Title
Risk Management and Combinatorial Optimization for Large-Scale Demand Response and Renewable Energy Integration

Author
Yang, Insoon

Publication Date
2015

Peer reviewed|Thesis/dissertation

Risk Management and Combinatorial Optimization for Large-Scale Demand Response and Renewable Energy Integration

by

Insoon Yang

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Engineering – Electrical Engineering and Computer Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Claire J. Tomlin, Chair
Assistant Professor Duncan S. Callaway
Professor Lawrence Craig Evans
Professor S. Shankar Sastry

Summer 2015

Risk Management and Combinatorial Optimization for Large-Scale Demand Response and Renewable Energy Integration

Abstract

Risk Management and Combinatorial Optimization for Large-Scale Demand Response and Renewable Energy Integration

by

Insoon Yang

Doctor of Philosophy in Engineering – Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Claire J. Tomlin, Chair

To decarbonize the electric power grid, there have been increased efforts to utilize clean renewable energy sources, as well as demand-side resources such as electric loads. This utilization is challenging because of uncertain renewable generation and inelastic demand. Furthermore, the interdependencies between system states of power networks or interconnected loads complicate several decision-making problems. Growing interactions between power and energy systems and human agents with advances in sensing, computing and communication technologies also increase the need for personalized operations.

In this dissertation, we present three control and optimization tools to help to overcome these challenges and improve the sustainability of electric power systems. The first tool is a new dynamic contract approach for direct load control that can manage the financial risks of utilities and customers, where the risks are generated by uncertain renewable generation. The key feature of the proposed contract method is its risk-limiting capability, which is achieved by formulating the contract design problem as mean-variance constrained risk-sensitive control. To design a globally optimal contract, we develop a dynamic programming solution method based on a novel dynamical system approach to track and limit risks. The performance of the proposed contract framework is demonstrated using data from the Electricity Reliability Council of Texas. The second tool is developed for combinatorial decision-making under system interdependencies, which are inherent in interconnected loads and power networks. For such decision-making problems, which can be formulated as optimization of combinatorial dynamical systems, we develop a linear approximation method that is scalable and has a provable suboptimality bound. The performance of the approximation algorithm is illustrated in ON/OFF control of interconnected supermarket refrigeration systems. The last tool seeks to provide a personalized control mechanism for electric loads, which can play an important role in demand-side management. We integrate Gaussian progress regression into a model predictive control framework to learn the customer's preference online and automatically customize the controller of electric loads that directly affect the customer's comfort. Finally, we discuss several future research directions in the opera-

tion of sustainable cyber-physical systems, including a unified risk management framework for electricity markets, a selective optimal control mechanism for resilient power grids, and contract-based modular management of cyber-physical infrastructure networks.

To my parents

Contents

List of Figures

List of Tables

Acknowledgments

This has been one of the most fantastic and unforgettable journeys of my life.

Professor Claire Tomlin has been the best academic advisor I could imagine in every respect. I am indebted to Claire for her wonderful guidance and unconditional support. First of all, I have reshaped my philosophy, not only in research, but also in life through many insightful and inspiring discussions with her over the past six years. I have learned that research can have both a practical impact and theoretical beauty. This invaluable lesson has made me seriously think about how I can contribute to our society with a novel idea. Furthermore, her unlimited support and patience have allowed me to perform my most desired but risky research projects, which produced exciting results as well as intriguing open problems that I want to explore in the future. This wonderful journey would not have been possible without the freedom and advice that Claire gave me. I am excited to pursue the next journey with my own research group having her as an ideal role model.

The main application area for the control and optimization tools in this dissertation is power and energy systems. I have learned how to synergistically develop theoretical and practical research programs in power systems and tightly integrate them through two different research projects, one with Professor Duncan Callaway, and another with Professor Ram Rajagopal. Working with Duncan has been a truly wonderful experience. I am extremely grateful to him for opening my eyes to several practical and interesting aspects of power systems. His broad and profound view on power systems has always motivated me to look at my research problems from different angles. I also appreciate his warm support and encouragement to my somewhat radical thoughts on combining engineering and economic methods in a project regarding contracts for demand response. The second part of this dissertation is based on interactions with Ram. I will never forget our many inspiring and pleasant discussions in Y2E2 (particularly in Coupa Café) at Stanford. His penetrating insight and passion for smart grids and optimization have stimulated my thoughts in novel, rigorous and practical ways. I have also learned from Ram that theoretical research can have an industrial impact – it was very exciting to initiate a collaboration with Coca-Cola to test our combinatorial optimization algorithm for supermarket refrigeration systems.

This dissertation is heavily influenced by an emerging research area called cyber-physical systems. Professor Shankar Sastry has been a great mentor for my exploration of this new area. I have had the privilege of interacting with Shankar in the NSF Foundation of Resilient Cyber-Physical Systems project. His view (from the top) has inspired me to set up several exciting research directions in cyber-physical systems combining engineering, operations research and economic methods. His advice, feedback and encouragement have meant a lot to me – because of them I have been able to learn and construct (near-) optimal strategies for me to move forward in this emerging research area.

Theoretically, this dissertation contributes to stochastic control and optimization. I have been very fortunate to interact with Professor Lawrence Craig Evans, my hero in mathematics. He shared his expertise in partial differential equations, stochastic control and dynamic games. His unbelievable mathematical insight and analysis have always mesmer-

ized me, and he has reinforced my joy of pursuing theoretical beauty. One valuable lesson learned from Craig is the importance of persistence in research. Persistence has been one of the major driving forces for me to complete the theoretical parts of this dissertation. Participating in discussions during his office hours has been one of the most enjoyable parts of my Ph.D., allowing me to chat with many math Ph.D. fellows. In particular, I would like to thank Christopher Miller for many useful and inspiring discussions in principal-agent problems and stochastic control and optimization. His brilliance and good character have always made our collaboration productive and enjoyable. I am very proud of our most recent work (which is not contained in this dissertation) and hope to work with more students like Chris in my future academic career. I also learned about many practical issues in financial contracts and risk management from Peter Vinella. His comments based on thirty years of experience in the financial industry have been more than valuable.

During my Ph.D., I have received great advice from a number of faculty members. I would like to thank Professors Marija Ilić and Ruzena Bajcsy for giving me critical advices on my career, research and teaching. Professor Scott Moura has not only been a great counsellor on any issue but also introduced me the research area of battery control, which I want to explore in the future. I give my gratitude to my "academic older brother," Professor Alexandre Bayen, for showing me his achievements on PDE control and its practical applications. I would also like to thank Professor Alexandre Chorin for introducing me to several stochastic techniques in numerical analysis, and Professor Alexandra von Meier for suggesting to me stimulating research directions using micro-synchrophasors in power distribution systems. I wish to express my gratitude to Professors Munther Dahleh, Domitilla Del Vecchio, Rahul Jain, Javad Lavaei, Steven Low, Richard Murray, Asuman Ozdaglar, Kameshwar Poolla, Luca Schenato, Bruno Sinopoli, Andrew Teel and Pravin Varaiya for inspiring scientific discussions.

Collaborations with folks at Berkeley have made my Ph.D. work more fruitful. I would like to thank Sam Burden for useful discussions on hybrid dynamical systems; Matthias Morzfeld for introducing me to implicit sampling and other stochastic techniques for optimal control and data assimilation; and Melanie Zeilinger for sharing her expertise in model predictive control. I am thankful to Sam Coogan and Lillian Ratliff for successfully co-organizing the EECS Control Theory Seminar series. I have also had the pleasure of interacting with folks around the Hybrid Systems Lab and the (Semi-)Autonomous Lab, which has made my graduate life more enjoyable.

On a personal note, I would like to thank my parents Hyun Gun and Bee Sook and sister Eui Young for their unlimited support and love. Thank you with all my heart.

Chapter 1

Introduction

1.1 Background

The legacy power grid, which is centralized and vertically integrated, is transitioning towards a decentralized and flexible future grid in which renewable energy sources and distributed demand-side resources are expected to be synergistically coordinated with sensing, computing and communication infrastructures. One of the most significant changes is the integration of clean renewable energy sources such as wind, solar, hydro and geothermal generation to decarbonize the electric power grid. Thirty three states in the US have set Renewable Portfolio Standards that require increased the energy production from renewable energy sources [107]. For example, California plans to increase the penetration of renewables up to 33% by the year 2020 [22].

Advances in sensing, communication and computing technologies such as smart meters and smart thermostats allow us to engage demand-side resources in the operation of power systems [23, 98, 56]. For example, the Pacific Gas & Electric (PG&E) Company offers a demand response program called "SmartAC" that pays customers certain rebates to gain the authority to control their air conditioners a few days per year to reduce peak demand [99]. The San Diego Gas & Electric Company is running a pilot project to develop a new pricing scheme for electric vehicles to shift customers' charging patterns in a way that's beneficial to its revenue in electricity markets.

Automation and information technologies also have a significant effect on power grid operations [37]. New sensors and power electronic actuators including micro synchrophasors and flexible AC transmission systems operated by automated sensing and control methods are expected to improve the resilience and efficiency of power distribution and transmission systems [89, 144]. For example, the California Institute for Energy and Environment and the Power Standards Laboratory are investigating the effect of micro synchrophasors on distribution and substation automations [21].

The aforementioned changes in transitioning to the future power grid sound promising. However, many challenges must be overcome to achieve a successful and seamless transition.

1.2 Challenges and solutions

1.2.1 Risks generated by uncertain renewable generation

One of the most important challenges is the risk generated by uncertain wind and solar generation. Power systems have traditionally operated in a way that balances supply and demand at any time, guaranteeing reliability. However, the high penetration of uncertain and intermittent solar and wind generation is expected to cause significant financial and operational risks. In particular, the uncertainty of solar and wind generation will increase errors in the forecast of net load, which is demand minus renewable generation. These errors make it difficult to balance supply and demand. The variability and uncertainty of wind and solar generation can also increase the volatility of wholesale electricity prices [135, 2]. Because of these two factors, utilities will face significant financial risks [137].

To mitigate the risks generated by renewables, it is important to balance out the uncertainties of wind and solar generation. One supply-side solution is to use fast-ramping generators to compensate for the variability of wind and solar generation [7]. Many researchers and system operators have also proposed new electricity market mechanisms and options to manage the risks (e.g., [81, 33, 108]). Another solution is to use demand-side resources such as air conditioners, refrigerators and electric vehicles. A fairly new concept called *demand response* aims to change electric usage by demand-side resources using incentive payments or different electric pricing schemes [128]. *Automated demand response* takes automatic control of the customers' loads using advanced information technologies [98]. This new technology significantly reduces customers' efforts to provide services to power systems because all control actions are automated and their comfort and service levels are guaranteed. In most programs, customers can opt-out from the automated demand response programs whenever they choose [34].

The first goal of this dissertation is to manage the financial risks generated by uncertain renewables using automated demand response technology. This is a nonstandard risk management problem because of the possibly complicated dynamics of loads such as refrigerators and air conditioners. We must take the dynamics of loads into account to guarantee customers' comfort. Another difficulty arises due to the customers' aversion to financial risks. If too many risks are transferred to customers, they may drop out of the demand response program. To overcome these challenges, we will propose a new contract framework that can handle complicated load dynamics and also limit the risk transferred to each customer. This objective is achieved by combining financial contracts and stochastic optimal control of engineered systems. More precisely, the proposed continuous-time contract framework has a risk-limiting capability. If a load-serving entity and a customer enter into such a contract, the load-serving entity can optimally manage its performance and risk with a guarantee that the customer's risk is less than or equal to a pre-specified level and that the customer's expected payoff is greater than or equal to another pre-specified threshold. We achieve such risk-management capabilities by formulating the contract design problem as mean-variance constrained risk-sensitive control. A dynamic programming-based method is developed to

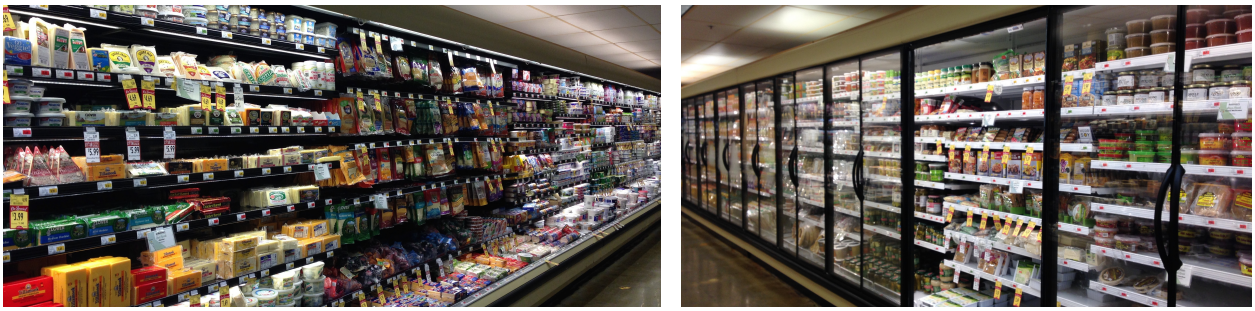

Figure 1.1: Examples of supermarket refrigerators.

solve the problem. The key idea of our proposed solution method is to reformulate the inequality constraints on the mean and the variance of the agent's payoff as dynamical system constraints by introducing new state and control variables. The reformulations use the martingale representation theorem. The proposed contract method enables us to develop a new direct load control method that provides the load-serving entity with financial risk management solutions in real-time electricity markets. We also propose an approximate decomposition of the optimal contract design problem for multiple customers into multiple low-dimensional contract problems for one customer. This allows the direct load control program to work for a large number of customers without any scalability issues. Furthermore, the contract design procedure can be completely parallelized. The performance and usefulness of the proposed contract method and its application to direct load control are demonstrated using data on the electric energy consumption of customers in Austin, Texas as well as the Electricity Reliability Council of Texas' locational marginal price data.

1.2.2 System interdependency

The next challenge considered in this dissertation is system interdependency or coupling. As an example of such, supermarket refrigeration systems are important demand response resources as they represent 7% of the total commercial energy consumption in the U.S. [129]. As shown in Figure 1.1, multiple display cases are interconnected in a supermarket refrigerator, leading to heat transfers between neighboring display cases. There are two factors that make it difficult to optimally control an aggregation of such refrigerators to provide services to power systems. First is the interdependency between display case temperature dynamics. Because of this interdependency, we cannot easily approximate or decouple the large-scale optimization problem for demand response when the total power consumption must be regulated. Another challenge occurs due to the discrete nature of the decision variable, which is the ON/OFF control of the expansion valve or cooling unit in each display case. The discrete control variables prevent us from using well-developed continuous optimization tools. Because of these two factors, demand response using supermarket refrigerators is a combinatorial decision-making problem when their total power consumption should be regulated

to provide services to the power grid. The combinatorial optimization problems associated with interdependent systems generally require significant computational efforts and some of them are not computationally tractable.

As another example, the combination of Kirchhoff's laws and Ohm's law in power networks makes power system states interdependent. Due to this interdependency, several important decision-making problems are combinatorial, including unit commitment and power network topology optimization.

To resolve computational challenges in these combinatorial optimization problems associated with interdependent systems, we will introduce a new linear approximation approach, which provides a fast and scalable algorithm with a provable suboptimality bound. This tool is particularly useful for the real-time optimization of large-scale interdependent systems. Specifically, we consider an optimization problem for a dynamical system whose evolution depends on a collection of binary decision variables. We develop scalable approximation algorithms with provable suboptimality bounds to provide computationally tractable solution methods even when the dimension of the system and the number of the binary variables are large. The proposed method employs a linear approximation of the objective function such that the approximate problem is defined over the feasible space of the binary decision variables, which is a discrete set. To define such a linear approximation, we propose two different variation methods: one uses continuous relaxation of the discrete space and the other uses convex combinations of the vector field and running payoff. The approximate problem is a 0–1 linear program, which can be exactly or approximately solved by existing polynomial-time algorithms with suboptimality bounds, and does not require the solution of the dynamical system. Furthermore, we characterize a sufficient condition ensuring the approximate solution has a provable suboptimality bound. We show that this condition can be interpreted as the concavity of the objective function. The performance and utility of the proposed algorithms are demonstrated with the ON/OFF control problems of interdependent refrigeration systems.

1.2.3 Personalized operation for customers

There have been growing interactions between physical infrastructures, including power systems, and human agents due to advances in sensing, computing and communication technologies. Hence, it is important to address the *personalized efficiency and resilience* in which the system performance level is measured by the agents' utility functions.

However, most existing control methods do not explicitly consider the human agents' utility functions, or the methods assume that the utility functions are given and fixed. These conventional approaches cannot address the personalized efficiency when the agents' utility functions vary over time and change because of adversarial events or environmental disturbances in an unmodeled manner.

Personalized operations of customers' loads are particularly important in demand-side management. For example, a customer may want to frequently tune her controller to find an optimal trade-off between energy efficiency and comfort when time-varying electricity

pricing is applied. However, tuning the controller by herself can be cumbersome. Under the direct load control program, in which an aggregator has the authority to control customers' loads, neglecting a customer's utility function can make her feel uncomfortable with and inconvenienced by her electric loads controlled for demand side management. Direct load control methods with a fixed utility function may result in a similar effect if the true utility deviates from the fixed utility. The resulting dissatisfaction with the management of electric loads by the aggregator can lead the customer to drop out of the direct load control program and is therefore undesirable for both parties. Under the indirect load control program, an aggregator or a load serving entity designs pricing or reward schemes to incentivize its customers to control their loads in a way that is desirable to the aggregator. However, a customer may not control her loads in an expected way if her true utility deviates over time from the nominal utility function adopted in the design of the pricing or incentive schemes.

To overcome this limitation, we develop a personalized control framework that tightly combines online learning of the agent's utility function from data (e.g., agent's rating of the control performance) and the control according to real-time updates of the utility. This new personalized control tool is called the *utility learning model predictive control*. This framework is particularly useful to automatically customize the controller of electric loads that directly affect the customer's comfort. Because the utility function is identified and predicted online using Gaussian process regression, the controller is capable of immediately setting its objective function to the learned utility function and of adjusting its control action to maximize the new objective. Furthermore, no separate training period to learn the customer's utility is needed. The performance of the proposed method is demonstrated by the application to a personalized thermostat controlling indoor temperature. In particular, the proposed integration of Gaussian progress regression into the model predictive control is shown to robustly learn the customer's utility and behavior from noisy data.

1.3 Organization of the dissertation

The rest of this dissertation is organized as follows.

1.3.1 Chapter 2: Risk-limiting dynamic contracts for direct load control

In Chapter 2, we address the challenge of risks in Section 1.2.1. Specifically, we propose a novel dynamic contract approach for direct load control that can manage the financial risks of utilities and customers, where the risks are generated by uncertain renewable generation. The problem setting for direct load control and the definition of the risk-limiting dynamic contract under symmetric information are presented in Section 2.2. We reformulate the constraint on the variance of the agent's payoff, which is called the risk-limiting condition, as a constraint on the compensation provided to the agent by introducing a new control variable in Section 2.4. Using the reformulated constraint, we propose the method

for designing a globally optimal risk-limiting dynamic contract and discuss its decentralized implementation in Section 2.5. Finally, the performance of the proposed contract method and its application to direct load control are demonstrated with data on the electric energy consumption of Austin customers as well as the Electricity Reliability Council of Texas' (ERCOT's) locational marginal price (LMP) data in Section 2.6.

1.3.2 Chapter 3: Approximation algorithms for optimization of combinatorial dynamical systems

In Chapter 3, we present an optimization tool developed for combinatorial decision-making under system interdependencies, which are inherent in interconnected loads and power networks. The problem setting for the optimization of combinatorial dynamical systems is specified in Section 3.2. In Section 3.3, the linear approximation approach for this problem is proposed. To achieve the linear approximation, we propose two different concepts of the derivative of the objective function. Furthermore, for each linear approximation, we provide a condition under which the proposed approximate problem gives a solution with a guaranteed suboptimality bound and show that the condition can be interpreted as the concavity of the objective function or that of a reformulated objective function. In Section 3.4, algorithms to solve the approximate problems with several types of linear inequality constraints are suggested. In Section 3.5, the proposed conditions for the suboptimality bounds to hold are compared with submodularity. Finally, the performance and usefulness of the proposed approximation algorithms are demonstrated with ON/OFF control problems for supermarket refrigeration systems in Section 3.6.

1.3.3 Chapter 4: Utility learning model predictive control for personal electric loads

In Chapter 4, we focus on the challenges in personalized operation of electric loads for customers. We introduce a new personalized control framework, called the utility learning model predictive control (MPC) that can learn its user's preference online. The setup of the utility learning MPC is proposed in Section 4.2. We then introduce the methods for learning the customer's behavior and her utility online and for combining them with the MPC in Section 4.3. The performance of the proposed method is demonstrated with numerical tests for personalized air conditioning in Section 4.4.

In Chapter 5, the contributions of this dissertation is summarized. It concludes with a discussion on future research directions.

Chapter 2

Risk-Limiting Dynamic Contracts for Direct Load Control

2.1 Electricity market risk management using demand-side resources

To reduce the greenhouse gases caused by electricity generation, there has been growing interest in and efforts toward integrating renewable energy sources into the electric power grid. As mentioned in Chapter 1, California plans to use renewable resources to serve 33% of the electricity load by 2020 [22]. In particular, the penetration of solar and wind energy resources is expected to significantly increase. However, the utilization of these resources is challenging because they are uncertain and intermittent. To absorb the uncertainty in solar and wind power, for example, the reserve capacity must be sufficiently large. In California, the increase in the reserve costs needs to be compensated by all customers and the amount of the compensation per customer may not be negligible when California achieves its goal of 33% renewable energy penetration [130].

Supply-side approaches have been proposed to address the uncertainty of renewable resources in economic dispatch and unit commitment using stochastic dynamic programming [108], mixed-integer stochastic programming [17] and robust optimization [14], among others. However, these aforementioned methods do not examine the potential of demand-side resources in managing uncertain renewable generation or loads. To investigate this potential, this chapter proposes a demand-side solution that manages the uncertainty of customers' solar and wind generation and loads. The load-serving entity or the aggregator for the customers procures power or generation reserves in a day-ahead market and the amount of procurement is determined based on a day-ahead load forecast. Because the actual total load deviates from the procured power, the load-serving entity must purchase the deviated amount of power in a real-time market to balance supply and demand. It is desirable for the load-serving entity to minimize the real-time purchase of energy because the energy price and the reserve cost in the real-time market are normally higher and more volatile than

those in the day-ahead market. As the penetration of customers' solar and wind generation increases, however, the electricity price in the real-time market is highly volatile and the customers' demand is very difficult to predict. Therefore, the risk of spending a substantial budget in the real-time market increases. If the load-serving entity bears this financial risk, the energy price in the customers' tariff would inevitably increase. To reduce this risk that the load-serving entity must face, we propose a contract approach for *direct load control* in which the customer transfers the authority to control his or her load to the load-serving entity. Once the load-serving entity and its customer enter into the contract, the load-serving entity can allocate a portion of the risk to the customer through the compensation scheme and the control strategy for the customer's load specified in the contract. A customer might reasonably worry that such compensation and control could increase the risk for large energy costs and a disruption of comfort. The proposed contract addresses this concern by guaranteeing that the risk in the customer's payoff, a weighted sum of energy costs and discomfort level, is limited by a pre-specified threshold and that the mean of the customer's payoff is greater than another pre-specified level. The former is called the *risk-limiting* condition, and the latter is called the *participation payoff* condition. The compensation scheme and the control strategy written in the contract must be designed such that their combination mitigates the load-serving entity's financial risk in the real-time market while satisfying the risk-limiting and participation payoff conditions for the customer.

The key element of the proposed contract for such a demand-side management is direct load control that allows the load-serving entity to actively use the customer's load to manage its financial risk. A number of direct control methods have been suggested for various types of electric loads such as thermostatically controlled loads [86, 50], electric vehicles [113, 83] and deferrable loads [97, 112]. The objectives of existing direct load control methods include shifting demand (e.g., 'valley-filling'), providing ancillary services (e.g., frequency regulation) and energy arbitrage [23, 145, 88]. To the best of the authors' knowledge, however, the potential of direct load control for financial risk management in real-time electricity markets has not yet been studied. We bridge the gap between direct load control and risk management by proposing a contract-based approach.

A variety of contract methods have been suggested for demand-side management in electric power systems. The use of contracts for reducing energy price risk in spot markets has been investigated [67]. Interruptible service contracts have been extensively studied, in which a customer takes risk of service interruption in return for a discount in the energy price [24, 124, 63]. More recently, deadline-differentiated deferrable energy contracts have been proposed to prevent the risk of a customer not receiving energy delivery by a pre-specified deadline [15]. A different contract approach associated with durations-differentiated loads is studied in [93]. None of the aforementioned contract methods, however, takes into account detailed electric load dynamics, which are essential in direct load control to guarantee the customer's comfort and the load's system constraints.

To incorporate dynamics of electric loads in contracts, we adopt a dynamic contract framework, also called a continuous-time principal-agent problem [53, 28]. In such a problem, a principal (e.g., a company) and an agent (e.g., a worker) make a contract that specifies a

compensation scheme and a control strategy in an uncertain environment. The setting we consider in this work is called the *first-best*, in which the principal and the agent have the same information and share the risk in the principal's revenue stream. The principal can monitor the agent's control or effort and, therefore, can enforce the control strategy written in the contract. In the electricity setting, regarding a load-serving entity as the principal and its customer as the agent, this first-best case is appropriate for direct load control because the load-serving entity has the authority to monitor and control its customer's electric loads. A dynamic contract problem for the first-best case was first considered in [90]. It uses a simple principal-agent model with exponential utility functions introduced by Holmstrom and Milgrom for moral hazard [53]. A more general class of the first-best case dynamic contract problems is addressed in [19] by using the martingale and convex duality methods [105, 26, 64, 66]. However, the proposed solution approach requires that the payoff functions be differentiable, strictly increasing and strictly concave and that the dynamical system be a stochastic integral equation. These restrictions are acceptable in many applications in economics and finance, but they may exclude some important engineering problems, including direct load control, because they often require dynamical systems and payoff functions be complicated. Furthermore, the aforementioned methods assume that the utility function of the agent with respect to his or her payoff is given. However, in practice, it is difficult to have complete knowledge of the agent's utility function. In particular, if the agent's utility function used in designing the contract deviates from his or her actual utility, the agent may not want to enter into the contract again.

In this chapter, we propose a novel dynamic contract method that overcomes the limitations of existing methods for the first-best case (i.e., symmetric information). The proposed method uses the variance of the agent's payoff as the risk measure for the agent. By imposing a constraint on the variance, we can limit the risk the agent needs to bear. We call this constraint the agent's *risk-limiting condition*. In addition, the proposed method guarantees that the mean of the agent's payoff exceeds some pre-specified threshold. From the principal's point of view, by executing an appropriately designed compensation scheme and control strategy specified in the contract, the principal can transfer some portion of its financial risk to the agent, respecting the agent's risk-limiting condition. This variance approach does not require complete knowledge of the agent's utility function, which is difficult to obtain in practice. Such a risk-limiting capability distinguishes our method from existing contract methods. To take into account the principal's risk aversion, we formulate the contract design problem as risk-sensitive control [60, 38]. Due to the constraints on the mean and the variance of the agent's payoff, however, dynamic programming is not directly applicable. One may be able to handle the constraints using the stochastic maximum principle or the duality method [102, 66]. However, these approaches do not, in general, allow a globally optimal solution.

The theoretic contribution of the chapter is to develop a method that gives a globally optimal solution of such mean-variance constrained stochastic optimal control problems. More specifically, using the martingale representation theorem, we reformulate the constraints on the mean and the variance of the agent's payoff function, which are difficult to handle, as

two new dynamical systems controlled by new control variables. The first new system state represents the agent's future expected payoff with a modified diffusion term. The second new system state can be interpreted as the remaining amount of risk the agent can bear. The former and latter systems are used to reformulate the constraint on the mean and the variance of the agent's payoff, respectively. It turns out that the reformulated problem is risk-sensitive control with a *stochastic target constraint* in the augmented state space of the original system and the new dynamical systems. A globally optimal solution to the reformulated problem can be obtained by using the dynamic programming principle in the augmented state space. The value function of the problem is computed by numerically solving an associated Hamilton-Jacobi-Bellman equation and is then used to synthesize optimal compensation and control strategy. The proposed solution method allows more general system models for loads and payoff functions for the principal and the agent than existing dynamic contract methods for the first-best case. This flexibility and the risk-limiting capability of the proposed contract method make it appropriate for direct load control policy which explicitly treats financial risk in real-time electricity markets. We also propose an approximate decomposition method for the contract design problem: the problem for n agents can be decomposed into n optimal contract design problems each for a single agent. This decomposition allows an approximate contract with a provable suboptimality bound. Due to the decomposition, the computational complexity of the proposed method increases linearly with the number of agents. Furthermore, the decomposed contract design problem for an agent is independent of that for another agent. Therefore, the contract design procedures for multiple customers can be completely parallelized.

2.2 Stochastic models

We consider a situation in which the load-serving entity wants to make a contract with n heterogeneous customers to directly control each customer's personal electric load, such as an air conditioner or a water heater. For simplicity, we assume that each customer allows the load-serving entity control over only one of his or her loads, although the proposed method is also applicable to the case of multiple loads per customer. The load-serving entity's goal is to manage the risk of spending a substantial budget in a real-time energy market by controlling the customers' loads in the direct load control program. We consider a finite time horizon contract: let $[0, T]$ be the period in which the contract is effective.

2.2.1 Total power consumption

Let $\eta_t^i \in \mathbb{R}$ be the energy consumption (in kWh) up to time $t \in [0, T]$ by customer i and $u_t^i \in \mathbb{R}$ be the power consumption (in kW) by customer i's electric load in the direct load control program. Note that even when $u_t^i = 0$, the total power consumption by customer i is not, in general, zero at time t due to the existence of the customer's other loads and possibly solar or wind generation (which can be considered as negative loads). If all the

customers enter into the contract, the load-serving entity has the authority to determine $u^i := \{u_t^i\}_{0\le t\le T}$ for $i = 1, \cdots, n$. The number, n, of customers is typically in the order of 10^3–10^5. Let $u := (u^1, \cdots, u^n)$. The uncertainty in the customers' loads and solar and wind power generation causes the energy consumption process $\{\eta_t\}_{0\le t\le T}$ to be stochastic. To describe the energy consumption process, we use a stochastic differential equation (SDE) model of the form

$$d\eta_t^i = \left(l_i(t) + u_t^i\right) dt + \tilde{\sigma}_i(t) dW_t^i, \tag{2.1}$$

where $l_i(t) \in \mathbb{R}$, $0 \le t \le T$, is the forecast of customer i's loads (in kW) other than those in the direct load control program. The effect of the load forecast error is modeled by the diffusion term, $\tilde{\sigma}_i(t)dW_t^i$, where $W^i := \{W_t^i\}_{0\le t\le T}$ is a one-dimensional standard Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and the diffusion coefficient $\tilde{\sigma}_i : [0, T] \to \mathbb{R}$ is a bounded function. We assume that W^i and W^j are independent for any $i, j \in \{1, \cdots, n\}$ such that $i \ne j$. The functions l_i and $\tilde{\sigma}_i$ can be estimated from data on the electric energy consumption of customers in Austin as explained in Section 4.4. Furthermore, the validity of the standard Brownian motion in the model for the proposed contract framework is tested using the data in Section 4.4.

2.2.2 Energy price and load-serving entity's revenue in real-time markets

Let $p(t) \in \mathbb{R}$, $0 \le t \le T$, be the amount of power procured by the load-serving entity in the day-ahead market. We assume that $p(t)$ is given. The energy price in the real-time market is chosen as the locational marginal price (LMP). Let λ_t be the LMP at time t. The dynamics of the LMP can be modeled as the following SDE [32, 63]:

$$d\lambda_t = r_0(\nu(t) - \ln \lambda_t)\lambda_t dt + \sigma_0(t)\lambda_t dW_t^0, \tag{2.2}$$

where $W^0 := \{W_t^0\}_{0\le t\le T}$ is a one-dimensional standard Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$ and the price volatility $\sigma_0 : [0, T] \to \mathbb{R}$ is a bounded function. For simplicity, we assume that W^0 is independent of W^i for $i = 1, \cdots, n$, but our contract method can easily be extended to the case in which they are dependent. This model is suitable to capture the *mean-reverting* behavior of energy prices in the real-time (spot) market: when the energy price is high (resp. low), the supply tends to increase (resp. decrease) and, therefore, causes the price to decrease (resp. increase) [31]. Let $w_t := \ln \lambda_t$, then $w := \{w_t\}_{0\le t\le T}$ satisfies

$$dw_t = r_0(\nu(t) - w_t)dt + \sigma_0(t)dW_t^0. \tag{2.3}$$

We estimate r_0, $\nu(t)$ and $\sigma_0(t)$ using the ERCOT LMP data in Section 4.4. In principle, the LMP is not completely exogenous because it is influenced by the power consumption of the customers' loads. In this work, however, we assume that this effect is negligible and, therefore, that the LMP is exogenous.

The load-serving entity's revenue in the real-time market up to time t, denoted as $z_t \in \mathbb{R}$, is given by

$$z_t = \int_0^t \lambda_s \left(p(s)ds - \sum_{i=1}^{n} d\eta_s^i \right).$$

Note that we assume that excess power is sold as easily as deficits are procured. This stochastic integral can be rewritten as the following SDE:

$$dz_t = \lambda_t \left(p(t) - \sum_{i=1}^{n} (l_i(t) + u_t^i) \right) dt - \sum_{i=1}^{n} \lambda_t \tilde{\sigma}_i(t) dW_t^i. \tag{2.4}$$

The load-serving entity's revenue is affected by the control u of the customers' loads in the direct load control program. The set of feasible controls is chosen as $\mathbb{U}^i := \{u^i : [0,T] \to \mathcal{U}^i \mid u^i \text{ progressively measurable with respect to } \mathcal{F}_t^{(i)}\}$, where $\mathcal{U}^i$ is a compact set in $\mathbb{R}$ and $\{\mathcal{F}_t^{(i)}\}_{0 \le t \le T}$ is the filtration generated by the two dimensional Brownian motion $W^{(i)} := (W^0, W^i)$. We also let $\mathbb{U} := \mathbb{U}^1 \times \cdots \times \mathbb{U}^n$.

2.2.3 Customers' loads

Consider customer i's load in the direct load control program, and let $x_t^i \in \mathbb{R}$ be the system state at time $t \in [0,T]$. If the load is an air conditioner unit, then x_t^i would represent the indoor temperature; if the load is a water heater, it would represent the water temperature. Then, the system dynamics can be modeled as the following stochastic differential equation:

$$dx_t^i = f_i(x_t^i, u_t^i)dt \tag{2.5}$$

with the initial condition $x_0^i = x^{0i}$ for $i = 1, \cdots, n$, where the control u_t^i is a stochastic process. Although our contract method can handle stochastic system models with a diffusion term, we use the model (2.5) for simplicity. Here, we assume that $f_i : \mathbb{R} \times \mathcal{U}^i \to \mathbb{R}$ is continuous and that $f_i(\boldsymbol{x}, \boldsymbol{u})$ is differentiable in $\boldsymbol{x}$ for any $\boldsymbol{u} \in \mathcal{U}^i$. We further assume that there exists a constant K such that for all $(\boldsymbol{x}, \boldsymbol{u}) \in \mathbb{R} \times \mathcal{U}_i$

$$\left| \frac{\partial f_i(\boldsymbol{x}, \boldsymbol{u})}{\partial \boldsymbol{x}} \right| \le K,$$
$$|f(\boldsymbol{x}, \boldsymbol{u})| \le K(1 + |\boldsymbol{x}| + |\boldsymbol{u}|).$$

Then, there exists a unique solution $x^i := \{x_t^i\}_{0 \le t \le T} \in \mathbb{L}^2(0,T)$ for $i = 1, \cdots n$, where $\mathbb{L}^2(0,T)$ denotes the space of all real-valued, progressively measurable stochastic processes $\mathbf{x}$ such that $\mathbb{E}\left[\int_0^T \mathbf{x}_t^2 dt\right] < \infty$. See [39] for the proof.

Example 1. *Let x_t^i denote the indoor temperature of customer i at time t, and let $\Theta_i(t)$ represent the corresponding outdoor temperature. Then, the dynamics of the indoor temperature can be described as the following equivalent thermal parameter (ETP) model [120]:*

$$dx_t^i = [\alpha_i(\Theta_i(t) - x_t^i) - \kappa_i u_t^i]dt \tag{2.6}$$

for $i = 1, \cdots, n$. Here, $\alpha_i = R_{1,i}/R_{2,i}$, where $R_{1,i}$ denotes the thermal conductance between the outdoor air and indoor air and $R_{2,i}$ is the thermal conductance between the indoor air and the thermal mass for customer i's room. The positive constant κ_i converts an increase in energy (kWh) to a reduction in temperature ($^\circ C$) for customer i's air conditioner.

2.2.4 Payoff functions

Load-serving entity's payoff

The load-serving entity's payoff function is chosen as its profit in the direct load control program. Let $C^i \in \mathbb{R}$ be the end-time compensation paid to customer i in the direct load control program and $\mu_i(t)$ be the energy price per unit kWh at time t specified in customer i's electricity tariff. We assume that $\mu_i : [0, T] \to \mathbb{R}$ is bounded. The load-serving entity's total payoff in real time, i.e., neglecting the cost of power procured in the day-ahead market, which is its revenue obtained from the customers, is then given by

$$\begin{aligned}
&\int_0^T dz_t + \sum_{i=1}^n \int_0^T \mu_i(t) d\eta_t^i - \sum_{i=1}^n C^i \\
&= \sum_{i=1}^n \int_0^T \left[(\mu_i(t) - \lambda_t)\left(l_i(t) + u_t^i\right) + \lambda_t p_i(t)\right] dt + \sum_{i=1}^n \int_0^T (\mu_i(t) - \lambda_t)\tilde{\sigma}_i(t) dW_t^i - \sum_{i=1}^n C^i,
\end{aligned}$$

where $\{p_1(t), \cdots, p_n(t)\}$ is a set satisfying $\sum_{i=1}^n p_i(t) = p(t)$ and the set of feasible compensation values is chosen as $\mathbb{C}^i := \{C^i \in \mathbb{R} \mid C^i \text{ is } \mathcal{F}_T^{(i)}\text{-measurable}\}$. Let $\mathbb{C} := \mathbb{C}^1 \times \cdots \times \mathbb{C}^n$. We define the payoff function of the load-serving entity as

$$J^P[C, u] := \sum_{i=1}^n \left(\int_0^T r_i^P(t, w_t, x_t^i, u_t^i) dt + \int_0^T \sigma_i^P(t, w_t) dW_t^i - C^i \right), \tag{2.7}$$

where $r_i^P : [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and $\sigma_i^P : [0, T] \times \mathbb{R} \to \mathbb{R}$ are such that

$$\begin{aligned}
r_i^P(t, w_t, x_t^i, u_t^i) &:= (\mu_i(t) - e^{w_t})(u_t^i + l_i(t)) + e^{w_t} p_i(t), \\
\sigma_i^P(t, w_t) &:= (\mu_i(t) - e^{w_t})\tilde{\sigma}_i(t)
\end{aligned} \tag{2.8}$$

for $i = 1, \cdots, n$. The superscript 'P' represents the fact that the load-serving entity plays the role of the principal in the contract. Note that in the direct load control application r_i^P is independent of x_t^i. We use the models (2.8) for direct load control but the proposed contract design method is applicable to more general models of running payoff and volatility. For notational simplicity, we will suppress the dependency of the functions on time.

