Dynamic Regret Bounds without Lipschitz Continuity: Online Convex Optimization with Multiple Mirror Descent Steps

Nima Eshraghi and Ben Liang

***Abstract*— We study the dynamic regret in online convex optimization (OCO), where the cost functions are revealed sequentially over time. Prior studies on the dynamic regret of OCO algorithms often require the cost functions to be Lipschitz continuous. However, the costs functions that arise in many applications may not satisfy this condition. In this work, we analyze the performance of Online Multiple Mirror Descent (OMMD), which can handle non-Lipschitz cost functions. OMMD is based on mirror descent but uses multiple mirror descent steps per online round. We first derive two upper bounds on the dynamic regret based on the path length and squared path length, and we further derive a third upper bound based on the cumulative optimal cost, which can be much smaller than the path length or squared path length especially when the sequence of minimizers fluctuates over time. We show that the dynamic regret of OMMD scales linearly with the minimum among the path length, squared path length, and cumulative optimal cost.**

I. Introduction

Online optimization refers to the design of sequential decisions where system parameters and cost functions vary with time. It has important applications to various classes of problems in control and learning [1]–[4]. In this work, we consider the problem of online convex optimization (OCO), which can be formulated as a discrete-time sequential learning process as follows. At each round t, the learner first makes a decision $x_t \in \mathcal{X}$, where $\mathcal{X}$ is a convex set representing the solution space. The learner then receives a convex cost function $f_t(\cdot) : \mathcal{X} \to \mathbb{R}$ and suffers the corresponding cost of $f_t(x_t)$ associated with the submitted decision. The goal of the online learner is to minimize the total accrued cost over a finite number of rounds, denoted by T. For performance evaluation, prior studies on online learning often focus on the *static* regret, defined as the difference between the learner's accumulated cost and that of an optimal fixed offline decision, which is made in hindsight with knowledge of $f_t(\cdot)$ for all t:

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

A successful online algorithm closes the gap between the online decisions and the offline counterpart when normalized by T, i.e., sustaining sublinear static regret in T. In the literature, there are various online algorithms [5]–[7] that guarantee a sublinear bound on the static regret.

N. Eshraghi and B. Liang are with the Department of Electrical and Computer Engineering, University of Toronto, Ontario, Canada. (`e-mail: {neshraghi, liang}@ece.utoronto.ca`).

This work was funded in part by the Natural Sciences and Engineering Research Council of Canada.

However, the static regret fails to accurately reflect the quality of decisions in many practical scenarios, e.g., the object tracking application where we aim to follow the movement of some target over time. Therefore, the *dynamic* regret has been proposed to allow comparison against an arbitrary comparator sequence [5]. In more recent literature, the learner performance is commonly measured relative to the best comparator sequence [1], [8]–[12], i.e.,

$$\mathrm{Reg}_T^{\mathrm{d}} = \sum_{t=1}^{T} f_t(x_t) - \sum_{t=1}^{T} f_t(x_t^*), \tag{2}$$

where $x_t^* = \mathrm{argmin}_{x\in\mathcal{X}} f_t(x)$ is a minimizer of the cost at round t.

It is well-known that the online optimization problem may be intractable in a dynamic setting, due to arbitrary fluctuation in the cost functions. Hence, achieving a sublinear bound on the dynamic regret may be impossible. However, it is possible to upper bound the dynamic regret in terms of certain regularity measures. The regularity measures reflect how fast an environment evolves as time progresses. Prior studies [5], [9]–[14] have utilized a variety of regularity measures to bound the dynamic regret. One of the measures to represent regularity is the *path length*, defined by

$$C_T = \sum_{t=2}^{T} \|x_t^* - x_{t-1}^*\|, \tag{3}$$

which illustrates the accumulative variation in the minimizer sequence. For instance, the dynamic regret of online gradient descent (OGD) for convex cost functions can be bounded by $O(\sqrt{T}(1 + C_T))$ [5]. For strongly convex cost functions, the dynamic regret of OGD can be reduced to $O(C_T)$ [10]. Another regularity measure based on the sequence of minimizers is the *squared path length*, i.e.,

$$S_T = \sum_{t=2}^{T} \|x_t^* - x_{t-1}^*\|^2, \tag{4}$$

which can be smaller than the path length when the distance between successive minimizers is small. Another metric to represent problem regularity is the cumulative cost of the optimizers. In particular, it has been shown that the static regret of online mirror descent (OMD) is bounded by $O(\sqrt{F_T})$, where $F_T = \sum_{t=1}^{T} f_t(x^*)$ is the cumulative cost of a fixed minimizer $x^* = \mathrm{argmin}_{x\in\mathcal{X}} \sum_{t=1}^{T} f_t(x)$ [15]. To bound the dynamic regret, a natural extension is to replace the fixed minimizer by a sequence of per round minimizers, defined by

$$F_T = \sum_{t=1}^{T} f_t(x_t^*). \quad (5)$$

The measure in (5) has been used under various names in the analysis of regret in the literature of online optimization [16].

All prior studies on the dynamic regret of OCO algorithms require the cost functions to be Lipschitz continuous [5]. Thus, their analyses fails to bound the dynamic regret of many important cost functions do not satisfy this condition, e.g., broad class of quadratic programming problems, and support vector machine training. Therefore, we investigate whether it is possible to bound the dynamic regret while relaxing the Lipschitz continuity condition on the cost functions.

To this end, in this work, we analyze the dynamic regret of the Online Multiple Mirror Descent (OMMD) algorithm, which is a natural extension to OMD but allows multiple steps of mirror descent per round. A salient feature of OMMD is that it does not require the cost functions to be Lipschitz continuous. We derive three upper bounds on the dynamic regret based on the path length C_T, the squared path length S_T, and the cumulative optimal cost F_T, for smooth and strongly convex functions. We note that F_T can be much smaller than both C_T and S_T, especially when the sequence of minimizers fluctuate drastically over time. Thus, we obtain an overall regret bound of $O(\min(C_T, S_T, F_T))$. Furthermore, we conduct numerical simulations to reveal that OMMD can achieve substantial improvement in the dynamic regret, in an object tracking application, compared with single-step dynamic mirror descent [13], online multiple gradient descent (OMGD) [11], and OGD [5].

II. Related Works

There is a rich body of works in the literature of online learning devoted to studying the dynamic regret of OGD and its variants [5], [8]–[12]. In contrast, only a few studies are based on the mirror descent method [13], [14]. Here we review the most relevant works on the dynamic regret of these two popular online optimization methods.

