Projection-Free Online Convex Optimization with Stochastic Constraints

Duksang Lee [1] Nam Ho-Nguyen [2] Dabeen Lee [1 *]

May 2, 2023

Abstract

This paper develops projection-free algorithms for online convex optimization with stochastic constraints. We design an online primal-dual projection-free framework that can take any projection-free algorithms developed for online convex optimization with no long-term constraint. With this general template, we deduce sublinear regret and constraint violation bounds for various settings. Moreover, for the case where the loss and constraint functions are smooth, we develop a primal-dual conditional gradient method that achieves $O(\sqrt{T})$ regret and $O(T^{3/4})$ constraint violations. Furthermore, for the setting where the loss and constraint functions are stochastic and strong duality holds for the associated offline stochastic optimization problem, we prove that the constraint violation can be reduced to have the same asymptotic growth as the regret.

1 Introduction

Online convex optimization (OCO) is a widely used framework for decision-making under uncertainty. In OCO, the decision-maker attempts to minimize a sequence of convex loss functions chosen by the adversarial environment. At each iteration, the decision-maker chooses a decision without knowing the loss function, after which the associated loss is revealed by the environment. Based on these repeated interactions, the decision-maker adapts to the environment so as to minimize the total cumulative loss. As the description suggests, the OCO framework is useful for designing iterative solution methods for optimizing a complex system even under limited information.

OCO with *long-term constraints* [15] extends the OCO framework. When the problem involves complex functional constraints, projection onto the feasible set can be difficult. For such scenarios, an alternate way is to take and aggregate the constraint functions and then require satisfying the aggregated constraint in the long run. For resource planning, there is a budget for each period, but budgets can be pooled over multiple time periods, over which

[1]Department of Industrial and Systems Engineering, KAIST, Daejeon 34141, Republic of Korea
[2]Discipline of Business Analytics, The University of Sydney, Sydney, NSW 2006, Australia

*Correspondence to <dabeenl@kaist.ac.kr>

resource allocation is flexible. The framework is given as the following online optimization model.

$$\min_{\boldsymbol{x_1},\ldots,\boldsymbol{x_T}\in\mathcal{X}} \quad \sum_{t=1}^{T} f_t(\boldsymbol{x_t}) \quad \text{s.t.} \quad \sum_{t=1}^{T} g_t(\boldsymbol{x_t}) \leq 0 \tag{1}$$

where $\{f_t\}_{t=1}^T$ and $\{g_t\}_{t=1}^T$ are the convex loss and constraint functions over T time periods. Here, the decision-maker selects $\boldsymbol{x_t}$ from a domain $\mathcal{X}$ based on the history up to time step t before observing f_t and g_t. The setting where the constraint functions $g_1, \ldots, g_T$ in (1) are independent and identically distributed (i.i.d.) with an unknown probability distribution is referred to as OCO with *stochastic constraints* [24].

The existing algorithms for OCO with long-term constraints and stochastic constraints, however, still require projections onto the domain $\mathcal{X}$. The online primal-dual augmented Lagrangian algorithm [12, 15] and the drift-plus-penalty algorithm [17, 23] provide the two main algorithmic frameworks for OCO with long-term constraints, and both are variants of the primal-dual projected gradient method, requiring projection onto the domain at each iteration. When $\mathcal{X}$ is given by linear inequalities, projection onto it boils down to solving a quadratic program in the general case. Matrix completion for recommender systems requires projection onto a spectahedron [10].

There has been a surge of interest in projection-free algorithms based on the famous Frank-Wolfe method, replacing each projection step by a linear optimization over $\mathcal{X}$ [10]. However, as far as we know, no projection-free algorithm exists for OCO with long-term constraints. Motivated by this, we develop projection-free algorithms for online convex optimization with stochastic constraints.

Our Contributions

1. We design an online primal-dual projection-free learning framework for OCO with stochastic constraints. The framework works as a general template, and it can take any projection-free algorithm developed for OCO with no long-term constraint. In particular, we apply the framework to different settings depending on (1) whether the loss and constraint functions are smooth or not and (2) whether the loss functions are arbitrary or stochastic. We provide sublinear regret and constraint violation bounds for various settings.

2. For the case where the loss and constraint functions are smooth, we develop Primal-Dual Meta-Frank-Wolfe, which is a variant of the Meta-Frank-Wolfe algorithm of [3]. Subject to a per-iteration cost of $O(\sqrt{T})$, the algorithm guarantees $O(\sqrt{T})$ regret and $O(T^{3/4})$ constraint violation.

3. When the loss functions are also stochastic and strong duality of the associated offline stochastic optimization problem is satisfied, we prove that the algorithms achieve smaller constraint violations. More precisely, we may reduce the constraint violation to have the same asymptotic growth as the regret if the strong duality assumption holds and the loss functions are i.i.d. with a probability distribution.

Our results are summarized in Table 1. In column “Setting", “Adversarial" means that the loss functions are adversarially chosen as the standard OCO framework, and “Stochastic" means that the loss functions are i.i.d. with an unknown probability distribution.

	Functions	Setting	Regret	Constraint violation	Per-round cost
Alg. 1	Non-smooth	Adversarial	$T^{5/6+\beta}$	$T^{11/12-\beta/2}$	1
	Non-smooth	Stochastic, SD	$T^{5/6}$	$T^{5/6}$	1
	Smooth	Adversarial	$T^{4/5+\beta}$	$T^{9/10-\beta/2}$	1
	Smooth	Stochastic	$T^{3/4+\beta}$	$T^{7/8-\beta/2}$	1
	Smooth	Stochastic, SD	$T^{3/4}$	$T^{3/4}$	1
Alg. 1	Non-smooth	Adversarial	$T^{3/4+\beta}$	$T^{7/8-\beta/2}$	T
	Non-smooth	Stochastic, SD	$T^{3/4}$	$T^{3/4}$	T
Alg. 2	Smooth	Adversarial	$T^{1/2+\beta}$	$T^{3/4-\beta/2}$	$\sqrt{T}$
	Smooth	Stochastic, SD	$T^{1/2}$	$T^{1/2}$	$\sqrt{T}$

Table 1: Regret and constraint violation bounds for various settings. Here SD means strong duality holds for the associated stochastic optimization problem; see Section 6.

2 Related Work

Our work is related to the literature on online convex optimization with long-term constraints and the projection-free online learning and optimization literature.

OCO with Long-Term Constraints The most general setting is where $\{g_t\}_{t=1}^T$ is adversarially chosen and the benchmark $\boldsymbol{x}^*$ is set to an optimal fixed solution of (1). However, for the general setting, it is known that it is impossible to simultaneously bound the regret and constraint violation by a sublinear function in T [16]. Mahdavi et al. [15] then considered the special case where $g_t = g$ for some fixed function g for all $t \in [T]$, and they provided an augmented-Lagrangian-based algorithm that achieves $O(\sqrt{T})$ regret and $O(T^{3/4})$ constraint violation. Jenatton et al. [12] modified the algorithm to obtain $O(T^{\max\{\beta,1-\beta\}})$ regret and $O(T^{1-\beta/2})$ constraint violation where $\beta \in (0,1)$ is an algorithm parameter.