Customer's payoff

Each customer's total payoff depends on (i) his or her economic profit and (ii) his or her comfort level. Customer i's profit can be computed as the compensation received minus the energy costs, i.e.,

$$\begin{aligned} & C^i - \int_0^T \mu_i(t) d\eta_t^i \\ & = - \int_0^T \mu_i(t)(l_i(t) + u_t^i)dt - \int_0^T \mu_i(t)\tilde{\sigma}_i(t)dW_t^i + C^i. \end{aligned}$$

Customer i's payoff function can be represented as

$$J_i^A[C^i, u^i] := \int_0^T r_i^A(x_t^i, u_t^i)dt + \int_0^T \sigma_i^A(t)dW_t^i + C^i, \tag{2.9}$$

where $r_i^A : [0,T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and $\sigma_i^A : [0,T] \to \mathbb{R}$ are such that

$$\begin{aligned} r_i^A(t, x_t^i, u_t^i) &:= -\mu_i(t)(l_i(t) + u_t^i) + r_i(x_t^i, u_t^i), \\ \sigma_i^A(t) &:= -\mu_i(t)\tilde{\sigma}_i(t) \end{aligned} \tag{2.10}$$

for $i = 1, \cdots, n$. Here, $r_i(x_t^i, u_t^i)$ represents customer i's comfort level given the system state x_t^i and control u_t^i. The superscript 'A' represents the fact that the customer is the agent in the contract. We assume that there exist constants B_0 and B_1 such that $|r_i^A(\boldsymbol{x}, \boldsymbol{u})| \leq B_0 + B_1|\boldsymbol{x}|$ for all $\boldsymbol{x} \in \mathbb{R}$ given any $\boldsymbol{u} \in \mathcal{U}^i$.

Example 2. *Customer i's discomfort level is zero if the indoor temperature, x_t^i, is within a desirable temperature range, $[\underline{\Theta}, \overline{\Theta}]$. The discomfort level increases as the indoor temperature increases above $\overline{\Theta}$ or drops below $\underline{\Theta}$. If we set the comfort level as the negative value of the discomfort level, then we can model customer i's comfort level as*

$$r_i(x_t^i, u_t^i) = -\omega_i \left[(x_t^i - \overline{\Theta})_+ + (\underline{\Theta} - x_t^i)_+ \right], \tag{2.11}$$

where the constant parameter ω_i represents the customer i's valuation of comfort and $(a)_+ := a$ if $a > 0$ and $(a)_+ := 0$ otherwise for any $a \in \mathbb{R}$.

2.3 Risk-limiting dynamic contracts

The load-serving entity (principal) offers customer i a contract that specifies the compensation scheme, C^i, and its control strategy, $u^i := \{u_t^i\}_{0 \leq t \leq T}$ for $i = 1, \cdots, n$. The contract is *dynamic* in the sense that the load-serving entity uses the state feedback control strategy written in the contract to dynamically choose the control action each customer must follow. Customer i (agent i) accepts the contract only if

1. (*participation-payoff condition*) the mean of customer i's payoff is greater than or equal to some threshold, $b_i \in \mathbb{R}$, i.e.,

$$\mathbb{E}[J_i^A[C^i, u^i]] \geq b_i, \tag{2.12}$$

and

2. (*risk-limiting condition*) the variance of customer i's payoff is less than or equal to some threshold, $S_i \in \mathbb{R}$, i.e.,

$$\mathrm{Var}[J_i^A[C^i, u^i]] \leq S_i. \tag{2.13}$$

Note that variance is used as the risk measure of the customer's payoff. We call b_i and S_i the *participation payoff* and the *risk share* of customer i, respectively.

Let $\Lambda := \{(\mathbf{b}_1, \mathbf{S}_1), \cdots, (\mathbf{b}_M, \mathbf{S}_M)\}$ be a set of given pairs of participation payoffs and risk shares. These pairs are designed by the load-serving entity and provided to the customers. Customer i selects a pair $(b_i, S_i) \in \Lambda$ and the contract determined from this pair. Once each customer enters into a contract, the load-serving entity directly controls each customer's load following the control strategy specified in the contract. At the end of the contract period, the load-serving entity compensates each customer according to the compensation scheme specified in the contract.

More specifically, once customer i agrees to enter into the contract, the load-serving entity company has the authority to control customer i's load for maximizing the load-serving entity's expected payoff under the constraints 1) and 2). Taking into account the risk of the load-serving entity's payoff being small as well, we formulate the problem of designing such a dynamic contract (C, u) as the following constrained *risk-sensitive control* problem:

$$\begin{aligned}
\max_{C\in\mathbb{C}, u\in\mathbb{U}} \quad & -\frac{1}{\theta}\log\mathbb{E}\left[\exp(-\theta J^P[C,u])\right] && \text{(2.14a)}\\
\text{subject to} \quad & dw_t = r_0(\nu(t) - w_t)dt + \sigma_0(t)dW_t^0 && \text{(2.14b)}\\
& dx_t^i = f_i(x_t^i, u_t^i)dt && \text{(2.14c)}\\
& \mathbb{E}[J_i^A[C^i, u^i]] \geq b_i && \text{(2.14d)}\\
& \mathrm{Var}[J_i^A[C^i, u^i]] \leq S_i, && \text{(2.14e)}
\end{aligned}$$

where $\theta \in \mathbb{R} \setminus \{0\}$ is a constant, called the coefficient of load-serving entity's *risk-aversion*. When θ is positive, the risk-sensitive objective function penalizes the risk of the load-serving entity's payoff being small because $-\exp(-\theta J^P)$ is concave increasing in J^P. Therefore, the load-serving entity can make a risk-averse decision by solving (3.32). If $\theta < 0$, the load-serving entity is risk-seeking. For intuition, note that the risk-sensitive objective function is well approximated by a weighted sum of the mean and the variance of the payoff when $|\theta|$ is small because the Taylor expansion of the risk-sensitive objective function is given by

$$-\frac{1}{\theta}\log\mathbb{E}\left[\exp(-\theta J^P)\right] = \mathbb{E}[J^P] - \frac{\theta}{2}\mathrm{Var}[J^P] + O(\theta^2) \tag{2.15}$$

as $\theta \to 0$. Note that the variance of the payoff is penalized when $\theta > 0$.

The solution, $(C^{\mathrm{OPT}}, u^{\mathrm{OPT}})$, to this problem is said to be the optimal *risk-limiting dynamic contract* for direct load control. This problem of risk-limiting dynamic contract design is a mean-variance constrained-stochastic optimal control problem that is not directly solvable via dynamic programming. In the following sections, we carefully characterize the necessary and sufficient conditions for the constraints on the mean and the variance of the customers' payoff functions. The characterizations allow us to show that the contract design problem can be reformulated as a risk-sensitive control problem with a stochastic target constraint that can be solved by dynamic programming.

The information availabilities to the load-serving entity and the customers in the direct load control program are symmetric, as opposed to the case of indirect load control, in which the load-serving entity has limited observation capability (e.g., [138]). More specifically, in the proposed framework, the load-serving entity can monitor the control and state of the customers' loads in the direct load control program as well as the energy price in the real-time market. Furthermore, the load-serving entity has all the parameters and functions needed to design an optimal contract. That is, it has the information of p, l_i, $\tilde{\sigma}_i$, μ_i, r_0, σ_0, ν, which can be estimated from data as shown in Section 4.4, and the customers' comfort functions r_i and load models for $i = 1, \cdots, n$. In practice, the comfort functions and load models can be identified using a training period. Each customer, in principle, can have the same information. However, customer i needs to know only p, l_i, $\tilde{\sigma}_i$, μ_i, r_0, σ_0, ν, his or her own comfort function and load model. The proposed framework assumes that each customer can monitor the control and state of his or her load in the direct load control program and the energy price in the real-time market. Another important feature of the proposed contract method is that the interactions between the load-serving entity and its customers can be decoupled from each other because one customer's load does not affect those of other customers and the participation payoff and risk-limiting conditions are personalized. This feature will allow us to decentralize the control of loads as shown in Section 2.5.2.

We assume that the contract period $[0, T]$ is a time interval within 24 hours, but the proposed method can handle arbitrary finite time horizons. Therefore, the customers and the load-serving entity, in principle, can renew the contracts every day. However, it may not be convenient for each customer to choose a contract or, equivalently, a participation payoff and a risk share pair every day. This issue can be resolved by automatically choosing the contract for the current day as that for the previous day unless the customer explicitly wants to change it. Daily contracts have a practical advantage: the day-ahead forecasts of the LMP model parameters and demand uncertainty (and outdoor temperature in the case of air conditioners) can be incorporated into the contracts. Therefore, the contracts can be designed using accurate models. We also assume that each customer does not strategically control other loads to modify the forecasted $\tilde{\sigma}_i$ by the load-serving entity. This assumption can be justified in two ways. First, the load-serving entity can make a contract to control multiple loads of a customer so that the customer has little flexibility to change $\tilde{\sigma}_i$. Second, even if the customer strategically affects $\tilde{\sigma}_i$ in one day, the customer's gain in the next day is marginal because the contract is renewed with a new estimate $\tilde{\sigma}_i$ that incorporates any strategic

behavior. Formally, this problem can be formulated as a Stackelberg differential game, in which the the load-serving entity chooses the estimates of l_i and $\tilde{\sigma}_i$ for the contract period $[0, T]$ assuming that the customer has no incentive to deviate from $\tilde{\sigma}_i$ in the contract period. A similar problem is considered in our previous work [138]. This Stackelberg differential game problem is out of the scope of this chapter and will be addressed in our future work on risk-limiting dynamic contracts for indirect load control.

2.4 Risk-limiting compensation

The risk-limiting condition (2.13) is an inequality constraint on the variance of each agent's payoff. This constraint hinders us from using the dynamic programming principle to solve the contract design problem (3.32). In this section, we characterize a condition on the end-time compensation, which is equivalent to the risk-limiting condition. It turns out that the new equivalent condition allows us to formulate the contract design problem as a risk-sensitive control problem that can be solved by dynamic programming.

We begin by defining the following set of stochastic processes: let Γ^i be the set of processes $\xi^i := \{\xi^i_t\}_{0\le t\le T}$, $\xi^i_t = (\xi^{i,1}_t, \xi^{i,2}_t) \in \mathbb{R}^{1\times 2}$ such that

(i) ξ^i_t is $\mathcal{F}^{(i)}_t$-progressively measurable;

(ii) $\mathbb{E}\left[\int_0^T \|\xi^i_t\|^2 dt\right] < \infty$

for $i = 1, \cdots, n$. We also let $\Gamma := \Gamma^1 \times \cdots \times \Gamma^n$. In the next lemma, we show that there exists a unique process in this set such that its integral over the Brownian motion $W^{(i)} := (W^0, W^i)$ corresponds to the difference between the agent's payoff and its mean value.

Lemma 1. *Fix $i \in \{1, \cdots, n\}$. Given $C^i \in \mathbb{C}^i$ and $u^i \in \mathbb{U}^i$, there exists a unique (up to a set of measure zero) stochastic process $\xi^i = \{\xi^i_t\}_{0\le t\le T} \in \Gamma^i$ such that*

$$J^A_i[C^i, u^i] - \mathbb{E}[J^A_i[C^i, u^i]] = \int_0^T \xi^i_t dW^{(i)}_t. \tag{2.16}$$

Proof. Fix $C^i \in \mathbb{C}^i$ and $u^i \in \mathbb{U}^i$. We introduce a new process

$$q^i_t := \mathbb{E}\left[\int_t^T r^A_i(x^i_s, u^i_s)ds + C^i \;\middle|\; \mathcal{F}^{(i)}_t\right].$$

Here, the expectation is conditioned over the filtration $\{\mathcal{F}^{(i)}_t\}_{0\le t\le T}$ generated by the Brownian motion $W^{(i)} = (W^0, W^i)$. We notice that the process

$$q^i_t + \int_0^t r^A_i(x^i_s, u^i_s)ds = \mathbb{E}\left[\int_0^T r^A_i(x^i_s, u^i_s)ds + C^i \;\middle|\; \mathcal{F}^{(i)}_t\right]$$

is martingale. Recall that there exist constants B_0 and B_1 such that

$$\left|r_i^A(\boldsymbol{x}, \boldsymbol{u})\right| \leq B_0 + B_1|\boldsymbol{x}|$$

for all $\boldsymbol{x} \in \mathbb{R}$ and $\boldsymbol{u} \in \mathcal{U}^i$. Therefore, we have

$$\left(\int_0^t r_i^A(x_s^i, u_s^i)ds\right)^2 \leq \int_0^t (B_0 + B_1|x_s^i|)^2 ds,$$

which implies that for $t \in [0, T]$

$$\mathbb{E}\left[\left(\int_0^t r_i^A(x_s^i, u_s^i)ds\right)^2\right] < \infty \tag{2.17}$$

because $x^i \in \mathbb{L}^2(0, T)$. From the definition of q_t^i, we deduce that

$$\mathbb{E}\left[(q_t^i)^2\right] < \infty. \tag{2.18}$$

Due to the inequalities (2.17) and (2.18), we obtain

$$\mathbb{E}\left[\left(q_t^i + \int_0^t r_i^A(x_s^i, u_s^i)ds\right)^2\right] < \infty.$$

The Martingale representation theorem (e.g., [65, 73]) suggests that there exists a unique (up to set of measure zero) process $\bar{\xi}^i = \{\bar{\xi}_t^i\}_{0 \leq t \leq T} \in \Gamma^i$ such that

$$q_T^i + \int_0^T r_i^A(x_t^i, u_t^i)dt = q_0^i + \int_0^T \bar{\xi}_t^i dW_t^{(i)}.$$

We also note that

$$\begin{aligned} q_T^i &= C^i, \\ q_0^i &= \mathbb{E}[J_i^A[C^i, u^i]]. \end{aligned}$$

Therefore, we obtain

$$J_i^A[C^i, u^i] = \mathbb{E}[J_i^A[C^i, u^i]] + \int_0^T \sigma_i^A(t)dW_t^i + \int_0^T \bar{\xi}_t^i dW_t^{(i)}.$$

Set $\xi_t^{1,i} := \bar{\xi}_t^{1,i}$ and $\xi_t^{2,i} := \bar{\xi}_t^{2,i} + \sigma_i^A(t)$, then ξ^i is in Γ^i and satisfies (2.16). □

This lemma represents the agent's payoff as the sum of its mean value and the Itô integral of the new process ξ^i along the Brownian motion $W^{(i)}$. The following theorem suggests that this representation allows reformulation of the risk-limiting condition (2.13).

Theorem 1. *Fix $i \in \{1, \cdots, n\}$ and $u^i \in \mathbb{U}^i$. The risk-limiting condition holds, i.e.,*

$$\mathrm{Var}[J_i^A[C^i, u^i]] \leq S_i \tag{2.19}$$

if and only if there exists a unique (up to set of measure zero) $\gamma^i \in \Gamma^i$ such that

$$C^i = \mathbb{E}[J_i^A[C^i, u^i]] - \int_0^T r_i^A(x_t^i, u_t^i)dt - \int_0^T \sigma_i^A(t)dW_t^i + \int_0^T \gamma_t^i dW_t^{(i)} \tag{2.20}$$

and

$$\mathbb{E}\left[\int_0^T \|\gamma_t^i\|^2 dt\right] \leq S_i. \tag{2.21}$$

Proof. Suppose that there exists $\gamma^i \in \Gamma^i$ such that (2.20) and (2.21) hold. Then,

$$J_i^A[C^i, u^i] - \mathbb{E}[J_i^A[C^i, u^i]] = \int_0^T \gamma_t^i dW_t^{(i)}.$$

Due to the Itô's isometry, we have

$$\mathrm{Var}[J_i^A[C^i, u^i]] = \mathbb{E}\left[\int_0^T \|\gamma_t^i\|^2 dt\right].$$

Combining this equality and (2.21), we obtain the risk-limiting condition (2.19).

Suppose now that the risk-limiting condition (2.19) holds. Lemma 2 suggests that there exists $\xi^i \in \Gamma^i$ such that

$$J_i^A[C^i, u^i] - \mathbb{E}[J_i^A[C^i, u^i]] = \int_0^T \xi_t^i dW_t^{(i)}.$$

The variance of agent i's payoff is given by

$$\mathrm{Var}[J_i^A[C^i, u^i]] = \mathbb{E}\left[\int_0^T \|\xi_t^i\|^2 dt\right].$$

Due to the risk-limiting condition (2.19), we have

$$\mathbb{E}\left[\int_0^T \|\xi_t^i\|^2 dt\right] \leq S_i.$$

Therefore, $\xi^i \in \Gamma^i$ satisfies both (2.20) and (2.21). Such a process must be unique up to a set of measure zero due to Lemma 2. □

The next corollary suggests a way to construct the end-time compensation given $u \in \mathbb{U}$ and $\gamma \in \Gamma$.

Corollary 1. *Fix $u \in \mathbb{U}$ and $\gamma \in \Gamma$ such that*

$$\mathbb{E}\left[\int_0^T \|\gamma_t^i\|^2 dt\right] \leq S_i \tag{2.22}$$

for $i = 1, \cdots, n$. The risk-limiting condition (2.13) *holds if and only if the end-time compensation, $C \in \mathbb{C}$, satisfies*

$$C^i = \mathbb{E}[J_i^A[C^i, u^i]] - \int_0^T r_i^A(x_t^i, u_t^i)dt - \int_0^T \sigma_i^A(t)dW_t^i + \int_0^T \gamma_t^i dW_t^{(i)}$$

for $i = 1, \cdots, n$.

Theorem 1 and Corollary 4 imply that, given $u^i \in \mathbb{U}^i$, determining $C^i \in \mathbb{C}^i$ is equivalent to choosing $\gamma^i \in \Gamma^i$ such that it satisfies (2.22). In the next section, we consider $\gamma^i \in \Gamma^i$ as a two-dimensional decision variable and then construct the end-time compensation using an optimal γ^i. We also show that the mean of agent i's payoff is given by the participation payoff b_i if an optimal contract is chosen. However, even if we reformulate the contract design problem (3.32) as a stochastic optimal control in which the decision variables are u^i and γ^i, $i = 1, \cdots, n$, the integral constraint (2.22) prohibits us from using dynamic programming to solve the reformulated problem. We resolve this issue in the following section by introducing new state and control variables.

2.5 Optimal contract design

We now propose the solution method for risk-limiting dynamic contract design problem (3.32) given $(b_i, S_i) \in \Lambda$ for $i = 1, \cdots, n$. Recall that the principal's objective is to maximize its risk-sensitive payoff with a guarantee that the participation payoff and risk-limiting conditions for all the agents are satisfied. The risk of each agent's payoff being too small is limited by the variance constraint (2.14e), which is the risk-limiting condition. On the other hand, small values of the principal's payoff are penalized by maximizing the risk-sensitive objective function (3.2a) when the risk-aversion coefficient θ is positive. In other words, if the principal is risk-averse, she transfers her financial risk to the agents as long as the agents' risk-limiting conditions are respected.

The contract design problem (3.32) is a constrained stochastic optimal control problem that cannot be directly solved by dynamic programming [10]. The stochastic maximum principle approach may be used to handle the constraints [102, 20]. However, it is not capable of finding a globally optimal solution unless the problem is concave in both C and u. To obtain a globally optimal solution, we reformulate the problem as a risk-sensitive control problem that can be solved by dynamic programming. The key idea is to introduce two new state variables. The first state variable's value at the terminal time allows us to construct the end-time compensation using Theorem 1. The second variable is used to reformulate the

constraint (2.21) on the new decision variable, γ, as an SDE. However, the dynamic programming approach, in general, has an inherent scalability issue: the computational complexity exponentially increases as the system dimension increases (e.g., [12]). We overcome this scalability issue by proposing an approximate decomposition of the contract design problem for all agents into n lower-dimensional contract design problems, each for a single agent, where n is the number of agents.

2.5.1 Dynamical system approach to limit risks

We show that the solution of the contract design problem (3.32) can be obtained by solving the following risk-sensitive control problem:

$$\begin{aligned}
\max_{u\in\mathbb{U},\gamma\in\Gamma,\zeta\in\Gamma} \quad & -\frac{1}{\theta}\log\mathbb{E}\left[\exp\left(-\theta\bar{J}^P[u,\gamma,\zeta]\right)\right] && \text{(2.23a)}\\
\text{subject to} \quad & dw_t = r_0(\nu(t)-w_t)dt+\sigma_0(t)dW_t^0 && \text{(2.23b)}\\
& dx_t^i = f_i(x_t^i,u_t^i)dt && \text{(2.23c)}\\
& dv_t^i = -r_i^A(x_t^i,u_t^i)dt+\gamma_t^{i,1}dW_t^0+(\gamma_t^{i,2}-\sigma_i^A(t))dW_t^i && \text{(2.23d)}\\
& v_0^i = b_i && \text{(2.23e)}\\
& dy_t^i = -\|\gamma_t^i\|^2dt+\zeta_t^i dW_t^{(i)} && \text{(2.23f)}\\
& y_0^i = S_i && \text{(2.23g)}\\
& y_T^i \geq 0 \quad \text{a.s.}, && \text{(2.23h)}
\end{aligned}$$

where $\bar{J}^P$ is the reformulated principal's payoff given by

$$\bar{J}^P[u,\gamma,\zeta] := \sum_{i=1}^n\left(\int_0^T r_i^P(w_t,x_t^i,u_t^i)+\int_0^T\sigma_i^P(w_t)dW_t^i-v_T^i\right).$$

Note that we now view $\gamma\in\Gamma$ as a decision variable instead of $C\in\mathbb{C}$. This is feasible due to Theorem 1 and Corollary 4. The new state y^i and new decision variable ζ^i handle the integral constraint (2.21) in Theorem 1. Intuitively speaking, the first new state variable v_t^i represents agent i's expected future payoff with a modified diffusion term. The second new state y_t^i can be interpreted as the remaining amount of risk that agent i can bear from time t. Having these interpretations, we can show that the terminal value of the first new state variable can be used to construct an optimal end-time compensation and that of the second new state variable must be greater than or equal to zero to satisfy agent i's risk-limiting condition. These two claims are shown in Theorem 2.

Another important observation is that the problem is now defined in the augmented state space of w_t, $x_t := (x_t^1,\cdots,x_t^n)\in\mathbb{R}^n$, $v_t := (v_t^1,\cdots,v_t^n)\in\mathbb{R}^n$ and $y_t := (y_t^1,\cdots,y_t^n)\in\mathbb{R}^n$. Therefore, the total system dimension is $3n+1$. This reformulated problem can be decentralized into n three dimensional risk-sensitive control problems, as shown in Section 2.5.2.

Theorem 2. *Let $(u^{OPT}, \gamma^{OPT}, \zeta^{OPT})$ be the solution to* (2.23)*. We also let x^{OPT}, v^{OPT} and y^{OPT} denote the processes driven by* (2.23c)*,* (2.23d) *and* (2.23f) *with $(u^{OPT}, \gamma^{OPT}, \zeta^{OPT})$, respectively. Define*

$$C^{OPT,i} := v_T^{OPT,i} \tag{2.24}$$

for $i = 1, \cdots, n$. Then, (C^{OPT}, u^{OPT}) is an optimal risk-limiting dynamic contract, i.e., it solves (3.32)*.*

Proof. We first observe that

$$\begin{aligned} J_i^A[C^{\mathrm{OPT},i}, u^{\mathrm{OPT},i}] &= \int_0^T r_i^A(x_t^{\mathrm{OPT},i}, u_t^{\mathrm{OPT},i})dt + \int_0^T \sigma_i^A(t)dW_t^i + v_T^{\mathrm{OPT},i} \\ &= b_i + \int_0^T \gamma_t^{\mathrm{OPT},i} dW_t^{(i)} \end{aligned} \tag{2.25}$$

due to the SDE (2.23d) with the initial condition (2.23e). Therefore, we have

$$\mathbb{E}[J_i^A[C^{\mathrm{OPT},i}, u^{\mathrm{OPT},i}]] = b_i, \tag{2.26}$$

which implies that the participation payoff condition (2.14d) holds. Furthermore, the variance of agent i's payoff with $(C^{\mathrm{OPT},i}, u^{\mathrm{OPT},i})$ is given by

$$\mathrm{Var}[J_i^A[C^{\mathrm{OPT},i}, u^{\mathrm{OPT},i}]] = \mathbb{E}\left[\int_0^T \|\gamma_t^{\mathrm{OPT},i}\|^2 dt\right]$$

due to the Itô's isometry. We also notice that

$$y_T^{\mathrm{OPT},i} = S_i - \int_0^T \|\gamma_t^{\mathrm{OPT},i}\|^2 dt + \int_0^T \zeta_t^{\mathrm{OPT},i} dW_t^{(i)},$$

which suggests that

$$\mathrm{Var}[J_i^A[C^{\mathrm{OPT},i}, u^{\mathrm{OPT},i}]] = \mathbb{E}\left[S_i - y_T^{\mathrm{OPT},i} + \int_0^T \zeta_t^{\mathrm{OPT},i} dW_t^{(i)}\right].$$

Hence, if $y_T^{\mathrm{OPT},i} \geq 0$ a.s., the risk-limiting condition (2.14e) also holds. Therefore, $(C^{\mathrm{OPT}}, u^{\mathrm{OPT}})$ is a feasible dynamic contract.

Suppose that $(C^{\mathrm{OPT}}, u^{\mathrm{OPT}})$ is not a solution of (3.32) and select a solution, $(\hat{C}, \hat{u})$, of (3.32). Theorem 1 suggests that there exists a unique (up to set of measure zero) $\hat{\gamma} \in \Gamma$ such that

$$\hat{C}^i = \mathbb{E}[J_i^A[\hat{C}^i, \hat{u}^i]] - \int_0^T r_i^A(\hat{x}_t^i, \hat{u}_t^i)dt - \int_0^T \sigma_i^A(t)dW_t^i + \int_0^T \hat{\gamma}_t^i dW_t^{(i)}$$

and

$$\mathbb{E}\left[\int_0^T \|\hat{\gamma}_t^i\|^2 dt\right] \leq S_i. \tag{2.27}$$

We claim that $(\hat{u}, \hat{\gamma})$ satisfies all the constraints of (2.23) with the stochastic process $\hat{v} := \{\hat{v}_t\}_{0\leq t\leq T}$ defined as

$$\hat{v}_t^i := \hat{v}_0^i - \int_0^t r_i^A(\hat{x}_s^i, \hat{u}_s^i)ds + \int_0^t (\hat{\gamma}_s^i - \sigma_i^A(s))dW_s^i$$

for some initial value $\hat{v}_0^i \in \mathbb{R}$ such that

$$\hat{v}_T^i = \hat{C}^i. \tag{2.28}$$

It is clear that the process $\hat{v}^i$ satisfies the SDE (2.23d) for $i = 1, \cdots, n$ by definition. Additionally, note that

$$\mathbb{E}[J_i^A[\hat{C}^i, \hat{u}^i]] = \hat{v}_0^i.$$

Suppose that $\hat{v}$ does not satisfy the initial condition (2.23e). We first assume that there exists $j \in \{1, \cdots, n\}$ such that $\hat{v}_0^j > b_j$. Define a new end-time compensation C' as

$$C'^i := \begin{cases} \hat{C}^j - (\hat{v}_0^j - b_j) & \text{if } i = j \\ \hat{C}^i & \text{otherwise.} \end{cases}$$

We then have

$$\begin{aligned} \mathbb{E}[J_j^A[C'^j, \hat{u}^j]] - C'^j &= \mathbb{E}[J_j^A[\hat{C}^j, \hat{u}^j]] - \hat{C}^j \\ &= \hat{v}_0^j - \hat{C}^j, \end{aligned}$$

which implies that

$$\mathbb{E}[J_j^A[C'^j, \hat{u}^j]] = b_j.$$

Therefore, $(C', \hat{u})$ satisfies the participation payoff condition (2.14d) for all $i = 1, \cdots, n$. The risk-limiting condition also holds with $(C', \hat{u})$ because the difference between C' and $\hat{C}$ is deterministic. On the other hand, we notice that

$$J^P[C', \hat{u}] > J^P[\hat{C}, \hat{u}]$$

because $C'^j < \hat{C}^j$ and $C'^i = \hat{C}^i$ for $i \neq j$. The contract $(C', \hat{u})$ satisfies all the constraints of (3.32) and is strictly better than $(\hat{C}, \hat{u})$. This is a contradiction because $(\hat{C}, \hat{u})$ solves the contract design problem (3.32). Therefore, $\hat{v}_0^i$ must be equal to the participation payoff b_i for $i = 1, \cdots, n$. Hence, $(\hat{u}, \hat{\gamma})$ satisfies all the constraints of (2.23) with the processes $\hat{y}$ and $\hat{x}$, where $\hat{x}$ solves (2.23c) with the control $\hat{u}$.

We define a stochastic process $\tilde{y}^i := \{\tilde{y}_t^i\}_{0\leq t\leq T}$ as

$$\tilde{y}_t^i := \mathbb{E}\left[\int_t^T \|\hat{\gamma}_s^i\|^2 ds \;\middle|\; \mathcal{F}_t^{(i)}\right].$$

Note that

$$\tilde{y}_t^i + \int_0^t \|\hat{\gamma}_s^i\|^2 ds = \mathbb{E}\left[\int_0^T \|\hat{\gamma}_t^i\|^2 dt \;\middle|\; \mathcal{F}_t^{(i)}\right]$$

is martingale. Furthermore, $\mathbb{E}\left[\left(\tilde{y}_t^i + \int_0^t \|\hat{\gamma}_s^i\|^2 ds\right)^2\right] < \infty$ because $\hat{\gamma}^i \in \Gamma^i$. Therefore, the martingale representation theorem suggests that there exists a unique (up to a set of measure zero) $\hat{\zeta}^i \in \Gamma^i$ such that

$$\tilde{y}_t^i + \int_0^t \|\hat{\gamma}_t^i\|^2 dt = \tilde{y}_0^i + \int_0^t \hat{\zeta}_t^i dW_t^{(i)}.$$

Therefore, the process $\tilde{y}^i$ solves (2.23f) with $(\hat{\gamma}, \hat{\zeta})$. Due to (2.27), we also have

$$\tilde{y}_0^i \leq S_i$$

and hence the constraint (2.23g) is satisfied. We define another stochastic process $\hat{y}^i := \{\hat{y}_t^i\}_{0\leq t\leq T}$ as

$$\hat{y}_t^i := \tilde{y}_t^i + S_i - \tilde{y}_0^i$$

for $t \in [0, T]$. Therefore, we have $\hat{y}_0^i = S_i$ and $\hat{y}_T^i \geq 0$ because $\tilde{y}_T^i = 0$ by definition. Hence, $\hat{y}^i$ satisfies the constraints (2.23g) and (2.23h).

Since $(\hat{C}, \hat{u})$ solves (3.32), while $(C^{\mathrm{OPT}}, u^{\mathrm{OPT}})$ does not, the following inequality holds:

$$\mathbb{E}[-\exp(-\theta J^P[\hat{C}, \hat{u}])] > \mathbb{E}[-\exp(-\theta J^P[C^{\mathrm{OPT}}, u^{\mathrm{OPT}}])].$$

This inequality can be rewritten as

$$\begin{aligned}
&\mathbb{E}\left[-\exp\left(-\theta \sum_{i=1}^n \left(\int_0^T r_i^P(w_t, \hat{x}_t^i, \hat{u}_t^i)dt + \int_0^T \sigma_i^P(w_t)dW_t^i - \hat{v}_T^i\right)\right)\right] \\
&> \mathbb{E}\left[-\exp\left(-\theta \sum_{i=1}^n \left(\int_0^T r_i^P(w_t, x_t^{\mathrm{OPT},i}, u_t^{\mathrm{OPT},i})dt + \int_0^T \sigma_i^P(w_t)dW_t^i - v_T^{\mathrm{OPT},i}\right)\right)\right]
\end{aligned}$$

due to (2.28) and (2.24). This is contradictory to the fact that $(u^{\mathrm{OPT}}, \gamma^{\mathrm{OPT}}, \zeta^{\mathrm{OPT}})$ is a solution to (2.23). Therefore, $(C^{\mathrm{OPT}}, u^{\mathrm{OPT}})$ should solve (3.32) and hence an optimal risk-limiting dynamic contract. □

We observe that the agents' expected payoff must be equal to their participation payoffs from (2.26), i.e., the inequalities (2.14d) for the participation payoff are always binding at an optimal contract. Intuitively speaking, if the agent's expected payoff is strictly greater than his or her participation payoff, the principal has an incentive to decrease the end-time compensation for the agent. The equality (2.25) also suggests that each agent's payoff can be completely characterized by his or her participation payoff and the new control variable γ if an optimal contract is executed.

Corollary 2. *Agent i's payoff with an optimal risk-limiting dynamic contract (C^{OPT}, u^{OPT}) is given by*

$$J_i^A[C^{OPT,i}, u^{OPT,i}] = b_i + \int_0^T \gamma_t^{OPT,i} dW_t^{(i)}$$

for $i = 1, \cdots, n$. Therefore,

$$\mathbb{E}\left[J_i^A[C^{OPT,i}, u^{OPT,i}]\right] = b_i.$$

2.5.2 Decoupled contract design and decentralized control

We propose an approximate decomposition of the contract design problem (3.32) into n low dimensional problems using the fact that the system dynamics (2.14c), the participation payoff condition (2.14d) and the risk-limiting condition (2.14e) for one agent are decoupled from those for other agents and that $W^1, \cdots, W^n$ are mutually independent. The approximate solution obtained using this decomposition has a guaranteed suboptimality bound. This decomposition enables the direct load control program with the proposed dynamic contracts to handle a large population of agents without scalability issues.

More specifically, for each $i \in \{1, \cdots, n\}$, the approximate risk-limiting dynamic contract for agent i can be obtained by solving the following risk-sensitive control problem:

$$\begin{aligned}
\max_{\substack{u^i \in \mathbb{U}^i, \\ \gamma^i \in \Gamma^i, \zeta^i \in \Gamma^i}} \quad & -\frac{1}{\theta} \log \mathbb{E}\left[\exp\left(-\theta \bar{J}_i^P[u^i, \gamma^i, \zeta^i]\right)\right] \\
\text{subject to} \quad & dw_t = r_0(\nu(t) - w_t)dt + \sigma_0(t) dW_t^0 \\
& dx_t^i = f_i(x_t^i, u_t^i)dt \\
& dy_t^i = -\|\gamma_t^i\|^2 dt + \zeta_t^i dW_t^{(i)} \\
& y_0^i = S_i \\
& y_T^i \geq 0 \quad \text{a.s.},
\end{aligned} \tag{2.29}$$

where

$$\begin{aligned}
\bar{J}_i^P[u^i, \gamma^i, \zeta^i] := -b_i &+ \int_0^T (r_i^P(w_t, x_t^i, u_t^i) + r_i^A(x_t^i, u_t^i))dt \\
&+ \int_0^T (\sigma_i^P(w_t) + \sigma_i^A(t))dW_t^i - \int_0^T \gamma_t^i dW_t^{(i)}.
\end{aligned}$$

Note that the system for v^i is absorbed into the modified payoff function $\bar{J}_i^P$. It is clear that this decomposition is exact when $\sigma_0 \equiv 0$ due to the mutual independence of $\{W^1, \cdots, W^n\}$. In addition, the following proposition suggests that the approximate contract obtained using the proposed decomposition has a provable suboptimality bound, which can be computed *a posteriori*.

Proposition 1. *Let* $(u^{OPT}, \gamma^{OPT}, \zeta^{OPT})$ *and* $(u^{*i}, \gamma^{*i}, \zeta^{*i})$ *be the solutions to* (3.32) *and* (2.29)*, respectively. We also let* $\bar{u}^i$ *be the solution to*

$$\begin{aligned} \max_{u^i \in \mathbb{U}^i} \quad & \mathbb{E}[\bar{J}_i^P[u^i, 0, 0]] \\ \textit{subject to} \quad & dw_t = r_0(\nu(t) - w_t)dt + \sigma_0(t)dW_t^0 \\ & dx_t^i = f_i(x_t^i, u_t^i)dt. \end{aligned} \tag{2.30}$$

Suppose that $\mathbb{E}[\bar{J}^P[u, 0, 0]] > 0$*, where* $\bar{u} := (\bar{u}^1, \cdots, \bar{u}^n)$*. Set*

$$\rho := \frac{-\frac{1}{\theta}\log \mathbb{E}\left[\exp\left(-\theta \bar{J}^P[u^*, \gamma^*, \zeta^*]\right)\right]}{\mathbb{E}[\bar{J}^P[\bar{u}, 0, 0]]}.$$

Then, the following suboptimality bound holds:

$$\rho\left(-\frac{1}{\theta}\log \mathbb{E}\left[\exp\left(-\theta \bar{J}^P[u^{OPT}, \gamma^{OPT}, \zeta^{OPT}]\right)\right]\right) \leq -\frac{1}{\theta}\log \mathbb{E}\left[\exp\left(-\theta \bar{J}^P[u^*, \gamma^*, \zeta^*]\right)\right] \tag{2.31}$$

for $\theta > 0$*.*

Proof. Using Jensen's inequality, we have

$$\begin{aligned} -\frac{1}{\theta}\log \mathbb{E}\left[\exp\left(-\theta \bar{J}^P[u^{\mathrm{OPT}}, \gamma^{\mathrm{OPT}}, \zeta^{\mathrm{OPT}}]\right)\right] &\leq -\frac{1}{\theta}\log \exp\left(-\theta \mathbb{E}[\bar{J}^P[u^{\mathrm{OPT}}, \gamma^{\mathrm{OPT}}, \zeta^{\mathrm{OPT}}]]\right) \\ &= \mathbb{E}[\bar{J}^P[u^{\mathrm{OPT}}, \gamma^{\mathrm{OPT}}, \zeta^{\mathrm{OPT}}]]. \end{aligned}$$

When the principal is risk-neutral, the principal's payoff is independent of (γ, ζ) and setting $(\gamma, \zeta) = (0, 0)$ always satisfies the risk-limiting condition. From this observation, we claim that $(\bar{u}, 0, 0)$ solves the problem (3.32) when $\theta = 0$, i.e., the objective function is replaced with $\mathbb{E}[\bar{J}^P[u, \gamma, \zeta]]$. We first note that $(\bar{u}, 0, 0)$ satisfies all the constraints in (3.32). Suppose that $(\bar{u}, 0, 0)$ does not solve (3.32) and choose a solution, $(\hat{u}, \hat{\gamma}, \hat{\zeta})$, of (3.32). Because $\mathbb{E}[\bar{J}_i^P[\hat{u}^i, \hat{\gamma}, \hat{\zeta}]] = \mathbb{E}[\bar{J}_i^P[\hat{u}^i, 0, 0]]$ and $\hat{u}$ satisfies the constraint of (2.30), the following inequality holds:

$$\mathbb{E}[\bar{J}_i^P[\hat{u}^i, \hat{\gamma}^i, \hat{\zeta}^i]] = \mathbb{E}[\bar{J}_i^P[\hat{u}^i, 0, 0]] \leq \mathbb{E}[\bar{J}_i^P[\bar{u}^i, 0, 0]].$$

Therefore, we have

$$\begin{aligned} \mathbb{E}[\bar{J}^P[\hat{u}, \hat{\gamma}, \hat{\zeta}]] &= \sum_{i=1}^n \mathbb{E}[\bar{J}_i^P[\hat{u}^i, \hat{\gamma}^i, \hat{\zeta}^i]] \\ &\leq \sum_{i=1}^n \mathbb{E}[\bar{J}_i^P[\bar{u}^i, 0, 0]] = \mathbb{E}[\bar{J}^P[\bar{u}, 0, 0]]. \end{aligned}$$

This inequality is contradictory to the fact that $(\bar{u}, 0, 0)$ does not solve (3.32). Therefore,

$$\mathbb{E}[\bar{J}^P[u^{\mathrm{OPT}}, \gamma^{\mathrm{OPT}}, \zeta^{\mathrm{OPT}}]] \leq \mathbb{E}[\bar{J}^P[\bar{u}, 0, 0]].$$

As a result, the suboptimality bound (2.31) holds. □

Note that the suboptimality bound can be computed by solving (2.29) and (2.30) for each i while it is not feasible to directly solve (3.32) for $n > 1$. This proposition implies that the proposed decomposition tends to be exact as the coefficient θ of the principal's risk aversion goes to zero because $\rho \to 1$ as $\theta \to 0$. Furthermore, the approximate contract (C^*, u^*) satisfies the participation-payoff condition (2.12) and the risk-limiting condition (2.13).

