2024

Provision of energy and regulation reserve services by buildings

BOSTON UNIVERSITY

COLLEGE OF ENGINEERING

Dissertation

PROVISION OF ENERGY AND REGULATION RESERVE SERVICES BY BUILDINGS

by

WALEED ASLAM

B.S., Lahore University of Management Sciences, 2013
M.S., North Carolina State University, 2017

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2024

Approved by

First Reader

Michael Caramanis, PhD
Professor of Mechanical Engineering
Professor of Systems Engineering

Second Reader

Panagiotis Andrianesis, PhD
Research Associate Professor of Systems Engineering

Third Reader

Ioannis Ch. Paschalidis, PhD
Distinguished Professor of Engineering
Professor of Electrical and Computer Engineering
Professor of Systems Engineering
Professor of Biomedical Engineering
Professor of Computing & Data Sciences

Fourth Reader

Pablo Ruiz, PhD
Research Associate Professor of Mechanical Engineering
Chief Executive Officer, NewGrid Inc.

Fifth Reader

Leslie Norford, PhD
George Macomber (1948) Professor in Construction Management
Department of Architecture
Massachusetts Institute of Technology

Acknowledgments

This dissertation would not have been possible without the help and support of many people. First and foremost, I would like to express my deepest gratitude to my PhD advisor Prof. Michael Caramanis for his support, encouragement and mentorship throughout my years at Boston University. He provided me the opportunity to work in different areas of power systems and ingrained in me a passion for research and excellence, and I would not have been where I am today without his guidance and patience. I am equally grateful to Prof. Panagiotis Andrianesis who co-advised my dissertation and collaborated on various research projects. He inspired me through his uncompromising work ethic and research process, and his attention to detail not only helped me hone this dissertation but also made me a better researcher and writer. I would also like to thank my dissertation committee members, professors Paschalidis, Ruiz and Norford, for providing insightful feedback and advice that helped improve this dissertation.

I am grateful for all my systems engineering friends including Majid, Kasra, Shirantha, Jessie, Wei, and Salomon, without whom my journey at BU would not have been possible. I also acknowledge my group mates, Selin and Alex, who always inspired and helped me. I would also like to appreciate my colleagues at the Electric Power Research Institute, Nikita Singhal and Erik Ela, who supported me and provided opportunities that furthered my understanding of electricity markets, better enabling me to contextualize the utility of this dissertation. I am also thankful to my friends Anas, Zaki, Usman, Kasim, Gulrayz, Ahmed and Waleed T, who were my family away from home.

Last but not the least, I would like to thank my parents who always supported and believed in me. Though Ammi is no longer with us, I am certain that she would be the happiest today with my accomplishment. I am thankful to my brothers, sister,

and their families for their prayers and continued support. Finally, I would like to thank my wife, Areeha, for her love, patience, and support through the years; this dissertation is as much yours as is mine.

PROVISION OF ENERGY AND REGULATION RESERVE SERVICES BY BUILDINGS

WALEED ASLAM

Boston University, College of Engineering, 2024

Major Professor: Michael Caramanis
Professor of Mechanical Engineering
Professor of Systems Engineering

ABSTRACT

Power consumption and generation in the electrical grid must be balanced at all times. This balance is maintained by the grid operator through the procurement of energy and regulation reserve services in the wholesale electricity market. Traditionally, these services could only be procured from generation resources. However, helped by the advances in the computational and communication infrastructure, the demand resources are increasingly being leveraged in this regard. In particular, the Heating, Ventilation and Air-Conditioning (HVAC) systems of buildings are gaining traction due to the consumption flexibility provided by their energy-shifting and fast-response capabilities.

The provision of energy and regulation reserve services in the wholesale market, from the perspective of a typical building's HVAC system, can be construed in terms of two synergistic problems: an hourly deterministic optimization problem, referred to as Scheduling Problem, and a real-time (seconds timescale) stochastic control problem, referred to as Deployment Problem. So far, the existing literature has synthesized the solutions of these two problems in a simplistic, sequential/iterative

manner without employing an integrated approach that captures explicitly their cost and constraint interactions. Moreover, the deployment problem has only been solved with classical controllers which are not optimal, whereas the non-convexities in the scheduling problem have been addressed with methods that are sensitive to initialization. The current approaches therefore do not fully optimize the decisions of the two problems either individually or collectively, and hence do not fully exploit the HVAC resource.

The goal of the proposed research is to have optimal decision-making across both the scheduling and deployment problems. Our approach proposes deriving an optimal control policy for the deployment problem, and expressing the corresponding expected sum of deployment costs over the hour (called 'expected intra-hour costs') as a closed-form analytic function of the scheduling decisions. The inclusion of these expected intra-hour costs into the scheduling problem allows the optimization of the hourly scheduling decisions, pursuant to the real-time use of the optimal deployment control policy. Thus, our approach captures the interaction of the two problems and optimizes decisions across timescales yielding a truly integrated solution. We investigate the estimation of the expected intra-hour costs (based on a myopic policy optimizing instantaneous tracking error and utility loss), and solve the integrated problem with tight relaxations, demonstrating the value and applicability of the approach. Further, we investigate the derivation of the optimal control policy for the deployment problem, formulating and solving it as discounted-cost infinite horizon Dynamic Program (DP) and Reinforcement Learning (RL) problems. We also characterize the optimal policy as a closed-form analytic mapping from state-space to action-space, and obtain the corresponding expected cost-to-go (a.k.a. expected intra-hour costs) as a closed-form analytic function of the scheduling decisions. We further illustrate that the optimal policy better captures the deployment costs compared to existing approaches. As

such, our work represents a structured, interpretable and automated way for the end-to-end consideration of energy and regulation reserve market participation, and can be extended to other demand side resources.

Contents

1 Introduction **1**

1.1 Power Balance in the Electric Grid 1

1.2 Wholesale Electricity Markets 2

1.3 Regulation Reserves 3

1.4 HVAC Systems as Demand Side Resources 5

1.5 Dissertation Outline 9

2 Optimal HVAC Energy and Regulation Reserve Scheduling in Day-ahead Electricity Markets **12**

2.1 Introduction 14

2.2 HVAC and Building Model 16

2.2.1 HVAC Model 16

2.2.2 Building Thermal Dynamics 19

2.2.3 Hourly Scheduling Decision and State Variables 20

2.3 Integrated Scheduling and Deployment (ISD) Problem 22

2.3.1 Nonlinear Programming (NLP) Formulation 22

2.3.2 Estimation of the Expected IHCs 25

2.4 ISD Problem MIQP Formulation 29

2.4.1 Mixed Integer Linear Relaxations 29

2.4.2 MIQP Problem Summary 34

2.4.3 Problem Variants 34

2.5 Case Study 37

2.5.1 Input Data 37
2.5.2 Parameter Values in J_t 37
2.6 Numerical Results 41
2.6.1 Base-Case 42
2.6.2 NLP vs. MIQP 45
2.6.3 Comparison with Existing Approaches 48
2.6.4 Sensitivity Analysis 50
2.6.5 Additional Results 52
2.7 Conclusions 55

3 Stochastic Optimal Control for Energy and Regulation Reserve Deployment in Buildings using HVAC 57

3.1 Introduction 57
3.2 Model 61
3.2.1 Building Thermal Dynamics 61
3.2.2 Power Consumption 62
3.2.3 Regulation Signal 63
3.3 Regulation Reserve Deployment Problem 64
3.4 Solution Approach 67
3.4.1 Discounted-Cost Dynamic Program 67
3.4.2 Linear Programming based Solution Approach 70
3.4.3 Reinforcement Learning Formulation 71
3.4.4 Actor-Critic Algorithm 72
3.5 Numerical Results 73
3.5.1 Input Data 75
3.5.2 Optimal Control with DP 76
3.5.3 Comparison with Existing Approaches 85

3.5.4 RL Implementation 86
3.6 Conclusions 88

4 Conclusion **90**

4.1 Contributions 90
4.2 Future Directions 92

References **95**

Curriculum Vitae **102**

List of Tables

2.1 Building, HVAC and Temperature Preference Data 38

2.2 Normalized Parameter Values in J_t 39

2.3 Tightness of Objective Bounds in Different Cases 47

List of Figures

1·1 Illustration of Scheduling and Deployment Problems in a Day-ahead Electricity Market. 7

2·1 Integrated Scheduling and Deployment (ISD) Problem — Considers the Expected Intra-hour Cost in Making the Hourly Scheduling Decisions 12

2·2 HVAC Model. 17

2·3 Resistance-Capacitance (RC) Model for the Building. 19

2·4 Piecewise Polyhedral Relaxation (PPR) — each domain partition has polyhedral representation of the feasible space. 32

2·5 Energy & Regulation Prices (PJM), and, Ambient Temperature & Heat (NSRDB, New Brunswick, NJ), for July 19th, 2017. 38

2·6 Decision and State Variables in the Base-Case. 41

2·7 Objective Function Terms in the Base-Case. 43

2·8 Breakdown of J_t in the Base-Case. 44

2·9 Regulation Reserve Deployment during hour 10. 45

2·10 Computational Time in Different Cases for the NLP and MIQP Formulations. 48

2·11 Decision Variables for the ISD and Worst-Case (WC) problems in the Base-Case. 49

2·12 PJM RegD Signal over 15-minute Intervals for July, 2019. 50

2·13 Impact of Regulation Reserve Prices, λ_t^r. 52

2·14 Impact of Ambient Temperature, θ_t^a. 53

2·15 Ambient Temperature, Irradiance and Prices for Other Representative Days. 54
2·16 Supply Air Flow Rate and Regulation Reserve Capacity for Other Representative Days. 55

3·1 Optimal Control Formulation of the Deployment Problem — Provides Optimal Control Policy and Cost-to-go over the Hour. 59
3·2 Two Dimensional Markov Chain (TDMC) Model for the Regulation Signal Dynamics with Signal Value y and Direction d. 63
3·3 HVAC Fan Power Consumption and Inside Air Temperature under the Optimal Control Policy in the Base Case. 77
3·4 HVAC Fan Power Consumption and Inside Air Temperature under the Optimal Control Policy for different Temperature Utility Loss Penalty Factors ($/°C^2) — $\epsilon'^1 = 1.5\text{e}^{-5}$ (base), $\epsilon'^2 = 1.5\text{e}^{-4}, \epsilon'^3 = 6\text{e}^{-4}, \epsilon'^4 = 1.5\text{e}^{-3}$. 78
3·5 Fan Power Consumption and Inside Air Temperature under the Optimal Control Policy for different Temperature Limit Penalty Factors ($) — $\gamma^1 = 5\text{e}^{-7}$ (base), $\gamma^2 = 5\text{e}^{-6}, \gamma^3 = 5\text{e}^{-5}, \gamma^4 = 1\text{e}^{-4}$. 78
3·6 Fan Power Consumption and Inside Air Temperature under the Optimal Control Policy for different hourly Regulation Reserve Capacities (kW) — $r_t^1 = 0.2, r_t^2 = 0.3, r_t^3 = 0.6$ (base), $r_t^4 = 0.9$. 79
3·7 Inside Air Temperature, θ_τ, under the Optimal Control Policy for different hourly Energy Consumption baselines (kW) — $p_t^1 = 1.5, p_t^2 = 1.7$ (base), $p_t^3 = 2.5, p_t^4 = 3.5$. 80
3·8 Fan Power Consumption, p_τ^f, under the Optimal Control Policy or different hourly Average Temperatures (°C) — $\bar{\theta_t}^1 = 26.5$, $\bar{\theta_t}^2 = 26$ (base), $\bar{\theta_t}^3 = 25$, $\bar{\theta_t}^4 = 24$. 81

3·9 Fan Power Consumption and Inside Air Temperature under the Optimal Control Policy for different initial Supply Air Flow Rates (kg/s) — $m_\tau^1 = 1.5, m_\tau^2 = 3.14\,(\text{base}), m_\tau^3 = 5.5$. 81

3·10 Fan Power Consumption and Inside Air Temperature under the Optimal Control Policy for different initial Inside Air Temperatures (°C) — $\theta_\tau^1 = 24$, $\theta_\tau^2 = 26\,(\text{base})$, $\theta_\tau^3 = 27$. 82

3·11 Fan Power Consumption and Inside Air Temperature under the Optimal Control Policy for different initial Building Mass Temperatures (°C) — $\theta_\tau^{b,1} = 23$, $\theta_\tau^{b,2} = 26\,(\text{base})$, $\theta_\tau^{b,3} = 27$. 83

3·12 Fan Power Consumption and Inside Air Temperature under the Optimal Control Policy for different initial Regulation Signal Values — $y_\tau^1 = -1$, $y_\tau^2 = 0$, $y_\tau^3 = 0.75\,(\text{base})$, $y_\tau^4 = 0.875$. 84

3·13 Fan Power Consumption and Inside Air Temperature under the Optimal Control Policy at different discretizations (using $\epsilon' = 1.5e^{-4}$) — $N^1 = 15552\,(\text{base})$, $N^2 = 25920$. 85

3·14 Regulation Signal Tracking and Inside Air Temperature under Open-loop, Switched and DP controllers (base with $\epsilon' = 1.5e^{-4}$). 87

3·15 Regulation Signal Tracking and Inside Air Temperature under the Open-loop, Switched and DP controllers (base with $\gamma = 1e^{-4}$). 87

3·16 Regulation Signal Tracking and Inside Air Temperature under RL — $\epsilon'^1 = 5e^{-5}$ (base), $\epsilon'^2 = 6e^{-4}$. 89

List of Abbreviations and Acronyms

AS	.....	Ancillary Services
CAISO	.....	California ISO
DA	.....	Day-Ahead
DER	.....	Distributed Energy Resource
DP	.....	Dynamic Program
ERCOT	.....	Electric Reliability Council of Texas
EV	.....	Electric Vehicle
FERC	.....	Federal Energy Regulatory Commission
HVAC	.....	Heating, Ventilation and Air-conditioning
IHC	.....	Intra-Hour Costs
ISD	.....	Integrated Scheduling and Deployment Problem
ISO	.....	Independent System Operator
ISO-NE	.....	ISO New England
LMP	.....	Locational Marginal Price
MIQP	.....	Mixed Integer Quadratic Program
MISO	.....	Midcontinent ISO
NLP	.....	Non-linear Program
NYISO	.....	New York ISO
PI	.....	Proportional-Integral
PID	.....	Proportional-Integral-Derivative
PJM	.....	Pennsylvania Jersey Maryland Interconnection
PPR	.....	Piecewise Polyhedral Relaxation
RC	.....	Resistance-Capacitance
RL	.....	Reinforcement Learning
RT	.....	Real-Time
RTO	.....	Regional Transmission Organization
SPP	.....	Southwest Power Pool

Chapter 1

Introduction

1.1 Power Balance in the Electric Grid

Active power balance is a prerequisite for stable operation of the electric grid as it maintains system frequency. Traditionally, the balance between active power generation and consumption was ensured by vertically integrated utilities, achieved through controlling the output of large, centralized generators. The process operated under the assumption of consumption being independent and binding, a phenomenon known as 'generation follows consumption'. Over the years, the electric grid has undergone significant structural changes including the introduction of competitive wholesale electricity markets, which has added a layer of complexity to power balancing. Further, the penetration of renewable energy resources in the electricity supply mix has drastically increased, leading to higher generation variability and further complicating the problem. Fortunately, recent advances in computational, communication and power electronics infrastructure, along with enhanced demand-side resource capabilities, have enabled Distributed Energy Resources (DERs) to provide the needed flexibility. In particular, DERs such as Electric Vehicles (EVs) and Heating, Ventilation and Air-Conditioning (HVAC) systems, are being deemed as effective solutions due to their load-shifting and fast-response capabilities (Beil et al., 2016; Wang et al., 2020; Kim et al., 2016).

1.2 Wholesale Electricity Markets

The introduction of wholesale electricity markets marked a paradigm shift in the traditional, vertically integrated electric grid model. The emergence of these markets globally in the 1980s and 1990s was motivated by a variety of technical, economic and policy reasons, including increased competition, improved efficiency, and enhanced transparency, as well as the promise of newer investment and potential for greater innovation. In the United States, these markets were created as a direct result of Federal Energy Regulatory Commission (FERC) Order 888 in 1996 that encouraged the grid operators to facilitate open grid access. Nine different Regional Transmission Organizations (RTOs) and Independent System Operators (ISOs) have since been established in different geographic regions across North America to operate wholesale electricity markets.

Although the established markets extend across different geographies and possess their own unique aspects, they have remarkable similarities as well. Most ISOs broadly operate three different types of auction-based markets:

- **Energy Market** that enables the buying and selling of energy for immediate or future use.
- **Ancillary Services (AS) Market** that facilitates the procurement of essential services to maintain the grid reliability and stability.
- **Capacity Market** that ensures sufficient generation capacity is available to meet the peak demand.

Of the above markets, energy and AS markets are related to day-to-day operations, whereas the capacity market is associated with longer term resource planning. Specifically, both the energy and AS markets encompass a day-ahead (DA) forward

market process and a real-time (RT) spot market process, where the DA process operates at an hourly and the RT process at 5-15 minute granularity. Each of these market processes leads to locational, marginal cost based prices that enable efficient grid operation. Generation and demand resources become part of the market process by offering price-quantity bids that indicate their generation cost and consumption utility, respectively. The market co-optimizes energy and AS by scheduling the resources such that the net costs are minimized while power balance and other relevant constraints are satisfied. Some resources may choose to participate in the market as price takers only.

The energy market includes a single product, namely, energy, whereas, the AS market has different types of reserves including regulation, spinning, and non-spinning reserves. Regulation reserve refers to the resource capacity that provides automatic, instantaneous response as part of secondary frequency control to keep the system frequency within limits. Spinning reserve includes the resource capacity that is online and synchronized to the grid frequency, and can start providing energy to the grid within 10 minutes of a dispatch instruction. Non-spinning reserve is the resource capacity that is offline, but can be ramped to capacity, synchronized to the grid, and start providing energy to the grid typically within 10 to 30 minutes of a dispatch instruction.

1.3 Regulation Reserves

Regulation involves continuous, automatic adjustments to the resource outputs that help maintain the system's power balance and counteract the short-term frequency deviations. The output adjustment is typically done every 2 to 6 seconds based on the automatic generation control (AGC) signals issued by the ISO, and may entail increasing or decreasing the output. Some ISOs, e.g., California ISO (CAISO), Mid-

continent ISO (MISO), and Southwest Power Pool (SPP) consider the output increase and decrease as separate services — termed appropriately as 'regulation up' and 'regulation down'. Other ISOs, e.g., Pennsylvania-Jersey-Maryland Interconnection (PJM), New York ISO (NYISO) and ISO New England (ISO-NE) consider regulation to be a single service, with the the maximum increase and decrease in resource output restricted to be equal. The regulation reserves in the latter model are sometimes referred to as 'symmetric', as opposed to 'asymmetric' in the former model.

The AGC signal that drives the resource outputs is calculated based on the system's current frequency deviation from the nominal value of 60 Hz. Traditionally, the signal was aimed at large, fuel-based generators that had unlimited energy available to them. However, the penetration of energy-limited resources (e.g., electrochemical batteries and thermal storage) has risen considerably. These resources have a fast response and are perfectly suited to providing regulation, but their energy-limited nature may present a barrier, as perfectly tracking the AGC signal may either completely deplete or overflow their energy store and impact their planned operations. Thus, many ISOs have separated the AGC signals for the traditional and fast, energy-limited resources. For example, PJM uses Reg-A signal for conventional generators such as steam and combustion turbines, and Reg-D for energy-limited resources such as energy storage and demand-side response. Similarly, ISO-NE uses three different types of signals, where the additional signal type addresses fast resources that may operate at full discharge, full charge or the mid-point. Note the distinguishing aspect of the signal tailored for energy-limited resources is that it prevents the resource's energy store from reaching limit constraints in either direction, i.e., going completely empty or fully charged. For the ISOs with symmetric reserves, this invariably means that the signal would be energy-neutral, which is indeed the case over 30 minutes or an hour.

The regulation performance of different resources is based on how well the AGC signal is tracked. Typically, this entails evaluating the error between the reference power (based on AGC signal) and the actual resource output, and calculating a performance score. PJM uses a slightly different measure, where the delay, correlation and accuracy of the resource's response are all used to calculate the performance score. Across most ISOs, the score is a factor between 0 and 1, with a certain minimum performance threshold required for continued market participation of the resource.

The compensation for a resource providing regulation can be divided into: (i) *capacity* component that depends on the maximum amount of regulation that could be provided, i.e., regulation reserve capacity, and the market's regulation reserve capacity price, and (ii) *performance* component that depends on the performance score, net movement or 'mileage' of the AGC signal over a certain period, and the market's regulation reserve performance price. The exact break-up of compensation may vary across ISOs, as some guarantee the capacity component of the payment, whereas others may adjust it based on the performance score as well.

1.4 HVAC Systems as Demand Side Resources

HVAC systems maintain environmental comfort within buildings and can use the thermal inertia of air and building structures to store heat. They can regulate their own power consumption by modulating the fan/pump speed to control the flow rate of air supplied into the building (Liu et al., 2017; Vrettos et al., 2018b; Vrettos et al., 2018a), or the flow rate of water to the chiller (Qureshi and Jones, 2018; Su and Norford, 2015). Although their primary purpose is space conditioning, HVAC systems can efficiently utilize their fans/pumps to track stochastic, albeit energy-neutral, intra-hour regulation reserve signals broadcasted by the ISO (PJM, 2022a). This enables HVAC systems to obtain revenues by offering regulation reserve capacity

in the day-ahead (or real-time) market, provided that they do not incur relatively inordinate Intra-Hour Costs (IHCs) during the ISO's deployment of regulation reserve signals (Zhao et al., 2013). In addition, HVAC systems can modify their power consumption to minimize the building's energy costs and reduce peak demand.

HVAC systems present an attractive value proposition for providing energy and regulation services, for a number of reasons. Firstly, they tend to be an abundant resource; there were about 116 million HVAC units around the world in 2018 and their number is expected to rise to 151 million by 2024 (de Best,). Secondly, the Energy Management Systems in buildings can easily be leveraged to control the HVACs, and respond to grid (pricing or other) signals. Thirdly, the HVACs are the largest contributor to building energy consumption, but possess unused energy storage-like capabilities, due to their associated thermal inertia (Verbeke and Audenaert, 2018). Most notably, the modulation of real-time HVAC power consumption in an energy-neutral manner (i.e. with zero net energy transfer over an hour) has a limited impact on thermal comfort in the building. Thus, HVAC management can benefit the consumer and the grid operator (ISO).

The use of HVACs for providing grid services in non-market settings, i.e., through utility and third-party programs, has been well studied. Most of the work has focused on energy minimization, peak curtailment and load shifting applications (Braun, 1990; Henze et al., 1997; Nagai, 2002; Ma et al., 2012; Turner et al., 2015; Chakraborty et al., 2017). Different aspects of these problems, such as strict adherence to thermal preferences and robustness to uncertain model inputs, have also been considered (Kim, 2018; Ma et al., 2019; Du et al., 2017). Typically within such programs, the HVACs are paid a fixed amount in exchange for allowing the utility to manage the HVAC power consumption, which could be on an ongoing basis or during certain events. Recently, there has been a lot of focus on the wholesale market participation

of Distributed Energy Resources (DERs) including HVACs, since the markets offer enhanced economic incentives and transparency.

The provision of energy and regulation reserves from a typical building's HVAC system in wholesale markets can be construed in terms of two synergistic problems: (i) an hourly (or sub-hourly) deterministic optimization problem, referred to as 'Scheduling Problem', where the hourly (or sub-hourly) energy consumption and regulation reserve capacity offered by the HVAC in the day-ahead (or real-time) electricity market are optimized, and (ii) a real-time stochastic control problem, referred to as 'Deployment Problem', where the scheduled energy consumption and regulation reserve capacity are deployed every few seconds, by modulating the HVAC power consumption, to track the reference regulation signal from the ISO as well as to minimize the building occupants' utility loss (thermal discomfort cost).

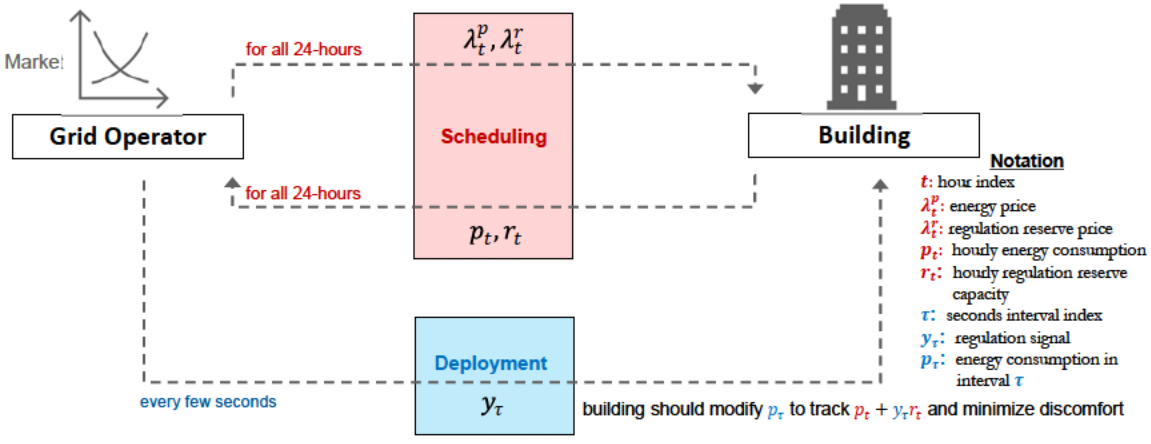


Figure 1·1: Illustration of Scheduling and Deployment Problems in a Day-ahead Electricity Market.

prices (Liu et al., 2017; Vrettos et al., 2018b; Maasoumy et al., 2014; Blum et al., 2016; Qureshi and Jones, 2018). In (Liu et al., 2017), the problem for determining power consumption and regulation reserve capacity schedules was considered in two-steps, where one corresponded to baseline power calculation and the other to regulation capacity determination. For computational efficiency reasons, the baseline calculation problem was further solved using sequential optimization. (Vrettos et al., 2018b) used a hierarchical approach where the high level problem determined the power consumption and regulation reserve capacity, with the lower level further modifying the power consumption to satisfy thermal comfort constraints. (Blum et al., 2016) considered a multi-zone building model and solved the optimization problem for scheduling energy, regulation and spinning reserves. (Maasoumy et al., 2014) and (Qureshi and Jones, 2018) assumed worst-case uncertainties to optimize the power consumption and regulation reserve capacities.

The deployment problem has also been considered in the literature, where different control architectures such as open-loop, feed-forward, proportional-integral, and switched control, have been successfully used for tracking the reference regulation signal and maintaining thermal comfort (MacDonald et al., 2014; Hao et al., 2013; Hao et al., 2014; Lin et al., 2015; Zhao et al., 2013; Qureshi and Jones, 2018; Liu et al., 2017; Vrettos et al., 2018b; Wang et al., 2022). Reference (MacDonald et al., 2014) used open-loop control whereas (Hao et al., 2013) and (Hao et al., 2014) used feed-forward control to modify the HVAC fan speed and power consumption, and thus to deploy regulation reserves and successfully track the reference regulation signal. Reference (Lin et al., 2015) used two distinct feedback control architectures; a proportional controller for HVAC fan speed and a lag-compensator for temperature set-points. Reference (Zhao et al., 2013) adopted a similar approach but instead used proportional-integral-derivative (PID) controllers in both cases, while (Qureshi and

Jones, 2018) only used a single proportional-integral (PI) controller for the temperature set-point. Reference (Liu et al., 2017) employed a slightly different approach by using a feed-forward controller for tracking the regulation signal, along with a PI controller for maintaining building temperature. Reference (Vrettos et al., 2018b) built upon the aforementioned control schemes by proposing a switched controller that selected between model-based feed-forward control and model-free feedback PI control depending on the regulation tracking error. Similarly, reference (Wang et al., 2022) also adopted a model-based controller.

1.5 Dissertation Outline

Although the existing literature addresses the scheduling and deployment problems, the solutions (including the parameters and control policies) for the two problems have often been synthesized separately. These separate solutions are then adjusted to work together, based on intuition and experimentation. In general, this approach of combining the separately derived solutions is not ideal, as it requires extensive trial-and-error testing to find an appropriate fit of scheduling parameters and deployment control policies. More importantly, this approach does not explicitly account for the cost and constraint interactions of the real-time deployment control dynamics with the (hourly) scheduling decisions. In addition to the lack of an integrated approach for the two problems, the individual problems have often themselves not been solved optimally. The deployment problem has been solved with classical controllers which are not optimal, whereas the scheduling problem has non-convexities which have been addressed with methods that can be sensitive to initialization. Thus, the use of HVAC resource, while effective, has been sub-optimal.

