UNIVERSITY OF CALIFORNIA,
IRVINE

Fairness-aware Machine Learning in Power Grids

THESIS

submitted in partial satisfaction of the requirements
for the degree of

MASTER OF SCIENCE

in Electrical and Computer Engineering

by

Ruijie Du

Thesis Committee:
Assistant Professor Yanning Shen, Chair
Professor Pramod Khargonekar
Professor A. Lee Swindlehurst

2023

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Page

LIST OF ALGORITHMS

ACKNOWLEDGMENTS

I would like to thank my chair and advisor, Dr. Yanning Shen, whose unreserved support and guidance made this thesis possible. It has been a great pleasure working closely with her over the past 2 years.

I would like to express my deep gratitude to Dr. Pramod Khargonekar for his extremely generous support. I have learned so much from him in research, teaching and many more.

I am very grateful to Dr. A. Lee Swindlehurst for the inspiring suggestions and very helpful feedback.

Special thanks also to Dr. Deepan Muthirayan for his helpful advice in research. Working with him has been an enjoyable experience.

I extend my gratitude to my friends, Karl Liu, Colin Liang and Annie Su, for their support and encouragement in my life.

This work is dedicated to my family.

VITA

Ruijie Du

EDUCATION

Master of Science in Electrical and Computer Engineering **2023**
University of California, Irvine *Irvine, California*

Bachelor of Science in Electrical and Computer Engineering **2020**
Rensselaer Polytechnic Institute *Troy, NY*

RESEARCH EXPERIENCE

Graduate Research Assistant **2021–Presnet**
University of California, Irvine *Irvine, California*

Undergraduate Research **2019–2020**
Rensselaer Polytechnic Institute *Troy, NY*

ABSTRACT OF THE THESIS

Fairness-aware Machine Learning in Power Grids

By

Ruijie Du

Master of Science in Electrical and Computer Engineering

University of California, Irvine, 2023

Assistant Professor Yanning Shen, Chair

With the development of intelligent measurement systems, power grids improved reliability and efficiency according to the vast amount of collected information. Machine learning(ML) techniques are increasingly used in smart grids since they are efficient in dealing with huge amounts of data and extracting valuable information. However, large-scale deployment of ML models relies on how trustworthy the model is. While sole pursuit of overall learning performance may lead to unfair results. Specifically, the model may unintentionally discriminate different subgroups. To mitigate the unfairness, we propose accuracy parity, equal opportunity and predictive equality regularizers, which can be used for different classification tasks in power grids to mitigate the corresponding discrepancy.

However, most tasks in power grids cannot be formulated as classification tasks. Instead, more practical tasks like regression and decision-making take precedence. When addressing the fairness of dynamic decision-making problems over a continuous time scale, two different fairness objectives naturally arise: the instantaneous fairness objective that aims to ensure fairness at every time slot and the long-term fairness objective that aims to sustain fairness over a period. Long-term fairness becomes increasingly crucial due to its adaptability and applicability. We formulate the problem as an online optimization problem with the long-term fairness constraint and propose an algorithm to tackle it. The proposed method analytically yields sub-linear dynamic regret and sub-linear accumulated fair violations.

Chapter 1

Introduction

With the development of intelligent measurement systems and smart metering technology, migrating to an electronically controlled grid has yielded significant improvements in reliability and efficiency [8]. The evolution of power grids has led to a more efficient data collection process. Machine learning techniques provide an efficient way to analyze a massive amount of data and extract valuable information. By analyzing the measurements, more information can be obtained more accurately: the status of the network, the actual detailed load patterns, etc. The increasing reliance on machine learning techniques has brought about significant benefits to power grids in various tasks, including predictions of consumption [33, 27], fault detection [14], etc.

However, as decision-making increasingly relies on machine learning and the utilization of data, the issue of fairness is receiving increasing attention. In classical machine learning, when considering a classification task, the objective is to minimize a loss function(e.g., cross entropy) that reflects the errors. Unfortunately, this approach is unable to control the distribution of errors across different subgroups. This lack of control over subgroup-specific errors has raised concerns about unintentional discrimination and unfairness in decision-making

processes. In recent years, research has pointed out plenty of evidence that decision making by machine learning models may unintentionally discriminate different subgroups and causes unfairness, especially in media and social studies. In settings such as loan approvals[15] or college admissions[2], fairness must be carefully taken into account in order to ensure the absence of discrimination. To tackle this issue, more and more fairness-aware machine learning solutions have been proposed [41, 40, 1, 42]. However, the issue of potential bias has not yet been extensively explored or considered in power grids. While the development of smart metering technology provides more convenient data collection, it also leads to potential bias since more detailed and private data may be revealed. Moreover, the prevalence of smart meters and sensors varies widely among users or locations, potentially leading to biased data collection practices. Consequently, machine Learning models trained from such datasets may henceforth inherit the bias in the collected data. In Chapter 3, a novel fairness-aware framework is introduced to eliminate potential bias in power grid user classification.

To extend fair machine learning to dynamic systems, such as network systems and power grids, some recent works study the long-term impact of fairness in various applications, such as [4, 10, 36, 44, 39, 38]. Most fairness-aware machine learning frameworks have predominantly focused on classification settings within ecosystems. However, in many real cyber-physical systems, the decisions in these systems are often bound by physical limitations and cannot be easily framed as classification tasks, such as [30, 6, 7]. For achieving long-term fairness in different systems, most approaches focus on the problems with dynamic costs but time-invariant fair constraints but do not consider the problem with time-varying fair constraints. However, the fair constraint might vary with the dynamics of the systems and decisions [44]. To address the need for long-term fairness in dynamic decision-making systems with time-varying fair constraints, Chapter 4 introduces a novel framework to ensure long-term fairness. By adapting to the dynamics of the system and considering time-varying fair constraints, this framework aims to provide sustainable fairness in dynamic environments.

The thesis is structured as follows: Chapter 1 provides an overview of the thesis. Chapter 2 provides the study background and the related literature. Chapter 3 addresses the fairness-aware user classification in power grids which introduces two regularizers to eliminate potential bias. Chapter 4 addresses the long-term fairness-aware decision-making in power grids and aims to ensure long-term fairness with time-varying fair constraints. Chapter 5 concludes the thesis and discusses the limitations and future directions.

Chapter 2

Background

This chapter presents the study background and the related literature about machine learning and fairness in power grids.

The power grid serves as a vital infrastructure, enabling the generation, transmission, and distribution of electricity to meet the growing energy demands of modern society [8]. As power systems become increasingly complex and face new challenges such as integrating renewable energy sources and managing fluctuating electricity demand, the adoption of machine learning techniques has emerged as a promising avenue to enhance grid operations, improve reliability, and optimize energy management [7, 31].

Two of my studies about fairness-aware machine learning in power grids, one on the fairness-aware classification (Chapter 3) and another one on long-term fairness-aware decision making (Chapter 4).

2.1 Machine Learning in Power Grids

Machine learning, a branch of artificial intelligence, encompasses a range of algorithms and methodologies that enable computers to learn from data and make predictions or take actions. In the context of power grids, machine learning techniques have the potential to revolutionize various aspects of grid operations, including predictive maintenance [35], demand response [24], energy management [6], etc.

Demand Response: Demand response programs play a crucial role in balancing electricity supply and demand, particularly with the increasing penetration of renewable energy sources and the variability of electricity consumption patterns. Machine learning techniques enable accurate demand forecasting by considering various factors such as weather patterns and historical consumption data. This enables utilities to optimize energy dispatch, load shifting, and effectively manage peak demand periods.

Energy Management: Efficient energy management is a key consideration in modern power grids. Machine learning algorithms can analyze historical energy consumption patterns, customer data, and external factors such as weather forecasts to develop predictive models for load forecasting [4]. These models help utilities in resource planning, grid expansion, and optimal resource allocation. Additionally, machine learning can support energy optimization strategies by optimizing the operation of distributed energy resources, energy storage systems, and microgrids.

The current power systems are undergoing a transition towards their more active, flexible, and intelligent counterpart smart grid, which brings about tremendous challenges in many domains, e.g., integration of various distributed renewable energy sources [18], demand-side management [24], and decision-making of system planning and operation. The concept of a smart grid, as Figure 2.1, has emerged as a complex cyber-physical system encompassing a huge electrical power network with the underlying information and communication system

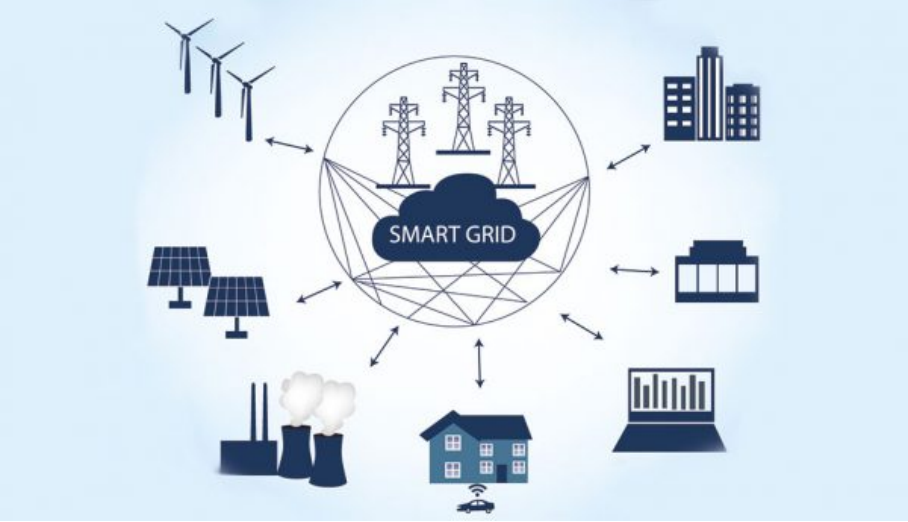


Figure 2.1: Concept of smart grid.

[11, 37]. The fulfillment of advanced functionalities in the smart grid firmly relies on the underlying information and communication infrastructure, and the efficient handling of a massive amount of data generated from various sources, e.g., smart meters and various forms of sensors [13].

The integration of machine learning techniques into power grids offers significant opportunities to enhance grid operations, improve reliability, optimize energy management, and strengthen cybersecurity. By leveraging advanced analytics, pattern recognition capabilities, and predictive modeling, machine learning empowers utilities to make data-driven decisions, automate processes, and enable more efficient utilization of grid resources. As the power grid continues to evolve and face new challenges, machine learning will play a crucial role in shaping the future of power systems and enabling a more sustainable and resilient energy infrastructure.

Machine learning techniques provide an efficient way to analyze the massive amount of data and extract valuable information. The real-time operational condition monitoring and efficient data analysis can greatly strengthen the system management for many aspects, including predictions of consumption [33, 27], fault detection [14], management of load demands

[24], etc.

2.2 Fairness in Power Systems

As power systems continue to evolve, there is a growing recognition of the importance of fairness in the design, operation, and regulation of these systems. Fairness encompasses the equitable allocation of resources, benefits, and costs among different stakeholders, ensuring that all individuals and communities have equal access to reliable and affordable electricity services. Achieving fairness in power systems is not only a matter of social justice but also crucial for promoting sustainability, inclusivity, and economic development.

As decision-making increasingly relies on machine learning and data, the issue of fairness is receiving increasing attention. In recent years, research has pointed out plenty of evidence that decision making by machine learning models may unintentionally discriminate different subgroups and causes unfairness, especially in media and social studies [41, 40, 1, 42].

The issue of potential bias also exists in power grids. With the increasing use of data analytics and machine learning in power systems, there is a need to address data bias and algorithmic fairness concerns. Data-driven decision-making processes should be carefully monitored to ensure that they do not inherit existing biases or lead to discriminatory outcomes. Power system decisions not only affect the overall system, but also individual end-users, such as generators, flexibility providers and end-consumers, depending on their location and characteristics [9]. Though the importance of fairness is increasingly recognized in the power system literature [9, 32], most machine learning models in power grids have not considered it yet. In order to achieve more fair results in decision making, such as energy management and resource allocation in power systems, existing machine learning schemes need to introduce additional bias-mitigating schemes before they can be readily applied to the grids.

Fairness is a fundamental principle that should underpin the design, operation, and regulation of power systems. By addressing challenges related to access, affordability, grid planning, renewable energy transitions, and algorithmic fairness, power systems can be transformed into more equitable and inclusive frameworks. Achieving fairness requires collaboration between policymakers, regulators, utilities, and community organizations to identify and implement strategies that promote equal access to reliable and affordable electricity services, regardless of socioeconomic status or geographic location.

Chapter 3

Fairness-aware User Classification in Power Grids

In this chapter, we examined and showcased the existing bias in directly applying neural network models, then two types of regularizers were introduced to promote the fairness in user classification tasks in power grids.

3.1 Introduction

Machine learning techniques are increasingly used in smart grids since they are efficient to deal with the huge amount of collected data and extract valuable information. The availability of large-scale data enables the employment of machine learning methods in various tasks in power grids [37]. However, large-scale deployment of machine learning model relies on how trustworthy the model is. Machine learning models for smart grids also have fairness concerns. Power consuming users and buildings with different power consumption patterns may be treated with different conditions. To mitigate the unfairness, we propose accuracy

parity, equal opportunity and predictive equality regularizers, which can be used for different classification tasks in power grids to mitigate the corresponding performance discrepancy. The chapter is organized as follows: Section 3.2 introduces the problem of high load indication, and proposes the fairness enhancing framework. Section 3.3 focuses on fairness-aware user type identification task. Numerical tests are presented in Section 3.4 to evaluate the performance of the proposed schemes.

3.2 Fairness-aware High-load Indication

High-load indication task refers to the task of predicting whether the load of the next time period is high-load state or not based on current and previous load information, see e.g., [19]. Knowing the loading status or behavioral categories of power consumers can better model the behavior forecast which is an important task for load balancing. The input information is the load information of T time frames, including the load information from now to $T-1$ time frames ago. And the output is the binary indicator $y \in \{0,1\}$ of high-load status at time $T+1$, where $y=1$ indicates high-load status.

Specifically, such task can be viewed as a binary classification problem, where the objective is to minimize a loss function that reflects the errors made by the classifier F. Specifically, the loss function considered in the present work is the cross-entropy loss

$$\mathcal{L}(\theta;\mathcal{D}) = -\frac{1}{|\mathcal{D}|}\sum_{\mathbf{x}\in\mathcal{D}}(y\log(F(\mathbf{x};\theta))+(1-y)\log(1-F(\mathbf{x};\theta)))\,, \tag{3.1}$$

where $\mathbf{x}\in\mathbb{R}^n$ is the input feature of a data sample and $y\in\{0,1\}$ is the ground truth label corresponding to $\mathbf{x}$. In high-load indication task, $\mathbf{x}$ represents time-series load data and y represents the load status. θ denotes the parameters of the classifier that will be trained to minimize the loss function. $F(\mathbf{x};\theta)$ is the output of the classifier which represents the

predicted probability of $\mathbf{x}$ being class 1. Because the classification is binary, $1 - F(\mathbf{x}; \theta)$ is the predicted probability of $\mathbf{x}$ being class 0. Such loss is widely used for classification tasks including but not limited to anomaly detection [43], quality prediction [35], fault detection [14], or load indication [19, 28].

3.2.1 Accuracy Parity Regularizer

Existing frameworks for load indication in power grids usually focus on how to design the classifier F with lower error rate, but they are usually bias-oblivious. This will lead to potential bias in the results. The classification of loading status is based on their consumption patterns which potentially have fairness issues due to the sensitive attributes, e.g. building sizes, geographical locations, user types.

Let $z \in \{1, 0\}$ denote the sensitive attribute, which can denote, e.g., building size or location. Let $d(\mathbf{x})$ represent the signed distance between the feature vectors of samples and the classifier decision boundary [40]. The covariance between the sensitive attributes z and $d(\mathbf{x})$ can be written as

$$\begin{aligned} \mathrm{Cov}(z, d(\mathbf{x})) =& E[(z - \bar{z})g(y, \mathbf{x})] - E[(z - \bar{z})]\bar{g}(y, \mathbf{x}) \\ \approx& \frac{1}{|\mathcal{D}|} \sum_{(\mathbf{x}, y, z) \in \mathcal{D}} (z - \bar{z})g(y, \mathbf{x}), \end{aligned} \tag{3.2}$$

where $\bar{z}$ and $\bar{g}(y, \mathbf{x})$ denote the average of the value z and $g(y, \mathbf{x})$ for all data samples in $\mathcal{D}$, $g(y, \mathbf{x}) := \max(0, (\frac{1}{2} - y)d(\mathbf{x}))$, $E[(z - \bar{z})]\bar{g}(y, \mathbf{x}) = 0$ since $E[(z - \bar{z})] = 0$. In neural networks for binary classification, the final decision is based on the output of last linear layer, denoted as $f(\mathbf{x})$. The final output of the classifier $F(\mathbf{x})$ is obtained as the output of the last logistic activation function with input $f(\mathbf{x})$. Hence, $\hat{y} = 1$ if $f(\mathbf{x}) > 0$, otherwise $y = 0$. The decision boundary is simply the hyperplane that $f(\mathbf{x}) = 0$. Hence $d(\mathbf{x}) = f(\mathbf{x}) - 0 = f(\mathbf{x})$

and $g(y, \mathbf{x}) := \max(0, (\frac{1}{2} - y)f(\mathbf{x}))$. Splitting the sum in Equation 3.2 into two terms with respect to the sensitive attribute z, we obtain

$$\sum_{\mathcal{D}}(z - \bar{z})g(y, \mathbf{x}) = \sum_{z=0}(0 - \bar{z})g(y, \mathbf{x}) + \sum_{z=1}(1 - \bar{z})g(y, x). \tag{3.3}$$

It can be readily observed that if the prediction matches the true label $\hat{y} = y$, we have $f(\mathbf{x}) > 0$ for $y = 1$ and $f(\mathbf{x}) < 0$ for $y = 0$. Hence, $(\frac{1}{2} - y)f(\mathbf{x}) < 0$. Due to the maximum operation, $g(y, \mathbf{x}) = 0$. Similarly, $g(y, \mathbf{x}) > 0$ if $\hat{y} \neq y$. Meanwhile, the term $(z - \bar{z})$ takes different signs for different sensitive groups: For the subgroup with $z = 0$, $(0 - \bar{z})g(y, \mathbf{x}) \leq 0$; for the subgroup with $z = 1$, $(1 - \bar{z})g(y, \mathbf{x}) \geq 0$. Hence, the addition of the two terms in Equation 3.3 characterizes the difference of the mismatch accumulation between two subgroups.

If a decision boundary satisfies accuracy parity (AP) in the sense that $P(\hat{y} \neq y|z = 0) = P(\hat{y} \neq y|z = 1)$, the covariance will be close to zero, $\mathrm{Cov}(z, d(\mathbf{x})) \approx 0$.Therefore, the accuracy parity regularizer can be introduced as

$$\begin{aligned}
\mathcal{R}_{\mathrm{AP}} &= \left(\frac{1}{|\mathcal{D}|}\sum(z - \bar{z})g(y, \mathbf{x})\right)^2 \\
&= \left(\frac{1}{|\mathcal{D}|}\sum(z - \bar{z})\max(0, (\frac{1}{2} - y)f(\mathbf{x}))\right)^2 \\
&= \left(\frac{1}{|\mathcal{D}|}\sum_{z=0}\max(0, (0 - \bar{z})(y - \frac{1}{2})f(\mathbf{x})) - \frac{1}{|\mathcal{D}|}\sum_{z=1}\max(0, (1 - \bar{z})(\frac{1}{2} - y)f(\mathbf{x}))\right)^2.
\end{aligned} \tag{3.4}$$

The penalty comes from the difference of the mismatch accumulation between the sensitive groups. If the model is more fair, the performance of two sensitive groups will be more similar and the penalty will be smaller. Hence, the problem can be formulated as:

$$\min_{\theta} \quad \mathcal{L}(\theta; \mathcal{D}) + \alpha\mathcal{R}_{\mathrm{AP}}, \tag{3.5}$$

where $\alpha > 0$ is a hyperparameter used to tradeoff between training accuracy and fairness in

terms of accuracy parity.