Dynamic regret of OGD: Dynamic regret was first introduced in [5] for the analysis of OGD, where an $O(\sqrt{T}(1 + C_T))$ bound on the dynamic regret was derived for convex functions[1]. To obtain stronger regret bounds, recent works often focus on the dynamic regret of form (2). Dynamic regret of different variants of OGD under various settings is studied in [8]–[10].

The above works make only a single query to the gradient of the cost functions in every round. Recent studies have shown that the dynamic regret of OGD can be improved when the learner access the gradients of the cost functions more than once [11], [12]. For strongly convex and smooth online optimization in an unconstrained setup, it has been proved that OGD with preconditioned gradients and multiple queries of gradients have regret bounds of $O(S_T)$ and $O(\min(C_T, V_T))$, respectively [12].

Finally, online multiple gradient descent was analyzed in [11], where the learner makes a fixed number of gradient queries in every round. It is shown that the dynamic regret is upper bounded by $O(\min(C_T, S_T))$ when the cost functions are Lipschitz continuous, strongly convex, and smooth [11]. In contrast, our analysis is focused on OMD, where the distance between two points in measured via Bregman divergence, which generalizes the Euclidean distance. Furthermore, we relax the Lipschitz continuity requirement. We show that even after such relaxation, the dynamic regret bound can be improved to $O(\min(C_T, S_T, F_T))$.

Dynamic Regret of OMD: The dynamic regret of online single-step mirror descent was studied in [13], where an upper bound of $O(\sqrt{T}(1 + C_T))$ was derived for convex and Lipschitz continuous cost functions. To take advantage of smoothness in cost functions, an adaptive algorithm based on optimistic mirror descent [17] was proposed in [14], which contains two steps of mirror descent per online round. However, different from our work, in that variant the learner is allowed to make only a single query about the gradient. The algorithm further requires some prior prediction of the gradient in each round, which is used in the first mirror descent step. The dynamic regret bound was given in terms of a combination of the path length C_T, deviation between the predictions and the actual gradients D'_T, and functional variation V_T under the condition that these regularity measures can be computed by the learner on-the-fly. In particular, to achieve this bound, the algorithm requires the design of a time-varying step size that depends on the optimal solution in the previous step.

All aforementioned works on OMD make only a single query to the gradient of the cost functions in every round. In contrast, in this work, we study the dynamic regret of OMD when the learner makes multiple gradient queries per round. Furthermore, we show that even after relaxing the Lipschitz continuity requirement that was commonly assumed in the aforementioned studies, the upper bound on the dynamic regret can be improved to $O(\min(C_T, S_T, F_T))$ when the learner can access the gradients of the cost functions multiple times.

III. Online Multiple Mirror Descent

In this section, we describe OMMD and discuss how the learner can improve the dynamic regret by performing multiple mirror descent steps per round.

We consider online optimization over a finite number of rounds, denoted by T. At the beginning of every round t, the learner submits a decision represented by x_t, which is taken from a convex and compact set $\mathcal{X}$. Then, the cost function $f_t(\cdot)$ is revealed and the learner suffers the corresponding cost $f_t(x_t)$. The learner then updates its decision in the next

[1] A more general definition of the dynamic regret was introduced in [5], which allows comparison against an arbitrary sequence $\{u_t\}_{t=1}^T$. We note that the regret bounds developed in [5] also hold for the specific case of $u_t = x_t^*$.

round. With standard mirror descent, this is given by

$$x_{t+1} = \underset{x\in\mathcal{X}}{\operatorname{argmin}}\Big\{\langle \nabla f_t(x_t), x\rangle + \frac{1}{\alpha}D_r(x, x_t)\Big\} \qquad (6)$$

where α is a fixed step size, and $D_r(\cdot,\cdot)$ is the Bregman divergence corresponding to the regularization function $r(\cdot)$. The update in (6) suggests that the learner aims to stay close to the current decision x_t as measured by the Bregman divergence, while taking a step in a direction close to the negative gradient to reduce the current cost at round t.

OMMD uses mirror descent in its core as the optimization workhorse. However, in contrast to classical OMD, where the learner queries the gradient of each cost function only once, OMMD is designed to take advantage of the curvature of cost functions by allowing the learner to make multiple queries to the gradient in each round. This is especially important when the successive cost functions have similar curvatures. In particular, in order to track x_t^* the learner needs to access the gradient of the cost function, i.e., $\nabla f_t(\cdot)$. Unfortunately, this information is not available until the end of round t. However, if the successive functions have similar curvatures, the gradient of $f_{t-1}(\cdot)$ is a reasonably accurate estimate for the gradient of $f_t(\cdot)$. In this case, every time that the learner queries the gradient of $f_{t-1}(\cdot)$, it finds a point that is likely to be closer to the minimizer of $f_t(\cdot)$. Hence, it may benefit the learner to perform multiple mirror descent steps in each round.

Thus, the learner generates a series of decisions, represented by $y_t^0, y_t^1, \ldots, y_t^m$, via the following updates:

$$y_t^0 = x_{t-1},$$
$$y_t^i = \underset{y\in\mathcal{X}}{\operatorname{argmin}}\Big\{\langle \nabla f_{t-1}(y_t^{i-1}), y\rangle + \frac{1}{\alpha}D_r(y, y_t^{i-1})\Big\}. \qquad (7)$$

Then, by setting $x_t = y_t^m$, the learner proceeds to the next round, and the procedure continues. Note that m is independent of T.

Applying multiple steps of mirror descent can reveal more information about the sequence of minimizers. It can reduce the dynamic regret, but only if the series of decisions in (7) helps decrease the distance to the minimizer x_t^*. Therefore, quantifying the benefit of OMMD over standard mirror descent requires careful analysis on the impact of the fluctuation of $f_t(\cdot)$ over time. To this end, we analyze and bound the dynamic regret of OMMD in the next section.

IV. Bounding the Dynamic Regret of OMMD

Following previous works on mirror descent [18], [19], we introduce the following assumption.

Assumption 1. The cost functions $f_t(\cdot)$ are λ-strongly convex and L-smooth with respect to a differentiable function $r(\cdot)$, i.e.,

$$f(y) + \langle \nabla f(y), x-y\rangle + \lambda D_r(x,y) \le f(x),\ \forall x, y \in \mathcal{X}.$$
$$f(y) \le f(x) + \langle \nabla f(x), y-x\rangle + L D_r(y,x),\ \forall x, y \in \mathcal{X}.$$

Furthermore, the regularization function $r(\cdot)$ is μ-strongly convex and μ'-smooth with respect to some norm.