Yu et al. [24] considered the case where g_t's are time-varying but are i.i.d. with an unknown probability distribution, i.e., stochastic constraints. They provided the drift-plus-penalty (DPP) algorithm which guarantees $O(\sqrt{T})$ expected regret and $O(\sqrt{T})$ expected constraint violation under Slater's condition. Wei et al. [20] gave a variant of DPP attaining the same asymptotic performance under a more general strong duality assumption.

Liakopoulos et al. [14], Neely and Yu [17], Valls et al. [19] consider adversarially chosen constraint functions but they set different benchmarks with more restrictions. Recently, Castiglioni et al. [2] came up with a unifying framework that works for both stochastic and adversarial constraint functions. Guo et al. [7], Yi et al. [22], Yuan and Lamperski [25] studied the notion of *cumulative constraint violation*, given by $\sum_{t=1}^T [g_t(\boldsymbol{x_t})]_+$ where $[a]_+ = \max\{a, 0\}$ over $a \in \mathbb{R}$, instead of imposing long-term constraints.

Projection-Free Online Learning and Optimization Hazan and Kale [10] developed the first online projection-free algorithm, based on the Frank-Wolfe method [5], that guarantees $O(T^{3/4})$ regret for non-smooth loss functions. Garber and Hazan [6] provided an improved algorithm for smooth and strongly convex loss functions. After these works, there has been many results on projection-free algorithms for online convex optimization with no long-term

Loss function	Setting	Regret	Per-round cost	Reference
Non-smooth	Adversarial	$T^{3/4}$	1	OCG [9]
Smooth	Adversarial	$T^{2/3}$	1	OSPF [11]
Smooth	Stochastic	$T^{1/2}$	1	ORGFW [21]
Non-smooth	Adversarial	$T^{1/2}$	T	SFTRL [11]
Smooth	Adversarial	$T^{1/2}$	$\sqrt{T}$	MFW [3]

Table 2: Regret bounds for projection-free algorithms for OCO with no long-term constraint.

constraint, and Table 2 summarizes the known regret bounds for various settings.

3 Online Convex Optimization with Stochastic Constraints

Let $\mathcal{X} \subseteq \mathbb{R}^d$ be a known fixed compact convex set. Let $f_1, \ldots, f_T : \mathbb{R}^d \to \mathbb{R}$ be a sequence of arbitrary convex loss functions. Let $\bar{g}(\boldsymbol{x}) = \mathbb{E}_{\boldsymbol{\omega}}\left[g(\boldsymbol{x}, \boldsymbol{\omega})\right] : \mathbb{R}^d \to \mathbb{R}$ be a function where $g(\boldsymbol{x}, \boldsymbol{\omega})$ is convex with respect to $\boldsymbol{x} \in \mathcal{X}$ and the expectation is taken with $\boldsymbol{\omega} \in \Omega$ from an unknown distribution.

Like Yu et al. [24], we take the benchmark decision $\boldsymbol{x}^*$ defined as an optimal solution to

$$\min_{\boldsymbol{x} \in \mathcal{X}} \sum_{t=1}^{T} f_t(\boldsymbol{x}) \ \text{ s.t. } \ \bar{g}(\boldsymbol{x}) = \mathbb{E}_{\boldsymbol{\omega}}\left[g(\boldsymbol{x}, \boldsymbol{\omega})\right] \le 0. \tag{2}$$

Instead of direct access to $\bar{g}$, we are presented constraint functions $g_1, \ldots, g_T : \mathbb{R}^d \to \mathbb{R}$ where $g_t(\boldsymbol{x}) := g(\boldsymbol{x}, \boldsymbol{\omega_t})$ for $t \in [T]$ where $\boldsymbol{\omega_1}, \ldots, \boldsymbol{\omega_T}$ are i.i.d. samples of $\boldsymbol{\omega}$.

Our goal is to design an algorithm for choosing $\boldsymbol{x_t}$, $t \in [T]$ that guarantees a sublinear regret against the benchmark $\boldsymbol{x}^*$ and a sublinear constraint violation at the same time, under the condition that each $\boldsymbol{x_t}$ only depends on functions from previous time steps $f_1, \ldots, f_{t-1}, g_1, \ldots, g_{t-1}$. Here, the regret and constraint violation are defined as follows:

$$\text{regret}(T) = \sum_{t=1}^{T} f_t(\boldsymbol{x_t}) - \sum_{t=1}^{T} f_t(\boldsymbol{x}^*), \quad \text{violation}(T) = \sum_{t=1}^{T} g_t(\boldsymbol{x_t}).$$

We focus on the single constraint setting for simplicity, but our framework easily extends to multiple constraints. We assume that each loss f_t is chosen adversarially, but is independent of $\boldsymbol{\omega_s}$ for $s \ge t+1$. In other words, f_t can be chosen with full knowledge of the history up to time t, but *not* future random realizations.

We also consider the case when the loss functions are stochastic i.i.d. realizations of some random function $f(\boldsymbol{x}, \boldsymbol{\omega})$, i.e., f_t is given by $f_t(\boldsymbol{x}) = f(\boldsymbol{x}, \boldsymbol{\omega_t})$ where $f(\boldsymbol{x}, \boldsymbol{\omega})$ is convex with respect to $\boldsymbol{x} \in \mathcal{X}$. Note here that the random variable $\boldsymbol{\omega}$ is the same as the one that appears in g_t, meaning that f_t and g_t are possibly dependent. A direct application of this setting is *stochastic constrained stochastic optimization*, that is formulated as the following optimization problem.

$$\min_{\boldsymbol{x} \in \mathcal{X}} \quad \bar{f}(\boldsymbol{x}) = \mathbb{E}_{\boldsymbol{\omega}}\left[f(\boldsymbol{x}, \boldsymbol{\omega})\right] \quad \text{s.t.} \quad \bar{g}(\boldsymbol{x}) \le 0.$$

An iterative algorithm would obtain an i.i.d. sample $\boldsymbol{\omega_t}$ of $\boldsymbol{\omega}$ at each iteration t and consider $f_t = f_t(\cdot, \boldsymbol{\omega_t})$ and $g_t = g_t(\cdot, \boldsymbol{\omega_t})$. Given $\{\boldsymbol{x_t}\}_{t=1}^T$, we may obtain $\bar{\boldsymbol{x}}_{\boldsymbol{T}} = (1/T)\sum_{t=1}^T \boldsymbol{x_t}$. By Jensen's inequality, the optimality gap

