Online Convex Optimization with Predictions

Static and Dynamic Environments

Pedro Zattoni Scroccaro

Master of Science Thesis

Delft University of Technology

Delft Center for Systems and Control

Online Convex Optimization with Predictions

Static and Dynamic Environments

Master of Science Thesis

For the degree of Master of Science in Systems and Control at Delft University of Technology

Pedro Zattoni Scroccaro

September 7, 2020

Faculty of Mechanical, Maritime and Materials Engineering (3mE) · Delft University of Technology

Delft University of Technology
Department of
Delft Center for Systems and Control (dcsc)

The undersigned hereby certify that they have read and recommend to the Faculty of Mechanical, Maritime and Materials Engineering (3mE) for acceptance a thesis entitled

Online Convex Optimization with Predictions

by

Pedro Zattoni Scroccaro

in partial fulfillment of the requirements for the degree of

Master of Science Systems and Control

Dated: September 7, 2020

Supervisor(s):

dr. P. Mohajerin Esfahani

Reader(s):

dr. S. Grammatico

dr. B. Atasoy

dr. A. S. Kolarijani

Abstract

In this thesis, we study Online Convex Optimization algorithms that exploit predictive and/or dynamical information about a problem instance. These features are inspired by recent developments in the Online Mirror Decent literature. When the Player's performance is compared with the best fixed decision in hindsight, we show that it is possible to achieve constant regret bounds under perfect gradient predictions and optimal minimax bounds in the worst-case, generalizing previous results from the literature. For dynamic environments, we propose a new algorithm, and show that it achieves dynamic regret bounds that exploit both gradient predictions and knowledge about the dynamics of the action sequence that the Player's performance is being compared with. We present results for both convex and strongly convex costs. Finally, we provide numerical experiments that corroborate our theoretical results.

Table of Contents

Acknowledgements

This document is the end result of my two years as a MSc student at the Delft Center for Systems and Control (DCSC). During my stay here, I was able to learn more than I could hope for in such a short period of time.

Of course, many people were involved in my learning experience. First of all, I would like to thank my supervisor dr. Peyman Mohajerin Esfahani. His kindness, excitement with research and extremely hard working mindset were an endless source of inspiration. The same can be said about dr. Arman Sharifi Kolarijani, who also spent many hours patiently teaching me how to properly transform an idea into an research paper. Moreover, I would like to thank the other members of my thesis committee: dr. Sergio Grammatico and dr. Bilge Atasoy.

Last but not least, I would like to express my gratitude towards my family and friends, for their much needed "non-technical" support.

Delft, University of Technology — Pedro Zattoni Scroccaro
September 7, 2020

Chapter 1

Introduction

1-1 Online Convex Optimization

In real-time decision making problems, actions must be taken and updated at the same time that new information about the problem is received. In this thesis, we will focus on a specific framework to tackle suck problems, called *Online Convex Optimization* (OCO). Abstractly, the standard OCO framework can be described as a game between a Player and Nature, played over T rounds. Let $\mathcal{A}$ be the Player's action space. Suppose that $\mathcal{X} \subseteq \mathcal{A}$ is a convex set representing the set of possible actions of the Player. Moreover, let $\mathcal{F}$ denote a set of convex functions available to Nature. At each round t, the Player chooses an action $x_t \in \mathcal{X}$. After the Player commits with an action, Nature reveals a convex cost $f_t : \mathcal{X} \to \mathbb{R}$ where $f_t \in \mathcal{F}$[1]. The Player suffers the loss $f_t(x_t)$. The goal of the Player is to perform as well as possible against the costs chosen by Nature. (The wonderful monographs [Haz16], [CBL06] and [SS+12] provide an in-depth study of fundamentals theories of OCO and its many applications). The protocol of the game is summarized as follows:

OCO Protocol

1. **for** $t = 1, \dots, T$
2. Player chooses an action $x_t \in \mathcal{X}$.
3. Nature chooses a cost $f_t \in \mathcal{F}$.
4. Player suffers the loss $f_t(x_t)$.
5. **end for**

[1] See Appendix C for a discussion on adversary types.

A common metric to evaluate the performance of the Player is the so-called *static regret*, defined as

$$\mathbf{Reg}_T^s := \sum_{t=1}^{T} f_t(x_t) - \min_{x \in \mathcal{X}} \sum_{t=1}^{T} f_t(x). \tag{1-1}$$

Intuitively, this metric quantifies how well the Player performs against the *best fixed* action computed in hindsight. Based on this notion of regret, OCO algorithms are designed such that the resulting action sequence $\{x_t\}_{t=1}^T$ guarantees a *sub-linear* regret w.r.t. T, i.e., $\lim_{T\to\infty}(\mathbf{Reg}_T^s/T) = 0$. In other words, such OCO strategies perform (on average) as well as the best fixed action in hindsight.

1-2 Application Example: Portfolio Selection

In this section, we give a concrete example of how one can use the OCO framework to model a real-world problem: the portfolio selection problem. For a more detailed description of this problem, please refer to [Haz16, Chapter 4].

In this application, the Player can be interpreted as an investor, and Nature as the stock market. At the beginning of day t, the Player chooses an action $x_t \in \Delta_n$, where $\Delta_n = \{x \in \mathbb{R}^n \mid x(i) \geq 0, \sum_{i=1}^n x(i) = 1\}$ is the n-dimensional simplex. This action corresponds to choosing a probability distribution over n assets. That is, the action x_t represents how the investor distributes his/her wealth over n assets. Nature (the stock market), on the other hand, chooses a strictly positive return vector $r_t \in \mathbb{R}_+^n$, where each entry corresponds to the return of an asset at day t, that is, the ratio of the value of the asset between days t and $t+1$. The ratio of the wealth of the Player (investor) between rounds t and $t+1$ is $r_t^\top x_t$, thus, the gain at round t can be defined as $\log(r_t^\top x_t)$. In a game of T rounds, the goal of the player is to maximize $\sum_{t=1}^T \log(r_t^\top x_t)$ (equivalently, minimize $\sum_{t=1}^T -\log(r_t^\top x_t)$), which nicely fits into the OCO framework. Below we rewrite the OCO protocol for this specific scenario.

OCO Protocol: Portfolio Selection

1. **for (days)** $t = 1, \ldots, T$
2. Player (investor) chooses $x_t \in \Delta_n$, that is, how he/she distributes his/her wealth over n assets.
3. Nature (stock market) chooses $r_t \in \mathbb{R}_+^n$, the return of each asset at the end of day t.
4. Player (investor) suffers loss $f_t(x_t) = -\log(r_t^\top x_t)$.
5. **end for**

In this application, using the static regret (1-1) as the Player's performance metric, the best fixed decision calculated in hindsight is called the *Best Constant Rebalanced Portfolio*

(BCRP), which is a powerful investing strategy[2]. Thus, if a sub-linear regret algorithm is used for the portfolio selection problem, the average loss of investor will approach the one of the BCRP, even though the investor chooses his/her actions "online".

1-3 Online Convex Optimization with Predictions

In the classical OCO protocol, the Player chooses its actions x_t based only on the information available up to round $t-1$. In other words, the Player uses only information about the *past*. This is the case because in most of the OCO literature, in the same way the Player aims to minimize its regret, Nature is assumed to choose the costs with the goal of *maximizing* the Player's regret (i.e. in a adversarial fashion). Therefore, it would not make sense to use information about *future* costs when choosing x_t.

However, in many real-world applications, it is natural assume that Nature would not be completely adversarial relative to the Player's actions. For example, recall our portfolio selection problem. In this scenario, it is reasonable to assume that Nature (the stock market) would not choose the return of each stock in order to maximize the regret of the Investor. Instead, the stock marked would follow its natural dynamics. Therefore, choosing the actions x_t as if Nature were completely adversarial can be, in many practical scenarios, overly pessimistic.

Thus, the main goal of this thesis is to tackle the following question: how can we exploit the "niceness" of Nature in OCO problems? In other words, in cases where the behavior of Nature is not completely adversarial, and the Player has some knowledge of this behavior, how can it choose actions x_t in order to exploit this knowledge? We call this approach *OCO with Predictions*. In these scenarios, we assume access to *prediction models* about the problem being studied, and we use OCO algorithms combined with these models in order to achieve improved regret rates. This approach is partially inspired by the classical control theory literature, in which dynamical models of the system being controlled are almost always assumed to exist.

1-4 Organization and Mathematical Notation

The organization of the thesis is as follows. In the remainder of Chapter 1, we introduce the mathematical definitions and notations that are used in the rest of the thesis. In Chapter 2, we introduce some central concepts used in this thesis, and also give a concise review of related works. The main results of this thesis are provided in Chapter 3. Numerical experiments are presented in Chapter 4. Chapter 5 concludes the thesis by discussing several future research directions. Technical proofs are provided in appendixes A and B.

We close this chapter by introducing several mathematical notions that are employed in the rest of the thesis. Let the action set $\mathcal{X} \subset \mathbb{R}^n$. We denote by $\|\cdot\|_*$ the dual norm of $\|\cdot\|$. Also, we denote $[T] = \{1, 2, \dots, T\}$. Next, we define the *Bregman Divergence*, a central concept for the algorithms studied in this thesis.

Definition 1 (Bregman divergence)**.** *Let $h : \mathcal{X} \to \mathbb{R}$ be a differentiable function. The Bregman divergence of $x, y \in \mathcal{X}$, w.r.t. the function h is $\mathcal{B}_h(x, y) := h(x) - h(y) - \langle \nabla h(y), x - y \rangle$.*

[2] For more details about Constant Rebalancing Portfolios, see [Haz16, Chapter 4.1.3]

The Bregman divergence can be interpreted as a *distance-like* function. In fact, for specific choices of h, the Bregman divergence is equivalent to well-known functions.

Example 1 (Euclidean setup)**.** *Let* $h(x) = \frac{1}{2}\|x\|_2^2$*. Then,* $\mathcal{B}_h(x, y) = \frac{1}{2}\|x - y\|_2^2$.

Example 2 (Simplex setup)**.** *Let* $\mathcal{X}$ *be the* n*-dimensional simplex* $\Delta_n = \{x \in \mathbb{R}^n \mid x \geq 0,\ \sum_{i=1}^n x_i = 1\}$ *and* $h(x) = \sum_{i=1}^n x_i \log(x_i)$*, the negative entropy function. Then,* $\mathcal{B}_h(x, y) = \sum_{i=1}^n x_i \log(\frac{x_i}{y_i}) - \sum_{i=1}^n (x_i - y_i)$*, also known as generalized Kullback-Leibler divergence.*

Finally, we define two classes of functions.

Definition 2 (α-Strong convexity)**.** *A function* $f : \mathcal{X} \to \mathbb{R}$ *is* α-strongly convex *w.r.t. a norm* $\|\cdot\|$ *if* $f(x) - f(y) \leq \langle \nabla f(x), x - y \rangle - \frac{\alpha}{2}\|x - y\|^2$*, for all* $x, y \in \mathcal{X}$.

Definition 3 (β-Smoothness)**.** *A function* $f : \mathcal{X} \to \mathbb{R}$ *is* β-smooth *w.r.t. a norm* $\|\cdot\|$ *if it is differentiable and* $\|\nabla f(x) - \nabla f(y)\|_* \leq \beta\|x - y\|$*, for all* $x, y \in \mathcal{X}$.

Chapter 2

Preliminaries

2-1 Online Mirror Descent

Returning to the discussion started in the beginning of Chapter 1, it was stated that a desired property of OCO algorithms would be to achieve *sub-linear* static regret. This begs the question: are there algorithms that achieve this regret rate for an arbitrary sequence of convex costs $\{f_t\}_{t=1}^T$ chosen by Nature. Turns out, the answer is yes. A famous algorithm that guarantees this regret rate is the so-called *Online Mirror Descent* (OMD) algorithm, which can be interpreted as an *online* version of the *Mirror Descent* (MD) algorithm introduced by Nemirovski and Yudin [NY83]. The following definition of this algorithm is based on the formulation introduced in [BT03][1]:

$$x_{t+1} = \underset{x \in \mathcal{X}}{\operatorname{argmin}} \ \{\eta_t \langle \nabla f_t(x_t), x \rangle + \mathcal{B}_h(x, x_t)\}, \tag{OMD}$$

where η_t is called the *step-size*. Interestingly, Algorithm (OMD) generalizes many classical algorithms from the literature, which are recovered for specific choices of h.

Example 1 (Euclidean setup - continued)**.** *In this case, the update step of Algorithm* (OMD) *can be shown to be equivalent to the so-called* Online Gradient Descent (OGD) *algorithm*

$$x_{t+1} = \Pi_{\mathcal{X}}(x_t - \eta_t \nabla f_t(x_t)),$$

where $\Pi_{\mathcal{X}}(x)$ *is the euclidean projection of* x *onto* $\mathcal{X}$*.*

Example 2 (Simplex setup - continued)**.** *In this case, the update step of Algorithm* (OMD) *can be shown to be equivalent to an online version of the so-called* Exponentiated Gradient Descent (EGD) *algorithm*

$$y_{t+1}(i) = x_t(i) \exp(-\eta_t [\nabla f_t(x_t)]_i), \quad i \in [n]$$
$$x_{t+1} = \frac{y_{t+1}}{\|y_{t+1}\|_1}.$$

[1] With an abuse of notation, when f_t is not differentiable, $\nabla f_t(x_t)$ denotes a subgradient of f_t at x_t.

By choosing η_t appropriately, Algorithm (OMD) guarantees $\mathbf{Reg}_T^s \leq O(\sqrt{T})$ [Bub11] or $\mathbf{Reg}_T^s \leq O(\log(T))$ [SSS07], for convex and strongly convex costs, respectively. Moreover, [ABRT08] showed that these regret rates are in fact optimal by the *minimax* formulation of OCO problems.

Next, we formally define the concept of *gradient predictions* and also introduce the notion of *dynamic environments*.

2-2 Gradient Predictions

The minimax regret bounds for OCO algorithms are derived assuming a worst-case (i.e. fully adversarial) cost sequence $\{f_t\}_{t=1}^T$. The cost sequence is however *not* completely adversarial in many practical OCO problems [RS13a]. In such problems, the Player can (partially) *predict* the *unseen* cost f_t at round t, before deciding its action x_t.[2] It is hence natural to expect that one can possibly exploit the predictability of an OCO problem to achieve tighter regret bounds.

A generic notion of the predictability of Nature's moves can be stated as follows [RS13a]. At the outset of each round $t \in [T]$, the Player has access to the value of a function

$$M_t : \mathcal{X}^{t-1} \times \mathcal{F}^{t-1} \times \mathcal{I}^{t-1} \to \mathcal{P},$$

where $\mathcal{I}$ denotes some information space provided to the Player via an exogenous source and $\mathcal{P}$ is the space to which each predictable entity belongs to. In particular, a certain class of OCO problems with predictability is the class of OCO problems with *gradient predictions*. Observe that here $\mathcal{P} \subseteq \mathcal{A}^*$, where $\mathcal{A}^*$ is the dual space of the action space $\mathcal{A}$. To exploit gradient predictions in OCO problems, [RS13a] proposed the Optimistic Mirror Descent (OptMD) algorithm

$$\begin{aligned} x_t &= \underset{x \in \mathcal{X}}{\operatorname{argmin}} \ \{\eta_t \langle M_t, x\rangle + \mathcal{B}_h(x, y_{t-1})\} \\ y_t &= \underset{y \in \mathcal{X}}{\operatorname{argmin}} \ \{\eta_t \langle \nabla f_t(x_t), y\rangle + \mathcal{B}_h(y, y_{t-1})\}, \end{aligned} \tag{OptMD}$$

where $\{M_t\}_{t=1}^T$ is a generic gradient prediction sequence[3]. In [RS13b], the authors further provided an adaptive step-size rule for Algorithm (OptMD) such that $\mathbf{Reg}_T^s \leq O(1 + \sqrt{D_T})$, where

$$D_T := \sum_{t=1}^T \|\nabla f_t(x_t) - M_t\|_*^2. \tag{2-1}$$

When the Player has access to $\nabla f_t(\cdot)$ *before* choosing x_t, we say that the Player has access to *perfect gradient predictions*. In this scenario, [HNKK19] showed that by setting $M_t := \nabla f_t(y_{t-1})$ and when $\mathcal{F}$ represents β-*smooth* functions, Algorithm (OptMD) guarantees $\mathbf{Reg}_T^s \leq O(1)$.

[2]This assumption deviates from the standard OCO protocol, where Nature reveals f_t *after* the Player chooses x_t.

[3]Notice that Algorithm (OptMD) reduces to Algorithm (OMD) when $M_t = 0$.

2-3 Dynamic Environments

In the regret notion (1-1), the Player's cumulative loss competes against the loss of the best fixed action in hindsight. There are, on the other hand, many OCO problems that the best fixed action is not accessible or does not exist [BGZ15]. Thus, in those cases, the use of the regret (1-1) is not convenient anymore. The term OCO problems in *dynamic environments* is used in the literature for such problems [HW15].

