University of California
Santa Barbara

Adaptive Algorithms for Dynamic Decision-Making: Bridging Online Learning and Non-Parametric Regression

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

Doctor of Philosophy
in
Computer Science

by

Dheeraj Baby

Committee in charge:

Professor Yu-Xiang Wang, Chair
Professor Daniel Lokshtanov
Professor Ramtin Pedarsani
Professor Ryan Tibshirani

March 2024

The Dissertation of Dheeraj Baby is approved.

Professor Daniel Lokshtanov

Professor Ramtin Pedarsani

Professor Ryan Tibshirani

Professor Yu-Xiang Wang, Committee Chair

March 2024

Adaptive Algorithms for Dynamic Decision-Making: Bridging Online Learning and Non-Parametric Regression

by

Dheeraj Baby

To the memory of my beloved father Baby.
To my loving mother Suseela.

Acknowledgements

I am most grateful to my advisor Professor Yu-Xiang Wang for shaping me as a researcher. His infinite wisdom and kindness are qualities I can only aspire to in my life. I remember the days when I was long stuck solving certain sub-problems, then after meeting him, went back to my lab with clear leads to tackle them. I greatly value the intellectual freedom he has granted me in working on hard research problems without being judged even if I fail to make progress. Such conducive research environment helped me to courageously pursue research ideas and make meaningful contributions to machine learning literature.

I am fortunate to have met my lab-mates Chong, Yuqing, Xuandong, Jianyu, Esha, Rachel and Ming. They have always been a source of inspiration to me both in research and life. I remember the BBQ parties, Hotpots and long walks through our beautiful UCSB campus I had with them where we reflected on both research and stuffs happening in our lives. I am grateful for my long-term roommate Nazmus and other friends such as Mert, Paul, Ishtiyaque and Momin. They were an integral part in making my campus life fun and exciting. I would also like to thank my friends Aman and Chaitanya from my undergraduate university. It is due to them I was inspired to pursue higher studies.

Last but not least, I thank my mom Suseela and my brother Dharmaj for the unfailing love, patience and support they have given me while pursuing my dreams even if it required me to move to the other part of the world.

Curriculum Vitæ
Dheeraj Baby

Education

2024	Ph.D. in Computer Science, University of California, Santa Barbara, USA
2016	B.Tech and M.Tech in Electrical Engineering, Indian Institute of Technology Madras, Chennai, India

Publications

NeurIPS 2023	*Online Label Shift: Optimal Dynamic Regret meets Practical Algorithms*, Proceedings of 37th Conference on Neural Information Processing Systems
AISTATS 2023	*Second Order Path Variationals in Non-Stationary Online Learning*, in Proceedings of 26th International Conference on Artificial Intelligence and Statistics
TMLR 2023	*Non-stationary Contextual Pricing with Safety Constraints*", in Journal of Transactions of Machine Learning Research
NeurIPS 2022	*Optimal Dynamic Regret in LQR Control*", in Proceedings of 36th Conference on Neural Information Processing Systems
AISTATS 2022	*Optimal Dynamic Regret in Proper Online Learning with Strongly Convex Losses and Beyond*", in Proceedings of 25th International Conference on Artificial Intelligence and Statistics
COLT 2021	*Optimal Dynamic Regret in Exp-Concave Online Learning*", in Proceedings of 34th Annual Conference On Learning Theory
AISTATS 2021	*An Optimal Reduction of TV-Denoising to Adaptive Online Learning*", in Proceedings of 24th International Conference on Artificial Intelligence and Statistics
NeurIPS 2020	*Adaptive Online Estimation of Piecewise Polynomial Trends*", in Proceedings of 34th Conference on Neural Information Processing Systems
NeurIPS 2019	*Online Forecasting of Total-Variation-bounded Sequences*", in Proceedings of 33rd Conference on Neural Information Processing Systems

Academic Service

Reviewer	ICML 19-21,23; JMLR 21-23; AISTATS 22-23; ICLR24

Professional Experience

AWS	Applied Scientist Intern, 2020 and 2023
Google	Student Researcher, 2022

Abstract

Adaptive Algorithms for Dynamic Decision-Making: Bridging Online Learning and Non-Parametric Regression

by

Dheeraj Baby

Making decisions in real-time by learning patterns in an online data-stream is an important problem in modern machine learning (ML). Applications that fall under this umbrella include domain adaptation, change-point detection, portfolio-optimization, optimally pricing airline tickets based on changing market features etc. The features of the environment where an ML model is deployed change from time to time. Designing decision-making algorithms that can quickly adapt to these environmental changes on-the-fly is a topic of great significance.

In this thesis, we will design and analyse information-theoretically optimal algorithms for online decision-making under non-stationarities. The presentation will also encompass the utilization of these algorithms across a wide spectrum of applications, spanning time series forecasting, dynamic pricing, non-parametric regression, LQR control and unsupervised domain adaptation.

A main challenge in the theoretical analysis of the algorithms is to exploit the curved geometry of loss functions while deriving fast dynamic regret rates. This is attained by connecting ideas from the domains of locally adaptive non-parametric regression and strongly adaptive online learning. These fields have been conventionally studied separately by researchers. In this thesis we provide new tools to bridge these two domains. A byproduct of this fusion are novel results that do not require observation models with stringent stochastic assumptions for non-parametric regression and online convex opti-

mization. Further, the developed algorithms are highly adaptive and do not require prior knowledge about the degree of non-stationarity in the environment. Our hope is that this thesis will inspire new collaborations between researchers from the communities of online learning and non-parametric regression.

List of Commonly Used Acronyms

BV	Bounded Variation
FLH	Follow the Leading History
FTL	Follow the Leader
GC	Geometric Cover
LQR	Linear Quadratic Regularator
ML	Machine Learning
OCO	Online Convex Optimization
OGD	Online Gradient Descent
ONS	Online Newton Step
SA	Strongly Adaptive
TSE	Total Squared Error
TV	Total Variation
WLOG	Without Loss of Generality

Contents

Curriculum Vitae **vi**

Abstract **viii**

1 Introduction **1**
1.1 Outline of the Thesis 2

Part I Theory and Algorithms under Stochastic Observation Model **8**

2 Online Forecasting of Total-Variation-bounded Sequences **9**
2.1 Setup, Assumptions and Contributions 10
2.2 Related Work 13
2.3 Main Results 16
2.4 Concluding Discussion 26

3 Adaptive Online Estimation of Piecewise Polynomial Trends **27**
3.1 Setup, Assumptions and Contributions 28
3.2 Main results 34
3.3 Extensions 39
3.4 Concluding Discussion 41

4 An Optimal Reduction of TV-Denoising to Adaptive Online Learning **43**
4.1 Setup, Assumptions and Contributions 47
4.2 Preliminaries 50
4.3 Main Results 53
4.4 Extensions 57
4.5 Experimental Results 58
4.6 Concluding Discussion 63

Part II Theory and Algorithms under Adversarial Observation Model 64

5 Optimal Dynamic Regret in Exp-Concave Online Learning 65
5.1 Performance guarantees for squared error losses 70
5.2 Performance guarantees for exp-concave losses 78
5.3 Concluding Discussion 82

6 Optimal Dynamic Regret in Proper Online Learning with Strongly Convex Losses and Beyond 83
6.1 Related Work 84
6.2 A gentle start: Squared loss games 86
6.3 Performance guarantees for strongly convex losses 95
6.4 Performance guarantees for exp-concave losses 100
6.5 Concluding Discussion 104

7 Second Order Path Variationals in Non-Stationary Online Learning 105
7.1 The Algorithm 108
7.2 Main Results 112
7.3 Concluding Discussion 122

Part III Applications 123

8 Online Unsupervised Domain Adaptation under Label Shift 124
8.1 Problem Setup 128
8.2 Unsupervised Online Label Shift 129
8.3 Supervised Online Label Shift 135
8.4 Experiments 138
8.5 Concluding Discussion 143

9 Non-Stationary Contextual Pricing with Safety Constraints 144
9.1 Introduction 144
9.2 Notations and Problem Setup 148
9.3 Main Results 153
9.4 Concluding Discussion 162

10 Optimal Dynamic Regret in LQR Control 163
10.1 Preliminaries 168
10.2 Non-stationary "mini-batch" linear regression 170
10.3 Instantiation for the LQR problem 177
10.4 Concluding Discussion 179

Part IV Appendix 181

A Supplementary Materials for Chapter 2 182
A.1 Additional Experiments . . . 182
A.2 Upper bounds of linear forecasters . . . 183
A.3 Proof of useful lemmas . . . 185
A.4 Proof of Theorem 1 . . . 191
A.5 Adaptive Optimality in Discrete Sobolev class . . . 196
A.6 Fast Computation . . . 197
A.7 Regret of AOMD . . . 198
A.8 Lower bound proof . . . 199
A.9 Optimality of linear forecasters in Discrete Sobolev class . . . 202

B Supplementary Materials for Chapter 3 204
B.1 Background . . . 204
B.2 Analysis . . . 208
B.3 Adapting to lower orders of k . . . 222
B.4 Problems with padding . . . 223
B.5 Technical Lemmas . . . 224

C Supplementary Materials for Chapter 4 233
C.1 Proofs of Technical Results . . . 233
C.2 Excluded details in Experimental section . . . 241

D Supplementary Materials for Chapter 5 248
D.1 Preliminaries . . . 248
D.2 Proofs for Section 5.1 . . . 249
D.3 Proofs for Section 5.2 . . . 257
D.4 Technical Lemmas . . . 279

E Supplementary Materials for Chapter 6 283
E.1 Proofs for Section 6.2 . . . 283
E.2 Proofs for Section 6.3 . . . 299
E.3 Proofs for Section 6.4 . . . 300
E.4 Reparametrization of certain polytopes to box . . . 321

F Supplementary Materials for Chapter 7 323
F.1 Analysis . . . 323
F.2 Proof of Proposition 65 . . . 365
F.3 Proof of Proposition 62 . . . 367

G Supplementary Materials for Chapter 8 **371**
G.1 Omitted proofs from Section 8.2 371
G.2 Design of low switching online regression algorithms 379
G.3 Omitted proofs from Section 8.3 381
G.4 More details on experiments 394
G.5 Additional Unsupervised Online Label Shift Experiments 400

H Supplementary Materials for Chapter 9 **405**
H.1 Detailed Proof 405

I Supplementary Materials for Chapter 10 **417**
I.1 Omitted Proofs 417

Bibliography **431**

Chapter 1

Introduction

Online learning is a powerful machine learning (ML) paradigm for sequential decision making. In this framework, a learning agent is faced with a data stream. At each round of the data stream the agent makes a decision / prediction solely based on the historical data observed so far. The main goal in online learning is to design algorithms for the agent that can quickly learn patterns present in the data stream and leverage them to make high quality predictions. Since a prediction needs to be made in real-time at each round, online algorithms are attractive only if it has low per-round computational complexity. On the other-hand the patterns that emerge in the stream are often transient. The agent needs to quickly detect and learn them for maximizing statistical efficiency of the predictions. These two challenges make the task of designing online learning algorithms highly non-trivial.

Numerous real-world ML applications can be cast into the framework of online learning. Some notable examples include forecasting weather / stock market trends, online portfolio optimization, pricing of airline tickets based on changing market features, adaptively controlling the oxygen flow rate in ICU inhalators based on real-time vital signals of the patient and recommending products in a retail website based on the constantly evolving preferences of the customer.

The main focus of this thesis is to develop principled algorithms for online learning under a non-stationary data stream. i.e, the data generating distribution at each round in the stream can drift across time. Rigorous proofs for the information-theoretic optimality of the developed algorithms will be presented. Further we will also feature the utilization of these algorithms for solving various real-world applications. Three main factors were kept in mind while developing these algorithms: 1) Low computational complexity; 2) Minimal assumptions on the nature of distribution shift across the data stream and 3) Attaining statistical efficiency without asking the user to tune hyper-parameters calibrated to the level of non-stationarity in the data stream.

1.1 Outline of the Thesis

Part I: Theory and Algorithms under Stochastic Observation Model

In this part we develop online forecasting algorithms where the observations are assumed to be noisy realisations of ground truth. The noise is taken to be distributed iid.

- In Chapter 2, we begin with an online denoising problem that inspired all the subsequent chapters of this thesis. In this problem, we sequentially observe noisy realisations of an unknown ground truth sequence of bounded total variation (TV). The noise is assumed to be sampled iid from a sub-gaussian distribution. The task is to estimate the ground truth sequence value at each round in an online manner. The quality of the predictions are measured using total squared error (TSE). The offline version of this problem is well studied in the locally adaptive non-parametric regression literature [1, 2, 3]. However, the online problem is harder, because unlike the offline version, the learner has access only to the past observations. Nevertheless we show that for the online estimation problem, one can attain a minimax TSE rate that is of similar order as that of the offline problem. A computationally efficient and minimax optimal online algorithm is designed based on a novel change point detection scheme that is constructed via building upon the theory of wavelet based non-parametric estimation [2].

- Chapter 3 is a generalisation of Chapter 2 where we estimate ground truth sequences with bounded higher order TV in an online fashion. Sequences with continuous piece-wise linear (or more generally polynomial) trends that are ubiquitously seen in time series forecasting problems [4] are examples of ground truths with bounded higher order TV. Sharp minimax estimation rates similar to what is seen in the offline statistical non-parametric estimation of higher order TV bounded sequences are obtained. The key ingredient in algorithm design is to adaptively restart online linear regression with monomial covariates of time. The restart rule is developed based on change point detection via statistics computed from denoised higher order wavelet coefficients.

- In Chapter 4, we revisit the same online estimation problem of Chapter 2 and rethink it via the lens of strongly adaptive (SA) online learning. SA algorithms have the nice property that their static regret in any continuous time window is controlled. Informally (under squared errror loss), this means that if we consider any time interval, the TSE of the online learner is less than or comparable to the TSE incurred by the mean of the observations within that interval. We show that SA algorithms can also attain minimax rates of estimating TV bounded sequences from noisy observations. This is the first time a connection between the fields of non-parametric regression and strongly adaptive online learning is established in literature. We observe that the developed algorithm leads to better performance than existing methods for the task of forecasting COVID-19 hospitalizations trends based on real-world CDC data. Further the SA algorithm also outperforms the algorithm developed in Chapter 2 in various simulation studies.

Part II: Theory and Algorithms under Adversarial Observation Model

In this part we develop online forecasting algorithms in a setting where the stringent stochastic assumptions of Part I are lifted. We allow for observations that can be perturbed from the ground truth in a fully adversarial manner. Consequently, the developments of this part can handle arbitrarily correlated observations and even allow for non-stationarities in the noise distribution. This agenda is realised by connecting ideas from offline convex optimization to online learning.

- Chapter 5 takes a significant step towards generalizing the observation model studied in Chapter 2. Specifically, we lift all stochastic assumptions on the noise and allow for a fully adversarial perturbation model. We adopt the performance measure of dynamic regret where the total loss of the agent is compared against the loss of a sequence of evolving decisions in hindsight. We show a surprising result that SA algorithms similar to what is used in Chapter 4 can lead to minimax optimal dynamic regret rates in the fully adversarial setting. This result subsumes the results in Chapters 2 and 4. Further, the result also allows one to perform non-parametric regression without imposing conventional iid based noise assumptions. We further generalise the setting to online convex optimization (OCO) and show that SA algorithms can yield optimal dynamic regret rates under the general exp-concave and gradient smooth family of losses. Squared loss, logistic and linear regression losses are some examples of popular exp-concave and gradient smooth losses. These losses stand out from usual convex losses through their additional curved geometry. The problem of attaining minimax optimal dynamic regret rate under exp-concave losses was long-standing in the literature dating back at-least to the 2003 work of [5]. Chapter 5 gives a definitive answer to this open problem.

- In Chapter 6, we address a drawback of Chapter 5. In the setting of *proper* OCO, the decisions of the learner need to obey certain prespecified physical constraints. Such constraints are modelled by requiring the decisions made by the online learner to belong to a given convex set. The algorithms in Chapter 5 can potentially violate this constraint making them only suitable for *improper* OCO. In Chapter 6, we develop new SA algorithms for proper OCO under strongly convex losses. Further results for proper OCO under exp-concave losses under a box constrained decision set (L_∞ ball) are also provided. Moreover, we relax the restriction of gradient smoothness of the losses from Chapter 5.

- Chapter 7 proceeds in a similar vein as in Chapter 5. We develop algorithms for improper OCO under exp-concave and gradient smooth losses with dynamic regret rates characterized by the second order TV of the comparator. The developed algorithm can simultaneously guarantee a rate that is the minimum of optimal dynamic regret rates that are measured using the number of change points in the comparator, first order TV and second order TV of the comparator sequence.

Part III: Applications

In this part, we demonstrate the applicability of the theory and algorithms developed thus far to three problems: unsupervised domain adaptation, dynamic pricing and LQR controller design.

- A standard assumption in learning theory is that the test distribution is same as the training distribution. However, this assumption can be violated in practical settings where the characteristics of the test environment can slowly drift apart from the training environment as time goes on. Consequently there is a real need for strategies that can slowly adapt an ML model to the changing test distribution. Chapter 8 studies the problem of how to systematically adapt a given probabilistic classifier trained on an offline data set to an online test data stream without seeing the labels. The test distribution at any time-stamp is assumed to be a label shifted version of the offline training distribution. An important feature of our solution that differentiates from prior work is that the proposed algorithm can support probabilistic classifiers without the need to impose convexity restrictions on the losses.

- Chapter 9 pushes the results of Chapter 6 one step further. We provide an algorithmic way to attain optimal dynamic regret rates under *proper* OCO with exp-concave losses that belong to the generalized linear family class. The developed techniques are applied to solve non-stationary dynamic pricing. Dynamic pricing studies the problem of optimally allocating prices to commodities based on changing market features and customer's internal evaluation of the products. Proper learning is important in dynamic pricing to satisfy various fairness and legal constraints when allocating prices.

- Chapter 10 studies the problem of non-stationary LQR controller design. The controller is expected to stabilize / navigate a linear dynamical system under environmental perturbations. Certain constraints are required on the actions taken by the controller to ensure desirable features of the system such as stability and less battery usage thereby necessitating *proper* online learning.

Algorithms that are featured across various chapters have the property of being adaptively minimax optimal: i.e, statistical optimality is attained without asking the user to tune hyper-parameters that are calibrated to the level of the non-staionarity in the world.

We hope that such algorithms can fuel further ML research and downstream applications without imposing restrictive modelling assumptions that are often unverifiable in practice.

Part I

Theory and Algorithms under Stochastic Observation Model

Chapter 2

Online Forecasting of Total-Variation-bounded Sequences

Nonparametric regression is a fundamental class of problems that has been studied for more than half a century in statistics and machine learning [6, 7, 8, 2, 9, 10, 11]. It solves the following problem:

- Let $y_i = f(u_i) + \text{Noise}$ for $i = 1, ..., n$. How can we estimate a function f using data points $(u_1, y_1), ..., (u_n, y_n)$ and the knowledge that f belongs to a function class $\mathcal{F}$?

Function class $\mathcal{F}$ typically imposes only weak regularity assumptions on the function f such as boundedness and smoothness, which makes nonparametric regression widely applicable to many real-life applications especially those with unknown physical processes.

A recent and successful class of nonparametric regression technique called trend filtering [12, 4, 3, 13] was shown to have the property of *local adaptivity* [14] in both theory and practice. We say a nonparametric regression technique is *locally adaptive* if it can cater to local differences in smoothness, hence allowing more accurate estimation of functions with varying smoothness and abrupt changes. For example, for functions with bounded total variation (when $\mathcal{F}$ is a total variation class), standard nonparametric regression techniques such as kernel smoothing and smoothing splines have a mean square error (MSE) of $O(n^{-1/2})$ while trend filtering has the optimal $O(n^{-2/3})$.

Trend filtering is, however, a batch learning algorithm where one observes the entire dataset ahead of the time and makes inference about the past. This makes it inapplicable to the many time series problems that motivate the study of trend filtering in the first place [4]. These include influenza forecasting, inventory planning, economic policy-making, financial market prediction and so on. In particular, it is unclear whether the advantage of trend filtering methods in estimating functions with heterogeneous smoothness (e.g., sharp changes) would carry over to the online forecasting setting. The focus of this work is in developing theory and algorithms for locally adaptive online forecasting which predicts the immediate future value of a function with heterogeneous smoothness using only noisy observations from the past.

2.1 Setup, Assumptions and Contributions

2.1.1 Problem Setup

We propose a model for nonparametric online forecasting as described in Figure 2.1. This model can be re-framed in the language of the online convex optimization model with three differences.

1. We consider only quadratic loss functions of the form $\ell_t(x) = (x - \theta_t)^2$.
2. The learner receives independent *noisy* gradient feedback, rather than the exact gradient.

1. Fix action time intervals $1, 2, ..., n$
2. The player declares a forecasting strategy $\mathcal{A}_i : \mathbb{R}^{i-1} \to \mathbb{R}$ for $i = 1, ..., n$.
3. An adversary chooses a sequence $\theta_{1:n} = [\theta_1, \theta_2, \dots, \theta_n]^T \in \mathbb{R}^n$.
4. For every time point $i = 1, ..., n$:
   (a) We play $x_i = \mathcal{A}_i(y_1, ..., y_{i-1})$.
   (b) We receive a feedback $y_i = \theta_i + Z_i$, where Z_i is a zero-mean, independent subgaussian noise.
5. At the end, the player suffers a cumulative error $\sum_{i=1}^{n} \big(x_i - \theta_i\big)^2$.

Figure 2.1: *Nonparametric online forecasting model. The focus of the proposed work is to design a forecasting strategy that minimizes the expected cumulative square error. Note that the problem depends a lot on the choice of the sequence* θ_i. *Our primary interest is on sequences with bounded total variation (TV) so that* $\sum_{i=2}^{n} |\theta_i - \theta_{i-1}| \le C_n$, *but we will also talk about the adaptivity of our method to easier problems such as forecasting Sobolev and Holder functions.*

3. The criterion of interest is redefined as the *dynamic regret* [5, 15]:

$$R_{\text{dynamic}}(\mathcal{A}, \ell_{1:n}) := \mathbb{E}\left[\sum_{t=1}^{n} \ell_t(x_t)\right] - \sum_{t=1}^{n} \inf_{x_t} \ell_t(x_t). \tag{2.1}$$

The new criterion is called a dynamic regret because we are now comparing to a stronger dynamic baseline that chooses an optimal x in every round. Of course in general, the dynamic regret will be linear in n [16]. To make the problem non-trivial, we restrict our attention to sequences of $\ell_1, ..., \ell_n$ that are *regular*, which makes it possible to design algorithms with *sublinear* dynamic regret. In particular, we borrow ideas from the nonparametric regression literature and consider sequences $[\theta_1, ..., \theta_n]$ that are discretizations of functions in the continuous domain. Regularity assumptions emerge naturally as we consider canonical functions classes such as the Holder class, Sobolev class and Total Variation classes (see eg, for a review [17]).

2.1.2 Assumptions

We consolidate all the assumptions used in this work and provide necessary justifications for them.

- (A1) The time horizon for the online learner is known to be n.
- (A2) The parameter σ^2 of subgaussian noise in the observations is known.
- (A3) The ground truth denoted by $\theta_{1:n} = [\theta_1, ..., \theta_n]^T$ has its total variation bounded by some positive C_n, i.e., we take $\mathcal{F}$ to be the total variation class $\mathrm{TV}(C_n) := \{\theta_{1:n} \in \mathbb{R}^n : \|D\theta_{1:n}\|_1 \leq C_n\}$ where D is the discrete difference operator. Here $D\theta_{1:n} = [\theta_2 - \theta_1, \ldots, \theta_n - \theta_{n-1}]^T$.
- (A4) $|\theta_1| \leq U$.

The knowledge of σ^2 in assumption (A2) is primarily used to get the optimal dependence of σ in minimax rate. This assumption can be relaxed in practice by using the Median Absolute Deviation estimator as described in Section 7.5 of [18] to estimate σ^2 robustly. Assumption (A3) features a samples from a large class of functions with spatially inhomogeneous degree of smoothness. The functions residing in this class need not even be continuous. Our goal is to propose a policy that is locally adaptive whose empirical mean squared error converges at the minimax rate for this function class. We stress that we do *not* assume that the learner knows C_n. The problem is open and nontrivial even when C_n is known. Assumption (A4) is very mild as it puts restriction only to the first value of the sequence. This assumption controls the inevitable prediction error for the first point in the sequence.

2.1.3 Our Results

The major contributions of this chapter are summarized below.

- It is known that the minimax MSE for *smoothing* sequences in the TV class is $\tilde{\Omega}(n^{-2/3})$. This implies a lowerbound of $\tilde{\Omega}(n^{1/3})$ for the dynamic regret in our setting. We present a policy Arrows (**A**daptive **R**estarting **R**ule for **O**nline averaging using **W**avelet **S**hrinkage) with a nearly minimax dynamic regret $\tilde{O}(n^{1/3})$ and a run-time complexity of $O(n \log n)$.

- We show that a class of forecasting strategies — including the popular Online Gradient Descent (OGD) with fixed restarts [15], moving averages (MA) [19] — are fundamentally limited by $\tilde{\Omega}(\sqrt{n})$ regret.

- We also provide a more refined lower bound that characterized the dependence of U, C_n and σ, which certifies the adaptive optimality of Arrows in all regimes. The bound also reveals a subtle price to pay when we move from the smoothing problem to the forecasting problem, which indicates the separation of the two problems when $C_n/\sigma \gg n^{1/4}$, a regime where the forecasting problem is *strictly* harder (See Figure 2.3).

- Lastly, we consider forecasting sequences in Sobolev classes and Holder classes and establish that Arrows can automatically *adapt* to the optimal regret of these *simpler* function classes as well, while OGD and MA cannot, unless we change their tuning parameter (to behave suboptimally on the TV class).

2.2 Related Work

The topic of this chapter sits well in between two amazing bodies of literature: non-parametric regression and online learning. Our results therefore contribute to both fields and hopefully will inspire more interplay between the two communities. Throughout this chapter when we refer $\tilde{O}(n^{1/3})$ as the optimal regret, we assume the parameters of the problem are such that it is acheivable (see Figure 2.3).

Nonparametric regression. As we mentioned before, our problem — online nonparametric forecasting — is motivated by the idea of using locally adaptive nonparametric regression for time series forecasting [14, 4, 3]. It is more challenging than standard nonparametric regression because we do not have access to the data in the future. While our proof techniques make use of several components (e.g., universal shrinkage) from the seminal work in wavelet smoothing [20, 2], the way we use them to construct and analyze our algorithm is new and more generally applicable for converting non-parametric regression methods to forecasting methods.

Adaptive Online Learning. Our problem is also connected to a growing literature on adaptive online learning which aims at matching the performance of a stronger time-varying baseline [5, 21, 15, 22, 16, 23, 24, 25, 26, 27, 28]. Many of these settings are highly general and we can apply their algorithms directly to our problem, but to the best of our knowledge, none of them achieves the optimal $\tilde{O}(n^{1/3})$ dynamic regret.

In the remainder of this section, we focus our discussion on how to apply the regret bounds in non-stationary stochastic optimization [15, 22] to our problem

Regret from Non-Stationary Stochastic Optimization The problem of non-stationary stochastic optimization is more general than our model because instead of considering only the quadratic functions, $\ell_t(x) = (x - \theta_t)^2$, they work with the more general class of strongly convex functions and general convex functions. They also consider both noisy gradient feedbacks (stochastic first order oracle) and noisy function value feedbacks (stochastic zeroth order oracle).

In particular, [15] define a quantity V_n which captures the total amount of "variation" of the functions $\ell_{1:n}$ using $V_n := \sum_{i=1}^{n-1} \|\ell_{i+1} - \ell_i\|_\infty$. [1] [22] generalize the notion to

[1]The V_n definition in [15] for strongly convex functions are defined a bit differently, the $\|\cdot\|_\infty$ is taken

$V_n(p,q) := \left(\sum_{i=1}^{n-1} \|\ell_{i+1} - \ell_i\|_p^q\right)^{1/q}$ for any $1 \le p, q \le +\infty$ where $\|\cdot\|_p := (\int |\cdot(x)|^p dx)^{1/p}$ is the standard L_p norm for functions[2]. Table 2.1 summarizes the known results under the non-stationary stochastic optimization setting.

Table 2.1: Summary of known minimax dynamic regret in the non-stationary stochastic optimization model. Note that the choice of q does not affect the minimax rate in any way, but the choice of p does. "-" indicates that the no upper or lower bounds are known for that setting.

	Noisy gradient feedback		Noisy function value feedback	
Assumptions on $\ell_{1:n}$	$p = +\infty$	$1 \le p < +\infty$	$p = +\infty$	$1 \le p < +\infty$
Convex & Lipschitz	$\Theta(n^{2/3}V_n^{1/3})$	$O(n^{\frac{2p+d}{3p+d}} V_n(p,q)^{\frac{p}{3p+d}})$	-	-
Strongly convex & Smooth	$\Theta(n^{1/2}V_n^{1/2})$	$\Theta(n^{\frac{2p+d}{4p+d}} V_n(p,q)^{\frac{2p}{4p+d}})$	$\Theta(n^{2/3}V_n^{1/3})$	$\Theta(n^{\frac{4p+d}{6p+d}} V_n(p,q)^{\frac{2p}{6p+d}})$

Our assumption on the underlying trend $\theta_{1:n} \in \mathcal{F}$ can be used to construct an upper bound of this quantity of variation V_n or $V_n(p,q)$. As a result, the algorithms in non-stationary stochastic optimization and their dynamic regret bounds in Table 2.1 will apply to our problem (modulo additional restrictions on bounded domain). However, our preliminary investigation suggests that this direct reduction does *not*, in general, lead to optimal algorithms. We illustrate this observation in the following example.

Example 1. Let $\mathcal{F}$ be the set of all bounded sequences in the total variation class $TV(1)$. It can be worked out that $V_n(p,q) = O(1)$ for all p, q. Therefore the smallest regret from [15, 22] is obtained by taking $p \to +\infty$, which gives us a regret of $O(n^{1/2})$. Note that we expect the optimal regret to be $\tilde{O}(n^{1/3})$ according to the theory of locally adaptive nonparametric regression.

In Example 1, we have demonstrated that one cannot achieve the optimal dynamic regret using known results in non-stationary stochastic optimization. We show in section

over the convex hull of minimizers. This creates some subtle confusions regarding our results which we explain in details in Appendix A.8.

[2] We define $V_n(p,q)$ to be a factor of $n^{-1/q}$ times bigger than the original scaling presented in [22] so the results become comparable to that of [15].

2.3.1 that "Restarting OGD" algorithm has a fundamental lower bound of $\tilde{\Omega}(\sqrt{n})$ on dynamic regret in the TV class.

Online nonparametric regression. As we finalize our manuscript, it comes to our attention that our problem of interest in Figure 2.1 can be cast as a special case of the "online nonparametric regression" problem [29, 30]. The general result of [29] implies the *existence* of an algorithm that enjoys a $\tilde{O}(n^{1/3})$ regret for the TV class without explicitly constructing one, which shows that $n^{1/3}$ is the minimax rate when $C_n = O(1)$. To the best of our knowledge, our proposed algorithm remains the first *polynomial time* algorithm with $\tilde{O}(n^{1/3})$ regret and our results reveal more precise (optimal) upper and lower bounds on all parameters of the problem (see Section 2.3.4).

2.3 Main Results

We are now ready to present our main results.

2.3.1 Limitations of Linear Forecasters

Restarting OGD as discussed in Example 1, fails to achieve the optimal regret in our setting. A curious question to ask is whether it is the algorithm itself that fails or it is an artifact of a potentially suboptimal regret analysis. To answer this, let's consider the class of linear forecasters — estimators that outputs a fixed linear transformation of the observations $y_{1:n}$. The following preliminary result shows that Restarting OGD is a linear forecaster . By the results of [2], linear smoothers are fundamentally limited in their ability to estimate functions with heterogeneous smoothness. Since forecasting is harder than smoothing, this limitation gets directly translated to the setting of linear forecasters.

Proposition 2. *Online gradient descent with a fixed restart schedule is a linear forecaster. Therefore, it has a dynamic regret of at least* $\tilde{\Omega}(\sqrt{n})$.

Proof. First, observe that the stochastic gradient is of form $2(x_t - y_t)$ where x_t is what the agent played at time t and y_t is the noisy observation θ_t+Independent noise. By the online gradient descent strategy with the fixed restart interval and an inductive argument, x_t is a linear combination of $y_1, ..., y_{t-1}$ for any t. Therefore, the entire vector of predictions $x_{1:t}$ is a fixed linear transformation of $y_{1:t-1}$. The fundamental lower bound for linear smoothers from [2] implies that this algorithm will have a regret of at least $\tilde{\Omega}(\sqrt{n})$. □

The proposition implies that we will need fundamentally new *nonlinear* algorithmic components to achieve the optimal $O(n^{1/3})$ regret, if it is achievable at all!

2.3.2 Policy

In this section, we present our policy Arrows (Adaptive Restarting Rule for Online averaging using Wavelet Shrinkage). The following notations are introduced for describing the algorithm.

- t_h denotes start time of the current bin and t be the current time point.
- $\bar{y}_{t_h:t}$ denotes the average of the y values for time steps indexed from t_h to t.
- $pad_0(y_{t_h}, ..., y_t)$ denotes the vector $(y_{t_h} - \bar{y}_{t_h:t}, ..., y_t - \bar{y}_{t_h:t})^T$ zero-padded at the end till its length is a power of 2. $i.e$, a re-centered and padded version of observations.
- $T(x)$ where x is a sequence of values, denotes the element-wise soft thresholding of the sequence with threshold $\sigma\sqrt{\beta \log(n)}$
- H denotes the orthogonal discrete Haar wavelet transform matrix of proper dimensions

- Let $Hx = \alpha = [\alpha_1, \alpha_2, ..., \alpha_k]^T$ where k being a power of 2 is the length of x. Then the vector $[\alpha_2, ..., \alpha_k]^T$ can be viewed as a concatenation of $\log_2 k$ contiguous blocks represented by $\alpha[l], l = 0, ..., \log_2(k) - 1$. Each block $\alpha[l]$ at level l contains 2^l coefficients.

ARROWS: inputs - observed y values, time horizon n, std deviation σ, $\delta \in (0, 1]$, a hyper-parameter $\beta > 24$

1. Initialize $t_h = 1$, $newBin = 1$, $y_0 = 0$
2. For $t = 1$ to n:
   (a) If $newBin == 1$, predict $x_t^{t_h} = y_{t-1}$, else predict $x_t^{t_h} = \bar{y}_{t_h:t-1}$
   (b) set $newBin = 0$, observe y_t and suffer loss $(x_t^{t_h} - \theta_t)^2$
   (c) Let $\tilde{y} = pad_0(y_{t_h}, ..., y_t)$ and k be the padded length.
   (d) Let $\hat{\alpha}(t_h : t) = T(H\tilde{y})$
   (e) Restart Rule: If $\frac{1}{\sqrt{k}} \sum_{l=0}^{\log_2(k)-1} 2^{l/2} \|\hat{\alpha}(t_h : t)[l]\|_1 > \frac{\sigma}{\sqrt{k}}$ then
      i. set $newBin = 1$
      ii. set $t_h = t + 1$

Our policy is the byproduct of following question: How can one lift a batch estimator that is minimax over the TV class to a minimax online algorithm?

Restarting OGD when applied to our setting with squared error losses reduces to partitioning the duration of game into fixed size chunks and outputting online averages. As described in Section 2.3.1, this leads to suboptimal regret. However, the notion of averaging is still a good idea to keep. If within a time interval, the Total Variation (TV) is adequately small, then outputting sample averages is reasonable for minimizing the cumulative squared error. Once we encounter a bump in the variation, a good strategy is to restart the averaging procedure. Thus we need to adaptively detect intervals with low TV. For accomplishing this, we communicate with an oracle estimator whose output can be used to construct a lowerbound of TV within an interval. The decision to restart online averaging is based on the estimate of TV computed using this oracle. Such a decision rule introduces non-linearity and hence breaks free of the suboptimal world of linear forecasters.

The oracle estimator we consider here is a slightly modified version of the soft thresholding estimator from [31]. We capture the high level intuition behind steps (d) and (e) as follows. Computation of Haar coefficients involves smoothing adjacent regions of a signal and taking difference between them. So we can expect to construct a lowerbound of the total variation $\|D\theta_{1:n}\|_1$ from these coeffcients. The extra thresholding step $T(.)$ in (d) is done to denoise the Haar coefficients computed from noisy data. In step (e), a weighted L1 norm of denoised coefficients is used to lowerbound the total variation of the true signal. The multiplicative factors $2^{l/2}$ are introduced to make the lowerbound tighter. We restart online averaging once we detect a large enough variation. The first coefficient $\hat{\alpha}(t_h : t)_1$ is zero due to the re-centering caused by pad_0 operation. The hyperparameter β controls the degree to which we shrink the noisy wavelet coefficients. For sufficiently small β, It is almost equivalent to the universal soft-thresholding of [31]. The optimal selection of β is described in Theorem 3.

We refer to the duration between two consecutive restarts inclusive of the first restart but exclusive of the second as a bin. The policy identifies several bins across time, whose width is adaptively chosen.

2.3.3 Dynamic Regret of Arrows

In this section, we provide bounds for non-stationary regret and run-time of the policy.

Theorem 3. *Let the feedback be $y_t = \theta_t + Z_t$, $t = 1, \ldots, n$ and Z_t be independent, σ-subgaussian random variables. If $\beta = 24 + \frac{8\log(8/\delta)}{\log(n)}$, then with probability at least $1-\delta$,* Arrows *achieves a dynamic regret of $\tilde{O}(n^{1/3}\|D\theta_{1:n}\|_1^{2/3}\sigma^{4/3} + |\theta_1|^2 + \|D\theta_{1:n}\|_2^2 + \sigma^2)$ where*

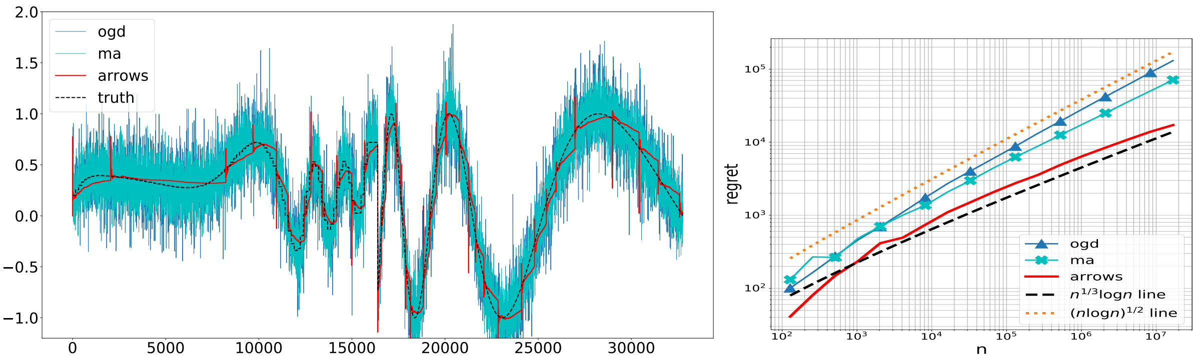


Figure 2.2: *An illustration of* Arrows *on a sequence with heterogeneous smoothness. We compare qualitatively (on the left) and quantitatively (on the right) to two popular baselines: (a) restarting online gradient descent [15]; (b) the moving averages [19] with optimal parameter choices. As we can see,* Arrows *achieves the optimal $\tilde{O}(n^{1/3})$ regret while the baselines are both suboptimal.*

$\tilde{O}$ hides a logarithmic factor in n and $1/\delta$.

Proof Sketch. Our policy is similar in spirit to restarting OGD but with an adaptive restart schedule. The key idea we used is to reduce the dynamic regret of our policy in probability roughly to a sum of squared error of a soft thresholding estimator and number of restarts. This was accomplished by using a Follow The Leader (FTL) reduction. For bounding the squared error part of the sum we modified the threshold value for the estimator in [31] and proved high probability guarantees for the convergence of its empirical mean. To bound the number of times we restart, we first establish a connection between Haar coefficients and total variation. This is intuitive since computation of Haar coefficients can be viewed as smoothing the adjacent regions of a signal and taking their difference. Then we exploit a special condition called "uniform shrinkage" of the soft-thresholding estimator which helps to optimally bound the number of restarts with high probability. □

Theorem 3 provides an upper bound of the minimax dynamic regret for forecasting the TV class.

Corollary 4. *Suppose the ground truth $\theta_{1:n} \in TV(C_n)$ and $|\theta_1| \leq U$. Then $\|D\theta_{1:n}\|_1 \leq C_n$. By noting that $\|D\theta_{1:n}\|_2 \leq \|D\theta_{1:n}\|_1$, under the setup in Theorem 3* Arrows *achieves a dynamic regret of $\tilde{O}(n^{1/3}C_n^{2/3}\sigma^{4/3} + U^2 + C_n^2 + \sigma^2)$ with probability at-least $1-\delta$.*

Remark 5 (Adaptivity to unknown parameters.)**.** Observe that Arrows does not require the knowledge of C_n.It adapts optimally to the unknown TV radius $C_n := \|D\theta_{1:n}\|_1$ of the ground truth $\theta_{1:n}$. The adaptivity to n can be achieved by a standard doubling trick. σ, if unknown, can be robustly estimated from the first few observations by a Median Absolute Deviation estimator (eg. Section 7.5 of [18]), thanks to the sparsity of wavelet coefficients of TV bounded functions.

2.3.4 A lower bound on the minimax regret

We now give a matching lower bound of the expected regret, which establishes that Arrows is adaptively minimax.

Proposition 6. *Assume $\min\{U, C_n\} > 2\pi\sigma$ and $n > 3$, there is a universal constant c such that*

$$\inf_{x_{1:n}} \sup_{\theta_{1:n} \in \mathrm{TV}(C_n)} \mathbb{E}\left[\sum_{t=1}^{n} \big(x_t(y_{1:t-1}) - \theta_t\big)^2\right] \geq c(U^2 + C_n^2 + \sigma^2 \log n + n^{1/3}C_n^{2/3}\sigma^{4/3}).$$

The proof is deferred to the Appendix A.8. The result shows that our result in Theorem 3 is optimal up to a logarithmic term in n and $1/\delta$ for almost all regimes (modulo trivial cases of extremely small $\min\{U, C_n\}/\sigma$ and n)[3].

[3]When both U and C_n are moderately small relative to σ, the lower bound will depend on σ a little differently because the estimation error goes to 0 faster than $1/\sqrt{n}$. We know the minimax risk exactly for that case as well but it is somewhat messy [32]. When they are both much smaller than σ, e.g., when $\min\{U, C_n\} \leq \sigma/\sqrt{n}$, then outputting 0 when we do not have enough information will be better than doing online averages.

Remark 7 (The price of forecasting)**.** The result also shows that *forecasting is strictly harder than smoothing.* Observe that a term with C_n^2 is required even if $\sigma = 0$, whereas in the case of a one-step look-ahead oracle (or the smoothing algorithm that sees all n observations) does not have this term. This implies that the total amount of variation that *any* algorithm can handle while producing a sublinear regret has dropped from $C_n = o(n)$ to $C_n = o(\sqrt{n})$. Moreover, the regime where the $n^{1/3}C_n^{2/3}\sigma^{4/3}$ term is meaningful only when $C_n = o(n^{1/4})$. For the region where $\sigma n^{1/4} \ll C_n \ll \sigma n^{1/2}$, the minimax regret is essentially proportional to C_n^2. This is illustrated in Figure 2.3.

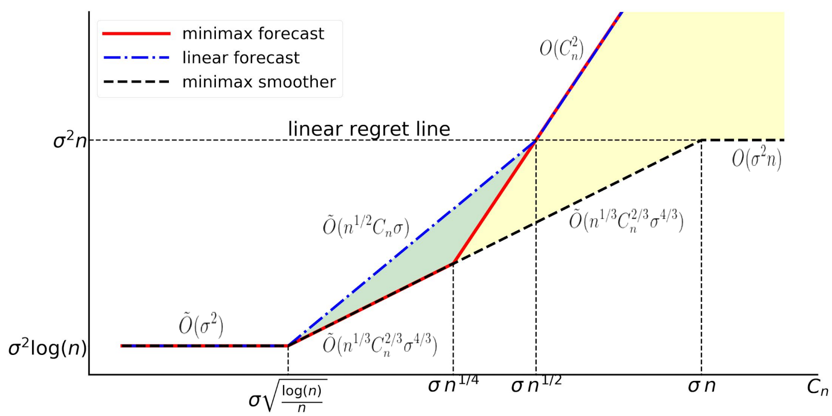


Figure 2.3: *An illustration of the minimax (dynamic) regret of forecasters and smoothers as a function of C_n. The non-trivial regime for forecasting is when C_n lies between $\sigma\sqrt{\frac{\log(n)}{n}}$ and $\sigma\, n^{1/4}$ where forecasting is just as hard as smoothing. When $C_n > \sigma\, n^{1/4}$, forecasting is harder than smoothing. The yellow region indicates the extra loss incurred by any minimax forecaster. The green region marks the extra loss incurred by a linear forecaster compared to minimax forecasting strategy. The figure demonstrates that linear forecasters are sub-optimal even in the non-trivial regime. When $C_n > \sigma\, n^{1/2}$, it is impossible to design a forecasting strategy with sub-linear regret. For $C_n > \sigma\, n$, identity function is optimal estimator for smoothing and when when $C_n < \sigma\sqrt{\frac{\log(n)}{n}}$, online averaging is optimal for both problems.*

We note that in much of the online learning literature, it is conventional to consider a slightly more restrictive setting with bounded domain, which could reduce the minimax regret. The following remark summarizes a variant of our results in this setting.

Remark 8 (Minimax regret in bounded domain)**.** If we consider predicting sequences from a subset of the $TV(C_n)$ ball having an extra boundedness condition $|\theta_i| \leq B$ for $i = 1 \ldots n$, it can be shown that (see Appendix A.8) minimax regret is $\tilde{\Omega}\left(\min\{nB^2, n\sigma^2, n^{1/3}C_n^{2/3}\sigma^{4/3}\} + B^2 + \min\{nB^2, BC_n\} + \sigma^2\right)$. In particular, forecasting is still strictly harder than smoothing due to the $\min\{nB^2, BC_n\}$ term in the bound. The discussion in Appendix A.8, shows a way of using Arrows whose regret can match this lower bound.

2.3.5 The adaptivity of Arrows to Sobolev and Holder classes

It turns out that Arrows is also adaptively optimal in forecasting sequences in the discrete Sobolev classes and the discrete Holder classes, which are defined as

$$\mathcal{S}(C'_n) = \{\theta_{1:n} : \|D\theta_{1:n}\|_2 \leq C'_n\}, \qquad \mathcal{H}(B'_n) = \{\theta_{1:n} : \|D\theta_{1:n}\|_\infty \leq B'_n\}. \tag{2.2}$$

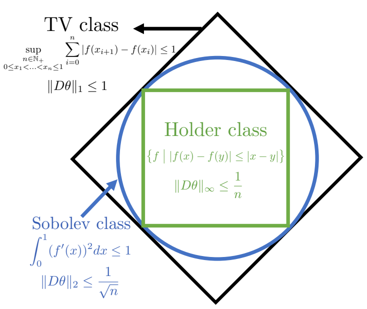
These classes feature sequences that are more spatially homogeneous than those in the TV class. The minimax cumulative error of nonparametric estimation in the discrete Sobolev class is $\Theta(n^{2/3}[C'_n]^{2/3}\sigma^{4/3})$ [33].

Corollary 9. *Let the feedback be $y_t = \theta_t + Z_t$ where Z_t is an independent, σ-subgaussian random variable. Let $\theta_{1:n} \in \mathcal{S}(C'_n)$ and $|\theta_1| \leq U$. If $\beta = 24 + \frac{8\log(8/\delta)}{\log(n)}$, then with probability at least $1-\delta$,* Arrows *achieves a dynamic regret of $\tilde{O}(n^{2/3}[C'_n]^{2/3}\sigma^{4/3} + U^2 + [C'_n]^2 + \sigma^2)$ where $\tilde{O}$ hides a logarithmic factor in n and $1/\delta$.*

Thus despite the fact that Arrows is designed for total variation class, it adapts to the optimal rates of forecasting sequences that are spatially regular. To gain some intuition, let's minimally expand the Sobolev ball to a TV ball of radius $C_n = \sqrt{n}C'_n$. The chosen scaling of C_n activates the embedding $\mathcal{S}(C'_n) \subset TV(C_n)$ (see the illustration

Table 2.2: *Minimax rates for cumulative error* $\sum_{i=1}^{n}(\hat{\theta}_i - \theta)^2$ *in various settings and policies that achieve those rates.* ARROWS *is adaptively minimax across all of the described function classes while linear forecasters fail to perform optimally over the TV class. For simplicity, we assume* U *is small and hide a* $\log n$ *factors in all the forecasting rates.*

Class		Minimax rate for Forecasting	Minimax rate for Smoothing	Minimax rate for Linear Forecasting
TV	$\|D\theta_{1:n}\|_1 \leq C_n$	$n^{1/3}C_n^{2/3}\sigma^{4/3} + C_n^2 + \sigma^2$	$n^{1/3}C_n^{2/3}\sigma^{4/3} + \sigma^2$	$n^{1/2}C_n\sigma + C_n^2 + \sigma^2$
Sobolev	$\|D\theta_{1:n}\|_2 \leq C_n'$	$n^{2/3}[C_n']^{2/3}\sigma^{4/3} + [C_n']^2 + \sigma^2$	$n^{2/3}[C_n']^{2/3}\sigma^{4/3} + \sigma^2$	$n^{2/3}[C_n']^{2/3}\sigma^{4/3} + [C_n']^2 + \sigma^2$
Holder	$\|D\theta_{1:n}\|_\infty \leq L_n$	$nL_n^{2/3}\sigma^{4/3} + nL_n^2 + \sigma^2$	$nL_n^{2/3}\sigma^{4/3} + \sigma^2$	$nL_n^{2/3}\sigma^{4/3} + nL_n^2 + \sigma^2$
Minimax Algorithm		ARROWS	Wavelet Smoothing Trend Filtering	Restarting OGD Moving Averages



Canonical Scaling[a]		Forecasting	Smoothing	Linear Forecasting
TV	$C_n \asymp 1$	$n^{1/3}$	$n^{1/3}$	$n^{1/2}$
Sobolev	$C_n' \asymp 1/\sqrt{n}$	$n^{1/3}$	$n^{1/3}$	$n^{1/3}$
Holder	$L_n \asymp 1/n$	$n^{1/3}$	$n^{1/3}$	$n^{1/3}$

[a] The "canonical scaling" are obtained by discretizing functions in canonical function classes. Under the canonical scaling, Holder class $\subset$ Sobolev class $\subset$ TV class, as shown in the figure on the left. ARROWS is optimal for the Sobolev and Holder classes inscribed in the TV class. MA and Restarting OGD on the other hand require different parameters and prior knowledge of variational budget (i.e C_n or C_n') to achieve the minimax linear rates for the TV class and the Sobolev/Holder class.

in Table 2.2) with both classes having same minimax rate in the batch setting. This implies that dynamic regret of ARROWS is simultaneously minimax optimal over $\mathcal{S}(C_n')$ and $TV(C_n)$ wrt the term containing n. It can be shown that ARROWS is optimal wrt to the additive $[C_n']^2, U^2, \sigma^2$ terms as well. Minimaxity in Sobolev class implies minimaxity in Holder class since it is known that a Holder ball is sandwiched between two Sobolev balls having the same minimax rate [34]. A proof of the Corollary and related experiments are presented in Appendix A.5 and A.9.

2.3.6 Fast computation

Last but not least, we remark that there is a fast implementation of ARROWS that reduces the overall time-complexity for n step from $O(n^2)$ to $O(n \log n)$.

Proposition 10. *The run time of* ARROWS *is* $O(n \log(n))$*, where n is the time horizon.*

The proof exploits the sequential structure of our policy and sparsity in wavelet transforms, which allows us to have $O(\log n)$ incremental updates in all but $O(\log n)$ steps. See Appendix A.6 for details.

2.3.7 Experimental Results

To empirically validate our results, we conducted a number of numerical simulations that compares the regret of ARROWS, (Restarting) OGD and MA. Figure 2.2 shows the results on a function with heterogeneous smoothness (see the exact details and more experiments in Appendix A.1) with the hyperparameters selected according to their theoretical optimal choice for the TV class (See Theorem 112, 113 for OGD and MA in Appendix A.2). The left panel illustrates that ARROWS is locally adaptive to heterogeneous smoothness of the ground truth. Red peaks in the figure signifies restarts. During the initial and final duration, the signal varies smoothly and ARROWS chooses a larger window size for online averaging. In the middle, signal varies rather abruptly. Consequently ARROWS chooses a smaller window size. On the other hand, the linear smoothers OGD and MA use a constant width and cannot adapt to the different regions of the space. This differences are also reflected in the quantitative evaluation on the right, which clearly shows that OGD and MA has a suboptimal $\tilde{O}(\sqrt{n})$ regret while ARROWS attains the $\tilde{O}(n^{1/3})$ minimax regret!

2.4 Concluding Discussion

In this chapter, we studied the problem of online nonparametric forecasting of bounded variation sequences. We proposed a new forecasting policy ARROWS and proved that it achieves a cumulative square error (or dynamic regret) of $\tilde{O}(n^{1/3}C_n^{2/3}\sigma^{4/3}+\sigma^2+U^2+C_n^2)$ with total runtime of $O(n \log n)$. We also derived a lower bound for forecasting sequences with bounded total variation which matches the upper bound up to a logarithmic term which certifies the optimality of ARROWS in all parameters. Through connection to linear estimation theory, we assert that no linear forecaster can achieve the optimal rate. ARROWS is highly adaptive and has essentially no tuning parameters. We show that it is adaptively minimax (up to a logarithmic factor) simultaneously for all discrete TV classes, Sobolev classes and Holder classes with unknown radius.

Chapter 3

Adaptive Online Estimation of Piecewise Polynomial Trends

In time series analysis, estimating and removing the trend are often the first steps taken to make the sequence "stationary". The non-parametric assumption that the underlying trend is a piecewise polynomial or a spline [35], is one of the most popular choices, especially when we do not know where the "change points" are and how many of them are appropriate. The higher order Total Variation (see Assumption A3) of the trend can capture in some sense both the sparsity and intensity of changes in underlying dynamics. A non-parametric regression method that penalizes this quantity — trend filtering [3] — enjoys a superior *local adaptivity* over traditional methods such as the Hodrick-Prescott Filter [36]. However, Trend Filtering is an *offline* algorithm which limits its applicability for the inherently *online* time series forecasting problem. In this chapter, we are interested in designing an online forecasting strategy that can essentially match the performance of the offline methods for trend estimation, hence allowing us to apply time series models forecasting on-the-fly. In particular, our problem setup (see Figure 3.1) and algorithm are applicable to all *online variants* of trend filtering problem such as predicting stock prices, server payloads, sales etc.

3.1 Setup, Assumptions and Contributions

3.1.1 Setup

Let's describe the notations that will be used throughout the chapter. All vectors and matrices will be written in bold face letters. For a vector $\boldsymbol{x} \in \mathbb{R}^m$, $\boldsymbol{x}[i]$ or $\boldsymbol{x}_i$ denotes its value at the i^{th} coordinate. $\boldsymbol{x}[a:b]$ or $\boldsymbol{x_{a:b}}$ is the vector $[\boldsymbol{x}[a], \ldots, \boldsymbol{x}[b]]$. $\|\cdot\|_p$ denotes finite dimensional L_p norms. $\|\boldsymbol{x}\|_0$ is the number of non-zero coordinates of a vector $\boldsymbol{x}$. $[n]$ represents the set $\{1, \ldots, n\}$. $\boldsymbol{D}^i \in \mathbb{R}^{(n-i)\times n}$ denotes the discrete difference operator of order i defined as in [3] and reproduced below.

$$\boldsymbol{D}^1 = \begin{bmatrix} -1 & 1 & 0 & \ldots & 0 & 0 \\ 0 & -1 & 1 & \ldots & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & \ldots & -1 & 1 \end{bmatrix} \in \mathbb{R}^{(n-1)\times n}, \tag{3.1}$$

and $\boldsymbol{D}^i = \tilde{\boldsymbol{D}}^1 \cdot \boldsymbol{D}^{i-1}\ \forall i \geq 2$ where $\tilde{\boldsymbol{D}}^1$ is the $(n-i) \times (n-i+1)$ truncation of $\boldsymbol{D}^1$.

The theme of this chapter builds on the non-parametric online forecasting model developed in [37]. We consider a sequential n step interaction process between an agent and an adversary as shown in Figure 3.1.

A forecasting strategy $\mathcal{S}$ is defined as an algorithm that outputs a prediction $\mathcal{S}(t)$ at time t only based on the information available after the completion of time $t-1$. Random variables ϵ_t for $t \in [n]$ are independent and subgaussian with parameter σ^2. This sequential game can be regarded as an online version of the non-parametric regression setup well studied in statistics community.

1. Fix a time horizon n.
2. Agent declares a forecasting strategy $\mathcal{S}$
3. Adversary chooses a sequence $\boldsymbol{\theta}_{1:n}$
4. For $t = 1, \ldots, n$:
   (a) Agent outputs a prediction $\mathcal{S}(t)$.
   (b) Adversary reveals $y_t = \boldsymbol{\theta}_{1:n}[t] + \epsilon_t$
5. After n steps, agent suffers a cumulative loss $\sum_{i=1}^{n} \big(\mathcal{S}(i) - \boldsymbol{\theta}_{1:n}[i]\big)^2$.

Figure 3.1: Interaction protocol

In this chapter, we consider the problem of forecasting sequences that obey $n^k \|D^{k+1}\boldsymbol{\theta}_{1:n}\|_1 \leq C_n$, $k \geq 0$ and $\|\boldsymbol{\theta}_{1:n}\|_\infty \leq B$. The constraint $n^k \|D^{k+1}\boldsymbol{\theta}_{1:n}\|_1 \leq C_n$ has been widely used in the rich literature of non-parametric regression. For example, the offline problem of estimating sequences obeying such higher order difference constraint from noisy labels under squared error loss is studied in [14, 2, 3, 38, 33, 39] to cite a few. We aim to design forecasters whose predictions are only based on past history and still perform as good as a batch estimator that sees the entire observations ahead of time.

Scaling of n^k. The family $\{\boldsymbol{\theta}_{1:n} \mid n^k \|D^{k+1}\boldsymbol{\theta}_{1:n}\|_1 \leq C_n\}$ may appear to be alarmingly restrictive for a constant C_n due to the scaling factor n^k, but let us argue why this is actually a natural construct. The continuous TV^k distance of a function $f : [0,1] \to \mathbb{R}$ is defined as $\int_0^1 |f^{(k+1)}(x)| dx$, where $f^{(k+1)}$ is the $(k+1)^{th}$ order (weak) derivative. A sequence can be obtained by sampling the function at $x_i = i/n$, $i \in [n]$. Discretizing the integral yields the TV^k distance of this sequence to be $n^k \|D^{k+1}\boldsymbol{\theta}_{1:n}\|_1$. Thus, the $n^k \|D^{k+1}\boldsymbol{\theta}_{1:n}\|_1$ term can be interpreted as the discrete approximation to continuous higher order TV distance of a function. See Figure 3.2 for an illustration for the case $k = 1$.

Non-stationary Stochastic Optimization. The setting above can also be viewed under the framework of non-stationary stochastic optimization as studied in [15, 22] with squared error loss and noisy gradient feedback. At each time step, the adversary chooses a loss function $f_t(x) = (x - \boldsymbol{\theta}_t)^2$. Since $\nabla f_t(x) = 2(x - \boldsymbol{\theta}_t)$, the feedback $\tilde{\nabla} f_t(x) = 2(x - y_t)$ constitutes an unbiased estimate of the gradient $\nabla f_t(x)$. [15, 22] quantifies the performance of a forecasting strategy $\mathcal{S}$ in terms of dynamic regret as follows.

$$R_{dynamic}(\mathcal{S}, \boldsymbol{\theta}_{1:n}) := \mathbb{E}\left[\sum_{t=1}^{n} f_t\left(\mathcal{S}(t)\right)\right] - \sum_{t=1}^{n} \inf_{x_t} f_t(x_t), = \mathbb{E}\left[\sum_{t=1}^{n} \left(\mathcal{S}(t) - \boldsymbol{\theta}_{1:n}[t]\right)^2\right], \quad (3.2)$$

where the last equality follows from the fact that when $f_t(x) = (x - \boldsymbol{\theta}_{1:n}[t])^2$, $\inf_x (x - \boldsymbol{\theta}_{1:n}[t])^2 = 0$. The expectation above is taken over the randomness in the noisy gradient feedback and that of the agent's forecasting strategy. It is impossible to achieve sublinear dynamic regret against arbitrary ground truth sequences. However if the sequence of minimizers of loss functions $f_t(x) = (x - \theta_t)^2$ obey a path variational constraint, then we can parameterize the dynamic regret as a function of the path length, which could be sublinear when the path-length is sublinear. Typical variational constraints considered in the existing work includes $\sum_t |\boldsymbol{\theta}_t - \boldsymbol{\theta}_{t-1}|$, $\sum_t |\boldsymbol{\theta}_t - \boldsymbol{\theta}_{t-1}|^2$, $(\sum_t \|f_t - f_{t-1}\|_p^q)^{1/q}$ [37]. These are all useful in their respective contexts, but do not capture higher order smoothness.

The purpose of this work is to connect ideas from batch non-parametric regression to the framework of online stochastic optimization and define a *natural family of higher order variational functionals* of the form $\|D^{k+1}\boldsymbol{\theta}_{1:n}\|_1$ to track a comparator sequence with piecewise polynomial structure. To the best of our knowledge such higher order path variationals for $k \geq 1$ are vastly unexplored in the domain of non-stationary stochastic optimization. In this work, we take the first steps in introducing such variational constraints to online non-stationary stochastic optimization and exploiting them to get sub-linear dynamic regret.

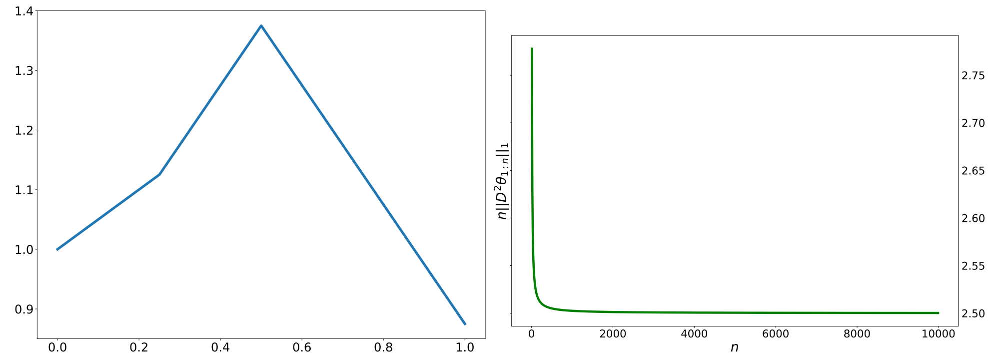


Figure 3.2: *A TV^1 bounded comparator sequence $\boldsymbol{\theta}_{1:n}$ can be obtained by sampling the continuous piecewise linear function on the left at points i/n, $i \in [n]$. On the right, we plot the TV^1 distance (which is equal to $n\|D^2\boldsymbol{\theta}_{1:n}\|_1$ by definition) of the generated sequence for various sequence lengths n. As n increases the discrete TV^1 distance converges to a constant value given by the continous TV^1 distance of the function on left panel.*

3.1.2 Assumptions

(A1) The time horizon is known to be n.

(A2) The parameter σ^2 of subgaussian noise in the observations is a known fixed positive constant.

(A3) The ground truth denoted by $\boldsymbol{\theta}_{1:n}$ has its k^{th} order total variation bounded by some positive C_n, i.e., we consider ground truth sequences that belongs to the class

$$\mathrm{TV}^k(C_n) := \{\boldsymbol{\theta}_{1:n} \in \mathbb{R}^n : n^k\|D^{k+1}\boldsymbol{\theta}_{1:n}\|_1 \leq C_n\}$$

We refer to $n^k\|D^{k+1}\boldsymbol{\theta}_{1:n}\|_1$ as TV^k distance of the sequence $\boldsymbol{\theta}_{1:n}$. To avoid trivial cases, we assume $C_n = \Omega(1)$.

(A4) The TV order k is a known *fixed* positive constant.

(A5) $\|\boldsymbol{\theta}_{1:n}\|_\infty \leq B$ for a known *fixed* positive constant B.

Though we require the time horizon to be known in advance in assumption (A1), this can be easily lifted using standard doubling trick arguments. The knowledge of time horizon helps us to present the policy in a most transparent way. If standard deviation of sub-gaussian noise is unknown, contrary to assumption (A2), then it can be robustly estimated by a Median Absolute Deviation estimator using first few observations, see for eg. [18]. This is indeed facilitated by the sparsity of wavelet coefficients of TV^k bounded sequences. Assumption (A3) characterizes the ground truth sequences whose forecasting is the main theme of this chapter. The $\mathrm{TV}^k(C_n)$ class features a rich family of sequences that can potentially exhibit spatially non-homogeneous smoothness. For example it can capture sequences that are piecewise polynomials of degree at most k. This poses a challenge to design forecasters that are *locally adaptive* and can efficiently detect and make predictions under the presence of the non-homogeneous trends. Though knowledge of the TV order k is required in assumption (A4), most of the practical interest is often limited to the lower orders $k = 0, 1, 2, 3$, see for eg. [4, 3] and we present (in Appendix B.3) a meta-policy based on exponential weighted averages [40] to adapt to these lower orders. Finally assumption (A5) is standard in the online learning literature.

3.1.3 Contributions

- When the revealed labels are noisy realizations of sequences that belong to $TV^k(C_n)$ we propose a *polynomial time* policy called `Ada-VAW` (**Ada**ptive **V**ovk **A**zoury **W**armuth forecaster) that achieves the nearly *minimax optimal* rate of $\tilde{O}\left(n^{\frac{1}{2k+3}} C_n^{\frac{2}{2k+3}}\right)$ for $R_{dynamic}$ with high probability. The proposed policy *optimally adapts to the unknown radius C_n*.

- We show that the proposed policy achieves optimal $R_{dynamic}$ when revealed labels are noisy realizations of sequences residing in higher order discrete Holder and discrete Sobolev classes.

- When the revealed labels are noisy realizations of sequences that obey $\|D^k \boldsymbol{\theta}_{1:n}\|_0 \leq J_n, \|\boldsymbol{\theta}_{1:n}\|_\infty \leq B$, we show that the same policy achieves the minimax optimal $\tilde{O}(J_n)$ rate for for $R_{dynamic}$ with high probability. The policy *optimally adapts to unknown* J_n.

Notes on key novelties. It is known that the VAW forecaster is an optimal algorithm for online polynomial regression with squared error losses [40]. With the side information of change points where the underlying ground truth switches from one polynomial to another, we can run a VAW forecaster on each of the stable polynomial sections to control the cumulative squared error of the policy. We use the machinery of wavelets to mimic an oracle that can provide side information of the change points. For detecting change points, a restart rule is formulated by exploiting connections between wavelet coefficients and locally adaptive regression splines. This is a *more general* strategy than that used in [37]. To the best of our knowledge, this is the *first* time an interplay between VAW forecaster and theory of wavelets along with its adaptive minimaxity [2] has been used in the literature.

Wavelet computations require the length of underlying data whose wavelet transform needs to be computed has to be a power of 2. In practice this is achieved by a padding strategy in cases where original data length is not a power of 2. We show that most commonly used padding strategies – eg. zero padding as in [37] – are not useful for the current problem and propose a novel *packing strategy* that alleviates the need to pad. This will be useful to many applications that use wavelets which can be well beyond the scope of the current chapter.

Our proof techniques for bounding regret use properties of the CDJV wavelet construction [41]. To the best of our knowledge, this is the *first* time we witness the ideas from a general CDJV construction scheme implying useful results in an online learning paradigm. Optimally controlling the bias of VAW demands to carefully bound the ℓ_2 norm of coefficients computed by polynomial regression. This is done by using ideas from number theory and symbolic determinant evaluation of polynomial matrices. This could be of independent interest in both offline and online polynomial regression.

3.2 Main results

We present below the main results of the chapter. All proofs are deferred to the appendix.

3.2.1 Limitations of linear forecasters

We exhibit a lower-bound on the dynamic regret that is implied by [2] in batch regression setting.

Proposition 11 (Minimax Regret)**.** *Let* $y_t = \boldsymbol{\theta}_{1:n}[t] + \epsilon_t$ *for* $t = 1, \ldots, n$ *where* $\theta_{1:n} \in TV^{(k)}(C_n)$*,* $|\boldsymbol{\theta}_{1:n}[t]| \leq B$ *and* ϵ_t *are iid* σ^2 *subgaussian random variables. Let* $\mathcal{A}_F$ *be the class of all forecasting strategies whose prediction at time* t *only depends on* $y_1, \ldots, y_{t-1}$*. Let* $\boldsymbol{s_t}$ *denote the prediction at time* t *for a strategy* $\boldsymbol{s} \in \mathcal{A}_F$*. Then,*

$$\inf_{\boldsymbol{s} \in \mathcal{A}_F} \sup_{\boldsymbol{\theta}_{1:n} \in TV^{(k)}(C_n)} \sum_{t=1}^{n} E\left[(\boldsymbol{s_t} - \boldsymbol{\theta}_{1:n}[t])^2\right] = \Omega\left(\min\{n, n^{\frac{1}{2k+3}} C_n^{\frac{2}{2k+3}}\}\right), \tag{3.3}$$

where the expectation is taken wrt to randomness in the strategy of the player and ϵ_t*.*

We define linear forecasters to be strategies that predict a fixed linear function of the history. This includes a large family of polices including the ARIMA family, Exponential Smoothers for Time Series forecasting, Restarting OGD etc. However in the presence of spatially inhomogeneous smoothness – which is the case with TV bounded sequences – these policies are doomed to perform sub-optimally. This can be made precise by providing a lower-bound on the minimax regret for linear forecasters. Since the offline problem of smoothing is easier than that of forecasting, a lower-bound on the minimax MSE of linear smoother will directly imply a lower-bound on the regret of linear forecasting strategies. By the results of [2], we have the following proposition:

Proposition 12 (Minimax regret for linear forecasters)**.** *Linear forecasters will suffer a dynamic regret of at least* $\Omega(n^{1/(2k+2)})$ *for forecasting sequences that belong to* $TV^k(1)$.

Thus we must look in the space of policies that are *non-linear* functions of past labels to achieve a minimax dynamic regret that can potentially match the lower-bound in Proposition 11.

3.2.2 Policy

In this section, we present our policy and capture the intuition behind its design. First, we introduce the following notations.

- The policy works by partitioning the time horizon into several bins. t_h denotes start time of the current bin and t be the current time point.
- $\boldsymbol{W}$ denotes the orthonormal Discrete Wavelet Transform (DWT) matrix obtained from a CDJV wavelet construction [41] using wavelets of regularity $k+1$.
- $T(\boldsymbol{y})$ denotes the vector obtained by elementwise soft-thresholding of $\boldsymbol{y}$ at level $\sigma\sqrt{\beta \log l}$ where l is the length of input vector.
- $\boldsymbol{x_t} \in \mathbb{R}^{(k+1)}$ denotes the vector $[1, t-t_h+k+1, \ldots, (t-t_h+k+1)^k]^T$.
- $A_t = \boldsymbol{I} + \sum_{s=t_h-k}^{t} \boldsymbol{x_s}\boldsymbol{x_s}^T$

- `recenter`($\boldsymbol{y}[s:e]$) function first computes the Ordinary Least Square (OLS) polynomial fit with features $\boldsymbol{x}_s, \ldots, \boldsymbol{x}_e$. It then outputs the residual vector obtained by subtracting the best polynomial fit from the input vector $\boldsymbol{y}[s:e]$.

- Let L be the length of a vector $\boldsymbol{u}_{1:t}$. `pack`($\boldsymbol{u}$) first computes $l = \lfloor \log_2 L \rfloor$. It then returns the pair
  ($\boldsymbol{u}_{1:2^l}, \boldsymbol{u}_{t-2^l+1:t}$). We call elements of this pair as segments of $\boldsymbol{u}$.

> `Ada-VAW`: inputs - observed y values, TV order k, time horizon n, sub-gaussian parameter σ, hyper-parameter $\beta > 24$ and $\delta \in (0, 1]$
>
> 1. For $t = 1$ to $k - 1$, predict 0
> 2. Initialize $t_h = k$
> 3. For $t = k$ to n:
>   (a) Predict $\hat{y}_t = \langle \boldsymbol{x_t}, A_t^{-1} \sum_{s=t_h-k}^{t-1} y_s \boldsymbol{x_s} \rangle$
>   (b) Observe y_t and suffer loss $(\hat{y}_t - \boldsymbol{\theta}_{1:n}[t])^2$
>   (c) Let $\mathbf{y}_r$ =`recenter`($\mathbf{y}[t_h - k : t]$) and L be its length
>   (d) Let $(\boldsymbol{y}_1, \boldsymbol{y}_2) =$ `pack`($\boldsymbol{y}_r$)
>   (e) Let $(\hat{\boldsymbol{\alpha}}_1, \hat{\boldsymbol{\alpha}}_2) = (T(\boldsymbol{W}\mathbf{y}_1), T(\boldsymbol{W}\mathbf{y}_2))$
>   (f) Restart Rule: If $\|\hat{\boldsymbol{\alpha}}_1\|_2 + \|\hat{\boldsymbol{\alpha}}_2\|_2 > \sigma$ then
>   i. set $t_h = t + 1$

The basic idea behind the policy is to adaptively detect intervals that have low TV^k distance. If the TV^k distance within an interval is guaranteed to be low enough, then outputting a polynomial fit can suffice to obtain low prediction errors. Here we use the polynomial fit from VAW [42] forecaster in step 3(a) to make predictions in such low TV^k intervals. Step 3(e) computes denoised wavelets coefficients. It can be shown that the expression on the LHS of the inequality in step 3(f) can be used to lower bound $\sqrt{L}$ times the TV^k distance of the underlying ground truth with high probability. Informally speaking, this is expected as the wavelet coefficents for a CDJV system with regularity k

are computed using higher order differences of the underlying signal. A restart is triggered when the scaled TV^k lower-bound within a bin exceeds the threshold of σ. Thus we use the energy of denoised wavelet coefficients as a device to detect low TV^k intervals. In Appendix B.4 we show that popular padding strategies such as zero padding, greatly inflate the TV^k distance of the recentered sequence for $k \geq 1$. This hurts the dynamic regret of our policy. To obviate the necessity to pad for performing the DWT, we employ a packing strategy as described in the policy.

3.2.3 Performance Guarantees

Theorem 13. *Consider the the feedback model* $y_t = \boldsymbol{\theta}_{1:n}[t] + \epsilon_t$ $t = 1, \ldots, n$ *where* ϵ_t *are independent* σ^2 *subguassian noise and* $|\boldsymbol{\theta}_{1:n}[t]| \leq B$*. If* $\beta = 24 + \frac{8\log(8/\delta)}{\log(n)}$*, then with probability at least* $1-\delta$*,* **`Ada-VAW`** *achieves a dynamic regret of* $\tilde{O}\left(n^{\frac{1}{2k+3}}\left(n^k\|D^{k+1}\boldsymbol{\theta}_{1:n}\|_1\right)^{\frac{2}{2k+3}}\right)$ *where* $\tilde{O}$ *hides poly-logarithmic factors of* n*,* $1/\delta$ *and constants* k*,*σ*,*B *that do not depend on* n*.*

Proof Sketch. Our proof strategy falls through the following steps.

1. Obtain a high probability bound of bias variance decomposition type on the total squared error incurred by the policy within a bin.

2. Bound the variance by optimally bounding the number of bins spawned.

3. Bound the squared bias using the restart criterion.

Step 1 is achieved by using the subgaussian behaviour of revealed labels y_t. For step 2, we first connect the wavelet coefficients of a recentered signal to its TV^k distance using ideas from theory of Regression Splines. Then we invoke the "uniform shrinkage" property of soft thresholding estimator to construct a lowerbound of the TV^k distance within a bin. Such a lowerbound when summed across all bins leads to an upperbound on the number of bins spawned. Finally for step 3, we use a reduction from the squared bias within a bin to the regret of VAW forecaster and exploit the restart criterion and adpative minimaxity of soft thresholding estimator [2] that uses a CDJV wavelet system. □

Corollary 14. *Consider the setup of Theorem 13. For the problem of forecasting sequences* $\boldsymbol{\theta}_{1:n}$ *with* $n^k\|D^{k+1}\boldsymbol{\theta}_{1:n}\|_1 \leq C_n$ *and* $\|\boldsymbol{\theta}_{1:n}\|_\infty \leq B$*,* `Ada-VAW` *when run with* $\beta = 24 + \frac{8\log(8/\delta)}{\log(n)}$ *yields a dynamic regret of* $\tilde{O}\left(n^{\frac{1}{2k+3}}(C_n)^{\frac{2}{2k+3}}\right)$ *with probability atleast* $1-\delta$*.*

Remark 15. (*Adaptive Optimality*) By combining with trivial regret bound of $O(n)$, we see that dynamic regret of `Ada-VAW` matches the lower-bound provided in Proposition 11. `Ada-VAW` optimally adapts to the variational budget C_n. Adaptivity to time horizon n can be achieved by the standard doubling trick.

Remark 16. (*Extension to higher dimensions*) Let the ground truth $\boldsymbol{\theta}_{1:n}[t] \in \mathbb{R}^d$ and let $\boldsymbol{v}_i = [\boldsymbol{\theta}_{1:n}[1][i], \ldots, \boldsymbol{\theta}_{1:n}[n][i]], \Delta_i = n^k\|D^{k+1}\boldsymbol{v}_i\|_1$ for each $i \in [d]$. Let $\sum_{i=1}^d \Delta_i \leq C_n$. Then by running d instances of `Ada-VAW` in parallel where instance i predicts ground truth sequence along co-ordinate i, a regret bound of $\tilde{O}\left(d^{\frac{2k+1}{2k+3}}n^{\frac{1}{2k+3}}C_n^{\frac{2}{2k+3}}\right)$ can be achieved.

Remark 17. (*Generalization to other losses*) Consider the protocol in Figure 3.1. Instead of squared error losses in step (5), suppose we use loss functions $f_t(x)$ such that $\operatorname{argmin} f_t(x) = \boldsymbol{\theta}_{1:n}[t]$ and $f_t'(x)$ is γ-Lipschitz. Under this setting, `Ada-VAW` yields a dynamic regret of $\tilde{O}\left(\gamma n^{\frac{1}{2k+3}}C_n^{\frac{2}{2k+3}}\right)$ with probability at least $1-\delta$. Concrete examples include (but not limited to):

1. Huber loss, $f_t^{(\omega)}(x) = \begin{cases} 0.5(x-\boldsymbol{\theta}_{[1:n]}[t])^2 & |x-\boldsymbol{\theta}_{[1:n]}[t]| \leq \omega \\ \omega(|x-\boldsymbol{\theta}_{[1:n]}[t]| - \omega/2) & \text{otherwise} \end{cases}$ is 1-Lipschitz in gradient.

2. Log-Cosh loss, $f_t(x) = \log(\cosh(x - \boldsymbol{\theta}_{[1:n]}[t]))$ is 1-Lipschitz in gradient.

3. ϵ-insensitive logistic loss [43], $f_t^{(\epsilon)}(x) = \log(1+e^{x-\boldsymbol{\theta}_{[1:n]}[t]-\epsilon}) + \log(1+e^{-x+\boldsymbol{\theta}_{[1:n]}[t]-\epsilon}) - 2\log(1+e^{-\epsilon})$ is 1/2-Lipschitz in gradient.

The rationale behind both Remark 16 and Remark 17 is described at the end of Appendix B.2.2

Proposition 18. *There exist an $O\left(((k+1)n)^2\right)$ run-time implementation of* `Ada-VAW`*.*

The run-time of $O(n^2)$ is larger than the $O(n \log n)$ run-time of the more specialized algorithm of [37] for $k = 0$. This is due to the more complex structure of higher order CDJV wavelets which invalidates their trick that updates the Haar wavelets in an amortized $O(1)$ time.

3.3 Extensions

In this section, we discuss the potential applications of the proposed algorithm which broadens its generalizability to several interesting use cases.

3.3.1 Optimality for Higher Order Sobolev and Holder Classes

So far we have been dealing with total variation classes, which can be thought of as ℓ_1-norm of the $(k+1)$th order derivatives. An interesting question to ask is "how does `Ada-VAW` behave under smoothness metric defined in other norms, e.g., ℓ_2-norm and ℓ_∞-norm?" Following [3], we define the higher order discrete Sobolev class $\mathcal{S}^{k+1}(C'_n)$ and discrete Holder class $\mathcal{H}^{k+1}(L'_n)$ as follows.

$$\mathcal{S}^{k+1}(C'_n) = \{\boldsymbol{\theta}_{1:n} : n^k \|D^{k+1}\boldsymbol{\theta}_{1:n}\|_2 \leq C'_n\}, \tag{3.4}$$

$$\mathcal{H}^{k+1}(L'_n) = \{\boldsymbol{\theta}_{1:n} : n^k \|D^{k+1}\boldsymbol{\theta}_{1:n}\|_\infty \leq L'_n\}, \tag{3.5}$$

where $k \geq 0$. These classes feature sequences that are *spatially more regular* in comparison to the higher order TV^k class. It is well known that (see for eg. [44]) the following embedding holds true:

$$\mathcal{H}^{k+1}\left(\frac{C_n}{n}\right) \subseteq \mathcal{S}^{k+1}\left(\frac{C_n}{\sqrt{n}}\right) \subseteq TV^k(C_n). \tag{3.6}$$

Here $\frac{C_n}{\sqrt{n}}$ and $\frac{C_n}{n}$ are respectively the maximal radius of a Sobolev ball and Holder ball enclosed within a $TV^k(C_n)$ ball. Hence we have the following Corollary.

Corollary 19. *Assume the observation model of Theorem 13 and that $\boldsymbol{\theta}_{1:n} \in \mathcal{S}^{k+1}(C'_n)$. If $\beta = 24 + \frac{8\log(8/\delta)}{\log(n)}$, then with probability at least $1-\delta$,* `Ada-VAW` *achieves a dynamic regret of $\tilde{O}\left(n^{\frac{2}{2k+3}}[C'_n]^{\frac{2}{2k+3}}\right)$.*

It turns out that this is the optimal rate for the Sobolev classes, even in the easier, offline non-parametric regression setting [44]. Since a Holder class can be sandwiched between two Sobolev balls of same minimax rates [44], this also implies the adaptive optimality for the Holder class. We emphasize that `Ada-VAW` does not need to know the C_n, C'_n or L'_n parameters, which implies that it will achieve the smallest error permitted by the right norm that captures the smoothness structure of the unknown sequence $\boldsymbol{\theta}_{1:n}$.

3.3.2 Optimality for the case of Exact Sparsity

Next, we consider the performance of `Ada-VAW` on sequences satisfying an ℓ_0-(pseudo)norm measure of the smoothness, defined as

$$\mathcal{E}^{k+1}(J_n) = \{\boldsymbol{\theta}_{1:n} : \|D^{k+1}\boldsymbol{\theta}_{1:n}\|_0 \leq J_n, \|\boldsymbol{\theta}_{1:n}\|_\infty \leq B\}. \tag{3.7}$$

This class captures sequences that has at most J_n jumps in its $(k+1)^{th}$ order difference, which covers (modulo the boundedness) kth order discrete splines [45] with exactly J_n

knots, and arbitrary piecewise polynomials with $O(J_n/k)$ polynomial pieces.

The techniques we developed in this chapter allows us to establish the following performance guarantee for `Ada-VAW`, when applied to sequences in this family.

Theorem 20. *Let $y_t = \boldsymbol{\theta}_{1:n}[t] + \epsilon_t$, for $t = 1, \ldots, n$ where ϵ_t are iid sub-gaussian with parameter σ^2 and $\|D^{k+1}\boldsymbol{\theta}_{1:n}\|_0 \le J_n$ with $|\boldsymbol{\theta}_{1:n}[t]| \le B$ and $J_n \ge 1$. If $\beta = 24 + \frac{8\log(8/\delta)}{\log(n)}$, then with probability at least $1-\delta$, `Ada-VAW` achieves a dynamic regret of $\tilde{O}(J_n)$ where $\tilde{O}$ hides polynmial factors of $\log(n)$ and $\log(1/\delta)$.*

We also establish an information-theoretic lower bound that applies to all algorithms.

Proposition 21. *Under the interaction model in Figure 3.1, the minimax dynamic regret for forecasting sequences in $\mathcal{E}^{k+1}(J_n)$ is $\Omega(J_n)$.*

Remark 22. Theorem 20 and Proposition 21 imply that `Ada-VAW` is optimal (up to logarithmic factors) for the sequence family $\mathcal{E}^k(J_n)$. It is noteworthy that the `Ada-VAW` is adaptive in J_n, so it is essentially performing as well as an oracle that knows *how many* knots are enough to represent the input sequence as a discrete spline and *where* they are in advance (which leaves only the J_n polynomials to be fitted).

3.4 Concluding Discussion

In this chapter, we considered the problem of forecasting TV^k bounded sequences and proposed the first efficient algorithm – `Ada-VAW`– that is adaptively minimax optimal. We also discussed the adaptive optimality of `Ada-VAW` in various parameters and other function classes. In establishing strong connections between the locally adaptive nonparametric regression literature to the adaptive online learning literature in a concrete problem, this chapter could serve as a stepping stone for future exchanges of ideas between the research communities, and hopefully spark new theory and practical algorithms.

Chapter 4

An Optimal Reduction of TV-Denoising to Adaptive Online Learning

Total variation (TV) denoising [46] is a classical algorithm originated in the signal processing community which removes noise from a noisy signal y by solving the following regularized optimization problem

$$\min_f \|f - y\|_2^2 + \lambda \mathrm{TV}(f). \tag{4.1}$$

where $\mathrm{TV}(\cdot)$ denotes the total variation functional which is equivalent to $\int |f'(x)|dx$ for weakly differentiable functions. In discrete time, TV denoising is known as "fused lasso" in the statistics literature [47, 48], which solves

$$\min_{\boldsymbol{\theta} \in \mathbb{R}^n} \sum_{i=1}^{n} (\theta_i - y_i)^2 + \lambda \sum_{i=2}^{n} |\theta_i - \theta_{i-1}|. \tag{4.2}$$

where θ_i is the element at index i of the vector $\boldsymbol{\theta}$. Unlike their L2-counterpart, the TV regularization functional is designed to promote sparsity in the number of change points, hence inducing a "piecewise constant" structure in the solution.

Over the three decades since the advent of TV denoising, it has seen many influential applications. Algorithms that use TV-regularization has been deployed in every cellphone, digital camera and medical imaging devices. More recently, TV denoising is recognized as a pivotal component in generating the first image of a super massive black hole [49]. Moreover, the idea of TV regularization has inspired a myriad of extensions to other tasks such as image debluring, super-resolution, inpainting, compression, rendering, stylization (we refer readers to a recent book [50] and the references therein) as well as other tasks beyond the context of images such as change-point detection, semisupervised learning and graph partitioning.

In this chapter, we focus on the *non-parametric statistical estimation* problem behind TV-denoising which aims to estimate a function $f : [0, 1] \to \mathbb{R}$ using observations of the following form:

$$y_i = f(x_i) + \epsilon_i, i \in [n] := \{1, \dots, n\}, \tag{4.3}$$

where ϵ_i are iid $N(0, \sigma^2)$ and the function f belongs to some fixed non-parametric function class $\mathcal{F}$. The exogenous variables x_i belongs to some subset $\mathcal{X}$ of $\mathbb{R}$. Similar to Chapter 2, we take $\mathcal{F}$ to be the Total Variation class: $\{f | \mathrm{TV}(f) \leq C_n\}$ or its discrete counterpart

$$\mathcal{F}(C_n) := \left\{ f \middle| \sum_{t=2}^{n} |f(x_t) - f(x_{t-1})| \leq C_n \right\}.$$

We are interested in finding algorithms that generate estimates $\hat{y}_t, t \in [n]$ such that the total square error

$$R_n(\hat{y}, f) := \sum_{t=1}^{n} \mathbb{E}[(\hat{y}_t - f(x_t))^2], \tag{4.4}$$

is minimized. Throughout this chapter, when we refer to *rate*, we mean the growth rate of R_n as a function of n and C_n. The family $\mathcal{F}(C_n)$ we consider here features a rich class of functions that exhibit spatially heterogeneous smoothness behavior. These functions can be very smoothly varying in certain regions of space, while in other regions, it can exhibit fast variations (see for eg. Fig. 4.5) or abrupt changes that may even be discontinuous. A good estimator should be able to detect such local fluctuations (which can be short lived) and adjust the amount of "smoothing" to apply according to the level of smoothness of the functions in each local neighborhood. Such estimators are referred as *locally adaptive estimators* by Donoho [2].

We are interested in algorithms that achieve the minimax optimal rates for estimating functions in $\mathcal{F}(C_n)$ defined as:

$$R_n^*(C_n) = \inf_{\{\hat{y}_t\}_{t=1}^n} \sup_{f \in \mathcal{F}(C_n)} R_n(\hat{y}, f),$$

which is known to be $\Theta(n^{1/3} C_n^{2/3})$[20, 51].

There is a body of work in *Strongly Adaptive online learning* that focuses on designing online algorithms such that its regret in any local time window is controlled [24]. Hence the notion of local adaptivity is built into such algorithms. This makes the problem of estimating TV bounded functions, a natural candidate to be amenable to techniques from Strongly Adaptive online learning. However, it is not clear that whether using Strongly Adaptive algorithms can lead to minimax optimal estimation rates. By formalizing the intuition above, we answer it affirmatively in this work.

We reserve the phrase *adaptive estimation* to describe the act of estimating TV bounded functions such that R_n of the estimator/algorithm can be bounded by a function of n and C_n without any prior knowledge of C_n. An *adaptively optimal* estimator $\hat{y}$ is able to estimate an arbitrary function f with an error

$$R_n(\hat{y}, f) = \tilde{O}\Big(\inf_{C_n \text{ such that } f \in \mathcal{F}(C_n)} R_n^*(C_n) \Big).$$

A TV bounded function will be referred as a Bounded Variation (BV) function henceforth for brevity.The notation $\tilde{O}(\cdot)$ hides poly-logarithmic factors of n.

It is well known that *all linear estimators* that output a linear transformation of the observations attain a suboptimal $\Omega(\sqrt{nC_n})$ rate [20]. This covers a large family of algorithms including the popular methods based on smoothing kernels, splines and local polynomials, as well as methods such as online gradient descent [37]. Wavelet smoothing [2] is known to attain the near minimax optimal rate of $\tilde{O}(n^{1/3}C_n^{2/3})$ for R_n without any prior information about C_n. Recently the same rate is shown to be achievable for the online forecasting setting by adding a wavelets-based adaptive restarting schedule to OGD [37].

In this chapter, we provide an alternative to wavelet smoothing by a novel reduction to a strongly adaptive regret minimization problem from the online learning literature. We show that the resulting algorithm achieves the same adaptive optimal rate of $\tilde{O}(n^{1/3}C_n^{2/3})$. The algorithm is more versatile than wavelet smoothing for three reasons:

1. Our algorithm is based on aggregating experts that performs local predictions. The experts we use perform online averaging. However, one may use more advanced algorithms such as kernel/spline smoothing, polynomial regression or even deep learning approaches as experts that can potentially lead to better performance in practice. Hence our algorithm is highly configurable.

2. Our algorithm accepts a learning rate parameter that can be set without prior knowledge of C_n to obtain the near optimal rate of $\tilde{O}(n^{1/3}C_n^{2/3})$ (see Theorem 27). However, this learning rate can also be tuned using heuristics that can lead to better practical performance (see Section 4.5).

3. It can also handle a more challenging setting where the data are streamed sequentially in an online fashion.

To the best of our knowledge, we are the first to formalize the connection between strongly adaptive online learning and the problem of local-adaptivity in nonparametric regression. By establishing this new perspective, we hope to encourage further collaboration between these two communities.

4.1 Setup, Assumptions and Contributions

4.1.1 Problem Setup

Though we are primarily motivated to solve the offline/batch estimation problem, our starting point is to consider a significant generalization of the batch problem as shown in Fig. 4.1. Any adaptively optimal algorithm to this online game immediately implies adaptive optimality in the batch/offline setting. For example, to solve the batch problem, adversary can be thought of as revealing the indices isotonically, i.e $i_t = t$. However, note that in the online game, adversary can even query the same index multiple times. The term "forecasting strategy" in step 1 of Fig. 4.1, is used to mean an algorithm that makes a prediction at current time point only based on the historical data.

Solving the online problem has an added advantage that the resulting algorithm can be applied to various instances of time series forecasting like financial markets, spread of contagious disease etc.

1. Player (we) declares a forecasting strategy
2. Adversary chooses an $\mathcal{X} = \{x_1 < x_2 < \ldots < x_n\}$ and reveals it to the player.
3. Adversary chooses $f(x_1), \ldots, f(x_n)$ such that $\sum_{t=2}^{n} |f(x_t) - f(x_{t-1})| \leq C_n$.
4. Adversary fixes an ordered set $\{i_1, \ldots, i_n\}$ where each $i_j \in [n]$.
5. For every time point $t = 1, ..., n$:
   (a) Adversary reveals i_t.
   (b) We play $\hat{y}_t$.
   (c) We receive a feedback
   $y_t = f(x_{i_t}) + \epsilon_t$,
   where ϵ_t is $N(0, \sigma^2)$.
   (d) We suffer loss $(\hat{y}_t - y_t)^2$
6. Our goal is to minimize
   $\sum_{t=1}^{n} \mathbb{E}[(\hat{y}_t - f(x_{i_t}))^2]$.

Figure 4.1: *Online interaction protocol*

Assumption 1 $|f(x_i)| \leq B, \forall i \in [n]$ for some known B.

Though this constraint is considered to be mild and natural, we note that standard non-parametric regression algorithms do not make this assumption.

4.1.2 Notes on novelty and contributions

To the best of our knowledge, in non-parametric regression literature, only wavelet smoothing [1] [2] is able to *provably* attain a near optimal $\tilde{O}(n^{1/3}C_n^{2/3})$ rate for estimating BV functions in batch setting without knowing the value of C_n. There are model-selection techniques based on information-criterion, which often either incurs significant practical overhead or comes with no optimal rate guarantees.

The contribution of this work is mainly theoretical. Our primary result is a novel

[1] Though [37] proposes a minimax policy for forecasting TV bounded sequences online, they heavily rely on the adaptive minimaxity of wavelet smoothing.

reduction from the problem of estimating BV functions to Strongly Adaptive online learning [24]. This reduction approach results in the development of a new $O(n \log n)$ time algorithm that is: 1) *minimax optimal* (modulo log factors) 2) *adaptive* to C_n and 3) can be used to tackle *both* online and offline estimation problems thereby providing new insights. To elaborate slightly, this is facilitated by few fundamentally different viewpoints than those adopted in the wavelet literature. In particular, we exhibit a specific partitioning of TV bounded function into consecutive chunks that incurs low total variation such that total number of chunks is $O(n^{1/3}C_n^{2/3})$. Then by designing a strongly adaptive online learner, we ensure an $\tilde{O}(1)$ cumulative squared error in each chunk of that partition. This immediately implies an estimation error rate of $\tilde{O}(n^{1/3}C_n^{2/3})$ when summed across all chunks. To the best of our knowledge, this is the *first* time a connection between strongly adaptive online learning and estimating BV functions has been exploited in literature.

Experimental results (see Section 4.5) indicate that our algorithm can outperform wavelet smoothing in terms of its cumulative squared error incurred in practice. We demonstrate that the proposed algorithm can be used without any hyper-parameter tuning and incurs very low computational overhead in comparison to model selection based approaches for the fused lasso problem (see Eq. (4.2)).

Before closing this section, we remind the reader that this work shouldn't be viewed only as providing yet another solution to a classical problem but rather one that provides *a fundamentally new set of tools* that adds new insight to this decades-old problem that might have a profound impact in many extensions of the basic setting we consider and other downstream tasks such as estimating higher-dimensional BV functions, fused lasso on graphs, image deblurring, trend filtering and so on.

4.2 Preliminaries

In this section, we briefly review the elements from online learning literature that are crucial to the development of our algorithm.

4.2.1 Geometric Cover

Geometric Cover (GC) proposed in [24] is a collection of intervals that belong to $\mathbb{N}$ defined below. In what follows $[a, b]$ denotes the set of natural numbers lie between a and b, both inclusive.

$$\mathcal{I} = \bigcup_{k \in \mathbb{N} \cup \{0\}} \mathcal{I}_k, \tag{4.5}$$

where $\forall k \in \mathbb{N} \cup \{0\}$, and $\mathcal{I}_k = \{[i \cdot 2^k, (i+1) \cdot 2^k - 1] : i \in \mathbb{N}\}$. Define AWAKE$(t) :=$ $\{I \in \mathcal{I} : t \in I\}$.By the construction of Geometric Cover $\mathcal{I}$, it holds that

$$|\text{AWAKE}(t)| = \lfloor \log t \rfloor + 1. \tag{4.6}$$

Let's denote $\mathcal{I}|_J := \{I \in \mathcal{I} : I \subseteq J\}$ for an interval $J \subseteq \mathbb{N}$. The GC has a very nice property recorded in the following Proposition.

Proposition 23. *[24] Let $I = [q, s] \subseteq \mathbb{N}$. Then the interval I can be partitioned into two finite sequences of disjoint consecutive intervals $(I_{-k}, \ldots, I_0) \subseteq \mathcal{I}|_I$ and $(I_1, \ldots, I_p) \subseteq \mathcal{I}|_I$ such that,*

$$\frac{|I_{-i}|}{|I_{-i+1}|} \leq \frac{1}{2}, \forall i \geq 1 \qquad \textit{and} \qquad \frac{|I_i|}{|I_{i-1}|} \leq \frac{1}{2}, \forall i \geq 2.$$

4.2.2 Sleeping Experts and Specialist Aggregation Algorithm (SAA)

In the problem of learning from expert advice with outcome space $\mathcal{O}$ and action space $\mathcal{A}$, there are K experts who provide a list of actions $a_{t,:} = [a_{t,1}, ..., a_{t,K}] \in \mathcal{A}^K$ at time $t = 1, ..., n$. The learner is supposed to takes an action $a_t \in \mathcal{A}$ based on the expert advice[2] before the outcome $o_t \in \mathcal{O}$ is revealed by an adversary. The player then incurs a loss given by $\ell(a_t, o_t)$, where ℓ is a loss function.

In the most basic setting, $\mathcal{A}, \mathcal{O}$ are discrete sets, ℓ can be described by a table, and we assign one constant expert to each $a \in \mathcal{A}$, then this becomes an online version of Von Neumann's linear matrix game. More generally, $\mathcal{A}$ can be a convex set, describing parameters of a classifier, $o \in \mathcal{O}$ could denote a feature-label pair in which case the loss could be a square loss or logistic loss that measures the performance of each classifier.

Our result leverages a variant of the learning from expert advice problem which assumes an arbitrary subset of K experts might be sleeping at time t and the learner needs to compete against an expert only during its awake duration. The learner chooses a distribution $\boldsymbol{w}_t$ over the awake experts and plays a weighted average over the actions of those awake experts. It then incurs a surrogate-loss called "MixLoss" which is a measure of how good the distribution $\boldsymbol{w}_t$ is. (See Figure 4.2 for details.) This setting is different from the classical prediction with experts advice problem in two aspects: 1) The adversary is endowed with more power of selecting an awake expert set in addition to the actual outcome o_t at each round. 2) Instead of the loss $\ell(a_t, o_t)$, the learner is incurred a surrogate loss on the distribution chosen by the learner at time t.

Consider the protocol of learning with sleeping experts shown in Fig. 4.2. Assume an expert pool of size K.

[2] Could be $a_{t,k}$ for some $k \in [K]$ or any other points in $\mathcal{A}$

For $t = 1, \ldots, n$

1. Adversary picks a subset $A_t \subset [K]$ of awake experts.
2. Learner choose a distribution $\boldsymbol{w}_t$ over A_t.
3. Adversary reveals loss of all *awake* experts, $\boldsymbol{\ell}_t \in (-\infty, \infty]^{|A_t|}$.
4. Learner suffers MixLoss: $-\log(\sum_{k \in A_t} w_{t,k} e^{-\ell_{t,k}})$.

Figure 4.2: *Interaction protocol with sleeping experts. The expert pool size is K.*

Initialize $u_{1,k} = 1/|\mathcal{S}|$ for all k in an index set $\mathcal{S}$ used to index the expert pool.
For $t = 1, \ldots, n$

1. Adversary reveals $A_t \subseteq \mathcal{S}$.
2. Play weighted average action wrt distribution: $w_{t,k} = \frac{u_{t,k} \mathbf{1}\{k \in A_t\}}{\sum_{j \in A_t} u_{t,j}}$.
3. Broadcast the weights $w_{t,k}$.
4. Receive losses $\ell_{t,k}$ for all $k \in A_t$.
5. Update:
   - $u_{t+1,k} = \frac{u_{t,k} e^{-\ell_{t,k}}}{\sum_{j \in A_t} u_{t,j} e^{-\ell_{t,j}}} \sum_{j \in A_t} u_{t,j}$ if $k \in A_t$.
   - $u_{t+1,k} = u_{t,k}$ if $k \notin A_t$.

Figure 4.3: *Specialist Aggregation Algorithm (SAA).*

Lemma 24. *[52] Regret R_n^j of SAA (Fig. 4.3) w.r.t. any fixed expert $j \in [K]$ satisfies,*

$$R_n^j := \sum_{t \in [n]} \mathbf{1}\{j \in A_t\} \left(-\log(\sum_{k \in A_t} w_{t,k} e^{-\ell_{t,k}}) - \ell_{t,j} \right) \leq \log K,$$

where $\mathbf{1}\{\cdot\}$ is the indicator function, $\ell_{t,k} := \mathcal{L}(a_{t,k}, o_t)$ and $a_{t,k}$ is the action taken by expert k at time t.

Note that $\ell_{t,j} = \text{MixLoss}(\boldsymbol{e}_j)$ where $\mathbf{e}_j$ selects j with probability 1. The regret measures the performance of the learner against any fixed expert in terms of the MixLoss in the sub-sequence where she is awake.

Definition 25. $\mathcal{L}(a,x)$ is η exp-concave in a for each x if $\sum_{k=1}^K w_k e^{-\eta \mathcal{L}(a_k,x)} \leq e^{-\eta \mathcal{L}(\sum_{k=1}^K w_k a_k, x)}$, for $w_k \geq 0$ and $\sum_{k=1}^K w_k = 1$.

A MixLoss regret bound is useful because it implies a regret bound on any exp-concave losses for learners playing the weighted average action $a_t = \sum_{k \in A_t} w_{t,k} a_{t,k}$. To see this, let $\mathcal{L}'(a,o)$ be η exp-concave in its first argument $a \in \mathcal{A}$. By the definition of exp-concavity it follows that if SAA is run with losses $\mathcal{L}(a,o) = \eta\mathcal{L}'(a,o)$, then,

$$\sum_{t \in [n]: j \in A_t} \left(\eta \mathcal{L}' \left(\sum_{k \in A_t} w_{t,k} a_{t,k} \ , \ o_t \right) - \eta \mathcal{L}'(a_{t,j}, o_t) \right) \leq R_n^j,$$

where $a_{t,k}$ is the action taken by expert k at time t.

We refer to Chapter 3 of [40] and [52] for further details on SAA.

4.3 Main Results

In this section, we present our algorithm and its performance guarantees.

4.3.1 Algorithm

Our goal is to explore the possibility that a Strongly Adaptive online learner can lead to minimax optimal estimation rate. Consequently the algorithm that we present is a fairly standard Strongly Adaptive online learner that can guarantee logarithmic regret in any interval.

Our algorithm Aligator (**A**ggregation of on**LI**ne avera**G**es using **A** geome**T**ric c**O**ve**R**) defined in Fig.4.4 can be used to tackle both online and batch estimation problems. The policy is based on learning with sleeping experts where expert pool is defined as follows.

Definition 26. The expert pool is $\mathcal{E} = \{\mathcal{A}_I : I \in \mathcal{I}|_{[n]}\}$, where $\mathcal{I}|_{[n]}$ is as defined in Section 4.2.1 and $\mathcal{A}_I$ is an algorithm that perform online averaging in interval I. Let $\mathcal{A}_I(t)$ denote the prediction of the expert $\mathcal{A}_I$ at time t, if $I \in \text{AWAKE}(t)$.

Due to relation (4.6), we have $|\mathcal{E}| \leq n \log n$. Our policy basically performs SAA over $\mathcal{E}$.

ALIGATOR:Inputs - time horizon n, learning rate η

1. Initialize SAA weights $u_{1,I} = 1/|\mathcal{E}|, \forall I \in \mathcal{I}|_{[n]}$.
2. For $t = 1$ to n:
   (a) Adversary reveals an arbitrary $x_{i_t} \in \mathcal{X}$.
   (b) Let $A_t = \text{AWAKE}(i_t)$. Pass A_t to SAA.
   (c) Receive $w_{t,I}$ from SAA for each $I \in A_t$.
   (d) Predict $\hat{y}_t = \sum_{I \in A_t} w_{t,I} \mathcal{A}_I(t)$.
   (e) Receive $y_t = f(x_{i_t}) + \epsilon_t$.
   (f) Pass losses $\ell_{t,I} = \eta(y_t - \mathcal{A}_I(t))^2$,
   for each $I \in A_t$ to the SAA.

Figure 4.4: *The* ALIGATOR *algorithm*

The precise definition of $\mathcal{A}_I(t)$ used in our algorithm is

$$\mathcal{A}_I(t) = \begin{cases} \frac{\sum_{s=1}^{t-1} y_s \mathbf{1}\{i_s \in I\}}{\sum_{s=1}^{t-1} \mathbf{1}\{i_s \in I\}} & \text{if } \sum_{s=1}^{t-1} \mathbf{1}\{i_s \in I\} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{4.7}$$

where i_s is the index of the exogenous variable x_{i_s} in step 2(a) of Fig. 4.4. This particular choice of experts is motivated by the fact that performing online averages lead to logarithmic static regret under quadratic losses. As shown later, this property when combined with the SAA scheme leads to logarithmic regret in *any* interval of $[n]$.

4.3.2 Performance Guarantees

Theorem 27. *Consider the online game in Fig. 4.1. Let* $\theta_t := f(x_{i_t})$. *Under Assumption 1, with probability atleast* $1-\delta$, ALIGATOR *forecasts* $\hat{y}_t$ *obtained by setting* $\eta = \frac{1}{8\left(B+\sigma\sqrt{\log(2n/\delta)}\right)^2}$, *incurs a cumulative error*

$$\sum_{t=1}^{n}(\hat{y}_t - \theta_t)^2 = \tilde{O}(n^{1/3}C_n^{2/3}),$$

where $\tilde{O}(\cdot)$ *hides the dependency of constants* B, σ *and poly-logarithmic factors of* n *and* δ.

Proof Sketch. We first show that ALIGATOR suffers logarithmic regret against any expert in the pool $\mathcal{E}$ during its awake period. Then we exhibit a particular partition of the underlying TV bounded function such that number of chunks in the partition is $O(n^{1/3}C_n^{2/3})$ (Lemma 157 in Appendix C.1). Following this, we cover each chunk with atmost $\log n$ experts and show that each expert in the cover suffers a $\tilde{O}(1)$ estimation error. The Theorem then follows by summing the estimation error across all chunks of the partition. In summary, the delicate interplay between Strongly Adaptive regret bounds and properties of the partition we exhibit leads to the adaptively minimax optimal estimation rate for ALIGATOR. We emphasize that existence of such partitions is a highly non-trivial matter. □

Remark 28. We note that under the above setting, ALIGATOR is minimax optimal in n and C_n, and adaptive to unknown C_n.

Remark 29. If the noise level σ is unknown, it can be robustly estimated from the wavelet coefficients of the observed data by a Median Absolute Deviation estimator [18]. This is facilitated by the sparsity of wavelet coefficients of BV functions .

Remark 30. In the offline problem where we have access to all observations ahead of time, the choice of $\eta = 1/(8\hat{\nu}^2)$ where $\hat{\nu} = \max\{|y_1|, \ldots, |y_n|\}$ results in the same near optimal rate for R_n as in Theorem 27. This is due to the fact that $B + \sigma\sqrt{\log(2n/\delta)}$ is nothing but a high probability bound on each $|y_t|$. Hence we don't require the prior knowledge of B and σ for the offline problem.

Remark 31. The authors of [2] use the error metric given by the L2 function norm in a compact interval $[0, 1]$ defined as $\int_0^1 \left(\hat{f}(x) - f(x)\right)^2 dx$ in an offline setting, where $\hat{f}(x)$ is the estimated function. A common observation model for non-parametric regression considers $x_{i_t} = t/n$ [3]. When $x_{i_t} = t/n$, Aligator guarantees that the empirical norm $\frac{1}{n}\sum_{t=1}^{n} (\hat{y}_t - f(t/n))^2$ decays at the rate of $\tilde{O}\left(n^{-2/3}C_n^{2/3}\right)$. For the TV class, it can be shown that the empirical norm and the function norm are close enough such that the estimation rates do not change (see Section 15.5 of [18]).

Remark 32. Note that conditioned on the past observations, the prediction of Aligator is deterministic in each round. So in the online setting, we can compete with an adversary who chooses the underlying ground truth in an adaptive manner based on the learner's past moves. With such an adaptive adversary, it becomes important to reveal the set of covariates $\mathcal{X}$ ahead of time. Otherwise there exists a strategy for the adversary to choose the covariates x_{i_t} that can enforce a linear growth in the cumulative squared error. We refer the readers to [53] for more details about such adversarial strategy.

Proposition 33. *The overall run-time of* Aligator *is* $O(n \log n)$.

Proof. On each round $|\text{AWAKE}(t)|$ is $O(\log n)$ by (4.6). So we only need to aggregate and update the weights of $O(\log n)$ experts per round which can be done in $O(\log n)$ time. □

4.4 Extensions

Motivated from a practical perspective, we discus two direct extensions to ALIGATOR below. These extensions highlight the versatility of ALIGATOR in adapting to each application.

Hedged Aligator. In our theoretical results, we found that choosing learning rate η conservatively according to Theorem 27 or Remark 30 ensures the minimax rates. In practice, however, one could use larger learning rates to adapt to the structure of every input sequence.

We propose to use a hedged ALIGATOR scheme that aggregates the predictions of ALIGATOR instantiated with different learning rates. In particular, we run different instances of ALIGATOR in parallel where an instance corresponds to a learning rate in the exponential grid $[\eta, 2\eta, \ldots, \max\{\eta, \log_2 n\}]$ which has a size of $O\left(\log\left((B^2+\sigma^2)\log n\right)\right)$. Here η is chosen as in Theorem 27 or Remark 30. Then we aggregate each of these instances by the Exponential Weighted Averages (EWA) algorithm [40]. The learning rate of this outer EWA layer is set according to the theoretical value.By exp-concavity of squared error losses, this strategy helps to match the performance of the best ALIGATOR instance. Since the theoretical choice of learning rate is included in the exponential grid, the strategy can also guarantee optimal minimax rate. We emphasize that Hedged ALIGATOR is adaptive to C_n and requires no hyper-parameter tuning.

Aligator with polynomial regression experts. This extension is motivated by the problem of identifying trends in time series. Though in Section 4.3.1 we use online averaging as experts, in practice one can consider using other algorithms. For example, if the trends in a time series are piecewise-linear, then experts based on online averaging can lead to poor practical performance because the TV budget C_n of piecewise linear signals can be very large. To alleviate this, in this extension, we propose to use Online

Polynomial Regression as experts where a polynomial of a fixed degree d is fitted to the data with time points as its exogenous variables. This is similar to the idea adopted in [54] where they construct a policy that performs restarted online polynomial regression where the restart schedule is adaptively chosen via wavelet based methods. They show that such a scheme can guarantee estimation rates that grow with (a scaled) L1 norm of higher order differences of the underlying trend which can be much smaller than its TV budget C_n. This extension can be viewed as a variant to the scheme in [54] where the "hard" restarts are replaced by "soft restarts" via maintaining distributions over the sleeping experts.

4.5 Experimental Results

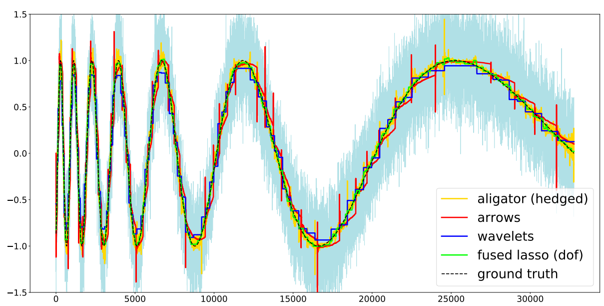


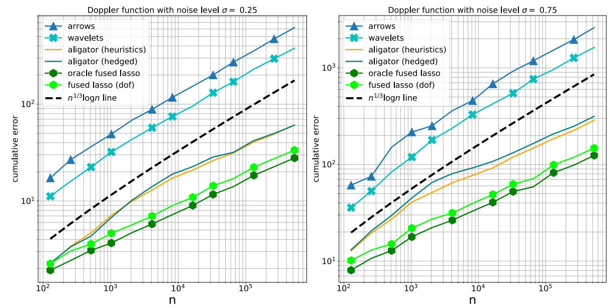
Figure 4.5: *Fitted signals for Doppler function with noise level* $\sigma = 0.25$

For empirical evaluation, we consider online and offline vesrions of the problems separately.

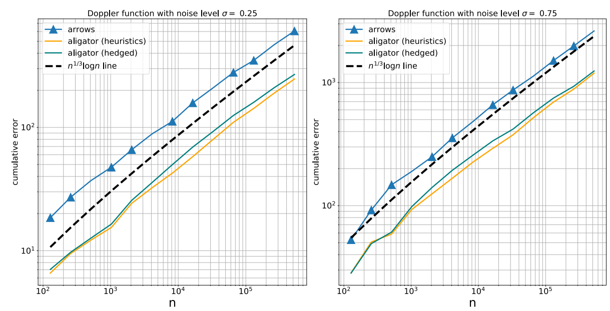
Description of policies. We begin by a description of each algorithm whose error curve is plotted in the figures.

ALIGATOR *(hedged):* This is the extension described in Section 4.4

ALIGATOR *(heuristics):* For this hueristics strategy, we divide the loss of each expert by $2(\sigma^2 + \sigma^2/m)$ where m is the number of samples whose running average is compued by



(a) Offline experiments



(b) Online experiments

Figure 4.6: *Cumulative squared error rate of various algorithms on offline setting and online setting.* Aligator *achieves the optimal* $\tilde{O}(n^{1/3})$ *rate while performing better than wavelet based methods. In particular, in the offline setting, it achieves a performance closer to that of dof based fused lasso while only incurring a cheap* $\tilde{O}(n)$ *run-time overhead.*

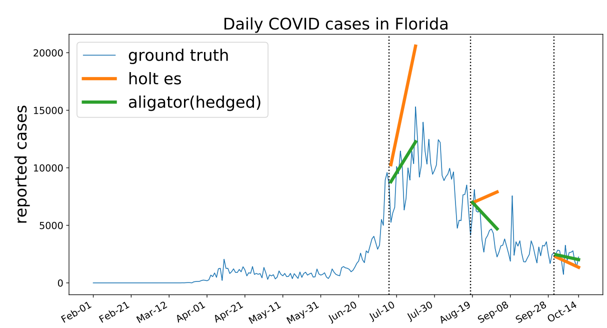


Figure 4.7: *A demo on forecasting COVID cases based on real world data. We display the two weeks forecasts of hedged* Aligator *and Holt ES, starting from the time points identified by the dotted lines. Both the algorithms are trained on a 2 month data prior to each dotted line. We see that hedged* Aligator *detects changes in trends more quickly than Holt ES. Further, hedged* Aligator *attains a 20% reduction in the average RMSE from that of Holt ES (see Section 4.5).*

the expert. This loss is proportional to the notion of (squared) z-score used in hypothesis testing. Intuitively, lower (squared) z-score corresponds to better experts. The multiplier 2 in the previous expression is found to provide good performnace across all signals we consider.

arrows: This is the the policy presented in [37], which runs online averaging with an adaptive restarting rule based on wavelet denoising results.

wavelets: This is the universal soft thresholding estimator from [2] based on Haar wavelets which is known to be minimax optimal for estimating BV functions.

oracle fused lasso: This estimator is obtained by solving (4.2) whose hyper-parameter is tuned by assuming access to an oracle that can compute the mean squared error wrt actual ground truth. The exact ranges used in the hyper-parameter grid search is described in Appendix C.2. Note that the oracle fused lasso estimator is purely hypothetical due to absence of such oracles described before in reality and is ultimately impractical. It is used here to facilitate meaningful comparisons.

fused lasso (dof): In this experiment, we maintain a list of λ for the fused lasso problem (Eq. (4.2)). Then we compute the Stein's Unbiased Risk estimator for the expected squared error incurred by each λ by estimating its degree of freedom (dof) [55] and select the λ with minimum estimated error.

Experiments on synthetic data. For the ground truth signal, we use the Doppler function of [56] whose waveform is depicted in Fig. 4.5. The observed data are generated by adding iid noise to the ground truth. For offline setting, we have access to all observations ahead of time. So we run Arrows and both versions of Aligator two times on the same data, once in isotonic order (i.e $i_t = t$ in Fig. 4.1) and other in reverse isotonic order and average the predictions to get estimates of the ground truth. For online setting such a forward-backward averaging is not performed. This process of generating the noisy data and computing estimates are repeated for 5 trials and the average cumulative error is plotted. As we can see from Fig.4.6 (a), ALIGATOR versions attains the $\tilde{O}(n^{1/3})$ rate and incurs much lower error than wavelet smoothing. Further, performance of hedged and heuristics versions of ALIGATOR is in the vicinity to that of the hypothetical fusedlasso estimator while the policies arrows and wavelets violate this property by a large margin. Even though the dof based fused lasso comes very close to the oracle counterpart, we emphasize that this strategy is not known to provide theoretical guarantees for its rate and requires heavy computational bottleneck since it requires to solve the fused lasso (Eq. 4.2) for many different values of λ.

For the online version of the problem, we consider the policy Arrows as the benchmark. This policy has been established to be minimax optimal for online forecasting of TV bounded sequences in [37]. We see from Fig.4.6 (b) that all the policies attains an $\tilde{O}(n^{1/3})$ rate while ALIGATOR variants enjoy lower cumulative errors.

Experiments on real data. Next we consider the task of forecasting COVID cases using the extension of Aligator with polynomial regression experts as in Section 4.4. The data are obtained from the CDC website ([57]).

We address a very relevant problem as follows: Given access to the historical data, forecast the evolution of COVID cases for the next 2 weeks. We compare the performance of hedged ALIGATOR and Holt Exponential Smoothing (Holt ES), on this problem, where the later is a common algorithm used in Time Series forecasting to detect underlying trends. For ALIGATOR, we use Online Linear Regression as experts where a polynomial of degree one is fitted to the data with time points as its exogenous variables. For each time point t in [Apr 20, Sep 27], we train both hedged ALIGATOR and Holt ES on a training window of past 2 months. Then we calculate a 2 week forecast for both algorithms. For ALIGATOR this is achieved by linearly extrapolating the predictions of experts awake at time t and aggregating them. Following this, we compute the Root Mean Squared Error (RMSE) in the interval $[t, t+14)$ for both algorithms. These RMSE are then averaged across all t in [Apr 20, Sep 27].

We choose data from the state of Florida, USA, as an illustrative example. We obtained an average RMSE of 1330.12 for hedged ALIGATOR and 1671.77 for Holt ES. Thus hedged ALIGATOR attains a 20% reduction in forecast error from that of Holt ES. A qualitative comparison of the forecasts is illustrated in Fig. 4.7. As we can see, the time series is non-stationary and has a varying degree of smoothness. ALIGATOR is able to adapt to the local changes quickly, while Holt ES fails to do so despite having a more sophisticated training phase. Similar experimental results for some of the other states are reported in Appendix C.2.

The training step of hedged ALIGATOR involves learning the weights of all experts by an online interaction protocol as shown in Fig. 4.1 with $i_t = t$. It is remarkable that *no* hyper-parameter tuning is required by ALIGATOR for its training phase. The slowest learning rate to be used in the grid for hedged ALIGATOR is computed as follows. First we calculate the maximum loss incurred by each expert for a one step ahead forecast in its awake duration. Then we take the maximum of this quantity across all experts in the pool. Let this quantity be β. The slowest learning rate in the grid is then set as $1/(2\beta)$. The learning rate of the outer layer of EWA is also set the same. This is justifiable because the quantity $4\left(B + \sigma\sqrt{\log(2n/\delta)}\right)$ in the denominator of the learning rate in Theorem 27 is a high probability bound on the loss incurred by any expert for a one step ahead forecast.

We defer further experimental results to Appendix C.2.

An important caveat for practitioners. Though ALIGATOR is able to detect non-stationary trends in the COVID data efficiently, we do *not* advocate using ALIGATOR *as is* for pandemic forecasting, which is a substantially more complex problem that requires input from domain experts.

However, ALIGATOR could have a role in this problem, and other online forecasting tasks. Estimating (and removing) trend is an important first step in many time series methods (e.g., Box-Jenkins method). Most trend estimation methods only apply to offline problems (e.g., Hodrick-Prescott filter or L1 Trend Filter) [4], while Holt ES is a common method used for online trend estimation. For instance, Holt ES is being used as a subroutine for trend estimation in a state-of-the-art forecasting method [58] for COVID cases that CDC is currently using. We expect that using Aligator instead in such models that use Holt ES will lead to more accurate forecasting, but that is beyond the scope of this chapter.

4.6 Concluding Discussion

In this work, we presented a novel reduction from estimating BV functions to Strongly Adaptive online learning. The reduction gives rise to a new algorithm Aligator that attains the near minimax optimal rate of $\tilde{O}(n^{1/3}C_n^{2/3})$ in $O(n \log n)$ run-time. The results form a parallel to wavelet smoothing in terms of optimal adaptivity to unknown variational budget C_n. However, our algorithm is more versatile than wavelets in terms of its configurability and practical performance. Further, for offline estimation, Aligator variants achieves a performance closer (than wavelets) to an oracle fused lasso while incurring only an $\tilde{O}(n)$ run-time with no hyper parameter tuning. This is in contrast to degree of freedom based approaches of tuning the fused lasso hyper parameter that requires significantly more computational overhead and is not known to provide guarantees on its rate.

Part II

Theory and Algorithms under Adversarial Observation Model

Chapter 5

Optimal Dynamic Regret in Exp-Concave Online Learning

We consider a generic online learning framework which is modelled as an interactive n step game between a learner and adversary. At each time step t, the learner predicts a $\boldsymbol{p}_t \in \mathcal{D} \subseteq \mathbb{R}^d$. Then the adversary reveals a loss function $f_t : \mathbb{R}^d \to \mathbb{R}$. The objective of the learner is to minimise its regret against a predefined set of strategies $\mathcal{W}$ that is known to the learner before the start of the game. We call a learning algorithm to be *proper* when $\mathcal{D} = \mathcal{W}$. Further when $\mathcal{D} = \mathcal{W}$ are convex sets and the losses f_t are convex in $\mathcal{D}$, the generic learning framework reduces to the one studied in Online Convex Optimization (OCO) [59]. On the other hand, we call the learning algorithm to be *improper* when $\mathcal{D} \supset \mathcal{W}$. A commonly used metric to measure the performance of the learner is its *static* regret defined as

$$R_n = \sum_{t=1}^{n} f_t(\boldsymbol{p}_t) - \inf_{\boldsymbol{w} \in \mathcal{W}} \sum_{t=1}^{n} f_t(\boldsymbol{w}). \tag{5.1}$$

A sub-linear static regret implies that the average loss incurred by the learner converges to that of the best comparator strategy in hindsight.

A canonical example of an improper algorithm can be found in an online linear regression setting where $f_t(\boldsymbol{u}) = (y_t - \boldsymbol{x}_t^T \boldsymbol{u})^2$ with $|y_t| \leq 1$, $\|\boldsymbol{x}_t\|_2 \leq 1$ and we are interested in controlling the static regret against against a set of linear predictors with bounded norm, $\mathcal{W} = \{\boldsymbol{w} \in \mathbb{R}^d : \|\boldsymbol{w}\|_2 \leq 1\}$. One popular learning algorithm in this framework is the Vovk-Azoury-Warmuth (VAW) forecaster [60, 61] (or see Section 11.8 in [40]). The VAW forecaster attains an $O(d \log n)$ static regret against $\mathcal{W}$. However predictions of VAW at time t denoted by $\boldsymbol{u}_t$ may not necessarily satisfy $\|\boldsymbol{u}_t\|_2 \leq 1$ hence making it an improper algorithm.

The notion of static regret is not befitting for non-stationary environments – such as financial markets – where it could be inappropriate to compete against a fixed comparator due to the changes in the dynamics of the environment. The work of [5] introduces the notion of *dynamic* regret defined as

$$R^n_{\boldsymbol{w}_1,\ldots,\boldsymbol{w}_n} := \sum_{t=1}^{n} f_t(\boldsymbol{p}_t) - f_t(\boldsymbol{w}_t), \tag{5.2}$$

for *any* sequence of comparators $\boldsymbol{w}_t$ in $\mathcal{W}$. The dynamic regret bounds are usually expressed in literature as a function of number of time steps and some path variation metric that captures the degree of non-stationarity in the comparator sequence. In this chapter, we study the following path variation:

$$TV(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n) := \sum_{t=2}^{n} \|\boldsymbol{w}_t - \boldsymbol{w}_{t-1}\|_1.$$

The maximum dynamic regret against all comparator sequences whose path variation is bounded by a number C_n can then be defined as

$$R_n(C_n) := \sup_{\substack{\boldsymbol{w}_1,\dots,\boldsymbol{w}_n \\ TV(\boldsymbol{w}_1,\dots,\boldsymbol{w}_n) \le C_n}} R^n_{\boldsymbol{w}_1,\dots,\boldsymbol{w}_n}. \tag{5.3}$$

There is a complementary body of work on Strongly Adaptive (SA) algorithms [24] where the static regret in *any* sub-interval of $[n] := \{1,\dots,n\}$ is controlled. Hence SA algorithms have the nice property of being globally and locally optimal. The work of [27] exploits this property of SA algorithms to control their dynamic regret in terms of a variational metric that measures how much the losses f_t change over time. In particular, whenever the losses have extra curvature properties such as strong convexity or exp-concavity, they show that one can get fast dynamic regret rates. However, it was unclear if SA methods can lead to optimal dynamic regret guarantees in terms of the path length of the comparator sequence — an open question raised in [27].

The works of [26] and [62] attains a dynamic regret of $O^*(\sqrt{n(1+C_n)})$ and $O^*(\sqrt{nC_n} \vee \log n)$ respectively, where $O^*(\cdot)$ hides dependence on the dimension and $(a \vee b) = \max\{a, b\}$. However, we show a lower bound of $\Omega^*(n^{1/3}C_n^{2/3} \vee \log n)$ in Proposition 44 applicable to the case when losses are strongly convex / exp-concave. Hence, there is a large gap between this lower bound and existing upper bounds. In this work, we show that whenever improper learning is allowed and when the loss functions are strongly convex / exp-concave, one can leverage SA algorithms to attain the *sharp* rate of $\tilde{O}^*(n^{1/3}C_n^{2/3} \vee \log n)$ for $R_n(C_n)$ where $\tilde{O}^*(\cdot)$ hides dependence in the dimension and factors of $\log n$ (see section 5.2 for formal statements and complete list of assumptions). Further, the SA algorithms need not require the apriori knowledge of C_n to attain this rate.

As a concrete use case, we show that our results have interesting implications to the problem of *online* Total Variation (TV) denoising. The offline version of TV-denoising problem has seen many influential applications in the signal processing community. For example, algorithms that use TV-regularization has been deployed in every cellphone, digital camera and medical imaging devices (we refer readers to the book [50] and the references therein) as well as other tasks beyond the context of images such as change-point detection, semisupervised learning and graph partitioning.

We proceed to formally introduce the non-paramteric regression problem behind TV-denoising. Define a non-parametric class of TV bounded sequences as

$$\mathcal{TV}(C_n) := \left\{ (w_1, \dots, w_n) : \sum_{t=2}^{n} |w_t - w_{t-1}| \leq C_n \right\},$$

where $\sum_{t=2}^{n} |w_t - w_{t-1}|$ is termed as the TV of the sequence $w_{1:n} := (w_1, \dots, w_n)$. In the offline TV-denoising problem we are given n observations of the form $y_t = w_t + \epsilon_t$ where ϵ_t are iid zero mean subgaussian noise, $t \in [n]$ and $w_{1:n}$ is an unknown sequence in $\mathcal{TV}(C_n)$. We are interested in coming up with estimates $\hat{w}_t$ such that $R^{\mathcal{TV}}(C_n) := \mathbb{E}\left[\sum_{t=1}^{n} (\hat{w}_t - w_t)^2\right]$ is controlled. Several non-parametric regression algorithms such as Trend Filtering [3] are known to achieve a near minimax optimal rate of $\tilde{O}(n^{1/3} C_n^{2/3})$ for $R^{\mathcal{TV}}(C_n)$ where $\tilde{O}(\cdot)$ hides dependence on factors of $\log n$.

We can instantiate an online version of the above non-parametric regression problem behind TV-denoising into our learning framework with slight modifications. We consider a TV class with bounded sequences

$$\mathcal{TV}^B(C_n) := \left\{ w_{1:n} : \sum_{t=2}^{n} |w_t - w_{t-1}| \leq C_n, \ |w_t| \leq B \ \forall t \in [n] \right\}. \tag{5.4}$$

When viewed through our online learning framework, we take $f_t(x) = (y_t - x)^2$ where $|y_t| \leq B$, $\mathcal{D} = \mathcal{W} = [-B, B]$. Labels $y_{1:n}$ is a fixed sequence in contrast to the stochastic noise setting discussed earlier, and we are hoping to compete with the best approximation from sequences in $\mathcal{TV}^B(C_n)$ for all $C_n \geq 0$ at the same time. We remark that to compete with the entire $\mathcal{TV}(C_n)$ class it is sufficient to compete with $\mathcal{TV}^B(C_n)$ due to the property $|y_t| \leq B$. We show in Section 5.1 that by using appropriate SA algorithms, one can attain a dynamic regret of $R_n(C_n) = \tilde{O}(n^{1/3}C_n^{2/3})$. This in turn implies the minimax estimation rate in the iid stochastic setting. Further our results have the added advantage of providing an oracle inequality. We conclude this section by summarizing our key contributions below.

- We show that Follow-the-Leading-History (FLH) algorithm [23] with Follow The Leader (FTL) as base learners can achieve the optimal minimax regret (modulo $\log n$ factors) of $\tilde{O}(n^{1/3}C_n^{2/3}B^{4/3} \vee B^2 \log n)$ for the problem of online non-parametric regression with TV bounded sequences – $\mathcal{TV}^B(C_n)$ – as the reference class. The policy is *adaptive* to the TV budget C_n. Further, we demonstrate that the same policy is minimax optimal for smoother non-parametric sequence classes such as Sobolev class or Holder class.

- When improper learning is allowed and when the loss functions revealed by the adversary are exp-concave, strongly smooth and Lipschitz on a box that encloses the set of comparators $\mathcal{W}$, (see Section 5.2) we show that FLH with ONS as base learners attains a dynamic regret of $\tilde{O}\left(d^{3.5}(n^{1/3}C_n^{2/3} \vee 1)\right)$ when $C_n \geq 1/n$ and $O(d^{1.5}\log n)$ otherwise, without prior knowledge of C_n – the path variation of the *comparator sequence*. This rate is shown to be minimax optimal modulo polynomial factors of $\log n$ and d.

- The proof of the regret bound is facilitated by exploiting a number of distinct structures of primal and dual variables in KKT conditions of the optimization problem solved by the offline oracle. We believe that this style of analysis can be useful in bounding the regret of online algorithms in a broader context.

5.1 Performance guarantees for squared error losses

In this section, we focus on the online TV-denoising problem which is a special case of our online learning framework with squared error losses. This will help to build the intuitions behind the analysis for general exp-concave losses as well. All unspecified proofs of this section are deferred to Appendix D.2. We consider the following interaction protocol.

- At time $t \in [n]$ learner predicts $x_t \in \mathcal{D} = [-B, B]$.
- Adversary reveals a label $y_t \in [-B, B]$.
- Learner suffers loss $(y_t - x_t)^2$.

We define the comparator class as the set of TV bounded sequences that takes values in $\mathcal{W} = [-B, B]$ as in Eq.(5.4). The performance of the learner is measured using dynamic regret against the sequences that belongs to $\mathcal{TV}^B(C_n)$, for all $C_n > 0$ simultaneously.

The main SA method that we will be relying on throughout this chapter is the FLH algorithm from [23]. We provide a description of this algorithm in Appendix D.1 for completeness. We have the following regret guarantee for FLH with Follow-the-Leader (FTL) as base learners (in this case, FTL is equivalent to simple online averaging).

Theorem 34. *Let x_t be the prediction at time t of* FLH *with learning rate $\zeta = 1/(8B^2)$ and base learners as FTL. Then for any compararator $(w_1, \ldots, w_n) \in \mathcal{TV}^B(C_n)$*

$$\sum_{t=1}^{n} (y_t - x_t)^2 - (y_t - w_t)^2 = \tilde{O}\left(n^{1/3} C_n^{2/3} B^{4/3} \vee B^2\right), \tag{5.5}$$

where the labels obey $|y_t| \leq B$, $\tilde{O}(\cdot)$ hides dependence on logarithmic factors of horizon n and $a \vee b := \max\{a, b\}$.

Remark 35 (Adaptivity to C_n and (non-stochastic) oracle inequality)**.** We remark that FLH-FTL does not require C_n as an input thus Theorem 34 implies the following oracle inequality

$$\sum_{t=1}^{n}(y_t - x_t)^2 \leq \min_{w_1,\ldots,w_n} \sum_{t=1}^{n}(y_t - w_t)^2 + \tilde{O}\left(n^{1/3}\mathrm{TV}(w_{1:n})^{2/3}B^{4/3} \vee B^2\right).$$

Such result is not known for any algorithm even in the offline case when $y_1, ..., y_n$ is known. Notice that w_t does not need to be constrained because $-B \leq y_t \leq B$.

The strongest oracle inequality for TV-denoising to our knowledge is that of [39, 63], which shows that the fused-lasso estimator with tuning parameter λ obeys $\sum_{t=1}^{n}(y_t - x_t)^2 \leq \min_{w_1,\ldots,w_n} \sum_{t=1}^{n}(y_t - w_t)^2 + O\left(\lambda \mathrm{TV}(w_{1:n})\right)$, under additional stochastic assumptions of y_t. Our results eliminate the need to choose hyperparameter λ all together and achieve the same rate achievable by the optimal choice of λ.

For the sake of clarity we next present the strategy we adopt for proving Theorem 34. We also highlight the main technical challenges that are needed to be overcome along the way. This is followed by some useful lemmas and proof of the main theorem in Section 5.1.2.

5.1.1 Proof strategy for Theorem 34

Let $u_1, \ldots, u_n$ be the *offline optimal* sequence (see Lemma 36) in $\mathcal{TV}^B(C_n)$ which attains the minimum cumulative squared error loss. Note that this offline optimal can depend on the entire sequence of labels $y_1, \ldots, y_n$ chosen by the adversary.

Consider a partitioning of $[n]$ into M sub-intervals $\{[i_s, i_t]\}_{i=1}^{M}$. We will also use the number i to refer to the interval $[i_s, i_t]$. For the interval i, define the quantities: $n_i = i_t - i_s + 1$, $\bar{y}_i = \frac{1}{n_i}\sum_{j=i_s}^{i_t} y_j$, $\bar{u}_i = \frac{1}{n_i}\sum_{j=i_s}^{i_t} u_j$.

We start by the following regret decomposition.

$$R_n = \sum_{i=1}^{M} \underbrace{\sum_{j=i_s}^{i_t} (x_j - y_j)^2 - (y_j - \bar{y}_i)^2}_{T_{1,i}} + \sum_{i=1}^{M} \underbrace{\sum_{j=i_s}^{i_t} (y_j - \bar{y}_i)^2 - (y_j - \bar{u}_i)^2}_{T_{2,i}} + \sum_{i=1}^{M} \underbrace{\sum_{j=i_s}^{i_t} (y_j - \bar{u}_i)^2 - (y_j - u_j)^2}_{T_{3,i}} \tag{5.6}$$

Now the task of bounding R_n reduces to bounding $T_{1,i}, T_{2,i}, T_{3,i}$ for each bin and adding them up across all M bins. Let C_i be the TV within bin i incurred by the offline optimal. In Lemma 38, we exhibit a partitioning $\mathcal{P}$ of $[n]$ into $M = O(n^{1/3} C_n^{2/3} B^{-2/3})$ bins such that $C_i \le B/\sqrt{n_i}$ for each bin.

Due to strong adaptivity of FLH, the term $T_{1,i} = O(B^2 \log n)$ since it is the static regret against the fixed comparator $\bar{y}_i$. Hence adding them across all bins in the partition $\mathcal{P}$ yields $\sum_{i=1}^{M} T_{1,i} = \tilde{O}(n^{1/3} C_n^{2/3} B^{4/3})$.

By exploiting the KKT conditions satisfied by the offline optimal and using strong smoothness, we show in Lemma 42 that $T_{3,i}$ can be at-most $O(n_i C_i^2 + \lambda C_i)$ in general. Here $\lambda \ge 0$ is the optimal dual variable arising from the KKT conditions (Lemma 36). Since $C_i = O(B/\sqrt{n_i})$ for bins in the partition $\mathcal{P}$, we have $n_i C_i^2 = O(B^2)$. However, it is not possible to bound $\lambda C_i = O(1)$ since λ can be even $\Theta(n)$ in some cases (See Example 167 in Appendix D.2).

This is where the term $T_{2,i}$ plays a crucial role. Note that since $\bar{y}_i$ is the minimizer of $g(x) = \sum_{j=i_s}^{i_t} (y_j - x)^2$, we conclude that $T_{2,i} \le 0$. For simplicity of exposition, let's assume that $T_{2,i} < 0$, deferring formal arguments for the general case to Section 5.1.2. We show that this negative term diminishes the λC_i arising from the bound on $T_{3,i}$ to a quantity that is $O(1)$. Specifically, $T_{2,i} + T_{3,i} = O(B^2)$ even though individually $|T_{2,i}|, |T_{3,i}|$ can be *very large*. The desired regret bound now follows by summing it across all $M = O(n^{1/3} C_n^{2/3} B^{-2/3})$ bins in $\mathcal{P}$.

5.1.2 Regret Analysis

Define the sign function as $\text{sign}(x) = 1$ if $x > 0$; -1 if $x < 0$; and some $u \in [-1, 1]$ if $x = 0$. For a vector $\boldsymbol{x} \in \mathbb{R}^d$, $\text{sign}(\boldsymbol{x}) \in \mathbb{R}^d$ is defined by the coordinate-wise application of this rule. We start by presenting a sequence of useful lemmas.

Lemma 36. *(**characterization of offline optimal**) Consider the following convex optimization problem (where $\tilde{z}_1, ..., \tilde{z}_{n-1}$ are introduced as dummy variables)*

$$\min_{\tilde{u}_1,\ldots,\tilde{u}_n,\tilde{z}_1,\ldots,\tilde{z}_{n-1}} \quad \frac{1}{2}\sum_{t=1}^{n}(y_t - \tilde{u}_t)^2 \tag{5.7a}$$

$$\text{s.t.} \quad \tilde{z}_t = \tilde{u}_{t+1} - \tilde{u}_t \; \forall t \in [n-1], \tag{5.7b}$$

$$\sum_{t=1}^{n-1} |\tilde{z}_t| \le C_n \tag{5.7c}$$

Let $u_1, \ldots, u_n, z_1, \ldots, z_{n-1}$ be the optimal primal variables and let $\lambda \ge 0$ be the optimal dual variable corresponding to the last constraint (5.7c)*. By the KKT conditions, we have*

- ***stationarity:*** $y_t = u_t - \lambda(s_t - s_{t-1})$*, where $s_t \in \partial|z_t|$ (a subgradient). Specifically, $s_t = \text{sign}(u_{t+1} - u_t)$ if $|u_{t+1} - u_t| > 0$ and s_t is some value in $[-1, 1]$ otherwise. For convenience of notations later, we also define $s_n = s_0 = 0$.*

- ***complementary slackness:*** $\lambda \left(\sum_{t=2}^{n} |u_t - u_{t-1}| - C_n\right) = 0$*.*

Remark 37. We enumerate some elementary observations about the optimal primal variables in Lemma 36 that will be used throughout.

P1 For any time point t, if the optimal solution $u_{t+1} > u_t$, then $s_t = 1$. Similarly $s_t = -1$ whenever $u_{t+1} < u_t$. If $u_t = u_{t+1}$, the s_t can be any number in $[-1, 1]$.

P2 Consider a sub-interval $[a, b]$ with $2 \le a \le n-1$ such that the optimal solution jumps at both the end points. i.e $u_k \neq u_{k-1}$ for $k \in \{b+1, a\}$. Define $\Delta s_{a \to b} :=$

$s_b - s_{a-1}$. Then either $|\Delta s_{a\to b}| = 0$ or $|\Delta s_{a\to b}| = 2$ since $s_{a-1} \in \{-1, 1\}$ and $s_b \in \{-1, 1\}$.

P3 Consider a sub-interval $[1, b]$ with $b < n$ such that $u_{b+1} \neq u_b$. Then $|\Delta s_{1\to b}| = 1$ since $s_0 = 0$ by convention (Lemma 36). Similarly for a sub-interval $[a, n]$ with $a > 1$, such that $u_{a-1} \neq u_a$, we have $|\Delta s_{a\to n}| = 1$.

Terminology. We will refer to the optimal primal variables $u_1, \ldots, u_n$ in Lemma 36 as the *offline optimal sequence* in this section.

Next, we exhibit a useful partitioning scheme of the interval $[n]$.

Lemma 38. ***(key partition)*** *Initialize $\mathcal{Q} \leftarrow \Phi$. Starting from time 1, spawn a new bin $[i_s, i_t]$ whenever $\sum_{j=i_s+1}^{i_t+1} |u_j - u_{j-1}| > B/\sqrt{n_i}$, where $n_i = i_t - i_s + 2$. Add the spawned bin $[i_s, i_t]$ to $\mathcal{Q}$. Consider the following post processing routine.*

1. *Initialize $\mathcal{P} \leftarrow \Phi$.*
2. *For $i \in [|\mathcal{Q}|]$:*
   - *if $u_{i_t} = u_{i_t+1}$:*
     (a) *Let p be the largest time point with $u_{p:i_t}$ being constant and let q be the smallest time point with $u_{i_t+1:q}$ being constant.*
     (b) *Add bin $[i_s, p-1]$ to $\mathcal{P}$.*
     (c) *If $(i+1)_t > q$ then add $[p, q]$ to $\mathcal{P}$ and set $(i+1)_s \leftarrow q+1$.*
     (d) *Goto Step 2.*
   - *Add $[i_s, i_t]$ to $\mathcal{P}$. Goto Step 2.*

Let $M := |\mathcal{P}|$. We have $M = O\left(1 \vee n^{1/3} C_n^{2/3} B^{-2/3}\right)$. Further for any bin $[i_s, i_t] \in \mathcal{P}$, it holds that $\sum_{j=i_s+1}^{i_t} |u_j - u_{j-1}| \leq B/\sqrt{n_i}$ where $n_i = i_t - i_s + 1$.

Remark 39. Consider a bin $[i_s, i_t] \in \mathcal{P}$. Let $\Delta s_i := s_{i_t} - s_{i_s - 1}$. By virtue of the post processing routine of Lemma 38, the bin $[i_s, i_t]$ will conform to either of the cases P2 or P3 in Remark 37. So we have $|\Delta s_i| > 0$ implies $|\Delta s_i| \geq 1$.

We emphasize that the bins $[i_s, i_t]$ we consider in Eq. (5.6) belong to the partition $\mathcal{P}$ of Lemma 38. We proceed to bound $T_{1,i}, T_{2,i}$ and $T_{3,i}$ in the regret decomposition of Eq.(5.6).

Lemma 40. *(**bounding** $T_{1,i}$) Assume that we run* FLH *with the settings described in Theorem 34. For any bin i we have $T_{1,i} = O\left(B^2 \log n\right)$*

Lemma 41. *(**bounding** $T_{2,i}$) Define $C_i := \sum_{j=i_s+1}^{i_t} |u_j - u_{j-1}|$, the TV within bin i incurred by the offline optimal solution. Let $\Delta s_i := s_{i_t} - s_{i_s - 1}$ and $n_i := i_t - i_s + 1$. We have $T_{2,i} \leq \frac{-\lambda^2 (\Delta s_i)^2}{n_i}$.*

Lemma 42. *(**bounding** $T_{3,i}$) Let C_i and Δs_i be as in Lemma 41.*

Case(a) *If $|\Delta s_i| > 0$ then $T_{3,i} \leq B^2 + 6\lambda C_i$.*

Case(b) *If $\Delta s_i = 0$ with $s_{i_s - 1} = s_{i_t} = 1$ and the offline optimal $\boldsymbol{u}$ is non-decreasing within bin i, then $T_{3,i} \leq B^2$.*

Case(c) *If $\Delta s_i = 0$ with $s_{i_s - 1} = s_{i_t} = -1$ and the offline optimal $\boldsymbol{u}$ is non-increasing within bin i, then $T_{3,i} \leq B^2$.*

Proof. **of Theorem 34** Tree diagrams that represent the flow of arguments in the proof is displayed in Fig.D.2 and D.3 in Appendix D.2. We start from the regret decomposition in Eq. (5.6).

Case (a) in Lemma 42. First we handle case(a) in Lemma 42 where $|\Delta s_i| > 0$. Define $T_i := T_{1,i} + T_{2,i} + T_{3,i}$. From Lemmas 40, 41 and 42 we have

$$
\begin{aligned}
T_i &\leq O\left(B^2 \log n\right) - \frac{\lambda^2 (\Delta s_i)^2}{n_i} + B^2 + 6\lambda C_i && (5.8)\\
&\leq O\left(B^2 \log n\right) - \frac{\lambda^2 (\Delta s_i)^2}{n_i} + 6\lambda C_i && (5.9)\\
&\leq_{(a)} O\left(B^2 \log n\right) + \frac{9 n_i C_i^2}{(\Delta s_i)^2} - \left(\frac{\lambda \Delta s_i}{\sqrt{n_i}} - \frac{3 C_i \sqrt{n_i}}{\Delta s_i}\right)^2 && (5.10)\\
&\leq_{(b)} O(B^2 \log n) + 9B^2 && (5.11)\\
&\leq O(B^2 \log n), && (5.12)
\end{aligned}
$$

where line (a) is obtained by completing the square. For line (b) we dropped the negative term used Remark 39 to conclude $|\Delta s_i| \geq 1$. Further $n_i C_i^2 \leq B^2$ for bins in the partition $\mathcal{P}$ of Lemma 38.

Case (b) and (c) in Lemma 42. To handle case (b) and case (c) in Lemma 42 where $\Delta s_i = 0$ and monotonic, we have $T_{1,i} = O(B^2 \log n)$ due to Lemma 40, $T_{2,i} \leq 0$ due to Lemma 41 and $T_{3,i} \leq B^2$ due to Lemma 42. So $T_i \leq O\left(B^2 \log n\right) + B^2 \leq O(B^2 \log n)$.

Other cases:

(A1) Consider the case when $\Delta s_i = 0$ with $s_{i_s - 1} = s_{i_t} = -1$ and the offline optimal $\boldsymbol{u}$ is non-decreasing within bin i. If the sequence is constant within the bin, then trivially we have $T_i = O\left(B^2 \log n\right)$ due to Strongly Adaptivity of FLH. Otherwise, we the split the original bin into two sub-bins $[i_s, k]$ and $[k+1, i_t]$ such that $s_k = 1$ with $u_{k+1} > u_k$. See config (a) in Fig.5.1 for an illustration. Then the two sub-bins falls into the category of case (a) in Lemma 42. By bounding the regret within each sub-bin separately by following the previous arguments for case (a) and adding them up, we can get $T_i \leq O\left(B^2 \log n\right)$ regret for the original bin. The arguments for the case when $\Delta s_i = 0$ with $s_{i_s - 1} = s_{i_t} = 1$ and the offline optimal $\boldsymbol{u}$ is non-increasing within bin i are similar.

(A2) To handle the case when $\Delta s_i = 0$ and the optimal sequence is *not* monotonic, we split the bin into two parts. Consider the case $s_{i_t} = s_{i_s - 1} = 1$. We can split $\boldsymbol{u}_{i_s:i_t}$ as $\boldsymbol{u}_{i_s:k}$ and $\boldsymbol{u}_{k+1:i_t}$ such that the sequence $\boldsymbol{u}_{i_s:k}$ is non-decreasing and $s_k = -1$ with $u_k > u_{k+1}$. See config (b) in Fig.5.1 for an illustration. Notice that both the sub-bins $\boldsymbol{u}_{i_s:k}$ and $\boldsymbol{u}_{k+1:i_t}$ now falls into the category of case(a) in Lemma 42. Adding the bounds within these sub-bins by following the treatment for case (a) above yields $T_i \leq O\left(B^2 \log n\right)$. The arguments for the scenario $s_{i_t} = s_{i_s - 1} = -1$ are similar.