Algorithm 1 Online Primal-Dual Projection-Free Learning Framework

Initialize: time horizon T, number of blocks Q, block size K, initial iterates $\boldsymbol{x_1} \in \mathcal{X}$, $\lambda_1 = 0$, the number of blocks Q, step size μ, augmentation parameter θ, and a projection-free convex optimization oracle $\mathcal{E}$.
for $q = 1$ **to** Q **do**
  for $k = 1$ **to** K **do**
    Observe f_t and g_t.
    Use oracle $\mathcal{E}$ to obtain $\boldsymbol{x_{t+1}}$ based on $h_t(\boldsymbol{x}) = f_t(\boldsymbol{x}) + \lambda_q g_t(\boldsymbol{x})$ and $\boldsymbol{x} = \boldsymbol{x_t}$.
    Set $t \leftarrow t+1$.
  end for
  Set $\lambda_{q+1} = \left[(1-\theta\mu)\lambda_q + \mu \sum_{t=(q-1)K+1}^{qK} g_t(\boldsymbol{x_t})\right]_+$.
end for

and constraint violation of $\bar{\boldsymbol{x}}_{\boldsymbol{T}}$ are given by

$$\mathbb{E}\left[\bar{f}(\bar{\boldsymbol{x}}_{\boldsymbol{T}})\right] - \bar{f}(\boldsymbol{x}) \leq \frac{1}{T}\mathbb{E}\left[\sum_{t=1}^{T} f_t(\boldsymbol{x_t}) - \sum_{t=1}^{T} f_t(\boldsymbol{x})\right], \quad \bar{g}(\bar{\boldsymbol{x}}_{\boldsymbol{T}}) \leq \frac{1}{T}\mathbb{E}\left[\sum_{t=1}^{T} g_t(\boldsymbol{x_t})\right].$$

4 Online Primal-Dual Projection-Free Learning Framework

Algorithm 1 provides a projection-free algorithmic framework for online convex optimization with stochastic constraints. The algorithm is a general template that can take any projection-free algorithm for online convex optimization with no long-term constraint. This allows us to use different projection-free algorithms depending on the structure of the loss and constraint functions.

The idea behind Algorithm 1 is as follows. First, we break the time horizon T into Q blocks. Each block has K time steps. Then for each of the Q blocks, we use as an oracle a projection-free algorithm developed for online convex optimization with no long-term constraint. To be specific, for block $q \in [Q]$, the oracle is applied to

$$f_t(\boldsymbol{x}) + \lambda_q g_t(\boldsymbol{x})$$

which is the loss function f_t at time t penalized by the constraint function g_t where λ_q is the penalty parameter for block q. Once iterations in block q are completed, we update the penalty parameter λ_q (the dual variable) based on the constraint function values realized in block q. The update rule is motivated by the online primal-dual augmented Lagrangian algorithm due to Jenatton et al. [12], Mahdavi et al. [15], and it makes use of the following augmented Lagrangian function

$$L_t(\boldsymbol{x}, \lambda) = f_t(\boldsymbol{x}) + \lambda g_t(\boldsymbol{x}) - \frac{\theta}{2K}\lambda^2.$$

Given functions f_t, g_t for time steps $t = (q-1)K+1, \ldots, qK$ observed in block q, Algorithm 1 then updates the penalty parameter using the gradient ascent-type update

$$\lambda_{q+1} = \left[\lambda_q + \mu \sum_{t=(q-1)K+1}^{qK} \nabla_\lambda L_t(\boldsymbol{x_t}, \lambda_q)\right]_+ = \left[(1-\theta\mu)\lambda_q + \mu \sum_{t=(q-1)K+1}^{qK} g_t(\boldsymbol{x_t})\right]_+,$$

where μ is a step size.

Throughout this section, we work over the ℓ_2 norm $\|\cdot\|_2$ in $\mathbb{R}^d$ for simplicity.

Definition 1 (Lipschitz continuity)**.** *We say that a function $h : \mathcal{X} \to \mathbb{R}$ is D-Lipschitz for some $D \geq 0$ if $\|\nabla h(\boldsymbol{x})\|_2 \leq D$ for all $\boldsymbol{x} \in \mathcal{X}$.*

Definition 2 (Smoothness)**.** *We say that a function $h : \mathcal{X} \to \mathbb{R}$ is L-smooth for some $L \geq 0$ if $\|\nabla h(\boldsymbol{x}) - \nabla h(\boldsymbol{y})\|_2 \leq L\|\boldsymbol{x} - \boldsymbol{y}\|_2$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$.*

We formally define the notion of oracle $\mathcal{E}$ used as a subroutine in Algorithm 1.

Definition 3 (Oracle)**.** *We say that $\mathcal{E}$ is an (α, C_0, C_1, C_2)-oracle for some $\alpha \in (0, 1)$ and $C_0, C_1, C_2 \geq 0$ if for any sequence of (adversarial or stochastic) convex loss functions $h_1, \ldots, h_K$ that are D-Lipschitz and L-smooth over domain $\mathcal{X} \subseteq \mathbb{R}^d$, oracle $\mathcal{E}$ guarantees*

$$\mathbb{E}\left[\sum_{k=1}^{K} h_k(\boldsymbol{x_k}) - \min_{x\in\mathcal{X}} \sum_{k=1}^{K} h_k(\boldsymbol{x})\right] \leq (C_0 + C_1 D + C_2 L)K^{\alpha}$$

where the expectation is taken over the randomness of oracle $\mathcal{E}$ itself and the randomness of the convex loss functions. Constants C_0, C_1, C_2 are independent of parameters D, L, K.

Assumption 1. *When a function is non-smooth, we assume that it is ∞-smooth. Moreover, we assume that $0 \cdot \infty = 0$.*

Remark 4.1. *Let $\mathcal{E}$ be an (α, C_0, C_1, C_2)-oracle for some $\alpha \in (0, 1)$ and $C_0, C_1, C_2 \geq 0$. If $C_2 = 0$, then $\mathcal{E}$ guarantees an $O(K^{\alpha})$ regret for any Lipschitz loss functions that can be non-smooth. If $C_2 > 0$, then $\mathcal{E}$ guarantees an $O(K^{\alpha})$ regret only if the loss functions are smooth.*

If we are given an (α, C_0, C_1, C_2)-oracle, then we set the parameters of Algorithm 1 as follows:

$$Q = T^{\frac{2-2\alpha}{3-2\alpha}}, K = T^{\frac{1}{3-2\alpha}}, \beta \in \left[0, \frac{1-\alpha}{3-2\alpha}\right], \theta = 3(C_1 D + C_2 L)T^{\frac{\alpha}{3-2\alpha}-\beta}, \mu = \frac{1}{\theta(Q+1)}. \tag{3}$$

We assume that both $T^{\frac{2-2\alpha}{3-2\alpha}}$ and $T^{\frac{1}{3-2\alpha}}$ are integers. Even if they are not, we can take the ceiling $\lceil\cdot\rceil$ as needed, and our framework still achieves the same asymptotic guarantees.