To "generalize" the standard regret notion in order to tackle these scenarios, [Zin03] proposed to compare the Player's performance against a general dynamical *reference sequence* $\{u_t\}_{t=1}^T \in \mathcal{X}^T$. The resulting metric is called the *dynamic regret*, defined as

$$\mathbf{Reg}_T^d := \sum_{t=1}^T f_t(x_t) - \sum_{t=1}^T f_t(u_t). \tag{2-2}$$

Unfortunately, it is impossible to achieve a sub-linear dynamic regret for an arbitrarily chosen $\{u_t\}_{t=1}^T$ [MSJR16]. Thus, in order to achieve *meaningful* dynamic regret bounds, it is common to place extra *regularity assumptions* on the costs and/or the reference sequence. For example, [HW13] consider the bounded variability of the reference sequence in terms of

$$C_T := \sum_{t=1}^T \|u_{t+1} - u_t\|. \tag{2-3}$$

For convex costs, the authors show that Algorithm (OMD) guarantees $\mathbf{Reg}_T^d \leq O(\sqrt{T}(1 + C_T))$. [HW13] further consider that the Player has access to a (possibly approximate) dynamical model $\Phi_t : \mathcal{X} \to \mathcal{X}$ of the behavior of the reference sequence, that is, $u_{t+1} \approx \Phi_t(u_t)$. They employ $\Phi_t(x_t)$ instead of x_t in Algorithm (OMD) and prove $\mathbf{Reg}_T^d \leq O(\sqrt{T}(1 + C_T'))$, where

$$C_T' := \sum_{t=1}^T \|u_{t+1} - \Phi_t(u_t)\|. \tag{2-4}$$

This modified OMD algorithm is called *Dynamic Mirror Descent* (DMD) algorithm. When Φ_t approximates the dynamics of the reference sequence well, we may have $C_T' \leq C_T$, which in turn implies a tighter dynamic regret bound. Subsequently, [JRSS15] studied dynamical environments to account for the cases with gradient predictions. The authors show that Algorithm (OptMD) guarantees $\mathbf{Reg}_T^d \leq O\left(\sqrt{1 + D_T}(1 + C_T)\right)$ in such cases.

2-4 The Problem with D_T

As already mentioned, [RS13b] and [JRSS15] use Algorithm (OptMD) to prove static and dynamic regret bounds in terms of D_T. In these works, it is argued that for "predictable sequences", external knowledge of the gradient sequence can be used to achieve tighter regret bounds. For example, in [JRSS15], it is stated that: "...one can get a tighter bound for regret once the learner advances a sequence of conjectures $\{M_t\}_{t=1}^T$ well-aligned with the gradients". However, in what follows we argue that in practice, regret bounds given in terms of D_T are not suitable for exploiting gradient predictions.

Consider the following scenario: at the beginning of round t, the Player has access to a prediction of $\nabla f_t(\cdot)$, namely $\nabla \hat{f}_t(\cdot)$. Now, based on those regret bounds given in terms of D_T, how would one set M_t when using Algorithm (OptMD)? Naturally, we want to choose M_t so that D_T is as small as possible (recall that $D_T = \sum_{t=1}^{T} \|\nabla f_t(x_t) - M_t\|_*^2$). However, since x_t **is not available** at the beginning of round t (see Algorithm (OptMD)), we cannot set $M_t = \nabla \hat{f}_t(x_t)$. Thus, from these regret bounds, it is not clear how one should choose M_t in order to exploit this type of gradient prediction. Moreover, [HNKK19] showed that when perfect gradient predictions are available (that is, $\nabla \hat{f}_t(\cdot) = \nabla f_t(\cdot)$), constant static regret is achievable. Still, this constant regret result is not recovered by the regret bound $\mathbf{Reg}_T^s \leq O(1 + \sqrt{D_T})$ given in [RS13b, Corollary 2], even when perfect gradient predictions are available. In fact, since [RS13b, Corollary 2] does not assume smoothness of the cost functions, if it was possible to choose M_t such that $D_T = 0$ (i.e., such that $\mathbf{Reg}_T^s \leq O(1)$), this would contradict the lower bound from [Nes04, Theorem 3.2.1] for any first-order optimization method (see Remark 1 in [YMJZ14]).

Therefore, we conclude that in order to exploit gradient predictions, a different approach must be used. In the next chapter, we will show that by carefully choosing an *adaptive* step-size η_t and assuming smoothness of the cost sequence is enough to achieve this goal.

2-5 Setup

In this thesis, we consider OCO problems with a certain of type gradient predictability, in both static and dynamic environments. Recall that [HNKK19] observed that perfect gradient predictability in the form of $M_t = \nabla f_t(y_{t-1})$ implies that Algorithm (OptMD) guarantees a constant static regret. Motivated by this observation and the discussion presented in section 2-4, we *extend* this idea to the case of an "imperfect" gradient predictability. To do so, we employ the entity

$$D'_t := \sum_{\tau=1}^{t} \|\nabla f_\tau(y_{\tau-1}) - M_\tau\|_*^2. \tag{2-5}$$

Moreover, for dynamic environments, we further suppose that the Player has access to a (possibly approximate) dynamical model Φ_t of the reference sequence $\{u_t\}_{t=1}^T$. This is a useful assumption, which has been used in practical applications of OCO algorithms (see [SJ17] and [WCSB19], for example).

We are now set to formally state the problem considered in this thesis.

Problem 1. *Design and/or analyse OCO algorithms such that the corresponding regret bounds exploit*

- *(possibly imperfect) gradient predictions of the cost sequence f_t;*
- *(possibly approximate) dynamical models Φ_t of the reference sequence $\{u_t\}_{t=1}^T$.*

2-6 Related Works

Other than the works already mentioned so far, several studies in the literature propose algorithms that take advantage of the predictability of the cost sequence. [CCL+16], [LQL18] and [LGW20] employ predictions in OCO problems with switching costs. In this scenario, at round t, the Player suffers the loss $f_t(x_t, x_{t-1}) = c_t(x_t) + \|x_t - x_{t-1}\|$, where c_t is a convex function. Notice that f_t is a function of both x_t and x_{t-1}. Thus, the results of these papers are not directly comparable to this thesis' results (here, we only consider the case where $f_t(x_t)$). [DFHJ17] study Online Linear Optimization. The authors suppose that at the outset of each round, the Player has access to a vector (or hint) that is correlated with the cost to be incurred to the Player. If *all* hints are sufficiently good and the action set possesses certain geometrical properties, the authors show $\mathbf{Reg}_T^s \leq O(\log(T))$. Recently, [BCKP20] extended this results to the case when not all hints are correlated with the true cost vector. In dynamic environments, [LLST20] show that tighter dynamic regret bounds can be achieved by only using predictions that meet certain quality conditions. In [RCT19], the authors employ gradient predictions in order to obtain possibly tighter dynamic regret bounds. However, the proposed approach yields regret bounds that lack worst-case guarantees.

Chapter 3

Main Results

In this section, we provide static and dynamic regret bounds that exploit gradient predictability and/or dynamical models of the reference sequence. We present results for two classes of costs: convex and strongly-convex. The corresponding proofs are postponed to Appendix B. First, we collect several assumptions which we will employ in the results to follow.

Assumption 1 (Regularity assumptions)**.** *Let $\mathcal{A}$ be a Banach space equipped with the norm $\|\cdot\|$. Suppose that*

- *The set $\mathcal{X}$ is a convex subset of $\mathcal{A}$;*
- *The map $h : \mathcal{A} \to \mathbb{R}$ is differentiable and 1-strongly convex on $\mathcal{X}$;*
- *Each member of the cost sequence $\{f_t\}_{t=1}^T$ is convex and β-smooth;*
- *For all $x, y \in \mathcal{X}$, it holds that $\mathcal{B}_h(x, y) \leq R^2$ where $R > 0$;*
- *For all $t \in [T]$, the gradient prediction M_t satisfies $\|\nabla f_t(y_{t-1}) - M_t\|_* \leq \sigma$ where $\sigma \geq 0$.*

3-1 Static Environments

Our first result concerns convex costs and static environments. In this case, we show that by assuming that the cost sequence is β-smooth and appropriately choosing an adaptive step-size η_t, we can use Algorithm (OptMD) to exploit gradient predictions.

Theorem 1 (Static regret: convex case)**.** *Suppose that Assumption 1 holds. Using the adaptive step-size $\eta_1 = \frac{1}{2\beta}$ and $\eta_t = \left(D'_{t-1} + 4\beta^2\right)^{-1/2}$ for all $t > 1$, Algorithm* (OptMD) *guarantees*

$$\mathbf{Reg}_T^s \leq \frac{\sigma^2}{\beta} + \left(2R^2 + 4\right)\sqrt{D'_T + 4\beta^2}. \tag{3-1}$$

Recall the discussion of Section 2-4 and that $D'_T = \sum_{t=1}^T \|\nabla f_t(y_{t-1}) - M_t\|_*^2$. Here we re-emphasize that we present regret bounds in terms of D'_T instead of D_T. Doing so, we solve the problem of how to choose M_t in order to exploit gradient predictions: simply set $M_t = \nabla \hat{f}_t(y_{t-1})$. This is possible because, differently from x_t, y_{t-1} **is available** at the beginning of round t (see Algorithm (OptMD)). Next, we discuss how the bound of Theorem 1 generalizes known optimal bounds from the literature.

Remark 1 (Generality of static regret bound: convex case)**.** *The static regret inequality* (3-1) *can be stated as* $\mathbf{Reg}_T^s \leq O(1+\sqrt{D'_T})$*. Notice that when perfect predictions are available, setting* $M_t = \nabla f_t(y_{t-1})$ *(and so* $D'_T = 0$*), the regret inequality* (3-1) *reduces to* $\mathbf{Reg}_T^s \leq O(1)$*, recovering the result of [HNKK19, Theorem 4]. On the other hand, in view of the last item in Assumption 1, the regret inequality* (3-1) *also recovers the minimax static regret* $\mathbf{Reg}_T^s \leq O(\sqrt{T})$ *in the worst case, that is, even if the gradient predictions are completely uncorrelated with the true gradients and* $D'_T = O(T)$*.*

Next, we state a static regret result for strongly convex costs. This stronger assumption on the costs allows us to achieve tighter bounds. For this result, we also need the following assumption.

Assumption 2 (Extra regularity assumption: Euclidean case)**.** *Consider that the hypotheses in assumption 1 hold. We further suppose that the action space* $\mathcal{A}$ *is a Euclidean space equipped with the* 2*-norm* $\|\cdot\|_2$ *and the mapping* $h(x)$ *is* $\frac{1}{2}\|x\|_2^2$*. In this case, the Bregman divergence* $\mathcal{B}_h(x,y) = \frac{1}{2}\|x-y\|_2^2$*.*

Assumption 2 formalizes the euclidean setup of Example 1.

Theorem 2 (Static regret: strongly convex case)**.** *Suppose that assumptions 1 and 2 hold and that the cost sequence* $\{f_t\}_{t=1}^T$ *is* α*-strongly convex w.r.t.* $\|\cdot\|_2$*. Using the adaptive step-size* $\eta_1 = \frac{1}{2\beta}$ *and* $\eta_t = \left(\frac{\alpha}{2\sigma^2}D'_{t-1} + 2\beta\right)^{-1}$ *for all* $t > 1$*, Algorithm* (OptMD) *guarantees*

$$\mathbf{Reg}_T^s \leq 2\beta R^2 + \frac{\sigma^2}{\beta} + \frac{4\sigma^2}{\alpha}\log\left(1 + \frac{\alpha}{4\beta\sigma^2}D'_T\right). \tag{3-2}$$

Remark 2 (Generality of static regret bound: strongly convex case)**.** *Observe that the static regret inequality* (3-2) *can be rewritten as* $\mathbf{Reg}_T^s \leq O(1+\log(1+D'_T))$*. Employing a similar line of arguments as in Remark 1, we now state two observations. With perfect gradient predictions, the inequality* (3-2) *becomes* $\mathbf{Reg}_T^s \leq O(1)$ *[HNKK19, Theorem 4]. Moreover, the minimax regret* $\mathbf{Reg}_T^s \leq O(\log(T))$ *is also recovered in the worst-case.*

Table 3-1 shows a summary of our static regret bounds using gradient predictions, together with their worst-case and perfect prediction counterparts.

3-2 Dynamic Environments

As previously mentioned, when working in dynamic environments, we would like to exploit both gradient predictions and knowledge of reference sequence's dynamics. Thus, we propose

	Worst-case	This thesis	Perfect prediction
Static regret + **convex costs**	[Bub11]	Theorem 1	[HNKK19]
	$\eta_t = O(1/\sqrt{t})$	$\eta_t = 1/\sqrt{D'_{t-1} + 4\beta^2}$	$\eta_t = \eta \leq 1/\beta$
	$\mathbf{Reg}_T^s \leq O(\sqrt{T})$	$\mathbf{Reg}_T^s \leq O(1 + \sqrt{D'_T})$	$\mathbf{Reg}_T^s \leq O(1)$
Static regret + **strongly convex costs**	[HAK07]	Theorem 2	[HNKK19]
	$\eta_t = 1/(\alpha t)$	$\eta_t = O(1/(D'_{t-1} + 2\beta))$	$\eta_t = \eta \leq 1/\beta$
	$\mathbf{Reg}_T^s \leq O(\log(T))$	$\mathbf{Reg}_T^s \leq O(1 + \log(D'_T + 1))$	$\mathbf{Reg}_T^s \leq O(1)$

Table 3-1: Summary of results presented in Section 3-1

the *Optimistic Dynamic Mirror Descent* (OptDMD) algorithm

$$\begin{aligned} x_t &= \operatorname*{argmin}_{x \in \mathcal{X}} \; \{\eta_t \langle M_t, x\rangle + \mathcal{B}_h(x, y_{t-1})\} \\ \tilde{y}_t &= \operatorname*{argmin}_{y \in \mathcal{X}} \; \{\eta_t \langle \nabla f_t(x_t), y\rangle + \mathcal{B}_h(y, y_{t-1})\} \\ y_t &= \Phi_t(\tilde{y}_t), \end{aligned} \tag{OptDMD}$$

This algorithm can be viewed as a combination of Algorithm (OptMD) and the DMD algorithm of [HW13]. To the best of our knowledge, no result in the literature has presented a regret analysis of an algorithm that combines both gradient predictions and knowledge about the dynamics of the reference sequence. In what follows, we assume that the Player has access to dynamical models Φ_t of $\{u_t\}_{t=1}^T$. Let us further make the following assumptions.

Assumption 3 (Lipschitz-likeness of $\mathcal{B}_h$)**.** *For all $x, y, z \in \mathcal{X}$, there exist a scalar $\gamma > 0$ such that the Bregman divergence satisfies the Lipschitz-like condition $\mathcal{B}_h(x, z) - \mathcal{B}_h(y, z) \leq \gamma \|x - y\|$.*

Remark 3 (Mildness of Assumption 3)**.** *It follows that Assumption 3 holds when the mapping h is Lipschitz on $\mathcal{X}$ [JRSS15, Section 3.1].*

Assumption 4 (Non-expansiveness of Φ_t)**.** *For all $x, y \in \mathcal{X}$ and $\mathcal{B}_h$, the mapping Φ_t is non-expansive, that is, $\mathcal{B}_h(\Phi_t(x), \Phi_t(y)) - \mathcal{B}_h(x, y) \leq 0$.*

Remark 4 (Necessity of Assumption 4)**.** *Observe that Assumption 4 is a restriction on the class of dynamical models Φ_t. The reason behind this assumption is to control the impact of a possibly unreliable prediction (made at some step), as the online game progresses. It is worth mentioning that this assumption has been used in the literature [HW15] [SJ17].*

We now present a dynamic regret bound for the Algorithm (OptDMD).

Theorem 3 (Dynamic regret: convex case using (OptDMD))**.** *Suppose that assumptions 1, 3 and 4 hold. Using the adaptive step-size $\eta_1 = \frac{1}{2\beta}$ and $\eta_t = \left(D'_{t-1} + 4\beta^2\right)^{-1/2}$ for all $t > 1$,* *Algorithm* (OptDMD) *guarantees*

$$\mathbf{Reg}_T^d \leq \frac{\sigma^2}{\beta} + \left(4 + 2R^2 + \gamma C_T'\right)\sqrt{D_T' + 4\beta^2}. \tag{3-3}$$

Notice that in static environments, $C_T' = 0$ and Theorem 3 recovers the bound of Theorem 1.

Remark 5 (Comparison with [JRSS15])**.** *Let us first rewrite the regret inequality* (3-3) *in the compact form* $\mathbf{Reg}_T^d \leq O\left(\sqrt{D_T' + 1}(C_T' + 1)\right)$*. Next, observe that when* Φ_t *approximates the true dynamics of the comparator sequence* $\{u_t\}_{t=1}^T$*, we may have* $C_T' \leq C_T$*. Moreover, we also recover* $C_T' = C_T$ *if we choose* Φ_t *as the identity map. Therefore, compared to the bound* $\mathbf{Reg}_T^d \leq O(\sqrt{D_T + 1}(C_T + 1))$ *provided by [JRSS15], our result improves it in the sense that it is given in terms of* C_T' *and* D_T'*, instead of* C_T *and* D_T *(recall the discussion of Section 2-4).*

Remark 6 (Comparison with [ZLZ18])**.** *Recall that we have* $\|\nabla f_t(y_{t-1}) - M_t\|_* \leq \sigma$ *by Assumption 1. Hence, it follows that* $O(\sqrt{1 + D_T'}(1 + C_T')) = O(\sqrt{T}(1 + C_T'))$ *in the worst-case, and we recover the bound given in [HW13]. However, [ZLZ18] proposed an algorithm called* Ader*, which achieves the* optimal *bound* $\mathbf{Reg}_T^d \leq O\left(\sqrt{T(1 + C_T')}\right)$*. Thus, in the worst-case, our regret bound* (3-3) *does not recover the optimal one. Nonetheless, in Section 4-3 we present numerical results, which show that Algorithm* (OptDMD) *can outperform* Ader *in practice. Moreover, Algorithm* (OptDMD) *uses general Mirror Descent updates, whereas* Ader *is based on Online Gradient Descent (OGD) updates (i.e., Algorithm* (OMD) *restricted by Assumption 2).*

In what follows, we present a different dynamic regret bound for OCO problems with convex costs. To do so, we introduce (i) a specific type of the reference sequence, (ii) an extra assumption and (iii) a modification of Algorithm (OptDMD). We consider that the Player's performance is now compared against the reference sequence $\{x_t^*\}_{t=1}^T$, defined as

$$x_t^* := \operatorname*{argmin}_{x \in \mathcal{X}} f_t(x).$$

In this case, the notation $\mathbf{Reg}_T^d(x_1^*, \ldots, x_T^*)$ indicates that the Player's performance is compared against $\{x_t^*\}_{t=1}^T$. For this particular reference sequence, the following assumption have been used in the literature to prove tighter dynamic regret bounds.