Now the theorem follows by summing $\sum_{i=1}^{M} T_i$ for the $M = O\left(1 \vee n^{1/3} C_n^{2/3} B^{-2/3}\right)$ bins in the partition $\mathcal{P}$ of Lemma 38. □

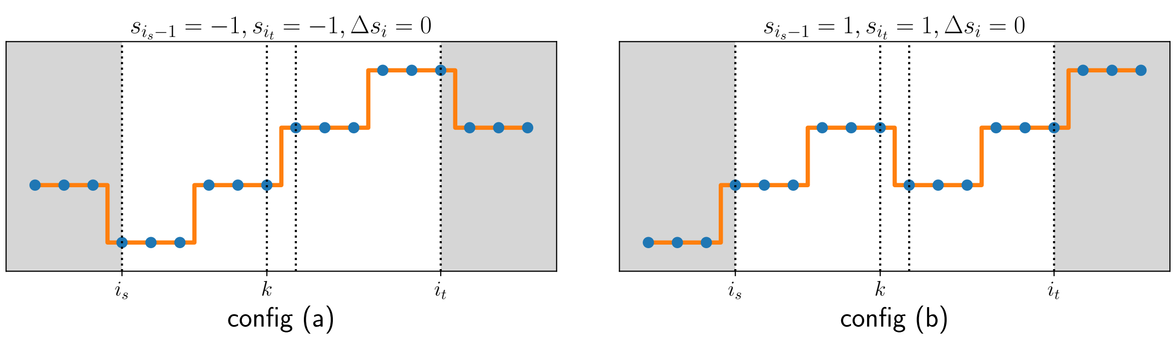


Figure 5.1: *Examples of configurations referred in the proof of Theorem 34. The blue dots corresponds to the offline optimal sequence.*

The previous results generalize to online TV-denoising framework in higher dimensions.

Proposition 43. ***(Extension to higher dimensions)*** *Consider a protocol where at each time the learner predicts a vector $\boldsymbol{x}_t \in \mathbb{R}^d$ after which the adversary reveals $\boldsymbol{y}_t$ such that $\|\boldsymbol{y}_t\|_\infty \leq B$. Consider a comparator sequence of vectors $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ such that $TV(\boldsymbol{w}_{1:n}) := \sum_{t=2}^{n} \|\boldsymbol{w}_t - \boldsymbol{w}_{t-1}\|_1 \leq C_n$. By running d instances of* FLH *with learning rate $\zeta = 1/(8B^2)$ and FTL as base learners, where instance i, $i \in [d]$, predicts $\boldsymbol{x}_t[i]$ at* *time t, we have*

$$R_n(\boldsymbol{w}_{1:n}) := \sum_{j=1}^{n} \|\boldsymbol{y}_t - \boldsymbol{x}_t\|_2^2 - \|\boldsymbol{y}_t - \boldsymbol{w}_t\|_2^2 = \tilde{O}\left(dB^2 \log n \vee d^{1/3} n^{1/3} C_n^{2/3} B^{4/3}\right).$$

Proposition 44. *(**Lower bound**) Assume the protocol and notations of Proposition 43. For any algorithm, we have*

$$\sup_{\boldsymbol{w}_{1:n}: TV(\boldsymbol{w}_{1:n}) \leq C_n} R_n(\boldsymbol{w}_{1:n}) = \Omega\left(dB^2 \log n \vee d^{1/3} n^{1/3} C_n^{2/3} B^{4/3}\right). \tag{5.13}$$

By comparing the upper and lower bounds, we conclude that the FLH-FTL strategy in Proposition 43 is minimax optimal (modulo log factors) wrt *all* parameters d, n, B and C_n.

Remark 45. Several other non-parametric sequence classes such as the Holder ball $\mathcal{H}^B(B_n') = \{w_{1:n} : \|Dw_{1:n}\|_\infty \leq B_n', \|w_{1:n}\|_\infty \leq B\}$ and Sobolev ball $\mathcal{S}^B(C_n') = \{w_{1:n} : \|Dw_{1:n}\|_2 \leq C_n', \|w_{1:n}\|_\infty \leq B\}$ can be shown to embedded inside a $\mathcal{TV}^B(C_n)$ ball for appropriate choices of C_n, B_n and B_n' (see [37]) with all classes having the same minimax rates of estimation in the iid setting. So the minimax optimality on TV ball for FLH with FTL as base learners implies minimax optimality on the embedded Holder and Sobolev balls as well.

5.2 Performance guarantees for exp-concave losses

We begin by listing all the assumptions we make about the loss functions.

EC-1 Without loss of generality, we assume $\mathbf{0} \in \mathcal{W}$. Let $B := \sup_{\boldsymbol{x} \in \mathcal{W}} \|\boldsymbol{x}\|_\infty$. Define $\mathcal{D}^- := \{\boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x}\|_\infty \leq B\}$. The loss functions $f_t(x) : \mathbb{R}^d \to \mathbb{R}$ are G Lipschitz in $\mathcal{D}^-$.

EC-2 The loss functions are β strongly smooth in $\mathcal{D} = \{\boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x}\|_\infty \leq B + G\}$. i.e $f_t(\boldsymbol{y}) \leq f_t(\boldsymbol{x}) + (\boldsymbol{y}-\boldsymbol{x})^T \nabla f_t(\boldsymbol{x}) + \frac{\beta}{2}\|\boldsymbol{x}-\boldsymbol{y}\|_2^2$, for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{D}$. We assume without loss of generality that $\beta \geq 1$.

EC-3 The loss functions are α exp-concave in $\mathcal{D}$. i.e $f_t(\boldsymbol{y}) \geq f_t(\boldsymbol{x}) + (\boldsymbol{y}-\boldsymbol{x})^T \nabla f_t(\boldsymbol{x}) + \frac{\alpha}{2}\left((\boldsymbol{y}-\boldsymbol{x})^T \nabla f_t(\boldsymbol{x})\right)^2$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{D}$.

EC-4 The loss functions $f_t(x) : \mathbb{R}^d \to \mathbb{R}$ are $G^\dagger$ Lipschitz in $\mathcal{D}$.

Below, we give an example of a family of loss functions that satisfy the above assumptions.

Example 46 (Generalized linear models)**.** Let $f_t(\boldsymbol{x}) = g(\boldsymbol{v}_t^T \boldsymbol{x})$, where $g : \mathbb{R} \to \mathbb{R}$ is a convex function and $\boldsymbol{v}_t$ is a feature vector. Let $\|\boldsymbol{v}_t\|_2 \leq R$. Assume that for all $\boldsymbol{x} \in \mathcal{D}^-$ we have $|g_t'(\boldsymbol{v}_t^T \boldsymbol{x})| \leq a$. Further for all $\boldsymbol{x} \in \mathcal{D}$, let $|g_t'(\boldsymbol{v}_t^T \boldsymbol{x})| \leq a^+$, $g_t''(\boldsymbol{v}_t^T \boldsymbol{x}) \leq b$, $g_t''(\boldsymbol{v}_t^T \boldsymbol{x}) \geq c > 0$. Then Assumptions EC 1-5 are satisfied by by the losses f_t with $G = aR$, $\beta = bR^2$, $\alpha = c/((a^+)^2)$ and $G^\dagger = Ra^+$.

We are interested in characterizing the maximum dynamic regret

$$R_n^+(C_n) := \sup_{\substack{\boldsymbol{w}_1,\ldots,\boldsymbol{w}_n \in \mathcal{D}^- \\ \sum_{t=2}^n \|\boldsymbol{w}_t - \boldsymbol{w}_{t-1}\|_1 \leq C_n}} \sum_{t=1}^n f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{w}_t), \tag{5.14}$$

where $\boldsymbol{x}_t$ are the predictions of the learner. Since $\mathcal{W} \subseteq \mathcal{D}^-$, the dynamic regret against comparators in $\mathcal{D}^-$ trivially upperbounds the dynamic regret against $\mathcal{W}$. The algorithms that we study throughout this section are improper in the sense that the predictions of the algorithms belong to $\mathcal{D} \supset \mathcal{W}$.

Before diving into the details, we remark that our main focus is to get optimal dependence on n and C_n. The dimension d is considered as a constant problem parameter and we do not try to optimize its polynomial dependence. All unspecified proofs of this section are given in Appendix D.3.

We have the following regret guarantee for exp-concave losses.

Theorem 47. *By using the base learner as ONS with parameter* $\zeta = \min\left\{\frac{1}{4G^\dagger(2B\sqrt{d}+2G/\beta)}, \alpha\right\}$, *decision set* $\mathcal{D}$ *and choosing learning rate* $\eta = \alpha$, FLH *obeys* $R_n^+(C_n) = \tilde{O}\left(d^{3.5}(n^{1/3}C_n^{2/3} \vee 1)\right)$ *if* $C_n > 1/n$ *and* $O(d^{1.5}\log n)$ *otherwise. Here* $a \vee b := \max\{a, b\}$ *and* $\tilde{O}(\cdot)$ *hides dependence on the constants* $B, G, G^\dagger, \alpha$ *and factors of* $\log n$.

proof sketch. Let $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ be the *offline optimal* sequence such that $\sum_{t=1}^n f_t(\boldsymbol{u}_t)$ is minimum across all sequences that obeys: (a) $\sum_{t=2}^n \|\boldsymbol{u}_t - \boldsymbol{u}_{t-1}\|_1 \le C_n$; (b) $\boldsymbol{u}_t \in \mathcal{D}^-$ for all $t \in [n]$ (see Lemma 175 in Appendix D.3 for more details).

Let $\mathcal{P}$ be a partition of $[n]$ into $M = O^*(n^{1/3}C_n^{2/3})$ bins obtained by a similar scheme in Lemma 38 where within each bin, we have $\sum_{j=i_s+1}^{i_t} \|\boldsymbol{u}_j - \boldsymbol{u}_{j-1}\|_1 \le B/\sqrt{n_i}$. Let $[i_s, i_t]$ denote the i^{th} bin in $\mathcal{P}$ and let n_i be its length. Define $\bar{\boldsymbol{u}}_i = \frac{1}{n_i}\sum_{j=i_s}^{i_t} \boldsymbol{u}_j$ and $\dot{\boldsymbol{u}}_i = \bar{\boldsymbol{u}}_i - \frac{1}{n_i\beta}\sum_{j=i_s}^{i_t} \nabla f_j(\bar{\boldsymbol{u}}_i)$ where β is as in Assumption EC-2. Let $\boldsymbol{x}_j$ be the prediction made by FLH at time j. We start with following regret decomposition.

$$R_n^+(C_n) \le \sum_{i=1}^{M} \underbrace{\sum_{j=i_s}^{i_t} f_j(\boldsymbol{x}_j) - f_j(\dot{\boldsymbol{u}}_i)}_{T_{1,i}} + \sum_{i=1}^{M} \underbrace{\sum_{j=i_s}^{i_t} f_j(\dot{\boldsymbol{u}}_i) - f_j(\bar{\boldsymbol{u}}_i)}_{T_{2,i}} + \sum_{i=1}^{M} \underbrace{\sum_{j=i_s}^{i_t} f_j(\bar{\boldsymbol{u}}_i) - f_j(\boldsymbol{u}_j)}_{T_{3,i}} \tag{5.15}$$

Unlike the squared error case, for the term $T_{1,i}$, we *do not* compete with the minimizer of $g(\boldsymbol{x}) := \sum_{j=i_s}^{i_t} f_j(\boldsymbol{x})$. Instead we compete with $\dot{\boldsymbol{u}}_i$ which is obtained by a one-step gradient descent of $g(\boldsymbol{x})$ from the point $\bar{\boldsymbol{u}}_i$ where the step size is set as $1/(n_i\beta)$.

Recall that the purpose of $\bar{y}_i$ in Eq. (5.6) was to make $T_{2,i}$ non-positive thereby facilitating potential cancellation of terms arising from the bound on $T_{3,i}$. Since $g(\boldsymbol{x})$ is $n_i\beta$ strongly smooth, by the well known *descent lemma* in first order optimization (eg. see Eq. 3.5 in [64]), we can bound $T_{2,i}$ in Eq. (5.15) with a "sufficiently negative" term

$-\frac{1}{2n_i\beta}\|\nabla g(\bar{\boldsymbol{u}}_i)\|^2$ as well. Also, observe that

$$\|\dot{\boldsymbol{u}}\|_\infty \leq \|\bar{\boldsymbol{u}}_i\|_\infty + \frac{\sum_{j=i_s}^{i_t}\|\nabla f_j(\bar{u})\|_\infty}{n_i\beta} \leq B + G, \tag{5.16}$$

where in the last line we used the fact $\bar{\boldsymbol{u}}_i \in \mathcal{D}^-$ and the Lipschitzness assumption in EC-1 along with $\beta > 1$ by assumption EC-2. So in $T_{1,i}$ the comparator term $\dot{\boldsymbol{u}}_i \in \mathcal{D}$. The base learners of the FLH produce predictions in $\mathcal{D}$ to compete with such a comparator hence making the overall algorithm improper. We do not project $\dot{\boldsymbol{u}}$ to the set $\mathcal{W}$, because doing so appears to make $T_{2,i}$ not negative enough to adequately diminish the terms arising from $T_{3,i}$.

Rest of the proof proceeds by introducing lemmas analogous to the squared error case, carefully bounding $T_{1,i} + T_{2,i} + T_{3,i}$ for each bin in $\mathcal{P}$ and summing them up across all bins. However, we remark that the analysis is significantly more involved in comparison to that of squared error case due to dual variables introduced by the additional constraint that $\boldsymbol{u}_t \in \mathcal{D}^-$.

We first present the proof for the 1D-exp-concave case in Appendix D.3.1, which illustrates how boundedness constraints are handled by the structures in the KKT-conditions (Lemma 169) and by discussing various combinations (see Fig. D.5-D.7). Then we present the full proof for the higher-dimensional exp-concave losses in Appendix D.3.2, where the structure becomes too complex for us to enumerate all combinations. We address this by constructing an iterative algorithm that generates bins and prove that the algorithm is guaranteed to find a partition with cardinality $O^*(n^{1/3}C_n^{2/3})$ that satisfies a number of additional properties that give rise to the regret bound we claim. □

Proposition 48. *For strongly convex losses, the regret bound can be improved to* $\tilde{O}\left(d^2(n^{1/3}C_n^{2/3} \vee 1)\right)$ *if* $C_n > 1/n$ *and* $O(\log n)$ *otherwise by using OGD as base learners* *in the FLH procedure. See Appendix D.3.2 for a proof.*

By comparing with the lower bound in Proposition 44 we conclude that the dynamic regret bound of Theorem 47 is minimax optimal (up to $\log n$ factors) in n and C_n.

Remark 49 (Implications in statistical methodology.)**.** Example 46 and Theorem 47 extends the locally-adaptive nonparametric regression theory that are typically studied for square loss to an arbitrary strongly convex / exp-concave loss while allowing covariates (exogenous variables) to be modeled. Moreover, the method enjoys strong oracle inequalities (e.g. Remark 35) that certifies the predictive performance in a fully agnostic / model-misspecified setting with no stochastic assumptions. In addition, the method does not introduce additional tuning parameters at all.

5.3 Concluding Discussion

In this chapter, we considered the problem of dynamic regret minimization with exp-concave losses and showed that SA methods are minimax optimal (modulo factors of $\log n$ and d) in a setting where improper learning is allowed. To the best of our knowledge this is the first work that attains optimal dynamic regret rates under this setting. The resulting algorithms are adaptive to the path variation of the comparator sequence. Further, our results have far reaching consequences in locally adaptive non-parametric regression as mentioned in Remark 49.

Chapter 6

Optimal Dynamic Regret in Proper Online Learning with Strongly Convex Losses and Beyond

A question that was left open in the previous chapter was whether improper learning is strictly necessary to achieve the optimal rates for exp-concave optimization. In this chapter, we answer this in the negative by showing that a proper version of the SA algorithms can attain the optimal (modulo log factors and dimension dependencies) dynamic regret rates whenever the losses are strongly convex.

We summarize our main contributions of this chapter below.

- We provide a new analysis that extends the results of [65] to proper strongly convex online learning to attain the near *optimal* dynamic regret rate of $\tilde{O}(d^{1/3}n^{1/3}C_n^{2/3} \vee d)$ for Strongly Adaptive methods (see Corollary 54). In contrast to [65], our results imply an important conclusion that improper learning is *not strictly necessary* for attaining such fast rates with general strongly convex losses. To the best of our knowledge, this is the *first* result that achieves near optimal dynamic regret in a setting of proper learning under strongly convex losses.

- For exp-concave losses, we prove an analogous result that Strongly Adaptive algorithms can attain a near optimal dynamic regret of $\tilde{O}^*((n^{1/3}C_n^{2/3} \vee 1))$ in the special case of L_∞ (box) constrained decision set, $\mathcal{D} = \{\boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x}\|_\infty \leq B\}$ (see Theorem 59).

- To facilitate these results we discover and exploit a number of new structures imposed by the KKT conditions that were not considered in [65], which could be of independent interest.

Notes on scope and relevance. Under exp-concave or strongly convex losses, the important question of finding an optimal (wrt universal dynamic regret) and proper algorithm has remained resistant to attacks in the non-stationary online learning literature for almost two decades since the work of [5]. In this work, we take the first steps in addressing this question by showing optimality of proper SA learner in proper learning settings. The fact that a proper version of Strongly Adaptive algorithms can lead to optimal rates was highly unclear from the analysis of [65]. Further, by lifting the gradient smoothness assumption for the revealed losses, we modestly enlarge the applicability of the results when compared to [65]. Though our proof techniques bear some semblance with that of [65] in terms of the usage of KKT conditions, this similarity is only superficial and we introduce several new non-trivial ideas in the analysis for attaining the new results (see Sections 6.2.2 and 6.4.1).

6.1 Related Work

In this section, we compare and contrast our work with several existing lines of research.

Dynamic regret minimization in non-stationary online learning. Apart from [65], our work fits into the broad literature of dynamic regret minimization in online learning such as [5, 15, 16, 25, 66, 22, 26, 27, 62, 67, 37, 68, 69, 54, 70, 71, 72, 73, 74]. However, to the best of our knowledge none of these works are known to attain the optimal dynamic regret rate for our setting in terms of path length of the arbitrary comparator sequence.

Adaptive online learning. There is a complementary body of work on Strongly Adaptive regret minimization such as [24, 75, 69, 76] and Adaptive regret minimization such as [23, 52] (which are in fact Strongly Adaptive wrt exp-concave losses) that aims at controlling the static regret in any local time interval. This work focuses on developing new guarantees for algorithms that are Strongly Adaptive (SA) wrt strongly convex / exp-concave losses. The base learners we use for SA methods are the static regret minimizing algorithms from [77].

Locally adaptive non-parametric regression. Our work is closely related to locally adaptive non-parametric regression literature from the statistics community such as [51, 1, 2, 4, 3, 13, 38, 39, 63]. This work supplements them by removing the statistical assumptions and enabling to go beyond squared error losses for the non-parametric function class of TV bounded functions.

Online non-parametric regression. The results of [29] certifies that the minimax rate for competing against a reference class of TV bounded functions with squared error losses is $O(n^{1/3})$. However this bound doesn't capture the correct dependence on C_n and is arrived via non-constructive arguments. On the other hand we arrive at the optimal dependence on both n and C_n via an efficient algorithm. Further, our results with squared error losses in Section 6.2 are more general than that of [65] (see Remark 51). Results on online non-parametric regression against reference class of Lipschitz functions, Sobolev functions and isotonic functions can be found in [30, 78, 53] respectively. However as noted in [37], these classes feature functions that are more regular than TV bounded functions. In fact they can be embedded inside a TV bounded function class. So the minimax optimally for TV class implies minimax optimality for the smoother function classes as well.

We refer the reader to [65] and references therein for a more elaborate survey on existing literature.

6.2 A gentle start: Squared loss games

To start with, we consider the following squared loss game which will later play a pivotal role in the generalization to strongly convex losses.

- At time $t \in [n] := \{1, \ldots, n\}$, player predicts $x_t \in [-B, B]$.
- Adversary reveals a label $y_t \in [-G, G]$
- Player suffers loss $(y_t - x_t)^2$.

We make the following assumption.

Assumption A1: We assume that $[-B, B] \subseteq [-G, G]$ with $B \geq 1$ without loss of generality.

Define a class of comparators as:

$$\mathcal{TV}^B(C_n) := \left\{ w_{1:n} \middle| \mathrm{TV}(w_{1:n}) := \sum_{t=2}^{n} |w_t - w_{t-1}| \leq C_n, |w_t| \leq B \ \forall t \in [n] \right\}. \tag{6.1}$$

We are interested in simultaneously controlling the dynamic regret against all sequences in $\mathcal{TV}^B(C_n)$. The main algorithm we use for this task is the Follow-the-Leading-History (FLH) from [23] with Online Gradient Descent (OGD) run on the decision set

$[-B, B]$ as base learners. This algorithm will be referred as *FLH-OGD* strategy henceforth. We have the following performance guarantee.

Theorem 50. *Suppose the labels y_t generated by the adversary belong to $[-G, G]$. Let x_t be the prediction at time t of FLH with learning rate $\zeta = 1/(2(G+B)^2)$, base learners as OGD with step sizes $1/(2t)$ and decision set $[-B, B]$. Then for any comparator sequence $(w_1, \ldots, w_n) \in \mathcal{TV}^B(C_n)$*

$$\sum_{t=1}^{n} (y_t - x_t)^2 - (y_t - w_t)^2 = \tilde{O}\left(n^{1/3} C_n^{2/3} \vee 1\right), \tag{6.2}$$

where $\tilde{O}(\cdot)$ hides dependence on logarithmic factors of horizon n, G, B and $a \vee b := \max\{a, b\}$.

Remark 51 (**Adaptivity to C_n and safe (non-stochastic) oracle inequality**)**.** The FLH-OGD strategy does not require C_n as an input. Further, Theorem 50 has implications in non-parametric regression under safety constraints. When the non-parametric estimator for the $\mathcal{TV}^B$ sequence class is required to obey a safety constraint that the estimator's outputs x_t must also lie in $[-B, B]$, Theorem 50 implies the following oracle inequality:

$$\sum_{t=1}^{n} (y_t - x_t)^2 + g(x_t) \le \min_{w_{1:n}} \sum_{t=1}^{n} (y_t - w_t)^2 + g(w_t) + \tilde{O}\left(n^{1/3} \mathrm{TV}(w_{1:n})^{2/3} \vee 1\right), \tag{6.3}$$

where $g(x)$ is a safety constraint such that $g(x) = \infty$ when $|x| > B$ and zero otherwise. This is a strict generalization of Remark 2 in [65].

6.2.1 Key insight behind the proof of Theorem 50

The insight we used in deriving regret rate in Theorem 50 for a proper learning setup is based on the following idea: Suppose that we need to compete against a comparator sequence that incurs a Total Variation (TV) of C_n. We observe that, this comparator sequence of decisions in hindsight requires to obey the TV constraint while the decisions of the Strongly Adaptive (SA) learner need not obey any such constraints. Consider a time interval I where the comparator sequence assumes a constant value (say v_1) in an *arbitrary* convex decision set D. There could be some other point in D (say v_2) which can incur better cumulative loss within that interval. Note that the comparator sequence may not assume the value v_2 in the interval I due to the global TV constraint. Due to the strongly adaptive property, the regret (against v_1) of the SA learner in interval I is then bounded by the regret (against v_1) of the static point v_2, which is less than or equal to zero, plus an extra log term. The presence of such non-positive terms can delicately offset the effect of the positive log terms when summed across all such intervals to get favorable dynamic regret rates. How small the non-positive terms are, when summed across all intervals, depends on the magnitude of C_n (and indirectly on n).

6.2.2 Detailed road map for the proof of Theorem 50

In this section, we focus on conveying the main ideas of our proof deferring the formal details to Appendix E.1. We start by briefly reviewing the proof strategy of [65] and then intuitively capture the points of similarities and differences in our analysis. Throughout the proof we use the shorthand $[a, b] := \{a, a+1, \ldots, b\}$ for two natural numbers $a < b$.

We start by characterizing the offline optimal. Define the sign function as $\text{sign}(x) = 1$ if $x > 0$; -1 if $x < 0$; and some $v \in [-1, 1]$ if $x = 0$.

Lemma 52. ***(characterization of offline optimal)*** *Consider the following convex* *optimization problem (where $\tilde{z}_1, ..., \tilde{z}_{n-1}$ are introduced as dummy variables)*

$$\min_{\tilde{u}_1,\dots,\tilde{u}_n,\tilde{z}_1,\dots,\tilde{z}_{n-1}} \quad \frac{1}{2}\sum_{t=1}^{n}(y_t - \tilde{u}_t)^2 \tag{6.4a}$$

$$\text{s.t.} \quad \tilde{z}_t = \tilde{u}_{t+1} - \tilde{u}_t \ \forall t \in [n-1], \tag{6.4b}$$

$$\sum_{t=1}^{n-1} |\tilde{z}_t| \le C_n, \tag{6.4c}$$

$$-B \le \tilde{u}_t \ \forall t \in [n], \tag{6.4d}$$

$$\tilde{u}_t \le B \ \forall t \in [n], \tag{6.4e}$$

Let $u_1, \dots, u_n, z_1, \dots, z_{n-1}$ be the optimal primal variables and let $\lambda \ge 0$ be the optimal dual variable corresponding to the constraint (6.4c)*. Further, let $\gamma_t^- \ge 0, \gamma_t^+ \ge 0$ be the optimal dual variables that correspond to constraints* (6.4d) *and* (6.4e) *respectively for all $t \in [n]$. By the KKT conditions, we have*

- ***stationarity:*** $u_t - y_t = \lambda\,(s_t - s_{t-1}) + \gamma_t^- - \gamma_t^+$*, where $s_t \in \partial|z_t|$ (a subgradient). Specifically, $s_t = \text{sign}(u_{t+1} - u_t)$ if $|u_{t+1} - u_t| > 0$ and s_t is some value in $[-1, 1]$ otherwise. For convenience of notations later, we also define $s_n = s_0 = 0$.*
- ***complementary slackness:*** *(a) $\lambda\,(\sum_{t=2}^{n} |u_t - u_{t-1}| - C_n) = 0$; (b) $\gamma_t^-(u_t + B) = 0$ and $\gamma_t^+(u_t - B) = 0$ for all $t \in [n]$*

Let the optimal solution constructed by the offline oracle be denoted by $u_{1:n}$ (termed as offline optimal henceforth). In [65], a partition $\mathcal{P} = \{[i_s, i_t], i \in [M]\}$ of $[n]$ is formed with cardinality $|\mathcal{P}| = M = O(n^{1/3}C_n^{2/3} \vee 1)$. The partition has an additional property that within each bin $[i_s, i_t] \in \mathcal{P}$, we have $C_i := \sum_{j=i_s+1}^{i_t} |u_j - u_{j-1}| \le B/\sqrt{i_t - i_s + 1}$ (see Lemma 184). Then for each bin, a three term regret decomposition is employed as follows:

$$\underbrace{\sum_{j=i_s}^{i_t}(y_j - x_j)^2 - (y_j - \bar{y}_i)^2}_{T_{1,i}} + \underbrace{\sum_{j=i_s}^{i_t}(y_j - \bar{y}_i)^2 - (y_j - \bar{u}_i)^2}_{T_{2,i}} + \underbrace{\sum_{j=i_s}^{i_t}(y_j - \bar{u}_i)^2 - (y_j - u_j)^2}_{T_{3,i}}, \tag{6.5}$$

where $\bar{u}_i = \sum_{j=i_s}^{i_t} u_j/(i_t - i_s + 1)$ and $\bar{y}_i = \sum_{j=i_s}^{i_t} y_j/(i_t - i_s + 1)$ and x_j are the predictions of the learner. They use online averaging as base learners for FLH. By strong adaptivity, they show $T_{1,i} = O(\log n)$. They show that $T_{3,i}$ can be $O(\lambda C_i)$ in general where λ is the dual variable arising from the KKT conditions (see Lemma 52) which can be even $\Theta(n)$ in the worst case. Since $\bar{y}_i$ is the static minimizer of $g(x) = \sum_{j=i_s}^{i_t}(y_j - x)^2$, they bound $T_{2,i}$ by a non-positive term which when added to $T_{3,i}$ can diminish into an $O(1)$ quantity. Thus regret within the bin $[i_s, i_t]$ is $T_{1,i} + T_{2,i} + T_{3,i} = O(\log n)$. This regret bound is added across all $O(n^{1/3}C_n^{2/3} \vee 1)$ bins of $\mathcal{P}$ to yield an $\tilde{O}(n^{1/3}C_n^{2/3} \vee 1)$ dynamic regret.

In our protocol of squared loss games, the labels $y_t \in [-G, G] \supseteq [-B, B]$. So we can't use online averages as base learner for constructing a proper learning algorithm. So in this work we use projected OGD as base learners with decision set $[-B, B]$. With such an algorithm, we may attempt to work with a slightly modified version of the three term regret decomposition of (6.5) as:

$$\underbrace{\sum_{j=i_s}^{i_t}(y_j - x_j)^2 - (y_j - \Pi(\bar{y}_i))^2}_{T'_{1,i}} + \underbrace{\sum_{j=i_s}^{i_t}(y_j - \Pi(\bar{y}_i))^2 - (y_j - \bar{u}_i)^2}_{T'_{2,i}} + \underbrace{\sum_{j=i_s}^{i_t}(y_j - \bar{u}_i)^2 - (y_j - u_j)^2}_{T'_{3,i}}, \tag{6.6}$$

where $\Pi(x)$ is the projection of $x \in \mathbb{R}$ to the interval $[-B, B]$. Unfortunately while doing so, the term $T'_{2,i}$ can be not negative enough to diminish $T'_{3,i}$ to an $O(1)$ quantity. We provide an empirical demonstration of this phenomenon in Fig.6.1. At this point, we hope that we have made a clear case on why the analysis of [65] cannot be directly extended to handle proper learning.

To get around this issue, we first identify two regimes for the dual variable λ. We show that when $\lambda = O(n^{1/3}/C_n^{1/3})$, one can still work with the same partitioning $\mathcal{P}$ of [65] (see Lemma 184) and use a decomposition similar to Eq.(6.6) to get the desired regret bound (see Lemma 186).

Before explaining the details of the regime $\lambda = \Omega(n^{1/3}/C_n^{1/3})$, we introduce the following definitions for convenience:

Definition 53.

- For a bin $[a, b] \subseteq \{2, \ldots, n-1\}$, the offline optimal solution is said to assume Structure 1 if $u_j = u_a \in (-B, B)$ for all $j \in [a, b]$ and $u_b > u_{b+1}$ and $u_a > u_{a-1}$.

- For a bin $[a, b] \subseteq \{2, \ldots, n-1\}$, the offline optimal solution is said to assume Structure 2 if $u_j = u_a \in (-B, B)$ for all $j \in [a, b]$ and $u_b < u_{b+1}$ and $u_a < u_{a-1}$.

- For a bin $[a, b]$, we define $\text{gap}_{\min}(\beta, [a, b]) := \min_{j \in [a,b]} |u_j - \beta|$ where $\beta \in \mathbb{R}$.

Consider the following two conditions.

Condition 1: For a bin $[i_s, i_t] \in \mathcal{P}$, the offline optimal satisfies $\text{gap}_{\min}(-B, [i_s, i_t]) \geq \text{gap}_{\min}(B, [i_s, i_t])$ and within at-least one sub-interval $[r, s] \subseteq [i_s, i_t]$, the offline optimal assumes the form of Structure 2.

Condition 2: For a bin $[i_s, i_t] \in \mathcal{P}$, the offline optimal satisfies $\text{gap}_{\min}(-B, [i_s, i_t]) < \text{gap}_{\min}(B, [i_s, i_t])$ and within at-least one sub-interval $[r, s] \subseteq [i_s, i_t]$, the offline optimal assumes the form of Structure 1.

Define:

$$\mathcal{Q} := \{[i_s, i_t] \in \mathcal{P} : \text{ the offline optimal satisfies Condition 1 or 2 in } [i_s, i_t]\}.$$

We refine a bin $[i_s, i_t] \in \mathcal{Q}$ that satisfy Condition 1 into smaller sub-intervals as shown in Fig.6.2, such that: for a style U sub-interval, the offline optimal takes the form of Structure 2 and for a style V sub-interval, the offline optimal has a non-decreasing section followed by an optional decreasing section. A similar refinement is also performed for bins in $\mathcal{Q}$ that satisfy Condition 2.

Our strategy is to bound:

$$\text{regret in style U sub-intervals} = O(\log n) + \text{a negative term}. \tag{6.7}$$

This is accomplished by a two term regret decomposition. Suppose $[a, b]$ is a style U sub-interval. We use the decomposition:

$$\underbrace{\sum_{j=a}^{b} (y_j - x_j)^2 - (y_j - w)^2}_{T_1} + \underbrace{\sum_{j=a}^{b} (y_j - w)^2 - (y_j - u_j)^2}_{T_2}, \tag{6.8}$$

with $w = \Pi\left(\sum_{j=a}^{b} y_j/(b-a+1)\right)$.

Next, we bound

$$\text{regret in style V sub-intervals} = O(\log n), \tag{6.9}$$

using a similar two term regret decomposition as in Eq.(6.8) with w replaced by a carefully chosen $w_j \in [(u_a \wedge \ldots \wedge u_b), (u_a \vee \ldots \vee u_b)]$ such that $\sum_{j=a+1}^{b} \mathbb{I}\{w_j \neq w_{j-1}\} \leq 6$ where $\mathbb{I}\{\cdot\}$ is the indicator function taking values in $\{0, 1\}$. We use the notation $x \wedge y =$

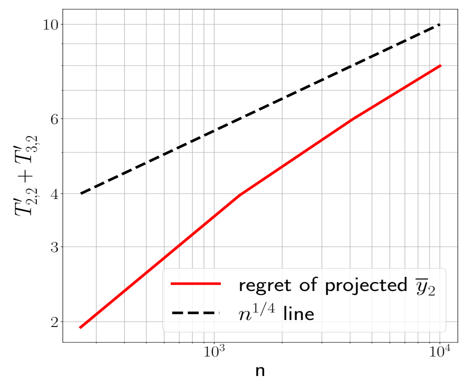


Figure 6.1: *Plot of* $T'_{2,2}+T'_{3,2}$ *(see Eq.(6.6) with* $i = 2$*) for the Example 182 in Appendix E.1. In this example,* $C_n = O(1/\sqrt{n})$ *and the partitioning procedure of [65] creates a partition* $\mathcal{P}$ *of* $[n]$ *containing two bins. We see that* $T'_{2,2} + T'_{3,2}$ *in the second bin grows roughly as* $O(n^{1/4})$*. However for applying the analysis of [65], we require this quantity for each bin in* $\mathcal{P}$ *to grow as* $O(1)$*. This makes the direct extension of the techniques in [65] with* $\bar{y}_i$ *replaced by* $\Pi(\bar{y}_i)$ *as in Eq.(6.6) inapplicable for the proper learning setting we study.*

$\min\{x, y\}$

We perform this task of refinement for every interval $[i_s, i_t]$ in $\mathcal{Q}$. Then we bound the regret in the resulting sub-intervals (as per Eq.(6.7) or (6.9)) and add the regret bounds across all such sub-intervals. Note that the total number of sub-intervals after refinement can be much larger than $|\mathcal{P}| = O(n^{1/3}C_n^{2/3} \vee 1)$. So if the bound in Eq.(6.7) is not tight enough, then there is a possibility that the resulting regret bound can be highly sub-optimal. This poses a major challenge in contrast to the analysis of [65] where they only need to work with a partition of size $O(n^{1/3}C_n^{2/3} \vee 1)$ and bound the regret in each interval of the partition by an $\tilde{O}(1)$ quantity.

To address this issue, we form tight bounds for Eq.(6.7) by exploiting certain structures in the KKT conditions that were previously unexplored in [65] via Lemmas 183, 189, 190 and 191. Of particular interest is Lemma 183 which highlights a fundamental way in which the adversary is constrained. Then we prove that if every bin $[i_s, i_t] \in \mathcal{Q}$ satisfies $\text{gap}_{\min}(-B, [i_s, i_t]) \vee \text{gap}_{\min}(B, [i_s, i_t]) \geq \mu_{\text{th}}$ where μ_{th} is as defined in Lemma 190, then the culmination of the negative terms in Eq.(6.7) can gracefully offset the effect of the positive $O(\log n)$ terms in Eq.(6.7) and Eq.(6.9) when summed across all refined intervals to obtain an $O(n^{1/3}C_n^{2/3} \vee 1)$ bound overall for $\sum_{[i_s,i_t] \in \mathcal{Q}} \sum_{j=i_s}^{i_t} (y_j - x_j)^2 - (y_j - u_j)^2$ (see proof of Lemma 191).

Further we show in Lemma 190 that when $\lambda = \Omega(n^{1/3}/C_n^{1/3})$ and $C_n = \tilde{O}(n)$, the criterion $\text{gap}_{\min}(-B, [i_s, i_t]) \vee \text{gap}_{\min}(B, [i_s, i_t]) \geq \mu_{\text{th}}$ is always satisfied for every bin $[i_s, i_t] \in \mathcal{Q}$. This can be seen informally as follows. Recall that the TV of the offline optimal within the bin $[i_s, i_t]$ is a "small" quantity that is at-most $(B/\sqrt{i_t - i_s + 1}) \leq B$. So if $\text{gap}_{\min}(-B, [i_s, i_t])$ is small, then due to this small TV constraint, we expect the quantity $\text{gap}_{\min}(B, [i_s, i_t])$ to be sufficiently large and vice versa.

Finally, for each bin in $\mathcal{R} := \mathcal{P} \setminus \mathcal{Q}$ we show (by using Lemma 187) that its regret contribution can be bounded by $O(\log n)$. Since $|\mathcal{R}| = O(n^{1/3}C_n^{2/3} \vee 1)$, such regret bounds lead to $\tilde{O}(n^{1/3}C_n^{2/3} \vee 1)$ bound overall when summed across all bins in $\mathcal{R}$.

Before closing this section, we capture the intuition behind the importance of the criterion $\text{gap}_{\min}(-B, [i_s, i_t]) \vee \text{gap}_{\min}(B, [i_s, i_t]) \geq \mu_{\text{th}}$ and why it can produce a sufficiently negative term in Eq.(6.7). Let's consider a style U sub-interval $[a, b]$ obtained by refining a bin $[i_s, i_t] \in \mathcal{Q}$ which satisfy Condition 1. Since $[a, b]$ is style U sub-interval, the offline optimal takes the form of Structure 2 in $[a, b]$. Suppose that $|B + u_a| \geq \text{gap}_{\min}(-B, [i_s, i_t]) \geq \text{gap}_{\min}(B, [i_s, i_t]) \geq \mu_{\text{th}}$. Here the first inequality holds by the definition of $\text{gap}_{\min}(-B, [i_s, i_t])$. Also, note that $u_j = u_a$ for all $j \in [a, b]$ by the definition of Structure 2. Let $\bar{y}_{a \to b} := \sum_{j=a}^{b} y_j/(b - a + 1)$. From the KKT conditions it can be shown that $\bar{y}_{a \to b} < u_a$. We provide intuitive explanation for the case $\Pi(\bar{y}_{a \to b}) = -B$. This can happen only when $\bar{y}_{a \to b} \leq -B$. Qualitatively in such a scenario, we expect the decision $-B$ to be much better than playing the decision u_a which is bigger than $-B$. Whenever there is sufficient gap (more formally a gap of at-least μ_{th}) between $-B$ and u_a, one can expect that u_a can be very sub-optimal in comparison to $-B$ $(= \Pi(\bar{y}_{a \to b}))$ which makes the term T_2 in Eq.(6.8) (with $w = -B$ and $u_j = u_a$) sufficiently negative.

When $\bar{y}_{a \to b} \in (-B, B)$, T_2 with $w = \bar{y}_{a \to b}$ can be shown to be sufficiently negative using the arguments of [65]. However, the interplay of this negative term with the sum of regret bounds in all refined intervals is more delicate as described in the proof of Lemma 191.

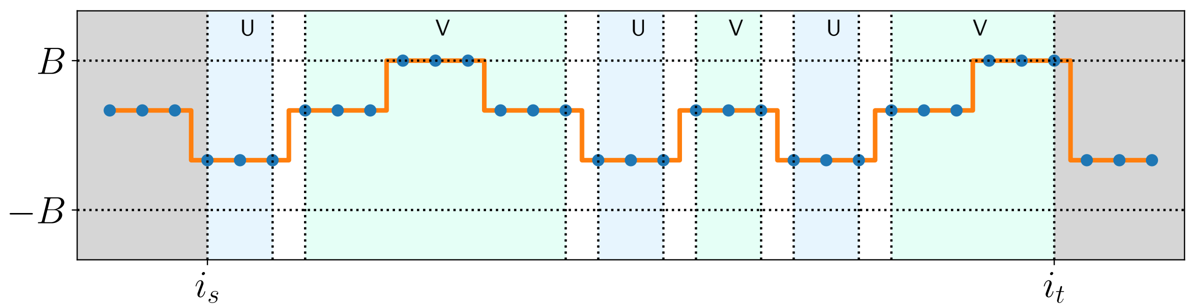


Figure 6.2: *Refinement of a bin* $[i_s, i_t] \in \mathcal{P}$ *that satisfy Condition 1 in Section 6.2.2 into smaller style U and style V sub-intervals. Blue dots represent the optimal sequence*

6.3 Performance guarantees for strongly convex losses

In this section, we extend the results on squared error losses to general strongly convex losses.

6.3.1 Strongly convex losses and box decision set

In this section, we show that the style of analysis presented for squared error losses directly generalizes to strongly convex losses in multi-dimensions whenever the decision set is an L_∞ norm ball. The main idea is to provide a reduction to the uni-variate squared loss games via standard surrogate loss tricks [77] and instantiate FLH-OGD appropriately. All unspecified proofs for this section are deferred to Appendix E.2. We consider the following protocol:

- At time $t \in [n]$ learner predicts $\boldsymbol{x}_t \in \mathbb{R}^d$ with $\|\boldsymbol{x}_t\|_\infty \leq B$.
- Adversary reveals loss f_t.
- Learner suffers loss $f_t(\boldsymbol{x}_t)$.

We have the following Corollary due to Theorem 50.

Corollary 54. *Let the loss functions f_t be H strongly convex in L_2 norm across the (box) domain $\mathcal{D} = \{\boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x}\|_\infty \leq B\}$. i.e, $f_t(\boldsymbol{y}) \geq f_t(\boldsymbol{x}) + \nabla f_t(\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{x}) + \frac{H}{2}\|\boldsymbol{y} - \boldsymbol{x}\|_2^2$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{D}$. Suppose $\|\nabla f_t(\boldsymbol{x})\|_\infty \leq G_\infty$ for all $\boldsymbol{x} \in \mathcal{D}$. For each $i \in [d]$, construct surrogate losses $\ell_t^{(i)} : \mathbb{R} \to \mathbb{R}$ as $\ell_t^{(i)}(x) = (x - (\boldsymbol{x}_t[i] - \nabla f_t(\boldsymbol{x}_t)[i]/H))^2$ where $\boldsymbol{x}_t$ is the prediction of the learner at time t. By running d instances of uni-variate FLH-OGD with decision set $[-B, B]$ and learning rate $\zeta = 1/(2(2B + G_\infty/H)^2)$ where instance i predicts $\boldsymbol{x}_t[i]$ at time t and suffers losses $\ell_t^{(i)}$, we have*

$$\sum_{t=1}^{n} f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{w}_t) = \tilde{O}\left(d^{1/3} n^{1/3} C_n^{2/3} \vee d\right), \tag{6.10}$$

for any comparator sequence $\boldsymbol{w}_{1:n}$ with $TV(\boldsymbol{w}_{1:n}) := \sum_{t=2}^{n} \|\boldsymbol{w}_t - \boldsymbol{w}_{t-1}\|_1 \leq C_n$. $\tilde{O}(\cdot)$ hides the dependence on factors of $\log n, B, H, G_\infty$.

When compared with the information theoretic lower bound of [65] (Proposition 11 there), we see that the rate of Theorem 50 is optimal (modulo log factors) wrt to n, C_n and d. The dependence of $\tilde{O}(d)$ for low C_n regimes is due to the fact that we only assume $\|\nabla f_t(\boldsymbol{x})\|_\infty = O(1)$ as opposed to assuming $\|\nabla f_t(\boldsymbol{x})\|_2 = O(1)$.

Remark 55. (relaxed assumptions & improvements) Unlike [65], we do not assume gradient Lipschitzness of the losses f_t. Further, for the box decision set, our results attain an optimal $O(d^{1/3})$ dimension dependence on regret in the non-trivial regime of $C_n \geq 1/n$ in comparison to the $O(d^2)$ dependence of [65] for strongly convex losses.

Remark 56. We emphasize that the theory developed in Section 6.2 is vital for extending the results with the surrogate losses as in Corollary 54. Consider squared losses $\ell_t(x) = (x-y_t)^2$ with labels y_t such that $|y_t| \leq Y$ for all t. [65] requires that the predictions x_t obey $x_t \in [-Y, Y]$. In our use case with surrogate losses $\ell_t^{(i)}(x) = (x - (\boldsymbol{x}_t[i] - \nabla f_t(\boldsymbol{x}_t)[i]/H))^2$ such a requirement can be not well defined. Here the labels can be regarded $y_t = \boldsymbol{x}_t[i] - \nabla f_t(\boldsymbol{x}_t)[i]/H$ which depends on $\boldsymbol{x}_t[i]$. As per the setup of Corollary 54, the i^{th} FLH-OGD instance uses losses $\ell_t^{(i)}$, $t \in [n]$ and its prediction at time t is $\boldsymbol{x}_t[i]$. So constructing a uniform bound Y to contain the predictions $\boldsymbol{x}_t[i]$ requires a uniform bound on the predictions $\boldsymbol{x}_t[i]$ itself for all t which is self conflicting. Hence the strategy of [65] for squared error losses is incompatible for using the surrogate losses $\ell_t^{(i)}$.

6.3.2 Strongly convex losses and general convex decision sets

In this section, we show how to convert an optimal algorithm described in Section 6.3 for the box decision set to an optimal (modulo factors of $\log n$ and dimensions dependencies) algorithm for any convex decision set via a black box reduction. This reduction is essentially due to the seminal work of [79].

We have the following guarantee for the scheme in Fig. 6.3.

Theorem 57. *Assume the notations in Fig. 6.3. Let the input decision set be $\mathcal{W}$. Let the losses be H strongly convex in L_2 norm across $\mathcal{D}$ and satisfy $\|\nabla f_t(\boldsymbol{x})\|_\infty \leq G$ for all $\boldsymbol{x} \in \mathcal{D}$. Then the reduction scheme in Fig. 6.3 guarantees that*

$$\sum_{t=1}^{n} f_t(\hat{\boldsymbol{x}}_t) - f_t(\boldsymbol{w}_t) = \tilde{O}\left(d^{1/3} n^{1/3} C_n^{2/3} \vee d\right), \tag{6.11}$$

for any comparator sequence $\boldsymbol{w}_{1:n} \in \mathcal{W}$ with $TV(\boldsymbol{w}_{1:n}) := \sum_{t=2}^{n} \|\boldsymbol{w}_t - \boldsymbol{w}_{t-1}\|_1 \leq C_n$. $\tilde{O}(\cdot)$ hides the dependence on factors of $\log n, H, G_\infty$.

Box to general convex set reduction: Inputs - Decision set $\mathcal{W}$, $G > 0$

1. Let $\mathcal{D}$ be the tightest box that circumscribes $\mathcal{W}$. i.e, $\mathcal{D} = \{\boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x}\|_\infty \leq \sup_{\boldsymbol{w} \in \mathcal{W}} \|\boldsymbol{w}\|_\infty\}$.
2. Let $\mathcal{A}$ be the algorithm attaining the guarantee in Corollary 54 with decision set $\mathcal{D}$ and $G_\infty = 2G$.
3. At round t, get iterate $\boldsymbol{x}_t$ from $\mathcal{A}$.
4. Play $\hat{\boldsymbol{x}}_t = \Pi_{\mathcal{W}}(\boldsymbol{x}_t) := \mathrm{argmin}_{\boldsymbol{y} \in \mathcal{W}} \|\boldsymbol{x}_t - \boldsymbol{y}\|_1$.
5. Get loss f_t.
6. Construct surrogate loss $\ell_t(\boldsymbol{x}) = f_t(\boldsymbol{x}) + G \cdot S(\boldsymbol{x})$, where $S(\boldsymbol{x}) := \|\boldsymbol{x} - \Pi_{\mathcal{W}}(\boldsymbol{x})\|_1$.
7. Send $\ell_t(\boldsymbol{x})$ to $\mathcal{A}$.

Figure 6.3: Black box reduction from box to arbitrary convex decision set. This technique is due to [79].

Proof. We start by listing several observations. First, note that the function $S(\boldsymbol{x})$ is convex and 1-Lipschitz across $\mathbb{R}^d$. (Proposition 1 in [79]).

Also, the sub-gradient $\partial S(\boldsymbol{x}) = \{\boldsymbol{y} \in \mathbb{R}^d : \boldsymbol{y}[j] = \mathrm{sign}\,(x[j] - \Pi_{\mathcal{W}}(\boldsymbol{x})[j]) \; j \in [d]\}$ (due to Theorem 4 in [79]). Here $\mathrm{sign}(a) = a/|a|$ if $|a| > 0$ and any number between $[-1, 1]$ otherwise.

Finally the surrogate losses ℓ_t are H strongly convex in $L2$ norm across $\mathcal{D}$, as adding a convex function to strongly convex function preserves strong convexity. However, ℓ_t are *not* gradient Lipschitz due to the component $G\|\boldsymbol{x} - \Pi_{\mathcal{W}}(\boldsymbol{x})\|_1$ being not smooth.

We have that for any $\boldsymbol{x} \in \mathcal{D}$,

$$\|\nabla \ell_t(\boldsymbol{x})\|_\infty \leq \|\nabla f_t(\boldsymbol{x})\|_\infty + G\|\partial S(\boldsymbol{x})\|_\infty \tag{6.12}$$

$$\leq 2G, \tag{6.13}$$

where the last line is due to the assumption that $\|\nabla f_t(\boldsymbol{x})\|_\infty \leq G$ and $\partial S(\boldsymbol{x})$ is just a vector of signs as established before.

Hence we have that the losses ℓ_t sent to algorithm $\mathcal{A}$ satisfy the conditions of Corollary 54 with $G_\infty = 2G$. Hence we have that

$$\sum_{t=1}^{n} \ell_t(\boldsymbol{x}_t) - \ell_t(\boldsymbol{w}_t) = \tilde{O}\left(d^{1/3} n^{1/3} C_n^{2/3} \vee d\right), \tag{6.14}$$

where $\boldsymbol{w}_{1:n}$ is as mentioned in the theorem statement.

By Taylor's theorem, we have that for some $\boldsymbol{z}$ in the line segment joining $\boldsymbol{x}_t$ and $\hat{\boldsymbol{x}}_t$

$$\begin{aligned} f_t(\hat{\boldsymbol{x}}_t) &= f_t(\boldsymbol{x}_t) + \nabla f_t(\boldsymbol{z})^T (\hat{\boldsymbol{x}}_t - \boldsymbol{x}_t) & (6.15) \\ &\leq f_t(\boldsymbol{x}_t) + G \|\hat{\boldsymbol{x}}_t - \boldsymbol{x}_t\|_1 & (6.16) \\ &= \ell_t(\boldsymbol{x}_t) & (6.17) \end{aligned}$$

where the inequality is due to Holder's inequality and the assumption that $\|\nabla f_t(\boldsymbol{x})\|_\infty \leq G$ for all $\boldsymbol{x} \in \mathcal{D}$.

Further for any $\boldsymbol{w}_t \in \mathcal{W}$, we have that $f_t(\boldsymbol{w}_t) = \ell_t(\boldsymbol{w}_t)$. Thus overall we obtain,

$$\sum_{t=1}^{n} f_t(\hat{\boldsymbol{x}}_t) - f_t(\boldsymbol{w}_t) \leq \sum_{t=1}^{n} \ell_t(\boldsymbol{x}_t) - \ell_t(\boldsymbol{w}_t). \tag{6.18}$$

Combining Eq.(6.14) and (6.18) now yeilds the theorem. □

Remark 58. We emphasize that the removal of gradient smoothness assumption for strongly convex losses (from [65]) as done in the current work was important to apply the reduction scheme of Fig.6.3 as the losses ℓ_t are not gradient smooth.

6.4 Performance guarantees for exp-concave losses

In this section, we control the dynamic regret with exp-concave and gradient smooth losses when the decision set is an L_∞ ball. All unspecified lemma statements and proofs are deferred to Appendix E.3. We make the following assumptions:

Assumption B1: The loss functions ℓ_t are α exp-concave in the box decision set $\mathcal{D} = \{\boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x}\|_\infty \leq B\}$ .ie, $\ell_t(\boldsymbol{y}) \geq \ell_t(\boldsymbol{x}) + \nabla \ell_t(\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{x}) + \frac{\alpha}{2}\left(\nabla \ell_t(\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{x})\right)^2$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{D}$.

Assumption B2: The loss functions ℓ_t satisfy $\|\nabla \ell_t(\boldsymbol{x})\|_2 \leq G$ and $\|\nabla \ell_t(\boldsymbol{x})\|_\infty \leq G_\infty$ for all $\boldsymbol{x} \in \mathcal{D}$. Without loss of generality, we let $G \wedge G_\infty \wedge B \geq 1$, where $a \wedge b := \min\{a, b\}$.

We consider the following protocol:

- At time $t \in [n]$ learner predicts $\boldsymbol{x}_t \in \mathbb{R}^d$ with $\|\boldsymbol{x}_t\|_\infty \leq B$.
- Adversary reveals the loss function ℓ_t.

In view of Assumption B1, following [77], one can define the surrogate losses:

$$f_t(\boldsymbol{x}) = \left(\sqrt{\alpha/2}\nabla \ell_t(\boldsymbol{x}_t)^T(\boldsymbol{x} - \boldsymbol{x}_t) + 1/\sqrt{2\alpha}\right)^2. \tag{6.19}$$

It follows that

$$\sum_{t=1}^{n} \ell_t(\boldsymbol{x}_t) - \ell_t(\boldsymbol{w}_t) \leq \sum_{t=1}^{n} f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{w}_t), \tag{6.20}$$

where $\boldsymbol{x}_t, \boldsymbol{w}_t \in \mathcal{D}$.

Further, we make two useful observations about surrogate losses f_t.

First for $\boldsymbol{x} \in \mathcal{D}$, since $\left|\sqrt{\alpha/2}\nabla \ell_t(\boldsymbol{x}_t)^T(\boldsymbol{x} - \boldsymbol{x}_t) + 1/\sqrt{2\alpha}\right| \leq 2GB\sqrt{\alpha d/2} + 1/\sqrt{2\alpha} := \gamma$, we have that f_t are $1/(2\gamma^2)$ exp-concave over $\mathcal{D}$ (see Section 3.3 in [40]).

Second, since $\nabla^2 f_t(\boldsymbol{x}) = \nabla \ell_t(\boldsymbol{x}_t)\nabla \ell_t(\boldsymbol{x}_t)^T \preccurlyeq G^2\boldsymbol{I}$, we have that the losses f_t are G^2 gradient Lipschitz over $\mathcal{D}$.

We are interested in controlling the regret:

$$R_n(C_n) := \sup_{\substack{\boldsymbol{w}_1,\ldots,\boldsymbol{w}_n \in \mathcal{D} \\ \sum_{t=2}^n \|\boldsymbol{w}_t - \boldsymbol{w}_{t-1}\|_1 \le C_n}} \sum_{t=1}^n \ell_t(\boldsymbol{x}_t) - \ell_t(\boldsymbol{w}_t), \tag{6.21}$$

where $\boldsymbol{x}_t$ is the decisions of the algorithm.

We have the following performance guarantee when the losses are exp-concave.

Theorem 59. *Suppose Assumptions B1-B2 are satisfied. Define* $\gamma := 2GB\sqrt{\alpha d/2} + 1/\sqrt{2\alpha}$. *By using the base learner as ONS with parameter* $\zeta = \min\left\{\frac{1}{16GB\sqrt{d}}, 1/(4\gamma^2)\right\}$, *decision set* $\mathcal{D}$, *loss at time* t *to be* f_t *and choosing learning rate of FLH as* $\eta = 1/(2\gamma^2)$, *FLH-ONS obeys*

$$\begin{aligned} R_n(C_n) &\le \sup_{\substack{\boldsymbol{w}_1,\ldots,\boldsymbol{w}_n \in \mathcal{D} \\ \sum_{t=2}^n \|\boldsymbol{w}_t - \boldsymbol{w}_{t-1}\|_1 \le C_n}} \sum_{t=1}^n f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{w}_t) && (6.22) \\ &= \tilde{O}\left(140 d^2 (8G^2B^2\alpha d + G^2B^2 + 1/\alpha)(n^{1/3}C_n^{2/3} \vee 1)\right) \mathbb{I}\{C_n > 1/n\} && (6.23) \\ &\quad + \tilde{O}\left(d(8G^2B^2\alpha d + 1/\alpha\right) \mathbb{I}\{C_n \le 1/n\}, && (6.24) \end{aligned}$$

where $\boldsymbol{x}_t$ *is the decision of the algorithm at time* t *and* $\tilde{O}(\cdot)$ *hides polynomial factors of* $\log n$. $\mathbb{I}\{\cdot\}$ *is the boolean indicator function assuming values in* $\{0,1\}$.

Remark 60. (**relaxed assumptions & improvements**) In [65], it is assumed that the losses are gradient Lipschitz and exp-concave over an enlarged set $\mathcal{D}^\dagger = \{\boldsymbol{x} : \|\boldsymbol{x}\|_\infty \le B + G\}$ where B and G are as in Assumptions B1-B2. While our proper learning results doesn't require gradient Lipschitzness and require exp-concavity to hold in the smaller constraint set $\mathcal{D}$ as in Assumption B1. Further [65] attains a worse dependence of $O(d^{3.5})$ in the non-trivial regime $C_n \ge 1/n$.

Further, we show in Appendix E.4 that when the decision set is a polytope satisfying certain conditions, we can reparametrize the original problem into the framework of box constrained online learning with exp-concave losses.

6.4.1 Road map for the proof of Theorem 59

The proof of Theorem 59 is facilitated by generalising the arguments used for proving Theorem 50. We first form a coarse partition of $[n]$ namely $\mathcal{P}$ in Lemma 193 by a direct extension of Lemma 184. For the regime where dual varaible $\lambda = O(d^{1.25}n^{1/3}/C_n^{1/3})$, we employ a two term regret decomposition for each bin $[i_s, i_t] \in \mathcal{P}$ as follows:

$$\underbrace{\sum_{j=i_s}^{i_t} f_j(\boldsymbol{x}_j) - f_j(\check{\boldsymbol{u}}_i)}_{T_{1,i}} + \underbrace{\sum_{j=i_s}^{i_t} f_j(\check{\boldsymbol{u}}_i) - f_j(\boldsymbol{u}_j)}_{T_{2,i}}, \tag{6.25}$$

where $\boldsymbol{x}_j$ is the prediction of the FLH-ONS algorithm and $\boldsymbol{u}_{1:n}$ is the offline optimal sequence in Lemma 192. We exhibit a choice of $\check{\boldsymbol{u}}_i \in \mathcal{D}$ in Lemma 196 so that $T_{1,i} + T_{2,i}$ when summed across all bins $[i_s, i_t] \in \mathcal{P}$ yield a total regret of $\tilde{O}^*(n^{1/3}C_n^{2/3} \vee 1)$.

For handling the alternate regime $\lambda = \Omega(d^{1.25}n^{1/3}/C_n^{1/3})$, we provide a refinement scheme `fineSplit` in Fig.E.4 in Appendix E.3. Specifically let $\mathcal{R}$ be the set of all intervals in $\mathcal{P}$ that satisfy the prerequisite of `fineSplit` procedure. Let $\mathcal{S} := \mathcal{P} \setminus \mathcal{R}$.

For each interval in $\mathcal{R}$, we invoke `fineSplit`. This refinement scheme splits the original interval into sub-bins that satisfy either the properties in Lemma 203 (which can be regarded as a generalization of style U sub-bins in Section 6.2.2) or Lemma 204 (which can be regarded as a generalization of style V sub-bins in Section 6.2.2). Sub-bins that satisfy condition in Lemma 203 is termed as style U^+ sub-bins and those that satisfy condition in Lemma 204 is termed as style V^+ sub-bins henceforth for brevity. Sub-bins satisfying conditions of both Lemmas 203 and 204 are regarded as style U^+ sub-bins. For each such sub-bin $[a, b]$, we employ a two term regret decomposition as follows:

$$\underbrace{\sum_{j=a}^{b} f_j(\boldsymbol{x}_j) - f_j(\check{\boldsymbol{u}}_j)}_{T_1} + \underbrace{\sum_{j=a}^{b} f_j(\check{\boldsymbol{u}}_j) - f_j(\boldsymbol{u}_j)}_{T_2}. \tag{6.26}$$

We term the sequence $\check{\boldsymbol{u}}_{a:b}$ as the *ghost sequence* as they are fictitious intermediate comparator sequence introduced solely for the purpose of analysis. We provide a mechanical way of generating an appropriate ghost sequence in the `generateGhostSequence` procedure in Fig.E.3 which satisfies the properties stated in Lemma 198. Of particular interest is how we choose the ghost sequence for style U^+ sub-bins. Suppose for a style U^+ sub-bin $[a, b]$, let $k \in [d]$ be the coordinate where the offline optimal takes the form of Structure 1 or Structure 2 (see Definition 200). Then we set for all $j \in [a, b]$:

$$\check{\boldsymbol{u}}_j[k] = \Pi\left(\boldsymbol{u}_a[k] - \frac{1}{(b-a+1)\beta}\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k]\right), \tag{6.27}$$

where $\Pi(\cdot)$ is the projection to $[-B, B]$ and $\beta := G^2$. This choice is very different from the unprojected gradient descent update used in [65]. It can be viewed as a lazy projected gradient descent like update (with step size $1/((b-a+1)\beta)$) where the update operation is performed only across coordinate k. Note that it is not exactly gradient descent across coordinate k since in the second term above we are using $\nabla f_j(\boldsymbol{u}_j)[k]$ instead of $\nabla f_j(\boldsymbol{u}_a)[k]$.

The choice of $\check{\boldsymbol{u}}_j[k']$ for $k' \neq k$ is more involved and is accomplished by carefully selecting a sequence that switches only $O(1)$ times and assumes values in $[(\boldsymbol{u}_a[k'] \wedge \ldots \wedge \boldsymbol{u}_b[k']), (\boldsymbol{u}_a[k'] \vee \ldots \vee \boldsymbol{u}_b[k'])]$ as mentioned in `generateGhostSequence` procedure in Fig.E.3 in Appendix E.3.

Next, by using similar gap criteria used in Section 6.2.2 and exploiting gradient Lipschitzness, we show that T_1+T_2 in Eq.(6.26) can be bounded by $O^*(\log n)+$ a negative term for each style U^+ sub-bin obtained by refining bins in $\mathcal{R}$. For each style V^+ sub-bin, the regret is bounded by $O^*(\log n)$ (see Lemma 199). When such bounds are added for all sub-bins generated by invoking `fineSplit` on every interval in $\mathcal{R}$, we show that the negative terms gracefully offset the culmination of $O^*(\log n)$ terms to result in a regret bound of $\tilde{O}^*(n^{1/3}C_n^{2/3} \vee 1)$ (see Proof of Lemma 207).

The regret contribution from all bins in $\mathcal{S}$ is bounded by $\tilde{O}^*(n^{1/3}C_n^{2/3} \vee 1)$ using Lemma 199. Finally summing the regret contributions from bins in $\mathcal{R}$ and $\mathcal{S}$ yield the theorem.

6.5 Concluding Discussion

In this work we presented a new analysis that extends the results of [65] and showed near optimal universal dynamic regret in a proper learning setting for strongly convex losses. Results on the special case of exp-concave losses and box decision set are also derived. Further we relaxed the gradient Lipschitzness assumption for losses revealed and derived regret rates with improved dependence on d.

An important open problem is to extend these results for exp-concave losses with general convex decision sets.

Chapter 7

Second Order Path Variationals in Non-Stationary Online Learning

In this chapter, we generalize the setting of Chapter 3 to the framework of OCO. We focus on the case of $\mathcal{TV}^{(1}$ class.

We recall the definition of discrete TV sequence class.

$$\mathcal{TV}^{(k)}(C_n) := \left\{ \boldsymbol{w}_{1:n} \middle| n^k \|D^{k+1} \boldsymbol{w}_{1:n}\|_1 \leq C_n \text{ where each } \boldsymbol{w}_t \in \mathcal{W} \right\}. \tag{7.1}$$

As noted in [54], this class features sequences such that along any coordinate $j \in [d]$, the sequence $\boldsymbol{w}_1[j], \ldots, \boldsymbol{w}_n[j]$ is obtained via sampling a function $f_j(x) \in \mathcal{F}_k(C_{n,j})$ at points $x = i/n$ for $i \in [n]$ with the property that $\sum_{j=1}^{d} C_{n,j} = C_n$.

Why is this useful? In this chapter, we focus on comparators that reside in the $\mathcal{TV}^{(1)}(C_n)$ class. Our goal will be to bound the dynamic regret against $\boldsymbol{w}_{1:n} \in \mathcal{TV}^{(1)}(C_n)$ as a function of n and C_n. We emphasize that our algorithm does not take C_n as an input and can *simultaneously* compete with the TV1 family described by *any* $C_n > 0$. As discussed earlier, comparators in $\mathcal{TV}^{(1)}(C_n)$ class exhibits a piece-wise linear structure across each coordinate (see Definition 61). The points where the sequence transition from one linear structure to other can be interpreted as abrupt changes or events in the underlying comparator dynamics. Many real world time series data are known to contain piece-wise linear trends. See for example Fig.7.2 or [4] for more examples. Hence controlling the dynamic regret against comparators from $\mathcal{TV}^{(1)}(C_n)$ class has *significant practical value.*

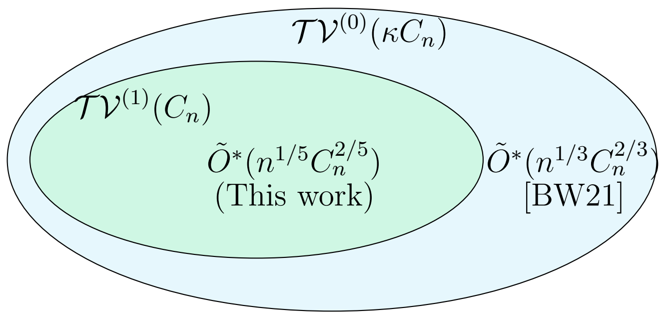


Figure 7.1: *Hierarchy of TV classes for the comparator sequence and the corresponding optimal dynamic regret rates under exp-concave and gradient smooth losses. Here $\tilde{O}^*$ hides dependencies on d and $\log n$. κ is a constant independent of n and C_n. We assume $C_n = \Omega(1)$. BW21 refers to the work of [65].*

Fast rate phenomenon. Sequence classes of the form $\mathcal{TV}^{(1)}(C_n)$ or more generally $\mathcal{TV}^{(k)}(C_n)$ have gained significant attention and have been the subject of extensive study in the non-parametric regression community for over two decades [1, 2, 4, 80, 38]. These works aim to estimate an unknown scalar (i.e $d = 1$) sequence $\theta_{1:n} \in \mathcal{TV}^{(k)}(C_n)$ from n noisy observations $y_t = \theta_t + \mathcal{N}(0, \sigma^2)$ in an offline setting. They propose algorithms that produce estimates $\hat{\theta}_t, t \in [n]$ such that the expected total squared error $\sum_{t=1}^n E[(\hat{\theta}_t - \theta_t)^2]$ is controlled. In particular, for the case when $\theta_{1:n} \in \mathcal{TV}^{(1)}(C_n)$ a (near) optimal estimation rate of $\tilde{O}(n^{1/5}C_n^{2/5})$ is shown to be attainable for the squared loss ($\tilde{O}(\cdot)$ hides poly-logarithmic factors of n). This rate is faster than the typical $O(\sqrt{nC_n})$ or $\tilde{O}(n^{1/3}C_n^{2/3})$ dynamic regret rates found in non-stationary online learning literature (see for eg. [26, 65]).

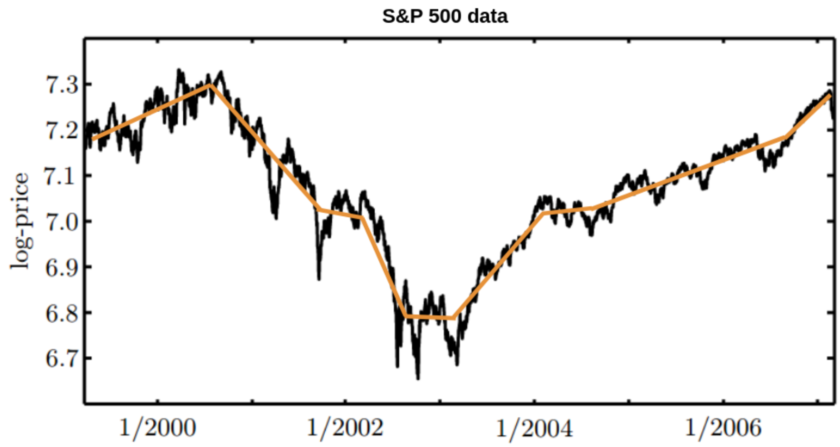
Central question and summary of results. A natural question that we ask here is:

Can we attain a universal dynamic regret of $\tilde{O}^*(n^{1/5}C_n^{2/5})$ when the comparators $\boldsymbol{w}_{1:n} \in \mathcal{TV}^{(1)}(C_n)$ and the losses being exp-concave?

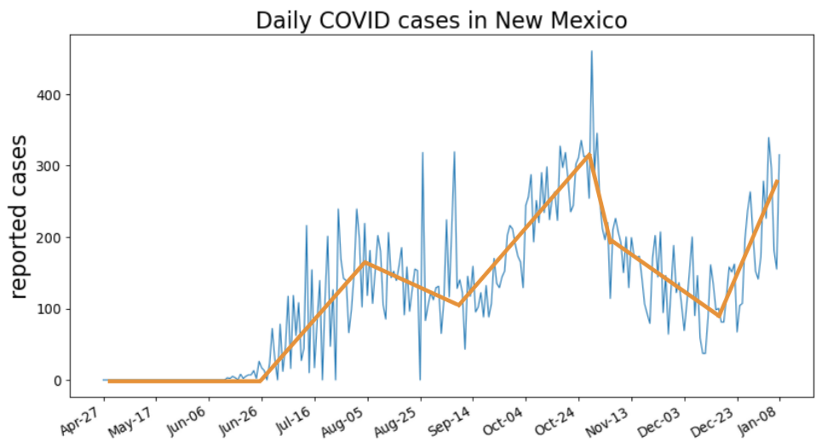
Here O^* hides dimension dependencies. A starting point in answering this question is to exploit the piece-wise linear structure of sequences in $\mathcal{TV}^{(1)}(C_n)$ across each coordinate. A sequence that is linear across each coordinate within some interval can be perfectly described using a *fixed* vector $\boldsymbol{u} \in \mathbb{R}^{2d}$ where $\boldsymbol{u}[2k-1:2k] \in \mathbb{R}^2$ specifies the slope and intercept along coordinate $k \in [d]$. We will call such $\boldsymbol{u}$ to be a *linear predictor.* If an algorithm guarantees that its *static regret* against fixed linear predictors within *any* interval is controlled, one can hope to perform nearly as well as the comparator sequence $\boldsymbol{w}_{1:n} \in \mathcal{TV}^{(1)}(C_n)$. This is precisely an application of Strongly Adaptive algorithms [23, 24, 52, 81] which aim to control their static regret in any interval and hence we can use them off-the-shelf to achieve our goal. We refer the reader to Section 7.1 for more details. Below, we briefly summarize our contributions:

- We show that by using appropriate Strongly Adaptive algorithms, one can attain the (near) *optimal* universal dynamic regret rate of $\tilde{O}(d^2 n^{1/5} C_n^{2/5} \vee d^2)$ (Theorem 63; $a \vee b = \max\{a, b\}$) whenever the comparators $\boldsymbol{w}_{1:n} \in \mathcal{TV}^{(1)}(C_n)$ and the losses are exp-concave and gradient smooth (see Section 7.2 for the list of Assumptions and associated definitions). Further this rate is attained *without prior knowledge* of C_n.

- To the best of our knowledge, we are the *first* to introduce path variationals based on second order differences to the setting of *adversarial* online learning. We show how to import the *fast rate* phenomenon observed in stochastic non-parametric regression problem under squared loss into the problem of controlling *universal* dynamic regret under general exp-concave losses with no stochastic assumptions.



(a) S&P500



(b) Daily COVID cases

Figure 7.2: *Fig.(a) displays S&P500 stock price data and Fig.(b) displays Daily COVID cases reported in the state of New Mexico, USA. In both scenarios we can see that the underlying trend (obtained via an L1 Trend Filter [4]) exhibits a weakly differentiable piece-wise linear structure (orange) which belongs to an appropriate* $\mathcal{TV}^{(1)}$ *class.*

7.1 The Algorithm

In this section, we formally describe the main algorithm FLH-SIONS (Follow the Leading History-Scale Invariant Online Newton Step) in Fig.7.3 and provide intuition on why it can favorably control the dynamic regret against comparators from $\mathcal{TV}^{(1)}$ class.

FLH-SIONS: inputs: exp-concavity factor σ and n SIONS base learners $E^1, \ldots, E^n$ initialized with parameters $\epsilon = 2$, $\eta = \sigma$ and $C = 20$. (see Fig. 7.4)

1. For each t, $v_t = (v_t^{(1)}, \ldots, v_t^{(t)})$ is a probability vector in $\mathbb{R}^t$. Initialize $v_1^{(1)} = 1$.
2. For any SIONS expert E_j with $j \leq t$, define $\boldsymbol{x}_j^{(t)} = [1, t-j+1]^T$ to be given to E_j at time t before making its prediction $E_j(t) \in \mathbb{R}^d$.
3. In round t, set $\forall j \leq t$, $\boldsymbol{y}_t^j \leftarrow E_j(t)$ (the prediction of the j^{th} base learner at time t). Play $\boldsymbol{p}_t = \sum_{j=1}^t v_t^{(j)} \boldsymbol{y}_t^{(j)}$.
4. After receiving f_t, set $\hat{v}_{t+1}^{(t+1)} = 0$ and perform update for $1 \leq i \leq t$:

$$\hat{v}_{t+1}^{(i)} = \frac{v_t^{(i)} e^{-\sigma f_t(\boldsymbol{x}_t^{(i)})}}{\sum_{j=1}^t v_t^{(j)} e^{-\sigma f_t(\boldsymbol{x}_t^{(j)})}} \tag{7.2}$$

5. Addition step - Set $v_{t+1}^{(t+1)}$ to $1/(t+1)$ and for $i \neq t+1$:

$$v_{t+1}^{(i)} = (1 - (t+1)^{-1}) \hat{v}_{t+1}^{(i)} \tag{7.3}$$

Figure 7.3: FLH algorithm of [23] with SIONS (see Fig.7.4) base experts

For the sake of simplicity, we capture the intuition in a uni-variate setting where the comparators $w_t \in \mathcal{W} \subset \mathbb{R}$ for all $t \in [n]$.

Definition 61. Within an interval $[a, b]$, we say that the comparator $w_{a:b}$ is a *linear signal* or assumes a *linear structure* if the slope $w_{t+1} - w_t$ is constant for all $t \in [a, b-1]$.

As described before, we are interested in competing against comparator sequences $w_{1:n}$ that have a piece-wise linear structure (across each coordinate in multi-dimensions). The durations / intervals of $[n]$ where the comparator is a fixed linear signal is unknown to the learner. Suppose that an ideal oracle provides us with the exact locations of these intervals of $[n]$. Consider an interval $[a, b]$ provided by the oracle where the comparator has a fixed linear structure given by $w_t = \boldsymbol{\mu}^T \boldsymbol{x}_a^{(t)}$ for the *co-variates* $\boldsymbol{x}_a^{(t)} := [1, t-a+1]^T$

SIONS: inputs: exp-concavity factor η, $\epsilon > 0$ and $C > 0$.

1. For any round t, we define $\tilde{f}_t(\boldsymbol{v}) = f_j(\boldsymbol{x}_t^T \boldsymbol{v}[1 : 2], \boldsymbol{x}_t^T \boldsymbol{v}[3 : 4], \ldots, \boldsymbol{x}_t^T \boldsymbol{v}[2k-1 : 2k])$ for any vector $\boldsymbol{v} \in \mathbb{R}^{2d}$.
2. At round $t+1$:
   (a) Receive co-variate $\boldsymbol{x}_{t+1} \in \mathbb{R}^2$.
   (b) Let $\mathcal{K}_{t+1} = \{\boldsymbol{w} \in \mathbb{R}^{2d} : |\boldsymbol{x}_{t+1}^T \boldsymbol{w}[2k-1 : 2k]| \leq C$ for all $k \in [d]\}$.
   (c) Let $\boldsymbol{A}_t = \epsilon \boldsymbol{I}_{2d} + \eta \sum_{j=1}^{t} \nabla \tilde{f}_j(\boldsymbol{v}_j) \nabla \tilde{f}_j(\boldsymbol{v}_j)^T$.
   (d) Let $\boldsymbol{u}_{t+1} = \boldsymbol{v}_t - \boldsymbol{A}_t^{-1} \nabla \tilde{f}_t(\boldsymbol{v}_t)$.
   (e) Let $\boldsymbol{v}_{t+1} = \operatorname{argmin}_{\boldsymbol{w} \in \mathcal{K}_{t+1}} \|\boldsymbol{w} - \boldsymbol{u}_{t+1}\|_{\boldsymbol{A}_t}$.
   (f) Play $\boldsymbol{w}_{t+1} \in \mathbb{R}^d$ such that $\boldsymbol{w}_{t+1}[k] = \boldsymbol{x}_{t+1}^T \boldsymbol{v}_{t+1}[2k-1 : 2k]$ for all $k \in [d]$.

Figure 7.4: An instance of SIONS algorithm from [82].

and $\boldsymbol{\mu}$ such that $|w_t|$ is $O(1)$ bounded for all $t \in [a, b]$. An effective strategy for the learner is to deploy an online algorithm E_a that starts from time a such that within the interval $[a, b]$ its regret:

$$R_{[a,b]}(\boldsymbol{\mu}) := \sum_{t=a}^{b} f_t\left(E_a(t)\right) - f_t(\boldsymbol{\mu}^T \boldsymbol{x}_a^{(t)}) \tag{7.4}$$

is controlled. Here $E_a(t)$ is the predictions of the algorithm E_a at time t. Under exp-concave losses, an $O(\log n)$ bound on the above regret can be achieved by the SIONS algorithm (Fig.7.4) from ([82], Theorem 2) run with co-variates $\boldsymbol{x}_a^{(t)}$.

In practice, the locations of such ideal intervals are unknown to us. So we maintain a pool of n base SIONS experts in Fig.7.3 where the expert E_τ starts at time τ with the monomial co-variate $\boldsymbol{x}_\tau^{(t)} = [1, t - \tau + 1]^T$ for all $t \geq \tau$. The adaptive regret guarantee of FLH with exp-concave losses (due to [23], Theorem 3.2) keeps the regret wrt *any* base expert to be small. In particular, FLH-SIONS satisfies that

$$\sum_{t=\tau}^{j} f_t(p_t) - f_t\left(E_\tau(t)\right) = O(\log n), \tag{7.5}$$

where p_t are the predictions of FLH-SIONS and $j \geq \tau$ for *any* $\tau \in [n]$. Hence for the interval $[a, b]$ given by the ideal oracle, it follows that

$$\sum_{t=a}^{b} f_t(p_t) - f_t(\boldsymbol{\mu}^T \boldsymbol{x}_a^{(t)}) \leq \sum_{t=a}^{b} f_t\left(E_a(t)\right) - f_t(\boldsymbol{\mu}^T \boldsymbol{x}_a^{(t)}) + O(\log n) \tag{7.6}$$

$$= R_{[a,b]}(\boldsymbol{\mu}) + O(\log n) = O(\log n), \tag{7.7}$$

where in the last equation, we appealed to the logarithmic static regret of SIONS from [82]. As a minor technical remark, we note that the original results of [82] assume that the losses are of the form $\tilde{f}_j(\boldsymbol{w}) = f_j(\boldsymbol{x}_j^T \boldsymbol{w})$ for a uni-variate function f_j. However, we show in Lemma 222 (in Appendix) that their regret bounds can be straightforwardly extended to handle multivariate losses $f_j$as in Line 1 of Fig.7.4 which is useful in our multi-dimensional setup.

Thus ultimately, the regret of the FLH-SIONS procedure is well controlled within each interval provided by the ideal oracle, thus allowing us to be competent against the piece-wise linear comparator sequence from a $\mathcal{TV}^1$ class. We remark that while both FLH and SIONS are well-known existing algorithms, our use of them with monomial co-variates is new. Our dynamic regret analysis is new too, which uncovers previously unknown properties of a particular combination of these existing algorithmic components using novel proof techniques.

7.2 Main Results

In this section, we explain the assumptions used and the main results of this chapter. Then we provide the proof summary for Theorem 63 in a uni-variate setting highlighting the technical challenges overcome along the way. Following which we explain how to handle multiple dimensions by constructing suitable reductions that will allow us to re-use much of the analytical machinery developed for the case of uni-variate setting. The following are the assumptions made.

A1. For all $t \in [n]$, the comparators $\boldsymbol{w}_t$ belongs to a given benchmark space $\mathcal{W} \subset \mathbb{R}^d$. Further we have $\mathcal{W} \subseteq [-1,1]^d$.

A2. The loss function $f_t : \mathbb{R}^d \rightarrow \mathbb{R}$ revealed at time t is 1-Lipschitz in $\| \cdot \|_2$ norm over the interval $[-20, 20]^d$.

A3. The losses f_t are 1-gradient Lipschitz over the interval $[-20, 20]^d$. This implies that $f_t(\boldsymbol{y}) \leq f_t(\boldsymbol{x}) + \nabla f_t(\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{x}) + \frac{L}{2}\|(\boldsymbol{y} - \boldsymbol{x})\|_2^2$ for all $\boldsymbol{x}, \boldsymbol{y} \in [-20, 20]^d$.

A4. The losses f_t are σ exp-concave over $[-20, 20]^d$. This implies that $f_t(\boldsymbol{y}) \geq f_t(\boldsymbol{x}) + \nabla f_t(\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{x}) + \frac{\sigma}{2}\left(\nabla f_t(\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{x})\right)^2$ for all $\boldsymbol{x}, \boldsymbol{y} \in [-20, 20]^d$.

Assumptions A3 and A4 ensure the smoothness and curvature of the losses which we crucially rely to derive fast regret rates. Assumptions about Lipschitzness as in A2 are usually standard in online learning. In assumption A1 we consider comparators that belong to an interval that is smaller than the intervals in other assumptions. This is due to the fact that we allow our algorithms to be improper in the sense that the decisions of the algorithm may lie outside the benchmark space $\mathcal{W}$.

We start with a lower bound on the dynamic regret which is obtained by adapting the arguments in [2] to the case of bounded sequences as in Assumption A1. See Appendix F.3 for a proof.

Proposition 62. *Under Assumptions A1-A4, any online algorithm necessarily suffers* $\sup_{\boldsymbol{w}_{1:n} \in \mathcal{TV}^{(1)}(C_n)} R_n(\boldsymbol{w}_{1:n}) = \Omega(d^{3/5} n^{1/5} C_n^{2/5} \vee d)$.

We have the following guarantee for FLH-SIONS.

Theorem 63. *Let* $\boldsymbol{p}_t$ *be the predictions of FLH-SIONS algorithm with parameters* $\epsilon = 2$, $C = 20$ *and exp-concavity factor* σ. *Under Assumptions A1-A4, we have that,*

$$\sum_{t=1}^{n} f_t(\boldsymbol{p}_t) - f_t(\boldsymbol{w}_t) = \tilde{O}(d^2 n^{1/5} C_n^{2/5} \vee d^2),$$

for any $C_n > 0$ *and any comparator sequence* $\boldsymbol{w}_{1:n} \in \mathcal{TV}^{(1)}(C_n)$. *Here* $\tilde{O}$ *hides poly-logarithmic factors of* n *and* $a \vee b = \max\{a, b\}$.

Remark 64. Compared with the lower bound in Proposition 62, we conclude that the regret rate of the above theorem is optimal modulo factors of d and $\log n$. We note that the guarantee of Theorem 63 is truly universal as no apriori knowledge of C_n is required.

Proposition 65. *It can be shown that the same algorithm FLH-SIONS under the setting of Theorem 220 enjoys optimal rates against comparators from the* $\mathcal{TV}^0(C_n)$ *class as well. When combined with Theorem 63 we conclude that under Assumptions A1-A4, FLH-SIONS attains an adaptive guarantee of*

$$\sum_{t=1}^{n} f_t(\boldsymbol{p}_t) - f_t(\boldsymbol{w}_t) = \tilde{O}(d^2 \min\{n^{1/3} \|D\boldsymbol{w}_{1:n}\|_1^{2/3}, n^{1/5}(n\|D^2\boldsymbol{w}_{1:n}\|_1)^{2/5}\} \vee d^2),$$

for any comparator sequence $\boldsymbol{w}_{1:n}$. *Here* $\tilde{O}$ *hides poly-logarithmic factors of* n *and* $a \vee b = \max\{a, b\}$. *See Appendix F.2 for a proof.*

Remark 66. One may ask if a simpler algorithm such as carefully tuned online gradient descend (OGD) can enjoy these fast rates too. However, Proposition 2 of [54] implies that properly tuned OGD algorithm which is optimal against comparators in $\mathcal{TV}^{(0)}$

class under convex losses, necessarily suffers a slower dynamic regret of $\Omega(n^{1/4})$ against comparators in $\mathcal{TV}^{(1)}(1)$ class under exp-concave losses [83].

7.2.1 Proof Summary of Theorem 63 for one dimension

In what follows, we present several useful lemmas and provide a running sketch on how to chain them to arrive at Theorem 63 in a uni-variate setting (i.e $d = 1$). Detailed proofs are deferred to Appendix F.1.1.

Suppose that we need to compete against comparators whose TV1 distance (i.e $n\|D^2 w_{1:n}\|_1$) is bounded by some number C_n. This quantity could be unknown to the algorithm. Consider the *offline oracle* who has access to the entire sequence of loss functions $f_1, \ldots, f_n$ and the TV1 bound C_n. It may then solve for the strongest possible comparator respecting the TV1 bound through the following convex optimization problem.

$$\min_{\tilde{u}_1, \ldots, \tilde{u}_n} \quad \sum_{t=1}^{n} f_t(\tilde{u}_t) \tag{7.8a}$$

$$\text{s.t.} \quad \|D^2 \tilde{u}_{1:n}\|_1 \leq C_n/n, \tag{7.8b}$$

$$-1 \leq \tilde{u}_t \; \forall t \in [n], \tag{7.8c}$$

$$\tilde{u}_t \leq 1 \; \forall t \in [n], \tag{7.8d}$$

Let $u_1, \ldots, u_n$ be the optimal solution of the above problem. This sequence will be referred as **offline optimal** hence-forth. Clearly we have that the regret against any comparator sequence $w_{1:n} \in \mathcal{TV}^1(C_n)$ obeys

$$\sum_{t=1}^{n} f_t(p_t) - f_t(w_t) \leq \sum_{t=1}^{n} f_t(p_t) - f_t(u_t), \tag{7.9}$$

and hence it suffices to bound the right side of the above inequality.

Lemma 67. ***(KKT conditions)*** *Let $u_1, \dots, u_n$ be the optimal primal variables and let $\lambda \geq 0$ be the optimal dual variable corresponding to the constraint* (7.8b)*. Further, let $\gamma_t^- \geq 0, \gamma_t^+ \geq 0$ be the optimal dual variables that correspond to constraints* (7.8c) *and* (7.8d) *respectively for all $t \in [n]$. By the KKT conditions, we have*

- ***stationarity:*** $\nabla f_t(u_t) = \lambda\left((s_{t-1} - s_t) - (s_{t-2} - s_{t-1})\right) + \gamma_t^- - \gamma_t^+$*, where* $s_t = \mathrm{sign}((u_{t+2} - u_{t+1}) - (u_{t+1} - u_t))$*. Here* $\mathrm{sign}(x) = x/|x|$ *if* $|x| > 0$ *and any value in* $[-1, 1]$ *otherwise. For convenience of notations, we also define* $s_{-1} = s_0 = s_{n-1} = s_n = 0$*.*

- ***complementary slackness:*** *(a)* $\lambda\left(\|D^2 u_{1:n}\|_1 - C_n/n\right) = 0$*; (b)* $\gamma_t^-(u_t + 1) = 0$ *and* $\gamma_t^+(u_t - 1) = 0$ *for all* $t \in [n]$

Next, we provide a partition of the horizon with certain useful properties.

Lemma 68. ***(key partition)*** *For some interval $[a, b] \in [n]$, define $\ell_{a \to b} := b - a + 1$. There exists a partitioning of the time horizon $\mathcal{P} := \{[1_s, 1_t], \dots, [i_s, i_t], \dots [M_s, M_t]\}$ where $M = |\mathcal{P}|$ such that for any bin $[i_s, i_t] \in \mathcal{P}$ we have: 1) $\|D^2 u_{i_s:i_t}\|_1 \leq 1/\ell_{i_s \to i_t}^{3/2}$; 2) $\|D^2 u_{i_s:i_t+1}\|_1 > 1/\ell_{i_s \to i_t+1}^{3/2}$ and 3) $M = O\left(n^{1/5} C_n^{2/5} \vee 1\right)$.*

Going forward, the idea is to bound the dynamic regret within each bin in $\mathcal{P}$ by an $\tilde{O}(1)$ quantity. Then we can add them up across all bins to arrive at the guarantee of Theorem 63 (with $d = 1$). We pause to remark that even-though this high-level idea resembles to that of [65], the underlying details of our analysis to materialize this idea requires highly non-trivial deviations from the path followed by [65].

First, we need some definitions. Consider a bin $[i_s, i_t] \in \mathcal{P}$ with length at-least 2. Let's define a co-variate $\boldsymbol{x}_j := [1, j - i_s + 1]^T$. Let $\boldsymbol{X}^T := [\boldsymbol{x}_{i_s}, \dots, \boldsymbol{x}_{i_t}]$ be the matrix of co-variates and $u_{i_s:i_t} := [u_{i_s}, \dots, u_{i_t}]^T$. Let $\boldsymbol{\beta} = \left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{X}^T u_{i_s:i_t}$ be the least square fit coefficient computed with co-variates $\boldsymbol{x}_j$ and labels u_j. Define a second moment matrix $\boldsymbol{A} = \sum_{j=i_s}^{i_t} \boldsymbol{x}_j\boldsymbol{x}_j^T$. Let $\boldsymbol{\alpha} := \boldsymbol{\beta} - \boldsymbol{A}^{-1}\sum_{j=i_s}^{i_t} \nabla f_j(\boldsymbol{\beta}^T\boldsymbol{x}_j)\boldsymbol{x}_j$. ($\boldsymbol{A}^{-1}$ is guaranteed to exist when length of the bin is at-least 2). We remind the reader that $\nabla f_j(\boldsymbol{\beta}^T\boldsymbol{x}_j)$ is a scalar as we consider uni-variate f_j in this section.

We connect these quantities via a *key regret decomposition* as follows:

$$\sum_{j=i_s}^{i_t} f_j(p_j) - f_j(u_j) = \underbrace{\sum_{j=i_s}^{i_t} f_j(p_j) - f_j(\boldsymbol{\alpha}^T\boldsymbol{x}_j)}_{T_1} + \underbrace{\sum_{j=i_s}^{i_t} f_j(\boldsymbol{\alpha}^T\boldsymbol{x}_j) - f_j(\boldsymbol{\beta}^T\boldsymbol{x}_j)}_{T_2} + \underbrace{\sum_{j=i_s}^{i_t} f_j(\boldsymbol{\beta}^T\boldsymbol{x}_j) - f_j(u_j)}_{T_3} \tag{7.10}$$

It can be shown that $|\boldsymbol{\alpha}^T\boldsymbol{x}_j| \le 20 = O(1)$. Hence the term T_1 can be controlled by an $O(\log n)$ bound due to Strong Adaptivity of FLH-SIONS as described in Section 7.1, Eq.(7.7). The quantity $\boldsymbol{\alpha}$ is obtained via moving in a direction reminiscent to that of Newton method. This is in sharp contrast to the one step gradient descent update used in [65]. More precisely, consider the function $F(\boldsymbol{\beta}) = \sum_{j=i_s}^{i_t} f_j(\boldsymbol{\beta}^T\boldsymbol{x}_j)$. Then $\boldsymbol{\alpha} = \boldsymbol{\beta} - \boldsymbol{A}^{-1}\nabla F(\boldsymbol{\beta})$. By exploiting gradient Lipschitzness of f_j, the correction matrix $\boldsymbol{A}$ can be shown to satisfy the Hessian dominance $\nabla^2 F(\boldsymbol{\beta}) \preccurlyeq \boldsymbol{A}$. This Newton style update is shown to keep the term T_2 to be negative through the following generalized descent lemma:

Lemma 69. *We have that* $T_2 \le -\frac{1}{2}\left\|\nabla F(\boldsymbol{\beta})\right\|_{\boldsymbol{A}^{-1}}^2$.

The negative descent term displayed in the above Lemma is similar to the standard (squared) Newton decrement [84] in the sense that it is also influenced by the local geometry through the norm induced by the inverse correction matrix $\boldsymbol{A}^{-1}$.

We then proceed to show that the negative T_2 can diminish the effect of T_3 by keeping $T_2 + T_3$ to be an $O(1)$ quantity. Thus the dynamic regret within the bin $[i_s, i_t] \in \mathcal{P}$ is controlled to $\tilde{O}(1)$. Adding the bound across all bins in $\mathcal{P}$ from Lemma 68 yields Theorem 63 in one dimension. However, the high level idea is that the smoothness of $\mathcal{TV}^1$ sequence class enables us to keep the regret within each bin to be $\tilde{O}(1)$ despite its larger width thus leading to faster rates when summed across a fewer number of bins.

A major challenge in the analysis is to prove that the term $T_2 + T_3 = O(1)$ without imposing restrictive assumptions such as Self-Concordance or Hessian Lipschitnzess as in the classical analysis of Newton method (see for eg.[84]). In the rest of this section, we outline the arguments leading to this result.

Lemma 70. *We have that $T_2 + T_3 = O(1)$ where T_2 and T_3 are as defined in Eq.*(7.10)

Proof Sketch. Here the main idea is $T_2 + T_3 = O(1)$ even-though $|T_2|$ and $|T_3|$ can be very large individually. Even-though this is the same observation as that in [65], our regret decomposition and the associated proof is more subtle and interesting as it wasn't apriori clear that T_2+T_3 can be possibly bound by $O(1)$ for the current problem. The key novelty is that we bound $T_2 + T_3$ by introducing an auxiliary function that is concave in its arguments which allows us to systematically explore the properties of its maximizers. We proceed to expand upon this proof summary further.

For the sake of explaining ideas, we consider a case where the offline optimal within a bin $[i_s, i_t] \in \mathcal{P}$ doesn't touch the boundary 1 but may touch boundary -1 at multiple time points. (In the full proof, we show that the partition $\mathcal{P}$ can be slightly modified so that in non-trivial cases, the offline optimal can only touch one of the boundaries due to the TV1 constraint within the bins described in Lemma 68.) Then by complementary slackness of Lemma 67 we conclude that $\gamma_j^+ = 0$ for all $j \in [i_s, i_t]$. Our analysis starts by considering a scenario where the offline optimal touches boundary -1 at precisely two points $r, w \in [i_s, i_t]$ with $r < w$ (see Fig.7.5). Again via complementary slackness, only γ_r^- and γ_w^- can be potentially non-zero in this case. Through certain careful bounding

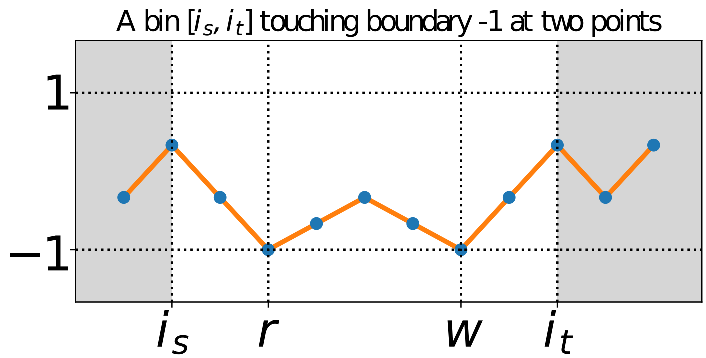


Figure 7.5: *A configuration referred in the proof sketch of Lemma 70. The blue dots represent the offline optimal sequence.*

steps, we show that:

$$T_2 + T_3 \leq -B(\lambda, \gamma_r^-, \gamma_w^-; r, w), \tag{7.11}$$

where B is a function jointly convex in its arguments $\lambda, \gamma_r^-, \gamma_w^-$. We treat r and w to be fixed parameters. The exact form of the function B is present at Eq.(F.142) in Appendix. Then we consider the following convex optimization procedure:

$$\min_{\lambda, \gamma_r^-, \gamma_w^-} \quad B(\lambda, \gamma_r^-, \gamma_w^-; r, w) \tag{7.12a}$$

$$\text{s.t.} \quad \lambda \geq 0 \tag{7.12b}$$

First, we perform a partial minimization wrt γ_r^- and γ_w^- keeping λ fixed. Note that even-though $\gamma_r^- \geq 0$ and $\gamma_w^- \geq 0$ via Lemma 67, we choose to perform an *unconstrained* minimization wrt these variables as doing so can only increase the bound on $T_2 + T_3$.

Let the optimal solutions of the partial minimization procedure be denoted by $\hat{\gamma}_r^-$

and $\hat{\gamma}_w^-$. We find that:

$$B(\lambda, \hat{\gamma}_r^-, \hat{\gamma}_w^-; r, w) = \mathcal{L}(\lambda), \tag{7.13}$$

where $\mathcal{L}(\lambda)$ is a linear function of λ that *doesn't depend* on r or w (Eq.(F.149) in Appendix). The constrained minimum of this linear function is then found to be attained at $\lambda = 0$ and we show that

$$-B(0, \hat{\gamma}_r^-, \hat{\gamma}_w^-; r, w) = O(1) \tag{7.14}$$

This leaves us with an important question on how to handle more than two boundary touches at -1 where many of γ_j^-, $j \in [i_s, i_t]$ can potentially be non-zero. One could perform a similar unconstrained optimization as earlier wrt all γ_j^-. However, deriving the closed form expressions for the optimal $\hat{\gamma}_j^-$ becomes very cumbersome as it involves solving for a complex system of linear equations. In the following, we argue that this general case can be handled via a reduction to the previous setting where only two dual variables γ_r^- and γ_w^- can be potentially non-zero. Specifically we show that the same auxiliary function B as in Eq.(7.11) can be used to obtain

$$T_2 + T_3 \leq -B(\lambda, \tilde{\gamma}_r^-, \tilde{\gamma}_w^-; \tilde{r}, \tilde{w}), \tag{7.15}$$

where $\tilde{r}, \tilde{w}, \tilde{\gamma}_r^-$ and γ_w^- can be computed from the sequence of dual variables $\gamma_{i_s:i_t}^-$. Now we can proceed to optimize similarly as in Eq.(7.12a) with the optimization variables being $\lambda, \tilde{\gamma}_r^-, \tilde{\gamma}_w^-$ and use the same arguments as earlier to bound $T_2 + T_3 = O(1)$. We remark that while doing so, it is an extremely fortunate fact that the partially minimized objective in Eq.(7.13) does not depend on the parameter values r and w. This fact in hindsight is what permitted us to fully eliminate the dependence of all γ_j^- where $j \in [i_s, i_t]$ on the bound via the method of reduction to the case of two non-zero dual variables considered earlier. □

7.2.2 Proof summary for Theorem 63 in multi-dimensions

In rest of this section, we focus on outlining the analysis ideas that facilitated the main result Theorem 63. The high-level idea is to construct a reduction that helps us to re-use much of the machinery developed in Section 7.2.1. We emphasize that this reduction happens only in the analysis, and we *do not* run d uni-variate FLH-SIONS algorithms for handling multi-dimensions. Following Lemma serves a key role in materializing the desired reduction.

Lemma 71. *Let $\boldsymbol{X}_j \in \mathbb{R}^{d \times 2d}$ be as defined as:*

$$\boldsymbol{X}_j^T = \begin{bmatrix} \boldsymbol{x}_j[1:2] & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{x}_j[3:4] & \dots & \mathbf{0} \\ \vdots & \ddots & & \vdots \\ \mathbf{0} & \dots & & \boldsymbol{x}_j[2d-1:2d] \end{bmatrix}, \tag{7.16}$$

where $\mathbf{0} = [0,0]^T$ and $\boldsymbol{x}_j \in \mathbb{R}^{2d}$. The entries $\boldsymbol{x}_j[2k-1:2k] \in \mathbb{R}^2$ for $k \in [d]$. Let $\tilde{f}_j(\boldsymbol{v}) = f_j(\boldsymbol{X}_j \boldsymbol{v})$ for some $\boldsymbol{v} \in \mathbb{R}^{2d}$ and let $\boldsymbol{\Sigma} := \boldsymbol{X}_j^T \boldsymbol{X}_j \in \mathbb{R}^{2d \times 2d}$ which is a block diagonal matrix. We have that

$$\nabla^2 \tilde{f}_j(\boldsymbol{v}) \preccurlyeq \boldsymbol{\Sigma}. \tag{7.17}$$

In multi-dimensions also we form a partition $\mathcal{P}$ of the offline optimal similar to Lemma

68. Then we consider following regret decomposition for any bin $[i_s, i_t] \in \mathcal{P}$.

$$\sum_{j=i_s}^{i_t} f_j(\boldsymbol{p}_j) - f_j(\boldsymbol{u}_j) = \underbrace{\sum_{j=i_s}^{i_t} f_j(\boldsymbol{p}_j) - f_j(\boldsymbol{X}_j \boldsymbol{\alpha}_j)}_{T_1} + \underbrace{\sum_{j=i_s}^{i_t} f_j(\boldsymbol{X}_j \boldsymbol{\alpha}_j) - f_j(\boldsymbol{X}_j \boldsymbol{\beta}_j)}_{T_2} + \underbrace{\sum_{j=i_s}^{i_t} f_j(\boldsymbol{X}_j \boldsymbol{\beta}_j) - f_j(\boldsymbol{u}_j)}_{T_3}, \tag{7.18}$$

where we shall shortly describe how to construct the quantities $\boldsymbol{X}_j \in \mathbb{R}^{d \times 2d}, \boldsymbol{\alpha}_j \in \mathbb{R}^{2d}$ and $\boldsymbol{\beta}_j \in \mathbb{R}^{2d}$. For compactness of notations later, let's define $\boldsymbol{\alpha}_{j,k} = \boldsymbol{\alpha}_j[2k-1:2k] \in \mathbb{R}^2$, $\boldsymbol{\beta}_{j,k} = \boldsymbol{\beta}_j[2k-1:2k] \in \mathbb{R}^2$ and $\boldsymbol{y}_{j,k} = \boldsymbol{x}_j[2k-1:2k] \in \mathbb{R}^2$ for some $\boldsymbol{x}_j \in \mathbb{R}^{2d}$ as in lemma 71. The Hessian dominance in Lemma 71 leads to:

$$\tilde{f}_j(\boldsymbol{\alpha}_j) - \tilde{f}_j(\boldsymbol{\beta}_j) \leq \sum_{k=1}^{d} \langle \nabla f_j(\boldsymbol{X}_j \boldsymbol{\beta}_j)[k] \boldsymbol{y}_{j,k}, \boldsymbol{\alpha}_{j,k} - \boldsymbol{\beta}_{j,k} \rangle + \frac{1}{2} \sum_{k=1}^{d} \|\boldsymbol{\alpha}_{j,k} - \boldsymbol{\beta}_{j,k}\|^2_{\boldsymbol{y}_{j,k}\boldsymbol{y}_{j,k}^T} \tag{7.19}$$

$$:= \sum_{k=1}^{d} t_{2,j,k}. \tag{7.20}$$

Further, due to gradient Lipschitzness of f_j,

$$\tilde{f}_j(\boldsymbol{\beta}_j) - f_j(\boldsymbol{u}_j) \leq \sum_{k=1}^{d} \nabla f_j(\boldsymbol{u}_j)[k] \cdot \left(\boldsymbol{\beta}_{j,k}^T \boldsymbol{y}_{j,k} - \boldsymbol{u}_j[k]\right) + \sum_{k=1}^{d} \frac{1}{2} \|\boldsymbol{\beta}_{j,k}^T \boldsymbol{y}_{j,k} - \boldsymbol{u}_j[k]\|_2^2 \tag{7.21}$$

$$:= \sum_{k=1}^{d} t_{3,j,k} \tag{7.22}$$

Combining Eq.(7.20) and (7.22), we see that $T_2 + T_3$ in any bin $[i_s, i_t]$ can be bounded coordinate-wise:

$$T_2 + T_3 \leq \sum_{k=1}^{d} \sum_{j=i_s}^{i_t} t_{2,j,k} + t_{3,j,k}.$$

This form allows one to bound $\sum_{j=i_s}^{i_t} t_{2,j,k} + t_{3,j,k} = O(1)$ separately for each coordinate by constructing $\boldsymbol{\alpha}_{j,k}, \boldsymbol{\beta}_{j,k}$ and $\boldsymbol{y}_{j,k}$ similar to Section 7.2.1. We then sum across all coordinates to bound $T_2 + T_3 = O(d)$. We remark that the situation is a bit more subtle here because in-order to handle certain combinatorial structures imposed by the KKT conditions, we had to use a sequence of comparators $\boldsymbol{\alpha}_{i_s}, \ldots, \boldsymbol{\alpha}_{i_t}$ (for linear predictors in Eq.(7.18)) that switches at-most $O(d)$ times . Finally by appealing to strong adaptivity of FLH-SIONS, we show that $T_1 = \tilde{O}(d^2)$ for each bin $[i_s, i_t] \in \mathcal{P}$ and Theorem 63 then follows by adding the $\tilde{O}(d^2)$ regret across all $O(n^{1/5} C_n^{2/5} \vee 1)$ bins in $\mathcal{P}$.

7.3 Concluding Discussion

In this chapter, we derived universal dynamic regret rate parametrized by a *novel* second-order path variational of the comparators. Such a path variational naturally captures the piecewise linear structures of the comparators and can be used to flexibly model many practical non-stationarities in the environment. Our results for the exp-concave losses achieved an adaptive universal dynamic regret of $\tilde{O}(d^2 n^{1/5} C_n^{2/5} \vee d^2)$ which matches our minimax lower bound up to a factor that depends on d and $\log n$. This is the first result of such kind in the adversarial setting and the first that works with general exp-concave family of losses. We conjecture that a similar algorithm as in Fig.7.3 based on degree k monomial co-variates $[1, t, \ldots, t^k]$ can lead to optimal dynamic regret for comparators from $\mathcal{TV}^{(k)}$ class.

Part III

Applications

Chapter 8

Online Unsupervised Domain Adaptation under Label Shift

Supervised machine learning algorithms are typically developed assuming independent and identically distributed (iid) data. However, real-world environments evolve dynamically [85, 86, 87, 88]. Absent further assumptions on the nature of the shift, such problems are intractable. One line of research has explored causal structures such as covariate shift [89], label shift [90, 91], and missingness shift [92], for which the optimal target predictor is identified from labeled source and unlabeled target data. Let's denote the feature-label pair of an example by (x, y). Label shift addresses the setting where the label marginal distribution $Q(y)$ may change but the conditional distribution $Q(x|y)$ remains fixed. Most prior work addresses the batch setting for unsupervised adaptation, where a single shift occurs between a source and target population [90, 91, 93, 94, 93, 95]. However, in the real world, shifts are more likely to occur continually and unpredictably, with data arriving in an *online* fashion. A nascent line of research tackles online distribution shift, typically in settings where labeled data is available in real time [65], seeking to minimize the *dynamic regret*.

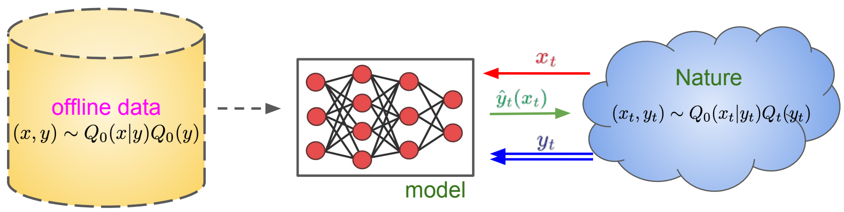


Figure 8.1: *UOLS and SOLS setup.* Dashed (double) arrows are exclusive to UOLS (SOLS) settings. Other objects are common to both setups. Central question: how to adapt the model in real-time to drifting label marginals based on all the available data so far?

Researchers have only begun to explore the role that structures like label shift might play in such online settings. Initial attempts to learn under unsupervised online label shifts were made by [96] and [97], both of which rely on reductions to Online Convex Optimization (OCO) [59, 98]. This line of research aims in updating a classification model based on online data so that the overall regret is controlled. However, [96] only control for *static regret* against a fixed classifier (or model) in hindsight and makes the assumption of the convexity (of losses), which is often violated in practice. In the face of online label shift, where the class marginals can vary across rounds, a more fitting notion is to control the *dynamic regret* against a sequence of models in hindsight. Motivated by this observation, [97] control for the dynamic regret. However, their approach is based on updating model parameters (of the classifier) with online gradient descent and relying on convex losses limits the applicability of their methods (e.g. algorithms in [97] can not be employed with decision tree classifiers).

In this chapter, we study the problem of learning classifiers under Online Label Shift (OLS) in both *supervised* and *unsupervised* settings (Fig.8.1). In both these settings, the distribution shifts are an online process that respects the label shift assumption. Our primary goal is to develop algorithms that side-step convexity assumptions and at the same time *optimally* adapt to the non-stationarity in the label drift. In the Unsupervised Online Label Shift (UOLS) problem, the learner is provided with a pool of labeled offline data sampled iid from the distribution $Q_0(x, y)$ to train an initial model f_0. Afterwards, at every online round t, few *unlabeled* data points sampled from $Q_t(x)$ are presented. The goal is to adapt f_0 to the non-stationary target distributions $Q_t(x, y)$ so that we can accurately classify the unlabelled data. By contrast, in Supervised Online Label Shift (SOLS), our goal is to learn classifiers from *only* the (labeled) samples that arrive in an online fashion from $Q_t(x, y)$ at each time step, while simultaneously adapting to the non-stationarity induced due to changing label proportions. While SOLS is similar to online learning under non-stationarity, UOLS differs from classical online learning as the test label is not seen during online adaptation. Below are the list of contributions of this chapter.

- **Unsupervised adaptation.** For the UOLS problem, we provide a reduction to online regression (see Defn. 72), and develop algorithms for adapting the initial classifier f_0 in a computationally efficient way leading to *minimax optimal* dynamic regret. Our approach achieves the best-of-both worlds of [96, 97] by controlling the dynamic regret while allowing us to use expressive black-box models for classification (Sec. 8.2).

- **Supervised adaptation.** We develop algorithms for SOLS problem that lead to *minimax optimal* dynamic regret without assuming convexity of losses (Sec. 8.3). Our theoretically optimal solution is based on weighted Empirical Risk Minimization (wERM) with weights tracked by online regression. Motivated by our theory, we also propose a *simple continual learning* baseline which achieves empirical performance competitive to the wERM from scratch at each time step across several semi-synthetic SOLS problems while being 15× more efficient in computation cost.

- **Low switching regressors.** We propose a black-box reduction method to convert an optimal online regression algorithm into another algorithm that switches decisions *sparingly* while *maintaining minimax optimality.* This method is relevant for online change point detection. We demonstrate its application in developing SOLS algorithms to train models only when significant distribution drift is detected, while maintaining statistical optimality (App. G.2 and Algorithm 8).

- **Extensive empirical study.** We corroborate our theoretical findings with experiments across numerous simulated and real-world OLS scenarios spanning vision and language datasets (Sec. 8.4). Our proposed algorithms often improve over the best alternatives in terms of both final accuracy and label marginal estimation. This advantage is particularly prominent with limited initial holdout data (in the UOLS problem) highlighting the *sample efficiency* of our approach.

Notes on technical novelties. Even-though online regression is a well studied technique, to the best of our knowledge, it is not used before to address the problem of online label shift. It is precisely the usage of regression which lead to tractable adaptation algorithms while side-stepping convexity assumptions thereby allowing us to use very flexible models for classification. This is in stark contrast to OCO based reductions in [96] and [97]. We propose new theoretical frameworks and identify the right set of assumptions for materializing the reduction to online regression. It was not evident initially that this link would lead to *minimax optimal* dynamic regret rates as well as *consistent* empirical improvement over prior works. Proof of the lower bounds requires adapting the ideas from non-stationary stochastic optimization [15] in a non-trivial manner. Further, none of the proposed methods require the prior knowledge of the extent of distribution drift. The material of this chapter closely follows [99].

8.1 Problem Setup

Let $\mathcal{X} \subseteq \mathbb{R}^d$ be the input space and $\mathcal{Y} = [K] := \{1, 2, \ldots, K\}$ be the output space. Let Q be a distribution over $\mathcal{X} \times \mathcal{Y}$ and let $q(\cdot)$ denotes the corresponding label marginal. Δ_K is the K-dimensional simplex. For a vector $v \in \mathbb{R}^K$, $v[i]$ is its i^{th} coordinate. We assume that we have a hypothesis class $\mathcal{H}$. For a function $f \in \mathcal{H} : \mathcal{X} \to \Delta_K$, we also use $f(i|x)$ to indicate $f(x)[i]$. With $\ell(f(x), y)$, we denote the loss of making a prediction with the classifier f on (x, y). L denotes the expected loss, i.e., $L = \mathbb{E}_{(x,y)\sim Q}\left[\ell(f(x), y)\right]$. $\tilde{O}(\cdot)$ hides dependencies in absolute constants and poly-logarithmic factors of horizon and failure probabilities.

In this work, we study online learning under distribution shift, where the distribution $Q_t(x, y)$ may continuously change with time. Throughout the chapter, we focus on the *label shift* assumption where the distribution over label proportions $q_t(y)$ can change arbitrarily but the distribution of the covariate conditioned on a label value (i.e., $Q_t(x|y)$) is assumed to be invariant across all time steps. We refer to this setting as Online Label Shift (OLS). Here, we consider settings of unsupervised and supervised OLS settings captured in Frameworks 1 and 3 respectively. In both settings, at round t a sample (x_t, y_t) is drawn from a distribution with density $Q_t(x_t, y_t)$. In the UOLS setting, the label is not revealed to the learner. However, we assume access to offline labeled data sampled iid from Q_0 which we use to train an initial classifier f_0. The goal is to adapt the initial classifier f_0 to drifting label distributions. In contrast, for the SOLS setting, the label is revealed to the learner after making a prediction and the goal is to learn a classifier $f_t \in \mathcal{H}$ for each time step.

Next, we formally define the concept of online regression which will be central to our discussions. Simply put, an online regression algorithm tracks a ground truth sequence from noisy observations.

Definition 72 (online regression)**.** Fix any $T > 0$. The following interaction scheme is defined to be the online regression protocol.

- At round $t \in [T]$, an algorithm predicts $\hat{\theta}_t \in \mathbb{R}^K$.
- A noisy version of ground truth $z_t = \theta_t + \epsilon_t$ is revealed where $\theta_t, \epsilon_t \in \mathbb{R}^K$, and $\|\epsilon_t\|_2, \|\theta_t\|_2 \leq B$. Further the noise ϵ_t are independent across time with $E[\epsilon_t] = 0$ and $\mathrm{Var}(\epsilon_t[i]) \leq \sigma^2 \; \forall i \in [K]$.

An online regression algorithm aims to control $\sum_{t=1}^{T} \|\hat{\theta}_t - \theta_t\|_2^2$. Moreover, the regression algorithm is defined to be adaptively minimax optimal if with probability at least $1 - \delta$, $\sum_{t=1}^{n} \|\hat{\theta}_t - \theta_t\|_2^2 = \tilde{O}(T^{1/3} V_T^{2/3})$ without knowing V_T ahead of time. Here $V_T := \sum_{t=2}^{T} \|\theta_t - \theta_{t-1}\|_1$ is termed as the Total Variation (TV) of the sequence $\theta_{1:T}$.

8.2 Unsupervised Online Label Shift

In this section, we develop a framework for handling the UOLS problem. We summarize the setup in Framework 1. Since in practice, we may need to work with classifiers such as deep neural networks or decision trees, we do not impose convexity assumptions on the (population) loss of the classifier as a function of the model parameters. Despite the absence of such simplifying assumptions, we provide performance guarantees for our label shift adaption techniques so that they are certified to be fail-safe.

Under the label shift assumption, we have $Q_t(y|x)$ as a re-weighted version of $Q_0(y|x)$:

$$Q_t(y|x) = \frac{Q_t(y)}{Q_t(x)} Q_t(x|y) = \frac{Q_t(y)}{Q_t(x)} Q_0(x|y) = \frac{Q_t(y) Q_0(x)}{Q_t(x) Q_0(y)} Q_0(y|x) \propto \frac{Q_t(y)}{Q_0(y)} Q_0(y|x), \tag{8.1}$$

where the second equality is due to the label shift assumption. Hence, a reasonable strategy is to re-weight the initial classifier f_0 with label proportions (estimate) at the

Framework 1 Unsupervised Online Label Shift (UOLS) protocol

Input: Initial classifier $f_0 : \mathcal{X} \to \Delta_K$ trained on offline labeled dataset $\{(x_i, y_i)\}_{i=1}^N$ sampled iid from Q_0;

$f_1 = f_0$
for each round $t \in [T]$ **do**
  Nature samples $x_t \in \mathcal{X}$ and $y_t \in \mathcal{Y}$, with $(x_t, y_t) \sim Q_t$; Only x_t is revealed to the learner.
  Learner predicts a label $i \sim f_t(x_t) \in \Delta_K$.
  $f_{t+1} = \mathcal{A}(f_0, x_{1:t})$, where $\mathcal{A}$ is strategy to adapt the classifier based on past data.
end for

Algorithm 2 `RegressAndReweight` to handle UOLS

Input: i) Online regression oracle ALG; ii) Initial classifier f_0; iii) The confusion matrix C; iv) The label marginal $q_0 \in \mathcal{D}$ of the training distribution;

At round t, get the classifier covariate x_t.
Let $\widehat{q}_t = \Pi_{\mathcal{D}}\left(\text{ALG}(s_{1:t-1})\right)$, where $\Pi_{\mathcal{D}}(x) = \text{argmin}_{y \in \mathcal{D}} \|y - x\|_2$.
Sample a label i with probability $\propto \frac{\widehat{q}_t(i)}{q_0(i)} f_0(i|x_t)$.
Let $s_t = C^{-1} f_0(x_t)$.
Update the online regression oracle with the estimate s_t.

current step, since we only have to correct the label distribution shift. This re-weighting technique is widely used for offline label shift correction [91, 94, 93] and for learning under label imbalance [100, 101, 102].

Our starting point in developing a framework is inspired by [96, 97] . For self-containedness, we briefly recap their arguments next. We refer interested readers to their chapters for more details. [96] considers a hypothesis class of re-weighted initial classifier f_0. The loss of a hypothesis is parameterised by the re-weighting vector. They use tools from OCO to optimise the loss and converge to a best fixed classifier. However as noted in [96], the losses are not convex with respect to the re-weight vector in practice. Hence usage of OCO techniques is not fully satisfactory in their problem formulation.

In a complementary direction, [97] abandons the idea of re-weighting. Instead, they update the parameters of a model at each round using online gradient descent and a loss function whose expected value is assumed to be convex with respect to model parameters. They provide dynamic regret guarantees against a sequence of changing model parameters in hindsight, and connects it to the variation of the true label marginals. More precisely, they provide algorithms with $\sum_{t=1}^{T} L_t(w_t) - L_t(w_t^*)$ to be well controlled where w_t^* is the best model parameter to be used at round t and L_t is a (population level) loss function. However, there are some scopes for improvement in this direction as well. For example, the convexity assumption can be easily violated when working with interpretable models based on decision trees, or if we want to retrain few final layers of a deep classifier based on new data. Further as noted in the experiments (Sec. 8.4), their methods based on retraining the classifier require more data than re-weighting based methods. Our experiments also indicate that re-weighting can be computationally cheaper than re-training without sacrificing the classifier accuracy.

Thus, on the one hand, the work of [96] allows us to use the power of expressive initial classifiers while only controlling the static regret against a fixed hypothesis. On the other hand, the work of [97] allows controlling the dynamic regret while limiting the flexibility of deployed models. We next provide our framework for handling label shifts that achieves the best of both worlds by controlling the dynamic regret while allowing the use of expressive *blackbox* models.

In summary, we estimate the sequence of online label marginals and leverage the idea of re-weighting an initial classifier as in [96]. In particular, given an estimate $\widehat{q}_t(y)$ of the true label marginal at round t, we compute the output of the re-weighted classifier f_t as $\frac{\widehat{q}_t(y)}{q_0(y)} f_0(y|x) / Z$ where $Z = \sum_y \frac{\widehat{q}_t(y)}{q_0(y)} f_0(y|x)$. However, to get around the issue of non-convexity, we separate out the process of estimating the re-weighting vectors via a reduction to online regression which is a well-defined and convex problem with computationally efficient off-the-shelf algorithms readily available. Second, and more importantly, [96] competes with the best *fixed* re-weighted hypothesis. However, in the problem setting of label shift, the true label marginals are in fact changing. Hence, we control the *dynamic regret* against a sequence of re-weighted hypotheses in hindsight. All proofs for the next sub-section are deferred to App. G.1.

8.2.1 Proposed algorithm and performance guarantees

We start by presenting our assumptions. This is followed by the main algorithm for UOLS and its performance guarantees. Similar to the treatment in [97], we assume the following.

Assumption 1. Assume access to the true label marginals $q_0 \in \Delta_K$ of the offline training data and the true confusion matrix $C \in \mathbb{R}^{K \times K}$ with $C_{ij} = E_{x \sim Q_0(\cdot|y=j)}, f_0(i|x)$. Further the minimum singular value $\sigma_{min}(C) = \Omega(1)$ is bounded away from zero.

As noted in prior work [91, 95], the invertibility of the confusion matrix holds whenever the classifier f_0 has good accuracy and the true label marginal q_0 assigns a non-zero probability to each label. Though we assume perfect knowledge of the label marginals of the training data and the associated confusion matrix, this restriction can be easily relaxed to their empirical counterparts computable from the training data. The finite sample error between the empirical and population quantities can be bounded by $O(1/\sqrt{N})$ where N is the number of initial training data samples. To this end, we operate in the regime where the time horizon obeys $T = O(\sqrt{N})$. However, similar to [97], we make this assumption mainly to simplify presentation without trivializing any aspect of the OLS problem.

Next, we present our assumptions on the loss function. Let $p \in \Delta_K$. Consider a classifier that predicts a label $\widehat{y}(x)$, by sampling $\widehat{y}(x)$ according to the distribution that assigns a weight $\frac{p(i)}{q_0(i)} f_0(i|x)$ to the label i. Define $L_t(p)$ to be any non-negative loss that ascertains the quality of the marginal p. For example, $L_t(p) = E[\ell(\hat{y}(x), y)]$ where the expectation is taken wrt the randomness in the draw $(x, y) \sim Q_t$ and in sampling $\hat{y}(x)$. Here ℓ is any classification loss (e.g. 0-1, cross-entropy).

Assumption 2 (Lipschitzness of loss functions)**.** Let $\mathcal{D}$ be a compact and convex domain. Assume that $L_t(p)$ is G Lipschitz with $p \in \mathcal{D} \subseteq \Delta_K$, i.e, $L_t(p_1) - L_t(p_2) \leq G\|p_1 - p_2\|_2$ for any $p_1, p_2 \in \mathcal{D}$. The constant G need not be known ahead of time.

We show in Lemmas 229 and 230 that the above assumption is satisfied under mild regularity conditions. Furthermore, the prior works such as [96] and [97] also require that losses are Lipschitz with a *known* Lipschitz constant apriori to set the step sizes for their OGD based methods.

The main goal here is to design appropriate re-weighting estimates such that the *dynamic regret*:

$$R_{\text{dynamic}}(T) = \sum_{t=1}^{T} L_t(\hat{q}_t) - L_t(q_t) \leq \sum_{t=1}^{T} G\|\hat{q}_t - q_t\|_2 \tag{8.2}$$

is controlled where $\hat{q}_t \in \Delta_K$ is the estimate of the true label marginal q_t. Thus we have reduced the problem of handling OLS to the problem of online estimation of the true label marginals.

Under label shift, we can get an unbiased estimate of the true marginals at any round via the techniques in [91, 94, 93]. More precisely, $s_t = C^{-1} f_0(x_t)$ has the property that $E[s_t] = q_t$ (see Lemma 233). Further, the variance of the estimate s_t is bounded by $1/\sigma^2_{min}(C)$. Unfortunately, these unbiased estimates can not be directly used to track the moving marginals q_t. This is because the total squared error $\sum_{t=1}^{T} E[\|s_t - q_t\|_2^2]$ grows linearly in T as the sum of the variance of the point-wise estimates accumulates unfavorably over time.

To get around these issues, one can use online regression algorithms such as FLH [23] with online averaging base learners or the Aligator algorithm [71]. These algorithms use ensemble methods to (roughly) output running averages of s_t where the variation in the *true* label marginals is small enough. The averaging within intervals where the true marginals change slowly helps to reduce the overall variance while injecting only a small bias. We use such *online regression oracles* to track the moving marginals and re-calibrate the initial classifier. Overall, Algorithm 2 summarizes our method which has the following performance guarantee.

Theorem 73. *Suppose we run Algorithm 2 with the online regression oracle ALG as FLH-FTL (App. G.4) or Aligator [71]. Then under Assumptions 1 and 2, we have*

$$E[R_{dynamic}(T)] = \tilde{O}\left(\frac{K^{1/6}T^{2/3}V_T^{1/3}}{\sigma_{min}^{2/3}(C)} + \frac{\sqrt{KT}}{\sigma_{min}(C)}\right), \tag{8.3}$$

where $V_T := \sum_{t=2}^{T} \|q_t - q_{t-1}\|_1$ *and the expectation is taken with respect to randomness in the revealed co-variates. Further, this result is attained without prior knowledge of* V_T*.*

Remark 74. We emphasize that any valid online regression oracle ALG can be plugged into Algorithm 2. This implies that one can even use transformer-based time series models to track the moving marginals q_t. Further, we have the flexibility of choosing the initial classifier to be any *black-box* model that outputs a distribution over the labels.

Remark 75. Unlike prior works such as [96, 97], we do not need a pre-specified bound on the gradient of the losses. Consequently Eq.(8.2) holds for the smallest value of the Lipschitzness coefficient G, leading to tight regret bounds. Further, the projection step in Line 2 of Algorithm 2 is done only to safeguard our theory against pathological scenarios with unbounded Lipschitz constant for losses. In our experiments, we do not perform such projections.

We next show that the performance guarantee in Theorem 73 is optimal (modulo factors of $\log T$) in a minimax sense.

Theorem 76. *Let* $V_T \leq 64T$*. There exists a loss function, a domain* $\mathcal{D}$ *(in Assumption 2), and a choice of adversarial strategy for generating the data such that for any algorithm,*

Framework 3 Supervised Online Label Shift (SOLS) protocol

Input: A hypothesis class $\mathcal{H}$.

1: **for** each round $t \in [T]$ **do**
2: Nature samples N iid data points $x_{t,1:N} \in \mathcal{X}$ and $y_{t,1:N} \in \mathcal{Y}$, with each $(x_{t,i}, y_{t,i}) \sim Q_t$; $x_{t,1:N}$ is revealed to the learner.
3: For each $i \in [N]$, learner predicts a label $f_t(x_{t,i})$.
4: The label $y_{t,i} \in \mathcal{Y}$ for each $i \in [N]$ is revealed.
5: $f_{t+1} = \mathcal{A}(f_t, \{x_{1:t,1:N}, y_{1:t,1:N}\})$ where algorithm $\mathcal{A}$ updates the classifier with past data.
6: **end for**

Algorithm 4 `TrainByWeights` to handle SOLS

Input: Online regression oracle ALG, hypothesis class $\mathcal{H}$

1: At round $t \in [T]$, get estimated label marginal $\hat{q}_t$ from ALG$(s_{1:t-1})$.
2: Update the hypothesis with weighted ERM:

$$f_t = \operatorname*{argmin}_{f \in \mathcal{H}} \sum_{i=1}^{t-1} \sum_{j=1}^{N} \frac{\hat{q}_t(y_{i,j})}{\hat{q}_i(y_{i,j})} \ell(f(x_{i,j}), y_{i,j}) \tag{8.4}$$

3: Get co-variates $x_{t,1:N}$ and make predictions with f_t
4: Get labels $y_{t,1:N}$
5: Compute $s_t[i] = \frac{1}{N} \sum_{j=1}^{N} \mathbb{I}\{y_{t,j} = i\}$ for all $i \in [K]$.
6: Update ALG with the empirical label marginals s_t.

we have $\sum_{t=1}^{T} E([L_t(\hat{q}_t)] - L_t(q_t)) = \Omega\left(\max\{T^{2/3} V_T^{1/3}, \sqrt{T}\}\right)$ *, where* $\hat{q}_t \in \mathcal{D}$ *is the weight estimated by the algorithm and* $q_t \in \mathcal{D}$ *is the label marginal at round* t *chosen by the adversary. Here the expectation is taken with respect to the randomness in the algorithm and the adversary.*

8.3 Supervised Online Label Shift

In this section, we focus on the SOLS problem where the labels are revealed to the learner after it makes decisions. Framework 3 summarizes our setup. Let $f_t^* := \operatorname{argmin}_{f \in \mathcal{H}} L_t(f)$ be the population minimiser. We aim to control the *dynamic regret* against the best sequence of hypotheses in hindsight:

$$R^{\mathcal{H}}_{\text{dynamic}}(T) =: \sum_{t=1}^{T} L_t(f_t) - L_t(f_t^*)\,. \tag{8.5}$$

If the SOLS problem is convex, it reduces to OCO [59, 98] and existing works provide $\tilde{O}(T^{2/3}V_T^{1/3})$ dynamic regret guarantees [27]. However, in practice, since loss functions are seldom convex with respect to model parameters in modern machine learning, the performance bounds of OCO algorithms cease to hold true. In our work, we extend the generalization guarantees of ERM from statistical learning theory [103] to the SOLS problem. All proofs of next sub-section are deferred to App. G.3.

8.3.1 Proposed algorithms and performance guarantees

We start by providing a simple initial algorithm whose computational complexity and flexibility will be improved later. Note that due to the label shift assumption, for any $j, t \in [T]$, we have $E_{(x,y)\sim Q_t}[\ell(f(x), y)] = E_{(x,y)\sim Q_j}\left[\frac{q_t(y)}{q_j(y)}\ell(f(x), y)\right]$. Here we assume that the true label marginals $q_t(y) > 0$ for all $t \in [T]$ and all $y \in [K]$. Based on this, we propose a simple weighted ERM approach (Algorithm 4) where we use an online regression oracle to estimate the label marginals from the (noisy) empirical label marginals computed with observed labeled data. With weighted ERM and plug-in estimates of importance weights, we can obtain our classifier f_t. One can expect that by adequately choosing the online regression oracle ALG, the risk of the hypothesis f_t computed will be close to that of f_t^*. Here the degree of closeness will also depend on the number of data points seen thus far. Consequently, Algorithm 4 controls the dynamic regret (Eq.(8.5)) in a graceful manner. We have the following performance guarantee:

Theorem 77. *Suppose the true label marginal satisfies* $\min_{t,k} q_t(k) \geq \mu > 0$. *Choose the online regression oracle in Algorithm 4 as FLH-FTL (App. G.4) or Aligator from [71] with its predictions clipped such that* $\hat{q}_t[k] \geq \mu$. *Then with probability at least* $1 - \delta$, *Algorithm 4 produces hypotheses with* $R^{\mathcal{H}}_{dynamic} = \tilde{O}\left(T^{2/3}V_T^{1/3} + \sqrt{T\log(|\mathcal{H}|/\delta)}\right)$, *where* $V_T = \sum_{t=2}^{T}\|q_t - q_{t-1}\|_1$. *Further, this result is attained without any prior knowledge of* *the variation budget* V_T*.*

The above rate contains the sum of two terms. The second term is the familiar rate seen in the supervised statistical learning theory literature under iid data [103]. The first term reflects the price we pay for adapting to distributional drift in the label marginals. While we prove this result for finite hypothesis sets, the extension to infinite sets is direct by standard covering net arguments [104].

Remark 78. Theorem 77 requires that the estimates of the label marginals to be clipped from below by μ. This is done only to facilitate theoretical guarantees by enforcing that the importance weights used in Eq.(8.4) do not become unbounded. However, note that only the labels we actually observe enters the objective in Eq.(8.4). In particular, if a label has very low probability of getting sampled at a round, then it is unlikely that it enters the objective. Due to this reason, in our experiments, we haven't used the clipping operation (see Section 8.4 and Appendix G.4 for more details).

The proof of the theorem uses concentration arguments to establish that the risk of the hypothesis f_t is close to the risk of the optimal f_t^*. However, unlike the standard offline supervised setting with iid data, for any fixed hypothesis, the terms in the summation of Eq.(8.4) are correlated through the estimates of the online regression oracle. We handle it by introducing uncorrelated surrogate random variables and bounding the associated discrepancy. Next, we show (near) minimax optimality of the guarantee in Theorem 77.

Theorem 79. *Let* $V_T \leq T/8$*. There exists a choice of hypothesis class, loss function, and adversarial strategy of generating the data such that* $R^{\mathcal{H}}_{dynamic} = \Omega\left(T^{2/3}V_T^{1/3} + \sqrt{T\log(|\mathcal{H}|)}\right)$ *, where the expectation is taken with respect to randomness in the algorithm and adversary.*

Remark 80. Though the rates in Theorems 76 and 79 are similar, we note that the corresponding regret definitions are different. Hence the minimax rates are not directly comparable between the supervised and unsupervised settings.

Even-though Algorithm 4 has attractive performance guarantees, it requires retraining with weighted ERM at every round. This can be computationally expensive. To alleviate this issue, we design a new online change point detection algorithm (Algorithm 7 in App. G.2) that can adaptively discover time intervals where the label marginals change slow enough. We show that the new online change point detection algorithm can be used to significantly reduce the number of retraining steps without sacrificing statistical efficiency (up to constants). We defer the exact details to App. G.2. We remark that our change point detection algorithm is applicable to general online regression problems and hence can be of independent interest to online learning community.

Remark 81. Algorithm 7 helps to reduce the run-time complexity. However, both Algorithms 4 and 7 have the drawback of storing all data points accumulated over the online rounds. This is reminiscent to FTL / FTRL type algorithms from online learning. We leave the task of deriving theoretical guarantees with reduced storage complexity under non-convex losses as an important future direction.

8.4 Experiments

Code is publicly available at `https://github.com/Anon-djiwh/OnlineLabelShift`.

8.4.1 UOLS Setup and Results

Setup Following the dataset setup of [97], we conducted experiments on synthetic and common benchmark data such as MNIST [105], CIFAR-10 [106], Fashion [107], EuroSAT [108], Arxiv [109], and SHL [110, 111]. For each dataset, the original data is split into labeled data available during offline training and validation, and the unlabeled data that we observe during online learning. We experiment with varying sizes of holdout offline data which is used to obtain the confusion matrix and update the model parameters

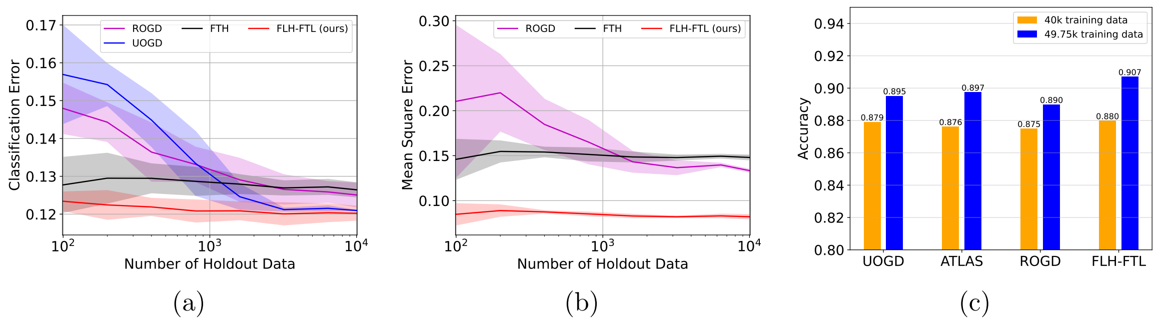


Figure 8.2: *Results on the UOLS problem.* **(a) and (b):** Ablation on CIFAR10 with monotone shift over sizes of holdout data used to update model parameters and compute confusion matrix, with amount of training data held fixed. FLH-FTL (ours) outperforms all other alternatives throughout in classification error and mean square error in label marginal estimation. Unlike the alternatives, the performance of FLH-FTL (ours) is unaffected by the decrease in amount of holdout data. **(c):** CIFAR10 results with monotone shift using varying amount of training data, with the remaining labeled data used as holdout (total number of samples fixed to 50k). The performance of FLH-FTL is minimally impacted by the reduction in the quantity of holdout data, thus yielding the greatest advantage from utilizing a larger volume of training data.

to adapt to OLS to probe the sample efficiency of all the methods. In contrast to previous works [97, 96], we have chosen to use a smaller amount of holdout offline data for our main experiments. We made this decision because the standard practice for deployment involves training and validating models on training and holdout splits, respectively (e.g., with k-fold cross-validation). Then, the final model is deployed by training on all available data (i.e., the union of train and holdout) with the identified hyperparameters. However, to employ UOLS techniques in practice, practitioners must hold out data that was not seen during training to update the model during online adaptation. Therefore, methods that are efficient with respect to the amount of offline holdout data required might be preferable.

For all datasets except SHL, we simulate online label shifts with four types of shifts studied in [97]: monotone shift, square shift, sinusoidal shift, and Bernoulli shift. For SHL locomotion, we use the real-world shift occurring over time. For architectures, we use

Methods	Synthetic		MNIST		CIFAR		EuroSAT		Fashion		ArXiv	
	Ber	Sin	Ber	Sin	Ber	Sin	Ber	Sin	Ber	Sin	Ber	Sin
Base	8.6±0.2	8.2±0.3	4.9±0.4	3.9±0.0	16±0	16±0	13±0	13±0	15±0	15±0	23±1	19±0
OFC	6.4±0.6	5.5±0.2	4.4±0.5	3.2±0.3	12±1	11±0	11±1	10±1	7.9±0.1	7.1±0.1	20±2	15±0
Oracle	3.7±0.8	3.9±0.2	2.5±0.5	1.5±0.1	5.4±0.5	5.8±0.1	3.9±0.3	4.1±0.1	3.7±0.2	3.6±0.1	7.7±1.0	5.1±0.1
FTH	6.5±0.6	5.7±0.3	4.5±0.6	**3.3±0.2**	11±0	**11±0**	10±0	**9.6±0.0**	8.5±0.3	**6.9±0.4**	20±1	**14±0**
FTFWH	6.6±0.5	5.7±0.3	4.5±0.6	**3.3±0.2**	11±1	**11±0**	9.8±0.4	**9.6±0.1**	8.2±0.6	**6.9±0.4**	20±1	**14±0**
ROGD	7.9±0.3	7.2±0.6	6.2±2.8	4.4±1.5	16±3	13±0	14±1	13±1	10±1	8.2±0.7	23±2	17±1
UOGD	8.1±0.6	7.5±0.6	5.4±0.6	4.0±0.0	14±0	14±1	10±1	9.8±0.7	11±2	11±2	21±1	17±1
ATLAS	8.0±1.0	7.5±0.6	5.2±0.6	3.7±0.2	13±0	13±1	10±1	9.9±0.7	12±2	12±2	21±1	16±0
FLH-FTL (ours)	**5.4±0.7**	**5.4±0.4**	**4.4±0.7**	**3.3±0.2**	**10±0**	**11±0**	**9.2±0.4**	**9.6±0.1**	**7.7±0.4**	7.0±0.0	**19±1**	**14±0**

	Synthetic		MNIST		CIFAR		EuroSAT		Fashion		ArXiv	
	Ber	Sin	Ber	Sin	Ber	Sin	Ber	Sin	Ber	Sin	Ber	Sin
FTH	0.19±0.01	0.10±0.00	0.27±0.00	0.14±0.00	0.27±0.01	0.14±0.00	0.27±0.00	0.14±0.00	0.29±0.01	**0.14±0.01**	0.29±0.01	**0.15±0.00**
FTFWH	0.19±0.02	0.09±0.00	0.26±0.02	0.13±0.00	0.25±0.02	**0.13±0.00**	0.25±0.01	**0.13±0.00**	0.25±0.04	**0.14±0.01**	0.27±0.02	**0.15±0.00**
ROGD	0.29±0.03	0.24±0.01	0.41±0.08	0.37±0.06	0.39±0.04	0.30±0.05	0.43±0.04	0.35±0.03	0.37±0.02	0.30±0.01	0.34±0.03	0.28±0.01
FLH-FTL (ours)	**0.10±0.01**	**0.08±0.00**	**0.15±0.01**	**0.12±0.00**	**0.17±0.01**	**0.13±0.00**	**0.16±0.01**	**0.13±0.00**	**0.18±0.02**	**0.14±0.01**	**0.23±0.01**	**0.15±0.00**

Table 8.1: *Results for UOLS problems under sinusoidal (Sin) and Bernoulli (Ber) shifts.* **Top:** Classification Error. **Bottom:** Mean-squared error in estimating label marginal. For both, lower is better. Across all datasets, we observe that FLH-FTL (ours) often improves over best alternatives.

an MLP for Fashion, SHL and MNIST, Resnets [112] for EuroSAT, CINIC, and CIFAR, and DistilBERT [113, 114] based models for arXiv. For alternate approaches, along with a base classifier (which does no adaptation) and oracle classifier (which reweight using the true label marginals), we make comparisons with adaptation algorithms proposed in prior works [96, 97]. In particular, we compare with ROGD, FTH, FTFWH from [96] and UOGD, ATLAS from [97]. For brevity, we refer to our method as FLH-FTL (though strictly speaking, our methods are based on FLH from [23] with online averages as base learners). We run all the online label shift experiments with the time horizon $T = 1000$ and at each step 10 samples are revealed. We repeat all experiments with 3 seeds to obtain means and standard deviations of the results. For other methods that perform re-weighting correction on softmax predictions, we use the labeled holdout data to calibrate the model with temperature scaling, which tunes one temperature parameter [115]. We provide exact details about the datasets, label shift simulations, models, and prior methods in App. G.4.

Results Overall, across all datasets, we observe that our method FLH-FTL performs better than alternative approaches in terms of both classification error and mean squared error for estimating the label marginal. Note that methods that directly update the model parameters (i.e., UOGD, ATLAS) do not provide any estimate of the label marginal (Table 8.1). UOGD and ATLAS also require offline holdout labeled data (i.e., from time step 0) to make online updates to the model parameters. For this purpose, we use the same labeled data that we use to compute the confusion matrix.

As we increase the holdout offline labeled dataset size for updating the model parameters (and to compute the confusion matrix), we observe that classification error and MSE with FLH-FTL stay (relatively) constant whereas the classification errors of other alternatives improve (Fig. 8.2). This highlights that FLH-FTL can be much more sample efficient with respect to the size of the hold-out offline labeled data. Motivated by this observation, we perform an additional experiment in which we increase the offline training data and observe that we can overall improve the classification accuracy significantly with FLH-FTL (Fig. 8.2). We present results on SHL dataset with similar findings on semi-synthetic datasets in App. G.5.5. Finally, we also experiment with a random forest model on the MNIST dataset. Note methods that update model parameters (e.g., UOGD and ATLAS) with OGD are not applicable here. Here, we also observe that we improve over existing applicable alternatives (Table 8.2).

	Base	Oracle	ROGD	FTH	FTFWH	FLH-FTL (ours)
Cl Err	18±1	6.3±1.3	19±3	14±2	14±2	**13±2**
MSE	NA	0.0±0.0	0.3±0.0	0.3±0.0	0.3±0.0	**0.2±0.0**

Table 8.2: *Results with a Random Forest classifier on MNIST dataset.* Note that methods that update model parameters are not applicable here. FLH-FTL outperforms existing alternatives for both accuracy and label marginal estimation.

	CT (base)	CT-RS (ours) w FTH	CT-RS (ours) w FLH-FTL	w-ERM (oracle)
Cl Err	20.0±0.5	18.38±0.4	**17.12±0.8**	16.32±0.7
MSE	NA	0.18±0.01	**0.12±0.01**	NA

Table 8.3: *Results on SOLS setup* on CIFAR10 SOLS with Bernoulli shift. CT with RS improves over the base model (CT) and achieves competitive performance with respect to weighted ERM oracle. MNIST results are similar (see App. G.4).

8.4.2 SOLS setup and results

Setup For the supervised problem, we experiment with MNIST and CIFAR datasets. We simulate a time horizon of $T = 200$. For each dataset, at each step, we observe 50 samples with Bernoulli shift. Motivated by our theoretical results with weighted ERM, we propose a simple baseline which continually trains the model at every step instead of starting ERM from scratch every time. We maintain a pool of all the labeled data received till that time step, and at every step, we randomly sample a batch with uniform label marginal to update the model. Finally, we re-weight the updated softmax outputs with estimated label marginal. We call this method Continual Training via Re-Sampling (CT-RS). Its relation as a close variant of weighted ERM is elaborated in App. G.4.1. To estimate the label marginal, we try FTH and ours FLH-FTL.

Results On both datasets, we observe that empirical performance with CT-RS improves over the naive continual training baseline. Additionally, CT-RS results are competitive with weighted ERM while being 5–15× faster in terms of computation cost (we include the exact computational cost in App. G.4.1). Moreover, as in UOLS setup, we observe that FLH-FTL improves over FTH for both target label marginal estimation and classification.

8.5 Concluding Discussion

In this work, we focused on unsupervised and supervised online label shift settings. For both settings, we developed algorithms with minimax optimal dynamic regret. Experimental results on both real and semi-synthetic datasets substantiate that our methods improve over prior works both in terms of accuracy and target label marginal estimation.

In future work, we aim to expand our experiments to more real-world label shift datasets. This chapter also motivates future work in exploiting other causal structures (e.g. covariate shift) for online distribution shift problems.

Chapter 9

Non-Stationary Contextual Pricing with Safety Constraints

In this Chapter we push the development of Chapter 6. We design algorithms for proper non-stationary OCO when the exp-concave loss belong to the generalized linear family. Examples of such losses include linear and logistic regression. The main results are featured via its applications to dynamic pricing.

9.1 Introduction

Feature-based dynamic pricing, or *contextual* pricing, is a problem where the seller sets prices for different products based on their features and aims to maximize revenue. In general, a customer will make her decision based on a comparison between the price and her own valuation of the product. Formally, many existing works [116, 117, 118, 119] adopt the following linear-feature valuation model:

Contextual pricing. For $t = 1, 2, ..., T$:

1. A context $x_t \in \mathbb{R}^d$ is revealed that describes a sales session (product, customer and context).
2. The customer valuates the product as $y_t = x_t^\top \theta_t^* + N_t$ using x_t.
3. The seller proposes a price $v_t > 0$ concurrently (according to x_t and historical sales records).
4. The transaction is successful if $v_t \leq y_t$, i.e., the seller gets a reward (payment) of $r_t = v_t \cdot \mathbb{I}(v_t \leq y_t)$.

Here T is the unknown time horizon, x_t's are adversarial features (which can be stochastic or non-stochastic series), θ_t^*'s are *hidden parameters* mapping features to valuations linearly, and N_t's are i.i.d. noises drawn from a known distribution $\mathbb{D}$. Denote $\mathbb{I}_t := \mathbb{I}(v_t \leq y_t)$ as the *Boolean-censored* feedback that equals 1 if $v_t \leq y_t$ and 0 otherwise, and we only observe $\mathbb{I}_t$ instead of the realized y_t at each round. Our goal is to maximize the cumulative expected reward, and the *regret* is defined as the difference of expected rewards between v_t and the best price at each round.

Time-variant Behavior and Dynamic Regret. Comparing with existing linear contextual pricing problem settings [116, 117, 118] where the linear valuation parameter θ_t^* is fixed as the same θ^* over all t, in this work we allow moderate changing of customers' valuations: i.e. θ_t^*'s can vary over time, and the *total variation* $\sum_{t=1}^{T-1} \|\theta_t^* - \theta_{t+1}^*\|_1$ is upper bounded by some C_T (which could be unknown to the seller). Here we adopt the L_1-norm bound because it is a reasonable metric for capturing the non-stationarity of the valuation mechanism: For instance, suppose each element of x_t indicates the amount of one component of this product, and therefore each element of θ_t^* indicates the unit price of this component. In this example, $\|\theta_t^* - \theta_{t+1}^*\|_1$ reflects the general price fluctuations on the market, i.e., the sum of market-wise price changes over all components.

