta = 96$ mHz.

5.6.2   Countermeasures and Robustness

Frequency oscillations with the δ and N_δ values reported above did not exist in the frequency data set that we used. However, such oscillations might appear in power systems with lower inertia constants, for example, in systems with large shares of RES that are connected via power electronics. In this case, countermeasures are needed to ensure the controller's robustness. A potential solution would be to define three different operation states, namely "Normal", "Alert" and "Emergency", depending on how close the refrigerator aggregation is to loss of controllability.

The state in which the system is depends on the value of the expression at the left hand side of (5.29), which is called "controllability index". If the controllability index is above a positive threshold, then the system is in the Normal state and the refrigerators respond to frequency

---

[4]In practice, the width of PFC deadband is also important because no reserve is activated as long as Δf_t is within the deadband, and therefore no refrigerators are locked despite the oscillations in frequency.

[5]$\Delta f_{\max} = 0.2$ Hz is the frequency deviation for full PFC reserve activation in ENTSO-E, whereas the other parameter values are typical for refrigerator aggregations. More details on the parameters are given in Section 5.11.
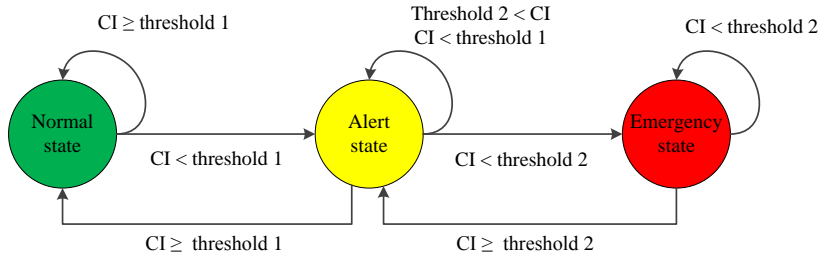
Figure 5.3: The three system states namely Normal, Alert and Emergency, depending on the controllability index (CI).

deviations according to (5.22). While being in the Normal state, the refrigerators also monitor the value of the controllability index.

If the controllability index gets smaller than the threshold, then each refrigerator assumes that the system has moved into the Alert state, and smoothes the local frequency deviation measurement with a filter before responding to it. This is expected to reduce the fraction of locked devices due to transient locking ($L_{\mathrm{on},t}$ and $L_{\mathrm{off},t}$). The refrigerators remain in the Alert state for as long as the controllability index is below the threshold, otherwise they return to the Normal state. In practice, a deadband around the threshold controllability index can be used to avoid frequent switching from the Normal to the Alert state.

On the other hand, if the controllability index continues to decrease and becomes smaller than a second threshold value, then the system moves to the Emergency state. In this state, it is critical to reduce the fraction of locked devices in order to avoid loss of controllability. This can be achieved if the micro-controller of each refrigerator forces its compressor to unlock. Of course, this might introduce wear on the compressor but it is a reasonable sacrifice if system stability is at stake. The three system states and the transition from one to the other is shown in Fig. 5.3.

5.7 Modification of Thermostat Limits

The probabilistic switching of Section 5.5 provides accurate PFC reserves immediately after a frequency deviation. However, if the frequency deviation remains, the accuracy of reserve provision will deteriorate because the aggregate power of the refrigerators will decay towards

the pre-disturbance value. In this section, we propose to modify the temperature limits of the refrigerators' thermostats to provide accurate reserves in case of prolonged frequency deviations.

5.7.1   Derivation of Temperature Resetting Factor

If the startup dynamics and lockout constraints are neglected, it is straightforward to verify that the refrigerators' aggregate power will decay towards the pre-disturbance value in case of a prolonged frequency deviation.

Let us denote by $s_{\text{off},t}$ and $s_{\text{on},t}$ the rates at which the refrigerators switch off and on at time step t, respectively. These rates can be approximated with

$$s_{\text{off},t} \approx D_t^{\text{d}} \cdot \dot{T}_{\text{d}}(\bar{T}), \; s_{\text{on},t} \approx (1 - D_t^{\text{d}}) \cdot \dot{T}_{\text{i}}(\bar{T}) \;, \qquad (5.30)$$

where $\dot{T}_{\text{d}}(\bar{T}) < 0$ and $\dot{T}_{\text{i}}(\bar{T}) > 0$ are the temperature decrease and increase rates evaluated at the mean temperature, respectively. $\dot{T}_{\text{d}}(\bar{T})$ can be obtained from (5.2) and $\dot{T}_{\text{i}}(\bar{T})$ from (5.1), and for the average parameter values used in this chapter, we get $\dot{T}_{\text{d}}(\bar{T}) = -0.0026°\text{C/s}$ and $\dot{T}_{\text{i}}(\bar{T}) = 0.0009°\text{C/s}$. Combining (5.30) with (5.7), the net switching rate can be expressed as

$$
\begin{aligned}
s_{\text{net},t} &= s_{\text{off},t} + s_{\text{on},t} \\
&\approx D^{\text{n}} \dot{T}_{\text{d}}(\bar{T}) + (1 - D^{\text{n}}) \dot{T}_{\text{i}}(\bar{T}) + D^{\text{r}} \frac{\Delta f_t}{\Delta f_{\text{max}}} \left[\dot{T}_{\text{d}}(\bar{T}) - \dot{T}_{\text{i}}(\bar{T}) \right] \\
&= s_{\text{off}} + s_{\text{on}} + D^{\text{r}} \frac{\Delta f_t}{\Delta f_{\text{max}}} \left[\dot{T}_{\text{d}}(\bar{T}) - \dot{T}_{\text{i}}(\bar{T}) \right] \\
&= D^{\text{r}} \frac{\Delta f_t}{\Delta f_{\text{max}}} \left[\dot{T}_{\text{d}}(\bar{T}) - \dot{T}_{\text{i}}(\bar{T}) \right] \;.
\end{aligned}
\qquad (5.31)
$$

In (5.31), $s_{\text{off}} = D^{\text{n}} \cdot \dot{T}_{\text{d}}(\bar{T})$ and $s_{\text{on}} = (1 - D^{\text{n}}) \cdot \dot{T}_{\text{i}}(\bar{T})$ are the switching rates at steady state, and $s_{\text{off}} + s_{\text{on}} = 0$ holds because the aggregate baseline power is approximately constant.

If $\Delta f_t > 0$, some refrigerators will switch on to provide PFC reserve, and shortly $s_{\text{net},t}$ will become negative because $\dot{T}_{\text{d}}(\bar{T}) - \dot{T}_{\text{i}}(\bar{T}) < 0$ in (5.31). This means that some refrigerators will start switching off, and therefore the aggregate power will start decreasing towards the baseline. A similar argument can be made for the case $\Delta f_t < 0$, when the initially lower aggregate power will start increasing towards the baseline.

As a result, the probabilistic switching approach of Section 5.5 cannot provide constant PFC reserve over time.

A simple solution to this problem is to dynamically modify the devices' thermostat limits (temperature deadband) in order to always keep $s_{\text{net},t}$ close to zero. For this purpose, we apply the temperature resetting factor

$$K_{\text{r},t}\ [^\circ\text{C}] = s_{\text{net},t} \cdot \Delta t = -D^{\text{r}}\Delta t \bar{\beta}\bar{P}_{\text{n}} \cdot \frac{\Delta f_t}{\Delta f_{\text{max}}}\ , \qquad (5.32)$$

where $\dot{T}_{\text{d}}(\bar{T}) - \dot{T}_{\text{i}}(\bar{T}) = -\bar{\beta}\bar{P}_{\text{n}}$ from (5.1) and (5.2). If $\Delta f_t > 0$, the resetting factor is negative, which means that both thermostat limits will decrease at a rate equal to $K_{\text{r},t}$ such that the population reaches an equilibrium at an aggregated power higher than the baseline. Similarly, if $\Delta f_t < 0$, both thermostat limits increase at a rate equal to $K_{\text{r},t}$.

If the frequency deviation is biased for a long period of time, applying (5.32) might change the refrigerator temperatures significantly. The implications of large changes in the average refrigerator temperature will be discussed in Section 5.8.

5.7.2   Inclusion of Startup Dynamics and Lockouts

To account for the startup dynamics and lockout constraints three modifications are needed in (5.32). The first modification is to substitute the term $D^{\text{r}} \cdot (\Delta f_t/\Delta f_{\text{max}})$ with x_t from (5.11) to account for the modified number of activated loads for PFC due to the startup dynamics.

A second modification is needed because the refrigerators that are locked (and cannot switch) do not contribute to the decay of the aggregate power towards the pre-disturbance value. To account for this, we incorporate the sequential unlocking of refrigerators after a frequency deviation event based on the CDFs of lockout times $F_{\text{on}}(t_{\text{on}}^l)$ and $F_{\text{off}}(t_{\text{off}}^l)$.

Consider a positive Δf_t starting at time step $t = 0$, which will induce a positive x_t. Instantaneously, some refrigerators will switch on to provide PFC and will lock in the on state. Therefore, at $t = 0$ fewer devices will be in the off state, $s_{\text{on},t}$ will start decreasing, and an instantaneous resetting factor of $-x_t\dot{T}_{\text{i}}(\bar{T})$ is needed to keep $s_{\text{net},t}$ close to zero.

However, as time elapses, some of the refrigerators that switched on will unlock based on $F_{\text{on}}(t_{\text{on}}^l)$, and larger changes in the temperature deadband limits will be required to maintain the same aggregate power.

Therefore, the resetting factor is determined by $x_t[\dot{T}_{\mathrm{d}}(\bar{T}) F_{\mathrm{on}}(t) - \dot{T}_{\mathrm{i}}(\bar{T})]\Delta t$ for $\Delta f_t > 0$, and by $x_t[\dot{T}_{\mathrm{d}}(\bar{T}) - \dot{T}_{\mathrm{i}}(\bar{T}) F_{\mathrm{off}}(t)]\Delta t$ for $\Delta f_t < 0$.

To generalize for multiple events, the resetting factor at time step t should accumulate all contributions from the previous time steps due to the recursive definition of x_t in (5.11). Therefore, the resetting factor is given by

$$K_{\mathrm{r},t} = \Delta t \cdot \sum_{k=0}^{t-1} x_k \cdot \big[c_k \cdot [\dot{T}_{\mathrm{d}}(\bar{T}) \cdot F_{\mathrm{on}}(t-k) - \dot{T}_{\mathrm{i}}(\bar{T})]$$
$$+ d_k \cdot [\dot{T}_{\mathrm{d}}(\bar{T}) - \dot{T}_{\mathrm{i}}(\bar{T}) \cdot F_{\mathrm{off}}(t-k)] \big] , \tag{5.33}$$

where c_k and d_k are defined in (5.19) and (5.20). Recall that F_{on}, F_{off} are CDFs, and so $F_{\mathrm{on}}(t-k) = 1$ for $t - k > N_{\mathrm{on}}$ and $F_{\mathrm{off}}(t-k) = 1$ for $t - k > N_{\mathrm{off}}$ in (5.35).

A last modification is needed because only the thermostat resetting of the unlocked devices affects the net switching rate during PFC reserve provision. For this reason, the temperature resetting is applied only to the unlocked refrigerators, and this is accounted for by including the normalization factor

$$r_t = \frac{1 - \mathbb{E}\left[L_{\mathrm{on}}^{\mathrm{st}}\right] - \mathbb{E}\left[L_{\mathrm{off}}^{\mathrm{st}}\right]}{1 - L_{\mathrm{on},t} - L_{\mathrm{off},t}} \tag{5.34}$$

in the temperature resetting factor, which is finally computed with

$$K_{\mathrm{r},t} = r_t \cdot \Delta t \cdot \sum_{k=0}^{t-1} x_k \cdot \big[c_k \cdot [\dot{T}_{\mathrm{d}}(\bar{T}) \cdot F_{\mathrm{on}}(t-k) - \dot{T}_{\mathrm{i}}(\bar{T})]$$
$$+ d_k \cdot [\dot{T}_{\mathrm{d}}(\bar{T}) - \dot{T}_{\mathrm{i}}(\bar{T}) \cdot F_{\mathrm{off}}(t-k)] \big] . \tag{5.35}$$

Note that $r_t > 1$ as long as $L_{\mathrm{on},t}^{\mathrm{tr}} \neq 0$ or $L_{\mathrm{off},t}^{\mathrm{tr}} \neq 0$, whereas $r_t = 1$ if $L_{\mathrm{on},t}^{\mathrm{tr}} = L_{\mathrm{off},t}^{\mathrm{tr}} = 0$, i.e., at steady state.

By modifying the thermostat limits using (5.35), we effectively transform the electric energy surplus or deficit due to the PFC reserve into thermal energy. In this way, we avoid the decay of aggregate power to the pre-disturbance value and achieve a constant reserve provision.

5.8 Robustness to Biased Frequency Deviations

If the frequency deviation is approximately zero-mean, the average refrigerator temperatures do not change significantly and the duty cycle

remains close to the nominal value D^{n}. However, in case of biased frequency deviations, the prolonged reduction or increase of the thermostat limits according to (5.35) results in unacceptable high or low refrigerator temperatures, which affect negatively the user utility.

Biased frequency deviations affect also the baseline power consumption. After a time interval with a biased frequency deviation, the refrigerators' thermostat limits, and consequently the average temperature of the population, will be different from the nominal values. However, the aggregate refrigerator power depends directly on the average temperature of the population. Specifically, if the average temperature decreases, the aggregate baseline power will increase due to higher thermal losses to the ambient, and vice versa. Therefore, biased frequency deviations will create steady-state deviations in the baseline power consumption. Since the activation of PFC reserve is measured with respect to the baseline, these steady-state deviations will increase the control error.

Without any reaction, the baseline deviation and control error will persist until the thermostat limits return to their nominal values, when the aggregation sustains an opposite bias in the frequency deviation. A simple approach to reduce the control error would be to filter the frequency deviation signal in order to eliminate the bias, as proposed in [94]. However, the disadvantage of this approach is that the bias needs to be absorbed by other available PFC resources or transferred to SFC. In this chapter, we propose a different approach that allows us to account for biased frequency deviations without relying on additional resources.

5.8.1 Corrective Temperature Control

Instead of relying on an opposite frequency deviation bias to eliminate the baseline deviation, we impose an opposite control action – called "corrective temperature control" – on the thermostat limits, such that they return to the nominal values in the long run. The corrective temperature control is designed as a proportional feedback control on the average temperature across the population using a gain K_{c}. Thus, the temperature limits of refrigerator i evolve in time according to

$$T_{\mathrm{min},i,t} = T_{\mathrm{min},i,t-1} + K_{\mathrm{r},t} - K_{\mathrm{c}} \cdot (\bar{T}_{t-1} - \bar{T}_{\mathrm{nom}}) \qquad (5.36)$$

$$T_{\mathrm{max},i,t} = T_{\mathrm{max},i,t-1} + K_{\mathrm{r},t} - K_{\mathrm{c}} \cdot (\bar{T}_{t-1} - \bar{T}_{\mathrm{nom}}) \ , \qquad (5.37)$$

where $\bar{T}_{t-1}$ is an estimate of the population's average temperature, and $\bar{T}_{\text{nom}}$ is the population's nominal mean temperature without PFC. Note that (5.36) and (5.37) are applied only to the unlocked refrigerators.

The goal of the corrective temperature control is to eliminate the persistent errors in activation of PFC due to steady-state baseline deviations resulting from biases in past frequency deviations. Recall that the temperature resetting factor $K_{\text{r},t}$ modifies the thermostat limits to improve the PFC accuracy. Although both the temperature resetting factor and the corrective temperature control modify the thermostat limits, the two control loops will not fight against each other, if the gains are appropriately selected. In particular, if K_{c} is sufficiently low, the corrective temperature control loop will act much slower than the resetting of thermostat limits with $K_{\text{r},t}$. In principle, this allows us to correct the deviations in the average temperature and the baseline consumption without compromising the accuracy of PFC. Tuning the gain of the corrective temperature control loop is the topic of Section 5.8.3.

5.8.2 Temperature and Duty Cycle Estimation

The mean temperature $\bar{T}_t$ depends on the past temperature resetting and corrective control actions, which in turn depend on the past frequency deviations, and can be estimated using

$$\bar{T}_t = \bar{T}_{t-1} + [K_{\text{r},t} - K_{\text{c}}(\bar{T}_{t-1} - \bar{T}_{\text{nom}})] \cdot (1 - L_{\text{on},t} - L_{\text{off},t}) . \quad (5.38)$$

The term $(1 - L_{\text{on},t} - L_{\text{off},t})$ in (5.38) is needed because the temperature resetting and corrective control actions are applied only to the unlocked devices, and it can be calculated from (5.21) based on $\mathbb{E}[t_{\text{on}}]$, $\mathbb{E}[t_{\text{off}}]$ and (5.18), (5.17). However, as the mean temperature changes, $\mathbb{E}[t_{\text{on}}]$ and $\mathbb{E}[t_{\text{off}}]$ change as well and can be approximated with

$$\mathbb{E}[t_{\text{on},t}] \approx \bar{R}\bar{C} \cdot \ln\left[\frac{\bar{T}_{\text{max},t-1} - \bar{T}_{\text{a}} + \bar{\eta}\bar{R}\bar{P}_{\text{n}}}{\bar{T}_{\text{min},t-1} - \bar{T}_{\text{a}} + \bar{\eta}\bar{R}\bar{P}_{\text{n}}}\right] \quad (5.39)$$

$$\mathbb{E}[t_{\text{off},t}] \approx \bar{R}\bar{C} \cdot \ln\left[\frac{\bar{T}_{\text{a}} - \bar{T}_{\text{min},t-1}}{\bar{T}_{\text{a}} - \bar{T}_{\text{max},t-1}}\right] , \quad (5.40)$$

where $\bar{R}, \bar{C}, \bar{\eta}, \bar{T}_{\text{a}}, \bar{P}_{\text{n}}$ are the expected values of the respective variables, and $\bar{T}_{\text{min},t-1} = \bar{T}_{t-1} - 0.5 \cdot \Delta T$, $\bar{T}_{\text{max},t-1} = \bar{T}_{t-1} + 0.5 \cdot \Delta T$, where ΔT is the population's mean (time-invariant) deadband width.

Using $\bar{T}_t$ it is also possible to keep track of the population's duty cycle with the approximation

$$\bar{D}_t^n \approx \frac{\mathbb{E}\left[t_{\text{on},t}\right]}{\mathbb{E}\left[t_{\text{on},t}\right] + \mathbb{E}\left[t_{\text{off},t}\right]} \ . \tag{5.41}$$

In this case, equation (5.12) that gives the actual duty cycle should be substituted with

$$D_t^a = D_{t-1}^a + x_t + (\bar{D}_t^n - \bar{D}_{t-1}^n) \ , \tag{5.42}$$

such that the fraction of loads activated for PFC (x_t) accounts for the change in baseline power.

5.8.3 Analytical Tuning of the Corrective Temperature Gain

From a user point of view, high K_c values are preferable to keep the refrigerator temperature deviations as low as possible. However, from a power system point of view, there is a tradeoff to consider when deciding the value of K_c. If the gain is very small, prolonged temperature deviations will change the aggregation's baseline consumption and introduce steady-state reserve errors. On the other hand, a very large gain might introduce oscillations in the aggregate power and consequently increased reserve errors. In fact, upper and lower bounds on the gain can be computed analytically in a simple case without startup dynamics and lockout constraints.

The correction of the average temperature with the gain K_c can be thought as a temperature resetting with the $K_{r,t}$ defined in (5.32), but for an opposite frequency deviation signal. Assuming that the application of the temperature resetting factor $K_{r,t}$ does not create oscillations in aggregate power, and if the opposite temperature change due to K_c is in absolute terms smaller than the average $\bar{K}_{r,t}$, then no oscillations in aggregate power should occur. Therefore, an upper bound on K_c can be obtained from

$$\left| K_c \cdot (\bar{T}_t - \bar{T}_{\text{nom}}) \right| \le \left| \bar{K}_{r,t} \right| \quad \Leftrightarrow$$

$$K_c \cdot |\bar{T}_t - \bar{T}_{\text{nom}}| \le \Delta t \bar{\beta} \bar{P}_n \left| \frac{D^r \bar{\Delta} f_t}{\Delta f_{\text{max}}} \right| = \Delta t \bar{\beta} \bar{P}_n \cdot |\bar{D}_t^n - D^n| \ , \tag{5.43}$$

where $\bar{\Delta}f_t$ is the average frequency deviation, and (5.7) was used.

From (5.41) we can see that $\bar{D}_t^{\mathrm{n}}$ is a function of the mean temperature $\bar{T}_t$. Therefore, the upper bound on K_c obtained from (5.43) is in general a function of $\bar{T}_t$ itself. However, a temperature-independent bound can be obtained by a first order Taylor expansion of (5.41) that gives

$$\bar{D}_t^{\mathrm{n}}(\bar{T}) \approx \bar{D}_t^{\mathrm{n}}\left(\bar{T}_{\mathrm{nom}}\right) + \frac{d\bar{D}_t^{\mathrm{n}}\left(\bar{T}\right)}{d\bar{T}}\bigg|_{\bar{T}_{\mathrm{nom}}} \cdot \left(\bar{T} - \bar{T}_{\mathrm{nom}}\right) \qquad \Leftrightarrow$$

$$\bar{D}_t^{\mathrm{n}} - D^{\mathrm{n}} = \bar{D}_t^{\mathrm{n}}(\bar{T}) - \bar{D}_t^{\mathrm{n}}(\bar{T}_{\mathrm{nom}}) \approx \frac{d\bar{D}_t^{\mathrm{n}}\left(\bar{T}\right)}{d\bar{T}}\bigg|_{\bar{T}_{\mathrm{nom}}} \cdot \left(\bar{T} - \bar{T}_{\mathrm{nom}}\right) . \quad (5.44)$$

Substituting now (5.44) in (5.43) gives us the upper bound

$$K_c \leq \left| \Delta t \bar{\beta} \bar{P}_{\mathrm{n}} \cdot \frac{d\bar{D}_t^{\mathrm{n}}\left(\bar{T}\right)}{d\bar{T}}\bigg|_{\bar{T}_{\mathrm{nom}}} \right| . \qquad (5.45)$$

A lower bound on K_c can be obtained based on the minimum/maximum acceptable mean temperature and the maximum acceptable settling time (i.e., the time needed to restore $\bar{T}_t$ sufficiently close to $\bar{T}_{\mathrm{nom}}$ after the biased frequency deviation event is over). Without lockout constraints ($L_{\mathrm{on},t} = L_{\mathrm{off},t} = 0$) equation (5.38) can be written as

$$\bar{T}_t = \lambda \cdot \bar{T}_{t-1} - \gamma \cdot \Delta f_t + (1 - \lambda) \cdot \bar{T}_{\mathrm{nom}} , \qquad (5.46)$$

where $\lambda = 1 - K_c$ and $\gamma = D^{\mathrm{r}} \Delta t \bar{\beta} \bar{P}_{\mathrm{n}} / \Delta f_{\mathrm{max}}$. With an initial condition $\bar{T}_0$, equation (5.46) has the solution

$$\bar{T}_t = \lambda^t \cdot \bar{T}_0 - \gamma \cdot \sum_{k=0}^{t-1} \lambda^k \cdot \Delta f_{t-k-1} + (1 - \lambda^t) \cdot \bar{T}_{\mathrm{nom}} . \qquad (5.47)$$

Assume that $\bar{T}_0 = \bar{T}_{\mathrm{nom}}$ and that the refrigerator aggregation faces the step frequency deviation

$$\Delta f_t = \begin{cases} \delta, & \text{if } 0 \leq t \leq N_{\mathrm{ev}} \\ 0, & \text{if } t > N_{\mathrm{ev}} \end{cases} , \qquad (5.48)$$

which represents a frequency deviation signal with a bias approximately equal to δ over a period N_{ev}. Let us denote by ε the tolerance in terms of the mean temperature during the frequency deviation event, i.e., $\bar{T}_t$ must satisfy $\bar{T}_{\mathrm{nom}} - \varepsilon \leq \bar{T}_t \leq \bar{T}_{\mathrm{nom}} + \varepsilon \; \forall t$. Furthermore, let us denote by

N_{rec} the maximum acceptable settling time, i.e, $\bar{T}_{\text{nom}} - \epsilon \leq \bar{T}_t \leq \bar{T}_{\text{nom}} + \epsilon$ should hold for $t \geq N_{\text{ev}} + N_{\text{rec}}$ after the frequency deviation event is over, where $\epsilon < \varepsilon$ is another tolerance. If δ, ε and ϵ are fixed, a lower bound on the corrective temperature gain can be computed using Proposition 2.

Proposition 2. *If $\bar{T}_0 = \bar{T}_{nom}$ and Δf_t is given by (5.48), then $\bar{T}_{nom} - \varepsilon \leq \bar{T}_t \leq \bar{T}_{nom} + \varepsilon$ $\forall t$ and $\bar{T}_{nom} - \epsilon \leq \bar{T}_t \leq \bar{T}_{nom} + \epsilon$ for $t \geq N_{ev} + N_{rec}$ hold if the corrective temperature gain is computed as $K_c = 1 - \lambda$, where λ satisfies*

$$\frac{1 - \lambda^{N_{ev}}}{1 - \lambda} \leq \frac{\varepsilon}{\gamma\delta} \tag{5.49}$$

$$\lambda^{N_{rec}} \cdot \frac{1 - \lambda^{N_{ev}}}{1 - \lambda} \leq \frac{\epsilon}{\gamma\delta} \ . \tag{5.50}$$

Proof. The proof is given in Appendix A. ∎

Remark 1. *Although (5.49) and (5.50) cannot be solved analytically, a lower bound on K_c can be obtained numerically by starting with a value of λ very close to 1, and gradually reducing it until it satisfies both (5.49) and (5.50).*

Equation (5.45), Proposition 2 and Remark 1 provide theoretical lower and upper bounds on K_c that can serve as an initial range when determining the final value of the gain via simulations with realistic Δf_t signals, startup dynamics, and lockout constraints.

5.9 Considering Thermostat Resolution

Typically, the necessary change in the thermostat's temperature limits calculated by (5.36) and (5.37) is very small, for example, 10^{-4} °C per second. Therefore, the temperature limits of some of the refrigerators will likely not change due to limited thermostat resolution or measurement noise, and so the aggregate power might decay towards its pre-disturbance value.[6]

We address this practical issue by allowing the temperature limits to change only by a sufficiently large fixed rate ΔT_{res} °C per second. This

---

[6]For example, the accuracy of temperature measurements in a refrigerator was ±0.2°C in the experimental work of [143].

is possible if we introduce an additional layer of randomization in the control algorithm. The underlying idea is simple: instead of distributing the temperature change over the whole unlocked population and requesting a small change from every device, we impose a much larger change on the temperature limits of fewer devices.

5.9.1   Probabilistic Implementation

At time step t, the required change in temperature limits of each refrigerator is

$$\Delta T_{\mathrm{lim},t} = K_{\mathrm{r},t} - K_{\mathrm{c}} \cdot (\bar{T}_{t-1} - \bar{T}_{\mathrm{nom}}) \ . \tag{5.51}$$

Equation (5.51) is applied by each refrigerator in a decentralized way using the locally obtained estimate of the average temperature.

If the temperature limits of an individual device change only by ΔT_{res}, we can get the same total change in temperature by applying ΔT_{res} to a fraction of the population equal to

$$\varrho_t = \frac{|\Delta T_{\mathrm{lim},t}|}{\Delta T_{\mathrm{res}}} \ . \tag{5.52}$$

The implementation of this approach is straightforward: each refrigerator draws a random number uniformly distributed between 0 and 1, and if it is smaller than ϱ_t and the refrigerator is unlocked, then it modifies its temperature limits by ΔT_{res}, otherwise it does not. The limits increase by ΔT_{res} if $\Delta T_{\mathrm{lim},t} > 0$, and decrease by ΔT_{res} if $\Delta T_{\mathrm{lim},t} < 0$.

5.9.2   Drawbacks and Countermeasure

Despite its simplicity, this approach introduces two sources of inaccuracy in the control. First, as ΔT_{res} increases, fewer devices are required to change their thermostat limits and therefore the control is exposed to inaccuracies from random number generation. Second, large ΔT_{res} values might synchronize the population. Without the fixed rate ΔT_{res}, the temperature limits change in a continuous fashion, the devices switch when they are very close to the thermostat limits, and this avoids synchronization. In contrast, if a large ΔT_{res} value is used, some devices will switch at a temperature considerably different from the nominal thermostat limits.

If a fixed rate ΔT_{res} is used, the change in temperature limits $\Delta T_{\lim,i,t}$ is different for each refrigerator i. Another disadvantage of this approach is that the variance of $\Delta T_{\lim,i,t}$ increases monotonically, as shown by Proposition 3 and Remark 2.

Proposition 3. *If the refrigerators' temperature limits change probabilistically according to (5.52), then the mean and variance of $\Delta T_{lim,i,t}$ are given by*

$$\mathbb{E}\left[\Delta T_{lim,i,t}\right] = \sum_{k=0}^{t-1} \Delta T_{lim,k} \tag{5.53}$$

$$\text{Var}\left[\Delta T_{lim,i,t}\right] = \sum_{k=0}^{t-1} |\Delta T_{lim,k}| \cdot \left(\Delta T_{res} - |\Delta T_{lim,k}|\right). \tag{5.54}$$

Proof. The proof is given in Appendix A. ∎

Remark 2. *If the fixed rate ΔT_{res} is not applied, all refrigerators change their temperature limits by $\sum_{k=0}^{t-1} \Delta T_{lim,k}$, and so the mean value is equal to $\sum_{k=0}^{t-1} \Delta T_{lim,k}$ and the variance is zero. Note that in our simulations, the mean value is slightly different and the variance is non-zero but small, because the thermostat limit changes are applied only to the unlocked devices.*

Equation (5.53) shows that using the fixed rate ΔT_{res} does not affect the mean value of the temperature limit change. In contrast, the variance monotonically increases over time (linearly with ΔT_{res}) because typically $\Delta T_{res} - |\Delta T_{lim,k}| \geq 0$ in (5.54).

Due to the monotonic increase in variance, some devices will likely sustain larger and prolonged deviations from the nominal temperature limits, which is undesirable. The variance can be bounded by imposing a restriction on the maximum deviation from the mean temperature of the population. To implement this, we require each unlocked device i that drew a random number smaller than ϱ_t at time step t to change its limits only if

$$\bar{T}_{t-1} - \Delta T_{\text{b}} \leq \Delta T_{\lim,i,t-1} + \zeta_i \Delta T_{\text{res}} \leq \bar{T}_{t-1} + \Delta T_{\text{b}} \tag{5.55}$$

holds, where ΔT_{b} is a temperature deviation bound, and ζ_i is equal to 1 or -1 depending on the sign of $\Delta T_{\lim,t}$.
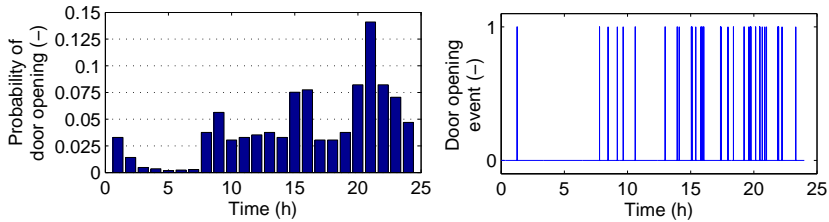
Figure 5.4: Left: The histogram of door openings during a day. Right:
Door opening events for a refrigerator generated based on the
histogram.

5.10   Accounting for Door Openings

The analysis so far neglected the door openings that significantly affect
the operation not only of individual refrigerators, but also of the aggre-
gation as a whole. In this section, we present a method to model door
openings and account for them in PFC design.

5.10.1   Stochastic Model for Door Opening Events

There exist only a few papers that consider door openings in power
system studies with refrigerators [89, 104]. In this chapter, we assume
that the number of door openings per day follows a normal distribution
with mean value μ_{op} and standard deviation σ_{op}. Furthermore, we
assume that the duration of a door opening event follows a normal
distribution with mean value μ_{d} and standard deviation σ_{d}. The door
opening events are distributed within the day based on the hourly profile
from [144], which is shown in the left plot of Fig. 5.4. Based on [144]
we select $\mu_{\mathrm{op}} = 40$, whereas we fix $\sigma_{\mathrm{op}} = 5$, $\mu_{\mathrm{d}} = 20$ s and $\sigma_{\mathrm{d}} = 3$ s [89].
The right plot of Fig. 5.4 shows a time-series example of door opening
events for a refrigerator, which is generated using the probability profile
of the left plot.

5.10.2   Modeling the Effect of Door Openings

The door openings increase the refrigerator's daily energy consumption
by $\xi \cdot 100\%$ (we assume $\xi = 0.22$ [144]). The effect of door openings
on refrigerator temperature can be modeled by reducing the thermal

resistance during each door opening event in (5.1) and (5.2), i.e., by increasing the thermal losses to the ambient. If μ_{op}, μ_{d} and ξ are known, then an upper bound to the thermal resistance can be obtained as shown by Proposition 4.

Proposition 4. *Recall that the refrigerator's thermal resistance with closed door is denoted by R, and let us denote by $N^d = 86,400$ seconds the duration of a day. An upper bound to the estimate of the thermal resistance R_{op} during a door opening event can be obtained from*

$$R_{op} \leq R \cdot \frac{1}{1 + \frac{N^d}{\mu_{op} \cdot \mu_d} \cdot \xi} \ . \tag{5.56}$$

Proof. The proof is given in Appendix A. ∎

Remark 3. *Evaluating (5.56) as an equality gives us an initial estimate of R_{op}. For the parameters assumed in this chapter, (5.56) gives $R_{op} \leq R/24.76$, and so we used the value $R_{op} = R/25$ in our simulations.*

5.10.3 Modifications in PFC Design

The proposed controller can be easily extended to account for door openings. Even without PFC, the aggregation's duty cycle and baseline power are not constant any more, but they depend on the distribution of door openings within the day. The duty cycle without PFC can be expressed as

$$D_t^{\mathrm{n}} = D^{\mathrm{n}} + \Delta D_t^{\mathrm{op}} \ , \tag{5.57}$$

where D^{n} is the nominal duty cycle and $\Delta D_t^{\mathrm{op}} \geq 0$ is the additional duty cycle due to door openings. In practice, a smoothed version of ΔD_t^{op} can be obtained by comparing historical data of aggregate power consumption with and without door openings.

If the duty cycle is now estimated using

$$\bar{D}_t^{\mathrm{n}} \approx \frac{\mathrm{E}\left[t_{\mathrm{on},t}\right]}{\mathrm{E}\left[t_{\mathrm{on},t}\right] + \mathrm{E}\left[t_{\mathrm{off},t}\right]} + \Delta D_t^{\mathrm{op}} \tag{5.58}$$

instead of (5.41), then the necessary information about the door openings is passed to the switching probability calculation (5.22) through (5.42), and no other modification in the control design is needed.
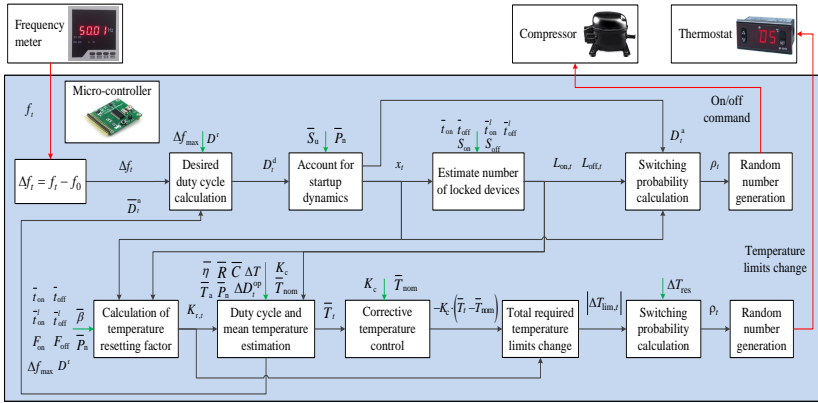
Figure 5.5: The proposed decentralized stochastic controller. Red arrows
indicate the controller's input/output signals, green arrows in-
dicate input from stored parameter values in the controller's
memory, and black arrows indicate internal control and feed-
back signals.

Figure 5.5 shows the control block of the decentralized stochastic con-
troller including all components presented so far. In the rest of this
chapter, we demonstrate the controller's performance in detailed simu-
lation studies.

5.11   Parameters and Performance Metrics

In this section, we present the refrigerator parameters and introduce
the metrics and benchmarks that we use to quantify the performance
of the proposed controller. Note that the thermal parameter β and the
lock-off time duration t_{off}^l do not follow a standard Probability Distri-
bution Function (PDF) because they are post-calculated based on the
imposed PDFs on other variables. The resulting PDFs resemble nor-
mal distributions with the mean value and standard deviation shown in
Table 5.7, but with positive skewness and kurtosis.

We fix $D^{\text{r}} = 0.15$ and the PFC reserve capacity is given by

$$P_{\text{res}} = N_{\text{r}} \bar{P}_{\text{n}} D^{\text{r}} . \tag{5.59}$$

Table 5.7: Refrigeration aggregation parameters

Parameter	Value	Parameter	Value
T_a (°C)	$\mathbb{U}[20, 24]$	α (s^{-1})	$\mathbb{U}[4, 6] \cdot 10^{-5}$
ΔT (°C)	$\mathbb{U}[1.7, 2.3]$	β (°C/kJ)	$(\mu, \sigma) = (4.4, 0.7) \cdot 10^{-5}$
T_set (°C)	$\mathbb{U}[4.5, 5.5]$	P_n (W)	$\mathbb{U}[70, 90]$
u (-)	$\mathbb{N}[0.25, 0.025]$	t_on^l (s)	$\mathbb{N}[60, 5]$
N_s (s)	$\mathbb{N}[30, 3]$	t_off^l (s)	$(\mu, \sigma) = (189, 31.5)$

The desired aggregate power at time step t with PFC is

$$P_{\mathrm{d},t} = \bar{P}_{\mathrm{b},t} + P_\mathrm{res} \cdot (\Delta f_t / \Delta f_\mathrm{max}) \ , \tag{5.60}$$

where $\bar{P}_{\mathrm{b},t}$ is the population's baseline power, i.e., a smoothed version of historical measurements of the aggregate power $P_{\mathrm{b},t}$ without PFC. If there are no door openings, then the baseline is constant and equal to $\bar{P}_\mathrm{b} = N_\mathrm{r} \bar{P}_\mathrm{n} D^\mathrm{n}$, whereas it is time-varying in the presence of door openings.

We define the control error as $P_{\mathrm{d},t} - P_{\mathrm{agg},t}$, where $P_{\mathrm{agg},t}$ is the actual aggregate power. We use the instantaneous percentage reserve error (5.61), the reserve Mean Absolute Percentage Error (MAPE) defined in (5.62), and the tracking MAPE defined in (5.63) as control performance metrics. In addition, we use the baseline MAPE (5.64) as a metric of the natural oscillations in an uncontrolled aggregation of refrigerators. Denoting by N_sim the simulation period, the metrics can be expressed as

$$e_{\mathrm{r},t} = 100 \cdot (P_{\mathrm{d},t} - P_{\mathrm{agg},t}) / P_\mathrm{res} \tag{5.61}$$

$$e_{\mathrm{r,mape}} = (100/N_\mathrm{sim}) \cdot \sum_{t=0}^{N_\mathrm{sim}-1} |e_{\mathrm{r},t}| \tag{5.62}$$

$$e_{\mathrm{t,mape}} = (100/N_\mathrm{sim}) \cdot \sum_{t=0}^{N_\mathrm{sim}-1} |(P_{\mathrm{d},t} - P_{\mathrm{agg},t})/P_{\mathrm{d},t}| \tag{5.63}$$

$$e_{\mathrm{b,mape}} = (100/N_\mathrm{sim}) \cdot \sum_{t=0}^{N_\mathrm{sim}-1} \left|(\bar{P}_{\mathrm{b},t} - P_{\mathrm{b},t})/P_\mathrm{res}\right| . \tag{5.64}$$

In the rest of this chapter, we demonstrate the performance of the proposed decentralized control scheme, which we call *"proposed controller"* hereafter, and benchmark it against two simpler controllers. The *"simple controller 1"* neglects the startup dynamics and lockout constraints, and is similar to the approach presented in [94]. The only difference

is that the temperature resetting factor $K_{\mathrm{r},t}$ is computed analytically from (5.32) and not heuristically as in [94]. The *"simple controller 2"* neglects the startup dynamics and lockout constraints, and additionally does not reset the thermostat limits. We compare the controllers via detailed simulations with actual frequency deviation data from the control area of Switzerland in 2009 and 2011.

5.12   Benchmarking and Aggregation Size

The control performance is expected to improve as the size of the refrigerator aggregation increases for three main reasons. First, the controller relies on probabilistic switching using random number generators, and thus the larger the aggregation the closer the resulting switching probabilities are to the desired values. The second reason is that the controller assumes a constant uncontrolled duty cycle D^{n} in the absence of door openings (which is increased by a smooth power profile ΔD_t^{op} to account for door openings). The assumption that D^{n} is constant (or that $D^{\mathrm{n}} + \Delta D_t^{\mathrm{op}}$ is time-varying but smooth) is reasonable only for sufficiently large aggregations, where the oscillations from individual load cycles cancel out. The third reason is that the controller is designed based on the average refrigerator parameters and, obviously, the larger the population the better it is described by these average values.

5.12.1   Reserve Error vs Aggregation Size

In this section, we investigate the dependence of the controller's performance on the aggregation size in simulations with different sizes using 30 five-hour samples of the frequency deviation signal. The simulations are performed with the proposed controller and the simple controller 1. Since the simple controller 1 is not robust to biased frequency deviations, we use only zero-mean frequency deviation signals in these simulations.

The left plot of Fig. 5.6 shows the reserve MAPE ($e_{\mathrm{r,mape}}$) for the proposed and the simple controller 1. A population of $1,000$ refrigerators results in a large reserve error, whereas increasing the size up to $10,000$ refrigerators drastically decreases the error. After this point the reserve error decreases asymptotically and it practically saturates at an aggregation size of $70,000$. For this reason, all simulations in Sections 5.13-5.16 were performed with a population of $70,000$ refrigerators. Observe
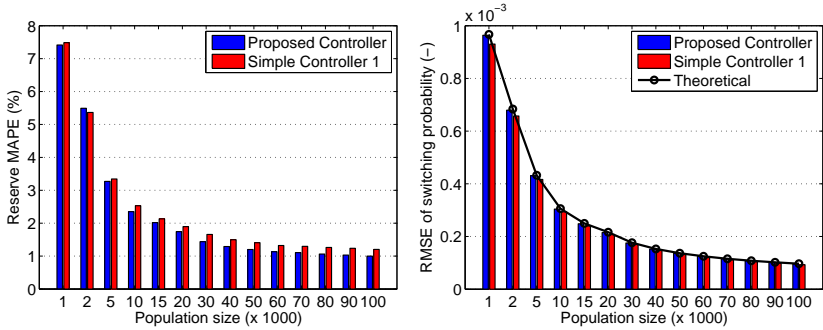
Figure 5.6: Dependence of control performance on the size of refrigerator aggregation for the proposed and the simple controller 1 using zero-mean frequency deviation signals. Left: reserve MAPE. Right: RMSE of switching probability (theoretical values and values obtained by simulations).

that the proposed controller outperforms the simple controller 1 for all aggregation sizes with the exception of a small aggregation of $2,000$ refrigerators. For populations of $70,000$ loads or more, the proposed controller reduces the reserve error by approximately 16% compared with the simple controller 1.

There are two error components hidden behind the reserve error: (i) the switching probability error, and (ii) the baseline error. In the rest of this section, we attempt to disaggregate the total reserve error into these two components.

5.12.2 Switching Probability Error vs Aggregation Size

The right plot of Fig. 5.6 shows the Root Mean Squared Error (RMSE) of switching probability ρ_t for the proposed controller and the simple controller 1, which is calculated with

$$e_{\mathrm{p,rmse}} = \sqrt{\frac{1}{N_{\mathrm{sim}}} \cdot \sum_{t=0}^{N_{\mathrm{sim}}-1} \left(\frac{\sum_{i=1}^{N_{\mathrm{r}}} \chi_{i,t}}{N_{\mathrm{r}}} - \rho_t \right)^2}. \qquad (5.65)$$

The switching action of each device i at time step t is a binary random variable $\chi_{i,t}$ with a Bernoulli distribution. Therefore, the term

$1/N_{\mathrm{r}} \cdot \sum_{i=1}^{N_{\mathrm{r}}} \chi_{i,t}$ denotes the percentage of loads that switch at each time step. Note that the success probability of $\chi_{i,t}$ is time-varying and equal to the switching probability ρ_t calculated with (5.22).

The dependence of the switching probability error on the aggregation size N_{r} can be derived analytically by making the simplifying assumption that the population applies the average value of the switching probability $\bar{\rho}$ for the whole simulation horizon. In this case, $\chi_{i,t}$ will have a constant success probability $p_{\mathrm{s}} = \bar{\rho}$ and a standard deviation $\sigma_{\chi_i} = \sqrt{\bar{p}_{\mathrm{s}} \cdot (1 - \bar{p}_{\mathrm{s}})} = \sqrt{\bar{\rho} \cdot (1 - \bar{\rho})}$. Therefore, the switching probability RMSE can be written as

$$e_{\mathrm{p,rmse}}\left(N_{\mathrm{r}}\right) = \sqrt{\frac{1}{N_{\mathrm{sim}}} \cdot \sum_{t=0}^{N_{\mathrm{sim}}-1} \left(\frac{\sum_{i=1}^{N_{\mathrm{r}}} \chi_i}{N_{\mathrm{r}}} - \bar{\rho}\right)^2} = \qquad (5.66\mathrm{a})$$

$$\mathrm{SD}\left(\frac{\sum_{i=1}^{N_{\mathrm{r}}} \chi_i}{N_{\mathrm{r}}}\right) = \mathrm{SD}\left(\bar{\chi}_i\right) = \frac{\sigma_{\chi_i}}{\sqrt{N_{\mathrm{r}}}} = \sqrt{\frac{\bar{\rho}(1 - \bar{\rho})}{N_{\mathrm{r}}}} \;, \qquad (5.66\mathrm{b})$$

where we used the square root law in (5.66b) and SD denotes the standard deviation of the sample mean $\bar{\chi}_i$.

According to (5.66), the switching probability error decreases with the square root of aggregation size. We calculate $e_{\mathrm{p,rmse}}$ for all considered population sizes using (5.66) and plot the theoretically expected curve in the right plot of Fig. 5.6, which is in agreement with the simulation results (bar plot) obtained with (5.65).

5.12.3   Baseline Error vs Aggregation Size

The left plot of Fig. 5.7 shows the reserve MAPE ($e_{\mathrm{r,mape}}$) together with the baseline MAPE ($e_{\mathrm{b,mape}}$). The baseline error decreases as the aggregation size increases because the uncontrolled aggregate power and duty cycle become smoother. This is graphically illustrated in the right plot of Fig. 5.7, where the duty cycle of a population of $1,000$ and $100,000$ refrigerators is shown. Observe that for aggregations up to $15,000$ loads the baseline error is dominant and sometimes even larger than the reserve error. This means that the PFC response of small aggregations is covered by the large natural oscillations of the uncontrolled baseline. Interestingly, even for an aggregation of $100,000$ loads the baseline error comprises a significant part (approximately 62%) of the reserve error. In other words, the control inaccuracy due to probabilistic switching contributes to only 38% of the reserve MAPE.
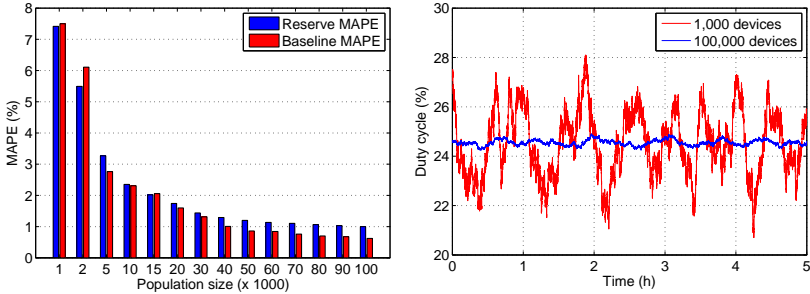
Figure 5.7: Left: comparison of the reserve and baseline MAPE of the proposed controller for different aggregation sizes. Right: The uncontrolled duty cycle of $1,000$ and $100,000$ refrigerators.

5.13 Sensitivity Analysis

In this section we investigate the sensitivity of the controller's performance to four important parameters: (i) the reserve capacity, (ii) the peak power during the startup phase, (iii) the lockout duration, and (iv) the PFC activation deadband. For this purpose, we run simulations over a period of 10 days for different values of these parameters. The frequency deviation is approximately zero-mean during the first 5 days of the data set, whereas it has a large bias during the remaining 5 days (up to 0.011 Hz).

5.13.1 Sensitivity to Reserve Capacity

Figures 5.8 and 5.9 show the sensitivity analysis results with respect to the reserve capacity D^{r}. Figure 5.8 corresponds to the part of the frequency deviation signal with small bias, whereas Fig. 5.9 is for the signal with large bias. In both figures, the left plot shows the reserve MAPE ($e_{\mathrm{r,MAPE}}$), whereas the right plot shows the tracking MAPE ($e_{\mathrm{t,MAPE}}$).[7]

A first observation is that as D^{r} increases, $e_{\mathrm{t,MAPE}}$ increases but $e_{\mathrm{r,MAPE}}$ decreases. This means that as the reserve capacity increases the abso-

---

[7]The difference between the proposed controller and the simple controller 1 is larger in Fig. 5.8 compared with Fig. 5.6. The reason is that there exists some small bias in the signal used for the simulations of Fig. 5.8, but there exists no bias in the signal used for the simulations of Fig. 5.6.
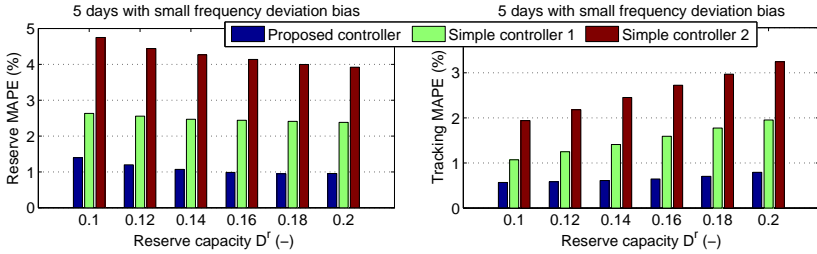
Figure 5.8: Sensitivity to reserve capacity in simulations with a small frequency deviation bias. Left: Reserve MAPE. Right: Tracking MAPE.
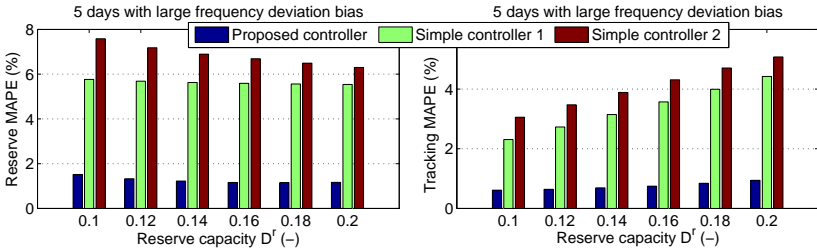
Figure 5.9: Sensitivity to reserve capacity in simulations with a large frequency deviation bias. Left: Reserve MAPE. Right: Tracking MAPE.

lute tracking error increases, but the relative error decreases. A second observation is that the proposed controller results in lower errors compared with the simple controller 1 for all cases. In addition, the simple controller 2 performs worse than the simple controller 1, which illustrates the need for resetting of the thermostat temperature limits, as explained in Section 5.7.1.

Observe that the difference between the proposed controller and simple controller 1 is much more pronounced for the biased frequency deviation signal. This shows the effectiveness of the corrective temperature control of Section 5.8. In fact, there is only a small difference between the simple controller 1 and simple controller 2 for biased frequency deviations. In this case, most of the reserve error is attributed to steady-state baseline errors, and so the improvement of the temperature resetting action is only marginal.
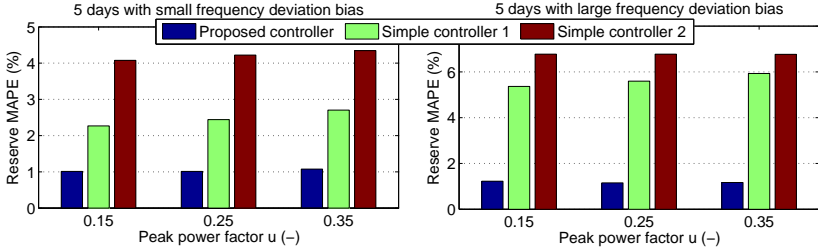
Figure 5.10: Sensitivity of reserve MAPE to the peak power during startup dynamics. Left: small frequency deviation bias. Right: large frequency deviation bias.
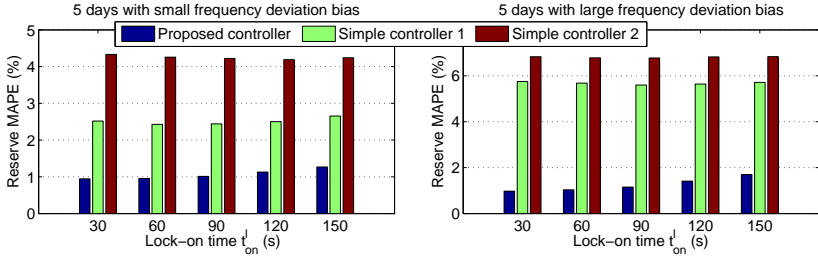
Figure 5.11: Sensitivity of reserve MAPE to the compressor's lock-on time. Left: small frequency deviation bias. Right: large frequency deviation bias.

5.13.2 Sensitivity to Startup Dynamics' Peak Power

The dependence of $e_{\mathrm{r,MAPE}}$ on the peak power factor during the startup phase (u) is presented in Fig. 5.10. The error of the proposed controller is not sensitive to u, whereas the error of the simple controller 1 clearly increases as u increases. This is expected because a larger u results in larger overshoots in the aggregate power consumption, if the effect of startup dynamics is not considered in the control design.

5.13.3 Sensitivity to Lockout Times

Figure 5.11 shows the dependence of the reserve error to the lock-on time t_{on}^l. The error of the proposed controller steadily increases as t_{on}^l increases, which indicates that the estimate of the number of locked
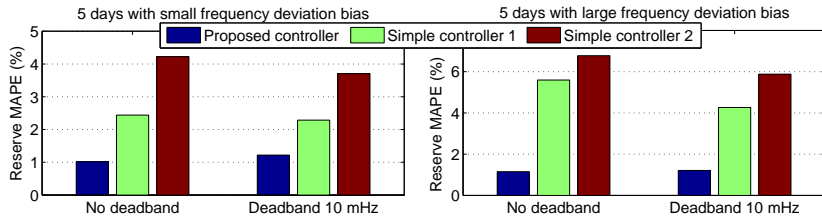
Figure 5.12: Reserve MAPE with and without a PFC activation deadband of ±10 mHz. Left: small frequency deviation bias. Right: large frequency deviation bias.

devices – obtained by (5.21) – deteriorates.

On the other hand, the error of the simple controller 1 is unaffected by t_{on}^l, and it remains constant at a level significant higher than that of the proposed controller. This can be explained because the simple controller 1 neglects both the lockout times and startup dynamics. With reference to (5.22), the simple controller 1 overestimates the value of the denominator (because it neglects $L_{\mathrm{off},t-1}$ and $L_{\mathrm{on},t-1}$) and so fewer loads than needed are activated for PFC. However, each of the activated loads consumes more power than expected due to the startup dynamics, which counteracts the error in the number of activated loads.

5.13.4   Effect of PFC Activation Deadband

In practice, a frequency deadband is used to avoid unnecessary PFC activation and to prevent PFC activation to the wrong direction due to frequency measurement error. For example, a deadband of ±10 mHz is used in the ENTSO-E system [145]. To investigate the effect of frequency deadband on the performance of the proposed controller, we repeated the simulations for our base case scenario with a deadband of ±10 mHz and we present the results in Fig. 5.12.

Even with a deadband, the proposed controller significantly reduces the reserve error compared with the benchmark controller for both zero-mean and biased frequency deviation signals. However, the difference between the controllers is smaller if a deadband is used. Note that when using a frequency deadband of ±10 mHz, the PFC is active only for approximately 58% of the time in our simulations. Figure 5.13 shows the desired and activated reserve for the proposed controller and simple
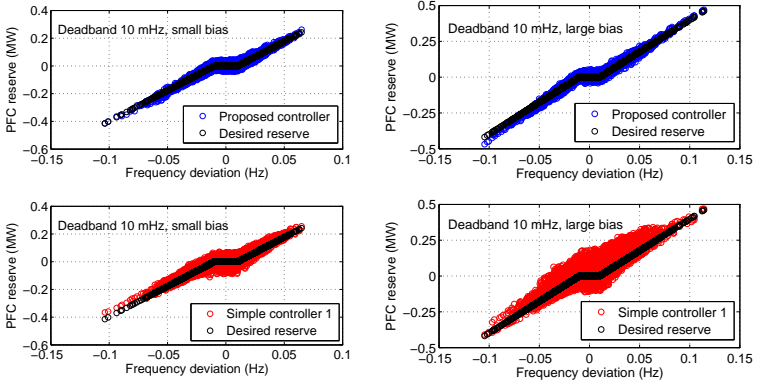
Figure 5.13: Droop characteristic with a PFC activation deadband of ± 10 mHz. Left: Small frequency deviation bias. Right: Large frequency deviation bias.

controller 1 with a PFC deadband for zero-mean and biased frequency deviations. The proposed controller follows the desired droop characteristic of PFC much closer than simple controller 1, in particular for biased frequency deviations.

5.14 Performance with Biased Frequency Deviations

In this section, we present results on tuning of the corrective temperature gain K_c via 10 day-long simulations. Figure 5.14 compares the reserve error $e_{r,t}$ and the temperature deviations for three different K_c values, as well as for a case without corrective control. The results correspond to simulations with a frequency deviation signal that is positively biased for the first 15 hours (with an average bias of $\delta = 0.0192$ Hz), and zero-mean for the rest of the day.

If $K_c = 0$, there is a steady-state error both in the temperature and reserve provision at the end of the day due to a permanent baseline error. On the other hand, all three non-zero gains bring the mean temperature close to the nominal value of 5°C and eliminate the steady-state reserve error. As expected, higher K_c values allow a faster temperature recovery but introduce oscillations in the baseline power, which translate into
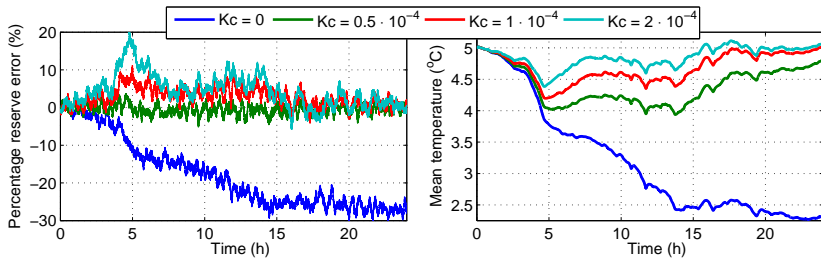
Figure 5.14: Time-domain simulation results of reserve error and mean temperature deviation for four different values of the corrective temperature gain.
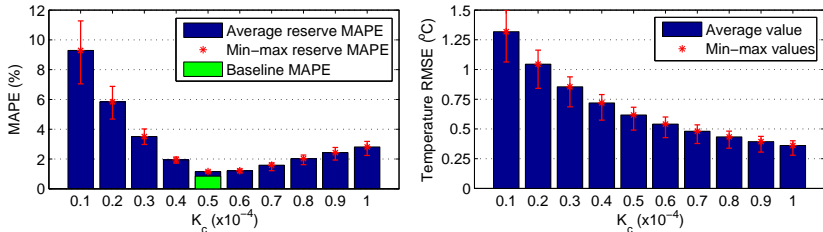
Figure 5.15: Dependence of reserve MAPE and temperature RMSE on the corrective temperature gain.

increased reserve errors at the beginning of the recovery period (around hour 5 in Fig. 5.14).

Results over the 10 day-long simulations for K_c in the range $[0.1, 1] \cdot 10^{-4}$ are presented in Fig. 5.15. The frequency deviation is positively biased in 4 of the days, negatively biased in another 4 of the days, and zero-mean in the remaining 2 days. The blue bars show the average values of the reserve MAPE and temperature RMSE, whereas the red lines indicate the range spanned from the minimum to the maximum value. In accordance with our intuition in Section 5.8.3, the reserve MAPE demonstrates a convex dependence on K_c and the minimum is achieved for $K_c = 0.5 \cdot 10^{-4}$. This gain results in a reserve MAPE of 1.15%, which is very close to the baseline MAPE of 0.85% due to the system's natural dynamics (shown with the green bar in Fig. 5.15). Of course, from a user point of view, the larger the temperature corrective gain is, the smaller the temperature deviations are.
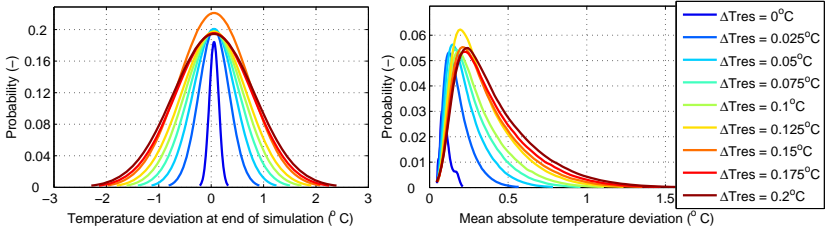
Figure 5.16: Left: The PDF of temperature deviations at the end of the simulation. Right: The PDF of mean absolute temperature deviations during the simulation.

Note that the results of Fig. 5.14 are in agreement with Proposition 2. With $\delta = 0.0192$ Hz, $N_{ev} = 15$ h, $N_{rec} = 9$ h, $\varepsilon = 1°$C, $\epsilon = 0.2°$C, $D^r = 0.15$, $\bar{P}_n = 80$ W and $\bar{\beta} = 4.4 \cdot 10^{-5}$, the lower bound on K_c that satisfies (5.49) and (5.50) is $K_c = 0.4863 \cdot 10^{-4}$. From Fig. 5.14 we can see that $K_c = 0.5 \cdot 10^{-4}$ is consistent with the design criteria: (i) the minimum temperature is approximately 4°C ($\varepsilon = 1°$C); and (ii) at the end of the day ($N_{rec} = 9$ h after the frequency bias has disappeared) the mean temperature is less than $\epsilon = 0.2°$C away from the nominal value of 5°C.

In addition, the results of Fig. 5.15 are in agreement with (5.45). For the considered values $\bar{\Delta}T = 2°$C, $\bar{T}_a = 22°$C, $\bar{\eta}\bar{R}\bar{P}_n = 70°$C and $\Delta t = 1$ s, (5.45) provides us with the upper bound $K_c = 0.5004 \cdot 10^{-4}$. Based on our simulations results, the theoretical lower and upper bounds on the gain of the corrective temperature control loop are very tight.

5.15 Limited Thermostat Resolution

To investigate the effect of minimum thermostat resolution ΔT_{res} on the temperature deviations and control performance, we perform 10 day-long simulations with 9 different ΔT_{res} values and present the results in Fig. 5.16. The left plot shows the PDFs of temperature deviations at the end of the simulation. The PDFs are fitted based on the histogram of observations and they match very closely normal distributions. The mean temperature deviation is approximately the same for all values of ΔT_{res}, whereas the variance increases as ΔT_{res} increases. This means that larger deviations are expected for higher ΔT_{res} values.
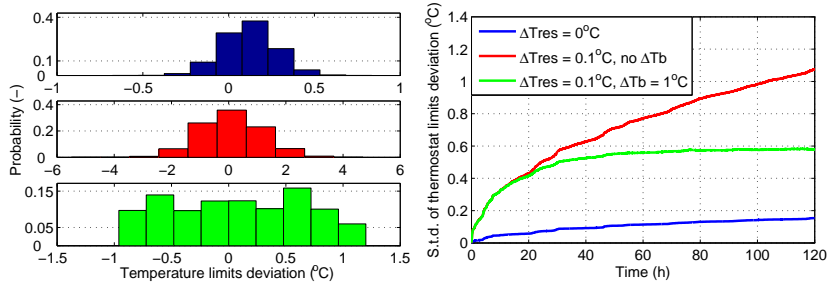
Figure 5.17: The effect of limited thermostat resolution on the temperature
deviations, with and without a temperature deviation bound.
Left: The PDFs of temperature deviations at the end of the
simulation. Right: The evolution of the standard deviation of
temperature deviations during the simulation.

The right plot of Fig. 5.16 shows the PDFs of the mean absolute temperature deviations during the simulation period. In this case, apart from the variance, the mean value of the distribution also increases for higher ΔT_{res}. This is expected because (i) the absolute temperature deviation ψ follows a half-normal distribution (since the temperature deviation ω follows a normal distribution) with a mean value $\mathbb{E}[\psi] = (\sigma_\omega \sqrt{2})/\sqrt{\pi}$, and (ii) σ_ω increases as ΔT_{res} increases. These simulation results are in agreement with Proposition 3 and Remark 2.

Figure 5.17 compares the results in terms of the temperature limit deviations for three cases: (i) $\Delta T_{\mathrm{res}} = 0°C$, (ii) $\Delta T_{\mathrm{res}} = 0.1°C$ and no temperature deviation bound ΔT_{b} is used, and (iii) $\Delta T_{\mathrm{b}} = 0.1°C$ and $\Delta T_{\mathrm{b}} = 1°C$. The left plot shows the histogram of temperature limit deviations at the end of a 5-day simulation, whereas the right plot shows the evolution of the standard deviation of temperature limit changes over time.

If $\Delta T_{\mathrm{res}} = 0°C$ (blue color), the histogram resembles a normal distribution and the standard deviation slowly increases and stabilizes to a small value, as explained in Remark 2. If the thermostat resolution is $\Delta T_{\mathrm{res}} = 0.1°C$ and no ΔT_{b} is used (red color), the histogram resembles again a normal distribution but the standard deviation monotonically increases, as shown in Proposition 3. However, if a bound $\Delta T_{\mathrm{b}} = 1°C$ is used (green color), the standard deviation is capped at a significantly lower value. Observe that the effect of bounding is a more uniform probability distribution of temperature deviations across the population.
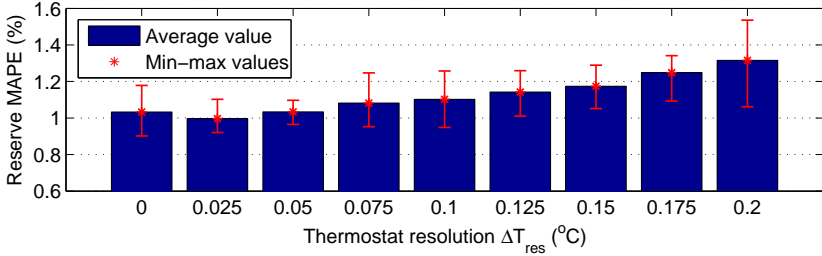
Figure 5.18: The dependence of reserve MAPE on the thermostat resolution.

The effect of minimum thermostat resolution on control accuracy is shown in Fig. 5.18. The bars are the mean values of reserve MAPE over 10 simulations, whereas the lines indicate the minimum and maximum values. The reserve MAPE generally increases as ΔT_{res} increases but the increase is rather small; therefore, the thermostat resolution limitations only slightly reduce the control accuracy. In addition, we compared the reserve MAPE for $\Delta T_{\mathrm{res}} = 0.1^\circ$C, with and without a bound $\Delta T_{\mathrm{b}} = 1^\circ$C, and no observable reduction in control accuracy was found.

5.16   Performance with Door Openings

In this section, we investigate the controller's performance in a realistic case with random door opening events. We performed simulations for the base case scenario with and without door openings and compare the control performance in Table 5.8. Due to the door openings, the reserve MAPE increases by approximately 40% for zero-mean frequency deviation signals and by 80% for biased signals. However, in both cases the error is still small, and therefore the proposed controller is applicable despite the random door openings.

Observe that the baseline MAPE decreases with door openings due to the smoother baseline power during the load ramps, especially around 3 pm and 7 pm. To illustrate this, Fig. 5.19 shows the baseline power with and without door openings, as well as the aggregate power without PFC. Notice that the natural fluctuations around the baseline are less pronounced in the presence of door openings.

Figure 5.20 shows the activation of PFC reserve ("desired power") around

Table 5.8: Baseline and reserve MAPE with and without door openings

Baseline MAPE (%)		Reserve MAPE (%) Small bias		Reserve MAPE (%) Large bias	
Without	With	Without	With	Without	With
0.90	0.47	1.01	1.41	1.15	2.06
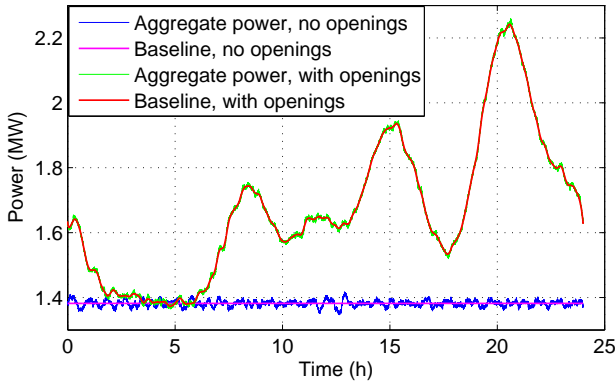
Figure 5.19: Aggregate power and baseline of the refrigerator population with and without door openings.

the baseline with and without door openings. The baseline with door openings has an increasing trend, but the PFC activation is the same as that without door openings. Furthermore, observe that the door openings do not have a significant effect on the tracking performance of the desired power.

5.17   Implementation Issues

In this section, we discuss a number of additional issues that need to be considered for a real-world implementation.

Refrigerator model: Our analysis is based on the commonly used first-order refrigerator model without a freezer. As shown in [143], using separate states for the temperature of the refrigerator compartment's air, refrigerator's content, and evaporator significantly improves the model's performance. Modeling a refrigerator with a freezer is possible following the approach of [104].
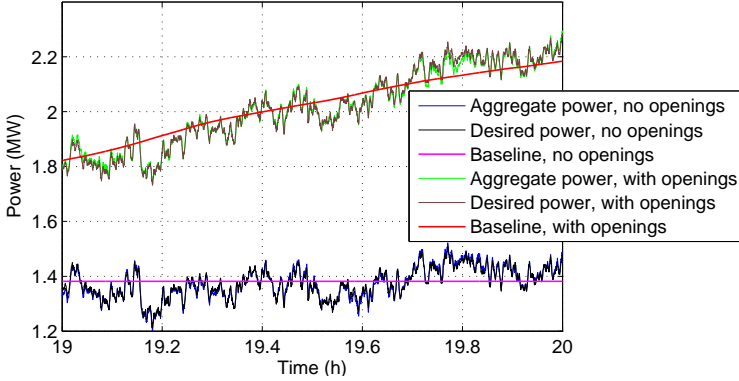
Figure 5.20: PFC activation around the baseline from 7 pm to 8 pm with and without door openings.

The above improved model versions can be used with only small changes in control design. Equations (5.39), (5.40) and (5.41) used to estimate the on/off times and the duty cycle and, possibly, the modeling of startup dynamics must be modified. Our analysis assumed that the profile of startup dynamics is time invariant; however, in reality it depends on the evaporator temperature when the refrigerator switches on. Because this temperature depends on the time elapsed since the previous on cycle, which is not constant when providing PFC, the startup dynamics profile will likely be time-varying. This dependence can be considered with a two-state refrigerator model.

Communication links: The control is decentralized and so there is no real-time communication required between the loads and an aggregator. However, in practice a communication channel with a low data transfer rate might be preferable leading to a semi-autonomous operation mode. For example, in several countries PFC is organized in a market setting where the reserve capacity is auctioned on a weekly or daily basis [70]. Whenever the market is cleared, a new value for D^r must be communicated to the refrigerators' micro-controllers.

The control design relies on mean values of several parameters in the population. Therefore, if the statistical properties of the population change (e.g., if at some point new refrigerators are added) the new mean values must be communicated to the refrigerators. Furthermore, the micro-controllers need estimates of the population's duty cycle and

average temperature. Although this can be done in a decentralized way, the estimates could be reset to the actual values periodically, if a low data transfer communication link is available. This will enhance the robustness to accumulated estimation errors (although this problem was not observed in our simulations).

Note that the above communication exchanges are not necessary for implementation, and if established, they are only needed very infrequently, e.g., once a day or once a week. This type of communication with the aggregator can be achieved via existing smart metering technologies, whereas the communication from the refrigerator's micro-controller to the smart meter is possible with a simple home area network [87].

Noisy frequency measurements: Since the control relies on frequency measurements at the wall outlet level, a valid question is how measurement noise affects the controller's performance. According to [85], a moving average filter can be used to reduce the measurement noise. The authors of [146] performed a correlation analysis of frequency measurements at the transmission level (230 kV) and at the wall outlet level (120 V). The results were promising for PFC and showed that the frequency measurements at the two voltage levels are highly correlated.

Cost and business cases: A common concern in load control schemes is that implementation cost should be low enough to enable customer adoption. According to [85], the estimated cost of a simple controller based on frequency thresholds is 20 € including the frequency meter.[8] We believe that the micro-controller would incur an additional cost of only a few € due to its simplicity.

We perform a simple revenue estimation for each refrigerator that provides PFC, using a price of 21.5 €/MW/h for the PFC reserve capacity (average price from January to September 2014 in Switzerland [150]) and neglecting the discount rate. According to our simulations, a population of 70,000 refrigerators can provide a PFC reserve capacity of

---

[8]The accuracy of this frequency meter was not reported in [85] and the references therein. Although high-accuracy meters of mains frequency are typically costly, some low-cost designs have been proposed; for example, [147] and [148] developed frequency meters with an accuracy of 1 mHz, but did not report the selling price. At the time when this thesis was written, a typical selling price for a frequency meter with accuracy 10 mHz was approximately 40 € [149]. Apart from the cost, another important aspect is the number of cycles (of the sinusoidal quantity in an AC system, typically the voltage) that are needed to achieve a certain measurement accuracy. In general, the higher the accuracy is, the more cycles are required, and therefore the longer it takes for the meter to produce a single frequency measurement.