Assumption 5 (Vanishing gradient [YZJY16])**.** *For all* $t \geq 1$*,* x_t^* *belongs to the relative interior of* $\mathcal{X}$*, that is,* $\nabla f_t(x_t^*) = 0$*.*

We now introduce a *modification* to the step-size with which $\tilde{y}_t$ is updated in Algorithm (OptDMD). Fixing it to a constant value $1/\omega$, we get (in the Euclidean case)

$$\begin{aligned} x_t &= \operatorname*{argmin}_{x \in \mathcal{X}} \; \{\eta_t \langle M_t, x\rangle + \frac{1}{2}\|x - y_{t-1}\|_2^2\} \\ \tilde{y}_t &= \operatorname*{argmin}_{y \in \mathcal{X}} \; \{\frac{1}{\omega}\langle \nabla f_t(x_t), y\rangle + \frac{1}{2}\|y - y_{t-1}\|_2^2\} \\ y_t &= \Phi_t(\tilde{y}_t). \end{aligned} \tag{OptDMD-mod}$$

A dynamic regret bound of Algorithm (OptDMD-mod) for convex costs reads as follows.

Theorem 4 (Dynamic regret: convex case using (OptDMD-mod))**.** *Suppose that Assumptions 1-5 hold. Using the adaptive step-size* $\eta_1 = \frac{1}{3\beta}$ *and* $\eta_t = (D'_{t-1} + 3\beta)^{-1}$ *for all* $t > 1$*, Algorithm* (OptDMD-mod) *with* $\omega = 2\beta$ *guarantees*

$$\mathbf{Reg}_T^d(x_1^*, \dots, x_T^*) \leq \frac{\sigma^2}{3\beta} + 4\beta R^2 + \log\left(1 + \frac{D'_T}{3\beta}\right) + 2\gamma\beta \sum_{t=1}^{T} \|x_{t+1}^* - \Phi_t(x_t^*)\|_2.$$

Remark 7 (Comparison with literature)**.** *The upper-bound of Theorem 4 can be stated as* $O(1 + \log(1 + D'_T) + \sum_{t=1}^{T} \|x_{t+1}^* - \Phi_t(x_t^*)\|_2)$*. In [YZJY16], the authors prove that using Assumption 5, the OGD algorithm (that is, Algorithm* (OMD) *in the euclidean case) with a constant step-size* $\eta = 0.5/\beta$ *guarantees* $\mathbf{Reg}_T^d(x_1^*, \dots, x_T^*) \leq O(1 + \sum_{t=1}^{T} \|x_{t+1}^* - x_t^*\|_2)$*. Compared to our result, their bound does not exploit the dynamical models* Φ_t*, whereas our bound has an extra logarithm additive term. In Section 4-4, we present numerical simulations which show that Algorithm* (OptDMD-mod) *can outperform the OGD algorithm with a constant step-size.*

In our final theoretical result, we present a tighter dynamic regret bound for the special (but popular) case of quadratic costs. Formally, we say that f_t is quadratic if it is α-strongly convex and β-smooth, with $\alpha = \beta$.

Theorem 5 (Dynamic regret: quadratic case)**.** *Suppose that Assumptions 1-5 hold and that the sequence* $\{f_t\}_{t=1}^T$ *consists of quadratic costs. Using the adaptive step-size* $\eta_1 = \frac{1}{3\beta}$ *and* $\eta_t = (D'_{t-1} + 3\beta)^{-1}$ *for all* $t > 1$*, Algorithm* (OptDMD-mod) $\omega = \beta$ *guarantees*

$$\begin{aligned}\mathbf{Reg}_T^d(x_1^*, \dots, x_T^*) \leq & \frac{\sigma^2}{6\beta} + \beta R^2 + \frac{1}{2}\log\left(1 + \frac{D'_T}{3\beta}\right) \\ & + \beta \min\left\{\frac{\gamma}{2}\sum_{t=1}^{T} \|x_{t+1}^* - \Phi_t(x_t^*)\|_2,\ \sum_{t=2}^{T} \|x_t^* - \Phi(x_{t-1}^*)\|_2^2\right\}.\end{aligned}$$

Remark 8 (Comparison with literature)**.** *[ZYY+17] also provides dynamic regret bounds in terms of* $O(\min\left\{\sum_{t=1}^{T} \|x_{t+1}^* - x_t^*\|_2,\ \sum_{t=2}^{T} \|x_t^* - x_{t-1}^*\|_2^2\right\})$*. Unlike the setup of Theorem 5, they do not exploit dynamical models* Φ_t *of the reference sequence and their algorithm requires many gradient descent iterations per round. On the other hand, their result is not restricted to quadratic costs.*

On a final note, we remark that a similar regret bound to the one of Theorem 5 can be proved for the more general case when $\beta \leq 3\alpha$ (instead of $\beta = \alpha$). For the sake of simplicity, we chose to present the result for quadratic costs.

Chapter 4

Numerical Experiments

4-1 Portfolio Selection

In this Section, we apply the result of Theorem 1 in a portfolio selection problem. Even though we already introduced this scenario in Section 1-2, for convenience, we repeat its description below.

Suppose that an investor (or the Player) has n assets in a Market (or Nature). Let the Player's action x be a probability distribution over n assets. The action set $\mathcal{X}$ is thus $\Delta_n := \{x \in \mathbb{R}^n : \; x(i) \geq 0, \sum_{i=1}^n x(i) = 1\}$. Let the return of an asset at round t be the ratio of the value of the asset between rounds t and $t+1$. At round t, Nature chooses a strictly positive return vector $r_t \in \mathbb{R}^n_{>0}$ such that each entry of r_t corresponds to the return of an asset. The Player's wealth ratio between rounds t and $t+1$ is $\langle r_t, x_t \rangle$. Let the Player's gain at round t be $\log(\langle r_t, x_t \rangle)$. In a game of T rounds, the goal of the Player is to maximize $\sum_{t=1}^T \log(\langle r_t, x_t \rangle)$ or, equivalently, to minimize $\sum_{t=1}^T -\log(\langle r_t, x_t \rangle)$. Hence, we have $f_t(x) = -\log(\langle r_t, x \rangle)$ and $\nabla f_t(x) = -r_t / \langle r_t, x \rangle$, for all $x \in \mathcal{X}$ (See [Haz16, Chapter 4] for a more detailed description of this problem.).

We assume that the Player has access to (possibly imperfect) prediction models of the return vector r_t, denoted by $\hat{r}_t$. Thus, in light of the approaches proposed in this thesis, we define

$$M_t = \nabla \hat{f}_t(y_{t-1}) := -\frac{\hat{r}_t}{\langle \hat{r}_t, y_{t-1} \rangle}.$$

In what follows, we show that how the Player can employ Algorithm (OptMD) to decide its action sequence $\{x_t\}_{t=1}^T$ considering the static regret (1-1). Since the costs are convex, the Player uses the step-size rule of Theorem 1 in Algorithm (OptMD). By assuming $0.5 \leq r_t \leq 2$ for all t (entry-wise), we can set the smoothness parameter $\beta = 16$. Since Δ_n is the n-dimensional simplex, we let $h(x)$ be the *negative entropy* function $\sum_{i=1}^n x(i) \log(x(i))$. Observe that the mapping h is 1-strongly convex w.r.t. $\| \cdot \|_1$ (see [Bub15, Section 4.3]). We consider the following four prediction models for the returns vector:

1. **perfect**: the perfect prediction model $\hat{r}_t := r_t$;

2. **previous**: a model that uses the previous return vector as its prediction $\hat{r}_t := r_{t-1}$;
3. **noisy**: a noisy predictor model $\hat{r}_t := \max(r_t + w_t, 0)$, where $w_t \sim \mathcal{N}(0, I)$;
4. **random**: a random predictor where, the entries of $\hat{r}_t$ are chosen uniformly between 0 and 2.

We use the NYSE dataset to simulate a stock market ([HSSW98], [AHKS06], [Cov91]). The number of assets n is 36. As a benchmark, we employ the *Constant Uniform Portfolio* (CUP) strategy, that is, a Player that chooses $x_t = [1/36, \dots, 1/36]$, for all $t \in [T]$. For the dataset considered in this experiment, the CUP strategy performed better than Algorithm (OMD), for any $\eta_t > 0$ and $x_0 = [1/36, \dots, 1/36]$.

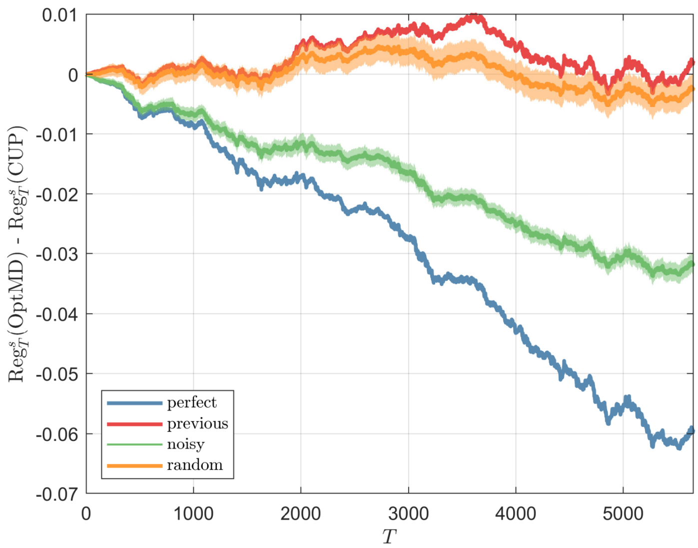


Figure 4-1: Regret difference between a Constant Uniform Portfolio (CUP) and Algorithm (OptMD) using different prediction models of the returns vector r_t.

Denote the regrets of Algorithm (OptMD) and CUP strategies by $\mathbf{Reg}^s_T(\text{OptMD})$ and $\mathbf{Reg}^s_T(\text{CUP})$, respectively. Figure 4-1 depicts the difference $\mathbf{Reg}^s_t(\text{OptMD}) - \mathbf{Reg}^s_t(\text{CUP})$. The experiment was repeated 20 times and the shaded areas correspond to one standard deviation. One can observe that using the adaptive step-size defined in Theorem 1, Algorithm (OptMD) with the **perfect** or **noisy** model achieves a smaller regret than the one achieved by the CUP strategy. In simple words, using our adaptive step-size, we were able to exploit the predictive information about the return vectors. On the other hand, when using the **previous** or **random** models (i.e., gradient predictions which do not approximate the true gradients), Algorithm (OptMD) performs almost as well as the benchmark strategy. This corroborates with the worst-case guarantee of Theorem 1 (see Remark 1).

4-2 Tracking Dynamical Parameters

In this Section, we employ Algorithm (OptDMD-mod) in a parameter tracking problem. The scenario presented in this Section is based on the numerical experiment of [SJ17]. Denote the parameters to be tracked by $x_t^* \in \mathbb{R}^2$. These parameters have dynamics described by the linear model $x_{t+1}^* = Ax_t^* + v_t$. For this experiment, we use

$$A = \begin{bmatrix} 1 & \epsilon \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \Sigma = 5 \begin{bmatrix} \frac{\epsilon^3}{3} & \frac{\epsilon^2}{2} \\ \frac{\epsilon^2}{2} & \epsilon \end{bmatrix},$$

where Σ is the covariance matrix of the Gaussian noise v_t, and $\epsilon = 0.1$ is the sampling interval. The cost at time t is defined as $f_t(x_t) = \frac{1}{2}\|x_t - x_t^*\|_2^2$, where x_t is the output of our tracking algorithm. The costs are quadratic, that is, f_t is α-strongly convex and β-smooth w.r.t. $\|\cdot\|_2$, with $\alpha = \beta = 1$. Moreover, we assume the Player has access to $\Phi_t(x) = Ax$, which is an approximate model of the dynamics of x_t^*.

In what follows, we show how the Player can employ Algorithm (OptDMD-mod) to choose its action sequence $\{x_t\}_{t=1}^T$. Notice that this parameter tracking problem fits the scenario of Theorem 5. Thus, the Player uses Algorithm (OptDMD-mod) with $\omega = \beta$ and step-size $\eta_t = \left(0.01D'_{t-1} + 3\beta\right)^{-1}$. Compared to the step-size defined in the Theorem 5, the only difference is a positive constant multiplying D'_{t-1}.[1] We consider the following gradient prediction models[2]:

1. **perfect**: the perfect prediction model $M_t := \nabla f_t(y_{t-1})$;
2. **noisy**: a noisy prediction model $M_t := \nabla f_t(y_{t-1}) + w_t$;
3. **noisy+bias**: a noisy prediction model plus a bias term $M_t := \nabla f_t(y_{t-1}) + w_t + 3$;
4. **previous**: a prediction model that uses the previous cost gradient $M_t := \nabla f_{t-1}(y_{t-1})$;
5. **random**: a random prediction model $M_t := w_t$,

where $w_t \sim \mathcal{N}(0, I)$. As a benchmark, we use the following algorithm

$$\begin{aligned} \tilde{x}_{t+1} &= \underset{x \in \mathcal{X}}{\operatorname{argmin}} \ \{\eta_t \langle \nabla f_t(x_t), x \rangle + \frac{1}{2}\|x - x_t\|_2^2\} \\ x_{t+1} &= \Phi_t(\tilde{x}_{t+1}), \end{aligned} \tag{4-1}$$

which is a variation of Algorithm (DMD) of [HW13]. According to [MSJR16], we set the step-size of the benchmark as $\eta_t = \eta = 1/\beta$.

Denote the regrets of Algorithm (OptDMD-mod) and Algorithm (4-1) by $\mathbf{Reg}_t^d$(OptDMD-mod) and $\mathbf{Reg}_t^d$(DMD), respectively. Figure 4-2 depicts the difference $\mathbf{Reg}_t^d$(OptDMD-mod)-$\mathbf{Reg}_t^d$(DMD). The experiment is repeated 100 times, and for each experiment, a new trajectory $\{x_t^*\}_{t=1}^T$ was generated. The shaded areas correspond to one standard deviation. One can observe that that all the models that use some kind of information about future gradients (**perfect**, **noisy**, **noisy+bias**) were able to perform better than the benchmark. This shows that indeed Algorithm (OptDMD-mod) was able to exploit predictive information about the problem. On the other hand, even though the **previous** and **random** models did not perform as well as the benchmark, their performance show robustness against inaccurate gradient predictions, which corroborates with the worst-case guarantee of Theorem 5.

[1] One can show that multiplying D'_{t-1} by a positive constant only changes the final regret bound up to multiplicative constants.

[2] Although we could have used the same linear model (i.e. A) in order to construct gradient predictions M_t, here we exemplify that these predictions may come from different sources.

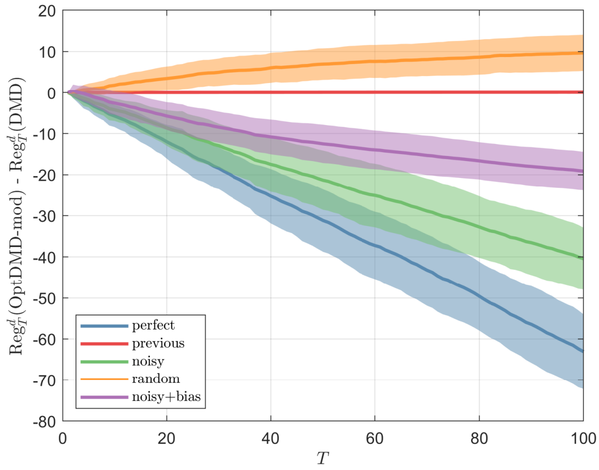


Figure 4-2: Regret difference between Algorithm (DMD) (benchmark) and Algorithm (OptDMD-mod) using different gradient prediction models.