We propose an integrated approach that optimizes decision-making across both the (hourly) scheduling and (real-time) deployment problems. Our approach pro-

poses deriving an optimal control policy for the deployment problem, and expressing the expected sum of the real-time deployment costs over the hour (called 'expected intra-hour costs (IHCs)') as a closed-form analytic function of the hourly scheduling decisions. Incorporating these expected IHCs into the scheduling problem ensures that the hourly decisions are optimized, subject to the use of optimal control policy (and the incurrence of corresponding real-time costs) in the deployment problem. Hence, the costs for both problems (and timescales) are optimized, and the interaction between the problems are captured. We demonstrate the approach for day-ahead markets.

We also address the individual gaps in the scheduling and deployment problems. Specifically, for the scheduling problem, we propose relaxations that circumvent the computational issues associated with the non-convexity and provide bounds on the optimal solution. As for the deployment problem, we propose optimal control formulations including dynamic programming and reinforcement learning (RL) formulations, which represent a structured and interpretable way to evaluate the control policies, and are amenable to data-driven and adaptive implementations.

This dissertation is organized as follows:

- Chapter 2 presents the Integrated Scheduling and Deployment (ISD) approach, where the expected intra-hour costs corresponding to a myopic control policy are incorporated in the scheduling problem. Relaxations for addressing non-convexity related issues in the ISD problem are also detailed. In addition, comparisons with existing scheduling approaches are illustrated and the applicability of the ISD approach to more general settings is discussed.

- Chapter 3 presents the stochastic dynamic programming and RL-based formulations of the deployment problem, along with their respective solution approaches. The performance and structure of the obtained optimal control policy

are thoroughly analyzed and compared with the existing deployment control approaches.

- Chapter 4 concludes by discussing the contributions of the dissertation and possible directions for future research.

Chapter 2

Optimal HVAC Energy and Regulation Reserve Scheduling in Day-ahead Electricity Markets

In this chapter, we present the scheduling problem that optimizes the HVAC's hourly energy consumption and regulation reserve capacity for each hour in the day-ahead market. The presented problem is referred to as *Integrated Scheduling and Deployment (ISD) Problem*, since it incorporates for each hour, the expected sum of real-time deployment costs over the hour (called 'expected IHCs') as a closed-form analytic function of the hourly scheduling decisions (see Fig. 2·1). Note that the expected intra-hour cost here corresponds to the use of myopic deployment control policy that

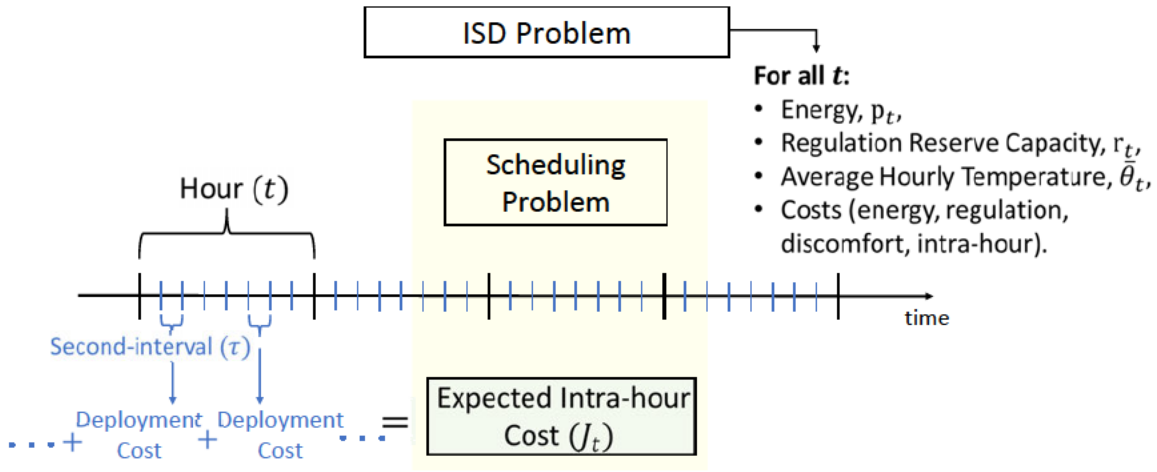


Figure 2·1: Integrated Scheduling and Deployment (ISD) Problem — Considers the Expected Intra-hour Cost in Making the Hourly Scheduling Decisions

Nomenclature	
Sets and Indices	
t	hour index, (Δt denotes interval duration)
τ	second-interval index ($\Delta\tau$ denotes interval duration)
n, k	partition index for $m, \bar{\theta}$ in bilinear term
n, n'	partition indices for m in quadratic term
Input Parameters	
c	specific heat capacity of air
R^b, R^w	thermal resistance (building mass, window)
C, C^b	thermal capacitance (inside air, building mass)
Q	heat transfer due to occupancy and irradiance
θ^a, θ^s	ambient temperature, supply air temperature
δ	valve position
η	coefficient of performance
α	coefficient of fan power consumption
λ^P, λ^R	energy price, regulation reserve price
$\hat{\theta}$	set-point temperature
ϵ	penalty factor for thermal discomfort
Variables	
decision	
m	supply air flow rate
r	regulation reserve capacity
state	
$\theta, \bar{\theta}$	inside air temperature, average θ
$\theta^b, \bar{\theta}^b$	building mass temperature, average θ^b
auxiliary	
ξ	equivalent linear variable for the bilinear term
ζ	equivalent linear variable for the quadratic term
w, u	weights for discretization points
x, y	indicators for interval activation
Functions of Decision and State Variables used for Compact Notation	
$\Delta\theta^c$	air temperature change at the cooling coil
p	HVAC power consumption
p^f, p^c	power consumption (fan, compressor)
Q^s	heat exchange between inside air and supply air
$J_{t(m_t, r_t, \theta_{t-1}, \theta^b_{t-1}, \bar{\theta}_t)}$	expected intra-hour costs (IHCs) for hour t

optimizes the real-time (seconds timescale) deployment costs associated with regulation signal tracking error and utility loss (thermal discomfort cost). We also present the mixed integer linear relaxations that are used for addressing the non-convexity related computational issues in the ISD problem. In addition, we compare the ISD problem with existing approaches that either consider the interaction between scheduling and deployment as hard constraints or assume worst-case scenarios. Lastly, we provide discussions and extensive numerical results to illustrate the various aspects and

the broader applicability of the proposed problem.

2.1 Introduction

As mentioned in Chapter 1, the HVAC scheduling problem has been considered in the literature (Maasoumy et al., 2014; Liu et al., 2017; Vrettos et al., 2018b; Vrettos et al., 2018a; Qureshi and Jones, 2018; Blum et al., 2016; Fabietti et al., 2018; Pavlak et al., 2014). However, these works do not model IHCs. They instead impose hard constraints that *either* incur large IHCs by over-committing on the regulation reserve capacity offer, *or*, assure negligible IHCs at the expense of limiting the regulation reserve capacity offer and the associated opportunity for additional revenues. For instance, the hard constraints are determined by considering a confidence-interval-based worst-case weather and regulation signal percentile (e.g., 95-th in (Liu et al., 2017), 97.5-th in (Vrettos et al., 2018b; Vrettos et al., 2018a)), or an uncertainty set based worst-case in (Qureshi and Jones, 2018; Maasoumy et al., 2014; Fabietti et al., 2018). Indeed, the size of the hourly reserve capacity offered, the HVAC/building state, and the level of the intra-hour regulation signal broadcasted by the ISO, may result in IHCs consisting of regulation signal tracking error penalties and occupant discomfort (Hao et al., 2013; Lin et al., 2015). Modeling of IHCs and anticipating them when hourly decisions are made, makes this chapter unique.

This chapter focuses on optimizing the *hourly* energy consumed and regulation reserve capacity offered by an HVAC system participating in the day-ahead electricity market. Its novelty is the incorporation into the objective function of the *expected IHCs* resulting from the HVAC's obligation to modulate its power consumption during the ISO's deployment of the hourly offered regulation reserve capacity. These IHCs internalize the expected regulation tracking error and incremental occupant discomfort incurred in the HVAC's best effort to follow ISO regulation reserve sig-

nals. We formulate an ISD problem by including into the optimized objective function the IHCs along with the hourly cost of power consumed, the revenue from regulation reserve capacity offered, and the space-conditioning derived occupant discomfort cost. As such, our ISD problem is unique in the optimal HVAC scheduling literature in the context of hourly day-ahead markets. We employ a pre-processor, exogenous to the ISD problem, which draws from and expands work reported in (Bilgin et al., 2016) and (Yanikara, 2020), to estimate the expected IHCs — i.e., the regulation signal tracking error and incremental building occupant discomfort *average cost*, incurred during each hour reflecting the optimal HVAC adaptation to intra-hour regulation signals — as a closed-form analytic function of the hourly decision variables (that refer to HVAC controls and offered regulation reserve capacity) and state variables (that refer to temperature).

Moreover, our ISD problem employs high fidelity models of the HVAC and the building structure's capability to store energy (Ma et al., 2012) that contain non-convexities. To deal with the non-convexities, in particular a bilinear term and a non-convex quadratic constraint, we relax the complicating terms using piecewise, convex polyhedra. Specifically, we employ a Special Ordered Set (SOS2) based formulation, known as Piecewise Polyhedral Relaxation (PPR) (Sundar et al., 2021), which is known to be tighter compared to standard and piecewise McCormick relaxations while incurring small computational overhead (Nagarajan et al., 2019). The resulting formulation is a Mixed Integer Quadratic Program (MIQP) problem, which can be efficiently solved using off-the-shelf solvers, and provides tight bounds on the ISD problem objective. We report numerical results of our formulation on a realistic office building to demonstrate our ISD problem formulation and quantify the tightness of the bounds on the ISD problem objective. We compare with existing approaches and we analyze the impact of different exogenous factors, namely of day-ahead market

prices, ambient temperature, and inside air temperature bounds, on the optimal ISD problem decision and state variables.

To summarize, the contribution of this chapter is three-fold. First and most important, we formulate an ISD problem which internalizes the trade-off captured by IHCs into the hourly problem. Second, we present an MIQP relaxation of the (originally NLP) ISD problem, which is computationally efficient and provides tight solution bounds. Third, we compare with existing works and illustrate how our approach has a more representative consideration of the deployment costs.

The remainder of the chapter is organized as follows. In Section 2.2, we present the HVAC and building model. In Section 2.3, the ISD problem formulation and the method for estimating the expected IHCs are presented. In Section 2.4, we discuss the relaxations for the ISD problem that result in a MIQP formulation. The input data and the functional form of expected IHCs are presented in Section 2.5, followed by relevant numerical results in Section 2.6 and conclusions in Section 2.7.

2.2 HVAC and Building Model

In this section, we detail the power consumption for the HVAC (in Subsection 2.2.1), thermal dynamics within the building (in Subsection 2.2.2), and the hourly decision and state variables under the state-space representation of the model (in Subsection 2.2.3).

2.2.1 HVAC Model

We consider a unitary HVAC system as illustrated in Fig. 2·2 (Ma et al., 2012), where, for simplicity and without loss of generality, we model the cooling operation. The HVAC mixes the ambient air with the return air from the building in a ratio that is governed by the valve position. This air mixture is then cooled by the cooling coil connected to a compressor (not shown) and circulated inside the building by the

fan. The supply air temperature, θ^s, thus satisfies:

$$\theta^s = \delta\,\theta + (1-\delta)\theta^a - \Delta\theta^c, \tag{2.1}$$

where δ is the valve position, θ is the inside air temperature that is also the return air temperature, θ^a is the ambient temperature, and $\Delta\theta^c$ is the air temperature change at the cooling coil. Typically, the values of the valve position, δ, and supply air temperature, θ^s, are assumed to be fixed (Qureshi and Jones, 2018). Note that we have dropped the time index for simplicity.

The HVAC total power consumption, p, is given by:

$$p = p^c + p^f, \tag{2.2}$$

where p^c is the power consumption of the compressor, and p^f is the power consumption of the fan.

The power consumption of the compressor, p^c, is related to the supply air flow rate, m, and the temperature change, $\Delta\theta^c$, as follows (Van Ness, 1983):

$$p^c = m\,\frac{c}{\eta}\,\Delta\theta^c, \tag{2.3}$$

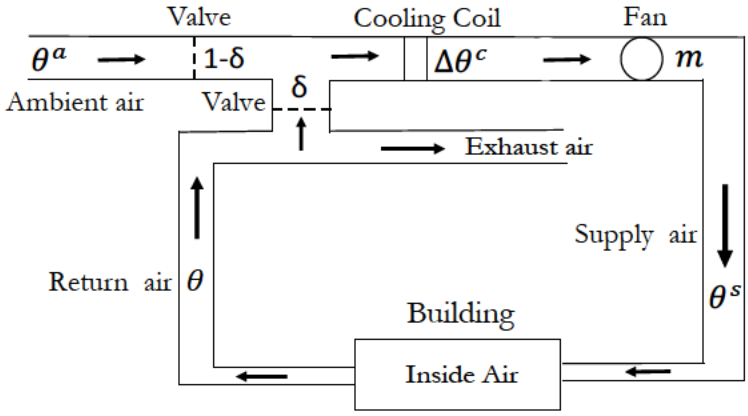


Figure 2·2: HVAC Model.

where c is the specific heat capacity of the air, and η the coefficient of performance of the HVAC. We make a simplifying assumption and consider η to be a constant. Replacing $\Delta\theta^c$ using (2.1), and rearranging terms, (2.3) becomes:

$$p^c = \frac{c}{\eta}\delta\, m\, \theta + \frac{c}{\eta}\big[(1-\delta)\theta^a - \theta^s\big]m. \tag{2.4}$$

Note that the first term in (2.4) contains a bilinear term $m\,\theta$, whereas the second term depends linearly on m.

The power consumption of the fan, p^f, is described by a quadratic function of the supply air flow rate, m, as follows:

$$p^f = \alpha_1 m + \alpha_2 m^2, \tag{2.5}$$

where α_1 and α_2 are parameters that are fitted using historical fan power data (Ma et al., 2012). The fast response of the fan, by adjusting the supply air flow rate, enables the HVAC to modulate its power consumption and deploy regulation reserves (Beil et al., 2016). Note the compressor's controller and equipment dynamics are slower than the fast variation in fan speed, thus preventing it from modulating its power consumption and deploying regulation reserves. Although the use of auxiliary technology can allow the compressor to modulate its power consumption as well, we do not assume that to be the case.

In this chapter, we consider that regulation reserve follows Eastern US ISOs' regulation deployment practices, e.g., ISO New England (ISO-NE) and Pennsylvania-Jersey-Maryland (PJM) (ISO-NE, 2020; PJM, 2022a). Specifically, the regulation reserve capacity r is symmetric — i.e., equal in the up and down directions around the hourly energy consumption — and the regulation signal from the grid operator that is used to modulate power and deploy regulation reserves is energy-neutral over an hour — so that its impact on the HVAC hourly energy consumption is negligible. Hence,

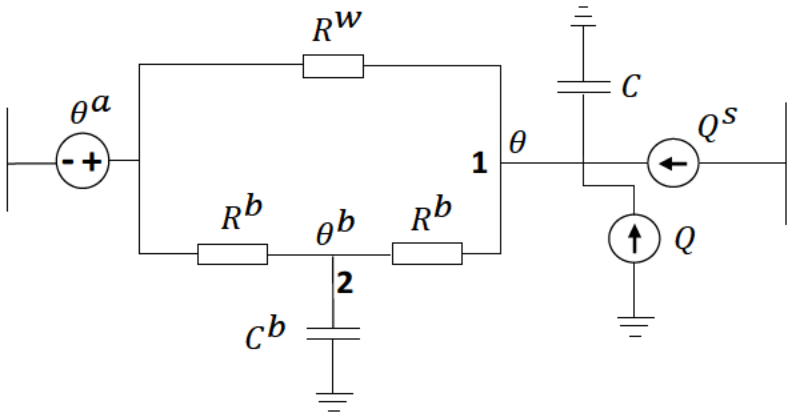


Figure 2·3: Resistance-Capacitance (RC) Model for the Building.

the regulation reserve capacity is constrained within the HVAC fan's capability, as follows:

$$p^f - r \geq p^{f-}, \qquad p^f + r \leq p^{f+}, \tag{2.6}$$

where p^{f-} and p^{f+} represent the minimum and maximum power consumption of the fan, which are determined by the fan capability and bounds on the supply air flow rate, m^- and m^+, respectively. Obviously, the HVAC maximum regulation reserve capability is $(p^{f+} - p^{f-})/2$, when the power consumption of the fan is $p^f = (p^{f+} + p^{f-})/2$.

The HVAC total power consumption (2.2), using (2.4) and (2.5), is given by:

$$p = \frac{c}{\eta}\,\delta\, m\,\theta + \Big[\alpha_1 + \frac{c}{\eta}\big[(1-\delta)\theta^a - \theta^s\big]\Big] m + \alpha_2 m^2. \tag{2.7}$$

Note that (2.5) includes a bilinear term $m\,\theta$, and a quadratic dependence on m.

2.2.2 Building Thermal Dynamics

The building thermal dynamics are described using the second-order Resistance-Capacitance (RC) model shown in Fig. 2·3 (Atam and Helsen, 2016), which represents temperature as a nodal voltage and heat exchange as a current injection. Let

R^b, C^b, and θ^b denote the building mass thermal resistance, thermal capacitance, and temperature, respectively, R^w the thermal resistance of windows, and C the thermal capacitance of inside air. Let Q^s denote the heat exchange between inside air and supply air, and Q the heat exchange between inside air and other sources (e.g., irradiance, occupancy).

Applying Kirchhoff's Current Law at nodes 1 and 2 of Fig. 2·3, we get the following thermal dynamics:

$$C\dot{\theta} = \frac{\theta^a - \theta}{R^w} + \frac{\theta^b - \theta}{R^b} + Q + Q^s, \tag{2.8a}$$

$$C^b\dot{\theta^b} = \frac{\theta^a - \theta^b}{R^b} + \frac{\theta - \theta^b}{R^b}, \tag{2.8b}$$

where $\dot{\theta}$ and $\dot{\theta^b}$ represents the rate of change of the inside air and the building mass temperatures, respectively.

2.2.3 Hourly Scheduling Decision and State Variables

Let us now consider the HVAC hourly scheduling. We introduce index t to denote an hour. The HVAC decides the amount of offered regulation reserve capacity, r_t, at each hour t, which is linked to the supply air flow rate m_t. Hence, m_t and r_t are the HVAC *decision variables* for hour t.

The building thermal dynamics determine the *state variables*, i.e., the inside air temperature, θ_t, and the building mass temperature θ_t^b. As can be seen in (2.8a), these state variables depend on the heat exchange Q_t^s between the inside air and supply air, which in turn is dependent on the decision variable m_t, as follows:

$$Q_t^s = m_t\, c\, (\theta^s - \theta_t). \tag{2.9}$$

Since the thermal time constant of the building (in the order of 30 hours) is much larger than the hourly granularity, we can accurately approximate the building dynamics using Euler's mid-point method, as follows.

Let $\bar{\theta}_t$ and $\bar{\theta}^b_t$ denote the average hourly temperatures of inside air and building mass, respectively. The inside air temperature dynamics (2.8a) are approximated by:

$$\theta_t - \theta_{t-1} = \frac{\theta^a_t - \bar{\theta}_t}{R^w C} + \frac{\bar{\theta}^b_t - \bar{\theta}_t}{R^b C} + \frac{Q_t}{C} + m_t \frac{c}{C}(\theta^s - \bar{\theta}_t), \tag{2.10}$$

where we also replaced Q^s_t using (2.9). Rearranging the terms in (2.10), we get:

$$\begin{aligned}\theta_t - \theta_{t-1} = \frac{c\theta^s}{C} m_t - \frac{c}{C} m_t \bar{\theta}_t - \big(\frac{1}{R^b C} + \frac{1}{R^w C}\big)\bar{\theta}_t \\ + \frac{1}{R^b C}\bar{\theta}^b_t + \frac{Q_t}{C} + \frac{\theta^a_t}{R^w C}.\end{aligned} \tag{2.11}$$

Similarly, the building mass temperature dynamics (2.8b) are approximated by:

$$\theta^b_t - \theta^b_{t-1} = \frac{1}{R^b C^b}\bar{\theta}_t - \frac{2}{R^b C^b}\bar{\theta}^b_t + \frac{\theta^a_t}{R^b C^b}. \tag{2.12}$$

Furthermore, the HVAC hourly energy consumption, denoted by p_t (since t refers to one hour), is given by:

$$p_t = \frac{c}{\eta}\,\delta\, m_t\, \bar{\theta}_t + \Big[\alpha_1 + \frac{c}{\eta}\big[(1-\delta)\theta^a - \theta^s\big]\Big] m_t + \alpha_2 m_t^2. \tag{2.13}$$

Notably, increasing participation of storage resources in the provision of frequency regulation is expected to favor the creation of energy-neutral (over an hourly period) regulation signals, *a la* PJM RegD. However, relaxing the energy-neutrality assumption, the deployment of regulation reserve may result in a different temperature at the end of the hour than the anticipated value. Hence, an offset to the energy consumption may be required to be applied in real-time, to counter potential persistence of the regulation signal in either direction. Again, similarly to storage resources, HVACs may update their economic basepoints every 5 minutes to adapt to real-time conditions, and correct any differences between the anticipated and actual temperatures.

2.3 Integrated Scheduling and Deployment (ISD) Problem

In this section, we detail the NLP formulation of the ISD problem (in Subsection 2.3.1), and we detail the estimation of the expected IHCs (in Subsection 2.3.2).

2.3.1 Nonlinear Programming (NLP) Formulation

Let $t \in \mathcal{T} = \{1, ..., 24\}$ denote the set of hours in a day. The NLP formulation is presented below.

Objective Function

We reiterate that the ISD problem optimizes the hourly energy consumption and regulation reserve capacity offered by the HVAC in the day-ahead market, while also considering the hourly discomfort cost, and the expected IHCs associated with regulation reserve deployment. The ISD problem objective is:

$$\textit{minimize}\ \{\text{Energy Cost} - \text{Regulation Revenue} + \text{Discomfort Cost} + \text{Total IHCs}\}. \tag{2.14}$$

We define each term in (2.14) below.

$$\text{Energy Cost} = \sum_t \lambda_t^p p_t, \tag{2.14a}$$

where λ_t^p is the energy price at hour t, and p_t is the HVAC energy consumption at hour t — see (2.13) — given by

$$p_t = \beta_1 m_t \bar{\theta}_t + \beta_2 m_t + \alpha_2 m_t^2, \quad \forall t, \tag{2.14b}$$

with $\beta_1 = \frac{c}{\eta}\delta$, and $\beta_2 = \alpha_1 + \frac{c}{\eta}[(1-\delta)\theta_t^a - \theta^s]$.

$$\text{Regulation Revenue} = \sum_t \lambda_t^r r_t, \tag{2.14c}$$

where λ_t^r is the regulation reserve price at hour t, and r_t is the HVAC offered regulation reserve capacity at hour t.

$$\text{Discomfort Cost} = \sum_t \epsilon_t(\bar{\theta}_t - \hat{\theta}_t)^2, \tag{2.14d}$$

where ϵ_t is the penalty factor for thermal discomfort at hour t, and $\hat{\theta}_t$ is the HVAC temperature set-point at hour t.

$$\text{Total IHCs} = \sum_t J_t(m_t, r_t, \theta_{t-1}, \theta_{t-1}^b, \bar{\theta}_t), \tag{2.14e}$$

where, J_t is the "expected IHCs" during hour t. We express J_t as a closed-form analytic function of the hourly decision variables (m_t and r_t) and state variables ($\theta_{t-1}, \theta_{t-1}^b, \bar{\theta}_t$). The expected IHCs, J_t, includes regulation signal tracking error and intra-hour occupant discomfort costs (incremental to the average hourly discomfort cost), which are incurred during regulation reserve deployment when the HVAC discomfort preferences limit regulation signal tracking capabilities. For example, scheduling high regulation reserve capacity will render tracking large regulation signals impossible without compromising occupant discomfort. Therefore, optimal deployment is achieved through the best trade-off between the tracking error and incremental occupant discomfort costs. We essentially capture this trade-off through the use of closed-form J_t acting as a soft constraint penalty.

We note that the decision variables appear in J_t's expression since they impact the intra-hour power modulation (and hence the tracking error), with air flow rate m_t also impacting the intra-hour thermal dynamics (and hence the intra-hour discomfort). The state variables are also included in J_t since they impact the intra-hour building occupant discomfort costs. Specifically, the intra-hour incremental occupant discomfort cost is calculated around the average inside air temperature $\bar{\theta}_t$, whereas the inside air and building mass temperatures at the start of the hour, θ_{t-1} and

θ^b_{t-1}, affect the intra-hour temperature dynamics (and thus the intra-hour discomfort costs). This should come as no surprise since J_t is in fact the expected cost-to-go over the hour when the system state at the beginning of the hour is described by θ_{t-1} and θ^b_{t-1}. We detail the estimation of J_t in Subsection 2.3.2.

Constraints

The ISD problem includes thermal dynamics, regulation and bound constraints, which are presented below.

Thermal Dynamic Constraints: The inside air and building mass temperatures are obtained by (2.11) and (2.12) as follows:

$$\theta_t = \theta_{t-1} + \gamma_1 m_t + \gamma_2 m_t \bar{\theta}_t + \gamma_3 \bar{\theta}_t + \gamma_4 \bar{\theta}_t{}^b + \gamma_{5,t}, \ \forall t, \tag{2.15}$$

$$\theta^b_t = \theta^b_{t-1} + \gamma_6 \bar{\theta}^b_t + \gamma_7 \bar{\theta}_t + \gamma_{8,t}, \quad \forall t, \tag{2.16}$$

where $\gamma_1 = c\theta^s/C$, $\gamma_2 = -c/C$, $\gamma_3 = -1/(R^bC) - 1/(R^wC)$, $\gamma_4 = 1/(R^bC)$ and $\gamma_{5,t} = Q_t/C + \theta^a_t/(R^wC)$, $\gamma_6 = -2/(R^bC^b)$, $\gamma_7 = 1/(R^bC^b)$ and $\gamma_{8,t} = \theta^a_t/(R^bC^b)$. For completeness, the average hourly temperatures for the inside air and the building mass are given by:

$$\bar{\theta}_t = \frac{\theta_{t-1} + \theta_t}{2}, \quad \bar{\theta}^b_t = \frac{\theta^b_{t-1} + \theta^b_t}{2}, \quad \forall t. \tag{2.17}$$

Regulation Constraints: Following (2.6), and using (2.5), the regulation reserve capacity, r_t, is constrained as given below:

$$\text{Regulation Up:} \qquad \alpha_1 m_t + \alpha_2 m_t^2 - r_t \geq p^{f-}, \quad \forall t, \tag{2.18}$$

$$\text{Regulation Down:} \quad \alpha_1 m_t + \alpha_2 m_t^2 + r_t \leq p^{f+}, \quad \forall t. \tag{2.19}$$

Bound Constraints: The fan capability and potential inside air temperature hard limits impose the following bound constraints on m_t and θ_t, respectively:

$$m^- \leq m_t \leq m^+, \quad \theta_t^- \leq \theta_t \leq \theta_t^+, \quad \forall t. \tag{2.20}$$

NLP Problem Summary

The NLP problem formulation is summarized as follows:

$$\textbf{NLP}: \quad \textit{minimize } (2.14), \ \textit{subject to: } (2.15) - (2.20), \tag{2.21}$$

with variables r_t non-negative, and $m_t, \theta_t, \theta_t^b, \bar{\theta}_t, \bar{\theta}_t^b$ real, $\forall t$.

Note that the NLP problem formulation includes the bilinear term $m_t\bar{\theta}_t$ in the J_t representation of the objective function (2.14b) and in constraint (2.15), as well as a non-convex quadratic inequality (2.18) that includes m_t^2. We treat these complicating terms and constraints with mixed integer linear relaxations in Section 2.4.