To characterize the performance of a pricing scheme under this non-stationary setting, we adopt the concept of *dynamic regret*. In this notion, we compare the performance of v_t we propose with that of the optimal pricing policy that knows the sequence of θ_t^* in advance. A rigorous definition of this dynamic regret will be presented in Section 9.2.3.

Proper Learning. Usually, the actions/strategies we are allowed to adopt are restricted in some specific *safe domains*. Taking any action/strategy outside this domain would probably cause risky, illegal or inconsistent outcomes. Our algorithm works by maintaining an estimate, θ_t, for the true valuation parameter θ_t^* at each round t, and we in turn take θ_t as a *parametric strategy* for proposing the price v_t according to a greedy policy (see Section 9.2.3 for more details). In this work, we require that the estimate θ_t must fall in a specific convex and closed domain $\mathcal{D}_t$ at each round t. Here $\mathcal{D}_t$ can be chosen adversarially with the constraint imposed by Assumption 82. As will be explained in Section 9.2.4, this is to address the fact that pricing strategies must conform to hard constraints due to safety restrictions.

Universal Dynamic Regret and Proper OCO with co-variates. Next, we take a digression and describe a general Online Convex Optimization (OCO) setting which will play a pivotal role in solving the contextual pricing problem.

Proper OCO with co-variates. For $t = 1, 2, ..., T$:

1. Adversary reveals a co-variate $x_t \in \mathbb{R}^d$.
2. Learner makes a decision $\hat{\theta}_t$ in a convex domain $\mathcal{D}_t \subset \mathbb{R}^d$.
3. Adversary reveals a convex loss function $\ell_t(\theta) = g_t(\theta^T x_t)$.

This setting embodies OCO under a wide range of loss functions from the generalized linear model (GLM) family for appropriate choices of g_t. The co-variates x_t can be thought of as a feature that encodes valuable information about the context in round t which can be used by the learner to make its predictions. Examples of this setting include (but are not limited to) linear regression and logistic regression.

The goal of the learner is to control its universal dynamic regret:

$$R(w_{1:T}) := \sum_{t=1}^{T} \ell_t(\hat{\theta}_t) - \ell_t(w_t), \tag{9.1}$$

where $w_{1:T} = \{w_1, w_2, \ldots, w_T\}$ is *any* comparator sequence satisfying $w_t \in \mathcal{D}_t$ for all $t \in [T]$. This is known to be a good metric in characterizing the performance of a learner in non-stationary environments [5]. Dynamic regret bounds are usually expressed in literature as functions of the time horizon T and a path length that captures the smoothness of the comparator sequence such as $C_T = \sum_{t=1}^{T-1} \|w_t - w_{t+1}\|_1$.

9.1.1 Summary of Contributions

Our main contributions are given below.

1. We present an algorithm ProDR (Algorithm 5) that attains an *optimal* $\tilde{O}(d^3(T^{\frac{1}{3}}C_T^{\frac{2}{3}} \vee 1))$ dynamic regret (modulo dependencies in d and $\log T$) for the setting of proper OCO with co-variates under exp-concave losses (see Section 9.3.1).

2. We construct an algorithm PDRP (Algorithm 6) with a base learner ProDR, which solves the non-stationary contextual pricing problem with strictly log-concave noise. We define the dynamic regret of contextual pricing as Eq.(9.5) and show that PDRP achieves a $\tilde{O}(d^3(T^{\frac{1}{3}}C_T^{\frac{2}{3}} \vee 1))$ dynamic regret guarantee (see Section 9.3.2).

3. We show that any algorithm must incur a dynamic regret of $\Omega(T^{\frac{1}{3}}C_T^{\frac{2}{3}} \vee 1)$ in the contextual pricing problem, which says that PDRP is minimax optimal up to d and $\log T$ factors (see Section 9.3.3).

Novelty. Owing to the reduction of [118], the non-stationary contextual pricing problem can be reduced to an OCO problem with co-variates and exp-concave losses. The key subroutine we developed — ProDR — is the first to achieve an *optimal* universal dynamic regret with *exp-concave* losses in the *proper OCO with covariate* setting. ProDR makes considerable progress towards addressing the open problem posed by [65] on the more general version of the above problem with *general exp-concave* losses (rather than GLM with known covariates) . The only existing attempt to this open problem requires the decision set to be an L_∞ ball [120], which cannot be used to handle arbitrary convex decision sets as we do.

Summary of techniques. The key technique in deriving ProDR is a novel "transfer theorem" which takes the algorithm of [120] (L_∞ ball decision set) and converts it to an optimal algorithm for the setting of proper OCO with co-variates under *arbitrary convex decision sets.* This idea is similar in spirit to the improper-proper reduction in the work of [121] where they consider general convex losses. However, a direct application of their reduction scheme cannot give fast rates for exp-concave losses. To circumvent this issue, we propose new reduction schemes that carefully take the curvature of the losses into account thereby allowing us to derive fast and optimal dynamic regret rates under exp-concave losses (see Section 9.3.1 for a list of technical challenges). Such a "transfer theorem" could be of independent interest and impactful in the general context of non-stationary online learning. That the non-stationary dynamic pricing problem can be optimally solved using ProDR is a testament to this fact. The material of this chapter closely follows [122].

9.2 Notations and Problem Setup

In this section, we specify necessary mathematical symbols and notations, and define functions for algorithm design and regret analysis. We also present three examples to illustrate the concept of proper learning in contextual pricing.

9.2.1 Symbols and Notations.

The pricing process consists of T rounds. $x_t, \theta_t^* \in \mathbb{R}^d$, $y_t \in \mathbb{R}$, $v_t \in \mathbb{R}_+$ and $N_t \in \mathbb{R}$ denote the feature vector, the linear valuation parameter, the customer's valuation, the seller's price and the noise at time t, sequentially. At each round, we receive a payoff (reward) $r_t = v_t \cdot \mathbb{I}_t$, where the binary variable $\mathbb{I}_t$ indicates the *customer's decision*, i.e., $\mathbb{I}_t = \mathbf{1}(v_t \leq y_t)$.

9.2.2 Technical Assumptions

Denote a norm-bounded domain family $\mathcal{D}_p^B = \{\theta \in \mathbb{R}^d, \|\theta\|_p \leq B\}$. We firstly present assumptions on domain constraints of x_t and θ_t^*:

Assumption 82 (Domain Constraints)**.** Assume $x_t \in \mathcal{D}_x$ where $\mathcal{D}_x \subseteq \mathcal{D}_2^1$ is convex and closed, and $\theta_t^* \in \mathcal{D}_t$ where every $\mathcal{D}_t \subseteq \mathcal{D}_2^B \subset \mathcal{D}_\infty^B$ is also convex and closed. Each $\mathcal{D}_t$ can be chosen adversarially and is known to the learner before time t.

Here we want the customers' valuations to be bounded. Equivalently, we may also assume that $\mathcal{D}_x \subseteq \mathcal{D}_2^{B_1}$ and $\mathcal{D}_t \subseteq \mathcal{D}_2^{B/B_1}$ for any $B_1 > 0$. With these assumptions, we know that $|x_t^\top \theta| \leq B, \forall \theta \in \mathcal{D}_2^B, t = 1, 2, \ldots, T$. Next, we make a reasonable assumption on customers' expected valuations:

Assumption 83 (Non-Negative Expected Valuation)**.** For a customer's valuation $y_t = x_t^\top \theta_t^*$, we assume the *expected valuation* $x_t^\top \theta_t^* \geq 0, t = 1, 2, \ldots, T$.

Now we make assumptions on the distribution of noise N_t. We firstly present the definitions of *log-concavity* and *strict log-concavity* on 1-dimensional distributions according to [123].

Definition 84 (Log-concavity and strict log-concavity)**.** A probability measure P defined on $\mathbb{R}$ is said to be *log-concave* if and only if for any pair $A, B \subset \mathbb{R}$ of intervals, it holds that

$$P(\lambda A + (1-\lambda)B) \geq \{P(A)\}^{\lambda}\{P(B)\}^{1-\lambda}, \forall \lambda \in (0,1). \tag{9.2}$$

Here "+" denotes Minkowski addition. Also, P is *strictly log-concave* if and only if for any pair $A, B \subset \mathbb{R}$ of intervals, $A \neq B$, it holds that

$$P(\lambda A + (1-\lambda)B) > \{P(A)\}^{\lambda}\{P(B)\}^{1-\lambda}, \forall \lambda \in (0,1). \tag{9.3}$$

Then we make the following assumption:

Assumption 85 (Valuation noise distribution)**.** At each time $t = 1, 2, \ldots, T$, the noise N_t is independently and identically sampled from a fixed *strictly log-concave* distribution $\mathbb{D}$ with a twice continuously differentiable cumulative distribution function (CDF) F. Furthermore, the first and second derivatives of the CDF, denoted as f and f', respectively are bounded by two finite constants $B_f := \sup_{\omega\in\mathbb{R}} f(\omega)$ and $B_{f'} := \sup_{\omega\in\mathbb{R}} |f'(\omega)|$.

According to Definition 84, let (i) $A = (-\infty, x], B = (-\infty, y]$ and (ii) $A = (x, +\infty), B = (y, +\infty)$ respectively, and we have F and $(1-F)$ are both strictly log-concave functions. Existing works on contextual pricing also adopt log-concavity assumptions [117]. For a detailed discussion on log-concave distributions, we kindly refer the audience to [124].

All of those three assumptions are supposed to hold throughout the section.

9.2.3 Functions and Key Quantities

Greedily Pricing. Here we adopt two functions defined by [118] and also make use of their properties. Firstly, we introduce an *expected reward* function $g(v, u) := \mathbb{E}[r_t | v_t = v, x_t^\top\theta^* = u] = v \cdot (1 - F(v-u))$ that is unimodal w.r.t. v. Secondly, we introduce a *greedily pricing* function $J(u) := \text{argmax}_{v\in\mathbb{R}} g(v, u)$. $J(u)$ has two important properties: On the one hand, $J(u)$ is strictly monotonically increasing, with $J'(u) \in (0,1)$. Therefore, $J(u)$ and $J^{-1}(v)$ are bijections, $\forall u \in \mathbb{R}, v > 0$. On the other hand, we have $\|\nabla_\theta J(x^\top\theta)\|_2 = |J'(x^\top\theta)| \cdot \|x\|_2 \leq 1$, which guarantees a low price-changing rate while modifying parameter θ.

Restrictions on Actions/Parametric Strategies. When we take an action by presenting a price v_t, there always exists an $\theta_t \in \mathbb{R}^d$ such that $x_t^\top\theta_t = J^{-1}(v_t)$. Therefore, for any price $v_t > 0$, it is equivalent to firstly propose a corresponding *parametric strategy* θ_t (satisfying $x_t^\top\theta_t = J^{-1}(v_t)$) and then set the price as $J(x_t^\top\theta_t)$. Since we are approaching the optimal price (which is $J(x_t^\top\theta^*)$) and that $\theta_t^* \in \mathcal{D}_t$, we may restrict the strategy θ_t to be taken within $\mathcal{D}_t$ at each time t. We will explain more on the motivation of the restrictions in Section 9.2.4.

Negative Log-likelihood. We define

$$\ell_t(\theta) = -\mathbb{I}_t \cdot \log\left(1 - F(v_t - x_t^\top\theta)\right) - (1 - \mathbb{I}_t)\log\left(F(v_t - x_t^\top\theta)\right) \tag{9.4}$$

as a negative log-likelihood function at round t. Also, we define an expected log-likelihood function $L_t := \mathbb{E}_{N_t}[\ell_t(\theta)|x_t, \theta_t^*]$. For the simplicity of notations in the following sections, we denote $h_t(\theta) := \frac{\partial \ell_t(\theta)}{\partial x_t^\top\theta} \in \mathbb{R}$, and we show a property of $h_t(\theta)$:

Lemma 86. *For $\theta \in \mathcal{D}_2^B$, there exist constants $0 < h_{\min} \leq h_{\max} < +\infty$ such that*

$$h_{\max} = \sup_{\theta\in\mathcal{D}_2^B} |h_t(\theta)|, h_{\min} = \inf_{\theta\in\mathcal{D}_2^B} |h_t(\theta)|, \forall t = 1, 2, \ldots, T.$$

We prove this by noticing that $h(\theta)$ is continuous and $\mathcal{D}_2^B$ is closed, and the details are in Appendix H.1.1. With this lemma, we may know that $\ell_t(\theta)$ is Lipschitz (see Lemma 97).

Dynamic Regret. Finally, we define the cumulative *dynamic regret*:

$$\mathbf{Reg}_T = \sum_{t=1}^{T} g(J(x_t^\top \theta_t^*), x_t^\top \theta_t^*) - g(v_t, x_t^\top \theta_t^*). \tag{9.5}$$

We usually measure the regret as a function of T, d and the total variation $C_T := \sum_{t=1}^{T-1} \|\theta_t^* - \theta_{t+1}^*\|_1$.

9.2.4 Examples

Here we present three examples where the nature requires the strategies to lie in a "safe domain", regarding risk-taking, legal or consistency concerns.

Risk Control Adopting strategies outside a pre-defined and protected decision set can be very risky in general. Concerning our contextual pricing problem, an extremely low price would lead to significant loss of profit. Therefore, we have to set a lower pricing bar for each item. At each time t, suppose the lower bar is $c_t > 0$, and therefore our parametric strategy θ_t should satisfy $c_t \leq J(x_t^\top \theta_t)$. Since $J(u)$ is monotonically increasing, we have $x_t^\top \theta_t \geq J^{-1}(c_t)$. By intersecting $\{\theta \in \mathbb{R}^d | x_t^\top \theta \geq J^{-1}(c_t)\}$ with the L_2-norm ball $\mathcal{D}_2^B$, we get a convex and compact set $\mathcal{D}_t$, in which any parametric strategy θ satisfies the norm bound and will lead to a price not less than c_t given the $J(x_t^\top \theta)$ greedy pricing policy.

Legal Concern There exist laws or regulations regarding the highest price of some specific products. For each item with feature x_t, suppose that we cannot set a price exceeding $c_t > 0$. Equivalently, the parametric strategy θ_t we take must satisfy $v_t = J(x_t^\top \theta_t) \leq c_t$. Since $J(u)$ is monotonically increasing, this is further equivalent to $x_t^\top \theta_t \leq J^{-1}(c_t)$. Therefore, the restricted strategy space $\mathcal{D}_t$ is the intersection of $\{\theta | x_t^\top \theta \leq J^{-1}(c_t)\}$ with the L_2-norm ball $\mathcal{D}_2^B$, which is a convex and compact set. Any parametric strategy falling out of $\mathcal{D}_t$ would lead to either $v_t > c_t$ or $\|\theta\| > B$.

Price Consistency It is important for the seller to be consistent on setting prices, or

Algorithm 5 Proper Dynamic Regret minimization (ProDR)

Input: Base algorithm $\mathcal{A}$, barrier multiplier $G' > 0$, exp-concavity factor β.
for $t = 1, 2, \ldots, T$: **do**
  Get iterate $\tilde{\theta}_t$ from $\mathcal{A}$.
  Feature x_t and proper domain $\mathcal{D}_t$ are revealed
  Output $\hat{\theta}_t = \text{argmin}_{\theta \in \mathcal{D}_t} |x_t^\top(\theta - \tilde{\theta}_t)|$.
  Loss ℓ_t is revealed.
  Construct $\hat{\ell}_t(\theta)$ as in Eq.(9.7) and set

$$f_t(\theta) = \hat{\ell}_t(\theta) + G' \cdot S_t(\theta),$$

  where $S_t(\theta) = \min_{\eta \in \mathcal{D}_t} |\nabla \hat{\ell}_t(\hat{\theta}_t)^\top (\eta - \theta)|$;
  Send $f_t(\theta)$ to $\mathcal{A}$ as loss at time t.
end for

otherwise it might cause pricing discrimination. Specifically, if two identical items with feature x occur at time t and $t+1$, then their prices must be close to each other. In other words, we require $|J(x^\top\theta_t) - J(x^\top\theta_{t+1})| \leq C, \forall x \in \mathcal{D}_x \subset \mathcal{D}_2^1$ for some constant $C > 0$. For each $x \in \mathcal{D}_x$, we may solve it and get

$$J^{-1}(J(x^\top\theta_t) - C) \leq x^\top\theta_{t+1} \leq J^{-1}(C + J(x^\top\theta_t)).$$

Denote this set as $\mathcal{S}_t(x)$, and we have $\mathcal{D}_{t+1} \subseteq \cap_{x \in \mathcal{D}_x} \mathcal{S}_t(x)$. Since $\theta_t \in \mathcal{S}_t(x), \forall x$, the intersection is non-empty.

9.3 Main Results

In this section, we present and analyse our algorithms. In Section 9.3.1, we first study the more general problem of universal dynamic regret (Eq.(9.1)) minimization in a proper OCO setting. Results of Section 9.3.1 will be applied in Section 9.3.2 to derive an optimal algorithm for the non-stationary pricing problem. All omitted proofs in this section are deferred to Appendix H.1.

9.3.1 Dynamic Regret of ProDR

In this section, we study the **Pro**per **D**ynamic **R**egret minimization (ProDR) algorithm (Algorithm 5). We consider the protocol of proper OCO with co-variates introduced in Section 9.1.

The goal of this section is to control the universal dynamic regret as defined in Eq.(9.1). We start by listing out the assumptions we made for the OCO problem.

Assumption 87. A constant $B > 0$ is known such that $\max_{\theta \in \mathcal{D}_t} \|\theta\|_\infty \leq B$ for all $t \in [n]$.

Assumption 88. The losses ℓ_t obey $\|\nabla \ell_t(\theta)\|_2 \leq G$ for all $t \in [n]$ and $\theta \in \mathcal{D}_t$ (recall that $\mathcal{D}_t \subseteq \mathcal{D}_2^B$ from Section 9.2.2).

Assumption 89. The losses are α exp-concave. i.e $\ell_t(y) \geq \ell_t(x) + \nabla \ell_t(x)^\top (y-x) + \frac{\alpha}{2}\left(\ell_t(x)^\top (y-x)\right)^2$, for $\alpha > 0$ and for all $x, y \in \mathcal{D}_2^B$.

Assumption 87 puts a relatively mild constraint that a box enclosing all the decision sets is known ahead of time. Lipschitzness assumptions like Assumption 88 are standard in online learning. Assumption 89 states that the loss ℓ_t exhibits a strong curvature in the direction of its gradients [77]. We will exploit this curvature to derive fast regret rates.

Qualitative description of ProDR. The base algorithm $\mathcal{A}$ in ProDR is expected to optimally control the dynamic regret under exp-concave losses and when the decision set is a box: $\mathcal{D}_\infty^B = \{x \in \mathbb{R}^d : \|x\|_\infty \leq B\}$, where B is as in Assumption 87. The idea is to perform a black-box reduction that can convert the base algorithm $\mathcal{A}$ to an algorithm that attains good dynamic regret guarantee on the domains $\mathcal{D}_t$. Though similar ideas have been already explored in the work of [121], our way of constructing such reductions for the current problem is new and interesting in its own right in the context of exp-concave online learning. Next, we expand upon this matter highlighting the differences from [121]. We construct losses f_t in Line 7 of ProDR where the $S_t(\theta)$ term acts as a regularizer that penalizes $\mathcal{A}$ for predicting points outside $\mathcal{D}_t$. We would like the losses f_t to be exp-concave as the base algorithm $\mathcal{A}$ expects. However, a direct application of the techniques of [121] does not satisfy this property. We address this issue by carefully constructing f_t as in Line 7 of Algorithm 5 such that: 1) gradients of both $\hat{\ell}_t(\theta)$ and $S_t(\theta)$ lie in the span of co-variate x_t and 2) $\hat{\ell}_t(\theta)$ is exp-concave, meaning that it exhibits strong curvature along the direction of x_t. Now, 1 and 2 together implies that the surrogate losses f_t still remains exp-concave as it exhibits strong curvature along the direction of its gradient (which is spanned by x_t). The particular choice of $\hat{\ell}_t(\theta)$ is found to be crucial in preventing the exp-concavity factor of losses f_t from collapsing to zero. We will show that the dynamic regret of ProDR w.r.t. losses $\hat{\ell}_t$ is upper bounded by the dynamic regret of the base algorithm $\mathcal{A}$ wrt losses f_t which is well controlled.

We next describe the dynamic regret guarantees of Algorithm 5. We inherit all the notations used in the algorithm description.

Theorem 90. *Let* $\beta = \min\{\alpha/2, 1/(8GB\sqrt{d})\}$ *and* $\gamma = \frac{1}{4\left(2GB\sqrt{d\beta}+1/(2\sqrt{\beta})\right)^2}$ *and* $G' = 1 + 2GB\beta\sqrt{d}$. *Let* $\mathcal{A}$ *in ProDR algorithm be FLH-ONS instantiated with parameters* $\zeta = 2\gamma/25$, $\mathcal{G} = GG'$ *and* $\phi = B$. *Then ProDR (Algorithm 5) satisfes*

$$\sum_{t=1}^{T} \ell_t(\hat{\theta}_t) - \ell_t(w_t) = \tilde{O}\left((d^3\gamma + \frac{d^2}{\gamma})(T^{1/3}C_T^{2/3} \vee 1)\right), \tag{9.6}$$

where $C_T := \sum_{t=2}^{T} \|w_t - w_{t-1}\|_1$ *with* $w_t \in \mathcal{D}_t$. $a \vee b := \max\{a, b\}$ *and* $\tilde{O}$ *hides dependence of constants* G, B, α *and poly-logarithmic factors of* T.

Remark 91 (Adaptivity to C_T)**.** In light of the $\Omega(dB^2 \log T \vee d^{1/3}T^{1/3}C_T^{2/3}B^{4/3})$ lower bound [65], we see that the ProDR algorithm adapts optimally to the path variation C_T of the comparator sequence, which may not be known ahead of time.

Proof. Due to the α exp-concavity of losses ℓ_t over the domain $\mathcal{D}_2^B$ and $\beta \leq \frac{\alpha}{2}$ we have that:

$$\ell_t(\theta) \ \geq \ell_t(\hat{\theta}_t) + \nabla \ell_t(\hat{\theta}_t)^\top(\theta - \hat{\theta}_t) + \beta \left(\nabla \ell_t(\hat{\theta}_t))^\top(\theta - \hat{\theta}_t)\right)^2,$$

for any $\theta \in \mathcal{D}_2^B$. Hence following [77], we consider the linear-regression-type surrogate losses:

$$\hat{\ell}_t(\theta) := \left(\nabla \ell_t(\hat{\theta}_t)^\top(\theta - \hat{\theta}_t)\sqrt{\beta} + \frac{1}{2\sqrt{\beta}}\right)^2. \tag{9.7}$$

Hence for any $\theta \in \mathcal{D}_2^B$ we have that

$$\ell_t(\hat{\theta}_t) - \ell_t(\theta) \leq \frac{1}{4\beta} - \hat{\ell}_t(\theta) = \hat{\ell}_t(\hat{\theta}_t) - \hat{\ell}_t(\theta). \tag{9.8}$$

where we used the fact that $\hat{\ell}_t(\hat{\theta}_t) = \frac{1}{4\beta}$.

Given that $S_t(\theta_t^*) = S_t(\hat{\theta}_t) = 0$ since $\theta_t^*, \hat{\theta}_t \in \mathcal{D}_t$, we have

$$f_t(\theta_t^*) = \hat{\ell}_t(\theta_t^*), f_t(\hat{\theta}_t) = \hat{\ell}_t(\hat{\theta}_t). \tag{9.9}$$

Let us denote $\nabla \ell_t(\theta) = h_t(\theta)x_t$ where $h_t(\theta) = g_t'(x_t^\top \theta)$. Now, according to the definition of $S_t(\theta)$ and $\hat{\theta}_t$, we have:

$$
\begin{aligned}
f_t(\tilde{\theta}_t) =& \hat{\ell}_t(\tilde{\theta}_t) + G' \cdot S_t(\tilde{\theta}_t) \\
=& \hat{\ell}_t(\tilde{\theta}_t) + G' \cdot \min_{\eta \in \mathcal{D}_t} |\nabla \ell_t(\hat{\theta}_t)^\top (\eta - \tilde{\theta}_t)| \\
=& \hat{\ell}_t(\tilde{\theta}_t) + G' \cdot \min_{\eta \in \mathcal{D}_t} |h_t(\hat{\theta}_t)||x_t^\top (\eta - \tilde{\theta}_t)| \\
=& \hat{\ell}_t(\tilde{\theta}_t) + G' \cdot |h_t(\hat{\theta}_t)||x_t^\top (\hat{\theta}_t - \tilde{\theta}_t)| \\
=& \hat{\ell}_t(\tilde{\theta}_t) + G' \cdot |\nabla \ell_t(\hat{\theta}_t)^\top (\hat{\theta}_t - \tilde{\theta}_t)|.
\end{aligned}
\tag{9.10}
$$

Next we proceed to upper bound the regret w.r.t. losses $\hat{\ell}_t$ by the regret w.r.t. losses f_t. We need the following lemma.

Lemma 92. *Under the assumptions of Theorem 90, we have that*

$$|\hat{\ell}_t(\theta) - \hat{\ell}_t(\hat{\theta}_t)| \leq G' |\nabla \ell_t(\hat{\theta}_t)^\top (\theta - \hat{\theta}_t)|, \tag{9.11}$$

for any $\theta \in \mathcal{D}_\infty^B$ where $G' := (1 + 2GB\beta\sqrt{d})$.

The proof is shown in Appendix H.1.2. With this lemma, we have

$$\hat{\ell}_t(\hat{\theta}_t) \leq \hat{\ell}_t(\tilde{\theta}_t) + G' \cdot |\nabla \ell_t(\hat{\theta}_t)^\top (\hat{\theta}_t - \tilde{\theta}_t)| = f_t(\tilde{\theta}_t). \tag{9.12}$$

Combining the above inequality with Eq.(9.9) we obtain

$$\hat{\ell}_t(\hat{\theta}_t) - \hat{\ell}_t(\theta_t^*) \leq f_t(\tilde{\theta}_t) - f_t(\theta_t^*). \tag{9.13}$$

Now using Eq.(9.8) along with the previous relation yields that

$$\sum_{t=1}^{T} \ell_t(\hat{\theta}_t) - \ell_t(\theta_t^*) \leq \sum_{t=1}^{T} f_t(\tilde{\theta}_t) - f_t(\theta_t^*). \tag{9.14}$$

The following lemma specifies how to compute the sub-gradient of the regularizer term $S_t(\theta)$ in Line 7 of Algorithm 5. Further it highlights an important property that a sub-gradient of $S_t(\theta)$ lies in the span of covariate x_t (recall that $\nabla \ell_t(\theta) = h(\theta) x_t$). This is also useful for proving the joint exp-concavity of the losses f_t.

Lemma 93. *The function $S_t(\theta)$ is convex across $\mathbb{R}^d$. Denote $\eta_t(\theta) := \operatorname{argmin}_\eta |x_t^\top(\eta - \theta)|$. When $\nabla \ell_t(\hat{\theta}_t)^\top(\eta_t(\theta) - \theta) \neq 0$, we have:*

$$\nabla S_t(\theta) = \begin{cases} \nabla \ell_t(\hat{\theta}_t), & \textit{if} \quad \nabla \ell_t(\hat{\theta}_t)^\top(\eta_t(\theta) - \theta) < 0 \\ -\nabla \ell_t(\hat{\theta}_t), & \textit{if} \quad \nabla \ell_t(\hat{\theta}_t)^\top(\eta_t(\theta) - \theta) > 0. \end{cases}$$

When $\nabla \ell_t(\hat{\theta}_t)^\top(\eta_t(\theta) - \theta) = 0$, we have $\mathbf{0} \in \partial S_t(\theta)$.

The proof of Lemma 93 is in Appendix H.1.3. In the next lemma, we show that the losses f_t remain exp-concave with appropriate exp-concavity factor **bounded away** from zero. This is the key lemma that helps to control the regret of ProDR.

Lemma 94. *Define $\gamma := \frac{1}{4\left(2GB\sqrt{d\beta} + 1/(2\sqrt{\beta})\right)^2}$. We have that the surrogate losses f_t are $2\gamma/25$ exp-concave and $2GG'$ Lipschitz in L_2 norm across $\mathcal{D}_\infty^B$.*

As is stated earlier in this section, the intuition of this lemma comes from two facts: (1) both $\nabla \hat{\ell}_t(\theta)$ and $\nabla S_t(\theta)$ are in the span of x_t, and (2) $\hat{\ell}_t(\theta)$ is exp-concave. As a result, the strong curvature of $\hat{\ell}_t(\theta)$ along the x_t direction "absorbs" the plain convexity of $S_t(\theta)$ and therefore guarantees the exp-concavity of $f_t(\theta)$. We defer the detailed proof to Appendix H.1.4. Hence from [120] , FLH-ONS algorithm run with parameters $\zeta = 2\gamma/25$,

Algorithm 6 Proper Dynamic Regret Pricing (PDRP)

1: **Input:** Noise distribution $\mathbb{D}$ (including its CDF F and PDF f).
ProDR algorithm $\mathcal{A}$ instantiated as in Theorem 90.
2: **for** $t = 1, 2, \ldots, T$: **do**
3: Feature x_t and proper domain $\mathcal{D}_t$ are revealed and sent to $\mathcal{A}$.
4: Get $\hat{\theta}_t \in \mathcal{D}_t$ from $\mathcal{A}$.
5: Seller proposes $v_t = J(x_t^\top \hat{\theta}_t)$ and receive $\mathbb{I}_t$.
6: Send loss $\ell_t(\theta)$ defined in Eq.(9.4) to $\mathcal{A}$.
7: **end for**

$\mathcal{G} = GG'$ and $\phi = B$ can be used to control

$$\sum_{t=1}^{T} f_t(\tilde{\theta}_t) - f_t(\theta_t^*) = \tilde{O}\left(d^2(G^2(G')^2B^2\gamma d + G^2(G')^2B^2 + \frac{1}{\gamma})(T^{1/3}C_T^{2/3} \vee 1)\right)$$
$$= \tilde{O}\left(d^3(T^{1/3}C_T^{2/3} \vee 1)\right), \tag{9.15}$$

where the last line is got by plugging in the values of γ and G' and upper bounding further. □

9.3.2 Dynamic regret of PDRP

In this section, we present our main algorithm for controlling the dynamic regret on contextual pricing problem, the **P**roper **D**ynamic **R**egret **P**ricing (PDRP) (Algorithm 6).

Qualitative description of PDRP. [118] observes that the pricing problem can be reduced to the setting of proper OCO with co-variates and exp-concave losses. This observation, armed with the ProDR algorithm, naturally lends itself to the algorithm PDRP for controlling dynamic regret of the pricing problem.

We are now ready to present regret guarantees for the non-stationary pricing problem.

Theorem 95. *Consider the linear noisy contextual pricing problem defined in Section 9.1. Assume that we know the noise distribution $\mathbb{D}$ exactly. By properly initializing β, γ*

and G' with pre-knowledge, PDRP (Algorithm 6) obeys $\mathbf{Reg}_T = \tilde{O}(d^3(T^{\frac{1}{3}}C_T^{\frac{2}{3}} \vee 1))$, where $\mathbf{Reg}_T$ is as defined in Eq.(9.5), $\tilde{O}$ hides poly-logarithmic factors of T and $(a \vee b) = \max\{a, b\}$.

Proof. We start with the lemmas that help us leverage the OCO framework of Section 9.3.1.

Lemma 96. *[118] Under the assumptions in Theorem 95, for $\theta \in \mathcal{D}_2^B$, we have:*

$$g(J(x_t^\top \theta_t^*), x_t^\top \theta_t^*) - g(J(x_t^\top \theta), x_t^\top \theta_t^*) \leq \frac{2C}{C_{down}} \left(E[\ell_t(\theta) - \ell_t(\theta_t^*)]\right), \tag{9.16}$$

where ℓ_t is defined in Eq.(9.4), $C = 2B_f + (B + J(0))B_{f'}$ and

$$C_{down} := \inf_{\omega \in [-B, B+J(0)]} \min\left\{\frac{d^2 \log(1 - F(\omega))}{d\omega^2}, \frac{d^2 \log(F(\omega))}{d\omega^2}\right\} > 0.$$

So we have

$$\mathbf{Reg}_T \leq \frac{2C}{C_{down}} \mathbb{E}[\ell_t(\hat{\theta}_t) - \ell_t(\theta_t^*)]. \tag{9.17}$$

Next, we record the curvature and smoothness properties of losses ℓ_t.

Lemma 97. *Let $G = h_{\max}$ defined in Lemma 86. Under the assumptions in Theorem 95, for $\theta \in \mathcal{D}_t$, we have: (1) $\ell_t(\theta)$ is G-Lipschitz in $\|\cdot\|_2$ norm, and (2) $\ell_t(\theta)$ $\frac{C_{down}}{C_{exp}}$-exp-concave. Here $C_{exp} := \sup_{\omega \in [-B, B+J(0)]} \max\left\{\frac{f(\omega)^2}{F(\omega)^2}, \frac{f(\omega)^2}{(1-F(\omega))^2}\right\}$ and C_{down} is defined in Lemma 96 .*

This lemma is derived from [118] Lemma 7, and we defer the proof to Appendix H.1.5. The lemma above implies that the losses satisfy Assumption 88 in Section 9.3.1.

Further they satisfy Assumption 89 with exp-concavity factor of C_{down}/C_{exp}. So we can use the ProDR algorithm (Algorithm 5) to control the dynamic regret wrt losses ℓ_t. Let $\beta = \min\{C_{down}/(2C_{exp}), 1/(8GB\sqrt{d})\}$ and $\gamma = \frac{1}{4\left(2GB\sqrt{d}\beta+1/(2\sqrt{\beta})\right)^2}$ and $G' = 1 + GB\sqrt{d}C_{down}/C_{exp}$. Hence continuing from Eq.(9.17), we apply Theorem 90 to obtain

$$\mathbf{Reg}_T \leq \tilde{O}\left(d^3(T^{1/3}C_T^{2/3} \vee 1)\right). \tag{9.18}$$

This completes the proof of the theorem. □

Remark 98. Although noise distributions are known as we assumed, the coefficient of our regret upper bound depends highly on the distribution. As is indicated by [118], when the noise N_t is an i.i.d. Gaussian noise with zero mean and σ standard deviation, this coefficient is exponentially large w.r.t. $\frac{1}{\sigma}$ as σ approaches 0, which is counter-intuitive.

9.3.3 Lower Bound on Dynamic Pricing Regret

So far, we have developed a ProDR algorithm that is suitable for domain-constraint optimization of generalized linear model, and have constructed a PDRP algorithm to solve the linear contextual pricing problem where PDRP achieves a $\tilde{O}(d^3(T^{\frac{1}{3}}C_T^{\frac{2}{3}} \vee 1))$ dynamic regret. This upper regret bound is optimal for online exp-concave optimization as is shown by [65], but is it still optimal for our feature-based dynamic pricing setting in specific? The answer is Yes. This dynamic regret is near-optimal up to d and $\log T$ factors, and here we present the following theorem.

Theorem 99 (Lower dynamic regret bound)**.** *Let $d = 1$ in the contextual pricing problem we consider. For any algorithm $\mathcal{A}$, there exists a specific problem setting where $\mathcal{A}$ has to suffer an $\Omega(T^{\frac{1}{3}}C_T^{\frac{2}{3}} \vee 1)$ expected dynamic regret.*

With this theorem, we may claim that our PDRP algorithm is near-optimal. We here show a proof sketch and defer the full proof to Appendix H.1.6.

Proof Sketch. The proof is developed in three steps: Firstly, we construct a hypothesis set $\mathbf{\Theta}$ in which there are N different $\{\theta_t^*\}_{t=1}^T$ series whose total variations are upper bounded by C_T. For any pair of two different series $\{\theta_t^*\}_{t=1}^T$'s in $\mathbf{\Theta}$, they are identical for $T/3$ out of T rounds in total, and are different by some small δ for the rest $2T/3$ rounds. Secondly, we show that their corresponding feedback distributions are also "similar" to each other under the metric of KL-divergence. Therefore, according to Fano's Inequality, any algorithm can hardly distinguish among these distributions. Finally, we show that a failure of correctly distinguish the underlying distribution (i.e., the real $\{\theta_t^*\}_{t=1}^T$ series) will result in an $\Omega(T^{\frac{1}{3}}C_T^{\frac{2}{3}} \vee 1)$ regret. □

9.4 Concluding Discussion

In this work, we studied the non-stationary contextual pricing problem under safety constraints. We first presented the ProDR algorithm for minimizing universal dynamic regret in the framework of proper OCO with co-variates and exp-concave losses. This contribution could be of independent interest in the context of non-stationary online learning. As a concrete application, we constructed our pricing algorithm, PDRP, by making use of ProDR as the base learner. We showed that PDRP attains a $\tilde{O}(d^3(T^{\frac{1}{3}}C_T^{\frac{2}{3}} \vee 1))$ dynamic regret in our pricing problem setting. Finally, we proved that this rate is information-theoretically optimal (modulo dependencies on d and $\log T$).

Chapter 10

Optimal Dynamic Regret in LQR Control

This chapter studies the linear quadratic regulator (LQR) control problem which is a specific instantiation of the more general RL framework where the evolution of states follows a predefined linear dynamics. At each round $t \in [n] := \{1, \dots, n\}$, the agent is at state $x_t \in \mathbb{R}^{d_x}$. Based on the state, the agent select a control input $u_t \in \mathbb{R}^{d_u}$. The next state evolves according to the law:

$$x_{t+1} = Ax_t + Bu_t + w_t, \tag{10.1}$$

where A and B are system matrices known to the agent. $w_t \in \mathbb{R}^{d_x}$ is a disturbance term that can be selected by a potentially adaptive adversary. We assume that $\|w_t\|_2 \leq 1$. This disturbance term reflects the perturbation from the ideal linear state transition arising due to environmental factors that could be difficult to model. The loss suffered by playing the control u at state x is given by $\ell(x, u) := x^T R_x x + u^T R_u u$, where $R_x, R_u \succcurlyeq 0$, that are apriori fixed and known.

Recently there has been a surge of interest in viewing this classical LQR problem under the lens of online learning [59]. The work of [125] places regret of the agent against a set of benchmark policies as the central notion to evaluate learner's performance. Following [125, 126] we adopt the class of disturbance action policies (DAP) as our benchmark class:

Definition 100. (Disturbance action policies, [126]). Let $M = (M^{[i]})_{i=1}^{m}$ denote a sequence of matrices $M^{[i]} \in \mathbb{R}^{d_u \times d_x}$. We define the corresponding disturbance action policies (DAP) π^M as:

$$\pi_t^M(x_t) = -K_\infty x_t - q^M(w_{1:t-1}), \tag{10.2}$$

where $q^M(w_{1:t-1}) = \sum_{i=1}^{m} M^{[i]} w_{t-i}$ and K_∞ as in Eq.(10.6). We are interested in DAPs for which the sequence M belongs to the set:

$$\mathcal{M}(m, R, \gamma) := \{M = (M^{[i]})_{i=1}^{m} : \|M^{[i]}\|_{\text{op}} \leq R\gamma^{i-1}\}, \tag{10.3}$$

where m, R and γ are algorithm parameters.

This class is known to be sufficiently rich to approximate many linear controllers. A policy takes in the past history and current state as input and produces a control signal as output. Let's denote $M_{1:n} := (M_1, \ldots, M_n)$ to be a sequence of DAP policies such that at time t, the control signal is selected using the policy parameterized by M_t (see Eq.(10.2)). We denote $x_t^{M_{1:n}}$ to be the state reached at round t by playing the sequence of policies defined by parameters $M_{1:t-1}$ in the past. Similarly $u_t^{M_{1:n}}$ is used to denote the control signal produced by the policy M_t. The *universal dynamic regret* of the learner against the policy sequence $M_{1:n}$ is defined as:

$$R(M_{1:n}) = \sum_{t=1}^{n} \ell(x_t^{\text{alg}}, u_t^{\text{alg}}) - \ell(x_t^{M_{1:n}}, u_t^{M_{1:n}}), \tag{10.4}$$

where $(x_t^{\text{alg}}, u_t^{\text{alg}})$ denotes the state and control signal of the learner at round t. Note that the policy sequence $M_{1:n}$ can be *any* valid sequence of DAP polices. The main focus of this section is to design algorithms that can control the dynamic regret against a sequence of reference policies as a function of the time horizon n and the a path variation of the DAP parameters of the comparator $M_{1:n}$. We remark that the comparator polices $M_{1:n}$ can be chosen in hindsight and potentially unknown to the learner.

Whenever $M_{1:n} = (M, \ldots, M)$ for a fixed parameter M, we recover the notion of static regret. However the notion of static regret is not befitting for non-stationary environments. For example consider the scenario of controlling a drone. Suppose during the initial half of the trajectory there is heavy wind eastwards and in the second half, wind blows westwards. For best performance, a controller has to choose different policies that can counter-act the wind and guide the motion properly in each half. Hence, we aim to control the dynamic regret which allows us to be competent against a sequence of potentially time-varying polices chosen in hindsight. We remark that our algorithm automatically adapts to the level of non-stationarity in the hindsight sequence of policies.

Next, we take a digression and discuss a desirable property for the design of algorithms for LQR control.

Proper learning in LQR control. Proper learning is an online learning paradigm where the decisions of the learner are required to obey some user specified physical limits. On the other hand, improper learning framework allows the learner to disregard such constraints. The paradigm of improper learning may not be attractive in certain applications where safety is a paramount concern. Improper algorithms can possibly take the system through trajectories that are deemed to be risky. It is desirable to avoid such behaviours in physical systems such as self driving cars, control of medical ventilators, robotic control [127] and cooling data centers [128]. A policy selects a control signal u_t depending on the current state x_t. Given the value of current state, there can be physical constraints on the allowable control actions. For example, imagine the situation where we want to maintain the velocity of a drone. Depending on the current position and other system and environmental factors (the state), one can only apply a range of allowable torque (the control action) to the blades. Not respecting this torque range can drain the battery quickly or can lead to catastrophic damages such as burnt rotors. In our framework, we model this set of allowable control actions at a state as $\mathcal{F}_t := \{u_t | u_t = \pi_t^M(x_t) \text{ for some } M \in \mathcal{M}(m, R, \gamma)\}$ (see Definition 100). So to ensure safety, at each round the learner plays a control signal from the feasible set $\mathcal{F}_t$ thus necessitating the need for proper learning.

Below are our contributions:

- We develop an optimal universal dynamic regret minimization algorithm for the general mini-batch linear regression problem (see Theorem 104).
- Applying the reduction of [126] from LQR problem to online linear regression, the above result lends itself to an algorithm for controlling the dynamic regret of the LQR problem (Eq.(10.4)) to be $\tilde{O}^*(n^{1/3}[\mathcal{TV}(M_{1:n})]^{2/3})$, where $\mathcal{TV}$ denotes the total variation incurred by the sequence of DAP policy parameters in hindsight (see Corollary 109). O^* hides the dependencies in dimensions and system parameters.
- We show that the aforementioned dynamic regret guarantee is minimax optimal modulo dimensions and factors of $\log n$ (see Theorem 110).
- The resulting algorithm is also strongly adaptive, in the sense that the static regret

against a DAP policy in any local time window is $O^*(\log n)$.

Notes on novelty and impact. As discussed before, the reduction of [126] casts LQR problem to an instance of proper online linear regression. In the context of regression, proper learning means that the decisions of the learner belongs to a user specified convex domain. The main challenge in developing aforementioned contributions rests on the design of an optimal universal dynamic regret minimization algorithm for online linear regression under the setting of *proper learning*. We are not aware of any such algorithms in the literature to-date and the problem remains open. However, there exists an improper algorithm from [65] for controlling the desired dynamic regret. Given this fact, the design of our algorithm is facilitated by coming up with *new* black-box reductions (see Section 10.2) that can convert an improper algorithm for non-stationary online linear regression to a proper one. There are improper to proper black-box reduction schemes given in the influential work of [79]. However, they are developed to support general convex or strongly convex (see Definition 103) losses. The linear regression losses arising in our setting are exp-concave (see Definition 102) which enjoy strong curvature only in the direction of the gradients as opposed to uniformly curved strongly convex losses. Hence the reduction scheme of [79] is inadequate to provide fast regret rates in our setting. In contrast, we develop novel reduction schemes that carefully take the non-uniform curvature of the linear regression losses into account so as to facilitate fast dynamic regret rates (see Section 10.2.2). We remark that the algorithm ProDR.control developed in Section 10.2 can be impactful in general online learning literature. That the non-stationary LQR problem can be *optimally* solved using ProDR.control is a testament to this fact. Further our algorithm is out-of-the-box applicable to more general settings such as non-stationary multi-task linear regression, which is beyond the current scope.

The lower bound we provide in Theorem 110 is also applicable to the more general problem of online non-parametric regression against a Besov space / class of Total Variation bounded functions [29] (see Section 10.3 for more details). The main contribution here is that we provide a new lower bounding strategy that characterizes the correct rate wrt both n and the radius (or path-variation) of the non-parametric function class. This is in contrast with [29] who establish the correct dependency only wrt n. Deriving dependencies wrt both n and the radius / path-variation is imperative in implying a dynamic regret lower bound for the LQR problem. The material of this chapter closely follows [129].

10.1 Preliminaries

We start with a brief overview of the LQR problem for the sake of completeness. The material of this section closely follows [126]. The definitions and notations introduced in this section will be used throughout the section.

A linear control law is given by $u_t = -Kx_t$ for a controller $K \in \mathbb{R}^{d_u \times d_x}$. A linear controller K is said to be stabilizing if $\rho(A - BK) < 1$ where $\rho(A - BK)$ is the maximum of the absolute values of the eigenvalues of $A - BK$. We assume that there exists a stabilizing controller for the system (A, B). For such systems, there exists a unique matrix P_∞ which is the solution to the equation:

$$P = A^T PA + R_x - A^T PB(R_u + B^T PB)^{-1} B^T PA. \tag{10.5}$$

The solution P_∞ is called the infinite horizon Lyapunov matrix. It is an intrinsic property of the system (A, B) and characterizes the optimal infinite horizon cost for control in the absence of noise [130]. We also define the optimal state feedback controller

$$K_\infty := (R_u + B^T P_\infty B)^{-1} B^T P_\infty A, \tag{10.6}$$

the steady state covariance matrix:

$$\Sigma_\infty := R_u + B^T P_\infty B, \tag{10.7}$$

and the closed loop dynamics matrix: $A_{\text{cl},\infty} := A - BK_\infty$.

[126] shows that the problem of controlling the regret in the LQR problem can be reduced to online linear regression problem with delays. Specifically we have the following fundamental result due to [126].

Proposition 101. *Suppose the learner plays policy of the form $\pi_t^{alg}(x) = -K_\infty x + q^{M_t^{alg}}(w_{1:t-1})$. Let the comparator policies take the form $\pi_t(x) = -K_\infty x + q^{M_t}(w_{1:t-1})$ for a sequence of matrices $M_{1:n}$ chosen in hindsight. Then the dynamic regret against the policies $\pi := (\pi_1, \ldots, \pi_n)$ satisfies:*

$$R_n(\pi) \le O(1) + \sum_{t=1}^{n} \hat{A}_t(M_t^{alg}, w_{t:t+h}) - \hat{A}_t(M_t, w_{t:t+h}), \tag{10.8}$$

where the parameters involved in the inequality are defined as below: $\hat{A}_t(M, w_{t:t+h}) := \|q^M(w_{1:t-1}) - q_{\infty;h}(w_{t:t+h})\|_{\Sigma_\infty}^2$. $q_{\infty;h}(w_{t:h+t}) := \sum_{i=t+1}^{t+h} \Sigma_\infty^{-1} B^T (A_{cl,\infty})^{i-1-t} P_\infty w_i$. $h := 2(1-\gamma_\infty)^{-1} \log(\kappa_\infty^2 \beta_*^2 \Psi_* \Gamma_*^2 n^2)$. $\gamma_\infty := \|I - P + \infty^{-1/2} R_x P_\infty^{1/2}\|_{op}^{1/2}$. $\kappa_\infty := \|P_\infty^{1/2}\|_{op} \|P_\infty^{-1/2}\|_{op}$. $\beta_* := \max\{1, \lambda_{min}^{-1}(R_u), \lambda_{min}^{-1}(r_x)\}$.
$\Psi_* = \max\{1, \|A\|_{op}, \|B\|_{op}, \|R_x\|_{op}, \|R_u\|_{op}\}$. $\Gamma_* := \max\{1, \|P_\infty\|_{op}\}$

Observe that the losses $\hat{A}_t(M, w_{t:t+h}) := \|q^M(w_{1:t-1}) - q_{\infty;h}(w_{t:t+h})\|_{\Sigma_\infty}^2 = \hat{A}_t(M, w_{t:t+h}) := \|\Sigma_\infty^{1/2} q^M(w_{1:t-1}) - \Sigma_\infty^{1/2} q_{\infty;h}(w_{t:t+h})\|_2^2$ are essentially linear regression losses. The quantity

$\Sigma_\infty^{1/2} q^M(w_{1:t-1})$ is a linear map from the matrix sequence M to $\mathbb{R}^{d_u}$. However, there is one caveat in that the bias vector at round t given by $\Sigma_\infty^{1/2} q_{\infty;h}(w_{t:t+h})$ is only available at round $t+h = t+O(\log n)$. This issue of delayed feedback can be directly handled using the delayed to non-delayed online learning reduction from [131].

10.2 Non-stationary "mini-batch" linear regression

In view of Proposition 101, the losses of interest are linear regression type losses. So we take a digression in this section and study the problem of controlling dynamic regret in a general linear regression setting.

10.2.1 Linear regression framework

Consider the following linear regression protocol.

- At round t, nature reveals a co-variate matrix $A_t \in \mathbb{R}^{p \times d}$.
- Learner plays $z_t \in \mathcal{D} \subset \mathbb{R}^d$.
- Nature reveals the loss $f_t(z) = \|A_t z - b_t\|_2^2$.

Under the above regression framework, we are interested in controlling the universal dynamic regret against an arbitrary sequence of predictors $u_1, \ldots, u_n \in \mathcal{D}$ (abbreviated as $u_{1:n}$) :

$$R_n(u_{1:n}) = \sum_{t=1}^{n} f_t(z_t) - f_t(u_t). \tag{10.9}$$

Dynamic regret is usually expressed as a function of n and a path variational that captures the smoothness of the comparator sequence. We will focus on the path variational defined by:

$$\mathcal{TV}(u_{1:n}) = \sum_{t=2}^{n} \|u_t - u_{t-1}\|_1. \tag{10.10}$$

Below are the list of assumptions made:

Assumption 1. Let $a_{t,i} \in \mathbb{R}^d$ be the i^{th} row vector of A_t. We assume that $\|a_{t,i}\|_1 \leq \alpha$ for all $t \in [n]$ and $i \in [p]$. Further $\|b_t\|_1 \leq \sigma$ for all t.

Assumption 2. For any $x \in \mathcal{D}$, $\|x\|_1 \leq \chi$ and $\|x\|_\infty \leq \tilde{R}$.

We refer this setting as mini-batch linear regression since the loss at round t can be written as a sum of a batch of quadratic losses: $f_t(z) = \sum_{i=1}^{p} \left(z^T a_{t,i} - b_t[i]\right)^2$.

Terminology. For a convex loss function f, we abuse the notation and take $\nabla f(x)$ to be a sub-gradient of f at x. We denote $\mathcal{D}_\infty(\tilde{R}) := \{x \in \mathbb{R}^d : \|x\|_\infty \leq \tilde{R}\}$.

Linear regression losses belong to a broad family of convex loss functions called exp-concave losses:

Definition 102. A convex function f is α exp-concave in a domain $\mathcal{D}$ if for all $x, y \in \mathcal{D}$ we have $f(y) \geq f(x) + \nabla f(x)^T(x-y) + \frac{\alpha}{2}(\nabla f(x)^T(x-y))^2$.

The losses $f_t(z) = \|A_t z - b_t\|_2^2$ are $(2R)^{-1}$ exp-concave if $f(z) \leq R$ for all $z \in \mathcal{D}$ (see Lemma 2.3 in [126]).

Definition 103. A convex function f is σ strongly convex wrt $\|\cdot\|_2$ norm in a domain $\mathcal{D}$ if for all $x, y \in \mathcal{D}$ we have $f(y) \geq f(x) + \nabla f(x)^T(x-y) + \frac{\sigma}{2}\|x-y\|_2^2$.

We note that if the matrix A_t is rank deficient, then the losses $f_t(z)$ cannot be strongly convex. Moving forward we do not impose any restrictive assumptions on the rank of A_t. As mentioned in Remark 111, the covariate matrix that arise in the reduction of the LQR problem to linear regression is not in general full rank. So we target a solution that can handle general covariate matrices irrespective of their rank.

10.2.2 The Algorithm

ProDR.control: Inputs - Decision set $\mathcal{D}$, $G > 0$

1. At round t, receive w_t from $\mathcal{A}$.
2. Receive co-variate matrix $A_t := [a_{t,1}, \ldots, a_{t,p}]^T$.
3. Play $\hat{w}_t \in \text{argmin}_{x \in \mathcal{D}} \max_{i=1,\ldots,p} |a_{t,i}^T(x - w_t)|$.
4. Let $\ell_t(w) = f_t(w) + G \cdot S_t(w)$, where $f_t(w) = \|A_t w - b_t\|_2^2$ and $S_t(w) = \min_{x \in \mathcal{D}} \max_{i=1,\ldots,p} |a_{t,i}^T(x - w)|$.
5. Send $\ell_t(w)$ to $\mathcal{A}$.

Figure 10.1: ProDR.control: An algorithm for non-stationary and proper linear regression.

Starting point of our algorithm design is the work of [120]. They provide an algorithm that attains optimal dynamic regret when the losses are exp-concave. However, their setting works only in a very restrictive setup where the decision set is an L_∞ constrained box. Consequently, we cannot directly apply their results to the linear regression problem of Section 10.2 whenever the decision set $\mathcal{D}$ is a general convex set.

An online learner is termed proper if the decisions of the learner are guaranteed to lie within the feasibility set $\mathcal{D}$. Otherwise it is called improper. A recent seminal work of [79] proposes neat reductions that can convert an improper online learner to a proper one, whenever the losses are convex. Following this line of research, we can aim to convert the algorithm of [120] that works exclusively on box decision set to one that can support arbitrary convex decision sets by coming up with suitable reduction schemes. However, the specific reduction scheme proposed in [79] is inadequate to yield fast dynamic rates for exp-concave losses. Our algorithm ProDR.control (Fig.10.1, **Pro**per **D**ynamic **R**egret.control) is a by-product of constructing new reduction schemes to circumvent the aforementioned problem for the case of linear regression losses. We expand upon these details below.

In ProDR.control, we maintain a surrogate algorithm $\mathcal{A}$, which is chosen to be the algorithm of [120] that produces iterates w_t in an L_∞ norm ball (box), $\mathcal{D}_\infty$, that encloses the actual decision set $\mathcal{D}$. Since w_t can be infeasible, we play $\hat{w}_t$ obtained via a special type of projection of w_t onto $\mathcal{D}$ which is formulated as a min-max problem in Line 3 of Fig.10.1. In Line 4, we construct surrogate losses ℓ_t to be passed to the algorithm $\mathcal{A}$. The surrogate loss penalises $\mathcal{A}$ for making predictions outside $\mathcal{D}$. We will show (see Lemma 245 in Appendix) that the instantaneous regret satisfies $f_t(\hat{w}_t) - f_t(u_t) \leq \ell_t(w_t) - \ell_t(u_t)$, where $u_t \in \mathcal{D}$ is the comparator at round t. Thus the dynamic regret of the proper iterates $\hat{w}_t$ wrt linear regression losses is upper bounded by the dynamic regret of the surrogate algorithm $\mathcal{A}$ on the losses ℓ_t and box decision set.

The design of the min-max barrier $S_t(w)$ is driven to ensure exp-concavity of the surrogate losses $\ell_t(w) = f_t(w) + G \cdot S_t(w)$. We capture its intuition as follows. We start by observing that since $\nabla^2 f_t(w) = 2A_t^T A_t$, the linear regression losses f_t exhibits strong curvature along the row-space of A_t, denoted by row(A_t). Further we have $\nabla f_t(w) = 2A_t^T(A_t w - b_t) \in \text{row}(A_t)$. So the loss f_t exhibits strong curvature along the direction of its gradient too. This is the fundamental reason behind the exp-concavity of f_t. The min-max barrier $S_t(w)$ is designed such that its gradient is guaranteed to lie in the row(A_t) (see Lemma 246 in Appendix for a formal statement). So the overall gradient $\nabla \ell_t(w)$ also lies in the row(A_t). Since the function f_t already exhibits strong curvature along row(A_t), we conclude that the sum $\ell_t(w) = f_t(w) + G \cdot S_t(w)$ exhibits strong curvature along its gradient $\nabla \ell_t(w)$. This maintains the exp-concavity of the losses ℓ_t over $\mathcal{D}_\infty$ (see Lemma 247 in Appendix). Such curvature considerations along with the fact that $S_t(w)$ has to be sufficiently large to facilitate the instantaneous regret bound $f_t(\hat{w}_t) - f_t(u_t) \leq \ell_t(w_t) - \ell_t(u_t)$ results in functional form for $S_t(w)$ displayed in Fig.10.1.

Consequently the fast dynamic regret rates derived in [120] becomes directly applicable. The reduction scheme used by [79] for producing proper iterates $\hat{w}_t$ and their accompanying surrogate loss design ℓ_t also allows one to upper bound the regret wrt linear regression losses f_t by the regret of the algorithm $\mathcal{A}$ wrt surrogate losses ℓ_t. However, the surrogate loss ℓ_t they construct is not guaranteed to be exp-concave and consequently not amenable to fast dynamic regret rates.

10.2.3 Main Results

We have the following guarantee for ProDR.control:

Theorem 104. *Let $u_{1:n} \in \mathcal{D}$ be any comparator sequence. In Fig.10.1, choose G such that $\sup_{w_1,w_2 \in \mathcal{D}_\infty(\tilde{R}), t\in[n]} \|A_t(w_1+w_2) - 2b_t\|_1 \le G$. Let α be as in Assumption 2. Let L be such that $\sup_{w\in\mathcal{D}_\infty(\tilde{R}), j\in[p]} 2\|A_t w - b_t\|_2^2 + 2G^2 \le L$ for all $t \in [n]$. Choose $\mathcal{A}$ as the algorithm from [120] with parameters $\gamma = 2G\alpha\tilde{R}\sqrt{d/8L} + \sqrt{2L}$ and $\zeta = \min\{\frac{1}{16G\alpha\tilde{R}\sqrt{d}}, 1/(4\gamma^2)\}$ and decision set $\mathcal{D}_\infty(\tilde{R})$. Under Assumptions 1 and 2, a valid of assignment of G and L are $2p\chi + 2\sigma$ and $6(p\chi+\sigma)^2$ respectively.*

Then the algorithm ProDR.control yields a dynamic regret rate of

$$\sum_{t=1}^{n} f_t(\hat{w}_t) - f_t(u_t) = \tilde{O}(d^3 n^{1/3}[\mathcal{TV}(u_{1:n})]^{2/3} \vee 1), \tag{10.11}$$

where $(a \vee b) := \max\{a, b\}$.

Remark 105. In view of Proposition 10 in [65], the dynamic regret guarantee in Theorem 104 is optimal modulo dependencies in d and $\log n$. Further the algorithm does not require apriori knowledge of the path length $\mathcal{TV}(u_{1:n})$.

Proof sketch for Theorem 104. First step is to show that $f_t(\hat{w}_t) \le \ell_t(w_t)$. This is accomplished by Lipschitzness type arguments. For any $u \in \mathcal{D}$, one observes that $\ell_t(u) = f_t(u)$. So the instantaneous regret of ProDR.control, $f_t(\hat{w}_t) - f_t(u_t)$, is upper bounded by the instantaneous regret, $\ell_t(w_t) - \ell_t(u_t)$ of the surrogate algorithm $\mathcal{A}$. The crucial step is to show the exp-concavity of the losses ℓ_t across $\mathcal{D}_\infty(\tilde{R})$. For this, we prove that there is a sub-gradient $\nabla S_t(w)$ that is aligned with $a_{t,j}$ for some $j \in [p]$. This observation followed by few algebraic manipulations (see proof of Lemma 247 in Appendix) allows us to show the exp-concavity of ℓ_t over $\mathcal{D}_\infty(\tilde{R})$. Now the overall regret can be controlled if the surrogate algorithm $\mathcal{A}$ provides optimal dynamic regret under exp-concave losses and box decision sets, $\mathcal{D}_\infty(\tilde{R})$. This is accomplished by choosing $\mathcal{A}$ as the algorithm in [120] which is also strongly adaptive. □

Since the surrogate algorithm $\mathcal{A}$ we used in Theorem 104 is strongly adaptive, we also have the following performance guarantee in terms of static regret:

Proposition 106. *Consider the instantiation of ProDR.control in Theorem 104. Then for any time window* $[a, b] \subseteq [n]$ *we have that:* $\sum_{t=a}^{b} f_t(\hat{w}_t) - \inf_{u \in \mathcal{D}} \sum_{t=a}^{b} f_t(u) = \tilde{O}(d^{1.5} \log n)$.

Remark 107. Theorem 104 and Proposition 106 together makes the algorithm ProDR.control a good candidate for performing proper online linear regression in non-stationary environments.

10.2.4 Linear regression with delayed feedback

In this section, we consider a linear regression protocol with feedback delayed by τ time steps.

- At round t, nature reveals a co-variate matrix $A_t \in \mathbb{R}^{p \times d}$.
- Learner plays $z_t \in \mathcal{D} \subset \mathbb{R}^d$.
- Nature reveals the loss $f_{t-\tau+1}(z) = \|A_{t-\tau+1} z - b_{t-\tau+1}\|_2^2$.

This delayed setting can be handled by the framework developed in [131]. Although these authors focus on bounding the regret as a function of time horizon n, the extension to dynamic regret bounds expressed in terms of both n and $\mathcal{TV}(u_{1:n})$ can be handled straight-forwardly in the analysis. We include the analysis in Appendix I.1 for the sake of completeness. The entire algorithm is as shown in Fig.10.2.

ProDR.control.delayed: Inputs- delay $\tau > 0$

- Maintain τ separate instances of ProDR.control (Fig.10.1). Enumerate them by $0, 1, \ldots, \tau - 1$.
- At time t:
  1. Update instance $(t-1) \mod \tau$ with loss $f_{t-\tau}$.
  2. Predict using instance $(t-1) \mod \tau$.

Figure 10.2: ProDR.control.delayed: An instance of delayed to non-delayed reduction from [131]

We have the following regret guarantee for Algorithm ProDR.control.delayed.

Theorem 108. *Let x_t be the prediction of the algorithm in Fig. 10.2 at time t. Instantiating each ProDR.control instance by the parameter setting described in Theorem 104. Let τ be the feedback delay. We have that*

$$\sum_{t=1}^{n} f_t(x_t) - f_t(u_t) = \tilde{O}(d^3 \tau^{2/3} n^{1/3} [\mathcal{TV}(u_{1:n})]^{2/3} \vee \tau). \tag{10.12}$$

Further for any interval $[a, b] \subseteq [n]$:

$$\sum_{t=a}^{b} f_t(x_t) - f_t(u) = O(d^{1.5} \tau \log n). \tag{10.13}$$

10.3 Instantiation for the LQR problem

In view of Proposition 101, the LQR problem is reduced to a mini-batch linear regression problem with delayed feedback, where the delay is given by $h = O(\log n)$ in Proposition 101. In this section, we provide explicit form of the linear regression losses arising in the LQR problem and instantiate Algorithm ProDR.control.delayed (Fig.10.2). First we need to define certain quantities:

For a sequence of matrices $(M^{[i]})_{i=1}^m$ define $\texttt{flatten}((M^{[i]})_{i=1}^m)$ as follows: Let $M_k^{[i]}$ be the k^{th} column of $M^{[i]}$.

Let's define

$$z^k = \begin{bmatrix} M_1^k \\ \vdots \\ M_{d_x}^k \end{bmatrix} \in \mathbb{R}^{d_u d_x}, \tag{10.14}$$

and

$$\texttt{flatten}((M^{[i]})_{i=1}^m) := \begin{bmatrix} z^1 \\ \vdots \\ z^m \end{bmatrix} \in \mathbb{R}^{m d_u d_x}. \tag{10.15}$$

For a sequence of DAP parameters $M_{1:n}$, let $\mathcal{TV}(M_{1:n}) := \sum_{t=2}^n \sum_{i=1}^m \|M_t^{[i]} - M_{t-1}^{[i]}\|_1$. We define `deflatten` as the natural inverse operation of `flatten`. We have the following Corollary of Theorem 108 and Proposition 101.

Corollary 109. *Assume the notations in Fig.10.1 and Section 10.1. Let $\Sigma_\infty = U_\infty^T \Lambda_\infty U_\infty$ be the spectral decomposition of the positive semi definite (PSD) matrix $\Sigma_\infty \in \mathbb{R}^{d_u \times d_u}$. . Let the covariate matrix $A_t := [w_{t-1}^T \dots w_{t-m}^T] \otimes \Lambda_\infty^{1/2} U_\infty \in \mathbb{R}^{d_u \times m d_u d_x}$, where $\otimes$ denotes the Kronecker product. Let the bias vector $b_t := \Lambda_\infty^{1/2} U_\infty q_{\infty;h}^*(w_{t:t+h})$. Let the delay factor* *of ProDR.control.delayed (Fig.10.2) be $\tau = h$ as defined in Proposition 101 and let the decision set given to the ProDR.control instances in Fig.10.2 be the DAP space defined in Eq.*(10.3)*. Let z_t be the prediction at round t made by the ProDR.control.delayed algorithm and let $M_t^{alg} :=$ **deflatten**(z_t). At round t, we play the control signal $u_t^{alg}(x_t) = \pi_t^{M_t^{alg}}(x_t)$ according to Eq.*(10.2)*. There exists a choice of input parameters for the ProDR.control instances in Fig.10.2 such that*

$$R(M_{1:n}) = \sum_{t=1}^{n} \ell(x_t^{alg}, u_t^{alg}) - \ell(x_t^{M_{1:n}}, u_t^{M_{1:n}}) \tag{10.16}$$

$$= \tilde{O}\left(m^3 d^4 d_x^5 (d_u \wedge d_x)(n^{1/3}[\mathcal{TV}(M_{1:n})]^{2/3} \vee 1)\right), \tag{10.17}$$

where $M_{1:n}$ is a sequence of DAP policies where each $M_t \in \mathcal{M}$ (eq.(10.3)*). Further the algorithm ProDR.control.delayed also enjoys a strongly adaptive regret guarantee for any interval $[a, b] \subseteq [n]$:*

$$\sum_{t=a}^{b} \ell(x_t^{alg}, u_t^{alg}) - \ell(x_t^{M}, u_t^{M}) = \tilde{O}((m d_u d_x)^{1.5} \log n), \tag{10.18}$$

for any fixed DAP policy $M \in \mathcal{M}$.

The following theorem provides a nearly matching lower bound.

Theorem 110. *There exists an LQR system, a choice of the perturbations w_t and a DAP policy class such that:*

$$\sup_{M_{1:n} \text{ with } \mathcal{TV}(M_{1:n}) \leq C_n} \mathbb{E}[R(M_{1:n})] = \Omega(n^{1/3} C_n^{2/3} \vee 1), \tag{10.19}$$

where the expectation is taken wrt randomness in the strategies of the agent and adversary.

The proof of the above lower bound given in Appendix I.1 is interesting in its own right. The proof is also applicable to the problem of online non-parametric regression against Total Variation (TV) bounded sequences [29, 65]. The lower bounding strategy in [29] goes through arguments based on sequential Rademacher complexity of the non-parametric class of TV bounded sequences. While they establish the rate wrt n as $n^{1/3}$, the correct dependency on the TV of the sequence was not provided in [29]. The work of [65] ameliorated this issue by appealing to the standard lowerbounds from *offline* non-parametric regression literature. This lower bounding route uses fairly sophisticated arguments based on characterizing the Bernstein width of the set of Haar wavelet coefficients of TV bounded sequences [20]. In contrast, we provide a lower bound capturing the correct rate wrt both n and TV of the sequence via more direct arguments based on constructing an explicit adversarial strategy. An elaborate outline of applying our lower bound to the online non-parametric regression framework is given in Appendix I.1.

Remark 111. The covariate matrix $A_t \in \mathbb{R}^{d_u \times m d_u d_x}$ that arise in Corollary 109 is rank deficient whenever $md_x > 1$. In such cases, the linear regression losses $f_t(w)$ as in Fig.10.1 cannot be strongly convex. So the proper universal dynamic regret minimizing algorithm for strongly convex losses from [120] is inapplicable in general except potentially for the particular setting of $m = d_x = 1$. Moreover, in the setting of $m = d_x = 1$ a non-zero strong convexity parameter can exist only if the magnitude of the perturbations $|w_t|$ are bounded away from zero which is restrictive in its scope.

10.4 Concluding Discussion

In this section, we designed a new algorithm for online non-stochastic LQR controller. The controller provably minimizes the regret with any oracle non-stationary sequence of Disturbance Action controllers chosen to handle the sequence of adversarially-chosen disturbances after they realized. We also show that no other algorithm is able to have smaller max-regret by more than a logarithmic factor, i.e., our proposed algorithm is optimal. The underlying algorithm is a new development in minimizing dynamic regret in non-stationary (minibatched) linear regression problem under the *proper* learning setup. The techniques developed in this work can be of independent interest in the broader literature of online learning. Future work include generalizing the family of loss functions to general strongly convex losses and exp-concave losses.

Part IV

Appendix

Appendix A

Supplementary Materials for Chapter 2

A.1 Additional Experiments

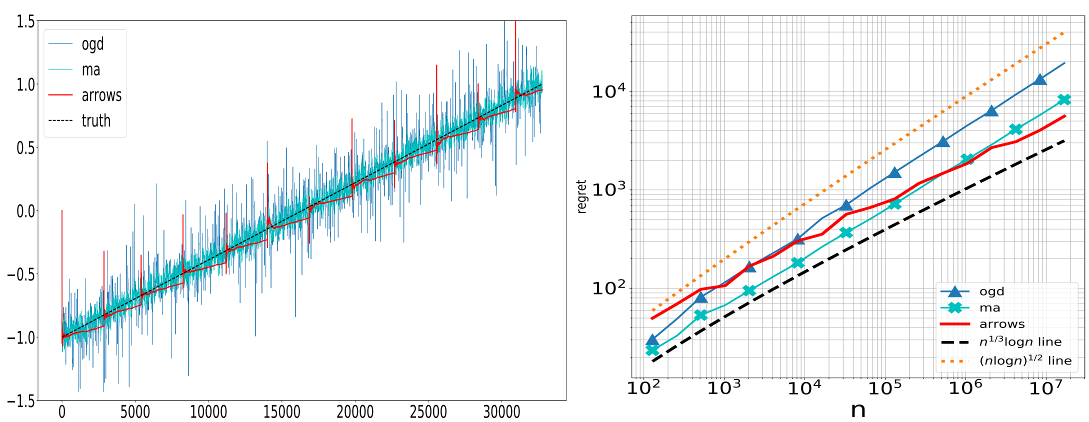


Figure A.1: *An illustration of* Arrows *on a linear trend which has homogeneous smoothness*

The function that we generated in Figure 2.2 is a hybrid function which in the first half is a "discretized cubic spline" with more knots closely placed towards the end. In the second half it is a Doppler function $f(t) = \sin\left(\frac{2\pi(1+\epsilon)}{t/n+0.38}\right)$ with n being the time horizon. We observe noisy data $y_i = f(i/n) + z_i$, $i = 1, ..., n$ and z_i are iid normal variables with $\sigma = 1$. The value of C_n for $n > 60K$ is around 17. Hence for all $n > 83521$, we are under the $n^{1/3}$ regime of $\sigma\sqrt{\log(n)/n} < C_n < \sigma n^{1/4}$.

The window size for moving averages and partition width of OGD were tuned optimally for the TV class (see Appendix C for details). Figure 2.2 depicts the estimated signals and dynamic regret averaged across 5 runs in a log log plot. The left panel illustrates that Arrows is locally adaptive to heterogeneous smoothness of the ground truth.

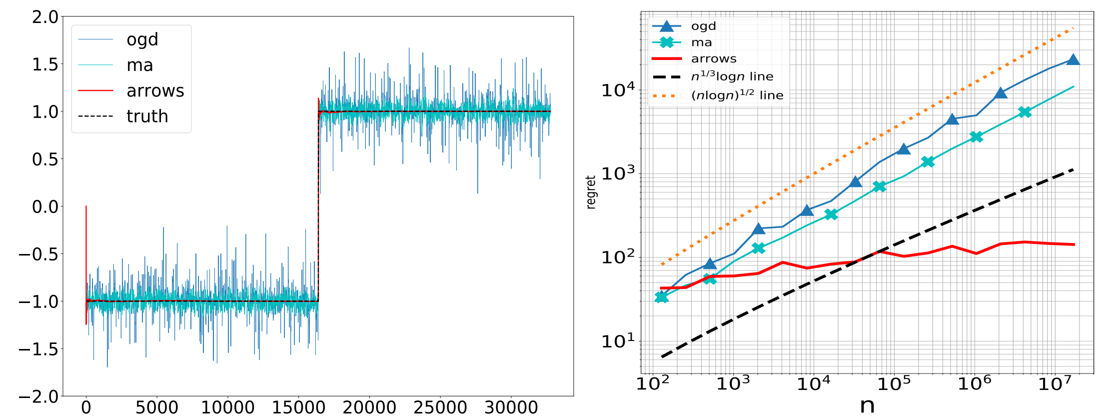


Figure A.2: *An illustration of* Arrows *on a step trend with abrupt inhomogeneity.*

Red peaks in the figure signifies restarts. During the initial and final duration, the signal varies smoothly and Arrows chooses a larger window size for online averaging. In the middle, signal varies rather abruptly. Consequently Arrows chooses a smaller window size. On the other hand, the linear smoothers OGD and MA attains a suboptimal $\tilde{O}(\sqrt{n})$ regret.

In Figure A.1 and A.2 we plot the estimates and log-log regret for two more functions: A linear function that is homogeneously smooth and less challenging and a step function which has an abrupt discontinuity making it more inhomogeneous than linear but have lesser inhomogeneity w.r.t hybrid signal considered in 2.3.7. Both OGD and MA were optimally tuned for the TV class as in Appendix A.2.

The red peaks corresponds to restarts by Arrows. For linear functions we can see that ARROWS chooses inter-restart duration/bin-widths that are constant throughout. This is expected as a linear trend is spatially homogeneous. For the step function, we see that Arrows restart only once since the start. Further, notice that it quickly restarts once the bump is hit. For both of these functions, necessary scaling is done so that we are in the $n^{1/3}$ regime quite early.

A.2 Upper bounds of linear forecasters

In this section we compute the optimal batch size for Restarting OGD and optimal window size for moving averages to yield the $\tilde{O}(\sqrt{n})$ regret rate.

Theorem 112. *Let the feedback be $y_t = \theta_t + Z_t$ where Z_t is an independent, σ-subgaussian random variable. Let $\theta_{1:n} \in \mathrm{TV}(C_n)$. Restarting OGD with batch size of $\sqrt{n \log n}\frac{\sigma}{C_n}$ achieves an expected dynamic regret of $\tilde{O}(U^2 + C_n^2 + \sigma C_n \sqrt{n})$.*

Proof. Note that in our setting with squared error losses $f_t(x) = (x - \theta_t)^2$, the update rule of restarting OGD reduces to computing online averages. Thus OGD essentially divides the time horizon n into fixed size batches and output online averages within each batch. Our objective here is to compute the optimal batch size that minimizes the dynamic regret.

We will bound the expected regret. Let x_t be the estimate of OGD at time t. Let batches be numbered as $1, ..., \lceil n/L \rceil$ where L is the fixed batch size. Let the total variation of ground truth within batch i be C_i. Time interval of batch i is denoted by $[t_h^{(i)}, t_l^{(i)}]$. Due to bias variance decomposition within a batch we have,

$$R_i = \sum_{t=t_h^{(i)}}^{t_l^{(i)}} E[(x_t - \theta_t)^2] = (\theta_{t_h^{(i)}-1} - \theta_{t_h^{(i)}})^2 + \sum_{t=t_h^{(i)}+1}^{t_l^{(i)}} (\theta_t - \bar{\theta}_{t_h^{(i)}:t-1})^2 + \frac{\sigma^2}{t - t_h^{(i)}}, \tag{A.1}$$

$$\leq (\theta_{t_h^{(i)}-1} - \theta_{t_h^{(i)}})^2 + LC_i^2 + \sigma^2(2 + \log L), \tag{A.2}$$

with the convention $\theta_0 = 0$ and at start of bin our prediction is just the noisy realization of the previous data point.

Summing across all bins gives,

$$\sum_{i=1}^{\lceil n/L \rceil} R_i \leq LC_n^2 + 2\sigma^2 \frac{n(2 + \log L)}{L} + U^2 + C_n^2. \tag{A.3}$$

where we have used assumption (A4) to bound the bias of the first prediction. The above expression can be minimized by setting $L = \sqrt{n \log n} \frac{\sigma}{C_n}$ to yield a regret bound of $O(U^2 + C^2 + \sigma C_n \sqrt{n \log n})$ □

Theorem 113. *Under the same setup as in Theorem 112, moving averages with window size* $\frac{\sigma\sqrt{n}}{C_n}$ *yields a dynamic regret of* $O(\sigma C_n \sqrt{n} + U^2 + C_n^2)$

Proof. Let the window size of moving averages be denoted by m. Consider the prediction at a time $x_t, t \geq m$. By bias variance decomposition we have,

$$E[(x_t - \theta_t)^2] = \left(\theta_i - \frac{\sum_{j=i-m}^{i-1} \theta_j}{m}\right)^2 + \frac{\sigma^2}{m}. \tag{A.4}$$

By Jensen's inequality,

$$\left(\theta_i - \frac{\sum_{j=i-m}^{i-1} \theta_j}{m}\right)^2 \leq \frac{\sum_{j=i-m}^{i-1} (\theta_j - \theta_i)^2}{m}, \tag{A.5}$$

$$\leq \frac{2\sum_{j=i-m}^{i-1} (j - i + 1 + m)(\theta_{j+1} - \theta_j)^2}{m} \text{ ,by } (a+b)^2 \leq 2a^2 + 2b^2. \tag{A.6}$$

Notice that the term $(\theta_i - \theta_{i-1})^2$ will be multiplied by a factor m in the above bias bound at time point i, $m-1$ times in the next time point $i+1$ and so on. By summing this bias bound across the times points, we obtain

$$\sum_{i=m}^{n} \frac{2\sum_{j=i-m}^{i-1}(j-i+1+m)(\theta_{j+1}-\theta_j)^2}{m} \leq 4m\sum_{i=1}^{n-1}(\theta_i-\theta_{i+1})^2 + U^2, \tag{A.7}$$

$$\leq 4mC_n^2 + U^2. \tag{A.8}$$

The squared bias for the initial points can be bounded by.

$$\sum_{i=1}^{m-1}(\theta_i - \hat{\theta}_{(1:i-1)})^2 \leq U^2 + C_n^2. \tag{A.9}$$

Summing the variance terms yields,

$$\sum_{t=1}^{n} \mathrm{Var}(x_t) = \sum_{t=1}^{m-1}\frac{\sigma^2}{t} + \sum_{t=m}^{n}\frac{\sigma^2}{m}, \tag{A.10}$$

$$\leq \frac{(1+\log m + n)\sigma^2}{m}. \tag{A.11}$$

Thus the total MSE can be minimized by setting $m = \frac{\sigma\sqrt{n}}{C_n}$, we obtain a dynamic regret bound of $O(\sigma C_n\sqrt{n} + U^2 + C_n^2)$

□

A.3 Proof of useful lemmas

We begin by recording an observation that follows directly from the policy.

Lemma 114. *For m^{th} bin that spans the interval $[t_h^{(m)}, t_l^{(m)}]$, discovered by the policy, let the lengths of $\hat{\alpha}(t_h^{(m)} : t_l^{(m)} - 1)$ and $\hat{\alpha}(t_h^{(m)} : t_l^{(m)})$ be k and k^+ respectively. Then $\sum_{l=0}^{\log_2(k)-1} 2^{l/2}\|\hat{\alpha}(t_h^{(m)} : t_l^{(m)} - 1)[l]\|_1 \leq \sigma$ and $\sum_{l=0}^{\log_2(k^+)-1} 2^{l/2}\|\hat{\alpha}(t_h^{(m)} : t_l^{(m)})[l]\|_1 > \sigma$*

Next we prove the marginal sub-gaussianity of the wavelet coefficients.

Lemma 115. *Consider the observation model $y_i = \theta_i + \sigma z_i$, where z_i is iid sub-gaussian with parameter 1, $i = 1,..,n$. Let α_i denote the wavelet coefficients of the sequence $z = pad_0(y_1, ...y_n)$. Then each α_i is sub-gaussian with parameter 2σ.*

Proof. Without loss of generaility let's charecterize α_1. Let $\boldsymbol{u} = [u_1, ...u_n, u_{n+1}, \ldots, u_{|z|}]^T$ denote the first row of the orthonormal wavelet transform matrix. Then,

$$\alpha_1 = \sum_{i=1}^{n} y_i \left(u_i(1-\frac{1}{n}) - \sum_{j=1, j\neq i}^{n} \frac{u_j}{n} \right). \tag{A.12}$$

Thus α_1 is a differentiable function of iid sub-gaussian noise z_i. We can find its Lipschtiz constant by bounding the gradient w.r.t z_i as follows,

$$\|\nabla \alpha_1(z_1, ..., z_n)\|_2 \leq \sigma \left(\sum_{i=1}^{n} 2u_i^2 (1 - \frac{1}{n})^2 + \frac{2}{n} \sum_{j=1, j\neq i}^{n} u_j^2 \right)^{\frac{1}{2}}, \quad \text{(A.13)}$$

$$\leq \sigma \left(2+2\right)^{\frac{1}{2}}, \quad \text{(A.14)}$$

$$= 2\sigma. \quad \text{(A.15)}$$

By proposition 2.12 in [18] we conclude that α_1 sub-gaussian with parameter 2σ. □

In the next lemma, we record the uniform shrinkage property of soft-thresholding estimator.

Lemma 116. *For any interval* $[t_h, t_l]$*, let* $Y = pad_0(y_{t_h}, ..., y_{t_l})$ *and* $\Theta = pad_0(\theta_{t_h}, ..., \theta_{t_l})$*. Then* $|(T(HY))_i| \leq |(H\Theta)_i|$ *with probability at-least* $1 - 2n^{3-\beta/8}$ *for each co-ordinate* i*.*

Proof. Consider a fixed bin $[\underline{l}, \bar{l}]$ with zero padded vector $Y \in R^k$. Due to sub-gaussian tail inequality, we have $|(HY)_i - (H\Theta)_i| \leq \sigma\sqrt{\beta \log(n)}$ with probability at-least $1 - 2/n^{\beta/8}$. Consider the case $(H\Theta)_i \geq \sigma\sqrt{\beta \log(n)}$. Then both the scenarios $(HY)_i \leq \sigma\sqrt{\beta \log(n)}$ and $(HY)_i > \sigma\sqrt{\beta \log(n)}$ leads to shrinkage to a value that is smaller than $|(H\Theta)_i|$ in magnitude due to soft-thresholding with threshold set to $\sigma\sqrt{\beta \log(n)}$. Now consider the case when $0 \leq (H\Theta)_i \leq \sigma\sqrt{\beta \log(n)}$. Again, soft-thresholding in both scenarios $(HY)_i \leq \sigma\sqrt{\beta \log(n)}$ and $\sigma\sqrt{\beta \log(n)} \leq (HY)_i \leq (H\Theta)_i + \sigma\sqrt{\beta \log(n)}$ leads to shrinkage to a value that is smaller than $|(H\Theta)_i|$ in magnitude. One can come up with a similar argument for the case where $(H\Theta)_i \leq 0$. Now applying a union bound across all $O(n)$ co-ordinates and all $O(n^2)$ bins, we get the statement of the lemma. □

Lemma 117. *The number of bins,* M*, discovered by the policy is at-most* $\max\{1, 2n^{1/3} C_n^{2/3} \sigma^{-2/3} \log(n)\}$ *with probability at-least* $1 - 2n^{3-\beta/2}$*.*

Proof. Let $\Theta_m = [\theta_1^{(m)}, \theta_2^{(m)}, ..., \theta_p^{(m)}]^T$ be the mean subtracted and zero padded ground truth sequence values in m^{th} bin $[\underline{l}, \bar{l}]$ discovered by our policy. $y^{(m)} = [y_1^{(m)}, y_2^{(m)}, ..., y_p^{(m)}]^T$ be the corresponding mean subtracted and zero padded observations. Note that due to zero padding $p \leq 2(\bar{l} - \underline{l})$ and some of the last values in the vector can be zeroes. Let $\alpha_m(\underline{l} : \bar{l}) = H\Theta$ denotes the discrete wavelet coefficient vector. We can view the computation of the Haar coefficients as a recursion. At each level l of the recursion, the entire length p, is divided into 2^l intervals. Let the sample averages of elements of Θ_m in these intervals be denoted by the sequence $\{\tilde{\theta}_1, \tilde{\theta}_2, ..., \tilde{\theta}_{2^l}\}$. Let $\alpha_m^{(l)} \in \mathbb{R}^{2^l}$ denotes the vector of Haar coefficients at level l.

First note that the Haar coefficient $\alpha_m^{(l)}(i) = \frac{1}{2}\sqrt{\frac{p}{2^l}}(\tilde{\theta}_{2i} - \tilde{\theta}_{2i-1})$ with $i = 1, ..., 2^l$.

$$\|\alpha_m^{(l)}\|_1^2 \leq \frac{p}{2^{l+2}} \left(\sum_{i=1}^{2^l} |\tilde{\theta}_{2i} - \tilde{\theta}_{2i-1}| \right)^2, \tag{A.16}$$

$$\leq \frac{pTV^2[\underline{l} - 1 : \bar{l}]}{2^l}, \tag{A.17}$$

where $TV[a, b]$ denotes the total variation of the true sequence in the interval $[a, b]$. The last inequality holds because the total variation of the smoothed sequence must be at-most four times the entire total variation of true sequence. The factor 4 is due to the fact that total variation when we pad a mean zero sequence with zeroes is at-most twice the total variation before zero padding.

We have,

$$\frac{1}{\sqrt{p}} \sum_{l=0}^{\log_2(p)-1} 2^{l/2} \|\alpha_m^{(l)}\|_1 \leq \log p \, TV[\underline{l} - 1 : \bar{l}]. \tag{A.18}$$

In the policy we compute $\hat{\alpha}_m(\underline{l} : \bar{l}) = T(Hy^{(m)})$ with the soft thresholding factor of $\sigma\sqrt{\beta \log(n)}$. From lemma 116, we have $|(T(Y))_i| \leq |(H\Theta)_i| \, \forall i \in [1, p]$ with probability at-least $1 - 2n^{3-\beta/8}$. Since $[\underline{l}, \bar{l}]$ is a bin discovered by policy, lemma 114 gives a lowerbound on $\|\alpha_m(\underline{l} : \bar{l})\|$. Putting it all together yields the relation,

$$\frac{\sigma}{\sqrt{p}} < \frac{1}{\sqrt{p}} \sum_{l=0}^{\log_2(p)-1} 2^{l/2} \|\hat{\alpha}_m^{(l)}(\underline{l} : \bar{l})\|_1 \leq \frac{1}{\sqrt{p}} \sum_{l=0}^{\log_2(p)-1} 2^{l/2} \|\alpha_m^{(l)}(\underline{l} : \bar{l})\|_1 \leq \log(p) \, TV[\underline{l} - 1 : \bar{l}], \tag{A.19}$$

with probability at-least $1 - 2n^{3-\beta/8}$.

Thus the total variation in the time interval $[\underline{l} - 1, \bar{l}]$ can be lower bounded in probability as

$$TV[\underline{l} - 1 : \bar{l}] > \frac{\sigma}{\sqrt{p} \log n}. \tag{A.20}$$

Due to assumption $(A3)$ we have,

$$\sum_{i=1}^{M} TV[\underline{l}^i - 1 : \bar{l}^i] = C_n, \tag{A.21}$$

where $[\underline{l}^i : \bar{l}^i]$ are the bins discovered by the policy.

Let p_i be the padded width of bin i discovered by the policy. Then,

$$C_n \log n \geq \sum_{i=1}^{M} \frac{\sigma}{\sqrt{p_i}}, \tag{A.22}$$

$$\geq \frac{M^2\sigma}{\sum_{i=1}^{M}\sqrt{p_i}}, \tag{A.23}$$

where the last line is obtained via Jensen's inequality. Now using Holder's inequality $\|x\|_\beta \leq d^{\frac{1}{\beta}-\frac{1}{\alpha}}\|x\|_\alpha$ for $0 < \beta \leq \alpha$, $x \in \mathbb{R}^d$ with $\alpha = 1/2$, $\beta = 1$ and noting that $\sum_{i=1}^{M} p_i \leq 2T$ gives,

$$\sigma M^2 \leq C_n \log n \sum_{i=1}^{M} \sqrt{p_i}, \tag{A.24}$$

$$\leq C_n \log n \sqrt{Mn}. \tag{A.25}$$

Hence we get $M \leq (2n)^{1/3}(C_n \log n)^{2/3}\sigma^{-2/3} \leq 2n^{1/3}C_n^{2/3}\sigma^{-2/3}\log(n)$.

When $C_n = 0$, (A.19) implies that our policy will not restart with probability at-least $1 - 2n^{3-\beta/8}$ making $M = 1$. □

We restate two useful results from [31].

Lemma 118. *Consider the observation model $y = \alpha + Z$, where $y \in R^k$ and $|Z_i| \leq \delta \forall i \in [1, k]$. Let $\hat{\alpha}_\delta$ be the soft thresholding estimator with input y and threshold δ, then*

$$\|\hat{\alpha}_\delta - \alpha\|^2 \leq \sum_{i=1}^{k} min\{\alpha_i^2, 4\delta^2\}. \tag{A.26}$$

Lemma 119. *Consider the observation model $y = \alpha + Z$, where $y \in R^k$, $\alpha \in A$ and each Z_i is sub-gaussian with parameter σ^2. If A is solid and orthosymmetric, then*

$$\inf_{\hat{\alpha}} \sup_{\alpha \in A} E[\|\hat{\alpha} - \alpha\|^2] \geq \frac{1}{2.22} \sup_{A} \sum_{i=1}^{k} min\{\alpha_i^2, \sigma^2\}. \tag{A.27}$$

Let's pause a moment to ponder how remarkable the above lemma is. The observations need not be even iid. Given A is solid and orthosymmetric, all that is required is the marginal sub-gaussianity as the soft-thresholding operation works co-ordinate wise. Now we reprove theorem 4.2 from [31] with a slight modification of threshold value in the estimator.

Theorem 120. *With probability at-least $1 - 2n^{-\beta/2}$, under the model in lemma 119, the* *soft thresholding estimator* $\hat{\alpha}_\delta$ *with* $\delta = \sigma\sqrt{\beta \log(n)}$ *obeys*

$$\|\hat{\alpha}_\delta - \alpha\|^2 \leq 8.88\beta(1+\log(n)) \inf_{\hat{\alpha}} \sup_{\alpha \in A} E[\|\hat{\alpha} - \alpha\|^2]. \tag{A.28}$$

Proof. Consider the soft thresholding estimator $\hat{\alpha}_\delta$. By Gaussian tail inequality we have $P(\sup_i |Z_i| \geq \delta) \leq 2n^{-\beta/2}$. Conditioning on the event $\sup_i |Z_i| \leq \delta$ and applying lemma 118,

$$\|\hat{\alpha}_\delta - \alpha\|^2 \leq \sum_{i=1}^{k} min\{\alpha_i^2, 4\delta^2\}, \tag{A.29}$$

$$= \sum_{i=1}^{k} min\{\alpha_i^2, 4\beta\sigma^2 \log(n)\}, \tag{A.30}$$

$$\leq max\{1, 4\beta \log(n)\} \sum_{i=1}^{k} min\{\alpha_i^2, \sigma^2\}, \tag{A.31}$$

$$\leq (1 + 4\beta \log(n)) \sup_{\alpha \in A} \sum_{i=1}^{k} min\{\alpha_i^2, \sigma^2\}, \tag{A.32}$$

$$\leq 4\beta(1+\log(n))\ 2.22\ \inf_{\hat{\alpha}} \sup_{\alpha \in A} E[\|\hat{\alpha} - \alpha\|^2], \tag{A.33}$$

where the last line follows from lemma 119. □

It can be shown that wavelet coefficients of functions residing in the TV class is solid and orthosymmetric. As shown in lemma 115, the noisy wavelet coefficients are marginally sub-gaussian. Thus in the coefficient space, we are under the same observation model as in lemma 119. Using a uniform bound argument across all $O(n^2)$ bins and all $O(n)$ points within a bin along with lemma 115 leads to the following corollary.

Corollary 121. *The soft-thresholded wavelet coefficients of re-centered and zero padded noisy data within any interval* $[t_h, t_l]$ *satisfy relation* (A.28) *with probability atleast* $1 - 2n^{3-\beta/8}$.

Next, we record an important preliminary bound that will be used in proving the main result.

Lemma 122. *With probability at-least* $1 - \frac{\delta}{2}$, *the total squared error for online averaging between two arbitrarily chosen time points* t_h *and* t_l *satisfies*

$$\sum_{t=t_h}^{t_l} (x_t^{t_h} - \theta_t)^2 \leq 4\sigma^2 \log(4n^3/\delta)(2 + \log(t_l - t_h + 1)) + 2(\theta_{t_h - 1} - \theta_{t_h})^2 + 2 \sum_{t=t_h+1}^{t_l} (\bar{\theta}_{t_h:t-1} - \theta_t)^2. \tag{A.34}$$

Proof. Throughout this lemma we assume the notation $\theta_0 = 0$. For proving this, first we bound the squared error for online sample averages within a bin, $b[\underline{l}, \bar{l}]$, that starts and ends at fixed times $\underline{l}$ and $\bar{l}$ respectively. Then a uniform bound argument will be used for bounding the squared error within any arbitrarily chosen bin. Note that $b[\underline{l}, \bar{l}]$ represents any fixed time interval and may not be even chosen by the policy. For $t \in [\underline{l}, \bar{l}]$, consider the prediction $x_t^{\underline{l}}$, with same notation as used in the policy. Define a random variable Z_t as

$$Z_t = \frac{(x_t^{\underline{l}} - \theta_t) - (\lambda_t - \theta_t)}{\sigma\sqrt{1/[t - \underline{l}]_{1+}}}, \tag{A.35}$$

where $[x]_{1+} = max\{1, x\}$, $\lambda_{\underline{l}} = \theta_{\underline{l}-1}$ and $\lambda_t = \bar{\theta}_{\underline{l}:t-1}, \forall t > \underline{l}$. Z_t is subgaussian with variance parameter 1 and mean 0. Hence by sub-gaussian tail inequality, we have $P(|Z_t| \geq \sqrt{2\log(4/\delta)}) \leq \delta/2$. By noting that length of a bin is $O(n)$ and applying uniform bound across all time points within the current bin we have

$$P\left(\sup_{\underline{l} \leq t \leq \bar{l}} |Z_t| \geq \sqrt{2\log(4n^3/\delta)}\right) \leq \delta/2n^2. \tag{A.36}$$

Hence with probability at-least $1 - \delta/2n^2$,

$$|x_t^{\underline{l}} - \theta_t| \leq |\lambda_t - \theta_t| + \sigma\sqrt{\frac{2\log(4n^3/\delta)}{[t - \underline{l}]_{1+}}}, \forall t \in [\underline{l}, \bar{l}]. \tag{A.37}$$

So the squared error within a bin can be bounded in probability as

$$\sum_{t=\underline{l}}^{\bar{l}} (x_t^{\underline{l}} - \theta_t)^2 \leq 2(\theta_{\underline{l}-1} - \theta_{\underline{l}})^2 + 2\sum_{t=\underline{l}+1}^{\bar{l}} (\bar{\theta}_{\underline{l}:t-1} - \theta_t)^2 + 2\sum_{t=\underline{l}}^{\bar{l}} \sigma^2 \frac{2\log(4n^3/\delta)}{[t - \underline{l}]_{1+}}. \tag{A.38}$$

Here we applied the inequality $(a+b)^2 \leq 2a^2 + 2b^2$ on (A.37). Ultimately we are interested in analyzing the MSE within a bin detected by the policy. However the observations within a bin satisfies the restarting criterion of the policy and cannot be regarded independent. To break free of this constraint, we uniformly bound the quantity of interest — MSE here — across all possible bins. Noting that number of bins is $O(n^2)$ and applying uniform bound across all bins gives the following single sided tail bound.

Let E denote the event: $\sup_{b[\underline{l}:\bar{l}]} (x_t^{\underline{l}} - \theta_t)^2 - 2(\theta_{\underline{l}-1} - \theta_{\underline{l}})^2 - 2\sum_{t=\underline{l}+1}^{\bar{l}} (\bar{\theta}_{\underline{l}:t-1} - \theta_t)^2 - 2\sum_{t=\underline{l}}^{\bar{l}} \sigma^2 \frac{2\log(4n^3/\delta)}{[t-\underline{l}]_{1+}} \geq 0$.

Then,

$$P(E) \leq \delta/2. \tag{A.39}$$

Hence with probability at-least $1 - \delta/2$, any bin $b[t_h : t_l]$ satisfies (A.34). □

Since (A.34) holds for any arbitrary interval of the time axis, it is particularly true for the bins discovered by the policy. Therefore the total squared error T of the policy is upper bounded in probability by the sum of bin bounds of the form,

$$T \leq \sum_{m=1}^{M} 4\sigma^2 \log(4n^3/\delta)(2+\log(t_l^{(m)} - t_h^{(m)} + 1)) + 2(\theta_{t_h^{(m)}-1} - \theta_{t_h^{(m)}})^2 + 2 \sum_{t=t_h^{(m)}+1}^{t_l^{(m)}} (\bar{\theta}_{t_h^{(m)}:t-1} - \theta_t)^2, \tag{A.40}$$

where the outer sum iterates over the bins and M is the total number of bins. The first term inside the outer summation can be controlled if we can upper bound M. Now we set out to prove our main theorem.

A.4 Proof of Theorem 1

From the discussion in section 2.1.1, the goal of bounding dynamic regret of the policy can be achieved by bounding the total squared error of its predictions. Our solution proceeds in two steps. First we upper bound the total squared error within a bin. Then we construct an upper bound for the number of bins spawned by the policy. With these two bounds in place, we bound the total squared error of the policy (A.40).

Let's first proceed to get a bound on the last summation term in (A.40). We use a reduction towards Follow The Leader (FTL) strategy. The term is basically the regret incurred by an FTL game with quadratic loss function for the duration $[t_h, t_l]$.

Let $\Theta(t_h : t_l - 1) = pad_0(\theta_{t_h}, ..., \theta_{t_l-1}) = [\Theta_{t_h}, ..., \Theta_{t_h+k-1}]^T$ denotes mean subtracted the zero padded true sequence in the interval $[t_h, t_l - 1]$. Then,

$$\sum_{t=t_h}^{t_l} (\bar{\theta}_{t_h:t-1} - \theta_t)^2 = (\bar{\theta}_{t_h:t_l-1} - \theta_{t_l})^2 + \sum_{t=t_h}^{t_l-1} (\bar{\theta}_{t_h:t-1} - \theta_t)^2, \tag{A.41}$$

$$\leq (\bar{\theta}_{t_h:t_l-1} - \theta_{t_l})^2 + \sum_{t=t_h}^{t_l-1} \frac{(\bar{\theta}_{t_h:t-1} - \theta_t)^2}{(t - t_h + 1)} + \sum_{t=t_h}^{t_l-1} (\bar{\theta}_{t_h:t_l-1} - \theta_t)^2, \tag{A.42}$$

$$= (\bar{\theta}_{t_h:t_l-1} - \theta_{t_l})^2 + \sum_{t=t_h}^{t_l-1} \frac{(\bar{\theta}_{t_h:t-1} - \theta_t)^2}{(t - t_h + 1)} + \|\Theta(t_h : t_l - 1)\|^2. \tag{A.43}$$

We have applied FTL reduction for online game of predicting the true sequence $\theta_{t_h}, ..., \theta_{t_l-1}$ to get (A.43).

In the discussion below we assume that $\|D\theta_{1:n}\|_1 \leq C_n$ and $|\theta_1| \leq U$.

Now let's try to bound the term $\|\Theta(t_h : t_l - 1)\|_2^2$. This is basically the regret of the best expert. By triangle inequality,

$$\|\Theta(t_h : t_l - 1)\|^2 \leq \|\hat{\alpha}(t_h : t_l - 1)\|_1^2 + \|\hat{\alpha}(t_h : t_l - 1) - \alpha(t_h : t_l - 1)\|_2^2, \tag{A.44}$$

$$\leq \left(\sum_{l=0}^{\log_2(p)-1} 2^{l/2} \|\hat{\alpha}(t_h : t_l - 1)[l]\|_1 \right)^2 + \|\hat{\alpha}(t_h : t_l - 1) - \alpha(t_h : t_l - 1)\|_2^2, \tag{A.45}$$

where p is the padded length.

We can base our online averaging restart rule on the output of wavelet smoother. Suppose we decide to restart when $\|\hat{\alpha}(t_h : t_l)\|_1 \geq Kn^{-1/3}$ for a constant K. Then the first term of (A.45) gives the optimal rate of $O(n^{1/3})$ when summed across all bins. But the estimation error term $\|\hat{\alpha}(t_h : t_l - 1) - \Theta(t_h : t_l - 1)\|^2$ should also be controlled. If the smoother is minimax over any bin $[t_h, t_l]$, then we can hope to get minimaxity over the entire horizon. However, the total variation inside the bin is not known to the smoother. This is where the adaptive minimaxity of wavelet smoother comes to rescue.

Suppose $\mathcal{F}$ denotes the class of functions f with total variation $TV(f) \leq C_n$. Let $\mathcal{A}$ denote the set of all coefficients of the continuous wavelet transform of functions $f \in \mathcal{F}$. Then $\mathcal{A} \subset \Theta^{1/2}_{1,\infty}(C_n)$, where $\Theta^{1/2}_{1,\infty}(C_n)$ is a Besov body as defined in [2]. The minimax rate of estimation in this Besov body is $O(n^{-2/3} C_n^{2/3} \sigma^{4/3})$ where n is the number of observations. However, this is the rate of convergence of the L_2 function norm instead of the discrete (input-averaged) norm that we consider here. Over the Besov spaces, these two norms are close enough that the rates do not change (see section 15.5 of [18]). Hence Corollary 121 can be used to control the bias.

Let $\hat{y}(t_h : t)$ denotes the soft-thresholding estimates of the vector $pad_0(y_{t_h:t})$. i.e $\hat{y}(t_h : t) = H^T T(H\, pad_0(y(t_h : t)))$.

$$(\bar{\theta}_{t_h:t_l-1} - \theta_{t_l})^2 \leq 2(\theta_{t_l-1} - \theta_{t_l})^2 + 2(\bar{\theta}_{t_h:t_l-1} - \theta_{t_l-1})^2, \tag{A.46}$$

$$\leq 2(\theta_{t_l-1} - \theta_{t_l})^2 + 4(\hat{y}(t_h : t_l - 1)[t_l - 1] - (\bar{\theta}_{t_h:t_l-1} - \theta_{t_l-1}))^2 + 4(\hat{y}(t_h : t_l - 1)[t_l - 1])^2. \tag{A.47}$$

Since L1 norm is greater than L2 norm, the policy's restart rule implies that

$$(\hat{y}(t_h : t_l - 1)[t_l - 1])^2 \leq \sigma^2 \tag{A.48}$$

Combining (A.47) and (A.48), we get

$$(\bar{\theta}_{t_h:t_l-1} - \theta_{t_l})^2 \leq 2(\theta_{t_l} - \theta_{t_l-1})^2 + \gamma_1 (t_l - t_h)^{1/3}\, TV^{2/3}[t_h : t_l]\, \sigma^{4/3} + \sigma^2, \tag{A.49}$$

where last line holds with probablity atleast $1 - 2n^{3-\beta/8}$ due to Corollary 121. Here γ_1 is a constant which can depend logarithmically on the width $t_l - t_h$.

Now let's bound the second term in (A.43). For any $t \in [t_h, t_l - 1]$ we have,

$$\sum_{t=t_h}^{t_l-1} \frac{(\bar{\theta}_{t_h:t-1} - \theta_t)^2}{(t - t_h + 1)} \leq \sum_{t=t_h}^{t_l-1} \frac{2(\theta_t - \theta_{t-1})^2 + 2(\bar{\theta}_{t_h:t-1} - \theta_{t-1})^2}{t - t_h + 1}, \tag{A.50}$$

$$\leq \sum_{t=t_h}^{t_l-1} 2(\theta_t - \theta_{t-1})^2 + \sum_{t=t_h}^{t_l-1} \frac{4(\hat{y}(t_h : t-1)[t-1] - (\bar{\theta}_{t_h:t-1} - \theta_{t-1}))^2 + 4(\hat{y}(t_h : t-1)[t-1])^2}{t - t_h + 1}, \tag{A.51}$$

$$\leq \sum_{t=t_h}^{t_l-1} 2(\theta_t - \theta_{t-1})^2 + (\gamma_2 (t_l - t_h)^{1/3}\, TV^{2/3}[t_h : t_l]\, \sigma^{4/3} + 4\sigma^2)(1 + \log n), \tag{A.52}$$

where the last line holds with probability at-least $1 - 2n^{3-\beta/8}$.

$$\|\Theta(t_h : t_l - 1)\|_2^2 \leq \left(\sum_{l=0}^{\log_2(p)-1} 2^{l/2} \|\hat{\alpha}(t_h : t_l - 1)[l]\|_1 \right)^2, + \gamma_3 (t_l - t_h)^{1/3}\, TV^{2/3}[t_h : t_l]\, \sigma^{4/3}, \tag{A.53}$$

$$\leq \sigma^2 + \gamma_3 (t_l - t_h)^{1/3}\, TV^{2/3}[t_h : t_l]\, \sigma^{4/3}, \tag{A.54}$$

with probability at-least $1 - 2n^{3-\beta/8}$ for some constant γ_3 which can depend logarithmically on the width $t_l - t_h$.

Due to Corollary 121 the bounds (A.49), (A.52), (A.54) all simultaneously holds with probability at-least $1 - 2n^{3-\beta/8}$. Combining these bounds, we get

$$\sum_{t=t_h}^{t_l} (\bar{\theta}_{t_h:t-1} - \theta_t)^2 \leq 2\|D\theta_{t_h:t_l}\|_2^2 + \gamma (t_l - t_h)^{1/3}\, TV^{2/3}[t_h : t_l]\, \sigma^{4/3} + 6\sigma^2(1 + \log(n)), \tag{A.55}$$

with probability at-least $1 - 2n^{3-\beta/8}$ and $\gamma = \gamma_1 + \gamma_2(1 + \log(n)) + \gamma_3$.

When summed across all bins as in (A.40), with probability at-least $1 - 2n^{3-\beta/8}$ we have,

$$\sum_{m=1}^{M} \sum_{t=t_h^{(m)}}^{t_l^{(m)}} (\bar{\theta}_{t_h^{(m)}:t-1} - \theta_t)^2 \leq U^2 + 2\|D\theta_{1:n}\|_2^2 + 6M\sigma^2(1+\log n) + \sum_{m=1}^{M} \gamma\, (k^{(m)})^{1/3}\, TV^{2/3}[t_h^{(m)} : t_l^{(m)}]\, \sigma^{4/3}, \tag{A.56}$$

$$\leq U^2 + 2\|D\theta_{1:n}\|_2^2 + 6M\sigma^2(1+\log n) + \gamma\sigma^{4/3} n^{1/3} \left(\sum_{m=1}^{M} \frac{k^{(m)}}{n}\right)^{\frac{1}{3}} \left(\sum_{m=1}^{M} TV[t_h^{(m)} : t_l^{(m)}]\right)^{\frac{2}{3}}, \tag{A.57}$$

$$\leq U^2 + 2\|D\theta_{1:n}\|_2^2 + 6M\sigma^2(1+\log n) + 2\gamma\sigma^{4/3} n^{1/3} C_n^{2/3}. \tag{A.58}$$

Here $k^{(m)}$ is the length of $\Theta(t_h^{(m)} : t_l^{(m)} - 1)$. The term $(\theta_{t_h^{(m)}-1} - \theta_{t_h^{(m)}})^2$ is at-most U^2 for the first bin. We arrive at (A.57) by applying Holder's inequality $x^T y \leq \|x\|_p \|y\|_q$ with $p = 3$ and $q = 3/2$. For both (A.57) and (A.58) we use the fact that $\sum_{m=1}^{M} k^{(m)} \leq 2n$ where the factor 2 is an artifact of zero-padding.

By appealing to lemma 117, we have with probability at-least $1 - 4n^{3-\beta/8}$,

$$\sum_{m=1}^{M} \sum_{t=t_h^{(m)}}^{t_l^{(m)}} (\bar{\theta}_{t_h^{(m)}:t-1} - \theta_t)^2 \leq U^2 + 2\|D\theta_{1:n}\|_2^2 + 12\sigma^2 \log n + 24(\log(n))^2 n^{1/3} C_n^{2/3} \sigma^{4/3} + \gamma\sigma^{4/3} n^{1/3} C_n^{2/3}. \tag{A.59}$$

Next, we proceed to bound the first summation terms in (A.40). For this, we upper-bound the number of bins to control the concentration terms in (A.40) when summed across all bins. Essentially our decision rule should not lead to over binning. Observe that the sum of total variations across all bins is C_n. If the decision rule guarantees (at-least in probability) that total variation inside any detected bin is $\tilde{\Omega}(n^{-1/3})$, then the number of bins is optimally $O(n^{1/3})$. Such a TV lower bounding property is satisfied by wavelet soft-thresholding as described in lemma 117. This is facilitated by the uniform shrinkage property of soft-thresholding estimator. More precisely,

Let's denote

$$V_m = 4\sigma^2 \log(2n^3/\delta)(2 + \log(t_l^{(m)} - t_h^{(m)} + 1)). \tag{A.60}$$

Then,

$$\sum_{m=1}^{M} V_m \leq 4\sigma^2 \log(4n^3/\delta)(2+\log(n)) \max\{1, 2n^{1/3}C_n^{2/3}\sigma^{-2/3}\log(n)\}, \tag{A.61}$$

$$\leq 4\sigma^2 \log(4n^3/\delta)(2+\log(n)) + 8n^{1/3}C_n^{2/3}\sigma^{4/3}\log(n)\log(4n^3/\delta)(2+\log(n)), \tag{A.62}$$

with probability at-least $1-2n^{3-\beta/8}$. Here $[t_h^m, t_l^m]$ corresponds to the m^{th} bin discovered by the policy. This relation follows due to Lemma 117.

Combining (A.62) and (A.59) we have with probability at-least $1-4n^{3-\beta/8}-\delta/2$

$$\begin{aligned} T \leq\ & 8n^{1/3}C_n^{2/3}\sigma^{4/3}(2+\log(n))\log(n) \\ &+ 4\sigma^2\log(4n^3/\delta)(2+\log(n)) \\ &+ U^2 + 2\|D\theta_{1:n}\|_2^2 + 12\sigma^2\log n \\ &\quad + 24(\log(n))^2 n^{1/3}C_n^{2/3}\sigma^{4/3} + 2\gamma\sigma^{4/3}n^{1/3}C_n^{2/3}. \end{aligned} \tag{A.63}$$

By observing that $\|D\theta_{1:n}\|_2 \leq \|D\theta_{1:n}\|_1 = C_n$ we get the bound,

$$\begin{aligned} T \leq\ & 8n^{1/3}C_n^{2/3}\sigma^{4/3}(2+\log(n))\log(n) \\ &+ 4\sigma^2\log(4n^3/\delta)(2+\log(n)) \\ &+ U^2 + 2C_n^2 + 12\sigma^2\log n \\ &\quad + 24(\log(n))^2 n^{1/3}C_n^{2/3}\sigma^{4/3} + 2\gamma\sigma^{4/3}n^{1/3}C_n^{2/3}. \end{aligned} \tag{A.64}$$

The above bounds holds with probability at-least $1-\delta$, if we set $\beta = 24 + \frac{8\log(8/\delta)}{\log(n)}$.

We conclude our proof by observing that the above arguments can be readily extended to any batch smoother that satisfy the following criteria.

- Adaptive minimaxity over any interval within the time horizon.
- The restart decision rule optimally lowerbounds the total variation of any spawned bin.

Thus our policy can be viewed as a meta-algorithm that lifts a “well behaved smoother” to an optimal forecaster in the online setting.

Next we remark how the proof can be adapted to the setting where an extra boundedness constraint is put on the ground truth. i.e, $\theta_{1:n} \in TV(C_n)$ and $|\theta_i| \leq B, i = 1, \ldots, n$. Then the U^2 term in (A.63) becomes B^2. The additive $\|D\theta_{1:n}\|_2^2$ term can be bounded as,

$$\|D\theta_{1:n}\|_2^2 = \sum_{i=2}^{n}(\theta_i - \theta_{i-1})^2, \tag{A.65}$$

$$\leq \sum_{i=2}^{n}(|\theta_i| + |\theta_{i-1}|)(|\theta_i - \theta_{i-1}|), \tag{A.66}$$

$$\leq 2BC_n. \tag{A.67}$$

With the boundedness constraint, we also have $\|D\theta_{1:n}\|_2^2 \leq 4nB^2$. This essentially implies that $\|D\theta_{1:n}\|_2^2 \leq \min\{4nB^2, 2BC_n\}$.

Thus when $\|\theta_{1:n}\|_\infty \leq B$ and if we set $\beta = 24 + \frac{8\log(8/\delta)}{\log(n)}$ then with probability at-least $1-\delta$,

$$\begin{aligned} T \leq\ & 8n^{1/3}C_n^{2/3}\sigma^{4/3}(2+\log(n))\log(n) \\ & + 4\sigma^2\log(4n^3/\delta)(2+\log(n)) \\ & + B^2 + 2\min\{4nB^2, 2BC_n\} + 12\sigma^2\log n \\ & \quad + 24(\log(n))^2 n^{1/3}C_n^{2/3}\sigma^{4/3} + 2\gamma\sigma^{4/3}n^{1/3}C_n^{2/3}. \end{aligned} \tag{A.68}$$

A.5 Adaptive Optimality in Discrete Sobolev class

In this section, we establish that despite the fact that Arrows is designed for the total variation class, it adapts to the optimal rates forecasting sequences that are more regular.

The discrete Sobelov class is defined as

$$\mathcal{S}(C_n') = \{\theta_{1:n} : \|D\theta_{1:n}\|_2 \leq C_n'\}. \tag{A.69}$$

The minimax cumulative error of nonparametric estimation in the discrete Sobolev class is $\theta_{1:n}(n^{2/3}[C_n']^{2/3}\sigma^{4/3})$ [33].

Recall that the discrete Total Variation class that we considered in this paper is defined as

$$\mathcal{T}(C_n) = \{\theta_{1:n} : \|D\theta_{1:n}\|_1 \leq C_n\}. \tag{A.70}$$

By the norm inequalities, we know that

$$\mathcal{T}(C_n') \subset \mathcal{S}(C_n') \subset \mathcal{T}(C_n'\sqrt{n}).$$

The following refinement of our main theorem establishes that Arrows also achieves the minimax rate in discrete Sobolev classes.

Theorem 123. *Let the feedback be $y_t = \theta_t + Z_t$ where Z_t is an independent, σ-subgaussian random variable. Let $\theta_{1:n} \in \mathcal{S}(C_n')$. If $\beta = 24 + \frac{8\log(8/\delta)}{\log(n)}$, then with probability at least $1-\delta$,* Arrows *achieves a dynamic regret of $\tilde{O}(n^{2/3}[C_n']^{2/3}\sigma^{4/3} + U^2 + [C_n']^2 + \sigma^2)$ where $\tilde{O}$ hides a logarithmic factor in n and $1/\delta$.*

Proof. Let's minimally expand the Sobolev ball to a TV ball of radius $C_n = \sqrt{n}C_n'$. This chosen radius of the TV ball is in accordance with the canonical scaling introduced in [33]. This activates the following embedding:

$$\mathcal{S}_1(C_n') \subseteq TV(C_n). \tag{A.71}$$

We can rewrite (A.63) as

$$\begin{aligned} T \leq\ & 8n^{1/3}\|D\theta_{1:n}\|_1^{2/3}\sigma^{4/3}(2+\log(n))\log(n) \\ &+ 4\sigma^2\log(4n^3/\delta)(2+\log(n)) \\ &+ U^2 + 2\|D\theta_{1:n}\|_2^2 + 12\sigma^2\log n \\ &\quad + 24(\log(n))^2 n^{1/3}\|D\theta_{1:n}\|_1^{2/3}\sigma^{4/3} + 2\gamma\sigma^{4/3}n^{1/3}\|D\theta_{1:n}\|_1^{2/3}. \end{aligned} \tag{A.72}$$

The above representation reveals the optimality of our policy over Sobolev class $S_1(C_n')$. Enlarging the Sobolev class to the TV class that contains it does not change the minimax rate in the smoothing setting. See, e.g., Theorem 5 and 6 of [33] and take $d=1$, and $C_n' = n^{-1/2}C_n$. By using $\|x\|_1 \leq n^{1/2}\|x\|_2$ for $x \in \mathbb{R}^n$,

$$\frac{\|D\theta_{1:n}\|_1}{n^{1/2}} \leq \|D\theta_{1:n}\|_2 \leq C_n' = \frac{C_n}{n^{1/2}}. \tag{A.73}$$

Plugging this bound on $\|D\theta_{1:n}\|_1$ in (A.72) recovers the minimax regret for the Sobolev class of radius C_n'. The additional term of $\|D\theta_{1:n}\|_2^2$ — similar to as shown in in appendix A.8 — is unavoidable in the online setting for predicting discrete Sobolev sequences.

□

Remark 124. Note that $\mathcal{T}(C_n') \subset \mathcal{S}(C_n')$, therefore our lower bound from Proposition 6 still applies, which suggests that the additional $[C_n']^2 + \sigma^2$ is required and that Arrows is an optimal forecaster for sequences in Sobolev classes as well.

A.6 Fast Computation

We describe the proof of $O(n\log n)$ runtime guarantee below.

We use an inductive argument. Without loss of generality let the start of current bin be at time 1. Suppose we know the wavelet transform of points upto time t. Let the next highest power of 2 for both t and $t+1$ be p. We identify this value as a pivot for time t and $t+1$. Zero padding is done to hit this pivot. We can view the pad_0 operation at time $t+1$ as the difference between the padded original data and and a step signal. This step signal assume the value $\bar{y}_{1:t+1}$ in time $[1, t+1]$ and 0 in $[t+2, p]$. For computing wavelet transform of the step, we need to update only $O(\log(p))$ coefficients. Inputs to the Haar transform of the padded data at times t and $t+1$ differs by just one co-ordinate. Hence coefficients of only $\log(p)$ wavelets need to be changed. Each such change can be performed in $O(1)$ time in an incremental fashion.

Now let's consider the case when the pivot for time $t+1$ is $2t$. Suppose we know the Haar wavelet coefficients upto time t. In this case, we need to compute the coefficients of $\log(t)$ newly introduced wavelets that span the interval $[t, 2t]$ since the zero padding will force most of the new wavelet coefficients to be zero. The computation of each of those new coefficients can be done in $O(1)$ due to sparsity of signal in interval $[t, 2t]$. We also need to change the first two wavelet coefficients which can be done again in in $O(1)$ time. In all these cases, we only need to do soft-thresholding to the newly updated coefficients. At the base case, when the pivot is just 2, then the computation can be in $O(1)$ time. Thus within a pivot p, the number of computations required is $O(p \log(p))$ which translates to $O(k^{(m)} \log(k^{(m)}))$ computations within the m^{th} bin. Summing across all the bins yields a runtime complexity of $O(n \log(n))$.

A.7 Regret of AOMD

In this section we prove that for any predictable sequence $\{M_t\}_{t=1}^n$, the AOMD algorithm has a dynamic regret of $\tilde{O}(\sqrt{n})$ when applied to our problem. As discussed in Section 2.2, consider loss functions $f_t(x) = (x - y_t)^2$ and comparator sequence $\{u_t\}_{t=1}^n$. First let's consider a deterministic noise setting [31]:

$$y_t = \theta_t + \delta\, \sigma \sqrt{20 \log(n)}, \tag{A.74}$$

where $|\delta| \leq 1$ is chosen by a clever adversary. Let's proceed to get a bound on the quantity D_n. The gradient of our loss function is $2(x - y_t)$. So after observing the values of x_t and M_t, an adversary can pick a suitable δ such that each term of D_n

$$D_n = \sum_{t=1}^{n} \|\nabla f_t(x_t) - M_t\|_*^2. \tag{A.75}$$

can be made $O(1)$. This gives an $O(n)$ bound for D_n.

We can show that V_n is $O(n)$ if we assume that $\mathcal{X}$ is compact and all of the y_t is bounded. Boundedness of y_t follows from the assumptions (A3) and (A4). By appealing to assumption (A3) we see that

$$C_n(u_1, u_2, ..., u_n) = \sum_{t=1}^{n} \|u_t - u_{t-1}\|. \tag{A.76}$$

$C_n(\theta_1, ..., \theta_n)$ is $O(1)$. Plugging this into the regret bound specified in [16] bounds the dynamic regret in our setting as $\tilde{O}(\sqrt{n})$.

We now relate this deterministic noise setting to the guassian setting where the observations are produced according to $y_t = \theta_t + Z_t$, where Z_t is a zero mean sub gaussian with parameter σ^2. As described in proof of theorem 120, $P(\sup_i |Z_i| \geq \sigma\sqrt{20\log(n)}) \leq 2n^{-9}$. Hence by conditioning on the event that $\sup_i |Z_i| \leq \sigma\sqrt{20\log(n)}$, the regret bound of the deterministic noise setting applies to gaussian setting with high probability.

A.8 Lower bound proof

Proof of Proposition 6. First, a lower bound of $\Omega(n^{1/3}C_n^{2/3}\sigma^{4/3})$ is given by [2] for the smoothing estimator $x_{1:n}$ that has more information than we do. The argument uses the fact that the TV-ball is sandwiched between two Besov-bodies with identical minimax rate. To the best of our knowledge, the dependence on C_n and σ is first made explicit in, e.g., [132].

By the fact that "the max is larger than the mean", we have that for any prior distribution $\mathcal{P}$,

$$\sup_{\theta_{1:n}\in \mathrm{TV}(C_n)} \mathbb{E}\left[\sum_{t=1}^n (x_t - \theta_t)^2\right] \geq \mathbb{E}_{\theta_{1:n}\sim\mathcal{P}}\left[\mathbb{E}[\sum_{t=1}^n (x_t - \theta_t)^2 | \theta_{1:n}]\right].$$

Take $\mathcal{P}$ such that

1. $\theta_1 = U$ with probability 0.5 and $-U$ otherwise.
2. $\theta_2 = \theta_1 + C_n$ with probability 0.5 and $\theta_1 - C_n$ otherwise.
3. $\theta_t = \theta_2$ for $t = 3, 4, ..., n$.

Note that x_1 does not observe anything yet, therefore $x_1 = 0$ is the Bayes optimal decision rule. This gives a trivial lower bound of $\mathbb{E}\left[(x_1 - \theta_1)^2\right] \geq U^2$. Now, let's reveal θ_1 to x_2 an additional information, then by the same argument, we have that $\mathbb{E}\left[(x_2 - \theta_2)^2\right] \geq C_n^2$.

Consider an alternative $\mathcal{P}$ when $\theta_1 = ... = \theta_n = \theta$. Let the noise be iid Gaussian with variance σ^2. In this case the problem reduces to a naive statistical estimation problem with $\theta \in [-U, U]$. For each t which observes $t-1$ iid samples from $\mathcal{N}(\theta, \sigma^2)$, then by [133], the minimax risk for this problem is

$$\inf_{\hat{\theta}} \sup_{\theta\in[-U,U]} \mathbb{E}(\hat{\theta} - \theta)^2 = \frac{\sigma^2}{t} - \frac{\pi^2\sigma^4}{tU^2} + o(\frac{\sigma^4}{tU^2}).$$

Summing over $t = 2, 3, ..., n$, and apply the upper/lower bounds of the harmonic series, we have a lower bound of

$$\mathbb{E}\left[\sum_{t=1}^{n}(x_t-\theta_t)^2\right] \geq \max\{0, \sigma^2\log(n+1) - \frac{\pi^2\sigma^4}{U^2}(1+\log(n))(1+o(1))\}.$$

Take the condition that $U > 2\pi\sigma$ and $n > 3$, the above expression can be further lower bounded by $0.5\sigma^2\log(n)$. Note that this bound applies even if $C_n = 0$.

Finally, we can similarly apply the same argument to the case when $\theta_1 = 0$ and $\theta_2 = ... = \theta_n = \theta$ and where the constraint is that $-C_n \leq \theta \leq C_n$. This gives us a lower bound of

$$\mathbb{E}\left[\sum_{t=2}^{n}(x_t-\theta_t)^2\right] \geq \max\{0, \sigma^2\log(n) - \frac{\pi^2\sigma^4}{C_n^2}(1+\log(n-1))(1+o(1))\}.$$

If $C_n > 2\pi\sigma$ and $n > 3$, we can again bound it below by $0.5\sigma^2\log(n)$. In other word, we get the $\sigma^2\log(n)$ lower bound provided that either C_n or U is greater than $2\pi\sigma$.

The proof is complete by taking the average of lower bounds above. We can take $c = 1/6$. □

A.8.1 Lower bound with extra boundedness constraint on ground truth

Suppose we assume $|\theta_i| \leq B, i = 1,\ldots,n$. Then we can adapt the proof presented above by considering a prior $\mathcal{P}$ such that $\theta_i = \epsilon_i B, i = 1,\ldots,\min\{n, 1+\lfloor C_n/2B\rfloor\}$. $\theta_i = \theta_{1+\lfloor C_n/2B\rfloor}, \forall i > \min\{n, 1+\lfloor C_n/2B\rfloor\}$. Here ϵ_i are independent random variables assuming value $+1$ with probability 0.5 and -1 with probability 0.5. Assume that we reveal to learner the probability law of observations θ_i. Under this setting we can see that $\mathbb{E}\left[\sum_{t=1}^{n}(x_t-\theta_t)^2\right] \geq B^2 + \min\{(n-1)B^2, BC_n/2\}$.

Under the setting of $y_i = \theta_i + \epsilon_i$ for iid σ^2 sub-gaussian ϵ_i, $|\theta_i| \leq B$ and $i = 1,\ldots,n$,[20] establishes that minimax total squared error scales as $n\min\{B^2,\sigma^2\}$. This along with previous discussions imply that in the bounded ground truth setting the minimax risk is $\tilde{\Omega}\left(\min\{nB^2, n\sigma^2, n^{1/3}C_n^{2/3}\sigma^{4/3}\} + B^2 + \min\{nB^2, BC_n\} + \sigma^2\right)$.

A.8.2 Minimax regret using Arrows for bounded ground truth

From (A.68) the regret of Arrows T_{Arrows} satisfy

$$T_{\text{Arrows}} = \tilde{O}(n^{1/3}C_n^{2/3}\sigma^{4/3} + \min\{nB^2, BC_n\} + \sigma^2). \quad \text{(A.77)}$$

Let T_1 be the regret of an algorithm, say $\mathcal{A}_1$, that predicts $p \sim N(0,\sigma^2)$ at time step

1 and zero for remaining times. Then it can seen that

$$T_1 = O(nB^2 + \sigma^2), \tag{A.78}$$
$$= O(nB^2 + \sigma^2 + \min\{nB^2, BC_n\}). \tag{A.79}$$

Let T_2 be the regret of an algorithm, say $\mathcal{A}_2$, that predicts y_{t-1} at time t. Then,

$$T_2 = O(n\sigma^2 + \min\{nB^2, BC_n\}). \tag{A.80}$$

Now consider running exponentially weighted average forecaster [40] with three experts: ARROWS, $\mathcal{A}_1$ and $\mathcal{A}_2$. Since squared error is exponentially concave, by Proposition 3.1 of [40] such a forecaster when run with $\eta = 2$ gives a regret T that satisfy,

$$T = O\left(\min\{T_{\text{ARROWS}}, T_1, T_2\} + \log 3\right), \tag{A.81}$$
$$= \tilde{O}\left(\min\{nB^2, n\sigma^2, n^{1/3}C_n^{2/3}\sigma^{4/3}\} + B^2 + \min\{nB^2, BC_n\} + \sigma^2 + \log 3\right). \tag{A.82}$$

Thus we acheive the optimal cumulative squared error upto a small additive term of $\log 3$. If we look at the per round regret this additive term contributes to a small $O(1/n)$ quantity.

A.8.3 Connections to other lower bounds in literature

[15] derived a lower bound of $O(n^{1/2}V_n^{1/2})$ by packing a sequence of quadratic loss functions. Note that this is larger than the upper bound that we attain with quadratic losses. Though this observation seems confusing, a careful study reveals that there is no contradiction. For constructing the lowerbound, [15] used a variational budget V_n as , $V_n = \sum_{t=2}^n \sup_{x \in conv(\theta_1,\ldots\theta_n)} |f_t(x) - f_{t-1}(x)| = \sum_{t=2}^n \sup_{x\in[\theta_{min},\theta_{max}]} |(x-\theta_t)^2 - (x-\theta_{t-1})^2|$, where $conv(.)$ denotes the convex hull of a sequence of points. This is different from the variational budget they use in section 2 of their paper and is also different from C_n that we use for the TV class. When applied to our setting this V_n is no longer proportional to our C_n, instead, it is proportional to $(\theta_{max} - \theta_{min})C_n$.

The packing set constructed through the functions defined in equation (A-12) of [15] obeys $(\theta_{max} - \theta_{min}) = \frac{1}{2}V_n^{1/4}n^{-1/4}$. So we have $C_n = \frac{V_n}{V_n^{1/4}n^{-1/4}} = V_n^{3/4}n^{1/4}$, where we have subsumed proportionality constants. Thus we see that $V_n = \frac{C_n^{4/3}}{n^{1/3}}$. Putting this into their lowerbound recovers exactly our $n^{1/3}C^{2/3}$ bound.

The additional C_n^2 term that appears in our upper bound is required for any methods that do online forecasting of sequences in the TV class. The reason why OGD appears to not require C_n^2 according to [15] is because they require the θ_t to be bounded for all t, while we only require θ_1 to be bounded by U (see Theorem 112).

The lowerbound discussed in [25] considers a more general setting of smooth non-strongly convex sequence of loss functions. Such a lowerbound will not apply in our more restrictive setting.

A.9 Optimality of linear forecasters in Discrete Sobolev class

In this section we first establish that just like Arrows, linear strategies such as OGD and MA are also optimal forecasters for sequences in Discrete Sobolev class. Then we substantiate it using experiments.

Theorem 125. *Let the feedback be* $y_t = \theta_t + Z_t$ *where* Z_t *is an independent,* σ*-subgaussian random variable. Let* $\theta_{1:n} \in \mathcal{S}(C'_n)$*. Restarting OGD with batch size of* $\frac{\sigma^{2/3}(n\log n)^{1/3}}{[C'_n]^{2/3}}$ *achieves an expected dynamic regret of* $\tilde{O}(U^2 + [C'_n]^2 + n^{2/3}[C'_n]^{2/3}\sigma^{4/3})$*.*

Proof. We stick to the same notations as in Appendix A.2. Let's start the analysis from (A.1). Let $t' = t - t_h^{(i)}$.

$$(\theta_t - \bar{\theta}_{t_h^{(i)}:t-1})^2 \leq \frac{\left(\sum_{i=t_h^{(i)}}^{t-1}(\theta_t - \theta_i)\right)^2}{[t']^2}, \tag{A.83}$$

$$\leq \frac{t'}{[t']^2}\sum_{i=t_h^{(i)}}^{t-1}(\theta_t - \theta_i)^2, \tag{A.84}$$

$$\lesssim L[C'_i]^2. \tag{A.85}$$

Hence summing across all points yields,

$$R_i \lesssim L^2[C'_i]^2 + \sigma^2 \log L. \tag{A.86}$$

So the total expected regret becomes,

$$\sum_{i=1}^{\lceil n/L \rceil} R_i \lesssim L^2[C'_n]^2 + \frac{n}{L}\sigma^2 \log L. \tag{A.87}$$

By choosing $L = \frac{\sigma^{2/3}(n\log n)^{1/3}}{[C'_n]^{2/3}}$ we get the theorem. The additive term $[C'_n]^2$ arises similarly as in proof of Theorem 112 □

The optimality of Moving Averages can be proved similarly.

Remark 126. Thus from Theorems 3, 9, 112, 125 we see that Arrows is minimax over both the classes $TV(C_n)$ and $\mathcal{S}(C_n/\sqrt{n})$ while linear forecasters such as OGD and MA require different tuning parameters to perform optimally in each class.

Next, we give numerical experiments substantiating the claims.

Experimental results: Here we consider a doppler function $f(t) = \sin\left(\frac{2\pi(1+\epsilon)}{t/n+0.01}\right)$ with

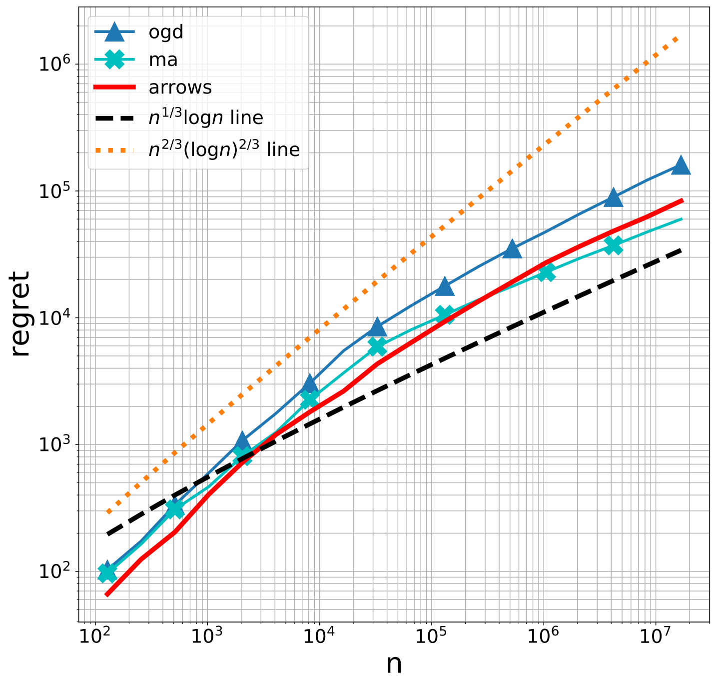


Figure A.3: Regret plot for policies calibrated according to Sobolev radius for a Doppler function

n being the time horizon. For this function $C'_n = \|D\theta\|_2 = O(C_n/\sqrt{n})$ when n is sufficiently large and $\|D\theta\|_2 = O(C_n)$ for small n for a TV bound $C_n = O(1)$. Thus for sufficiently large n, this sequence belong to a small Sobolev ball with radius $O(1/\sqrt{n})$ while the TV class that encloses that Sobolev ball as per Theorem 123 has radius $O(1)$.

We observe noisy data $y_i = f(i/n) + z_i$, $i = 1, ..., n$ and z_i are iid normal variables with $\sigma = 1$.Figure A.3 plots the regret averaged across 5 runs in a log log scale. The necessary input calibration was made as per Remark 124 while running Arrows. We can see that in this case all the algorithms perform in an optimal manner.

Specifically we identify two regimes one for small n and other for larger n. When n is large, we obtain the minimax regret rate $\tilde{O}(n^{1/3})$ due to small C'_n which can be considered as $O(1/\sqrt{n})$. Numerically for $n > 10^5$, C'_n is less than 0.1% of C_n. For smaller values of n where C'_n can be not too small, we attain a regret in accordance with the $\tilde{O}(n^{2/3})$ minimax rate. Numerically when $n < 10^4$, C'_n is atleast 8.5% of C_n which can be considered as $O(C_n) = O(1)$.

Appendix B

Supplementary Materials for Chapter 3

B.1 Background

In this section, we compile some preliminary results well established in literature. For brevity we only discuss the essential aspects that lead to design of our algorithm and its proof.

B.1.1 Non-parametric regression

A popular model studied in non-parametric regression is

$$y_i = f(i/n) + \epsilon_i, i \in [n], \tag{B.1}$$

where ϵ_i are iid subgaussian noise and for unknown $f : [0, 1] \to \mathbb{R}$. The idea is to recover the underlying ground truth f from the observations y_i. Let $\boldsymbol{\theta}_{1:n} = [f(1/n), \ldots, f(1)] \in \mathbb{R}^n$ be the ground truth sequence. We constraint the ground truth to belong to some non-parametric class. A well studied (dating back since 90s atleast) non-parametric family is the class of TV^k bounded sequences defined below.

$$\mathrm{TV}^k(C_n) := \{\boldsymbol{\theta}_{1:n} \in \mathbb{R}^n : n^k \|D^{k+1}\boldsymbol{\theta}_{1:n}\|_1 \le C_n\}. \tag{B.2}$$

The sequences in this class have a piecewise (discrete) polynomial structure. Each stable section features a polynomial of degree atmost k. However the number of polynomial sections and positions where the sequence transitions from one polynomial to another is unknown. This makes the task of estimating ground truth from noisy observations quite challenging. Moreover as noted in [4], such sequences can be used to model a wide spectrum of real world phenomena. TV^k sequences can be obtained by sampling the function whose continuous TV^k distance is bounded. An illustration for $k = 2$ is given in Figure B.1.

The purpose of a non-parametric regression algorithm $\mathcal{A}$ is to estimate $\boldsymbol{\theta}_{1:n}$ given the noisy observations y_i. The most common metric used to ascertain the performance of an algorithm in non-parametric regression literature is the squared error loss. Let the estimates of the algorithm be $\hat{\boldsymbol{y}}_{1:n}$. The empirical risk is defined as

$$R_n = E\left[\sum_{t=1}^{n} (\hat{\boldsymbol{y}}_{1:n}[t] - \boldsymbol{\theta}_{1:n}[t])^2\right], \tag{B.3}$$

and the minimax risk for estimating sequences in $TV^k(C_n)$ is formulated as

$$R_n^* = \min_{\mathcal{A}} \max_{\boldsymbol{\theta} \in TV^k(C_n)} R_n, \tag{B.4}$$

where $\mathcal{A}$ is an estimation of algorithm. It is well established (see eg. [2]) that

$$R_n^* = \Omega(n^{\frac{1}{2k+3}} C_n^{\frac{2}{2k+3}}). \tag{B.5}$$

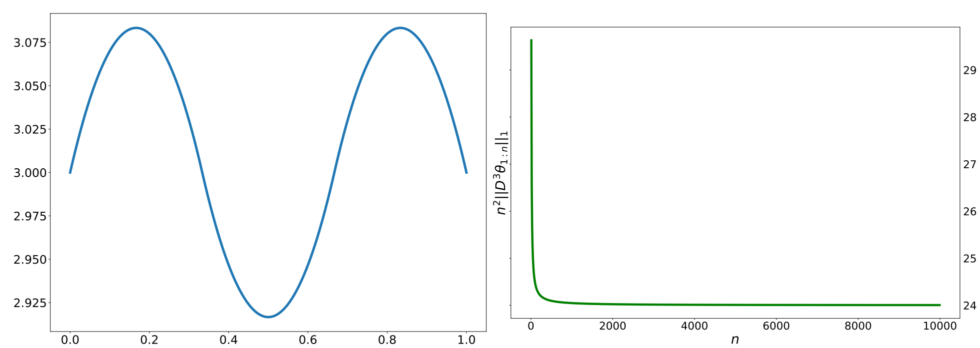


Figure B.1: *A TV^2 bounded sequence $\boldsymbol{\theta}_{1:n}$ can be obtained by sampling the continuous piecewise quadratic function on the left at points i/n, $i \in [n]$. On the right, we plot the TV^2 distance of the generated sequence for various sequence lengths n. As n increases the discrete TV^2 distance converges to a constant value given by the continous TV^2 distance of the function on left panel.*

B.1.2 Wavelet Smoothing

Let $\mathbb{Z}_+ = \mathbb{N} \cup \{0\}$ and $L_2[0,1]$ be the space of all square integrable functions defined in $[0,1]$.

Definition 127. A Multi Resolution Analysis (MRA) on interval [0,1] is a sequence of subspaces $\{V_j, j \in \mathbb{Z}_+\}$ satisfying

1. $V_j \subset V_{j+1}$
2. $f(x) \in V_j$ if and only if $f(2x) \in V_{j+1}$
3. $\bigcap_{j\in\mathbb{Z}_+} = \{0\}$ and $\bigcup_{j\in\mathbb{Z}_+}$ spans $L_2[0,1]$.
4. There exists a function $\phi \in V_0$ such that $\{\phi(x-k) : k \in \mathbb{Z}$ such that $\phi(x-k)$ is supported in [0,1]$\}$ is an orthonormal basis for V_0

The spaces V_j form an increasing sequence of approximations to $L_2[0,1]$. Let $\phi_{jk}(x) = 2^{j/2}\phi(2^j x - k)$. In what follows we define $\phi_{jk}(x) = 0$ if it is not supported entirely within $[0,1]$. Due to properties 2 and 4 it follows that $\{\phi_{jk}(x), k \in \mathbb{Z}\}$ is an orthonormal basis for V_j. The function $\phi(x)$ is called the *scale function.*

Now let's define wavelets. *Detail subpace* $W_j \subset L_2[0,1]$ is defined as the orthogonal complement of V_j in V_{j+1}. A function $\psi(x)$ is defined to be a *wavelet* (or mother wavelet) function if $\{\psi_{jk}(x) = 2^{j/2}\psi(2^j x - k), k \in \mathbb{Z}_+$ such that $\psi_{jk}(x)$ is supported in $[0,1]\}$ is an orthonormal basis for W_j $\forall j \in \mathbb{Z}_+$.

Definition 128. A wavelet function $\psi(x)$ has regularity r if

$$\int_0^1 x^p \psi(x)dx = 0, p = 0, \dots, r-1. \tag{B.6}$$

The CDJV construction in [41] is an algorithm that provides a scale function $\phi(x)$ and wavelet function $\psi(x)$ of a given regularity r. We record an important property of this construction.

Proposition 129. *The CDJV construction with regularity r satisfy*

1. *Let $L = \lceil \log 2r \rceil$. Then V_L contains polynomials of degree $\leq r-1$.*
2. *The functions $\psi_{jk}(x), j \geq L, k \in \mathbb{Z}$ are orthogonal to polynomials of degree atmost $r-1$.*

Let $n = 2^J$ and $L < J$. A discrete Wavelet Transform (DWT) matrix $\boldsymbol{W} \in \mathbb{R}^{n\times n}$ is generated by sampling the basis functions that make up V_L and $W_L, \dots, W_{J-1}$ at points $i/n, i \in [n]$ and scaling them by a factor of $n^{-1/2}$. The obtained matrix $\boldsymbol{W}$ can be shown to be orthonormal. The total number of basis functions that make up the space V_J is n.

Now to provide a clearer picture, we orchestrate all the above ideas with the help of the simple *Haar wavelets.*

Definition 130. The Haar MRA on [0,1] is defined by

1. The scale function $\phi(x) = 1$
2. The mother wavelet $\psi(x) = -1$ if $x \leq 1/2; 1$ otherwise.

3. Both $\phi(x), \psi(x)$ are zero outside $[0, 1]$

Here V_0 is the space of constant signals in $[0, 1]$. W_0 is the functions of the form $c\psi(x)$ for $c \in \mathbb{R}$. W_1 is spanned by $\psi_{10}(x)$ and $\psi_{11}(x)$ and so on. It is clear that regularity of Haar wavelet $\psi(x)$ is 1. In fact Haar system is a special case of CDJV construction for regularity 1. Hence $L = \lceil \log 2r \rceil = 1$. The space V_1 is spanned by $\{\phi(x), \psi(x)\}$. It is easy to verify that space V_1 contains all polynomials of degree $r - 1 = 0$ as asserted by Proposition 129. Furthermore property 2 stated in Proposition 129 is also true.

Now let's construct the orthonormal Haar DWT matrix $\boldsymbol{W} \in \mathbb{R}^{n \times n}$. Let $J = \log n$ We need to sample sample basis functions of $V_1, W_1, \ldots W_{J-1}$ at points $i/n, i \in [n]$ and scale them by $n^{-1/2}$. For simplicity we illustrate this for $n = 4$.

$$\boldsymbol{W} = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 & -1/2 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 & 0 \\ 0 & 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}. \tag{B.7}$$

It is noteworthy that general CDJV wavelets for regularity $r \geq 2$ do not have a closed form expression like the Haar system. The filter coefficients are computed by an efficient iterative algorithm.

Define the soft thresholding operator as

$$T_\lambda(x) = \begin{cases} 0 & |x| \leq \lambda \\ x - \lambda & x > \lambda \\ x + \lambda & x < \lambda \end{cases} \tag{B.8}$$

If the input is a vector the operation is done co-ordinate wise.

Now we are ready to discuss the famous universal soft thresholding algorithm of [2].

WaveletSoftThreshold: Inputs - observations $\boldsymbol{y}_{1:n}$, subgaussian parameter σ of noise in (B.1), TV order k

1. Let $\boldsymbol{W} \in \mathbb{R}^{n \times n}$ be a CDJV DWT matrix of regularity $k + 1$.
2. Output $\hat{\boldsymbol{y}}_{1:n} = \boldsymbol{W}^T T_{\sigma\sqrt{2 \log n}}(\boldsymbol{W} y)$.

We have the following proposition due to [2].

Proposition 131. *The risk of the wavelet soft thresholding scheme satisfy*

$$R_n = \tilde{O}(n^{\frac{1}{2k+3}} C_n^{\frac{2}{2k+3}}). \tag{B.9}$$

Comparing with equation (B.5) we see that WaveletSoftThreshold is a near minimax algorithm for estimating sequences in $TV^k(C_n)$. It optimally adapts to the unknown radius C_n as well.

B.1.3 Vovk Azoury Warmuth (VAW) forecaster

The VAW algorithm is shown in Figure B.2. For a more elaborate discussion on this algorithm, refer to chapter 11 of [40]. The VAW forecaster is defined as follows.

VAW algorithm

1. Adversary reveals $\boldsymbol{x}_t \in \mathbb{R}^d$.
2. Agent predicts $\hat{p}_t = \hat{\boldsymbol{w}}_{t-1}^T \boldsymbol{x}_t$ with $\hat{\boldsymbol{w}}_t = (\boldsymbol{I} + \sum_{s=1}^{t} \boldsymbol{x}_s \boldsymbol{x}_s^T)^{-1} \sum_{s=1}^{t-1} y_s \boldsymbol{x}_s$.
3. Adversary reveals y_t.
4. Incur loss $(\hat{p}_t - y_t)^2$.

Figure B.2: The VAW algorithm

We have the following guarantee on the regret bound of VAW.

Proposition 132. *If the VAW forecaster is run on a sequence* $(\boldsymbol{x}_1, y_1), \dots, (\boldsymbol{x}_n, y_n) \in \mathbb{R}^d \times \mathbb{R}$, *then for all* $\boldsymbol{u} \in \mathbb{R}^d$ *and* $n \geq 1$,

$$\sum_{t=1}^{n} (y_t - \hat{p}_t)^2 - (y_t - \boldsymbol{u}^T \boldsymbol{x}_t) \leq \frac{1}{2} \|\boldsymbol{u}\|_2^2 + \frac{dY^2}{2} \log(1 + \frac{nX^2}{d}), \tag{B.10}$$

where $\|\boldsymbol{x}_t\|_2 \leq X$, *and* $|y_t| \leq Y, t \in [n]$.

B.2 Analysis

B.2.1 Connecting wavelet coefficients and higher order TV^k distance

Lemma 133. *Let* $\tilde{\boldsymbol{\theta}}_{1:t} = \texttt{recenter}(\boldsymbol{\theta}_{1:t})$ *and* $(\boldsymbol{a}, \boldsymbol{b}) = \texttt{pack}(\tilde{\boldsymbol{\theta}}_{1:t})$. *For an orthonormal DWT matrix* $\boldsymbol{W}$,

$$\frac{\|\boldsymbol{W}\boldsymbol{a}\|_2 + \|\boldsymbol{W}\boldsymbol{b}\|_2}{\sqrt{t}} \lesssim t^k \|D^{k+1} \boldsymbol{\theta}_{1:t}\|_1, \tag{B.11}$$

where we have subsumed constants that depend only on k.

Proof. Consider the truncated power basis with knots at points $\frac{1}{n}, \frac{2}{n}, \ldots, 1$ defined as follows:

$$g_1(x) = 1,\ g_2(x) = x, \ldots,\ g_k(x) = x^k \tag{B.12}$$

$$g_{k+1+j}(x) = \left(x - \frac{j}{n}\right)_+^k,\ j = 1, \ldots, n-k-1, \tag{B.13}$$

$x_+ = \max\{x, 0\}$. Since an $t \times t$ matrix $\boldsymbol{G}$ with entries $g_j(\frac{i}{t})$ at the position (i, j) is invertible, we can write any sequence $\boldsymbol{\theta}_{1:t}$ as

$$\boldsymbol{\theta}_{1:t}[i] = \sum_{j=1}^{t} \beta_j g_j(\frac{i}{t}), \tag{B.14}$$

for $i = 1, \ldots, t$. From the above equation we see that,

$$t^k \|D^{k+1} \boldsymbol{\theta}_{1:t}\|_1 = k! \sum_{j=k+2}^{t} |\beta_j| \tag{B.15}$$

Let $\tilde{\boldsymbol{\theta}}_{1:t} = \texttt{recenter}(\boldsymbol{\theta}_{1:t})$. Let $\tilde{\boldsymbol{g}}_j = \texttt{recenter}(\boldsymbol{g}_j)$ where $\tilde{\boldsymbol{g}}_j$ is the j^{th} column of the matrix $\boldsymbol{G}$. Since $\|\boldsymbol{g}_j\|_\infty \leq 1$ we have $\|\tilde{\boldsymbol{g}}_j\|_\infty = O(1)$ where the hidden constant only depends on k.