Theorem 1. *Suppose that loss functions $f_1, \ldots, f_T$ and stochastic constraint functions $g_1, \ldots, g_T$ are D-Lipschitz and L-smooth. If $\mathcal{E}$ is an (α, C_0, C_1, C_2)-oracle for some $\alpha \in (0, 1)$ and $C_0, C_1, C_2 \geq 0$, then Algorithm 1 whose parameters are set as in (3) guarantees that*

$$\mathbb{E}\left[\sum_{t=1}^{T} f_t(\boldsymbol{x_t}) - \min_{x\in\mathcal{X}} \sum_{t=1}^{T} f_t(\boldsymbol{x})\right] = O\left(T^{\frac{2-\alpha}{3-2\alpha}+\beta}\right), \quad \mathbb{E}\left[\sum_{t=1}^{T} g_t(\boldsymbol{x_t})\right] = O\left(T^{\frac{5-3\alpha}{6-4\alpha}-\frac{\beta}{2}}\right)$$

where the expectations are taken over the randomness of oracle $\mathcal{E}$ and the randomness of the loss and constraint functions.

In particular, based on projection-free algorithms for OCO with no long-term constraint as in Table 2, we deduce results in Table 1. The setting where functions are smooth and $O(\sqrt{T})$ per-round cost is allowed and the settings under strong duality are considered in the next sections.

Algorithm 2 Primal-Dual Meta-Frank-Wolfe (PDMFW)

Initialize: initial iterates $\boldsymbol{x_1} \in \mathcal{X}$, $\lambda_1 = 0$, step size μ, augmentation parameter θ, inner loop length K, projection-free linear optimization oracles $\mathcal{E}^1, \ldots, \mathcal{E}^K$, and step sizes $\gamma_1, \ldots, \gamma_K$.

for $t = 1$ **to** T **do**

Observe f_t and g_t.

Primal update: obtain $\boldsymbol{x_{t+1}}$ as follows

Set $\boldsymbol{x^1_{t+1}} = \boldsymbol{x_1}$

for $k = 1$ **to** K **do**

Use oracle $\mathcal{E}^k$ (Algorithm 3) to obtain $\boldsymbol{v^k_{t+1}}$.

Update

$$\boldsymbol{x^{k+1}_{t+1}} = \boldsymbol{x^k_{t+1}} + \gamma_k \left(\boldsymbol{v^k_{t+1}} - \boldsymbol{x^k_{t+1}}\right)$$

end for

Set

$$\boldsymbol{x_{t+1}} = \boldsymbol{x^{K+1}_{t+1}}$$

Dual update:

$$\lambda_{t+1} = [(1 - \theta\mu)\lambda_t + \mu g_t(\boldsymbol{x_t})]_+$$

end for

5 Primal-Dual Meta-Frank-Wolfe for Smooth Functions

For the setting where the loss and constraint functions are smooth, Algorithm 1 guarantees $O(T^{3/4+\beta})$ regret and $O(T^{7/8-\beta/2})$ constraint violation. In this section, we develop Algorithm 2 which provides $O(T^{1/2+\beta})$ regret and $O(T^{3/4-\beta/2})$ constraint violation, where we use $O(\sqrt{T})$ gradient evaluations per time step.

Algorithm 2 is a combination of Meta-Frank-Wolfe [3] for projection-free online convex optimization (with no long-term constraint) and the online primal-dual gradient method [12, 15]. We refer to Algorithm 2 as Primal-Dual Meta-Frank-Wolfe (PDMFW). In constrast to Algorithm 1 that updates the dual variable only when a new block starts, Algorithm 2 updates the dual variable for every time step.

As in [12, 15], PDMFW works over the following *augmented Lagrangian function*. Upon observing f_t and g_t at time t, we take

$$L_t(\boldsymbol{x}, \lambda) = f_t(\boldsymbol{x}) + \lambda g_t(\boldsymbol{x}) - \frac{\theta}{2}\lambda^2 \tag{4}$$

to compute the next iterate $\boldsymbol{x_{t+1}}$. At a high level, PDMFW is an online primal-dual framework based on the augmented Lagrangian function (4) that applies a Frank-Wolfe subroutine for the primal update and gradient ascent for the dual update.

The Frank-Wolfe subroutine replaces the projection-based primal update of the online primal-dual gradient method. Starting from $\boldsymbol{x^1_{t+1}} = \boldsymbol{x_t}$ at time t, the Frank-Wolfe procedure runs with K steps and generates $\boldsymbol{x^2_{t+1}}, \ldots, \boldsymbol{x^{K+1}_{t+1}}$. Then we set $\boldsymbol{x_{t+1}} = \boldsymbol{x^{K+1}_{t+1}}$. The Frank-Wolfe update at each step k is given by $\boldsymbol{x^{k+1}_{t+1}} = \boldsymbol{x^k_{t+1}} + \gamma_k \left(\boldsymbol{v^k_{t+1}} - \boldsymbol{x^k_{t+1}}\right)$

Algorithm 3 Follow-The-Perturbed-Leader for $\mathcal{E}^k$

Choose a random perturbation vector $\boldsymbol{p}$ from $[0, \delta]^d$ uniformly at random where

$$\delta = \frac{1}{2D\sqrt{d}T^{\frac{1}{2}+\beta}}$$

for $t = 1$ **to** T **do**

Choose

$$\boldsymbol{v}_{\boldsymbol{t+1}}^{\boldsymbol{k}} \in \operatorname*{argmin}_{\boldsymbol{v}\in\mathcal{X}} \left\{ \boldsymbol{p}^\top \boldsymbol{v} + \sum_{s=1}^{t-1} h_s^k(\boldsymbol{v}) \right\}$$

where

$$h_s^k(\boldsymbol{v}) := \left(\nabla f_s(\boldsymbol{x}_{\boldsymbol{s}}^{\boldsymbol{k}}) + \lambda_s \nabla g_s(\boldsymbol{x}_{\boldsymbol{s}}^{\boldsymbol{k}})\right)^\top \boldsymbol{v}.$$

end for

for some step size γ_k. Here, the direction $\boldsymbol{v}_{\boldsymbol{t+1}}^{\boldsymbol{k}}$ would ideally be a vector $\boldsymbol{v} \in \mathcal{X}$ minimizing

$$\nabla_{\boldsymbol{x}} L_{t+1}(\boldsymbol{x}_{\boldsymbol{t+1}}^{\boldsymbol{k}}, \lambda_{t+1})^\top \boldsymbol{v} = \left(\nabla f_{t+1}(\boldsymbol{x}_{\boldsymbol{t+1}}^{\boldsymbol{k}}) + \lambda_{t+1} \nabla g_{t+1}(\boldsymbol{x}_{\boldsymbol{t+1}}^{\boldsymbol{k}})\right)^\top \boldsymbol{v}.$$