3.2.2 Equal Opportunity and Predictive Equality Regularizers

In addition to the accuracy parity criterion presented in the previous subsection, in certain scenarios, we may be interested in different fairness criteria. For example, in anomaly detection, where costly immediate actions may be taken towards the high-load states, more emphasis needs to be put on the classes that are predicted positive. In this case, fairness criteria such as equal opportunity or predictive equality need to be incorporated. Specifically, equal opportunity $\Delta_{\text{EO}} := |P(\hat{y} = 1|y = 1, z = 0) - P(\hat{y} = 1|y = 1, z = 1)|$ characterizes the difference of true positive rate $TRP := P(\hat{y} = 1|y = 1)$, while predictive equality $\Delta_{\text{PE}} := |P(\hat{y} = 1|y = 0, z = 0) - P(\hat{y} = 1|y = 0, z = 1)|$ measures the discrepancy between the false positive rate $FPR := P(\hat{y} = 1|y = 0)$ see e.g., [1]. Based on these two criteria, we can obtain the corresponding regularizers as

$$\mathcal{R}_{\text{PE}} = \left(\frac{\sum_{z=0,y=0} f(\mathbf{x})}{N_{z=0,y=0}} - \frac{\sum_{z=1,y=0} f(\mathbf{x})}{N_{z=1,y=0}} \right)^2, \tag{3.6}$$

$$\mathcal{R}_{\text{EO}} = \left(\frac{\sum_{z=0,y=1} f(\mathbf{x})}{N_{z=0,y=1}} - \frac{\sum_{z=1,y=1} f(\mathbf{x})}{N_{z=1,y=1}} \right)^2. \tag{3.7}$$

Similarly, $f(\mathbf{x})$ is the output of the last linear layer. β_1 and β_2 are the weights of regularization terms which trade off between fairness and accuracy. Adding the equal opportunity and predictive equality regularizers, the problem can be formulated as:

$$\min_{\theta} \mathcal{L}(\theta; \mathcal{D}) + \beta_1 \mathcal{R}_{\text{PE}} + \beta_2 \mathcal{R}_{\text{EO}}. \tag{3.8}$$

3.3 Fairness-aware User Type Identification

There exist large numbers of users in power grids, which fall into different classes, e.g., building types, safety or quality status. Hence, user type identification plays an important role in many practical problems, including but not limited to accurate pricing, abnormal behavior detection. Accurate identification of the consumer types can also help estimate their future consumption which can further help the power companies balance the load of power grids and manage the demand and supply. However, existing user type classification frameworks mainly focus on imputing classification accuracy, see e.g., [34]. While the user consumption patterns can be used for classifying the user types, it may also lead to potential bias due to underlying correlation between users' consumption patterns and their sensitive attributes such as locations, building sizes. Since user type identification typically faces with customers from more than one category, the corresponding cross-entropy loss can be written as

$$\mathcal{L}(\theta; \mathcal{D}) = -\frac{1}{|\mathcal{D}|} \sum_{i}^{K} \sum_{\mathbf{x} \in \mathcal{D}} \mathbb{1}_{y=i} log(F_i(\mathbf{x}; \theta)), \tag{3.9}$$

where $F(\mathbf{x}; \theta) \in \mathbb{R}^K$ and $y \in \{1, \ldots, K\}$, with K denoting the number of classes. And $\mathbb{1}_{y=i}$ is the identify function, which returns value 1 if $y = i$, meaning the input data belongs to class i. And $F_i(\mathbf{x}; \theta)$ denotes the ith entry of $F(\mathbf{x}; \theta)$ which provides the predicted probability of $\mathbf{x}$ being class i. In user type identification task, $\mathbf{x}$ represents the load time-series data and y represents user types. The sensitive attribute z denotes the geographical location of the user. In this case the output of last fully connected layer is a vector of size K, $f(\mathbf{x}) \in \mathbb{R}^K$. Hence the distance to the decision boundary for class i, $1 \leq i \leq K$, is the ith entry of $f(\mathbf{x})$, denoted as $f_i(\mathbf{x})$.

3.3.1 Accuracy Parity Regularizer

Faced with multiple user classes, the accuracy parity regularizer also needs to be designed for multiple user classes. Specifically, the accuracy parity regularizer can be written as follows:

$$\mathcal{R}_{\text{AP}} = \left(\frac{1}{|D|}\sum(z-\bar{z})g(y,\mathbf{x})\right)^2 = \left[\frac{1}{|\mathcal{D}|}\sum(z-\bar{z})\max\left(0,(f_{\hat{y}(\mathbf{x})}-f_y(\mathbf{x}))\right)\right]^2. \quad (3.10)$$

Since the final decision is based on the maximum value of $f(\mathbf{x})$, $f_{\hat{y}}(\mathbf{x})$, denotes the $\hat{y}$th entry of $f(\mathbf{x})$, is greater than any other value in the vector, $f_{\hat{y}}(\mathbf{x}) \geq f_i(\mathbf{x}), \forall i \in \{1,\ldots,K\}$. The term contributes to the sum only when the predicted label mismatches the true label. Due to the fact that $f_{\hat{y}}(\mathbf{x}) - f_y(\mathbf{x}) \geq 0$, and the equality holds if $\hat{y} = y$, meaning the predicted label correct. Therefore, $\max(0,(f_{\hat{y}}(\mathbf{x}) - f_y(\mathbf{x}))) = f_{\hat{y}}(\mathbf{x}) - f_y(\mathbf{x})$, and the max operation can be removed. The problem can be formulated as Equation 3.5 with loss function in Equation 3.9 and the regularization $\mathcal{R}_{\text{AP}}$ in Equation 3.10.

3.3.2 Equal Opportunity and Predictive Equality Regularizers

For equal opportunity and predictive equality regularizers, we treat every class as a one versus $K-1$ binary classification problem and aggregate the regularization term for K classes. For each class, the two penalizers are similar to Equation 3.6 and Equation 3.7 except that the distance to the decision boundary is $f_i(\mathbf{x})$ instead of $f(\mathbf{x})$. Then the EO and PE regularizer for multiple classes can be written as:

$$\mathcal{R}_{\text{PE,i}} = \left(\frac{\sum_{z=0,y\neq i} f_i(\mathbf{x})}{N_{z=0,y\neq i}} - \frac{\sum_{z=1,y\neq i} f_i(\mathbf{x})}{N_{z=1,y\neq i}}\right)^2, \quad \mathcal{R}_{\text{PE}} = \frac{\beta_1}{K}\sum_{i=0,\ldots,K-1}\mathcal{R}_{\text{PE,i}}; \quad (3.11)$$

$$\mathcal{R}_{\text{EO,i}} = \left(\frac{\sum_{z=0,y=i} f_i(\mathbf{x})}{N_{z=0,y=i}} - \frac{\sum_{z=1,y=i} f_i(\mathbf{x})}{N_{z=1,y=i}}\right)^2, \quad \mathcal{R}_{\text{EO}} = \frac{\beta_2}{K}\sum_{i=0,\ldots,K-1}\mathcal{R}_{\text{EO,i}}. \quad (3.12)$$

Upon adding the EO and PE regularizers, the problem can be formulated as Equation 3.8 by employing the loss in Equation 3.9 and the regularization $\mathcal{R}_{\text{EO}}$, $\mathcal{R}_{\text{PE}}$ in Equation 3.11, Equation 3.12.

3.4 Experiments

In this section, experimental results for high-load indication and user type identification are presented to evaluate the proposed fairness-aware framework. Specifically, details of data preprocessing and experimental settings will be clarified.

Dataset: The data was obtained from the database "Commercial and Residential Hourly Load Profiles for TMY3 Locations in the United States" [23]. It consists of hourly collected load profile data for 16 different commercial building types and residential buildings. The commercial buildings data is based on the DOE commercial reference building models and the residential buildings data is based on the Building America House Simulation Protocols. Specifically, the data consists 7 types (classes) of commercial buildings in several cities in the US. The input feature $\mathbf{x}$ is the load information of a building.

3.4.1 Fariness-aware Measurements

In order to evaluate fairness performance, the equal opportunity (EO), predictive equality (PE) and accuracy parity (AP) measures are used, with $\Delta_{EO} := |P(\hat{y} = 1|y = 1, z = 1) - P(\hat{y} = 1|y = 1, z = 0)|$, $\Delta_{PE} := |P(\hat{y} = 1|y = 0, z = 1) - P(\hat{y} = 1|y = 0, z = 0)|$. Similarly, AP measures the difference between two subgroups' accuracy: $\Delta_{AP} = |P(\hat{y} = y|z = 0) - P(\hat{y} = y|z = 1)|$. For the user type identification task, although the number of classes increases, the Δ_{AP} stays the same. In order to measure the EO fairness among multiple classes, taking every class as a binary classification case and weighted aggregate the

difference between two subgroups' TPR as EO measurement: $\Delta_{EO} = \sum_i^K w_i \Delta_{EO,i}$ where $\Delta_{EO,i} = |TPR_{y=i,z=0} - TPR_{y=i,z=1}|$ and $TPR_i = P(\hat{y} = i, y = i)/P(y = i)$. Similarly, the PE measurement for multiple classes is defined as $\Delta_{PE} := \sum_i^K w_i \Delta_{PE,i}$ where $\Delta_{PE,i} = |FPR_{y=i,z=0} - FPR_{y=i,z=1}|$ and $FPR_i = P(\hat{y} = i, y \neq i)/P(y \neq i)$. The weight of each class, $w_i := \frac{N_i}{N}$, is based on the distribution of each class across the dataset, N_i is the number of samples in class i and N is the total number of samples.

3.4.2 High Load Indication

The task here is predicting next hour load status of the building based on its history loading pattern. We use the long short-term memory(LSTM) network as the first layer of the model. LSTM is an artificial recurrent neural network (RNN) architecture that can provide an internal memory for the networks. The output of the LSTM goes through 3 fully connected layers with ReLU activation function in hidden layers and Sigmoid in the last layer to get the final result. The load information of a building in the last 12 hours is denoted as $\mathbf{x}$. The high load status of water heat load in the next time hour is indicated as y: If it is high load, $y = 1$; otherwise, $y = 0$.

Table 3.1 and Table 3.2 list the results where building type and building location are treated as sensitive attributes respectively. Specifically, In Table 3.1, $z = 0$ represents large building; $z = 1$ represents small building. In Table 3.2, z denotes which state the building is in: $z = 1$ indicates the building is in New York, $z = 0$ for buildings in California. All algorithms were run with training and test ratio of $4 : 1$ where a random set of periodic loading sequences is sampled from the original dataset. Experimental results were averaged over 4 random runs.

Table 3.1 and Table 3.2 list the results evaluated on testing set in terms of accuracy(Acc), Δ_{PE}, Δ_{EO} and Δ_{AP}. In each row, the number in the first column refers to the equation used for training. Equation 3.1 represents the vanilla logistic classifier. Equation 3.5 and

Equation 3.8 are trained with the corresponding regularizers. It can be observed from Table 3.1 and 3.2 that the proposed regularizers improve the fairness without degrading the classification performance.

Table 3.1: Performance of high load indication with building type as the sensitive attribute.

	Acc(%)	$\Delta_{PE}(e^{-2})$	$\Delta_{EO}(e^{-2})$	$\Delta_{AP}(e^{-2})$
Equation 3.1	87.04±0.5	5.6±1.2	8.3±3.3	1.25±0.7
Equation 3.5	87.6±1.2	4.7±1.5	9.8±5.1	0.45 ±0.4
Equation 3.8	87.8±0.5	5.1 ±2.1	7.7 ±4.0	1.3±0.6

Table 3.2: Performance of high load indication with building location as the sensitive attribute.

	Acc(%)	$\Delta_{PE}(e^{-2})$	$\Delta_{EO}(e^{-2})$	$\Delta_{AP}(e^{-2})$
Equation 3.1	87.04±0.5	6.03±1.2	8.6±3.2	3.5±1.4
Equation 3.5	87.12±1.5	5.79±1.9	8.4±2.5	2.2±1.3
Equation 3.8	86.76±1.1	5.3±1.5	5.5±3.7	2.3±1.0

3.4.3 User Type Identification

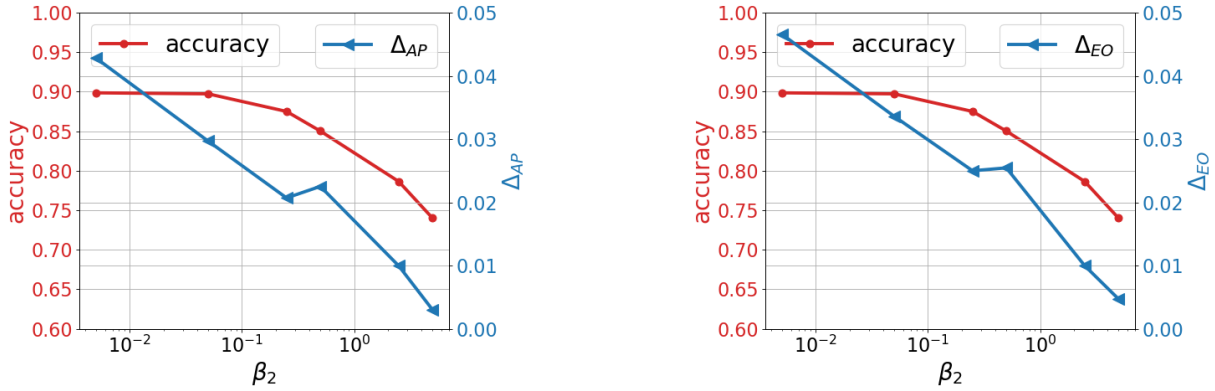
The user type identification task is using the loading pattern of the buildings' previous T hours to determine the type of the building. Correct user type identification can assist the smart power system for multiple tasks, such as the power management, demand prediction[34]. The neural network contains 4 fully connected layers with ReLU activation function in hidden layers and Softmax in the last output layer.

In Table 3.3, z denotes which state the building is in: $z = 1$ means that the building is in New York, $z = 0$ means that the building is in California. In this task, $\mathbf{x}$ is the load information of a building in the last 8 hours and $y \in \{0, \ldots, 6\}$ represents 7 types of buildings.

The vanilla logistic classifier in Equation 3.9 results more unfairness compared with the regularized model. In Table 3.3, the regularizers ensure the fairness during the training without performance loss. The AP regularizer in Equation 3.5 even hugely improves the overall performance.

Table 3.3: Performance of building classification with building location as the sensitive attribute.

	Acc(%)	$\Delta_{PE}(e^{-3})$	$\Delta_{EO}(e^{-2})$	$\Delta_{AP}(e^{-2})$
Equation 3.9	88.97±3.4	9.4 ±3.1	4.55±1.9	4.18±1.3
Equation 3.5	92.9±2.4	3.1±0.6	1.22±0.5	0.95±0.51
Equation 3.8	89.72±3.4	7.0±3.0	3.36±1.8	2.97±1.3



(a) Accuracy and Δ_{AP} vs β_2.

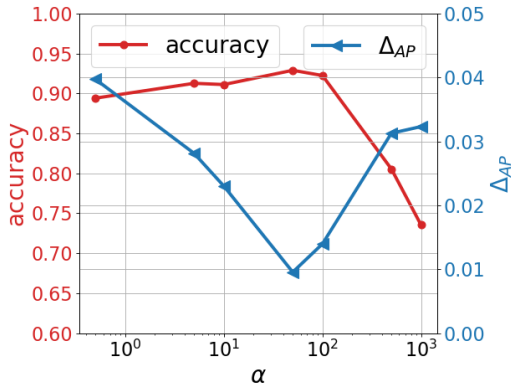
(b) Accuracy and Δ_{EO} vs β_2.

Figure 3.1: Impact of equal opportunity and predictive equality regularizers. The weight of $\mathcal{R}_{\text{EO}}$ is represented as β_2, and $\beta_1 = 0.8\beta_2$.

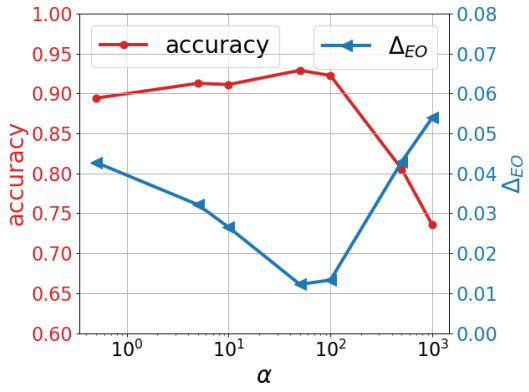
3.4.4 Impact of Fairness Regularizers

Though fine-tuning the weights of two regularizers could let the model ensure the fairness while having little impacts on final performance, increasing the weights which strengthens the fair constraints will probably hurt the overall performance. In this section, we will tune the weights of regularizers and study the effect of fairness regularizers on the accuracy performance.

The impact of EO regularizer shows as Figure 3.1a and Figure 3.1b. While the weight of EO regularizer increases, the fairness measurements(Δ_{AP} and Δ_{EO}) decrease and the accuracy also decreases. The impact of AP regularizer shows as Figure 3.2a and Figure 3.2b. The fairness measures firstly decreases with increasing weight while the accuracy performance is not hurt at all. However, when the weight is larger than 50, the accuracy decreases and



(a) Accuracy and Δ_{AP} vs α.



(b) Accuracy and Δ_{EO} vs α.

Figure 3.2: Impact of accuracy parity regularizer. The weight of $\mathcal{R}_{\text{AP}}$ is represented as α.

two fairness measurements both increase. The AP regularizer has less impact than EO regularizer when the accuracy is high due to the following possible reason: the penalty of AP regularizer of a subgroup only considers the cases that are mis-classifed; based on the expression of Equation 3.7, the EO regularizer takes all cases into consideration no matter the prediction is correct or not. When the model has very high accuracy, the penalty from AP regularizer is small which makes its weight less sensitive than EO regularizer.

3.5 Summary

In order to achieve more fair results in power systems, existing machine learning schemes need to introduce additional bias-mitigating schemes before they can be readily applied to the grids. To this end, the present work first examined and showcased the existing bias in directly applying Neural Network models, then two types of regularizers were introduced to promote fairness in user classification tasks. The proposed regularizers could indeed improve the fairness without significant performance loss. The proposed regularizers could be readily used in other tasks not limited to user classification, e.g., abnormality detection, theft detection.

Chapter 4

Long-term Fairness for Real-time Decision Making

In this chapter, we investigate the problem of energy management in power grids, which is formulated as a constrained online optimization problem where the overall utility is optimized in the presence of the long-term fairness constraint. We focus on a special case in power grids' peer-to-peer (P2P) electricity market [25, 22, 18]. We develop an online algorithm that solves the problem 'on the fly'. While we use the energy management problem in P2P electricity market as a case study, our method can be extended to other decision-making scenarios with long-term fairness concerns, such as the resource allocation problem in network systems [30].

The chapter is organized as follows: Section 4.2 introduces the problem formulation and performance metrics of the decision-making system. Section 4.3 proposes the online algorithm that is proven to achieve sub-linear dynamic regret and sub-linear accumulated unfairness under convexity assumptions. Section 4.5 introduces the energy management problem in P2P electricity market and the numerical tests are presented in Section 4.7 to evaluate the performance of the proposed method.

4.1 Introduction

As machine learning (ML) techniques are increasingly incorporated within power grids there is a growing need to understand the fair behaviors of the ML-based decision-making systems. In applications that require real-time decision-making, different fairness objectives naturally arise, including the instantaneous fairness objective that aims to ensure fairness at every time slot, and the long-term fairness objective that aims to ensure fairness over a time horizon. At present, most approaches focused on the fairness implications of decisions made instantaneously in which long-term effects and system level dynamics are not considered. However, in practice, decision-making systems are usually operating in a dynamic manner such that the decision-maker makes decisions over a period of time. The difficulty of achieving long-term fairness is understandable: Long-term dynamics are hard to assess and they do not align with the traditional supervised ML framework that uses fixed data sets [4]. There is a growing need to understand long-term fairness due to its flexibility and applicability to real-world systems [31, 22, 18, 38].

Some recent works study the long-term impact of fairness in various applications, such as [4, 10, 36, 44, 39]. Most fairness definitions are studied in the ecosystems and focus on classification settings. However, in many real cyber-physical systems, the decisions are strictly constrained by physical limitations and most of them cannot be represented as a classification task, such as [30, 6, 7]. For achieving long-term fairness, most approaches focus on the problems with dynamic costs but time-invariant fairness constraints but do not consider the problem with time-varying fairness constraints. However, the fairness constraint might vary with the dynamics of the systems and decisions [44]. The present work aims to introduce a novel framework to ensure long-term fairness in dynamic decision-making systems with time-varying fairness constraints.