Thus

$$\|\tilde{\boldsymbol{\theta}}_{1:t}\|_\infty = \left\| \sum_{j=k+2}^{t} \beta_j \tilde{\boldsymbol{g}}_j \right\|_\infty, \tag{B.16}$$

$$\leq \sup_{k+2 \leq i \leq t} \|\tilde{\boldsymbol{g}}_i\|_\infty \sum_{j=k+2}^{t} |\beta_j|, \tag{B.17}$$

$$\lesssim t^k \|D^{k+1} \boldsymbol{\theta}_{1:t}\|_1, \tag{B.18}$$

where the last line follows from (B.15). We subsume a constant that only depends on k. Now using $\|\boldsymbol{x}\|_2 \leq \sqrt{m}\|\boldsymbol{x}\|_\infty$ for $\boldsymbol{x} \in \mathbb{R}^m$, we have

$$\frac{\|\tilde{\boldsymbol{\theta}}_{1:t}\|_2}{\sqrt{t}} \lesssim t^k \|D^{k+1} \boldsymbol{\theta}_{1:t}\|_1. \tag{B.19}$$

We have thus established a lower-bound on the TV using the energy of the OLS residuals. For a vector $\boldsymbol{z}$ let $(\boldsymbol{x}, \boldsymbol{y}) = \texttt{pack}(\boldsymbol{z})$. We have the following relations,

$$\|\boldsymbol{z}\|_2 \geq \sqrt{\frac{\|\boldsymbol{x}\|_2^2 + \|\boldsymbol{y}\|_2^2}{2}}, \tag{B.20}$$

$$\geq \frac{\|\boldsymbol{x}\|_2 + \|\boldsymbol{y}\|_2}{2}, \tag{B.21}$$

where the last line follows from Jensen's inequality and the concavity of $\sqrt{\cdot}$ function.

□

B.2.2 Bounding the Regret

Our proof strategy falls through the following steps.

1. Obtain a high probability bound of bias variance decomposition type on the total squared error incurred by the policy within a bin.

2. Bound the variance by optimally bounding the number of bins spawned.

3. Bound the bias using the restart criterion and adaptive minimaxity of soft-thresholding estimator [2].

Lemma 134. ***(bias-variance bound)**)* *Let $E[\hat{y}_t] = p_t$. For any bin $[t_h, t_l]$ with $t_h \geq k$ discovered by the policy, we have with probability atleast $1 - \delta/2$*

$$\sum_{t=t_h}^{t_l} (\hat{y}_t - \boldsymbol{\theta}_{1:n}[t])^2 \leq \sum_{t=t_h}^{\bar{t}_l} 2(p_t - \boldsymbol{\theta}_{1:n}[t])^2 + 4\sigma^2(k+1)\log\left(1 + \frac{n^{2k+3}}{k+1}\right)\log(4n^3/\delta). \tag{B.22}$$

Proof. First let's consider an arbitrary interval $[\underline{l}, \bar{l}]$ such that $\underline{l} \geq k$. We proceed to bound the bias and variance of predictions made by a VAW forecaster. Note that the bin $[\underline{l}, \bar{l}]$ is arbitrary and may not be an interval discovered by the policy. The predictions made by VAW forecaster at time $t \in [\underline{l}, \bar{l}]$ is given by,

$$\hat{y}_t = \langle x_t, \tilde{\boldsymbol{A}}_{\boldsymbol{t}}^{-1} \sum_{s=\underline{l}-k}^{t-1} y_s \boldsymbol{x}_{\boldsymbol{s}} \rangle, \tag{B.23}$$

where $\tilde{\boldsymbol{A}}_{\boldsymbol{t}} = \boldsymbol{I} + \sum_{s=\underline{l}-k}^{t} \boldsymbol{x}_{\boldsymbol{s}} \boldsymbol{x}_{\boldsymbol{s}}^T$.

Let

$$p_t = E[\hat{y}_t], \tag{B.24}$$

$$= \langle x_t, \tilde{\boldsymbol{A}}_{\boldsymbol{t}}^{-1} \sum_{s=\underline{l}-k}^{t-1} \boldsymbol{\theta}_{1:n}[s \boldsymbol{x}_{\boldsymbol{s}} \rangle. \tag{B.25}$$

For notational convenience, define

$$\boldsymbol{X_t} = [\boldsymbol{x_{\underline{l}-k}}, \ldots, \boldsymbol{x_t}]^T. \tag{B.26}$$

Let

$$\begin{aligned}
\mathrm{Var}(\hat{y}_t) &= \sigma^2 \boldsymbol{x_t}^T \tilde{\boldsymbol{A}}_{\boldsymbol{t}}^{-1} \boldsymbol{X_t}^T \boldsymbol{X_t} \tilde{\boldsymbol{A}}_{\boldsymbol{t}}^{-1} \boldsymbol{x_t}, && \text{(B.27)}\\
&\le \sigma^2 \boldsymbol{x_t}^T \tilde{\boldsymbol{A}}_{\boldsymbol{t}}^{-1} \boldsymbol{x_t}, && \text{(B.28)}\\
&= \sigma_t^2 && \text{(B.29)}
\end{aligned}$$

where the last line is due to $\boldsymbol{X_t}^T \boldsymbol{X_t} \preccurlyeq \tilde{\boldsymbol{A}}_{\boldsymbol{t}}$, where $\boldsymbol{U} \preccurlyeq \boldsymbol{V}$ means $\boldsymbol{V} - \boldsymbol{U}$ is a Positive Semi Definite matrix.

Define a normalized random variable

$$Z_t = \frac{\hat{y}_t - p_t}{\sigma_t}. \tag{B.30}$$

Thus Z_t is a sub-gaussian random variable with variance parameter 1. By sub-gaussian tail inequality we have,

$$P\left(|Z_t| \ge \sqrt{2\log(4n^3/\delta)}\right) \le \delta/2n^3, \tag{B.31}$$

for some $\delta \in (0,1]$. Noting that length of a bin is atmost n, an application of uniform bound yields

$$P\left(\sup_{\underline{l} \le t \le l} |Z_t| \ge \sqrt{2\log(4n^3/\delta)}\right) \le \delta/2n^2. \tag{B.32}$$

Adding and subtracting a $\boldsymbol{\theta}_{1:n}[t]$ to the numerator of (B.30), we get that with probability atleast $1 - \delta/2n^2$,

$$|\hat{y}_t - \boldsymbol{\theta}_{1:n}[t]| \le |p_t - \boldsymbol{\theta}_{1:n}[t]| + \sigma_t\sqrt{2\log(4n^3/\delta)}, \forall t \in [\underline{l}, \bar{l}]. \tag{B.33}$$

Hence the squared error within a bin can be bounded in probability as

$$\sum_{t=\underline{l}}^{\bar{l}} (\hat{y}_t - \boldsymbol{\theta}_{1:n}[t])^2 \le \sum_{t=\underline{l}}^{\bar{l}} 2(p_t - \boldsymbol{\theta}_{1:n}[t])^2 + 4\sigma_t^2 \log(4n^3/\delta), \tag{B.34}$$

where we used $(a+b)^2 \le 2a^2 + 2b^2$.

Let's focus on the second term in (B.34). By lemma 11.11 of [40] and by following the arguments of proof of Theorem 11.7 there, we get

$$\sum_{t=\underline{l}}^{\bar{l}} \sigma_t^2 \leq \sigma^2 \sum_{d=1}^{k+1} \log(1+\lambda_d), \tag{B.35}$$

where λ_d are the eigenvalues of the $(k+1)\times(k+1)$ matrix $\tilde{\boldsymbol{A}}_{\bar{l}} - \boldsymbol{I}$. It is well known that $\tilde{\boldsymbol{A}}_{\bar{l}} - \boldsymbol{I}$ has the same nonzero eigenvalues as the Gram matrix $\boldsymbol{G}$ with entries $G_{i.j} = \boldsymbol{x_i}^T \boldsymbol{x_j}$. Note that $\|\boldsymbol{x_t}\|_2^2 \leq n^{2k+2}, \forall t \in [1,n]$. Since the product $\Pi_{d=1}^{k+1}(1+\lambda_d)$ is maximised when $\lambda_d = (\underline{l} - \bar{l})n^{2k+2}/(k+1) \leq n^{2k+3}/(k+1)$ we have,

$$\sigma^2 \sum_{d=1}^{k+1} \log(1+\lambda_d) \leq \sigma^2 (k+1) \log(1 + \frac{n^{2k+3}}{k+1}). \tag{B.36}$$

Thus with probability atleast $1 - \delta/n^2$

$$\sum_{t=\underline{l}}^{\bar{l}} (\hat{y}_t - \boldsymbol{\theta}_{1:n}[t])^2 \leq \sum_{t=\underline{l}}^{\bar{l}} 2(p_t - \boldsymbol{\theta}_{1:n}[t])^2 + 4\sigma^2(k+1)\log\left(1 + \frac{n^{2k+3}}{k+1}\right)\log(4n^3/\delta). \tag{B.37}$$

As mentioned earlier, the bin $[\underline{l}, \bar{l}]$ can be arbitrary and may not be discovered by policy. However, we want to analyze the Total Squared Error (TSE) incurred within true bins spawned by the policy. A small caveat here is that observations within such true bins satisfy the restart criteria and can't be regarded as independent random variables. To get rid of this problem, we use a uniform bound argument to bound the TSE incurred in all possible $O(n^2)$ bins. This leads to

$$P\Bigg(\sup_{[\underline{l},\bar{l}]} \sum_{t=\underline{l}}^{\bar{l}} (\hat{y}_t - \boldsymbol{\theta}_{1:n}[t])^2 - \sum_{t=\underline{l}}^{\bar{l}} 2(p_t - \boldsymbol{\theta}_{1:n}[t])^2 \tag{B.38}$$

$$- 4\sigma^2(k+1)\log\left(1 + \frac{n^{2k+3}}{k+1}\right)\log(4n^3/\delta) \geq 0 \Bigg) \leq \delta/2. \tag{B.39}$$

□

Lemma 135. ***(subgaussian wavelet coefficients)*** *Let* $(\boldsymbol{y_1}, \boldsymbol{y_2}) = \texttt{pack}(\texttt{recenter}(\boldsymbol{y}))$ *for a vector* $\boldsymbol{y}$ *of observations of length L. Let* $(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}) = (\boldsymbol{W y_1}, \boldsymbol{W y_2})$ *for an orthonormal DWT matrix* $\boldsymbol{W}$*. Then both* $\boldsymbol{\alpha_1}$ *and* $\boldsymbol{\alpha_2}$ *are marginally subgaussian with parameter* $4\sigma^2$*.*

Proof. From the theory of least squares regression,

$$\texttt{recenter}(\boldsymbol{y}) = \boldsymbol{y} - \boldsymbol{X_L}(\boldsymbol{X_L}^T \boldsymbol{X_L})^{-1} \boldsymbol{X_L}^T \boldsymbol{y}, \tag{B.40}$$

where $\boldsymbol{X_L}$ is defined as in (B.26). Since $L \geq k+1$, $\boldsymbol{X_L}^T\boldsymbol{X_L}$ can be shown to be invertible. (see for eg. lemma 150)

Without loss of generality, we proceed to characterize the sub-gaussian behaviour of the *first* wavelet coefficient of $\boldsymbol{y_1}$. The extension to other wavelet coefficients is straight forward.

Let $\boldsymbol{u}^T$ be the first row of the wavelet transform matrix $\boldsymbol{W}$ whose dimension is compatible to $\boldsymbol{y_1}$. Let's augment $\boldsymbol{u}^T$ as follows.

$$\tilde{\boldsymbol{u}}^T = [\boldsymbol{u}^T, \boldsymbol{0}^T], \tag{B.41}$$

such that length of $\tilde{\boldsymbol{u}}$ is L.

We have,

$$\boldsymbol{\alpha_1}[0] = \tilde{\boldsymbol{u}}^T y - \tilde{\boldsymbol{u}}^T \boldsymbol{X_L}(\boldsymbol{X_L}^T\boldsymbol{X_L})^{-1}\boldsymbol{X_L}^T\boldsymbol{y}. \tag{B.42}$$

(B.42) along with noisy feedback implies that $\boldsymbol{\alpha_1}[0]$ is a Lipschitz function of L iid subgaussian random variables. Then by Proposition 2.12 from [18], $\boldsymbol{\alpha_1}[0]$ is also subgaussian with variance parameter given by the square of Lipschitz constant ℓ^2 times σ^2. Since $\boldsymbol{\alpha_1}[0]$ is a linear function of the iid subgaussians we have,

$$\begin{aligned}
\ell &= \|\tilde{\boldsymbol{u}} - \boldsymbol{X_L}(\boldsymbol{X_L}^T\boldsymbol{X_L})^{-1}\boldsymbol{X_L}^T\tilde{\boldsymbol{u}}\|_2, && \text{(B.43)}\\
&\leq \|\tilde{\boldsymbol{u}}\|_2 + \|\boldsymbol{X_L}(\boldsymbol{X_L}^T\boldsymbol{X_L})^{-1}\boldsymbol{X_L}^T\tilde{\boldsymbol{u}}\|_2, && \text{(B.44)}\\
&\leq_{(a)} \|\boldsymbol{u}\|_2 + \|\boldsymbol{X_L}(\boldsymbol{X_L}^T\boldsymbol{X_L})^{-1}\boldsymbol{X_L}^T\|_2\|\boldsymbol{u}\|_2, && \text{(B.45)}\\
&=_{(b)} 2. && \text{(B.46)}
\end{aligned}$$

In (a) we used $\|\boldsymbol{Ax}\|_2 \leq \|\boldsymbol{A}\|_2\|\boldsymbol{x}\|_2$ where $\|\boldsymbol{A}\|_2$ is the induced matrix norm and the fact that $\|\tilde{\boldsymbol{u}}\|_2 = \|\boldsymbol{u}\|_2$. In (b) we notice that $\|\boldsymbol{u}\|_2 = 1$ as the DWT matrix $\boldsymbol{W}$ is orthonormal and $\|\boldsymbol{X_L}(\boldsymbol{X_L}^T\boldsymbol{X_L})^{-1}\boldsymbol{X_L}^T\|_2 = 1$ since $\boldsymbol{X_L}(\boldsymbol{X_L}^T\boldsymbol{X_L})^{-1}\boldsymbol{X_L}^T$ is a projection matrix.

Similarly it can be shown that $\boldsymbol{\alpha_2}$ is marginally subgaussian with parameter $4\sigma^2$. □

Lemma 136. ***(uniform shrinkage)*** *Assume the setting of lemma 135. Let* $(\hat{\boldsymbol{\alpha}}_{\boldsymbol{1}}, \hat{\boldsymbol{\alpha}}_{\boldsymbol{2}}) = (T(\boldsymbol{\alpha_1}), T(\boldsymbol{\alpha_2}))$ *where* $T(\cdot)$ *is the soft-thresholding operator with threshold* $\sigma\sqrt{\beta \log n}$*. Then with probability atleast* $1 - 2n^{3-\beta/8}$*,* $|(\hat{\boldsymbol{\alpha}}_{\boldsymbol{r}})_i| \leq |E\left[(\boldsymbol{\alpha_r})_i\right]|$ *for each co-ordinate* i *and* $r = 1, 2$*. The expectation is taken wrt to randomness in the observations.*

Proof. Consider a fixed bin $[\underline{l}, \bar{l}]$. Due to results of lemma 135 and subgaussian tail inequality,

$$P\left(|(\hat{\boldsymbol{\alpha}}_{\boldsymbol{r}})_i - E\left[(\boldsymbol{\alpha_r})_i\right]| \geq \sigma\sqrt{\beta \log n}\right) \leq 2n^{-\beta/8}. \tag{B.47}$$

Then arguing in the similar lines as in the proof of lemma 15 of [37], the result follows. □

Lemma 137. ***(bin control)*** *With probability atleast $1 - 2n^{3-\beta/8}$, the number of bins M, spawned by the policy is atmost*
$\min\left\{n, \max\{1, \tilde{O}(n^{\frac{1}{2k+3}}\|n^k D^{(k+1)}\boldsymbol{\theta}_{1:n}\|_1^{\frac{2}{2k+3}})\}\right\}$ *where $\tilde{O}$ hides factors that depend on wavelet function, constants that only depend on TV order k and polynomial factors of $\log n$.*

Proof. Let L_i be the length of the i^{th} bin. Let $\hat{\boldsymbol{\alpha}}_{1i}, \hat{\boldsymbol{\alpha}}_{2i}$ be the denoised wavelet coefficient segments of the re-centered observations within a bin i as described in the policy and $\boldsymbol{\theta_i}$ be the ground truth vector in bin i.

By the policy's restart rule,

$$\frac{\sigma}{\sqrt{L_i}} \leq \frac{1}{\sqrt{L_i}}\left(\|\hat{\boldsymbol{\alpha}}_{1i}\|_2 + \|\hat{\boldsymbol{\alpha}}_{2i}\|_2\right). \tag{B.48}$$

Due to the uniform shrinkage property specified in lemma 136, we have with probability atleast $1 - 2n^{3-\beta/8}$

$$\frac{\sigma}{\sqrt{L_i}} \leq \frac{1}{\sqrt{L_i}}\left(\|\boldsymbol{\alpha}_{1i}\|_2 + \|\boldsymbol{\alpha}_{2i}\|_2\right), \tag{B.49}$$

$$\lesssim_{(a)} 2^k L_i^k \|D^{k+1}\boldsymbol{\theta_i}\|_1, \tag{B.50}$$

where (a) follows due to lemma 133. The factor of 2^k is due to the fact that length of vectors $\boldsymbol{\alpha}_{1i}$ or $\boldsymbol{\alpha}_{2i}$ is atmost $2L_i$. The last line implies that when the TV^k distance is zero, `Ada-VAW` doesn't restart with high probability making $M = 1$.

Rearranging and summing across all bins yields

$$\sum_{i=1}^{M} \frac{\sigma}{L_i^{k+1/2}} \lesssim \|D^{k+1}\boldsymbol{\theta}_{1:n}[t]\|_1. \tag{B.51}$$

Now applying Jensen's inequality for the convex function $f(x) = \frac{1}{x^{k+1/2}}, x > 0$, we get

$$\sigma M^{\frac{2k+3}{2}} n^{\frac{-(2k+1)}{2}} \lesssim \|D^{k+1}\boldsymbol{\theta}_{1:n}\|_1, \tag{B.52}$$

where $\lesssim$ subsumes constants that depend only on wavelet functions, TV order k and polynomial factors of $\log n$.

Rearranging the last expression yields the lemma. □

Lemma 138. ***(Vovk-Azoury-Warmuth regret)*** *If the Vovk-Azoury-Warmuth forecaster with output denoted by $\hat{v}_j$ at time j, is run on a sequence*

$(\boldsymbol{w_1}, v_1), \ldots, (\boldsymbol{w_n}, v_n) \in \mathbb{R}^{k+1} \times \mathbb{R}$, *then for all* $\boldsymbol{u} \in \mathbb{R}^{k+1}$,

$$\sum_{j=1}^{t} (\hat{v}_j - v_j)^2 - (\boldsymbol{u}^T \boldsymbol{w_j} - v_j)^2 \leq \frac{1}{2}\|\boldsymbol{u}\|_2^2 + \frac{(k+1)B^2}{2} \log\left(1 + \frac{t^{k+2}}{k+1}\right), \tag{B.53}$$

$$= \tilde{O}(B^2), \tag{B.54}$$

where $B = \max_{i=1,\ldots,t} |y_i|$ *and* $\boldsymbol{w_j} = [1, j, \ldots, j^k]^T$.

Proof. The first inequality is due to Theorem 11.8 of [40]. The second equality follows because under the given choice of monomial features, it is shown in Corollary 154 that when $\boldsymbol{u}$ is the coefficient vector of OLS fit, $\|\boldsymbol{u}\|_2^2 = O(B^2)$. □

Next we characterize the optimality of soft-thresholding estimator on TV^k class. The key to this is the Theorem 19 from [37].

Theorem 139. *[37] Consider the observation model* $\breve{\boldsymbol{y}} = \breve{\boldsymbol{\alpha}} + \boldsymbol{Z}$, *where* $\breve{\boldsymbol{y}} \in \mathbb{R}^n$, $\boldsymbol{Z}$ *is marginally subgaussian with parameter* σ^2 *and* $\breve{\boldsymbol{\alpha}} \in \boldsymbol{A}$ *for some solid and orthosymmetric* $\boldsymbol{A}$. *Let* $\hat{\boldsymbol{\alpha}}_\delta$ *be the soft thresholding estimator with input* $\breve{\boldsymbol{y}}$ *and threshold* δ. *When* $\delta = \sigma\sqrt{\beta \log n}$, *with probability atleast* $1 - 2n^{1-\beta/2}$ *the estimator* $\hat{\boldsymbol{\alpha}}_\delta$ *satisfies*

$$\|\hat{\alpha}_\delta - \alpha\|^2 \leq 8.88\beta(1 + \log(n)) \inf_{\hat{\alpha}} \sup_{\alpha \in A} E[\|\hat{\alpha} - \alpha\|^2]. \tag{B.55}$$

We are interested in the case where $\boldsymbol{A}$ is the space of wavelet coefficients for TV^k bounded fucntions. Since TV^k class is sandwiched between two Besov spaces, it can be shown that $\boldsymbol{A}$ is solid and orthosymmetric (see for eg. [18], section 4.8). Note that subtracting a polynomial of degree k has no effect on the TV^k distance. It has been established in lemma 135 that OLS residual are subgaussian with parameter $4\sigma^2$. Hence we are under the observation model of Theorem 139. By the results of [2], we have $\inf_{\hat{\alpha}} \sup_{\alpha \in A} E[\|\hat{\alpha} - \alpha\|^2] = \tilde{O}(n^{\frac{1}{2k+3}}(n^k D\boldsymbol{\theta}_{1:n}\|_1)^{\frac{2}{2k+3}}\sigma^{\frac{4k+4}{2k+3}})$. This along with using a uniform bound across all $O(n^2)$ bins leads to the following Corollary.

Corollary 140. *Under the observation model and notations in Theorem 139 but with a subgassuan parameter* $4\sigma^2$ *when* $\boldsymbol{A}$ *is the wavelet coefficients of re-centered ground truth within a bin discovered by the policy, then with probability atleast* $1 - 2n^{3-\beta/8}$

$$\|\hat{\alpha}_\delta - \alpha\|^2 = \tilde{O}(n^{\frac{1}{2k+3}}(n^k D\boldsymbol{\theta}_{1:n}\|_1)^{\frac{2}{2k+3}}\sigma^{\frac{4k+4}{2k+3}}). \tag{B.56}$$

Lemma 141. ***(bias control)*** *Let* $E[\hat{y}_t] = p_t$. *For any bin* $[t_h, t_l]$, $L = t_l - t_h$, *with* $t_h \geq k$ *discovered by the policy, we have with probability atleast* $1 - 2n^{3-\beta/8}$

$$\sum_{t=t_h}^{\bar{t}_l - 1} (p_t - \boldsymbol{\theta}_{1:n}[t])^2 = \tilde{O}(1) + \tilde{O}\left(L^{\frac{2k+1}{2k+3}} \|D^{k+1}\boldsymbol{\theta}_{t_h-k:t_l-1}\|_1^{\frac{2}{2k+3}}\right) + (p_{t_l} - \boldsymbol{\theta}_{1:n}[t_l])^2. \tag{B.57}$$

Proof. For a bin $[t_h, t_l]$ let

$$T = \sum_{t=t_h}^{t_l} (p_t - \boldsymbol{\theta}_{1:n}[t])^2. \tag{B.58}$$

Note that T is the squared error incurred by the VAW forecaster when run with the sequence $\boldsymbol{\theta}_{t_h:t_l}$. Let $\boldsymbol{u}$ be the coefficient of the OLS fit using monomial features for the ground truth $[\boldsymbol{\theta}_{t_h-k:t_l-1}]$. Further let's recall/adopt the following notations:

1 $(\boldsymbol{g_1}, \boldsymbol{g_2}) = \texttt{pack}\left(\texttt{recenter}(\boldsymbol{\theta}_{t_h-k:t_l-1})\right)$;

2 $(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}) = (\boldsymbol{W}\boldsymbol{g_1}, \boldsymbol{W}\boldsymbol{g_2})$;

1. [37] $(\boldsymbol{y_1}, \boldsymbol{y_2}) = \texttt{pack}\left(\texttt{recenter}(\boldsymbol{y}_{t_h-k:t_l-1})\right)$;

4 $L = t_l - t_h + k$;

5 $(\hat{\boldsymbol{\alpha_1}}, \hat{\boldsymbol{\alpha_2}}) = (T(\boldsymbol{W}\boldsymbol{y_1}), T(\boldsymbol{W}\boldsymbol{y_2}))$ where $T(\cdot)$ is soft-thresholding operator at threshold $\sigma\sqrt{\beta \log n}$.

$$T - (p_{t_l} - \boldsymbol{\theta}_{1:n}[t_l])^2 \leq_{(a)} \sum_{j=t_h-k}^{t_l-1} (\boldsymbol{u}^T \boldsymbol{x_j} - \boldsymbol{\theta}_{1:n}[j])^2 + \tilde{O}(B^2), \tag{B.59}$$

$$\leq_{(b)} \|\boldsymbol{\alpha_1}\|_2^2 + \|\boldsymbol{\alpha_2}\|_2^2 + \tilde{O}(B^2), \tag{B.60}$$

$$\leq_{(c)} \|\hat{\boldsymbol{\alpha_1}}\|_2^2 + \|\hat{\boldsymbol{\alpha_2}}\|_2^2 + \|\hat{\boldsymbol{\alpha_1}} - \boldsymbol{\alpha_1}\|_2^2 + \|\hat{\boldsymbol{\alpha_2}} - \boldsymbol{\alpha_2}\|_2^2 + \tilde{O}(B^2), \tag{B.61}$$

$$\leq_{(d)} \|\hat{\boldsymbol{\alpha_1}}\|_2^2 + \|\hat{\boldsymbol{\alpha_2}}\|_2^2 + \tilde{O}\left(L^{\frac{2k+1}{2k+3}} \|D^{k+1}\boldsymbol{\theta}_{t_h-k:t_l-1}\|_1^{\frac{2}{2k+3}} \sigma^{\frac{4k+4}{2k+3}}\right) + \tilde{O}(B^2), \tag{B.62}$$

$$\leq_{(e)} \frac{\sigma^2}{L} + \tilde{O}\left(L^{\frac{2k+1}{2k+3}} \|D^{k+1}\boldsymbol{\theta}_{t_h-k:t_l-1}\|_1^{\frac{2}{2k+3}} \sigma^{\frac{4k+4}{2k+3}}\right) + \tilde{O}(B^2), \tag{B.63}$$

$$= \tilde{O}(1) + \tilde{O}\left(L^{\frac{2k+1}{2k+3}} \|D^{k+1}\boldsymbol{\theta}_{t_h-k:t_l-1}\|_1^{\frac{2}{2k+3}}\right), \tag{B.64}$$

with probability atleast $1 - 2n^{3-\beta/8}$. Inequality (a) is due to lemma 138, (b) is due to orthonormality of wavelet transform matrix $\boldsymbol{W}$, (c) by triangle inequality, (d) by Corollary 140 and (e) is due to the fact that restart condition is not satisfied in the interior of a bin. □

Theorem 13. *Consider the the feedback model $y_t = \boldsymbol{\theta}_{1:n}[t] + \epsilon_t$ $t = 1, \ldots, n$ where ϵ_t are independent σ^2 subguassian noise and $|\boldsymbol{\theta}_{1:n}[t]| \leq B$. If $\beta = 24 + \frac{8\log(8/\delta)}{\log(n)}$, then with probability at least $1-\delta$,* **`Ada-VAW`** *achieves a dynamic regret of $\tilde{O}\left(n^{\frac{1}{2k+3}}\left(n^k\|D^{k+1}\boldsymbol{\theta}_{1:n}\|_1\right)^{\frac{2}{2k+3}}\right)$*

where $\tilde{O}$ hides poly-logarithmic factors of n, $1/\delta$ and constants k,σ,B that do not depend on n.

Proof. Let L_i be the length of the i^{th} bin $[t_h^{(i)}, t_l^{(i)}]$ discovered by the policy. Let

$$T_i = \sum_{t=t_h^{(i)}}^{t_l^{(i)}} (p_t - \boldsymbol{\theta}_{1:n}[t])^2. \tag{B.65}$$

From lemma 141 we have with with probability atleast $1 - 2n^{3-\beta/8}$,

$$T_i = \tilde{O}(1) + \tilde{O}\left(L_i^{\frac{2k+1}{2k+3}} \|D^{k+1}\boldsymbol{\theta}_{t_h^{(i)}-k:t_l^{(i)}-1}\|_1^{\frac{2}{2k+3}}\right) + (p_{t_l^{(i)}} - \boldsymbol{\theta}_{1:n}[t_l^{(i)}])^2 \tag{B.66}$$

$$= \tilde{O}(1) + \tilde{O}\left(L_i^{\frac{2k+1}{2k+3}} \|D^{k+1}\boldsymbol{\theta}_{t_h^{(i)}-k:t_l^{(i)}-1}\|_1^{\frac{2}{2k+3}}\right), \tag{B.67}$$

where in the last line we used the fact that ground truths are bounded by B.

Now summing the squared bias across all M bins discovered by the policy yields

$$T = \sum_{i=1}^{M} T_i, \tag{B.68}$$

$$=_{(a)} O(\tilde{M}) + \sum_{i=1}^{M} \tilde{O}\left(L_i^{\frac{2k+1}{2k+3}} \|D^{k+1}\boldsymbol{\theta}_{t_h^{(i)}-k:t_l^{(i)}-1}\|_1^{\frac{2}{2k+3}}\right), \tag{B.69}$$

$$=_{(b)} \tilde{O}\left(n^{\frac{1}{2k+3}} \|n^k D^{(k+1)}\boldsymbol{\theta}_{\mathbf{1:n}}\|_1^{\frac{2}{2k+3}}\right) + \sum_{i=1}^{M} \tilde{O}\left(L_i^{\frac{2k+1}{2k+3}} \|D^{k+1}\boldsymbol{\theta}_{t_h^{(i)}-k:t_l^{(i)}-1}\|_1^{\frac{2}{2k+3}}\right), \tag{B.70}$$

$$=_{(c)} \tilde{O}\left(n^{\frac{1}{2k+3}} \|n^k D^{(k+1)}\boldsymbol{\theta}_{\mathbf{1:n}}\|_1^{\frac{2}{2k+3}}\right) + \tilde{O}\Bigg(\left(\sum_{i=1}^{M} L_i\right)^{\frac{2k+1}{2k+3}} \times \tag{B.71}$$

$$\left(\sum_{i=1}^{M} \|D^{k+1}\boldsymbol{\theta}_{t_h^{(i)}-k:t_l^{(i)}-1}\|_1\right)^{\frac{2}{2k+3}}\Bigg), \tag{B.72}$$

$$= \tilde{O}\left(n^{\frac{1}{2k+3}} \|n^k D^{(k+1)}\boldsymbol{\theta}_{\mathbf{1:n}}\|_1^{\frac{2}{2k+3}}\right) + \tilde{O}\left(n^{\frac{1}{2k+3}} \|n^k D^{(k+1)}\boldsymbol{\theta}_{\mathbf{1:n}}\|_1^{\frac{2}{2k+3}}\right), \tag{B.73}$$

with probability atleast $1-4n^{3-\beta/8}$. Line (a) holds with probability atleast $1-2n^{3-\beta/8}$. For (b) we used lemma 137 and it holds with probability atleast $\left(1 - 2n^{3-\beta/8}\right)^2 \geq 1 - 4n^{3-\beta/8}$. For (c) we used Holder's inequality $\boldsymbol{x}^T\boldsymbol{y} \leq \|\boldsymbol{x}\|_p\|\boldsymbol{y}\|_q$ with $p = \frac{2k+3}{2k+1}$ and $q = \frac{2k+3}{2}$.

Since the variance within a bin is $\tilde{O}(\sigma^2)$ as indicated by lemma 134, when summed across all bins we get a total variance of $\tilde{O}(\sigma^2 M)$ which is $\tilde{O}\left(n^{\frac{1}{2k+3}} \|n^k D^{(k+1)} \boldsymbol{\theta_{1:n}}\|_1^{\frac{2}{2k+3}}\right)$ by lemma 137.

A trivial upperbound for T is

$$T \leq n(B^2 + \sigma^2), \tag{B.74}$$

$$= O(n). \tag{B.75}$$

Combining (B.73) (B.75) and the variance summed across all terms yields

$$T = \tilde{O}\left(\max\left\{n, n^{\frac{1}{2k+3}} \|n^k D^{(k+1)} \boldsymbol{\theta_{1:n}}\|_1^{\frac{2}{2k+3}}\right\}\right), \tag{B.76}$$

with probability atleast $1 - 4n^{3-\beta/8} - \delta/2$ where the dependence of δ in the failure probability is due to that fact that bias variance decomposition in lemma 134 holds with probability atleast $1 - \delta/2$. By setting $\beta = 24 + \frac{8\log(8/\delta)}{\log(n)}$, we get the Theorem 13. □

Remark 142. (*Specialization to* $k = 0$) When specialized to the case $k = 0$, we recover the optimal rate established in [37] for the bounded ground truth setting upto constants B and σ. When $k = 0$, our policy predicts $\frac{y_{t_h} + \ldots + y_{t-1}}{t - t_h + 2}$ at time t. This is similar to online averaging except that the denominator is now $t - t_h + 2$ instead of $t - t_h$. [37] also considers the scenario where the point-wise bound on ground truth can increase in time as $O(C_n)$. As hinted by the similarity of `Ada-VAW` with that of [37] for $k = 0$ along with the fact that our restart rule also lower-bounds the Total Variation of ground truth with high probability, it is possible to get a regret bound of $\tilde{O}(n^{1/3} C_n^{2/3} + C_n^2)$ for `Ada-VAW` in this stronger setting.

Proposition 11 (Minimax Regret)**.** *Let* $y_t = \boldsymbol{\theta}_{1:n}[t] + \epsilon_t$ *for* $t = 1, \ldots, n$ *where* $\theta_{1:n} \in TV^{(k)}(C_n)$, $|\boldsymbol{\theta}_{1:n}[t]| \leq B$ *and* ϵ_t *are iid* σ^2 *subgaussian random variables. Let* $\mathcal{A}_F$ *be the class of all forecasting strategies whose prediction at time t only depends on* $y_1, \ldots, y_{t-1}$*. Let* $\boldsymbol{s_t}$ *denote the prediction at time t for a strategy* $\boldsymbol{s} \in \mathcal{A}_F$*. Then,*

$$\inf_{\boldsymbol{s} \in \mathcal{A}_F} \sup_{\boldsymbol{\theta}_{1:n} \in TV^{(k)}(C_n)} \sum_{t=1}^{n} E\left[(\boldsymbol{s_t} - \boldsymbol{\theta}_{1:n}[t])^2\right] = \Omega\left(\min\{n, n^{\frac{1}{2k+3}} C_n^{\frac{2}{2k+3}}\}\right), \tag{3.3}$$

where the expectation is taken wrt to randomness in the strategy of the player and ϵ_t*.*

Proof. Since a batch non-parametric regression algorithm is allowed to see the entire observations ahead of time, lower bound in the batch setting directly translates to lower bound for $R_{dynamic}$. Let $\mathcal{A}_B$ be the set of all offline regression algorithms. The minimax rates of estimation of TV^k bounded sequences under squared error losses from [2] gives,

$$\inf_{\boldsymbol{s} \in \mathcal{A}_B} \sup_{\boldsymbol{\theta}_{1:n} \in TV^{(k)}(C_n)} \sum_{t=1}^{M} E\left[(\boldsymbol{s_t} - \boldsymbol{\theta}_{1:n}[t])^2\right] \tag{B.77}$$

$$= \Omega\left(n^{\frac{1}{2k+3}} C_n^{\frac{2}{2k+3}}\right). \tag{B.78}$$

□

From [20], minimax rates of estimation under squared error losses of sequences that satisfy $|\boldsymbol{\theta}_i| \leq B$ scales as $\min\{nB^2, n\sigma^2\}$. Combining the two bounds yields Proposition 11.

Proposition 18. *There exist an $O\left(((k+1)n)^2\right)$ run-time implementation of* `Ada-VAW`.

Proof. Let's describe the computational requirement at each time step. As outlined in Section 11.8 of [40], we can use Sherman-Morrison formula to compute A_t^{-1} in $O((k+1)^2)$ time. Using the same logic we can compute $(\boldsymbol{X_t}^T \boldsymbol{X_t})^{-1}$ needed by `recenter` operation incrementally in $O((k+1)^2)$ time. Re-centering operation and computation of wavelet coefficients requires $O(n)$ time per round. Since there are n rounds, the total run-time complexity becomes $O((k+1)^2 n^2)$. □

Extension to higher dimensions Consider a variational measure and the setup described in Remark 16. Let $\hat{y}_t^{(i)}$ be the prediction of instance i of `Ada-VAW` at time t. For each $i \in [d]$, we've

$$\sum_{t=1}^{n} (\hat{y}_t^{(i)} - \boldsymbol{\theta}_{1:n}[t][i])^2 = \tilde{O}\left(n^{\frac{1}{2k+3}} \Delta_i^{\frac{2}{2k+3}}\right), \tag{B.79}$$

by Theorem 13. Summing across all dimensions yields,

$$R_n = \sum_{i=1}^{d} \tilde{O}\left(n^{\frac{1}{2k+3}} \Delta_i^{\frac{2}{2k+3}}\right) \tag{B.80}$$

$$= \tilde{O}\left(d^{\frac{2k+1}{2k+3}} n^{\frac{1}{2k+3}} C_n^{\frac{2}{2k+3}}\right), \tag{B.81}$$

where the last inequality follows from applying Holder's inequality $\boldsymbol{x}^T \boldsymbol{y} \leq \|\boldsymbol{x}\|_p \|\boldsymbol{y}\|_q$ to $\sum_{i=1}^{d} 1^{\frac{2k+1}{2k+3}} \Delta_i^{\frac{2}{2k+3}}$ with norms $p = \frac{2k+3}{2k+1}$ and $q = \frac{2k+3}{2}$.

Extension to general losses Assume the interaction model in Figure 3.1. Instead of squared error losses, let the losses be f_t as discussed in Remark 17. Since f_t is gamma smooth, we have

$$f_t(b) \leq f_t(a) + f_t'(a)(b-a) + \frac{\gamma}{2}(b-a)^2. \tag{B.82}$$

Let $\hat{y}_t$ be the prediction of `Ada-VAW` at time t and $\boldsymbol{\theta}_t := \theta_{1:n}[t]$. Then regret with this loss function is

$$\sum_{t=1}^{n} f_t(\hat{y}_t) - f_t(\theta_t) \leq \sum_{t=1}^{n} \frac{\gamma}{2}(\hat{y}_t - \theta_t)^2, \tag{B.83}$$

by (B.82) and using the fact $f_t'(\theta_t) = 0$. Now the statement in Remark 17 is immediate by appealing to Theorem 13.

B.2.3 Exact sparsity

We start by the observation that an exact sparsity (i.e sparsity in the $\|\cdot\|_0$ sense) in the number of jumps of $\|D^{k+1}\boldsymbol{\theta}_{1:n}\|_0$ translates to an exact sparsity in the wavelet coefficients. This is made precise by the following lemma.

Lemma 143. *Consider a sequence with $\|D^{k+1}\boldsymbol{\theta}_{1:n}\|_0 = J$. Then both the signals $\boldsymbol{\theta}_{1:n}$ and $\tilde{\boldsymbol{\theta}}_{1:n} =$ `recenter`$(\boldsymbol{\theta}_{1:n})$ can be represented using $O(k + J\log n)$ wavelet coefficients of a CDJV system of regularity $k+1$.*

Proof. Throughout this proof when we say jumps, we refer to jumps in $\|D^{k+1}\boldsymbol{\theta}_{1:n}\|_0$. Let $L = 2^{\lceil \log_2(k+1) \rceil}$. Consider splitting the coefficients $\boldsymbol{\alpha}$ of the DWT transform into two parts: $\boldsymbol{\alpha}_{1:L}$ and $\boldsymbol{\alpha}_{L+1:n}$. By CDJV construction, the wavelets corresponding to indices $L+1,\ldots,n$ are all orthogonal to polynomials to degree atmost k. The space of polynomials of degree atmost k is contained in the span of wavelets identified by the indices $1,\ldots,L$. Though the span of the first L wavelets can also generate other waveforms which are not polynomials as well.

Notice that between two jumps, the underlying signal is a polynomial of degree atmost k. By orthogonality property discussed above, wavelet coefficients from the group $\boldsymbol{\alpha}_{L+1:n}$ assume the value zero if the support of corresponding wavelet is a region where the signal behaves as a polynomial. Since there are J jump points and each point is covered by $\log n$ wavelets by the Multi Resolution property, there can be atmost $O(J\log n)$ non zero coefficients from the group $\boldsymbol{\alpha}_{L+1:n}$.

When we subtract the best polynomial fit due to the re-centering operation, it is only going to affect the first L coefficients and keep the remaining unchanged. Hence the re-centered signal can have atmost $O(k + J\log n)$ nonzero coefficients.

□

Due to lemmas 133 and 136, the expression in the LHS of restart rule of the policy lower-bounds the TV^k distance within a bin with high probability. So if a bin lies entirely between two jumps, we do not restart with high probability as the TV^k distance is zero. This lead to the following Corollary.

Corollary 144. *Let $y_t = \boldsymbol{\theta}_t + \epsilon_t$, for $t = 1,\ldots,n$ where ϵ_t are sub-gaussian with parameter σ^2 and $\|D^{k+1}\boldsymbol{\theta}_{1:n}\|_0 = J$ with $|\boldsymbol{\theta}_t| \leq B$. Then with probability at-least $1 - 2n^{3-\beta/8}$* `Ada-VAW` *restarts $O(J)$ times.*

In the next Theorem, we characterize the optimality of soft-thresholding estimator in the exact sparsity case.

Theorem 145. *Under the setup of Corollary 144, the soft thresholding estimator whose estimates denoted by* $\hat{\boldsymbol{\alpha}}_{1:n}$ *with threshold set to* $\sigma\sqrt{\log n}$ *satisfy,*

$$\|\hat{\boldsymbol{\alpha}}_{1:n} - \boldsymbol{\theta}_{1:n}\|_2^2 = \tilde{O}(J\sigma^2), \tag{B.84}$$

with probability atleast $1 - 2n^{1-\beta/2}$ *where* $\tilde{O}$ *hides logarithmic factors of* n.

Proof. Let $\boldsymbol{\alpha}$ denote the DWT coefficients of $\boldsymbol{\theta}_{1:n}$. By Gaussian tail inequality and union bound we have $P(\sup_t |\epsilon_t| \geq \sigma\sqrt{\log n}) \leq 2n^{1-\beta/2}$. Conditioning on the event $\sup_t |\epsilon_t| \leq \sigma\sqrt{\log n}$ we are under the observation model in lemma 17 of [37]. Following the results there, with probability atleast $1 - 2n^{1-\beta/2}$ we have,

$$\|\hat{\boldsymbol{\alpha}}_{1:n} - \boldsymbol{\theta}_{1:n}\|_2^2 = \sum_{i=1}^{n} \min\left\{\boldsymbol{\alpha}[i]^2, 16\sigma^2 \log n\right\}, \tag{B.85}$$

$$= \tilde{O}(J\sigma^2), \tag{B.86}$$

where the last line follows from lemma 143 and the fact that $O(k+J\log n) = O(KJ\log n) = O(J\log n)$. □

Now using a uniform bound argument across all $O(n^2)$ bins yields the following Corollary.

Corollary 146. *Under the observation model and notations in Corollary 144 but with a subgassuan parameter* $4\sigma^2$ *when* $\boldsymbol{\theta}_{1:n}$ *is the re-centered ground truth within a bin discovered by the policy, then with probability atleast* $1 - 2n^{3-\beta/8}$

$$\|\hat{\alpha}_\delta - \alpha\|^2 = \tilde{O}(J\sigma^2). \tag{B.87}$$

With Corollaries 144 and 146, the proof of Theorem 13 can be readily adapted to give Theorem 20.

Proposition 21. *Under the interaction model in Figure 3.1, the minimax dynamic regret for forecasting sequences in* $\mathcal{E}^{k+1}(J_n)$ *is* $\Omega(J_n)$.

Proof. Let $U\{a,b,c\}$ denote a uniform sample from set $\{a,b,c\}$. Consider a ground truth sequence as follows:

1. For t=1, $\boldsymbol{\theta}_1 = U\{-B, 0, B\}$
2. For t = 2 to $J_n + 1$:
    - if $\boldsymbol{\theta}_{t-1} = -B$, $\boldsymbol{\theta}_t = U\{0, B\}$
    - if $\boldsymbol{\theta}_{t-1} = 0$, $\boldsymbol{\theta}_t = U\{-B, B\}$

- if $\boldsymbol{\theta}_{t-1} = B$, $\boldsymbol{\theta}_t = U\{-B, 0\}$

3. For $t > J_n + 1$, output $\boldsymbol{\theta}_t = \boldsymbol{\theta}_{t-1}$

Such a signal will have $\|D^{k+1}\boldsymbol{\theta}_{1:n}\|_0 \leq J_n$. Let's assume that we reveal this sequence generating process to the learner. Then the Bayes optimal algorithm will suffer a regret of $\Omega(J_n)$. □

Extension to higher dimensions Let the ground truth $\boldsymbol{\theta}_{1:n}[t] \in \mathbb{R}^d$ and let $\boldsymbol{v}_i = [\boldsymbol{\theta}_{1:n}[1][i], \ldots, \boldsymbol{\theta}_{1:n}[n][i]], \|D^{k+1}\boldsymbol{v}_i\|_1| \leq J_n, \forall i \in [d]$. Then run d instances of `Ada-VAW` where instance i is dedicated to track the sequence v_i. By appealing to Theorem 20 for each co-ordinate and summing across all d dimensions yields a regret bound of $\tilde{O}(dJ_n)$.

B.3 Adapting to lower orders of k

Though the theory of offline non parametric regression with squared error loss is well developed for the complete spectrum of function classes $TV^k(C_n)$ with $k \geq 0$, most of the practical interest is often limited to lower orders of k namely $k = 0, 1, 2, 3$ (see for eg. [4, 3]). This motivates us to design policies that can perform optimally for these lower TV orders without requiring the knowledge of k beforehand.

Let $\mathcal{E}$ be the event that $|\epsilon_t| \leq \sigma\sqrt{2\log(2n^2)}$ for all $t = 1, \ldots, n$ where ϵ_t are as presented in Figure 3.1. By using subgaussian tail inequality and a union bound across all time points, it can be shown that the event $\mathcal{E}$ happens with probability atleast $1 - \frac{1}{n}$.

The basic idea to achieve adaptivity to k is as follows:

Meta-Policy:

- Instantiate `Ada-VAW` for $k = 0, 1, 2, 3$ and run them in parallel.
- Forecast according to an Exponentially Weighted Averages (EWA) ([40]) over the predictions made by each of the instances. Set the parameter η of EWA to $1/4(B + \sqrt{2\log(2n^2)})^2$.

We condition on the event $\mathcal{E}$. The arguments in the proof of Theorem 13 still goes through even if we condition on $\mathcal{E}$. Let the dynamic regret of `Ada-VAW` for a particular value of k be the random variable $R_n^{(k)}$. The maximum possible value of $R_n^{(k)}$ is κn for some constant κ. We have,

$$\mathbb{E}[R_n^{(k)}|\mathcal{E}] = \int_{-\infty}^{\kappa n} r d\mathbb{P}(r), \tag{B.88}$$

$$\leq \gamma n^{\frac{1}{2k+3}} C_n^{\frac{2}{2k+3}} + \int_{\gamma n^{\frac{1}{2k+3}} C_n^{\frac{2}{2k+3}}}^{\kappa n} r d\mathbb{P}(r), \tag{B.89}$$

$$\leq \gamma n^{\frac{1}{2k+3}} C_n^{\frac{2}{2k+3}} + \kappa n \cdot \delta, \tag{B.90}$$

for some constant γ, where last line follows due to Theorem 13. By choosing $\delta = 1/n$ we get

$$\mathbb{E}[R_n^{(k)}|\mathcal{E}] = \tilde{O}\left(n^{\frac{1}{2k+3}} C_n^{\frac{2}{2k+3}}\right). \tag{B.91}$$

Let $\hat{y}_t$, be the output of any forecasting strategy at time t. Each expert in the meta-policy suffers a loss $(y_t - \hat{y}_t)^2$ for appropriate value of $\hat{y}_t$. Let $\theta_t := \boldsymbol{\theta}_{1:n}[t]$. we have

$$\sum_{t=1}^{n} \mathbb{E}[(y_t - \hat{y}_t)^2|\mathcal{E}] - \mathbb{E}[(y_t - \theta_t)^2|\mathcal{E}], =_{(a)} \sum_{t=1}^{n} \mathbb{E}[(\theta_t - \hat{y}_t)^2|\mathcal{E}] - \mathbb{E}[(\hat{y}_t - \theta_t)^2|\mathcal{E}]\mathcal{E}[\epsilon_t|\mathcal{E}], \tag{B.92}$$

$$= \sum_{t=1}^{n} \mathbb{E}[(\theta_t - \hat{y}_t)^2|\mathcal{E}], \tag{B.93}$$

where the last line is simply the expected dynamic regret of the strategy and line (a) is due to independence of ϵ_t with $\hat{y}_t$.

Let the dynamic regret of the meta-policy be denoted as R_{meta}. Since squared error loss $(y_t - \hat{y}_t)^2$ is exponentially concave with parameter $1/4(B + \sqrt{2\log(2n^2)})^2$, Proposition 3.1 of [40] along with (B.91) and (B.93) guarantees that,

$$\mathbb{E}[R_{meta}|\mathcal{E}] = \log 4 + \tilde{O}\left(\min_{k=0,1,2,3} n^{\frac{1}{2k+3}} \left(n^k \|D^{k+1}\theta_{1:n}\|_1\right)^{\frac{2}{2k+3}}\right) \tag{B.94}$$

Thus we see that expected dynamic regret of the meta-policy adapts to TV order k upto a additive constant of $\log 4$. This additive constant only contributes to a small $O(1/n)$ term if we consider the per round regret.

B.4 Problems with padding

In this section, we explain why some commonly used padding schemes can potentially inflate the TV^k distance of the resulting sequence.

B.4.1 Zero padding

Consider a sequence $\boldsymbol{\theta}_{1:t}$ such that best polynomial fit of this sequence is uniformly zero. Let $\boldsymbol{\gamma}$ be the zero padded version of $\boldsymbol{\theta}_{1:t}$ such that length of $\boldsymbol{\gamma}$ is a power of 2. Let $\tilde{\boldsymbol{\theta}} = [\boldsymbol{\theta}_{t-k}, \ldots, \boldsymbol{\theta}_t, 0, \ldots, 0]^T \in \mathbb{R}^{2k+2}$. We have,

$$(D^{k+1}\boldsymbol{\gamma})^T = [(D^{k+1}\boldsymbol{\theta}_{1:t})^T, (D^{k+1}\tilde{\boldsymbol{\theta}})^T, 0, 0, \ldots, 0]. \tag{B.95}$$

Due to (B.18), we have $\|\boldsymbol{\theta}_{1:t}\|_\infty = O(t^k \|D^{k+1}\boldsymbol{\theta}_{1:t}\|_1)$. Hence the existence of $\tilde{\boldsymbol{\theta}}$ term makes $\|D^{k+1}\boldsymbol{\gamma}\|_1 = O(t^k\|D^{k+1}\boldsymbol{\theta}_{1:t}\|_1)$.

B.4.2 Mirror image padding

Let $\boldsymbol{\gamma}$ be the mirror image padded version of the re-centered sequence, $\boldsymbol{\theta}_{1:t}$. i.e $\boldsymbol{\gamma} = [\theta_1, \ldots, \theta_t, \theta_t, \theta_{t-1}, \ldots]$ such that its length becomes a power of 2. Then,

$$\|D^{k+1}\boldsymbol{\gamma}\|_1 = 2\|D^{k+1}\boldsymbol{\theta}_{1:t}\|_1 + D^{k+1}[\boldsymbol{\theta}_{t-k}, \ldots, \boldsymbol{\theta}_{t-1}, \boldsymbol{\theta}_t, \boldsymbol{\theta}_t, \boldsymbol{\theta}_{t-1}, \ldots, \boldsymbol{\theta}_{t-k}]^T, \tag{B.96}$$
$$= 2\|D^{k+1}\boldsymbol{\theta}_{1:t}\|_1 + O(t^k\|D^{k+1}\boldsymbol{\theta}_{1:t}\|_1), \tag{B.97}$$

where the last line follows from (B.18).

B.5 Technical Lemmas

Lemma 147. *The procedure `CalcDetRecurse` in [134] is sound.*

Proof. We use induction on the dimension of the input square matrix.

Base case: when $d = 3$. Assume that $e[0][0]$ is non-zero. Let the matrix be given by

$$\boldsymbol{X} = \begin{bmatrix} e_{00} & e_{01} & e_{02} \\ e_{10} & e_{11} & e_{12} \\ e_{20} & e_{21} & e_{22} \end{bmatrix} \tag{B.98}$$

The idea is to convert $\boldsymbol{X}$ to an upper triangular matrix. Define:

$$\boldsymbol{Y} = \begin{bmatrix} 1 & \frac{e_{01}}{e_{00}} & \frac{e_{02}}{e_{00}} \\ e_{10} & e_{11} & e_{12} \\ e_{20} & e_{21} & e_{22} \end{bmatrix} \tag{B.99}$$

So that $\det(\boldsymbol{Y}) = \frac{\det(\boldsymbol{X})}{e_{00}}$. Applying elementary row operations we get

$$\det(\boldsymbol{Y}) = \begin{vmatrix} 1 & \frac{e_{01}}{e_{00}} & \frac{e_{02}}{e_{00}} \\ 0 & e_{11} - e_{10}\frac{e_{01}}{e_{00}} & e_{12} - e_{10}\frac{e_{01}}{e_{00}} \\ 0 & e_{21} - e_{20}\frac{e_{01}}{e_{00}} & e_{22} - -e_{20}\frac{e_{01}}{e_{00}} \end{vmatrix} \tag{B.100}$$

The inner loop in the procedure `CalcDetRecurse` computes the determinant of the inner 2×2 sub-matrix by considering the numerator of the fractional terms. Hence the value v return by the recursive call is $\det(\boldsymbol{Y}[1:][1:])e_{00}^2$. So $\det(X) = e_{00}\frac{v}{e_{00}^2} = \frac{v}{e_{00}}$. This is precisely the value returned by the procedure after the final division loop.

When e_{00} is zero, we can swap it with the row whose first element is non-zero and apply the arguments above. If such a swap is not possible, the procedure correctly recognizes the determinant as zero.

Inductive case: Assume that procedure is sound for matrices upto dimension n. Now define $\boldsymbol{Y}$ as before to set the element e_{00} to one. By similar arguments we obtain that value v returned by the recursive call is $\det(\boldsymbol{Y}[1:][1:])e_{00}^{n}$. Thus we obtain $\det(\boldsymbol{X}) = \frac{v}{e_{00}^{n-1}}$. This division is performed at the final loop of the procedure.

Here also when e_{00} is zero, the swapping argument similar to the base case can be applied.

□

Consider OLS fit on the inputs $(\boldsymbol{x_1}, y_1), \ldots, \boldsymbol{x_t}, y_t)$ where the features $\boldsymbol{x_j} = [1, j, \ldots, j^m]^T$ and the responses obey $\max_{i=1,\ldots,t} |y_i| = B$. Let the design matrix be

$$\boldsymbol{X}_t = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_t]^T. \tag{B.101}$$

Lemma 148. $\det(\boldsymbol{X_t}^T \boldsymbol{X_t})$ *is a polynomial in* t *with degree atmost* $(k+1)^2$*.*

Proof. The procedure `CalcDegreeOfDet` in [134] can be used to upperbound the degree of determinant. It assumes that while doing the subtractions in procedure `CalcDetRecurse`, the highest degree terms in the corresponding polynomials do not cancel out.

Let $m = k + 1$. Observe that $\boldsymbol{X}_t^T \boldsymbol{X_t}$ can be compactly written as

$$\boldsymbol{X}_t^T \boldsymbol{X_t} = \begin{bmatrix} S_0(t) & S_1(t) & \ldots & S_{m-1}(t) \\ \vdots & \vdots & \ddots & \vdots \\ S_{m-1}(t) & S_m(t) & \ldots & S_{2m-2}(t) \end{bmatrix}, \tag{B.102}$$

where $S_p(t) = \sum_{n=1}^{t} n^p$.

Let's run procedure `CalcDegreeOfDet` on an $m \times m$ matrix $\boldsymbol{D}$ of degrees arising from $\boldsymbol{X}_t^T \boldsymbol{X}_t$ as below.

$$\boldsymbol{D} = \begin{bmatrix} 1 & 2 & \ldots & m \\ \vdots & \vdots & \ddots & \vdots \\ m & m+1 & \ldots & 2m-1 \end{bmatrix} \tag{B.103}$$

Let's define a seed sequence $\{s\}_i$ as the sequence of numbers that can be found the main diagonal of a given matrix, excluding the element at the bottom right corner. The seed sequnece of $\boldsymbol{D}$ is simply $1, 3, \ldots, 2m-3$. Let T_i be the element at index $(0,0)$ for the matrix in the i^{th} recursive call. Note that $T_1 = 1$. Tracing the steps through the recursion we get

$$T_2 = s_2 + T_1 \tag{B.104}$$
$$T_3 = s_2 + T_2 + T_1 \tag{B.105}$$
$$\vdots$$
$$T_{m-1} = s_{k-1} + T_{k-2} + \ldots + T_1 \tag{B.106}$$

In $m-1$ calls, we will be left with a 2×2 matrix whose entries are

$$\begin{bmatrix} T_{m-1} & 1+T_{m-1} \\ 1+T_{m-1} & 2+T_{m-1} \end{bmatrix} \tag{B.107}$$

Now let's start with the winding up procedure. There are $k-3$ wind-ups that need to be performed. Let u_t be the wound up value from the t^{th} winding up step. We have,

$$u_{m-2} = 2 + 2T_{m-1} - T_{m-2} \tag{B.108}$$

$$u_{m-3} = u_{m-2} - 2T_{m-3} \tag{B.109}$$

$$u_{m-4} = u_{m-3} - 3T_{m-4} \tag{B.110}$$

$$\vdots$$

$$u_1 = u_2 - (m-2)T_1 \tag{B.111}$$

Note that u_1 is the final output produced by the topmost call to `CalcDegreeOfDet` procedure. These systems can be unrolled to get

$$u_1 = 2 + 2T_{m-1} - (T_{m-2} + 2T_{m-3} + \ldots + (m-2)T_1 \tag{B.112}$$

$$= 2 + s_{m-1} + \sum_{i=1}^{m-1} s_i \tag{B.113}$$

Now using explicit expressions for seed sequence $\{s\}_i$ we get

$$u_1 = 2 + 2m - 3 + (m-1)^2 \tag{B.114}$$

$$= m^2 \tag{B.115}$$

$$= (k+1)^2 \tag{B.116}$$

□

Lemma 149. *Let $S_p(t)$ be a polynomial in t defined as $S_p(t) = \sum_{n=1}^{t} n^p$ where p is a non-negative integer. Then,*

$$(-1)^{p-1} S_p(t-1) = S_p(-t) \tag{B.117}$$

Proof. For $a(t) = \frac{t(t+1)}{2}$, Faulhaber's formula states that

$$\sum_{n=1}^{t} n^p = \sum_{i=1}^{(p-1)/2} c_i a(t)^{(p+1)/2}, \tag{B.118}$$

when p is odd and

$$\sum_{n=1}^{t} n^p = \frac{t+0.5}{p+1} \sum_{i=1}^{p/2} (i+1) c_i a(t)^{p/2}, \tag{B.119}$$

when p is even. The the explicit form of c_i can be expressed in terms of Bernoulli numbers.

Note that $a(-t) = a(t-1)$. Substituting this in the formulas yields the lemma. □

Lemma 150. *For a universal constant $H(m)$ that depends only on $m = k+1$,*

$$\det(\boldsymbol{X}_t^T \boldsymbol{X}_t) = H(m)\, t^m \prod_{i=2}^{m} \left(t^2 - (i-1)^2\right)^{m-i+1} \tag{B.120}$$

Proof. The strategy is to characterize the roots of determinant. For brevity let's denote $\boldsymbol{Z}_t = \boldsymbol{X}_t^T \boldsymbol{X}_t$. Observe that

$$\boldsymbol{Z}_t = \sum_{i=1}^{t} \boldsymbol{x}_i \boldsymbol{x}_i^T, \tag{B.121}$$

where $x_i = [1, \ldots, i^{m-1}]$. Each update $\boldsymbol{x}_i \boldsymbol{x}_i^T$ increases the rank by atmost 1. After m such updates $\boldsymbol{X}_m$ becomes a square Vandermonde matrix formed by the sequence $\{1, 2, \ldots, m\}$. Since all of the elements in the sequence are distinct $\boldsymbol{X}_m$ is full rank and so is $\boldsymbol{Z}_m$. This implies that each such update $\boldsymbol{x}_i \boldsymbol{x}_i^T$ for $i \leq m$ increased the rank by exactly one.

We can view the equation (B.121) as a quantity that evolves in time. For $1 \leq i \leq m-1$, there exists $m-i$ rows in $\boldsymbol{Z}_i$ that are linearly dependent. This means $t = i$ is a root of $\det(\boldsymbol{Z}_t)$ with multiplicity $(m-i)$. By defining $\boldsymbol{x}_0 = [0, \ldots, 0]^T$ for the initial case $t = 0$, all the rows are simply zeroes and multiplicity of the root $t = 0$ is m. Thus we have established that $t^m \prod_{i=2}^{m} (t-(i-1))^{m-i+1}$ is a sub-expression of $\det(\boldsymbol{Z}_t)$.

Let's view $\boldsymbol{Z}_t$ as a function of t with $t \in \mathbb{R}$ as displayed in (B.102). Put $t = -t'$ in (B.102). Then we have,

$$\boldsymbol{Z}(t') = \begin{bmatrix} S_0(-t') & S_1(-t') & \ldots & S_{m-1}(-t') \\ \vdots & \vdots & \ddots & \vdots \\ S_{m-1}(-t') & S_m(-t') & \ldots & S_{2m-2}(-t') \end{bmatrix}. \tag{B.122}$$

Hence showing $t' = a$ is a root of $\boldsymbol{Z}(t')$ implies that $t = -a$ is a root of $\boldsymbol{Z}_t$. We have

$$\det(\boldsymbol{Z}(t')) = (-1)^m \begin{vmatrix} -S_0(-t') & -S_1(-t') & \ldots & -S_{m-1}(-t') \\ \vdots & \vdots & \ddots & \vdots \\ -S_{m-1}(-t') & -S_m(-t') & \ldots & -S_{2m-2}(-t') \end{vmatrix} \tag{B.123}$$

Consider

$$\det(\tilde{\boldsymbol{Z}}(t')) = \begin{vmatrix} -S_0(-t') & -S_1(-t') & \dots & -S_{m-1}(-t') \\ \vdots & \vdots & \ddots & \vdots \\ -S_{m-1}(-t') & -S_m(-t') & \dots & -S_{2m-2}(-t') \end{vmatrix} \tag{B.124}$$

When t' is a non-negative integer, lemma 149 implies that the elements in the matrix above are result of the summation:

$$\sum_{i=0}^{t'-1}(-i)^p = (-1)^p S_p(t'-1) \tag{B.125}$$

$$= -S_p(-t'), \tag{B.126}$$

where we adopt the convention $0^0 = 1$.

Thus we have,

$$\tilde{\boldsymbol{Z}}(t') = \sum_{i=1}^{t'} \boldsymbol{x'}_i \boldsymbol{x'}_i^T, \tag{B.127}$$

where $\boldsymbol{x'}_i = [1, -(i-1), \dots, (-(i-1))^{m-1}]$. Let $\boldsymbol{X'}_t = [\boldsymbol{x'}_1, \dots, \boldsymbol{x'}_t]^T$.

After m updates, we have that $\boldsymbol{X'}_m$ is a square Vandermonde matrix defined by the sequence $\{0, -1, \dots, -(m-1)\}$. Since each of the elements are distinct, this a full rank matrix and so each update $\boldsymbol{x'}_i \boldsymbol{x'}_i^T$ for $i \le m$ increased the rank by exactly one leading to $\tilde{\boldsymbol{Z}}(m)$ being full rank.

Using similar arguments as above we see that $t' = i$ is a root of $\det(\tilde{\boldsymbol{Z}}(t'))$ with multiplicity $(m-i)$. This in turn imply that $t = -i$ is a root of $\det(\boldsymbol{Z}_t)$ with multiplicity $(m-i)$. Now we have established that $t^m \prod_{i=2}^{m} (t^2 - (i-1)^2)^{m-i+1}$ is a sub-expression of $\det(\boldsymbol{Z}_t)$. By lemma 148 we conclude that we have found all roots of the determinant and no further terms depending t can be there.

□

Remark 151. We conjecture that the universal constant $H(m)$ in lemma 150 is the determinant of Hilbert matrix of order m.

Definition 152. Let $\boldsymbol{H}(t)$ be a square matrix with each entry $r_{ij}(t) = \frac{n_{ij}(t)}{d_{ij}(t)}$ for polynomials $n_{ij}(t)$ and $d_{ij}(t)$. We say $r_{ij}(t)$ is *Hilbert-like* if $r_{ij}(t) = O\left(\frac{1}{t^{i+j-1}}\right)$ for all i, j.

Lemma 153. *All the elements of* $\left(\boldsymbol{X}_t^T \boldsymbol{X}_t\right)^{-1}$ *are Hilbert-like when* $t \ge m = k+1$.

Proof. Computation of inverse is essentially a computation of determinants of the matrix and its minors. Each element (i, j) of an inverse matrix is a rational function with numerator being determinant of minor M_{ij} and denominator being the determinant of the original symmetric matrix.

Let $\boldsymbol{Z}_t = \boldsymbol{X}_t^T \boldsymbol{X}_t$ When $t \geq m$ we have from lemma 150 that $\det(\boldsymbol{Z}_t) = \Omega(t^{m^2})$. So it is sufficient to show that $\det(M_{ij}$ is $O(t^{m^2+1-i-j})$. The strategy we follow is same of that in lemma 148.

We follow a 1 based indexing. Since $\boldsymbol{Z}_t$ is symmetric, it is enough to compute the minors when $1 \leq i \leq j \leq m$.

case 1: Consider $\det(M_{ij})$ when $1 < i < j < m-1$. Following the same notations as in the prood of lemma 150, after $m-2$ calls to `CalDegreeOfDet` we end up with a matrix below.

$$\boldsymbol{F} = \begin{bmatrix} T_{m-2} & 1 + T_{m-2} \\ 1 + T_{m-2} & 2 + T_{m-2} \end{bmatrix} \tag{B.128}$$

The corresponding seed sequence $\{s\}_i$ is $\{1, 3, 5, \ldots, 2i-3, 2i, 2i+2, \ldots, 2j-2, 2j+1, 2j+3, \ldots, 2m-3\}$. The jumps in the progression is attributed to the deletion of row i and column j for obtaining minor M_{ij}.

The final output u_1, from the topmost call to `CalDegreeOfDet` is then given by

$$\begin{aligned} u_1 &= s_{m-2} + \sum_{i=1}^{m-2} s_i && \text{(B.129)} \\ &= 2 + (2m-3) + (i-1)^2 + (j-i)(j+i-1) + (m+j-1)(m-j-1), && \text{(B.130)} \\ &= m^2 + 1 - i - j. && \text{(B.131)} \end{aligned}$$

So $\det(M_{ij})$ is $O(t^{m^2+1-i-j})$ where the constant in the big-oh only dependents on m.

case 2: $(1 < i < j = m-1)$. After $m-2$ recursion calls we get the matrix below.

$$\boldsymbol{F} = \begin{bmatrix} T_{m-2} & 2 + T_{m-2} \\ 1 + T_{m-2} & 3 + T_{m-2} \end{bmatrix} \tag{B.132}$$

The seed sequence $\{s\}_i$ is $\{1, 3, \ldots, 2i-3, 2i, \ldots, 2j-2\}$. So

$$\begin{aligned} u_1 &= 3 + s_{m-2} + \sum_{i=1}^{m-2} s_i, && \text{(B.133)} \\ &= 3 + (2m-4) + (i-1)^2 + (j-i)(j+i-1), && \text{(B.134)} \\ &= m^2 + 1 - i - j. && \text{(B.135)} \end{aligned}$$

So $\det(M_{ij})$ is $O(t^{m^2+1-i-j})$.

case 3: $(1 < i = j < k-1)$.

The seed sequence $\{s\}_i$ is $\{1, 3, \ldots, 2i-3, 2i+1, \ldots, 2m-3\}$. At the last step we get a matrix as in equation (B.128). Hence,

$$u_1 = 2 + s_{m-2} + \sum_{i=1}^{m-2} s_i, \tag{B.136}$$
$$= 2 + (2m-3) + (i-1)^2 + (m-i-1)(2i+1+m-i-2), \tag{B.137}$$
$$= m^2 + 1 - i - j. \tag{B.138}$$

So $\det(M_{ij})$ is $O(t^{m^2+1-i-j})$.
case 4: $(i = j = m-1)$.
The seed sequence $\{s\}_i$ is $\{1, 3, \ldots, 2i-3\}$. At the last step we get a matrix below.

$$\boldsymbol{F} = \begin{bmatrix} T_{m-2} & 2 + T_{m-2} \\ 2 + T_{m-2} & 3 + T_{m-2} \end{bmatrix} \tag{B.139}$$

So,

$$u_1 = 4 + s_{m-2} + \sum_{i=1}^{m-2} s_i, \tag{B.140}$$
$$= 2 + (2i-3) + (i-1)^2, \tag{B.141}$$
$$= m^2 + 1 - i - j. \tag{B.142}$$

So $\det(M_{ij})$ is $O(t^{m^2+1-i-j})$.
case 5: $(i = j = m)$.
The seed sequence $\{s\}_i$ is $\{1, 3, \ldots, 2m-5\}$. At the last step we get a matrix as in equation (B.128).

$$u_1 = 2 + s_{m-2} + \sum_{i=1}^{m-2} s_i, \tag{B.143}$$
$$= 2 + (2m-5) + (m-2)^2, \tag{B.144}$$
$$= m^2 + 1 - i - j. \tag{B.145}$$

So $\det(M_{ij})$ is $O(t^{m^2+1-i-j})$.
case 6: $(i = j = 1)$.
The seed sequence $\{s\}_i$ is $\{3, \ldots, 2m-3\}$. At the last step we get a matrix as in equation (B.128).

$$u_1 = 2 + s_{m-2} + \sum_{i=1}^{m-2} s_i, \tag{B.146}$$
$$= 2 + (2m-3) + (m-1)^2 - 1, \tag{B.147}$$
$$= m^2 + 1 - i - j. \tag{B.148}$$

So $\det(M_{ij})$ is $O(t^{m^2+1-i-j})$.

case 7: $(1 < i < k-1 < j = k)$.

The seed sequence $\{s\}_i$ is $\{1, \ldots, 2i-3, 2i, \ldots, 2m-4\}$. At the last step we get a matrix as in equation (B.128).

So,

$$u_1 = 2 + s_{m-2} + \sum_{i=1}^{m-2} s_i, \tag{B.149}$$
$$= 2 + (2m-4) + (i-1)^2 + (m-i-1)(2i+k-i-2), \tag{B.150}$$
$$= m^2 + 1 - i - j. \tag{B.151}$$

So $\det(M_{ij})$ is $O(t^{m^2+1-i-j})$.

case 8: $(i = 1, j = m)$.

The seed sequence $\{s\}_i$ is $\{2, \ldots, 2m-4\}$. At the last step we get a matrix as in equation (B.128).

So,

$$u_1 = 2 + s_{m-2} + \sum_{i=1}^{m-2} s_i, \tag{B.152}$$
$$= 2 + (2m-4) + (m-2)(2+m-3), \tag{B.153}$$
$$= m^2 + 1 - i - j. \tag{B.154}$$

So $\det(M_{ij})$ is $O(t^{m^2+1-i-j})$.

case 9: $(i = 1, j = m-1)$.

The seed sequence $\{s\}_i$ is $\{2, \ldots, 2m-4\}$. At the last step we get a matrix as in equation (B.132).

So,

$$u_1 = 3 + s_{m-2} + \sum_{i=1}^{m-2} s_i, \tag{B.155}$$
$$= 3 + (2m-4) + (m-2)(2+m-3), \tag{B.156}$$
$$= m^2 + 1 - i - j. \tag{B.157}$$

So $\det(M_{ij})$ is $O(t^{m^2+1-i-j})$.

case 10: $(i = 1 < j < m-1)$.

The seed sequence $\{s\}_i$ is $\{2, \dots, 2j-2, 2j+1, \dots, 2m-3\}$. At the last step we get a matrix as in equation (B.128).

So,

$$u_1 = 2 + s_{m-2} + \sum_{i=1}^{m-2} s_i, \tag{B.158}$$
$$= 2 + (2m-3) + (j-1)(2+j-2) + (m-j-1)(2j+1+(m-j-2)), \tag{B.159}$$
$$= m^2 + 1 - i - j. \tag{B.160}$$

So $\det(M_{ij})$ is $O(t^{m^2+1-i-j})$.

□

With the above lemma, the following Corollary can be readily verified.

Corollary 154. *When $\boldsymbol{\theta}_{1:n}$ is such that $\|\boldsymbol{\theta}_{1:n}\|_\infty \leq B = O(1)$, we have $\| \left(\boldsymbol{X}_t^T \boldsymbol{X}_t\right)^{-1} \boldsymbol{X}_t^T \boldsymbol{\theta}_{1:n}\|_2 = O(1)$.*

Appendix C

Supplementary Materials for Chapter 4

C.1 Proofs of Technical Results

For the sake of clarity, we present a sequence of lemmas and sketch how to chain them to reach the main result in Section C.1.1. This is followed by proof of all lemmas in Section C.1.2 and finally the proof of Theorem 27 in Section C.1.3.

C.1.1 Proof strategy for Theorem 27

We first show that ALIGATOR suffers logarithmic regret against any expert in the pool $\mathcal{E}$ during its awake period. Then we exhibit a particular partition of the underlying TV bounded function such that number of chunks in the partition is $O(n^{1/3}C_n^{2/3})$. Following this, we cover each chunk with atmost $\log n$ experts and show that each expert in the cover suffers a $\tilde{O}(1)$ estimation error. The Theorem then follows by summing the estimation error across all chunks.

Some notations. In the analysis thereafter, we will use the following notations. Let $\tilde{\sigma} = \sigma\sqrt{2\log(4n/\delta)}$, $R_\sigma = 16(B+\tilde{\sigma})^2$ and $\mathcal{T}(I) = \{t \in [n] : i_t \in I\}$ for any $I \in \mathcal{I}|_{[n]}$, where $\mathcal{I}|_{[n]}$ is defined according to the terminology in Section 4.2.1. Let $\theta_t := f(x_{i_t})$.

First, we show that ALIGATOR is competitive against any expert in the pool $\mathcal{E}$.

Lemma 155. *For any interval $I \in \mathcal{I}|_{[n]}$ such that $\mathcal{T}(I)$ is non-empty, the predictions made by* ALIGATOR *$\hat{y}_t$ satisfy*

$$\sum_{t\in\mathcal{T}(I)} (\hat{y}_t - \theta_t)^2 \le \frac{e-1}{3-e} \sum_{t\in\mathcal{T}(I)} (\mathcal{A}_I(t) - \theta_t)^2 + \frac{\log(n\log n)R_\sigma + 2R_\sigma^2 \log(2n\log n/\delta)}{3-e}, \tag{C.1}$$

with probability atleast $1-\delta$.

Corollary 156. *Let $\mathbb{S} = \{P_1, \ldots, P_M\}$ be an arbitrary ordered set of consecutive intervals in $[n]$. For each $i \in [n]$ let $\mathcal{U}_i$ be the set containing elements of the GC that covers the interval P_i according to Proposition 23. Denote $\lambda := \frac{\log(n \log n) R_\sigma + 2R_\sigma^2 \log(2n \log n/\delta)}{3-e}$. Then* Aligator *forecasts $\hat{y}_t$ satisfy*

$$\sum_{t=1}^{n} (\hat{y}_t - \theta_t)^2 \leq \min_{\mathbb{S}} \sum_{i=1}^{M} \sum_{I \in \mathcal{U}_i} \mathbf{1}\{|\mathcal{T}(I)| > 0\} \left(\frac{e-1}{3-e} \sum_{t \in \mathcal{T}(I)} (\mathcal{A}_I(t) - \theta_t)^2 + \lambda \right), \tag{C.2}$$

with probability atleast $1 - \delta$.

The minimum across all partitions in the Corollary above hints to the novel ability of Aligator to incur potentially very low estimation errors.

Next, we proceed to exhibit a partition of the set of exogenous variables queried by the adversary that will eventually lead to the minimax rate of $\tilde{O}(n^{1/3} C_n^{2/3})$. The existence of such partitions is a non-trivial matter.

Lemma 157. *Let $\mathbb{S} = \{x_{k_1} < \ldots, < x_{k_m}\} \subseteq \mathcal{X}$ be the exogenous variables queried by the adversary over n rounds where each $k_i \in [n]$. Denote $\theta^{(i)} := f(x_{k_i})$ and $p(i) := \#\{t : x_{i_t} = x_{k_i}\}$ for each $i \in [m]$. Denote $[x_i, x_j] := \{x_{k_i}, x_{k_{i+1}}, \ldots, x_{k_j}\}$. For any $[x_i, x_j] \subseteq \mathbb{S}$, define $V(x_i, x_j) = \sum_{k=i}^{j-1} |\theta^{(i)} - \theta^{(i+1)}|$. There exists a partitioning $\mathcal{P} = \{[x_1, x_{r_1}], [x_{r_1+1}, x_{r_2}], \ldots, [x_{r_{M-1}+1}, x_m]\}$ of $\mathbb{S}$ that satisfies*

1. *For any $[x_i, x_j] \in \mathcal{P} \setminus \{[x_{r_{M-1}+1}, x_m]\}$, $V(x_i, x_j) \leq \frac{B}{\sqrt{\sum_{k=i}^{j} p(k)}}$.*
2. *$V(x_{r_{M-1}+1}, x_{m-1}) \leq \frac{B}{\sqrt{\sum_{k=r_{M-1}+1}^{m-1} p(k)}}$.*
3. *Number of partitions $M \leq \max\{3n^{1/3} C_n^{2/3} B^{-2/3}, 1\}$.*

The next lemma controls the estimation error incurred by an expert during its awake period.

Lemma 158. *Let $\{\underline{x}, < \ldots, < \bar{x}\}$ be the exogenous variables queried by the adversary over n rounds in an arbitrary interval $I \in \mathcal{I}|_{[n]}$. Then with probability atleast $1 - \delta$*

$$\sum_{t \in \mathcal{T}(I)} (\theta_t - \mathcal{A}_I(t))^2 \leq 2V(\underline{x}, \bar{x})^2 |\mathcal{T}(I)| + 2\sigma^2 \log(2n^3 \log n/\delta) \log(|\mathcal{T}(I)|), \tag{C.3}$$

where $V(\cdot, \cdot)$ is defined as in Lemma 157.

To prove Theorem 27, our strategy is to apply Corollary 156 to the partition in Lemma 157. By the construction of the GC, each chunk in the partition can be covered using atmost $\log n$ intervals. Now consider the estimation error incurred by an expert corresponding to one such interval. Due to statements 1 and 2 in Lemma 157 the

$V(\underline{x}, \bar{x})^2|\mathcal{T}(I)|$ term of error bound in Lemma 158 can be shown to $O(1)$. When summed across all intervals that cover a chunk, the total estimation error within a chunk becomes $\tilde{O}(1)$. Now appealing to statement 3 of Lemma 157, we get a total error of $\tilde{O}(n^{1/3}C_n^{2/3})$ when the error is summed across all chunks in the partition.

C.1.2 Omitted Lemmas and Proofs

Lemma 159. *Let $\mathcal{V}$ be the event that for all $t \in [n]$, $|\epsilon_t| \leq \sigma\sqrt{2\log(4n/\delta)}$. Then $\mathbb{P}(\mathcal{V}) \geq 1 - \delta/2$.*

Proof. By gaussian tail inequality, we have for a fixed t $P(|\epsilon_t| > \sigma\sqrt{2\log(4n/\delta)}) \leq \delta/2n$. By taking a union bound we get $P(|\epsilon_t| \geq \sigma\sqrt{2\log(4n/\delta)}) \leq \delta/2$ for all $t \in [n]$. □

Some notations. In the analysis thereafter, we will use the following filtration.

$$\mathcal{F}_j = \sigma((i_1, y_{i_1}), \ldots, (i_{j-1}, y_{i_{j-1}})). \tag{C.4}$$

Let's denote $\mathbb{E}_j[\cdot] := \mathbb{E}[\cdot|\mathcal{F}_j]$ and $\mathrm{Var}_j[\cdot] := \mathrm{Var}[\cdot|\mathcal{F}_j]$. Let $\theta_j = f(x_{i_j})$ and $\tilde{\sigma} = \sigma\sqrt{2\log(4n/\delta)}$. Let $R_\sigma = 16(B + \tilde{\sigma})^2$ and $\mathcal{T}(I) = \{t \in [n] : i_t \in I\}$

Lemma 160. *(Freedman type inequality, [135]) For any real valued martingale difference sequence $\{Z_t\}_{t=1}^T$ with $|Z_t| \leq R$ it holds that,*

$$\sum_{t=1}^T Z_t \leq \eta(e-2)\sum_{t=1}^T \mathrm{Var}_t[Z_t] + \frac{R\log(1/\delta)}{\eta}, \tag{C.5}$$

with probability atleast $1 - \delta$ for all $\eta \in [0, 1/R]$.

Lemma 161. *For any $j \in [n]$, we have*

1. $\mathbb{E}_j[(y_j - \mathcal{A}_I(j))^2 - (y_j - \theta_j)^2|\mathcal{V}] = \mathbb{E}_j[(\mathcal{A}_I(j) - \theta_j)^2|\mathcal{V}]$.
2. $\mathrm{Var}_j[(y_j - \mathcal{A}_I(j))^2 - (y_j - \theta_j)^2|\mathcal{V}] \leq R_\sigma\mathbb{E}_j[(\mathcal{A}_I(j) - \theta_j)^2|\mathcal{V}]$.

Proof. We have,

$$\mathbb{E}_j[(y_j - \mathcal{A}_I(j))^2 - (y_j - \theta_j)^2|\mathcal{V}] =_{(a)} \mathbb{E}_j[(\mathcal{A}_I(j) - \theta_j)^2|\mathcal{V}] - 2\mathbb{E}_j[\epsilon_j|\mathcal{V}]\mathbb{E}_j[(\mathcal{A}_I(j) - \theta_j)|\mathcal{V}], \tag{C.6}$$

$$= \mathbb{E}_j[(\mathcal{A}_I(j) - \theta_j)^2|\mathcal{V}], \tag{C.7}$$

where line (a) is due to the independence of ϵ_j with the past. Since $(\mathcal{A}_I(j) + \theta_j - 2y_j)^2 \leq 16(B + \tilde{\sigma})^2$ under the event $\mathcal{V}$, it holds that

$$\mathrm{Var}_j[(y_j - \mathcal{A}_I(j))^2 - (y_j - \theta_j)^2|\mathcal{V}] \leq \mathbb{E}_j[(y_j - \mathcal{A}_I(j))^2 - (y_j - \theta_j)^2|\mathcal{V}]^2, \tag{C.8}$$

$$\leq 16(B + \tilde{\sigma})^2\mathbb{E}_j[(\mathcal{A}_I(j) - \theta_j)^2|\mathcal{V}]. \tag{C.9}$$

□

Lemma 162. *For any interval $I \in \mathcal{I}$, it holds with probability atleast $1-\delta$ that*

1. $\sum_{j \in \mathcal{T}(I)} (y_j - \mathcal{A}_I(j))^2 - (y_j - \theta_j)^2 \leq \sum_{j \in \mathcal{T}(I)} (e-1)(\mathcal{A}_I(j) - \theta_j)^2 + R_\sigma^2 \log(2n \log n/\delta)$,
2. $\sum_{j \in \mathcal{T}(I)} (y_j - \hat{y}_j)^2 - (y_j - \theta_j)^2 \geq \sum_{j \in \mathcal{T}(I)} (3-e)(\hat{y}_j - \theta_j)^2 - R_\sigma^2 \log(2n \log n/\delta)$.

Proof. Define $Z_j = (y_j - \mathcal{A}_I(j))^2 - (y_j - \theta_j)^2 - (\mathcal{A}_I(j) - \theta_j)^2$.

Condition on the event $\mathcal{V}$ that $|\epsilon_t| \leq \sigma\sqrt{2\log(4n/\delta)}. \forall t \in [n]$ which happens with probability atleast $1-\delta/2$ by Lemma 159. By Lemma 161, we have $\{Z_j\}_{j \in \mathcal{T}(I)}$ is a martingale difference sequence and $|Z_j| \leq 16(B+\tilde{\sigma})^2 = R_\sigma$. Note that once we condition on the filtration $\mathcal{F}_j$, there is no randomness remaining in the terms $(\mathcal{A}_I(j) - \theta_j)^2$ and $(\hat{y}_j - \theta_j)^2$. Hence $\mathbb{E}_j[(\mathcal{A}_I(j) - \theta_j)^2 | \mathcal{V}] = (\mathcal{A}_I(j) - \theta_j)^2$ and $\mathbb{E}_j[(\hat{y}_j - \theta_j)^2 | \mathcal{V}] = (\hat{y}_j - \theta_j)^2$. Using Lemma 160 and taking $\eta = 1/R_\sigma$ we get,

$$\sum_{j \in \mathcal{T}(I)} (y_j - \mathcal{A}_I(j))^2 - (y_j - \theta_j)^2 \leq \sum_{j \in \mathcal{T}(I)} (e-1)(\mathcal{A}_I(j) - \theta_j)^2 + R_\sigma^2 \log(4n \log n/\delta), \tag{C.10}$$

with probability atleast $1 - \delta/(4n \log n)$ for a fixed expert A_I. Taking a union bound across all $O(n \log n)$ experts in $\mathcal{E}$ leads to,

$$\mathbb{P}\left(\sum_{j \in \mathcal{T}(I)} (y_j - \mathcal{A}_I(j))^2 - (y_j - \theta_j)^2 \geq \sum_{j \in \mathcal{T}(I)} (e-1)(\mathcal{A}_I(j) - \theta_j)^2 + R_\sigma^2 \log(2n \log n/\delta) | \mathcal{V} \right) \tag{C.11}$$

$$\leq \delta/4, \tag{C.12}$$

for any expert $\mathcal{A}_I$.

By similar arguments on the martingale difference sequence $(\hat{y}_j - \theta_j)^2 - (y_j - \hat{y}_j)^2 - (y_j + \theta_j)^2$, it can be shown that

$$\mathbb{P}\left(\sum_{j \in \mathcal{T}(I)} (y_j - \hat{y}_j)^2 - (y_j - \theta_j)^2 \leq \sum_{j \in \mathcal{T}(I)} (3-e)(\hat{y}_j - \theta_j)^2 - R_\sigma^2 \log(2n \log n/\delta) | \mathcal{V} \right) \leq \delta/4, \tag{C.13}$$

for any interval $I \in \mathcal{I}|_{[n]}$. Taking union bound across the previous two bad events and multiplying the probability of noise boundedness event $\mathcal{V}$ leads to the lemma. □

Lemma 155. *For any interval $I \in \mathcal{I}|_{[n]}$ such that $\mathcal{T}(I)$ is non-empty, the predictions made by* ALIGATOR *$\hat{y}_t$ satisfy*

$$\sum_{t \in \mathcal{T}(I)} (\hat{y}_t - \theta_t)^2 \leq \frac{e-1}{3-e} \sum_{t \in \mathcal{T}(I)} (\mathcal{A}_I(t) - \theta_t)^2 + \frac{\log(n \log n) R_\sigma + 2R_\sigma^2 \log(2n \log n/\delta)}{3-e}, \tag{C.1}$$

with probability atleast $1-\delta$.

Proof. Condition on the event $\mathcal{V}$. Then the losses $f_t(x) = (y_t - x)^2$ are $\frac{1}{4(B+\sigma\sqrt{\log(2n/\delta)})^2} := \eta$ exp-concave [136, 40]. Since we pass $\eta \cdot f_t(x)$ as losses to SAA in ALIGATOR, Lemma 24 gives

$$\sum_{t \in \mathcal{T}(I)} -\log\left(\sum_{J \in A_t} w_{t,J} e^{-\eta f_t(\mathcal{A}_J(t))}\right) - \eta f_t(\mathcal{A}_I(t)) \leq \log(n \log n). \tag{C.14}$$

By η exp-concavity of $f_t(x)$, we have

$$-\log\left(\sum_{J \in A_t} w_{t,J} e^{-\eta f_t(\mathcal{A}_J(t)))}\right) \geq \eta f_t\left(\sum_{J \in A_t} w_{t,J} \mathcal{A}_I(t)\right), \tag{C.15}$$

$$= \eta f_t(\hat{y}_t). \tag{C.16}$$

Combining (C.14) and (C.16) gives,

$$\sum_{t \in \mathcal{T}(I)} f_t(\hat{y}_t) - f_t(\mathcal{A}_I(t)) \leq \frac{\log(n \log n)}{\eta}, \tag{C.17}$$

$$\leq \log(n \log n) R_\sigma. \tag{C.18}$$

So,

$$\sum_{t \in \mathcal{T}(I)} (y_t - \hat{y}_t)^2 - (y_t - \theta_t)^2 \leq \sum_{t \in \mathcal{T}(I)} (y_t - \mathcal{A}_I(t))^2 - (y_t - \theta_t)^2 + \log(n \log n) R_\sigma, \tag{C.19}$$

Now invoking Lemma (162) followed by a trivial rearrangement completes the proof. □

Lemma 157. *Let* $\mathbb{S} = \{x_{k_1} < \ldots, < x_{k_m}\} \subseteq \mathcal{X}$ *be the exogenous variables queried by the adversary over* n *rounds where each* $k_i \in [n]$. *Denote* $\theta^{(i)} := f(x_{k_i})$ *and* $p(i) := \#\{t : x_{i_t} = x_{k_i}\}$ *for each* $i \in [m]$. *Denote* $[x_i, x_j] := \{x_{k_i}, x_{k_{i+1}}, \ldots, x_{k_j}\}$. *For any* $[x_i, x_j] \subseteq \mathbb{S}$, *define* $V(x_i, x_j) = \sum_{k=i}^{j-1} |\theta^{(i)} - \theta^{(i+1)}|$. *There exists a partitioning* $\mathcal{P} = \{[x_1, x_{r_1}], [x_{r_1+1}, x_{r_2}], \ldots, [x_{r_{M-1}+1}, x_m]\}$ *of* $\mathbb{S}$ *that satisfies*

1. *For any* $[x_i, x_j] \in \mathcal{P} \setminus \{[x_{r_{M-1}+1}, x_m]\}$, $V(x_i, x_j) \leq \frac{B}{\sqrt{\sum_{k=i}^{j} p(k)}}$.
2. $V(x_{r_{M-1}+1}, x_{m-1}) \leq \frac{B}{\sqrt{\sum_{k=r_{M-1}+1}^{m-1} p(k)}}$.
3. *Number of partitions* $M \leq \max\{3n^{1/3} C_n^{2/3} B^{-2/3}, 1\}$.

Proof. We provide below a constructive proof. Consider the following scheme of partitioning $\mathbb{S}$.

1. Set pings $= p(1), \mathrm{TV} = 0, M = 1$.
2. Start a partition from x_1.
3. For i = 2 to m
   (a) If $\mathrm{TV} + |\theta^{(i)} - \theta^{(i-1)}| > \frac{B}{\sqrt{\text{pings}+p(i)}}$:
      i. pings $= p(i), \mathrm{TV} = 0$ // start a new bin (partition) from position x_i.
      ii. $M = M + 1$ // increase the bin counter
   (b) Else:
      i. pings $=$ pings $+ p(i), \mathrm{TV} = \mathrm{TV} + |\theta^{(i)} - \theta^{(i-1)}|$

Statements 1 and 2 of the Lemma trivially follows from the strategy. Next, we provide an upper bound on number of bins M spawned by the above scheme. Let $[x_1, x_{r_1}], [x_{r_1+1}, x_{r_2}], \ldots, [x_{r_{M-1}, x_{r_M}}]$ be the partition of $\mathbb{S}$ discovered by the above scheme.

Define the quantity $\mathrm{TV}_1 := \sum_{i=1}^{r_1} |\theta^{(i)} - \theta^{(i+1)}|$ associated with bin 1. Similarly define $\mathrm{TV}_2, \ldots, \mathrm{TV}_{M-1}$ for other bins.

Define $N(1) = \sum_{i=1}^{r_1+1} p(i)$. Similarly define $N(2), \ldots N(M-1)$. It is immediate that $\sum_{i=1}^{M-1} N(i) \leq 2n$.

We have,

$$C_n \geq \sum_{i=1}^{M-1} \mathrm{TV}_i, \tag{C.20}$$

$$\geq_{(1)} \sum_{i=1}^{M-1} \frac{B}{\sqrt{N(i)}}, \tag{C.21}$$

$$\geq_{(2)} \frac{(M-1)^{3/2} \cdot B}{\sqrt{2n}}, \tag{C.22}$$

where (1) follows from step 3(a) of the partitioning scheme and (2) is due to convexity of $1/\sqrt{x}, x > 0$ and applying Jensen's inequality. Rearranging and noting that $M - 1 \geq M/2$, when $M > 1$, we obtain

$$M \leq 3n^{1/3} C_n^{2/3} B^{-2/3}. \tag{C.23}$$

Note that when $C_n = 0$, M will remain 1 as a result of the partitioning scheme.

□

Lemma 158. *Let $\{\underline{x}, < \ldots, < \bar{x}\}$ be the exogenous variables queried by the adversary over n rounds in an arbitrary interval $I \in \mathcal{I}|_{[n]}$. Then with probability atleast $1 - \delta$*

$$\sum_{t \in \mathcal{T}(I)} (\theta_t - \mathcal{A}_I(t))^2 \leq 2V(\underline{x}, \bar{x})^2 |\mathcal{T}(I)| + 2\sigma^2 \log(2n^3 \log n/\delta) \log(|\mathcal{T}(I)|), \tag{C.3}$$

where $V(\cdot,\cdot)$ is defined as in Lemma 157.

Proof. Let $q(t) = \sum_{s=1}^{t-1} \mathbf{1}\{i_s \in I\}$. Assume $q(t) > 0$. Fix a particular expert $\mathcal{A}_I$ and a time t. Since $y_t \sim N(\theta_t, \sigma^2)$ by gaussian tail inequality we have,

$$\mathbb{P}\left(\left|\frac{\sum_{s=1}^{t-1}(y_s - \theta_s)\mathbf{1}\{i_s \in I\}}{\sum_{s=1}^{t-1}\mathbf{1}\{i_s \in I\}}\right| \geq \frac{\sigma}{\sqrt{q(t)}}\sqrt{\log\left(\frac{2n^3\log n}{\delta}\right)}\right) \leq \frac{\delta}{(n^3 \log n)}. \tag{C.24}$$

Applying a union bound across all time points and all experts implies that for any expert $\mathcal{A}_I$ and $t \in \mathcal{T}(I)$ with q(t) ¿ 0,

$$\left|\mathcal{A}_I(t) - \frac{\sum_{s=1}^{t-1}\theta_s\mathbf{1}\{i_s \in I\}}{q(t)}\right| \leq \frac{\sigma}{\sqrt{q(t)}}\sqrt{\log\left(\frac{2n^3\log n}{\delta}\right)} \tag{C.25}$$

with probability atleast $1 - \delta$.

Now adding and subtracting θ_t inside the $|\cdot|$ on LHS and using $|a - b| \geq |a| - |b|$ yields,

$$|\mathcal{A}_I(t) - \theta_t| \leq \left|\theta_t - \frac{\sum_{s=1}^{t-1}\theta_s\mathbf{1}\{i_s \in I\}}{q(t)}\right| + \frac{\sigma}{\sqrt{q(t)}}\sqrt{\log\left(\frac{2n^3\log n}{\delta}\right)}. \tag{C.26}$$

Hence,

$$\sum_{t\in\mathcal{T}(I)}(\theta_t - \mathcal{A}_I(t))^2 \leq_{(a)} \sum_{t\in\mathcal{T}(I)} 2\left(\theta_t - \frac{\sum_{s=1}^{t-1}\theta_s\mathbf{1}\{i_s \in I\}}{q(t)}\right)^2 + 2\frac{\sigma^2}{q(t)}\log\left(\frac{2n^3\log n}{\delta}\right) \tag{C.27}$$

$$\leq \sum_{t\in\mathcal{T}(I)} 2\left(\theta_t - \frac{\sum_{s=1}^{t-1}\theta_s\mathbf{1}\{i_s \in I\}}{q(t)}\right)^2 + 2\sigma^2\log(|\mathcal{T}(I)|)\log\left(\frac{2n^3\log n}{\delta}\right), \tag{C.28}$$

with probability atleast $1 - \delta$. In (a) we used the relation $(a + b)^2 \leq 2a^2 + 2b^2$.

Further we have,

$$\sum_{t\in\mathcal{T}(I)} 2\left(\theta_t - \frac{\sum_{s=1}^{t-1}\theta_s\mathbf{1}\{i_s \in I\}}{q(t)}\right)^2 \leq 2V(\underline{x}, \bar{x})^2|\mathcal{T}(I)|. \tag{C.29}$$

Combining (C.28) and (C.29) completes the proof.

□

C.1.3 Proof of the main result: Theorem 27

Proof. Throughout the proof we carry forward all notations used in Lemmas 157 and 158.

We will apply Corollary 156 to the partition in Lemma 157. Take a specific partition $[x_i, x_j] \in \mathcal{P}$ with $j \neq m$. Consider a set of indices $F = \{k_i, k_i + 1, \dots, k_j\}$ of consecutive natural numbers between k_i and k_j. By Proposition 23 F can be covered using elements in $\mathcal{I}|_{[n]}$. Let this cover be $\mathcal{U}$. For any $I \in \mathcal{U}$, we have

$$\sum_{t \in T(I)} (\theta_t - \mathcal{A}_I(t))^2 \leq_{(a)} 2V(\underline{x}, \bar{x})^2 |\mathcal{T}(I)| + 2\sigma^2 \log(2n^3 \log n/\delta) \log(|\mathcal{T}(I)|) \tag{C.30}$$

$$\leq 2V(\underline{x}, \bar{x})^2 |\mathcal{T}(F)| + 2\sigma^2 \log(2n^3 \log n/\delta) \log(|\mathcal{T}(I)|) \tag{C.31}$$

$$\leq_{(b)} 2B^2 + 2\sigma^2 \log(2n^3 \log n/\delta) \log(n), \tag{C.32}$$

, with probability atleast $1-\delta$. Step (a) is due to Lemma 158 and (b) is due to statement 1 of Lemma 157.

Using Lemma 155 and a union bound on the bad events in Lemmas 155 and 158 yields,

$$\sum_{t \in \mathcal{T}(I)} (\hat{y}_t - \theta_t)^2 \leq \frac{e-1}{3-e} \left(2B^2 + 2\sigma^2 \log(2n^3 \log n/\delta) \log(n)\right) + \lambda, \tag{C.33}$$

with probability atleast $1 - 2\delta$ and λ is as defined in Corollary 156. Due to the property of exponentially decaying lengths as stipulated by Proposition 23, there are only atmost $2 \log |F| \leq 2 \log n$ intervals in $\mathcal{U}$. So,

$$\sum_{t \in \mathcal{T}(F)} (\hat{y}_t - \theta_t)^2 \leq 2 \log n \left(\frac{e-1}{3-e} \left(2B^2 + 2\sigma^2 \log(2n^3 \log n/\delta) \log(n)\right) + \lambda \right). \tag{C.34}$$

Similar bound can be obtained for the last bin $[x_{r_{M-1}+1}, x_m]$ in $\mathcal{P}$. There are two cases to consider. In case 1, we consider the scenario when $V(x_{r_{M-1}+1}, x_m)$ obeys relation 1 of Lemma 157. Then the analysis is identical to the one presented above. In case 2, we consider the scenario when $V(x_{r_{M-1}+1}, x_{m-1})$ obeys relation 2 of Lemma 157 while $V(x_{r_{M-1}+1}, x_m)$ doesn't. Then the error incurred within the interior $[x_{r_{M-1}+1}, x_{m-1}]$ can be bounded as before. To bound the error at last point, we only need to bound the error of expert that performs mean estimation of iid gaussians. It is well known that the cumulative squared error for this problem is atmost $\sigma^2 \log(n/\delta)$ with probability atleast $1 - \delta$.

By Lemma 157, $|\mathcal{P}| = \max\{3n^{1/3} C_n^{2/3} B^{-2/3}, 1\}$. Hence the total error summed across all partitions in $\mathcal{P}$ becomes,

$$\begin{aligned}\sum_{t=1}^{n}(\hat{y}_t - \theta_t)^2 \leq 2\log n \Bigg(& \frac{e-1}{3-e}\left(4n^{1/3}C_n^{2/3}B^{4/3}\right. \\ & \left. + 4\sigma^2 \log(2n^3 \log n/\delta)\log(n) n^{1/3}C_n^{2/3}B^{-2/3}\right)\Bigg) \\ & + 4\log(n)\frac{e-1}{3-e}\lambda n^{1/3}C_n^{2/3}B^{-2/3} \\ & + 2\log(n)\left(\frac{e-1}{3-e}\left(2B^2 + 2\sigma^2\log(2n^3\log n/\delta)\log(n)\right) + \lambda\right) \\ & + \sigma^2\log(n/\delta), \\ & = \tilde{O}(n^{1/3}C_n^{2/3}),\end{aligned} \tag{C.35}$$

with probability atleast $1-2\delta$. A change of variables from $2\delta \to \delta$ completes the proof. As a closing note, we remark that the aggressive dependence of B in (C.35) on cases when B is too small can be dampened by using a threshold of $\frac{1}{\sqrt{\text{pings}+p(i)}}$ in the partition scheme presented in proof of Lemma 157. □

C.2 Excluded details in Experimental section

Waveforms. The waveforms shown in Fig. C.1 and C.2 are borrowed from [56]. Note that both functions exhibit spatially inhomogeneous smoothness behaviour.

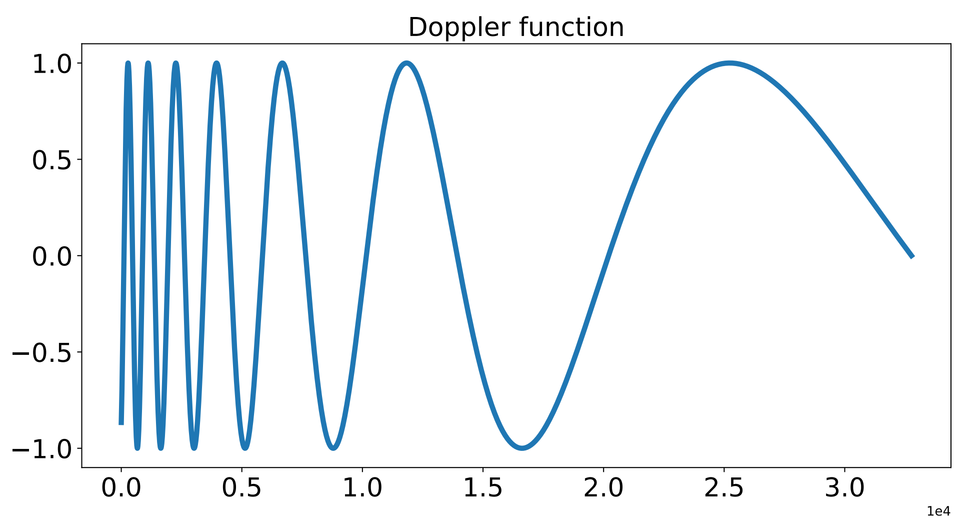


Figure C.1: *Doppler function, TV = 27*

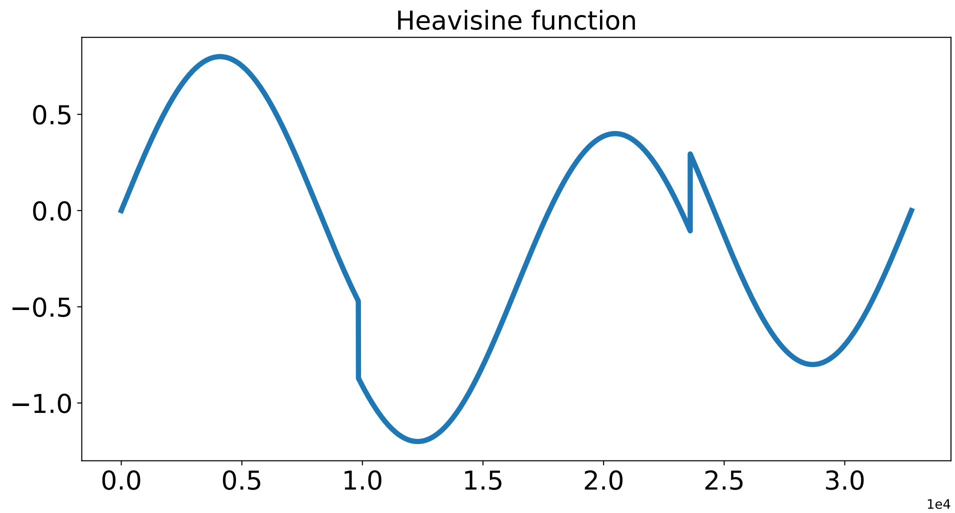


Figure C.2: *Heavisine function, TV = 7.2*

Hyper-parameter search. Initially we used a grid search on an exponential grid to realize that the optimal λ across all experiments fall within the range [0.125, 8]. Then we used a fine-tuned grid
$[0.125, 0.25, 0.5, 0.75, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 7.5, 8, 10, 12, 14, 16]$ to search for the final hyper parameter value. For Aligator *(heuristics)*, we searched for different noise levels in order to find best learning rate. We set search method as Loss/$(para * (\sigma^2 + \sigma^2/m))$. As Fig. C.7 shows, $para = 2$ is found to provide good results across all signals we consider.

Padding for wavelets. For "wavelet" estimator in Fig. 4.6, when data length is not a power of 2, we used the reflect padding mode in [137], though the results are similar for other padding schemes.

Experiments on Real Data. We follow the experimental setup described in Section 4.5. A qualitative comparison of the forecasts for the state of New Mexico, USA is illustrated in Fig. C.8. The average RMSE of Aligator and Holt ES for all states in USA is reported in Table C.1.

State	**RMSE Aligator**	**RMSE Holt ES**	**% improvement**
New Jersey	411.87	546.89	24.69
Ohio	216.24	280.24	22.84
Florida	1330.33	1671.23	20.4
Alabama	290.71	362.13	19.72
New York	876.35	1054.2	16.87
Rhode Island	85.11	98.23	13.35
Vermont	7.59	8.7	12.76

Kansas	142.17	162.16	12.33
New Mexico	57.88	65.99	12.29
Connecticut	206.79	235.6	12.23
California	1456.48	1650.25	11.74
Pennsylvania	258.21	290.6	11.14
Kentucky	145.61	163.59	10.99
New Hampshire	25.16	27.99	10.1
Minnesota	161.41	179.12	9.89
Michigan	315.86	350.24	9.82
Hawaii	30.24	33.18	8.86
Texas	1510.42	1650.73	8.5
South Dakota	56.83	61.8	8.04
Utah	118.97	128.96	7.74
Alaska	17.54	18.96	7.52
Washington	188.8	202.74	6.88
North Carolina	265.74	284.47	6.58
Nebraska	98.49	105.41	6.56
Montana	28.31	30.28	6.51
Missouri	224.51	239.9	6.42
Iowa	205.77	219.28	6.16
District of Columbia	33.58	35.74	6.04
Virginia	194.29	206.44	5.89
Nevada	159.88	168.92	5.35
Wyoming	16.43	17.25	4.73
Georgia	493.93	518.27	4.7
Oregon	55.48	58.21	4.68
Louisiana	562.89	590.49	4.67
Maryland	209.95	218.22	3.79
Illinois	475.49	492.09	3.37
West Virginia	37.34	38.63	3.33
Delaware	64.1	66.26	3.26
Tennessee	384.55	396.95	3.12
Arizona	481.91	493.73	2.39
South Carolina	271.87	277.42	2.0
Idaho	93.83	95.44	1.68
Colorado	142.58	144.53	1.35
Mississippi	206.67	209.11	1.16
Arkansas	164.83	164.88	0.03
Massachusetts	302.79	301.8	-0.32
Oklahoma	151.82	146.65	-3.41

Indiana	185.1	178.2	-3.73
North Dakota	42.14	40.49	-3.92
Wisconsin	219.04	203.37	-7.15
Maine	14.59	13.37	-8.36

Table C.1: Average RMSE across all states in USA. The experimental setup and computation of error metrics are as described in Section 4.5. The % improvement tab is computed as follows. Let x_1 and x_2 be the RMSE of ALIGATOR and Holt ES respectively. Then % improvement = $(x_2 - x_1)/\max\{x_1, x_2\}$.

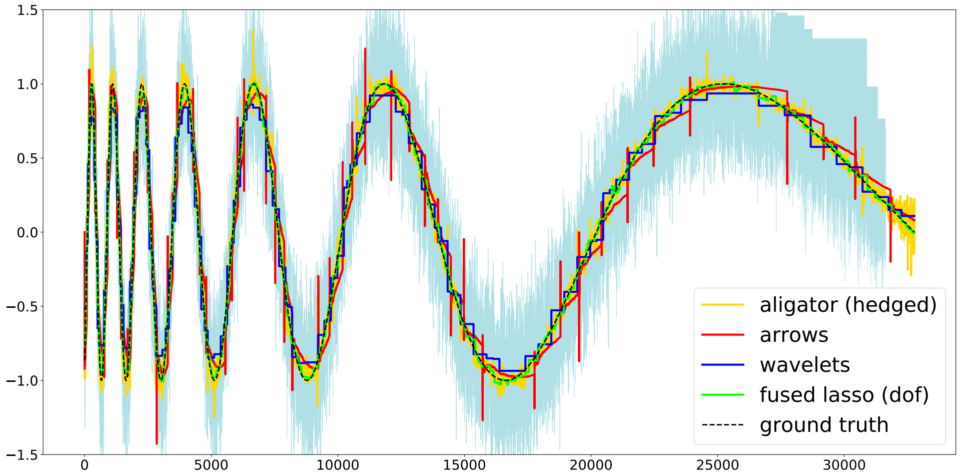


Figure C.3: *Fitted signals for Doppler function with noise level* $\sigma = 0.35$

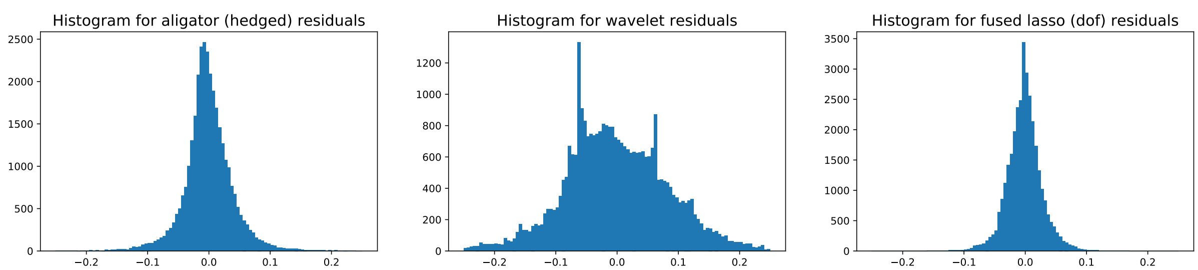


Figure C.4: *Histogram of residuals for various algorithms when run on Doppler function with noise level* $\sigma = 0.35$. *Note that they are residuals w.r.t to ground truth.* ALIGATOR *incurs lower bias than wavelets. The bias incurred by dof fused lasso is roughly comparable to* ALIGATOR *while former is more compute intensive.*

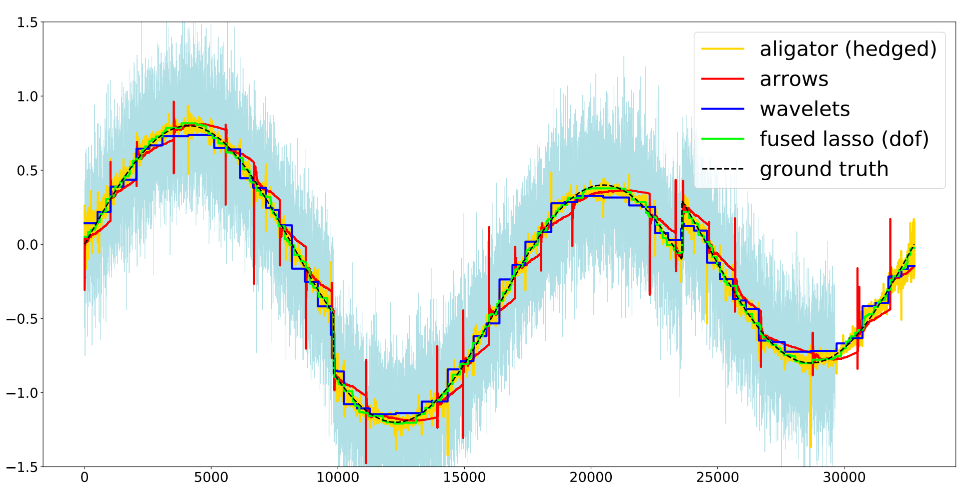


Figure C.5: *Fitted signals for Heavisine function with noise level* $\sigma = 0.35$

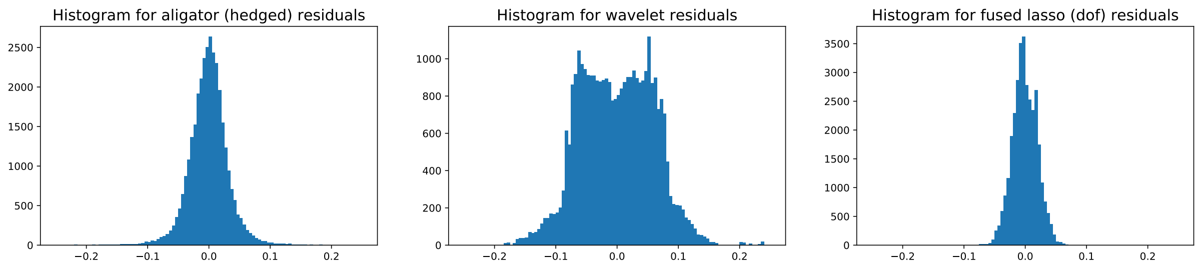


Figure C.6: *Histogram of residuals for various algorithms when run on Heavisine function with noise level $\sigma = 0.35$. Note that they are residuals w.r.t to ground truth.* Aligator *incurs lower bias than wavelets. The bias incurred by dof fused lasso is roughly comparable to* Aligator *while former is more compute intensive.*

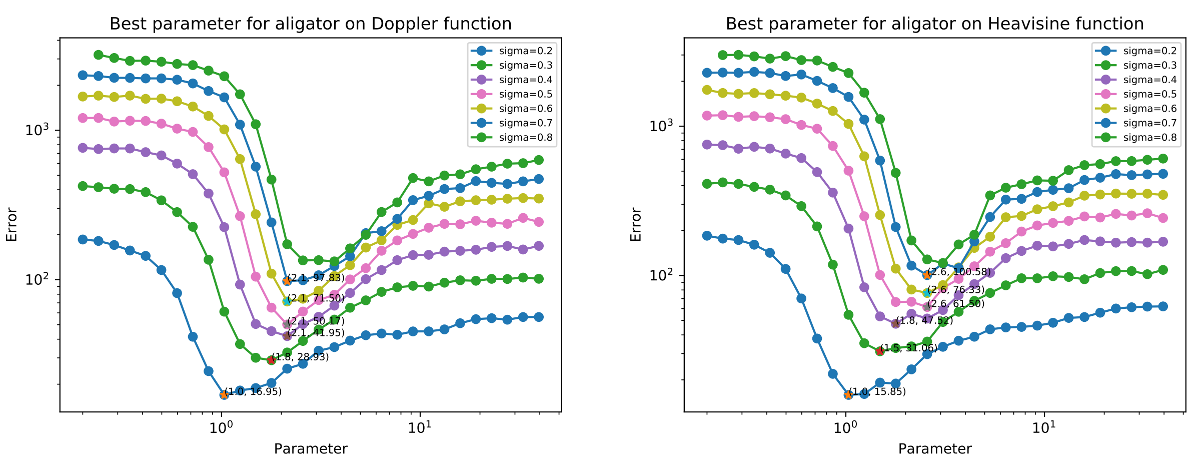


Figure C.7: *Hyper-parameter search for learning rate in* Aligator *(heuristics).*

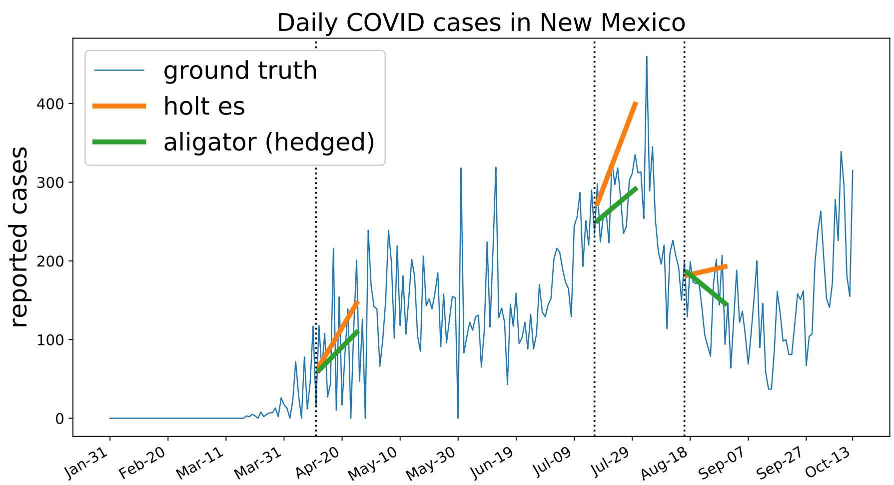


Figure C.8: *A demo on forecasting COVID cases based on real world data. We display the two weeks forecasts of hedged* Aligator *and Holt ES, starting from the time points identified by the dotted lines. Both the algorithms are trained on a 2 month data prior to each dotted line. We see that hedged* Aligator *detects changes in trends more quickly than Holt ES. Further, hedged* Aligator *attains a 12% reduction in the average RMSE from that of Holt ES (see Table C.1).*

Appendix D

Supplementary Materials for Chapter 5

D.1 Preliminaries

In this section, we recall the Follow-the-Leading-History (FLH) algorithm from [23] along with some basic definitions.

Definition 163. (**Strong convexity**) Loss functions f_t are said to be H strongly convex in the domain $\mathcal{D}$ if it satisfies

$$f_t(\boldsymbol{y}) \geq f_t(\boldsymbol{x}) + (\boldsymbol{y} - \boldsymbol{x})^T \nabla f_t(\boldsymbol{x}) + \frac{H}{2}\|\boldsymbol{x} - \boldsymbol{y}\|^2,$$

for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{D}$.

FLH enjoys the following guarantee against any base learner.

Proposition 164. *[23] Suppose the loss functions are exp-concave with parameter α. For any interval $I = [r, s]$ in time, the algorithm FLH Fig.D.1 with learning rate $\zeta = \alpha$ gives $O(\alpha^{-1}(\log r + \log |I|))$ regret against the base learner in hindsight.*

Definition 165. ([24]) An algorithm is said to be Strongly Adaptive (SA) if for every contiguous interval $I \subseteq [n]$, the static regret incurred by the algorithm is $O(\text{poly}(\log n)\Gamma^*(|I|))$ where $\Gamma^*(|I|)$ is the value of minimax static regret incurred in an interval of length $|I|$.

It is known from [77] that OGD and ONS achieves static regret of $O(\log n)$ and $O(d \log n)$ for strongly convex and exp-concave losses respectively. Hence in view of Proposition 164 and Definition 165, we can conclude that:

- FLH with OGD as base learners is an SA algorithm for strongly convex losses.
- FLH with ONS as base learners is an SA algorithm for exp-concave losses. (We treat dimension d as a constant problem parameter and consider minimaxity only wrt n.)

FLH: inputs - Learning rate ζ and n base learners $E^1, \ldots, E^n$

1. For each t, $v_t = (v_t^{(1)}, \ldots, v_t^{(t)})$ is a probability vector in $\mathbb{R}^t$. Initialize $v_1^{(1)} = 1$.
2. In round t, set $\forall j \leq t$, $x_t^j \leftarrow E^j(t)$ (the prediction of the j^{th} bas learner at time t). Play $x_t = \sum_{j=1}^t v_t^{(j)} x_t^{(j)}$.
3. After receiving f_t, set $\hat{v}_{t+1}^{(t+1)} = 0$ and perform update for $1 \leq i \leq t$:

$$\hat{v}_{t+1}^{(i)} = \frac{v_t^{(i)} e^{-\zeta f_t(x_t^{(i)})}}{\sum_{j=1}^t v_t^{(j)} e^{-\zeta f_t(x_t^{(j)})}} \tag{D.1}$$

4. Addition step - Set $v_{t+1}^{(t+1)}$ to $1/(t+1)$ and for $i \neq t+1$:

$$v_{t+1}^{(i)} = (1 - (t+1)^{-1})\hat{v}_{t+1}^{(i)} \tag{D.2}$$

Figure D.1: FLH algorithm

We have the following guarantee on runtime.

Proposition 166. *[23] Let ρ be the per round run time of base learners and r_n be the static regret suffered by the base learners over n rounds. Then FLH procedure has a runtime of $O(\rho n)$ per round. To improve the runtime one can use AFLH procedure from [23] that incurs $O(\rho \log n)$ runtime overhead per round and suffers $O(r_n \log n)$ static regret in any interval.*

D.2 Proofs for Section 5.1

We start by providing an example of a scenario where λ in Lemma 36 can scale linearly with n.

Example 167. Consider the $\mathcal{TV}(C_n)$ class with $C_n = 1$ and $n \geq 6$. Let the offline optimal be given by the step sequence $u_1 = \ldots = u_{(n/2)-1} = 0$ and $u_{n/2} = \ldots = u_n = 1$. Our aim is to generate a sequence of labels y_t such that this sequence $\boldsymbol{u}$ is indeed the offline optimal in the class $\mathcal{TV}(1)$ along with the property that the optimal dual variable λ scales linearly with the horizon n.

Clearly we must have $s_{(n/2)-1} = 1$. For some appropriate parameter ϵ, consider the following sign assignment:

- $s_{(n/2)-2} = 1 - \epsilon, s_{(n/2)-3} = 1 - 2\epsilon, \ldots, s_1 = 1 - ((n/2) - 2)\epsilon$,
- $s_{n-1} = \epsilon, s_{n-2} = 2\epsilon, \ldots, s_{n/2} = (n/2)\epsilon$.

By setting $\epsilon = 2/n$ for $n \geq 6$, we get a consistent sign assignment because $s_t \in [-1, 1]$ for all $1 \leq t \leq (n/2) - 2$ which corresponds to the portion where $u_t = 0$; $s_{(n/2)-1} = 1$; and $s_t \in [-1, 1]$ for all $n/2 \leq t \leq n - 1$ which corresponds to the portion where $u_t = 1$.

By taking $\lambda = n/2$ the adversary can generate labels y_t according to the stationarity condition in Lemma 36 as follows:

- $y_1 = -2,$
- $y_t = -1, \text{ for } 2 \leq t \leq (n/2) - 1,$
- $y_{n/2} = 1,$
- $y_t = 2, \text{ for } (n/2) + 1 \leq t \leq n.$

Since the TV of the sequence $\boldsymbol{u}$ is 1, the complementary slackness is also satisfied. Thus we conclude that if the labels $y_t \in [-2, 2]$ are generated as above, the offline optimal sequence in $\mathcal{TV}(1)$ class is given by the step sequence $\boldsymbol{u}$. Furthermore, the optimal dual variable $\lambda = n/2$ scales linearly with the horizon.

Lemma 36. ***(characterization of offline optimal)*** *Consider the following convex optimization problem (where $\tilde{z}_1, ..., \tilde{z}_{n-1}$ are introduced as dummy variables)*

$$\min_{\tilde{u}_1,\ldots,\tilde{u}_n,\tilde{z}_1,\ldots,\tilde{z}_{n-1}} \quad \frac{1}{2}\sum_{t=1}^{n}(y_t - \tilde{u}_t)^2 \tag{5.7a}$$

$$\text{s.t.} \quad \tilde{z}_t = \tilde{u}_{t+1} - \tilde{u}_t \; \forall t \in [n-1], \tag{5.7b}$$

$$\sum_{t=1}^{n-1} |\tilde{z}_t| \leq C_n \tag{5.7c}$$

Let $u_1, \ldots, u_n, z_1, \ldots, z_{n-1}$ be the optimal primal variables and let $\lambda \geq 0$ be the optimal dual variable corresponding to the last constraint (5.7c)*. By the KKT conditions, we have*

- ***stationarity:*** $y_t = u_t - \lambda(s_t - s_{t-1})$*, where $s_t \in \partial|z_t|$ (a subgradient). Specifically, $s_t = \mathrm{sign}(u_{t+1} - u_t)$ if $|u_{t+1} - u_t| > 0$ and s_t is some value in $[-1, 1]$ otherwise. For convenience of notations later, we also define $s_n = s_0 = 0$.*
- ***complementary slackness:*** $\lambda\left(\sum_{t=2}^{n} |u_t - u_{t-1}| - C_n\right) = 0.$

Proof. We can form the Lagrangian of the optimization problem as:

$$\mathcal{L}(\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{z}}, \tilde{\boldsymbol{v}}, \tilde{\lambda}) = \frac{1}{2}\sum_{t=1}^{n}(y_t - \tilde{u}_t)^2 + \tilde{\lambda}\left(\sum_{t=1}^{n-1}|\tilde{z}_t| - C_n\right) + \sum_{t=1}^{n-1}\tilde{v}_t(\tilde{u}_{t+1} - \tilde{u}_t - \tilde{z}_t), \tag{D.3}$$

for dual variables $\tilde{\lambda} > 0$ and $\tilde{\boldsymbol{v}} \in \mathbb{R}^{n-1}$ unconstrained. Let the $(\boldsymbol{u}, \boldsymbol{z}, \boldsymbol{v}, \lambda)$ be the optimal primal and dual variables. By stationarity conditions, we have

$$u_t - y_t = v_t - v_{t-1}, \tag{D.4}$$

where we take $v_0 = v_n = 0$ and

$$v_t = \lambda s_t \tag{D.5}$$

Combining the above two equations and the complementary slackness rule yields the lemma. □

Lemma 38. ***(key partition)*** *Initialize $\mathcal{Q} \leftarrow \Phi$. Starting from time 1, spawn a new bin $[i_s, i_t]$ whenever $\sum_{j=i_s+1}^{i_t+1} |u_j - u_{j-1}| > B/\sqrt{n_i}$, where $n_i = i_t - i_s + 2$. Add the spawned bin $[i_s, i_t]$ to $\mathcal{Q}$. Consider the following post processing routine.*

1. *Initialize $\mathcal{P} \leftarrow \Phi$.*
2. *For $i \in [|\mathcal{Q}|]$:*
   - *if $u_{i_t} = u_{i_t+1}$:*
     - *(a) Let p be the largest time point with $u_{p:i_t}$ being constant and let q be the smallest time point with $u_{i_t+1:q}$ being constant.*
     - *(b) Add bin $[i_s, p-1]$ to $\mathcal{P}$.*
     - *(c) If $(i+1)_t > q$ then add $[p, q]$ to $\mathcal{P}$ and set $(i+1)_s \leftarrow q+1$.*
     - *(d) Goto Step 2.*
   - *Add $[i_s, i_t]$ to $\mathcal{P}$. Goto Step 2.*

Let $M := |\mathcal{P}|$. We have $M = O\left(1 \vee n^{1/3} C_n^{2/3} B^{-2/3}\right)$. Further for any bin $[i_s, i_t] \in \mathcal{P}$, it holds that $\sum_{j=i_s+1}^{i_t} |u_j - u_{j-1}| \leq B/\sqrt{n_i}$ where $n_i = i_t - i_s + 1$.

Proof. Let's use the notation $TV[a, b]$ to denote the TV incurred by the optimal solution sequence in the interval $[a, b]$. Let $\mathcal{Q} = \{[\underline{t}_1, \bar{t}_1], \ldots, [\underline{t}_N, \bar{t}_N]\}$ with $\underline{t}_1 := 1$ and $\bar{t}_N := n$. Let $n_j := \bar{t}_j - \underline{t}_j + 1$ We have,

$$\sum_{j=1}^{N-1} TV[\underline{t}_j, \bar{t}_j + 1] \leq C_n. \tag{D.6}$$

By construction we have $TV[\underline{t}_j, \bar{t}_j + 1] > \nu/\sqrt{n_j}$. So,

$$C_n \geq \sum_{j=1}^{N-1} \nu/\sqrt{n_j} \tag{D.7}$$

$$\geq (N-1)^{3/2}\nu/\sqrt{n}, \tag{D.8}$$

where the last line follows by Jensen's inequality. Rearranging gives the bound on $N = O\left(1 \vee n^{1/3}C_n^{2/3}B^{-2/3}\right)$. Now the post processing step only increases the number of bins by $O(N)$. Thus we get $M = O\left(1 \vee n^{1/3}C_n^{2/3}B^{-2/3}\right)$. □

Lemma 40. *(**bounding** $T_{1,i}$) Assume that we run* FLH *with the settings described in Theorem 34. For any bin i we have $T_{1,i} = O\left(B^2 \log n\right)$*

Proof. Note that FTL with squared error losses outputs predictions which are online averages of the past labels that the algorithm has seen so far. Hence the predictions of all base learners as well as FLH belong to the interval $[-B, B]$. It is known that (see for eg. [40], Chapter 3) squared error losses are $1/(8B^2)$ exp-concave in the interval $[-B, B]$. Further FTL with squared error losses suffers only logarithmic regret of $O(B^2 \log n)$ ([40], Chapter 3).

Hence due to the adaptive regret bound of FLH (Theorem 3.2 in [23]) by setting the learning rate $\zeta = 1/(8B^2)$, we have that the static regret of FLH in any interval $[i_s, i_t]$ is also $O(\log n)$. This proves the lemma. □

Lemma 41. *(**bounding** $T_{2,i}$) Define $C_i := \sum_{j=i_s+1}^{i_t} |u_j - u_{j-1}|$, the TV within bin i incurred by the offline optimal solution. Let $\Delta s_i := s_{i_t} - s_{i_s-1}$ and $n_i := i_t - i_s + 1$. We have $T_{2,i} \leq \frac{-\lambda^2(\Delta s_i)^2}{n_i}$.*

Proof. From the stationarity conditions in Lemma 36, we can write

$$\bar{u}_i - \bar{y}_i = \frac{\lambda \Delta s_i}{n_i}. \tag{D.9}$$

Further,

$$\sum_{j=i_s}^{i_t}(y_j - \bar{y}_i)^2 - (y_j - \bar{u}_i)^2 = n_i(\bar{u}_i - \bar{y}_i)^2 + 2\sum_{j=i_s}^{i_t}(y_j - \bar{u}_i)(\bar{u}_i - \bar{y}_i) \tag{D.10}$$

$$= -n_i(\bar{u}_i - \bar{y}_i)^2 \tag{D.11}$$

Now plugging in Eq. (D.9) yields the lemma.

□

Lemma 42. *(**bounding** $T_{3,i}$) Let C_i and Δs_i be as in Lemma 41.*

Case(a) *If $|\Delta s_i| > 0$ then $T_{3,i} \leq B^2 + 6\lambda C_i$.*

Case(b) *If $\Delta s_i = 0$ with $s_{i_s-1} = s_{i_t} = 1$ and the offline optimal $\boldsymbol{u}$ is non-decreasing within bin i, then $T_{3,i} \le B^2$.*

Case(c) *If $\Delta s_i = 0$ with $s_{i_s-1} = s_{i_t} = -1$ and the offline optimal $\boldsymbol{u}$ is non-increasing within bin i, then $T_{3,i} \le B^2$.*

Proof. Applying stationarity conditions, we have

$$T_{3,i} = \sum_{j=i_s}^{i_t} (y_j - \bar{u}_i)^2 - (y_j - u_j)^2 \tag{D.12}$$

$$= \sum_{j=i_s}^{i_t} (u_j - \bar{u}_i)(2y_j - \bar{u}_i - u_j) \tag{D.13}$$

$$= \sum_{j=i_s}^{i_t} (u_j - \bar{u}_i)(2y_j - 2u_j + u_j - \bar{u}_i) \tag{D.14}$$

$$= \sum_{j=i_s}^{i_t} (u_j - \bar{u}_i)^2 + 2\lambda(u_j - \bar{u}_i)(s_{j-1} - s_j) \tag{D.15}$$

$$\le n_i C_i^2 + \sum_{j=i_s}^{i_t} 2\lambda(u_j - \bar{u}_i)(s_{j-1} - s_j), \tag{D.16}$$

where in the last line we used $|u_j - \bar{u}_i| \le C_i$. Also observe that $n_i C_i^2 \le B^2$ for bins in the partition $\mathcal{P}$ by Lemma 38. Now by expanding the second term followed by a regrouping of the terms in the summation, we can write

$$\sum_{j=i_s}^{i_t} 2\lambda(u_j - \bar{u}_i)(s_{j-1} - s_j) = 2\lambda \left(s_{i_s-1}(u_{i_s} - \bar{u}_i) - s_{i_t}(u_{i_t} - \bar{u}_i)\right) + 2\lambda \sum_{j=i_s+1}^{i_t} |u_j - u_{j-1}| \tag{D.17}$$

$$= 2\lambda C_i + 2\lambda \left(s_{i_s-1}(u_{i_s} - \bar{u}_i) - s_{i_t}(u_{i_t} - \bar{u}_i)\right) \tag{D.18}$$

Now we discuss the three cases.

Case (a) When $|\Delta s_i| > 0$, then by triangle inequality we have $2\lambda \left(s_{i_s-1}(u_{i_s} - \bar{u}_i) - s_{i_t}(u_{i_t} - \bar{u}_i)\right) \le 4\lambda C_i$.

Case (b) In this case we have $2\lambda \left(s_{i_s-1}(u_{i_s} - \bar{u}_i) - s_{i_t}(u_{i_t} - \bar{u}_i)\right) = \lambda(u_{i_s} - u_{i_t}) = -2\lambda C_i$ since the sequence is non-decreasing within the bin. Hence this term cancels with the corresponding additive term of $2\lambda C_i$ in Eq. (D.18).

Case (c) By similar logic as in case (b) we can once again write $2\lambda \left(s_{i_s-1}(u_{i_s} - \bar{u}_i) - s_{i_t}(u_{i_t} - \bar{u}_i)\right) = -2\lambda C_i$.

Substituting the bound of each case into (D.16). we obtain the expression as stated. □

Proposition 43. ***(Extension to higher dimensions)*** *Consider a protocol where at each time the learner predicts a vector $\boldsymbol{x}_t \in \mathbb{R}^d$ after which the adversary reveals $\boldsymbol{y}_t$ such that $\|\boldsymbol{y}_t\|_\infty \leq B$. Consider a comparator sequence of vectors $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ such that $TV(\boldsymbol{w}_{1:n}) := \sum_{t=2}^n \|\boldsymbol{w}_t - \boldsymbol{w}_{t-1}\|_1 \leq C_n$. By running d instances of* FLH *with learning rate $\zeta = 1/(8B^2)$ and FTL as base learners, where instance i, $i \in [d]$, predicts $\boldsymbol{x}_t[i]$ at time t, we have*

$$R_n(\boldsymbol{w}_{1:n}) := \sum_{j=1}^n \|\boldsymbol{y}_t - \boldsymbol{x}_t\|_2^2 - \|\boldsymbol{y}_t - \boldsymbol{w}_t\|_2^2 = \tilde{O}\left(dB^2 \log n \vee d^{1/3} n^{1/3} C_n^{2/3} B^{4/3}\right).$$

Proof. Let $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ be the offline optimal sequence. Let $C_n[k] = \sum_{t=2}^n |\boldsymbol{u}_t[k] - \boldsymbol{u}_{t-1}[k]|$ be its TV allocated to coordinate k. WLOG, let's assume the FLH for coordinates $k \in [k']$ for $k' \leq d$ incurs $\tilde{O}\left(n^{1/3}(C_n[k])^{2/3} B^{4/3}\right)$ regret and the regret incurred by FLH for coordinates $k > k'$ is $O(\log n)$. Since squared error losses decomposes coordinate-wise, we have

$$\begin{aligned}
R_n(\boldsymbol{w}_{1:n}) &\leq \sup_{\boldsymbol{w}_{1:n}: TV(\boldsymbol{w}_{1:n}) \leq C_n} R_n(\boldsymbol{w}_{1:n}) && \text{(D.19)}\\
&= R_n(\boldsymbol{u}_{1:n}) && \text{(D.20)}\\
&= (d-k')B^2 \log n + \sum_{k=1}^{k'} \tilde{O}\left(n^{1/3}(C_n[k])^{2/3} B^{4/3}\right) && \text{(D.21)}\\
&\leq (d-k')B^2 \log n + \tilde{O}\left(n^{1/3}(k')^{1/3} B^{4/3} \left(\sum_{k=1}^{k'} C_n[k]\right)^{2/3}\right), && \text{(D.22)}
\end{aligned}$$

where the last line follows by Holder's inequality $\boldsymbol{x}^T \boldsymbol{y} \leq \|x\|_3 \|y\|_{3/2}$, where we treat $\boldsymbol{x}$ as just a vector of ones in $\mathbb{R}^{k'}$. The above expression can be further upper bounded by $\tilde{O}\left(2dB^2 \log n \vee 2d^{1/3} n^{1/3} C_n^{2/3} B^{4/3}\right)$.

□

Proposition 44. ***(Lower bound)*** *Assume the protocol and notations of Proposition 43. For any algorithm, we have*

$$\sup_{\boldsymbol{w}_{1:n}: TV(\boldsymbol{w}_{1:n}) \leq C_n} R_n(\boldsymbol{w}_{1:n}) = \Omega\left(dB^2 \log n \vee d^{1/3} n^{1/3} C_n^{2/3} B^{4/3}\right). \quad (5.13)$$

Proof. Consider a fixed (but unknown) sequence $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ such that $TV(\boldsymbol{u}_{1:n}) \leq C_n$ with $\|\boldsymbol{u}_t\|\infty \leq B/2$ and TV along the coordinate $k \in [d]$, $TV(\boldsymbol{u}_{1:n}[k]) \leq C_n/d$ for all k. Let the labels be $\boldsymbol{y}_t = \boldsymbol{u}_t + \boldsymbol{\epsilon}_t$ where each coordinate of $\boldsymbol{\epsilon}_t$ is generated by iid $U[-B/2, B/2]$. Further $\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_n$ are also iid. Then by the results of [2], for any predic-

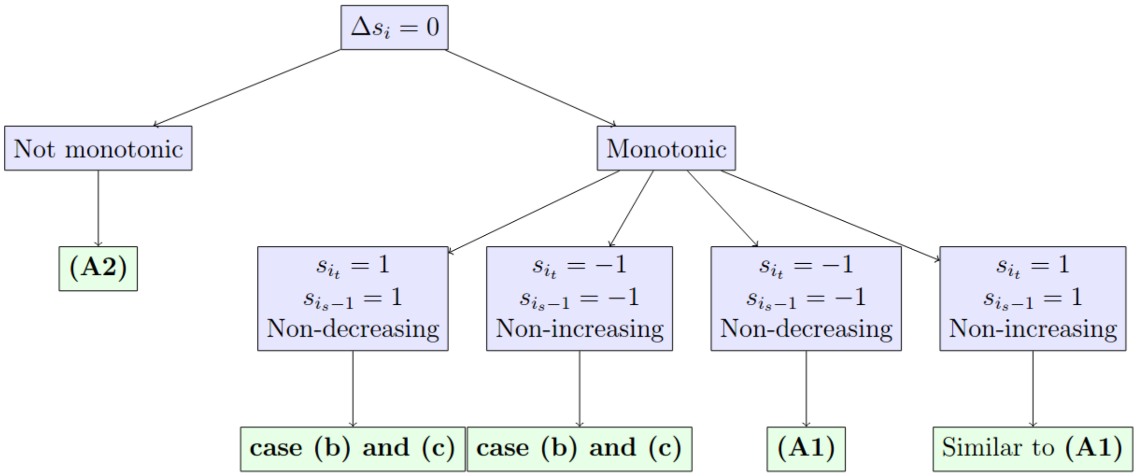


Figure D.2: *Various configurations of the optimal sequence within a bin* $[i_s, i_t]$ *with* $\Delta s_i = 0$. *The leaf nodes indicate the labels of the paragraphs in the Proof of Theorem 34 to handle each scenario.*

tion strategy that produces outputs $\boldsymbol{x}_t$, we have

$$\sup_{\boldsymbol{w}_{1:n}:TV(\boldsymbol{w}_{1:n})\leq C_n} R_n(\boldsymbol{w}_{1:n}) \geq \sum_{k=1}^{d}\sum_{t=1}^{n} E\left[(\boldsymbol{y}_t[k]-\boldsymbol{x}_t[k])^2 - (\boldsymbol{y}_t[k]-\boldsymbol{u}_t[k])^2\right] \tag{D.23}$$

$$=_{(a)} \sum_{k=1}^{d}\sum_{t=1}^{n} E\left[(\boldsymbol{u}_t[k]-\boldsymbol{x}_t[k])^2\right] \tag{D.24}$$

$$= \sum_{k=1}^{d} \Omega(n^{1/3}(C_n/d)^{2/3}B^{4/3}) \tag{D.25}$$

$$= \Omega(d^{1/3}n^{1/3}C_n^{2/3}B^{4/3}), \tag{D.26}$$

where in line (a) we used the fact that $\boldsymbol{x}_t[k]$ is independent of $\boldsymbol{y}_t[k]$ and $\boldsymbol{y}_t[k]-\boldsymbol{u}_t[k] \sim U[-B/2, B/2]$.

The $dB^2 \log n$ part of the lower bound is implied by the lower bound construction of Vovk [42] (or cf. proof of Theorem 11.9 in [40]).

□

Close comparison to lower bound in [37]. For the case of 1D forecasting of TV bounded sequences, [37] consider a stochastic setting where the labels obey $y_t = w_t + \epsilon_t$ for some iid σ subgaussian noise ϵ_t and $w_t \in \mathcal{TV}^B(C_n)$. They provide a lower bound of $\tilde{\Omega}\left((nB^2 \wedge n\sigma^2 \wedge n^{1/3}C_n^{2/3}\sigma^{4/3}) + (nB^2 \wedge BC_n) + B^2\right)$ where $(a \wedge b) = min\{a, b\}$. In accordance with the proof of Proposition 44, we can take $\sigma = B/2$ and $\boldsymbol{w}_{1:n} \in \mathcal{TV}^{B/2}(C_n)$

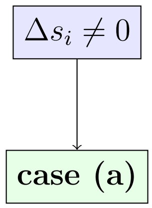


Figure D.3: *A configuration of optimal sequence within a bin $[i_s, i_t]$ with $|\Delta s_i| \neq 0$. The leaf node indicate the label of the paragraph in the Proof of Theorem 34 to handle this scenario.*

to translate this lower bound into our setting for 1D case to get a lower bound of:

$$R_n(C_n) = \tilde{\Omega}\left((nB^2 \wedge n^{1/3}C_n^{2/3}B^{4/3}) + (nB^2 \wedge BC_n) + B^2\right). \tag{D.27}$$

Any learner must have to incur $O(B^2)$ loss in the first round. Combining this with the upper bound in Theorem 34 along with the trvial regret bound of $O(nB^2)$ we can get a refined regret upper bound of:

$$R_n(C_n) = \tilde{O}\left((nB^2 \wedge n^{1/3}C_n^{2/3}B^{4/3}\right) + B^2). \tag{D.28}$$

Comparing Eq.(D.27) and (D.28) seems to falsely suggest that during the regime where $n^{1/3}C_n^{2/3}B^{4/3} < BC_n < nB^2$ upper bound in Eq.(D.28) is smaller than the lower bound in Eq.(D.27). But $n^{1/3}C_n^{2/3}B^{4/3} < BC_n$ happens when $C_n > nB$, in which case $BC_n < nB^2$ is not satisfied. Hence we conclude that this regime is not realisable implying no contradictions.

Close comparison to lower bound in [62]. Proposition 1 of [62] considers squared error losses in 1D and show that when $C_n = n^{\frac{2+\gamma}{4-\gamma}}$ for all $\gamma \in (0,1)$, the dynamic regret obeys

$$R_n(C_n) = \Omega\left(\log n \vee (nC_n)^{\gamma/2}\right). \tag{D.29}$$

We proceed to show that our lower bound of $\Omega(\log n \vee n^{1/3}C_n^{2/3})$ is tighter than this. Whenever $C_n = n^{\frac{2+\gamma}{4-\gamma}}$, we have

$$(nC_n)^{\gamma/2} = n^{\frac{3\gamma}{4-\gamma}}, \tag{D.30}$$

and,

$$n^{1/3}C_n^{2/3} = n^{\frac{8+\gamma}{12-3\gamma}}. \tag{D.31}$$

It can be verified that for all $\gamma \in (0,1)$, $n^{\frac{3\gamma}{4-\gamma}} \leq n^{\frac{8+\gamma}{12-3\gamma}}$ making our lower bound tighter.

D.3 Proofs for Section 5.2

D.3.1 One dimensional setting

In the section, we adopt all the notations used in Section 5.1. For the sake of simplicity of exposition, we first present the results in one dimensional setting and extend it later to higher dimensions. We have the following guarantee in one dimension.

Theorem 168. *($d = 1$) By using the base learner as ONS with parameter $\zeta = \min\left\{\frac{1}{4G^\dagger(2B+2G/\beta)}, \alpha\right\}$ and decision set $\mathcal{D}$ and choosing learning rate $\eta = \alpha$,* FLH *guarantees a dynamic regret $R_n(C_n) = \tilde{O}\left(n^{1/3}C_n^{2/3} \vee \log n\right)$.*

We start the analysis by inspecting the KKT conditions.

Lemma 169. *(**characterization of offline optimal**) Consider the following convex optimization problem.*

$$\min_{\tilde{u}_1,\dots,\tilde{u}_n,\tilde{z}_1,\dots,\tilde{z}_{n-1}} \quad \sum_{t=1}^{n} f_t(\tilde{u}_t) \tag{D.32a}$$

$$\text{s.t.} \quad \tilde{z}_t = \tilde{u}_{t+1} - \tilde{u}_t \; \forall t \in [n-1], \tag{D.32b}$$

$$\sum_{t=1}^{n-1} |\tilde{z}_t| \le C_n, \tag{D.32c}$$

$$-B \le \tilde{u}_t \; \forall t \in [n], \tag{D.32d}$$

$$\tilde{u}_t \le B \; \forall t \in [n], \tag{D.32e}$$

Let $u_1, \dots, u_n, z_1, \dots, z_{n-1}$ be the optimal primal variables and let $\lambda \ge 0$ be the optimal dual variable corresponding to the constraint (D.32c)*. Further, let $\gamma_t^- \ge 0, \gamma_t^+ \ge 0$ be the optimal dual variables that correspond to constraints* (D.32d) *and* (D.32e) *respectively for all $t \in [n]$. By the KKT conditions, we have*

- ***stationarity:*** *$\nabla f_t(u_t) = \lambda\,(s_t - s_{t-1}) + \gamma_t^- - \gamma_t^+$, where $s_t \in \partial|z_t|$ (a subgradient). Specifically, $s_t = \mathrm{sign}(u_{t+1} - u_t)$ if $|u_{t+1} - u_t| > 0$ and s_t is some value in $[-1, 1]$ otherwise. For convenience of notations later, we also define $s_n = s_0 = 0$.*
- ***complementary slackness:*** *(a) $\lambda\left(\sum_{t=2}^{n} |u_t - u_{t-1}| - C_n\right) = 0$; (b) $\gamma_t^-(u_t + B) = 0$ and $\gamma_t^+(u_t - B) = 0$ for all $t \in [n]$*

Terminology. We will refer to the optimal primal variables $u_1, \dots, u_n$ in Lemma 169 as the *offline optimal sequence* in this section.

Next, we record an easy corollary of Lemma 38.