4-3 Algorithm (OptDMD) versus Ader

In this Section, we numerically compare the performance of Algorithm (OptDMD) and the Ader algorithm of [ZLZ18] (see Remark 6 for a discussion on their theoretical guarantees). We consider a 1-D scenario with costs defined as

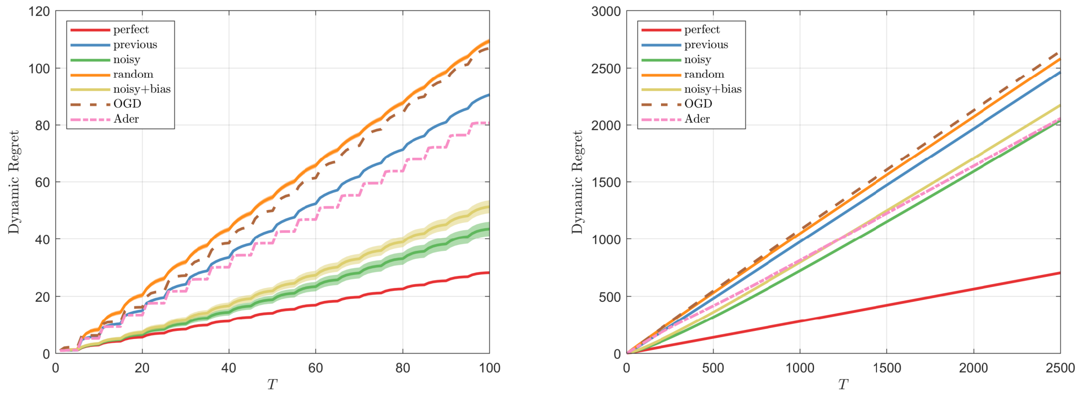
$$f_t(x_t) := \begin{cases} (x_t - u_t)^2 & \text{if } |x_t - u_t| \leq 2 \\ 4|x_t - u_t| - 4 & \text{otherwise,} \end{cases}$$

for some choice of u_t. This function is known as *Huber loss*. In order to use Algorithm (OptDMD), we use $\beta = 2$, set the adaptive step-size η_t according to Theorem 3 and consider the following gradient prediction models:

1. **perfect**: the perfect prediction model $M_t := \nabla f_t(y_{t-1})$;
2. **noisy**: a noisy prediction model $M_t := \nabla f_t(y_{t-1}) + w_t$;
3. **noisy+bias**: a noisy prediction model plus a bias term $M_t := \nabla f_t(y_{t-1}) + w_t + 0.5$;
4. **previous**: a prediction model that uses the previous cost gradient $M_t := \nabla f_{t-1}(y_{t-1})$;
5. **random**: a random prediction model $M_t := w_t$,

where $w_t \sim \mathcal{N}(0, 0.5)$. For simplicity, we will use Φ_t as the identity map, i.e., $\Phi_t(u_t) = u_t$. In this example, we will consider the scenario where $u_t = -\kappa^{-\lfloor t/5 \rfloor}$, for two different choices of κ:

- $\kappa = 1$: in this case, u_t switches between -1 and 1 every 5 iterations;
- $\kappa = 1.1$: in this case, u_t switches between positive and negative values every 5 iterations, but now these values converge to zero as t increases.



(a) $T = 100$.

(b) $T = 2500$.

Figure 4-3: Comparison of Ader [ZLZ18], Online Gradient Descent (OGD) and Algorithm (OptDMD) using different gradient prediction models, with $\kappa = 1$.

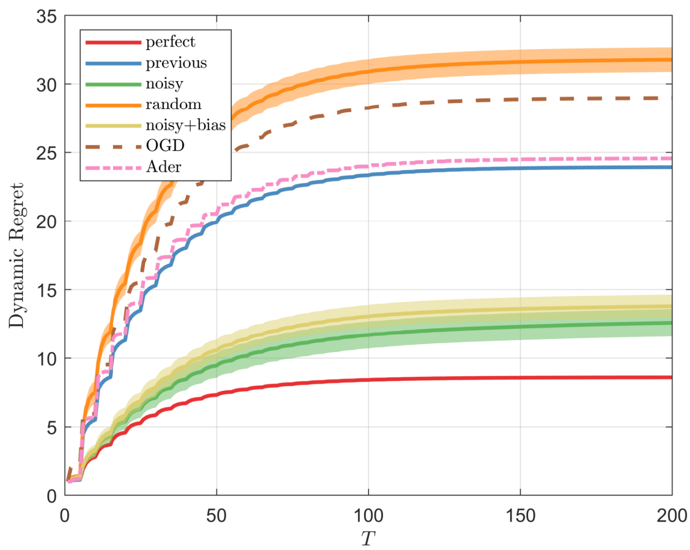


Figure 4-4: Comparison of Ader [ZLZ18], Online Gradient Descent (OGD) and Algorithm (OptDMD) using different gradient prediction models, with $\kappa = 1.1$.

First, we compare the Ader algorithm [ZLZ18] and Algorithm (OptDMD) for the case where $\kappa = 1$. For completeness, we also show the performance of the OGD algorithm using $\eta_t = 1/t$. Figure 4-3 depicts the numerical performance of these algorithms. The experiment was repeated 100 times, and the shaded areas correspond to one standard deviation. In Figure 4-3(a), we can see that Algorithm (OptDMD) with models that use some kind of information about future gradients (**perfect**, **noisy**, **noisy+bias**) outperformed Ader and OGD. However, as discussed in Remark 6, Ader has better *asymptotically* regret guarantees. Thus, in Figure 4-3(b), we increase T from 100 to 2500. In this case, at $T \approx 1250$ and at $T \approx 2500$, Ader starts to outperform the **noisy+bias** and **noisy** models, respectively.

However, even though for $\kappa = 1$, Ader and the **perfect** model have the same regret rate of $O(T)$, Algorithm (OptDMD) shows much better numerical performance. Finally, even though the **previous** and **random** models did not perform as well as Ader, they showed robustness against inaccurate gradient predictions, with performance compared to the one using the OGD algorithm.

Figure 4-4 depicts the case of $\kappa = 1.1$, that is, when u_t converges to 0 as t increases. In this scenario, we see that the Ader algorithms shows poor performance compared to Algorithm (OptDMD), for all prediction models expect the **random** one. This corroborates with the observation drawn from Figure 4-3: even though Ader presents better asymptotic regret bounds, it suffers from slow adaptability. Thus, by the time Ader "adapted" to the problem, u_t already converged to a small value, an it is not able to leverage its better asymptotic bounds anymore. Overall, we conclude that, even though the Ader algorithm has better asymptotic worst-case guarantees, it suffers from slow adaptability compared to Algorithm (OptDMD), which may be prohibitive in practical applications.

4-4 Algorithm (OptDMD-mod) versus OGD with constant step-size

In this Section, we numerically compare the performance of the OGD algorithm (with constant step size $\eta = 0.5/\beta$) [YZJY16] with the performance of Algorithm (OptDMD-mod), according to Theorem 4 (see Remark 7 for comments on the theoretical guarantees of these algorithms). We consider the same scenario as described in Section 4-3 with $\kappa = 1$, that is, u_t switches between -1 and 1 every 5 iterations. Figure 4-5 depicts their numerical performance. As one can see, with the exception of the **random** model, Algorithm (OptDMD-mod) achieved better results than the OGD with constant step-size. Moreover, the **random** model presented performance very similar to the OGD algorithm, which again corroborates with the worst-case theoretical guarantees of Algorithm (OptDMD-mod).

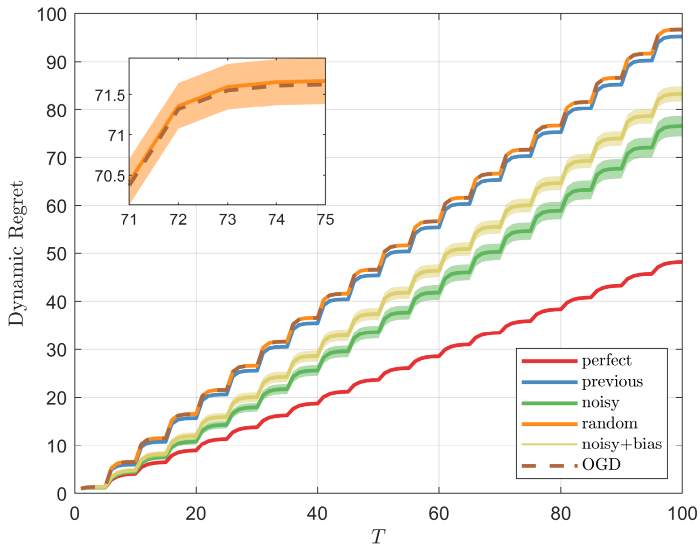


Figure 4-5: Comparison between Online Gradient Descent (OGD) with constant step-size [YZJY16] and Algorithm (OptDMD-mod) using different gradient prediction models.

Chapter 5

Future Directions

There are several directions that one can follow to extend the results of this thesis. One possible direction is to design *meta-learning* algorithms that learn predictive models of ∇f_t and/or dynamical models Φ_t of the reference sequence, while simultaneously playing the game. In this regard, see for example [vEK16], [ZLZ18] and [RCT19]. Another direction is to extend the results of this thesis to bandit (a.k.a. zeroth order) problems, where the Player receives feedback in the form of function evaluations $f_t(\cdot)$, instead of gradients $\nabla f_t(\cdot)$. Based on the multi-point feedback idea proposed by [AD10], if one is allowed to query $2(n+1)$ points per round ($n+1$ points to estimate each $\nabla f_t(x_t)$ and $\nabla f_t(y_{t-1})$), the regret bounds in this thesis still hold up to multiplicative/additive constants. The question is then the possibility of achieving similar results while using only 1-point or 2-point bandit feedback. Next, one could extend the results of this thesis by relaxing the smoothness assumption on the costs. For example, one can borrow the ideas from [YMJZ14, Section 3] to consider the class of costs that each member function is a summation of a varying smooth part and a fixed non-smooth part. Finally, it would be interesting to investigate the possibility of combining the ideas presented in this thesis with other OCO algorithms, e.g., the *Online Newton Step* algorithm of [HAK07].

Appendix A

Auxiliary Lemmas

In what follows, we collect several auxiliary lemmas that are used in the proofs of the results presented in this thesis.

Lemma 1. *Consider the Bregman divergence $\mathcal{B}_h$ as given in Definition 1. If the mapping h is 1-strongly convex w.r.t. a norm $\|\cdot\|$, then,*

$$-\mathcal{B}_h(x, y) \leq -\frac{1}{2}\|x - y\|^2.$$

Proof. Recall that by the definition of the Bregman divergence, we have

$$-\mathcal{B}_h(x, y) = h(y) - h(x) + \langle \nabla h(y), x - y \rangle.$$

On the other hand, the 1-strong convexity of h implies that

$$h(y) - h(x) + \langle \nabla h(y), x - y \rangle \leq -\frac{1}{2}\|x - y\|^2.$$

The lemma's claim is an immediate consequence of the above relations. □

Lemma 2. *Let $h : \mathcal{X} \to \mathbb{R}$ be differentiable 1-strongly convex w.r.t. $\|\cdot\|$. Given $x_t \in \mathcal{X}$, $x_{t+1} = \operatorname{argmin}_{x\in\mathcal{X}} \ \eta_t\langle x, g_t\rangle + \mathcal{B}_h(x, x_t)$ and any $z \in \mathcal{X}$, we have*

$$\langle g_t, x_t - z \rangle \leq \frac{1}{\eta_t}\left(\mathcal{B}_h(z, x_t) - \mathcal{B}_h(z, x_{t+1})\right) + \frac{1}{2}\eta_t\|g_t\|_*^2.$$

Proof. From the optimality of x_{t+1}, we have [BBV04, Section 4.2.3]

$$\langle \eta_t g_t + \nabla h(x_{t+1}) - \nabla h(x_t), z - x_{t+1} \rangle \geq 0,$$

for any $z \in \mathcal{X}$. Manipulating this inequality, one can arrive at

$$\begin{aligned}
\eta_t\langle g_t, x_t - z\rangle \leq\ & h(z) - h(x_t) - \langle \nabla h(x_t), z - x_t \rangle \\
& - h(z) + h(x_{t+1}) + \langle \nabla h(x_{t+1}), z - x_{t+1} \rangle \\
& - h(x_{t+1}) + h(x_t) + \langle \nabla h(x_t), x_{t+1} - x_t \rangle + \eta_t\langle g_t, x_t - x_{t+1} \rangle.
\end{aligned}$$

Using the definition of the Bregman divergence (Definition 1), we have

$$\begin{aligned}
\eta_t\langle g_t, x_t - z\rangle &\leq \mathcal{B}_h(z, x_t) - \mathcal{B}_h(z, x_{t+1}) - \mathcal{B}_h(x_{t+1}, x_t) + \eta_t\langle g_t, x_t - x_{t+1}\rangle \\
&\leq \mathcal{B}_h(z, x_t) - \mathcal{B}_h(z, x_{t+1}) - \frac{1}{2}\|x_{t+1} - x_t\|^2 + \eta_t\langle g_t, x_t - x_{t+1}\rangle && \text{Lemma 1} \\
&\leq \mathcal{B}_h(z, x_t) - \mathcal{B}_h(z, x_{t+1}) - \frac{1}{2}\|x_{t+1} - x_t\|^2 + \eta_t\|g_t\|_*\|x_t - x_{t+1}\| \\
& && \text{generalized Cauchy–Schwarz} \\
&\leq \mathcal{B}_h(z, x_t) - \mathcal{B}_h(z, x_{t+1}) + \frac{1}{2}\eta_t^2\|g_t\|_*^2. && -a^2 + 2ab \leq b^2
\end{aligned}$$

□

Lemma 3. *Suppose that the mapping $f : \mathcal{X} \to \mathbb{R}$ is α-strongly convex w.r.t. a norm $\|\cdot\|$. Then, for all $x, y \in \mathcal{X}$,*

$$\alpha\|x - y\|^2 \leq \langle\nabla f(x) - \nabla f(y), x - y\rangle.$$

Proof. Since f is *α-strongly convex*, Definition 2 implies that for all $x, y \in \mathcal{X}$,

$$\begin{aligned}
f(x) - f(y) &\leq \langle\nabla f(x), x - y\rangle - \frac{\alpha}{2}\|x - y\|^2 \\
f(y) - f(x) &\leq \langle\nabla f(y), y - x\rangle - \frac{\alpha}{2}\|y - x\|^2.
\end{aligned}$$

The claim immediately follows by adding the above two inequalities. □

Lemma 4. *Suppose that $\mathcal{X}$ is a closed convex set in a Euclidean space $\mathbb{E}$ equipped with $\|\cdot\|_2$. Let $w \in \mathbb{E}$, $v \in \mathcal{X}$, and $\eta > 0$. Define*

$$u := \operatorname*{argmin}_{x\in\mathcal{X}} \|(v - \eta w) - x\|_2^2 = \operatorname*{argmin}_{x\in\mathcal{X}} \left\{\eta\langle w, x\rangle + \frac{1}{2}\|x - v\|_2^2\right\}.$$

It then holds that, for all $z \in \mathcal{X}$,

$$\langle w, v - z\rangle \leq \frac{1}{2\eta}\left(\|v - z\|_2^2 - \|u - z\|_2^2\right) + \frac{\eta}{2}\|w\|_2^2.$$

Proof. By definition, u is the Euclidean projection of $(v - \eta w)$ onto $\mathcal{X}$. It follows from the Pythagorean theorem that, for all $z \in \mathcal{X}$,

$$\|u - z\|_2^2 \leq \|v - \eta w - z\|_2^2,$$

and as a result,

$$\|u - z\|_2^2 \leq \eta^2\|w\|_2^2 - 2\eta\langle w, v - z\rangle + \|v - z\|_2^2.$$

Rearranging the above inequality concludes the proof. □

Lemma 5. *Suppose that $\mathcal{X}$ is a closed convex set in a Banach space $\mathbb{S}$. Let $w \in \mathbb{S}^*$, $v \in \mathcal{X}$, and $\eta > 0$. Define*

$$u := \underset{x \in \mathcal{X}}{\operatorname{argmin}} \Big\{ \eta \langle w, x \rangle + \mathcal{B}_h(x, v) \Big\}.$$

It follows that, for all $z \in \mathcal{X}$,

$$\langle w, u - z \rangle \leq \frac{1}{\eta} (\mathcal{B}_h(z, v) - \mathcal{B}_h(z, u) - \mathcal{B}_h(u, v)).$$

Proof. Let $g(x) := \eta \langle w, x \rangle + \mathcal{B}_h(x, v)$ and observe that $\nabla g(x) = \eta w + \nabla h(x) - \nabla h(v)$. Since u is an optimal solution to the above problem, $\langle \nabla g(u), z - u \rangle \geq 0$ for all feasible $z \in \mathcal{X}$, i.e.,

$$\langle \eta w + \nabla h(u) - \nabla h(v), z - u \rangle \geq 0, \ \forall z \in \mathcal{X},$$

and as a result,

$$\eta \langle w, u - z \rangle \leq \langle \nabla h(v) - \nabla h(u), u - z \rangle. \tag{A-1}$$

In what follows, we properly reformulate the right-hand side of the above inequality, that is

$$\langle \nabla h(v), u - z \rangle + \langle \nabla h(u), z - u \rangle =: a.$$

Notice that by the definition of the Bregman divergence (see Definition 1),

$$\langle \nabla h(u), z - u \rangle = h(z) - h(u) - \mathcal{B}_h(z, u) =: a_1.$$

Now observe that

$$\begin{aligned} \langle \nabla h(v), u - z \rangle &= \langle \nabla h(v), u - z + v - v \rangle \\ &= \langle \nabla h(v), u - v \rangle - \langle \nabla h(v), z - v \rangle \\ &=: a_2 - a_3. \end{aligned}$$