2.3.2 Estimation of the Expected IHCs

We employ a pre-processor which is run off line, i.e. prior to the solution of the ISD problem, to estimate parameters of a quadratic approximation of the expected IHCs, J_t, expressed in terms of state and decision variables $(m_t, r_t, \theta_{t-1}, \theta_{t-1}^b, \bar{\theta}_t)$ optimized in the ISD problem. The pre-processor executes the following tasks.

A reasonable seconds time scale regulation reserve deployment controller is used to simulate different l indexed, $l = 1, 2, ..., L$, values of expected IHCs, J_t^l, associated with similarly indexed state and decision variable levels $(m_t^l, r_t^l, \theta_{t-1}^l, \theta_{t-1}^{b,l}, \bar{\theta}_t^l)$.

A range of different state and decision variable levels, $l = 1, 2, ..., L$, that span the possible hourly state and decision variables which may be encountered in the solution of the ISD problem, is selected and for each of them the average IHCs, J_t^l,

are simulated.

We finally estimate the unknown parameters to fit the quadratic function of $(m_t^l, r_t^l, \theta_{t-1}^l, \theta_{t-1}^{b,l}, \bar{\theta}_t^l)$ to the J_t^l values.

The fully specified quadratic, which as noted below is a convex second-degree polynomial (with a single cross-term $m_t\bar{\theta}_t$), is then included in the ISD problem's objective function to integrate expected deployment costs incurred for each contemplated scheduling decision. Estimated parameters for a medium-sized office building are listed in Subsection 2.5.2.

The pre-processor simulation of the regulation deployment controller and the resulting intra-hour costs depend on intra-hour time scale dynamics that involve a new state variable representing the regulation reserve signal broadcasted by the ISO after 2, 4 or 6 second intervals (PJM, 2022a; ISO-NE, 2020; NYISO, 2022). In this work, we consider 4 second deployment intervals. Accounting for intra-hour time by τ, the regulation signal state is denoted by $S\tau$ and the control decision representing the modulation in the airflow is denoted by Δm_τ. $S\tau$ is associated with random dynamics described by a Markov Chain whose transition probabilities are estimated from historical data (Bilgin et al., 2016) and can be simulated accordingly. As detailed next, hourly state and decision variables are also associated with the short time scale τ in the intra-hour simulator of IHCs. The simulator of IHCs: (i) starts from the instance l of the HVAC state at the beginning of the hour, (ii) employs Monte Carlo to simulate the regulation signal dynamics to determine $S\tau$, (iii) applies a myopic controller to select a feasible Δm_τ to minimizes current period τ costs by trading off incremental impact on occupant thermal comfort and regulation signal tracking, and as such instantiates a myopic controller (iv) propagates state variables to the next τ period, and (v) sums period costs to arrive to the hourly total, J_t^l.

State Variables

Intra-hour simulator state variables include the supply air flow rate m_τ, inside air temperature, θ_τ and building mass temperature θ^b_τ, as well and the regulation signal $S\tau$. The regulation signal is modeled using the approach in (Bilgin et al., 2016). For simplicity of exposition we disregard a second variable used only in the simulation of the underlying Markov chain, focusing instead on $S\tau$ which takes values in the interval [-1, +1] and explicitly impacting each τ period cost. The variables m_τ, θ_τ and θ^b_τ are subject to the bound constraints, same as the ones in the scheduling problem, i.e., (2.20), which can be enforced as hard or soft constraints. Collectively, all the state variables are denoted as x_τ.

Control Variable

The change in air flow rate, Δm_τ, is the control variable which modifies the fan power consumption p^f_τ – see (2.5) – and allows the tracking of the reference regulation signal. The control variable is limited by the ramp rate and thus allowed to vary within a range $[\Delta m^-, \Delta m^+]$. Note that the compressor consumption is not impacted by the control variable since the compressor has a slow response and the net change in the control variable over a longer timescale can be considered negligible (due to energy-neutrality assumption).

State Dynamics

Dynamics for the air flow rate are given by $m_\tau = m_{\tau-1} + \Delta m_\tau$. Temperature dynamics for θ_τ and θ^b_τ are governed by equation (2.15) and (2.16), where the hourly interval t is replaced by the shorter, seconds scale interval τ. As noted, Regulation signal dynamics are simulated using the procedure described in (Bilgin et al., 2016) which employs a second variable associated with increasing or decreasing regulation signal.

Fixed hour-specific Information

Hour specific Information that is considered to be approximately constant during the simulation of IHCs includes the scheduling variables, i.e., air flow rate m_t, regulation reserve capacity r_t, inside air temperature θ_{t-1}, building mass temperature θ^b_{t-1}, and average inside air temperature $\bar{\theta}_t$. Other information that is intra-hour constant includes the temperature set-point $\hat{\theta}_t$ and allowable temperature range (θ^-_t, θ^+_t), as well as the thermal discomfort penalty factor $\tilde{\epsilon}$ and tracking error penalty factor κ. Ambient temperature θ^a_τ and heat Q_τ for each interval may either be considered deterministic, or stochastic, but in any case independent of ISD problem decisions.

Controller Objective

The intra-hour controller's objective is to minimize the cost of regulation reserve deployment tracking error plus the incremental intra-hour thermal discomfort. In particular:

$$\begin{aligned} \underset{\Delta m_t}{minimize} \; \underset{x_{\tau+1}|x_\tau}{\mathbb{E}} \Bigg[& \sum_\tau \kappa \big(p_t + S_\tau r_t - p_\tau\big)^2 \\ & + \tilde{\epsilon} \Big(2(\bar{\theta}_t - \hat{\theta}_t)(\theta_\tau - \bar{\theta}_t) + (\theta_\tau - \bar{\theta}_t)^2 \Big) \Bigg]. \end{aligned} \tag{2.22}$$

The first term penalizes (by κ) the square of the tracking error during intra-hour period τ, defined as the deviation from the regulation signal specified power consumption above or below the ISD scheduled power consumption $(p_t + S_\tau r_t)$ and the actual power output p_τ. The actual power output is calculated as $p_\tau = p^c_t + p^f_\tau$, with the compressor power p^c_t considered fixed and the fan power p^f_τ dependent on the control Δm_τ. The second term penalizes (by $\tilde{\epsilon}$) the incremental thermal discomfort incurred during intra-hour period τ.

Each l indexed instance of the pre-processor's estimation of IHCs has been solved

by Monte Carlo simulation of the Regulation reserve signal using the underlying Markov Process and the myopic or greedy heuristic controller described above. It is interesting to note that the myopic controller trades off tracking error cost for thermal comfort, which, when the internal air temperature is way above the maximum comfort set point, will be biased towards positive tracking errors resulting in AC power consumption that exceeds the level scheduled in the hourly ISD problem. The opposite will be true when the internal air temperature is below the max comfort set point. This bias may be welcome in the short term mitigating the impact of unplanned behavior such as open windows that bring hot air into the building, but may it may also interfere with pre-cooling strategies implemented by the ISD problem. Alternative controller designs aiming at discovering optimal or near optimal controllers would have to rely on much more cumbersome stochastic Dynamic Programming formulations of the IHCs pre-processor, and are detailed in the next chapter.

2.4 ISD Problem MIQP Formulation

In this section, we describe the relaxations that are used to handle the non-convexities in the NLP formulation of the ISD problem (in Subsection 2.4.1), and we summarize the resulting MIQP formulation (in Subsection 2.4.2).

2.4.1 Mixed Integer Linear Relaxations

Unlike existing works that deal with an NLP formulation of the ISD problem (Maasoumy et al., 2014; Liu et al., 2017; Vrettos et al., 2018b; Vrettos et al., 2018a; Qureshi and Jones, 2018; Blum et al., 2016; Fabietti et al., 2018; Pavlak et al., 2014), we consider an SOS2-based formulation, known as Piecewise Polyhedral Relaxation (PPR), to relax the bilinear and quadratic terms using piecewise, convex polyhedra (Sundar

et al., 2021). The PPR relaxation provides a piecewise convex hull representation for the feasible region of the NLP problem. It is more powerful than linearization techniques – e.g., a linear approximation of the bilinear or quadratic terms – and provides bounds to the NLP problem objective. It also offers a tighter relaxation and similar/better computational performance compared to other convex relaxations, such as the standard and piecewise McCormick relaxation (Nagarajan et al., 2019). In what follows, we detail the PPR relaxation employed to treat the bilinear term $m_t\bar{\theta}_t$ in the objective function and (2.15), and the quadratic term m_t^2 of the non-convex quadratic inequality (2.18).

Bilinear Term ($m_t\bar{\theta}_t$)

We consider an SOS2-based formulation, known as Piecewise Polyhedral Relaxation (PPR), to relax the bilinear term using piecewise, convex polyhedra (Sundar et al., 2021). PPR uses a spatial disjunction on the variables and a vertex representation for the bilinear term. It associates a spatial disjunction with each variable, m_t and $\bar{\theta}_t$, by defining a set of discretization points and intervals that partition the domain of the variable. Note that each pair of discretization points from the domain of variables m_t and $\bar{\theta}_t$ represents a vertex. Also note that for a given partition, the convex combination of the bilinear term $m_t\bar{\theta}_t$ values at the vertices represents a convex polyhedron, which is a relaxation of the bilinear term within that partition. Therefore, the convex combination of the bilinear term values at the vertices, considered disjunctively for different domain partitions using special ordered sets, helps characterize the piecewise, convex polyhedral relaxation of the bilinear term.

Let the feasible regions of variables m_t and $\bar{\theta}_t$ be partitioned into N and K intervals using $(N+1)$ and $(K+1)$ discretization points respectively. Let $[\mu_n, \mu_{n+1}]$ and $[\phi_k, \phi_{k+1}]$ denote the discretization values associated with n-th and k-th intervals of variables m_t and $\bar{\theta}_t$ respectively. Let $w_{t,n,k}$ denote a (non-negative) weight associated

with vertex (n, k), with

$$\sum_{n=1}^{N+1}\sum_{k=1}^{K+1} w_{t,n,k} = 1, \quad \forall t, \tag{2.23}$$

so that the domains of m_t and $\bar{\theta}_t$ are spanned as follows:

$$m_t = \sum_{n=1}^{N+1}\sum_{k=1}^{K+1} w_{t,n,k}\,\mu_n,\ \bar{\theta}_t = \sum_{n=1}^{N+1}\sum_{k=1}^{K+1} w_{t,n,k}\,\phi_k,\ \forall t. \tag{2.24}$$

Hence, the domain of the bilinear term $m_t\bar{\theta}_t$, which for ease of exposition is denoted by ξ_t, is spanned as follows:

$$\xi_t = \sum_{n=1}^{N+1}\sum_{k=1}^{K+1} w_{t,n,k}\,\mu_n\,\phi_k, \quad \forall t. \tag{2.25}$$

Let $x_{t,n}$ be a binary indicator denoting if the n-th interval of variable m_t is active, and $y_{t,k}$ a binary indicator denoting if the k-th interval of variable $\bar{\theta}_t$ is active. The following constraints ensure that only one interval per variable is active:

$$\sum_{n=1}^{N} x_{t,n} = 1, \quad \sum_{k=1}^{K} y_{t,k} = 1, \quad \forall t. \tag{2.26}$$

Using indicators $x_{t,n}$ and $y_{t,k}$, we ensure that weights $w_{t,n,k}$ that correspond to vertices of non-active intervals are forced to zero as follows:

$$w_{t,n,k} \le x_{t,n-1} + x_{t,n}, \quad n = 2, ..., N, \quad k = 1, .., K, \tag{2.27}$$

$$w_{t,1,k} \le x_{t,1}, \quad w_{t,N+1,k} \le x_{t,N}, \quad k = 1, .., K, \tag{2.28}$$

$$w_{t,n,k} \le y_{t,k-1} + y_{t,k}, \quad n = 1, ..., N, \quad k = 2, ..., K, \tag{2.29}$$

$$w_{t,n,1} \le y_{t,1}, \quad w_{t,n,K+1} \le y_{t,K}, \quad n = 1, ..., N. \tag{2.30}$$

Note that when the two adjoining intervals for a weight are not active, the *rhs* of its corresponding constraints in (2.27) and (2.29) will be zero, hence also forcing the *lhs* to zero. Constraints (2.28) and (2.30) account for the first and last discretization

$$\sum_{k=1}^{K+1} w_{t,n,k}\,\phi_k, \quad \forall t.$$

$$\phi_k, \quad \forall t.$$

$$= 1, \quad \forall t.$$

$$N, \quad k = 1, .., K,$$

$$_N, \quad k = 1, .., K,$$

$$\text{N}, \quad k = 2, ..., K,$$

$$\kappa, \quad n = 1, ..., N.$$

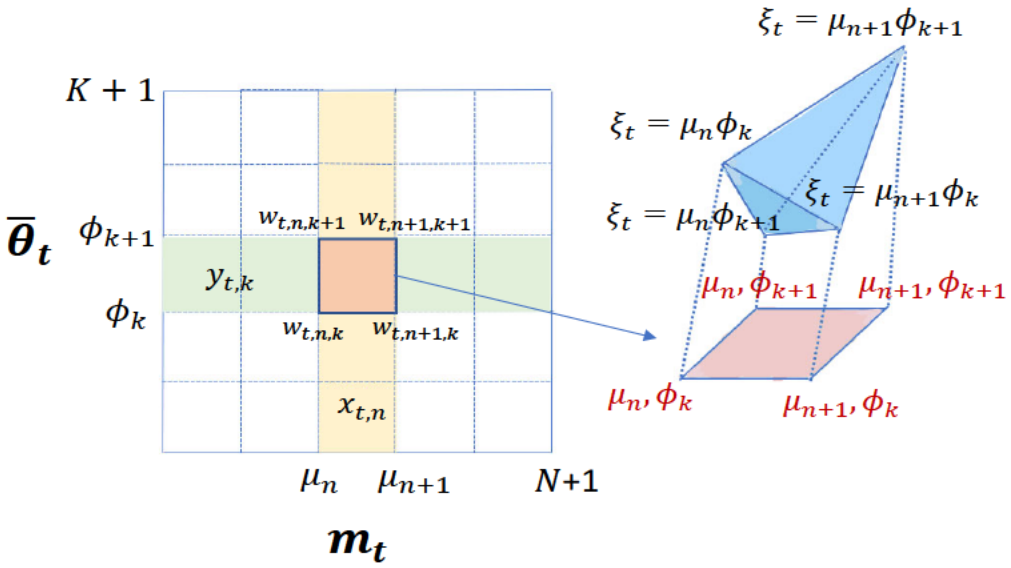


Figure 2·4: Piecewise Polyhedral Relaxation (PPR) — each domain partition has polyhedral representation of the feasible space.

point. Observe that constraints (2.26)-(2.30) enforce adjacency conditions on the weights akin to SOS2 constraints (Beale and Tomlin, 1970). Overall, (2.23)-(2.30) exploit the vertices to have for each domain partition, a convex polyhedral relaxation of the bilinear term — through (2.23)-(2.25), and use special ordered sets to have a disjunctive union of these polyhedral relaxations — through (2.26)-(2.30). PPR is visually illustrated in Fig. 2·4.

Quadratic Term (m_t^2) in (2.18)

We similarly treat the quadratic term m_t^2 using PPR and the aforementioned partition of variable m_t into N intervals and $(N+1)$ discretization points, yielding a total of $(N+1) \times (N+1)$ vertices. Let $u_{t,n,n'}$ be the weight associated with vertex (n, n'), with

$$\sum_{n=1}^{N+1} \sum_{i=1}^{N+1} u_{t,n,i} = 1, \quad \forall t. \tag{2.31}$$

The domain of the quadratic term m_t^2, which for ease of exposition is denoted by ζ_t, is spanned as follows:

$$m_t = \sum_{n=1}^{N+1} \sum_{n'=1}^{N+1} u_{t,n,n'}\, \mu_n, \ \forall t. \tag{2.32}$$

$$\zeta_t = \sum_{n=1}^{N+1} \sum_{n'=1}^{N+1} u_{t,n,n'}\, \mu_n\, \mu_{n'}, \quad \forall t. \tag{2.33}$$

Using the aforementioned indicator $x_{t,n}$, we ensure that weights $u_{t,n,n'}$ that correspond to vertices of non-active intervals are forced to zero as follows:

$$u_{t,n,n'} \leq x_{t,n-1} + x_{t,n},\ n = 2, ..., N,\ n' = 1, ..., N, \tag{2.34}$$

$$u_{t,1,n'} \leq x_{t,1},\ u_{t,N+1,n'} \leq x_{t,N},\ n' = 1, ..., N, \tag{2.35}$$

$$u_{t,n,n'} \leq x_{t,n'-1} + x_{t,n'},\ n = 1, ..., N,\ n' = 2, ..., N, \tag{2.36}$$

$$u_{t,n,1} \leq x_{t,1},\ u_{t,n,N+1} \leq x_{t,N},\ n = 1, ..., N. \tag{2.37}$$

Usefulness of the Relaxations

The above relaxations are not only tractable and tight, but they add further value in the context of the ISD problem since they allow some degree of non-linearity in the expected IHC expression — the corresponding non-linear terms in expected IHC are convexified as a by-product of relaxing the bilinear and quadratic terms in the constraints. This can prove to be even more useful in some ISD problem variants (see 2.4.3), where other closed-form expressions may be needed in thermal dynamic constraints. In addition, the convex structure makes the ISD problem much more amenable to distributed optimization (Sun and Dai, 2018), which can be useful in multi-zone problems. As such, our use of these relaxations is unique in the context of an HVAC system participating in a day-ahead electricity market.

2.4.2 MIQP Problem Summary

Replacing the bilinear term $m_t\bar{\theta}_t$ and the quadratic term m_t^2 by ξ_t and ζ_t, respectively, the NLP problem objective function and constraints are modified as follows.

The objective function term (2.14b) becomes:

$$p_t = \beta_1\xi_t + \beta_2 m_t + \alpha_2\zeta_t, \quad \forall t. \tag{2.14b$'$}$$

and similar changes are made in the functional form of J_t in (2.14e). We refer to this modified objective as (2.14$'$).

Constraints (2.15), (2.18) and (2.19) are written as follows:

$$\theta_t = \theta_{t-1} + \gamma_1 m_t + \gamma_2\xi_t + \gamma_3\bar{\theta}_t + \gamma_4\bar{\theta}_t^b + \gamma_{5,t}, \quad \forall t, \tag{2.38}$$

$$\alpha_1 m_t + \alpha_2\zeta_t - r_t \geq p^{f-}, \quad \forall t, \tag{2.39}$$

$$\alpha_1 m_t + \alpha_2\zeta_t + r_t \leq p^{f+}, \quad \forall t. \tag{2.40}$$

Hence, the ISD NLP problem is transformed to an MIQP problem which is summarized below:

$$\begin{aligned} &\textbf{MIQP:} \quad \textit{minimize} \ \ (2.14'), \\ &\textit{subject to:} \ (2.16), (2.17), (2.20), (2.23)\text{–}(2.40), \end{aligned} \tag{2.41}$$

with variables m_t, θ_t, θ_t^b, $\bar{\theta}_t$, $\bar{\theta}_t^b$, ξ_t, ζ_t real, r_t, $u_{t,n,n'}$, $w_{t,n,k}$ non-negative, and $x_{t,n}$, $y_{t,k}$ binary.

2.4.3 Problem Variants

We look at some aspects of the ISD problem that may be considered differently. For all these variants, the process to estimate the parameters in the approximation of the expected IHCs would remain the same, i.e., using the pre-processor, with trivial

modifications required for how the hourly state and decision variables are specified and the relevant aspect is modeled by the regulation reserve deployment controller. However, there would be some changes required within the ISD problem, which are detailed below.

Energy-Neutrality Assumption for the Regulation Signal

The implication of an energy-neutral regulation signal is that the hourly change in inside air temperature θ_t and building mass temperature θ_t^b has negligible dependence on regulation reserve capacity r_t — see (2.15) and (2.16). In case the energy-neutrality assumption does not hold, regulation reserve deployment may result in a different temperature at the end of the hour than the anticipated value. The expected difference between the anticipated and actual temperatures can be captured in (2.15) and (2.16) as closed-form analytic functions of decision and state variables. The functions can be determined, similarly to the expected IHCs, using the preprocessor.

Asymmetric Regulation Reserves

Some ISOs operate distinct up and down regulation reserve markets (CAISO, 2023), where the reserve capacity in each market can be different for the hour, i.e., asymmetric. The ISD problem can handle asymmetric reserves by writing out the regulation reserve capacity r_t as separate up and down variables, r_t^u and r_t^d. Note these variables appear in the objective function through (2.14c) and (2.14e), and in their corresponding regulation constraints — i.e., (2.18) and (2.19). In addition, just as with the absence of energy-neutral signal, asymmetric reserves require closed-form analytic functions in (2.15) and (2.16) to capture the expected difference between the anticipated and actual hourly temperatures.

Heating Operation

The ISD problem is equally applicable to the heating operation. In this case, the supply air temperature θ^s would be larger than the inside air temperature θ_t. Thus, (2.1) would be written as $\theta^s = \delta\theta + (1-\theta)\theta^a + \Delta\theta^h$, where $\Delta\theta^h$ would be the air temperature change at the heating coil. Moreover, the temperature bounds (θ_t^-, θ_t^+) may be different.

Use of Compressor for Providing Regulation Reserves

The provision of regulation reserves from the compressor can be considered by modifying the regulation constraints (2.18) and (2.19) to include the compressor's power consumption, p_t^c. This modification would introduce the complicating bilinear term, $m_t\bar{\theta}_t$, to these constraints — see (2.4) — which may potentially exacerbate the non-convexity issues in the NLP formulation. However, the MIQP formulation would be unaffected as our relaxations already encompass the bilinear term. Note that the compressor has significantly larger power consumption than the fan, typically ranging between 60%-80% of the nominal HVAC power consumption, hence the offered regulation reserve capacity and the corresponding revenues would be much higher when using the compressor for regulation. From a practical standpoint, it is crucial that the HVAC has the required auxiliary technology to modulate the compressor's speed, which in turn would help modulate the power consumption during regulation reserve deployment.

Multi-Zone HVAC System

The ISD problem is extendable to a multi-zone system, though the exact formulation depends on the fidelity of the considered model. For example, (Yang et al., 2020) considers additional valves in Fig. 2·2 to control the supply air flow rate to each zone, and would involve one bilinear term per zone (with different temperatures per

zone), and a quadratic term for the HVAC system. The thermal discomfort cost and IHCs would also be expressed as the sum of the respective costs per zone.

2.5 Case Study

In this section, we present a Case Study, which uses an office building on a summer day. In Subsection 2.5.1, we list the input data. In Subsection 2.5.2, we present the parameter values in the closed-form expression of the expected IHCs, J_t.

2.5.1 Input Data

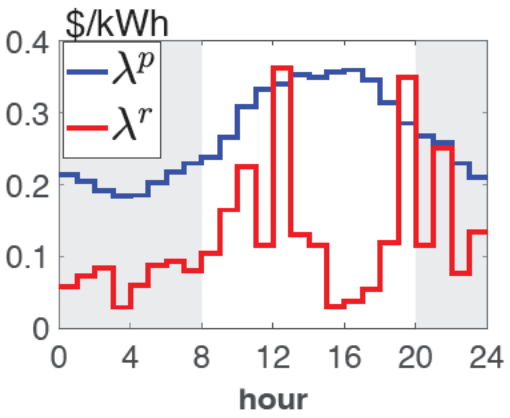
We consider a medium-sized office building using data from the DOE commercial reference building library ("US Department of Energy: Office of Energy Efficiency and Renewable Energy", 2020). The data for the building and HVAC parameters as well as temperature related preferences is given in Table 2.1. The energy and regulation reserve prices are obtained from the PJM market (PJM, 2022b), and the ambient temperature and irradiance (used along with occupancy and other miscellaneous sources to calculate heat Q_t) from the National Solar Radiation Database (NSRDB) (National Renewable Energy Laboratory (NREL), 2020) for a summer day (July 19th, 2017) in New Brunswick, New Jersey; they are shown in Fig. 2·5. The optimization problems are implemented with YALMIP toolbox in MATLAB (Löfberg, 2004) on a laptop Intel i7 − 6500U at 2.5 GHz with 16 GB RAM. The NLP and MIQP formulations are solved using IPOPT 3.12.10 (Wächter and Biegler, 2006), and GUROBI 9.0.1 (Gurobi Optimization, LLC, 2020), respectively. For the MIQP formulation, the number of discretization intervals used is $N = 10$ and $K = 4$.

2.5.2 Parameter Values in J_t

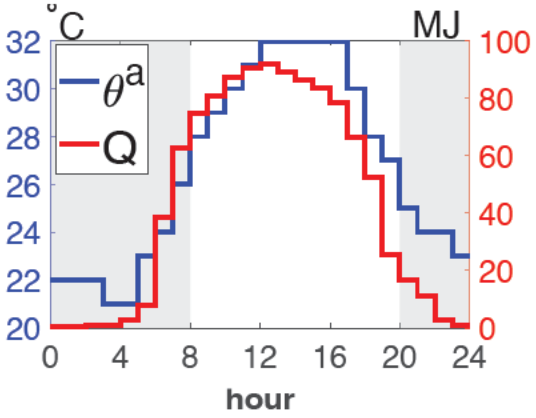
In this subsection, we derive the functional form of the expected IHCs, J_t, distinguishing between low and high-occupancy hours. We remind the reader that J_t is a

Table 2.1: Building, HVAC and Temperature Preference Data

Building	
Thermal Resistance of Building Mass, R^b	2×10^{-7} K.h/J
Thermal Resistance of Window, R^w	3.61×10^{-7} K.h/J
Thermal Capacitance of Inside Air, C	6.91×10^6 J/K
Thermal Capacitance of Building Mass, C^b	1.94×10^8 J/K
HVAC	
Supply Air Temperature, θ^s	17°C
Valve Position, δ	0.8
Supply Air Flow Rate, m_t	1–6 kg/s
Fan parameter, α_1	0.234 kJ/kg
Fan parameter, α_2	0.0975 (kJ.s)/kg^2
Coefficient of Performance, η	3.07
Preferences	
Temperature Set-point, $\hat{\theta}_t$	25°C
Low-Occupancy Hours	*(8pm–8am)*
Acceptable Temperature Range, (θ_t^-, θ_t^+)	(18°C, 28°C)
Discomfort Penalty Factor, ϵ_t	\$0.014/°C^2
High-Occupancy Hours	*(8am–8pm)*
Acceptable Temperature Range, (θ_t^-, θ_t^+)	(23°C, 27°C)
Discomfort Penalty Factor, ϵ_t	\$0.090/°C^2
Other	
Specific Heat Capacity of Air, c	1 kJ/(kg.K)



(a) Energy prices λ_t^p; regulation prices λ_t^r.



(b) Ambient Temperature, θ_t^a (left axis); Heat, Q_t (right axis).

Figure 2·5: Energy & Regulation Prices (PJM), and, Ambient Temperature & Heat (NSRDB, New Brunswick, NJ), for July 19th, 2017.

Table 2.2: Normalized Parameter Values in J_t

Parameters	**Occupancy (Hours)**	
	Low (8pm−8am)	High (8am−8pm)
m^2	–	1.518
$\bar{\theta}_t^2$	–	0.185
$m_t\bar{\theta}_t$	0.412	0.832
r_t	0.031	0.175
m_t	0.036	−1.212
θ_{t-1}	−0.221	−0.587
θ^b_{t-1}	−0.140	−0.146
$\bar{\theta}_t$	0.461	0.548

second-degree polynomial in terms of the hourly decision variables, m_t, r_t, and the state variables, $\theta_{t-1}, \theta^b_{t-1}, \bar{\theta}_t$, which captures the regulation signal tracking error and intra-hour incremental occupant discomfort costs.