4.2 Problem Formulation

Machine learning has been increasingly incorporated within decision-making systems, especially in ecosystems and cyber-physical systems. Such increasing popularity also leads to a growing need arises to understand the trustworthiness of ML-based decision-making systems [29]. Most approaches to understanding and improving fairness focus on instantaneous settings, which do not consider long-term concepts. However, in applications that require real-time decision-making, a more general long-term fairness criterion needs to be developed and taken into consideration.

Specifically, at each time slot t, a decision maker makes an decision $\mathbf{x}_t$ from a set $\mathcal{X}_t \subseteq \mathcal{X} \in \mathbb{R}^n$, and observes a loss function $f_t(\cdot)$: $\mathbb{R}^n \to \mathbb{R}$. In addition, there are two groups $\mathcal{D}_0$ and $\mathcal{D}_1$ determined by the sensitive attribute $z \in \{0, 1\}$ of individuals in each group. The decision maker incurs a loss $f_t(\mathbf{x}_t)$. The set $\mathcal{X}$ is assumed to be known and fixed. Fairness is ensuring unbiased and equitable outcomes for different groups, $\mathcal{D}_0$ and $\mathcal{D}_1$. Specifically, it can be formulated as a time-varying constraint made by a time-varying function $g_t(\cdot) : \mathbb{R}^n \to \mathbb{R}$, which is driven by the unknown dynamics in various applications and computed from specific fairness metrics. For instance,

$$g_t^{\kappa}(\mathbf{x}_t) \coloneqq \left(\Omega(\mathcal{D}_0, \mathbf{x}_t) - \Omega(\mathcal{D}_1, \mathbf{x}_t)\right)^2 - \kappa, \tag{4.1}$$

where $\Omega(\mathcal{D}_0, \mathbf{x}_t)$ denotes the performance metric computed on group $\mathcal{D}_0$ and the decision $\mathbf{x}_t$. And $g_t^{\kappa}(\mathbf{x}_t)$ can be used to define an instantaneous time-varying constraint $g_t^{\kappa}(\mathbf{x}_t) \leq 0$ which bounds the disparity induced by the decision $\mathbf{x}_t$ between two groups by $\kappa > 0$. For example, in ecosystems such as bank loans and college admissions, the decision makers aim to decide whether approve or reject applications from a stream of individuals [4]. In this case, $\mathbf{x}_t$ denotes the set of binary decisions for individuals considered at t. The decision maker minimizes loss function $f_t(\mathbf{x}_t)$ subjects to instantaneous demographic parity(DP) constraints

at every time slot [5], [39]. Specifically, at each time slot t, decisions should be made such that individuals from different sensitive groups have similar probabilities of being approved, i.e., $g_t^\kappa(\mathbf{x}_t) \coloneqq (P(q_t(d \in \mathcal{D}_0, \mathbf{x}_t) = 1) - P(q_t(d \in \mathcal{D}_1, \mathbf{x}_t) = 1))^2 - \kappa \leq 0$, where $q_t(d \in \mathcal{D}_0, \mathbf{x}_t)$ denotes the predicted outcome of individual belonging to sensitive group $\mathcal{D}_0$ based on current decision $\mathbf{x}_t$, $\Omega(\mathcal{D}_i, \mathbf{x}_t) = P(q_t(d \in \mathcal{D}_i, \mathbf{x}_t) = 1)$ denotes the probability of being approved for individuals $\mathcal{D}_i$ and $\kappa \in [0, 1]$. The sensitive attributes can denote the properties of individuals such as male/female and white/non-white.

In the instantaneous fair setting, the decision maker minimizes cost function subjects to instantaneous fairness constraints at every time slot.

$$\min_{\mathbf{x}_t \in \mathcal{X}_t, \forall t \in \{1, \cdots, T\}} \sum_{t=1}^{T} f_t(\mathbf{x}_t) \qquad \text{s.t. } [\mathbf{x}_t]_{t=1}^T \in \mathcal{G}_T^I, \tag{4.2}$$

where T is the time horizon, $[\mathbf{x}_t]_{t=1}^T = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_T]$ denotes the decision trajectory which contains a sequence of instantaneous decisions over the time horizon T, and $\mathcal{G}_T^I$ denotes the instantaneous fairness constraint defined as

$$\mathcal{G}_T^I = \left\{ [\mathbf{x}_t]_{t=1}^T \middle| g_t^\kappa(\mathbf{x}_t) \leq 0, \forall t \in \{1, \cdots, T\} \right\}. \tag{4.3}$$

To consider long-term fairness, one intuitive attempt would be aggregating the instantaneous fairness violation as the constraint, $\sum_{t=1}^T g_t^\kappa(\mathbf{x}_t) \leq 0$. However, such direct aggregation cannot ensure fairness in the long term. In the example of college admission, where one tries to equalize DP measure between two groups, the constraint $\sum_{t=1}^T g_t^\kappa(\mathbf{x}_t) \leq 0$ with $g_t^\kappa(\mathbf{x}_t) \coloneqq (P(q_t(d \in \mathcal{D}_0, \mathbf{x}_t) = 1) - P(q_t(d \in \mathcal{D}_1, \mathbf{x}_t) = 1))^2 - \kappa$ cannot ensure that the average DP goes to zero over time: $\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^T [P(q_t(d \in \mathcal{D}_0, \mathbf{x}_t) = 1) - P(q_t(d \in \mathcal{D}_1, \mathbf{x}_t) = 1)] = 0$. To address this problem, we redesign the constraint set from a long-term perspective as

$$\mathcal{G}_T^L = \left\{ [\mathbf{x}_t]_{t=1}^T \middle| \sum_{t=1}^T g_t(\mathbf{x}_t) = 0 \right\}, \tag{4.4}$$

where

$$g_t(\mathbf{x}_t) \coloneqq \Omega(\mathcal{D}_0, \mathbf{x}_t) - \Omega(\mathcal{D}_1, \mathbf{x}_t).$$

For the college admission example, $g_t(\mathbf{x}_t) = P(q_t(d \in \mathcal{D}_0, \mathbf{x}_t) = 1) - P(q_t(d \in \mathcal{D}_1, \mathbf{x}_t) = 1)$.

Note that enforcing $\sum_{t=1}^T g_t(\mathbf{x}_t) = 0$ is different from restricting $\mathbf{x}_t$ in every time slot to satisfy $g_t(\mathbf{x}_t) = 0$. The long-term fairness constraint allows the online decision maker to adapt online decisions to the environment dynamics, in the sense that it is tolerable to instantaneous violations, e.g., $g_t(\mathbf{x}_t) \leq 0$ or $g_t(\mathbf{x}_t) \geq 0$. This implies that the instantaneous fairness violations can be compensated by future decisions.

In the long-term setting, the goal is to find a set of consequent decisions that minimizes the aggregate loss and ensures the fairness constraints are satisfied in the long term, shown as

$$\min_{\mathbf{x}_t \in \mathcal{X}_t, \forall t \in \{1, \cdots, T\}} \sum_{t=1}^T f_t(\mathbf{x}_t) \qquad \text{s.t. } [\mathbf{x}_t]_{t=1}^T \in \mathcal{G}_T^L. \tag{4.5}$$

However, solving the problem in (4.5) offline requires the information over the entire time horizon which is not practical in real scenarios. And at time slot t, previous decisions $[\mathbf{x}_0, \cdots, \mathbf{x}_{t-1}]$ cannot be changed after receiving latest information, $f_t(\mathbf{x}_t)$. Instead, the problem needs to be solved online: At each time slot t, the decision $\mathbf{x}_t$ needs to be made only according to current and inherited information.

4.3 Method

The present section proposes an online algorithm that outputs the decision $\mathbf{x}_t$ only requires the previous time slot information.

Note that long-term fairness constraint $[\mathbf{x}_t]_{t=1}^T \in \mathcal{G}_T^L$ (4.4) can be equivalently written as two inequality constraints and the problem (4.5) is equivalent to

$$\min_{\mathbf{x}_t \in \mathcal{X}_t, \forall t} \sum_{t=1}^{T} f_t(\mathbf{x}_t) \qquad \text{s.t.} \quad \sum_{t=1}^{T} g_t(\mathbf{x}_t) \leq 0, \quad \sum_{t=1}^{T} -g_t(\mathbf{x}_t) \leq 0. \tag{4.6}$$

Hence, the instantaneous problem at time t can be written as

$$\min_{\mathbf{x}_t \in \mathcal{X}_t,} f_t(\mathbf{x}_t) \qquad \text{s.t.} \quad g_t(\mathbf{x}_t) \leq 0, \quad -g_t(\mathbf{x}_t) \leq 0. \tag{4.7}$$

Let $\lambda_{1,t}, \lambda_{2,t} \geq 0$ denote the Lagrange multipliers associated with the time-varying constraints. The online Lagrangian of Equation 4.7 can be written as

$$\mathcal{L}_t(\mathbf{x}_t, \lambda_{1,t}, \lambda_{2,t}) = f_t(\mathbf{x}_t) + \lambda_{1,t} g_t(\mathbf{x}_t) + \lambda_{2,t}(-g_t(\mathbf{x}_t)). \tag{4.8}$$

Given the previous primal iterate $\mathbf{x}_{t-1}$ and current dual iterates $\lambda_{1,t}, \lambda_{2,t}$ at each time slot t, instead of taking the exact gradient step, the current decision $\mathbf{x}_t$ can be obtained by solving the following optimization problem:

$$\begin{aligned} \mathbf{x}_t &= \min_{\mathbf{x} \in \mathcal{X}_t} \nabla f_{t-1}(\mathbf{x}_{t-1})^\top (\mathbf{x} - \mathbf{x}_{t-1}) + \lambda_{1,t} g_{t-1}(\mathbf{x}) + \lambda_{2,t}(-g_{t-1}(\mathbf{x})) + \frac{||\mathbf{x} - \mathbf{x}_{t-1}||^2}{2\alpha} \\ &= \min_{\mathbf{x} \in \mathcal{X}_t} \nabla f_{t-1}(\mathbf{x}_{t-1})^\top (\mathbf{x} - \mathbf{x}_{t-1}) + \boldsymbol{\lambda}_t^\top \bar{\mathbf{g}}_{t-1}(\mathbf{x}) + \frac{||\mathbf{x} - \mathbf{x}_{t-1}||^2}{2\alpha}, \end{aligned} \tag{4.9}$$

where α is a positive step size, $\boldsymbol{\lambda}_t \coloneqq [\lambda_{1,t}, \lambda_{2,t}]^\top$, $\bar{\mathbf{g}}_t(\mathbf{x}) \coloneqq [g_t(\mathbf{x}), -g_t(\mathbf{x})]^\top$, and $\nabla f_{t-1}(\mathbf{x}_{t-1})$ is the gradient of primal objective $f_{t-1}(\mathbf{x})$ at $\mathbf{x}_{t-1}$. Equation 4.9 tries to minimize the fair

constraint violation by taking a modified descent step, and $\frac{||\mathbf{x}-\mathbf{x}_{t-1}||^2}{2\alpha}$ is a proximal term. After the decision $\mathbf{x}_t$ is made, $f_t(\mathbf{x}_t)$ and $g_t(\mathbf{x}_t)$ are revealed. Then the dual ascent step updates as:

$$\lambda_{1,t+1} = \max(0, \lambda_{1,t} + \mu\nabla_{\lambda_{1,t}}\mathcal{L}_t(\mathbf{x}_t, \lambda_{1,t}, \lambda_{2,t})) = \max(0, \lambda_{1,t} + \mu g_t(\mathbf{x}_t)), \tag{4.10}$$

$$\lambda_{2,t+1} = \max(0, \lambda_{1,t} + \mu\nabla_{\lambda_{2,t}}\mathcal{L}_t(\mathbf{x}_t, \lambda_{1,t}, \lambda_{2,t})) = \max(0, \lambda_{2,t} - \mu g_t(\mathbf{x}_t)), \tag{4.11}$$

where μ is a positive step size, $\lambda_{1,t}, \lambda_{2,t} \geq 0$. And $\nabla_{\lambda_{1,t}}\mathcal{L}_t(\mathbf{x}_t, \lambda_{1,t}, \lambda_{2,t})) = g_t(\mathbf{x}_t)$ and $\nabla_{\lambda_{2,t}}\mathcal{L}_t(\mathbf{x}_t, \lambda_{1,t}, \lambda_{2,t}) = -g_t(\mathbf{x}_t)$ are the gradients of (4.8) with respect to $\lambda_{1,t}$ and $\lambda_{2,t}$.

Note that updating $\boldsymbol{\lambda}_t$ and making decision $\mathbf{x}_t$ only require the previous time slot's information, $f_{t-1}(\mathbf{x}_{t-1})$ and $g_{t-1}(\mathbf{x}_{t-1})$. The overall algorithm is summarized as Algorithm 1.

Algorithm 1

for $t = 1, 2, \ldots$ **do**
  $\boldsymbol{\lambda}_t = [\lambda_{1,t}, \lambda_{2,t}]^\top$, $\bar{\mathbf{g}}_t(\mathbf{x}) = [g_t(\mathbf{x}), -g_t(\mathbf{x})]^\top$
  observe constraint $g_{t-1}(\mathbf{x}_{t-1})$
  $\mathbf{x}_t = \min_{\mathbf{x}\in\mathcal{X}_t} \nabla f_{t-1}(\mathbf{x}_{t-1})^\top(\mathbf{x} - \mathbf{x}_{t-1}) + \boldsymbol{\lambda}_t^\top \bar{\mathbf{g}}_{t-1}(\mathbf{x}) + \frac{||\mathbf{x}-\mathbf{x}_{t-1}||^2}{2\alpha}$
  observe current cost $f_t(\mathbf{x}_t)$ and constraint $g_t(\mathbf{x}_t)$
  $\lambda_{1,t+1} = \max(0, \lambda_{1,t} + \mu g_t(\mathbf{x}_t))$
  $\lambda_{2,t+1} = \max(0, \lambda_{2,t} - \mu g_t(\mathbf{x}_t))$
end for

4.4 Performance Analysis

In this section, two metrics will be introduced to measure the performance and fair violations of online decisions in the dynamic setup: dynamic regret and dynamic fairness [20], [3]. Then we will show that the dynamic regret and dynamic fairness are both sub-linearly increasing.

4.4.1 Performance Metrics

Firstly, we introduce dynamic regret which measures the performance:

$$\mathcal{R}_T \coloneqq \sum_{t=1}^{T} f_t(\mathbf{x}_t) - \sum_{t=1}^{T} f_t(\mathbf{x}_t^*), \tag{4.12}$$

where $[\mathbf{x}_t^*]_{t=1}^T = [\mathbf{x}_1^*, \cdots, \mathbf{x}_t^*, \cdots, \mathbf{x}_T^*]$ are obtained as:

$$\mathbf{x}_t^* = \arg\min_{\mathbf{x}_t \in \mathcal{X}_t} f_t(\mathbf{x}_t) \quad \text{s.t. } g_t(\mathbf{x}_t) = 0. \tag{4.13}$$

To evaluate the feasibility of online decisions, dynamic fairness is introduced to measure the accumulated fairness violation of constraints, which is defined as:

$$\mathcal{F}_T \coloneqq \sum_{t=1}^{T} g_t(\mathbf{x}_t). \tag{4.14}$$

The definition in Equation 4.14 assumes that the instantaneous fair violation(e.g. $g_t(\mathbf{x}_t) \leq 0$) can be compensated by future decisions with opposite violations(e.g. $g_{t+1}(\mathbf{x}_t) \geq 0$). In the cases of cyber-physical systems, such as energy management in power grids and resource allocation in networks, since there usually exist other types of constraints, such as transmission limitations for power grids [12], introducing long-term constraint offers the flexibility to avoid significant increases in $f_t(\mathbf{x}_t)$ compared with meeting the instantaneous constraint at each time slot.

An ideal algorithm should achieve both sub-linear dynamic regret and sub-linear dynamic fairness on the long-term average, i.e., $\lim_{T\to\infty} \frac{\mathcal{R}_T}{T} = 0$ and $\lim_{T\to\infty} \frac{\mathcal{F}_T}{T} = 0$.

4.4.2 Performance Bounds

We assume that the following conditions are satisfied:

Assumption 1: For every t, the cost function $f_t(\mathbf{x})$ and the time-varying fairness constraint $g_t(\mathbf{x})$ are convex.

Assumption 2: For every t, $f_t(\mathbf{x})$ has bounded gradient on $\mathcal{X}$, $||\nabla f_t(\mathbf{x})|| \leq G, \forall \mathbf{x} \in \mathcal{X}$; and $g_t(\mathbf{x})$ is bounded, $||\bar{\mathbf{g}}_t(\mathbf{x})|| \leq \sqrt{2}M$.

Assumption 3: The radius of the convex feasible set $\mathcal{X}$ is bounded, $||\mathbf{x}-\mathbf{y}|| \leq R, \forall \mathbf{x}, \mathbf{y} \in \mathcal{X}$. And the decision $\mathbf{x}_t$ is selected from a convex set $\mathcal{X}_t \subseteq \mathcal{X} \in \mathbb{R}^n$.

Assumption 4: There exists a constant $\epsilon > 0$, and two interior points $\tilde{\mathbf{x}}_t$ and $\tilde{\tilde{\mathbf{x}}}_t$ such that $g_t(\tilde{\mathbf{x}}_t) \leq -\epsilon$ and $g_t(\tilde{\tilde{\mathbf{x}}}_t) \geq \epsilon, \forall t$.

Assumption 5: The constant ϵ in Assumption 4 satisfies $\epsilon > \bar{V}(g)$, which is the point-wise maximum variation of the time-varying constraints and is defined as

$$\bar{V}(g) = \max_{t \in T} \max_{\mathbf{x} \in \mathcal{X}} |g_{t+1}(\mathbf{x}) - g_t(\mathbf{x})|.$$

Based on assumptions 1 and 3, problem (4.6) is an online convex optimization (OCO) problem. Firstly, the following theorem states the upper bound and lower bound of dynamic fairness for Algorithm 1.

Theorem 1: Under Assumptions 1-5, using Algorithm 1 with dual variable initialization as $\boldsymbol{\lambda}_1 = \mathbf{0}$, the $\lambda_{1,t}$ and $\lambda_{2,t}$ are bounded by

$$\lambda_{1,t}, \lambda_{2,t} \leq \bar{\lambda} \coloneqq 2\mu M + \frac{2GR + \frac{R^2}{2\alpha} + \mu M^2}{\epsilon - \bar{V}(g)}, \tag{4.15}$$

and the dynamic fairness of Equation 4.14 is bounded by

$$-2M - \frac{\frac{2GR}{\mu} + \frac{R^2}{2\mu\alpha} + M^2}{\epsilon - \bar{V}(g)} \leq \mathcal{F}_T \leq 2M + \frac{\frac{2GR}{\mu} + \frac{R^2}{2\mu\alpha} + M^2}{\epsilon - \bar{V}(g)}. \tag{4.16}$$

Proof of Theorem 1: See Appendix A.3.

Under the conditions on the time-varying constraint, Theorem 1 shows that $\mathcal{F}_T$ depends on primal and dual stepsizes and $|\mathcal{F}_T|$ is in the order of $\mathcal{O}(\frac{1}{\mu})$ with fixed primal stepsize α. Thus, a larger dual stepsize is beneficial for the long-term fairness constraint.

We further bound the dynamic regret of Equation 4.12 in the next theorem.

Theorem 2: Under Assumptions 1-5, using Algorithm 1 with dual variable initialization as $\boldsymbol{\lambda}_1 = \mathbf{0}$, the dynamic regret of Equation 4.12 is upper bounded by

$$\begin{aligned}\mathcal{R}_T =& \sum_{t=1}^{T} f_t(\mathbf{x}_t) - \sum_{t=1}^{T} f_t(\mathbf{x}_t^*) \\ &\leq \frac{R}{\alpha} V(\{\mathbf{x}_t^*\}_{t=1}^T) + \frac{R^2}{2\alpha} + |\bar{\lambda}| V(\{\bar{\mathbf{g}}_t\}_{t=1}^T) + \mu M^2 T + \frac{\alpha G^2 T}{2} + \frac{\mu M^2}{2}, \end{aligned} \tag{4.17}$$

where $V(\{\mathbf{x}_t^*\}_{t=1}^T)$ and $V(\{\bar{\mathbf{g}}_t\}_{t=1}^T)$ denote the accumulated variations of the minimizers $\mathbf{x}_t^*$ and the fairness constraints at every time slot, which are defined as follows:

$$V(\{\mathbf{x}_t^*\}_{t=1}^T) \coloneqq \sum_{t=1}^{T} V(\mathbf{x}_t^*) = \sum_{t=1}^{T} ||\mathbf{x}_t^* - \mathbf{x}_{t-1}^*||, \tag{4.18}$$

$$V(\{\bar{\mathbf{g}}_t\}_{t=1}^T) \coloneqq \sum_{t=1}^{T} V(\bar{\mathbf{g}}_t) = \sum_{t=1}^{T} \max_{\mathbf{x}} ||\bar{\mathbf{g}}_{t+1}(\mathbf{x}) - \bar{\mathbf{g}}_t(\mathbf{x})||. \tag{4.19}$$

Proof of Theorem 2: See Appendix A.4.