1.103 MW on average (this value corresponds to $D^{\mathrm{r}} = 0.2$); therefore, approximately $63,500$ refrigerators are needed to provide 1 MW of capacity. The average revenue from selling 1 MW of capacity is $21.5\ \text{€}/\text{MW}/\text{h} \cdot 1\ \text{MW} \cdot 8760\ \text{h}/\text{year} = 188,340\ \text{€}/\text{year}$. Thus, the average revenue per refrigerator is as low as $188,340/63,500 \approx 3\ \text{€}/\text{year}$. Assuming an average lifetime of 14 years [151], each refrigerator will recover approximately $3 \cdot 14 = 42\ \text{€}$, which is slightly higher than the implementation cost. Although the implementation cost per device is low, investing in the proposed controller is not a profitable investment due to the large number of devices that are needed to provide a meaningful amount of reserve capacity.

Retrofits are possible but standardization will help reduce the costs because new refrigerators could be sold with the necessary hardware already integrated. If the revenues per device from selling PFC reserves are still very low, a different business case is needed to enable customer adoption. One possibility is to offer an electricity price rebate to the customers who participate in PFC.

5.18 Conclusion

In this chapter, we presented a decentralized stochastic control scheme to enable PFC reserve provision from an aggregation of refrigerators. The control is based on probabilistic switching to avoid load synchronization, and accounts for refrigerator startup dynamics and lockout constraints. Reliable reserve provision during persistent frequency deviations is achieved by introducing a deterministic resetting of thermostat temperature limits. The controller's implementation at the device level is simple and does not require communication with an aggregator. Furthermore, the controller is robust to biased frequency deviations, limited thermostat resolution and door openings. Extensive simulation results showed that the proposed control scheme outperforms a relevant benchmark controller from the literature, and allows a refrigerator aggregation to provide fast and reliable PFC without communication.

Chapter 6

Secondary Frequency Control with Electric Water Heaters

6.1   Introduction

A large amount of literature has focused on provision of Secondary Frequency Control (SFC) reserves from Electric Water Heaters (EWHs) or other types of Thermostatically Controlled Loads (TCLs). A detailed literature review was provided in Chapter 4. In this chapter, we design four rule-based controllers with different complexity and communication requirements for SFC with EWHs. Furthermore, we develop a statistical model for a population of EWHs based on the thermally stratified EWH model of Section 2.2.3, and use the statistical model to evaluate the performance of the rule-based controllers in a simulation study.

The remainder of this chapter is organized as follows. In Section 6.2, we present a statistical model for EWH populations, and in Section 6.3 we define the State of Charge (SoC) for thermally stratified EWHs. The rule-based controllers for EWH aggregations are presented in Section 6.4 and evaluated based on different indicators in Section 6.5. Last, Section 6.6 concludes this chapter that is based on [21]. The nomenclature of this chapter is summarized in Tables 6.1, 6.2 and 6.3.

Table 6.1: Nomenclature of Chapter 6: symbols

Symbol	Unit	Description
b	-	Binary signal used in stochastic blocking control
c	J/kg$^\circ$C	Specific heat capacity of water
E	Wh or -	Energy
m	kg	Water mass
$\boldsymbol{m}_{\mathrm{cat}}$	-	EWH water mass categories
$\dot{m}_{\mathrm{norm}}$	-	Normalized mass flow rate during a water draw
$\dot{m}_{\mathrm{draw}}$	l/min	Mass flow rate during a water draw
$\mathcal{N}$	-	Set of EWHs
N_{d}	-	Number of EWH devices
$n_{\mathrm{hourly}}^{\mathrm{avg}}$	-	Average number of water draws for each hour
n_{daily}	-	Total number of water draws per day
P	W	Electric power
$\boldsymbol{P}_{\mathrm{el,cat}}^{\mathrm{rated}}$	W	EWH rated power categories
$P_{\mathrm{el},i}$	W	Rated power of EWH i
$\mathbf{p}_{\mathrm{cat}}$	-	Probability vector of mass categories
p_{draw}	-	Probability of draw event for each hour
R	$^\circ$C or -	EWH thermal energy content
u	-	Switching signal in direct and indirect control
s	-	Signal for deadband crossing event
T	$^\circ$C	Temperature
t_{draw}	-	Time instance when a water draw event starts
U	W/(m$^2\,^\circ$C)	EWH thermal loss coefficient
w	-	Normalized SFC signal
γ	-	Control parameter used in stochastic blocking control
ΔP	W	Difference between P_{ref} and P_{agg}
ΔP_{ef}	W	Effective ΔP after accounting for natural switching
Δt_{draw}	s	Time between two consecutive water draw events
$\Delta t_{\mathrm{duration}}$	s	Duration of a water draw event
μ	-	Mean value of a random variable
ρ	-	Random number uniformly distributed in $[0,1]$
σ	-	Standard deviation of a random variable
ω	-	Weighting factor for SoC calculation

Table 6.2: Nomenclature of Chapter 6: probability operators

Operator	Description
$\mathbb{N}$	Normal probability distribution
$\mathbb{U}$	Uniform probability distribution
EXP	Exponential probability distribution

6.2   Water Heater Population Model

In this section, we 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 an EWH population with and without

Table 6.3: Nomenclature of Chapter 6: subscripts and superscripts

Symbol	Description
agg	Aggregate power
av,off	EWH that is available to switch off
av,on	EWH that is available to switch on
base	Baseline power
cat	Category vector
cb	SFC control band
cold	EWH with a temperature below the minimum thermostat limit
daily	Daily water consumption
day	Daily energy consumption with SFC
db	Thermostat temperature deadband
el	Water layer of the heating element
est	Estimated temperature
hot	EWH with a temperature above the maximum thermostat limit
i	EWH index
j	Water layer index within an EWH
long	Long water draw
off	EWH that is in the off state
on	EWH that is in the on state
r	Reference energy content of an EWH
ref	Target power profile (reference)
rel	Relative energy content of an EWH
min/max	Minimum/maximum value of a variable or parameter
set	Thermostat temperature setpoint
short	Short water draw
sim	Simulation horizon
sw,off	EWH that is switched off by the external controller
sw,on	EWH that is switched on by the external controller
t	Time index
th,off	EWH that is switched off by its thermostatic controller
th,on	EWH that is switched on by its thermostatic controller
top	Top layer of the water tank
0	Initial temperature vector

external control actions from a central coordination entity. The EWH population model consists of two parts: (a) the EWH thermal model and (b) the water draw model.

6.2.1 Modeling of Population Parameters

We describe a routine that creates an heterogeneous set of parameters such as tank volume, heater power rating, and daily water consumption for any desired number of EWHs. We consider a certain correlation between these parameters, e.g., a small EWH is more likely to have a low-power heating element. The methodology consists of the following steps:
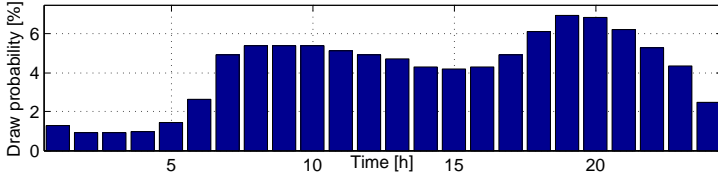
Figure 6.1: Profile describing the probability of a water draw occurring in the course of the day ($\boldsymbol{p}_{\mathrm{draw}}$).

1. We define the EWH category vector $\boldsymbol{m}_{\mathrm{cat}}$ that contains the considered tank sizes. The individual EWH sizes are drawn from $\boldsymbol{m}_{\mathrm{cat}}$ based on a predefined probability vector $\boldsymbol{p}_{\mathrm{cat}}$.

2. We define the matrix $\boldsymbol{P}_{\mathrm{el,cat}}^{\mathrm{rated}}$, the column l of which contains possible power ratings of EWHs of category l. The rated power $P_{\mathrm{el},i}$ of EWH i of category l is drawn from column l of $\boldsymbol{P}_{\mathrm{el,cat}}^{\mathrm{rated}}$ based on a discrete uniform distribution.

3. For each EWH category, we define an interval $[m_{\mathrm{daily,cat}}^{\mathrm{min}}, m_{\mathrm{daily,cat}}^{\mathrm{max}}]$ for the daily water consumption in kg. The actual daily water consumption $m_{\mathrm{daily},i}$ of EWH i is drawn from a uniform distribution between these bounds.

4. The center and width of the thermostat's deadband of EWH i, $T_{\mathrm{set},i}$ and $T_{\mathrm{db},i}$ respectively, are drawn from uniform distributions in the intervals $[T_{\mathrm{set}}^{\mathrm{min}}, T_{\mathrm{set}}^{\mathrm{max}}]$ and $[T_{\mathrm{db}}^{\mathrm{min}}, T_{\mathrm{db}}^{\mathrm{max}}]$.

5. The thermal loss coefficient U_i is drawn from a uniform distribution in the interval $[U^{\mathrm{min}}, U^{\mathrm{max}}]$.

6.2.2   Modeling of Water Draws

In order to model random water draw events induced by customer behavior, we generate a draw scenario for each EWH by taking random values for the draw starting time, the draw duration, and the water flow rate from predefined probability distributions. In the rest of this section, we describe how we assemble the probability distributions and show an example of a water draw time series.

1. To determine the probability of a draw event, we utilize the probability profile (vector) from [152] depicted in Fig. 6.1 and denoted

by $\boldsymbol{p}_{\mathrm{draw}} \in \mathbb{R}^{24}$. Furthermore, we determine the total number of water draws per day by drawing from

$$n_{\mathrm{daily},i} \sim \mathbb{U}\left(n_{\mathrm{daily}}^{\min}, n_{\mathrm{daily}}^{\max}\right) \ , \tag{6.1}$$

where $\mathbb{U}$ denotes the uniform distribution between a minimum and a maximum value. We multiply the hourly draw probability vector with the total number of draws per day to determine a vector with the average number of water draws for each hour

$$\boldsymbol{n}_{\mathrm{hourly},i}^{\mathrm{avg}} = \boldsymbol{p}_{\mathrm{draw}} \cdot n_{\mathrm{daily},i} \ . \tag{6.2}$$

2. We determine the time intervals $\Delta t_{\mathrm{draw},i}$ between water draws in the course of the day by drawing from the distribution

$$\Delta t_{\mathrm{draw},i} \ \sim \mathrm{EXP}\left(\frac{1}{n_{\mathrm{hourly},i}^{\mathrm{avg}}}\right) \ , \tag{6.3}$$

where EXP is the exponential distribution with the expected value as a parameter, and $n_{\mathrm{hourly},i}^{\mathrm{avg}}$ is the appropriate element of $\boldsymbol{n}_{\mathrm{hourly},i}^{\mathrm{avg}}$ depending on the hour. By calculating the cumulative sum of $\Delta t_{\mathrm{draw},i}$, we assemble a vector containing the time instances $t_{\mathrm{draw},i}$ when a water draw event starts.

3. We distinguish between two types of water draws: long draws that correspond to showers or baths, and shorter draws related to activities such as washing hands or cooking. We assume that a water draw is more likely to be long during times of the day when a lot of water is used. The draw duration times $\Delta t_{\mathrm{duration},i}^{\mathrm{long}}$ and $\Delta t_{\mathrm{duration},i}^{\mathrm{short}}$ are random variables drawn from normal distributions, and the draw durations $\Delta t_{\mathrm{duration},i}$ are assembled in a vector.

4. In order to induce a variation in the water flow rate, we draw a normalized flow rate for every draw event from the normal distribution

$$\dot{m}_{\mathrm{draw},i} \sim \mathbb{N}\left(\mu_{\dot{m}_{\mathrm{draw}}}, \sigma_{\dot{m}_{\mathrm{draw}}}\right) \ . \tag{6.4}$$

5. The resulting normalized mass flow time series $\dot{m}_{\mathrm{norm},i}$ is assembled with the information contained in $t_{\mathrm{draw},i}$, $\dot{m}_{\mathrm{draw},i}$, and $\Delta t_{\mathrm{duration},i}$. The draw time series is then scaled so that the total daily consumption equals $m_{\mathrm{daily},i}$.

Table 6.4: EWH population parameters

Symbol	Parameter	Value	Unit
$\boldsymbol{m}_{\mathrm{cat}}$	Water mass categories	[50, 100, 200, 300]	[kg]
$\boldsymbol{p}_{\mathrm{cat}}$	Category probability	[0.1, 0.25, 0.25, 0.4]	[-]
$\boldsymbol{P}^{\mathrm{rated}}_{\mathrm{el,cat}}$	Rated power per cat.	$\begin{bmatrix} 2 & 3 & 4 & 4 \\ 2.5 & 3.5 & 4.5 & 5 \\ 3 & 4 & 5 & 6 \end{bmatrix}$	[kW]
$\left[T^{\min}_{\mathrm{set}}, T^{\max}_{\mathrm{set}}\right]$	Temp. set-point bounds	[55, 65]	[°C]
$\left[T^{\min}_{\mathrm{db}}, T^{\max}_{\mathrm{db}}\right]$	Deadband size bounds	[5, 15]	[°C]
$\left[U^{\min}, U^{\max}\right]$	Heat loss coefficient bounds	[0.2, 1]	$\left[\frac{\mathrm{W}}{\mathrm{m}^2\,\mathrm{K}}\right]$
T_0	Initial temperature	60	[°C]

Table 6.5: Parameters of the probabilistic water draw model

Variable	Distr.	Parameters
$n_{\mathrm{daily},i}$	$\mathbb{U}$	$(n^{\min}_{\mathrm{daily}}, n^{\max}_{\mathrm{daily}}) = (50, 80)$
$\Delta t^{\mathrm{long}}_{\mathrm{duration},i}$	$\mathbb{N}$	$\mu = 600$ s, $\sigma = 60$ s
$\Delta t^{\mathrm{short}}_{\mathrm{duration},i}$	$\mathbb{N}$	$\mu = 60$ s, $\sigma = 6$ s
$\dot{m}_{\mathrm{draw},i}$	$\mathbb{N}$	$\mu = 0.7$ l/min, $\sigma = 0.3$ l/min

6.2.3   Population and Water Draw Parameters

Table 6.4 shows the parametrization for the water heater population used in the simulations of this chapter. The temperature of all EWHs at the beginning of the simulation is assumed to be equal to 60°C (all water layers of each EWH are at 60°C, i.e, there is no stratification). The daily water draw volume is assumed to be equal to 200 l for all EWHs, i.e., $m^{\min}_{\mathrm{daily,cat}} = m^{\max}_{\mathrm{daily,cat}} = 200$ l. The parameters of the probability distributions of the water draw model are given in Table 6.5, whereas an example of a water draw scenario is shown in Fig. 6.2.
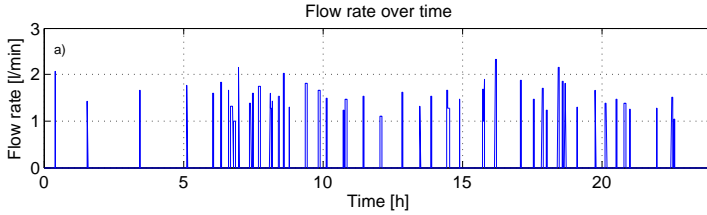
Figure 6.2: An example of a water draw scenario.

6.3 SoC Definition

In this section we discuss three approaches to quantify the energy stored in the thermal mass of an EWH. A simple way to do so is by comparing the difference between the reading of the temperature sensor and the lower temperature limit of the thermostat $T_{\min,i}$, with the size of the thermostat's deadband $T_{\max,i} - T_{\min,i}$. Let us denote by $T_{\mathrm{el},i,t}$ the temperature measurement of a sensor placed close to the heating element (which is typically the case) of EWH i at time step t. In this case, we can quantify the relative energy stored in the EWH according to [103, 153]

$$E_{\mathrm{rel},i,t} = \frac{T_{\mathrm{el},i,t} - T_{\min,i}}{T_{\max,i} - T_{\min,i}} \ .$$
(6.5)

An alternative simple way to calculate the relative energy is based on the temperature perceived by the user, when water is drawn from the top of the water tank ($T_{\mathrm{top},i,t}$), according to[1]

$$E_{\mathrm{rel},i,t} = \frac{T_{\mathrm{top},i,t} - T_{\min,i}}{T_{\max,i} - T_{\min,i}} \ .$$
(6.6)

Despite their simplicity, (6.5) and (6.6) do not capture the distribution of temperature within the water tank. A more detailed calculation of the stored energy should account for the thermal stratification along the vertical axis of the water tank. To this end, we define the reference energy content $E_{\mathrm{r},i}$ as the thermal energy stored in the EWH when

---

[1]Typically, cold water enters an EWH at the bottom of the tank, whereas hot water is drawn from the top of the tank, to maintain the thermal stratification.

the temperature of all water layers is equal to the thermostat's upper
temperature limit $T_{\mathrm{max},i}$

$$E_{\mathrm{r},i} = \sum_j m_j c(T_{\mathrm{max},i} - T_{\mathrm{min},i}) \ , \tag{6.7}$$

where m_j is the water mass contained in each layer, and c is the thermal
capacity of water.

The stored energy in a thermally stratified EWH can be expressed us-
ing the SoC concept, which is typically used for batteries. Using the
thermostat's lower temperature limit $T_{\mathrm{min},i}$ as a reference temperature,
the SoC is given by

$$SOC_{i,t} = \frac{n \sum_j \omega_j m_j c \cdot \max\left[(T_{j,i} - T_{\mathrm{min},i}), 0\right]}{E_{\mathrm{r},i}} \ . \tag{6.8}$$

Since $T_{\mathrm{min},i}$ is considered as a reference energy value, the water layers
with temperature $T_{j,i}$ below $T_{\mathrm{min},i}$ have zero thermal energy content
and do not contribute to the total SoC. The weighting factors ω_j are
used in (6.8) to reflect the dependency of the value of thermal energy
on layer position along the vertical axis of the tank, and they satisfy
$\sum_j \omega_j = 1$. In the simpler case, this dependency can be neglected and
all ω_j can be set equal to $1/n$. Alternatively, the layers close to the
top of the tank can be given a larger weight, because they affect the
temperature perceived by the user during a water draw. According to
(6.8), the lowest SoC value is 0 and the maximum value can be greater
than 1, which corresponds to average water temperatures greater than
$T_{\mathrm{max},i}$.

6.4   Control Strategies

We use the models and definitions introduced in Sections 2.2.3, 6.2, and
6.3 to develop four strategies for EWH aggregate power control. The
goal is to track a given power reference $P_{\mathrm{ref},t}$ that depends on the SFC
signal (the way to construct $P_{\mathrm{ref},t}$ is presented in Section 6.5.2). The
strategies differ with respect to: (i) the amount of information that
is exchanged between each device and the central coordination unit;
and (ii) the degree to which the central controller can interact with the
device's internal duty cycle.

6.4.1 Stochastic Blocking Control (C1)

In this strategy, only the reference power $P_{\text{ref},t}$, the measured aggregate load $P_{\text{agg},t}$, and the maximum power demand of the population P_{max} (the total power when all EWHs are on) are available to the central controller. The controller can only block (or release from blocking) individual EWHs, but not actively switch on an EWH. A blocking command overrides a potential on signal from the internal hysteresis controller, which means that an EWH might be off even if the temperature is outside the deadband. On the other hand, if an EWH is not blocked, then the on/off state of the heating element is determined by the internal hysteresis controller.

This approach builds on the work of [104, 153] and is based on the broadcast control parameter γ_t, which indicates the fraction of appliances that shall be blocked (or released from blocking) at each time step

$$\gamma_t = \frac{P_{\text{agg},t} - P_{\text{ref},t}}{P_{\text{max}}} \ . \tag{6.9}$$

If $0 < \gamma_t \leq 1$ holds, a blocking action is required, whereas if $-1 \leq \gamma_t < 0$ holds, some of the EWHs that were previously blocked should be released. The decision about blocking or releasing devices is made by generating a uniformly distributed random number $\rho_{i,t}$ for each device i ($0 \leq \rho_{i,t} \leq 1$). At time step t, a binary block signal $b_{i,t}$ is assigned to each EWH according to the rule (the value $b_{i,t} = 1$ means that the EWH is blocked)

$$b_{i,t} = \begin{cases} 1 & \text{if } (b_{i,t-1} = 0) \wedge (\gamma_t > 0) \wedge (\rho_{i,t} \leq \gamma_t) \\ 0 & \text{if } (b_{i,t-1} = 1) \wedge (\gamma_t < 0) \wedge (-\rho_{i,t} \geq \gamma_t) \ . \\ b_{i,t-1} & \text{otherwise} \end{cases} \tag{6.10}$$

If $\gamma_t > 0$, only the EWHs that are unblocked at time step $t - 1$ draw a random number $\rho_{i,t}$, and therefore the uniform distribution of $\rho_{i,t}$ ensures that the percentage of EWHs that block at time step t is close to γ_t. Similarly, if $\gamma_t < 0$, the percentage of EWHs that unblock at time step t is close to $|\gamma|$.

The control error $P_{\text{agg},t} - P_{\text{ref},t}$ in (6.9) is normalized by the maximum power demand of the population P_{max}. Although this is a very simple approach, it introduces some inaccuracy in the control because not all EWHs are available to respond to the broadcast signal γ_t. A better

approach would be to normalize $P_{\text{agg},t} - P_{\text{ref},t}$ with an estimate of the total power of the unblocked devices if $\gamma_t > 0$, and with an estimate of the total power of the blocked devices if $\gamma_t < 0$.

However, even with an improved calculation of γ_t, the control will not be perfect because this strategy does not actively switch on or off the EWHs. More specifically, blocking an EWH at time step t does not necessarily reduce the aggregate power at time step t, because the EWH might be off already from time step $t-1$. In addition, unblocking a blocked EWH does not guarantee that the EWH will switch on, because this depends on the EWH temperature and its internal controller.

Overall, the stochastic blocking strategy C1 has a very simple control implementation, it is based on broadcasting, and it does not require any feedback from the loads or estimation of unmeasured quantities. We include strategy C1 in our analysis to investigate the potential of simple control approaches in SFC.

6.4.2   Direct Temperature Feedback Control (C2)

In this approach, the controller receives information about the on/off state and the temperature of each EWH through individual communication links, and it also knows the thermostat's temperature limits $T_{\text{min},i}$ and $T_{\text{max},i}$ $\forall i$. Unlike approach C1, the controller can actively switch both on and off individual devices. This strategy consists of two parts: (i) the selection of the subset of EWHs that is in principle available for external control, and (ii) the ranking of this subset to specify the EWHs that will actually provide the reserve. The algorithm is outlined in the rest of this section.

Selection

At each time step t, the external controller divides the EWHs into sub-populations based on their temperature and on/off state. Let $\mathcal{N}_{\text{off},t}$ denote the set of EWHs that are currently off, and by $\mathcal{N}_{\text{on},t}$ the set of EWHs that are currently on. $\mathcal{N}_{\text{cold},t}$ and $\mathcal{N}_{\text{hot},t}$ are the sets of EWHs that are not available for external control actions because their temperature is either below the minimum setpoint or above the maximum setpoint. Therefore, the set of EWHs that are within the temperature deadband is defined as

$$\mathcal{N}_{\text{db},t} = \neg(\mathcal{N}_{\text{cold},t} \cup \mathcal{N}_{\text{hot},t}) \ . \tag{6.11}$$

Let $\mathcal{N}_{\text{th,on},t}$ and $\mathcal{N}_{\text{th,off},t}$ denote the sets of EWHs that will be switched on and off by their internal thermostatic controllers due to very low and very high temperatures, respectively

$$\mathcal{N}_{\text{th,on},t} = \mathcal{N}_{\text{off},t} \cap \mathcal{N}_{\text{cold},t} \tag{6.12}$$

$$\mathcal{N}_{\text{th,off},t} = \mathcal{N}_{\text{on},t} \cap \mathcal{N}_{\text{hot},t} \ . \tag{6.13}$$

The sets of EWHs that are available for SFC actions are defined as

$$\mathcal{N}_{\text{av,off},t} = \mathcal{N}_{\text{on},t} \cap \mathcal{N}_{\text{db},t} \tag{6.14}$$

$$\mathcal{N}_{\text{av,on},t} = \mathcal{N}_{\text{off},t} \cap \mathcal{N}_{\text{db},t} \ . \tag{6.15}$$

Ranking

The control error at each time step is equal to the difference between the reference and the actual aggregate power of the population

$$\Delta P_t = P_{\text{ref},t} - P_{\text{agg},t} \ . \tag{6.16}$$

Taking into account the thermostat switching actions of (6.12) and (6.13), the effective difference in power that needs to be achieved by additional switching actions is equal to

$$\Delta P_{\text{ef},t} = \Delta P_t + P_{\text{th,off},t} - P_{\text{th,on},t} \ , \tag{6.17}$$

where $P_{\text{th,off},t}$ is the total power of the EWHs that belong to the set $\mathcal{N}_{\text{th,off},t}$, and $P_{\text{th,on},t}$ is the total power of the EWHs in $\mathcal{N}_{\text{th,on},t}$.

If $\Delta P_{\text{ef},t} < 0$ additional switching off action is required, whereas if $\Delta P_{\text{ef},t} > 0$ switching on action is desired. The list of EWHs that will switch is constructed depending on the sign of $\Delta P_{\text{ef},t}$ as follows:

- If $\Delta P_{\text{ef},t} < 0$, the EWHs in the set $N_{\text{av,off},t}$ are ranked in descending order of the thermal energy content $R_{i,t}$, such that the hotter an EWH is the higher its switching off priority. The set of EWHs that will receive a switch off signal $\mathcal{N}_{\text{sw,off},t}$ is determined by selecting from the ordered list the number $N_{\text{d},t}$ of devices that minimizes

$$\left| \Delta P_{\text{ef},t} + \sum_{i=0}^{N_{\text{d},t}} P_{\text{el},i} \right| \ , \tag{6.18}$$

where $P_{\text{el},i}$ is the power of the ith heater in the list.

- If $\Delta P_{\mathrm{ef},t} > 0$, the EWHs in the set $\mathcal{N}_{\mathrm{av,on},t}$ are ranked in ascending order of thermal energy content, such that the colder an EWH is the higher its switching on priority. The set of EWHs that will receive a switch on signal $\mathcal{N}_{\mathrm{sw,on},t}$ is determined by selecting from the ordered list the number $N_{\mathrm{d},t}$ of devices that minimizes

$$\left| \Delta P_{\mathrm{ef},t} - \sum_{i=0}^{N_{\mathrm{d},t}} P_{\mathrm{el},i} \right| . \tag{6.19}$$

Depending on the definition of the sets $\mathcal{N}_{\mathrm{cold},t}$ and $\mathcal{N}_{\mathrm{hot},t}$, as well as the thermal energy content $R_{i,t}$, different variations of control strategy C2 can be considered. In this chapter, we investigate the six different variations C2a – C2f for which $\mathcal{N}_{\mathrm{cold},t}$, $\mathcal{N}_{\mathrm{hot},t}$ and $R_{i,t}$ are defined as

$$\mathcal{N}_{\mathrm{cold},t} = \begin{cases} \{i \mid T_{\mathrm{el},i,t} \leq T_{\min,i}\} & \text{for C2a, C2b, C2c} \\ \{i \mid T_{\mathrm{top},i,t} \leq T_{\min,i}\} & \text{for C2d} , \\ \{i \mid SOC_{i,t} \leq SOC_{\min,i}\} & \text{for C2e, C2f} \end{cases} \tag{6.20}$$

$$\mathcal{N}_{\mathrm{hot},t} = \begin{cases} \{i \mid T_{\mathrm{el},i,t} \geq T_{\max,i}\} & \text{for C2a, C2b, C2c} \\ \{i \mid T_{\mathrm{top},i,t} \geq T_{\max,i}\} & \text{for C2d} , \\ \{i \mid SOC_{i,t} \geq SOC_{\max,i}\} & \text{for C2e, C2f} \end{cases} \tag{6.21}$$

$$R_{i,t} = \begin{cases} T_{\mathrm{el},i,t} - T_{\min,i} & \text{for C2a and } \Delta P_{\mathrm{ef},t} < 0 \\ T_{\mathrm{el},i,t} - T_{\max,i} & \text{for C2a and } \Delta P_{\mathrm{ef},t} > 0 \\ T_{\mathrm{top},i,t} - T_{\min,i} & \text{for C2b and } \Delta P_{\mathrm{ef},t} < 0 \\ T_{\mathrm{top},i,t} - T_{\max,i} & \text{for C2b and } \Delta P_{\mathrm{ef},t} > 0 \\ SOC_{i,t} & \text{for C2c, C2d, C2e, C2f} \end{cases} \tag{6.22}$$

In strategies C2a, C2b and C2c, we determine the EWHs that are outside the deadband (the sets $\mathcal{N}_{\mathrm{cold},t}$ and $\mathcal{N}_{\mathrm{hot},t}$), as well as the thermal energy content $R_{i,t}$, based on the temperature measurement from the sensor close to the heating element. In strategy C2d, we determine the sets $\mathcal{N}_{\mathrm{cold},t}$ and $\mathcal{N}_{\mathrm{hot},t}$, as well as the thermal energy content $R_{i,t}$, based on the temperature measurement from the sensor at the top of the water tank.

In contrast, in strategies C2e and C2f, we use the SoC defined in (6.8) to identify the sets $\mathcal{N}_{\mathrm{cold},t}$ and $\mathcal{N}_{\mathrm{hot},t}$, as well as to rank the EWHs based on the stored energy. The only difference between strategies C2e and C2f is the selection of the boundary value $SOC_{\min,i}$ of the set $\mathcal{N}_{\mathrm{cold},t}$, which

is fixed to 0.3 for strategy C2e and 0.4 for strategy C2f. Strategy C2e is more aggressive than strategy C2f, because it utilizes the flexibility of EWHs with lower energy content due to the lower $SOC_{\min,i}$ value. On the other hand, the boundary value $SOC_{\max,i}$ of the set $\mathcal{N}_{\text{hot},t}$ is fixed to 0.9 for both strategies, because preliminary simulations showed that the results are not sensitive to this parameter.

The central controller sends the switching signal $u_{i,t}$ to each EWH according to the rule

$$u_{i,t} = \begin{cases} 1 & \text{if } i \in \mathcal{N}_{\text{sw,on},t} \cup \mathcal{N}_{\text{th,on},t} \\ 0 & \text{if } i \in \mathcal{N}_{\text{sw,off},t} \cup \mathcal{N}_{\text{th,off},t} \\ u_{i,t-1} & \text{otherwise} \end{cases} . \tag{6.23}$$

From (6.12), (6.13), (6.20) and (6.21) one can see that, in contrary to C2a, C2b and C2c, in variations C2d, C2e and C2f the external SFC controller overrides the internal thermostatic controllers because $\mathcal{N}_{\text{th,on},t}$ and $\mathcal{N}_{\text{th,off},t}$ are not based on $T_{\text{el},i,t}$.

6.4.3 Indirect Temperature Feedback Control (C3)

In this approach, the controller does not receive any direct temperature information. The controller is aware of the on/off state of each EWH and the temperature deadband crossing events, which can be communicated by transmitting the signal

$$s_{i,t} = \begin{cases} 1 & \text{if } T_{\text{el},i,t} \leq T_{\min,i} \\ 0 & \text{if } T_{\min,i} \leq T_{\text{el},i,t} \leq T_{\max,i} \\ -1 & \text{if } T_{\text{el},i,t} \geq T_{\max,i} \end{cases} . \tag{6.24}$$

The signal $s_{i,t}$ is sent only when a deadband crossing occurs, thus the communication burden is much lower compared with strategy C2. The selection and ranking steps of Section 6.4.2 are applied also here and the subsets $\mathcal{N}_{\text{cold},t}$, $\mathcal{N}_{\text{hot},t}$ and $\mathcal{N}_{\text{db},t}$ are defined as

$$\mathcal{N}_{\text{cold},t} = \{i \mid s_{i,t} = 1\} \tag{6.25}$$
$$\mathcal{N}_{\text{hot},t} = \{i \mid s_{i,t} = -1\} \tag{6.26}$$
$$\mathcal{N}_{\text{db},t} = \{i \mid s_{i,t} = 0\} . \tag{6.27}$$

With these definitions, (6.16) and (6.17) can be used to calculate $\Delta P_{\text{ef},t}$.

The central controller models each EWH in the aggregation using (2.25)-(2.28) and historical data of the hourly water consumption. During real-time operation, the controller uses these data to simulate the population and estimate the temperature of each EWH $\mathbf{T}_{i,t}^{\text{est}}$.[2] In addition, whenever a signal $s_{i,t}$ from an individual EWH is received, the vector $\mathbf{T}_{i,t}^{\text{est}}$ is updated according to

$$
\mathbf{T}_{i,t}^{\text{est}} = \begin{cases} T_{\max,i} \cdot [0.32\ 1\ 0.995\ 0.992\ 0.991\ 0.99\ 0.989\ 0.989\ 0.988\ 0.988]^{\top} \\ \qquad \text{if } (s_{i,t-1} = 0 \wedge s_{i,t} = -1) \vee (s_{i,t-1} = -1 \wedge s_{i,t} = 0) \ , \\ T_{\min,i} \cdot [0.37\ 1\ 1.14\ 1.16\ 1.16\ 1.16\ 1.16\ 1.16\ 1.16\ 1.16]^{\top} \\ \qquad \text{if } (s_{i,t-1} = 0 \wedge s_{i,t} = 1) \vee (s_{i,t-1} = 1 \wedge s_{i,t} = 0) \ . \end{cases}
$$

$$(6.28)$$

The numeric values of (6.28) were estimated from simulation results with a population of EWHs without external control.

In this algorithm, the thermal energy content is $R_{i,t} = SOC_{i,t}$ and it is calculated using the temperature estimates $\mathbf{T}_{i,t}^{\text{est}}$. With this information, the sets $\mathcal{N}_{\text{sw,off},t}$ and $\mathcal{N}_{\text{sw,on},t}$ are calculated applying (6.18) and (6.19), whereas the switching signals are computed with (6.23).

6.4.4   Aggregate Power Feedback Control (C4)

Similarly to C1, in strategy C4 the central controller receives measurements of $P_{\text{ref},t}$ and $P_{\text{agg},t}$, but it has no information about the on/off state or temperature of individual EWHs. The main difference with C1 is that in strategy C4 bidirectional switching actions are possible by broadcasting control signals to all EWHs. Due to the absence of state feedback from individual EWHs, the controller in strategy C4 has a model for each EWH and historical data of the hourly water consumption to estimate $\mathbf{T}_{i,t}^{\text{est}}$, similarly to strategy C3.

At each time step, the central controller computes ΔP_t with (6.16). The controller sets upper and lower limits on SoC ($SOC_{\max}$ and $SOC_{\min}$) to avoid using very hot or very cold devices for SFC. Based on these limits, $\mathcal{N}_{\text{cold},t}$ and $\mathcal{N}_{\text{hot},t}$ are defined similarly to strategies C2e and C2f in (6.20) and (6.21), whereas $\mathcal{N}_{\text{db},t}$ is defined as in (6.11). All EWHs

---

[2]$\mathbf{T}_{i,t}^{\text{est}}$ denotes the temperature distribution along the vertical axis of the hot water storage tank.

in $\mathcal{N}_{\text{cold},t}$ are expected to switch on with a total power $P_{\text{cold},t}$, whereas all EWHs in $\mathcal{N}_{\text{hot},t}$ are expected to switch off with a total power $P_{\text{hot},t}$. The controller calculates $\Delta P_{\text{ef},t}$ using

$$\Delta P_{\text{ef},t} = \Delta P_t + P_{\text{hot},t} - P_{\text{cold},t} \ . \tag{6.29}$$

Equation (6.29) is similar to (6.17) with the difference that $P_{\text{cold},t}$ and $P_{\text{hot},t}$ are estimates of the aggregate power, instead of the actual values, because there is no real-time feedback from the EWHs to the controller.

The EWHs in the set $\mathcal{N}_{\text{db},t}$ are ranked in descending SoC order if $\Delta P_{\text{ef},t} < 0$, and in ascending order if $\Delta P_{\text{ef},t} > 0$. The controller builds the list of the EWHs that have to be switched by finding the number of devices $N_{\text{d},t}$ that minimizes (6.18) or (6.19). The SoC of the last EWH that enters the list is called "threshold SoC" and is denoted by $SOC_{\text{th},t}$. The values of SOC_{min}, SOC_{max} and $SOC_{\text{th},t}$ are broadcasted to all devices along with an auxiliary binary signal defined as

$$s_t = \begin{cases} 1 & \text{if } \Delta P_{\text{ef},t} > 0 \\ 0 & \text{if } \Delta P_{\text{ef},t} < 0 \end{cases} \ . \tag{6.30}$$

Each EWH responds to the broadcasted signal according to (6.31), where $u_{i,t}$ denotes the actual on/off state of EWH i at time t.

$$u_{i,t} = \begin{cases} \text{on} & \text{if } (s_t = 1 \wedge SOC_{i,t} \leq SOC_{\text{th},t}) \vee SOC_{i,t} \leq SOC_{\text{min}} \\ \text{off} & \text{if } (s_t = 0 \wedge SOC_{i,t} \geq SOC_{\text{th},t}) \vee SOC_{i,t} \geq SOC_{\text{max}} \\ u_{i,t-1} & \text{otherwise} \end{cases} \ . \tag{6.31}$$

6.5 Evaluation of Control Strategies

In this section, we evaluate the proposed rule-based control strategies with respect to the control accuracy, device operation and user comfort. We quantify the control accuracy using the Root Mean Squared Error (RMSE) defined as

$$\text{RMSE} = \sqrt{\frac{1}{N_{\text{sim}}} \cdot \sum_{t=0}^{N_{\text{sim}}-1} \left(P_{\text{agg},t} - P_{\text{ref},t} \right)^2} \ , \tag{6.32}$$

where N_{sim} is the simulation horizon. We quantify the effect on user comfort by computing the amount of time that an EWH operates outside the temperature deadband. Finally, we quantify the effect of SFC on the EWHs by counting the number of switching actions per device.

We run simulations with a population of 500 EWHs and the goal of tracking a 24-hour target power profile with a resolution of 10 seconds. Two different case studies are considered: a constructed target power profile and a target profile based on an actual SFC signal.

6.5.1   Constructed Target Power Profile

In this case, the power profile includes ramps, step increases and decreases, as well as a sinusoidal part, and it is shown in Fig. 6.3. This figure shows also the tracking performance of the four control strategies using the variation C2e for strategy C2.[3] The tracking RMSE, effect on user comfort and number of switching actions for the four strategies (including all variations of strategy C2) are shown in Fig. 6.4.

None of the strategies manages to fully track the sinusoidal part of the signal, not even strategy C2e that has full state feedback from each EWH in the population. The reason is that the target power profile lies above the population's baseline (shown with the green curve in Fig. 6.3) for most of the day. Therefore, the thermal energy storage of the population is full, i.e., all EWHs are at the higher end of the temperature deadband, and there is no more controllability for SFC.

Disregarding the sinusoidal part, the stochastic blocking strategy C1 tracks the target profile accurately and results in a small RMSE. However, it does not perform well with respect to user comfort, since approximately 40% of the devices do not have warm water for at least 30 minutes during the day.

Among the variations of strategy C2, the lowest tracking error is achieved by C2d. However, this comes with a high impact on customer comfort because approximately 35% of the EWHs do not have warm water for more than 30 minutes. Strategies C2a, C2b and C2c demonstrate the same performance in terms of tracking error and user comfort. An interesting result is the superior performance of strategies C2e and C2f

---

[3]There is an initial transient phase at the beginning of the simulation period in Fig. 6.3 and Fig. 6.5 due to the random initialization of EWH states. This transient period is disregarded when evaluating the control performance.
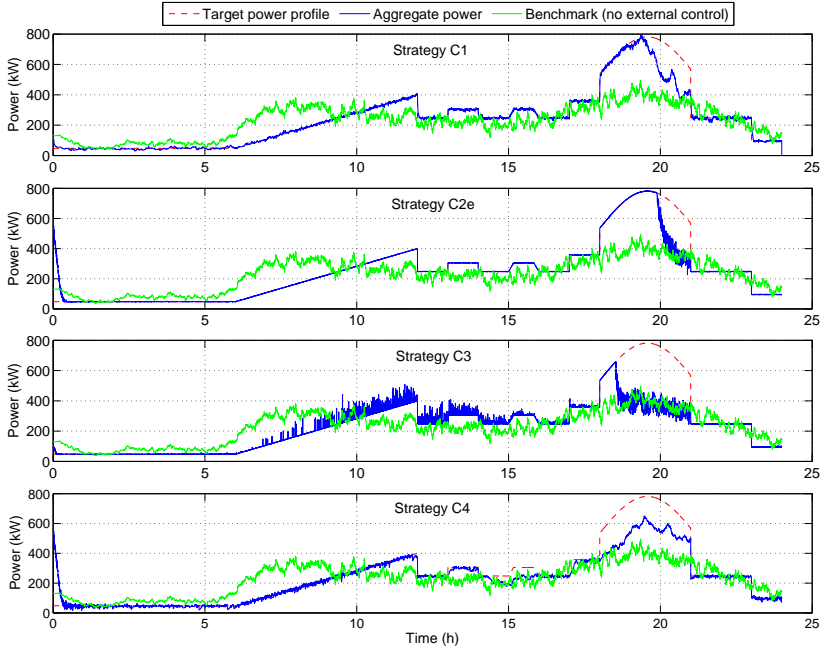
Figure 6.3: Tracking performance of strategies C1, C2e, C3 and C4 when using a constructed target power profile.

over strategies C2a, C2b and C2c. Therefore, using the SoC defined in (6.8), instead of the relative energy content defined in (6.5) and (6.6), reduces the tracking error significantly with very little compromise on user comfort. Comparing strategies C2e and C2f, we conclude that the parameters SOC_{min} and SOC_{max} can be tuned to achieve the desired tradeoff between tracking accuracy and customer comfort. Strategy C2e is more aggressive than C2f (due to a lower SOC_{min} value) and results in a lower tracking error, but with a higher user discomfort.

Strategy C3 respects the control signals from the EWH internal thermostatic controllers, therefore the effects on user comfort are negligible. Due to the absence of continuous temperature feedback, the tracking quality deteriorates resulting in power spikes especially during the ramp-up part of the target power profile.

It is interesting to note that strategy C4 has a much lower RMSE compared with C3, which is counterintuitive since the controller in C4 re-
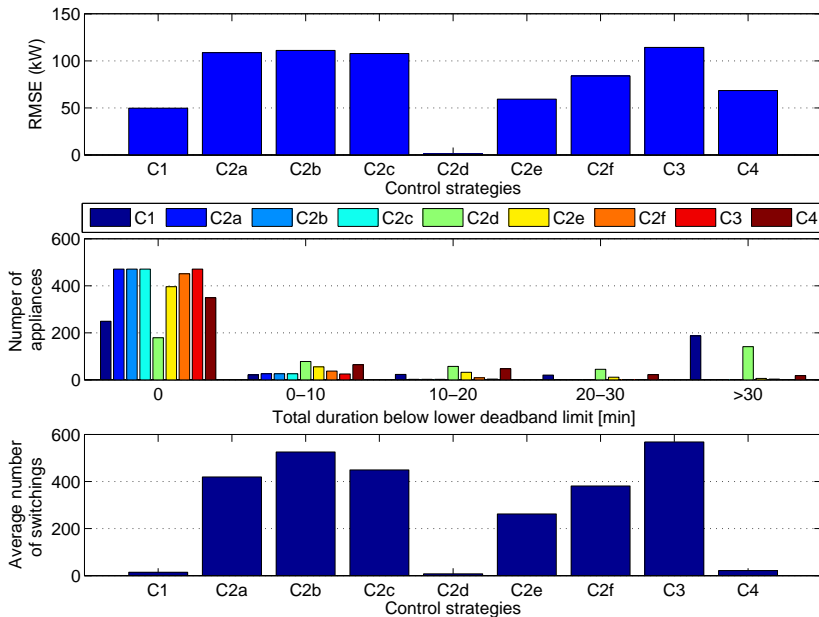
Figure 6.4: Comparison of control strategies with respect to tracking performance, user comfort, and device operation using a constructed target power profile.

ceives less information in real time. This is because there are no spikes in the aggregate power consumption in strategy C4.

According to Fig. 6.4, strategies C1, C2d and C4 result in the lowest average number of switching actions, whereas strategy C3 in the highest. Strategy C3 increases the number of switching actions because the decisions of the central controller are very sensitive to temperature estimation errors. Note that the direct temperature feedback control approaches C2e and C2f, which are based on the SoC defined in (6.8), reduce the number of switching actions compared with strategies C2a-C2d, which use the relative energy content defined in (6.6) or (6.5).

6.5.2   Target Power Profile based on a SFC Signal

To assess the algorithms' performance in a more realistic case, we repeat the simulations for strategies C1, C2e, C3 and C4 using an extract of

an actual SFC signal from the control area of Switzerland in 2009. The SFC signal for 24 hours is shown with the black curve in Fig. 6.6 as a percentage of the control band, i.e., a value of $+100\%$ denotes full activation of positive SFC reserve. We obtain the target power profile for the EWH population applying the following procedure.

We calculate the daily energy consumption without external control E_{day}, and construct a baseline power profile P_{base} using E_{day} and the water draw probability profile of Fig. 6.1. We assume that the population can participate in SFC with hourly reserve bids and a control band defined as $P_{\mathrm{cb}} = \pm 20\% \cdot P_{\mathrm{base}}$ (both P_{base} and P_{cb} are vectors with dimension 24). We obtain the target power profile (reference power $P_{\mathrm{ref},t}$) by superimposing the reserve request on the base power profile

$$P_{\mathrm{ref},t} = P_{\mathrm{base}} + w_t \cdot P_{\mathrm{cb}} \ , \tag{6.33}$$

where w_t is the normalized SFC signal, and the value of P_{base} and P_{cb} depends on the hour of the day.

Figure 6.5 shows the tracking performance, Fig. 6.6 the relative instantaneous tracking error, and Fig. 6.7 the evaluation results for the four control strategies. Note that in case of control without temperature feedback (strategies C1 and C4), there are intervals with very high tracking errors that can exceed 50%. As expected, strategy C2e achieves the lowest tracking error with a normalized RMSE of $= 1.18\%$. Strategy C3 performs significantly better with the actual SFC signal than the constructed target profile. In this case, the maximum relative error is approximately 30%, whereas the normalized RMSE is 8.7%. Note also that the impact of both C2e and C3 on customer comfort is minimal.

Based on these results, the direct temperature feedback control strategy C2 is in principle suitable for SFC with aggregations of EWHs, but it requires more communication than the other strategies. Indirect temperature feedback control (strategy C3) demonstrates a limited potential for SFC applications. From Fig. 6.5 and Fig. 6.6, one can observe that larger tracking errors with strategy C3 occur during the two peak power intervals, i.e., between hours $8 - 10$ am and $6 - 7$ pm. This indicates that the performance of strategy C3 decreases with larger control bands.

Based on our simulations, the stochastic blocking strategy C1 and the aggregate power feedback strategy C4 are not very appropriate for SFC. Strategy C4 might lead to inaccurate SFC signal tracking because the control decisions rely on EWH temperature information, even though
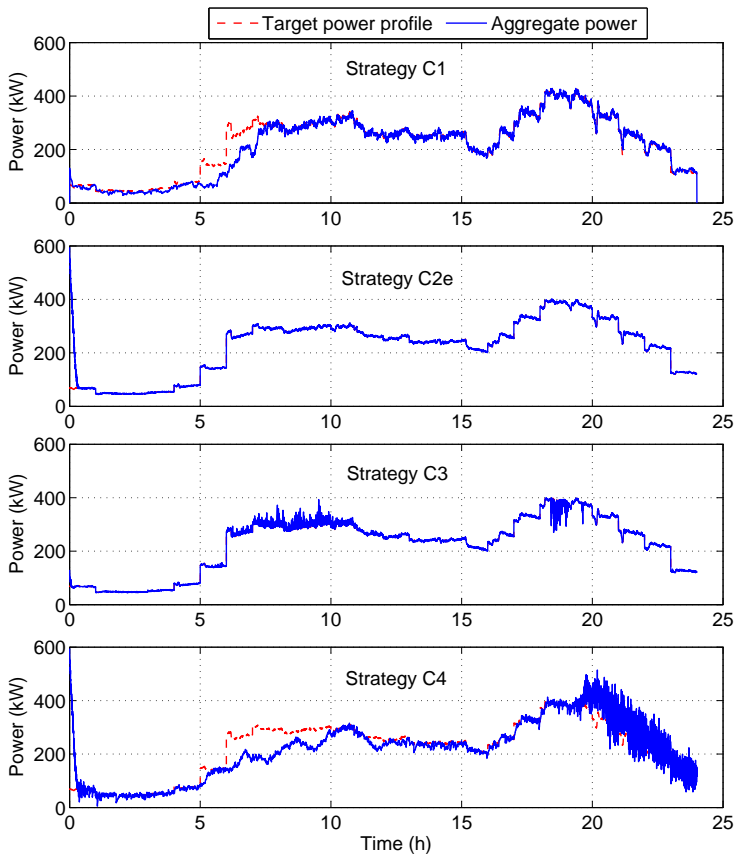
Figure 6.5: Tracking performance of strategies C1, C2e, C3 and C4 when
using a target power profile based on an actual SFC signal.

there is no feedback from the loads. On the other hand, strategy C1
generally achieves good tracking performance[4], but its limitation is that
control decisions are taken randomly and without considering the EWH
states. Although a relevant probabilistic control approach was success-
fully used for PFC in Chapter 5, strategy C1 results in significant user

---

[4]The tracking performance of strategy C1 is low during the interval $5 - 7$ am
when step changes in the target power profile occur due to large changes in the
aggregate water draw. Since these step changes are due to the hourly resolution of
P_{base}, better tracking performance is expected if a smoother P_{base} profile is used.
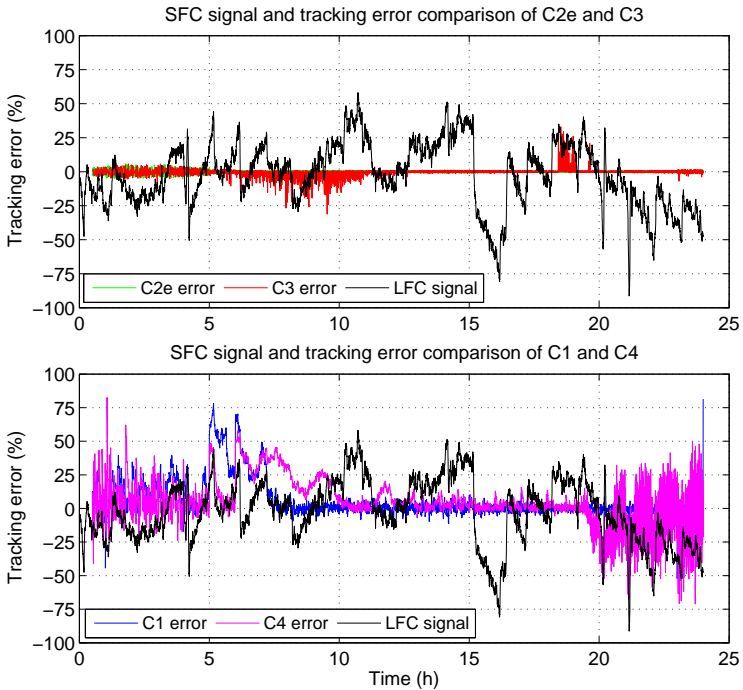
Figure 6.6: Normalized SFC signal and tracking error for strategies C1, C2e, C3 and C4.

discomfort (see Fig. 6.4 and Fig. 6.7) because, in contrast to the frequency deviation signal, the SFC signal is not zero-mean.

6.6 Conclusion

In this chapter, we presented four rule-based control strategies to manage the aggregate power consumption of a population of EWHs. The control strategies differ with respect to complexity and communication requirements from individual EWHs to the central controller, and their suitability for SFC was evaluated using the stratified EWH thermal model of Chapter 2.2.3.

Our simulation results show that information about EWH states, namely temperature and on/off state, is necessary to track the SFC signal accu-
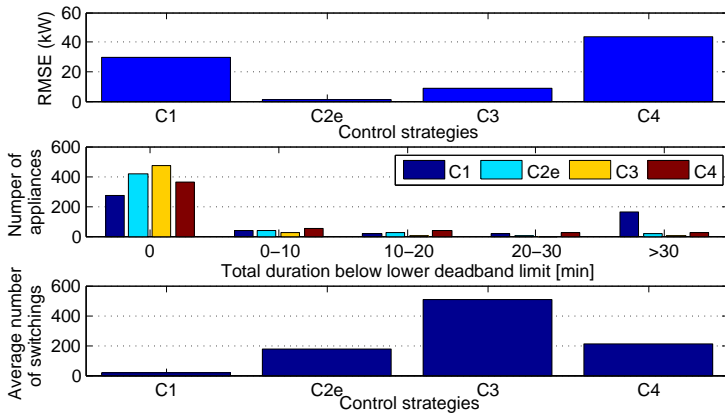
Figure 6.7: Comparison of control strategies C1, C2e, C3 and C4 with respect to tracking performance, user comfort and device operation using a target power profile based on an actual SFC signal.

rately. This information can be available through dedicated communication links between each EWH and the central controller. If information is available in real time, the control performance is very high, especially if the SoC of each EWH is defined taking into account the temperature distribution within the water tank. To reduce the communication burden, information may be transmitted from the EWHs to the central controller only when temperature deadband crossing events occur. The reduced communication leads to a relatively small reduction in the control performance.

SFC approaches with EWHs based only on aggregate power feedback might result in poor tracking performance or significant user discomfort. However, such approaches might be appropriate for other power system applications with lower accuracy requirements, e.g., load shifting to low-price periods.

Chapter 7

Combined Secondary Frequency Control and Voltage Regulation

7.1   Introduction

In Chapter 6 we presented methods to provide Secondary Frequency Control (SFC) reserves using the flexibility of Electric Water Heaters (EWHs). Apart from system-wide frequency deviations, supply-demand imbalances due to large shares of Renewable Energy Resources (RES) are expected to increase the voltage deviations in Distribution Networks (DNs). Traditionally, the DNs voltages are controlled using tap changers in distribution transformers or by reactive power compensation.

In this chapter, we develop a method that allows an aggregation of Thermostatically Controlled Loads (TCLs) to provide SFC reserves and at the same time regulate the DN voltage. Multi-tasking with controllable loads, i.e., the concurrent provision of two or more Ancillary Services (AS) to the grid, will leverage the full potential of demand-side resources for power system control tasks. On the other hand, offering each service separately may have adverse effects due to possibly conflicting objectives.

Table 7.1: Nomenclature of Chapter 7: symbols

Symbol	Unit	Description
A_l	m^2	Total external EWH surface
c	J/kg$^\circ$C	Specific heat capacity of water
c_{el}	€/MWh	Electricity cost
D	m^3/s	Water draw
$\mathbf{d}$	-	Vector of rated power of EWHs in $\mathcal{N}_{\mathrm{CL}}$
$E_{\mathrm{wc}}^{\mathrm{d}}$	%	Worst case down-regulation energy (% of control band)
$E_{\mathrm{wc}}^{\mathrm{u}}$	%	Worst case up-regulation energy (% of control band)
m	kg	Water mass
$\mathcal{N}_{\mathrm{CL}}$	-	Set of EWHs within their deadband
N_{h}	h	Horizon of day-ahead scheduling
n_{ap}	-	Number of EWHs
n_{b}	-	Number of buses
P	W	Active power
P_{n}	W	Rated active power
$P_{\mathrm{n,a}}^{\mathrm{tot}}$	W	Aggregated active power rating
Q	Var	Reactive power
S	VA	Apparent power
$\mathcal{S}$	-	Set of EWHs connected to a substation
$\mathbf{s}^-, \mathbf{s}^+$	-	Vectors of slack variables
T	$^\circ$C	Temperature
U	W/(m$^2 {}^\circ$C)	Inverse of tank thermal resistance
u	-	On/off state
V	V	Voltage
w_1, w_2	-	Weighting factors
$\mathbf{w}$	-	Allocation factor of reserve energy per bus
α	%	Normalized SFC reserve capacity
Δt	s	Simulation time step
δ	degrees	Phase angle
η	-	EWH efficiency
ρ	kg/m^3	Water density

For instance, consider the case where a load aggregation situated at a Medium-Voltage (MV) network with a large Photovoltaic (PV) penetration participates in SFC. In cases when down regulation is requested for a prolonged period of time, the reduced power consumption of the aggregation in combination with a large PV production may lead to overvoltages in certain network locations. In this case, a more sophisticated reserve allocation algorithm that takes into account the DN topology could minimize the adverse effects. This is the motivation of this chapter: the development of an algorithm that allows a TCL aggregation to provide frequency and voltage regulation simultaneously.

Although many works investigated how TCL aggregations can provide either voltage or frequency control, the concurrent provision of both services is not explored in the literature. However, relevant approaches

Table 7.2: Nomenclature of Chapter 7: subscripts and superscripts

Symbol	Description
CL	Controllable load (EWH)
cb	SFC control band
G	Generator
i	Bus index
in	Incoming water temperature
inj	Power injection
j	EWH index at a specific bus
L	Uncontrollable load
min/max	Minimum/maximum value of a variable or parameter
out	Ambient temperature
P	Active power
Q	Reactive power
t	Time index
tp	Target power (reference)
0	Initial value of a variable

have been proposed for electric vehicles. In [154], a co-optimization of smart-charging and frequency regulation for an aggregation of electric vehicles was proposed. It consists of a day-ahead multi-period DC Optimal Power Flow (DC-OPF) for scheduling, and a decentralized approach for the real-time dispatch of the regulation signal. The combined provision of frequency and voltage regulation by electric vehicle aggregations was explicitly addressed in [155], but without including the DN constraints.

The contribution of this chapter is the development of a hierarchical algorithm that enables multi-tasking with EWHs.[1] The algorithm consists of a day-ahead scheduling phase, which determines the optimal load dispatch by solving a multi-period AC-OPF, and a real-time operation phase to dynamically allocate setpoints to the available resources. The setpoints satisfy specific requirements for SFC provision, RES integration, and user comfort. Although the underlying optimization problems are demanding, the algorithm is designed in a computationally efficient way. Moreover, the algorithm is robust to uncertainties in demand and RES predictions, as well as model mismatches, and achieves a successful reserve provision exploiting statistical information about the SFC signal obtained from historical data.

The remainder of this chapter, which is based on the methods and re-

---

[1]Although EWHs are used in the simulations of this chapter, the hierarchical algorithm for combined frequency and voltage control is generic and applies also to other types of TCLs.

sults presented in [156], is organized as follows. In Section 7.2 simplified EWH models are introduced, Section 7.3 presents the control algorithm, Section 7.4 analyzes simulation results from a case study, whereas Section 7.5 concludes. The nomenclature of this chapter is summarized in Tables 7.1 and 7.2, and bold symbols indicate vectors and matrices.

7.2   Simplified EWH Modeling

In this section we introduce two simplified EWH models: the first model is for an individual device, whereas the second one is for a population of EWHs.

7.2.1   Individual EWH Model

An EWH is modeled as a hybrid system with a hysteresis controller based on a deadband. The temperature evolution is described by [29]

$$T_{t+1} = a_t T_t + b_t u_t + e_t \tag{7.1}$$

$$a_t = \exp(-\Delta t/RC) \, , \ b_t = \eta R P_n (1 - a_t) \tag{7.2}$$

$$e_t = (GRT_{out} + BRT_{in})(1 - a_t) \tag{7.3}$$

$$R = 1/(G + B) \, , \ B = \rho D_t c \, , \ G = A_l U \, , \ C = mc \, , \tag{7.4}$$

where T_t is the water temperature, $u_t \in \{0,1\}$ is the on/off state, Δt is the simulation time step, P_n is the rated EWH electric power, η is the EWH efficiency, T_{out} is the ambient temperature, T_{in} is the incoming water temperature, ρ is the water density, D_t is the water draw at time step t, c is the specific heat capacity of water, A_l is the total EWH surface, U is the inverse of the tank thermal resistance, and m is the mass of the contained water.

Although a more detailed stratified EWH model was developed in Chapter 2, it is not used here because it is overly complex for real-time optimization-based control.

7.2.2   EWH Population Model

Aggregate EWH models are needed for scheduling and control purposes, especially if the control involves solving optimization problems. Two

different aggregate models are considered in this chapter. The first one is obtained by simply stacking individual EWH models one on top of the other. If we denote by n_{ap} the total number of EWHs and apply (7.1)-(7.4) for each of them, we result in the following aggregate model

$$\mathbf{T}_{t+1} = \mathbf{A}_t \mathbf{T}_t + \mathbf{B}_t \mathbf{U}_t + \mathbf{E}_t \tag{7.5}$$

$$\mathbf{A}_t = diag\left([a_t^1 \ldots a_t^{n_{\mathrm{ap}}}]\right), \ \mathbf{B}_t = diag\left([b_t^1 \ldots b_t^{n_{\mathrm{ap}}}]\right) \tag{7.6}$$

$$\mathbf{E}_t = diag\left([e_t^1 \ldots e_t^{n_{\mathrm{ap}}}]\right) \tag{7.7}$$

$$\mathbf{T}_t = [T_t^1 \ldots T_t^{n_{\mathrm{ap}}}]^\top, \ \mathbf{U}_t = [u_t^1 \ldots u_t^{n_{\mathrm{ap}}}]^\top, \tag{7.8}$$

where the *diag* operator transforms a vector to a diagonal matrix with this vector in the diagonal.

If the model (7.5)-(7.8) with n_{ap} in the order of a few hundreds or thousands was included in a day-ahead multi-period AC-OPF, this would lead to a large scale nonlinear optimization problem with binary variables, which is computationally intractable even for relatively small network sizes. To render the optimization problem tractable, a simpler aggregate EWH model is needed. Construction of aggregate state-space models for heterogeneous populations of thousands of TCLs was studied in [3,157]. In this chapter, each aggregate model represents a few tenths or hundreds of EWHs connected to a MV/Low-Voltage (LV) substation, and therefore a different approach is required.

Let $\mathcal{S}^i$ denote the set of all EWHs connected to substation i. Each set is represented by an equivalent or "parent" EWH with a tank volume and a water draw equal to the sum of the volumes and the sum of the water draws of the "children" EWHs, respectively. The deadband limits of the parent EWH i, $T_{\min}^i$ and $T_{\max}^i$, are calculated as a weighted average of the deadband limits of the children EWHs, where their volumes are used as weighting coefficients. The state of parent EWH i at time step t is defined using the State of Charge (SoC) concept

$$\mathrm{SOC}_t^i := \frac{(T_t^i - T_{\min}^i)}{(T_{\max}^i - T_{\min}^i)}, \tag{7.9}$$

where T_t^i is its water temperature [21]. Let us define the input for each parent EWH as the total power consumption of the group, and denote it by $P_{\mathrm{CL},t}^i$. By defining $\mathbf{SOC}_t := [\mathrm{SOC}_t^1 \ldots \mathrm{SOC}_t^{n_{\mathrm{b}}}]^\top$ and $\mathbf{P}_{\mathrm{CL},t} := [P_{\mathrm{CL},t}^1 \ldots P_{\mathrm{CL},t}^{n_{\mathrm{b}}}]^\top$, where n_{b} is the number of buses of the MV network, the aggregated EWH dynamics can be written in the compact form

$$\mathbf{SOC}_{t+1} = \tilde{\mathbf{A}}_t \mathbf{SOC}_t + \tilde{\mathbf{B}}_t \mathbf{P}_{\mathrm{CL},t} + \tilde{\mathbf{E}}_t \ , \qquad (7.10)$$

where $\tilde{\mathbf{A}}_t$, $\tilde{\mathbf{B}}_t$ and $\tilde{\mathbf{E}}_t$ are appropriate matrices calculated applying (7.2)-(7.4) for the parent EWHs. Model (7.10) was compared to model (7.5) to assess its accuracy in terms of predicting the aggregated SoC of EWH groups at each substation. The Mean Absolute Percentage Error (MAPE) between the two models is 9.7%.

7.3   Scheduling and Control Algorithm

We consider a population of EWHs managed by an aggregator in order to participate in the SFC market. The same aggregator is responsible for dispatching the EWHs in a way that maximizes RES penetration without stresses on the local DN. Although the voltage and power flow constraints of the MV network are explicitly modeled, the underlying LV networks are not considered in this chapter.

Two-way real-time communication is assumed between the aggregator and each EWH through direct wired or wireless communication links. These links can be used to send control signals and receive information about EWH temperatures and on/off states. Although alternative approaches based on signal broadcasting exist [3, 21, 104, 154], their applicability in our problem is limited since they cannot easily handle the requirement of spatial allocation of the control actions. Note that the requirement for real-time load measurements can be relaxed using estimates, provided that a reliable state estimation method is available. State estimation for Demand Response (DR) applications will be the topic of Chapter 8.

The communication is assumed perfect, i.e., there are neither bandwidth constraints nor communication delays. Even in the presence of communication delays, the proposed algorithm can be effective in practice because the Transmission System Operators (TSOs) usually allow a sufficient time for the response to the SFC signal. For example, swissgrid, the Swiss TSO, allows a response time of 20 seconds [158].

The algorithm consists of two phases: a day-ahead scheduling phase that determines the optimal EWH dispatch at each substation, and a real-time control phase that allocates control actions to individual EWHs. A schematic of the algorithm is shown in Fig. 7.1.
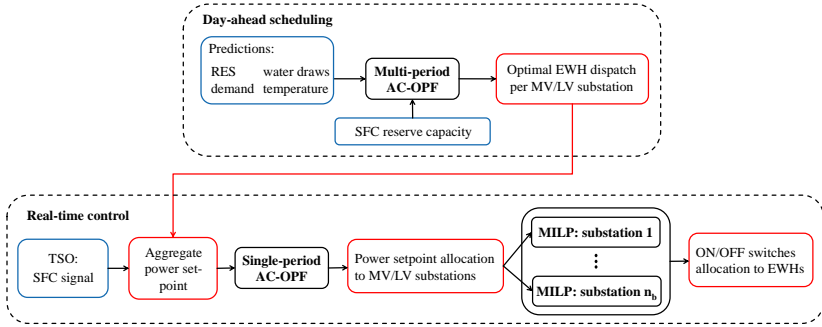
Figure 7.1: The scheduling and control algorithm for combined provision of SFC reserves and voltage regulation in DNs with EWHs.

7.3.1 Day-Ahead Scheduling

At the end of each day the aggregator predicts the demand, RES production and water draws, and builds aggregate EWH models for each MV/LV substation in the form of (7.10). The aggregator also knows the SFC reserve capacity of the EWH aggregation, which is defined as a percentage $\alpha\%$ of the hourly scheduled power consumption.[2] All this information is used in a multi-period AC-OPF to determine the optimal EWH dispatch, i.e., the aggregate hourly EWH consumption at each MV/LV substation that minimizes electricity costs by maximizing local RES energy absorption, while leaving enough slack for provision of SFC reserves. The optimization problem is formulated as

$$\min_{\mathbf{P}_{\mathrm{G},t},\mathbf{P}_{\mathrm{CL},t},\mathbf{V}_t,\boldsymbol{\delta}_t} \sum_{t=1}^{N_{\mathrm{h}}} \mathbf{c}_{\mathrm{el,P},t}^{\top}\mathbf{P}_{\mathrm{G},t} + \mathbf{c}_{\mathrm{el,Q},t}^{\top}\mathbf{Q}_{\mathrm{G},t} \tag{7.11a}$$

$$\text{s.t.} \quad \mathbf{P}_{\mathrm{G},t} - \mathbf{P}_{\mathrm{CL},t} - \mathbf{P}_{\mathrm{L},t} = \mathbf{P}_{\mathrm{inj},t} \quad \forall\, t \tag{7.11b}$$

$$\mathbf{Q}_{\mathrm{G},t} - \mathbf{Q}_{\mathrm{L},t} = \mathbf{Q}_{\mathrm{inj},t} \quad \forall\, t \tag{7.11c}$$

$$\mathbf{P}_{\mathrm{G},t}^{\min} \leq \mathbf{P}_{\mathrm{G},t} \leq \mathbf{P}_{\mathrm{G},t}^{\max} \quad \forall\, t \tag{7.11d}$$

---

[2]An AS market with hourly reserve bids is implicitly assumed to facilitate the demand-side participation. The reserve capacity factor α shows the relative size of the hourly reserve capacity with respect to the uncontrolled hourly load consumption, and it is fixed to $\alpha = 40\%$ based on preliminary simulations. If the market requires the reserve providers to bid the same reserve capacity for each hour of the day, then the formulation (7.11) should be modified accordingly.

$$\mathbf{Q}_{\mathrm{G},t}^{\min} \leq \mathbf{Q}_{\mathrm{G},t} \leq \mathbf{Q}_{\mathrm{G},t}^{\max} \quad \forall\, t \tag{7.11e}$$

$$\mathbf{V}^{\min} \leq \mathbf{V}_t \leq \mathbf{V}^{\max} \quad \forall\, t \tag{7.11f}$$

$$\boldsymbol{\delta}_t(i) = 0 \quad \text{(for the slack bus),} \quad \forall\, t \tag{7.11g}$$

$$\mathbf{S}_t \leq \mathbf{S}^{\max} \quad \forall\, t \tag{7.11h}$$

$$\mathbf{SOC}_{t+1} = \tilde{\mathbf{A}}_t \mathbf{SOC}_t + \tilde{\mathbf{B}}_t \mathbf{P}_{\mathrm{CL},t} + \tilde{\mathbf{E}}_t \quad \forall\, t \tag{7.11i}$$

$$\mathbf{0} \leq \mathbf{P}_{\mathrm{CL},t} \leq \mathbf{P}_{\mathrm{CL}}^{\max} \quad \forall\, t \tag{7.11j}$$

$$\mathbf{SOC}_0 = \mathbf{SOC}_{N_{\mathrm{h}}} \tag{7.11k}$$

$$\alpha(\mathbf{1}^{\top}\mathbf{P}_{\mathrm{CL},t}) \geq \beta P_{\mathrm{n},a}^{\mathrm{tot}} \quad \forall\, t \tag{7.11l}$$

$$\mathbf{SOC}_{t+1} + \alpha E_{\mathrm{wc}}^{\mathrm{u}} \tilde{\mathbf{B}}_t \mathbf{w}(\mathbf{1}^{\top}\mathbf{P}_{\mathrm{CL},t}) \leq \mathrm{SOC}^{\max} \cdot \mathbf{1} \,\forall\, t \tag{7.11m}$$

$$\mathbf{SOC}_{t+1} - \alpha E_{\mathrm{wc}}^{\mathrm{d}} \tilde{\mathbf{B}}_t \mathbf{w}(\mathbf{1}^{\top}\mathbf{P}_{\mathrm{CL},t}) \geq \mathrm{SOC}^{\min} \cdot \mathbf{1} \,\forall\, t \tag{7.11n}$$

$$(1+\alpha)(\mathbf{1}^{\top}\mathbf{P}_{\mathrm{CL},t}) \leq P_{\mathrm{n},a}^{\mathrm{tot}} \quad \forall\, t \,, \tag{7.11o}$$

where n_{l} is the number of lines in the network, N_{h} is the number of steps for the day-ahead AC-OPF ($N_{\mathrm{h}} = 24$ if hourly time steps are used), $P_{\mathrm{n},a}^i$ is the total installed power of the EWHs connected to substation i, and $P_{\mathrm{n},a}^{\mathrm{tot}}$ is the total installed power of all EWHs in the DN. In (7.11a)-(7.11o), $\mathbf{P}_{\mathrm{G},t} \in \mathbb{R}^{n_{\mathrm{b}}}$ and $\mathbf{Q}_{\mathrm{G},t} \in \mathbb{R}^{n_{\mathrm{b}}}$ are the generation active and reactive power vectors, respectively; $\mathbf{P}_{\mathrm{L},t} \in \mathbb{R}^{n_{\mathrm{b}}}$ and $\mathbf{Q}_{\mathrm{L},t} \in \mathbb{R}^{n_{\mathrm{b}}}$ are the uncontrollable active and reactive power demand vectors, respectively; $\mathbf{P}_{\mathrm{CL},t} \in \mathbb{R}^{n_{\mathrm{b}}}$ is the controllable active power demand vector (EWH aggregation); $\mathbf{P}_{\mathrm{inj},t} \in \mathbb{R}^{n_{\mathrm{b}}}$ and $\mathbf{Q}_{\mathrm{inj},t} \in \mathbb{R}^{n_{\mathrm{b}}}$ are the active and reactive power injection vectors, respectively; $\mathbf{V}_t \in \mathbb{R}^{n_{\mathrm{b}}}$ and $\boldsymbol{\delta}_t \in \mathbb{R}^{n_{\mathrm{b}}}$ are the voltage magnitudes and angles, respectively; $\mathbf{S}_t \in \mathbb{R}^{n_{\mathrm{l}}}$ is the power flow vector; $\mathbf{c}_{\mathrm{el,P},t} \in \mathbb{R}^{N_{\mathrm{h}}}$ and $\mathbf{c}_{\mathrm{el,Q},t} \in \mathbb{R}^{N_{\mathrm{h}}}$ are the cost vectors for active and reactive power, respectively. Note that $\mathbf{P}_{\mathrm{G},t}^{\min}$, $\mathbf{P}_{\mathrm{G},t}^{\max}$, $\mathbf{Q}_{\mathrm{G},t}^{\min}$ and $\mathbf{Q}_{\mathrm{G},t}^{\max}$ are determined by the generator capabilities and $\mathbf{S}^{\max}$ is determined by the line properties. The voltage limits are fixed to $\mathbf{V}^{\min} = 0.9 \cdot \mathbf{1}^{\top}$ pu and $\mathbf{V}^{\max} = 1.1 \cdot \mathbf{1}^{\top}$ pu.

