Fairness, Efficiency and Privacy in Energy Resource Allocation: Incentive Design and Machine Learning for Distributed Agents

Jiayi Li

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2024

Reading Committee:
Baosen Zhang, Chair
Lillian Ratliff
June Lukuyu

Program Authorized to Offer Degree:
Electrical and Computer Engineering

University of Washington

Abstract

Fairness, Efficiency and Privacy in Energy Resource Allocation: Incentive Design and Machine Learning for Distributed Agents

Jiayi Li

Chair of the Supervisory Committee:
Baosen Zhang
Department of Electrical and Computer Engineering

The rapid proliferation of distributed energy resources (DERs) has fundamentally reshaped the energy landscape. While these technologies present opportunities for decentralized energy management, they have also introduced diverse and heterogeneous user groups, surpassing the capabilities of traditional efficiency-oriented allocation schemes. This growing complexity often leads to fairness concerns, with disproportionate payments or allocations disadvantaging certain user groups due to their utility formats, budgets or group sizes. Balancing fairness, efficiency, and protecting user privacy in energy resource allocation is a pressing challenge, essential for developing equitable, resilient, and stable energy systems, particularly in the context of decentralized agents and competitive energy markets.

This dissertation develops principled approaches for coordinating distributed agents in energy systems through the design of decision-making frameworks leveraging optimization and learning. Drawing on methods from game theory, machine learning and dynamical systems, this work explores how fairness, efficiency, and privacy objectives can be achieved while ensuring system stability.

The contributions of this dissertation are organized across three key areas:

1. **Privacy-Preserving Adaptive Pricing Mechanisms**: A two-time-scale incentive mechanism, that alternatively updates between the users and a system operator, is proposed to align user behavior with system-level social objectives that induce socially optimal energy usage without requiring private user information. The iterative pricing updates are proven to converge to the social welfare solution under

mild assumptions, accommodating user behaviors driven by machine learning-based load control algorithms.

2. **Fairness in Energy Systems through Aggregator Market Structures**: This thesis formalizes the problem of fair energy resource allocation by introducing an aggregator-based framework that balances fairness and efficiency through principled trade-offs. By jointly optimizing total resources and individual allocations, this work develops schemes that achieve Pareto-optimal outcomes for diverse user groups. Extending to a multi-aggregator setting, a game-theoretical model is proposed to analyze strategic interactions among aggregators in energy markets, proving the existence of a Nash equilibrium. These contributions demonstrate how aggregator structures stabilize market outcomes, ensure equitable resource distribution, and optimize user surplus, providing a comprehensive foundation for fairness in energy systems.

3. **Safe Learning-Based Control for Distributed Systems**: This thesis develops a decentralized reinforcement learning framework for designing neural network-based controllers that ensure exponential stability and safety in distributed systems. By imposing Lipschitz constraints on control policies and engineering their structure to satisfy these constraints by design, the framework guarantees reliable performance under dynamic system conditions. This approach enables decentralized decision-making without requiring real-time centralized coordination, making it scalable and robust for modern energy systems.

By integrating fairness, efficiency, and privacy into the design of energy allocation mechanisms, this dissertation advances the theoretical and practical understanding of decentralized energy systems. The proposed methods contribute to a sustainable, fair, and stable energy future by addressing fundamental challenges in modern energy systems.

Contents

List of Figures

Chapter 1

ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest gratitude to my advisor, Professor Baosen Zhang, for his unwavering support and guidance throughout my PhD journey. I feel incredibly fortunate and honored to have conducted research under his supervision. He has taught me to approach research methodically—beginning with identifying the right questions, breaking down complex problems into manageable pieces, experimenting with toy models, and then generalizing to larger systems, all the way through to writing and presenting our findings.

His mentorship extended far beyond research; even during the challenges of the COVID lockdown, he ensured we stayed connected with regular meetings. When I had to undergo a year-long surgery during the pandemic, his patience and strong support helped me navigate that difficult time while continuing my academic progress. His dedication, encouragement, and passion for excellence have been a constant source of motivation for me. I cannot fully express my appreciation, but I will always look to Baosen as an extraordinary role model as I move forward in my career.

I would also like to sincerely thank the members of my thesis committee members, Professor Chaoyue Zhao, Professor June Lukuyu, and Professor Lillian Ratliff for their willingness to join my committee quickly and for accommodating my defense despite their extremely busy schedules and the tight timeline. I deeply appreciate their commitment to making this process possible. Their insightful and intellectually stimulating questions have played a critical role in refining both my research motivation and the technical aspects of my work. I am very grateful for the time and effort they have invested in helping me improve my research.

I am deeply grateful to Professor Lillian Ratliff for her guidance in game theory, both through the Learning in Games course and our research discussions, which were

instrumental in identifying and developing the game-theoretical frameworks for my work. Additionally, I want to thank Professor June Lukuyu for her invaluable feedback on energy fairness and energy poverty that helped clarify the definitions, motivations, and scope of these key aspects of my research. I would also like to extend my sincere appreciation to Professor Daniel Kirschen for his insightful discussions on the practical implementation of my research, particularly in convergence of dynamic pricing for electricity markets. I also want to extend my sincere thanks to Prof Maryam Fazel, Prof Sam Burden and Prof Eli Shlizerman for serving on my qualifying exam committee. I also appreciate all the professors and mentors I've taken classes and interacted with.

It has been a privilege to work alongside an exceptional group of collaborators and labmates at the University of Washington throughout my graduate years. I am especially grateful to Wenqi Cui, who has been both a mentor and a friend. Wenqi's expertise in dynamical systems, voltage control and reinforcement learning has profoundly shaped my understanding of these fields. Together, we explored many new problems, had countless thought-provoking discussions about research and life, and collaborated on paper writing and career development. I also want to express my heartfelt gratitude to Yuanyuan Shi and Yize Chen for their invaluable guidance and many enlightening discussions that extended beyond research to career and life decisions. Their encouragement and suggestions have been instrumental in shaping my journey. I want to thank Matt Motoki for his thoughtful insights and collaborative spirit during our brainstorming and paper writing sessions, which helped bring many new research ideas to life. Additionally, I am deeply appreciative of Yan Jiang, whose deep knowledge of control theory enriched my research journey, and whose engaging discussions on research and life were always a source of inspiration. I also really appreciate Liyuan Zheng for his guidance and inspiring perspective on conducting impactful research and navigating career development. I also want to thank Yao Long and Yushi Tan for their invaluable career advice.

My gratitude extends to my labmates (in no particular order): Daniel Tabas, Ling Zhang, Zixiao Ma, Trager Joswig-Jones, Lane Smith, Adhyyan Narang and Benjamin Chasnov. Their camaraderie, collaborative efforts, and lively discussions created a supportive and stimulating research environment that I will always treasure. I also want to thank my colleagues (in no particular order), including Kun Su, Evan Faulkner, Jingyuan Li, Mingfei Chen, Daniel Calderone, Yang Zheng, Sarah Li, Xiangyu Gao and Amber Chou. I am grateful for the professional and personal support we have shared in co-creating a nurturing environment for growth and collaboration.

My research career was partially supported by the National Science Foundation and the University of Washington Clean Energy Institute, whose funding played a vital role in

enabling this work. I am deeply grateful for their support, which provided the resources and opportunities necessary to pursue my research goals.

During my PhD, I was fortunate to intern at Microsoft Research and the Reinforcement Learning Group at Amazon. At Microsoft Research, I am deeply grateful to my mentor, Weiwei Yang, for introducing me to real-world applications of machine learning, data science and brain-computer interface. Weiwei provided invaluable guidance and the research freedom to explore new directions and experiment with various peak pricing and forecasting models. I also want to thank Kate Lytvynets for her hands-on support with day-to-day implementation and model deployment. At Amazon, I truly appreciate the mentorship of Prof. Eric Laber and Dr. Zachary Hervieux-Moore, whose expertise and guidance helped me develop business-constrained generative models and apply reinforcement learning to build and improve recommendation systems. Their mentorship was instrumental in bridging theoretical concepts with practical applications. I am also thankful to my colleague, Hunyong Cho, for valuable discussions on developing recommendation systems and implementation details, which deepened my understanding and improved my technical skills. These internships significantly enriched my PhD journey by broadening my research horizons, deepening my understanding of the intersection between academia and industry, and leaving a lasting impact on my professional development.

This dissertation would not have been possible without the support and encouragement of my family and friends. I am deeply grateful to my mother, Jingrong, and my father, Zhiyong, whose unconditional love and belief in me have been the foundation of all my achievements and personal growth. Their guidance, devotion and trust in my potential have been the driving force behind every milestone of my journey. I also want to thank my friends for support and companionship throughout the years, no matter where life has taken us or the challenges we've faced. I am especially grateful to Ziyi, Jing, Kaiyu, Ling, Steve, Yike, Xingjian, Jiayu, Dantong, Daniel, Liwei, Fangyu, Theophile, Ben and all my friends for standing by me during the highs and lows of this journey. I am also deeply grateful to Dr. Kronstrom and Dr. Takada from UW School of Dentistry, ensuring the success of my surgery and recovery while guiding me through numerous challenging situations throughout the second-year of PhD with expertise and support during the peak of the COVID-19 pandemic.

This journey has been a collective effort, shaped by the guidance, support, and encouragement of countless individuals, to whom I am profoundly grateful. I will carry all the lessons, wisdom, inspiration and kindness as I move forward into the next chapter of my life.

DEDICATION

to my parents,
whose steadfast love and support have made this journey possible.

Chapter 2

Introduction

The growing integration of renewable energy and distributed energy resources (DERs) is transforming modern energy systems, creating new opportunities for sustainability and efficiency while introducing complex challenges in ensuring fairness, privacy, and stability. Distributed energy resources (DERs), such as solar panels, electric vehicles, and batteries, are at the forefront of this revolution, enabling decentralized energy management and empowering users to actively participate in energy markets. This decentralization marks a significant departure from traditional centralized systems, but it also brings complex challenges in achieving fairness-efficiency trade-offs, protecting privacy, and maintaining stability within energy resource allocation frameworks. Addressing these challenges is essential to design energy systems that are fair, resilient, economically viable and stable.

Maintaining stability, particularly in voltage control, has become increasingly difficult due to the variability and intermittency of renewable energy generation. The fast response capabilities of inverter-based DERs offer a promising solution for voltage regulation, but designing decentralized controllers that ensure stability and safety across distributed networks remains a significant technical challenge. Existing methods often lack scalability and fail to account for the dynamic behaviors of modern power systems, necessitating advanced approaches that combine reinforcement learning and robust control.

Traditional energy systems prioritized efficiency but often overlooked fairness, resulting in disparities across user groups. The rise of DERs has introduced a heterogeneous mix of users, ranging from small-scale households to industrial consumers. Without mechanisms to explicitly address fairness, economically disadvantaged groups face disproportionate burdens, exacerbating inequalities in energy access and costs. Meanwhile, larger consumers such as data centers are driving unprecedented demand, straining infrastructure and contributing to imbalanced resource distribution.

Aggregators, as intermediaries between users and wholesale markets, offer a promising avenue for addressing these challenges. By pooling resources and coordinating energy usage, they can allocate resources fairly among users and stabilize market dynamics through strategic interactions. However, their role introduces additional complexities, including privacy concerns and economic interdependencies, which require innovative frameworks to balance fairness, efficiency, with system stability guarantee.

This thesis explores decision-making frameworks that integrate game theory, machine learning, optimization and dynamical system to address these multifaceted challenges. The methodologies developed herein aim to balance fairness and efficiency, provide privacy-preserving mechanisms, and ensure stable system operations. These approaches offer theoretical guarantees and practical solutions for modern energy systems, contributing to a more sustainable and inclusive energy future.

2.0.1 Motivation

The increasing integration of distributed energy resources (DERs) has redefined the structure of energy systems, presenting both challenges and opportunities:

Rising Electricity Demand

The rapid expansion of data centers, coupled with increased electrification, is driving a significant rise in electricity demand. By 2030, data centers alone could account for up to 26% of electricity consumption in the United States, putting unprecedented stress on regional grids and necessitating sustainable energy solutions [2, 3, 4]. As data centers and other large consumers scale operations, their energy needs exacerbate grid challenges, particularly during peak demand periods. The Department of Energy highlights the importance of clean energy resources, such as solar, wind, and battery storage, to address this surging demand [2].

Unlocking Efficiency through Autonomous Systems

Advances in artificial intelligence and machine learning are transforming energy systems. For example, back to 2016, DeepMind's AI-based cooling system reduced data center cooling costs by over 40%, demonstrating the potential of autonomous agents to optimize energy usage efficiently [5]. Similarly, autonomous smart home devices are widely used to adjust consumption patterns in real-time based on dynamic electricity prices and user behaviors. However, despite their capabilities, autonomous agents often operate in isolation, resulting in suboptimal system-wide outcomes.

Coordination Among Distributed Agents

DERs and autonomous agents increasingly optimize local objectives without considering global system requirements. This lack of coordination can lead to demand-supply mismatches, such as renewable energy oversupply during low-demand periods or shortages at peak times. Addressing these mismatches requires incentive mechanisms that align local decisions with global objectives while ensuring stability and efficiency in market operations [6].

Fairness in Resource Allocation

The increasing diversity of user groups—ranging from large-scale data centers to individual households—exacerbates inequities in energy access and cost distribution. Larger users often benefit disproportionately due to economies of scale, while smaller users face higher and unpredictable costs. For example, in the Pacific Northwest, the data center boom has strained hydropower resources, raising concerns about the equitable distribution of energy resources and pricing impacts on local communities [4]. Protecting disadvantaged communities from being priced out of essential energy services is critical to ensuring fairness in energy systems.

Efforts to address fairness in energy systems have led to the introduction of community-based market structures, such as energy collectives, which aim to promote fairness by pooling resources among participants. However, these methods often fail to model individual user surpluses explicitly, limiting their ability to achieve optimal trade-offs between fairness and efficiency. Additionally, many fairness-oriented frameworks rely on oversimplified utility assumptions, such as quadratic functions, which fail to capture the diverse needs and behaviors of modern energy users. These limitations underscore the need for more comprehensive frameworks that address user heterogeneity while providing scalable and adaptive solutions to ensure fair resource allocation.

Market Complexity and Economic Interdependencies

Aggregators act as intermediaries between users and wholesale markets, facilitating resource allocation and stabilizing energy markets. However, their strategic interactions introduce additional complexities. Aggregators must navigate trade-offs between fair resource allocation and market stability while contending with dynamic pricing relationships. Existing frameworks often assume aggregators act as price-takers or operate in homogeneous markets, overlooking the strategic behaviors and diversity observed in

real-world systems. These assumptions can result in significant disparities in resource distribution and user surplus, particularly for smaller or disadvantaged users.

Moreover, current fairness frameworks often focus on static utility functions, failing to account for the dynamic relationship between resource allocation and pricing. In energy systems, prices are directly linked to demand and total allocated resources, meaning fixed utility assumptions overlook critical interactions that influence pricing strategies and market stability. Designing pricing mechanisms that account for both local utility functions and global constraints (e.g., total system usage and physical capacity limits) is essential to address these challenges effectively. Designing pricing mechanisms that account for both local utility functions and global constraints (e.g., total system usage and physical capacity limits) is essential to address these challenges effectively [3, 6]. By addressing these gaps, fairness frameworks can better promote fair resource allocation while enhancing the efficiency and stability of energy systems.

Stability in Voltage Control with Distributed Energy Resources

The growing reliance on distributed energy resources (DERs) such as rooftop solar panels, electric vehicles, and battery storage introduces significant challenges in maintaining voltage stability within power distribution networks. These challenges are exacerbated by the intermittent nature of renewable energy generation and the variability in energy consumption patterns, which occur at timescales faster than conventional mechanical control devices can handle. Inverter-based DERs provide an opportunity for fast and flexible voltage control by leveraging power-electronic interfaces. However, designing controllers that ensure system stability, particularly in decentralized settings, remains a complex task.

Existing control methods often rely on centralized communication and linear control strategies, which are not scalable or adaptable to the heterogeneity of modern energy systems. Reinforcement learning (RL) approaches have emerged as a promising solution, enabling controllers to adapt to dynamic conditions without the need for detailed system models. Nevertheless, most RL methods lack formal guarantees on system stability, a critical requirement for practical deployment in real-world energy systems. Addressing these limitations requires developing decentralized frameworks that enforce stability constraints while optimizing voltage control actions in a model-free setting. Incorporating such advanced voltage control mechanisms is essential for enabling the seamless integration of DERs into modern energy systems.

2.1 Contributions

This thesis addresses the challenges of fairness, efficiency, privacy, and stability in energy resource allocation through novel incentive design, game-theoretic modeling, and decentralized control frameworks. The key contributions are outlined as follows:

2.1.1 Socially Optimal Energy Usage via Adaptive Pricing

Coordinating the energy consumption of distributed users poses significant challenges due to the lack of private user information available to the system operator. Privacy concerns often prevent operators from directly optimizing resource allocation, which can lead to inefficiencies. To address this, the first part of this thesis develops a two-time-scale incentive mechanism that iteratively updates prices based on observed consumption patterns. Users respond by optimizing their consumption according to the prices. This decentralized interaction eliminates the need for the operator to access or learn private user information while converging to a socially optimal solution that balances individual preferences with system-level efficiency.

The framework's flexibility is demonstrated through its ability to handle non-convex cost functions and incorporate machine learning-based load control algorithms, making it scalable and adaptable to diverse energy systems. By aligning individual consumption behaviors with global objectives, this contribution lays the foundation for privacy-preserving, efficient energy usage in decentralized systems.

2.1.2 Balancing Fairness and Efficiency in Energy Resource Allocations

The rise of distributed energy resources (DERs) has introduced diverse and heterogeneous user groups, exposing disparities in energy access and cost distribution. Traditional allocation schemes, focused solely on efficiency, often fail to address the needs of disadvantaged users, exacerbating inequities in resource allocation. This thesis formalizes the problem of fair energy resource allocation and introduces a principled framework for balancing fairness and efficiency using α-fairness metrics.

The proposed framework constructs allocation schemes that lie on the Pareto front, allowing for optimal trade-offs between fairness and efficiency. By jointly optimizing total resources and individual allocations, this approach ensures fairness is explicitly accounted for while maintaining system performance. This contribution highlights how fairness can be embedded into energy systems without compromising efficiency, particularly for underserved communities.

2.1.3 Strategic Aggregator Interactions in Energy Markets

In systems with multiple aggregators acting as intermediaries between users and energy markets, strategic interactions add layers of complexity to resource allocation. The thesis extends the fairness-efficiency framework to a multi-aggregator setting, modeling the interactions among aggregators as a quasi-concave game. This analysis rigorously proves the existence of Nash equilibria, providing theoretical guarantees for market stability and resource allocation outcomes.

By incorporating game-theoretic modeling, this work demonstrates how aggregators stabilize energy markets, improve fairness in resource distribution, and optimize user surplus. The framework also provides actionable insights into managing fairness and efficiency in competitive energy markets, addressing the interdependencies between aggregators and market dynamics. This contribution advances the understanding of strategic interactions in decentralized energy systems and their role in achieving equitable and efficient outcomes.

2.1.4 Decentralized Safe Reinforcement Learning for Voltage Control

The increasing reliance on distributed energy resources, particularly inverter-based systems, introduces new challenges in maintaining voltage stability in distribution networks. Existing control methods often fail to guarantee safety and stability in dynamic environments, especially when reinforcement learning (RL) approaches are used. The final contribution of this thesis addresses these challenges by developing a decentralized safe RL framework for voltage control.

This framework ensures system stability by imposing Lipschitz constraints on neural network-based controllers, guaranteeing exponential stability across the network. The decentralized approach enables local controllers to be trained in a model-free setting without requiring real-time communication between nodes. By scaling efficiently to large distribution networks, this contribution bridges the gap between advanced machine learning techniques and practical requirements for grid stability, offering a robust solution for modern energy systems.

2.2 Structure of the Thesis

Chapter 3 introduces the incentive mechanisms developed for adaptive pricing in energy markets and presents theoretical and empirical results demonstrating their efficacy in achieving socially optimal energy usage. The focus is on how users' privacy concerns are

addressed while still achieving global efficiency objectives.

Chapter 4 discusses the formalization of fairness and efficiency trade-offs in energy resource allocations. The chapter introduces the aggregator framework, which jointly optimizes total resources and individual allocations, ensuring the balance between fairness and system efficiency.

Chapter 5 extends the fair resource allocation problem to a multi-agent setting, exploring the strategic interactions between multiple aggregators in a central energy market. The chapter includes the theoretical proof of the quasi-concave game structure and the existence of a Nash equilibrium, followed by simulations that demonstrate practical outcomes in competitive energy markets.

Chapter 6 focuses on decentralized safe reinforcement learning for voltage control, presenting the safe RL framework and its application to inverter-based distributed energy resources. The chapter includes a detailed discussion on the design of Lipschitz-constrained neural network controllers and the guarantees of system stability.

Chapter 7 concludes the thesis by summarizing the contributions from each chapter and discussing potential avenues for future research in energy resource allocation, incentive mechanism design, and decentralized control systems.

Chapter 3

Socially Optimal Energy Usage via Adaptive Pricing

3.1 Introduction

We study the coordination of the electricity usage of a group of users by an operator. This question has been studied extensively by the community, for example, in the context of demand response, customer aggregation, and virtual power plants (see [7, 8, 9, 10] and the references within). The common setup is where each user is endowed with a cost (or utility) function, and the operator seeks to minimize a global cost function that is made up of the individual costs of the users and social welfare considerations.

A central challenge in these problems is that the cost functions of the users may not be known to the operator. Furthermore, users themselves may not be able to provide an analytical description. For example, using machine learning to schedule and manage demand has been a very active area of research (see, e.g. [11, 12, 13, 14, 15] and the references within). However, most learning algorithms minimize some cost or discomfort by changing the behavior of the users or their devices and do not provide a cost function in terms of power that can be easily optimized. It's important to note that we do not mean that a cost function does not exist, rather, it is often not revealed when learning algorithms are applied. In many existing algorithms, however, the operator needs to somehow learn the cost functions of the users [16, 17, 18]. This restricts the users to having simple (typically quadratic) cost functions and is often too restrictive to be implemented in practice.

In this chapter, we overcome this challenge by deploying a two-time-scale incentive mechanism that alternatively updates between the operator and the users. More concretely,

the actions of the users are characterized by their electricity consumption and the action of the operator is in selecting a price vector. This price is an externality to the users, and a user optimizes its actions by minimizing its own internal cost and the external cost induced by the presence of a price. The operator adaptively updates the price to influence the behavior of the users to achieve a socially beneficial solution. The adaptive updates can be treated as a negotiation process: the customer and the system operator are negotiating on a future price that optimizes the social welfare of the system. Note that as long as users can optimize their actions given a price, it suffices for the operator to just observe the actions of the users, and the operator does not need to know or learn users' private information.

These types of incentive designs where an operator sets a payment (or sometimes called a tax) have received significant interest. Typically, the goal is to provide incentives to selfish users such that the resulting equilibrium is closer to the socially optimal solution [19, 20, 21]. In the context of electricity markets, a large number of demand response strategies have been proposed, ranging from alert/text-based signals [22], to pricing [23], to direct load control [24]. In this chapter, we focus on developing pricing mechanisms. The incentive mechanism we adopt is from [25], which is a type of dynamic incentive that evolves with the actions of the users. The key technical questions in these types of mechanisms are twofold: whether the price update converges, and if so, whether it leads to socially optimal user behaviors.

Existing results in answering these two questions often require strong assumptions that do not readily apply to electricity markets. In [26], the cost function of the users and the system cost are assumed to be linear. The results in [27] generalize the class of cost functions but assume the users' strategies are convex in the price. This assumption often does not hold even in relatively simple games induced by the price. The work of [28] is similar to ours but makes much stronger assumptions that the system and users' costs are convex and smooth. Moreover, our algorithm circumvents the need for each user to determine the incremental demand at each step, which is challenging, especially for learning-based methods. The work in [25], which acts as the impetus to this chapter, shows that the price converges and leads to socially optimal behavior assuming strict convexity of the cost functions.

In this chapter, we relax the assumptions needed in [26, 27, 25], and show that the two-time-scale algorithm leads to price dynamics that both converge and induce optimal social behavior of the users. For example, we show that in some settings, the only assumption needed is that a user's action (electricity consumption) is decreasing in price. This assumption is intuitive because higher price naturally leads to lower consumption.

This result would then apply to users who use learning algorithms to manage their demands. We also generalize the type of social costs that operators could consider. For example, in addition to widely used quadratic functions and time-of-use pricing, our framework provides a way to implement peak pricing through designing the social cost.

In our proof of convergence, we construct a continuous-time dynamical system and its associated Lyapunov functions. Then we show that the price leads to user actions that are optimal for the global optimization problem by analyzing the optimality conditions. The chapter is organized as follows. In Section 3.2, we describe the setup of our model. In Section 3.3, we provide the theoretical results, and Section 3.4 illustrates these results with several simulated case studies. We conclude and give some thoughts in future directions in Section 3.5.

3.2 Problem Formulation and Pricing Updates

In this section, we formally state the problem of obtaining socially optimal energy usage and describe an iterative procedure to compute the price seen by each of the individual users.

3.2.1 Global Optimization Problem

We consider a system with N users. The action of each user is their energy consumption over T time periods. For example, $T = 24$ if we are considering hourly loads of the users in a day. Let $\mathbf{x}_i \in \mathbb{R}^T$ denote the electricity consumption of user i.[1] We assume $\mathbf{x}_i$ takes value in a compact set $\mathcal{X}_i \subset \mathbb{R}^T$.[2] The notation $x_{i,t}$ denote the energy consumption of user i at the t'th period. Each user is interested in optimizing its cost function, $f_i : \mathbb{R}^T \to \mathbb{R}$. For example, if $f_i(\mathbf{x}_i) = \frac{1}{2}(\mathbf{x}_i - \bar{\mathbf{x}}_i)^2$, the user can be interpreted as minimizing its energy usage to some predetermined setpoint $\bar{\mathbf{x}}_i$. Note that in this paper we do not assume any symmetry between the users, nor do we restrict f_i to be in particular parametric forms. Of course, some assumptions on f_i's are required, and these will be stated closer to the technical results section.

We also place a cost on the system in serving the total energy demanded by the users. We do not distinguish between which user is using the energy, and the system cost is in the form of $g\left(\sum_{i=1}^N \mathbf{x}_i\right)$. Note the sum in g is over the users, and g is a function from $\mathbb{R}^T$ to $\mathbb{R}$. Throughout this paper, we will assume that g is convex or strictly convex. Several examples of g include $g\left(\sum_i \mathbf{x}_i\right) = c\left(\sum_i \mathbf{x}_i\right)$, where $c(\cdot)$ is the cost of procuring energy;

[1] We use bold notation to indicate vectors and matrices.

[2] The constraints that a user has can be encoded in this set.

$g\left(\sum_i \mathbf{x}_i\right) = ||\sum_i \mathbf{x}_i||^2$, where the system tries to reduce its total energy consumption; and $g\left(\sum_i \mathbf{x}_i\right) = \max\left(\sum_i x_{i,1}, \ldots, \sum_i x_{i,T}\right)$, which represent a cost on the peak energy demand [29].

The system operator is interested in solving a global social welfare problem:

$$\min_{\mathbf{x}} \quad \sum_i f_i(\mathbf{x}_i) + g\left(\sum_i \mathbf{x}_i\right). \tag{3.1}$$

The first term represents the sum of the users' costs, while the second represents a social or system cost. This problem has been studied in many settings. When the operator has full information about f_i's and g, and can set the demand of the users, this is essentially a variant of the economic (load) dispatch problem. This has been extensively studied and readers can refer to [7, 30] and the references within.

In practice, it is often the case that the system operator does not know the exact cost functions of the users. This could stem from privacy concerns, where the user does not communicate their cost to the operator. Another way that incomplete information can arise is that the users themselves do not have a closed-form description of their own costs. Learning algorithms are becoming increasingly popular for demand management, where a user's action is trained without explicitly learning a cost model [1, 14, 13]. In this setting, an iterative algorithm is often used, where local steps and global steps alternate and try to converge to the solution of the social welfare problem in (3.1).

When f_i's are convex (or strictly convex), differentiable and frequent communications are possible, primal/dual type of gradient algorithms can be used to solve it. More efficient algorithms exist when f_i's are assumed to be in particular parametric forms (mostly quadratic [16, 17]). Drawbacks of these approaches appear when the local problem is more complicated. For example, suppose user i uses tabular Q-learning to determine its actions. Then it is difficult to apply existing distributed optimization algorithms.

In this paper, we study a strategy to iteratively solve (3.1), where we can weaken the existing assumptions on the users. We do this through price updates and more details are given next.

3.2.2 User's Optimization Problem

From a user's perspective, it receives a price vector, $\mathbf{p} \in \mathbb{R}^T$ from the system operator. We will define how the operator arrives at this price in the next section, but for now, we

treat it as a given vector. User i solves the following optimization problem:

$$\min_{\mathbf{x}_i \in \mathcal{X}_i} \quad f_i(\mathbf{x}_i) + \mathbf{p}^\intercal \mathbf{x}_i. \tag{3.2}$$

We denote the solution of this problem as $\mathbf{x}_i^*(\mathbf{p})$. We do not consider the details of how user i solves (3.2) to obtain $\mathbf{x}_i^*(\mathbf{p})$. If f_i is available, the user may explicitly solve (3.2). In [1], a user represents a water heater and has the goal of balancing the discomfort of receiving underheated or overheated water with the cost of power. Since the discomfort is easily expressed in terms of the temperature of the water but not in terms of power, a Q-learning algorithm is used to look up $\mathbf{x}_i^*(\mathbf{p})$.

In this paper, we make the following assumption:

Assumption 1. *Given a price* $\mathbf{p}$*, the solution* $\mathbf{x}_i^*(\mathbf{p})$ *to* (3.2) *is unique for all users* i*.*

This says that the behavior of a user is uniquely determined by the price it sees from the system operator. This condition is true under a wide range of conditions, and we note it is much weaker than the ones made in existing literature. For example, it holds if f_i is strictly convex (as assumed in [25, 26, 31, 28], or if the user used a deterministic algorithm regardless of convexity (not covered by the conditions in [25, 26, 31, 28]). If a probabilistic algorithm is used, the results in the paper hold with respect to the averaged solution, but the analysis becomes much more cumbersome. A more refined understanding of possible stochastic behaviors is an important future direction for us.

It's worth mentioning that this assumption is not strictly necessary for the adaptive algorithm to find socially optimal incentives. However, without this assumption, the updates may oscillate between several different solutions. When the adaptive algorithm does not converge to a unique solution, the notion of convergence becomes more difficult to characterize. In response to the potential oscillations that destabilize the system, the system operator could develop a consistent tie-breaking procedure (e.g. consuming at the earliest hour) so the system would converge to a fixed point. In addition, we could expand the allowed behavior of the agents by considering ideas such as convergence in expectation.

3.2.3 Price Update

Now we consider how $\mathbf{p}$ should be updated by the system operator. Following the scheme in [25], we define the externality $\mathbf{e}$ as the marginal social cost arising from the term g:

$$\mathbf{e}(\mathbf{x}) = \nabla_{\mathbf{z}} g(\mathbf{z})|_{\mathbf{z}=\sum_i \mathbf{x}_i} \tag{3.3}$$

where $\mathbf{x} = \{\mathbf{x}_i\}_{i=1}^N$. The price is updated as

$$\mathbf{p}_{k+1} = (1 - \beta_k)\mathbf{p}_k + \beta_k \mathbf{e}(\mathbf{x}^*(\mathbf{p}_k)). \tag{3.4}$$

where $\mathbf{x}^*(\mathbf{p}_k) = \{\mathbf{x}_i^*(\mathbf{p}_k)\}_{i=1}^N$ and the following conditions hold $\sum_{k=1}^\infty \beta_k = \infty$ and $\sum_{k=1}^\infty \beta_k^2 < \infty$; e.g., $\beta_k = 1/k$. These assumptions about the step size are standard assumptions that allow us to prove the convergence of the price updates.[3]

The main technical results of the paper are to show that under mild assumptions, the dynamics in (3.4) converge to a unique $\mathbf{p}^*$ and $\mathbf{x}_i^*(\mathbf{p}^*)$ is the minimizer of (3.1). This price update scheme is attractive because the operator doesn't need to know or learn about the users.[4] As long as the sum of decisions $\mathbf{x}_i$'s are observed, a unique $\mathbf{p}^*$ can be found to induce globally optimal behavior.