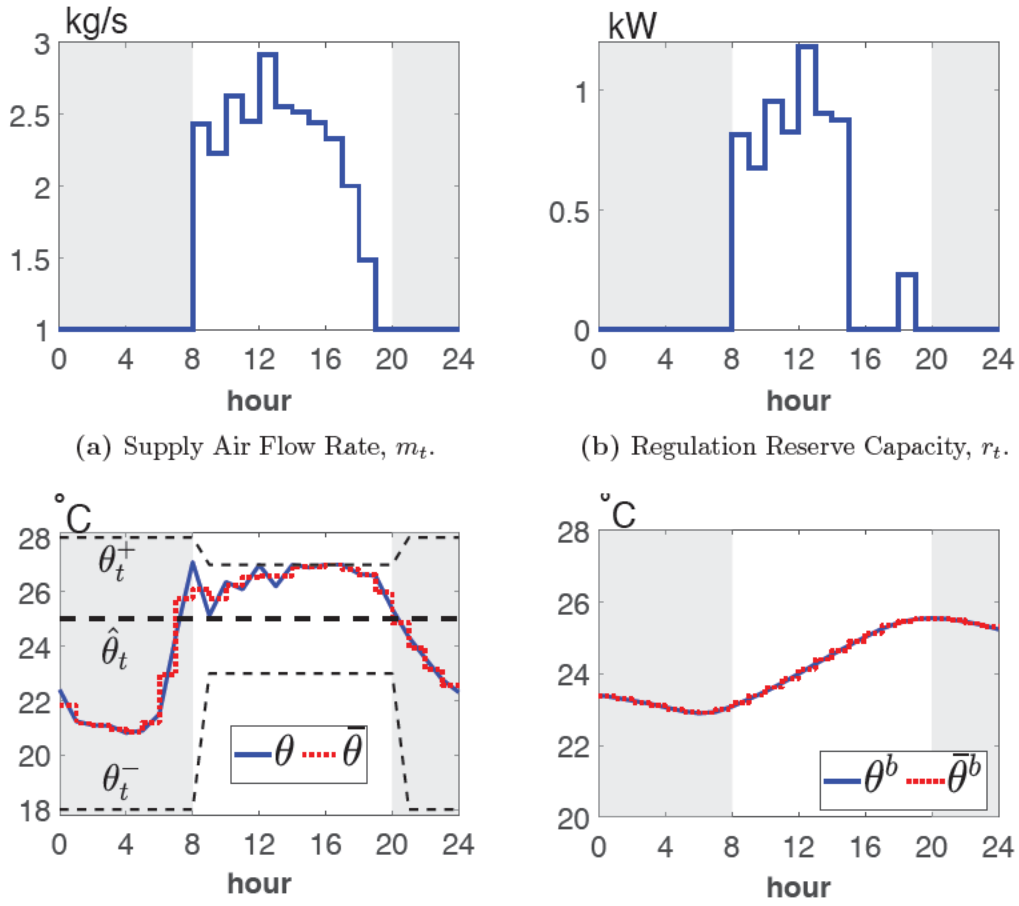
Table 2.2 presents normalized J_t parameters. The costs calculated using these parameters compare well to the actual costs from the preprocessor; R^2 is high ($>$ 0.90), and normalized root-mean-square errors are low ($< 7\%$). Note that a second-degree polynomial is required only for the high-occupancy hours — see values for m_t^2 and $\bar{\theta}_t^2$ in Table 2.2; however, both m_t and $\bar{\theta}_t$ appear in the bilinear term and in linear terms, which renders their interpretation not straightforward. Next, we proceed to some remarks for the parameters in Table 2.2.

First, we observe that the parameters of r_t are positive and their values are larger in high-occupancy hours. Indeed, positive parameters imply that if an HVAC increased the regulation reserve capacity, r_t, the IHCs would increase, because tracking the regulation signal would become more difficult, which would in turn result in higher regulation signal tracking error cost. Larger values in high-occupancy hours are explained by the fact that the stricter temperature bounds limit the HVAC's signal tracking ability.

Second, we observe that if an HVAC increased the supply air flow rate, m_t, the IHCs would increase in low-occupancy hours (positive parameters are associated with

terms containing m_t), but would either increase or decrease in high-occupancy hours (both positive and negative parameters are associated with terms containing m_t). During most of the low-occupancy hours, the ambient temperature θ_t^a and heat Q_t are smaller, thus a higher supply air flow rate would result in intra-hour inside air temperature that would be lower and further away from the preferred set-point, which would in turn result in higher intra-hour incremental occupant discomfort cost. During high-occupancy hours, however, a higher supply air flow rate would move the intra-hour inside air temperature either close to or away from the preferred set-point, depending on the HVAC's hourly operating point, which would in turn result in lower or higher, respectively, intra-hour incremental occupant discomfort cost.

Third, we observe that an increase in the average inside air temperature $\bar{\theta}_t$ would increase the IHCs, with a higher impact during high-occupancy hours (see the $\bar{\theta}_t^2$ term), implying that a higher average inside air temperature would considerably reduce the intra-hour temperature buffer (to the upper bound of the inside air temperature — notably stricter during high-occupancy hours) for performing signal tracking, and would hence increase the IHCs. On the other hand, an increase in the inside air temperature and the building mass temperature at the start of the hour, θ_{t-1} and θ_{t-1}^b, would decrease the IHCs. Collectively, the parameters associated with the temperature terms also account for the intra-hour incremental occupant discomfort cost. For example, if the starting temperatures were higher than the average inside air temperature, it would be possible to maintain the intra-hour temperature closer to the preferred set-point, without compromising signal tracking performance; hence, the actual discomfort incurred during that hour would be less than the average discomfort calculated in (2.14d), which would in turn result in negative IHCs.



(a) Supply Air Flow Rate, m_t.

(b) Regulation Reserve Capacity, r_t.

(c) Inside Air Temperature, θ_t.

(d) Building Mass Temperature, θ_t^b.

Figure 2·6: Decision and State Variables in the Base-Case.

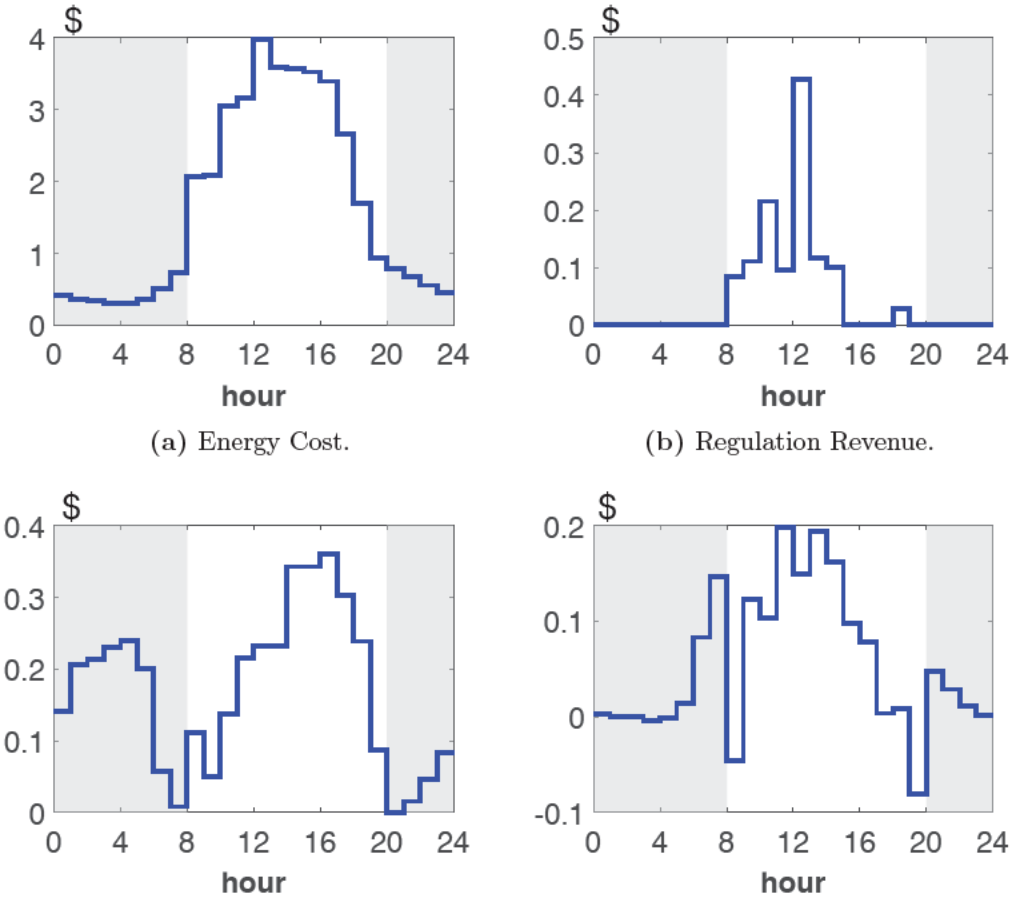
2.6 Numerical Results

In this section, we discuss the numerical results for the Case Study. In Subsection 2.6.1, we present a base-case using the NLP formulation. In Subsection 2.6.2, we compare the NLP and MIQP formulations. In Subsection 2.6.4, we analyze the impact of different exogenous factors.

2.6.1 Base-Case

In this subsection, we present the solution of the ISD problem, using the NLP formulation, for a base-case that refers to the input data and IHC parameter values listed in Section 2.5. The decision and state variable trajectories are shown in Fig. 2·6. The supply air flow rate, m_t (see Fig. 2·6a), rises above its minimum value only during the high-occupancy hours, when the high ambient temperature, θ_t^a, and heat, Q_t, as well as the stricter temperature bounds, (θ_t^-, θ_t^+), and larger discomfort penalty factor, ϵ_t, necessitate larger cooling. The regulation reserve capacity, r_t (see Fig. 2·6b), which is related to m_t through (2.18) and (2.19), also takes non-zero values during high-occupancy hours, when it is cost-efficient to do so. The inside air temperature, θ_t (see Fig. 2·6c), is lower and stays further away from the set-point, $\hat{\theta}_t$, during most of the low-occupancy hours; it is higher than the set-point during the high-occupancy (warmer) hours. The building mass temperature, θ_t^b (see Fig. 2·6d), follows a smoother trajectory, due to the large thermal capacitance of the building mass.

The trajectories for the different objective function terms in (2.14) are shown in Fig. 2·7. The energy cost (see Fig. 2·7a) is in general much higher during the high-occupancy hours, due to the higher energy prices and consumption. The regulation revenue (see Fig. 2·7b) is only earned during the high-occupancy hours, when the regulation reserve capacity has been scheduled. The aggregate discomfort cost (see Fig. 2·7c) during the high-occupancy hours is higher compared to the low-occupancy hours, due to the larger penalty factor, ϵ_t, even though the average inside air temperature stays mostly closer to the set-point. The expected IHC (see Fig. 2·7d) takes overall larger values during the high-occupancy hours, when the average inside air temperature, $\bar{\theta}_t$, is close to its upper bound and the offered regulation reserve capacity, r_t, is high (see, e.g., hours $10 - 15$ in Figs. 2·6b and 2·6c). When the inside



(a) Energy Cost.

(b) Regulation Revenue.

(c) Discomfort Cost.

(d) Expected IHC, (J_t).

Figure 2·7: Objective Function Terms in the Base-Case.

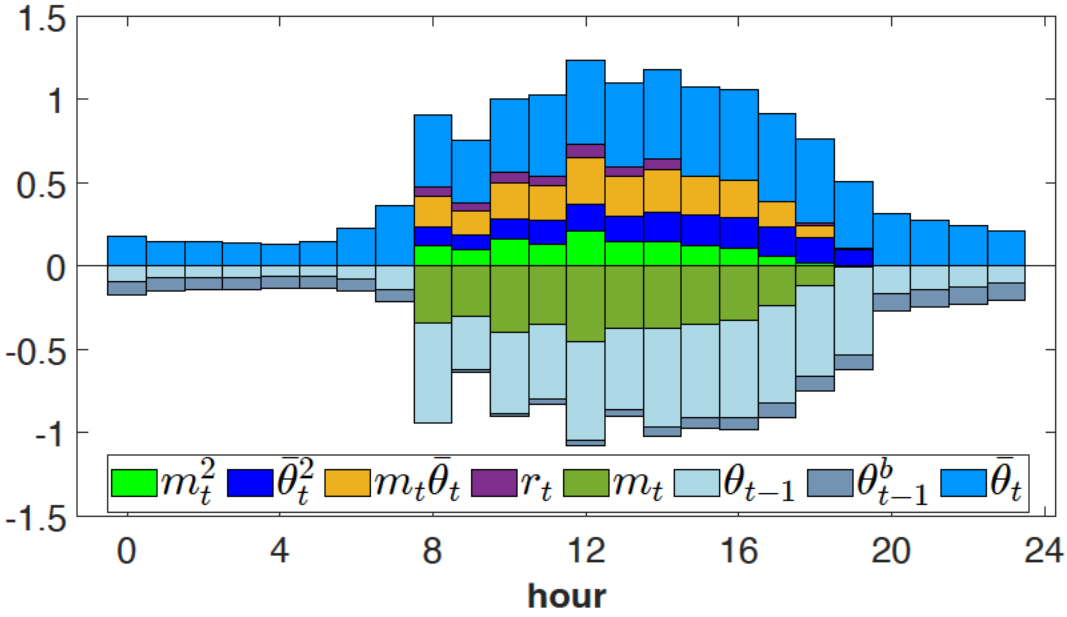
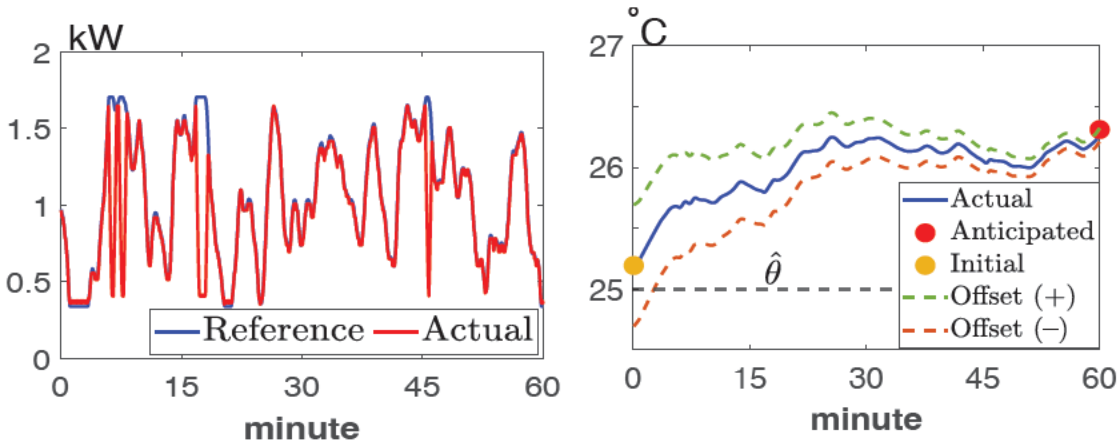


Figure 2·8: Breakdown of J_t in the Base-Case.

air temperature is decreasing and is further away from the bounds (see, e.g., hours 9 and 20 in Fig. 2·6c), the intra-hour temperature stays much closer to the set-point compared to the average inside air temperature, which results in intra-hour discomfort cost that is lower than the average discomfort calculated in (2.14d), and hence negative IHC.

The contribution of different terms towards the expected IHC, J_t, is shown in Fig. 2·8. The quadratic terms, m_t^2 and $\bar{\theta}_t^2$, add to the cost during the high-occupancy hours, whereas the cross-term, $m_t\bar{\theta}_t$, does not appear during the low-occupancy hours since the air flow rate, m_t, is at its minimum value (i.e., its normalized value is 0). Likewise, the terms m_t and r_t only appear during the high-occupancy hours (when m_t is above its minimum value). Finally, the temperature terms $\theta_{t-1}, \theta_{t-1}^b$, and $\bar{\theta}_t$ appear during all hours, with their net contribution being larger during the high-occupancy hours.

The regulation reserve deployment, which occurs under the control policy for which the expected IHCs is determined and incorporated into the ISD problem, is



(a) Regulation Signal Tracking.

(b) Intra-Hour Temperature.

Figure 2·9: Regulation Reserve Deployment during hour 10.

presented in Fig. 2·9. The deployment is shown for hour 10, where it occurs every 4 seconds by modulating the fan power consumption to track the reference regulation signal (Fig. 2·9a). The signal is tracked fairly well and the corresponding inside air temperature change over the hour is the same as the anticipated value (Fig. 2·9b), which is expected due to the symmetric regulation reserve provision and energy-neutral regulation signal. Notably, in Fig. 2·9b, we also illustrate two cases where the initial temperature was 0.5^oC lower/higher from the predicted temperature — see offset(+/-). Interestingly, the controller will mitigate this difference, which is reduced to 0.05^oC at the end of the hour.

2.6.2 NLP vs. MIQP

In this subsection, we compare the NLP and MIQP formulations, in terms of accuracy and computational performance.

We first consider the base-case, where the NLP and MIQP objective values are 43.681 and 43.666, respectively; their difference is about 0.03%. Notably, the MIQP objective value provides a lower bound (lb) to the NLP objective value. An upper bound (ub) can be straightforwardly calculated by plugging the MIQP solution into

the NLP objective, where m_t is determined using (2.15) and r_t using (2.18) and (2.19). The upper bound is 43.687 (i.e., about 0.013% higher than the NLP objective value). Note the bounds on the objective values are tight since the gaps between the bilinear and quadratic terms ($m_t\bar{\theta}_t$ and m_t^2), and their respective relaxations (ξ_t and ζ_t), are extremely small in the MIQP formulation. The average gap (expressed as percentage of the bilinear/quadratic term) between the bilinear term ($m_t\bar{\theta}_t$) and its relaxation (ξ_t) is 0.02%, whereas between the quadratic term (m_t^2) and its relaxation (ζ_t) is 0.11%. In terms of computational time, the NLP formulation took an order of magnitude more to solve (129 seconds) than the MIQP formulation (11 seconds).

We further compare the NLP and MIQP formulations by perturbing the base-case in terms of the regulation reserve price, ambient temperature and inside air temperature bounds. The considered cases are shown in Table 2.3, where: (i) we scale the regulation reserve price, λ_t^r, by a factor that varies from 0.5 to 5 in increments of 0.5; (ii) we decrease/increase the ambient temperature θ_t^a by a constant value, in steps of 0.5°C up to 3°C, and (iii) we modify the inside air temperature bounds (θ_t^-, θ_t^+), in steps of 0.25°C, up to 1°C. Table 2.3 shows that the average gap (and its standard deviation) between the bilinear/quadratic terms and their relaxations are small (less that 1%), and the bounds on the objective values are extremely tight (the largest difference between the *lb* and *ub* is 0.22%).

In terms of computational time, the results in Fig. 2·10 indicate that the NLP formulation is highly sensitive to input data and can take significantly longer to solve. For example, if the regulation reserve price λ^r is scaled by 0.5 (case R1) or if the ambient temperature θ^a is decreased/increased by 0.5°C (cases A1, A2) across all hours, the solution time for the NLP formulation is more than 1 hour. On the other hand, the MIQP formulation has a consistent solution time in the order of seconds. Notably, high computational times associated with the NLP formulation

Table 2.3: Tightness of Objective Bounds in Different Cases

Case	Parameters	Avg. Gap (std) in (%)		Diff. (%)
		$m_t\bar{\theta}_t$	m_t^2	$\frac{ub-lb}{lb}$
Base	$\lambda_t^r, \theta_t^a, (\theta_t^-, \theta_t^+)$	0.02 (0.04)	0.11 (0.27)	0.05
	Reg. Reserve Price			
R1	$\lambda_t^r \times 0.5$	0.02 (0.03)	0.15 (0.31)	0.05
R2	$\times 1.5$	0.03 (0.05)	0.33 (0.84)	0.07
R3	$\times 2.0$	0.03 (0.06)	0.24 (0.47)	0.06
R4	$\times 2.5$	0.05 (0.06)	0.26 (0.47)	0.10
R5	$\times 3.0$	0.04 (0.06)	0.27 (0.47)	0.10
R6	$\times 3.5$	0.04 (0.06)	0.27 (0.40)	0.12
R7	$\times 4.0$	0.05 (0.06)	0.19 (0.26)	0.17
R8	$\times 4.5$	0.04 (0.06)	0.21 (0.31)	0.18
R9	$\times 5.0$	0.04 (0.06)	0.21 (0.29)	0.22
	Ambient Temperature			
A1	$\theta_t^a - 0.5$	0.03 (0.05)	0.21 (0.38)	0.07
A2	$+ 0.5$	0.02 (0.04)	0.11 (0.28)	0.04
A3	$- 1.0$	0.04 (0.07)	0.35 (0.63)	0.11
A4	$+ 1.0$	0.03 (0.06)	0.19 (0.40)	0.06
A5	$- 1.5$	0.06 (0.09)	0.47 (0.88)	0.13
A6	$+ 1.5$	0.03 (0.05)	0.28 (0.50)	0.08
A7	$- 2.0$	0.07 (0.10)	0.46 (0.77)	0.14
A8	$+ 2.0$	0.03 (0.06)	0.38 (0.57)	0.11
A9	$- 2.5$	0.06 (0.08)	0.39 (0.57)	0.13
A10	$+ 2.5$	0.04 (0.06)	0.49 (0.91)	0.12
A11	$- 3.0$	0.05 (0.07)	0.34 (0.56)	0.13
A12	$+ 3.0$	0.05 (0.07)	0.41 (0.84)	0.11
	Temperature Bounds			
B1	$\theta_t^- - 0.25$, $\theta_t^+ + 0.25$	0.04 (0.07)	0.21 (0.36)	0.08
B2	$- 0.50$, $+ 0.50$	0.08 (0.10)	0.34 (0.46)	0.14
B3	$- 0.75$, $+ 0.75$	0.10 (0.13)	0.43 (0.56)	0.16
B4	$- 1.00$, $+ 1.00$	0.10 (0.14)	0.45 (0.60)	0.16
B5	$+ 0.25$, $- 0.25$	0.01 (0.04)	0.11 (0.27)	0.04
B6	$+ 0.50$, $- 0.50$	0.01 (0.02)	0.18 (0.31)	0.06
B7	$+ 0.75$, $- 0.75$	0.01 (0.01)	0.18 (0.28)	0.06
B8	$+ 1.00$, $- 1.00$	0.01 (0.03)	0.15 (0.33)	0.04

are undesirable, especially in a market-based setting, since the HVAC system may be required to solve the ISD problem multiple times before market clearing (Caramanis et al., 2016). We also note that the solution time for the deployment controller was about 3ms, which is appropriate for the fast, few-second deployment of regulation reserve. The pre-processor solves many 3ms problems, however, this time is not critical as it can be done offline to estimate the parameters of the expected IHCs.

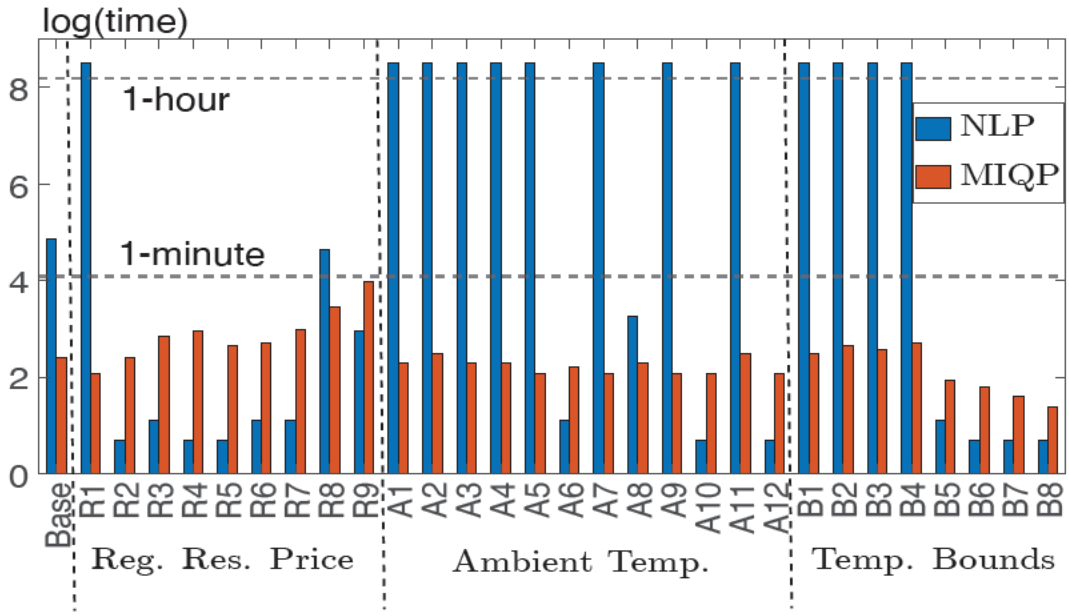
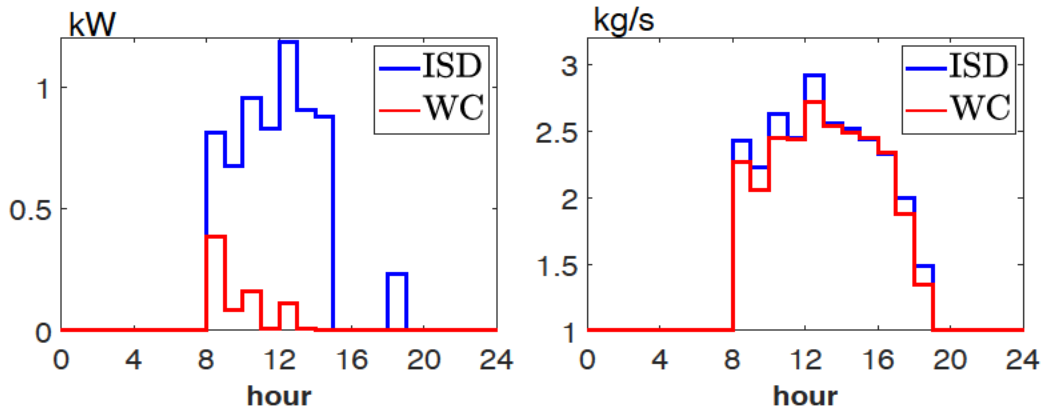


Figure 2·10: Computational Time in Different Cases for the NLP and MIQP Formulations.

2.6.3 Comparison with Existing Approaches

In this subsection, we compare our ISD problem to existing approaches that consider hard constraints, and worst-case-based approaches (Liu et al., 2017; Vrettos et al., 2018b) and (Maasoumy et al., 2014), which, essentially, consider a worst case deployment of regulation reserve (in either direction) and hard temperature limits, without allowing the HVAC to overbid. A comparison for the Base Case with our approach is shown in Fig. 2·11. Evidently, (Maasoumy et al., 2014) leads to conservative esti-

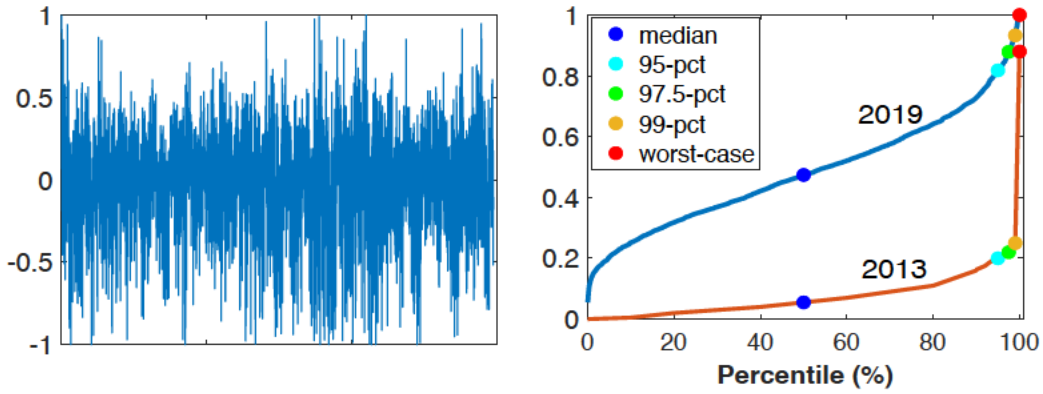


(a) Regulation Reserve Capacity, r_t. (b) Supply Air Flow Rate, m_t.

Figure 2·11: Decision Variables for the ISD and Worst-Case (WC) problems in the Base-Case.

mates for the regulation reserve capacity (Fig. 2·11a), although the differences in the supply air flow rates are small (Fig. 2·11b).

The conservativeness can potentially be overcome by offering more capacity than allowed by the HVAC's capability, based on regulation signal statistics. (Liu et al., 2017) and (Vrettos et al., 2018b) observe the average change in the signal over 15 minutes to be within 10% and 25% of the capacity, and thus overbid the HVAC capability. However, these numbers were observed using the PJM RegD signal characteristics from $2012-2013$, which no longer hold. PJM has made significant changes since then. Based on recent data, the potential for overbidding is drastically reduced, and the regulation reserve capacity offers should be much more in line with the HVAC capability. Fig. 2·12 illustrates the average change over 15-minute intervals for the PJM RegD signal in July 2019 (Fig. 2·12a), and the cumulative distribution of the signal's energy content over 15 minute intervals (Fig. 2·12). Fig. 2·12b also illustrates the comparison with the data provided (Vrettos et al., 2018b). Indicatively, in 2013, the average change requested by RegD over 15 minutes was less than 25% of the regulation capacity with 97.5% probability, whereas in 2019, it was less than 88%



(a) Average Change. (b) Energy Content.