There are three types of constraints in (7.11). Constraints (7.11b)-(7.11h) are the standard AC-OPF constraints. In particular, constraints (7.11b) and (7.11c) are the power balance equations, (7.11d) and (7.11e) represent the generation limits, whereas (7.11f) and (7.11h) describe the voltage and thermal limits, respectively. The apparent power flow vector $\mathbf{S}_t$ in (7.11h) is a nonlinear function of the voltage magnitudes and angles, which is omitted here for simplicity but can be found in [159]. The problem formulation (7.11) assumes that the active power injection at each bus can be controlled by the aggregator, for example

by curtailing the available PV power.

Constraints (7.11i)-(7.11k) correspond to the controllable aggregation of EWHs: (7.11i) models the aggregate thermal dynamics, (7.11j) sets limits to the aggregated power, and (7.11k) requires that the SoC between two consecutive days is the same. In general, it suffices that $\mathbf{SOC}_0$ and $\mathbf{SOC}_{N_{\mathrm{h}}}$ are given, i.e., it is not necessary that they are equal.

Constraints (7.11l)-(7.11o) guarantee a secure provision of SFC reserves. Constraint (7.11l) requires that a minimum amount of reserves, defined as $\beta\%$ of $P_{\mathrm{n,a}}^{\mathrm{tot}}$, is procured throughout the whole day. The parameter β can be tuned to avoid offering a very small reserve capacity; of course, if $\beta = 0$, the aggregator is free to schedule zero capacity for some parts of the day. Constraint (7.11o) guarantees that the aggregator can increase the aggregate EWH power as much as the scheduled reserve capacity. To ensure that the aggregate EWH power can be decreased as much as the scheduled reserve capacity, it suffices to fix α to a value less than 1.

Constraints (7.11m) and (7.11n) ensure that the EWH virtual storage will not be full (depleted) even in the worst-case scenario of prolonged down (up) regulation.[3] We calculated the worst-case scenarios for prolonged up and down regulation using SFC data from the Swiss control area in 2009. We identified the intervals during which the SFC signal did not change sign for up to an hour, and calculated the corresponding energy requirements. Based on this analysis, the worst-case amounts of energy that can be requested within an hour during down and up regulation are equal to $E_{\mathrm{wc}}^{\mathrm{d}} = 96.2\%$ and $E_{\mathrm{wc}}^{\mathrm{u}} = 97.8\%$ of the control band, respectively.

There are two assumptions implicitly made by (7.11i), (7.11m) and (7.11n). The first assumption is that the effect of the worst-case reserve energy request during time step t does not carry over to the EWH aggregate SoC at time step $t + 1$. This assumes that if the worst-case up or down regulation is requested at time step t, then the SFC signal at time step $t + 1$ will be such that the aggregation's SoC will be restored to the nominal value without SFC. This will be the case if the SFC signal is energy-constrained and crosses zero sufficiently often. The second assumption is that the worst-case reserve energy is allocated to the MV/LV substations proportionally to their EWH installed load. This is performed in (7.11m) and (7.11n) by defining $\mathbf{P}_{\mathrm{CL}}^{\max} := [P_{\mathrm{n,a}}^1 \ldots P_{\mathrm{n,a}}^{n_{\mathrm{b}}}]^\top$ and $\mathbf{w} := \mathbf{P}_{\mathrm{CL}}^{\max}/P_{\mathrm{n,a}}^{\mathrm{tot}}$. In Chapter 9, we will

---

[3]Up-regulation denotes the increase of generator power or decrease of load power. Down-regulation is the decrease of generator power or increase of load power.

present a relevant reserve scheduling formulation for commercial building aggregations, where these two assumptions are dropped.

The values of SOC^{max} and SOC^{min} in (7.11m) and (7.11n) depend on the desired level of robustness to uncertainties introduced by the water draw predictions and the aggregate model mismatch. We can increase the robustness by reducing SOC^{max} and increasing SOC^{min}. Note that problem (7.11a)-(7.11o) is a multi-period AC-OPF, because constraint (7.11i) couples different time periods.

The outcome of the day-ahead AC-OPF consists of $\{\mathbf{P}_{G,t}^*, \mathbf{P}_{CL,t}^*, \mathbf{V}_t^*, \boldsymbol{\delta}_t^*\}$, $\forall\, t \in [1, N_h]$, which are the optimal values for each bus and each time step. $\mathbf{P}_{CL,t}^*$ is the baseline part of the target profile for the next day at each time step $t \in [1, N_s]$, where $N_s = N_h \cdot (3600/\Delta t)$. The variable part of this profile is determined during real-time operation by the normalized SFC signal $P_{SFC,t}$ and the control band of the EWH aggregation $P_{cb,t} = \alpha(\mathbf{1}^\top \mathbf{P}_{CL,t}^*)$. Therefore, for an arbitrary SFC signal the target power profile (reference) is

$$
\begin{aligned}
P_{tp,t} &= \mathbf{1}^\top \mathbf{P}_{CL,t}^* + P_{SFC,t} P_{cb,t} \\
&= (1 + \alpha P_{SFC,t})(\mathbf{1}^\top \mathbf{P}_{CL,t}^*) \quad \forall\, t \in [1, N_s] .
\end{aligned}
\tag{7.12}
$$

7.3.2   Real-Time Coordinated Control

During real-time operation, the aggregator receives measurements of the current RES production, the demand, the temperatures of EWHs, as well as the normalized SFC signal, and it calculates the power reference applying (7.12). Next, the aggregator needs to control the on/off states of the EWHs such that the aggregate power tracks the power reference. The on/off control actions must satisfy the following requirements: (a) the scheduled SFC reserves must be provided with nearly 100% reliability, (b) the adverse effects on the DN operation should be minimal, and (c) user comfort should be respected.

Ideally, a single-period AC-OPF for the whole MV network incorporating the dynamics of individual EWHs as in (7.5) would provide the optimal on/off control actions; however, the corresponding optimization problem would be intractable. To enable real-time implementation, we devise a two-step optimization-based method as an approximate scheme for the original problem. We address requirements (a) and (b) in the first step by solving a simplified version of the single-period AC-OPF

problem, while we address requirement (c) in the second step by solving a number of computationally tractable Mixed-Integer Linear Programs (MILPs). The two-step method, which is solved whenever the SFC signal is received (typically every $2 - 4$ seconds), is explained in the rest of this section.

Single-period AC-OPF

First, the aggregator divides the EWHs into subsets based on their SoC using the approach presented in Section 6.4.2. Specifically, the aggregator identifies the subset of EWHs at each MV/LV substation that will be automatically switched by their internal controllers because their temperature is outside the deadband. It also identifies the subset of EWHs that are within their deadbands and thus available for external control actions, which is denoted by $\mathcal{N}^i_{\mathrm{CL},t} = \{j \in \mathcal{S}^i \,|\, 0 \leq \mathrm{SOC}_j(t) \leq 1\}$ for substation i.

Second, the aggregator determines the aggregate EWH power at each substation that minimizes the electricity cost at the current time step by solving the following single-period AC-OPF problem

$$\min_{\mathbf{P}_{\mathrm{G},t}, \mathbf{P}_{\mathrm{CL},t}, \mathbf{V}_t, \boldsymbol{\delta}_t} \mathbf{c}^\top_{\mathrm{el,P},t} \mathbf{P}_{\mathrm{G},t} + \mathbf{c}^\top_{\mathrm{el,Q},t} \mathbf{Q}_{\mathrm{G},t} \tag{7.13a}$$

$$\text{s.t.} \quad \text{constraints (7.11b)} - \text{(7.11h)}$$

$$P_{\mathrm{tp},t} = \mathbf{1}^\top \mathbf{P}_{\mathrm{CL},t} \tag{7.13b}$$

$$\mathbf{0} \leq \mathbf{P}_{\mathrm{CL},t} \leq \mathbf{P}^{\max}_{\mathrm{CL},t} \,, \;\; \mathbf{P}^{\max}_{\mathrm{CL},t}(i) = \sum_{j \,\in\, \mathcal{N}^i_{\mathrm{CL},t}} P^j_{\mathrm{n}} \,, \tag{7.13c}$$

where P^j_{n} is the power rating of EWH j. Constraint (7.13b) ensures that the SFC reserves are provided successfully.

The outcome of the single-period AC-OPF is the set $\{\mathbf{P}^*_{\mathrm{G},t}, \mathbf{P}^*_{\mathrm{CL},t}, \mathbf{V}^*_t, \boldsymbol{\delta}^*_t\}$, where $\mathbf{P}^*_{\mathrm{CL},t}$ represents the optimal allocation of the power reference $P_{\mathrm{tp},t}$ among the MV/LV substations at the current time step. Problem (7.13) is a nonlinear optimization problem, but it does not involve integer variables because it uses the aggregate EWH model.

Independent MILPs

The algorithm further allocates $\mathbf{P}^*_{\mathrm{CL},t}$ to individual EWHs by solving the following MILP separately for each MV/LV substation i

$$\forall i \min_{\mathbf{u}_t, \mathbf{s}^-, \mathbf{s}^+} \quad w_1 |\mathbf{d}^\top \mathbf{u}_t - \mathbf{P}^*_{\mathrm{CL},t}(i)| + w_2(|\mathbf{s}^-| + |\mathbf{s}^+|) \tag{7.14a}$$

$$\text{s.t.} \qquad \mathbf{T}_{t+1} = \mathbf{A}_t \mathbf{T}_t + \mathbf{B}_t \mathbf{u}_t + \mathbf{E}_t \tag{7.14b}$$

$$\mathbf{T}^{\min} - \mathbf{s}^- \leq \mathbf{T}_{t+1} \leq \mathbf{T}^{\max} + \mathbf{s}^+ \tag{7.14c}$$

$$\mathbf{s}^- \geq \mathbf{0} , \, \mathbf{s}^+ \geq \mathbf{0} , \tag{7.14d}$$

where $\mathbf{u}_t \in \{0, 1\}$ is the vector of on/off control actions, $\mathbf{d}$ is the vector of the power ratings of the EWHs that belong to $\mathcal{N}^i_{\mathrm{CL},t}$, $\mathbf{A}_t$, $\mathbf{B}_t$ and $\mathbf{E}_t$ are taken from (7.5), $\mathbf{T}^{\min}$ and $\mathbf{T}^{\max}$ are the vectors of upper and lower limits of the deadbands, respectively, and w_1, w_2 are weighing factors. Note that problems (7.14) are independent across substations, and therefore can be solved in parallel.

We introduce the slack variables $\mathbf{s}^-$ and $\mathbf{s}^+$ to obtain soft constraints on temperature, which guarantee feasibility of the problem under all circumstances. The EWH states at the end of each time step are predicted internally in the optimization using (7.14b), i.e., the problem is basically an MPC with a prediction horizon of 2 steps. Since the absolute value of the slack variables is penalized in the objective function, solutions that would result in significant violations of the temperature limits or undesired on/off cycling can be excluded.

7.4   Performance of Control Algorithm

7.4.1   Case Study Description

We investigate the effectiveness of the proposed algorithm in simulation studies using the benchmark MV DN shown in Fig. 7.2 [2]. The network has a nominal voltage of 20 kV and consists of 11 buses and 10 lines. Although the network has a meshed structure, we assume a radial topology in our simulations (see Fig. 7.2). The network supplies a rural area with both residential and industrial customers and a peak demand of 3.55 MW. The network parameters are given in Appendix B.

A large PV penetration is assumed in the network. According to [160], the average PV potential in rural areas in Germany is 13.7 kWp for
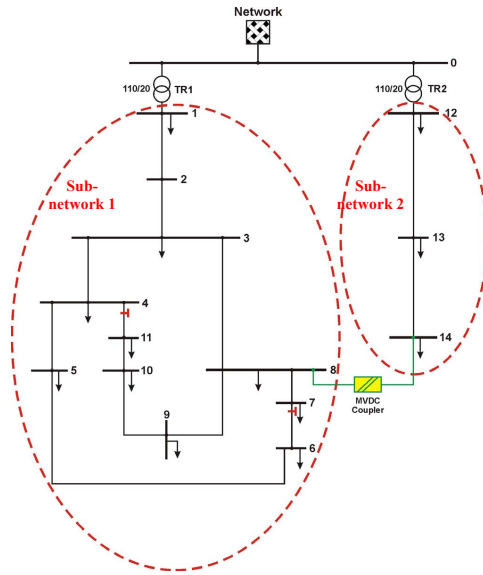
Figure 7.2: The benchmark MV grid used in the simulation studies (taken from [2]). Only subnetwork 1 is considered and the disconnectors between buses 4-11 and 6-7 are open, i.e., a radial grid topology is used.

residential buildings and 53.9 kWp for agricultural buildings. Based on this information, we consider an optimistic scenario of 1.5 MWp installed PV power at MV/LV substations 3, 4, 5, 8, 9 and 11, adding up to a total installed power of 9 MWp. The standard deviation of the PV day-ahead prediction errors is taken equal to 12% of the predicted value for each hour of the next day. Furthermore, we assume that the PV inverters operate at a constant unity power factor and do not control reactive power to contribute to voltage control.

In our simulations, we consider a population of 500 EWHs uniformly distributed among the MV/LV substations. We use the statistical method presented in Section 6.2 to generate the EWH population parameters and the water draw time series. We simulated the EWH population without external control actions and calculated the aggregate power consumption profile. We subtracted this profile from the total load profile to obtain the profile of the uncontrollable loads. We fix the two control parameters of problem (7.11) to $\alpha = 40\%$ and $\beta = 2\%$.

7.4.2   Simulation Results

Day-Ahead Scheduling

We start the analysis of the simulation results with the day-ahead scheduling phase of the algorithm considering the three cases of Table 7.3. Case A is the uncontrolled case where only constraints (7.11b)-(7.11h) are considered in the day-ahead AC-OPF and there is no real-time control. In case A, a part of the PV energy is curtailed if voltage deviations occur, whereas the EWHs operate based on their internal temperature controllers. In cases B and C, the EWHs are controlled to provide SFC reserves and at the same time regulate the DN voltages to minimize the curtailed PV energy. Different SOC^{min} and SOC^{max} values are used in cases B and C to illustrate the effect of different levels of robustness on algorithm performance.

Table 7.3: Simulation cases

Case	Description
A	No external control
B	External control with $SOC^{min} = 0.2$ and $SOC^{max} = 0.8$
C	External control with $SOC^{min} = 0.3$ and $SOC^{max} = 0.7$

Figure 7.3 shows the optimal EWH dispatch, Fig. 7.4 shows the resulting SoC profile, and Fig. 7.5 shows the total PV power that can be injected in the grid without violating the network constraints. To maximize the PV integration, a major part of the EWH consumption is shifted to hours $13.00 - 15.00$, when the available PV energy is at a maximum. Note that at hour 13.00 the scheduled EWH power in cases B and C is more than 2 times higher than that of case A. The significant energy consumption in the interval $20.00 - 24.00$ is not only due to the large amount of water draws, but also due to constraint (7.11k). The scheduled SoC keeps far from SOC^{min} and SOC^{max} to satisfy constraints (7.11m) and (7.11n). In case C the SoC keeps closer to 50%, which provides additional robustness to aggregate EWH modeling errors. In cases B and C there are periods of time when the scheduled power is equal to the minimum allowed by constraint (7.11l) (e.g., hours $01.00 - 07.00$).

Table 7.4 compares the absorbed PV energy in the three cases. The day-ahead scheduling algorithm increases the PV energy yield by 3.80%
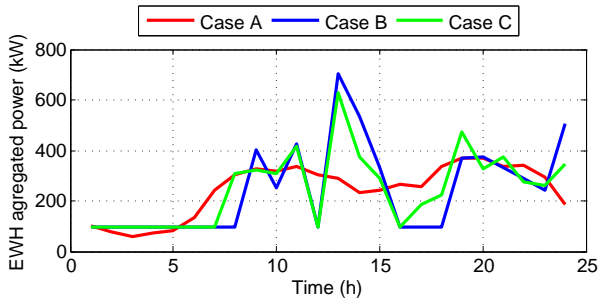
Figure 7.3: Day-ahead optimal dispatch of the EWH aggregation.
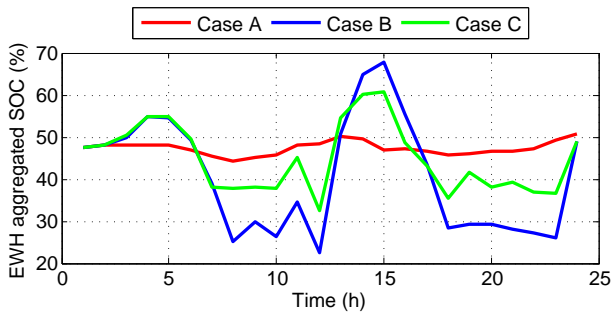
Figure 7.4: Day-ahead optimal SoC of the EWH aggregation.

and 3.04% in cases B and C, respectively, compared with case A. This is achieved with an increase of only 0.34% and 1.53% in EWH energy consumption for cases B and C, respectively.

Real-Time Operation

In the rest of this section, we present results from the real-time operation algorithm during a typical day. Figure 7.6 shows the SFC signal requested from the EWH aggregation as a percentage of its control band.[4] Figure 7.7 shows the corresponding target power trajectory (reference) and the aggregate EWH power of the whole MV network for cases B and C. The MAPE between the target trajectory and the aggregate

---

[4]Although the SFC signal is available with a resolution of 10 seconds, we downsample it to 1 minute in the simulations to decrease the computation time.
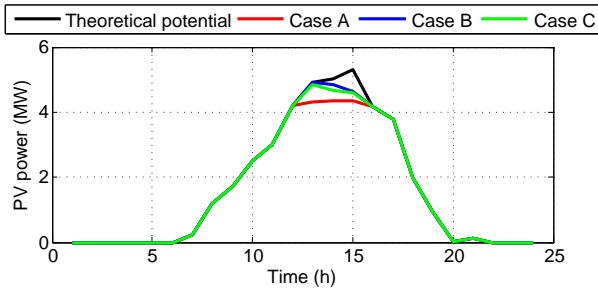
Figure 7.5: Day-ahead optimal dispatch of PV.

EWH power is given in Table 7.4, and the tracking accuracy is very high in both cases. In Case B, the tracking error increases during hours 7 and 11 because the energy storage level of the EWH aggregation is low. Observe that the tracking error increases whenever the aggregate SoC falls below 30% and the target power is much lower than the EWH aggregate power in the uncontrolled case A. On the other hand, Case C results in a lower MAPE due to the increased robustness to aggregate EWH modeling errors and water draw prediction errors.

Figure 7.8 shows the evolution of the aggregate EWH SoC per substation in case C. The solid curves correspond to buses with installed PV power, the dashed curves correspond to buses with load only, whereas the black curve indicates the SoC of the whole aggregation. Although the black curve generally follows the day-ahead optimal SoC curve from Fig. 7.4, the two curves are not identical for two reasons: (a) SFC reserves are provided, and (b) the aggregate EWH model is not perfect. Observe that the EWHs connected to remote buses with installed PV power (e.g., buses 9 and 11) are at around 100% SoC during most of the daytime to avoid overvoltages.

To maximize PV integration, the contribution of remote buses with PV to up regulation is zero for many hours during daytime. In addition, since the SoC of these buses is close to 100% during most of the daytime, their contribution to down regulation is also limited. The maximum potential for up and down regulation for a remote bus with PV (bus 11) is shown in Fig. 7.9, and it is contrasted with the larger potential of a bus without PV (bus 7).

To assess the impact of the proposed algorithm on user comfort and device operation, we show the temperature deviations from the deadband
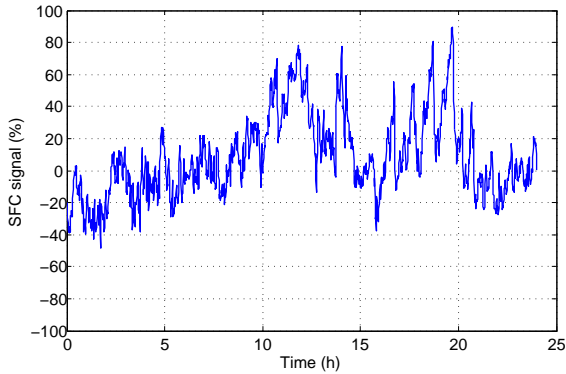
Figure 7.6: The SFC signal of the Swiss control area in 2009 during a sample day.

and the average number of switching actions in Fig. 7.10 and Table 7.4. In cases B and C, the worst case deviations from the lower and upper deadband limits are $-0.006°C$ and $+0.027°C$, respectively. The deviations are kept negligible by considering only the EWHs within the deadband in problem (7.13a)-(7.13c), and by formulating the predictive optimization problem (7.14). Although the user comfort is nearly always respected, many EWHs have to operate at their upper deadband limit for several hours per day to maximize the PV energy integration. This, in combination with the SFC provision, drastically increases the average number of switching actions per EWH, which might increase the wear of the devices. Furthermore, the switching actions are not uniformly distributed among the EWHs, but instead the EWHs that are connected to substations with PV switch more often. This distinction should be taken into account when designing remuneration schemes for the participants in such DR programs.

Table 7.4: Comparison of the controllers' performance

Case	Tracking MAPE (%)	PV energy (MWh)	EWH energy (MWh)	Switching actions
A	N/A	36.89	5.89	7
B	0.92	38.29	5.91	217
C	0.15	38.01	5.98	242
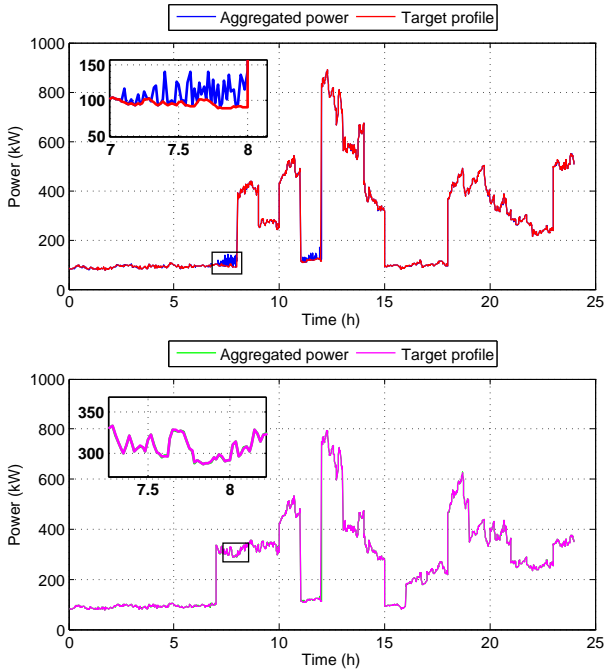
Figure 7.7: Tracking performance of the target power trajectory. The top plot corresponds to case B, and the bottom plot corresponds to case C.

7.4.3   Implementation Issues

Based on our results, SFC and active DN management for PV integration can be simultaneously provided by the same TCL aggregation using the proposed algorithm. The more robust the algorithm is, the higher the SFC tracking accuracy, but the less the PV energy yield and the higher the EWH consumption and the number of switching actions.

The computation time of the algorithm is low enough to allow a real-world implementation. We solve problems (7.11) and (7.13a)-(7.13c) using the solver IPOPT, whereas we solve problem (7.14) using the solver GLPK. We call both solvers through a MATLAB-YALMIP interface [43] using a 4 core machine (2.83 GHz) with 8 GB RAM. Problem (7.11) takes 20 minutes on average, problem (7.13a)-(7.13c) takes around 2 seconds, whereas each of the problems (7.14) is solved in ap-
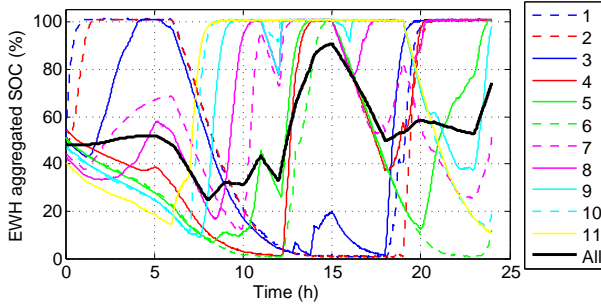
Figure 7.8: Evolution of the aggregated SoC per substation for case C.
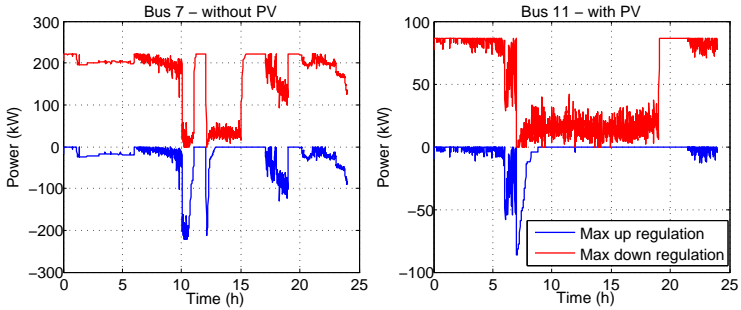
Figure 7.9: Maximum potential for up and down regulation for a bus without PV (left) and a bus with PV (right). Note that the two plots have different scales.

proximately 0.1 seconds. Note that single-period AC-OPF problems, like (7.13a)-(7.13c), can be solved using specialized packages, for example MATPOWER, in less than a second even for networks with hundreds of buses [161]. Moreover, problems (7.14) are independent and thus can be solved in parallel, which indicates the scalability of the approach for larger networks.

In our simulations, the residual $|\mathbf{d}^\top \mathbf{u}_t - \mathbf{P}^*_{\mathrm{CL},t}(i)|$ in (7.14a) is 0.025 kW on average with its maximum value being 1.256 kW. Therefore, problems (7.13a)-(7.13c) and (7.14) are a good approximation of the original real-time operation problem.
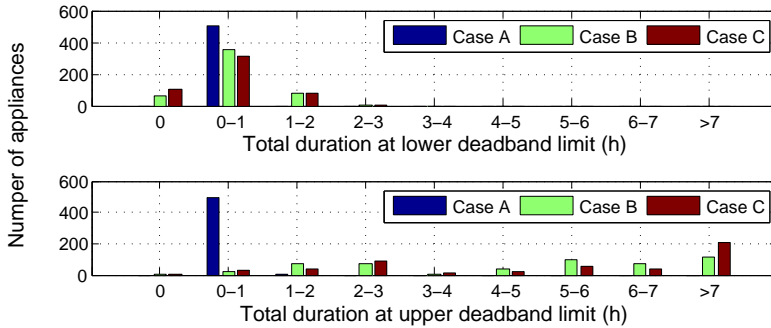
Figure 7.10: Effect of control algorithm on user comfort.

7.5   Conclusion

In this chapter, we presented a hierarchical control algorithm to enable multi-tasking with TCLs, namely frequency and voltage regulation, assuming a real-time, two-way communication infrastructure between an aggregator and individual loads. The scheduling phase of the algorithm determines the optimal dispatch of the loads and is robust to uncertainties related to prediction errors and model mismatches. The real-time operation phase allocates control actions to loads to securely provide frequency reserves while respecting DN and user comfort constraints. Moreover, the optimization problems are tractable and therefore the proposed algorithm is implementable for realistic network sizes. The algorithm can be used to provide SFC with TCL aggregations without adverse side-effects such as overvoltages in DNs with large RES shares.

Chapter 8

Reduced Communication Needs in Demand Response by State Estimation

8.1   Introduction

Chapter 6 illustrated the importance of communicating state measurements from Thermostatically Controlled Loads (TCLs) to the central controller when providing Secondary Frequency Control (SFC) reserves. However, real-time communication with every load entails significant implementation costs. A number of state estimation methods have been proposed to reduce communication costs, e.g., [3, 109–111]. Nevertheless, these approaches rely on aggregate load models and estimate the distribution of TCLs in a normalized temperature state space, which does not perfectly capture the states of individual loads.

In contrast to the above work, we attempt to estimate *individual TCL states*, i.e. temperatures and on/off modes. Estimating the states of individual TCLs instead of the states of aggregate models could improve the control performance. Individual state estimates help us better estimate the effect of TCLs' internal controllers and aggregator's external control actions on the aggregate power consumption of TCLs.

Table 8.1: Nomenclature of Chapter 8: symbols

Symbol	Unit	Description
C	kWh/$^\circ$C	Thermal capacitance
C_{p}	-	Coefficient of performance
M	$^\circ$C	Upper limit of temperature deadband
M_u, m_u	-	Upper and lower bound on u (for the MLD model)
M_x, m_x	-	Upper and lower bound on x_{c} (for the MLD model)
m	$^\circ$C	Lower limit of temperature deadband
$m_1 - m_5$	-	Weighting factors
N	-	Estimation horizon (number of steps)
N_{sim}	s	Simulation horizon
n_{ap}	-	Number of TCLs
n_{sub}	-	Number of substations
P	W	Active power
Q	$^\circ$C^2	Covariance matrix of process noise
R	$^\circ$C/kW	Thermal resistance
R	W^2	Covariance matrix of measurement noise
r_{s}	-	Bernoulli random variable with success probability P_{s}
r_{soc}	-	Uniform random variable
SOC_{th}	-	Threshold SoC value (control signal)
s	-	Auxiliary binary control signal
T_{db}	$^\circ$C	Thermostat deadband width
T_m	min	Period of TCL measurements
T_{sp}	$^\circ$C	Thermostat temperature setpoint
T_α	$^\circ$C	Ambient temperature
u	-	Fraction of rated power consumed at the on state
V	-	Known statistic of measurement noise
v	W	Measurement noise
W	-	Known statistic of process noise
w	$^\circ$C	Process noise
x_{c}	$^\circ$C	TCL temperature (continuous variable)
x_{l}	-	TCL on/off state (binary variable)
x_{th}	$^\circ$C	Temperature threshold corresponding to SOC_{th}
Y	-	Normalized SFC signal
z	-	Continuous auxiliary variable

Separate state estimation and control is not optimal for stochastic hybrid systems (such as TCLs), because the separation principle does not apply to such systems [162]. However, it is a common practice to heuristically separate the two tasks in the interest of simplicity. The control performance generally improves with better state estimates, but this must be verified through experiments or simulation. State estimation problems for stochastic hybrid systems are often solved with "multiple model" estimation schemes that involve a filter for each mode [163]. However, with large numbers of modes these approaches become intractable [164].

In this chapter we follow a different approach. First, we transform

Table 8.2: Nomenclature of Chapter 8: Greek letters

Symbol	Unit	Description
α	-	SFC control band
ΔP	W	Power change
ΔP_{int}	W	Power change due to thermostatic actions
Δt	s	Simulation time step
$\delta_1 - \delta_5$	-	Binary auxiliary variables
ε_{agg}	W	Aggregate power estimation error
$\varepsilon_{\text{on/off}}$	-	Estimation error of on/off states
ε_{soc}	-	SoC estimation error
ϵ	-	Tolerance for the MLD model
μ	-	Mean value of a random variable
σ	-	Standard deviation of a random variable

Table 8.3: Nomenclature of Chapter 8: subscripts and superscripts

Symbol	Description
agg	Aggregate power in the system
CL	Load cluster
i	TCL index
l	Substation index
meas	Measured power
n	Rated power (nominal)
set	Setpoint power (reference)
sub	Aggregate power at a substation
t	Time index
UL	Power of uncontrollable loads
$(\hat{\cdot})$	Estimate of a variable

our stochastic hybrid system into a Mixed Logical Dynamical (MLD) system [165], which allows us to represent a TCL aggregation as a system of mixed-integer linear inequalities. Second, we propose a Moving Horizon State Estimation (MHSE) method based on a Mixed-Integer Linear Program (MILP) to estimate the states of individual TCLs from aggregate power measurements and periodic TCL state measurements under process and measurement noise. Third, we combine the estimator with a scalable hierarchical broadcast controller to enable provision of SFC with minimal investment in new infrastructure. We investigate the performance of the proposed control-estimator loop with respect to estimation and control accuracy. Finally, we benchmark the proposed control-estimator loop against a controller with a simpler "predictor" considering different aggregation sizes, process and measurement noise levels, estimation horizons, and control trajectory characteristics.

The remainder of this chapter is organized as follows. Section 8.2 provides a high-level introduction to the problem and describes the control

and communication architecture. Section 8.3 introduces the MLD system model, Section 8.4 presents the controller, and Section 8.5 develops the MHSE scheme. Simulation results under different conditions are analyzed in Section 8.6 in order to benchmark the proposed estimator against the simpler predictor, and assess its usefulness in SFC. Finally, Section 8.7 concludes this chapter that is based on the methods and results presented in [166, 167]. The nomenclature of this chapter is introduced in Tables 8.1-8.3, and bold symbols indicate vectors or matrices.

8.2   Problem Description

We consider the problem of an aggregator managing a population of TCLs in a Distribution Network (DN) with the goal of providing power system Ancillary Services (AS). We assume that aggregate power measurements from MV/LV distribution substations are available to the aggregator at every time step. We also assume there exists a two-way communication network with bandwidth constraints. Specifically, the TCLs interact with home energy management systems, which in turn communicate with smart meters that transmit data to the aggregator. The TCL state measurements are obtained by the home energy management system with a high resolution (e.g., every 10 seconds [168][1]), but they are transmitted from the smart meters to the aggregator at a lower rate (e.g., every 15 minutes) due to communication network constraints. Note that TCL state measurements are sent in packets that include all past state measurements since the last transmission. We also assume that the control signals are broadcasted to all TCLs, i.e., TCLs are not individually addressed by the aggregator.

Our goal is to estimate the states of individual TCLs between consecutive state measurements with the overarching goal of improving control performance. This is a challenging task due to process and measurement noise. Process noise includes plant-model mismatch (TCL modeling errors) and errors in predictions of external forcing, such as ambient temperature and consumer behavior. Measurement noise includes errors in aggregate power measurements. To obtain aggregate power measurements, we subtract the predicted uncontrolled load from the measured

---

[1]Apart from enabling estimation of individual TCL states, high resolution monitoring of TCLs is also necessary for auditing purposes.
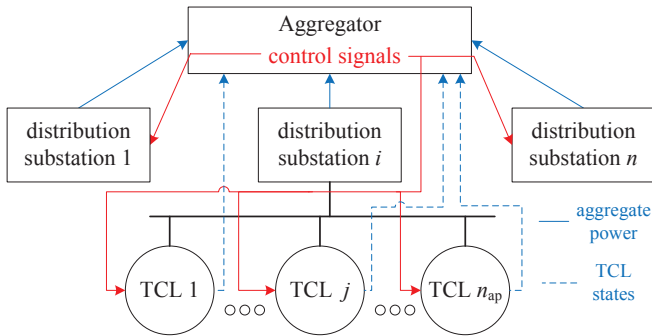
Figure 8.1: The proposed control and communication architecture that re-
lies on MHSE of individual load states.

load at the substation, which results in a noisy estimate of the aggregate
power consumption of the TCLs.

The proposed MHSE method is efficient and computationally tractable
for small TCL aggregations. In practice, larger TCL aggregations can be
managed by solving parallel MHSE problems at several substations but
as a single control problem. With this architecture, the proposed MHSE
method is scalable and allows us to offer a significant amount of AS. A
major difference between this work and the previous works [3,4,109,110]
is that the MHSE method efficiently utilizes measurements from MV/LV
substations, instead of only aggregate measurements from high-medium
voltage substations.

The control and communication architecture is shown in Fig. 8.1. Solid
lines represent high-frequency communication flows (e.g., every 10 sec-
onds), whereas dashed lines show low-frequency communication flows
(e.g., every 15 minutes). The blue lines represent the measurements,
and the red lines the control signals. Note that the control signal is
broadcast to all substations, whereas estimation is performed indepen-
dently for each substation.

The developed controller/estimation loop can cope with constrained
communication networks with limited bandwidth. In parallel to this
work, we developed a related method based on Model Predictive Control
(MPC) and Kalman filtering that can additionally handle communica-
tion delays and packet losses. The interested reader is referred to [169]
for relevant results.

8.3   Modeling

8.3.1   Individual TCL Modeling

Processes that evolve according to continuous dynamics, discrete dynamics, and logic rules can be modeled as hybrid systems [165]. Here, we use the two-state hybrid TCL model developed in [80, 170, 171]. Let us denote the TCL temperature at time step t by $x_{c,t} \in \mathbb{R}$ and the on/off state at time step t by $x_{l,t} \in \{0, 1\}$. A heating TCL's stochastic discrete-time dynamics can be expressed as

$$x_{c,t+1} = ax_{c,t} + bu_t x_{l,t} + fT_{\alpha,t} + w_t \,, \tag{8.1}$$

$$x_{l,t+1} = \begin{cases} 0 & \text{if } x_{c,t+1} \geq M \\ 1 & \text{if } x_{c,t+1} \leq m \\ x_{l,t} & \text{otherwise} \end{cases} , \tag{8.2}$$

where $a = e^{-\Delta t/(CR)}$, $b = (1-a)RC_p P_n$, $f = (1-a)$, Δt is the discretization time step, C is the thermal capacitance, R is the thermal resistance, C_p is the Coefficient of Performance (COP), P_n is the rated power, $u_t \in [0, 1]$ is the fraction of the rated power consumed by the TCL if it is on, $T_{\alpha,t}$ is the ambient temperature, and w_t is the process noise. Additionally, $M = T_{sp} + 0.5T_{db}$ and $m = T_{sp} - 0.5T_{db}$ are the upper and lower temperature deadband limits, respectively, where T_{sp} is the thermostat temperature setpoint and T_{db} is the deadband width.

The stochastic hybrid system (8.1)-(8.2) can be described using the MLD framework. Following the approach proposed in [165], we introduce the auxiliary binary variables $\delta_{1,t}$, $\delta_{2,t}$, $\delta_{3,t}$ and $\delta_{4,t}$ defined as

$$[\delta_{1,t} = 1] \leftrightarrow [x_{c,t} \geq M] \,, \tag{8.3}$$

$$[\delta_{2,t} = 1] \leftrightarrow [x_{c,t} \leq m] \,, \tag{8.4}$$

$$\delta_{3,t} = x_{l,t+1} \,, \tag{8.5}$$

$$\delta_{4,t} = x_{l,t} \,. \tag{8.6}$$

Additionally, since (8.1) is bilinear between $x_{l,t}$ and u_t, we introduce the auxiliary continuous variable $z_t = x_{l,t}u_t = \delta_{4,t}u_t$ to linearize the model. By defining $x_t := [x_{c,t} \; x_{l,t}]^\top$ and $\delta_t := [\delta_{1,t} \; \delta_{2,t} \; \delta_{3,t} \; \delta_{4,t}]^\top$, equation

(8.1) is rewritten as

$$x_{t+1} = \underbrace{\begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix}}_{A} x_t + \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{B_2} \delta_t + \underbrace{\begin{bmatrix} b \\ 0 \end{bmatrix}}_{B_3} z_t + \underbrace{\begin{bmatrix} f \\ 0 \end{bmatrix}}_{F_t} T_{a,t} + \begin{bmatrix} w_t \\ 0 \end{bmatrix}.$$

(8.7)

Note that (8.7) is similar to (11a) in [165], except that we include two additional terms: ambient temperature and process noise.

The TCL internal hysteresis controller (8.2) can be described by the logic relations

$$[\delta_{1,t} = 1] \rightarrow [\delta_{3,t} = 0] \ , \tag{8.8}$$

$$[\delta_{2,t} = 1] \rightarrow [\delta_{3,t} = 1] \ , \tag{8.9}$$

$$[\delta_{1,t} = 0] \wedge [\delta_{2,t} = 0] \rightarrow [\delta_{3,t} = \delta_{4,t}] \ , \tag{8.10}$$

which are transformed into the mixed-integer linear inequalities

$$E_2 \delta_t + E_3 z_t \leq E_1 u_t + E_4 x_t + E_5 \ . \tag{8.11}$$

Equation (8.11) describes the deadband crossing events, and therefore represents the switching actions imposed by the thermostatic hysteresis controller of each TCL. More specifically, the dynamics (8.7) together with the inequalities (8.11) fully describe the evolution of the on/off mode and temperature of a TCL. The matrices $E_1, E_2, E_3, E_4,$ and E_5 are defined as follows

$$E_1 = [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1 \ -1]^\top \ , \tag{8.12}$$

$$E_2 = \begin{bmatrix} 0 & M_x - m & 0 & 0 \\ 0 & m_x - m - \epsilon & 0 & 0 \\ M - m_x & 0 & 0 & 0 \\ M - M_x - \epsilon & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 0 & 1 & 0 \\ -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -M_u \\ 0 & 0 & 0 & m_u \\ 0 & 0 & 0 & -m_u \\ 0 & 0 & 0 & M_u \end{bmatrix} \ , \tag{8.13}$$

$$E_3 = [0 \; 0 \; 0 \; 0 \; 0 \; 0 \; 0 \; 0 \; 0 \; 0 \; 1 \; -1 \; 1 \; -1]^\top \; , \tag{8.14}$$

$$E_4 = \begin{bmatrix} -1 & 1 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}^\top , \tag{8.15}$$

$$E_5 = [M_x \; (-m - \epsilon) \; - m_x \; (M - \epsilon) \; 1 \; 0 \; 1 \; 0 \; 0 \; 0 \; 0 \; 0 \; - m_u \; M_u]^\top \; , \tag{8.16}$$

where $m_x = 0$ and $M_x = 100$ are lower and upper bounds on the continuous state $x_{c,t}$, $m_u = 1$ and $M_u = 30$ are lower and upper bounds on u_t, and $\epsilon = 10^{-6}$ is a small tolerance.

Simplification for Constant Power Loads

If the power consumption is constant when the TCL is on ($u_t = 1$), the auxiliary variable z_t is equal to $\delta_{4,t}$ and thus it is redundant. Therefore, (8.7) and (8.11) can be simplified to

$$x_{t+1} = \underbrace{\begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix}}_{A} x_t + \underbrace{\begin{bmatrix} 0 & 0 & 0 & b \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{B} \delta_t + \underbrace{\begin{bmatrix} f \\ 0 \end{bmatrix} T_{\alpha,t}}_{F_t} + \begin{bmatrix} w_t \\ 0 \end{bmatrix} , \tag{8.17}$$

$$\tilde{E}_1 \delta_t \leq \tilde{E}_2 x_t + \tilde{E}_3 \; , \tag{8.18}$$

where $\tilde{E}_2 = E_4$, $\tilde{E}_3 = E_1 + E_5$. The first three columns of $\tilde{E}_1$ are identical to the first three columns of E_2. The fourth column of $\tilde{E}_1$ is denoted by $\tilde{E}_1(:, 4)$ and is calculated as

$$\tilde{E}_1(:, 4) = E_2(:, 4) + E_3 \; , \tag{8.19}$$

where $E_2(:, 4)$ denotes the fourth column of E_2. In the rest of this chapter we make this simplification, which is typical for many TCLs.

8.3.2   TCL Aggregation Modeling

An heterogeneous aggregation of n_{ap} TCLs can be modeled by drawing the parameters R, C, C_{p}, P_{n}, M and m from suitable distributions and stacking together the models of individual TCLs. This leads to the following state-space representation

$$\boldsymbol{x}_{t+1} = \boldsymbol{A}\boldsymbol{x}_t + \boldsymbol{B}\boldsymbol{\delta}_t + \boldsymbol{F}_t + \boldsymbol{w}_t \tag{8.20}$$

$$\boldsymbol{E}_1 \boldsymbol{\delta}_t \leq \boldsymbol{E}_2 \boldsymbol{x}_t + \boldsymbol{E}_3 \, , \tag{8.21}$$

where $\boldsymbol{x}_t$, $\boldsymbol{\delta}_t$, $\boldsymbol{w}_t$, $\boldsymbol{E}_3$ are stacked vectors, e.g., $\boldsymbol{x}_t = [(x_t^1)^\top \dots (x_t^{n_{\mathrm{ap}}})^\top]^\top$, and $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{F}_t$, $\boldsymbol{E}_1$ and $\boldsymbol{E}_2$ are block diagonal matrices with the matrices of the individual TCLs on the diagonals.

The output depends upon whether full state information or only noisy aggregate power measurements are available and is given by

$$y_t = \begin{cases} \boldsymbol{C}_1 \boldsymbol{x}_t, & \text{if } t = jT_m, \ j \in \mathbb{N} \\ \boldsymbol{C}_2 \boldsymbol{x}_t + v_t, & \text{otherwise} \end{cases} , \qquad (8.22)$$

where $\boldsymbol{C}_1 = I$ (identity matrix), $\boldsymbol{C}_2 = \begin{bmatrix} 0 \ P_{\mathrm{n}}^1 \ 0 \ P_{\mathrm{n}}^2 \dots 0 \ P_{\mathrm{n}}^{n_{\mathrm{ap}}} \end{bmatrix}$, v_t is the measurement noise, and T_m is the period of the TCL measurements.

8.4 Controller Design

The controller broadcasts signals to the TCLs with the goal of tracking a power trajectory. At each time step t, the controller calculates the required change in power

$$\Delta P_t = P_{\mathrm{set},t} - \hat{P}_{\mathrm{agg},t} - \Delta\hat{P}_{\mathrm{int},t} , \qquad (8.23)$$

where $P_{\mathrm{set},t}$ is the desired setpoint, $\hat{P}_{\mathrm{agg},t}$ is the measured or estimated TCL aggregate power, and $\Delta\hat{P}_{\mathrm{int},t}$ is the estimated change in power resulting from the TCLs' internal hysteresis controller actions, computable from the TCL state estimates $\hat{\mathbf{x}}_{\mathrm{l},t}$. Let $\mathbf{P}_{\mathrm{n}}$ denote the vector of TCL power ratings, n_{sub} the number of MV/LV substations in the network, $P_{\mathrm{agg},t}^l$ the actual TCL aggregate power, $\bar{P}_{\mathrm{UL},t}^l$ the predicted uncontrolled load, $P_{\mathrm{UL},t}^l = \bar{P}_{\mathrm{UL},t}^l + v_t$ the actual uncontrolled load, and $P_{\mathrm{meas},t}^l = P_{\mathrm{agg},t}^l + P_{\mathrm{UL},t}^l$ the noisy aggregate power measurement, all at substation l. There are two ways to calculate $\hat{P}_{\mathrm{agg},t}$. The first one ("Method A") is based on the state estimates, namely

$$\hat{P}_{\mathrm{agg},t} = \mathbf{P}_{\mathrm{n}} \hat{\mathbf{x}}_{\mathrm{l},t} . \qquad (8.24)$$

The second one ("Method B") is based on the aggregate power measurements, namely

$$\hat{P}_{\mathrm{agg},t} = \sum_{l=1}^{n_{\mathrm{sub}}} \left(P_{\mathrm{meas},t}^l - \bar{P}_{\mathrm{UL},t}^l \right) . \qquad (8.25)$$

If $\Delta P_t < 0$ additional off switching is required, whereas if $\Delta P_t > 0$ additional on switching is required. The TCLs that will be switched are determined according to a priority list based on their estimated SoC

$$S\hat{O}C_t = \frac{\hat{x}_{c,t} - m}{M - m} \, , \tag{8.26}$$

where $\hat{x}_{c,t}$ is the TCL's estimated temperature.

If $\Delta P_t < 0$, the TCLs that are on and within the deadband are candidates for switching off actions. An ordered list of these TCLs is obtained by ranking them in descending order based on $S\hat{O}C_t$. The TCLs at the top of this list are hotter, and therefore have a higher switching off priority. Starting from the top of the list, we determine the number of TCLs that will be actively switched off such that their aggregate power consumption is as close as possible to $|\Delta P_t|$. Analogously, if $\Delta P_t > 0$, the TCLs that are off and within the deadband are ranked in ascending order based on $S\hat{O}C_t$. We select the appropriate number of TCLs that will be actively switched on from the top of the ordered list, since they are colder and thus have a higher switching on priority. The $S\hat{O}C_t$ of the last TCL of the priority list that is actively switched on or off is called threshold SoC and is denoted by $SOC_{\text{th},t} \in [0,1]$.

At each time step, the controller broadcasts a pair $[SOC_{\text{th},t}, s_t]$, where $s_t \in \{0,1\}$ is a signal indicating whether an increase in consumption ($s_t = 1$ if $\Delta P_t > 0$) or a decrease in consumption ($s_t = 0$ if $\Delta P_t < 0$) is required. The TCLs that are outside of their deadband are not controllable and ignore the control signal, whereas the rest respond based on their SoC. For each TCL, $SOC_{\text{th},t}$ can be mapped to the temperature threshold

$$x_{\text{th},t} = SOC_{\text{th},t}(M - m) + m \, . \tag{8.27}$$

The desired control actions are described by the state transitions in Tables 8.4a and 8.4b, which can be incorporated into the MLD framework of Section 8.3 by introducing a new auxiliary variable $\delta_{5,t}$ defined as

$$[\delta_{5,t} = 1] \leftrightarrow [x_{c,t} \le x_{\text{th},t}] \, . \tag{8.28}$$

However, an equivalent formulation can be obtained without adding a new auxiliary variable in the following way. Set $\tilde{M} = x_{\text{th},t}$ and $\tilde{m} = m$ if $s_t = 0$, and $\tilde{M} = M$ and $\tilde{m} = x_{\text{th},t}$ if $s_t = 1$. With this notation and given that $SOC_{\text{th},t} \in [0,1]$, Tables 8.4a and 8.4b are equivalent to Table 8.4c by inspection.

Table 8.4: State transition tables

(a) State transitions if $s_t = 0$

$x_{c,t} \leq m$	$m < x_{c,t} < x_{th,t}$	$x_{th,t} \leq x_{c,t} < M$	$M \leq x_{c,t}$
$\delta_{3,t} = 1$	$\delta_{3,t} = \delta_{4,t}$	$\delta_{3,t} = 0$	$\delta_{3,t} = 0$

(b) State transitions if $s_t = 1$

$x_{c,t} \leq m$	$m < x_{c,t} \leq x_{th,t}$	$x_{th,t} < x_{c,t} < M$	$M \leq x_{c,t}$
$\delta_{3,t} = 1$	$\delta_{3,t} = 1$	$\delta_{3,t} = \delta_{4,t}$	$\delta_{3,t} = 0$

(c) Equivalent transitions

$x_{c,t} \leq \tilde{m}$	$\tilde{m} < x_{c,t} < \tilde{M}$	$\tilde{M} \leq x_{c,t}$
$\delta_{3,t} = 1$	$\delta_{3,t} = \delta_{4,t}$	$\delta_{3,t} = 0$
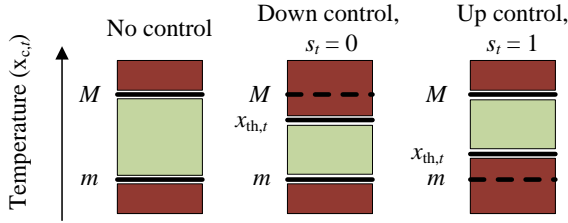
Figure 8.2: Dynamic tightening of TCL deadband under external control actions (the light color denotes the active deadband).

The state transitions of Table 8.4c are described by the logic relations

$$[\delta_{1,t} = 1] \leftrightarrow [x_{c,t} \geq \tilde{M}] \tag{8.29}$$

$$[\delta_{2,t} = 1] \leftrightarrow [x_{c,t} \leq \tilde{m}], \tag{8.30}$$

along with (8.5), (8.6), (8.8)–(8.10). Therefore, the external control actions can be directly incorporated into the MLD framework of Section 8.3 using (8.11). The only difference is that matrix $\tilde{E}_1$ and vector $\tilde{E}_3$ now depend on $\tilde{M}$, $\tilde{m}$ and are time-varying. The external control can be seen as a dynamic tightening of a TCL's deadband, which is visualized in Fig. 8.2. This control approach is similar to approaches based on temperature setpoint control [100]; however, in our approach user comfort is guaranteed because $\tilde{M} \leq M$ and $\tilde{m} \geq m$.

The control loop is shown in Fig. 8.3. Note that the aggregator applies the MHSE method independently for each substation i.
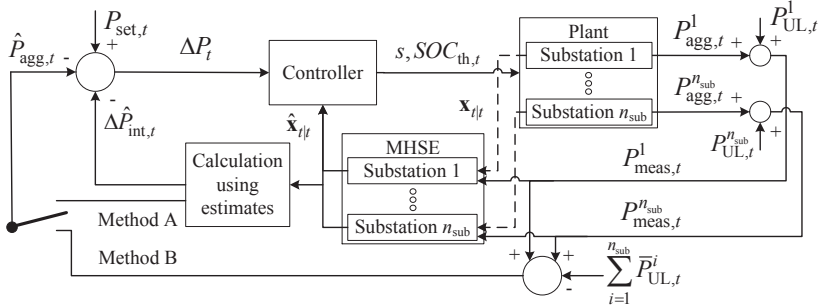
Figure 8.3: Control loop including the broadcast controller and the moving horizon state estimator for each substation.

8.5   State Estimator Design

We propose an MHSE method to estimate the TCL states when current TCL state measurements are unavailable. At each time step $t \neq jT_m$, $j \in \mathbb{N}$ we solve a multi-period MILP with the TCL temperatures, on/off modes, auxiliary binary variables, as well as process and measurement noise as optimization variables. Let us define the optimization vector as

$$\boldsymbol{x}_t^{\text{opt}} := [\hat{\boldsymbol{x}}_{t-N+1|t}, \hat{\boldsymbol{\psi}}_{t-N+1|t}, \dots, \hat{\boldsymbol{\psi}}_{t|t}] \tag{8.31}$$

$$\hat{\boldsymbol{\psi}}_{k|t} := [\hat{\boldsymbol{\delta}}_{k|t}, \hat{\boldsymbol{w}}_{k|t}, \hat{v}_{k|t}] \ , \tag{8.32}$$

where $\hat{\boldsymbol{\delta}}_{k|t} \in \{0,1\}^{4n_{\text{ap}}}$ are the auxiliary binary variables, $\hat{\boldsymbol{w}}_{k|t} \in \mathbb{R}^{n_{\text{ap}}}$ are the process noise variables, $\hat{v}_{k|t} \in \mathbb{R}$ are the measurement noise variables, $k \in [t-N+1, t]$, N is the estimation horizon, and $\hat{\cdot}_{k|t}$ is the estimate of $\cdot$ at time step k using measurements up to time step t. Note that $\hat{\boldsymbol{x}}_{t-N+2|t}, \dots, \hat{\boldsymbol{x}}_{t|t}$ can be determined from $\hat{\boldsymbol{\delta}}_{k|t}$, $k \in [t-N+1, t]$, and therefore these additional optimization variables are not needed.

With this notation, the estimation problem can be written as

$$\min_{\boldsymbol{x}_t^{\text{opt}}} \sum_{k=t-N+1}^{t} \Big(m_1 \, |\hat{y}_{k|t} - y_k| + m_2 \, ||Q^{-1}\hat{\boldsymbol{w}}_{k|t}||_1$$

$$+ m_3 \, |R^{-1}\hat{v}_{k|t}| \Big) + \sum_{k=t-N+1}^{t-1} \Big(m_4 \, ||Q^{-1} \big(\hat{\boldsymbol{x}}_{\text{c},k|t} - \hat{\boldsymbol{x}}_{\text{c},k|t-1} \big) ||_1 \Big)$$

$$+ m_5 \sum_{i=1}^{n_{\text{ap}}} \left(\left| \hat{W}^i_{k|t} - W^i \right| \right) + m_6 \left| \hat{V}_{k|t} - V \right|, \tag{8.33}$$

$$\text{s.t.} \quad \hat{\boldsymbol{x}}_{k+1|t} = \boldsymbol{A}\hat{\boldsymbol{x}}_{k|t} + \boldsymbol{B}\hat{\boldsymbol{\delta}}_{k|t} + \boldsymbol{F}_t + \hat{\boldsymbol{w}}_{k|t}, \tag{8.34}$$

$$\hat{y}_{k|t} = \boldsymbol{C}_2 \hat{\boldsymbol{x}}_{k|t} + \hat{v}_{k|t}, \tag{8.35}$$

$$\boldsymbol{E}_1 \hat{\boldsymbol{\delta}}_{k|t} \leq \boldsymbol{E}_2 \hat{\boldsymbol{x}}_{k|t} + \boldsymbol{E}_3, \tag{8.36}$$

$$\hat{\boldsymbol{x}}_{k|t} = \boldsymbol{x}_{k|t}, \quad \forall t \in [jT_m + 1, jT_m + N - 1],$$
$$\forall k \in [t - N + 1, jT_m], \tag{8.37}$$

where Q^{-1} and R^{-1} are the inverses of the process and measurement noise covariance matrices, respectively, $m_1 - m_6$ are weighting factors, W^i is a known statistic on the process noise of TCL i (e.g., the mean value), V is a known statistic on the measurement noise, and $\hat{W}^i_{k|t}$, $\hat{V}_{k|t}$ are estimates of those statistics computed from $\hat{\boldsymbol{w}}_{k|t}$ and $\hat{v}_{k|t}$.

The first term of (8.33) minimizes the difference between the measured aggregate power and the output calculated from the estimated states. The second and third terms penalize the process and measurement noise. The fourth term is needed to link the current estimation problem to the results of the previous estimation problems. Note that only the continuous states are included in the fourth term. The fifth and sixth terms require that the current noise statistics are close to their known values. Equations (8.34)-(8.36) describe the TCL hybrid dynamics. For $t \in [jT_m+1, jT_m+N-1]$, $j \in \mathbb{N}$, noise-free TCL state measurements are available (because the estimation horizon includes the last time when a packet of TCL measurements was received), which are taken into account by introducing the equality constraint (8.37) and setting $m_4 = 0$ in (8.33). For $t \geq jT_m + N$, (8.37) is not considered and $m_4 \neq 0$.

Note that the terms of (8.33) that are related to either w or v are normalized by either Q or R, which are assumed to be known. This ensures that the numeric values of the penalties on process and measurement noise are in the same range, which is essential for good performance. However, no assumption on the Probability Distribution Functions (PDFs) of the noises is needed.

The MHSE performance can be improved by tuning the weighting factors $m_1 - m_6$. For example, for longer estimation horizons higher m_5 and m_6 values might be preferable, because the statistics W^i and V can be estimated more accurately due to the larger sample size. In addition, m_4 can be time-varying, i.e., a function of the quality of recent estimates. If no noise statistics are available, the last two terms of

(8.33) are dropped, and m_2, m_4 can be tuned based on the magnitude of w, whereas m_3 based on the magnitude of v.

Let $n_{\mathrm{d}} = 4$ denote the number of auxiliary binary variables per TCL, $n_{\mathrm{w}} = 1$ denote the number of process noise variables per TCL, and $n_{\mathrm{v}} = 1$ denote the number of measurement noise variables per time step. The total number of variables is $n_{\mathrm{var}} = 2n_{\mathrm{ap}} + N[n_{\mathrm{ap}}(n_{\mathrm{d}} + n_{\mathrm{w}}) + n_{\mathrm{v}}]$, of which $n_{\mathrm{bi}} = n_{\mathrm{ap}} + N n_{\mathrm{ap}} n_{\mathrm{d}}$ are binary variables. Therefore, the MHSE problem is a large scale mixed-integer optimization problem, even for small TCL aggregations and estimation horizons. For example, if $n_{\mathrm{ap}} = 20$ and $N = 10$, the estimation problem for each substation will have 820 binary variables.

Due to the size of the problem, we use a 1-norm minimization in (8.33), which can be reformulated as a MILP, instead of a least-squares minimization as in [165, 172, 173], which would lead to a quadratic integer program. Using the 1-norm enhances the tractability of our approach, but it might reduce the performance compared with a least-squares formulation. Adopting the hierarchical architecture of Section 8.2, the estimation problem can be solved independently and in parallel for each substation, which makes it tractable in real time since the number of consumers per substation and/or feeder is typically low, e.g., 20 to 40 consumers.

8.6   Performance of the MHSE Method

In this section, we demonstrate the performance of the MHSE method in autonomous operation (i.e., when the TCLs are solely controlled by their hysteresis controllers), and under external control with and without measurement noise. Furthermore, we benchmark the MHSE method against a simple "model-based predictor", which receives TCL measurements whenever they are available and evolves the TCL dynamics assuming zero process noise, but does not utilize the noisy aggregate power measurements. In addition, we investigate the value of using MHSE in SFC with TCLs.

For all case studies we consider space heaters and parameterize them by drawing their parameters from the uniform probability distributions of Table 8.5. All simulations are performed with a time step of 10 seconds in MATLAB using a 4 core machine (2.83 GHz) with 8 GB RAM, and the MHSE problem was solved using CPLEX.

Table 8.5: Space heater parameters, adapted from [3, 4]

$P_\mathrm{n} \sim U(5,9)$ kW	$C \sim U(5,9)$ kWh/°C	$R \sim U(1.5, 2.5)$ °C/kW
$C_\mathrm{p} \sim U(2,3)$	$T_\mathrm{sp} \sim U(19,23)$ °C	$T_\mathrm{db} \sim U(0.25,1)$ °C

8.6.1   Autonomous Operation without Measurement Noise

In the first investigation, we consider the autonomous operation of a TCL aggregation connected to a single distribution substation, and estimate the states without external control actions to gain insight into the estimation process. In this case, the TCLs are controlled only by their internal thermostats. To understand how process noise affects the estimation quality, we set the measurement noise to zero ($v_t = 0$). The weighting factors m_1 to m_4 in (8.33) are all chosen equal to 1, which was shown to perform well. We also disregard the last two terms of (8.33) by setting $m_5 = m_6 = 0$, i.e., we do not assume knowledge of noise statistics.

We consider 20 TCLs and use only integer values of the rated power P_n, which complicates the estimation process since different TCLs appear similar to the estimator. The process noise for each TCL follows a normal distribution with mean value $\mu = 0$ and standard deviation $\sigma = 10^{-3}$. We run hourly simulations with an estimation horizon of $N = 10$ steps (100 seconds), while TCL measurements are received every $T_m = 20$ minutes.

Figure 8.4 shows the actual, estimated, and predicted temperature trajectories of a single TCL. During the first 36 minutes, there is practically no difference between the MHSE and the model-based predictor. At this point, the estimator notices a significant mismatch between the expected and actual aggregate power. Afterwards, the temperature estimate converges to the actual value and the MHSE clearly outperforms the predictor. Notice that even small process noise can lead to significant discrepancy between the actual and predicted trajectories due to the on/off type of operation of TCLs.

The quality of estimates provided by the MHSE and the predictor at each time step t is quantified using the Mean Absolute Error (MAE) for the temperatures and the 1-norm error (L_1) for the on/off states,
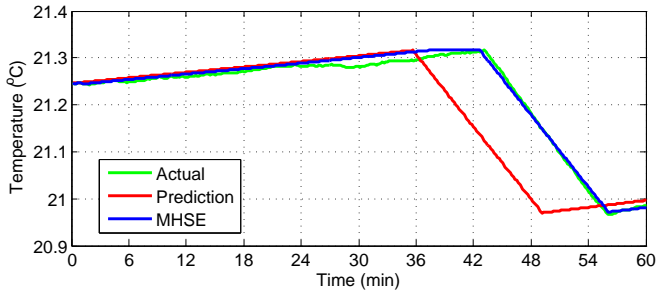
Figure 8.4: The evolution of the actual, predicted and estimated tempera-
ture trajectories of a single TCL.

which are defined according to

$$\text{MAE}_t = \frac{1}{n_{\text{ap}}} \sum_{i=1}^{n_{\text{ap}}} |x_{\text{c},t}^i - \hat{x}_{\text{c},t}^i| \,, \tag{8.38}$$

$$L_{1,t} = \sum_{i=1}^{n_{\text{ap}}} |x_{\text{l},t}^i - \hat{x}_{\text{l},t}^i| \,. \tag{8.39}$$

The evolution of estimation errors over time is shown in Fig. 8.5. The
on/off state of each TCL is almost always correctly estimated by the
MHSE, whereas the predictor produces significant errors. Although the
on/off estimate diverges from the actual state around the 42nd minute,
afterwards the estimation error goes to zero. Therefore, even in case of
temporary wrong estimates, the on/off state estimates eventually con-
verge to the actual states. The temperature prediction error generally
increases with time, whereas the MHSE temperature estimation error
is kept at lower values. A perfect temperature estimation is not pos-
sible because the estimator retrieves information only when differences
between the predicted and measured power occur. In between, the tem-
perature estimation error might grow depending on the process noise.

The optimal estimation period N depends on the aggregation size and,
in practical applications, would likely be determined by real-time com-
putational limitations. For these reasons, we do not choose N based
on the observability tests proposed in [165], but select it empirically.
Figure 8.6 shows results from simulations with $n_{\text{ap}} = \{10, 20, 30\}$ and
$N = \{5, 10, 15, 20\}$. To ensure a fair comparison, exactly the same pro-
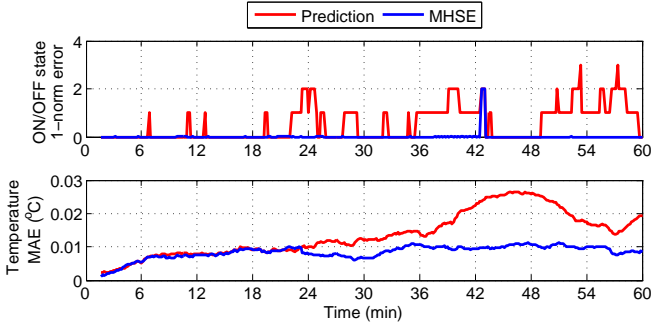
Figure 8.5: Evolution of the estimation errors for a population of 20 TCLs.
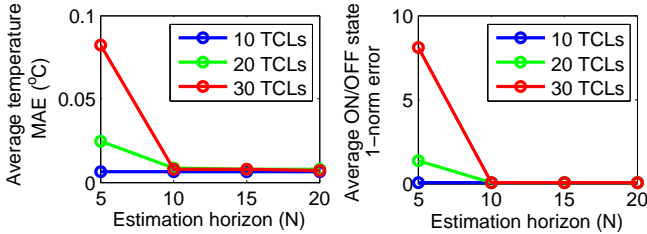
Figure 8.6: Dependence of the estimation error on the estimation horizon N for different TCL aggregation sizes.

cess noise time series were used to simulate a given aggregation for each value of N.

In all cases, an horizon of $N = 10$ performs very well, while further increasing N does not reduce the estimation errors. Shorter horizons degrade the estimation quality for aggregations of 20 or 30 TCLs; however, for very small aggregations of 10 TCLs even an horizon of $N = 5$ is enough. Note that since we have considered specific TCL aggregations and process noise time series, these results are only meant to provide intuition. More general conclusions regarding the best choice of N could be obtained by Monte Carlo simulations.

The problem (8.33)-(8.37) can be solved quickly for the aggregation sizes considered here. Table 8.6 shows how the average computation time (first number) and the maximum computation time (second number) depend on n_{ap} and N. Not surprisingly, increasing n_{ap} or N leads to

Table 8.6: Mean/max computation time (in seconds)

n_{ap}	$N = 5$	$N = 10$	$N = 15$	$N = 20$
10	0.06/0.23	0.09/0.36	0.12/0.40	0.17/0.49
20	0.10/1.93	0.15/0.47	0.28/0.76	0.42/1.01
30	0.20/4.15	0.27/0.96	0.47/1.34	0.72/1.74

longer solution times. Note that increasing N has a larger impact on the computation time since this introduces more binary variables in the problem. The reported computation times are likely acceptable for real-time estimation applications.

8.6.2 Controlled Operation without Measurement Noise

In the second investigation, we perform state estimation on a TCL aggregation that is externally controlled to track a power trajectory. The purpose of this investigation is to quantify the improvement of controller's performance with better TCL state estimates. We consider the same 20 TCLs of Section 8.6.1 and run a simulation over 20 minutes with m_1 to m_4 equal to 1, $m_5 = m_6 = 0$, $N = 10$, $T_m = 20$, $v_t = 0$, and the same process noise parameters as in Section 8.6.1. Since T_m and the simulation horizon are both equal to 20 minutes, there are no TCL measurements received apart from the initial measurements.

We control the TCL aggregation using the controller of Section 8.4. We calculate the power trajectory P_{set} by superimposing a random signal upon the predicted TCL aggregation baseline, i.e., the power trajectory without external control actions, as in [21, 156].

The model-based predictor leads to an MAE of 6.40 kW, whereas the MHSE method reduces the MAE to 4.04 kW, which represents an improvement of about 37%. The absolute tracking errors for both cases are shown in Fig. 8.7a. Note that we have smoothed the error time series using a moving window of 1 minute to enhance the visibility of the trend. Fig. 8.7b shows the broadcasted control signal SOC_{th} for both cases. One can see that the improved state estimates with MHSE also lead to less aggressive control signals compared with the model-based prediction approach.
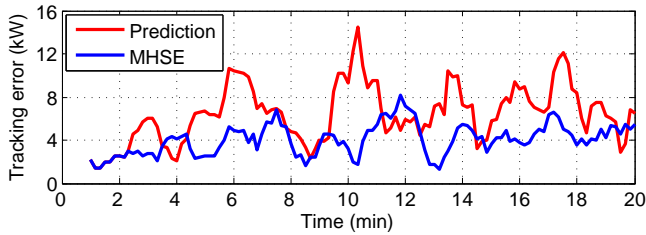
(a) Power trajectory tracking error.
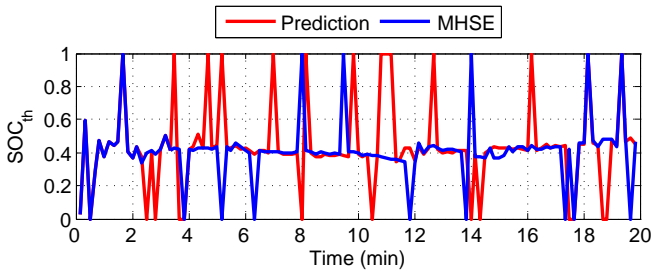
(b) Broadcasted control signal.
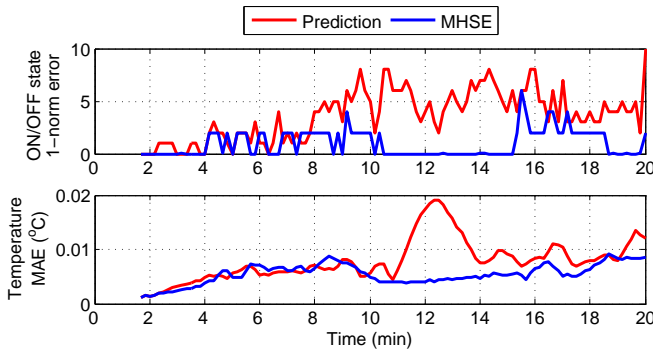
(c) Estimation and prediction errors.

Figure 8.7: Closed-loop control of 20 TCLs with MHSE, with process noise but without measurement noise.

Figure 8.7c shows the evolution of estimation errors over time. The on/off state estimation error of the MHSE method is generally larger than the values reported in Section 8.6.1 due to the external control actions. There exist intervals where the estimates of the on/off state

and temperature produced by the MHSE method are worse than those produced by the model-based predictor. However, the controller performs better with MHSE than with prediction during the same intervals, which indicates that the MHSE improves controller performance even with temporarily poor TCL-level state estimates.

8.6.3   Controlled Operation with Measurement Noise

In the third investigation, apart from the process noise (w_t), we additionally consider measurement noise (v_t). We assume that w_t and v_t follow zero-mean normal distributions with known variances, no autocorrelation, and no correlation between TCLs. In this case, the last two terms of (8.33) become

$$m_5 \sum_{i=1}^{n_{\text{ap}}} \frac{1}{\sigma_{\text{w}}^i} \left| \hat{\mu}_{\text{w},k|t}^i - \mu_{\text{w}}^i \right| + m_6 \, \frac{1}{\sigma_{\text{v}}} \left| \hat{\mu}_{\text{v},k|t} - \mu_{\text{v}} \right| , \qquad (8.40)$$

where σ_{w}^i is the process noise standard deviation for TCL i, and σ_{v} is the measurement noise standard deviation. The mean values of the process and measurement noise are fixed to $\mu_{\text{w}}^i = \mu_{\text{v}} = 0$. In this case Q is diagonal.

Measurements of TCL states are assumed to be available every $T_m = 15$ minutes. We fix $m_1 = 10^8$ such that the first term of (8.33) is a soft constraint, and we set m_2 to m_6 equal to 1. The choice of m_4 is critical for the convergence of the estimator [173], and an empirical investigation showed that $m_4 = 1$ leads to a good performance. Similarly to the previous case studies, we set the estimation horizon to $N = 10$.

Consider a population of 20 TCLs connected to a substation along with other uncontrolled loads, and assume a TCL coincident load equal to 20% of the total substation load P_{sub}. Assume $\sigma_{\text{w}} = 10^{-3}$ and $\sigma_{\text{v}} = 0.02 \cdot P_{\text{sub}}$, where the measurement noise is given as a percentage of P_{sub} since it represents the prediction error of the uncontrolled demand.

Open-Loop Controller with Random Control Signals

We run simulations for 15 minutes using an open-loop controller that broadcasts the random control signals

$$s_t = |s_{t-1} - r_s| \qquad (8.41a)$$

$$SOC_{\text{th},t} = SOC_{\text{th},t-1} - r_{\text{soc}}(-1)^{s_t} , \qquad (8.41b)$$

where $r_{\mathrm{s}} \in \{0, 1\}$ is a Bernoulli random variable with success probability P_{s}, and $r_{\mathrm{soc}} \in [0, R_{\mathrm{soc}}]$ is a continuous uniformly distributed random variable. The parameter P_{s} is the probability with which the control direction changes between two consecutive time steps. The parameter R_{soc} describes the change in the control signal magnitude between two consecutive time steps. Therefore, higher values of P_{s} and/or R_{soc} result in more aggressive control actions. In this chapter, two different forcing levels are investigated: (a) a low forcing scenario with $P_{\mathrm{s}} = 0.1$, and (b) a high forcing scenario with $P_{\mathrm{s}} = 0.5$. In both cases, we fix $R_{\mathrm{soc}} = 0.025$.

TCL Clustering for Estimation Evaluation

The estimation quality is evaluated using as indicators the estimation errors of temperature, on/off mode, and aggregate power consumption at each time step. Depending on the measurement noise level, a group of TCLs with similar power ratings might not be distinguishable by the MHSE method. However, what matters most is to estimate the aggregate power and the SoC of the group.