It can be observed that the dynamic regret $\mathcal{R}_T$ is bounded by the primal stepsize, dual stepsize and accumulated variations of instantaneous minimizers $V(\{\mathbf{x}_t^*\}_{t=1}^T)$ and time-varying constraints $V(\{\bar{\mathbf{g}}_t\}_{t=1}^T)$. The upper bound could be small while choosing appropriate primal and dual stepsizes.

Under Assumptions 1-5 and without knowledge of variations($V(\{\mathbf{x}_t^*\}_{t=1}^T)$ and $V(\{\bar{\mathbf{g}}_t\}_{t=1}^T)$), according to Theorems 1-2, if the primal and dual stepsizes are chosen as $\alpha = \mu = \mathcal{O}(T^{-\frac{1}{3}})$,

then the dynamic fairness and regret can be bounded as

$$|\mathcal{F}_T| = \mathcal{O}(T^{\frac{2}{3}}),$$
$$\mathcal{R}_T = \mathcal{O}\left(\max\left\{V(\{\mathbf{x}_t^*\}_{t=1}^T)T^{\frac{1}{3}}, V(\{\bar{\mathbf{g}}_t\}_{t=1}^T)T^{\frac{1}{3}}, T^{\frac{2}{3}}\right\}\right).$$

If the variations, $V(\{\mathbf{x}_t^*\}_{t=1}^T)$ and $V(\{\bar{\mathbf{g}}_t\}_{t=1}^T)$, are known, the primal and dual stepsizes can be chosen as

$$\alpha = \mu = \sqrt{\frac{\max\left\{V(\{\mathbf{x}_t^*\}_{t=1}^T), V(\{\bar{\mathbf{g}}_t\}_{t=1}^T)\right\}}{T}},$$

then the regret can be bounded by

$$\mathcal{R}_T = \mathcal{O}\left(\max\left\{\sqrt{V(\{\mathbf{x}_t^*\}_{t=1}^T)T}, \sqrt{V(\{\bar{\mathbf{g}}_t\}_{t=1}^T)T}\right\}\right),$$

and the dynamic fairness is bounded as

$$\mathcal{F}_T = \mathcal{O}\left(\max\left\{\frac{T}{V(\{\mathbf{x}_t^*\}_{t=1}^T)}, \frac{T}{V(\{\bar{\mathbf{g}}_t\}_{t=1}^T)}\right\}\right).$$

Proof: See Appendix A.4.1.

Although the dynamic benchmark $\mathbf{x}_t^*$ for regret in (4.12) is got through Equation 4.13, $[\mathbf{x}_t^*]_{t=1}^T$ is not the optimal solution to problem (4.5). Instead, the optimal solutions can be obtained as the offline optimal solutions of Equation 4.5, denoted as $[\mathbf{x}_t^{\text{off}}]_{t=1}^T$, which requires the information over the entire time horizon. In the implementations, especially the **time-varying** setting, solving Equation 4.13 frequently fails, so we use $\mathbf{x}_t^{\text{off}}$ as the dynamic benchmark for some experiments and the offline dynamic regret is computed as

$$\mathcal{R}_T^{\text{off}} = \sum_{t=1}^T f_t(\mathbf{x}_t) - \sum_{t=1}^T f_t(\mathbf{x}_t^{\text{off}}),$$

where the superscript denotes the dynamic benchmark is $\mathbf{x}_t^{\text{off}}$. Since $\mathcal{R}_T^{\text{off}} = \mathcal{R}_T + \sum_{t=1}^T f_t(\mathbf{x}_t^*) - \sum_{t=1}^T f_t(\mathbf{x}_t^{\text{off}}) \geq \mathcal{R}_T$, the upperbound of $\mathcal{R}_T^{\text{off}}$ is larger than the upperbound shown in Equation 4.17. We will further discuss the upperbound of the offline dynamic regret in Appendix A.5.

In this section, we present that dynamic regret and cumulative fair violation increase sublinearly under Assumptions 1-5 while choosing suitable primal and dual stepsizes.

4.5 Application: Peer-to-peer Electricity Market

Algorithm 1 can be applied in various applications which require making decisions on the fly while taking into account long-term fairness. For example, resource allocation in networks and energy management in power grids. In this section, we focus on a special case in power grids' peer-to-peer (P2P) electricity market [25],[31],[18]. We formulate the energy management problem in the peer-to-peer electricity market as an online optimization problem like (4.5). Besides the objective function and long-term fairness constraints, there are additional constraints inherited from the physical limits of the system.

Specifically, consider the P2P network with N nodes which are responsive to grid conditions such as energy prices, local energy supply and demand [18]. All nodes are divided into 2 groups two $\mathcal{D}_0$ and $\mathcal{D}_1$ based on their sensitive attributes. Per time t, node i, corresponding to client i, has demand d_t^i and supply s_t^i. Each client has surplus or deficit as $h_t^i = s_t^i - d_t^i$. The client i will be considered as a producer at time t if $h_t^i \geq 0$; otherwise client i will be considered as a consumer since $h_t^i \leq 0$. The decision $\mathbf{X}_t$ is an $N \times N$ matrix with the (i,j)th entry representing the trade between node i and j. Specifically, if $\mathbf{X}_t^{ij} > 0$, it denotes client i sells $\mathbf{X}_t^{ij}$ units of energy to client j; otherwise, client i buys energy from client j. Assuming the price of unit energy traded in P2P electricity market $e_t^{ij}, \forall i, j \in N$ is cheaper than the

price of utility company p, consumers will buy electricity from producers in the P2P market first then buy electricity from utility companies.

The goal is to minimize the total cost to satisfy the supply-demand balance for all consumers. Therefore, the decision maker tends to satisfy consumers with lower costs first due to their convenient locations or advanced power transmission system, and the advantageous group would potentially have a higher satisfaction rate in P2P market transactions. To ensure fairness in the long run, each group should have a similar satisfaction rate on average. Hence, the set $\mathcal{G}_T^I$ is formulated as

$$\mathcal{G}_T^I = \left\{ [\mathbf{X}_t]_{t=1}^T \middle| g_t^\kappa(\mathbf{X}_t) = \left(\frac{1}{|\mathcal{D}_0|} \sum_{i \in \mathcal{D}_0 \cap w_t^c} \frac{\Delta_t^{i,-}}{d_t^i} - \frac{1}{|\mathcal{D}_1|} \sum_{i \in \mathcal{D}_1 \cap w_t^c} \frac{\Delta_t^{i,-}}{d_t^i} \right)^2 - \kappa \leq 0 \right\}.$$

The set $\mathcal{G}_T^L$ is formulated as

$$\mathcal{G}_T^L = \left\{ [\mathbf{X}_t]_{t=1}^T \middle| \sum_{t=1}^T g_t(\mathbf{X}_t) = \sum_{t=1}^T \left(\frac{1}{|\mathcal{D}_0|} \sum_{i \in \mathcal{D}_0 \cap w_t^c} \frac{\Delta_t^{i,-}}{d_t^i} - \frac{1}{|\mathcal{D}_1|} \sum_{i \in \mathcal{D}_1 \cap w_t^c} \frac{\Delta_t^{i,-}}{d_t^i} \right) = 0 \right\}.$$

And the cost function $f_t(\mathbf{X}_t)$ is

$$f_t(\mathbf{X}_t) = \frac{1}{|w_t^c|} \left(- \sum_{i \in w_t^c} \left((h_t^i - \sum_{j \in N} \mathbf{X}_t) p + \sum_{j \in N} e_t^{ij} \mathbf{X}_t \right) - \sum_{j \in w_t^c} \sum_{i \in N} \gamma \mathbf{D}^{ji} \mathbf{X}_t^{ji} \right),$$

where p is the unit price of energy trading with the utility company, e^{ij} denotes the trading price between client i and j in the P2P electricity market, $\mathbf{D}$ denotes the time-invariant power transfer distance matrix can be obtained from the grid (Appendix A.6) and γ denotes the unit price of the transmission line utilization fee. The whole turn $-\sum_{j \in w_t^c} \sum_{i \in N} \gamma \mathbf{D}^{ji} \mathbf{X}_t^{ji}$ computed the total transmission line utilization fee. The objective function attempts to minimize the total cost to purchase enough energy to satisfy the supply-demand balance for all consumers. The overall online optimization problem can be viewed as the following:

$$\min_{\mathbf{X}_t \in \mathcal{X}_t, \forall t} \sum_{t=1}^{T} f_t(\mathbf{X}_t) \quad \text{s.t.} \qquad [\mathbf{X}_t]_{t=1}^{T} \in \mathcal{G}_T^L, \tag{4.20}$$

where

$$\mathcal{X}_t := \left\{ \begin{array}{lll} \mathbf{X}_t \in \mathbb{R}^{N \times N} & & (4.20a) \\ 0 \leq \mathbf{X}_t^{ij} \leq h_t^i, & \forall i \in w_t^p, \forall j \in \{1, \cdots, n\} & (4.20b) \\ h_t^i \leq \mathbf{X}_t^{ij} \leq 0, & \forall i \in w_t^c, \forall j \in \{1, \cdots, n\} & (4.20c) \\ 0 \leq \sum_{j \in N_i} \mathbf{X}_t^{ij} \leq h_t^i, & \forall i \in w_t^p & (4.20d) \\ h_t^i \leq \sum_{j \in N_i} \mathbf{X}_t^{ij} \leq 0, & \forall i \in w_t^c & (4.20e) \\ \mathbf{X}_t + \mathbf{X}_t^\top = \mathbf{0} & & (4.20f) \\ -\rho_{\max}^l \leq \rho^l(\mathbf{X}_t) \leq \rho_{\max}^l & \forall l \in \{1, \cdots, N_l\} & (4.20g) \\ g_t^\kappa(\mathbf{X}_t) <= \tau & & (4.20h) \end{array} \right\}.$$

The definition of $\mathcal{X}_t$ is the physical constraints of the P2P electricity market. Constraints (4.20b) and (4.20c) guarantee that producers only sell energy and consumers only buy energy, and (4.20d) and (4.20e) make sure that producers will not sell more than their redundancy and consumers will not buy more than their deficit. Thus, after P2P trading, for consumer $i \in w_t^c$, it has deficit $\Delta_t^{i,-} = s_t^i - d_t^i - \sum_{j \in N} = h_t^i - \sum_{j \in N} \mathbf{X}_t^{ij} \leq 0$. Constraint (4.20f) represents the equal trading volume and offsetting purchase and sell. Constraint (4.20g) denotes the flow constraint which avoids the energy flow exceeding transmission line capacity (details in Appendix A.6). Constraint (4.20h) prevents very unfair outcomes, τ is a larger constant than κ, $\tau > \kappa$.

The numerical experiments on synthetic data and real dataset are presented in Section 4.7.

4.6 Conclusion

In practice, decision-making systems are usually operating in a dynamic manner and potential unfairness is a concern. The sole pursuit of the overall performance may lead to unfair results. In order to achieve fair results in decision making scenarios, such as energy management, there is a growing need to understand long-term fairness due to its flexibility and applicability. In this paper, we investigate the problem of long-term fairness for decision making where the overall utility is optimized in the presence of the long-term fairness constraint. The present paper proposes an online algorithm that is proven to achieve sub-linear dynamic regret and sub-linear dynamic fairness. The experimental results for the energy management case show that the proposed method could indeed guarantee long-term fairness without significant performance loss.

4.7 Experiments

In this section, we test the proposed algorithm on the P2P electricity market application in the 14-bus grid system.

Datasets and Settings: The experiments are carried out on synthetic data and real data from California ISO Open Access Same-time Information System (OASIS). Fourteen nodes are divided into 2 groups by their locations in the power grids. In the synthetic data, we assume that the average trading price within the same group is a little cheaper than the average trading price between the two groups.

1. Random: The supply and demand of each client at every t are randomly generated according to predefined random distributions. The supply and demand of each client at every t are randomly generated according to predefined time-invariant distributions. The supply

comes from two renewable energy sources: Solar power and wind power. More specifically, solar power generation follows a normal distribution and wind power generation follows a uniform distribution.

2. Time-varying: The supply and demand of each client at every t are generated as cosine functions, e.g., $s_t^i := a^i + b^i cos(\pi t) + c_t$ where a^i and b^i are two constants for client i and $c_t \sim \mathcal{N}(0, \sigma^2)$ is a small noise centered at zero. This setting of supply and demand satisfies the assumptions in section 4.4.

3. OASIS: The supply and demand of each client at every t are pre-processed real data from OASIS. OASIS contains hourly total demand and supply. At each time t, based on the Dirichlet distribution, three distributions for demand, wind power supply and solar power supply are randomly generated. Then assign demand and supply to each node according to distributions.

Baseline: In each setting, we compare the results of three different approaches:

•**Long-term:** Solve the problem in Equation 4.5 through Algorithm 1.

•**Instantaneous:** Solve the problem in Equation 4.2 at each time slot:

$$\mathbf{x}_t = \arg\min_{\mathbf{x} \in \mathcal{X}_t} f_t(\mathbf{x}) \quad \text{s.t. } g_t^\kappa(\mathbf{x}) \leq 0.$$

•**Offline:** Solve the problem in Equation 4.5 offline.

Evaluation Metrics

To demonstrate our method, we will present the time-average cost, dynamic regret and

dynamic fairness. We further introduce the averaged absolute fair violation(A^2FV) as

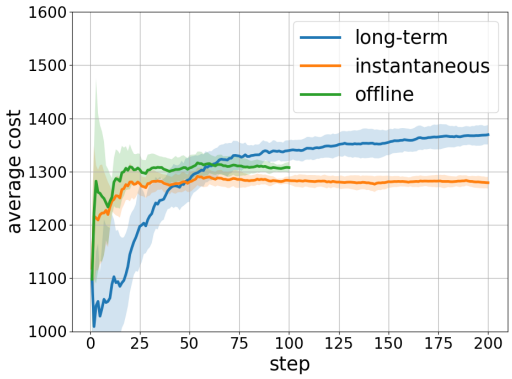
$$A^2FV = \frac{1}{T}\sum_{t=1}^{T}|g_t(\mathbf{x}_t)|.$$

4.7.1 Random Setting

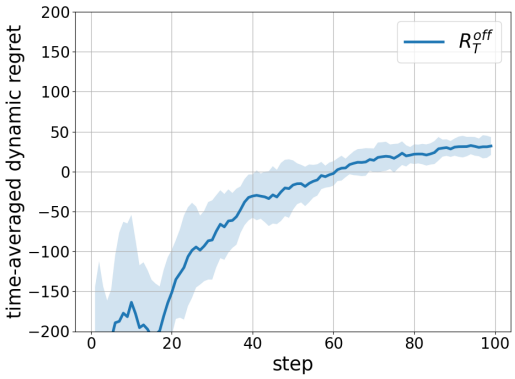
Firstly, the performance of the proposed method is evaluated on randomly generated synthetic data. With time horizon $T = 200$ and constant stepsizes, μ and α in Algorithm 1, the results are shown as Figure 4.1 and Figure 4.2. It shows that our method can achieve both sub-linear regret and dynamic fairness on the long-term average. In this case, the time average cost of our method is slightly higher than the instantaneous solution. Firstly, the instantaneous optimal solution benefits from the $\kappa > 0$ but it leads to the linear increase on $\mathcal{F}_T$ as Figure 4.2a. Secondly, it is worth mentioning that the dynamic regret may not be sub-linear in this case theoretically since there is no guarantee that *Assumption 5* holds. But experimental results still show that our method still can ensure long-term fairness with acceptable performance loss in this case. Figure 4.2a also confirms that aggregating instantaneous fairness constraints cannot ensure the fairness in the long term. Figure 4.2b shows that our method allows violations in each time slot but enforces long-term fairness by adapting the decisions to the environment dynamics.

4.7.2 Time-varying Setting

In this section, we investigate how our method performs on time-varying demand and supply of each client at every t. With time horizon $T = 100$ and constant stepsize, the results are shown as Figure 4.3 and Figure 4.4.

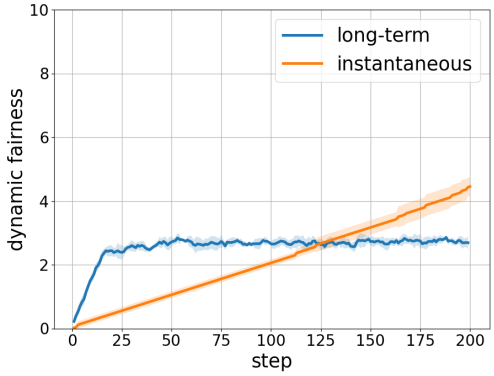


(a) Results of the random setting, time-averaged costs, $\mu = 1e^4, \alpha = 0.5$.

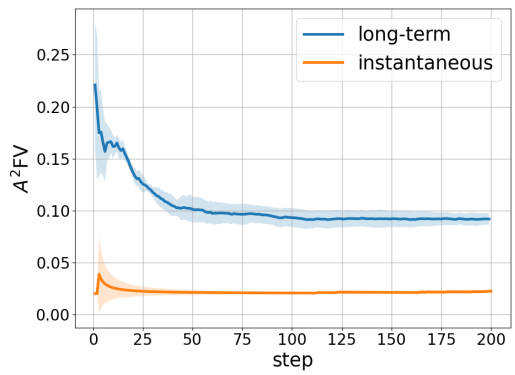


(b) Results of the random setting, dynamic regrets, $\mu = 1e^4, \alpha = 0.5$.

Figure 4.1: Experimental results of random setting, time-averaged costs and dynamic regrets, $\mu = 1e^4, \alpha = 0.5$.

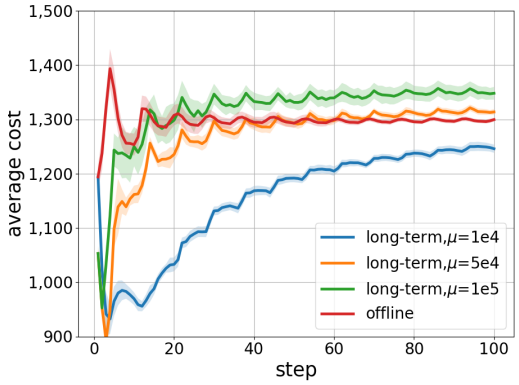


(a) Results of the random setting, dynamic fairness, $\mu = 1e^4, \alpha = 0.5$.

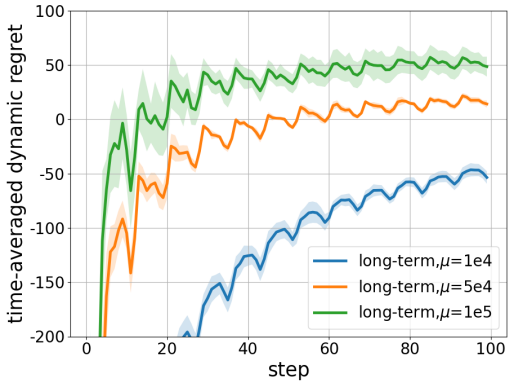


(b) Results of the random setting, dynamic A^2FV, $\mu = 1e^4, \alpha = 0.5$.

Figure 4.2: Experimental results of random setting, dynamic fairness and A^2FV, $\mu = 1e^4, \alpha = 0.5$.

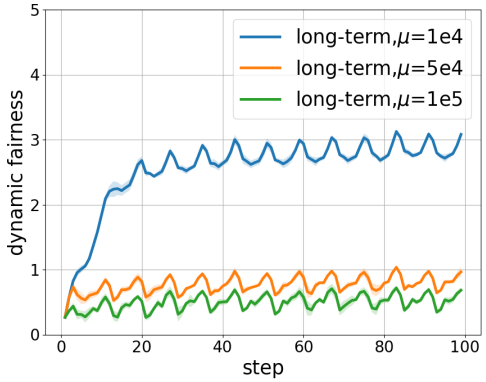


(a) Results of the time-varying setting, time-averaged costs, $\alpha = 0.5$.