However, as functions f_{t+1} and g_{t+1} are only revealed after choosing $\boldsymbol{v}_{\boldsymbol{t+1}}^{\boldsymbol{k}}$, we use a projection-free online linear optimization oracle $\mathcal{E}^k$ to obtain $\boldsymbol{v}_{\boldsymbol{t+1}}^{\boldsymbol{k}}$ based on the history up to t. This idea was first introduced by Chen et al. [4]. For $k = 1, \ldots, K$, we set

$$\gamma_k = \frac{2}{k+1}.$$

For the dual update, we follow the update rule of the online primal-dual gradient method, that is,

$$\lambda_{t+1} = \left[\lambda_t + \mu \nabla_\lambda L_t(\boldsymbol{x}_{\boldsymbol{t}}^{\boldsymbol{k}}, \lambda_t)\right]_+ = \left[(1-\theta\mu)\lambda_t + \mu g_t(\boldsymbol{x}_{\boldsymbol{t}})\right]_+$$

where μ is a step size. For a fixed $\beta \in (0, 1/2)$, we set the parameters as follows:

$$K = \lfloor T^{\frac{1}{2}+\beta} \rfloor, \quad \theta = \frac{12RD\sqrt{d}}{T^{\frac{1}{2}+\beta}}, \quad \mu = \frac{1}{\theta(T+2)}. \tag{5}$$

Here, we may set any value between 0 and 1 for β. We will show that the (expected) regret of PDMFW is $O(T^{\frac{1}{2}+\beta})$ and the (expected) constraint violation is $O(T^{\frac{3}{4}-\frac{\beta}{2}})$.

For a projection-free online linear optimization oracle, we use the *Follow-The-Perturbed-Leader (FTPL)* algorithm [8, 13], given as in Algorithm 3.

The regret of FTPL for online linear optimization has a dependence on T bounded above by $O(\sqrt{T})$ [13]. However, the coefficient $\nabla_{\boldsymbol{x}} L_t(\boldsymbol{x}_{\boldsymbol{t}}, \lambda_t) = \nabla f_t(\boldsymbol{x}_{\boldsymbol{t}}^{\boldsymbol{k}}) + \lambda_t \nabla g_t(\boldsymbol{x}_{\boldsymbol{t}}^{\boldsymbol{k}})$ for time t has dual variable λ_t, so the regret grows as a function of $\lambda_1, \ldots, \lambda_T$. Therefore, we need a refined regret analysis of FTPL to show how the regret grows as a function of the dual variables $\lambda_1, \ldots, \lambda_T$.

We remark that Algorithm 2 is similar to the OSPHG algorithm of Sadeghi and Fazel [18] developed for online DR-submodular maximization, but modified to be projection-free.

Throughout this section, we work over the ℓ_1 norm $\|\cdot\|_1$ in $\mathbb{R}^d$ and its dual $\|\cdot\|_\infty$, the ℓ_∞ norm.

Assumption 2 (Basic assumptions)**.** *There are positive constants D, G, R satisfying the following.*

- $\|\nabla f_t(\boldsymbol{x})\|_\infty, \|\nabla g_t(\boldsymbol{x})\|_\infty \leq D$ *and* $g_t(\boldsymbol{x}) \leq G$ *for all* $t \in [T]$ *and* $\boldsymbol{x} \in \mathcal{X}$.
- $\|\boldsymbol{x} - \boldsymbol{y}\|_1 \leq R$ *for all* $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$.

Assumption 3 (Smoothness)**.** *There exists a positive constant L such that*

$$\|\nabla f_t(\boldsymbol{x}) - \nabla f_t(\boldsymbol{y})\|_\infty, \|\nabla g_t(\boldsymbol{x}) - \nabla g_t(\boldsymbol{y})\|_\infty \leq L \|\boldsymbol{x} - \boldsymbol{y}\|_1$$

for all $t \in [T]$ *and* $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$.

We first provide an upper bound on the expected regret of FTPL (Algorithm 3).

Lemma 5.1. *For each $k = 1, \ldots, K$, the expected regret of FTPL (Algorithm 3) under linear functions $h_1^k(\boldsymbol{v}), \ldots, h_T^k(\boldsymbol{v})$ (defined in Algorithm 3) is bounded above by*

$$\mathcal{R}^{\mathcal{E}}(T) := RD\sqrt{d}\left(3T^{\frac{1}{2}+\beta} + \frac{1}{T^{\frac{1}{2}+\beta}}\mathbb{E}\left[\sum_{t=1}^{T} \lambda_t^2\right]\right).$$

Theorem 2 gives bounds on the expected regret and the expected long-term constraint violation under Algorithm 2, respectively. Let constants C_1, C_2 be defined as

$$C_1 = 4RD + 5R^2L + \frac{R(4D + 5RL)^2}{4D\sqrt{d}} + 3RD\sqrt{d} + \frac{G^2}{12RD\sqrt{d}},$$
$$C_2 = \sqrt{96RD\sqrt{d}\,(RD + C_1)}.$$

Theorem 2. *Suppose that the loss functions $f_1, \ldots, f_T$ and the stochastic constraint functions $g_1, \ldots, g_T$ satisfy Assumptions 2 and 3. Then Algorithm 2 with parameters set according to (5) guarantees that*

$$\mathbb{E}\left[\sum_{t=1}^{T} f_t(\boldsymbol{x_t}) - \min_{x \in \mathcal{X}} \sum_{t=1}^{T} f_t(\boldsymbol{x})\right] \leq C_1 T^{\frac{1}{2}+\beta}, \quad \mathbb{E}\left[\sum_{t=1}^{T} g_t(\boldsymbol{x_t})\right] \leq C_2 T^{\frac{3}{4}-\frac{\beta}{2}}$$

where the expectations are taken over the randomness of $\mathcal{E}^1, \ldots, \mathcal{E}^K$ and the randomness of the loss and constraint functions.

6 Stochastic Loss Functions under Strong Duality

In this section, we focus on stochastic loss functions, i.e., $f_t(\boldsymbol{x}) = f(\boldsymbol{x}, \boldsymbol{\omega_t})$ in addition to stochastic constraint functions $g_t(\boldsymbol{x}) = g(\boldsymbol{x}, \boldsymbol{\omega_t})$. Recall that $\bar{f}(\boldsymbol{x}) = \mathbb{E}[f(\boldsymbol{x}, \boldsymbol{\omega})]$ and $\bar{g}(\boldsymbol{x}) = \mathbb{E}[g(\boldsymbol{x}, \boldsymbol{\omega})]$. We show that we can obtain improved bounds under strong Lagrangian duality of the following optimization problem:

$$\min_{\boldsymbol{x} \in \mathcal{X}} \left\{\bar{f}(\boldsymbol{x}) : \ \bar{g}(\boldsymbol{x}) \leq 0\right\}. \tag{6}$$

The Lagrangian of this is

$$\bar{L}(\boldsymbol{x}, \lambda) := \bar{f}(\boldsymbol{x}) + \lambda \bar{g}(\boldsymbol{x})$$

where $\boldsymbol{x} \in \mathcal{X}$ and $\lambda \geq 0$, and we assume that there exist $\boldsymbol{x}^* \in \mathcal{X}$ and $\lambda^* \geq 0$ such that

$$\bar{L}(\boldsymbol{x}^*, \lambda) \leq \bar{L}(\boldsymbol{x}^*, \lambda^*) \leq \bar{L}(\boldsymbol{x}, \lambda^*), \quad \forall \boldsymbol{x} \in \mathcal{X}, \lambda \geq 0, \tag{7}$$

i.e., strong duality holds for (6). This is satisfied, for example, under Slater constraint qualification, when there exists $\hat{\boldsymbol{x}} \in \mathcal{X}$ with $\bar{g}(\hat{\boldsymbol{x}}) = \mathbb{E}\left[g(\hat{\boldsymbol{x}}, \boldsymbol{\omega})\right] < 0$ (see Bertsekas [1, Proposition 5.1.6]), but our analysis allows for more general settings where strong duality holds but Slater constraint qualification may not hold.