For this reason, we cluster the TCLs according to their power ratings and use these clusters to assess the estimation quality. Let us denote the number of clusters by n_{cl}, cluster i by CL_i, and define its SoC as

$$ SOC_{\mathrm{CL}_i, t} = \frac{\sum_{j \in \mathrm{CL}_i} (x_{\mathrm{c}, t}^j - m^j)}{\sum_{j \in \mathrm{CL}_i} (M^j - m^j)} \ . \tag{8.42} $$

The estimated SoC of cluster i can be defined similarly using $\hat{x}_{\mathrm{c}, t}^j$. We define the SoC and on/off mode estimation errors as

$$ \varepsilon_{\mathrm{soc}, t} = \sum_{i=1}^{n_{\mathrm{cl}}} \left| SOC_{\mathrm{CL}_i, t} - S\hat{O}C_{\mathrm{CL}_i, t} \right| \tag{8.43} $$

$$ \varepsilon_{\mathrm{on/off}, t} = \sum_{i=1}^{n_{\mathrm{cl}}} \left| \sum_{j \in \mathrm{CL}_i} u_t^j - \sum_{j \in \mathrm{CL}_i} \hat{u}_t^j \right| \ . \tag{8.44} $$

We define the TCL aggregate power estimation error as

$$ \varepsilon_{\mathrm{agg}, t} = \boldsymbol{C}_2 \boldsymbol{x}_t - \boldsymbol{C}_2 \hat{\boldsymbol{x}}_t \ . \tag{8.45} $$

The MAE over the simulation horizon N_{sim} is computed with MAE $= (1/N_{\mathrm{sim}}) \cdot \sum_{t=1}^{N_{\mathrm{sim}}} |\varepsilon_t|$, where ε_t is any of $\varepsilon_{\mathrm{soc}, t}$, $\varepsilon_{\mathrm{on/off}, t}$, and $\varepsilon_{\mathrm{agg}, t}$.
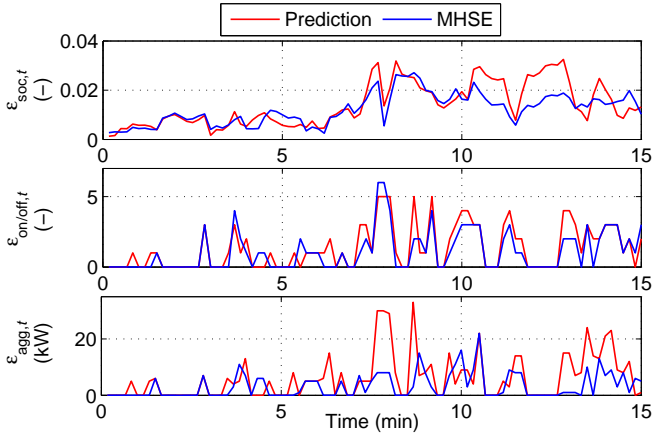
Figure 8.8: Estimation errors for the MHSE and the model-based predictor in a simulation with the external open-loop controller (8.41).

Indicative Simulation Results

In Fig. 8.8 we present estimation results for the MHSE and the model-based predictor from a simulation with the external open-loop controller (8.41). High controller forcing ($P_s = 0.5$) and $n_{cl} = 5$ clusters are assumed in the simulation. The MHSE method provides better estimates compared with the model-based predictor most of the time, and it reduces MAE_{soc} by 13%, $\text{MAE}_{on/off}$ by 19%, and MAE_{agg} by 47%.

8.6.4   Benchmarking of MHSE Method

In this section, we benchmark the MHSE method against the model-based predictor for different process and measurement noise levels, and different controller forcing levels. The goal is to provide insights on the noise levels that could be handled by the estimator in real-world applications.

We consider the three process noise levels $\sigma_w = \{5 \cdot 10^{-4}, \, 10^{-3}, \, 5 \cdot 10^{-3}\}$, referred to as "low", "medium", and "high" process noise, and the four measurement noise levels $\sigma_v = \{0, \, 0.02, \, 0.05, \, 0.1\} \cdot P_{sub}$, referred to as "zero", "low", "medium", and "high" measurement noise. In addition, we consider a low controller forcing level with $P_s = 0.1$ and a high forcing
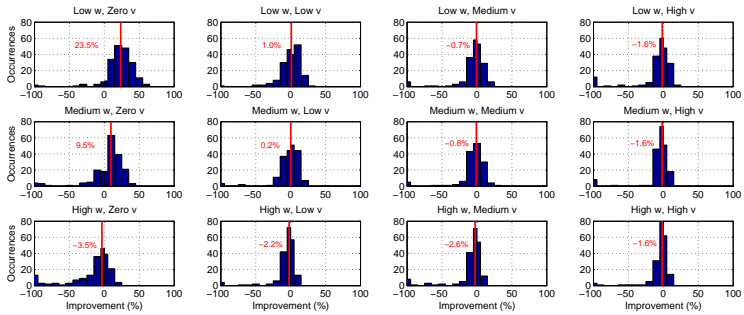
(a) Histograms of SoC estimation improvement.
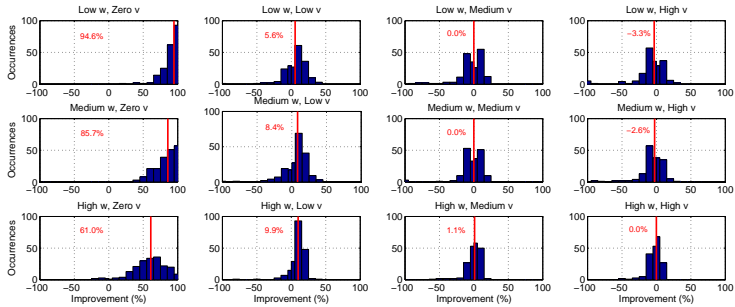
(b) Histograms of on/off mode estimation improvement.
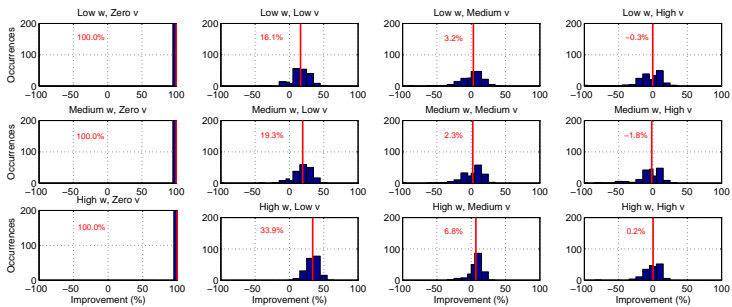
(c) Histograms of aggregate power estimation improvement.

Figure 8.9: Comparison of MHSE method versus the model-based prediction approach for high forcing control. The light-colored lines indicate the medians of performance improvement.
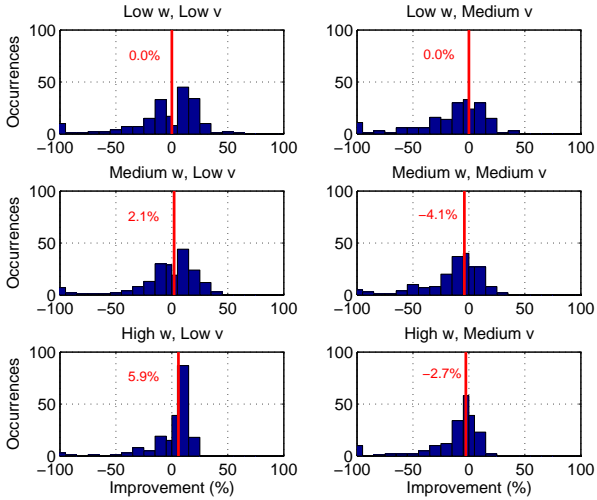
(a) Histograms of on/off mode estimation improvement.
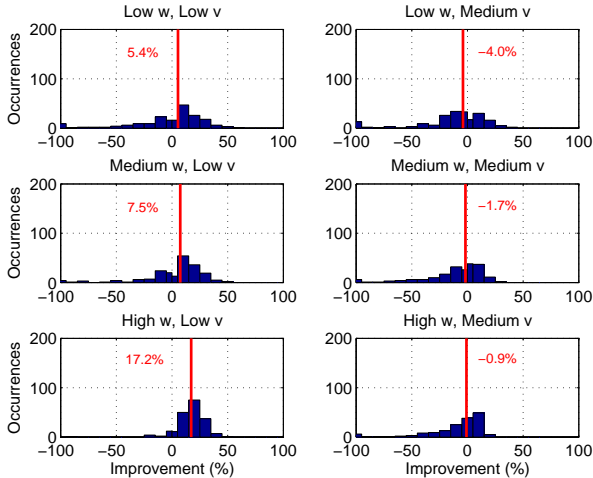
(b) Histograms of aggregate power estimation improvement.

Figure 8.10: Comparison of MHSE method versus the model-based prediction approach for low forcing control. The light-colored lines indicate the medians of performance improvement.

level with $P_s = 0.5$. This leads to a total combination of 24 simulation cases for each of which we run 200 simulations with randomly generated TCL populations and noise realizations.

Figure 8.9 shows the histograms of estimation improvement for the high forcing case, where the light-colored lines correspond to the medians. We quantify the improvement of estimation performance according to

$$-100 \cdot \frac{\text{MAE}_{\text{est}} - \text{MAE}_{\text{pred}}}{\text{MAE}_{\text{pred}}} \% , \qquad (8.46)$$

where MAE_{pred} is the MAE of the model-based predictor and MAE_{est} is the MAE of the estimator. With the exception of the cases {low w, zero v} and {medium w, zero v}, the MHSE method does not improve the SoC estimates. This is because the TCL temperatures cannot be observed directly, but only through the on/off mode estimates. However, even with poor SoC estimates, the MHSE estimates the on/off modes and aggregate power better than the model-based predictor for most of the cases.

If there is no measurement noise ($v = 0$), the MHSE method drastically improves the estimation quality, especially in terms of aggregate power. For all cases with low v and for the case {high w, medium v}, the estimates of the on/off modes and aggregate power generally improve, resulting in positive medians in the range $1.1\% - 33.9\%$. Interestingly, for the case {high w, low v} the MHSE method achieves better aggregate power estimates for all 200 scenarios, and better on/off mode estimates for 87% of them. On the other hand, for all cases with high v, the MHSE method performance is poor.

Figure 8.10 shows the results for on/off mode and aggregate power estimation for the low forcing case. For SoC estimates and for all cases with zero or high v, we observed patterns similar to those seen in the high forcing case and thus those histograms are omitted. The MHSE method outperforms the model-based prediction only for the cases {medium w, low v}, and {high w, low v}. Also, note that the resulting improvement is worse than in the high forcing cases. This provides useful intuition: the higher the forcing level is, the more information is retrieved from the aggregate power measurements improving the estimation quality. Overall, our results indicate that the MHSE method performs better in applications with high process and low measurement noise, irrespective of the forcing level.
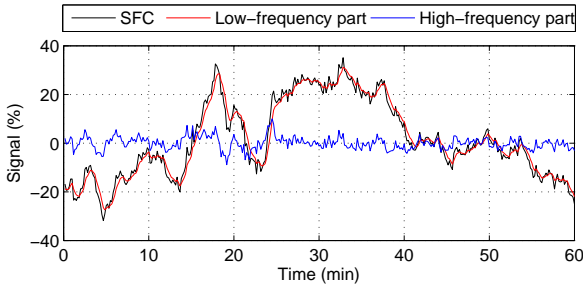
Figure 8.11: The standard SFC as well as its high and low frequency components.

8.6.5 Application of MHSE in SFC

In the last case study we investigate the value of MHSE in SFC. Consider a population of n_{ap} TCLs evenly distributed among n_{sub} distribution substations. Based on the results of Section 8.6.4, we select $\sigma_w = 5 \cdot 10^{-3}$ (high process noise) and $\sigma_v = 0.02 \cdot P_{sub}$ (low measurement noise), and fix the noise realizations. The desired setpoint is a function of the baseline consumption of the TCL population $P_{b,t}$, the control band α, and the normalized control signal Y_t (it takes values in the range $[-1, 1]$)

$$P_{set,t} = P_{b,t} + \alpha Y_t P_{b,t} . \tag{8.47}$$

We test two control signals: (a) an extract of the Swiss SFC signal from 2009 with $\alpha = 0.5$, and (b) the high-frequency component of the same signal obtained by applying a high-pass filter with cutoff frequency $1/30$ Hz and $\alpha = 3$. Both signals are shown in Fig. 8.11. The reason why we investigate (b) is that a larger control band can be offered without energy constraint violations, and more forcing will be necessary, which is expected to improve the estimation/control performance according to Section 8.6.4.

A controller with MHSE is compared against a controller with the model-based predictor, using Methods A and B to calculate $\hat{P}_{agg,t}$ – see (8.24) and (8.25). We consider a reference case with perfect TCL state information that gives us a performance bound. Two aggregation sizes are considered: $n_{ap} = 40$ TCLs distributed among $n_{sub} = 2$ substations, and $n_{ap} = 400$ TCLs distributed among $n_{sub} = 20$ substations. All simulations are performed for 1 hour with a time step of 10 seconds.
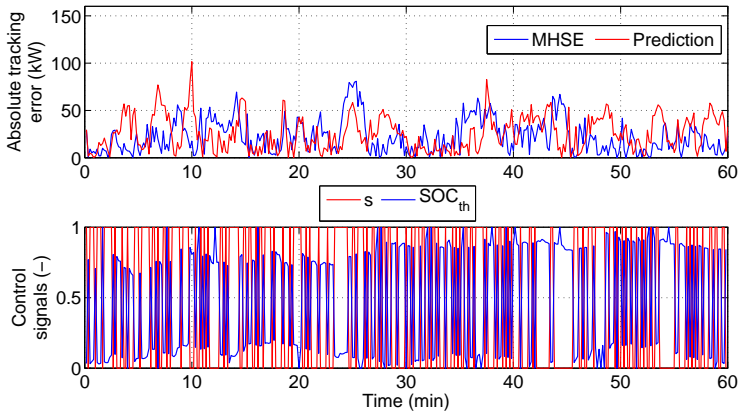
Figure 8.12: Tracking performance and broadcasted control signal for the standard SFC signal.
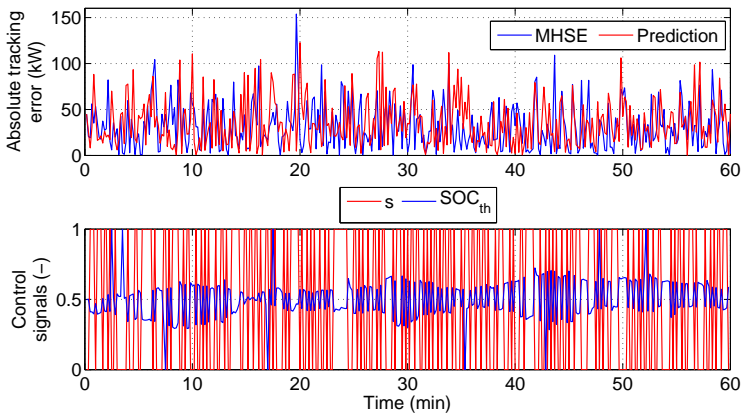
Figure 8.13: Tracking performance and broadcasted control signal for the high frequency part of the SFC signal.

We assess the controller's performance by calculating the Root Mean Squared Error (RMSE) between the desired setpoint and the TCL aggregate power consumption, and we summarize the results in Table 8.7. For all cases, the MHSE method outperforms the model-based predictor resulting in lower RMSE values. The improvement is more pronounced for the high-frequency signal due to higher forcing, which is consistent

Table 8.7: RMSE values for setpoint tracking in SFC

	2 Substations		20 Substations	
$\hat{P}_{\text{agg},t}$ Method	A	B	A	B
Standard SFC				
Reference (kW)	2.21	N/A	2.13	N/A
Prediction (kW)	11.84	16.40	31.66	36.67
MHSE (kW)	**10.26**	14.56	**28.55**	35.09
Improvement (%)	13.34	11.22	9.82	4.31
High frequency part of SFC				
Reference (kW)	2.29	N/A	2.18	N/A
Prediction (kW)	23.29	18.26	75.75	44.08
MHSE (kW)	18.13	**15.67**	55.18	**40.18**
Improvement (%)	22.16	14.18	27.16	8.85

with the results of Section 8.6.4.

Interestingly, the best way to calculate $\hat{P}_{\text{agg},t}$ depends on the signal. For the standard SFC signal Method A (based on the estimated states) is preferable, whereas for the high-frequency signal Method B (based on the aggregate power measurements) performs better. The reason is that the higher the frequency content of the signal, the higher the signal-to-measurement noise ratio; therefore, the aggregate power measurements $P^i_{\text{meas},t}$ can be trusted more.

The above observations are valid for both $n_{\text{sub}} = 2$ and $n_{\text{sub}} = 20$. The best results for each case are highlighted with bold font in Table 8.7. Overall, the MHSE method reduces the RMSE by approximately 14% for $n_{\text{sub}} = 2$ and 9% for $n_{\text{sub}} = 20$.

The tracking performance and the broadcasted control signals for $n_{\text{sub}} = 20$ are shown in Figs. 8.12 and 8.13. Note that the high-frequency signal results in a more aggressive s_t time series, and therefore the SOC_{th} values are much closer to 0.5, compared with that of the standard SFC signal.

Table 8.8 shows the average number of switching actions per TCL for the highlighted cases of Table 8.7. Similarly to Table 8.7, in the reference case the TCL states are known at every time step. For the standard SFC signal, MHSE slightly increases the number of switching actions compared to the model-based predictor. However, for the high-frequency

Table 8.8: Average number of switching actions per TCL

	2 Substations	20 Substations
Standard SFC		
Reference	3.82	2.56
Prediction	3.93	2.57
MHSE	4.63	2.63
High frequency part of SFC		
Reference	10.40	9.57
Prediction	23.23	10.69
MHSE	17.68	9.43

signal, MHSE significantly reduces the number of switching actions, in particular for $n_{\mathrm{sub}} = 2$. In this case, MHSE not only improves the controller's performance, but also decreases the wear on TCLs.

Beyond filtering the SFC signal, there are additional ways to increase the forcing levels to improve the MHSE/controller performance. For example, the TCL population could be divided into two groups, each tasked with following a high frequency signal that sum to the original SFC signal. Additionally, one could add an artificial signal to the SFC signal, and that signal could be balanced by another resource.

8.7 Conclusion

In this chapter, we presented an MHSE method to extract the temperatures and on/off states of individual loads in a TCL aggregation. The method assumes a two-way constrained communication infrastructure between an aggregator and each TCL, and is designed to work with realistic measurements, i.e., real-time noisy aggregate power measurements from distribution substations and TCL state measurements from smart meters that arrive at lower frequency intervals. We also proposed a scalable hierarchical closed-loop controller for SFC provision with TCLs, which uses the MHSE method and is based on signal broadcasts.

We demonstrated the performance of the MHSE method under different process and measurement noise characteristics and controller forcing levels, and benchmarked it against a simpler model-based predictor.

Simulations showed that the MHSE method provides accurate estimates for a certain spectrum of process and measurement noise. Additionally, our investigations revealed some interesting observations: (i) the MHSE method works better with aggressive control trajectories; (ii) the controller can achieve good tracking even with temporarily poor state estimates for individual TCLs; and (iii) for aggregation sizes up to 30 TCLs, an estimation horizon $N = 10$ is a reasonable tradeoff between estimation accuracy and computation time.

There are several interesting avenues for future work. For example, one can analyze the convergence properties of the proposed MHSE method from a theoretical point of view. Moreover, the proposed method can be extended to include different types of noise distributions and autocorrelation.

Chapter 9

Secondary Frequency Control with Commercial Buildings

9.1   Introduction

The goal of this chapter is to develop methods to allow aggregations of commercial buildings to participate in Secondary Frequency Control (SFC). The Heating, Ventilation and Air Conditioning (HVAC) systems of commercial buildings are widely considered for Demand Response (DR) applications in the literature because they provide a number of advantages. First, commercial buildings are responsible for a significant percentage of the total energy consumption (for example, 20% in the USA), whereas the HVAC system represents a large percentage of a building's overall energy consumption (for example, 50% in the USA) [174]. Second, commercial buildings typically have a large thermal inertia and therefore changes in heating and cooling power are not instantaneously realized by the occupants. Third, many buildings (for example, one third of them in the USA [120]) are equipped with a Building Automation System (BAS) that facilitates DR implementation.

Chapter 4 provided a detailed literature review of early work on modeling [112, 113], load shifting [9–13], and recent work on SFC track-

Table 9.1: Nomenclature of Chapter 9: symbols

Symbol	Unit	Description
$\bar{\mathbf{A}}_{\mathrm{w}}$	-	Matrix of polyhedral constraint
a	-	Auxiliary variable to describe L_2
$\bar{\mathbf{b}}_{\mathrm{w}}$	-	Vector of polyhedral constraint
b	-	Auxiliary variable to describe L_2
c	CHF/kWh	Electricity price
d	-	Binary variable to denote the on/off HP state
H	-	Affine disturbance feedback matrix
h	W/m^2	Scheduled heat flux by the HP if $w = 0$
L	-	Number of buildings in the aggregation
L_1, L_2	Wh/m^2	Lost thermal energy due to HP dynamics
N_{s}	-	Total number of samples for chance constraints
N_{st}	s	Prediction horizon of short-term capacity allocation
N_1, N_2	hours	Prediction horizon of reserve scheduling and MPC
$n_{\mathrm{u}}, n_{\mathrm{d}}$	-	Up and down rate limits of the HP
P	W	HP electric power
r	W/m^2	Symmetric reserve capacity in the heat flux domain
r_{e}	W/m^2	Symmetric electric reserve capacity
r_+, r_-	W/m^2	Down- and up-reserve capacity in heat flux domain
S	-	Sample size for chance constraints
su, sd	-	Auxiliary continuous variables to model HP constraints
TD, TU	h	Minimum HP down- and up-time
u	W/m^2 or -	Vector of HVAC system control inputs
$u_{\mathrm{s}}, u_{\mathrm{th}}$	W/m^2	HP thermal power setpoint and actual value
u_{th}^l	W/m^2	HP thermal power based on the linear model
v	°C or W/m^2	Vector of weather and occupancy disturbances
w	-	Uncertain SFC signal (scalar)
$\mathbf{w}$	-	Uncertain SFC signal along an horizon (vector)
x	°C	Vector of building states
T	h	Averaging period of energy-constrained SFC signals
α, σ	-	Auxiliary variables for the poles of the HP model
Δu	W/m^2	Heat flux disturbance due to SFC
$\varepsilon_+, \varepsilon_-$	-	Bias coefficients of energy-constrained SFC signals
ϵ	-	Violation probability for the chance constraints
λ	-	Dual variable of an optimization problem
$\lambda^{\mathrm{c}}, \lambda^{\mathrm{e}}$	CHF/kW/h	Reserve capacity and reserve energy payment
λ_1, λ_2	-	Poles of the HP thermal model
ρ	-	Correlation of two random variables
τ_{d}	s	Dead-time of the HP model
τ_f	s	Lower cutoff period of the band-pass filter
$\tau_1 - \tau_4$	s	Time constants of the HP model
ξ	-	Efficiency vector for HVAC control inputs

ing [118–122, 131]. More relevant to this chapter are [123–125] that proposed methods to estimate the SFC reserve capacity of individual commercial buildings.

We follow this line of research and investigate how SFC reserves can be provided by *aggregations* of commercial buildings. The problem we

Table 9.2: Nomenclature of Chapter 9: sets and probability operators

Operator	Description
$\mathbb{U}$	Constraint set for HVAC inputs
$\mathbb{X}$	Constraint set for building states
$\mathbb{W}_{PC}$	Uncertainty set for conventional SFC signals
$\mathbb{W}_{PEC}$	Uncertainty set for energy-constrained SFC signals
$\text{cov}[\cdot]$	Covariance of random variables
$\mathbb{E}[\cdot]$	Expected value of a random variable
$\mathbb{P}[\cdot]$	Probability of a random variable

Table 9.3: Nomenclature of Chapter 9: subscripts and superscripts

Symbol	Description
b	Building index within the aggregation
d	Desired power of HP aggregation
i	Row index in a matrix constraint
k, j, l	Time indexes
Lv2, Lv3	HVAC setpoints in level 2 and level 3
min, max	Minimum and maximum values of a variable
s	Short-term power schedule of HP aggregation
$*$	Optimal value of a variable
$+, -$	Down- and up-reserve in levels 1 and 2
$\uparrow, \downarrow$	Down- and up-reserve in the short-term capacity allocation
$\overline{(\cdot)}, \underline{(\cdot)}$	Upper and lower bound on a variable
$[\cdot]_+, [\cdot]_-$	Caps a variable to non-negative and non-positive values

consider is how a building aggregation can determine the amount of reserves it can collaboratively offer, and how the buildings should be controlled to achieve reserve provision without (or with minimal) occupant discomfort.

For this purpose, we develop a control framework for scheduling and provision of SFC reserves that relies on a three-level hierarchical control scheme. In level 1 an aggregator determines the aggregation's reserve capacity, in level 2 a robust Model Predictive Control (MPC) scheme optimizes the building's HVAC system setpoints, and in level 3 a feedback controller tracks the SFC signal. The proposed framework can handle energy-constrained SFC signals, i.e., signals that are not zero-mean but their integral over time is bounded, and it additionally allows us to model the SFC signal with scenarios generated from available statistics.

The remainder of this chapter is organized as follows. In Section 9.2 we provide a high-level introduction to the control framework and in Section 9.3 we introduce the building models. Sections 9.4, 9.5, and 9.6 present the control formulations in the three levels of the hierarchy. Sections 9.7, 9.8, 9.9, and 9.10 present extensive simulation and sensi-

tivity analysis results, Section 9.11 discusses practical implementation aspects, and Section 9.12 concludes. The methods and results of this chapter are based on [175–178]. The nomenclature of this chapter is given in Tables 9.1 - 9.2, and bold symbols correspond to vectors and matrices.

9.2   Reserve Scheduling and Provision Framework

9.2.1   Problem Statement and the Role of Aggregation

The framework's goals are (a) to enable an aggregation of buildings to schedule the reserve capacity that they can collaboratively offer to the Transmission System Operator (TSO), and (b) to operate the buildings such that the reserves are provided with sufficient accuracy. This is a challenging task for a number of reasons. First, the reserve capacity must be reliably available at any time. Second, to provide reserves the buildings must deviate from the energy-optimal operation [177], and so there exists a tradeoff between maximizing the amount of reserves and minimizing energy consumption. Third, reserve provision should not compromise the primary goal of building control, which is to maintain an acceptable comfort level for the occupants. And last, tracking the SFC signal should not rely on excessive communication between the TSO and the buildings to facilitate real-world implementations.

In contrast to [122, 123, 127, 133, 179] that assumed individual buildings participating in the reserve market, we are interested in building aggregations for two main reasons. First, aggregators act as intermediaries between the reserve market and the buildings, and can take over the following responsibilities: (i) determination of the reserve capacity and bidding in the market, (ii) interaction with the TSO during reserve activation, and (iii) a-posteriori financial settlement with the TSO. Without an aggregator, these tasks would be a large overhead for the building managers.

Second, many reserve markets have requirements on the minimum size of the reserve bid, which cannot be met by individual buildings [70]. Moreover, reserve markets usually have additional requirements such as
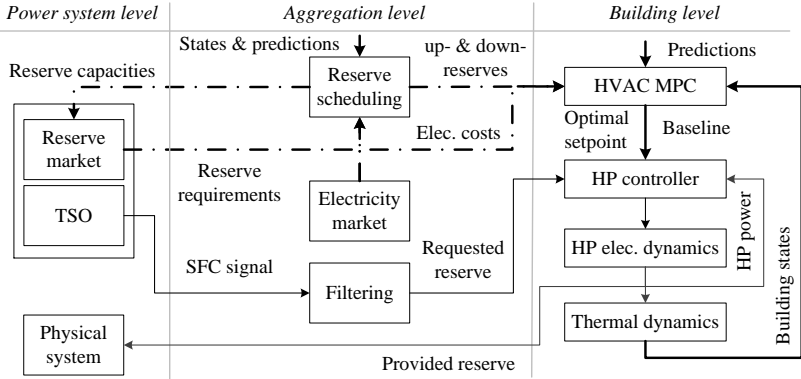
Figure 9.1: Overview of the control framework. Thin arrows are real-time signals (level 3), thick arrows are signals every 1 hour (level 2), and thick-dashed arrows are daily signals (level 1).

symmetric bids (equal up- and down-reserve capacities)[1], and minimum duration of the bid capacity. If each building calculated its capacity separately and without any coordination, it would be very unlikely that the aggregate reserve capacity meets the aforementioned requirements. In contrast, an aggregator can exploit the different building characteristics and coordinate the reserve allocation to meet these requirements and maximize the reserve capacity.

9.2.2 Hierarchical Control and High-Level Problem Formulation

To address the aforementioned challenges, we develop the hierarchical control scheme of Fig. 9.1, which consists of three levels. Level 1 is related to reserve scheduling, level 2 to building control, and level 3 to SFC signal tracking.

---

[1]In the context of SFC, the term up-reserves denotes increase of a generator's production or decrease of a load's consumption to increase system frequency. Similarly, the term down-reserves denotes decrease of a generator's production or increase of a load's consumption to decrease system frequency.

Level 1 - Reserve Scheduling

On a daily basis, an aggregator determines the optimal reserve capacity of the building aggregation, as well as the contribution of each building to the total capacity. The aggregator's goal is to minimize the total cost along the scheduling horizon defined as the sum of electricity cost and profit in the reserve market. The reserve capacity is determined by solving a centralized optimization problem given electricity and reserve costs, the current building states, and predictions of disturbances. The reserve scheduling optimization problem incorporates the building dynamics, the HVAC system and comfort constraints, and market requirements such as symmetry of the reserve capacity.

Level 2 - Building Control

In this level, control setpoints for the HVAC system are determined locally at each building typically every 15 minutes to 1 hour. This is achieved by an MPC controller with the goal of minimizing energy consumption, while leaving enough slack to provide reserves, if requested. The MPC formulation is similar to level 1 and uses as inputs the electricity cost, the current building states, predictions of disturbances, as well as the reserve capacities fixed by the aggregator in level 1.

Level 3 - SFC Signal Tracking

In real time (typically every few seconds), the TSO sends the SFC signal to the aggregator who passes it to the buildings, possibly after appropriate filtering. Each building modifies the power consumption of the HVAC system to track the signal using a feedback controller. For this purpose, we rely on control of the Heat Pump (HP) electric power by modifying the refrigerant's flow rate via valves [118]. The reference power of the HP controller at each time step is calculated based on (i) the MPC setpoint from level 2, (ii) the building's reserve capacity from level 1, and (iii) the SFC signal.

It is important to note that there is no communication between the aggregator and the buildings in level 2. In level 3 the only necessary communication is the SFC signal broadcast, but no state feedback from the buildings to the aggregator is needed. This is a merit of the centralized formulation in level 1, which reduces the communication burden in

real time to a minimum. Detailed mathematical formulations for levels 1, 2 and 3 of the control framework will be presented in Sections 9.4, 9.5 and 9.6.

9.3   Modeling

An appropriate building model for our application should consider both the slow thermal dynamics, i.e., the heat transfer and temperature evolution within the building, as well as the HP fast electric dynamics, i.e., the HP's response to step changes in the power reference during reserve provision. Developing a single model that accounts for both the slow and fast dynamics would be rather impractical, and it is not necessary.

The thermal model is needed in levels 1 and 2 to predict the temperature evolution using time steps in the range 15 min to 1 hour. The HP electric dynamics are much faster and, due to the time-scale separation, we can assume that the HP operates at steady-state during each time step of levels 1 and 2. Therefore, we develop a fast HP model for use in level 3, and a separate thermal model for use in levels 1 and 2 in which the HP's steady-state depends on the properties of the fast HP model.

9.3.1   Slow Thermal Building Dynamics

We consider buildings where heating, cooling, ventilation, blinds, and lighting are jointly controlled with MPC, and represent their thermal dynamics using the 12th order multiple-input-multiple-output bilinear model presented in Section 2.2.1. Using an established sequential linearization approach [24], the model can be transformed to a linear time-varying system of the form

$$x_{k+1} = Ax_k + B_{u,k}u_k + B_{v,k}v_k + R\Delta u_k \ . \tag{9.1}$$

The vector $x_k \in \mathbb{R}^{n_x}$ denotes the model states at time step k, i.e., the room, wall, floor and ceiling temperatures, all measured in °C. The vector $u_k \in \mathbb{R}^{n_u}$, denotes the HVAC control inputs, i.e., heating and cooling power, ventilation, blind position, and lighting. We model the effect of heating and cooling on system states using the heat fluxes as control inputs in the thermal model instead of the HP control command, which simplifies the model significantly and is justified by the time-scale separation explained before. The heat fluxes and the lighting power are

normalized by the floor area and measured in $\mathrm{W/m^2}$, whereas the blind position is a number between 0 (fully closed) and 1 (fully open).

There are two sources of disturbances that affect the building states. First, weather and occupancy disturbances (denoted by $v_k \in \mathbb{R}^{n_v}$), namely the ambient temperature in $^\circ$C, solar radiation in $\mathrm{W/m^2}$, and internal heat gains by occupants and equipment in $\mathrm{W/m^2}$. Second, the SFC signal w_k is also a disturbance because it modifies the HP consumption during reserve provision and consequently perturbs the heat fluxes in the building. We denote by $\Delta u_k \in \mathbb{R}^{n_r}$ the heat flux disturbance due to the SFC signal, where $n_r \leq n_u$ is the number of actuators that provide reserves.

A building can provide *up-reserves* by decreasing its consumption, and *down-reserves* by increasing it. Let us denote by $r_{e,k}$ the electric reserve capacity of the building at time step k. Since the HVAC control input for heating and cooling is defined as a heat flux, it is convenient to define also the "thermal" reserve capacity r_k (in $\mathrm{W/m^2}$). The electric reserve capacity $r_{e,k}$ can be obtained from r_k by division with the Coefficient of Performance (COP). For notational convenience, we use the variable r_k in the problem formulations and call it simply reserve capacity, keeping in mind that it is actually the "thermal" reserve capacity. Typically, the TSO requests the reserve energy from the reserve providers as a percentage of the reserve capacity by broadcasting a normalized SFC signal $w_k \in [-1, 1]$. Therefore, the heat flux disturbance can be parameterized as $\Delta u_k = r_k w_k$.

The thermal model (9.1) is used in level 1 to determine the capacities r_k, and in the MPC of level 2 to calculate the optimal HVAC control inputs u_k. Reserve provision is subject to HVAC input and comfort zone constraints of the form $u_{\min} \leq u_k + \Delta u_k \leq u_{\max}$ and $x_{\min} \leq x_k \leq x_{\max}$, respectively, that we write compactly as

$$\mathbb{U}_k = \{u_k : G(u_k + \Delta u_k) \leq g\} \tag{9.2}$$
$$\mathbb{X}_k = \{x_k : Fx_k \leq f\}, \tag{9.3}$$

for appropriate matrices G and F and vectors g and f. Note that the bounds on x_k can be time-varying (F_k, f_k) to account for comfort zone setbacks during non-working hours.

To allow for more flexibility when determining u_k and r_k, we parameterize the heating/cooling control inputs based on past reserve requests

with the affine disturbance feedback on w_k

$$u_k = h_k + \sum_{j=0}^{k-1} H_{k,j} w_j \ , \tag{9.4}$$

where h_k is the scheduled input (the heat fluxes that would occur if the SFC signal was zero), and $H_{k,j} \in \mathbb{R}^{n_u}$ for $j \in [0, k-1]$ are column vectors that comprise the feedback policy matrix $H \in \mathbb{R}^{n_u \times k}$. Although an affine disturbance feedback could be also applied for the weather and occupancy disturbances v_k, this is not done here because it has been widely explored in the literature of building climate control [24].

9.3.2 Fast Heat Pump Dynamics

We consider buildings with water-based HVAC systems, where water is heated or cooled centrally with an HP, and then distributed to different zones via pipes.[2] The desired heat flux u_k in (9.1) is obtained by controlling the HP's thermal power output, which depends on the compressor's electric power consumption. Using the COP, we can calculate the necessary compressor power to obtain the desired heat flux. On the device level, the compressor power can be controlled by modifying the refrigerant's flow rate via valves or inlet guide vanes.[3]

In this chapter, we do not rely on a first-principles model to describe the relationship between the refrigerant's flow rate and HP power. Instead, we assume that a low-level controller controls the flow given a compressor power reference, and we derive a data-driven model for the closed-loop system with the power reference as an input and the actual electric and thermal power as outputs. The model is non-linear and consists of linear terms and dead-times that capture the delays in mechanical and electrical parts (valve, compressor, motor), as well as a rate limiter to account for the non-linearity in the response. The

---

[2]Water-based HVAC systems are very common in Europe. The consumption of water circulation pumps is typically small, and the fresh air flow rate is usually constant to avoid occupants' discomfort, thus reserves are offered by the HP. However, the control framework of Sections 9.4, 9.5 and 9.6 applies also to air-based systems.

[3]Using inlet guide vanes is a standard practice to control the power of fixed-speed centrifugal compressors. In case of compressors with variable speed drives, higher efficiency can be achieved by combined control of the inlet guide vane position and compressor speed [180]. The part load operation of the compressor affects its efficiency and the overall COP in a non-linear way [118]. However, a constant COP is assumed in this thesis because a non-linear COP model would result in non-convex optimization problems.
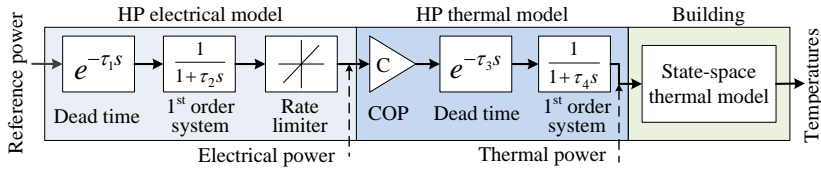
Figure 9.2: Overview of the heat pump and building thermal dynamics.
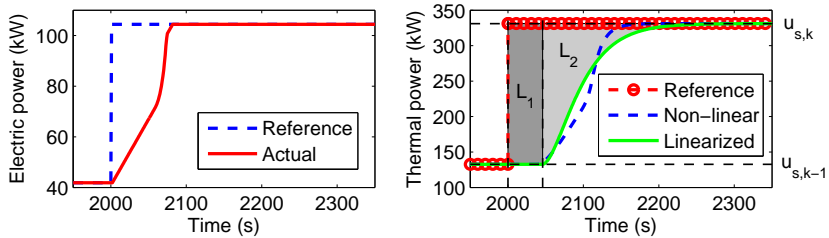
Figure 9.3: HP step response. Left: Electric power from the non-linear model. Right: Thermal power from both the non-linear and the linearized model.

model is shown in Fig. 9.2, where $\tau_1 - \tau_4$ are time constants fitted using experimental data from [118].

In Fig. 9.3, we show the HP's electric power response $P_k(t)$ and the thermal power response $u_{\text{th},k}(t)$ to a step-up reference[4] $u_{\text{s},k}$. Notice that there is a significant delay between $u_{\text{s},k}$ and $u_{\text{th},k}(t)$: the HP needs 200 seconds to reach the new thermal power setpoint. The building thermal model (9.1) assumes that the HP operates in steady-state, i.e., that the setpoint $u_{\text{s},k}$ has been reached. Since the time it takes to reach the setpoint is not negligible, and because the effective thermal power provided to the building is less than the desired setpoint during the transient phase, the model (9.1) will overestimate the building temperature. Therefore, the HP delay must be considered when determining the reserve capacities and HP schedules in level 1 and level 2 to ensure occupants' comfort.

However, it is not necessary to account for the exact dynamic evolution of $u_{\text{th},k}(t)$. This would require a time step in the order of seconds in level 1 and level 2, and would drastically increase the computational

---

[4]The thermal power setpoint $u_{\text{s},k}$ is the heat flux created by the HP, and it is part of the HVAC control input vector u_k in (9.1).

complexity. Instead, it is sufficient to estimate the amount of "lost thermal energy" in the right plot of Fig. 9.3, i.e., the area between the reference power (curve with circles) and the actual thermal power (dashed curve), and account for it in level 1 and level 2. An approach to do so is presented in the next section.

9.3.3 Accounting for Heat Pump Dynamics in the Thermal Model

The lost thermal energy due to HP dynamics consists of two parts $L_1(t)$ and $L_2(t)$. The first part $L_1(t)$ is due to the HP dead-times and corresponds to the dark-grey area in Fig. 9.3, whereas $L_2(t)$ is due to the HP delays and rate limits and corresponds to the light-grey area in Fig. 9.3. We propose to account for $L_1(t)$ and $L_2(t)$ in three steps: (a) linear approximation of the HP thermal dynamics, (b) analytical expression for $L_1(t)$ and $L_2(t)$ based on the linear model, and (c) augmentation of the state vector of building model (9.1) to include $L_1(t)$ and $L_2(t)$.

The first step is to approximate the thermal HP dynamics with an over-damped 2nd order linear model with a dead-time τ_{d} and two negative real poles λ_1 and λ_2 given by

$$\lambda_{1,2} = -\sigma \pm \alpha \tag{9.5}$$

$$\sigma = \omega_0 \zeta \tag{9.6}$$

$$\alpha = \omega_0 \sqrt{\zeta^2 - 1} \; , \tag{9.7}$$

where $\zeta > 1$ is the damping ratio and ω_0 is the natural frequency. We expose the non-linear HP model of Fig. 9.2 to a sequence of step commands and identify the parameters ζ and ω_0 that maximize the fit between the linear and non-linear models. The step response of the identified linear model is shown with the solid curve in Fig. 9.3. Neglecting the dead-time, the linear model's thermal power output $y(t) = u_{\text{th},k}^l(t)$ for a step input $u_{\text{s},k}$ and with initial conditions $y(0) = u_{\text{s},k-1}$ and $\dot{y}(0) = 0$ is the solution of the Ordinary Differential Equation (ODE)

$$\ddot{y}(t) + 2\zeta\omega_0\dot{y}(t) + \omega_0^2 y(t) = \omega_0^2 u_{\text{s},k} \; . \tag{9.8}$$

In a second step, the ODE is solved and $y(t) = u_{\text{th},k}^l(t)$ is analytically derived. With reference to Fig. 9.3, the lost thermal energy $L_2(t)$ is

calculated as the area between $u_{\text{th},k}^l(t)$ and $u_{\text{s},k}$, which results in the expression (the calculation steps are given in Appendix A)

$$L_2(t) = \int_0^t \left[u_{\text{th},k}^l(\tau) - u_{\text{s},k} \right] d\tau = a(t) \cdot u_{\text{s},k-1} + b(t) \cdot u_{\text{s},k} , \qquad (9.9)$$

where the functions $a(t)$ and $b(t)$ are defined based on the linear model parameters and are given by

$$a(t) = \frac{\lambda_2 \cdot (e^{\lambda_1 t} - 1)}{\lambda_1 \cdot (\lambda_2 - \lambda_1)} - \frac{\lambda_1 \cdot (e^{\lambda_2 t} - 1)}{\lambda_2 \cdot (\lambda_2 - \lambda_1)} \qquad (9.10)$$

$$b(t) = \frac{1}{\sigma^2 - \alpha^2} \cdot \left[2\sigma \left[1 - e^{-\sigma t} \cosh(\alpha t) \right] - \left(\frac{\sigma^2 + \alpha^2}{\alpha} \right) e^{-\sigma t} \sinh(\alpha t) \right] . \qquad (9.11)$$

Equation (9.9) describes the time evolution of the lost thermal energy during the transient phase. Denote by Δt the discretization time step of level 1 and level 2. The cumulative lost thermal energy during Δt is obtained by evaluating the function $L_2(t)$ at $t = \Delta t$. For notational simplicity, we denote $L_2(\Delta t)$, $a(\Delta t)$, and $b(\Delta t)$ by L_2, a, and b, respectively, and so

$$L_2 = au_{\text{s},k-1} + bu_{\text{s},k} . \qquad (9.12)$$

The dead-time τ_{d} is approximately 50 seconds and typically smaller than Δt. Therefore, $L_1(t)$ for $t = \Delta t$ is equal to the area of the dark-grey rectangle in Fig. 9.3

$$L_1 = (u_{\text{s},k-1} - u_{\text{s},k})\tau_{\text{d}} . \qquad (9.13)$$

The effective thermal power provided to the building can be written as

$$u_{\text{th},k} = u_{\text{s},k} + L_1 + L_2 , \qquad (9.14)$$

where $L_1, L_2 \leq 0$ because they represent lost thermal energy. The input matrix of system (9.1) is $B_{u,k} = [B_k^{\text{r}} \ R]$, where R corresponds to the HP input and B_k^{r} corresponds to the rest of the actuators. Focusing only on the HP, neglecting the disturbance terms, and using the derived expressions for L_1 and L_2, the building thermal dynamics will evolve according to

$$\begin{aligned} x_{k+1} &= Ax_k + Ru_{\text{th},k} = Ax_k + R \cdot (u_{\text{s},k} + L_1 + L_2) \\ &= Ax_k + R \cdot (1 - \tau_{\text{d}} + b) \cdot u_{\text{s},k} + R \cdot (\tau_{\text{d}} + a)u_{\text{s},k-1} . \qquad (9.15) \end{aligned}$$

The last step of the procedure is to define

$$\bar{R} = R \cdot (\tau_{\mathrm{d}} + a) \tag{9.16}$$

$$\tilde{R} = R \cdot (1 - \tau_{\mathrm{d}} + b) , \tag{9.17}$$

and augment the state of the building thermal model as

$$\hat{x}_k = [x_k^\top \ u_{\mathrm{s},k-1}^\top]^\top , \tag{9.18}$$

which leads to the linear model

$$\hat{x}_{k+1} = \hat{A}x_k + \hat{B}_k u_k + \hat{E}v_k + \hat{R}\Delta u_k, \tag{9.19}$$

$$\hat{A} = \begin{bmatrix} A & \bar{R} \\ 0 & 0 \end{bmatrix}, \ \hat{B}_k = \begin{bmatrix} B_k^{\mathrm{r}} & \tilde{R} \\ 0 & 1 \end{bmatrix}, \ \hat{E} = \begin{bmatrix} E \\ 0 \end{bmatrix}, \ \hat{R} = \begin{bmatrix} R \\ 0 \end{bmatrix} . \tag{9.20}$$

The presented approach allows us to introduce the lost thermal energy due to fast HP dynamics in the building thermal model of level 1 and level 2 in a tractable way. For simplicity, we drop the $\hat{\cdot}$ notation for the augmented state in the rest of the chapter.

9.3.4 Modeling of Uncertain SFC Reserves

The SFC signal is a disturbance for the buildings, and we treat it as uncertain because it is unknown when the reserve capacities are determined. The SFC signal uncertainty affects both the HVAC input constraints (9.2) and the comfort constraints (9.3). In this section, we discuss various approaches to model the reserve uncertainty.

HVAC Input Constraints

The TSO typically requests the reserves via a normalized frequency signal $w_k \in [-1, 1]$, where $w_k = -1$ and $w_k = 1$ indicate full activation of up- and down-reserves, respectively. In practice, the TSOs require 100% availability of the reserve capacity because reserves are critical for power system security [72]. Thus, the heating and/or cooling actuators should be always able to modify their power consumption as much as the reserve capacity. Given a scheduling horizon N, the reserve uncertainty can be mathematically described by the box set

$$\mathbb{W}_{\mathrm{PC}} = \{\mathbf{w}_k \in \mathbb{R}^N : ||\mathbf{w}_k||_\infty \le 1\} , \tag{9.21}$$

which represents the signal's power bounds and is called Power Constraints (PC). To comply with the TSO requirements, we treat the HVAC input constraints as robust constraints that need to be satisfied for any $\mathbf{w}_k \in \mathbb{W}_{\text{PC}}$.

SFC Signal with Energy Constraints

In order to facilitate the participation of energy-constrained resources in frequency control, some TSOs offer reserve products for energy-constrained SFC signals [73]. We model energy constraints as linear constraints on the signal's mean value along an averaging period T

$$-\varepsilon_- \leq \frac{1}{T} \cdot \sum_{k=t}^{t+T-1} w_k \leq \varepsilon_+ \ , \tag{9.22}$$

where $\varepsilon_+, \varepsilon_-$, are the bias coefficients for down- and up-reserves, respectively. Equation (9.22) bounds the bias of the signal based on the triplet $(T, \varepsilon_-, \varepsilon_+)$, which is fixed by the TSO for a given product.

If we stack (9.22) along N, we get the polyhedral constraint on the uncertainty $\mathbf{A}_w \mathbf{w}_t \leq \mathbf{b}_w$, where $\mathbf{A}_w$ is a matrix with entries -1, 0 or 1, and $\mathbf{b}_w$ is a vector with entries $\varepsilon_- T$ or $\varepsilon_+ T$. The power constraints (9.21) are still present, since the full reserve capacity could be requested anytime. Denote by $\mathbf{I}_N$ the N-dimensional identity matrix, and by $\mathbf{1}_N$ the N-dimensional vector with ones. Defining $\bar{\mathbf{A}}_w = [\mathbf{A}_w; \mathbf{I}_N; -\mathbf{I}_N]$ and $\bar{\mathbf{b}}_w = [\mathbf{b}_w; \mathbf{1}_N; \mathbf{1}_N]$, the uncertainty set can be written as

$$\mathbb{W}_{\text{PEC}} = \{\mathbf{w}_t \in \mathbb{R}^N : \bar{\mathbf{A}}_w \mathbf{w}_t \leq \bar{\mathbf{b}}_w\} \ , \tag{9.23}$$

which we call Power and Energy Constraints (PEC).

Comfort Constraints

Different modeling approaches can be used for the comfort constraints depending on (a) how critical it is to respect the comfort zone, and (b) whether the SFC signal is energy-constrained or not. In certain cases (e.g., laboratories with sensitive equipment) the temperature might be strictly constrained within a range. In such cases, the comfort constraints can be modeled as robust using (9.21) for conventional unconstrained SFC signals or (9.22) for energy-constrained signals.

In most cases, however, respecting the comfort zone is not a critical requirement. Indeed, it is typical in office building control to interpret comfort constraints as chance (probabilistic) constraints that can be violated with a predefined small probability [24]. In such cases, one can think of $\mathbf{w}_k$ as a random variable with an unknown probability distribution, which can be approximated by generating scenarios based on historical SFC signals (conventional or energy-constrained).

If comfort constraints are modeled as robust constraints, then in the worst case $\mathbf{w}_k$ will be an extreme point of the uncertainty set (9.21) or (9.22). However, the realized $\mathbf{w}_k$ will usually lie in the interior of the set, and thus minimization for the worst-case realization will likely result in conservative solutions. In contrast, a probabilistic approach is expected to reduce the conservatism by disregarding scenarios that are hardly ever realized.

9.3.5   Weather, Occupancy, and Price Uncertainty

Apart from the SFC signal, there are two other sources of uncertainty in our problem: (a) weather and occupancy uncertainties, i.e., deviations from the predicted values, and (b) electricity and reserve price uncertainties. For simplicity, weather and occupancy predictions are assumed perfect in the formulations of Sections 9.4 and 9.5. There exist approaches in the literature to build robustness to model prediction errors [181] and to imperfect weather/occupancy predictions [24, 182], which are easy to integrate in our framework. Note, however, that we also report simulation results under weather uncertainty in Section 9.8.3.

We assume that the building aggregation acquires energy in the retail market, and since the retail tariffs do not change frequently, there is no electricity price uncertainty. However, the situation is different for the reserve prices. Although the capacity bid's price is chosen by the aggregator, the bid might not be accepted in the auction. Incorporating this uncertainty requires a reserve market model, which is beyond the scope of this thesis. In Switzerland, the reserve energy payments are coupled with the spot price. We assume perfect spot price forecasts and expect imperfect forecasts to have minor effects on the results because the capacity payments are dominant, especially for energy-constrained SFC signals. If necessary though, spot price uncertainty is straightforward to incorporate in the framework.

9.4   Level 1: Reserve Scheduling

Recall that the goal of level 1 is to determine the optimal reserve capacity of each building b within the aggregation $(r_k^{b,*})$ for each time step k of the next day. The reserve scheduling problem is solved centrally by the aggregator once a day, and the reserve capacities are communicated to the buildings. Since we use an affine disturbance feedback on the uncertain SFC signal w_k, the scheduled input h_k and the policy matrix $H_{k,j}$ $\forall j$ are also used as optimization variables. Although the optimal values of these variables $h_k^{b,*}$ and $H_{k,:}^{b,*}$ are computed, they are not used in level 1 of the control hierarchy, but instead they are recomputed in level 2.

In this section, we present four different formulations for the day-ahead reserve scheduling problem, as well as an additional short-term capacity allocation process. First, we assume symmetric reserve capacities and neglect HP compressor constraints. The formulations will be extended in Sections 9.4.5 and 9.4.6 to account for asymmetric reserves and HP constraints.

9.4.1   Robust Scheduling with Power Constraints

The first formulation models both the input constraints (9.2) and the comfort constraints (9.3) using the uncertainty set (9.21), and without considering energy-constrained SFC signals. This formulation is called Robust Problem with Power Constraints (RPC) and can be written as

$$\min_{h_k^b, H_{k,:}^b, r_k^b} \sum_{k=0}^{N_1-1} \left[c_k \sum_{b=1}^{L} \xi_b^\top u_k^b - \left(\lambda_k^c - \lambda_k^e \mathbb{E}[w_k] \right) \sum_{b=1}^{L} \xi_b^\top r_k^b \right] \qquad (9.24a)$$

$$\text{s.t. } x_{k+1}^b = A^b x_k^b + B_{u,k}^b u_k^b + B_v^b v_k^b + R^b w_k r_k^b, \ \forall b, \forall k \qquad (9.24b)$$

$$u_k^b = h_k^b + \sum_{j=0}^{k-1} H_{k,j}^b w_j, \ \forall b, \ \forall k \qquad (9.24c)$$

$$(u_k^b + w_k r_k^b) \in \mathbb{U}_k^b, \ \forall \mathbf{w}_k \in \mathbb{W}_{\text{PC}}, \ \forall b, \ \forall k \qquad (9.24d)$$

$$x_{k+1}^b \in \mathbb{X}_{k+1}^b, \ \forall \mathbf{w}_k \in \mathbb{W}_{\text{PC}}, \ \forall b, \ \forall k \qquad (9.24e)$$

$$\sum_{b=1}^{L} \xi_b^\top (r_{b,k} - r_{b,k+j}) = 0, \forall j \in \{1, \ldots, T_r\}, \forall k \qquad (9.24f)$$

$$r_k^b \geq 0, \ \forall b, \ \forall k \ , \qquad (9.24g)$$

where $b \in \{1 \ldots L\}$ is the building index, N_1 is the prediction horizon, c_k is the electricity cost, λ_k^c is the reserve capacity payment, λ_k^e is the reserve energy payment, and ξ_b is an efficiency vector that maps the heat fluxes to electric power.

The objective function (9.24a) optimizes the tradeoff between electricity costs and reserve profits along the prediction horizon by minimizing the expected value of the total cost. The expected reserve energy profit is equal to the product of the reserve energy payment λ_k^e and the average activated reserve energy. The latter is equal to the product of the reserve capacity $\sum_{b=1}^{L} \xi_b^\top r_k^b$ and the average normalized SFC signal $\mathbb{E}[w_k]$ (recall that $-1 \leq w_k \leq 1$). Constraints (9.24b) and (9.24c) are the building thermal dynamics and the affine feedback policy on w_k. Equation (9.24d) models the robust HVAC input constraints, (9.24e) models the robust comfort zone constraints, (9.24f) requires the capacity to be constant over a period T_r, whereas (9.24g) is a non-negativity constraint on the capacities. From (9.24c) and (9.24d), one can see that

$$u_k^b + w_k r_k^b = h_k^b + \sum_{j=0}^{k} H_{k,j}^b w_j \ , \tag{9.25}$$

with $H_{k,k}^b = r_k^b$, i.e., the reserve capacities can be seen as the diagonal entries of the policy matrix H. In order to solve problem (9.24) we must reformulate the robust constraints.

The robust constraint (9.24d) is reformulated applying standard techniques from robust optimization. Recall from (9.2) that $u_k^b + w_k r_k^b \in \mathbb{U}_k^b$ is equivalent to $G(u_k^b + w_k r_k^b) \leq g$, and denote by $G_i^\top$ the ith row of G and by g_i the ith element of g. The ith robust input constraint is reformulated according to

$$\max_{\mathbf{w}_k \in \mathbb{W}_{\text{PC}}} \left[G_i^\top (u_k^b + w_k r_k^b) \right] \leq g_i \quad \Rightarrow$$

$$G_i^\top h_k^b + \sum_{j=0}^{k} \left| \left(H_{k,j}^b \right)^\top G_i \right| \leq g_i \ . \tag{9.26}$$

By doing the same for all rows of G and repeating the same procedure for the state constraints, we result in a problem with l_1-norm constraints, which can be expressed as a Linear Program (LP) [183, 184].

9.4.2   Robust Scheduling with Power and Energy Constraints

In this formulation we consider energy-constrained SFC signals, and the comfort constraints (9.3) are modeled with the uncertainty set (9.23) in a robust way according to

$$x_{k+1}^b \in \mathbb{X}_{k+1}^b, \forall \, \mathbf{w}_k \in \mathbb{W}_{\text{PEC}} \; \Leftrightarrow \; F x_{k+1}^b \leq f, \forall \, \mathbf{w}_k \in \mathbb{W}_{\text{PEC}}. \quad (9.27)$$

To derive the robust counterpart problem, we first combine the dynamics (9.24b) with the policy equation (9.24c), and write the state at time step $k+1$ as

$$x_{k+1} = A^{k+1} x_0 + \sum_{j=0}^{k} A^{k-j} B_v v_j + \sum_{j=0}^{k} A^{k-j} B_{u,k} h_j$$
$$+ \sum_{j=0}^{k} \Big[\sum_{l=j+1}^{k} A^{k-l} B_{u,k} H_{l,j} + A^{k-j} R r_j \Big] w_j \ , \qquad (9.28)$$

where the building index b is dropped for clarity. Combining (9.27) with (9.28), the ith row of the state constraint can be written in the form

$$\overbrace{F_i^\top \Big[A^{k+1} x_0 + \sum_{j=0}^{k} A^{k-j} B_v v_j + \sum_{j=0}^{k} A^{k-j} B_u h_j \Big]}^{\tilde{f}_i}$$
$$+ \underbrace{\sum_{j=0}^{k} F_i^\top \Big[\sum_{l=j+1}^{k} A^{k-l} B_u H_{l,j} + A^{k-j} R r_j \Big] w_j}_{\mathbf{w}_k^\top Y_i \tilde{r}_k} \leq f_i \ . \qquad (9.29)$$

Denote by $\tilde{f}_i$ the first term of (9.29), which does not depend on $\mathbf{w}_k$. The second term of (9.29) can be written as $\mathbf{w}_k^\top Y_i \tilde{r}_k$, where $\tilde{r}_k$ is a column-wise vectorized version of H and Y_i is an appropriate matrix. The interested reader is referred to [177] for details. Since (9.29) must hold $\forall \, \mathbf{w}_k \in \mathbb{W}_{\text{PEC}}$, we write (9.29) as the following maximization problem over $\mathbf{w}_k$

$$\max_{\mathbf{w}_k} \big(\mathbf{w}_k^\top Y_i \tilde{r}_k \big) + \tilde{f}_i \leq f_i \qquad (9.30a)$$

$$\text{s.t.} \quad \bar{\mathbf{A}}_w \mathbf{w}_k \leq \bar{\mathbf{b}}_w \ . \qquad (9.30b)$$

Using standard procedures [183, 184] we obtain the dual of (9.30)

$$\min_{\lambda_i} \quad \bar{\mathbf{b}}_w^\top \lambda_i + \tilde{f}_i - f_i \leq 0 \tag{9.31a}$$

$$\text{s.t.} \quad \bar{\mathbf{A}}_w^\top \lambda_i = Y_i \tilde{r}_k \tag{9.31b}$$

$$\lambda_i \geq 0 , \tag{9.31c}$$

where λ_i are the dual variables. Finally, we drop the minimization in (9.31a) and substitute each row of (9.27) with the constraints (9.31a)–(9.31c), i.e., we relax the optimization problem (9.31). Any feasible λ_i in (9.31) will provide an upper bound of the maximization in (9.30) and, due to strong duality of LP, this relaxation will be tight.

Therefore, for each building b and time step k, the robust constraint (9.27) is replaced by the inequality and equality constraints

$$\bar{\mathbf{b}}_w^\top \lambda_{i,k}^b + \tilde{f}_i^b - f_i^b \leq 0 \qquad \forall i, \ \forall b, \ \forall k \tag{9.32a}$$

$$\bar{\mathbf{A}}_w^\top \lambda_{i,k}^b = Y_i^b \tilde{r}_k^b \qquad \forall i, \ \forall b, \ \forall k \tag{9.32b}$$

$$\lambda_{i,k}^b \geq 0 \qquad \forall i, \ \forall b, \ \forall k , \tag{9.32c}$$

where i indexes the rows of matrix F of the state constraints. The number of constraints per building and time step depends on the dimensions of $\bar{\mathbf{A}}_w$ and $\bar{\mathbf{b}}_w$, which in turn depend on the prediction horizon N_1 and the averaging period T. In addition, the dimensions of $\bar{\mathbf{A}}_w$ and $\bar{\mathbf{b}}_w$ depend on the dimension of $\mathbf{w}_k$, and so they are different for each k. The input constraints are reformulated using (9.26), exactly as was done for RPC. The resulting problem is called Robust Problem with Power and Energy Constraints (RPEC).

9.4.3 Hybrid Robust and Stochastic Scheduling with Power Constraints

In this hybrid formulation the input constraints are still modeled as robust constraints with (9.24d), but the state constraints are modeled as chance constraints that need to be satisfied with a predefined probability. Therefore, we substitute (9.24e) with

$$\mathbb{P}\big[x_{k+1}^b \in \mathbb{X}_{k+1}^b\big] \geq 1 - \epsilon, \ \forall b, \ \forall k , \tag{9.33}$$

where the probability $\mathbb{P}$ is over the uncertain SFC signal w_k, and ϵ is the violation probability. The chance constraint (9.33) limits the

number of comfort zone violations, but not the size of the violations. In practice, large comfort zone violations while providing SFC reserves are unlikely due to thermal inertia (especially if the SFC signal is energy constrained).

Reformulating the chance constraint (9.33) into a tractable constraint is challenging. One way of doing this is by using the so-called scenario approach, which approximates the chance constraint by randomly drawing a finite number of independent and identically distributed (i.i.d.) samples of the uncertain variable. An advantage of the scenario approach is that no specific knowledge of the underlying probability distribution of the uncertain variable (w_k in our case) is required. The samples of the uncertain variable are used to obtain a finite number of sampled state constraints.

The challenge in this so-called scenario approach is to select the number of samples such that the chance constraints are satisfied with a probability of at least $1 - \epsilon$. General bounds for convex programs were given in [185], whereas [186] proposed tighter bounds for multi-stage optimization problems with affine uncertainty.

Denote by n_f the number of state constraints for each time step k. If the chance constraints are applied separately for each building, then we have $n_f = 2$ to describe the upper and lower comfort zone limits. If the sample size at time step k satisfies

$$S_k \geq \frac{n_f(k+2)}{\epsilon} - 1 \;, \tag{9.34}$$

then the optimal solution of the sampled problem satisfies the chance constraint. For a proof of (9.34) and technical details, the reader is referred to [176]. We generate a total number of

$$N_{\mathrm{s}} = \sum_{k=0}^{N_1-1} S_k \tag{9.35}$$

samples (scenarios) using historical SFC signals, where the nth sample $\mathbf{w}_k^{(n)}$ contains values for the time steps from 0 to k. Each sample $\mathbf{w}_k^{(n)}$ is inserted into (9.28) and generates a sampled state trajectory $x_{k+1}^{b,n}$. Substituting each sampled state trajectory into $x_{k+1}^b \in \mathbb{X}_{k+1}^b$, the chance constraint (9.24e) can be replaced by N_{s} linear inequalities

$$\mathbb{P}\big[x_{k+1}^b \in \mathbb{X}_{k+1}^b\big] \geq 1 - \epsilon, \; \forall b, \; \forall k \quad \Leftrightarrow$$
$$F^b x_{k+1}^{b,n} \leq f^b, \; \forall b, \; \forall k, \; \text{for } n = [1, S_k] \;. \tag{9.36}$$

By replacing (9.24d) with (9.26), and (9.24e) with (9.36), we result in a tractable LP, which we call Stochastic Problem with Power Constraints (SPC) in the rest of this chapter.

9.4.4 Hybrid Robust and Stochastic Scheduling with Power and Energy Constraints

This formulation, which is called Stochastic Problem with Power and Energy Constraints (SPEC), is mathematically identical with the one of Section 9.4.3. The only difference is that the scenarios are now for energy-constrained SFC signals. If sufficient historical data exist, they can be used directly to generate scenarios. Otherwise, one can pass the historical unconstrained SFC signals through a high-pass filter with a time constant equal to the averaging period T in order to generate the scenarios.

In the rest of this section we show how asymmetric reserves and HP constraints can be introduced into (9.24).

9.4.5 Asymmetric reserve capacities

RPC Formulation

It is straightforward to introduce asymmetric reserves for each building in (9.24). Let us define $r^b_{+,k} \geq 0$ as the down-reserve capacity and $w_{+,k} \geq 0$ as the positive part of the SFC signal. Similarly, we define $r^b_{-,k} \geq 0$ as the up-reserve capacity and $w_{-,k} \leq 0$ as the negative part of the SFC signal. If we fix the diagonal elements of the policy matrix H to zero, the robust input constraint (9.26) will be

$$G_i^\top h_k^b + \sum_{j=0}^{k} \left| \left(H_{k,j}^b \right)^\top G_i \right| + \max_{\mathbf{w}_k \in \mathbb{W}_{\mathrm{PC}}} \left[\Delta u_k \right] \leq g_i \ . \qquad (9.37)$$

Due to non-negativity of $r^b_{-,k}$ and $r^b_{+,k}$, the maximization in (9.37) is trivial: the worst case for the $u_k + \Delta u_k \leq u_{\max}$ part of the constraints is $\max \left[\Delta u_k \right] = r^b_{+,k}$, whereas the worst case for the $u_{\min} \leq u_k + \Delta u_k$ part of the constraints is $\max \left[\Delta u_k \right] = r^b_{-,k}$. Exactly the same procedure is followed for the robust state constraints.

RPEC Formulation

With energy constraints, the problem becomes more complicated because the non-linear constraint $w_{-,k} \cdot w_{+,k} = 0$ is needed to ensure that up- and down- reserves are not requested simultaneously. Without including this constraint, the worst-case reserve request might consist of $w_{-,k}$ and $w_{+,k}$ that are both non-zero, which cannot happen in reality.[5] However, adding this constraint to (9.23) leads to an uncertainty set which is not a polyhedron any more, and so the dualization technique of Section 9.4.2 cannot be applied to derive the robust counterpart problem.

SPC and SPEC Formulations

Approximating the chance constraints with sampling allows us to easily account for asymmetric reserves. The sampled state constraints are still in the form of (9.36), where $x_{k+1}^{b,n}$ is now given by

$$x_{k+1} = A^{k+1}x_0 + \sum_{j=0}^{k} A^{k-j}B_v v_j + \sum_{j=0}^{k} A^{k-j}B_{u,k}h_j \qquad (9.38)$$

$$+ \sum_{j=0}^{k}\sum_{l=j+1}^{k} A^{k-l}BH_{l,j}w_j^{(n)} + A^{k-j}R\Big(r_{+,j}\big[w_j^{(n)}\big]_+ + r_{-,j}\big[w_j^{(n)}\big]_-\Big) ,$$

where $[\cdot]_+ \geq 0$ and $[\cdot]_- \leq 0$ cap their arguments to non-negative and non-positive values, respectively, and the index b is dropped for clarity.

Apart from increasing the number of optimization variables, considering asymmetric reserves does not change the structure of the reserve scheduling problem. Note that even with asymmetric reserve capacities on the building level, symmetric reserves in aggregate can be achieved imposing constraints of the form

$$\sum_{b=1}^{L} \xi_b^\top r_{+,k}^b = \sum_{b=1}^{L} \xi_b^\top r_{-,k}^b . \qquad (9.39)$$

---

[5]It is easy to show that without energy constraints, one of $w_{-,k}$ and $w_{+,k}$ will always be zero in the worst-case reserve request. For this reason, the non-linear constraint $w_{-,k} \cdot w_{+,k} = 0$ is not required in the RPC formulation.

9.4.6 Modeling of Heat Pump Constraints

Minimum compressor loading constraints can be expressed by introducing binary auxiliary variables $d_k^b \in \{0,1\}$ to keep track of the on/off state of each HP b. The input constraints (9.24d) are now written as

$$d_k^b \ u_{\min}^b \leq u_k^b + w_k r_k^b \leq d_k^b \ u_{\max}^b, \forall \mathbf{w}_k \in \mathbb{W}_{\mathrm{PC}}, \ \forall b, \forall k \ . \qquad (9.40)$$

The binary variables d_k^b do not alter the way the uncertainty w_k enters the constraints. Therefore, the reformulation (9.26) still holds [184] with the only difference that g_i at the right hand side is substituted with $[-u_{\max}^b \ u_{\min}^b]^\top d_k^b$.

Frequent HP cycling reduces its COP and, possibly, the compressor's lifetime. For these reasons, minimum up-times (TU) and/or down-times (TD) might be required. The auxiliary binary variables su_k^b and sd_k^b that represent the start-up and shut-down events of the HPs, respectively, are needed to model TU and TD [187]. More specifically, $su_k^b \in [0,1]$ takes the value 1 if HP b starts up at time step k and 0 otherwise, whereas $sd_k^b \in [0,1]$ takes the value 1 if HP b shuts down at time step k and 0 otherwise. Using su_k^b, sd_k^b, we model TU and TD as

$$\sum\nolimits_{k-TU+1}^{k} su_k^b \leq d_k^b, \ \sum\nolimits_{k-TD+1}^{k} sd_k^b \leq 1 - d_k^b \qquad (9.41)$$

$$d_k^b - d_{k-1}^b = su_k^b - sd_k^b \ . \qquad (9.42)$$

Equation (9.42) forces su_k^b and sd_k^b to take binary values, since d_k^b is defined as a binary variable. Therefore, su_k^b and sd_k^b can be declared as continuous variables in the formulation of the optimization problem. Constraints (9.40)-(9.42) will turn the reserve scheduling problem into a Mixed-Integer Linear Program (MILP).

9.4.7 Short-Term Capacity Allocation

The result of (9.24) is a sequence of reserve capacities with step changes, for example every 1 hour. Such step changes at the beginning of each hour will likely result in poor SFC signal tracking due to the HP ramping constraints [188]. To avoid this, we adjust the short-term reserve capacities (at the beginning of each hour) taking into account the HP ramping constraints, but without considering the building thermal dynamics. Note that the adjusted capacities of all buildings sum up to the capacities calculated by the day-ahead reserve scheduling problem.

The short-term capacity allocation is formulated as the convex optimization problem (9.43), where N_{st} is the prediction horizon; $P_{\text{s},t}^b, r_{\uparrow,t}^b, r_{\downarrow,t}^b$ denote the short-term HP schedules, and down- and up-reserve capacities; $P_{\text{d},t}$ is the desired aggregate HP power; $\hat{w} = 0.85$ and $\check{w} = 0.5$ are "bad case" normalized SFC signals; $\Delta\hat{w} = 0.005$ is a "bad case" increase/decrease rate of the signal between two consecutive time steps; and $n_{\text{u}}^b, n_{\text{d}}^b$ are the maximum increase and decrease rates of HP power, respectively.

Note that (9.43) is solved for every scheduling interval k of the day-ahead reserve scheduling problem of Sections 9.4.1 to 9.4.4. The desired power $P_{\text{d},t}$ is defined as an 8-minute ramp from the total scheduled HP power at step $k-1$ to the total scheduled HP power at step k.

We choose $\hat{w} = 0.85$ and $\check{w} = 0.5$ instead of the worst case values $\hat{w} = \check{w} = 1$ to allow for more flexibility (actually $\check{w} \leq 0.5$ holds for 90% of the time). In general, n_{u} and n_{d} are non-linear functions of $P_{\text{s},t}^b$, which can be approximated in a piece-wise linear fashion. Due to the minimum loading constraints, and in case of symmetric reserve capacities, two pieces are enough to model the non-linearity: $n_{\text{u}} = 0.25\%/\text{s}$, $n_{\text{d}} = 0.2\%/\text{s}$ (% of HP rated power) for low HP operating points, and $n_{\text{u}} = 0.5\%/\text{s}$, $n_{\text{d}} = 0.8\%/\text{s}$ for high HP operating points.

$$\min_{P_{\text{s},t}^b, r_{\uparrow,t}^b, r_{\downarrow,t}^b} \sum_{t=1}^{N_{\text{st}}} \left[\left| P_{\text{d},t} - \sum_{b=1}^{L} P_{\text{s},t}^b \right| + \sum_{b=1}^{L} \left(P_{\text{s},t}^b - P_{\text{s},t-1}^b \right)^2 \right] \tag{9.43a}$$

$$\text{s.t.} \quad \underline{P}_t^b \leq P_{\text{s},t}^b \leq \overline{P}_t^b, \ \forall b, \forall t \tag{9.43b}$$

$$\underline{r}_{\uparrow,t}^b \leq r_{\uparrow,t}^b \leq \overline{r}_{\uparrow,t}^b, \ \underline{r}_{\downarrow,t}^b \leq r_{\downarrow,t}^b \leq \overline{r}_{\downarrow,t}^b, \ \forall b, \forall t \tag{9.43c}$$

$$P_{\text{s},t}^b + \hat{w} r_{\uparrow,t}^b \leq P_{\max}^b, \ P_{\text{s},t}^b - \hat{w} r_{\downarrow,t}^b \geq P_{\min}^b, \ \forall b, \forall t \tag{9.43d}$$

$$\sum_{b=1}^{L} r_{\uparrow,t}^b = \sum_{b=1}^{L} r_{+,k}^b, \sum_{b=1}^{L} r_{\downarrow,t}^b = \sum_{b=1}^{L} r_{-,k}^b, \forall t \tag{9.43e}$$

$$P_{\text{s},t}^b - P_{\text{s},t-1}^b \leq n_{\text{u}}^b P_{\max}^b - \Delta\hat{w} \max\{r_{\uparrow,t}^b, r_{\downarrow,t}^b\}$$
$$- \check{w} \max\{r_{\uparrow,t}^b - r_{\uparrow,t-1}^b, r_{\downarrow,t-1}^b - r_{\downarrow,t}^b, 0\}, \forall b, \forall t \tag{9.43f}$$

$$P_{\text{s},t}^b - P_{\text{s},t-1}^b \geq n_{\text{d}}^b P_{\max}^b + \Delta\hat{w} \max\{r_{\uparrow,t}^b, r_{\downarrow,t}^b\}$$
$$+ \check{w} \max\{r_{\uparrow,t-1}^b - r_{\uparrow,t}^b, r_{\downarrow,t}^b - r_{\downarrow,t-1}^b, 0\}, \forall b, \forall t \ . \tag{9.43g}$$

The first term of (9.43a) minimizes the deviation of the HP aggregate power from $P_{\text{d},t}$, while the second term keeps the resulting short-term HP schedules smooth. Constraints (9.43b) and (9.43c) impose lower and

upper bounds on $P_{s,t}^b, r_{\uparrow,t}^b, r_{\downarrow,t}^b$, which depend on the solution of the day-ahead reserve scheduling problem, on HP ramp limits, and on whether a HP starts up, shuts down, or simply changes its setpoint. Constraints (9.43d) ensure that the HP schedules and reserve capacities respect the minimum ($P_{\min}^b$) and maximum ($P_{\max}^b$) HP power. Constraints (9.43e) enforce that the short-term reserve capacities add up to the fixed capacities of the day-ahead reserve scheduling problem. Finally, (9.43f) and (9.43g) ensure that the short-term reserve allocation respects the HP ramp limits for the considered bad cases of the SFC signal and its rate of change.