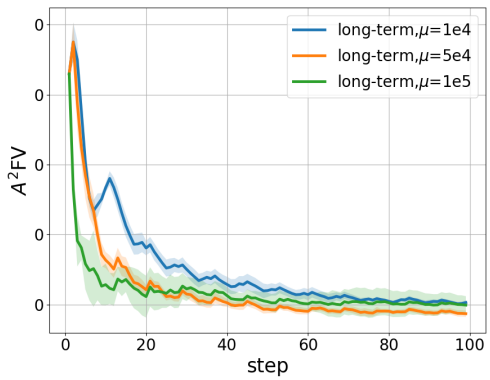


(b) Results of the time-varying setting, time-averaged dynamic regrets, $\alpha = 0.5$.

Figure 4.3: Experimental results of time-varying setting, time-averaged costs and dynamic regrets, $\alpha = 0.5$.



(a) Results of the time-varying setting: dynamic fairness, $\alpha = 0.5$.



(b) Results of the time-varying setting, A^2FV, $\alpha = 0.5$.

Figure 4.4: Experimental results of time-varying setting dynamic fairness and A^2FV, $\alpha = 0.5$.

In this setting, the instantaneous approach that solves the problem in Equation 4.2 fails frequently and usually cannot survive over 15 iterations. While increasing the value of μ, the average cost increases and $|\mathcal{F}_T|$ decreases. It can be concluded that the parameter μ controls the strength enforcing the long-term fairness. If the stepsize is initialized with a small value and $\boldsymbol{\lambda}$ is initialized as $\mathbf{0}$, the cost could be better than the offline decisions in a short-term period due to the fair violations at the first few iterations. For better long-term fairness (smaller upper bound of $|\mathcal{F}_T|$), the algorithm can sacrifice the utility by increasing dual stepsize. The experimental results show that the sub-linear dynamic regret and sub-linear dynamic fairness are guaranteed in the long run with fixed stepsizes.

4.7.3 OASIS

In addition to synthetic data generated under 2 different settings, we also consider real-world data in simulating. We collect one week's demand, supply and price starting on 2023 January 1st at 7:00 am and ending on 2023 January 8th at 7:00 am from 14 nodes on California ISO Open Access Same-time Information System (OASIS). Fourteen nodes are randomly divided into 2 groups. The supply comes from 2 kinds of renewable energy sources (RES): Solar power and wind power [18]. The data contains hourly information about the total demand, supply and unit price of each node buying energy. The total demand is much larger than the supply of RES since the energy market is still dominated by traditional energy, so we only count 15% of it as the demand of the P2P electricity market. With the one-week time horizon and decreasing stepsize, the results are shown as Figure 4.5 and Figure 4.6.

In Figure 4.5, the sharp increase in average cost is due to a huge decrease in wind power supply during that period. It implies that even the long-term fairness constraint, $\sum_{t=1}^{T} g_t(\mathbf{x}_t) = 0$, is still very restricted in some real scenarios even solving it offline. In order to achieve more fair results in dynamic decision-making systems with a lot of uncertainty, there is a

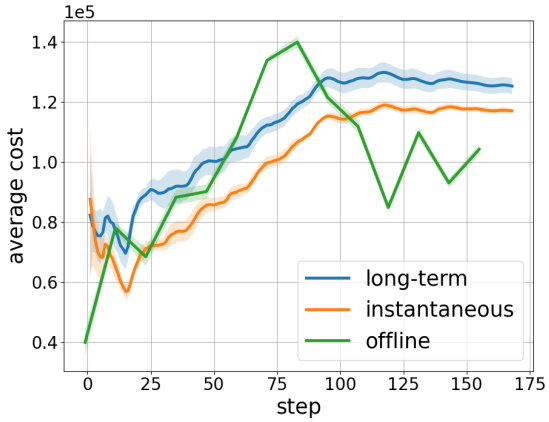
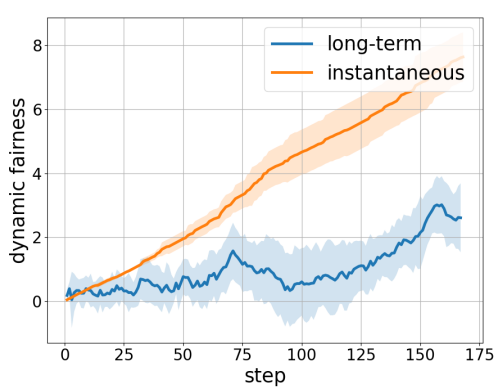
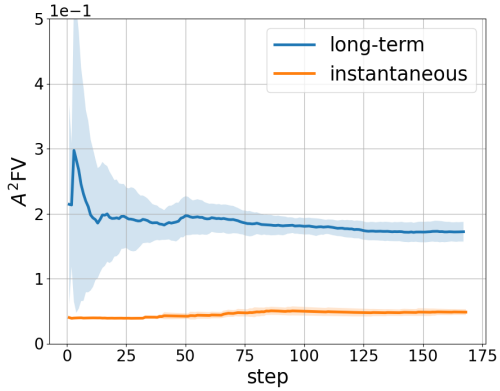


Figure 4.5: Results of OASIS data, time-averaged costs.



(a) Results of OASIS data, dynamic fairness.



(b) Results of OASIS data, A^2FV.

Figure 4.6: Experimental results of OASIS, dynamic fairness and A^2FV.

growing need to understand long-term fairness due to its flexibility and applicability. In the Figure 4.6a, we can observe that $g_t(\mathbf{x}_t)$ oscillates around 0 in some periods because we cannot identify which group of nodes is the advantaged group in this setting. Though the violation of each time slot is much huger than the instantaneous solution with $g_t^\kappa(\mathbf{x}_t)$, our method still can achieve sub-linear increase in dynamic fairness which guarantees the long-term fairness constraint.

4.8 Conclusion

In practice, decision-making systems are usually operating in a dynamic manner and potential unfairness is a concern. The sole pursuit of the overall performance may lead to unfair results. In order to achieve fair results in decision making scenarios, such as energy management, there is a growing need to understand long-term fairness due to its flexibility and applicability. In this paper, we investigate the problem of long-term fairness for decision making where the overall utility is optimized in the presence of the long-term fairness constraint. The present paper proposes an online algorithm that is proven to achieve sub-linear dynamic regret and sub-linear dynamic fairness. The experimental results on the energy management case show that the proposed method could indeed guarantee long-term fairness without significant performance loss.

4.9 Summary

In this chapter, we investigate the problem of energy management in power grids where the overall utility is optimized in the presence of the long-term fairness constraint. In order to achieve fair results in decision making in power systems, such as energy management, there is a growing need to understand long-term fairness due to its flexibility and applicability. We propose an online algorithm that is proven to achieve sub-linear dynamic regret and sub-linear dynamic fairness. The experimental results show that the proposed method could indeed guarantee the long-term fairness constraint without significant performance loss.

Chapter 5

Discussion and Conclusion

Fairness-aware machine learning has emerged as a critical consideration in various domains. Ensuring fairness becomes crucial as machine learning algorithms are utilized for decision-making processes that affect the distribution and management of electricity resources. In the context of power grids, fairness-aware machine learning aims to address possible disparities and biases that arise in the allocation of energy resources, demand response programs, etc.

To summarize, this thesis is a study to explore a suite of fairness-aware machine learning in power grids. By integrating fairness metrics and considerations into the design and training of machine learning models, it becomes possible to promote equitable outcomes and mitigate the risk of discriminatory practices. Power grid datasets may suffer from inherent biases and imbalances, such as historical inequalities in electricity access or biased data collection practices. Fairness-aware machine learning techniques need to address these biases during data preprocessing and model training. Another challenge in fairness-aware machine learning for power grids is defining appropriate fairness metrics that align with the specific context and societal values. The choice of fairness metric depends on various factors, such as different goals of power grid application, the types of available data and various ethical considerations

related to the resource distribution and opportunities.

In the first work, we propose accuracy parity, equal opportunity and predictive equality regularizers to mitigate the unfairness, which can be used for different classification tasks in power grids.

In order to achieve more fair results in dynamic decision making in power systems, there is a growing need to understand long-term fairness due to its flexibility and applicability. We propose an online algorithm that is proven to achieve sub-linear dynamic regret and sub-linear accumulated unfairness under convexity assumptions. We focus on a simple time-varying P2P electricity market are presented to validate the effectiveness of the proposed framework by testing it on synthetic and real data.

In conclusion, fairness-aware machine learning is still at the beginning stage in power grids. By addressing bias and developing algorithms, fairness-aware machine learning can help build more efficient and trustworthy smart power systems.

5.1 Future Work

As the importance of fairness is increasingly recognized in the power system, it is a challenge to integrate the existing schemes in power grids with advanced fairness-aware machine learning techniques. Due to the varieties of applications in power grids, [24, 32, 37, 27, 22], task-specific fairness metrics are necessary. The invented algorithms should account for the unique characteristics and constraints of power grid systems, such as real-time constraints and network constraints.

Combing fairness-aware machine learning with optimization techniques, game theory, or economic models is another important direction worthy of research. Such approaches could

optimize fairness objectives while considering system-level constraints, resource allocation efficiency, and economic incentives.

Furthermore, defining fairness in power grids is an open question. Measuring fairness by metrics may be an oversimplification of the problem. In reality, fairness is not a technical nor statistical concept; satisfying fairness metrics does not necessarily meet equity goals in practice [38].

There is an increasing number of applications in power systems, where data are collected from non-Euclidean domains and represented as graph-structured data with high-dimensional features and interdependence among nodes. The complexity of graph-structured data has brought significant challenges to the existing machine learning frameworks in power systems defined in Euclidean domains. Graph neural network(GNN) is a powerful framework for analyzing and modeling graph-structured data, providing benefits for various applications in power grids involving graph-structured data [16]. However, current GNN frameworks in power grids do not consider fairness issues. Many fairness learning techniques cannot be directly applied to the GNN frameworks due to the complexity of graph-structured data. How to make GNNs contribute to fairness in power grids is a problem worthy of research.

Large-scale deployment of machine learning models relies on how trustworthy the model is. To achieve this goal, fairness is a key component of building trust in systems. Ensuring fairness in algorithms and decision-making processes is crucial for establishing trustworthiness. In the power grids, such a complex large system, many issues deserve further investigation in order to make fair-aware machine learning systems more useful. Fairness-aware machine learning has the potential to revolutionize decision-making processes in power grids, promoting equitable energy distribution, load management, and resource allocation.

Bibliography

[1] Y. Bechavod and K. Ligett. Penalizing unfairness in binary classification. *arXiv preprint arXiv:1707.00044*, 2017.

[2] S. Bird, K. Kenthapadi, E. Kiciman, and M. Mitchell. Fairness-aware machine learning: Practical challenges and lessons learned. In *Proceedings of the ACM International Conference on Web Search and Data Mining*, page 834–835, 2019.

[3] T. Chen, Q. Ling, and G. B. Giannakis. An online convex optimization approach to proactive network resource allocation. *IEEE Transactions on Signal Processing*, 65(24):6350–6364, 2017.

[4] A. D'Amour, H. Srinivasan, J. Atwood, P. Baljekar, D. Sculley, and Y. Halpern. Fairness is not static: deeper understanding of long term fairness via simulation studies. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pages 525–534, 2020.

[5] C. Dwork, M. Hardt, T. Pitassi, O. Reingold, and R. Zemel. Fairness through awareness. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, pages 214–226, 2012.

[6] Z. Guo, W. Wei, L. Chen, Z. Wang, J. P. Catalão, and S. Mei. Optimal energy management of a residential prosumer: A robust data-driven dynamic programming approach. *IEEE Systems Journal*, 16(1):1548–1557, 2020.

[7] Z. Guo, W. Wei, M. Shahidehpour, Z. Wang, and S. Mei. Optimisation methods for dispatch and control of energy storage with renewable integration. *IET Smart Grid*, 5(3):137–160, 2022.

[8] T. Hatzakis, R. Rodrigues, and D. Wright. Smart grids and ethics. *ORBIT Journal*, 2(2):Article–2, 2019.

[9] E. Heylen, M. Ovaere, S. Proost, G. Deconinck, and D. Van Hertem. Fairness and inequality in power system reliability: Summarizing indices. *Electric Power Systems Research*, 168:313–323, 2019.

[10] Y. Hu and L. Zhang. Achieving long-term fairness in sequential decision making. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 9549–9557, 2022.

[11] M. S. Ibrahim, W. Dong, and Q. Yang. Machine learning driven smart electric power systems: Current trends and new perspectives. *Applied Energy*, 272:115237, 2020.

[12] M. Khorasany, Y. Mishra, and G. Ledwich. A decentralized bilateral energy trading system for peer-to-peer electricity markets. *IEEE Transactions on Industrial Electronics*, 67(6):4646–4657, 2020.

[13] A. Kumbhar, P. G. Dhawale, S. Kumbhar, U. Patil, and P. Magdum. A comprehensive review: Machine learning and its application in integrated power system. *Energy Reports*, 7:5467–5474, 2021.

[14] A. E. Labrador Rivas and T. Abrão. Faults in smart grid systems: Monitoring, detection and classification. *Electric Power Systems Research*, 189:106602, 2020.

[15] E. L. Lee, J.-K. Lou, W.-M. Chen, Y.-C. Chen, S.-D. Lin, Y.-S. Chiang, and K.-T. Chen. Fairness-aware loan recommendation for microfinance services. In *Proceedings of the International Conference on Social Computing*, page 1–4, 2014.

[16] W. Liao, B. Bak-Jensen, J. R. Pillai, Y. Wang, and Y. Wang. A review of graph neural networks and their applications in power systems. *Journal of Modern Power Systems and Clean Energy*, 10(2):345–360, 2021.

[17] M. Liu and G. Gross. Role of distribution factors in congestion revenue rights applications. *IEEE Transactions on Power Systems*, 19(2):802–810, 2004.

[18] Y. Liu, L. Wu, and J. Li. Peer-to-peer (p2p) electricity trading in distribution systems of the future. *The Electricity Journal*, 32(4):2–6, 2019. Special Issue on Strategies for a sustainable, reliable and resilient grid.

[19] M. Macedo, J. Galo, L. de Almeida, and A. de C. Lima. Demand side management using artificial neural networks in a smart grid environment. *Renewable and Sustainable Energy Reviews*, 41:128–133, 2015.

[20] A. Mokhtari, S. Shahrampour, A. Jadbabaie, and A. Ribeiro. Online optimization in dynamic environments: Improved regret rates for strongly convex problems. In *2016 IEEE 55th Conference on Decision and Control (CDC)*, pages 7195–7201. IEEE, 2016.

[21] Y. Nesterov. *Introductory Lectures on Convex Optimization: A Basic Course*, volume 87. Springer Science & Business Media, 2013.

[22] E. Oh and S.-Y. Son. Peer-to-peer energy transaction mechanisms considering fairness in smart energy communities. *IEEE Access*, 8, 2020.

[23] Ong, Sean, Clark, and Nathan. Commercial and residential hourly load profiles for all tmy3 locations in the united states, 2014.

[24] P. Palensky and D. Dietrich. Demand side management: Demand response, intelligent energy systems, and smart loads. *IEEE transactions on industrial informatics*, 7(3):381–388, 2011.

[25] C. Park and T. Yong. Comparative review and discussion on p2p electricity trading. *Energy Procedia*, 128:3–9, 2017. International Scientific Conference "Environmental and Climate Technologies", CONECT 2017, 10-12 May 2017, Riga, Latvia.

[26] A. Paudel, L. P. M. I. Sampath, J. Yang, and H. B. Gooi. Peer-to-peer energy trading in smart grid considering power losses and network fees. *IEEE Transactions on Smart Grid*, 11(6):4727–4737, 2020.

[27] F. Rahimi and A. Ipakchi. Demand response as a market resource under the smart grid paradigm. *IEEE Trans. on Smart Grid*, 1(1):82–88, 2010.

[28] A. Reinhardt, P. Baumann, D. Burgstahler, M. Hollick, H. Chonov, M. Werner, and R. Steinmetz. On the accuracy of appliance identification based on distributed load metering data. In *Sustainable Internet and ICT for Sustainability*, pages 1–9, 2012.

[29] L. Robert, G. Bansal, and C. Lutge. ICIS 2019 SIGHCI Workshop Panel Report: Human computer interaction challenges and opportunities for fair, trustworthy and ethical artificial intelligence. 2020.

[30] T. Si Salem, G. Iosifidis, and G. Neglia. Enabling long-term fairness in dynamic resource allocation. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 6(3):1–36, 2022.

[31] T. Sousa, T. Soares, P. Pinson, F. Moret, T. Baroche, and E. Sorin. Peer-to-peer and community-based markets: A comprehensive review. *Renewable and Sustainable Energy Reviews*, 104:367–378, 2019.

[32] M. Tan, Y. Ren, R. Pan, L. Wang, and J. Chen. Fair and efficient electric vehicle charging scheduling optimization considering the maximum individual waiting time and operating cost. *IEEE Transactions on Vehicular Technology*, 72(8):9808–9820, 2023.

[33] K. Tornai, A. Olah, and M. Lőrincz. Forecast based classification for power consumption data. In *International Conference on Intelligent Green Building and Smart Grid*, pages 1–5, 2016.

[34] K. Tornai, A. Oláh, R. Drenyovszki, L. Kovács, I. Pinté, and J. Levendovszky. Recurrent neural network based user classification for smart grids. In *IEEE Power Energy Society Innovative Smart Grid Technologies Conference*, pages 1–5, 2017.

[35] T. Vantuch, S. Mišák, and J. Stuchlý. Power quality prediction designed as binary classification in AC coupling off-grid system. In *IEEE International Conference on Environment and Electrical Engineering*, pages 1–6, 2016.

[36] A. Weber, B. Metevier, Y. Brun, P. S. Thomas, and B. C. da Silva. Enforcing delayed-impact fairness guarantees. 2022.

[37] J. Xie, I. Alvarez-Fernandez, and W. Sun. A review of machine learning applications in power system resilience. In *2020 IEEE Power & Energy Society General Meeting (PESGM)*, pages 1–5. IEEE, 2020.

[38] A. Yan and B. Howe. Fairness-aware demand prediction for new mobility. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(01):1079–1087, Apr. 2020.

[39] T. Yin, R. Raab, M. Liu, and Y. Liu. Long-term fairness with unknown dynamics. *arXiv preprint arXiv:2304.09362*, 2023.

[40] M. B. Zafar, I. Valera, M. Gomez Rodriguez, and K. P. Gummadi. Fairness beyond disparate treatment & disparate impact: Learning classification without disparate mistreatment. In *Proceedings of the 26th international conference on world wide web*, pages 1171–1180, 2017.

[41] M. B. Zafar, I. Valera, M. G. Rogriguez, and K. P. Gummadi. Fairness Constraints: Mechanisms for Fair Classification. In *Proceedings of the International Conference on Artificial Intelligence and Statistics*, volume 54, pages 962–970. PMLR, 2017.

[42] R. Zemel, Y. Wu, K. Swersky, T. Pitassi, and C. Dwork. Learning fair representations. In *Proceedings of the International Conference on Machine Learning*, volume 28, pages 325–333. PMLR, 2013.

[43] L. Zhang, X. Shen, F. Zhang, M. Ren, B. Ge, and B. Li. Anomaly detection for power grid based on time series model. In *IEEE International Conference on Computational Science and Engineering and IEEE International Conference on Embedded and Ubiquitous Computing*, pages 188–192, 2019.

[44] X. Zhang and M. Liu. Fairness in learning-based sequential decision algorithms: A survey. In *Handbook of Reinforcement Learning and Control*, pages 525–555. Springer, 2021.