We again employ Definition 1 and arrive at

$$\begin{aligned} a_2 &= h(u) - h(v) - \mathcal{B}_h(u, v) \\ a_3 &= h(z) - h(v) - \mathcal{B}_h(z, v). \end{aligned}$$

Now, since $a = a_1 + a_2 - a_3$ and in the light of inequality (A-1), the claim of the lemma follows. □

Lemma 6. *Suppose that $\mathcal{X}$ is a closed convex set in a Banach space $\mathbb{S}$ equipped with a norm $\|\cdot\|$. Let $w_1, w_2 \in \mathbb{S}^*$, $v \in \mathcal{X}$, and $\eta > 0$. Define*

$$u_1 := \underset{x_1 \in \mathcal{X}}{\operatorname{argmin}} \Big\{ \eta \langle w_1, x_1 \rangle + \mathcal{B}_h(x_1, v) \Big\} \text{ and } u_2 := \underset{x_2 \in \mathcal{X}}{\operatorname{argmin}} \Big\{ \eta \langle w_2, x_2 \rangle + \mathcal{B}_h(x_2, v) \Big\}.$$

Then, it holds that

$$\|u_1 - u_2\| \leq \eta \|w_1 - w_2\|_*.$$

Proof. Let us first state two inequalities that are implied by equation (A-1) in the proof of Lemma 5. In equation (A-1), set $w = w_1$, $u = u_1$, and $z = u_2$. We thus have

$$\eta\langle w_1, u_2 - u_1\rangle \geq \langle \nabla h(u_1) - \nabla h(v), u_1 - u_2\rangle.$$

Next, set $w = w_2$, $u = u_2$, and $z = u_1$ in equation (A-1), we now get

$$\eta\langle -w_2, u_2 - u_1\rangle \geq \langle \nabla h(v) - \nabla h(u_2), u_1 - u_2\rangle.$$

Adding the last two inequalities up, we arrive at

$$\eta\langle w_1 - w_2, u_2 - u_1\rangle \geq \langle \nabla h(u_1) - \nabla h(u_2), u_1 - u_2\rangle. \tag{A-2}$$

Since h is 1-strongly convex, it follows from Lemma 3 that

$$\langle \nabla h(u_1) - \nabla h(u_2), u_1 - u_2\rangle \geq \|u_1 - u_2\|^2. \tag{A-3}$$

Combining (A-2) and (A-3), and using the Cauchy-Schwarz inequality, we have

$$\|u_1 - u_2\|^2 \leq \eta\|w_1 - w_2\|_* \|u_2 - u_1\|.$$

As a result, the claim follows. □

The next lemma is a generalization of [SM10, Lemma 1].

Lemma 7. *Let c be a positive real and n be a positive integer. For all sequences of non-negative reals $\{a_t\}_{t=1}^n$, it holds that*

$$\sum_{j=1}^n \frac{a_j}{\sqrt{c + \sum_{i=1}^j a_i}} \leq 2\sqrt{c + \sum_{j=1}^n a_j}.$$

Proof. The proof is by induction. For $n = 1$, the claim holds trivially. Fix an integer $n > 1$. Suppose now that the claim holds for $n - 1$. It follows that

$$\begin{aligned}\sum_{j=1}^n \frac{a_j}{\sqrt{c + \sum_{i=1}^j a_i}} &\leq 2\sqrt{c + \sum_{j=1}^{n-1} a_j} + \frac{a_n}{\sqrt{c + \sum_{i=1}^n a_i}} \\ &=: 2\sqrt{c + Z - x} + \frac{x}{\sqrt{c + Z}} =: g(x),\end{aligned}$$

where $Z := \sum_{i=1}^n a_i$ and $x := a_n$. Observe that $\mathrm{d}g(x)/\mathrm{d}x < 0$, for all $x > 0$. (In simple words, the function $g(x)$ is strictly decreasing, for all $x > 0$.) As a result, we have

$$\operatorname*{argmax}_{x \geq 0}\; g(x) = 0,$$

and $g(x) \leq 2\sqrt{c + Z}$. This concludes the proof. □

Lemma 8. *Given two positive reals a and b, it holds that*

$$b\left(\frac{1}{b} - \frac{1}{a}\right) \leq \log\left(\frac{b^{-1}}{a^{-1}}\right).$$

Proof. Let us first recall the identity $\log(\xi) \leq \xi - 1$, for any $\xi > 0$. Set $\xi = a^{-1}/b^{-1}$. Notice that

$$-\log\left(\frac{b^{-1}}{a^{-1}}\right) = \log\left(\frac{a^{-1}}{b^{-1}}\right) \leq \frac{a^{-1}}{b^{-1}} - 1 = b\left(\frac{1}{a} - \frac{1}{b}\right).$$

Thus, the claim is an immediate consequence of the above relation. □

Appendix B

Technical Proofs

B-1 Proof of Theorem 1

In Theorem 3, set $u_t = \text{argmin}_{x\in\mathcal{X}} \sum_{\tau=1}^{T} f_\tau(x)$ and let Φ_t be the the identity mapping (i.e., $\Phi_t(x) = x$), for all $t \in [T]$. The claim immediately follows. □

B-2 Proof of Theorem 2

In order to prove Theorem 2, first we will prove a version of this theorem for general Bregman divergences (Lemma 9). This result is achieved by exploiting a certain technical assumption (Assumption 6). Then, we will show that for the euclidean case (i.e. $\mathcal{B}_h(x,y) = \frac{1}{2}\|x-y\|_2^2$), this technical assumption always holds and Theorem 2 follows.

Assumption 6 (Technical assumption)**.** *For $\eta_t = \left(\frac{\lambda}{\sigma^2}D'_{t-1} + 2\beta\right)^{-1}$ and $\beta \geq \alpha > 0$, there exists a constant $\lambda > 0$ such that $\lambda\mathcal{B}_h(x^*, y_t) - \frac{1}{\eta_t}\mathcal{B}_h(y_t, x_t) - \frac{\alpha}{2}\|x^* - x_t\|^2 \leq 0$ for all $t > 0$.*

Before stating the general version of Theorem 2, we make a short remark on Assumption 6.

Remark 9 (Mildness of Assumption 6)**.** *Notice that η_t, α, and $\mathcal{B}_h(x,y)$ are all non-negative, for all t and $x, y \in \mathcal{X}$. Moreover, $\{\eta_t\}_{t=1}^T$ is a non-increasing sequence. Then, for a general choice of h, one should be able to choose a small enough λ to ensure that the inequality in Assumption 6 holds for all $t > 0$. In particular, when $\mathcal{B}_h(x,y) = \frac{1}{2}\|x-y\|_2^2$, we will show that Assumption 6 holds for $\lambda = \alpha/2$.*

Lemma 9 (Static regret: Strongly convex case with general divergence)**.** *Suppose that Assumptions 1 and 6 hold and that the cost sequence $\{f_t\}_{t=1}^T$ is α-strongly convex. Using the adaptive step-size $\eta_1 = \frac{1}{2\beta}$ and $\eta_t = \left(\frac{\lambda}{\sigma^2}D'_{t-1} + 2\beta\right)^{-1}$ for all $t > 1$, Algorithm* (OptMD) *guarantees*

$$\textbf{Reg}_T^s \leq 2\beta R^2 + \frac{\sigma^2}{\beta} + \frac{2\sigma^2}{\lambda}\log\left(1 + \frac{\lambda}{2\beta\sigma^2}D'_T\right).$$

Proof. Recall that $x^* = \text{argmin}_{x\in\mathcal{X}} \sum_{t=1}^T f_t(x)$ is the optimal action in hindsight. Since f_t is α-strongly convex, we have

$$f_t(x_t) - f_t(x^*) \leq \langle \nabla f_t(x_t), x_t - x^* \rangle - \frac{\alpha}{2}\|x^* - x_t\|^2.$$

We next follow the same steps taken in the proof of Theorem 3 to derive equation (B-6) and arrive at

$$\begin{aligned} f_t(x_t) - f_t(x^*) \leq \langle \nabla f_t(x_t) - M_t, x_t - y_t \rangle - \frac{\alpha}{2}\|x^* - x_t\|^2 \\ + \frac{1}{\eta_t}(\mathcal{B}_h(x^*, y_{t-1}) - \mathcal{B}_h(x^*, y_t) - \mathcal{B}_h(y_t, x_t) - \mathcal{B}_h(x_t, y_{t-1})). \end{aligned}$$

Define

$$\begin{aligned} A_t &:= \frac{1}{\eta_t}\left(\mathcal{B}_h(x^*, y_{t-1}) - \mathcal{B}_h(x^*, y_t) - \mathcal{B}_h(y_t, x_t)\right) - \frac{\alpha}{2}\|x^* - x_t\|^2 \\ B_t &:= \langle \nabla f_t(x_t) - M_t, x_t - y_t \rangle - \frac{1}{\eta_t}\mathcal{B}_h(x_t, y_{t-1}). \end{aligned}$$

With the above notations at hand, it follows that

$$\mathbf{Reg}_T^s = \sum_{t=1}^T (f_t(x_t) - f_t(x^*)) \leq \sum_{t=1}^T A_t + \sum_{t=1}^T B_t. \tag{B-1}$$

We proceed by bounding $\sum_{t=1}^T A_t$ and $\sum_{t=1}^T B_t$, separately.

(Upper-bounding $\sum_{t=1}^T A_t$) Observe that

$$\begin{aligned} \sum_{t=1}^T A_t &= \sum_{t=1}^T \left(\frac{1}{\eta_t}\left(\mathcal{B}_h(x^*, y_{t-1}) - \mathcal{B}_h(x^*, y_t)\right)\right) - \sum_{t=1}^T \left(\frac{1}{\eta_t}\mathcal{B}_h(y_t, x_t) + \frac{\alpha}{2}\|x^* - x_t\|^2\right) \\ &\leq \frac{\mathcal{B}_h(x^*, y_0)}{\eta_1} + \sum_{t=1}^T \left(\left(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t}\right)\mathcal{B}_h(x^*, y_t)\right) - \sum_{t=1}^T \left(\frac{1}{\eta_t}\mathcal{B}_h(y_t, x_t) + \frac{\alpha}{2}\|x^* - x_t\|^2\right). \end{aligned}$$

Recall that $\eta_1 = \frac{1}{2\beta}$ in Theorem 2. The fourth item in Assumption 1 implies that

$$\frac{\mathcal{B}_h(x^*, y_0)}{\eta_1} \leq 2\beta R^2. \tag{B-2a}$$

Based on the definition of η_t in Theorem 2, we get

$$\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t} = \frac{\lambda}{\sigma^2}\|\nabla f_t(y_{t-1}) - M_t\|_*^2, \tag{B-2b}$$

where scalars λ and σ are explained in Assumptions 6 and 1, respectively. Hence, we obtain

$$\begin{aligned} \sum_{t=1}^T \left(\left(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t}\right)\mathcal{B}_h(x^*, y_t)\right) &= \lambda \sum_{t=1}^T \left(\frac{\|\nabla f_t(y_{t-1}) - M_t\|_*^2}{\sigma^2}\mathcal{B}_h(x^*, y_t)\right) \\ &\leq \lambda \sum_{t=1}^T \mathcal{B}_h(x^*, y_t), \end{aligned} \tag{B-2c}$$

where the inequality follows from the fifth item in Assumption 1. In light of the upper-bounds derived in equation (B-2), we then infer that

$$\sum_{t=1}^{T} A_t \leq 2\beta R^2 + \sum_{t=1}^{T}\left(\lambda\mathcal{B}_h(x^*, y_t) - \frac{1}{\eta_t}\mathcal{B}_h(y_t, x_t) + \frac{\alpha}{2}\|x^* - x_t\|^2\right) \leq 2\beta R^2, \tag{B-3}$$

where the second inequality follows from Assumption 6.

(Upper-bounding $\sum_{t=1}^{T} B_t$) Following the same steps taken in the proof of Theorem 3 to obtain equation (B-10), we arrive at

$$\sum_{t=1}^{T} B_t \leq \frac{\sigma^2}{\beta} + 2\sum_{t=1}^{T}\left(\eta_{t+1}\|\nabla f_t(y_{t-1}) - M_t\|_*^2\right).$$

Notice that by the definition of the adaptive step size in Theorem 2, we have

$$\|\nabla f_t(y_{t-1}) - M_t\|_*^2 = \frac{\sigma^2}{\lambda}\left(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t}\right),$$

and as a result,

$$\sum_{t=1}^{T} B_t \leq \frac{\sigma^2}{\beta} + \frac{2\sigma^2}{\lambda}\sum_{t=1}^{T}\left(\eta_{t+1}\left(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t}\right)\right).$$

We next use Lemma 8 to upper-bound the right-hand side of the above inequality and arrive at

$$\begin{aligned}\sum_{t=1}^{T} B_t \leq \frac{\sigma^2}{\beta} + \frac{2\sigma^2}{\lambda}\sum_{t=1}^{T}\log\left(\frac{\eta_{t+1}^{-1}}{\eta_t^{-1}}\right) &= \frac{\sigma^2}{\beta} + \frac{2\sigma^2}{\lambda}\left(\log\left(\frac{1}{\eta_{T+1}}\right) - \log\left(\frac{1}{\eta_1}\right)\right)\\ &= \frac{\sigma^2}{\beta} + \frac{2\sigma^2}{\lambda}\log\left(\frac{\eta_1}{\eta_{T+1}}\right).\end{aligned}$$

We now use use the step-size rule defined in Theorem 2 to upper-bound the right-hand side of the last equality. By doing so, we get

$$\sum_{t=1}^{T} B_t \leq \frac{\sigma^2}{\beta} + \frac{2\sigma^2}{\lambda}\log\left(1 + \frac{\lambda}{2\beta\sigma^2}D_T'\right), \tag{B-4}$$

where D_T' is defined in (2-5).

(Regret upper-bound) In light of equations (B-1), (B-3) and (B-4), we get

$$\mathbf{Reg}_T^s \leq 2\beta R^2 + \frac{\sigma^2}{\beta} + \frac{2\sigma^2}{\lambda}\log\left(1 + \frac{\lambda}{2\beta\sigma^2}D_T'\right).$$

The claim immediately follows. □

Finally, for the euclidean case (i.e. $\mathcal{B}_h(x, y) = \frac{1}{2}\|x - y\|_2^2$) and choosing $\lambda = \alpha/2$, we have

$$\begin{aligned}\lambda\mathcal{B}_h(x^*, y_t) - \frac{1}{\eta_t}\mathcal{B}_h(y_t, x_t) - \alpha\mathcal{B}_h(x^*, x_t) &= \frac{\alpha}{4}\|x^* - y_t\|_2^2 - \frac{1}{2\eta_t}\|y_t - x_t\|_2^2 - \frac{\alpha}{2}\|x^* - x_t\|\\ &\leq \frac{\alpha}{2}\|x^* - x_t\|_2^2 + \frac{\alpha}{2}\|x_t - y_t\|_2^2 - \frac{1}{2\eta_t}\|y_t - x_t\|_2^2 - \frac{\alpha}{2}\|x^* - x_t\|_2^2\\ &\leq 0,\end{aligned}$$

where the second inequality follows from $\|a-b\|^2 \leq 2\|a-c\|^2 + 2\|c-b\|^2$ and the third inequality follows from $\eta_t^{-1} \geq \beta \geq \alpha$. Thus, we have shown that Assumption 6 holds for all t, and Theorem 2 follows from Lemma 9. □

B-3 Proof of Theorem 3

Let $u_t \in \mathcal{X}$. Since f_t is convex, it holds that

$$f_t(x_t) - f_t(u_t) \leq \langle \nabla f_t(x_t), x_t - u_t \rangle.$$

We add $\pm\langle \nabla f_t(x_t), \tilde{y}_t \rangle$ to the right-hand side of the above inequality. So,

$$f_t(x_t) - f_t(u_t) \leq \langle \nabla f_t(x_t), x_t - \tilde{y}_t \rangle + \langle \nabla f_t(x_t), \tilde{y}_t - u_t \rangle.$$

We next add $\pm\langle M_t, x_t - \tilde{y}_t \rangle$ to the right-hand side of the last inequality and arrive at

$$f_t(x_t) - f_t(u_t) \leq \langle \nabla f_t(x_t) - M_t, x_t - \tilde{y}_t \rangle + \langle M_t, x_t - \tilde{y}_t \rangle + \langle \nabla f_t(x_t), \tilde{y}_t - u_t \rangle. \tag{B-5}$$

In Algorithm (OptDMD), recall that

$$\begin{aligned} x_t &= \operatorname*{argmin}_{x \in \mathcal{X}} \ \{\eta_t \langle M_t, x \rangle + \mathcal{B}_h(x, y_{t-1})\} \\ \tilde{y}_t &= \operatorname*{argmin}_{y \in \mathcal{X}} \ \{\eta_t \langle \nabla f_t(x_t), y \rangle + \mathcal{B}_h(y, y_{t-1})\}. \end{aligned}$$

It follows from Lemma 5 with $w = M_t$, $u = x_t$, $z = \tilde{y}_t$, $v = y_{t-1}$, and $\eta = \eta_t$ that

$$\langle M_t, x_t - \tilde{y}_t \rangle \leq \frac{1}{\eta_t}(\mathcal{B}_h(\tilde{y}_t, y_{t-1}) - \mathcal{B}_h(\tilde{y}_t, x_t) - \mathcal{B}_h(x_t, y_{t-1})).$$

Similarly, Lemma 5 with $w = \nabla f_t(x_t)$, $u = \tilde{y}_t$, $z = u_t$, $v = y_{t-1}$, and $\eta = \eta_t$ implies that

$$\langle \nabla f_t(x_t), \tilde{y}_t - u_t \rangle \leq \frac{1}{\eta_t}(\mathcal{B}_h(u_t, y_{t-1}) - \mathcal{B}_h(u_t, \tilde{y}_t) - \mathcal{B}_h(\tilde{y}_t, y_{t-1})).$$

In light of the last two inequalities, the right-hand side of inequality (B-5) can be upper-bounded as follows

$$\begin{aligned} &f_t(x_t) - f_t(u_t) \\ &\leq \langle \nabla f_t(x_t) - M_t, x_t - \tilde{y}_t \rangle + \frac{1}{\eta_t}(\mathcal{B}_h(u_t, y_{t-1}) - \mathcal{B}_h(u_t, \tilde{y}_t) - \mathcal{B}_h(\tilde{y}_t, x_t) - \mathcal{B}_h(x_t, y_{t-1})) \\ &\leq \langle \nabla f_t(x_t) - M_t, x_t - \tilde{y}_t \rangle + \frac{1}{\eta_t}(\mathcal{B}_h(u_t, y_{t-1}) - \mathcal{B}_h(u_t, \tilde{y}_t) - \mathcal{B}_h(x_t, y_{t-1})), \end{aligned} \tag{B-6}$$

where we made use of the fact that $\mathcal{B}_h(\tilde{y}_t, x_t) \geq 0$ in the second inequality. Define now

$$\begin{aligned} A_t &:= \frac{1}{\eta_t}(\mathcal{B}_h(u_t, y_{t-1}) - \mathcal{B}_h(u_t, \tilde{y}_t)) \\ B_t &:= \langle \nabla f_t(x_t) - M_t, x_t - \tilde{y}_t \rangle - \frac{1}{\eta_t}\mathcal{B}_h(x_t, y_{t-1}). \end{aligned} \tag{B-7}$$

Thus, we have

$$\mathbf{Reg}_T^d = \sum_{t=1}^T (f_t(x_t) - f_t(u_t)) \leq \sum_{t=1}^T A_t + \sum_{t=1}^T B_t. \tag{B-8}$$

In what follows, we seek to upper-bound the terms $\sum_{t=1}^T A_t$ and $\sum_{t=1}^T B_t$ in order to prove the theorem's claim.