Due to this decomposition, the contract design for agent i only requires the state space, $\mathbb{R}^3$, of (w_t, x_t^i, y_t^i) rather than the full joint state space, $\mathbb{R}^{3n+1}$, of (w_t, x_t, v_t, y_t). Therefore, the computational complexity of designing a risk-limiting contract for an agent is independent of the total number of agents. The decomposed problem for agent i is solved via dynamic programming over the reduced state space, $\mathbb{R}^3$, of (w_t, x_t^i, y_t^i) as follows. We set the feasible set of control as

$$\Omega^i := \{(u^i, \gamma^i, \zeta^i) \in \mathbb{U}^i \times \Gamma^i \times \Gamma^i \mid y_T^i \geq 0 \text{ a.s.}\}.$$

To synthesize a risk-limiting dynamic contract for agent i, we first define the value function of (2.29) associated with agent i as

$$\begin{aligned}\phi_i(\boldsymbol{w}, \boldsymbol{x}_i, \boldsymbol{y}_i, t) := \max_{(u^i,\gamma^i,\zeta^i)\in\Omega^i} & -\frac{1}{\theta}\log \mathbb{E}_{\boldsymbol{w},\boldsymbol{x}_i,\boldsymbol{y}_i,t} \\ & \left[\exp\left(-\theta\left(\int_t^T R_i(w_s, x_s^i, u_s^i)ds + \int_t^T G_i(w_s, \gamma_s^i)dW_s^{(i)} - b_i\right)\right)\right],\end{aligned} \tag{2.32}$$

where $\mathbb{E}_{\boldsymbol{w},\boldsymbol{x}_i,\boldsymbol{y}_i,t}[A]$ denotes the expectation of A conditioned on $(w_t, x_t^i, y_t^i) = (\boldsymbol{w}, \boldsymbol{x}_i, \boldsymbol{y}_i)$, and

$$\begin{aligned}R_i(\boldsymbol{w}, \boldsymbol{x}_i, \boldsymbol{u}) &:= r_i^P(\boldsymbol{w}, \boldsymbol{x}_i, \boldsymbol{u}) + r_i^A(\boldsymbol{x}_i, \boldsymbol{u}), \\ G_i(t, \boldsymbol{w}, \boldsymbol{\gamma}) &:= \begin{bmatrix}-\boldsymbol{\gamma}_1 & -\boldsymbol{\gamma}_2 + \sigma_i^P(\boldsymbol{w}) + \sigma_i^A(t))\end{bmatrix}.\end{aligned}$$

To handle the constraint $y_T^i \geq 0$ a.s., which is often called the *stochastic target constraint*, we use the Hamilton-Jacobi-Bellman (HJB) characterization proposed in [16]. This characterization converts the target constraint into a 'classical' state constraint using the *geometric dynamic programming principle* [121]. The reformulated constraint is embedded in an auxiliary value function. This auxiliary value function is a viscosity solution of an HJB equation. Applying the dynamic programming principle on the original value function (2.32), one can derive a constrained-HJB equation in which the stochastic target constraint is reformulated as the constraints on the auxiliary value function and the control. In our case, the auxiliary value function is a zero function. Let

$$U_i(\boldsymbol{y}_i) := \{(\boldsymbol{u}, \boldsymbol{\gamma}, \boldsymbol{\zeta}) \in \mathcal{U}^i \times \mathbb{R}^2 \times \mathbb{R}^2 \mid \boldsymbol{\gamma} = \boldsymbol{\zeta} = 0 \text{ if } \boldsymbol{y}_i \leq 0\}.$$

Then, the stochastic target constraint is simply incorporated into the following constrained-

HJB equation:

$$\begin{aligned}\frac{\partial \phi_i}{\partial t} + \max_{\substack{(\boldsymbol{u},\boldsymbol{\gamma},\boldsymbol{\zeta}) \\ \in U_i(\boldsymbol{y}_i)}} \Big\{ & (F_i(\boldsymbol{w},\boldsymbol{x}_i,\boldsymbol{u},\boldsymbol{\gamma}) - \theta\Sigma(\boldsymbol{\zeta})G_i(\boldsymbol{w},\boldsymbol{\gamma})^\top)^\top D\phi_i \\ & + R_i(\boldsymbol{w},\boldsymbol{x}_i,\boldsymbol{u}) - \frac{\theta}{2}\|G_i(\boldsymbol{w},\boldsymbol{\gamma})\|^2 - \frac{\theta}{2}\|\Sigma(\boldsymbol{\zeta})^\top D\phi_i\|^2 + \frac{1}{2}\mathrm{tr}(\Sigma(\boldsymbol{\zeta})\Sigma(\boldsymbol{\zeta})^\top D^2\phi_i) \Big\} = 0, \\ \phi_i(\boldsymbol{w},\boldsymbol{x}_i,\boldsymbol{y}_i,T) = & -b_i,\end{aligned} \tag{2.33}$$

whose viscosity solution corresponds to the value function (2.32) [27, 39, 16], where

$$F_i(t,\boldsymbol{w},\boldsymbol{x}_i,\boldsymbol{u},\boldsymbol{\gamma}) := \begin{bmatrix} r_0(\nu(t) - \boldsymbol{w}) \\ f_i(\boldsymbol{x}_i,\boldsymbol{u}) \\ -\|\boldsymbol{\gamma}\|^2 \end{bmatrix},$$

$$\Sigma(t,\boldsymbol{\zeta}) := \begin{bmatrix} \sigma_0(t) & 0 \\ 0 & 0 \\ \boldsymbol{\zeta}_1 & \boldsymbol{\zeta}_2 \end{bmatrix}.$$

2.5.3 Optimal contract as a state feedback strategy

In general, an analytic solution of the HJB equation (2.33) is not available. Therefore, we grid up the state space, which corresponds to the domain of the PDE, and numerically evaluate the solution at the grid points using convergent schemes, e.g., [9, 77]. After solving the HJB equation, we can use the value function to construct an optimal contract (C_i^*, u_i^*) as the following state feedback strategy:

1. set $(w_0, x_0^{*i}, y_0^{*i}) = (\ln \lambda^0, x^{0i}, S_i)$;
2. given $\mathbf{x}_s^{*i} := (w_s^*, x_s^{*i}, y_s^{*i}) \in \mathbb{R}^3$ for $s \in [0,t]$, determine an optimal control action $(u_t^{*i}, \gamma_t^{*i}, \zeta_t^{*i})$ as an element of the arg max in the HJB equation (2.33) at $(w_t^*, x_t^{*i}, y_t^{*i})$;
3. the processes (w^*, x^{*i}, y^{*i}) are evolved with the optimal $(u_t^{*i}, \gamma_t^{*i}, \zeta_t^{*i})$ for $(t, t+dt]$ (the agent performs u_t^{*i});
4. repeat 2) and 3) until $t = T$.
5. the agent is paid by $C^{*i} = v_T^{*i}$.

Note that this rule, the value function ϕ and the initial values $(\ln \lambda^0, x^{0i}, S_i)$ must be specified in the contract. The agent must follow u^* in the proposed contract under symmetric information. The auxiliary control variables γ^* and ζ^* are used to satisfy the agent's participation payoff and the risk-limiting conditions. A more detailed discussion regarding how to synthesize an optimal control using a viscosity solution of an associated HJB equation can be found in [8] even when the viscosity solution is not differentiable.

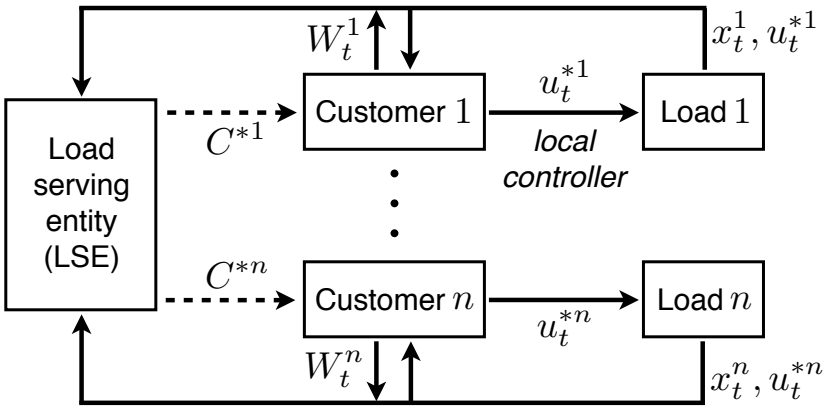

Figure 2.1: Implementation of the proposed contracts: the controls of loads can be decentralized with a broadcast of price (LMP) information, while a centralized monitoring is required. The compensations are provided at the end of the contract period.

Note that the control action for one load does not affect the state of another load. Furthermore, the optimal control for a load is given as state feedback, where the state variables only require the energy price in the real-time market and the local information of the load. Therefore, the proposed control can be decentralized with a broadcast of the price information, i.e., the local controller in which the optimal control strategy is programmed is sufficient for the implementation of the contract as depicted in Figure 2.1. On the other hand, the load-serving entity still needs to monitor the state, the control and the forecast error for each customer to ensure that each customer follows the optimal control strategy written in the contract. The total power consumption of each customer monitored by a smart meter can be used to compute the forecast error. In addition, the sensors for loads such as thermostats provide the state and control information. The monitored information could be transferred to the load-serving entity through a one-way data connection such as the Internet. The information gathered by the monitoring is also used to compute the optimal compensation provided to each customer at the end of the contract period.

2.6 Application to direct load control for financial risk management

In this section, we apply the proposed risk-limiting dynamic contracts to direct load control. The performance and usefulness of the novel direct load control program for financial risk management are demonstrated using the data of LMPs in the Electricity Reliability Council of Texas (ERCOT) and the electric energy consumption of customers in Austin, Texas.

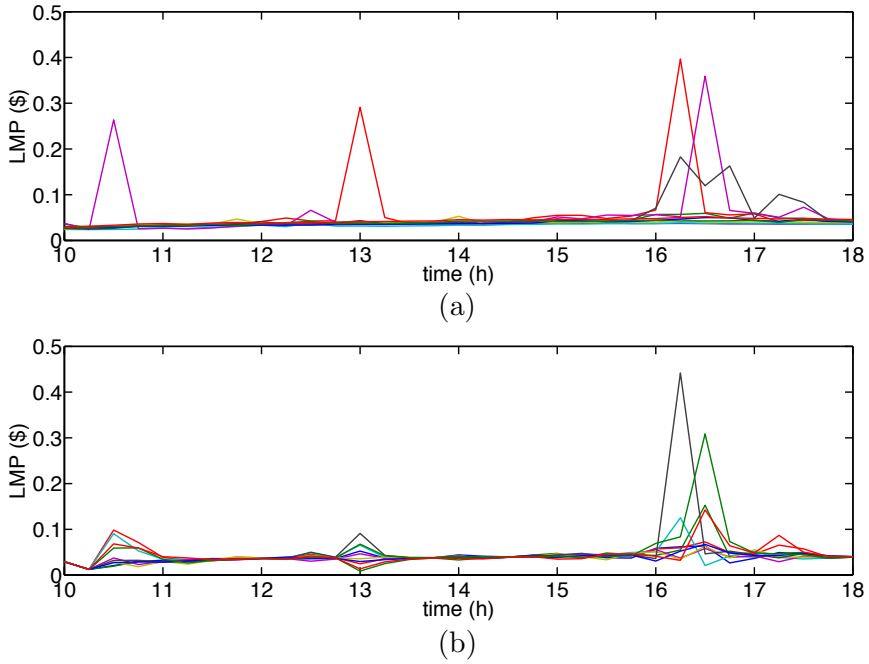

Figure 2.2: (a) Locational marginal price (LMP) in Austin, Texas, from July 1, 2013 to July 10, 2013. (b) Ten sampled trajectories of LMP, $\{\lambda_t\}_{0\le t\le T}$, generated by our identified price model (2.2).

2.6.1 Data assimilation

We consider a scenario in which each customer provides one air conditioner for the proposed direct load control program. We use the ETP model (2.6) for the customer's indoor temperature dynamics given in Example 1. The set of feasible control values is chosen as $\mathcal{U}^i := \{0, 2\}$, assuming that customer i's air conditioner consumes 0kW in its OFF state and 2kW in its ON state. The model parameters are chosen as $\alpha_i = 0.1$ and $\kappa_i = 1.5$, which are calculated based on the Residential module user's guide from GridLAB-D and are physically reasonable [46]. Customer i's comfort level is chosen as (2.11) in Example 2, with $\omega_i = 0.15$. We set the customer's desirable indoor temperature range, $[\underline{\Theta}, \overline{\Theta}]$, as $[20°\text{C}, 22°\text{C}]$. We choose the contract period as $[10\text{h}, 18\text{h}]$.

We use customers' electric energy consumption data in Austin, Texas [134] to estimate the load profile $l_i(t)$ and the diffusion coefficient $\tilde{\sigma}_i(t)$ in (2.1). The load $l_i(t)$ is chosen as the mean value of the customer i's power consumption at t other than the air conditioner. For the estimation of $\tilde{\sigma}_i$, we apply the Kalman filter [62, 76] over the data set for the summer period, from June to September 2013, assuming that customer i's energy consumption profile other than the air conditioner for one day in the period represents one sampled trajectory. We then scale the estimated diffusion coefficient by a constant factor such that $\int_0^T \tilde{\sigma}_i(t)^2 dt$ is equal to the variance of the energy consumption data. This scaling guarantees that $\text{Var}[J_i^A[0, 0]]$ is

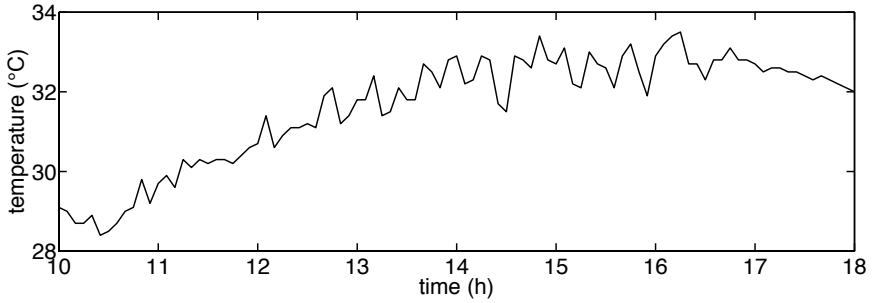

Figure 2.3: Outdoor air temperature, $\Theta(t)$, in Austin, Texas, on July 5, 2013.

equal to the variance of the customer's energy cost.

The price model (2.2) is identified using the ERCOT LMP data at the settlement point, AUSTIN PLANT, from July 1, 2013 to July 10, 2013 [35]. We estimate the parameters, r_0, $\nu(t)$ and $\sigma_0(t)$, by applying the Kalman filter on the transformed linear model (2.3). The LMP data and samples of the price profile generated by the identified model are shown in Figure 2.2.

We use the National Oceanic and Atmospheric Administration Quality Controlled Local Climatological Data in Austin, Texas for the outdoor temperature profile [92]. The outdoor temperature, $\Theta(t)$, at time $t \in [10\text{h}, 18\text{h}]$ is chosen as the temperature on July 5, 2013, at time $t \in [10\text{h}, 18\text{h}]$ and is shown in Figure 2.3.

2.6.2 Comparison to optimal load control by customers

Suppose that customer i does not participate in the direct load control program and has the following payoff:

$$\hat{J}_i^A[u^i] := \int_0^T r_i^A(x_t^i, u_t^i)dt + \int_0^T \sigma_i^A(t)dW_t^i, \tag{2.34}$$

where r_i^A and σ_i^A are given by (2.10). The solution of the following optimal control problem maximizes customer i's expected payoff:

$$\begin{aligned} \max_{u^i \in \mathbb{U}^i} \quad & \mathbb{E}[\hat{J}_i^A[u^i]] := \int_0^T r_i^A(x_t^i, u_t^i)dt \\ \text{subject to} \quad & dx_t^i = f_i(x_t^i, u_t^i)dt. \end{aligned} \tag{2.35}$$

The optimal control can be obtained using the viscosity solution of the following HJB equation [8]:

$$\begin{aligned} &\frac{\partial \hat{\phi}_i(\boldsymbol{x}_i, t)}{\partial t} + \max_{\boldsymbol{u} \in \mathcal{U}^i} \{f(\boldsymbol{x}_i, \boldsymbol{u}) D_{\boldsymbol{x}_i} \hat{\phi}_i(\boldsymbol{x}_i, t) + r_i^A(\boldsymbol{x}_i, \boldsymbol{u})\} = 0, \\ &\hat{\phi}_i(\boldsymbol{x}_i, T) = 0. \end{aligned}$$

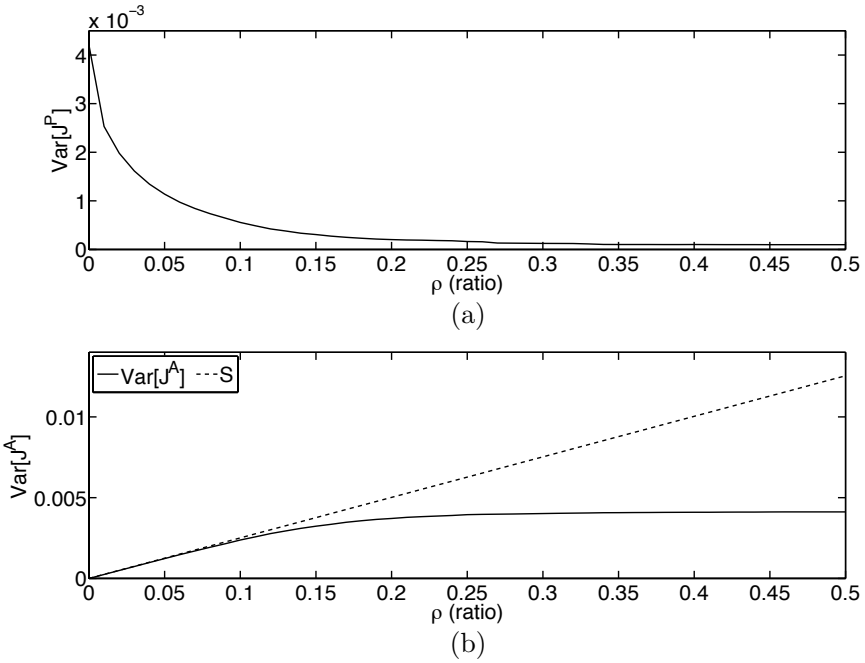

Figure 2.4: The simulation results with the contract: (a) variance of the load-serving entity's payoff. (b) variance of the customer's payoff and the risk limit $S_i = \rho \bar{S}_i$.

Let $\hat{u}^{*i}$ be an optimal control, i.e., a solution to (2.35). We let $\bar{b}_i := \mathbb{E}[\hat{J}_i^A[\hat{u}^{*i}]]$ and $\bar{S}_i := \mathrm{Var}[\hat{J}_i^A[\hat{u}^{*i}]] = \int_0^T \sigma_i^A(t)^2 dt$ be the nominal expected payoff and risk of customer i, respectively.

We now compare the performance of the proposed contracts to that of this optimal load control without a contract. We choose the customer's electricity price as the flat price, $\mu_i \equiv \bar{\mu} = \$0.11$, specified in Austin Energy's electricity tariff for summer [5]. In the absence of a contract, the mean and variance of customer i's payoff are $\bar{b}_i$ and $\bar{S}_i$, respectively. We set $b_i = \bar{b}_i$ and $S_i = \rho \bar{S}_i$ and vary ρ from 0 to 0.3. If customer i enters into the contract, which is the solution of (2.29), then the mean value of the customer's payoff is guaranteed to be greater than or equal to $\bar{b}_i$, and the variance of the customer's payoff is guaranteed to be less than or equal to $\rho \bar{S}_i$.

The results of numerical experiments presented in Figure 2.4 verify the performance of the proposed contract. The coefficient of load-serving entity's risk aversion is chosen as $\theta = 10^{-2}$. As shown in Figure 2.4 (a), The variance of the load-serving entity's payoff decreases as the ratio ρ of the amount of the risk that the customer is willing to bear to the customer's nominal risk increases. On the other hand, the variance of the customer's payoff (blue) increases as the ratio ρ of the amount of the risk that the customer is willing to bear to the customer's nominal risk increases, as shown in Figure 2.4 (b). More importantly, it is less than or equal to the risk limit $S_i = \rho \bar{S}_i$ (red). Therefore, we confirm that the

risk-limiting condition is satisfied.

When the customer does not enter into in the contract, the variance of the load-serving entity's payoff is

$$\begin{aligned}\mathrm{Var}[\hat{J}_i^P[\hat{u}^{*i}]] &:= \mathrm{Var}\left[\int_0^T r_i^P(w_t, \hat{x}_t^{*i}, \hat{u}_t^{*i})dt + \int_0^T \sigma_i^P(w_t)dW_t^i\right] \\ &= 0.0108.\end{aligned}$$

By comparing this variance with the variance of the load-serving entity's payoff when the contract is executed (Figure 2.4 (a)), we note that the contract reduces the load-serving entity's risk by more than 50% even when the customer is extremely risk-averse, i.e., $S_i = 0$. If the customer chooses $S_i \geq 0.2\bar{S}_i$ in the contract, the load-serving entity's risk is decreased by more than 95%.

The mean values of the load-serving entity's payoff with and without the contract are given by

$$\mathbb{E}[J_i^P[C^{*i}, u^{*i}]] = 1.324, \quad \mathbb{E}[\hat{J}_i^P[\hat{u}^{*i}]] = 1.297,$$

respectively. Therefore, the load-serving entity can pay \$0.027 more for the customer without reducing its mean payoff if the customer enters into a contract. In other words, the load-serving entity can incentivize the customer to enter into the contract by increasing the customer's expected payoff by \$0.027.

Figure 2.5 shows the effect of the contract on the control and the indoor temperature. In this set of experiments, we set $S_i = 0.1\bar{S}_i$ and $\theta = 10^{-2}$. When the customer does not enter into the contract, the customer regularly turns on and off the air conditioner such that the indoor temperature is kept near 22°C. This is because the customer wants to save the energy cost by only taking into account the energy price μ_i, which is fixed for all time. On the other hand, if the customer enters into the contract, the room is pre-cooled before 12:30pm when the LMP λ_t is low and not volatile. Another pre-cooling interval is from 3pm to 4pm. This pre-cooling allows to save energy purchase in the real-time market from 4pm to 5pm when the LMP is highly volatile. In other words, the contract properly manages the price risk in the spot market.

2.6.3 Validation of the Brownian motion model using data in terms of the closed-loop performance

The energy consumption process of a customer is modeled by the SDE (2.1). In practice, the load forecast error may not be exactly captured by the diffusion term, $\tilde{\sigma}_i(t)dW_t^i$, with a standard Brownian motion. We test the robustness of the proposed contract method with respect to the deviation of the demand forecast errors in the data from the Brownian motion model. More specifically, we execute the optimal contract synthesized using the Brownian motion model over the data. Then, we compute how much the load-serving entity's and the customer's resulting payoffs differ from their optimal payoffs obtained under the Brownian

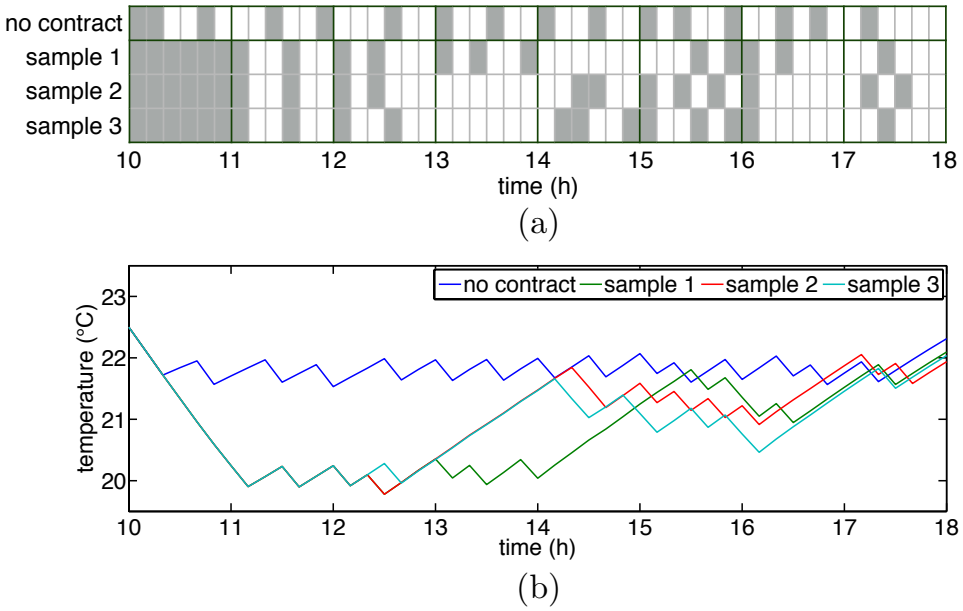

Figure 2.5: The simulation results without and with the contract (three samples are presented in the case with the contract): (a) control (gray: ON, white: OFF); and (b) indoor temperature.

motion assumption. Setting $S_i = \rho \bar{S}_i$, we averaged the percentage deviations over $\rho = [0, 1]$. The average deviations in the mean of the load-serving entity's and the customer's payoffs are 0.010% and 0.012%, respectively. Furthermore, the risk-limiting condition is not violated for $\rho > 0.14$; furthermore, for $\rho \leq 0.14$, the condition is never violated by more than 12% [137]. This preliminary test of the proposed contract framework with respect to errors in the Brownian motion model suggests that the framework is robust to load model errors, however more data is need for accurate approximations of the mean and variance values. Further experiments will be performed in the future to rigorously test the validity of Brownian motion model in the proposed contracts.

2.6.4 Real-time pricing in retail tariff

We now consider the case in which the energy price in the customer's electric tariff is chosen as

$$\mu(t) = e^{w_t} + \mu_0,$$

where $\mu_0 = \bar{\mu} - \mathbb{E}\left[\frac{1}{T}\int_0^T e^{w_t} dt\right] = 0.0692$. This real-time pricing scheme gives $\mathbb{E}\left[\frac{1}{T}\int_0^T \mu(t) dt\right] = \bar{\mu}$. In this case, the customer's optimal control needs to take into account the LMP e^{w_t} even when there is no contract. The optimal control, denoted as $\tilde{u}^{*i}$, can be obtained as the

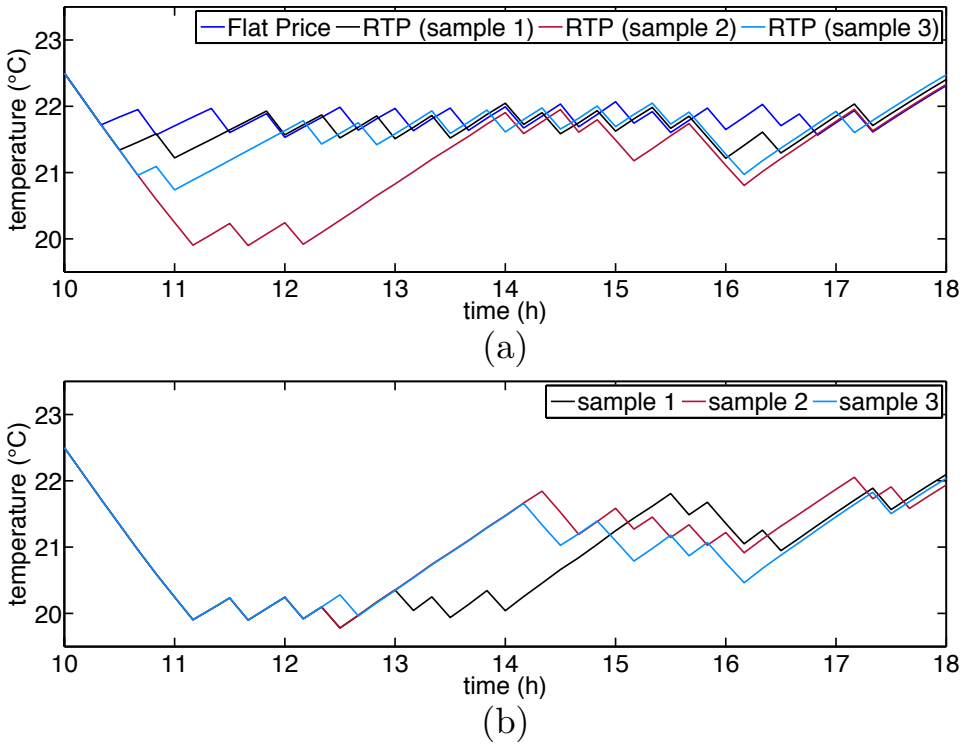

Figure 2.6: (a) Optimally controlled indoor temperatures without any contract in the case of flat price (blue) and real-time pricing (RTP) (black, red, cyan), and (b) optimally controlled indoor temperatures with the contract.

solution to

$$\begin{aligned} \max_{u^i \in \mathbb{U}^i} \quad & \mathbb{E}[\tilde{J}_i^A[u^i]] \\ \text{subject to} \quad & dw_t = r_0(\nu(t) - w_t)dt + \sigma_0(t)dW_t^0 \\ & dx_t^i = f_i(x_t^i, u_t^i)dt, \end{aligned}$$

where

$$\tilde{J}_i^A[u^i] := \int_0^T -(e^{w_t} + \mu_0)(l_i + u_t^i) + r_i(x_t^i, u_t^i)dt - \int_0^T (e^{w_t} + \mu_0)\tilde{\sigma}_i dW_t^i. \tag{2.36}$$

As shown in Figure 2.6 (a), the customer's optimal control does take into account the LMP dynamics even without any contract. In particular, the customer infrequently uses the air conditioner from 4:10 pm to 5:20 pm when the LMP is high (see Figure 2.2) as in the case under the contract. However, without a contract overall shifting of the consumption pattern from the flat price case is less significant than the case in which the contract is executed (Figure 2.6 (b)). This difference occurs because the risk-limiting dynamic contract takes into

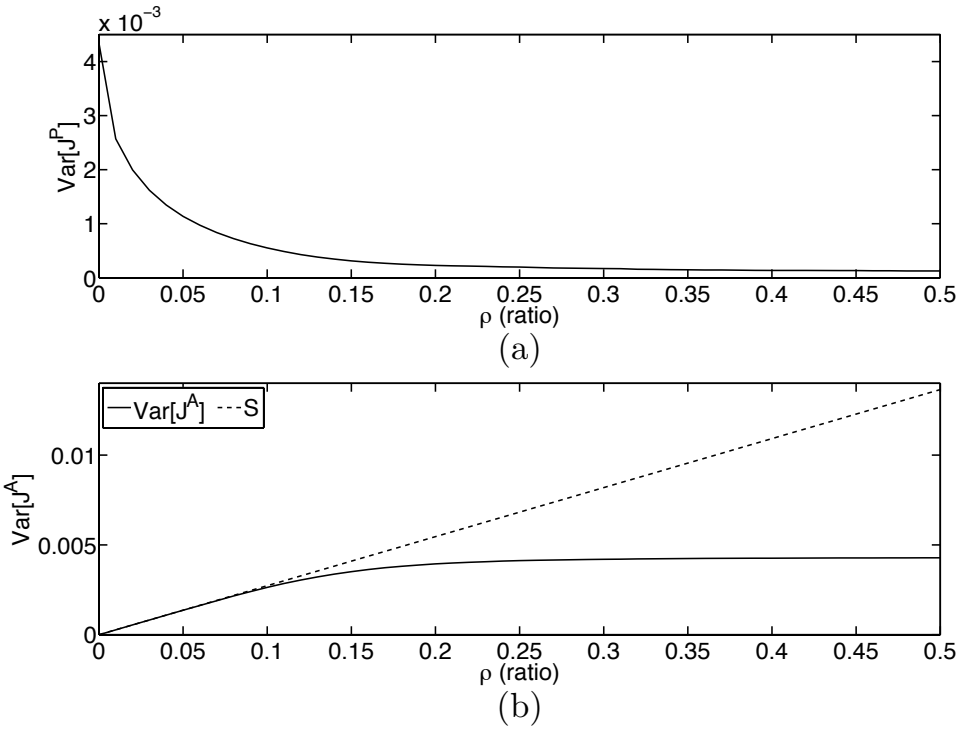

Figure 2.7: The simulation results with the contract in the case of real-time pricing: (a) variance of the load-serving entity's payoff, (b) variance of the customer's payoff and the risk share $S_i = \rho S_i$.

consideration the mean and variance of the customer's payoff and those of the load-serving entity's payoff, while the customer's optimal control only maximizes the mean value of the customer's payoff.

As before, we set $\bar{b} := \mathbb{E}[\tilde{J}_i^A[\tilde{u}^{*i}]] = -1.5050$ and $\bar{S} := \mathrm{Var}[\tilde{J}_i^A[\tilde{u}^{*i}]] = 0.0273$ as the customer's nominal expected payoff and risk, respectively. When $p_i \equiv l_i$, the load-serving entity's risk is given by $\mathrm{Var}[J_i^P[0, \tilde{u}^{*i}]] = 0.0107$. Note that this variance value can be further reduced by carefully choosing the procured power p_i other than the load forecast l_i. When the risk-limiting dynamic contract is executed, the load-serving entity's risk is efficiently managed even in the case of real-time pricing as shown in Figure 2.7. Furthermore, the customer's risk-limiting condition is satisfied. Under real-time pricing, the variance of the customer's payoff depends highly on the price volatility σ_0. Therefore, the customer may face high risk when the LMP is extremely volatile. The proposed contract framework complements the real-time pricing by explicitly limiting the customer's risk by a pre-specified level.

Chapter 3

Approximation Algorithms for Optimization of Combinatorial Dynamical Systems

3.1 Optimization of large-scale interdependent systems

The dynamics of critical infrastructures and their system elements—for instance, electric grid infrastructure and their electric load elements—are *interdependent*, meaning that the state of each infrastructure or its system elements influences and is influenced by the state of the others [111]. Such dynamic interdependencies can be classified as follows: (*i*) infrastructure–infrastructure interdependency; (*ii*) infrastructure–system interdependency; and (*iii*) system–system interdependency. All three classes of interdependencies must be addressed when making decisions that improve the performance metrics, such as efficiency, resilience and reliability, of infrastructures and their system elements. For an example of (*i*), consider the placement of power electronic actuators, such as high-voltage direct current links, on transmission networks. Such placement requires consideration of the interconnected swing dynamics of transmission grid infrastructures. As an example of (*ii*), it is important to consider the interdependency between the dynamics of grid frequency and those of (aggregate) loads when selecting the set of loads for frequency regulation service. Furthermore, the ON/OFF control of a large population of electric loads whose system dynamics are coupled with each other, e.g., supermarket refrigeration systems, must take into account their system-system interdependency (*iii*). These decision-making problems under dynamic interdependencies combine the combinatorial optimization problems of network actuator placement, load subset selection and ON/OFF control with the time evolution of continuous system states. Therefore, we seek decision-making techniques that unify combinatorial optimization and dynamical systems theory.

This chapter examines a fundamental problem that supports such combinatorial decision-

making involving dynamical systems. Specifically, we consider an optimization problem associated with a dynamical system whose state evolution depends on binary decision variables, which we call the *combinatorial dynamical system*. In our problem formulation, the binary decision variables do not change over time, unlike in the optimal control or predictive control of switched systems [18, 136, 131, 11]. Our focus is to develop scalable methods for optimizing the binary variables associated with a dynamical system when the number of the variables is too large to enumerate all possible system 'modes' and when the dimension of the system state is large. However, the optimization problem for combinatorial dynamical system presents a computational challenge because: (*i*) it is a 0–1 nonlinear program, which is generally NP-hard [100]; and (*ii*) it requires the solution of a system of ordinary differential equations (ODEs). To provide a computationally tractable solution method that can address large-scale problems, we propose scalable approximation algorithms with provable suboptimality bounds.

The key idea of the proposed methods is to linearize the objective function in the feasible space of binary decision variables. Our first contribution is to propose a linear approximation method for nonlinear optimization of combinatorial dynamical systems. The approximate 0–1 optimization can be efficiently solved because it is a linear 0–1 program and it does not require the solution of the dynamical system. The proposed approximation method allows us to employ polynomial-time exact or approximation algorithms including those for problems with l_0-norm constraints or linear inequality constraints. In particular, the proposed algorithms for an l_0-norm constrained problem are computationally more efficient than a greedy algorithm for the same problem because our algorithms are *one-shot*, i.e., do not require multiple iterations.

The proposed linear approximation approach requires the *derivative* of the objective function, but this is nontrivial to construct because the function's domain is a discrete space, in general. The second contribution of this work is to propose two different derivative concepts. The first concept uses a natural relaxation of the discrete space, whereas for the second concept a novel relaxation method in a function space using convex combinations of the vector fields and running payoffs is developed. We refer to the former construction as the *standard derivative* because it is the same as the derivative concept in continuous space, and the latter as the *nonstandard derivative*. We show the existence and the uniqueness of the nonstandard derivative, and provide an adjoint-based formula for it. The nonstandard derivative is well-defined even when the vector field and the payoff function are undefined on interpolated values of the binary decision variables. Because the two derivatives are different in general, we can solve two instances of the approximate problem (if the problem is well-defined on intermediate values in addition to a 0–1 lattice), one with the standard derivative and another with the nonstandard derivative and then choose the better solution. If the problem is defined only on a 0–1 lattice, we can utilize the nonstandard derivative.

The third contribution of this chapter is to characterize conditions under which the proposed algorithms have guaranteed suboptimality bounds. We show that the concavity of the original problem gives a sufficient condition for the suboptimality bound to hold if the approximation is performed using the standard derivative. On the other hand, the same

concavity condition does not suffice when the nonstandard derivative is employed in the approximation. To resolve this difficulty, we propose a reformulated problem and show that its concavity guarantees the suboptimality bound to hold. We validate the performance of the proposed approximation algorithms by solving ON/OFF control problems of commercial refrigeration systems, which consume approximately 7% of the total commercial energy consumption in the United States [129].

In operations research, 0–1 nonlinear optimization problems have been extensively studied over the past five decades, although the problems are not generally associated with dynamical systems. In particular, 0–1 polynomial programming, in which the objective function and the constraints are polynomials in the decision variables, has attracted great attention. Several exact methods that can transform a 0–1 polynomial program into a 0–1 linear program have been developed by introducing new variables that represent the cross terms in the polynomials (e.g., [133, 42]). Roof duality suggests approximation methods for 0–1 polynomial programs [48]. It constructs the best linear function that upperbounds the objective function (in the case of maximization) by solving a dual problem. Its size can be significantly bigger than that of the primal problem because it introduces $O(m^k)$ additional variables, where m and k denote the number of binary variables and the degree of polynomial, respectively. This approach is relevant to our proposed method in the sense that both methods seek a linear function that bounds the objective function. However, the proposed method explicitly constructs such a linear function without solving any dual problems. Furthermore, whereas all the aforementioned methods assume that the objective function is a polynomial in the decision variables, our method does not require a polynomial representation of the objective function. This is a considerable advantage because constructing a polynomial representation of a given function, $J : \{0,1\}^m \to \mathbb{R}$, generally requires 2^m calculations (e.g., via multi-linear extension [49]). Even when the polynomial representations of the vector field and the objective function in the decision variables, $\alpha \in \{0,1\}^m$, are given, a polynomial representation of the objective function in α is not readily available because the state of a dynamical system is not, in general, a polynomial in α with a finite degree. For more general 0–1 nonlinear programs, branch-and-bound methods (e.g., [78]) and penalty/smoothing methods (e.g., [91]) have been suggested. However, the branch-and-bound methods cannot, in general, find a solution in polynomial time. The penalty and smoothing methods do not provide any performance guarantee although they perform well in many data sets, whereas our proposed methods guarantee suboptimality bounds.