Corollary 170. *(**key partition**) Assume the notations of Lemma 38. Create a partition of $\mathcal{P}$ of $[n]$ with the procedure mentioned in Lemma 38 . Then for any $[i_s, i_t] \in \mathcal{P}$, we have*

- *(TV constraint)* $\sum_{j=i_s+1}^{i_t} |u_j - u_{j-1}| \le B/\sqrt{n_i}$,
- *(Bins bound)* $M := |\mathcal{P}| = O(n^{1/3} C_n^{2/3})$.
- *(Structural property) If* $i_s > 1$ *then* $u_{i_s} \neq u_{i_s-1}$. *Similarly if* $i_t < n$ *then* $u_{i_t} \neq u_{i_t+1}$.

Now we make an important observation regrading the dual variables γ_j^- and γ_j^+. The following property will be used several times in the proofs to follow.

Lemma 171. *Define* $\Gamma_i^+ := \sum_{j=i_s}^{i_t} \gamma_j^+$ *and* $\Gamma_i^- := \sum_{j=i_s}^{i_t} \gamma_j^-$. *Consider a bin* $[i_s, i_t] \in \mathcal{P}$, *where* $\mathcal{P}$ *is the partition of* $[n]$ *constructed in Corollary 170. Then at-least one of the following is always satisfied.*

- $\gamma_j^- = 0$ *for all* $j \in [i_s, i_t]$.
- $\gamma_j^+ = 0$ *for all* $j \in [i_s, i_t]$.

Consequently we have $\sum_{j=i_s}^{i_t} |\gamma_j^-| + |\gamma_j^+| = \left|\Gamma_i^- - \Gamma_i^+\right|$, *for any bin* $[i_s, i_t] \in \mathcal{P}$.

Proof. From the properties of the partition $\mathcal{P}$ in Corollary 170, we have that the TV of the offline optimal incurred within each bin is at-most $B/\sqrt{n_i} \le B$. Hence within bin $[i_s, i_t] \in \mathcal{P}$, if the optimal sequence attains the value $-B$ at some time point, it can never attain the value B and vice-versa. So due to complementary slackness rule in Lemma 169, either $\gamma_j^+ = 0$ or $\gamma_j^- = 0$ uniformly for all $j \in [i_s, i_t]$. The last line in the statement of lemma follows by recalling that $\gamma_j^- \ge 0$ and $\gamma_j^+ \ge 0$ from Lemma 169.

□

For convenience, we recall here the regret decomposition of Eq.(5.15) specified to one dimensional setting. Let $\mathcal{P}$ be a partition of $[n]$ into M bins as specified in Corollary 170. Let $[i_s, i_t]$ denote the i^{th} bin in $\mathcal{P}$ and let n_i be its length. Define $\bar{u}_i = \frac{1}{n_i} \sum_{j=i_s}^{i_t} u_j$ and $\dot{u}_i = \bar{u}_i - \frac{1}{n_i \beta} \sum_{j=i_s}^{i_t} \nabla f_j(\bar{u}_i)$ where β is as in Assumption EC-2. Let x_j be the prediction made by FLH at time j. We start with following regret decomposition.

$$R_n(C_n) \le \sum_{i=1}^{M} \underbrace{\sum_{j=i_s}^{i_t} f_j(x_j) - f_j(\dot{u}_i)}_{T_{1,i}} + \sum_{i=1}^{M} \underbrace{\sum_{j=i_s}^{i_t} f_j(\dot{u}_i) - f_j(\bar{u}_i)}_{T_{2,i}} + \sum_{i=1}^{M} \underbrace{\sum_{j=i_s}^{i_t} f_j(\bar{u}_i) - f_j(u_j)}_{T_{3,i}}. \tag{D.33}$$

We proceed to bound the terms $T_{1,i}, T_{2,i}, T_{3,i}$ for the bins that belong to the partition $\mathcal{P}$.

Lemma 172. ***(bounding** $T_{1,i}$**)*** *Let the experts in* FLH *be the ONS algorithms with parameter* $\zeta = \min\left\{\frac{1}{4G^\dagger(2B+2G)}, \alpha\right\}$ *and decision set* $\mathcal{D}$. *Also choose learning rate* $\eta = \alpha$, *for* FLH*. Then for any bin* $[i_s, i_t]$ *we have,*

$$\sum_{j=i_s}^{i_t} f_j(x_j) - f_j(\dot{u}_i) = O\left(BG^\dagger \log n + GG^\dagger \log n + \frac{\log n}{\alpha}\right) \tag{D.34}$$

$$= O(\log n). \tag{D.35}$$

Proof. First we proceed to bound $|\dot{u}_i|$. Since $|\nabla f_j(u_j)| \leq G$ by Assumption EC-1, we have

$$|\dot{u}_i| \leq |\bar{u}_i| + \frac{G}{\beta} \tag{D.36}$$

$$\leq B + G, \tag{D.37}$$

since $\beta \geq 1$ by Assumption EC-2. For any $x \in \mathcal{D}$, we have $|x - \dot{u}_i| \leq 2B + 2G$ by triangle inequality.

By Assumption EC-4 we have $|\nabla f_j(x)| \leq G^\dagger$ for any $x \in \mathcal{D}$. Also, recall that by Assumption EC-3, the loss functions f_j are α exp-concave in the domain $\mathcal{D}$. Let p_j be the predictions of ONS in the interval $[i_s, i_t]$. If we choose $\zeta = \min\left\{\frac{1}{4G^\dagger(2B+2G)}, \alpha\right\}$ as the parameter of the ONS, Theorem 2 of [77] implies that

$$\sum_{j=i_s}^{i_t} f_j(p_j) - f_j(\dot{u}_i) = O\left(BG^\dagger \log n + GG^\dagger \log n\right) \tag{D.38}$$

$$= O\left(\log n\right). \tag{D.39}$$

Now the Lemma is implied by the SA regret bound of FLH (Theorem 3.2 of [23]). □

Lemma 173. ***(bounding*** $T_{2,i}$*). For a bin* $[i_s, i_t] \in \mathcal{P}$*, let* C_i, n_i *and* Δs_i *be as in Lemma 41 and* Γ_i^+, Γ_i^- *be as in Lemma 171. We have*

$$\sum_{j=i_s}^{i_t} f_j(\dot{u}_i) - f_j(\bar{u}_i) \leq \frac{-\left(\lambda \Delta s_i + \Gamma_i^- - \Gamma_i^+\right)^2}{2n_i\beta} + \lambda|\Delta s_i|C_i + |\Gamma_i^- - \Gamma_i^+|C_i. \tag{D.40}$$

Proof. We start with the short proof the descent lemma. Let $g(x)$ be a L strongly smooth function. Let $x^+ = x - \mu\nabla f(x)$ for some $\mu > 0$. Then we have

$$g(x^+) - g(x) \leq (\nabla g(x))^2 \left(\frac{L}{2}\mu^2 - \mu\right) \tag{D.41}$$

$$= \frac{-(\nabla g(x))^2}{2L}, \tag{D.42}$$

by choosing $\mu = 1/L$. By taking $g(x) = \sum_{j=i_s}^{i_t} f_j(x)$ and noting that g is $n_i\beta$ gradient

Lipschitz due to Assumption EC-2, we get

$$T_{2,i} := \sum_{j=i_s}^{i_t} f_j(\dot{u}_i) - f_j(\bar{u}_i) \tag{D.43}$$

$$\leq \frac{-\left(\sum_{j=i_s}^{i_t} \nabla f_j(\bar{u}_i)\right)^2}{2n_i\beta} \tag{D.44}$$

$$= \frac{-1}{2n_i\beta}\left(\sum_{j=i_s}^{i_t} \nabla f_j(u_j) + \nabla f_j(\bar{u}_i) - \nabla f_j(u_j)\right)^2 \tag{D.45}$$

$$\leq \frac{-1}{2n_i\beta}\left(\sum_{j=i_s}^{i_t} \nabla f_j(u_j)\right)^2 + \frac{1}{n_i\beta}\left|\sum_{j=i_s}^{i_t} \nabla f_j(u_j)\right|\left|\sum_{j=i_s}^{i_t} \nabla f_j(\bar{u}_i) - \nabla f_j(u_j)\right|. \tag{D.46}$$

From the KKT conditions in Lemma 169 we have $\sum_{j=i_s}^{i_t} \nabla f_j(u_j) = \lambda\Delta s_i + \Gamma_i^- - \Gamma_i^+$. Since f_j are β-gradient Lipschitz and $|\bar{u}_i - u_j| \leq C_i$, we also have

$$\left|\sum_{j=i_s}^{i_t} \nabla f_j(\bar{u}_i) - \nabla f_j(u_j)\right| \leq n_i\beta C_i. \tag{D.47}$$

Substituting these we get,

$$T_{2,i} \leq \frac{-\left(\lambda\Delta s_i + \Gamma_i^- - \Gamma_i^+\right)^2}{2n_i\beta} + \lambda|\Delta s_i|C_i + |\Gamma_i^- - \Gamma_i^+|C_i. \tag{D.48}$$

□

Lemma 174. ***(bounding*** $T_{3,i}$*)* *For a bin* $[i_s, i_t] \in \mathcal{P}$*, let* C_i, n_i *and* Δs_i *be as in Lemma 41 and* Γ_i^+, Γ_i^- *be as in Lemma 171.*
case(a) *If* $|\Delta s_i| > 0$ *then we have,*

$$\sum_{j=i_s}^{i_t} f_j(\bar{u}_i) - f_j(u_j) \leq \frac{\beta n_i C_i^2}{2} + 3\lambda C_i + |\Gamma_i^- - \Gamma_i^+|C_i. \tag{D.49}$$

case(b) *If* $\Delta s_i = 0$ *with* $s_{i_s-1} = s_{i_t} = 1$ *and the offline optimal* $\boldsymbol{u}$ *is non-decreasing within bin i with* $-B < u_i < B$ *for all* $i \in [i_s, i_t]$*, then*

$$\sum_{j=i_s}^{i_t} f_j(\bar{u}_i) - f_j(u_j) \leq \frac{\beta n_i C_i^2}{2}. \tag{D.50}$$

case(c) *If* $\Delta s_i = 0$ *with* $s_{i_s-1} = s_{i_t} = -1$ *and the offline optimal* $\boldsymbol{u}$ *is non-increasing* *within bin i with $-B < u_i < B$ for all $i \in [i_s, i_t]$, then*

$$\sum_{j=i_s}^{i_t} f_j(\bar{u}_i) - f_j(u_j) \leq \frac{\beta n_i C_i^2}{2}. \tag{D.51}$$

Proof. Due to strong smoothness, we have

$$T_{3,i} := \sum_{j=i_s}^{i_t} f_j(\bar{u}_i) - f_j(u_j) \tag{D.52}$$

$$\leq \sum_{j=i_s}^{i_t} \nabla f_j(u_j)(\bar{u}_i - u_j) + \frac{\beta}{2}(\bar{u}_i - u_j)^2 \tag{D.53}$$

$$\leq \frac{\beta n_i C_i^2}{2} + \sum_{j=i_s}^{i_t} \nabla f_j(u_j)(\bar{u}_i - u_j). \tag{D.54}$$

Now by expanding the second term and using the structure of gradients as in Lemma 169 followed by a regrouping of the terms in the summation we can write,

$$\sum_{j=i_s}^{i_t} \nabla f_j(u_j)(\bar{u}_i - u_j) = \lambda \left(s_{i_s-1}(u_{i_s} - \bar{u}_i) - s_{i_t}(u_{i_t} - \bar{u}_i)\right) + \lambda \sum_{j=i_s+1}^{i_t} |u_j - u_{j-1}| \tag{D.55}$$

$$+ \sum_{j=i_s}^{i_t} (\gamma_j^- - \gamma_j^+)(\bar{u}_i - u_j) \tag{D.56}$$

$$\leq \lambda \left(s_{i_s-1}(u_{i_s} - \bar{u}_i) - s_{i_t}(u_{i_t} - \bar{u}_i)\right) + \lambda C_i + |\Gamma_i^- - \Gamma_i^+| C_i, \tag{D.57}$$

where the last line follows due to Lemma 171 and $|\bar{u}_i - u_j| \leq C_i$ for all $j \in [i_s, i_t]$.

Now we consider three cases in the statement of the lemma.

case (a) When $|\Delta s_i| > 0$, then by triangle inequality we have $\lambda \left(s_{i_s-1}(u_{i_s} - \bar{u}_i) - s_{i_t}(u_{i_t} - \bar{u}_i)\right) \leq 2\lambda C_i$.

case (b) In this case we have $\lambda \left(s_{i_s-1}(u_{i_s} - \bar{u}_i) - s_{i_t}(u_{i_t} - \bar{u}_i)\right) = \lambda(u_{i_s} - u_{i_t}) = -\lambda C_i$ since the sequence is non-decreasing within the bin. Hence this term cancels with the corresponding additive term of λC_i in Eq. (D.57). Further $\gamma_j^- = \gamma_j^+ = 0$ since $-B < u_j < B$ for all $j \in [i_s, i_t]$.

case (c) By similar logic as in case (b) we can once again write $\lambda \left(s_{i_s-1}(u_{i_s} - \bar{u}_i) - s_{i_t}(u_{i_t} - \bar{u}_i)\right) = -\lambda C_i$.

Putting everything together now yields the lemma.

□

Proof. **of Theorem** 168. The strategy of the proof is to bound the regret incurred within each time interval $[i_s, i_t] \in \mathcal{P}$ where $\mathcal{P}$ is as in Corollary 170 and add them up towards the end. We annotate several key paragraphs for the purposes of referring the arguments contained in them at later points.

If the the partition $\mathcal{P}$ contains only one bin, then we split it into at-most two bins $[1, a]$ and $[a+1, n]$ such that the optimal sequence is constant within $[1, a]$ and hence regret incurred within this bin is $\tilde{O}(1)$ by Strong Adaptivity of FLH. The regret incurred in the bin $[a+1, i_t]$ can be bounded by using the arguments below. So in what follows we assume for a bin $[i_s, i_t]$ either $i_s > 1$ or $i_t < n$.

By virtue of Lemma 171, any bin $[i_s, i_t] \in \mathcal{P}$ will have either $\gamma_j^- = 0$ for all $j \in [i_s, i_t]$ or $\gamma_j^+ = 0$ for all $j \in [i_s, i_t]$. Below we bound the regret for bins with $\gamma_j^+ = 0$ uniformly for all $j \in [i_s, i_t]$. The arguments for the alternate case where $\gamma_j^- = 0$ follows similarly. Figures D.5, D.6 and D.7 sketch the floor plan of the proof pictorially. Throughout the proof, we will use the properties in Corollary 170 in conjunction with the observations in Remark 39.

(S1): Consider a bin with $\Delta s_i = 0$ with $s_{i_t} = s_{i_s-1} = 1$ and the optimal sequence is non-decreasing within the bin. By the structural property of Corollary 170, this happens when $u_{i_s} > u_{i_s-1}$, $u_{i_t+1} > u_{i_t}$ where $1 < i_s < i_t < n$. Since the sequence is non-decreasing, it never attains $-B$ within this bin. Hence this is the same situation as in case (b) of Lemma 174. We have $T_{1,i} = \tilde{O}(1)$ due to Lemma 172. $T_{2,i} = 0$ due to Lemma 173 as $\Gamma_i^+ = \Gamma_j^- = 0$ since the sequence never attains $\pm B$ within the current bin combined with the fact that $\Delta s_i = 0$. $T_{3,i} = O(1)$ due to Lemma 174 combined with the fact that $C_i \leq B/\sqrt{n_i}$ due to Corollary 170. So the total regret within the current bin is bounded by $T_{1,i} + T_{2,i} + T_{3,i} = \tilde{O}(1)$.

The total regret for a bin satisfying case (c) of Lemma 174 can be bound using similar arguments as above.

The three cases where (i) $\Delta s_i = 0$ with $s_{i_s-1} = s_{i_t} = -1$ and the offline optimal $\boldsymbol{u}$ is non-decreasing within bin i; (ii) $\Delta s_i = 0$ with $s_{i_s-1} = s_{i_t} = 1$ and the offline optimal $\boldsymbol{u}$ is non-increasing within bin i and (iii) $\Delta s_i = 0$ and $\boldsymbol{u}$ is not monotonic will be covered shortly in the arguments to follow.

Consider a bin with $|\Delta s_i| > 0$ and $\gamma_j^+ = 0$ uniformly. From Lemmas 173 and 174 and using the fact that $|\Delta s_i| \leq 2$ we have,

$$T_{2,i} + T_{3,i} \leq \frac{\beta n_i C_i^2}{2} + \underbrace{\frac{-\lambda^2 (\Delta s_i)^2}{2 n_i \beta} + 7\lambda C_i}_{(1)} + \underbrace{\frac{\left(-\Gamma_i^-\right)^2}{2 n_i \beta} + 2\Gamma_i^- C_i}_{(2)} - \frac{\lambda \Delta s_i \Gamma_i^-}{n_i \beta}. \tag{D.58}$$

By completing the squares with the terms (1) and (2) in the above display and dropping the negative terms, we get

$$\begin{aligned} T_{2,i} + T_{3,i} &\leq \frac{\beta n_i C_i^2}{2} + \frac{49 C_i^2 n_i \beta}{2(\Delta s_i)^2} + 2\beta n_i C_i^2 - \frac{\lambda \Delta s_i \Gamma_i^-}{n_i \beta} && \text{(D.59)} \\ &\leq 27 B^2 \beta - \frac{\lambda \Delta s_i \Gamma_i^-}{n_i \beta}, && \text{(D.60)} \end{aligned}$$

where in last line we used the facts that $C_i \leq B/\sqrt{n_i}$ by Corollary 170 and $|\Delta s_i| > 1$ whenever $|\Delta s_i| \neq 0$ by Remark 39.

Define $T_i := \sum_{j=i_s}^{i_t} f_j(x_j) - f_j(u_j)$. Notice that:

- **(A1):** When $\Gamma_i^- = 0$ and $|\Delta s_i| > 0$, combining Lemma 172 we have $T_i = \tilde{O}(1)$;
- **(A2):** Similarly when $\Delta s_i > 0$, we get $T_i = \tilde{O}(1)$ as $\Gamma_i^- \geq 0$ by Lemma 169.

In what follows, we try to split an original bin $[i_s, i_t]$ with $\Delta s_i < 0$ into sub-bins that satisfy the above conditions (A1) or (A2).

If optimal sequence is uniformly constant, we can appeal to the static regret guarantee of FLH to get logarithmic regret over n rounds. So we assume that the optimal sequence is not constant uniformly in the analysis below.

Next, we consider the case when $\Delta s_i < 0$. We start with the following observation.

(B1): Consider a bin $[i_s, i_t]$ that satisfies the structural property in Corollary 170. When either $i_s > 1$ or $i_t < n$ and $\Delta s_i < 0$, then $s_{i_t} \in \{-1, 0\}$ and $s_{i_s - 1} \in \{0, 1\}$ with at-least one of them being non-zero.

Since by our assumption $|\mathcal{P}| > 1$, i_s and i_t can't be 1 and n simultaneously. So for any bin $[i_s, i_t]$ with $\Delta s_i < 0$, observation (B1) has to be satisfied.

When $\Delta s_i < 0$, we can have three cases as follows.

Case (1): If the optimal solution is constant (i.e $C_i = 0$) within the bin i. Then we trivially get $T_i = \tilde{O}(1)$.

Case (2): If the optimal solution is monotonic within bin i (see config (a) in Fig.D.4 for an example of this configuration). Then we split the original bin $[i_s, i_t]$ into at-most 2 bins. Let j_1, j_2 be such that $u_m = -B \forall m \in [i_s, j_1 - 1] \cup [j_2 + 1, i_t]$ and $u_{j_1} > -B, u_{j_2} > -B$. If $u_{i_s} > -B$, then $j_1 = i_s$ and $[i_s, j_1 - 1]$ is viewed as an empty interval. Similar logic applies for the right interval $[j_2 + 1, i_t]$. Since the optimal sequence is monotonic within $[i_s, i_t]$, either $j_1 = i_s$ or $j_2 = i_t$. Without loss of generality let's assume that $j_2 = i_t$. We proceed to bound the regret incurred within each of the two sub-bins separately.

Let's annotate bin $[i_s, j_1 - 1]$ by $i^{(1)}$ and bin $[j_1, i_t]$ by $i^{(2)}$. For the bin $i^{(1)}$, the optimal solution is constant and hence the regret $T_{i^{(1)}} = \tilde{O}(1)$. For the bin $i^{(2)}$, notice that $\gamma_j^- = 0 \, \forall j \in [j_1, i_t]$ since the sequence is monotonic with $u_{j_1} > -B$ and since our assumption $j_2 = i_t$ implies $u_{j_2} > -B$. Hence we have $\Gamma_{i^{(2)}}^- = 0$. Since $s_{j_1 - 1} \in \{0, 1\}$ and by observation (B1), $s_{i_t} \in \{-1, 0\}$ with at-least one of them being non-zero, we have $|\Delta s_i| \neq 0$. Hence the bin $i^{(2)}$ falls into the category (A1). So $T_{i^{(2)}} = \tilde{O}(1)$. Adding the regret incurred in each sub-bin separately yields $T_i = \tilde{O}(1)$.

Case (3): Consider the alternate case where we have $\Delta s_i < 0$ and the sequence is not monotonic (see config (b) in Fig. D.4 for an example of this configuration). We split the original bin $[i_s, i_t]$ into at-most three sub-bins $[i_s, j_1 - 1], [j_1, j_2], [j_2 + 1, i_t]$ such that (i) If $u_{i_s} = -B$, then $u_m = -B \, \forall m \in [i_s, j_1 - 1]$ and $u_{j_1} > -B$. If $u_{i_s} > -B$, then we take $j_1 = i_s$ and view $[i_s, j_1 - 1]$ as empty interval. (ii) j_2 is the smallest point in $[j_1, i_t]$ such that $s_{j_2} = -1$ and $u_{j_2} > u_{j_2+1}$.

Let's annotate bins $[i_s, j_1 - 1], [j_1, j_2], [j_2 + 1, i_t]$ by $i^{(1)}, i^{(2)}, i^{(3)}$ respectively. If bin $i^{(1)}$ is not empty, then we have $T_{i^{(1)}} = \tilde{O}(1)$ since $\boldsymbol{u}$ is constant within that bin.

Since $\Delta s_i < 0$, we must have $s_{j_1-1} \in \{0,1\}$ even if $j_1 = i_s$. By construction the sequence $\boldsymbol{u}$ never attains the value $-B$ in the bin $i^{(2)}$ since $u_{j_1} > -B$ and j_2 is the first time point since j_1 after which the optimal sequence jumps downwards. So we have $\Gamma_{i^{(2)}} = 0$. Further we also have $|\Delta s_{i^{(2)}}| > 0$ within bin $i^{(2)}$. So we get $T_{i^{(2)}} = \tilde{O}(1)$ since $i^{(2)}$ falls into category (A1)

For simplicity let's assume that $u_{i_t} > -B$, otherwise we can create another bin that ends at time i_t where optimal solution assumes a constant value of $-B$ and proceed with similar arguments as before to bound the regret in the constant interval.

(S2): If the sequence $\boldsymbol{u}$ is not monotonic in $i^{(3)}$, we split the bin $i^{(3)}$ into two parts $[j_2+1, j_3], [j_3+1, i_t]$ such that j_3 is the largest point in $[j_2+1, i_t]$ with $s_{j_3} = 1$ and $u_{j_3} < u_{j_3+1}$. Let's annotate the bins $[j_2+1, j_3], [j_3+1, i_t]$ by $q^{(1)}, q^{(2)}$ respectively. We have $\Delta s_{q^{(1)}} > 0$ since $s_{j_3} = 1$ and $s_{j_2} = -1$. Hence the bins $q^{(1)}$ falls into the category(A2) mentioned before and we get $T_{q^{(1)}} = \tilde{O}(1)$. Notice that $s_{i_t} \in \{-1, 0\}$ as $\Delta s_i < 0$. Since j_3 is the largest point in $[j_2+1, i_t]$ with $s_{j_3} = 1$ and it is assumed before that $u_{i_t} > -B$, we conclude that the sequence in the interval $q^{(2)}$ is a non-increasing sequence that never attains the value $-B$. So $\Gamma^-_{q^{(2)}} = 0$. Further we have $|\Delta s_{q^{(2)}}| > 1$. So $T_{q^{(2)}} = \tilde{O}(1)$ since $q^{(2)}$ falls into the category (A1). We pause to remark that the arguments we used to bound the regret in the bin $i^{(3)}$ can be used to bound the regret of any bin $[r_s, r_t] \in \mathcal{P}$ with $\Delta s_r = 0$ and the sequence $\boldsymbol{u}$ being not monotonic within bin r.

Note that since $u_{j_2+1} < u_{j_2}$, bin $i^{(3)}$ satisfies the structural property of Corollary 170. So if the sequence $\boldsymbol{u}$ is non-increasing in bin $i^{(3)}$ and $s_{i_t} = -1$, it fits into case (c) of Lemma 174. So we can bound $T_{i^{(3)}} = \tilde{O}(1)$ using arguments presented in (S1).

If the sequence $\boldsymbol{u}$ is monotonic in bin $i^{(3)}$ and $s_{i_t} = 0$ (which happens when $i_t = n$), then we have $\Delta s_{i^{(3)}} = 0-(-1) = 1 > 0$. So bin $i^{(3)}$ falls into the category(A2) mentioned before. Hence the regret $T_{i^{(3)}} = \tilde{O}(1)$.

(S3): If the sequence $\boldsymbol{u}$ is non-decreasing in bin $i^{(3)}$, we split the bin into two intervals $[j_2+1, k], [k+1, i_t]$ such that k is any point in $[j_2+1, i_t]$ with $s_k = 1$ and $u_{k+1} > u_k$. (This configuration is similar to that of config (a) in Fig.5.1). Annotate $[j_2+1, k], [k+1, i_t]$ by $q^{(1)}, q^{(2)}$ respectively. In bin $q^{(1)}$ we have $\Delta s_{q^{(1)}} = 2$ and hence $T_{q^{(1)}} = \tilde{O}(1)$ since $q^{(1)}$ falls into the category (A2). Within bin $q^{(2)}$ due to the assumption that $u_{i_t} > -B$, we have $\Gamma^-_{q^{(2)}} = 0$. We also have $|\Delta s_{q^{(2)}}| > 0$ and consequently $q^{(2)}$ falls into category (A1). So we have $T_{q^{(2)}} = \tilde{O}(1)$. We pause to remark that the arguments we used to bound the regret in bin $i^{(3)}$ for the case where $\boldsymbol{u}$ is non-decreasing, can also be used to bound the regret of any bin $[r_s, r_t]$ with $\Delta s_r = 0$ and $s_{r_t} = s_{r_s-1} = -1$ and the sequence $\boldsymbol{u}$ is non-decreasing. The regret for the alternate case where $\Delta s_r = 0$ and $s_{r_t} = s_{r_s-1} = 1$ and the sequence $\boldsymbol{u}$ is non-increasing can be bounded similarly using a mirrored argument.

So summarizing, in case (3) we get $T_i = \tilde{O}(1)$. Since the intermediate splitting operations can only increase the number of bins to at-most $6M$, adding the regret across all $O(M)$ bins in Corollary 170 yields the Theorem.

□

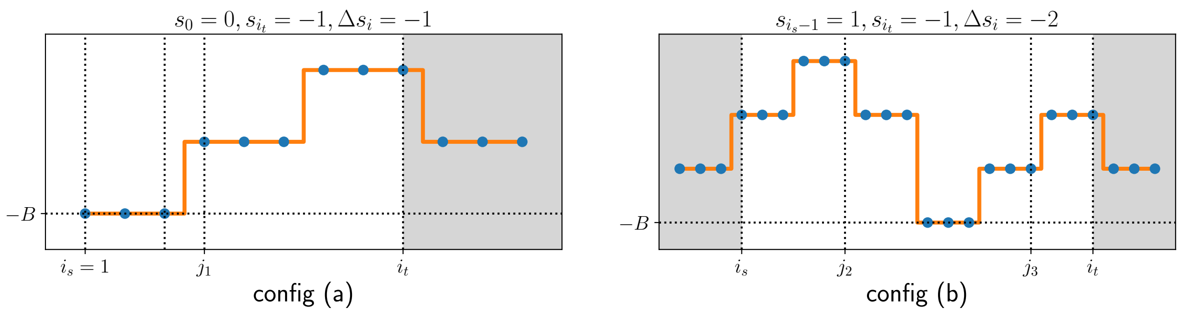


Figure D.4: *Examples of configurations referred in the proof of Theorem 168. The blue dots corresponds to the offline optimal sequence.*

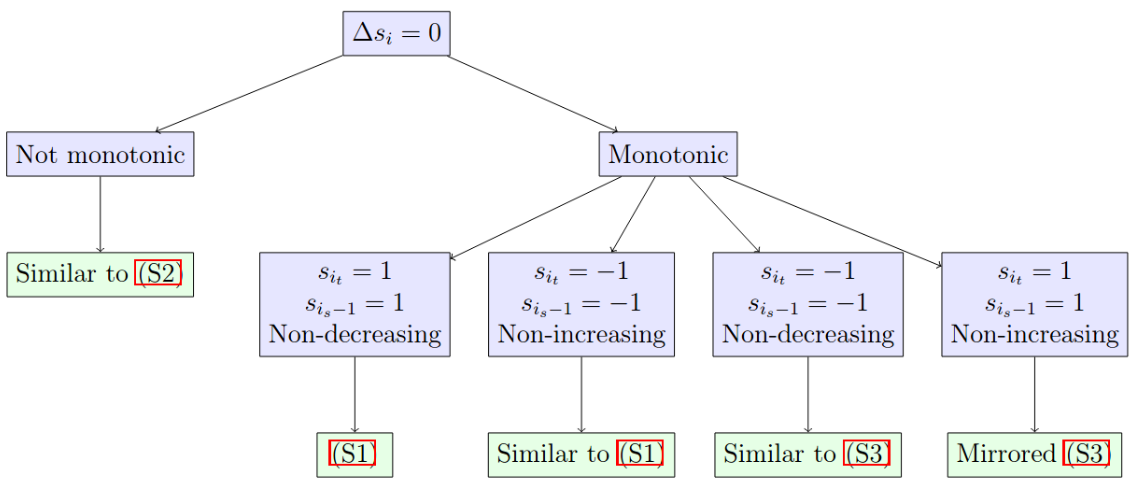


Figure D.5: *Various configurations of the optimal sequence within a bin* $[i_s, i_t]$ *with* $\Delta s_i = 0$. *The leaf nodes indicate the arguments used in the proof of Theorem 168 to handle each scenario.*

D.3.2 Multi dimensional setting

We start by inspecting the KKT conditions.

Lemma 175. *(**characterization of offline optimal**) Consider the following convex*

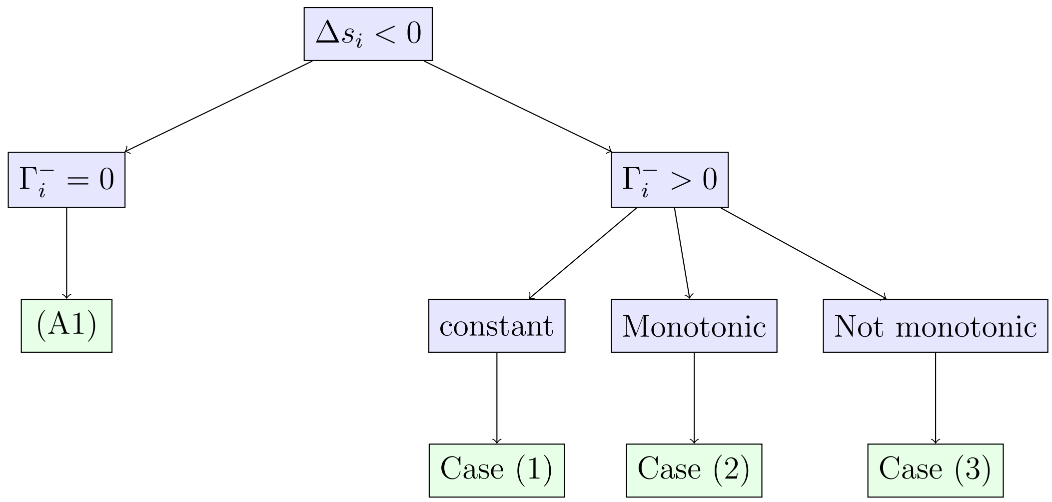


Figure D.6: *Various configurations of optimal sequence within a bin* $[i_s, i_t]$ *with* $\Delta s_i < 0$. *The leaf nodes indicate the arguments used in the proof of Theorem 168 to handle each scenario.*

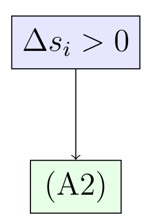


Figure D.7: *A configuration of optimal sequence within a bin* $[i_s, i_t]$ *with* $\Delta s_i > 0$. *The leaf node indicate the arguments used in the proof of Theorem 168 to handle each scenario.*

optimization problem.

$$\min_{\tilde{\boldsymbol{u}}_1,\dots,\tilde{\boldsymbol{u}}_n,\tilde{\boldsymbol{z}}_1,\dots,\tilde{\boldsymbol{z}}_{n-1}} \quad \sum_{t=1}^{n} f_t(\tilde{\boldsymbol{u}}_t) \tag{D.61a}$$

$$\text{s.t.} \quad \tilde{\boldsymbol{z}}_t = \tilde{\boldsymbol{u}}_{t+1} - \tilde{\boldsymbol{u}}_t \,\forall t \in [n-1], \tag{D.61b}$$

$$\sum_{t=1}^{n-1} \|\tilde{\boldsymbol{z}}_t\|_1 \le C_n, \tag{D.61c}$$

$$\|\tilde{\boldsymbol{u}}_t\|_\infty \le B \,\forall t \in [n], \tag{D.61d}$$

Let $\boldsymbol{u}_1,\dots,\boldsymbol{u}_n,\boldsymbol{z}_1,\dots,\boldsymbol{z}_{n-1} \in \mathbb{R}^d$ be the optimal primal variables and let $\lambda \ge 0$ be the optimal dual variable corresponding to the constraint (D.61c)*. Further, let $\boldsymbol{\gamma}_t^+, \boldsymbol{\gamma}_t^- \in \mathbb{R}^d$ with $\boldsymbol{\gamma}_t^+ \ge 0$ and $\boldsymbol{\gamma}_t^- \ge 0$ be the optimal dual variables that correspond to constraint* (D.61d)*. Specifically for $k \in [d]$, $\boldsymbol{\gamma}_t^+[k]$ corresponds to the dual variable for the constraint $\boldsymbol{u}_t[k] \le B$ induced by the relation* (D.61d)*. Similarly $\boldsymbol{\gamma}_t^-[k]$ corresponds to the constraint $-B \le \boldsymbol{u}_t[k]$. By the KKT conditions, we have*

- ***stationarity:*** $\nabla f_t(\boldsymbol{u}_t) = \lambda(\boldsymbol{s}_t - \boldsymbol{s}_{t-1}) + \boldsymbol{\gamma}_t^- - \boldsymbol{\gamma}_t^+$*, where $\boldsymbol{s}_t \in \partial|\boldsymbol{z}_t|$ (a subgradient). Specifically, $\boldsymbol{s}_t[k] = \operatorname{sign}(\boldsymbol{u}_{t+1}[k] - \boldsymbol{u}_t[k])$ if $|\boldsymbol{u}_{t+1}[k] - \boldsymbol{u}_t[k]| > 0$ and $\boldsymbol{s}_t[k]$ is some value in $[-1,1]$ otherwise. For convenience of notations later, we also define* $\boldsymbol{s}_n = \boldsymbol{s}_0 = \mathbf{0}$.
- ***complementary slackness:*** *(a) $\lambda\left(\sum_{t=2}^{n} \|\boldsymbol{u}_t - \boldsymbol{u}_{t-1}\|_1 - C_n\right) = 0$; (b) $\boldsymbol{\gamma}_t^-[k](\boldsymbol{u}_t[k]+B) = 0$ and $\boldsymbol{\gamma}_t^+[k](\boldsymbol{u}_t[k]-B) = 0$ for all $t \in [n]$ and all $k \in [d]$.*

The proof of the above lemma is similar to the 1D case and hence omitted.
Terminology. We will refer to the optimal primal variables $\boldsymbol{u}_1,\dots,\boldsymbol{u}_n$ in Lemma 175 as the *offline optimal sequence* in this section.

Next, we claim the existence of a partitioning of $[n]$ with some useful properties.

Lemma 176. ***(key partition)*** *There exist a partitioning $\mathcal{P}$ of $[n]$ into $M = O(dn^{1/3}C_n^{2/3})$ intervals viz $\{[i_s,i_t]\}_{i=1}^M$ such that for any interval $[i_s,i_t] \in \mathcal{P}$, $C_i \le B/\sqrt{n_i}$ where $C_i := \sum_{j=i_s+1}^{i_t} \|\boldsymbol{u}_j - \boldsymbol{u}_{j-1}\|_1$ and n_i is the length of the interval.*

Define $\boldsymbol{\Gamma}_i^+ := \sum_{j=i_s}^{i_t} \boldsymbol{\gamma}_j^+$ and $\boldsymbol{\Gamma}_i^- := \sum_{j=i_s}^{i_t} \boldsymbol{\gamma}_j^-$. Let $\Delta\boldsymbol{s}_i = \boldsymbol{s}_{i_t} - \boldsymbol{s}_{i_s-1}$, where $\boldsymbol{s}$ is as defined in Section 5.1.2. We also have that each bin $[i_s,i_t] \in \mathcal{P}$ satisfies at-least one of the following properties.

Property 1 *Across each coordinate $k \in [d]$, the sequence $\boldsymbol{u}_j[k], j \in [i_s,i_t]$ is either non-decreasing or non-increasing.*

Property 2 $\|\lambda\Delta\boldsymbol{s}_i + \boldsymbol{\Gamma}_i^- - \boldsymbol{\Gamma}_i^+\|_2 \ge \lambda/4$.

The proof of the above lemma is deferred to Section D.4.

We recall Eq.(5.15) here for convenience. Let $\mathcal{P}$ be a partition of $[n]$ into M bins obtained in Lemma 176 Let $[i_s, i_t]$ denote the i^{th} bin in $\mathcal{P}$ and let n_i be its length. Define $\bar{\boldsymbol{u}}_i = \frac{1}{n_i}\sum_{j=i_s}^{i_t} \boldsymbol{u}_j$ and $\dot{\boldsymbol{u}}_i = \bar{\boldsymbol{u}}_i - \frac{1}{n_i\beta}\sum_{j=i_s}^{i_t} \nabla f_j(\bar{\boldsymbol{u}}_i)$ where β is as in Assumption EC-2. Let $\boldsymbol{x}_j$ be the prediction made by FLH at time j. We start with following regret decomposition.

$$R_n(C_n) \leq \sum_{i=1}^{M} \underbrace{\sum_{j=i_s}^{i_t} f_j(\boldsymbol{x}_j) - f_j(\dot{\boldsymbol{u}}_i)}_{T_{1,i}} + \sum_{i=1}^{M} \underbrace{\sum_{j=i_s}^{i_t} f_j(\dot{\boldsymbol{u}}_i) - f_j(\bar{\boldsymbol{u}}_i)}_{T_{2,i}} + \sum_{i=1}^{M} \underbrace{\sum_{j=i_s}^{i_t} f_j(\bar{\boldsymbol{u}}_i) - f_j(\boldsymbol{u}_j)}_{T_{3,i}}. \tag{D.62}$$

Lemma 177. *(**bounding** $T_{1,i}$) Let the experts in* FLH *be the ONS algorithms with parameter $\zeta = \min\left\{\frac{1}{4G^\dagger(2B\sqrt{d}+2G\sqrt{d})}, \alpha\right\}$ and decision set $\mathcal{D}$. Also choose learning rate $\eta = \alpha$, for* FLH*. Then for any bin $[i_s, i_t]$ we have,*

$$\sum_{j=i_s}^{i_t} f_j(\boldsymbol{x}_j) - f_j(\dot{\boldsymbol{u}}_i) = O\left(d^{3/2} B G^\dagger \log n + d^{3/2} G G^\dagger \log n + \frac{\log n}{\alpha}\right) \tag{D.63}$$

$$= O(d^{3/2} \log n), \tag{D.64}$$

where $\boldsymbol{x}_j \in \mathbb{R}^d$ are the outputs of FLH*.*

Proof. First we proceed to bound $\|\dot{\boldsymbol{u}}_i\|_\infty$. Since $\|\nabla f_j(\boldsymbol{u}_j)\|_2 \leq G$ by Assumption EC-1, we have

$$\|\dot{\boldsymbol{u}}_i\|_\infty \leq \|\bar{\boldsymbol{u}}_i\|_\infty + \frac{G}{\beta} \tag{D.65}$$

$$\leq B + \frac{G}{\beta} \tag{D.66}$$

$$\leq B + G, \tag{D.67}$$

where we used $\beta > 1$ from Assumption EC-2.

For any $\boldsymbol{x} \in \mathcal{D}$, we have $\|\boldsymbol{x} - \dot{\boldsymbol{u}}_i\|_2 \leq 2B\sqrt{d} + 2G\sqrt{d}$ by triangle inequality and the fact $\|\boldsymbol{y}\|_2 \leq \sqrt{d}\|\boldsymbol{y}\|_\infty$.

By Assumption EC-4 we have $\|\nabla f_j(\boldsymbol{x})\|_2 \leq G^\dagger$ for any $\boldsymbol{x} \in \mathcal{D}$. Also, recall that by Assumption EC-3, the loss functions f_j are α exp-concave in the domain $\mathcal{D}$. Let $\boldsymbol{p}_j$, $j \in [i_s, i_t]$ be the predictions of an ONS algorithm when run in the interval $[i_s, i_t]$. If we choose $\zeta = \min\left\{\frac{1}{4G^\dagger(2B\sqrt{d}+2G\sqrt{d})}, \alpha\right\}$ as the parameter of the ONS, Theorem 2 of [77] implies that

$$\sum_{j=i_s}^{i_t} f_j(\boldsymbol{p}_j) - f_j(\dot{\boldsymbol{u}}_i) = O\left(d^{3/2}BG^\dagger \log n + d^{3/2}GG^\dagger \log n + \frac{\log n}{\alpha}\right) \tag{D.68}$$

$$= O\left(d^{3/2}\log n\right). \tag{D.69}$$

Now the Lemma is implied by the SA regret bound of FLH (Theorem 3.2 of [23]). □

For strongly convex, losses the tern $T_{1,i}$ can enjoy a better bound.

Lemma 178. *(**bounding** $T_{1,i}$ **for strongly convex losses**) Suppose that the losses are H strongly convex. Take experts in* FLH *as OGD with step size* $1/(Hn)$ *and decision set* $\mathcal{D}$*. Also choose learning rate* $\eta = H/(G^\dagger)^2$*, for* FLH*. Then for any bin* $[i_s, i_t]$ *we have,*

$$\sum_{j=i_s}^{i_t} f_j(\boldsymbol{x}_j) - f_j(\dot{\boldsymbol{u}}_i) = O\left(\frac{(G^\dagger)^2 \log n}{H}\right), \tag{D.70}$$

where $\boldsymbol{x}_j \in \mathbb{R}^d$ *are the outputs of* FLH.

Proof Sketch. The lemma follows by using the regret bound of OGD with strongly convex losses from [77] and following similar lines of arguments as in Lemma 177. □

We state the next lemma to be generically valid for any bin which is not necessarily a member of $\mathcal{P}$.

Some notations. For a bin $[a,b]$, introduce the notations $\Delta\boldsymbol{s}_{a\to b} := s(\boldsymbol{u}_{b+1} - \boldsymbol{u}_b) - s(\boldsymbol{u}_a - \boldsymbol{u}_{a-1})$, $\boldsymbol{\Gamma}^+_{a\to b} = \sum_{j=a}^b \boldsymbol{\gamma}^+_j$ and $\boldsymbol{\Gamma}^-_{a\to b} := \sum_{j=a}^b \boldsymbol{\gamma}^-_j$. $n_{a\to b} := b - a + 1$. $\bar{\boldsymbol{u}}_{a\to b} = \frac{1}{n_{a\to b}} \sum_{j=a}^b \boldsymbol{u}_j$ and $\dot{\boldsymbol{u}}_{a\to b} = \bar{\boldsymbol{u}}_{a\to b} - \frac{1}{\beta n_{a\to b}} \sum_{j=a}^b \nabla f_j(\bar{\boldsymbol{u}}_{a\to b})$.

Lemma 179. *For any bin* $[a,b]$*, we have*

$$T_{2,[a,b]} := \sum_{j=a}^{b} f_j(\dot{\boldsymbol{u}}_{a\to b}) - f_j(\bar{\boldsymbol{u}}_i) \tag{D.71}$$

$$\le \frac{-\|\lambda\Delta\boldsymbol{s}_{a\to b} + \boldsymbol{\Gamma}^-_{a\to b} - \boldsymbol{\Gamma}^+_{a\to b}\|_2^2}{2n_{a\to b}\beta} + \|\lambda\Delta\boldsymbol{s}_{a\to b} + \boldsymbol{\Gamma}^-_{a\to b} - \boldsymbol{\Gamma}^+_{a\to b}\|_1 C_{a\to b}. \tag{D.72}$$

Proof. Let $g(\boldsymbol{x})$ be a α-strongly smooth function. Let $\boldsymbol{x}^+ = \boldsymbol{x} - \mu\nabla f(\boldsymbol{x})$ for some $\mu > 0$. Then we have

$$g(\boldsymbol{x}^+) - g(\boldsymbol{x}) \le \|\nabla g(\boldsymbol{x})\|_2^2 \left(\frac{\alpha}{2}\mu^2 - \mu\right) \tag{D.73}$$

$$= \frac{-\|\nabla g(\boldsymbol{x})\|_2^2}{2\alpha}, \tag{D.74}$$

by choosing $\mu = 1/\alpha$. By taking $g(x) = \sum_{j=a}^{b} f_j(\boldsymbol{x})$ and noting that g is $n_i\beta$ gradient Lipschitz due to Assumption SC-2, we get

$$T_{2,[a,b]} := \sum_{j=a}^{b} f_j(\dot{\boldsymbol{u}}_{a\to b}) - f_j(\bar{\boldsymbol{u}}_{a\to b}) \tag{D.75}$$

$$\leq \frac{-\left\|\sum_{j=a}^{b} \nabla f_j(\bar{\boldsymbol{u}}_{a\to b})\right\|_2^2}{2n_{a\to b}\beta} \tag{D.76}$$

$$= \frac{-1}{2n_{a\to b}\beta}\left\|\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j) + \nabla f_j(\bar{\boldsymbol{u}}_{a\to b}) - \nabla f_j(\boldsymbol{u}_j)\right\|_2^2 \tag{D.77}$$

$$\leq \frac{-1}{2n_{a\to b}\beta}\left\|\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)\right\|_2^2 + \frac{1}{n_{a\to b}\beta}\left\|\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)\right\|_1 \left\|\sum_{j=a}^{b} \nabla f_j(\bar{\boldsymbol{u}}_{a\to b}) - \nabla f_j(\boldsymbol{u}_j)\right\|_2, \tag{D.78}$$

where we used $\langle \boldsymbol{x}, \boldsymbol{y}\rangle \leq \|\boldsymbol{x}\|_2\|\boldsymbol{y}\|_2 \leq \|\boldsymbol{x}\|_1\|\boldsymbol{y}\|_2$ and dropped a negative term from expanding the squared norm. From the KKT conditions in Lemma 175 we have $\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j) = \lambda\Delta\boldsymbol{s}_{a\to b} + \boldsymbol{\Gamma}_{a\to b}^{-} - \boldsymbol{\Gamma}_{a\to b}^{+}$. Since f_j are β-gradient Lipschitz and $\|\bar{\boldsymbol{u}}_{a\to b} - \boldsymbol{u}_j\|_2 \leq \|\bar{\boldsymbol{u}}_{a\to b} - \boldsymbol{u}_j\|_1 \leq C_{a\to b}$, we also have

$$\left\|\sum_{j=a}^{b} \nabla f_j(\bar{\boldsymbol{u}}_{a\to b}) - \nabla f_j(\boldsymbol{u}_j)\right\|_2 \leq n_{a\to b}\beta C_{a\to b}. \tag{D.79}$$

Substituting these we get the statement of the lemma. □

Lemma 180. *For any bin $[i_s, i_t] \in \mathcal{P}$, we have*

$$\sum_{j=i_s}^{i_t} f_j(\bar{\boldsymbol{u}}_i) - f_j(\boldsymbol{u}_j) \leq \frac{\beta n_i C_i^2}{2} + 5\lambda C_i + \left\|\lambda\Delta\boldsymbol{s}_i + \boldsymbol{\Gamma}_i^{-} - \boldsymbol{\Gamma}_i^{+}\right\|_2 C_i. \tag{D.80}$$

Proof. Due to strong smoothness, we have

$$T_{3,i} := \sum_{j=i_s}^{i_t} f_j(\bar{\boldsymbol{u}}_i) - f_j(\boldsymbol{u}_j) \tag{D.81}$$

$$\leq_{(a)} \sum_{j=i_s}^{i_t} \langle \nabla f_j(\boldsymbol{u}_j), \bar{\boldsymbol{u}}_i - \boldsymbol{u}_j\rangle + \frac{\beta}{2}\|\bar{\boldsymbol{u}}_i - \boldsymbol{u}_j\|_1^2 \tag{D.82}$$

$$\leq \frac{\beta n_i C_i^2}{2} + \sum_{j=i_s}^{i_t} \langle \nabla f_j(\boldsymbol{u}_j), \bar{\boldsymbol{u}}_i - \boldsymbol{u}_j\rangle, \tag{D.83}$$

where in line (a) we used $\|\bar{\boldsymbol{u}}_i - \boldsymbol{u}_j\|_2 \le \|\bar{\boldsymbol{u}}_i - \boldsymbol{u}_j\|_1$.

Further,

$$\sum_{j=i_s}^{i_t} \langle \nabla f_j(\boldsymbol{u}_j), \bar{\boldsymbol{u}}_i - \boldsymbol{u}_j \rangle = \lambda \left(\langle \boldsymbol{s}_{i_s - 1}, \boldsymbol{u}_{i_s} - \bar{\boldsymbol{u}}_i \rangle - \langle \boldsymbol{s}_{i_t}, \boldsymbol{u}_{i_t} - \bar{\boldsymbol{u}}_i \rangle \right) \tag{D.84}$$

$$+ \lambda \sum_{j=i_s+1}^{i_t} \|\boldsymbol{u}_j - \boldsymbol{u}_{j-1}\|_1 + \sum_{j=i_s}^{i_t} \langle \boldsymbol{\gamma}_j^- - \boldsymbol{\gamma}_j^+, \bar{\boldsymbol{u}}_i - \boldsymbol{u}_j \rangle \tag{D.85}$$

By triangle and Holder's inequalities, the first two terms can be bounded by $3\lambda C_i$ (recall that $\|\boldsymbol{u}_t - \bar{\boldsymbol{u}}_i\|_1 \le C_i$ for all $t \in [i_s, i_t]$). Let's proceed to bound the last term in the above display. From Lemma 176, we have $C_i \le B/\sqrt{n_i}$. So the TV incurred across each coordinate of the optimal solution is at-most B. Using similar arguments as in Lemma 171, the complementary slackness in Lemma 175 implies that for each $k \in [d]$, if $\boldsymbol{\gamma}_j^-[k] > 0$ for at-least one $j \in [i_s, i_t]$ then $\boldsymbol{\gamma}_j^+[k] = 0$ for all $j \in [i_s, i_t]$. Similarly for each $k \in [d]$, if $\boldsymbol{\gamma}_j^+[k] > 0$ for at-least one $j \in [i_s, i_t]$ then $\boldsymbol{\gamma}_j^-[k] = 0$ for all $j \in [i_s, i_t]$. This observation allows us to write,

$$\sum_{j=i_s}^{i_t} |\boldsymbol{\gamma}_j^-[k] - \boldsymbol{\gamma}_j^+[k]| = \left| \boldsymbol{\Gamma}_i^-[k] - \boldsymbol{\Gamma}_i^+[k] \right| \tag{D.86}$$

Define $C_i^k := \sum_{j=i_s+1}^{i_t} |\boldsymbol{u}_j[k] - \boldsymbol{u}_{j-1}[k]|$. We have,

$$\begin{aligned}
&\sum_{j=i_s}^{i_t} \langle \boldsymbol{\gamma}_j^- - \boldsymbol{\gamma}_j^+, \bar{\boldsymbol{u}}_i - \boldsymbol{u}_j \rangle = \sum_{k=1}^{d} \sum_{j=i_s}^{i_t} (\boldsymbol{\gamma}_j^-[k] - \boldsymbol{\gamma}_j^+[k])(\bar{\boldsymbol{u}}_i[k] - \boldsymbol{u}_j[k]) \\
&\le_{(a)} \sum_{k=1}^{d} \left| \boldsymbol{\Gamma}_i^-[k] - \boldsymbol{\Gamma}_i^+[k] \right| C_i^k \\
&= \sum_{k=1}^{d} \left(\lambda \Delta \boldsymbol{s}_i[k] \operatorname{sign}\!\left(\boldsymbol{\Gamma}_i^-[k] - \boldsymbol{\Gamma}_i^+[k] \right) + \operatorname{sign}\!\left(\boldsymbol{\Gamma}_i^-[k] - \boldsymbol{\Gamma}_i^+[k] \right) \left(\boldsymbol{\Gamma}_i^-[k] - \boldsymbol{\Gamma}_i^+[k] \right) \right) C_i^k \\
&\qquad - \sum_{k=1}^{d} \lambda \Delta \boldsymbol{s}_i[k] \operatorname{sign}\!\left(\boldsymbol{\Gamma}_i^-[k] - \boldsymbol{\Gamma}_i^+[k] \right) C_i^k \\
&\le_{(b)} \left\| \lambda \Delta \boldsymbol{s}_i + \boldsymbol{\Gamma}_i^- - \boldsymbol{\Gamma}_i^+ \right\|_2 C_i + 2\lambda C_i,
\end{aligned}$$

where in line (a) we applied $\langle \boldsymbol{x}, \boldsymbol{y} \rangle \le \|\boldsymbol{x}\|_1 \|\boldsymbol{y}\|_\infty$ along with the Eq. (D.86). In line (b) we applied $\langle \boldsymbol{x}, \boldsymbol{y} \rangle \le \|\boldsymbol{x}\|_2 \|\boldsymbol{y}\|_1$ for the first term and $\langle \boldsymbol{x}, \boldsymbol{y} \rangle \le \|\boldsymbol{x}\|_\infty \|\boldsymbol{y}\|_1$ for the second term. Putting everything together yields the Lemma.

□

`splitMonotonic`: Inputs - (1) an interval $[i_s, i_t]$ such that the offline optimal is monotonic across each coordinate $k \in [d]$; (2) offline optimal sequences $\boldsymbol{u}_{1:n}$ and the sequence of subgradients (dual variables) $\boldsymbol{s}_{1:n-1}$ (recall that $\boldsymbol{s}_0 = \boldsymbol{s}_n = 0$ by convention.).

1. Initialize $\mathcal{T} \leftarrow \Phi$, $\mathcal{S} \leftarrow \Phi$.
2. Add i_s, i_t to $\mathcal{T}$.
3. For each coordinate $k \in [d]$:
   (a) If $\boldsymbol{u}[k]$ is constant in $[i_s, i_t]$, then skip the current coordinate.
   (b) Initialize $z_1 \leftarrow i_s, z_2 \leftarrow i_t$.
   (c) If $\boldsymbol{u}_{i_s}[k] = \pm B$, let z_1 be the first time point in $[i_s, i_t]$ where $\boldsymbol{u}_{z_1}[k] \neq \pm B$. Add $z_1 - 1, z_1$ to $\mathcal{T}$.
   (d) If $\boldsymbol{u}_{i_t}[k] = \pm B$, let z_2 be the last time point in $[i_s, i_t]$ where $\boldsymbol{u}_{z_2}[k] \neq \pm B$. Add $z_2, z_2 + 1$ to $\mathcal{T}$.
   (e) If $\boldsymbol{u}[k]$ is non-decreasing in $[i_s, i_t]$ then let $p \geq z_1$ be the first point with $\boldsymbol{s}_{p-1}[k] = 1$. If $p > z_1$, add $p-1, p$ to $\mathcal{T}$.
   (f) If $\boldsymbol{u}[k]$ is non-decreasing in $[i_s, i_t]$ then let $q \leq z_2$ be the last point with $\boldsymbol{s}_q[k] = 1$. If $q < z_2$, add $q, q+1$ to $\mathcal{T}$.
   (g) If $\boldsymbol{u}[k]$ is non-increasing in $[i_s, i_t]$ then let $p \geq z_1$ be the first point with $\boldsymbol{s}_{p-1}[k] = -1$. If $p > z_1$, add $p-1, p$ to $\mathcal{T}$.
   (h) If $\boldsymbol{u}[k]$ is non-increasing in $[i_s, i_t]$ then let $q \leq z_2$ be the last point with $\boldsymbol{s}_q[k] = -1$. If $q < z_2$, add $q, q+1$ to $\mathcal{T}$.
4. For each entry t in $\mathcal{T}$:
   (a) If t appears more than 2 times, delete some occurences of t such that t only appears 2 times in $\mathcal{T}$.
5. Sort $\mathcal{T}$ in non-decreasing order. For each consecutive points $s, t \in \mathcal{T}$, add $[s, t]$ to $\mathcal{S}$.
6. Return the partition $\mathcal{S}$.

Figure D.8: `splitMonotonic` procedure

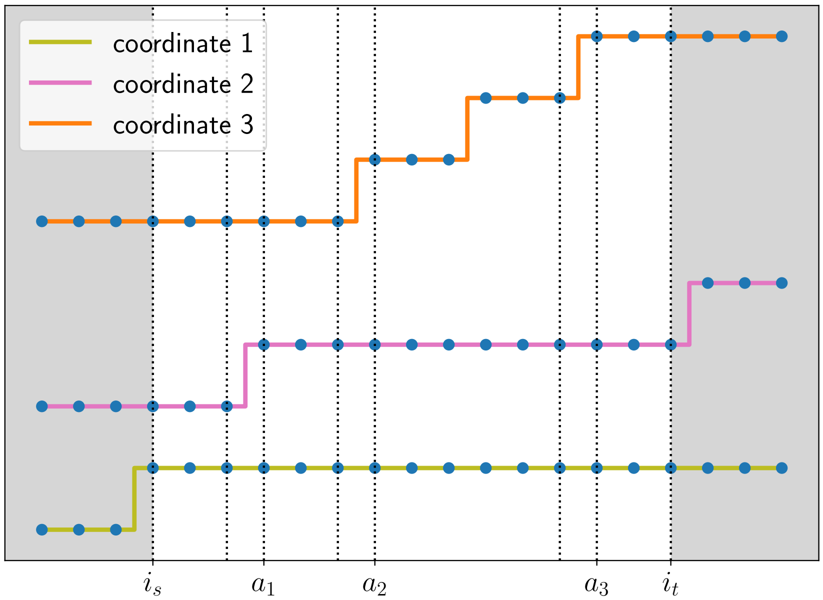


Figure D.9: *An example of the partitioning created by* `splitMonotonic` *(See Fig. D.8). The partition $\mathcal{S}$ returned by* `splitMonotonic` is $\{[i_s, a_1 - 1], [a_1, a_2 - 1], [a_2, a_3 - 1], [a_3, i_t]\}$. Blue dots indicate the offline optimal sequence.

Lemma 181. *Let* `splitMonotonic` *in Fig.D.8 be run with an input $[i_s, i_t]$. Then the partition $\mathcal{S}$ it return obeys $|\mathcal{S}| = O(d)$.*

Proof. From the psuedo-code in Fig. D.8 it is obvious that each coordinate can contribute to increasing the bin count by $O(1)$. Hence the overall bin count in $\mathcal{S}$ is $O(d)$. □

An illustrative example of the input and output of `splitMonotonic` is given in Fig. D.9.

Theorem 47. *By using the base learner as ONS with parameter $\zeta = \min\left\{\frac{1}{4G^\dagger(2B\sqrt{d}+2G/\beta)}, \alpha\right\}$, decision set $\mathcal{D}$ and choosing learning rate $\eta = \alpha$, FLH obeys $R_n^+(C_n) = \tilde{O}\left(d^{3.5}(n^{1/3}C_n^{2/3} \vee 1)\right)$ if $C_n > 1/n$ and $O(d^{1.5}\log n)$ otherwise. Here $a \vee b := \max\{a, b\}$ and $\tilde{O}(\cdot)$ hides dependence on the constants $B, G, G^\dagger, \alpha$ and factors of $\log n$.*

Proof. Consider a bin $[i_s, i_t] \in \mathcal{P}$. By Lemma 176, the bin has to satisfy one of the two Properties. Let's first focus on the scenario where $[i_s, i_t]$ satisfies Property 2.

Combining the results of Lemmas 179, 180 we can write,

$$T_{2,i} + T_{3,i} \leq \frac{-\|\lambda\Delta \boldsymbol{s}_i + \boldsymbol{\Gamma}_i^- - \boldsymbol{\Gamma}_i^+\|_2^2}{2n_i\beta} + \|\lambda\Delta \boldsymbol{s}_i + \boldsymbol{\Gamma}_i^- - \boldsymbol{\Gamma}_i^+\|_2 C_i \tag{D.87}$$

$$+ \frac{\beta n_i C_i^2}{2} + 5\lambda C_i + \left\|\lambda\Delta \boldsymbol{s}_i + \boldsymbol{\Gamma}_i^- - \boldsymbol{\Gamma}_i^+\right\|_2 C_i \tag{D.88}$$

$$\leq_{(a)} \frac{\beta B^2}{2} - \frac{\|\lambda\Delta \boldsymbol{s}_i + \boldsymbol{\Gamma}_i^- - \boldsymbol{\Gamma}_i^+\|_2^2}{2n_i\beta} + 7\left(\left\|\lambda\Delta \boldsymbol{s}_i + \boldsymbol{\Gamma}_i^- - \boldsymbol{\Gamma}_i^+\right\|_2 \vee \lambda\right) C_i \tag{D.89}$$

$$= \frac{\beta B^2}{2} - \left(\frac{\|\lambda\Delta \boldsymbol{s}_i + \boldsymbol{\Gamma}_i^- - \boldsymbol{\Gamma}_i^+\|_2}{\sqrt{2n_i\beta}} - \frac{7C_i\sqrt{n_i\beta}\left(\left\|\lambda\Delta \boldsymbol{s}_i + \boldsymbol{\Gamma}_i^- - \boldsymbol{\Gamma}_i^+\right\|_2 \vee \lambda\right)}{\|\lambda\Delta \boldsymbol{s}_i + \boldsymbol{\Gamma}_i^- - \boldsymbol{\Gamma}_i^+\|_2\sqrt{2}} \right)^2 \tag{D.90}$$

$$+ \frac{49n_i\beta C_i^2}{2}\left(\frac{\left\|\lambda\Delta \boldsymbol{s}_i + \boldsymbol{\Gamma}_i^- - \boldsymbol{\Gamma}_i^+\right\|_2 \vee \lambda}{\|\lambda\Delta \boldsymbol{s}_i + \boldsymbol{\Gamma}_i^- - \boldsymbol{\Gamma}_i^+\|_2} \right)^2 \tag{D.91}$$

$$\leq_{(b)} \frac{\beta B^2}{2} + 392\beta n_i C_i^2 \tag{D.92}$$

$$\leq 393\beta B^2, \tag{D.93}$$

where in line (a) we used $C_i \leq B/\sqrt{n_i}$ for partitions in $\mathcal{P}$ (Lemma 176). In line (b) we used $\frac{\|\lambda\Delta \boldsymbol{s}_i + \boldsymbol{\Gamma}_i^- - \boldsymbol{\Gamma}_i^+\|_2 \vee \lambda}{\|\lambda\Delta \boldsymbol{s}_i + \boldsymbol{\Gamma}_i^- - \boldsymbol{\Gamma}_i^+\|_2} \geq 4$ since $\|\lambda\Delta \boldsymbol{s}_i + \boldsymbol{\Gamma}_i^- - \boldsymbol{\Gamma}_i^+\|_2 \geq \lambda/4$ by Property 2 of Lemma 176.

Now using Lemma 177, for the bins $[i_s, i_t]$ that satisfy property 2, we can write

$$T_{1,i} + T_{2,i} + T_{3,i} = \tilde{O}(d^{1.5}). \tag{D.94}$$

Now suppose that the bin $[\underline{t}, \bar{t}]$ satisfies Property 1 in Lemma 176. In this case, via a call to `splitMonotonic` function with the input interval as $[\underline{t}, \bar{t}]$, we split the original bin into $O(d)$ sub-bins (see Lemma 181). Further for a fixed k, if $\boldsymbol{u}_j[k]$, $j \in [\underline{t}, \bar{t}]$ is non-decreasing, then we can group those consecutive sub-bins into at-most three categories: (a) a section of time where $\boldsymbol{u}_j[k]$ is constant; (b) a section of time where $\boldsymbol{u}_j[k]$ is non-decreasing; (c) a section of time where $\boldsymbol{u}_j[k]$ is constant.

We proceed to define these sections formally (where p, m, q are indices defined for convenience)

- For section (a) let $\mathcal{A} = \{[\underline{t}, \underline{t}_{-p} - 1], [\underline{t}_{-p}, \bar{t}_{-p}], \ldots [\underline{t}_0, \bar{t}_0]\}$
- For section (b) let $\mathcal{B} = \{[\underline{t}_1, \bar{t}_1], \ldots, [\underline{t}_m, \bar{t}_m]\}$
- For section (c) let $\mathcal{C} = \{[\underline{t}_{m+1}, \bar{t}_{m+1}], \ldots, [\bar{t}_q + 1, \bar{t}]\}$

As mentioned before, these sections are constructed so that the offline optimal satisfy the following properties.

(i) $\boldsymbol{u}_j[k]$ $j \in [\underline{t}, \bar{t}_0]$ is constant.

(ii) $\boldsymbol{u}_{\underline{t}_1}[k] > \boldsymbol{u}_{\underline{t}_1-1}[k]$ and $\boldsymbol{u}_{\underline{t}_{m+1}}[k] > \boldsymbol{u}_{\underline{t}_{m+1}-1}[k]$.

(iii) $\boldsymbol{u}_j[k]$ $j \in [\underline{t}_1, \bar{t}_m]$ is non-decreasing.

(iv) $\boldsymbol{u}_j[k]$ $j \in [\underline{t}_{m+1}, \bar{t}]$ is constant.

We remark that the grouping may be different for different coordinates k. Further some of $\mathcal{A}, \mathcal{B}$ or $\mathcal{C}$ can be empty. In the example we gave in Fig. D.9:

- For coordinate 1 $\mathcal{A} = \phi$, $\mathcal{B} = \phi$, $\mathcal{C} = \{[i_s, a_1 - 1], [a_1, a_2 - 1], [a_2, a_3 - 1], [a_3, i_t]\}$.
- For coordinate 2 $\mathcal{A} = [i_s, a_1 - 1]$, $\mathcal{B} = \phi$, $\mathcal{C} = \{[a_1, a_2 - 1], [a_2, a_3 - 1], [a_3, i_t]\}$.
- For coordinate 3 $\mathcal{A} = \{[i_s, a_1 - 1], [a_1, a_2 - 1]\}$, $\mathcal{B} = \{[a_2, a_3 - 1]\}$, $\mathcal{C} = \{[a_3, i_t]\}$

We fill focus on the aforementioned scenario where $\boldsymbol{u}_j[k]$, $j \in [\underline{t}, \bar{t}]$ is non-decreasing. The arguments for the case where $\boldsymbol{u}_j[k]$, $j \in [\underline{t}, \bar{t}]$ is non-increasing are similar. Further similar to the proof of Theorem 168, we give arguments for the case where $\gamma_j^+[k] = 0$ for all j in the interval $[\underline{t}, \bar{t}]$ stating that arguments for the case $\gamma_j^-[k] = 0$ uniformly in $[\underline{t}, \bar{t}]$ are similar.

From Lemma 179, we have

$$\sum_{j=a}^{b} f_j(\dot{\boldsymbol{u}}_{a\to b}) - f_j(\bar{\boldsymbol{u}}_{a\to b}) \leq \frac{-\|\lambda \Delta \boldsymbol{s}_{a\to b} + \boldsymbol{\Gamma}^-_{a\to b} - \boldsymbol{\Gamma}^+_{a\to b}\|_2^2}{2n_{a\to b}\beta} + \|\lambda \Delta \boldsymbol{s}_{a\to b} + \boldsymbol{\Gamma}^-_{a\to b} - \boldsymbol{\Gamma}^+_{a\to b}\|_1 C_{a\to b}. \tag{D.95}$$

Observe that the relation in Eq. (D.83) holds for any generic bin $[a, b]$ that may not be a member of $\mathcal{P}$ (replacing $C_i, n_i, \bar{\boldsymbol{u}}_i$ with $C_{a\to b}, n_{a\to b}, \bar{\boldsymbol{u}}_{a\to b}$). So

$$T_{3,[a,b]} := \sum_{j=a}^{b} f_j(\bar{\boldsymbol{u}}_{a\to b}) - f_j(\boldsymbol{u}_j) \leq \sum_{k=1}^{d} \frac{\beta n_{a\to b} C^2_{a\to b}}{2d} + \sum_{j=a}^{b} \langle \nabla f_j(\boldsymbol{u}_j), \bar{\boldsymbol{u}}_{a\to b} - \boldsymbol{u}_j \rangle. \tag{D.96}$$

Note that Eq. (D.95) and (D.96) decompose coordinate-wise. So for the bin $[\underline{t}, \bar{t}] \in \mathcal{P}$ where the optimal sequence is monotonic across each coordinate, our strategy is to bound

$$\begin{aligned}\boldsymbol{S}_{a\to b}[k] := & \frac{-\left(\lambda \Delta \boldsymbol{s}_{a\to b}[k] + \boldsymbol{\Gamma}^-_{a\to b}[k] - \boldsymbol{\Gamma}^+_{a\to b}[k]\right)^2}{2n_{a\to b}\beta} + |\lambda \Delta \boldsymbol{s}_{a\to b}[k] + \boldsymbol{\Gamma}^-_{a\to b}[k] - \boldsymbol{\Gamma}^+_{a\to b}[k]| C_{a\to b} \\ & + \frac{\beta n_{a\to b} C^2_{a\to b}}{2d} + \sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k](\bar{\boldsymbol{u}}_{a\to b}[k] - \boldsymbol{u}_j[k]),\end{aligned} \tag{D.97}$$

for each $k \in [d]$ and $[a,b] \in \mathcal{A} \cup \mathcal{B} \cup \mathcal{C}$ and finally adding them across all coordinates to bound $\sum_{k=1}^{d} \boldsymbol{S}_{a\to b}[k]$. Doing so will result in a bound on $T_{2,[a,b]} + T_{3,[a,b]}$. Further, $T_{1,[a,b]}$ can be bound by strongly adaptive regret. This enables us to bound $\sum_{[a,b]\in\mathcal{A}\cup\mathcal{B}\cup\mathcal{C}} T_{1,[a,b]} + T_{2,[a,b]} + T_{3,[a,b]}$ thereby leading to a regret bound in the parent bin $[\underline{t}, \bar{t}] \in \mathcal{P}$ which was the input interval for the call to `splitMonotonic` that we started with.

Let $C_{[a,b]}[k]$ be the TV of offline optimal incurred in the interval any interval $[a,b]$ along coordinate k. First we focus on the bins in $\mathcal{B}$. If $\mathcal{B}$ is not empty, then $\boldsymbol{\gamma}_j[k] = 0 \,\forall j \in [\underline{t}_1 \to \bar{t}_m]$ due to property (ii) and (iii) above. By using the stationarity conditions in Lemma 175, we can write

$$\sum_{[a,b]\in\mathcal{B}} \sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k](\bar{\boldsymbol{u}}_{a\to b}[k] - \boldsymbol{u}_j[k]) = \lambda C_{\underline{t}_1\to\bar{t}_m}[k] + \lambda\left(s_{\underline{t}_1-1}[k]\boldsymbol{u}_{\underline{t}_1}[k] - s_{\bar{t}_m}[k]\boldsymbol{u}_{\bar{t}_m}[k]\right) \tag{D.98}$$

$$+ \sum_{[a,b]\in\mathcal{B}} \lambda \bar{\boldsymbol{u}}_{a\to b}[k]\Delta\boldsymbol{s}_{a\to b}[k]. \tag{D.99}$$

So we have,

$$\sum_{[a,b]\in\mathcal{B}} \boldsymbol{S}_{a\to b}[k] \leq \frac{\beta n_{\underline{t}_1\to\bar{t}_m} C^2_{\underline{t}_1\to\bar{t}_m}}{2d} + \lambda C_{\underline{t}_1\to\bar{t}_m}[k] + \lambda\left(s_{\underline{t}_1-1}[k]\boldsymbol{u}_{\underline{t}_1}[k] - s_{\bar{t}_m}[k]\boldsymbol{u}_{\bar{t}_m}[k]\right) \tag{D.100}$$

$$+ \sum_{[a,b]\in\mathcal{B}} \frac{-\lambda^2(\Delta\boldsymbol{s}_{a\to b}[k])^2}{2n_{a\to b}\beta} + \lambda|\Delta\boldsymbol{s}_{a\to b}[k]|C_{\underline{t}_1\to\bar{t}_m} + \lambda\bar{\boldsymbol{u}}_{a\to b}[k]\Delta\boldsymbol{s}_{a\to b}[k] \tag{D.101}$$

$$\leq_{(a)} \frac{\beta n_{\underline{t}\to\bar{t}} C^2_{\underline{t}\to\bar{t}}}{2d} + \lambda C_{\underline{t}_1\to\bar{t}_m}[k] + \lambda\left(s_{\underline{t}_1-1}[k]\boldsymbol{u}_{\underline{t}_1}[k] - s_{\bar{t}_m}[k]\boldsymbol{u}_{\bar{t}_m}[k]\right) \tag{D.102}$$

$$+ \sum_{[a,b]\in\mathcal{B}} \frac{-\lambda^2(\Delta\boldsymbol{s}_{a\to b}[k])^2}{2n_{a\to b}\beta} + \lambda|\Delta\boldsymbol{s}_{a\to b}[k]|C_{\underline{t}\to\bar{t}} + \lambda\bar{\boldsymbol{u}}_{a\to b}[k]\Delta\boldsymbol{s}_{a\to b}[k] \tag{D.103}$$

$$\leq_{(b)} \frac{\beta B^2}{2d} + \sum_{[a,b]\in\mathcal{B}} \frac{-\lambda^2(\Delta\boldsymbol{s}_{a\to b}[k])^2}{2n_{a\to b}\beta} + \lambda|\Delta\boldsymbol{s}_{a\to b}[k]|C_{\underline{t}\to\bar{t}} + \lambda\bar{\boldsymbol{u}}_{a\to b}[k]\Delta\boldsymbol{s}_{a\to b}[k], \tag{D.104}$$

where in line (a) we used the fact that $C_{\underline{t}_1\to\bar{t}_m} \leq C_{\underline{t}\to\bar{t}}$ and $n_{\underline{t}_1\to\bar{t}_m} \leq n_{\underline{t}\to\bar{t}}$ since $[\underline{t}_1, \bar{t}_m]$ is contained within $[\underline{t}, \bar{t}]$. In line (b) we used $C_{\underline{t}\to\bar{t}} \leq B/\sqrt{n_{\underline{t}\to\bar{t}}}$ (since $[\underline{t}, \bar{t}] \in \mathcal{P}$) along with the fact that
$\lambda\left(\boldsymbol{s}_{\underline{t}_1-1}[k]\boldsymbol{u}_{\underline{t}_1}[k] - \boldsymbol{s}_{\bar{t}_m}[k]\boldsymbol{u}_{\bar{t}_m}[k]\right) = -\lambda C_{\underline{t}_1\to\bar{t}_m}$ since
$\boldsymbol{s}_{\underline{t}_1-1}[k] = \boldsymbol{s}_{\bar{t}_m} = 1$ due to property (ii) and (iii) above.

Define $\check{\boldsymbol{u}}_{\mathcal{B}} := \frac{1}{|\mathcal{B}|}\sum_{[a,b]\in\mathcal{B}} \bar{\boldsymbol{u}}_{a\to b}$. Observe that since $\boldsymbol{s}_{\underline{t}_1-1}[k] = \boldsymbol{s}_{\bar{t}_m}[k] = 1$, we can write $\sum_{[a,b]\in\mathcal{B}} \Delta\boldsymbol{s}_{a\to b}[k] = 0$ by the telescoping structure.

By noting that we can subtract $0 = \lambda\check{\boldsymbol{u}}_{\mathcal{B}}[k]\sum_{(a,b)\in\mathcal{B}} \Delta\boldsymbol{s}_{a\to b}[k]$ and that $|\bar{\boldsymbol{u}}_{a\to b}[k] - \check{\boldsymbol{u}}_{\mathcal{B}}[k]| \le C_{\underline{t}\to\bar{t}}$, we have

$$\begin{aligned}
\sum_{[a,b]\in\mathcal{B}} \boldsymbol{S}_{a\to b}[k] &\le \frac{\beta B^2}{2d} + \sum_{[a,b]\in\mathcal{B}} \frac{-\lambda^2(\Delta\boldsymbol{s}_{a\to b}[k])^2}{2n_{a\to b}\beta} + \lambda|\Delta\boldsymbol{s}_{a\to b}[k]|C_{\underline{t}\to\bar{t}} \\
&\quad + \lambda\left(\bar{\boldsymbol{u}}_{a\to b}[k] - \check{\boldsymbol{u}}_{\mathcal{B}}[k]\right)\Delta\boldsymbol{s}_{a\to b}[k] \\
&\le \frac{\beta B^2}{2d} + \sum_{[a,b]\in\mathcal{B}} \frac{-\lambda^2(\Delta\boldsymbol{s}_{a\to b}[k])^2}{2n_{a\to b}\beta} + 2\lambda|\Delta\boldsymbol{s}_{a\to b}[k]|C_{\underline{t}\to\bar{t}} \\
&= \frac{\beta B^2}{2d} + \sum_{[a,b]\in\mathcal{B}} -\left(\frac{\lambda\Delta\boldsymbol{s}_{a\to b}[k]}{\sqrt{2n_{a\to b}\beta}} - C_{\underline{t}\to\bar{t}}\sqrt{2n_{a\to b}\beta}\right)^2 + 2\beta n_{a\to b}C^2_{\underline{t}\to\bar{t}} \\
&\le \frac{\beta B^2}{2d} + 2\beta n_{\underline{t}\to\bar{t}}C^2_{\underline{t}\to\bar{t}} \\
&\le \frac{\beta B^2}{2d} + 2\beta B^2 \\
&\le 3\beta B^2.
\end{aligned}$$

Next, we address bins present in $\mathcal{A}$ and $\mathcal{C}$. We provide the arguments for bounding $\sum_{[a,b]\in\mathcal{A}} \boldsymbol{S}_{a\to b}[k]$. Bounding the sum for bins in $\mathcal{C}$ can be done using similar arguments.

Observe that by property (i) above, the sequence $\boldsymbol{u}_j[k]$ for $j \in [i_s, \bar{t}_0]$ is a constant. So the last term in Eq. (D.97) is zero for any $\boldsymbol{S}_{a\to b}[k]$ where $[a,b]\in\mathcal{A}$. Now proceeding similar to above by completing the squares and dropping the negative terms, we get

$$\sum_{[a,b]\in\mathcal{A}} \boldsymbol{S}_{a\to b}[k] \leq \sum_{[a,b]\in\mathcal{A}} \left(\frac{-\left(\lambda\Delta \boldsymbol{s}_{a\to b}[k] + \boldsymbol{\Gamma}^-_{a\to b}[k] - \boldsymbol{\Gamma}^+_{a\to b}[k]\right)^2}{2n_{a\to b}\beta} \right. \tag{D.105}$$

$$\left. + |\lambda\Delta \boldsymbol{s}_{a\to b}[k] + \boldsymbol{\Gamma}^-_{a\to b}[k] - \boldsymbol{\Gamma}^+_{a\to b}[k]| C_{a\to b} + \frac{\beta n_{a\to b} C^2_{a\to b}}{2d} \right) \tag{D.106}$$

$$= \sum_{[a,b]\in\mathcal{A}} \left(-\left(\frac{\lambda\Delta \boldsymbol{s}_{a\to b}[k] + \boldsymbol{\Gamma}^-_{a\to b}[k] - \boldsymbol{\Gamma}^+_{a\to b}[k]}{\sqrt{2n_{a\to b}\beta}} - C_{a\to b}\sqrt{\frac{n_{a\to b}\beta}{2}} \right)^2 \right. \tag{D.107}$$

$$\left. + \frac{n_{a\to b}\beta C^2_{a\to b}}{2} + \frac{\beta n_{a\to b} C^2_{a\to b}}{2d} \right) \tag{D.108}$$

$$\leq \sum_{[a,b]\in\mathcal{A}} n_{a\to b}\beta C^2_{\underline{t}\to\bar{t}} \tag{D.109}$$

$$\leq n_{\underline{t}\to\bar{t}}\beta C^2_{\underline{t}\to\bar{t}} \tag{D.110}$$

$$\leq \beta B^2. \tag{D.111}$$

Similarly it can be shown that $\sum_{[a,b]\in\mathcal{C}} \boldsymbol{S}_{a\to b}[k] = O(1)$. Recalling that $|\mathcal{A}| + |\mathcal{B}| + |\mathcal{C}| = O(d)$ we have

$$T_{2,[\underline{t},\bar{t}]} + T_{3,[\underline{t},\bar{t}]} \leq \sum_{k=1}^{d} \sum_{[a,b]\in\mathcal{A}\cup\mathcal{B}\cup\mathcal{C}} \boldsymbol{S}_{a\to b}[k] = O(d). \tag{D.112}$$

From Lemma 177 we have

$$T_{1,[\underline{t},\bar{t}]} = \tilde{O}(d^{2.5}), \tag{D.113}$$

for bins $[\underline{t},\bar{t}] \in \mathcal{P}$ that satisfy property 2 in Lemma 176.

Comparing Eq. (D.94) and (D.113) we conclude that

$$T_{1,i} + T_{2,i} + T_{3,i} = \tilde{O}(d^{2.5}), \tag{D.114}$$

for all bins $[i_s, i_t]$ in the partition $\mathcal{P}$ of Lemma 176. Since $|\mathcal{P}| = O(dn^{1/3}C_n^{2/3})$, adding the above bound across all bins leads to the theorem.

If $C_n \leq 1/n$, then we have

$$\sum_{t=1}^{n} f_t(\boldsymbol{x}_j) - f_t(\boldsymbol{u}_t) \leq \sum_{t=1}^{n} f_t(\boldsymbol{x}_j) - f_t(\boldsymbol{u}_1) + \sum_{t=1}^{n} f_t(\boldsymbol{u}_1) - f_t(\boldsymbol{u}_t) \tag{D.115}$$

$$\leq_{(a)} \tilde{O}(d^{1.5}) + G^{\dagger} n C_n \tag{D.116}$$

$$= \tilde{O}(d^{1.5}) \tag{D.117}$$

where line (a) follows from the fact that f_t is $G^{\dagger}$ Lipschitz in $\mathcal{D}$.

□

Proposition 48. *For strongly convex losses, the regret bound can be improved to* $\tilde{O}\left(d^2(n^{1/3}C_n^{2/3} \vee 1)\right)$ *if* $C_n > 1/n$ *and* $O(\log n)$ *otherwise by using OGD as base learners in the FLH procedure. See Appendix D.3.2 for a proof.*

Proof Sketch. First we consider the case where the offline optimal in monotonic in each coordinate of a bin in $\mathcal{P}$. The static regret in any bin for strongly convex losses is $O(\log n)$ by Lemma 178 (as opposed to $\tilde{O}(d^{1.5})$ for exp-concave losses). Hence Eq.(D.113) can be re-written as $T_{1,[\underline{t},\bar{t}]} = \tilde{O}(d)$. By following similar arguments as in proof of Theorem 47, we can re-write Eq.(D.114) as

$$T_{1,i} + T_{2,i} + T_{3,i} = \tilde{O}(d). \tag{D.118}$$

If the offline optimal is not monotonic, in each coordinate, we can write

$$T_{1,i} + T_{2,i} + T_{3,i} = \tilde{O}(1), \tag{D.119}$$

by following similar arguments for the corresponding case in the proof of Theorem 47.

Finally we sum across all $|\mathcal{P}| = O(dn^{1/3}C_n^{2/3})$. The case $C_n \leq 1/n$ can be handled similar to that of the exp-concave case. □

D.4 Technical Lemmas

We start by describing a partitioning procedure namely `generateBins`.

`generateBins`: Inputs - the offline optimal sequence.

Step 1 Initialize $\mathcal{Q} \leftarrow \Phi$. Starting from time 1, spawn a new bin $[i_s, i_t]$ whenever $\sum_{j=i_s+1}^{i_t+1} \|\boldsymbol{u}_j - \boldsymbol{u}_{j-1}\|_1 > B/\sqrt{n_i}$, where $n_i = i_t - i_s + 1$. Add the spawned bin $[i_s, i_t]$ to $\mathcal{Q}$.

Step 2 Initialize $\mathcal{P} \leftarrow \Phi, \mathcal{R} \leftarrow \Phi$.

Step 3 For each bin $[i_s, i_t] \in \mathcal{Q}$:

(a) Let $\Delta \boldsymbol{s}_i = \boldsymbol{s}_{i_t} - \boldsymbol{s}_{i_s-1}$. $\boldsymbol{\Gamma}_i^+ = \sum_{j=i_s}^{i_t} \boldsymbol{\gamma}_j^+$. $\boldsymbol{\Gamma}_i^- = \sum_{j=i_s}^{i_t} \boldsymbol{\gamma}_j^-$.

(b) If for each $k \in [d]$, the sequence $\boldsymbol{u}_k$ is monotonic in $[i_s, i_t]$, then remove $[i_s, i_t]$ from $\mathcal{Q}$ and add it to $\mathcal{P}$.

(c) If there exists one coordinate $k \in [d]$ such that $\boldsymbol{s}_{i_s-1}[k] \in [-1, -1/4]$ and $\boldsymbol{s}_{i_t}[k] \in [0, 1]$ and $\boldsymbol{\gamma}_j^+[k] = 0 \, \forall j \in [i_s, i_t]$, then remove $[i_s, i_t]$ from $\mathcal{Q}$ and add it to $\mathcal{P}$. Goto Step 3.

(d) If there exists one coordinate $k \in [d]$ such that $\boldsymbol{s}_{i_s-1}[k] \in [-1/4, 0]$ and $\boldsymbol{s}_{i_t}[k] \in [1/4, 1]$ and $\boldsymbol{\gamma}_j^+[k] = 0 \, \forall j \in [i_s, i_t]$, then remove $[i_s, i_t]$ from $\mathcal{Q}$ and add it to $\mathcal{P}$. Goto Step 3.

(e) If there exists one coordinate $k \in [d]$ such that $\boldsymbol{u}_k$ is non-monotonic in $[i_s, i_t]$ and $\boldsymbol{s}_{i_s-1}[k] \in [-1/4, 1]$ and $\boldsymbol{s}_{i_t}[k] \in [-1/4, 1]$ and $\boldsymbol{\gamma}_j^+[k] = 0 \, \forall j \in [i_s, i_t]$ then:

   i. Initialize $z \leftarrow i_s$. Remove $[i_s, i_t]$ from $\mathcal{Q}$.

   ii. if $\boldsymbol{u}_{i_s}[k] = -B$, then split $[i_s, i_t]$ into $[i_s, a]$ and $[a+1, i_t]$ where a is the first time point within $[i_s, i_t]$ such that $\boldsymbol{u}_a[k] > -B$. Add $[i_s, a-1]$ to $\mathcal{Q}$. Set $z \leftarrow a$.

   iii. Let j be the first time in $[z, i_t]$ such that $\boldsymbol{s}_{j-1}[k] = -1$ with $\boldsymbol{u}_j[k] < \boldsymbol{u}_{j-1}[k]$. Add $[z, j-1]$ and $[j, i_t]$ to $\mathcal{P}$. Goto Step 3.

(f) If there exists one coordinate $k \in [d]$ such that $\boldsymbol{u}_k$ is non-monotonic in $[i_s, i_t]$ and $\boldsymbol{s}_{i_s-1}[k], \boldsymbol{s}_{i_t}[k] \in [-1/4, 1/4]$ and $\boldsymbol{\gamma}_j^+[k] = 0 \, \forall j \in [i_s, i_t]$ then:

   i. Initialize $z \leftarrow i_s$. Remove $[i_s, i_t]$ from $\mathcal{Q}$.

   ii. if $\boldsymbol{u}_{i_s}[k] = -B$, then split $[i_s, i_t]$ into $[i_s, a]$ and $[a+1, i_t]$ where a is the first time point within $[i_s, i_t]$ such that $\boldsymbol{u}_a[k] > -B$. Add $[i_s, a]$ to $\mathcal{Q}$. Set $z \leftarrow a+1$.

   iii. Let j be the first time in $[z, i_t]$ such that $\boldsymbol{s}_{j-1}[k] = -1$ with $\boldsymbol{u}_j[k] < \boldsymbol{u}_{j-1}[k]$. . Add $[z, j-1]$ and $[j, i_t]$ to $\mathcal{P}$. Goto Step 3.

(g) If there exists one coordinate $k \in [d]$ $\boldsymbol{u}_k$ is non-monotonic in $[i_s, i_t]$ and such that $\boldsymbol{s}_{i_s-1}[k] \in [-1, -1/4]$ and $\boldsymbol{s}_{i_t}[k] \in [-1, 1/4]$ and $\boldsymbol{\gamma}_j^+[k] = 0 \, \forall j \in [i_s, i_t]$ then:

   i. Initialize $z \leftarrow i_t$. Remove $[i_s, i_t]$ from $\mathcal{Q}$.

   ii. If $\boldsymbol{u}_{i_t}[k] = -B$, then split $[i_s, i_t]$ into $[i_s, a]$ and $[a+1, i_t]$ where a is the last time point within $[i_s, i_t]$ such that $\boldsymbol{u}_a[k] > -B$. Add $[a+1, i_t]$ to $\mathcal{Q}$. Set $z \leftarrow a$.

   iii. Let j be the last time in $[i_s, z]$ such that $\boldsymbol{s}_{j-1}[k] = 1$ with $\boldsymbol{u}_j[k] > \boldsymbol{u}_{j-1}[k]$.. Add $[i_s, j-1]$ and $[j, z]$ to $\mathcal{P}$. Goto Step 3.

(h) If there exists a coordinate $k \in [d]$ such that $\boldsymbol{u}_k$ is non-monotonic in $[i_s, i_t]$ and $\boldsymbol{s}_{i_s-1}[k] \in [-1/4, 1]$ and $\boldsymbol{s}_{i_t}[k] \in [-1, 1/4]$ and $\boldsymbol{\gamma}_j^+[k] = 0, \forall j \in [i_s, i_t]$ then:

   i. Initialize $p \leftarrow i_s - 1$. Remove $[i_s, i_t]$ from $\mathcal{Q}$.

   ii. If $\boldsymbol{u}_{i_s}[k] = -B$, then let p be the largest point in $[i_s, i_t]$ such that $\boldsymbol{u}_t[k] = -B \, \forall t \in [i_s, p]$. Add $[i_s, p]$ to $\mathcal{Q}$.

iii. Let j be the first point in $[p+1, i_t]$ with $\boldsymbol{s}_{j-1}[k] = -1$ with $\boldsymbol{u}_{j-1}[k] > -B$ and $\boldsymbol{u}_j[k] < \boldsymbol{u}_{j-1}[k]$. Add $[p+1, j-1]$ to $\mathcal{P}$.
iv. If $\boldsymbol{u}_r[k]$ is monotonic in $[j, i_t]$, add $[j, i_t]$ to $\mathcal{Q}$. Goto Step 3.
v. Initialize $q \leftarrow i_t + 1$.
vi. If $\boldsymbol{u}_{i_t}[k] = -B$, let q be smallest point in $[j, i_t]$ such that $\boldsymbol{u}_r[k] = -B \, \forall r \in [q, i_t]$. Add $[q, i_t]$ to $\mathcal{Q}$.
vii. Let h be the last time point in $[j, q-1]$ such that $\boldsymbol{s}_{h-1}[k] = 1$ with $\boldsymbol{u}_h[k] > \boldsymbol{u}_{h-1}[k]$. Add $[j, h-1]$ to $\mathcal{P}$.
viii. If $h < q-1$, add $[h, q-1]$ to $\mathcal{P}$.
ix. Goto Step 3.

(i) If there exists one coordinate $k \in [d]$ such that $\boldsymbol{s}_{i_s-1}[k] \in [0, 1]$ and $\boldsymbol{s}_{i_t}[k] \in [-1, -1/4]$ and $\boldsymbol{\gamma}_j^- = 0 \, \forall j \in [i_s, i_t]$, then remove $[i_s, i_t]$ from $\mathcal{Q}$ and add it to $\mathcal{P}$. Goto Step 3.

(j) If there exists one coordinate $k \in [d]$ such that $\boldsymbol{s}_{i_s-1}[k] \in [1/4, 1]$ and $\boldsymbol{s}_{i_t}[k] \in [-1/4, 0]$ and $\boldsymbol{\gamma}_j^- = 0 \, \forall j \in [i_s, i_t]$, then remove $[i_s, i_t]$ from $\mathcal{Q}$ and add it to $\mathcal{P}$. Goto Step 3.

(k) If there exists one coordinate $k \in [d]$ such that $\boldsymbol{u}_k$ is non-monotonic in $[i_s, i_t]$ and $\boldsymbol{s}_{i_s-1}[k], \boldsymbol{s}_{i_t}[k] \in [-1, 1/4]$ and $\boldsymbol{\gamma}_j^- = 0 \, \forall j \in [i_s, i_t]$ and there exists a coordinate $j \in [i_s, i_t]$ such that $\boldsymbol{s}_{j-1}[k] = 1$ and $\boldsymbol{u}_{j-1}[k] < B$ then:
i. Initialize $p \leftarrow i_s$. Remove $[i_s, i_t]$ from $\mathcal{Q}$.
ii. If $\boldsymbol{u}_{i_s}[k] = B$, then let p be the largest point in $[i_s, i_t]$ such that $\boldsymbol{u}_t[k] = B \, \forall t \in [i_s, p]$. Add $[i_s, p]$ to $\mathcal{Q}$.
iii. Let j be the first point in $[p+1, i_t]$ such that $\boldsymbol{s}_{j-1}[k] = 1$ with $\boldsymbol{u}_{j-1}[k] < \boldsymbol{u}_j[k]$. Add $[p+1, j-1]$ to $\mathcal{P}$. Add $[j, i_q]$ to $\mathcal{Q}$. Goto Step 3.

(l) If there exists one coordinate $k \in [d]$ such that $\boldsymbol{u}_k$ is non-monotonic in $[i_s, i_t]$ and $\boldsymbol{s}_{i_s-1}[k], \boldsymbol{s}_{i_t}[k] \in [-1/4, 1]$ and $\boldsymbol{\gamma}_j^- = 0 \, \forall j \in [i_s, i_t]$ and there exists a $j \in [i_s, i_t]$ such that $\boldsymbol{u}_j[k] - \boldsymbol{u}_{j-1}[k] = -1$ and $u_j[k] < B$ then:
i. Initialize $z_1 \leftarrow i_s, z_2 \leftarrow i_t$. Remove $[i_s, i_t]$ from $\mathcal{Q}$.
ii. If $\boldsymbol{u}_{i_s}[k] = B$, then let p_1 be the last point in $[i_s, i_t]$ such that $\boldsymbol{u}_t[k] = B \, \forall t \in [i_s, p_1]$. Set $z_1 \leftarrow p_1 + 1$. Add $[i_s, p]$ to $\mathcal{Q}$.
iii. If $\boldsymbol{u}_{i_t}[k] = B$, then let p_2 be the smallest point in $[i_s, i_t]$ such that $\boldsymbol{u}_t[k] = B \, \forall t \in [p_2, i_t]$. Set $z_2 \leftarrow p_2 - 1$. Add $[p_2, i_t]$ to $\mathcal{Q}$.
iv. Let j be the last point in $[z_1, z_2]$ such that $\boldsymbol{s}_{j-1}[k] = -1$ and $\boldsymbol{u}_j[k] < B$ with $\boldsymbol{u}_{j-1}[k] > \boldsymbol{u}_j[k]$. Add $[z_1, j-1]$ and $[j, z_2]$ to $\mathcal{P}$. Goto Step 3.

(m) If there exists one coordinate $k \in [d]$ such that $\boldsymbol{u}_k$ is non-monotonic in $[i_s, i_t]$ and $\boldsymbol{s}_{i_s-1}[k] \in [-1, 1/4]$ and $\boldsymbol{s}_{i_t}[k] \in [-1/4, 1]$ and $\boldsymbol{\gamma}_j^- = 0 \, \forall j \in [i_s, i_t]$ then:
i. Initialize $p \leftarrow i_s - 1$. Remove $[i_s, i_t]$ from $\mathcal{Q}$.

ii. If $\boldsymbol{u}_{i_s}[k] = B$, then let p be the last time point such that $\boldsymbol{u}_t[k] = B \, \forall t \in [i_s, p]$. Add $[i_s, p]$ to $\mathcal{Q}$.

iii. Let j be the first point in $[p+1, i_t]$ such that $\boldsymbol{s}_{j-1}[k] = 1$ with $\boldsymbol{u}_{j-1}[k] < B$ with $\boldsymbol{u}_{j-1}[k] < \boldsymbol{u}_j[k]$. Add $[p+1, j-1]$ to $\mathcal{P}$.

iv. If $\boldsymbol{u}_r[k]$ is monotonic in $[j, i_t]$, add $[j, i_t]$ to $\mathcal{Q}$. Goto Step 3.

v. Initialize $q \leftarrow i_t + 1$.

vi. If $\boldsymbol{u}_{i_t}[k] = B$, let q be smallest point in $[j, i_t]$ such that $\boldsymbol{u}_r[k] = B \, \forall r \in [q, i_t]$. Add $[q, i_t]$ to $\mathcal{Q}$.

vii. Let h be the last time point in $[j, q-1]$ such that $\boldsymbol{s}_{h-1}[k] = -1$ with $\boldsymbol{u}_{h-1}[k] > \boldsymbol{u}_h[k]$. Add $[j, h-1]$ to $\mathcal{P}$.

viii. If $h < q-1$, add $[h, q-1]$ to $\mathcal{P}$.

ix. Goto Step 3.

Step 4 Return $\mathcal{P}$.

Appendix E

Supplementary Materials for Chapter 6

E.1 Proofs for Section 6.2

We start by characterizing the offline optimal. Define the sign function as $\text{sign}(x) = 1$ if $x > 0$; -1 if $x < 0$; and some $v \in [-1, 1]$ if $x = 0$. We start by presenting a sequence of useful lemmas.

Lemma 52. *(**characterization of offline optimal**) Consider the following convex optimization problem (where $\tilde{z}_1, ..., \tilde{z}_{n-1}$ are introduced as dummy variables)*

$$
\begin{aligned}
\min_{\tilde{u}_1, \ldots, \tilde{u}_n, \tilde{z}_1, \ldots, \tilde{z}_{n-1}} \quad & \frac{1}{2} \sum_{t=1}^{n} (y_t - \tilde{u}_t)^2 && (6.4a) \\
\text{s.t.} \quad & \tilde{z}_t = \tilde{u}_{t+1} - \tilde{u}_t \; \forall t \in [n-1], && (6.4b) \\
& \sum_{t=1}^{n-1} |\tilde{z}_t| \leq C_n, && (6.4c) \\
& -B \leq \tilde{u}_t \; \forall t \in [n], && (6.4d) \\
& \tilde{u}_t \leq B \; \forall t \in [n], && (6.4e)
\end{aligned}
$$

Let $u_1, \ldots, u_n, z_1, \ldots, z_{n-1}$ be the optimal primal variables and let $\lambda \geq 0$ be the optimal dual variable corresponding to the constraint (6.4c)*. Further, let $\gamma_t^- \geq 0, \gamma_t^+ \geq 0$ be the optimal dual variables that correspond to constraints* (6.4d) *and* (6.4e) *respectively for all $t \in [n]$. By the KKT conditions, we have*

- ***stationarity:*** $u_t - y_t = \lambda (s_t - s_{t-1}) + \gamma_t^- - \gamma_t^+$*, where $s_t \in \partial |z_t|$ (a subgradient). Specifically, $s_t = \text{sign}(u_{t+1} - u_t)$ if $|u_{t+1} - u_t| > 0$ and s_t is some value in $[-1, 1]$ otherwise. For convenience of notations later, we also define $s_n = s_0 = 0$.*

- ***complementary slackness:*** *(a)* $\lambda\left(\sum_{t=2}^{n} |u_t - u_{t-1}| - C_n\right) = 0$; *(b)* $\gamma_t^-(u_t + B) = 0$ *and* $\gamma_t^+(u_t - B) = 0$ *for all* $t \in [n]$

Proof. We can form the Lagrangian of the optimization problem as:

$$\mathcal{L}(\tilde{u}_{1:n}, \tilde{z}_{1:n-1}, \tilde{\boldsymbol{v}}, \tilde{\lambda}, \tilde{\gamma}_{1:n}^+, \tilde{\gamma}_{1:n}^-) = \frac{1}{2}\sum_{t=1}^{n}(y_t - \tilde{u}_t)^2 + \tilde{\lambda}\left(\sum_{t=1}^{n-1}|\tilde{z}_t| - C_n\right) + \sum_{t=1}^{n-1}\tilde{v}_t(\tilde{u}_{t+1} - \tilde{u}_t - \tilde{z}_t) \tag{E.1}$$

$$+ \sum_{t=1}^{n} \tilde{\gamma}_t^-(-B - \tilde{u}_t) + \tilde{\gamma}_t^+(\tilde{u}_t - B), \tag{E.2}$$

for dual variables $\tilde{\lambda} > 0$, $\tilde{v}_{1:n}$ unconstrained, $\tilde{\gamma}_{1:n}^- \geq 0$ and $\tilde{\gamma}_{1:n}^+ \geq 0$. Let $(u_{1:n}, z_{1:n}, v_{1:n}, \lambda, \gamma_{1:n}^-, \gamma_{1:n}^+)$ be the optimal primal and dual variables. By stationarity conditions (via the derivative wrt u_t), we have:

$$u_t - y_t + v_{t-1} - v_t - \gamma_t^- + \gamma_t^+ = 0, \tag{E.3}$$

where we take $v_0 = v_n = 0$. Stationarity conditions via derivative wrt z_t yields

$$v_t = \lambda s_t. \tag{E.4}$$

Combining the above two equations and the complementary slackness rules yields the lemma. □

Example 182. We describe the example used to create Fig.6.1. We adopt the notations of Lemma 52.

- $G = 4$ and $B = 2$.
- For each $k \in [0, \frac{n^{1/4}}{2} - 1]$, $u_j = B - \frac{1}{2n^{3/4}}$ for all $j \in [2kn^{3/4} + 1, (2k+1)n^{3/4}]$.
- For each $k \in [0, \frac{n^{1/4}}{2} - 1]$, $u_j = B$ for all $j \in [(2k+1)n^{3/4} + 1, (2k+2)n^{3/4}]$.
- $y_1 = y_{n^{3/4}} = B - \frac{1}{2n^{3/4}} - \frac{n^{3/4}-2}{n}$. $y_j = B - \frac{1}{2n^{3/4}} - (1 - 2/n)$ for all $j \in [2, n^{3/4} - 1]$.
- For each $k \in [1, \frac{n^{1/4}}{2} - 1]$, $y_{2kn^{3/4}+1} = y_{(2k+1)n^{3/4}} = B - \frac{1}{2n^{3/4}} - \frac{n^{3/4}-2}{n}$. $y_j = B - \frac{1}{2n^{3/4}} - (1 - 2/n)$ for all $j \in [2kn^{3/4} + 2, (2k+1)n^{3/4} - 1]$.
- For each $k \in [0, \frac{n^{1/4}}{2} - 1]$, $y_j = G$ for all $j \in [(2k+1)n^{3/4} + 1, (2k+2)n^{3/4}]$.
- $\gamma_j^- = 0$ for all $j \in [n]$.
- For each $k \in [0, \frac{n^{1/4}}{2} - 1]$, $\gamma_j^+ = 0$ for all $j \in [2kn^{3/4} + 1, (2k+1)n^{3/4}]$.

- For each $k \in [0, \frac{n^{1/4}}{2} - 2]$, $\gamma^+_{(2k+1)n^{3/4}+1} = \gamma^+_{(2k+2)n^{3/4}} = G - B - \frac{n^{3/4}-2}{n}$. $\gamma^+_j = G - B - 2(1-1/n)$ for all $j \in [(2k+1)n^{3/4}+2, (2k+2)n^{3/4}-1]$.
- $\gamma^+_{n-n^{3/4}+1} = \gamma^+_n = G - B - \frac{n^{3/4}-2}{n}$.
- $\lambda = n^{3/4} - 2$.
- $s_t = 1/n + (t-1)\frac{1-2/n}{n^{3/4}-2}$ for $1 \le t \le n^{3/4}-1$. $s_{n^{3/4}} = 1$.
- For each $k \in [0, \frac{n^{1/4}}{2} - 2]$, $s_t = 1 - 1/n + (t-1-(2k+1)n^{3/4})\frac{2/n-2}{n^{3/4}-2}$ for $(2k+1)n^{3/4}+1 \le t \le (2k+2)n^{3/4}-1$. $s_{(2k+2)n^{3/4}} = -1$.
- For each $k \in [1, \frac{n^{1/4}}{2} - 1]$, $s_t = -1 + 1/n + (t-1-2kn^{3/4})\frac{2-2/n}{n^{3/4}-2}$. $s_{(2k+1)n^{3/4}} = 1$.
- $s_t = 1 - 1/n + (t-1-n+n^{3/4})\frac{2/n-1}{n^{3/4}-1}$ for $n - n^{3/4} + 1 \le t \le n-1$. $s_n = 0$.

Terminology. We will refer to the optimal primal variables $u_1, \ldots, u_n$ in Lemma 52 as the *offline optimal solution* in this section. For two natural numbers $a < b$, we denote $[a, b] = \{a, a+1, \ldots, b\}$.

Definition 53.

- For a bin $[a, b] \subseteq \{2, \ldots, n-1\}$, the offline optimal solution is said to assume Structure 1 if $u_j = u_a \in (-B, B)$ for all $j \in [a, b]$ and $u_b > u_{b+1}$ and $u_a > u_{a-1}$.
- For a bin $[a, b] \subseteq \{2, \ldots, n-1\}$, the offline optimal solution is said to assume Structure 2 if $u_j = u_a \in (-B, B)$ for all $j \in [a, b]$ and $u_b < u_{b+1}$ and $u_a < u_{a-1}$.
- For a bin $[a, b]$, we define $\text{gap}_{\min}(\beta, [a, b]) := \min_{j \in [a,b]} |u_j - \beta|$ where $\beta \in \mathbb{R}$.

The following Lemma plays a central role in the analysis. Qualitatively, it captures a fundamental way in which the adversary is constrained.

Lemma 183. ***(λ-length lemma)*** *Suppose that the offline optimal solution sequence takes the form of Structure 1 or Structure 2 in an interval $[j, j+\ell-1]$ for some $\ell > 0$ and $j \in \{2, \ldots, n-1\}$. Then $\lambda \le \frac{(B+G)\ell}{2}$.*

Proof. We consider the case of Structure 2. Arguments are similar for case of Structure 1. Let the optimal sign assignments be written as $s_{j+k-1} = -1 + \epsilon_k$ where $\epsilon_k \in [0, 2]$ for all $k \in [\ell - 1]$. From the KKT conditions, we have

$$
\begin{aligned}
y_j &= u - \lambda\epsilon_1 \\
y_{j+1} &= u - \lambda(\epsilon_2 - \epsilon_1) \\
&\vdots \\
y_{j+\ell-2} &= u - \lambda(\epsilon_{\ell-1} - \epsilon_{\ell-2}) \\
y_{j+\ell-1} &= u - \lambda(2 - \epsilon_{\ell-1})
\end{aligned}
$$

Consider a vector $\boldsymbol{z} = [\epsilon_1, \epsilon_2 - \epsilon_1, \ldots, 2 - \epsilon_{\ell-1}]^T$. Note that the condition $\|\boldsymbol{z}\|_\infty > 0$ is always satisfied. Otherwise we must have $2 = \epsilon_{\ell-1} = \ldots = \epsilon_1$. But $\epsilon_1 = 2$ makes $\|\boldsymbol{z}\|_\infty > 0$ yielding a contradiction.

Let k^* be such that $|\boldsymbol{z}[k^*]| = \|\boldsymbol{z}\|_\infty$. Since $\lambda \geq 0$, we can write $\lambda = \frac{|y_{j+k^*-1} - u|}{\|\boldsymbol{z}\|_\infty}$. Since $|y_{j+k^*-1} - u|$ is bounded, a lower bound on $\|\boldsymbol{z}\|_\infty$ will yield an upper bound on λ. To this end, we consider the following optimization problem:

$$\min_{t, \epsilon_1, \ldots, \epsilon_{\ell-1}} \quad t \tag{E.5a}$$
$$\text{s.t.} \quad 0 \leq \epsilon_i \leq 2 \, \forall i \in [\ell - 1], \tag{E.5b}$$
$$\epsilon_1 \leq t, \tag{E.5c}$$
$$|\epsilon_{i+1} - \epsilon_i| \leq t \, \forall i \in [\ell - 2], \tag{E.5d}$$
$$2 - \epsilon_{\ell-1} \leq t \tag{E.5e}$$

We can form the Lagrangian as:

$$\mathcal{L}(t, \epsilon_{1:\ell-1}, a_{1:\ell-1}, b_{1:\ell-1}, c_{1:\ell-2}, d_{1:\ell-2}, e_1, e_{\ell-1}) = t - \sum_{i=1}^{\ell-1} a_i \epsilon_i + \sum_{i=1}^{\ell-1} b_i(\epsilon_i - 2) \tag{E.6}$$
$$+ \sum_{i=1}^{\ell-2} c_i(-t - \epsilon_{i+1} + \epsilon_i) + \sum_{i=1}^{\ell-2} d_i(\epsilon_{i+1} - \epsilon_i - t) \tag{E.7}$$
$$+ e_1(\epsilon_1 - t) + e_2(2 - \epsilon_{\ell-1} - t) \tag{E.8}$$

Stationarity conditions are:

$$\frac{\partial \mathcal{L}}{\partial t} = 0 \implies 1 + \sum_{i=1}^{\ell-2} -c_i - d_i - e_1 - e_2 = 0 \tag{E.9}$$
$$\frac{\partial \mathcal{L}}{\partial \epsilon_1} = 0 \implies -a_1 + b_1 + c_1 - d_1 + e_1 = 0 \tag{E.10}$$
$$\frac{\partial \mathcal{L}}{\partial \epsilon_{\ell-1}} = 0 \implies -a_{\ell-1} + b_{\ell-1} - c_{\ell-2} + d_{\ell-2} - e_2 = 0 \tag{E.11}$$
$$\frac{\partial \mathcal{L}}{\partial \epsilon_i} = 0 \implies -a_i + b_i - c_{i-1} + c_i + d_{i-1} - d_i = 0, \text{ where } i \in \{2, \ldots, \ell - 2\} \tag{E.12}$$

Complementary slackness conditions are:

$$a_i \epsilon_i = 0,\ i \in [\ell - 1] \tag{E.13}$$
$$b_i(\epsilon_i - 2) = 0,\ i \in [\ell - 1] \tag{E.14}$$
$$c_i(-t - \epsilon_{i+1} + \epsilon_i) = 0,\ i \in [\ell - 2] \tag{E.15}$$
$$d_i(\epsilon_{i+1} - \epsilon_i - t) = 0,\ i \in [\ell - 2] \tag{E.16}$$
$$e_1(\epsilon_1 - t) = 0 \tag{E.17}$$
$$e_2(2 - \epsilon_{\ell-1} - t) = 0 \tag{E.18}$$

Dual feasibility conditions are $a_i \geq 0,\ b_i \geq 0$ for $i \in [\ell - 1]$ and $c_i \geq 0,\ d_i \geq 0$ for $i \in [\ell - 2]$ and $e_1 \geq 0,\ e_2 \geq 0$.

Primal feasibility conditions are given by the constraint set of the optimization problem.

Now we form a guess for optimal primal and dual variables as $t = 2/\ell$ and $\epsilon_i = 2i/\ell$ for $i \in [\ell - 1]$ and $a_i = b_i = 0$ for $i \in [\ell - 1]$ and $c_i = 0$ for $i \in [\ell - 2]$ and $e_1 = e_2 = d_1 = \ldots = d_{\ell-2} = 1/\ell$. All the KKT conditions can be readily verified for this solution guess.

Recall that, earlier we defined $\boldsymbol{z} = [\epsilon_1, \epsilon_2 - \epsilon_1, \ldots, 2 - \epsilon_{\ell-1}]^T$ and $\lambda = \frac{|y_{j+k^*-1} - u|}{\|\boldsymbol{z}\|_\infty}$ where k^* is such that $|\boldsymbol{z}[k^*]| = \|\boldsymbol{z}\|_\infty$. By the previous optimization problem we deduce that $\|\boldsymbol{z}\|_\infty \geq 2/\ell$. Since $|y_{j+k^*-1} - u| \leq B + G$, we conclude that $\lambda \leq (B + G)\ell/2$

□

Next, we exhibit a useful partitioning scheme of the interval $[n]$.

Lemma 184. *([65])(**key partition**) Initialize $\mathcal{P} \leftarrow \Phi$. Starting from time 1, spawn a new bin $[i_s, i_t]$ whenever $\sum_{j=i_s+1}^{i_t+1} |u_j - u_{j-1}| > B/\sqrt{n_i}$, where $n_i = i_t - i_s + 2$. Add the spawned bin $[i_s, i_t]$ to $\mathcal{P}$.*

Let $M := |\mathcal{P}|$. We have $M = O\left(1 \vee n^{1/3} C_n^{2/3} B^{-2/3}\right)$.

Notations. For bin $[i_s, i_t] \in \mathcal{P}$ we define: $n_i = i_t - i_s + 1$, $\bar{u}_i = \frac{1}{n_i}\sum_{j=i_s}^{i_t} u_j$, $\bar{y}_i = \frac{1}{n_i}\sum_{j=i_s}^{i_t} y_j$, $\Gamma_i^+ = \sum_{j=i_s}^{i_t} \gamma_j^+$, $\Gamma_i^- = \sum_{j=i_s}^{i_t} \gamma_j^-$, $\Delta s_i = s_{i_t} - s_{i_s-1}$, $C_i = \sum_{j=i_s+1}^{i_t} |u_j - u_{j-1}|$.

For any general bin $[a, b]$ define the quantities $n_{a\to b}, \bar{u}_{a\to b}, \bar{y}_{a\to b}, \Gamma_{a\to b}^+, \Gamma_{a\to b}^-, \Delta s_{a\to b}, C_{a\to b}$ analogously as above.

Next we calculate the static regret guarantee of the FLH-ONS strategy.

Lemma 185. *([77], [23]) Consider a bin $[a, b] \subseteq [n]$ and a point $w \in [-B, B]$. Under the setting of Theorem 50 we have*

$$\sum_{t=a}^{b} (y_t - x_t)^2 - (y_t - w)^2 \leq 10(B + G)^2 \log n \tag{E.19}$$
$$= \tilde{O}(1), \tag{E.20}$$

where x_t are the predictions of FLH-OGD.

Proof. The losses $(y_t - x)^2$ are strongly convex with parameter 2. Further the gradients are bounded by $2(G+B)$. Hence by Theorem 1 in [77] we have the static regret guarantee of OGD being $4(G+B)^2 \cdot (2\log n)/4 = 2(G+B)^2 \log n$.

The losses $(y_t - x)^2$ are $1/(2(G+B)^2)$ exp-concave. So by applying Theorem 3.2 in [23] we have the regret of FLH against any base experts bounded as $8(G+B)^2 \log n$.

Adding these regret bounds yields the lemma.

□

Lemma 186. ***(low λ regime)*** *If the optimal dual variable $\lambda = O\left(\frac{n^{1/3}}{C_n^{1/3}}\right)$, we have the regret of FLH-OGD strategy bounded as*

$$\sum_{t=1}^{n} (y_t - x_t)^2 - (y_t - u_t)^2 = \tilde{O}(n^{1/3} C_n^{2/3} \vee 1), \tag{E.21}$$

where x_t is the prediction of FLH-OGD at time t.

Proof. Throughout this proof, the bins $[i_s, i_t]$ we consider belong to the partition $\mathcal{P}$.

Case 1: When the offline optimal solution touches the boundary B within a bin $[i_s, i_t]$. We use a three term regret decomposition as follows.

$$\underbrace{\sum_{j=i_s}^{i_t} (y_j - x_j)^2 - (y_j - B)^2}_{T_{1,i}} + \underbrace{\sum_{j=i_s}^{i_t} (y_j - B)^2 - (y_j - \bar{u}_i)^2}_{T_{2,i}} + \underbrace{\sum_{j=i_s}^{i_t} (y_j - \bar{u}_i)^2 - (y_j - u_j)^2}_{T_{3,i}} \tag{E.22}$$

Now $T_{1,i} = O(\log n)$ by strong adaptivity of FLH. Observe that due to complementary slackness, $\gamma_j^- = 0$ uniformly within the bin since the TV within the bin is at-most $B/\sqrt{n_i} < 2B$ and hence the solution never touches $-B$ boundary within this bin. By using the KKT conditions, we have $y_j = u_j - \lambda(s_j - s_{j-1}) + \gamma_j^+$. So

$$\begin{aligned}
T_{2,i} &= \sum_{j=i_s}^{i_t} (\bar{u}_i - B)^2 + 2(y_j - \bar{u}_i)(\bar{u}_i - B) && \text{(E.23)}\\
&= n_i(\bar{u}_i - B)^2 + 2n_i(\bar{y}_i - \bar{u}_i)(\bar{u}_i - B) && \text{(E.24)}\\
&\leq_{(a)} B^2 + 2(\bar{u}_i - B)(\Gamma_i^+ - \lambda \Delta s_i) && \text{(E.25)}\\
&\leq_{(b)} B^2 + 4\lambda C_i + 2\Gamma_i^+(\bar{u}_i - B) && \text{(E.26)}
\end{aligned}$$

where in line (a) we used KKT conditions and $|\bar{u}_i - B| \leq B/\sqrt{n_i}$ due to the TV constraint within bin and in line (b) we used: (i) $|\bar{u}_i - B| \leq C_i$ as the optimal solution assumes the value B at some time point in $[i_s, i_t]$ (ii) $|\Delta s_i| \leq 2$.

We have

$$T_{3,i} = \sum_{j=i_s}^{i_t} (u_j - \bar{u}_i)^2 + 2(y_j - u_j)(u_j - \bar{u}_i) \tag{E.27}$$

$$\leq n_i C_i^2 + 2\sum_{j=i_s}^{i_t} (-\lambda(s_j - s_{j-1}) + \gamma_j^+)(u_j - \bar{u}_i) \tag{E.28}$$

$$=_{(a)} n_i C_i^2 + 2\lambda(s_{i_s-1}(u_{i_s} - \bar{u}_i) - s_{i_t}(u_{i_t} - \bar{u}_i)) + 2\lambda C_i + 2\sum_{j=i_s}^{i_t} \gamma_j^+(u_j - \bar{u}_i) \tag{E.29}$$

$$=_{(b)} n_i C_i^2 + 2\lambda(s_{i_s-1}(u_{i_s} - \bar{u}_i) - s_{i_t}(u_{i_t} - \bar{u}_i)) + 2\lambda C_i + 2\Gamma_i^+(B - \bar{u}_i) \tag{E.30}$$

$$\leq_{(c)} B^2 + 6\lambda C_i + 2\Gamma_i^+(B - \bar{u}_i), \tag{E.31}$$

where line (a) is obtained by a rearrangement of the sum and line (b) is obtained by the complementary slackness condition which states that $\gamma_j^+ = 0$ if $u_j < B$. Line (c) is obtained by $|u_j - \bar{u}_i| \leq C_i$ for any $j \in [i_s, i_t]$ and by applying triangle inequality.

So overall we can bound the regret within this bin by adding Eq.(E.26) and (E.31) with $T_{1,i} = O(\log n)$ as

$$T_{1,i} + T_{2,i} + T_{3,i} \leq O(\log n) + 2B^2 + 10\lambda C_i. \tag{E.32}$$

Case 2: When the offline optimal solution touches boundary $-B$ within a bin $[i_s, i_t]$. This case can be treated similar to Case 1.

Case 3: When the offline optimal solution doesn't touch either boundaries within a bin $[i_s, i_t]$. Here we use a two term regret decomposition as

$$\underbrace{\sum_{j=i_s}^{i_t} (y_j - x_j)^2 - (y_j - \bar{u}_i)^2}_{T_{1,i}} + \underbrace{\sum_{j=i_s}^{i_t} (y_j - \bar{u}_i)^2 - (y_j - u_j)^2}_{T_{2,i}}. \tag{E.33}$$

By following the analysis used in obtaining the bound of Eq.(E.31) (where we use $\gamma_j^- = \gamma_j^+ = 0$ due to complementary slackness), we obtain

$$T_{1,i} + T_{2,i} \leq O(\log n) + B^2 + 6\lambda C_i \tag{E.34}$$

By summing up the regret bounds which assumes the form in Eq.(E.32) (for Case 1 and 2) or Eq.(E.34) (for Case 3) across all bins in the partition $\mathcal{P}$, we obtain the overall regret as

$$\sum_{t=1}^{n}(y_t - x_t)^2 - (y_t - u_t)^2 \leq O(|\mathcal{P}| \log nB^2) + 2B^2|\mathcal{P}| + 10\lambda C_n \tag{E.35}$$

$$= \tilde{O}(n^{1/3}C_n^{2/3} \vee 1), \tag{E.36}$$

where in the last line we used the fact that $|\mathcal{P}| = O(n^{1/3}C_n^{2/3} \vee 1)$ and $\lambda = O((n/C_n)^{1/3})$ by the premise of the lemma. □

Lemma 187. ***(monotonic sequence)*** *Consider a bin $[i_s, i_t] \in \mathcal{P}$ such that the offline optimal solution is monotonic within this bin. Then the regret of FLH-OGD strategy within this bin is at-most $31(B+G)^2 \log n = O(\log n)$.*

Proof. When the optimal sequence is monotonic within a bin $[i_s, i_t] \in \mathcal{P}$, it is always possible to form *at-most* 3 bins: $[i_s, r_1]$, $[r_1+1, r_2]$, $[r_2+1, i_t]$ such that the offline optimal solution is constant within bins $[i_s, r_1]$ and $[r_2+1, i_t]$ alongside the condition that the bin $[r_1+1, r_2]$ satisfies one of the following properties: a) $s_{r_1} = s_{r_2} = 1$ and the offline optimal solution is non-decreasing within bin $[r_1+1, r_2]$ or b) $s_{r_1} = s_{r_2} = -1$ and the offline optimal solution is non-increasing within bin $[r_1+1, r_2]$. (see for eg. Fig.E.1).

Due to Lemma 185, the regret within bins $[i_s, r_1]$ and $[r_2+1, i_t]$ is at-most $10(B+G)^2 \log n$ each. Note that this three sub-bin refinement can make sure that the offline optimal solution doesn't touch the boundaries $\pm B$ within the bin $[r_1+1, r_2]$. We bound the regret within bin $[r_1+1, r_2]$ via a two term regret decomposition as follows.

$$\underbrace{\sum_{j=r1+1}^{r_2} (y_j - x_j)^2 - (y_j - \bar{u}_{r1+1\to r_2})^2}_{T_1} + \underbrace{\sum_{j=r1+1}^{r_2} (y_j - \bar{u}_{r1+1\to r_2})^2 - (y_j - u_j)^2}_{T_2}. \tag{E.37}$$

We have $T_1 \leq 10(B+G)^2 \log n$. Further due to KKT conditions we have,

$$T_2 = \sum_{j=r1+1}^{r_2} (u_j - \bar{u}_{r1+1\to r_2})(2y_j - u_j - \bar{u}_{r1+1\to r_2}) \tag{E.38}$$

$$= \sum_{j=r1+1}^{r_2} (u_j - \bar{u}_{r1+1\to r_2})(2y_j - 2u_j + u_j - \bar{u}_{r1+1\to r_2}) \tag{E.39}$$

$$= \sum_{j=r1+1}^{r_2} (u_j - \bar{u}_{r1+1\to r_2})^2 + 2\lambda(u_j - \bar{u}_{r1+1\to r_2})(s_{j-1} - s_j) \tag{E.40}$$

$$\leq n_i C_i^2 + \sum_{j=r1+1}^{r_2} 2\lambda(u_j - \bar{u}_{r1+1\to r_2})(s_{j-1} - s_j), \tag{E.41}$$

where in the last line we used $|u_j - \bar{u}_{r1+1\to r_2}| \le C_i$. We also have $n_i C_i^2 \le B^2$ by the construction in Lemma 184.

By expanding the second term followed by a regrouping of terms in the summation, we can write

$$\sum_{j=r1+1}^{r_2} 2\lambda(u_j - \bar{u}_{r1+1\to r_2})(s_{j-1} - s_j) = 2\lambda\left(s_{r_1}(u_{r_1+1} - \bar{u}_{r1+1\to r_2}) - s_{r_2}(u_{r_2} - \bar{u}_{r1+1\to r_2})\right) \tag{E.42}$$

$$+ 2\lambda \sum_{j=r_1+2}^{r_2} |u_j - u_{j-1}| \tag{E.43}$$

$$= 2\lambda C_{r_1+1\to r_2} \tag{E.44}$$

$$+ 2\lambda\left(s_{r_1}(u_{r_1+1} - \bar{u}_{r1+1\to r_2}) - s_{r_2}(u_{r_2} - \bar{u}_{r1+1\to r_2})\right). \tag{E.45}$$

Since $s_{r_1} = s_{r_2} = 1$ if the offline optimal is non-decreasing in $[r_1+1, r_2]$ or $s_{r_1} = s_{r_2} = -1$ if the offline optimal is non-increasing in $[r_1+1, r_2]$, we have $s_{r_1} u_{r_1+1} - s_{r_2} u_{r_2} = -|u_{r_1+1} - u_{r_2}| = -C_{r_1+1\to r_2}$. Hence we see that the second term exactly cancels with the first term in Eq.(E.45).

Thus overall we have shown that the total regret in $[i_s, i_t]$ is at-most $31(B+G)^2 \log n$. □

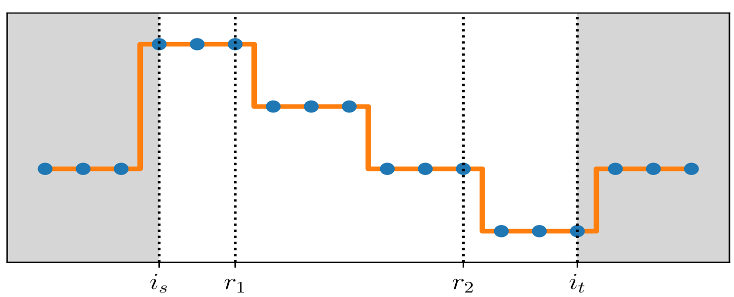


Figure E.1: *An example of a configuration referred in the proof of Lemma 187. Here $s_{r_1} = s_{r_2} = -1$ and the sequence is non-increasing within $[r_1+1, r_2]$.*

Lemma 188. *Suppose there exists an interval $[a, b]$ (which may not belong to $\mathcal{P}$) with length ℓ such that the optimal sequence takes the form of Structure 1 or Structure 2 within $[a, b]$. Assume that $\bar{y}_{a\to b} \in [-B, B]$. Then the regret of FLH-OGD within the bin $[a, b]$ at-most $10(B+G)^2 \log n - \frac{4\lambda^2}{\ell}$.*

Proof. We use a two term regret decomposition as follows:

$$\underbrace{\sum_{j=a}^{b}(y_j - x_j)^2 - (y_j - \bar{y}_{a\to b})^2}_{T_1} + \underbrace{\sum_{j=a}^{b}(y_j - \bar{y}_{a\to b})^2 - (y_j - u_j)^2}_{T_2}. \tag{E.46}$$

By the Definition 53 of Structure 1 and 2, the offline optimal solution is constant within bin $[a, b]$. We denote $u_j = u$ for all $j \in [a, b]$. Further $|\Delta s_{a\to b} = 2|$. We have,

$$T_2 = -\ell(\bar{y}_{a\to b} - u)^2 - 2\sum_{j=a}^{b}(y_j - \bar{y}_{a\to b})(\bar{y}_{a\to b} - u) \tag{E.47}$$

$$= -\ell(\bar{y}_{a\to b} - u)^2 \tag{E.48}$$

$$=_{(a)} -\frac{-\lambda^2(\Delta s_{a\to b})^2}{\ell} \tag{E.49}$$

$$= -\frac{4\lambda^2}{\ell}, \tag{E.50}$$

where line (a) is obtained by the KKT conditions $y_j = u - \lambda(s_j - s_{j-1})$ for all $j \in [a, b]$ and hence $\bar{y}_{a\to b} = u - \frac{\lambda \Delta s_{a\to b}}{\ell}$.

Due to Lemma 185, we have $T_1 \leq 10(B + G)^2 \log n$. Combining both bounds yields the lemma. □

Lemma 189. *Consider a bin $[a, b]$ with length ℓ.*

Case 1: *When offline optimal takes the form of Structure 1 within this bin and $\bar{y}_{a\to b} \geq B$, then*

$$\sum_{j=a}^{b}(y_j - x_j)^2 - (y_j - u_j)^2 \leq 10(B + G)^2 \log n - \ell(B - u_a)^2, \tag{E.51}$$

and

Case 2: *When offline optimal takes the form of Structure 2 within this bin and $\bar{y}_{a\to b} \leq -B$, then*

$$\sum_{j=a}^{b}(y_j - x_j)^2 - (y_j - u_j)^2 \leq 10(B + G)^2 \log n - \ell(B + u_a)^2, \tag{E.52}$$

where x_j are the predictions of the FLH-OGD algorithm.

Proof. We consider Case 2. Arguments for Case 1 are similar. We employ a two term regret decomposition as follows.

$$\underbrace{\sum_{j=a}^{b}(y_j - x_j)^2 - (y_j + B)^2}_{T_1} + \underbrace{\sum_{j=a}^{b}(y_j + B)^2 - (y_j - u_j)^2}_{T_2}. \tag{E.53}$$

By Definition 53, the offline optimal solution is constant within bin $[a, b]$. So we have $u_j = u_a$ for all $j \in [a, b]$. From the KKT conditions, we have

$$T_2 = \sum_{j=a}^{b}(u_a + B)^2 + 2(y_j - u_a)(u_a + B) \tag{E.54}$$

$$= \ell(u_a + B)^2 - 2\lambda\Delta s_{a\to b}(u_a + B) \tag{E.55}$$

$$= \ell(u_a + B)^2 - 4\lambda(u_a + B), \tag{E.56}$$

where in the last line we used $\Delta s_{a\to b} = 2$ for Structure 2. From the premise of the lemma for Case 2, we have $\bar{y}_{a\to b} \le -B$. Since $\bar{y}_{a\to b} = u_a - 2\lambda/\ell$, we must have

$$\bar{y}_{a\to b} \le -B \implies \lambda \ge \frac{\ell}{2}(u_a + B). \tag{E.57}$$

Plugging this lower bound to Eq.(E.56) and noting that $u_a + B \ge 0$, we get

$$T_2 \le -\ell(u_a + B)^2. \tag{E.58}$$

By Lemma 185, we have $T_1 \le 10(B + G)^2 \log n$. Now summing T_1 and T_2 results in the lemma.

□

Lemma 190. ***(large margin bins)*** *Assume that $\lambda \ge \phi\frac{n^{1/3}}{C_n^{1/3}}$ for some constant ϕ that do not depend on n and C_n. Consider a bin $[i_s, i_t] \in \mathcal{P}$ within which the offline optimal solution takes the form of Structure 1 or Structure 2 (or both) for some appropriate sub-intervals of $[i_s, i_t]$. Let $\mu_{th} = \sqrt{\frac{36(B+G)^3C_n^{1/3}\log n}{\phi n^{1/3}}}$. Then $gap_{min}(-B, [i_s, i_t]) \vee gap_{min}(B, [i_s, i_t]) \ge \mu_{th}$ whenever $C_n \le \left(\frac{B^2\phi}{144(B+G)^3\log n}\right)^3 n = \tilde{O}(n)$.*

Proof. Suppose $\text{gap}_{\min}(-B, [i_s, i_t]) < \mu_{\text{th}}$. Then the largest value of offline optimal attained within this bin $[i_s, i_t]$ is at-most $-B + \mu_{\text{th}} + B/\sqrt{n_i}$ (recall $n_i := i_t - i_s + 1$ and TV within this bin is at-most $B/\sqrt{n_i}$ by Lemma 184). So $\text{gap}_{\min}(B, [i_s, i_t]) \ge 2B - \mu_{\text{th}} - B/\sqrt{n_i}$. Our goal is to show that whenever C_n obeys the constraint stated in the lemma, we must have

$$2B - \mu_{\text{th}} - B/\sqrt{n_i} \ge \mu_{\text{th}}. \tag{E.59}$$

Let ℓ_i be the length of a sub-interval of $[i_s, i_t]$ where the offline optimal solution assumes the form of Structure 1 or Structure 2. Due to Lemma 183, we have

$$n_i \geq \ell_i \geq \frac{2\lambda}{(G+B)} \geq \frac{2\phi}{(G+B)} \frac{n^{1/3}}{C_n^{1/3}}, \tag{E.60}$$

where the last inequality follows due to the condition on λ assumed in the current lemma. So a sufficient condition for Eq.(E.59) to be true is

$$2B \geq 2\left(2\sqrt{\frac{36(G+B)^3 C_n^{1/3} \log n}{\phi n^{1/3}}} \vee B\sqrt{\frac{(G+B)C_n^{1/3}}{2\phi n^{1/3}}}\right). \tag{E.61}$$

Recall that by Assumption A1 in Section 6.2, we have $G \geq B \geq 1$ WLOG. So the above maximum will be attained by the first term and can be further simplified as

$$2B \geq 4\sqrt{\frac{36(G+B)^3 C_n^{1/3} \log n}{\phi n^{1/3}}}. \tag{E.62}$$

The above condition is always satisfied whenever $C_n \leq \left(\frac{B^2\phi}{144(B+G)^3 \log n}\right)^3 n$.

At this point, we have shown that $\text{gap}_{\min}(-B, [i_s, i_t]) < \mu_{\text{th}} \implies \text{gap}_{\min}(B, [i_s, i_t]) \geq \mu_{\text{th}}$ under the conditions of the lemma. Taking the contrapositive yields $\text{gap}_{\min}(B, [i_s, i_t]) < \mu_{\text{th}} \implies \text{gap}_{\min}(-B, [i_s, i_t]) \geq \mu_{\text{th}}$. □

Lemma 191. *(**high λ regime**) If the optimal dual variable $\lambda \geq \phi\frac{n^{1/3}}{C_n^{1/3}} = \Omega\left(\frac{n^{1/3}}{C_n^{1/3}}\right)$ for some constant $\phi > 0$ that doesn't depend on n and C_n, we have the regret of FLH-OGD strategy bounded as*

$$\sum_{t=1}^{n} (y_t - x_t)^2 - (y_t - u_t)^2 = \tilde{O}(n^{1/3} C_n^{2/3} \vee 1), \tag{E.63}$$

where x_t is the prediction of FLH-OGD at time t.

Proof. Throughout the proof, we consider only the regime where $C_n \leq \left(\frac{B^2\phi}{84(B+G)^3 \log n}\right)^3 n = \tilde{O}(n)$. In the alternate regime where $C_n = \tilde{\Omega}(n)$, the trivial regret bound of $\tilde{O}(n)$ is near minimax optimal.

Reminiscent to the road-map in Section 6.2.2, it is useful to define the following condition:

Condition (A): Let a bin $[a, b]$ be given such that $C_{a\to b} \leq B/\sqrt{b-a+1}$. It satisfies at-least one of the following criteria. (i) $\text{gap}_{\min}(B, [a, b]) \geq \text{gap}_{\min}(-B, [a, b])$ and the optimal solution takes the form of Structure 1 in at-least one sub-interval $[r, s] \subseteq [a, b]$;

or (ii) $\text{gap}_{\min}(-B,[a,b]) \geq \text{gap}_{\min}(B,[a,b])$ and the optimal solution takes the form of Structure 2 in at-least one sub-interval $[r,s] \subseteq [a,b]$.

Consider a bin $[i_s, i_t] \in \mathcal{P}$ that satisfies Condition (A). We refine $[i_s, i_t]$ into a partition that contains smaller sub-intervals as follows:

$$\mathcal{P}_i := \{[i_s, \underline{i}_1 - 1], [\underline{i}_1, \bar{i}_1], [\underline{i}'_1, \bar{i}'_1], \ldots, [\underline{i}_{m^{(i)}}, \bar{i}_{m^{(i)}}], [\underline{i}'_{m^{(i)}}, \bar{i}'_{m^{(i)}} := i_t]\}, \tag{E.64}$$

such that:

1. If $\text{gap}_{\min}(B,[i_s,i_t]) > \text{gap}_{\min}(-B,[i_s,i_t])$, then the offline optimal in the intervals $[\underline{i}_j, \bar{i}_j]$, $j \in [m^{(i)}]$ takes the form of Structure 1. Further, let k be the largest value in $[i_s, i_t]$ such that $u_{i_s:k}$ is constant. If $u_{i_s} > u_{i_s-1}$ and $u_k > u_{k+1}$, then we treat the first sub-interval in $\mathcal{P}_i$ as empty by putting $\underline{i}_1 = i_s$. Similarly let k be smallest value in $[i_s, i_t]$ such that $u_{k:i_t}$ is constant. If $u_{k-1} < u_k$ and $u_{i_t} > u_{i_t+1}$ then we treat the last sub-interval in $\mathcal{P}_i$ as empty by putting $\underline{i}'_{m^{(i)}} = i_t + 1$.

2. If $\text{gap}_{\min}(B,[i_s,i_t]) \leq \text{gap}_{\min}(-B,[i_s,i_t)$, then the offline optimal in the intervals $[\underline{i}_j, \bar{i}_j]$, $j \in [m^{(i)}]$ takes the form of Structure 2. Further, let k be the largest value in $[i_s, i_t]$ such that $u_{i_s:k}$ is constant. If $u_{i_s} < u_{i_s-1}$ and $u_k < u_{k+1}$, then we treat the first sub-interval in $\mathcal{P}_i$ as empty by putting $\underline{i}_1 = i_s$. Similarly let k be smallest value in $[i_s, i_t]$ such that $u_{k:i_t}$ is constant. If $u_{k-1} > u_k$ and $u_{i_t} < u_{i_t+1}$ then we treat the last sub-interval in $\mathcal{P}_i$ as empty by putting $\underline{i}'_{m^{(i)}} = i_t + 1$.

3. In all sub-intervals $[\underline{i}'_j, \bar{i}'_j]$, $j \in [m^{(i)}]$, the offline optimal sequence can be split into piece-wise monotonic sections with at-most 2 pieces.

An illustration of this refinement scheme is given in Fig.E.2.

Let there be $m_1^{(i)}$ bins among $\{[\underline{i}_1, \bar{i}_1], \ldots, [\underline{i}_{m^{(i)}}, \bar{i}_{m^{(i)}}]\}$ which satisfy the property in Lemma 188. Let their lengths be denoted by $\{\ell^{(1)}_{1^{(i)}}, \ldots, \{\ell^{(1)}_{m_1^{(i)}}\}$. These bins will be referred as *Type 1* bins henceforth.

Similarly let there be $m_2^{(i)}$ bins among $\{[\underline{i}_1, \bar{i}_1], \ldots, [\underline{i}_{m^{(i)}}, \bar{i}_{m^{(i)}}]\}$ which satisfy either Case 1 or Case 2 in Lemma 189. Let their lengths be denoted by $\{\ell^{(2)}_{1^{(i)}}, \ldots, \{\ell^{(2)}_{m_2^{(i)}}\}$. These bins will be referred as *Type 2* bins henceforth.

Each bin in Type 1 and Type 2 can be paired with one adjacent bin (if non-empty) in $\mathcal{P}_i$ where the optimal sequence displays a piece-wise monotonic behaviour with at-most 2 pieces. (For example the bin $[\underline{i}_1, \bar{i}_1]$ can be paired with $[\underline{i}'_1, \bar{i}'_1]$ where in the later the optimal sequence displays a piece-wise monotonic behaviour. See Fig.E.2 for example.) To see why this is true, consider the case $\text{gap}_{\min}(B,[i_s,i_t]) \leq \text{gap}_{\min}(-B,[i_s,i_t)$. By construction, the optimal solution must preclude the form of Structure 2 in the bin $[\underline{i}'_k, \bar{i}'_k]$ where $k \in [m^{(i)}]$. This means the offline optimal can either take a non-increasing form in $[\underline{i}'_k, \bar{i}'_k]$ or it can monotonically increase and then optionally monotonically decrease. In both the cases, it can be split into at-most 2 sections where the solution is purely monotonic. Similar arguments apply for the case $\text{gap}_{\min}(B,[i_s,i_t]) > \text{gap}_{\min}(-B,[i_s,i_t)$.

Similarly, if bin $[i_s, \underline{i}_1 - 1]$ is non-empty then the offline optimal must assume a piecewise monotonic structure with at-most 2 pieces. Then applying Lemma 187 to each of the 2 pieces separately and adding the regret bounds yields

$$\sum_{j=i_s}^{\underline{i}_1-1} f_j(x_j) - f_j(u_j) = \tilde{O}(1). \tag{E.65}$$

Note that $m_1^{(i)} + m_2^{(i)} = m^{(i)}$. Let the total regret contribution from Type 1 bins along with their pairs and Type 2 bins along with their pairs be referred as $R_1^{(i)}$ and $R_2^{(i)}$ respectively.

Since a sub-bin that is paired with a Type 1 or Type 2 bin can be split into at-most 2 sub-intervals where the optimal sequence is purely monotonic (see Fig. E.2), we can bound the regret within such sub-bins $[\underline{i}'_k, \bar{i}'_k]$, $k \in [m^{(i)}]$ by at-most $62(G+B)^2 \log n$ by Lemma 187.

For a Type 2 bin $[a, b] \subseteq [i_s, i_t]$, we can have two possible configurations: If $\text{gap}_{\min}(B, [i_s, i_t]) > \text{gap}_{\min}(-B, [i_s, i_t])$ then $B - u_a \geq \text{gap}_{\min}(B, [i_s, i_t]) \geq \mu_{\text{th}}$ where the first inequality follows by the definition of $\text{gap}_{\min}(B, [i_s, i_t])$ and the last inequality follows by Lemma 190. Similarly If $\text{gap}_{\min}(B, [i_s, i_t]) \leq \text{gap}_{\min}(-B, [i_s, i_t])$ then $B + u_a \geq \text{gap}_{\min}(-B, [i_s, i_t]) \geq \mu_{\text{th}}$. With this observation and using the results of Lemma 189, we can bound the regret contribution from any Type 2 bin and its pair as:

$$R_2^{(i)} \leq \sum_{j=1^{(i)}}^{m_2^{(i)}} \left(\left(10(G+B)^2 \log n - \ell_j^{(2)} \mu_{\text{th}}^2 \right) + 62(G+B)^2 \log n \right) \tag{E.66}$$

$$\leq 72 m_2^{(i)} (G+B)^2 \log n - \mu_{\text{th}}^2 \left(\sum_{j=1^{(i)}}^{m_2^{(i)}} \ell_j^{(2)} \right). \tag{E.67}$$

From Eq.(E.60), we have $\ell_j^{(2)} \geq \frac{2\phi n^{1/3}}{(G+B)C_n^{1/3}}$ for $j \in \{1^{(i)}, \ldots, m_2^{(i)}\}$. So we can continue as

$$R_2^{(i)} \leq 72 m_2^{(i)} (G+B)^2 \log n - \mu_{\text{th}}^2 \frac{2\phi n^{1/3}}{(G+B)C_n^{1/3}} m_2^{(i)} \tag{E.68}$$

$$= 0, \tag{E.69}$$

where the last line is obtained by plugging in the value of μ_{th} from Lemma 190.

So by refining every interval in $\mathcal{P}$ that satisfy Condition (A) and summing the regret contribution from all Type 2 bins and their pairs across all refined intervals in $\mathcal{P}$ yields

$$\sum_{i=1}^{M} R_2^{(i)} \leq 0, \tag{E.70}$$

where we recall that $M := |\mathcal{P}| = O(n^{1/3} C_n^{2/3} \vee 1)$ and assign $R_2(i) = 0$ for intervals in $\mathcal{P}$ that do not satisfy Condition (A).

For any Type 1 bin, its regret contribution can be bounded by Lemma 188. So we have the regret contribution from Type 1 bins and their pairs bounded as

$$R_1^{(i)} \leq \sum_{j=1}^{m_1^{(i)}} \left(\left(10(G+B)^2 \log n - \frac{4\lambda^2}{\ell_{j^{(i)}}^{(1)}} \right) + 62(G+B)^2 \log n \right) \tag{E.71}$$

$$= 72 m_1^{(i)} (G+B)^2 \log n - 4\lambda^2 \sum_{j=1}^{m_1^{(i)}} \frac{1}{\ell_{j^{(i)}}^{(1)}} \tag{E.72}$$

By refining every interval in $\mathcal{P}$ that satisfies Condition (A) and summing the regret contribution from all Type 2 bins and their pairs across all refined intervals in $\mathcal{P}$ yields

$$\sum_{i=1}^{M} R_1^{(i)} \leq 72(G+B)^2 \log n \sum_{i=1}^{M} m_1^{(i)} - 4\lambda^2 \sum_{i=1}^{M} \sum_{j=1}^{m_1^{(i)}} \frac{1}{\ell_{j^{(i)}}^{(1)}} \tag{E.73}$$

$$\leq 72(G+B)^2 M_1 \log n - 4\lambda^2 \frac{M_1^2}{n}, \tag{E.74}$$

where in the last line: a) we define $M_1 := \sum_{i=1}^{M} m_1^{(i)}$ with the convention that $m_1^{(i)} = 0$ if the bin $[i_s, i_t] \in \mathcal{P}$ doesn't satisfy Condition (A); b) applied AM-HM inequality and noted that $\sum_{i=1}^{M} \sum_{j=1}^{m_1^{(i)}} \ell_{j^{(i)}}^{(1)} \leq n$.

To further bound Eq.(E.74), we consider two separate regimes as follows.

Recall that $\lambda \geq \phi \frac{n^{1/3}}{C_n^{1/3}}$. So continuing from Eq.(E.74),

$$\sum_{i=1}^{M} R_1^{(i)} \leq 72(G+B)^2 M_1 \log n - 4\phi^2 \frac{n^{2/3}}{C_n^{2/3}} \frac{M_1^2}{n} \tag{E.75}$$

$$\leq 0, \tag{E.76}$$

whenever $M_1 \geq \frac{18(G+B)^2 \log n}{\phi^2} n^{1/3} C_n^{2/3} = \tilde{\Omega}(n^{1/3} C_n^{2/3})$.

In the alternate regime where $M_1 \leq \frac{18(G+B)^2 \log n}{\phi^2} n^{1/3} C_n^{2/3} = \tilde{O}(n^{1/3} C_n^{2/3} \vee 1)$, we trivially obtain:

$$\sum_{i=1}^{M} R_1^{(i)} = \tilde{O}(n^{1/3}C_n^{2/3} \vee 1) \tag{E.77}$$

Putting everything together by combining the bounds in Eq.(E.65), (E.70), (E.76) and (E.77), we can bound the total regret contribution from the bins that satisfy Condition (A) as:

$$\sum_{i=1}^{M} R_1^{(i)} + R_2^{(i)} + \tilde{O}(1) = \tilde{O}(n^{1/3}C_n^{2/3} \vee 1), \tag{E.78}$$

where we have assigned $R_1^{(i)} = R_2^{(i)} = 0$ for bins that don't satisfy Condition (A).

Throughout the proof till now, we have only considered bins $[i_s, i_t] \in \mathcal{P}$ which satisfy Condition (A). Not meeting this criterion will only make the arguments easier as explained below.