Algorithm 1 Adaptive Pricing: An iterative method that solves the global optimization problem (3.1).

0: Initialize $\mathbf{p} \in \mathbb{R}^T$
0: **for** $k = 1, 2, \ldots$ **do**
0: *// Solve Local Optimization*
0: **for** $i = 1, 2, \ldots, N$ **do**
0: $\mathbf{x}_i^*(\mathbf{p}) = \arg\min_{\mathbf{x}_i \in \mathcal{X}_i}\ f_i(\mathbf{x}_i) + \mathbf{p}^\intercal \mathbf{x}_i$
0: **end for**
0: *// Price Update*
0: Choose step size: $\beta = 1/k$
0: Compute externality: $\mathbf{e} = \nabla g\Big(\sum_{i=1}^N \mathbf{x}_i^*(\mathbf{p})\Big)$
0: Compute new price: $\mathbf{p} = (1 - \beta)\mathbf{p} + \beta\mathbf{e}$
0: **end for**=0

3.3 Theoretical Results

The first step to analyzing the system in (3.4) is to approximate it as a continuous system. A standard result in dynamical systems is that as $k \to \infty$ and the step-sizes β_k are asymptotically going to zero, (3.4) behaves as (3.5) [25, 32]:

$$\dot{\mathbf{p}} = \mathbf{e}(\mathbf{x}^*(\mathbf{p})) - \mathbf{p}. \tag{3.5}$$

[3] We note that step sizes of $1/k$ are required for theoretical reasons; however, we find that experimentally that other step sizes can lead to the same fixed point with fewer iterations.

[4] In fact, the operator does not even need to know how many users there are in the system since it simply broadcasts the price to all users.

We want (3.5) to have two properties. The first is that it has a unique equilibrium, and the second is that this equilibrium induces the solutions to the local optimization problem in (3.2) that's simultaneously solving the global problem in (3.1).

The proof of these results involves constructing a Lyapunov function based on suitable assumptions on f_i's and g. We feel it is easier to break the proof into two parts, and first consider the case of $T = 1$ and then $T > 1$. When $T = 1$, the proof is fairly short and the intuition carries over to the $T > 1$ case, where the math becomes more cumbersome.

3.3.1 Single time-period case (T = 1)

For a single time-dimensional case, $x_i \in \mathbb{R}$, g is a scalar function, and the dynamical system in (3.5) is also a scalar system. We have the following result

Theorem 1. *Suppose Assumption 1 holds and $x_i^*(p)$ is decreasing for each i and g is convex and differentiable. Then the scalar dynamical system* (3.5) *is globally asymptotically stable and has a unique fixed point p^*. The solutions $x_i^*(p^*)$ solve the global optimization problem in* (3.1).

The decreasing condition on $x_i^*(p)$ says that if the price increases, then the energy consumption of user i will not increase (users will not consume more if energy becomes more expensive). This condition is satisfied for most types of cost functions. For example, if f_i's are convex. The decreasing condition is also satisfied for situations where f_i's are nonconvex [33]. Therefore, our pricing algorithm is applicable to a wider range of problems than existing pricing schemes that require both f_i's and g to be convex [25].

Proof. We first show that a unique equilibrium of (3.5) exists. When $T = 1$, (3.3) simplifies to $e(x) = g'(\sum_i x_i)$. If g is convex, its derivative is increasing and e is an increasing function. Since $x_i^*(p)$ is decreasing in p, the function of p obtained by summing over i, $\sum_i x_i^*(p)$, is also decreasing in p. Using the fact that for scalar functions, the composition of an increasing function with a decreasing function is decreasing, we have $e(\sum_i x_i^*(p))$ being decreasing in p.

Looking at the right hand side of (3.5), $e(x^*(p)) - p$ is a strictly decreasing function ranging from ∞ to $-\infty$ as p goes from $-\infty$ to ∞. Therefore it takes the value 0 at exactly one point, and we denote that point by p^*, and it is an equilibrium point of the dynamical system in (3.5).

Next, we show that the dynamical system is globally asymptotically stable around p^*.

We do this by constructing the following Lyapunov function:

$$V(p) = -\int_{p^*}^{p} [e(x^*(q)) - q]\, dq.$$

Using the fact that the integrand is strictly decreasing, is easy to verify that $V(p) > 0$ if $p \neq p^*$ and $V(p^*) = 0$. The time derivative of V is

$$\dot{V}(p) = V'(p)\dot{p} = -[e(x(p)) - p]^2,$$

which is negative except when $p = p^*$. Therefore, p^* is globally asymptotically stable.

Finally, we show that the p^* induces globally optimal behavior. That is, the solutions of the local problems $x_i^*(p^*)$ solves the global optimization problem in (3.1). To show this, let $\hat{x}$ denote an optimal solution to (3.1). For notational simplicity, define $\hat{x}_s = \sum_i \hat{x}_i$ and $x_s^* = \sum_i x_i^*(p^*)$. Suppose $\hat{x} \neq x^*(p^*)$, and $\hat{x}$ achieves a lower social welfare cost. Then we have

$$\sum_i f_i(\hat{x}_i) + g(\hat{x}_s) < \sum_i f_i(x_i^*(p^*)) + g(x_s^*). \tag{3.6}$$

But $x_i^*(p^*)$ are the minimizers of the local optimization problem in (3.2), we have

$$f_i(x_i^*(p^*)) + g'(x_s^*)x_i^*(p^*) < f_i(\hat{x}_i) + g'(x_s^*)\hat{x}_i,\ \forall i, \tag{3.7}$$

where we use the fact that $p^* = g'(x_s^*)$ and the inequality is strict because of the uniqueness assumption in Assumption 1. Adding (3.6) and (3.7) and simplifying the expressions, we get

$$g(\hat{x}_s) < g(x_s^*) + g'(x_s^*)(\hat{x}_s - x_s^*).$$

This is the first-order condition of a strictly *concave* function, and it contradicts the fact that g is *convex*. Hence the assumption that there exists a more optimal solution to (3.1) than x^* is not possible. □

3.3.2 Multi-Time Period Case (T > 1)

The analysis in the single time period case relies on the monotonicity of several functions. To extend that to the multi-time period case, we extend the notion of monotonicity to vector-valued functions:

Definition 1. *Consider a vector-valued function $h : \mathbb{R}^n \to \mathbb{R}^n$. We say that h is increasing if $[h(\mathbf{x}) - h(\mathbf{x}')]^\intercal(\mathbf{x} - \mathbf{x}') \geq 0$. It is strictly increasing if the inequality is strict when $\mathbf{x} \neq \mathbf{x}'$. A function is (strictly) decreasing if its negation is (strictly) increasing.*

We adopt Definition 1 since it is a natural generalization of monotonicity for a scalar function in the following sense: the gradient of a differentiable convex function is increasing.[5]

In the $T = 1$ setting, the proof for Theorem 1 has three parts. First, we showed that a unique equilibrium price exists, then we show that the dynamical system in (3.5) is asymptotically stable around this price, and finally that this price induces socially optimal behavior. In the $T > 1$ setting, the first and last parts generalize directly from the $T = 1$ setting with minimal changes. However, asymptotic stability is considerably more difficult since the composition of two monotone functions is no longer monotone.[6] Therefore, we first state a lemma about the uniqueness and social optimality of the equilibrium price. Then we provide two conditions where the iterative algorithm converges asymptotically to this equilibrium.

Lemma 2. *Consider the dynamical system in* (3.5)*. If* $\mathbf{x}_i^*(\mathbf{p})$ *is a decreasing function of* $\mathbf{p}$*, then there exists a unique equilibrium* $\mathbf{p}^*$*. Furthermore, under Assumption 1,* $\mathbf{x}^*(\mathbf{p}^*)$ *is the optimal solution to social welfare problem in* (3.1)*.*

Proof. The proof of these two results is analogous to the $T = 1$ case. By sweeping $\mathbf{p}$ from ∞ to $-\infty$ and using the fact that $\mathbf{e}(\mathbf{x}^*(\mathbf{p})) - \mathbf{p}$ is strictly decreasing, we can conclude that there exists a unique $\mathbf{p}^*$ such that $\mathbf{e}(\mathbf{x}^*(\mathbf{p}^*)) - \mathbf{p}^* = 0$. For the second part of the lemma, it suffices to change a derivative to a gradient in (3.7) and the product between scalars to a dot product between vectors. All other steps remain the same. □

Next, we state two conditions where the dynamical system in (3.5) converges:

Theorem 3. *Suppose the local solutions* $\mathbf{x}_i^*(\mathbf{p})$ *are decreasing in* $\mathbf{p}$*. The dynamical system in* (3.5) *is asymptotically stable if one of the following conditions are true*

1. $g(\mathbf{z}) = \frac{1}{2}\mathbf{z}^\intercal \mathbf{B}\mathbf{z}$ *for some positive definite* $\mathbf{B}$*.*
2. *Each* f_i *is twice differentiable and strictly convex and* g *is strictly convex and differentiable.*

Before presenting the proof, we comment on why these settings are potentially interesting and practical. A quadratic cost is the most common form used in the literature [34] since it includes standard pricing schemes such as flat rates and time-of-use

[5] We note there exists other definition of monotonicity for vector-valued functions. For example, sometimes a function is said to be monotone if it is monotone in every coordinate. This definition, however, is too restrictive for our purposes.

[6] This is not even true for linear functions in the vector-valued case.

(ToU) rates. The assumption on each of the users is minimal since we just require that an increase in energy price leads to a decrease in energy consumption.

The second setting generalizes the price function by assuming a more strict form of the local costs f_i's. The reason this result is useful is that it includes peak pricing, which is being adopted by a large number of users [35, 36, 37, 38]. More specifically, peak pricing has the form of $g(\mathbf{z}) = c \cdot \max(z_1, \ldots, z_T)$, where c is some constant. This function is convex, but not strictly convex nor differentiable. But it can be approximated by the so-called LogSumExp function, defined as

$$(z_1, \ldots, z_T) = \frac{1}{\alpha} \log(\exp(\alpha z_1) + \cdots + \exp(\alpha z_T)). \tag{3.8}$$

The LSE function is strictly convex if at least one of its arguments is positive and approximates the max function arbitrarily as α grows [39].

Proof. We prove Theorem 3 by constructing appropriate Lyapunov functions. In the first case, we use a quadratic Lyapunov function

$$V(\mathbf{p}) = \frac{1}{2}(\mathbf{p} - \mathbf{p}^*)^\intercal \mathbf{B}^{-1}(\mathbf{p} - \mathbf{p}^*). \tag{3.9}$$

It is clear that $V(\mathbf{p}) = 0$ if $\mathbf{p} = \mathbf{p}^*$ and $V(\mathbf{p}) > 0$ otherwise.

In the second setting, we use a different Lyapunov function, defined as

$$\begin{aligned} V(p) = {} & g\left(\sum_i \mathbf{x}_i^*(\mathbf{p})\right) - g\left(\sum_i \mathbf{x}_i^*(\mathbf{p}^*)\right) \\ & - \nabla g\left(\sum_i \mathbf{x}_i^*(\mathbf{p}^*)\right)^\intercal \left(\sum_i (\mathbf{x}_i^*(\mathbf{p}) - \mathbf{x}_i^*(\mathbf{p}^*))\right) \\ & - \sum_i (\mathbf{x}_i^*(\mathbf{p}) - \mathbf{x}_i^*(\mathbf{p}^*))^\intercal (\mathbf{p} - \mathbf{p}^*). \end{aligned} \tag{3.10}$$

This function is based on the Bregman's divergence of g, and using the fact that g is strictly convex, $V(\mathbf{p}) = 0$ if $\mathbf{p} = \mathbf{p}^*$ and $V(\mathbf{p}) > 0$ otherwise.

In Appendix A, we show that the time derivative of both Lyapunov functions are negative when $\mathbf{p} \neq \mathbf{p}^*$. □

The proof above illustrates the difficulty of showing a general result when $T > 1$ since we need two different Lyapunov functions. However, we believe this is a consequence of our proof technique, and we conjecture that the twice differentiable and strictly convex condition in Theorem 3 can be relaxed to be just $\mathbf{x}_i^*(\mathbf{p})$ is strictly decreasing for all i.

3.4 Simulation Results

In this section, we conduct numerical studies that illustrate the theoretical results in this paper, in particular the convergence properties of the update in (3.4). Our code for the numerical simulations can be found at `https://github.com/socially-optimal-energy`.

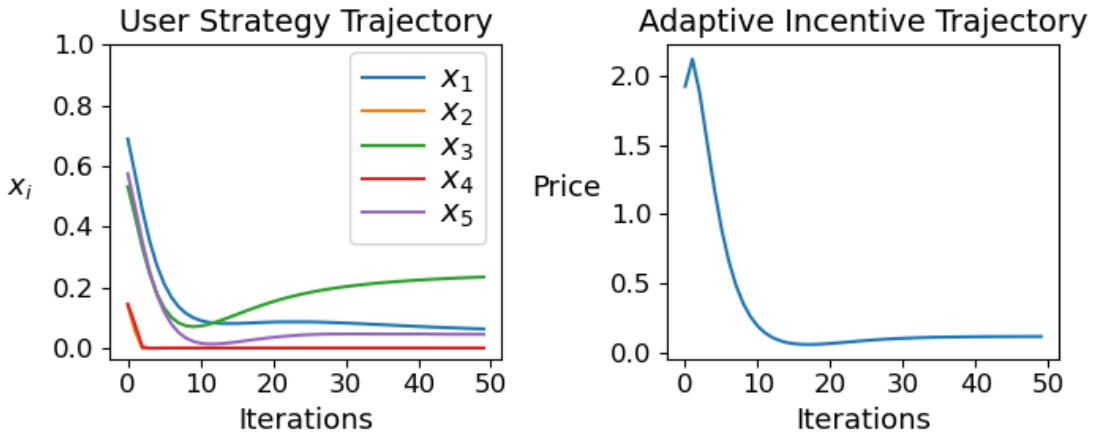


Figure 3.1: Convergence of user actions and price incentive for a system with 5 users and a single time-period. Both the actions and the price converge quickly.

3.4.1 Single Time-Period

The first is a simple single time-period ($T = 1$) example that clearly shows the convergence to the global solution (see Fig. 3.1). Suppose there are 5 users ($N = 5$) that have actions $x_i \in \mathbb{R}$ which can automatically adapt to changes in price.[7] The user cost functions are $f_i(x_i) = \frac{1}{2}(x_i - \bar{x}_i)^2$ and the global cost is $g(\sum_i x_i) = (\sum_i x_i)^2$. Both actions and the price converge quickly, and it is easy to check that they converge to the optimal values (which can be computed in this simple setup).

3.4.2 Peak Pricing of Multiple Time-Periods

Here, we consider multiple periods, with $T = 24$. Then user i's action is $\mathbf{x}_i \in \mathbb{R}^{24}$. For simplicity, let $f_i(\mathbf{x}_i) = \frac{1}{2}(\mathbf{x}_i - \bar{\mathbf{x}}_i)^2$ for some given $\bar{\mathbf{x}}_i$'s. Suppose the system adopts peak pricing, approximated as $g(\sum_i \mathbf{x}_i) = \lambda(\sum_i \mathbf{x}_i)$ for some λ, where the LSE function was

[7] In this example, we assume that the users and utility are in a negotiating a future price (e.g., a day-ahead price) and the algorithm converges by the time the price needs to be set and the load is realized.

defined in (3.8). Following (3.3), the externality of each user is calculated as:

$$\mathbf{e}(\mathbf{x}) = \nabla_{\mathbf{z}} g(\mathbf{z})|_{\mathbf{z}=\sum_i \mathbf{x}_i} = \frac{1}{\sum_{t=1}^{T} \exp(\alpha \mathbf{z}_t)} \exp(\alpha \mathbf{z}).$$

For 10 users, the convergence of price and total demand is shown in Fig. 3.2, where both converge quickly to their final values. The comparison of initial and final price and load profiles are shown in Fig. 3.3. The initial prices (chosen uniformly at random) induce a total (summed over the users at each time period) load profile that is uneven. After convergence, load profiles become much more even, which is what we expect to see when the peak system load is minimized. Note the price that achieves this still shows variations across time periods.

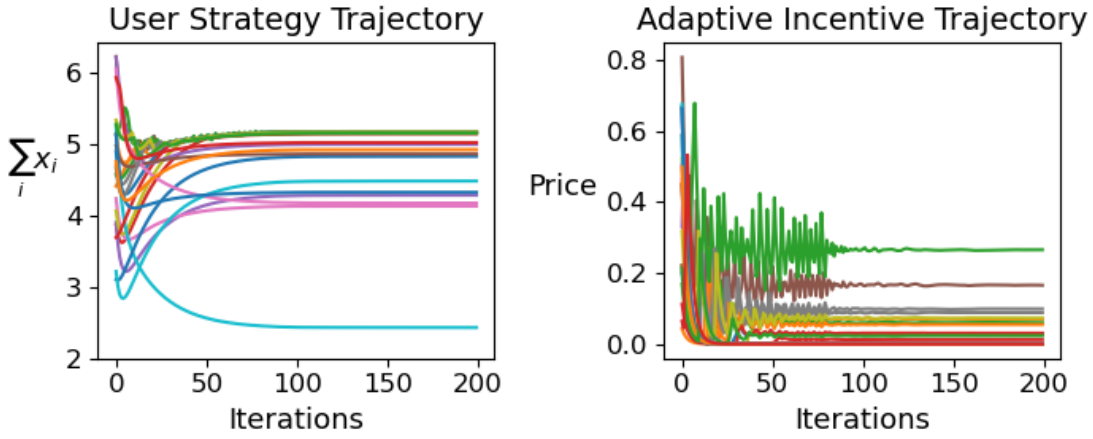


Figure 3.2: The top figure shows the convergence in the sum of the users' actions at each time period. The bottom figure shows the convergence of price.

3.4.3 Water Heater Load Optimization with Q-Learning

Now we consider a case where users' actions are determined by a learning algorithm and show that the f_i's need not be differentiable or convex. We consider a group of water heaters where $\mathbf{x}_i \in \mathbb{R}^{96}$ represents its daily load profile measured every 15-minutes, f_i represents user discomfort, and g is the function. In this example, users can flexibly shift their energy consumption earlier to avoid high prices and store the energy (hot water) until it is needed. We model and optimize each user following the learning algorithm in [1] and numerically check that $\mathbf{x}_i^*$ is decreasing in $\mathbf{p}$.

For simplicity, we assume that each user's demand for hot water is a binary vector drawn from a sequence of Bernoulli random variables. We assume the demand and all

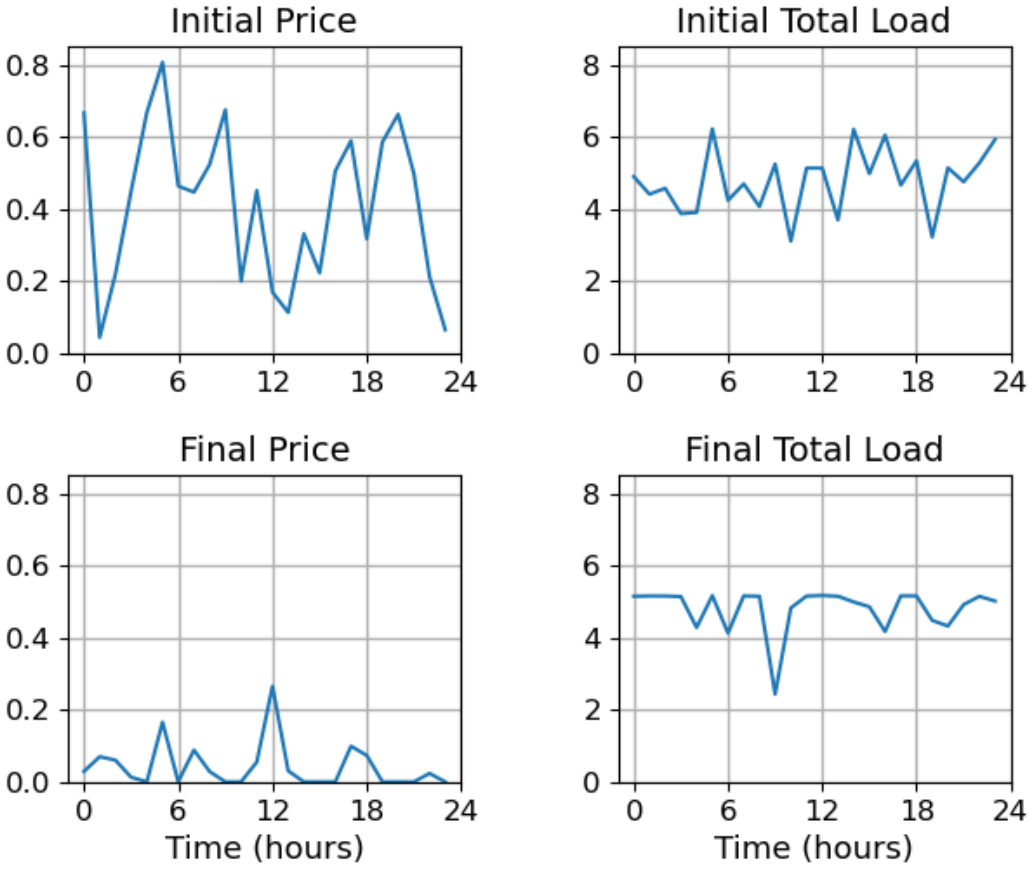


Figure 3.3: The initial and converged price and demand profiles demonstrate that this adaptive pricing framework is effective in reducing system peak.

system parameters are known and so given a price $\mathbf{p}$, we can solve for $\mathbf{x}^*(\mathbf{p})$ exactly using dynamic programming. Fig. 3.4, shows the initial and final price and the corresponding distributions of total energy consumption of all the users. Initially, the system cost is 12.781 and the user cost is 0.630. At convergence, the system cost decreases to 10.397 and the user cost to 0.627. Overall, the adaptive pricing procedure reduces the social welfare cost by 2.387 (17.8%).

3.5 Conclusion

In this chapter, we consider the problem of using price signals to coordinate the electricity consumption of a group of users. We develop a two-time-scale incentive mechanism that alternately updates between the users and a system operator. We assume that users can

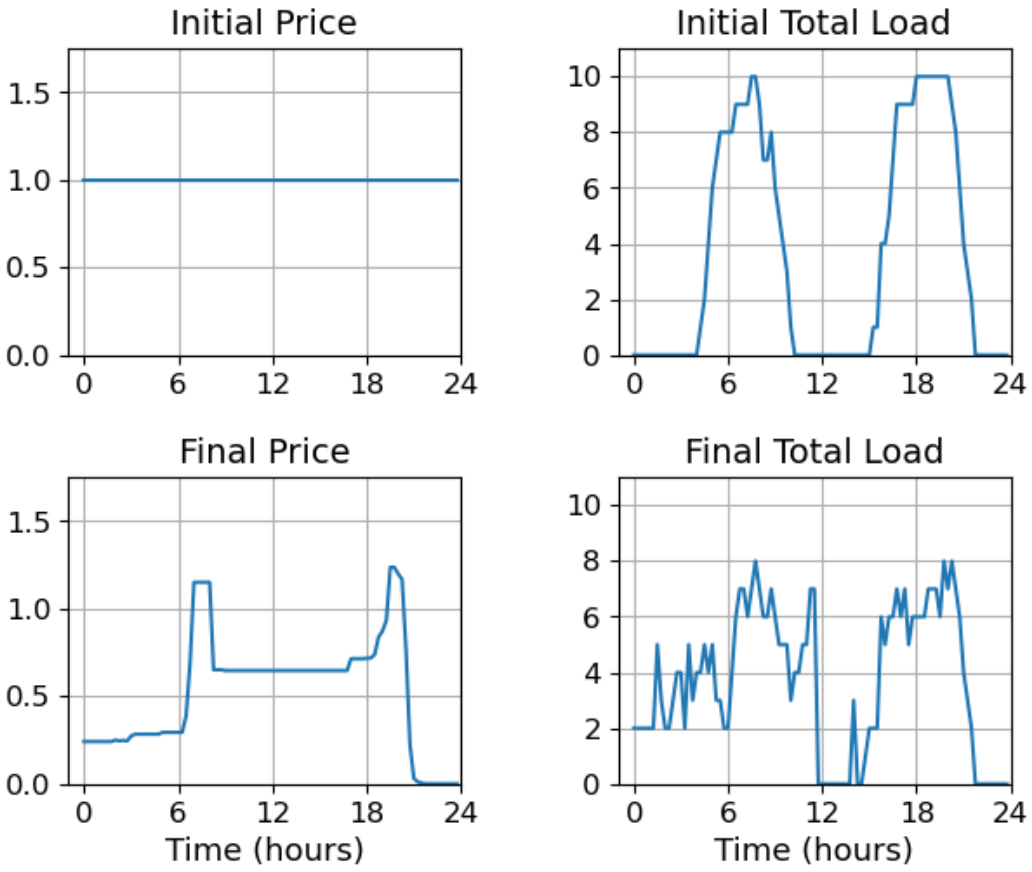


Figure 3.4: Simulation results for water heater optimization. There are 10 users minimizing discomfort using a Q-learning algorithm [1], and the system operator tries to minimize the peak load. The initial price is chosen to be flat, which leads to a profile with high peaks. After several iterations, the price becomes uneven, and the final load has much lower peaks.

optimize their actions given a price, and no private information is exchanged between the users and the system operator. In turn, the system operator does not need to learn the cost of the users. We show that the iterative algorithm converges to the socially optimal solution.

The algorithms proposed in this chapter can be implemented by any smart device that can receive and respond to information from a system operator. For example, a home with a Nest thermostat could easily negotiate with a system to reach the price equilibrium without any manual intervention of a user. A key assumption in this chapter is that each of the users are price takers, in the sense that they lack any market power to explicitly manipulate the prices. Thus an important future direction for us is to consider the price-anticipatory behavior of users.

Chapter 4

Balancing Fairness and Efficiency in Energy Resource Allocations

4.1 Introduction

Distributed energy resources (DERs), such as small-scale solar and wind generation, electric vehicles, and batteries, are crucial components of the clean energy transition; they enable end-users to actively participate in the energy market by generating, storing, and potentially selling electricity back to the grid [40]. However, individual users often cannot directly interact with the larger electricity market, facing barriers because of the complexity of energy markets, lack of economies of scale, and high transaction costs [41]. Therefore, energy aggregators help overcome the barriers faced by individuals by negotiating for power on their behalf and distributing both the costs and benefits amongst the group.

The goal of an aggregator is typically to maximize the efficiency of the group of users. More specifically, it maximizes the total surplus (utility minus payment) of the users [42, 43, 44, 45]. A number of algorithms, both distributed and centralized (see e.g., [28, 46, 47] and the references within), have been proposed over the years. However, focusing solely on efficiency may lead to large asymmetries in the allocation and surplus of the users. In short, the allocation can be unfair. For example, the results in [48, 49], as well as our own findings in this paper, demonstrate that maximizing efficiency may result in disproportionately more energy being allocated to households or businesses with a higher willingness and ability to pay, leaving fewer resources for those with lower incomes.

The need for fairness consideration in the energy domain has been recognized and has gained importance in recent years. This highlights the need for a more comprehensive

approach that balances efficiency and fairness in energy resource allocation. For example, [50] highlights the importance of fairness by estimating the impact the Department of Energy's Office of Energy Efficiency and Renewable Energy's investments have on disadvantaged communities and minority-serving universities. In the energy justice literature, distributional justice examines the fair allocation of energy benefits and burdens [51, 52]. While these studies have provided valuable insights into fairness and equity issues in energy systems, they have primarily focused on qualitative evaluations of the outcomes of particular allocation policies. There is a need for a quantitative framework that enables rigorous analysis and optimization of different allocation strategies.

This paper makes two main contributions towards this goal. First, we formalize the problem of fair energy resource allocation, providing a framework for studying fairness in the context of energy systems. This framework allows aggregators to trace out a portion of the Pareto front and explore the optimal trade-offs between efficiency and fairness. Second, we generalize the resource allocation problem to involve jointly optimizing the total resources to allocate and the allocation to individual users. This generalization leads to new theoretical and computational challenges. In particular, the joint optimization problem is, in general, not jointly convex, which makes it difficult to solve directly using standard optimization techniques; however, we show that it can be solved effectively by searching over convex subproblems.

Our work is similar in spirit to [53], which introduced the concept of energy collectives—a community-based market structure—that can be used to encourage fairness among market participants. However, [53] did not explicitly model users' surplus and can lead to suboptimal tradeoffs between different fairness measures. In addition, the participants are restricted to quadratic utilities. Our approach in this paper takes a broader scope and addresses the challenges of finding optimal fairness and efficiency tradeoffs between the users.

Fair resource allocation has been widely studied in various domains, including wireless communications [54, 55], networks [56], and machine learning [57]. However, fair resource allocation in the energy domain has received relatively less attention. The unique characteristics of energy systems, such as the price being used in actual payments (instead of shadow prices in communication networks), make the problem of fair energy resource allocation particularly interesting and challenging. Specifically, resource allocation problems have traditionally focused on fixed users' utilities; however, we focus on users' surpluses where the price for the resource is a function of the total resources allocated. Optimizing over surpluses allows for a more complete analysis that captures the relationship between resource allocation and pricing decisions, which is particularly relevant in the energy

sector.

The remainder of the paper is structured as follows: Section 5.2 formalizes the fair energy resource allocation problem; Section 4.3 reviews the concept of α-fairness and the price of fairness and price of efficiency metrics; Section 4.4 discusses our theoretical results; Section 5.4 provides numerical results and analysis; and Section 5.5 concludes the paper and offers future research directions.

4.2 Problem Formulation

We study the problem of allocating a group of users' surplus to reach a desirable level of fairness and efficiency. Consider a system involving N users and a central decision maker: the central decision maker decides on the total resources to purchase and allocates the resources to users with the objective of maximizing the chosen system performance metric. This central decision maker is also called the aggregator and we use these two terms interchangeably in the paper.

Denote the utility of user i to be U_i and the amount of energy allocated to user i as x_i. The unit price of energy faced by the users is denoted as p. Given p, the user surplus of user i is defined as

$$s_i = U_i(x_i) - px_i. \tag{4.1}$$

For N users, we define the surplus profile $\mathbf{s}$ to be the vector of each user's surplus s_i, and the allocation profile $\mathbf{x}$ to be the vector of each user's allocated energy x_i.[1]

Let $l = \sum_i x_i$ be the total energy purchased by the aggregator and let $C(l)$ be the cost of procuring this amount. As standard practice, we assume that C is differentiable and set the price of energy to be to the marginal cost $p = C'(l)$ [58]. In this paper, we actually only make use of $C'(l)$. Therefore, our results apply to markets where only the price of energy is given. The total payment from the users to the aggregator, and from the aggregator to the market, is $l \cdot C'(l)$. Quadratic functions are commonly used to model costs, and these lead to prices that are affine in demand.

In this paper, we make the following assumption:

Assumption 2. *Each utility function $U_i(x_i)$ is concave in x_i, and $U_i(0) = 0, \forall i \in \{1, \ldots, N\}$. Furthermore, the function $C(l)$ is convex in l and $C(0) = 0$.*

This assumption is very common in economic modeling for networked systems [59].