Lemma 6.1. *Let $\boldsymbol{x_1}, \ldots, \boldsymbol{x_T}$ and $\lambda_1, \ldots, \lambda_Q$ be the decisions and dual variables chosen by Algorithm 1. When (7) holds we have*

$$\mathbb{E}\left[\sum_{t=1}^{T} g_t(\boldsymbol{x_t})\right] \leq \mathbb{E}\left[\sum_{q=1}^{Q}\sum_{k=1}^{K}\left(f_{(q-1)K+k}(\boldsymbol{x_{(q-1)K+k}}) + (\lambda^* + 1) g_{(q-1)K+k}(\boldsymbol{x_{(q-1)K+k}})\right)\right] - \mathbb{E}\left[\sum_{q=1}^{Q}\sum_{k=1}^{K}\left(f_{(q-1)K+k}(\boldsymbol{x}^*) + \lambda_q g_{(q-1)K+k}(\boldsymbol{x}^*)\right)\right].$$

Lemma 6.2. *Let $\boldsymbol{x_1}, \ldots, \boldsymbol{x_T}$ and $\lambda_1, \ldots, \lambda_T$ be the decisions and dual variables chosen by Algorithm 2. When (7) holds we have*

$$\mathbb{E}\left[\sum_{t=1}^{T} g_t(\boldsymbol{x_t})\right] \leq \mathbb{E}\left[\sum_{t=1}^{T}\left(f_t(\boldsymbol{x_t}) + (\lambda^* + 1) g_t(\boldsymbol{x_t})\right) - \sum_{t=1}^{T}\left(f_t(\boldsymbol{x}^*) + \lambda_t g_t(\boldsymbol{x}^*)\right)\right].$$

Using these bounds, we deduce the following results.

Theorem 3. *Suppose that the stochastic loss functions $f_1, \ldots, f_T$ and stochastic constraint functions $g_1, \ldots, g_T$ are D-Lipschitz and L-smooth. Furthermore, (7) is satisfied. If $\mathcal{E}$ is an (α, C_0, C_1, C_2)-oracle for some $\alpha \in (0, 1)$ and $C_0, C_1, C_2 \geq 0$, then Algorithm 1 whose parameters are set as in (3) with $\beta = 0$ guarantees that*

$$\mathbb{E}\left[\sum_{t=1}^{T} f_t(\boldsymbol{x_t}) - \min_{x \in \mathcal{X}} \sum_{t=1}^{T} f_t(\boldsymbol{x})\right] = O\left(T^{\frac{2-\alpha}{3-2\alpha}}\right), \quad \mathbb{E}\left[\sum_{t=1}^{T} g_t(\boldsymbol{x_t})\right] = O\left(T^{\frac{2-\alpha}{3-2\alpha}}\right)$$

where the expectations are taken over the randomness of oracle $\mathcal{E}$ and the randomness of the loss and constraint functions.

Theorem 4. *Suppose that the stochastic loss functions $f_1, \ldots, f_T$ and the stochastic constraint functions $g_1, \ldots, g_T$ satisfy Assumptions 2 and 3. Furthermore, (7) is satisfied. Then Algorithm 2 whose parameters are set as in (5) with $\beta = 0$ guarantees that*

$$\mathbb{E}\left[\sum_{t=1}^{T} f_t(\boldsymbol{x_t}) - \min_{x \in \mathcal{X}} \sum_{t=1}^{T} f_t(\boldsymbol{x})\right] \leq C_1 T^{\frac{1}{2}}, \quad \mathbb{E}\left[\sum_{t=1}^{T} g_t(\boldsymbol{x_t})\right] \leq C_2 T^{\frac{1}{2}}$$

where the expectations are taken over the randomness of $\mathcal{E}^1, \ldots, \mathcal{E}^K$ and the randomness of the loss and constraint functions.

7 Numerical Experiments

In this section, we present our experimental results to test the numerical performance of our projection-free algorithms, Algorithms 1 and 2, for online convex optimization with stochastic constraints. For Algorithm 1, we use the algorithms listed in Table 2. We consider the online matrix completion problem as in [3, 9]. For our experiments, we generate instances with synthetic simulated data. Here, to test our framework, we impose stochastic constraints.

We are given an $m \times n$ matrix M, and at each iteration t, we observe a subset $B_t \subseteq \{(i,j) : 1 \leq i \leq m, 1 \leq j \leq n\}$ of the entries of M. Here, M may encode the preferences of users over certain media items, in which case, B_t corresponds to the ratings inputted by some users at time t. Therefore, based on the sequence of subsets $B_1, \ldots, B_T$ that we observe over time, we want to infer the underlying matrix M. To be specific, we consider the following problem formulation:

$$\min \quad \frac{1}{2}\sum_{t=1}^{T}\sum_{(i,j)\in B_t}(X_{ij}^t - M_{ij})^2 \quad \text{s.t.} \quad \|X^t\|_* \leq k \quad \forall t \in [T], \quad \sum_{t=1}^{T}\mathrm{Tr}\left(G^t X^t\right) \leq 0.$$

Here, $X^1, \ldots, X^T$ are the sequence of $m \times n$ matrices we choose online. Constraint $\|X^t\|_* \leq k$ where $\|\cdot\|_*$ is the nuclear norm induces that each X^t has a low rank. Matrix G^t is randomly generated, and $\sum_{t=1}^{T}\mathrm{Tr}\left(G^t X^t\right) \leq 0$ is the long-term constraint that we impose. Furthermore, each B_t consists of b entries of M.