An important class of 0–1 nonlinear programs is the minimization or the maximization of a submodular set-function, which has the property of *diminishing returns*. Unconstrained submodular function minimization can be solved in polynomial time using a convex extension (e.g., [47]) or a combinatorial algorithm (e.g., [116, 58]). However, constrained submodular function minimization is NP-hard in general, and approximation algorithms with performance guarantees are available only in special cases (e.g., [43, 59, 61]). On the other hand, our proposed method can handle a large class of linear constraints with a guaranteed suboptimality bound. In the case of submodular function maximization, a greedy algorithm can obtain a provably near-optimal solution [95]. As mentioned, our proposed algorithm

for l_0-norm constrained problems has, in general, lower computational complexity than the greedy algorithm. We also show that the concavity conditions for our proposed suboptimality bounds to hold are not equivalent to submodularity nor does either imply the other.

3.2 Problem statement

Consider the following dynamical system in the continuous state space $\mathcal{X} \subseteq \mathbb{R}^n$:

$$\dot{x}(t) = f(x(t), \alpha), \quad x(0) = \mathbf{x} \in \mathcal{X}, \tag{3.1}$$

where the vector field depends on an m-dimensional binary vector variable $\alpha := \{\alpha_1, \cdots, \alpha_m\} \in \{0,1\}^m$ and $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$. We call (3.1) a *combinatorial dynamical system* with a binary vector variable α. We later view α as a decision variable that does not change over time in a given time interval $[0, T]$. Let $x^{\mathbf{a}} := (x_1^{\mathbf{a}}, \cdots, x_n^{\mathbf{a}})$ denote the solution of the ordinary differential equation (3.1) given $\mathbf{a} \in \mathbb{R}^m$. We consider the following assumptions on the vector field.

Assumption 1. *For each $\alpha \in \{0,1\}^m$, $f(\,\cdot\,, \alpha) : \mathbb{R}^n \to \mathbb{R}^n$ is twice differentiable, has a continuous second derivative and is globally Lipschitz continuous in $\mathcal{X}$.*

Assumption 2. *For any $\boldsymbol{x} \in \mathcal{X}$, $f(\boldsymbol{x}, \cdot) : \mathbb{R}^m \to \mathbb{R}^n$ is continuously differentiable in $[0,1]^m$.*

Under Assumption 1, the solution of (3.1) satisfies the following property (Proposition 5.6.5 in [106]): for any $\alpha \in \{0,1\}^m$,

$$\|x^\alpha\|_2 := \left(\int_0^T \|x^\alpha(t)\|^2 dt \right)^{\frac{1}{2}} < \infty.$$

In other words, $x^\alpha : [0, T] \to \mathbb{R}^n$ is such that $x^\alpha \in L^2([0, T]; \mathbb{R}^n)$. Furthermore, Assumption 1 guarantees that the system admits a unique solution, which is continuous in time, for each $\alpha \in \{0,1\}^m$.

3.2.1 Optimization of combinatorial dynamical systems

Our aim is to determine the binary vector $\alpha \in \{0,1\}^m$ that maximizes the payoff (or utility) function, $J : \mathbb{R}^m \to \mathbb{R}$, associated with the dynamical system (3.1). More specifically, we want to solve the following combinatorial optimization problem:

$$\max_{\alpha \in \{0,1\}^m} \quad J(\alpha) := \int_0^T r(x^\alpha(t), \alpha) dt + q(x^\alpha(T)) \tag{3.2a}$$

$$\text{subject to} \quad \mathbf{A}\alpha \leq \mathbf{b}, \tag{3.2b}$$

where x^α is the solution of (3.1) and $r : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ and $q : \mathbb{R}^n \to \mathbb{R}$ are running and terminal payoff functions, respectively. Here, $\mathbf{A}$ is an $l \times m$ matrix, $\mathbf{b}$ is an l-dimensional vector and the inequality constraint (3.32b) holds entry-wise.

This optimization problem, in general, presents a computational challenge because (*i*) it is NP-hard; and (*ii*) it requires the solution to the system of ODEs (3.1). Therefore, we seek a scalable approximation method that gives a suboptimal solution with a guaranteed suboptimality bound. The key idea of our proposed method is to take a first-order linear approximation of the objective function (3.2a) with respect to the binary decision variable α. This linear approximation should also take into account the dependency of the state on the binary decision variable. If the payoff function in (3.2a) is replaced with its linear approximation, which is linear in the decision variable, the approximate problem is a 0–1 linear optimization. Therefore, existing polynomial-time exact and approximation algorithms for 0–1 linear programs can be employed, as shown in Section 3.4. To obtain the linear approximations of the payoff function J, in the following section we formulate two different derivatives of J with respect to the discrete decision variable. Furthermore, we suggest a sufficient condition under which the approximate solution has a guaranteed suboptimality bound in Section 3.3.3.

3.3 Linear approximation for optimization of combinatorial dynamical systems

Suppose for a moment that the derivative of the objective function with respect to the binary decision variable is given, and that the derivative is well-defined in $\{0,1\}^m$, which is the feasible space of the decision variable. The derivative can be used to obtain the first-order linear approximation of the objective function, i.e., for $\alpha \in \{0,1\}^m$,

$$J(\alpha) \approx J(\bar{\alpha}) + DJ(\bar{\alpha})^\top(\alpha - \bar{\alpha}). \tag{3.3}$$

If the objective function in (3.32) is substituted with the right-hand side of (3.3), then we obtain the approximate problem:

$$\max_{\alpha \in \{0,1\}^m} \quad DJ(\bar{\alpha})^\top \alpha \tag{3.4a}$$

$$\text{subject to} \quad \mathbf{A}\alpha \leq \mathbf{b}. \tag{3.4b}$$

This approximate problem is a 0–1 linear program, which can be solved by several polynomial-time exact or approximation algorithms (see Section 3.4). We characterize a bound on the suboptimality of the approximate solution in Section 3.3.3.

We propose two different variation approaches for defining the derivatives in the discrete space $\{0,1\}^m$. The first uses the variation of the binary decision variable in a relaxed continuous space (Figure 3.1 (a)); the second uses the variation of the vector field of dynamical systems (Figure 3.1 (b)). The first and second concepts of the derivatives are

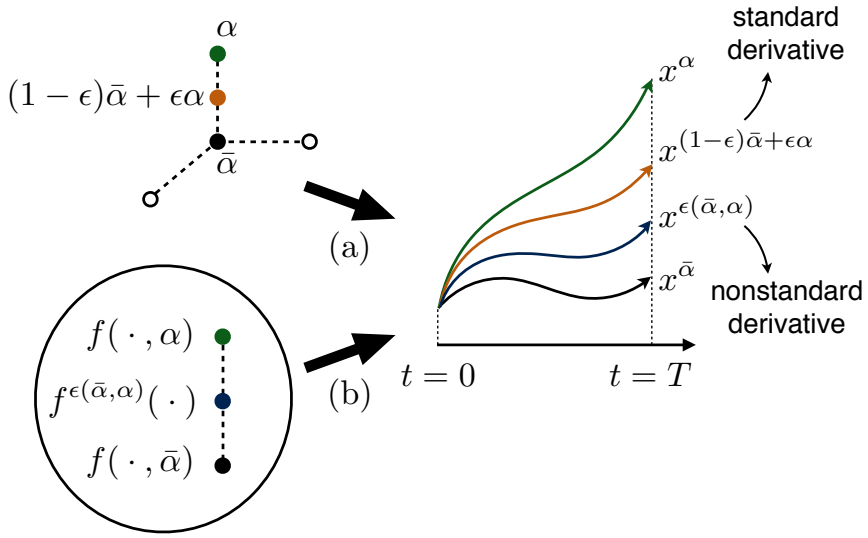

Figure 3.1: Two variation methods: (a) the variation $(1-\epsilon)\bar{\alpha}+\epsilon\alpha$ of the binary variable produces the trajectory $x^{(1-\epsilon)\bar{\alpha}+\epsilon\alpha}(t)$, $t\in[0,T]$; and (b) the variation $f^{\epsilon(\bar{\alpha},\alpha)}$ of the vector field, to be defined, generates another trajectory $x^{\epsilon(\bar{\alpha},\alpha)}(t)$, $t\in[0,T]$. These two new system trajectories are used to define the standard and nonstandard derivatives, respectively.

called the *standard* and *nonstandard* derivatives, respectively. It is advantageous to have two different derivative concepts: we solve the approximate problem (3.4) twice, one with the standard derivative $D^{\mathrm{S}}J$ and another with the nonstandard derivative $D^{\mathrm{NS}}J$ and then choose the better solution. The one of two approximate solutions that outperforms another is problem-dependent, in general. We also show that the nonstandard derivative requires fewer assumptions than the standard derivative.

Remark 1. *As we will see in the following subsection, the nonstandard derivative requires less restrictive assumptions than the standard derivative. One important distinction is that the nonstandard derivative can be well-defined even when the problem is only defined on the* 0–1 *lattice* $\{0,1\}^m$*, i.e.,* $f:\mathbb{R}^n\times\{0,1\}^m\to\mathbb{R}^n$ *and* $J:\{0,1\}^m\to\mathbb{R}$ *(i.e.,* $r:\mathbb{R}^n\times\{0,1\}^m\to\mathbb{R}$*). However, we can use the standard derivative only if the problem is well-defined on the relaxed space* $[0,1]^m$*.*

3.3.1 Standard and nonstandard derivatives

We first define the derivative of the payoff function, J, with respect to discrete variation of the decision variable by relaxing the discrete space $\{0,1\}^m$ into the continuous space $\mathbb{R}^m$. This definition of derivatives in discrete space is exactly the same as the standard definition of derivatives in continuous space. Therefore, it requires the differentiability of the vector field and the running payoff with respect to α.

Assumption 3. *The functions $r(\,\cdot\,,\alpha):\mathbb{R}^n\to\mathbb{R}$ and $q:\mathbb{R}^n\to\mathbb{R}$ are continuously differentiable for any $\alpha\in\{0,1\}^m$.*

Assumption 4. *For any $\boldsymbol{x}\in\mathcal{X}$, $r(\boldsymbol{x},\,\cdot\,):\mathbb{R}^m\to\mathbb{R}^n$ is continuously differentiable in $[0,1]^m$.*

More precisely, Assumptions 3 and 4 are needed for the standard derivative while the nonstandard derivative does not require Assumption 4. Throughout this chapter, we let $\mathbf{1}_i$ denote the m-dimensional vector whose ith entry is one and all other entries are zero. For notational convenience, we introduce a functional, $\mathcal{J}:L^2([0,T];\mathbb{R}^n)\times\mathbb{R}^m\to\mathbb{R}$, defined as

$$\mathcal{J}(z,\beta):=\int_0^T r(z(t),\beta)dt+q(z(T)). \tag{3.5}$$

Note that $J(\alpha)=\mathcal{J}(x^\alpha,\alpha)$, where x^α is defined as the solution to the ODE (3.1) with α.

Definition 1. *Suppose that Assumptions 1, 2, 3 and 4 hold. Given $\bar\alpha\in\{0,1\}^m$, the standard derivative, $D^SJ:\{0,1\}^m\to\mathbb{R}^m$, of the payoff function J in (3.2a) is defined as*

$$[D^SJ(\bar\alpha)]_i:=\lim_{\epsilon\to0}\frac{1}{\epsilon}\left[\mathcal{J}(x^{\bar\alpha+\epsilon\mathbf{1}_i},\bar\alpha+\epsilon\mathbf{1}_i)-\mathcal{J}(x^{\bar\alpha},\bar\alpha)\right]$$

for $i=1,\cdots,m$, where the functional $\mathcal{J}:L^2([0,T];\mathbb{R}^n)\times\mathbb{R}^m\to\mathbb{R}$ is defined in (3.5) and $x^{\bar\alpha}$ is the solution of (3.1) with $\bar\alpha$.

The standard derivative can be computed by direct and adjoint-based methods [72, 106]. We summarize the adjoint-based method in the following proposition.

Proposition 2. *Suppose that Assumptions 1, 2, 3 and 4 hold. The derivative in Definition 1 can be obtained as*

$$D^SJ(\bar\alpha)=\int_0^T\left(\frac{\partial f(x^{\bar\alpha}(t),\bar\alpha)}{\partial\boldsymbol{\alpha}}^\top\lambda^{\bar\alpha}(t)+\frac{\partial r(x^{\bar\alpha}(t),\bar\alpha)}{\partial\boldsymbol{\alpha}}^\top\right)dt,$$

where $x^{\bar\alpha}$ is the solution of (3.1) with $\bar\alpha$ and $\lambda^{\bar\alpha}$ solves the following adjoint system:

$$\begin{aligned}-\dot\lambda^{\bar\alpha}(t)&=\frac{\partial H(x^{\bar\alpha}(t),\lambda^{\bar\alpha}(t),\bar\alpha)}{\partial\boldsymbol{x}}^\top\\ \lambda^{\bar\alpha}(T)&=\frac{\partial q(x^{\bar\alpha}(T))}{\partial\boldsymbol{x}}^\top\end{aligned} \tag{3.6}$$

with the Hamiltonian $H:\mathbb{R}^n\times\mathbb{R}^n\times\{0,1\}^m\to\mathbb{R}$,

$$H(\boldsymbol{x},\boldsymbol{\lambda},\alpha):=\boldsymbol{\lambda}^\top f(\boldsymbol{x},\alpha)+r(\boldsymbol{x},\alpha).$$

We now define the derivative of the payoff function using variations in vector fields and running payoffs. The proposed nonstandard definition of derivatives does not require Assumptions 2 and 4, i.e., the differentiability of the vector field and the running payoff with respect to α. Furthermore, the nonstandard derivative is well-defined even when the vector field and the payoff function are not defined on the interpolated values of the binary decision variable, i.e., $f(\cdot, \alpha)$ and $r(\cdot, \alpha)$ are defined only at $\alpha \in \{0,1\}^m$. This is a practical advantage of the nonstandard derivative over the standard derivative. The proposed variation procedure is as follows.

(i) The 0–1 vector variable $\bar{\alpha}$ in the discrete space $\{0,1\}^m$ is mapped to $x^{\bar{\alpha}}$ in the continuous metric space $L^2([0,T];\mathbb{R}^n)$ via the original dynamical system (3.1);

(ii) In $L^2([0,T];\mathbb{R}^n)$, we construct a new state $x^{\epsilon(\bar{\alpha},\alpha)}$ as the solution to the *ϵ-variational system* associated with $(\bar{\alpha}, \alpha)$ for $\epsilon \in [0,1]$,

$$\dot{x}(t) = f^{\epsilon(\bar{\alpha},\alpha)}(x(t)), \quad x(0) = \mathbf{x} \in \mathcal{X}, \tag{3.7}$$

where the new vector field is obtained as the convex combination of the two vector fields with $\bar{\alpha}$ and α, i.e.,

$$f^{\epsilon(\bar{\alpha},\alpha)}(\,\cdot\,) := (1-\epsilon) f(\,\cdot\,, \bar{\alpha}) + \epsilon f(\,\cdot\,, \alpha).$$

Set the distance between α and its ϵ-variation $\epsilon(\bar{\alpha}, \alpha)$ as ϵ; and

(iii) The nonstandard derivative of J is defined in the following:

Definition 2. *Suppose that Assumptions 1 and 3 hold. Given $\bar{\alpha} \in \{0,1\}^m$, we define the (nonstandard) derivative, $D^{NS}J : \{0,1\}^m \to \mathbb{R}^m$ of J as*

$$[D^{NS}J(\bar{\alpha})]_i := \begin{cases} \lim\limits_{\epsilon \to 0^+} \frac{1}{\epsilon} \left[\mathcal{J}^{\epsilon(\bar{\alpha}, \bar{\alpha}+\mathbf{1}_i)}(x^{\epsilon(\bar{\alpha},\bar{\alpha}+\mathbf{1}_i)}) - \mathcal{J}(x^{\bar{\alpha}}, \bar{\alpha}) \right] & \text{if } \bar{\alpha}_i = 0 \\ \lim\limits_{\epsilon \to 0^+} \frac{1}{\epsilon} \left[\mathcal{J}(x^{\bar{\alpha}}, \bar{\alpha}) - \mathcal{J}^{\epsilon(\bar{\alpha}, \bar{\alpha}-\mathbf{1}_i)}(x^{\epsilon(\bar{\alpha},\bar{\alpha}-\mathbf{1}_i)}) \right] & \text{if } \bar{\alpha}_i = 1, \end{cases} \tag{3.8}$$

where $\mathcal{J} : L^2([0,T];\mathbb{R}^n) \times \mathbb{R}^m \to \mathbb{R}$ is given by (3.5) *and $\mathcal{J}^{\epsilon(\bar{\alpha},\alpha)} : L^2([0,T];\mathbb{R}^n) \to \mathbb{R}$ is given by*

$$\mathcal{J}^{\epsilon(\bar{\alpha},\alpha)}(\,\cdot\,) := (1-\epsilon)\mathcal{J}(\,\cdot\,, \bar{\alpha}) + \epsilon \mathcal{J}(\,\cdot\,, \alpha). \tag{3.9}$$

Here, $x^{\bar{\alpha}}$ is the solution of (3.1) *with $\bar{\alpha}$ and $x^{\epsilon(\bar{\alpha},\alpha)}$ is the solution of* (3.7)*.*

Note that we separately consider the cases with $\bar{\alpha}_i = 0$ and $\bar{\alpha}_i = 1$. This is because $\bar{\alpha} + \mathbf{1}_i$ is out of the feasible space of the binary decision variable when $\bar{\alpha}_i = 1$ and similarly for $\bar{\alpha} - \mathbf{1}_i$ when $\bar{\alpha}_i = 0$. Unlike a classical derivative with respect to continuous variable, the allowed directions for discrete variation depend on the base point $\bar{\alpha}$. Here, the new payoff functional uses the convex combination of the running payoff because

$$\mathcal{J}^{\epsilon(\bar{\alpha},\alpha)}(z) = \int_0^T (1-\epsilon) r(z, \bar{\alpha}) + \epsilon r(z, \alpha) dt + q(z(T)).$$

The ϵ-variational system is used as a continuation tool of the discrete variation from one decision variable to another. The properties of its solution are discussed in our previous work [142] and summarized in Appendix A.

This nonstandard definition of derivatives raises the two following questions: (*i*) *is the nonstandard derivative well-defined?*; and (*ii*) *is there a method to compute the nonstandard derivative?* We answer these two questions using the adjoint system (3.6) associated with the combinatorial optimization problem (3.32).

Theorem 3. *Suppose that Assumptions 1 and 3 hold. The nonstandard derivative* $D^{NS}J : \{0,1\}^m \to \mathbb{R}^m$ *satisfies*

$$[D^{NS}J(\bar{\alpha})]_i := \int_0^T \left(f(x^{\bar{\alpha}}(t), \bar{\alpha} + \mathbf{1}_i) - f(x^{\bar{\alpha}}(t), \bar{\alpha})\right)^\top \lambda^{\bar{\alpha}}(t) + r(x^{\bar{\alpha}}(t), \bar{\alpha} + \mathbf{1}_i) - r(x^{\bar{\alpha}}(t), \bar{\alpha}) dt$$

if $\bar{\alpha}_i = 0$ *and*

$$[D^{NS}J(\bar{\alpha})]_i := \int_0^T \left(f(x^{\bar{\alpha}}(t), \bar{\alpha}) - f(x^{\bar{\alpha}}(t), \bar{\alpha} - \mathbf{1}_i)\right)^\top \lambda^{\bar{\alpha}}(t) + r(x^{\bar{\alpha}}(t), \bar{\alpha}) - r(x^{\bar{\alpha}}(t), \bar{\alpha} - \mathbf{1}_i) dt$$

if $\bar{\alpha}_i = 1$. *Here* $x^{\bar{\alpha}}$ *and* $\lambda^{\bar{\alpha}}$ *are the solutions of* (3.1) *and* (3.6) *with* $\bar{\alpha}$, *respectively. The derivative uniquely exists and is bounded.*

The proof of Theorem 3 is contained in Appendix B. The detailed comparisons between the standard and nonstandard derivative concepts are provided in Appendix C.

3.3.2 Complexity of computing derivatives

To solve the 0–1 linear program (3.4), we first need to compute the standard derivative $D^{\mathrm{S}}J(\bar{\alpha})$ or the nonstandard derivative $D^{\mathrm{NS}}J(\bar{\alpha})$. Recall that the dimensions of the system state and the binary decision variable are n and m, respectively. Let N_T be the number of time points in the time interval $[0, T]$ used to integrate the dynamical system (3.1) and the adjoint system (3.6). Then the complexity of computing the trajectories of $x^{\bar{\alpha}}$ and $\lambda^{\bar{\alpha}}$ is $O(nN_T)$ if the first-order forward Euler scheme is employed (e.g., [1]). Note that the computation of the adjoint state trajectory $\lambda^{\bar{\alpha}}$ requires the state trajectory $x^{\bar{\alpha}}$ in $[0, T]$. Given $x^{\bar{\alpha}}$ and $\lambda^{\bar{\alpha}}$, calculating all the entries of either the standard derivative or the nonstandard derivative requires $O(mnN_T)$ if a first-order approximation scheme for the integral over time is used. Therefore, the total complexity of computing either the standard derivative or the nonstandard derivative is $O(mnN_T)$. Note that the complexity is linear in the dimension, m, of the decision variable α.

3.3.3 Suboptimality bounds

We now characterize the condition in which the solution to the approximate problem (3.4) has a guaranteed suboptimality bound. The suboptimality bound is obtained by showing

that the optimal value of the payoff function is bounded by an affine function of the solution to the approximate problem (3.4). This motivates the following concavity-like assumption:

Assumption 5. *Let $\bar{\alpha} \in \{0,1\}^m$ be the point at which the original problem (3.32) is linearized. The following equality holds*

$$DJ(\bar{\alpha})^\top(\alpha - \bar{\alpha}) \geq J(\alpha) - J(\bar{\alpha}) \quad \forall \alpha \in \{0,1\}^m. \tag{3.10}$$

Here, DJ represents $D^{\mathrm{S}}J$ if the standard derivative used in the approximate problem, and it represents $D^{\mathrm{NS}}J$ if the nonstandard derivative is adopted in (3.4).

For notational convenience, we let $\mathcal{A}$ denote the feasible set of the optimization problem (3.32), i.e.,

$$\mathcal{A} := \{\alpha \in \{0,1\}^m \mid \mathbf{A}\alpha \leq \mathbf{b}\}.$$

By subtracting $J(\bar{\alpha})$ from the payoff function, we normalize the payoff function such that, given $\bar{\alpha} \in \{0,1\}^m$ at which the original problem (3.32) is linearized,

$$J(\bar{\alpha}) = 0.$$

Note that $J(\alpha^{\mathrm{OPT}}) \geq 0$, where α^{OPT} is a solution of the original optimization problem (3.32), if $\bar{\alpha} \in \mathcal{A}$.

Theorem 4 (Performance Guarantee)**.** *Suppose that Assumption 5 holds. Let*

$$\begin{aligned} \alpha^{OPT} &\in \arg\max_{\alpha \in \mathcal{A}} \ J(\alpha), \\ \alpha^* &\in \arg\max_{\alpha \in \mathcal{A}} \ D^S J(\bar{\alpha})^\top \alpha, \\ \hat{\alpha}^* &\in \arg\max_{\alpha \in \mathcal{A}} \ D^{NS} J(\bar{\alpha})^\top \alpha. \end{aligned} \tag{3.11}$$

If $D^S J(\bar{\alpha})^\top(\alpha^* - \bar{\alpha}) \neq 0$ and $D^{NS} J(\bar{\alpha})^\top(\hat{\alpha}^* - \bar{\alpha}) \neq 0$, set

$$\begin{aligned} \rho &:= \frac{J(\alpha^*)}{D^S J(\bar{\alpha})^\top(\alpha^* - \bar{\alpha})}, \\ \hat{\rho} &:= \frac{J(\hat{\alpha}^*)}{D^{NS} J(\bar{\alpha})^\top(\hat{\alpha}^* - \bar{\alpha})}. \end{aligned} \tag{3.12}$$

and we have the following suboptimality bounds for the solutions of the approximate problems, i.e., α^* and $\hat{\alpha}^*$:

$$\begin{aligned} \rho J(\alpha^{OPT}) &\leq J(\alpha^*), \\ \hat{\rho} J(\alpha^{OPT}) &\leq J(\hat{\alpha}^*). \end{aligned} \tag{3.13}$$

Otherwise,

$$J(\alpha^{OPT}) = J(\bar{\alpha}) = 0,$$

i.e., $\bar{\alpha}$ is an optimal solution.

Proof. Due to Assumption 5, we have

$$J(\alpha^{\mathrm{OPT}}) = J(\alpha^{\mathrm{OPT}}) - J(\bar{\alpha}) \leq D^{\mathrm{S}}J(\bar{\alpha})^{\top}(\alpha^{\mathrm{OPT}} - \bar{\alpha}). \tag{3.14}$$

On the other hand, because $\alpha^* \in \arg\max_{\alpha\in\mathcal{A}} D^{\mathrm{S}}J(\bar{\alpha})^{\top}\alpha$ and $\alpha^{\mathrm{OPT}} \in \mathcal{A}$,

$$D^{\mathrm{S}}J(\bar{\alpha})^{\top}\alpha^{\mathrm{OPT}} \leq D^{\mathrm{S}}J(\bar{\alpha})^{\top}\alpha^*. \tag{3.15}$$

Suppose that $D^{\mathrm{S}}J(\bar{\alpha})^{\top}(\alpha^* - \bar{\alpha}) \neq 0$. Combining (3.14) and (3.15), we obtain the first inequality in (3.13); the second inequality can be derived using a similar argument. If $D^{\mathrm{S}}J(\bar{\alpha})^{\top}(\alpha^* - \bar{\alpha}) = 0$ or $D^{\mathrm{NS}}J(\bar{\alpha})^{\top}(\hat{\alpha}^* - \bar{\alpha}) = 0$, we have

$$J(\alpha^{\mathrm{OPT}}) \leq 0 = J(\bar{\alpha}).$$

Due to the optimality of α^{OPT}, the inequality must be binding. □

The coefficients ρ and $\hat{\rho}$ must be computed *a posteriori* because they require the solutions, α^* and $\hat{\alpha}^*$, respectively, of the approximate problems. They do not require the solution, α^{OPT}, of the original optimization problem. Note that ρ is, in general, different from $\hat{\rho}$. If $\bar{\alpha}$ is feasible, i.e., $\bar{\alpha} \in \mathcal{A}$, then we can improve the approximate solution by a simple post-processing that replaces it with $\bar{\alpha}$ if it is worse than $\bar{\alpha}$. The payoff functions evaluated at the post-processed approximate solutions are guaranteed to be greater or than equal to zero because $J(\bar{\alpha}) = 0$.

Corollary 3 (Post-Processing)**.** *Suppose that Assumption 5 holds and $\bar{\alpha} \in \mathcal{A}$. Let α^{OPT}, α^* and $\hat{\alpha}^*$ be given by* (3.11)*. Assume that $D^{S}J(\bar{\alpha})^{\top}(\alpha^* - \bar{\alpha}) \neq 0$ and $D^{NS}J(\bar{\alpha})^{\top}(\hat{\alpha}^* - \bar{\alpha}) \neq 0$. Define*

$$\begin{aligned} \alpha_* &= \arg\max\{J(\alpha^*), J(\bar{\alpha})\}, \\ \hat{\alpha}_* &= \arg\max\{J(\hat{\alpha}^*), J(\bar{\alpha})\}. \end{aligned} \tag{3.16}$$

and

$$\begin{aligned} \rho_* &:= \max\{\rho, 0\}, \\ \hat{\rho}_* &:= \max\{\hat{\rho}, 0\}, \end{aligned} \tag{3.17}$$

where ρ and $\hat{\rho}$ are given by (3.13)*. Then, we have the following suboptimality bounds for α_* and $\hat{\alpha}_*$:*

$$\begin{aligned} \rho_* J(\alpha^{OPT}) &\leq J(\alpha_*), \\ \hat{\rho}_* J(\alpha^{OPT}) &\leq J(\hat{\alpha}_*). \end{aligned} \tag{3.18}$$

The complexity of checking (3.10) in Assumption 5 for all $\alpha \in \{0,1\}^m$ increases exponentially as the dimension of the decision variable α increases. Therefore, we provide sufficient conditions, which are straightforward to check in some applications of interest, for Assumption 5. Note that the inequality condition (3.10) with $DJ = D^{\mathrm{S}}J$ is equivalent to the concavity of the payoff function at $\bar{\alpha}$ if the space in which α lies is $[0,1]^m$ instead of $\{0,1\}^m$. This observation is summarized in the following proposition.

Proposition 3. *Suppose that Assumption 1, 2, 3 and 4 hold. We also assume that the payoff function $J : \mathbb{R}^m \to \mathbb{R}$ in* (3.2a) *with x^α defined by* (3.1) *is concave in $[0,1]^m$, i.e.,*

$$J(\alpha) := \int_0^T r(x^\alpha(t), \alpha)dt + q(x^\alpha(T)),$$

with x^α satisfying

$$\dot{x}^\alpha(t) = f(x^\alpha(t), \alpha), \quad x^\alpha(0) = \mathbf{x} \in \mathcal{X},$$

is concave for all $\alpha \in [0,1]^m$. Then, the inequality condition (3.10) *with $DJ = D^S J$ holds for any $\bar{\alpha} \in \{0,1\}^m$.*

Recall that we view x^α as a function of α. Therefore, the concavity of J is affected by how the system state depends on α.

The inequality condition (3.10) with $DJ = D^{\mathrm{NS}} J$ is difficult to interpret due to the nonstandard derivative. We reformulate the dynamical system and the payoff function such that (i) the standard derivative of the reformulated payoff function corresponds to the nonstandard derivative of the original payoff function and (ii) the reformulated and original payoff functions have the same values at any $\alpha \in \{0,1\}^m$. Then, the concavity of the reformulated payoff function guarantees the inequality (3.10). To be more precise, we begin by considering the following *reformulated vector field and running payoff*:

$$\begin{aligned} \hat{f}(\,\cdot\,, \alpha) &:= f(\,\cdot\,, 0) + \sum_{i=1}^m \alpha_i (f(\,\cdot\,, \mathbf{1}_i) - f(\,\cdot\,, 0)), \\ \hat{r}(\,\cdot\,, \alpha) &:= r(\,\cdot\,, 0) + \sum_{i=1}^m \alpha_i (r(\,\cdot\,, \mathbf{1}_i) - r(\,\cdot\,, 0)). \end{aligned} \tag{3.19}$$

In general, $\hat{f}(\,\cdot\,, \alpha)$ (resp. $\hat{r}(\,\cdot\,, \alpha)$) and $f(\,\cdot\,, \alpha)$ (resp. $r(\,\cdot\,, \alpha)$) are different even when α is in the discrete space $\{0,1\}^m$. One can show that they are the same when $\alpha \in \{0,1\}^m$ if the following additivity assumption holds.

Assumption 6. *The functions $f(\boldsymbol{x}, \,\cdot\,)$ and $r(\boldsymbol{x}, \,\cdot\,)$ are additive in the entries of α for all $\boldsymbol{x} \in \mathcal{X}$, i.e.,*

$$\begin{aligned} f(\,\cdot\,, \alpha) &= f(\,\cdot\,, 0) + \sum_{i=1}^m (f(\,\cdot\,, \alpha_i \mathbf{1}_i) - f(\,\cdot\,, 0)), \\ r(\,\cdot\,, \alpha) &= r(\,\cdot\,, 0) + \sum_{i=1}^m (r(\,\cdot\,, \alpha_i \mathbf{1}_i) - r(\,\cdot\,, 0)). \end{aligned}$$

Note that these additivity conditions are less restrictive than the conditions that both of the functions are affine in α as shown in Example 7 in Appendix C.

This reformulation and Assumption 6 play an essential role in interpreting the nontrivial inequality condition (3.10) (with $DJ = D^{\mathrm{NS}} J$) as the concavity of a reformulated payoff

function, $\hat{J}$, defined in the next theorem. The standard derivative of the reformulated payoff function is equivalent to the nonstandard derivative of the original payoff function under Assumption 6, i.e.,

$$D^{\mathrm{S}}\hat{J} \equiv D^{\mathrm{NS}}J.$$

Furthermore, the two payoff functions have the same values when α is in the discrete space $\{0,1\}^m$, i.e.,

$$J|_{\{0,1\}^m} \equiv \hat{J}|_{\{0,1\}^m}.$$

Therefore, the inequality condition (3.10) with nonstandard derivative can be interpreted as the concavity of the reformulated payoff function.

Theorem 5. *Suppose that Assumptions 1, 3 and 6 hold. Define the reformulated payoff function $\hat{J} : \mathbb{R}^m \to \mathbb{R}$ as*

$$\hat{J}(\alpha) := \int_0^T \hat{r}(y^\alpha(t), \alpha)dt + q(y^\alpha(T)), \tag{3.20}$$

with y^α satisfying

$$\dot{y}^\alpha(t) = \hat{f}(y^\alpha(t), \alpha), \quad y^\alpha(0) = \mathbf{x} \in \mathcal{X},$$

where $\hat{f}$ and $\hat{r}$ are the reformulated vector field and running payoff, respectively, given in (3.19)*. If the reformulated payoff function $\hat{J}$ is concave in $[0,1]^m$, then the inequality condition* (3.10) *with $DJ = D^{NS}J$ holds for any $\bar{\alpha} \in \{0,1\}^m$.*

Proof. Fix $\boldsymbol{x} \in \mathbb{R}^n$ and $i \in \{1, \cdots, m\}$. If $\alpha_i = 0$, then

$$\alpha_i(f(\boldsymbol{x}, \mathbf{1}_i) - f(\boldsymbol{x}, 0)) = 0 = f(\boldsymbol{x}, \alpha_i\mathbf{1}_i) - f(\boldsymbol{x}, 0).$$

If $\alpha_i = 1$, then

$$\alpha_i(f(\boldsymbol{x}, \mathbf{1}_i) - f(\boldsymbol{x}, 0)) = f(\boldsymbol{x}, \alpha_i\mathbf{1}_i) - f(\boldsymbol{x}, 0).$$

On the other hand, due to Assumption 6, we have

$$f(\boldsymbol{x}, \alpha) = f(\boldsymbol{x}, 0) + \sum_{i=1}^{m}(f(\boldsymbol{x}, \alpha_i\mathbf{1}_i) - f(\boldsymbol{x}, 0)).$$

Therefore, $\hat{f}(\boldsymbol{x}, \alpha) = f(\boldsymbol{x}, \alpha)$ for any $\alpha \in \{0,1\}^m$. Using a similar argument, we can show that $\hat{r}(\boldsymbol{x}, \alpha) = r(\boldsymbol{x}, \alpha)$ for any $\alpha \in \{0,1\}^m$. These imply that

$$\hat{J}(\alpha) = J(\alpha) \quad \forall \alpha \in \{0,1\}^m. \tag{3.21}$$

Furthermore, using the adjoint-based formula in Proposition 2 for the standard derivative of the reformulated payoff function $\hat{J}$, we obtain

$$[D^{\mathrm{S}}\hat{J}(\alpha)]_i := \int_0^T \left(f(x^\alpha(t), \mathbf{1}_i) - f(x^\alpha(t), 0)\right)^\top \lambda^\alpha(t) + r(x^\alpha(t), \mathbf{1}_i) - r(x^\alpha(t), 0)dt. \tag{3.22}$$

On the other hand, under Assumption 6, the adjoint-based formula for the nonstandard derivative of the original payoff function J can be rewritten as

$$[D^{\mathrm{NS}}J(\alpha)]_i := \int_0^T \left(f(x^\alpha(t),(\alpha_i+1)\mathbf{1}_i) - f(x^\alpha(t),\alpha_i\mathbf{1}_i)\right)^\top \lambda^\alpha(t) + r(x^\alpha(t),(\alpha_i+1)\mathbf{1}_i) - r(x^\alpha(t),\alpha_i\mathbf{1}_i)dt$$

if $\alpha_i = 0$ and

$$[D^{\mathrm{NS}}J(\alpha)]_i := \int_0^T \left(f(x^\alpha(t),\alpha_i\mathbf{1}_i) - f(x^\alpha(t),(\alpha_i-1)\mathbf{1}_i)\right)^\top \lambda^\alpha(t) + r(x^\alpha(t),\alpha_i\mathbf{1}_i) - r(x^\alpha(t),(\alpha-1)\mathbf{1}_i)dt$$

if $\alpha_i = 1$. Plugging $\alpha_i = 0$ and $\alpha_i = 1$ into the two formulae, respectively and comparing them with (3.22), we conclude that

$$D^{\mathrm{S}}\hat{J}(\alpha) = D^{\mathrm{NS}}J(\alpha) \quad \forall \alpha \in \{0,1\}^m. \tag{3.23}$$

Suppose now that $\hat{J}$ is concave in $[0,1]^m$. Then, for any $\bar{\alpha}, \alpha \in \{0,1\}^m$,

$$D^{\mathrm{S}}\hat{J}(\bar{\alpha})^\top(\alpha - \bar{\alpha}) \geq \hat{J}(\alpha) - \hat{J}(\bar{\alpha}).$$

Combining this inequality with (3.21) and (3.23), we confirm that the inequality condition (3.10) with $DJ = D^{\mathrm{NS}}J$ holds for any $\bar{\alpha} \in \{0,1\}^m$. □

3.4 Algorithms

We now propose approximation algorithms for the optimization of combinatorial dynamical systems (3.32) using the linear approximation proposed in the previous section. Formulating the approximate problem (3.4) only requires the computation of the standard or nonstandard derivative with computational complexity $O(mnN_T)$ as suggested in Section 3.3.2, i.e., it is linear in the dimension of the decision variable. Because the approximate problem (3.4) is a 0–1 linear program, several polynomial time exact or approximation algorithms can be employed. Another advantage of the proposed approximation is that the approximate problem no longer depends on the dynamical system. Therefore, we do not need to compute the solution of the dynamical system once the derivative has been calculated.

We begin by proposing an efficient algorithm for the l_0-norm constrained problem. We then consider linear constraints (3.32b). Depending on the types of the linear constraints, several exact and approximation algorithms can be employed to solve the derivative-based approximate problem (3.4), which is a 0–1 linear program.