Appendix A

A.1 Proof of Lemma 1

Lemma 1 The difference between the dual variables of two successive iterates $|\lambda_{1,t+1}-\lambda_{2,t+1}|$ can be upper bounded by

$$|\lambda_{1,t+1} - \lambda_{2,t+1}| \leq \bar{\lambda} := 2\mu M + \frac{2GR + \frac{R^2}{2\alpha} + \mu M^2}{\epsilon - \bar{V}(g)}. \tag{A.1}$$

Proof:

Based on the updating rule of Equation 4.10 and Equation 4.11, we have

$$\begin{aligned}
|\lambda_{1,t+1}|^2 &= |\max(0, \lambda_{1,t} + \mu g_t(\mathbf{x}_t))|^2 \\
&\leq |\lambda_{1,t} + \mu g_t(\mathbf{x}_t)|^2 \\
&= |\lambda_{1,t}|^2 + 2\mu\lambda_{1,t} g_t(\mathbf{x}_t) + \mu^2 |g_t(\mathbf{x}_t)|^2, \\
|\lambda_{2,t+1}|^2 &= |\max(0, \lambda_{2,t} - \mu g_t(\mathbf{x}_t))|^2 \\
&\leq |\lambda_{2,t} + \mu g_t(\mathbf{x}_t)|^2 \\
&= |\lambda_{2,t}|^2 - 2\mu\lambda_{2,t} g_t(\mathbf{x}_t) + \mu^2 |g_t(\mathbf{x}_t)|^2.
\end{aligned}$$

Reorganizing the terms leads to

$$\Delta(\lambda_{1,t}) \coloneqq \frac{|\lambda_{1,t+1}|^2 - |\lambda_{1,t}|^2}{2} \leq \mu\lambda_{1,t} g_t(\mathbf{x}_t) + \frac{\mu^2}{2}|g_t(\mathbf{x}_t)|^2, \tag{A.2a}$$

$$\Delta(\lambda_{2,t}) \coloneqq \frac{|\lambda_{2,t+1}|^2 - |\lambda_{2,t}|^2}{2} \leq -\mu\lambda_{2,t} g_t(\mathbf{x}_t) + \frac{\mu^2}{2}|g_t(\mathbf{x}_t)|^2. \tag{A.2b}$$

Combining the two inequalities results in

$$\Delta(\boldsymbol{\lambda}_t) \coloneqq \frac{||\boldsymbol{\lambda}_{t+1}||^2 - ||\boldsymbol{\lambda}_t||^2}{2} \leq \mu\boldsymbol{\lambda}_t^\top \bar{\mathbf{g}}_t(\mathbf{x}_t) + \mu^2|g_t(\mathbf{x}_t)|^2. \tag{A.3}$$

where $\boldsymbol{\lambda}_t = [\lambda_{1,t}, \lambda_{2,t}]^\top$ and $\bar{\mathbf{g}}_t(\mathbf{x}) = [g_t(\mathbf{x}), -g_t(\mathbf{x})]^\top$. Based on the update in Equation 4.9, i.e., $\mathbf{x}_{t+1} = \min_{\mathbf{x}} \nabla f_t(\mathbf{x}_t)^\top(\mathbf{x} - \mathbf{x}_t) + \boldsymbol{\lambda}_t^\top \bar{\mathbf{g}}_t(\mathbf{x}) + \frac{||\mathbf{x}-\mathbf{x}_t||^2}{2\alpha}$, the following inequality hold for $\tilde{\mathbf{x}}$

$$\begin{aligned}
&\nabla f_t(\mathbf{x}_t)^\top(\mathbf{x}_{t+1} - \mathbf{x}_t) + \boldsymbol{\lambda}_{t+1}^\top \bar{\mathbf{g}}_t(\mathbf{x}_{t+1}) + \frac{1}{2\alpha}||\mathbf{x}_{t+1} - \mathbf{x}_t||^2 \\
\leq &\nabla f_t(\mathbf{x}_t)^\top(\tilde{\mathbf{x}}_t - \mathbf{x}_t) + \boldsymbol{\lambda}_{t+1}^\top \bar{\mathbf{g}}_t(\tilde{\mathbf{x}}_t) + \frac{1}{2\alpha}||\tilde{\mathbf{x}}_t - \mathbf{x}_t||^2.
\end{aligned}$$

In the case when $\lambda_{1,t+1} \geq \lambda_{2,t+1}$, the right-hand side can be equivalently written and bounded as follows

$$\begin{aligned}
&\nabla f_t(\mathbf{x}_t)^\top(\tilde{\mathbf{x}}_t - \mathbf{x}_t) + \boldsymbol{\lambda}_{t+1}^\top \bar{\mathbf{g}}_t(\tilde{\mathbf{x}}_t) + \frac{1}{2\alpha}||\tilde{\mathbf{x}}_t - \mathbf{x}_t||^2 \\
&= \nabla f_t(\mathbf{x}_t)^\top(\tilde{\mathbf{x}}_t - \mathbf{x}_t) + (\lambda_{1,t+1} - \lambda_{2,t+1}) g_t(\tilde{\mathbf{x}}_t) + \frac{1}{2\alpha}||\tilde{\mathbf{x}}_t - \mathbf{x}_t||^2 \\
&\overset{(a)}{\leq} \nabla f_t(\mathbf{x}_t)^\top(\tilde{\mathbf{x}}_t - \mathbf{x}_t) - \epsilon(\lambda_{1,t+1} - \lambda_{2,t+1}) + \frac{1}{2\alpha}||\tilde{\mathbf{x}}_t - \mathbf{x}_t||^2 \\
&= \nabla f_t(\mathbf{x}_t)^\top(\tilde{\mathbf{x}}_t - \mathbf{x}_t) - \epsilon|\lambda_{1,t+1} - \lambda_{2,t+1}| + \frac{1}{2\alpha}||\tilde{\mathbf{x}}_t - \mathbf{x}_t||^2,
\end{aligned} \tag{A.4}$$

where (a) is true do to assumption 4. Rearranging terms in Equation A.4,

$$
\begin{aligned}
&\boldsymbol{\lambda}_{t+1}^{\top}\bar{\mathbf{g}}_t(\mathbf{x}_{t+1}) \\
\leq &\nabla f_t(\mathbf{x}_t)^{\top}(\tilde{\mathbf{x}}_t - \mathbf{x}_t) - \nabla f_t(\mathbf{x}_t)^{\top}(\mathbf{x}_{t+1} - \mathbf{x}_t) - \epsilon|\lambda_{1,t+1} - \lambda_{2,t+1}| \\
&+ \frac{1}{2\alpha}||\tilde{\mathbf{x}}_t - \mathbf{x}_t||^2 - \frac{1}{2\alpha}||\mathbf{x}_{t+1} - \mathbf{x}_t||^2 \\
\overset{(a)}{\leq} &||\nabla f_t(\mathbf{x}_t)||||\tilde{\mathbf{x}}_t - \mathbf{x}_t|| - ||\nabla f_t(\mathbf{x}_t)||||\mathbf{x}_{t+1} - \mathbf{x}_t|| - \epsilon|\lambda_{1,t+1} - \lambda_{2,t+1}| \\
&+ \frac{1}{2\alpha}||\tilde{\mathbf{x}}_t - \mathbf{x}_t||^2 - \frac{1}{2\alpha}||\mathbf{x}_{t+1} - \mathbf{x}_t||^2 \\
\overset{(b)}{\leq} &2GR - \epsilon|\lambda_{1,t+1} - \lambda_{2,t+1}| + \frac{R^2}{2\alpha},
\end{aligned} \tag{A.5}
$$

where (a) comes from Cauchy–Schwarz inequality and (b) results from assumptions 2 and 3.

Similarly, in the case when $\lambda_{1,t+1} \leq \lambda_{2,t+1}$, the following holds

$$
\begin{aligned}
&\nabla f_t(\mathbf{x}_t)^{\top}(\mathbf{x}_{t+1} - \mathbf{x}_t) + \boldsymbol{\lambda}_{t+1}^{\top}\bar{\mathbf{g}}_t(\mathbf{x}_{t+1}) + \frac{1}{2\alpha}||\mathbf{x}_{t+1} - \mathbf{x}_t||^2 \\
&\leq \nabla f_t(\mathbf{x}_t)^{\top}(\tilde{\tilde{\mathbf{x}}}_t - \mathbf{x}_t) + \boldsymbol{\lambda}_{t+1}^{\top}\bar{\mathbf{g}}_t(\tilde{\tilde{\mathbf{x}}}_t) + \frac{1}{2\alpha}||\tilde{\tilde{\mathbf{x}}}_t - \mathbf{x}_t||^2 \\
&= \nabla f_t(\mathbf{x}_t)^{\top}(\tilde{\tilde{\mathbf{x}}}_t - \mathbf{x}_t) + (\lambda_{2,t+1} - \lambda_{1,t+1})(-g_t(\tilde{\tilde{\mathbf{x}}}_t)) + \frac{1}{2\alpha}||\tilde{\tilde{\mathbf{x}}}_t - \mathbf{x}_t||^2 \\
&\leq \nabla f_t(\mathbf{x}_t)^{\top}(\tilde{\tilde{\mathbf{x}}}_t - \mathbf{x}_t) - \epsilon(\lambda_{2,t+1} - \lambda_{1,t+1}) + \frac{1}{2\alpha}||\tilde{\tilde{\mathbf{x}}}_t - \mathbf{x}_t||^2 \\
&\leq \nabla f_t(\mathbf{x}_t)^{\top}(\tilde{\tilde{\mathbf{x}}}_t - \mathbf{x}_t) - \epsilon|\lambda_{1,t+1} - \lambda_{2,t+1}| + \frac{1}{2\alpha}||\tilde{\tilde{\mathbf{x}}}_t - \mathbf{x}_t||^2.
\end{aligned} \tag{A.6}
$$

Rearranging terms in Equation A.6

$$
\begin{aligned}
&\boldsymbol{\lambda}_{t+1}^{\top}\bar{\mathbf{g}}_t(\mathbf{x}_{t+1}) \\
\leq&\nabla f_t(\mathbf{x}_t)^{\top}(\tilde{\tilde{\mathbf{x}}}_t - \mathbf{x}_t) - \nabla f_t(\mathbf{x}_t)^{\top}(\mathbf{x}_{t+1} - \mathbf{x}_t) - \epsilon|\lambda_{1,t+1} - \lambda_{2,t+1}| \\
&+ \frac{1}{2\alpha}||\tilde{\tilde{\mathbf{x}}}_t - \mathbf{x}_t||^2 - \frac{1}{2\alpha}||\mathbf{x}_{t+1} - \mathbf{x}_t||^2 \\
\leq&||\nabla f_t(\mathbf{x}_t)||||\tilde{\tilde{\mathbf{x}}}_t - \mathbf{x}_t|| - ||\nabla f_t(\mathbf{x}_t)||||\mathbf{x}_{t+1} - \mathbf{x}_t|| - \epsilon|\lambda_{1,t+1} - \lambda_{2,t+1}| \\
&+ \frac{1}{2\alpha}||\tilde{\tilde{\mathbf{x}}}_t - \mathbf{x}_t||^2 - \frac{1}{2\alpha}||\mathbf{x}_{t+1} - \mathbf{x}_t||^2 \\
\leq&2GR - \epsilon|\lambda_{1,t+1} - \lambda_{2,t+1}| + \frac{R^2}{2\alpha}.
\end{aligned} \tag{A.7}
$$

Combining Equation A.5 and Equation A.7 leads to

$$
\boldsymbol{\lambda}_{t+1}^{\top}\bar{\mathbf{g}}_t(\mathbf{x}_{t+1}) \leq 2GR - \epsilon|\lambda_{1,t+1} - \lambda_{2,t+1}| + \frac{R^2}{2\alpha}. \tag{A.8}
$$

Hence, Equation A.3 can be re-written as follows by adding and subtracting $\mu\boldsymbol{\lambda}_{t+1}^{\top}\bar{\mathbf{g}}_t(\mathbf{x}_{t+1})$

$$
\begin{aligned}
&\Delta(\boldsymbol{\lambda}_{t+1}) \\
\leq&\mu\boldsymbol{\lambda}_{t+1}^{\top}\bar{\mathbf{g}}_{t+1}(\mathbf{x}_{t+1}) - \mu\boldsymbol{\lambda}_{t+1}^{\top}\bar{\mathbf{g}}_t(\mathbf{x}_{t+1}) + \mu\boldsymbol{\lambda}_{t+1}^{\top}\bar{\mathbf{g}}_t(\mathbf{x}_{t+1}) + \mu^2|g_{t+1}(\mathbf{x}_{t+1})|^2 \\
\overset{(a)}{\leq}&\mu\boldsymbol{\lambda}_{t+1}^{\top}(\bar{\mathbf{g}}_{t+1}(\mathbf{x}_{t+1}) - \bar{\mathbf{g}}_t(\mathbf{x}_{t+1})) + \mu(2GR - \epsilon|\lambda_{1,t+1} - \lambda_{2,t+1}| + \frac{R^2}{2\alpha}) + \mu^2M^2 \\
=&\mu(\lambda_{1,t+1} - \lambda_{2,t+1})(g_{t+1}(\mathbf{x}_{t+1}) - g_t(\mathbf{x}_{t+1})) - \epsilon\mu|\lambda_{1,t+1} - \lambda_{2,t+1}| + 2\mu GR + \frac{\mu R^2}{2\alpha} + \mu^2M^2 \\
\overset{(b)}{\leq}&\mu\bar{V}(g)|\lambda_{1,t+1} - \lambda_{2,t+1}| - \mu\epsilon|\lambda_{1,t+1} - \lambda_{2,t+1}| + 2\mu GR + \frac{\mu R^2}{2\alpha} + \mu^2M^2,
\end{aligned} \tag{A.9}
$$

where (a) results because of Equation A.8 and (b) holds due to Cauchy-Schwarz inequality and Assumption 5.

Note that Lemma 1 holds at $t = 1$ since $\boldsymbol{\lambda}_1 = \mathbf{0}$. Assume that (A.1) holds for all time slots

up till $t+1$, and $t+2$ is the first time slot that (A.1) does not hold. Then we have

$$|\lambda_{1,t+1} - \lambda_{2,t+1}| \leq 2\mu M + \frac{2GR + \frac{R^2}{2\alpha} + \mu M^2}{\epsilon - \bar{V}(g)}, \tag{A.10a}$$

$$|\lambda_{1,t+2} - \lambda_{2,t+2}| \geq 2\mu M + \frac{2GR + \frac{R^2}{2\alpha} + \mu M^2}{\epsilon - \bar{V}(g)}. \tag{A.10b}$$

Hence, it holds that

$$\begin{aligned}
&|\lambda_{1,t+1} - \lambda_{2,t+1}| \\
\geq &|\lambda_{1,t+2} - \lambda_{2,t+2}| - |(\lambda_{1,t+1} - \lambda_{2,t+1}) - (\lambda_{1,t+2} - \lambda_{2,t+2})| \\
\geq &|\lambda_{1,t+2} - \lambda_{2,t+2}| - |(\lambda_{1,t+1} + \mu g_{t+1}(\mathbf{x}_{t+1}) - \lambda_{1,t+1}) - (\lambda_{2,t+1} - \mu g_{t+1}(\mathbf{x}_{t+1}) - \lambda_{2,t+1})| \\
= &|\lambda_{1,t+2} - \lambda_{2,t+2}| - |2\mu g_{1,t+1}(\mathbf{x}_{t+1})| \\
\overset{(a)}{\geq} &\frac{2GR + \frac{R^2}{2\alpha} + \mu M^2}{\epsilon - \bar{V}(g)},
\end{aligned}$$

where (a) holds due to Equation A.10b and Assumption 2. Substituting it into (A.9) and combining with Assumption 5 arrives at $\Delta(\boldsymbol{\lambda}_{t+1}) = \frac{\lambda_{1,t+2}^2 + \lambda_{2,t+2}^2 - \lambda_{1,t+1}^2 - \lambda_{2,t+1}^2}{2} \leq 0$, which can be written as

$$(\lambda_{1,t+2} + \lambda_{1,t+1})(\lambda_{1,t+2} - \lambda_{1,t+1}) \leq (\lambda_{2,t+2} + \lambda_{2,t+1})(\lambda_{2,t+1} - \lambda_{2,t+2}). \tag{A.11}$$

If $g_{t+1}(\mathbf{x}_{t+1}) > 0$, the updating rule in Equation 4.10 leads to $\lambda_{1,t+2} - \lambda_{1,t+1} = \mu g_{t+1}(\mathbf{x}_{t+1})$ and Equation 4.11 implies $0 \leq \lambda_{2,t+1} - \lambda_{2,t+2} \leq \mu g_{t+1}(\mathbf{x}_{t+1})$. Substituting them into Equation A.11 leads to

$$(\lambda_{1,t+2} + \lambda_{1,t+1})\mu g_{t+1}(\mathbf{x}_{t+1}) \leq (\lambda_{2,t+2} + \lambda_{2,t+1})\mu g_{t+1}(\mathbf{x}_{t+1}),$$

which is equivalent to $\lambda_{1,t+2} + \lambda_{1,t+1} \leq \lambda_{2,t+2} + \lambda_{2,t+1}$. Furthermore, since $\lambda_{1,t+2} + \lambda_{1,t+1} =$

$2\lambda_{1,t+1} + \mu g_{t+1}(\mathbf{x}_{t+1})$ and $\lambda_{2,t+1} + \lambda_{2,t+2} \leq 2\lambda_{2,t+2} + \mu g_{t+1}(\mathbf{x}_{t+1})$, we have

$$2\lambda_{1,t+1} + \mu g_{t+1}(\mathbf{x}_{t+1}) \leq 2\lambda_{2,t+2} + \mu g_{t+1}(\mathbf{x}_{t+1}),$$

which implies $\lambda_{1,t+1} \leq \lambda_{2,t+2}$ and henceforth $\lambda_{1,t+2} \leq \lambda_{2,t+1}$.
Due to $\lambda_{1,t+1} \leq \lambda_{2,t+2} \leq \lambda_{2,t+1}$, $\lambda_{1,t+1} \leq \lambda_{1,t+2} \leq \lambda_{2,t+1}$ and non-negativity of $\lambda_{1,t+1}, \lambda_{2,t+1}$, $\lambda_{1,t+2}, \lambda_{2,t+2}$, we have

$$|\lambda_{1,t+1} - \lambda_{2,t+1}| \geq |\lambda_{1,t+2} - \lambda_{2,t+2}|,$$

which contradicts the $|\lambda_{1,t+1} - \lambda_{2,t+1}| \leq |\lambda_{1,t+2} - \lambda_{2,t+2}|$.

Similarly, if $g_{t+1}(\mathbf{x}_{t+1}) < 0$, we have $\lambda_{2,t+1} \leq \lambda_{1,t+2} \leq \lambda_{1,t+1}$ and $\lambda_{2,t+1} \leq \lambda_{2,t+2} \leq \lambda_{1,t+1}$. Due to the non-negativity of $\lambda_{1,t+1}, \lambda_{2,t+1}, \lambda_{1,t+2}, \lambda_{2,t+2}$, we can have $|\lambda_{1,t+1} - \lambda_{2,t+1}| \geq |\lambda_{1,t+2} - \lambda_{2,t+2}|$ which contradicts the $|\lambda_{1,t+1} - \lambda_{2,t+1}| \leq |\lambda_{1,t+2} - \lambda_{2,t+2}|$. To sum up, Lemma 1 can be proved by contradiction.

A.2 Proof of Lemma 2

Let $\bar{\lambda}_1, \bar{\lambda}_2$ denote the max value of $\lambda_{1,t}, \lambda_{2,t}$ across the whole time horizon, i.e., $\bar{\lambda}_1 \coloneqq \max_{t\in\{1,\cdots,T\}} \lambda_{1,t}$, $\bar{\lambda}_2 \coloneqq \max_{t\in\{1,\cdots,T\}} \lambda_{2,t}$.

Lemma 2: Let $\bar{\delta} := \max_{t\in\{1,\cdots,T\}} |\lambda_{1,t+1} - \lambda_{2,t+1}|$, the following equality holds

$$\bar{\delta} = \max(\bar{\lambda}_1, \bar{\lambda}_2), \tag{A.12}$$

which implies that $\bar{\lambda}_1 \leq \bar{\delta} \leq \bar{\lambda}$, and $\bar{\lambda}_2 \leq \bar{\delta} \leq \bar{\lambda}$.

Proof of Lemma 2: By mathematical induction

Proposition:

For time slot t, $\bar{\delta}_t = \max_t |\lambda_{1,t+1} - \lambda_{2,t+1}| = \max(\bar{\lambda}_{1,t}, \bar{\lambda}_{2,t})$. $\bar{\lambda}_{1,t}$ and $\bar{\lambda}_{2,t}$ denote the maximum value of $\lambda_{1,t+1}$ and $\lambda_{2,t+1}$ at time t and before time t.

Base case:

Since $\boldsymbol{\lambda}$ is initialized as $\mathbf{0}$, $\bar{\delta}_t = \max(\bar{\lambda}_1, \bar{\lambda}_2)$ holds for $t = 1, 2$.