Figure 2·12: PJM RegD Signal over 15-minute Intervals for July, 2019.

of the regulation capacity with the same probability. Notably, 88% was the worst case in (Vrettos et al., 2018b), which was deemed to be too conservative. Hence, the characteristics of the signal have changed, which do not favor the application of approaches such as (Liu et al., 2017) and (Vrettos et al., 2018b) that would still lead to conservative regulation reserve provision.

2.6.4 Sensitivity Analysis

In this subsection, we perform sensitivity analysis *w.r.t.* the regulation reserve capacity prices, the ambient temperature, and the inside air temperature bounds.

Regulation Reserve Prices

Fig. 2·13 compares the base-case with case R9, in which the regulation reserve prices are 5 times higher. Unsurprisingly, case R9 offers, in general, a higher amount of regulation reserve capacity r_t (see Fig. 2·13a). Case R9 also exhibits an increase in the supply air flow rate, m_t (see Fig. 2·13b, which follows a pattern that is similar to Fig. 2·13a). Indeed, m_t and r_t are related by (2.18) and (2.19), and the capability for offering r_t is maximized when the fan energy consumption, p_t^f, expressed

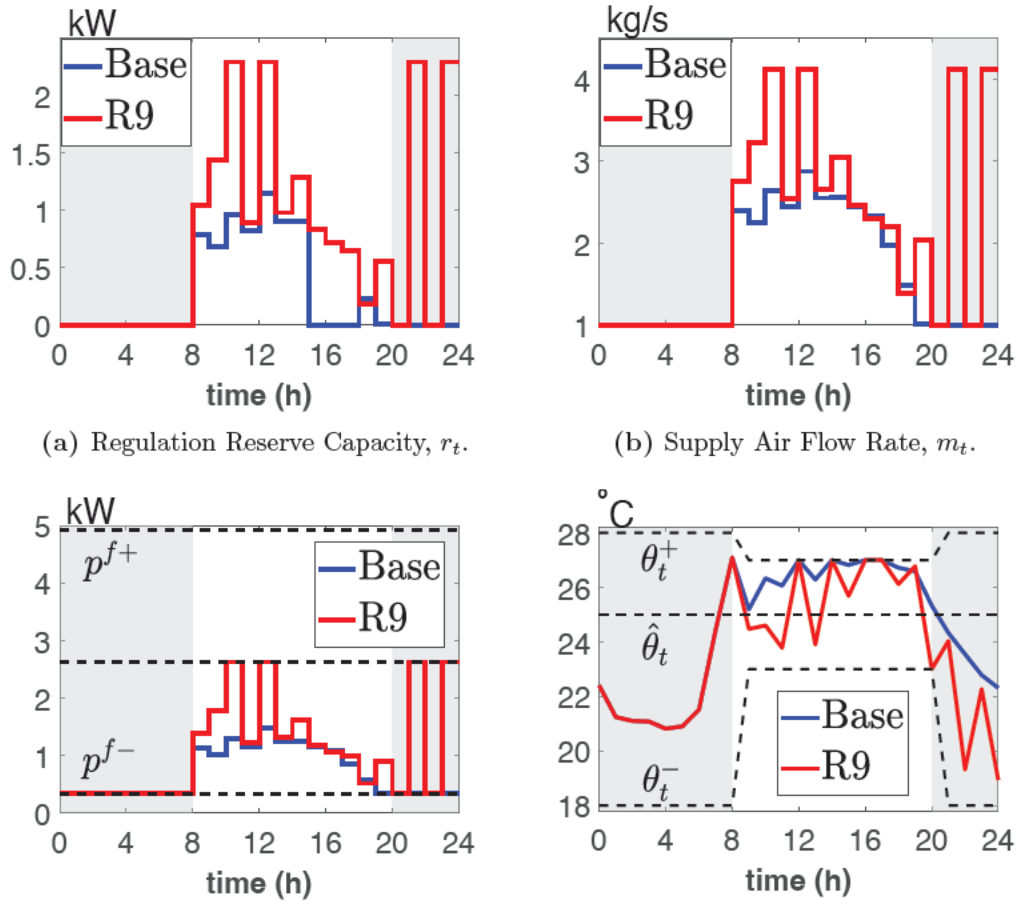
as $\alpha_1 m_t + \alpha_2 m_t^2$, equals the mid-point $(p^{f-} + p^{f+})/2$. For example, see Fig. 2·13c, hours 11, 13, 22, and 24, during which p_t^f is at the mid-point and the HVAC provides the maximum amount of r_t (see Fig. 2·13a). Notably, in the base-case, p_t^f is lower than the mid-point (see Fig. 2·13c), and hence, an increase of m_t (up to the mid-point) in case R9 enables the provision of higher r_t. Evidently, the increase of m_t translates to more cooling and, in general, a lower inside air temperature θ_t (see Fig. 2·13d). Hour 20 (the last high-occupancy hour of the day) is an interesting hour, since θ_t drops to its lower bound, θ_t^-. Taking a closer look, we note that hour 20 has a similar (high) regulation reserve price to hour 13 (see Fig. 2·13a), however, the amount of offered r_t is much lower in hour 20 compared to hour 13. The reason is that if m_t were to increase more in hour 20 (which would enable higher r_t), θ_t would drop below its lower bound (which is not allowed).

Ambient Temperature

Fig. 2·14 compares the base-case with case A12, in which the ambient temperature, θ_t^a, is increased by 3°C. Evidently, a warmer day would require more cooling, i.e., a higher supply air flow rate, m_t (see Fig. 2·14b), which in turn enables the provision of more regulation reserve capacity, r_t (see Fig. 2·14a). The fan power consumption, p_t^f, increases due to the higher m_t (see Fig. 2·14c), however, since the energy prices dominate the regulation reserve prices, p_t^f increases only to the extent required to keep the inside air temperature θ_t (which is in general elevated compared to the base-case) within the bounds during the high occupancy hours (see Fig. 2·14d).

Temperature Bounds

Stricter temperature bounds (θ_t^-, θ_t^+) require an increase in the supply air flow rate, m_t, to satisfy thermal discomfort preferences during the high-occupancy hours when the inside air temperature, θ_t, is close to the upper bound. Because the HVAC fan



(a) Regulation Reserve Capacity, r_t.

(b) Supply Air Flow Rate, m_t.

(c) Fan Energy Consumption, p_t^f.

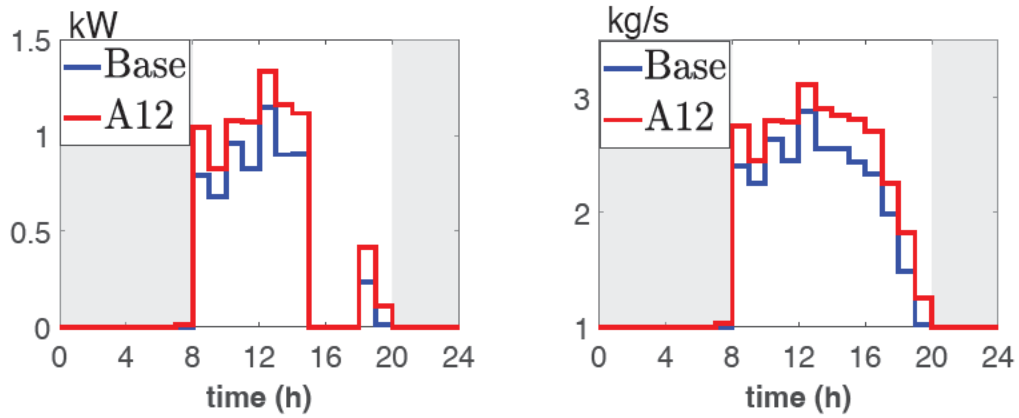
(d) Inside Air Temperature, θ_t.

Figure 2·13: Impact of Regulation Reserve Prices, λ_t^r.

operation is in the lower half of its energy consumption, higher m_t due to stricter bounds also leads to higher regulation reserve capacity, r_t. For example, case B8, which has stricter bounds by 1°C, i.e., $(\theta_t^- + 1, \theta_t^+ - 1)$, results in the provision of about 21% more regulation reserve capacity compared to the base-case.

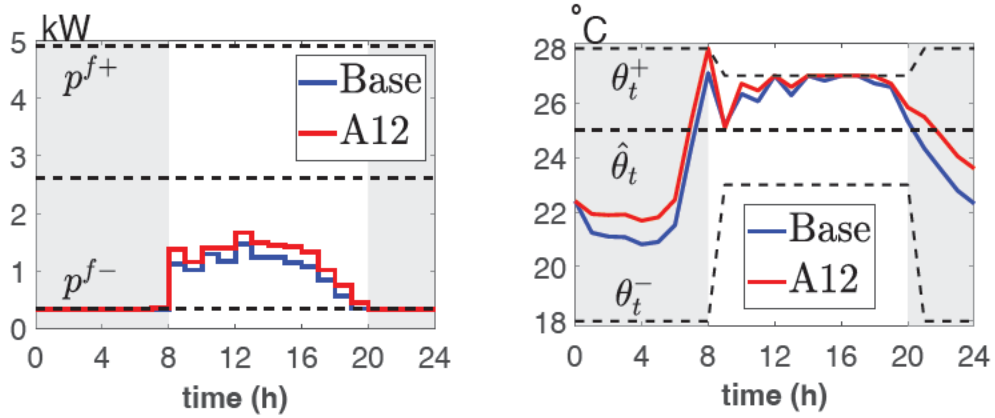
2.6.5 Additional Results

In this subsection, we provide additional numerical results for seven summer days from 2017, which span different values for ambient temperature, irradiance, energy and regulation reserve prices, as shown in Fig. 2·15.



(a) Regulation Reserve Capacity, r_t.

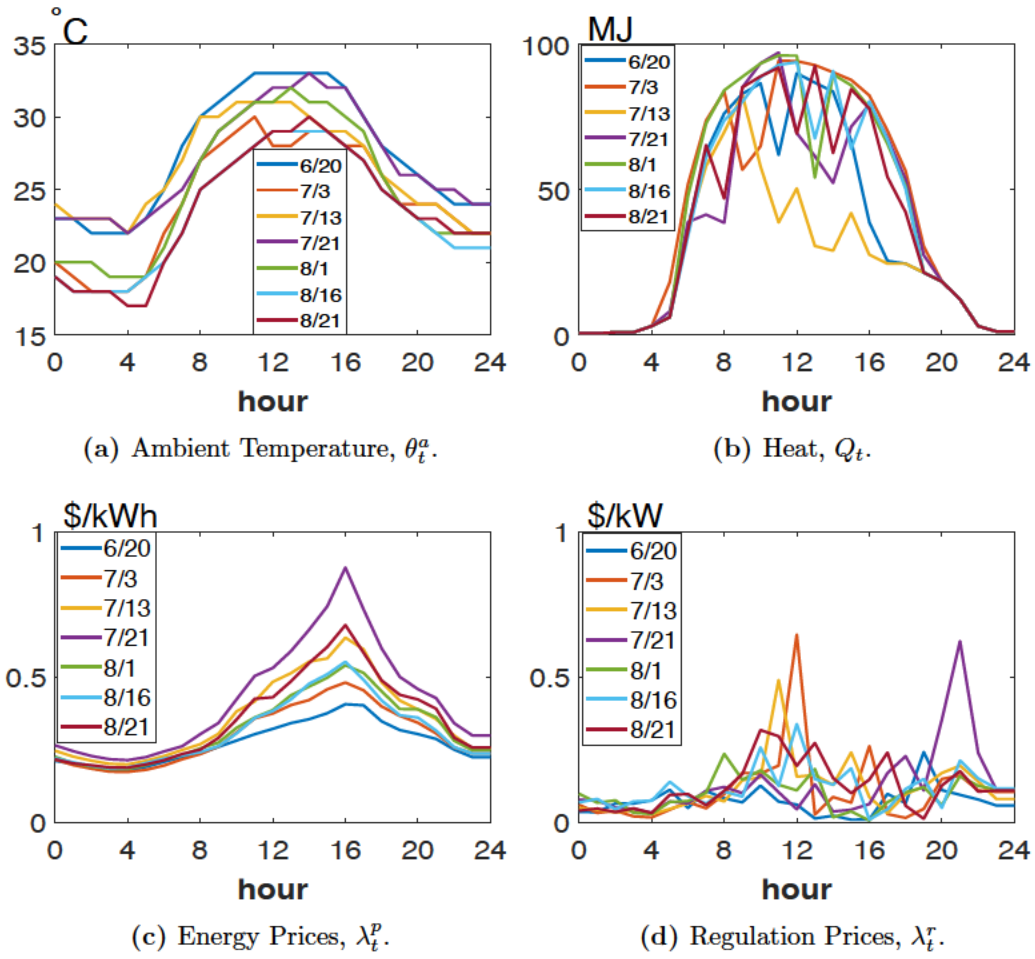
(b) Supply Air Flow Rate, m_t.



(c) Fan Energy Consumption, p_t^f.

(d) Inside Air Temperature, θ_t.

Figure 2·14: Impact of Ambient Temperature, θ_t^a.



(a) Ambient Temperature, θ_t^a.

(b) Heat, Q_t.

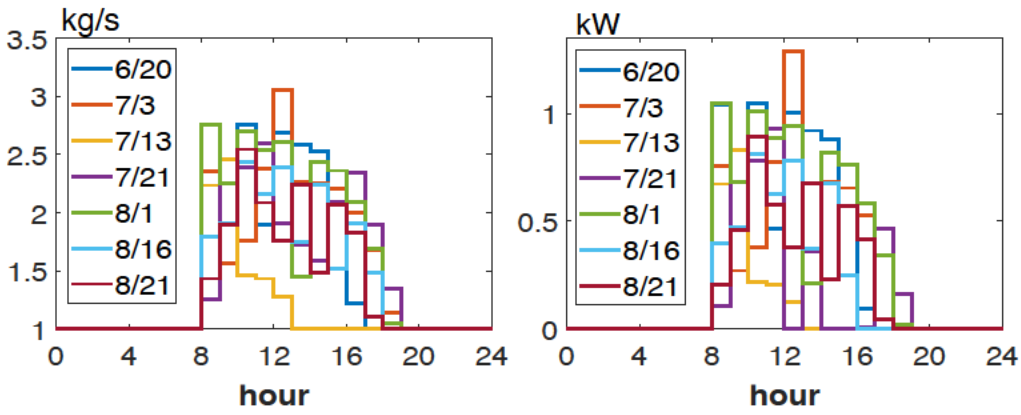
(c) Energy Prices, λ_t^p.

(d) Regulation Prices, λ_t^r.

Figure 2·15: Ambient Temperature, Irradiance and Prices for Other Representative Days.

We estimated the normalized J_t parameter, and calculated the costs which compared well to the actual costs from the pre-processor; average R^2 about 0.90, and normalized root-mean-square errors about 7%. Computational times for the MIQP formulation ranged from 2.2 to 5.0 seconds, indicating a robust performation of the proposed relaxation. The bounds of the relaxation were also tight with the difference between the upper and lower bounds ranging between 0.06% and 0.15%.

The energy consumption and offered regulation reserve capacity are presented in Fig. 2·16. We make a few interesting observations that illustrate the various fac-



(a) Supply Air Flow Rate, m_t. (b) Regulation Reserve Capacity, r_t.

Figure 2·16: Supply Air Flow Rate and Regulation Reserve Capacity for Other Representative Days.

tors at play in determining the regulation reserve capacity. For example, July 21st and August 1st are similar days in terms of the cooling requirements (similar ambient temperatures and heat). However, lower energy prices and relatively higher regulation prices in August 1st result in the HVAC offering significantly more regulation capacity (aggregate 7.3 kW over the entire day compared with 3.5 kW in July 21st). Interestingly, in July 21st, the HVAC is unable to profit from the spike in regulation prices during hour 21 because it is already operating at its minimum fan power consumption. On the other hand, July 3rd has moderate cooling requirements and relatively smaller energy prices, and hence, it is able to offer the highest regulation capacity observed in single hour during hour 13 profiting from the spike in the regulation reserve price.

2.7 Conclusions

In this chapter, we considered an integrated scheduling and deployment problem for optimizing the hourly energy consumption and regulation reserve capacity of an HVAC system in a day-ahead market. Our proposal internalized the intra-hour trade-

off between regulation signal tracking error and incremental occupant discomfort into the hourly decision-making, by using a pre-processor estimated closed-form representation of the expected IHCs. Computationally efficient piecewise convex relaxations for the complicating non-convex terms in the original NLP problem were also presented, which resulted in an MIQP problem formulation that provided tight bounds for the original NLP problem's objective. We also elaborated on the different variants to which our approach may be applied, and demonstrated its efficacy and relevance pertaining to existing approaches.

Chapter 3

Stochastic Optimal Control for Energy and Regulation Reserve Deployment in Buildings using HVAC

In this chapter, we present optimal control formulations of the deployment problem, including a discounted-cost infinite horizon stochastic Dynamic Program (DP) and a RL formulation. Each formulation minimizes the regulation signal tracking error and utility loss (thermal discomfort) over an hour, while also respecting the intra-hour temperature bounds (see Fig. 3·1). The solution to the formulation provides the dynamic time-invariant optimal deployment control policy and the associated cost-to-go, for a specified hourly energy consumption baseline, regulation reserve capacity and hourly average inside air temperature. Note that the value of the calculated cost-to-go is in fact the expected IHC for the specified energy, reserve capacity and temperature. We solve the formulations and evaluate the optimal policy for various cases, in addition to providing insights about its structure. We further identify the usefulness of the proposed approach in comparison to the classical controllers.

3.1 Introduction

Regulation reserve deployment for HVAC systems has previously been considered in the literature where its potential has been demonstrated using both numerical simulations and experimental test-beds (MacDonald et al., 2014; Hao et al., 2013;

Nomenclature	
Sets and Indices	
t	hour index, (Δt denotes interval duration)
τ	second-interval index ($\Delta\tau$ denotes interval duration)
Parameters	
c	specific heat capacity of air
R^b, R^w	thermal resistance (building mass, window)
C, C^b	thermal capacitance (inside air, building mass)
Q	heat transfer due to occupancy and irradiance
θ^a, θ^s	ambient temperature, supply air temperature
α_1, α_2	coefficient of fan power consumption
κ	tracking error penalty factor
ϵ'	temperature utility loss (thermal discomfort) penalty factor
γ	temperature limit penalty factor
ρ	discounting factor
$\hat{\theta}_t$	hourly set-point temperature
p_t	hourly average energy consumption baseline
r_t	hourly regulation reserve capacity
$\bar{\theta}_t$	hourly average temperature
θ_t^-, θ_t^+	hourly lower temperature limit, upper temperature limit
$\boldsymbol{v}$, $\mathbf{w}$	parameters for policy, value functions
$\alpha^{\boldsymbol{v}}$, $\alpha^{\mathbf{w}}$	learning rates for policy, value functions
Variables	
control	
Δm	change in supply air flow rate
state	
m	supply air flow rate
θ	inside air temperature
θ^b	building mass temperature
y	regulation signal value
d	regulation signal direction
Other	
x	state
u	control
$g(x,u)$	period cost
$J_\rho(x)$	discounted cost function
$J_{t(m_t, r_t, \theta_{t-1}, \theta^b_{t-1}, \bar{\theta}_t)}$	expected intra-hour cost
$\bar{J}$	average period cost
$h(x)$	differential cost function
$\mathcal{Q}(x,u)$	Q-function
$V(x)$	value function
$\mu(x)$	optimal policy function

Hao et al., 2014; Lin et al., 2015; Zhao et al., 2013; Qureshi and Jones, 2018; Liu et al., 2017; Vrettos et al., 2018b; Wang et al., 2022; Su and Norford, 2015). As mentioned in Chapter 1, this encompasses a range of control approaches, including open-loop, PI, PID, and switched controllers. The prior works provide a promising

solution but do not address certain aspects of regulation reserve deployment problem.

Firstly, there has been no effort to derive the optimal control policy, which can be used by the HVAC system to optimally deploy regulation reserves. The existing works use classical controllers which often require cumbersome manual gain tuning, based on trial and error approaches and assuming prior domain knowledge. Thus, the controller design may be prone to errors and not always yield optimal or near-optimal control policies. An optimal control policy is more representative of the deployment costs associated with regulation signal tracking error and utility loss (thermal discomfort).

Secondly, there has been little insight into the structure of the control policies, and of the total cost incurred over the horizon of the deployment problem. Such an approach would be beneficial in capturing the impact of different variables towards regulation reserve deployment, and in understanding the interaction between the control/cost of the deployment problem and the fixed parameters from the scheduling problem. Hence, it can help map the seconds timescale deployment costs to an hourly

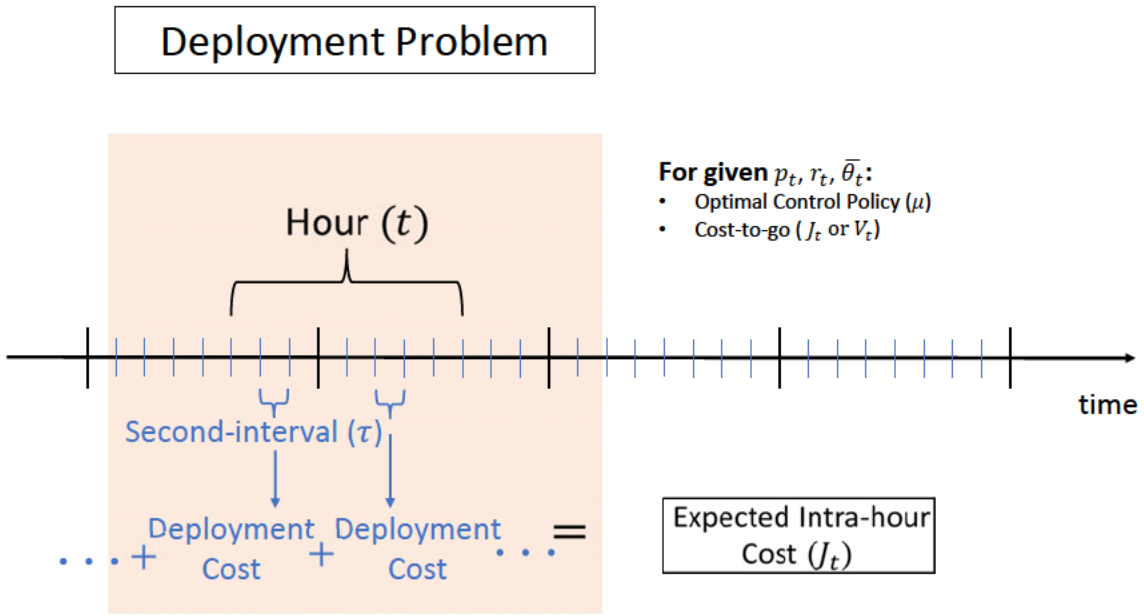


Figure 3·1: Optimal Control Formulation of the Deployment Problem — Provides Optimal Control Policy and Cost-to-go over the Hour.

cost that may be used to make better scheduling decisions, whilst also making the deployment problem more interpretable (Aslam et al., 2024). Overall, the approach may benefit not only the building HVAC systems but also the aggregators and grid operators.

Thirdly, the current approaches require knowledge of the building models, and may further entail extensive trial and error testing to be effective. Therefore, they are not easily replicable to newer buildings where the model parameters are unknown, and may also not adapt to the changes in a building model that occur over time, e.g., due to degradation and/or modification. Hence, there is a need for versatile and adaptable data-driven approaches that can be applied more broadly. Some past works have used data-driven methods, not all necessarily model-free approaches, for addressing HVAC system related problems, e.g. thermal comfort satifaction (Yang et al., 2021; Gao et al., 2020; Raman et al., 2020; Vázquez-Canteli et al., 2017; Ding et al., 2019), cost and energy reduction (Wei et al., 2017; Lissa et al., 2020; Zhang et al., 2018; Yu et al., 2020; Murugesan et al., 2020; Xu et al., 2020; Faddel et al., 2020b; Faddel et al., 2020a), resilience (Raman et al., 2021), and demand response (Azuatalam et al., 2020). However, none have done so in the context of regulation reserve deployment, which is typically more granular than the considered problems.

We look to fill some of the aforementioned gaps in the literature. To this end, we formulate the regulation reserve deployment problem for an HVAC system as discounted-cost infinite-horizon stochastic DP and RL formulations. We first solve the DP to obtain the optimal control policy and the associated costs in the form of a look-up table, and then thoroughly assess the performance of the control policy for a range of use-cases. We further compare the policy against some existing control approaches, and illustrate how the optimal control accurately considers the costs from a consumer's perspective, and provides the correct signals to the market operator

regarding the available long-term flexibility from HVACs. In addition, we analyze the obtained data (in conjunction with Chapter 2 results) to derive parametric forms of the policy and cost functions that are interpretable, and enable the RL formulation to be solved efficiently.

The remainder of the chapter is organized as follows. In Section 3.2, we present the model for building thermal dynamics, HVAC power consumption and regulation signal dynamics. We then model the regulation reserve deployment problem as a Markov Decision Process (MDP) in Section 3.3. The DP and RL formulations and their respective solution approaches are presented in Section 3.4. In Section 3.5, relevant numerical results and discussions are presented, followed by conclusions in Section 3.6.

3.2 Model

We model the building thermal dynamics, HVAC power consumption and reference regulation signal, in order to address the energy and regulation reserve deployment problem. During deployment, the power consumption of HVAC is modulated in $\Delta\tau$ second intervals over the hour t to follow the reference regulation signal from the grid operator for a specified hourly energy consumption baseline and regulation reserve capacity. Note that the power modulation is achieved by changing the HVAC supply air flow rate, which also impacts the building thermal dynamics.

3.2.1 Building Thermal Dynamics

The building is represented by a second-order Resistance-Capacitance model, using which the discrete-time dynamics at interval τ for inside air temperature θ_τ and building mass temperature θ^b_τ are compactly written as follows (Atam and Helsen,

2016):

$$\begin{aligned}\begin{bmatrix}\theta_{\tau+1}\\ \theta^b_{\tau+1}\end{bmatrix} =&\mathbf{A_1}\begin{bmatrix}\theta_\tau\\ \theta^b_\tau\end{bmatrix} + \mathbf{A_2}\left[m_\tau + \Delta m_\tau\right] + \\ &\mathbf{A_3}\begin{bmatrix}\theta_\tau\\ \theta^b_\tau\end{bmatrix}\left[m_\tau + \Delta m_\tau\right] + \mathbf{A_4}\begin{bmatrix}\theta^a_\tau\\ Q_\tau\end{bmatrix},\end{aligned} \tag{3.1}$$

where,

$$\mathbf{A_1} = \begin{bmatrix}1 - \frac{1}{R^bC} - \frac{1}{R^wC} & \frac{1}{R^bC}\\ \frac{1}{R^bC^b} & 1 - \frac{2}{R^bC^b}\end{bmatrix}, \quad \mathbf{A_2} = \begin{bmatrix}\frac{c\theta^s}{C}\\ 0\end{bmatrix},$$
$$\mathbf{A_3} = \begin{bmatrix}-\frac{c}{C} & 0\\ 0 & 0\end{bmatrix}, \quad \mathbf{A_4} = \begin{bmatrix}\frac{1}{R^wC} & \frac{1}{C}\\ \frac{1}{R^bC^b} & 0\end{bmatrix}.$$

Note above that m_τ is the initial supply air flow rate which can be changed by Δm_τ to adjust the inside air temperature. Further, θ^a_τ and Q_τ can be considered as exogenous disturbances, where θ^a_τ is the ambient air temperature and Q_τ is the heat from miscellaneous sources such as occupancy and irradiance. The matrices above contain different parameters, which include the fixed supply air temperature θ^s, specific heat capacity of air c, as well as thermal resistance (building mass R^b, windows R^w) and thermal capacitance (inside air C, building mass C^b).

3.2.2 Power Consumption

HVAC has two components which contribute towards the power consumption, namely the fan and the compressor. We only consider regulation reserve deployment through the fan since it has fast response and can track the regulation signal fairly well (Beil et al., 2016). The compressor's power consumption, without the use of auxiliary technology, remains nearly constant over $\Delta\tau$ seconds, which is due to the slower dynamics of its controller and equipment than the variation in fan speed. Thus, the compressor is unable to modulate the power consumption and deploy regulation reserves.