9.5 Level 2: Building Climate Control

In level 2, an MPC controller determines the HVAC control inputs that minimize energy consumption, while ensuring that the scheduled reserves can be provided. The MPC runs typically every 15 minutes to 1 hour, and it has four main differences with the reserve scheduling problem of Section 9.4. First, the MPC problem is solved separately and locally at each building. Second, the optimization is performed only over the HVAC control inputs h_k^b and feedback policy $H_{k,:}^b$, because the reserve capacities are fixed to $r_k^{b,*}$ from level 1. Third, the only objective is to minimize the electricity cost. And fourth, the constraints (9.24f), (9.24g) on the reserve structure are not present, because the reserve is fixed. The MPC formulation is dependent on the level 1 formulation and can be any of RPC, RPEC, SPC, and SPEC.

RPC Formulation

The robust MPC formulation with power constraints (RPC) is

$$\min_{h_k^b, H_{k,:}^b} \sum_{k=0}^{N_2-1} c_k \cdot \left(\xi_b^\top u_k^b\right) \tag{9.44a}$$

$$\text{s.t. } x_{k+1}^b = A^b x_k^b + B_{u,k}^b u_k^b + B_v^b v_k^b + R^b w_k r_k^b, \forall k \tag{9.44b}$$

$$u_k^b = h_k^b + \sum_{j=0}^{k-1} H_{k,j}^b w_j, \ \forall k \tag{9.44c}$$

$$u_k^b + w_k r_k^b \in \mathbb{U}_k^b, \ \forall \mathbf{w}_k \in \mathbb{W}_{\text{PC}}, \ \forall k \tag{9.44d}$$

$$x_{k+1}^b \in \mathbb{X}_{k+1}^b, \ \forall \mathbf{w}_k \in \mathbb{W}_{\text{PC}}, \ \forall k \ . \tag{9.44e}$$

The robust counterpart of (9.44) can be obtained by applying exactly the same procedure as in Section 9.4.1. Since the reserve capacities are fixed, the diagonal elements of the policy matrix H in (9.26) are also fixed in this case.

RPEC Formulation

In the presence of energy constraints, a tractable robust MPC formulation (RPEC) can be obtained by applying the dualization technique of Section 9.4.2. The constraint $x_{k+1}^b \in \mathbb{X}_{k+1}^b$, $\forall \mathbf{w}_k \in \mathbb{W}_{\mathrm{PEC}}$ is reformulated to

$$\min_{\lambda_i} \quad \bar{\mathbf{b}}_{w,t}^\top \lambda_i + \tilde{f}_i - f_i \leq 0 \tag{9.45a}$$

$$\text{s.t.} \quad \bar{\mathbf{A}}_{w,t}^\top \lambda_i = Y \tilde{r}_k^* \tag{9.45b}$$

$$\lambda_i \geq 0 \ , \tag{9.45c}$$

where $\bar{\mathbf{A}}_{w,t}, \bar{\mathbf{b}}_{w,t}$ are time-varying. For a time step t in the averaging interval $[t_1, t_2]$ of length T, (9.22) can be written as

$$-\varepsilon_- T - w_{\mathrm{p},t} \leq \sum_{k=t}^{t_2} w_t \leq \varepsilon_+ T - w_{\mathrm{p},t} \tag{9.46}$$

$$w_{\mathrm{p},t} = \sum_{k=t_1}^{t-1} w_k \ , \tag{9.47}$$

where $w_{\mathrm{p},t}$ is known because the uncertainty up to $t - 1$ is realized. Thus, the coupling constraint on $\{w_t, \ldots, w_{t_2}\}$ depends on the energy content of the SFC signal in the previous time steps of the averaging interval, and this is reflected in the time-varying matrices $\bar{\mathbf{A}}_{w,t}$ and $\bar{\mathbf{b}}_{w,t}$.

SPC and SPEC Formulations

If the state constraints are modeled as chance constraints, the MPC problem is very similar to that of level 1 for both unconstrained SFC signals (SPC) and energy-constrained signals (SPEC). Of course, the SFC scenarios might be updated at each MPC time step, if new information becomes available.

Declaration of Baseline Power

The first input of the optimal control sequence of the MPC determines the HVAC system setpoints for the next time step

$$u_k^{b,\text{Lv2}} = h_k^{b,*} \ . \tag{9.48}$$

Although a consumption schedule $h_k^{b,*}$ is calculated already in level 1, the MPC of level 2 can reduce both the cost due to less uncertainty (recent reserve requests are known), as well as the constraint violations in case of plant-model mismatch. The HP thermal power setpoint $u_{\text{s},k}^{b,\text{Lv2}}$ is an element of $u_k^{b,\text{Lv2}}$ and can be translated to an electric power setpoint according to

$$P_k^{b,\text{Lv2}} = \frac{u_{\text{s},k}^{b,\text{Lv2}}}{\text{COP}} \ . \tag{9.49}$$

Equation 9.49 provides us with an estimate of the building's baseline power consumption (i.e., the power consumption without providing SFC) ahead of time, which is beneficial for the financial settlement with the aggregator or the TSO.

Note that the MPC problem might be infeasible for two reasons. The first reason is imperfections in the building model used within the MPC, or imperfect predictions of weather and/or occupancy. The second reason is that the comfort constraints might be violated with a small probability due to the probabilistic interpretation of the comfort constraints in SPC and SPEC. To ensure that a feasible solution can always be found and achieve continuous control without interruptions, we model the comfort constraints as soft constraints and penalize the constraint violation with a very high penalty to ensure that the MPC will not choose to violate the constraints.

9.6   Level 3: Frequency Signal Filtering and Tracking

In level 3, the HP electric power is controlled around $P_k^{b,\text{Lv2}}$ to track the SFC signal. The power reference of HP b is

$$P_k^{b,\text{Lv3}} = P_k^{b,\text{Lv2}} + w_k r_{e,k}^{b,*} \tag{9.50}$$

$$r_{e,k}^{b,*} = \frac{r_k^{b,*}}{\text{COP}} \; , \qquad (9.51)$$

where the electric reserve capacity $r_{e,k}^{b,*}$ is fixed from level 1. The set-point $P_k^{b,\text{Lv3}}$ is tracked controlling the refrigerant's flow rate with a Proportional-Integral (PI) controller, which creates the heat flux

$$u_k^{b,\text{Lv3}} = u_k^{b,\text{Lv2}} + w_k r_k^{b,*} \; . \qquad (9.52)$$

If there is no mismatch between the model and the real building, the room temperature will be within the comfort zone always for the RPC and RPEC formulations, and with probability at least $1 - \epsilon$ for the SPC and SPEC formulations.

The tracking quality depends on HP delays and ramping limits, as well as on how fast the SFC signal is changing. Two options to improve the tracking quality are envisaged. First, the TSO could split the SFC signal into a number of products with different ramping requirements using an appropriate filter bank, and the buildings could choose the reserve product with the appropriate ramping requirements. Second, if specialized reserve products are unavailable, the buildings could form a pool with faster resources (e.g., an aggregation of EWHs), exclude very high frequency components of the SFC signal using an appropriate band-pass filter, and send them to the faster resources. We use a band-pass Chebyshev filter with a lower cutoff period τ_f to consider these two options in our simulations.

9.7 Investigation Setup

9.7.1 Building Parameters

We investigate the performance of the proposed control framework in simulations with an aggregation of typical Swiss office buildings. Two HVAC systems are considered: in system A, heating is performed via radiators, and cooling with cooled ceilings; in system B, both heating and cooling are performed using Thermally Activated Building Systems (TABS). We also differentiate between heavy or light building envelope, high or low window area fraction, and high or low internal gains. In our simulations, we consider an aggregation of 6 large buildings ($15,000$ m^2 each) with the characteristics of Tables 9.4 and 9.5. More information regarding the buildings can be found in [24, 175, 177, 178].

Table 9.4: Commercial building configurations

Building type	HVAC system	Building envelope	Window area fraction	Internal gains
A1	A	heavy	high	high
A2	A	heavy	low	low
A3	A	light	low	low
B1	B	heavy	high	high
B2	B	heavy	low	low
B3	B	light	low	low

Table 9.5: Heat pump parameters ($U[a, b]$ denotes the uniform probability distribution with bounds a and b)

Parameter	Value	Parameter	Value
Rated power	$U[95, 205]$ kW	τ_1, τ_2	$U[0.5, 3]$ s
Minimum loading	$U[15, 30]$ %	τ_3	$U[25, 75]$ s
COP	$U[2.5, 3.5]$	τ_4	$U[10, 20]$ s

Typical occupancy profiles are used and weather data are provided by Meteoswiss (the Swiss federal office of meteorology and climatology). The temperature comfort zone during working hours is $[21, 24]°$C in winter and $[22, 25]°$C in summer, which is in accordance with the ASHRAE 55-2013 standard [189]. During non-working hours and weekends, the comfort zone is relaxed to $[12, 35]°$C in both seasons.[6] The reserve energy payments in (9.24) were modeled using spot price data from the Swiss market in winter 2013.

We perform the simulations in MATLAB and solve the optimization problems with CPLEX. The reserve scheduling and MPC problems are solved with prediction horizons $N_1 = 2$ days and $N_2 = 1$ day, respectively. The optimization time step is either 30 minutes or 1 hour depending on the simulation case. The short-term capacity allocation problem (9.43) is solved with an horizon of $N_{st} = 600$ seconds and a time step of 1 second.

---

[6]Such a wide comfort zone is typically used for unoccupied periods of time in MPC controllers for energy-efficient building operation [24]. Nevertheless, in practice the temperature remains much higher than $12°$C and much lower than $35°$C.

Table 9.6: Bias coefficients (ε) for the normalized SFC signal ($-1 \leq w_k \leq 1$) and its high-frequency component (HF) for different averaging periods (T)

T	1 h	2 h	4 h	6 h	8 h	12 h
SFC 2009	1.000	0.989	0.952	0.796	0.592	0.505
HF 2009	0.528	0.467	0.337	0.273	0.384	0.229
SFC 2012	0.927	0.781	0.674	0.624	0.553	0.448
HF 2012	0.382	0.300	0.317	0.290	0.237	0.203

9.7.2   Parameters of Energy Constraints

To apply energy constraints in the form of (9.22), we consider six different averaging periods ($T = \{1, 2, 4, 6, 8, 12\}$ hours) and determine the respective bias coefficients ε_- and ε_+ based on the historical normalized SFC signals from the Swiss control area in 2009 and 2012. For this purpose, we use the Chebyshev filter

$$H(z) = \frac{\sum_{i=0}^{n_{\mathrm{f}}} b_i z^{-i}}{1 + \sum_{i=1}^{n_{\mathrm{f}}} a_i z^{-i}} \ , \tag{9.53}$$

where n_{f} is the filter's order (fixed to $n_{\mathrm{f}} = 3$ that showed a good performance in preliminary simulations), and a_i, b_i are its coefficients that can be computed based on the pass-band edge frequency f_{c} [190], which is related to the averaging period with $f_{\mathrm{c}} = 1/T$.

For each value of T, we filter the historical SFC signals using the corresponding filter (9.53) and get the high-frequency (HF) part of the signals for 2009 and 2012. For each of the four signals and for each value of T, we calculate ε_+ and ε_- based on the largest and the smallest average values of the signal w_k, respectively, over any time interval with duration T according to

$$\varepsilon_+ = \max \left(\frac{1}{T} \cdot \sum_{k=1}^{T} w_k \right) \tag{9.54}$$

$$\varepsilon_- = -\min \left(\frac{1}{T} \cdot \sum_{k=1}^{T} w_k \right) \ . \tag{9.55}$$

The results are summarized in Table 9.6, where the reported ε values are computed using

$$\varepsilon = \max \left(\varepsilon_-, \varepsilon_+ \right) \ . \tag{9.56}$$
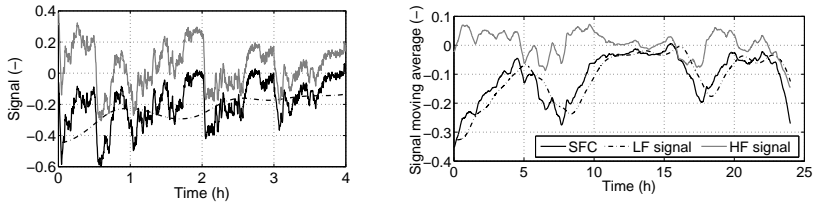
Figure 9.4: Left: A 4-hour extract of the original SFC signal, its low-frequency (LF), and its high-frequency (HF) components for a sample day. Right: The moving averages of the signals for the whole sample day.

Observe that the original SFC signals are in the worst case biased over periods of several hours. Note that the biases of the HF signals are significantly lower than those of the original SFC signals, and that the signals in 2012 are generally less biased than in 2009. This is because from March 2012 the ACE of Switzerland is combined with that of other European countries before generating the SFC signal.

In Fig. 9.4, we show a 4-hour extract of the SFC, the low-frequency part of it (LF signal), and the high-frequency part of it (HF signal) for a sample day, as well as the 2-hour moving averages of the signals for the whole day. Although the original SFC signal is mostly negative, the bias is absorbed by the LF signal and so the HF signal crosses zero often. As shown in the right plot of Fig. 9.4, the largest bias of the SFC signal for the sample day is approximately 0.3 (to the negative direction), i.e., the worst-case reserve request averaged over periods of two hours is negative and approximately equal to 30% of the reserve capacity. On the other hand, the energy-constrained HF signal has a bias less than 0.15 during the same day, i.e., the worst-case average reserve request is less than 15% of the reserve capacity.

9.7.3   SFC Signal Scenarios

The SFC scenarios to reformulate the chance constraints are obtained based on historical SFC signal data in 2009. The data are available with a resolution of 10 seconds, but are averaged over 1 hour intervals
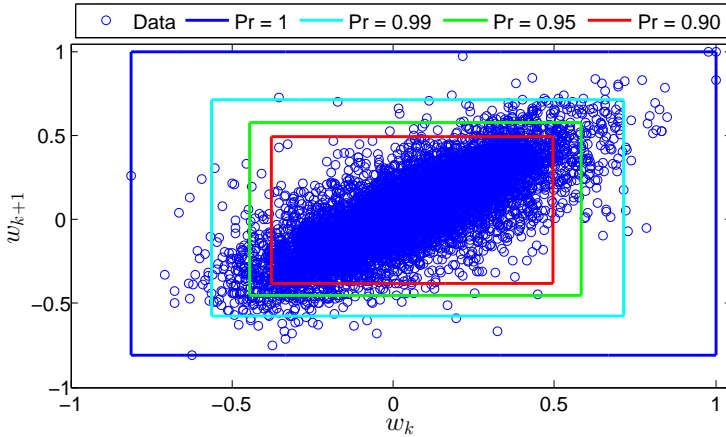
Figure 9.5: Scatter plot of historical SFC signals. Colored boxes correspond to different probability levels.

to fit our reserve scheduling framework.[7] Figure 9.5 shows a scatter plot of all measured SFC signal pairs. Each circle corresponds to a pair (w_k, w_{k+1}), where $k \in \{1, \ldots, 8760\}$ indexes each hour during the year. The data are strongly positively correlated and centered around $(0, 0)$, with a few outliers taking extreme values.

The blue box in Fig. 9.5 contains all realized regulation signals. It is interesting to notice that from the four theoretical worst-case combinations of (w_k, w_{k+1}) of the form $(\pm 1, \pm 1)$, only the combination $(1, 1)$ is observed in the data. The cyan, green, and red boxes contain 99%, 95%, and 90% of all historical data, respectively. Observe that the volumes of the boxes shrink rapidly if we accept a small risk of not satisfying the comfort zone requirements for all reserve demands. For example, if we consider the cyan box, any SFC signal pair (w_k, w_{k+1}) will lie in the rectangle defined by the vertices $(-0.56, -0.57)$, $(-0.56, 0.71)$, $(0.71, 0.71)$, $(0.71, -0.57)$ with a probability of 99%. Although Fig. 9.5 shows only consecutive signal pairs, our analysis can be repeated in higher dimensions, where the volume of the probability-boxes will shrink even faster. This indicates that if we allow small violations of the state (comfort) constraints, we could restrict ourselves to a smaller subset of the original

---

[7]Recall that the scenarios are used in level 1 and level 2 of the control hierarchy. The necessary information at this level of control is the bias of the SFC signal during each scheduling interval.
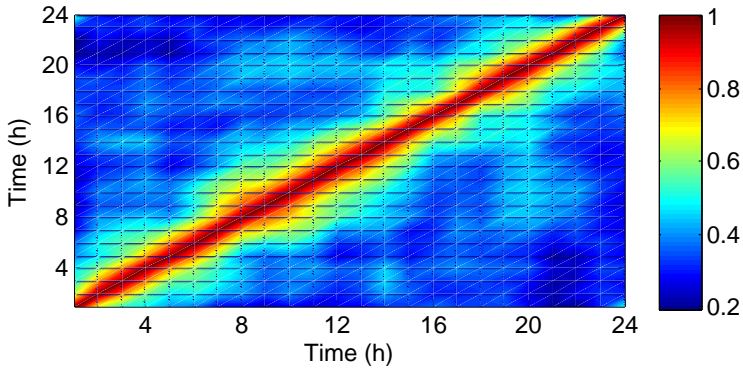
Figure 9.6: Empirical correlation matrix of the SFC signal.

uncertainty set, which in turn allows us to increase the reserve capacity and obtain a higher reserve profit.

Apart from the size of the probability-boxes and the frequency of occurrence of extreme events, the temporal interdependence of the SFC signals within a day is important. Figure 9.6 visualizes the autocorrelation matrix of the random variable $\mathbf{w}$. The correlation coefficient between the random variables w_i and w_j at two different time steps i and j is calculated with

$$\rho_{w_i,w_j} = \frac{\text{cov}(w_i, w_j)}{\sigma_i \sigma_j} \ , \tag{9.57}$$

where $\text{cov}(\cdot)$ denotes the covariance operator, and σ_i denotes the standard deviation of w_i. We see from Fig. 9.6 that the correlation values sharply decrease as the gap between the time steps increases. Moreover, the patterns indicate a complex interdependence of the SFC signal along the horizon, which can vary significantly during the day.

From this analysis we can conclude that a hybrid robust/stochastic reserve formulation, which takes into account the probability distribution of the SFC signal when enforcing chance constraints on the states, may potentially lead to significantly higher reserve capacities compared with a fully robust approach, where the state constraints are required to be satisfied for all admissible reserve demands. There are two main reasons for this. First, we may ignore highly unlikely SFC signals that form the extreme points of the uncertainty set, and therefore limit the
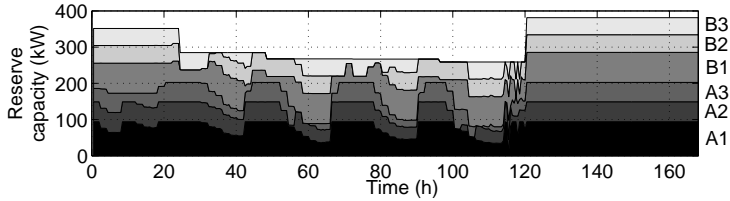
Figure 9.7: Optimal reserve allocation among the buildings during a winter
week (the plot starts on Monday).

reserve capacity. Second, the use of historical data naturally preserves
the correlation information between reserve requests at different times,
motivating a sampling-based reformulation of the chance-constrained
problem from a practical point of view.

9.8   Building Aggregation Simulation Results

9.8.1   Reserve Scheduling and Building Control

In this section, we demonstrate how the control framework can be used
to estimate the amount of reserves of a building aggregation, assess
the quality of SFC signal tracking, and investigate the effect of reserve
provisioning on building control.

Reserve Capacity Scheduling and Allocation

Fully Robust Formulation (RPEC):

In Fig. 9.7 we present the optimal reserve capacity of the aggregation
and its allocation among the buildings for a winter week. These results
correspond to simulations with the RPEC formulation assuming sym-
metric reserves at each building with a daily duration (constant reserve
capacity over a day). An extract of the high-frequency part of the SFC
signal of 2012 is used, which is obtained using the averaging period
$T = 2$ h and the bias coefficient $\varepsilon = 0.3$. The capacity payments are
10% higher than the electricity price ($\lambda_k^{\mathrm{c}} = 1.1 \cdot c_k$), whereas no reserve
energy payments are considered ($\lambda_k^{\mathrm{e}} = 0$).
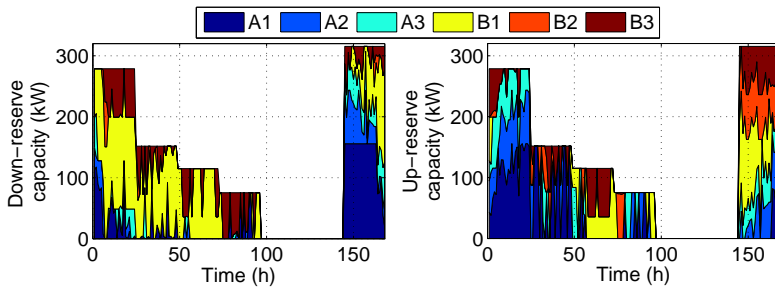
Figure 9.8: Optimal allocation of down-reserves (left) and up-reserves (right) during a winter week (the plot starts on Monday).

The capacity is constant for each day and ranges from approximately 260 kW on Friday to 380 kW on Saturday with an average weekly value of 313 kW. Note that the capacity is shifted among the buildings in a way that maximizes the total capacity of the aggregation. Interestingly, the buildings offer higher reserve capacities when they normally consume less power. For example, buildings with HVAC system A contribute mainly at night, whereas buildings with HVAC system B offer more reserves during working hours, because they prefer to preheat at night. The scheduled energy consumption of the aggregation is spread throughout the whole week to maximize the reserve potential. During the weekend, the buildings are not occupied and the comfort zone is relaxed, which results in higher reserve capacities compared with working days.

Hybrid Robust/Stochastic Formulation (SPEC):

Figure 9.8 shows the reserve scheduling results for the same winter week, but using the SPEC formulation and a violation probability $\epsilon = 0.2$. The reserve capacities are now asymmetric at each building but symmetric in aggregate. The capacity is assumed constant over a day, the capacity payments are fixed to a value 30% higher than the electricity price ($\lambda_k^c = 1.3 \cdot c_k$), and reserve energy payments are considered.

The total capacity has an average weekly value of 134 kW, but it highly depends on the building and ambient conditions, and varies from zero to approximately 300 kW. This illustrates the importance of daily SFC reserve bidding that gives more flexibility to buildings compared with weekly reserve auctions, which are still applied in several countries, e.g., in Switzerland [70].

Although higher capacity payments are selected for the simulations with the SPEC formulation compared with the simulations with the RPEC formulation, the average reserve capacity in Fig. 9.8 is lower than that of Fig. 9.7. This happens because the minimum loading constraints of HP compressors are taken into account in the simulations with the SPEC formulation, but not in the simulations with the RPEC formulation.

Estimation of Necessary Aggregation Size:

Based on our simulations in winter, the considered 6 buildings with floor area 15,000 m² and average rated heating power 27 W/m² provide a reserve capacity equal to 134 kW or 313 kW on average, depending on the assumptions. Using simple linear extrapolation, we find that the required minimum reserve capacity of 5 MW in Switzerland can be offered by approximately $100 - 225$ (similar) buildings.

Note that this is only a rough estimate because: (a) it is based on the average reserve capacity values; and (b) the extrapolation might overestimate the required number of buildings to meet the 5 MW limit, because the larger the aggregation the more flexibility exists in allocating reserves among buildings. However, it is clear that the amount of reserves per building and the necessary aggregation size depend highly on the formulation and the underlying assumptions. In Sections 9.9 and 9.10 we will investigate these dependencies in detail.

Building Climate Control

We show results from the daily operation of building A1 in Figs. 9.9 and 9.10 using the SPEC formulation. Figure 9.9 shows the heating power of the HVAC system and Fig. 9.10 shows the temperature trajectories. The blue-dashed and black-solid curves indicate the scheduled HVAC power and temperature in levels 1 and 2, respectively, the red-solid curves correspond to the values after tracking the SFC signal (level 3), and the green-solid curves are the HVAC and comfort constraints.

Since building A1 offers mainly up-reserves (consumption decrease), the room temperature in level 3 is lower than the scheduled value in level 1 and level 2. The grey envelope in Fig. 9.9 is the robust region, meaning that the HVAC power can move anywhere within this region without violating the constraints. The grey envelope in Fig. 9.10 shows the temperature trajectories that correspond to 50,000 randomly generated SFC signal scenarios. The envelope spanned by the scenarios is narrow because the energy content of the SFC signal is limited. The comfort
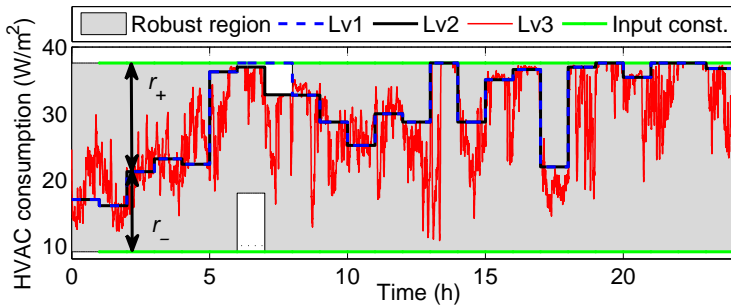
Figure 9.9: HVAC consumption in levels 1, 2 and 3 during a winter day.
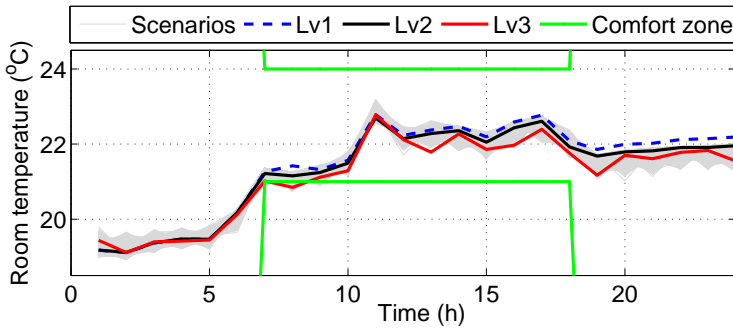
Figure 9.10: Room temperature in levels 1, 2 and 3 during a winter day.

zone is violated only for 1.9% of the scenarios, thus the chance constraint is satisfied with a probability much higher than the design probability of $1 - \epsilon = 0.8$.

SFC Signal Tracking

In Fig. 9.11, we show the SFC signal tracking over a 2-hour period using a filter with lower cutoff period $\tau_f = 40$ seconds. The tracking is generally good but it deteriorates due to HP ramping limits in two cases: (a) when the signal changes very rapidly, and (b) when the HP schedules change at the beginning of each hour. The three hour changes are indicated with a rectangle in Fig. 9.11; in the first two, the short-term capacity allocation problem (9.43) is not used, whereas in the third one it is considered. In the latter case, the performance is significantly
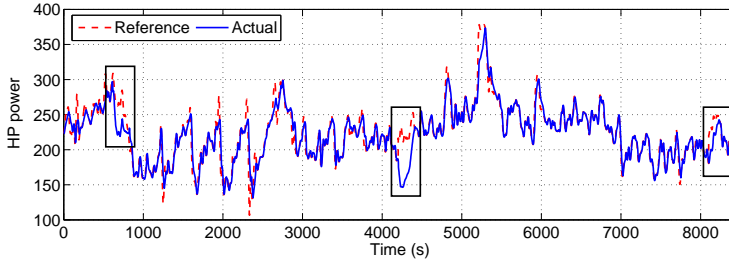
Figure 9.11: Tracking of the SFC signal by the aggregation's HPs over a period of two hours.

better and the tracking Mean Absolute Percentage Error (MAPE) is less than 5% during the second half of the 2-hour period.

9.8.2   Effect on Energy Consumption

To maximize the reserve capacity, the buildings operate in a less energy-efficient way. For example, during the considered winter week reserve provision using the SPEC formulation resulted in a consumption 60% higher than that of energy-efficient building control, i.e., an MPC with the objective of minimizing electricity cost without offering reserves. In order to provide reserves in both directions, the buildings try to operate close to the middle of the heating/cooling device's power range. This is in contrast to an energy efficient operation, where the power consumption would be as close as possible to the minimum value. However, the increase in energy consumption does not mean that the building control is suboptimal. For the given electricity and reserve prices, this building operation minimizes the total cost defined as the sum of electricity cost and reserve profit.

In fact, there is a tradeoff between increasing the reserve capacity and reducing the energy consumption, which depends on the capacity payment. To illustrate this, we report the increase in energy consumption for different reserve capacity levels (expressed as a percentage of the maximum possible capacity) in Fig. 9.12a. Note that reserve capacities up to 15% of the maximum capacity do not practically increase the energy consumption, that 50% of the maximum capacity can be offered with approximately 30% increase in energy consumption, and that in
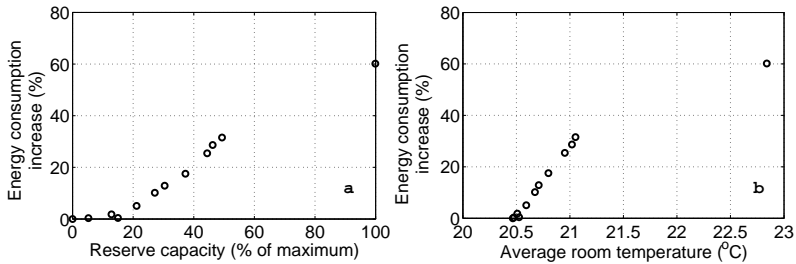
Figure 9.12: Fig. 9.12a (left): The dependence of increase in energy consumption on the reserve capacity offered by the buildings. Fig. 9.12b (right): The effect of increase in energy consumption on the average room temperature.

order to maximize the reserve capacity the energy consumption must be increased by 60%.

The effect of increased energy consumption is a higher average room temperature during the considered winter week, as shown in Fig. 9.12b. In theory, the increase in energy consumption is proportional to the room temperature increase with a proportionality constant equal to the building's thermal resistance. This linearity is reflected in the data points of Fig. 9.12b. The rightmost data point of Fig. 9.12b does not fall on the same line as the other points because the blinds are set to different positions, which affects the thermal resistance.

Note that the reported increase in energy consumption is in comparison with a building with an energy-efficient MPC. If the comparison was performed considering supervisory rule-based controllers, which are used in most buildings today, a smaller increase in energy consumption would be expected.

In addition, note that the increase in energy consumption is qualitatively different to the results of [137]. The authors of [137] calculated the round-trip efficiency of buildings while offering fast DR services, i.e., the efficiency loss due to the *utilization* of reserves. On the contrary, the reported increase in energy consumption in this chapter is the result of scheduling the building consumption in a way that large amounts of reserves can be offered reliably. The energy consumption increase that we report will occur even if no reserve is actually called, i.e., it is the price to pay for the *availability* of reserves.

9.8.3   Effect of Imperfect Disturbance Predictions

The SPEC formulation guarantees the satisfaction of the comfort zone with a given probability for any admissible reserve request and perfect weather and occupancy predictions. Although this probabilistic guarantee does not hold in case of imperfect predictions, some intuition on the expected comfort zone violations can be provided.

As shown in Fig. 9.10, the room temperature is close to the middle of the comfort zone under reserve provision in order to maximize the potential for up and down regulation. For this reason, we expect weather and occupancy prediction errors to lead to fewer comfort zone violations compared with an MPC-based energy-efficient building control, where the temperature remains close to one of the comfort zone limits. Indeed, we repeated the simulations with imperfect weather predictions and observed comfort zone violations for 9.3% of the time when the building operated under reserve provision, and 12.9% of the time when the building operated with an energy-efficient MPC.

9.9   Comparison of Reserve Scheduling Methods

The goal of this section is to compare the four reserve scheduling formulations with respect to the reserve capacity, baseline energy consumption, total cost (sum of electricity cost, capacity and reserve energy profits), and computation time. For this purpose, we run simulations for two weeks in winter assuming that each building offers symmetric reserve capacities that are constant for each day. The capacity payments are fixed to 110% of the electricity price ($\lambda_k^c = 1.1 \cdot c_k$). Furthermore, no HP minimum loadings or up- and down-times are considered, and the violation probability is fixed to $\epsilon = 0.1$. An energy-constrained SFC signal with averaging period $T = 2$ h and bias coefficients $(\varepsilon_+, \varepsilon_-) = (0.336, 0.467)$ is assumed. The simulation results are presented in Table 9.7.

For conventional SFC signals, the hybrid robust/stochastic SPC formulation increases the reserves by 5.6% and reduces the cost by 5% compared with the robust RPC formulation. The improvement is less pronounced for energy-constrained signals, where the SPEC formulation increases the reserves by 3.6% and reduces the cost by 1.2% compared

Table 9.7: Comparison of robust and stochastic formulations

Method	Reserve (MW)	Consumption (MWh)	Cost (CHF)	Time (s) Mean	Max
RPC	109.5	143.6	10719	10	112
RPEC	115.2 (+5.2%)	148.9 (+3.7%)	9960 (−7.1%)	23	134
SPC	115.6 (+5.6%)	149.6 (+4.1%)	10177 (−5.0%)	290	518
SPEC	119.2 (+8.8%)	152.9 (+6.5%)	9831 (−8.3%)	216	454

with the RPEC formulation. Not surprisingly, higher reserve capacities require a higher baseline energy consumption; however, the energy consumption increases at a lower rate than the reserve capacity. The performance improvement achieved by the SPC and SPEC formulations comes at the cost of longer computation times. The average computation time of SPC is approximately 30 times higher than that of RPC, whereas it takes approximately 10 times longer to solve SPEC compared with RPEC. The SPEC formulation achieves the maximum savings of 8.3% without further increase in computation time compared with SPC.

We also compared closed-loop formulations (where all diagonal and lower-diagonal elements of the policy matrix H in (9.4) are free optimization variables) against simpler open-loop formulations without affine disturbance feedback (in this case the policy matrix H is constrained to be a diagonal matrix, because all lower-diagonal elements are fixed to zero). Our results show that the closed-loop formulation slightly outperforms the open-loop formulation for the RPC case, and achieves a small cost reduction of 0.6% with only a marginal increase in computation time. On the other hand, the closed-loop formulation does not reduce the cost compared with the open-loop formulation for the RPEC, SPC, and SPEC cases, while it increases computation time by up to a factor of 3.

9.10   Sensitivity Analysis

9.10.1   Capacity and Reserve Energy Payments

The proposed reserve scheduling methods identify the optimal tradeoff between minimizing energy consumption and leaving enough slack for
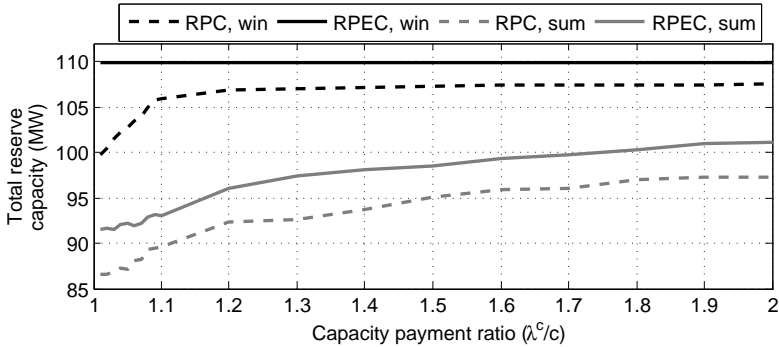
Figure 9.13: Bid curves of the building aggregation in winter and summer for the RPC and RPEC formulations and ratios $\lambda^c/c > 1$.

reserve provision. The buildings would not have any incentive to deviate from the energy-efficient control and offer reserves if the additional electricity cost occurring due to this deviation was higher than the reward received for providing reserves. The reward comes from capacity payments λ_k^c (the remuneration for the *availability* of each kW of reserve capacity) and reserve energy payments λ_k^e (the remuneration for the *utilization* of each kWh of reserves by the TSO).

In this section, we investigate the relationship between reserve payments and reserve capacity by running simulations over 2-week periods in winter and summer and for various λ^c/c ratios. Figures 9.13 and 9.14 correspond to the fully robust formulations RPC and RPEC, whereas Fig. 9.15 is for the SPEC formulation. These plots represent the aggregation's bid curves because they communicate how much capacity the aggregation is willing to bid in the reserve market depending on the payment it receives for each kW of the capacity.

Figure 9.13 presents the results for $\lambda^c/c > 1$, where the black curves correspond to the winter weeks ("win") and the grey curves to the summer weeks ("sum"), whereas the dashed curves are for RPC and the solid curves for RPEC. Our simulations show that the buildings are willing to offer up to 10% more reserves if energy constraints are considered. Note that the gap between RPC and RPEC is generally larger for lower λ^c/c ratios, particularly in winter. In winter, the capacity saturates at its maximum value at $\lambda^c/c = 1.1$, whereas in summer it increases monotonically as the ratio increases up to 2.
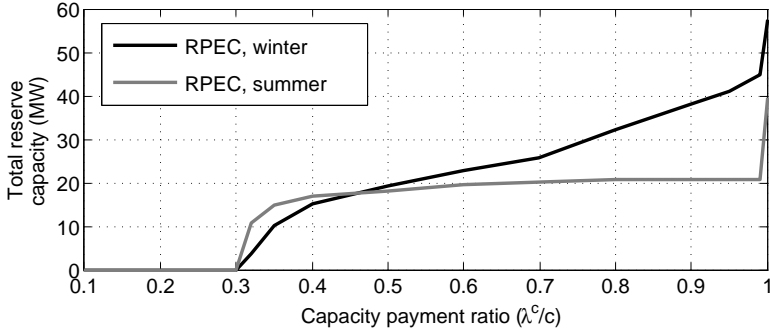
Figure 9.14: Bid curves of the building aggregation in winter and summer for the RPEC formulation and ratios $\lambda^c/c \leq 1$.

The analysis of Fig. 9.13 considered only ratios $\lambda^c/c > 1$, which is a necessary condition for reserve provision with the RPC formulation. If $\lambda^c < c$, the optimal solution is $(u_k^*, r_k^*) = (u_k^{\mathrm{opt}}, 0) \; \forall k$, where u_k^{opt} is the energy optimal scheduling and no reserves are provided. If $\lambda^c = c$, the total cost (cost from electricity consumption plus profit from reserve provision) is 0, and any solution within the feasible range of r_k will be optimal (degenerate case). If $\lambda^c > c$, the optimal solution is $(r_k^{\mathrm{max}}, r_k^{\mathrm{max}})$, where r_k^{max} is the upper limit of r, and the optimal cost will be negative, i.e., the aggregation earns profit. The limit r_k^{max} depends on the HVAC and comfort zone constraints, and therefore different solutions are obtained for different $\lambda^c/c > 1$ ratios.[8]

However, with energy-constrained SFC signals and the RPEC formulation, reserves can be provided also with ratios $\lambda^c/c < 1$. It is easier to explain this with an example. Assume that the buildings have declared a capacity $r(t)$ for day d, and that up-regulation (consumption decrease) is mainly requested during day d. If the signal is energy constrained, only a fraction of the worst-case reserve energy $\int r(t)dt$ will be requested as consumption decrease from the buildings. The part of $\int r(t)dt$ that is not requested as reserve energy will be stored as thermal energy in the buildings, and will reduce the required heating/cooling energy (and the respective costs) during day $d + 1$. For this reason, the

---

[8]In case of daily reserves, $\lambda^c/c > 1$ need not to be satisfied point-wise throughout the whole day. Instead, $\lambda^c/c < 1$ can be chosen during daytime, and $\lambda^c/c > 1$ at night when electricity prices might be lower. Reserve provision will be triggered if the capacity payment is on average higher than the electricity price.
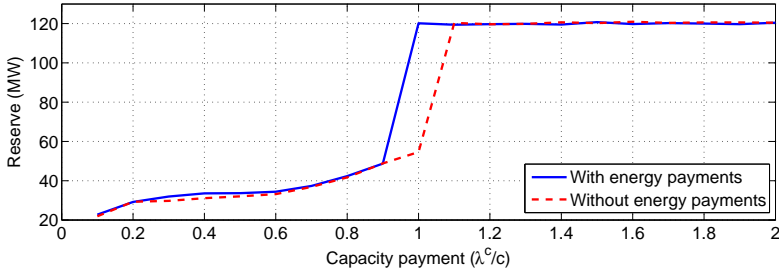
Figure 9.15: Bid curves of building aggregation in winter for the SPEC formulation, with and without reserve energy payments.

buildings are willing to provide reserves even if $\lambda^c/c < 1$. We present simulation results for the RPEC formulation in winter and summer in Fig. 9.14. The capacity is zero for $\lambda^c/c < 0.32$, it increases slowly in the ratio range $[0.32 - 0.99]$, and it suddenly jumps to higher values as λ^c/c approaches to 1.

Figure 9.13 shows that for unconstrained SFC signals a ratio $\lambda^c/c = 1.01$ already taps most of the reserve potential. Assuming an average electricity price of 146.6 CHF/MWh[9], with a ratio $\lambda^c/c = 1.01$ capacity payments around 148 CHF/MW/h are needed. This is significantly lower than the most expensive accepted bids, but approximately 4 times higher than the average capacity payment in 2013 [191]. Figure 9.14 shows that energy-constrained SFC signals can reduce the necessary capacity payments down to 32% of the retail electricity price ($\lambda^c/c = 0.32$), but with a large reduction in the reserve capacity.

Figure 9.15 presents the bid curves for the SPEC formulation with only capacity payments (dashed curve), as well as with both capacity and reserve energy payments (solid curve). The bid curve with reserve energy payments has two distinct regions: for $\lambda^c/c < 1$ the reserve capacity increases slowly with λ^c/c, at a ratio equal to 1 there is a sudden jump in the capacity, and for $\lambda^c/c > 1$ the capacity is constant. Neglecting the reserve energy payments results in only a slight underestimation of the reserve capacity for ratios $\lambda^c/c < 0.9$. However, the consideration of reserve energy payments affects the critical price ratio which is needed for the aggregation to maximize the reserve capacity. If reserve energy payments are considered the critical ratio is $\lambda^c/c = 1$, otherwise

---

[9]This is the case for consumers who consume more than 60 MWh/year in Zurich.

it increases to 1.1.

Our results show that reserves are actually costly for buildings already equipped with MPC for *energy-efficient* (optimal) control, especially if the SFC signal is not energy-constrained. However, note that the calculations are based on the prevailing case where the buildings acquire energy in the retail electricity market. In another market setting where the buildings acquired energy directly in the spot market, the buildings could offer reserves at more competitive prices, because the retail electricity prices are typically significantly higher than the wholesale spot electricity prices.

The analysis of this section provides intuition on the relationship between the amount of reserves from building aggregations and the capacity payments. In practice, estimating the capacity payment is a challenging task that needs to consider additional costs (e.g., due to device wear) but also the competition (the expected bid prices of generators and/or other load aggregations).

9.10.2 Reserve Product Characteristics

To identify appropriate reserve products from a building aggregation point of view, we perform simulations for: (i) reserve capacities that are constant over a day or over hourly intervals; and (ii) symmetric reserves at each building, symmetric reserves in aggregate (but possibly asymmetric at each building), or asymmetric reserves in aggregate. More specifically, we consider the following cases: Daily-Symmetric Reserves at each Building (DSB), Daily-Symmetric Reserves in Aggregate (DSA), Daily-Asymmetric Reserves at each Building (DA), Hourly-Symmetric Reserves at each Building (HSB), Hourly-Symmetric Reserves in Aggregate (HSA), and Hourly-Asymmetric Reserves at each Building (HA).

In addition, we investigate five different combinations of $(T, \varepsilon_+, \varepsilon_-)$ for energy-constrained SFC signals. In total, we run 30 simulations and summarize the results in Tables 9.8 and 9.9. The simulations are performed with the SPEC formulation and $\epsilon = 0.1$.

For the symmetric cases (DSB, DSA, HSB, and HSA), the capacity shown is in one direction only, so the total capacity is the double of the reported value. For the asymmetric cases (DA and HA), positive values denote up-reserves and negative values denote down-reserves. There is a benefit in allowing asymmetric reserves, because moving from DSB

Table 9.8: Maximum reserve capacity for daily reserve products with different symmetry requirements

Parameters			Daily reserves (MW)		
T (h)	ε_+ (-)	ε_- (-)	DSB	DSA	DA
2	0.336	0.467	119.2	119.4	+33.9/-216.3
4	0.337	0.237	118.9	119.1	+33.9/-215.4
6	0.273	0.252	114.9	119.1	+33.5/-214.2
8	0.384	0.343	118.2	118.5	+32.3/-213.9
12	0.197	0.229	118.4	118.0	+30.5/-214.2

Table 9.9: Maximum reserve capacity for hourly reserve products with different symmetry requirements

Parameters			Hourly reserves (MW)		
T (h)	ε_+ (-)	ε_- (-)	HSB	HSA	HA
2	0.336	0.467	121.0	123.8	+39.3/-234.1
4	0.337	0.237	120.7	122.8	+39.0/-232.9
6	0.273	0.252	120.7	123.3	+36.8/-233.9
8	0.384	0.343	120.5	122.2	+35.4/-233.2
12	0.197	0.229	120.5	122.3	+32.9/-232.3

to DSA increases the reserve capacities by approximately 1%, whereas moving from DSA to DA the increase is approximately 7%.

If asymmetric reserves are allowed, the aggregation offers much more down- than up-reserves. This happens because the buildings provide down-reserves (HP power increase) while operating close to the minimum-energy control trajectory and without increasing the buildings' baseline consumption. For any $(T, \varepsilon_+, \varepsilon_-)$ and symmetry case, adopting hourly instead of daily reserves increases the capacities in the range $2-9\%$. Finally, we observe that $T = 2$ h maximizes the capacities for all products and that the capacity generally decreases for longer T.

9.10.3 Comfort Zone and Violation Probability

The buildings are able to provide reserves because of their thermal inertia, and the fact that the occupants are not disturbed as long as the
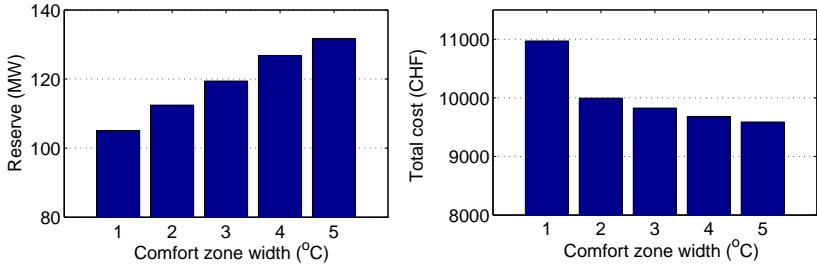
Figure 9.16: Dependence of reserve capacity (left) and total cost (right) on the comfort zone's width.

temperature is kept within the comfort zone. Therefore, the reserve capacity depends directly on the comfort zone's width. To investigate this dependence we conducted simulations with the SPEC formulation for five different comfort zones, namely $[22, 23]°C$, $[21.5, 23.5]°C$, $[21, 24]°C$, $[20.5, 24.5]°C$ and $[20, 25]°C$, and present the resulting reserve capacities and total costs in Fig. 9.16. The capacity scales linearly with the comfort zone's width, whereas the cost decays exponentially, i.e., it decreases significantly as the width of an initially narrow comfort zone increases, but it saturates for wider comfort zones.

Recall that the violation probability ϵ determines the necessary number of SFC signal scenarios to approximate the chance constraints based on (9.34). As ϵ decreases, more scenarios are needed and more sampled state (comfort) constraints are added in (9.24). This means that with lower ϵ the problem becomes more conservative, and thus higher costs should be expected. Interestingly, our simulations show that this is not always the case. Figure 9.17a shows the cost over a 2-week simulation period for $\epsilon = \{0.2, 0.15, 0.1, 0.05, 0.025\}$, which corresponds to $\{250, 333, 500, 1000, 2000\}$ scenarios. Clearly, more scenarios do not always increase the cost.

We have identified three main reasons for this: (i) with more scenarios the expected energy payments $\lambda_k^e \mathbb{E}[w_k]$ in the objective function are approximated better; (ii) since the reserve scheduling is a finite horizon optimization problem, the results of previous problems affect the aggregation's availability to provide reserves in the future; and (iii) the comfort zone $[21, 24]°C$ is wide enough to make the HVAC input constraints usually active instead of the comfort constraints.[10] To illustrate this, we

---

[10] An additional reason why the cost does not monotonically increase with the
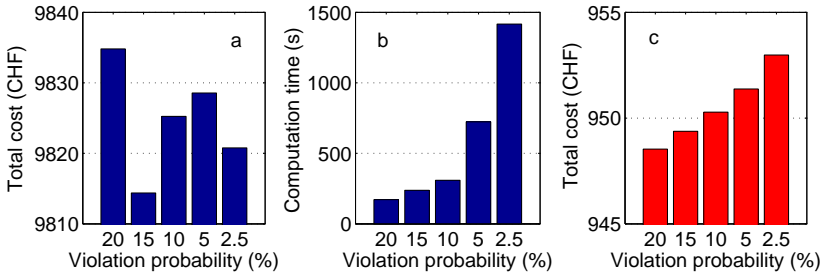
Figure 9.17: Cost (left) and computation time (middle) for a 2-week simulation period. Right: Cost for a simpler case and simulation over 1 day.

show in Fig. 9.17c the cost from a 1-day simulation with the narrower comfort zone $[22, 23]°$C and without energy payments. In this case, smaller values of ϵ indeed increase the cost, but still only marginally. As expected and shown in Fig. 9.17b, the computation time increases exponentially as ϵ decreases.

9.10.4   HP Dynamics and Constraints

In this section, we investigate the sensitivity of reserve capacity to: (i) the HP minimum loading constraints ($P_{\min}$), expressed as a percentage of the HP rated power; and (ii) the minimum up- (TU) and downtime (TD) HP constraints. Assuming $\lambda^{c} = 1.3 \cdot c$ and $\epsilon = 0.2$, we consider four different $P_{\min}$ values and compare them with the case of an unconstrained HP in terms of reserve capacity, total cost (electricity cost and reserve profit), and computation time. Table 9.10 presents results from a simulation period of two weeks. The results show that as $P_{\min}$ increases, the reserves decrease rapidly and the cost and computation

---

number of scenarios is that the scenarios are randomly drawn. Say that a set of 250 scenarios is available, which corresponds to the violation probability $\epsilon = 0.2$. When increasing the number of scenarios to 333 in order to model a violation probability of $\epsilon = 0.15$, a completely new set of scenarios is created, instead of extending the existing set of 250 scenarios with 83 new scenarios. Although the scenarios are i.i.d., it is still possible that a smaller set of scenarios contains "worse" scenarios than a larger set. Of course, if the chance-constrained optimization problem is solved many times and the results are averaged, the effect of drawing scenarios randomly will be eliminated.

Table 9.10: Effect of HP constraints

$P_{\min}$	TU	TD	Reserve	Cost	Time (s)	
(%)	(h)	(h)	(MW)	(CHF)	Mean	Max
0	0	0	120.1	5352	19	65
10	0	0	92.3	10819	37	88
20	0	0	34.9	16570	227	711
30	0	0	24.7	18641	237	1132
20	2	2	34.7	16822	235	852

Table 9.11: Effect of HP dynamics

τ_f	Symmetry	MAPE
(s)	(-)	(%)
N/A	DSA	4.12
20	DSA	3.73
40	DSA	2.91
60	DSA	2.53
40	DSB	2.12

time increase. On the other hand, TU and TD do not have a significant impact on the results.

In Table 9.11, we show the dependence of tracking MAPE on the filter's lower cutoff period τ_f. As expected, larger τ_f reduce MAPE because the reference power signal to be tracked is smoother. For fixed τ_f, we observe that symmetric reserves on the building level (DSB) result in better tracking performance than asymmetric reserves at each building but symmetric in aggregate (DSA). The reason is that with DSB the HPs usually operate at a power significantly higher than $P_{\min}$. Since the ramping limits are tighter at a low operating power, DSB results in smaller errors than DSA, where some HPs operate close to $P_{\min}$.

9.11   Implementation Aspects

In this section, we discuss a number of additional issues that need to be considered for a real-world implementation of the proposed framework. Since the reserve scheduling formulation relies on building models, the

building managers need to identify these models and share them with the aggregator. This might raise privacy concerns and discourage building managers from participating in such programs, and a potential solution could be to solve the reserve scheduling problem in a distributed way. In addition, the buildings might need to periodically send updated versions of the models to the aggregator, e.g., if the models are seasonal. Moreover, a systematic way to fairly divide the profits from reserve provision between the aggregator and the buildings is needed.

The simulation results presented in this chapter are based on four main assumptions: (a) there is no plant-model mismatch; (b) the predictions of weather and occupancy are perfect; (c) all building states can be measured; and (d) the reaction of heating/cooling devices is fast and does not cause any wear. For a real implementation of the proposed control framework, additional care must be taken for (a)-(d): accounting for modeling and weather/occupancy prediction errors, use of state estimators, and modeling of the fast dynamics of heating/cooling devices. These aspects will be discussed in Chapter 10.

An important question is if participating in reserve markets is a viable option for the buildings. There are three relevant aspects that need to be considered. First, the operational profits are highly dependent on the capacity payments (which reflect the cost of reserves in the system) and the cost of electricity. The reserves from buildings will become important as the shares of Renewable Energy Sources (RES) increase, because fewer dispatchable power plants will exist in the energy mix. With increasing RES shares, the reserves from buildings will also become more cost-competitive due to the expected drop in electricity cost and increase in reserve cost. Second, the buildings will need to modify the existing building management system to integrate the proposed controllers. The associated cost is expected to be manageable for buildings already equipped with MPC for energy-efficient control; however, the cost might be significant for buildings with conventional controllers. Third, reserve provision might cause wear or even reduce the lifetime of HPs or chillers. The buildings would need to verify if the wear is significant and, if yes, to consider the associated cost when making the investment decision.

Although the proposed framework is designed for SFC, it can be applied to other ancillary services (e.g., tertiary control, synchronized reserves, or day-ahead scheduling reserves) with reasonable modifications. Building aggregations could participate also in load management prod-

ucts that exist in some capacity markets, e.g., in PJM. One opportunity would be the limited DR program [192], where the buildings must be available to curtail load for at least 10 times during the summer period and maintain the load reduction for at least 6 hours with a 2-hour notification. Such DR events can be incorporated into the framework considering relevant load reduction scenarios and payments from the capacity market.

9.12 Conclusion

In this chapter, we developed a hierarchical control framework to allow an aggregation of commercial buildings to participate in SFC reserve markets with HVAC system control. The framework has three control layers and it is based on techniques from robust and stochastic optimization and MPC. A main attribute of the framework is that it estimates the reserve capacity that can be offered reliably to the TSO without occupant discomfort.

To illustrate the framework's performance, we considered a case study with typical Swiss office buildings and ran simulations during a winter and a summer week. The results show that an aggregation of $100 - 225$ buildings is needed to meet the minimum bid size requirement of 5 MW of the Swiss ancillary service market. The case study results indicate that asymmetric reserves are preferable for buildings in case of traditional unconstrained SFC signals. In addition, we observed that energy-constrained SFC signals can reduce the necessary capacity payments and increase the reserves significantly (for example, by up to ~10% in our case study) compared with traditional SFC signals. Furthermore, reducing the duration of reserve products from 1 day to 1 hour also increases the reserves significantly (for example, by up to ~9% in our case study).

Chapter 10

Experimental Verification of Frequency Regulation with Commercial Buildings

10.1 Introduction

Chapter 4 reviewed the few experimental demonstrations of frequency regulation with commercial buildings that have taken place so far.

To the best of our knowledge, this chapter presents the first experimental demonstration of frequency regulation from a commercial building that simultaneously addresses the following challenges: (i) a priori determination of reserve capacity, (ii) a priori declaration of base-point power (i.e., short-term operating power around which we provide frequency regulation), (iii) guarantees on satisfaction of occupant comfort, and (iv) accurate tracking of the regulation signal. This is achieved by adopting the hierarchical control framework of Chapter 9, which is extended and adjusted in order to fit the needs of the demonstration project.

Table 10.1: Nomenclature of Chapter 10: symbols

Symbol	Unit	Description
$a_{11} - a_{22}$	-	Entries of the state matrix of building model
b	-	Entry of the input matrix of building model
C	kWh/$^\circ$C	Thermal capacitance
c	€/kWh	Electricity cost
c_p	J/kg$^\circ$C	Specific heat capacity of water
$d_{11} - d_{13}$	-	Entries of the disturbance matrix of building model
E	Wh	Energy consumption
e	W	Control error metric
$e_\mathrm{old}, e_\mathrm{new}$	W	Control errors for the PI controller
f	-	Function mapping air flow rate to fan power
g	-	Function mapping fan speed to fan power
h	-	Function mapping fan speed to air flow rate
G	W/m^2	Solar radiation
I_g	W	Internal heat gain
K	-	Kalman gain
K_p, K_i	-	Proportional and integral gains of the PI controller
$\dot{m}, u$	cmf	Air flow rate (cubic feet per minute)
$\dot{m}_\mathrm{cw}, u$	gpm	Chilled water flow rate (gallons per minute)
N_f	%	Fan speed (% of rated)
N_1, N_2	hours	Prediction horizons of levels 1 and 2
P	W	Fan power
$P_\mathrm{e}^-, P_\mathrm{e}$	-	A priori and a posteriori covariance matrix
Q_c	W	Cooling power from the HVAC system
R_ra	$^\circ$C/kW	Thermal resistance (room to ambient)
R_rm	$^\circ$C/kW	Thermal resistance (room to lumped mass)
R	W	Electric reserve
r	W	Thermal reserve
S_c	-	PJM correlation score
S_d	-	PJM delay score
S_p	-	PJM precision score
S_tot	-	PJM total score
T	$^\circ$C	Temperature
v	$^\circ$C or W/m^2	Vector of weather disturbances
w	-	Normalized SFC signal
w_lim	-	Normalized energy content of SFC signal
x	$^\circ$C	Vector of building states

The demonstration is unique in that it balances energy consumption minimization and profit maximization from selling the reserve capacity in a way that the net profit is maximum. In other words, the control scheme we designed chooses to favor either energy efficiency or regulation capacity depending on weather and occupancy conditions, building states, as well as electricity and reserve prices. Furthermore, the reserve capacity is determined in a day-ahead fashion using techniques from robust optimization to ensure that it can be provided without compromising the occupants' comfort.

Table 10.2: Nomenclature of Chapter 10: Greek letters

Symbol	Unit	Description
$\alpha_0 - \alpha_3$	-	Parameters of the air flow to fan power model
$\beta_0 - \beta_3$	-	Parameters of the fan speed to fan power model
γ_0, γ_1	-	Parameters of the fan speed to air flow model
γ	-	Solar radiation absorption factor
Δt	min	Discretization time step
Δu	W/m^2	Perturbation around u to provide SFC reserve
ε	W	Threshold for switching between controllers in level 3
η	-	Efficiency
λ	€/kW/h	Reserve capacity payment

Table 10.3: Nomenclature of Chapter 10: subscripts and superscripts

Symbol	Description
a	Ambient temperature
c	Control error
ch,r	Return chilled water
ch,s	Supply chilled water
d	Down-reserve
k	Time index
m	Lumped thermal mass temperature
mae	Mean absolute control error
me	Mean control error
min, max	Minimum and maximum values of a variable
r	Room temperature (T_r)
r	Reserve percentage error (e_r)
rmse	Root mean square error
r,mape	Mean absolute percentage reserve error
s	Supply air temperature
t	Tracking percentage error
t,mape	Mean absolute percentage tracking error
u	Up-reserve
$*, \star$	Optimal value of a variable
$(\bar{\cdot}), (\underline{\cdot})$	Upper and lower bound on a variable
$(\hat{\cdot})$	Estimate of a variable

The building climate control is formulated as a robust Model Predictive Control (MPC) problem, which attempts to minimize the energy consumption while leaving enough slack for reserve provision. The MPC uses a simple 2nd order building model that can be obtained easily from historical data, and an extended Kalman filter. The regulation signal is tracked in real-time using a novel switched controller comprised of a feedforward and a feedback element. The control variable is the fan speed, which is shown to achieve a much better tracking performance compared with previous approaches that relied on control of the static duct pressure setpoint, air flow rate setpoint, Supply Air Tem-

perature (SAT) setpoint, or zone temperature setpoints.

Another unique characteristic of the demonstration is that it is performed on a test building comprised of two identical zones. We use one zone for the frequency regulation experiment, whereas the other zone is used as a real-time benchmark to evaluate the effect of our control actions.

The experiments were performed over a period of two weeks and the results are very encouraging. A significant amount of regulation reserves is extracted from the test building, the tracking performance is excellent, and the effect on occupant comfort and on the Heating, Ventilation and Air Conditioning (HVAC) system is minimal. Most of the tests were performed with historical RegD signals from the Pennsylvania, Jersey, and Maryland Power Market (PJM). In addition, we established a connection with PJM, received, and tracked the actual RegD signal in real time for a small part of the experiment.

The remainder of this chapter, which is based on [193, 194], is organized as follows. Section 10.2 introduces the goals of the experiment and the test facility, whereas Section 10.3 presents the control and communication architecture. Section 10.4 presents the developed building models and Section 10.5 the hierarchical controller. Extensive experimental results are analyzed in Section 10.6, some important observations and suggestions are summarized in Section 10.7, whereas Section 10.8 concludes. The nomenclature of this chapter is given in Tables 10.1 - 10.3.

10.2   Experiment Goals and Facility

10.2.1   Goals of this Experiment

The main goal of this experiment is to demonstrate that frequency regulation can be provided by commercial buildings at a high quality. In particular, we attempt to show the following:

- A building can determine the reserve capacity that it can offer and bid it in a day-ahead Ancillary Service (AS) market.

- Accurate tracking of the frequency regulation signal is possible with fan control.

- With proper control design, frequency regulation has no or minimal adverse effect on occupant comfort.

Figure 10.1: Front (left) and rear (right) view of FLEXLAB.

- A hierarchical controller can operate the building favoring either frequency regulation or energy efficiency depending on external parameters and conditions.

We believe that the proposed hierarchical controller is appropriate for frequency regulation in commercial buildings, and the purpose of this chapter is to evaluate its performance in a small but highly controllable test building. This is a necessary first step before field testing in larger commercial buildings and widespread implementation.

10.2.2   Description of FLEXLAB

The experiment was performed at the Facility for Low Energy eXperiments (FLEXLAB), a new facility for energy efficiency research in buildings located at the Lawrence Berkeley National Laboratory (LBNL). The facility (shown in Fig. 10.1) is comprised of 4 buildings that are called "bays", and each of the bays consists of 2 building zones that are called "cells". Each pair of cells is designed to be thermally identical, constructed with the same materials and dimensions.

Although the two cells are semi-attached, i.e., they are built side-by-side, the walls at the side where the cells are physically connected are adiabatic. As a result, there is no heat exchange between the two cells and therefore they can be considered as independent. This is a unique feature of FLEXLAB that allows us to perform frequency regulation experiments in one of the two identical cells, while using the other one as a benchmark to evaluate the effect of our control actions in real time and under the same external conditions.

One of the bays stands on a rotating platform that can be used to change the buildings orientation, whereas the other three bays have a
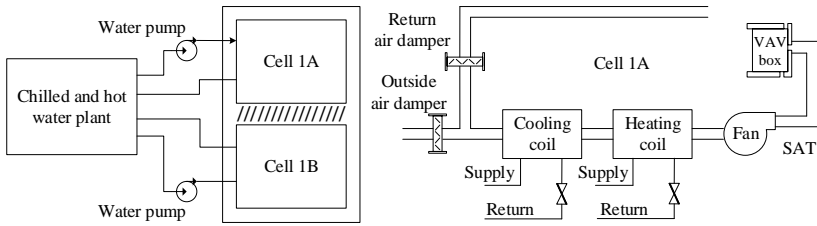
Figure 10.2: The HVAC system of building cells 1A and 1B of FLEXLAB.

fixed south orientation. For our experiments we used the first of the bays with south orientation, which has a total floor area of 120 m^2 (60 m^2 per building cell). In the rest of this chapter, we will call cell "1A" the cell where the reserve experiments were performed, and cell "1B" the cell that was used as a benchmark. Although small, the building is a good representation of commercial buildings with Variable Air Volume (VAV) systems constructed in the 1980's.

Three major cascade control loops are present in a VAV HVAC system: chilled water temperature control, SAT control, and zone (room) temperature control. A chiller plant cools down water that is then piped to the building's Air Handling Unit (AHU). The chilled water decreases the temperature of a mixture of return and outside air in the AHU using a heat exchanger, and the flow of the chilled water is controlled to maintain a constant SAT. The cooled air is circulated to the building zones through the duct system using a fan. The temperature of each zone is maintained close to the desired setpoint by controlling the damper position of the VAV box.

Typically, an AHU provides several building zones with cooled air, whereas reheating is done locally at the VAV box of each zone. However as shown in Fig. 10.2, the cells of FLEXLAB are served by dedicated AHUs that contain also a heating coil, apart from the fan and cooling coil. Another particularity of FLEXLAB is that the inlet dampers in the zones are controlled manually, hence, the zone temperature is regulated by fan control.

10.2.3   Preparation of the Experiment

Since FLEXLAB has no occupants, we had to emulate the internal heat gains from occupants and equipment. There were two ways to do so:

(i) using an existing floor heating system that actuates in the building slab; and (ii) using electric heaters as plug loads. The second option was used because it additionally allowed us to emulate hotter days by plugging in more heaters to create a higher temperature differential with the ambient. The total internal heat gain in both cells was kept lower than the chiller cooling capacity.

Before the start of the experiment, the manually controlled inlet dampers in the rooms where fixed to fully open positions, and the SAT was fixed to 17°C. In addition, the return air damper was fixed to a 100% opening and the outside air damper to a 0% opening, i.e., the return air was fully recirculated. The speeds of the primary and secondary chilled water pumps were fixed to 75% and 100% of their rated speeds, respectively. Moreover, both the floor heating system and the reheating coil at the AHU were deactivated.

To verify that the the two cells are indeed sufficiently similar, we applied exactly the same control inputs (fan speeds) in both cells and recorded the room temperatures. The comparison showed that the two cells are sufficiently similar, but it also revealed a systematic discrepancy between the two cells. Analyzing the measurements we found that this discrepancy was due to calibration differences between the two AHUs, i.e., the same fan speed command resulted in different supply air flow rates in the two cells. We used these data to fit a correction model that slightly modifies the fan speed in cell 1B to achieve the same air flow rate as in cell 1A.

10.3 Control and Communication Architecture

10.3.1 Hierarchical Control

FLEXLAB is controlled by a Central Working Station (CWS) based on a predefined control sequence programmed in LabVIEW. From a TestStand National Instruments user interface, the operator can modify the setpoints of various control loops and observe the system's behavior in the chilled water loop, hot water loop, AHU, and conditioned room. A screenshot of the user interface for the AHU is shown in Fig. 10.3. This interface is very useful for infrequent update of control setpoints
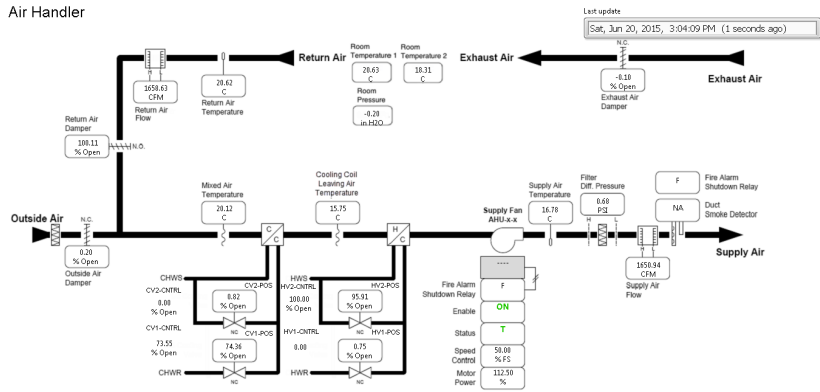
Figure 10.3: Screenshot of the TestStand user interface for the AHU.

or monitoring purposes, but it cannot be used for frequency regulation because it would require constant human intervention.

One possibility to automate the frequency regulation process would be to integrate our controllers within the existing LabVIEW environment. However, the option that was finally adopted was to use an existing scripting environment to send the control setpoints directly to the CWS. This allowed us to develop the fan controller for frequency regulation externally, and therefore to minimize potential conflicts with the Lab-VIEW code. The only modifications we did in the LabVIEW code were settings that did not change during the experiment (e.g., deactivation of the floor heating system). In addition, the SAT was fixed via the Test-Stand interface because there existed no relevant control point within the scripting environment. Apart from the fan speed, all other control variables (e.g., setpoints in the chilled water loop) were calculated by the default controller in the CWS. Therefore, the designed controllers work complementarily to the existing ones and do not substitute them.

The hierarchical controller designed for the experiment consists of three levels: Level 1: reserve scheduler, Level 2: room climate controller, and Level 3 frequency regulation controller. In this section, we provide a high-level introduction to the three control levels and explain the underlying communication architecture. The control sequences are shown in Fig. 10.4 and detailed mathematical formulations for the controllers will be provided in Section 10.5.
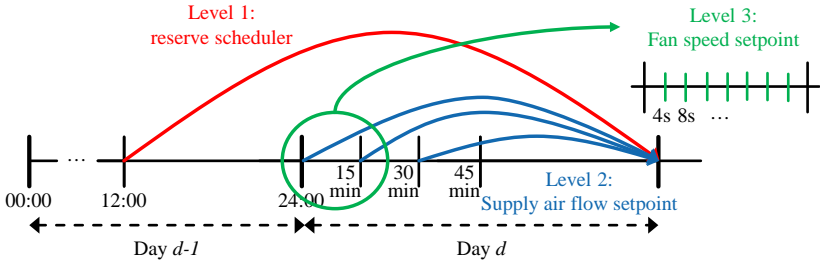
Figure 10.4: Control sequences of the three levels of the hierarchical control for frequency regulation.

Level 1: Reserve Scheduler

The goal of the reserve scheduler is to determine the reserve capacity that the building can offer to the Transmission System Operator (TSO). Depending on the reserve market requirements, the reserve scheduling might be performed on a daily or on an hourly basis, but certainly ahead of real-time operation. In the experiment we assumed a day-ahead reserve scheduling occurring at 12.00 of each day to determine the reserve capacity for the next day. In order to reliably provide the reserve capacity for any possible reserve request, the reserve scheduling is formulated as a multi-period robust optimization problem and is solved in Matlab.

Level 2: Room Climate Controller

This controller calculates the supply air flow rate setpoints that minimize the energy consumption, while ensuring occupant comfort under reserve provision. It is important to note that this is a zonal controller because it regulates the room temperature in the air loop, but not the control setpoints in the chilled water loop, which are determined by the default FLEXLAB controller. This controller is implemented as a robust MPC in Matlab and runs every 15 minutes.

Level 3: Frequency Regulation Controller

The goal of this controller is to track the frequency regulation signal nearly in real time by modifying the fan power via fan speed control with a Variable Frequency Drive (VFD). For this purpose, we design a switched controller comprised of a feedforward model-based controller
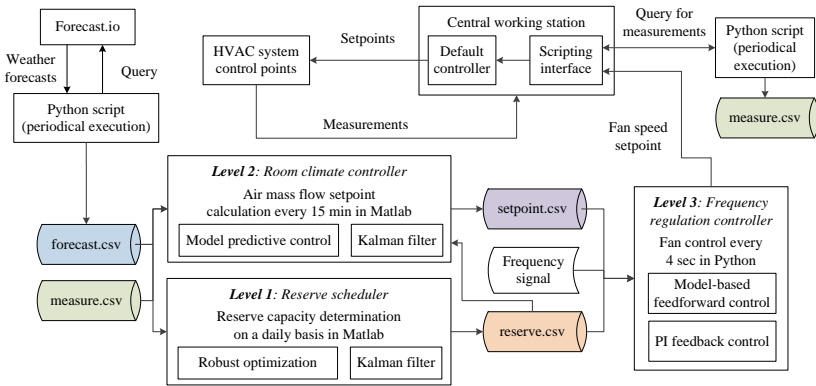
Figure 10.5: The developed control and communication architecture for the frequency regulation experiment in FLEXLAB.

and a feedback Proportional-Integral (PI) controller with gain scheduling. The controller is implemented in Python and is executed every 4 seconds.

10.3.2   Communication Architecture

A simple communication architecture was designed to coordinate the operation of the three levels of the hierarchical controller. At the lower level, an existing scripting environment was used to send control setpoints directly to the CWS by executing ssh-based proxy commands from a Unix shell. This allowed us to override the default FLEXLAB controller at the cost of introducing some additional delay in the control loop.