[1] Throughout the paper, we use bold to indicate vectors and matrices.

To account for fairness among the users, we make use of a fairness measure, denoted as $\Phi : \mathbb{R}^N \to \mathbb{R}$. We will introduce the format of Φ in more detail in the next section. In the following, we first introduce the optimization problem of interest:

$$\begin{aligned} \max_{\mathbf{s}} \quad & \Phi(\mathbf{s}) \\ \text{s.t.} \quad & \mathbf{s} \geq 0 \end{aligned} \tag{4.2}$$

The optimization problem in (4.2) is maximizing the fairness measure over the feasible set of nonnegative surplus. We sometime use $\mathcal{S} = \{s_i \geq 0,\ \forall i\}$ to denote this feasible set.

For the convenience of the follow-up analysis, it is easier to work with the allocations $\mathbf{x}$ and the load l, rather than directly with the surpluses. Define the feasible set of $(\mathbf{x}, l)$ be $\mathcal{X} = \{(\mathbf{x}, l) \mid \sum_i x_i = l, \quad \mathbf{x} \geq 0, \quad \mathbf{s}(\mathbf{x}, l) \geq 0\}$.

An interesting fact is that $\mathcal{S}$ is not necessarily convex, even for quadratic utility and quadratic cost functions (see Fig. 4.1). Therefore, it is not immediately clear that the problem (4.2) can be solved, regardless of the fairness measure. These types of issues have been part of the reason that it is not trivial to apply results about fairness from other domains to energy. It turns out we can solve it by directly working with allocations $\mathbf{x}$ and the load l. Define the feasible set of $(\mathbf{x}, l)$ be $\mathcal{X} = \{(\mathbf{x}, l) \mid \sum_i x_i = l, \quad \mathbf{x} \geq 0, \quad \mathbf{s}(\mathbf{x}, l) \geq 0\}$, and we study

$$\begin{aligned} \max_{\mathbf{x}, l} \quad & \Phi(\mathbf{s}(\mathbf{x}, l)) \\ \text{s.t.} \quad & \sum_i x_i = l \\ & \mathbf{s}(\mathbf{x}, l) \geq 0 \\ & \mathbf{x} \geq 0. \end{aligned} \tag{4.3}$$

We denote the optimal solution to the problem in (4.3) as $(\mathbf{x}^*, l^*)$, and use $\mathbf{s}^*$ to denote the optimal solution to (4.2). We show that (4.3) can be efficiently solved in the next section.

The format of Φ trades off fairness and efficiency and the central decision maker chooses Φ depending on the performance requirements of the system. In the next section, we specify Φ using the notion of α-fairness.

4.3 Fairness Measures

4.3.1 α-fairness

In this section, we detail the form of Φ that we use in this paper. We use the notion of α-fairness [60], which provides a parametric family of functions that includes three widely used fairness measures: social welfare, proportional fairness, and max-min fairness. The idea of α-fairness is to provide a unified framework in which the aggregator can tune the level of fairness by adjusting the α parameter, with higher α values producing more "fair" allocations. The α-fairness is defined as

$$\Phi(\mathbf{s}) = \begin{cases} \sum_i \frac{s_i^{1-\alpha}}{1-\alpha} & \text{for } \alpha \geq 0,\ \alpha \neq 1, \\ \sum_i \log(s_i) & \text{for } \alpha = 1. \end{cases} \tag{4.4}$$

In the following subsections, we describe the three common special cases of α-fairness.

Social Welfare

When $\alpha = 0$, $\Phi(\mathbf{s}) = \sum_i s_i$, which is called the utilitarian objective, and the corresponding solutions are called the social welfare solution, denoted as $\mathbf{x}^{\mathcal{SW}}$. This solution is when the central decision-maker maximizes the sum of the surplus. This is considered to be the most "efficient solution" [60, 58].

Proportional Fairness

When $\alpha = 1$, the resulting optimization problem is said to be maximizing the proportional fairness of the surpluses. It is also the generalized Nash bargaining solution for multiple users [61]. Proportional fairness can be intuitively understood as giving each user a proportional share of the resources based on their surpluses. It provides a compromise between efficiency and fairness, as it balances the total surplus with the individual surpluses of the users. Proportional fair allocation of user surplus, denoted as $\mathbf{s}^{\mathcal{PF}}$, should satisfy: compared to any other feasible allocation of user surplus, the aggregated proportional change is less than or equal to 0. In mathematical terms,

$$\sum_i \frac{s_i - s_i^{\mathcal{PF}}}{s_i^{\mathcal{PF}}} \leq 0, \forall \mathbf{s} \in \mathcal{S}. \tag{4.5}$$

We state a simple lemma that shows that setting $\alpha = 1$ in (5.3) does give solutions that satisfy (4.5), and the corresponding energy allocation $\mathbf{x}$ such that $\mathbf{s}(\mathbf{x}, \sum_i x_i) = \mathbf{s}^{\mathcal{PF}}$ is

denoted as $\mathbf{x}^{\mathcal{PF}}$.

Lemma 4. *The proportional fair surplus profile denoted as $\mathbf{s}^{\mathcal{PF}}$, can be obtained as the optimal solution to the following optimization [61, 62]:*

$$\begin{aligned} \max_{\mathbf{s}} \quad & \sum_i \log(s_i) \\ s.t. \quad & \mathbf{s} \in \mathcal{S} \end{aligned} \tag{4.6}$$

Proof. By first-order optimality condition, the objective of optimization problem (4.6) can be written as

$$\left\langle \mathbf{s} - \mathbf{s}^{\mathcal{PF}}, \nabla \sum_i \log \mathbf{s}_i^{\mathcal{PF}} \right\rangle = \sum_i \frac{s_i - s_i^{\mathcal{PF}}}{s_i^{\mathcal{PF}}} \leq 0, \forall \mathbf{s} \in \mathcal{S}$$

This inequality matches (4.6). □

Max-min Fairness

When $\alpha = \infty$, we obtain the max-min allocation is denoted as $\mathbf{x}^{MaxMin}$. This solution maximizes the worst-case surplus for each user in the system, and is sometimes called the egalitarian solution since it is considered to be the most "fair".

4.3.2 Pareto Efficiency

To signify how each different α leads to different fair surplus allocations, we first present the definition of Pareto optimality and Pareto front [63, 64].

Definition 2 (Pareto Optimality)**.** *A feasible surplus profile $\mathbf{s} \in \mathcal{S}$ is Pareto optimal if there doesn't exist another set of feasible $\bar{\mathbf{s}}$ such that*

$$s_i \leq \bar{s}_i, \forall i$$

with at least one inequality strict.

Definition 3 (Pareto Front)**.** *The Pareto front is the set of all Pareto optimal surpluses.*

Pareto optimality captures the notion of optimal trade-offs: no user can improve its surplus without decreasing other users' surpluses. Because Φ is an increasing function for all $\alpha \geq 0$, if a surplus profile is not on the Pareto front, it is neither efficient nor fair, since there are strictly better solutions for all values of α.

To illustrate how different fairness notions lead to different surplus allocations within $\mathcal{S}$, as shown in Fig. 4.1 we visualize the feasible set of surplus for an example 2-user systems with quadratic utility functions and plot the optimal surplus profile under a selected set of fair objective. As shown in the plots, the feasible surplus region is not always convex. Depending on the formats of users' utilities, the shape of the Pareto front, as well as the distribution of optimal α-fairness solutions along the Pareto front, differ from one another.

Such Pareto front [65] trade-off curves allow the aggregator to make informed decisions on how to balance fairness and efficiency among the users in designing an appropriate objective. It's important to note that we only explore certain regions of the Pareto front. Some points on the Pareto front are neither efficient nor fair under our chosen metric. In particular, the aggregator should only explore between utilitarian and max-min points by choosing different α.

4.3.3 Efficiency Measures

We leverage the price of fairness (PoF) and the price of efficiency (PoE) [66, 67] to quantitatively measure the efficiency-fairness trade-offs in our systems.

Given the feasible surplus set $\mathcal{S}$, denote the optimal utilitarian objective value as $SYSTEM(\mathcal{S})$, the fair objective under Φ as $FAIR(\mathcal{S}, \alpha)$; that is, $SYSTEM(\mathcal{S}) = FAIR(\mathcal{S}, 0)$. By definition,

$$PoF(\mathcal{S}, \alpha) = \frac{SYSTEM(\mathcal{S}) - FAIR(\mathcal{S}, \alpha)}{SYSTEM(\mathcal{S})}$$

The price of fairness quantifies the relative decrease in total user surplus when using a fair allocation compared to the utilitarian solution. In other words, it measures the relative reduction in overall efficiency. Denote the Max-Min allocation as $\max_{\mathbf{s}\in\mathcal{S}} \min_i s_i$ and for each α-fair surplus allocation, denoted as $\mathbf{z}(\alpha)$

$$PoE(\mathcal{S}, \alpha) = \frac{\max_{\mathbf{s}\in\mathcal{S}} \min_i s_i - \min_i z_i(\alpha)}{\max_{\mathbf{s}\in\mathcal{S}} \min_i s_i}$$

The price of efficiency is the relative decrease in the minimum surplus of the users under a given allocation compared to the max-min fair allocation (the most "fair" allocation [67, 63]). The price of efficiency measures the relative reduction in the surplus of the worst-off user.

4.4 Optimization and Exploring the Pareto Front

4.4.1 Fairness Metric and Feasible Surplus Region

Note that for any value of $\alpha \geq 0$, the Φ function is concave and monotonically increasing [68]. The feasible surplus region $\mathcal{S}$ is typically assumed to be convex, thus making optimizing Φ over $\mathcal{S}$ a convex optimization problem. However, as shown in Fig. 4.1, $\mathcal{S}$ is not convex even for simple utility and cost functions. In the following, we work with (4.3) and optimize directly over $\mathbf{x}$ and l, which leads to more tractable problems.

4.4.2 Optimization Characterization

Optimizing over $\mathbf{x}, l$ jointly is not a convex optimization because of the product between $C'(l)$ and x_i's. We propose to optimize over $\mathbf{x}$ and l separately in an iterative fashion. Given l, the optimization problem in (4.7) optimizes over $\mathbf{x}$:

$$\begin{aligned} J(l) = \max_{\mathbf{x}} \quad & \Phi(\mathbf{s}(\mathbf{x}, l)) \\ \text{s.t.} \quad & \sum_i x_i = l \\ & \mathbf{s}(\mathbf{x}, l) \geq 0 \\ & \mathbf{x} \geq 0 \end{aligned} \tag{4.7}$$

The above optimization problem is clearly convex. Because l is a scalar, a grid search would find the optimal l without much difficulty, as shown in Algorithm 2.

Algorithm 2 Grid Search: Searches through a discretized set of values for l to approximate the solution for the global optimization problem (4.3).

Require: step size Δl and maximum value $l_{\max}$
0: **initialize** $\phi^* \leftarrow -\infty$ and $l^* \leftarrow 0$
0: **for** l in $[\Delta l, 2\Delta l, \ldots, l_{\max}]$ **do**
0: $\phi \leftarrow J(l)$
0: **if** $\phi > \phi^*$ **then**
0: $\phi^* \leftarrow s$ and $l^* \leftarrow l$
0: **end if**
0: **end for**
return l^* =0

In a networked setting where l is a vector, grid search could become computationally expensive. However, we make the following conjecture:

Conjecture 5. *The function $J(l)$, as defined in (4.7), is quasiconvex. In particular, it remains quasiconvex when l is a vector, as long as the relationship between x_i and l is affine.*

We empirically validated this conjecture for a large number of settings. Providing a rigorous proof is an important future direction for us.

Last, since quadratic utility functions are commonly adopted in practice and in the literature [58, 53], we provide a result on when the optimization problem is jointly convex in $\mathbf{x}$ and l under this setting.

Theorem 6. *If the utility functions of each user are concave and quadratic, C and C' are convex, and C' is twice differentiable. Then the optimization problem in (4.3) jointly concave in $\mathbf{x}, l$ for $\alpha = 0, 1, \infty$ (that is, the social welfare, proportional fair, and the max-min fairness cases).*

Proof. For quadratic utility functions we have $U_i(x_i) = -a_i x_i^2 + b_i x_i$, with $a_i > 0$. We can factor out x_i from the surplus to obtain

$$\begin{aligned} s_i(x_i, l) &= -a_i x_i^2 + b_i x_i - C'(l) x_i \\ &= x_i(-a_i x_i + b_i - C'(l)). \end{aligned}$$

Thus, each non-negativity surplus constraint can be decoupled into two separate constraints $x_i \geq 0$ and $-a_i x_i + b_i - C'(l) \geq 0$. Both of these constraints are convex, hence the feasible set is convex. Next, we look at the objective function for different values of α.

1. When $\alpha = 0$, the objective in (4.3) can be written as

$$\begin{aligned} \Phi(\mathbf{s}(\mathbf{x}, l)) &= \sum_i x_i(-a_i x_i + b_i - C'(l)) \\ &= \sum_i x_i(-a_i x_i + b_i) - l \cdot C'(l). \end{aligned}$$

   In the second line, we used $\sum_i x_i = l$. Since $\sum_i x_i(-a_i x_i + b_i)$ is concave in $\mathbf{x}$ by assumption, we focus on showing $l \cdot C'(l)$ is convex in l. As both C and C' are convex, we have $C''(l) \geq 0$ and $C'''(l) \geq 0$. Since we only consider $l \geq 0$, we have $[2](l \cdot C'(l))l = 2C''(l) + l \cdot C'''(l) \geq 0$.

2. When $\alpha = 1$, the objective function is

$$\begin{aligned}\Phi(\mathbf{s}(\mathbf{x}, l)) &= \sum_i \log(x_i(-a_i x_i + b_i - C'(l))) \\ &= \sum_i \log(x_i) + \log(-a_i x_i + b_i - C'(l)).\end{aligned}$$

   The composition of a concave function (log) and a concave function $(-a_i x_i + b_i - C'(l))$ is also concave.

3. When $\alpha = \infty$, we recognize that

$$\max_{(\mathbf{x},l)\in\mathcal{X}} \min_i a_i x_i^2 + b_i x_i - C'(l) x_i$$

   is equivalent to

$$\max_{(\mathbf{x},l)\in\mathcal{X}} \min_i \log(-a_i x_i^2 + b_i x_i - C'(l) x_i).$$

   As proved in previous case, $\log(-a_i x_i^2 + b_i x_i - C'(l) x_i)$ is concave on x_i and l. So the minimum of concave functions is a concave function on $\mathbf{x}, l$.

□

4.4.3 Pareto Efficiency

Here we state a lemma about the Pareto optimality, which allows us to restrict our attention to the Pareto front.

Lemma 7. *The optimal solution* $\mathbf{s}^*$ *to* (4.2) *(or it's the corresponding* $\mathbf{x}^*, l^*$ *to* (4.3)*), lies at the Pareto front of* $\mathcal{S}$*, which is the upper right side boundary of the set of feasible surplus.*

Proof. A function $f : \mathbb{R}^N \to \mathbb{R}$ is component-wise strictly increasing if for $\bar{\mathbf{y}} \geq \mathbf{y}$, where the inequality is strict in at least one component, we have $f(\bar{\mathbf{y}}) > f(\mathbf{y})$. The function Φ is component-wise strictly increasing for all $\alpha \geq 0$ [60].

Now suppose the solution to optimization (4.2), $\mathbf{s}^*$, doesn't lie at the upper right boundary of $\mathcal{S}$. Then there exist another feasible point $\mathbf{s}^* + \boldsymbol{\epsilon}$ where $\epsilon_i > 0$ for at least one i. Since Φ is component-wise strictly increasing, we have $\Phi(\mathbf{s}^* + \boldsymbol{\epsilon}) > \Phi(\mathbf{s}^*)$, which contradicts our assumption that $\mathbf{s}^*$ is an optimal solution to optimization problem (4.2). □

4.5 Simulation Results

In this section, we demonstrate through simulations how our modeling framework could enable the aggregator to make fair allocations. We first present a simple two-user example that shows how different fairness criteria lead to different allocations and surpluses. Next, we examine how the price of fairness and price of efficiency scale with the number of users. Finally, we provide a two-class example that demonstrates how fair allocation mechanisms (specifically proportional fairness) can help reduce disparities amongst different user groups.[2]

4.5.1 Two-user example

We now demonstrate in a simple two-user example how the social welfare solution can produce an unfair allocation while the max-min solution results in a more even allocation at the expense of efficiency and the proportionally fair solution provides a compromise between the two. In this example, the users have quadratic utilities $U_1(x_1) = -x_1^2 + 3x_1$ and $U_2(x_2) = -x_2^2 + 6x_2$. As shown in Fig. 4.1 and Table 4.1, we see that under the social welfare solution, user 1 receives most of the allocation while user 2 receives almost nothing. On the other hand, optimizing max-min fairness results in a relatively even allocation, but the efficiency (total surplus) of the system is greatly reduced. Optimizing proportional fairness gives an allocation that is more even than the social welfare solution and has a higher efficiency than that of the max-min fairness solution.

	Allocation		**Surplus**		
Criterion	**User 1**	**User 2**	**User 1**	**User 2**	**Total**
$\alpha = 0.0$ (SW)	0.187	1.125	0.281	3.375	3.656
$\alpha = 0.5$	0.427	0.911	0.527	3.003	3.530
$\alpha = 1.0$ (PF)	0.535	0.682	0.668	2.564	3.232
$\alpha = 2.0$	0.620	0.435	0.822	1.867	2.689
$\alpha = \infty$ (MM)	0.691	0.204	0.977	0.977	1.954

Table 4.1: Allocations and surpluses for the two-user example.

4.5.2 Price of Fairness and Efficiency

This section demonstrates how the PoF and PoE scale with the number of users. We run 100 experiments for various number of users. Each experiment has users with quadratic

[2]Our code for the numerical simulations can be found at `https://github.com/lijiayi9712/fair_resource_allocation`.

utilities $U_i(x_i) = -\frac{1}{2}a_i x_i^2 + b_i x_i$ with $a_i \sim 1 + \text{Unif}(0,1)$ and $b_i \sim 1 + 10 \cdot (a_i + 1) + 10 \cdot \text{Unif}(0,1)$.[3] Fig. 4.2 and Fig. 4.3 both show that the PoF and PoE increase as the number of users increases. As the PoF increases, the system becomes less efficient compared to the socially optimal solution. However, it is important to note that the PoF does not converge to 1, even as the number of users N grows large. This implies that the efficiency loss due to fairness considerations remains bounded. Similarly, an increasing PoE indicates that the system becomes less equitable as the number of users increases. For $\alpha \neq 0$, the PoE also does not approach 1, suggesting that the fairness loss compared to the max-min fair solution is limited, even as the number of users increases.

4.5.3 Two-class Example: How Fair Objectives Help

We now provide a simple example that splits users into two classes and demonstrates that the social welfare solution produces a large asymmetry in the allocation while the proportional fairness scheme produces allocations that are less one-sided.

To define the classes, first suppose energy was free. For quadratic utilities $U_i(x_i) = -\frac{1}{2}a_i x_i^2 + b_i x_i$, user i would want to consume $x_i = b_i/a_i$ energy (set the derivative of $U_i(x_i)$ to zero and solve for x_i). In this example, we assume that all users have the same desired consumption when energy is free $\bar{x} = b_i/a_i$. For a non-zero price p, user i would want to consume energy to maximize their surplus $s_i(x_i) = -\frac{1}{2}a_i x_i^2 + b_i x_i - p x_i$ or equivalently $x_i = (b_i - p)/a_i = \bar{x} - p/a_i$. Thus, the larger a_i is, the more user i would prefer not to deviate from $\bar{x}$. For the first class, we sample $a_i^{(1)} \sim \text{Unif}(1,2)$ and for the second class, we sample $a_i^{(2)} \sim \text{Unif}(3,4)$.

We run 1000 experiments with 10 users in each class and compare the distribution of allocations and surpluses under the social welfare (SW) solution and the proportionally fair (PF) solution. In Fig. 4.4, we see that under the SW solution, the users in Class 1 receive almost no resources and have close to zero surpluses while the users in Class 2 receive most of the resources and large surpluses. On the other hand, the PF solution gives almost equal resources to users in both classes and the difference between the surpluses in Class 1 and Class 2 is reduced.

We refer to the gain in allocation/surplus as how much a user's allocation/surplus changed when going from the SW solution to the PF solution. In Fig. 4.4, we see the users in Class 1 almost always benefit from moving from the SW solution to the PF solution. Some users in Class 2 also benefit from the PF solution; however, many of the users in Class 2 receive smaller allocations and lower surpluses under the PF solution.

[3]To avoid experiments with small feasible regions, we chose b_i to be large relative to a_i. This is because the constraints $x_i \geq 0$ and $s_i = -\frac{1}{2}a_i x_i^2 + b_i x_i - l x_i \geq 0$ imply that $0 \leq x_i \leq 2(b_i - l)/a_i$.

4.6 Conclusion and Future work

In this chapter, we formalized the problem of fair energy resource allocation in the context of distributed energy resources (DERs) and energy aggregators. We generalized the resource allocation problem to involve jointly optimizing the total resources to allocate and the allocation itself. By doing so, we provided a principled framework that allows aggregators to explore the trade-offs between efficiency and fairness by tracing out a portion of the Pareto front. The theoretical results, numerical simulations, and analysis presented in this paper demonstrate the effectiveness of the proposed approach in achieving fair energy resource allocation.

Our work opens up several avenues for future research. In this work, we assume the aggregator knows the users' utility functions, which may be unrealistic in some scenarios. Future research could focus on developing fair resource allocation schemes for cases where the aggregator has only partial or no knowledge of each user's utility. Additionally, applying our framework to real-world datasets and exploring methods to learn utility functions from historical data could provide valuable insights and improve the practicality of our approach. Furthermore, investigating decentralized algorithms for solving the fair energy resource allocation problem could lead to more scalable and privacy-preserving solutions.

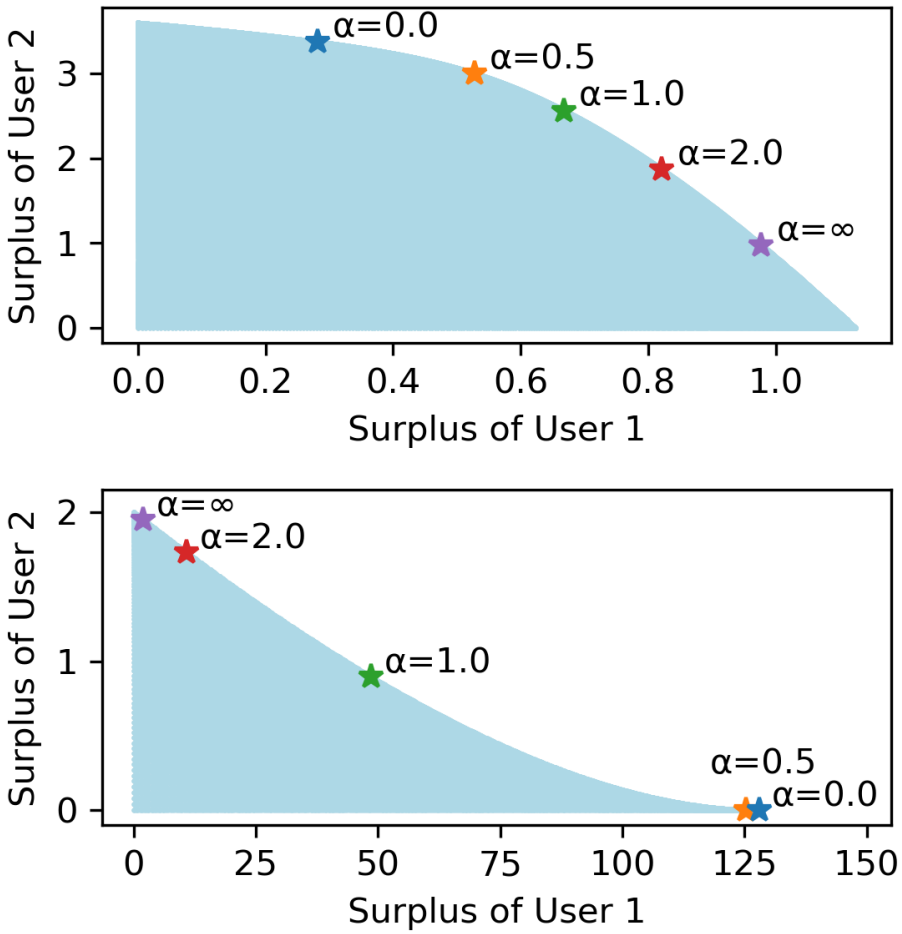


Figure 4.1: The figure illustrates the feasible regions and Pareto fronts for two-user systems with quadratic utilities. The top panel ($U_1(x_1) = -x_1^2 + 3x_1$ and $U_2(x_2) = -x_2^2 + 6x_2$) shows a convex feasible region, while the bottom panel ($U_1(x_1) = -x_1^2 + 40x_1$ and $U_2(x_2) = -x_2^2 + 4x_2$) shows a non-convex feasible region. Optimal α-fairness solutions lie on the Pareto front (the upper right boundary of the feasible region) and increasing α traces out a portion of the Pareto front starting with the least fair social welfare solution ($\alpha = 0$) to the most fair max-min solution ($\alpha = \infty$).

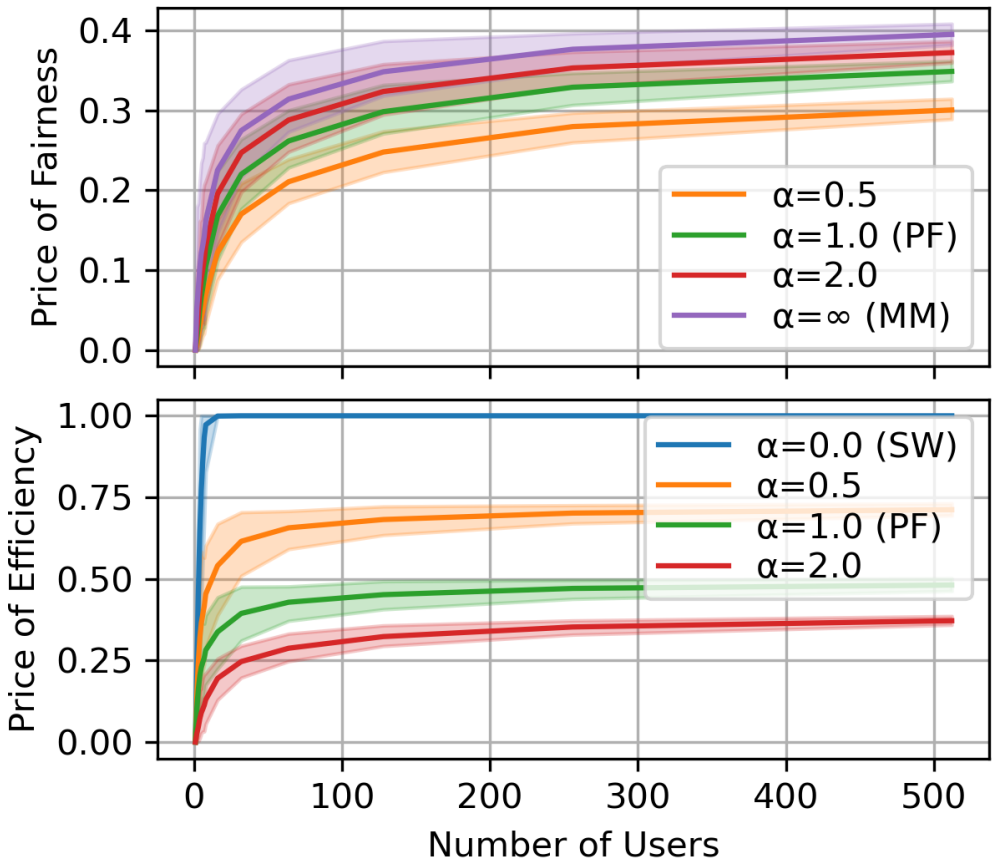


Figure 4.2: This graph plots the PoF and PoE for various α-fairness criteria as a function of the number of users. The shaded areas represent the 90% confidence interval (from the 5th to the 95th percentile) for each parameter setting. Fairness parameters closer to the socially optimal ($\alpha = 0.0$) tend to have a lower PoF and higher PoE. On the other hand, fairness parameters closer to the max-min solution ($\alpha = \infty$) tend to have higher PoF and lower PoE

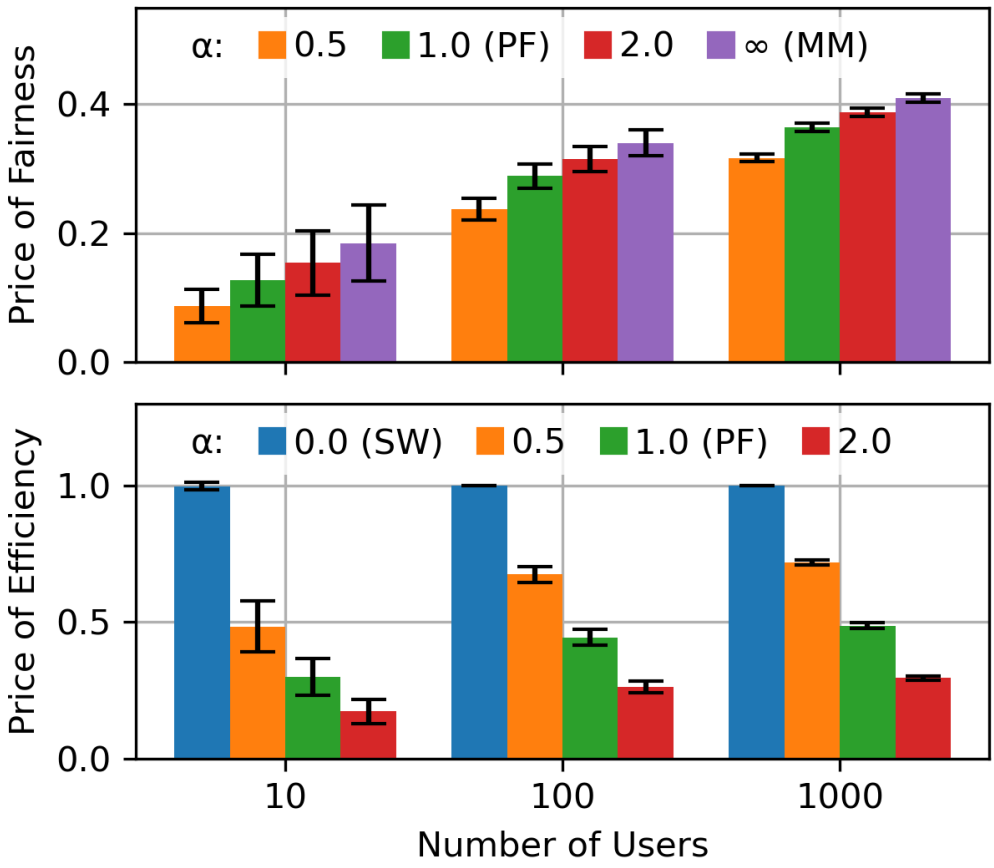


Figure 4.3: This graph plots the PoF and PoE for various α-fairness criteria and number of users. Each bar represents the mean value of PoF/PoE for a specific number of users, with error bars indicating the standard deviation. Fairness parameters closer to the socially optimal ($\alpha = 0.0$) tend to have a lower PoF and higher PoE. On the other hand, fairness parameters closer to the max-min solution ($\alpha = \infty$) tend to have higher PoF and lower PoE.