Algorithm 1 Online Multiple Mirror Descent

Input: Arbitrary initialization of $x_1 \in \mathcal{X}$; step size α; time horizon T.

Output: Sequence of decisions $\{x_t : 1 \le t \le T\}$.

for $t = 1, 2, \ldots, T$ **do**
  submit $x_t \in \mathcal{X}$ and receive $f_t(\cdot)$
  set $y_t^0 = x_{t-1}$
  for $i = 1, 2, \ldots, m$ **do**
    $y_t^i = \underset{y\in\mathcal{X}}{\operatorname{argmin}}\{\langle \nabla f_{t-1}(y_t^{i-1}), y\rangle + \frac{1}{\alpha}D_r(y, y_t^{i-1})\}$
  end for
  set $x_t = y_t^m$
end for

We note that these are standard assumptions commonly used in the literature after the group of studies began by [6], [20], to provide stronger regret bounds by constraining the curvature of cost functions. Example cost functions that are not Lipschitz continuous but satisfy the conditions in Assumption 1 are given in [21], [22].

Assumption 2. We further make a standard assumption that the Bregman divergence is Lipschitz continuous as follows:

$$|D_r(x,z) - D_r(y,z)| \le K\|x-y\|,\ \forall x, y, z \in \mathcal{X}, \qquad (8)$$

where K is a positive constant. When the function $r(\cdot)$ is Lipschitz continuous on the feasibility domain, the Lipschitz condition on the Bregman divergence is automatically satisfied. Popular choices of $r(\cdot)$ that satisfy the condition in Assumption 2 are provided in [1]. This condition is commonly assumed in the study of OMD [1], [14], [23], [24].

In this work, we do not require the cost functions to be Lipschitz continuous. Thus, the condition stated in (8) can be viewed as a replacement for imposing Lipschitz continuity on the cost functions. Since the sequence of cost functions are revealed to the learner, the learner has no control over it. In contrast, the learner can design and control the regularization function. Hence, if the cost functions happen to not meet the Lipschitz condition, the learner can benefit from carefully choosing a regularization function that is Lipschitz continuous. Therefore, the condition in (8) is indeed a milder requirement compared with the assumption of Lipschitz continuity of the cost functions. In the context of this work, it can be viewed as a small price that the learner pays to deal with non-Lipschitz cost functions.

The following lemma paves the way for the proposed analysis on the dynamic regret of OMMD. It bounds the distance of the learner's future decision from the current optimal solution, after a single step of mirror descent.

Lemma 1: Assume that $f_t(\cdot)$ is λ-strongly convex and L-smooth with respect to a differentiable function $r(\cdot)$. Single-step mirror descent with a fixed step size $\alpha \le \frac{1}{L}$ guarantees the following:

$$D_r(x_t^*, x_{t+1}) \le \beta D_r(x_t^*, x_t),$$

where x_t^* is the minimizer of $f_t(\cdot)$, and $\beta = 1 - \frac{2\alpha\lambda}{1+\alpha\lambda}$.

Remark 1. Lemma 1 states that a mirror descent step reduces the distance (measured by the Bregman divergence) of the learner's decisions to the current minimizer. This generalizes the results in [10], [11], where similar bounds were derived for OGD when the distance was measured in Euclidean norms. In particular, those results correspond to the special choice of $r(x) = \|x\|_2^2$, which reduces the Bregman divergence to Euclidean distance, i.e., $D_r(x, y) = \|x - y\|^2$.

Lemma 1 indicates that the distance between the next decision x_{t+1} and the minimizer x_t^* is strictly smaller than the distance between the current decision x_t and x_t^*. This implies that in a slowly changing environment, where the minimizers of the functions $f_t(\cdot)$ and $f_{t+1}(\cdot)$ are not far from each other, applying mirror descent multiple times enables the online learner to more accurately track the sequence of optimal solutions x_t^*.

The following theorems provide three separate upper bounds on the dynamic regret of OMMD, based on path length C_T, squared path length S_T, and cumulative optimal cost F_T.

Theorem 2: Under Assumptions 1 and 2, let x_t be the sequence of decisions generated by OMMD with a fixed step size $\frac{1}{2\lambda} < \alpha \leq \frac{1}{L}$. When $m \geq \lceil \frac{1+\alpha\lambda}{2\alpha\lambda} \log(\frac{\mu'}{\mu}) \rceil$, the dynamic regret is upper bounded by

$$\mathrm{Reg}_T^{\mathrm{d}} \leq \frac{\lambda K}{2\alpha\lambda - 1}\Big(\frac{1 + \sqrt{\beta^m \mu'/\mu}}{1 - \sqrt{\beta^m \mu'/\mu}}\Big)\Big(\|x_1 - x_1^*\| + C_T\Big).$$

Remark 2. It has been shown in [13] that single-step mirror descent guarantees an upper bound of $O(\sqrt{T}(1 + C_T))$ on the dynamic regret for convex and Lipschitz cost functions. With that bound, a sublinear path length is not sufficient to guarantee sublinear dynamic regret. In contrast, Theorem 2 implies that OMMD reduces the upper bound to $O(C_T)$ when the cost functions are strongly convex and smooth, which implies that a sublinear path length is sufficient to yield sublinear dynamic regret.

Theorem 3: Under Assumptions 1 and 2, let x_t be the sequence of decisions generated by OMMD with a fixed step size $\alpha \leq \frac{1}{L}$. When $m \geq \lceil \frac{1+\alpha\lambda}{2\alpha\lambda} \log(\frac{\mu'}{2\mu}) \rceil$, the dynamic regret satisfies

$$\mathrm{Reg}_T^{\mathrm{d}} \leq \sum_{t=1}^{T} \frac{\|\nabla f_t(x_t^*)\|_*^2}{2L} + \frac{L\mu(\mu' + 1)}{\mu - 2\mu'\beta^m}\Big(\frac{\|x_1 - x_1^*\|^2}{2} + S_T\Big).$$

Since the gradient at x_t^* is zero if x_t^* is in the relative interior of the feasibility set $\mathcal{X}$, i.e., $\|\nabla f_t(x_t^*)\| = 0$, the above theorem can be simplified to the following corollary.

Corollary 4: If x_t^* belongs to the relative interior of the feasibility set $\mathcal{X}$ for all t, the dynamic regret bound in Theorem 3 is of order $O(S_T)$.

When the cost functions drift slowly, the distances between successive minimizers are small. Hence, the squared path length S_T, which relies on the square of those distances, can be significantly smaller than the path length C_T. In this case, Theorem 3 and Corollary 4 can provide a tighter regret bound than Theorem 2.