(Upper-bounding $\sum_{t=1}^T A_t$) Adding $\pm\frac{1}{\eta_t}\mathcal{B}_h(u_{t+1}, y_t)$ and $\pm\frac{1}{\eta_t}\mathcal{B}_h(\Phi_t(u_t), y_t)$ to A_t and summing the result over $t = 1, \ldots, T$, we get

$$\sum_{t=1}^T A_t = \sum_{t=1}^T \Big(\frac{1}{\eta_t}\Big(\mathcal{B}_h(u_t, y_{t-1}) - \mathcal{B}_h(u_{t+1}, y_t) + \mathcal{B}_h(u_{t+1}, y_t) \\ - \mathcal{B}_h(\Phi_t(u_t), y_t) + \mathcal{B}_h(\Phi_t(u_t), \Phi_t(\tilde{y}_t)) - \mathcal{B}_h(u_t, \tilde{y}_t)\Big)\Big),$$

where we made use of $y_t = \Phi_t(\tilde{y}_t)$ (see Algorithm (OptDMD)). By Assumption 3, it holds for some positive real γ that

$$\mathcal{B}_h(u_{t+1}, y_t) - \mathcal{B}_h(\Phi_t(u_t), y_t) \leq \gamma\|u_{t+1} - \Phi_t(u_t)\|.$$

By Assumption 4, it further holds that

$$\mathcal{B}_h(\Phi_t(u_t), \Phi_t(\tilde{y}_t)) - \mathcal{B}_h(u_t, \tilde{y}_t) \leq 0.$$

By virtue of the last two inequalities, we arrive at

$$\sum_{t=1}^T A_t \leq \sum_{t=1}^T \Big(\frac{1}{\eta_t}(\mathcal{B}_h(u_t, y_{t-1}) - \mathcal{B}_h(u_{t+1}, y_t) + \gamma\|u_{t+1} - \Phi_t(u_t)\|)\Big). \tag{B-9}$$

Recall the definition of η_t in Theorem 3 and observe that

$$\begin{aligned}
&\sum_{t=1}^T \left(\frac{1}{\eta_t}(\mathcal{B}_h(u_t, y_{t-1}) - \mathcal{B}_h(u_{t+1}, y_t))\right) \\
&\quad \leq \frac{1}{\eta_1}\mathcal{B}_h(u_1, y_0) - \frac{1}{\eta_T}\mathcal{B}_h(u_{T+1}, y_T) + \sum_{t=2}^T \left(\mathcal{B}_h(u_t, y_{t-1})\left(\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}}\right)\right) \\
&\quad \overset{\text{(i)}}{\leq} \frac{R^2}{\eta_T} + R^2\sum_{t=2}^T \left(\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}}\right) \\
&\quad = \frac{R^2}{\eta_T} + R^2\left(\frac{1}{\eta_T} - \frac{1}{\eta_1}\right) \\
&\quad \overset{\text{(ii)}}{\leq} \frac{2R^2}{\eta_T},
\end{aligned}$$

where we made use of the facts that $\eta_T \leq \eta_1$, $\mathcal{B}_h(u_{T+1}, y_T) \geq 0$, and of the fourth item in Assumption 1 in the inequality (i), and the inequality (ii) is implied by $\eta_1 > 0$, respectively. Considering inequality (B-9), one can conclude based on the above arguments that

$$\sum_{t=1}^T A_t \leq \frac{2R^2}{\eta_T} + \sum_{t=1}^T \left(\frac{\gamma}{\eta_t}\|u_{t+1} - \Phi_t(u_t)\|\right) \leq \frac{2R^2}{\eta_T} + \frac{\gamma}{\eta_T}\sum_{t=1}^T \|u_{t+1} - \Phi_t(u_t)\| = \frac{1}{\eta_T}(2R^2 + \gamma C_T'),$$

where the second inequality follows from $\eta_t \geq \eta_{t+1}$ and C'_T is defined in equation (2-4). Lastly, we use the step-size η_T given in Theorem 3 to obtain

$$\sum_{t=1}^{T} A_t \leq \left(2R^2 + \gamma C'_T\right) \sqrt{D'_T + 4\beta^2}, \tag{B-10}$$

where D'_T is defined in equation (2-5).

(Upper-bounding $\sum_{t=1}^{T} B_t$) We proceed by bounding $\sum_{t=1}^{T} B_t$ in the sequel, where B_t is given in (B-7). The Cauchy–Schwarz inequality implies that

$$B_t \leq \|\nabla f_t(x_t) - M_t\|_* \|x_t - \tilde{y}_t\| - \frac{1}{\eta_t}\mathcal{B}_h(x_t, y_{t-1}).$$

We now employ Lemma 6 to provide an upper-bound on $\|x_t - \tilde{y}_t\|$. Recall that x_t and $\tilde{y}_t$ are constructed based on Algorithm (OptDMD). Take $\eta = \eta_t$, $u_1 = \tilde{y}_t$, $w_1 = \nabla f_t(x_t)$, $u_2 = x_t$, and $w_2 = M_t$ in Lemma 6. So, we arrive at

$$\|\tilde{y}_t - x_t\| \leq \eta_t \|\nabla f_t(x_t) - M_t\|_*,$$

and as a result,

$$B_t \leq \eta_t \|\nabla f_t(x_t) - M_t\|_*^2 - \frac{1}{\eta_t}\mathcal{B}_h(x_t, y_{t-1}).$$

Notice that

$$\begin{aligned}
B_t &\leq 2\eta_t \|\nabla f_t(x_t) - \nabla f_t(y_{t-1})\|_*^2 + 2\eta_t \|\nabla f_t(y_{t-1}) - M_t\|_*^2 - \frac{1}{\eta_t}\mathcal{B}_h(x_t, y_{t-1}) \\
&\leq 2\eta_t \beta^2 \|x_t - y_{t-1}\|^2 + 2\eta_t \|\nabla f_t(y_{t-1}) - M_t\|_*^2 - \frac{1}{\eta_t}\mathcal{B}_h(x_t, y_{t-1}).
\end{aligned}$$

where we made use of the identity $\|a-b\|^2 \leq 2\|a-c\|^2 + 2\|c-b\|^2$ in the first inequality, and the β-smoothness of f_t implies the second inequality. Recall now that h is 1-strongly convex w.r.t. the norm $\|\cdot\|$ (see Assumption 1, second item). We employ Lemma 1 to conclude that

$$B_t \leq \left(2\eta_t \beta^2 - \frac{1}{2\eta_t}\right) \|x_t - y_{t-1}\|^2 + 2\eta_t \|\nabla f_t(y_{t-1}) - M_t\|_*^2.$$

Based on the definition of η_t in Theorem 3, the sequence $\{\eta_t\}_{t=1}^{T}$ is a non-increasing sequence. Thus, we have $\eta_t \leq 1/(2\beta)$, and as a result,

$$2\eta_t \beta^2 - \frac{1}{2\eta_t} \leq 0, \ \ \forall t \geq 1.$$

In view of this observation, we get

$$B_t \leq 2\eta_t \|\nabla f_t(y_{t-1}) - M_t\|_*^2.$$

Summing B_t over $t = 1, \ldots, T$ yields

$$\begin{aligned}
\sum_{t=1}^{T} B_t &\leq 2\sum_{t=1}^{T} \left(\eta_t \|\nabla f_t(y_{t-1}) - M_t\|_*^2\right) \\
&\leq 2\sum_{t=1}^{T} \left(\eta_{t+1} \|\nabla f_t(y_{t-1}) - M_t\|_*^2\right) + 2\sum_{t=1}^{T} \left((\eta_t - \eta_{t+1}) \|\nabla f_t(y_{t-1}) - M_t\|_*^2\right).
\end{aligned}$$

Recall the definition of η_t in Theorem 3. Set $n = T$, $a_t = \|\nabla f_t(y_{t-1}) - M_t\|_*^2$, and $c = 4\beta^2$ in Lemma 7. Hence, we get

$$\sum_{t=1}^{T} \left(\eta_{t+1}\|\nabla f_t(y_{t-1}) - M_t\|_*^2\right) \leq 2\sqrt{4\beta^2 + D'_T},$$

where D'_T is defined in (2-5). By virtue of the fifth item in Assumption 1, it also follows that

$$\begin{aligned}\sum_{t=1}^{T} \left((\eta_t - \eta_{t+1})\|\nabla f_t(y_{t-1}) - M_t\|_*^2\right) &\leq \sigma^2 \sum_{t=1}^{T} (\eta_t - \eta_{t+1}) \\ &= \sigma^2 (\eta_1 - \eta_{T+1}) \leq \sigma^2 \eta_1 = \frac{\sigma^2}{2\beta}.\end{aligned}$$

Based on the above analyses, it is straightforward to see that

$$\sum_{t=1}^{T} B_t \leq 4\sqrt{D'_T + 4\beta^2} + \frac{\sigma^2}{\beta}. \tag{B-11}$$

(Regret upper-bound) Considering equations (B-8), (B-10), and (B-11), it holds that

$$\mathbf{Reg}_T^d \leq \left(2R^2 + \gamma C'_T\right)\sqrt{D'_T + 4\beta^2} + 4\sqrt{D'_T + 4\beta^2} + \frac{\sigma^2}{\beta}.$$

This concludes the proof. □

B-4 Proof of Theorem 4

Recall that $x_t^* = \text{argmin}_{x \in \mathcal{X}} f_t(x)$. Since f_t is convex,

$$f_t(x_t) - f_t(x_t^*) \leq \langle \nabla f_t(x_t), x_t - x_t^* \rangle.$$

Adding $\pm\langle \nabla f_t(x_t), y_{t-1}\rangle$ and $\pm\langle M_t, x_t - y_{t-1}\rangle$ to the right-hand side of the above inequality, we get

$$f_t(x_t) - f_t(x_t^*) \leq \langle \nabla f_t(x_t) - M_t, x_t - y_{t-1}\rangle + \langle M_t, x_t - y_{t-1}\rangle + \langle \nabla f_t(x_t), y_{t-1} - x_t^*\rangle.$$

Let us now mention that $\mathcal{B}(p,q) = \frac{1}{2}\|p - q\|_2^2$, for all $p, q \in \mathcal{X}$. Invoking Lemma 5 with $\eta = \eta_t$, $w = M_t$, $u = x_t$, and $z = v = y_{t-1}$, we get

$$\langle M_t, x_t - y_{t-1}\rangle \leq -\frac{1}{\eta_t}\|x_t - y_{t-1}\|_2^2,$$

and so,

$$f_t(x_t) - f_t(x_t^*) \leq \langle \nabla f_t(x_t) - M_t, x_t - y_{t-1}\rangle - \frac{1}{\eta_t}\|x_t - y_{t-1}\|_2^2 + \langle \nabla f_t(x_t), y_{t-1} - x_t^*\rangle.$$

In Algorithm (OptDMD-mod), recall that

$$\tilde{y}_t = \underset{y \in \mathcal{X}}{\text{argmin}} \ \{\frac{1}{\omega}\langle \nabla f_t(x_t), y\rangle + \frac{1}{2}\|y - y_{t-1}\|_2^2\}.$$

We now use Lemma 4 with $w = \nabla f_t(x_t)$, $v = y_{t-1}$, $z = x_t^*$, $\eta = 1/\omega$ and $u = \tilde{y}_t$ to obtain

$$\langle \nabla f_t(x_t), y_{t-1} - x_t^* \rangle \leq \frac{\omega}{2}\|y_{t-1} - x_t^*\|_2^2 - \frac{\omega}{2}\|\tilde{y}_t - x_t^*\|_2^2 + \frac{1}{2\omega}\|\nabla f_t(x_t)\|_2^2,$$

and as a result,

$$\begin{aligned} f_t(x_t) - f_t(x_t^*) \leq \langle \nabla f_t(x_t) - M_t, x_t - y_{t-1} \rangle - \frac{1}{\eta_t}\|x_t - y_{t-1}\|_2^2 \\ + \frac{\omega}{2}\|y_{t-1} - x_t^*\|_2^2 - \frac{\omega}{2}\|\tilde{y}_t - x_t^*\|_2^2 + \frac{1}{2\omega}\|\nabla f_t(x_t)\|_2^2. \end{aligned}$$

For the sake of notational simplicity, define

$$\begin{aligned} A_t &:= \frac{\omega}{2}\|x_t^* - y_{t-1}\|_2^2 - \frac{\omega}{2}\|x_t^* - \tilde{y}_t\|_2^2 \\ B_t &:= \langle \nabla f_t(x_t) - M_t, x_t - y_{t-1} \rangle - \frac{1}{\eta_t}\|x_t - y_{t-1}\|_2^2. \end{aligned} \tag{B-12}$$

With the above notations in mind, we have

$$f_t(x_t) - f_t(x_t^*) \leq A_t + B_t + \frac{1}{2\omega}\|\nabla f_t(x_t)\|_2^2,$$

and so

$$\mathbf{Reg}_T^d = \sum_{t=1}^T (f_t(x_t) - f_t(x_t^*)) \leq \sum_{t=1}^T A_t + \sum_{t=1}^T B_t + \frac{1}{2\omega}\sum_{t=1}^T \|\nabla f_t(x_t)\|_2^2. \tag{B-13}$$

(Upper-bounding $\sum_{t=1}^T B_t$) Let us now find an upper-bound on B_t. Add $\pm\langle \nabla f_t(y_{t-1}), x_t - y_{t-1} \rangle$ to B_t. A straightforward application of the Cauchy-Schwartz inequality implies that

$$\begin{aligned} B_t \leq \|\nabla f_t(x_t) - \nabla f_t(y_{t-1})\|_2\|x_t - y_{t-1}\|_2 \\ + \|\nabla f_t(y_{t-1}) - M_t\|_2\|x_t - y_{t-1}\|_2 - \frac{1}{\eta_t}\|x_t - y_{t-1}\|_2^2. \end{aligned}$$

Since f_t is β-smooth w.r.t. $\|\cdot\|_2$, it holds that

$$B_t \leq \left(\beta - \frac{1}{2\eta_t}\right)\|x_t - y_{t-1}\|_2^2 + \|\nabla f_t(y_{t-1}) - M_t\|_2\|x_t - y_{t-1}\|_2 - \frac{1}{2\eta_t}\|x_t - y_{t-1}\|_2^2.$$

Using the identity $2ab - a^2 \leq b^2$ yields

$$\|\nabla f_t(y_{t-1}) - M_t\|_2\|x_t - y_{t-1}\|_2 - \frac{1}{2\eta_t}\|x_t - y_{t-1}\|_2^2 \leq \frac{\eta_t}{2}\|\nabla f_t(y_{t-1}) - M_t\|_2^2,$$

and thus, it follows that

$$\sum_{t=1}^T B_t \leq \sum_{t=1}^T \left(\left(\beta - \frac{1}{2\eta_t}\right)\|x_t - y_{t-1}\|_2^2\right) + \frac{1}{2}\sum_{t=1}^T \left(\eta_t\|\nabla f_t(y_{t-1}) - M_t\|_2^2\right). \tag{B-14a}$$