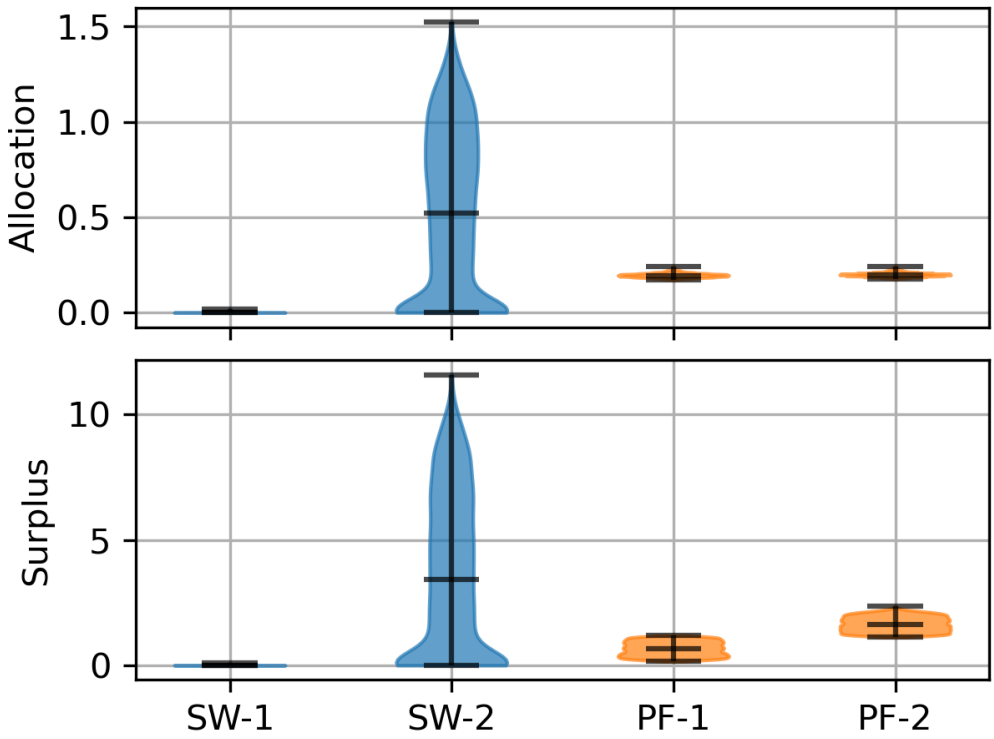


Figure 4.4: These plots compare the distribution of allocations (top) and surpluses (bottom) under the social welfare solution (SW) and the proportional fairness solution (PF) for Class 1 in (blue) and Class 2 (orange). The probability densities are shown in the shaded regions and the black lines indicate the minimum, median, and values.

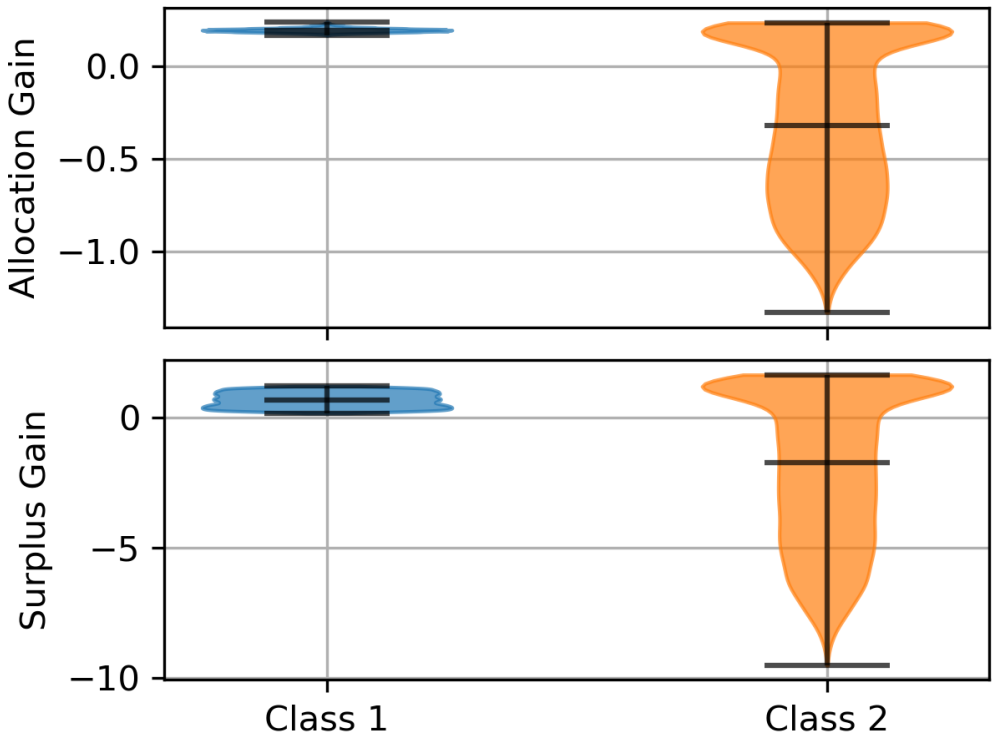


Figure 4.5: These plots illustrate the gains in allocation (top) and surplus (bottom) when switching from the social welfare solution to the proportional fairness solutions for Class 1 in (blue) and Class 2 (orange). Positive values indicate that the proportional fairness solution provides higher allocations or surpluses compared to the social welfare solution, while negative values show the opposite. The probability densities are shown in the shaded regions and the black lines indicate the minimum, median, and values.

Chapter 5

Strategic and Fair Aggregator Interactions in Energy Markets: Multi-agent Dynamics and Quasi-concave Games

5.1 Introduction

In recent years, the barriers for small users to enter the wholesale markets have been successively lowered. On the regulatory side, rulings such as the Federal Energy Regulatory Commission (FERC) Order 2222 [69] allow users to participate in electricity markets. On the technology side, users have much more flexibility and resources to change their actions to maximize their own benefits. Aggregators play a crucial role in this context by bridging the gap between individual users and the wholesale market [41, 45, 53]. They help mitigate the complexities and transaction costs that individual users would face if they directly participated in the market. At the same time, they reduce the number of agents that a system operator needs to coordinate. Therefore, a system with aggregators can be thought of as having three layers as shown in Fig. 5.1. In this chapter, we address two questions: 1) How should an aggregator allocate resources within its own group and 2) How would different aggregators compete with each other in the market.

The question of how to allocate resources within an aggregation is becoming more important since the growth of distributed energy resources (DERs) has resulted in more diverse and varied individuals. Existing research often makes the assumption that users

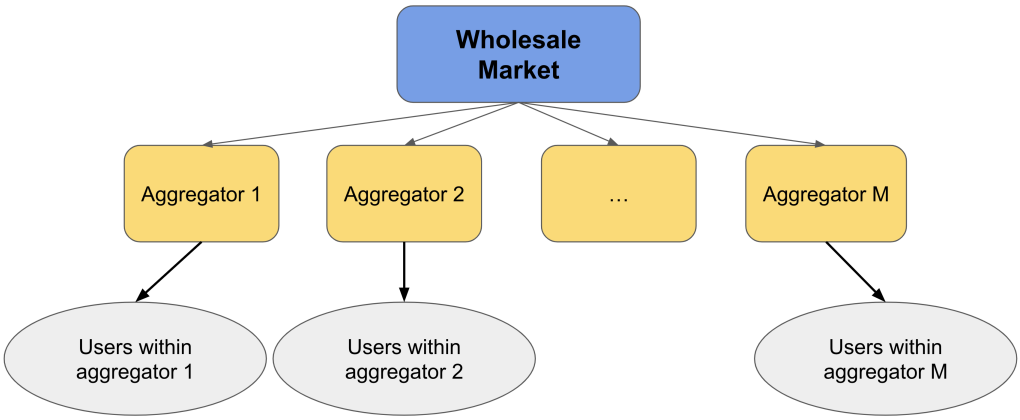


Figure 5.1: Markets with aggregators can be thought as having a three-layer architecture. The aggregators interact with the market and compete strategically with each other. Within an aggregation, the resources (or benefits) are allocated to each of the users.

within an aggregation are homogeneous or small enough that their differences can be ignored [70, 41]. However, this assumption may not be valid since users can have drastically different resources and utilities. For example, some may only be able to shift small amount of load, while others may have access to substantial solar and storage. Treating everyone the same could lead to significantly unfair division of benefits, as shown in [71, 48, 49]. To mitigate these undesirable consequences, aggregators need to take a more sophisticated approach to how they allocate resources [71].

At the level of the aggregators, they simultaneously purchase energy from a wholesale market, where the price of purchasing depends on the total purchase by all of the aggregators. Because the aggregators are (by design) not small entities, we consider the strategic interactions among them. However, moving away from the price taker assumption can lead to challenging technical difficulties since we cannot use the notion of competitive equilibrium, which is the standard solution concept in electricity markets [58, 72]. In particular, as we shown later in this chapter, even under the simplest possible settings, the interaction between the aggregators is not a concave game. In particular, the joint feasible set of actions is not convex, and this rules out a number of game-theoretic tools that are commonly employed to study the electricity market (see, e.g. [73, 74, 75] and the references within). Therefore, it is not easy to characterize whether the interaction between users would even lead to an equilibrium.

However, the proliferation of distributed energy resources (DERs) has resulted in

a more diverse set of user groups [40], challenging traditional allocation schemes that prioritize efficiency. Aggregators' exclusive focus on efficiency can lead to significant disparities in resource allocation and user surplus, as demonstrated in [71, 48, 49]. Disparities in the current market structure encourage users to join an aggregator, allowing them to benefit from better-negotiated terms [71, 42, 43, 44, 45].

5.1.1 Summary of Results and Contributions

This chapter extends our previous work [71] to a multi-aggregator setting. In [71], we considered the allocation of energy resources to users within a single aggregator under different fairness metrics. In particular, we showed how the notion of α-fairness [60, 76] can be used. In this chapter, we show that a pure strategy Nash equilibrium exists when there are multiple aggregator all strategically competing against each other.

Showing the existence of a pure Nash equilibrium is nontrivial from the inherent complexities arising from two main factors: the price dependency on total market purchases in the wholesale market, and the need for aggregators to allocate purchased energy to its users while balancing fairness and efficiency. In fact, the resource allocation problem, even when restricted to a single aggregator, is nonconvex. In this chapter, through a careful analysis of the optimization problems, we show that while the game between the aggregators is not concave, it is quasiconcave. This allows us to conclude that the game has a pure Nash equilibrium using well-known tools.

Next, we design simulations to demonstrate the interactions among aggregators in the wholesale market and the effectiveness of having an aggregator structure from both the market and the users' perspectives. These simulations provide valuable information on how aggregators can stabilize the market, ensure fair resource distribution, and optimize user surplus. In particular, we show how to balance the need for efficiency and fairness, since neither extreme leads to desirable equilibria.

The remainder of the chapter is structured as follows. Section 5.2 reviews the fair energy resource allocation problem, introduces the market structures and formalizes the multi-aggregator interaction problems; Section 5.3 analyzes the game and provides the main theoretical result; Section 5.4 provides simulation results and interpretations; Finally, Section 5.5 concludes the chapter and offers future research directions.

5.1.2 Limitations

This chapter has the following limitations in scope and results. Firstly, we establish the existence, but not the uniqueness of the Nash equilibrium. We believe that this

is mostly a technical difficulty, as we conjecture that the Nash equilibrium is, in fact, unique. Secondly, we do not consider how each aggregator is formed nor how stable they are, instead the aggregations are treated as given "as is". We plan to incorporate the work in coalition formations in electricity markets [77, 78] in future works. Third, the aggregator passes through its revenue and cost to the individual users in this chapter. This is consistent with setups where the aggregator takes a fixed "cut" of the profits, but perhaps not with setups where the aggregator delivers a promised payout and can retain the rest of the profits [79].

5.2 Problem Formulation and Preliminaries

We assume that the power system market has three layers as shown in Fig. 5.1: an upper-level wholesale market, a middle-level with aggregators, and the lower-level with individual users. This model is fairly standard, with or without middle-level aggregators [58, 41, 80].

Suppose that there are M aggregators. For the j'th aggregator, suppose that it has N_j users. We denote the energy consumption of the i'th user in the j'th aggregation as x_{ji}. Then the energy consumption of the j'th aggregator is $y_j = \sum_i x_{ji}$. The wholesale market then determines a price of electricity, p, as a function of the consumption of the aggregators $y_1, \ldots, y_M$. In this chapter, we use $\mathbf{y} = (y_1, \ldots, y_M)$ to denote the vector of aggregator consumption, and sometimes write $p(\mathbf{y})$ when we want to emphasize the dependence of p on $\mathbf{y}$. We also use the standard game-theoretic notation of $\mathbf{y}_{-j} = (y_1, \ldots, y_{j-1}, y_{j+1}, \ldots, y_M)$ to denote the consumption of the rest of the aggregators except j.

Given a price $p(\mathbf{y})$, an aggregator passes this price to its users. Instead of being price-takers as in some studies (see, e.g. [81]), we assume users solve their own optimization problems to maximize their utility. That is, a user has a utility function U, which represents the value of consuming electricity [70, 82]. Then the surplus of user i in the aggregation j is defined as

$$s_{ji} = U_i(x_{ji}) - p(\mathbf{y})x_{ji}, \tag{5.1}$$

which is a function of its own actions x_{ji} and the system-level consumption $\mathbf{y}$. Next, we will look at the allocation problem within an aggregator and the game between aggregators.

5.2.1 Allocation Problems within an Aggregation

We think of an aggregator as performing an allocation: given y_j, determine the consumption $\{x_{j1}, \ldots, x_{jN_j}$ for each of the users. Of course, a constraint is that $y_j = \sum_{i=1}^{N_j} x_{ji}$. In

addition, the allocation should be optimal in some sense, given the surplus of the users. Let $\mathbf{s}_j = (s_{j1}, \ldots, s_{jN_1}$, we model the allocation by an aggregator as solving the following optimization problem:

$$J(y_j, \mathbf{y}_{-j}) = \max_{\mathbf{x}_j} \Phi(\mathbf{s}_j) \tag{5.2a}$$

$$\text{s.t.} \sum_i x_{ji} = y_j \tag{5.2b}$$

$$s_{ji} = U_i(x_{ji}) - p(\mathbf{y})x_{ji}, \ \forall i \tag{5.2c}$$

$$s_j \geq 0 \tag{5.2d}$$

$$\mathbf{x}_j \geq 0. \tag{5.2e}$$

We will go into much more detail about the function Φ in the next paragraph. We use J as the value of the optimization problem, and it depends on the action of the other aggregators (through the price $p(\mathbf{y})$). The other constraints define the surplus and restrict the surplus and consumption to be nonnegative.

The objective function Φ greatly controls the allocation. The most popular form is the summation of surpluses, defined as $\Phi(\mathbf{s}_j) = \sum_{i=1}^{N_j} s_{ji}$. This is sometimes called social welfare [60, 76],[1] since it maximizes the total (sum) surplus. However, as we have discussed in our prior work [71], the social welfare tend tend to be unfair, in the sense that it is disproportionately biased towards users with higher utilities. In this chapter, we will again use the notion of α-fairness, defined as

$$\Phi_\alpha(\mathbf{s}_j) = \begin{cases} \sum_i^{N_j} \frac{s_{ji}^{1-\alpha}}{1-\alpha} & \text{for } \alpha \geq 0, \alpha \neq 1, \\ \sum_i^{N_j} \log(s_{ji}) & \text{for } \alpha = 1, \end{cases} \tag{5.3}$$

where we think of α as a nonegative parameter varying from 0 to ∞. When $\alpha = 0$, $\Phi_0(\mathbf{s}_j) = \sum_{i=1}^{N_j} s_{ji}$, recovering the social welfare objective; when $\alpha = 1$, $\Phi_1(\mathbf{s}_j) = \sum_{i=1}^{N_j} \log(s_{ji})$, which is the proportional fair objective; and when $\alpha = \infty$, $\Phi_\infty(\mathbf{s}_j) = \min(s_{j1}, \ldots, s_{jN_j}$, which maximizes the minimum of the surpluses. We interpret α as smoothly trading off efficiency with fairness, with social welfare being the most efficient but the least fair, and MaxMin as the most fair but the least efficient. The notion of α-fairness has not been explored in electricity markets, but it has a long history and was first developed in the networking, and interested users can refer to [60, 76].

Here we repeat an example in [71] to illustrate some properties of the optimization

[1] This objective has several names, for example, it is sometimes called the total welfare. In this chapter, we follow the literature on networking and economics and call it social welfare.

problems in (5.2). Consider two users in an aggregation where the total consumption is fixed. The aggregator solves (5.2) with respect to different values of α. Figure 4.1 shows two examples of the underlying feasible surplus values and how different α explores the Pareto-front. In particular, even for simple utility functions, the underlying feasible sets can be nonconvex.

5.2.2 Game Between Aggregators

In this chapter, we assume that each aggregator is running a fair allocation optimization problem with some α, although not necessarily all the same. They naturally form a game, where the j'th aggregator's action is y_j, and its payoff is J_j. Naturally, we restrict y_j to be nonegative. This payoff depends on the actions of all other players because of the price depends on the actions of all aggregators. We are interested in the (pure) Nash equilibria of the game, given by a point $\mathbf{y}^*$ where

$$J_j(y_j^*, \mathbf{y}_{-j}^*) \geq J_j(y_j, \mathbf{y}_{-j}^*),\ \forall\ y_j \geq 0,\ \forall\ j = 1, \ldots, M. \tag{5.4}$$

At these points, no aggregator can increase its payout by unilaterally changing its action.

Because the payoffs are value of optimization problems in (5.2), it is not clear what its equilibrium behavior is. For example, we cannot directly use some of the standard game theoretical tools to conclude whether the game has a unique equilibrium, isolated (pure) equilibria, or whether any equalibria exist. This is the case even for simple prices.

For example, suppose that $p(\mathbf{y})$ is the sum of y_j': $p(\mathbf{y}) = \sum_j y_j$. Then the payoff of the j'th aggregator, with the objective being social welfare (i.e., $\alpha = 0$), is

$$J_j(y_j, \mathbf{y}_{-j}) = \max_{\mathbf{x}_j} \sum_{i=1}^{N_j} U_{ji}(x_{ji}) - \left(\sum_j y_j\right) x_{ji} \tag{5.5a}$$

$$\text{s.t.} \sum_i x_{ji} = y_j \tag{5.5b}$$

$$U_i(x_{ji}) - p(\mathbf{y}) x_{ji} \geq 0 \tag{5.5c}$$

$$\mathbf{x}_j \geq 0. \tag{5.5d}$$

Because $\mathbf{y}$ appears in the objective as a product and is also in the constraints, J_j appears to have a complicated dependence on actions $\mathbf{y}$. In the next section, however, we show that we can get a good understanding of the equilibrium of the game through a careful analysis of the allocation problems.

5.2.3 Limiting case: each user being its own aggregator

When the number of aggregator increases to the point that it is equal to the size of the users, each user would be its own aggregator and therefore fairness-efficiency trade-off turn into its surplus optimization. The formulation for the aggregator j is detailed as:

$$\max_{\mathbf{y}_j} J_j(y_j, \mathbf{y}_{-j}) = U_j(y_j) - p(\mathbf{y})y_j \tag{5.6a}$$

$$\text{s.t. } U_j(y_j) - p(\mathbf{y})y_j \geq 0 \tag{5.6b}$$

$$\mathbf{y}_j \geq 0 \tag{5.6c}$$

$p(\mathbf{y})$ is a function on the sum of $\mathbf{y}$.

5.3 Nash Equilibria of the Aggregator Game

In this section, we study the Nash equilibria of the game between the aggregators. We do this by using the following classical result on quasiconcave games: [Existence of Nash Equilibrium for Quasiconcave Games] Given M players, with actions $y_1, \ldots, y_M$ taking values in compact intervals. Let $J_1(\mathbf{y}), \ldots, J_M(\mathbf{y})$ be the payoff functions. Then if for any player j, J_j is quasiconcave in y_j for any fixed values of $\mathbf{y}_{-j}$, then the game has a pure Nash equilibrium. This proposition dates back to the work in [83, 84, 85] and can be stated in much more general forms than the above proposition. We note that it is often given in terms of concave games, but in this chapter we do need the more general notion of quasiconcavity.

We make several standard assumptions on the utility functions of users and the price in the system.

Assumption 3 (Utility Function and System Price)**.** *a) The utility functions $U_{ji}(x_{ji})$ are concave for all users and $U_{ji}(0) = 0$.*

b) The system price $p(\mathbf{y}) = p(\sum_j^M y_j)$ and $p(\cdot) = C'(\cdot)$ for some convex function C. Equivalently, p is an continuous and increasing function.

The concavity assumption is standard and $U(0) = 0$ is a convenient way to "center" the utility functions. The assumption on p being an increasing function of total consumption is also a natural assumption. We do note that in this chapter the prices are restricted to depend on the sum of aggregators' consumption. This models the case where the system solves an economic dispatch problem to find the electricity price in the system [58]. In

this chapter, we do not consider the impact of constraints, although we believe that they do not present fundamental difficulties, and we will consider them in future work.

We state the main result in the following theorem:

Theorem 8. *Consider the game between the aggregators with payoff functions $J_1(\mathbf{y}), \ldots, J_M(\mathbf{y})$, defined using* (5.2)*. Suppose Assumption 3 holds. Then $J_j(y_j, \mathbf{y}_{-j})$ is quasiconcave in y_j for any fixed $\mathbf{y}_{-j}$, and the game has a pure Nash equilibrium.*

In fact, we make a stronger conjecture:

Conjecture 9 (Uniqueness of the Nash Equilibrium)**.** *With the same setup as in Theorem 8, the Nash equilibrium is unique.*

This conjecture is supported by extensive simulation evidence, but we are not able to obtain a rigorous proof.

Due to the fact that the definition of Φ_α takes on two different forms, we prove Theorem 8 in two steps, when $\alpha = 1$ and $\alpha \geq 0, \alpha \neq 1$. Both would be done through an analysis of the dual multipliers.

5.3.1 Proof of Theorem 8 when $\alpha = 1$

There are many equivalent ways to define a quasiconcave function. A convenient one for our purposes is

Definition 4. *Consider a continuous function $f(x) : \mathcal{I} \to \mathbb{R}$, where $\mathcal{I}$ is a bounded interval in $\mathbb{R}$. Then f is quasiconcave if there exist a number x^* such that f is nondecreasing on $\{x \in \mathcal{I} : x < x^*$ and f is nonincreasing on $\{x \in \mathcal{I} : x \geq x^*$.*[2]

For more background on quasiconcave functions, the interested reader can consult references such as [86].

When $\alpha = 1$, the payoff function is given by

$$J_j(y_j, \mathbf{y}_{-j}) = \max_{\mathbf{x}_j} \sum_{i=1}^{N_j} \log\left(U_{ji}(x_{ji}) - p(\mathbf{y})x_{ji}\right) \tag{5.7a}$$

$$\text{s.t.} \sum_i x_{ji} = y_j, \tag{5.7b}$$

[2]Note we allow the function to be monotonically increasing or decreasing, where we can take x^* to be ∞ or $-\infty$.

where we can drop the positivity constraints because of the log in the objective. Note that J_j might not be differentiable. But we first assume it is, since it illustrates the main idea of the proof.

Suppose J_j is differentiable. For a given $\mathbf{y}_{-j}$, consider y_j and $\hat{y}_j$. To show the quasiconcavity of J_j, it suffices to show: 1) if $\hat{y}_j > y_j$ and $J_j'(\hat{y}_j, \mathbf{y}_{-j}) \geq 0$, then $J_j'(y_j, \mathbf{y}_{-j}) \geq 0$; or 2) if $\hat{y}_j < y_j$ and $J_j'(\hat{y}_j, \mathbf{y}_{-j}) \leq 0$, then $J_j'(y_j, \mathbf{y}_{-j}) \leq 0$. These conditions are tantamount to saying that if the function J_j starts to decrease, it cannot increase again. Then the point y_j^* required in the definition of quasiconave functions is taken to be when the derivative becomes 0.

Taking the dual of (5.7), we have

$$\mathcal{L} = \sum_{i=1}^{N_j} \log\left(U_{ji}(x_{ji}) - p(\mathbf{y})x_{ji}\right) + \lambda(y_j - \sum_i x_{ji}), \tag{5.8}$$

where λ is the Lagrange multiplier associated with (5.7b). The first order optimality conditions are:

$$\frac{U_{ji}'(x_{ji}) - p(\mathbf{y})}{U(x_{ji}) - p(\mathbf{y})x_{ji}} = \lambda, \ \forall i = 1, \ldots, N_j,$$

and $J_j'(y_j, \mathbf{y}_{-j}) = \lambda$. Now suppose that we have $\hat{y}_j > y_j$, define $\hat{\mathbf{y}} = (\hat{y}_j, \mathbf{y}_{-j}$, $\hat{\lambda} = J_j'(\hat{y}_j, \mathbf{y}_{-j})$ to be the associated Lagrange multiplier, and $\hat{x}_{ji}$ to the associated primal variables.

Due to $\sum_i \hat{x}_{ji} = \hat{y}_j > y_j = \sum_i x_{ji}$, and all x and $\hat{x}$ are nonnegative, there exist some i such that $\hat{x}_{ji} > x_{ji}$. Now suppose $\hat{\lambda} \geq 0$, or equivalently,

$$\frac{U'(\hat{x}_{ji}) - p(\hat{y}_j, \mathbf{y}_{-j})}{U(x_{ji}) - p(\hat{y}_j, \mathbf{y}_{-j})} \geq 0.$$

But by Assumption 3, U is concave and hence U' is nonincreasing, implying that $U'(x_{ji}) \geq U'(\hat{x}_{ji})$. Since p is increasing, $p(\hat{y}_j, \mathbf{y}_{-j}) \geq p(y_j, \mathbf{y}_{-j})$. Combining the two, we have $U'(x_{ji}) - p(y_j, \mathbf{y}_{-j}) \geq U'(\hat{x}_{ji}) - p(\hat{y}_j, \mathbf{y}_{-j})$. Since the denominator $U(x_{ji}) - p(\mathbf{y})x_{ji}$ is always positive at optimality, we have $\lambda \geq 0$ if $\hat{\lambda} \geq 0$. This shows that if $\hat{y}_j > y_j$ and $J_j'(\hat{y}_j, \mathbf{y}_{-j}) \geq 0$, then $J_j'(y_j, \mathbf{y}_{-j}) \geq 0$. An analogous argument shows that $\hat{y}_j < y_j$ and $J_j'(\hat{y}_j, \mathbf{y}_{-j}) \leq 0$, then $J_j'(y_j, \mathbf{y}_{-j}) \leq 0$.

If J is not differentiable, we work with the subgradients, denoted as $\partial_s J_j(y_j, \mathbf{y}_{-j}) = \{s : J_j(\hat{y}_j, \mathbf{y}_{-j}) \geq J_j(y_j, \mathbf{y}_{-j}) + s(\hat{y}_j - y_j)\}$. Since the optimization problem in (5.7) is convex in x, $\lambda \in \partial_s J_j(y_j, \mathbf{y}_{-j})$. From the above, we have $\hat{\lambda} > 0$ implies $\lambda > 0$ when $\hat{y}_j > y_j$. Then using the subgradient definition, $J_j(\hat{y}_j, \mathbf{y}_{-j}) \geq J_j(y_j, \mathbf{y}_{-j})$, or the

function is nondecreasing. Similarly, when $\hat{\lambda} < 0$, we have $\lambda < 0$, and $\hat{y}_j > y_j$ implies $J_j(\hat{y_j}, \mathbf{y}_{-j}) \leq J_j(y_j, \mathbf{y}_{-j})$, or the function is nonincreasing. We can then find the required y_j^* required in the definition of quasiconcave functions by selecting a point that has 0 as a subgradient. This finishes the proof.

5.3.2 Proof of Theorem 8 when $\alpha \neq 1$

When $\alpha \neq 1$, the payoff function is given by

$$J_j(y_j, \mathbf{y}_{-j}) = \max_{\mathbf{x}_j} \sum_{i=1}^{N_j} \frac{(U_{ji}(x_{ji}) - p(\mathbf{y})x_{ji})^{1-\alpha}}{1-\alpha} \tag{5.9a}$$

$$\text{s.t.} \sum_i x_{ji} = y_j, \tag{5.9b}$$

$$U_{ji}(x_{ji}) - p(\mathbf{y})x_{ji} \geq 0, \quad \forall i \tag{5.9c}$$

$$x_{ji} \geq 0, \quad \forall i \tag{5.9d}$$

where we incorporate positivity constraints and similar to the $\alpha = 1$ case, we first assume J_j being differentiable to convey the main idea of the proof.

Suppose J_j is differentiable. For a given $\mathbf{y}_{-j}$, consider y_j and $\hat{y}_j$. To show the quasiconcavity of J_j, it suffices to show: 1) if $\hat{y}_j > y_j$ and $J_j'(\hat{y}_j, \mathbf{y}_{-j}) \geq 0$, then $J_j'(y_j, \mathbf{y}_{-j}) \geq 0$; or 2) if $\hat{y}_j < y_j$ and $J_j'(\hat{y}_j, \mathbf{y}_{-j}) \leq 0$, then $J_j'(y_j, \mathbf{y}_{-j}) \leq 0$. These conditions are tantamount to saying that if the function J_j starts to decrease, it cannot increase again. Then the point y_j^* required in the definition of quasi-concave functions is taken to be when the derivative becomes 0.

Taking the dual of (5.9), we have

$$\mathcal{L} = \sum_{i=1}^{N_j} \frac{(U_{ji}(x_{ji}) - p(\mathbf{y})x_{ji})^{1-\alpha}}{1-\alpha} \tag{5.10a}$$

$$+ \lambda(y_j - \sum_i x_{ji}) + \sum_{i=1}^{N_j} [\sigma_i(U_{ji}(x_{ji}) - p(\mathbf{y})x_{ji}) + \mu_i x_{ji}], \tag{5.10b}$$

where λ is the Lagrange multiplier associated with (5.9b), σ_i is the Lagrange multiplier associated with each (5.9c) and μ_i is the Lagrange multiplier associated with each (5.9d) . The first order optimality conditions are:

$$\frac{U_{ji}'(x_{ji}) - p(\mathbf{y})}{(U_{ji}(x_{ji}) - p(\mathbf{y})x_{ji})^{\alpha}} - \lambda + \mu_i + \sigma_i \left(U_{ji}'(x_{ji}) - p(\mathbf{y})\right) = 0,$$

$$\forall i = 1, \ldots, N_j.$$

For $\mathbf{x}_{ji} > 0$, we have $\mu_i = 0$ by Complementary Slackness and Dual Feasibility of the KKT conditions. In assumption 3, we assume that the utility functions $U_{ji}(x_{ji})$ are concave for all users, and $U_{ji}(0) = 0$. Following the assumptions, we have $U_{ji}(0) - 0 * p(\mathbf{y}) = 0$ and $U'_{ji}(x_{ji})$ is decreasing in x_{ji}. Moreover, each user's utility is constrained to be positive to achieve non-negative surplus, which indicates that for any feasible $p(\mathbf{y})$, $U'_{ji}(0) - p(\mathbf{y}) > 0$ and $U'_{ji}(0) > 0$.

Similarly, without loss of generality, suppose that we have $\hat{y}_j > y_j$, define $\hat{\mathbf{y}} = (\hat{y}_j, \mathbf{y}_{-j})$, $\hat{\lambda} = J'_j(\hat{y}_j, \mathbf{y}_{-j})$, $\hat{\sigma_i}$ and $\hat{\hat{\mu}}_i$ to be the associated Lagrange multipliers and $\hat{x}_{ji}$ to the associated primal variables.

Because of $\sum_i \hat{x}_{ji} = \hat{y}_j > y_j = \sum_i x_{ji}$, and all x and $\hat{x}$ are non-negative, there exist some i such that $\hat{x}_{ji} > x_{ji} \geq 0$. By Complementary Slackness, we have $\hat{\mu}_i = 0$.