Theorem 5: Under Assumptions 1 and 2, let x_t be the sequence of decisions generated by OMMD with arbitrary $m \geq 1$ and a fixed step size $\frac{1}{2\lambda} < \alpha < \sqrt{\frac{\mu}{2L\mu'}}$. The dynamic regret satisfies

$$\mathrm{Reg}_T^{\mathrm{d}} \leq \frac{2L\mu'\alpha^2}{\mu - 2L\mu'\alpha^2} F_T.$$

Remark 3. Interestingly, Theorem 5 implies that the dynamic regret can be bounded by the cumulative cost of a sequence of minimizers, i.e., $F_T = \sum_{t=1}^{T} f_t(x_t^*)$. This bound is especially important when the sequence of minimizers drastically fluctuates over time. In this scenario, the path length C_T and squared path length S_T may grow linearly, whereas the third regret bound based on F_T could be smaller.

Theorem 2, Corollary 4, and Theorem 5, respectively, state that the dynamic regret of OMMD is upper bounded *linearly* by path length C_T, squared path length S_T, and cumulative optimal cost F_T. This immediately leads to the following result.

Corollary 6: Under the conditions stated in Assumptions 1 and 2, the dynamic regret of OMMD with suitably chosen α has an upper bound of $O(\min(C_T, S_T, F_T))$.

Remark 4. We note that [10] and [11] provide upper bounds of $O(C_T)$ and $O(\min(C_T, S_T))$, respectively, on the dynamic regret of OGD with single and multiple gradient queries, while [13] presents an upper bound of $O(\sqrt{T}(1 + C_T))$ on the dynamic regret of OMD with a single gradient query per round. Corollary 6 shows that OMMD can improve the dynamic regret bound to $O(\min(C_T, S_T, F_T))$. Furthermore, in contrast to the previous studies, our analysis does not require the cost functions to be Lipschitz continuous.

V. OMMD with Time-Varying Number of Mirror Descent Steps

In practice, it may not be possible to maintain the same number of gradient computation in every round due to, e.g., the potentially time-varying processing speed in computing systems. In this case, the number of queries to the gradients of the cost functions and mirror descent updates could vary over time. We denote by m_t the number of mirror descent steps at round t and assume $m_t \geq 1$ for all t. In this case, m_t can be modeled by a random variable, which depends on the amount of available processing resources to compute the gradients of the cost function at round t. The following theorem bounds the expected dynamic regret of OMMD with time-varying mirror descent steps.

Theorem 7: Under Assumptions 1 and 2, let x_t be the sequence of decisions generated by OMMD with a fixed step size $\frac{1}{2\lambda} < \alpha \leq \frac{1}{L}$. Let $m_{\min} = \min\{m_t\}$ be the minimum number of mirror descent steps. If $\mathbb{E}[\beta^{\frac{m_t}{2}}] < \sqrt{\mu/\mu'}$, the expected dynamic regret satisfies

$$\mathbb{E}\Big[\mathrm{Reg}_T^{\mathrm{d}}\Big] \leq \frac{\lambda K}{2\alpha\lambda - 1}\left(\frac{1 + \sqrt{\beta^{m_{\min}}\mu'/\mu}}{1 - \mathbb{E}[\beta^{\frac{m_t}{2}}]\sqrt{\mu'/\mu}}\right)\Big(\|x_1 - x_1^*\| + C_T\Big).$$

Furthermore, if $\mathbb{E}[\beta^{m_t}] < \mu/2\mu'$, the expected dynamic regret is upper bounded by

$$\mathbb{E}\Big[\mathrm{Reg}_T^{\mathrm{d}}\Big] \leq \sum_{t=1}^{T} \frac{\|\nabla f_t(x_t^*)\|_*^2}{2L} + \frac{L\mu(\mu'+1)}{\mu - 2\mu'\mathbb{E}[\beta^{m_t}]}\left(\frac{\|x_1 - x_1^*\|^2}{2} + S_T\right).$$

The proof scheme is similar to the proof of Theorems 2 and 3, and is omitted for brevity.

Remark 5. An instructive example for the above conditions is when m_t has a Poisson distribution with mean ν. In this case, both conditions can be equivalently expressed in terms of a lower bound on ν.

The results of Theorem 7 shows that even when it is not possible to compute the same number of gradients in every round, the dynamic regret of OMMD is still bounded in expectation. Furthermore, Theorem 5 show that the dynamic regret of OMMD is upper bounded by $O(F_T)$ when $m \geq 1$. It can be shown that the same upper bound holds on the expected dynamic regret of OMMD with time-varying number of mirror descent steps as long as $m_t \geq 1$. This immediately leads to the following results.

Corollary 8: Under the conditions stated in Theorem 7, the expected dynamic regret of OMMD with time-varying mirror descent steps has an upper bound of $O(\min(C_T, S_T, F_T))$.

VI. Simulations

We investigate the performance of OMMD via numerical experiments, in two different scenarios. We compare OMMD with the following alternatives: Online Gradient Descent (OGD) [5], Online Multiple Gradient Descent (OMGD) [11], and Dynamic Mirror Descent (DMD) [13].

First, we consider an online object tracking problem, where the object can randomly move on a 2-D plane. Similar to [1], we have used the Euclidean distance as the Bregman divergence, which turns our framework to state estimation and tracking. We note that although OMMD reduces to OMGD [11] in this special case, but unlike [11], our analysis does not require the cost functions to be Lipschitz continuous, and our bound is stronger.

Let us consider a maneuvering target in the 2-D plane and assume that each position component of the online target evolves independently according to a constant-velocity model [25]. The state of the target at each round consists of four components: horizontal position, vertical position, horizontal velocity, and vertical velocity. Therefore, with $x_t^* \in \mathbb{R}^4$ representing the state at round t, the state-space model takes the form

$$x_{t+1}^* = Ax_t^* + \nu_t,$$

where $\nu_t \in \mathbb{R}^4$ is the system noise, and

$$A = I_2 \otimes \begin{bmatrix} 1 & \epsilon \\ 0 & 1 \end{bmatrix}$$

where I_2 is the identity matrix, $\otimes$ is the Kronecker product, and ϵ represents the sampling interval. For our experiment,

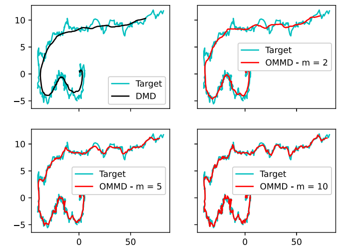

Fig. 1. Trajectory of the moving object is compared with the trajectory of the estimator obtained by DMD and OMMD for different values of m.