Recall that $\eta_1 = 1/(3\beta)$. One can conclude that

$$
\begin{aligned}
&\frac{1}{2}\sum_{t=1}^{T}\left(\eta_t\|\nabla f_t(y_{t-1}) - M_t\|_2^2\right) \\
&\quad = \frac{1}{2}\sum_{t=1}^{T}\left((\eta_t - \eta_{t+1})\|\nabla f_t(y_{t-1}) - M_t\|_2^2\right) + \frac{1}{2}\sum_{t=1}^{T}\left(\eta_{t+1}\|\nabla f_t(y_{t-1}) - M_t\|_2^2\right) \\
&\quad \overset{\text{(i)}}{\leq} \frac{\sigma^2}{2}\sum_{t=1}^{T}(\eta_t - \eta_{t+1}) + \frac{1}{2}\sum_{t=1}^{T}\left(\eta_{t+1}\|\nabla f_t(y_{t-1}) - M_t\|_2^2\right) \qquad \text{(B-14b)} \\
&\quad = \frac{\sigma^2}{2}(\eta_1 - \eta_{T+1}) + \frac{1}{2}\sum_{t=1}^{T}\left(\eta_{t+1}\|\nabla f_t(y_{t-1}) - M_t\|_2^2\right) \\
&\quad \overset{\text{(ii)}}{\leq} \frac{\sigma^2}{6\beta} + \frac{1}{2}\sum_{t=1}^{T}\left(\eta_{t+1}\|\nabla f_t(y_{t-1}) - M_t\|_2^2\right),
\end{aligned}
$$

where the fifth item in Assumption 1 implies the inequality (i) and we used the fact that $\eta_{T+1} > 0$ in the inequality (ii), respectively. By the definition of the adaptive step size in Theorem 4, we have

$$
\|\nabla f_t(y_{t-1}) - M_t\|_2^2 = \frac{1}{\eta_{t+1}} - \frac{1}{\eta_t},
$$

and thus,

$$
\sum_{t=1}^{T}\left(\eta_{t+1}\|\nabla f_t(y_{t-1}) - M_t\|_2^2\right) = \sum_{t=1}^{T}\eta_{t+1}\left(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t}\right).
$$

Recall that D'_t is defined in (2-5). Observe that

$$
\sum_{t=1}^{T}\eta_{t+1}\left(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t}\right) \overset{\text{(i)}}{\leq} \sum_{t=1}^{T}\log\left(\frac{\eta_{t+1}^{-1}}{\eta_t^{-1}}\right) = \log\left(\frac{\eta_1}{\eta_{T+1}}\right) \overset{\text{(ii)}}{=} \log\left(1 + \frac{D'_T}{3\beta}\right),
$$

where we used Lemma 8 in the inequality (i) and the definition of step-size in Theorem 4 equality (ii). In light of the relations in equation (B-14), we infer that

$$
\sum_{t=1}^{T} B_t \leq \sum_{t=1}^{T}\left(\left(\beta - \frac{1}{2\eta_t}\right)\|x_t - y_{t-1}\|_2^2\right) + \frac{\sigma^2}{6\beta} + \frac{1}{2}\log\left(1 + \frac{D'_T}{2\beta}\right). \qquad \text{(B-15)}
$$

In what follows, we proceed to upper-bound the term $\sum_{t=1}^{T}(A_t + B_t)$.

(Upper-bounding $\sum_{t=1}^{T}(A_t + B_t)$) From equation (B-12), we have

$$
\sum_{t=1}^{T} A_t = \frac{\omega}{2}\sum_{t=1}^{T}(\|x_t^* - y_{t-1}\|_2^2 - \|x_t^* - \tilde{y}_t\|_2^2). \qquad \text{(B-16)}
$$

Adding $\pm\|x_{t+1}^* - y_t\|_2^2$ and $\pm\|\Phi_t(x_t^*) - y_t\|_2^2$ to each A_t, we arrive at

$$
\begin{aligned}
\sum_{t=1}^{T} A_t = \frac{\omega}{2}\sum_{t=1}^{T}\Big(&\|x_t^* - y_{t-1}\|_2^2 - \|x_{t+1}^* - y_t\|_2^2 + \|x_{t+1}^* - y_t\|_2^2 - \|\Phi_t(x_t^*) - y_t\|_2^2 \\
&+ \|\Phi_t(x_t^*) - \Phi_t(\tilde{y}_t)\|_2^2 - \|x_t^* - \tilde{y}_t\|_2^2\Big),
\end{aligned}
$$

where we used $y_t = \Phi_t(\tilde{y}_t)$. One can see that

$$\sum_{t=1}^{T}(\|x_t^* - y_{t-1}\|_2^2 - \|x_{t+1}^* - y_t\|_2^2) = \|x_1^* - y_0\|_2^2 - \|x_{T+1}^* - y_T\|_2^2 \le \|x_1^* - y_0\|_2^2 \le 2R^2, \tag{B-17a}$$

where the last inequality is implied by the fourth item in Assumption 1. Next, recall that $h(\cdot) = \frac{1}{2}\|\cdot\|_2^2$. Observe that there exists a positive real γ such that

$$\|x_{t+1}^* - y_t\|_2^2 - \|\Phi_t(x_t^*) - y_t\|_2^2 \le \gamma\|x_{t+1}^* - \Phi_t(x_t^*)\|_2, \tag{B-17b}$$

where we made use of Assumptions 3 in the above inequality. Moreover, Assumption 4 implies that

$$\|\Phi_t(x_t^*) - \Phi_t(\tilde{y}_t)\|_2^2 - \|x_t^* - \tilde{y}_t\|_2^2 \le 0, \ \forall t \in [T]. \tag{B-17c}$$

By virtue of equations (B-16) and (B-17), we thus obtain

$$\sum_{t=1}^{T} A_t \le \omega R^2 + \frac{\gamma\omega}{2}\sum_{t=1}^{T}\|x_{t+1}^* - \Phi_t(x_t^*)\|_2. \tag{B-18}$$

It follows from equations (B-15) and (B-18) that

$$\begin{aligned}\sum_{t=1}^{T}(A_t + B_t) &\le \omega R^2 + \frac{\gamma\omega}{2}\sum_{t=1}^{T}\|x_{t+1}^* - \Phi_t(x_t^*)\|_2 \\ &\quad + \sum_{t=1}^{T}\left(\left(\beta - \frac{1}{2\eta_t}\right)\|x_t - y_{t-1}\|_2^2\right) + \frac{\sigma^2}{6\beta} + \frac{1}{2}\log\left(1 + \frac{D_T'}{3\beta}\right) \\ &\le \omega R^2 + \frac{\gamma\omega}{2}\sum_{t=1}^{T}\|x_{t+1}^* - \Phi_t(x_t^*)\|_2 + \frac{\sigma^2}{6\beta} + \frac{1}{2}\log\left(1 + \frac{D_T'}{3\beta}\right),\end{aligned} \tag{B-19}$$

where in the last inequality we made use of the fact that $\beta - \frac{1}{2\eta_t} \le 0$, implied by the definition of η_t. Thus, we arrive at

$$\mathbf{Reg}_T^d \le \frac{\sigma^2}{6\beta} + \omega R^2 + \frac{1}{2}\log\left(1 + \frac{D_T'}{3\beta}\right) + \frac{1}{2\omega}\sum_{t=1}^{T}\|\nabla f_t(x_t)\|_2^2 + \frac{\gamma\omega}{2}\sum_{t=1}^{T}\|x_{t+1}^* - \Phi_t(x_t^*)\|_2. \tag{B-20}$$

To finish the proof, we need to upper-bound $\|\nabla f_t(x_t)\|_2^2$. To do so, first notice that the β-smoothness of f_t imply that $f_t(u) \le f_t(v) + \langle \nabla f_t(v), u - v\rangle + \frac{\beta}{2}\|u - v\|_2^2$. Setting $u = x_t - \frac{1}{\beta}\nabla f_t(x_t)$ and $v = x_t$, we get

$$\frac{1}{2\beta}\|\nabla f_t(x_t)\|_2^2 \le f_t(x_t) - f_t(u). \tag{B-21}$$

On the other hand, from the convexity of f_t, we have

$$f_t(x_t^*) - f_t(u) \le \langle \nabla f_t(x_t^*), x_t^* - u\rangle \implies -f_t(u) \le -f_t(x_t^*), \tag{B-22}$$

where the implication follows from Assumption 5. Combining (B-21) and (B-22), we get

$$\|\nabla f_t(x_t)\|_2^2 \leq 2\beta(f_t(x_t) - f_t(x_t^*)). \tag{B-23}$$

Substituting (B-23) in our regret bound, setting $\omega = 2\beta$ and rearranging, we finally arrive at

$$\mathbf{Reg}_T^d \leq \frac{\sigma^2}{3\beta} + 4\beta R^2 + \log\left(1 + \frac{D'_T}{3\beta}\right) + 2\gamma\beta \sum_{t=1}^{T} \|x_{t+1}^* - \Phi_t(x_t^*)\|_2.$$

The claim of the theorem hence follows. □

B-5 Proof of Theorem 5

First, since for quadratic costs have $\alpha = \beta$, we will use β for both the smoothness and strong convexity parameters. In this section, we begin with providing an auxiliary results that are used in the proof of Theorem 5. This result gives an upper-bound on the summation $\sum_{t=1}^{T} \|x_t^* - y_{t-1}\|_2^2$, where the sequences $\{x_t\}_{t=1}^{T}$ and $\{y_t\}_{t=0}^{T-1}$ are chosen based on Algorithm (OptDMD-mod).

Lemma 10. *Suppose that Assumptions 1 and 4 hold. Consider $\{f_t\}_{t=1}^{T}$ is a sequence of quadratic costs w.r.t. $\|\cdot\|_2$. Moreover, let*

$$\tilde{y}_t = \underset{y \in \mathcal{X}}{\operatorname{argmin}} \left\{ \langle \nabla f_t(x_t), y - y_{t-1} \rangle + \frac{\beta}{2}\|y - y_{t-1}\|_2^2 \right\} \quad \text{and} \quad x_t^* = \underset{x \in \mathcal{X}}{\operatorname{argmin}}\ f_t(x).$$

Then, it holds that

$$\sum_{t=1}^{T} \|x_t^* - y_{t-1}\|_2^2 \leq 2R^2 + 2\sum_{t=2}^{T} \|x_t^* - \Phi_t(x_{t-1}^*)\|_2^2 + \frac{4}{2\beta}\sum_{t=1}^{T} \langle \nabla f_t(x_t) - \nabla f_t(y_{t-1}), x_t^* - \tilde{y}_t \rangle.$$

Let us mention that $\tilde{y}_t$ is the second iterate given in Algorithm (OptDMD-mod), where we multiplied the corresponding objective function by scalar β. The proof is as follows.

Proof. We first use the identity $\|a - c\|_2^2 \leq 2\|a - b\|_2^2 + 2\|b - c\|_2^2$ to arrive at

$$\begin{aligned}
\sum_{t=1}^{T} \|x_t^* - y_{t-1}\|_2^2 &= \|x_1^* - y_0\|_2^2 + \sum_{t=2}^{T} \|x_t^* - y_{t-1}\|_2^2 \\
&\leq \|x_1^* - y_0\|_2^2 + 2\sum_{t=2}^{T} \|x_t^* - \Phi_t(x_{t-1}^*)\|_2^2 + 2\sum_{t=2}^{T} \|\Phi_t(x_{t-1}^*) - y_{t-1}\|_2^2.
\end{aligned}$$

Recall that $y_{t-1} = \Phi_t(\tilde{y}_{t-1})$ in Algorithm (OptDMD-mod). Since we have $h = \frac{1}{2}\|\cdot\|^2$, it holds that $\mathcal{B}_h(p, q) = \frac{1}{2}\|p - q\|_2^2$, for all $p, q \in \mathcal{X}$. Hence, Assumption 4 reads as

$$\|\Phi_t(x_{t-1}^*) - y_{t-1}\|_2 = \|\Phi_t(x_{t-1}^*) - \Phi_t(\tilde{y}_{t-1})\|_2 \leq \|x_{t-1}^* - \tilde{y}_{t-1}\|_2.$$

Based on the above fact, we have

$$\begin{aligned}\sum_{t=1}^{T}\|x_t^*-y_{t-1}\|_2^2 &\leq \|x_1^*-y_0\|_2^2+2\sum_{t=2}^{T}\|x_t^*-\Phi_t(x_{t-1}^*)\|_2^2+2\sum_{t=2}^{T}\|x_{t-1}^*-\tilde{y}_{t-1}\|_2^2\\ &\leq \|x_1^*-y_0\|_2^2+2\sum_{t=2}^{T}\|x_t^*-\Phi_t(x_{t-1}^*)\|_2^2+2\sum_{t=1}^{T}\|x_t^*-\tilde{y}_t\|_2^2.\end{aligned}$$

Recalling again $h(\cdot)=\frac{1}{2}\|\cdot\|_2^2$, the fourth item in Assumption 1 is equivalent to

$$\|x_1^*-y_0\|_2^2\leq 2R^2,$$

and as a result,

$$\sum_{t=1}^{T}\|x_t^*-y_{t-1}\|_2^2\leq 2R^2+2\sum_{t=2}^{T}\|x_t^*-\Phi_t(x_{t-1}^*)\|_2^2+2\sum_{t=1}^{T}\|x_t^*-\tilde{y}_t\|_2^2. \tag{B-24}$$

We next obtain an upper-bound on the term $\sum_{t=1}^{T}\|x_t^*-\tilde{y}_t\|_2^2$. Recall that $\mathcal{B}_h(p,q)=\frac{1}{2}\|p-q\|_2^2$, for all $p,q\in\mathcal{X}$. Take $\eta=\frac{1}{\beta}$, $w=\nabla f_t(x_t)$, $u=\tilde{y}_t$, $z=x_t^*$, and $v=y_{t-1}$ in Lemma 5. It thus holds that

$$\langle\nabla f_t(x_t),\tilde{y}_t-x_t^*\rangle\leq\frac{\beta}{2}\|x_t^*-y_{t-1}\|_2^2-\frac{\beta}{2}\|x_t^*-\tilde{y}_t\|_2^2-\frac{\beta}{2}\|\tilde{y}_t-y_{t-1}\|_2^2.$$

In view of the above inequality, we have

$$\frac{\beta}{2}\|\tilde{y}_t-y_{t-1}\|_2^2\leq\langle\nabla f_t(x_t),x_t^*-\tilde{y}_t\rangle+\frac{\beta}{2}\|x_t^*-y_{t-1}\|_2^2-\frac{\beta}{2}\|x_t^*-\tilde{y}_t\|_2^2. \tag{B-25a}$$

On the other hand, the β-strong convexity of f_t implies that

$$f_t(y_{t-1})\leq f_t(x_t^*)+\langle\nabla f_t(y_{t-1}),y_{t-1}-x_t^*\rangle-\frac{\beta}{2}\|x_t^*-y_{t-1}\|_2^2. \tag{B-25b}$$

Moreover, recall that $x_t^*=\operatorname{argmin}_{x\in\mathcal{X}}\ f_t(x)$, that is x^* is an optimal decision. Coupled with the β-strong convexity of f_t, we hence have

$$f_t(x_t^*)\leq f_t(\tilde{y}_t)-\frac{\beta}{2}\|x_t^*-\tilde{y}_t\|_2^2. \tag{B-25c}$$

Notice that

$$\begin{aligned}f_t(\tilde{y}_t)&\leq f_t(y_{t-1})+\langle\nabla f_t(y_{t-1}),\tilde{y}_t-y_{t-1}\rangle+\frac{\beta}{2}\|\tilde{y}_t-y_{t-1}\|_2^2\\ &\overset{\text{(B-25a)}}{\leq} f_t(y_{t-1})+\langle\nabla f_t(y_{t-1}),\tilde{y}_t-y_{t-1}\rangle+\langle\nabla f_t(x_t),x_t^*-\tilde{y}_t\rangle+\frac{\beta}{2}\|x_t^*-y_{t-1}\|_2^2-\frac{\beta}{2}\|x_t^*-\tilde{y}_t\|_2^2\\ &\overset{\text{(B-25b)}}{\leq} f_t(x_t^*)+\langle\nabla f_t(x_t)-\nabla f_t(y_{t-1}),x_t^*-\tilde{y}_t\rangle-\frac{\beta}{2}\|x_t^*-\tilde{y}_t\|_2^2\\ &\overset{\text{(B-25c)}}{\leq} f_t(\tilde{y}_t)+\langle\nabla f_t(x_t)-\nabla f_t(y_{t-1}),x_t^*-\tilde{y}_t\rangle-\beta\|x_t^*-\tilde{y}_t\|_2^2,\end{aligned}$$

where the first inequality follows from the β-smoothness of f_t. Hence, we get

$$\|x_t^*-\tilde{y}_t\|_2^2\leq\frac{1}{\beta}\langle\nabla f_t(x_t)-\nabla f_t(y_{t-1}),x_t^*-\tilde{y}_t\rangle. \tag{B-26}$$

In light of inequality (B-26), one can infer from inequality (B-24) that

$$\sum_{t=1}^{T} \|x_t^* - y_{t-1}\|_2^2 \leq 2R^2 + 2\sum_{t=2}^{T} \|x_t^* - \Phi(x_{t-1}^*)\|_2^2 + \frac{2}{\beta}\sum_{t=1}^{T} \langle \nabla f_t(x_t) - \nabla f_t(y_{t-1}), x_t^* - \tilde{y}_t \rangle.$$

The claim immediately follows. □

Proof of Theorem 5

Recall that $x_t^* = \operatorname{argmin}_{x\in\mathcal{X}} f_t(x)$. Since f_t is β-strongly convex,

$$f_t(x_t) - f_t(x_t^*) \leq \langle \nabla f_t(x_t), x_t - x_t^* \rangle - \frac{\beta}{2}\|x_t - x_t^*\|_2^2.$$

Notice that since quadratic costs are a special case of convex costs, we can proceed to bound $\langle \nabla f_t(x_t), x_t - x_t^* \rangle$ the same way we did in the proof of Theorem 4. From Equation (B-20), we get

$$\mathbf{Reg}_T^d \leq \frac{\sigma^2}{6\beta} + \beta R^2 + \frac{1}{2}\log\left(1 + \frac{D_T'}{3\beta}\right) + \frac{\gamma\beta}{2}\sum_{t=1}^{T} \|x_{t+1}^* - \Phi_t(x_t^*)\|_2 + H_T, \tag{B-27}$$

where

$$H_T := \sum_{t=1}^{T}\left(\frac{1}{2\beta}\|\nabla f_t(x_t)\|_2^2 - \frac{\beta}{2}\|x_t - x_t^*\|_2^2\right).$$

In what follows, using the same notation/definitions of the proof of Theorem 4, we proceed to upper-bound the term $\sum_{t=1}^{T}(A_t + B_t)$ using a different approach, exploiting the fact that $\alpha = \beta$. We then take the minimum of the upper bounds as the desired upper-bound on $\sum_{t=1}^{T}(A_t + B_t)$.