Compared to the $\alpha = 1$ case, the tricky part is to handle conditions (5.9c) and (5.9d) and we handle the complexity through case-by-case analysis.

Case 1

when $U'(x_{ji}) - p(y_j, \mathbf{y}_{-j}) > 0$, $U(x_{ji}) - p(\mathbf{y})x_{ji}$ is increasing as x_{ji} increases. For $x_{ji} > 0$, $U(x_{ji}) - p(\mathbf{y})x_{ji} > 0$ and the corresponding dual variable $\sigma_i = 0$. For $\hat{x}_{ji} > x_{ji} > 0$, we also have $U(\hat{x}_{ji}) - p(\mathbf{y})\hat{x}_{ji} > 0$ with $\hat{\sigma}_i = 0$

Case 2

when $U'(x_{ji}) - p(y_j, \mathbf{y}_{-j}) < 0$, for $\hat{x}_{ji} > x_{ji} > 0$, $0 > U'(x_{ji}) - p(y_j, \mathbf{y}_{-j}) > U'(\hat{x}_{ji}) - p(\hat{y}_j, \mathbf{y}_{-j})$: $U(x_{ji}) - p(\mathbf{y})x_{ji}$ is decreasing as x_{ji} increases. Since $U(\hat{x}_{ji}) - p(\hat{\mathbf{y}})\hat{x}_{ji}$ is constrained to be positive at optimality, $U(x_{ji}) - p(\mathbf{y})x_{ji} > 0$. The corresponding dual variable $\sigma_i = 0$ and $\hat{\sigma}_i = 0$.

In both cases, at optimality point, we have $\lambda \geq 0$,

$$\lambda = \frac{U'_{ji}(x_{ji}) - p(\mathbf{y})}{(U_{ji}(x_{ji}) - p(\mathbf{y})x_{ji})^{\alpha}}$$

Now suppose $\hat{\lambda} \geq 0$, or equivalently,

$$\frac{U'(\hat{x}_{ji}) - p(\hat{y}_j, \mathbf{y}_{-j})}{(U(\hat{x}_{ji}) - p(\hat{y}_j, \mathbf{y}_{-j}))^{\alpha}} \geq 0.$$

By Assumption 3, U is concave and therefore U' is non-increasing, implying that $U'(x_{ji}) \geq U'(\hat{x}_{ji})$. Since p is increasing, $p(\hat{y}_j, \mathbf{y}_{-j}) \geq p(y_j, \mathbf{y}_{-j})$. Combining the two, we have $U'(x_{ji}) - p(y_j, \mathbf{y}_{-j}) \geq U'(\hat{x}_{ji}) - p(\hat{y}_j, \mathbf{y}_{-j})$. This shows that if $\hat{y}_j > y_j$ and $J'_j(\hat{y}_j, \mathbf{y}_{-j}) \geq 0$, then $J'_j(y_j, \mathbf{y}_{-j}) \geq 0$. An analogous argument shows that $\hat{y}_j < y_j$ and $J'_j(\hat{y}_j, \mathbf{y}_{-j}) \leq 0$, then $J'_j(y_j, \mathbf{y}_{-j}) \leq 0$.

If J is not differentiable, we work with the subgradients, denoted as $\partial_s J_j(y_j, \mathbf{y}_{-j}) = \{s : J_j(\hat{y}_j, \mathbf{y}_{-j}) \geq J_j(y_j, \mathbf{y}_{-j}) + s(\hat{y}_j - y_j)\}$. Since the optimization problem in (5.9) is convex in x, $\lambda \in \partial_s J_j(y_j, \mathbf{y}_{-j})$. From the above, we have $\hat{\lambda} > 0$ implies $\lambda > 0$ when $\hat{y}_j > y_j$. Then using the subgradient definition, $J_j(\hat{y}_j, \mathbf{y}_{-j}) \geq J_j(y_j, \mathbf{y}_{-j})$, or the function is nondecreasing. Similarly, when $\hat{\lambda} < 0$, we have $\lambda < 0$, and $\hat{y}_j > y_j$ implies $J_j(\hat{y}_j, \mathbf{y}_{-j}) \leq J_j(y_j, \mathbf{y}_{-j})$, or the function is nonincreasing. We can then find the required y_j^* required in the definition of quasiconcave functions by selecting a point that has 0 as a subgradient. This finishes the proof.

5.4 Simulation Results

In this section, we present simulation results that illustrate the impact of aggregator structures on market dynamics, particularly focusing on the interactions between large and small users, as well as the interactions among multiple aggregators. The simulations aim to demonstrate how aggregators can enhance market outcomes for small users, especially in the presence of larger users with significant market influence.

For each of the following scenarios, we simulate the long-term multi-agent interactions, including those of aggregators, large-scale users, and small-scale users. Over multiple rounds of market interactions, all participants adjust their strategies according to their optimization schemes. We create visualizations to show the evolution of equilibrium and the convergence of strategies among market participants. Finally, we analyze the outcomes in terms of different types of user surplus and the distribution of total consumption at equilibrium. The code for all simulations is available at: `https://github.com/lijiayi9712/aggregator-game`.

5.4.1 Utility Function of Users

To perform the simulations, we adopt a quadratic utility for the users, which is commonly used in the literature [87, 17, 88]. Other differentiable concave utility functions can be used and do not change the qualitative conclusions.

More concretely, user i has utility function $U_i(x_i) = -a_i x_i^2 + b_i x_i$, where a_i and b_i are

Algorithm 3 Best-Response Dynamics

0: Initialize the strategy profile $\mathbf{y} = (y_1, y_2, \ldots, y_M)$ for all players
0: **while** $\mathbf{y}$ is not a pure Nash Equilibrium **do**
0: For each player $i \in \{1, 2, \ldots, H\}$:
0: Calculate player i's best response y_i' to $\mathbf{y}_{-i}$,
0: **if** y_i' provides a better outcome for player i **then**
0: Update $y_i \leftarrow y_i'$
0: **end if**
0: **end while**
0: Return the equilibrium strategy profile $\mathbf{y}$ =0

positive numbers. Without other constraints, this function is maximized at $\frac{b_i}{2a_i}$, and we think of the ratio $\frac{b_i}{a_i}$ as representing the size of the user.

To illustrate how aggregator structure impacts smaller users' performance at equilibrium, we simulate two types of users: small-scale users (e.g. residential users) vs. large-scale users (eg. data centers, industrial factories) whose consumption is large enough to appreciably change the market equilibrium.

For small users, we draw a and b's from uniform distributions: $a_i \sim \text{Uniform}(0.1, 0.9)$ and $b_i \sim \text{Uniform}(0, 10)$. For larger users, we scale the range of the b parameter so it is much larger than a and use a variable k to denote the magnitude of the market power of the large user. Our goal is to study a market that has both large and small users and to understand how large users would impact the smaller ones. As a general rule, we usually set the large user to be as large as the "sum" of all small users. In this regime, the aggregation of users makes a significant difference in their allocations, and we observe the increase in market power of the small users when they are in an aggregate.

5.4.2 Dynamics of the Game

As proven in the previous section, each aggregator's optimization is quasi-concave, ensuring that the aggregator-aggregator interaction forms a well-defined quasi-concave game. We simulate a market with H participants, where each participant decides on the amount of energy to purchase from the wholesale market in each round. The market price is dynamically determined by the total demand of all the market participants. To understand the behavior of the system, we first investigate whether the game would converge to a pure Nash equilibrium under best-response dynamics.

Best-response dynamics simulate the process when both aggregators and individual users (large and small) play the game repeatedly, and the market dynamics evolve over

time as the game progresses through multiple iterations. In the setting, the current purchasing profile $\mathbf{y}$ represents the strategy of each participant in the market: while $\mathbf{y}$ is not a pure Nash equilibrium: There exists at least one user, say user i, who can make a beneficial new purchasing strategy y_i', given $\mathbf{y}_{-i}$, that improves its objective value $J_i(y_i', \mathbf{y}_{-i})$ in this round. Therefore, a given strategy $\mathbf{y}$ would lead to a new strategy $\mathbf{y}'$, stopping only at a pure Nash equilibrium.

In Algorithm 3, we summarize the best-response dynamics process, which iteratively updates the purchasing strategies of participants until a pure Nash equilibrium is reached. We note that if a game has multiple Nash equilibria, there is no guarantee to which one the best response dynamics will converge to. However, in all of our simulations, the best response dynamic always converges to the same Nash equilibrium, regardless of how it is initiated. This provides some numerical evidence that the Nash equilibrium is unique (as in Conjecture 1).

Given the action of the rest of the aggregators, aggregator i solves

$$\max_{y_i, \mathbf{x}} \ \Phi(\mathbf{s}) \tag{5.11a}$$

$$\text{s.t.} \ \sum_j x_j = y_i \tag{5.11b}$$

$$s_j = U_j(x_j) - p(y_i, \mathbf{y}_{-i}) x_j, \ \forall j \tag{5.11c}$$

$$s_j \geq 0 \tag{5.11d}$$

$$x_j \geq 0, \tag{5.11e}$$

where the optimization is jointly over y_i and the allocations $\mathbf{x}$. This optimization problem is not jointly convex in y_i and $\mathbf{x}$. However, it can be easily solved in a two-step process. Firstly, given y_i, the problem is convex in $\mathbf{x}$ and can be solved efficiently. Next, because of the quasiconcavity of the objective value with respect to y_i, we can use a simple bisection method to optimize over y_i [89]. Essentially, we can discretize the search space and only need to check a logarithmic number of points in the search space. Together, (5.11) can be efficiently solved.

We consider the following simulation settings:

Baseline: Small Individual Users

We first simulate a scenario in which only small users are present. Each user interacts with the market independently and we analyze the resulting equilibrium in terms of energy consumption, market price, and economic surplus. This is the limiting case where

each aggregator has only one user. This baseline setup serves as a reference point for understanding how the presence of a large user and the introduction of aggregators impact market dynamics.

We adopt the linear pricing model, that is, $p(\mathbf{y}) = c\left(\sum_j y_j\right)$, for some constant c ($c = 0.001$ in our simulations). Linear prices are commonly used in the literature [70] and other price functions can be used, as long as they are increasing, the qualitative conclusions in this section do not change. We simulate a system of 400 small users. Figure 5.2 shows the evolution of the consumption level (the x_i's) and the surplus of the users. They quickly converges to an equilibrium after a few iterations.

multi-aggregator interaction

Figure 5.3 presents an example of 2 aggregators playing iteratively: the market dynamics converge smoothly and stably to an equilibrium point very fast through the best-response dynamics update.

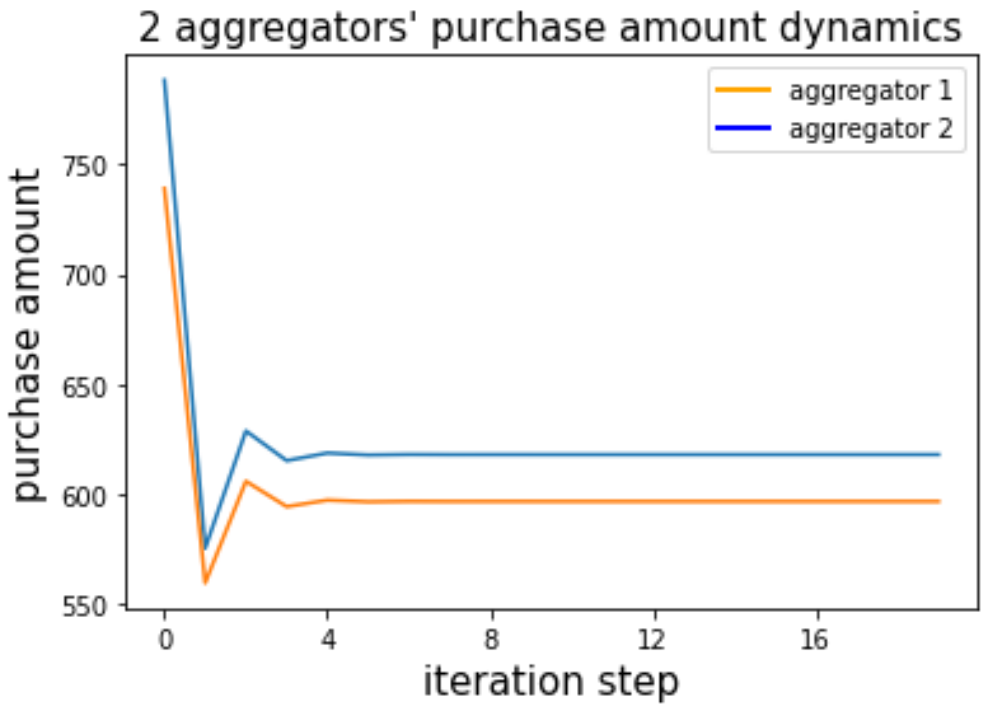


Figure 5.3: Convergence of strategic purchasing amounts for two aggregators using best-response dynamics: This figure illustrates the convergence of purchase amounts for two aggregators, each representing around 100 small users, as they interact with one large user in the market. The aggregators adjust their strategies iteratively using best-response dynamics, and the market dynamics converge smoothly to an equilibrium point within a few iteration steps. The rapid convergence demonstrates the efficiency of best-response updates in achieving equilibrium in this multi-aggregator setting.

For the rest of the simulation section, we focus on characterizing the converged market equilibrium, explore how aggregator can help the users within, and analyze how aggregators' structure, including their fairness-efficiency trade-offs, competition among aggregators, and the quantity of aggregators impact the resource allocation in the market.

5.4.3 Impact of Large Users

Next, following the baseline system that consists of only small users, we introduce a large user in the system: the large user directly competes with the small users. First, we consider the setting where the small users do not form any aggregation and there is one large user. This large user has a significant consumption level that influences the market dynamics, introduces disparities and affects the equilibrium results for the remaining small users. We compare the market's resulting surplus distribution, in the presence of

a large user of different scales, to highlight the disparities that arise when a large user operates without aggregation.

To investigate how the magnitude of the large user impacts small users at the market equilibrium, we increase the variable K from 0 to 400 to adjust the size of the single large user that, together with the $N = 200$ small users, directly participate in the market, and presents a bar-plot in Fig. 5.4. More concretely, the large user has the utility function $U(x) = -ax^2 + bx$, where a and b are positive numbers. Without other constraints, this function is maximized at $\frac{b}{2a}$, and we set $K = \frac{b}{a}$ as the measurement of the market power.

Next, we look at what happens when small users participate in the market through joining 2 aggregators that negotiate with the market on their behalf. Each aggregator aims to choose the total amount of consumption that maximizes the α-fairness objectives. We then simulate the market interactions among these two aggregators with the large user. Using different fairness measures (e.g., social welfare, proportional fairness, max-min fairness) in the aggregators' decision-making, the aggregators make different purchasing and allocation decisions under each fairness measure, leading to different market outcomes.

As demonstrated in the comparison bar-plots 5.5 and 5.6, adopting social welfare leads to the highest average surplus for smaller users, with proportional fairness leading to slightly lower values, and MaxMin fairness giving very small surplus values. A natural question is whether an aggregator would ever choose an allocation scheme other than maximizing social welfare, or put in another way, would a user ever join an aggregation that does not use social welfare?

To answer this question, we look at the distribution of the surplus and consumption of small users. Fig 5.7 shows how they change by comparing no aggregation and an aggregator that uses social welfare for allocation. With an aggregator, the distribution of the consumption is more spread out, although most of the users still receive relative little allocation, and the distribution has a long tail, showing that a few users are much better off than others.

Fig. 5.8 shows the distribution of users' consumption and surpluses when the aggregators use a proportional fair allocation. Here, the distribution is much more "uniform", with most users' consumption concentrated around the averaged value and a much shorter tail. Therefore, both social welfare and proportional fair-based allocation have their merits, and users might join an aggregator with these schemes depending on their preferences.

5.4.4 Fairness and Competition between Aggregators

We also investigate how different fairness measures affect the competition between aggregators. That is, what happens when different aggregators run different fairness schemes. We consider two aggregators in the market, and tune the α values of one or both of them from 0 to ∞.

First, Fig. 5.9 shows how the average surplus of users changes when both aggregators have the same α and the value of α increases. As α increases (the allocation with an aggregate becomes more fair), the average surplus decreases. Fig. 5.10 shows what happens when one aggregator holds its α at 0 while the other increases its α value. It shows that the α values should not be too different between the aggregators. These simulations raise an interesting question about how the α values should be set in practice, which we do not answer in this chapter but is an important issue to be addressed in the future.

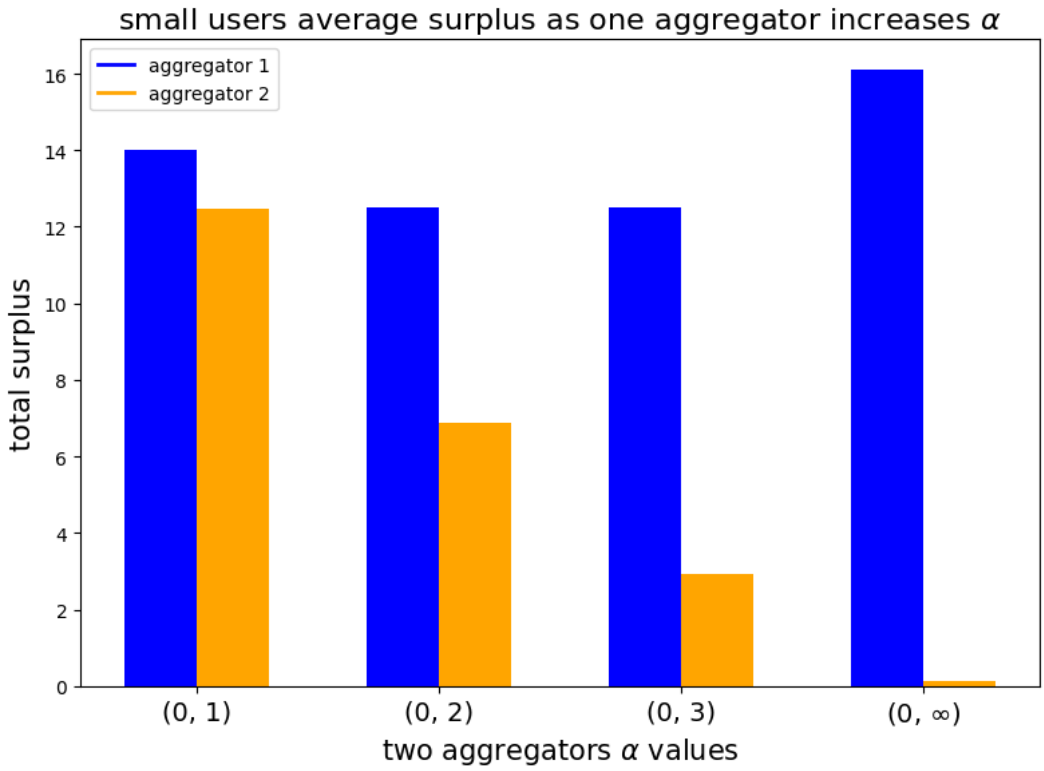


Figure 5.10: When the first aggregator's resource allocation scheme stays the same, while the second aggregator increases its fairness considerations, the more fairness second aggregator considers, the smaller the average surplus that users within the second aggregators would experience.

The results show the user surplus under each fairness measure, the total system surplus and its distribution, and the fairness metrics. The results will be visualized in plots showing the user surplus distribution under different fairness measures, the total system surplus versus the fairness measure parameter, and the fairness index versus fairness measure parameter.

5.4.5 Impact of Number of Aggregators

Understanding the effect of the number of aggregators on market outcomes is essential for optimal market design and regulatory oversight. To explore this, we conducted simulations where the number of aggregators in the market varied (e.g., 2, 4, 8, ..., N), while keeping their fairness objectives constant (set to be social welfare or proportional fairness for all aggregators). These simulations offer valuable information on how the number of aggregators influences the distribution of average surplus and consumption among small users at the equilibrium point. Where large users are present, the number of aggregators has a noticeable impact on small users' average surplus, as demonstrated in Fig. 5.11 for the social welfare objective and Fig. 5.12 for the proportional fairness objective. Although no clear monotonic trend is observed as the number of aggregators increases, the data suggest a trade-off between small users' surplus and consumption at the market equilibrium point. This trade-off is particularly relevant for market operators and regulators, who must balance efficiency with fairness while considering the average consumption upper limit imposed by generation capacity.

For market operators, these results imply that the optimal configuration of the aggregators is context dependent and may require local market experimentation to identify. The number of aggregators can influence market dynamics in ways that are not immediately apparent, affecting both the stability of market prices and the distribution of economic benefits among users. Therefore, understanding the implications of aggregator configurations is critical for designing market mechanisms that improve both efficiency and equity.

5.5 Conclusion and Future work

In this chapter, we have extended the fair energy resource allocation problem to a multi-aggregator setting, where multiple aggregators interact within an electricity market. We prove that the strategic optimization faced by aggregators forms a quasiconcave game, and we demonstrate the existence of a Nash equilibrium and conjecture on the uniqueness

of the converged Nash equilibrium. We also presents valuable insights into how aggregator market structures can promote market stability while realizing fairness and efficiency in resource allocation. Our theoretical framework, supported by simulations, illustrates how aggregators stabilize market dynamics, shape users' surplus/consumption distributions, and manage system fairness-efficiency trade-offs, particularly for small-scale users.

The simulations further validated the theoretical results, showing how different market configurations and fairness objectives influence small users' surplus and consumption distribution. The introduction of large users was shown to significantly impact the surplus of small users, which is mitigated when aggregators are introduced. Aggregators operating under proportional fairness or social welfare objectives were seen to improve the outcomes for smaller users, balancing fairness and efficiency. Furthermore, the simulations demonstrated how the number of aggregators and their fairness considerations shape the equilibrium behavior of the market.

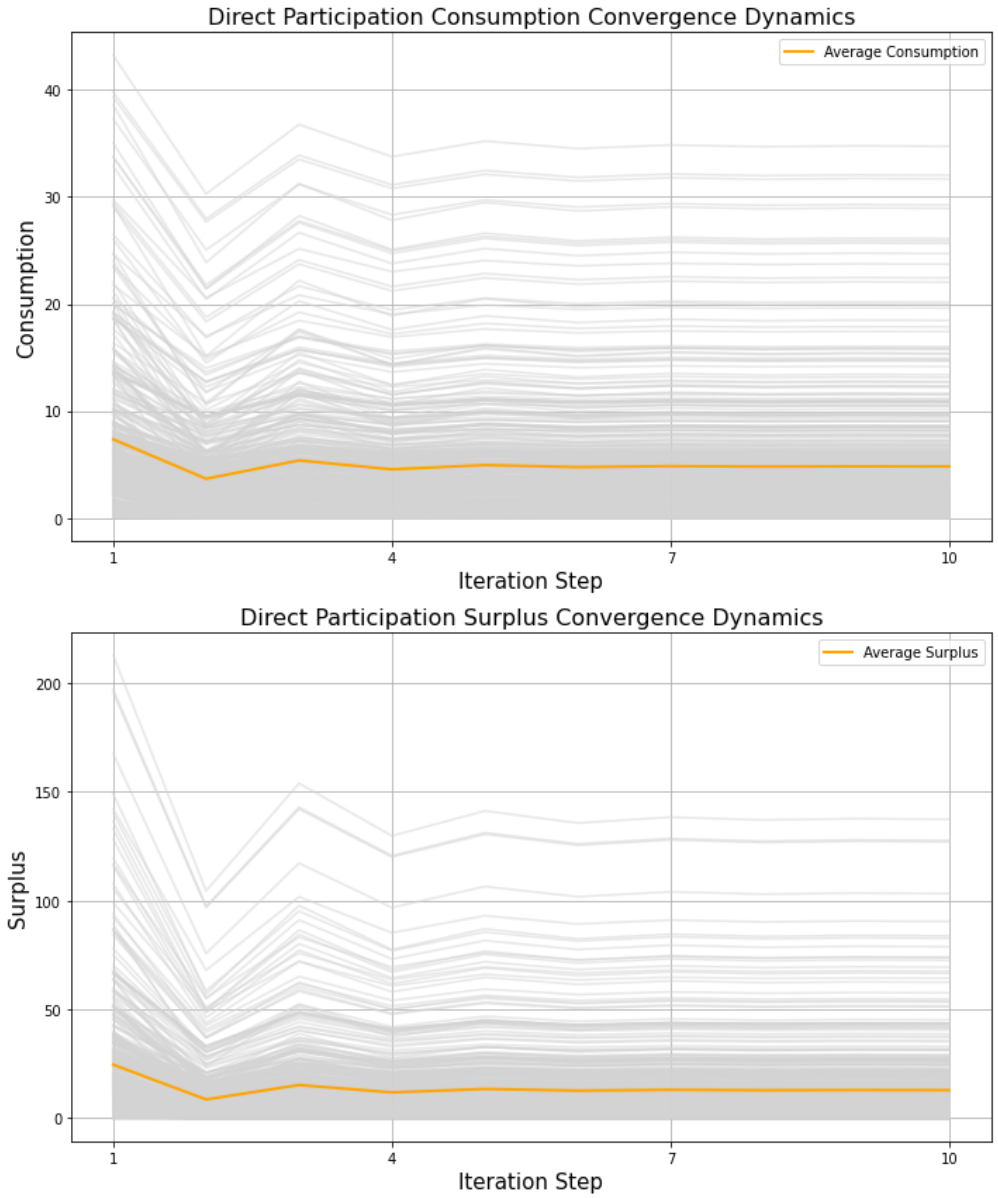


Figure 5.2: Evolution of small users' consumption and surplus under direct market participation: the top plot illustrates the convergence of consumption level over multiple iteration steps, while the bottom plot shows the convergence of surplus. Both plots display the results for 400 small users directly participating in the market. The figures demonstrate how the consumption and surplus dynamics quickly stabilize, converging to a market equilibrium.

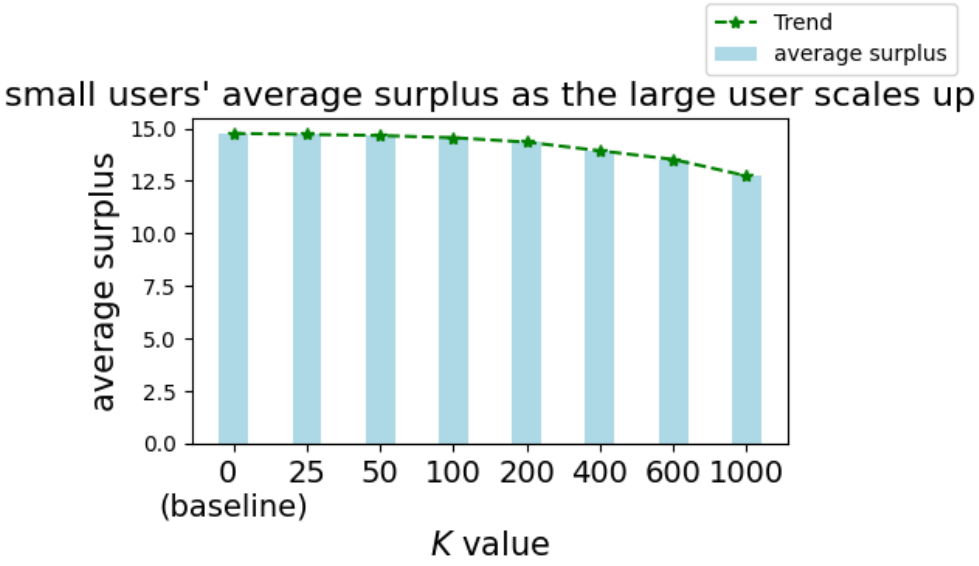


Figure 5.4: This figure illustrate how the average surplus of the N small users change as the magnitude of the large user increases by changing the value M from 0 to 400. The x-axis is the K value that signifies and y-axis represents the average surplus amount. We observe that the average surplus decreases as M increases.

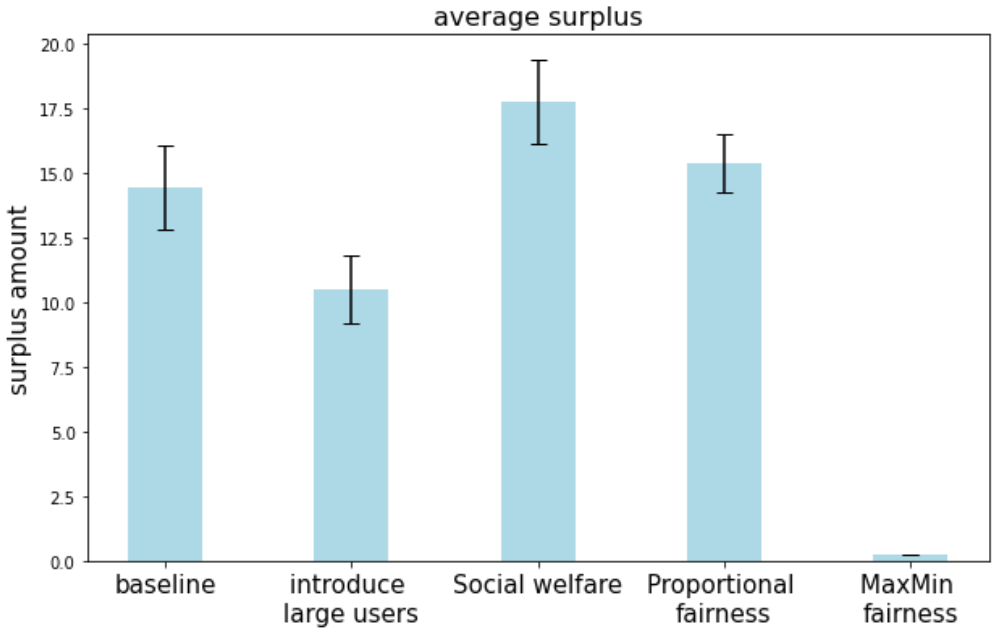


Figure 5.5: Average surplus for small users under different market configurations and fairness objectives: This figure compares the average surplus achieved by small users across various scenarios: baseline (200 small users), introduction of large users and without aggregators (with the large user having market power equivalent to 200 small users), and three fairness objectives applied within aggregators—social welfare, proportional fairness, and max-min fairness. The bars represent the average values obtained from 50 simulation runs. The results highlight how the introduction of large users and the choice of fairness objectives impact the distribution of surplus among small users, with social welfare objective leading to the highest surplus.

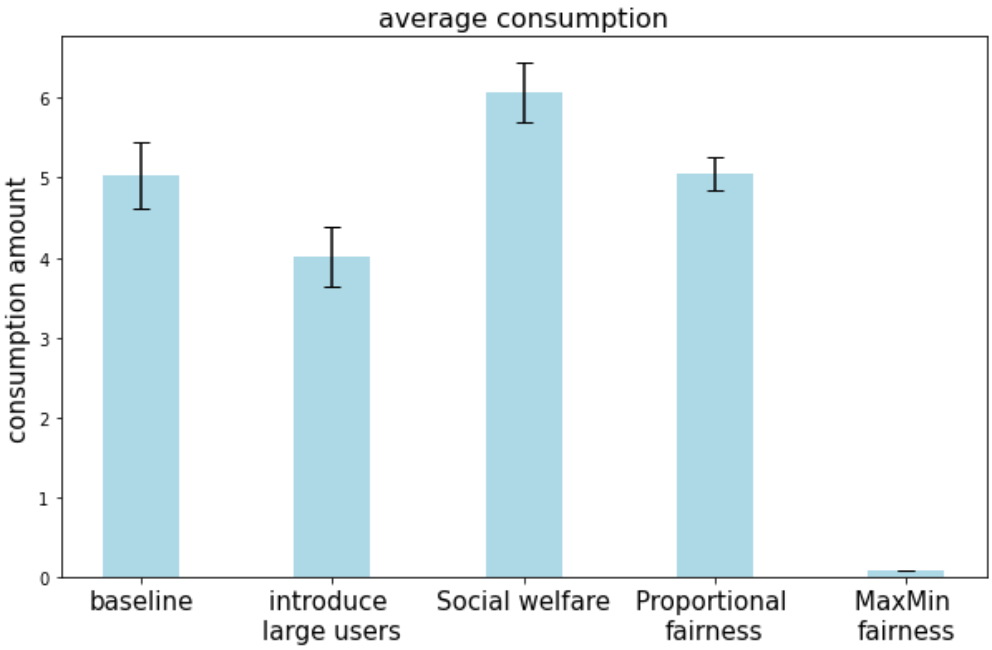


Figure 5.6: Average consumption for small users under different market configurations and fairness objectives under the same setup as in Fig. 5.5. The results demonstrate the impact of market configurations and fairness objectives on energy consumption, with social welfare and proportional fairness leading to higher consumption levels, while max-min fairness results in significantly lower consumption.