If a bin $[i_s, i_t] \in \mathcal{P}$ doesn't satisfy Condition (A), by taking a logical negation of Condition (A), we conclude that this can only happen if the optimal solution precludes the form of either Structure 1 or Structure 2 (or both) within some sub-interval of $[i_s, i_t]$. Consequently by applying similar arguments we used to handle the bins $[\underline{i}'_k, \bar{i}'_k]$, $k \in [m^{(i)}]$, we can split the offline optimal sequence $u_{i_s:i_t}$ into at-most 2 piece-wise monotonic sections and use Lemma 187 to bound the regret in $[i_s, i_t]$ as $\tilde{O}(1)$. Since $|\mathcal{P}| = O(n^{1/3}C_n^{2/3} \vee 1)$, we conclude that the total regret from all bins that don't satisfy Condition (A) is $\tilde{O}(n^{1/3}C_n^{2/3} \vee 1)$. □

Proof. **of Theorem 50**. The proof is now immediate from Lemmas 186 and 191. □

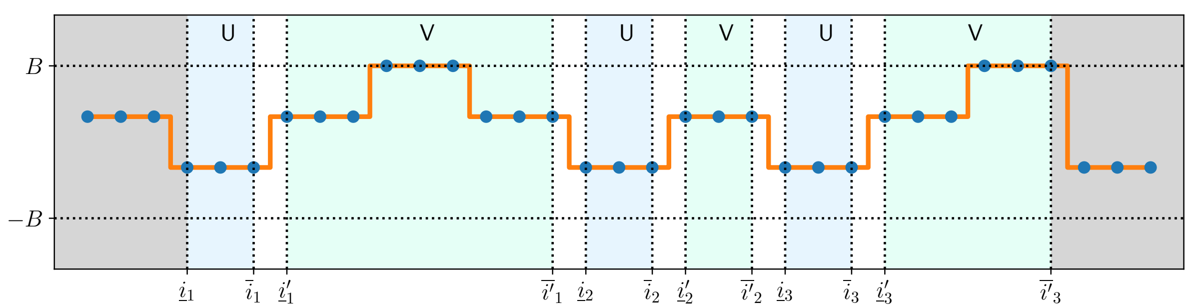


Figure E.2: *Refinement of a bin that satisfy Condition (A) in the proof of Lemma 191 with* $gap_{min}(-B, [i_s, i_t]) \geq gap_{min}(B, [i_s, i_t])$. *Here we assign* $i_s = \underline{i}_1$ *and* $i_t = \bar{i}'_3$ *The following pairs are formed in the proof of Lemma 191:* $\mathcal{P}_i = ([\underline{i}_1, \bar{i}_1], [\underline{i}'_1, \bar{i}'_1]), ([\underline{i}_2, \bar{i}_2], [\underline{i}'_2, \bar{i}'_2]), ([\underline{i}_3, \bar{i}_3], [\underline{i}'_3, \bar{i}'_3])$. *Blue dots represent the optimal sequence*

E.2 Proofs for Section 6.3

Corollary 54. *Let the loss functions f_t be H strongly convex in L_2 norm across the (box) domain $\mathcal{D} = \{\boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x}\|_\infty \le B\}$. i.e, $f_t(\boldsymbol{y}) \ge f_t(\boldsymbol{x}) + \nabla f_t(\boldsymbol{x})^T(\boldsymbol{y}-\boldsymbol{x}) + \frac{H}{2}\|\boldsymbol{y}-\boldsymbol{x}\|_2^2$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{D}$. Suppose $\|\nabla f_t(\boldsymbol{x})\|_\infty \le G_\infty$ for all $\boldsymbol{x} \in \mathcal{D}$. For each $i \in [d]$, construct surrogate losses $\ell_t^{(i)} : \mathbb{R} \to \mathbb{R}$ as $\ell_t^{(i)}(x) = (x - (\boldsymbol{x}_t[i] - \nabla f_t(\boldsymbol{x}_t)[i]/H))^2$ where $\boldsymbol{x}_t$ is the prediction of the learner at time t. By running d instances of uni-variate FLH-OGD with decision set $[-B, B]$ and learning rate $\zeta = 1/(2(2B+G_\infty/H)^2)$ where instance i predicts $\boldsymbol{x}_t[i]$ at time t and suffers losses $\ell_t^{(i)}$, we have*

$$\sum_{t=1}^{n} f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{w}_t) = \tilde{O}\left(d^{1/3} n^{1/3} C_n^{2/3} \vee d\right), \tag{6.10}$$

for any comparator sequence $\boldsymbol{w}_{1:n}$ with $TV(\boldsymbol{w}_{1:n}) := \sum_{t=2}^n \|\boldsymbol{w}_t - \boldsymbol{w}_{t-1}\|_1 \le C_n$. $\tilde{O}(\cdot)$ hides the dependence on factors of $\log n, B, H, G_\infty$.

Proof. Due to strong convexity, we have for any $\boldsymbol{w}_t \in \mathbb{R}^d$,

$$\begin{aligned}
f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{w}_t) &\le -\langle \nabla f_t(\boldsymbol{x}_t), \boldsymbol{w}_t - \boldsymbol{x}_t\rangle - \frac{H}{2}\|\boldsymbol{w}_t - \boldsymbol{x}_t\|^2 &\text{(E.79)}\\
&= H\left(\langle \nabla f_t(\boldsymbol{x}_t)/H, \boldsymbol{x}_t - \boldsymbol{x}_t\rangle + (1/2)\|\boldsymbol{x}_t - \boldsymbol{x}_t\|^2)\right) &\text{(E.80)}\\
&\quad - H\left(\langle \nabla f_t(\boldsymbol{x}_t)/H, \boldsymbol{w}_t - \boldsymbol{x}_t\rangle + (1/2)\|\boldsymbol{w}_t - \boldsymbol{x}_t\|^2)\right) &\text{(E.81)}\\
&= \sum_{i=1}^{d} H\left(\nabla f_t(\boldsymbol{x}_t)[i](\boldsymbol{x}_t[i] - \boldsymbol{x}_t[i])/H + (1/2)(\boldsymbol{x}_t[i] - \boldsymbol{x}_t[i])^2\right) &\text{(E.82)}\\
&\quad - H\left(\nabla f_t(\boldsymbol{x}_t)[i](\boldsymbol{w}_t[i] - \boldsymbol{x}_t[i])/H + (1/2)(\boldsymbol{w}_t[i] - \boldsymbol{x}_t[i])^2\right) &\text{(E.83)}\\
&= (H/2)\left(\sum_{i=1}^{d} \ell_t^{(i)}(\boldsymbol{x}_t[i]) - \ell_t^{(i)}(\boldsymbol{w}_t[i]),\right) &\text{(E.84)}
\end{aligned}$$

where the last line is obtained by completing the squares. Let $\boldsymbol{u}_t \in \mathbb{R}^d$ for $t \in [n]$ be defined as the offline optimal sequence corresponding to the optimization problem:

$$\begin{aligned}
\min_{\tilde{\boldsymbol{u}}_1, \ldots, \tilde{\boldsymbol{u}}_n, \tilde{z}_1, \ldots, \tilde{\boldsymbol{z}}_{n-1}} \quad & \sum_{t=1}^{n}\sum_{i=1}^{d} \ell_t^{(i)}(\tilde{\boldsymbol{u}}_t[i]) &\text{(E.85a)}\\
\text{s.t.} \quad & \tilde{\boldsymbol{z}}_t = \tilde{\boldsymbol{u}}_{t+1} - \tilde{\boldsymbol{u}}_t \ \forall t \in [n-1], &\text{(E.85b)}\\
& \sum_{t=1}^{n-1} \|\tilde{\boldsymbol{z}}_t\|_1 \le C_n, &\text{(E.85c)}\\
& \|\tilde{\boldsymbol{u}}_t\|_\infty \le B \ \forall t \in [n], &\text{(E.85d)}
\end{aligned}$$

Let $C_n[i] = \sum_{t=2}^n |\boldsymbol{u}_t[i] - \boldsymbol{u}_{t-1}[i]|$ be its TV allocated to coordinate i. By Theorem

1, the FLH-OGD instance i with learning rate $\zeta = 1/(2(2B + G_\infty/H)^2)$ attains the regret of $\tilde{O}\left(n^{1/3}(C_n[i])^{2/3} \vee 1\right)$ regret. WLOG, let's assume that FLH-OGD instances for coordinates $i \in [k]$, $k \le d$ incurs $\tilde{O}\left(n^{1/3}(C_n[i])^{2/3}\right)$ regret wrt losses $\ell_t^{(i)}$ and the regret incurred by FLH-OGD instances for coordinates $k > k'$ is $O(\log n)$. Let $R_n(\boldsymbol{w}_{1:n}) := \sum_{t=1}^n f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{w}_t)$ and $R'_n(\boldsymbol{w}_{1:n}) := (H/2)\left(\sum_{t=1}^n \sum_{i=1}^d \ell_t^{(i)}(\boldsymbol{x}_t[i]) - \ell_t^{(i)}(\boldsymbol{w}_t[i])\right)$. From Eq.(E.84) $R_n(\boldsymbol{w}_{1:n}) \le R'_n(\boldsymbol{w}_{1:n})$. We have,

$$
\begin{aligned}
R_n(\boldsymbol{w}_{1:n}) &\le R'_n(\boldsymbol{w}_{1:n}) && \text{(E.86)}\\
&\le \sup_{\substack{\boldsymbol{w}_1,\dots,\boldsymbol{w}_n \in \mathcal{D} \\ \sum_{t=2}^n \|\boldsymbol{w}_t - \boldsymbol{w}_{t-1}\|_1 \le C_n}} R'_n(\boldsymbol{w}_{1:n}) && \text{(E.87)}\\
&= R'_n(\boldsymbol{u}_{1:n}) && \text{(E.88)}\\
&= (d-k)\tilde{O}(1) + \sum_{i=1}^k \tilde{O}\left(n^{1/3}(C_n[i])^{2/3}\right) && \text{(E.89)}\\
&\le (d-k)\tilde{O}(1) + \tilde{O}\left(n^{1/3}(k)^{1/3}\left(\sum_{i=1}^k C_n[i]\right)^{2/3}\right), && \text{(E.90)}
\end{aligned}
$$

where the last line follows by Holder's inequality $\boldsymbol{x}^T\boldsymbol{y} \le \|\boldsymbol{x}\|_3\|\boldsymbol{y}\|_{3/2}$, where we treat $\boldsymbol{x}$ as just a vector of ones in $\mathbb{R}^k$. The above expression can be further upper bounded by $\tilde{O}\left(2d \vee 2d^{1/3}n^{1/3}C_n^{2/3}\right)$. □

E.3 Proofs for Section 6.4

We start by inspecting the KKT conditions.

Lemma 192. *(**characterization of offline optimal**) Consider the following convex optimization problem (where $\tilde{\boldsymbol{z}}_1, ..., \tilde{\boldsymbol{z}}_{n-1}$ are introduced as dummy variables).*

$$
\begin{aligned}
\min_{\tilde{\boldsymbol{u}}_1,\dots,\tilde{\boldsymbol{u}}_n,\tilde{\boldsymbol{z}}_1,\dots,\tilde{\boldsymbol{z}}_{n-1}} \quad & \sum_{t=1}^n f_t(\tilde{\boldsymbol{u}}_t) && \text{(E.91a)}\\
\text{s.t.} \quad & \tilde{\boldsymbol{z}}_t = \tilde{\boldsymbol{u}}_{t+1} - \tilde{\boldsymbol{u}}_t \;\forall t \in [n-1], && \text{(E.91b)}\\
& \sum_{t=1}^{n-1} \|\tilde{\boldsymbol{z}}_t\|_1 \le C_n, && \text{(E.91c)}\\
& \|\tilde{\boldsymbol{u}}_t\|_\infty \le B \;\forall t \in [n], && \text{(E.91d)}
\end{aligned}
$$

Let $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n, \boldsymbol{z}_1, \dots, \boldsymbol{z}_{n-1} \in \mathbb{R}^d$ be the optimal primal variables and let $\lambda \ge 0$ be the optimal dual variable corresponding to the constraint (E.91c)*. Further, let $\boldsymbol{\gamma}_t^+, \boldsymbol{\gamma}_t^- \in \mathbb{R}^d$ with $\boldsymbol{\gamma}_t^+ \ge \boldsymbol{0}$ and $\boldsymbol{\gamma}_t^- \ge \boldsymbol{0}$ be the optimal dual variables that correspond to constraint* (E.91d)*. Specifically for $k \in [d]$, $\boldsymbol{\gamma}_t^+[k]$ corresponds to the dual variable for the constraint*

$\boldsymbol{u}_t[k] \leq B$ induced by the relation (E.91d). Similarly $\boldsymbol{\gamma}_t^-[k]$ corresponds to the constraint $-B \leq \boldsymbol{u}_t[k]$. By the KKT conditions, we have

- ***stationarity:*** $\nabla f_t(\boldsymbol{u}_t) = \lambda\left(\boldsymbol{s}_t - \boldsymbol{s}_{t-1}\right) + \boldsymbol{\gamma}_t^- - \boldsymbol{\gamma}_t^+$*, where* $\boldsymbol{s}_t \in \partial|\boldsymbol{z}_t|$ *(a subgradient). Specifically,* $\boldsymbol{s}_t[k] = \mathrm{sign}(\boldsymbol{u}_{t+1}[k] - \boldsymbol{u}_t[k])$ *if* $|\boldsymbol{u}_{t+1}[k] - \boldsymbol{u}_t[k]| > 0$ *and* $\boldsymbol{s}_t[k]$ *is some value in* $[-1, 1]$ *otherwise. For convenience of notations later, we also define* $\boldsymbol{s}_n = \boldsymbol{s}_0 = \boldsymbol{0}$.
- ***complementary slackness:*** *(a)* $\lambda\left(\sum_{t=2}^{n} \|\boldsymbol{u}_t - \boldsymbol{u}_{t-1}\|_1 - C_n\right) = 0$*; (b)* $\boldsymbol{\gamma}_t^-[k](\boldsymbol{u}_t[k] + B) = 0$ *and* $\boldsymbol{\gamma}_t^+[k](\boldsymbol{u}_t[k] - B) = 0$ *for all* $t \in [n]$ *and all* $k \in [d]$.

The proof of the above lemma is similar to that of Lemma 52 and hence omitted.

Terminology. We will refer to the optimal primal variables $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ in Lemma 192 as the *offline optimal sequence* in this section. We reserve the term *FLH-ONS* for the instantiation of FLH with ONS as base learners with parameters as in Theorem 59.

Notations. For bin $[i_s, i_t] \in \mathcal{P}$ we define: $n_i = i_t - i_s + 1$, $\bar{\boldsymbol{u}}_i = \frac{1}{n_i}\sum_{j=i_s}^{i_t} \boldsymbol{u}_j$, $\boldsymbol{\Gamma}_i^+ = \sum_{j=i_s}^{i_t} \boldsymbol{\gamma}_j^+$, $\boldsymbol{\Gamma}_i^- = \sum_{j=i_s}^{i_t} \boldsymbol{\gamma}_j^-$, $\Delta \boldsymbol{s}_i = \boldsymbol{s}_{i_t} - \boldsymbol{s}_{i_s - 1}$, $C_i = \sum_{j=i_s+1}^{i_t} \|\boldsymbol{u}_j - \boldsymbol{u}_{j-1}\|_1$.

For any general bin $[a, b]$ define the quantities $n_{a\to b}, \bar{\boldsymbol{u}}_{a\to b}, \boldsymbol{\Gamma}_{a\to b}^+, \boldsymbol{\Gamma}_{a\to b}^-, \Delta\boldsymbol{s}_{a\to b}, C_{a\to b}$ analogously as above.

The following is a direct extension for Lemma 184.

Lemma 193. ***(key partition)*** *Initialize* $\mathcal{P} \leftarrow \Phi$*. Starting from time 1, spawn a new bin* $[i_s, i_t]$ *whenever* $\sum_{j=i_s+1}^{i_t+1} \|\boldsymbol{u}_j - \boldsymbol{u}_{j-1}\|_1 > B/\sqrt{n_i}$*, where* $n_i = i_t - i_s + 2$*. Add the spawned bin* $[i_s, i_t]$ *to* $\mathcal{P}$.

Let $M := |\mathcal{P}|$*. We have* $M = O\left(1 \vee n^{1/3} C_n^{2/3} B^{-2/3}\right)$.

Proposition 194. *The losses* f_t *defined in Eq.*(6.19) *are:*

- G^2 *gradient Lipschitz over the domain* $\mathcal{D}$ *in Assumption B1*
- *Define* $\gamma := 2GB\sqrt{\alpha d/2} + 1/\sqrt{2\alpha}$*. Then the losses* f_t *are* $\alpha' := 1/(2\gamma^2)$ *exp-concave across* $\mathcal{D}$.
- f_t *are* $G' := 2\alpha G^2 B\sqrt{d} + G$ *Lipschitz in L2 norm across* $\mathcal{D}$.

Proof. The first two statements have been already proved in Section 6.4. For the last statement we have that

$$\nabla f_t(\boldsymbol{x}) = (\alpha \nabla \ell_t(\boldsymbol{x}_t)^T (\boldsymbol{x} - \boldsymbol{x}_t) + 1)\nabla \ell_t(\boldsymbol{x}_t). \tag{E.92}$$

So by triangle inequality we obtain that $\|\nabla f_t(\boldsymbol{x})\|_2 \leq 2\alpha G^2 B\sqrt{d} + G$. □

Lemma 195. ***(Strongly Adaptive regret)*** *([77], [23]) Consider any bin $[a, b]$ and a comparator $\boldsymbol{w} \in \mathcal{D}$. Under Assumptions B1-2 in Section 6.4, the static regret of the FLH-ONS with losses f_t obeys*

$$\sum_{j=a}^{b} f_j(\boldsymbol{x}_j) - f_j(\boldsymbol{w}) \leq 10d(8G^2B^2\alpha d + 1/\alpha)\log n \tag{E.93}$$

where $\boldsymbol{x}_j$ are predictions of FLH-ONS and γ is as defined in Theorem 59.

Proof. Let $\alpha' = 1/(2\gamma^2)$. The static regret of ONS is $5d(G'D + 1/\alpha')\log n$ for α' exp-concave losses (Theorem 2 in [77]) where D is the diameter of the decision set. We have $D = 2B\sqrt{d}$ for the box decision set. the static regret of ONS in our setting is at-most $5d(2G'B\sqrt{d} + 1/\alpha')\log n$.

The regret of the FLH against any of its base experts is at-most $(4/\alpha')\log n$ for α' exp-concave losses (Theorem 3.2 in [23]). Adding both these regret bounds, using Proposition 194 and further upper bounding the sum results in the lemma.

□

Lemma 196. ***(low $\boldsymbol{\lambda}$ regime)*** *If the optimal dual variable $\lambda = O\left(\frac{d^{1.5}n^{1/3}}{C_n^{1/3}}\right)$, we have the regret of FLH-ONS strategy bounded as*

$$\sum_{t=1}^{n} f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{u}_t) = \tilde{O}\left(10d(8G^2B^2\alpha d + 1/\alpha)(n^{1/3}C^{2/3} \vee 1)\right), \tag{E.94}$$

where $\boldsymbol{x}_t$ is the prediction of FLH-ONS at time t.

Proof. Consider a bin $[i_s, i_t] \in \mathcal{P}$. Note that for any $j \in [i_s, i_t]$ and $k \in [d]$, both $\boldsymbol{\gamma}_j^+[k]$ and $\boldsymbol{\gamma}_j^-[k]$ can't be simultaneously non-zero due to complementary slackness and the fact that $C_i \leq B/\sqrt{n_i} < 2B$ by the construction in Lemma 193. For some fixed $\check{\boldsymbol{u}} \in \mathcal{D}$, we have

$$\underbrace{\sum_{j=i_s}^{i_t} f_j(\boldsymbol{x}_j) - f_j(\check{\boldsymbol{u}})}_{T_{1,i}} + \underbrace{\sum_{j=i_s}^{i_t} f_j(\check{\boldsymbol{u}}) - f_j(\boldsymbol{u}_j)}_{T_{2,i}}. \tag{E.95}$$

By virtue of Lemma 195, we have $T_{1,i} = \tilde{O}(d^2)$. Due to gradient Lipschitzness in Proposition 194

$$T_{2,i} \leq \sum_{j=i_s}^{i_t} \langle \nabla f_j(\boldsymbol{u}_j), \check{\boldsymbol{u}} - \boldsymbol{u}_j \rangle + \frac{G^2}{2}\|\check{\boldsymbol{u}} - \boldsymbol{u}_j\|_2^2. \tag{E.96}$$

We construct $\check{\boldsymbol{u}}$ as follows:

- If there exists a $j \in [i_s, i_t]$ and $k \in [d]$ such that $\boldsymbol{u}_j[k] = B$, set $\check{\boldsymbol{u}}[k] = B$.
- If there exists a $j \in [i_s, i_t]$ and $k \in [d]$ such that $\boldsymbol{u}_j[k] = -B$, set $\check{\boldsymbol{u}}[k] = -B$.
- If the optimal solution doesn't touch either boundaries $\pm B$ in $[i_s, i_t]$ across a coordinate, set $\check{\boldsymbol{u}}[k] = \boldsymbol{u}_{i_s}[k]$.

It is easy to see that $\check{\boldsymbol{u}} \in \mathcal{D}$ and $\|\check{\boldsymbol{u}} - \boldsymbol{u}_j\|_2 \leq \|\check{\boldsymbol{u}} - \boldsymbol{u}_j\|_1 \leq C_i$ for all $j \in [i_s, i_t]$. Using this observation along with the KKT conditions, we continue from Eq.(E.96) as

$$T_{2,i} \leq G^2 n_i C_i^2 + \sum_{j=i_s}^{i_t} \langle \nabla f_j(\boldsymbol{u}_j), \check{\boldsymbol{u}} - \boldsymbol{u}_j \rangle, \tag{E.97}$$

$$\leq G^2 n_i C_i^2 + \sum_{j=i_s}^{i_t} \lambda \langle \boldsymbol{s}_j - \boldsymbol{s}_{j-1}, \check{\boldsymbol{u}} - \boldsymbol{u}_j \rangle + \langle \boldsymbol{\gamma}_j^- - \boldsymbol{\gamma}_j^+, \check{\boldsymbol{u}} - \boldsymbol{u}_j \rangle \tag{E.98}$$

$$\leq_{(a)} G^2 B^2 + \lambda \langle \boldsymbol{s}_{i_s - 1}, \boldsymbol{u}_{i_s} - \check{\boldsymbol{u}} \rangle - \lambda \langle \boldsymbol{s}_{i_t}, \boldsymbol{u}_{i_t} - \check{\boldsymbol{u}} \rangle + \lambda \sum_{j=i_s+1}^{i_t} \|\boldsymbol{u}_j - \boldsymbol{u}_{j-1}\|_1 \tag{E.99}$$

$$+ \sum_{j=i_s}^{i_t} \sum_{k=1}^{d} \boldsymbol{\gamma}_j^-[k](\check{\boldsymbol{u}}[k] - \boldsymbol{u}_j[k]) - \gamma_j^+[k](\check{\boldsymbol{u}}[k] - \boldsymbol{u}_j[k]) \tag{E.100}$$

$$\leq_{(b)} G^2 B^2 + 3\lambda C_i, \tag{E.101}$$

where line (a) is obtained by using that fact that $C_i \leq B/\sqrt{n_i}$ and a rearrangement of the summations and line (b) is obtained by noting that $\boldsymbol{\gamma}_j^-[k] = 0$ when $\boldsymbol{u}_j[k] > -B$ via complementary slackness and $\check{\boldsymbol{u}}[k] - \boldsymbol{u}_j[k]$ is zero when $\boldsymbol{u}_j[k] = -B$ since by construction of $\check{\boldsymbol{u}}$: $\check{\boldsymbol{u}}[k] = -B$ if $\boldsymbol{u}_j[k] = -B$ for some $j \in [i_s, i_t]$. Similar arguments are applied to show the terms including $\boldsymbol{\gamma}_j^+$ also sums to zero. In line (b) we also used the fact that $\langle \boldsymbol{s}_{i_s - 1}, \boldsymbol{u}_{i_s} - \check{\boldsymbol{u}} \rangle \leq \|\boldsymbol{s}_{i_s-1}\|_\infty \|\boldsymbol{u}_{i_s} - \check{\boldsymbol{u}}\|_1 \leq C_i$. Similarly $\langle \boldsymbol{s}_{i_t}, \boldsymbol{u}_{i_t} - \check{\boldsymbol{u}} \rangle \leq C_i$

Hence summing $T_{1,i}$ and $T_{2,i}$ across all bins in $\mathcal{P}$ yields

$$\sum_{i=1}^{M} T_{1,i} + T_{2,i} \leq_{(a)} \tilde{O}(Md^2) + \lambda C_n \tag{E.102}$$

$$\leq \tilde{O}\left(10d(8G^2B^2\alpha d + 1/\alpha)(n^{1/3}C^{2/3} \vee 1)\right), \tag{E.103}$$

where we recall that $M := |\mathcal{P}| = O(1 \vee n^{1/3} C_n^{2/3})$ by Lemma 193 and in line (a) we used $\sum_{i=1}^{M} C_i \leq C_n$ and $\lambda = O(d^{1.5} n^{1/3}/C_n^{1/3})$ by the premise of the current Lemma.

□

Definition 197. For a bin $[a, b]$, the offline optimal is said to be **piece-wise maximally monotonic in $[\mathbf{a}, \mathbf{b}]$ with $\boldsymbol{m}$ pieces across some coordinate $\mathbf{k} \in [\mathbf{d}]$**, if we can split

$[a, b]$ into m disjoint consecutive bins $[a_1, b_1], \ldots, [a_m, b_m]$ such that the offline optimal sequence within each $[a_i, b_i]$ is purely monotonic across coordinate k'. Further, right-extending any interval $[a_i, b_i]$ to $[a_i, b_i+1]$ if $b_i+1 \in [a, b]$ makes $\boldsymbol{u}_{a_i:b_i+1}[k']$ non-monotonic. The sections $[a_i, b_i]$ for $i \in [m]$ are termed **maximally monotonic sections**.

Lemma 198. *The sequence returned at Step 3 of* `generateGhostSequence` *in Fig.E.3 has the following properties:*

Property 1 *The elements in the sequence changes only at-most* $3d$ *times. i.e,* $\sum_{j=a+1}^{b} \mathbb{I}(\check{\boldsymbol{u}}_j \neq \check{\boldsymbol{u}}_{j-1}) \leq 7d$, *where* $\mathbb{I}(\cdot)$ *is the indicator function.*

Property 2 *Every member of the sequence lie in the box decision set* $\mathcal{D}$.

Property 3 *For any* $j \in [a, b]$, $\sum_{\substack{k=1 \\ k \neq k_{fix}}}^{d} |\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]| \leq C_{a\to b}$, *where* $C_{a\to b}$ *is the TV of the offline optimal in bin* $[a, b]$.

Proof. Observe that in the procedure detailed in Fig.E.3, we split the bin $[a, b]$ into at-most 7 bins across any coordinate. The value of the comparator across that coordinate stays unchanged in each of the new sub-bins. This implies that number of distinct comparators in $\{\check{\boldsymbol{u}}_a, \ldots, \check{\boldsymbol{u}}_b\}$ is at-most $7d$. It is also easy to see that each $\check{\boldsymbol{u}}_j$, $j \in [a, b]$ stays inside the decision set $\mathcal{D}$.

Note that for any $j \in [a, b]$ and any $k \in [d] \setminus \{k_{\text{fix}}\}$, $\check{\boldsymbol{u}}_j[k]$ coincides with the value of $\boldsymbol{u}_{j'}[k]$ for some $j' \in [a, b]$. This implies that $|\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]| \leq C_{a\to b}[k]$ for any $j \in [a, b]$, where $C_{a\to b}[k]$ is the TV of the optimal solution across coordinate k in bin $[a, b]$. So $\sum_{\substack{k=1 \\ k \neq k_{\text{fix}}}}^{d} |\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]| \leq \sum_{\substack{k=1 \\ k \neq k_{\text{fix}}}}^{d} C_{a\to b}[k] \leq C_{a\to b}$. Thus Property 3 is true. □

Lemma 199. ***(monotonic bins)*** *Consider a bin* $[a, b]$ *with length* ℓ *where the offline optimal sequence is piece-wise maximally monotonic in* $[a, b]$ *across any coordinate with at-most 4 pieces. Let the TV of the optimal solution within bin* $[a, b]$ *denoted by* $C_{a\to b}$ *be at-most* $B/\sqrt{\ell}$. *Then we have the regret of FLH-ONS strategy in this bin bounded as*

$$\sum_{j=a}^{b} f_j(\boldsymbol{x}_j) - f_j(\boldsymbol{u}_j) \leq 70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n, \tag{E.104}$$

where $\boldsymbol{x}_j$ *are the predictions of the FLH-ONS lagorithm.*

Proof. We first construct a useful sequence of comparators:
$\check{\boldsymbol{u}}_{a:b} =$ `generateGhostSequence` $(\boldsymbol{u}_{1:n}, k_{\text{fix}} = 0, u_{\text{fix}} = 0, [a, b])$.

We remark that as $k_{\text{fix}} = 0 \notin [d]$, the condition in Step 2(a) of Fig.E.3 is never satisfied.

`generateGhostSequence`: Inputs- (1) offline optimal sequence (2) two numbers $k_{\text{fix}} \in [d] \cup \{0\}$ and $u_{\text{fix}} \in [-B, B]$ (3) an interval $[a, b] \subseteq [n]$ where the offline optimal is piece-wise maximally monotonic with at-most 4 pieces across any coordinate $k \in [d]$.

1. Initialize $\mathcal{Q} \leftarrow \Phi$.
2. For each coordinate $k \in [d]$:
   (a) If k is same as k_{fix}, then set $\check{\boldsymbol{u}}_t[k] = u_{\text{fix}}$ for all $t \in [a, b]$. Goto Step 2.
   (b) If the optimal solution is constant across coordinate k, set $\check{\boldsymbol{u}}_t[k] = \boldsymbol{u}_a[k]$ for all $t \in [a, b]$. Goto Step 2
   (c) If the optimal solution monotonically increases (decreases) first across coordinate k, then:
      i. Split $[a, b]$ into at-most 7 sub-bins – $[\underline{r}_i, \bar{r}_i]$, $i \in [7]$ – with the following properties:
         - $\underline{r}_1 = a$. $\underline{r}_2$ is the largest value in $[a, b]$ such that $\boldsymbol{u}_{\underline{r}_1:\underline{r}_2}$ is monotonically increasing (decreasing) and $\boldsymbol{u}_{\underline{r}_2}[k] \underset{(<)}{>} \boldsymbol{u}_{\underline{r}_2-1}[k]$. Set $\bar{r}_1 = \underline{r}_2 - 1$.
         - $\bar{r}_2$ is the largest value in $[a, b]$ such that $\boldsymbol{u}_{\underline{r}_2:\bar{r}_2}[k]$ is constant.
         - If $\bar{r}_2 = b$, then set $[\underline{r}_i, \bar{r}_i]$, $i \in [3, 7]$ to be empty. Goto Step 2(c)(ii).
         - $\underline{r}_3 = \bar{r}_2 + 1$.
         - If $\boldsymbol{u}_{\underline{r}_3:b}[k]$ is a constant, set $\bar{r}_3 = b$. Set $[\underline{r}_i, \bar{r}_i]$, $i \in [4, 7]$ to be empty. Goto Step 2(c)(ii).
         - $\underline{r}_4$ is the largest point in $[\underline{r}_3, b]$ such that $\boldsymbol{u}_{\underline{r}_3:\underline{r}_4}[k]$ is monotonically decreasing (increasing) and $u_{\underline{r}_4}[k] \underset{(>)}{<} u_{\underline{r}_4-1}[k]$. Set $\bar{r}_3 = \underline{r}_4 - 1$.
         - $\bar{r}_4$ is the largest point in $[\underline{r}_4, b]$ such that $\boldsymbol{u}_{\underline{r}_4:\bar{r}_4}[k]$ is constant.
         - If $\bar{r}_4 = b$, Set $[\underline{r}_i, \bar{r}_i]$, $i \in [5, 7]$ to be empty. Goto Step 2(c)(ii).
         - $\underline{r}_5 = \bar{r}_4 + 1$.
         - If $\boldsymbol{u}_{\underline{r}_5:b}[k]$ is constant, then set $\bar{r}_5 = b$ and $[\underline{r}_i, \bar{r}_i]$, $i \in [6, 7]$ to be empty. Goto Step 2(c)(ii).
         - $\underline{r}_6$ is the largest point such that $\boldsymbol{u}_{\underline{r}_5:\underline{r}_6}[k]$ is monotonically increasing (decreasing) and $\boldsymbol{u}_{\underline{r}_6}[k] \underset{(<)}{>} \boldsymbol{u}_{\underline{r}_6-1}[k]$. Set $\bar{r}_5 = \underline{r}_6 - 1$.
         - $\bar{r}_6$ is the largest point in $[\underline{r}_6, b]$ such that $\boldsymbol{u}_{\underline{r}_6:\bar{r}_6}$ is constant.
         - If $\bar{r}_6 = b$, set $[\underline{r}_7, \bar{r}_7]$ as empty. Goto Step 2(c)(ii).
         - Set $\underline{r}_7 = \bar{r}_6 + 1$ and $\bar{r}_7 = b$.

      ii. Assign $\check{\boldsymbol{u}}_t[k] = \boldsymbol{u}_{\underline{r}_1}[k]$ for all $t \in [\underline{r}_1, \bar{r}_1]$; $\check{\boldsymbol{u}}_t[k] = \boldsymbol{u}_{\underline{r}_2}[k]$ for all $t \in [\underline{r}_2, \bar{r}_2]$ if non-empty; $\check{\boldsymbol{u}}_t[k] = \boldsymbol{u}_{\underline{r}_3}[k]$ for all $t \in [\underline{r}_3, \bar{r}_3]$ if non-empty; $\check{\boldsymbol{u}}_t[k] = \boldsymbol{u}_{\underline{r}_4}[k]$ for all $t \in [\underline{r}_4, \bar{r}_4]$ if non-empty; $\check{\boldsymbol{u}}_t[k] = \boldsymbol{u}_{\underline{r}_5}[k]$ for all $t \in [\underline{r}_5, \bar{r}_5]$ if non-empty; $\check{\boldsymbol{u}}_t[k] = \boldsymbol{u}_{\underline{r}_6}[k]$ for all $t \in [\underline{r}_6, \bar{r}_6]$ if non-empty; $\check{\boldsymbol{u}}_t[k] = \boldsymbol{u}_{\bar{r}_7}[k]$ for all $t \in [\underline{r}_7, \bar{r}_7]$ if non-empty;
3. Return $\{\check{\boldsymbol{u}}_a, \ldots, \check{\boldsymbol{u}}_b\}$.

Figure E.3: *generateGhostSequence procedure. If line 2(c) is replaced by "If the optimal solution monotonically decreases first across coordinate k, then", then we propagate that change by replacing the phrases increasing/decreasing and $>$ / $<$ in the lines below 2(c)(i) by the bracketed statements next to it.*

Next, we employ a two term regret decomposition as follows

$$\underbrace{\sum_{j=a}^{b} f_j(\boldsymbol{x}_j) - f_j(\check{\boldsymbol{u}}_j)}_{T_1} + \underbrace{\sum_{j=a}^{b} f_j(\check{\boldsymbol{u}}_j) - f_j(\boldsymbol{u}_j)}_{T_2}. \tag{E.105}$$

By noting that there are only at-most $7d$ change points in the comparator sequence (see Lemma 198), we can sum up the SA regret guarantee from Lemma 195 against each of the constant sections of $\check{\boldsymbol{u}}_{a:b}$ to obtain

$$T_1 \le 70d^2(8G^2B^2\alpha d + 1/\alpha)\log n. \tag{E.106}$$

To bound T_2 we use gradient Lipschitzness in Proposition 194 and look at a coordinate-wise decomposition.

$$T_2 \le \sum_{j=a}^{b} \langle \nabla f_j(\boldsymbol{u}_j), \check{\boldsymbol{u}}_j - \boldsymbol{u}_j \rangle + \frac{G^2}{2}\|\check{\boldsymbol{u}}_j - \boldsymbol{u}_j\|_2^2 \tag{E.107}$$

$$\le \frac{\ell G^2 C_{a\to b}^2}{2} + \sum_{k=1}^{d}\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k](\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]), \tag{E.108}$$

where in the last line we used that fact that $\|\check{\boldsymbol{u}}_j - \boldsymbol{u}_j\|_2^2 \le \|\check{\boldsymbol{u}}_j - \boldsymbol{u}_j\|_1^2 \le C_{a\to b}^2$ by Property 3 of Lemma 198, where $C_{a\to b}$ is the TV of the optimal solution within bin $[a, b]$.

Since $C_{a\to b} \le B\sqrt{\ell}$, we have the first term in Eq.(E.108) bounded by $\frac{G^2B^2}{2}$. Next we proceed to bound the second term in Eq.(E.108) coordinate-wise. Consider a coordinate $k \in [d]$. We have two cases:

Case 1: When the optimal solution across coordinate k in bin $[a, b]$ has a structure described in Step 2(b) of the `generateGhostSequence` procedure of Fig.E.3. In this case $\check{\boldsymbol{u}}_j[k] = \boldsymbol{u}_j[k] = \boldsymbol{u}_a[k]$ for $j \in [a, b]$. So

$$\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k](\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]) = 0. \tag{E.109}$$

Case 2: When the optimal solution across coordinate k in bin $[a, b]$ has a structure described in Step 2(c) of the `generateGhostSequence` procedure of Fig.E.3. In this case, we can write

$$\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k](\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]) = \sum_{i=1}^{7}\sum_{j=\underline{r}_i}^{\bar{r}_i} \nabla f_j(\boldsymbol{u}_j)[k](\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]), \tag{E.110}$$

where $[\underline{r}_i, \bar{r}_i]$, $i \in [7]$ are as defined in `generateGhostSequence` of Fig.E.3.

From Step 2(c)(ii) we have for each $i \in \{2, 4, 6\}$, $\check{\boldsymbol{u}}_j[k] = \boldsymbol{u}_j[k] = \boldsymbol{u}_{\underline{r}_i}[k]$ for all

$j \in [\underline{r}_i, \bar{r}_i]$ if non-empty. So $\sum_{j=\underline{r}_i}^{\bar{r}_i} \nabla f_j(\boldsymbol{u}_j)[k](\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]) = 0$ for each $i \in \{2, 4, 6\}$.

Next we consider the interval $[\underline{r}_1, \bar{r}_1]$. If within bin $[\underline{r}_1, \bar{r}_1]$, the optimal solution across coordinate k is constant, then $\sum_{j=\underline{r}_1}^{\bar{r}_1} \nabla f_j(\boldsymbol{u}_j)[k](\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]) = 0$. Otherwise let $[\underline{r}_1, \bar{r}_1] = [\underline{r}_1, p] \cup [p+1, \bar{r}_1]$ such that the optimal solution is constant in $[\underline{r}_1, p]$ and non-decreasing (non-increasing) within $[p+1, \bar{r}_1]$ across coordinate k. Recall from Fig.E.3 that $\underline{r}_1 = a$. Since $\check{\boldsymbol{u}}_j[k] = \boldsymbol{u}_a[k]$ for all $j \in [\underline{r}_1, p]$ we get $\sum_{j=\underline{r}_1}^{p} \nabla f_j(\boldsymbol{u}_j)[k](\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]) = 0$. Further note that due to the presence of bins $[\underline{r}_1, p]$ and $[\underline{r}_2, \bar{r}_2]$ the solution $\boldsymbol{u}_j[k]$ for $j \in [p+1, \bar{r}_1]$ will never touch the boundaries $\pm B$. So by the KKT conditions and using $\check{\boldsymbol{u}}_j[k] = \boldsymbol{u}_a[k]$ for $j \in [p+1, \bar{r}_1]$, we have

$$\sum_{j=p+1}^{\bar{r}_1} \nabla f_j(\boldsymbol{u}_j)[k](\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]) = \sum_{j=p+1}^{\bar{r}_1} \lambda(\boldsymbol{s}_j[k] - \boldsymbol{s}_{j-1}[k])(\boldsymbol{u}_a[k] - \boldsymbol{u}_j[k]) \tag{E.111}$$

$$= \lambda\left(\boldsymbol{s}_p[k](\boldsymbol{u}_{p+1}[k] - \boldsymbol{u}_a[k]) - \boldsymbol{s}_{\bar{r}_1}[k](\boldsymbol{u}_{\bar{r}_1}[k] - \boldsymbol{u}_a[k])\right) \tag{E.112}$$

$$+ \lambda \sum_{j=p+2}^{\bar{r}_1} |\boldsymbol{u}_j[k] - \boldsymbol{u}_{j-1}[k]| \tag{E.113}$$

$$= 0, \tag{E.114}$$

where the last line is obtained as follows: Observe that $\boldsymbol{s}_p[k] = \boldsymbol{s}_{\bar{r}_1}[k] = 1$ (or -1) and $\boldsymbol{s}_p[k]\boldsymbol{u}_{p+1}[k] - \boldsymbol{s}_{\bar{r}_1}[k]\boldsymbol{u}_{\bar{r}_1}[k] = -C_{p+1 \to \bar{r}_1}$ due to monotonicity of $\boldsymbol{u}_{p+1:\bar{r}_1}$

By using similar arguments we used to show Eq.(E.114), it can be proved that

$$\sum_{j=\underline{r}_i}^{\bar{r}_i} \nabla f_j(\boldsymbol{u}_j)[k](\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]) = \sum_{j=\underline{r}_i}^{\bar{r}_i} \nabla f_j(\boldsymbol{u}_j)[k](\boldsymbol{u}_{\underline{r}_i}[k] - \boldsymbol{u}_j[k]) \tag{E.115}$$

$$= 0, \tag{E.116}$$

for $i \in \{3, 5\}$.

Further, by using similar arguments we used to handle $[\underline{r}_1, \bar{r}_1]$, it can be shown that

$$\sum_{j=\underline{r}_7}^{\bar{r}_7} \nabla f_j(\boldsymbol{u}_j)[k](\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]) = 0. \tag{E.117}$$

Thus overall by combining Case 1 and 2 and continuing from Eq.(E.108), we have $T_2 \le G^2B^2/2$. Thus the total regret

$$T_1 + T_2 \le 70d^2(8G^2B^2\alpha d + 1/\alpha)\log n + G^2B^2/2 \tag{E.118}$$

$$\le 70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n, \tag{E.119}$$

which concludes the proof.

□

Definition 200. We introduce the following definitions for convenience.

- For a bin $[a, b] \subseteq \{2, \ldots, n-1\}$, the offline optimal solution is said to assume Structure 1 across coordinate k if $\boldsymbol{u}_j[k] = \boldsymbol{u}_a[k] \in (-B, B)$ for all $j \in [a, b]$ and $\boldsymbol{u}_b[k] > \boldsymbol{u}_{b+1}[k]$ and $\boldsymbol{u}_a[k] > \boldsymbol{u}_{a-1}[k]$.
- For a bin $[a, b] \subseteq \{2, \ldots, n-1\}$, the offline optimal solution is said to assume Structure 2 across coordinate k if $\boldsymbol{u}_j[k] = \boldsymbol{u}_a[k] \in (-B, B)$ for all $j \in [a, b]$ and $\boldsymbol{u}_b[k] < \boldsymbol{u}_{b+1}[k]$ and $\boldsymbol{u}_a[k] < \boldsymbol{u}_{a-1}[k]$.
- A bin $[r, s]$ is said to contain Structure 1 and Structure 2 if across some coordinate k, the offline optimal solution assumes the form of Structure 1 in an interval $[a, b] \subset [r, s]$ and Structure 2 in some interval $[a', b'] \subset [r, s]$ with $[a, b] \cap [a', b'] = \Phi$.
- For a bin $[a, b]$, we define $\text{GAP}_{\min}(\beta, [a, b])[k] := \min_{j \in [a,b]} |\boldsymbol{u}_j[k] - \beta|$, where $\beta \in \mathbb{R}$.

Next we provide a lemma analogous to Lemma 188.

Lemma 201. *Consider a bin $[a, b]$ with length ℓ where the TV of the offline optimal obeys $C_{a \to b} \le B/\sqrt{\ell}$. Assume that for some coordinate $k' \in [d]$, $\boldsymbol{u}_{a:b}[k']$ takes the form of Structure 1 or Structure 2. Further suppose that across all coordinates, the offline optimal solution is piece-wise maximally monotonic in $[a, b]$ with at-most 4 pieces. If $\left|\boldsymbol{u}_a[k'] - \frac{1}{\ell G^2} \sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k']\right| \le B$, then*

$$\sum_{j=a}^{b} f_j(\boldsymbol{x}_j) - f_j(\boldsymbol{u}_j) \le 70 d^2 (8 G^2 B^2 \alpha d + G^2 B^2 + 1/\alpha) \log n - \frac{2\lambda^2}{\ell G^2}, \tag{E.120}$$

where $\boldsymbol{x}_j$ are the predictions of FLH-ONS.

Proof. Let $k_{\text{fix}} = k'$ and $u_{\text{fix}} = \boldsymbol{u}_a[k'] - \frac{1}{\ell G^2} \sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k']$. Consider a comparator sequence
$\check{\boldsymbol{u}}_{a:b} = \texttt{generateGhostSequence}(\boldsymbol{u}_{1:n}, k_{\text{fix}}, u_{\text{fix}}, [a, b])$. We use a two term regret decomposition

$$\underbrace{\sum_{j=a}^{b} f_j(\boldsymbol{x}_j) - f_j(\check{\boldsymbol{u}}_j)}_{T_1} + \underbrace{\sum_{j=a}^{b} f_j(\check{\boldsymbol{u}}_j) - f_j(\boldsymbol{u}_j)}_{T_2}. \tag{E.121}$$

By Properties 1 and 2 in Lemma 198, we know that the comparator $\check{\boldsymbol{u}}_{a:b}$ changes only at-most $7d$ times and every single point in the sequence belongs to $\mathcal{D}$. Hence by strong adaptivity (Lemma 195), we have

$$T_1 \leq 70d^2(8G^2B^2\alpha d + 1/\alpha)\log n. \tag{E.122}$$

Further via gradient Lipschitzness in Proposition 194,

$$T_2 \leq \sum_{j=a}^{b} \langle \nabla f_j(\boldsymbol{u}_j), \check{\boldsymbol{u}}_j - \boldsymbol{u}_j \rangle + \frac{G^2}{2}\|\check{\boldsymbol{u}}_j - \boldsymbol{u}_j\|_2^2 \tag{E.123}$$

$$= \sum_{j=a}^{b} \left(\nabla f_j(\boldsymbol{u}_j)[k'](\check{\boldsymbol{u}}_j[k'] - \boldsymbol{u}_j[k']) + \frac{G^2}{2}(\check{\boldsymbol{u}}_j[k'] - \boldsymbol{u}_j[k'])^2 \right) \tag{E.124}$$

$$+ \sum_{\substack{k=1 \\ k\neq k'}}^{d} \sum_{j=a}^{b} \left(\nabla f_j(\boldsymbol{u}_j)[k](\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]) + \frac{G^2}{2}(\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k])^2 \right) \tag{E.125}$$

$$\leq \frac{G^2B^2}{2} + \sum_{\substack{k=1 \\ k\neq k'}}^{d} \sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k](\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]) \tag{E.126}$$

$$+ \sum_{j=a}^{b} \left(\nabla f_j(\boldsymbol{u}_j)[k'](\check{\boldsymbol{u}}_j[k'] - \boldsymbol{u}_j[k']) + \frac{G^2}{2}(\check{\boldsymbol{u}}_j[k'] - \boldsymbol{u}_j[k'])^2 \right), \tag{E.127}$$

where in the last line we have used the facts that $\sum_{\substack{k=1 \\ k\neq k'}}^{d}(\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k])^2 \leq \left(\sum_{\substack{k=1 \\ k\neq k'}}^{d} |\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]|\right)^2 \leq C_{a\to b}^2 \leq B^2/\ell$ by Property 3 of Lemma 198 and the TV constraint assumed in the premise of the current lemma.

Since the optimal solution across any coordinate is piece-wise maximally monotonic with at-most 4 pieces, by following the same arguments used in Case 1 and 2 in the proof of Lemma 199, we can write

$$\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k](\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]) = 0, \tag{E.128}$$

for any $k \neq k'$.

Recall that $\boldsymbol{u}_j[k'] = \boldsymbol{u}_a[k'] \in (-B, B)$ for all $j \in [a, b]$. Further by our construction, $\check{\boldsymbol{u}}_j[k'] = \boldsymbol{u}_a[k'] - \frac{1}{\ell G^2}\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k']$, for all $j \in [a, b]$. The key observation is to realize that $(\check{\boldsymbol{u}}_j[k'] - \boldsymbol{u}_j[k'])$ stays at a constant value for all $j \in [a, b]$. So we have

$$\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k'](\check{\boldsymbol{u}}_j[k'] - \boldsymbol{u}_j[k']) = (\check{\boldsymbol{u}}_a[k'] - \boldsymbol{u}_a[k']) \sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k'] \tag{E.129}$$

$$= \frac{-1}{\ell G^2} \left(\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k'] \right)^2. \tag{E.130}$$

Further we have,

$$\sum_{j=a}^{b} \frac{G^2}{2} (\check{\boldsymbol{u}}_j[k'] - \boldsymbol{u}_j[k'])^2 = \frac{1}{2\ell G^2} \left(\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k'] \right)^2. \tag{E.131}$$

Combining Eq.(E.130) and (E.131), we get

$$\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k'](\check{\boldsymbol{u}}_j[k'] - \boldsymbol{u}_j[k']) + \frac{G^2}{2} (\check{\boldsymbol{u}}_j[k'] - \boldsymbol{u}_j[k'])^2 = \frac{-1}{2\ell G^2} \left(\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k'] \right)^2 \tag{E.132}$$

$$=_{(a)} \frac{-1}{2\ell G^2} \left(\lambda \Delta \boldsymbol{s}_{a\to b}[k'] \right)^2 \tag{E.133}$$

$$= (b) \frac{-2\lambda^2}{\ell G^2}, \tag{E.134}$$

where line (a) is due to the KKT conditions and the fact that $\boldsymbol{u}_j[k'] \in (-B, B)$ thus making $\boldsymbol{\gamma}_j^+[k'] = \boldsymbol{\gamma}_j^-[k'] = 0$ and line (b) is due to the fact that $|\Delta \boldsymbol{s}_{a\to b}[k']| = 2$ for Structure 1 and Structure 2.

Hence overall we have shown that $T_2 \leq \frac{G^2 B^2}{2} - \frac{2\lambda^2}{\ell G^2}$. Combining with Eq.(E.122) we conclude that the total regret of the FLH-ONS strategy within the bin $[a, b]$ is bounded by

$$T_1 + T_2 \leq 70d^2(8G^2B^2\alpha d + 1/\alpha) \log n + \frac{G^2B^2}{2} - \frac{2\lambda^2}{\ell G^2} \tag{E.135}$$

$$\leq 70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha) \log n - \frac{2\lambda^2}{\ell G^2}. \tag{E.136}$$

□

Lemma 202. *Consider a bin $[a, b]$ with length ℓ where the TV of the offline optimal obeys $C_{a\to b} \leq B/\sqrt{\ell}$. Assume that for some coordinate $k' \in [d]$, $\boldsymbol{u}_{a:b}[k']$ takes the form of Structure 1 or Structure 2. Further suppose that across all coordinates, the offline optimal solution is piece-wise maximally monotonic in $[a, b]$ with at-most 2 pieces.*

Case 1: *When $\boldsymbol{u}_{a:b}[k']$ takes the form of Structure 1 and $\boldsymbol{u}_a[k'] - \frac{1}{\ell G^2} \sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k'] \geq$*

B, then

$$\sum_{j=a}^{b} f_j(\boldsymbol{x}_j) - f_j(\boldsymbol{u}_j) \leq 70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n - \frac{\ell G^2}{2}(B - \boldsymbol{u}_a[k'])^2, \tag{E.137}$$

and

Case 2: *When $\boldsymbol{u}_{a:b}[k']$ takes the form of Structure 2 and $\boldsymbol{u}_a[k'] - \frac{1}{\ell G^2}\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k'] \leq -B$, then*

$$\sum_{j=a}^{b} f_j(\boldsymbol{x}_j) - f_j(\boldsymbol{u}_j) \leq 70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n - \frac{\ell G^2}{2}(B + \boldsymbol{u}_a[k'])^2, \tag{E.138}$$

where $\boldsymbol{x}_j$ are the predictions of FLH-ONS.

Proof. We consider Case 1. The arguments for the alternate case are similar. We proceed in a similar way as in the proof of Lemma 201. Let $k_{\text{fix}} = k'$ and $u_{\text{fix}} = B$. Consider a comparator sequence $\check{\boldsymbol{u}}_{a:b} = \texttt{generateGhostSequence}(\boldsymbol{u}_{1:n}, k_{\text{fix}}, u_{\text{fix}}, [a,b])$. We use a two term regret decomposition as in Eq.(E.121). Using similar argumets as in the proof of Lemma 201, we have

$$T_1 \leq 70d^2(8G^2B^2\alpha d + 1/\alpha)\log n. \tag{E.139}$$

Bounding T_2 in a similar fashion as in the proof of Lemma 201, we have

$$T_2 \leq \frac{G^2B^2}{2} + \sum_{j=a}^{b}\left(\nabla f_j(\boldsymbol{u}_j)[k'](\check{\boldsymbol{u}}_j[k'] - \boldsymbol{u}_j[k']) + \frac{G^2}{2}(\check{\boldsymbol{u}}_j[k'] - \boldsymbol{u}_j[k'])^2\right), \tag{E.140}$$

where we have used Eq.(E.128) for bounding the cross terms for coordinates $k \neq k'$

The main difference is in how we handle the last term of Eq.(E.140). Recall that $\check{\boldsymbol{u}}_j[k'] = B$ and $\boldsymbol{u}_j[k'] = \boldsymbol{u}_a[k']$ for all $j \in [a,b]$. So

$$\sum_{j=a}^{b}(\nabla f_j(\boldsymbol{u}_j)[k'](\check{\boldsymbol{u}}_j[k'] - \boldsymbol{u}_j[k']) + \frac{G^2}{2}(\check{\boldsymbol{u}}_j[k'] - \boldsymbol{u}_j[k'])^2 \tag{E.141}$$

$$= \frac{G^2\ell}{2}(B - \boldsymbol{u}_a[k'])^2 - 2\lambda(B - \boldsymbol{u}_a[k']), \tag{E.142}$$

where the last line is obtained via the KKT conditions and the fact that $\Delta \boldsymbol{s}_{a\to b}[k'] = -2$ for Case 1. (Recall that $|\boldsymbol{u}_a[k']| < B$ by the definition of Structure 1. So by complementary slackness $\boldsymbol{\gamma}_j^+[k'] = \gamma_j^-[k'] = 0$.)

By the premise of the lemma for Case 1, we have $\boldsymbol{u}_a[k'] - \frac{1}{\ell G^2}\sum_{j=a}^{b}\nabla f_j(\boldsymbol{u}_j)[k'] \geq B$.

Again by using the KKT conditions and noting that $\Delta s_{a\to b}[k'] = -2$, we conclude that

$$\lambda \geq \frac{(B - \boldsymbol{u}_a[k'])\ell G^2}{2}. \tag{E.143}$$

Plugging this lower bound for λ to Eq.(E.142) and noting that $(B - \boldsymbol{u}_a[k']) \geq 0$, we get

$$\sum_{j=a}^{b} \left(\nabla f_j(\boldsymbol{u}_j)[k'](\check{\boldsymbol{u}}_j[k'] - \boldsymbol{u}_j[k']) + \frac{G^2}{2}(\check{\boldsymbol{u}}_j[k'] - \boldsymbol{u}_j[k'])^2 \right) \leq \frac{-\ell G^2}{2}(B - \boldsymbol{u}_a[k'])^2. \tag{E.144}$$

Hence overall, we conclude that

$$T_1 + T_2 \leq 70d^2(8G^2B^2\alpha d + 1/\alpha)\log n + \frac{G^2B^2}{2} - \frac{\ell G^2}{2}(B - \boldsymbol{u}_a[k'])^2 \tag{E.145}$$

$$\leq 70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n - \frac{\ell G^2}{2}(B - \boldsymbol{u}_a[k'])^2. \tag{E.146}$$

□

Lemma 203. *Suppose* ***fineSplit*** *is invoked with input $[r, s]$ such that $C_{r\to s} \leq B/\sqrt{s-r+1}$. The offline optimal solution within any bin $[a, b] \in \mathcal{Q}$ at Step 7 of* ***fineSplit*** *procedure in Fig.E.4 is piece-wise maximally monotonic in $[a, b]$ with at-most 4 pieces across any coordinate $k \in [d]$. Further there exists a coordinate $k \in [d]$ that satisfy one of the following conditions:*

1. *The offline optimal within bin $[a, b]$ takes the form of Structure 1 across coordinate k and $B - \boldsymbol{u}_a[k] \geq GAP_{min}(B, [r, s])[k] \geq GAP_{min}(-B, [r, s])[k]$.*
2. *The offline optimal within bin $[a, b]$ takes the form of Structure 2 across coordinate k and $B + \boldsymbol{u}_a[k] \geq GAP_{min}(-B, [r, s])[k] \geq GAP_{min}(B, [r, s])[k]$.*

Proof. We start by a basic observation.

FACT 1: Note that $C_{r\to s} \leq B/\sqrt{s-r+1} \leq B$. So the $\boldsymbol{u}_{r:s}[k']$ cannot touch both B and $-B$ boundaries.

Consider a bin $[a, b] \in \mathcal{Q}$. By the construction of `fineSplit`, there exists a coordinate $k \in [d]$ across which the optimal solution stays constant within $[a, b]$ and assumes the form of Structure 1 or 2. For the sake of contradiction, let's assume that for some $k' \in [d]$, with $k' \neq k$, the optimal solution is maximally monotonic in $[a, b]$ with at-least 5 pieces across the coordinate k'. This can happen only when the optimal solution increases (decreases) then decreases (increases) then increases (decreases) then decreases (increases) and finally increase (decrease) again within bin $[a, b]$ and evolve arbitrarily there on-wards. Combined with FACT 1, such a behaviour can result in one of the following configurations across the coordinate k':

`fineSplit`: Input - (1) offline optimal sequence $\boldsymbol{u}_{1:n}$ (2) an interval $[r, s] \subseteq [n]$. Across some coordinate $k \in [d]$, the offline optimal solution must take the form of both Structure 1 and 2 or either one of them at-least two times within some appropriate sub-intervals of $[r, s]$.

1. Initialize $\mathcal{Q} \leftarrow \Phi$, $\mathcal{Q}' \leftarrow \Phi$.
2. For each coordinate $k \in [d]$ across which the optimal solution takes the form of Structures 1 and 2 or either one of them at-least two times within some appropriate sub-intervals of $[r, s]$:
   (a) if $\text{GAP}_{\min}(B, [r, s])[k] > \text{GAP}_{\min}(-B, [r, s])[k]$ then add intervals $[a, b] \subset [r, s]$ where the offline optimal across coordinate k assumes the form of Structure 1 to $\mathcal{Q}$.
   (b) if $\text{GAP}_{\min}(B, [r, s])[k] \leq \text{GAP}_{\min}(-B, [r, s])[k]$ then add intervals $[a, b] \subset [r, s]$ where the offline optimal across coordinate k assumes the form of Structure 2 to $\mathcal{Q}$.
3. For each bin $[a, b] \in \mathcal{Q}$ if there exists another interval $[p, q] \in \mathcal{Q}$ with $[p, q] \subseteq [a, b]$, then remove $[a, b]$ from $\mathcal{Q}$.
4. Sort intervals in $\mathcal{Q}$ in increasing order of the left endpoints. (i.e $[a, b] < [p, q]$ if $a < p$).
5. Starting from the first bin, for each bin $[a, b] \in \mathcal{Q}$:
   (a) if there exists an interval $[p, q] \in \mathcal{Q}$ such that $a < p$ and $b < q$, then remove $[p, q]$ from $\mathcal{Q}$
6. Add disjoint and maximally continuous intervals that are the subsets of $[r, s] \backslash \{\cup_{[a,b] \in \mathcal{Q}} [a, b]\}$ to $\mathcal{Q}'$ such that the interval $[r, s]$ can be fully covered by disjoint intervals from $\mathcal{Q}$ and $\mathcal{Q}'$.
7. Return $(\mathcal{Q}, \mathcal{Q}')$.

Figure E.4: ***fineSplit*** *procedure.*

- Both Structure 1 and Structure 2 are formed.
- Only Structure 2 is formed at-least two times. This means that if $[x, y] \subset [a, b]$ is a maximally monotonic section with $\boldsymbol{u}_{x:y}[k']$ increasing, then $\boldsymbol{u}_y[k'] = B$. Then $\text{GAP}_{\min}(-B, [r, s])[k'] > \text{GAP}_{\min}(B, [r, s])[k'] = 0$.
- Only Structure 1 is formed at-least two times. This means that if $[x, y] \subset [a, b]$ is

a maximally monotonic section with $\boldsymbol{u}_{x:y}[k']$ decreasing, then $\boldsymbol{u}_y[k'] = -B$. Then $\mathrm{GAP}_{\min}(B, [r,s])[k'] > \mathrm{GAP}_{\min}(-B, [r,s])[k'] = 0$.

In all of the above cases, at-least one sub-interval of $[a,b]$ will be added to $\mathcal{Q}$ at Step 2(a) or 2(b). This would imply that at Step 3, the bin $[a,b]$ is removed from $\mathcal{Q}$ and never added again resulting in a contradiction.

The last statement of the Lemma is immediate from Steps 2(a)-(b) of `fineSplit`. □

Lemma 204. *Suppose* ***fineSplit*** *is invoked with input $[r,s]$ such that $C_{r\to s} \le B/\sqrt{s-r+1}$. The offline optimal solution within any interval $[p,q] \in \mathcal{Q}'$ at Step 7 of* ***fineSplit*** *procedure in Fig.E.4 is piece-wise maximally monotonic in $[p,q]$ with at-most 4 pieces across any coordinate.*

Proof. Consider a coordinate $k \in [d]$ and a bin $[p,q] \in \mathcal{Q}'$. We provide the arguments for the case when $\mathrm{GAP}_{\min}(-B, [r,s])[k] \ge \mathrm{GAP}_{\min}(B, [r,s])[k]$. The arguments for the complementary case are similar. We start by stating two facts.

FACT 1: $\mathrm{GAP}_{\min}(-B, [p,q])[k] > 0$.

To see this, assume for the sake of contradiction that $\mathrm{GAP}_{\min}(-B, [p,q])[k] = 0$. Then this means that $\mathrm{GAP}_{\min}(-B, [p,q])[k] = \mathrm{GAP}_{\min}(B, [r,s])[k] = 0$. So the optimal solution across coordinate k, $\boldsymbol{u}_{r:s}[k]$ must touch both B and $-B$ at distinct time points in $[r,s]$. This would violate the TV constraint that $C_{p\to q} \le B/\sqrt{s-r+1} \le B$, thus yielding a contradiction.

FACT2: It is not the case that there exists two intervals $[p_1,q_1],[p_2,q_2] \subset [p,q]$ within which the offline optimal takes the form of Structure 2 across the coordinate $k \in [d]$.

Let's prove the above fact via contradiction. Assume that there exists $[p_1,q_1],[p_2,q_2] \subset [p,q] \in \mathcal{Q}'$ such that the offline optimal takes the form of Structure 2 within them across the coordinate $k \in [d]$. Then $[p_i,q_i]$ $(i=1,2)$ must have been added to $\mathcal{Q}$ in step 2(b) of `fineSplit`. Since intervals in $\mathcal{Q}$ don't overlap with intervals in $\mathcal{Q}'$ due to Step 6, this would mean that the interval $[p_i,q_i]$ $(i=1,2)$ got removed from $\mathcal{Q}$ later.

Case 1: Consider the case where $[p_i,q_i]$ $(i=1,2)$ has been removed at Step 5(a). This means that there exists an interval $[a,b] \subseteq [r,s]$ where the offline optimal has Structure 1 or 2 across some coordinate $k' \ne k$ and $[p_i,q_i] \cap [a,b] \ne \Phi$. Observe that $[a,b]$ is never removed from $\mathcal{Q}$ since we are processing bins in sorted order at Step 4-5. This would contradict the fact that intervals in $\mathcal{Q}$ don't overlap with intervals in $\mathcal{Q}'$ due to Step 6.

Case 2: Consider the case where $[p_i,q_i]$ $(i=1,2)$ has been removed at Step 3. This means that there exists an interval $[x,y] \subseteq [p_i,q_i]$ where the offline optimal assumes Structure 1 or 2 across some coordinate $k' \ne k$. If $[x,y]$ is present in the final $\mathcal{Q}$ in Step 7, then this would again warrant a contradiction to the non-overlapping property between the intervals of $\mathcal{Q}$ and $\mathcal{Q}'$. If $[x,y]$ is removed at a later point through Step 5(a), by using similar arguments as in Case 1 yields a contradiction. Thus we conclude that the FACT 2 is true.

FACT 3: It is not the case that there exists two intervals $[p_1,q_1],[p_2,q_2] \subset [p,q]$ within the offline optimal takes the form of Structure 1 in $[p_1,q_1]$ and Structure 2 in $[p_2,q_2]$ across the coordinate $k \in [d]$.

The above fact can be proven using similar arguments that are used in proving FACT 2.

In light of FACT 1, FACT 2 and FACT 3, we conclude the statement of the lemma. □

Next we introduce a structural lemma analogous to Lemma 183.

Lemma 205. *(**λ-length lemma**) Consider a bin $[a,b] \subseteq \{2,\ldots,n-1\}$ with length ℓ. Suppose that within this bin, the offline optimal solution sequence assumes the form of Structure 1 or Structure 2 across some coordinate $k \in [d]$, then $\lambda \leq \frac{G_\infty \ell}{2}$, where G_∞ is as in Assumption B2.*

Proof Sketch. The arguments for this proof are almost identical to that used for proving Lemma 183. We outline the parts where there are differences. We provide the arguments for Structure 2. Structure 1 can be handled similarly. Let the optimal sign assignments across coordinate k be written as $\boldsymbol{s}_j[k] = -1 + \epsilon_j$ where $\epsilon_j \in [0,2]$ and $j \in [a,b]$. From the KKT conditions, we can write:

$$\nabla f_a(\boldsymbol{u}_a)[k] = \lambda \epsilon_a \tag{E.147}$$

$$\nabla f_{a+1}(\boldsymbol{u}_{a+1})[k] = \lambda(\epsilon_{a+1} - \epsilon_a) \tag{E.148}$$

$$\vdots \tag{E.149}$$

$$\nabla f_{b-1}(\boldsymbol{u}_{b-1})[k] = \lambda(\epsilon_{b-1} - \epsilon_{b-1}) \tag{E.150}$$

$$\nabla f_b(\boldsymbol{u}_b)[k] = \lambda(2 - \epsilon_{b-1}) \tag{E.151}$$

Define the vector $\boldsymbol{z} = [\epsilon_a, \epsilon_{a+1} - \epsilon_a, \ldots, 2 - \epsilon_{b-1}]^T$. As noted in the proof of Lemma 183, we must have $\|\boldsymbol{z}\|_\infty > 0$. Let j^* be such that $\|\boldsymbol{z}\|_\infty = |\boldsymbol{z}[j^*]|$. Then $\lambda = \nabla f_{a+j^*-1}(\boldsymbol{u}_{a+j^*-1})[k]/\|\boldsymbol{z}\|_\infty$. From the optimization problem considered in the proof of Lemma 183, we have $\|\boldsymbol{z}\|_\infty \geq 2/\ell$. Since $\|\nabla f_j(\boldsymbol{u}_j)\|_\infty \leq G_\infty$ for all $j \in [n]$ by Assumption B2, we have $\lambda = \nabla f_{a+j^*-1}(\boldsymbol{u}_{a+j^*-1})[k]/\|\boldsymbol{z}\|_\infty \leq (G_\infty \ell)/2$.

□

Lemma 206. *(**large margin bins**) Assume that $\lambda \geq d^{1.5}\phi \frac{n^{1/3}}{C_n^{1/3}}$ for a constant $\phi = \sqrt{70(8G^2B^2\alpha + G^2B^2 + 1/\alpha)}$ that does not depend on n and C_n. Consider a bin $[i_s, i_t] \in \mathcal{P}$ within which the offline optimal solution takes the form of Structure 1 or Structure 2 (or both) across a coordinate $k \in [d]$ for some appropriate sub-intervals of $[i_s, i_t]$. Let $\mu_{th} = \sqrt{\frac{140 d^{1.5}(8G^2B^2\alpha + G^2B^2 + 1/\alpha) G_\infty C_n^{1/3} \log n}{G^2 \phi n^{1/3}}}$. Then*
$GAP_{min}(-B, [i_s, i_t])[k] \vee GAP_{min}(B, [i_s, i_t])[k] \geq \mu_{th}$,
whenever $C_n \leq \left(\frac{B^2 G^2 \phi}{560 d^{1.5}(8G^2B^2\alpha + G^2B^2 + 1/\alpha) G_\infty \log n}\right)^3 n = \tilde{O}(n)$.

Proof. Suppose $\text{GAP}_{\min}(-B, [i_s, i_t])[k] < \mu_{\text{th}}$. Then the largest value of the optimal solution across coordinate k attained within this bin $[i_s, i_t]$ is at-most $-B + \mu_{\text{th}} + B/\sqrt{n_i}$

(recall $n_i := i_t - i_s + 1$ and $C_i \leq B/\sqrt{n_i}$ due to Lemma 193). So $\text{GAP}_{\min}(B, [i_s, i_t])[k] \geq 2B - \mu_{\text{th}} - B/\sqrt{n_i}$. Our goal is to show that whenever C_n obeys the constraint stated in the lemma, we must have

$$2B - \mu_{\text{th}} - B/\sqrt{n_i} \geq \mu_{\text{th}}. \tag{E.152}$$

Let ℓ_i be the length of a sub-interval of $[i_s, i_t]$ where the offline optimal solution assumes the form of Structure 1 or Structure 2. Due to Lemma 205, we have

$$n_i \geq \ell_i \geq \frac{2\lambda}{G_\infty} \geq \frac{2d^{1.5}\phi n^{1/3}}{G_\infty C_n^{1/3}} \tag{E.153}$$

where the last inequality follows due to the condition assumed in the current lemma. So a sufficient condition for Eq.(E.152) to be true is

$$2B \geq 2\left(2\sqrt{\frac{140d^{1.5}(8G^2B^2\alpha + G^2B^2 + 1/\alpha)G_\infty C_n^{1/3}\log n}{G^2\phi n^{1/3}}} \vee B\sqrt{\frac{G_\infty C_n^{1/3}}{2d^{2.5/2}\phi n^{1/3}}}\right). \tag{E.154}$$

Recall that by Assumption B2, we have $G \wedge G_\infty \wedge B \geq 1$. So the above maximum will be attained by the first term and can be further simplified as

$$2B \geq 4\sqrt{\frac{140d^{1.5}(8G^2B^2\alpha + G^2B^2 + 1/\alpha)G_\infty C_n^{1/3}\log n}{G^2\phi n^{1/3}}}. \tag{E.155}$$

The above condition is always satisfied whenever $C_n \leq \left(\frac{B^2G^2\phi}{560d^{1.5}(8G^2B^2\alpha+G^2B^2+1/\alpha)G_\infty\log n}\right)^3 n$.

At this point, we have shown that $\text{GAP}_{\min}(-B, [i_s, i_t])[k] < \mu_{\text{th}} \implies \text{GAP}_{\min}(B, [i_s, i_t])[k] \geq \mu_{\text{th}}$ under the conditions of the lemma. Taking the contrapositive yields $\text{GAP}_{\min}(B, [i_s, i_t])[k] < \mu_{\text{th}} \implies \text{GAP}_{\min}(-B, [i_s, i_t])[k] \geq \mu_{\text{th}}$. □

Lemma 207. *(**high λ regime**) Suppose the optimal dual variable $\lambda \geq d^{1.5}\phi\frac{n^{1/3}}{C_n^{1/3}} = \Omega\left(\frac{n^{1/3}}{C_n^{1/3}}\right)$ for $\phi = \sqrt{70(8G^2B^2\alpha + G^2B^2 + 1/\alpha)}$ that does not depend on n and C_n. We have the regret of FLH-ONS strategy bounded as*

$$\sum_{t=1}^{n} f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{u}_t) = \tilde{O}\left(140d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)(n^{1/3}C_n^{2/3} \vee 1)\right)\mathbb{I}\{C_n > 1/n\} \tag{E.156}$$

$$+ \tilde{O}\left(d(8G^2B^2\alpha d + 1/\alpha)\,\mathbb{I}\{C_n \leq 1/n\}\right), \tag{E.157}$$

where $\boldsymbol{x}_t$ is the prediction of FLH-ONS at time t and $\mathbb{I}\{\cdot\}$ is the boolean indicator function taking values in $\{0,1\}$.

Proof. Throughout the proof we assume that $C_n \left(\frac{B^2G^2\phi}{560d^{1.5}(8G^2B^2\alpha+G^2B^2+1/\alpha)G_\infty \log n}\right)^3 n$. Otherwise the trivial regret bound of $\tilde{O}(n)$ is near minimax optimal.

First we consider the regime where $C_n \geq 1/n$. It is useful to define the following annotated condition.

Condition (A): Let a bin $[r,s]$ be given. For some coordinate $k' \in [d]$, there exists disjoint intervals $[r_1,s_1],[r_2,s_2] \subset [r,s]$ that satisfy at-least one of the following: (i) $\boldsymbol{u}_{r_1:s_1}[k']$ has the form of Structure 1 and $\boldsymbol{u}_{r_2:s_2}[k']$ has the form of Structure 2; (ii) Both $\boldsymbol{u}_{r_1:s_1}[k']$ and $\boldsymbol{u}_{r_2:s_2}[k']$ have the form of Structure 1; (iii) Both $\boldsymbol{u}_{r_1:s_1}[k']$ and $\boldsymbol{u}_{r_2:s_2}[k']$ have the form of Structure 2.

The above condition is basically the prerequisite for the `fineSplit` procedure of Fig.E.4.

Let $[i_s,i_t] \in \mathcal{P}$ be a bin that satisfy Condition (A) for a coordinate $k' \in [d]$. Here $\mathcal{P}$ is the partition obtained in Lemma 193.

Let $(\mathcal{Q},\mathcal{Q}')$ be the collections of intervals obtained by invoking the `fineSplit` procedure with the bin $[i_s,i_t]$ as input. Let's write $\mathcal{Q}\cup\mathcal{Q}'\cup\{\Phi\}$ as a collection of disjoint consecutive intervals as follows:

$$\mathcal{Q}\cup\mathcal{Q}'\cup\{\Phi\} := \{[i_s,\underline{i}_1-1],[\underline{i}_1,\bar{i}_1],[\underline{i}'_1,\bar{i}'_1],\ldots,[\underline{i}_{m^{(i)}},\bar{i}_{m^{(i)}}],[\underline{i}'_{m^{(i)}},\bar{i}'_{m^{(i)}}]\}, \tag{E.158}$$

with $\bar{i}'_{m^{(i)}} = i_t$.

Here we follow the convention that the bins $[\underline{i}_p,\bar{i}_p] \in \mathcal{Q}$ and $[\underline{i}'_p,\bar{i}'_p] \in \mathcal{Q}'\cup\{\Phi\}$ for all $p \in [m^{(i)}]$. Similar to the proof of Lemma 191, for enforcing this convention, we may have to set either of the bins $[i_s,\underline{i}_1-1]$ or $[\underline{i}'_{m^{(i)}},\bar{i}'_{m^{(i)}}]$ to be empty. More precisely, if i_s belongs to some interval in $\mathcal{Q}$, then we set the first sub-interval $[i_s,\bar{i}_1-1]$ to be empty by setting $\bar{i}_1 = i_s$. Similarly, if i_t belongs to some interval in $\mathcal{Q}$, we treat the sub-interval $[\underline{i}'_{m^{(i)}},\bar{i}'_{m^{(i)}}]$ as empty by setting $\underline{i}'_{m^{(i)}} = i_t+1$. Further some of the intervals: $[\underline{i}'_k,\bar{i}'_k]$, $k \in [m^{(i)}]$ can be empty. For example if $\underline{i}_{k+1} = \bar{i}_k+1$, then $[\underline{i}'_k,\bar{i}'_k]$ is treated as empty.

Note that if the first sub-interval $[i_s,\underline{i}_1-1]$ is non-empty then it must belong to $\mathcal{Q}'$ according to our convention. By Lemma 204 and Lemma 199,

$$\sum_{j=i_s}^{\underline{i}_1-1} f_j(\boldsymbol{x}_j) - f_j(\boldsymbol{u}_j) = \tilde{O}\left(70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n\right). \tag{E.159}$$

We proceed to bound the regret in $[\underline{i}_1,\bar{i}'_{m^{(i)}}]$. Let $\mathcal{P}_1^{(i)}$ denote the collection of bins among $\mathcal{Q} = \{[\underline{i}_1,\bar{i}_1],\ldots,[\underline{i}_{m^{(i)}},\bar{i}_{m^{(i)}}]\}$ which satisfy the property in Lemma 201. Let $|\mathcal{P}_1^{(i)}| := m_1^{(i)}$ and their lengths be denoted by $\{\ell^{(1)}_{1^{(i)}},\ldots,\ell^{(1)}_{m_1^{(i)}}\}$. These bins will be referred as *Type 1* bins henceforth.

Similarly let $\mathcal{P}_2^{(i)} = \mathcal{Q}\setminus\mathcal{P}_1^{(i)}$ which satisfy either of the properties in Lemma 202.

Let $|\mathcal{P}_2^{(i)}| := m_2^{(i)}$ and their lengths be denoted by $\{\ell^{(2)}_{1^{(i)}}, \ldots, \ell^{(2)}_{m_2^{(i)}}\}$. These bins will be referred as *Type 2* bins henceforth. A bin $[a,b] \in \mathcal{P}_2^{(i)}$ satisfy at-least one of the following properties

P1: For some coordinate $k \in [d]$, the offline optimal satisfy the condition of Case 1 in Lemma 202 and $B - \boldsymbol{u}_a[k] \geq \mu_{th}$.

P2: For some coordinate $k \in [d]$, the offline optimal satisfy the condition of Case 2 in Lemma 202 and $B + \boldsymbol{u}_a[k] \geq \mu_{th}$.

To see this, let's inspect the way in which the bin $[a,b]$ has been added to $\mathcal{Q}$ when we invoke `fineSplit` with the input bin $[i_s, i_t]$. If $[a,b]$ has been added via Step 2-(a), then we have $\text{GAP}_{\min}(B, [i_s, i_t])[k] > \text{GAP}_{\min}(-B, [i_s, i_t])[k]$ for a coordinate k. By Lemma 206 it holds that $\text{GAP}_{\min}(B, [i_s, i_t])[k] \geq \mu_{\text{th}}$ under the C_n regime we consider. So $B - \boldsymbol{u}_a[k] \geq \text{GAP}_{\min}(B, [i_s, i_t])[k] \geq \mu_{\text{th}}$ where the first inequality follows by the definition of GAP (see Definition 200). Further, observe that $\boldsymbol{u}_a[k'] - \frac{1}{\ell G^2}\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k'] < -B$ is never satisfied, where $\ell = b - a + 1$. Otherwise it will imply that $-\frac{1}{\ell G^2}\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k'] = \frac{2\lambda}{\ell G^2} < -B - \boldsymbol{u}_a[k'] \leq 0$ which is not true as $\lambda \geq 0$. We must also have $\boldsymbol{u}_a[k'] - \frac{1}{\ell G^2}\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k'] \notin [-B, B]$. Otherwise, bin $[a,b]$ would have been already added to $\mathcal{P}_1^{(i)}$ and would have never present in $\mathcal{P}_2^{(i)}$. So we conclude that property P1 follows. Property P2 can also be shown to be true using similar arguments when the bin $[a,b]$ has been added to $\mathcal{Q}$ via Step 2-(b) of `fineSplit`.

Each bin $[\underline{i}_k, \bar{i}_k]$, $k \in [m^{(i)}]$ of Type 1 and Type 2 can be paired with an adjacent bin $[\underline{i}'_k, \bar{i}'_k] \in \mathcal{Q}' \cup \{\Phi\}$, $k \in [m^{(i)}]$ which is either empty or the optimal sequence displays a piece-wise maximally monotonic behaviour in $[\underline{i}'_k, \bar{i}'_k]$ across all coordinates as recorded in Lemma 204.

Note that $m_1^{(i)} + m_2^{(i)} = m^{(i)}$. Let the total regret contribution from Type 1 bins along with their pairs and Type 2 bins along with their pairs be referred as $R_1^{(i)}$ and $R_2^{(i)}$ respectively.

For a bin $[a,b] \in \mathcal{P}_2^{(i)}$, in either of the cases covered by the properties P1 and P2, we have by Lemma 202 that

$$\sum_{j=a}^{b} f_j(\boldsymbol{x}_j) - f_j(\boldsymbol{u}_j) \leq 70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n - \frac{\ell G^2}{2}\mu_{\text{th}}^2, \tag{E.160}$$

Let $[a', b'] \in \mathcal{Q}' \cup \{\Phi\}$ be the pair assigned to $[a,b]$. If it is non-empty, then due to Lemma 204 and Lemma 199 the regret from the bin $[a', b']$ is at-most $70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n$.

So we can bound $R_2^{(i)}$ as

$$R_2^{(i)} \leq \sum_{j=1^{(i)}}^{m_2^{(i)}} \left(\left(70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n - \frac{\ell_j^{(2)}G^2}{2}\mu_{\text{th}}^2 \right) \right. \tag{E.161}$$

$$\left. + 70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n \right) \tag{E.162}$$

$$\leq m_2^{(i)} 140d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n - \frac{G^2\mu_{\text{th}}^2}{2}\left(\sum_{j=1^{(i)}}^{m_2^{(i)}} \ell_j^{(2)}\right), \tag{E.163}$$

From Eq.(E.153), we have $\ell_j^{(2)} \geq \frac{2\phi d^{1.5} n^{1/3}}{G_\infty C_n^{1/3}}$ for $j \in \{1^{(i)}, \ldots, m_2^{(i)}\}$ under the regime of λ we consider. So we can continue as

$$R_2^{(i)} \leq 140 m_2^{(i)} d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n - G^2\mu_{\text{th}}^2 m_2^{(i)} \frac{\phi d^{1.5} n^{1/3}}{G_\infty C_n^{1/3}} \tag{E.164}$$

$$\leq 140 m_2^{(i)} d^3(8G^2B^2\alpha + G^2B^2 + 1/\alpha)\log n - G^2\mu_{\text{th}}^2 m_2^{(i)} \frac{\phi d^{1.5} n^{1/3}}{G_\infty C_n^{1/3}} \tag{E.165}$$

$$= 0, \tag{E.166}$$

where the last line is obtained by plugging in the value of μ_{th} as in Lemma 206.

So by refining every interval in $\mathcal{P}$ (recall that $\mathcal{P}$ is from Lemma 193) that satisfy Condition (A) and summing the regret contribution from all Type 2 bins and their pairs across all refined intervals in $\mathcal{P}$ yields

$$\sum_{i=1}^{M} R_2^{(i)} \leq 0, \tag{E.167}$$

where we recall that $M := |\mathcal{P}| = O(n^{1/3}C_n^{2/3} \vee 1)$ and assign $R_2(i) = 0$ for intervals in $\mathcal{P}$ that do not satisfy Condition (A).

For any Type 1 bin, its regret contribution can be bounded by Lemma 201. The regret contribution from its pair can be bounded by Lemma 199 as before. So we have

$$R_1^{(i)} \leq \sum_{j=1}^{m_1^{(i)}} \left(\left(70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n - \frac{2\lambda^2}{\ell_{j^{(i)}}^{(1)} G^2} \right) \right. \tag{E.168}$$

$$\left. + 70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n \right) \tag{E.169}$$

$$= 140 m_1^{(i)} d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n - \frac{2\lambda^2}{G^2}\sum_{j=1}^{m_1^{(i)}} \frac{1}{\ell_{j^{(i)}}^{(1)}}. \tag{E.170}$$

So by refining every interval in $\mathcal{P}$ that satisfy Condition (A) and summing the regret contribution from all Type 2 bins and their pairs across all refined intervals in $\mathcal{P}$ yields

$$\sum_{i=1}^{M} R_1^{(i)} \leq 140d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n \sum_{i=1}^{M} m_1^{(i)} - \frac{2\lambda^2}{G^2}\sum_{i=1}^{M}\sum_{j=1}^{m_1^{(i)}} \frac{1}{\ell_{j^{(i)}}^{(1)}} \tag{E.171}$$

$$\leq 140d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)M_1 \log n - \frac{2\lambda^2}{G^2}\frac{M_1^2}{n}, \tag{E.172}$$

$$\leq 140d^3(8G^2B^2\alpha + G^2B^2 + 1/\alpha)M_1 \log n - \frac{2\lambda^2}{G^2}\frac{M_1^2}{n} \tag{E.173}$$

where in the last line: a) we define $M_1 := \sum_{i=1}^{M} m_1^{(i)}$ with the convention that $m_1^{(i)} = 0$ if the i^{th} bin in $\mathcal{P}$ doesn't satisfy Condition (A); b) applied AM-HM inequality and noted that $\sum_{i=1}^{M}\sum_{j=1}^{m_1^{(i)}} \ell_{j^{(i)}}^{(1)} \leq n$.

To further bound Eq.(E.173), we consider two separate regimes as follows.

Recall that $\lambda \geq d^{1.5}\phi\frac{n^{1/3}}{C_n^{1/3}}$. So continuing from Eq.(E.173),

$$140d^3(8G^2B^2\alpha + G^2B^2 + 1/\alpha)M_1 \log n - 2\lambda^2\frac{M_1^2}{G^2 n} \leq 140d^3(8G^2B^2\alpha + G^2B^2 + 1/\alpha)\times \tag{E.174}$$

$$M_1 \log n - 2d^{2.5}\phi^2\frac{n^{2/3}}{C_n^{2/3}}\frac{M_1^2}{G^2 n} \tag{E.175}$$

$$\leq 0, \tag{E.176}$$

whenever $M_1 \geq \frac{70(8G^2B^2\alpha + G^2B^2 + 1/\alpha)\log n}{\phi^2} n^{1/3}C_n^{2/3} = \tilde{\Omega}(n^{1/3}C_n^{2/3})$.

In the alternate regime where $M_1 \leq \left(\frac{70(8G^2B^2\alpha + G^2B^2 + 1/\alpha)\log n}{\phi^2} n^{1/3}C_n^{2/3} \vee 1\right)$ $= \tilde{O}(n^{1/3}C_n^{2/3} \vee 1)$, we trivially obtain $\sum_{i=1}^{M} R_1^{(i)}$ $= \tilde{O}\left(140d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)(n^{1/3}C_n^{2/3} \vee 1)\right)$.

The regret contribution from all sub-bins that starts at i_s $i \in [M]$ which are not paired in Eq.(E.158) is only at-most $\tilde{O}(d^{2.5}(n^{1/3}C_n^{2/3} \vee 1))$ by adding the bound of Eq.(E.159) across all $O(n^{1/3}C_n^{2/3} \vee 1)$ bins in $\mathcal{P}$.

Throughout the entire proof we have assumed that $m_1^{(i)}$ and $m_2^{(i)}$ are non-zero for some bin $[i_s, i_t] \in \mathcal{P}$. Not meeting this criterion will only make the arguments easier as explained below.

We have shown that the total regret contribution from the refined bins $\sum_{i=1}^{M} R_1^{(i)} + R_2^{(i)} = \tilde{O}(n^{1/3}C_n^{2/3} \vee 1)$, we trivially obtain $\sum_{i=1}^{M} R_1^{(i)} = \tilde{O}\left(140d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)(n^{1/3}C_n^{2/3} \vee 1)\right)$ under the conditions of the lemma, where we have taken $R_1^{(i)} = R_2^{(i)} = 0$ if the i^{th} bin $[i_s, i_t] \in \mathcal{P}$ doesn't satisfy Condition (A) across any coordinate.

If a bin doesn't satisfy Condition (A) across any coordinate, then the offline optimal solution within that bin assumes a piece-wise maximally monotonic structure with at-most 4 pieces across any coordinate. By Lemma 199, the regret within such bins is $\tilde{O}\left(70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\right)$. Since there can be at-most $O(n^{1/3}C_n^{2/3} \vee 1)$ such bins in $\mathcal{P}$, the total regret contribution from those bins is again $\tilde{O}\left(70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)(n^{1/3}C_n^{2/3} \vee 1)\right)$. Now putting everything together yields the lemma.

If $C_n \leq 1/n$, then we have

$$\sum_{t=1}^{n} f_t(\boldsymbol{x}_j) - f_t(\boldsymbol{u}_t) \leq \sum_{t=1}^{n} f_t(\boldsymbol{x}_j) - f_t(\boldsymbol{u}_1) + \sum_{t=1}^{n} f_t(\boldsymbol{u}_1) - f_t(\boldsymbol{u}_t) \tag{E.177}$$

$$\leq_{(a)} \tilde{O}\left(10d(8G^2B^2\alpha d + 1/\alpha)\log n\right) + GnC_n \tag{E.178}$$

$$= \tilde{O}\left(d(8G^2B^2\alpha d + 1/\alpha\right) \tag{E.179}$$

where line (a) follows from the fact that f_t is G Lipschitz. □

Proof. **of Theorem** 59. The proof is immediate from the results of Lemmas 196 and 207.

□

E.4 Reparametrization of certain polytopes to box

Proposition 208. *Consider an online problem with losses f_t that are α exp-concave on the decision set $\mathcal{D} = \{\boldsymbol{x} \in \mathbb{R}^d : \boldsymbol{c} \leq \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}\}$ such that $\boldsymbol{A}$ is full rank and $\boldsymbol{0} < \boldsymbol{b} - \boldsymbol{c}$.*

We can reparametrize this into an equivalent online learning problem with losses $\tilde{f}_t(\boldsymbol{z}) = f_t(\boldsymbol{A}^{-1}(\boldsymbol{D}^{-1}(\boldsymbol{z} + \boldsymbol{1}) + \boldsymbol{c}))$ that are α exp-concave on the decision set $\tilde{\mathcal{D}} = \{\ \boldsymbol{z} \in \mathbb{R}^d : \|\boldsymbol{z}\|_\infty \leq 1\}$, where $\boldsymbol{D} = \mathrm{diag}(2/(\boldsymbol{b}[1] = \boldsymbol{c}[1]), \ldots, 2/(\boldsymbol{b}[d] - \boldsymbol{c}[d]))$ and $\boldsymbol{1}$ is the vector of ones in $\mathbb{R}^d$.

Further if the losses f_t are G Lipschitz in $\mathcal{D}$, then the losses $\tilde{f}_t$ are $\|\boldsymbol{A}^{-1}\boldsymbol{D}^{-1}\|_{op}G$ Lispchitz in $\tilde{\mathcal{D}}$.

Proof. We have,

$$\boldsymbol{c} \leq \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b} \tag{E.180}$$

$$\iff \boldsymbol{0} \leq \boldsymbol{A}\boldsymbol{x} - \boldsymbol{c} \leq \boldsymbol{b} - \boldsymbol{c}. \tag{E.181}$$

Then we have $\boldsymbol{0} \leq \boldsymbol{D}(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{c}) \leq (2)\boldsymbol{1}$. This equivalent to $-\boldsymbol{1} \leq \boldsymbol{D}(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{c}) - \boldsymbol{1} \leq \boldsymbol{1}$. By putting $\boldsymbol{z} = \boldsymbol{D}(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{c}) - \boldsymbol{1}$ we can rewrite the original decision set as $\|\boldsymbol{z}\|_\infty \leq 1$.

Since $\boldsymbol{A}$ is full rank, there is a one-one mapping between the original decision set $\mathcal{D}$ and the new decision set $\tilde{\mathcal{D}} := \{\boldsymbol{z} \in \mathbb{R}^d : \|\boldsymbol{z}\|_\infty \leq 1\}$. Given a $\boldsymbol{z} \in \tilde{\mathcal{D}}$, we can find the corresponding point $\boldsymbol{x} \in \mathcal{D}$ as $\boldsymbol{x} = \boldsymbol{A}^{-1}(\boldsymbol{D}^{-1}(\boldsymbol{z} + \boldsymbol{1}) + \boldsymbol{c})$. So the losses in the new parametrization becomes $\tilde{f}_t(\boldsymbol{z}) = f_t(\boldsymbol{A}^{-1}(\boldsymbol{D}^{-1}(\boldsymbol{z} + \boldsymbol{1}) + \boldsymbol{c}))$.

Let $B := \boldsymbol{A}^{-1}\boldsymbol{D}^{-1}$ and $\boldsymbol{d} := \boldsymbol{A}^{-1}\boldsymbol{D}^{-1}\boldsymbol{1} + \boldsymbol{A}^{-1}\boldsymbol{c}$ so that $\tilde{f}_t(\boldsymbol{z}) = f_t(\boldsymbol{B}\boldsymbol{z} + \boldsymbol{d})$. Then we have

$$\nabla \tilde{f}_t(\boldsymbol{z}) = \boldsymbol{B}^T \nabla f_t(\boldsymbol{B}\boldsymbol{z} + \boldsymbol{d}) \tag{E.182}$$

$$= \boldsymbol{B}^T \nabla f_t(\boldsymbol{x}), \tag{E.183}$$

for a point $\boldsymbol{x} = (\boldsymbol{B}\boldsymbol{z} + \boldsymbol{d}) \in \mathcal{D}$.

Similarly

$$\nabla^2 \tilde{f}_t(\boldsymbol{z}) = \boldsymbol{B}^T \nabla^2 f_t(\boldsymbol{B}\boldsymbol{z} + \boldsymbol{d})\boldsymbol{B} \tag{E.184}$$

$$= \boldsymbol{B}^T \nabla^2 f_t(\boldsymbol{x})\boldsymbol{B}. \tag{E.185}$$

From the above two equations we can easily verify that $\nabla^2 \tilde{f}_t(\boldsymbol{z}) \succcurlyeq \alpha \nabla \tilde{f}_t(\boldsymbol{z}) \nabla \tilde{f}_t(\boldsymbol{z})^T$ as the functions f_t itself are α exp-concave in $\mathcal{D}$.

Further by Holder's inequality we have
$\|\nabla \tilde{f}_t(\boldsymbol{z})\| \leq \|\boldsymbol{A}^{-1}\boldsymbol{D}^{-1}\|_{\text{op}} \|\nabla f_t(\boldsymbol{x})\|_2 \leq \|\boldsymbol{A}^{-1}\boldsymbol{D}^{-1}\|_{\text{op}} G$.

□

Appendix F

Supplementary Materials for Chapter 7

F.1 Analysis

We start with the analysis in the uni-variate setting followed by the proof in multi-dimensions. The analysis requires very clumsy algebraic manipulations in certain places. We used Python's open-source simplification engine SymPy [138] to assist with the algebraic manipulations.

A remark. The constants occurring in the proofs may be optimized further though we haven't aggressively focused on doing so.

F.1.1 One dimensional setting

Lemma 67. ***(KKT conditions)*** *Let $u_1, \ldots, u_n$ be the optimal primal variables and let $\lambda \geq 0$ be the optimal dual variable corresponding to the constraint* (7.8b)*. Further, let $\gamma_t^- \geq 0, \gamma_t^+ \geq 0$ be the optimal dual variables that correspond to constraints* (7.8c) *and* (7.8d) *respectively for all $t \in [n]$. By the KKT conditions, we have*

- ***stationarity:*** $\nabla f_t(u_t) = \lambda \left((s_{t-1} - s_t) - (s_{t-2} - s_{t-1}) \right) + \gamma_t^- - \gamma_t^+$*, where* $s_t = \mathrm{sign}((u_{t+2} - u_{t+1}) - (u_{t+1} - u_t))$*. Here* $\mathrm{sign}(x) = x/|x|$ *if* $|x| > 0$ *and any value in* $[-1, 1]$ *otherwise. For convenience of notations, we also define* $s_{-1} = s_0 = s_{n-1} = s_n = 0$*.*

- ***complementary slackness:*** *(a)* $\lambda \left(\|D^2 u_{1:n}\|_1 - C_n/n \right) = 0$*; (b)* $\gamma_t^-(u_t + 1) = 0$ *and* $\gamma_t^+(u_t - 1) = 0$ *for all* $t \in [n]$

Proof. By introducing auxiliary variables, we can re-write the offline optimization problem as:

$$\min_{\tilde{u}_1, \ldots, \tilde{u}_n} \quad \sum_{t=1}^{n} f_t(\tilde{u}_t) \tag{F.1a}$$

$$\text{s.t.} \quad \tilde{z}_t = \tilde{u}_{t+2} - 2\tilde{u}_{t+1} + \tilde{u}_t \; \forall t \in [n-2] \tag{F.1b}$$

$$\sum_{t=1}^{n-2} |\tilde{z}_t| \leq C_n/n, \tag{F.1c}$$

$$-1 \leq \tilde{u}_t \; \forall t \in [n], \tag{F.1d}$$

$$\tilde{u}_t \leq 1 \; \forall t \in [n], \tag{F.1e}$$

The Lagrangian of the optimization problem can be written as

$$\mathcal{L}(\tilde{u}_{1:n}, \tilde{z}_{1:n-2}, \tilde{v}_{n-2}, \tilde{\gamma}^-_{1:n}, \tilde{\gamma}^+_{1:n}\tilde{\lambda}) = \sum_{t=1}^{n} f_t(\tilde{u}_t) + \tilde{\lambda}\left(\sum_{t=1}^{n-2} |\tilde{z}_t| - C_n/n\right) \tag{F.2}$$

$$+ \sum_{t=2}^{n-2} \tilde{v}_t(\tilde{u}_{t+2} - 2\tilde{u}_{t+1} + \tilde{u}_t - \tilde{z}_t) + \sum_{t=1}^{n} \gamma_t^+(\tilde{u}_t - 1) - \gamma_t^-(\tilde{u}_t + 1). \tag{F.3}$$

Due to stationary conditions wrt u_t, we have

$$\nabla f_t(u_t) = 2v_{t-1} - v_t - v_{t-2} + \gamma_t^- - \gamma_t^+, \tag{F.4}$$

where we define $v_{-1} = v_0 = v_{n-1} = v_n = 0$ and, due to staionarity conditions wrt v_t we have

$$v_t = \lambda \text{sign}(z_t). \tag{F.5}$$

Combining the above two equations and the complementary slackness rule now yields the Lemma.

□

Terminology. In what follows, we refer to $u_{1:n}$ from the Lemma above to be the offline optimal sequence.

Lemma 68. ***(key partition)*** *For some interval $[a, b] \in [n]$, define $\ell_{a \to b} := b - a + 1$. There exists a partitioning of the time horizon $\mathcal{P} := \{[1_s, 1_t], \ldots, [i_s, i_t], \ldots [M_s, M_t]\}$ where $M = |\mathcal{P}|$ such that for any bin $[i_s, i_t] \in \mathcal{P}$ we have: 1) $\|D^2 u_{i_s:i_t}\|_1 \leq 1/\ell^{3/2}_{i_s \to i_t}$; 2) $\|D^2 u_{i_s:i_t+1}\|_1 > 1/\ell^{3/2}_{i_s \to i_t+1}$ and 3) $M = O\left(n^{1/5} C_n^{2/5} \vee 1\right)$.*

Proof. Let the total number of bins formed be M. Consider the case where $M > 1$. We have that

$$\|D^2 u_{1:n}\|_1 \geq \sum_{i=1}^{M-1} \|D^2 u_{i_s \to i_t+1}\|_1 \tag{F.6}$$

$$\geq_{(a)} 1/\ell^{3/2}_{i_s \to i_t+1} \tag{F.7}$$

$$\geq_{(b)} \frac{(M-1)^{5/2}}{n^{3/2}}, \tag{F.8}$$

where line (a) follows due to the construction of the partition and line (b) is due to Jensen's inequality applied to the convex fucntion $f(x) = 1/x^{3/2}$ for $x > 0$.

Rearranging and including the trivial case where $M = 1$ yields the lemma.

□

Proposition 209. *In the following analysis we will often use a useful represent offline optimal within a bin* $[a, b]$ *to be* $m_a, m_a + m_{a+1}, \ldots, \sum_{t=a}^{b} m_t$ *WLOG. We can view this sequence to be samples obtained from a piece-wise linear signal that is continuous at every sampling point.*

Lemma 210. *(**residual bound**) Consider a bin* $[a, b]$*. Let* $\ell := b - a + 1$*. Define:*

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ \vdots & \\ 1 & \ell \end{bmatrix} \tag{F.9}$$

Let $\boldsymbol{\beta} = (\boldsymbol{X}^T \boldsymbol{X})^{-1} \boldsymbol{X}^T u_{a:b}$ *be the least square fit coefficient computed with labels* u_t *and co-variates* $\boldsymbol{x}_t = [1, t - a + 1]^T$ *where* $t \in [a, b]$*. Then we have that the residuals satisfy*

$$|\boldsymbol{\beta}^T \boldsymbol{x}_t - u_t| \leq 20\ell \|D^2 u_{a:a+\ell-1}\|_1, \tag{F.10}$$

whenever $\ell \geq 6$*.*

Proof. We follow the notations of Proposition 209 for representing the offline optimal $u_a, \ldots, u_b$. The residual at time $i \in [a, b]$ can be computed through straight forward algebra as:

$$u_i - \boldsymbol{\beta}^T \boldsymbol{x}_i = \frac{1}{(\ell^2 - 1)} \sum_{j=2}^{\ell} \Bigg(6(1 + (1 - 2i)/\ell)(\ell - j + 1/)(\ell + j)/2 \tag{F.11}$$

$$+ (6i + 6i/\ell - 4\ell - 6 - 2/\ell)(\ell - j + 1) + (\ell^2 - 1)\mathbb{I}\{j \leq i\} \Bigg) m_{a+j-1}, \tag{F.12}$$

where $\mathbb{I}\{\cdot\}$ is the indicator function assuming value 1 if the argument evaluates true and 0 otherwise. Now we note that if all m_k for $k \in [a+1, b]$ are same, then the residuals $u_i - \boldsymbol{\beta}^T \boldsymbol{x}_i$ must be zero for all i as the least square fit exactly matches the labels in this case. In particular, this implies from Eq.(F.12) that

$$\frac{1}{(\ell^2-1)} \sum_{j=2}^{\ell} \Bigg(6(1+(1-2i)/\ell)(\ell-j+1/)(\ell+j)/2 \tag{F.13}$$

$$+ (6i + 6i/\ell - 4\ell - 6 - 2/\ell)(\ell - j + 1) + (\ell^2 - 1)\mathbb{I}\{j \le i\} \Bigg) m_{a+1} = 0. \tag{F.14}$$

Subtracting Eq.(F.14) from (F.12) we get,

$$u_i - \boldsymbol{\beta}^T \boldsymbol{x}_i = \frac{1}{(\ell^2-1)} \sum_{j=2}^{\ell} \Bigg(6(1+(1-2i)/\ell)(\ell-j+1/)(\ell+j)/2 \tag{F.15}$$

$$+ (6i + 6i/\ell - 4\ell - 6 - 2/\ell)(\ell - j + 1) + (\ell^2 - 1)\mathbb{I}\{j \le i\} \Bigg) (m_{j+a-1} - m_{a+1}) \tag{F.16}$$

$$\le \frac{1}{(\ell^2-1)} \max_{j \in [a+2,b]} |m_j - m_{a+1}| \sum_{j=3}^{\ell} \Bigg| 6(1+(1-2i)/\ell)(\ell-j+1/)(\ell+j)/2 \tag{F.17}$$

$$+ (6i + 6i/\ell - 4\ell - 6 - 2/\ell)(\ell - j + 1) + (\ell^2 - 1)\mathbb{I}\{j \le i\} \Bigg|, \tag{F.18}$$

where the last line is due to Holder's inequality. Further, we have $|m_j - m_{a+1}| \le \sum_{t=j}^{a+2} |m_j - m_{j-1}| \le \|D^2 u_{a:b}\|_1$ by the definition of the discrete difference operator D^2.

Now applying triangle inequality and the crude bounds $1+(1-2i)/\ell \le 3, (\ell-j+1) \le \ell, (\ell+j) \le 2\ell, i/\ell \le 1, 2\ell \ge 2/\ell$ and $-2/\ell \le 0$ we obtain

$$\Bigg| 6(1+(1-2i)/\ell)(\ell-j+1/)(\ell+j)/2 \tag{F.19}$$

$$+ (6i + 6i/\ell - 4\ell - 6 - 2/\ell)(\ell - j + 1) + (\ell^2 - 1)\mathbb{I}\{j \le i\} \Bigg| \le 19\ell^2 + 2\ell. \tag{F.20}$$

So,

$$|u_i - \boldsymbol{\beta}^T \boldsymbol{x}_i| \leq \ell \cdot \frac{19\ell^2 + 2\ell}{\ell^2 - 1} \|D^2 u_{a:b}\|_1 \tag{F.21}$$
$$\leq 20\ell \|D^2 u_{a:b}\|_1, \tag{F.22}$$

where the last line is due to $19\ell^2 + 2\ell \leq 20\ell^2 - 20$ for all $\ell \geq 6$.

□

Lemma 211. ***(bounding** T_3)* *Consider a bin $[a, b]$ with length $\ell = b - a + 1$ obtained from the scheme in Lemma 68. Assume the notations in Lemma 210. Let's represent the residual as $r_t := \boldsymbol{\beta}^T \boldsymbol{x}_t - u_t = (t-a+1)M_{t-1} + C_{t-1}$ for $t > a$ and $r_1 := \boldsymbol{\beta}^T \boldsymbol{x}_a - u_a = M_a + C_a$ with $M_b := M_{b-1} = M_{a+\ell-2}$ and $C_b := C_{b-1} = C_{a+\ell-2}$. Suppose $\|D^2 u_{a:b}\|_1 \leq \ell^{-3/2}$. We have,*

$$\sum_{t=a}^{b} f_t(\boldsymbol{\beta}^T \boldsymbol{x}_t) - f_t(u_t) \leq 200 + \lambda \Bigg((s_{a-1} - s_{a-2})(M_a + C_a) - (s_b - s_{b-1})(\ell M_b + C_b) \tag{F.23}$$
$$- s_{a-1} M_a + s_{b-1} M_{b-1} - \sum_{t=a+1}^{b} |M_t - M_{t-1}| \Bigg) \tag{F.24}$$
$$+ 20\ell^{-1/2} \sum_{t=a}^{b} |\gamma_t^- - \gamma_t^+| \tag{F.25}$$

Further we have $|M_a| \leq \|D^2 u_{a:b}\|_1$ and $|M_b| \leq \|D^2 u_{a:b}\|_1$ whenever $\ell \geq 2$.

Here the semantics is that each $M_t = r_{t+1} - r_t$ for all $t > a$ and $M_a = r_{a+1} - r_a$. Any two points r_t and r_{t+1} can be joined using a unique line segment which in turn defines C_t appropriately.

Proof. By gradient Lipschitzness of f we have

$$\sum_{t=a}^{b} f_t(\boldsymbol{\beta}^T \boldsymbol{x}_t) - f_t(u_t) \leq \sum_{t=a}^{b} \langle \nabla f_t(u_t), \boldsymbol{\beta}^T \boldsymbol{x}_t - u_t \rangle + \sum_{t=a}^{b} \frac{1}{2} (\boldsymbol{\beta}^T \boldsymbol{x}_t - u_t)^2. \tag{F.26}$$

Now will focus on bounding the last two terms above.

From the construction of bins in Lemma 68, we know that $\ell \|D^2 u_{a:b}\|_1 \leq 1/\sqrt{\ell}$. Hence we obtain using Lemma 210 that

$$\sum_{t=a}^{b} \frac{1}{2} (\boldsymbol{\beta}^T \boldsymbol{x}_t - u_t)^2 \leq 200. \tag{F.27}$$

Recall the representation of the residual $\boldsymbol{\beta}^T\boldsymbol{x}_t - u_t = tM_t + C_t$ mentioned in the lemma statement. Observe that in accordance with Proposition 209 this residual can also be viewed as samples of a piece-wise linear signal that is continuous at every sampled point. In particular observe that for every $t \in [a, b]$ we have:

$$(t-a+1)M_{t-1} + C_{t-1} = (t-a+1)M_t + C_t \tag{F.28}$$

Consequently

$$C_t - C_{t-1} = (t-a+1)(M_{t-1} - M_t) \tag{F.29}$$

From KKT conditions of Lemma 67 we have

$$\sum_{t=a}^{b}\langle \nabla f_t(u_t), \boldsymbol{\beta}^T\boldsymbol{x}_t - u_t\rangle = \underbrace{\sum_{t=a}^{b}\lambda\left(\left((s_{t-1}-s_{t-2}) - (s_t - s_{t-1})\right)\left((t-a+1)M_t + C_t\right)\right)}_{X_1} \tag{F.30}$$

$$+\underbrace{\sum_{t=a}^{b}(\gamma_t^- - \gamma_t^+)(\boldsymbol{\beta}^T\boldsymbol{x}_t - u_t)}_{X_2} \tag{F.31}$$

$$\frac{X_1}{\lambda} = (s_{a-1} - s_{a-2})(M_a + C_a) - (s_b - s_{b-1})(\ell M_b + C_b) \tag{F.32}$$

$$+\sum_{t=a}^{b-1}(s_t - s_{t-1})\left((t-a+2)M_{t+1} + C_{t+1} - ((t-a+1)M_t + C_t)\right) \tag{F.33}$$

$$=_{(a)}\ (s_{a-1} - s_{a-2})(M_a + C_a) - (s_b - s_{b-1})(\ell M_b + C_b) + \sum_{t=a}^{b-1}(s_t - s_{t-1})M_{t+1} \tag{F.34}$$

$$= (s_{a-1} - s_{a-2})(M_a + C_a) - (s_b - s_{b-1})(\ell M_b + C_b) + \sum_{t=a}^{b-1}(M_{t+1} - M_{t+2})s_t - \tag{F.35}$$

$$s_{a-1}M_2 + s_{b-1}M_\ell \tag{F.36}$$

$$=_{(b)}\ (s_{a-1} - s_{a-2})(M_a + C_a) - (s_b - s_{b-1})(\ell M_b + C_b) - s_{a-1}M_a + s_{b-1}M_{b-1} - \tag{F.37}$$

$$\sum_{t=a+1}^{b}|M_t - M_{t-1}|, \tag{F.38}$$

where in line (a) we used Eq.(F.29) and in line (b) we used the fact that

$s_t = \text{sign}((u_{t+2} - u_{t+1}) - (u_{t+1} - u_t)) = \text{sign}(M_{t+2} - M_{t+1})$ along with the fact that $M_a = M_{a+1}$ and $M_{b-1} = M_b$.

By Holder's inequality and Lemma 210, we have

$$X_2 \leq 20\ell \|D^2 u_{a:b}\|_1 \sum_{t=a}^{b} |\gamma_t^- - \gamma_t^+| \tag{F.39}$$

$$\leq 20\ell^{-1/2} \sum_{t=a}^{b} |\gamma_t^- - \gamma_t^+|, \tag{F.40}$$

where the last line is due to $\|D^2 u_{a:b}\|_1 \leq \ell^{-3/2}$ as assumed in the lemma's statement. Putting everything together completes the proof.

Next, we proceed to give useful bounds on $|M_a|$ and $|M_{b-1}|$.

Since $M_a = M_{a+1}$ and $C_a = C_{a+1}$, we have $M_a = (u_{a+1} - \boldsymbol{\beta}^T \boldsymbol{x}_{a+1}) - (u_a - \boldsymbol{\beta}^T \boldsymbol{x}_a)$.So Eq.(F.12) we have,

$$|M_a| = \left| \sum_{j=2}^{\ell} \frac{6(\ell - j + 1)(1 - j)}{\ell^3 - \ell} (m_{j+a-1} - m_{a+1}) \right| \tag{F.41}$$

$$\leq \|D^2 u_{a:b}\|_1 \sum_{j=3}^{\ell} \frac{6(\ell - j + 1)(j - 1)}{\ell^3 - \ell} \tag{F.42}$$

$$= \|D^2 u_{a:b}\|_1 \frac{\ell^2 + \ell - 6}{\ell(\ell + 1)} \tag{F.43}$$

$$\leq \|D^2 u_{a:b}\|_1, \tag{F.44}$$

where in the last line we used $\frac{\ell^2+\ell-6}{\ell(\ell+1)} \leq 1$ for all $\ell \geq 2$.

Similarly $M_{b-1} = u_b - \boldsymbol{\beta}^T \boldsymbol{x}_b - (u_{b-1} - \boldsymbol{\beta}^T \boldsymbol{x}_{b-1})$ by recalling that $M_b = M_{b-1}$ and $C_b = C_{b-1}$. Proceeding from Eq.(F.12) we obtain,

$$|M_{b-1}| = \left| \sum_{j=2}^{\ell-1} \frac{6(\ell - j + 1)(1 - j)}{\ell^3 - \ell} (m_{j+a-1} - m_b) \right| \tag{F.45}$$

$$\leq \|D^2 u_{a:b}\|_1 \sum_{j=2}^{\ell-1} \frac{6(\ell - j + 1)(j - 1)}{\ell^3 - \ell} \tag{F.46}$$

$$= \|D^2 u_{a:b}\|_1 \frac{\ell^2 + \ell - 6}{\ell(\ell + 1)} \tag{F.47}$$

$$\leq \|D^2 u_{a:b}\|_1. \tag{F.48}$$

□

Lemma 212. *Consider a bin $[a,b] \in \mathcal{P}$ of length ℓ from Lemma 68. Suppose $|u_a| < 1$. Then either $\gamma_j^- = 0$ or $\gamma_j^+ = 0$ for all $j \in [a,b]$.*

Proof. We will provide arguments for the case when the offline optimal first hits -1 before hitting 1 for some point in $[a,b]$. The arguments for the alternate case where it hits 1 first are similar.

If the offline optimal hits -1 at some point in $[a,b]$ it can only rise upto at-most $-1 + 1/\sqrt{l}$ afterwards. This is due to the constraint $\|D^2 u_{a:b}\|_1 \leq 1/\ell^{3/2}$.

Since $-1 + 1/\sqrt{l} < 1$ as $\ell > 1/4$, we have that the offline optimal never touches 1 within the bin $[a,b]$. Consequently $\gamma_j^+ = 0$ for all $j \in [a,b]$.

□

Definition 213. The **slope** of the optimal solution at a time point t is defined to be $u_{t+1} - u_t$ for all $t \in [n-1]$.

Proposition 214. *The bins in $\mathcal{P}$ can be further refined in such a way that each bin either satisfy the condition in Lemma 212 or has constant slope, meaning the L1 TV distance is zero. Further in doing so the size of partition $\mathcal{P}$ only gets increased by at-most 2.*

Proof. Suppose for a bin $[a,b] \in \mathcal{P}$, if the offline optimal starts at 1. Then we can split that bin into two bins $[a,c]$ and $[c+1,b]$ such that $u_c > -1$ and $\|D^2 u_{a:c}\|_1 = 0$. Similar splitting can also be done for bins that start from -1. Observe that this refinement only increases the partition size only by at-most 2. □

Corollary 215. *One powerful consequence of Lemma 212 and Proposition 214 when combined with the fact that γ_t^- and γ_t^+ are both non-negative (Lemma 67) is that $\sum_{t=a}^{b} |\gamma_t^- - \gamma_t^+|$ is either equal to $\sum_{t=a}^{b} \gamma_t^-$ or $\sum_{t=a}^{b} \gamma_t^+$ for all bins $[a,b]$ in the refined partition of Proposition 214 whenever the $\|D^2 u_{a:b}\|_1 > 0$.*

From here on WLOG we will assume that the bins $[a,b]$ in partition $\mathcal{P}$ will satisfy the conditions:

- $\|D^2 u_{a:b}\|_1 \leq 1/l^{3/2}$, where $\ell = b - a + 1$.
- It satisfies the conditions mentioned in Proposition 214 and consequently satisfying the condition in Corollary 215.
- $|\mathcal{P}| = O(n^{1/5} C_n^{2/5})$.

Lemma 216. ***(bounding** T_2**)** *Consider a bin $[a,b] \in \mathcal{P}$ with length $\ell = b - a + 1$ that doesn't touch boundary 1. Let $\Gamma = \sum_{j=a}^{b} \gamma_j^-$ and $\tilde{\Gamma} = \sum_{j=a}^{b} j' \gamma_j^-$ where $j' := j - a + 1$. Let $\boldsymbol{\beta}$ be as in Lemma 210.*

Let $F(\boldsymbol{\beta}) := \sum_{j=a}^{b} f_j(\boldsymbol{x}_j^T \boldsymbol{\beta})$. Define:

$$\boldsymbol{A} := \sum_{j=a}^{b} \boldsymbol{x}_j \boldsymbol{x}_j^T \tag{F.49}$$

Consider the following update:

$$\boldsymbol{\alpha} = \boldsymbol{\beta} - \boldsymbol{A}^{-1} \sum_{j=a}^{b} f_j'(\boldsymbol{x}_j^T \boldsymbol{\beta}) \boldsymbol{x}_j \tag{F.50}$$

$$= \boldsymbol{\beta} - \boldsymbol{A}^{-1} \nabla F(\boldsymbol{\beta}) \tag{F.51}$$

We have,

$$2L\left(F(\boldsymbol{\alpha}) - F(\boldsymbol{\beta})\right) \leq -\|\boldsymbol{g}\|^2_{\boldsymbol{A}^{-1}} - \|\boldsymbol{h}\|^2_{\boldsymbol{A}^{-1}} - 2 < \boldsymbol{g}, \boldsymbol{A}^{-1}\boldsymbol{h}\rangle \tag{F.52}$$

$$+ 2\langle \boldsymbol{A}^{-1}(\boldsymbol{g}+\boldsymbol{h}), \sum_{j=a}^{b} \boldsymbol{x}_j \left(f_j'(\boldsymbol{x}_j^T\boldsymbol{\beta}) - f_j'(u_j)\right), \tag{F.53}$$

where $\boldsymbol{g} = \lambda[-s_{a-2} + s_{a-1} + s_{b-1} - s_b, -s_{a-2} + (\ell+1)s_{b-1} - \ell s_b]^T$ *and* $\boldsymbol{h} = [\Gamma, \tilde{\Gamma}]^T$ *so that* $\sum_{j=a}^{b} f_j'(u_j)\boldsymbol{x}_j = \boldsymbol{g} + \boldsymbol{h}$.

Further we have:

- $\|\boldsymbol{g}\|^2_{\boldsymbol{A}^{-1}}$ *as in Eq.* (F.67)
- $\|\boldsymbol{h}\|^2_{\boldsymbol{A}^{-1}}$ *as in Eq.* (F.90)
- $\langle \boldsymbol{A}^{-1}\boldsymbol{g}, \boldsymbol{h}\rangle$ *as in Eq.* (F.92)
- $\langle \boldsymbol{A}^{-1}\boldsymbol{g}, \sum_{j=a}^{b} \boldsymbol{x}_j(f_j'(\boldsymbol{x}_j^T\boldsymbol{\beta}) - f_j'(u_j))\rangle$ *bounded above by Eq.*(F.83)
- $\langle \boldsymbol{A}^{-1}\boldsymbol{h}, \sum_{j=a}^{b} \boldsymbol{x}_j(f_j'(\boldsymbol{x}_j^T\boldsymbol{\beta}) - f_j'(u_j))\rangle$ *bounded above by Eq.*(F.89)

Similar expressions can be derived for bins $[a, b]$ *that may touch boundary 1 instead of -1.*

Proof. We note that due to gradient Lipschitnzess of f,

$$\nabla^2 F(\boldsymbol{\beta}) = \sum_{j=a}^{b} f_j''(\boldsymbol{x}_j^T\boldsymbol{\beta})\boldsymbol{x}_j\boldsymbol{x}_j^T \preccurlyeq \boldsymbol{A} \tag{F.54}$$

So by Taylor's theorem we have for some $\boldsymbol{z} = t\boldsymbol{\alpha} + (1-t)\boldsymbol{\beta}$

$$F(\boldsymbol{\alpha}) - F(\boldsymbol{\beta}) = -\langle \nabla F(\boldsymbol{\beta}), \boldsymbol{A}^{-1}\nabla F(\boldsymbol{\beta})\rangle + \frac{1}{2}\|\boldsymbol{A}^{-1}\nabla F(\boldsymbol{\beta})\|^2_{\nabla^2 F(\boldsymbol{z})} \tag{F.55}$$

$$\leq -\langle \nabla F(\boldsymbol{\beta}), \boldsymbol{A}^{-1}\nabla F(\boldsymbol{\beta})\rangle + \frac{1}{2}\|\boldsymbol{A}^{-1}\nabla F(\boldsymbol{\beta})\|^2_{\boldsymbol{A}} \tag{F.56}$$

$$= -\frac{1}{2}\|\nabla F(\boldsymbol{\beta})\|^2_{\boldsymbol{A}^{-1}}, \tag{F.57}$$

where

$$\boldsymbol{A}^{-1} = \frac{2}{(\ell-1)\ell}\begin{bmatrix} 2\ell+1 & -3 \\ -3 & \frac{6}{\ell+1} \end{bmatrix} \tag{F.58}$$

Next we turn to lower bounding the above RHS

$$\|\nabla F(\boldsymbol{\beta})\|^2_{\boldsymbol{A}^{-1}} = \|\sum_{j=a}^{b} f_j'(u_j)\boldsymbol{x}_j + \sum_{j=a}^{b}(f_j'(\boldsymbol{x}_j^T\boldsymbol{\beta}) - f_j'(u_j))\boldsymbol{x}_j\|^2_{\boldsymbol{A}^{-1}} \tag{F.59}$$

$$\geq \|\sum_{j=a}^{b} f_j'(u_j)\boldsymbol{x}_j\|^2_{\boldsymbol{A}^{-1}} - 2\langle \boldsymbol{A}^{-1}\sum_{j=a}^{b} f_j'(u_j)\boldsymbol{x}_j, \sum_{j=a}^{b}(f_j'(\boldsymbol{x}_j^T\boldsymbol{\beta}) - f_j'(u_j))\boldsymbol{x}_j\rangle \tag{F.60}$$

From the KKT conditions in Lemma 67, we have

$$\sum_{j=a}^{b} f_j'(u_j)\boldsymbol{x}_j = [\lambda(-s_{a-2}+s_{a-1}+s_{b-1}-s_b)+\Gamma, \lambda(-s_{a-2}+(\ell+1)s_{b-1}-\ell s_b)+\tilde{\Gamma}]^T, \tag{F.61}$$

where Γ and $\tilde{\Gamma}$ are as defined in the statement of the lemma.

For the sake of brevity let's denote $\boldsymbol{g} = \lambda[-s_{a-2}+s_{a-1}+s_{b-1}-s_b, -s_{a-2}+(\ell+1)s_{b-1}-\ell s_b]^T$ and $\boldsymbol{h} = [\Gamma, \tilde{\Gamma}]^T$ so that $\sum_{j=a}^{b} f_j'(u_j)\boldsymbol{x}_j = \boldsymbol{g}+\boldsymbol{h}$.

We have

$$\boldsymbol{A}^{-1}\boldsymbol{g} = \frac{2\lambda}{(\ell-1)\ell}[(2-2\ell)s_{a-2} + (2\ell+1)s_{a-1} - (\ell+2)s_{b-1} + (\ell-1)s_b, \tag{F.62}$$

$$\frac{3(\ell-1)}{\ell+1}s_{a-2} - 3s_{a-1} + 3s_{b-1} - \frac{3(\ell-1)}{\ell+1}s_b]^T, \tag{F.63}$$

and so

$$\|\boldsymbol{g}\|^2_{\boldsymbol{A}^{-1}} = \frac{2\lambda^2}{(\ell-1)\ell(\ell+1)}\Bigg((2\ell^2-3\ell+1)s_{a-2}^2 + (4-4\ell^2)s_{a-2}s_{a-1} - (2\ell^2-6\ell+4)s_{a-2}s_b + \tag{F.64}$$

$$(2\ell^2-2)s_{a-2}s_{b-1} + (2\ell^2+3\ell+1)s_{a-1}^2 + \tag{F.65}$$

$$(2\ell^2-2)s_{a-1}s_b - (2\ell^2+6\ell+4)s_{a-1}s_{b-1} + \tag{F.66}$$

$$(2\ell^2-3\ell+1)s_b^2 + (4-4\ell^2)s_{b-1}s_b + (2\ell^2+3\ell+1)s_{b-1}^2\Bigg) \tag{F.67}$$

Using Eq. (F.63) we get

$$\langle \boldsymbol{A}^{-1}\boldsymbol{g}, \sum_{j=a}^{b} \boldsymbol{x}_j (f_j'(\boldsymbol{x}_j^T\boldsymbol{\beta}) - f_j'(u_j))\rangle = \frac{2\lambda}{(\ell-1)\ell}\Bigg((2-2\ell)s_{a-2} + (2\ell+1)s_{a-1} \tag{F.68}$$

$$- (\ell+2)s_{b-1} + (\ell-1)s_b\Bigg)\sum_{j=a}^{b}(f_j'(\boldsymbol{x}_j^T\boldsymbol{\beta}) - f_j'(u_j)) \tag{F.69}$$

$$+ \frac{2\lambda}{(\ell-1)\ell}\Bigg(\frac{3(\ell-1)}{\ell+1}s_{a-2} - 3s_{a-1} \tag{F.70}$$

$$+ 3s_{b-1} - \frac{3(\ell-1)}{\ell+1}s_b\Bigg)\sum_{j=a}^{b} j'(f_j'(\boldsymbol{x}_j^T\boldsymbol{\beta}) - f_j'(u_j)). \tag{F.71}$$

Using gradient Lipschitzness, triangle inequality and Lemma 210 we have

$$\sum_{j=a}^{b}(f_j'(\boldsymbol{x}_j^T\boldsymbol{\beta}) - f_j'(u_j)) \le \sum_{j=a}^{b} |\boldsymbol{x}_j^T\boldsymbol{\beta} - u_j| \tag{F.72}$$

$$\le 20\ell^2 \|D^2 u_{a:b}\|_1, \tag{F.73}$$

and similarly

$$\sum_{j=a}^{b} j'(f_j'(\boldsymbol{x}_j^T\boldsymbol{\beta}) - f_j'(u_j)) \le \sum_{j=a}^{b} Lj'|\boldsymbol{x}_j^T\boldsymbol{\beta} - u_j| \tag{F.74}$$

$$\le 20\ell^3 \|D^2 u_{a:b}\|_1. \tag{F.75}$$

So continuing from Eq. (F.71),

$$\langle \boldsymbol{A}^{-1}\boldsymbol{g}, \sum_{j=a}^{b} \boldsymbol{x}_j (f_j'(\boldsymbol{x}_j^T\boldsymbol{\beta}) - f_j'(u_j))\rangle \leq \frac{40\lambda\ell \|D^2 u_{a:b}\|_1}{(\ell-1)} \Big| (2-2\ell)s_{a-2} + (2\ell+1)s_{a-1} \tag{F.76}$$

$$- (\ell+2)s_{b-1} + (\ell-1)s_b \Big| \tag{F.77}$$

$$+ \frac{40\lambda\ell^2 \|D^2 u_{a:b}\|_1}{(\ell-1)} \Big| \frac{3(\ell-1)}{\ell+1} s_{a-2} - 3s_{a-1} + 3s_{b-1} \tag{F.78}$$

$$- \frac{3(\ell-1)}{\ell+1} s_b \Big| \tag{F.79}$$

$$\leq \frac{40\lambda\ell^{-1/2}}{(\ell-1)} \Big| (2-2\ell)s_{a-2} + (2\ell+1)s_{a-1} \tag{F.80}$$

$$- (\ell+2)s_{b-1} + (\ell-1)s_b \Big| \tag{F.81}$$

$$+ \frac{40\lambda\ell^{1/2}}{(\ell-1)} \Big| \frac{3(\ell-1)}{\ell+1} s_{a-2} - 3s_{a-1} \tag{F.82}$$

$$+ 3s_{b-1} - \frac{3(\ell-1)}{\ell+1} s_b \Big|, \tag{F.83}$$

where we used $\|D^2 u_{a:b}\|_1 \leq \ell^{-3/2}$.

We have:

$$\langle \boldsymbol{A}^{-1}\boldsymbol{h}, \sum_{j=a}^{b} \boldsymbol{x}_j \left(f_j'(\boldsymbol{\beta}^T\boldsymbol{x}_j) - f_j'(u_j)\right) = \frac{2}{(\ell-1)\ell}\Bigg(((2\ell+1)\Gamma - 3\tilde{\Gamma}) \sum_{j=a}^{b} f_j'(\boldsymbol{\beta}^T\boldsymbol{x}_j) - f_j'(u_j) \tag{F.84}$$

$$+ \left(\frac{6\tilde{\Gamma}}{\ell+1} - 3\Gamma\right) \times \tag{F.85}$$

$$\sum_{j=a}^{b} (j-a+1)\left(f_j'(\boldsymbol{\beta}^T\boldsymbol{x}_j) - f_j'(u_j)\right)\Bigg) \tag{F.86}$$

$$\leq \frac{40\ell \|D^2 u_{a:b}\|_1}{(\ell-1)} |(2\ell+1)\Gamma - 3\tilde{\Gamma}| \tag{F.87}$$

$$+ \frac{40\ell^2 \|D^2 u_{a:b}\|_1}{(\ell-1)} \left|\frac{6\tilde{\Gamma}}{\ell+1} - 3\Gamma\right| \tag{F.88}$$

$$\leq \frac{40\ell^{-1/2}}{(\ell-1)} |(2\ell+1)\Gamma - 3\tilde{\Gamma}| + \frac{40\ell^{1/2}}{(\ell-1)} \left|\frac{6\tilde{\Gamma}}{\ell+1} - 3\Gamma\right|, \tag{F.89}$$

where the last line is obtained by using similar arguments used for obtaining Eq.(F.83).

By substituting the expression for $\boldsymbol{A}^{-1}$ and simplifying,

$$\|\boldsymbol{h}\|^2_{\boldsymbol{A}^{-1}} = \frac{2}{(\ell-1)\ell(\ell+1)} \left((2\ell+1)(\ell+1)\Gamma^2 - 6\Gamma\tilde{\Gamma}(\ell+1) + 6\tilde{\Gamma}^2\right). \tag{F.90}$$

Using Eq.(F.63), we obtain

$$\langle \boldsymbol{A}^{-1}\boldsymbol{g}, \boldsymbol{h}\rangle = \sum_{j=a}^{b} \frac{2\lambda\gamma_j^-}{(\ell-1)\ell}\Bigg(\frac{-3j+3\ell j' - 2\ell^2 + 2}{\ell+1} s_{a-2} + (-3j' + 2\ell + 1)s_{a-1} \tag{F.91}$$

$$+ (3j' - \ell - 2)s_{b-1} + \frac{-3\ell j + 3j' + \ell^2 - 1}{\ell+1} s_b\Bigg) \tag{F.92}$$

□

Lemma 217. ***(bounding** T_1)* *Consider a bin $[a,b]$. Let $\boldsymbol{p}_t$ be the predictions of FLH-SIONS algorithm with parameters $\epsilon = 2$, $C = 20$ and exp-concavity factor σ. Suppose $\boldsymbol{\alpha}$*

and $\boldsymbol{\beta}$ are as defined in Lemma 216. For any $\boldsymbol{\mu} \in \{\boldsymbol{\alpha}, \boldsymbol{\beta}\}$ FLH-SIONS satsifies:

$$\sum_{t=a}^{b} f_t(p_t) - f_t(\boldsymbol{\mu}^T \boldsymbol{x}_t) \leq 256 + \frac{1}{2\sigma} \log(1 + \sigma n/2) + \frac{4}{\sigma} \log n. \tag{F.93}$$

Proof. We will derive the guarantee for $\boldsymbol{\mu} = \boldsymbol{\alpha}$. The guarantee for $\boldsymbol{\mu} = \boldsymbol{\alpha}$ follows similarly.

Let's begin by calculating $\boldsymbol{v} := A^{-1} \sum_{j=a}^{b} f_j'(\boldsymbol{x}_j^T \boldsymbol{\beta}) \boldsymbol{x}_j$.

We have,

$$|\boldsymbol{v}[1]| = \left| \frac{2}{(\ell-1)\ell} \sum_{j=1}^{\ell} (2\ell + 1 - 3j) f'_{(j+a-1)}(\boldsymbol{x}_{j+a-1}^T \boldsymbol{\beta}) \right| \tag{F.94}$$

$$_{(a)} \leq \frac{2}{(\ell-1)\ell} \cdot 2\ell(\ell-1) \tag{F.95}$$

$$= 4, \tag{F.96}$$

where line (a) is obtained via Lipschitzness and Holder's inequality $\boldsymbol{x}^T \boldsymbol{y} \leq \|\boldsymbol{x}\|_1 \|\boldsymbol{y}\|_\infty$ and the fact that $|2\ell + 1 - 3j| \leq 2(\ell - 1)$ for all $j \in [1, \ell]$.

Similarly

$$|\boldsymbol{v}[2]| = \left| \frac{2}{(\ell-1)\ell(\ell+1)} \sum_{j=1}^{\ell} (-3(\ell+1) + 6j) f'_{(j+a-1)}(\boldsymbol{x}_{j+a-1}^T \boldsymbol{\beta}) \right| \tag{F.97}$$

$$\leq \frac{2}{(\ell-1)\ell(\ell+1)} \cdot 3\ell(\ell-1) \tag{F.98}$$

$$= \frac{6}{(\ell+1)}, \tag{F.99}$$

where we used $|-3(\ell+1) + 6j| \leq 3(\ell-1)$ for all $j \in [1, \ell]$.

Combining Eq.(F.96) and (F.99) we conclude that

$$|\boldsymbol{v}^T \boldsymbol{x}_j| \leq 4 + (j - a + 1) \frac{6}{(\ell+1)} \tag{F.100}$$

$$\leq 10, \tag{F.101}$$

where the last line follows due to the fact $(j - a + 1) \leq \ell$.

Hence by Triangle inequality we have,

$$|\boldsymbol{\alpha}^T \boldsymbol{x}_j| \leq |\boldsymbol{\beta}^T \boldsymbol{x}_j| + 10. \tag{F.102}$$

Further note that

$$\|\boldsymbol{v}\|_2 \leq 8 \tag{F.103}$$

Notice that $\boldsymbol{\beta} = \boldsymbol{A}^{-1}\sum_{j=a}^{\ell} u_j \boldsymbol{x}_j$ which have similar functional form as $\boldsymbol{v}$. Since $|u_j| \leq B$ for all $j \in [n]$, by following similar arguments used in bounding $\boldsymbol{v}$ we obtain $|\boldsymbol{\beta}^T \boldsymbol{x}_j| \leq 10$ and

$$\|\boldsymbol{\beta}\|_2 \leq 8. \tag{F.104}$$

Continuing from (F.102) we get

$$|\boldsymbol{\alpha}^T \boldsymbol{x}_j| \leq 20. \tag{F.105}$$

Further,

$$\begin{aligned} \|\boldsymbol{\alpha}\|_2 &\leq \|\boldsymbol{\beta}\|_2 + \|\boldsymbol{v}\|_2 && \text{(F.106)} \\ &\leq 16 && \text{(F.107)} \end{aligned}$$

Since the losses f_t are σ exp-concave in $[-1, 1]$, b+y Theorem 2 in [82] and Lemma 3.3 in [23], FLH-SIONS with parameters set as in the statement of the Lemma yields a regret of

$$\sum_{t=a}^{b} f_t(p_t) - f_t(\boldsymbol{\alpha}^T \boldsymbol{x}_t) \leq 256 + \frac{1}{2\sigma}\log(1 + \sigma n/2) + \frac{4}{\sigma}\log n. \tag{F.108}$$

□

Lemma 218. ***(monotonic slopes)*** *Consider a bin $[i_s, i_t] \in \mathcal{P}$ such that the slopes are monotonic (i.e either non-decreasing or non-increasing). Let p_j be the predictions made by the FLH-SIONS algorithm with parameters as set in Lemma 217. Then we have,*

$$\begin{aligned} \sum_{j=i_s}^{i_t} f_j(p_j) - f_j(u_j) &\leq O\left(\frac{1}{2\sigma}\log(1 + \sigma n/2) + \frac{4}{\sigma}\log n + 210408\right) && \text{(F.109)} \\ &= \tilde{O}(1) && \text{(F.110)} \end{aligned}$$

Proof. We will consider the case of non-decreasing slopes. The alternate case can be handled similarly.

Assume that the slope within the bin is not constant, otherwise we trivially get logarithmic regret as we need only to compete with the best fixed linear fit which is handled by the static regret of FLH-SIONS in any interval ($\boldsymbol{\mu} = \boldsymbol{\beta}$ in Lemma 217).

The optimal solution within a bin of $\mathcal{P}$ obtained via Proposition 214 which doesn't have constant slope may touch either -1 or 1 but not both. Consider the case where the optimal touches -1. Then as the slopes are non-decreasing, once it leaves -1, it never touches -1 again. So we can split the bin $[i_s, i_t]$ into at-most 3 bins $[a, b]$, $[b+1, c]$ and $[c+1, d]$ such that the optimal touches -1 only within $[b+1, c]$. (This bin can be empty if the optimal doesn't touch -1 anywhere within $[i_s, i_t]$).

Now we will bound the regret within bin $[a, b]$.

Suppose that $s_{a-1} = 1$ and $s_b = 1$. If this condition is not satisfied, we can refine the bin $[a, b]$ into at-most 3 bins $[a_1, b_1], [a_2, b_2], [a_3, b_3]$ such that the optimal has constant slope in the first and last bins and $s_{a_2 - 1} = s_{b_2} = 1$. This is possible because the slopes in $[a, b]$ are non-decreasing.

Let $\Delta := \|D^2 u_{a:b}\|_1$ and $\ell := b - a + 1$. Let p and q be two numbers in $[0, 2]$. Substituting $s_{a-2} = 1 - p$, $s_{a-1} = 1$, $s_{b-1} = 1 - q$ and $s_b = 1$ into Lemma 211 and using the fact that $|jM_j + C_j| \leq 20\ell\Delta$ for all $j \in [a, b]$ due to Lemma 210, we get

$$T_3 \leq 40\lambda(p+q)\ell\Delta + 200, \tag{F.111}$$

where we observed that a term arising from Lemma 211: $-M_a + M_{b-1} - \sum_{t=a+1}^{b} |M_t - M_{t-1}| = 0$ as the slopes are non-decreasing.

By making similar sign substitutions in Lemma 216 and noting that $\boldsymbol{h} = \boldsymbol{0}$, we get

$$\begin{aligned}
T_2 &\leq \frac{-2\lambda^2}{2(\ell-1)\ell(\ell+1)} \left((2\ell^2 - 3\ell + 1)p^2 + (2\ell^2 + 3\ell + 1)q^2 + 2(\ell^2 - 1)pq\right) && \text{(F.112)}\\
&\quad + \frac{40\lambda\ell\Delta}{\ell - 1}\left(2p(\ell-1) + q(\ell+2)\right) + \frac{40\lambda\ell^2\Delta}{(\ell-1)(\ell+1)}\left(p(\ell-1) + q(\ell+1)\right) && \text{(F.113)}\\
&\leq \frac{-2\lambda^2}{2(\ell-1)\ell(\ell+1)} \left((2\ell^2 - 3\ell + 1)p^2 + (2\ell^2 + 3\ell + 1)q^2 + 2(\ell^2 - 1)pq\right) && \text{(F.114)}\\
&\quad + 160\lambda\ell\Delta(p+q) + 160\lambda\ell\Delta(p+q), && \text{(F.115)}
\end{aligned}$$

where in the last line we used the fact that $\ell - 1 \geq \ell/2$ and $\ell + 2 \leq 2\ell$ for all $\ell \geq 2$.

Now consider the case where $p \geq q$. Then,

$$\begin{aligned}
(2\ell^2 - 3\ell + 1)p^2 + (2\ell^2 + 3\ell + 1)q^2 + 2(\ell^2 - 1)pq &\geq (2\ell^2 - 3\ell + 1)p^2 && \text{(F.116)}\\
&\geq \ell^2 p^2, && \text{(F.117)}
\end{aligned}$$

where the last line holds for all $\ell \geq 3$. (If $\ell \leq 3$, the regret within the bin is trivially $O(1)$ appealing to the Lipschitzness of the losses f_t and the boundedness of the predictions and the comparators (see proof of Lemma 217)). Thus by using $\ell - 1 \leq \ell$ and $\ell + 1 \leq 2\ell$, we get

$$\frac{-2\lambda^2}{2(\ell-1)\ell(\ell+1)}\left((2\ell^2-3\ell+1)p^2+(2\ell^2+3\ell+1)q^2+2(\ell^2-1)pq\right)\le\frac{-\lambda^2p^2}{2\ell}. \quad \text{(F.118)}$$

Combining Eq. (F.115) and (F.118) and using the fact that $p \geq q$, we have

$$T_2 \le \frac{-\lambda^2p^2}{2\ell} + 640\lambda\ell\Delta p. \quad \text{(F.119)}$$

Similarly from (F.111) using $p \geq q$ we get

$$T_3 \le 40\lambda(p+q)\ell\Delta + 200 \quad \text{(F.120)}$$
$$\le 80\lambda p\ell\Delta + 200 \quad \text{(F.121)}$$

Combining Eq. (F.119) and (F.121) we have

$$T_2 + T_3 \le \frac{-\lambda^2p^2}{2\ell} + 648\lambda p\ell\Delta + 200 \quad \text{(F.122)}$$
$$= -\left(\frac{\lambda p}{\sqrt{2\ell}} - 648\sqrt{2}\ell^{3/2}\Delta\right) + 209952\ell^3\Delta^2 + 200 \quad \text{(F.123)}$$
$$\le 210152, \quad \text{(F.124)}$$

where in the last line we dropped the negative term and used the facts that $\Delta \leq 1/\ell^{3/2}$.

For the case of $q \geq p$, we have

$$(2\ell^2-3\ell+1)p^2+(2\ell^2+3\ell+1)q^2+2(\ell^2-1)pq \ge (2\ell^2-3\ell+1)q^2 \quad \text{(F.125)}$$
$$\ge \ell^2q^2, \quad \text{(F.126)}$$

where the last line holds for all $\ell \geq 3$. This is the same expression as in Eq.(F.117) with p replaced by q. By replacing p with q in the arguments we detailed for the case of $p \geq q$ earlier, we arrive at the same conclusion that $T_2 + T_3 \leq 210152$ even when $q \geq p$. (If $\ell \leq 3$, the regret within the bin is trivially $O(1)$ appealing to the Lipschitzness of the losses f_t and the boundedness of the predictions and the comparators (see proof of Lemma 217))

Similar bound on $T_2 + T_3$ can be shown for bin $[c+1, d]$ by essentially the same arguments.

Hence through Lemma 217 we have the dynamic regret in bins $[a, b]$ to be:

$$\sum_{t=a}^{b} f_t(p_t) - f_t(u_t) \leq 256 + \frac{1}{2\sigma}\log(1+\sigma n/2) + \frac{4}{\sigma}\log n + 210152 \tag{F.127}$$

$$= \tilde{O}(1) \tag{F.128}$$

Similarly, the regret within bin $[c+1, d]$ is also bounded by the above expression.

As the slope within bin $[b+1, c]$ is constant, the regret incurred within this bin is trivially bounded by $256 + \frac{1}{2\sigma}\log(1+\sigma n/2) + \frac{4}{\sigma}\log n$ due to Lemma 217.

Adding the regret incurred across the bins $[a, b]$, $[b+1, c]$ and $[c+1, d]$ together yields the lemma.

□

Next, we will focus on bounding $T_2 + T_3$ for general non-monotonic bins in $\mathcal{P}$.

Lemma 219. ***(non-monotonic slopes)*** *Consider a bin $[i_s, i_t] \in \mathcal{P}$ such that the slopes are not monotonic. Let p_j be the predictions made by the FLH-SIONS algorithm with parameters as set in Lemma 217. Then we have,*

$$\sum_{j=i_s}^{i_t} f_j(p_j) - f_j(u_j) \leq O\left(\frac{1}{\sigma}\log(1+\sigma n) + \frac{12}{\sigma}\log n + 1\right) \tag{F.129}$$

$$= \tilde{O}(1) \tag{F.130}$$

Proof. Let $[a, b] \in \mathcal{P}$ be a bin where the slope is not monotonic and not constant.

Assume that $|s_{a-1}| = |s_b| = 1$. Otherwise we can split the original bin into at-most 3 bins $[a, b_1 - 1], [b_1, b_2], [b_2+1, b]$ such that $|s_{b_1-1}| = |s_{b_2}| = 1$ and slopes are constant in the the other two bins. This is possible because slope in $[a, b]$ is not constant or monotonic.

For a bin $[a, b]$ we define **boundary signs** to be $s_{a-2}, s_{a-1}, s_{b-1}$ and s_b.

First, we will study the case where the offline optimal touches the boundary -1 at two point r and w with $r < w$. The case of arbitrary number of boundary touches will be discussed towards the end. (All arguments can be mirrored appropriately for the case where optimal touches boundary 1).

In what follows we use the notations in the proof of Lemma 216. From Eq.(F.61) we have

$$\boldsymbol{g} + \boldsymbol{h} = \lambda\boldsymbol{\mu} + \gamma_r^- \boldsymbol{x}_r + \gamma_w^- \boldsymbol{x}_w, \tag{F.131}$$

where $\boldsymbol{\mu} \in \mathbb{R}^2$ is a vector depending on the boundary signs and the length $\ell := b - a + 1$. $\boldsymbol{x}_r = [1, r - a + 1]^T$ and $\boldsymbol{x}_w$ defined similarly.

Since $\boldsymbol{g} + \boldsymbol{h}$ is an affine map of $[\lambda, \gamma_r^-, \gamma_w^-]^T$ and since $\boldsymbol{A}$ is positive definite for $\ell \geq 2$, we conclude that $\|\boldsymbol{g} + \boldsymbol{h}\|^2_{\boldsymbol{A}^{-1}}$ is jointly convex in $\lambda, \gamma_r^-, \gamma_w^-$ via appealing to the convexity of squared L2 norm.

First let's focus on the case where boundary signs obey $s_{a-1} = 1$ and $s_b = -1$. Let $s_{a-2} = 1 - p$ and $s_{b-1} = -1 + q$ for some $p, q \in [0, 2]$.

Making these sign substitutions in Lemma 216, we get:

$$\|\boldsymbol{g}\|^2_{\boldsymbol{A}^{-1}} = \frac{2\lambda^2}{(\ell-1)\ell(\ell+1)}\Bigg((2\ell^2 - 3\ell + 1)p^2 + (2\ell^2 + 3\ell + 1)q^2 - (2\ell^2 - 2)pq + 12(\ell-1)p - 12(\ell+1)q + 24\Bigg). \tag{F.132}$$

$$\begin{aligned}\langle \boldsymbol{g}, \boldsymbol{A}^{-1}\boldsymbol{h}\rangle &= \frac{\lambda}{(\ell-1)\ell(\ell+1)}\left(-24 - 6\ell(p-q) + 6(p+q)\right)(r'\gamma_r^- + w'\gamma_w^-) + \\ &+ \frac{\lambda}{(\ell-1)\ell(\ell+1)}\left(2\ell^2(2p-q) - 6\ell q - 4(p+q) + 12(\ell+1)\right)(\gamma_r^- + \gamma_w^-),\end{aligned} \tag{F.133}$$

where $r' = r - a + 1$ and $w' = w - a + 1$.

Let $\Delta := \ell^{-3/2}$. By using equation (F.83) and the facts $\ell - 1 \geq \ell/2$, $\ell + 1 < \ell$, $p, q \in [0, 2]$ and triangle inequality, we bound

$$\begin{aligned}\langle \boldsymbol{A}^{-1}\boldsymbol{g}, \sum_{j=a}^{b} \boldsymbol{x}_j (f'_j(\boldsymbol{x}_j^T\boldsymbol{\beta}) - f'_j(u_j))\rangle &\leq_{(a)} \frac{40\lambda\ell\Delta}{\ell-1}\,|2p(\ell-1) - q(\ell+2) + 6| \\ &\quad + \frac{40\lambda\ell^2\Delta}{\ell-1}\,|3q(\ell+1) + p(1-\ell) - 4| \\ &{}_{(b)} \leq 640\lambda\ell\Delta(p+q) + 800\lambda\Delta,\end{aligned} \tag{F.134}$$

where the line (a) is obtained by equation (F.83) and making the boundary sign substitutions. Line (b) is obtained using the facts $\ell - 1 \geq \ell/2$, $\ell + 2 \leq 2\ell$ whenever $\ell \geq 2$ and $p, q \in [0, 2]$ along with triangle inequality.

From Eq.(F.89), by using similar triangle inequality based arguments and the fact that $|\tilde{\Gamma}| \leq \ell|\Gamma|$ by Holder's inequality and Corollary 215 in as above we obtain

$$\langle \boldsymbol{A}^{-1}\boldsymbol{h}, \sum_{j=a}^{b} \boldsymbol{x}_j (f'_j(\boldsymbol{x}_j^T\boldsymbol{\beta}) - f'_j(u_j))\rangle \leq 1200\ell\Delta(\gamma_r^- + \gamma_w^-). \tag{F.135}$$

To bound T3 we observe from Lemma 211

$$
\begin{aligned}
T_3 &= 200 + \lambda\Bigg((s_{a-1} - s_{a-2})(M_a + C_a) - (s_b - s_{b-1})(\ell M_b + C_b) \\
&- s_{a-1}M_a + s_{b-1}M_{b-1} - \sum_{t=2}^{\ell} |M_t - M_{t-1}| \Bigg) + 20\ell\Delta(\gamma_r^- + \gamma_w^-) \\
&\le 200 + \lambda\Bigg(|(s_{a-1} - s_{a-2})(M_a + C_a)| + |(s_b - s_{b-1})(\ell M_b + C_b)| \\
&+ |M_a| + |M_{b-1}| + \Delta \Bigg) + 20\ell\Delta(\gamma_r^- + \gamma_w^-) \\
&\le 200 + 80\lambda\ell\Delta(p+q) + 3\lambda\Delta + 20\ell\Delta(\gamma_r^- + \gamma_w^-), \qquad \text{(F.136)}
\end{aligned}
$$

where in the last line we used the fact that $|(j-a+1)M_j + C_j| \le 20\ell\Delta$ from Lemma 210.