3.4.1 l_0-norm constraints

An important class of combinatorial optimization problems relevant to (3.32) is to maximize the payoff function, given that the l_0-norm of the decision variable is bounded. More specifically, instead of the original linear constraint (3.32b), we consider the constraint,

$$\underline{K} \leq \|\alpha\|_0 \leq \overline{K}, \tag{3.24}$$

where $\underline{K}$ and $\overline{K}$ are given constants. We consider the following first-order approximation of the combinatorial optimization problem:

$$\begin{aligned} \max_{\alpha \in \{0,1\}^m} \quad & DJ(\bar{\alpha})^\top \alpha \\ \text{subject to} \quad & \underline{K} \leq \|\alpha\|_0 \leq \overline{K}. \end{aligned} \tag{3.25}$$

A simple algorithm to solve (3.25) can be designed based on the ordering of the entries of $DJ(\bar{\alpha})$, where DJ is equal to either $D^{\mathrm{S}}J$ or $D^{\mathrm{NS}}J$. Let $\mathbf{d}(\cdot)$ denote the map from $\{1, \cdots, m\}$ to $\{1, \cdots, m\}$ such that

$$[DJ(\bar{\alpha})]_{\mathbf{d}(i)} \geq [DJ(\bar{\alpha})]_{\mathbf{d}(j)} \tag{3.26}$$

for any $i, j \in \{1, \cdots, m\}$ such that $i \leq j$. Such a map can be constructed using a sorting algorithm with $O(m \log m)$ complexity (e.g., [117]). Note that such a map may not be unique. We let $\alpha_{\mathbf{d}(i)} = 1$ for $i = 1, \cdots, \underline{K}$. We then assign 1 on $\alpha_{\mathbf{d}(i)}$ if $[DJ(\bar{\alpha})]_{\mathbf{d}(i)} > 0$ and $\underline{K} + 1 \leq i \leq \overline{K}$. Therefore, the total computational complexity to solve the approximate problem (3.25) requires $O(mnN_T) + O(m \log m)$. A more detailed algorithm to solve the l_0-norm constrained problem (3.25) is presented in Algorithm 1.

3.4.2 Totally unimodular matrix constraints

A totally unimodular (TU) matrix is defined as an integer matrix for which the determinant of every square non-singular sub-matrix is either $+1$ or -1. TU matrices play an important role in integer programs because they are invertible over the integers (e.g., Chapter III.1. of [94]). Suppose that $\mathbf{A}$ is TU and $\mathbf{b}$ is integral. Let

$$\bar{\mathbf{A}} := \begin{bmatrix} \mathbf{A} \\ I_{m \times m} \end{bmatrix} \quad \text{and} \quad \bar{\mathbf{b}} := \begin{bmatrix} \mathbf{b} \\ \mathbf{1} \end{bmatrix},$$

where $\mathbf{1}$ is the m-dimensional vector whose entries are all 1's. The new matrix $\bar{\mathbf{A}}$ is also TU. The approximate optimization problem is equivalent to the following integer linear program:

$$\begin{aligned} \max_{\alpha \in \mathbb{Z}^m} \quad & DJ(\bar{\alpha})^\top \alpha \\ \text{subject to} \quad & \bar{\mathbf{A}}\alpha \leq \bar{\mathbf{b}}. \end{aligned}$$

Algorithm 1: Algorithm for the l_0-norm constrained problem (3.25)

Initialization:
Given $\bar{\alpha}, \underline{K}, \overline{K}$;
$\alpha \leftarrow 0$;

Construction of d:
Compute $DJ(\bar{\alpha})$;
Sort the entries of $DJ(\bar{\alpha})$ in descending order;
Construct $\mathbf{d} : \{1, \cdots, m\} \to \{1, \cdots, m\}$ satisfying (3.26);

Solution of (3.25)**:**
for $i = 1 : \underline{K}$ **do**
  $\alpha_{\mathbf{d}(i)} \leftarrow 1$;
end
$i \leftarrow \underline{K} + 1$;
while $[DJ(\bar{\alpha})]_{\mathbf{d}(i)} > 0$ *and* $i \leq \overline{K}$ **do**
  $\alpha_{\mathbf{d}(i)} \leftarrow 1$;
  $i \leftarrow i + 1$;
end

Because $\bar{\mathbf{A}}$ is TU and $\mathbf{b}$ is integral, the solution of this problem can be obtained as the solution to the linear program, whose feasible region is relaxed to $\mathbb{R}^m$, of the form

$$\begin{aligned} \max_{\alpha \in \mathbb{R}^m} \quad & DJ(\bar{\alpha})^\top \alpha \\ \text{subject to} \quad & \bar{\mathbf{A}}\alpha \leq \bar{\mathbf{b}}. \end{aligned} \tag{3.27}$$

The proof of the exactness of this continuous relaxation can be found in [94]. The linear program (3.27) can be solved by a simplex algorithm (e.g., [29]), interior-point methods (e.g., [96]), and several others. Note that this approach does not require any rounding or thresholding of the solution because the solution of the relaxed problem lies in the original feasible space $\{0, 1\}^m$.

3.4.3 General linear constraints

Suppose that $l = 1$, i.e., $\mathbf{A} \in \mathbb{R}^{1 \times l}$ is a vector and $\mathbf{b} \in \mathbb{R}$ is a scalar and that all the entries of $\mathbf{A}$ and $[DJ(\bar{\alpha})]_i$ are non-negative.[1] In this case, the approximate problem (3.4) is a 0–1 *knapsack problem*, which has been extensively studied in the past six decades. A popular solution method is the greedy algorithm based on the linear programming (LP) relaxation proposed by Dantzig [30], which replaces the feasible region $\{0, 1\}^m$ with $[0, 1]^m$. A simple post-processing on the solution of the LP gives a 0.5-approximate solution of the knapsack problem. Such

[1] The latter non-negativity assumption can easily be relaxed by fixing $\alpha_j = 0$ for j such that $[DJ(\bar{\alpha})]_j < 0$.

an approximate solution can be computed with complexity of $O(m) + O(m \log m)$ using a greedy algorithm (e.g., pp. 28–29 of [85]). Other approximation algorithms have been proposed including a polynomial time approximation [57]. 0-1 knapsack problems with a large number of variables can be exactly solved by branch-and-bound algorithms (e.g., [54], [6]). Another classic exact method for knapsack problems is via dynamic programming (e.g., [127]). Several other algorithms and computational experiments can be found in the monograph [85] and the references therein. If $l > 1$ and $\mathbf{A}_{i,j} \geq 0$ (and $[DJ(\bar{\alpha})]_j \geq 0$) for $i = 1, \cdots, l$ and $j = 1, \cdots, m$, then the approximate problem is called the *multidimensional 0–1 knapsack problem*. Several exact and approximation algorithms have been developed and can be found in the review [40], as well as among the references therein. If no assumptions are imposed, i.e., the approximate problem (3.4) with general linear inequality constraints is considered, then successive linear or semidefinite relaxation methods for a 0–1 polytope can provide approximation algorithms with suboptimality bounds [82, 119, 80].

Remark 2. *Note that our proposed* 0–1 *linear program approximation does not have any dynamical system constraints, while the original problem* (3.32) *does. This is advantageous because the approximate problem does not require any computational effort to solve the dynamical system once the standard or nonstandard derivative is calculated. In other words, the complexity of any algorithm applied to the approximate problem is independent of the time horizon* $[0, T]$ *of the dynamical system or the number,* N_T*, of discretization points in* $[0, T]$ *used to approximate* $DJ(\bar{\alpha})$*.*

3.5 Comparison with submodularity

Submodularity of a set function has attracted significant attention due to its usefulness in combinatorial optimization. As summarized in Section 3.1, several algorithms have been proposed for minimizing or maximizing a submodular function. Its application includes sensor placement [74, 75], actuator placement (based on the controllability Grammian) [123], network inference [44], dynamic state estimation [84], and leader selection under link noise [25].

Consider a set Ω with m elements, $\Omega := \{1, \cdots, m\}$. We define a set indicator function $\mathbb{I} : 2^{\Omega} \to \{0, 1\}^m$ as

$$[\mathbb{I}(X)]_i := \begin{cases} 0 & \text{if } i \notin X \\ 1 & \text{if } i \in X. \end{cases}$$

The set function $J(\mathbb{I}(\cdot)) : 2^{\Omega} \to \mathbb{R}$ is said to be *submodular* provided that for any $X \subset Y \subseteq \Omega$ and any $s \in \Omega \setminus Y$

$$J(\mathbb{I}(X \cup \{s\})) - J(\mathbb{I}(X)) \leq J(\mathbb{I}(Y \cup \{s\})) - J(\mathbb{I}(Y)).$$

If, in addition, it is monotone, i.e., for any $X \subset Y \subseteq \Omega$

$$J(\mathbb{I}(X)) \leq J(\mathbb{I}(Y)),$$

then the problem of maximizing (3.2a) with l_0-norm constraint (3.24) admits a $(1-1/e)$ approximation algorithm [95]. Minimizing (3.2a) with the l_0-norm constraint is NP-hard, while several polynomial time algorithms can solve unconstrained submodular minimization problems as mentioned in Section 3.1.

Recall that the concavity of J (resp. $\hat{J}$) guarantees the suboptimality bound to hold if the standard (resp. nonstandard) derivative is employed (see Proposition 3 and Theorem 5). We investigate *sufficient conditions* for the concavity of J and $\hat{J}$ and the submodularity of $J(\mathbb{I}(\cdot))$. It turns out that the concavity of J or $\hat{J}$ does not imply the submodularity of $J(\mathbb{I}(\cdot))$; furthermore, the submodularity of $J(\mathbb{I}(\cdot))$ does not imply the concavity of J or $\hat{J}$.

3.5.1 Conditions for concavity and submodularity

We begin by providing examples to show that concavity and submodularity do not imply one another.

Example 3 (Concavity does not imply submodularity)**.** *Consider the following vector field and running payoff:*

$$\begin{aligned} f(\boldsymbol{x},\alpha) &= (\boldsymbol{x}_1+\alpha_1+2, \boldsymbol{x}_2+\alpha_2),\\ r(\boldsymbol{x},\alpha) &= -(\boldsymbol{x}_1-\boldsymbol{x}_2)^2. \end{aligned}$$

Then, we have $r(x^\alpha(t),\alpha) = -\Big(\int_0^t e^{t-\tau}(\alpha_1-\alpha_2+2)d\tau\Big)^2$. *The terminal payoff is set to* $q\equiv 0$. *Since the following equalities hold*

$$\begin{aligned} J(\mathbb{I}(\{2\})) - J(\mathbb{I}(\emptyset)) &= 3\left(\int_0^t e^{t-\tau}d\tau\right)^2,\\ J(\mathbb{I}(\{1,2\})) - J(\mathbb{I}(\{1\})) &= 5\left(\int_0^t 2e^{t-\tau}d\tau\right)^2, \end{aligned}$$

$J(\mathbb{I}(\cdot))$ is not submodular. On the other hand, $J=\hat{J}$ is concave in $\alpha\in[0,1]^2$.

Example 4 (Submodularity does not imply concavity)**.** *Suppose that all the assumptions in previous example hold except that the running payoff is given by*

$$r(\boldsymbol{x},\alpha) = (\boldsymbol{x}_1-\boldsymbol{x}_2)^2.$$

In this case, $J=\hat{J}$ is not concave in α, while $J(\mathbb{I}(\cdot))$ is submodular.

For comparison, we consider the case in which the vector field is linear in state and decision variable and the payoff function has a particular structure. In this case, the solution of the dynamical system is affine in the decision variable.

Proposition 4. *Suppose that r is separable as*

$$r(\boldsymbol{x}, \alpha) = r_1(\boldsymbol{x}) + r_2(\alpha).$$

Consider the vector field of the form

$$f(\boldsymbol{x}, \alpha) = A\boldsymbol{x} + B\alpha,$$

where A is an $n \times n$ matrix and B is an $n \times m$ matrix. Then,

1. *J is concave if r_1, r_2 and q are concave;*
2. *$\hat{J}$ is concave if r_1 and q are concave;*
3. *$J(\mathbb{I}(\cdot))$ is submodular if: $r_2(\mathbb{I}(\cdot))$ is submodular; r_1 and q are separable such that $r_1(\boldsymbol{x}) = \sum_{i=1}^m r_{1,i}(\boldsymbol{x}_i)$ and $q(\boldsymbol{x}) = \sum_{i=1}^m q_i(\boldsymbol{x}_i)$ with $r_{1,i}$ and q_i concave for all i; and given i, for any $X \subset Y \subseteq \Omega$, either*

$$x_i^{\mathbb{I}(X)}(t) \leq x_i^{\mathbb{I}(Y)}(t) \quad \forall t \in [0, T],$$

or

$$x_i^{\mathbb{I}(X)}(t) \geq x_i^{\mathbb{I}(Y)}(t) \quad \forall t \in [0, T].$$

Proof. The ODE (3.1) admits a unique solution,

$$x^\alpha(t) = e^{At}\mathbf{x} + \int_0^t e^{A(t-\tau)}B\alpha d\tau.$$

Therefore, $x^\alpha(t)$ is affine in α for all $t \in [0, T]$. This implies that $r_1(x^\alpha(t))$ and $q(x^\alpha(t))$ are concave in α. Furthermore, because r_2 is concave in α, so is J.

In this linear system case, the reformulated vector field $\hat{f}$ in (3.19) is equivalent to f and therefore the reformulated ODE admits the same solution, i.e., $y^\alpha \equiv x^\alpha$ for all $\alpha \in [0, 1]$. The reformulated running payoff in (3.19) is given by

$$\begin{aligned}\hat{r}(y^\alpha(t), \alpha) &= r(y^\alpha(t), 0) + \sum_{i=1}^m \alpha_i (r(y^\alpha(t), \mathbf{1}_i) - r(y^\alpha(t), 0)) \\ &= r(y^\alpha(t), 0) + \sum_{i=1}^m \alpha_i (r_2(\mathbf{1}_i) - r_2(0)).\end{aligned}$$

Therefore, it is concave in α and so is $\hat{J}$. Note that it does not require the concavity of r_2.

Given $i \in \{1, \cdots, n\}$, we notice that for any $X \subset Y \subseteq \Omega$ and for any $s \in \Omega \setminus Y$, either

$$x_i^{\mathbb{I}(X \cup \{s\})} - x_i^{\mathbb{I}(X)} = x_i^{\mathbb{I}(Y \cup \{s\})} - x_i^{\mathbb{I}(Y)} \geq 0, \;\; x_i^{\mathbb{I}(X)} \leq x_i^{\mathbb{I}(Y)}$$

or

$$x_i^{\mathbb{I}(X\cup\{s\})} - x_i^{\mathbb{I}(X)} = x_i^{\mathbb{I}(Y\cup\{s\})} - x_i^{\mathbb{I}(Y)} \leq 0, \;\; x_i^{\mathbb{I}(X)} \geq x_i^{\mathbb{I}(Y)}$$

For both cases, the concavity of $r_{1,i}$ implies that

$$r_{1,i}(x_i^{\mathbb{I}(X\cup\{s\})}) - r_{1,i}(x_i^{\mathbb{I}(X)}) \geq r_{1,i}(x_i^{\mathbb{I}(Y\cup\{s\})}) - r_{1,i}(x_i^{\mathbb{I}(Y)}).$$

A similar inequality holds for $q_{1,i}$. Since r_2 is submodular, we also have for any $X \subset Y \subseteq \Omega$ and for any $s \in \Omega \setminus Y$

$$r_2(\mathbb{I}(X \cup \{s\})) - r_2(\mathbb{I}(X)) \geq r_2(\mathbb{I}(Y \cup \{s\})) - r_2(\mathbb{I}(Y)).$$

Therefore, we obtain that for any $X \subset Y \subseteq \Omega$ and for any $s \in \Omega \setminus Y$

$$\begin{aligned}
&J(\mathbb{I}(X \cup \{s\})) - J(\mathbb{I}(X)) \\
&= \int_0^T r(x^{\mathbb{I}(X\cup\{s\})}, \mathbb{I}(X \cup \{s\})) - r(x^{\mathbb{I}(X)}, \mathbb{I}(X))dt + q(x^{\mathbb{I}(X\cup\{s\})}) - q(x^{\mathbb{I}(X)}) \\
&\geq \int_0^T r(x^{\mathbb{I}(Y\cup\{s\})}, \mathbb{I}(Y \cup \{s\})) - r(x^{\mathbb{I}(Y)}, \mathbb{I}(Y))dt + q(x^{\mathbb{I}(Y\cup\{s\})}) - q(x^{\mathbb{I}(Y)}) \\
&= J(\mathbb{I}(Y \cup \{s\})) - J(\mathbb{I}(Y)),
\end{aligned}$$

which implies that $J(\mathbb{I}(\cdot))$ is submodular. □

Note that these are not necessary but sufficient conditions. We observe that the concavity of $\hat{J}$ does not require the concavity of r_2. However, a sufficient condition proposed in Theorem 5 for Assumption 5, which is essential for the suboptimality bound, requires the additivity of r_2 (Assumption 6) in addition to the concavity of $\hat{J}$. In Section 4.4, the payoff function of the proposed direct load control problem satisfies all the conditions and therefore it is both concave and submodular. In nonlinear system cases, we admit that it is nontrivial to check the concavity of J or $\hat{J}$ and the submodularity of $J(\mathbb{I}(\cdot))$ unless an analytical solution of the system is available. Further studies on characterizing the conditions for the concavity and the submodularity in the case of nonlinear systems will be performed in the future.

3.5.2 Computational complexity

We now compare our proposed algorithm (Algorithm 1) for the l_0-norm constrained problem with the greedy algorithm for maximizing a submodular function with the same constraint, assuming that the payoff function is submodular and satisfies Assumption 5. Our algorithm is *one-shot* in the sense that, after computing the derivative and ordering its entries only once, the solution is obtained. On the other hand, the greedy algorithm chooses a locally optimal solution at each stage. In other words, this iterative greedy choice approach requires one to find an entry that maximizes the increment in the current payoff at every stage. Its complexity is $O(m^2 n N_T)$, quadratic in m. Therefore, our proposed algorithm is computationally more efficient as the number m of binary decision variables grows because it requires $O(mnN_T) + O(m \log m)$ calculations.

3.6 Application to real-time regulation of supermarket refrigeration systems

Commercial refrigeration systems account for approximately 7% of the total commercial energy consumption in the United States [129]. With this high power consumption capacity of commercial refrigeration systems, such as supermarket refrigeration systems, there have been growing efforts toward engaging their aggregations in the operation of the electric grid [118, 101]. We propose a new optimization-based control approach for aggregate supermarket refrigeration systems to provide real-time services to the grid. We formulate the real-time regulation problem as a centralized *one-step look-ahead optimization* problem explicitly taking into account the interdependency of refrigerator display case temperatures. The formulation generates ON/OFF control signals for all the refrigerator evaporators in the aggregation to achieve the following two objectives: (i) to minimize the deviation of each display case temperature from a desirable range, and (ii) to satisfy the real-time request of a system operator. The proposed ON/OFF control approach has a fixed duty cycle and is thus easy to implement compared with other methods that control duty cycles [115] or temperature set points and cooling capacities [36, 55, 118].

The optimization program formulated in Section 3.6.2 is combinatorial because the display case temperature dynamics are interdependent and the total power consumption is limited by the real-time regulation signal. This problem presents computational challenges because it is NP-hard in general and requires repeated simulation of the temperature dynamic model. The most relevant formulations to our problem are the model predictive control (MPC)-based methods in [79] and [101]. Because the control variables are binary (ON or OFF), these MPC problems are combinatorial and NP-hard. [79] uses a hybrid MPC scheme that does not scale well with the number of control variables. [101] proposes a heuristic algorithm that gives an approximate solution, which shows good performance in the simulation results. However, the approximate solution does not have a guaranteed suboptimality bound. To overcome the computational challenges, we utilize the proposed algorithm.

3.6.1 Supermarket refrigeration systems

We consider an aggregation of supermarket refrigerators in which multiple display cases are interconnected with one another. For example, Figure 3.2 shows two refrigerators, each of which has 10 display cases. The temperature of each display case is controlled by an evaporator unit, where the refrigerant evaporates absorbing heat from the display case. Let evaporator i be in charge of display case i for $i = 1, \cdots, n$, where n is the number of display cases in all the refrigerators. Expansion valve i controls the refrigerant injection into evaporator i and decreases the pressure of the refrigerant if it is open, as shown in Figure 3.3. We say that evaporator i is in the ON state when expansion valve i is open; otherwise, we say that it is in the OFF state.

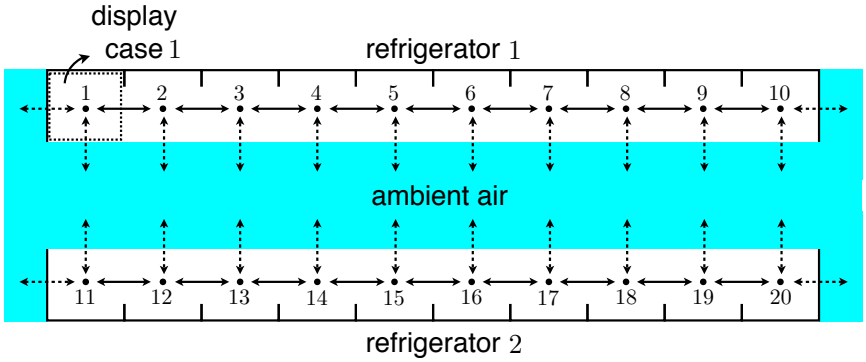

Figure 3.2: Two supermarket refrigerators, each of which has 10 evaporators. Evaporator i controls the temperature of display case i.

Evaporators and display case temperatures

Let $x_i(t)$, $\theta_i(t)$ and $u_i(t)$ be the temperature of display case i, the ambient temperature near display case i and the ON/OFF control signal of evaporator i at time t, respectively. More precisely, we set $u_i(t) = 0$ if evaporator i is OFF at time t and $u_i(t) = 1$ otherwise. The temperature of display case i can be modeled as the following first-order differential equation:

$$\dot{x}_i = \begin{cases} -a_{ii}(x_i - \theta_i) - \sum_{j=1}^{n} a_{ij}(x_i - x_j) - b_i & \text{if } u_i = 1 \\ -a_{ii}(x_i - \theta_i) - \sum_{j=1}^{n} a_{ij}(x_i - x_j) & \text{if } u_i = 0. \end{cases} \tag{3.28}$$

The thermal parameters are chosen as $a_{ii} = \frac{\bar{U}_{i-amb}}{M_i C_i}, a_{ij} = \frac{\bar{U}_{ij}}{M_i C_i}, b_i = \frac{\bar{U}_{i-ref}}{M_i C_i}\bar{b}$, where M_i and C_i denote the mass and heat capacity of the goods in display case i, respectively, and $\bar{U}_{i-amb}$, $\bar{U}_{ij}$ and $\bar{U}_{i-ref}$ are the coefficients of the heat transfer between display case i and ambient air, the heat transfer between display cases i and j and the heat transfer between display case i and its refrigerant, respectively. Higher-order models can be employed for detailed description of display case temperature dynamics [115, 55]. Let A be an $n \times n$ connectivity matrix, whose (i, j)–th entry is given by

$$[A]_{ij} := \begin{cases} -\sum_{k=1}^{n} a_{ik} & \text{if } i = j \\ a_{ij} & \text{otherwise,} \end{cases}$$

B be an $n \times n$ diagonal matrix, whose ith diagonal element is given by $-b_i$, and Θ be an n dimensional vector, whose ith element is given by $a_{ii}\theta_i$. Then, the display case temperature dynamics of multiple refrigerators can be compactly written as

$$\dot{x} = Ax + Bu + \Theta,$$

where $x := (x_1, \cdots, x_n)$ and $u := (u_1, \cdots, u_n)$.

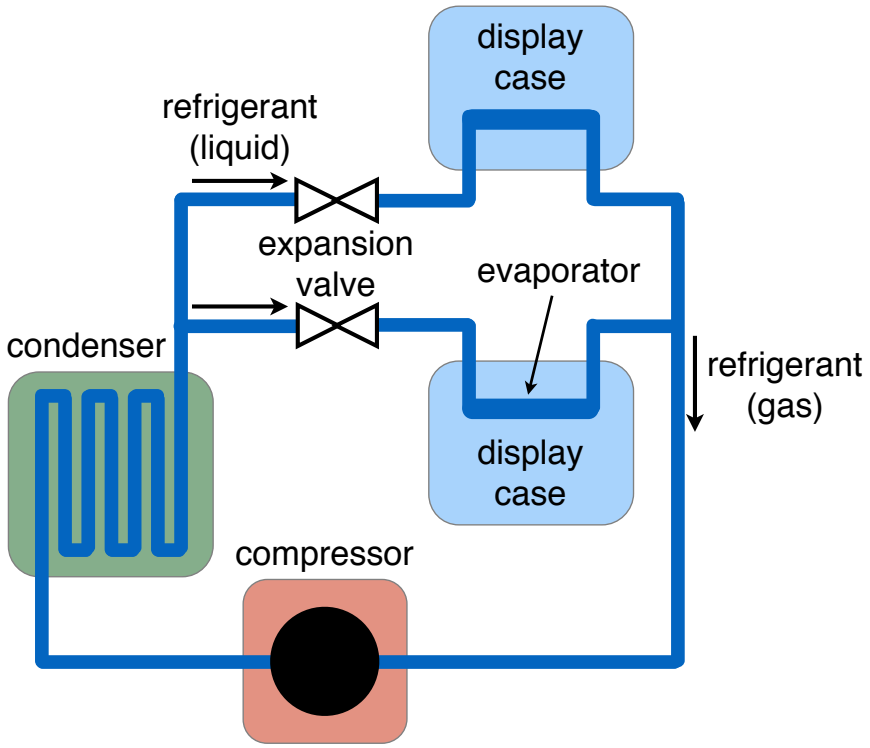

Figure 3.3: Schematic diagram of the supermarket refrigerators.

Compressors and power consumptions

As shown in Figure 3.3, the evaporated refrigerant with low pressure from the outlet of the evaporator is compressed by the electric motors in the compressor bank. Each refrigerator could have multiple compressors and each compressor is switched ON or OFF. For example, all the compressors are turned ON when maximal compression is needed. We assume that the number of compressors that are ON is approximately proportional to the total mass flow of incoming refrigerants. Conventionally, the compressor bank is controlled by a PI controller to maintain the suction pressure within a bandwidth [115]. Our assumption is a good approximation of this conventional PI control when the pressure bandwidth is small. The compressors of a refrigerator account for a major portion of its power consumption. Therefore, modeling the power consumption of all the refrigerators, as in the following, is reasonable: $\sum_{j=1}^{n_c} \bar{c}_j \bar{u}_j$, where n_c is the number of compressors in all the refrigerators, and $\bar{c}_j$ and $\bar{u}_j$ denote the power consumption and the ON/OFF control signal of compressor j, respectively, for $j = 1, \cdots, n_c$. Under our assumption that the number of ON compressors is proportional to the number of ON evaporators in the same refrigerator, we can approximate the power consumption of all the refrigerators as

$$P(u) := \sum_{i=1}^{n} c_i u_i, \tag{3.29}$$

where the coefficient c_i can be interpreted as the contribution of evaporator i on the aggregate power consumption. In Section 3.6.2, we propose an optimization-based algorithm to regulate the aggregate power consumption $P(u)$ in real time.

The compressed refrigerant flows to the condenser and is liquified by generating heat, as shown in Figure 4.1. The liquified refrigerant flows to the expansion valve, and as a result, the refrigeration circuit is closed.

Conventional control

Conventionally the following set-point based control is used for evaporators:

$$u_i(t) := \begin{cases} 0 & \text{if } x_i(t) > \overline{\theta}_i \\ 1 & \text{if } x_i(t) < \underline{\theta}_i \\ u_i(t^-) & \text{otherwise,} \end{cases} \tag{3.30}$$

where $[\underline{\theta}_i, \overline{\theta}_i]$ is the desirable temperature range for display case i. Note that this control approach is decentralized: the control decision for evaporator i depends only on its local temperature. Intuitively, this decentralized control is suboptimal because it does not actively consider the heat transfer between neighboring display cases. This inefficiency of the conventional control approach has motivated us to develop a new optimization-based control method that takes into account the interdependency of display case temperature dynamics.

Another disadvantage of conventional control is that it is inappropriate for the real-time regulation of the total power consumption of aggregate refrigeration systems. To provide services to the electric grid, such as peak demand reduction and spinning reserve, an aggregator may want to limit the total power consumption $P(u)$ in (3.29) with a real-time regulation signal provided by a system operator. For real-time regulations, we propose a one-step look-ahead optimization program that explicitly constrains $P(u)$ with a regulation signal.

As mentioned, each compressor is conventionally controlled by a PI controller. Advanced approaches, such as MPC, have been recently proposed to reduce the number of compressor switchings [115]. In this study, we do not design a new controller for compressors, but the proposed method can be extended to an optimization-based control of compressors. We assume that a PI controller is employed for compressors with a small pressure bandwidth, so (3.29) is a good approximation of the aggregate power consumption.

3.6.2 Real-time regulation

Direct Load Control Setting

We consider an aggregator or a utility, one of whose demand response (DR) programs is to use aggregate refrigeration systems for providing services to the electric grid. The aggregator receives a regulation signal, y^k, from a system operator at time step k, i.e., at time $t = kh$ for $k = 0, \cdots, K-1$, where h denotes the duration of the time step. The regulation signal

y^k may be unavailable before time step k. Our setup assumes that all the temperatures of the display cases are measured at time $t = kh$ for $k = 1, \cdots, K$ and provided to the aggregator. The task of the aggregator at time $t = kh$ is then to determine the ON/OFF control decision, $\alpha^k := (\alpha_1^k, \cdots, \alpha_n^k) \in \{0,1\}^n$, of the evaporators. The control $u_i(t)$ is chosen as α_i^k for $t \in [kh, (k+1)h)$. The aggregator has the following two objectives: (i) each display case maintains the freshness of its goods, such as fruits, vegetables, eggs and dairy products, by minimizing the deviation of each display case temperature from its desirable range, and (ii) the total power consumption (3.29) is limited by y^k, i.e., $P(\alpha^k) \leq y^k$. If $c_i \equiv \bar{c}$ for all i, then this constraint can be rewritten as

$$\sum_{i=1}^{n} \alpha_i^k = \|\alpha^k\|_0 \leq \frac{y^k}{\bar{c}} =: z^k. \tag{3.31}$$

One-step look ahead optimization

To achieve these objectives of the aggregator, we propose the following binary optimization program whose solution is employed as the control signal for $[kh, (k+1)h)$:

$$\begin{aligned}
\max_{\alpha^k \in \{0,1\}^n} \quad & J(\alpha^k) := -\int_{kh}^{(k+1)h} \sum_{i=1}^{n} \mathbf{P}_i(x_i(t))dt && \text{(3.32a)}\\
\text{subject to} \quad & \|\alpha^k\|_0 \leq z^k && \text{(3.32b)}\\
& u_i(t) = \alpha_i^k, \quad t \in [kh, (k+1)h) && \text{(3.32c)}\\
& \dot{x} = Ax + Bu + \Theta, \quad t \in [kh, (k+1)h) && \text{(3.32d)}\\
& x(kh) = x_{meas}^k, && \text{(3.32e)}
\end{aligned}$$

where $\mathbf{P}_i(x_i) := (\underline{\theta}_i - x_i)^2 + (x_i - \overline{\theta}_i)^2$. The function $\mathbf{P}_i(x_i)$ can be interpreted as the penalty for the temperature deviation from $[\underline{\theta}_i, \overline{\theta}_i]$. In other words, we use a soft constraint for the desirable temperature range. The constraint (3.32b) guarantees that the aggregate power consumption respects the system operator's regulation signal. In (3.32e), x_{meas}^k represents the display case temperature vector measured at time $t = kh$. Note that the proposed control method has a fixed duty cycle and is therefore very easy to implement. By choosing a large enough h, we can decelerate the mechanical wear of the compressors. The proposed method is applicable to the practical situation in which only a subset of the refrigerators is available to provide services at any given time, if the control laws (e.g., set-point based control) for unavailable refrigerators are known. This case can be handled by choosing the decision variable α^k as the ON/OFF control decision of the evaporators only in available refrigerators at time $t = kh$.

Given the information y^k and x_{meas}^k at time $t = kh$, we solve (3.32) and immediately use its solution as the control signal for $[kh, (k+1)h)$. Therefore, it is important to solve (3.32) in a very fast way such that the computational time is negligible compared with the duration h. However, the combinatorial optimization problem (3.32) presents two computational

challenges. First, it is NP-hard in general. Second, it requires us to repeatedly solve the dynamical system (3.32d). Suppose that there are 25 supermarkets in the aggregation and each supermarket has four refrigerators. If each refrigerator has 10 evaporator units, then $n = 25 \times 4 \times 10 = 1000$, i.e., (3.32) has 1,000 binary decision variables and depends on a 1,000-dimensional dynamical system. Furthermore, this problem cannot be decoupled as multiple lower-dimensional problems, each of which is associated with one supermarket, due to the power consumption constraint (3.32b). To rapidly solve this large-scale combinatorial optimization problem, we develop a scalable approximation algorithm.

3.6.3 Simulation results: performance validation

Case I: target profile

At the beginning of time step k (i.e., at time $t = (k-1)h$), the aggregator is requested by the SO to maintain the total power consumption by the n refrigerator units in the target range $[\underline{y}^k, \overline{y}^k]$ (kW) for $h = 15$ minutes. In other words, the following inequality must be guaranteed:

$$\underline{y}^k \leq \sum_{i=1}^{n} c_i \alpha_i^k \leq \overline{y}^k,$$

where c_i is the power consumption (kW) by refrigerator unit i when it is in the ON state. To reduce the energy consumption during the period of high demand, the profile is chosen as $\overline{y}_k = 5500$kW for $k = 9, \cdots, 16$; and $\overline{y}_k = 5000$kW otherwise in the numerical experiments.

Taking into account the penalty for temperature deviation and the constraint on the refrigerators' total power consumption, the aggregator determines the ON/OFF control for time step k as the solution of the following combinatorial optimization problem:

$$\max_{\alpha^k \in \{0,1\}^m} \quad J_k(\alpha^k) := -\int_{(k-1)h}^{kh} \sum_{i=1}^{n} \mathbf{P}_i(x_i(t))dt \tag{3.33a}$$

$$\text{subject to} \quad \underline{y}^k \leq \sum_{i=1}^{n} c_i \alpha_i^k \leq \overline{y}^k \tag{3.33b}$$

$$u(t) = \alpha^k, \quad t \in [(k-1)h, kh). \tag{3.33c}$$

Remark 3. *The inequality constraint* (3.33b) *cannot be decomposed. Therefore, although we can decompose the full n-dimensional system into N subsystems such that any two subsystems are independent of each other, the optimization problem* (3.33) *cannot be decomposed into N subproblems such that each subproblem is associated only with one of the subsystems.*

Remark 4. *Since the dynamical system* (3.28) *is defined only at $u(t) \in \{0,1\}^m$, the standard derivative is ill-defined. Therefore, we use the nonstandard derivative in the linear approximation of* (3.33)*. We can prove that the reformulated payoff function, $\hat{J}_k$, defined in*

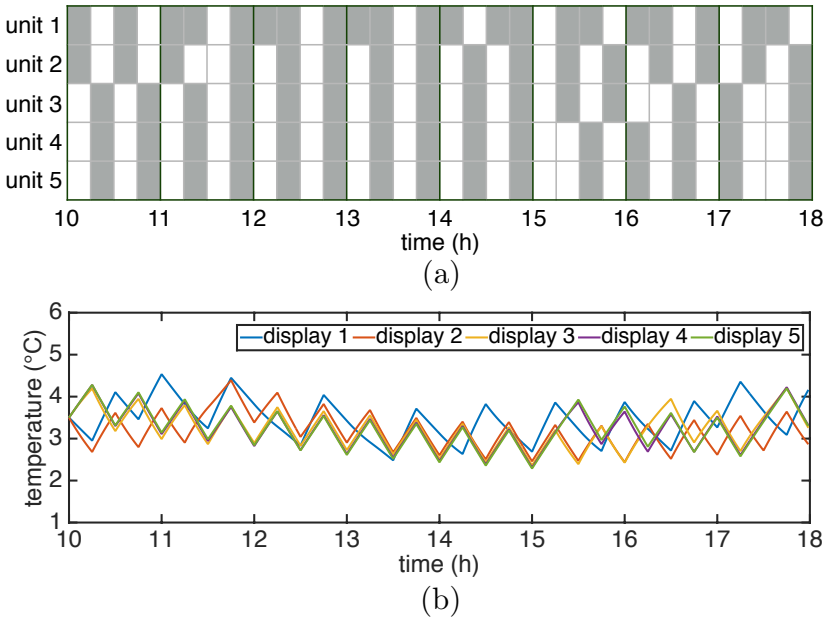

Figure 3.4: The simulation results with approximate solution α_*^k, $k = 1, \cdots, 32$ for 1000 units: (a) control signals, α_{*i}^k, $i = 1, \cdots, 5$ (grey: ON, white: OFF) (b) controlled display case temperatures, x_i, $i = 1, \cdots, 5$.

(3.20), *is concave. Therefore, the proposed suboptimality bound holds. Furthermore,* $J_k(\mathbb{I}(\cdot))$ *is submodular due to Proposition 4.*

We first set the number of evaporator units to $m = 1000$ and every 10 units have the configuration in Figure 3.2. This problem approximately takes into account 25 supermarket stores. The model parameters of the first 10 units are selected as the nominal parameter set. The model parameters of the remaining 990 units are chosen by perturbing the nominal parameter set by $\pm 10\%$ with a uniform random distribution. The power consumption by unit i is set as $c_i = 10$kW. We solve the approximate problem of (3.33) for $k = 1, \cdots, 32$ using the proposed algorithm. The first five entries of the approximate solution are shown in Figure 3.4 (a), in which we linearize the objective function at $\bar{\alpha}^k = 0$. The alternating pattern of the control induces that the display case temperatures do not deviate significantly from $[0°\text{C}, 4°\text{C}]$ as shown in Figure 3.4 (b). The suboptimality bound, ρ_*, provided in Corollary 3 is computed at $k = 1, \cdots, 32$. The computed values suggest that the approximate solution is at least 0.7-optimal solution for all time as shown in Figure 3.5 (a). This suboptimality bound is better than that of the multi-linear relaxation-based local search algorithm in [132] for non-monotone submodular maximization with knapsack constraints, which gives at least a $(3 - \sqrt{5})/2 \approx 0.309$-optimal solution. Note that the optimization problem (3.33) can also be considered as a quadratic knapsack problem, whose objective function can be formulated as $(\alpha^k)^\top \mathbf{Q} \alpha^k$ for a profit matrix $\mathbf{Q}$. When all the entries of $\mathbf{Q}$ are nonnegative,

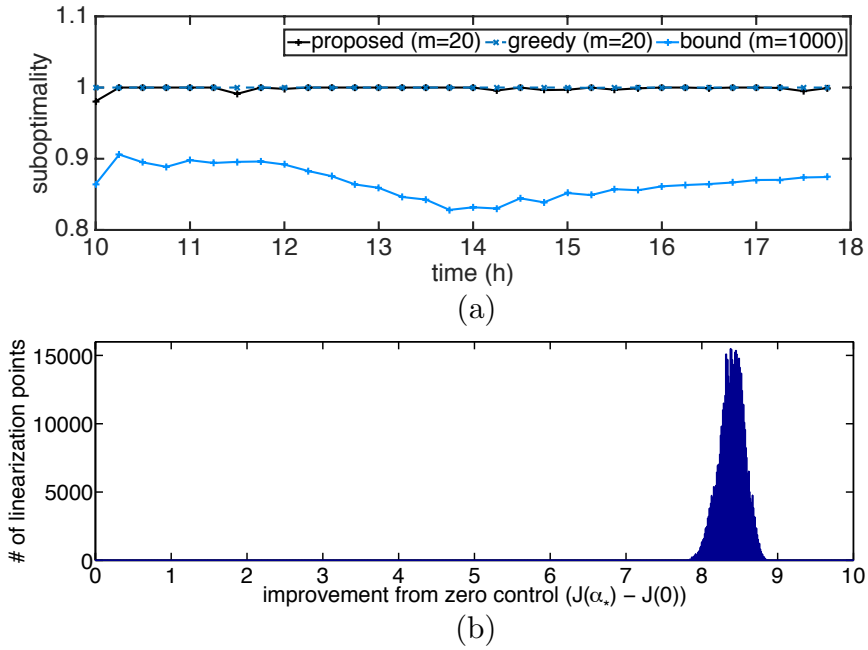

Figure 3.5: (a) The suboptimality bound $\rho_* = \hat{\rho}_*$ in the simulation with $m = 1000$; and the performance comparison of the proposed algorithm and the greedy algorithm to the oracle when $m = 20$; (b) Robustness test for the performance of the proposed algorithm with respect to the linearization point $\bar{\alpha}_1$ (with $m = 20$).

several upper bounds have been studied (see [104] and the references therein). In our case, however, $\mathbf{Q}$ contains negative entries. For such a general case, semidefinite programming (SDP) relaxation approaches have provided powerful tools for computing upper bounds for quadratic knapsack problems (e.g., [52]). We use the SDP relaxation (SQK3) in [52], which is shown to be tight [52, 104]. As it provides an upper bound of the objective function, given a feasible solution we can compute a suboptimality bound. Note that the solution $\hat{\alpha}_*^k$ obtained by our linear approximation method is feasible. We use this solution to compute the suboptimality bound ρ^{SDP} of the SDP relaxation for $m = 1000$. The calculated ρ^{SDP} is approximately 0.6, which suggests that our proposed approach gives a tighter suboptimality bound than the SDP relaxation method.