The scripting environment was also used to query the measurement points via the CWS. To obtain synchronized measurements and minimize the measurement delay, we created three "users". The "control user" includes the measurement points related to the fan and the electric power measurement points, which are queried every 1 second. The "feedback user" includes all measurement points that are needed as feedback within the MPC, namely the supply air flow rate, SAT, room temperature, plug loads and ambient conditions, and it queries these points every 1 minute. And finally, the "monitor user" includes all other

measurement points that are useful for post-processing and analysis of the experimental results, and it queries these points every 5 minutes.

The fan speed control in level 3 and the formation of the appropriate CWS proxy commands are performed in Python, whereas the optimization problem of level 1 and the MPC of level 2 are implemented in Matlab. We created a simple communication interface between Python and Matlab based on csv files.

The reserve scheduler stores the computed reserve capacity in the reserve.csv file. A Python script periodically queries the three measurement users and stores the measurements in the measure.csv file. Another Python script periodically queries the publicly available weather forecast database of forecast.io and stores the necessary forecasts in the forecast.csv file. The input to the MPC is obtained by reading the building measurement points of the feedback user from measure.csv and the weather forecasts from forecast.csv. The optimal air flow rate setpoint calculated by Matlab every 15 minutes is stored in the setpoint.csv file. The setpoint.csv and reserve.csv files are accessed by Python every 4 seconds to determine the fan speed reference, of course based also on the received frequency regulation signal. The communication architecture is graphically shown in Fig. 10.5.

10.4 Modeling

10.4.1 Building Thermal Model

We model the building thermal dynamics using the 2-state Resistance-Capacitance (RC) network of Fig. 10.6. If the heating coil of the AHU is deactivated, the cooling power of the HVAC system is given by $Q_c = \dot{m}c_p(T_s - T_r)$, where T_r is the room temperature, $\dot{m}$ is the mass air flow rate in the room, c_p is the specific heat capacity of air, and T_s is the SAT. The differential equations of the RC network are

$$C_r\frac{dT_r}{dt} = \dot{m}c_p(T_s - T_r) + \frac{T_a - T_r}{R_{ra}} + \frac{T_m - T_r}{R_{rm}} + \gamma G + I_g \quad (10.1)$$

$$C_m\frac{dT_m}{dt} = \frac{T_r - T_m}{R_{rm}} \ , \quad (10.2)$$

where T_m is the temperature of the lumped thermal mass; C_r is the room's thermal capacitance; C_m is thermal capacitance of the lumped
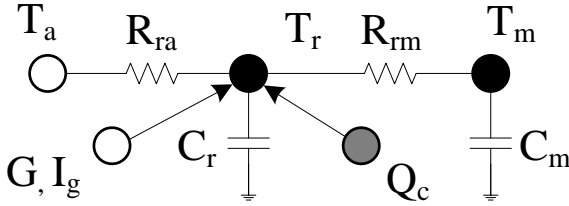
Figure 10.6: The equivalent RC network of the considered building thermal model. The black nodes are the building states, the grey node is the control input, and the white nodes are the disturbances.

mass; T_a is the ambient temperature; R_{ra} is the thermal resistance between the room and the ambient; R_{rm} is the thermal resistance between the room and the lumped mass; G is the solar radiation; γ is the solar radiation absorption factor; and I_g represents the internal heat gains. The model is bilinear between the control input $\dot{m}$ and the state T_r. In general, there exists also a bilinearity between $\dot{m}$ and T_s; however, T_s is controlled to a fixed value in our experiment and thus it is treated as a constant in the model.

By reordering (10.1) and (10.2), the continuous-time model can be written in the state-space form

$$
\begin{bmatrix} \dot{T}_r \\ \dot{T}_m \end{bmatrix} = \begin{bmatrix} -\left(\frac{1}{C_r R_{ra}} + \frac{1}{C_r R_{rm}} \right) & \frac{1}{C_r R_{rm}} \\ \frac{1}{C_m R_{rm}} & -\frac{1}{C_m R_{rm}} \end{bmatrix} \begin{bmatrix} T_r \\ T_m \end{bmatrix} +
$$

$$
\begin{bmatrix} \frac{c_p T_s}{C_r} \\ 0 \end{bmatrix} \dot{m} + \begin{bmatrix} -\frac{c_p}{C_r} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} T_r \\ T_m \end{bmatrix} \dot{m} + \begin{bmatrix} \frac{1}{C_r R_{ra}} & \frac{\gamma}{C_r} & \frac{1}{C_r} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} T_a \\ G \\ I_g \end{bmatrix} . \quad (10.3)
$$

Applying a first-order Euler discretization, the structure of the continuous-time matrices is maintained [195], and therefore the discrete-time model can be written in the form

$$
\begin{bmatrix} T_{r,k+1} \\ T_{m,k+1} \end{bmatrix} = \overbrace{\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}}^{A} \begin{bmatrix} T_{r,k} \\ T_{m,k} \end{bmatrix} + \overbrace{\begin{bmatrix} b T_{s,k} \\ 0 \end{bmatrix}}^{B_{u,k}} \dot{m}_k +
$$

$$
\overbrace{\begin{bmatrix} -b & 0 \\ 0 & 0 \end{bmatrix}}^{B_{xu}} \begin{bmatrix} T_{r,k} \\ T_{m,k} \end{bmatrix} \dot{m}_k + \overbrace{\begin{bmatrix} d_{11} & d_{12} & d_{13} \\ 0 & 0 & 0 \end{bmatrix}}^{B_v} \begin{bmatrix} T_{a,k} \\ G_k \\ I_{g,k} \end{bmatrix} , \quad (10.4)
$$

where the state, input and disturbance vectors are defined as

$$x_k = [T_{r,k} \quad T_{m,k}]^\top \tag{10.5}$$

$$u_k = \dot{m}_k \tag{10.6}$$

$$v_k = [T_{a,k} \quad G_k \quad I_{g,k}]^\top . \tag{10.7}$$

The discrete-time model parameters $a_{11}, a_{12}, a_{21}, a_{22}, b, d_{11}, d_{12}, d_{13}$ can be identified by regression on available data of $T_{r,k}, \dot{m}_k, T_{a,k}, G_k, I_{g,k}$.[1] Since the state $T_{m,k}$ is not directly measured, the resulting regression is a non-linear optimization problem that involves multiplications of the optimization variables (the model parameters)

$$\min_{a_{11}-a_{22}, b, d_{11}-d_{13}, \hat{T}_{r,k}, \hat{T}_{m,k}} \sum_k \left(\hat{T}_{r,k} - T_{r,k} \right)^2 \tag{10.8a}$$

$$\text{s.t.} \begin{bmatrix} \hat{T}_{r,k+1} \\ \hat{T}_{m,k+1} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} T_{r,k} \\ \hat{T}_{m,k} \end{bmatrix} + \begin{bmatrix} bT_{s,k} \\ 0 \end{bmatrix} \dot{m}_k +$$

$$\begin{bmatrix} -b & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} T_{r,k} \\ \hat{T}_{m,k} \end{bmatrix} \dot{m}_k + \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} T_{a,k} \\ G_k \\ I_{g,k} \end{bmatrix} , \ \forall k \tag{10.8b}$$

$$a_{12} \geq 0, \ a_{21} \geq 0, \ b \geq 0, \ d_{11} \geq 0, \ d_{12} \geq 0, \ d_{13} \geq 0 \tag{10.8c}$$

$$\hat{T}_{m,k}^{\min} \leq \hat{T}_{m,k} \leq \hat{T}_{m,k}^{\max}, \ \forall k \tag{10.8d}$$

$$|\text{eig}(A)| \leq 1, \ |\text{eig}(A + B_{xu}\dot{m}_k)| \leq 1, \ \forall k . \tag{10.8e}$$

Let us denote by A^c, B_u^c, and B_v^c the continuous-time state matrix, input matrix and disturbance matrix, respectively. Due to the first-order discretization, the discrete-time matrices are obtained from the continuous-time matrices as $A = I + \Delta t \cdot A^c$, $B_u = \Delta t \cdot B_u^c$ and $B_v = \Delta t \cdot B_v^c$, where Δt is the discretization step. For this reason, the positive elements of A^c, B_u^c, and B_u^c in (10.3) remain positive in the discrete-time matrices of (10.4), which is represented in the regression by constraints (10.8c). Constraints (10.8d) are lower and upper bounds on the estimated unmeasured state $T_{m,k}$ to avoid unreasonable values. In this experiment, the bounds $\hat{T}_{m,k}^{\min} = 0.01 \cdot T_{r,k}$ and $\hat{T}_{m,k}^{\max} = 2.5 \cdot T_{r,k}$ were used.

---

[1]The entries d_{21}, d_{22}, d_{23} of the second row of the disturbance matrix B_v are fixed to zero, which gives more reasonable results. Therefore, the external disturbances act directly only on the room temperature, which is a practice followed also in previous works, e.g., in [196].

Table 10.4: Comparison of building model variants

Model	RMSE (°C)
1-state model with 1-step ahead prediction	0.92
1-state model with 1-day ahead prediction	0.67
2-state model with 1-step ahead prediction	0.89
2-state model with 1-day ahead prediction	0.42

Constraints (10.8e) impose stability of the identified model. More specifically, $|\text{eig}(A)| \leq 1$ is a standard stability requirement for linear systems, whereas $|\text{eig}(A + B_{xu}\dot{m}_k)| \leq 1$ is needed because the system is bilinear. Note that for a sufficiently large data set (10.8e) are not expected to be active constraints. However, we include these constraints to avoid over-fitting when a building model is identified using a relatively small data set or a data set without sufficient excitation in terms of control inputs and states.

The optimization problem (10.8) identifies a model based on a "1-step ahead prediction", meaning that the estimate $\hat{T}_{r,k+1}$ at each time step depends directly on the measurement $T_{r,k}$ at the previous time step. This identification problem formulation is a standard practice for MPC applications, where what matters the most is the quality of the one-step ahead prediction of the model. In contrast, in our problem the building model is also used within the day-ahead reserve scheduler, and thus high-quality day-ahead predictions are important.

To address this requirement, a more sophisticated "1-day ahead prediction" model can be obtained by substituting $T_{r,k}$ with the optimization variable $\hat{T}_{r,k}$ in (10.8b), which results in a more complex non-linear regression problem. Note that a 1-state model can be used, if the lumped thermal mass of the room is neglected. In that case the regression problem is simplified significantly, especially if the "1-step ahead prediction" approach is used, which results in a least squares problem with linear constraints.

Building Model Identification Results

We fitted two sets of building model parameters. The first one used data from $17 - 21$ June, $22 - 25$ June, and $4 - 5$ July 2015. During these periods, the system was sufficiently excited by operating it with
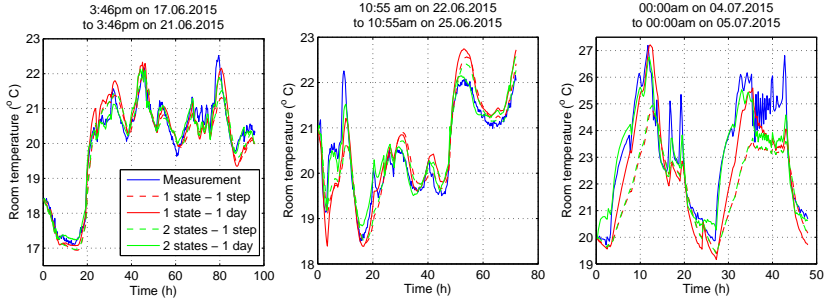
Figure 10.7: The results of the thermal model identification. The 2-state model with 1-day ahead prediction matches closer the experimental data.

Table 10.5: Parameters of the 2-state model with 1-day ahead prediction identified with data from June-July 2015

Parameter	a_{11}	a_{12}	a_{21}	a_{22}
Value	0.8665	0.0918	0.0374	0.9703
Parameter	b	d_{11}	d_{12}	d_{13}
Value	0.2996	0.0230	$2.0156 \cdot 10^{-4}$	$1.4242 \cdot 10^{-4}$

different combinations of air flow rate, SAT, and heater power. Four different model variants were compared: (i) 1-state model with 1-step ahead prediction, (ii) 1-state model with 1-day ahead prediction, (iii) 2-state model with 1-step ahead prediction, and (iv) 2-state model with 1-day ahead prediction. The identification results are shown in Fig. 10.7 and the Root Mean Squared Errors (RMSEs) are given in Table 10.4.

As expected, increasing the number of states from one to two reduces the RMSE. The error reduces also by increasing the prediction horizon from one step to one day. Based on Table 10.4, increasing the prediction horizon reduces the error more effectively than increasing the number of states. We use the 2-state model with 1-day ahead prediction in the frequency regulation experiments because it achieves the lowest RMSE. The identified model parameters are shown in Table 10.5.[2]

The second set of model parameters was fitted using data from $12 - 18$

---

[2]Different models were identified for cells 1A and 1B and they were similar to each other because the two cells are nearly identical. Only the model of cell 1A is shown here.

Table 10.6: Parameters of the 2-state model with 1-day ahead prediction identified with data from November 2015

Parameter	a_{11}	a_{12}	a_{21}	a_{22}
Value	0.6344	0.2661	0.1021	0.9170
Parameter	b	d_{11}	d_{12}	d_{13}
Value	0.4716	0.0405	0.0028	$3.3686 \cdot 10^{-4}$

November 2015, and the identified parameters are given in Table 10.6. The reason why we used two different parameter sets is to investigate the importance of periodic calibration of the building model.

10.4.2   Fan Model

A steady-state fan model is required both in the MPC controller to determine the optimal mass air flow setpoints and in the fan controller to track the frequency regulation signal. In the MPC controller the mapping between air flow rate and fan electric power must be considered, whereas in the fan controller the desired electric power setpoint must be converted into a fan speed reference, which is the control variable. According to the fan laws, the mass air flow rate u is proportional to the fan speed N_f, and the fan power P increases with the cube of the fan speed. Therefore, a steady-state model can be obtained by fitting the parameters of

$$P = f(u) = \alpha_3 u^3 + \alpha_2 u^2 + \alpha_1 u + \alpha_0 \qquad (10.9)$$

$$P = g(N_f) = \beta_3 N_f^3 + \beta_2 N_f^2 + \beta_1 N_f + \beta_0 \qquad (10.10)$$

$$u = h(N_f) = \gamma_1 N_f + \gamma_0 \ , \qquad (10.11)$$

using measurements of N_f, u and P.

For this purpose, we vary the fan speed setpoint from 10% to 90% with a step of 5%, and record the air flow rate and electric power.[3] Each fan speed setpoint is kept for 6 minutes, but the first 20 seconds of the data after each step change are discarded to allow enough time for the fan to reach the new steady state and account for actuation and measurement delays. The identified parameters are given in Table 10.7.

---

[3]The minimum and maximum fan speed setpoints of 10% and 90% were suggested by the FLEXLAB building manager. The nominal fan power is 2500 W.

Table 10.7: Fan model parameters

Parameter	α_3	α_2	α_1	α_0
Value	$2.5882 \cdot 10^3$	$-1.4580 \cdot 10^3$	630.8961	28.7249
Parameter	β_3	β_2	β_1	β_0
Value	0.0032	-0.0151	1.4521	55.7634
Parameter	-	-	γ_1	γ_0
Value	-	-	0.0133	0.0606
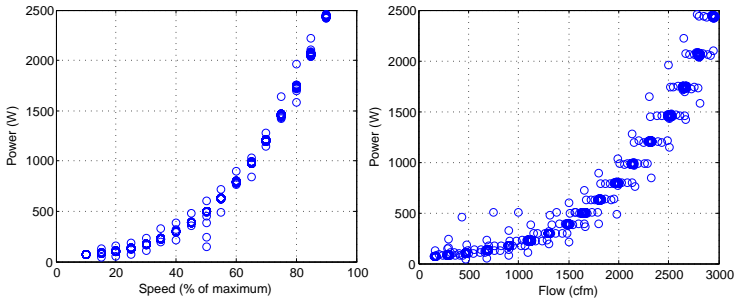
Figure 10.8: The raw measurements of fan speed, air flow and power from the fan identification experiment. The data points are scattered due to the measurement delays (average value of 2.89 seconds) and the response time of the fan to step changes.

Note that to avoid very small model parameters, the air flow rate data are normalized by the rated flow rate of 2700 cfm before identifying the model, and so the flow rate u in (10.9) is a normalized value. The fan speed N_{f} in (10.10) is expressed as percentage of the maximum speed, whereas the fan power is in W for both (10.9) and (10.10). The raw measurements are shown in Fig. 10.8, whereas the fit of the identified models on the filtered fan measurements is shown in Fig. 10.9. The fitting performance is very high and the RMSE is only 5 W for the speed-to-power model and 21 W for the flow-to-power model.
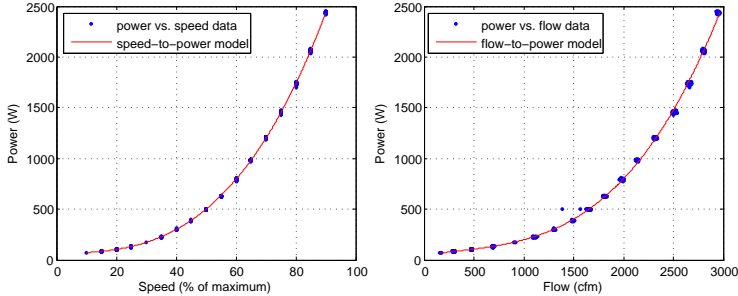
Figure 10.9: The fitting performance of the identified steady-state fan models on the filtered measurements.

10.5   Hierarchical Control Design

10.5.1   Level 1: Reserve Scheduler

The goal of level 1 is to compute the reserve capacities that the building can offer to the power system during the following day. The day-ahead reserve scheduling is formulated as a multi-period robust optimization problem. Before presenting the mathematical formulation, we discuss some important assumptions on building operation.

Building Operation Assumptions

We perform the experiment while FLEXLAB operates in cooling mode, i.e., both the floor heating and the reheating coil at the AHU are deactivated. Furthermore, we make the following assumption on the SAT.

Assumption 1. *The SAT is controlled to a setpoint $T_{s,k}$ (via cooling valve control) that is always less or equal to the lower bound of the building's comfort zone $x_{min,k}$:*

$$T_{\mathrm{s},k} \leq x_{\min,k} \leq T_{\mathrm{r},k}, \ \forall k \ . \tag{10.12}$$

Under this assumption, pushing more air into the room will always decrease the room temperature.

The duct static pressure is not modeled and no pressure constraints are used in the problem formulation. The duct system of FLEXLAB is short because a dedicated AHU supplies each zone, and thus drastic changes in pressure are very unlikely. However, additional pressure constraints might be necessary in large buildings with longer duct systems.

Robust Reserve Scheduling Formulation

For generality, we differentiate between the electric up- and down-reserve capacity of the building at time step k, which we denote by $R_{u,k}$ and $R_{d,k}$, respectively. The up-reserve capacity $R_{u,k}$ is the maximum decrease in electric power that can be requested, whereas $R_{d,k}$ is the maximum increase in electric power.[4]

It is convenient to define also the thermal up- and down-reserve capacities, $r_{u,k}$ and $r_{d,k}$, as the maximum changes in the mass air flow rate due to reserve provision. In cooling operation, a request for regulation up results in a reduction in air mass flow rate, such that $R_{u,k}$ is related to $r_{d,k}$. On the other hand, regulation down results in an increase in air mass flow rate ($R_{d,k}$ is related to $r_{u,k}$).

The relationship between the electric and the thermal reserve capacities is obtained from the air flow-to-power fan model (10.9). Given an operating point of air flow u_k, $R_{u,k}$ and $R_{d,k}$ are expressed as

$$R_{u,k} = f(u_k) - f(u_k - r_{d,k}) \tag{10.13}$$
$$R_{d,k} = f(u_k + r_{u,k}) - f(u_k) . \tag{10.14}$$

This nonlinear relationship is very important for the reserve scheduling formulation and it is graphically shown in Fig 10.10.

The objective of the reserve scheduler is to minimize the total cost defined as the sum of energy consumption cost and reserve profit

$$c_k P_k - \lambda_k \left(R_{d,k} + R_{u,k} \right) , \tag{10.15}$$

where $P_k = f(u_k)$ is the electric power consumption at the scheduled air flow rate u_k. Assuming the same payment for up- and down-reserves,

---

[4]In the context of frequency regulation, up-reserve means increase of generation or decrease of consumption, whereas down-reserve means decrease of generation or increase of consumption.
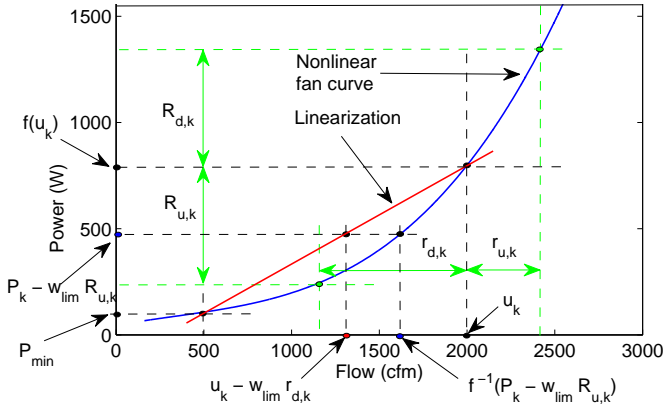
Figure 10.10: The nonlinear fan curve and the simple linearization for opti-
mization purposes. The plot shows the thermal reserves ($r_{u,k}$
and $r_{d,k}$) and the electric reserves ($R_{u,k}$ and $R_{d,k}$) around
an air flow operating point (u_k). When providing reserve the
new actual operating point $f^{-1}(P_k - w_{lim}R_{u,k})$ is higher than
the one assumed by the linearization ($u_k - w_{lim}r_{d,k}$).

which is denoted by λ_k, the reserve profit is given by

$$\lambda_k \left(R_{d,k} + R_{u,k} \right) = \lambda_k \left[\left[f(u_k + r_{u,k}) - f(u_k) \right] + \left[f(u_k) - f(u_k - r_{d,k}) \right] \right]$$
$$= \lambda_k \left[f(u_k + r_{u,k}) - f(u_k - r_{d,k}) \right] . \qquad (10.16)$$

Typically, the TSO requests the reserve energy as a fraction of the
reserve capacity using a normalized frequency regulation signal [72],
which we denote by $w_k \in [-1, 1]$. Therefore, the reserve request at time
step k is

$$R_k = \begin{cases} w_k R_{u,k}, & \text{if } w_k < 0 \\ w_k R_{d,k}, & \text{if } w_k \geq 0 . \end{cases} \qquad (10.17)$$

The electric reserve request can be translated to a perturbation around
u_k using the fan curve

$$\Delta u_k = f^{-1} \left(P_k + R_k \right) - u_k . \qquad (10.18)$$

With the above notation, the multi-period robust reserve scheduling

problem can be written as

$$\min_{u_k, r_{\mathrm{u},k}, r_{\mathrm{d},k}} \sum_{k=0}^{N_1-1} c_k f(u_k) - \lambda_k \big[f(u_k + r_{\mathrm{u},k}) - f(u_k - r_{\mathrm{d},k}) \big] \quad (10.19\mathrm{a})$$

$$\text{s.t. } x_{k+1} = A x_k + B_u T_{\mathrm{s},k} \cdot (u_k + \Delta u_k) +$$
$$B_{xu} x_k \cdot (u_k + \Delta u_k) + B_v v_k, \ \forall k \quad (10.19\mathrm{b})$$
$$u_{\min,1} \le u_k + \Delta u_k \le u_{\max,1}, \ \forall w_k \in [-1,1], \ \forall k \quad (10.19\mathrm{c})$$
$$x_{\min,k} \le x_k \le x_{\max,k}, \ \forall w_k \in [-w_{\mathrm{lim}}, w_{\mathrm{lim}}], \ \forall k \ . \quad (10.19\mathrm{d})$$

Equation (10.19b) represents the building dynamics, whereas (10.19c) and (10.19d) set upper and lower bounds on the air mass flow rate and the temperature, respectively. The limits $u_{\min,1}$ and $u_{\max,1}$ are calculated based on the fan speed limits of 20% and 80% and the speed-to-flow fan model according to

$$u_{\min,1} = h\,(20\%), \quad u_{\max,1} = h\,(80\%) \ . \quad (10.20)$$

The time-varying temperature limits $x_{\min,k}$ and $x_{\max,k}$ reflect a time-varying comfort zone (different limits for working and non-working hours).

The reserve scheduling problem is subject to uncertainty because the regulation signal w_k is unknown. The formulation (10.19) builds robustness to any possible w_k throughout the whole scheduling horizon by imposing the robust input and state constraints (10.19c) and (10.19d).

In terms of instantaneous power, the worst case is either full up-reserve activation or full down-reserve activation. For this reason, we use the uncertainty set $w_k \in [-1,1] \ \forall k$ in (10.19c). In terms of energy, though, one has to consider that the regulation signal has usually energy limits because of activation of tertiary control reserves or redispatch. We account for these energy limits in the formulation by using the uncertainty set $w_k \in [-w_{\mathrm{lim}}, w_{\mathrm{lim}}] \ \forall k$ in (10.19d), where $0 < w_{\mathrm{lim}} \le 1$.

Reformulation and Approximation of the Robust Problem

Problem (10.19) is not directly solvable, but a deterministic robust counterpart problem can be obtained by exploiting the problem's structure. It is sufficient to consider only the boundaries of the uncertainty w_k and formulate the input and state constraints based on them, which results in the robust counterpart problem (10.21), where $\overline{x}_k$ and $\underline{x}_k$ are

the worst-case higher and lower state trajectories, respectively. Proposition 5 states that problems (10.21) and (10.19) are equivalent. Before presenting and proving Proposition 5, we first prove two useful lemmas.

$$\min_{u_k, r_{\mathrm{u},k}, r_{\mathrm{d},k}} \sum_{k=0}^{N_1-1} c_k f(u_k) - \lambda_k \big[f(u_k + r_{\mathrm{u},k}) - f(u_k - r_{\mathrm{d},k}) \big] \quad (10.21\mathrm{a})$$

$$\text{s.t. } \overline{x}_{k+1} = A\overline{x}_k + B_u T_{\mathrm{s},k} \cdot f^{-1} \left(P_k - w_{\lim} R_{\mathrm{u},k} \right) +$$
$$B_{xu} \overline{x}_k \cdot f^{-1} \left(P_k - w_{\lim} R_{\mathrm{u},k} \right) + B_v v_k, \ \forall k \quad (10.21\mathrm{b})$$

$$\underline{x}_{k+1} = A\underline{x}_k + B_u T_{\mathrm{s},k} \cdot f^{-1} \left(P_k + w_{\lim} R_{\mathrm{d},k} \right) +$$
$$B_{xu} \underline{x}_k \cdot f^{-1} \left(P_k + w_{\lim} R_{\mathrm{d},k} \right) + B_v v_k, \ \forall k \quad (10.21\mathrm{c})$$

$$u_{\min,1} \le u_k - r_{\mathrm{d},k}, \forall k \quad (10.21\mathrm{d})$$

$$u_k + r_{\mathrm{u},k} \le u_{\max,1}, \forall k \quad (10.21\mathrm{e})$$

$$x_{\min,k} \le \underline{x}_k, \ \forall k \quad (10.21\mathrm{f})$$

$$\overline{x}_k \le x_{\max,k}, \ \forall k \ . \quad (10.21\mathrm{g})$$

Lemma 1. *Function $f(u)$ is both monotonic and convex in its domain for the parameters of Table 10.7.*

Proof. The proof is given in Appendix A.   ∎

Lemma 2. *If $w_k \in [-w_{lim}, w_{lim}]$ with $0 < w_{lim} \le 1$, the following statements are true:*

$$\min_{w_k} \left(u_k + \Delta u_k \right) = f^{-1} \left(P_k - w_{lim} R_{u,k} \right) \ \text{for } w_{lim} \le 1 \quad (10.22)$$

$$\max_{w_k} \left(u_k + \Delta u_k \right) = f^{-1} \left(P_k + w_{lim} R_{d,k} \right) \ \text{for } w_{lim} \le 1 \quad (10.23)$$

$$\min_{w_k} \left(u_k + \Delta u_k \right) = u_k - r_{d,k}, \ \text{for } w_{lim} = 1 \quad (10.24)$$

$$\max_{w_k} \left(u_k + \Delta u_k \right) = u_k + r_{u,k}, \ \text{for } w_{lim} = 1 \ . \quad (10.25)$$

Proof. The proof is given in Appendix A.   ∎

Proposition 5. *Under assumption 1, optimization problems (10.19) and (10.21) are equivalent.*

Proof. The proof is given in Appendix A.   ∎

The dynamics in (10.21) involve the inverse of a polynomial combination of optimization variables, which is very complicated. We approximate (10.21) by a simple linearization of the inverse function that leads to the following problem:

$$\min_{u_k, r_{\mathrm{u},k}, r_{\mathrm{d},k}} \sum_{k=0}^{N_1-1} c_k f(u_k) - \lambda_k \big[f(u_k + r_{\mathrm{u},k}) - f(u_k - r_{\mathrm{d},k}) \big] \quad (10.26\mathrm{a})$$

$$\text{s.t. } \overline{x}_{k+1} = A\overline{x}_k + B_u T_{\mathrm{s},k} \cdot (u_k - w_{\lim} \cdot r_{\mathrm{d},k}) +$$
$$B_{xu}\overline{x}_k \cdot (u_k - w_{\lim} \cdot r_{\mathrm{d},k}) + B_v v_k, \ \forall k \quad (10.26\mathrm{b})$$

$$\underline{x}_{k+1} = A\underline{x}_k + B_u T_{\mathrm{s},k} \cdot (u_k + w_{\lim} \cdot r_{\mathrm{u},k}) +$$
$$B_{xu}\underline{x}_k \cdot (u_k + w_{\lim} \cdot r_{\mathrm{u},k}) + B_v v_k, \ \forall k \quad (10.26\mathrm{c})$$

$$(10.21\mathrm{d}), (10.21\mathrm{e}), (10.21\mathrm{f}), (10.21\mathrm{g}) \ .$$

For the general case $0 < w_{\lim} \leq 1$, problems (10.21) and (10.26) are not equivalent. In fact, (10.26) overestimates both the maximum state trajectory (highest room temperature) and the minimum state trajectory (lowest room temperature), as shown by Proposition 6. In the special case $w_{\lim} = 1$ though, the approximation is exact and problems (10.21) and (10.26) are equivalent, as shown by Proposition 7. The value $w_{\lim} = 1$ corresponds to the most conservative case where the energy limits of the frequency regulation signal are neglected.

Proposition 6. *Let us denote by $\overline{x}_k^\star$ and $\underline{x}_k^\star$ the maximum and minimum state trajectories of the original problem (10.21). Furthermore, let us denote by $\overline{x}_k^*$ and $\underline{x}_k^*$ the maximum and minimum state trajectories obtained by (10.26). If w_{lim} satisfies $0 < w_{lim} \leq 1$, the inequalities $\overline{x}_k^* \geq \overline{x}_k^\star$ and $\underline{x}_k^* \geq \underline{x}_k^\star$ hold for any time step k.*

Proof. The proof is given in Appendix A. ∎

Proposition 7. *If w_{lim} is equal to 1, then problems (10.21) and (10.26) are equivalent.*

Proof. The proof is given in Appendix A. ∎

Remark 4. *The overestimation of maximum and minimum temperature trajectories by (10.26) builds additional robustness to temperature excursions above $x_{max,k}$, but it reduces the robustness to temperature excursions below $x_{min,k}$. However, the reduction of robustness at the lower*

*temperature range is not expected to create significant comfort zone vio-
lations. This is because the state trajectory will generally remain closer
to $x_{max,k}$ than $x_{min,k}$ due to minimization of the energy consumption
costs in (10.26a).*

Problem (10.26) is a deterministic non-linear optimization problem with
cubic objective function, bilinear equality constraints, and linear in-
equality constraints. Although this is a non-convex problem, it is pos-
sible to solve it in due time using a non-linear solver because of its
relatively small size. The problem is compiled using YALMIP [43] and
solved with IPOPT.

Although not shown in (10.26) for simplicity, the state constraints are
modeled as soft constraints using slack variables with high penalties.
This is done to avoid infeasible solutions due to plant-model mismatch
or prediction errors while running the experiment.

The main outcome of (10.26) is the up-reserve $R_{u,k}$ and down-reserve
$R_{d,k}$ capacity for each time slot of the scheduling horizon $k \in [0, N_1 - 1]$.
Although a schedule for the air mass flow rate over the scheduling hori-
zon u_k is also obtained, this is not used at this point. Determining the
air mass flow rate setpoints is the task of the controller of level 2.

Worst-Case Energy Limit of Regulation Signal

The energy limit w_{lim} of the regulation signal corresponds to the worst-
case normalized reserve request. In this chapter, we analyze 2-month
historical data of the RegD signal from PJM (December 2012 to January
2013). Figure 10.11 shows the cumulative distribution of the RegD
signal's energy content over 15 minute intervals, as well as the actual
worst case, the median, 95%, 97.5%, and 99% percentiles. It is apparent
that considering the actually observed worst case of $w_{lim} = 0.88$ would
lead to very conservative solutions. Instead, we define the worst case
as the 97.5% percentile of the distribution ($w_{lim} = 0.25$), because it is
a reasonable trade-off between robustness and conservatism. In other
words, the requested reserve energy by the RegD signal over 15 minutes
will be less than 25% of the reserve capacity with probability 97.5%.

Note that the worst-case reserve request of $w_{lim} = 0.25$ for each time
step is not conditional on the reserve requests during the previous time
steps. However, due to the predictive nature of reserve scheduler, the
worst-case reserve request has to be defined across the prediction hori-
zon. We define the horizon's worst-case reserve request by taking the
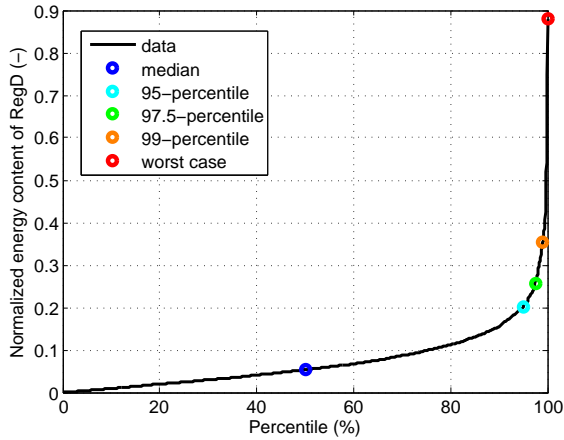
Figure 10.11: Cumulative distribution of the RegD signal's energy content over 15 minute intervals.

worst-case reserve request for each time step independently, as shown in (10.19). Modeling the reserve request as an uncertainty that lives in a multidimensional box constraint is inspired by [175].

Arguably, this approach will lead to conservative solutions, because it is unlikely that the worst case will be realized for every time step. However, we adopt this approach because it is simple and allows us to derive a tractable approximation of the robust reserve scheduling problem. Another advantage of this approach is that it provides some additional degree of robustness that might be desirable in the presence of other uncertainties, such as modeling and/or forecast errors, as it will be shown in our experimental results section.

Another option would be to determine the worst-case reserve energy request over multiple 15 minute intervals (e.g., over hourly intervals), and use an approach similar to the one proposed in [177] to derive the robust counterpart problem. The challenge however would be that the building model is bilinear in our case, whereas it was linear in [177].

Modeling of Reserve Product Constraints

Problem (10.26) allows us to select different reserve capacities for different time slots, as well as different $R_{u,k}$ and $R_{d,k}$ for the same time

slot. In practice, many reserve markets have requirements on the structure of the reserve product, in particular reserve blocks with minimum duration and/or symmetric up- and down-reserves.

Reserve blocks with minimum duration: Reserve blocks with minimum duration of $T_{\text{res}} \in \mathbb{N}$ time steps with a duration of 15 minutes each can be modeled by adding in (10.26) the constraints

$$R_{\text{u},k} = R_{\text{u},k+j}, \tag{10.27}$$
$$R_{\text{d},k} = R_{\text{d},k+j}, \; \forall k = n T_{\text{res}} + 1, \; \forall j \in \{1, \ldots T_{\text{res}}\} \; , \tag{10.28}$$

where $n \in \mathbb{N}$ and $n \leq (N_1 - 1)/T_{\text{res}}$. In this chapter, we select $T_{\text{res}} = 4$ to require the reserve capacities to be constant over periods of 1 hour. This is a reasonable assumption, especially for power systems where the reserve market is co-optimized with the energy market.

Symmetric reserve capacities: Symmetric reserve capacities can be enforced by introducing in (10.26) the constraints

$$R_{\text{u},k} = R_{\text{d},k} \; \forall k \; . \tag{10.29}$$

These symmetry constraints might limit the amount of reserves significantly due to the nonlinear flow-to-power fan model. Note that modeling symmetric reserves and/or reserve blocks with minimum duration increases the complexity because (10.27), (10.28) and (10.29) are nonlinear equality constraints on the optimization variables $r_{\text{u},k}$ and $r_{\text{d},k}$.

Asymmetric reserve capacities: In general, allowing different values for the up- and down-reserve capacity increases the flexibility and, possibly, the total amount of reserves extracted from a building. As shown in [175, 177], down-reserves are preferable for buildings because they can be offered without increasing the baseline energy consumption. Therefore, if asymmetric electric reserves are allowed by removing constraint (10.29), the optimal solution of (10.26) is likely to include zero up-reserves, especially if the same profit is assumed for up- and down-reserves. Since this is a limiting factor for our experimental demonstration, a different formulation for asymmetric reserves is used.

We impose the following symmetry constraint on the thermal reserves

$$r_{\text{u},k} = r_{\text{d},k} \; \forall k \; , \tag{10.30}$$

but the electric reserves are still asymmetric due to the nonlinear fan curve and (10.13), (10.14). In this case, the building will still favor down- over up-reserves, but it has to provide both of them at the same time. Instead of adding (10.30) in (10.26), the problem can be reformulated by substituting the redundant variables $r_{u,k}$ and $r_{d,k}$ with a new variable r_k, which reduces the total number of optimization variables and, possibly, the computation time. Note that (10.29) and (10.30) cannot occur at the same time due to the nonlinear fan curve, i.e., either the thermal or the electric reserves have to be asymmetric.

Limitations, Alternatives, and Extensions

Although solving (10.26) in due time is possible for FLEXLAB, this might not be the case for buildings with many zones or in case of reserve scheduling within building aggregations [175,177,178]. In this case, $f(u)$ can be approximated by a piecewise affine function, which will remove the fan curve non-linearity but will introduce binary variables.

The only remaining non-linearity will be the bilinear building dynamics that can be handled using sequential convex optimization, which is a standard practice in energy-efficient building control [24]. If energy consumption minimization is the only objective, sequential convex optimization usually converges to a good feasible solution. However, if maximization of reserve profits is an additional objective, it is unclear whether sequential convex optimization converges smoothly. This is an interesting direction for further research.

The conservatism of (10.26) can be reduced by using scenario-based optimization. The reserve uncertainty can be modeled with regulation signal scenarios obtained from historical data, and $r_{u,k}$, $r_{d,k}$ can be selected such that the comfort zone is satisfied for all scenarios. However, selecting the number of scenarios to satisfy the comfort constraints with a given probability (the so-called chance constraints) is still an open problem for the general case of non-convex problems.

Note that the disturbances v_k are included in (10.26) using the predicted values, but without accounting for prediction errors. Therefore, we build robustness only to w_k, but not to plant-model mismatch and disturbance prediction errors. An alternative approach would be to model the weather uncertainty using scenarios, which would also result in a scenario program.

10.5.2   Level 2: Room Climate Controller

MPC Formulation

The goal of level 2 controller is to determine the air mass flow rate set-point u_k, such that the scheduled reserves $R_{u,k}$ and $R_{d,k}$ from level 1 can be provided while satisfying room comfort. We use the following robust MPC formulation for this purpose, where the robustness is against the uncertain reserve w_k:

$$\min_{u_k, r_{u,k}, r_{d,k}} \sum_{k=0}^{N_2-1} c_k f(u_k) \tag{10.31a}$$

$$\text{s.t. } \overline{x}_{k+1} = A\overline{x}_k + B_u T_{s,k} \cdot (u_k - w_{\lim} \cdot r_{d,k}) + \\ B_{xu}\overline{x}_k \cdot (u_k - w_{\lim} \cdot r_{d,k}) + B_v v_k, \ \forall k \tag{10.31b}$$

$$\underline{x}_{k+1} = A\underline{x}_k + B_u T_{s,k} \cdot (u_k + w_{\lim} \cdot r_{u,k}) + \\ B_{xu}\underline{x}_k \cdot (u_k + w_{\lim} \cdot r_{u,k}) + B_v v_k, \ \forall k \tag{10.31c}$$

$$u_{\min,2} \leq u_k - r_{d,k}, \forall k \tag{10.31d}$$

$$u_k + r_{u,k} \leq u_{\max,2}, \forall k \tag{10.31e}$$

$$x_{\min,k} \leq \underline{x}_k, \ \forall k \tag{10.31f}$$

$$\overline{x}_k \leq x_{\max,k}, \ \forall k \tag{10.31g}$$

$$R_{u,k}^* = f(u_k) - f(u_k - r_{d,k}), \ \forall k \tag{10.31h}$$

$$R_{d,k}^* = f(u_k + r_{u,k}) - f(u_k), \ \forall k \ . \tag{10.31i}$$

Problem (10.31) is similar to (10.26) with the main differences being (i) the electric reserve capacities $R_{u,k}^*$ and $R_{d,k}^*$ are fixed from level 1, and (ii) the only objective is to minimize the energy cost.

Note that the MPC is free to select a pair $(u_k, r_{u,k}, r_{d,k})$ different to the one calculated in level 1, provided that the electric reserves $R_{u,k}^*$ and $R_{d,k}^*$ can be provided, which is enforced by the nonlinear equality constraints (10.31h) and (10.31i). Similarly to (10.26), the state constraints are modeled as soft constraints in (10.31) to avoid infeasible solutions while running the experiment. In addition, similarly to (10.26), predictions for the disturbances v_k are used in (10.31), which are updated at every time step.

Observe that the upper and lower bounds on the air flow rate $u_{\min,2}$ and $u_{\max,2}$ in the MPC controller are different to the ones used in the reserve scheduler. We select the following less tight bounds in the MPC

controller to allow more freedom when setting the air flow rate to meet the comfort zone constraints:

$$u_{\min,2} = h\,(10\%)\,, \quad u_{\max,2} = h\,(90\%) \ . \tag{10.32}$$

Kalman Filter

Recall that the MPC controller uses the 2-state building model (10.4) and that the temperature of the lumped thermal mass cannot be measured directly. Furthermore, although the room temperature is measured with a sensor, the resulting measurement is noisy. For these two reasons, we developed an extended Kalman filter (because the building model is bilinear) to provide a state estimate $\hat{x}_k$ to the MPC controller. Assuming additive process and measurement noise, the a priori error covariance $P_{\mathrm{e},k}^-$, a posteriori error covariance $P_{\mathrm{e},k}$, and Kalman gain K_k are given by [197]

$$P_{\mathrm{e},k}^- = F_k P_{\mathrm{e},k} F_k^\top + Q \tag{10.33}$$

$$K_k = P_{\mathrm{e},k}^- H_k^\top (H_k P_{\mathrm{e},k}^- H_k^\top + R)^{-1} \tag{10.34}$$

$$P_{\mathrm{e},k} = (I - K_k H_k)P_{\mathrm{e},k}^- \ , \tag{10.35}$$

where F_k is the Jacobian matrix of partial derivatives of the system dynamics, and H_k is the Jacobian matrix of partial derivatives of the output (measurement) function $y_k = Cx_k$, both with respect to the state x_k; Q and R are the process and measurement noise covariance matrices, respectively; and I is the identity matrix. Let us denote by f_x the bilinear dynamics and by f_y the linear output equation. The matrices F_k and H_k are obtained from the building dynamics and the output equation as

$$F_k = \left.\frac{\partial f_x}{\partial x}\right|_{\hat{x}_{k-1},u_{k-1}} = A + B_{xu}\hat{x}_{k-1}u_{k-1} \tag{10.36}$$

$$H_k = \left.\frac{\partial f_y}{\partial x}\right|_{\hat{x}_{k-1},u_{k-1}} = C \ . \tag{10.37}$$

The process and measurement noise covariance matrices are fixed to[5]

---

[5]The measurement noise covariance is set to $R = 0.1$ based on the accuracy of the temperature sensors. The RMSE of the building model provides an estimate of the process noise covariance. With reference to Table 10.4, the diagonal elements of Q can be set to $0.42^2 = 0.1764$. However, the larger value 0.4 was chosen because the model's out-of-sample RMSE will be higher than 0.42.
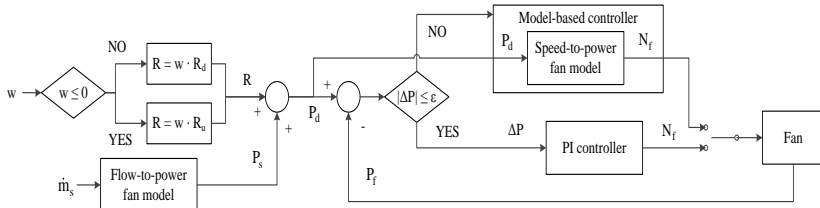
Figure 10.12: The control block of the level 3 switched controller.

$$Q = \begin{bmatrix} 0.4 & 0 \\ 0 & 0.4 \end{bmatrix}, \quad R = 0.1 . \tag{10.38}$$

Note that the same Kalman filter is also used before each execution of the reserve scheduler in level 1.

10.5.3   Level 3: Frequency Regulation Controller

The goal of level 3 is to control the fan speed such that the fan power tracks the frequency regulation signal. This approach is different from [132] that used the frequency of the VFD as a control variable. Controlling the fan speed setpoint avoids interference with the VFD's internal control logic, and it leverages the existing fan speed feedback controller. In addition, our approach is different from [127], where a fan speed command was superimposed on the output of the fan controller.

There are four important design requirements for the frequency regulation controller: (i) fast response, (ii) minimal computation effort, (iii) accuracy, and (iv) stability. Two classical control approaches are envisaged. The first approach is a feedforward (open-loop) controller that uses the static speed-to-power fan model (10.10). The advantage of this controller is that it is inherently stable due to the absence of feedback. The disadvantage, of course, is that there will most certainly be a steady-state control error. The second approach is a model-free, feedback (closed-loop) controller, for example, a Proportional-Integral (PI) controller. A PI controller can reduce the steady-state error, but its stability is not guaranteed and requires gain tuning.

In order to combine the advantages of these two control approaches, we propose to implement the frequency regulation controller as a switched controller consisting of two blocks: (i) Ctrl1: a model-based, feedforward controller, and (ii) Ctrl2: a model-free, feedback PI controller.

The control block of the switched controller is shown in Fig. 10.12, whereas its discrete-time implementation is described by Algorithm 3.

Algorithm 3 Implementation of the switched controller

1: initialize old tracking error: $e_{old} = 0$
2: initialize fan speed: N_f
3: **while** experiment is running **do**
4:      calculate baseline power: $P_s = f(\dot{m}_s)$
5:      calculate reserve request: $R = w \cdot R_d, \quad$ if $w \geq 0, \quad$ and
6:                                    $R = w \cdot R_u, \quad$ if $w < 0$
7:      calculate desired fan power: $P_d = P_s + R$
8:      **repeat**
9:          measure fan power P_f
10:          calculate new tracking error: $e_{new} = P_d - P_f$
11:          **if** $|e_{new}| \leq \varepsilon$ **then**
12:              calculate PI output: $N_{f,pi} = N_f + K_p(e_{new} - e_{old}) + K_i \Delta t e_{new}$
13:              cap fan speed: $N_f = \min[\max(N_{f,pi}, N_{f,min}), N_{f,max}]$
14:              set fan speed to N_f
15:              set old tracking error to: $e_{old} = e_{new}$
16:          **else**
17:              set fan speed to: $N_f = g(P_d)$
18:              set old tracking error to: $e_{old} = 0$
19:          **end if**
20:      **until** elapsed time is equal to control loop duration
21: **end while**

Step 4 of Algorithm 3 uses the flow-to-power fan model (10.9) to translate the scheduled flow rate by the MPC of level 2 to baseline power consumption. The desired fan power P_d is computed at step 7 based on the baseline, the reserve capacity of level 1, and the received regulation signal. The new control error e_{new} is calculated at step 10 as the difference between P_d and the measured fan power P_f. At step 11 the condition $|e_{new}| \leq \varepsilon$ is checked to decide whether Ctrl1 or Ctrl2 will be used at the current time step (ε is a parameter that represents the accuracy of the steady-state fan model). If $|e_{new}| \leq \varepsilon$ holds, then Ctrl2 is used, and the PI controller's discrete-time implementation is from step 12 to step 15. On the other hand, if $|e_{new}| > \varepsilon$, Ctrl1 is used, and the fan speed is determined at step 17 based on the speed-to-power model (10.10).

After a large power setpoint change, Ctrl1 becomes active and tries to

restore the fan power close enough to the new setpoint using the fan speed-to-power model. Ctrl1 remains active for as long as the absolute tracking error is larger than the tolerance ε. By analyzing the model fitting results of Section 10.4.2 we observe that the largest deviation between the model and the filtered data points is approximately 30 W, and so we fix $\varepsilon = 30$ W.

At the moment when the absolute tracking error becomes smaller than ε the controller switches from Ctrl1 to Ctrl2, and it remains like this for as long as the error is less than ε. At the moment when the error exceeds ε again, we switch back to Ctrl1 and reset the integral error to zero, as shown at step 18. This practice improves the control performance because it avoids large overshoots due to accumulated errors when setpoints change suddenly. In addition, whenever the output of Ctrl2 is larger than 90% (percentage of the maximum fan speed) or smaller than 10%, we cap or floor the fan speed to these values.

Due to the nonlinear relationship between the fan speed and the fan power, gain scheduling was used to tune the PI controller. Distinct operating regions in terms of fan power were defined and different sets of gains were calculated for each region. Motivated by the speed-to-power curve of the fan we chose 5 operating regions. We ran several step response tests and tuned the proportional (K_p) and integral gains (K_i) for each region using the Ziegler-Nichols method [198]. Following this method, we initially set $K_i = 0$ and gradually increase K_p until a critical value is reached, where stable and consistent oscillations are observed in the output. The critical proportional gain and the period of the oscillations are used to set the values of K_p and K_i according to empirical rules.

Note that the gains obtained with the Ziegler-Nichols method served only as an initial guess because the controller is switched, rather than a simple PI controller. The final gain values were determined with trial and error, and are presented in Table 10.8. The selected integral gains are significantly higher than the ones suggested by the Ziegler-Nichols method, whereas the selected proportional gains are slightly lower than the ones proposed by the method. The reason why the integral gains turned out to be more important than the proportional gains is because the goal of the PI controller (Ctrl2) is to correct the bounded steady-state error of Ctrl1, but not to recover the system after a large setpoint change.

This switched controller implementation provides advantages in terms

Table 10.8: Tuning results of the PI controller gains

Operating region (W)	Proportional gain (K_p)		Integral gain (K_i)	
	Ziegler-Nichols	Selected	Ziegler-Nichols	Selected
$[0, 500)$	N/A	0.004	N/A	0.01
$[500, 1000)$	0.009	0.004	0.001106	0.0035
$[1000, 1500)$	0.010125	0.004	0.001245	0.003
$[1500, 2000)$	0.00675	0.0045	0.00083	0.0025
$[2000, 2500)$	0.0045	0.004	0.000559	0.002

of stability and performance compared with the open-loop controller of [132] or the closed-loop controller of [127]. The controlled system is inherently bounded and stable: even if the PI loop tends to be unstable, whenever the tracking error is greater than ε, the control will be passed to the (stable) open-loop controller. Ctrl1 allows us to track fast and sudden power setpoint changes. To track such changes accurately with a simple PI controller, much larger gains would be needed at the expense of compromising stability. In contrast, the role of Ctrl2 in the proposed implementation is to correct steady-state errors and therefore relatively small gains can be chosen that achieve a stable response.

10.6 Experimental Results

10.6.1 Experiment Plan

The experiment was organized into two parts. The first part of the experiment took place in the period $15-18$ November 2015 and relied on the building model identified with data from June-July 2015 (see Table 10.5). On 19 November the experiment was paused and the so far collected data were used to identify a new building model (see Table 10.6), which was then used in the second part of the experiment from 20 to 21 November. In both parts of the experiment the 2-state model with 1-day ahead prediction was used. In the rest of this section we present a selection of the most interesting results.
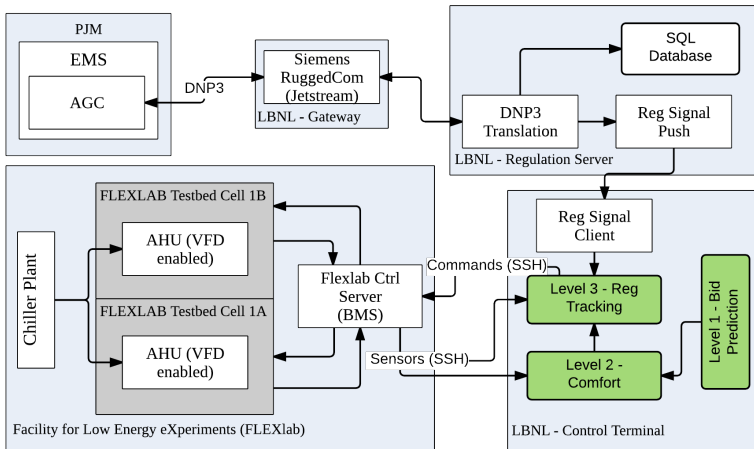
Figure 10.13: The complete communication architecture from PJM to FLEXLAB. The LBNL Control Terminal is the developed hierarchical controller.

10.6.2   Data Sources and Experiment Settings

Before starting with the analysis of the results, we briefly mention the sources of external data and summarize some important experiment settings. Most of the experiment was performed using archived data of PJM's RegD signal from December 2012 to January 2013. Although the historical signal was available with a resolution of 2 seconds, we down-sampled it to 4 seconds due to the communication delays of FLEXLAB. During the course of the experiment, a connection with PJM was established based on the DNP3 protocol and using a Siemens Jetstream gateway, which provided us with the RegD signal in real time. At the FLEXLAB side, the received data were translated based on the DNP3 protocol, saved in an SQL database, and pushed by a "RegD signal server" to a "RegD signal client" (the hierarchical controller of this chapter). The complete communication architecture from PJM to FLEXLAB is graphically shown in Fig. 10.13. However, the connection was not very reliable at the FLEXLAB side and the RegD signal was often not received. For this reason, we ran the experiment with the real-time PJM signal for only one hour.
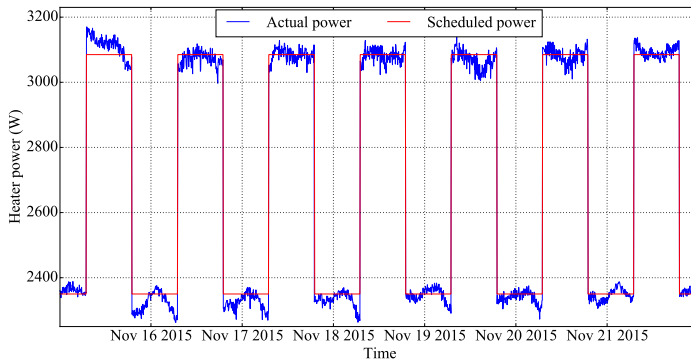
Figure 10.14: The heater schedule and the actual power consumption during one week. The heater consumption is high during working hours and low during non-working hours.

The ambient temperature forecasts were obtained from the *forecast.io* database. The solar radiation forecasts were obtained from a clear-sky radiation model, which turned out to be sufficient for the weather conditions during the experiment. Forecast.io provides also cloud cover forecasts that were used within a neural network to improve the predictions of the clear-sky model. However, the results were not satisfactory (perhaps due to the limited spatial resolution of the cloud cover forecast) and therefore the clear-sky predictions were used.

Two electric heaters were used to emulate the internal heat gain. Their consumption profile was fixed according to the red curve of Fig. 10.14 using digital timer sockets. The actual heater power (blue curve in Fig. 10.14) generally follows the consumption profile, but it fluctuates around it due to variations in the socket's voltage. The temperature comfort zone was set to $21 - 25°$C during working hours.

Recall that the SAT is regulated to $17°$C with a PI controller that controls the position of a cooling valve. This PI controller was tuned by FLEXLAB engineers before the start of the experiment. However, since the tuning was performed for conventional building operation, the SAT fluctuated a lot when the building was providing regulation. To reduce the fluctuations we empirically re-tuned the gains of the PI controller, which resulted in the SAT behavior shown in Fig. 10.15. The mean deviation from the SAT setpoint is $0.05°$C, i.e., there is a small bias to larger SAT values, and the mean absolute deviation is $0.37°$C.
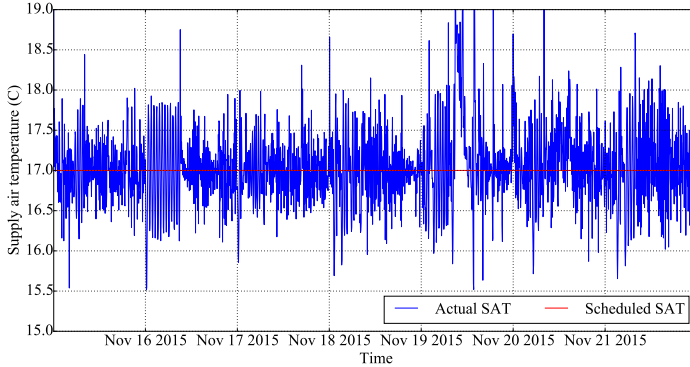
Figure 10.15: The SAT setpoint and actual values during one week.

The electricity cost was assumed constant and equal to $c_k = 0.18 \, €/\text{kWh}$, whereas the reserve capacity payment was fixed to a 10% higher value, i.e., $\lambda_k = 0.198 \, €/\text{kWh}$. The goal of this work is to demonstrate the technical viability of reserve provision from commercial buildings; therefore, a relatively high capacity payment was chosen to incentivize reserve provision. Finally, the penalty for the soft constraints of the optimization problems of level 1 and level 2 was set to 10^5.

10.6.3   Reserve Scheduling Results (Level 1)

In this section, we present results relevant to the reserve scheduler. Two main factors that determine the amount of reserves are the building's energy capacity and the symmetry of reserve capacity. Apart from the physical properties of the building, the energy capacity depends also on the comfort zone's width. In this experiment, we specifically address the effect of the so-called night setback by enlarging the comfort zone during unoccupied hours to $12 - 35°\text{C}$. We defined four combinations, namely symmetric or asymmetric reserves and with or without setback, and performed six full-day experiments. Hourly reserve capacities were used in all cases.

Figure 10.16 shows results for 20 November when symmetric reserve capacities were assumed, and for 21 November when asymmetric capacities were used (in both dates night setback was applied). The capacities are reported in % of the fan rated power that is 2500 W. Observe that
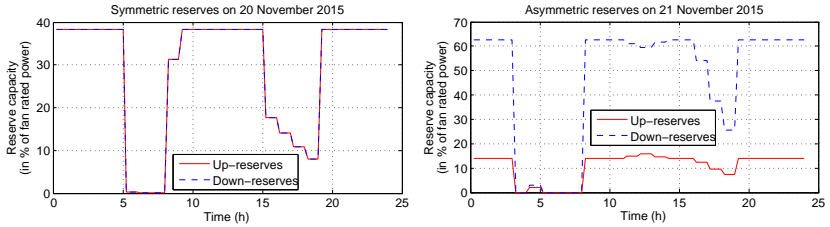
Figure 10.16: The hourly reserve capacities as a percentage of the nominal fan power for 20 November 2015 (left) and 21 November 2015 (right). Note that the two plots have different scales.
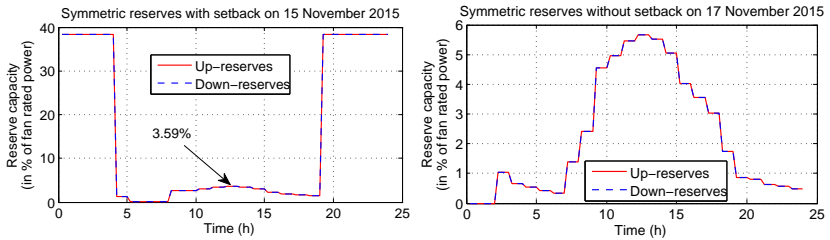
Figure 10.17: The hourly reserve capacities as a percentage of the nominal fan power for 15 November 2015 (left) and 17 November 2015 (right). Note that the two plots have different scales.

the reserve capacity is maximized at night when the comfort zone is enlarged, and during the hottest part of daytime. In case of symmetric reserves, the maximum capacity is slightly less than 40% of the rated fan power. In case of asymmetric reserves, the maximum up-reserve capacity is approximately 15% of the rated fan power, whereas the down-reserve capacity is more than 60%. These experimental results are in agreement with relevant simulation results reported in [175, 177], and show that down-reserves are indeed preferable for commercial buildings. The reason is that down-reserves (consumption increase) can be provided without increasing the building's baseline and the respective energy costs.

Figure 10.17 compares the experimental results of 15 November, when setback was used, with those of 17 November, when no setback was applied (symmetric reserve capacities were used in both dates). Observe that the capacity curve is qualitatively different: with setback most of

Table 10.9: Reserve capacities for different combinations of night setback and reserve symmetry. The capacities are daily average values expressed in % of the fan's nominal power (2500 W).

Date	Symmetric, Setback		Asymmetric, Setback		Symmetric, No setback		Asymmetric, No setback	
	R_u	R_d	R_u	R_d	R_u	R_d	R_u	R_d
15/11	**15.61**	**15.61**	6.45	26.15	1.91	1.91	1.64	4.42
16/11	9.09	9.09	**3.90**	**14.88**	0.74	0.74	0.90	1.85
17/11	11.44	11.44	5.39	21.21	**2.24**	**2.24**	1.85	4.59
18/11	16.70	16.70	7.78	31.75	3.94	3.94	**3.28**	**9.91**
20/11	**28.95**	**28.95**	11.89	49.66	15.07	15.07	7.82	28.81
21/11	22.10	22.10	**10.72**	**46.55**	13.60	13.60	6.79	26.30

the reserve is provided at night, whereas without setback the reserve provision coincides with the highest cooling load in the middle of the day. Of course, the reserve capacity depends on the weather conditions and this is why the capacity profile is considerably different during daytime on 15 and 20 November, although the experiment was conducted with setback and symmetric reserves on both dates.[6]

To allow for a more thorough comparison, we simulated the reserve capacity scheduling for all combinations of symmetry and setback and summarize the results in Table 10.9. The capacities are daily average values in % of the fan's nominal power. The values in bold indicate the experimental results of each day, whereas the values in normal font correspond to the simulation results. The capacity ranges from low values below 1% to high values nearly 50%, and it heavily depends on the reserve symmetry, setback, and weather conditions.[7]

Analyzing the values of Table 10.9 we find that the night setback increases the capacity by 177.0% on average for symmetric reserves, by 107.0% for asymmetric up-reserves, and by 150.7% for asymmetric down-reserves. Reversely, if setback is already used, adopting asymmetric capacities instead of symmetric capacities reduces the up-reserves by 55.6%, but increases the down-reserves by 83.1%, and therefore the net

---

[6]Another reason why the reserve capacities are quite different is that different building models were used in these two days (see Section 10.6.1)

[7]In the symmetric case, the maximum reserve capacity corresponds to 50% of the fan's nominal power.

effect is an increase of 13.7% in the total capacity. Overall, our results show that asymmetric reserves are preferable for commercial buildings and that using a night setback significantly increases the reserve potential.

10.6.4   Room Climate Control Results (Level 2)

In this section, we present results from the operation of cell 1A while providing frequency regulation, and from cell 1B that is used as a real-time benchmark. Our analysis focuses on the comfort zone satisfaction, the performance of the building model and Kalman filter, the effect of frequency regulation on SAT and energy consumption, and the computation time of the MPC of level 2.

Comfort Satisfaction

Experimental results for 17, 18, 20 and 21 November are shown in Figs. 10.18 to 10.21. In all figures, the top plot shows the temperature trajectories in cells 1A and 1B, the middle plot presents the forecasts and the actual values for ambient temperature and solar radiation[8], whereas the bottom plot shows the SAT and the air flow rate in cell 1A. The comfort zone is indicated with red, and two values for the upper limit can be seen. The actual upper limit (red solid line) is 25°C, but a tighter limit of 24°C (red dashed line) is used within the MPC to account for modeling and prediction errors.

In Figures 10.18 and 10.19 the cell 1B is under energy-efficient operation and this is reflected in the temperature, which remains constantly close to the upper limit of the comfort zone. On the other hand, in Figs. 10.20 and 10.21 the cell 1B is in a "regulation-ready" operation mode, namely the HVAC consumption is scheduled identically to cell 1A to allow reserve provision, but no regulation signal is received. For this reason, the temperature trajectories of the two cells are very close to each other for most of the time on 20 and 21 November. This result is also a validation that the two cells are indeed nearly identical.

---

[8]The middle plot shows the total global radiation, which includes the long-wave radiation losses from the building envelope to the atmosphere. The long-wave radiation losses become dominant at night and may result in a negative total radiation. This effect is known as nighttime radiation cooling [199].
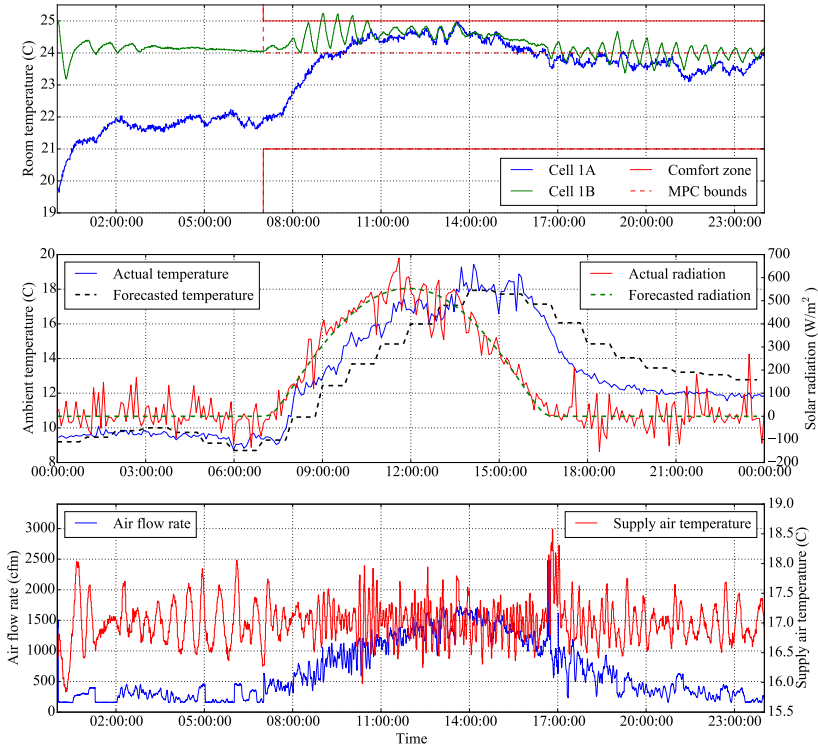
Figure 10.18: Experimental results for the room climate controller under frequency regulation on 17 November.

However, there exist discrepancies between the state trajectories from 12.00 to 19.00 on 20 November and from 07.00 to 17.00 on 21 November. The discrepancies on 20 November are due to calibration differences of the fan models in the two cells. As a result, small changes in fan speed result in significant temperature differences due to the high heating load from the heaters. The discrepancies on 21 November are because of interruptions in the hierarchical control in cell 1B due to server connection timeout error from approximately 07.00 to 11.00. When the server was unresponsive, the cell was controlled by an existing fallback controller.

Observe that the temperature trajectory of cell 1B in Figs. 10.18 and 10.19 remains in the band $[24 - 25]°C$ for most of the day. This illustrates the necessity of the applied constraint tightening in the MPC to
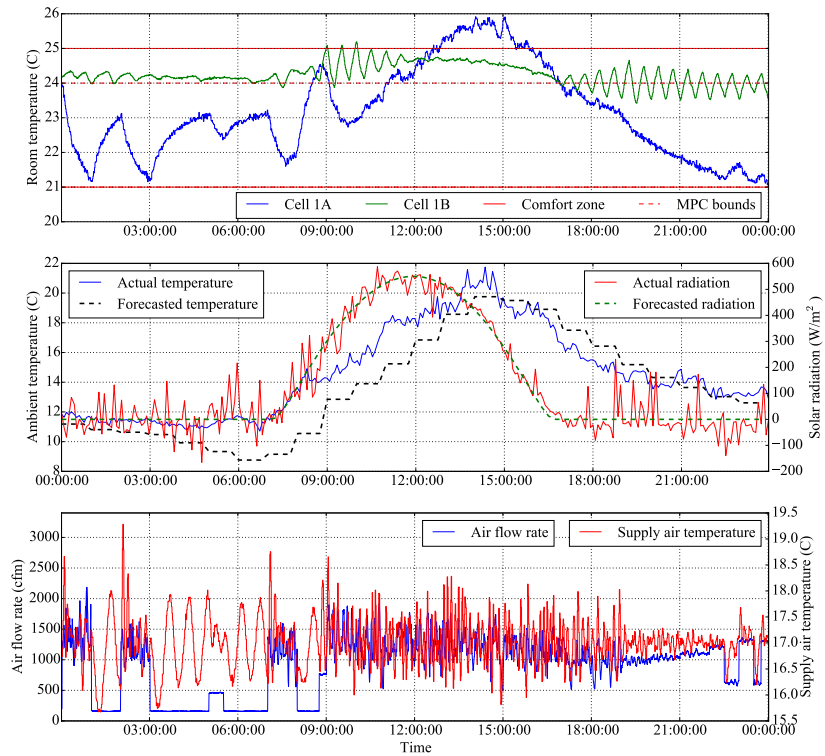
Figure 10.19: Experimental results for the room climate controller under frequency regulation on 18 November.

compensate for modeling errors. In the same figures, and particularly in Fig. 10.19, the temperature trajectory of cell 1A is more variable compared with cell 1B, and it follows the scheduled reserve and air flow rate. In Fig. 10.18 frequency regulation is provided while respecting the comfort zone.

However, on 18 November (Fig. 10.19) the comfort zone is violated approximately from 13.00 to 16.00 in cell 1A, but not in cell 1B. There are two main reasons for this: (i) the actual ambient temperature is constantly higher than the day-ahead forecast from the beginning of the day until 15.00; and (ii) asymmetric reserves are used on 18 November, instead of symmetric reserves as on 17 November. The asymmetry allows for a more aggressive scheduling with a larger down-reserve capacity on
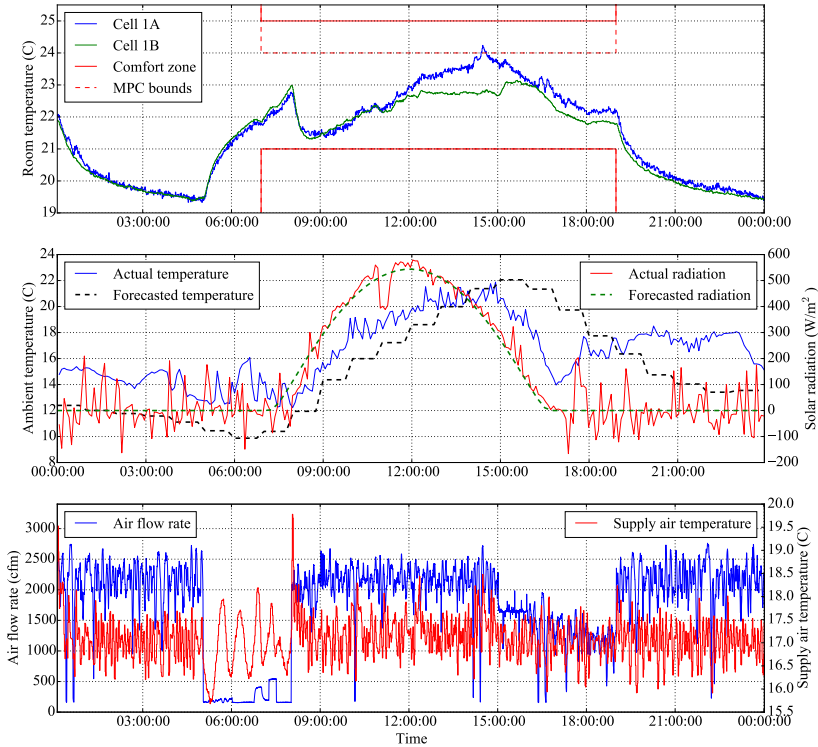
Figure 10.20: Experimental results for the room climate controller under frequency regulation on 20 November.

18 November in comparison with 17 November (see Table 10.9). This clearly shows that if the combined modeling and prediction error exceeds the robustness margin of the controller, frequency regulation with day-ahead bidding of the reserve capacity might have an adverse effect on occupant comfort.