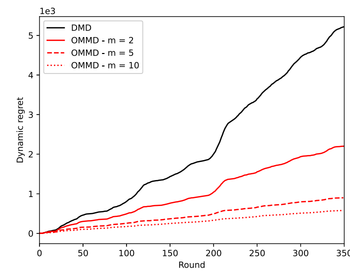

Fig. 2. Dynamic regret comparison for object tracking.

we generate noise according to a zero-mean Gaussian distribution with unit variance, and we set $\epsilon = 0.1$ seconds. At round t, the learner observes a noisy version of x_t^* as follows:

$$z_t = x_t^* + n_t,$$

where n_t denotes the observation noise, generated independently from a uniform distribution on $[-0.5, 0.5]$. We use a squared cost function $f_t(x) = \|z_t - x\|^2$. Then, the dynamic regret reflects the accumulated tracking error over time.

Fig. 1 plots the target trajectory versus the estimator trajectory obtained by DMD, and OMMD, for $m = 2, 5, 10$. The trajectory estimators obtained by OMMD with higher m values closely follow the moving target trajectory. Furthermore, in Fig. 2, we compare the performance of OMMD and DMD in terms of the dynamic regret. We see that, for $m = 10$, OMMD can reduce the dynamic regret up to 70% after 350 rounds.

Next, we study the performance of OMMD in solving a sequence of quadratic programming problems of the form $f_t(x_1, x_2) = \rho\|x_1 - a_t\|^2 + \|x_2 - b_t\|^2$, where ρ is a positive constant, a_t and b_t are time-variant vectors, and the decision variable are $x_1 \in \mathbb{R}^{d_1}$, $x_2 \in \mathbb{R}^{d_2}$, such that $d_1 + d_2 = d$. In our experiment, we set $\rho = 10$, $d_1 = 500$, and $d = 1000$. We assume that b_t is time-invariant and for all rounds t we have $b_t = 2$, while a_t satisfies the recursive formula $a_{t+1} = a_t + 1/\sqrt{t}$ with initial value $a_1 = -1.5$. For this example, we consider a fixed number M of mirror descent steps. We further set the step size $\alpha = 0.03$. We use the same regularization function and constraint set as in the previous experiment.

Fig. 3 shows the dynamic regret of a sequence of slowly moving cost functions for OGD, DMD, OMGD and OMMD with $m = 2, 5, 10$. As time progresses and the difference between the successive cost functions becomes less significant,

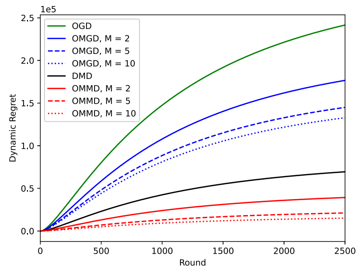

Fig. 3. Dynamic regret comparison for slowly drifting cost functions.

the difference between the minimizers decreases. In this case, OMMD can significantly improve the performance of online optimization by reducing the gap between the learner's decisions and the minimizers sequence. In particular, compared with DMD, OMMD with $m = 10$ reduces the dynamic regret up to 60% after 2500 rounds.

VII. Conclusion

We have studied online mirror descent in dynamic settings, when the mirror descent step can be applied multiple times per round. We derive three upper bounds on the dynamic regret based on the path length C_T, squared path length S_T, and cumulative optimal cost F_T, for strongly convex and smooth cost functions. The main benefit of the third bound is that it can be much smaller than both C_T and S_T when the sequence of minimizers drastically fluctuates over time. In contrast to prior studies [10], [11], [13], our analysis does not require the cost functions to be Lipschitz continuous. Our numerical results further reveal that OMMD can offer substantial improvement on the dynamic regret compared with existing alternatives.

References

[1] S. Shahrampour and A. Jadbabaie, "Distributed online optimization in dynamic environments using mirror descent," *IEEE Transactions on Automatic Control*, vol. 63, no. 3, pp. 714–725, 2018.

[2] R. Dixit, A. S. Bedi, K. Rajawat, and A. Koppel, "Distributed online learning over time-varying graphs via proximal gradient descent," in *Proc. IEEE International Conference on Decision and Control*, 2019.

[3] N. Eshraghi and B. Liang, "Improving dynamic regret in distributed online mirror descent using primal and dual information," in *Proc. Learning for Dynamic and Control Conference*, 2022.

[4] R. Zhang, Y. Li, and N. Li, "On the regret analysis of online LQR control with predictions," in *Proc. American Control Conf.*, 2021.

[5] M. Zinkevich, "Online convex programming and generalized infinitesimal gradient ascent," in *Proc. International Conference on Machine Learning*, 2003.

[6] E. Hazan, A. T. Kalai, S. Kale, and A. Agarwal, "Logarithmic regret algorithms for online convex optimization," in *Proc. Conference on Learning Theory*, 2006.

[7] S. Shalev-Shwartz, "Online learning and online convex optimization," *Foundations and Trends in ML*, vol. 4, no. 2, pp. 107–194, 2012.

[8] O. Besbes, Y. Gur, and A. Zeevi, "Non-stationary stochastic optimization," *Operations Research*, vol. 63, no. 5, pp. 1227–1244, 2015.

[9] T. Yang, L. Zhang, R. Jin, and J. Yi, "Tracking slowly moving clairvoyant: Optimal dynamic regret of online learning with true and noisy gradient," in *Proc. International Conference on Machine Learning*, 2016.

[10] A. Mokhtari, S. Shahrampour, A. Jadbabaie, and A. Ribeiro, "Online optimization in dynamic environments: Improved regret rates for strongly convex problems," in *Proc. IEEE Conference on Decision and Control*, 2016.

[11] L. Zhang, T. Yang, J. Yi, R. Jin, and Z.-H. Zhou, "Improved dynamic regret for non-degenerate functions," in *Proc. Advances in Neural Information Processing Systems*, 2017.

[12] T.-J. Chang and S. Shahrampour, "Unconstrained online optimization: Dynamic regret analysis of strongly convex and smooth problems," in *Proc. Artificial Intelligence and Statistics*, 2020.

[13] E. Hall and R. Willett, "Online convex optimization in dynamic environments," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 4, pp. 647–662, 2015.

[14] A. Jadbabaie, A. Rakhlin, S. Shahrampour, and K. Sridharan, "Online optimization:Competing with dynamic comparators," in *Proc. Artificial Intelligence and Statistics*, 2015.