Inductive step:

Assume τ is the last time $\lambda_{1,\tau+1} = 0$ or $\lambda_{2,\tau+1} = 0$ holds, meaning $\lambda_{1,\tilde{t}} > 0$ and $\lambda_{2,\tilde{t}} > 0$ for all $\tau < \tilde{t} < t$. Hence, $\bar{\lambda}_{1,\tau} = \max_{t\in\{1,\cdots,\tau\}} \lambda_{1,t+1}$ and $\bar{\lambda}_{2,\tau} = \max_{t\in\{1,\cdots,\tau\}} \lambda_{2,t+1}$ which are the maximum value of $\lambda_{1,t}$ and $\lambda_{2,t}$ at time τ and before time τ. We also have $\bar{\delta}_\tau = \max_{t\in\{1,\cdots,\tau\}} |\lambda_{1,t+1} - \lambda_{2,t+1}| = \max(\bar{\lambda}_{1,\tau}, \bar{\lambda}_{2,\tau})$. Below, we will discuss 3 different cases: $\lambda_{2,t+1} = 0$, $\lambda_{1,t+1} = 0$ and $\lambda_{1,t+1}, \lambda_{2,t+1} > 0$.

Case 1: $\lambda_{1,t+1} > 0$ and $\lambda_{2,t+1} > 0$, then the maximum operations in Equation 4.10 and Equation 4.11 would not be activated. So $\lambda_{1,t+1} = \lambda_{1,\tau+1} + \sum_{\tilde{t}=\tau+1}^{t} \mu g_{\tilde{t}}$ and $\lambda_{2,t+1} = \lambda_{2,\tau+1} - \sum_{\tilde{t}=\tau+1}^{t} \mu g_{\tilde{t}}$.

If $\lambda_{2,\tau+1} = 0$, then $\sum_{\tilde{t}=\tau+1}^{t} \mu g_{\tilde{t}} < 0$ which leads to

$$\lambda_{1,t+1} \leq \lambda_{1,\tau+1},$$
$$\lambda_{2,t+1} \leq \lambda_{1,\tau+1}.$$

Since $\lambda_{1,\tau+1} \leq \bar{\lambda}_{1,\tau}$ and $\bar{\lambda}_{1,\tau} \leq \bar{\delta}_\tau$, we have

$$\lambda_{1,t+1} \leq \bar{\delta}_\tau, \lambda_{2,t+1} \leq \bar{\delta}_\tau.$$

Due to their nonegativity,

$$|\lambda_{1,t+1} - \lambda_{2,t+1}| \leq \bar{\delta}_\tau.$$

Similarly, if $\lambda_{1,\tau+1} = 0$, then $\sum_{t=\tau+1}^{t} \mu g_t > 0$ which leads to

$$\lambda_{1,t+1} \leq \lambda_{2,\tau+1},$$
$$\lambda_{2,t+1} \leq \lambda_{2,\tau+1},$$

and hence $\lambda_{1,t+1} \leq \bar{\delta}_\tau$, $\lambda_{2,t+1} \leq \bar{\delta}_\tau$, and $|\lambda_{1,t+1} - \lambda_{2,t+1}| \leq \bar{\delta}_\tau$ also holds due to the non-negativity.

So $\bar{\delta}_t = \bar{\delta}_\tau = \max(\bar{\lambda}_{1,t}, \bar{\lambda}_{2,t})$ holds for t while $\lambda_{1,t+1} > 0$ and $\lambda_{2,t+1} > 0$. Since $\lambda_{1,\tilde{t}+1} > 0$ and $\lambda_{2,\tilde{t}+1} > 0$ for all $\tau < \tilde{t} < t$, $\bar{\delta}_{\tilde{t}} = \bar{\delta}_\tau$ also holds.

Case 2: $\lambda_{1,t+1} = 0$, then $\lambda_{1,t+1} \leq \bar{\lambda}_{1,\tau+1}$ due to the non-negativity of λ_1. So we have $\bar{\lambda}_{1,t} = \bar{\lambda}_{1,\tau}$ Based on the definition of $\bar{\lambda}_{2,t}$, we have

$$\bar{\lambda}_{2,t} = \max(\bar{\lambda}_{2,\tau}, \lambda_{2,t+1}).$$

Then $\bar{\delta}_t$ can be written as

$$\begin{aligned}\bar{\delta}_t = \max(|\lambda_{1,t+1} - \lambda_{2,t+1}|, \bar{\delta}_\tau) &\overset{(a)}{=} \max(\lambda_{2,t+1}, \bar{\delta}_\tau) \\ &\overset{(b)}{=} \max(\lambda_{2,t+1}, \bar{\lambda}_{2,\tau}, \bar{\lambda}_{1,\tau}) \\ &\overset{(c)}{=} \max(\bar{\lambda}_{1,t}, \bar{\lambda}_{2,t}),\end{aligned}$$

where (a) holds due to $\lambda_{1,t+1} = 0$, (b) holds due to $\bar{\delta}_\tau = \max(\bar{\lambda}_{2,\tau}, \bar{\lambda}_{1,\tau})$, and (c) holds since $\bar{\lambda}_{2,t} = \max(\bar{\lambda}_{2,\tau}, \lambda_{2,t+1}), \bar{\lambda}_{1,t} = \bar{\lambda}_{1,\tau}$.

So $\bar{\delta}_t = \max(\bar{\lambda}_{1,t}, \bar{\lambda}_{2,t})$ holds while $\lambda_{1,t+1} = 0$.

Case 3: $\lambda_{2,t+1} = 0$, similar to case 2, we have

$$\begin{aligned}
\bar{\lambda}_{2,t} &= \bar{\lambda}_{2,\tau}, \\
\bar{\lambda}_{1,t} &= \max(\bar{\lambda}_{1,\tau}, \lambda_{1,t+1}), \\
\bar{\delta}_t &= \max(|\lambda_{1,t+1} - \lambda_{2,t+1}|, \bar{\delta}_\tau) = \max(\lambda_{1,t+1}, \bar{\delta}_\tau) \\
&= \max(\lambda_{1,t+1}, \bar{\lambda}_{1,\tau}, \bar{\lambda}_{2,\tau}) = \max(\bar{\lambda}_{1,t}, \bar{\lambda}_{2,t}).
\end{aligned}$$

So $\bar{\delta}_t = \max(\bar{\lambda}_{1,t}, \bar{\lambda}_{2,t})$ holds while $\lambda_{1,t+1} = 0$.

Combining the three cases, $\bar{\delta}_t = \max(\bar{\lambda}_{1,t}, \bar{\lambda}_{2,t})$ also holds for t.

Conclusion:

Since both the base case and the inductive step hold, the statement $\bar{\lambda}_d = \max(\bar{\lambda}_1, \bar{\lambda}_2)$ holds across the whole time horizon $\{1, \cdots, T\}$.

A.3 Proof of Theorem 1

Using the recursion in (4.10) and (4.11), we have

$$\lambda_{1,T+1} \geq \lambda_{1,T} + \mu g_T(\mathbf{x}_T) \geq \lambda_{1,1} + \sum_{t=1}^{\top} \mu g_t(\mathbf{x}_t), \tag{A.13}$$

$$\lambda_{2,T+1} \geq \lambda_{2,T} - \mu g_T(\mathbf{x}_T) \geq \lambda_{2,1} - \sum_{t=1}^{T} \mu g_t(\mathbf{x}_t). \tag{A.14}$$

Since $\bar{\lambda}_1, \bar{\lambda}_2$ denote the max value of $\lambda_{1,t}, \lambda_{2,t}$ across the whole time horizon, i.e., $\bar{\lambda}_1 \coloneqq \max_{t\in\{1,\cdots,T\}} \lambda_{1,t}$, $\bar{\lambda}_2 \coloneqq \max_{t\in\{1,\cdots,T\}} \lambda_{2,t}$, the following inequalities hold

$$\sum_{t=1}^{T} g_t(\mathbf{x}_t) \leq \frac{\lambda_{1,T+1}}{\mu} - \frac{\lambda_{1,1}}{\mu} \leq \frac{\lambda_{1,T+1}}{\mu} \leq \frac{\bar{\lambda}_1}{\mu}, \tag{A.15}$$

$$\sum_{t=1}^{T} g_t(\mathbf{x}_t) \geq -\frac{\lambda_{2,T+1}}{\mu} + \frac{\lambda_{2,1}}{\mu} \geq -\frac{\lambda_{2,T+1}}{\mu} \geq -\frac{\bar{\lambda}_2}{\mu}. \tag{A.16}$$

Furthermore, from **Lemma 2**, we have $\bar{\lambda}_1 \leq \bar{\lambda}$ and $\bar{\lambda}_2 \leq \bar{\lambda}$. In addition, based on **Lemma 1**, $\bar{\lambda} := 2\mu M + \frac{2GR + \frac{R^2}{2\alpha} + \mu M^2}{\epsilon - \bar{V}(g)}$. Substituting $\bar{\lambda}$ into Equation A.15 and Equation A.16 leads to

$$\sum_{t=1}^{T} g_t(\mathbf{x}_t) \leq \frac{\bar{\lambda}_1}{\mu} \leq \frac{\bar{\lambda}}{\mu} = 2M + \frac{\frac{2GR}{\mu} + \frac{R^2}{2\mu\alpha} + M^2}{\epsilon - \bar{V}(g)}, \tag{A.17}$$

$$\sum_{t=1}^{T} g_t(\mathbf{x}_t) \geq -\frac{\bar{\lambda}_2}{\mu} \geq -\frac{\bar{\lambda}}{\mu} = -2M - \frac{\frac{2GR}{\mu} + \frac{R^2}{2\mu\alpha} + M^2}{\epsilon - \bar{V}(g)}. \tag{A.18}$$

Based on he definition in Equation 4.14, $\mathcal{F}_T = \sum_{t=1}^{T} g_t(\mathbf{x}_t)$, the proof is completed: $-2M - \frac{\frac{2GR}{\mu} + \frac{R^2}{2\mu\alpha} + M^2}{\epsilon - \bar{V}(g)} \leq \mathcal{F}_T \leq 2M + \frac{\frac{2GR}{\mu} + \frac{R^2}{2\mu\alpha} + M^2}{\epsilon - \bar{V}(g)}$.

A.4 Proof of Theorem 2

$h_t(\mathbf{x}) = \nabla f_t(\mathbf{x}_t)^\top(\mathbf{x} - \mathbf{x}_t) + \boldsymbol{\lambda}_{t+1}^\top \bar{\mathbf{g}}_t(\mathbf{x}) + \frac{||\mathbf{x} - \mathbf{x}_t||^2}{2\alpha}$ is $1/\alpha$-strongly convex [21], which implies that

$$h_t(\mathbf{y}) - h_t(\mathbf{x}) \geq \nabla h_t(\mathbf{x})^\top(\mathbf{y} - \mathbf{x}) + \frac{1}{2\alpha}||\mathbf{y} - \mathbf{x}||^2.$$

Since $\mathbf{x}_{t+1}$ is the minimizer of the problem $\min_{\mathbf{x}} h_t(\mathbf{x})$:

$$\nabla h_t(\mathbf{x}_{t+1})^\top(\mathbf{y} - \mathbf{x}_{t+1}) \geq 0, \forall \mathbf{y}.$$

Set $\mathbf{y} = \mathbf{x}_t^*$ which is the optimal solution at time t as Equation 4.13:

$$\mathbf{x}_t^* = \arg\min_{\mathbf{x}_t \in \mathcal{X}_t} f_t(\mathbf{x}_t) \quad \text{s.t. } g_t(\mathbf{x}_t) = 0.$$

Then it leads to

$$h_t(\mathbf{x}_t^*) \geq h_t(\mathbf{x}_{t+1}) + \frac{1}{2\alpha}||\mathbf{x}_t^* - \mathbf{x}_{t+1}||^2,$$

which is equivalent to

$$\begin{aligned}
&\nabla f_t(\mathbf{x}_t)^\top(\mathbf{x}_{t+1} - \mathbf{x}_t) + \boldsymbol{\lambda}_{t+1}^\top \bar{\mathbf{g}}_t(\mathbf{x}_{t+1}) + \frac{||\mathbf{x}_{t+1} - \mathbf{x}_t||^2}{2\alpha} \\
\leq &\nabla f_t(\mathbf{x}_t)^\top(\mathbf{x}_t^* - \mathbf{x}_t) + \boldsymbol{\lambda}_{t+1}^\top \bar{\mathbf{g}}_t(\mathbf{x}_t^*) + \frac{||\mathbf{x}_t^* - \mathbf{x}_t||^2}{2\alpha} - \frac{||\mathbf{x}_t^* - \mathbf{x}_{t+1}||^2}{2\alpha}.
\end{aligned} \tag{A.19}$$

Adding $f_t(\mathbf{x}_t)$ on both sides

$$\begin{aligned}
&f_t(\mathbf{x}_t) + \nabla f_t(\mathbf{x}_t)^\top(\mathbf{x}_{t+1} - \mathbf{x}_t) + \boldsymbol{\lambda}_{t+1}^\top \bar{\mathbf{g}}_t(\mathbf{x}_{t+1}) + \frac{||\mathbf{x}_{t+1} - \mathbf{x}_t||^2}{2\alpha} \\
\leq &f_t(\mathbf{x}_t) + \nabla f_t(\mathbf{x}_t)^\top(\mathbf{x}_t^* - \mathbf{x}_t) + \boldsymbol{\lambda}_{t+1}^\top \bar{\mathbf{g}}_t(\mathbf{x}_t^*) + \frac{||\mathbf{x}_t^* - \mathbf{x}_t||^2}{2\alpha} - \frac{||\mathbf{x}_t^* - \mathbf{x}_{t+1}||^2}{2\alpha} \\
\overset{(a)}{\leq} &f_t(\mathbf{x}_t^*) + \boldsymbol{\lambda}_{t+1}^\top \bar{\mathbf{g}}_t(\mathbf{x}_t^*) + \frac{||\mathbf{x}_t^* - \mathbf{x}_t||^2}{2\alpha} - \frac{||\mathbf{x}_t^* - \mathbf{x}_{t+1}||^2}{2\alpha} \\
\overset{(b)}{\leq} &f_t(\mathbf{x}_t^*) + \frac{||\mathbf{x}_t^* - \mathbf{x}_t||^2}{2\alpha} - \frac{||\mathbf{x}_t^* - \mathbf{x}_{t+1}||^2}{2\alpha},
\end{aligned} \tag{A.20}$$

where (a) is due to the convexity of $f_t(\mathbf{x}_t)$ and (b) comes from $g_t(\mathbf{x}_t^*) = 0$. Meanwhile, the term $\nabla f_t(\mathbf{x}_t)^\top(\mathbf{x}_{t+1} - \mathbf{x}_t)$ is given by

$$\begin{aligned}
-\nabla f_t(\mathbf{x}_t)^\top(\mathbf{x}_{t+1} - \mathbf{x}_t) &\leq ||\nabla f_t(\mathbf{x}_t)||||\mathbf{x}_{t+1} - \mathbf{x}_t|| \\
&\leq \frac{||\nabla f_t(\mathbf{x}_t)||^2}{2\eta} + \frac{\eta}{2}||\mathbf{x}_{t+1} - \mathbf{x}_t||^2 \\
&\overset{(c)}{\leq} \frac{G^2}{2\eta} + \frac{\eta}{2}||\mathbf{x}_{t+1} - \mathbf{x}_t||^2,
\end{aligned} \tag{A.21}$$

where η is an arbitrary positive constant, and (c) holds due to Assumption 2. Combining Equation A.20 and Equation A.21 and rearranging the terms, we have

$$\begin{aligned} f_t(\mathbf{x}_t) + \boldsymbol{\lambda}_{t+1}^\top \mathbf{g}_t(\mathbf{x}_{t+1}) \leq & f_t(\mathbf{x}_t^*) + (\frac{\eta}{2} + \frac{1}{2\alpha})||\mathbf{x}_{t+1} - \mathbf{x}x_t||^2 \\ & + \frac{1}{2\alpha}(||\mathbf{x}_t^* - \mathbf{x}_t||^2 - ||\mathbf{x}_t^* - \mathbf{x}_{t+1}||^2) + \frac{G^2}{2\eta}. \end{aligned} \tag{A.22}$$

Setting $\eta = \frac{1}{\alpha}$, we have $\frac{\eta}{2} - \frac{1}{2\alpha} = 0$, hence

$$f_t(\mathbf{x}_t) + \boldsymbol{\lambda}_{t+1}^\top \mathbf{g}_t(\mathbf{x}_{t+1}) \leq f_t(\mathbf{x}_t^*) + \frac{1}{2\alpha}(||\mathbf{x}_t^* - \mathbf{x}_t||^2 - ||\mathbf{x}_t^* - \mathbf{x}_{t+1}||^2) + \frac{\alpha G^2}{2}. \tag{A.23}$$

Based on $\Delta(\boldsymbol{\lambda}_t) \leq \mu \boldsymbol{\lambda}_t^\top \bar{\mathbf{g}}_t(\mathbf{x}_t) + \mu^2 |g_t(\mathbf{x}_t)|^2$ (A.3):

$$\begin{aligned} \frac{\Delta(\boldsymbol{\lambda}_{t+1})}{\mu} + f_t(\mathbf{x}_t) \leq & f_t(\mathbf{x}_t) + \boldsymbol{\lambda}_{t+1}^\top \bar{\mathbf{g}}_{t+1}(\mathbf{x}_{t+1}) + \boldsymbol{\lambda}_{t+1}^\top \bar{\mathbf{g}}_t(\mathbf{x}_{t+1}) - \boldsymbol{\lambda}_{t+1}^\top \bar{\mathbf{g}}_t(\mathbf{x}_{t+1}) + \mu^2 |g_t(\mathbf{x}_t)|^2 \\ \leq & f_t(\mathbf{x}_t^*) + \frac{1}{2\alpha}(||\mathbf{x}_t^* - \mathbf{x}_t||^2 - ||\mathbf{x}_t^* - \mathbf{x}_{t+1}||^2) + \\ & \boldsymbol{\lambda}_{t+1}^\top (\bar{\mathbf{g}}_{t+1}(\mathbf{x}_{t+1}) - \bar{\mathbf{g}}_t(\mathbf{x}_{t+1})) + \mu^2 |g_t(\mathbf{x}_t)|^2 + \frac{\alpha G^2}{2} \\ \overset{(a)}{\leq} & f_t(\mathbf{x}_t^*) + \frac{1}{2\alpha}(||\mathbf{x}_t^* - \mathbf{x}_t||^2 - ||\mathbf{x}_t^* - \mathbf{x}_{t+1}||^2) \\ & + |\lambda_{1,t+1} - \lambda_{2,t+1}| V(g) + \mu M^2 + \frac{\alpha G^2}{2} \\ \overset{(b)}{\leq} & f_t(\mathbf{x}_t^*) + \frac{1}{2\alpha}(||\mathbf{x}_t^* - \mathbf{x}_t||^2 - ||\mathbf{x}_t^* - \mathbf{x}_{t+1}||^2) + |\bar{\lambda}| V(g) + \mu M^2 + \frac{\alpha G^2}{2}, \end{aligned} \tag{A.24}$$

where $V(g) = \max(0, g_{t+1}(\mathbf{x}_{t+1}) - g_t(\mathbf{x}_{t+1}))$, (a) holds because of Equation A.8, (b) is true due to Lemma 1. By adding and subtracting $||\mathbf{x}_t - \mathbf{x}_{t-1}^*||^2$ in the $||\mathbf{x}_t^* - \mathbf{x}_t||^2 - ||\mathbf{x}_t^* - \mathbf{x}_{t+1}||^2$,

we have that

$$
\begin{aligned}
&||\mathbf{x}_t^* - \mathbf{x}_t||^2 - ||\mathbf{x}_t^* - \mathbf{x}_{t+1}||^2 \\
=&||\mathbf{x}_t^* - \mathbf{x}_t||^2 - ||\mathbf{x}_t - \mathbf{x}_{t-1}^*||^2 + ||\mathbf{x}_t - \mathbf{x}_{t-1}^*||^2 - ||\mathbf{x}_t^* - \mathbf{x}_{t+1}||^2 \\
=&||\mathbf{x}_t^* - \mathbf{x}_{t-1}^*||||\mathbf{x}_t^* - 2\mathbf{x}_t + \mathbf{x}_{t-1}^*|| + ||\mathbf{x}_t - \mathbf{x}_{t-1}^*||^2 - ||\mathbf{x}_t^* - \mathbf{x}_{t+1}||^2 \\
\overset{(c)}{\leq}&2R||\mathbf{x}_t^2 - \mathbf{x}_{t-1}^*|| + ||\mathbf{x}_t - \mathbf{x}_{t-1}^*||^2 - ||\mathbf{x}_t^* - \mathbf{x}_{t+1}||^2,
\end{aligned} \tag{A.25}
$$

where (c) holds since $||\mathbf{x}_t^* - 2\mathbf{x}_t + \mathbf{x}_{t-1}^*|| \leq ||\mathbf{x}_t^* - \mathbf{x}_t|| + ||\mathbf{x}_t - \mathbf{x}_{t-1}^*|| \leq 2R$. Combining Equation A.25 with Equation A.24 can arrive at

$$
\begin{aligned}
\frac{\Delta(\boldsymbol{\lambda}_{t+1})}{\mu} + f_t(\mathbf{x}_t) \leq& f_t(\mathbf{x}_t^*) + |\bar{\lambda}|V(g) + \mu M^2 + \frac{\alpha G^2}{2} \\
&+ \frac{1}{2\alpha}(2R||\mathbf{x}_t^2 - \mathbf{x}_{t-1}^*|| + ||\mathbf{x}_t - \mathbf{x}_{t-1}^*||^2 - ||\mathbf{x}_t^* - \mathbf{x}_{t+1}||^2).
\end{aligned} \tag{A.26}
$$