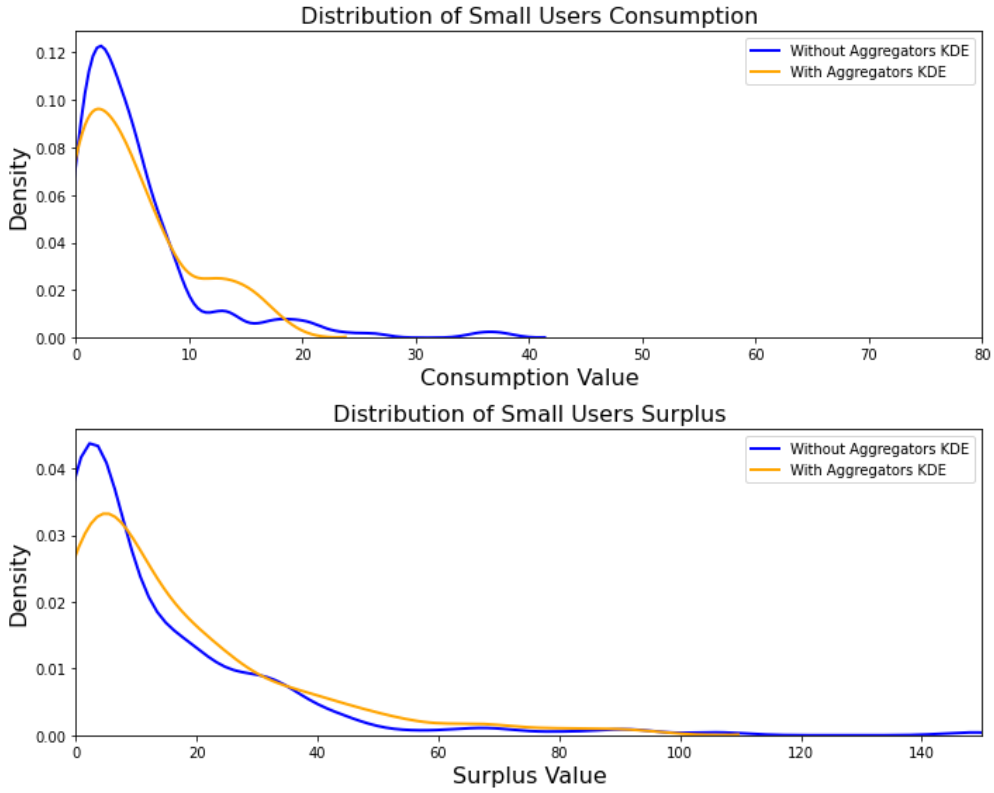


Figure 5.7: Comparison of Small Users' Consumption and Surplus Distributions With and Without Aggregators optimizing social welfare. The top plot shows the smoothed kernel density estimates (KDE) of small users' consumption values, while the bottom plot displays the KDE of their surplus values. Both distributions are compared under two scenarios: without aggregators (blue) and with aggregators which are optimizing social welfare (orange). The distribution under social welfare is more concentrated, although it still has a long tail.

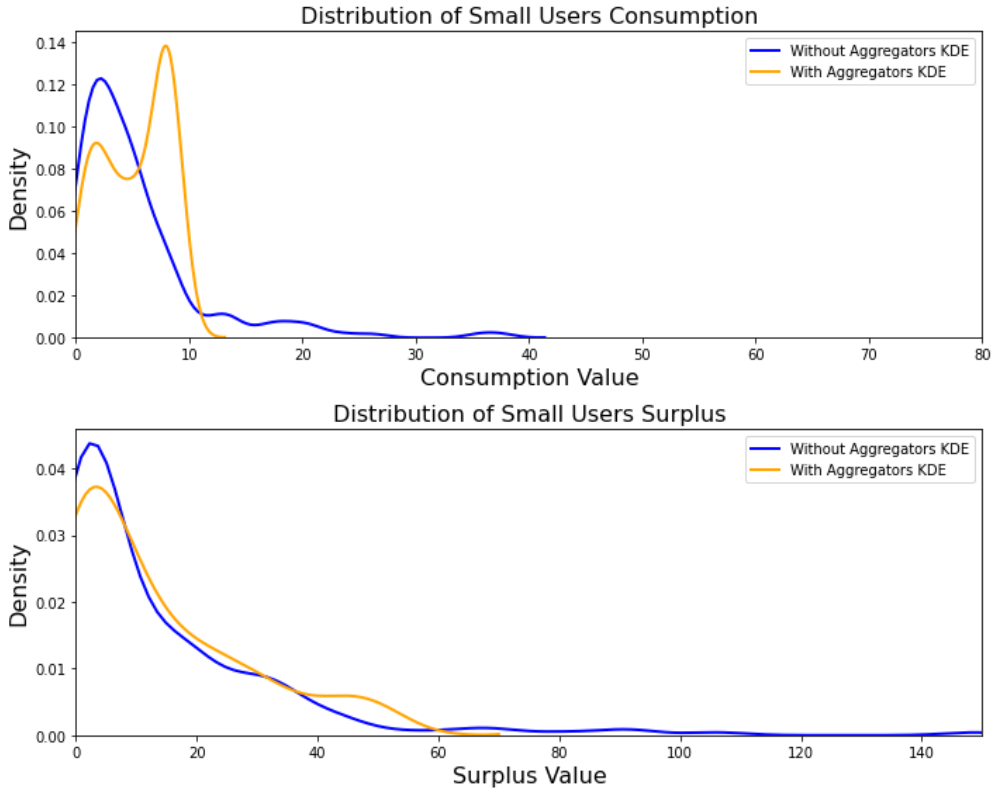


Figure 5.8: Comparison of Small Users' Consumption and Surplus Distributions With and Without Aggregators optimizing proportional fairness: The top plot shows the smoothed kernel density estimates (KDE) of small users' consumption values, while the bottom plot displays the KDE of their surplus values. Both distributions are compared under two scenarios: without aggregators (blue) and with aggregators which are optimizing proportional fairness (orange). The distribution under proportional fairness is much more uniform, with most values concentrating around the mean.

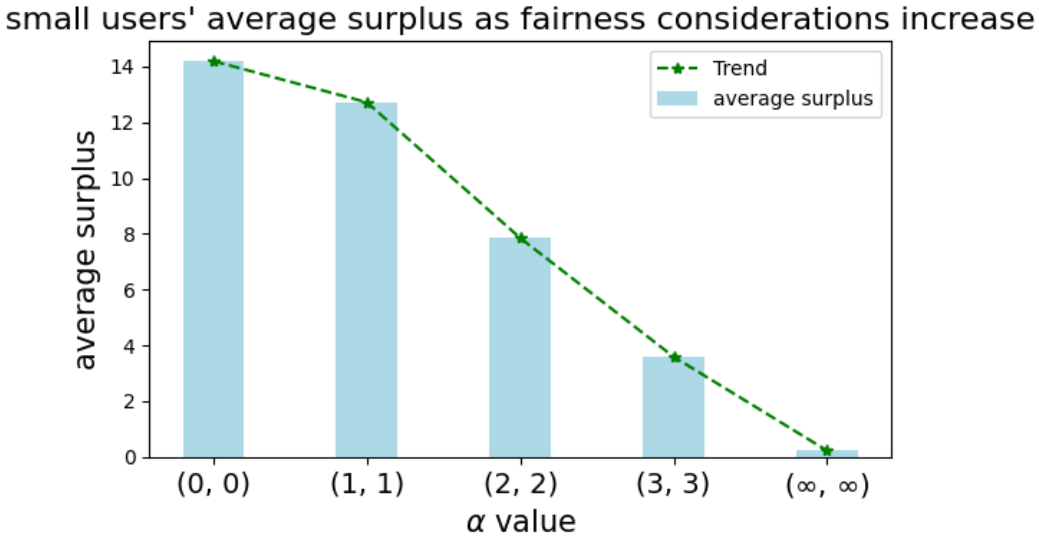


Figure 5.9: When both aggregators increase their fairness consideration in resource allocation schemes, small users' average surplus decreases.

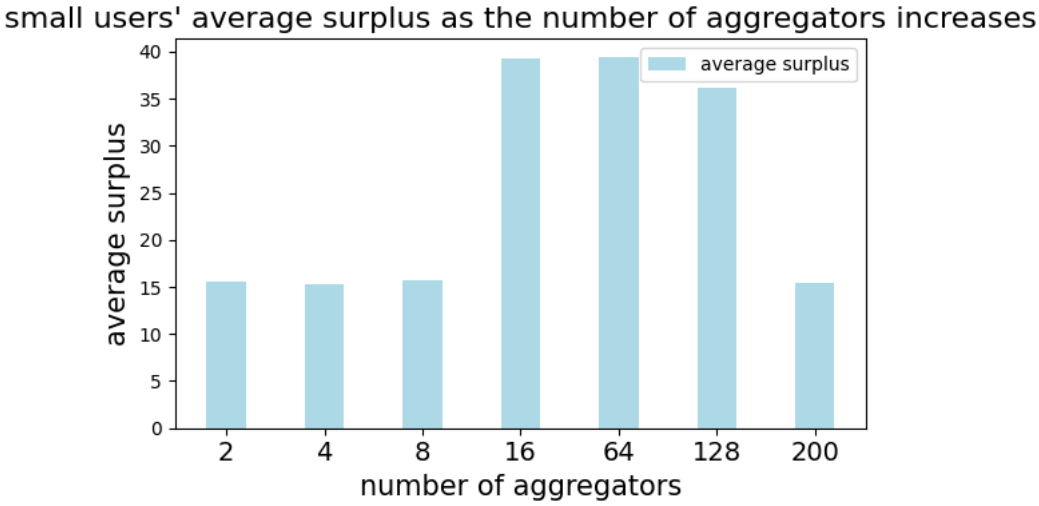


Figure 5.11: How the number of aggregators impact the small users average surplus at equilibrium when the aggregators are optimizing **social welfare**.

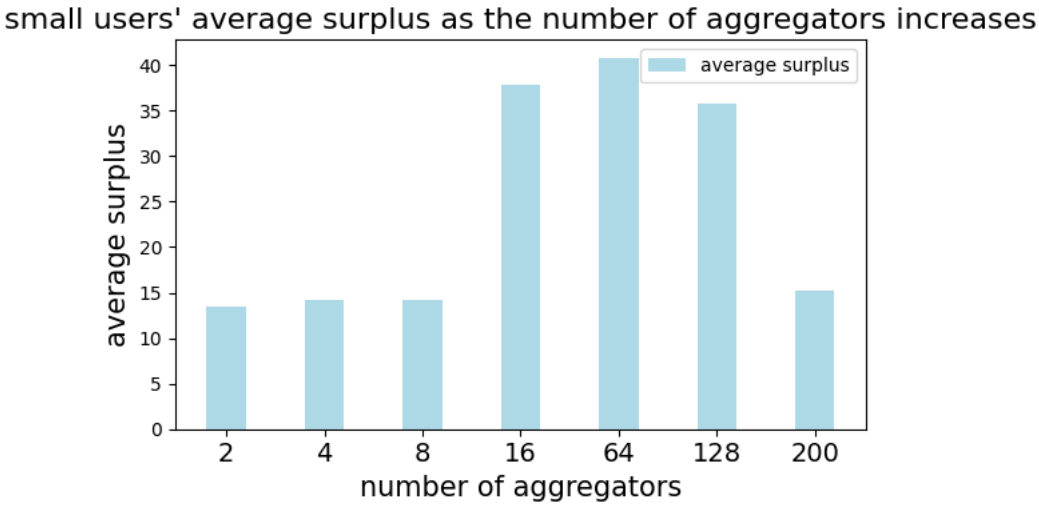


Figure 5.12: How the number of aggregators impact the small users average surplus at equilibrium when the aggregators are optimizing **proportional fairness**.

Chapter 6

Decentralized Safe Reinforcement Learning for Voltage Control

6.1 Introduction

Distributed energy resources (DERs) such as rooftop solar PV, electric vehicles and battery storage are growing at an increasing pace. For example, solar capacity had almost 50% yearly growth in 2021 [90], which is by far the fastest among all renewable resources. Most of these growth are occurring in the distribution network, the low voltage network that connects customers to substations.

High variability of solar PV and sudden change in load due to electric vehicles and storage can lead to large voltage fluctuations. These fluctuations occur at timescales much faster than the conventional mechanical control devices such as tap-changing transformers. Instead, power electronic devices allow flexible and frequent control actions without degrading lifetime. Consequently, there have been growing interests to use the power electronic inverters on the DERs themselves to provide voltage control [91, 92, 93, 94, 95].

Since most distribution networks are not yet equipped with real-time communication infrastructure, voltage control strategies should use local measurements available at each bus. More specifically, controllers need to operate at an iterative fashion [96, 97], successively updating their control actions based on each measurement. Designing such decentralized controller is a nontrivial problem. Linear controllers can be far from optimal, even for quadratic costs. Therefore, neural networks have been used to parametrize the controllers to fully utilize the capabilities of the inverters [98, 99, 100].

Reinforcement learning algorithms are proposed to train the neural network controllers with trajectory measurements. This provides the advantages of updating neural networks

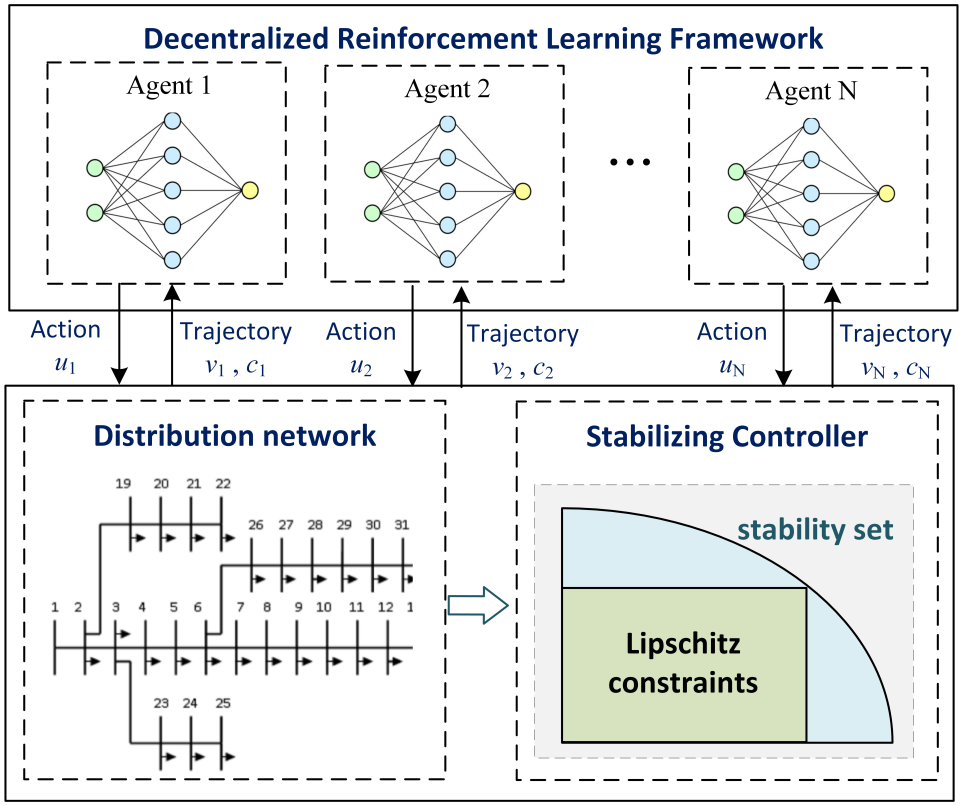


Figure 6.1: Proposed decentralized safe RL approach for optimal voltage control. We prove that the system is guaranteed to be exponentially stable if each controller satisfies certain Lipschitz constraints. The neural network controllers are engineered to satisfy these Lipschitz constraints by design, and is updated from local trajectories with a decentralized RL framework.

in a model-free setting, i.e., eliminating the requirement on system parameters and communications [101]. Many algorithms, such as deep Q learning [102], actor-critic [103], DDPG [104], have been applied to the control of tap-changing transformers or inverter based resources. Since the control actions are taken in an iterative fashion, it creates a dynamical system, whose transition depends on the actions and the underlying physical distribution network. The key constraint on the controllers is that they do not destabilize the system. However, most works neglect the stability requirement and currently this stability condition is checked through simulations [102, 103]. Considering that voltage control is implemented locally without real-time communication, formal guarantees on stability are required in practice.

This chapter presents a decentralized safe learning method, which guarantees the learned neural network would maintain the stability of iterative voltage control dynamics. We prove that the system is guaranteed to be exponentially stable if each controller

satisfies certain Lipschitz constraints. We optimize the set of Lipschitz bounds to enlarge the search space of controllers. On this basis, we propose to engineer the structure of neural network controllers such that they can satisfy the Lipschitz constraints by design. A decentralized RL framework is constructed to train neural network controller locally at each bus with policy gradient algorithm. The structure of the proposed approach is illustrated in Fig. 6.1.

Case studies show that the controllers learned with stability constraints outperform those with linear controllers and unconstrained neural network controllers. Interestingly, we also observe good learning convergence of the controllers in a model-free setting, even though they interact through the underlying distribution network. Code and data are available at https://github.com/Safe-RL-Power-Systems-Control/Voltage-Control.

The chapter is organized as follows. Section 6.2 introduces the model and the optimal voltage control problem. Section 6.3 gives the main theorems governing the structure of a stabilizing controller and derives the optimal Lipschitz bounds. Section 6.4 illustrates the decentralized safe RL framework for training a stabilizing neural network controller locally at each bus. Section 6.5 shows the simulation results and Section 6.6 concludes the chapter.

6.2 Model

A standard requirement for distribution network is that voltages should deviate no more than 5% from their rated values at all buses [105]. For example, if the rated voltage is 110 V, then the actual voltages should be in the interval from 104.5 V to 115.5 V. For simplicity, we normalize the units such that the reference value for voltage is 1 p.u. For a power network with N buses, let $\boldsymbol{v}$ be the voltage vector where v_i is the voltage at bus i. Let $\boldsymbol{p}$ be active power and $\boldsymbol{q}$ be reactive power. The voltage of the system follows the LinDistFlow model:

$$\boldsymbol{v} = \mathbf{R}\boldsymbol{p} + \mathbf{X}\boldsymbol{q} + 1 \tag{6.1}$$

where 1 is the all one's vector and $\mathbf{R}$ and $\mathbf{X}$ are positive definite matrices describing the network [96]. The active power depends on external environment and is uncertain and variable. The reactive power comes from phase offsets and is controllable, subject to some actuation constraints [91].

This work focuses on optimizing the control of $\boldsymbol{q}$ through inverter-based resources. The aim of voltage regulation is to control $\boldsymbol{q}$ such that $\boldsymbol{v}$ is close to its reference value. Due to the lack of communication in many distribution systems, $\boldsymbol{q}$ needs to be successively updated

based on the local voltage measurements. Denote $u_i(v_i)$ as the control law for each bus $i = 1, \cdots, N$, which is a mapping from the voltage to reactive power. Let $v_{i,t}$ be the local voltage at the bus i at the t-th iteration step, and denote $\boldsymbol{u}_t = (u_1(v_{1,t}), \cdots, u_N(v_{N,t}))$. We update $\boldsymbol{q}$ and $\boldsymbol{v}$ iteratively as

$$\boldsymbol{q}_{t+1} = \boldsymbol{q}_t - \boldsymbol{u}_t, \tag{6.2a}$$

$$\boldsymbol{v}_{t+1} = \mathbf{R}\boldsymbol{p} + \mathbf{X}\left(\boldsymbol{q}_t - \boldsymbol{u}_t\right) + 1, \tag{6.2b}$$

6.2.1 Optimal voltage control

Our objective is to optimize the $\boldsymbol{u}_t$ to minimize cost in $\boldsymbol{v}$ and $\boldsymbol{q}$ defined as $C(\boldsymbol{u})$, subject to the iterative update rule and the saturation limit on $\boldsymbol{u}_t$. The optimization problem is

$$\min_{\mathbf{u}} \quad C(\boldsymbol{u}) \tag{6.3a}$$

$$\text{s.t. } \boldsymbol{q}_{t+1} = \boldsymbol{q}_t - \boldsymbol{u}_t \tag{6.3b}$$

$$\boldsymbol{v}_{t+1} = \mathbf{R}\boldsymbol{p} + \mathbf{X}\left(\boldsymbol{q}_t - \boldsymbol{u}_t\right) + 1 \tag{6.3c}$$

$$\underline{\boldsymbol{u}}_t \leq \boldsymbol{u}_t \leq \overline{\boldsymbol{u}}_t \tag{6.3d}$$

$$\boldsymbol{u}_t \text{ is stabilizing} \tag{6.3e}$$

where constraints (6.3b)-(6.3e) hold for the iteration step t from 0 to T. The cost typically trades off between driving voltage to the reference value and the control effort. The deviation of voltage can typically be quantified as two-norm, one-norm or infinity-norm of the sequence of $\boldsymbol{v}_t$ [93, 106, 107]. The control effort depends on the type of resources and can be both quadratic [108, 109] and non quadratic ones [110, 106, 107]. For example, control effort from batteries is commonly defined as one-norm of actions since charging/discharging power affects cycle-depth linearly [110, 107]. The proposed safe RL approach works for all types of cost functions listed above. The lower and upper bound for the control action at bus i are $\underline{u}_{t,i}$ and $\overline{u}_{t,i}$, respectively. The subscript t signifies that these bounds can be time-varying as active power changes.

The controllers $\boldsymbol{u}$ are conventionally designed to be linear (up to a thresholding by (6.3d)), which does not leverage the capability of inverter-based resources in implementing almost arbitrary control laws [111]. To design a flexible non-linear control law for inverter-based resources, we parameterize each controller $u_i(v_i)$ as a neural network with weight θ_i, sometimes written as $u_{\theta_i}(v_i)$.

However, there remain two challenges. First, due to the lack of communication, neural network controller needs to be trained decentralizely in each bus with local observations

of voltages. Second, even if the controller is optimized and implemented locally, they need to be "safe" in the sense that the controller stabilizes the entire system, as defined by (6.3b) and (6.3c). In the next sections, we show how to design the local neural network controllers that guarantee the stability of this system, and how to train the controllers through decentralized reinforcement learning.

In this chapter, we assume that the topology and parameter information of the distribution system is available. That is, we know $\mathbf{X}$, but there is no real-time communication between the buses. This assumption comes from the fact that $\mathbf{X}$ (and $\mathbf{R}$) can be estimated using smart meter data collected over a period of time [112, 113, 114], where the communication rate can be quite slow (e.g., once per day [115]). Therefore, design of the controllers can depend on $\mathbf{X}$, but the dependence must be determined offline. The system parameters are not required for real-time training and implementation.

6.3 Stabilizing controller

In this section, we derive the properties of a stabilizing local controller from the Lyapunov stability theory and standard nonlinear system theory. We engineer the structure of neural network to satisfy these structure properties and thus guarantee the stability of the system.

6.3.1 Reduced-order system

We can simplify the dynamics in (6.2) by shifting the origin of the system. Denote $\hat{\boldsymbol{v}}_t$ as the difference between voltage and its reference value 1 at time t. We assume that the active power $\boldsymbol{p}$ remains constant during one iteration period. Then we have

$$\begin{aligned}\hat{\boldsymbol{v}}_t &= \boldsymbol{v}_t - 1 \\ &= \boldsymbol{R}\boldsymbol{p} + \mathbf{X}\boldsymbol{q}_t \\ &= \boldsymbol{R}\boldsymbol{p} + \mathbf{X}(\boldsymbol{q}_{t-1} - \boldsymbol{u}_{t-1}) \\ &= (\boldsymbol{R}\boldsymbol{p} + \mathbf{X}\boldsymbol{q}_{t-1}) - \mathbf{X}\boldsymbol{u}_{t-1} \\ &= \hat{\boldsymbol{v}}_{t-1} - \mathbf{X}\boldsymbol{u}_{t-1}\end{aligned} \tag{6.4}$$

Therefore, instead of the iteration with both $\boldsymbol{q}_t$ and $\boldsymbol{v}_t$ being the state variables, it suffices to study the dynamics in $\hat{\boldsymbol{v}}_t$.

6.3.2 Structure property of a stabilizing controller

The structure property of a stabilizing controller is obtained from Theorem 10. It shows that as long as each controller u_i satisfies the Lipschitz constraints, the system is guaranteed to be locally exponentially stable.

Theorem 10. *Suppose a vector $\boldsymbol{k} = (k_1, \cdots, k_N)$ satisfies $0 \prec diag(\boldsymbol{k}) \prec 2\mathbf{X}^{-1}$. Then if the derivative of controller satisfies $u_i(0) = 0$ and $0 < \frac{\mathrm{d}u_i(\hat{v}_i)}{\mathrm{d}\hat{v}_i} < k_i$ for all $i = 1, \cdots, N$, the equilibrium point $\boldsymbol{v} = 0$ of the dynamic system in (6.4) is locally exponentially stable.*

Proof. The Jacobian of the state transition dynamics in (6.4) is

$$\mathbf{J}(\hat{\boldsymbol{v}}) = \mathbf{I} - \mathbf{X}\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u} \tag{6.5}$$

where $\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u}$ is the gradient of control action $\boldsymbol{u}$ with respect to $\hat{\boldsymbol{v}}$ defined as

$$\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u} = \begin{bmatrix} \frac{\mathrm{d}u_1(\hat{v}_1)}{\mathrm{d}\hat{v}_1} & & \\ & \ddots & \\ & & \frac{\mathrm{d}u_N(\hat{v}_N)}{\mathrm{d}\hat{v}_N} \end{bmatrix}. \tag{6.6}$$

To guarantee an exponentially stable system around the equilibrium, the goal is to show that all the eigenvalues of $\mathbf{J}(\hat{\boldsymbol{v}})$ have magnitude less than 1. To this end, we first show that the eigenvalues of $\mathbf{J}(\hat{\boldsymbol{v}})$ are the same as that of $\mathbf{I} - (\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u})^{\frac{1}{2}}\mathbf{X}(\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u})^{\frac{1}{2}}$.

Let (λ, w) be an eigenpair for $\mathbf{I} - \mathbf{X}(\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u})$. That is, $(\mathbf{I} - \mathbf{X}(\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u}))w = \lambda w$. Then, we have

$$\begin{aligned} &(\mathbf{I} - (\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u})^{\frac{1}{2}}\mathbf{X}(\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u})^{\frac{1}{2}})(\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u})^{\frac{1}{2}}w \\ =&(\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u})^{\frac{1}{2}}w - (\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u})^{\frac{1}{2}}\mathbf{X}(\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u})w \\ =&(\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u})^{\frac{1}{2}}(\mathbf{I} - \mathbf{X}(\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u}))w \\ =&\lambda(\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u})^{\frac{1}{2}}w \end{aligned} \tag{6.7}$$

Therefore, $(\lambda, (\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u})^{\frac{1}{2}}w)$ is an eigenpair for $\mathbf{I} - (\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u})^{\frac{1}{2}}\mathbf{X}(\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u})^{\frac{1}{2}}$. To prove that the eigenvalue of $\mathbf{J}(\hat{\boldsymbol{v}})$ to be strictly smaller than 1, it suffices to show that $-\mathbf{I} \prec \mathbf{I} - (\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u})^{\frac{1}{2}}\mathbf{X}(\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u})^{\frac{1}{2}} \prec \mathbf{I}$.

By picking the controller $\boldsymbol{u}$ such that $0 < \frac{\mathrm{d}u_i(\hat{v}_i)}{\mathrm{d}\hat{v}_i} < k_i$ for all $i = 1, \cdots, N$ and $0 \prec \mathrm{diag}(\boldsymbol{k}) \prec 2\mathbf{X}^{-1}$, we have $0 \prec \nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u} \prec 2\mathbf{X}^{-1}$ and thus $(\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u})^{-1} \succ \frac{1}{2}\mathbf{X}$. Since $\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u} \succ 0$ is diagonal, we then have $(\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u})^{\frac{1}{2}}\mathbf{X}(\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u})^{\frac{1}{2}} \prec 2\mathbf{I}$ and thus $-\mathbf{I} \prec \mathbf{I} - (\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u})^{\frac{1}{2}}\mathbf{X}(\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u})^{\frac{1}{2}} \prec \mathbf{I}$. The right side inequality holds because $\mathbf{X} \succ 0$. □

6.3.3 Optimizing search space for neural network controllers

Note that all the feasible stabilizing $\boldsymbol{u}$ are in a convex set described by $\mathcal{S} = \left\{\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u} | 0 \prec \nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u} \prec 2\mathbf{X}^{-1}\right\}$. Since there is no communication between buses during training, each $\frac{\mathrm{d}u_i(\hat{v}_i)}{\mathrm{d}\hat{v}_i}$ needs to be bounded by a separate k_i for bus $i = 1, \cdots, N$. Therefore, the search space for neural network controllers is constrained by the selection of $\boldsymbol{k}$. A uniform bound $k_i = \frac{2}{\lambda_{max}(\mathbf{X})}$ can be found in literatures [96], but it might be too conservative since $\mathcal{S}$ may be much larger than the region described by $\mathcal{D} = \left\{\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u} | 0 \prec \nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u} \prec \frac{2}{\lambda_{max}(\mathbf{X})}1\right\}$.

Here we show an illustration on a three-bus system (with the first bus as the feeder) where $\mathbf{X} = \begin{bmatrix} 0.20 & -0.16 \\ -0.16 & 0.97 \end{bmatrix}$. For different Lipschitz bounds on controllers, feasible regions for $\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u}$ are shown in Fig. 6.2.

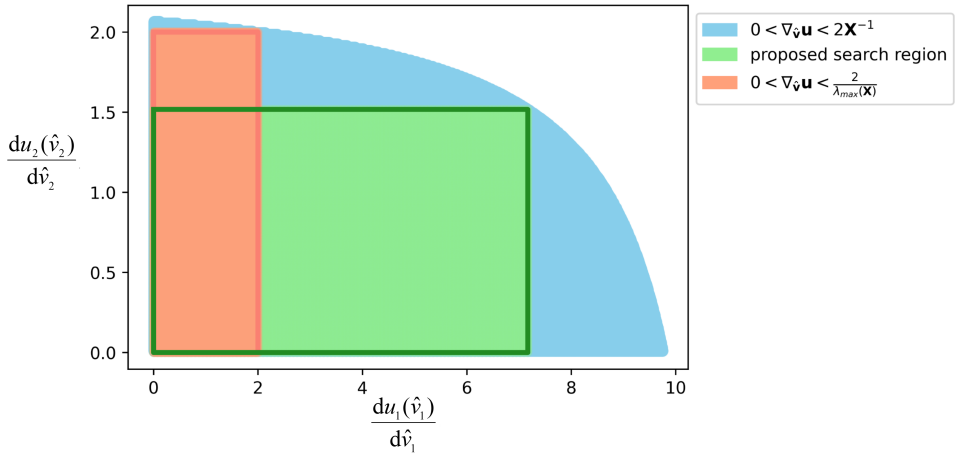


Figure 6.2: Feasible search space comparisons for controllers. The blue area is the set of all feasible $\boldsymbol{u}$ in $\mathcal{S} = \left\{\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u} | 0 \prec \nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u} \prec 2\mathbf{X}^{-1}\right\}$. The orange area is the search space with uniform Lipschitz bounds defined as $\mathcal{D} = \left\{\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u} | 0 \prec \nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u} \prec \frac{2}{\lambda_{max}(\mathbf{X})}\mathbf{1}\right\}$, which is the largest square within blue region but is only a very small subset of $\mathcal{S}$. With each controller being trained independently, it is natural to consider some larger non-uniform search space such as the green area.