[15] N. Srebro, K. Sridharan, and A. Tewari, "Smoothness, low noise and fast rates," *Proc. Adv. in Neural Information Processing Systems*, 2010.

[16] N. Littlestone and M. K. Warmuth, "The weighted majority algorithm," *Information and computation*, vol. 108, no. 2, pp. 212–261, 1994.

[17] S. Rakhlin and K. Sridharan, "Optimization, learning, and games with predictable sequences," in *Proc. Advances in Neural Information Processing Systems*, 2013.

[18] H. Hendrikx, L. Xiao, S. Bubeck, F. Bach, and L. Massoulie, "Statistically preconditioned accelerated gradient method for distributed optimization," *arXiv preprint, arXiv:2002.10726*, 2020.

[19] E. Hazan and S. Kale, "Beyond the regret minimization barrier: optimal algorithms for stochastic strongly-convex optimization," *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 2489–2512, 2014.

[20] S. Shalev-Shwartz and Y. Singer, *Logarithmic Regret Algorithms for Strongly Convex Repeated Games*. The Hebrew University, 2007.

[21] H. Lu, R. M. Freund, and Y. Nesterov, "Relatively smooth convex optimization by first-order methods, and applications," *SIAM Journal on Optimization*, vol. 28, no. 1, pp. 333–354, 2018.

[22] H. H. Bauschke, J. Bolte, and M. Teboulle, "A descent lemma beyond lipschitz gradient continuity: first-order methods revisited and applications," *Mathematics of Operations Research*, vol. 42, no. 2, pp. 330–348, 2017.

[23] N. Eshraghi and B. Liang, "Dynamic regret of online mirror descent for relatively smooth convex cost functions," *IEEE Control Systems Letters, Early Access*, 2022.

[24] N. Eshraghi and B. Liang, "Distributed online optimization over a heterogeneous network with any-batch mirror descent," in *Proc. International Conference on Machine Learning*, 2020.

[25] Y. Bar-Shalom and T. E. Fortmann, *Tracking and Data Association*. Academic Press, 1988.

[26] A. Beck and M. Teboulle, "Mirror descent and nonlinear projected subgradient methods for convex optimization," *Operations Research Letters*, vol. 31, no. 3, pp. 167–175, 2003.

[27] C.-K. Chiang, T. Yang, C.-J. Lee, M. Mahdavi, C.-J. Lu, R. Jin, and S. Zhu, "Online optimization with gradual variations," in *Proc. Conference on Learning Theory*, 2012.

Appendix

A. Helper Lemmas

We first present several lemmas that will be used in our regret analysis.

Lemma 9: The following equality holds for any $x, y, z \in \mathcal{X}$, and the Bregman divergence $D_r(\cdot, \cdot)$ with respect to any function $r(\cdot)$:

$$\langle \nabla r(z) - \nabla r(y), x - y \rangle = D_r(x, y) - D_r(x, z) + D_r(y, z).$$

The proof of Lemma 9 is given in earlier studies [26].

Lemma 10: With a single-step of OMD with update

$$x_{t+1} = \underset{x \in \mathcal{X}}{\text{argmin}} \Big\{ \langle \nabla f_t(x_t), x \rangle + \frac{1}{\alpha} D_r(x, x_t) \Big\} \quad (9)$$

the following inequality holds:

$$\langle x_{t+1} - y, \nabla f_t(x_t) \rangle \leq \frac{1}{\alpha} \Big\{ D_r(y, x_t) - D_r(y, x_{t+1}) - D_r(x_{t+1}, x_t) \Big\}, \ \forall y \in \mathcal{X}.$$

Proof. By applying the first-order optimality condition of (9), we have

$$\langle x_{t+1} - y, \nabla f_t(x_t) + \frac{1}{\alpha} \nabla r(x_{t+1}) - \frac{1}{\alpha} \nabla r(x_t) \rangle \leq 0, \ \forall y \in \mathcal{X}. \quad (10)$$

We re-arrange the terms in (10) to obtain

$$\langle x_{t+1} - y, \nabla f_t(x_t)\rangle \le \frac{1}{\alpha}\langle x_{t+1} - y, \nabla r(x_t) - \nabla r(x_{t+1})\rangle \le \frac{1}{\alpha}\Big\{D_r(y, x_t) - D_r(y, x_{t+1}) - D_r(x_{t+1}, x_t)\Big\}, \quad (11)$$

where we have applied Lemma 9 to obtain the right hand-side of (11). □

Lemma 11: Suppose $r(x)$ is strongly convex and y satisfies $\nabla r(y) = \nabla r(u) - \alpha \nabla f(l)$ for some convex function $f(x)$ and step size α. We have

$$\operatorname*{argmin}_{x\in\mathcal{X}}\Big\{\langle \nabla f(l), x\rangle + \frac{1}{\alpha}D_r(x,u)\Big\} = \operatorname*{argmin}_{x\in\mathcal{X}}\ D_r(x,y).$$

The proof of Lemma 11 is given in *Proposition 17* in [27].

Furthermore, we will also use the following Bregman divergence projection inequality to bound the dynamic regret of OMMD. If $y = \operatorname{argmin}_{x\in\mathcal{X}}\ D_r(x,z)$, i.e., y is the Bregman projection of z into the set $\mathcal{X}$, then for any arbitrary point $d \in \mathcal{X}$, we have

$$D_r(d,z) \ge D_r(d,y) + D_r(y,z). \quad (12)$$

B. Proof of Lemma 1

Consider single-step OMD update as follows:

$$x_{t+1} = \operatorname*{argmin}_{x\in\mathcal{X}}\Big\{\langle \nabla f_t(x_t), x\rangle + \frac{1}{\alpha}D_r(x, x_t)\Big\}. \quad (13)$$

By applying the result of Lemma 10, we have

$$\langle x_{t+1} - x_t^*, \nabla f_t(x_t)\rangle \le \frac{1}{\alpha}\Big\{D_r(x_t^*, x_t) - D_r(x_t^*, x_{t+1}) - D_r(x_{t+1}, x_t)\Big\}, \quad (14)$$

where x_t^* is a minimizer of the cost function, i.e., $x_t^* = \operatorname{argmin}_{x\in\mathcal{X}} f_t(x)$. Furthermore, from the smoothness of $f_t(\cdot)$, we have

$$f_t(x_{t+1}) \le f_t(x_t) + \langle \nabla f_t(x_t), x_{t+1} - x_t\rangle + L D_r(x_{t+1}, x_t). \quad (15)$$