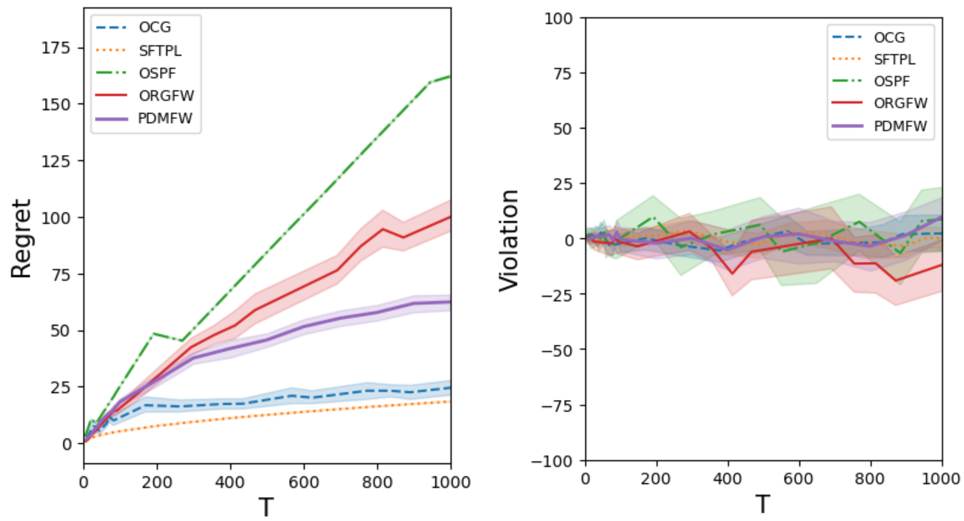
For our experiments, we test instances with $(m, n, k, b) = (50, 50, 5, 100)$. For each of the instances, the underlying matrix M is randomly chosen to have nuclear norm 1, and thus, it is contained in the domain $\mathcal{X} = \{X \in \mathbb{R}^{m\times n} : \|X\|_* \leq k\}$. At each iteration t, we sample matrix G^t from the uniform distribution over $[-1, 1]^{m\times n}$ and sample subset B_t from $\{B \subseteq \{1, \ldots, m \times n\} : |B| = b\}$ uniformly at random. Therefore, loss and constraint functions are stochastic and smooth. We test instances with time horizon $T \in \{10, 20, \ldots, 90, 100, 200, \ldots, 900, 1000\}$, and for each value of T, we generate 30 instances.

Figure 1a shows the regret values of the algorithms. PDMFW is Algorithm 2 while the others correspond to Algorithm 1 with oracles listed in Table 2. Note that each algorithm exhibits a sublinear growth in regret, as expected from our theoretical results. In particular, Algorithm 1 with OCG and Algorithm 1 with SFTPL achieve lower regret values than the others. Figure 1b summarizes the constraint violations of the algorithms. We observe that constraint violation values are centered around 0. In fact, there exists some instances where the constraint violation is below 0. This is possible because the benchmark $\boldsymbol{x}^*$ is set to a solution with $\bar{g}(\boldsymbol{x}^*) \leq 0$. Nevertheless, the results shown in Figure 1b support the theoretical results that Algorithms 1 and 2 attain sublinear constraint violations.

Acknowledgements This research is supported, in part, by the KAIST Starting Fund (KAIST-G04220016), the FOUR Brain Korea 21 Program (NRF-5199990113928), the National Research Foundation of Korea (NRF-2022M3J6A1063021).

References

[1] Dimitri P. Bertsekas. *Nonlinear Programming*. Athena Scientific, Cambridge, Massachusetts, 1999.



(a) (b)

Figure 1: Results from testing Algorithms 1 and 2 on matrix completion instances

[2] Matteo Castiglioni, Andrea Celli, Alberto Marchesi, Giulia Romano, and Nicola Gatti. A unifying framework for online optimization with long-term constraints. In Alice H. Oh, Alekh Agarwal, Danielle Belgrave, and Kyunghyun Cho, editors, *Advances in Neural Information Processing Systems*, 2022. URL https://openreview.net/forum?id=DhHqObn2UW.

[3] Lin Chen, Christopher Harshaw, Hamed Hassani, and Amin Karbasi. Projection-free online optimization with stochastic gradient: From convexity to submodularity. In Jennifer Dy and Andreas Krause, editors, *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 814–823. PMLR, 10–15 Jul 2018. URL https://proceedings.mlr.press/v80/chen18c.html.

[4] Lin Chen, Hamed Hassani, and Amin Karbasi. Online continuous submodular maximization. In Amos Storkey and Fernando Perez-Cruz, editors, *Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics*, volume 84 of *Proceedings of Machine Learning Research*, pages 1896–1905. PMLR, 09–11 Apr 2018. URL https://proceedings.mlr.press/v84/chen18f.html.

[5] Marguerite Frank and Philip Wolfe. An algorithm for quadratic programming. *Naval Research Logistics*, 3 (1-2):95–110, 1956.

[6] Dan Garber and Elad Hazan. A linearly convergent variant of the conditional gradient algorithm under strong convexity, with applications to online and stochastic optimization. *SIAM Journal on Optimization*, 26(3): 1493–1528, 2016. doi: 10.1137/140985366. URL https://doi.org/10.1137/140985366.

[7] Hengquan Guo, Xin Liu, Honghao Wei, and Lei Ying. Online convex optimization with hard constraints: Towards the best of two worlds and beyond. In Alice H. Oh, Alekh Agarwal, Danielle Belgrave, and Kyunghyun

Cho, editors, *Advances in Neural Information Processing Systems*, 2022. URL https://openreview.net/forum?id=rwdpFgfVpvN.

[8] James F. Hannan. *Approximation to bayes risk in repeated play*, pages 97–140. Princeton University Press, 1957. ISBN 9780691079363. URL http://www.jstor.org/stable/j.ctt1b9x26z.8.

[9] Elad Hazan. Introduction to online convex optimization. *Found. Trends Optim.*, 2(3–4):157–325, aug 2016. ISSN 2167-3888. doi: 10.1561/2400000013. URL https://doi.org/10.1561/2400000013.

[10] Elad Hazan and Satyen Kale. Projection-free online learning. In *Proceedings of the 29th International Coference on International Conference on Machine Learning*, ICML'12, page 1843–1850, Madison, WI, USA, 2012. Omnipress. ISBN 9781450312851.

[11] Elad Hazan and Edgar Minasyan. Faster projection-free online learning. In Jacob Abernethy and Shivani Agarwal, editors, *Proceedings of Thirty Third Conference on Learning Theory*, volume 125 of *Proceedings of Machine Learning Research*, pages 1877–1893. PMLR, 09–12 Jul 2020. URL https://proceedings.mlr.press/v125/hazan20a.html.

[12] Rodolphe Jenatton, Jim Huang, and Cedric Archambeau. Adaptive algorithms for online convex optimization with long-term constraints. In Maria Florina Balcan and Kilian Q. Weinberger, editors, *Proceedings of The 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pages 402–411, New York, New York, USA, 20–22 Jun 2016. PMLR. URL https://proceedings.mlr.press/v48/jenatton16.html.