The power consumption of the fan at second-interval τ is given as a quadratic function of the initial supply air flow rate m_τ and the change Δm_τ as follows:

$$p_\tau^f = \alpha_1(m_\tau + \Delta m_\tau) + \alpha_2(m_\tau + \Delta m_\tau)^2, \tag{3.2}$$

where α_1 and α_2 are parameters fitted using historical fan power data (Ma et al., 2012). Note that the power consumption of the compressor is assumed to stay constant for each interval τ in the hour t, and is denoted as p_t^c.

3.2.3 Regulation Signal

The regulation signal is determined by the grid operator through a proportional-integral-derivative filter of system-wide frequency deviation (PJM Staff, 2017). It is issued every $\Delta\tau$ seconds to the resources participating in regulation reserve market, which must deploy to track its value and maintain system frequency. The signal value used for deploying bi-directional regulation reserves stochastically varies between $[-1, 1]$ and is typically energy-neutral over an hour. We follow the approach presented in (Bilgin et al., 2016) to model the signal, where the monotonic change in

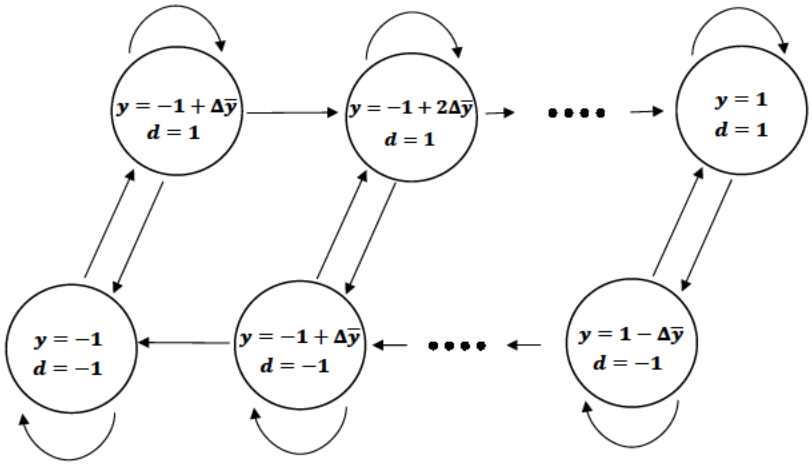


Figure 3·2: Two Dimensional Markov Chain (TDMC) Model for the Regulation Signal Dynamics with Signal Value y and Direction d.

signal sub-trajectories is used to represent the signal dynamics by a two-dimensional Markov chain (TDMC). The chain dimensions include the signal value $y_\tau \in [-1, 1]$ and direction $d_\tau = sign(y_\tau - y_{\tau-1})$ at an interval τ, with the change in signal value between intervals specified as $\Delta y_\tau = (y_\tau - y_{\tau-1}) \in \{-\Delta\bar{y}, 0, \Delta\bar{y}\}$. Note that the direction d_τ is only allowed to take values $\{-1, 1\}$, which is ensured by the condition $d_\tau = d_{\tau-1}$ in case $\Delta y_\tau = 0$. Therefore, the modeled TDMC has discrete state-space $\big(y, d\big) \in \big(\{-1, -1+\Delta\bar{y}, .., 0, .., 1-\Delta\bar{y}, 1\}, \{-1, 1\}\big)$ as shown in figure 3·2. The transition probabilities in the Markov chain can be calculated using historical regulation signal data from the grid operator.

3.3 Regulation Reserve Deployment Problem

We model the regulation reserve deployment control problem as a MDP, where the goal is to track the reference regulation signal from the grid operator and minimize the temperature utility loss over an hourly horizon. Note this is accomplished for a given hourly average energy consumption baseline, regulation reserve capacity and average hourly temperature, all of which are specified to the deployment controller by the regulation reserve capacity scheduler (i.e. ISD problem).

State, $x_\tau = (m_\tau, \theta_\tau, \theta^b_\tau, y_\tau, d_\tau)$

The state variable x_τ is a 5-tuple, whose elements include the supply air flow rate m_τ, inside air temperature θ_τ and building mass temperature θ^b_τ, as well as the regulation signal value y_τ and direction d_τ. The level of discretization for an element of the state variable is denoted by n^i, where i refers to the element index (e.g. $i = 1$ for m_τ, $i = 2$ for θ_τ and so on). Total number of possible states can then be represented as $|x_\tau| = \prod_i n^i = N$.

Control, $u_\tau = (\Delta m_\tau)$

The change in supply air flow rate, Δm_τ, is the control variable, which modifies the fan power consumption p_τ^f in (3.2) to allow the tracking of reference regulation signal. The control variable is limited by the ramp rate of HVAC fan, which corresponds to change in air flow rate values between $[-\Delta\bar{m}, \Delta\bar{m}]$. The number of discrete control levels within the aforementioned range is denoted as $|u_\tau| = z$. Note that, for a state x, the range of permissible controls may be smaller than the maximum range given above and is represented as $U(x) = \big[\max(m^- - m, -\Delta\bar{m}), \min(m^+ - m, \Delta\bar{m})\big]$.

Exogenous Variables and Parameters

The hourly energy consumption baseline p_t and regulation reserve capacity r_t are specified to the deployment controller by the scheduling problem/scheduler, and stay fixed for the hour (i.e. the horizon of DP). In addition, the scheduler specifies the average inside air temperature for the hour $\bar{\theta}_t$, around which the deployment controller must calculate the temperature utility loss that is incremental/residual (additive or subtractive) to the utility loss determined by the scheduler. Other known information for the hour includes the temperature set-point $\hat{\theta}_t$ and allowable temperature range (θ_t^-, θ_t^+). Different penalty factors for the $\Delta\tau$ second-interval are known, including the temperature utility loss penalty factor ϵ', temperature limit penalty factor γ and tracking error penalty factor κ. In addition, the ambient temperature θ_τ^a and heat Q_τ are specified for each $\Delta\tau$ second interval, and can be considered deterministic or probabilistic. Furthermore, system parameters including supply air temperature θ^s, thermal resistances (R^b, R^w) and thermal capacitances (C, C^b) are known.

State Transitions

Dynamics for the air flow rate are given by $m_{\tau+1} = m_\tau + \Delta m_\tau$, whereas temperature dynamics for θ_τ and θ_τ^b are governed by (3.1). Regulation signal dynamics are obtained

through the TDMC. Based on the solution approach used, state transitions may be sampled or the transition probabilities may be calculated using the aforementioned dynamics.

Objective Function

The objective function seeks to minimize the deployment signal tracking error cost, and the temperature utility loss that is residual to the utility loss from the scheduling problem. It also includes a logarithmic barrier term to enforce the temperature limit constraints, as follows:

$$\underset{\Delta m_\tau}{\text{minimize}} \sum_{\tau=1}^{T} \left[\kappa D_\tau + \epsilon' L_\tau + \gamma B_\tau\right], \tag{3.3}$$

where,

- The first term penalizes (by κ) the quadratic tracking error between the reference power ($p_t + y_\tau r_t$) — which we also denote as p_τ^{ref} — and the actual power output p_τ in each time-interval, as follows:

$$D_\tau = (p_t + y_\tau r_t - p_\tau)^2, \tag{3.3a}$$

  The reference power is determined for each time-interval using the fixed hourly energy consumption p_t and regulation reserve capacity r_t as well as the time-varying regulation signal value y_τ, whereas the actual power output in the time-interval is the sum of the compressor and fan power i.e. $p_\tau = p_t^c + p_\tau^f$. Note that the compressor power p_t^c appears in both p_t and p_τ, thus in effect we have $p_t - p_\tau = p_t^f - p_\tau^f$. Hence, for regulation reserve deployment, only the fan power consumption is of interest, and from here onwards, p_t and p_t^f would be used interchangeably.

- The second term penalizes (by ϵ') the residual temperature utility loss L_t in

each time-interval, which has a fixed functional form expressed as:

$$L_\tau = 2(\bar{\theta}_t - \hat{\theta}_t)(\theta_\tau - \bar{\theta}_t) + (\theta_\tau - \bar{\theta}_t)^2, \tag{3.3b}$$

The hourly temperature utility loss calculated by the regulation reserve capacity scheduler is given as $(\bar{\theta}_t - \hat{\theta}_t)^2$, hence the expression in (3.3b) captures the incremental/residual (additive or subtractive) utility loss for each $\Delta\tau$ second time-interval. Note that the penalty factor ϵ' is related to the hourly thermal discomfort penalty factor ϵ_t as $\epsilon' = \epsilon_t \Delta\tau / \Delta t$.

- The third term penalizes (by γ) the proximity of temperature θ_τ in each time-interval, to the lower and upper limits, θ_t^- and θ_t^+ respectively, as follows:

$$B_\tau = -\log[(\theta_t^+ - \theta_\tau).(\theta_\tau - \theta_t^-)]. \tag{3.3c}$$

Note that the total penalty costs for each of the aforementioned terms are dependent on the error/loss/distance, as well as on the corresponding penalty factor (κ, ϵ' or γ).

3.4 Solution Approach

The MDP representing the regulation reserve deployment problem can be formulated as a stochastic dynamic programming or Reinforcement Learning (RL) problem to derive the optimal control policy and cost. We formulate the discounted-cost Dynamic Program (DP) in 3.4.1, and solve it using the Linear Programming based approach in 3.4.2. We then present the RL formulation in 3.4.3, and solve it using actor-critic algorithm in 3.4.4.

3.4.1 Discounted-Cost Dynamic Program

The control decision in the aforementioned MDP is made every $\Delta\tau$ seconds, which is a much shorter duration than the one-hour horizon. Thus, the MDP can be modeled

as a finite state, discounted cost infinite horizon DP using the following Bellman Equation (Bertsekas, 2017):

$$J_\rho(x\,|p_t, r_t, \bar{\theta}_t) = \min_{u\in U(x)} \mathop{\mathbb{E}}_{x'|x,u,p_t,r_t,\bar{\theta}_t} \Big(g(x,u) + \rho J_\rho(x'\,|p_t, r_t, \bar{\theta}_t) \Big), \tag{3.4}$$

where,

- J_ρ is the discounted cost, representing the expected cost-to-go for the current state x, given the hourly values of energy consumption baseline p_t, regulation reserve capacity r_t, and average inside air temperature $\bar{\theta}_t$,
- u (or Δm) is the control that can take discretized values in the range defined by $U(x)$,
- x' is the new state reached from the current state x,
- The transition from x to x' is probabilistic due to: (i) stochastic dynamics of the regulation signal (y_τ, d_τ), (ii) state and control space discretization, as it leads to an approximation of the continuous dynamics for the remaining states $(m_\tau, \theta_\tau, \theta^b_\tau)$, introducing a degree of uncertainty. We note that exogenous disturbances, i.e., ambient temperature θ^a_τ and heat Q_τ impact the transition of the inside air temperature θ_τ — see (3.1), and may also contribute to the probabilistic transition if they are themselves considered stochastic. However, in this chapter, we assume these disturbances to be deterministic and constant,
- $g(x,u)$ is the one-step cost function given as follows:

$$g(x,u) = \kappa D(x,u) + \epsilon' L(x) + \gamma B(x), \tag{3.5}$$

Note from (3.3a) that the $D(x,u)$ is also dependent on the hourly energy con-

sumption p_t and regulation reserve capacity r_t, whereas from (3.3b) that $L(x)$ is dependent on the hourly average inside air temperature $\bar{\theta}_t$. The value of period cost can be negative but its absolute value is bounded from above by some finite number M,

- ρ is the discounting factor.

We use $\mu(x)$ to denote the optimal steady-state infinite horizon control policy that maps states to control actions. Further $x_0 = (\theta_{t-1}, \theta^b_{t-1} | p_t(m_t), r_t, \bar{\theta}_t)$ denotes the state at τ_0, where $\theta_{\tau_0} = \theta_{t-1}$ and $\theta^b_{\tau_0} = \theta^b_{t-1}$, with $m_{\tau_0}, y_{\tau_0}, d_{\tau_0}$ free to take any value for the given hourly p_t, r_t and $\bar{\theta}_t$. Note the hourly power p_t is related to the hourly supply air flow rate m_t, used in expected IHC J_t's expression in the previous chapter, through $p_t = p^f_t = \alpha_1 m_t + \alpha_2 m_t^2$.

We define the optimal average cost per $\Delta\tau$ period as $\bar{J}_\mu = \lim_{T\to\infty} \frac{1}{T} J^T_\mu(x_0)$, where $J^T_\mu(x_0) = \mathbb{E}[\sum_{\tau=0}^{T} g(x_\tau, \mu(x_\tau)]$ represents the expected total cost incurred under the optimal policy. Note that $\bar{J}_\mu$ is a scalar that is independent of the state x_0. In addition, we define $h_\mu(x_0) = \lim_{T\to\infty} (J^T_\mu(x_0) - T\bar{J}_\mu)$ to be the differential cost. For large T, i.e., $T = T^{large} = 1$ hour $= t - (t-1) = \frac{3600}{\Delta\tau}$, we can approximately write $h_\mu(x_0) \approx J^{T^{large}}_\mu(x_0) - T^{large}\bar{J}_\mu$. By letting $J_t(x_0) = J^{T^{large}}_\mu(x_0)$, we can further write the following relationship:

$$J_t(x_0) = J^{T^{large}}_\mu(x_0) \approx h_\mu(x_0) + T^{large}\bar{J}_\mu = h_\mu(x_0) + \frac{3600}{\Delta\tau}\bar{J}_\mu, \tag{3.6}$$

A benefit of using $\bar{J}_\mu$ and $h_\mu(x)$, compared to directly deriving J_t using Monte-Carlo method as shown in the previous chapter, is that it eliminates the need to consider different values of θ_{t-1} and θ^b_{t-1} separately, as the cost for all the values of θ_{τ_0} and $\theta^b_{\tau_0}$ are provided at once by considering $h_\mu(x_0)$, which reduces the overall simulation time for expected IHC (although different values of $p_t, r_t, \bar{\theta}_t$ are still considered).

The discounted cost infinite horizon DP with $\rho \approx 1$ resembles the above described average cost infinite horizon problem. In fact, the optimal policy μ_ρ for the discounted cost DP with $\rho \approx 1$ would be Blackwell optimal, and equal to the average infinite horizon optimal policy μ. Further note that as ρ gets closer to 1, the discounted cost-to-go J_ρ under the optimal policy becomes an increasingly flatter function of the state x_0 and is related to the optimal average period cost through $\bar{J}_\mu = \lim_{\rho \to 1}(1-\rho)J_\rho$. Thus, by solving the discounted cost DP and using its solution (i.e., μ_ρ and J_ρ), we can find the optimal policy μ and average cost $\bar{J}_\mu$, as well as the differential cost h_μ through simulation, all of which help determine the expected IHC J_t. From here onwards, we will drop the conditional notation in J_ρ for brevity. We present our approach for solving (3.4) next.

3.4.2 Linear Programming based Solution Approach

We solve the discounted cost DP using the following Linear Program (LP) (Bertsekas, 2017).

$$\begin{aligned} &\max_{J_\rho(1),\ldots,J_\rho(N)} \sum_{x=1}^{N} J_\rho(x), \\ J_\rho(x) &\leq g(x,u) + \rho \sum_{x'=1}^{N} P_{xx'}(u) J_\rho(x'), \quad \forall x, \forall u. \end{aligned} \tag{3.7}$$

The solution to the above LP provides the optimal control policy and costs. The optimal control policy $\mu(x)$ is the mapping between a state x and the optimal control u^*, whereas the optimal cost $J_\rho(x)$ is the expected cost-to-go incurred when starting in a state x and following the optimal control policy. Specifically, the optimal control policy can be obtained using the binding constraints in the solution of the LP. For each state x, the specific control u corresponding to the binding constraint (out of the z different constraints) represents the optimal control u^*.

The number of constraints in the above LP is a product of the number of states

and control. Thus, with granular discretization, the size of the problem can increase quickly, taking multiple hours to solve and potentially becoming intractable. However, there are various modeling tricks that may be employed to improve the computational efficiency of the LP approach, e.g. (Bilgin et al., 2016) used policy parameterization along with using approximate policy iteration, which enabled solving a reasonably tractable LP. We present a different, model-free approach to overcome such issues and solve the deployment problem using RL.

3.4.3 Reinforcement Learning Formulation

The MDP can be also be solved using RL which is inspired by Q-learning. The underlying idea in Q-learning is to apply a control decision u for only the current interval, and transition from the current state x to a new state x', with the assumption that control policy $\mu(x)$ would be followed thereafter. The expected incurred cost from the new state over the hour would be the cost-to-go under the given policy. Rather than considering the exact expected cost-to-go, we use an estimate that is referred as the Q-function and denoted $\mathcal{Q}(x, u)$. By repeatedly applying the control to optimize the current interval cost and the Q-function estimate, and then updating the estimates, one can converge to the optimal solution. The value iteration for Q-learning can be written as follows:

$$\mathcal{Q}^{new}(x, u) = \mathbb{E}\Big(g(x, u) + \rho \min_{u}(\mathcal{Q}^{old}(x', u)\Big) \tag{3.8}$$

The premise with RL is to enable a model-free method that is amenable to sampling and online implementations. Since the transition is sampled and no transition probabilities are used, the above Q-learning equation becomes:

$$\mathcal{Q}^{new}(x, u) = \mathcal{Q}^{old}(x, u) + \alpha^{\mathcal{Q}}_{(x,u)}\Big(g(x, x', u) + \rho \min_{u}\mathcal{Q}^{old}(x', u) - \mathcal{Q}^{old}(x, u)\Big) \tag{3.9}$$

where $\alpha^{\mathcal{Q}}_{(x,u)}$ is the step-size or learning rate that satisfies the Robbins Monro conditions for stochastic convergence, e.g., $\alpha^{\mathcal{Q}}_{(x,u)} = \frac{b}{n(x,u)}$ where b is a constant and n is the count of the visits to the pair (x, u) during the simulation period.

The Q-function captures the cost for state-action pairs, allowing us to define cost estimates only for the states, which are referred to as the value function $V(x)$, similar to the cost-to-go used in DP. Note that the estimate of expected IHC J_t through RL can be obtained similarly to the DP as mentioned in Section 3.4.1. However, the RL approach differs from DP in two distinct ways. Firstly, it does not include a model for transitioning from state x to x'. Instead, it uses an actual realization of the building HVAC system and samples the state transition from one time step to the next. Secondly, the update of the value function is performed only for the sampled states, as opposed to the update of J_ρ for all the states in the DP approach. Moreover, when determining the value function, approximations for the optimal control policy can be employed. The use of continuous, differentiable value function and policy approximations enables a continuous mapping from state to control space, and can help overcome the dimensionality issues associated with discrete states and controls in DP. Additionally, performing the value function updates for a subset of states through sampling allows for faster updates, and is well-suited for adaptive and online control in regulation reserve deployment. However, the flexibility in modeling requirements and computational tractability in RL may come at the expense of convergence guarantees and sample efficiency.

3.4.4 Actor-Critic Algorithm

The RL formulation is solved using actor-critic algorithm, as it strikes a valuable balance between various solution considerations, e.g., modeling complexity, computational tractability, convergence, and sample efficiency. The algorithm represents a control policy structure (called 'actor') that maps the states to controls, as well as a

value function (called 'critic') that evaluates the cumulative cost of being in a state under the policy. The actor implements the control policy with the goal of optimizing the expected cumulative cost over time, whereas the critic provides feedback on the quality of the implemented controls and helps guide the actor towards controls that further optimize this goal, whilst also improving its own value function estimates.

For example, during each episode of the learning process of one-step actor-critic shown in Algorithm 1 (Sutton and Barto, 2018), the actor implements the controls based on the control policy and the current state. As a result, the building HVAC system transitions to the next state and incurs a cost. The critic evaluates the value function for the two states alongside the observed transition cost to determine the impact of the transition. This impact is used to update the actor's control policy and the critic's value function estimates accordingly. By repeating this process iteratively, the actor and critic enable policy and value function estimation improvements.

3.5 Numerical Results

In this section, we present numerical results of a case study to demonstrate different solution approaches for the regulation reserve deployment problem. In subsection 3.5.1, we provide the input data. In subsection 3.5.2, we discuss the performance of the optimal control policy and the discounted cost function for the DP formulation, derived using the LP based solution approach. Various aspects of the formulation are explored including the impact of temperature penalty factors, hourly scheduling parameters, initial states and discretization. In subsection 3.5.3, we compare the performance of the derived policy with existing control approaches. In subsection 3.5.4, we present the policy and value function results for the RL formulation, obtained using the actor-critic algorithm.

Algorithm 1: One-step Actor-Critic Algorithm

Inputs:
policy parameterization, $\mu_v = \mu(x, \boldsymbol{v})$
value function parameterization, $V_{\mathbf{w}} = V(x, \mathbf{w})$
Algorithm Parameters:
learning rate or step size, $\alpha^{\boldsymbol{v}} \geq 0$, $\alpha^{\mathbf{w}} \geq 0$
Steps:
initialize parameters, $\boldsymbol{v} \in \mathbb{R}^{d'}$ and $\mathbf{w} \in \mathbb{R}^{d}$
for *episode* **do**
  initialize state and parameters:
  state, $x \leftarrow x_0$
  $I \leftarrow 1$
  while $x \neq x_{terminal}$ **do**
    sample:
    control, $u \sim \mu(x, \boldsymbol{v})$
    new state, x', and one-step cost, $g(x, u)$
    compute the value function correction:
    $\delta \leftarrow g(x, u) + \rho V(x', \mathbf{w}) - V(x, \mathbf{w})$
    update parameters and state:
    $\mathbf{w} \leftarrow \mathbf{w} + \alpha^{\mathbf{w}} \delta \nabla_{\mathbf{w}} V(x, \mathbf{w})$
    $\boldsymbol{v} \leftarrow \boldsymbol{v} + \alpha^{\boldsymbol{v}} I \delta \nabla_{\boldsymbol{v}} \mu(x, \boldsymbol{v})$
    $I \leftarrow \rho I$
    $x \leftarrow x'$

3.5.1 Input Data

The exogenous information is first specified. We consider a medium-sized office building based on data from DOE commercial reference building library ("US Department of Energy: Office of Energy Efficiency and Renewable Energy", 2020). The building parameters are $R^b = 4 \times 10^{-7}\,\text{K.h/J}, R^w = 3.61 \times 10^{-7}\,\text{K.h/J}, C = 6.91 \times 10^6\,\text{J/K}$, and $C^b = 1.94 \times 10^8\,\text{J/K}$. The supply air flow temperature is $\theta^s = 17\,^\circ\text{C}$ and the fan parameters are $\alpha_1 = 0.234\,\text{kJ/kg}$ and $\alpha_2 = 0.0975\,(\text{kJ.s})/\text{kg}^2$. The interval length is $\Delta\tau = 4$ seconds, and the specific heat capacity of air is $c = 1\,\text{kJ}/(\text{kg.K})$. The hourly average energy consumption baseline is $p_t = 1.7$ kW, regulation reserve capacity is $r_t = 0.6$ kW, and average inside air temperature is $\bar{\theta}_t = 26\,^\circ\text{C}$. The temperature utility loss penalty factor is $\epsilon' = 1.5 \times 10^{-5}\,\$/^\circ\text{C}^2$, temperature limit penalty factor is $\gamma = 5 \times 10^{-7}\,\$$, tracking error penalty factor is $\kappa = 1.5 \times 10^{-3}\,\$/\text{kW}^2$ and discounting factor is $\rho = 0.9955$. The preferred temperature set-point is $\hat{\theta}_t = 25\,^\circ\text{C}$, and the ambient temperature θ^a_τ and heating Q_τ for each interval in the considered hour are $32\,^\circ\text{C}$ and $84\,\text{MJ}$ respectively.

Different elements of the state variable x_τ are specified as follows: supply air flow rate m_τ is allowed to vary between $1 - 6$ kg/s, whereas the inside air and building mass temperatures, θ_τ and θ^b_τ, can range between $23 - 27^\circ$C and $22.5 - 27.5^\circ$C, respectively. Regulation signal value y_τ can vary between $[-1, 1]$ and its direction d_τ can take values in $\{-1, 1\}$. The number of discrete levels for the elements of x_t are $(n^1, n^2, n^3, n^4, n^5) = (81, 3, 2, 16, 2)$, so the total number of states is $N = 15552$. Note the regulation signal data used to derive the transition probabilities in the two-dimensional markov chain is taken from PJM market (PJM, 2022b). Regarding the control variable u_τ, the change in the air flow rate corresponding to the ramp rate of HVAC fan is limited to $[-\Delta\tilde{m}, \Delta\tilde{m}] = [-0.125, 0.125]$ kg/s, and the number of discrete control levels within this range is $z = 5$. Note the aforementioned discretizations apply

only to the DP, since the RL formulation uses continuous representation for the state and control variables.

All the implementations are performed in MATLAB on a laptop Intel $i7-12700H$ at 2.30 Ghz with 16 GB ram. We implement the LP formulation using the YALMIP toolbox (Wächter and Biegler, 2006), where the optimization problem is solved using Gurobi (Gurobi Optimization, LLC, 2020).

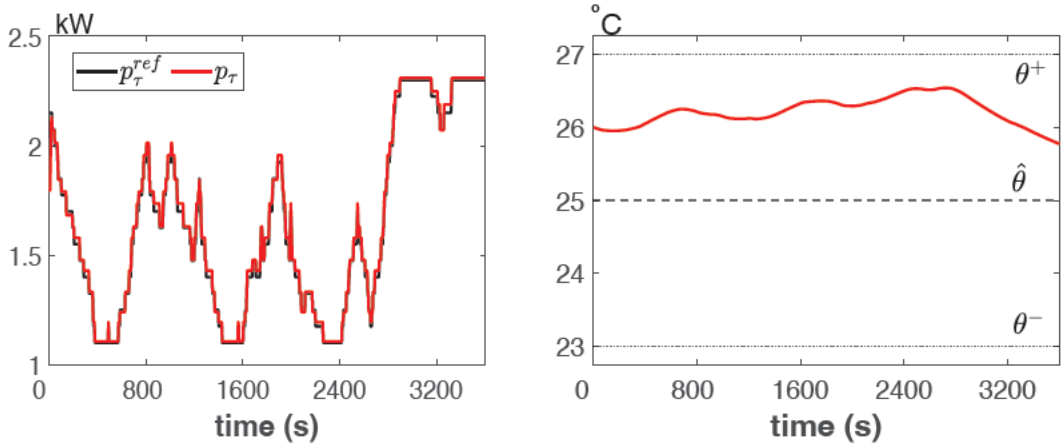
3.5.2 Optimal Control with DP

By solving the LP in (3.7), we are able to obtain the optimal control policy for the DP formulation in the form of a look-up table, where the optimal control for each state is the control corresponding to its binding constraint. In Fig. 3·3a, we present the regulation signal tracking performance of the HVAC under the obtained optimal control policy; p_τ^{ref} denotes the reference power based on the regulation signal from the grid operator, whereas p_τ the actual HVAC power consumption. It can be seen that the control policy is effectively able to modulate the HVAC power and track the signal. The resulting inside air temperature is shown in Fig. 3·3b, where it stays within the upper half of the acceptable range and away from the specified limits.

To better assess the usefulness of the approach, we derive the optimal control policy and test its performance for various sensitivities around the aforementioned base-case. This assessment encompasses different temperature penalty factors, hourly scheduling parameters, initial states, and discretization levels.