By combining (14) and (15) we obtain

$$f_t(x_{t+1}) \le f_t(x_t) + \langle \nabla f_t(x_t), x_t^* - x_t\rangle + \Big(L - \frac{1}{\alpha}\Big)D_r(x_{t+1}, x_t) + \frac{1}{\alpha}\Big\{D_r(x_t^*, x_t) - D_r(x_t^*, x_{t+1})\Big\}. \quad (16)$$

The strong convexity of the cost function $f_t(\cdot)$ implies

$$f_t(x_t) + \langle \nabla f_t(x_t), x_t^* - x_t\rangle + \lambda D_r(x_t^*, x_t) \le f_t(x_t^*). \quad (17)$$

By substituting (17) into (16), and using the fact that $D_r(\cdot,\cdot)$ is non-negative when the regularization function $r(\cdot)$ is convex, and also $\alpha \le \frac{1}{L}$, we obtain

$$f_t(x_{t+1}) \le f_t(x_t^*) - \lambda D_r(x_t^*, x_t) + \frac{1}{\alpha}D_r(x_t^*, x_t) - \frac{1}{\alpha}D_r(x_t^*, x_{t+1}). \quad (18)$$

Next, we use the result of [19], which states that for every λ-strongly convex function $f_t(\cdot)$, the following inequality holds:

$$f_t(x) - f_t(x_t^*) \ge \lambda D_r(x_t^*, x),\ \forall x \in \mathcal{X}, \quad (19)$$

We set $x = x_{t+1}$ in (19) and combine it with (18) to complete the proof. □

C. Proof of Theorem 2

The following lemma paves the way for our regret analysis leading to Theorem 2.

Lemma 12: Under the same convexity and smoothness condition stated in Theorem 2, let x_t be the sequence of decisions generated by OMMD. Then, the following bound holds:

$$\|x_{t+1} - x_t^*\| \le \sqrt{\frac{\mu'}{\mu}\beta^m}\|x_t - x_t^*\|, \quad (20)$$

where m represents the number of mirror descent steps per round, parameters μ and μ' denote the strong convexity and smoothness factors of the regularization function $r(\cdot)$, and β is the shrinking factor introduced in Lemma 1.

Proof. Using the result of Lemma 1, OMMD with m mirror descent steps guarantees

$$D_r(x_t^*, x_{t+1}) \le \beta^m D_r(x_t^*, x_t). \quad (21)$$

Since the regularization function $r(\cdot)$ is μ-strongly convex, we have

$$\frac{\mu}{2}\|x_t^* - x_{t+1}\|^2 \le r(x_t^*) - r(x_{t+1}) - \langle \nabla r(x_{t+1}), x_t^* - x_{t+1}\rangle. \quad (22)$$

In addition, smoothness of $r(\cdot)$ implies

$$r(x_t^*) - r(x_t) - \langle \nabla r(x_t), x_t^* - x_t\rangle \le \frac{\mu'}{2}\|x_t^* - x_t\|^2. \quad (23)$$

Combining the above with (21), and (22), and using the definition of Bregman divergence completes the proof. □

Now, we are ready to present the proof of Theorem 2. To bound the dynamic regret, we begin by using the strong convexity of the cost function $f_t(\cdot)$, i.e.,

$$\begin{aligned}
f_t(x_t) - f_t(x_t^*) &\le \langle \nabla f_t(x_t), x_t - x_t^*\rangle - \lambda D_r(x_t^*, x_t)\\
&\le \frac{1}{\alpha}\langle \nabla r(x_t) - \nabla r(z_{t+1}), x_t - x_t^*\rangle - \lambda D_r(x_t^*, x_t)\\
&\le \frac{1}{\alpha}\Big(D_r(x_t^*, x_t) - D_r(x_t^*, x_{t+1}) - D_r(x_{t+1}, z_{t+1})\\
&\quad + D_r(x_t, z_{t+1})\Big) - \lambda D_r(x_t^*, x_t)\\
&\le \frac{1}{\alpha}\Big(D_r(x_t, z_{t+1}) - D_r(x_{t+1}, z_{t+1})\Big)\\
&\quad + \Big(\frac{1}{\alpha} - \lambda\Big)D_r(x_t^*, x_t)\\
&\le \frac{1}{\alpha}\Big(D_r(x_t, z_{t+1}) - D_r(x_{t+1}, z_{t+1})\Big)\\
&\quad + \Big(\frac{1}{\alpha} - \lambda\Big)\Big(\frac{f_t(x_t) - f_t(x_t^*)}{\lambda}\Big),
\end{aligned} \quad (24)$$

where in the second line we have used Lemma 11, the third line is obtained by applying Lemma 9 and the fact of the Bregman projection property in (12), and the last line is obtained using the result of [19] stated in (19). Thus, if

$\alpha > \frac{1}{2\lambda}$, from (24) we have

$$f_t(x_t) - f_t(x_t^*) \leq \frac{\lambda}{2\alpha\lambda - 1}\Big(D_r(x_t, z_{t+1}) - D_r(x_{t+1}, z_{t+1})\Big)$$
$$\leq \frac{\lambda K}{2\alpha\lambda - 1}\|x_{t+1} - x_t\|$$
$$\leq \frac{\lambda K}{2\alpha\lambda - 1}\Big(\|x_{t+1} - x_t^*\| + \|x_t - x_t^*\|\Big)$$
$$\leq \frac{\lambda K}{2\alpha\lambda - 1}\Big(1 + \sqrt{\frac{\mu'}{\mu}\beta^m}\Big)\|x_t - x_t^*\|, \quad (25)$$

where we have used inequality (8) in the second line, and the last line follows from the result of Lemma 12. Summing (25) over time, we have

$$\mathrm{Reg}_T^{\mathrm{d}} \leq \frac{\lambda K}{2\alpha\lambda - 1}\sum_{t=1}^{T}\Big(1 + \sqrt{\frac{\mu'}{\mu}\beta^m}\Big)\|x_t - x_t^*\|. \quad (26)$$

Now, we proceed to bound $\sum_{t=1}^{T}\|x_t - x_t^*\|$ as follows:

$$\sum_{t=1}^{T}\|x_t - x_t^*\| = \|x_1 - x_1^*\| + \sum_{t=2}^{T}\|x_t - x_t^*\|$$
$$\leq \|x_1 - x_1^*\| + \sum_{t=2}^{T}\|x_t - x_{t-1}^*\| + \|x_{t-1}^* - x_t^*\|$$
$$\leq \|x_1 - x_1^*\| + \sum_{t=2}^{T}\sqrt{\frac{\mu'}{\mu}\beta^m}\|x_{t-1} - x_{t-1}^*\| + \|x_t^* - x_{t-1}^*\|. \quad (27)$$