The blue area demonstrates the space of controllers constrained by $\mathcal{S} = \left\{\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u} | 0 \prec \nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u} \prec 2\mathbf{X}^{-1}\right\}$. The orange area is the space defined by $\mathcal{D} = \left\{\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u} | 0 \prec \nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u} \prec \frac{2}{\lambda_{max}(\mathbf{X})}1\right\}$, which is the largest square within blue region but is only a very small subset of blue area for $\mathcal{S}$. Note that the axes are scaled so the orange one does not look like a square. With each controller being trained independently, it is natural to consider some larger non-uniform search space such as the green area by choosing different $\boldsymbol{k}$. We may choose a $\boldsymbol{k}^*$ such that the search space $\{\nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u} | 0 \prec \nabla_{\hat{\boldsymbol{v}}}\boldsymbol{u} \prec \boldsymbol{k}^*\}$ is the largest rectangular volume inside blue space, denoted as $\prod_{i=1}^{N} k_i$.

The volume is not a convex function in $\boldsymbol{k}$, but we can apply a simple log trick and

solve the following optimization problem:

$$\max_{\mathbf{k}} \sum_{i=1}^{N} w_i \log(k_i) \tag{6.8a}$$

$$\text{s.t. } 0 \prec \begin{bmatrix} k_1 & & \\ & \ddots & \\ & & k_N \end{bmatrix} \prec 2\mathbf{X}^{-1} \tag{6.8b}$$

where $w_1, \cdots, w_N$ are the coefficients to represent the relative importance of buses. For example, if bus j has none or very limited capacity for voltage regulation, w_j is set to be small. If bus j is the source node of a branch, w_j can be set to be larger to speed up the convergence of voltage at the source node and thus help the convergence of following branches.

In practice, this set of coefficients can be adjusted according to the solutions of the optimization problem and the training of controllers. For the controller u_i whose derivative $\frac{\mathrm{d}u_i(\hat{v}_i)}{\mathrm{d}\hat{v}_i}$ is far from being bounded by k_i, its coefficient w_i can be adjusted to be smaller to encourage larger control action at the other buses. We envision that the system operator has the capability to communicate with each bus at a slower timescale (e.g., once a day) and collect the above information. Accordingly, the operator adjusts coefficients w_i, solves (6.8) and issues the bounds k_i to each bus at this the slower timescale.

6.3.4 Design of stabilizing neural network controllers

From Theorem 10, the structural property of locally exponentially stabilizing controllers is derived in Corollary 10.1. We aim to engineer the neural networks to satisfy these structural property in Corollary 10.1 by design.

Corollary 10.1. *The condition for a locally exponentially stabilizing controller in Theorems 10 is equivalent to:*

1. $u_{\theta_i}(\hat{v}_i)$ *has the same sign as* $\hat{v}_i$
2. $u_{\theta_i}(\hat{v}_i)$ *is monotonically increasing*
3. $\frac{\mathrm{d}u_{\theta_i}(\hat{v}_i)}{\mathrm{d}\hat{v}_i} < k_i$.

The first two requirements are equivalent to designing a monotonically increasing function through the origin. This is constructed by decomposing the function into a positive and a negative part as $f_i(\hat{v}_i) = f_i^{+(\hat{v}_i)+f_i^{-(\hat{v}_i)}}$, where $f_i^{+(\hat{v}_i)}$ is monotonically increasing for $\hat{v}_i > 0$ and zero when $\hat{v}_i \leq 0$; $f_i^{-(\hat{v}_i)}$ is monotonically increasing for $\hat{v}_i < 0$

and zero when $\hat{v}_i \geq 0$. To this end, we formulate the controller with a stacked-ReLU structure shown in Fig. 6.3, which is developed in [116]. This design is a piecewise linear function where the slope of each piece is equal to the summation of weights in activated neurons. Then the requirement 3) can be satisfied by directly thresholding the slope. The neural network controller is constructed as (6.9)

$$u_i(\hat{v}_i) = \boldsymbol{s}_i \text{ReLU}(1\hat{v}_i + \boldsymbol{b}_i) + \boldsymbol{z}_i \text{ReLU}(-1\hat{v}_i + \boldsymbol{d}_i) \tag{6.9a}$$

$$\text{where } 0 < \sum_{j=1}^{l} s_i^j < k_i, \quad \forall l = 1, 2, \cdots, m \tag{6.9b}$$

$$-k_i < \sum_{j=1}^{l} z_i^j < 0, \quad \forall l = 1, 2, \cdots, m \tag{6.9c}$$

$$b_i^1 = 0, b_i^l \leq b_i^{(l-1)}, \quad \forall l = 2, 3, \cdots, m \tag{6.9d}$$

$$d_i^1 = 0, d_i^l \leq d_i^{(l-1)}, \quad \forall l = 2, 3, \cdots, m \tag{6.9e}$$

where m is the number of neurons and $1 \in \mathbb{R}^m$ is the all 1's column vector. Variables $\boldsymbol{s}_i = [s_i^1 \quad s_i^2 \quad \cdots \quad s_i^m]$ and $\boldsymbol{z}_i = [z_i^1 \quad z_i^2 \quad \cdots \quad z_i^m]$ are the weight vector of bus i; $\boldsymbol{b}_i = [b_i^1 \quad b_i^2 \quad \cdots \quad b_i^m]^\intercal$ and $\boldsymbol{d}_i = [d_i^1 \quad d_i^2 \quad \cdots \quad d_i^m]^\intercal$ are the corresponding bias vector. The variables to be trained are weights $\boldsymbol{\theta} = \{\boldsymbol{s}, \boldsymbol{b}, \boldsymbol{z}, \boldsymbol{d}\}$ in (6.9). The saturation limits can be satisfied by hard thresholding the output of the neural network. Note that (6.9) is a single-layer neuron network and m determines the number of pieces for the piece-wise linear function. We tune m according to the testing performances of the controllers and we find that $m = 20$ is generally enough in most settings.

After one iteration, the Lipschitz constraints guarantee that $||\hat{\boldsymbol{v}}_t|| \leq ||\hat{\boldsymbol{v}}_{t-1}||$ and thus the magnitude of voltage deviation will not be worse after a control action. If there is an abrupt change in active power, the voltage will experience an abrupt change and its magnitude may be larger than before. Therefore, the "safety" in our proposed method is in the sense of stability, where the voltage deviations would go to zero if active power changes relatively slowly. In addition, this guarantees the neural network controller does not degrade the the voltage performance compared to an uncontrolled system.

6.4 Decentralized Safe Reinforcement Learning

In this section, we construct a decentralized reinforcement learning framework to optimize neural network controller in each bus locally with observation of trajectories. By having the constraints in (6.9), the neural network controller is guaranteed to stabilize the system.

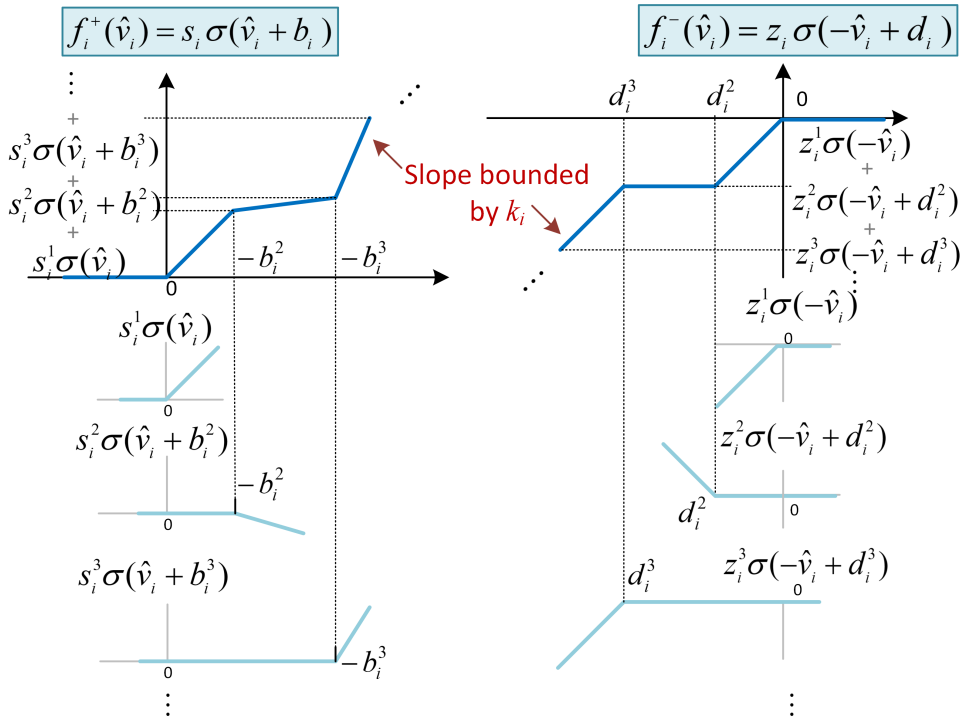


Figure 6.3: Stacked ReLU neural network to formulate a controller satisfying the stabilizing constraint

Most reinforcement learning algorithms, including Q-learning, actor-critic and DDPG, rely on learning a value function (Q-function) satisfying the Bellman equations. Q-function assumes an infinite-horizon formulation where the states follow a stationary probability distribution, which is generally not true for the voltage control problem in this paper. Instead, REINFORCE policy gradient algorithm adopts the log probability trick and avoids learning the value function [101]. Therefore, we use REINFORCE policy gradient algorithm to obtain sampled gradient for updating the weights of neural network controllers.

Notably, there are natural noises in the system coming from the changes in active power $\boldsymbol{p}$, which enable us to implement REINFORCE policy gradient with equivalent stochastic policy. Specifically, we assume that the distribution of noise on the system can be estimated. By incorporating noise term into control action, each action $u_{i,t}$ comes from an equivalent stochastic policy with probability distribution $\pi_\theta(u_{i,t}|\hat{v}_{i,t})$. The gradient for

updating weights of neural network controller at bus i is obtained by [101]

$$\nabla J(\theta) = \mathbb{E}[\sum_{t=1}^{T} \nabla_\theta \log \pi_\theta(u_{i,t}|\hat{v}_{i,t}) \sum_{t=1}^{T} C_i(u_{i,t})] \tag{6.10}$$

The pseudo-code for the decentralized RL framework is given in Algorithm 1. Each bus i has its local RL agent for training in a batch-updating style. Let H be the number of batches. At each episode, each agent collects trajectory $\left\{\hat{v}_{i,1}^h, u_{i,1}^h, \cdots, \hat{v}_{i,T}^h, u_{i,T}^h\right\}$ and the corresponding cost $c_i^h = \sum_{t=1}^{T} C_i(u_{i,t}^h)$ for $h = 1, \cdots, H$. Adam algorithm is adopted to update weights of neural network controllers with gradient computed through batch average of (6.10). We would also like to emphasize that any model-free RL algorithms can be readily utilized in our framework by replacing the REINFORCE algorithm. We use the standard REINFORCE algorithm in this paper to illustrate our contributions: the design of stabilizing neural network controllers and training them in a decentralized manner.

Algorithm 4 Decentralized Reinforcement Learning algorithm with Policy Gradient

Require: Learning rate α, batch size H, trajectory length T, number of episodes E
Input: Initial weights θ for control network

1: **for** *episode* $= 1$ to E **do**
2: **for** agent $i = 1$ to N **do**
3: Collect trajectories $\left\{\hat{v}_{i,1}^h, u_{i,1}^h, \cdots, \hat{v}_{i,T}^h, u_{i,T}^h\right\}$ and the corresponding cost $c_i^h = \sum_{t=1}^{T} C_i(u_{i,t}^h)$ for $h = 1, \cdots, H$
4: Compute the gradient $\nabla J_i(\boldsymbol{\theta}_i) = \frac{1}{H} \sum_{h=1}^{H} \sum_{t=1}^{T} \nabla_\theta \log \pi_\theta(u_{i,t}^h|\hat{v}_{i,t}^h) c_i^h$
5: Update weights in the neural network by passing $J_i(\boldsymbol{\theta}_i)$ to Adam optimizer: $\boldsymbol{\theta}_i \leftarrow \boldsymbol{\theta}_i - \alpha \nabla J_i(\boldsymbol{\theta}_i)$
6: **end for**
7: **end for**=0

6.5 Numerical Results

We verify the performance of the proposed safe RL approach on IEEE 33-bus test feeders [117]. We first show that unconstrained neural network controllers learned by RL might lead to an unstable system, while the controllers trained by safe RL approach are guaranteed to stabilize the system. Then, we show that the proposed decentralized RL framework can learn flexible non-linear controllers for different buses that outperform conventional linear control law.

6.5.1 Simulation setup

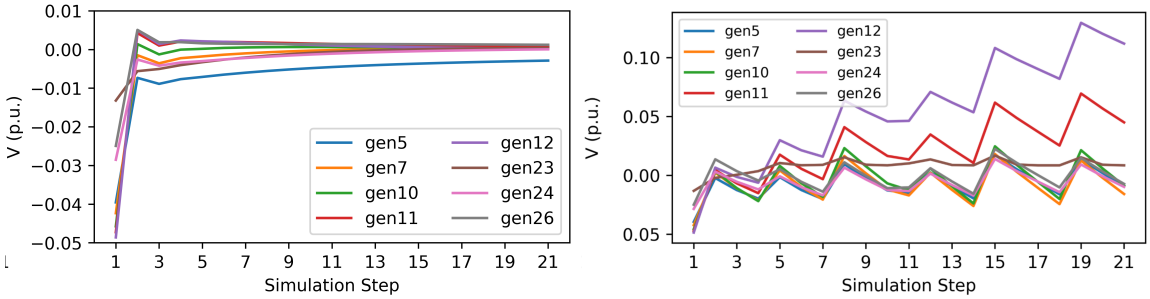
The cost function that each controller collectively optimizes is $C(\boldsymbol{u}) = \sum_{t=1}^{T} (||\boldsymbol{v}_t||_1 + \gamma||\boldsymbol{u}_t||_1)$, where γ acts as a trade-off parameter and is set to be 0.01. The base unit for power and voltage is 100kVA and 12.66kV, respectively. The bound on action $\bar{\boldsymbol{u}}$ is generated to be uniformly distributed in $[0.01, 0.05]$. We assume other voltage regulation equipment, such as tap-changing transformers and discrete switching capacitor backs, operate at much slower timescales than the inverters. Therefore, in the simulations, we only consider the operation of the inverters, as learned by an agent running RL algorithms [118, 93]. We use TensorFlow 2.0 framework to build the reinforcement learning environment. The episode number, batch size and the number of neurons are 500, 500, 20, respectively. Parameters of neural network controllers are updated using Adam with learning rate initialized to be 0.003 and decayed every 100 steps with a base of 0.6. We compare the performance of neural network controller designed with and without the safe RL approach, as well as conventional linear controller. All of them are trained using the decentralized RL framework.

6.5.2 Necessity of the stabilizing requirement

Intuitively speaking, if a controller achieves a low loss function after training converges, one might hope that it naturally leads to a stabilizing controller since the trajectory does not blow up to a high cost. Fig. 6.4 shows the dynamics of voltage deviation under the neural network controllers trained with and without the safe RL approach. The one without safe RL approach is *unstable* and leads to very large state oscillations (Fig. 6.4(b)). In contrast, the controller with safe RL approach shows good performance in Fig. 6.4(a). Therefore, explicitly constraining the controller structure is necessary.

6.5.3 Performance comparison

To investigate the convergence of the safe RL approach, Fig. 6.6(a) shows the normalized cost on the test set along episodes for training of neural network controllers and linear controllers. All the losses converge, with the proposed neural network controllers achieving the lowest cost. Fig. 6.6(b) shows the cost on selected buses along the episodes of training. It is interesting to observe that training the controllers in a decentralized fashion did not impact convergence or performance. Namely, during training, u_i is updated based only on the trajectory of v_i, even though the control action impacts the voltage at all neighboring buses.



(a) safe RL approach **(b)** without safe RL approach

Figure 6.4: Dynamics of voltage deviation for safe RL approach(left) and without safe RL approach(right). The controller designed without the safe RL approach approach leads to unstable trajectories

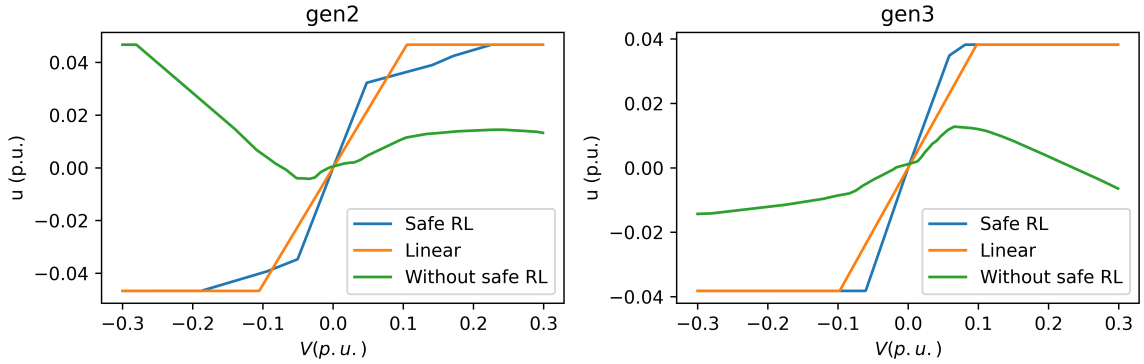
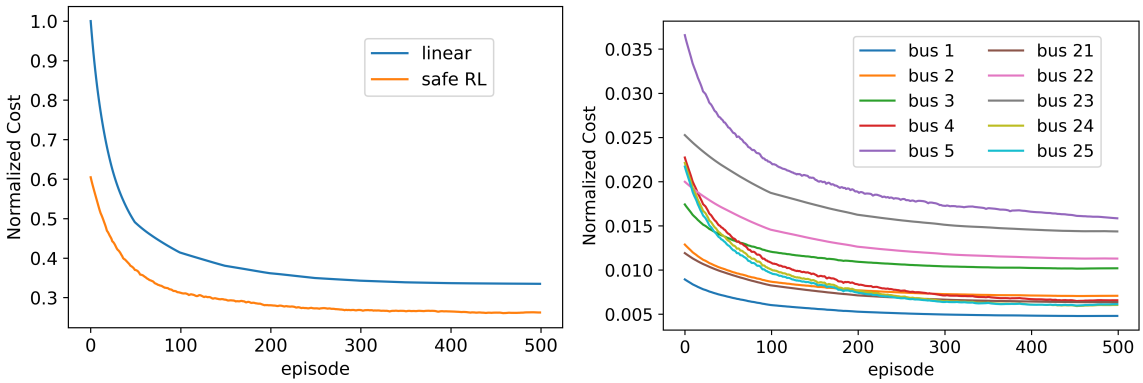


Figure 6.5: Voltage control law obtained by linear controller with optimal linear coefficient, neural network controllers designed with safe RL approach and without safe RL approach. The neural network controllers learn flexible non-linear control laws for different buses, with the slope of controller obtained by safe RL approach bounded by Lipschitz constraints.

The control law for neural network controller learned with safe RL, without safe RL approach and linear controller with optimal linear coefficient are shown in Fig. 6.5. The neural network controllers learn flexible non-linear control law for different generators, with the safe RL approach guaranteeing a stabilizing controller by bounding the slope with Lipschitz constraints. Fig. 6.7 illustrates the dynamics of voltage deviation $\boldsymbol{v}$ and corresponding control action $\boldsymbol{u}$ under optimal linear controller and neural network controller trained by safe RL approach. The neural network controller generally leads to faster decay of voltage deviation.

In the test set with random initial states, the distribution of cost in selected buses is shown in Fig 6.8. The average costs of the linear controller, the neural network controller



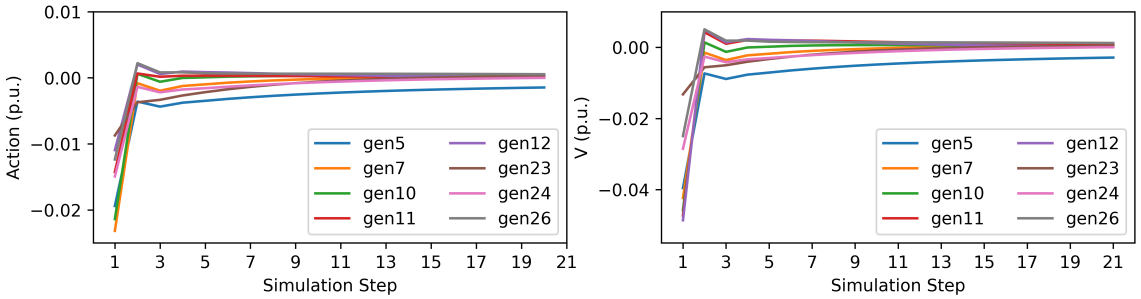
(a) Total cost **(b)** Cost for selected buses

Figure 6.6: Normalized cost on test set along the episode of training. (a) Total cost during training of neural network controller and linear controller. Neural network controller designed with safe RL approach achieves lower cost than conventional linear controller. (b) Cost on selected generator buses during the training of neural network controller. All learning trajectories converge well in the decentralized model-free setting, even though they interact through the underlying distribution network.

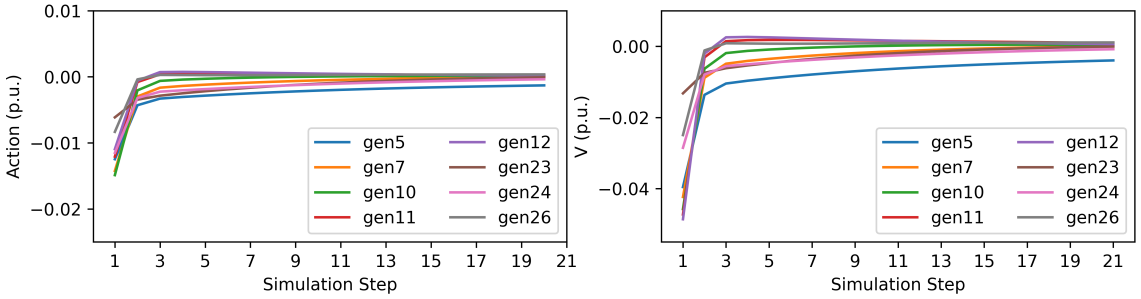
bounded by $\frac{2}{\lambda_{max}(\mathbf{X})}$, and the neural network controller with optimal Lipschiz bound obtained in (6.8) are 0.44, 0.38 and 0.36, respectively. Therefore, the proposed approach can learn a stabilizing controller that reduces the cost by approximately 18.18% compared to conventional linear control law. Moreover, safe RL with the optimal Lipschiz bound also reduces the cost by approximately 5.26% compared to safe RL with the uniform Lipschiz bound $\frac{2}{\lambda_{max}(\mathbf{X})}$.

6.6 Conclusions

This paper proposes a safe RL approach for optimal voltage control. The exponential stability of the system is guaranteed by controllers constrainted by Lipschitz bounds, which are optimized to enlarge the search space. The neural network controllers are parameterized by a staked ReLU neural network to satisfy stabilizing constraints implicitly. Each bus updates weights locally with the decentralized RL framework. Case studies show that RL without stability constraints can lead to unstable controllers, while the proposed safe learning approach will lead to a stabilizing controller. The neural network controllers outperform conventional linear controllers by speeding up the convergence of voltages to reference values with relatively low control effort. Rigorously analyzing the difference between decentralized and centralized training, their convergence behaviors



(a) Dynamics of $\hat{v}$ (left) and u (right) for neural network controller obtained through safe RL approach



(b) Dynamics of $\hat{v}$ (left) and u (right) for linear control

Figure 6.7: Dynamics of the voltage deviation $\hat{\boldsymbol{v}}$ and the control action $\boldsymbol{u}$ in selected generator buses corresponding to (a) neural network controller trained with safe RL approach (b) Linear control obtained by the same decentralized RL algorithm. The neural network controller generally leads to faster decay of voltage deviation.

and generalizing the proposed methods to nonlinear power flow models are important future directions for us.

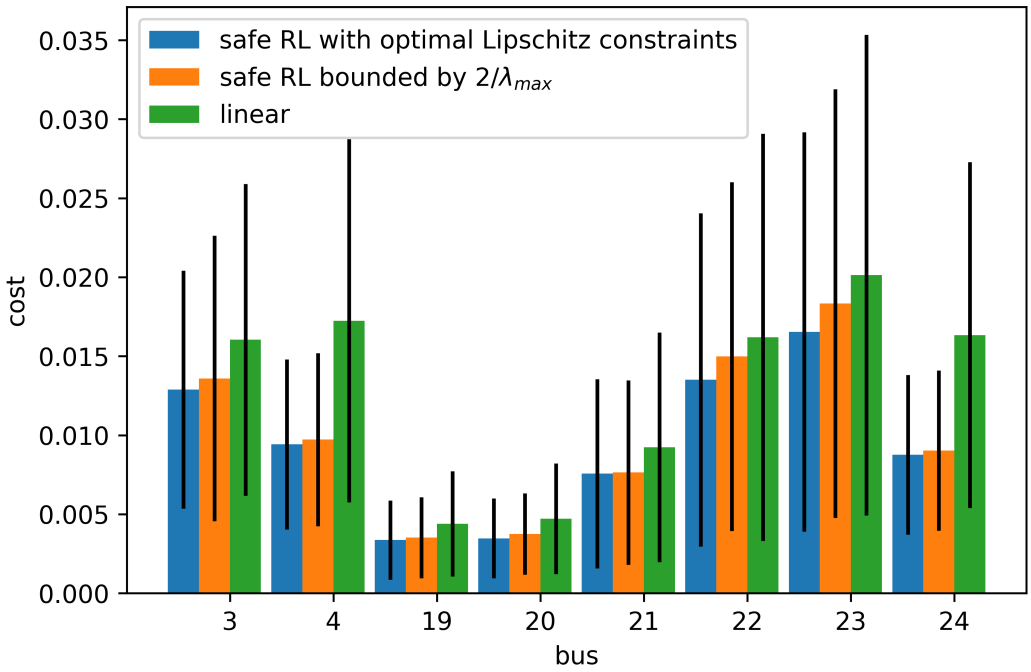


Figure 6.8: Distribution of cost in selected generator buses with random initial states corresponding to safe RL with proposed optimal Lipschiz constraints, safe RL bounded by $\frac{2}{\lambda_{max}(\mathbf{X})}$ and optimal linear control. Compared to uniform bound $\frac{2}{\lambda_{max}(\mathbf{X})}$ and linear controller, the proposed approach reduces the average cost by approximately 5.26%, 18.18%, respectively.

Chapter 7

Conclusions

The increasing integration of distributed energy resources (DERs) into modern energy systems has brought about transformative changes, offering opportunities for decentralized energy management and user participation. However, these changes also introduce critical challenges in addressing fairness, efficiency, privacy, and stability. This thesis tackles these challenges by developing principled frameworks that integrate game theory, machine learning and control theory to achieve fair, efficient and resilient energy resource allocation.

7.1 Summary of Contributions

This thesis has made the following key contributions:

- **Socially Optimal Energy Usage via Adaptive Pricing**: A privacy-preserving two-time-scale incentive mechanism was developed to coordinate distributed users' energy consumption without requiring access to private user data. The framework achieves socially optimal solutions that align individual preferences with global efficiency goals, demonstrating flexibility through its ability to handle non-convex cost functions and incorporate machine learning-based load control algorithms.
- **Fairness in Energy Resource Allocations**: A fairness framework using α-fairness metrics was introduced to address disparities in energy access and cost distribution. By jointly optimizing total resource purchased and individual allocations, the framework constructs allocation strategies that lie on the Pareto front, achieving a principled trade-off between fairness and efficiency. This achieves fair resource allocation with quantifiable metrics to track system performance.

- **Strategic Aggregator Interactions in Energy Markets**: The fairness framework was extended to a multi-aggregator setting, modeling their interactions as a quasi-concave game. The analysis provided theoretical guarantees for the existence of Nash equilibria, offering insights into how aggregators stabilize markets, ensure fair resource distribution, and optimize user surplus.
- **Decentralized Safe Reinforcement Learning for Voltage Control**: A decentralized reinforcement learning framework was developed for voltage control in inverter-based systems. By deriving and imposing Lipschitz constraints on neural network-based controllers, the framework guarantees system stability while enabling scalable, model-free training for distributed networks.

7.2 Future Research Directions

While this thesis addresses several critical challenges in energy systems, there remain open questions and opportunities for future work.

7.2.1 Extending Incentive Mechanisms to Strategic and Networked Energy Systems

Our preliminary work [119] presents a privacy-preserving, two-time-scale incentive mechanism that achieves socially optimal energy consumption coordination. While the current approach assumes that all users are price takers, an important direction for future research is to account for price-anticipatory behaviors. Extending the framework to incorporate strategic user behavior will involve leveraging game-theoretic tools to design mechanisms that remain efficient and equitable even when users influence prices.

Additionally, applying the incentive mechanism to networked systems, such as the Direct Current Optimal Power Flow (DCOPF) problem, offers a promising avenue for exploration. DCOPF is a critical tool in power system operation, particularly under the growing integration of renewable resources and uncertain load-generation scenarios. Investigating the convergence behavior of the proposed mechanism in these networked environments will provide valuable insights into its scalability and adaptability. These extensions will enhance the applicability of the mechanism to more complex, real-world energy systems, further supporting decentralized decision-making and system-wide efficiency.

7.2.2 Advancing Fairness and Efficiency in Energy Systems: Future Directions

This thesis addresses the challenge of balancing fairness and efficiency in energy resource allocation through principled frameworks and aggregator-based models. While significant progress has been made, several important directions remain for future exploration.

A critical assumption in the current work is that aggregators have full knowledge of users' utility functions. However, this may not hold in real-world scenarios, where such information might be incomplete or unavailable. Future research could focus on developing fair resource allocation schemes that operate under partial or no knowledge of users' utility functions. This could involve integrating machine learning techniques to infer utility functions from historical data, enabling aggregators to make informed decisions based on limited information.

Another promising direction is to explore decentralized algorithms for fair resource allocation. Decentralized approaches would enhance scalability and provide privacy-preserving solutions by reducing the need for centralized data collection. These algorithms could allow users to interact directly with aggregators or other users, aligning local decisions with global fairness-efficiency objectives in a distributed setting.

Extending the framework to incorporate individual users' budget constraints offers another important avenue for future research. Accounting for budget limitations could better address real-world disparities in energy access and affordability, particularly for small-scale or economically disadvantaged users. This addition would further enhance the practical applicability of fairness frameworks by ensuring that resource allocation strategies remain equitable across diverse user groups.

In the multi-aggregator setting, future work could focus on refining the game-theoretic framework to address unresolved questions, such as the uniqueness of Nash equilibria. Investigating the influence of fairness considerations on the strategic behavior of aggregators, especially in competitive markets, could provide deeper insights into how aggregators stabilize market outcomes while balancing fairness and efficiency. Moreover, applying these models to real-world datasets would validate their effectiveness and offer actionable recommendations for policy and market design.