Temperature Penalty Factors

The values of temperature penalty factors are determined by the thermal comfort preferences of the building occupants. In Fig. 3·4a, we look at the impact of temperature utility loss penalty factor, ϵ', on the regulation signal tracking performance. The case numbers are indicated as superscripts, and align with rising ϵ' values to signify



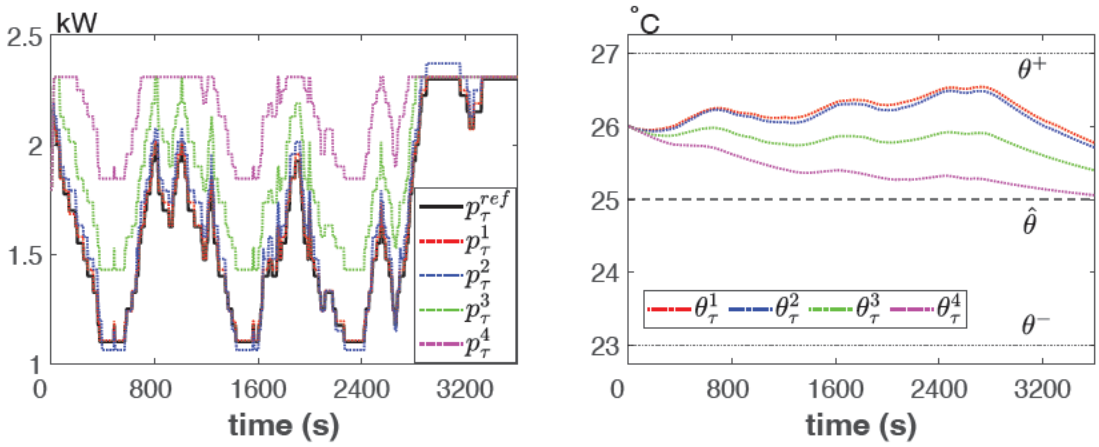
(a) Fan Power Consumption, p_τ^f.

(b) Inside Air Temperature, θ_τ.

Figure 3·3: HVAC Fan Power Consumption and Inside Air Temperature under the Optimal Control Policy in the Base Case.

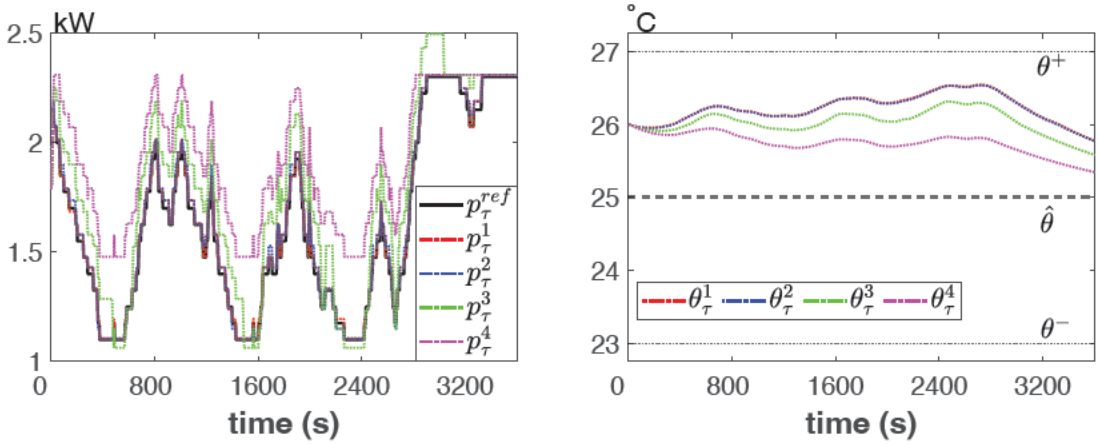
an increasing preference for the temperature set-point, $\hat{\theta}_t$, – see (3.3) and (3.3b). As the case numbers increase, the tracking performance deteriorates since the optimal control policy prioritizes the proximity of the inside air temperature to the set-point, thus sacrificing better signal tracking. Indeed, in Fig. 3·4b, it can be observed that the inside air temperature, θ_τ, under the optimal control policy is much closer to the set-point when the case number, and hence ϵ' is higher.

The impact of temperature limit penalty factor, γ, which penalizes the proximity of the inside air temperature to the temperature limits – see (3.3) and (3.3c), is presented in Fig. 3·5a. The case numbers correspond to increasing γ values and exhibit tracking performance degradation. This is because the inside air temperature, θ_τ, as seen in Fig. 3·5b, is in the upper half of the acceptable range. Increasing the γ value pushes θ_τ further from the allowable upper temperature limit, θ_t^+, which requires the HVAC fan to operate consistently at a higher power consumption level than normal, and compromises its ability to effectively track the reference regulation signal.



(a) Fan Power Consumption, p_τ^f. (b) Inside Air Temperature, θ_τ.

Figure 3·4: HVAC Fan Power Consumption and Inside Air Temperature under the Optimal Control Policy for different Temperature Utility Loss Penalty Factors ($/°C^2) — $\epsilon'^1 = 1.5\text{e}^{-5}$ (base), $\epsilon'^2 = 1.5\text{e}^{-4}$, $\epsilon'^3 = 6\text{e}^{-4}$, $\epsilon'^4 = 1.5\text{e}^{-3}$.

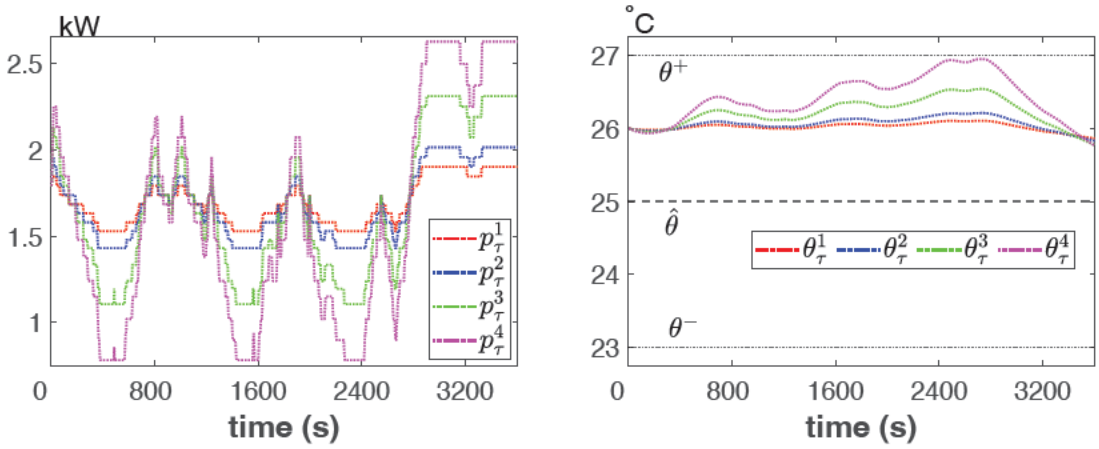


(a) Fan Power Consumption, p_τ^f. (b) Inside Air Temperature, θ_τ.

Figure 3·5: Fan Power Consumption and Inside Air Temperature under the Optimal Control Policy for different Temperature Limit Penalty Factors ($) — $\gamma^1 = 5\text{e}^{-7}$ (base), $\gamma^2 = 5\text{e}^{-6}$, $\gamma^3 = 5\text{e}^{-5}$, $\gamma^4 = 1\text{e}^{-4}$.

Hourly Scheduling Parameters

The hourly scheduling parameters, namely the hourly energy consumption baseline p_t, regulation reserve capacity r_t and average temperature $\bar{\theta}_t$, are specified to the



(a) Fan Power Consumption, p_τ^f.

(b) Inside Air Temperature, θ_τ.

Figure 3·6: Fan Power Consumption and Inside Air Temperature under the Optimal Control Policy for different hourly Regulation Reserve Capacities (kW) — $r_t^1 = 0.2, r_t^2 = 0.3, r_t^3 = 0.6$ (base), $r_t^4 = 0.9$.

regulation reserve deployment controller by the regulation reserve capacity scheduler. p_t and r_t contribute to defining the reference power in (3.3a) that the controller must track, whereas $\bar{\theta}_t$ impacts the residual temperature utility loss in (3.3b). The impact of different values of r_t on the HVAC power consumption and the corresponding inside air temperature is shown in Fig. 3·6a and Fig. 3·6b, respectively. The variation in the power consumption, p_τ, increases with r_t, since it aims to track the more variable reference power p_τ^{ref} – see (3.3a). Although the tracking performance is only moderately affected, this results in lower overall supply air flow rate and lesser cooling provided by the HVAC – see (3.1) and (2.5) – hence the larger inside air temperatures. On the contrary, a higher p_t value leads to a consistently higher reference power that can be difficult to track, and leads to higher supply air flow rate and lower inside air temperature as shown in Fig. 3·7. The average hourly temperature, $\bar{\theta}_t$, does not impact the trade-offs between different terms in (3.3), but only helps facilitate the accurate accounting of the residual temperature utility loss. Thus, in Fig. 3·8, the power consumption for all the cases with different $\bar{\theta}_t$ is the same.

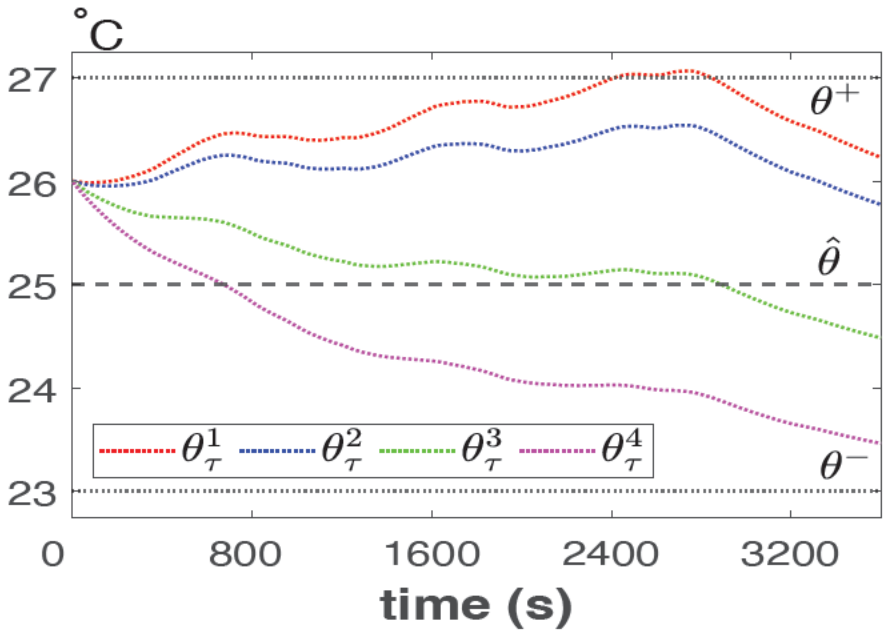


Figure 3·7: Inside Air Temperature, θ_τ, under the Optimal Control Policy for different hourly Energy Consumption baselines (kW) — $p_t^1 = 1.5, p_t^2 = 1.7$ (base), $p_t^3 = 2.5, p_t^4 = 3.5$.

Initial State

The performance of the optimal control policy may be impacted by the initial state of the system. First, the impact of different supply air flow rates, m_τ, at the start of the hour is considered. Depending on the initial m_τ value, the regulation signal tracking error may be large for a few intervals at the beginning, as shown in Fig. 3·9a. However, beyond these few intervals the tracking performance is similar across all cases. Note the difference in HVAC power consumption over the starting intervals translates only into a small difference in inside air temperature θ_τ for the cases, as shown in Fig. 3·9b. Thus, the control policy's performance is fairly robust to different initial m_τ.

The impact of different initial inside air temperatures, θ_τ, on HVAC power consumption and inside air temperature is presented in Fig. 3·10a and Fig. 3·10b, respectively. The regulation tracking performance is only marginally impacted in dif-

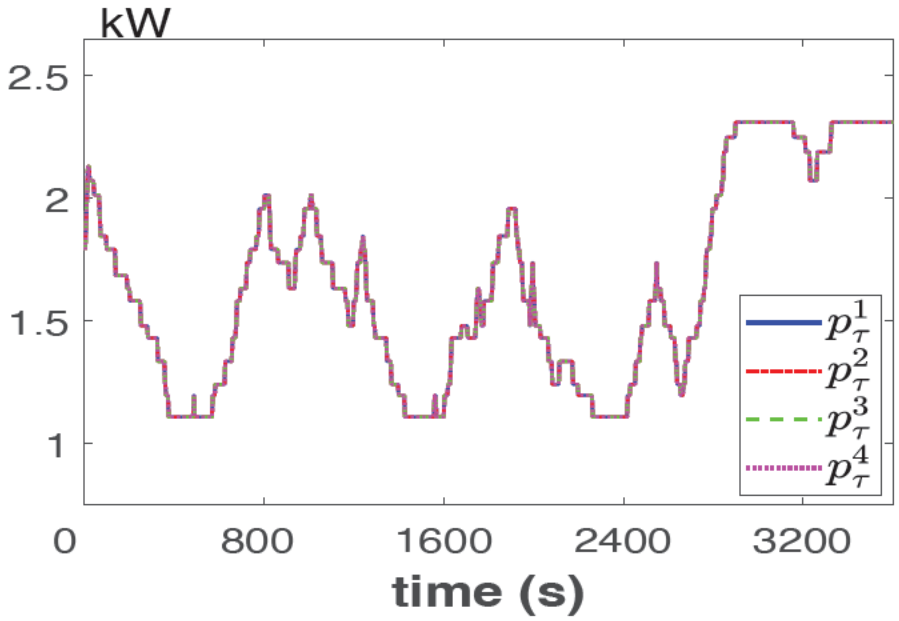
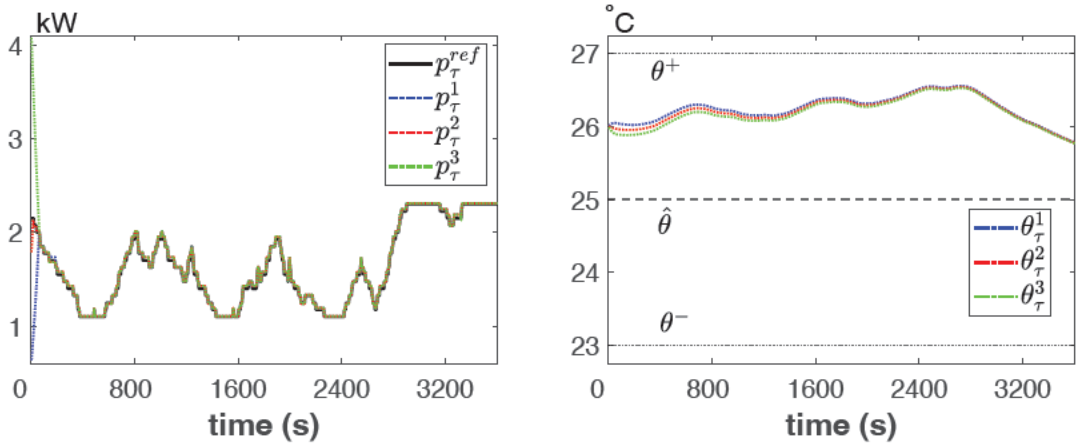


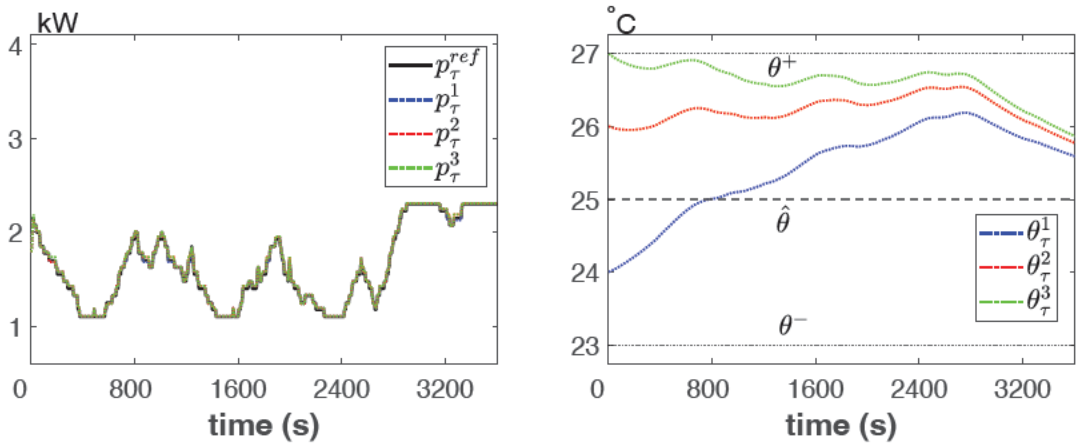
Figure 3·8: Fan Power Consumption, p_τ^f, under the Optimal Control Policy or different hourly Average Temperatures (°C) — $\bar{\theta}_t{}^1 = 26.5$, $\bar{\theta}_t{}^2 = 26$ (base), $\bar{\theta}_t{}^3 = 25$, $\bar{\theta}_t{}^4 = 24$.



(a) Fan Power Consumption, p_τ^f.

(b) Inside Air Temperature, θ_τ.

Figure 3·9: Fan Power Consumption and Inside Air Temperature under the Optimal Control Policy for different initial Supply Air Flow Rates (kg/s) — $m_\tau^1 = 1.5, m_\tau^2 = 3.14$ (base), $m_\tau^3 = 5.5$.



(a) Fan Power Consumption, p_τ^f.

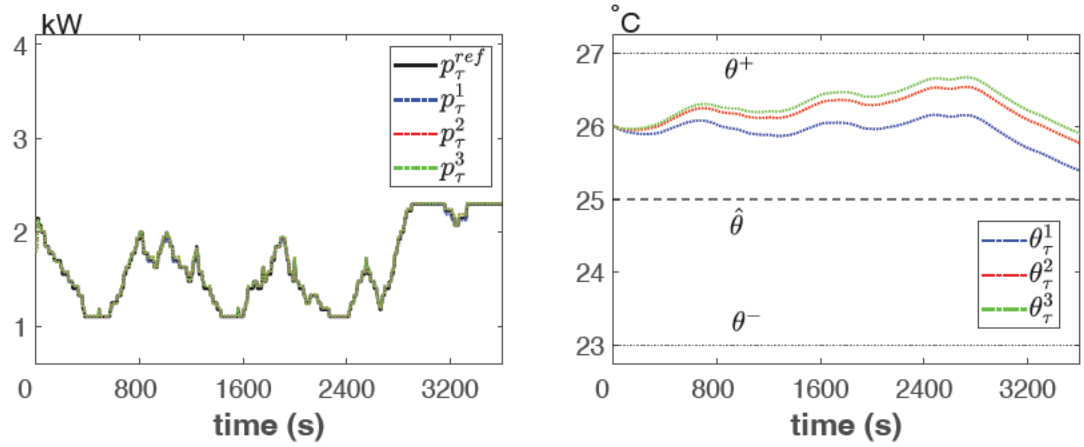
(b) Inside Air Temperature, θ_τ.

Figure 3·10: Fan Power Consumption and Inside Air Temperature under the Optimal Control Policy for different initial Inside Air Temperatures (°C) — $\theta_\tau^1 = 24$, $\theta_\tau^2 = 26$ (base), $\theta_\tau^3 = 27$.

ferent cases, which is due to the higher penalty factor for tracking as compared to temperature. An interesting consequence of similar power consumption in different cases is that the temperatures converge to within close proximity of each other by the end of the hour.

The impact of different initial values of building mass temperature, θ_τ^b, on power consumption and tracking performance is similar to that of θ_τ, as can be seen in Fig. 3·11a. However, the inside air temperature is affected differently. For example, in Fig. 3·11b, a higher initial θ_τ^b leads to a consistently higher inside air temperature, which further increases over the hour. Similarly, a lower initial θ_τ^b leads to a lower, decreasing inside air temperature.

The impact of different initial regulation signal values, y_τ, is illustrated in Fig. 3·12a. The power consumption is initially different in all the cases since each case has different reference power levels to track. However, the regulation signal (y_τ) and reference power for all the cases converge at approximately 400 seconds, hence we see similar power consumption beyond that point. Note the figure shows just one



(a) Fan Power Consumption, p_τ^f.

(b) Inside Air Temperature, θ_τ.

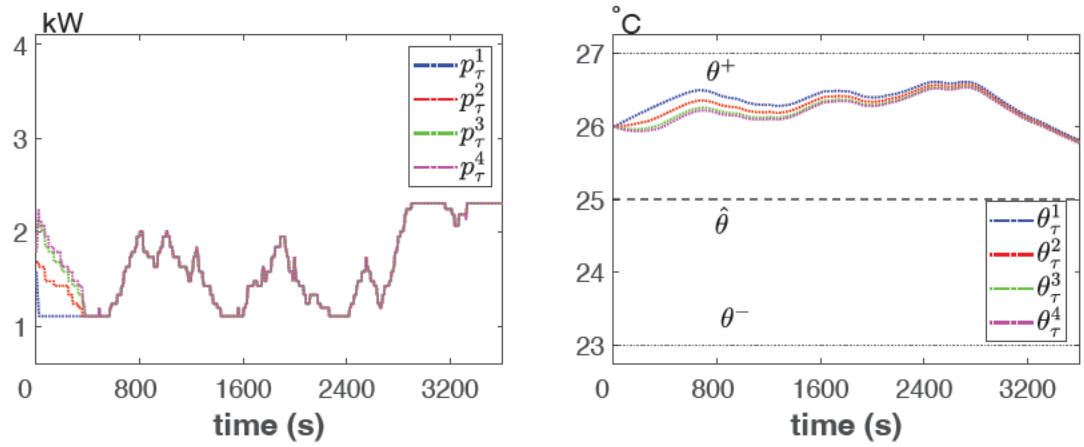
Figure 3·11: Fan Power Consumption and Inside Air Temperature under the Optimal Control Policy for different initial Building Mass Temperatures (°C) — $\theta_\tau^{b,1} = 23$, $\theta_\tau^{b,2} = 26$ (base), $\theta_\tau^{b,3} = 27$.

realization of the regulation signal, and that the convergence in other realizations may occur earlier or later than depicted here. The temperature trajectories for all the cases follow consistently in Fig. 3·12b, with a higher initial power consumption leading to a lower temperature at the start, followed by the temperatures converging over the hour.

The impact of different initial regulation signal directions, d_τ, is similar to y_τ, and thus not shown here. The different cases would exhibit similar regulation signal, power and temperature trajectories right from the beginning, indicating convergence would occur at the start. Only in rare instances would we observe different regulation signal (and hence power and temperature) trajectories with the convergence still occurring at some later point during the hour.

Discretization

The level of discretization for state and control variables has an impact on the accuracy and computational complexity of the solution. Finer discretization helps capture

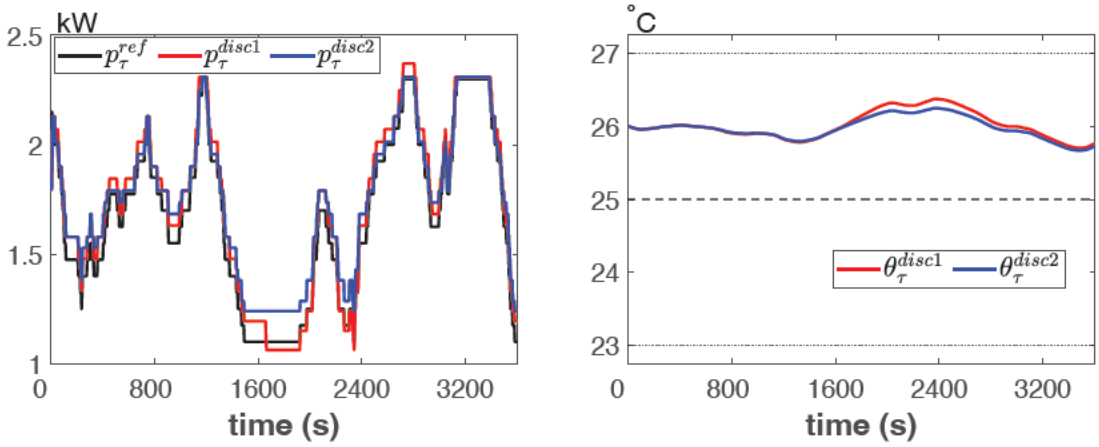


(a) Fan Power Consumption, p_τ^f.

(b) Inside Air Temperature, θ_τ.

Figure 3·12: Fan Power Consumption and Inside Air Temperature under the Optimal Control Policy for different initial Regulation Signal Values — $y_\tau^1 = -1$, $y_\tau^2 = 0$, $y_\tau^3 = 0.75$ (base), $y_\tau^4 = 0.875$.

the problem's dynamics and constraints more precisely, leading to a more accurate calculation of the optimal control policy, $\mu^*(x)$, and corresponding cost over the hour, $J^*(x)$. However, the finer representation may come at the expense of increased computational time and memory requirements. In Fig. 3·13a, the tracking performance for the case with finer discretization is slightly worse since it is better able to represent the temperature utility, which allows it to optimize the total hourly cost more accurately. Indeed, this can be verified in Fig. 3·13b, where the inside air temperature is slightly closer to the set-point across the entire hour when using finer discretization. The solution time for the finer discretization case is much larger though, i.e., approximately 40 minutes versus 7 minutes for the coarser discretization case. The trade-offs between accuracy and computational complexity can become more pronounced based on the values of different parameters, including ϵ', γ, ρ and $\pm\Delta\tilde{m}$, leading to solution times of more than 4 hours and difference in total costs of up to 10%.



(a) Fan Power Consumption, p_τ^f.

(b) Inside Air Temperature, θ_τ.

Figure 3·13: Fan Power Consumption and Inside Air Temperature under the Optimal Control Policy at different discretizations (using $\epsilon' = 1.5e^{-4}$) — $N^1 = 15552$ (base), $N^2 = 25920$.

3.5.3 Comparison with Existing Approaches

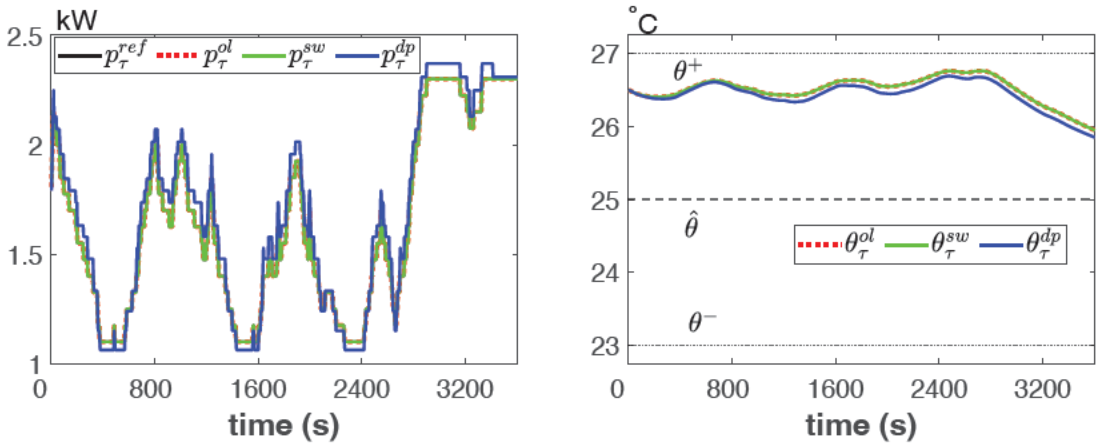
The performance of the optimal control policy, derived using the DP formulation, is compared with open-loop (MacDonald et al., 2014) and switched controllers (Vrettos et al., 2018b). The rationale for choosing these works is that they span a range of different control strategies, including open-loop, feedforward, feedback, and PI controls. Fig. 3·14a shows tracking for the controllers, where it can be seen that the optimal control policy with DP sacrifices the tracking performance to keep the corresponding inside air temperature in Fig. 3·14b slightly closer to the set-point and optimize the overall problem costs. The difference is more pronounced in another case shown in Fig. 3·15a, where tracking the signal well enough would have increased the temperature above its allowable upper limit. Thus, while the open-loop and switched controllers achieve perfect tracking, they end up with temperature limit violations for an aggregate of 10 minutes over the hour, as shown in Fig 3·15b. Although, the tracking performance for DP is perhaps only adequate in this case, it is able to satisfy

the thermal comfort constraints and incur significantly lesser overall cost, i.e., \$0.69 compared to \$1.05 with the other controllers. These examples illustrate that the DP formulation is better able to account for the expected total costs incurred over the hour. Thus, from the perspective of the building occupants, it is more representative of the actual costs of providing load-shifting and regulation reserve services, and lends interpretability to the various preference dependent control parameters at their disposal (e.g. penalty factors). For the grid operator, this would provide a better insight into the long-term availability and performance of HVAC systems participating in electricity markets. For example, with the other controllers, the building occupants may likely override the current settings during the operating hour to improve their thermal comfort, leading to unanticipated poor tracking performance and reserve deliverability issues for the grid operator. On the other hand, if building occupants are not able to do so, and if instances that lead to temperature limit violations are frequent or if comfort preferences are not adequately satisfied, it may increase churn for the market participating HVAC systems. The proposed DP approach provides a fairer assessment of the long-term flexibility that would be available.