Summing up over $t = 1, 2, \cdots, T$

$$
\begin{aligned}
&\sum_{t=1}^{T} \frac{\Delta(\boldsymbol{\lambda}_{t+1})}{\mu} + \sum_{t=1}^{T} f_t(\mathbf{x}_t) \\
\leq& \sum_{t=1}^{T} f_t(\mathbf{x}_t^*) + \frac{1}{2\alpha}\sum_{t=1}^{T}(||\mathbf{x}_t - \mathbf{x}_{t-1}^*||^2 - ||\mathbf{x}_t^* - \mathbf{x}_{t+1}||^2) + \frac{R}{\alpha}V(\{\mathbf{x}_t^*\}_{t=1}^T) \\
&+ \sum_{t=1}^{T} |\bar{\lambda}|V(g_t(\mathbf{x}_t)) + \mu M^2 T + \frac{\alpha G^2 T}{2} \\
=& \sum_{t=1}^{T} f_t(\mathbf{x}_t^*) + \frac{1}{2\alpha}(||\mathbf{x}_1 - \mathbf{x}_0^*||^2 - ||\mathbf{x}_T^* - \mathbf{x}_{T+1}||) + \frac{R}{\alpha}V(\{\mathbf{x}_t^*\}_{t=1}^T) \\
&+ |\bar{\lambda}|\sum_{t=1}^{T} V(g_t(\mathbf{x}_t)) + \mu M^2 T + \frac{\alpha G^2 T}{2} \\
\leq& \sum_{t=1}^{T} f_t(\mathbf{x}_t^*) + \frac{1}{2\alpha}(||\mathbf{x}_1 - \mathbf{x}_0^*||^2) + \frac{R}{\alpha}V(\{\mathbf{x}_t^*\}_{t=1}^T) + |\bar{\lambda}|V(\{g_t\}_{t=1}^T) + \mu M^2 T + \frac{\alpha G^2 T}{2}.
\end{aligned} \tag{A.27}
$$

Recall that $V(\{\mathbf{x}_t^*\}_{t=1}^T)$ and $V(\{\bar{\mathbf{g}}_t\}_{t=1}^T)$ are defined as Equation 4.19 and Equation 4.18, and the fact that $\mathcal{R}_T = \sum_{t=1}^T f_t(\mathbf{x}_t) - \sum_{t=1}^T f_t(\mathbf{x}_t^*)$, by rearranging the terms of Equation A.27, we have

$$\begin{aligned}
\mathcal{R}_T \leq & \frac{R}{\alpha} V(\{\mathbf{x}_t^*\}_{t=1}^T) + \frac{||\mathbf{x}_1 - \mathbf{x}_0^*||^2}{2\alpha} + |\bar{\lambda}| V(\{g_t\}_{t=1}^T) \\
& + \mu M^2 T + \frac{\alpha G^2 T}{2} - \sum_{t=1}^T \frac{\Delta(\boldsymbol{\lambda}_{t+1})}{\mu} \\
\leq & \frac{R}{\alpha} V(\{\mathbf{x}_t^*\}_{t=1}^T) + \frac{||\mathbf{x}_1 - \mathbf{x}_0^*||^2}{2\alpha} + |\bar{\lambda}| V(\{g_t\}_{t=1}^T) \\
& + \mu M^2 T + \frac{\alpha G^2 T}{2} - \frac{||\boldsymbol{\lambda}_{T+2}||^2}{2\mu} + \frac{||\boldsymbol{\lambda}_2||^2}{2\mu} \\
\overset{(a)}{\leq} & \frac{R}{\alpha} V(\{\mathbf{x}_t^*\}_{t=1}^T) + \frac{R^2}{2\alpha} + |\bar{\lambda}| V(\{\bar{\mathbf{g}}_t\}_{t=1}^T) + \mu M^2 T + \frac{\alpha G^2 T}{2} + \frac{\mu M^2}{2},
\end{aligned} \tag{A.28}$$

where (a) holds since $||\mathbf{x}_1 - \mathbf{x}_0^*|| \leq R$, $|\boldsymbol{\lambda}_{T+2}|^2 \geq 0$, $\boldsymbol{\lambda}_1 = \mathbf{0}$ and $||\boldsymbol{\lambda}_2||^2 = |\lambda_{1,2}|^2 + |\lambda_{2,2}|^2 = (\max(0, \mu g_1(\mathbf{x}_1)))^2 + (\max(0, -\mu g_1(\mathbf{x}_1)))^2 \leq \mu^2 (g_1(\mathbf{x}_1))^2 \leq \mu^2 M^2$.

A.4.1 Upper bounds

Let $V^+ = \max\left\{V(\{\mathbf{x}_t^*\}_{t=1}^T), V(\{\bar{\mathbf{g}}_t\}_{t=1}^T)\right\}$. If the primal and dual stepsizes are chosen as $\alpha = \mu = \sqrt{\frac{V^+}{T}}$. Then

$$\begin{aligned}
|\bar{\lambda}| &:= 2\mu M + \frac{2GR + \frac{R^2}{2\alpha} + \mu M^2}{\epsilon - \bar{V}(g)} = \mathcal{O}\left(\sqrt{\frac{T}{V^+}}\right), \\
|\mathcal{F}_T| &\leq \frac{|\bar{\lambda}|}{\mu} = \mathcal{O}(\frac{T}{V^+}) = \mathcal{O}\left(\max\left\{\frac{T}{V(\{\mathbf{x}_t^*\}_{t=1}^T)}, \frac{T}{V(\{\bar{\mathbf{g}}_t\}_{t=1}^T)}\right\}\right).
\end{aligned}$$

Replace the $\bar{\lambda}$ in Equation 4.17

$$\begin{aligned}
\mathcal{R}_T &\leq \frac{R}{\alpha}V(\{\mathbf{x}_t^*\}_{t=1}^T) + \frac{R^2}{2\alpha} + |\bar{\lambda}|V(\{\bar{\mathbf{g}}_t\}_{t=1}^T) + \mu M^2 T + \frac{\alpha G^2 T}{2} + \frac{\mu M^2}{2} \\
&\leq \frac{R}{\alpha}V(\{\mathbf{x}_t^*\}_{t=1}^T) + \frac{R^2}{2\alpha} + (2\mu M + \frac{2GR + \frac{R^2}{2\alpha} + \mu M^2}{\epsilon - \bar{V}(g)})V(\{\bar{\mathbf{g}}_t\}_{t=1}^T) \\
&+ \mu M^2 T + \frac{\alpha G^2 T}{2} + \frac{\mu M^2}{2} \\
&\leq R\sqrt{V^+ T} + \frac{R^2}{2}\sqrt{\frac{T}{V^+}} + \mathcal{O}(\sqrt{\frac{T}{V^+}})V^+ + \sqrt{V^+ T}M^2 + \sqrt{V^+ T}\frac{G^2}{2} + \sqrt{\frac{V^+}{T}}\frac{M^2}{2} \\
&:= \mathcal{O}(\sqrt{V^+ T}) = \mathcal{O}\left(\max\left\{\sqrt{V(\{\mathbf{x}_t^*\}_{t=1}^T)T}, \sqrt{V(\{\bar{\mathbf{g}}_t\}_{t=1}^T)T}\right\}\right).
\end{aligned}$$

A.5 Upperbound of $\mathcal{R}_T^{\text{off}}$

Based on the definition of $\mathcal{R}_T^{\text{off}}$, we have

$$\begin{aligned}
\mathcal{R}_T^{\text{off}} =& \sum_{t=1}^T f_t(\mathbf{x}_t) - \sum_{t=1}^T f_t(\mathbf{x}_t^{\text{off}}) \\
=& \sum_{t=1}^T f_t(\mathbf{x}_t) - \sum_{t=1}^T f_t(\mathbf{x}_t^*) + \sum_{t=1}^T f_t(\mathbf{x}_t^*) - \sum_{t=1}^T f_t(\mathbf{x}_t^{\text{off}}) \\
=& \mathcal{R}_T + \sum_{t=1}^T f_t(\mathbf{x}_t^*) - \sum_{t=1}^T f_t(\mathbf{x}_t^{\text{off}}).
\end{aligned} \tag{A.29}$$

The term $\sum_{t=1}^T f_t(\mathbf{x}_t^*) - \sum_{t=1}^T f_t(\mathbf{x}_t^{\text{off}})$ represents the difference between the performance of instantaneous minimizers and the offline optimal solutions.

Consider the dual function of the instantaneous primal problem (4.7), which can be expressed

by minimizing the online Lagrangian in (4.8) at time t

$$\begin{aligned}\mathcal{J}_t(\boldsymbol{\lambda}) &\coloneqq \min_{\mathbf{x}_t \in \mathcal{X}_t} \mathcal{L}_t(\mathbf{x}_t, \boldsymbol{\lambda}) \\ &= \min_{\mathbf{x}_t \in \mathcal{X}_t} f_t(\mathbf{x}_t) + \boldsymbol{\lambda}_t^\top \bar{\mathbf{g}}_{t-1} \\ &= \min_{\mathbf{x}_t \in \mathcal{X}_t} f_t(\mathbf{x}_t) + \lambda_1 g_t(\mathbf{x}_t) + \lambda_2(-g_t(\mathbf{x}_t)). \end{aligned} \tag{A.30}$$

The dual function of (4.6) over the entire horizon is

$$\begin{aligned}\mathcal{J}(\boldsymbol{\lambda}) &\coloneqq \min_{\mathbf{x}_t \in \mathcal{X}_t, \forall t} \sum_{t=1}^{T} \mathcal{L}_t(\mathbf{x}_t, \boldsymbol{\lambda}) \\ &= \sum_{t=1}^{T} \min_{\mathbf{x}_t \in \mathcal{X}_t} \mathcal{L}_t(\mathbf{x}_t, \boldsymbol{\lambda}) = \sum_{t=1}^{T} \mathcal{J}_t(\boldsymbol{\lambda}). \end{aligned} \tag{A.31}$$

Since the problems (4.7) and (4.6) are both convex, Assumption 4 implies that strong duality holds. Accordingly, $\sum_{t=1}^{T} f_t(\mathbf{x}_t^*) - \sum_{t=1}^{T} f_t(\mathbf{x}_t^{\text{off}})$ in Equation A.29 can be written as

$$\sum_{t=1}^{T} f_t(\mathbf{x}_t^*) - \sum_{t=1}^{T} f_t(\mathbf{x}_t^{\text{off}}) = \sum_{t=1}^{T} \max_{\boldsymbol{\lambda}_t} \mathcal{J}_t(\boldsymbol{\lambda}_t) - \max_{\boldsymbol{\lambda}} \sum_{t=1}^{T} \mathcal{J}_t(\boldsymbol{\lambda}). \tag{A.32}$$

Then we establish $\sum_{t=1}^{T} f_t(\mathbf{x}_t^*) - \sum_{t=1}^{T} f_t(\mathbf{x}_t^{\text{off}})$ can be bounded by the variation of the dual function.

Define the variation of the dual function (A.30) from time t to $t+1$ as

$$V(\mathcal{J}_t) \coloneqq \max_{\boldsymbol{\lambda}} |\mathcal{J}_{t+1}(\boldsymbol{\lambda}) - \mathcal{J}_t(\boldsymbol{\lambda})|, \tag{A.33}$$

and the total variation over the time horizon T as

$$V(\{\mathcal{J}_t\}_{t=1}^{T}) \coloneqq \sum_{t}^{T} V(\mathcal{J}_t). \tag{A.34}$$

Let $\tilde{t}$ denote any time slot in $\{1, \cdots, T\}$, we have

$$\begin{aligned}
&\sum_{t=1}^{T} \max_{\boldsymbol{\lambda}_t} \mathcal{J}_t(\boldsymbol{\lambda}_t) - \max_{\boldsymbol{\lambda}} \sum_{t=1}^{T} \mathcal{J}_t(\boldsymbol{\lambda}) \\
&\leq \sum_{t}^{T} (\mathcal{J}_t(\boldsymbol{\lambda}_t^*) - \mathcal{J}_t(\boldsymbol{\lambda}_{\tilde{t}}^*)) \leq T \max_{t} \{\mathcal{J}_t(\boldsymbol{\lambda}_t^*) - \mathcal{J}_t(\boldsymbol{\lambda}_{\tilde{t}}^*)\}. \qquad \text{(A.35)}
\end{aligned}$$

where $\boldsymbol{\lambda}_t^* \in \arg\max_{\boldsymbol{\lambda}} \mathcal{J}_t(\boldsymbol{\lambda})$ which is the instantaneous best solution for Equation A.30.

Then we will show that

$$\max_{t} \{\mathcal{J}_t(\boldsymbol{\lambda}_t^*) - \mathcal{J}_t(\boldsymbol{\lambda}_{\tilde{t}}^*)\} \leq 2V(\{\mathcal{J}_t\}_{t=1}^{T}).$$

Assume there exists a slot t_0 that $\mathcal{J}_{t_0}(\boldsymbol{\lambda}_{t_0}^*) - \mathcal{J}_{t_0}(\boldsymbol{\lambda}_{\tilde{t}}^*) > 2V(\{\mathcal{J}_t\}_{t=1}^{T})$. Then it implies that

$$\mathcal{J}_{\tilde{t}}(\boldsymbol{\lambda}_{\tilde{t}}^*) \overset{(a)}{\leq} \mathcal{J}_{t_0}(\boldsymbol{\lambda}_{\tilde{t}}^*) + V(\{\mathcal{J}_t\}_{t=1}^{T}) < \mathcal{J}_{t_0}(\boldsymbol{\lambda}_{t_0}^*) - V(\{\mathcal{J}_t\}_{t=1}^{T}) \overset{(b)}{\leq} \mathcal{J}_{\tilde{t}}(\boldsymbol{\lambda}_{t_0}^*), \quad \forall \tilde{t} \in T,$$

where (a) and (b) come from $\max_{t_1,t_2} |\mathcal{J}_{t_1}(\boldsymbol{\lambda}) - \mathcal{J}_{t_2}(\boldsymbol{\lambda})| \leq V(\{\mathcal{J}_t\}_{t=1}^{T})$ since $V(\{\mathcal{J}_t\}_{t=1}^{T})$ is the accumulated variation over T. $\mathcal{J}_{\tilde{t}}(\boldsymbol{\lambda}_{\tilde{t}}^*) \leq \mathcal{J}_{\tilde{t}}(\boldsymbol{\lambda}_{t_0}^*)$ contradict that $\boldsymbol{\lambda}_{\tilde{t}}^*$ is the maximizer of $\mathcal{J}_{\tilde{t}}(\boldsymbol{\lambda})$. Therefore we have $\max_t \{\mathcal{J}_t(\boldsymbol{\lambda}_t^*) - \mathcal{J}_t(\boldsymbol{\lambda}_{\tilde{t}}^*)\} \leq 2V(\{\mathcal{J}_t\}_{t=1}^{T})$. Substituting to Equation A.32 and Equation A.35 leads to

$$\sum_{t=1}^{T} f_t(\mathbf{x}_t^*) - \sum_{t=1}^{T} f_t(\mathbf{x}_t^{\text{off}}) \leq 2V(\{\mathcal{J}_t\}_{t=1}^{T}).$$

Replacing the term $\sum_{t=1}^{T} f_t(\mathbf{x}_t^*) - \sum_{t=1}^{T} f_t(\mathbf{x}_t^{\text{off}})$ in Equation A.29 leads to

$$\mathcal{R}_T^{\text{off}} = \mathcal{R}_T + \sum_{t=1}^{T} f_t(\mathbf{x}_t^*) - \sum_{t=1}^{T} f_t(\mathbf{x}_t^{\text{off}}) \leq \mathcal{R}_T + 2V(\{\mathcal{J}_t\}_{t=1}^{T}).$$

A.6 Flow constraint and Transmission utilization fee

The appropriate P2P market design should make it possible to have different prices for different transaction which implies that all market clients may account for preferences, different valuations of electricity and differentiated network charges through the market. Hence, the impact of each transaction on line flow constraints and cost are considered. Line flow constraints are added as a constraint of the objective function to model physical network in the energy trading,

$$-\rho_l^{\max} \leq \rho_l \leq \rho_l^{\max}, \tag{A.36}$$

where $\rho_l^{\max}$ is the maximum capacity of line l.

A method to specify the share of each transaction in the line flow constraint is employing Power transfer distribution factor (PTDF). PTDF is defined as "linear approximations of the sensitivities of the active power line flows with respect to various variables" [17]. PTDF is used to calculate flow of lines and indicate lines that used for power transfer in each individual transaction [12, 26]. PTDF for line l shown by ϕ_l^{ij} and indicates part of the generated energy by seller i, which is transferred to buyer j through line l. ϕ_l^{ij} can be obtained using

$$\phi_l^{ij} = \phi_l^i - \phi_l^j, \tag{A.37}$$

where ϕ_l^i, ϕ_l^j are injection shift factors (ISF) in line l for node i and j respectively. ISF is an approximation of the sensitivity matrix and quantifies the redistribution of power through each branch following a change in generation or load in a particular node. The ISF matrix is shown as $\Phi \in \mathbb{R}^{N_l \times N}$ where N_l denotes the number of lines. This matrix can be obtained from diagonal branch susceptance matrix (B'), branch to node incidence matrix (A) and reduced nodal susceptance matrix (C). The PTDF matrix depends on the network structure and is independent of the power flows through the network .

$$\Phi = B'AC^{-1}, \tag{A.38}$$

$$B' = diag[b_1, \ldots, b_{N_l}] \in \mathbb{R}^{N_l \times N_l}, \tag{A.39}$$

$$A = [a_1, a_2, \cdots, a_{N_l}]^\top \in \mathbb{R}^{N_l \times N}, \tag{A.40}$$

$$C = A^T B' A \in \mathbb{R}^{N \times N}. \tag{A.41}$$

In the matrix A, $a_l^\top \overset{e.g.}{=} [1, -1, 0, \cdots, 0]$ is the lth row, in which a line exists between the first and second node.

With PTDF, the power transfer distance between consumer j and producer i can be obtained as

$$\mathbf{D}^{ij} = \sum_{l \in N_l} \phi_l^{ij}, \quad \mathbf{D} \in \mathbb{R}^{N \times N}. \tag{A.42}$$

With traded energy decision ($\mathbf{X}_t$), the power flow in line l can be obtained by

$$\rho_l(\mathbf{X}_t) = \sum_{i \in w_t^p} \sum_{j \in w_t^c} \phi_l^{ij} \mathbf{x}_t^{ij}, \tag{A.43}$$

and the total line utilization fee can be calculated as

$$\sum_{i \in w_t^p} \sum_{j \in N} \gamma \mathbf{D}^{ij} \mathbf{X}_t^{ij} = - \sum_{j \in w_t^c} \sum_{i \in N} \gamma \mathbf{D}^{ji} \mathbf{X}_t^{ji}.$$

where w_t^p denotes the clients who are producers in time slot t and $\mathbf{X}_t^{ij} \geq 0, \forall i \in w_t^p, j \in N$, w_t^c denotes the clients who are consumers in time slot t and $\mathbf{X}_t^{ji} \geq 0, \forall j \in w_t^p, i \in N$.