By addressing these directions, future work can expand the scope and impact of fairness frameworks, enabling energy systems that are not only efficient and equitable but also scalable, privacy-preserving, and robust to real-world constraints.

7.2.3 Advancing Multi-agent Decentralized Learning and Stability in Voltage Control

Our preliminary work [120] presents a decentralized reinforcement learning (RL) framework for safe and efficient voltage control, but several avenues for future exploration remain. A critical direction is to rigorously analyze the convergence of decentralized algorithms to a Nash equilibrium. The framework trains local neural network controllers at each bus in a model-free setting without requiring real-time communication. Preliminary findings indicate that structural properties of stabilizing controllers, such as convexity, are pivotal for achieving convergence. Further research is needed to compare decentralized and centralized training approaches to uncover insights into their relative efficiency, scalability, and robustness.

Additionally, the framework currently assumes linearized power flow models, which simplify system dynamics. Extending it to nonlinear power flow models would significantly broaden its applicability, as real-world systems often exhibit nonlinear behaviors. Finally, reformulating the decentralized voltage control problem as a non-cooperative game could provide deeper insights into the existence and uniqueness of Nash equilibria, offering a robust theoretical foundation for dynamic, distributed energy networks. Addressing these directions will enhance the applicability and scalability of decentralized RL frameworks, supporting the integration of distributed energy resources in modern power systems.

Appendix A

Appendix to Chapter 2

Consider the first condition in Theorem 3. Again we let $\mathbf{x}_s$ denote the sum $\sum_i \mathbf{x}_i$. The time derivative of the Lyapunov function in (3.9) is

$$\begin{aligned}
\dot{V}(\mathbf{p}) &= \nabla V(\mathbf{p})^\intercal \dot{\mathbf{p}} \\
&= (\mathbf{p} - \mathbf{p}^*)^\intercal \mathbf{B}^{-1}(\mathbf{e}(\mathbf{x}_s^*(\mathbf{p})) - \mathbf{p}) \\
&\overset{(a)}{=} (\mathbf{p} - \mathbf{p}^*)^\intercal \mathbf{B}^{-1}(\mathbf{e}(\mathbf{x}_x^*(\mathbf{p})) - \mathbf{p} - \mathbf{e}(\mathbf{x}_s^*(\mathbf{p}^*)) + \mathbf{p}^*) \\
&\overset{(b)}{=} (\mathbf{p} - \mathbf{p}^*)^\intercal \mathbf{B}^{-1}(\mathbf{B}(\mathbf{x}_s^*(\mathbf{p}) - \mathbf{x}_s^*(\mathbf{p}^*)) - \mathbf{p} + \mathbf{p}^*) \\
&= (\mathbf{p} - \mathbf{p}^*)^\intercal (\mathbf{x}_s^*(\mathbf{p}) - \mathbf{x}_s^*(\mathbf{p}^*)) \\
&\quad - (\mathbf{p} - \mathbf{p}^*)^\intercal \mathbf{B}^{-1}(\mathbf{p} - \mathbf{p}^*) \\
&\overset{(c)}{\leq} 0,
\end{aligned}$$

where (a) follows from the fact that $\mathbf{p}^*$ is the equilibrium; (b) follows from the definition $\mathbf{e} = \nabla g$; (c) follows from the assumption that each $\mathbf{x}_i^*(\mathbf{p}^*)$ is decreasing (and hence their sum is) and $\mathbf{B}$ is positive definite. Furthermore, $V(\mathbf{p}) = 0$ only if $\mathbf{p} = \mathbf{p}^*$.

Next, we consider the second condition in Theorem 3. The time derivative of the

Lyapunov function in (3.10) is

$$\begin{aligned}\dot{V}(\mathbf{p}) &= \nabla V(\mathbf{p})^\intercal \dot{\mathbf{p}} \\ &= \nabla V(\mathbf{p})^\intercal (\mathbf{e}(\mathbf{x}_s^*(\mathbf{p})) - \mathbf{e}(\mathbf{x}_s^*(\mathbf{p}^*)) - \mathbf{p} + \mathbf{p}^*) \\ &= (\mathbf{e}(\mathbf{x}_s^*(\mathbf{p})) - \mathbf{e}(\mathbf{x}_s^*(\mathbf{p}^*)) - \mathbf{p} + \mathbf{p}^*)^\intercal \left[\nabla_{\mathbf{p}} \mathbf{x}_s^*(\mathbf{p})\right] \\ &\quad (\mathbf{e}(\mathbf{x}_s^*(\mathbf{p})) - \mathbf{e}(\mathbf{x}_s^*(\mathbf{p}^*)) - \mathbf{p} + \mathbf{p}^*) \\ &\quad - (\mathbf{x}_s^*(\mathbf{p}) - \mathbf{x}_s^*(\mathbf{p}^*))^\intercal (\mathbf{e}(\mathbf{x}_s^*(\mathbf{p})) - \mathbf{e}(\mathbf{x}_s^*(\mathbf{p}^*))) \\ &\quad + (\mathbf{x}_s^*(\mathbf{p}) - \mathbf{x}_s^*(\mathbf{p}^*))^\intercal (\mathbf{p} - \mathbf{p}^*).\end{aligned} \tag{A.1}$$

It's easy to see that $\dot{V}(\mathbf{p}) = 0$ when $\mathbf{p} = \mathbf{p}^*$. We first look at the last two terms. Since $\mathbf{e} = \nabla g$ and g is strictly convex, $\mathbf{e}$ is strictly increasing, and $(\mathbf{x}_s^*(\mathbf{p}) - \mathbf{x}_s^*(\mathbf{p}^*))^\intercal (\mathbf{e}(\mathbf{x}_s^*(\mathbf{p})) - \mathbf{e}(\mathbf{x}_s^*(\mathbf{p}^*))) > 0$ if $\mathbf{p} \neq \mathbf{p}^*$. By assumption, $\mathbf{x}_s^*(\mathbf{p})$ is decreasing and $(\mathbf{x}_s^*(\mathbf{p}) - \mathbf{x}_s^*(\mathbf{p}^*))^\intercal (\mathbf{p} - \mathbf{p}^*) \leq 0$.

Next, consider the term in (A.1). The object $\nabla_{\mathbf{p}} \mathbf{x}_s^*(\mathbf{p})$ is the gradient of the vector $\mathbf{x}_s^*(\mathbf{p})$, hence it is a $T \times T$ matrix and it suffices to show $\nabla_{\mathbf{p}} \mathbf{x}_s^*(\mathbf{p})$ is negative definite. Using the fact that summations of negative definite matrices are negative definite, it is enough to show that $\nabla_{\mathbf{p}} \mathbf{x}_i^*(\mathbf{p})$ is negative definite for all i. Recall $\mathbf{x}_i^*(\mathbf{p})$ is the optimal solution to the local optimization problem, and it satisfies the first-order condition

$$\nabla_{\mathbf{x}_i} f_i(\mathbf{x}_i) + \mathbf{p} = 0.$$

Taking the derivative of $\mathbf{p}$ on both sides, applying the chain rule, and rearranging leads to:

$$\mathbf{I} = -\left[\nabla_{\mathbf{p}} \mathbf{x}_i^*(\mathbf{p})\right] \left[\mathbf{H}_{\mathbf{x}_i}(\mathbf{x}_i^*(\mathbf{p}))\right]$$

where $[\mathbf{H}_{\mathbf{x}_i}(\mathbf{x}_i^*(\mathbf{p}))]$ is the Hessian of f_i evaluated at $\mathbf{x}_i^*(\mathbf{p})$. Since f_i are assumed to be strictly convex and twice differentiable, $[\mathbf{H}_{\mathbf{x}_i}(\mathbf{x}_i^*(\mathbf{p}))]$ is positive definite. Rearranging the above equation gives $\nabla_{\mathbf{p}} \mathbf{x}_i^*(\mathbf{p}) = -\left[\mathbf{H}_{\mathbf{x}_i}(\mathbf{x}_i^*(\mathbf{p}))\right]^{-1}$. Therefore $\nabla_{\mathbf{p}} \mathbf{x}_i^*(\mathbf{p})$ is negative definite and $\dot{V}(\mathbf{p}) < 0$ when $\mathbf{p} \neq \mathbf{p}^*$.

Appendix B

Bibliography

[1] M. Motoki, M. Umeda, M. Fripp, and A. Kuh, "Approximate dynamic programming for control of a residential water heater," in *2015 International Joint Conference on Neural Networks (IJCNN)*, 2015, pp. 1–8.

[2] U. D. of Energy, "Clean energy resources meet data center electricity demand," 2023. [Online]. Available: https://www.energy.gov/policy/articles/clean-energy-resources-meet-data-center-electricity-demand

[3] B. . Company, "Utilities must reinvent themselves to harness the ai-driven data center boom," 2023. [Online]. Available: https://www.bain.com/insights/utilities-must-reinvent-themselves-to-harness-the-ai-driven-data-center-boom

[4] S. Times, "Power hungry: How the data center boom drained wa of hydropower," 2023. [Online]. Available: https://www.seattletimes.com/seattle-news/times-watchdog/power-hungry-how-the-data-center-boom-drained-wa-of-hydropower/

[5] DeepMind, "Deepmind ai reduces data center cooling costs by 40%," 2017. [Online]. Available: https://deepmind.google/discover/blog/safety-first-ai-for-autonomous-data-centre-cooling-and-industrial-control/

[6] McKinsey and Company, "Demand-based pricing stabilizes the electricity market of the future," 2024. [Online]. Available: https://www.mckinsey.com/industries/electric-power-and-natural-gas/our-insights/demand-based-pricing-stabilizes-the-electricity-market-of-the-future

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

[8] C. Zhang, Y. Xu, Z. Y. Dong, and K. P. Wong, "Robust coordination of distributed generation and price-based demand response in microgrids," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 4236–4247, 2017.

[9] S. H. Tindemans, V. Trovato, and G. Strbac, "Decentralized control of thermostatic loads for flexible demand response," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 5, pp. 1685–1700, 2015.

[10] M. R. Sarker, Y. Dvorkin, and M. A. Ortega-Vazquez, "Optimal participation of an electric vehicle aggregator in day-ahead energy and reserve markets," *IEEE transactions on power systems*, vol. 31, no. 5, pp. 3506–3515, 2015.

[11] D. O'Neill, M. Levorato, A. Goldsmith, and U. Mitra, "Residential demand response using reinforcement learning," in *2010 First IEEE international conference on smart grid communications*. IEEE, 2010, pp. 409–414.

[12] H. Li, Z. Wan, and H. He, "Constrained ev charging scheduling based on safe deep reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2427–2439, 2019.

[13] J. R. Vázquez-Canteli and Z. Nagy, "Reinforcement learning for demand response: A review of algorithms and modeling techniques," *Applied energy*, vol. 235, pp. 1072–1089, 2019.

[14] B. Wang, Y. Li, W. Ming, and S. Wang, "Deep reinforcement learning method for demand response management of interruptible load," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3146–3155, 2020.

[15] Y. Shi and B. Zhang, "Multi-agent reinforcement learning in cournot games," in *2020 59th IEEE Conference on Decision and Control (CDC)*. IEEE, 2020, pp. 3561–3566.

[16] P. Li, H. Wang, and B. Zhang, "A distributed online pricing strategy for demand response programs," *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 350–360, 2019.

[17] K. Khezeli and E. Bitar, "Risk-sensitive learning and pricing for demand response," *IEEE Transactions on Smart Grid*, vol. 9, no. 6, pp. 6000–6007, 2017.

[18] S. Zheng, Y. Sun, B. Li, B. Qi, K. Shi, Y. Li, and X. Tu, "Incentive-based integrated demand response for multiple energy carriers considering behavioral coupling effect of consumers," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3231–3245, 2020.

[19] T. Başar, "Affine incentive schemes for stochastic systems with dynamic information," *SIAM Journal on Control and Optimization*, vol. 22, no. 2, pp. 199–210, 1984.

[20] Y.-C. Ho, P. B. Luh, and G. J. Olsder, "A control-theoretic view on incentives," *Automatica*, vol. 18, no. 2, pp. 167–179, 1982.

[21] D. Paccagnan, R. Chandan, B. L. Ferguson, and J. R. Marden, "Incentivizing efficient use of shared infrastructure: Optimal tolls in congestion games," *arXiv preprint arXiv:1911.09806*, 2019.

[22] M. Peplinski and K. T. Sanders, "Residential electricity demand on caiso flex alert days: a case study of voluntary emergency demand response programs," *Environmental Research: Energy*, vol. 1, no. 1, 2023.

[23] J. S. Vardakas, N. Zorba, and C. V. Verikoukis, "A survey on demand response programs in smart grids: Pricing methods and optimization algorithms," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 152–178, 2014.

[24] C. Chen, J. Wang, and S. Kishore, "A distributed direct load control approach for large-scale residential demand response," *IEEE Transactions on Power Systems*, vol. 29, no. 5, pp. 2219–2228, 2014.

[25] C. Maheshwari, K. Kulkarni, M. Wu, and S. S. Sastry, "Inducing social optimality in games via adaptive incentive design," in *Conference on Decision and Control (CDC)*. IEEE, 2022, pp. 2864–2869.

[26] L. J. Ratliff and T. Fiez, "Adaptive incentive design," *IEEE Transactions on Automatic Control*, vol. 66, no. 8, pp. 3871–3878, 2020.

[27] B. Liu, J. Li, Z. Yang, H.-T. Wai, M. Hong, Y. M. Nie, and Z. Wang, "Inducing equilibria via incentives: Simultaneous design-and-play finds global optima," *arXiv:2110.01212*, 2021.

[28] N. Li, L. Chen, and S. H. Low, "Optimal demand response based on utility maximization in power networks," in *2011 IEEE power and energy society general meeting*. IEEE, 2011, pp. 1–8.

[29] X. Yan, D. Wright, S. Kumar, G. Lee, and Y. Ozturk, "Enabling consumer behavior modification through real time energy pricing," in *2015 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*. IEEE, 2015, pp. 311–316.

[30] W. Huang, N. Zhang, C. Kang, M. Li, and M. Huo, "From demand response to integrated demand response: Review and prospect of research and application," *Protection and Control of Modern Power Systems*, vol. 4, pp. 1–13, 2019.

[31] Z. Liu, I. Liu, S. Low, and A. Wierman, "Pricing data center demand response," *ACM SIGMETRICS Performance Evaluation Review*, vol. 42, no. 1, pp. 111–123, 2014.

[32] V. S. Borkar, "Stochastic approximation with two time scales," *Systems & Control Letters*, vol. 29, no. 5, pp. 291–294, 1997. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167691197900153

[33] Y. Wan, T. Kober, and M. Densing, "Nonlinear inverse demand curves in electricity market modeling," *Energy Economics*, vol. 107, 2022.

[34] D. S. Kirschen and G. Strbac, *Fundamentals of power system economics*. John Wiley & Sons, 2018.

[35] Federal Energy Management Program, "Demand response and time-variable pricing programs: Western states," U.S. Department of Energy, Tech. Rep., 2024.

[36] SmartGrid.gov, "Recovery act: Time based rate programs," U.S. Department of Energy, Tech. Rep., 2024.

[37] G. R. Newsham and B. G. Bowker, "The effect of utility time-varying pricing and load control strategies on residential summer peak electricity use: A review," *Energy policy*, vol. 38, no. 7, pp. 3289–3296, 2010.

[38] Y. Wang and L. Li, "Time-of-use electricity pricing for industrial customers: A survey of us utilities," *Applied Energy*, vol. 149, pp. 89–103, 2015.

[39] F. Nielsen and G. Hadjeres, "Monte carlo information-geometric structures," *Geometric Structures of Information*, pp. 69–103, 2019.

[40] M. F. Akorede, H. Hizam, and E. Pouresmaeil, "Distributed energy resources and benefits to the environment," *Renewable and Sustainable Energy Reviews*, vol. 14, no. 2, pp. 724–734, 2010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032109002561

[41] S. Burger, J. P. Chaves-Ávila, C. Batlle, and I. J. Pérez-Arriaga, "A review of the value of aggregators in electricity systems," *Renewable and Sustainable Energy Reviews*, vol. 77, pp. 395–405, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032117305191

[42] M. R. Sarker, Y. Dvorkin, and M. A. Ortega-Vazquez, "Optimal participation of an electric vehicle aggregator in day-ahead energy and reserve markets," *IEEE Transactions on Power Systems*, vol. 31, no. 5, pp. 3506–3515, 2016.

[43] J. E. Contreras-Ocana, M. A. Ortega-Vazquez, and B. Zhang, "Participation of an energy storage aggregator in electricity markets," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 1171–1183, 2017.

[44] L. Xie, T. Huang, P. Kumar, A. A. Thatte, and S. K. Mitter, "On an information and control architecture for future electric energy systems," *Proceedings of the IEEE*, vol. 110, no. 12, pp. 1940–1962, 2022.

[45] C. Chen, A. S. Alahmed, T. D. Mount, and L. Tong, "Competitive der aggregation for participation in wholesale markets," in *Hawaii International Conference on System Sciences*, 2023. [Online]. Available: https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/fb77470a-ef2a-43a3-9ffc-61906dd13000/content

[46] A. M. Carreiro, H. M. Jorge, and C. H. Antunes, "Energy management systems aggregators: A literature survey," *Renewable and sustainable energy reviews*, vol. 73, pp. 1160–1172, 2017.

[47] J. Li, M. Motoki, and B. Zhang, "Socially optimal energy usage via adaptive pricing," *arXiv preprint arXiv:2310.13254*, 2023.

[48] Y. Yang, G. Hu, and C. J. Spanos, "Optimal sharing and fair cost allocation of community energy storage," *IEEE Transactions on Smart Grid*, vol. 12, no. 5, pp. 4185–4194, 2021.

[49] Z. Fornier, V. Leclère, and P. Pinson, "Fairness by design in shared-energy allocation problems," *arXiv preprint arXiv:2402.00471*, 2024.

[50] Office of Energy Efficiency and Renewable Energy, "EERE impact evaluation method guide for justice 40, equity, and workforce diversity goals," *US Department of Energy*, 2023.

[51] K. Jenkins, D. McCauley, R. Heffron, H. Stephan, and R. Rehner, "Energy justice: A conceptual review," *Energy Research & Social Science*, vol. 11, pp. 174–182, 2016.

[52] W. Ren, Y. Guan, F. Qiu, T. Levin, and M. Heleno, "A literature review of energy justice," *arXiv: 2312.14983*, 2023.

[53] F. Moret and P. Pinson, "Energy collectives: A community and fairness based approach to future electricity markets," *IEEE Transactions on Power Systems*, vol. 34, no. 5, pp. 3994–4004, 2019.

[54] J. Brehmer and W. Utschick, "On proportional fairness in nonconvex wireless systems," in *Proceedings of the International ITG Workshop on Smart Antennas, Berlin, Germany*, 2009.

[55] A. Sinha and A. Anastasopoulos, "Incentive mechanisms for fairness among strategic agents," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 2, pp. 288–301, 2017.

[56] J. Chen, Y. Wang, and T. Lan, "Bringing fairness to actor-critic reinforcement learning for network utility optimization," in *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications*, 2021, pp. 1–10.

[57] T. Li, S. Hu, A. Beirami, and V. Smith, "Ditto: Fair and robust federated learning through personalization," in *International conference on machine learning*. PMLR, 2021, pp. 6357–6368.

[58] D. Kirschen and G. Strbac, *Fundamentals of Power System Economics*. United Kingdom: John Wiley & Sons Ltd, 2004.

[59] R. A. Berry and R. Johari, "Economic modeling in networking: A primer," *Foundations and Trends® in Networking*, vol. 6, no. 3, pp. 165–286, 2013. [Online]. Available: http://dx.doi.org/10.1561/1300000011

[60] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Transactions on networking*, vol. 8, no. 5, pp. 556–567, 2000.

[61] H. Boche and M. Schubert, "A generalization of nash bargaining and proportional fairness to log-convex utility sets with power constraints," *IEEE Transactions on Information Theory*, vol. 57, no. 6, pp. 3390–3404, 2011.

[62] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan, "Rate control for communication networks: shadow prices, proportional fairness and stability," *Journal of the Operational Research society*, vol. 49, pp. 237–252, 1998.

[63] T. Lan, D. Kao, M. Chiang, and A. Sabharwal, *An axiomatic theory of fairness in network resource allocation.* IEEE, 2010.

[64] C. Joe-Wong, S. Sen, T. Lan, and M. Chiang, "Multiresource allocation: Fairness-efficiency tradeoffs in a unifying framework," *IEEE/ACM Transactions on Networking*, vol. 21, no. 6, pp. 1785–1798, 2013.

[65] L. Jia and L. Tong, "Dynamic pricing and distributed energy management for demand response," *IEEE Transactions on Smart Grid*, vol. 7, no. 2, pp. 1128–1136, 2016.

[66] D. Bertsimas, V. F. Farias, and N. Trichakis, "The price of fairness," *Oper. Res.*, vol. 59, no. 1, p. 17–31, jan 2011. [Online]. Available: https://doi.org/10.1287/opre.1100.0865

[67] ——, "On the efficiency-fairness trade-off," *Manage. Sci.*, vol. 58, no. 12, p. 2234–2250, dec 2012. [Online]. Available: https://doi.org/10.1287/mnsc.1120.1549

[68] E. Altman, K. Avrachenkov, and A. Garnaev, "Generalized α-fair resource allocation in wireless networks," in *2008 47th IEEE Conference on Decision and Control.* IEEE, 2008, pp. 2414–2419.

[69] FERC, "Participation of distributed energy resource aggregations in markets operated by regional transmission organizations and independent system operators," https://www.govinfo.gov/content/pkg/FR-2021-03-30/pdf/2021-06089.pdf, 20201.

[70] B. Zhang, R. Johari, and R. Rajagopal, "Competition and coalition formation of renewable power producers," *IEEE Transactions on Power Systems*, vol. 30, no. 3, pp. 1624–1632, 2015.

[71] J. Li, M. Motoki, and B. Zhang, "Balancing fairness and efficiency in energy resource allocations," in *IEEE Conference on Decision and Control*, 2024.

[72] B. Hobbs, U. Helman, and J.-S. Pang, "Equilibrium market power modeling for large scale power systems," in *2001 Power Engineering Society Summer Meeting. Conference Proceedings (Cat. No. 01CH37262)*, vol. 1. IEEE, 2001, pp. 558–563.

[73] E. Wei, A. Malekian, and A. Ozdaglar, "Competitive equilibrium in electricity markets with heterogeneous users and price fluctuation penalty," in *53rd IEEE Conference on Decision and Control*, 2014, pp. 6452–6458.

[74] R. Rodríguez, M. Negrete-Pincetic, N. Figueroa, Á. Lorca, and D. Olivares, "The value of aggregators in local electricity markets: A game theory based comparative analysis," *Sustainable Energy, Grids and Networks*, vol. 27, p. 100498, 2021.

[75] K. Bruninx, H. Pandžić, H. Le Cadre, and E. Delarue, "On the interaction between aggregators, electricity markets and residential demand response providers," *IEEE Transactions on Power Systems*, vol. 35, no. 2, pp. 840–853, 2019.

[76] S. H. Low, F. Paganini, and J. C. Doyle, "Internet congestion control," *IEEE control systems magazine*, vol. 22, no. 1, pp. 28–43, 2002.

[77] T. Pinto, H. Morais, P. Oliveira, Z. Vale, I. Praça, and C. Ramos, "A new approach for multi-agent coalition formation and management in the scope of electricity markets," *Energy*, vol. 36, no. 8, pp. 5004–5015, 2011.

[78] T. Wolff and A. Nieße, "Dynamic overlapping coalition formation in electricity markets: An extended formal model," *Energies*, vol. 16, no. 17, p. 6289, 2023.

[79] X. Lu, K. Li, H. Xu, F. Wang, Z. Zhou, and Y. Zhang, "Fundamentals and business model for resource aggregator of demand response in electricity markets," *Energy*, vol. 204, p. 117885, 2020.

[80] C. Chen, S. Bose, T. D. Mount, and L. Tong, "Wholesale market participation of deras: Dso-dera-iso coordination," *IEEE Transactions on Power Systems*, 2024.

[81] S. P. Mathur, A. Arya, and M. Dubey, "Optimal bidding strategy for price takers and customers in a competitive electricity market," *Cogent Engineering*, vol. 4, no. 1, p. 1358545, 2017.

[82] D. Zhao, M. Jafari, A. Botterud, and A. Sakti, "Strategic energy storage investments: A case study of the caiso electricity market," *Applied Energy*, vol. 325, p. 119909, 2022.

[83] G. Debreu, "A social equilibrium existence theorem," *Proceedings of the national academy of sciences*, vol. 38, no. 10, pp. 886–893, 1952.

[84] M. J. Osborne, *A Course in Game Theory*. MIT Press, 1994.

[85] R. A. Berry, R. Johari *et al.*, "Economic modeling in networking: A primer," *Foundations and Trends® in Networking*, vol. 6, no. 3, pp. 165–286, 2013.

[86] A. Cambini and L. Martein, *Generalized convexity and optimization: Theory and applications*. Springer Science & Business Media, 2008, vol. 616.

[87] P. Li, H. Wang, and B. Zhang, "A distributed online pricing strategy for demand response programs," *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 350–360, 2017.

[88] B. S. K. Patnam and N. M. Pindoriya, "Demand response in consumer-centric electricity market: Mathematical models and optimization problems," *Electric Power Systems Research*, vol. 193, p. 106923, 2021.

[89] A. Agrawal and S. Boyd, "Disciplined quasiconvex programming," *Optimization Letters*, vol. 14, no. 7, pp. 1643–1657, 2020.

[90] M. Davis, "Solar market insight report," *Wood Mackenzie Power and Renewables*, 2021.

[91] K. Turitsyn, P. Sulc, S. Backhaus, and M. Chertkov, "Options for control of reactive power by distributed photovoltaic generators," *Proceedings of the IEEE*, vol. 99, no. 6, pp. 1063–1073, 2011.

[92] H.-G. Yeh, D. F. Gayme, and S. H. Low, "Adaptive var control for distribution circuits with photovoltaic generators," *IEEE Transactions on Power Systems*, vol. 27, no. 3, pp. 1656–1663, 2012.

[93] B. Zhang, A. Y. Lam, A. D. Domínguez-García, and D. Tse, "An optimal and distributed method for voltage regulation in power distribution systems," *IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 1714–1726, 2014.

[94] N. Li, G. Qu, and M. Dahleh, "Real-time decentralized voltage control in distribution networks," in *2014 52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2014, pp. 582–588.

[95] S. Bolognani and S. Zampieri, "A distributed control strategy for reactive power compensation in smart microgrids," *IEEE Transactions on Automatic Control*, vol. 58, no. 11, pp. 2818–2833, 2013.

[96] H. Zhu and H. J. Liu, "Fast local voltage control under limited reactive power: Optimality and stability analysis," *IEEE Transactions on Power Systems*, vol. 31, no. 5, pp. 3794–3803, 2015.

[97] G. Qu and N. Li, "Optimal distributed feedback voltage control under limited reactive power," *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 315–331, 2019.

[98] Y.-Y. Hsu and F.-C. Lu, "A combined artificial neural network-fuzzy dynamic programming approach to reactive power/voltage control in a distribution substation," *IEEE transactions on Power Systems*, vol. 13, no. 4, pp. 1265–1271, 1998.

[99] S. Toma, T. Senjyu, Y. Miyazato, A. Yona, K. Tanaka, and C.-H. Kim, "Decentralized voltage control in distribution system using neural network," in *2008 IEEE 2nd International Power and Energy Conference*. IEEE, 2008, pp. 1557–1562.

[100] X. Shen, H. Wang, J. Li, Q. Su, and L. Gao, "Distributed secondary voltage control of islanded microgrids based on rbf-neural-network sliding-mode technique," *IEEE Access*, vol. 7, pp. 65 616–65 623, 2019.

[101] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[102] Q. Yang, G. Wang, A. Sadeghi, G. B. Giannakis, and J. Sun, "Two-timescale voltage control in distribution grids using deep reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2313–2323, 2019.

[103] Y. Gao, W. Wang, and N. Yu, "Consensus multi-agent reinforcement learning for volt-var control in power distribution networks," *IEEE Transactions on Smart Grid*, 2021.

[104] J. Duan, D. Shi, R. Diao, H. Li, Z. Wang, B. Zhang, D. Bian, and Z. Yi, "Deep-reinforcement-learning-based autonomous voltage control for power grid operations," *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 814–817, 2019.

[105] N. E. M. Association *et al.*, *American National Standard for Electric Power Systems and Equipment-Voltage Ratings (60 Hertz)*. National Electrical Manufacturers Association, 1996.

[106] A. Vaccaro, G. Velotto, and A. F. Zobaa, "A decentralized and cooperative architecture for optimal voltage regulation in smart grids," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 10, pp. 4593–4602, 2011.

[107] M. Jafari, T. O. Olowu, and A. I. Sarwat, "Optimal smart inverters volt-var curve selection with a multi-objective volt-var optimization using evolutionary algorithm approach," in *2018 North American Power Symposium (NAPS)*. IEEE, 2018, pp. 1–6.

[108] C. Zhao, U. Topcu, N. Li, and S. Low, "Design and stability of load-side primary frequency control in power systems," *IEEE Transactions on Automatic Control*, vol. 59, no. 5, pp. 1177–1189, 2014.

[109] E. Mallada, C. Zhao, and S. Low, "Optimal load-side control for frequency regulation in smart grids," *IEEE Transactions on Automatic Control*, vol. 62, no. 12, pp. 6294–6309, 2017.

[110] Y. Shi, B. Xu, D. Wang, and B. Zhang, "Using battery storage for peak shaving and frequency regulation: Joint optimization for superlinear gains," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 2882–2894, 2017.

[111] B. B. Johnson, S. V. Dhople, A. O. Hamadeh, and P. T. Krein, "Synchronization of parallel single-phase inverters with virtual oscillator control," *IEEE Transactions on Power Electronics*, vol. 29, no. 11, pp. 6124–6138, 2013.

[112] J. Yu, Y. Weng, and R. Rajagopal, "Patopaem: A data-driven parameter and topology joint estimation framework for time-varying system in distribution grids," *IEEE Transactions on Power Systems*, vol. 34, no. 3, pp. 1682–1692, 2018.

[113] J. Zhang, Y. Wang, Y. Weng, and N. Zhang, "Topology identification and line parameter estimation for non-pmu distribution network: A numerical method," *IEEE Transactions on Smart Grid*, vol. 11, no. 5, pp. 4440–4453, 2020.

[114] S. Lin and H. Zhu, "Data-driven modeling for distribution grids under partial observability," *arXiv preprint arXiv:2108.08350*, 2021.

[115] R. Kavet, "Characterization of radiofrequency emissions from two models of wireless smartmeters," *Project Manager Electric Power Research Institute, EPRI*, 2011.

[116] W. Cui and B. Zhang, "Reinforcement learning for optimal frequency control: A lyapunov approach," *arXiv preprint arXiv:2009.05654*, 2020.

[117] M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Power Engineering Review*, vol. 9, no. 4, pp. 101–102, 1989.

[118] B. A. Robbins, C. N. Hadjicostis, and A. D. Domínguez-García, "A two-stage distributed architecture for voltage control in power distribution systems," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1470–1482, 2012.

[119] J. Li, M. Motoki, and B. Zhang, "Socially optimal energy usage via adaptive pricing," *Electric Power Systems Research*, vol. 235, p. 110640, 2024.

[120] W. Cui, J. Li, and B. Zhang, "Decentralized safe reinforcement learning for inverter-based voltage control," *Electric Power Systems Research*, vol. 211, p. 108609, 2022.