3.5.4 RL Implementation

The one-step actor-critic algorithm presented in subsection 3.4.4 is used to solve the RL formulation. The value function parameterization used here is based on (Aslam et al., 2024), and as discussed in chapter 2, assumed to be a second-degree polynomial. For completeness, we include all the states in the expression below:

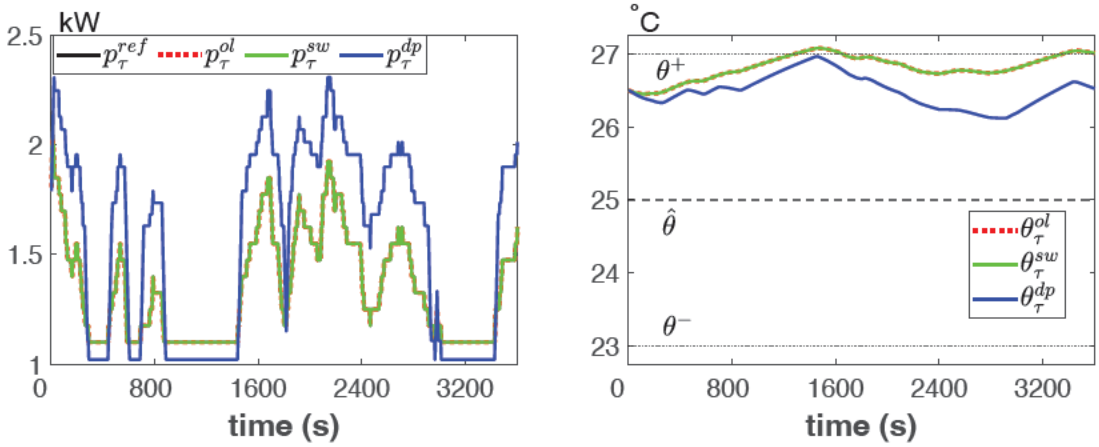
$$\begin{aligned} V_{\mathbf{w}} = V(x, \mathbf{w}) = {} & w_1 m_t^2 + w_2 m_t + w_3 m + w_4 \theta + w_5 \theta^b + w_6 y + w_7 d + w_8 m_t \bar{\theta}_t + w_9 \bar{\theta}_t^2 \\ & + w_{10} \bar{\theta}_t + w_{11} r_t, \end{aligned} \tag{3.10}$$



(a) Fan Power Consumption, p_τ^f.

(b) Inside Air Temperature, θ_τ.

Figure 3·14: Regulation Signal Tracking and Inside Air Temperature under Open-loop, Switched and DP controllers (base with $\epsilon' = 1.5e^{-4}$).



(a) Fan Power Consumption, p_τ^f.

(b) Inside Air Temperature, θ_τ.

Figure 3·15: Regulation Signal Tracking and Inside Air Temperature under the Open-loop, Switched and DP controllers (base with $\gamma = 1e^{-4}$).

where, m_t is the hourly supply air flow rate that is determined using the inverse of the hourly power consumption p_t.

The policy can be represented using a hyperbolic tangent function, where the output is determined by the adjusted sum of tracking and temperature errors along

with the state variables, as shown below:

$$\mu_v = \mu(x, \boldsymbol{v}) = \frac{1 - e^{-a(x)}}{1 + e^{-a(x)}}, \tag{3.11}$$

where,

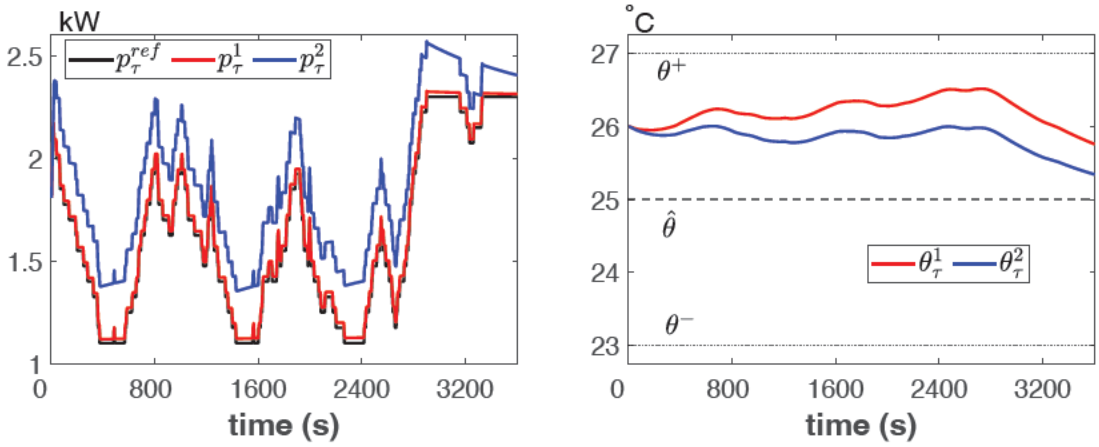
$$\begin{aligned} a(x) =& v_1(p_t + yr_t - \alpha_1 m - \alpha_2 m^2) + v_2(\theta - \hat{\theta}_t) + v_3(\frac{1}{\theta_t^+ - \theta}) \\ & - v_4(\frac{1}{\theta - \theta_t^-}) + v_5 m + v_6\theta + v_7\theta^b + v_8 y + v_9 d \end{aligned} \tag{3.12}$$

The above policy function parameterization provides several benefits such as smoothness, bounded output, symmetry around zero, and effective representation of various underlying policy shapes, e.g., linear, sigmoid, and step functions.

Using the specified parameterizations, the RL formulation is efficiently solved within a few seconds, averaging less than 0.9 seconds per episode. This marks a significant improvement compared to the DP formulation, which required multiple minutes to an hour for the sensitivity cases and faced intractability at high discretization levels. The tracking performance of the derived optimal policy is illustrated in Fig. 3·16a, and is fairly similar to the DP results presented in the previous subsections. This can be further verified by observing the corresponding inside air temperatures in Fig. 3·16b. While the total hourly costs with RL-based optimal policy can be lower than the DP costs, they are often sensitive to the parameter values. Nevertheless, the RL formulation offers the benefits of DP without the tractability concerns, and represents a potentially viable option for online implementations of the deployment problem.

3.6 Conclusions

In this chapter, we developed formulations for the regulation reserve deployment problem using both infinite-horizon DP and RL. The performance of the optimal



(a) Fan Power Consumption, p_τ^f.

(b) Inside Air Temperature, θ_τ.

Figure 3·16: Regulation Signal Tracking and Inside Air Temperature under RL — $\epsilon'^1 = 5e^{-5}$ (base), $\epsilon'^2 = 6e^{-4}$.

control policy, derived from solving the DP, was demonstrated across various use-cases. Furthermore, we compared the optimal policy with existing control approaches, showing it to be more representative of the deployment costs from regulation signal tracking error and temperature utility loss (thermal discomfort). Finally, parametric forms of value and policy functions were derived to have a continuous representation for cost-to-go and controls as well as to facilitate the RL solution.

Chapter 4

Conclusion

This chapter concludes the dissertation. In section 4.1, we highlight the contributions of our research, while section 4.2 outlines various avenues for future research.

4.1 Contributions

In this dissertation, we have addressed issues pertaining to the assessment of HVAC cost efficiency and performance, in the context of providing energy and regulation reserves in day-ahead electricity markets. These issues arise due to the absence of interaction between the seconds timescale deployment and the hourly scheduling problems, and ultimately impact both the HVAC owners and market operators (ISOs). For HVAC owners, these issues can result in sub-optimal schedules and control policies, limiting the realization of economic potential or compromising the thermal comfort of building occupants. Meanwhile, for market operators, these issues can lead to inaccurate signals regarding the short and long-term available flexibility from HVACs.

The contributions of this dissertation can be summarized as follows:

- An ISD problem that explicitly accounts for the expected IHCs, incurred during deployment, as a function of hourly HVAC decision and state variables. The problem allows internalizing the trade-off captured by the IHCs, between regulation signal tracking error and intra-hour occupant discomfort, into the hourly problem, and enables optimization across timescales.

- An MIQP relaxation of the (originally NLP) ISD problem, which provides tight bounds for the NLP optimal solution. We demonstrate the computational performance and tightness of the relaxation through extensive numerical experimentation. To the best of our knowledge this is the first work to provide a computationally efficient method for the ISD problem of a high fidelity HVAC system model.

- Comparison of the ISD problem with existing works that employ hard constraints and worst-case-based approaches which are shown to be overly conservative and not adaptive to the frequency regulation signal characteristics, which our approach inherently considers when estimating the expected IHCs.

- Modeling regulation reserve deployment as an optimal control problem, including as stochastic DP and RL formulations. We demonstrate that these formulations yield interpretable optimal control policies, which are applicable across a broad set of use-cases and more representative of the underlying costs than classical control approaches.

- Parameterizing policy and cost functions in the DP and RL formulations, which facilitate efficient solution approaches and offer valuable insights into the regulation reserve deployment problem, in addition to capturing the interaction between regulation reserve deployment and ISD problems.

- Identifying and discussing the usefulness and applicability of the proposed ISD problem and optimal control approaches, keeping in view the diverse range of possible use-cases including various wholesale market rules, heterogeneity among buildings, and different modes of HVAC operation.

4.2 Future Directions

- **Extension to Real-time Markets:** The ISD problem was primarily considered in the context of symmetric reserves in DA markets, and it would be interesting to explore and extend the approach in the RT market. The expected IHCs from regulation reserve deployment would still be reflected within the ISD problem, although the deployment horizon would be much smaller (i.e., 5-15 minutes) and the deployment problem may need to be formulated as a finite horizon DP or RL. In addition, the energy neutrality assumption regarding the AGC signal would not hold, thus the temperature difference over each RT interval would need to be explicitly accounted in the ISD problem.

- **Extensive Case Studies:** It would be useful to evaluate the impact of the ISD problem and the optimal control formulations under different real-world settings, e.g., in CAISO where the largest revenue contributor for storage is energy, as opposed to ERCOT where faster reserves contribute the most. It may also be particularly helpful to conduct studies in reference to extreme events, e.g., winter storms and summer heat waves, where the utility of the proposed approaches would likely be more evident from both the market operator and building owner perspective.

- **Aggregations** The proposed approaches assumed wholesale market price exposure where the aggregator only serves as a gateway for transmitting the offered schedules. However, in practice, this may lead to market convergence issues. Thus, it would be useful to consider the aggregator problem, evaluate the various game-theoretic considerations for aggregation participants, and assess the implications on the provision of energy and regulation reserves.

- **Multi-Zone Buildings** The proposed approaches are extendable to multi-zone

buildings and it would helpful to quantify the benefits, which would likely be higher, as the benefits of optimal control formulation and ISD problem are added over each zone. Further, the proposed approaches are amenable to distributed and data-driven implementations, and thus can help address the tractability and automation issues associated with highly resolved modeling.

- **Application to Other DERs** The optimal control formulations and the ISD problem can be applied to other DERs as well. Particularly, HVACs are analogous to energy storage (e.g. batteries). Thus, a similar deployment problem may be formulated for batteries where state-of-charge would be a state variable instead of temperature. The ISD problem would work in the same way as well, albeit with battery related costs, variables and constraints.

- **HVAC Synergies with Other Resources** The provision of energy and regulation reserves from a co-located or hybrid HVAC system would be an interesting extension. Potential questions would be whether it is beneficial to consider the different resources in the same deployment and ISD problem, or solve these problems for each resource individually with exogenous consideration for the other resource.

- **Data-Driven Extensions and Transfer Learning** The proposed approaches are amenable to data-driven paradigms such as RL and deep learning. A possible extension for the deployment problem may be to assess the use of different deep learning architectures for policy and value functions in the RL formulation, and evaluate the improvements, if any. For the ISD problem, various machine learning approaches may be employed to continually estimate the building and HVAC parameters involved in the optimization problem. Alternatively, the ISD problem could be formulated as a RL problem. Modeling the ISD problem as

such would allow an end-to-end machine learning implementation for ISD and regulation reserve deployment, which offers several benefits including improved cost efficiency, enhanced performance, and better scalability. Another benefit of the proposed approaches is the ability to use transfer learning, i.e., applying the policies learnt from one model to another. This could help automate and simplify the provision of market services from HVACs as a broader group, which can be often difficult due to the wide-ranging differences in buildings, HVACs and occupant preferences.

References

Aslam, W., Andrianesis, P., and Caramanis, M. (2024). Optimal hvac energy and regulation reserve scheduling in power markets. *IEEE Transactions on Sustainable Energy*, 15(1):201–214. doi:10.1109/TSTE.2023.3279060.

Atam, E. and Helsen, L. (2016). Control-oriented thermal modeling of multizone buildings: Methods and issues: Intelligent control of a building system. *IEEE Control Systems Magazine*, 36(3):86–111. doi:10.1109/MCS.2016.2535913.

Azuatalam, D., Lee, W.-L., de Nijs, F., and Liebman, A. (2020). Reinforcement learning for whole-building HVAC control and demand response. *Energy and AI*, 2:100020. doi:10.1016/j.egyai.2020.100020.

Beale, E. M. L. and Tomlin, J. A. (1970). Special facilities in a general mathematical programming system for non-convex problems using ordered sets of variables. In *J. Lawrence, Ed., Proceedings of the Fifth International Conference on Operational Research*, pages 447–454. Tavistock Publications, London.

Beil, I., Hiskens, I., and Backhaus, S. (2016). Frequency regulation from commercial building HVAC demand response. *Proceedings of the IEEE*, 104(4):745–757. doi:10.1109/JPROC.2016.2520640.

Bertsekas, D. P. (2017). *Dynamic programming and optimal control 4th edition, volume II*. Belmont, MA: Athena Scientific.

Bilgin, E., Caramanis, M. C., Paschalidis, I. C., and Cassandras, C. G. (2016). Provision of regulation service by smart buildings. *IEEE Transactions on Smart Grid*, 7(3):1683–1693. doi:10.1109/TSG.2015.2501428.

Blum, D. H., Xu, N., and Norford, L. K. (2016). A novel multi-market optimization problem for commercial heating, ventilation, and air-conditioning systems providing ancillary services using multi-zone inverse comprehensive room transfer functions. *Science and Technology for the Built Environment*, 22(6):783–797. doi:10.1080/23744731.2016.1197718.

Braun, J. E. (1990). Reducing energy costs and peak electrical demand through optimal control of building thermal storage. *ASHRAE Transactions*, 96:876–887.

CAISO (2023). Business practice manual for market operations. https://bpmcm.caiso.com/Pages/BPMDetails.aspx?BPM=Market%20Operations.

Caramanis, M., Ntakou, E., Hogan, W., Chakrabortty, A., and Schoene, J. (2016). Co-optimization of power and reserves in dynamic T&D power markets with nondispatchable renewable generation and distributed energy resources. *Proceedings of the IEEE*, 104(4):807–836. doi:10.1109/JPROC.2016.2520758.

Chakraborty, N., Mondal, A., and Mondal, S. (2017). Intelligent scheduling of thermostatic devices for efficient energy management in smart grid. *IEEE Transactions on Industrial Informatics*, 13(6):2899–2910. doi:10.1109/TII.2017.2695241.

de Best, R. *HVAC industry - Statistics & Facts*. https://www.statista.com/topics/5225/hvac-industry.

Ding, X., Du, W., and Cerpa, A. (2019). Octopus: Deep reinforcement learning for holistic smart building control. In *Proceedings of the 6th ACM international conference on systems for energy-efficient buildings, cities, and transportation*, pages 326–335. doi:10.1145/3360322.3360857.

Du, Y., Jiang, L., Duan, C., Li, Y., and Smith, J. (2017). Energy consumption scheduling of HVAC considering weather forecast error through the distributionally robust approach. *IEEE Transactions on Industrial Informatics*, 14(3):846–857. doi:10.1109/TII.2017.2702009.

Fabietti, L., Qureshi, F. A., Gorecki, T. T., Salzmann, C., and Jones, C. N. (2018). Multi-time scale coordination of complementary resources for the provision of ancillary services. *Applied Energy*, 229:1164–1180. doi:10.1016/j.apenergy.2018.08.045.

Faddel, S., Tian, G., Zhou, Q., and Aburub, H. (2020a). Data driven q-learning for commercial HVAC control. In *2020 SoutheastCon*, pages 1–6. IEEE. doi:10.1109/SoutheastCon44009.2020.9249737.

Faddel, S., Tian, G., Zhou, Q., and Aburub, H. (2020b). On the performance of data-driven reinforcement learning for commercial HVAC control. In *2020 IEEE Industry Applications Society Annual Meeting*, pages 1–7. IEEE. doi:10.1109/IAS44978.2020.9334865.

Gao, G., Li, J., and Wen, Y. (2020). Deepcomfort: Energy-efficient thermal comfort control in buildings via reinforcement learning. *IEEE Internet of Things Journal*, 7(9):8472–8484. doi:10.1109/JIOT.2020.2992117.

Gurobi Optimization, LLC (2020). Gurobi optimizer reference manual. https://www.gurobi.com/documentation/current/refman/index.html.

Hao, H., Kowli, A., Lin, Y., Barooah, P., and Meyn, S. (2013). Ancillary service for the grid via control of commercial building HVAC systems. In *2013 American Control Conference*, pages 467–472. doi:10.1109/ACC.2013.6579881.

Hao, H., Lin, Y., Kowli, A. S., Barooah, P., and Meyn, S. (2014). Ancillary service to the grid through control of fans in commercial building HVAC systems. *IEEE Transactions on Smart Grid*, 5(4):2066–2074. doi:10.1109/TSG.2014.2322604.

Henze, G. P., Dodier, R. H., and Krarti, M. (1997). Development of a predictive optimal controller for thermal energy storage systems. *HVAC Research*, 3(3):233–264. doi:10.1080/10789669.1997.10391376.

ISO-NE (2020). Transmission, Markets, and Services Tariff: Section III. https://www.iso-ne.com/participate/rules-procedures/tariff/.

Kim, Y.-J. (2018). Optimal price based demand response of HVAC systems in multizone office buildings considering thermal preferences of individual occupants buildings. *IEEE Transactions on Industrial Informatics*, 14(11):5060–5073. doi:10.1109/TII.2018.2790429.

Kim, Y. J., Blum, D. H., Xu, N., Su, L., and Norford, L. K. (2016). Technologies and magnitude of ancillary services provided by commercial buildings. *Proceedings of the IEEE*, 104(4):758–779. doi:10.1109/JPROC.2016.2520678.

Lin, Y., Barooah, P., Meyn, S., and Middelkoop, T. (2015). Experimental evaluation of frequency regulation from commercial building HVAC systems. *IEEE Transactions on Smart Grid*, 6(2):776–783. doi:10.1109/TSG.2014.2381596.

Lissa, P., Schukat, M., and Barrett, E. (2020). Transfer learning applied to reinforcement learning-based HVAC control. *SN Computer Science*, 1:1–12. doi:10.1007/s42979-020-00146-7.

Liu, Y., Yu, N., Shi, J., Dong, B., Ren, W., and Guan, X. (2017). Evaluation of frequency regulation provision by commercial building HVAC systems. In *2017 13th IEEE Conference on Automation Science and Engineering (CASE)*, pages 888–893. doi:10.1109/COASE.2017.8256215.

Löfberg, J. (2004). Yalmip : A toolbox for modeling and optimization in MATLAB. In *2004 IEEE International Conference on Robotics and Automation*, pages 284–289, Taipei, Taiwan. doi:10.1109/CACSD.2004.1393890.

Ma, K., Yu, Y., Yang, B., and Yang, J. (2019). Demand-side energy management considering price oscillations for residential building heating and ventilation systems. *IEEE Transactions on Industrial Informatics*, 15(8):4742–4752. doi:10.1109/TII.2019.2901306.

Ma, Y., Kelman, A., Daly, A., and Borrelli, F. (2012). Predictive control for energy efficient buildings with thermal storage: Modeling, simulation, and experiments. *IEEE Control Systems Magazine*, 32(1):44–64. doi:10.1109/MCS.2011.2172532.

Maasoumy, M., Rosenberg, C., Vincentelli, A. S., and Callaway, D. S. (2014). Model predictive control approach to online computation of demand-side flexibility of commercial buildings HVAC systems for supply following. In *2014 American Control Conference*, pages 1082–1089. doi:10.1109/ACC.2014.6858874.

MacDonald, J., Kiliccote, S., Boch, J., Chen, J., and Nawy, R. (2014). Commercial building loads providing ancillary services in PJM. In *2014 ACEEE Summer Study on Energy Efficiency in Buildings*, pages 3–192–3–206. American Council for an Energy-Efficient Economy. https://www.aceee.org/files/proceedings/2014/data/papers/3-811.pdf.

Murugesan, S., Jiang, Z., Risbeck, M., Amores, J., Zhang, C., Ramamurti, V., Drees, K., and Lee, Y. (2020). Less is more: Simplified state-action space for deep reinforcement learning based HVAC control. In *Proceedings of the 1st International Workshop on Reinforcement Learning for Energy Management in Buildings & Cities*, pages 20–23. Association for Computing Machinery. doi:10.1145/3427773.3427864.

Nagai, T. (2002). Optimization method for minimizing annual energy, peak energy demand, and annual energy cost through use of building thermal storage. *ASHRAE Transactions*, 108:43.

Nagarajan, H., Lu, M., Wang, S., Bent, R., and Sundar, K. (2019). An adaptive, multivariate partitioning algorithm for global optimization of nonconvex programs. *Journal of Global Optimization*, 74(4):639–675. doi:10.1007/s10898-018-00734-1.

National Renewable Energy Laboratory (NREL) (2020). National solar radiation database. https://nsrdb.nrel.gov/data-viewer.

NYISO (2022). Ancillary services manual. https://www.nyiso.com/manuals-tech-bulletins-user-guides.

Pavlak, G. S., Henze, G. P., and Cushing, V. J. (2014). Optimizing commercial building participation in energy and ancillary service markets. *Energy and Buildings*, 81:115 – 126. doi:10.1016/j.enbuild.2014.05.048.

PJM (2022a). Manual 11: Energy and ancillary service market operations. https://www.pjm.com/-/media/documents/manuals/m11.ashx.

PJM (2022b). Market and operations. https://www.pjm.com/markets-and-operations.

PJM Staff (2017). Implementation and rationale for PJM's conditional neutrality regulation signals. https://pjm.com/-/media/committees-groups/task-forces/rmistf/postings/regulation-market-whitepaper.ashx.

Qureshi, F. A. and Jones, C. N. (2018). Hierarchical control of building HVAC system for ancillary services provision. *Energy and Buildings*, 169:216 – 227. doi:10.1016/j.enbuild.2018.03.004.

Raman, N. S., Devraj, A. M., Barooah, P., and Meyn, S. P. (2020). Reinforcement learning for control of building HVAC systems. In *2020 American Control Conference (ACC)*, pages 2326–2332. IEEE. doi:10.23919/ACC45564.2020.9147629.

Raman, N. S., Gaikwad, N., Barooah, P., and Meyn, S. P. (2021). Reinforcement learning-based home energy management system for resiliency. In *2021 American Control Conference (ACC)*, pages 1358–1364. IEEE. doi:10.23919/ACC50511.2021.9483162.

Su, L. and Norford, L. K. (2015). Demonstration of HVAC chiller control for power grid frequency regulation—part 1: Controller development and experimental results. *Science and Technology for the Built Environment*, 21(8):1134–1142. doi:10.1080/23744731.2015.1072449.

Sun, C. and Dai, R. (2018). Distributed optimization for convex mixed-integer programs based on projected subgradient algorithm. In *2018 IEEE Conference on Decision and Control (CDC)*, pages 2581–2586. doi:10.1109/CDC.2018.8618670.

Sundar, K., Nagarajan, H., Linderoth, J., Wang, S., and Bent, R. (2021). Piecewise polyhedral formulations for a multilinear term. *Operations Research Letters*, 49(1):144–149. doi:10.1016/j.orl.2020.12.002.

Sutton, R. S. and Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Turner, W., Walker, I., and Roux, J. (2015). Peak load reductions: Electric load shifting with mechanical pre-cooling of residential buildings with low thermal mass. *Energy*, 82:1057 – 1067. doi:10.1016/j.energy.2015.02.011.

"US Department of Energy: Office of Energy Efficiency and Renewable Energy" (2020). Commercial reference buildings. https://www.energy.gov/eere/buildings/commercial-reference-buildings.

Van Ness, H. C. (1983). *Understanding thermodynamics*. Courier Corp.

Vázquez-Canteli, J., Kämpf, J., and Nagy, Z. (2017). Balancing comfort and energy consumption of a heat pump using batch reinforcement learning with fitted q-iteration. *Energy Procedia*, 122:415–420. doi:10.1016/j.egypro.2017.07.429.

Verbeke, S. and Audenaert, A. (2018). Thermal inertia in buildings: A review of impacts across climate and building use. *Renewable and Sustainable Energy Reviews*, 82:2300 – 2318. doi:10.1016/j.rser.2017.08.083.

Vrettos, E., Kara, E. C., MacDonald, J., Andersson, G., and Callaway, D. S. (2018a). Experimental demonstration of frequency regulation by commercial buildings-part II: Results and performance evaluation. *IEEE Transactions on Smart Grid*, 9(4):3224–3234. doi:10.1109/TSG.2016.2628893.

Vrettos, E., Kara, E. C., MacDonald, J., Andersson, G., and Callaway, D. S. (2018b). Experimental demonstration of frequency regulation by commercial buildings—part I: Modeling and hierarchical control design. *IEEE Transactions on Smart Grid*, 9(4):3213–3223. doi:10.1109/TSG.2016.2628897.

Wächter, A. and Biegler, L. T. (2006). On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming*, 106(1):25–57. doi:10.1007/s10107-004-0559-y.

Wang, W., Tian, G., Sun, Q. Z., and Liu, H. (2022). A control framework to enable a commercial building HVAC system for energy and regulation market signal tracking. *IEEE Transactions on Power Systems*, 38(1):290–301. doi:10.1109/TPWRS.2022.3156867.

Wang, X., Wang, J., and Liu, J. (2020). Vehicle to grid frequency regulation capacity optimal scheduling for battery swapping station using deep Q-network. *IEEE Transactions on Industrial Informatics*, 17(2):1342–1351. doi:10.1109/TII.2020.2993858.

Wei, T., Wang, Y., and Zhu, Q. (2017). Deep reinforcement learning for building HVAC control. In *Proceedings of the 54th annual design automation conference 2017*, pages 1–6. doi:10.1145/3061639.3062224.

Xu, X., Jia, Y., Xu, Y., Xu, Z., Chai, S., and Lai, C. S. (2020). A multi-agent reinforcement learning-based data-driven method for home energy management. *IEEE Transactions on Smart Grid*, 11(4):3201–3211. doi:10.1109/TSG.2020.2971427.

Yang, Y., Hu, G., and Spanos, C. J. (2020). HVAC energy cost optimization for a multizone building via a decentralized approach. *IEEE Transactions on Automation Science and Engineering*, 17(4):1950–1960. doi:10.1109/TASE.2020.2983486.

Yang, Y., Hu, G., and Spanos, C. J. (2021). Stochastic optimal control of HVAC system for energy-efficient buildings. *IEEE Transactions on Control Systems Technology*, 30(1):376–383. doi:10.1109/TCST.2021.3057630.

Yanikara, F. S. (2020). *Decentralized scheduling of EV energy and regulation reserve services in distribution network markets*. PhD thesis, Boston University. https://open.bu.edu/handle/2144/41047.

Yu, L., Sun, Y., Xu, Z., Shen, C., Yue, D., Jiang, T., and Guan, X. (2020). Multi-agent deep reinforcement learning for HVAC control in commercial buildings. *IEEE Transactions on Smart Grid*, 12(1):407–419. doi:10.1109/TSG.2020.3011739.

Zhang, Z., Chong, A., Pan, Y., Zhang, C., Lu, S., and Lam, K. P. (2018). A deep reinforcement learning approach to using whole building energy model for HVAC optimal control. In *2018 Building Performance Analysis Conference and SimBuild*, volume 3, pages 22–23. https://publications.ibpsa.org/conference/?id=simbuild2018.

Zhao, P., Henze, G. P., Plamp, S., and Cushing, V. J. (2013). Evaluation of commercial building HVAC systems as frequency regulation providers. *Energy and Buildings*, 67:225 – 235. doi:10.1016/j.enbuild.2013.08.031.

CURRICULUM VITAE