The control performance is significantly better on 20 and 21 November (Figs. 10.20 and 10.21) despite the large discrepancies between the day-ahead ambient temperature forecasts and the actual values. No comfort zone violations occur and moreover the temperature keeps even below the MPC constraint of 24°C for most of the time. The performance improves mainly because a recently calibrated building model is used in these experiments (see Table 10.6). Therefore, periodic calibration
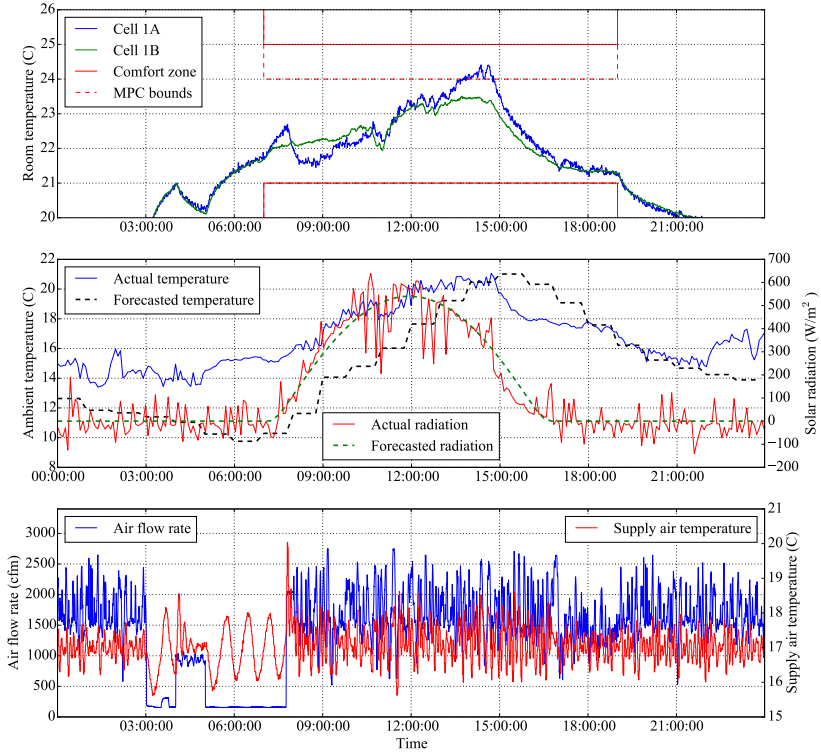
Figure 10.21: Experimental results for the room climate controller under frequency regulation on 21 November.

of the building model is important to account for seasonality and eliminate any systematic errors. The calibration can be also done in an online fashion, where the model parameters are fitted on a daily basis using recent measurements. From Figs. 10.20 and 10.21 we conclude that if the model and weather predictions are sufficiently accurate, the robustness provided by the reserve scheduling formulation allows a commercial building to bid in day-ahead markets for frequency regulation products.

The pattern of the temperature trajectory in Figs. 10.20 and 10.21 is typical for a building with night setback. The reserve capacity is high at night resulting in high cooling power and low temperatures, whereas the capacity is lower during the working hours (with a higher heating load)
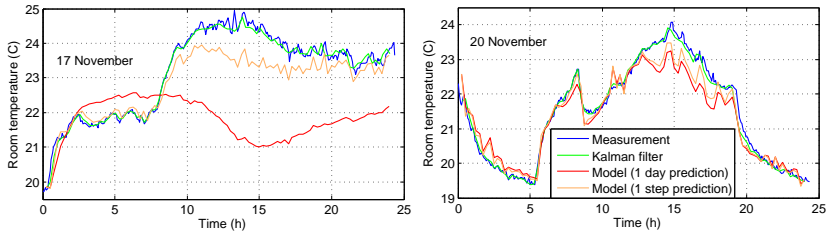
Figure 10.22: Model performance in predicting the future room temperature and estimating the current room temperature with a Kalman filter. Left: Results for 17 November with the older model. Right: Results for 20 November with the new model.

resulting in higher temperatures. A comparison of the temperature trajectories in cells 1A and 1B shows that tracking the RegD signal has little effect on the room temperature due to the signal's limited energy content.

Model and Estimator Performance

The importance of the building model is further discussed in this section. Figure 10.22 compares the out-of-sample performance of the older model (left plot) and the new model (right plot). The blue curve is the room temperature measurement, and the green curve is the estimated temperature using the building model and a Kalman filter. The red curve corresponds to a day-ahead model prediction, whereas the orange curve shows the step-ahead temperature predictions. Clearly, the new model outperforms the older one, especially for the day-ahead predictions. This explains why the performance of the level 2 controller is much better on 20 November than on 17 November in terms of comfort zone violations.

Figure 10.23 shows the effect of model accuracy on the MPC operation. On 17 November the model mismatch is large (because the older model is used), which results in a significant discrepancy in the scheduled air flow rate and fan power between level 1 and level 2. The MPC reacts on the modeling error in an interesting way: compared with level 1, the cooling power in level 2 is reduced during night hours and increased during daytime. By shifting cooling energy to the middle of the day, the MPC can provide the same amount of electric reserve with less
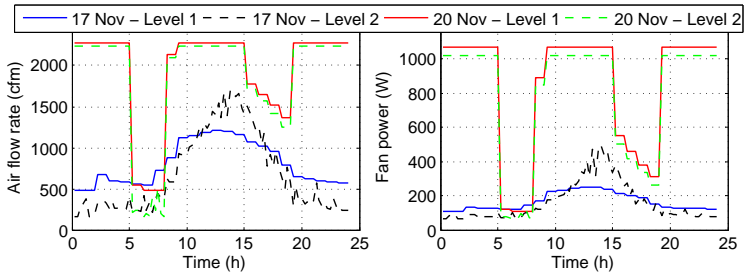
Figure 10.23: Left: Air flow rate schedule in level 1 and level 2. Right: Fan power schedule in level 1 and level 2.

change in air flow rate by taking advantage of the nonlinear fan curve. On the other hand, the model mismatch is small on 20 November (because the new model is used) and therefore the air flow and fan power schedules of level 1 and level 2 are close to each other. In fact, level 2 consistently schedules less cooling power than level 1. This is possible because constraints (10.32) are more relaxed than (10.20), and the day-ahead scheduling in level 1 is robust and thus conservative.

Fan Heat Gain at High Speeds

In this section, we present results on the dependence of SAT and cooling valve opening on fan speed. The blue points in Fig. 10.24 are measurements and the red trend represents a polynomial fit on them. The trend in cooling valve opening is increasing for the whole fan speed operation range. This is reasonable and it means that the higher the fan speed is, the more cooling is required from the chilled water loop. On the other hand the trend in SAT depends on the fan speed. For fan speeds up to 50% the SAT trend is practically a flat line. However, for speeds above 50% (and especially above 70%) there is a clear increasing SAT trend despite the increased cooling valve opening.

These results lead to an interesting observation: the heat gain due to the rotation of the fan is significant at high speeds, and it cannot be effectively rejected by exchanging heat with the chilled water loop. In other words, if the fan operates at a high speed (70% or higher), then the SAT will have a steady-state deviation from the setpoint 17°C, which can be as high as 1°C according to Fig. 10.24. Steady-state SAT deviations will likely result in comfort zone violations, because the
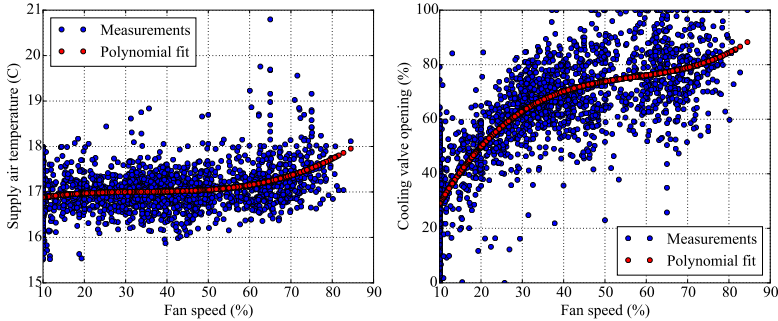
Figure 10.24: The dependence of SAT and cooling valve opening on fan speed.

scheduling is performed assuming an SAT fixed to 17°C. This did not create problems in our experiment because the scheduled fan speed by the MPC was at most 70%. Nevertheless, the heat gain by fan rotation might be important to consider in similar experimental demonstrations or real-word implementations.

Effect on Energy Consumption

A major concern when providing AS with commercial buildings is the effect on energy consumption. Reference [137] reported a round-trip efficiency of approximately 46% in an experiment with a commercial building. As explained in Chapter 9, there are two types of efficiency losses relevant to frequency regulation: "reserve availability efficiency loss" and "reserve utilization efficiency loss". The first one is the efficiency loss due to scheduling the consumption in an energy suboptimal way to be able to provide frequency reserves, if requested. The second one is the additional efficiency loss while tracking the frequency regulation signal.

In this section, we briefly report some efficiency results from our experiment; however, a detailed energy efficiency analysis is outside the scope of this chapter. The efficiency loss is calculated comparing the energy consumption of cell 1A (E_{1A}) with that of the benchmark cell 1B (E_{1B}) according to

$$\eta = \frac{E_{1A} - E_{1B}}{E_{1B}} \cdot 100\% \ . \tag{10.39}$$

Table 10.10: Effect of frequency regulation on energy consumption

1B operation mode	Energy efficient 15/11-18/11		Regulation-ready			
			20/11 (0-24 h)		20/11 (0-12 h)	
Cell	1A	1B	1A	1B	1A	1B
Fan energy (kWh)	27.22	16.23	19.24	20.91	9.40	10.19
Cooling (gpm · F)	1989.92	1800.55	772.16	788.69	328.78	331.09
Mean temp. (C)	22.85	24.43	21.45	21.26	21.05	21.05

This is a challenging task because there is a single chiller providing with cooling power both cells, and it is not obvious how to allocate the total chiller electric consumption to the two cells. We use two different definitions of energy consumption in (10.39): (i) electric energy consumption of the fan, and (ii) thermal cooling power consumption of each cell. The later is calculated based on the chilled water loop using

$$P_{\text{cool}} = \dot{m}_{\text{cw}} \cdot (T_{\text{ch,r}} - T_{\text{ch,s}}) \ , \qquad (10.40)$$

where $\dot{m}_{\text{cw}}$ is the chilled water flow rate, whereas $T_{\text{ch,s}}$ and $T_{\text{ch,r}}$ are the supply and return chilled water temperatures, respectively.

We report efficiency results in Table 10.10 for: (i) the four-day period from 15 to 18 November, when cell 1B was under energy-efficient operation, to quantify the "reserve availability efficiency loss"; and (ii) 20 November, when cell 1B was in regulation-ready operation mode, to quantify the "reserve utilization efficiency loss".

Based on our results, the "reserve availability efficiency loss" is approximately 68% in terms of fan consumption and 11% in terms of cooling power from the chiller.[9] However, one has to keep in mind that the additional consumption in cell 1A is not entirely wasted because it results in a lower average temperature, as shown in Table 10.10. Therefore, the reserve scheduling entails a significant efficiency loss compared with

---

[9]Note that the efficiency definition used here is different to the one used in [137], and thus the results are not directly comparable. In [137], a power increase with a duration of 15 minutes was requested from the building, then a power decrease of the same magnitude and duration was requested, and in the end a recovery period of 90 minutes was allowed when no demand response event occurred. The efficiency was calculated according to $\frac{E_\downarrow}{E_\uparrow + E_{\text{rec}}} \cdot 100\%$, where $E_\downarrow$ is the energy consumption during the power decrease, $E_\uparrow$ is the energy consumption during the power increase, and E_{rec} is the energy consumption during the recovery period.
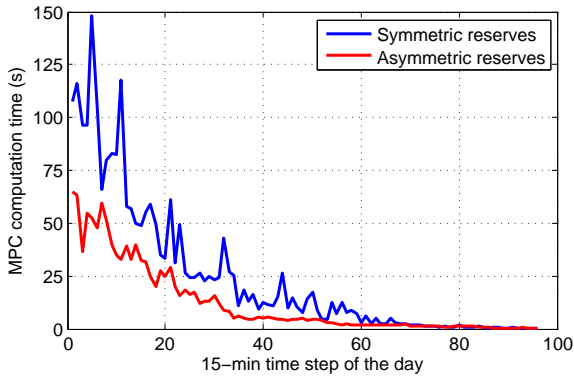
Figure 10.25: The average computation time for each MPC problem depending on the time of the day and on reserve symmetry.

an energy-optimal building operation, even if no reserve is actually requested.

When cell 1B is in the regulation-ready mode it applies the air flow setpoints calculated by the MPC of cell 1A, but it does not track the regulation signal. Table 10.10 shows that cell 1A has actually less energy consumption than cell 1B on 20 November despite frequency regulation. The non-negligible difference in the average temperature of the two cells is due to imperfections in fan model calibration and the limited temperature sensor accuracy. However, even if we consider only the interval 00.00 - 12.00 when the average cell temperatures are identical, the consumption of cell 1A is still lower than that of cell 1B. This result indicates that the "reserve utilization efficiency loss" is negligible while tracking a fast-moving regulation signal like RegD.

MPC Computation Effort

The MPC computation time is relatively low and acceptable for our demonstration. As shown in Fig. 10.25, the longest computation time is approximately 150 seconds for symmetric reserves and 65 seconds for asymmetric reserves. The computation is faster for asymmetric reserves because the total number of optimization variables is lower (see Section 10.5.1) and the nonlinear equality constraints (10.29) are not used.

Furthermore, the computation time generally decreases as we move from the beginning to the end of each day, because a reducing horizon is used in the MPC problem. After the 70th time step, i.e., when the MPC prediction horizon is 26 time steps (6.5 hours) at most, the computation time is less than 2 seconds for both types of reserves. Of course, this would not be the case if an MPC with fixed prediction horizon was used. Figure 10.25 can be used to select the prediction horizon's length depending on the maximum allowable computation time. The computation time grows exponentially with the number of optimization variables due to the nonlinear nature of the optimization problem. Therefore, for larger buildings a lower prediction horizon or model reduction might be necessary to keep the computation time reasonable.

10.6.5 Regulation Signal Tracking Results (Level 3)

Control Performance Metrics

In this section, we present results from level 3 and analyze the tracking performance of the regulation signal. We use different metrics to quantify the control performance: the metrics $e_{c,k}$, $e_{t,k}$ and $e_{r,k}$ describe the instantaneous errors, whereas the metrics e_{me}, e_{mae}, e_{rmse}, $e_{t,mape}$ and $e_{r,mape}$ evaluate the control performance over the experiment's duration. Note that $e_{t,k}$ and $e_{r,k}$ are both instantaneous relative (percentage) errors, but the normalization is performed differently. In $e_{t,k}$ it is done using the current desired fan power, whereas in $e_{r,k}$ using the current up- or down-reserve capacity.

$$
\text{Control error:} \qquad e_{c,k} \;=\; P_{\text{d},k} - P_{\text{f},k} \qquad\qquad (10.41)
$$

$$
\text{Tracking percentage error:} \; e_{t,k} \;=\; \frac{e_{c,k}}{P_{\text{d},k}} \qquad\qquad (10.42)
$$

$$
\text{Reserve percentage error:} \quad e_{r,k} \;=\; \begin{cases} \frac{e_{c,k}}{R_{\text{u},k}}, & \text{if } w_k < 0 \\ \frac{e_{c,k}}{R_{\text{d},k}}, & \text{if } w_k \geq 0 \end{cases} \qquad (10.43)
$$

$$
\text{Mean control error:} \qquad e_{me} \;=\; \frac{1}{N_{\text{exp}}} \sum_{k=0}^{N_{\text{exp}}-1} e_{c,k} \qquad (10.44)
$$

$$
\text{MAE:} \qquad e_{mae} \;=\; \frac{1}{N_{\text{exp}}} \sum_{k=0}^{N_{\text{exp}}-1} |e_{c,k}| \qquad (10.45)
$$

$$\text{RMSE:} \qquad e_{\text{rmse}} = \sqrt{\frac{1}{N_{\text{exp}}} \sum_{k=0}^{N_{\text{exp}}-1} e_{\text{c},k}^2} \qquad (10.46)$$

$$\text{MAPE (tracking):} \qquad e_{\text{t,mape}} = \frac{1}{N_{\text{exp}}} \sum_{k=0}^{N_{\text{exp}}-1} |e_{\text{t},k}| \qquad (10.47)$$

$$\text{MAPE (reserve):} \qquad e_{\text{r,mape}} = \frac{1}{N_{\text{exp}}} \sum_{k=0}^{N_{\text{exp}}-1} |e_{\text{r},k}| \qquad (10.48)$$

In addition to the generic metrics (10.41) - (10.48), we also use the specific score applied by PJM to assess the performance of the resources that provide frequency regulation. The total score S_{tot} consists of three parts, namely the correlation score S_{c}, the delay score S_{d} and the precision score S_{p}, which are defined as [79]

$$S_{\text{c}} = \max_{\tau \in [0,5 \text{ min}]} (R_{\text{cor}}) \qquad (10.49)$$

$$\tau^* = \operatorname*{argmax}_{\tau \in [0,5 \text{ min}]} (R_{\text{cor}}) \qquad (10.50)$$

$$S_{\text{d}} = \left| \frac{\tau^* - 5 \text{ min}}{5 \text{ min}} \right| \qquad (10.51)$$

$$S_{\text{p}} = 1 - \frac{1}{n} \sum_{k=0}^{n-1} \left| \frac{e_{\text{c},k}}{\bar{P}_{\text{d,h}}} \right| \qquad (10.52)$$

$$S_{\text{tot}} = \frac{1}{3} S_{\text{c}} + \frac{1}{3} S_{\text{d}} + \frac{1}{3} S_{\text{p}} \ . \qquad (10.53)$$

The correlation score is the maximum correlation of the reference signal (desired power $P_{\text{d},k}$) and fan power $P_{\text{f},k}$, and τ^* is the time shift at which the correlation is maximized (τ takes a value from 0 to 5 minutes with a step of 10 seconds). The delay score is calculated based on the time shift with maximum correlation. The absolute control error is normalized by the average hourly value of the reference signal $\bar{P}_{\text{d,h}}$ in the precision score calculation, whereas the total score is just a weighted sum of the individual scores.

Experimental Time-Series Results

In Fig. 10.26 we present experimental results from the operation of level 3 controller from 18.30 to 19:30 on 20 November 2015. With the

exception of the duct pressure measurements that are available with a resolution of 1 minute, all other measurements are available with a resolution of 1 second.

During frequency regulation, the air flow rate is proportional to the fan speed and the duct pressure is quadratic to the fan speed, as expected from the fan laws. This is an indication that if the duct system of FLEXLAB is designed to sustain the pressure corresponding to the maximum fan speed, then frequency regulation will not create any pressure problems, and thus pressure constraints are not necessary in the reserve scheduling and MPC formulations.

The direction of the RegD signal changes very often and the signal has a limited energy content, which is considered in the reserve scheduling formulation. Observe that during periods of time when the RegD signal is relatively flat, or the reserve capacity is low, the PI controller (Ctrl2) is active because relatively smooth changes in fan power are requested. On the other hand, whenever the changes in fan power are rapid, the control switches to the model-based feedforward control Ctrl1.

The tracking of the RegD signal is generally very good. However, when large and rapid changes in fan power are requested, overshoots or undershoots might appear. In addition, when the reserve capacities change at the beginning of every full hour, temporarily high errors might occur depending on how different the capacities of the two consecutive hours are.

Similar observations are made on the control error plots. The instantaneous control error $e_{c,k}$ is generally higher when the reserve capacity is high, as well as at the hour change. On the other hand, the instantaneous percentage errors $e_{t,k}$ and $e_{r,k}$ are higher at a low operating fan power and when the reserve capacity is low. The control error $e_{c,k}$ has a negative bias equal to -5.9 W during this 1-hour period. Therefore, the actual fan power is more often higher than the desired setpoint, i.e., the overshoots are more significant than the undershoots.

Evaluation of Tracking Performance

The performance of RegD signal tracking during the 6 days of Table 10.9 is evaluated using metrics (10.41)-(10.48), and the results are presented in Table 10.11. Observe that $e_{r,\text{mape}}$ is significantly larger than $e_{t,\text{mape}}$. However, one has to consider that $e_{r,k}$ and $e_{r,\text{mape}}$ directly depend on the magnitude of the reserve capacity.
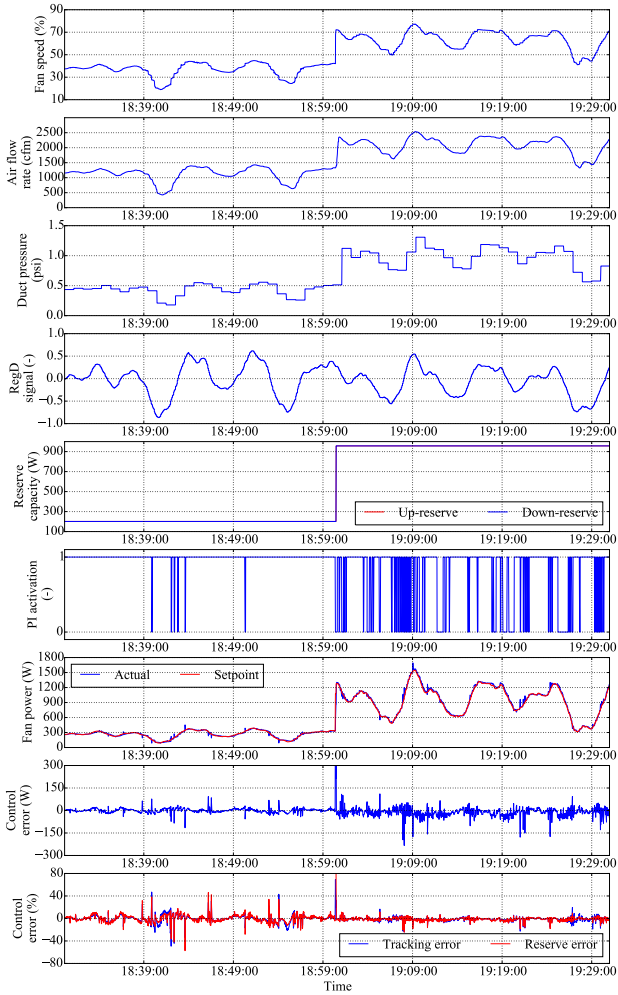
Figure 10.26: Control results from level 3 for a 1-hour period.

We investigate the dependence of control performance on the minimum reserve capacity that can be bid during each hour, which we call "reserve threshold" and denote by R_{thr}. The metrics e_{mae}, e_{rmse}, $e_{\text{t,mape}}$ and $e_{\text{r,mape}}$ are recalculated but the summation is now performed only over the time steps when $R_{\text{u},k} \geq R_{\text{thr}}$ holds if $w_k < 0$, or $R_{\text{d},k} \geq R_{\text{thr}}$ holds

Table 10.11: Performance metrics for tracking the RegD signal during the experiment

Metric	e_{me}	e_{mae}	e_{rmse}	$e_{\text{t,mape}}$	$e_{\text{r,mape}}$
Value	-5.66 W	12.45 W	27.00 W	3.58%	8.23%
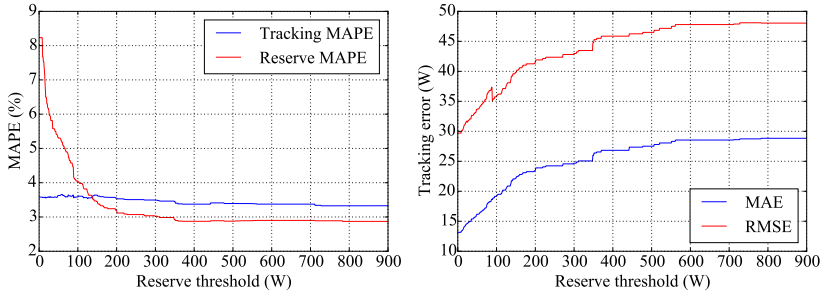
Figure 10.27: Dependence of tracking and reserve errors on the reserve threshold, i.e., the minimum reserve bid for each hour.

if $w_k \geq 0$. We repeat this procedure for different values of R_{thr} and present the results in Fig. 10.27.

In contrast to $e_{\text{t,mape}}$ that remains practically constant (in fact, it decreases at a small rate), $e_{\text{r,mape}}$ decreases rapidly as R_{thr} increases in the range $[0, 200]$ W. This happens because (for the same absolute control error) $e_{\text{r},k}$ decreases, if $R_{\text{u},k}$ or $R_{\text{d},k}$ increase. On the other hand, e_{mae} and e_{rmse} generally increase as R_{thr} increases. This is intuitively expected because the higher the reserve capacity the larger the requested changes in fan power, and thus the higher the errors due to the overshoots and undershoots. Results similar to those of Fig. 10.27 can be used to set minimum or maximum thresholds on the reserve capacity from a tracking performance point of view.

In Table 10.12 we report the PJM scores calculated for the four-day period 15 - 18 November and the two-day period 20 - 21 November. The scores are calculated separately for each hour (only for the hours when the reserve capacity is non-zero) [79], and therefore the values in Table 10.12 are hourly average values. The quality of frequency regulation is exceptional during the whole experiment. For comparison, the highest possible total score is 1 and the minimum total score accepted by PJM is $S_{\text{tot}} = 0.75$. Note that the performance scores are slightly

Table 10.12: PJM scores for tracking the RegD signal (the minimum acceptable score is 0.75 and the maximum possible score is 1)

Score	S_c	S_d	S_p	S_{tot}
$15 - 18$ Nov.	0.89	0.97	0.96	0.94
$20 - 21$ Nov.	0.96	0.99	0.98	0.98
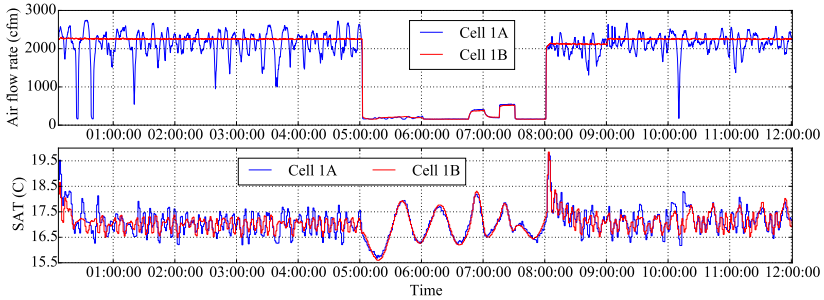
Figure 10.28: Air flow rate (top) and SAT (bottom) in cells 1A and 1B from 00.00 to 12.00 on 20 November. The effect of frequency regulation is high-frequency oscillations on SAT.

higher for $20-21$ November, when the building provides a larger reserve capacity compared with $15 - 18$ November.

Effect on Supply Air Temperature

Tracking the fast-moving RegD signal introduces high-frequency oscillations on SAT, as shown in Fig. 10.28. Cell 1A provides frequency regulation, whereas cell 1B operates in the regulation-ready mode. Observe that the SAT of cell 1A oscillates more than that of 1B, especially after sudden changes in the regulation signal that induce sudden changes in the air flow rate. In addition, large excursions in SAT occur in both cells when the MPC controller changes the air flow setpoint significantly, for example at hour 08.00. Moreover, the magnitude of the SAT oscillations is high even without frequency regulation at low air flow rates, for example from 05.00 to 08.00.
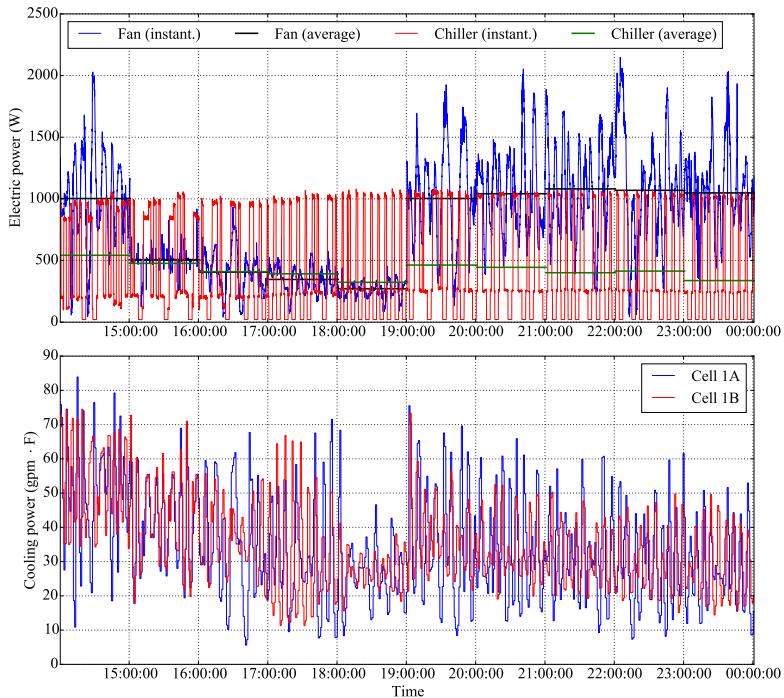
Figure 10.29: Effect of frequency regulation with the fan on the power consumption of the chiller. Top: Instantaneous and average power of the fan in cell 1A and the chiller. Bottom: Cooling power in the chilled water loop for cells 1A and 1B.

Effect of Fan Control on Chiller Power

Since the fan and the chiller are thermally coupled through the chilled water loop, it is worth investigating if the operation of the chiller is affected while providing regulation with the fan.

In Fig. 10.29 we present relevant experimental results for a duration of 10 hours. The top plot shows the instantaneous and hourly-average electric power of the fan in cell 1A and the chiller (recall that cells 1A and 1B are served by the same chiller). The bottom plot shows the thermal cooling power P_{cool} in the chilled water loop for cells 1A and 1B calculated with (10.40).

The chiller has two stages and the electric power consumption is rela-

tively constant at each stage. Of course, the chiller's cycling depends on the cooling load, which in turn depends on the fan power and ambient conditions. In general, as the fan power increases, the chiller is expected to cycle more often and to remain longer at the on state. This can be seen in Fig. 10.29 where the average chiller power (green line) generally follows the average fan power (black line).

During these 10 hours the fan in cell 1A tracks the RegD signal, whereas the fan in cell 1B tracks the setpoint calculated by the MPC controller of cell 1A, which is constant for each hour. The effect of regulation is visible on the thermal cooling power, which fluctuates much more in cell 1A compared with cell 1B. Whenever the fan power increases, the cooling load also increases and the SAT tends to decrease. This is sensed by the SAT controller that opens the cooling valve to compensate for the SAT decrease, which in turn increases the cooling power in the chilled water loop. Therefore, following a fan power change, the thermal cooling power is expected to change within some time frame that depends on the time constant of the cooling valve's controller.

On the other hand, there is no observable effect on the electric power consumption of the chiller, i.e., the cycling within each hour does not practically change. There are two reasons for this: (i) the chilled water is stored in a tank and from there circulated to the building cells, which provides some inertia; and (ii) the RegD signal is approximately zero-mean, and so the hourly-average fan power (black line) is very close to the scheduled value by the MPC of level 2. Note that the gradual reduction in the hourly-averaged electric power consumption of the chiller from 19.00 to 00.00 in Fig. 10.29 is not a side-effect of frequency regulation. Instead, this is the result of the negative trend in cooling power that reflects lower cooling needs, e.g., due to ambient temperature drop.

These results show that frequency regulation can be provided with fan control without observable side-effects on the profile of the chiller's electric power consumption. However, this does not necessarily hold if the regulation signal has a larger energy content, e.g., if RegA is used instead of RegD. In addition, the chiller's consumption will be likely affected during frequency regulation if the chiller is controlled in a continuous manner using a VFD, especially if there is no chilled water storage tank. If the effect on chiller power is non-negligible, the control in level 3 should be revised to mitigate it. This is an interesting direction for future work.
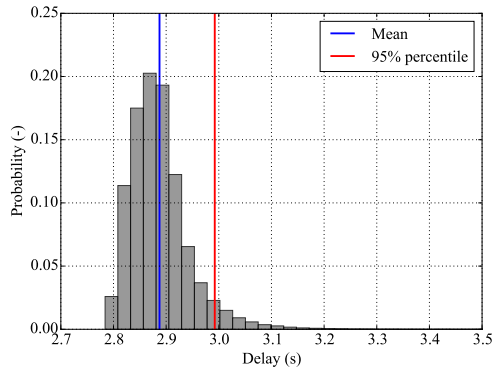
Figure 10.30: Histogram of communication delays during the experiment.

Analysis of Communication Delays

A challenge in this experiment was the communication delays in measurements and actuation. In Fig. 10.30 we present a histogram of the experienced delays during the whole experiment. The probability distribution of the delay is positively skewed with a mean value of 2.89 seconds and a 95%-percentile of 2.99 seconds. In fact, there exist a few very large delays in excess of 5 seconds due to temporary unresponsiveness of the CWS, which are not included in Fig. 10.30. Despite the fact that the average delay is large compared with the time step of the level 3 controller (4 seconds), the tracking performance of RegD signal is good. Since some of the overshoots and undershoots in fan power in Fig. 10.26 are due to the large communication delays, the tracking performance is expected to improve if the average communication delay decreases.

10.7 Lessons Learned and Outlook

In this section, we summarize the most important findings from the experiment and provide suggestions for real-world implementation projects. In addition, we mention a few practical aspects that require more investigation in future work.

10.7.1 Lessons Learned

Hierarchical control is an efficient way to provide frequency regulation with commercial buildings because time-separated tasks are considered

individually. Three control layers are essential: (i) a day-ahead scheduling of the reserve capacity, (ii) a building climate controller to satisfy comfort while leaving enough slack for reserves, and (iii) a frequency regulation controller to track the regulation signal.

Frequency regulation accuracy: High-quality frequency regulation can be provided by controlling the fan speed of AHUs in commercial buildings. The fan power response is very fast and the RegD signal tracking is excellent, even in the presence of communication delays with an average value of approximately 3 seconds in the BAS. However, note that the fan of FLEXLAB has a low rated power of 2500 W and a low inertia; therefore, the control response might not be as fast for larger fans with a higher inertia. A switched controller consisting of a feedforward controller and a PI feedback controller with gain scheduling provides a fast response without compromising stability. This results in a total PJM score as high as 0.98, which is well above PJM's limit of 0.75.

Means to increase reserve capacity: In our experiment, the fan provided $0.74 - 49.66\%$ of its rated power as reserve capacity, depending on ambient and building conditions and reserve assumptions. Allowing asymmetric reserve capacities and using a night setback are effective ways to increase the reserve potential from commercial buildings. In fact, down-reserves are preferable for buildings because the capacity can be offered without increasing the baseline energy consumption.

Occupant comfort: If the building bids in a day-ahead AS market, respecting occupant comfort might be challenging. In our experiment, the comfort is satisfied when the building model and weather forecasts are relatively accurate. Although asymmetric reserves increase the reserve capacity, they result in a more aggressive scheduling that might increase the comfort zone violations.

Building model: The accuracy of the building thermal model is the most important factor in terms of comfort satisfaction. Models with a good day-ahead prediction performance are necessary in order to bid in a day-ahead AS market. It is essential to calibrate the building model periodically to account for seasonality and eliminate any offsets in modeling error.

Performance of MPC: an MPC controller has some important advantages as a building climate controller in frequency regulation applications. Perhaps the most important advantage is that it identifies the optimal balance between leaving enough slack for reserve provision and favouring energy efficiency. Another advantage is that it provides us

with a baseline consumption ahead of real-time operation, which is beneficial from a practical point of view. Moreover, due to its predictive closed-loop nature it reacts to modeling and weather prediction errors in a smart way, and minimizes the adverse effects on occupant comfort.

Robustness measures: It is important to consider the uncertainty of the regulation signal when scheduling the reserve capacity. This can be done by characterizing the worst-case energy content of the signal based on historical data, and accounting for it using robust sets in reserve scheduling and building climate control. Our experience shows that being conservative in modeling of the regulation signal uncertainty is beneficial, because it builds robustness to modeling and weather prediction errors. Additional robustness can be obtained by tightening the comfort zone constraints in the MPC according to the expected modeling error, and allowing a larger fan speed control band in the MPC compared with the day-ahead reserve scheduling.

Effects of frequency regulation on building control: The experimental results show that there are two main side-effects from frequency regulation on building control. First, high-frequency oscillations in SAT are possible, which might require re-tuning of the existing cooling valve controller to obtain tighter SAT control. And second, if the MPC schedules the fan speed at very high values, the cooling loop might not be able to reject the additional heat gain due to fan rotation. However, our results show that these effects are manageable, and that there is no observable effect on the electric power consumption of the chiller, at least while tracking fast-moving frequency regulation signals.

Energy consumption: Provision of frequency reserves entails some efficiency loss. Our results show that the efficiency loss is mainly due to scheduling the HVAC consumption in a suboptimal way compared with an energy efficient building control. This source of efficiency loss can be as high as 67%. On the other hand, the efficiency loss while tracking a frequency regulation signal with limited energy content is negligible.

10.7.2 Outlook

There are several avenues for future experimental work in this field. Some direct extensions of the work presented here is to repeat the experiments with the zone reheating loop activated and/or with the RegA signal of PJM, which is slower but has more energy content. Furthermore, the reliability of the connection between FLEXLAB and PJM can

be improved, and more experiments with the actual real-time signal of PJM can be performed.

A very interesting research direction is to perform frequency regulation experiments with multiple buildings to verify the scalability of the hierarchical control approach. This will leverage the full potential of hierarchical control and is in principle possible to do with FLEXLAB using the remaining three buildings.

In large HVAC systems, a duct pressure controller situated downstream of the fan typically regulates the pressure to a fixed setpoint. When the fan speed setpoint increases, the dampers will close to regulate the temperature of the zones, and this will increase the duct pressure. The pressure controller will reduce the fan speed in an attempt to reduce the duct pressure, effectively rejecting the frequency regulation action. This is an important challenge in large systems that could not be addressed in our experiment due to the small size of FLEXLAB.

The control approach presented here is in principle also applicable to large commercial buildings with many zones. A simple work-around is to (i) deactivate the duct pressure controller, and (ii) command the flow rate setpoint of each VAV box simultaneously with the fan speed command. Doing so will avoid rejecting the frequency regulation control actions, but it will likely increase the variations in duct pressure. This might necessitate the introduction of simplified pressure constraints in the optimization problems of level 1 and level 2, and/or an emergency control layer that falls back to the traditional pressure controller when the pressure exceeds a safety threshold.

10.8 Conclusion

In this chapter, we reported results from a frequency regulation experiment in a commercial building test facility. We developed a hierarchical controller capable of determining the reserve capacity in a day-ahead fashion, scheduling the HVAC systems setpoints accordingly, and tracking the regulation signal in real time. Experimental results show that frequency regulation can be provided from commercial buildings reliably, with high accuracy, and with minimal occupant discomfort. We believe that the proposed hierarchical control can be used in field tests and real-world implementations in larger buildings.

Chapter 11

Effects of Large Shares of Loads in Frequency Control

11.1 Introduction

In this chapter we bring together the approaches developed in Chapters 5, 6, and 9 for Primary Frequency Control (PFC) and Secondary Frequency Control (SFC), and perform dynamic frequency studies with a two-area power system model. PFC is provided by generators and refrigerators, whereas the SFC resources are generators, commercial buildings and Electric Water Heaters (EWHs). The SFC signal sent to the demand-side resources is split between the buildings and EWHs using an appropriate filter, whereas the reserve capacities are allocated among the buildings using techniques from robust optimization. We use the developed model to investigate the effects of large load shares in PFC and SFC on system frequency for different inertia constant values. This chapter is based on [200] and its nomenclature is given in Tables 11.1 - 11.3.

Table 11.1: Nomenclature of Chapter 11: symbols

Symbol	Unit	Description
B	MW/Hz	Frequency bias factor
C_{th}	Wh/°C	Thermal capacitance
C_{p}	-	Proportional gain of SFC
$1/D$	MW/Hz	Damping
F_{HP}	-	Parameter of the turbine transfer function
f_0	Hz	Nominal system frequency
H	s	System inertia constant
$\mathbb{N}$	-	Normal probability distribution
P_{HP}	kW	HP power
P_{n}	W	Refrigerator rated power
P_{PFC}	MW	Total PFC reserve capacity
P_{prim}	MW	Activated PFC power
P_{sec}	MW	Activated SFC power
P_{SFC}	MW	Total SFC reserve capacity
$1/R$	MW/Hz	Droop of primary control
R_{th}	°C/W	Thermal resistance
r	kW	Electric reserve capacity
s	-	Laplace operator
T_{a}	°C	Ambient temperature
T_{G}	s	Time constant of the governor's transfer function
T_{N}	s	Integral time constant of SFC
T_{CH}	s	Time constant of the turbine's transfer function (steam chest)
T_{RH}	s	Time constant of the turbine's transfer function (reheater)
t_{on}^l	s	Refrigerator lock-on time
$\mathbb{U}$	-	Uniform probability distribution
w_{f}	%	Fast component of the SFC signal
w_{s}	%	Slow component of the SFC signal

11.2   Frequency Control Scheme

To design an efficient frequency control scheme, thermal loads can be distinguished based on thermal/electrical dynamics (e.g., ramp rates, response times, and switching constraints), local device controllers, power and energy capacity, and user interaction. We propose to use refrigerators for PFC, as well as an aggregation of EWHs and Heating, Ventilation, and Air Conditioning (HVAC) systems of commercial buildings for SFC.

We combine EWHs with commercial buildings because SFC signals can be significantly biased to the positive or negative direction. Typically, the HVAC systems of commercial buildings have large thermal inertia, and the optimization methods of Chapter 9 can be used to robustly schedule the reserve capacities. However, the dynamic response of Heat Pumps (HPs) or chillers might be slow. EWHs are controlled by simple on/off switches and have fast response times, but their consumption

Table 11.2: Nomenclature of Chapter 11: Greek letters

Symbol	Unit	Description
α	-	Load participation factor in PFC
β	-	Load participation factor in SFC
Δf	Hz	Frequency deviation from f_0
ΔP_{e}	MW	Total electric power deviation
ΔP_{L}	MW	Load disturbance
$\Delta P_{\mathrm{L}}^{\mathrm{f}}$	MW	Response of frequency-dependent loads
ΔP_{m}	MW	Power contribution of PFC and SFC
ΔP_{EWH}	MW	Contribution of EWHs in SFC
ΔP_{HP}	MW	Contribution of HPs in SFC
ΔP_{GEN}	MW	Contribution of generators in PFC and SFC
ΔP_{REF}	MW	Contribution of refrigerators in PFC
ΔP_{T12}	MW	Tie-line power deviation
η	-	Refrigerator's coefficient of performance
τ_{f}	s	Filter time constant
$\tau_1 - \tau_4$	s	HP time constants

Table 11.3: Nomenclature of Chapter 11: subscripts and superscripts

Symbol	Description
b	Building index
t	Time index
$\uparrow$	Up-reserve capacity
$\downarrow$	Down-reserve capacity
$(\bar{\cdot})$	Scheduled power

heavily depends on time-varying water draws. We propose to pool aggregations of commercial buildings and EWHs to collaboratively provide SFC reserves. By splitting the SFC signal into a fast component sent to the EWHs and a slow component sent to the HPs, we can obtain the desired response without user comfort violations.

Remark 5. *Chapter 10 demonstrated that the fans of Air Handling Units (AHUs) in commercial buildings can track fast-moving SFC signals very well. However, many commercial buildings have water-based HVAC systems where water is heated or cooled centrally using an HP or a chiller, and then it is piped to the building zones. The electric power consumption of these water circulation pumps is typically very small (e.g., 1 W of installed power per m^2 of floor area), and therefore the HPs or chillers are considered for SFC in buildings with water-based HVAC systems. This chapter focuses on these types of commercial buildings because the response times of HPs and chillers are longer than those of fans.*

The proposed frequency control scheme is shown in Fig. 11.1. The
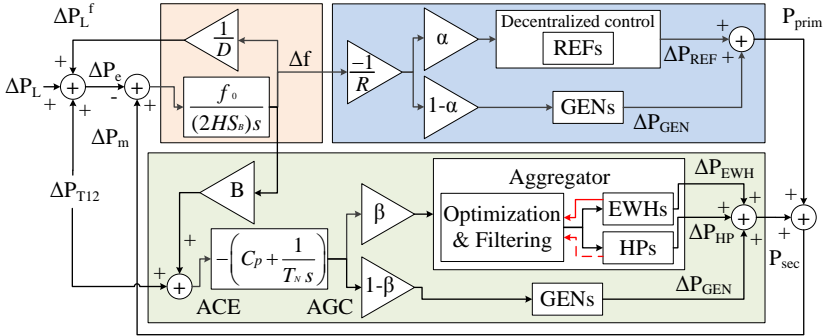
Figure 11.1: The proposed frequency control scheme for area 1 of a two-area power system. The red-solid arrow denotes real-time feedback, whereas the red-dashed arrow denotes daily feedback. Parameter values are given in Table 11.4.

orange part is a linearized power system model incorporating the aggregate system inertia and load damping [71]. The blue and green parts denote the PFC and SFC loops, respectively, where α and β determine the share of thermal loads in frequency reserves. Turbine and governor models for generators are included in the box "GENs", whereas the boxes "REFs", "EWHs", and "HPs" contain the load controllers and models for the refrigerators, the EWHs and the HPs from Chapters 5, 6 and 9.

The HPs are the main SFC reserve resources, and the total reserve capacity is allocated among them day-ahead in a robust way using the formulations of Section 9.4. In real time, the SFC signal w_t is split into a slow ($w_{s,t}$) and a fast component ($w_{f,t}$) using a low-pass Chebyshev filter. The slow component $w_{s,t}$ is broadcast to all buildings, whereas the fast component $w_{f,t}$ is sent to the EWH aggregator. The HP of each building b uses a Proportional-Integral (PI) controller to track the reference power

$$P_{HP,t}^b = \bar{P}_{HP,t}^b + w_{s,t} r_t^b , \tag{11.1}$$

where $\bar{P}_{HP,t}^b$ is the scheduled HP power (i.e., the operating point without SFC), and r_t^b is the electric reserve capacity for SFC.

The EWH aggregator controls the total EWH demand in order to track $w_{f,t}$ by allocating on/off commands based on the device states and priority lists. The aggregator receives measurements of the on/off state

of each EWH (see Fig. 11.1). The EWH temperatures are not communicated to the aggregator, but are estimated applying the method of Chapter 6.4.3 to reduce data transfer. To ensure user comfort, the hysteresis controller overrides any external commands if an EWH is outside its deadband.

Preliminary simulations showed that EWH aggregations can track fast-moving power setpoints that vary $\pm 50\%$ around their baseline consumption without loss of user comfort. Note that there is a tradeoff between increasing the filter's cutoff period τ_f, which increases the amount of EWHs' reserve capacity, and decreasing τ_f, which results in a more volatile signal for the HPs. We observed that $\tau_f = 40$ seconds is a good tradeoff and results in reserve requests from EWHs with a magnitude less than 15% of the buildings' total reserve capacity.

11.3 Case Study

We perform simulations with a two-area power system model using the parameters of Table 11.4. Area 1 is similar to the Swiss system, whereas area 2 is similar to the rest of the interconnected European system. The generators are modeled as steam turbines with reheat and are modeled with the simplified transfer function [71]

$$G_t = \frac{1 + sF_{HP}T_{RH}}{(1 + sT_{CH}) \cdot (1 + sT_{RH})} \ , \tag{11.2}$$

where s is the Laplace operator, F_{HP} is the fraction of the total turbine power generated by the high pressure section, T_{RH} is the time constant of the reheater, and T_{CH} is the time constant of the main inlet volume and steam chest. The governor is modeled as a first-order lag with time constant T_G using the transfer function [71]

$$G_g = \frac{1}{(1 + sT_G)} \ . \tag{11.3}$$

We consider two types of disturbances: (a) large contingencies, e.g., a sudden loss of 3 GW of generation in area 2 that will trigger full activation of PFC reserves; and (b) smaller disturbances, e.g., random load fluctuations in both areas. Disturbances (a) and (b) are considered to assess the quality of PFC and SFC reserves from Demand Response (DR) resources, respectively. We investigate the effect of DR shares,

Table 11.4: Parameters of the two-area power system

Parameter	Variable	Area 1 value	Area 2 value
Base power	S_B	8 GW	240 GW
Damping	$1/D$	120 MW/Hz	3600 MW/Hz
Steam turbine parameter	F_{HP}	0.3	0.3
Steam chest time constant	T_{CH}	0.3 s	0.3 s
Reheat time constant	T_{RH}	7 s	7 s
Governor time constant	T_G	0.2 s	0.2 s
Primary reserves	P_{PFC}	71 MW	2929 MW
Primary control droop	$1/R$	355 MW/Hz	14645 MW/Hz
SFC reserves	P_{SFC}	400 MW	14000 MW
SFC parameters	(C_p, T_N)	$(0.17, 120 \text{ s})$	$(0.17, 240 \text{ s})$

Table 11.5: Refrigeration population parameters

Parameter	Value	Parameter	Value
T_a (°C)	$\mathbb{U}[20, 24]$	$1/(R_{th}C_{th})$ (s^{-1})	$\mathbb{U}[4,6] \cdot 10^{-5}$
Temp. deadband (°C)	$\mathbb{U}[1.7, 2.3]$	R_{th} (°C/W)	$\mathbb{U}[1.6, 2]$
Temp. setpoint (°C)	$\mathbb{U}[4.5, 5.5]$	P_n (W)	$\mathbb{U}[70, 90]$
t_{on}^l (s)	$\mathbb{N}(90, 5)$	η (−)	$\mathbb{U}[1.5, 1.7]$

Table 11.6: Electric water heater and water draw parameters

EWH parameter	Value	Draw parameter	Value
Volume	$\mathbb{U}[100, 400]$ l	long duration	$\mathbb{N}(5, 1)$ min
Rated power	$[3, 6]$ kW	long flow rate	$\mathbb{N}(8, 1)$ l/min
Temp. setpoint	$\mathbb{U}[55, 65]$°C	medium/low duration	$\mathbb{N}(1, 0.1)$ min
Temp. deadband	$\mathbb{U}[5, 15]$°C	medium flow rate	$\mathbb{N}(6, 1)$ l/min
Number of draws	$[10, 60]$	low flow rate	$\mathbb{N}(1, 2)$ l/min

power system inertia levels, and the filter's lower cutoff period on system performance in simulations with a time step of 0.001 seconds for the power system dynamics and 1 second for the thermal load dynamics.

The refrigerator aggregation consists of 100,000 refrigerators with the parameters of Table 11.5. An aggregation of 500 EWHs is created using the parameters of Table 11.6. For each EWH, the tank volume, temperature setpoint, and temperature deadband are sampled according to uniform distributions with the bounds of Table 11.6, whereas the rated

Table 11.7: Commercial building and heat pump parameters

Parameter	Value	Parameter	Value
Floor area	$\mathbb{U}[10000, 20000]$ m^2	COP	$\mathbb{U}[2.5, 3.5]$
Comfort zone	$[21, 24]°$C	τ_1, τ_2	$\mathbb{U}[0.5, 3]$ s
Rated power	$\mathbb{U}[100, 260]$ kW	τ_3	$\mathbb{U}[25, 75]$ s
Minimum loading	$\mathbb{U}[15, 30]$ %	τ_4	$\mathbb{U}[10, 20]$ s

power and the daily number of water draws depend on the tank volume. We differentiate among three types of water draws: high (e.g., shower), medium (e.g., dishwasher), and low (e.g., hand washing) draws. The draw duration and the water flow rate follow normal distributions with the mean values and standard deviations of Table 11.6.[1]

We consider an aggregation of 16 typical Swiss office buildings. The buildings differ with respect to the construction type (heavy or light), the window area fraction (high or low), and the internal heat gains (high or low). We also consider two possible HVAC systems: (a) heating with radiators and cooling with cooled ceilings, and (b) heating and cooling using Thermally Activated Building Systems (TABS). The floor area, HP rated power, Coefficient of Performance (COP), and the HP time constants ($\tau_1 - \tau_4$) follow uniform distributions with the bounds of Table 11.7, whereas the HP minimum loading is a percentage of its rated power, and the comfort zone is the same for all buildings. Typical occupancy profiles are used, whereas weather data for a winter day are provided by Meteoswiss (the Swiss federal office of meteorology and climatology). For more information, the reader is referred to Sections 2.2.1, 9.3 and 9.7.

The reserve capacities are allocated among the buildings by solving the Robust Problem with Power Constraints (RPC) formulated in (9.24). An hourly time step is used, constant reserve capacities over a day and capacity payments 50% higher than the electricity price are assumed (i.e., $\lambda_k^c/c = 1.5$ using the notation of Chapter 9), and the optimization problem is solved using CPLEX. The short-term HP capacity allocation problem is performed applying (9.43) with an horizon of $N_2 = 600$ seconds, it is compiled using YALMIP [43] and solved with CPLEX. Simulations are performed in MATLAB with the aforementioned load aggre-

---

[1]The water draw model is very similar to the one presented in Chapter 6.2.2. The only difference is that three types of water draws are considered here, in contrast to Chapter 6.2.2, where only two types were considered.
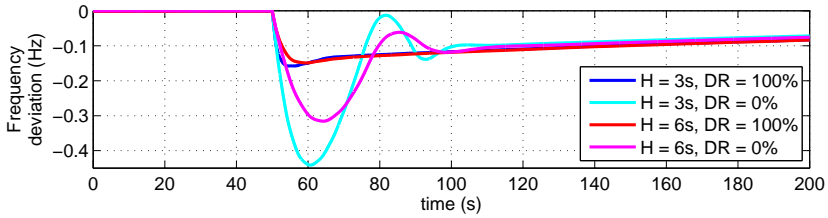
Figure 11.2: Frequency trajectory after a sudden loss of 3 GW of generation in area 2, with and without DR and for two inertia levels.
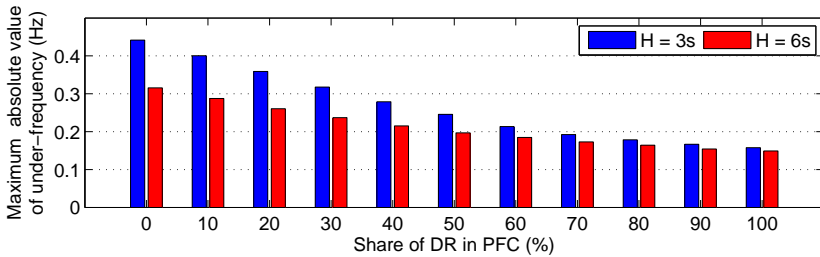
Figure 11.3: Absolute value of maximum under-frequency for different DR shares in PFC and for two inertia levels.

gations, and the results are scaled up to match the reserve requirements of each power system area.

11.4   Simulation Results

11.4.1   Large Disturbances and PFC Performance

In the first set of simulations, we subject the system to a sudden loss of 3 GW of generation in area 2 and investigate the frequency trajectories for the inertia values $H = 3$ seconds and $H = 6$ seconds, as well as for 0% and 100% DR shares. We assume DR resources in PFC and conventional resources in SFC in both areas. Figure 11.2 shows that in case of slow conventional generators (in these simulations we used steam turbines with reheat), the contingency leads to a significant transient under-frequency. This is more pronounced for low inertia values, e.g., in systems with large shares of Renewable Energy Resources (RES),
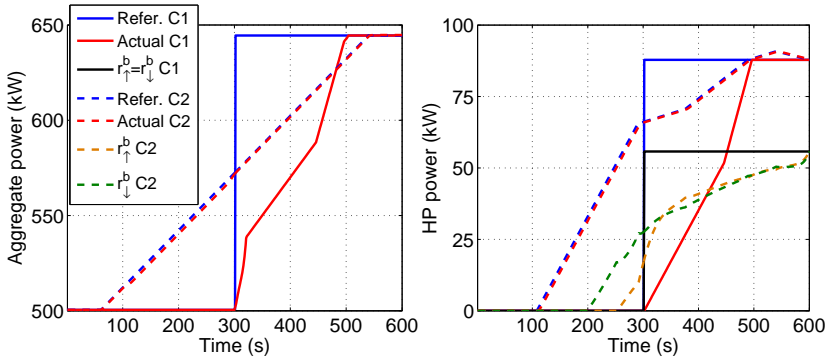
Figure 11.4: Left: Baseline power of the HP aggregation with and without ramping optimization. Right: Baseline and reserve capacities of a HP with and without ramping optimization.

where the frequency approaches the statutory limit of 49.5 Hz below which load shedding occurs. The load control significantly reduces the magnitude of under-frequency for both inertia levels (the frequency is above 49.84 Hz), which is due to the refrigerators' fast response times.

We repeated the same simulations for DR shares α in the range $[0\%, 100\%]$ with a step of 10%, and present the resulting maximum under-frequency values in Fig. 11.3. Due to the refrigerators' fast response times, the maximum under-frequency decreases as the DR share increases. Note that the under-frequency decreases at a faster rate as the DR share increases up to 60%, compared with larger shares. Since the main contribution of refrigerators is at the beginning of the disturbance, even small shares can be effective due to their fast reaction. Observe that the gap between the low and high inertia values decreases as the DR share increases, and is practically zero for a 100% share. Therefore, integrating fast PFC resources such as refrigerators is an efficient measure to counteract inertia reduction in power systems with large RES shares.

11.4.2 SFC Reserve Capacity Allocation

First, we allocate the SFC reserve capacity for each hour of the day among the 16 buildings by solving problem (9.24). Then, we apply the HP ramping optimization problem (9.43) to reallocate the reserve capacities among the HPs during the hour changes.
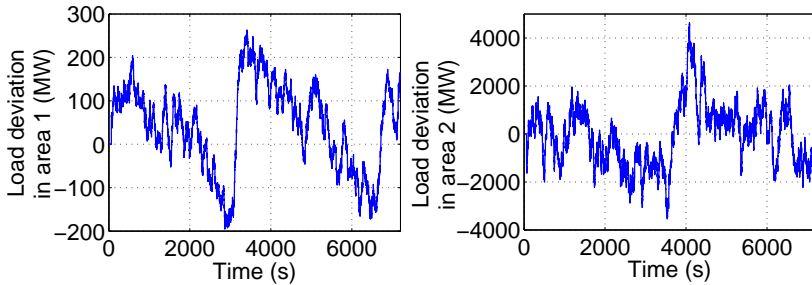
Figure 11.5: Load deviations in areas 1 and 2 constructed such that the resulting SFC signals resemble actual SFC signals from the Swiss power system.

The left plot of Fig. 11.4 shows the aggregation's baseline (red solid curve) without ramping optimization (case C1), where step changes on the HP schedules and reserve capacities are applied (blue solid curve). The left plot also shows the baseline (red dashed curve) with ramping optimization (case C2), which is smoother and very close to the desired ramp with a duration of eight minutes (blue dashed curve). The right plot focuses on the ramping process of an HP that switches on at the hour change. The solid step curves indicate the baseline and reserve capacity trajectories without HP ramping optimization, whereas the dashed curves are the respective values with HP ramping optimization. In the latter case, the HP switches on slightly before the hour change and different short-term up- and down-reserve capacities are allocated to achieve a smoother aggregate baseline and accurate reserve provision.

11.4.3   Small Disturbances and SFC Performance

In the second set of simulations, we investigate the system performance during two hours under regular load fluctuations in both areas that activate SFC (see Fig. 11.5). We focus our analysis on area 1 that has a 50% DR share in both PFC and SFC, whereas conventional PFC and SFC resources are assumed in area 2 (the inertia constant is set to $H = 6$ seconds in both areas). The original SFC signal and its slow $(w_{s,t})$ and fast components $(w_{f,t})$ are shown in Fig. 11.6. The slow component $w_{s,t}$ absorbs the bias of the original SFC signal. The fast component $w_{f,t}$ absorbs the signal's high-frequency components, it
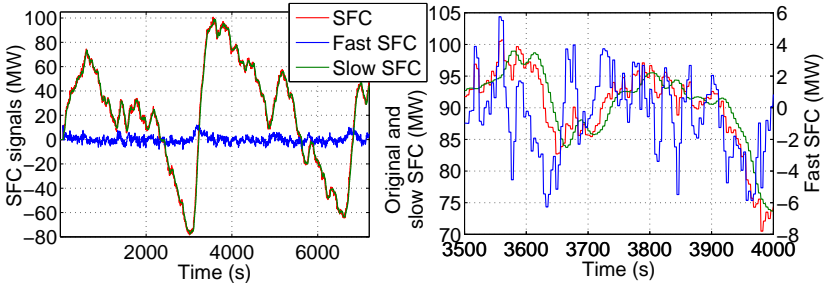
Figure 11.6: Left: The original SFC signal, as well as its slow and fast components for the 2-hour period. Right: Close-up view of the signals during the hour change (there are two y-axes in this plot).
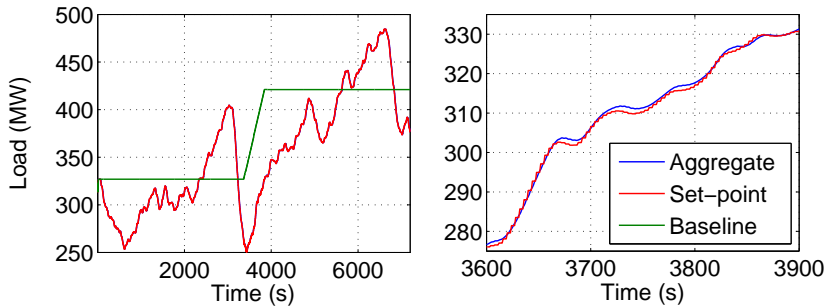
Figure 11.7: Left: Baseline power, setpoint power, and aggregate power of the HPs. Right: Close-up view of the setpoint and aggregate power during the hour change.

is approximately zero-mean, and its magnitude is generally very small compared with $w_{s,t}$ (notice the two y-axes in the right plot of Fig. 11.6).

The left plot of Fig. 11.7 shows the resulting aggregate HP power after reserve provision as well as the aggregation's baseline, i.e., the consumption without reserve provision. The HPs follow $w_{s,t}$ very accurately throughout the whole simulation period with very small tracking errors, as shown in the right plot of Fig. 11.7.

Similar results are presented for the EWHs and $w_{f,t}$ in Fig. 11.8. Since the dynamic response of EWHs is practically instantaneous, the small tracking errors are only due to control with partial state information.
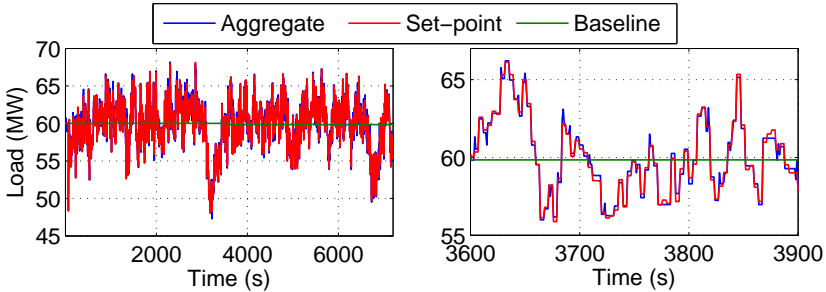
Figure 11.8: Left: Baseline power, setpoint power, and aggregate power of the EWHs. Right: Close-up view of the setpoint and aggregate power during the hour change.
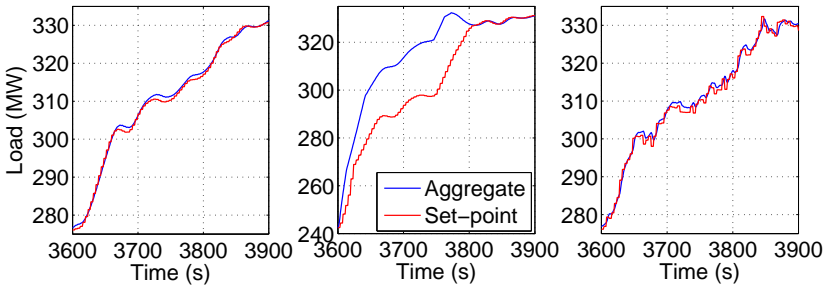
Figure 11.9: Setpoint tracking by the HPs for three cases. Left: Base case. Middle: Without ramping optimization. Right: Without filtering.

Recall that only the on/off states of the EWHs are communicated to the controller, whereas the temperatures are estimated to reduce data transfer. The temperature estimation errors result in the control overshoots or undershoots of Fig. 11.8.

To quantify the importance of filtering and HP ramping optimization, we compare the following cases: (a) both filtering and HP ramping optimization are applied (base case), (b) filtering is applied but not HP ramping optimization, and (c) HP ramping optimization is applied but not filtering. We present tracking results at the hour change for these three cases in Fig. 11.9.

The tracking in case (a) is very good with a Mean Absolute Percentage Error (MAPE) of 0.22% (left plot). Without HP ramping optimization
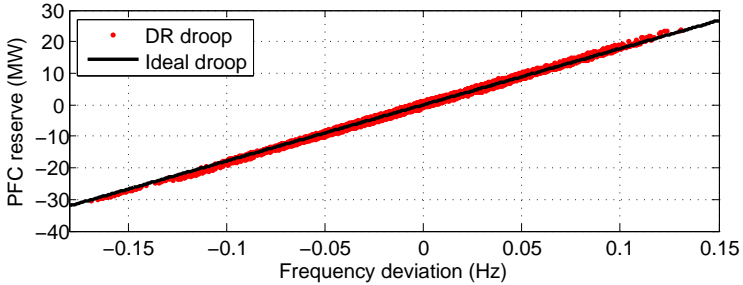
Figure 11.10: Comparison of the ideal PFC droop with the droop characteristic of the aggregation of refrigerators under PFC reserve provision.
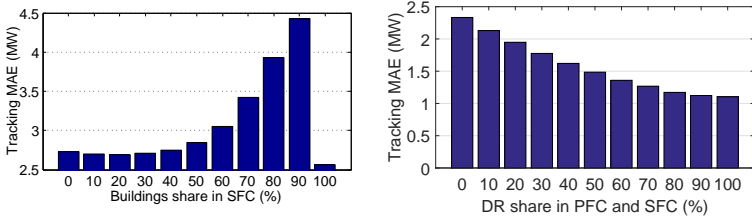
Figure 11.11: Mean Absolute Error while tracking the SFC signal. Left: Different shares of commercial buildings in SFC. Right: Different shares of refrigerators in PFC, and EWHs and commercial buildings in SFC.

(middle plot), large tracking errors occur for approximately 200 seconds, because the ramping constraints prevent the HPs from following the step changes in the consumption schedules and reserve capacities. Overall, the MAPE is 4.18% for the considered period. The effect of filtering on tracking quality is less pronounced than that of HP ramping optimization, but non-negligible. Without filtering (right plot), the setpoint power is more volatile, the tracking error is higher due to the HP latencies and ramping limits, and the MAPE is equal to 0.33%.

An important challenge when offering PFC and SFC reserves with distinct load aggregations is to avoid interactions that could endanger system stability. For example, one needs to make sure that accumulated errors from tracking the SFC signal do not translate into frequency deviations that cannot be handled by the PFC resources. In our simulations

the integration of HPs and EWHs in SFC had no noticeable effect on the performance of refrigerators in PFC. Figure 11.10 shows that the empirical droop characteristic of refrigerators is very close to the ideal one, and therefore refrigerators can efficiently substitute generators in PFC.[2]

We also simulated the system with DR shares in PFC and SFC from 0% to 100%, with a step of 10%, and present the tracking Mean Absolute Error (MAE) for each case in Fig. 11.11. The left plot corresponds to simulations where PFC is provided only by generators, and SFC by commercial buildings and generators. The right plot is for simulations where refrigerators and generators provide PFC, whereas EWHs, commercial buildings and generators provide SFC. Since the two plots correspond to slightly different simulation conditions, we compare the results of the two plots only qualitatively.

If commercial buildings is the only DR resource, then a low-pass filter with a cutoff period of 40 seconds is used to get the SFC signal for the buildings, whereas the filter's residual is sent to the generators. Of course, for a 100% share of buildings, the SFC signal is sent to the buildings without filtering. If only generators provide reserves ($\beta = 0\%$), the tracking MAE is 2.73 MW, which is slightly reduced to 2.56 MW if only buildings provide reserves ($\beta = 100\%$). Since a smoothed version of the SFC signal is sent to the buildings, one would expect a decrease in the SFC tracking error as the share of buildings increases. However, this is not the case in the left plot of Fig. 11.11. Although the tracking error on the buildings' side indeed decreases, the total error increases because the filter's residual signal is too fast for the generators.

The results are qualitatively different in the right plot of Fig. 11.11. If EWHs are used to track the filter's high-frequency residual instead of the generators, the SFC tracking error monotonically decreases as the DR shares increase. Similarly to Fig. 11.3, the error decreases rapidly for low DR shares and then saturates at its minimum value at a 80% share. This illustrates the benefit of pooling DR resources with large thermal inertia but slow response (HPs of commercial buildings), with DR resources with small inertia but fast response (EWHs).

---

[2]We observed that using a PFC deadband might lead to oscillatory behavior, if a single transfer function represents the aggregate dynamics of all generators connected to the system, as in our case study. Therefore, we did not model the PFC deadband in these simulations, in contrast to the simulations of Chapter 5, where no dynamic power system model was used.
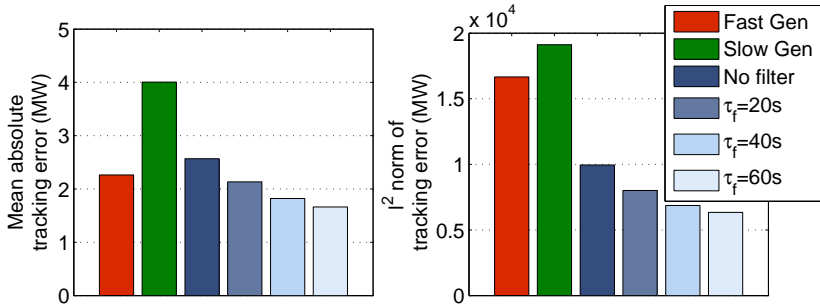
Figure 11.12: Comparison of DR resources against generators. Left: Mean absolute tracking error. Right: Tracking l^2 norm error.

11.4.4 Simulations with a Historical SFC signal

We observed that the SFC signal obtained by the two-area power system model is slightly smoother than the historical SFC signal. This is because in a real system there exist: (a) several areas and tie-lines; and (b) additional control loops, e.g., automatic voltage regulators and/or power system stabilizers. To test the control scheme's performance in a more realistic case, we simulate the tracking of a historical SFC signal from the Swiss power system in 2012 by DR resources. We consider a case with HPs only and no filtering, and three cases with both HPs and EWHs but with different filter time constants (τ_f). We compare the DR tracking errors against that of slow reheat steam turbine generators and faster generators (the model parameters are adopted from [71, p. 427]) in Fig. 11.12. The left plot shows the tracking MAEs, whereas the right plot shows the l^2 norm errors.

All the DR cases achieve much smaller errors compared with the slow generators, but when comparing them against the fast generators different error metrics lead to qualitatively different results. In terms of MAE, the performance of fast generators is similar to that of DR; however, in terms of the l^2 norm error, all DR cases clearly outperform the fast generators. This indicates that the proposed control scheme reduces the magnitude of the worst-case tracking errors due to the faster response times. Observe also that increasing the value of τ_f improves the performance; however, this comes at the cost of increasing the EWHs' reserve capacity.

11.5   Conclusion

In this chapter, we presented a frequency control scheme that allows large shares of heterogeneous thermal loads (residential refrigerators and EWHs, as well as HPs of commercial buildings) to participate in PFC and SFC. Refrigerators provide reserves in a decentralized fashion based on local frequency measurements, whereas SFC reserves are allocated among EWHs and HPs using techniques from robust optimization and signal filtering. The scheme accounts for the loads' electrical and thermal dynamics, it provides high-quality reserves without impact on user comfort, and it is compatible with classical power system control at the generation side. Dynamic simulations with a two-area power system model led to a number of interesting observations and recommendations, which we highlight in the following.

Due to their fast response times, refrigerators significantly reduce the magnitude of frequency excursions in case of large contingencies. In particular, it is shown that PFC by refrigerators improves the frequency response in low-inertia power systems (e.g., systems with large amounts of RES). HPs of commercial buildings can cope with biased SFC signals, but tracking errors increase in case of fast-moving signals due to HP delays and ramping limits. The fast components of the SFC signal can be followed by EWH aggregations without excessive communication requirements.

By pooling HPs and EWHs, smaller tracking errors compared with conventional generators can be achieved. Furthermore, in our simulations the performance of refrigerators in PFC did not deteriorate when HPs and EWHs were integrated in SFC. This indicates that the control scheme does not introduce any adverse interactions between distinct load aggregations. Overall, our results show that if load dynamics and constraints are understood and modeled properly, thermal loads can effectively substitute conventional generators in PFC and SFC.

Chapter 12

Conclusions and Outlook

12.1 Summary of this Thesis

This thesis investigated the potential of providing power system services with control of residential and commercial loads. The first part of the thesis focused on optimal control of loads to minimize the electricity cost of individual consumers (Chapter 2), and maximize the self-consumption of Photovoltaic (PV) energy (Chapter 3).

The second part of the thesis focused on provision of power system Ancillary Services (AS) with loads, and was organized in eight chapters. Chapter 4 provided a detailed literature review of existing load control approaches for AS. Chapters 5-8 considered aggregations of small residential Thermostatically Controlled Loads (TCLs) as an AS resource, whereas Chapters 9 and 10 investigated the potential from the Heating, Ventilation and Air Conditioning (HVAC) systems of commercial buildings.

More specifically, Chapter 5 presented a method that allows a large aggregation of refrigerators to provide Primary Frequency Control (PFC) in a decentralized way without real-time communication. Chapter 6 introduced rule-based control algorithms with different complexity and communication requirements to coordinate the aggregate power of an aggregation of Electric Water Heaters (EWHs) in order to track Secondary Frequency Control (SFC) signals. A method to provide SFC

without adverse effects on the Distribution Network (DN) was proposed in Chapter 7, thereby addressing a topic that is widely neglected in the literature. Chapter 8 presented a method to estimate the states of individual TCLs while providing SFC reserves in order to improve control performance and reduce communication cost.

Reserve scheduling, allocation and provision within an aggregation of commercial buildings was the topic of Chapter 9, which proposed a hierarchical control framework for this purpose. This control framework was used as a basis for an experimental demonstration of SFC provision from a commercial building test facility, which was presented in Chapter 10.

Finally, Chapter 11 presented a framework to combine different types of thermal loads in PFC and SFC. Some of the controllers developed in previous chapters were put together, and dynamic power system frequency studies with different load shares were performed.

12.2   Conclusions

The main conclusions and findings of this thesis are summarized as follows:

- Model Predictive Control (MPC) has a significant potential for energy cost reduction in residential households in the presence of dynamic end-customer electricity prices.

- Batteries are well-suited for PV self-consumption maximization in buildings, and even simple control algorithms can be effective.

- Large aggregations of refrigerators can provide PFC reserves more accurately than generators and in a decentralized fashion without real-time communication.

- EWH aggregations can track SFC signals accurately and with reasonable information feedback from the loads to the controller.

- With proper control design, TCLs can offer SFC without adverse effects on the DN, such as overvoltages and transformer or cable overloadings.

- Moving Horizon State Estimation (MHSE) can be used to estimate the states of individual TCLs in SFC applications resulting in a performance improvement.