In addition, if $m \geq \lceil\frac{1+\alpha\lambda}{2\alpha\lambda}\log(\frac{\mu'}{\mu})\rceil$, we have

$$\beta^{\frac{m}{2}} \leq \Big(1 - \frac{2\alpha\lambda}{1+\alpha\lambda}\Big)^{\frac{m}{2}} \leq \exp\left(\frac{-\alpha\lambda m}{1+\alpha\lambda}\right) < \sqrt{\mu/\mu'}. \quad (28)$$

Thus, $\sqrt{\beta^m\mu'/\mu} < 1$. Combining (26), (27), and (28), completes the proof. □

D. Proof of Theorem 3

In order to bound the dynamic regret, we begin by the generalized smoothness of the cost function $f_t(.)$, i.e.,

$$f_t(x_t) - f_t(x_t^*) \leq \langle\nabla f_t(x_t^*), x_t - x_t^*\rangle + LD_r(x_t, x_t^*)$$
$$\leq \|\nabla f_t(x_t^*)\|_*\|x_t - x_t^*\| + LD_r(x_t, x_t^*). \quad (29)$$

Next, we use the fact

$$\|\nabla f_t(x_t^*)\|_*\|x_t - x_t^*\| \leq \frac{\|\nabla f_t(x_t^*)\|_*^2}{2\theta} + \frac{\theta\|x_t - x_t^*\|^2}{2}, \quad (30)$$

for any arbitrary positive constant $\theta > 0$.

Furthermore, smoothness of the regularization function $r(\cdot)$ implies

$$D_r(x_t, x_t^*) \leq \frac{\mu'}{2}\|x_t^* - x_t\|^2. \quad (31)$$

By combining (29), (30), and (31), with $\theta = L$ we have

$$f_t(x_t) - f_t(x_t^*) \leq \frac{\|\nabla f_t(x_t^*)\|_*^2}{2L} + \frac{L(\mu'+1)\|x_t - x_t^*\|^2}{2}. \quad (32)$$

Summing (32) over time, we obtain

$$\mathrm{Reg}_T^{\mathrm{d}} \leq \sum_{t=1}^{T}\frac{\|\nabla f_t(x_t^*)\|_*^2}{2L} + \frac{L(\mu'+1)}{2}\sum_{t=1}^{T}\|x_t - x_t^*\|^2. \quad (33)$$

Now, we proceed by bounding $\sum_{t=1}^{T}\|x_t - x_t^*\|^2$ as follows:

$$\sum_{t=1}^{T}\|x_t - x_t^*\|^2 \leq \|x_1 - x_1^*\|^2$$
$$+ \sum_{t=2}^{T}\big(2\|x_t - x_{t-1}^*\|^2 + 2\|x_{t-1}^* - x_t^*\|^2\big)$$
$$\leq \|x_1 - x_1^*\|^2$$
$$+ 2\sum_{t=2}^{T}\big(\frac{\mu'}{\mu}\beta^m\|x_{t-1} - x_{t-1}^*\|^2 + \|x_{t-1}^* - x_t^*\|^2\big) \quad (34)$$

where we have applied the result of Lemma 12 in the third line. Furthermore, if $m \geq \lceil\frac{1+\alpha\lambda}{2\alpha\lambda}\log(\frac{\mu'}{2\mu})\rceil$, since $\beta < 1$ we have

$$\beta^m \leq \Big(1 - \frac{2\alpha\lambda}{1+\alpha\lambda}\Big)^m \leq \exp\left(\frac{-2\alpha\lambda m}{1+\alpha\lambda}\right) < \frac{\mu}{2\mu'}. \quad (35)$$

Thus, combining (33), (34), and (35) completes the proof. □

E. Proof of Theorem 5

Before presenting the proof of Theorem 5, we first state the following lemma.

Lemma 13: For H-smooth and non-negative function $g(x)$, the following condition holds:

$$\|\nabla g(x)\|_* \leq \sqrt{4Hf(x)}, \ \forall x \in \mathcal{X}. \quad (36)$$

Lemma 13 is proved in [15].

Now we continue to prove Theorem 5. The proof of Theorem 5 initially follows the first half of the proof of Theorem 2 until (25). Then, from (25) we have

$$f_t(x_t) - f_t(x_t^*) \leq \frac{\lambda}{2\alpha\lambda - 1}\left(D_r(x_t, z_{t+1}) - D_r(x_{t+1}, z_{t+1})\right). \quad (37)$$

Now we continue to bound $D_r(x_t, z_{t+1})$. By the definition of Bregman divergence, we have

$$D_r(x_t, z_{t+1}) + D_r(z_{t+1}, x_t)$$
$$= \langle\nabla r(x_t) - \nabla r(z_{t+1}), x_t - z_{t+1}\rangle = \langle\alpha\nabla f_t(x_t), x_t - z_{t+1}\rangle$$
$$\leq \|\alpha\nabla f_t(x_t)\|_*\|x_t - z_{t+1}\|$$
$$\leq \frac{\alpha^2}{2\mu}\|\nabla f_t(x_t)\|_*^2 + \frac{\mu}{2}\|x_t - z_{t+1}\|^2. \quad (38)$$

The strong convexity of the regularization function implies

$$\frac{\mu}{2}\|x_t - z_{t+1}\|^2 \leq D_r(z_{t+1}, x_t). \quad (39)$$

Combining the above with (38), we obtain

$$D_r(x_t, z_{t+1}) \leq \frac{\alpha^2}{2\mu}\|\nabla f_t(x_t)\|_*^2 \leq \frac{2L\mu'\alpha^2}{\mu}f_t(x_t), \quad (40)$$

where to obtain the right hand-side of (40) we have applied Lemma 13.

We substitute (40) into (37) to obtain

$$f_t(x_t) - f_t(x_t^*) \leq \frac{2L\mu'\alpha^2}{\mu}f_t(x_t). \quad (41)$$

By setting $\alpha < \sqrt{\frac{\mu}{2L\mu'}}$, from (41) we have

$$f_t(x_t) - f_t(x_t^*) \leq \frac{2L\mu'\alpha^2}{\mu - 2L\mu'\alpha^2}f_t(x_t^*). \quad (42)$$

Summing (42) over time completes the proof. □