Recall that $\Delta = \ell^{-3/2}$. Combining all the above equations / inequalities above and Eq. (F.90), define:

$$
\begin{aligned}
T(\lambda, \gamma_r^-, \gamma_w^-) := &\left((2\ell+1)(\ell+1)(\gamma_r^- + \gamma_w^-)^2 - 6(\gamma_r^- + \gamma_w^-)(r'\gamma_r^- + w'\gamma_w^-)(\ell+1) + 6(r'\gamma_r^- + w'\gamma_w^-)^2\right) \\
&+ \left((2\ell+1)(\ell+1)(\gamma_r^- + \gamma_w^-)^2 - 6(\gamma_r^- + \gamma_w^-)(r'\gamma_r^- + w'\gamma_w^-)(\ell+1) + 6(r'\gamma_r^- + w'\gamma_w^-)^2\right) \\
&+ \left(-24 - 6\ell(p-q) + 6(p+q)\right)(r'\gamma_r^- + w'\gamma_w^-) \\
&+ \left(2\ell^2(2p-q) - 6\ell q - 4(p+q) + 12(\ell+1)\right)(\gamma_r^- + \gamma_w^-) \\
&- ((\ell-1)\ell(\ell+1))\left(720\lambda\ell\ell^{-3/2}(p+q) + 803\lambda\ell^{-3/2} + 1220\ell\ell^{-3/2}(\gamma_r^- + \gamma_w^-)\right). \qquad \text{(F.137)}
\end{aligned}
$$

We have,

$$
T_2 + T_3 \le -\frac{T(\lambda, \gamma_r^-, \gamma_w^-)}{(\ell-1)\ell(\ell+1)} + 200. \qquad \text{(F.138)}
$$

The expression in Eq.(F.137) can be compactly written as:

$$
\begin{aligned}
T(\lambda, \check{\gamma}_r^-, \check{\gamma}_w^-) &= 0.5 \cdot (\ell-1)\ell(\ell+1)\|\boldsymbol{g} + \boldsymbol{h}\|_{\boldsymbol{A}^{-1}}^2 + \Phi(\lambda, \check{\gamma}_r^- + \check{\gamma}_w^-) \qquad \text{(F.139)} \\
&:= Q(\lambda, \gamma_r^- + \gamma_w^-, r\gamma_r^- + w\gamma_w^-), \qquad \text{(F.140)}
\end{aligned}
$$

where $\boldsymbol{g}+\boldsymbol{h}$ is as in Eq.(F.131) (which only depends on the boundary signs and $\lambda, \gamma_r^- + \gamma_w^-$

and $r\gamma_r^- + w\gamma_w^-$) and $\Phi(\lambda, \check{\gamma}_r^-, \check{\gamma}_w^-)$ is a linear function of its arguments namely,

$$\Phi(\lambda, \check{\gamma}_r^- + \check{\gamma}_w^-) = -(\ell-1)\ell(\ell+1)\left(20\lambda\ell\ell^{-3/2}(p+q) + 803\lambda\ell^{-3/2} + 1220\ell\ell^{-3/2}(\gamma_r^- + \gamma_w^-)\right) \tag{F.141}$$

Since we have established earlier that $\|\boldsymbol{g}\|^2_{\boldsymbol{A}^{-1}}$ is convex in $\lambda, \gamma_r^-, \gamma_w^-$ we will certainly have $T(\lambda, \gamma_r^-, \gamma_w^-)$ as a function jointly convex in its arguments.

The function B referred in main text is defined to be:

$$B(\lambda, \gamma_r^-, \gamma_w^-; r, w) := -\frac{T(\lambda, \gamma_r^-, \gamma_w^-)}{(\ell-1)\ell(\ell+1)} + 200, \tag{F.142}$$

with r' and w' in Eq.(F.137) to be taken as $r' = r - i_s + 1$ and $w' = w - i_s + 1$, $\ell = i_t - i_s + 1$ and $T(\lambda, \gamma_r^-, \gamma_w^-)$ is as in Eq.(F.137).

So we consider the following convex optimization problem:

$$\min_{\lambda, \gamma_r^-, \gamma_w^-} \quad T(\lambda, \gamma_r^-, \gamma_w^-) \tag{F.143a}$$

$$\text{s.t.} \quad \lambda \geq 0 \tag{F.143b}$$

Note that in the program above we do *unconstrained* minimization over γ_r^- and γ_w^-. Doing so can only further decrease the objective function leading to a valid upper bound on $T_2 + T_3$.

First we will perform a partial minimization wrt the variables γ_r^- and γ_w^-. Differentiating the objective wrt γ_r^- and setting to zero yields:

$$\begin{aligned}
&\left(2(2\ell^2+3\ell+1) - 12(\ell+1)r' + 12(r')^2\right)\hat{\gamma}_r^- \\
&+ \left(2(2\ell^2+3\ell+1) - 6(\ell+1)(r'+w') + 12r'w'\right)\hat{\gamma}_w^- \\
&= \lambda r'\left(24 + 6\ell(p-q) - 6(p+q)\right) - \lambda\left(2\ell^2(2p-q) - 6\ell q - 4(p+q) + 12(\ell+1)\right) \\
&+ 1220\ell^2(\ell^2-1)\ell^{-3/2}.
\end{aligned} \tag{F.144}$$

Similarly differentiating the objective wrt γ_w^- and setting to zero yields:

$$\begin{aligned}
&\left(2(2\ell^2+3\ell+1) - 12(\ell+1)w' + 12(w')^2\right)\hat{\gamma}_w^- \\
&+ \left(2(2\ell^2+3\ell+1) - 6(\ell+1)(r'+w') + 12r'w'\right)\hat{\gamma}_r^- \\
&= \lambda w'\left(24 + 6\ell(p-q) - 6(p+q)\right) - \lambda\left(2\ell^2(2p-q) - 6\ell q - 4(p+q) + 12(\ell+1)\right) \\
&+ 1220\ell^2(\ell^2-1)\ell^{-3/2}.
\end{aligned} \tag{F.145}$$

Solving the above two equations yields:

$$\hat{\gamma}_r^- = \frac{\begin{aligned} &w'\lambda\ell^2 p + w'\lambda\ell^2 q - w'\lambda p - w'\lambda q \\ &+ 1220w'\ell^{0.5} - 1220w'\ell^{2.5} - \lambda\ell^3 q - \lambda\ell^2 p \\ &- \lambda\ell^2 q + 2\lambda\ell^2 + \lambda\ell q + \lambda p \\ &+ \lambda q - 2\lambda - 610\ell^{0.5} - 610\ell^{1.5} + 610\ell^{2.5} + 610\ell^{3.5} \end{aligned}}{r'\ell^2 - r' - w'\ell^2 + w'}, \tag{F.146}$$

and

$$\hat{\gamma}_w^- = \frac{\begin{aligned} &- r'\lambda\ell^2 p - r'\lambda\ell^2 q + r'\lambda p + r'\lambda q \\ &- 1220r'\ell^{0.5} + 1220r'\ell^{2.5} + \lambda\ell^3 q + \lambda\ell^2 p \\ &+ \lambda\ell^2 q - 2\lambda\ell^2 - \lambda\ell q - \lambda p - \lambda q \\ &+ 2\lambda + 610\ell^{0.5} + 610\ell^{1.5} - 610\ell^{2.5} - 610\ell^{3.5} \end{aligned}}{r'\ell^2 - r' - w'\ell^2 + w'}. \tag{F.147}$$

Substituting the above two expression we get:

$$T(\lambda, \hat{\gamma}_r^-, \hat{\gamma}_w^-) = \frac{\begin{aligned} &- 797\lambda\ell^{2.0} - 1780\lambda\ell^{3.0}p - 1780\lambda\ell^{3.0}q \\ &+ 2391\lambda\ell^{4.0} + 5340\lambda\ell^{5.0}p + 5340\lambda\ell^{5.0}q - 2391\lambda\ell^{6.0} \\ &- 5340\lambda\ell^{7.0}p - 5340\lambda\ell^{7.0}q + 797\lambda\ell^{8.0} + 1780\lambda\ell^{9.0}p + 1780\lambda\ell^{9.0}q \\ &+ 744200\ell^{3.5} - 2232600\ell^{5.5} + 2232600\ell^{7.5} - 744200\ell^{9.5} \end{aligned}}{\ell^{2.5} - 2\ell^{4.5} + \ell^{6.5}} \tag{F.148}$$

Looking at Eq.(F.148) we notice that it is a linear fucntion of λ which defined the function $\mathcal{L}(\lambda)$ mentioned in Section 7.2.1 of the main text:

$$\mathcal{L}(\lambda) = \frac{\begin{aligned} &- 797\lambda\ell^{2.0} - 1780\lambda\ell^{3.0}p - 1780\lambda\ell^{3.0}q \\ &+ 2391\lambda\ell^{4.0} + 5340\lambda\ell^{5.0}p + 5340\lambda\ell^{5.0}q - 2391\lambda\ell^{6.0} \\ &- 5340\lambda\ell^{7.0}p - 5340\lambda\ell^{7.0}q + 797\lambda\ell^{8.0} + 1780\lambda\ell^{9.0}p + 1780\lambda\ell^{9.0}q \\ &+ 744200\ell^{3.5} - 2232600\ell^{5.5} + 2232600\ell^{7.5} - 744200\ell^{9.5} \end{aligned}}{\ell^{2.5} - 2\ell^{4.5} + \ell^{6.5}} \tag{F.149}$$

We observe that the leading term (i.e terms whose magnitude is biggest) in the denominator is a positive quantity namely $\ell^{6.5}$. The leading term in the numerator that contains λ grows as $1780\lambda\ell^9(p+q)+797\lambda\ell^8$. So the unconstrained minimum of this linear function is attained at $\lambda = -\infty$.

Hence the constrained minimum (with constraint $\lambda \geq 0$) of the optimization problem F.143a is attained at $\lambda = 0$. We calculate the optimal objective to the constrained problem via Eq.(F.148) as

$$T(0, \hat{\gamma}_r^-, \hat{\gamma}_w^-) = \frac{744200\,(\ell^{1.5} - 3\ell^{3.5} + 3\ell^{5.5} - \ell^{7.5})}{\ell^{0.5} - 2\ell^{2.5} + \ell^{4.5}}, \tag{F.150}$$

where we consider bins with length $\ell \geq 14$.

Since $\ell^4 \geq 2\ell^2$ for all $\ell \geq 2$, we continue from the previous display to obtain:

$$T(0, \hat{\gamma}_r^-, \hat{\gamma}_w^-) \geq -744200 \cdot (1 + 3 + 3 + 1)\frac{\ell^{7.5}}{\ell^{\ell.5}} \tag{F.151}$$

$$= -5953600\ell^3, \tag{F.152}$$

Hence we have

$$\frac{T(0, \hat{\gamma}_r^-, \hat{\gamma}_w^-)}{(\ell-1)\ell(\ell+1)} \geq \frac{-5953600\ell^3}{(\ell-1)\ell(\ell+1)} \tag{F.153}$$

$$\geq -11907200, \tag{F.154}$$

where in the last line we used the fact that $\ell - 1 \geq \ell/2$ is satisfied for all $\ell \geq 14$ and $\ell + 1 > \ell$.

Hence continuing from Eq.(F.138) we conclude that

$$T_2 + T_3 \leq (11907200 + 200) \tag{F.155}$$

$$= 11907400. \tag{F.156}$$

The term T_1 can be bound as

$$T_1 \leq 256 + \frac{1}{2\sigma}\log(1 + \sigma n/2) + \frac{4}{\sigma}\log n \tag{F.157}$$

$$= \tilde{O}(1), \tag{F.158}$$

by Lemma 217.

Now suppose that the offline optimal within bin $[a, b]$ touches boundary -1 more than two times. In this case we propose a reduction to the previous type of analysis where only γ_r^- and γ_w^- are potentially non-zero.

The reduction is facilitated by two observations:

1. While performing the minimization of function $T(\lambda, \gamma_r^-, \gamma_w^-)$ in Eq.(F.137) via the optimization problem F.143a we neither used the fact that r and w are integers nor constrained any bounds on them as well

2. The partially minimized objective in Eq.(F.148) fortunately doesn't depend on neither r nor w.

Now let's consider the case where arbitrary number of γ_j^-, $j \in [a, b]$ can be non-zero.

We can then write,

$$\Gamma = \sum_{j=a}^{b} \gamma_j^- \tag{F.159}$$

$$= \check{\gamma}_r^- + \check{\gamma}_w^-, \tag{F.160}$$

where $\check{\gamma}_r^- := \gamma_1^-$ and $\check{\gamma}_w^- = \Gamma - \check{\gamma}_1^-$.

Define $r' := 1$ and $w' := \frac{\sum_{j=a}^{b} j'\gamma_j^- - \check{\gamma}_r^-}{\check{\gamma}_w^-} = \frac{\tilde{\Gamma} - \check{\gamma}_r^-}{\check{\gamma}_w^-}$ where we assume that $\check{\gamma}_w^- > 0$ (otherwise, we fall back to the earlier analysis).

With these re-definitions we note that

$$T_2 + T_3 \leq -\frac{T(\lambda, \check{\gamma}_r^-, \check{\gamma}_w^-)}{(\ell-1)\ell(\ell+1)}, \tag{F.161}$$

still holds. Further, $T(\lambda, \check{\gamma}_r^-, \check{\gamma}_w^-)$ is jointly convex in its arguments. This can be seen as follows: Note that $T(\lambda, \check{\gamma}_r^-, \check{\gamma}_w^-)$ assumes the form

$$T(\lambda, \check{\gamma}_r^-, \check{\gamma}_w^-) = 0.5 \cdot (\ell-1)\ell(\ell+1)\|\boldsymbol{g} + \boldsymbol{h}\|_{\boldsymbol{A}^{-1}}^2 + \Phi(\lambda, \check{\gamma}_r^- + \check{\gamma}_w^-), \tag{F.162}$$

where $\Phi(\lambda, \check{\gamma}_r^- + \check{\gamma}_w^-)$ is an affine function of its arguments and

$$\boldsymbol{h} = [\Gamma, \tilde{\Gamma}]^T \tag{F.163}$$

$$= [\check{\gamma}_r^- + \gamma_w^-, r'\check{\gamma}_r^- + w'\check{\gamma}_w^-]^T, \tag{F.164}$$

where the last line follows due to our re-parametrizations. By following essentially same arguments as earlier for proving convexity of $T(\lambda, \gamma_r^-, \gamma_w^-)$ we conclude that $T(\lambda, \check{\gamma}_r^-, \check{\gamma}_w^-)$ is also jointly convex in its arguments.

This completes our reduction to the case of two-boundary touches and rest of analysis proceeds by minimizing $T(\lambda, \check{\gamma}_r^-, \check{\gamma}_w^-)$ as earlier.

We now consider the case where $s_{a-1} = s_b = 1$. We can split the original bin $[a, b]$ into two sub-bins $[a_1, b_1]$ and $[a_2, b_2]$ with $a_2 = b_1 + 1$ such that (i) $s_{b_1} = -1$ with $u_{b_1+1} - u_{b_1} > u_{a_2+1} - u_{a_2}$ and (ii) the slopes are non-decreasing within $[a_2, b_2]$. This can be achieved by picking b_1 as the last point within $[a, b]$ where $u_{b_1+1} - u_{b_1} > u_{b_1+2} - u_{b_1+1}$.

In the bin $[a_1, b_1]$ we apply the previous analysis to bound regret by $\tilde{O}(1)$. For the bin $[a_2, b_2]$ we resort to Lemma 218 to bound regret by $\tilde{O}(1)$.

The analysis for the case of boundary signs assignments $s_{a-1} = -1$ and $s_b = 1$ as well as $s_{a-1} = -1$ and $s_b = -1$ can be done similarly.

Adding the regret bounds across all newly formed bins due to potential splitting yields the lemma.

□

Next, we provide the full regret guarantee in a uni-variate setting.

Theorem 220. *Let p_t be the predictions of FLH-SIONS algorithm with parameters $\epsilon = 2$, $C = 20$ and exp-concavity factor σ. Under Assumptions A1-A4, we have that,*

$$\sum_{t=1}^{n} f_t(p_t) - f_t(w_t) = \tilde{O}(n^{1/5} C_n^{2/5} \vee 1), \tag{F.165}$$

for any comparator sequence $w_{1:n} \in \mathcal{TV}^{(1)}(C_n)$. Here $\tilde{O}$ hides poly-logarithmic factors of n and $a \vee b = \max\{a, b\}$.

Proof. The proof is complete by adding the $\tilde{O}(1)$ dynamic regret bound from Lemmas 218 and 219 across $O(n^{1/5} C_n^{2/5} \vee 1)$ bins in the partition $\mathcal{P}$.

□

The proof of Lemma 69 stated in the main text is similar to the arguments used to derive Eq.(F.57). We record it for the sake of completeness.

Lemma 69. *We have that $T_2 \leq -\frac{1}{2} \|\nabla F(\boldsymbol{\beta})\|^2_{\boldsymbol{A}^{-1}}$.*

Proof. We follow the same notations used in defining Lemma 69 in the main text.

Let's begin by calculating $\boldsymbol{v} := \boldsymbol{A}^{-1} \sum_{j=a}^{b} f_j'(\boldsymbol{x}_j^T \boldsymbol{\beta}) \boldsymbol{x}_j$.

We have,

$$|\boldsymbol{v}[1]| = \left| \frac{2}{(\ell-1)\ell} \sum_{j=1}^{\ell} (2\ell + 1 - 3j) f'_{(j+a-1)}(\boldsymbol{x}_{j+a-1}^T \boldsymbol{\beta}) \right| \tag{F.166}$$

$$_{(a)} \leq \frac{2}{(\ell-1)\ell} \cdot 2\ell(\ell-1) \tag{F.167}$$

$$= 4, \tag{F.168}$$

where line (a) is obtained via Lipschitzness and Holder's inequality $\boldsymbol{x}^T \boldsymbol{y} \leq \|\boldsymbol{x}\|_1 \|\boldsymbol{y}\|_\infty$ and the fact that $|2\ell + 1 - 3j| \leq 2(\ell - 1)$ for all $j \in [1, \ell]$.

Similarly

$$|\boldsymbol{v}[2]| = \left| \frac{2}{(\ell-1)\ell(\ell+1)} \sum_{j=1}^{\ell} (-3(\ell+1) + 6j) f'_{(j+a-1)}(\boldsymbol{x}_{j+a-1}^T \boldsymbol{\beta}) \right| \tag{F.169}$$

$$\leq \frac{2}{(\ell-1)\ell(\ell+1)} \cdot 3\ell(\ell-1) \tag{F.170}$$

$$= \frac{6}{(\ell+1)}, \tag{F.171}$$

where we used $|-3(\ell+1) + 6j| \leq 3(\ell - 1)$ for all $j \in [1, \ell]$.

Combining Eq.(F.168) and (F.171) we conclude that

$$|\boldsymbol{v}^T \boldsymbol{x}_j| = 4 + (j - a + 1)\frac{6}{(\ell + 1)} \tag{F.172}$$

$$\leq 10, \tag{F.173}$$

where the last line follows due to the fact $(j - a + 1) \leq \ell$.

Hence by Triangle inequality we have

$$|\boldsymbol{\alpha}^T \boldsymbol{x}_j| \leq |\boldsymbol{\beta}^T \boldsymbol{x}_j| + 10. \tag{F.174}$$

Now we bound $|\boldsymbol{\beta}^T \boldsymbol{x}_j|$ using similar arguments. We have $\boldsymbol{v}' := \boldsymbol{A}^{-1} \sum_{j=a}^{b} u_j \boldsymbol{x}_j$. Now noting that $|u_j| \leq 1$ by Assumption A1 and using similar arguments used to obtain Eq.(F.173) we conclude that

$$|\boldsymbol{\beta}^T \boldsymbol{x}_j| \leq 10. \tag{F.175}$$

So continuing from Eq.(F.174) we have $|\boldsymbol{\alpha}^T \boldsymbol{x}_j| \leq 20$.

For some $\boldsymbol{z} = t\boldsymbol{\alpha} + (1 - t)\boldsymbol{\beta}$, $t \in [0, 1]$ we have by Taylor's theorem that

$$F(\boldsymbol{\alpha}) - F(\boldsymbol{\beta}) = -\langle \nabla F(\boldsymbol{\beta}), \boldsymbol{A}^{-1} \nabla F(\boldsymbol{\beta}) \rangle + \frac{1}{2} \|\boldsymbol{A}^{-1} \nabla F(\boldsymbol{\beta})\|^2_{\nabla^2 F(\boldsymbol{z})} \tag{F.176}$$

$$\leq -\langle \nabla F(\boldsymbol{\beta}), \boldsymbol{A}^{-1} \nabla F(\boldsymbol{\beta}) \rangle + \frac{1}{2} \|\boldsymbol{A}^{-1} \nabla F(\boldsymbol{\beta})\|^2_{\boldsymbol{A}} \tag{F.177}$$

$$= -\frac{1}{2} \|\nabla F(\boldsymbol{\beta})\|^2_{\boldsymbol{A}^{-1}}, \tag{F.178}$$

where in the first inequality we used that fact that $\nabla^2 F(\boldsymbol{z}) \preccurlyeq \boldsymbol{A}$ due to the fact that the functions f_j are 1 gradient Lipschitz in $[-20, 20]^d$ via Assumption A3. □

F.1.2 Multi-dimensional setting

Lemma 221. *Consider the following convex optimization problem.*

$$\min_{\tilde{\boldsymbol{u}}_1, \ldots, \tilde{\boldsymbol{u}}_n, \tilde{\boldsymbol{z}}_1, \ldots, \tilde{\boldsymbol{z}}_{n-1}} \quad \sum_{t=1}^{n} f_t(\tilde{\boldsymbol{u}}_t) \tag{F.179a}$$

$$\text{s.t.} \quad \tilde{\boldsymbol{z}}_t = \tilde{\boldsymbol{u}}_{t+2} - 2\tilde{\boldsymbol{u}}_{t+1} + \tilde{\boldsymbol{u}}_t \ \forall t \in [n-2], \tag{F.179b}$$

$$\sum_{t=1}^{n-2} \|\tilde{\boldsymbol{z}}_t\|_1 \leq C_n/n, \tag{F.179c}$$

$$-1 \leq \tilde{\boldsymbol{u}}_t[k] \ \forall t \in [n], \ \forall k \in [d] \tag{F.179d}$$

$$\tilde{\boldsymbol{u}}_t[k] \leq 1 \ \forall t \in [n], \ \forall k \in [d] \tag{F.179e}$$

`splitMonotonic`: Inputs- (1) offline optimal sequence (2) A bin $[i_s, i_t]$ (3) A coordinate $k \in [d]$

1. Compute $\boldsymbol{z}_j[k] = \boldsymbol{u}_{j+1}[k] - \boldsymbol{u}_j[k]$
2. If $\boldsymbol{z}[k]$ is constant in $[i_s, i_t]$ return $\{i_s, i_t\}$.
3. If $\boldsymbol{z}[k]$ is non-decreasing (non-increasing) across $[i_s, i_t]$: `//ensure equal boundary signs (see caption) for bin` $[b+1, c]$ `below`.
   (a) Split $[i_s, i_t]$ into at-most three bins $[i_s, b], [b+1, c], [c+1, i_t]$ such that $\boldsymbol{z}_j[k]$ remains constant in the first and last bins. Further $\boldsymbol{z}_{b+1}[k] > (<) \boldsymbol{z}_b[k]$ and $\boldsymbol{z}_{c+1}[k] > (<) \boldsymbol{z}_c[k]$.
   (b) Return $\{i_s, b, b+1, c, c+1, i_t\}$

Figure F.1: *`splitMonotonic` procedure. If line 3 is replaced by "If $\boldsymbol{z}[k]$ is non-increasing ...", then we propagate that change by replacing the symbols $>$ / $<$ in the lines below 3 by the bracketed statements next to it. For a bin $[a, b]$, we refer to s_{a-1} and s_b as the boundary signs.*

Let $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n, \boldsymbol{z}_1, \ldots, \boldsymbol{z}_{n-2}$ be the optimal primal variables and let $\lambda \geq 0$ be the optimal dual variable corresponding to the constraint (F.283c). *Further, let $\boldsymbol{\gamma}_t^- \geq 0, \boldsymbol{\gamma}_t^+ \geq 0$ (coordinate-wise) be the optimal dual variables that correspond to constraints* (F.283d) *and* (F.179e) *respectively for all $t \in [n]$. Note that $\boldsymbol{\gamma}_t^-, \boldsymbol{\gamma}_t^+ \in \mathbb{R}^d$. By the KKT conditions, we have*

- ***stationarity:*** $\nabla f_t(\boldsymbol{u}_t) = \lambda \left((\boldsymbol{s}_{t-1} - \boldsymbol{s}_t) - (\boldsymbol{s}_{t-2} - \boldsymbol{s}_{t-1}) \right) + \boldsymbol{\gamma}_t^- - \boldsymbol{\gamma}_t^+$, *where $\boldsymbol{s}_t[k] \in \partial |\boldsymbol{z}_t[k]|$ (a subgradient) for $k \in [d]$. Specifically, $\boldsymbol{s}_t[k] = \mathrm{sign}((\boldsymbol{u}_{t+2}[k] - \boldsymbol{u}_{t+1}[k]) - (\boldsymbol{u}_{t+1}[k] - \boldsymbol{u}_t[k]))$ if $|(\boldsymbol{u}_{t+2}[k] - \boldsymbol{u}_{t+1}[k]) - (\boldsymbol{u}_{t+1}[k] - \boldsymbol{u}_t[k])| > 0$ and $\boldsymbol{s}_t[k]$ is some value in $[-1, 1]$ otherwise. For convenience of notations, we also define $\boldsymbol{s}_{-1} = \boldsymbol{s}_0 = \boldsymbol{0}$.*
- ***complementary slackness:*** *(a) $\lambda \left(\sum_{t=1}^{n-2} \|\boldsymbol{z}_t\|_1 - C_n/n \right) = 0$; (b) $\boldsymbol{\gamma}_t^-[k](\boldsymbol{u}_t[k] + 1) = 0$ and $\boldsymbol{\gamma}_t^+[k](\boldsymbol{u}_t[k] - 1) = 0$ for all $t \in [n]$.*

The proof of above Lemma is similar to that of Lemma 67 and hence omitted.

Lemma 222. *[82] Consider an online learning setting where at each round t, we are given a feature vector $\boldsymbol{x}_t \in \mathbb{R}^2$. Define $\tilde{f}_t(\boldsymbol{v}) = f_t(\boldsymbol{x}_t^T \boldsymbol{v}[1:2], \ldots, \boldsymbol{x}_t^T \boldsymbol{v}[2d-1:2d])$ for some vector $\boldsymbol{v} \in \mathbb{R}^{2d}$. Let the function $f(\boldsymbol{r})$ be σ exp-concave and G Lipschitz for $\boldsymbol{r} \in \mathbb{R}^d$ with $\|\boldsymbol{r}\|_\infty \leq C$. Define $\mathcal{K}_t := \{\boldsymbol{w} \in \mathbb{R}^{2d} : |\boldsymbol{x}_t^T \boldsymbol{w}[2k-1:2k]| \leq C \ \forall k \in [d]\}$. Let $\mathcal{K} := \cap_{t=1}^T \mathcal{K}_t$ and $\boldsymbol{g}_t := \nabla \tilde{f}_t(\boldsymbol{p}_t)$. Consider a variant of the algorithm proposed by [82]* `generateBins`: Input- (1) offline optimal sequence

1. Form consecutive bins $[i_s, i_t]$ such that: `// coarse partition based on TV1 distance`
   (a) $\|D^2 \boldsymbol{u}_{i_s:i_t}\|_1 \leq 1/\ell_{i_s \to i_t}^{3/2}$
   (b) $\|D^2 \boldsymbol{u}_{i_s:i_t+1}\|_1 > 1/\ell_{i_s \to i_t+1}^{3/2}$,

   where $\ell_{a \to b} := b - a + 1$.
2. Let the partition of the time horizon be represented as $\c{P}' := \{[1_s, 1_t], \ldots, [i_s, i_t], \ldots [M_s, M_t]\}$ where $M = |\mathcal{P}'|$.
3. Initialize $\mathcal{R} \leftarrow \Phi$.
4. For each bin $[i_s, i_t] \in \mathcal{P}'$: `// ensuring` $\boldsymbol{\gamma}_j^+[k]\gamma_j^-[k] = 0$ `for all` $k \in [d]$
   (a) $\mathcal{R} = \mathcal{R} \cup \{i_s, i_t\}$.
   (b) For each coordinate $k \in [d]$:
      i. If $\boldsymbol{u}_{i_s}[k] = 1(-1)$ and there exists a point $p \in [i_s, i_t]$ such that $\boldsymbol{u}_p = -1(1)$ then $\mathcal{R} \leftarrow \mathcal{R} \cup \{p-1, p\}$
      ii. If $\boldsymbol{u}_{i_t}[k] = 1(-1)$ and there exists a point $p \in [i_s, i_t]$ such that $\boldsymbol{u}_p = -1(1)$ then $\mathcal{R} \leftarrow \mathcal{R} \cup \{p-1, p\}$
5. Remove duplicates from $\mathcal{R}$ and form a partition $\mathcal{P}$ by splitting at each point in $\mathcal{R}$
6. Return $\mathcal{P}$

Figure F.2: *generateBins procedure. If line 7(d) is replaced by "If $\boldsymbol{z}_p[k] < \boldsymbol{z}_{p-1}[k]$", then we propagate that change by replacing the symbols > / < in the lines below 7(d) by the bracketed statements next to it. For a bin $[a, b]$, we refer to s_{a-1} and s_b as the boundary signs.*

where the algorithm makes a prediction $\hat{\boldsymbol{p}}_{t+1} \in \mathbb{R}^d$ at round $t+1$ as:

$$\boldsymbol{w}_{t+1} = \boldsymbol{p}_t - \boldsymbol{A}_t^{-1}\boldsymbol{g}_t \tag{F.180}$$

$$\boldsymbol{p}_{t+1} = \underset{\boldsymbol{w} \in \mathcal{K}_{t+1}}{\operatorname{argmin}} \|\boldsymbol{w} - \boldsymbol{w}_{t+1}\|_{\boldsymbol{A}_t} \tag{F.181}$$

$$\hat{\boldsymbol{p}}_{t+1} = \left[\boldsymbol{x}_{t+1}^T \boldsymbol{p}_{t+1}[1:2], \ldots, \boldsymbol{x}_{t+1}^T \boldsymbol{p}_{t+1}[2d-1:2d]\right]^T \tag{F.182}$$

where $\boldsymbol{A}_t = \epsilon \boldsymbol{I} + \sum_{s=1}^t \sigma \boldsymbol{g}_s \boldsymbol{g}_s^T$ with $\boldsymbol{I}$ is the identity matrix and ϵ is an input parameter.

Then for any $\boldsymbol{w} \in \mathcal{K}$ we have the regret controlled as

$$\sum_{t=1}^{T} f_t(\hat{\boldsymbol{p}}_t) - \tilde{f}_t(\boldsymbol{w}) = \sum_{t=1}^{T} \tilde{f}_t(\boldsymbol{p}_t) - \tilde{f}_t(\boldsymbol{w}) \tag{F.183}$$

$$\leq \frac{\epsilon \|\boldsymbol{w}\|_2^2}{2} + \frac{2d}{\sigma} \log\left(1 + \frac{\sigma T G^2}{d\epsilon}\right). \tag{F.184}$$

We will call this algorithm as SIONS (Scale Invariant Online Newton Step).

Proof. First we show that exp-concavity is invariant to affine transforms. Since f_t is σ exp-concave, we have

$$\begin{aligned}\tilde{f}_t(\boldsymbol{w}) \geq \tilde{f}_t(\boldsymbol{v}) + \Big\langle &\nabla f_t(\boldsymbol{x}_t^T \boldsymbol{v}[1:2], \ldots, \boldsymbol{x}_t^T \boldsymbol{v}[2d-1:2d]), \\ &[\boldsymbol{x}_t^T(\boldsymbol{w}[1:2] - \boldsymbol{v}[1:2]), \ldots, \boldsymbol{x}_t^T(\boldsymbol{w}[2d-1:2d] - \boldsymbol{v}[2d-1:2d])]^T] \Big\rangle \\ &+ \frac{\sigma}{2}\Big(\Big\langle \nabla f_t(\boldsymbol{x}_t^T \boldsymbol{v}[1:2], \ldots, \boldsymbol{x}_t^T \boldsymbol{v}[2d-1:2d]), \\ &[\boldsymbol{x}_t^T(\boldsymbol{w}[1:2] - \boldsymbol{v}[1:2]), \ldots, \boldsymbol{x}_t^T(\boldsymbol{w}[2d-1:2d] - \boldsymbol{v}[2d-1:2d])]^T] \Big\rangle\Big)^2.\end{aligned} \tag{F.185}$$

For the sake of brevity let's denote $f_t^{(k)} := \nabla f_t(\boldsymbol{x}_t^T \boldsymbol{v}[1:2], \ldots, \boldsymbol{x}_t^T \boldsymbol{v}[2d-1:2d])[k]$ for $k \in [d]$. Then we have

$$\nabla \tilde{f}_t(\boldsymbol{v}) = \left[f_t^{(1)} \boldsymbol{x}_t^T, \ldots, f_t^{(d)} \boldsymbol{x}_t^T\right]^T. \tag{F.186}$$

Let

$$\begin{aligned}A = \Big\langle &\nabla f_t(\boldsymbol{x}_t^T \boldsymbol{v}[1:2], \ldots, \boldsymbol{x}_t^T \boldsymbol{v}[2d-1:2d]), \\ &[\boldsymbol{x}_t^T(\boldsymbol{w}[1:2] - \boldsymbol{v}[1:2]), \ldots, \boldsymbol{x}_t^T(\boldsymbol{w}[2d-1:2d] - \boldsymbol{v}[2d-1:2d])]^T] \Big\rangle.\end{aligned} \tag{F.187}$$

With this, we observe that,

$$A = (\boldsymbol{w} - \boldsymbol{v})^T \nabla \tilde{f}_t(\boldsymbol{v}). \tag{F.188}$$

Thus, we obtain the affine invariance of exp-concavity as:

$$\tilde{f}_t(\boldsymbol{w}) \geq \tilde{f}_t(\boldsymbol{v}) + (\boldsymbol{w}-\boldsymbol{v})^T \nabla \tilde{f}_t(\boldsymbol{v}) + \frac{\sigma}{2}\left((\boldsymbol{w}-\boldsymbol{v})^T \nabla \tilde{f}_t(\boldsymbol{v})\right)^2. \tag{F.189}$$

Note that the set $\mathcal{K}_t$ is convex. This can be seen as follows: if $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{K}_t$, then we have $|\boldsymbol{x}_t^T \boldsymbol{v}[2k-1:2k]| \leq C$ and $|\boldsymbol{x}_t^T \boldsymbol{w}[2k-1:2k]| \leq C$ for all $k \in [d]$. Now for any $t \in [0,1]$ let $\boldsymbol{z} = t\boldsymbol{v} + (1-t)\boldsymbol{w}$. Then we have for any $k \in [d]$ that

$$|\boldsymbol{x}_t^T \boldsymbol{z}[2k-1:2k]| \leq t|\boldsymbol{x}_t^T \boldsymbol{v}[2k-1:2k]| + (1-t)|\boldsymbol{x}_t^T \boldsymbol{2}[2k-1:2k]| \tag{F.190}$$

$$\leq C, \tag{F.191}$$

where the first inequality is via triangle inequality. Thus $\boldsymbol{z} \in \mathcal{K}_t$ so the set $\mathcal{K}_t$ is convex.

So by the properties of projection to convex sets (see for example, Lemma 16 in [77]) and the definition of the algorithm, we have that

$$\|\boldsymbol{p}_{t+1} - \boldsymbol{w}\|_{\boldsymbol{A}_t}^2 \leq \|\boldsymbol{w}_{t+1} - \boldsymbol{w}\|_{\boldsymbol{A}_t}^2 \tag{F.192}$$

$$= \|\boldsymbol{p}_t - \boldsymbol{w}\|_{\boldsymbol{A}_t}^2 + \boldsymbol{g}_t^T \boldsymbol{A}_t^{-1} \boldsymbol{g}_t - 2\boldsymbol{g}_t^T(\boldsymbol{p}_t - \boldsymbol{w}). \tag{F.193}$$

Let $R_T(\boldsymbol{w}) := \sum_{t=1}^T \tilde{f}_t(\boldsymbol{p}_t) - \tilde{f}_t(\boldsymbol{w})$. Since each f_t is exp-concave, we have by Eq.(F.189) and the previous inequality that

$$2R_T(\boldsymbol{w}) \leq \sum_{t=1}^T 2\boldsymbol{g}_t^T(\boldsymbol{p}_t - \boldsymbol{w}) - \sigma(\boldsymbol{g}_t^T(\boldsymbol{p}_t - \boldsymbol{w}))^2 \tag{F.194}$$

$$\leq \sum_{t=1}^T \boldsymbol{g}_t^T \boldsymbol{A}_t^{-1} \boldsymbol{g}_t + \|\boldsymbol{p}_t - \boldsymbol{w}\|_{\boldsymbol{A}_t}^2 - \|\boldsymbol{p}_{t+1} - \boldsymbol{w}\|_{\boldsymbol{A}_t}^2 - \sigma(\boldsymbol{g}_t^T(\boldsymbol{p}_t - \boldsymbol{w}))^2 \tag{F.195}$$

$$\leq \|\boldsymbol{w}\|_{\boldsymbol{A}_0}^2 + \sum_{t=1}^T \boldsymbol{g}_t^T \boldsymbol{A}_t^{-1} \boldsymbol{g}_t + (\boldsymbol{p}_t - \boldsymbol{w})^T(\boldsymbol{A}_t - \boldsymbol{A}_{t-1} - \sigma \boldsymbol{g}_t \boldsymbol{g}_t^T)(\boldsymbol{p}_t - \boldsymbol{w}) \tag{F.196}$$

$$= \|\boldsymbol{w}\|_{\boldsymbol{A}_0}^2 + \sum_{t=1}^T \boldsymbol{g}_t^T \boldsymbol{A}_t^{-1} \boldsymbol{g}_t, \tag{F.197}$$

where the last line is by the definition of $\boldsymbol{A}_t$.

By using the arguments of Lemma 12 of [77] we have

$$\sum_{t=1}^T \boldsymbol{g}_t^T \boldsymbol{A}_t^{-1} \boldsymbol{g}_t \leq \frac{2d}{\sigma} \log\left(1 + \frac{\sigma T G^2}{d\epsilon}\right). \tag{F.198}$$

Thus overall we have,

$$R_T(\boldsymbol{w}) \leq \frac{\epsilon \|\boldsymbol{w}\|_2^2}{2} + \frac{2d}{\sigma} \log\left(1 + \frac{\sigma T G^2}{d\epsilon}\right) \tag{F.199}$$

□

Corollary 223. *[23] Consider the FLH algorithm from [23] with SIONS from Lemma 222 as the base experts with parameter $\epsilon = 2$ as described in Fig.7.3. Consider an arbitrary interval $[a, b] \subseteq [n]$. Then the regret of FLH-SIONS within this interval is controlled as:*

$$\sum_{j=a}^{b} f_j(\boldsymbol{y}_j) - \tilde{f}_j(\boldsymbol{w}) \leq \|\boldsymbol{w}\|_2^2 + \frac{2d}{\sigma} \log\left(1 + \frac{\sigma n^3 G^2}{d\epsilon}\right) + \frac{4 \log n}{\sigma}, \tag{F.200}$$

where $\boldsymbol{w} \in \cap_{j=a}^{b} \mathcal{K}_j$ and $\tilde{f}$ is as defined in Lemma 222.

Proof. Since the loss functions f_j are σ exp-concave, by Lemma 3.3 in [23] we have that

$$\sum_{j=a}^{b} f_j(\boldsymbol{y}_j) \leq \frac{4 \log n}{\sigma} + \sum_{j=a}^{b} f_j(E_a(j)). \tag{F.201}$$

Subtracting $\tilde{f}_j(\boldsymbol{w})$ from both sides and using Lemma 222 now yields the result. □

Corollary 224. *The number of bins $M := |\mathcal{P}|$ formed via a call to **`generateBins`**$(\boldsymbol{u}_{1:n})$ is at-most $O(n^{1/5} C_n^{2/5} \vee 1)$.*

Proof. The proof is similar to that of Lemma 68. □

Lemma 225. *Let $[i_s, i_t] \in \mathcal{P}$ where $\mathcal{P}$ is the partition produced via the **`generateBins`** procedure. We have that the dynamic rgeret of FLH-SIONS within this bin controlled as*

$$\sum_{j=i_s}^{i_t} f_j(\hat{\boldsymbol{p}}_j) - f_j(\boldsymbol{u}_j) = \tilde{O}(d^2), \tag{F.202}$$

where $\hat{\boldsymbol{p}}_j \in \mathbb{R}^d$ are the predictions of the algorithm.

Proof. Consider a bin $[i_s, i_t]$. Let $\mathcal{Q} = \texttt{refineSplit}([i_s, i_t])$. Define $\tilde{f}_j(\boldsymbol{v}) := \tilde{f}_j(\boldsymbol{y}_j^T \boldsymbol{v})$ for $v \in \mathbb{R}^{2d}$.

Next, we proceed to construct the details of a regret decomposition within a bin

$[i_s, i_t]$:

$$\sum_{j=i_s}^{i_t} f_j(\hat{\boldsymbol{p}}_j) - f_j(\boldsymbol{u}_j) = \underbrace{\sum_{j=i_s}^{i_t} f_j(\hat{\boldsymbol{p}}_j) - f_j(\boldsymbol{X}_j \boldsymbol{\alpha}_j)}_{T_1} + \underbrace{\sum_{j=i_s}^{i_t} f_j(\boldsymbol{X}_j \boldsymbol{\alpha}_j) - f_j(\boldsymbol{X}_j \boldsymbol{\beta}_j)}_{T_2} \quad \text{(F.203)}$$

$$+ \underbrace{\sum_{j=i_s}^{i_t} f_j(\boldsymbol{X}_j \boldsymbol{\beta}_j) - f_j(\boldsymbol{u}_j)}_{T_3}, \quad \text{(F.204)}$$

where we will construct appropriate $\boldsymbol{y}_j, \boldsymbol{\alpha}_j, \boldsymbol{\beta}_j \in \mathbb{R}^{2d}$ and $\boldsymbol{X}_j \in \mathbb{R}^{d \times 2d}$ in what follows.

`AssignCo-variatesAndSlopes1`: Inputs- (1) offline optimal sequence (2) A bin $[a, b]$ (3) A coordinate $k \in [d]$

1. Let $\boldsymbol{\beta}_k$ be the least square fit coefficient computed with labels being $\boldsymbol{u}_a[k], \ldots, \boldsymbol{u}_b[k]$ and co-variates $\boldsymbol{x}_j := [1, j - a + 1]^T$ so that the fitted value at time j is given by $\hat{\boldsymbol{u}}_j[k] = \boldsymbol{\beta}^T \boldsymbol{x}_j$.
2. Set $\boldsymbol{\beta}_j[2k-1 : 2k] \leftarrow \beta_k$ for all $j \in [a, b]$.
3. Set $\boldsymbol{\alpha}_k \leftarrow \beta_k$
4. Set $\boldsymbol{\alpha}_j[2k-1 : 2k] \leftarrow \boldsymbol{\alpha}_k$ for all $j \in [a, b]$.
5. Set $\boldsymbol{y}_j[2k-1 : 2k] = \boldsymbol{x}_j$ for all $j \in [a, b]$.

Figure F.3: *AssignCo-variatesAndSlopes1 used to set the parameters in the regret decomposition of Eq.*(F.204) *whenever the offline optimal is constant across the specified coordinate k within the interval* $[a, b]$. *We use a 1-based indexing. i.e* $\boldsymbol{v}[1]$ *refers the first element of a vector* $\boldsymbol{v}$.

(A1): Consider a coordinate $k \in [d]$ such that $\boldsymbol{u}[k]$ is not monotonic in $[i_s, i_t]$ and do not touch boundary 1. Let $[i_s, i_t] = [i_s, a-1] \cup [a, b] \cup [b+1, c] \cup [c+1, i_t]$ such that $\boldsymbol{u}[k]$ is constant in bins $[i_s, a-1]$ and $[c+1, i_t]$. Further we consider the case where $s_{a-1} = 1$ and $s_b = -1$ with $\boldsymbol{u}[k]$ non-decreasing within $[b+1, c]$. (Note that this can be guaranteed by picking b as the last point with $\boldsymbol{u}_{b+1}[k] - \boldsymbol{u}_b[k] > \boldsymbol{u}_{b+2}[k] - \boldsymbol{u}_{b+1}[k]$.) The alternate case where $s_{a-1} = -1$ and $s_b = 1$ with $\boldsymbol{u}[k]$ non-increasing within $[b+1, c]$ can be handled similarly. All the arguments we explain for the case of offline optimal touching the boundary -1 can be mirrored to handle the case where the offline optimal touches the boundary 1. (The offline optimal can't touch both boundaries simultaneously along a coordinate, see Lemma 212)

We will use 1-based indexing. (i.e $\boldsymbol{v}[1]$ denotes the first element of a vector). For each $k \in [d]$:

`AssignCo-variatesAndSlopes2`: Inputs- (1) offline optimal sequence (2) A bin $[a, b]$ (3) A coordinate $k \in [d]$

1. Let $\boldsymbol{\beta}_k$ be the least square fit coefficient computed with labels being $\boldsymbol{u}_a[k], \ldots, \boldsymbol{u}_b[k]$ and co-variates $\boldsymbol{x}_j := [1, j - a + 1]^T$ so that the fitted value at time j is given by $\hat{\boldsymbol{u}}_j[k] = \boldsymbol{\beta}^T \boldsymbol{x}_j$.
2. Set $\boldsymbol{\beta}_j[2k-1 : 2k] \leftarrow \beta_k$ for all $j \in [a, b]$.
3. Set $\boldsymbol{y}_j[2k-1 : 2k] \leftarrow \boldsymbol{x}_j$ for all $j \in [a, b]$.
4. Define $\boldsymbol{A}_k := \sum_{j=a}^{b} \boldsymbol{x}_j \boldsymbol{x}_j^T$, $\tilde{f}_j(\boldsymbol{v}) := \tilde{f}_j(\boldsymbol{y}_j^T \boldsymbol{v})$ for some $v \in \mathbb{R}^{2d}$.
5. Set $\boldsymbol{\alpha}_k \leftarrow \boldsymbol{\beta}_k - \boldsymbol{A}_k^{-1} \sum_{j=a}^{b} \nabla \tilde{f}(\boldsymbol{\beta}_j)[2k-1 : 2k]$.
6. Set $\boldsymbol{\alpha}_j[2k-1 : 2k] \leftarrow \boldsymbol{\alpha}_k$ for all $j \in [a, b]$.

Figure F.4: *`AssignCo-variatesAndSlopes2` used to set the parameters in the regret decomposition of Eq.*(F.204) *whenever the offline optimal may not be constant across the specified coordinate k within the interval $[a, b]$. We use a 1-based indexing. i.e $\boldsymbol{v}[1]$ refers the first element of a vector $\boldsymbol{v}$.*

- Call `AssignCo-variatesAndSlopes1`$(\boldsymbol{u}_{1:n}, [i_s, a-1], k)$.
- Call `AssignCo-variatesAndSlopes2`$(\boldsymbol{u}_{1:n}, [a, b], k)$.
- Let $[b+1, t_1-1], [t_1, t_2], [t_2+1, c]$ be the bins returned by a call to `splitMonotonic`$(\boldsymbol{u}_{1:n}, [b+1, c], k)$.
- Call `AssignCo-variatesAndSlopes1`$(\boldsymbol{u}_{1:n}, [b+1, t_1-1], k)$.
- Call `AssignCo-variatesAndSlopes2`$(\boldsymbol{u}_{1:n}, [t_1, t_2], k)$.
- Call `AssignCo-variatesAndSlopes1`$(\boldsymbol{u}_{1:n}, [t_2+1, c], k)$.
- Call `AssignCo-variatesAndSlopes1`$(\boldsymbol{u}_{1:n}, [c+1, i_t], k)$.

For a vector $\boldsymbol{y}$ we treat $\boldsymbol{y}[m : n] = [\boldsymbol{y}[m], \ldots, \boldsymbol{y}[n]]^T$. Define $\boldsymbol{X}_j \in \mathbb{R}^{d \times 2d}$ as

$$\boldsymbol{X}_j^T = \begin{bmatrix} \boldsymbol{y}_j[1:2] & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{y}_j[3:4] & \ldots & \boldsymbol{0} \\ \vdots & \ddots & & \vdots \\ \boldsymbol{0} & \ldots & & \boldsymbol{y}_j[2d-1:2d] \end{bmatrix}, \tag{F.205}$$

where $\mathbf{0} = [0, 0]^T$ and $\boldsymbol{y}_j$ is set according to various calls of `AssignCo-variatesAndSlopes1` and
`AssignCo-variatesAndSlopes2` as done previously.

We proceed to bound $T_2 + T_3$ in Eq.(F.204). First notice that due to Taylor's theorem,

$$\tilde{f}_j(\boldsymbol{\alpha}_j) - \tilde{f}_j(\boldsymbol{\beta}_j) = \langle \nabla \tilde{f}_j(\boldsymbol{\beta}_j), \boldsymbol{\alpha}_j - \boldsymbol{\beta}_j \rangle + \frac{1}{2} \|\boldsymbol{\alpha}_j - \boldsymbol{\beta}_j\|^2_{\nabla^2 \tilde{f}_j(\boldsymbol{v})}, \tag{F.206}$$

where $\boldsymbol{v} = t\boldsymbol{\alpha}_j + (1-t)\boldsymbol{\beta}_j$ for some $t \in [0, 1]$. Now we use Lemma 226 to obtain,

$$\tilde{f}_j(\boldsymbol{\alpha}_j) - \tilde{f}_j(\boldsymbol{\beta}_j) \leq \langle \nabla \tilde{f}_j(\boldsymbol{\beta}_j), \boldsymbol{\alpha}_j - \boldsymbol{\beta}_j \rangle + \frac{1}{2} \sum_{k'=1}^{d} \|\boldsymbol{\alpha}_j[2k'-1:2k'] - \boldsymbol{\beta}_j[2k'-1:2k']\|^2_{\boldsymbol{y}_j[2k'-1:2k']\boldsymbol{y}_j[2k'-1:2k']^T} \tag{F.207}$$

$$= \sum_{k'=1}^{d} \langle \nabla f_j(\boldsymbol{X}_j\boldsymbol{\beta}_j)[k']\boldsymbol{y}_j[2k'-1:2k'], \boldsymbol{\alpha}_j[2k'-1:2k'] - \boldsymbol{\beta}_j[2k'-1:2k'] \rangle \tag{F.208}$$

$$+ \frac{1}{2} \sum_{k'=1}^{d} \|\boldsymbol{\alpha}_j[2k'-1:2k'] - \boldsymbol{\beta}_j[2k'-1:2k']\|^2_{\boldsymbol{y}_j[2k'-1:2k']\boldsymbol{y}_j[2k'-1:2k']^T} \tag{F.209}$$

Further, due to gradient Lipschitzness,

$$\tilde{f}_j(\boldsymbol{\beta}_j) - f_j(\boldsymbol{u}_j) = f_j(\boldsymbol{X}_j\boldsymbol{\beta}_j) - f_j(\boldsymbol{u}_j) \tag{F.210}$$

$$\leq \langle \nabla f_j(\boldsymbol{u}_j), \boldsymbol{X}_j\boldsymbol{\beta}_j - \boldsymbol{u}_j \rangle + \frac{1}{2} \|\boldsymbol{X}_j\boldsymbol{\beta}_j - \boldsymbol{u}_j\|_2^2 \tag{F.211}$$

$$= \sum_{k'=1}^{d} \nabla f_j(\boldsymbol{u}_j)[k'] \cdot \left(\boldsymbol{\beta}_j[2k'-1:2k']^T \boldsymbol{y}_j[2k'-1:2k'] - \boldsymbol{u}_j[k']\right) \tag{F.212}$$

$$+ \sum_{k'=1}^{d} \frac{1}{2} \|\boldsymbol{\beta}_j[2k'-1:2k']^T \boldsymbol{y}_j[2k'-1:2k'] - \boldsymbol{u}_j[k']\|_2^2 \tag{F.213}$$

Looking at Eq.(F.209) and (F.213), we see that they decompose across each coordinate k'. So we can bound $T_2 + T_3$ in any bin $[i_s, i_t]$ coordinate wise:

$$T_2 + T_3 = \sum_{k'=1}^{d} \sum_{j=i_s}^{i_t} \langle \nabla f_j(\boldsymbol{X}_j \boldsymbol{\beta}_j)[k'] \boldsymbol{y}_j[2k'-1:2k'], \boldsymbol{\alpha}_j[2k'-1:2k'] - \boldsymbol{\beta}_j[2k'-1:2k'] \rangle \tag{F.214}$$

$$+ \frac{1}{2} \sum_{k'=1}^{d} \|\boldsymbol{\alpha}_j[2k'-1:2k'] - \boldsymbol{\beta}_j[2k'-1:2k']\|^2_{\boldsymbol{y}_j[2k'-1:2k']\boldsymbol{y}_j[2k'-1:2k']^T} \tag{F.215}$$

$$+ \nabla f_j(\boldsymbol{u}_j)[k'] \cdot \left(\boldsymbol{\beta}_j[2k'-1:2k']^T \boldsymbol{y}_j[2k'-1:2k'] - \boldsymbol{u}_j[k']\right) \tag{F.216}$$

$$+ \frac{1}{2} \|\boldsymbol{\beta}_j[2k'-1:2k']^T \boldsymbol{y}_j[2k'-1:2k'] - \boldsymbol{u}_j[k']\|_2^2 \tag{F.217}$$

$$:= \sum_{k'=1}^{d} \sum_{j=i_s}^{i_t} t_{2,j,k'} + t_{3,j,k'}, \tag{F.218}$$

where in the last line we define:

$$t_{2,j,k'} := \langle \nabla f_j(\boldsymbol{X}_j \boldsymbol{\beta}_j)[k'] \boldsymbol{y}_j[2k'-1:2k'], \boldsymbol{\alpha}_j[2k'-1:2k'] - \boldsymbol{\beta}_j[2k'-1:2k'] \rangle \tag{F.219}$$

$$+ \frac{1}{2} \|\boldsymbol{\alpha}_j[2k'-1:2k'] - \boldsymbol{\beta}_j[2k'-1:2k']\|^2_{\boldsymbol{y}_j[2k'-1:2k']\boldsymbol{y}_j[2k'-1:2k']^T}, \tag{F.220}$$

and

$$t_{3,j,k'} := \nabla f_j(\boldsymbol{u}_j)[k'] \cdot \left(\boldsymbol{\beta}_j[2k'-1:2k']^T \boldsymbol{y}_j[2k'-1:2k'] - \boldsymbol{u}_j[k']\right) \tag{F.221}$$

$$+ \frac{1}{2} \|\boldsymbol{\beta}_j[2k'-1:2k']^T \boldsymbol{y}_j[2k'-1:2k'] - \boldsymbol{u}_j[k']\|_2^2. \tag{F.222}$$

Next, we proceed to bound $\sum_{j=i_s}^{i_t} t_{2,j,k} + t_{3,j,k}$ for the coordinate k with a structure as mentioned in Paragraph (A1).

Recall that $[i_s, i_t] = [i_s, a-1] \cup [a, b] \cup [b+1, t_1-1] \cup [t_1, t_2] \cup [t_2+1, c] \cup [c+1, i_t]$. So we will consider each of these sub-bins separately.

For bin $[i_s, a-1]$ we have $\boldsymbol{\alpha}_j[2k-1:2k] = \boldsymbol{\beta}_j[2k-1:2k]$ and $\boldsymbol{\beta}_j[2k-1:2k]^T \boldsymbol{y}_j[2k-1:2k] = \boldsymbol{u}_j[k]$. So we trivially have

$$\sum_{j=i_s}^{a-1} t_{2,j,k} + t_{3,j,k} = 0. \tag{F.223}$$

Next, we focus on the bin $[a, b]$. We note that by construction, $\boldsymbol{\alpha}_j[2k-1:2k]$ and $\boldsymbol{\beta}_j[2k-1:2k]$ are fixed for all $j \in [a, b]$. Let's denote these fixed values by $\boldsymbol{\alpha}_k$ and $\boldsymbol{\beta}_k$ respectively. For the sake of brevity let's denote $\boldsymbol{x}_j := \boldsymbol{y}_j[2k-1:2k]$ and

$\boldsymbol{A}_k = \sum_{j=a}^{b} \boldsymbol{x}_j \boldsymbol{x}_j^T$. We have the relation,

$$\boldsymbol{\alpha}_k = \boldsymbol{\beta}_k - \boldsymbol{A}_k^{-1} \sum_{j=a}^{b} \nabla \tilde{f}(\boldsymbol{\beta}_j)[2k-1:2k] \tag{F.224}$$

$$= \boldsymbol{\beta}_k - \boldsymbol{A}_k^{-1} \sum_{j=a}^{b} \nabla f_j(\boldsymbol{X}_j \boldsymbol{\beta}_j)[k] \boldsymbol{x}_j. \tag{F.225}$$

By the new compact notations, we have

$$t_{2,j,k} = \langle \nabla f_j(\boldsymbol{X}_j \boldsymbol{\beta}_j)[k] \boldsymbol{x}_j, \boldsymbol{\alpha}_k - \boldsymbol{\beta}_k \rangle + \frac{1}{2} \|\boldsymbol{\alpha}_k - \boldsymbol{\beta}_k\|^2_{\boldsymbol{x}_j \boldsymbol{x}_j^T}, \tag{F.226}$$

and

$$t_{3,j,k} = \nabla f_j(\boldsymbol{u}_j)[k] \cdot \left(\boldsymbol{\beta}_k^T \boldsymbol{x}_j - \boldsymbol{u}_j[k]\right) + \frac{1}{2} \|\boldsymbol{\beta}_k^T \boldsymbol{x}_j - \boldsymbol{u}_j[k]\|_2^2. \tag{F.227}$$

From Eq.(F.225) we have,

$$\sum_{j=a}^{b} t_{2,j,k} = -\left\| \sum_{j=a}^{b} \nabla f_j(\boldsymbol{X}_j \boldsymbol{\beta}_j)[k] \boldsymbol{x}_j \right\|^2_{\boldsymbol{A}_k^{-1}} + \frac{1}{2} \left\| \boldsymbol{A}_k^{-1} \sum_{j=a}^{b} \nabla f_j(\boldsymbol{X}_j \boldsymbol{\beta}_j)[k] \boldsymbol{x}_j \right\|^2_{\boldsymbol{A}_k} \tag{F.228}$$

$$= -\frac{1}{2} \left\| \sum_{j=a}^{b} \nabla f_j(\boldsymbol{X}_j \boldsymbol{\beta}_j)[k] \boldsymbol{x}_j \right\|^2_{\boldsymbol{A}_k^{-1}} \tag{F.229}$$

$$\leq -\frac{1}{2} \left\| \sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k] \boldsymbol{x}_j \right\|^2_{\boldsymbol{A}_k^{-1}} + 2 \langle \boldsymbol{A}_k^{-1} \sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k] \boldsymbol{x}_j, \tag{F.230}$$

$$\sum_{j=a}^{b} \boldsymbol{x}_j \left(\nabla f_j(\boldsymbol{X}_j \boldsymbol{\beta}_j)[k] - \nabla f_j(\boldsymbol{u}_j)[k]\right) \rangle. \tag{F.231}$$

Now we define $\boldsymbol{g}_k = \lambda[-\boldsymbol{s}_{a-2}[k] + \boldsymbol{s}_{a-1}[k] + \boldsymbol{s}_{b-1}[k] - \boldsymbol{s}_b[k], -\boldsymbol{s}_{a-2}[k] + (\ell+1)\boldsymbol{s}_{b-1}[k] - \ell \boldsymbol{s}_b[k]]^T$ and $\boldsymbol{h}_k = [\boldsymbol{\Gamma}[k], \tilde{\boldsymbol{\Gamma}}[k]]^T$ where $\boldsymbol{\Gamma} = \sum_{j=a}^{b} \boldsymbol{\gamma}_j^- - \boldsymbol{\gamma}_j^+$ and $\tilde{\boldsymbol{\Gamma}} = \sum_{j=a}^{b} j'(\boldsymbol{\gamma}_j^- - \boldsymbol{\gamma}_j^+)$ where $j' = j - a + 1$ so that $\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k] \boldsymbol{x}_j = \boldsymbol{g}_k + \boldsymbol{h}_k$ via the KKT conditions in Lemma 221.

With these, we can bound:

$$\sum_{j=a}^{b} 2 \cdot t_{2,j,k} \leq -\|\boldsymbol{g}_k\|^2_{\boldsymbol{A}_k^{-1}} - \|\boldsymbol{h}_k\|^2_{\boldsymbol{A}_k^{-1}} - 2 < \boldsymbol{g}_k, \boldsymbol{A}_k^{-1}\boldsymbol{h}_k\rangle \tag{F.232}$$

$$+ 2\langle \boldsymbol{A}_k^{-1}(\boldsymbol{g}_k + \boldsymbol{h}_k), \sum_{j=a}^{b} \boldsymbol{x}_j \left(\nabla f_j(\boldsymbol{X}_j\boldsymbol{\beta}_j)[k] - \nabla f_j(\boldsymbol{u}_j)[k]\right)\rangle \tag{F.233}$$

Proceeding similarly to Eq.(F.73) and (F.75) by gradient Lipschitzness we obtain,

$$\sum_{j=a}^{b} \nabla f_j(\boldsymbol{X}_j\boldsymbol{\beta}_j)[k] - \nabla f_j(\boldsymbol{u}_j)[k] \leq \sum_{j=a}^{b} \|\boldsymbol{X}_j\boldsymbol{\beta}_j - \boldsymbol{u}_j\|_1 \tag{F.234}$$

$$\leq 20\ell^2\ell^{-3/2}, \tag{F.235}$$

where in the last line we used Lemma 210 coordinate-wise and the fact that $\|D^2\boldsymbol{u}_{a:b}\|_1 \leq \ell^{-3/2}$.

Similarly,

$$\sum_{j=a}^{b} j'(\nabla f_j(\boldsymbol{X}_j\boldsymbol{\beta}_j)[k]) - \nabla f_j(\boldsymbol{u}_j)[k] \leq \sum_{j=a}^{b} j'\|\boldsymbol{X}_j\boldsymbol{\beta}_j - \boldsymbol{u}_j\|_1 \tag{F.236}$$

$$\leq 20\ell^3\ell^{-3/2}. \tag{F.237}$$

Hence by KKT conditions in Lemma 221, we can further bound

$$\langle \boldsymbol{A}_k^{-1}\boldsymbol{g}_k, \sum_{j=a}^{b} \boldsymbol{x}_j \left(\nabla f_j(\boldsymbol{X}_j\boldsymbol{\beta}_j)[k] - \nabla f_j(\boldsymbol{u}_j)[k]\right)\rangle \leq \frac{40\lambda\ell^{-1/2}}{(\ell-1)}\Bigg|(2-2\ell)\boldsymbol{s}_{a-2}[k]+ \tag{F.238}$$

$$(2\ell+1)\boldsymbol{s}_{a-1}[k] \tag{F.239}$$

$$-(\ell+2)\boldsymbol{s}_{b-1}[k] + (\ell-1)\boldsymbol{s}_b[k]\Bigg| \tag{F.240}$$

$$+\frac{40\lambda\ell^{1/2}}{(\ell-1)}\Bigg|\frac{3(\ell-1)}{\ell+1}\boldsymbol{s}_{a-2}[k] - 3\boldsymbol{s}_{a-1}[k] \tag{F.241}$$

$$+3\boldsymbol{s}_{b-1}[k] - \frac{3(\ell-1)}{\ell+1}\boldsymbol{s}_b[k]\Bigg|, \tag{F.242}$$

and

$$\langle \boldsymbol{A}_k^{-1}\boldsymbol{h}_k, \sum_{j=a}^{b} \boldsymbol{x}_j \left(\nabla f_j(\boldsymbol{X}_j\boldsymbol{\beta}_j)[k] - \nabla f_j(\boldsymbol{u}_j)[k]\right)\rangle \leq \frac{40\ell^{-1/2}}{(\ell-1)}|(2\ell+1)\boldsymbol{\Gamma}[k] - 3\tilde{\boldsymbol{\Gamma}}[k]| \quad \text{(F.243)}$$

$$+ \frac{40\ell^{1/2}}{(\ell-1)}\left|\frac{6\tilde{\boldsymbol{\Gamma}}[k]}{\ell+1} - 3\boldsymbol{\Gamma}[k]\right|. \quad \text{(F.244)}$$

We observe that Eq.(F.233),(F.242),(F.244) are semantically same as Eq.(F.218), (F.73) and (F.75) respectively in the 1D case.

Next, we proceed to setup a similar observation for bounding $\sum_{j=a}^{b} t_{3,j,k}$. From KKT conditions in Lemma 221 and proceeding similar to the arguments in Lemma 211 we get,

$$\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k] \cdot \left(\boldsymbol{\beta}_j[2k-1:2k]^T \boldsymbol{y}_j[2k-1:2k] - \boldsymbol{u}_j[k]\right) \quad \text{(F.245)}$$

$$= \sum_{j=a}^{b} \lambda \Bigg(((\boldsymbol{s}_{j-1}[k] - \boldsymbol{s}_{j-2}[k]) - (\boldsymbol{s}_j[k] - \boldsymbol{s}_{j-1}[k])) \times \quad \text{(F.246)}$$

$$((j-a+1)\boldsymbol{M}_j[k] + \boldsymbol{C}_j[k])\Bigg) \quad \text{(F.247)}$$

$$+ \sum_{j=a}^{b} (\boldsymbol{\gamma}_j^-[k] - \boldsymbol{\gamma}_j^+[k]) \times \quad \text{(F.248)}$$

$$(\boldsymbol{\beta}_j[2k-1:2k]^T \boldsymbol{y}_j[2k-1:2k] - \boldsymbol{u}_j[k]) \quad \text{(F.249)}$$

$$\leq \sum_{j=a}^{b} \lambda \Bigg(((\boldsymbol{s}_{j-1}[k] - \boldsymbol{s}_{j-2}[k]) - (\boldsymbol{s}_j[k] - \boldsymbol{s}_{j-1}[k])) \times \quad \text{(F.250)}$$

$$((j-a+1)\boldsymbol{M}_j[k] + \boldsymbol{C}_j[k])\Bigg) \quad \text{(F.251)}$$

$$+ 20\ell^{-1/2} \sum_{j=a}^{b} \left|\boldsymbol{\gamma}_j^-[k] - \boldsymbol{\gamma}_j^+[k]\right|, \quad \text{(F.252)}$$

where similar to Lemma 211, we represent $\boldsymbol{\beta}_j[2k-1:2k]^T\boldsymbol{y}_j[2k-1:2k] - \boldsymbol{u}_j[k] = (j-a+1)\boldsymbol{M}_j[k] + \boldsymbol{C}_j[k]$ with $\boldsymbol{M}_a[k] = \boldsymbol{M}_{a+1}[k]$, $\boldsymbol{C}_a[k] = \boldsymbol{C}_{a+1}[k]$, $\boldsymbol{M}_b[k] = \boldsymbol{M}_{b-1}[k]$ and $\boldsymbol{C}_b[k] = \boldsymbol{C}_{b-1}[k]$. The last line is obtained due to Lemma 210.

Further, by using Lemma 210 we obtain,

$$\sum_{k'=1}^{d} \sum_{j=a}^{b} \frac{1}{2} \|\boldsymbol{\beta}_j[2k'-1:2k']^T \boldsymbol{y}_j[2k'-1:2k'] - \boldsymbol{u}_j[k']\|_2^2 \leq 200. \quad \text{(F.253)}$$

Combining the last two inequalities yields,

$$\sum_{j=a}^{b} t_{3,j,k} \leq 200 + \sum_{j=a}^{b} \lambda \Bigg(\left((\boldsymbol{s}_{j-1}[k] - \boldsymbol{s}_{j-2}[k]) - (\boldsymbol{s}_j[k] - \boldsymbol{s}_{j-1}[k]) \right) \left((j-a+1)\boldsymbol{M}_j[k] \right. \tag{F.254}$$

$$\left. + \boldsymbol{C}_j[k] \right) \Bigg) + 20\ell^{-1/2} \sum_{j=a}^{b} \left| \boldsymbol{\gamma}_j^-[k] - \boldsymbol{\gamma}_j^+[k] \right|. \tag{F.255}$$

We observe that the last inequality is semantically similar to Eq.(F.25) for 1D case. Recall that Eq.(F.233),(F.242),(F.244) are also semantically same as Eq.(F.218), (F.83) and (F.89) respectively in the 1D case.

Hence we can proceed to bound

$$\sum_{j=a}^{b} t_{2,j,k} + t_{3,j,k} = O(1), \tag{F.256}$$

using the same arguments as in Lemma 219.

Observe that by construction, the slopes across coordinate k are constant in the bins $[b+1, t_1-1], [t_2+1, c]$ and $[c+1, i_t]$. So by using similar arguments used for handling the bin $[i_s, a-1]$ we obtain,

$$\sum_{j \in \mathcal{I}} t_{2,j,k} + t_{3,j,k} = 0, \tag{F.257}$$

where $\mathcal{I} \in \{[b+1, t_1-1], [t_2+1, c], [c+1, i_t]\}$.

By appealing to our reduction to 1D case facilitated by Eq.(F.233) and (F.255) and using similar arguments used to handle the monotonic slopes case as in Lemma 218 we obtain,

$$\sum_{j=t_1}^{t_2} t_{2,j,k} + t_{3,j,k} = O(1). \tag{F.258}$$

So far we have discussed bounding $\sum_{j=i_s}^{i_t} t_{2,j,k} + t_{3,j,k}$ for a bin with structure across coordinate k as described in Paragraph (A1). We remark that if the slopes across a coordinate k assumes a monotonic structure across $[i_s, i_t]$, we can handle it in the same way as we handled the sub-bin $[t_1, t_2]$ above.

We pause to remark that Eq.(F.223),(F.256),(F.257) and (F.258) together gives a way to bound to $\sum_{j=i_s}^{i_t} t_{2,j,k'} + t_{3,j,k'}$ across any coordinate k' as we comprehensively considered all the possible structures across a coordinate k'. (The alternate cases where where $s_{a-1} = -1$ and $s_b = 1$ with $\boldsymbol{u}[k']$ non-increasing within $[b+1, c]$ can be handled similarly to the case described in Paragraph (A1). Finally the case where the offline optimal touches boundary 1 instead of -1 can be handled using similar arguments.)

Thus overall we obtain that for any bin $[i_s, i_t] \in \mathcal{P}$ we have:

$$T_2 + T_3 \leq \sum_{k'=1}^{d} \sum_{j=i_s}^{i_t} t_2 t_{2,j,k'} + t_{3,j,k'} \tag{F.259}$$

$$= O(d), \tag{F.260}$$

where T_2 and T_3 are as defined in Eq.(F.204).

Next, we proceed to control T_1. Recall that

$$T_1 = \sum_{j=i_s}^{i_t} f_j(\boldsymbol{p}_j) - f_j(\boldsymbol{X}_j \boldsymbol{\alpha}_j). \tag{F.261}$$

Let's revisit bin $[i_s, i_t]$ with structure as described in Paragraph (A1) across coordinate k. First we consider the bin $[a, b]$. Through the call to `AssignCo-variatesAndSlopes2`$(\boldsymbol{u}_{1:n}, [a,b], k)$ we set $\boldsymbol{\alpha}_k$. Further $\boldsymbol{\alpha}_j[2k-1:2k] = \boldsymbol{\alpha}_k$ for all $j \in [a,b]$. By using similar arguments as in the proof of Lemma 217 which lead to Eq.(F.105), we have that $|\boldsymbol{y}_j[2k-1:2k]^T \boldsymbol{\alpha}_j| \leq 20$. For other bins such as $[i_s, a-1], [b+1, t_1-1], [t_2+1, c], [c+1, i_t]$ where the slope of the offline optimal across coordinate k remains constant, we set $\alpha_j[2k-1:2k]$ for $j \in \mathcal{I}$ with $\mathcal{I} \in \{[i_s, a-1], [b+1, t_1-1], [t_2+1, c], [c+1, i_t]\}$ to be a constant value obtained as the least square fit coefficients with co-variates $\boldsymbol{y}_j[2k-1:2k]$ and labels set appropriately via the call to `AssignCo-variatesAndSlopes1`. Hence in this case also we have $|\boldsymbol{y}_j^T[2k-1:2k]\boldsymbol{\alpha}_j[2k-1:2k]| \leq 10$ via the arguments in Lemma 217.

For the alternate cases (i) where $s_{a-1} = -1$ and $s_b = 1$ with $\boldsymbol{u}[k']$ non-increasing within $[b+1, c]$ as described in Paragraph (A1) (ii) case where the offline optimal touches boundary 1 instead of -1 (iii) The offline optimal across coordinate k is non-decreasing within $[i_s, i_t]$ and (iv) The offline optimal across coordinate k is non-increasing within $[i_s, i_t]$. In all these cases we can set the quantities $\boldsymbol{\alpha}_j[2k-1:2k], \boldsymbol{y}_j[2k-1:2k]$ by similar calls to `AssignCo-variatesAndSlopes1` or `AssignCo-variatesAndSlopes2` such that $\boldsymbol{y}_j[2k-1:2k]^T \boldsymbol{\alpha}_j[2k-1:2k] \leq 20$ for all $j \in [i_s, i_t]$. For example, for case (iii) we can resort to similar arguments used for handling sub-bin $[t_1, t_2]$ which is again similar to how we handled the bin $[a, b]$. (see Paragraph (A1)).

Further, even-though we create at-most 6 sub-bins across each coordinate for an interval $[i_s, i_t] \in \mathcal{P}$ (see Paragraph (A1) and the sequence of calls beneath), doing so for each coordinate can result in at-most $6d$ partitions of $\boldsymbol{u}_{i_s:i_t}$ overall. However, if we consider any sub-bin $[p, q]$ of this partition, we have that $\boldsymbol{\alpha}_j[2k-1:2k]$ is fixed and $\boldsymbol{\beta}_j[2k-1:2k]$ is fixed for all $j \in [p,q]$ across any coordinate $k \in [d]$ and $y_j[2k-1:2k][2]$ is monotonically increasing wrt $j \in [p,q]$ for all coordinates $k \in [d]$. Now suppose that $k' \in [d]$ is such that $\boldsymbol{y}_j[2k'-1:2k'][1] \leq \boldsymbol{y}_j[2k-1:2k][1]$ for all $k \neq k'$ and for all $j \in [p,q]$. With a change of variables we have that $\tilde{\boldsymbol{\alpha}}_j[2k-1:2k]^T \boldsymbol{y}_j[2k'-1:2k'] =$

$\boldsymbol{\alpha}_j[2k-1:2k]^T\boldsymbol{y}_j[2k-1:2k]$ by setting $\tilde{\boldsymbol{\alpha}}_j[2k-1:2k][2] = \boldsymbol{\alpha}_j[2k-1:2k][2]$ and $\tilde{\boldsymbol{\alpha}}_j[1] = \boldsymbol{\alpha}_j[1] + (\boldsymbol{y}_j[2k-1:2k][2] - \boldsymbol{y}_j[2k'-1:2k'][2])\boldsymbol{\alpha}_j[2k-1:2k][2]$ for $k \neq k'$ within the bin $[p,q]$. Since $(\boldsymbol{y}_j[2k-1:2k][2] - \boldsymbol{y}_j[2k'-1:2k'][2]) \leq \boldsymbol{y}_j[2k-1:2k][2]$ by Eq.(F.99) we have that

$$|(\boldsymbol{y}_j[2k-1:2k][2] - \boldsymbol{y}_j[2k'-1:2k'][2])\boldsymbol{\alpha}_j[2k-1:2k][2]| \leq 6. \tag{F.262}$$

Further we have from Eq.(F.105) that $|\boldsymbol{\alpha}_j[2k-1:2k][2]\boldsymbol{y}_j[2k'-1:2k'][2] + \boldsymbol{\alpha}_j[2k-1:2k][1]| \leq 20$ due to the fact that $\boldsymbol{\alpha}_j[2k-1:2k]$ remains fixed from a time point $j^* \leq p$ such that $\boldsymbol{y}_{j^*}[2k-1:2k] = [1,1]^T$. Further we have that

$$\begin{aligned}
\|\tilde{\boldsymbol{\alpha}}_j[2k-1:2k]\|_2^2 &= (\boldsymbol{\alpha}_j[2k-1:2k][2])^2 && \text{(F.263)}\\
&\quad + (\boldsymbol{\alpha}_j[2k-1:2k][1] + (\boldsymbol{y}_j[2k-1:2k][2] - \boldsymbol{y}_j[2k'-1:2k'][2])\times && \text{(F.264)}\\
&\quad \boldsymbol{\alpha}_j[2k-1:2k][2])^2 && \text{(F.265)}\\
&\leq 2\left(\boldsymbol{\alpha}_j[2k-1:2k][2] + \boldsymbol{\alpha}_j[2k-1:2k][1]\right)^2 && \text{(F.266)}\\
&\quad + 2\left((\boldsymbol{y}_j[2k-1:2k][2] - \boldsymbol{y}_j[2k'-1:2k'][2])\boldsymbol{\alpha}_j[2k-1:2k][2]\right)^2 && \text{(F.267)}\\
&\leq 584, && \text{(F.268)}
\end{aligned}$$

where the last line is due to Eq.(F.107) and (F.262).

Let's represent $\boldsymbol{\mu} \in \mathbb{R}^{2d}$ such that $\boldsymbol{\mu}[2k'-1:2k'] = \boldsymbol{\alpha}[2k'-1:2k']$ and $\boldsymbol{\mu}[2k-1:2k] = \boldsymbol{\alpha}[2k-1:2k]$ for all other $k \in [d]$.

Thus within the sub-bin $[p,q]$, we have that $|\boldsymbol{\mu}^T[2k-1:2k]\boldsymbol{y}_j[2k'-1:2k']| \leq 20$ for all $k \in [d]$. Further, due to Eq.(F.268) we have that $\|\boldsymbol{\mu}\|_2^2 \leq 584d$. Hence we can use a base expert that starts at time p which gives the co-variate $\boldsymbol{y}_j[2k'-1:2k']$ to all coordinates where $j \in [p,q]$. Note that the sub-bin $[p,q]$ must have been resulted via a splitting across coordinate k' at time p. So by the calls to `AssignCo-variatesAndSlopes1` or `AssignCo-variatesAndSlopes2` we set $\boldsymbol{y}_p[2k'-1:2k'] = [1,1]^T$. Thus there exists a base expert in FLH-SIONS (Fig.7.3) that provides the co-variate $\boldsymbol{y}_j[2k'-1:2k']$ to all coordinates where $j \in [p,q]$.

This expert will have a regret of $\tilde{O}(d)$ against $\boldsymbol{\mu}$ via Lemma 222. By using Strong Adaptivity from Corollary 223 (set $\boldsymbol{w} = \boldsymbol{\mu}$ there and recall that $\|\boldsymbol{\mu}\|_2^2 \leq 584d$) and adding the regret across all $6d$ sub-bins of $[i_s, i_t]$ lead to an $\tilde{O}(d^2)$ on T_1 in Eq.(F.204). Thus for any bin in $\mathcal{P}$ produced by generate bins procedure, we have its dynamic regret bounded by $\tilde{O}(d^2)$.

□

***Proof of Theorem** 63.* The proof is now complete by adding the $\tilde{O}(d^2)$ dynamic regret bound across all $O(n^{1/5}C_n^{2/5} \vee 1)$ bins in $\mathcal{P}$ from Corollary 224.

□

The proof of Lemma 71 is same as that of the lemma below, albeit with slightly different notations for $\boldsymbol{X}_j$.

Lemma 226. *Let $\boldsymbol{X}_j$ be as defined in Eq.(F.205). Let $\tilde{f}_j(\boldsymbol{v}) = f_j(\boldsymbol{X}_j\boldsymbol{v})$ for some $\boldsymbol{v} \in \mathbb{R}^{2d}$ and let $\boldsymbol{\Sigma} := \boldsymbol{X}_j^T\boldsymbol{X}_j \in \mathbb{R}^{2d\times 2d}$. We have,*

$$\nabla^2 \tilde{f}_j(\boldsymbol{v}) \preccurlyeq \boldsymbol{\Sigma} \tag{F.269}$$

Proof. We have,

$$\tilde{f}_j(\boldsymbol{v}) = f_j\left(\langle \boldsymbol{y}_j[1:2], \boldsymbol{v}[1:2]\rangle, \ldots, \langle \boldsymbol{y}_j[2d-1:2d], \boldsymbol{v}[2d-1:2d]\rangle\right). \tag{F.270}$$

Let

$$f''_{jk} := \nabla^2 f_j\left(\langle \boldsymbol{y}_j[1:2], \boldsymbol{v}[1:2]\rangle, \ldots, \langle \boldsymbol{y}_j[2d-1:2d], \boldsymbol{y}_j[2d-1:2d]\rangle\right)[j][k], \tag{F.271}$$

be the Hessian of f evaluated at the vector
$[\langle \boldsymbol{y}[1:2], \boldsymbol{v}[1:2]\rangle, \ldots, \langle \boldsymbol{y}[2d-1:2d], \boldsymbol{v}[2d-1:2d]\rangle]^T \in \mathbb{R}^d$.

By straightforward calculations, we obtain

$$\nabla^2 \tilde{f}_j(\boldsymbol{v}) = \begin{bmatrix} f''_{11}\boldsymbol{y}_j[1:2]\boldsymbol{y}_j[1:2]^T & \ldots & f''_{1d}\boldsymbol{y}_j[1:2]\boldsymbol{y}_j[2d-1:2d]^T \\ \vdots & \ddots & \vdots \\ f''_{d1}\boldsymbol{y}_j[2d-1:2d]\boldsymbol{y}_j[1:2]^T & \ldots & \boldsymbol{f}''_{dd}\boldsymbol{y}_j[2d-1:2d]\boldsymbol{y}_j[2d-1:2d]^T \end{bmatrix}, \tag{F.272}$$

Let $\boldsymbol{I} \in \mathbb{R}^{d\times d}$ be the identity matrix and $\boldsymbol{1} \in \mathbb{R}^{2\times 2}$ be the matrix of all ones. Further let's denote $\boldsymbol{b} := [\langle \boldsymbol{y}[1:2], \boldsymbol{v}[1:2]\rangle, \ldots, \langle \boldsymbol{y}[2d-1:2d], \boldsymbol{v}[2d-1:2d]\rangle]^T$ We can succinctly write:

$$\boldsymbol{\Sigma} - \nabla^2 \tilde{f}_j(\boldsymbol{v}) = \left(\left(\boldsymbol{I} - \nabla^2 f(\boldsymbol{b})\right) \otimes \boldsymbol{1}\right) \circ \boldsymbol{y}_j\boldsymbol{y}_j^T, \tag{F.273}$$

where $\otimes$ denotes the Kronecker product and $\circ$ denotes the Hadamard product.

Recall that the loss functions f_j are 1-gradient Lipschitz. So we have $(\boldsymbol{I} - \nabla^2 f(\boldsymbol{b}))$ is Positive Semi Definite (PSD). The matrices $\boldsymbol{1}$ and $\boldsymbol{y}_j\boldsymbol{y}_j^T$ are also PSD. Since both Kronecker and Hadamard products preserves positive semidefiniteness, we have $\nabla^2 \tilde{f}_j(\boldsymbol{v}) \preccurlyeq \boldsymbol{\Sigma}$ which proves the lemma.

□

Proposition 227. *Consider the sequence class $\mathcal{TV}^1(C_n)$ as per Eq.(7.1). Under Assumption A1 (see Section 7.2) we have that $\mathcal{TV}^{(1)}(C_n) \subseteq \mathcal{TV}^{(0)}(2C_n + 20d)$.*

Proof. We start by considering a 1D setting. Consider a sequence $w_{1:n} \in \mathcal{TV}^{(1)}(C_n)$. We can represent it as sum (point-wise) of two sequences as

$$w_{1:n} = p_{1:n} + q_{1:n}, \tag{F.274}$$

where $q_{1:n} = \boldsymbol{\beta}^T \boldsymbol{x}_t$ where $\boldsymbol{x}_t = [1, t]^T$ and $\boldsymbol{\beta}$ is the least square fit coefficnts computed by using covariates $\boldsymbol{x}_t$ and labels w_t, $t \in [n]$. Here the $p_{1:n}$ is the residual sequence obtained by subtracting the least square fit sequence from the true sequence.

Following the terminology in Lemma 211, we can represent $p_t = tM_t + C_t$. Further, due to Eq.(F.29) (with $a = 1$) we have that $p_{t+1} - p_t = M_{t+1}$.

Applying triangle inequality to Eq.(F.274) we have

$$\|Dw_{1:n}\|_1 \leq \|Dp_{1:n}\|_1 + \|Dq_{1:n}\|_1. \tag{F.275}$$

Further,

$$\begin{aligned}
\|Dp_{1:n}\|_1 &= \sum_{t=2}^{n} |M_t| && \text{(F.276)}\\
&= \sum_{t=2}^{n} \left| M_1 + \sum_{j=1}^{t-1} M_{j+1} - M_j \right| && \text{(F.277)}\\
&\leq \sum_{t=2}^{n} |M_1| + D^2 \|p_{1:n}\|_1 && \text{(F.278)}\\
&=_{(a)} n|M_1| + nD^2 \|w_{1:n}\|_1 && \text{(F.279)}\\
&\leq_{(b)} 2nD^2 \|w_{1:n}\|_1, && \text{(F.280)}
\end{aligned}$$

where in line (a) we used the fact that $\|D^2 p_{1:n}\|_1 = \|D^2 w_{1:n}\|_1$ as subtracting a linear sequence doesn't affect the TV1 distance. In line (b) we applied $|M_1| \leq \|D^2 w_{1:n}\|_1$ as shown in Lemma 211.

It remains to bound $\|Dq_{1:n}\|_1$. For this we note that $\|q_t\| \leq 10$ for all $t \in [n]$ due to Eq.(F.175). Since $q_{1:n}$ is a monotonic sequence we have that its variation $\|Dq_{1:n}\|_1 \leq 20$.

Thus overall we obtain that

$$\begin{aligned}
\|Dw_{1:n}\|_1 &\leq 2nD^2 \|w_{1:n}\|_1 + 20 && \text{(F.281)}\\
&\leq 2C_n + 20. && \text{(F.282)}
\end{aligned}$$

For multiple dimensions we apply the same argument across each dimension and add them up to yield the lemma.

□

F.2 Proof of Proposition 65

In this section, we first prove the following result.

Theorem 228. *Let $\boldsymbol{p}_t$ be the predictions of FLH-SIONS algorithm with parameters $\epsilon = 2$,*

$C = 20$ *and exp-concavity factor* σ*. Under Assumptions A1-A4, we have that,*

$$\sum_{t=1}^{n} f_t(\boldsymbol{p}_t) - f_t(\boldsymbol{w}_t) = \tilde{O}(d^2 n^{1/3} C_n^{2/3} \vee d^2),$$

for any $C_n > 0$ *and any comparator sequence* $\boldsymbol{w}_{1:n} \in \mathcal{TV}^{(0)}(C_n)$*. Here* $\tilde{O}$ *hides poly-logarithmic factors of* n *and* $a \vee b = \max\{a, b\}$*.*

Proof. The proof follows almost directly from the arguments in [65]. First, we use the partition $\mathcal{P}$ mentioned in Lemma 30 in [65]. Let the partition be $cP = \{[1_s, 1_t], \ldots, [M_s, M_t]\}$, with $|\mathcal{P}| = M$.

Consider the following convex optimization problem.

$$\min_{\tilde{\boldsymbol{u}}_1, \ldots, \tilde{\boldsymbol{u}}_n, \tilde{\boldsymbol{z}}_1, \ldots, \tilde{\boldsymbol{z}}_{n-1}} \quad \sum_{t=1}^{n} f_t(\tilde{\boldsymbol{u}}_t) \tag{F.283a}$$

$$\text{s.t.} \quad \tilde{\boldsymbol{z}}_t = \tilde{\boldsymbol{u}}_{t+1} - \tilde{\boldsymbol{u}}_t \; \forall t \in [n-1], \tag{F.283b}$$

$$\sum_{t=1}^{n-1} \|\tilde{\boldsymbol{z}}_t\|_1 \leq C_n, \tag{F.283c}$$

$$\|\tilde{\boldsymbol{u}}_t\|_\infty \leq B \; \forall t \in [n], \tag{F.283d}$$

Let $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ be the optimal solution to the above problem. Let $\boldsymbol{w}_j$ be the prediction of the FLH-SIONS algorithm at time j. Define:

$$R_n(C_n) = \sum_{t=1}^{n} f_j(\boldsymbol{w}_t) - f_t(\boldsymbol{u}_t). \tag{F.284}$$

Define $\bar{\boldsymbol{u}}_i = \frac{1}{n_i} \sum_{j=i_s}^{i_t} \boldsymbol{u}_j$ and $\dot{\boldsymbol{u}}_i = \bar{\boldsymbol{u}}_i - \frac{1}{n_i} \sum_{j=i_s}^{i_t} \nabla f_j(\bar{\boldsymbol{u}}_i)$. We can use the regret decomposition of [65].

$$R_n(C_n) \leq \sum_{i=1}^{M} \underbrace{\sum_{j=i_s}^{i_t} f_j(\boldsymbol{w}_j) - f_j(\dot{\boldsymbol{u}}_i)}_{T_{1,i}} + \sum_{i=1}^{M} \underbrace{\sum_{j=i_s}^{i_t} f_j(\dot{\boldsymbol{u}}_i) - f_j(\bar{\boldsymbol{u}}_i)}_{T_{2,i}} + \sum_{i=1}^{M} \underbrace{\sum_{j=i_s}^{i_t} f_j(\bar{\boldsymbol{u}}_i) - f_j(\boldsymbol{u}_j)}_{T_{3,i}}. \tag{F.285}$$

For any bin $[i_s, i_t] \in \mathcal{P}$, we can bound $T_{2,i} + T_{3,i} = O(1)$ by using the arguments in the proof of Theorem 14 of [65] since the losses in our case are also gradient-Lipschitz as per Assumption A3. So we only need to consider the term $T_{1,i}$. Observe that

$$\|\dot{\boldsymbol{u}}_i\|_\infty \le \|\bar{\boldsymbol{u}}_i\|_\infty + \frac{1}{n_i}\sum_{j=i_s}^{i_t} \|\nabla f_j(\bar{\boldsymbol{u}}_i)\|_\infty \tag{F.286}$$

$$\le 2, \tag{F.287}$$

as per Assumptions A1-A2. Further we can view the comparator $\dot{\boldsymbol{u}}_i$ as a linear predictor with slope zero. The output of this linear predictor is bounded in magnitude by 2 which is less that 20. Hence FLH-SIONS under the setting of the current theorem leads to $T_{1,i} = \tilde{O}(d)$. Since $M = O(dn^{1/3}C_n^{2/3} \vee d)$ for the partition in Lemma 30 of adding the regret across all bins results in the theorem. □

Theorem228 when combined with Theorem 63 now directly leads to Proposition 65.

F.3 Proof of Proposition 62

The result proven in this section is mainly due to the geometric arguments in [20, 2] (or see [18] for a comprehensive monograph) with an extra technicality of handling boundedness constraint as per Assumption A1 (in Section 7.2).

In the proof we make extensive use of wavelet theory and refer readers to [18] for necessary preliminaries.

Proposition 62. *Under Assumptions A1-A4, any online algorithm necessarily suffers* $\sup_{\boldsymbol{w}_{1:n}\in\mathcal{TV}^{(1)}(C_n)} R_n(\boldsymbol{w}_{1:n}) = \Omega(d^{3/5}n^{1/5}C_n^{2/5} \vee d)$.

Proof. We consider a uni-variate setting with the losses $f_t(w) = (d_t - w)^2$ where $d_t = u_t + \mathcal{N}(0,1)$ with $u_{1:n} \in \mathcal{TV}^{(1)}(C_n)$. At each step, d_t is revealed to the learner as doing so can only make learning easier.

Let $\mathbb{W}$ be the set of whole numbers. For the purposes of analysis, we start with an abstract observation model:

$$y_j = \theta_j + \epsilon\mathcal{N}(0,1),\ j \in \mathbb{W} \tag{F.288}$$

where θ_j are the wavelet coefficients in a regularity-three CDJV multi-resolution basis [41] of a function in $\mathcal{F}_1(C_n)$ from which the discrete samples $u_{1:n}$ are generated.

In what follows we will show that for any procedure estimating the wavelet coefficients (let the estimate be $\hat{\theta}_j$, $j \in \mathbb{W}$) we have that

$$\sum_{j\in\mathbb{W}} (\hat{\theta}_j - \theta_j)^2 = \Omega(C^{2/5}\epsilon^{8/5}). \tag{F.289}$$

Due to Section 15.5 of [18], by taking $\epsilon = 1/\sqrt{n}$, such a guarantee will then imply a lower bound of $\Omega(n^{-4/5}C^{2/5})$ for $\frac{1}{n}\sum_{t=1}^n (u_t - \hat{u}_t)^2$, where $\hat{u}_t$ is the estimate produced by observing the data d_t (assume $C = \Omega(1/\sqrt{n})$ for now). This rate will finally imply a dynamic regret lower bound in the following manner:

$$E\left[\sup_{r_{1:n}\in\mathcal{TV}^{(1)}(C)} \sum_{t=1}^{n} f_t(\hat{u}_t) - f_t(r_t)\right] \geq \sup_{r_{1:n}\in\mathcal{TV}^{(1)}(C)} E\left[\sum_{t=1}^{n} f_t(\hat{u}_t) - f_t(r_t)\right] \tag{F.290}$$

$$=_{(a)} \sup_{r_{1:n}\in\mathcal{TV}^{(1)}(C)} \sum_{t=1}^{n} E[(\hat{u}_t - u_t)^2] - (r_t - u_t)^2 \tag{F.291}$$

$$= \sum_{t=1}^{n} E[(\hat{u}_t - u_t)^2], \tag{F.292}$$

where in line (a) we used the bias variance decomposition and the fact that $\hat{u}_t$ is independent of d_t for online algorithms.

In what follows we use a dyadic indexing scheme for referring to wavelet coefficients in Eq.(F.288) as θ_{jk} which means the k^{th} wavelet coefficient in resolution $j \geq 0$. There are 2^j wavelet coefficients in resolution j. We will also use $\theta_{j\cdot}$ to denote a sequence of 2^j wavelet coefficients at resolution j.

Let β be the subset of wavelet coefficients at resolutions less than or equal to 2. i.e, $\beta = [\theta_{0\cdot}, \theta_{1\cdot}, \theta_{2\cdot}]$ which has a length of 7.

Define a Besov norm as follows:

$$\|\theta\|_{b^{3/2}_{1,1}} := \|\beta\|_1 + \sum_{j\geq 3} 2^{3j/2}\|\theta_{j\cdot}\|_1. \tag{F.293}$$

Define a Besov space as:

$$\mathcal{A}(B) := \{\theta : \|\theta\|_{b^{3/2}_{1,1}} \leq B\}. \tag{F.294}$$

It is known that $A(\kappa C) \subseteq \mathcal{F}_1(C)$ for some constant $0 < \kappa \leq 1$. (see for eg. Eq.(33) in [80] along with Theorem 1 in [2]).

Since the space $\mathcal{A}(B)$ is solid and orthosymmetric (see Section 4.8 in [18]) we have that the risk of estimating coefficients from $\mathcal{A}$ is lower bounded by the risk (i.e $\sum_{j\geq 0}(\hat{\theta}_j - \theta_j)^2$) of the hardest rectangular sub-problem as shown by [20].

A hyper-rectangle is defined as follows:

$$\Theta(\tau) = \{\theta : |\theta_j| \leq \tau_j,\ j \geq 0\}. \tag{F.295}$$

From [20], the minimax risk over a hyper-rectangle under the observation model

Eq.(F.288) is known to be:

$$R^*(\tau) := \min_{\hat{\theta}} \max_{\theta \in \Theta(\tau)} \sum_{j \geq 0} (\hat{\theta}_j - \theta)^2 \tag{F.296}$$

$$\geq \sum_{j \geq 0} \min\{\tau_j^2, \epsilon^2\}. \tag{F.297}$$

So all we need to show is an appropriate hyper-rectangle (which is identified by τ) within $\mathcal{A}(B)$ whose minimax risk is sufficiently large.

We next proceed to give such a hyper-rectangle. Let $j_* \in \mathbb{W}$ be the smallest number such that

$$2^{j_*} \geq \frac{C^{2/5}}{\epsilon^{2/5}}. \tag{F.298}$$

For simplicity, from now on-wards, let's assume that j_* is an integer that satisfy $2^{j_*} = \frac{C^{2/5}}{\epsilon^{2/5}}$.

Define the hyper-rectangle coordinates by

$$\tau_{j_* k} = \frac{\kappa C}{2^{5j_*/2}}, \tag{F.299}$$

for all $k = 0, 1, \ldots, 2^{j_*} - 1$ and $\tau_{j\cdot} = 0$ for all other resolutions.

Note that $\frac{\kappa C}{2^{5j_*/2}} = \epsilon$. The minimax risk over such a hyper-rectangle then becomes

$$R^*(\tau) = 2^{j_*}\epsilon \tag{F.300}$$

$$= (\kappa C)^{2/5} \epsilon^{8/5}. \tag{F.301}$$

Now it remains to verify that

1. The hyper-rectangle in Eq.(F.299) is indeed in $\mathcal{A}(\kappa C)$.
2. The function produced by the coefficients in that hyper rectangle is bounded by 1 point-wise in magnitude.

First we notice that by taking $\epsilon = 1/\sqrt{n}$ as mentioned earlier, we have

$$2^{j_*} > 4, \tag{F.302}$$

whenever $C > 4^{5/2}/\sqrt{n}$. We first consider the case where C is within this regime.

For the first item, we have that

$$\|\tau\|_{b_{1,1}^{3/2}} = 2^{3j_*/2} \cdot 2^{j_*} \frac{\kappa C}{2^{5j_*/2}} \tag{F.303}$$

$$= \kappa C, \tag{F.304}$$

where we used the fact that $j_* > 2$ in the regime $C > 4^{5/2}/\sqrt{n}$.

Hence $\Theta(\tau) \subseteq \mathcal{A}(\kappa C)$.

For the second item, we notice that due to Lemma B.18 in [18], it is sufficient to show that $2^{j_*/2}\|\theta_{j_*\cdot}\|_\infty = O(1)$. Taking $\epsilon = 1/\sqrt{n}$ as mentioned earlier, we have that

$$2^{j_*/2}\|\theta_{j_*\cdot}\|_\infty = \frac{\kappa C}{2^{2j_*}} \tag{F.305}$$

$$= \kappa^{1/5} C^{1/5} \epsilon^{4/5} \tag{F.306}$$

$$= \frac{\kappa^{1/5} C^{1/5}}{n^{2/5}} \tag{F.307}$$

$$\leq 1, \tag{F.308}$$

in the non-trivial regime of $C \leq n^2$ where we recall that $\kappa \leq 1$.

For the regime where $C \leq 1/\sqrt{n}$, the trivial lower bound of $\Omega(1)$ estimation error kicks in. Thus overall we have shown that for any online algorithm producing estimates $\hat{u}_t$ we have that

$$\sum_{t=1}^{n} E[(\hat{u}_t - u_t)^2] = \Omega(n^{1/5} C^{2/5} \vee 1), \tag{F.309}$$

thus obtaining a lower bound on the dynamic regret as per Eq.(F.292).

In multiple-dimensions we can consider a similar setup as before with losses $f_t(\boldsymbol{w}) = \|\boldsymbol{d}_t - \boldsymbol{w}\|_2^2$ with $\boldsymbol{d}_t[k] = \boldsymbol{u}_t[k] + \mathcal{N}(0,1)$ where $\boldsymbol{u}_{1:n} \in \mathcal{TV}^{(1)}(C)$. We can consider a sequence $\boldsymbol{u}_{1:n}$ such that $\|nD^2 \boldsymbol{u}_{1:n}[k]\|_1 = C/d$ across each coordinate $k \in [d]$.

$$\min_{\boldsymbol{p}_{1:n}} \max_{\boldsymbol{w}_{1:n} \in \mathcal{TV}^{(1)}(C)} \sum_{t=1}^{n} f_t(\boldsymbol{p}_t) - f_t(\boldsymbol{w}_t) = \sum_{k=1}^{d} \sum_{t=1}^{n} \Omega(n^{1/5} (C/d)^{2/5} \vee 1) \tag{F.310}$$

$$= \Omega(d^{3/5} n^{1/5} C^{2/5} \vee d). \tag{F.311}$$

This completes the proof of the proposition.

□

Appendix G

Supplementary Materials for Chapter 8

G.1 Omitted proofs from Section 8.2

In the next two lemmas, we verify Assumption 2 for some important loss functions.

Lemma 229 (cross-entropy loss)**.** *Consider a sample $(x, y) \sim Q$. Let $p \in \mathbb{R}_+^K$ and $\tilde{p}(x) \in \Delta_K$ be a distribution that assigns a weight proportional $\frac{p(i)}{q_0(i)} f_0(i|x)$ to the label i . Let $\ell(\tilde{p}(x), y) = \sum_{i=1}^K \mathbb{I}\{y = i\} \log(1/p(x)[i])$ be the cross-entropy loss. Let $L(p) := E_{(x,y)\sim Q}[\ell(p(x), y)]$ be its population analogue. Then $L(p)$ is $2\sqrt{K}/\mu$ Lipschitz in $\|\cdot\|_2$ norm over the clipped box $\mathcal{D} := \{p \in \mathbb{R}_+^K : \mu \le p(i) \le 1\ \forall i \in [K]\}$ which is compact and convex. Further, the true marginals $q_t \in \mathcal{D}$ whenever $q_t(i) \ge \mu$ for all $i \in [K]$.*

Proof. We have

$$L(p) = -\sum_{i=1}^{K} E[E[Q_t(i|x) \log(\tilde{p}(x)[i])|x]] \tag{G.1}$$

$$= E[\log(\sum_{i=1}^{K} w_i p(i))] - \sum_{i=1}^{K} E[E[Q_t(i|x) \log(w_i p(i))|x]], \tag{G.2}$$

where we define $w_i := f_0(i|x)/q_0(i)$. Then we can see that

$$\nabla L(p)[i] = E\left[\frac{w_i}{\sum_{j=1}^K w_j p(j))}\right] - E\left[\frac{Q_t(i|x)}{p(i)}\right]. \tag{G.3}$$

So if $\min_i p(i) \ge \mu$, we have that $\frac{w_i}{\sum_{j=1}^K w_j p(j)} \le 1/\mu$ and $Q_t(i|x)/p(i) \le 1/\mu$. So by triangle inequality, $|\nabla L(p)[i]| \le 1/\mu + 1/\mu$.

□

Lemma 230 (binary 0-1 loss)**.** *Consider a sample $(x, y) \sim Q$. Let $p \in \mathbb{R}_+^K$ and $\tilde{p}(x) \in \Delta_K$ be a distribution that assigns a weight proportional $\frac{p(i)}{q_0(i)} f_0(i|x)$ to the label i. Let $\hat{y}(x)$ be a sample obtained from the distribution $\tilde{p}(x)$. Consider the binary 0-1 loss $\ell(\hat{y}(x), y) = \mathbb{I}(\hat{y}(x) \neq y)$. Let $L(p) := E_{(x,y)\sim Q, \hat{y}(x)\sim\tilde{p}(x)} I(\hat{y}(x) \neq y)$ be its population analogue. Let $q_0(i) \geq \alpha > 0$. Then $L(p)$ is $2K^{3/2}/(\alpha\tau)$ Lipschitz in $\|\cdot\|_2$ norm over the domain $\mathcal{D} := \{p \in \mathbb{R}_+^K : \sum_{i=1}^K p(i) f_0(i|x) \geq \tau,\ p(i) \leq 1\ \forall i \in [K]\}$ which is compact and convex. Further, the true marginals $q_t \in \mathcal{D}$ whenever $q_t(i) \geq \mu$ for all $i \in [K]$.*

Proof. We have that

$$L(p) = \sum_{i=1}^{K} E[Q(y \neq i|x)\tilde{p}(x)[i]]. \tag{G.4}$$

Denote $\tilde{p}(x)[i] = p(i)w_i / \sum_{j=1}^K p(j)w_j$ with $w_j := f_0(i|x)/q_0(i)$. Then we see that

$$\left|\frac{\partial \tilde{p}(x)[i]}{\partial p(i)}\right| = \left| \frac{w_i}{\sum_{j=1}^K w_j p(j)} - \frac{(w_i p(i))w_i}{\left(\sum_{j=1}^K w_j p(j)\right)^2} \right| \tag{G.5}$$

$$\leq \frac{1}{\alpha\tau} + \frac{w_i}{\sum_{j=1}^K w_j p(j)} \tag{G.6}$$

$$\leq 2/(\alpha\tau). \tag{G.7}$$

Similarly,

$$\left|\frac{\partial \tilde{p}(x)[i]}{\partial p(j)}\right| = \frac{w_i p(i) w_j}{\left(\sum_{j=1}^K w_j p(j)\right)^2} \tag{G.8}$$

$$\leq 1/(\alpha\tau). \tag{G.9}$$

Thus we conclude that $\|\nabla \tilde{p}(x)[i]\|_2 \leq 2\sqrt{K}/(\alpha\tau)$, where the gradient is taken with respect to $p \in \mathbb{R}_+^K$.

Therefore,

$$\|\nabla L(p)\|_2 \leq \sum_{i=1}^{K} \|\nabla \tilde{p}(x)[i]\|_2 \tag{G.10}$$

$$\leq 2K^{3/2}/(\alpha\tau). \tag{G.11}$$

□

Remark 231. The condition $\sum_{i=1}^K f_0(i|x)p(i) \geq \tau$ is closely related to Condition 1 of [95]. Note that this is strictly weaker than imposing the restriction that the distribution $p(i) \geq \mu$ for each i.

Remark 232. We emphasize that the conditions in Lemmas 229 and 230 are only sufficient conditions that imply bounded gradients. However, they are not necessary for satisfying bounded gradients property.

Lemma 233. *Let $\mu, \nu \in \Delta_K$ be such that $\mu[i] = q_t(i)$. Let $s_t = C^{-1} f_0(x_t)$, where C is the confusion matrix defined in Assumption 1. We have that $E[s_t] = \mu$ and $Var(s_t) \leq 1/\sigma^2_{min}(C)$*

Proof. Let $\tilde{q}_t(\hat{y}_t) = E_{x_t \sim Q_t^X, \hat{y}(x_t) \sim f_0(x_t)} \mathbb{I}\{\hat{y}(x_t) = \hat{y}_t\}$ be the probability that the classifier f_0 predicts the label $\hat{y}_t$. Here $Q_t^X(x) := \sum_{i=1}^K Q_t(x, i)$. Let's denote $Q_t(\hat{y}(x_t) = \hat{y}_t | y_t = i) := E_{x_t \sim Q_t(\cdot | y=i), \hat{y}(x_t) \sim f_0(x_t)} \mathbb{I}\{\hat{y}(x_t) = \hat{y}_t\}$. By law of total probability, we have that

$$\tilde{q}_t(\hat{y}_t) = \sum_{i=1}^K Q_t(\hat{y}(x_t) = \hat{y}_t | y_t = i) q_t(i) \tag{G.12}$$

$$= \sum_{i=1}^K Q_0(\hat{y}(x_t) = \hat{y}_t | y_t = i) q_t(i), \tag{G.13}$$

where the last line follows by the label shift assumption.

Let $\mu, \nu \in \mathbb{R}^K$ be such that $\mu[i] = q_t(i)$ and $\nu[i] = \tilde{q}_t(i)$. Then the above equation can be represented as $\nu = C\mu$. Thus $\mu = C^{-1}\nu$.

Given a sample $x_t \in Q_t$, the vector $f_0(x_t)$ forms an unbiased estimate of ν. Hence we have that the vector $\hat{\mu} := C^{-1} f_0(x_t)$ is an unbiased estimate of μ. Moreover,

$$\|\hat{\mu}\|_2 \leq \|C^{-1}\|_2 \|f_0(x_t)\| \tag{G.14}$$

$$\leq 1/\sigma_{min}(C). \tag{G.15}$$

Hence the variance of the estimate $\hat{\mu}$ is bounded by $1/\sigma^2_{min}(C)$.

□

We have the following performance guarantee for online regression due to [71].

Proposition 234 ([71])**.** *Let $s_t = C^{-1} f_0(x_t)$. Let $\hat{q}_t := ALG(s_{1:t-1})$ be the online estimate of the true label marginal q_t produced by the Aligator algorithm by taking $s_{1:t-1}$ as input at a round t. Then we have that*

$$\sum_{t=1}^T E\left[\|\hat{q}_t - q_t\|_2^2\right] = \tilde{O}(K^{1/3} T^{1/3} V_T^{2/3} (1/\sigma^{4/3}_{min}(C)) + K), \tag{G.16}$$

where $V_T := \sum_{t=2}^T \|q_t - q_{t-1}\|_1$. Here $\tilde{O}$ hides dependencies in absolute constants and poly-logarithmic factors of the horizon. Further this result is attained without prior knowledge of the variation V_T.

By following the arguments in [65], a similar statement can be derived also for the FLH-FTL algorithm of [23] (Algorithm 9).

Theorem 73. *Suppose we run Algorithm 2 with the online regression oracle ALG as FLH-FTL (App. G.4) or Aligator [71]. Then under Assumptions 1 and 2, we have*

$$E[R_{dynamic}(T)] = \tilde{O}\left(\frac{K^{1/6}T^{2/3}V_T^{1/3}}{\sigma_{min}^{2/3}(C)} + \frac{\sqrt{KT}}{\sigma_{min}(C)}\right), \tag{8.3}$$

where $V_T := \sum_{t=2}^{T} \|q_t - q_{t-1}\|_1$ and the expectation is taken with respect to randomness in the revealed co-variates. Further, this result is attained without prior knowledge of V_T.

Proof. Owing to our carefully crafted reduction from the problem of online label shift to online regression, the proof can be conducted in just a few lines. Let $\tilde{q}_t$ be the value of $\mathrm{ALG}(s_{1:t-1})$ computed at line 2 of Algorithm 2. Recall that the dynamic regret was defined as:

$$R_{\text{dynamic}}(T) = \sum_{t=1}^{T} L_t(\hat{q}_t) - L_t(q_t) \leq \sum_{t=1}^{T} G\|\hat{q}_t - q_t\|_2 \tag{G.17}$$

Continuing from Eq.(I.27), we have

$$\begin{aligned}
E[R_{\text{dynamic}}(T)] &\leq \sum_{t=1}^{T} G \cdot E[\|\hat{q}_t - q_t\|_2] && \text{(G.18)}\\
&\leq \sum_{t=1}^{T} G \cdot E[\|\tilde{q}_t - q_t\|_2] && \text{(G.19)}\\
&\leq \sum_{t=1}^{T} G\sqrt{E\|\tilde{q}_t - q_t\|_2^2} && \text{(G.20)}\\
&\leq G\sqrt{T\sum_{t=1}^{T} E[\|\tilde{q}_t - q_t\|_2^2]} && \text{(G.21)}\\
&= \tilde{O}\left(K^{1/6}T^{2/3}V_T^{1/3}(1/\sigma_{min}^{2/3}(C)) + \sqrt{KT}/\sigma_{min}(C)\right), && \text{(G.22)}
\end{aligned}$$

where the second line is due to non-expansivity of projection, the third line is due to Jensen's inequality, fourth line by Cauchy-Schwartz and last line by Proposition 234. This finishes the proof.

□

Next, we provide matching lower bounds (modulo log factors) for the regret in the unsupervised label shift setting. We start from an information-theoretic result which will play a central role in our lower bound proofs.

Proposition 235 (Theorem 2.2 in [17]). *Let $\mathbb{P}$ and $\mathbb{Q}$ be two probability distributions on $\mathcal{H}$, such that $KL(\mathbb{P}||\mathbb{Q}) \leq \beta < \infty$, Then for any $\mathcal{H}$-measurable real function $\phi : \mathcal{H} \to \{0, 1\}$,*

$$\max\{\mathbb{P}(\phi = 1), \mathbb{Q}(\phi = 0)\} \geq \frac{1}{4}\exp(-\beta). \tag{G.23}$$

Theorem 76. *Let $V_T \leq 64T$. There exists a loss function, a domain $\mathcal{D}$ (in Assumption 2), and a choice of adversarial strategy for generating the data such that for any algorithm, we have $\sum_{t=1}^{T} E([L_t(\hat{q}_t)] - L_t(q_t)) = \Omega\left(\max\{T^{2/3}V_T^{1/3}, \sqrt{T}\}\right)$, where $\hat{q}_t \in \mathcal{D}$ is the weight estimated by the algorithm and $q_t \in \mathcal{D}$ is the label marginal at round t chosen by the adversary. Here the expectation is taken with respect to the randomness in the algorithm and the adversary.*

Proof. We start with a simple observation about KL divergence. Consider distributions with density $P(x, y) = P_0(x|y)p(y)$ and $Q(x, y) = P_0(x|y)q(y)$ where $(x, y) \in \mathbb{R} \times [K]$. Note that these distributions are consistent with the label shift assumption. We note that

$$\text{KL}(P||Q) = \sum_{i=1}^{K} \int_{\mathbb{R}} P_0(x|i)p(i) \log\left(\frac{P_0(x|i)p(i)}{P_0(x|i)q(i)}\right) dx \tag{G.24}$$

$$= \sum_{i=1}^{K} \int_{\mathbb{R}} P_0(x|i)p(i) \log\left(\frac{p(i)}{q(i)}\right) dx \tag{G.25}$$

$$= \sum_{i=1}^{K} p(i) \log\left(\frac{p(i)}{q(i)}\right) \tag{G.26}$$

Thus we see that under the label shift assumption, the KL divergence is equal to the KL divergence between the marginals of the labels.

Next, we define a problem instance and an adversarial strategy. We focus on a binary classification problem where the labels is either 0 or 1. As noted before, the KL divergence only depends on the marginal distribution of labels. So we fix the density $Q_0(x|y)$ to be any density such that under the uniform label marginals ($q_0(1) = q_0(0) = 1/2$) we can find a classifier with invertible confusion matrix (recall from Fig. 1 that Q_0 corresponds to the data distribution of the training data set).

Divide the entire time horizon T is divided into batches of size Δ. So there are $M := T/\Delta$ batches (we assume divisibility). Let $\Theta = \left\{\frac{1}{2} - \delta, \frac{1}{2} + \delta\right\}$ be a set of success probabilities, where each probability can define a Bernoulli trial. Here $\delta \in (0, 1/4)$ which will be tuned later.

The problem instance is defined as follows:

- For batch $i \in [M]$, adversary selects a probability $\mathring{q}_i \in \Theta$ uniformly at random.
- For any round t that belongs to the i^{th} batch, sample a label $y_t \sim \text{Ber}(q_t)$ and co-variate $x_t \sim Q_0(\cdot|y_t)$. Here $q_t = \mathring{q}_i$. The co-variate x_t is revealed.
- Let $\hat{q}_t$ be any estimate of q_t at round t. Define the loss as $L_t(\hat{q}_t) := \mathbb{I}\{q_t \geq 1/2\}(1 - \hat{q}_t) + \mathbb{I}\{q_t < 1/2\}\hat{q}_t$.

We take the domain $\mathcal{D}$ in Assumption 2 as $[1/2 - \delta, 1/2 + \delta]$. It is easy to verify that $L_t(\hat{q}_t)$ is Lipschitz over $\mathcal{D}$. Note that unlike [15], we do not have an unbiased estimate of the gradient of loss functions.

Let's compute an upperbound on the total variation incurred by the true marginals. We have

$$\sum_{t=2}^{T} |q_t - q_{t-1}| = \sum_{i=2}^{M} |\mathring{q}_i - \mathring{q}_{i-1}| \tag{G.27}$$

$$\leq 2\delta M \tag{G.28}$$

$$\leq V_T, \tag{G.29}$$

where the last line is obtained by choosing $\delta = V_T/(2M) = V_T\Delta/(2T)$.

Since at the beginning of each batch, the sampling probability is chosen uniformly at random, the loss function in the current batch is independent of the history available at the beginning of the batch. So only the data in the current batch alone is informative in minimising the loss function in that batch. Hence it is sufficient to consider algorithms that only use the data within a batch alone to make predictions at rounds that falls within that batch.

Now we proceed to bound the regret incurred within batch 1. The computation is identical for any other batches.

Let $\mathbb{P}$ be the joint probability distribution in which labels $(y_1, \ldots, y_\Delta)$ within batch 1 are sampled with success probability $1/2 - \delta$ (i.e $q_t = 1/2 - \delta$)

$$\mathbb{P}(y_1, \ldots, y_\Delta) = \Pi_{i=1}^{\Delta}(1/2 - \delta)^{y_i}(1/2 + \delta)^{1-y_i}. \tag{G.30}$$

Define an alternate distribution $\mathbb{Q}$ such that

$$\mathbb{Q}(y_1, \ldots, y_\Delta) = \Pi_{i=1}^{\Delta}(1/2 + \delta)^{y_i}(1/2 - \delta)^{1-y_i}. \tag{G.31}$$

According to the above distribution the data are independently sampled from Bernoulli trials with success probability $1/2 + \delta$. (i.e $q_t = 1/2 + \delta$)

Moving forward, we will show that by tuning Δ appropriately, any algorithm won't be able to detect between these two alternate worlds with constant probability resulting in sufficiently large regret.

We first bound the KL distance between these two distributions. Let

$$\mathrm{KL}(1/2-\delta||1/2+\delta) := (1/2+\delta)\log\left(\frac{1/2+\delta}{1/2-\delta}\right) + (1/2-\delta)\log\left(\frac{1/2-\delta}{1/2+\delta}\right) \tag{G.32}$$

$$\leq_{(a)} (1/2+\delta)\frac{2\delta}{1/2+\delta} - (1/2-\delta)\frac{2\delta}{1/2+\delta} \tag{G.33}$$

$$= \frac{16\delta^2}{1-4\delta^2} \tag{G.34}$$

$$\leq_{(b)} \frac{64\delta^2}{3}, \tag{G.35}$$

where in line (a) we used the fact that $\log(1+x) \leq x$ for $x > -1$ and observed that $-4\delta/(1+2\delta) > -1$ as $\delta \in (0, 1/4)$. In line (b) we used $\delta \in (0, 1/4)$.

Since $\mathbb{P}$ and $\mathbb{Q}$ are product of the marginals due to independence we have that

$$\mathrm{KL}(\mathbb{P}||\mathbb{Q}) = \sum_{t=1}^{\Delta} \mathrm{KL}(1/2-\delta||1/2+\delta) \tag{G.36}$$

$$\leq (64\Delta/3)\cdot\delta^2 \tag{G.37}$$

$$= 16/3 \tag{G.38}$$

$$:= \beta, \tag{G.39}$$

where we used the choices $\delta = \Delta V_T/(2T)$ and $\Delta = (T/V_T)^{2/3}$.

Suppose at the beginning of batch, we reveal the entire observations within that batch $y_{1:\Delta}$ to the algorithm. Note that doing so can only make the problem easier than the sequential unsupervised setting. Let $\hat{q}_t$ be any measurable function of $y_{1:\Delta}$. Define the function $\phi_t := \mathbb{I}\{\hat{q}_t \geq 1/2\}$. Then by Proposition 235, we have that

$$\max\{\mathbb{P}(\phi_t = 1), \mathbb{Q}(\phi_t = 0)\} \geq \frac{1}{4}\exp(-\beta), \tag{G.40}$$

where β is as defined in Eq.(G.39).

Notice that if $q_t = 1/2-\delta$, then $L_t(\hat{q}_t) \geq 1/2$ for any $\hat{q}_t \geq 1/2$. Similarly if $q_t = 1/2+\delta$, we have that $L_t(\hat{q}_t) \geq 1/2$ for any $\hat{q}_t < 1/2$.

Further note that $L_t(q_t) = 1/2 - \delta$ by construction.

For notational clarity define $L_t^p(x) := x$ and $L_t^q(x) := 1 - x$. We can lower-bound the instantaneous regret as:

$$E[L_t(\hat{q}_t)] - L_t(q_t) =_{(a)} \frac{1}{2}(E_{\mathbb{P}}[L^p_t(\hat{q}_t)] - L^p_t(1/2-\delta)) + \frac{1}{2}(E_{\mathbb{Q}}[L^q_t(\hat{q}_t)] - L^q_t(1/2+\delta)) \tag{G.41}$$

$$\geq_{(b)} \frac{1}{2}(E_{\mathbb{P}}[L^p_t(\hat{q}_t)|\hat{q}_t \geq 1/2] - L^p_t(1/2-\delta)\mathbb{P}(\phi_t = 1) \tag{G.42}$$

$$+ \frac{1}{2}(E_{\mathbb{Q}}[L^q_t(\hat{q}_t)|\hat{q}_t < 1/2] - L^q_t(1/2+\delta)\mathbb{Q}(\phi_t = 0) \tag{G.43}$$

$$\geq_{(c)} \frac{1}{2}\delta\mathbb{P}(\phi_t = 1) + \frac{1}{2}\delta\mathbb{Q}(\phi_t = 0) \tag{G.44}$$

$$\geq \delta/2 \max\{\mathbb{P}(\phi_t = 1), \mathbb{Q}(\phi_t = 0)\} \tag{G.45}$$

$$\geq_{(d)} \frac{\delta}{8}\exp(-\beta), \tag{G.46}$$

where in line (a) we used the fact the success probability for a batch is selected uniformly at random from Θ. In line (b) we used the fact that $L^p_t(\hat{q}_t) - L^p_t(1/2-\delta) \geq 0$ since $\hat{q}_t \in \mathcal{D} = [1/2-\delta, 1/2+\delta]$. Similarly term involving L^q_t is also handled. In line (c) we applied $(E_{\mathbb{P}}[L^p_t(\hat{q}_t)|\hat{q}_t \geq 1/2] - L^p_t(1/2-\delta)) \geq \delta$ since $E_{\mathbb{P}}[L^p_t(\hat{q}_t)|\hat{q}_t \geq 1/2] \geq 1/2$ and $L^p_t(1/2-\delta) = 1/2-\delta$. Similar bounding is done for the term involving $E_{\mathbb{Q}}$ as well. In line (d) we used Eq.(G.40).

Thus we get the total expected regret within batch 1 as

$$\sum_{t=1}^{\Delta} E[L_t(\hat{q}_t)] - L_t(q_t) \geq \frac{\delta\Delta}{8}\exp(-\beta) \tag{G.47}$$

The total regret within any batch $i \in [M]$ can be lower bounded using exactly the same arguments as above. Hence summing the total regret across all batches yields

$$\sum_{t=1}^{T} E[L_t(\hat{q}_t)] - L_t(q_t) \geq \frac{T}{\Delta} \cdot \frac{\delta\Delta}{8}\exp(-\beta) \tag{G.48}$$

$$= \frac{V_T\Delta}{16} \cdot \exp(-\beta) \tag{G.49}$$

$$= T^{2/3}V_T^{1/3}\exp(-\beta)/16. \tag{G.50}$$

The $\Omega(\sqrt{T})$ part of the lowerbound follows directly from Theorem 3.2.1 in [59] by choosing $\mathcal{D}$ with diameter bounded by $\Omega(1)$.

□

Algorithm 7 `LPA`: a black-box reduction to produce a low-switching online regression algorithm

Input: Online regression oracle ALG, failure probability δ, maximum standard deviation σ (see Definition 72).

Initialize prev $= 0 \in \mathbb{R}^K$, $b = 1$
Get estimate $\tilde{\theta}_t$ from ALG$(z_{1:t-1})$
Output $\hat{\theta}_t =$ prev
Receive an observation z_t
`// test to detect non-staionarity`
if $\sum_{j=b+1}^{t} \|\text{prev} - \tilde{\theta}_j\|_2^2 > 5K\sigma^2 \log(2T/\delta)$ **then**
  Set $b = t + 1$, prev $= z_t$
  Restart ALG
else if $t - b + 1$ is a power of 2 **then**
  Set prev $= \sum_{j=b}^{t} z_j / t-b+1$
end if
Update ALG with z_t

G.2 Design of low switching online regression algorithms

Even-though Algorithm 4 has attractive performance guarantees, it requires retraining with weighted ERM at every round. This is not satisfactory since the retraining can be computationally expensive. In this section, we aim to design a version of Algorithm 4 with few retraining steps while not sacrificing the statistical efficiency (up to constants). To better understand why this goal is attainable, consider a time window $[1, n] \subseteq [T]$ where the true label marginals remain constant or drift very slowly. Due to the slow drift, one reasonable strategy is to re-train the model (with weighted ERM) using the past data only at time points within $[1, n]$ that are powers of 2 (i.e via a doubling epoch schedule). For rounds $t \in [1, n]$ that are not powers of 2, we make predictions with a previous model h_{prev} computed at $t_{\text{prev}} := 2^{\lfloor \log_2 t \rfloor}$ which is trained using data seen upto the time t_{prev}. Observe that this constitutes at least half of the data seen until round t. This observation when combined with the slow drift of label marginals implies that the performance of the model h_{prev} at round t will be comparable to the performance of a model obtained by retraining using entire data collected until round t.

To formalize this idea, we need an efficient online change-point-detection strategy that can detect intervals where the TV of the *true* label marginals is low and retrain only (modulo at most $\log T$ times within a low TV window) when there is enough evidence for sufficient change in the TV of the true marginals. We address this problem via a two-step approach. In the first step, we construct a generic black-box reduction that takes an online regression oracle as input and converts it into another algorithm with the property that the number of switches in its predictions is controlled without sacrificing the statistical performance. Recall that the purpose of the online regression oracles is to track the true label marginals. The output of our low-switching online algorithm remains the same as long as the TV of the *true* label marginals (TV computed from the time point of the last switch) is sufficiently small. Then we use this low-switching online regression algorithm to re-train the classifier when a switch is detected.

We next provide the **L**ow switching through **P**hased **A**veraging (LPA) (Algorithm 7), our black-box reduction to produce low switching regression oracles. We remark that this algorithm is applicable to the much broader context of *online regression* or *change point detection* and can be of independent interest.

We now describe the intuition behind Algorithm 7. The purpose of Algorithm 7 is to denoise the observations z_t and track the underlying ground truth θ_t in a statistically efficient manner while incurring low switching cost. Hence it is applicable to the broader context of online non-parametric regression [37, 139, 71] and offline non-parametric regression [3, 140].

Algorithm 7 operates by adaptively detecting low TV intervals. Within each time window it performs a phased averaging in a doubling epoch schedule. i.e consider a low TV window $[b, n]$. For a round $t \in [b, n]$ let $t_{\text{prev}} := 2^{\lfloor \log_2 (t-b+1) \rfloor}$. In round t, the algorithm plays the average of the observations $z_{b:t_{\text{prev}}}$. So we see that in any low TV window, the algorithm changes its output only at-most $O(\log T)$ times.

For the above scheme to not sacrifice statistical efficiency, it is important to efficiently detect windows with low TV of the true label marginals. Observe that the quantity `prev` computes the average of at-least half of the observations within a time window that start at time b. So when the TV of the ground truth within a time window $[b, t]$ is small, we can expect the average to be a good enough representation of the entire ground truth sequence within that time window. Consider the quantity $R_t := \sum_{j=b+1}^{t} \|\text{prev} - \theta_j\|_2^2$ which is the total squared error (TSE) incurred by the fixed decision `prev` within the current time window. Whenever the TV of the ground truth sequence $\theta_{b:t}$ is large, there will be a large bias introduced by `prev` due to averaging. Hence in such a scenario the TSE will also be large indicating non-stationarity. However, we can't compute R_t due to the unavailability of θ_j. So we approximate R_t by replacing θ_j with the estimates $\tilde{\theta}_j$ coming from the input online regression algorithm that is not constrained by switching cost restrictions. This is the rationale behind the non-stationarity detection test at Step 5. Whenever a non-staionarity is detected we restart the input online regression algorithm as well as the start position for computing averages (in Step 6).

We have the following guarantee for Algorithm 7.

Theorem 236. *Suppose the input black box ALG given to Algorithm 7 is adaptively minimax optimal (see Definition 72). Then the number of times Algorithm 7 switches its decision is at most $\tilde{O}(T^{1/3} V_T^{2/3})$ with probability at least $1 - \delta$. Further, Algorithm 7 satisfies $\sum_{t=1}^{T} \|\hat{\theta}_t - \theta_t\|_2^2 = \tilde{O}(T^{1/3} V_T^{2/3})$ with probability at least $1 - \delta$, where $V_T = \sum_{t=2}^{T} \|\theta_t - \theta_{t-1}\|_1$.*

Remark 237. Since Algorithm 7 is a black-box reduction, there are a number of possible candidates for the input policy ALG that are adaptively minimax. Examples include FLH with online averages as base learners [23] or Aligator algorithm [71].

Armed with a low switching online regression oracle `LPA`, one can now tweak Algorithm 4 to have sparse number of retraining steps while not sacrificing the statistical efficiency (up to multiplicative constants). The resulting procedure is described in Algorithm 8 (in App. G.3) which enjoys similar rates as in Theorem 77 (see Theorem 240).

G.3 Omitted proofs from Section 8.3

First we recall a result from [71].

Proposition 238 (Theorem 5 of [71])**.** *Consider the online regression protocol defined in Definition 72. Let $\hat{\theta}_t$ be the estimate of the ground truth produced by the Aligator algorithm from [71]. Then with probability at-least $1-\delta$, the total squared error (TSE) of Aligator satisfies*

$$\sum_{t=1}^{T} \|\theta_t - \hat{\theta}_t\|_2^2 = \tilde{O}(T^{1/3} V_T^{2/3} + 1), \tag{G.51}$$

where $V_T = \sum_{t=2}^{T} \|\theta_t - \theta_{t-1}\|_1$. This bound is attained without any prior knowledge of the variation V_T.

The high probability guarantee also implies that

$$\sum_{t=1}^{T} E[\|\theta_t - \hat{\theta}_t\|_2^2] = \tilde{O}(T^{1/3} V_T^{2/3} + 1), \tag{G.52}$$

where the expectation is taken with respect to randomness in the observations.

By following the arguments in [65], a similar statement can be derived also for the FLH-FTL algorithm of [23] (Algorithm 9).

Next, we verify that the noise condition in Definition 72 is satisfied for the empirical label marginals computed at Step 5 of Algorithm 4.

Lemma 239. *Let s_t be as in Step 5 of Algorithm 4. Then it holds that $s_t = q_t + \epsilon_t$ with ϵ_t being independent across t and $\mathrm{Var}(\epsilon_t) \leq 1/N$.*

Proof. Since s_t is simply the empirical label proportions, it holds that $E[s_t] = q_t$. Further $\mathrm{Var}(s_t) \leq 1$ as the indicator function is bounded by $1/N$. This concludes the proof. □

Theorem 77. *Suppose the true label marginal satisfies $\min_{t,k} q_t(k) \geq \mu > 0$. Choose the online regression oracle in Algorithm 4 as FLH-FTL (App. G.4) or Aligator from [71] with its predictions clipped such that $\hat{q}_t[k] \geq \mu$. Then with probability at least $1-\delta$,* *Algorithm 4 produces hypotheses with* $R^{\mathcal{H}}_{dynamic} = \tilde{O}\left(T^{2/3}V_T^{1/3} + \sqrt{T\log(|\mathcal{H}|/\delta)}\right)$ *, where* $V_T = \sum_{t=2}^{T} \|q_t - q_{t-1}\|_1$*. Further, this result is attained without any prior knowledge of the variation budget* V_T*.*

Proof. In the proof we first proceed to bound the instantaneous regret at round t. Rewrite the population loss as:

$$L_t(h) = \frac{1}{N(t-1)} \sum_{i=1}^{t-1} \sum_{j=1}^{N} E\left[\frac{q_t(y_{ij})}{q_i(y_{ij})} \ell(h(x_{ij}), y_{ij})\right], \tag{G.53}$$

where the expectation is taken with respect to randomness in the samples.

We define the following quantities:

$$L_t^{\text{emp}}(h) := \frac{1}{N(t-1)} \sum_{i=1}^{t-1} \sum_{j=1}^{N} \frac{q_t(y_{ij})}{q_i(y_{ij})} \ell(h(x_{ij}), y_{ij}), \tag{G.54}$$

$$\tilde{L}_t(h) := \frac{1}{N(t-1)} \sum_{i=1}^{t-1} \sum_{j=1}^{N} E\left[\frac{\hat{q}_t(y_{ij})}{\hat{q}_i(y_{ij})} \ell(h(x_{ij}), y_{ij})\right], \tag{G.55}$$

and

$$\tilde{L}_t^{\text{emp}}(h) := \frac{1}{N(t-1)} \sum_{i=1}^{t-1} \sum_{j=1}^{N} \frac{\hat{q}_t(y_{ij})}{\hat{q}_i(y_{ij})} \ell(h(x_{ij}), y_{ij}). \tag{G.56}$$

We decompose the regret at round t as

$$L_t(h_t) - L_t(h_t^*) = L_t(h_t) - \tilde{L}_t(h_t) + \tilde{L}_t(h_t) - \tilde{L}_t^{\text{emp}}(h_t) + L_t^{\text{emp}}(h_t^*) - L_t(h_t^*) \tag{G.57}$$

$$+ \tilde{L}_t^{\text{emp}}(h_t) - L_t^{\text{emp}}(h_t^*) \tag{G.58}$$

$$\leq \underbrace{L_t(h_t) - \tilde{L}_t(h_t)}_{\text{T1}} + \underbrace{\tilde{L}_t(h_t) - \tilde{L}_t^{\text{emp}}(h_t)}_{\text{T2}} + \underbrace{L_t^{\text{emp}}(h_t^*) - L_t(h_t^*)}_{\text{T3}} \tag{G.59}$$

$$+ \underbrace{\tilde{L}_t^{\text{emp}}(h_t^*) - L_t^{\text{emp}}(h_t^*)}_{\text{T4}}, \tag{G.60}$$

where in the last line we used Eq.(8.4). Now we proceed to bound each terms as note above.

Note that for any label m,

$$\left|\frac{q_t(m)}{q_i(m)} - \frac{\hat{q}_t(m)}{\hat{q}_i(m)}\right| \leq \left|\frac{q_t(m)}{q_i(m)} - \frac{q_t(m)}{\hat{q}_i(m)}\right| + \left|\frac{q_t(m)}{\hat{q}_i(m)} - \frac{\hat{q}_t(m)}{\hat{q}_i(m)}\right| \tag{G.61}$$

$$\leq \frac{1}{\mu^2}\left(|q_i(m) - \hat{q}_i(m)| + |q_t(m) - \hat{q}_t(m)|\right), \tag{G.62}$$

where in the last line, we used the assumption that the minimum label marginals (and hence of the online estimates via clipping) is bounded from below by μ. So by applying triangle inequality and using the fact that the losses are bounded by B in magnitude, we get

$$T1 \leq \frac{B}{N(t-1)\mu^2}\sum_{i=1}^{t-1}\sum_{j=1}^{N} E\left[\|\hat{q}_i - q_i\|_1 + \|\hat{q}_t - q_t\|_1\right] \tag{G.63}$$

$$\leq \frac{B\sqrt{K}}{(t-1)\mu^2}\sum_{i=1}^{t-1} E\left[\|\hat{q}_i - q_i\|_2 + \|\hat{q}_t - q_t\|_2\right] \tag{G.64}$$

$$\leq_{(a)} \frac{B\sqrt{K}}{\mu^2}\left(E[\|\hat{q}_t - q_t\|_2] + \sqrt{\frac{\sum_{i=1}^{t-1} E[\|q_i - \hat{q}_i\|_2^2]}{t-1}}\right) \tag{G.65}$$

$$\leq_{(b)} \frac{B\sqrt{K}}{\mu^2}\left(E[\|\hat{q}_t - q_t\|_2] + \phi \cdot \frac{V_T^{1/3}}{(t-1)^{1/3}}\right), \tag{G.66}$$

where line (a) is a consequence of Jensen's inequality. In line (b) we used the following fact: by Lemma 239 and Proposition 234, the expected cumulative error of the online oracle at any step is bounded by $\phi t^{1/3} V_t^{2/3}$ for some multiplier ϕ which can contain poly-logarithmic factors of the horizon (see Proposition 238).

Proceeding in a similar fashion, the term $T4$ can be bounded by Eq.(G.66).

Next, we proceed to handle T3. Let $h \in \mathcal{H}$ be any fixed hypothesis. Then each summand in Eq.(G.54) is an independent random variable assuming values in $[0, B/\mu]$ (recall that the losses lie within $[0, B]$). Hence by Hoeffding's inequality we have that

$$L_t^{\text{emp}}(h) - L_t(h) \leq \frac{B}{\mu}\sqrt{\frac{\log(3T|\mathcal{H}|/\delta)}{N(t-1)}}, \tag{G.67}$$

$$\leq \frac{B}{\mu}\sqrt{\frac{\log(3T|\mathcal{H}|/\delta)}{(t-1)}}, \tag{G.68}$$

with probability at-least $1 - \delta/(3T|\mathcal{H}|)$. Now taking union bound across all hypotheses in $\mathcal{H}$, we obtain that:

$$T3 \leq \frac{B}{\mu}\sqrt{\frac{\log(3|\mathcal{H}|/\delta)}{(t-1)}}, \tag{G.69}$$

with probability at-least $1-\delta/(3T)$.

To bound T2, we notice that it is not possible to directly apply Hoeffding's inequality because the summands in Eq.(G.55) are correlated through the estimates of the online algorithm. So in the following, we propose a trick to decorrelate them. For any hypothesis $h \in \mathcal{H}$, we have that

$$\frac{\hat{q}_t(y_{ij})}{\hat{q}_i(y_{ij})}\ell(h(x_{ij}, y_{ij})) - E\left[\frac{\hat{q}_t(y_{ij})}{\hat{q}_i(y_{ij})}\ell(h(x_{ij}, y_{ij}))\right] \tag{G.70}$$

$$= \underbrace{\left(\frac{\hat{q}_t(y_{ij})}{\hat{q}_i(y_{ij})} - \frac{q_t(y_{ij})}{q_i(y_{ij})}\right)\ell(h(x_{ij}, y_{ij})) -}_{U_{ij}} \tag{G.71}$$

$$\underbrace{E\left[\left(\frac{\hat{q}_t(y_{ij})}{\hat{q}_i(y_{ij})} - \frac{q_t(y_{ij})}{q_i(y_{ij})}\right)\ell(h(x_{ij}, y_{ij}))\right]}_{V_{ij}} + \tag{G.72}$$

$$\underbrace{\frac{q_t(y_{ij})}{q_i(y_i j)}\ell(h(x_{ij}, y_{ij})) - E\left[\frac{q_t(y_{ij})}{q_i(y_i j)}\ell(h(x_{ij}, y_{ij}))\right]}_{W_{ij}}. \tag{G.73}$$

Now using Eq.(G.62) and proceeding similar to the bouding steps of Eq.(G.66), we obtain

$$\frac{1}{N(t-1)}\sum_{i=1}^{t-1}\sum_{j=1}^{N} U_{ij} \leq \frac{B}{N(t-1)\mu^2}\sum_{i=1}^{t-1}\sum_{j=1}^{N} \|\hat{q}_i - q_i\|_1 + \|\hat{q}_t - q_t\|_1 \tag{G.74}$$

$$\leq \frac{B\sqrt{K}}{\mu^2(t-1)}\sum_{i=1}^{t-1} \|\hat{q}_i - q_i\|_2 + \|\hat{q}_t - q_t\|_2 \tag{G.75}$$

$$\leq_{(a)} \frac{B\sqrt{K}}{\mu^2}\left(\|\hat{q}_t - q_t\|_2 + \sqrt{\frac{\sum_{i=1}^{t-1}\|q_i - \hat{q}_i\|_2^2}{t-1}}\right) \tag{G.76}$$

$$\leq_{(b)} \frac{B\sqrt{K}}{\mu^2}\left(\|\hat{q}_t - q_t\|_2 + \phi \cdot \frac{V_T^{1/3}}{(t-1)^{1/3}}\right), \tag{G.77}$$

with probability at-least $1-\delta/3$. In line (a) we used Jensen's inequlaity and in the last line we used the fact the the online oracle attains a high probability bound on the total squared error (TSE) (see Proposition 238).

$\frac{1}{N(t-1)} \sum_{i=1}^{t-1} \sum_{j=1}^{N} V_{ij}$ can be bounded using the same expression as above using similar logic.

To bound $\frac{1}{N(t-1)} \sum_{i=1}^{t-1} \sum_{j=1}^{N} W_{ij}$, we note that it is the sum of independent random variables. Hence using the same arguments used to obtain Eq.(G.68), we have that

$$\frac{1}{N(t-1)} \sum_{i=1}^{t-1} \sum_{j=1}^{N} W_{ij} \leq \frac{B}{\mu} \sqrt{\frac{\log(3T|\mathcal{H}|/\delta)}{(t-1)}}, \tag{G.78}$$

with probability at-least $1-\delta/(3T|\mathcal{H}|)$. Hence taking a union bound across all hypothesis classes and across the high probability event of low TSE for the online algorithm yields that

$$T2 \leq \frac{2B\sqrt{K}}{\mu^2} \left(\|\hat{q}_t - q_t\|_2 + \phi \cdot \frac{V_T^{1/3}}{(t-1)^{1/3}} \right) + \frac{B}{\mu} \sqrt{\frac{\log(3T|\mathcal{H}|/\delta)}{(t-1)}}, \tag{G.79}$$

with probability at-least $1 - 2\delta/(3T)$.

Combining the bounds developed for T1,T2,T3 and T4 and by taking a union bound across the event that resulted in Eq.(G.69), we obtain the following bound on instantaneous regret.

$$L_t(h_t) - L_t(h_t^*) \leq \frac{2B\sqrt{K}}{\mu^2} \left(\|\hat{q}_t - q_t\|_2 + E[\|\hat{q}_t - q_t\|_2] + \phi \cdot \frac{V_T^{1/3}}{(t-1)^{1/3}} + \sqrt{\frac{\log(3T|\mathcal{H}|/\delta)}{(t-1)}} \right), \tag{G.80}$$

with probability at-least $1 - \delta/T$.

Note that via Jensen's inequality:

$$\sum_{t=1}^{T} E[\|q_t - \hat{q}_t\|_2] \leq \sqrt{T \sum_{t=1}^{T} E[\|q_t - \hat{q}_t\|_2^2]} \tag{G.81}$$

$$\leq \phi T^{2/3} V_T^{1/3}, \tag{G.82}$$

where in the last line we used Proposition 238.

Similarly it can be shown that

$$\sum_{t=1}^{T} \|q_t - \hat{q}_t\|_2 \leq \phi T^{2/3} V_T^{1/3}, \tag{G.83}$$

under the event that resulted in Eq.(G.80).

Observe that

$$\sum_{t=1}^{T} \frac{V_T^{1/3}}{t^{1/3}} \leq 2T^{2/3}V_T^{1/3}. \tag{G.84}$$

Finally note that

$$\sum_{t=1}^{T} \frac{1}{\sqrt{t}} \leq 2\sqrt{T}. \tag{G.85}$$

Hence combining the above bounds and adding Eq.(G.80) across all time steps, followed by a union bound across all rounds, we obtain that

$$\sum_{t=1}^{T} L_t(h_t) - L_t(h_t^*) \leq \frac{4B\sqrt{K}}{\mu^2}\left(3\phi T^{2/3}V_T^{1/3} + \sqrt{T\log(3T|\mathcal{H}|/\delta)}\right), \tag{G.86}$$

with probability at-least $1-\delta$.

□

Next, we prove Theorem 236.

Theorem 236. *Suppose the input black box ALG given to Algorithm 7 is adaptively minimax optimal (see Definition 72). Then the number of times Algorithm 7 switches its decision is at most* $\tilde{O}(T^{1/3}V_T^{2/3})$ *with probability at least* $1-\delta$. *Further, Algorithm 7 satisfies* $\sum_{t=1}^{T}\|\hat{\theta}_t - \theta_t\|_2^2 = \tilde{O}(T^{1/3}V_T^{2/3})$ *with probability at least* $1-\delta$, *where* $V_T = \sum_{t=2}^{T}\|\theta_t - \theta_{t-1}\|_1$.

Proof. First we proceed to bound the number of switches. Observe that between two time points where condition in Line 5 of Algorithm 7 evaluates true, we can have at-most $\log T$ switches due to the doubling epoch schedule in Line 8.

We first bound the number of times, condition in Line 5 is satisfied. Suppose for some some time t, we have that $\sum_{j=b+1}^{t}\|\text{prev} - \tilde{\theta}_j\|_2^2 > 4K\sigma^2\log(T/\delta)$. Suppose throughout the run of the algorithm, this is i^{th} time the previous condition is satisfied. Let $n_i := t-b+1$ and let $C_i = \text{TV}[b \to t]$ where $\text{TV}[p \to q] = \sum_{t=p+1}^{q}\|\theta_t - \theta_{t-1}\|_1$. Due to the doubling epoch schedule, we have that that $\text{prev} = \frac{1}{\ell}\sum_{j=b}^{\ell} y_j$ and $E[\text{prev}] = \frac{1}{\ell}\sum_{j=b}^{\ell}\theta_j$ for some $n_i \geq \ell \geq (t-b+1)/2 = n_i/2$.

So we have

$$\sum_{j=b+1}^{t} \|\text{prev} - \tilde{\theta}_j\|_2^2 \leq \sum_{j=b+1}^{t} 2\|\text{prev} - \theta_j\|_2^2 + 2\|\tilde{\theta}_j - \theta_j\|_2^2 \tag{G.87}$$

$$\leq \sum_{j=b+1}^{t} 2\|E[\text{prev}] - \theta_j\|_2^2 + 2\|\text{prev} - E[\text{prev}]\|_2^2 + 2\|\tilde{\theta}_j - \theta_j\|_2^2 \tag{G.88}$$

$$\leq_{(a)} 2(\ell C_i^2 + 2\sigma^2 K \log(2T/\delta)) + 2\phi n_i^{1/3} C_i^{2/3} \tag{G.89}$$

$$\leq 4 \max\{n_i C_i^2, \phi n_i^{1/3} C_i^{2/3}\} + 4\sigma^2 K \log(2T/\delta)), \tag{G.90}$$

with probability at-least $1 - \delta/(T)$. In line (a) we used the following facts: i) Due to Hoeffding's inequality, $\|\text{prev} - E[\text{prev}]\|_2^2 \leq \sigma^2 K \log(4T/\delta))/\ell \leq 2\sigma^2 K \log(2T/\delta))/n_i$ with probability at-least $1-\delta/(2T)$; ii) $\|E[\text{prev}]-\theta_j\|_2 = \|\frac{1}{\ell}\sum_{i=b}^{\ell} \theta_i - \theta_j\|_2 \leq \frac{1}{\ell}\sum_{i=b}^{\ell} \|\theta_i - \theta_j\|_2] \leq C_i$; iii) $\|\tilde{\theta}_j - \theta_j\|_2^2 \leq \phi n_i^{1/3} C_i^{2/3}$ with probability at-least $1-\delta/(2T)$ due to condition in Theorem 236; iv) Union bound over the events in (i) and (iii).

Since the condition in Line 5 is satisfied at round t, Eq.(G.90) will imply that $5K\sigma^2 \log(2T/\delta) \leq 4 \max\{n_i C_i^2, \phi n_i^{1/3} C_i^{2/3}\} + 4\sigma^2 K \log(2T/\delta))$. Rearranging the above, we find that

$$C_i \gtrsim K/\sqrt{n_i}, \tag{G.91}$$

where we suppress the dependence on constants and $\log T$.

Let the condition in Line 5 be satisfied M number of times. By union bound, we have that with probability at-least $1 - \delta$

$$V_T \geq \sum_{i=1}^{M} C_i \tag{G.92}$$

$$\gtrsim \sum_{i=1}^{M} K/\sqrt{n_i} \tag{G.93}$$

$$\gtrsim_{(a)} KM \frac{1}{\sqrt{(1/M)\sum_{i=1}^{M} n_i}} \tag{G.94}$$

$$\gtrsim KM^{3/2}/\sqrt{T}, \tag{G.95}$$

where in Line (a) we used Jensen's inequality. Rearranging we get that

$$M = \tilde{O}(T^{1/3} V_T^{2/3} K^{-2/3}), \tag{G.96}$$

with probability at-least $1 - \delta$.

Now we proceed to bound the total squared error (TSE) incurred by Algorithm 7. Let $\hat{\theta}_j$ be the output of Algorithm 7 at round j. Suppose at times $b-1$ and $c+1$, the condition in Line (5) is satisfied. Observe that the condition in Line 5 is not satisfied for any times in $[b,c]$. Then we can conclude that within the interval $[b,c]$ we have that $\sum_{j=b}^{c} \|\hat{\theta}_j - \tilde{\theta}_j\|_2^2 \le 5K\sigma^2 \log(4T/\delta)\log(T)$, since there are only at-most $\log T$ times within $[b,c]$ where condition in Line 9 is satisfied. So we have that

$$\sum_{j=b}^{c} \|\hat{\theta}_j - \theta_j\|_2^2 \le \sum_{j=b}^{c} \|\hat{\theta}_j - \tilde{\theta}_j\|_2^2 + \|\theta_j - \tilde{\theta}_j\|_2^2 \tag{G.97}$$

$$\le 5K\sigma^2 \log(2T/\delta)\log(T) + \phi \cdot n_i^{1/3} C_i^{2/3}, \tag{G.98}$$

with probability at-least $1-\delta/T$. Here $n_i := b-c+1$ and $C_i := \text{TV}[b \to c]$. Further we have that $\|\hat{\theta}_{c+1} - \theta_{c+1}\|_2^2 \le 2B^2$ due to the boundedness condition in Definition 72.