(Upper-bounding $\sum_{t=1}^{T}(A_t + B_t)$ for quadratic costs) Recall again that A_t is defined in (B-12). We invoke Lemma 10 and obtain

$$\begin{aligned}\sum_{t=1}^{T} A_t &= \frac{\beta}{2}\sum_{t=1}^{T} \|x_t^* - y_{t-1}\|_2^2 - \frac{\beta}{2}\sum_{t=1}^{T} \|x_t^* - \tilde{y}_t\|_2^2 \\ &\leq \beta R^2 + \beta\sum_{t=2}^{T} \|x_t^* - \Phi(x_{t-1}^*)\|_2^2 + \sum_{t=1}^{T} \langle \nabla f_t(x_t) - \nabla f_t(y_{t-1}), x_t^* - \tilde{y}_t \rangle - \frac{\beta}{2}\sum_{t=1}^{T} \|x_t^* - \tilde{y}_t\|_2^2.\end{aligned}$$

Notice that

$$\begin{aligned}\langle \nabla f_t(x_t) - \nabla f_t(y_{t-1}), x_t^* - \tilde{y}_t \rangle - \frac{\beta}{2}\|x_t^* - \tilde{y}_t\|_2^2 &\overset{\text{(i)}}{\leq} \|\nabla f_t(x_t) - \nabla f_t(y_{t-1})\|_2 \|x_t^* - \tilde{y}_t\|_2 - \frac{\beta}{2}\|x_t^* - \tilde{y}_t\|_2^2 \\ &\overset{\text{(ii)}}{\leq} \beta\|x_t - y_{t-1}\|_2 \|x_t^* - \tilde{y}_t\|_2 - \frac{\beta}{2}\|x_t^* - \tilde{y}_t\|_2^2 \\ &\overset{\text{(iii)}}{\leq} \frac{\beta}{2}\|x_t - y_{t-1}\|_2^2,\end{aligned}$$

where the inequalities (i)-(iii) follow from the Cauchy-Schwartz inequality, the β-smoothness of f_t, and the identity $2ab - a^2 \leq b^2$, respectively. In view of the above two arguments, we thus have

$$\sum_{t=1}^{T} A_t \leq \beta R^2 + \beta\sum_{t=2}^{T} \|x_t^* - \Phi(x_{t-1}^*)\|_2^2 + \frac{\beta}{2}\sum_{t=1}^{T} \|x_t - y_{t-1}\|_2^2. \tag{B-28}$$

One can deduce from equations (B-15) and (B-28) that

$$\begin{aligned}\sum_{t=1}^{T}(A_t + B_t) \leq \beta R^2 + \beta \sum_{t=2}^{T} \|x_t^* - \Phi(x_{t-1}^*)\|_2^2 \\ + \sum_{t=1}^{T}\left(\left(\beta - \frac{1}{2\eta_t} + \frac{\beta}{2}\right)\|x_t - y_{t-1}\|_2^2\right) + \frac{\sigma^2}{3\beta} + \frac{1}{2}\log\left(1 + \frac{D_T'}{3\beta}\right) \\ \leq \beta R^2 + \beta \sum_{t=2}^{T} \|x_t^* - \Phi(x_{t-1}^*)\|_2^2 + \frac{\sigma^2}{6\beta} + \frac{1}{2}\log\left(1 + \frac{D_T'}{3\beta}\right),\end{aligned} \tag{B-29}$$

where the second inequality follows from the definition of η_t.

(Regret upper-bound) Considering the upper-bounds provided in equations (B-27) and (B-29), we thus conclude that

$$\mathbf{Reg}_T^d \leq \frac{\sigma^2}{6\beta} + \beta R^2 + \frac{1}{2}\log\left(1 + \frac{D_T'}{3\beta}\right) + H_T + \beta \min\left\{\frac{\gamma}{2}\sum_{t=1}^{T}\|x_{t+1}^* - \Phi_t(x_t^*)\|_2,\ \sum_{t=2}^{T}\|x_t^* - \Phi(x_{t-1}^*)\|_2^2\right\}.$$

Finally, we show that the fact that $\nabla f_t(x_t^*) = 0$ implies that $H_T \leq 0$.

$$\begin{aligned}H_T &= \frac{1}{2\beta}\sum_{t=1}^{T}\|\nabla f_t(x_t)\|_2^2 - \frac{\beta}{2}\sum_{t=1}^{T}\|x_t - x_t^*\|_2^2 \\ &= \frac{1}{2\beta}\sum_{t=1}^{T}\|\nabla f_t(x_t) - \nabla f_t(x_t^*)\|_2^2 - \frac{\beta}{2}\sum_{t=1}^{T}\|x_t - x_t^*\|_2^2 \\ &\leq \frac{\beta}{2}\sum_{t=1}^{T}\|x_t - x_t^*\|_2^2 - \frac{\beta}{2}\sum_{t=1}^{T}\|x_t - x_t^*\|_2^2 \\ &\leq 0.\end{aligned}$$

The claim of the theorem hence follows. □

Appendix C

Adversary Types

In the standard OCO literature, the various flavors of adversaries (a.k.a. Nature) differ in the amount of information they have access to, when choosing the cost $f_t \in \mathcal{F}$ at round t.

C-1 Standard Definitions

- An **Oblivious Adversary** does not use the information of the Player's actions (past or present) when choosing the cost f_t. That is, this adversary does not have access to $x_1, \ldots, x_t$ at round t. This is equivalent to an adversary that chooses the cost sequence $\{f_t\}_{t=0}^T$ before the game starts.
- An **Adaptive Adversary** can use all the information from the Player's past actions $x_0, \ldots, x_{t-1}$ when choosing the cost f_t. Alternative definitions allow the adversary to also use information from the player's current action x_t.

C-2 The Problem with Adaptive Adversaries

When using the standard definition of regret (1-1) as a metric of performance against Adaptive Adversaries, the meaning of this metric becomes difficult to interpret. The excerpt bellow is taken from [CBL06], but adapted so the notation matches to one used in this thesis:

> "It is important to point out that in the definition of the standard regret (1-1), the cumulative loss $\sum_{t=0}^T f_t(x)$ associated with the "constant" action x corresponds to the sequence of costs $f_0, \ldots, f_T$ of the Adversary. The costs chosen by the Adversary may depend on the player's actions, which, in this case, are $x_0, ..., x_T$. Therefore, it is important to keep in mind that if the Adversary is adaptive, then $\sum_{t=0}^T f_t(x)$ is not the same as the cumulative loss the Player would have suffered had he played action $x_t = x$ for all t."

The next excerpt is adapted from [ADT12], where the meaning of $\min_{x\in\mathcal{X}} \sum_{t=0}^{T} f_t(x)$, (second term of Eq. (1-1)) is discussed for Adaptive Adversaries:

> "We can attempt to articulate the meaning of this term: it is the loss in the peculiar situation where the Adversary reacts to the Player's original sequence $(x_0, \ldots, x_T)$, but the Player somehow manages to secretly play the sequence $x_t = x$ for all t. This is not a feasible situation and it is unclear why this quantity is an interesting baseline for comparison."

Therefore, we can conclude that the standard regret (1-1) is not a suitable metric to measure the performance of the Player against Adaptive Adversaries, and only makes sense against oblivious ones. In order to cope with this issue, the notion of **policy-regret** was introduced. See [ADT12] and [AHM15] for further discussions on this topic.

Bibliography

[ABRT08] Jacob Abernethy, Peter Bartlett, Alexander Rakhlin, and Ambuj Tewari. Optimal strategies and minimax lower bounds for online convex games. In *Conference on Learning Theory (COLT 2008)*, pages 415–423, 2008.

[AD10] Alekh Agarwal and Ofer Dekel. Optimal algorithms for online convex optimization with multi-point bandit feedback. In *Conference of Learning Theory (COLT 2010)*, pages 28–40, 2010.

[ADT12] Raman Arora, Ofer Dekel, and Ambuj Tewari. Online bandit learning against an adaptive adversary: from regret to policy regret. *arXiv preprint arXiv:1206.6400*, 2012.

[AHKS06] Amit Agarwal, Elad Hazan, Satyen Kale, and Robert E Schapire. Algorithms for portfolio management based on the newton method. In *Proceedings of the 23rd International Conference on Machine Learning (ICML 2006)*, pages 9–16, 2006.

[AHM15] Oren Anava, Elad Hazan, and Shie Mannor. Online learning for adversaries with memory: price of past mistakes. In *Advances in Neural Information Processing Systems (NIPS 2015)*, pages 784–792, 2015.

[BBV04] Stephen Boyd, Stephen P Boyd, and Lieven Vandenberghe. *Convex optimization.* Cambridge university press, 2004.

[BCKP20] Aditya Bhaskara, Ashok Cutkosky, Ravi Kumar, and Manish Purohit. Online learning with imperfect hints. In *Proceedings of the 37th International Conference on Machine Learning (ICML 2020)*, 2020.

[BGZ15] Omar Besbes, Yonatan Gur, and Assaf Zeevi. Non-stationary stochastic optimization. *Operations research*, 63(5):1227–1244, 2015.

[BT03] Amir Beck and Marc Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Operations Research Letters*, 31(3):167–175, 2003.

[Bub11] Sébastien Bubeck. Introduction to online optimization. *Lecture notes*, 2011.

[Bub15] S. Bubeck. *Convex Optimization: Algorithms and Complexity.* Foundations and Trends in Machine Learning. Now Publishers, 2015.

[CBL06] Nicolo Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games.* Cambridge University Press, 2006.

[CCL$^+$16] Niangjun Chen, Joshua Comden, Zhenhua Liu, Anshul Gandhi, and Adam Wierman. Using predictions in online optimization: Looking forward with an eye on the past. *ACM SIGMETRICS Performance Evaluation Review*, 44(1):193–206, 2016.

[Cov91] Thomas M Cover. Universal portfolios. *Mathematical Finance*, 1(1):1–29, 1991.

[DFHJ17] Ofer Dekel, Arthur Flajolet, Nika Haghtalab, and Patrick Jaillet. Online learning with a hint. In *Advances in Neural Information Processing Systems (NIPS 2017)*, pages 5299–5308, 2017.

[HAK07] Elad Hazan, Amit Agarwal, and Satyen Kale. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2-3):169–192, 2007.

[Haz16] Elad Hazan. Introduction to online convex optimization. *Foundations and Trends in Optimization*, 2(3-4):157–325, 2016.

[HNKK19] Nam Ho-Nguyen and Fatma Kılınç-Karzan. Exploiting problem structure in optimization under uncertainty via online convex optimization. *Mathematical Programming*, 177(1-2):113–147, 2019.

[HSSW98] David P Helmbold, Robert E Schapire, Yoram Singer, and Manfred K Warmuth. On-line portfolio selection using multiplicative updates. *Mathematical Finance*, 8(4):325–347, 1998.

[HW13] Eric Hall and Rebecca Willett. Dynamical models and tracking regret in online convex programming. In *Proceedings of the 30th International Conference on Machine Learning (ICML 2013)*, pages 579–587, 2013.

[HW15] Eric C Hall and Rebecca M Willett. Online convex optimization in dynamic environments. *IEEE Journal of Selected Topics in Signal Processing*, 9(4):647–662, 2015.

[JRSS15] Ali Jadbabaie, Alexander Rakhlin, Shahin Shahrampour, and Karthik Sridharan. Online optimization: Competing with dynamic comparators. In *Proceedings of the 18th International Conference on Artificial Intelligence and Statistics (AISTATS 2015)*, pages 398–406, 2015.

[LGW20] Yiheng Lin, Gautam Goel, and Adam Wierman. Online optimization with predictions and non-convex losses. *Proc. ACM Meas. Anal. Comput. Syst.*, 4(1), 2020.

[LLST20] Antoine Lesage-Landry, Iman Shames, and Joshua A Taylor. Predictive online convex optimization. *Automatica*, 113:108771, 2020.

[LQL18] Y. Li, G. Qu, and N. Li. Using predictions in online optimization with switching costs: A fast algorithm and a fundamental limit. In *American Control Conference (ACC 2018)*, pages 3008–3013, 2018.

[MSJR16] Aryan Mokhtari, Shahin Shahrampour, Ali Jadbabaie, and Alejandro Ribeiro. Online optimization in dynamic environments: Improved regret rates for strongly convex problems. In *55th IEEE Conference on Decision and Control (CDC 2016)*, pages 7195–7201, 2016.

[Nes04] Yurii Nesterov. *Introductory lectures on convex optimization: A basic course.* Springer, 2004.

[NY83] Arkadi Semenovich Nemirovski and David Berkovich Yudin. *Problem Complexity and Method Efficiency in Optimization.* Wiley, New York, 1983.

[RCT19] R. J. Ravier, A. R. Calderbank, and V. Tarokh. Prediction in online convex optimization for parametrizable objective functions. In *58th IEEE Conference on Decision and Control (CDC 2019)*, pages 2455–2460, 2019.

[RS13a] Alexander Rakhlin and Karthik Sridharan. Online learning with predictable sequences. In *Conference on Learning Theory (COLT 2013)*, pages 993–1019, 2013.

[RS13b] Sasha Rakhlin and Karthik Sridharan. Optimization, learning, and games with predictable sequences. In *Advances in Neural Information Processing Systems (NIPS 2013)*, pages 3066–3074, 2013.

[SJ17] Shahin Shahrampour and Ali Jadbabaie. Distributed online optimization in dynamic environments using mirror descent. *IEEE Transactions on Automatic Control*, 63(3):714–725, 2017.

[SM10] Matthew Streeter and H Brendan McMahan. Less regret via online conditioning. *preprint arXiv:1002.4862*, 2010.

[SS$^+$12] Shai Shalev-Shwartz et al. Online learning and online convex optimization. *Foundations and Trends in Machine Learning*, 4(2):107–194, 2012.

[SSS07] Shai Shalev-Shwartz and Yoram Singer. Logarithmic regret algorithms for strongly convex repeated games. *The Hebrew University*, 2007.

[vEK16] Tim van Erven and Wouter M Koolen. Metagrad: Multiple learning rates in online learning. In *Advances in Neural Information Processing Systems (NIPS 2016)*, pages 3666–3674, 2016.

[WCSB19] Nolan Wagener, Ching-An Cheng, Jacob Sacks, and Byron Boots. An online learning approach to model predictive control. *Proceedings of Robotics: Science and Systems (RSS)*, 2019.

[YMJZ14] Tianbao Yang, Mehrdad Mahdavi, Rong Jin, and Shenghuo Zhu. Regret bounded by gradual variation for online convex optimization. *Machine learning*, 95(2):183–223, 2014.

[YZJY16] Tianbao Yang, Lijun Zhang, Rong Jin, and Jinfeng Yi. Tracking slowly moving clairvoyant: Optimal dynamic regret of online learning with true and noisy gradient. In *Proceedings of the 33rd International Conference on International Conference on Machine (ICML 2016)*, 2016.

[Zin03] Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th International Conference on Machine Learning (ICML 2003)*, pages 928–936, 2003.

[ZLZ18] Lijun Zhang, Shiyin Lu, and Zhi-Hua Zhou. Adaptive online learning in dynamic environments. In *Advances in Neural Information Processing Systems (NIPS 2018)*, pages 1323–1333, 2018.

[ZYY+17] Lijun Zhang, Tianbao Yang, Jinfeng Yi, Rong Jin, and Zhi-Hua Zhou. Improved dynamic regret for non-degenerate functions. In *Advances in Neural Information Processing Systems (NIPS 2017)*, pages 732–741, 2017.