To compute the actual suboptimality, we compare the approximate solution with the optimal solution by considering a problem with 20 refrigerator evaporator units. As shown in Figure 3.5 (a), the performance of the proposed approximation algorithm is greater than 95% of the oracle's performance. In this case, the greedy algorithm performs optimally; however, we will see in the next subsection that it can get stuck at a local optimum in the presence of a more complicated constraint. The proposed algorithm takes 0.015 seconds to solve this problem while the greedy algorithm and exhaustive search take 0.57 seconds and 3112 seconds, respectively.

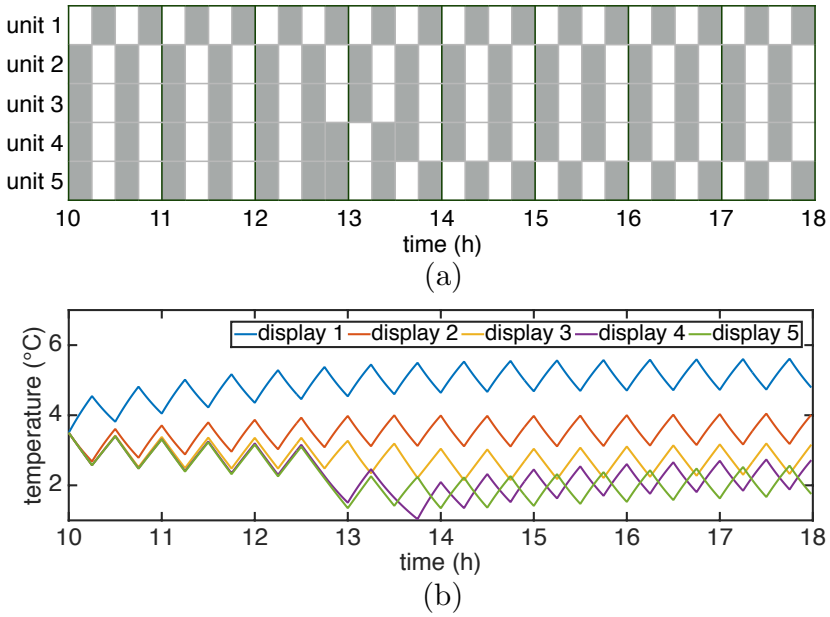

Figure 3.6: The simulation results with approximate solution α^{k*}, $k = 1, \cdots, 32$ for 1000 units with TU constraints: (a) control signals, α_i^{k*}, $i = 1, \cdots, 5$ (grey: ON, white: OFF) (b) controlled display case temperatures, x_i, $i = 1, \cdots, 5$.

Considering 10 refrigerator evaporator units with a single time step, i.e., $m = 10$ and $k = 1$, we compare the performance of the proposed algorithm and that of the greedy algorithm with 4^{10} initial values such that $\mathbf{x}_i = 2, \cdots, 5$ for $i = 1, \cdots, 10$. The ratio, $(J(\alpha_*) - J(0))/(J(\alpha^{\text{greedy}}) - J(0))$, is within $[0.99, 1.01]$ for over 99% of the initial values. Lastly, we confirm that the performance of the proposed algorithm is robust with respect to the linearization point $\bar{\alpha}^1$ as shown in Figure 3.5 (b) by solving the approximate problem for $m = 20$ and $k = 1$ with all possible 2^{20} values of $\bar{\alpha}^1$.

Case II: customized operation

In practice, a customer may specify constraints on the operation of the refrigerators. We consider the situation in which the constraint can be represented as

$$\bar{Q}\alpha^k \leq \bar{r}, \tag{3.34}$$

where $\bar{Q}$ is an $m \times n$ totally unimodular (TU) matrix and $\bar{r}$ is an m dimensional vector with integer entries. The usefulness of TU constraints is demonstrated in the following example.

Example 5. *Suppose that the power consumption by unit* 1 *is comparable to the sum of the power consumptions by units* 2 *and* 3. *The customer has a limited budget to operate the*

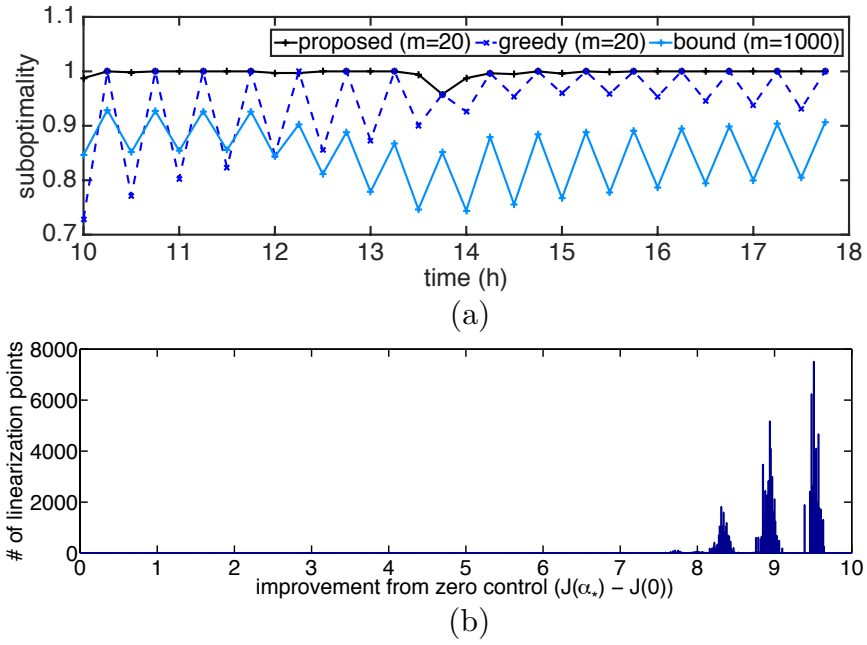

Figure 3.7: a) The suboptimality bound $\rho_* = \hat{\rho}_*$ in the simulation with $m = 1000$; and the performance comparison of the proposed algorithm and the greedy algorithm to the oracle when $m = 20$; (b) Robustness test for the performance of the proposed algorithm with respect to the linearization point $\bar{\alpha}^1$ (with $m = 20$).

refrigerators and therefore requests the following constraints to the aggregator:

$$\alpha^k_{1+10(l-1)} + \alpha^k_{2+10(l-1)} \leq 1, \;\; \alpha^k_{1+10(l-1)} + \alpha^k_{3+10(l-1)} \leq 1,$$
$$\alpha^k_{10l} + \alpha^k_{9+10(l-1)} \leq 1, \;\; \alpha^k_{10l} + \alpha^k_{8+10(l-1)} \leq 1$$

for $l = 1, \cdots, 100$. *The rest of the units satisfy the following constraint:*

$$\sum_{l=1}^{100} \sum_{j=4}^{7} \alpha^k_{j+10(l-1)} \leq \bar{z}^k,$$

where $\bar{z}_k$ *is an integer. Note that these constraints can be formulated as the inequality* (3.34) *with a TU matrix* $\bar{Q}$ *and a integer vector* $\bar{r}$.

We now consider the optimization problem (3.33) with TU constraint (3.34) instead of (3.33b). The decomposability of the problem depends on that of the TU constraint (3.34). We use the same system model as that of Case I and impose the TU constraint in Example 5 with $\bar{z}^k = 2500$ for $k = 9, \cdots, 16$ and $\bar{z}^k = 2000$ otherwise. As shown in Figure 3.6 (a), the constraints specified on units 1, 2 and 3 are satisfied by the solution of the approximate problem. We also note that unit 1 is used less frequently than in the previous case because if

unit 1 is OFF, then units 2 and 3 can be used. On the other hand, the greedy algorithm uses unit 1 frequently whenever the gain obtained by turning on unit 1 is greater than that by turning on unit 2 or unit 3. This behavior is not desirable because units 2 and 3 cannot be used when unit 1 is ON. Therefore, the proposed algorithm outperforms the greedy algorithm as shown in Figure 3.7 (a). We compute the suboptimality bound, ρ_*, provided in Corollary 3 for $k = 1, \cdots, 32$ as shown in Figure 3.7 (a). The calculated values suggest that the approximate solution is at least 0.64-optimal solution for all time.

We again consider the problem with 20 units to compare the approximate solution with the optimal solution. As shown in Figure 3.7, the performance of the proposed approximation algorithm is at least 90% of the oracle. On the other hand, the greedy algorithm achieves only $70 - 85\%$ of the oracle 7 times out of 32. To compare their performances with multiple initial values, we solve the problem using the proposed approximation algorithm and the greedy algorithm for one time step, i.e., $K = 1$, with 4^{10} initial values such that $\mathbf{x}_i = 2, \cdots, 5$ for $i = 1, \cdots, 10$ considering 10 refrigerator evaporator units. The ratio, $(J(\alpha_*) - J(0))/(J(\alpha^{\text{greedy}}) - J(0))$, is greater than 1.1 for over 99% of the initial values, i.e., the proposed algorithm performs at least 10% better for over 99% of the initial values. Furthermore, the average performance of the proposed algorithm is twice as high as that of the greedy algorithm. Finally, we perform the robustness test for the proposed algorithm with respect to the liberalization point $\bar{\alpha}^1$ as in the previous subsection by solving the problem with all possible 2^{20} values of $\bar{\alpha}^1$. As shown in Figure 3.7 (b), the performance does not deviate more than 15% from its average.

Comparison of standard and nonstandard derivatives

We now compare the performance of approximation methods based on the standard and nonstandard derivatives. For comparison purpose, we assume that the dynamical system is given by (3.28) for $i \in \mathcal{I}_1 := 2, 4, 6, 8, 12, 14, 16, 18$ and

$$\dot{x}_i = -a_{ii}(x_i - \theta_i) - \sum_{j=1}^{n} a_{ij}(x_i - x_j) - b_i e^{-\xi_i(1-u_i)t} \tag{3.35}$$

for $i \in \mathcal{I}_2 := \{1, 2, \cdots, 20\} \setminus \mathcal{I}_1$, where we set $\xi_i = 100$. The modified term in the dynamical system (3.35) models transient shutdown behavior of refrigerators after the OFF control signal is given. For $i \in \mathcal{I}_2$, the standard and nonstandard derivatives are given by

$$[D^{\text{S}} J(\bar{\alpha})]_i = -b_i \xi_i \int_0^T t e^{-\xi_i(1-\bar{\alpha}_i)t} \lambda_i^{\bar{\alpha}}(t) dt,$$
$$[D^{\text{NS}} J(\bar{\alpha})]_i = -b_i \int_0^T (1 - e^{-\xi_i t}) \lambda_i^{\bar{\alpha}}(t) dt,$$

respectively. Setting $m = 20$ and $k = 1$, we solve the the approximation problems based on the two derivatives with all possible 2^{20} values of the linearization point. Recall that $\hat{\alpha}^*$ and

(per supermarket)	energy (8h)	peak load	reserve	arbitrage
conventional	784kWh	200kW	N/A	N/A
proposed	752kWh	94kW	60kW	$35.08

Table 3.1: Summary of the simulation results per supermarket: conventional control is not applicable to a spinning reserve service or energy arbitrage.

α^* denote the solution of the approximate problems based on the nonstandard and standard derivatives, respectively. The average of $J(\hat{\alpha}_*) - J(0)$ over all the linearization points solving the approximate problem using the nonstandard derivative is 13.87, which is greater than the average 13.68 of $J(\alpha_*) - J(0)$ obtained using the standard derivative. Therefore, the approximation algorithm using the nonstandard derivative performs better than that using the standard derivative on average in this problem. This result can be explained as follows. Intuitively, the optimal solution should preferentially turn on refrigerator units in $\mathcal{I}_1$ because the transient behavior of refrigerator unit $i \in \mathcal{I}_2$ provides a refrigeration even when it is OFF. Note that the approximate solution using the nonstandard derivative preferential selects to turn on refrigerators in $\mathcal{I}_1$ because $[\hat{D}J(\bar{\alpha})]_i$ for $i \in \mathcal{I}_2$ is deflated from the case of Section 3.6.3. However, the standard derivative $[DJ(\bar{\alpha})]_i$ is inflated for $i \in \mathcal{I}_2$ and, therefore, its approximate solution preferentially turns on refrigerators in $\mathcal{I}_2$. As a result, the approximate solution using the nonstandard derivative slightly outperforms the standard derivative. In general, the one of two approximate solutions that outperforms another is problem-dependent (see also Appendix C.2).

3.6.4 Simulation results: demand response

We consider the situation in which 25 supermarkets are enrolled in a direct load control program provided by an aggregator. Each supermarket has four refrigerators and each refrigerator contains 10 evaporator units. Therefore, the number of control variables or display cases is $n = 25 \times 4 \times 10 = 1000$. The aggregator receives a real-time regulation signal y^k every $h = 15$ minutes from a system operator. We use the following parameters in the simulations: $a_{ii} = 0.5$ for display cases in the two ends of each refrigerator, $a_{ii} = 0.3$ for the rest of the display cases, $a_{ij} = 1.5$ for $i \neq j$ for the first 10 display cases and a_{ij}'s of the remaining 990 display cases are chosen by perturbing the nominal parameter set by $\pm 10\%$. Furthermore, we set $b_i = 12$, $c_i \equiv \bar{c} = 5$kW, $\underline{\theta} = 0°$C, $\overline{\theta} = 4.5°$C and $\theta_i = 20°$C. We use Algorithm 1 to generate the control signals. The linearization point is chosen as $\bar{\alpha} = 0$. All tests were performed on a 2.3GHz Intel Core i7 processor with 16GB RAM. The computational time of the proposed algorithm for $n = 1000$ display cases is 0.81 seconds with its MATLAB implementation. In all the simulations below, the suboptimality bound ρ of the proposed method is greater than 0.65.

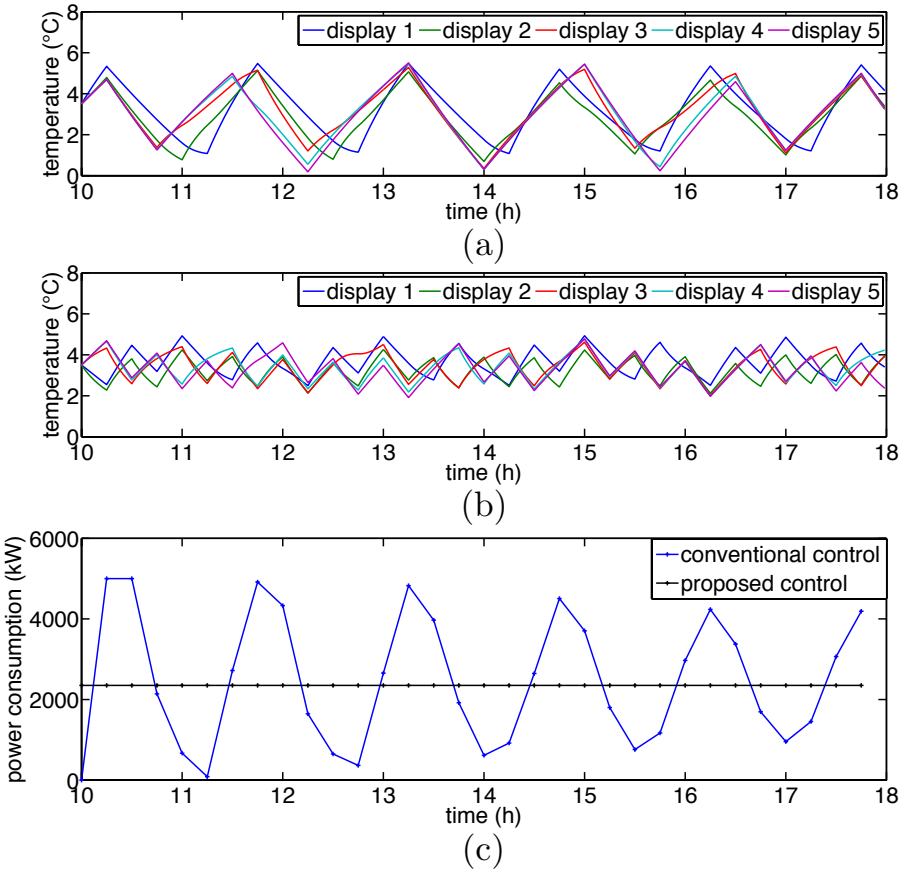

Figure 3.8: Temperature profiles of display cases $1, \cdots, 5$ controlled by (a) the conventional set-point based control; (b) the proposed optimization-based control; and (c) total power consumption profiles.

Comparison with conventional control

We compare the proposed optimization-based control with the conventional set-point based control (3.30). We use $\underline{\theta} = 1.5°\text{C}$ in the simulation for conventional control because of the strong temperature undershoot of the conventional control approach. Figure 3.8 (a) and (b) show the temperatures for the first five display cases controlled by the conventional control and the proposed control, respectively. Because our method actively takes into account the heat transfer between neighboring display cases, it achieves a better energy efficiency (saves 4.5% of energy) than the conventional control approach. Our result confirms a well-known fact that the conventional approach induces the synchronization of evaporators and therefore causes high peak power consumption as shown in Figure 3.8 (c). By contrast, the proposed method induces the desynchronization of evaporators and therefore flattens power consumption profile. The simulations are summarized in Table 3.1.

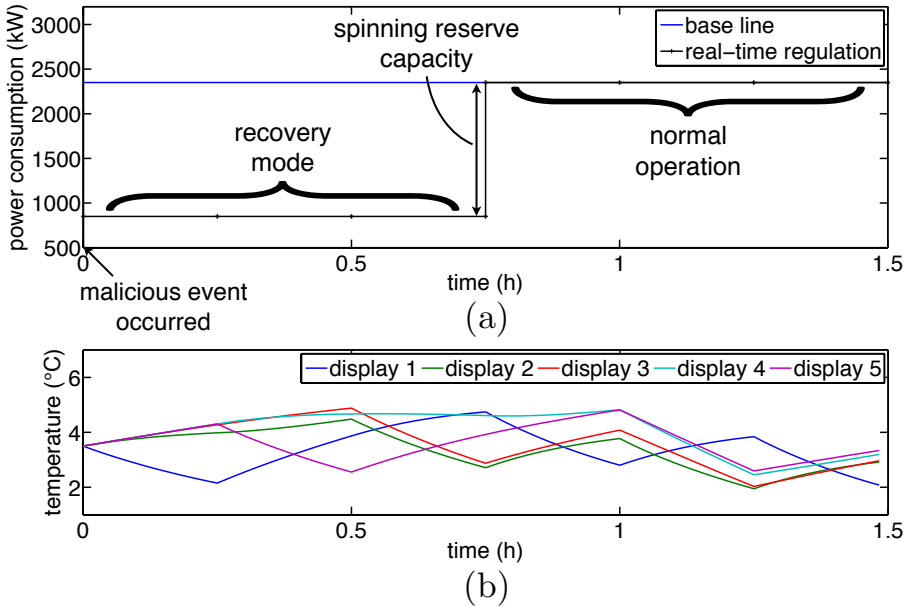

Figure 3.9: (a) A real-time regulation signal for a spinning reserve; and (b) temperature profiles of display cases $1, \cdots, 5$ controlled by the proposed algorithm.

Real-time regulation I: power system resilience

We study the potential of using supermarket refrigeration systems as a spinning reserve (e.g., [69]) for power system resilience. In particular, we consider a scenario in which supply shortage occurs because of an incorrect forecast of renewables, a cyber-attack to the supervisory control and data acquisition systems for the electric grid, or a deceptive manipulation of electricity markets. By providing the real-time regulation signal y^k in Figure 3.9 (a), the system operator requests the aggregation of supermarkets to serve as a spinning reserve of 1,500kW capacity for 45 minutes. To meet this request, we assume that the supermarkets use display case covers, which are normally utilized during off-hours. If the covers are used, the heat transfer between each display case and its ambient air is reduced: we set $a_{ii} = 0.3$ for display cases in the two ends of each refrigerator and $a_{ii} = 0.2$ for the rest. In this setting, the aggregation can work as the requested spinning reserve of 15kW per refrigerator and the temperature deviation is less than 0.5°C (see Figure 3.9 (b)).

Real-time regulation II: energy arbitrage

We now consider the case of energy arbitrage (e.g., [87]). Based on the energy price in a 15 minute-ahead wholesale electricity market, the system operator provides a real-time regulation signal that induces the aggregator to buy energy when its price is low and to sell it when the price is high. Figure 3.10 (a) shows the 15 minute-ahead wholesale electricity market at the Austin node in ERCOT [35], and Figure 3.10 (b) shows the base-line power procured by the aggregator in a day-ahead market or through a long-term contract and the

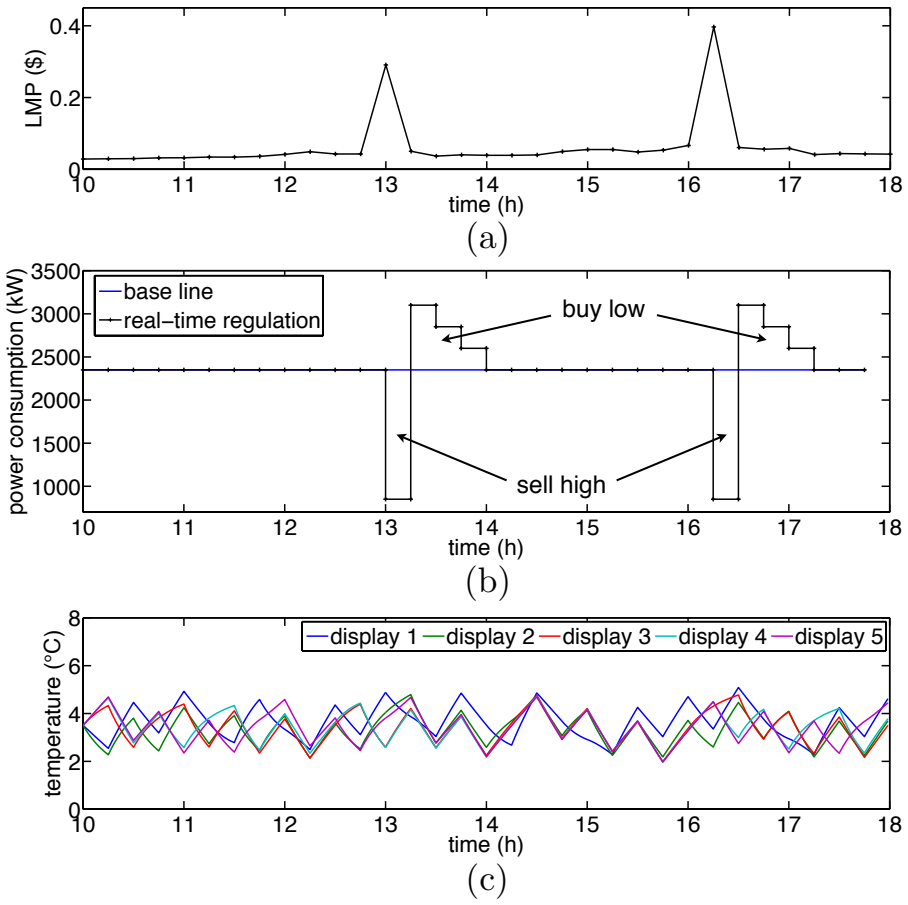

Figure 3.10: (a) Locational marginal price on July 3, 2013 at the Austin node in the ERCOT; (b) a real-time regulation signal for energy arbitrage; and (c) temperature profiles of display cases $1, \cdots, 5$ controlled by the proposed algorithm.

real-time regulation signal y^k provided by the system operator from 10 am to 6 pm. The controlled temperatures for the first five display cases are presented in Figure 3.10 (c). The temperature deviation from $[\underline{\theta}, \overline{\theta}]$ is less than 0.6°C because right after the reduction in total power consumption, the regulation signal encourages enough cooling to recover the desired temperature levels. The profit of the aggregator through this arbitraging is \$877.36, i.e., \$8.77 per refrigerator.

Chapter 4

Utility Learning Model Predictive Control for Personal Electric Loads

4.1 Personalized control for electric loads and demand-side management

One of the major changes expected for the future electric power grid is a transition towards *demand side management*, with objectives including (*i*) the enhancement of *energy efficiency* of a customer's electric loads and (*ii*) the guiding of a customer's electricity usage patterns or flexibility via energy pricing and/or incentive payment schemes [122]. The latter is also called *demand response*. Increasing controllability of a customer's loads with appropriate control schemes is expected to play a key role in achieving these objectives [23]. A number of sophisticated control and scheduling methods have been proposed for various types of personal loads, such as electric vehicles [113, 83], deferrable loads [41, 112] and thermostatically controlled loads [86, 51, 139, 141]. In addition to achieving good energy efficiency or demand response, a main task of the control method is to ensure that each customer's comfort does not deteriorate. Personalized control of loads taking into account the customer's satisfaction has recently shaped up as a successful concept for addressing this goal. The Nest smart thermostat [56] can be taken as one example.

This work proposes a novel personalized control framework that is suitable for customer's electric loads. The proposed method has the capability of learning online the customer's utility function that represents her satisfaction (e.g., comfort and energy saving) with the controlled personal loads. More specifically, the customer reports her satisfaction with the control performance, e.g., in the form of a simple rating, say a number between 1 and 5, through a given interface between the customer and the controller at any convenient time. For example, the customer can rate the controller of her air conditioner judging from the amount of energy savings and her comfort with the indoor temperature and report it to the controller via her smartphone. Using the satisfaction data, the proposed method infers the customer's utility function that is used to automatically customize the controller for the

customer.

In the proposed framework, the customer's utility is learned online from her satisfaction data and the control objective is changed accordingly. The approach is suitable even for the situation in which the customer's behavior affects the system controlled by her loads. For example, opening a window affects the indoor temperature controlled by the customer's air conditioner. To achieve online learning of the customer's utility and behavior in this setting, we propose a *system manager*, which is interfaced with the customer and the electric load. The system manager consists of a *behavior learning module*, a *controller* and a *utility learning module*. Intuitively speaking, the behavior and utility learning modules infer the customer's behavior and utility from the state measurements and the satisfaction data, respectively, and deliver the inferred values to the controller that uses them to compute personalized control actions. More specifically, the behavior learning module estimates and predicts online the effect of the customer's behavior, which is not known *a priori* or cannot be modeled explicitly, on the system state controlled by her loads, e.g., indoor temperature. A Gaussian process (GP) model is used to learn the effect from online state measurements and construct a stochastic prediction [110, 103, 70]. The prediction of the customer's behavior enhances the performance of the model predictive control (MPC) [71, 45, 3, 125], which is used as a controller in the proposed framework. The objective of the MPC controller is chosen as the customer's utility function, which is inferred by the utility learning module. The identification of the customer's utility is performed online by solving a convex optimization problem whenever the customer reports a new satisfaction data to the system manager. Furthermore, the module provides an option to learn the user's utility online as a time-varying function using GP regression and therefore to generate an estimate of the utility at all sampling times, i.e., also when no satisfaction data is available. The MPC controller immediately sets the updated customer's utility as its objective function and thereby modifies the resulting control law according to her preferences. We call this framework the *utility learning model predictive control*.

The most distinctive feature of the proposed framework is that the customer's utility is learned online during closed-loop control. Therefore, (*i*) the controller is capable of immediately updating its objective function as the identified or predicted customer's utility and of controlling the system to maximize the new objective, and (*ii*) no separate training period to learn the customer's utility is needed. These are advantages of the proposed framework, which tightly integrates the utility learning and system control, over offline utility learning approaches (e.g., [68, 13]).

4.2 The setup

We begin by describing the setting of the utility learning model predictive control framework, which consists of the customer, the load (and the system controlled by the load), and the system manager.

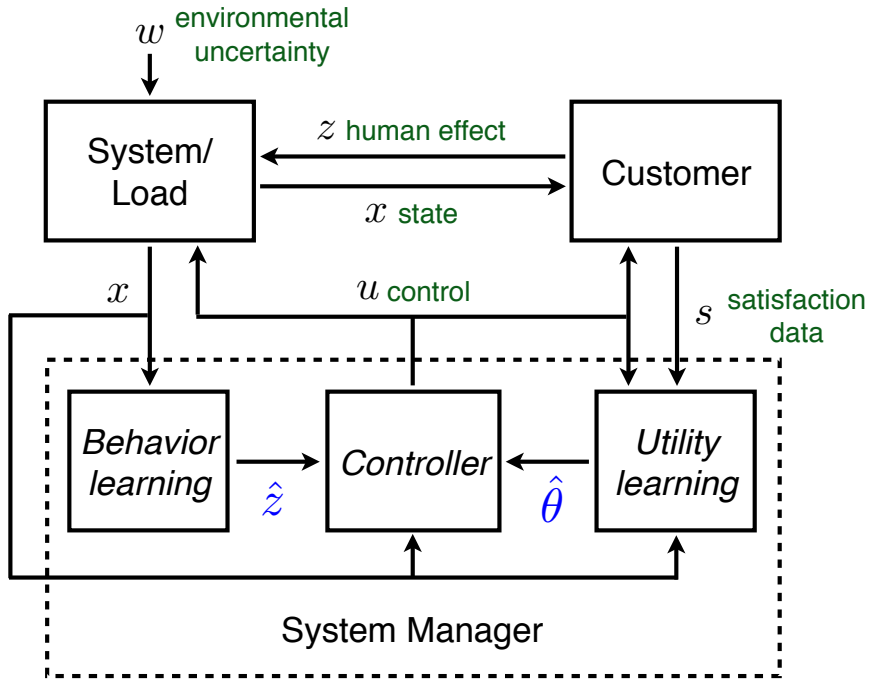

Figure 4.1: Information flow among the customer, the system/load and the system manager.

4.2.1 System model

Consider an electric load of a customer, e.g., a water heating system, an air conditioner or a CO_2 controller. Let $x_\tau \in \mathbb{R}^n$ be the system state, e.g., the water temperature, the indoor temperature, or the indoor CO_2 level, at time $\tau\Delta t$, $\tau = 0, 1, \cdots$ for some positive time step Δt. We let $u_\tau \in \mathbb{R}^m$ be the control input, e.g., the power consumption by the load, and assume that the system can be represented or at least well approximated by linear system dynamics of the form

$$x_{\tau+1} = Ax_\tau + Bu_\tau + z_\tau + w_\tau, \tag{4.1}$$

where $w_\tau \in \mathbb{R}^n$ models the exogenous (or environmental) uncertainty (with a known probability distribution) that affects the system. Here, $z_\tau \in \mathbb{R}^n$ represents an unknown effect of the customer's behavior on the system dynamics. The customer's behavior is assumed to be independent of the system state but to be dependent on time. This effect can make the system model arbitrarily wrong if we do not have good knowledge (or estimation) of the customer's behavior. Neglecting this effect can cause a significant model mismatch and deteriorate the closed-loop performance of the MPC controller.

4.2.2 Customer's utility function

The customer has multiple objectives to control her load. Most commonly, she wants to maximize her utility associated with energy savings and, at the same time, maximize her utility associated with her comfort/convenience. To mathematically model the ith objective of the customer, we choose a basis utility function, $\beta_i : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, for each $i = 1, \cdots, K$. Then, the total utility function of the customer (at time $\tau\Delta t$) can be modeled as the weighted sum of the basis utility functions, i.e.,

$$J(x_\tau, u_\tau; \theta) := \sum_{i=1}^{K} \theta^i \beta_i(x_\tau, u_\tau),$$

where $\theta^i \in \mathbb{R}$ is the weight of the ith basis. The vector of weights is denoted by $\theta \in \mathbb{R}^K$. Here, the basis functions are assumed to be concave. This assumption is valid in many practical problems.

Example 6. *Let x_τ and u_τ be the indoor temperature and the power consumption by the customer's air conditioner at time $\tau\Delta t$, respectively. Let c_τ (dollar/kWh) be the energy price; then, the customer's utility function with respect to energy costs can be modeled as the negative value of energy costs, i.e., $\beta_1(x_\tau, u_\tau) := -c_\tau u_\tau \Delta t$. Assuming that the user feels most comfortable at the temperature $\bar{x}$, the customer's utility function with respect to her comfort can be formulated as $\beta_2(x_\tau, u_\tau) := -(x_\tau - \bar{x})^2 \Delta t$. This is a modified version of the comfort metric defined in terms of temperature deviation proposed in the ANSI/ASHRAE standards [126]. Note that these two basis utility functions are concave. These basis functions are employed in the numerical tests in Section 4.4.*

The task is then to learn the weights to identify and infer the customer's utility. Learning the weights online from appropriate data is essential to align the system manager's objective with the customer's actual objective. Given the basis weights, the system manager synthesizes a controller that maximizes the utility function in a receding horizon approach as will be shown in Section 4.3.

4.2.3 Information flow

The information flow between the customer, the system controlled by the load, and the system manager is depicted in Figure 4.1. The customer's behavior affects the system state (e.g., indoor temperature) and this effect is represented by z as proposed. The system is also perturbed by the exogenous disturbance or noise, w, and the current state information, x, is delivered to all three modules of the system manager: (i) the behavior learning module; (ii) the controller; and (iii) the utility learning module.

In the behavior learning module, the customer's effect, z, is learned by Gaussian process (GP) regression. The inferred value of z is denoted as $\hat{z}$. Weakly periodic behavior of the customer is informative when designing a kernel of the GP model with good performance.

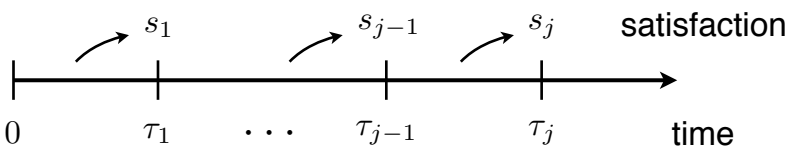

Figure 4.2: The satisfaction data, $s_j \in \mathbb{R}^K$, is provided at $t_j := \tau_j \Delta t$ to the system manager.

Judging from the system state and control during $(\tau_{j-1}\Delta t, \tau_j \Delta t]$ (and/or even before the interval), the customer can report her satisfaction, $s_j \in \mathbb{R}$, to the system manager, where $\tau_{j-1}\Delta t$ is the time at which satisfaction was last reported and $\tau_j \Delta t$ is the current time as illustrated in Figure 4.2. Therefore, $\{\tau_j\}_{j=1,\cdots}$ is a subsequence of the MPC sampling time steps $\tau = 0, 1, \cdots$. Note that the reporting time can be arbitrarily chosen by the customer. For example, the user rates the performance of her smart thermostat as a number, s_j, between 1 and 5 based on the amount of energy saving and her comfort with the indoor temperature during $(\tau_{j-1}\Delta t, \tau_j \Delta t]$ and reports it to the thermostat via her smartphone. The utility learning module then identifies the basis weights given the satisfaction data $\{s_j, s_{j-1} \cdots, s_{j-M+1}\}$, where M can be chosen by the customer. The identified basis weight vector is denoted as θ^*. The module also provides an option to learn the basis weight vector as a function in time using GP regression. The inferred value of θ is denoted as $\hat{\theta}$. The utility identification and learning procedures will be discussed in Section 4.3.3.

The controller of the system manager receives the information on the system state, x, the inferred behavioral effect, $\hat{z}$, and the identified (or inferred) basis weights, θ^* (or $\hat{\theta}$). It then generates a control signal u using an MPC approach that maximizes the updated user's utility.

4.3 Utility learning model predictive control

We now propose the utility learning model predictive control method under the setting described in the previous section. The proposed method builds on the three main building blocks: the controller, and the behavior and utility learning modules. The details for each of these components are discussed in the following.

4.3.1 Model predictive control interfaced with learned customer's utility function and behavior

A model predictive control scheme is proposed, which allows to tightly couple the three blocks and compute a control input minimizing the cost (or maximizing the utility) obtained from the utility learning module, subject to the imposed system constraints and the prediction of the model dynamics inferred in the behavior learning module. The system (4.1) is stochastic due to the uncertain customer's effect as well as the exogenous uncertainty. To take into

account the stochastic dynamics, a stochastic MPC problem is formulated:

$$\begin{aligned} \min_{\mathbf{u}_\tau \in \mathbb{R}^{mN}} \quad & -\sum_{l=\tau}^{\tau+N-1} \mathbb{E}[J(x_l, u_l; \hat{\theta})] \\ \text{subject to} \quad & x_\tau = x_\tau^* \\ & x_{\tau+i+1} = Ax_{\tau+i} + Bu_{\tau+i} + \hat{z}_{\tau+i} + w_{\tau+i}, \\ & u_{\tau+i} \in \mathcal{U}, \quad i = 0, \cdots, N-1, \end{aligned} \tag{4.2}$$

where x_τ^* is the system state measured/estimated at time step τ, $\mathbf{u}_\tau := \{u_l\}_{l=\tau}^{\tau+N-1}$ is the sequence of control inputs and N is the prediction horizon. $\mathcal{U}$ is a compact set, defining the set of admissible control values. It should be noted that state constraints are not explicitly incorporated in the MPC problem. In the application context of electric loads, it is assumed that state constraints are introduced as soft constraints and included as a basis function. Note that with a slight abuse of notation we use $\hat{\theta}$ in the MPC cost as the weight provided by the utility learning module for both updating strategies proposed in Section 4.3.3.

By using the mean and variance of $\hat{z}_{\tau+i}$ over the prediction horizon provided by the behavior learning module (see Section 4.3.2) and the given probability distribution of $w_{\tau+i}$, the stochastic MPC problem can be reduced to a deterministic formulation. In the case of a quadratic stage cost on the states and inputs, for instance, the stochastic MPC problem corresponds to employing the expected values $\bar{z}_{\tau+i} = \mathbb{E}[\hat{z}_{\tau+i}], \bar{w}_{\tau+i} = \mathbb{E}[w_{\tau+i}]$ in the dynamics. This can be easily seen from $\mathbb{E}[x_l^\top Q x_l] = \mathrm{tr}(Q\mathrm{Var}[x_l]) + \mathbb{E}[x_l]^\top Q\mathbb{E}[x_l]$ and the fact that $\mathrm{Var}[x_l] = \mathrm{Var}[\hat{z}_l] + \mathrm{Var}[w_l]$ is known or provided by the prediction and hence a given constant at the time of computation. Let $\mathbf{u}_\tau^*$ be the solution to the optimization problem at time step τ. The MPC control law, $\kappa : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^m$, is defined in a receding horizon fashion by applying the first control input, i.e., $\kappa(x, \tau) = u_\tau^*$.