Thus overall we have that $\sum_{j=b}^{c+1} = \tilde{O}(K + n_i^{1/3} C_i^{2/3})$, with probability at-least $1-\delta$ for any interval [b,c+1] such that condition in Line 5 is satisfied at times $b-1$ and $c+1$. Thus we have that

$$\sum_{t=1}^{T} \|\hat{\theta}_j - \theta_j\|_2^2 \precsim \sum_{i=1}^{M} K + n_i^{1/3} C_i^{2/3} \tag{G.99}$$

$$\precsim_{(a)} T^{1/3} V_T^{2/3} K^{1/3} + \sum_{i=1}^{M} n_i^{1/3} C_i^{2/3} \tag{G.100}$$

$$\precsim_{(b)} T^{1/3} V_T^{2/3} K^{1/3} + \left(\sum_{i=1}^{M} n_i\right)^{1/3} \left(\sum_{i=1}^{M} C_i\right)^{2/3} \tag{G.101}$$

$$\precsim T^{1/3} V_T^{2/3} K^{1/3}, \tag{G.102}$$

with probability at-least $1-\delta$. In line (a) we used Eq.(G.96). In line (b) we used Holder's inequality with the dual norm pair $(3, 3/2)$. This concludes the proof.

□

We now present the tweak of Algorithm 4 by instantiating ALG with Algorithm 7 and prove its regret guarantees. The resulting algorithm is described in Algorithm 8.

Theorem 240. *Assume the same notations as in Theorem 77. Suppose we run Algorithm 8 (see Appendix G.3) with ALG instantiated using Algorithm 7 with* $\sigma^2 = 1/N$ *and predictions clipped as in Theorem 77. Further let the online regression oracle given to Algorithm 7 be chosen as one of the candidates mentioned in Remark 237. Then with*

Algorithm 8 `Lazy-TrainByWeights`: handling label shift with sparse ERM calls

Input: Instance ALG of Algorithm 7, A hypothesis Class $\mathcal{H}$

1: At round $t \in [T]$, get estimated label marginal $\hat{q}_t \in \mathbb{R}^K$ from ALG$(s_{1:t-1})$.
2: **if** $\hat{q}_t == \hat{q}_{t-1}$ **then**
3: $\quad h_t = h_{t-1}$
4: **else**
5: $\quad$ Update the hypothesis by calling a weighted-ERM oracle:

$$h_t = \operatorname*{argmin}_{h \in \mathcal{H}} \sum_{i=1}^{t-1} \sum_{j=1}^{N} \frac{\hat{q}_t(y_{i,j})}{\hat{q}_i(y_{i,j})} \ell(h(x_{i,j}), y_{i,j}) \tag{G.103}$$

6: **end if**
7: Get N co-variates $x_{t,1:N}$ and make predictions according to h_t
8: Get labels $y_{t,1:N}$
9: Compute $s_t[i] = \frac{1}{N} \sum_{j=1}^{N} \mathbb{I}\{y_{t,j} = i\}$ for all $i \in [K]$.
10: Update ALG with the empirical label marginals s_t.

probability at-least $1-\delta$, we have that

$$R_{dynamic}^{\mathcal{H}} = \tilde{O}\left(T^{2/3} V_T^{1/3} + \sqrt{T \log(|\mathcal{H}|/\delta)}\right). \tag{G.104}$$

Further, the number of number of calls to ERM oracle (via Step 5) is at-most $\tilde{O}(T^{1/3} V_T^{2/3})$ with probability at-least $1-\delta$.

Sketch. The proof of this theorem closely follows the steps fused for proving Theorem 77. So we only highlight the changes that need to be incorporated to the proof of Theorem 77.

Replace the use of Proposition 238 in the proof of Theorem 77 with Theorem 236.

For any round t, where Step 5 of Algorithm 8 is triggered, we can use the same arguments as in the Proof of Theorem 240 to bound the instantaneous regret by Eq.(G.80). i.e:

$$L_t(h_t) - L_t(h_t^*) \leq \frac{2B\sqrt{K}}{\mu^2} \left(\|\hat{q}_t - q_t\|_2 + E[\|\hat{q}_t - q_t\|_2] + \phi \cdot \frac{V_T^{1/3}}{(t-1)^{1/3}} + \sqrt{\frac{\log(3T|\mathcal{H}|/\delta)}{(t-1)}} \right), \tag{G.105}$$

with probability at-least $1-\delta/T$.

For a round t, where Step 5 is not triggered, we proceed as follows:

Let t' be the most recent time step prior to t when Step 5 is executed. Notice that the population loss can be equivalently represented as

$$L_t(h) = \frac{1}{N(t'-1)} \sum_{i=1}^{t'-1} \sum_{j=1}^{N} E\left[\frac{q_t(y_{ij})}{q_i(y_{ij})} \ell(h(x_{ij}), y_{ij})\right], \tag{G.106}$$

where the expectation is taken with respect to randomness in the samples.

We define the following quantities:

$$L_t^{\text{emp}}(h) := \frac{1}{N(t'-1)} \sum_{i=1}^{t'-1} \sum_{j=1}^{N} \frac{q_t(y_{ij})}{q_i(y_{ij})} \ell(h(x_{ij}), y_{ij}), \tag{G.107}$$

$$\tilde{L}_t(h) := \frac{1}{N(t'-1)} \sum_{i=1}^{t'-1} \sum_{j=1}^{N} E\left[\frac{\hat{q}_t(y_{ij})}{\hat{q}_i(y_{ij})} \ell(h(x_{ij}), y_{ij})\right], \tag{G.108}$$

and

$$\tilde{L}_t^{\text{emp}}(h) := \frac{1}{N(t'-1)} \sum_{i=1}^{t'-1} \sum_{j=1}^{N} \frac{\hat{q}_t(y_{ij})}{\hat{q}_i(y_{ij})} \ell(h(x_{ij}), y_{ij}). \tag{G.109}$$

We decompose the regret at round t as

$$\begin{aligned}
L_t(h_t) - L_t(h_t^*) &= L_t(h_t) - \tilde{L}_t(h_t) + \tilde{L}_t(h_t) - \tilde{L}_t^{\text{emp}}(h_t) + L_t^{\text{emp}}(h_t^*) - L_t(h_t^*) &\text{(G.110)}\\
&+ \tilde{L}_t^{\text{emp}}(h_t) - L_t^{\text{emp}}(h_t^*) &\text{(G.111)}\\
&\le \underbrace{L_t(h_t) - \tilde{L}_t(h_t)}_{\text{T1}} + \underbrace{\tilde{L}_t(h_t) - \tilde{L}_t^{\text{emp}}(h_t)}_{\text{T2}} + \underbrace{L_t^{\text{emp}}(h_t^*) - L_t(h_t^*)}_{\text{T3}} &\text{(G.112)}\\
&+ \underbrace{\tilde{L}_t^{\text{emp}}(h_t^*) - L_t^{\text{emp}}(h_t^*)}_{\text{T4}}, &\text{(G.113)}
\end{aligned}$$

where in the last line we used Eq.(8.4). Now we proceed to bound each terms as note above.

By using the same arguments as in Proof of Theorem 77 and replacing the use of Proposition 238 with Theorem 236, we can bound T1-4. This will result in an instantaneous regret bound at round t (which doesn't trigger step 5) as:

$$L_t(h_t) - L_t(h_t^*) \leq \frac{2B\sqrt{K}}{\mu^2}\left(\|\hat{q}_t - q_t\|_2 + E[\|\hat{q}_t - q_t\|_2] + \phi \cdot \frac{V_T^{1/3}}{(t'-1)^{1/3}} \right. \tag{G.114}$$

$$\left. + \sqrt{\frac{\log(3T|\mathcal{H}|/\delta)}{(t'-1)}} \right), \tag{G.115}$$

$$\leq \frac{2B\sqrt{K}}{\mu^2}\left(\|\hat{q}_t - q_t\|_2 + E[\|\hat{q}_t - q_t\|_2] \right. \tag{G.116}$$

$$\left. + \phi \cdot 4^{1/3} \cdot \frac{V_T^{1/3}}{(t-1)^{1/3}} + \sqrt{\frac{4\log(3T|\mathcal{H}|/\delta)}{(t-1)}} \right), \tag{G.117}$$

with probability at-least $1-\delta/T$. In the last line we used the fact that $t'-1 \geq (t/2)-1 \geq (t-1)/4$ for all $t \geq 3$.

Now adding Eq.(G.105) and (G.117) across all rounds and proceeding similar to the proof of Theorem 77 (and replacing the use of Proposition 238 with Theorem 236) completes the argument.

□

We next prove the matching (up to factors of $\log T$) lower bound.

Theorem 79. *Let $V_T \leq T/8$. There exists a choice of hypothesis class, loss function, and adversarial strategy of generating the data such that $R_{dynamic}^{\mathcal{H}} = \Omega\left(T^{2/3}V_T^{1/3} + \sqrt{T\log(|\mathcal{H}|)}\right)$, where the expectation is taken with respect to randomness in the algorithm and adversary.*

Proof. First we fix the hypothesis class and the data generation strategy. In the problem instance we consider, there are no co-variates. The hypothesis class is defined as

$$\mathcal{H} := \{h_p : h_p \text{ predicts a label } y \sim \text{Ber}(p);\ p \in [|\mathcal{H}|]\}. \tag{G.118}$$

Further we design the hypothesis class such that both $h_0, h_1 \in \mathcal{H}$. Next we fix the data generation strategy:

- Divide the time horizon into batches of length Δ.
- At the beginning of a batch i, the adversary picks $\mathring{q}_i$ uniformly at random from $\{1/2-\delta, 1/2+\delta\}$.
- For all rounds t that falls within batch i, the label $y_t \sim \text{Ber}(q_t)$ is sampled with $q_t := \mathring{q}_i$.
- Learner predicts a label $\hat{y}_t \in \{0,1\}$ and then the actual label y_t is revealed (hence $N=1$ in the protocol of Fig.3).

- Learner suffers a loss given by $\ell_t(\hat{y}_t) = \mathbb{I}\{\hat{y}_t \neq y_t\}$.

It is easy to see that the losses are bounded in $[0, 1]$. Now let's examine the two possibilities of generating labels within a batch. Let's upper bound the variation incurred by the label marginals:

$$\sum_{t=2}^{T} |q_t - q_{t-1}| = \sum_{i=2}^{M} |\mathring{q}_i - \mathring{q}_{i-1}| \tag{G.119}$$

$$\leq 2\delta M \tag{G.120}$$

$$\leq V_T, \tag{G.121}$$

where the last line is obtained by choosing $\delta = V_T/(2M) = V_T\Delta/(2T)$.

Since at the beginning of each batch, the sampling probability of true labels is independently renewed, the historical data till the beginning of a batch is immaterial in minimising the loss within the batch. So we can lower bound the regret within each batch separately and add them up. Below, we focus on lower bounding the regret in batch 1 and the computations are similar for any other batch.

Suppose that the probability that an algorithm predict label $y_t = 1$ is $\hat{q}_t$, where $\hat{q}_t$ is a measurable function of the past data $y_{1:t-1}$. Then we have that the population loss $L_t(\hat{q}_t) := E[\ell_t(\hat{y}_t)] = (1-\hat{q}_t)q_t + \hat{q}_t(1-q_t)$. Here we abuse the notation $L(q_t) := L(h_{q_t})$. We see that the population loss $L_t(\hat{q}_t)$ are convex and its gradient obeys $\nabla L_t(\hat{q}_t) = 1 - 2q_t = E[1 - 2y_t]$ since by our construction $y_t \sim \text{Ber}(q_t)$. Thus the population losses are convex and its gradients can be estimated in an unbiased manner from the data.

We use the following Proposition due to [15].

Proposition 241 (due to Lemma A-1 in [15])**.** *Let $\tilde{\mathbb{P}}$ denote the joint probability of the label sequence $y_{1:\Delta}$ within a batch when they are generated using Ber$(1/2 - \delta)$. So*

$$\tilde{\mathbb{P}}(y_1, \ldots, y_\Delta) = \Pi_{i=1}^{\Delta}(1/2 - \delta)^{y_i}(1/2 + \delta)^{1-y_i}. \tag{G.122}$$

Similarly define $\tilde{\mathbb{Q}}$ as

$$\tilde{\mathbb{Q}}(y_1, \ldots, y_\Delta) = \Pi_{i=1}^{\Delta}(1/2 + \delta)^{y_i}(1/2 - \delta)^{1-y_i}. \tag{G.123}$$

According to the above distribution the data are independently sampled from Bernoulli trials with success probability $1/2 + \delta$. Let $\hat{q}_t$ be the decision of the online algorithm qt round t so that the algorithm predicts label 1 with probability $\hat{q}_t$.

Let $\mathbb{P}$ denote the joint probability distribution across the decisions $\hat{q}_{1:\Delta}$ of any online algorithm under the sampling model $\tilde{\mathbb{P}}$. Similarly define $\mathbb{Q}$. Note that any online algorithm can make decisions at round t only based on the past observed data $y_{1:t-1}$. Further after making the decision $\hat{q}_t$ at round t, an unbiased estimate of the population loss can be *constructed due to the fact that* $\nabla L_t(\hat{q}_t) = E[1-2y_t]$. *Under the availability of unbiased gradient estimates of the losses, it holds that*

$$KL(\mathbb{P}||\mathbb{Q}) \leq 4\Delta\delta^2. \tag{G.124}$$

By choosing $\delta = V_T/(2M) = V_T\Delta/(2T)$ *and* $\Delta = (T/V_T)^{2/3}$, *we get that* $KL(\mathbb{P}||\mathbb{Q}) \leq 1$.

Since $V_T \leq T/8$, the above choice implies that $\delta \in (0, 1/4)$.

For notational clarity, define $L^{\mathbb{P}}(q) = (1-q)(1/2-\delta) + q(1/2+\delta)$ and $L^{\mathbb{Q}}(q) = (1-q)(1/2+\delta)+q(1/2-\delta)$. These corresponds to the population losses according to the sampling models $\mathbb{P}$ and $\mathbb{Q}$ respectively. Observe that $\min_q L^{\mathbb{P}}(q) = \min_q L^{\mathbb{Q}}(q) = 1/2-\delta$. The minimum of $L^{\mathbb{P}}$ and $L^{\mathbb{Q}}$ are achieved at 0 and 1 respectively. Note that both $h_0, h_1 \in \mathcal{H}$. So there is always a hypothesis in $\mathcal{H}$ that corresponds the minimiser of the loss.

Further whenever $\hat{q} \geq 1/2$ we have that

$$L^{\mathbb{P}}(q) = (1/2-\delta) + q(2\delta) \tag{G.125}$$
$$\geq 1/2. \tag{G.126}$$

Similarly whenever $q < 1/2$ we have $L^{\mathbb{Q}}(q) \geq 1/2$. So we define the selector function as $\phi_t := \mathbb{I}\{\hat{q}_t \geq 1/2\}$. Let $q_t^* \in \{0,1\}$ be the minimiser of the loss at round t. Now we can lower bound the instantaneous regret similar as

$$E[L_t(\hat{q}_t) - L_t(q_t^*)] = \frac{1}{2}(E_{\mathbb{P}}[L_t^{\mathbb{P}}(\hat{q}_t) - L_t^{\mathbb{P}}(0)] + \frac{1}{2}(E_{\mathbb{Q}}[L_t^{\mathbb{Q}}(\hat{q}_t) - L_t^{\mathbb{Q}}(1)] \tag{G.127}$$
$$\geq \frac{1}{2}(E_{\mathbb{P}}[L_t^{\mathbb{P}}(\hat{q}_t) - L_t^{\mathbb{P}}(0)|\phi_t = 1]\mathbb{P}(\phi_t = 1) \tag{G.128}$$
$$+ \frac{1}{2}(E_{\mathbb{Q}}[L_t^{\mathbb{Q}}(\hat{q}_t) - L_t^{\mathbb{Q}}(1)|\phi_t = 0]\mathbb{Q}(\phi_t = 0) \tag{G.129}$$
$$\geq \delta/2\max\{\mathbb{P}(\phi_t = 1), \mathbb{Q}(\phi_t = 0)\} \tag{G.130}$$
$$\geq (\delta/8)e^{-1}, \tag{G.131}$$

where the last line is obtained by Propositions 241 and 235.

Thus we get a total lower bound on the instantanoeus regret as

$$\sum_{t=1}^{T} E[L_t(\hat{q}_t) - L_t(q_t^*)] \geq T\delta/(8e) \tag{G.132}$$
$$= \Delta V_T/(16e) \tag{G.133}$$
$$= T^{2/3}V_T^{1/3}/(16e), \tag{G.134}$$

where the last line is obtained by using our choices of $\delta V_T\Delta/(2T)$ and $\Delta = (T/V_T)^{2/3}$.

The second term of of $\Omega(\sqrt{T \log |\mathcal{H}|})$ can be obtained from the existing results on statistical learning theory without distribution shifts. (see for example Theorem 3.23 in [141]).

□

G.4 More details on experiments

In Algorithm 9, we describe the FLH-FTL algorithm from [23] when specialised to squared error losses. When specialized to squared error losses, this algorithm runs FLH with online averages as the base experts.

Algorithm 9 An instance of FLH-FTL from [23] with squared error losses

1: Parameter α is defined to be a learning rate

```
// initializations and definitions
```

2: For FLH-FTL instantiations within UOLS algorithms (as in Algorithm 2), we set $\alpha \leftarrow \sigma^2_{min}(C)/(8K)$, where C is the confusion matrix as in Assumption 1. For instantiations within SOLS algorithms (as in Algorithm 4) we set $\alpha \leftarrow 1/(8K)$

3: For each round $t \in [T]$, $v_t := (v_t^{(1)}, \ldots, v_t^{(t)})$ is a probability vector in $\mathbb{R}^t$. Initialize $v_1^{(1)} \leftarrow 1$

4: For each $j \in [T]$, define a base learner E^j. For each $t > j$, the base expert outputs $E^j(t) := \frac{1}{t-j} \sum_{i=j}^{t-1} z_j$, where z_j to be specified as below. Further $E^j(j) := 0 \in \mathbb{R}^K$

```
// execution steps
```

5: In round $t \in [T]$, set $\forall j \leq t$, $x_t^j \leftarrow E^j(t)$ (the prediction of the j^{th} base learner at time t). Play $x_t = \sum_{j=1}^{t} v_t^{(j)} x_t^{(j)}$.

6: Receive feedback z_t, set $\hat{v}_{t+1}^{(t+1)} \leftarrow 0$ and perform update for $1 \leq i \leq t$:

$$\hat{v}_{t+1}^{(i)} \leftarrow \frac{v_t^{(i)} e^{-\alpha \|x_t^{(i)} - z_t\|_2^2}}{\sum_{j=1}^{t} v_t^{(j)} e^{-\alpha \|x_t^{(j)} - z_t\|_2^2}} \tag{G.135}$$

7: Addition step - Set $v_{t+1}^{(t+1)}$ to $1/(t+1)$ and for $i \neq t+1$:

$$v_{t+1}^{(i)} \leftarrow (1 - (t+1)^{-1}) \hat{v}_{t+1}^{(i)} \tag{G.136}$$

Rationale behind the learning rate setting at Line 2 of Algorithm 9 The loss that is incurred by Algorithm 9 and any of its base learners at round t is defined to be the squared error loss $\ell_t(x) = \|z_t - x\|_2^2$. Whenever $\|z_t\|_2^2 \leq B^2$ and $\|x\|_2^2 \leq B^2$, the losses $\ell_t(x)$ are $1/(8B^2)$ exp-concave (see for eg. Chapter 3 of [40]). The notion of exp-concavity is crucial for FLH-FTL algorithm since the learning rate is set to be equal to the exp-concavity factor of the loss functions (see Theorem 3.1 in [23]).

For the UOLS problem, from Algorithm 2, we have
$\|z_t\|_2 = \|C^{-1}f_0(x_t)\|_2 \leq \sqrt{K}/\sigma_{min}(C)$. Since the decisions of the algorithm is a convex combination of the previously seen z_t, we conclude that the losses $\ell_t(x)$ are $\sigma^2_{min}(C)/(8K)$ exp-concave.

For the SOLS problem, let $z_t = s_t$ where s_t is as defined in Algorithm 4. We have that $\|z_t\|_2 \leq \sqrt{K}$. Hence arguing in a similar fashion as above, we conclude that the losses $\ell_t(x)$ are $1/(8K)$ exp-concave for the SOLS problem.

This is the motivation behind Line 2 in Algorithm 9, where the learning rates are set according to the problem setting.

Dataset and model details.

- Synthetic: For the synthetic data, we generated 72k samples as described in [97]. There are three classes each with 24k samples generated from three Gaussian distributions in $\mathbb{R}^{12}$. Each Gaussian distribution is defined by a randomly generated unit-norm centre v and covariance matrix $0.215 \cdot I$. 60k samples are used as source data, and 12k samples are used as target data to be sampled from during online learning. We used logistic regression to train a linear model. It is trained for a single epoch with learning rate 0.1, momentum 0.9, batch size 200, and l_2 regularization 1×10^{-4}.

- MNIST [105]: An image dataset of 10 types of handwritten digits. 60k samples are used as source data and 10k as target data. We used an MLP for prediction with three consecutive hidden layers of sizes 100, 100, and 20. It is trained for a single epoch with a learning rate 0.1, momentum 0.9, batch size 200, and l_2 regularization 1×10^{-4}.

- CIFAR-10 [106]: A dataset of colored images of 10 items: airplane, automobile, bird, cat, deer, dog, frog, horse, ship, and truck. 50k samples are used as source data and 10k as target data. We train a ResNet18 model ([112]) from scratch. It is finetuned for 70 epochs with learning rate 0.1, momentum 0.9, batch size 200, and l_2 regularization 1×10^{-4}. The learning rate decayed by 90% at the 25th and 40th epochs.

- Fashion [107]: An image dataset of 10 types of fashion items: T-shirt, trouser, pullover, dress, coat, sandals, shirt, sneaker, bag, and ankle boots. 60k samples are used as source data and 10k as target data. We trained an MLP for prediction. It is trained for 50 epochs with learning rate 0.1, momentum 0.9, batch size 200, and l_2 regularization 1×10^{-4}.

- EuroSAT [108]: An image dataset of 10 types of land uses: industrial buildings, residential buildings, annual crop, permanent crop, river, sea & lake, herbaceous vegetation, highway, pasture, and forest. 60k samples are used as source data and 10k as target data. We cropped the images to the size $(3, 64, 64)$. We train a

ResNet18 model for 50 epochs with learning rate 0.1, momentum 0.9, batch size 200, and l_2 regularization 1×10^{-4}.

- Arxiv [109]: A natural language dataset of 23 classes over different publication subjects. 198k samples are used as source data and 22k as target data. We trained a DistilBERT model ([113]) for 50 epochs with learning rate 2×10^{-5}, batch size 64, and l_2 regularization 1×10^{-2}.

- SHL [110, 111]: A tabular locomotion dataset of 6 classes of human motion: still, walking, run, bike, car and bus. 30k samples are used as source data and 70k as target data. We trained an MLP for prediction for 50 epochs with learning rate 0.1, momentum 0.9, batch size 200, and l_2 regularization 1×10^{-4}.

For all the datasets above, the initial offline data are further split by 80 : 20 into training and holdout data, where the former is used for offline training of the base model and the latter for computing the confusion matrix and retraining (e.g. updating the linear head parameters with UOGD or updating the softmax prediction with our FLT-FTL) during online learning. To examine how well the algorithms adapt when holdout data is limited, we use 10% of the holdout data (i.e., 2% of the initial offline data) in the main chapter unless stated otherwise. In App. G.5.2, we ablate with full hold-out data.

Types of Simulated Shifts. We simulate four kinds of label shifts to capture different non-stationary environments. These shifts are similar to the ones used in [97]. For each shift, the label marginals are a mixture of two different constant marginals $\mu_1, \mu_2 \in \Delta_K$ with a time-varying coefficient α_t: $\mu_{y_t} = (1 - \alpha_t)\mu_1 + \alpha_t\mu_2$, where μ_{y_t} denotes the label distribution at round t and α_t controls the shift non-stationarity and patterns. In particular, we have: *Sinusoidal Shift (Sin)*: $\alpha_t = \sin \frac{i\pi}{L}$, where $i = t \mod L$ and L is a given periodic length. In the experiments, we set $L = \sqrt{T}$. *Bernoulli Shift (Ber)*: at every iteration, we keep the $\alpha_t = \alpha_{t-1}$ with probability $p \in [0, 1]$ and otherwise set $\alpha_t = 1 - \alpha_{t-1}$. In the experiments, the parameter is set as $p = 1/\sqrt{T}$, which implies $V_t = \Theta(\sqrt{T})$. *Square Shift (Squ)*: at every L rounds we set $\alpha_t = 1-\alpha_{t-1}$. *Monotone Shift (Mon)*: we set $\alpha_t = t/T$. Square, sinusoidal, and Bernoulli shifts simulate fast-changing environments with periodic patterns.

Methods for UOLS Adaptation.

- Base: the base classifier without any online modifications.

- OFC: the optimal fixed classifier predicts with an optimal fixed re-weighting in hindsight as in [96].

- Oracle: re-weight the base model's predictions with the true label marginal of the unlabeled data at each time step.

		CT (base)	CT-RS (ours) w FLH	CT-RS (ours) w FLT-FTL	w-ERM (oracle)
MNIST	Cl Err	5.0±0.5	4.71±0.2	**4.53±0.1**	3.2±0.4
	MSE	NA	0.12±0.01	**0.08±0.01**	NA

Table G.1: *Results on SOLS setup.* We report results with MNIST SOLS setup runs for $T = 200$ steps. We observe that continual training with re-sampling improves over the base model which continually trains on the online data and achieves competitive performance with respect to weighted ERM oracle.

	CT-RS (ours)	w-ERM (oracle)
CIFAR	**145±3.7**	1882±14
MNIST	**20±2.7**	107±3.6

Table G.2: *Comparison on computation time (in minutes).* We report results with MNIST and CIFAR SOLS setup runs for $T = 200$ steps. We observe that continual training with re-sampling is approximately 5–15× more efficient than weighted ERM oracle.

- FTH: proposed by [96], follow-the-history classifier re-weights the base model's predictions with a simple online average of all marginal estimates seen thus far.
- FTFWH: proposed by [96], follow-the-fixed-window-history classifier is a version of FTH that tracks only the k most recent marginal estimates. We choose $k = 100$ throughout the experiments in this work.
- ROGD: proposed by [96], ROGD uses online gradient descent to update its re-weighting vector based on current marginal estimate.
- UOGD: proposed by [97], retrains the last linear layer of the model based on current marginal estimate.
- ATLAS: proposed by [97] is a meta-learning algorithm that has UOGD as its base learners.

The learning rates of ROGD, UOGD, and ATLAS are set according to suggestions in the original works. The learning rate of FLH-FTL is set to $1/K$. This corresponds to a faster rate than the theoretically optimal learning rate given in Line 2 of Algorithm 9. It has been observed in prior works such as [71] that the theoretical learning rate is often too conservative and faster rates lead to a better performance.

G.4.1 Supervised Online Label Shift Experiment Details

For each dataset, we first fix the number of time steps and then simulate the label marginal shift. To train the learner with all the methods, we store all the online data observed giving the storage complexity of $\mathcal{O}(T)$. We observe $N = 50$ examples at every iteration and we split the observed labeled examples into 80:20 split for training and validation. The validation examples are used to decide the number of gradient steps at every time step, in particular, we take gradient steps till the validation error continues to decrease.

Dataset and model details.

- MNIST [105]: An image dataset of 10 types of handwritten digits. At each step, we sample 50 samples with the label marginal that step without replacement and reveal the examples to the learner. We used an MLP for prediction with three consecutive hidden layers of sizes 100, 100, and 20. It is trained for a single epoch with a learning rate 0.1, momentum 0.9, batch size 200, and l_2 regularization 1×10^{-4}.
- CIFAR-10 [106]: A dataset of colored images of 10 items: airplane, automobile, bird, cat, deer, dog, frog, horse, ship, and truck. At each step, we sample 50 samples with the label marginal that step without replacement and reveal the examples to the learner. It is finetuned for 70 epochs with learning rate 0.1, momentum 0.9, batch size 200, and l_2 regularization 1×10^{-4}.

We simulate Bernoulli label shifts to capture different non-stationary environments.

Connection of CT-RS to weighted ERM Before making the connection, we first re-visit the CT-RS algorithm. **Step 1**: Maintain a pool of all the labeled data received till that time step, and at every iteration, we randomly sample a batch with uniform label marginal to update the model. **Step 2**: Re-weight the softmax outputs of the updated model with estimated label marginal. Below we show that it is equivalent to wERM:

$$\begin{aligned} f_t &= \operatorname*{argmin}_{f \in \mathcal{H}} \sum_{i=1}^{t-1} \sum_{j=1}^{N} \frac{\hat{q}_t(y_{i,j})}{\hat{q}_i(y_{i,j})} \ell(f(x_{i,j}), y_{i,j}) \\ &= \operatorname*{argmin}_{f \in \mathcal{H}} \sum_{k=1}^{K} \hat{q}_t(k) \sum_{i=1}^{t-1} \sum_{j=1}^{N} \frac{\mathbb{I}\left(y_{i,j} = k\right)}{\hat{q}_i(k)} \ell(f(x_{i,j}), k) \\ &= \operatorname*{argmin}_{f \in \mathcal{H}} \sum_{k=1}^{K} \frac{\hat{q}_t(k)}{(1/K)} \sum_{i=1}^{t-1} \sum_{j=1}^{N} \frac{\mathbb{I}\left(y_{i,j} = k\right)}{K \cdot \hat{q}_i(k)} \ell(f(x_{i,j}), k) \\ &= \operatorname*{argmin}_{f \in \mathcal{H}} \sum_{k=1}^{K} \hat{\mu}_{t,k} \underbrace{\sum_{i=1}^{t-1} \sum_{j=1}^{N} \frac{\mathbb{I}\left(y_{i,j} = k\right)}{\hat{\mu}_{i,k}} \ell(f(x_{i,j}), k)}_{L_{t-1,k}} \end{aligned}$$

where $\hat{\mu}_{t,k} = {}^{\hat{q}_t(k)}/_{(1/K)}$ is the importance ratio at time i with respect to uniform label marginal. Similarly, we define $\hat{\mu}_{i,k} = {}^{\hat{q}_i(k)}/_{(1/K)}$. Here, $L_{t-1,k}$ is the aggregate loss at t-th time step for k-th class such that at each step the sampling probability of that class is uniform. By continually training a classifier with CT-RS, Step 1 approximates the classifier $\widetilde{f}_t$ trained to minimize the average of $L_{t-1,k}$ over all classes with uniform proportion for each class. To update the classifier $\widetilde{f}_t$ according to label proportions at time t, we update the softmax output of $\widetilde{f}_t$ according to $\hat{\mu}_t$.

The primary benefit of CT-RS over wERM is to avoid re-training from scratch at every iteration. Instead, we can leverage the model trained in the previous iteration to warm-start training in the next iteration.

G.5 Additional Unsupervised Online Label Shift Experiments

G.5.1 Additional results with Monotone and Square Shifts and Low Amount of Holdout Data

Methods	Synthetic		MNIST		CIFAR		EuroSAT		Fashion		ArXiv	
	Mon	Squ	Mon	Squ	Mon	Squ	Mon	Squ	Mon	Squ	Mon	Squ
Base	8.7±0.1	8.5±0.2	4.7±0.0	4.4±0.2	17±0	17±0	13±0	13±0	15±0	15±0	22±0	21±0
OFC	6.9±0.1	6.6±0.3	4.1±0.1	3.9±0.2	14±0	14±0	11±1	11±0	9.0±0.0	9.6±0.5	18±1	18±0
Oracle	5.2±0.2	3.6±0.2	2.5±0.1	2.2±0.1	7.7±0.1	6.8±0.2	5.3±0.2	4.4±0.0	5.1±0.1	4.1±0.1	6.9±0.3	6.6±0.2
FTH	7.1±0.3	6.8±0.4	4.1±0.1	**4.0±0.0**	13±1	13±0	11±0	11±0	9.3±0.6	8.9±0.4	19±1	18±0
FTFWH	**6.3±0.2**	7.0±0.0	**4.0±0.0**	4.1±0.1	**12±0**	13±0	**9.9±0.2**	11±0	**8.4±0.3**	9.1±0.5	**18±1**	18±0
ROGD	7.8±0.3	7.8±0.3	4.5±0.2	5.4±1.7	14±1	15±0	11±0	14±1	8.9±0.4	10±1	19±1	21±1
UOGD	8.1±0.3	8.1±0.5	4.9±0.1	4.8±0.4	15±1	15±0	10±1	11±1	11±2	12±2	20±1	19±0
ATLAS	8.0±0.0	8.2±0.5	4.6±0.2	4.5±0.3	15±1	15±0	10±1	11±1	12±2	12±1	20±1	19±1
FLH-FTL (ours)	**6.3±0.3**	**5.6±0.4**	**4.0±0.0**	**4.0±0.0**	**12±0**	**12±0**	10±0	**10±0**	8.6±0.4	**8.4±0.4**	**18±1**	**17±0**

	Synthetic		MNIST		CIFAR		EuroSAT		Fashion		ArXiv	
	Mon	Squ	Mon	Squ	Mon	Squ	Mon	Squ	Mon	Squ	Mon	Squ
FTH	0.11±0.00	0.21±0.01	0.14±0.00	0.27±0.00	0.15±0.01	0.28±0.00	0.14±0.01	0.28±0.00	0.16±0.02	0.28±0.01	0.18±0.00	0.30±0.00
FTFWH	**0.05±0.00**	0.23±0.01	**0.07±0.00**	0.30±0.00	**0.07±0.00**	0.30±0.00	**0.07±0.00**	0.30±0.00	**0.08±0.01**	0.31±0.01	**0.09±0.00**	0.32±0.00
ROGD	0.18±0.01	0.29±0.01	0.28±0.05	0.41±0.04	0.22±0.04	0.37±0.03	0.27±0.03	0.41±0.02	0.21±0.01	0.37±0.01	0.21±0.01	0.36±0.01
FLH-FTL (ours)	**0.05±0.00**	**0.11±0.00**	**0.07±0.00**	**0.15±0.00**	0.09±0.01	**0.17±0.00**	0.08±0.01	**0.17±0.01**	0.09±0.01	**0.18±0.02**	0.11±0.00	**0.24±0.00**

Table G.3: *Results for UOLS problems with monotone and square shifts using low amount of holdout data.* **Top:** Classification Error. **Bottom:** Mean-squared error in estimating label marginal.

G.5.2 Additional results with All of Holdout Data

Methods	Synthetic		MNIST		CIFAR		EuroSAT		Fashion		ArXiv	
	Ber	Sin	Ber	Sin	Ber	Sin	Ber	Sin	Ber	Sin	Ber	Sin
Base	8.6±0.3	8.2±0.3	3.8±0.3	3.9±0.0	17±0	16±0	13±0	13±0	15±0	15±0	23±0	19±0
OFC	6.7±0.2	5.5±0.2	3.4±0.4	3.4±0.2	13±0	11±0	11±1	9.8±1.3	8.3±0.5	6.8±0.2	21±1	14±0
Oracle	3.7±0.1	3.7±0.1	1.7±0.2	1.5±0.1	6.3±0.1	5.9±0.1	4.0±0.0	4.1±0.1	3.5±0.2	3.6±0.1	7.8±0.2	5.1±0.1
FTH	6.8±0.2	5.5±0.3	**3.2±0.2**	**3.2±0.2**	12±0	**10±0**	11±0	9.5±0.1	8.0±0.0	**6.8±0.2**	20±0	**14±0**
FTFWH	6.6±0.3	5.5±0.2	3.3±0.2	**3.2±0.1**	12±0	**10±0**	10±0	9.4±0.2	7.9±0.0	6.9±0.2	20±0	**14±0**
ROGD	7.8±0.3	7.2±0.3	4.7±0.3	3.3±0.2	15±0	11±0	11±0	10±0	14±5	8.2±0.2	23±0	16±1
UOGD	7.6±0.4	7.0±0.0	**3.2±0.2**	**3.2±0.2**	**11±0**	**10±0**	**7.7±0.0**	**7.3±0.2**	9.6±0.2	8.6±0.1	**19±0**	**14±0**
ATLAS	7.5±0.3	6.8±0.3	**3.2±0.3**	**3.2±0.2**	12±0	11±0	9.1±0.0	8.3±0.2	12±0	11±0	21±0	16±0
FLH-FTL (ours)	**5.4±0.3**	**5.3±0.2**	**3.2±0.2**	3.3±0.2	**11±0**	**10±0**	9.4±0.2	9.3±0.1	**7.5±0.1**	7.0±0.0	**19±0**	**14±0**

	Synthetic		MNIST		CIFAR		EuroSAT		Fashion		ArXiv	
	Ber	Sin	Ber	Sin	Ber	Sin	Ber	Sin	Ber	Sin	Ber	Sin
FTH	0.20±0.00	0.10±0.00	0.25±0.00	0.14±0.00	0.28±0.00	0.14±0.00	0.27±0.00	0.14±0.00	0.27±0.00	0.14±0.00	0.29±0.00	0.15±0.00
FTFWH	0.19±0.00	0.09±0.00	0.24±0.00	0.13±0.00	0.24±0.00	**0.13±0.00**	0.26±0.00	**0.13±0.00**	0.24±0.00	**0.13±0.00**	0.28±0.00	**0.15±0.00**
ROGD	0.29±0.00	0.23±0.00	0.43±0.00	0.33±0.00	0.31±0.00	0.21±0.00	0.41±0.00	0.34±0.00	0.45±0.08	0.31±0.00	0.34±0.00	0.28±0.00
FLH-FTL (ours)	**0.09±0.00**	**0.08±0.00**	**0.13±0.00**	**0.12±0.00**	**0.15±0.00**	**0.13±0.00**	**0.15±0.00**	**0.13±0.00**	**0.15±0.00**	**0.13±0.00**	**0.22±0.00**	**0.15±0.00**

Table G.4: *Results for UOLS problems using all hold-out data.* **Top:** Classification Error. **Bottom:** Mean-squared error in estimating label marginal. Compared to the result in the main paper (Table 8.1), we observe that the performances of ROGD, UOGD, and ATLAS depend more on availability of holdout data that FLH-FTL. Notably, UOGD becomes competitive in the majority of the datasets when abundant holdout data are available.

G.5.3 Ablation over Number of Online Samples

Here we examine how different algorithms perform as the number of online samples varies. We introduce an additional baseline **BBSE**, which simply uses the label marginal estimate provided by black box shift estimator to reweight the predictions of classifiers. Figure G.1 shows an interesting trend that as number of online samples increases, the

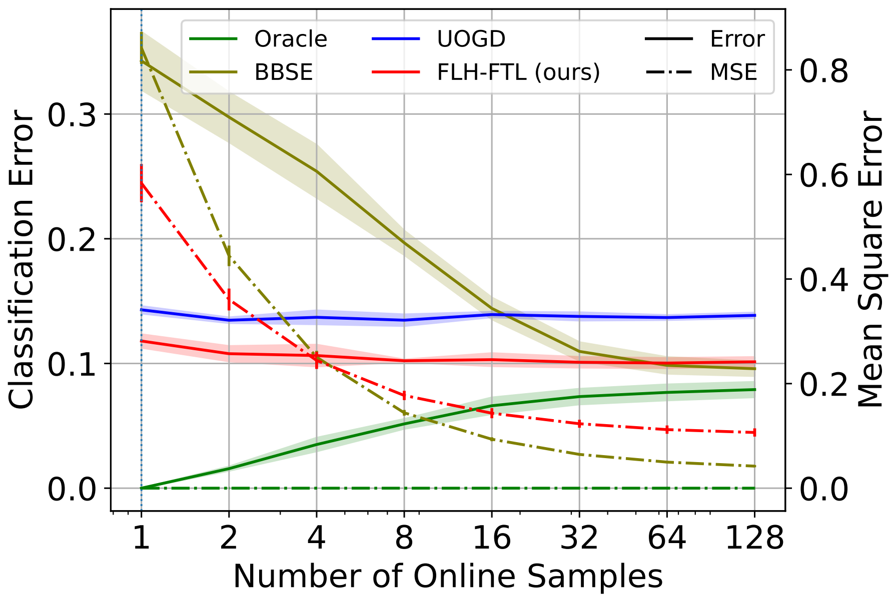


Figure G.1: *Performances of online learning algorithms with different number of online samples.* CIFAR-10 results with bernouli shift and limited holdout data. Solid line is the classification error (Error) and the dotted line is the marginal estimation mean squared error (MSE).

simple baseline BBSE becomes more competitive and eventually outperforms UOGD, whereas our algorithm remains competitive.

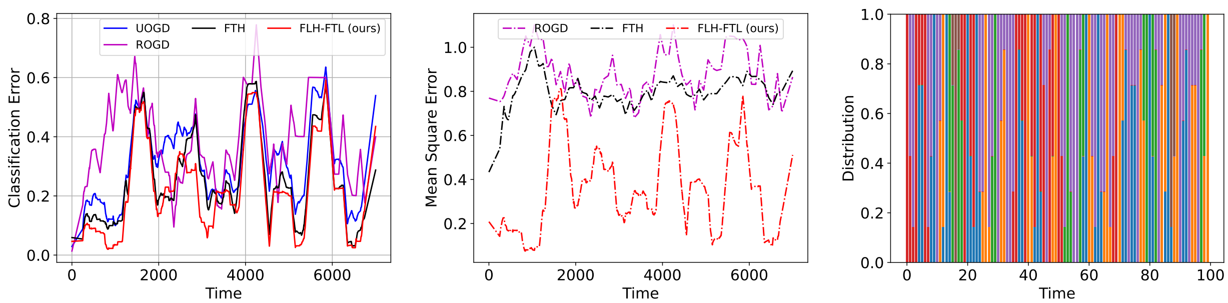
G.5.4 Ablation over Types of Marginal Estimates

All the algorithms examined in this work use black box shift estimate (BBSE) [91] to obtain an unbiased estimate of the target label marginal. However, two alternative methods exist: Maximum Likelihood Label Shift (MLLS) [90] and Regularized Learning under Label Shift (RLLS) [94]. Table G.5 presents additional results using these two estimates. The results shows using the alternative estimates do not substantially change the performances of the algorithms considered in this work.

Datasets		Synthetic			CIFAR			Fashion		
MARG EST.		BBSE	MLLS	RLLS	BBSE	MLLS	RLLS	BBSE	MLLS	RLLS
Bernouli	Base	8.6±0.2	8.6±0.2	8.6±0.2	16±0	16±0	16±0	15±0	15±0	15±0
	OFC	6.4±0.6	6.4±0.6	6.4±0.6	12±1	12±1	12±1	7.9±0.1	7.9±0.1	7.9±0.1
	FTH	6.5±0.6	6.5±0.7	6.5±0.7	11±0	11±1	11±0	8.5±0.3	8.0±0.0	8.6±0.2
	FTFWH	6.6±0.5	6.7±0.5	6.6±0.5	11±1	11±1	11±1	8.2±0.6	7.9±0.2	8.3±0.6
	ROGD	7.9±0.3	7.9±0.3	7.9±0.2	16±3	16±3	15±2	10±1	10±1	9.6±1.2
	UOGD	8.1±0.6	8.0±1.0	8.0±1.0	14±0	13±0	14±0	11±2	10±1	11±1
	ATLAS	8.0±1.0	7.9±0.5	8.0±1.0	13±0	13±0	13±0	12±2	11±1	12±1
	FLH-FTL (ours)	**5.4±0.7**	**5.4±0.8**	**5.4±0.7**	**10±0**	**10±1**	**10±0**	**7.7±0.4**	**7.3±0.3**	**7.6±0.3**
Sinusoidal	Base	8.2±0.3	8.2±0.3	8.2±0.3	16±0	16±0	16±0	15±0	15±0	15±0
	OFC	5.5±0.2	5.5±0.2	5.5±0.2	11±0	11±0	11±0	7.1±0.1	7.1±0.1	7.1±0.1
	FTH	5.7±0.3	5.7±0.2	5.7±0.2	**11±0**	11±0	**11±0**	**6.9±0.4**	**6.6±0.2**	**6.8±0.4**
	FTFWH	5.7±0.3	5.6±0.2	5.7±0.3	**11±0**	11±0	**11±0**	**6.9±0.4**	**6.6±0.3**	6.9±0.4
	ROGD	7.2±0.6	7.2±0.6	7.2±0.6	13±0	13±0	13±0	8.2±0.7	8.9±0.6	8.2±0.3
	UOGD	7.5±0.6	7.4±0.5	7.4±0.5	14±1	13±1	14±1	11±2	9.4±0.9	11±2
	ATLAS	7.5±0.6	7.4±0.6	7.4±0.6	13±1	13±1	13±1	12±2	11±1	12±2
	FLH-FTL (ours)	**5.4±0.4**	**5.4±0.3**	**5.4±0.4**	**11±0**	**10±0**	**11±0**	7.0±0.0	**6.6±0.2**	6.9±0.4

Table G.5: *Performances of online learning algorithms with different types of marginal estimates with low amount of holdout data.*

G.5.5 Additional results and details on the SHL dataset



(a) (b) (c)

Figure G.2: *Additional results and details on the SHL datasets with real shift.* **(a) and (b):** The accuracies and mean square errors in label marginal estimation on SHL dataset over 7,000 time steps with limited amount of holdout data. **(c):** Label marginals of the six classes of SHL dataset. Each time step here shows the marginals over 700 samples.

G.5.6 Reweighting Versus Retraining Linear Layer

Here we compare the efficacies of re-weighting (RW-FLH-FTL) and retraining (RT-FLH-FTL) given the marginal estimate provided by FLH-FTL; the latter retrains the last linear layer on the loss of the holdout data re-weighted by the marginal estimate. Note that RW-FLH-FTL corresponds to FLH-FTL in the main text. We retrain RT-FLH-FTL for 50 epochs at each time step. To compare against the best possible retrained classifiers, we used all the holdout data for retraining. Table G.6 shows that retraining is often worse and at best similar to re-weighting in performance, despite greater computational cost and need for holdout data.

Datasets	Synthetic				CIFAR			
Shift	Mon	Sin	Ber	Squ	Mon	Sin	Ber	Squ
Base	8.7±0.1	8.2±0.3	8.6±0.3	8.5±0.2	17±0	16±0	17±0	17±0
OFC	6.8±0.1	5.5±0.2	6.7±0.2	6.7±0.3	14±0	11±0	13±0	13±1
FTFWH	7.0±0.0	5.5±0.3	6.8±0.2	6.8±0.3	13±0	10±0	12±0	13±0
UOGD	7.4±0.1	7.0±0.0	7.6±0.4	7.6±0.2	**12±0**	**10±0**	**11±0**	13±0
RW-FLH-FTL (ours)	**6.3±0.2**	**5.3±0.2**	**5.4±0.3**	**5.5±0.2**	**12±0**	**10±0**	**11±0**	**12±0**
RT-FLH-FTL (ours)	6.7±0.1	6.2±0.2	6.0±0.0	6.3±0.4	**12±0**	**10±0**	**11±0**	**12±0**

Table G.6: *Comparison of performances of re-weighting and retraining strategies with high amount of holdout data.*

Appendix H

Supplementary Materials for Chapter 9

H.1 Detailed Proof

H.1.1 Proof of Lemma 86

Proof. To begin with, we know that

$$h_t(\theta) = -\mathbb{I}_t \cdot \frac{f(\omega)}{1-F(\omega)} + (1-\mathbb{I}_t) \cdot \frac{f(\omega)}{F(\omega)},$$

where $\omega = v_t - x_t^\top \theta$. Since $\exists \theta_t \in \mathcal{D}_t$ such that $v_t = J(x_t^\top \theta_t)$, given that $J'(u) \in (0,1)$ [118], we know that $\omega \in [J(-B) - B, J(B) + B]$ is bounded in a closed interval. Since we assume that $f(\omega) > 0, \forall \omega \in \mathbb{R}$, we know that $f_{\min} = \inf_{\omega \in [J(-B)-B, J(B)+B]} f(\omega) > 0$ and $F(\omega) \in [F(J(-B) - B), F(J(B) + B)] \subset (0,1)$. Remember that we denote $B_f := \sup_{\omega \in \mathbb{R}} f(\omega) < +\infty$. As a result, we have

$$\begin{aligned} 0 < f_{\min} \le & \frac{f(\omega)}{1-F(\omega)} \le \frac{B_f}{1-F(J(B)+B)} < +\infty \\ 0 < f_{\min} \le & \frac{f(\omega)}{F(\omega)} \le \frac{B_f}{F(J(-B)-B)} < +\infty. \end{aligned} \tag{H.1}$$

Since $h_t(\theta) = \frac{f(\omega)}{1-F(\omega)}$ for $\mathbb{I}_t = 1$ or $h(t) = \frac{f(\omega)}{F(\omega)}$ for $\mathbb{I}_t = 0$, we know that $|h_t(\theta)| \in [f_{\min}, \frac{B_f}{\min\{1-F(J(B)+B), F(J(-B)-B)\}}]$. Let $h_{\max} = \frac{B_f}{\min\{1-F(J(B)+B), F(J(-B)-B)\}}$ and $h_{\min} = f_{\min}$, and the lemma is therefore proved. □

H.1.2 Proof of Lemma 92

Proof. We have that for any $\theta \in \mathcal{D}_\infty^B$,

$$|\hat{\ell}_t(\theta) - \hat{\ell}_t(\hat{\theta}_t)| = \left|1/\sqrt{\beta} + \sqrt{\beta} \cdot \nabla \ell_t(\hat{\theta}_t)^\top(\theta - \hat{\theta}_t)\right| \cdot \left|\sqrt{\beta}\nabla \ell_t(\hat{\theta}_t)^\top(\theta - \hat{\theta}_t)\right| \tag{H.2}$$

$$\leq \left(1 + 2GB\beta\sqrt{d}\right) |\nabla \ell_t(\hat{\theta}_t)^\top(\theta - \hat{\theta}_t)|, \tag{H.3}$$

where in the last line we apply triangle inequality and the facts that $|\nabla \ell_t(\hat{\theta}_t)^\top(\theta - \hat{\theta}_t)| \leq G\|\theta - \hat{\theta}_t\|_2$ with $\|\theta - \hat{\theta}_t\|_2 \leq 2B\sqrt{d}$.

Putting $G' = 1 + 2GB\beta\sqrt{d}$ completes the lemma.

□

H.1.3 Proof of Lemma 93

Proof. For the simplicity of notation, we denote $\nabla_t := \nabla \ell_t(\hat{\theta}_t)$, and we have: $S_t(\theta) = \min_{x\in\mathcal{D}_t} |\nabla_t^\top(x - \theta)|$. Since $S_t(\theta)$ is convex in $\mathbb{R}^d$, we have:

$$S_t(\theta_2) \geq S_t(\theta_1) + \langle \nabla S_t(\theta_1), \theta_2 - \theta_1 \rangle, \forall \theta_1, \theta_2 \in \mathcal{D}_\infty^B.$$

Now we conduct an orthogonal decomposition: $\nabla S_t(\theta_1) = \mu_1 \nabla_t + \nabla_t^\perp$, where $\nabla_t^\top \nabla_t^\perp = 0$. Let $\theta_3 = \theta_2 + \mu_2 \nabla_t^\perp$, and we have $|\nabla_t^\top(x - \theta_2)| = |\nabla_t^\top(x - \theta_3)|, \forall x \in \mathbb{R}^d$. In other words, we have $S_t(\theta_2) = S_t(\theta_3)$ and therefore we have:

$$\begin{aligned} S_t(\theta_2) = S_t(\theta_3) &\geq S_t(\theta_1) + \langle \nabla S_t(\theta_1), \theta_3 - \theta_1 \rangle \\ &= S_t(\theta_1) + \langle \mu_1 \nabla_t + \nabla_t^\perp, \theta_2 + \mu_2 \nabla_t^\perp - \theta_1 \rangle \\ &= S_t(\theta_1) + \langle \nabla S_t(\theta_1), \theta_2 - \theta_1 \rangle + \mu_2 \langle \nabla_t^\perp, \nabla_t^\perp \rangle, \forall \theta_2 \in \mathbb{R}^d, \mu_2 \in \mathbb{R} \end{aligned}$$

In other words, $\mu_2 \|\nabla_t^\perp\|_2^2 \leq S_t(\theta_2) - S_t(\theta_1) - \langle \nabla S_t(\theta_1), \theta_2 - \theta_1 \rangle$. Denote $\eta_1 = \operatorname{argmin}_{x\in\mathcal{D}_t} |\nabla_t^\top(x - \theta_1)|$, and $\eta_2 = \operatorname{argmin}_{x\in\mathcal{D}_t} |\nabla_t^\top(x - \theta_2)|$. Notice that

$$\begin{aligned} S_t(\theta_2) - S_t(\theta_1) =& |\nabla_t^\top(\eta_2 - \theta_2)| - |\nabla_t^\top(\eta_1 - \theta_1)| \\ \leq& |\nabla_t^\top(\eta_1 - \theta_2)| - |\nabla_t^\top(\eta_1 - \theta_1)| \\ \leq& |\nabla_t^\top(\theta_1 - \theta_2)| \\ \leq& \|\nabla_t\|_2 \cdot \|\theta_1 - \theta_2\|_2 \\ \leq& G \cdot \|\theta_1 - \theta_2\|_2. \end{aligned} \tag{H.4}$$

Here the first inequality comes from the definition of η_2, the second inequality is an application of the triangular inequality, the third inequality is derived from Cauchy-Schwarz Inequality, and the last inequality is from Assumption A5 on the Lipschitzness of $\ell_t(\theta)$ over $\mathcal{D}_t$. Therefore, $S_t(\theta)$ is G-Lipschitz as well over $\mathcal{D}_\infty^B$, and we have:

$$\begin{aligned}\mu_2\|\nabla_t^{\perp}\|_2^2 &\leq S_t(\theta_2) - S_t(\theta_1) - \langle \nabla S_t(\theta_1), \theta_2 - \theta_1\rangle \\ &\leq 2G\|\theta_2 - \theta_1\|_2.\end{aligned}$$

This holds for any $\theta_1, \theta_2 \in \mathcal{D}_\infty^B$. However, we may fix θ_1 and θ_2 while also let $\mu_2 \to +\infty$ since it holds for any $\mu_2 \in \mathbb{R}$. If $\|\nabla_t^{\perp}\|_2 \neq 0$ then it will fall into a contradiction. Therefore, we know that $\nabla_t^{\perp} = \mathbf{0}$ and $\nabla S_t(\theta)$ is always in the same direction of ∇_t.

Without losing generality, denote $\nabla S_t(\theta_1) := \lambda \cdot \nabla_t$. In the following, we will prove that $\lambda = \pm 1$ or 0. From Eq. (H.4) line 3, we know that $S_t(\theta_2) - S_t(\theta_1) \leq |\nabla_t^\top(\theta_1 - \theta_2)|$. Combined with the convexity of $S_t(\theta)$, we have:

$$\begin{aligned}|\nabla_t^\top(\theta_1 - \theta_2)| \geq & S_t(\theta_2) - S_t(\theta_1) \\ \geq & \nabla S_t(\theta_1)^\top(\theta_2 - \theta_1) \\ = & \lambda \cdot \nabla_t^\top(\theta_2 - \theta_1).\end{aligned} \tag{H.5}$$

Notice that we can choose arbitrary θ_2 without changing λ, we may let $\theta_2 = 0$ and $\theta_2 = 2\theta_1$ in Eq. (H.5):

$$\pm\lambda \cdot \nabla_t^\top \theta_1 \leq |\nabla_t^\top \theta_1| \tag{H.6}$$

If $\nabla_t^\top \theta_1 \neq 0$, then we have $\lambda \in [-1, 1]$. Otherwise, we know from Eq. (H.5) that $|\nabla_t^\top \theta_2| \geq \lambda \cdot \nabla_t^\top \theta_2, \forall \theta_2$, and similarly we have $\lambda \in [-1, 1]$. Now we denote $\theta_4 := \frac{\theta_1 + \eta_1}{2}$, and we have:

$$\langle \nabla S_t(\theta_1), \theta_4 - \theta_1\rangle + S_t(\theta_1) \leq S_t(\theta_4) \tag{H.7}$$

from the convexity of S_t. And we also have:

$$\begin{aligned}S_t(\theta_4) &= \min_{x \in \mathcal{D}_t} |\nabla_t^\top(x - \theta_4)| \\ &\leq |\nabla_t^\top(\eta_1 - \theta_4)| \\ &= |\nabla_t^\top \frac{\theta_1 - \eta_1}{2}| \\ &= \frac{1}{2} S_t(\theta_1) \\ &= |\nabla_t^\top(\theta_1 - \theta_4)| \\ &= S_t(\theta_1) - |\nabla^\top(\theta_1 - \theta_4)|.\end{aligned} \tag{H.8}$$

Combine Eq. (H.7) and (H.8), we have:

$$\langle \nabla S_t(\theta_1), \theta_4 - \theta_1\rangle \leq S_t(\theta_4) - S_t(\theta_1) = -|\nabla_t^\top(\theta_1 - \theta_4)| \tag{H.9}$$

Plug in $\nabla S_t(\theta_1) = \lambda \nabla_t$ to Eq. (H.9), and we have:

$$\lambda \cdot \nabla_t^\top(\theta_4 - \theta_1) \leq -|\nabla_t^\top(\theta_1 - \theta_4)|. \tag{H.10}$$

According to Eq. (H.10), if $\nabla_t^\top(\theta_4 - \theta_1) > 0$, then we have $\lambda \leq -1$; if $\nabla_t^\top(\theta_4 - \theta_1) < 0$, then we have $\lambda \geq 1$. Since we already know that $\lambda \in [-1, 1]$, then for the two case we should have $\lambda = -1$ or $\lambda = 1$.

Finally, what if $\nabla_t^\top(\theta_4 - \theta_1) = 0$? In this case, it means that $\nabla_t^\top(\eta_1 - \theta_1)/2 = 0$. Since $\eta_1 = \text{argmin}_{x \in \mathcal{D}_t} |\nabla_t^\top(x - \theta_1)|$, we know that $S_t(\theta_1) = 0$ at this time. Since $S_t(\theta) \geq 0, \forall \theta \in \mathbb{R}^d$, we know that $S_t(\theta) \geq S_t(\theta_1) + \mathbf{0}^\top(\theta - \theta_1)$ and as a result $0 \in \partial S_t(\theta_1)$. This in fact holds the lemma.

□

H.1.4 Proof of Lemma 94

Proof. We begin by noticing that $\hat{\ell}_t(\theta)$ is exp-concave over $\mathcal{D}_\infty^B$. To see this, note that by the triangular inequality and Cauchy-Schwarz Inequality,

$$|\nabla \ell_t(\hat{\theta}_t)^\top(\theta - \hat{\theta}_t)\sqrt{\beta} + 1/(2\sqrt{\beta})| \leq |\nabla \ell_t(\hat{\theta}_t)^\top(\theta - \hat{\theta}_t)|\sqrt{\beta} + 1/(2\sqrt{\beta}) \leq 2GB\sqrt{d\beta} + 1/(2\sqrt{\beta}), \tag{H.11}$$

where we use the fact that $\|\nabla \ell_t(\hat{\theta}_t)\|_2 \leq G$ by Assumption A5 and $\|\theta - \hat{\theta}_t\|_2 \leq 2B\sqrt{d}$ as $\theta \in \mathcal{D}_\infty^B$ and $\hat{\theta}_t \in \mathcal{D}_t \subset \mathcal{D}_\infty^B$.

With γ as defined in the statement of the lemma, we have that the losses $\hat{\ell}_t(\theta)$ are 2γ exp-concave over $\mathcal{D}_\infty^B$. [**?**, see]Section 3.3]BianchiBook2006.

Now we proceed to show that the losses $f_t(\theta)$ are in-fact exp-concave with appropriate exp-concavity factor.

For the sake of brevity, let us denote

$$\nabla \hat{\ell}_t(u) = 2\sqrt{\beta}\left(\nabla \ell_t(\hat{\theta}_t)^\top(u - \hat{\theta}_t)\sqrt{\beta} + \frac{1}{2\sqrt{\beta}}\right)\nabla \ell_t(\hat{\theta}_t) \tag{H.12}$$

$$:= p(u)\nabla \ell_t(\hat{\theta}_t). \tag{H.13}$$

We have that for any $u, v \in \mathcal{D}_\infty^B$,

$$\hat{\ell}_t(v) \geq \hat{\ell}_t(u) + p(u)\nabla \ell_t(\hat{\theta}_t)^\top(v - u) \tag{H.14}$$

$$+ \gamma\left(p(u)\nabla \ell_t(\hat{\theta}_t)^\top(v - u)\right)^2. \tag{H.15}$$

Due to convexity, we have

$$S_t(v) \geq S_t(u) + \lambda \nabla \ell_t(\hat{\theta}_t)^\top(v - u), \tag{H.16}$$

for some $\lambda \in \{-1, 0, 1\}$ as per Lemma 93.

Adding Eq.(H.15) and (H.16), we obtain

$$f_t(v) \geq f_t(u) + \nabla f_t(u)^\top (u - v) \tag{H.17}$$

$$+ \gamma p(u)^2 \left(\nabla \ell_t(\hat{\theta}_t)^\top (v - u)\right)^2 \tag{H.18}$$

$$= f_t(u) + \nabla f_t(u)^\top (u - v) \tag{H.19}$$

$$+ \gamma \left(\frac{p(u)}{\lambda + p(u)}\right)^2 \left(\nabla f_t(u)^\top (v - u)\right)^2. \tag{H.20}$$

Next, we proceed to obtain a lower bound on the exp-concavity factor. Note that

$$p(u) \geq 2\sqrt{\beta}\left(-2GB\sqrt{d\beta} + \frac{1}{2\sqrt{\beta}}\right) \geq 2\sqrt{\beta} \cdot \frac{1}{4\sqrt{\beta}} = \frac{1}{2} \tag{H.21}$$

where the first inequality is via Cauchy-Schwarz Inequality and the second inequality holds due to the fact that $\beta \leq 1/(8GB\sqrt{d})$ due to the setting in Theorem 90

Similarly we have that

$$|p(u) + \lambda| \leq 4GB\beta\sqrt{d} + 2 \leq 5/2, \tag{H.22}$$

where in the last line we used $\beta \leq 1/(8GB\sqrt{d})$.

Combining the last two displays, we have that

$$\gamma \left(\frac{p(u)}{\lambda + p(u)}\right)^2 \geq \gamma/25. \tag{H.23}$$

Applying this lower bound to Eq.(H.20) now yields the exp-concavity of $f_t(\theta)$ claimed in the lemma.

Next, we proceed to calculate the Lipschitz constant of f_t. Since $\|\nabla \ell_t(\hat{\theta}_t)\|_2 \leq G$, by Lemma 93 we conclude that $G'S_t(\theta)$ is $G'G$ Lipschitz in L2 norm across $\mathbb{R}^d$. Now using Lemma 92 we conclude that the losses f_t are $2G'G$ are Lipschitz in L2 norm across $\mathcal{D}_\infty^B$. □

H.1.5 Proof of Lemma 97

[**?**,]Lemma 7]xu2021logarithmic has proved the $\frac{C_{down}}{C_{exp}}$-exp-concavity. Here we prove the other claim on Lipschitzness.

Proof. Notice that $\ell_t(\theta)$ is a continuous function. Therefore, for any $\theta_1, \theta_2 \in \mathcal{D}_t$, there exists a $\theta_3 = \epsilon\theta_1 + (1-\epsilon)\theta_2$ for some $\epsilon \in [0,1]$ such that

$$\begin{aligned}\ell_t(\theta_1) - \ell_t(\theta_2) &= \nabla\ell_t(\theta_3)^\top(\theta_1 - \theta_2)\\ &= h_t(\theta_3)x_t^\top(\theta_1 - \theta_2)\\ &\leq h_{\max}\|x_t\|_2\|\theta_1 - \theta_2\|_2\\ &= h_{\max}\|\theta_1 - \theta_2\|_2\\ &= G\|\theta_1 - \theta_2\|_2\end{aligned} \tag{H.24}$$

where $h_{\max}$ is defined in Appendix H.1.1. In Eq.(H.24), the first equality is by Lagrange interpolation, the second equality is by definition of $h_t(\theta)$, the third inequality is by Cauchy-Schwarz Inequality, the fourth equality is by the assumption that $x_t \in \mathcal{D}_2^1$, and the last inequality is from the fact that $h_{\max} = G$. Since $\mathcal{D}_t$ is convex, we know that $\theta_3 \in \mathcal{D}_t$. Therefore, the lemma is proven.

□

H.1.6 Lower Bound Proof (Proof of Theorem 99)

Here we present and prove the following theorem, which is sufficient to show a $\Omega(T^{\frac{1}{3}}C_T^{\frac{2}{3}})$ lower bound for $C_T > \frac{1}{\sqrt{T}}$.

Theorem 242. *Consider a feature-based dynamic pricing problem with $d = 1, x_t = 1, N_t \sim_{i.i.d.} \mathcal{N}(0,1), t = 1,2,\ldots,T$ and $C_T > \frac{1}{\sqrt{T}}$ For any algorithm $\mathcal{A}$ there exists a specific setting such that $\mathcal{A}$ suffer $\Omega(T^{\frac{1}{3}}C_T^{\frac{2}{3}})$ expected regret even with y_t observable.*

The sufficiency comes from the fact that we cannot observe any realized y_t's in the pricing problem (but a binary indicator instead). In comparison, the lower bound in Theorem 242 even works for observable y_t's, which is sufficient to derive the binary feedback (by comparing y_t with v_t).

Proof. To summarize the procedure of proof: Denote $[n] := \{1,2,\ldots,n\}$ for any positive integer n. Define $\theta_0 = 1, \theta_1 = 1 + \delta(T, C_T)$ where $\delta = \frac{1}{40}(\frac{C_T}{T})^{\frac{1}{3}}$ is an additional amount. Then we construct a set $S \subset \{0,1\}^T$ consisting of randomly-sampled $\beta^{(i)} \in \{0,1\}^T, i = 1,2,\ldots,N$ that we will use to construct $\theta_t^*(i)$ series (each i indicating a specific $\{\theta_t^*\}$ series) later. Afterward, we will show that the $\{\theta_t^*(i)\}$ and the $\{\theta_t^*(j)\}$ series are hard to distinguish by any algorithm, and we will further show that a large enough regret caused by this misspecification. In this way, we can prove an expected lower regret bound (where the expectation is also taken over different $\{\theta_t^*(i)\}$).

The process to sample each $\beta^{(i)}$ is as follows: We split $[T]$ uniformly into $m = \frac{C_T}{4\delta}$ intervals, with each length $\frac{4T\delta}{C_T}$. Since $\delta = \frac{1}{40}(\frac{C_T}{T})^{\frac{1}{3}}$ and $C_T \geq \frac{1}{\sqrt{T}}$, we know that $m \geq 10$. Denote these intervals as $I_1, I_2, \ldots, I_m$. For any $\beta^{(i)} \in S$, we construct it in a stochastic process: For each index interval $I_k, k = 1,2,\ldots,m$, we generate a random variable $Z_k^{(i)} \sim Ber(\frac{1}{2})$ independently, and then let $\beta_l^{(i)} = Z_k^{(i)}, \forall l \in I_k$. Denote the vector $Z^{(i)} = [Z_1^{(i)}, Z_2^{(i)}, \ldots, Z_m^{(i)}]^\top \in \{0,1\}^m$, and we know that $\mathbb{E}[\|Z^{(i)} - Z^{(j)}\|_1] = \frac{m}{2}$. Accordingly, we have $\mathbb{E}[\|\beta^{(i)} - \beta^{(j)}\|_1] = \frac{m}{2} \cdot \frac{4T\delta}{C_T} = \frac{T}{2}$.

Therefore, according to Hoeffding's inequality, we have:

$$\begin{aligned} \Pr[|\|Z^{(i)} - Z^{(j)}\|_1 - \frac{m}{2}| \leq \frac{m}{6}] \geq& 1 - 2 \cdot e^{-\frac{(\frac{m}{6})^2}{2m}} \\ \Leftrightarrow \Pr[|\|\beta^{(i)} - \beta^{(j)}\|_1 - \frac{T}{2}| \leq \frac{T}{6}] \geq& 1 - 2 \cdot e^{-\frac{m}{72}}, \forall i,j \in [N]. \end{aligned} \tag{H.25}$$

By applying a union bound over all $\binom{N}{2}$ pairs of $i,j \in [N]$, we have:

$$\Pr[|\|\beta^{(i)} - \beta^{(j)}\|_1 - \frac{T}{2}| \leq \frac{T}{6}, \forall i,j \in [N]] \geq 1 - N^2 \cdot e^{-\frac{m}{72}}. \tag{H.26}$$

Also, we know that $\Pr[\beta^{(i)} \neq \beta^{(j)}] = \Pr[Z^{(i)} \neq Z^{(j)}] = 1 - \frac{1}{2^m}$ for $i \neq j$. By applying a union bound over all $\binom{N}{2}$ pairs of i,j, we have $\Pr[\beta^{(i)} \neq \beta^{(j)}] \geq 1 - \frac{N^2}{2^{m+1}}$. Combining these two probability bounds, we know that in this way we can find a satisfactory set S with probability at least $\Pr \geq 1 - N^2 \cdot (e^{-\frac{m}{72}} + 2^{-(m+1)})$. Let $N = e^{\frac{m}{200}}$ (and therefore $\log N = \frac{m}{200} = \frac{C_T}{800\delta}$), and then $\Pr \geq 1 - N^2 \cdot (e^{-\frac{m}{72}} + 2^{-(m+1)}) \geq 1 - (e^{-\frac{m}{300}} + e^{-\frac{3}{5}m})$. Since the total number of possible S (i.e., any set consisting N (repeatable) vectors $\beta \in \{0,1\}^T$) is $(2^m)^N$ and we are uniformly sampling from this whole family, the expected total number of satisfactory S is at least $(2^m)^N \cdot (1 - (e^{-\frac{m}{300}} + e^{-\frac{3}{5}m}))$. Since $m \geq 10$ as we showed above, we have $(2^m)^N \cdot (1 - (e^{-\frac{m}{300}} + e^{-\frac{3}{5}m})) \geq 2^{10\times 1} \cdot (1 - e^{-\frac{1}{30}} - e^{-6}) = 31.0325 > 1$. As a result, there must exist at least one satisfactory S in the whole possible set family, such that: (1) $\frac{T}{3} \leq \|\beta^{(i)} - \beta^{(j)}\|_1 \leq \frac{2T}{3}$, and (2) $\beta^{(i)} \neq \beta^{(j)}, \forall i \neq j \in [N]$. We here pick this satisfactory S and in the following we use it for further proof.

Now, for each $\beta^{(i)} \in S$, we generate a sequence of parameter $\{\theta_t^*(i)\}$ according to $\beta^{(i)}$: For $t = 1, 2, \ldots, T$, we let $\theta_t^*(i) = 1 + \delta \cdot \beta_t^{(i)}$, i.e., if $\beta^{(i)} = 0$, then $\theta_t^*(i) = \theta_0 = 1$; if $\beta^{(i)} = 1$, then $\theta_t^*(i) = 1 + \delta$. Therefore, we have the following result:

$$\mathrm{TV}(\{\theta_t^*(i)\}) \leq m \cdot \delta = \frac{C_T}{4} < C_T. \tag{H.27}$$

This is because $\|\theta_t^*(i) - \theta_{t+1}^*(i)\| > 0$ only if $\exists k \in [m]$ s.t. $t \in I_k, t+1 \in I_{k+1}$. As a result, the total variation of this $\{\theta_t^*(i)\}$ satisfies the upper bound C_T.

Now, let us consider the realized valuation sequence $\{y_t\}$. For any $i \in [N]$, denote

$$\mathbf{y}(i) := [x_1(1 + \beta_1^{(i)}\delta) + N_1, x_2(1 + \beta_2^{(i)}\delta) + N_2, \ldots, x_T(1 + \beta_T^{(i)}) + N_T]^\top$$

Let us denote the distribution of $\mathbf{y}(i)$ as $\mathbb{P}_i, i = 1, 2, \ldots, N$. Recall that $x_t = 1$ and $N_t \sim \mathcal{N}(0,1), \forall t$, and we have $\mathbb{P}_i = [\mathcal{N}(1 + \beta_1^{(i)}\delta, 1), \mathcal{N}(1 + \beta_2^{(i)}\delta, 1), \ldots, \mathcal{N}(1 + \beta_T^{(i)}\delta)]^\top$. Consider the difference between $\mathbb{P}_i$ and $\mathbb{P}_j$ while fixing $\beta^{(i)}$ and $\beta^{(i)}$, and for any $i, j \in$

$[N], i \neq j$ we have:

$$
\begin{aligned}
KL[\mathbb{P}_i||\mathbb{P}_j] =& \sum_{t=1}^{T} KL[\mathcal{N}(1+\beta_t^{(i)}\delta, 1)||\mathcal{N}(1+\beta^{(j)}\delta, 1)] \\
=& \sum_{t=1}^{T} \frac{(\beta_t^{(i)} - \beta_t^{(j)})^2\delta^2}{2} \\
=& \frac{\delta^2}{2} \cdot \|\beta^{(i)} - \beta^{(j)}\|_2^2 \\
=& \frac{\delta^2}{2} \cdot \|\beta^{(i)} - \beta^{(j)}\|_1.
\end{aligned} \tag{H.28}
$$

Again, the KL-divergence is conditioning on $\beta^{(i)}$ and $\beta^{(j)}$. Here the first line is from the fact that y_t's are independent for every t, the second line is by $x_t = 1$, the third line is from the fact that $KL[\mathcal{N}(\mu_1, \sigma_1)||\mathcal{N}(\mu_2, \sigma_2)] = \log\frac{\sigma_2}{\sigma_1} + \frac{\sigma_1^2+(\mu_1-\mu_2)^2}{2\sigma_2^2} - \frac{1}{2}$, the fourth line is by calculation and the fifth line is from that $|\beta_t^{(i)} - \beta_t^{(j)}| \in \{0, 1\}$.

Here we introduce a Fano's Inequality as the following proposition:

Proposition 243 (Fano's Inequality). *Let $X_1, X_2, \ldots, X_n \sim_{i.i.d.} \mathbb{P}$ where $\mathbb{P} \in \{\mathbb{P}_1, \mathbb{P}_2, \ldots, \mathbb{P}_N\}$ is a distribution family. Let Ψ be any function of $X_1, X_2, \ldots, X_n$ taking values in $\{1, 2, \ldots, N\}$. Let $\alpha = \max_{j\neq k} KL(\mathbb{P}_j||\mathbb{P}_k)$.*[1] *Then*

$$
\frac{1}{N}\sum_{j=1}^{N} \mathbb{P}_j(\Psi \neq j) \geq 1 - \frac{n\alpha + \log 2}{\log N}. \tag{H.29}
$$

According to Fano's Inequality, we have:

$$
\inf_{\Psi:\mathbb{R}^T\rightarrow\{1,2,\ldots,N\}} \sup_{i\in\{1,2,\ldots,N\}} \mathbb{P}_i(\Psi \neq i) \geq \inf_{\Psi} \frac{1}{N}\sum_{i=1}^{N} \mathbb{P}_i(\Psi \neq i) \geq 1 - \frac{n\alpha + \log 2}{\log N} \geq \frac{1}{2} - \frac{n\alpha}{\log N}. \tag{H.30}
$$

Here $n = 1$ since only one specific $\mathbf{y}(i)$ covers the whole time series and is only sampled once, and $\alpha = \max_{i,j\in[N],i\neq j} KL[\mathbf{y}(i)||\mathbf{y}(j)] = \max_{i,j\in[N],i\neq j} \frac{\delta^2}{2} \cdot \|\beta^{(i)} - \beta^{(j)}\|_1 \leq \frac{\delta^2 T}{3}$ is the upper bound of KL-divergences on different distributions. Now we specify the function $\Psi_{\mathcal{A}}$ for any pricing algorithm $\mathcal{A}$: At each round $t = 1, 2, \ldots, T$, suppose the algorithm $\mathcal{A}$ proposes a price $v_t^{\mathcal{A}}$. Define a vector $\mathbf{w} = [w_1, w_2, \ldots, w_T]^\top$ where $w_t = \mathbb{I}[v_t^{\mathcal{A}} \geq \frac{J(\theta_0)+J(\theta_1)}{2}]$

[1] Usually it is denoted as β, but here we denote it as α for clarity, since we have already defined $\beta^{(i)}$ as vectors in S.

is a Boolean value. Then we let $\Psi_{\mathcal{A}} = \text{argmin}_i \in [N]\|\mathbf{w} - \beta^{(i)}\|_1$. Therefore, we have:

$$\begin{aligned} 2 \cdot \|\mathbf{w} - \beta^{(j)}\|_1 \geq & \|\beta^{(\Psi_{\mathcal{A}})} - \mathbf{w}\|_1 + \|\mathbf{w} - \beta^{(j)}\|_1 \\ \geq & \|\beta^{(\Psi_{\mathcal{A}})} - \beta^{(j)}\|_1, \forall j \in [N], j \neq \Psi_{\mathcal{A}} \\ \geq & \frac{T}{6} \end{aligned} \tag{H.31}$$

Here the first inequality is from the optimality of $\Psi_{\mathcal{A}}$, the second inequality is from the triangular inequality, and the third inequality is from the Hoeffding bound in Eq. (H.25). Therefore we know that if $\Psi_{\mathcal{A}} \neq i$ then we have $\|\mathbf{w} - \beta^{(i)}\|_1 \geq \frac{T}{12}$, which further leads to

$$\begin{aligned} & \sum_{t=1}^{T} (v_t^{\mathcal{A}} - J(x_t\theta_t^*(i)))^2 \\ \geq & \sum_{t=1}^{T} (\mathbb{I}[w_t = 1]\mathbb{I}[\beta_t^{(i)} = 0] + \mathbb{I}[w_t = 0]\mathbb{I}[\beta_t^{(i)} = 1])(v_t^{\mathcal{A}} - J(x_t\theta_t^*(i)))^2 \\ = & \sum_{t=1}^{T} \mathbb{I}[v_t^{\mathcal{A}} \geq \frac{J(\theta_0) + J(\theta_1)}{2}]\mathbb{I}[\beta_t^{(i)} = 0](v_t^{\mathcal{A}} - J(\theta_0))^2 + \\ & \mathbb{I}[v_t^{\mathcal{A}} < \frac{J(\theta_0) + J(\theta_1)}{2}]\mathbb{I}[\beta_t^{(i)} = 1](J(\theta_1 - v_t^{\mathcal{A}}))^2 \\ \geq & \sum_{t=1}^{T} \mathbb{I}[|w_t - \beta_t^{(i)}| = 1](\frac{J(\theta_1) - J(\theta_0)}{2})^2 \\ = & \|\mathbf{w} - \beta^{(i)}\|_1(\frac{J(\theta_1) - J(\theta_0)}{2})^2 \\ \geq & \frac{T}{12} \cdot (\frac{J(\theta_0) - J(\theta_1)}{2})^2. \end{aligned} \tag{H.32}$$

Here the first line is because we omit the case where $\mathbb{I}[w_t = \beta_t^{(i)}]$, the second line is from the definition of w_t, the third line is from the facts of $\theta_0 < \theta_1$ and $J'(u) > 0, \forall u \in \mathbb{R}$, the fourth line is by the definition of L_1-norm and the last line is from the fact we mentioned prior to this equation. Now we propose a lemma of properties:

Lemma 244 (Properties of $g(v,u)$ and $J(u)$)**.** *For $g(v,u)$ and $J(u)$ with $N_t \sim \mathcal{N}(0,1)$, we have:*

1. *$J(u) > u$ when $u \in (0, \sqrt{\frac{\pi}{2}})$ and $J(u) < u$ when $u \in (\sqrt{\frac{\Pi}{2}}, +\infty)$.*
2. *$\exists c_J > 0$ s.t. $J'(u) \geq c_J, \forall u \in [-B, B]$.*
3. *$\exists c_g > 0$ s.t. $g(J(u), u) - g(v, u) \geq c_g(J(u) - v)^2, \forall v \in [0, B + J(B)]$.*

We will show the proof of Lemma 244 by the end of this section. With Lemma 244, when $\Psi_A \neq i$, we have:

$$\begin{aligned}
\mathbf{Reg}_{\mathcal{A}} =& \sum_{t=1}^{T} g(J(x_t\theta_t^*(i)), x_t\theta_t^*(i)) - g(v_t^{\mathcal{A}}, x_t\theta_t^*(i)) \\
\geq& \sum_{t=1}^{T} c_g(v_t^{\mathcal{A}} - J(x_t\theta_t^*(i)))^2 \\
\geq& c_g \cdot \frac{T}{12} \cdot (\frac{J(\theta_0) - J(\theta_1)}{2})^2 \\
\geq& c_g \cdot \frac{T}{12} \cdot \frac{c_J^2}{4} \cdot (\theta_1 - \theta_0)^2 \\
\geq& \frac{c_g c_J^2 \cdot T\delta^2}{48}.
\end{aligned} \tag{H.33}$$

Finally, let $\delta = \frac{1}{40}(\frac{C_T}{T})^{\frac{1}{3}}$, and according to Eq. (H.28),(H.30) and (H.33), we have:

$$\begin{aligned}
\mathbb{E}[\mathbf{Reg}_{\mathcal{A}}] \geq& \sup_{i\in[N]} \mathbb{P}_i(\Psi_{\mathcal{A}} \neq i) \cdot (\sum_{t=1}^{T} g(J(x_t\theta_t^*(i)), x_t\theta_t^*(i)) - g(v_t^{\mathcal{A}}, x_t\theta_t^*(i))) \\
\geq& \sup_{i\in[N]} \mathbb{P}_i(\Psi_{\mathcal{A}} \neq i) \cdot \frac{c_g c_J^2 \cdot T\delta^2}{48} \\
\geq& (\frac{1}{2} - \frac{n\alpha}{\log N}) \cdot \frac{c_g c_J^2 \cdot T\delta^2}{48} \\
=& (\frac{1}{2} - \frac{\frac{\delta^2 T}{3}}{\frac{C_T}{800\delta}}) \cdot c_g c_J^2 \frac{\cdot T\delta^2}{48} \\
=& c_g c_J^2 (\frac{1}{2} - \frac{800}{3} \cdot \frac{\delta^3 T}{C_T}) \cdot \frac{\cdot T\delta^2}{48} \\
=& \frac{c_g c_J^2}{48} (\frac{1}{2} - \frac{1}{240}) \frac{T \cdot (\frac{C_T}{T})^{\frac{2}{3}}}{144} \\
\geq& \frac{c_g c_J^2}{307200} \cdot C_T^{\frac{2}{3}} T^{\frac{1}{3}}.
\end{aligned}$$

This holds the theorem. □

Also, since our upper regret bound w.r.t. T and C_T is $\tilde{O}(1)$ when $C_T \leq \frac{1}{\sqrt{T}}$, which is trivial up to $\log T$ and d factors, we may conclude that our upper regret bound of $\tilde{O}(T^{\frac{1}{3}} C_T^{\frac{2}{3}} \vee 1)$ is optimal with respect to T and C_T.

Proof of Lemma 244. We here prove each of them.

1. According to [118], we know that $u - J(u)$ is monotonically increasing since $J'(u) \in (0, 1)$. Also, since $\frac{\partial g(v,u)}{\partial v}|_{v=J(u)} = 1 - F(J(u) - u) - J(u) \cdot f(J(u) - u) = 0$, we have

$J(\sqrt{\frac{\pi}{2}}) = \sqrt{\frac{\pi}{2}}$. Therefore, $u - J(u) > 0$ when $u > \sqrt{\frac{\pi}{2}}$ and $u - J(u) < 0$ when $0 < u < \sqrt{\frac{\pi}{2}}$.

2. From [118], we know that $J'(u) = 1 + \frac{1}{\phi'(\phi^{-1}(u))} \in (0,1), \forall u \in \mathbb{R}$ where $\phi(\omega) = \frac{1-F(\omega)}{f(\omega)} - \omega$ is invertible and smooth for standard Gaussian distribution. Therefore, we know that $J'(u)$ is continuous. Therefore, $\exists c_J > 0$ such that $\inf_{u\in[-B,B]} J'(u) = c_J$.

3. From the optimality of $J(u)$ we know that $\frac{\partial g(v,u)}{\partial v}|_{v=J(u)} = 1 - F(J(u)-u) - J(u) \cdot f(J(u)-u) = 0$. Define $q(u) := 1 - F(J(u)-u) - J(u) \cdot f(J(u)-u)$. Since $q(u) = 0, \forall u \in \mathbb{R}$, we have:

$$\begin{aligned}
&\frac{\partial q(u)}{\partial u} = 0 \\
\Leftrightarrow &\left(J'(u)(J(u)^2 - u \cdot J(u) - 2) - (J(u)^2 - u \cdot J(u) - 1)\right) f(J(u)-u) = 0 \\
\Leftrightarrow &J'(u) = 1 + \frac{1}{J(u)^2 - u \cdot J(u) - 2}.
\end{aligned} \tag{H.34}$$

The second line is by standard Gaussian noises and some calculations, and the third line is from the fact that $f(x) > 0$ for standard Gaussian distribution. Since we already know that $J'(u) \in (0,1)$, we may then realized that $J(u)^2 - u \cdot J(u) - 2 < -1$. Notice that $\frac{\partial^2 g(v,u)}{\partial v^2} = (v^2 - vu - 2)f(v-u)$ for standard gaussian noise. Therefore, we have $\frac{\partial^2 g(v,u)}{\partial v^2} = (J(u)^2 - u \cdot J(u) - 2)f(J(u)-u) \leq (-1) \cdot f_{\min}$¡0 where $f_{\min}$ has been defined in Appendix H.1.1 as the universal lower bound of f. This means that $g(v,u)$ is $f_{\min}$-strongly concave at $v = J(u)$, which further leads to the fact that there exists a neighborhood $v \in [J(u) - B_u, J(u) + B_u]$ with constant[2] B_u such that $\frac{\partial^2 g(v,u)}{\partial v^2} \leq -\frac{f_{\min}}{2}$. As a result, for $v \in [J(u) - B_u, J(u) + B_u]$ we have

$$\begin{aligned}
g(J(u),u) - g(v,u) &= -\frac{\partial g(v,u)}{\partial v}|_{v=J(u)}(J(u) - u) \\
&\quad - \frac{1}{2} \cdot \frac{\partial^2 g(v,u)}{\partial v^2}|_{v=v'\in[J(u),v] \text{ or } [v,J(u)]}(J(u) - v)^2 \\
&\geq -\frac{1}{2}(-\frac{f_{\min}}{2})(J(u) - v)^2 \\
&= \frac{f_{\min}}{4}(J(u) - v)^2.
\end{aligned}$$

Now, let us consider the case when $v \in [0, B + J(B)]$ but $v \notin [J(u) - B_u, J(u) + B_u]$. On the one hand, $(J(u) - v)^2 \leq (B + J(B) - (-B))^2 = (2B + J(B))^2$. On the other hand, $g(J(u),u) - g(v,u) \geq g(J(u),u) - \max\{g(J(u) - B_u, u), g(J(u) + B(u), u)\} > 0$. Denote $c_u := \inf_{u\in[-B,B]}\{g(J(u),u) - \max\{g(J(u) - B_u, u), g(J(u) + B(u), u)\}\}$,

[2] B_u can be defined as the inferior of all B_u over all $u \in [-B, B]$ and is still a positive constant.

and we have $c_u > 0$. Therefore, we have:

$$g(J(u), u) - g(v, u) \geq c_u \geq \frac{c_u}{(2B + J(B))^2}(2B + J(B))^2 \geq \frac{c_u}{(2B + J(B))^2}(J(u) - v)^2. \tag{H.35}$$

Finally, let $c_g = \min\{\frac{f_{\min}}{4}, \frac{c_u}{(2B+J(B))^2}\}$, and we have proved the lemma.

□

Appendix I

Supplementary Materials for Chapter 10

I.1 Omitted Proofs

In this section we use the notations defined in Fig.10.1.

The lemma below shows how the surrogate losses ℓ_t can be used to upper bound the regression losses f_t.

Lemma 245. *Assume the notations in Fig.10.1. Let G be such that $\sup_{w_1,w_2 \in \mathcal{D}_\infty(\tilde{R})} \|A_t(w_1 + w_2) - 2b_t\|_1 \le G$ for all $t \in [n]$. We have that:*

- $f_t(\hat{w}_t) \le \ell_t(w_t)$,
- $f_t(u) = \ell_t(u)$ *for all* $u \in \mathcal{D}$

Proof. For any $w_1, w_2 \in \mathcal{D}_\infty(\tilde{R})$

$$
\begin{aligned}
f_t(w_1) - f_t(w_2) &= \|A_t w_1 - b_t\|_2^2 - \|A_t w_2 - b_t\|_2^2 && \text{(I.1)}\\
&= \left(A_t(w_1 + w_2) - 2b_t\right)^T \left(A_t(w_1 - w_2)\right) && \text{(I.2)}\\
&\le \|A_t(w_1 + w_2) - 2b_t\|_1 \|A_t(w_1 - w_2)\|_\infty && \text{(I.3)}\\
&\le G \max_{i=1,\ldots,p} |a_{t,i}^T(w_1 - w_2)|, && \text{(I.4)}
\end{aligned}
$$

for a G such that $\sup_{w_1,w_2 \in \mathcal{D}_\infty(\tilde{R})} \|A_t(w_1 + w_2) - 2b_t\|_1 \le G$ holds true.

In particular we have that:

$$
f_t(\hat{w}_t) \le f_t(w_t) + G \max_{i=1,\ldots,p} |a_{t,i}^T(\hat{w}_t - w_t)| := \ell_t(w_t) \qquad \text{(I.5)}
$$

For any $u \in \mathcal{D}$, we have that $S_t(u) = 0$. Hence $f_t(u) = \ell_t(u)$.

□

The lemma below establishes certain useful properties of the barrier function $S_t(w)$.

Lemma 246. *The function $S_t(w)$ satisfies the following properties:*

1. $S_t(w) = \max_{i=1,\dots,p} \min_{x\in\mathcal{D}} |a_{i,t}^T(x-w)|$.
2. $S_t(w)$ *is convex over* $\mathbb{R}^d$.
3. *Let i^* be such that $S_t(w) = \min_{x\in\mathcal{D}} |a_{i^*,t}^T(x-w)|$. Let $\Pi(w) \in \operatorname{argmin}_{x\in\mathcal{D}} |a_{i^*,t}^T(x-w)|$. Let $g_t \in \partial S_t(w)$, When $a_{i^*,t}^T(\Pi(w)-w) \neq 0$ we have:*

$$g_t = \begin{cases} a_{i^*,t}, & \text{if} \quad a_{i^*,t}^\top(\Pi(w)-w) < 0 \\ -a_{i^*,t}, & \text{if} \quad a_{i^*,t}^\top(\Pi(w)-w) > 0. \end{cases}$$

If $a_{i^*,t}^T(\Pi(w)-w) = 0$ then we take $g_t = 0$.

Proof. We set out to prove the first statement. Let Δ_p be the p dimensional simplex. We have that

$$S_t(w) = \min_{x\in\mathcal{D}} \max_{i=1,\dots,p} |a_{i,t}^T(x-w)| \tag{I.6}$$

$$=_{(a)} \min_{x\in\mathcal{D}} \max_{v\in\Delta_p} \sum_{i=1}^{p} v_i |a_{i,t}^T(x-w)| \tag{I.7}$$

$$=_{(b)} \max_{v\in\Delta_p} \min_{x\in\mathcal{D}} \sum_{i=1}^{p} v_i |a_{i,t}^T(x-w)|. \tag{I.8}$$

For line (a) we observed that for a given x $\max_{v\in\Delta_p} \sum_{i=1}^p v_i |a_{i,t}^T(x-w)|$ is attained by putting all the weights of v to an $i^* \in \operatorname{argmax}_{i=1,\dots,p} |a_{i,t}^T(x-w)|$.

For line (b) we observe that the function $r(x,v) = \sum_{i=1}^p v_i |a_{i,t}^T(x-w)|$ is a convex function of x and concave function of p. So by applying Sion's minimax theorem we arrive at line (b).

Next we set out to prove that:

$$\max_{v\in\Delta_p} \min_{x\in\mathcal{D}} r(x,v) = \max_{i=1,\dots,p} \min_{x\in\mathcal{D}} |a_{i,t}^T(x-w)| \tag{I.9}$$

Let (x^*, v^*) be a solution that attains $\max_{v\in\Delta_p} \min_{x\in\mathcal{D}} r(x,v)$. Further, for the sake of contradiction, let's assume that $v^* \neq e_k$ for any $k \in [p]$. (e_k is the unit vector with 1 at entry k). Let the index j be such that $|a_{j,t}^T(x^*-w)| > |a_{i,t}^T(x^*-w)|$ for all $i \in [p]\setminus\{j\}$. Then we can find a solution e_j such that $r(x^*, e_j) > r(x^*, v^*)$. This contradicts the fact that (x^*, v^*) is a valid solution.

In the alternate case let j be an index in $[p]$ such that $|a_{j,t}^T(x^*-w)| \geq |a_{i,t}^T(x^*-w)|$ for all $i \in [p]\setminus\{j\}$. Suppose for all $i \in Q \subseteq [p]\setminus\{j\}$ we have $|a_{j,t}^T(x^*-w)| = |a_{i,t}^T(x^*-w)|$. By earlier arguments, we must have $v^*[k]$ must be equal to zero for all $k \in [p]\setminus(Q\cup\{j\})$.

Then putting all the weight to j produces an equally valid solution in the sense that $r(x^*, e_j) = r(x^*, v^*)$

Combining the above two cases, we conclude that there exists maximizers v^* such that $v^* = e_k$ for some $k \in [p]$. This leads to Eq.(I.9).

Next we prove statement 2. For any given i we have that $|a_{i,t}^T(x - w)|$ is a convex function of both x and w. Hence the point-wise maximum $\max_{i=1,\ldots,p} |a_{i,t}^T(x - w)|$ is also convex in both x and w. Since partial minimisation preserves convexity, we have that $\min_{x\in\mathcal{D}} \max_{i=1,\ldots,p} |a_{i,t}^T(x - w)|$ remains convex in $w \in \mathbb{R}^d$.

Next we prove statement 3. We know that sub-gradient set of point-wise maximum of convex functions is the convex hull of sub-gradients of the active functions. Applying this result along with the sub-gradient characterization of the function $\min_{x\in\mathcal{D}} |a_{i,t}^T(x-w)|$ in Lemma 248 leads to the third statement.

□

The next lemma establishes the exp-concavity of the surrogate losses ℓ_t over the decision domain of the surrogate algorithm $\mathcal{A}$.

Lemma 247. *Assume the notations in Fig.10.1. Let L be such that $\sup_{w\in\mathcal{D}_\infty(\tilde{R}),j\in[p]} 2\|A_t w - b_t\|_2^2 + 2G^2 \leq L$ for all $t \in [n]$. Then the losses ℓ_t are exp-concave over $\mathcal{D}_\infty(\tilde{R})$ with parameter $1/4L$.*

Proof. Observe that $\nabla f_t(w) = 2A_t^T(A_t w - b_t)$ and $\nabla^2 f_t(w) = 2A_t^T A_t$.

We have that for any $w_1, w_2 \in \mathbb{R}^d$

$$f_t(w_2) = f_t(w_1) + \langle \nabla f_t(w_1), w_2 - w_1 \rangle + \frac{1}{2}\|w_2 - w_1\|_{2A_t^T A_t}^2. \tag{I.10}$$

Due to the convexity of $S_t(w)$ over $\mathbb{R}^d$ from Lemma 246, we have that

$$S_t(w_2) \geq S(w_1) + \langle \nabla S_t(w_1), w_2 - w_1 \rangle. \tag{I.11}$$

Combining Eq.(I.10) and (I.11) we have that

$$\ell_t(w_2) \geq \ell_t(w_1) + \langle \nabla \ell_t(w_1), w_2 - w_1 \rangle + \frac{1}{2}\|w_2 - w_1\|_{2A_t^T A_t}^2 \tag{I.12}$$

Observe that $\nabla \ell_t(w_1) = 2A_t^T(A_t w_t - b_t) + GhA_t^T e_j$, for some $h \in \{-1, 0, 1\}$ and $j \in [p]$ due to Lemma 246. Now, let's focus on points $w_1, w_2 \in \mathcal{D}_\infty(\tilde{R})$. We have

$$\begin{aligned}
\nabla \ell_t(w_1) \nabla \ell_t(w_1)^T &= 4A_t^T(A_t w_1 - b_t + Ghe_j)(A_t w_1 - b_t + Ghe_j)^T A_t && \text{(I.13)}\\
&\preccurlyeq 4LA_t^T A_t, && \text{(I.14)}
\end{aligned}$$

L is such that:

$$\sup_{w \in \mathcal{D}_\infty(\tilde{R}), j \in [p]} \|(A_t w - b_t + Ghe_j)\|_2^2 \leq L. \tag{I.15}$$

Hence for all $w_1, w_2 \in \mathcal{D}_\infty(\tilde{R})$, we have the relation

$$\ell_t(w_2) \geq \ell_t(w_1) + \langle \nabla \ell_t(w_1), w_2 - w_1 \rangle + \frac{1}{4L} \|w_2 - w_1\|^2_{\nabla \ell_t(w_1) \nabla \ell_t(w_1)^T}. \tag{I.16}$$

Thus the losses ℓ_t remains exp-concave over $\mathcal{D}_\infty(\tilde{R})$ with parameter $1/4L$.

□

We are now ready to prove Theorem 104.

Theorem 104. *Let $u_{1:n} \in \mathcal{D}$ be any comparator sequence. In Fig.10.1, choose G such that $\sup_{w_1, w_2 \in \mathcal{D}_\infty(\tilde{R}), t \in [n]} \|A_t(w_1 + w_2) - 2b_t\|_1 \leq G$. Let α be as in Assumption 2. Let L be such that $\sup_{w \in \mathcal{D}_\infty(\tilde{R}), j \in [p]} 2\|A_t w - b_t\|_2^2 + 2G^2 \leq L$ for all $t \in [n]$. Choose $\mathcal{A}$ as the algorithm from [120] with parameters $\gamma = 2G\alpha\tilde{R}\sqrt{d/8L} + \sqrt{2L}$ and $\zeta = \min\{\frac{1}{16G\alpha\tilde{R}\sqrt{d}}, 1/(4\gamma^2)\}$ and decision set $\mathcal{D}_\infty(\tilde{R})$. Under Assumptions 1 and 2, a valid of assignment of G and L are $2p\chi + 2\sigma$ and $6(p\chi + \sigma)^2$ respectively.*

Then the algorithm ProDR.control yields a dynamic regret rate of

$$\sum_{t=1}^{n} f_t(\hat{w}_t) - f_t(u_t) = \tilde{O}(d^3 n^{1/3} [\mathcal{TV}(u_{1:n})]^{2/3} \vee 1), \tag{10.11}$$

where $(a \vee b) := \max\{a, b\}$.

Proof. From Eq.(I.4) we have that for any $w_1, w_2 \in \mathcal{D}_\infty(\tilde{R})$

$$f_t(w_1) - f_t(w_2) \leq G\alpha \|w_1 - w_2\|_2, \tag{I.17}$$

for a G such that $\sup_{w_1, w_2 \in \mathcal{D}_\infty(\tilde{R})} \|A_t(w_1 + w_2) - 2b_t\|_1 \leq G$ holds true.

From Lemma 246 we have for any subgradient $\|\nabla S_t(w)\|_2 \leq \alpha$ (where α is as in Assumption 1). Thus the losses ℓ_t are $2G\alpha$-Lipschitz in L2 norm over $\mathcal{D}_\infty(\tilde{R})$. Now combining Lemma 247 and Theorem 10 in [120] we have that

$$\sum_{t=1}^{n} \ell_t(w_t) - \ell_t(u_t) = \tilde{O}\left((d^3 G^2 \alpha^2 \tilde{R}^2 / L + d^2 G^2 \alpha^2 \tilde{R}^2 + d^2 L)(n^{1/3} [\mathcal{TV}(u_{1:n})]^{2/3} \vee 1) \right)) \tag{I.18}$$

$$= \tilde{O}(d^3 n^{1/3} [\mathcal{TV}(u_{1:n})]^{2/3} \vee 1). \tag{I.19}$$

Applying Lemma 245 now concludes the proof. □

Lemma 248. *Let $f(x) = min_{u \in \mathcal{D}} |a^T(u - x)|$ for a compact and convex set $\mathcal{D}$. Let $0 \in \mathcal{D}$. f(x) is convex. Let $s \in argmin_{u \in \mathcal{D}} |a^T(u - x)|$.*

$$\nabla f(x) = \begin{cases} -a & a^T(s-x) > 0 \\ a & a^T(s-x) < 0 \\ 0 & o.w \end{cases} \tag{I.20}$$

Proof. First we argue the convexity of f. Observe that

$$f(x) = \min_{u \in \mathcal{D}} |a^T(u - x)| \tag{I.21}$$
$$= \min_{u \in \mathcal{D}} \|u - x\|_{aa^T}. \tag{I.22}$$

The norm $\|u - x\|_{aa^T}$ is convex in both u and x across $\mathbb{R}^d$. So we have that $f(x)$ which is obtained by partial minimization of a convex function across a convex domain remains convex over $\mathbb{R}^d$. It follows that for any $x, y \in \mathbb{R}^d$,

Now let x be such that $\nabla f(x) = 0$. Existence of such a point is guaranteed since $\mathcal{D}$ in the definition of f is compact.

$$f(y) \geq f(x) + \nabla f(x)^T(y - x). \tag{I.23}$$

We proceed to show the Lipschitzness of f. Let $w \in argmin_{u \in \mathcal{D}} |a^T(u - x)|$. We have

$$f(y) - f(x) = \min_{u \in \mathcal{D}} |a^T(u - y)| - \min_{u \in \mathcal{D}} |a^T(u - x)| \tag{I.24}$$
$$\leq |a^T(w - x)| - |a^T(w - y)| \tag{I.25}$$
$$\leq |a^T(x - y)| \tag{I.26}$$
$$\leq \|a\|_2 \|x - y\|_2. \tag{I.27}$$

Since $\|a\|_2 \leq \kappa$, we conclude that the function f is κ Lipschitz.

We argue that $\nabla f(x) = \lambda a$ for some scalar λ. Let b be a such that $a^T b = 0$. Let $z = y + \sigma b$. Notice that by the definition of f, we have that $f(y) = f(z)$. So,

$$f(z) = f(y) \tag{I.28}$$
$$\geq f(x) + \nabla f(x)^T(z - x) \tag{I.29}$$
$$= f(x) + \nabla f(x)^T(y - x) + \sigma \nabla f(x)^T b. \tag{I.30}$$

The above inequality must hold for any σ. Note that both $f(y)$ and $f(x)$ is bounded for any two points in $x, y\mathbb{R}^d$. Further, $\nabla f(x)^T(y - x)$ is also bounded due to the Lipschitzness of f. So if $\nabla f(x)^T b$ is not zero, we can choose a σ such that inequality is violated, leading to a contradiction in the convexity of f across $\mathbb{R}^d$.

So $\nabla f(x)^T b = 0$. This implies that $\nabla f(x) = \lambda(x) a$ for some scalar $\lambda(x)$ and for any $x \in \mathbb{R}^d$.

Next, we argue that $\lambda(x) \in [-1, 1]$. Combining Eq.(I.23) and (I.27) we have

$$|a^T(x-y)| \geq \nabla f(x)^T (y-x), \tag{I.31}$$

for all $x, y \in \mathbb{R}^d$. So taking $y = 0$ followed by $y = 2x$ leads to

$$|a^T x| \geq \pm \lambda(x) a^T x. \tag{I.32}$$

Suppose x is chosen such that $a^T x \neq 0$. Then the above inequality implies that $\lambda(x) \in [-1, 1]$.

Let $w \in argmin_{u \in \mathcal{D}} |a^T(u-x)|$. Let $s = (x+w)/2$. We have that

$$f(s) \geq f(x) + \lambda(x) a^T (s-x). \tag{I.33}$$

Moroever,

$$\begin{aligned} f(s) &\leq |a^T(w-s)| && \text{(I.34)} \\ &= \frac{1}{2}|a^T(x-w)| && \text{(I.35)} \\ &= f(x) - |a^T(x-s)|. && \text{(I.36)} \end{aligned}$$

Combining Eq.(I.33) and (I.36), we obtain

$$-|a^T(s-x)| \geq \lambda(x) a^T (s-x). \tag{I.37}$$

Recall that when $a^T x \neq 0$, $\lambda(x) \in [-1, 1]$.

So we conclude that if $a^T x \neq 0$ and $a^T(s-x) > 0$, then $\lambda(x) \leq -1$. This implies that $\lambda(x) = -1$ as $\lambda(x) \in [-1, 1]$ holds true.

Similarly if $a^T x \neq 0$ and $a^T(s-x) < 0$, then $\lambda(x) \geq 1$. This implies that $\lambda(x) = 1$ as $\lambda(x) \in [-1, 1]$ holds true.

Now if $a^T x \neq 0$ and $a^T(s-x) = 0$, we can choose $\lambda(x) = 0$ as $f(z) \geq f(x) + \lambda(x) a^T (z-x) = 0$ holds true for any z.

If $a^T x = 0$, $0 \in argmin_{u \in \mathcal{D}} |a^T(u-x)|$ as $0 \in \mathcal{D}$ is assumed to be true. So by using the previous line of arguments we conclude that $\lambda(x) = 0$. □

Theorem 108. *Let x_t be the prediction of the algorithm in Fig. 10.2 at time t. Instantiating each ProDR.control instance by the parameter setting described in Theorem 104.*

Let τ be the feedback delay. We have that

$$\sum_{t=1}^{n} f_t(x_t) - f_t(u_t) = \tilde{O}(d^3 \tau^{2/3} n^{1/3} [\mathcal{TV}(u_{1:n})]^{2/3} \vee \tau). \tag{10.12}$$

Further for any interval $[a, b] \subseteq [n]$:

$$\sum_{t=a}^{b} f_t(x_t) - f_t(u) = O(d^{1.5} \tau \log n). \tag{10.13}$$

Proof. By following the arguments in [131], we have that

$$\sum_{t=1}^{n} f_t(x_t) - f_t(u_t) = \sum_{i=1}^{\tau} \sum_{k=1}^{\lfloor 1+\frac{n-i}{\tau} \rfloor} f_t(x_{i+(k-1)\tau}) - f_t(u_{i+(k-1)\tau}). \tag{I.38}$$

The second summation in the above expression is the dynamic regret of instance i wrt comparator sequence $\{u_{i+(k-1)\tau}\}$ with k ranging from 1 to $\lfloor 1 + \frac{n-i}{\tau} \rfloor$. Now by triangle inequality we have that

$$\sum_{k=2}^{\lfloor 1+\frac{n-i}{\tau} \rfloor} \|u_{i+(k-1)\tau} - i + (k-2)\tau\|_1 \leq \sum_{t=2}^{n} \|u_t - u_{t-1}\|_1 = \mathcal{TV}(u_{1:n}). \tag{I.39}$$

Thus by Theorem 104 we have

$$\begin{aligned} \sum_{t=1}^{n} f_t(x_t) - f_t(u_t) &\leq \sum_{i=1}^{\tau} \tilde{O}(d^3 (n/\tau)^{1/3} \vee 1) && \text{(I.40)} \\ &\leq \tilde{O}(d^3 \tau^{2/3} n^{1/3} [\mathcal{TV}(u_{1:n})]^{2/3} \vee \tau). && \text{(I.41)} \end{aligned}$$

□

Next, we provide the version of Corollary 109 indicating the closed form expression for all the algorithm parameters.

Corollary 249. *Let $\Sigma_\infty = U_\infty^T \Lambda_\infty U_\infty$ be the spectral decomposition of the positive semi definite (PSD) matrix $\Sigma_\infty \in \mathbb{R}^{d_u \times d_u}$. Assume the notations in Fig.10.1. Let the covariate matrix $A_t := [w_{t-1}^T \dots w_{t-m}^T] \otimes \Lambda_\infty^{1/2} U_\infty$, where $\otimes$ denotes the Kronecker product. Let the bias vector $b_t := \Lambda_\infty^{1/2} U_\infty q_{\infty;h}^*(w_{t:t+h})$. For a sequence of DAP parameters $M_{1:n}$, let $\mathcal{TV}(M_{1:n}) := \sum_{t=2}^{n} \sum_{i=1}^{m} \|M_t^{[i]} - M_{t-1}^{[i]}\|_1$. For a sequence of matrices $(M^{[i]})_{i=1}^m$ define* $\texttt{flatten}((M^{[i]})_{i=1}^m)$ *as follows: Let $M_k^{[i]}$ be the k^{th} column of $M^{[i]}$.*

Let's define

$$z^k = \begin{bmatrix} M_1^k \\ \vdots \\ M_{d_x}^k \end{bmatrix} \in \mathbb{R}^{d_u d_x}, \tag{I.42}$$

and

$$\texttt{flatten}((M^{[i]})_{i=1}^m) := \begin{bmatrix} z^1 \\ \vdots \\ z^m \end{bmatrix} \in \mathbb{R}^{m d_u d_x}. \tag{I.43}$$

Let the decision set given to the ProDR.control (Fig.10.1) algorithm be the DAP space defined in Eq.(10.3)*. Let* $G = 2md_u d_x R\gamma\sqrt{d_x \wedge d_u}\|\Lambda^{1/2}U_\infty\|_1 + 2\frac{\|\Lambda^{-1/2}U_\infty B^T\|_2\|P_\infty\|_2\sqrt{d_u}}{1-\gamma}$*. Let the delay factor of ProDR.control.delayed (Fig.10.2) be* $\tau = h$ *as defined in Proposition 101. Choose* $\alpha = \sqrt{m\|\Sigma_\infty\|_{op}}$ *and* $L = 4G^2$*. Let* $\tilde{R}$ *in Theorem 104 be chosen as* $\tilde{R} = R\gamma\sqrt{d_u \wedge d_x}$*. Let* z_t *be the prediction at round* t *made by the ProDR.control.delayed algorithm. Let* $M_t^{alg} := \texttt{deflatten}(z_t)$*, where* **`deflatten`** *is the natural inverse operation of* **`flatten`** *defined above. Let* $\pi := (M_1, \ldots, M_n)$ *define a sequence of DAP policies. For a sequence of matrices* M*, define* $\|M\|_1 := \sum_{i=1}^m \|M^{[i]}\|_1$*. By playing a control* $u_t^{alg}(x_t) = \pi_t^{M_t^{alg}}(x_t)$ *according to Eq.*(10.2)*, we have that*

$$R_n(M_{1:n}) = \sum_{t=1}^n \ell(x_t^{alg}, u_t^{alg}) - \ell(x_t^{M_{1:n}}, u_t^{M_{1:n}}) = \tilde{O}\left(m^3 d^4 d_x^5 (d_u \wedge d_x)(n^{1/3}[\mathcal{TV}(M_{1:n})]^{2/3} \vee 1)\right), \tag{I.44}$$

where $M_{1:n}$ *is a sequence of DAP policies where each* $M_t \in \mathcal{M}$ *(eq.*(10.3)*). Further the algorithm ProDR.control.delayed also enjoys a strongly adaptive regret guarantee for any interval* $[a, b] \subseteq [n]$*:*

$$\sum_{t=a}^b \ell(x_t^{alg}, u_t^{alg}) - \ell(x_t^M, u_t^M) = \tilde{O}((md_u d_x)^{1.5} \log n), \tag{I.45}$$

for any fixed DAP policy $M \in \mathcal{M}$*.*

Proof. Define

$$X_t = [w_{t-1}^T \ldots w_{t-m}^T] \otimes I_{d_u}, \tag{I.46}$$

where $I_{d_u} \in \mathbb{R}^{d_u \times d_u}$ is the identity matrix and $\otimes$ denotes the Kronecker product. Clearly $X_t \in \mathbb{R}^{d_u \times md_u d_x}$.

With these definitions, it is easy to verify that

$$q^M(w_{t-1}) = X_t z. \tag{I.47}$$

Now we return back to losses $\hat{A}_t$ mentioned in Proposition 101. Let $\Sigma_\infty = U_\infty^T \Lambda_\infty U_\infty$ be the spectral decomposition of the positive semi definite (PSD) matrix $\Sigma_\infty \in \mathbb{R}^{d_u \times d_u}$. We have that

$$\hat{A}_t(M; w_{t+h}) = \|\Lambda_\infty^{1/2} U_\infty q^M(w_{t-1}) - \Lambda_\infty^{1/2} U_\infty q^*_{\infty;h}(w_{t:t+h})\|_2^2 \tag{I.48}$$
$$= \|\Lambda_\infty^{1/2} U_\infty X_t z - \Lambda_\infty^{1/2} U_\infty q^*_{\infty;h}(w_{t:t+h})\|_2^2. \tag{I.49}$$

Define

$$A_t := \Lambda_\infty^{1/2} U_\infty X_t \tag{I.50}$$
$$= [w_{t-1}^T \dots w_{t-m}^T] \otimes \Lambda_\infty^{1/2} U_\infty \tag{I.51}$$

Next, we proceed to compute a box that encloses all DAP policies of interest. We have for each $i \in [m]$,

$$\|z^i\|_\infty^2 \leq \|z^i\|_2^2 \tag{I.52}$$
$$= \|M^{[i]}\|_F^2 \tag{I.53}$$
$$\leq (d_u \wedge d_x)\|M^{[i]}\|_{op}^2 \tag{I.54}$$
$$\leq (d_u \wedge d_x) R^2 \gamma^2, \tag{I.55}$$

where the last line is due to the DAP policy set that we are interested in.

Thus the box $\mathcal{D}_\infty(R\gamma\sqrt{d_u \wedge d_x}) := \mathcal{D}_\infty(\tilde{R})$ encapsulates the DAP policy space that we are interested in.

We need to compute the parameters in Theorem 104. First, let's focus on computing G. We have for any $z_1, z_2 \in$

$$\|A_t(z_1 + z_2) - 2b_t\|_1 \leq 2\|A_t\|_1 m d_u d_x \tilde{R} + 2\|b_t\|_1, \tag{I.56}$$

where $b_t = \Lambda_\infty^{1/2} U_\infty q^*_{\infty;h}(w_{t:t+h})$.

We have

$$\|A_t\|_1 = \max_{i=1,\dots,m} \|w_{t-i}\|_\infty \|\Lambda^{1/2} U_\infty\|_1 \tag{I.57}$$
$$\leq \|\Lambda^{1/2} U_\infty\|_1, \tag{I.58}$$

as the disturbances obey $\|w_t\|_2 \leq 1$.

We have

$$\|b_t\|_2 \leq \sum_{i=t}^{t+h} \|\Lambda^{-1/2} U_\infty B^T (A_{cl,\infty})^{i-t} P_\infty w_i\|_2 \tag{I.59}$$

$$\leq \sum_{i=t}^{t+h} \|\Lambda^{-1/2} U_\infty B^T\|_2 \|(A_{cl,\infty})^{i-t}\|_2 \|P_\infty\|_2 \|w_i\|_2 \tag{I.60}$$

$$(a) \leq \|\Lambda^{-1/2} U_\infty B^T\|_2 \|P_\infty\|_2 \sum_{i=1}^{h} \gamma^{i-1} \tag{I.61}$$

$$\leq \|\Lambda^{-1/2} U_\infty B^T\|_2 \|P_\infty\|_2 \cdot \frac{1}{1-\gamma}, \tag{I.62}$$

where in line (a) we used the strong stability criterion and the fact that $\|w_t\|_2 \leq 1$. Thus we have

$$\|b_t\|_1 \leq \sqrt{d_u} \|b_t\|_2 \tag{I.63}$$

$$\leq \frac{\|\Lambda^{-1/2} U_\infty B^T\|_2 \|P_\infty\|_2 \sqrt{d_u}}{1-\gamma}. \tag{I.64}$$

Putting together Eq.(I.56).(I.58) and (I.64) we arrive at

$$\|A_t(z_1 + z_2) - 2b_t\|_1 \leq 2 m d_u d_x R \gamma \sqrt{d_x \wedge d_u} \|\Lambda^{1/2} U_\infty\|_1 + 2 \frac{\|\Lambda^{-1/2} U_\infty B^T\|_2 \|P_\infty\|_2 \sqrt{d_u}}{1-\gamma} \tag{I.65}$$

$$:= G \tag{I.66}$$

Next we proceed to calculate α in Theorem 104. Denote by U_j the j^{th} column of the matrix U_∞. The squared norm of the i^{th} row of the covariate matrix A_t is given by

$$\sum_{k=1}^{m} \|w_{t-k}\|_2^2 \sum_{j=1}^{d_u} \lambda_j u_j^2[i] \leq \|\Sigma_\infty\|_{op} \sum_{k=1}^{m} \sum_{j=1}^{d_u} u_j^2[i] \tag{I.67}$$

$$= m \|\Sigma_\infty\|_{op}, \tag{I.68}$$

where we used the fact the matrix U_∞ is orthogonal. Thus we choose

$$\alpha = \sqrt{m \|\Sigma_\infty\|_{op}}. \tag{I.69}$$

By similar arguments used to reach Eq.(I.66), we choose

$$L = 4G^2 \tag{I.70}$$

For a sequence of policies $M_1, \ldots, M_n$, observe that $\sum_{t=2}^{n} \|\texttt{flatten}(M_t) - \texttt{flatten}(M_{t-1})\|_1 \leq d_x \sum_{t=2}^{n} \|M_t - M_{t-1}\|_1$. The last relation expresses the dynamic regret incurred by ProDR.control.delayed in terms of total variation of $\texttt{flatten}(M_t)$ to be bounded by total variation of the matrices themselves.

Putting all the constants together and applying Theorem 108 and Theorem 104 yields the Corollary.

□

Theorem 110. *There exists an LQR system, a choice of the perturbations w_t and a DAP policy class such that:*

$$\sup_{M_{1:n} \text{ with } \mathcal{TV}(M_{1:n}) \leq C_n} \mathbb{E}[R(M_{1:n})] = \Omega(n^{1/3} C_n^{2/3} \vee 1), \tag{10.19}$$

where the expectation is taken wrt randomness in the strategies of the agent and adversary.

Proof. Consider a system with matrices $A = 0 \in \mathbb{R}^{2\times 2}$, $B = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$, $R_x = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $R_u = 0 \in \mathbb{R}^{2\times 2}$. In this setting $K_\infty = 0$ as per Eq.(10.6). We consider DAP polices (see Definition 100) with $m = 1$. Let the starting state be $x_1 = 0 \in \mathbb{R}^{2\times 2}$.

Let $y_t = \pm 1$ with probability half each. Let $w_t = [y_t, 1]^T$. For a policy that chooses a control signal u_t at time t, its next state is given by $x_{t+1} = w_t - u_t$ and $\ell_{t+1}(x_{t+1}, u_{t+1}) = (u_t[1] - y_t)^2$. Hence for any algorithm, the loss is given by:

$$\sum_{t=1}^{n} \ell_t(x_t, u_t) = \sum_{t=1}^{n-1} (u_t^{\text{alg}}[1] - y_t)^2. \tag{I.71}$$

Divide the time horizon into bins of width W. Let the number of bins be $M := n/W$. We assume that n/W is an integer for simplicity. Let the i^{th} be denoted by $[s_i, e_i]$ for $i \in [M]$. Define

$$a_i := \frac{1}{W} \sum_{t=s_i}^{e_i} y_t. \tag{I.72}$$

We will uniformly use the same DAP policy within a bin i as the comparator. This policy will be parameterized by the matrix $M_i := \begin{bmatrix} 0 & -a_i \\ 0 & 0 \end{bmatrix}$

By Hoeffding's inequality and a union bound across all M bins, we arrive at

$$a_i \in \left[-\sqrt{\frac{\log(nM/\delta)}{2W}}, \sqrt{\frac{\log(nM/\delta)}{2W}} \right], \tag{I.73}$$

with probability at-least $1 - \delta$. We will call this high probability event as $\mathcal{E}$. Due to symmetry we have that $P(y_t = 1|\mathcal{E}) = 1/2$. So under the event $\mathcal{E}$, the Bayes optimal online prediction of any algorithm as per Eq.(I.71) will be to set $u = [0, 0]^T$. So within a bin we have that

$$\sum_{t=s_i}^{s_e} E[\ell_t(x_t, u_t)|\mathcal{E}] \geq W. \tag{I.74}$$

Now we need to upper bound the cumulative loss of the comparator within a bin. Since the policy within a bin is parameterized by M_i, we have that $u_t = -M_t w_{t-1} = [a_i, 0]^T$ for all $t \in [s_i, e_i]$.

So we have:

$$E[(y_t - u_t)^2|\mathcal{E}] = \frac{E[(y_t - u_t)^2] - E[(y_t - u_t)^2|\mathcal{E}^c]P(\mathcal{E}^c)}{P(\mathcal{E})} \tag{I.75}$$

$$\leq \frac{E[(y_t - u_t)^2]}{1-\delta}, \tag{I.76}$$

where $\mathcal{E}^c$ denotes complement of event $\mathcal{E}$.

By bias variance decomposition, we have that

$$E[(y_t - u_t)^2] = 1 - 1/W. \tag{I.77}$$

So the overall regret is lower bounded by

$$\sum_{i=1}^{M} \sum_{t=s_i}^{e_i} E[(y_t - u_t^{\text{alg}}[1])^2|\mathcal{E}] - E[(y_t - a_i)^2|\mathcal{E}] \geq \sum_{i=1}^{M} W(1 - \frac{1}{1-\delta}) + \frac{1}{1-\delta} \tag{I.78}$$

$$\geq M/(1-\delta) - W\delta/(1-\delta) \tag{I.79}$$

$$\geq M/2, \tag{I.80}$$

where the last line is obtained by setting $\delta = 1/n^2$

Under the event $\mathcal{E}$ with $\delta = 1/n^2$, the total variation (TV) of the sequence $a_{1:n}$ is given by:

$$\text{TV}(a_{1:n}) \leq \frac{n\sqrt{2\log(n^4)}}{W^{3/2}}. \tag{I.81}$$

Now setting $W = \frac{n^{2/3}(8\log n)^{1/3}}{C_n^{2/3}}$ we obtain $\text{TV}(a_{1:n}) \leq C_n$ with probability at-least $1 - 1/n^2$.

For the sake of brevity let's denote $R(M_{1:n})$ (Eq.(10.4)) by R_n.

Continuing from Eq.(I.80), we obtain that

$$E[R_n|\mathcal{E}] := \sum_{i=1}^{M}\sum_{t=s_i}^{e_i} E[(y_t - u_t^{\text{alg}}[1])^2|\mathcal{E}] - E[(y_t - a_i)^2|\mathcal{E}] \tag{I.82}$$

$$\geq \frac{n^{1/3}C_n^{2/3}}{2(8\log n)^{1/3}}, \tag{I.83}$$

where the event $\mathcal{E}$ occurs with probability at-least $1 - 1/n^2$.

Now consider the event $\mathcal{E}^c$. For the purpose of obtaining a lower bound we can restrict our attention to comparators $a_{1:n}$ such that $|a_i| \leq 1$ for all $i \in [n]$ and $\text{TV}(a_{1:n}) \leq C_n$. Using the DAP policy given by $M_i := \left[\begin{smallmatrix} 0 & -a_i \\ 0 & 0 \end{smallmatrix}\right]$ as comparators, we have that under the event $\mathcal{E}^c$

$$R_n \geq -\sum_{t=1}^{n}(y_t - a_t)^2 \tag{I.84}$$

$$\geq -4n \tag{I.85}$$

So overall we have that

$$E[R_n] \geq E[R_n|\mathcal{E}]p(\mathcal{E}) + E[R_n|\mathcal{E}^c]p(\mathcal{E}^c) \tag{I.86}$$

$$\geq \Omega(n^{1/3}C_n^{2/3})(1 - 1/n^2) - 4n \cdot (1/n^2) \tag{I.87}$$

$$= \Omega(n^{1/3}C_n^{2/3}). \tag{I.88}$$

When $C_n \leq 1/\sqrt{n}$, the static regret bound of $\Omega(\log n)$ (see Theorem 11.9 in [40]). This completes the proof of the theorem.

□

Connections to online non-parametric regression framework of [29]. In the work of [29], they study the following online regression framework (simplified here without affecting the information-theoretic rates):

- At each round t, learner plays a decision $x_t \in \mathbb{R}$.
- Nature reveals a label y_t such that $|y_t| \leq 1$.
- Learner suffers loss $(y_t - x_t)^2$.

One is interested in finding the min-max rate of regret against a non-parametric sequence class. We define the space of total variation (TV) bounded sequences as:

$$\text{TV}(C_n) := \{\theta_{1:n}|\text{TV}(\theta_{1:n}) \leq C_n\}. \tag{I.89}$$

Translated into the setup of [29], one can aim to control the regret against $\mathrm{TV}(C_n)$ which is:

$$R_n := \sum_{t=1}^{n} (y_t - x_t)^2 - \inf_{\theta_{1:n} \in \mathrm{TV}(C_n)} \sum_{t=1}^{n} (y_t - \theta_t)^2. \tag{I.90}$$

The TV class is known to be sandwiched between two Besov spaces having the same minimax rate (see for eg. [142]). So the results of [29] based on characterizing the sequential Rademacher complexity of the Besov class leads to $O(n^{1/3})$ as the minimax rate of R_n wrt n. The rate wrt C_n was not provided in their work. However, we remark that they establish an $O(n^{1/3})$ upper bound also via non-constructive arguments.

In contrast, the lower bound we provided in the proof of Theorem 110 is for $\sum_{t=1}^{n} E[(y_t - u_t^{\mathrm{alg}}[1])^2 - (y_t - a_t)^2 | \mathcal{E}]$ (Eq.(I.83)) where $\mathrm{TV}(a_{1:n}) \leq C_n$ under the high probability event $\mathcal{E}$ trivially lower bounds R_n in Eq.(I.90) with high probability.

Bibliography

[1] S. van de Geer, *Estimating a regression function*, *Annals of Statistics* **18** (1990), no. 2 907—924.

[2] D. L. Donoho and I. M. Johnstone, *Minimax estimation via wavelet shrinkage*, *The annals of Statistics* **26** (1998), no. 3 879–921.

[3] R. J. Tibshirani, *Adaptive piecewise polynomial estimation via trend filtering*, *Annals of Statistics* **42** (2014), no. 1 285–323.

[4] S.-J. Kim, K. Koh, S. Boyd, and D. Gorinevsky, ℓ_1 *trend filtering*, *SIAM Review* **51** (2009), no. 2 339–360.

[5] M. Zinkevich, *Online convex programming and generalized infinitesimal gradient ascent*, in *International Conference on Machine Learning (ICML-03)*, pp. 928–936, 2003.

[6] E. A. Nadaraya, *On estimating regression*, *Theory of Probability & Its Applications* **9** (1964), no. 1 141–142.

[7] C. De Boor, C. De Boor, E.-U. Mathématicien, C. De Boor, and C. De Boor, *A practical guide to splines*, vol. 27. Springer-Verlag New York, 1978.

[8] G. Wahba, *Spline models for observational data*, vol. 59. Siam, 1990.

[9] S. Mallat, *A wavelet tour of signal processing*. Elsevier, 1999.

[10] B. Scholkopf and A. J. Smola, *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT press, 2001.

[11] C. E. Rasmussen and C. K. Williams, *Gaussian processes for machine learning*. MIT Press, 2006.

[12] G. Steidl, S. Didas, and J. Neumann, *Splines in higher order TV regularization*, *International Journal of Computer Vision* **70** (2006), no. 3 214–255.

[13] Y.-X. Wang, A. Smola, and R. Tibshirani, *The falling factorial basis and its statistical applications*, in *International Conference on Machine Learning (ICML-14)*, pp. 730–738, 2014.

[14] E. Mammen and S. van de Geer, *Locally adaptive regression splines*, *Annals of Statistics* **25** (1997), no. 1 387–413.

[15] O. Besbes, Y. Gur, and A. Zeevi, *Non-stationary stochastic optimization*, *Operations research* **63** (2015), no. 5 1227–1244.

[16] A. Jadbabaie, A. Rakhlin, S. Shahrampour, and K. Sridharan, *Online optimization: Competing with dynamic comparators*, in *Artificial Intelligence and Statistics*, pp. 398–406, 2015.

[17] A. B. Tsybakov, *Introduction to Nonparametric Estimation.* Springer Publishing Company, Incorporated, 1st ed., 2008.

[18] I. M. Johnstone, *Gaussian estimation: Sequence and wavelet models.* 2017.

[19] G. E. Box and G. M. Jenkins, *Time series analysis: forecasting and control.* John Wiley & Sons, 1970.

[20] D. Donoho, R. Liu, and B. MacGibbon, *Minimax risk over hyperrectangles, and implications*, *Annals of Statistics* **18** (1990), no. 3 1416–1437.

[21] E. Hall and R. Willett, *Dynamical models and tracking regret in online convex programming*, in *International Conference on Machine Learning (ICML-13)*, pp. 579–587, 2013.

[22] X. Chen, Y. Wang, and Y.-X. Wang, *Non-stationary stochastic optimization under lp, q-variation measures*, .

[23] E. Hazan and C. Seshadhri, *Adaptive algorithms for online decision problems*, in *Electronic colloquium on computational complexity (ECCC)*, vol. 14, 2007.

[24] A. Daniely, A. Gonen, and S. Shalev-Shwartz, *Strongly adaptive online learning*, in *International Conference on Machine Learning*, pp. 1405–1411, 2015.

[25] T. Yang, L. Zhang, R. Jin, and J. Yi, *Tracking slowly moving clairvoyant: optimal dynamic regret of online learning with true and noisy gradient*, in *International Conference on Machine Learning (ICML-16)*, pp. 449–457, 2016.

[26] L. Zhang, S. Lu, and Z.-H. Zhou, *Adaptive online learning in dynamic environments*, in *Advances in Neural Information Processing Systems (NeurIPS-18)*, pp. 1323–1333, 2018.

[27] L. Zhang, T. Yang, Z.-H. Zhou, *et. al.*, *Dynamic regret of strongly adaptive methods*, in *International Conference on Machine Learning (ICML-18)*, pp. 5877–5886, 2018.

[28] N. Chen, G. Goel, and A. Wierman, *Smoothed online convex optimization in high dimensions via online balanced descent*, in *Conference on Learning Theory (COLT-18)*, 2018.

[29] A. Rakhlin and K. Sridharan, *Online non-parametric regression*, in *Conference on Learning Theory*, pp. 1232–1264, 2014.

[30] P. Gaillard and S. Gerchinovitz, *A chaining algorithm for online nonparametric regression*, in *Conference on Learning Theory*, pp. 764–796, 2015.

[31] D. L. Donoho, *De-noising by soft-thresholding*, *IEEE transactions on information theory* **41** (1995), no. 3 613–627.

[32] L. Wasserman, *All of Nonparametric Statistics*. Springer, New York, 2006.

[33] V. Sadhanala, Y.-X. Wang, and R. Tibshirani, *Total variation classes beyond 1d: Minimax rates, and the limitations of linear smoothers*, *Advances in Neural Information Processing Systems (NIPS-16)* (2016).

[34] R. Tibshirani, *Nonparametric Regression: Statistical Machine Learning, Spring 2015*, 2015. URL: `http://www.stat.cmu.edu/~larry/=sml/nonpar.pdf`. Last visited on 2019/04/29.

[35] C. de Boor, *A Practical Guide to Splines*. Springer, New York, 1978.

[36] R. J. Hodrick and E. C. Prescott, *Postwar us business cycles: an empirical investigation*, *Journal of Money, credit, and Banking* (1997) 1–16.

[37] D. Baby and Y.-X. Wang, *Online forecasting of total-variation-bounded sequences*, in *Neural Information Processing Systems (NeurIPS)*, 2019.

[38] Y.-X. Wang, J. Sharpnack, A. Smola, and R. J. Tibshirani, *Trend filtering on graphs*, *Journal of Machine Learning Research* **17** (2016), no. 105 1–41.

[39] A. Guntuboyina, D. Lieu, S. Chatterjee, and B. Sen, *Adaptive risk bounds in univariate total variation denoising and trend filtering*, arXiv:1702.0511.

[40] N. Cesa-Bianchi and G. Lugosi, *Prediction, Learning, and Games*. Cambridge University Press, New York, NY, USA, 2006.

[41] A. Cohen, I. Daubechies, and P. Vial, *Wavelets on the interval and fast wavelet transforms*, *Applied and Computational Harmonic Analysis* **1** (1993), no. 1 54 – 81.

[42] V. Vovk, *Competitive on-line statistics*, *International Statistical Review* **69** (2001), no. 2 213–248.

[43] O. Dekel, S. Shalev-Shwartz, and Y. Singer, *Smooth ϵ-insensitive regression by loss symmetrization*, *J. Mach. Learn. Res.* (2005).

[44] L. Gyorfi, M. Kohler, A. Krzyzak, and H. Walk, *A Distribution-Free Theory of Nonparametric Regression.* Springer, New York, 2002.

[45] L. Schumaker, *Spline functions: basic theory.* Cambridge University Press, 2007.

[46] L. Rudin, S. Osher, and E. Faterni, *Nonlinear total variation based noise removal algorithms*, *Physica D: Nonlinear Phenomena* **60** (1992) 259–268.

[47] R. Tibshirani, M. Saunders, S. Rosset, J. Zhu, and K. Knight, *Sparsity and smoothness via the fused lasso*, *Journal of the Royal Statistical Society: Series B* **67** (2005), no. 1 91–108.

[48] H. Hoefling, *A path algorithm for the fused lasso signal approximator*, *Journal of Computational and Graphical Statistics* **19** (2010), no. 4 984–1006.

[49] K. Akiyama, K. Bouman, and D. Woody, *First m87 event horizon telescope results. i. the shadow of the supermassive black hole*, *Astrophysical Journal Letters* (2019).

[50] A. Chambolle, V. Caselles, D. Cremers, M. Novaga, and T. Pock, *An introduction to total variation for image analysis*, *Theoretical foundations and numerical methods for sparse recovery* **9** (2010), no. 263-340 227.

[51] E. Mammen, *Nonparametric regression under qualitative smoothness assumptions*, *Annals of Statistics* **19** (1991), no. 2 741—759.

[52] D. Adamskiy, W. M. Koolen, A. Chernov, and V. Vovk, *A closer look at adaptive regret*, *Journal of Machine Learning Research* (2016).

[53] W. Kotłowski, W. M. Koolen, and A. Malek, *Online isotonic regression*, in *Annual Conference on Learning Theory (COLT-16)*, vol. 49, pp. 1165–1189, PMLR, 2016.

[54] D. Baby and Y.-X. Wang, *Adaptive online estimation of piecewise polynomial trends*, *Neural Information Processing Systems (NeurIPS)* (2020).

[55] R. J. Tibshirani and J. Taylor, *Degrees of freedom in lasso problems*, *Annals of Statistics* **40** (2012), no. 2 1198–1232.

[56] D. Donoho and I. Johnstone, *Ideal spatial adaptation by wavelet shrinkage*, *Biometrika* **81** (1994), no. 3 425–455.

[57] "Centers for disease control and prevention (cdc)." `https://www.cdc.gov/`.

[58] X. Jin, Y.-X. Wang, and X. Yan, *Inter-series attention model for covid-19 forecasting*, *SIAM International Conference on Data Mining (to appear)* (2021).

[59] E. Hazan, *Introduction to online convex optimization*, *Foundations and Trends® in Optimization* **2** (2016), no. 3-4 157–325.

[60] V. Vovk, *Competitive on-line linear regression*, in *NIPS*, 1997.

[61] K. S. Azoury and M. K. Warmuth, *Relative loss bounds for on-line density estimation with the exponential family of distributions*, *Machine Learning* **43** (2004) 211–246.

[62] J. Yuan and A. Lamperski, *Trading-off static and dynamic regret in online least-squares and beyond*, in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, pp. 6712–6719, 2020.

[63] F. Ortelli and S. van de Geer, *Prediction bounds for (higher order) total variation regularized least squares*, *arXiv preprint arXiv:1904.10871* (2019).

[64] S. Bubeck, *Convex optimization: Algorithms and complexity*, *Found. Trends Mach. Learn.* **8** (2015) 231–357.

[65] D. Baby and Y.-X. Wang, *Optimal dynamic regret in exp-concave online learning*, in *Conference on Learning Theory*, pp. 359–409, PMLR, 2021.

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

[67] G. Goel and A. Wierman, *An online algorithm for smoothed regression and lqr control*, in *Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics*, 2019.

[68] P. Zhao, Y. Zhang, L. Zhang, and Z.-H. Zhou, *Dynamic regret of convex and smooth functions*, *NeurIPS* (2020).

[69] A. Cutkosky, *Parameter-free, dynamic, and strongly-adaptive online learning*, in *Proceedings of the 37th International Conference on Machine Learning*, 2020.

[70] P. Zhao and L. Zhang, *Improved analysis for dynamic regret of strongly convex and smooth functions*, *L4DC* (2021).

[71] D. Baby, X. Zhao, and Y.-X. Wang, *An optimal reduction of tv-denoising to adaptive online learning*, *AISTATS* (2021).

[72] P. Zhao, Y.-X. Wang, and Z.-H. Zhou, *Non-stationary online learning with memory and non-stochastic control*, *AISTATS* (2022).

[73] T.-J. Chang and S. Shahrampour, *On online optimization: Dynamic regret analysis of strongly convex and smooth problems*, *Proceedings of the AAAI Conference on Artificial Intelligence* (2021).

[74] D. Baby, H. Hasson, and Y. Wang, *Dynamic regret for strongly adaptive methods and optimality of online krr*, *ArXiv* **abs/2111.11550** (2021).

[75] K.-S. Jun, F. Orabona, S. Wright, and R. Willett, *Improved Strongly Adaptive Online Learning using Coin Betting*, in *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, 2017.

[76] L. Zhang, G. Wang, W.-W. Tu, and Z.-H. Zhou, *Dual adaptivity: A universal algorithm for minimizing the adaptive regret of convex functions*, *NeurIPS* (2021).

[77] E. Hazan, A. Agarwal, and S. Kale, *Logarithmic regret algorithms for online convex optimization*, .

[78] W. M. Koolen, A. Malek, P. L. Bartlett, and Y. Abbasi, *Minimax time series prediction*, in *Advances in Neural Information Processing Systems (NIPS'15)*, pp. 2557–2565. 2015.

[79] A. Cutkosky and F. Orabona, *Black-box reductions for parameter-free online learning in banach spaces*, in *COLT*, 2018.

[80] R. J. Tibshirani, *Adaptive piecewise polynomial estimation via trend filtering*, *Annals of Statistics* **42** (2014), no. 1 285–323.

[81] A. Cutkosky, *Parameter-free, dynamic, and strongly-adaptive online learning*, in *ICML*, 2020.

[82] H. Luo, A. Agarwal, N. Cesa-Bianchi, and J. Langford, *Efficient second order online learning by sketching*, in *NIPS*, 2016.

[83] D. Baby and Y.-X. Wang, *Second order path variationals in non-stationary online learning*, *AISTATS* (2023).

[84] Y. Nesterov, *Introductory lectures on convex optimization - a basic course*, in *Applied Optimization*, 2004.

[85] J. Quinonero-Candela, M. Sugiyama, A. Schwaighofer, and N. D. Lawrence, *Dataset shift in machine learning*. Mit Press, 2008.

[86] A. Torralba and A. A. Efros, *Unbiased look at dataset bias*, in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pp. 1521–1528, IEEE, 2011.

[87] P. W. Koh, S. Sagawa, H. Marklund, S. M. Xie, M. Zhang, A. Balsubramani, W. Hu, M. Yasunaga, R. L. Phillips, I. Gao, T. Lee, E. David, I. Stavness, W. Guo, B. A. Earnshaw, I. S. Haque, S. Beery, J. Leskovec, A. Kundaje, E. Pierson, S. Levine, C. Finn, and P. Liang, *WILDS: A benchmark of in-the-wild distribution shifts*, in *International Conference on Machine Learning (ICML)*, 2021.

[88] S. Garg, S. Balakrishnan, Z. Lipton, B. Neyshabur, and H. Sedghi, *Leveraging unlabeled data to predict out-of-distribution performance*, in *International Conference on Learning Representations (ICLR)*, 2022.

[89] H. Shimodaira, *Improving predictive inference under covariate shift by weighting the log-likelihood function*, *Journal of statistical planning and inference* **90** (2000), no. 2 227–244.

[90] M. Saerens, P. Latinne, and C. Decaestecker, *Adjusting the Outputs of a Classifier to New a Priori Probabilities: A Simple Procedure*, *Neural Computation* (2002).

[91] Z. Lipton, Y.-X. Wang, and A. Smola, *Detecting and correcting for label shift with black box predictors*, in *International Conference on Machine Learning*, 2018.

[92] H. Zhou, S. Balakrishnan, and Z. C. Lipton, *Domain adaptation under missingness shift*, in *Artificial Intelligence and Statistics (AISTATS)*, 2022.

[93] A. Alexandari, A. Kundaje, and A. Shrikumar, *Adapting to label shift with bias-corrected calibration*, in *International Conference on Machine Learning (ICML)*, 2021.

[94] K. Azizzadenesheli, A. Liu, F. Yang, and A. Anandkumar, *Regularized learning for domain adaptation under label shifts*, in *International Conference on Learning Representations (ICLR)*, 2019.

[95] S. Garg, Y. Wu, S. Balakrishnan, and Z. Lipton, *A unified view of label shift estimation*, in *Advances in Neural Information Processing Systems (NeurIPS)*, 2020.

[96] R. Wu, C. Guo, Y. Su, and K. Q. Weinberger, *Online adaptation to label distribution shift*, in *Advances in Neural Information Processing Systems*, 2021.

[97] Y. Bai, Y.-J. Zhang, P. Zhao, M. Sugiyama, and Z.-H. Zhou, *Adapting to online label shift with provable guarantees*, in *Advances in Neural Information Processing Systems*, 2022.

[98] F. Orabona, *A modern introduction to online learning*, 2019.

[99] D. Baby, S. Garg, T.-C. Yen, S. Balakrishnan, Z. C. Lipton, and Y.-X. Wang, *Online label shift: Optimal dynamic regret meets practical algorithms*, in *Advances in Neural Information Processing Systems (NeurIPS)*, 2023.

[100] C. Huang, Y. Li, C. C. Loy, and X. Tang, *Learning deep representation for imbalanced classification*, in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 5375–5384, 2016.

[101] Y.-X. Wang, D. Ramanan, and M. Hebert, *Learning to model the tail*, *Advances in neural information processing systems* **30** (2017).

[102] Y. Cui, M. Jia, T.-Y. Lin, Y. Song, and S. Belongie, *Class-balanced loss based on effective number of samples*, in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 9268–9277, 2019.

[103] O. Bousquet, S. Boucheron, and G. Lugosi, *Introduction to statistical learning theory.*, in *Advanced Lectures on Machine Learning*, Lecture Notes in Computer Science, Springer, 2003.

[104] R. Vershynin, *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge University Press, 2018.

[105] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, *Gradient-Based Learning Applied to Document Recognition*, *Proceedings of the IEEE* **86** (1998).

[106] A. Krizhevsky and G. Hinton, *Learning Multiple Layers of Features from Tiny Images*, tech. rep., Citeseer, 2009.

[107] H. Xiao, K. Rasul, and R. Vollgraf, *Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms*, *arXiv preprint arXiv:1708.07747* (2017).

[108] P. Helber, B. Bischke, A. Dengel, and D. Borth, *Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification*, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **12** (2019), no. 7 2217–2226.

[109] C. B. Clement, M. Bierbaum, K. P. O'Keeffe, and A. A. Alemi, *On the use of arxiv as a dataset*, *arXiv preprint arXiv:1905.00075* (2019).

[110] H. Gjoreski, M. Ciliberto, L. Wang, F. J. O. Morales, S. Mekki, S. Valentin, and D. Roggen, *The university of sussex-huawei locomotion and transportation dataset for multimodal analytics with mobile devices*, *IEEE Access* **6** (2018) 42592–42604.

[111] L. Wang, H. Gjoreski, M. Ciliberto, S. Mekki, S. Valentin, and D. Roggen, *Enabling reproducible research in sensor-based transportation mode recognition with the sussex-huawei dataset*, *IEEE Access* **7** (2019) 10870–10891.

[112] K. He, X. Zhang, S. Ren, and J. Sun, *Deep Residual Learning for Image Recognition*, in *Computer Vision and Pattern Recognition (CVPR)*, 2016.

[113] V. Sanh, L. Debut, J. Chaumond, and T. Wolf, *Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter*, *arXiv preprint arXiv:1910.01108* (2019).

[114] T. Wolf, L. Debut, V. Sanh, J. Chaumond, C. Delangue, A. Moi, P. Cistac, T. Rault, R. Louf, M. Funtowicz, *et. al.*, *Huggingface's transformers: State-of-the-art natural language processing*, *arXiv preprint arXiv:1910.03771* (2019).

[115] C. Guo, G. Pleiss, Y. Sun, and K. Q. Weinberger, *On calibration of modern neural networks*, in *International conference on machine learning*, pp. 1321–1330, PMLR, 2017.

[116] M. C. Cohen, I. Lobel, and R. Paes Leme, *Feature-based dynamic pricing*, *Management Science* **66** (2020), no. 11 4921–4943.

[117] A. Javanmard and H. Nazerzadeh, *Dynamic pricing in high-dimensions*, *The Journal of Machine Learning Research* **20** (2019), no. 1 315–363.

[118] J. Xu and Y.-X. Wang, *Logarithmic regret in feature-based dynamic pricing*, *Advances in Neural Information Processing Systems* **34** (2021) 13898–13910.

[119] Y. Luo, W. W. Sun, *et. al.*, *Distribution-free contextual dynamic pricing*, *arXiv preprint arXiv:2109.07340* (2021).

[120] D. Baby and Y.-X. Wang, *Optimal dynamic regret in proper online learning with strongly convex losses and beyond*, *AISTATS* (2022).

[121] A. Cutkosky and F. Orabona, *Black-box reductions for parameter-free online learning in banach spaces*, in *Thirty-First Annual Conference on Learning Theory (COLT-18)*, pp. 1493–1529, PMLR, 2018.

[122] D. Baby, J. Xu, and Y.-X. Wang, *Non-stationary contextual pricing with safety constraints*, in *TMLR*, 2023.

[123] A. Prékopa, *On logarithmic concave measures and functions*, *Acta Scientiarum Mathematicarum* **34** (1973) 335–343.

[124] M. Bagnoli and T. Bergstrom, *Log-concave probability and its applications*, in *Rationality and Equilibrium: A Symposium in Honor of Marcel K. Richter*, pp. 217–241, Springer, 2006.

[125] N. Agarwal, B. Bullins, E. Hazan, S. Kakade, and K. Singh, *Online control with adversarial disturbances*, in *Proceedings of the 36th International Conference on Machine Learning*, 2019.

[126] D. Foster and M. Simchowitz, *Logarithmic regret for adversarial online control*, in *International Conference on Machine Learning*, pp. 3211–3221, PMLR, 2020.

[127] S. Levine, C. Finn, T. Darrell, and P. Abbeel, *End-to-end training of deep visuomotor policies*, .

[128] A. Cohen, A. Hassidim, T. Koren, N. Lazic, Y. Mansour, and K. Talwar, *Online linear quadratic control*, in *ICML*, 2018.

[129] D. Baby and Y.-X. Wang, *Optimal dynamic regret in LQR control*, in *Advances in Neural Information Processing Systems*, 2022.

[130] D. P. Bertsekas, *Dynamic Programming and Optimal Control.* Athena Scientific, 3rd ed., 2005.

[131] P. Joulani, A. György, and C. Szepesvari, *Online learning under delayed feedback*, in *ICML*, 2013.

[132] L. Birge and P. Massart, *Gaussian model selection*, *Journal of the European Mathematical Society* **3** (2001), no. 3 203–268.

[133] P. Bickel *et. al.*, *Minimax estimation of the mean of a normal distribution when the parameter space is restricted*, *The Annals of Statistics* **9** (1981), no. 6 1301–1309.

[134] B. M. Dingle, *Symbolic determinants:calculating the degree*, *Technical Report, Texas A&M University* (2005).

[135] A. Beygelzimer, J. Langford, L. Li, L. Reyzin, and R. Schapire, *Contextual bandit algorithms with supervised learning guarantees*, in *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, 2011.

[136] D. Haussler, J. Kivinen, and M. K. Warmuth, *Sequential prediction of individual sequences under general loss functions*, *IEEE Trans. Inf. Theor.* (1998).

[137] G. Lee, R. Gommers, F. Waselewski, K. Wohlfahrt, and A. O. Leary, *Pywavelets: A python package for wavelet analysis*, *Journal of Open Source Software* (2019).

[138] A. Meurer, C. P. Smith, M. Paprocki, O. Čertík, S. B. Kirpichev, M. Rocklin, A. Kumar, S. Ivanov, J. K. Moore, S. Singh, T. Rathnayake, S. Vig, B. E. Granger, R. P. Muller, F. Bonazzi, H. Gupta, S. Vats, F. Johansson, F. Pedregosa, M. J. Curry, A. R. Terrel, v. Roučka, A. Saboo, I. Fernando, S. Kulal, R. Cimrman, and A. Scopatz, *Sympy: symbolic computing in python*, *PeerJ Computer Science* (Jan., 2017).

[139] A. Raj, P. Gaillard, and C. Saad, *Non-stationary online regression*, 2020.

[140] Y.-X. Wang, J. Sharpnack, A. Smola, and R. Tibshirani, *Trend filtering on graphs*, in *AISTATS'15*, pp. 1042–1050, 2015.

[141] M. Mohri, A. Rostamizadeh, and A. Talwalkar, *Foundations of Machine Learning.* The MIT Press, 2012.

[142] R. A. DeVore and G. G. Lorentz, *Constructive approximation*, in *Grundlehren der mathematischen Wissenschaften*, 1993.