- There exists a large potential for SFC from the HVAC systems of commercial buildings, and techniques from robust/stochastic optimization and MPC can be used to reliably estimate the reserve capacity.

- An experimental demonstration verified the technical feasibility of providing SFC with commercial buildings and showed the advantages of a hierarchical control design.

- Dynamic frequency studies showed that the frequency quality improves with large shares of heterogeneous loads in PFC and SFC, if they are properly coordinated and controlled.

12.3 Outlook

There are many avenues for future work on the topics addressed in this thesis. In the following, we briefly point out the most important questions that remain unanswered or require further investigation:

- The analysis in Chapter 5 assumed that each refrigerator measures the frequency locally without measurement error. However, this will not be the case in a real-world implementation due to the limited accuracy of the frequency meter and the frequency noise accumulation in the DN. It is therefore important to study the robustness of the decentralized controller to frequency measurement errors.

- As shown in Chapter 9, a priori reserve capacity estimation is possible within aggregations of commercial buildings. However, doing the same with TCLs is more involved due to the much larger aggregation size. Therefore, low-order but accurate aggregate TCL models are needed to enable a priori reserve scheduling. In addition, calibration of these models should be possible with available measurements.

- Although MHSE was shown to improve the SFC performance in Chapter 8, adopting this method in a real-world implementation is not straightforward because it is computationally intensive.

Therefore, a cost-benefit analysis and comparison with alternative estimation approaches from the literature is needed to identify the suitability of MHSE for practical implementation.

- The MHSE method of Chapter 8 accounts for communication networks with limited bandwidth. However, extensions are necessary to account also for communication delays and packet losses, which will likely occur in practice, and [169] took a step to this direction.

- Dynamic frequency studies should be performed considering communication delays and packet losses while providing SFC. Furthermore, the power system model can be extended to include more areas and voltage dynamics.

- The literature on load control for power system AS lacks experimental demonstrations and field tests, which are absolutely necessary to build confidence for widespread implementation. Chapter 10 of this thesis took one of the first steps to this direction, but we believe that more experimental work is needed. For example, experiments similar to those of Chapter 10 should be performed in larger commercial buildings (preliminary results from a follow-up work are reported in [201]). In addition, experiments with aggregations of commercial buildings are required to verify the scalability of the approach. More importantly, there is very little experimental demonstration of AS with large TCL aggregations, and much more work is needed in this area.

- A common concern for DR is profitability and realistic business cases. Various works have shown that the profit margin per participant in load control schemes is too small to attract a large interest. One solution would be to reduce the implementation cost of load control, for example by standardization that will allow new devices to be sold with the necessary controllers already integrated. An alternative would be to come up with new remuneration schemes or to support load control with state subsidies, similarly to the support of RES.

Appendix A

Proofs

A.1 Proof of Proposition 1

Proof. The estimated aggregate power at time step t after reserve activation is

$$\hat{P}_{\text{agg},t} = N_{\text{r,a}} \bar{P}_{\text{n}} \bar{S}_{\text{u}}(t) \ , \qquad (A.1)$$

where $N_{\text{r,a}}$ is the number of activated devices. Assuming independence of random variables $P_{\text{n},i}$, u_i and $N_{\text{s},i}$, the expected actual aggregate power is computed as

$$
\begin{aligned}
\bar{P}_{\text{agg},t} &= \mathbb{E}\Big[\sum\nolimits_{i=1}^{N_{\text{r,a}}} P_{i,t} \Big] = \sum\nolimits_{i=1}^{N_{\text{r,a}}} \mathbb{E}\left[P_{i,t}\right] \\
&= \sum_{i=1}^{N_{\text{r,a}}} \mathbb{E}\left[P_{\text{n},i}\left[1 + u_i\left[1 - \frac{t}{N_{\text{s},i}} \right]_+ \right] \right] \\
&= \sum_{i=1}^{N_{\text{r,a}}} \bar{P}_{\text{n}}\left[1 + \bar{u} \cdot \mathbb{E}\left[1 - \frac{t}{N_{\text{s},i}} \right]_+ \right] .
\end{aligned} \qquad (A.2)
$$

<u>*Part 1:*</u> If $t \leq N_{\text{s,min}}$, then $[\cdot]_+$ is redundant and from (A.2) it follows

$$
\begin{aligned}
\bar{P}_{\text{agg},t} &= \sum\nolimits_{i=1}^{N_{\text{r,a}}} \bar{P}_{\text{n}}\left[1 + \bar{u}\left[1 - \mathbb{E}\left[\frac{t}{N_{\text{s},i}} \right] \right] \right] \\
&\leq \sum\nolimits_{i=1}^{N_{\text{r,a}}} \bar{P}_{\text{n}}\left[1 + \bar{u}\left(1 - \frac{t}{\bar{N}_{\text{s}}} \right) \right] = N_{\text{r,a}} \bar{P}_{\text{n}} \bar{S}_{\text{u}}(t) = \hat{P}_{\text{agg},t}, \qquad (A.3)
\end{aligned}
$$

where Jensen's inequality $\phi(\mathbb{E}[N_{s,i}]) \leq \mathbb{E}[\phi(N_{s,i})]$ is invoked with $\phi(N_{s,i})$ $= t/N_{s,i}$, which is a convex function since $t, N_{s,i} \geq 0$. Therefore, it holds that $\bar{P}_{\text{agg},t} \leq \hat{P}_{\text{agg},t}$ for $t \leq N_{s,\min}$.

<u>Part 2:</u> For $t \geq \bar{N}_s$, we obtain $\hat{P}_{\text{agg},t} = 0$ from (5.10) and (A.1). Since the expected aggregate power $\bar{P}_{\text{agg},t}$ is non-negative, $\hat{P}_{\text{agg},t}$ is clearly a lower bound.

For $N_{s,\min} < t \leq \bar{N}_s$, the estimated aggregate power is also given by (A.2). Assume that $N_{s,i}$ follows the uniform probability distribution $N_{s,i} \sim \mathbb{U}[N_{s,\min}, N_{s,\max}]$. For a given t, the probability that a randomly selected refrigerator contributes to the aggregate startup dynamics is equal to $\mathbb{P}(N_{s,i} \geq t) = (N_{s,\max} - t) / (N_{s,\max} - N_{s,\min})$. Denote by $\tilde{N}_{s,i}$ the random startup duration that follows the uniform distribution $\tilde{N}_{s,i} \sim [t, N_{s,\max}]$. The expectation from the last step of (A.2) can be written as

$$\mathbb{E}\left[1 - \frac{t}{N_{s,i}}\right]_+ = \mathbb{E}\left[1 - \frac{t}{\tilde{N}_{s,i}}\right] \cdot \frac{N_{s,\max} - t}{N_{s,\max} - N_{s,\min}} \tag{A.4a}$$

$$\leq \left(1 - \frac{t}{\mathbb{E}[\tilde{N}_{s,i}]}\right) \cdot \frac{N_{s,\max} - t}{N_{s,\max} - N_{s,\min}} \tag{A.4b}$$

$$= \left[1 - \frac{t}{0.5 \cdot (t + N_{s,\max})}\right] \cdot \frac{N_{s,\max} - t}{N_{s,\max} - N_{s,\min}} \tag{A.4c}$$

$$= \frac{(N_{s,\max} - t)^2}{(N_{s,\max} - N_{s,\min}) \cdot (N_{s,\max} + t)} . \tag{A.4d}$$

We used Jensen's inequality in (A.4b) and the fact that $\tilde{N}_{s,i}$ follows the uniform distribution $\tilde{N}_{s,i} \sim [t, N_{s,\max}]$ in step (A.4c).

The inequality $\hat{P}_{\text{agg},t} \geq \bar{P}_{\text{agg},t}$ holds if and only if

$$1 - \frac{t}{\bar{N}_s} \geq \mathbb{E}\left[1 - \frac{t}{N_{s,i}}\right]_+ \Leftrightarrow \tag{A.5}$$

$$\frac{N_{s,\max} + N_{s,\min} - 2t}{N_{s,\max} + N_{s,\min}} \geq \frac{(N_{s,\max} - t)^2}{(N_{s,\max} - N_{s,\min}) \cdot (N_{s,\max} + t)} , \tag{A.6}$$

where we used the fact that $\bar{N}_s = 0.5 \cdot (N_{s,\max} + N_{s,\min})$. By tedious but simple algebraic operations (A.6) can be equivalently written as

$$(N_{s,\min} - 3N_{s,\max}) \cdot t^2 + (N_{s,\max}^2 + 4N_{s,\min}N_{s,\max} - N_{s,\min}^2) \cdot t$$
$$- (N_{s,\min}^2 N_{s,\max} + N_{s,\min} N_{s,\max}^2) \geq 0 . \tag{A.7}$$

The quadratic form in (A.7) has two real solutions t_1 and $t_{\lim}$. The first one is $t_1 = N_{s,\min}$ by inspection. For the second one, we use Vieta's formula $t_1 \cdot t_{\lim} = c/a$ with $a = N_{s,\min} - 3N_{s,\max}$ and $c = -(N_{s,\min}^2 N_{s,\max} + N_{s,\min} N_{s,\max}^2)$ and find

$$t_{\lim} = \frac{N_{s,\max}(N_{s,\min} + N_{s,\max})}{3N_{s,\max} - N_{s,\min}} \ . \tag{A.8}$$

The quadratic inequality (A.7) holds for $t \in [t_1, t_{\lim}]$. Therefore, when $t > N_{s,\min}$ the inequality $\hat{P}_{agg,t} \geq \bar{P}_{agg,t}$ holds for $t \leq t_{\lim}$, which proves the claim of Proposition 1. ∎

A.2 Proof of Proposition 2

Proof. <u>*Part 1:*</u> For $0 \leq t \leq N_{ev}+1$ using $\bar{T}_0 = \bar{T}_{nom}$ and (5.48), equation (5.47) gives

$$\bar{T}_t = \bar{T}_{nom} - \gamma\delta \cdot \sum_{k=0}^{t-1} \lambda^k = \bar{T}_{nom} - \gamma\delta \cdot \frac{1 - \lambda^{t-1}}{1 - \lambda} \ , \tag{A.9}$$

where a standard property of geometric series is invoked. The minimum of (A.9) is obtained for $t = N_{ev}+1$, and thus if $\bar{T}_{nom} - \varepsilon \leq \bar{T}_t \leq \bar{T}_{nom} + \varepsilon$ holds for $t = N_{ev} + 1$ then it holds $\forall t$. Assuming $\delta > 0$ without loss of generality, it should hold

$$\bar{T}_{t=N_{ev}+1} \geq \bar{T}_{nom} - \varepsilon \quad \Rightarrow \quad -\gamma\delta \cdot \frac{1 - \lambda^{N_{ev}}}{1 - \lambda} \geq -\varepsilon \ , \tag{A.10}$$

which is equivalent to (5.49).

<u>*Part 2:*</u> For $t \geq N_{ev} + 1$, $\Delta f_t = 0$ and the mean temperature recovers towards $\bar{T}_{nom}$. We reset the time index such that $t' = 0 \Leftrightarrow t = N_{ev} + 1$. The new initial condition is $\bar{T}_{(t'=0)} = \bar{T}_{(t=N_{ev}+1)}$ and is computed from (A.9). Equation (5.47) gives

$$\bar{T}_{t'} = \bar{T}_{nom} - \gamma\delta \cdot \frac{1 - \lambda^{N_{ev}}}{1 - \lambda} \cdot \lambda^{t'} \ . \tag{A.11}$$

At time $t' = N_{rec}$, $\bar{T}_{nom} - \epsilon \leq \bar{T}_{N_{rec}} \leq \bar{T}_{nom} + \epsilon$ must hold. Assuming $\delta > 0$ without loss of generality, we get

$$\bar{T}_{t'=N_{rec}} \geq \bar{T}_{nom} - \epsilon \Rightarrow -\gamma\delta \cdot \lambda^{N_{rec}} \cdot \frac{1 - \lambda^{N_{ev}}}{1 - \lambda} \geq -\epsilon \ , \tag{A.12}$$

which is equivalent to (5.50). ∎

A.3   Proof of Proposition 3

Proof. The change in temperature limits of a randomly selected refrigerator i can be expressed as

$$\Delta T_{\text{lim},i,t} = \sum_{k=0}^{t-1} \Delta T_{\text{res}} \cdot \zeta_k \cdot z_k \ , \tag{A.13}$$

where ζ_k is equal to 1 or -1 depending on the sign of $\Delta T_{\text{lim},k}$, and z_k is the Bernoulli random variable

$$f_{\text{z}}(z_k) = \begin{cases} \varrho_k, & \text{if } z_k = 1 \\ 1 - \varrho_k, & \text{if } z_k = 0 \end{cases} . \tag{A.14}$$

Using the fact $\mathbb{E}[z_k] = \varrho_k$, the mean value is computed as

$$\mathbb{E}[\Delta T_{\text{lim},i,t}] = \sum_{k=0}^{t-1} \Delta T_{\text{res}} \zeta_k \mathbb{E}[z_k] = \sum_{k=0}^{t-1} \Delta T_{\text{res}} \zeta_k \varrho_k$$
$$= \sum_{k=0}^{t-1} \zeta_k |\Delta T_{\text{lim},k}| = \sum_{k=0}^{t-1} \Delta T_{\text{lim},k} \ . \tag{A.15}$$

The variance is computed as

$$\text{Var}[\Delta T_{\text{lim},i,t}] = \text{Var}\Big[\sum_{k=0}^{t-1} \Delta T_{\text{res}} \cdot \zeta_k \cdot z_k\Big] \tag{A.16a}$$

$$= \sum_{k=0}^{t-1} \Delta T_{\text{res}}^2 \cdot \zeta_k^2 \cdot \text{Var}[z_k] = \sum_{k=0}^{t-1} \Delta T_{\text{res}}^2 \cdot \varrho_k \cdot (1 - \varrho_k) \tag{A.16b}$$

$$= \sum_{k=0}^{t-1} \Delta T_{\text{res}}^2 \cdot \frac{|\Delta T_{\text{lim},k}|}{\Delta T_{\text{res}}} \cdot \Big(1 - \frac{|\Delta T_{\text{lim},k}|}{\Delta T_{\text{res}}}\Big) \tag{A.16c}$$

$$= \sum_{k=0}^{t-1} |\Delta T_{\text{lim},k}| \cdot \big(\Delta T_{\text{res}} - |\Delta T_{\text{lim},k}|\big) \ , \tag{A.16d}$$

where the random variables z_k and z_j are assumed to be uncorrelated $\forall k \neq j$ in (A.16a), the fact that $\text{Var}[z_k] = \varrho_k \cdot (1 - \varrho_k)$ is used in (A.16b) since z_k is a Bernoulli random variable, and the definition of ϱ_k from (5.52) is used in (A.16c). ∎

A.4   Proof of Proposition 4

Proof. Denote by E_{op}^{d} and E_{cl}^{d} the refrigerator's daily energy consumption with and without door openings, respectively. Consider a refrigerator subject to door openings and denote by $\tilde{T}_{\text{op}}$ and $\bar{T}_{\text{cl}}$ the time-averaged temperature while the door is open and closed, respectively.

The refrigerator operates on average $\mu_{\mathrm{op}} \cdot \mu_{\mathrm{d}}$ seconds with open door, and $N^{\mathrm{d}} - \mu_{\mathrm{op}} \cdot \mu_{\mathrm{d}}$ seconds with closed door. The refrigerator's time constant with closed door is $\alpha = 1/(RC)$, whereas with open door it is $\alpha_{\mathrm{op}} = 1/(R_{\mathrm{op}}C)$.

The refrigerator's energy consumption is equal to the integral of instantaneous thermal losses over time, because its temperature is regulated. Assuming that $\bar{T}_{\mathrm{cl}}$ and $\bar{T}_{\mathrm{op}}$ are constant, the daily energy consumption with and without door openings can be approximated with (5.1) and (5.2) in the following way

$$E_{\mathrm{cl}}^{\mathrm{d}} \approx \frac{\alpha}{\beta} \cdot (T_{\mathrm{a}} - \bar{T}_{\mathrm{cl}}) \cdot N^{\mathrm{d}} \tag{A.17}$$

$$E_{\mathrm{op}}^{\mathrm{d}} \approx \frac{\alpha}{\beta} \cdot (T_{\mathrm{a}} - \bar{T}_{\mathrm{cl}}) \cdot (N^{\mathrm{d}} - \mu_{\mathrm{op}}\mu_{\mathrm{d}}) + \frac{\alpha_{\mathrm{op}}}{\beta} \cdot (T_{\mathrm{a}} - \bar{T}_{\mathrm{op}}) \cdot \mu_{\mathrm{op}}\mu_{\mathrm{d}}, \tag{A.18}$$

whereas $E_{\mathrm{op}}^{\mathrm{d}}$ and $E_{\mathrm{cl}}^{\mathrm{d}}$ are related according to

$$E_{\mathrm{op}}^{\mathrm{d}} = (1 + \xi) \cdot E_{\mathrm{cl}}^{\mathrm{d}} . \tag{A.19}$$

Combining (A.17), (A.18) and (A.19), we get

$$\frac{\alpha}{\beta} \cdot (T_{\mathrm{a}} - \bar{T}_{\mathrm{cl}}) \cdot (N^{\mathrm{d}} - \mu_{\mathrm{op}}\mu_{\mathrm{d}}) + \frac{\alpha_{\mathrm{op}}}{\beta} \cdot (T_{\mathrm{a}} - \bar{T}_{\mathrm{op}}) \cdot \mu_{\mathrm{op}}\mu_{\mathrm{d}} =$$
$$(1 + \xi) \cdot \frac{\alpha}{\beta} \cdot (T_{\mathrm{a}} - \bar{T}_{\mathrm{cl}}) \cdot N^{\mathrm{d}} , \tag{A.20}$$

which when solved for α_{op} gives

$$\alpha_{\mathrm{op}} = \alpha \cdot \frac{T_{\mathrm{a}} - \bar{T}_{\mathrm{cl}}}{T_{\mathrm{a}} - \bar{T}_{\mathrm{op}}} \cdot \left(1 + \frac{N^{\mathrm{d}}}{\mu_{\mathrm{op}} \cdot \mu_{\mathrm{d}}} \cdot \xi\right) \quad \Rightarrow \tag{A.21}$$

$$\alpha_{\mathrm{op}} \geq \alpha \cdot \left(1 + \frac{N^{\mathrm{d}}}{\mu_{\mathrm{op}} \cdot \mu_{\mathrm{d}}} \cdot \xi\right) , \tag{A.22}$$

because $\bar{T}_{\mathrm{cl}} \leq \bar{T}_{\mathrm{op}}$. Inequality (5.56) now directly follows from (A.22) using $\alpha = (1/RC)$ and $\alpha_{\mathrm{op}} = (1/R_{\mathrm{op}}C)$. ∎

A.5 Derivation of Equation (9.9)

We derive (9.9) by solving the non-homogeneous second order ODE (9.8). The solution $y(t)$ is the sum of the solution of the respective

homogeneous ODE $y_h(t)$ and a specific solution of the non-homogeneous
ODE $y_n(t)$

$$y(t) = y_h(t) + y_n(t) \; . \tag{A.23}$$

The homogeneous solution can be obtained directly using the two distinct real poles λ_1 and λ_2 of the ODE

$$y_h(t) = A \cdot e^{\lambda_1 t} + B \cdot e^{\lambda_2 t} \; . \tag{A.24}$$

The following specific solution of the non-homogeneous ODE can be obtained applying a Laplace transformation

$$y_n(t) = u_{s,k} \cdot \left[1 - e^{-\sigma t} \cdot \cosh\left(\alpha t\right) - \frac{\sigma}{\alpha} \cdot e^{-\sigma t} \cdot \sinh\left(\alpha t\right) \right] \; . \tag{A.25}$$

The parameters A and B in (A.24) can be determined by using the initial conditions $y(0) = u_{s,k-1}$ and $\dot{y}(0) = 0$. It is easy to verify that $y_n(0) = 0$ and $\dot{y}_n(0) = 0$, therefore it is sufficient to impose

$$y_h(0) = 0 \quad \Rightarrow u_{s,k-1} = A + B \tag{A.26}$$
$$\dot{y}_h(0) = 0 \quad \Rightarrow 0 = A \cdot \lambda_1 + B \cdot \lambda_2 \; . \tag{A.27}$$

Solving the system of equations (A.26) and (A.27) we result in the following expressions

$$A = u_{s,k-1} \cdot \frac{\lambda_2}{\lambda_2 - \lambda_1} \tag{A.28}$$

$$B = -u_{s,k-1} \cdot \frac{\lambda_1}{\lambda_2 - \lambda_1} \; . \tag{A.29}$$

Therefore, the final solution of the ODE is

$$\begin{aligned} y(t) = & u_{s,k-1} \cdot \frac{\lambda_2}{\lambda_2 - \lambda_1} \cdot e^{\lambda_1 t} - u_{s,k-1} \cdot \frac{\lambda_1}{\lambda_2 - \lambda_1} \cdot e^{\lambda_2 t} + \\ & u_{s,k} \cdot \left[1 - e^{-\sigma t} \cdot \cosh\left(\alpha t\right) - \frac{\sigma}{\alpha} \cdot e^{-\sigma t} \cdot \sinh\left(\alpha t\right) \right] \; . \end{aligned} \tag{A.30}$$

The lost thermal energy $L_2(t)$, i.e., the area between $u_{th,k}^l(t)$ and $u_{s,k}$,

can be now written as follows

$$
L_2(t) = \int_0^t \left[u_{\text{th},k}^l(\tau) - u_{\text{s},k} \right] d\tau = \int_0^t \left[y(\tau) - u_{\text{s},k} \right] d\tau
$$

$$
= \int_0^t u_{\text{s},k-1} \cdot \frac{\lambda_2}{\lambda_2 - \lambda_1} \cdot e^{\lambda_1 \tau} d\tau - \int_0^t u_{\text{s},k-1} \cdot \frac{\lambda_1}{\lambda_2 - \lambda_1} \cdot e^{\lambda_2 \tau} d\tau
$$

$$
- \int_0^t u_{\text{s},k} e^{-\sigma\tau} \cdot \cosh(\alpha\tau) d\tau - \int_0^t u_{\text{s},k} \frac{\sigma}{\alpha} \cdot e^{-\sigma\tau} \cdot \sinh(\alpha\tau) d\tau
$$

$$
= u_{\text{s},k-1} \cdot \left[\frac{\lambda_2}{\lambda_1 \cdot (\lambda_2 - \lambda_1)} \cdot (e^{\lambda_1 t} - 1) - \frac{\lambda_1}{\lambda_2 \cdot (\lambda_2 - \lambda_1)} \cdot (e^{\lambda_2 t} - 1) \right]
$$

$$
- \frac{u_{\text{s},k}}{\sigma^2 - \alpha^2} \cdot \left[-e^{-\sigma t} \cdot [\sigma \cdot \cosh(\alpha t) + \alpha \cdot \sinh(\alpha t)] + \sigma \right]
$$

$$
- \frac{u_{\text{s},k} \cdot \sigma}{\alpha \cdot (\sigma^2 - \alpha^2)} \cdot \left[-e^{-\sigma t} [\alpha \cosh(\alpha t) + \sigma \sinh(\alpha t)] + \alpha \right] . \quad \text{(A.31)}
$$

After tedious but simple mathematical operations (A.31) can be written in the form $a(t)u_{\text{s},k-1} + b(t)u_{\text{s},k}$ with the $a(t)$ and $b(t)$ given in (9.10) and (9.11).

A.6 Proof of Lemma 1

Proof. The monotonicity and convexity can be verified by inspection of the fan curve in Fig. 10.9. For a formal proof, it is sufficient to show that the first and second order derivatives $f'(u)$ and $f''(u)$ are non-negative $\forall u$ in the domain of f. From (10.9) we have

$$
f'(u) = 3\alpha_3 u^2 + 2\alpha_2 u + \alpha_1 \quad \text{(A.32)}
$$

$$
f''(u) = 6\alpha_3 u + 2\alpha_2 . \quad \text{(A.33)}
$$

The discriminant of $f'(u)$ has no real roots, and therefore $f'(u) \geq 0 \ \forall u$ because $\alpha_3 \geq 0$. Furthermore, $f''(u) \geq 0$ holds if $u \geq -\frac{\alpha_2}{3\alpha_3}$, which is satisfied for the parameters of Table 10.7. ∎

A.7 Proof of Lemma 2

Proof. From the definition of Δu_k in (10.18) and (10.17) we get

$$
\min_{w_k} (u_k + \Delta u_k) = \min_{w_k \in [-w_{\text{lim}}, 0)} \left[f^{-1} (P_k + w_k R_{\text{u},k}) \right] . \quad \text{(A.34)}
$$

Due to monotonicity of f, it holds

$$\operatorname{argmin}\left[f^{-1}\left(P_k + w_k R_{u,k}\right)\right] = \operatorname{argmin}\left(P_k + w_k R_{u,k}\right) = -w_{\lim}.$$
$$\text{(A.35)}$$

Substituting the minimizer $-w_{\lim}$ in (A.34) we get (10.22). Equation (10.24) is a special case of (10.22) derived as

$$f^{-1}\left(P_k - w_{\lim}R_{u,k}\right) = f^{-1}\left[f(u_k) - R_{u,k}\right] =$$
$$f^{-1}\left[f(u_k) - f(u_k) + f(u_k - r_{d,k})\right] = u_k - r_{d,k} \;, \qquad \text{(A.36)}$$

where (10.13) is used. The maximization case for (10.23) and (10.25) can be proved analogously but the proof is omitted for brevity. ∎

A.8   Proof of Proposition 5

Proof. It is sufficient to show that: (i) the set of input constraints (10.21d) and (10.21e) is equivalent to (10.19c); and (ii) the set of state constraints (10.21b), (10.21c), (10.21f) and (10.21g) is equivalent to the set of constraints (10.19b) and (10.19d).

Input constraints: The equivalence follows directly from (10.24) and (10.25).

State constraints: We first write (10.19b) as

$$x_{k+1} = Ax_k + (B_u T_{s,k} + B_{xu} x_k) \cdot (u_k + \Delta u_k) + B_v v_k \;. \qquad \text{(A.37)}$$

Constraint (10.19d) is applied only to the first state $T_{r,k} = Cx_k$ of the state vector x_k, where $C = [1 \quad 0]$ is the output matrix. Observing that $CB_u T_{s,k} = bT_{s,k}$ and $CB_{xu} x_k = -bT_{r,k}$, constraint (10.19d) can be written as

$$\min_{w_k}\left[CAx_k + b(T_{s,k} - T_{r,k}) \cdot (u_k + \Delta u_k) + CB_v v_k\right] \geq x_{\min,k} \qquad \text{(A.38)}$$

$$\max_{w_k}\left[CAx_k + b(T_{s,k} - T_{r,k}) \cdot (u_k + \Delta u_k) + CB_v v_k\right] \leq x_{\max,k} \;. \qquad \text{(A.39)}$$

Due to assumption 1, we have $b \cdot (T_{s,k} - T_{r,k}) \leq 0$; thus, the left hand side of (A.38) is minimized when $u_k + \Delta u_k$ is maximized. From (10.23), this is achieved when $u_k + \Delta u_k = f^{-1}\left(P_k + w_{\lim}R_{d,k}\right)$ holds. Denote by $\underline{x}_k$ the minimum state trajectory at time step k. The time evolution of $\underline{x}_k$ is therefore obtained by $\underline{x}_{k+1} = CA\underline{x}_k + b(T_{s,k} - C\underline{x}_k) \cdot$

$f^{-1}\left(P_k + w_{\lim}R_{\mathrm{d},k}\right) + CB_v v_k$, which is essentially (10.21c). Thus, if (10.21f) holds (if $x_{\min,k} \leq \underline{x}_k$), then the left hand side of (10.19d) also holds because $x_k \leq \underline{x}_k \ \forall w_k$.

Analogous arguments can be used to show that (A.39) results in (10.21b) and (10.21g), which are equivalent to the right hand side of (10.19d), but this is omitted for brevity. ∎

A.9 Proof of Proposition 6

Proof. Using definition (10.13) and the fact that $f(u)$ is convex we get

$$
\begin{aligned}
P_k - w_{\lim}R_{\mathrm{u},k} &= f(u_k) - w_{\lim}\left[f(u_k) - f(u_k - r_{\mathrm{d},k})\right] \\
&= (1 - w_{\lim})f(u_k) + w_{\lim}f(u_k - r_{\mathrm{d},k}) \\
&\geq f\left[(1 - w_{\lim})u_k + w_{\lim}(u_k - r_{\mathrm{d},k})\right] \ . \qquad \text{(A.40)}
\end{aligned}
$$

Using the monotonicity of $f(u)$ on (A.40) we get

$$
\begin{aligned}
f^{-1}\left(P_k - w_{\lim}R_{\mathrm{u},k}\right) &\geq (1 - w_{\lim})u_k + w_{\lim}(u_k - r_{\mathrm{d},k}) \\
&= u_k - w_{\lim}r_{\mathrm{d},k} \ . \qquad \text{(A.41)}
\end{aligned}
$$

$\overline{x}_k^* \geq \overline{x}_k^\star$ is now obtained by combining (A.41), (10.21b), (10.26b), and using the same arguments related to $b(T_{\mathrm{s},k} - T_{\mathrm{r},k}) \leq 0$ as in the proof of Proposition 5.

Similarly, one can show that $u_k + w_{\lim}r_{\mathrm{u},k} \leq f^{-1}\left(P_k + w_{\lim}R_{\mathrm{d},k}\right)$ holds and so that $\underline{x}_k^* \geq \underline{x}_k^\star$ also holds. Fig. 10.10 provides a graphical interpretation of (A.41). ∎

A.10 Proof of Proposition 7

Proof. The proof follows directly by rewriting (10.21b) and (10.21c) using (10.24), (10.25) and (A.36) from Lemma 2. ∎

Appendix B

Model Parameters

B.1 Model Parameters of Chapter 2

Symbol	Value
C_f	$4.550 \cdot 10^7$ W/K
C_wr	$5.156 \cdot 10^6$ W/K
C_ws	$2.006 \cdot 10^5$ W/K
c_w	$4.185 \cdot 10^3$ J/(Kg K)
c_0	5.593
c_1	0.0569 K^{-1}
c_2	-0.0661 K^{-1}
k_wf	$1.160 \cdot 10^3$ W/K
$\dot{m}_\mathrm{w}$	0.266 kg/s

B.2 Parameters of Distribution Network of Chapter 7

The resistance, reactance and branches' line length of the DN of Fig. 7.2 are based on the benchmark DN of [2]. However, the capacitance pa-

Table B.2: Battery model parameters of Section 2.2.4

Symbol	Value	Symbol	Value
b_1	0.315	$SOC_{\min}$	40%
b_2	1.24 (1/h)	α_1	1505.89
$E_{\max}$	5 kWh	α_2	9687.24
n_{c}	0.8624	α_3	4.90
n_{d}	0.8532	α_4	9845.09
n_{conv}	0.93	α_5	6.59

Table B.3: PV model parameters of Section 2.2.5

Symbol	Value	Symbol	Value
A	1	q	$1.6 \cdot 10^{-19}$ C
G_{s}	1000 W/m^2	T_{NOCT}	47°C
I_{p}	34.478 V	T_{s}	298.15 K
k_{b}	$1.38 \cdot 10^{-23}$ J/K	U_{p}	4.65 A
K_1	-10.01 V	ΔI_{p}	$2.42 \cdot 10^{-3}$
K_2	-1850.80 V	ΔU_{p}	$-4.58 \cdot 10^{-3}$

rameters are modified based on typical values from Italian DNs reported in [202]. The final DN parameters are given in Table B.4.

The network's load is allocated to the buses analogously to [2], while typical residential and industrial load profiles are adopted from [203].

Table B.4: Parameters of distribution network of Chapter 7

Bus From	Bus to	Resistance (Ω/km)	Reactance (Ω/km)	Capacitance (nF/km)	Length (km)
1	2	0.579	0.367	9	2.82
2	3	0.164	0.113	500	4.42
3	4	0.262	0.121	420	0.61
4	5	0.354	0.129	350	0.56
5	6	0.336	0.126	350	1.54
7	8	0.294	0.123	350	1.67
8	9	0.339	0.13	350	0.32
9	10	0.399	0.133	350	0.77
10	11	0.367	0.133	350	0.33
3	8	0.172	0.115	420	1.3

Bibliography

[1] F. Borrelli, A. Bemporad, and M. Morari, *Predictive control for linear and hybrid systems.* Lecture notes available online at http://www.mpc.berkeley.edu/mpc-course-material, 2015.

[2] K. Rudion, A. Orths, A. Styczynski, and K. Strunz, "Design of benchmark of medium voltage distribution network for investigation of DG integration," in *IEEE PES GM*, Montreal, Canada, June 2006.

[3] J. Mathieu, S. Koch, and D. Callaway, "State estimation and control of electric loads to manage real-time energy imbalance," *IEEE Transactions on Power Systems*, vol. 28, no. 1, pp. 430–440, 2013.

[4] J. L. Mathieu and D. S. Callaway, "State estimation and control of heterogenous thermostatically controlled loads for load following," in *Proceedings of the 45th Hawaii International Conference on Sytem Sciences*, Hawaii, USA, January 2012, pp. 2002–2011.

[5] Y. Makarov, C. Loutan, J. Ma, and P. De Mello, "Operational impacts of wind generation on California power systems," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 1039–1050, 2009.

[6] F. Schweppe, R. Tabors, J. Kirtley Jr, H. Outhred, F. Pickel, and A. Cox, "Homeostatic utility control," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-99, no. 3, pp. 1151–1163, 1980.

[7] D. Callaway and I. Hiskens, "Achieving controllability of electric loads," *Proceedings of the IEEE*, vol. 99, no. 1, pp. 184–199, 2011.

[8] A. Ulbig and G. Andersson, "Towards variable end-consumer electricity tariffs reflecting marginal costs: A benchmark tariff," in *7th International European Energy Market Conference (EEM)*, Madrid, Spain, January 2010, pp. 1–6.

[9] J. E. Braun, "Load control using building thermal mass," *Journal of solar energy engineering*, vol. 125, no. 3, pp. 292–301, 2003.

[10] ——, "Reducing energy costs and peak electrical demand through optimal control of building thermal storage," *ASHRAE Transactions*, vol. 96, no. 2, pp. 876–887, 1990.

[11] F. B. Morris, J. E. Braun, and S. J. Treado, "Experimental and simulated performance of optimal control of building thermal storage," *ASHRAE Transactions*, vol. 100, no. 1, pp. 402–414, 1994.

[12] J. E. Braun, "A comparison of chiller-priority, storage-priority, and optimal control of an ice-storage system," *ASHRAE Transactions*, vol. 98, no. 1, pp. 893–902, 1992.

[13] G. P. Henze, R. H. Dodier, and M. Krarti, "Development of a predictive optimal controller for thermal energy storage systems," *HVAC&R Research*, vol. 3, no. 3, pp. 233–264, 1997.

[14] A. J. Conejo, J. M. Morales, and L. Baringo, "Real-time demand response model," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 236–242, 2010.

[15] D. J. Livengood, "The energy box: Comparing locally automated control strategies of residential electricity consumption under uncertainty," Ph.D. dissertation, Massachusetts Institute of Technology (MIT), 2011.

[16] F. Oldewurtel, A. Parisio, C. N. Jones, M. Morari, D. Gyalistras, M. Gwerder, V. Stauch, B. Lehmann, and K. Wirth, "Energy efficient building climate control using stochastic model predictive control and weather predictions," in *American Control Conference*, Baltimore, USA, June 2010.

[17] Y. Ma, F. Borrelli, B. Hencey, B. Coffey, S. Bengea, and P. Haves, "Model predictive control for the operation of building cooling systems," *IEEE Transactions on Control Systems Technology*, vol. 20, no. 3, pp. 797–803, 2012.

[18] Y. Zong, D. Kullmann, A. Thavlov, O. Gehrke, and H. W. Bindner, "Application of model predictive control for active load management in a distributed power system with high wind penetration," *IEEE Transactions on Smart Grid*, vol. 3, no. 2, pp. 1055–1062, 2012.

[19] Y. Guo, M. Pan, and Y. Fang, "Optimal power management of residential customers in the smart grid," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 9, pp. 1593–1606, 2012.

[20] E. Vrettos, K. Lai, F. Oldewurtel, and G. Andersson, "Predictive control of buildings for demand response with dynamic day-ahead and real-time prices," in *European Control Conference (ECC), Zurich, Switzerland*, October 2013.

[21] E. Vrettos, S. Koch, and G. Andersson, "Load frequency control by aggregations of thermally stratified electric water heaters," in *IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe)*, Berlin, Germany, October 2012.

[22] E. Vrettos and S. Papathanassiou, "Operating policy and optimal sizing of a high penetration RES-BESS system for small isolated grids," *IEEE Transactions on Energy Conversion*, vol. 26, no. 3, pp. 744–756, 2011.

[23] D. Gyalistras and M. Gwerder, "Use of weather and occupancy forecasts for optimal building climate control (OptiControl): Two years progress report," Terrestrial Systems Ecology, ETH Zurich, and Building Technologies Division, Siemens, Tech. Rep., 2010.

[24] F. Oldewurtel, "Stochastic model predictive control for energy efficient building climate control," Ph.D. dissertation, ETH Zürich, 2011.

[25] B. Lehmann, D. Gyalistras, M. Gwerder, K. Wirth, and S. Carl, "Intermediate complexity model for Model Predictive Control of Integrated Room Automation," *Energy and Buildings*, vol. 58, pp. 250–262, 2013.

[26] D. Sturzenegger, D. Gyalistras, M. Gwerder, C. Sagerschnig, M. Morari, and R. Smith, "Model predictive control of a Swiss office building," in *REHVA World Congress Clima*, Prague, Czech Republic, June 2013.

[27] C. Verhelst, D. Axehill, C. Jones, and L. Helsen, "Impact of the cost function in the optimal control formulation for an air-to-water heat pump system," in *8th International Conference on System Simulation in Buildings (SSB), Liege, Belgium*, 2010.

[28] M. H. Nehrir, R. Jia, D. A. Pierre, and D. Hammerstrom, "Power management of aggregate electric water heater loads by voltage control," in *Proceedings of the IEEE PES General Meeting*, Tampa, USA, June 2007, pp. 1–6.

[29] L. Paull, H. Li, and L. Chang, "A novel domestic electric water heater model for a multi-objective demand side management program," *Electric Power Systems Research*, vol. 80, no. 12, pp. 1446–1451, 2010.

[30] J. Kondoh, N. Lu, and D. Hammerstrom, "An evaluation of the water heater load potential for providing regulation service," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1309–1316, 2011.

[31] A. H. Fanney and B. P. Dougherty, "The thermal performance of residential electric water heaters subjected to various off-peak schedules," *Journal of Solar Energy Engineering*, vol. 118, pp. 73–80, 1996.

[32] I. Dinçer and M. A. Rosen, *Thermal Energy Storage: Systems and Applications.* John Wiley and Sons, 2002.

[33] W. F. Ames, *Numerical Methods for Partial Differential Equations.* Computer Science and Scientific Computing, 1992.

[34] J. M. Varah, "Stability restrictions on second order, three level finite difference schemes for parabolic equations," *SIAM Journal on Numerical Analysis*, vol. 17, no. 2, pp. 300–309, 1980.

[35] W. M. Healy, J. D. Lutz, and A. B. Lekov, "Variability in energy factor test results for residential electric water heaters," *HVAC&R Research*, vol. 9, no. 4, pp. 435–449, 2003.

[36] J. Manwell and J. McGowan, "Lead acid battery storage model for hybrid energy systems," *Solar Energy*, vol. 50, pp. 399–405, 1993.

[37] M. R. Jongerden and B. R. Haverkort, *Battery modeling*, 2008.

[38] J. F. Manwell, A. Rogers, G. Hayman, C. T. Avelar, J. G. Mc-Gowan, U. Abdulwahid, and K.Wu, *Hybrid2 - a hybrid system simulation model - theory manual*. Renewable Energy Research Laboratory, University of Massachusetts, 2006.

[39] H. Fakham, P. Degobert, and B. François, "Control system and power management for a PV based generation unit including batteries," in *International Aegean Conference on Electrical Machines and Power Electronics*, Bodrum, Turkey, September 2007, pp. 141–146.

[40] H. S. Rauschenbach, "Solar cell array design handbook - the principles and technology of photovoltaic energy conversion," 1980.

[41] R. Ross Jr., "Flat-plate photovoltaic array design optimization," in *14th Photovoltaic Specialists Conference*, San Diego, USA, January 1980, pp. 112–1132.

[42] E. Camacho and C. Bordon, *Model predictive control*. Springer, 2013.

[43] J. Löfberg, "Yalmip: A toolbox for modeling and optimization in MATLAB," in *IEEE International Symposium on Computer Aided Control Systems Design*, Taipei, Taiwan, September 2004, pp. 284–289.

[44] R. Jia, M. H. Nehrir, and D. A. Pierre, "Voltage control of aggregate electric water heater load for distribution system peak load shaving using field data," in *39th North American Power Symposium*, Las Cruces, USA, September 2007.

[45] J. Z. Kolter and M. J. Johnson, "REDD: A public data set for energy disaggregation research," in *SustKDD workshop on Data Mining Applications in Sustainability*, San Diego, USA, August 2011.

[46] H. Chen, T. N. Cong, W. Yang, C. Tan, Y. Li, and Y. Ding, "Progress in electrical energy storage system: A critical review," *Progress in Natural Science*, vol. 19, no. 3, pp. 291–312, 2009.

[47] J.-Y. Joo and M. D. Ilić, "Adaptive load management (ALM) in electric power systems," in *International Conference on Networking, Sensing and Control (ICNSC)*, Chicago, USA, April 2010, pp. 637–642.

[48] P. Frankl, S. Nowak, M. Gutschner, S. Gnos, and T. Rinke, "Technology roadmap: solar photovoltaic energy," *International Energy Association*, May 2010.

[49] "German rooftops dominate global photovoltaic capacity," Available online at http://marketrealist.com/2015/02/german-rooftops-domniate-global-photovoltaic-capacity, last accessed 15.04.2016.

[50] P. Trichakis, P. Taylor, P. Lyons, and R. Hair, "Predicting the technical impacts of high levels of small-scale embedded generators on low-voltage networks," *IET Renewable Power Generation*, vol. 2, no. 4, pp. 249–262, 2008.

[51] "Bundesnetzagentur, website," Available online at http://www.bundesnetzagentur.de, last accessed 29.09.2013.

[52] M. Fulton, R. Capalino, and J. Auer, "The german feed-in tariff: recent policy changes," *report of the Deutsche Bank Group*, September 2012.

[53] "Polysun simulation software by Vela Solaris AG," Available online at www.polysun.ch, last accessed on 29.09.2013.

[54] "Neu in Polysun 6: Optimieren Sie den Photovoltaik-Eigenverbrauch mit dem neuen Batterie-Rechenmodell (in German)," Polysun 6 press release, available online at http://www.velasolaris.com, last accessed in February 2013.

[55] R. Suzuki, Y. Fujimoto, and Y. Hayashi, "Determination method of optimal planning and operation for residential PV system and storage battery based on weather forecast," in *IEEE International Conference on Power and Energy (PECon)*, Kota Kinabalu Sabah, Malaysia, December 2012.

[56] E. L. Ratnam, S. R. Weller, and C. M. Kellett, "An optimization-based approach for assessing the benefits of residential battery storage in conjunction with solar PV," in *9th IREP Symposium*, Rethymon, Greece, August 2013.

[57] M. Braun, K. Büdenbender, D. Magnor, and A. Jossen, "Photovoltaic self-consumption in Germany: using lithium-ion storage to increase self-consumed photovoltaic energy," in *24th Euro-*

pean Photovoltaic Solar Energy Conference, Hamburg, Germany, September 2009.

[58] M. Osawa, K. Yoshimi, D. Yamashita, R. Yokohama, T. Masuda, H. Kondou, and T. Hirota, "Increase the rate of utilization of residential photovoltaic generation by EV charge-discharge control," in *Innovative Smart Grid Technologies - Asia (ISGT Asia)*, Tianjin, China, May 2012.

[59] M. Castillo-Cagigal, E. Caamano-Martín, E. Matallanas, D. Masa-Bote, A. Gutiérrez, F. Monasterio-Huelin, and J. Jiménez-Leube, "PV self-consumption optimization with storage and active DSM for the residential sector," *Solar Energy*, vol. 85, no. 9, pp. 2338–2348, 2011.

[60] E. Vrettos, A. Witzig, R. Kurmann, S. Koch, and G. Andersson, "Maximizing local PV utilization using small-scale batteries and flexible thermal loads," in *Proceedings of the 28th European Photovoltaic Solar Energy Conference*, Paris, France, October 2013.

[61] H. G. Beyer, J. Betcke, A. Drews, D. Heinemann, E. Lorenz, G. Heilscher, and S. Bofinger, "Identification of a general model for the MPP performance of PV-modules for the application in a procedure for the performance check of grid connected systems," in *28th European Photovoltaic Solar Energy Conference*, Paris, France, June 2004.

[62] J. A. Duffie and W. A. Beckman, *Solar engineering of thermal processes.* 2nd Edition, Wiley Interscience, New York, 1991.

[63] S. H. Rezaei, A. Witzig, and J. Marti, "Design methodology for combined solar and geothermal heating systems," in *European Solar Thermal Energy Conference (ESTEC)*, Munich, Germany, May 2009.

[64] S. A. Klein, J. A. Duffie, and W. A. Beckman, *TRNSYS a transient system simulation program (user's manual).* Solar Energy Laboratory University of Wisconsin-Madison.

[65] "Meteonorm handbook, part II: theory," available online at http://meteonorm.com, last accessed on 29.09.2013.

[66] N. Pflugradt and B. Platzer, "Behavior based load profile generator for domestic hot water and electricity use," in *International Conference on Energy Storage (Innostock)*, Lleida, Spain, May 2012.

[67] "Load profile generator," available online at http://www-user.tu-chemnitz.de, last accessed on 29.09.2013.

[68] "Recent facts about photovoltaics in Germany," Report from Fraunhofer Institute for Solar Energy Systems, Germany, September 2013.

[69] "Ancillary services: Technical and commercial insights," Available online at http://www.consultkirby.com/files/Ancillary_Services_-_Technical_And_Commercial_Insights_EXT_.pdf, last accessed 19.04.2016.

[70] "Survey on ancillary services procurement and electricity balancing market design," Available online at www.entsoe.eu/publications/position-papers, last accessed on 23.01.2015.

[71] P. Kundur, *Power system stability and control.* Edited by N. J. Balu and M. G. Lauby, New York: McGraw-hill, 1994.

[72] "Basic principles of ancillary service products," Available online at www.swissgrid.ch, last accessed on 23.01.2015.

[73] "PJM manual 11: energy & ancillary services market operations," Available online at www.pjm.com/~/media/documents/manuals/m11.ashx, last accessed on 23.01.2015.

[74] O. Leitermann and J. L. Kirtley, "Energy storage for use in load frequency control," in *Innovative Technologies for an Efficient and Reliable Electricity Supply (CITRES)*, Waltham, USA, September 2010.

[75] ——, "Coordinating regulation and demand response in electric power grids using multirate model predictive control," in *Innovative Smart Grid Technologies (ISGT)*, Anaheim, USA, January 2011.

[76] Y. Makarov, P. Du, M. Kintner-Meyer, C. Jin, and H. Illian, "Sizing energy storage to accommodate high penetration of variable energy resources," *IEEE Transactions on Sustainable Energy*, vol. 3, no. 1, pp. 1039–1050, 2009.

[77] I. Avramiotis-Falireas, T. Haring, G. Andersson, and M. Zima, "Redesign of the automatic generation control scheme in the Swiss power system," in *IEEE PES General Meeting*, Washington DC, USA, July 2014.

[78] P. Barooah, A. Busic, and S. Meyn, "Spectral decomposition of demand-side flexibility for reliable ancillary services in a smart grid," in *48th Annual Hawaii International Conference on System Sciences*, Kauai, USA, January 2015.

[79] "PJM manual 12: Balancing operations," Available online at www.pjm.com/∼/media/documents/manuals/m12.ashx, last accessed 22.02.2016.

[80] S. Ihara and F. Schweppe, "Physically based modeling of cold load pickup," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-100, no. 9, pp. 4142–4150, 1981.

[81] C. Chong and A. Debs, "Statistical synthesis of power system functional load models," in *18th IEEE Conference on Decision and Control including the Symposium on Adaptive Processes*, no. 18, San Diego, USA, January 1979, pp. 264–269.

[82] R. Malhame and C.-Y. Chong, "Electric load model synthesis by diffusion approximation of a high-order hybrid-state stochastic system," *IEEE Transactions on Automatic Control*, vol. 30, no. 9, pp. 854–860, 1985.

[83] F. Constantopoulos, P. Schweppe and R. Larson, "Estia: A real-time consumer control scheme for space conditioning usage under spot electricity pricing," *Computers and Operations Research*, vol. 18, no. 8, pp. 751–765, Jan. 1991.

[84] J. Short, D. Infield, and L. Freris, "Stabilization of grid frequency through dynamic demand control," *IEEE Transactions on Power Systems*, vol. 22, no. 3, pp. 1284–1293, Aug. 2007.

[85] Z. Xu, J. Østergaard, and M. Togeby, "Demand as frequency controlled reserve," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1062–1071, 2011.

[86] A. Molina-Garcia, F. Bouffard, and D. S. Kirschen, "Decentralized demand-side contribution to primary frequency control," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 411–419, 2011.

[87] K. Samarakoon, J. Ekanayake, and N. Jenkins, "Investigation of domestic load control to provide primary frequency response using smart meters," *IEEE Transactions on Smart Grid*, vol. 3, no. 1, pp. 282–292, 2012.

[88] D. Angeli and P. A. Kountouriotis, "A stochastic approach to dynamic-demand refrigerator control," *IEEE Transactions on Control Systems Technology*, vol. 20, no. 3, pp. 581–592, May 2012.

[89] E. Kremers, J. Mari, and O. Barambones, "Emergent synchronisation properties of a refrigerator demand side management system," *Applied Energy*, vol. 101, pp. 709–717, 2013.

[90] S. H. Tindemans, V. Trovato, and G. Strbac, "Frequency control using thermal loads under the proposed ENTSO-E demand connection code," in *PowerTech Conference*, Eindhoven, the Netherlands, June 2015, pp. 1–6.

[91] D. Hammerstrom *et al.*, "Pacific northwest gridwise testbed demonstration projects - Part II: Grid friendly appliance project," Pacific Northwest National Laboratory, Tech. Rep., 2007.

[92] V. Moghadam, R. Mohammad, R. T. Ma, and R. Zhang, "Distributed frequency control in smart grids via randomized demand response," *IEEE Transactions on Smart Grid*, vol. 5, no. 6, pp. 2798–2809, 2014.

[93] S. H. Tindemans, V. Trovato, and G. Strbac, "Decentralized control of thermostatic loads for flexible demand response," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 5, pp. 1685–1700, 2015.

[94] T. Borsche, U. Markovic, and G. Andersson, "A new algorithm for primary frequency control with cooling appliances," *Computer Science-Research and Development*, pp. 1–7, 2015.

[95] B. Biegel, L. H. Hansen, P. Andersen, and J. Stoustrup, "Primary control by on/off demand-side devices," *IEEE Transactions on Smart Grid*, vol. 4, no. 4, pp. 2061–2071, 2013.

[96] A. Latiers, F. Glineur, and E. De Jaeger, "Toward semi-autonomous decentralized primary control from demand response," in *CIGRE Belgium Conference*, Brussels, Belgium, March 2014.

[97] S. Weckx, J. Driesen, and R. D'Hulst, "Optimal frequency support by dynamic demand," in *PowerTech Conference*, Grenoble, France, June 2013, pp. 1–6.

[98] C. Zhao, U. Topcu, and S. H. Low, "Frequency-based load control in power systems," in *American Control Conference (ACC), 2012*, Montreal, Canada, June 2012, pp. 4423–4430.

[99] C. Zhao, U. Topcu, N. Li, and S. Low, "Design and stability of load-side primary frequency control in power systems," *IEEE Transactions on Automatic Control*, vol. 59, no. 5, pp. 1177–1189, 2014.

[100] D. S. Callaway, "Tapping the energy storage potential in electric loads to deliver load following and regulation, with application to wind energy," *Energy Conversion and Management*, vol. 50, no. 5, pp. 1389–1400, 2009.

[101] C. Perfumo, E. Kofman, J. H. Braslavsky, and J. K. Ward, "Load management: Model-based control of aggregate power for populations of thermostatically controlled loads," *Energy Conversion and Management*, vol. 55, pp. 36–48, 2012.

[102] S. Kundu, N. Sinitsyn, S. Backhaus, and I. Hiskens, "Modeling and control of thermostatically controlled loads," *arXiv preprint arXiv:1101.2157*, 2011.

[103] S. Koch, M. Zima, and G. Andersson, "Potentials and applications of coordinated groups of thermal household appliances for power system control purposes," in *IEEE-PES/IAS Conference on Sustainable Alternative Energy*, Valencia, Spain, September 2009.

[104] T. M. Keep, F. E. Sifuentes, D. M. Auslander, and D. S. Callaway, "Using load switches to control aggregated electricity demand for

load following and regulation," in *IEEE PES General Meeting*, Detroit, USA, July 2011.

[105] S. Koch, J. L. Mathieu, and D. S. Callaway, "Modeling and control of aggregated heterogeneous thermostatically controlled loads for ancillary services," in *17th Power Systems Computation Conference*, Stockholm, Sweden, August 2011.

[106] L. Totu and R. Wisniewski, "Demand response of thermostatic loads by optimized switching-fraction broadcast," in *19th IFAC World Congress*, Cape Town, South Africa, August 2014.

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

[108] W. Zhang, J. Lian, C.-Y. Chang, and K. Kalsi, "Aggregated modeling and control of air conditioning loads for demand response," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4655–4664, Nov 2013.

[109] E. Can Kara, Z. Kolter, M. Berges, B. Krogh, G. Hug, and T. Yuksel, "A moving horizon state estimator in the control of thermostatically controlled loads for demand response," in *IEEE SmartGridComm Conference*, Vancouver, Canada, October 2013.

[110] T. Borsche, F. Oldewurtel, and G. Andersson, "Minimizing communication cost for demand response using state estimation," in *Proceedings of the PowerTech Conference*, Grenoble, France, June 2013.

[111] S. Moura, J. Bendtsen, and V. Ruiz, "Observer design for boundary coupled PDEs: Application to thermostatically controlled loads in smart grids," in *Proceedings of the IEEE Conference on Decision and Control*, Florence, Italy, Dec. 2013.

[112] J. A. Crabb, N. Murdoch, and J. M. Penman, "A simplified thermal response model," *Building Services Engineering Research and Technology*, vol. 8, no. 1, pp. 13–19, 1987.

[113] R. Kramer, J. van Schijndel, and H. Schellen, "Simplified thermal and hygric building models: a literature review," *Frontiers of Architectural Research*, vol. 1, no. 4, pp. 318–325, 2012.

[114] F. Oldewurtel, A. Ulbig, A. Parisio, G. Andersson, and M. Morari, "Reducing peak electricity demand in building climate control using real-time pricing and model predictive control," in *IEEE Conference on Decision and Control (CDC), Atlanta, USA*, December 2010.

[115] S. Kiliccote, "Open automated demand response communications in demand response for wholesale ancillary services," *Lawrence Berkeley National Laboratory*, 2010.

[116] B. Kirby, J. Kueck, T. Laughner, and K. Morris, "Spinning reserve from hotel load response," *The Electricity Journal*, vol. 21, no. 10, pp. 59–66, 2008.

[117] D. H. Blum and L. K. Norford, "Dynamic simulation and analysis of ancillary service demand response strategies for variable air volume HVAC systems," *HVAC&R Research*, vol. 20, no. 8, pp. 908–921, 2014.

[118] S. Kawachi, H. Hagiwara, and J. Baba, "Modeling and simulation of heat pump air conditioning unit intending energy capacity reduction of energy storage system in microgrid," in *European Conference on Power Electronics and Applications*, Birmingham, UK, August 2011.

[119] P. Zhao, G. P. Henze, S. Plamp, and V. J. Cushing, "Evaluation of commercial building HVAC systems as frequency regulation providers," *Energy and Buildings*, vol. 67, pp. 225–235, 2013.

[120] H. Hao, T. Middelkoop, P. Barooah, and S. Meyn, "How demand response from commercial buildings will provide the regulation needs of the grid," in *50th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, Illinois, USA, October 2012.

[121] H. Hao, A. Kowli, Y. Lin, P. Barooah, and S. Meyn, "Ancillary service for the grid via control of commercial building HVAC systems," in *American Control Conference (ACC)*, Washington DC, USA, June 2013.

[122] Y. Lin, P. Barooah, and S. Meyn, "Low-frequency power grid ancillary services from commercial building HVAC systems," in *IEEE SmartGridComm Conference*, Vancouver, Canada, October 2013.

[123] M. Maasoumy, C. Rosenberg, A. Sangiovanni-Vincentelli, and D. S. Callaway, "Model predictive control approach to online computation of demand-side flexibility of commercial buildings HVAC systems for supply following," in *American Control Conference (ACC)*, Portland, USA, June 2014.

[124] G. S. Pavlak, G. P. Henze, and V. J. Cushing, "Optimizing commercial building participation in energy and ancillary service markets," *Energy and Buildings*, vol. 81, pp. 115–126, 2014.

[125] J. T. Hughes, A. D. Dominguez-Garcia, and K. Poolla, "Identification of virtual battery models for flexible loads," *IEEE Transactions on Power Systems (early access article)*, 2016.

[126] J. L. Mathieu, D. S. Callaway, and S. Kiliccote, "Examining uncertainty in demand response baseline models and variability in automated responses to dynamic pricing," in *50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, Orlando, USA, December 2011, pp. 4332–4339.

[127] Y. Lin, P. Barooah, S. Meyn, and T. Middelkoop, "Experimental evaluation of frequency regulation from commercial building HVAC system," *IEEE Transactions on Smart Grid*, vol. 6, no. 2, pp. 776–783, 2015.

[128] M. Maasoumy, B. M. Sanandaji, A. Sangiovanni-Vincentelli, and K. Poolla, "Model predictive control of regulation services from commercial buildings to the smart grid," in *American Control Conference (ACC)*, Portland, USA, June 2014.

[129] F. Baccino, F. Conte, S. Massucco, F. Silvestro, and S. Grillo, "Frequency regulation by management of building cooling systems through model predictive control," in *Power Systems Computation Conference (PSCC), 2014*, Wroclaw, Poland, August 2014, pp. 1–7.

[130] D. He, J. Mei, R. Harley, and T. Habeter, "Utilizing building-level demand response in frequency regulation of actual microgrids," in *IEEE Conference on Industrial Electronics Society*, Vienna, Austria, November 2013.

[131] S. Kiliccote, P. Price, M. Piette, G. Bell, S. Pierson, E. Koch, J. Carnam, H. Pedro, J. Hernandez, and A. Chiu, "Field testing of

automated demand response for integration of renewable resources in California's ancillary services market for regulation products," Lawrence Berkeley National Laboratory, Tech. Rep. LBNL-5556E, 2007.

[132] J. MacDonald, S. Kiliccote, J. Boch, J. Chen, and R. Nawy, "Commercial building loads providing ancillary services in PJM," in *ACEEE Summer Study on Energy Efficiency in Buildings*, Pacific Grove, CA, USA, August 2014.

[133] Y. Kim, L. K. Norford, and J. L. Kirtley Jr., "Modeling and analysis of a variable speed heat pump for frequency regulation through direct load control," *IEEE Transactions on Power Systems*, vol. 30, no. 1, pp. 397–408, 2015.

[134] Y.-J. Kim, E. Fuentes, and L. K. Norford, "Experimental study of grid frequency regulation ancillary service of a variable speed heat pump," *IEEE Transactions on Power Systems (early access article)*, 2016.

[135] L. Su and L. K. Norford, "Demonstration of HVAC chiller control for power grid frequency regulation – Part 1: Controller development and experimental results," *Science and Technology for the Built Environment*, vol. 21, no. 8, pp. 1134–1142, 2015.

[136] ——, "Demonstration of HVAC chiller control for power grid frequency regulation – Part 2: Discussion of results and considerations for broader deployment," *Science and Technology for the Built Environment*, vol. 21, no. 8, pp. 1143–1153, 2015.

[137] I. Biel, I. Hiskens, and S. Backhaus, "Round-trip efficiency of fast demand response in a large commercial air conditioner," *Energy and Buildings*, vol. 97, pp. 47–55, 2015.

[138] C. Ziras, E. Vrettos, and G. Andersson, "Primary frequency control with refrigerators under startup dynamics and lockout constraints," in *IEEE PES General Meeting*, Denver, USA, July 2015.

[139] E. Vrettos, C. Ziras, and G. Andersson, "Fast and reliable primary frequency reserves from refrigerators with decentralized stochastic control," *IEEE Transactions on Power Systems (Early Access Article)*, 2016.

[140] M. Stadler, W. Krause, M. Sonnenschein, and U. Vogel, "Modelling and evaluation of control schemes for enhancing load shift of electricity demand for cooling devices," *Environmental Modelling & Software*, vol. 24, no. 2, pp. 285–295, 2009.

[141] J. M. Goncalves, C. Melo, C. J. L. Hermes, and J. R. Barbosa Jr., "Experimental mapping of the thermodynamic losses in vapor compression refrigeration systems," *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. 33, pp. 159 – 165, 2011.

[142] G. Van Kempen and L. Van Vliet, "Mean and variance of ratio estimators used in fluorescence ratio imaging," *Cytometry*, vol. 39, no. 4, pp. 300–305, 2000.

[143] F. Sossan, V. Lakshmanan, G. T. Costanzo, M. Marinelli, P. J. Douglass, and H. Bindner, "Grey-box modelling of a household refrigeration unit using time series data in application to demand side management," *Sustainable Energy, Grids and Networks*, vol. 5, pp. 1–12, 2016.

[144] D. Parker and T. Stedman, "Measured electricity savings of refrigerator replacement: Case study and analysis," *Summer Study on Energy Efficiency in Buildings, American Council for an Energy Efficient Economy*, vol. 3, pp. 199–211, 1992.

[145] "Test for primary control capability," Available online at www.swissgrid.ch, last accessed on 04.02.2016.

[146] M. Donnelly, D. Harvey, R. Munson, and D. Trudnowski, "Frequency and stability control using decentralized intelligent loads: Benefits and pitfalls," in *Power and Energy Society General Meeting*, Minneapolis, USA, July 2010, pp. 1–6.

[147] D. Ibrahim, "Accurate measurement of the mains electricity frequency," in *International Conference on Electrical and Electronics Engineering*, Istanbul, Turkey, December 2011.

[148] "Mains frequency display," Available online at https://fowkc.wordpress.com/2013/12/10/mains-frequency-display, last accessed on 21.04.2016.

[149] "Digital AC power generator frequency meter guage," Available online at http://www.ebay.com/itm/Digital-AC-Power-Generator-Frequency-Meter-Guage-/, last accessed on 21.04.2016.

[150] "Primary control power," Available online at www.swissgrid.ch, last accessed on 19.04.2016.

[151] "Appliance life expectancy," Available online at http://www.mrappliance.com/expert-tips/appliance-life-guide/, last accessed 19.04.2016.

[152] J. D. Lutz, "Water heaters and hot water distribution systems," Lawrence Berkeley National Laboratory. California Energy Commission, PIER Buildings End-Use Energy Effciency. CEC-500-2005-082., Tech. Rep., 2008.

[153] M. Pfeiffer, "Load control strategies for electric water heaters with thermal stratification," Semester Thesis, ETH No. PSL1108 (available online from ETH Zurich, Switzerland), 2011.

[154] M. González Vaya and G. Andersson, "Combined smart-charging and frequency regulation for fleets of plug-in electric vehicles," in *IEEE PES General Meeting*, Vancouver, Canada, July 2013.

[155] C. Wu, H. Mohsenian-Rad, J. Huang, and J. Jatskevich, "PEV-based combined frequency and voltage regulation for smart grid," in *IEEE PES ISGT*, Washington, USA, January 2012.

[156] E. Vrettos and G. Andersson, "Combined load frequency control and active distribution network management with thermostatically controlled loads," in *IEEE International Conference on Smart Grid Communications*, Vancouver, Canada, October 2013.

[157] W. Zhang, K. Kalsi, J. Fuller, M. Elizondo, and D. Chassin, "Aggregate model for heterogeneous thermostatically controlled loads with demand response," in *IEEE PES General Meeting*, San Diego, USA, July 2012.

[158] D. Schlipf, M. Scherer, and M. Haller, "Test for secondary control capability," Available online at www.swissgrid.ch, last accessed 25.04.2013.

[159] G. Andersson, *Power System Analysis.* Lecture notes at ETH Zurich, 2015.

[160] M. Lödl and R. Witzmann, "Operation strategies of hybrid energy storages in low-voltage distribution grids with a high degree of decentralized generation," in *IEEE PES ISGT Europe*, Berlin, Germany, October 2012.

[161] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER: steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, Feb. 2011.

[162] W. Heemels, M. Lazar, N. van de Wouw, and A. Pavlov, "Observer-based control of discrete-time piecewise affine systems: Exploiting continuity twice," in *IEEE Conference on Decision and Control*, Cancun, Mexico, December 2008.

[163] I. Hwang, H. Balakrishnan, and C. Tomlin, "State estimation for hybrid systems: applications to aircraft tracking," *IEE Proceedings of Control Theory and Applications*, vol. 153, no. 5, pp. 556–566, 2006.

[164] M. Hofbaur and B. Williams, "Hybrid estimation of complex systems," *IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics*, vol. 34, no. 5, pp. 2178–2191, 2004.

[165] A. Bemporad, G. Ferrari-Trecate, and M. Morari, "Observability and controllability of piecewise affine and hybrid systems," *IEEE Transactions on Automatic Control*, vol. 45, no. 10, pp. 1864–1876, October 2000.

[166] E. Vrettos, J. L. Mathieu, and G. Andersson, "Demand response with moving horizon estimation of individual thermostatic load states from aggregate power measurements," in *Proceedings of the American Control Conference*, Portland, USA, June 2014.

[167] ——, "Control of thermostatic loads using moving horizon estimation of individual load states," in *Power Systems Computation Conference (PSCC)*, Wroclaw, Poland, August 2014.

[168] F. Chen, J. Dai, B. Wang, S. Sahu, M. Naphade, and C. T. Lu, "Activity analysis based on low sample rate smart meters," in *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Diego, USA, August 2011.

[169] G. S. Ledva, E. Vrettos, S. Mastellone, G. Andersson, and J. L. Mathieu, "Applying networked estimation and control algorithms to address communication bandwidth limitations and latencies in demand response," in *48th Hawaii International Conference on System Sciences (HICSS)*, Kauai, USA, January 2015, pp. 2645–2654.

[170] R. Mortensen and K. Haggerty, "Dynamics of heating and cooling loads: models, simulation, and actual utility data," *IEEE Transactions on Power Systems*, vol. 5, no. 1, pp. 243–249, 1990.

[171] C. Uçak and Çağlar, "The effect of load parameter dispersion and direct load control actions on aggregated load," in *POWERCON*, Beijing, China, August 1998, pp. 280–284.

[172] A. Bemporad, D. Mignone, and M. Morari, "Moving horizon estimation for hybrid systems and fault detection," in *American Control Conference (ACC)*, San Diego, USA, June 1999, pp. 2471–2475.

[173] ——, "Moving horizon estimation for hybrid systems," *IEEE Transactions on Automatic Control*, vol. 47, no. 10, pp. 1663–1676, October 2002.

[174] L. Pérez-Lombard, J. Ortiz, and C. Pout, "A review on buildings energy consumption information," *Energy and buildings*, vol. 40, no. 3, pp. 394–398, 2008.

[175] E. Vrettos, F. Oldewurtel, F. Zhu, and G. Andersson, "Robust provision of frequency reserves by office building aggregations," in *World Congress of the International Federation of Automatic Control (IFAC)*, Cape Town, South Africa, August 2014.

[176] X. Zhang, E. Vrettos, M. Kamgarpour, G. Andersson, and J. Lygeros, "Stochastic frequency reserve provision by chance-constrained control of commercial buildings," in *Control Conference (ECC), 2015 European*, Linz, Austria, July 2015, pp. 1134–1140.

[177] E. Vrettos, F. Oldewurtel, and G. Andersson, "Robust energy-constrained frequency reserves from aggregations of commercial buildings," *IEEE Transactions on Power Systems*, vol. 31, no. 6, pp. 4272–4285, 2016.

[178] E. Vrettos and G. Andersson, "Scheduling and provision of secondary frequency reserves by aggregations of commercial buildings," *IEEE Transactions on Sustainable Energy*, vol. 7, no. 2, pp. 850–864, 2016.

[179] H. Hao, Y. Lin, A. Kowli, P. Barooah, , and S. Meyn, "Ancillary service to the grid through control of fans in commercial building HVAC systems," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 2066–2074, 2014.

[180] J. Brasz and L. Tetu, "Variable-speed centrifugal chiller control for variable primary flow (VPF) applications," in *International Compressor Engineering Conference*, West Lafayette, USA, July 2008.

[181] Y. Lin, P. Barooah, and J. Mathieu, "Ancillary services to the grid from commercial buildings through demand scheduling and control," in *American Control Conference (ACC)*, Chicago, USA, July 2015.

[182] X. Zhang, G. Schildbach, D. Sturzenegger, and M. Morari, "Scenario-based MPC for energy-efficient building climate control under weather and occupancy uncertainty," in *IEEE European Control Conference*, Zurich, Switzerland, July 2013.

[183] S. P. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[184] J. Löfberg, "Automatic robust convex programming," *Optimization methods and software*, vol. 27, no. 1, pp. 115–129, 2012.

[185] M. C. Campi, S. Garatti, and M. Prandini, "The scenario approach for systems and control design," *Annual Reviews in Control*, vol. 33, pp. 149–157, 2009.

[186] X. Zhang, S. Grammatico, G. Schildbach, P. Goulart, and J. Lygeros, "On the sample size of random convex programs with structured dependence on the uncertainty," *Automatica*, vol. 60, pp. 182–188, 2015.

[187] G. Morales-España, J. Latorre, and A. Ramos, "Tight and compact MILP formulation of start-up and shut-down ramping in unit commitment," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1288–1296, 2013.

[188] C. Ziras, "Integrating large shares of thermal loads in frequency reserves," Master Thesis, ETH No. PSL1427 (available online from ETH Zurich, Switzerland), 2015.

[189] "CBE thermal comfort tool," Available online at http://comfort.cbe.berkeley.edu/, last accessed on 08.06.2015.

[190] "Chebyshev type I filter design, MathWorks," Available online at http://www.mathworks.com/help/signal/ref/cheby1.html, last accessed 18.04.2016.

[191] "Secondary control power (weekly)," Available online at www.swissgrid.ch, last accessed on 23.01.2015.

[192] "PJM manual 18: PJM capacity market," Available online at www.pjm.com/~/media/documents/manuals/m18.ashx, last accessed 24.07.2015.

[193] E. Vrettos, E. C. Kara, J. MacDonald, G. Andersson, and D. Callaway, "Experimental demonstration of frequency regulation by commercial buildings – Part I: Modeling and hierarchical control design," *IEEE Transactions on Smart Grid (Early Access Article)*, 2016.

[194] ——, "Experimental demonstration of frequency regulation by commercial buildings – Part II: Results and performance evaluation," *IEEE Transactions on Smart Grid (Early Access Article)*, 2016.

[195] J. Široký, F. Oldewurtel, J. Cigler, and S. Prívara, "Experimental analysis of model predictive control for an energy efficient building heating system," *Applied Energy*, vol. 88, no. 9, pp. 3079–3087, 2011.

[196] Y. Ma, S. Richter, and F. Borrelli, "DMPC for building temperature regulation," in *Control and optimization with differential-algebraic constraints*, L. T. Biegler, S. L. Campbell, and V. Mehrmann, Eds. Philadelphia: Society for Industrial and Applied Mathematics, 2012.

[197] G. Welch and G. Bishop, "An introduction to the Kalman filter. 2006," *University of North Carolina: Chapel Hill, North Carolina, US*, 2006.

[198] J. G. Ziegler and N. B. Nichols, "Optimum settings for automatic controllers," *ASME Transactions*, vol. 64, no. 11, pp. 759–765, 1942.

[199] B. Givoni, *Passive low energy cooling of buildings*. John Wiley & Sons, 1994.

[200] E. Vrettos, C. Ziras, and G. Andersson, "Integrating large shares of heterogeneous thermal loads in power system frequency control," in *IEEE Powertech Conference*, Eindhoven, the Netherlands, June 2015.

[201] Q. Hu, F. Oldewurtel, M. Balandat, E. Vrettos, D. P. Zhou, and C. J. Tomlin, "Model identification of commercial building HVAC systems during regular operation - empirical results and challenges," in *American Control Conference (ACC)*, Boston, USA, July 2016.

[202] C. Schwaegerl *et al.*, "Advanced architectures and control concepts for more microgrids," European Commission Project, Tech. Rep., 2008.

[203] S. Papathanassiou, N. Hatziargyriou, and K. Strunz, "A benchmark low voltage microgrid network," in *CIGRE Symposium Power systems with dispersed generation: technologies, impacts on development, operation and performance*, Athens, Greece, April 2005.

Curriculum Vitae

09/2011 – 06/2016	Research assistant and PhD student at the Power Systems Laboratory, ETH Zurich, Switzerland. Advisor: Prof. Göran Andersson.
03/2015 – 06/2015	Visiting researcher at the Energy and Resources Group, University of California, Berkeley, USA. Advisor: Assistant Prof. Duncan Callaway.
02/2010 – 07/2011	Research assistant at the Electric Machines and Power Electronics Laboratory, National Technical University of Athens, Greece. Advisor: Associate Prof. Stavros Papathanassiou.
02/2010	Diploma in Electrical and Computer Engineering (Dipl.-Ing.), National Technical University of Athens, Greece.
07/2009 – 09/2009	IAESTE internship at the Faculty of Technology and Metallurgy, University of Belgrade, Serbia.
06/2004	High School Diploma (Apolytirion), Athens, Greece.
March 29, 1987	Born in Athens, Greece.