4.3.2 Learning the customer's behavior

The effect of the customer's behavior is learned from data collected during online control. We make use of a Gaussian process model, providing a general non-parametric modeling framework and a posterior uncertainty description. Informally, a Gaussian process (GP) can be thought of as describing a probability distribution over functions. Due to the fact that human behavior generally follows a daily routine, the customer's effect can be assumed to exhibit dynamics with periodic characteristics, which may, however, change over multiple period lengths. A GP model for capturing 'locally periodic' dynamical effects has been recently proposed and incorporated in an MPC controller in [70]. A function g is here said to be 'locally periodic' provided that $g(t) \approx g(t + \mathbf{n}\omega)$ if $\mathbf{n}\omega \ll l$ but $g(t) \not\approx g(t + \mathbf{n}\omega)$ if $\mathbf{n}\omega \gg l$, where ω is the period length, l is a measure of locality and $\mathbf{n} = 1, 2, \cdots$. An example is shown in Figure 4.3. This approach can be directly employed to learn the customer's effect online as described in the following. We only discuss the main steps of GP regression, more details about Gaussian processes can be found in [110], see also [103] for a more general overview of kernel methods for system identification.

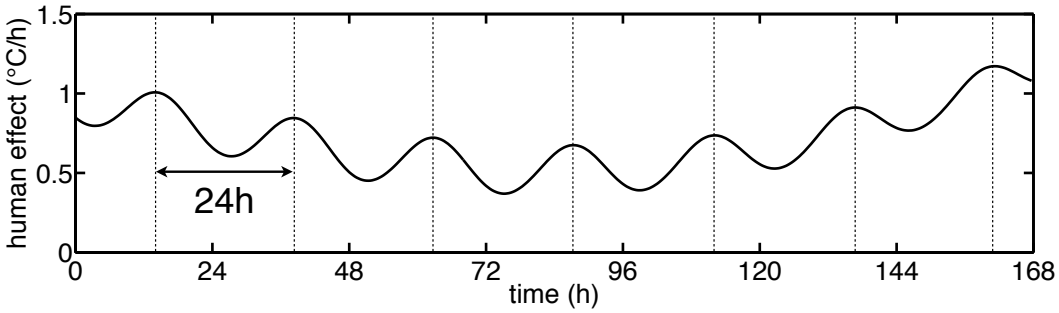

Figure 4.3: An example of quasi-periodic human effect.

Consider the case of learning a scalar function $g_i : \mathbb{R} \to \mathbb{R}$ for $z_\tau^i = g_i(\tau \Delta t)$, where z_τ^i denotes the ith element of the vector z. A Gaussian process model is defined by a mean function $\mu : \mathbb{R} \to \mathbb{R}$ and a covariance function $k : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, which together form the GP prior. The choice of this prior is an important modeling assumption as it determines the properties of the functions considered for z_τ^i and in turn affects convergence and extrapolation properties. In the particular case of time varying functions, extrapolation is crucial because future function values have to be predicted from past observations.

A prior covariance function focusing probability mass on periodic functions has been recently presented in [70], which is composed of a square exponential kernel $k_{SE}(t, t')$ and a periodic kernel $k_P(t, t')$ for $t, t' \in \mathbb{R}$:

$$k(t, t') = \sigma^2 \cdot k_{SE}(t, t') \cdot k_P(t, t'), \tag{4.3}$$

where $k_{SE}(t, t') = \exp(-(t - t')^2 / 2 l_{SE}^2)$ with length scale l_{SE}, $k_P(t, t') = \exp(-2 \sin^2(\frac{\pi}{\omega}(t - t'))/l_P^2)$ with length scale l_P and period length ω, and σ^2 is the signal variance. Intuitively speaking, this kernel considers two input times similar if they are similar under both the square exponential and the periodic kernel. The square exponential kernel admits the function to be not strictly periodic and for $l_{SE} \gg \omega$ to vary over a longer time scale. As shown in [70], this kernel has the significant advantage that it offers good extrapolation properties for functions with this property. A function drawn from a GP with this kernel is shown for the application example in Figure 4.7, a more detailed discussion and illustration of the different parameters can be found in [70]. The role of the mean function is less critical. Since no prior hypothesis about the mean of the customer behavior is available, the mean function is here chosen to zero.

After choosing the prior hypothesis class, GP regression is performed by inferring the function from observed data. Samples of the customer effect z_τ are obtained from the dynamical model in (4.1) by measuring the state x_τ at two consecutive sampling times. We denote by $\tilde{\mathbf{z}}_T^i \in \mathbb{R}^D$ the vector of observations at time points $T \in \mathbb{R}^D$. If data points are observed with Gaussian noise $\epsilon_z \sim \mathcal{N}(0, \sigma_z^2)$, i.e., $\tilde{z}_\tau^i = z_\tau^i + \epsilon_z$, $i = 1, \cdots, n$, then, given the Gaussian process prior, the posterior distribution is also Gaussian [110]. For a given test

point $t = j\Delta t$, the resulting predictive distribution with mean $\bar{z}_j^i$ and variance $\mathbb{V}(z_j^i)$ is given by:

$$\bar{z}_j^i = \mu(t) + K(t,T)^\top (K(T,T) + \sigma_z^2 I)^{-1} (\tilde{\mathbf{z}}_T - \bar{\mu}(T)), \tag{4.4a}$$

$$\mathbb{V}(z_j^i) = K(t,t) - K(t,T)^\top (K(T,T) + \sigma_z^2 I)^{-1} K(T,t), \tag{4.4b}$$

where $[K(T,T')]_{mn} := k(T_m, T'_n)$, $[\bar{\mu}(T)]_m := \mu(T_m)$. The posterior mean and variance completely characterize the inferred value, $\hat{z}_j^i$, of z_j^i. By means of (4.4), a prediction of the mean and variance of $\hat{z}_j^i$, $j = \tau, \cdots, \tau + N - 1$, is constructed for $i = 1, \ldots, n$ and is incorporated in the MPC problem.

The covariance function has several free hyperparameters, which may be difficult to choose in practice. While in the case of human behavior the period length can be chosen from intuition as 24h, choosing good values for the remaining hyperparameters is important to provide a good model. In GP regression, hyperparameters are often inferred from the available training data. In the considered case of online learning, inference is performed recurrently during closed-loop operation, after a new batch of L data points has been collected, where L is a tuning parameter. Different techniques for hyperparameter inference are available, see [110] for a detailed discussion, or [70] for a customized approach for periodic kernels.

4.3.3 Learning the customer's utility

As proposed in Section 4.2.3, the customer can provide the satisfaction data to the system manager at any sampling time. We propose two techniques for utility learning in the following. The first approach is based on an immediate identification of the weights as soon as new satisfaction data is available, after which the weight is kept constant until the next update. The second approach additionally learns time varying functions for the basis weights, which are used to predict the weights in the future and are updated with every new satisfaction data point. As shown in Figure 4.2, let $s_j \in \mathbb{R}$ be the customer's satisfaction with the system operation during $(t_{j-1}, t_j] = (\tau_{j-1}\Delta t, \tau_j \Delta t]$. We also denote

$$d_j := \frac{\sum_{\tau=\tau_{j-1}+1}^{\tau_j} (\beta_1(x_\tau^*, u_\tau^*), \cdots, \beta_K(x_\tau^*, u_\tau^*))}{\tau_j - \tau_{j-1} - 1}. \tag{4.5}$$

In other words, $d_j \in \mathbb{R}^K$ represents the *averaged basis function* with the MPC control over $(t_{j-1}, t_j]$.

At time step τ_j, the system manager considers the following data:

$$\{(d_{j-M+1}, s_{j-M+1}), \cdots, (d_j, s_j)\}, \tag{4.6}$$

where the parameter M determines how many data points in the past are taken into account, which can be chosen by the customer. If the customer wants the system manager to consider her most recent previous satisfaction data and the current satisfaction data, for example, then she should choose $M = 2$.

Online identification of the customer's utility

From the data, the system manager estimates the basis weights θ^* at time step τ_j by solving the following optimization problem:

$$\begin{aligned} \theta^*_{\tau_j} = \arg\min_{\theta\in\mathbb{R}^K} \quad & \sum_{i=j-M+1}^{j} (d_i^\top \theta - s_i)^2 + R(\theta) \\ \text{subject to} \quad & \theta_{\mathbf{L}} \le \theta \le \theta_{\mathbf{U}}, \end{aligned} \tag{4.7}$$

where the lower and upper bounds $\theta_{\mathbf{L}}$ and $\theta_{\mathbf{U}}$ of the basis weights, respectively, can be chosen by the customer. Here, $R : \mathbb{R}^K \to \mathbb{R}$ is a strictly convex function, which can be interpreted as a regularizer to guarantee the uniqueness of the solution. Let $D := (d_{j-M+1}, \cdots, d_j)^\top \in \mathbb{R}^{M\times K}$, and $s := (s_{j-M+1}, \cdots, s_j) \in \mathbb{R}^M$. Then, the cost function can be rewritten as

$$\sum_{i=j-M+1}^{j} (d_i^\top \theta - s_i)^2 + R(\theta) = \|D\theta - s\|^2 + R(\theta).$$

If $\text{rank}(D) \ge K$, then the optimization problem (4.7) has a unique solution even if $R \equiv 0$. However, the rank condition is not guaranteed, even when $M \ge K$, because of the possibility that the linearly independent data in (4.6) are less than K. Therefore, an appropriate strictly convex function R needs to be chosen to guarantee the uniqueness of the solution to (4.7). In the examples in Section 4.4, we choose R as a 'smoothness-inducing function'. More specifically, set

$$R(\theta) = \lambda\|\theta - \theta^*_{\tau_{j-1}}\|^2,$$

where $\theta^*_{\tau_{j-1}}$ indicates the basis weight vector identified at the last utility learning time, i.e., $\tau_{j-1}\Delta t$, and λ is a positive constant. This regularization function encourages the smoothness in the variation of the basis weights, i.e., it discourages abrupt jumps in the objective (utility) function in the MPC module. It is important to note that the utility identification is performed online. Therefore, the proposed method is ideal when the customer's utility function changes over time and these changes are critical in controlling the load.

Online learning of the customer's utility

Using the identified weights, a time-varying function for θ can be learned online, which allows to predict the weight vector for all time points, i.e., even when no satisfaction data is available. A GP model as described in Section 4.3.2 is employed, where the observed data points in this case are given by the identified weight vectors, $\{\theta^{*i}_{\tau_1}, \cdots, \theta^{*i}_{\tau_j}\}$, defined in (4.7). Due to the overall periodicity of human behavior, the customer's preference can again be assumed to vary periodically and the kernel in (4.3) is employed. All weights are initially assumed to be equivalently important by choosing a constant mean equal to one, i.e. $\mu(t) = 1, \forall t$. Let $T_j := (t_1, \cdots, t_j) \in \mathbb{R}^j$, where $t_l := \tau_l\Delta t$, $l = 1, 2, \cdots$. For $i = 1, \cdots, K$, let $\tilde{\theta}^{*i}_j \in \mathbb{R}^j$ denote the vector of observations for the ith weight at times T_j,

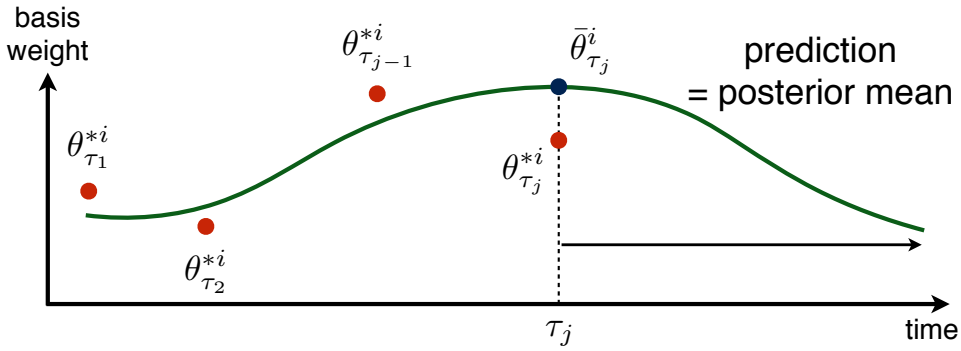

Figure 4.4: The identified basis weights $\theta^{*i}_{\tau_j}$'s (red dots) can be obtained by solving (4.7). The prediction, $\bar{\theta}^{i}_{\tau_j}$, (blue dot), is taken as the mean value of $\hat{\theta}^{i}_{\tau_j}$ inferred from the data $\{\theta^{*i}_{\tau_1}, \cdots, \theta^{*i}_{\tau_j}\}$. At step τ_j+l, $\{\bar{\theta}^{i}_{\tau_j+l}, \cdots, \bar{\theta}^{i}_{\tau_j+l+N-1}\}$ is used as the basis weight in the objective function of the MPC (4.2) for the prediction horizon.

i.e., $\tilde{\theta}^{*i}_j := (\theta^{*i}_{\tau_1}, \cdots, \theta^{*i}_{\tau_j})$. We also assume that $\theta^{*i}_{\tau_l} = \theta^{i}_{\tau_l} + \epsilon^i$, $l = 1, 2, \cdots$, where $\epsilon^i \sim \mathcal{N}(0, \sigma_i^2)$. An estimation of the weight vector at time $t = (\tau_j + l)\Delta t \in \mathbb{R}$, $l = 0, 1, 2, \cdots$ is obtained by using the mean predictive equation for the inferred value $\hat{\theta}^{i}_{\tau_j+l}$:

$$\begin{aligned}\bar{\theta}^{i}_{\tau_j+l} =& K_i(t, T_j)^\top (K_i(T_j, T_j) + \sigma_i^2 I)^{-1} (\tilde{\theta}^{*i}_j - \bar{\mu}(T_j)) \\ &+ \mu(t),\end{aligned} \tag{4.8}$$

where $k_i(\cdot, \cdot)$ is the covariance function in (4.3), $\mu(t)$ is the mean function of θ^i, $[K_i(T_j, T'_j)]_{mn} := k_i(t_m, t'_n)$ and $\bar{\mu}_m(T_j) := \mu(t_m)$. Equation (4.8) is used to generate the predictions, $\bar{\theta}^{i}_{\tau_j+l}$'s, and the GP model is updated whenever a new satisfaction data is reported. At each sampling time, $\tau_j + l$, we use the basis weight in the MPC objective function (4.2) as $\{\bar{\theta}^{i}_{\tau_j+l}, \cdots, \bar{\theta}^{i}_{\tau_j+l+N-1}\}$ for the prediction horizon, provided that $\tau_j + l < \tau_{j+1}$, as shown in Figure 4.4.

4.3.4 Algorithm

We summarize the description of the proposed utility learning model predictive control in Algorithm 2. The behavior learning module generates the forecast of the customer's effect at $N-1$ future time points by performing the Gaussian process regression discussed in Section 4.3.2. The prediction of the customer's behavior is used in the MPC (4.2) to compute the control signal in a receding horizon fashion. As soon as the customer reports her satisfaction, the utility learning module kicks in and solves the problem (4.7) to identify the customer's utility. The identified basis weights are used in the MPC from the next time step until a new satisfaction data point is provided if the prediction of the weights is not employed. If the basis weight vector is learned as a time-varying function by GP regression (optional),

Algorithm 2: Utility learning model predictive control

1 **Initialization:**
2 $\theta^* \leftarrow \theta_0$;
3 L, N, M selected by the customer;
4 **for** $\tau = 0, 1, \cdots$ **do**
5 $x_\tau^* \leftarrow$ current measured/estimated state
6 **Behavior learning module:**
7 **if** $mod(\tau, L) = 0$ **then**
8 Perform hyperparameter estimation;
9 **end**
10 $\{\hat{z}_\tau, \cdots, \hat{z}_{\tau+N-1}\} \leftarrow$ prediction of the customer's effect from GP model (4.4);
11 **Controller:**
12 $\mathbf{u}_\tau^* := \{u_i^\circ\}_{i=\tau}^{\tau+N-1} \leftarrow$ solution of the MPC (4.2);
13 $u_\tau^* \leftarrow u_\tau^\circ$;
14 **Utility learning module:**
15 **if** *the customer reports the* j*th satisfaction,* s_j*, at time* $\tau\Delta t$ **then**
16 $\tau_j \leftarrow \tau$;
17 $d_j \leftarrow$ averaged basis function data (4.5);
18 $\theta_{\tau_j}^* \leftarrow$ solution of the utility identification (4.7);
19 **end**
20 (optional) $\{\bar{\theta}_\tau, \cdots, \bar{\theta}_{\tau+N-1}\} \leftarrow$ prediction of the basis weights from GP model (4.8);
21 **end**

on the other hand, the predicted basis weight vector, $\{\bar{\theta}_\tau, \cdots, \bar{\theta}_{\tau+N-1}\}$, which is the mean of the inferred value $\hat{\theta}_\tau$, is computed at each sampling time step (line 21) and used as the weights in the MPC objective function (4.2) as shown in Figure 4.4.

4.4 Application to personalized air conditioning

We consider the application of the proposed utility learning MPC scheme to the personalized thermostat that controls the customer's room temperature. In this scenario, the thermostat plays the role of the system manager.

4.4.1 Indoor temperature dynamics and utility functions

Let $x_\tau \in \mathbb{R}$ be the indoor room temperature at time $\tau\Delta t$, $\tau = 0, 1, \cdots$. We set $u_\tau \in \mathbb{R}$ to be the ratio between the duration in which the air conditioner (AC) is ON and the period Δt.

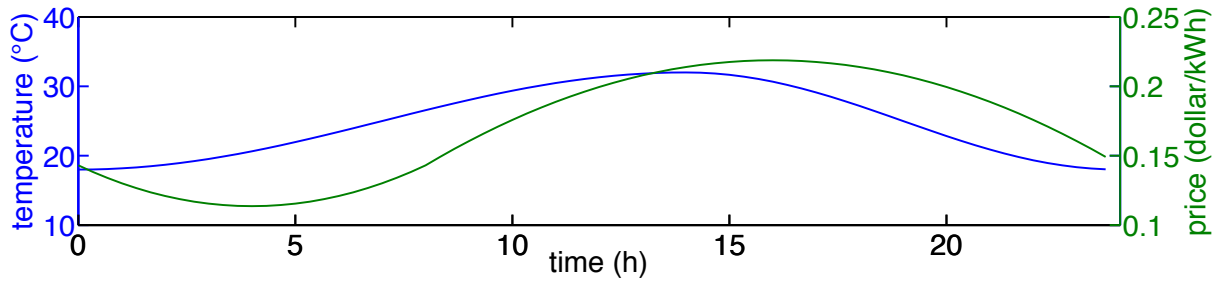

Figure 4.5: The profiles of outdoor temperature (blue) and time-of-use energy price (red).

By definition, $0 \leq u_\tau \leq 1$, i.e., $\mathcal{U} = [0, 1]$. Frequent ON/OFF switching of the AC is not desirable because (i) it can result in physical damage to the AC; and (ii) each switching ON of the AC generates a transient spike of power consumption, which is higher than steady state power consumption [4]. To avoid frequent switching, the time step is chosen as 20 minutes, i.e., $\Delta t = 1/3$ h in our numerical tests. Given the (forecasted) outdoor temperature $\Theta_\tau \in \mathbb{R}$, $\tau = 0, 1, \cdots$, the indoor temperature dynamics can be modeled as

$$x_{\tau+1} = x_\tau + [\alpha(\Theta_\tau - x_\tau) - \kappa u_\tau + z_\tau^o + w_\tau^o]\Delta t,$$

where $\alpha(\Theta_\tau - x_\tau)$ and $w_\tau^o \in \mathbb{R}$ model the temperature fluctuation due to the heat transfer from outside and the effect of forecast error, respectively. Here, $z_\tau^o \in \mathbb{R}$ represents an unknown effect from the customer's behavior on temperature, which will be learned. As explained in [120], $\alpha := R/C$, where R is the conductance between the outdoor air and indoor air and C is the conductance between the indoor air and the thermal mass. Furthermore, the positive constant κ depends on the efficiency of the AC.

We assume that the customer's objective for the AC operation is twofold: (i) to maximize the energy cost savings and (ii) to minimize the discomfort level that is affected by the indoor temperature. Therefore, we use the basis utility functions in Example 6.

The profiles of the time-of-use energy price c_τ and the outdoor temperature Θ_τ used in the simulation are shown in Figure 4.5. We assume that the forecast error $\bar{w}_\tau$ is distributed with $\mathcal{N}(0, 0.03^2)$ and that the customer reports her satisfaction every two hours. We choose $\alpha = 0.4834$, $\kappa = 2.5$, $\bar{x} = 22$, $\lambda = 10^{-10}$, $\theta_{\mathbf{L}}^1 = \theta_{\mathbf{L}}^2 = 0$, $\theta_{\mathbf{U}}^1 = \theta_{\mathbf{U}}^2 = 2$ and $\theta_0 = (1, 1)$. The hyperparameter estimation for the GP is performed every five steps, i.e., $L = 5$ using the GPML toolbox [109]. The MPC prediction horizon, N, is chosen as 20 and the number of data used in the utility learning module is set to $M = 2$.

4.4.2 Numerical tests

To test the performance of the proposed utility learning model predictive control, we consider the case in which the profiles of the ground truth basis weights in the customer's utility function are given by the step functions (blue) in Figure 4.6 (a) and (b). The weights

(red) θ_1^* and θ_2^* optimized by solving (4.7) well identify the ground truth values after 12h. However, their accuracy is poor at some points in which there is a step change in the ground truth values. This inaccurate identification undesirably generates sharp temperature peaks as shown in Figure 4.6 (c).

We then apply the GP learning module to learn and predict the basis weights, whose mean predictive equation is given by in (4.8). As shown in Figure 4.6 (a) and (b), it rejects the outliers produced by the optimization-based identification. This outlier rejection is due to the fact that our Gaussian process is designed to adapt a quasi-periodic change of the basis weights by setting the prior covariance function as the 'locally periodic' kernel given by (4.3). This result validates the usefulness of the sophisticatedly designed kernel in learning a quasi-periodic utility function. As a result, by using the predicted weights in the MPC module, we can remove the temperature peaks occurred by the outliers in the optimized weights as shown in Figure 4.6 (c). Further simulation results can be found in [140].

The predicted effect of the customer's behavior by the GP regression is shown in Figure 4.7. The true effect is shown together with the noisy samples and the mean prediction of the customer effect over the prediction horizon obtained from the GP model at every sampling time. After having collected (noisy) data of the human behavior for only about 24h, the GP prediction already provides a very good estimate of the customer effect. Furthermore, the prediction accuracy increases as time goes by.

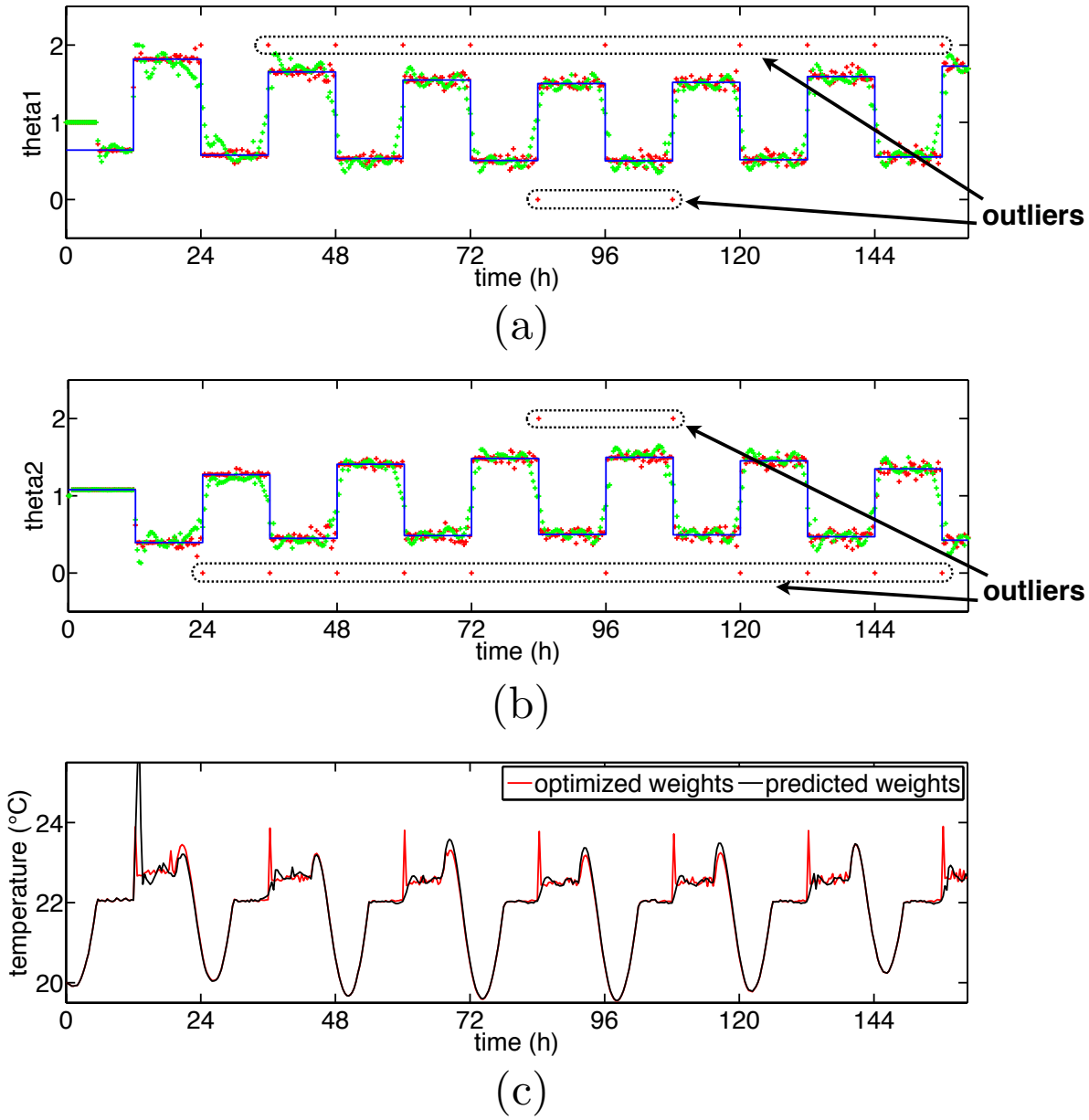

Figure 4.6: Ground truth basis weights in the customer's utility function (blue); optimized basis weights (red); and predicted (learned) basis weights (green) for (a) θ_1 and (b) θ_2. (c) the indoor temperature controlled by the optimized basis weights (black) and the predicted weights (red).

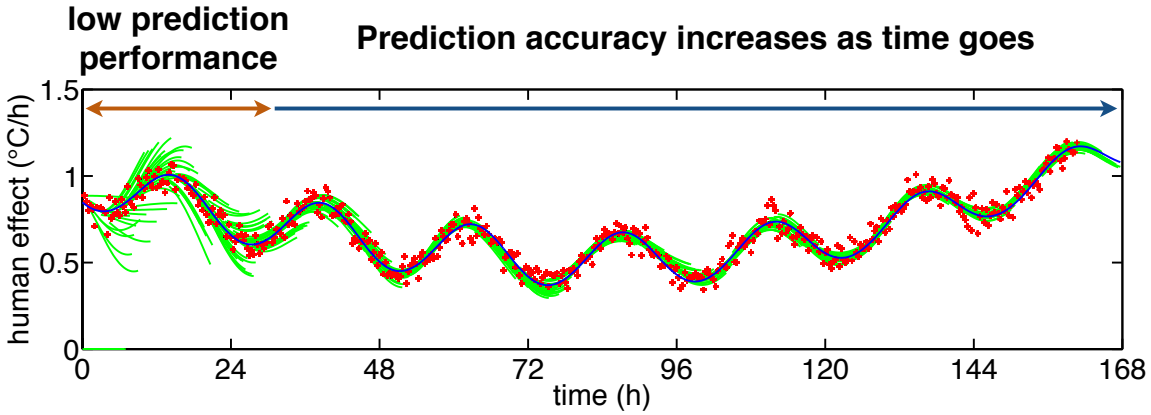

Figure 4.7: Predicted mean human effect (green), the ground truth (blue) and the noisy samples of the human effect (red) at each sampling time as used for MPC.

Chapter 5

Conclusion

5.1 Summary and contributions

This dissertation presented three new control mechanisms that can improve the sustainability of power and energy systems. These mechanisms also contribute to (dynamic) contract theory, stochastic control, combinatorial optimization and learning-based control.

The first proposed mechanism is a new continuous-time dynamic contract framework that has a risk-limiting capability. The key feature of the proposed contract is that the variance of the agent's payoff is bounded by a threshold specified in the contract. This feature enables the contract framework, by combining with direct load control, to provide financial risk management solutions for real-time electricity markets. To obtain a globally optimal contract, a dynamic programming-based method is developed. Difficulty arises in dealing with the constraints on the mean and the variance of the agent's payoff. We resolve this issue by developing a dynamical system approach to track and limit risks. We also proposed an approximate decomposition of the contract design problem for n agents into n low-dimensional problems for each agent. The approximate contract obtained using the proposed decomposition has a guaranteed suboptimality bound. This decomposability allows the direct load control program based on this contract framework to handle a large number of customers without any scalability issue. Using the data of the ERCOT LMPs and electric energy consumption by customers in Austin, Texas, we perform numerical experiments to validate the performance and usefulness of the proposed contracts. The numerical experiment results suggest that the proposed contract reduces the utility's financial risks by more than 58%.

The second tool is developed to optimize large-scale interdepedent systems in real time. We have proposed approximation algorithms for optimization of combinatorial dynamical systems, in which the decision variable is a binary vector and the cost is evaluated along the solution of the systems. The key idea of the approximation is to linearize the objective function using its derivative, which is well-defined in the feasible space of the binary decision variable. We proposed two different variation methods to define such derivatives. The

approximate problem has three major advantages: (*i*) the approximate problem is a 0–1 linear program and, therefore, can be exactly or approximately solved by polynomial time algorithms with suboptimality bounds; (*ii*) it does not require us to repeatedly solve the dynamical system; and (*iii*) its solution has a provable suboptimality bound under certain concavity conditions. In our numerical experiments in direct load control, the suboptimality bound is greater than 74% though in practice the performance of the proposed approximation algorithm is greater than 90% of the oracle's performance.

The third proposed mechanism is a novel personalized control framework, called the utility learning model predictive control. It integrates online learning of the customer's utility into the control of the customer's electric loads taking into account real-time updates of the utility. The proposed method is suitable to automatically customize the controller when the customer's utility associated with the controlled system changes over time. The numerical experiment results suggest the proposed Gaussian process regression approach with a locally periodic kernel is effective to learn a quasi-periodically changing utility function by rejecting outliers occurred in a convex optimization-based approach.

We believe that these control mechanisms can play a critical role in the seamless transition from the centralized and vertically integrated legacy power grid to a decentralized and flexible future grid in which distributed renewable energy sources and electric loads are synergistically coordinated with sensing, computing and communication infrastructures.

5.2 Future work

5.2.1 Unified risk management frameworks for electricity markets

To simultaneously manage risks in the real-time markets for multiple distribution systems (nodes), we must properly take into account the coupling between the locational marginal prices (LMPs) at the nodes: the price volatility at one node can influence and be influenced by those of other nodes. The interdependent LMP dynamics increase the size of contract design problems and hinder us from employing dynamic programming if the dimension of the system state is large. To overcome this scalability issue, we will pursue an approach using a grid-free Monte Carlo method for evaluating the value function of stochastic optimal control, combined with an efficient sampling technique called *implicit sampling* [143]. To build a unified risk management framework for day-ahead and real-time electricity markets, we will develop optimal dispatch and unit commitment methods for the aforementioned contract frameworks jointly with system operators. To this end, we will rigorously project the long-term impact of the proposed frameworks on system operators and electricity markets through risk analytics using societal-scale data. We eventually aim to implement these risk management tools to facilitate the integration of renewable energy sources and demand-side resources into the electric grid, which in turn contributes to improve power system sustainability.

5.2.2 Foundational optimization and control tools for resilient infrastructures

When undesirable and malicious events occur, resilient infrastructures should recover acceptable performance levels given a limited amount of recovery resources. This resource constraint motivates *selective control*, which optimally targets the points of control where the resources are invested. This new automatic control algorithm unifies combinatorial optimization and optimal control theory. For example, such a method can optimally select and control flexible alternating current transmission systems and loads to mitigate cascading failures, such as blackouts, in a power transmission system and to recover from them as quickly as possible. The performance of such a selective control strategy can be enhanced by allowing the selection to change over time. We will explore theoretic and algorithmic developments for this advanced selective control method with application to *dynamic load/customer targeting* for services to the electric grid in the event of supply shortages, cyber-attacks and electricity market manipulations.

To handle large-scale infrastructures, such as the electric grid, we will propose a novel *contract-based modular management* framework utilizing the contract methods in Chapter 2. This contract-based incentive mechanism will enhance the resilience of multi-scale infrastructure networks in both reactive and proactive ways. Successful implementation of the incentive mechanism will be supported by my research program on the estimation of utility functions in Chapter 4. We expect that this modular approach will play a critical role in the operation of micro-grids. It seeks a mechanism such that each micro-grid operates with a desired performance level which is measured by reliability, resilience and economic efficiency while each micro-grid operation is beneficial to the system operator. To make a societal impact that contributes to achieving resilient and sustainable power systems, I aim to implement the aforementioned control and optimization tools and operational strategies into the electric grid and demand-side resources in collaboration with utility companies and system operators.

5.3 Concluding remark: towards sustainable cyber-physical systems interfaced with human decision-makers

The control, computing, sensing and communication capabilities of cyber-physical systems (CPS) enable us to engage human agents in the system operations as shown in the example of automated demand response for power systems. Many modern infrastructure systems are CPS including power grids, home energy management systems, water grids, transportation networks and health care systems. As sensor and communication networks are pervasive in physical infrastructures that affect our daily lives, the sustainable interaction between CPS and human agents are becoming more important than ever before. On one hand, we can

customize the system operation for the human agents by learning their behaviors and preferences. On the other hand, we can incentivize the human agents to cooperate for the system operation. The control and optimization tools produced in my future research programs will enable the sustainable feedback loop between human agents and CPS. Therefore, the control mechanisms will contribute to sustainable CPS which can be interfaced with human decision-makers.

Appendix A

The ϵ-variational systems

For given $\bar{\alpha}, \alpha \in \{0,1\}^m$, the ϵ-variational system associated with $(\bar{\alpha}, \alpha)$ is defined as (3.7), where its vector field is given by the convex combination of the two vector fields with $\bar{\alpha}$ and α. The state trajectory of (3.7) is unique for given $\bar{\alpha}, \alpha \in \{0,1\}^m$ and is bounded on a finite time interval by Theorem 5.5 in [8], as shown in the following lemma.

Lemma 2. *Suppose that Assumption 1 holds. For any $\epsilon \in [0,1]$ and any $\bar{\alpha}, \alpha \in \{0,1\}^m$, the ϵ-variational system (3.7) associated with $(\bar{\alpha}, \alpha)$ admits a unique solution, $x^{\epsilon(\bar{\alpha},\alpha)}$. In addition, $\|x^{\epsilon(\bar{\alpha},\alpha)}(t)\|$ is bounded by some constant independent of ϵ for all $t \in [0,T]$.*

To use the ϵ-variational system for defining the derivative of the payoff function, it is important to address how the ϵ-variational system behaves as ϵ tends to zero compared to the original dynamical system. The following lemma shows that the difference $x^{\epsilon(\bar{\alpha},\alpha)}(t) - x^{\alpha}(t)$ is Lipschitz continuous in $\epsilon \in [0,1]$ for any $\bar{\alpha}, \alpha \in \{0,1\}^m$ (e.g., Lemma 5.6.7 in [106]).

Lemma 3. *Suppose that Assumption 1 holds. For any $\epsilon \in [0,1]$ and any $\bar{\alpha}, \alpha \in \{0,1\}^m$, there exists a constant L independent of ϵ such that for all $t \in [0,T]$*

$$\|x^{\epsilon(\bar{\alpha},\alpha)}(t) - x^{\bar{\alpha}}(t)\| \leq L\epsilon.$$

Combining the two lemmas and the dominated convergence theorem (e.g., [114]), we have the following corollary.

Corollary 4. *Suppose that Assumption 1 holds. For any $\bar{\alpha}, \alpha \in \{0,1\}^m$, the following equality holds:*

$$\lim_{\epsilon \to 0^+} \frac{1}{\epsilon} \int_0^T \|x^{\epsilon(\bar{\alpha},\alpha)}(t) - x^{\bar{\alpha}}(t)\|^2 dt = 0.$$

This corollary is essential to show that our proposed nonstandard derivative is well-defined and can be computed using an adjoint-based formula.

Appendix B

Proof of Theorem 3

We will show a more general equality,

$$\lim_{\epsilon\to 0^+}\frac{1}{\epsilon}\left[\mathcal{J}^{\epsilon(\bar{\alpha},\alpha)}(x^{\epsilon(\bar{\alpha},\alpha)})-\mathcal{J}(x^{\bar{\alpha}},\bar{\alpha})\right]= \int_0^T \left(f(x^{\bar{\alpha}}(t),\alpha)-f(x^{\bar{\alpha}}(t),\bar{\alpha})\right)^\top \lambda^{\bar{\alpha}}(t)+r(x^{\bar{\alpha}}(t),\alpha)-r(x^{\bar{\alpha}}(t),\bar{\alpha})dt. \tag{B.1}$$

Substituting $\alpha=\bar{\alpha}+\mathbf{1}_i$ and $\alpha=\bar{\alpha}-\mathbf{1}_i$ into the above equality, we obtain the formulae in Theorem 3 for $\bar{\alpha}_i=0$ and $\bar{\alpha}_i=1$, respectively.

Proof. Fix $\bar{\alpha},\alpha\in\{0,1\}^m$. For notational simplicity, we let $\hat{x}(\cdot):=x^{\epsilon(\bar{\alpha},\alpha)}(\cdot)-x^{\bar{\alpha}}(\cdot)$. Then, it satisfies the following ODE:

$$\dot{\hat{x}}(t)=f(\hat{x}(t)+x^{\bar{\alpha}}(t),\bar{\alpha})-f(x^{\bar{\alpha}}(t),\bar{\alpha})+\epsilon(f(\hat{x}(t)+x^{\bar{\alpha}}(t),\alpha)-f(\hat{x}(t)+x^{\bar{\alpha}}(t),\bar{\alpha}))$$

with $\hat{x}(0)=0$. The dynamical system can be rewritten as

$$\dot{\hat{x}}(t)=\frac{\partial f(x^{\bar{\alpha}}(t),\bar{\alpha})}{\partial \boldsymbol{x}}\hat{x}(t)+\epsilon(f(\hat{x}(t)+x^{\bar{\alpha}}(t),\alpha)-f(\hat{x}(t)+x^{\bar{\alpha}}(t),\bar{\alpha}))+\sigma(\hat{x}(t),x^{\bar{\alpha}}(t)), \tag{B.2}$$

where $\sigma:=(\sigma_1,\cdots,\sigma_n)$ is given by

$$\sigma_i(\hat{\boldsymbol{x}},\boldsymbol{x}):=\mathbf{H}_i(\hat{\boldsymbol{x}},\boldsymbol{x},\bar{\alpha})+\epsilon\left(\frac{\partial f_i(\boldsymbol{x},\alpha)}{\partial \boldsymbol{x}}-\frac{\partial f_i(\boldsymbol{x},\bar{\alpha})}{\partial \boldsymbol{x}}\right)\hat{\boldsymbol{x}}+\epsilon(\mathbf{H}_i(\hat{\boldsymbol{x}},\boldsymbol{x},\alpha)-\mathbf{H}_i(\hat{\boldsymbol{x}},\boldsymbol{x},\bar{\alpha}))$$

and $\mathbf{H}_i(\hat{\boldsymbol{x}},\boldsymbol{x},\alpha)$ denotes the higher-order terms in the Taylor expansion of $f_i(\hat{\boldsymbol{x}}+\boldsymbol{x},\alpha)$ at $\boldsymbol{x}$, i.e., by applying the mean value theorem, $\mathbf{H}_i(\hat{\boldsymbol{x}},\boldsymbol{x},\alpha):=\int_0^1(1-s)\hat{\boldsymbol{x}}^\top D^2_{\boldsymbol{x}}f_i(\boldsymbol{x}+s\hat{\boldsymbol{x}},\alpha)\hat{\boldsymbol{x}}ds$. Due to Lemma 3 or Corollary 4, for all $t\in[0,T]$,

$$\lim_{\epsilon\to 0^+}\frac{1}{\epsilon}\sigma_i(\hat{x}(t),x^{\bar{\alpha}}(t))=0.$$

We now consider the difference

$$\begin{aligned}&\mathcal{J}^{\epsilon(\bar{\alpha},\alpha)}(x^{\epsilon(\bar{\alpha},\alpha)}) - \mathcal{J}(x^{\bar{\alpha}},\bar{\alpha})\\ &= \int_0^T r(x^{\epsilon(\bar{\alpha},\alpha)}(t),\bar{\alpha}) - r(x^{\bar{\alpha}}(t),\bar{\alpha}) + \epsilon(r(x^{\epsilon(\bar{\alpha},\alpha)}(t),\alpha) - r(x^{\epsilon(\bar{\alpha},\alpha)}(t),\bar{\alpha}))dt\\ &\quad + q(x^{\epsilon(\bar{\alpha},\alpha)}(T)) - q(x^{\bar{\alpha}}(T)).\end{aligned}$$

The difference can be rewritten as

$$\begin{aligned}&\mathcal{J}^{\epsilon(\bar{\alpha},\alpha)}(x^{\epsilon(\bar{\alpha},\alpha)}) - \mathcal{J}(x^{\bar{\alpha}},\bar{\alpha})\\ &= \int_0^T \frac{\partial r(x^{\bar{\alpha}}(t),\bar{\alpha})}{\partial \boldsymbol{x}}\hat{x}(t) + \epsilon(r(x^{\bar{\alpha}}(t),\alpha) - r(x^{\bar{\alpha}}(t),\bar{\alpha}))dt + \frac{\partial q(x^{\bar{\alpha}}(T))}{\partial \boldsymbol{x}}\hat{x}(T) + \eta(\hat{x},x^{\bar{\alpha}}),\end{aligned}$$

where

$$\begin{aligned}\eta(\hat{x},x) := \int_0^T &\mathbf{I}(\hat{x}(t),x(t),\bar{\alpha}) + \epsilon\left(\frac{\partial r(x(t),\alpha)}{\partial \boldsymbol{x}} - \frac{\partial r(x(t),\bar{\alpha})}{\partial \boldsymbol{x}}\right)\hat{x}(t)\\ &+ \epsilon(\mathbf{I}(\hat{x}(t),x(t),\alpha) - \mathbf{I}(\hat{x}(t),x(t),\bar{\alpha}))dt + \mathbf{J}(\hat{x}(T),x(T))\end{aligned}$$

and $\mathbf{I}(\hat{\boldsymbol{x}},\boldsymbol{x},\alpha)$ and $\mathbf{J}(\hat{\boldsymbol{x}},\boldsymbol{x})$ denote the higher-order terms in the Taylor expansions of $r(\hat{\boldsymbol{x}}+\boldsymbol{x},\alpha)$ and $q(\hat{\boldsymbol{x}}+\boldsymbol{x})$ at $\boldsymbol{x}$, respectively, i.e., $\mathbf{I}(\hat{\boldsymbol{x}},\boldsymbol{x},\alpha) := \int_0^1 (1-s)\hat{\boldsymbol{x}}^\top D^2_{\boldsymbol{x}} r(\boldsymbol{x}+s\hat{\boldsymbol{x}},\alpha)\hat{\boldsymbol{x}}ds$ and $\mathbf{J}(\hat{\boldsymbol{x}},\boldsymbol{x}) := \int_0^1 (1-s)\hat{\boldsymbol{x}}^\top D^2_{\boldsymbol{x}} q(\boldsymbol{x}+s\hat{\boldsymbol{x}})\hat{\boldsymbol{x}}ds$. Due to Lemma 3 or Corollary 4, we have

$$\lim_{\epsilon\to 0^+}\frac{1}{\epsilon}\eta(\hat{x},x^{\bar{\alpha}}) = 0.$$

Adding the inner product between the adjoint state and the system (B.2), which is zero, to the difference, we have

$$\begin{aligned}&\mathcal{J}^{\epsilon(\bar{\alpha},\alpha)}(x^{\epsilon(\bar{\alpha},\alpha)}) - \mathcal{J}(x^{\bar{\alpha}},\bar{\alpha}) =\\ &\int_0^T \left(\frac{\partial r(x^{\bar{\alpha}}(t),\bar{\alpha})}{\partial \boldsymbol{x}}\hat{x}(t) + \epsilon(r(x^{\bar{\alpha}}(t),\alpha) - r(x^{\bar{\alpha}}(t),\bar{\alpha}))\right)dt + \frac{\partial q(x^{\alpha}(T))}{\partial \boldsymbol{x}}\hat{x}(T)\\ &+ \int_0^T (\lambda^{\bar{\alpha}}(t))^\top\left(-\dot{\hat{x}}(t) + \frac{\partial f(x^{\bar{\alpha}}(t),\bar{\alpha})}{\partial \boldsymbol{x}}\hat{x}(t) + \epsilon(f(\hat{x}(t)+x^{\bar{\alpha}}(t),\alpha) - f(\hat{x}(t)+x^{\bar{\alpha}}(t),\bar{\alpha}))\right)dt\\ &+ \Theta(\hat{x},x^{\bar{\alpha}}),\end{aligned} \tag{B.3}$$

where $\Theta(\hat{x},x^{\bar{\alpha}}) := \int_0^T (\lambda^{\bar{\alpha}})^\top \sigma(\hat{x}(t),x^{\bar{\alpha}}(t))dt + \eta(\hat{x},x^{\bar{\alpha}})$. Using integration by parts, we have

$$\begin{aligned}\int_0^T (\lambda^{\bar{\alpha}})^\top\dot{\hat{x}}dt &= \lambda^{\bar{\alpha}}(T)^\top\hat{x}(T) - \lambda^{\bar{\alpha}}(0)^\top\hat{x}(0) - \int_0^T(\dot{\lambda}^{\bar{\alpha}}(t))^\top\hat{x}(t)dt\\ &= \frac{\partial q(x^{\bar{\alpha}}(T))}{\partial \boldsymbol{x}}\hat{x}(T) - \int_0^T(\dot{\lambda}^{\bar{\alpha}}(t))^\top\hat{x}(t)dt.\end{aligned} \tag{B.4}$$

Combining (B.3) and (B.4), we obtain

$$\begin{aligned}
&\mathcal{J}^{\epsilon(\bar{\alpha},\alpha)}(x^{\epsilon(\bar{\alpha},\alpha)}) - \mathcal{J}(x^{\bar{\alpha}},\bar{\alpha}) \\
&= \int_0^T \left((\lambda^{\bar{\alpha}}(t))^\top \frac{\partial f(x^{\bar{\alpha}}(t),\bar{\alpha})}{\partial \boldsymbol{x}} + \frac{\partial r(x^{\bar{\alpha}}(t),\bar{\alpha})}{\partial \boldsymbol{x}} + (\dot{\lambda}^{\bar{\alpha}}(t))^\top \right) \hat{x}(t) dt \\
&+ \epsilon \int_0^T r(x^{\bar{\alpha}}(t),\alpha) - r(x^{\bar{\alpha}}(t),\bar{\alpha}) + (\lambda^{\bar{\alpha}}(t))^\top (f(x^{\bar{\alpha}}(t),\alpha) - f(x^{\bar{\alpha}}(t),\bar{\alpha})) dt \\
&+ \Theta(\hat{x}, x^{\bar{\alpha}}),
\end{aligned}$$

where the first integral term on the right-hand side is equal to zero due to the definition of the adjoint system (3.6). Since

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon} \Theta(\hat{x}, x^{\bar{\alpha}}) = 0,$$

we obtain (B.1) as desired.

The existence and the uniqueness of the state $x^{\bar{\alpha}}(t)$ and the adjoint state $\lambda^{\bar{\alpha}}(t)$ guarantee the existence and uniqueness of the nonstandard derivative. Furthermore, the boundedness of $x^{\bar{\alpha}}(t)$ and $\lambda^{\bar{\alpha}}(t)$ for $t \in [0, T]$ imply that the nonstandard derivative is bounded. □

Appendix C

Comparison of standard and nonstandard derivatives

We first characterize a condition under which both derivatives are the same.

Proposition 5. *Suppose that Assumptions 1, 2, 3 and 4 hold. If $f(\boldsymbol{x}, \cdot) : \mathbb{R}^m \to \mathbb{R}^n$ and $r(\boldsymbol{x}, \cdot) : \mathbb{R}^m \to \mathbb{R}$ are affine functions, then the two derivatives, $D^S J$ and $D^{NS} J$, are equivalent to each other.*

Proof. Since $f(\boldsymbol{x}, \cdot)$ and $r(\boldsymbol{x}, \cdot)$ are differentiable and affine, we have $\frac{\partial f(\boldsymbol{x},\alpha)}{\partial \boldsymbol{\alpha}_i} = f(\boldsymbol{x}, \mathbf{1}_i)$. A similar inequality holds for r. Comparing the adjoint-based formulae for $D^{\mathrm{S}} J$ and $D^{\mathrm{NS}} J$ in Proposition 2 and Theorem 3, respectively, with the assumption that $f(\boldsymbol{x}, \alpha)$ and $r(\boldsymbol{x}, \alpha)$ are affine in α, we deduce that the two derivatives are equivalent to each other. □

In general, $D^{\mathrm{S}} J$ and $D^{\mathrm{NS}} J$ are different from each other because they use different variation methods in their definitions. We present a concrete example in which the standard derivative is different the nonstandard derivative.

Example 7. *Suppose that $n = 1$, $m > 1$, $f(\boldsymbol{x}, \alpha) = \boldsymbol{x} + \sum_{i=1}^{m} e^{-\alpha_i}$, $r(\boldsymbol{x}, \alpha) = \boldsymbol{x}$ and the terminal payoff q is set to be zero. Note that the vector field is not affine but additive in α. Then, the standard and nonstandard derivatives are given by*

$$[D^S J(\bar{\alpha})]_i = -\int_0^T \lambda^{\bar{\alpha}}(t) e^{-\bar{\alpha}_i} dt$$
$$[D^{NS} J(\bar{\alpha})]_i = \int_0^T \lambda^{\bar{\alpha}}(t)(e^{-1} - e^0) dt,$$

respectively. We notice that $[D^S J(\bar{\alpha})]_i$ and $[D^{NS} J(\bar{\alpha})]_i$ are not equal to each other.

C.1 Differentiability issue

Recall that the standard derivative $D^{\mathrm{S}}J$ requires the differentiability of $f(\boldsymbol{x}, \cdot\,)$ and $r(\boldsymbol{x}, \cdot\,)$, which can be restrictive in many applications. One can reformulate $f(\boldsymbol{x}, \cdot\,)$ and $r(\boldsymbol{x}, \cdot\,)$ as the following polynomials in α using the multi-linear polynomial extension:

$$\tilde{f}(\boldsymbol{x}, \alpha) = \sum_{V \subseteq \Omega} f(\boldsymbol{x}, \mathbb{I}(V)) \prod_{i \in V} \alpha_i \prod_{i \in \Omega \setminus V} (1 - \alpha_i),$$
$$\tilde{r}(\boldsymbol{x}, \alpha) = \sum_{V \subseteq \Omega} r(\boldsymbol{x}, \mathbb{I}(V)) \prod_{i \in V} \alpha_i \prod_{i \in \Omega \setminus V} (1 - \alpha_i),$$

where $\Omega := \{1, \cdots, m\}$ and $\mathbb{I} : 2^{\Omega} \to \{0, 1\}^m$ is the set indicator function. However, each of these representations requires 2^m calculations in the worst case. Therefore, it is not computationally tractable to construct the multi-linear polynomial representations of f and r.

The nonstandard derivative $D^{\mathrm{NS}}J$ is a good alternative to resolve this differentiability issue. Note that this nonstandard derivative fully takes advantage of the fact that the problem is associated with a dynamical system: the construction of the nonstandard derivative is possible because we are able to utilize the vector field of the dynamical system as a relaxation tool. This convex combination approach for vector fields and running payoffs naturally resolves the differentiability issue.

C.2 Performance comparison

As suggested in Section 3.3, we solve the approximate problem (3.4) twice: once using the standard derivative $D^{\mathrm{S}}J$ and again using the nonstandard derivative $D^{\mathrm{NS}}J$. Between the two approximate solutions, the solution that gives a larger payoff is chosen. Despite this practical advantage of using the two derivative concepts, it is still valuable to have an insight on the comparison of their effects on the proposed approximation. We consider a simple example, where $n = 1$, $m = 2$, and the vector field and the running payoff are given by

$$f(\boldsymbol{x}, \alpha) = \boldsymbol{x} + \alpha_1^3 + 2\alpha_2, \quad r(\boldsymbol{x}, \alpha) = \boldsymbol{x}^2$$

and the terminal payoff is set to be zero. The solutions of the primal and adjoint systems are given by $x^{\alpha}(t) = e^t \mathbf{x} + \int_0^t e^{t-\tau}(\alpha_1^3 + 2\alpha_2)d\tau$ and $\lambda^{\alpha}(t) = \int_0^{T-t} e^{T-t-\tau} 2x^{\alpha}(T - t - \tau)d\tau$, respectively. Suppose that the initial value $\mathbf{x}$ is positive. Then, $x^{\alpha}(t)$ is positive for any $t \in [0, T]$ and for any $\alpha \in \{0, 1\}^2$ and, therefore, so $\lambda^{\alpha}(t)$ is. The adjoint-based formulae in Proposition 2 and Theorem 3 for the two derivatives imply that

$$D^{\mathrm{S}}J(\bar{\alpha}) = \int_0^T \begin{bmatrix} 3\bar{\alpha}_1 \\ 2 \end{bmatrix} \lambda^{\bar{\alpha}}(t)dt,$$
$$D^{\mathrm{NS}}J(\bar{\alpha}) = \int_0^T \begin{bmatrix} 1 \\ 2 \end{bmatrix} \lambda^{\bar{\alpha}}(t)dt.$$

Suppose that the constraint $\|\alpha\|_0 \leq 1$ is imposed. In this case, the optimal solution is $(0, 1)$. If we linearize the optimization problem at $\bar{\alpha} = (1, 1)$, then the approximate solution based on $DJ(\bar{\alpha})$ is $(1, 0)$, while that based on $\hat{D}J(\bar{\alpha})$ is $(0, 1)$, which corresponds to the optimal solution. The reason why the first derivative gives a wrong solution is that it introduces a 'bias' in its first entry due to the cubic term in α_1. Here, we do not overstate that the second derivative performs better than the first because this bias might help find an optimal solution in other cases. Nevertheless, it is worth noting that the second derivative does not introduce this bias. We believe that this observation can stimulate further research on the performance comparison of the two derivatives and theoretic investigation on this bias in the future.

Bibliography

[1] U. M. Ascher and L. Petzold. *Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations*. Philadelphia, PA: SIAM, 1998.

[2] M. F. Astaneh and Z. Chen. "Price volatility in wind dominant electricity markets". In: *Proceedings of 2013 IEEE EUROCON*. 2013, pp. 770–775.

[3] A. Aswani et al. "Provably safe and robust learning-based model predictive control". In: *Automatica* 49 (2013), pp. 1216–1226.

[4] A. Aswani et al. "Reducing Transient and Steady State Electricity Consumption in HVAC Using Learning-Based Model-Predictive Control". In: *Proc. of the IEEE* 100.1 (2012), pp. 240–253.

[5] Austin Energy. *Residential Electric Rates and Line Items*. URL: `http://www.austinenergy.com/wps/portal/ae/residential/rates`.

[6] E. Balas and E. Zemel. "An algorighm for large zero-one knapsack problems". In: *Operations Research* 28.5 (1980), pp. 1130–1154.

[7] L. Balling. "Fast cycing and rapid start-up: new generation of plants achieves impressive results". In: *Modern Power Systems* (2011), pp. 35–41.

[8] M. Bardi and I. Capuzzo-Dolcetta. *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman equations*. Boston, MA: Birkhäuser, 1997.

[9] G. Barles and P. E. Souganidis. "Convergence of approximation schemes for fully nonlinear second order equations". In: *Asymptotic Analysis* 4.3 (1991), pp. 271–283.

[10] R. Bellman. "Dynamic programming and Lagrange multipliers". In: *Proceedings of the National Academy of Sciences* 42.10 (1956), pp. 767–769.

[11] A. Bemporad and M. Morari. "Control of systems integrating logic, dynamics, and constraints". In: *Automatica* 35 (1999), pp. 407–427.

[12] D. P. Bertsekas. *Dynamic Programming and Optimal Control*. 3rd. Belmont, MA: Athena Scientific, 2005.

[13] D. Bertsimas, V. Gupta, and I. Ch. Paschalidis. "Data-driven estimation in equilibrium using inverse optimization". In: *Mathematical Programming* (2014).

[14] D. Bertsimas et al. "Adaptive robust optimization for the security constrained unit commitment problem". In: *IEEE Transactions on Power Systems* 28.1 (2013), pp. 52–63.

[15] E. Bitar and S. Low. "Deadline differentiated pricing of deferrable electric power service". In: *Proceedings of the 51st IEEE Conference on Decision and Control*. 2012, pp. 4991–4997.

[16] B. Bouchard, R. Elie, and C. Imbert. "Optimal control under stochastic target constraints". In: *SIAM Journal on Control and Optimization* 48.5 (2010), pp. 3501–3531.

[17] F. Bouffard and F. D. Galiana. "Stochastic security for operations planning with significant wind power generation". In: *IEEE Transactions on Power Systems* 23.2 (2008), pp. 306–316.

[18] M. S. Branicky, V. S. Borkar, and S. K. Mitter. "A unified framework for hybrid control: model and optimal control theory". In: *IEEE Transactions on Automatic Control* 43.1 (1998), pp. 31–45.

[19] A. Cadenillas, J. Cvitanić, and F. Zapatero. "Optimal risk-sharing with effort and project choice". In: *Journal of Economic Theory* 133 (2007), pp. 403–440.

[20] A. Cadenillas and I. Karatzas. "The stochastic maximum principle for linear, convex optimal control with random coefficients". In: *SIAM Journal on Control and Optimization* 33.2 (1995), pp. 590–624.

[21] California Institute for Energy and Environment and Power Standards Laboratory. *Micro-Synchrophasors for Distribution Systems*. URL: `http://www.pqubepmu.com`.

[22] California Public Utilities Commission. *California Renewables Portfolio Standard (RPS)*. URL: `http://www.cpuc.ca.gov/PUC/energy/Renewables/`.

[23] D. S. Callaway and I. A. Hiskens. "Achieving controllability of electric loads". In: *Proc. of the IEEE* 99.1 (2011), pp. 184–199.

[24] H.-P. Chao and R. Wilson. "Priority service: pricing, investment, and market organization". In: *The American Economic Reivew* 77.5 (1987), pp. 899–916.

[25] A. Clark, L. Bushnell, and R. Poovendran. "A supermodular optimization framework for leader selection under link noise in linear multi-agent systems". In: *IEEE Transactions on Automatic Control* 59.2 (2014), pp. 283–296.

[26] J. C. Cox and C.-F. Huang. "Optimal consumption and portfolio policies when asset prices follow a diffusion process". In: *Journal of Economic Theory* 49 (1989), pp. 33–83.

[27] M. G Crandall, H. Ishii, and P.-L. Lions. "User's guide to viscosity solutions of second order partial differential equations". In: *Bulletin of the American Mathematical Society* 27.1–67 (1992).

[28] J. Cvitanić and J. Zhang. *Contract Theory in Continuous-Time Models*. Berlin; New York: Springer, 2012.

[29] G. Dantzig. *Linear Programming and Extensions*. Princeton, NJ: Princeton University Press, 1998.

[30] G. B. Dantzig. "Discrete-variable extremum problems". In: *Operations Research* 5.2 (1957), pp. 266–277.

[31] S. J. Deng. "Stochastic models of energy commodity prices and their applications: mean-reversion with jumps and spikes". In: *POWER working paper, University of California Energy Institute, Berkeley* (1999).

[32] S.-J. Deng, B. Johnson, and A. Sogomonian. "Exotic electricity options and the valuation of electricity generation and transmission assets". In: *Decision Support Systems* 30 (2001), pp. 383–392.

[33] S. J. Deng and S. S. Oren. "Electricity derivatives and risk management". In: *Energy* 31 (2006), pp. 940–953.

[34] Electric Power Research Institute. "Automated Demand Response Today". In: (2012).

[35] Electric Reliability Council of Texas. *Real-time prices reports*. URL: `http://www.ercot.com/mktinfo`.

[36] M. S. Elliott and B. P. Rasmussen. "Model-Based Predictive Control of a multi-evaporator vapor compression cooling cycle". In: *Proceedings of 2008 American Control Conference*. 2008, pp. 1463–1468.

[37] H. Farhangi. "The Path of the Smart Grid". In: *IEEE Power & Energy Magazine* (2010), pp. 18–28.

[38] W. H. Fleming and W. M. McEneaney. "Risk-sensitive control on an infinite time horizon". In: *SIAM Journal on Control and Optimization* 33.6 (1995), pp. 1881–1915.

[39] W. H. Fleming and H. M. Soner. *Controlled Markov Processes and Viscosity Solutions*. 2nd. Springer-Verlag, 2006.

[40] A. Fréville. "The multidimensional 0-1 knapsack problem". In: *European Journal of Operational Research* 155 (2004), pp. 1–21.

[41] L. Gan et al. "Real-time deferable load control: handling the uncertainties of renewable generation". In: *Proceedings of the 4th International Concerence on Future Energy Systems*. 2013, pp. 113–124.

[42] F. Glover and E. Woolsey. "Converting the 0–1 polynomial programming problem to a 0–1 linear program". In: *Operations Research* 22.1 (1974), pp. 180–182.

[43] G. Goel et al. "Approximability of combinatorial problems with multi-agent submodular cost functions". In: *Proceedings of the 50th Annual IEEE symposium on Foundations of Computer Science*. 2009, pp. 755–764.

[44] M. Gomez-Rodriguez, J. Leskovec, and A. Krause. "Inferring networks of diffusion and influence". In: *ACM Transactions on Knowledge Discovery from Data* 5.4 (2012), 21:1–21:37.

[45] A. Grancharova, J. Kocijan, and T. A. Johansen. "Explicit stochastic predictive control of combustion plants based on Gaussian process models". In: *Automatica* 44 (2008), pp. 1621–1631.

[46] GridLAB-D. *Residential Module User's Guide*. URL: `http://gridlabd.me.uvic.ca/gridlabd/wiki/index.php/`.

[47] M. Grötschel, L. Lovász, and A. Schrijver. "The ellipsoid method and its consequences in combinatorial optimization". In: *Combinatorica* 1.2 (1981), pp. 169–197.

[48] P. L. Hammer, P. Hansen, and B. Simeone. "Roof duality, complementation and persistency in quadratic 0–1 optimization". In: *Mathematical Programming* 28 (1984), pp. 121–155.

[49] P. L. Hammer and S. Rudeanu. *Boolean Methods in Operations Research*. New York: Springer-Verlag, 1968.

[50] H. Hao et al. "A generalized battery model of a collection of thermostatically controlled loads for providing ancillary service". In: *Proceedings of the 51st Annual Allerton Conference on Communication, Control and Computing*. 2013, pp. 551–558.

[51] H. Hao et al. "Aggregate flexibility of thermostatically controlled loads". In: *IEEE Transactions on Power Systems* (to appear).

[52] C. Helmberg, F. Rendl, and R. Weismantel. "A semidefinite programming approach to the quadratic knapsack problem". In: *Journal of Combinatorial Optimization* 4 (2000), pp. 197–215.

[53] B. Holmstrom and P. Milgrom. "Aggregation and linearity in the provision of intertemporal incentives". In: *Econometrica* 55.2 (1987), pp. 303–328.

[54] E. Horowitz and S. Sahni. "Computing partitions with applications to the knapsack problem". In: *Journal of the Association for Computing Machinery* 21.2 (1974), pp. 277–292.

[55] T. G. Hovgaard et al. "Nonconvex model predictive control for commercial refrigeration". In: *International Journal of Control* 86.8 (2013), pp. 1349–1366.

[56] *https://nest.com/*.

[57] O. H. Ibarra and C. E. Kim. "Fast approximation algorithms for the knapsack and sum of subset problems". In: *Journal of the Association for Computing Machinery* 22.4 (1975), pp. 463–468.

[58] S. Iwata, L. Fleischer, and S. Fujishige. "A combinatorial strongly polynomial algorighm for minimizing submodular functions". In: *Journal of the Association for Computing Machinery* 48.4 (2001), pp. 761–777.

[59] S. Iwata and K. Nagano. "Submodular function minimization under convering constraints". In: *Proceedings of the 50th Annual IEEE symposium on Foundations of Computer Science*. 2009, pp. 671–680.

[60] D. H. Jacobson. "Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games". In: *IEEE Transactions on Automatic Control* AC-18.2 (1973), pp. 124–131.

[61] S. Jegelka and J. Bilmes. "Approximation bounds for inference using cooperative cuts". In: *Proceedings of the 28th International Conference on Machine Learning*. 2011.

[62] R. E. Kalman and R. S. Bucy. "New results in linear filtering and prediction theory". In: *Journal of Fluids Engineering* 83.1 (1961), pp. 95–108.

[63] R. Kamat and S. S. Oren. "Exotic options for interruptible electricity supply contracts". In: *Operations Research* 50.5 (2002), pp. 835–850.

[64] I. Karatzas, J. P. Lehoczky, and S. E. Shreve. "Optimal portfolio and consumption decisions for a "small investor" on a finite horizon". In: *SIAM Journal on Control and Optimization* 25.6 (1987), pp. 1557–1586.

[65] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. 2nd. New York: Springer-Verlag, 1998.

[66] I. Karatzas et al. "Martingale and duality methods for utility maximization in an incomplete market". In: *SIAM Journal on Control and Optimization* 29.3 (1991), pp. 703–730.

[67] R. J. Kaye, H. R. Outhred, and C. H. Bannister. "Forward contracts for the operation of an electricity industry under spot pricing". In: *IEEE Transactions on Power Systems* 5.1 (1990), pp. 46–52.

[68] A. Keshavarz, Y. Wang, and S. Boyd. "Imputing a convex objective function". In: *Proceedings of 2011 IEEE International Symposium on Intelligent Control*. 2011, pp. 613–619.

[69] B. J. Kirby. "Spinning Reserve From Responsive Loads". Oak Ridge National Laboratory, ORNL/TM-2003/19. 2003.

[70] Edgar D. Klenske et al. "Nonparametric dynamics estimation for time periodic systems". In: *Proceedings of the 51st Annual Allerton Conference on Communication, Control and Computing*. 2013, pp. 486–493.

[71] J. Kocijan and R. Murray-Smith. "Gaussian process model based predictive control". In: *Proceedings of the 2004 American Control Conference*. 2004, pp. 2214–2219.

[72] P. Kokotović and J. Heller. "Direct and adjoint sensitivity equations for parameter optimization". In: *IEEE Transactions on Automatic Control* 12.5 (1967), pp. 609–610.

[73] R. Korn and E. Korn. *Option Pricing and Portfolio Optimization*. Providence, RI: American Mathematical Society, 2001.

[74] A. Krause, A. Singh, and C. Guestrin. "Near-optimal sensor placements in Gaussian processes: theory, efficient algorithms and empirical studies". In: *Journal of Machine Learning Research* 9 (2008), pp. 235–284.

[75] A. Krause et al. "Simulatneous optimization of sensor placements and balanced schedules". In: *IEEE Transactions on Automatic Control* 56.10 (2011), pp. 2390–2405.

[76] N. R. Kristensen, H. Madsen, and S. B. Jørgensen. "Parameter estimation in stochastic grey-box models". In: *Automatica* 40 (2004), pp. 225–237.

[77] H. Kushner and P. G. Dupuis. *Numerical Methods for Stochastic Control Problems in Continuous Time*. New York: Springer, 2001.

[78] A. H. Land and A. Doig. "An automatic method of solving discrete programming problems". In: *Econometrica* 28.3 (1960), pp. 497–520.

[79] L. F. S. Larsen, T. Geyer, and M. Morari. "Hybrid model predictive control in supermarket refrigeration systems". In: *Proceedings of the 16th World Congress of the International Federation of Automatic Control*. 2005.

[80] J. B. Lasserre. "An explicit equivalent positive semidefinite program for nonlinear 0–1 programs". In: *SIAM Journal on Optimization* 12.3 (2002), pp. 756–769.

[81] M. Liu and F. F. Wu. "Managing price risk in a multimarket environment". In: *IEEE Transactions on Power Systems* 21.4 (2006), pp. 1512–1519.

[82] L. Lovász and A. Schrijver. "Cones of matrices and set-fuctions and 0–1 optimization". In: *SIAM Journal on Optimization* 1.2 (1991), pp. 166–190.

[83] Z. Ma, D. S. Callaway, and I. A. Hiskens. "Decentralized charging control of large populations of plug-in electric vehicles". In: *IEEE Trans. on Control Systems Technology* 21.1 (2013), pp. 67–78.

[84] Q. Maillet et al. "Dynamic state estimation in distributed aircraft electric control systems via adaptive submodularity". In: *Proceedings of the 52nd IEEE Conference on Decision and Control*. 2013, pp. 5497–5503.

[85] S. Martello and P. Toth. *Knapsack Problems*. New York: John Wiley & Sons, 1990.

[86] J. L. Mathieu, S. Koch, and D. S. Callaway. "State estimation and control of electric loads to manage real-time energy imbalance". In: *IEEE Transactions on Power Systems* 28.1 (2013), pp. 430–440.

[87] J. L. Mathieu et al. "Arbitraging intraday wholesale energy market prices with aggregations of thermostatic loads". In: *IEEE Transaction on Power Systems* (to appear).

[88] J. L. Mathieu et al. "Energy arbitrage with thermostatically controlled loads". In: *Proceedings of 2013 European Control Conference*. 2013, pp. 2519–2526.

[89] A. von Meier et al. "Micro-Synchrophasors for Distribution Systems". In: *Proceedings of 2014 IEEE PES Innovative Smart Grid Tecnologies Conference*. 2014, pp. 1–5.

[90] H. M. Müller. "The first-best sharing rule in the continuous-time principal-agent problem with exponential utility". In: *Journal of Economic Theory* 79 (1998), pp. 276–280.

[91] W. Murray and K.-M. Ng. "An algorithm for nonlinear optimization problems with binary variables". In: *Computational Optimization and Applications* 47.2 (2010), pp. 257–288.

[92] National Ocenic and Atmospheric Administration. *Quality Controlled Local Climatological Data (QCLCD)*. URL: `http://cdo.ncdc.noaa.gov/qclcd`.

[93] A. Nayyar et al. "Duration-differentiated services in electricity". In: *arXiv:1404.1112 [cs.SY]* (2014).

[94] G. L. Nemhauser and L. A. Wolsey. *Integer and Combinatorial Optimization*. New York: Wiley, 1988.

[95] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher. "An analysis of approximations for maximizing submodular set functions–I". In: *Mathematical Programming* 14 (1978), pp. 265–294.

[96] Y. Nesterov and A. Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*. Philadelphia, PA: SIAM, 1994.

[97] G. O'Brien and R. Rajagopal. "A method for automatically scheduling notified deferrable loads". In: *Proceedings of 2013 American Control Conference*. 2013, pp. 5080–5085.

[98] *OpenADR Alliance*. URL: `http://www.openadr.org/`.

[99] Pacific Gas & Electric. *SmartAC Program*. URL: `http://www.pge.com/en/myhome/saveenergymoney/plans/smartac`.

[100] C. H. Papadimitriou and K. Steiglitz. *Combinatorial Optimization: Algorithms and Complexity*. Dover, 1998.

[101] R. Pedersen et al. "Aggregation and control of supermarket refrigeration systems in a smart grid". In: *Proceedings of the 19th World Congress of the International Federation of Automatic Control*. 2014.

[102] S. Peng. "A general stochastic maximum principle for optimal control problems". In: *SIAM Journal on Control and Optimization* 28.4 (1990), pp. 966–979.

[103] Gianluigi Pillonetto et al. "Kernel methods in system identification, machine learning and function estimation: A survey". In: *Automatica* 50.3 (2014), pp. 657–682.

[104] D. Pisinger. "The quadratic knapsack problem – a survey". In: *Discrete Applied Mathematics* 155 (2007), pp. 623–648.

[105] S. R. Pliska. "A stochastic calculus model of continuous trading: optimal portfolios". In: *Mathematics of Operations Research* 11.2 (1986), pp. 371–382.

[106] E. Polak. *Optimization: algorithms and consistent approximations*. New York: Springer-Verlag, 1997.

[107] B. Rabe. "Race to the top: the expanding role of U.S. state renewable portfolio standards". In: *Sustainable Development Law and Policy* 7.3 (2007), pp. 10–16.

[108] R. Rajagopal et al. "Risk-limiting dispatch for integrating renewable power". In: *Electric Power and Energy Systems* 44 (2013), pp. 615–628.

[109] C. E. Rasmussen and H. Nickisch. "Gaussian Processes for Machine Learning (GPML) toolbox". In: *Journal of Machine Learning Research* 11 (2010), pp. 3011–3015.

[110] C.E. Rasmussen and C.K.I Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006.

[111] S. M. Rinaldi, J. P. Peerenboom, and T. K. Kelly. "Identifying, understanding, and analyzing critical infrastructure interdepencies". In: *IEEE Control Systems* 21.6 (2001), pp. 11–25.

[112] M. Roozbehani et al. "Robust and optimal consumption policies for deadline-constrained deferrable loads". In: *IEEE Transactions on Smart Grid* 5.4 (2014), pp. 1823–1834.

[113] N. Rotering and M. Ilic. "Optimal charge control of plug-in hybrid electric vehicles in deregulated electricity markets". In: *IEEE Transactions on Power Systems* 26.3 (2011), pp. 1021–1029.

[114] H. L. Royden and P. M. Fitzpatrick. *Real Analysis*. 4th. Boston, MA: Prentice Hall, 2010.

[115] D. Sarabia et al. "Hybrid NMPC of supermarket display cases". In: *Control Engineering Practice* 17 (2009), pp. 428–441.

[116] A. Schrijver. "A combinatorial algorithm minimizing submodular functions in strongly polynomial time". In: *Journal of Combinatorial Theory, Series B* 80 (2000), pp. 346–355.

[117] R. Sedgewick. *Algorithms*. Reading, MA: Addison-Wesley, 1983.

[118] S. E. Shafiei et al. "A decentralized control method for direct smart grid control of refrigeration systems". In: *Proceedings of the 52nd IEEE Conference on Decision and Control*. 2013, pp. 6934–6939.

[119] H. D. Sherali and W. P. Adams. "A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems". In: *SIAM Journal on Discrete Mathematics* 3.3 (1990), pp. 411–430.

[120] R. C. Sonderegger. "Dynamic Models of House Heating Based on Equivalent Thermal Parameters". PhD thesis. Princeton University, 1978.

[121] M. Soner and N. Touzi. "Stochastic target problems, dynamic programming, and viscosity solutions". In: *SIAM Journal on Control and Optimization* 41 (2002), pp. 404–424.

[122] G. Strbac. "Demand side management: benefits and challenges". In: *Energy Policy* 36.12 (2008), pp. 4419–4426.

[123] T. H. Summers, F. L. Cortesi, and J. Lygeros. "On submodularity and controllability in complex dynamical networks". In: *arXiv:1404.7665 [math.OC]* (2014).

[124] C.-W. Tan and P. Varaiya. "Interruptible electric power service contracts". In: *Journal of Economic Dynamics and Control* 17 (1993), pp. 495–517.

[125] M. Tanaskovic et al. "Adaptive model predictive control for constrained linear systems". In: *Proceedings of 2013 European Control Conference*. 2013, pp. 382–387.

[126] *Thermal Environmental Conditions for Human Occupancy*. Standard 55, ANSI/ASHRAE, 2013.

[127] P. Toth. "Dynamic programming algorithms for the zero-one knapsack problem". In: *Computing* 25 (1980), pp. 29–45.

[128] U.S. Department of Energy. "Benefits of demand response in electricity markets and recommendations for achieving them". In: (2006).

[129] U.S. Department of Energy. "Commercial real estate energy alliance". In: (2012).

[130] P. P. Varaiya, F. F. Wu, and J. W. Bialek. "Smart operation of smart grid: risk-limiting dispatch". In: *Proceedings of the IEEE* 99.1 (2011), pp. 40–57.

[131] R. Vasudevan et al. "Consistent approximations for the optimal control of constrained switched systems–Part 1: A conceptual algorithm". In: *SIAM Journal on Control and Optimization* 51.6 (2013), pp. 4463–4483.

[132] J. Vondrák. "Symmetry and approximability of submodular maximization problems". In: *SIAM Journal on Computing* 42.1 (2013), pp. 265–304.

[133] L. J. Watters. "Reduction of integer polynomial programming problems to zero-one linear programming problems". In: *Operations Research* 15.6 (1967), pp. 1171–1174.

[134] Wiki Energy. URL: `http://wiki-energy.org/`.

[135] C. K. Woo et al. "The impact of wind generation on the electricity spot-market price level and variance: The Texas experience". In: *Energy Policy* 39.7 (2011), pp. 3939–3944.

[136] X. Xu and P. J. Antsaklis. "Optimal control of switched systems based on parametrization of the switching instants". In: *IEEE Transactions on Automatic Control* 49.1 (2004), pp. 2–16.

[137] I. Yang, D. S. Callaway, and C. J. Tomlin. "Direct load control for electricity market risk management via risk-limiting dynamic contracts". In: *Proceedings of the 52nd Annual Allerton Conference on Communication, Control and Computing*. 2014, pp. 1058–1063.

[138] I. Yang, D. S. Callaway, and C. J. Tomlin. "Dynamic contracts with partial observations: application to indirect load control". In: *Proceedings of 2014 American Control Conference*. 2014, pp. 1224–1230.

[139] I. Yang, D. S. Callaway, and C. J. Tomlin. "Risk-limiting dynamic contracts for direct load control". In: *arXiv:1409.1994 [math.OC]* (2014).

[140] I. Yang, M. N. Zeilinger, and C. J. Tomlin. "Utility learning model predictive control for personal electric loads". In: *Proceedings of the 53rd IEEE Conference on Decision and Control*. 2014, pp. 4868–4874.

[141] I. Yang et al. "Approximation algorithms for optimization of combinatorial dynamical systems". In: *arXiv:1409.7861 [math.OC]* (2014).

[142] I. Yang et al. "Infinitesimal interconnection variations in nonlinear networked systems". In: *Proceedings of the 52nd IEEE Conference on Decision and Control*. 2013, pp. 1417–1422.

[143] I. Yang et al. "Path integral formulation of stochastic optimal control with generalized costs". In: *Proceedings of the 19th World Congress of the International Federation of Automatic Control*. 2014, pp. 6994–7000.

[144] X.-P. Zhang, C. Rehtanz, and B. Pal. *Flexible AC Transmission Systems: Modelling and Control*. Berlin Heidelberg: Springer-Verlag, 2012.

[145] C. Zhao et al. "Design and stability of load-side primary frequency control in power systems". In: *IEEE Transactions on Automatic Control* 59.5 (2014), pp. 1177–1189.