[13] Adam Kalai and Santosh Vempala. Efficient algorithms for online decision problems. *Journal of Computer and System Sciences*, 71(3):291–307, 2005. ISSN 0022-0000. doi: https://doi.org/10.1016/j.jcss.2004.10.016. URL https://www.sciencedirect.com/science/article/pii/S0022000004001394. Learning Theory 2003.

[14] Nikolaos Liakopoulos, Apostolos Destounis, Georgios Paschos, Thrasyvoulos Spyropoulos, and Panayotis Mertikopoulos. Cautious regret minimization: Online optimization with long-term budget constraints. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 3944–3952. PMLR, 09–15 Jun 2019. URL https://proceedings.mlr.press/v97/liakopoulos19a.html.

[15] Mehrdad Mahdavi, Rong Jin, and Tianbao Yang. Trading regret for efficiency: Online convex optimization with long term constraints. *Journal of Machine Learning Research*, 13(81):2503–2528, 2012. URL http://jmlr.org/papers/v13/mahdavi12a.html.

[16] Shie Mannor, John N. Tsitsiklis, and Jia Yuan Yu. Online learning with sample path constraints. *Journal of Machine Learning Research*, 10(20):569–590, 2009. URL http://jmlr.org/papers/v10/mannor09a.html.

[17] Michael J. Neely and Hao Yu. Online convex optimization with time-varying constraints, 2017. URL https://arxiv.org/abs/1702.04783.

[18] Omid Sadeghi and Maryam Fazel. Online continuous dr-submodular maximization with long-term budget constraints. In Silvia Chiappa and Roberto Calandra, editors, *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pages 4410–4419. PMLR, 26–28 Aug 2020. URL https://proceedings.mlr.press/v108/sadeghi20a.html.

[19] Victor Valls, George Iosifidis, Douglas Leith, and Leandros Tassiulas. Online convex optimization with perturbed constraints: Optimal rates against stronger benchmarks. In Silvia Chiappa and Roberto Calandra, editors, *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pages 2885–2895. PMLR, 26–28 Aug 2020. URL https://proceedings.mlr.press/v108/valls20a.html.

[20] Xiaohan Wei, Hao Yu, and Michael J. Neely. Online primal-dual mirror descent under stochastic constraints. In *Abstracts of the 2020 SIGMETRICS/Performance Joint International Conference on Measurement and Modeling of Computer Systems*, SIGMETRICS '20, page 3–4, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450379854. doi: 10.1145/3393691.3394209. URL https://doi.org/10.1145/3393691.3394209.

[21] Jiahao Xie, Zebang Shen, Chao Zhang, Boyu Wang, and Hui Qian. Efficient projection-free online methods with stochastic recursive gradient. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(04): 6446–6453, Apr. 2020. doi: 10.1609/aaai.v34i04.6116. URL https://ojs.aaai.org/index.php/AAAI/article/view/6116.

[22] Xinlei Yi, Xiuxian Li, Tao Yang, Lihua Xie, Tianyou Chai, and Karl Johansson. Regret and cumulative constraint violation analysis for online convex optimization with long term constraints. In Marina Meila and Tong Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 11998–12008. PMLR, 18–24 Jul 2021. URL https://proceedings.mlr.press/v139/yi21b.html.

[23] Hao Yu and Michael J. Neely. A low complexity algorithm with o(â^št) regret and o(1) constraint violations for online convex optimization with long term constraints. *Journal of Machine Learning Research*, 21(1):1–24, 2020. URL http://jmlr.org/papers/v21/16-494.html.

[24] Hao Yu, Michael Neely, and Xiaohan Wei. Online convex optimization with stochastic constraints. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017. URL https://proceedings.neurips.cc/paper/2017/file/da0d1111d2dc5d489242e60ebcbaf988-Paper.pdf.

[25] Jianjun Yuan and Andrew Lamperski. Online convex optimization for cumulative constraints. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018. URL https://proceedings.neurips.cc/paper/2018/file/9cb9ed4f35cf7c2f295cc2bc6f732a84-Paper.pdf.

A Performance Analysis of Online Primal-Dual Projection-Free Framework (Algorithm 1)

Let us define filtration $\{\mathcal{F}_t : t \geq 0\}$ where $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_t = \sigma(\boldsymbol{\omega_1}, \ldots, \boldsymbol{\omega_t})$ being the σ-algebra generated by the set of random samples $\{\boldsymbol{\omega_1}, \ldots, \boldsymbol{\omega_t}\}$. Note that $\boldsymbol{x_t}$ is $\mathcal{F}_{t-1}$-measurable for all $t \geq 1$.

For simplicity, we define vector $\boldsymbol{x_k^q}$ and functions $f_k^q, g_k^q, h_k^q, L_k^q$ for $q \in [Q]$ and $k \in [K]$ as

$$\boldsymbol{x_k^q} = \boldsymbol{x_t}, \quad f_k^q(\boldsymbol{x}) = f_t(\boldsymbol{x}), \quad g_k^q(\boldsymbol{x}) = g_t(\boldsymbol{x}), \quad h_k^q(\boldsymbol{x}) = h_t(\boldsymbol{x}), \quad L_k^q(\boldsymbol{x}) = L_t(\boldsymbol{x}, \lambda)$$

for any $\boldsymbol{x} \in \mathcal{X}$ and $\lambda \geq 0$ where $t = k + (q-1)K$. Then it follows that

$$\begin{aligned}
h_k^q(\boldsymbol{x}) &= f_k^q(\boldsymbol{x}) + \lambda_q g_k^q(\boldsymbol{x}), \\
L_k^q(\boldsymbol{x}, \lambda) &= f_k^q(\boldsymbol{x}) + \lambda_q g_k^q(\boldsymbol{x}) - \frac{\theta}{2K}\lambda^2 \\
&= h_k^q(\boldsymbol{x}) - \frac{\theta}{2K}\lambda^2.
\end{aligned}$$

Moreover, note that

$$\begin{aligned}
\sum_{t=(q-1)K+1}^{qK} (L_t(\boldsymbol{x_t}, \lambda_q) - L_t(\boldsymbol{x}, \lambda_q)) &= \sum_{k=1}^{K} (L_k^q(\boldsymbol{x_k^q}, \lambda_q) - L_k^q(\boldsymbol{x}, \lambda_q)) \\
&= \sum_{k=1}^{K} (h_k^q(\boldsymbol{x_k^q}) - h_k^q(\boldsymbol{x}))
\end{aligned} \tag{8}$$

Lemma A.1. *Let $\boldsymbol{x_k^q}$ for $k = 1, \ldots, K$ be decisions determined by an (α, C_0, C_1, C_2)-oracle. Then*

$$\mathbb{E}\left[\sum_{k=1}^{K} (h_k^q(\boldsymbol{x_k^q}) - h_k^q(\boldsymbol{x})) \mid \mathcal{F}_{(q-1)K}\right] \leq (C_0 + C_1 D(1+\lambda_q) + C_2 L(1+\lambda_q)) T^{\frac{\alpha}{3-2\alpha}}$$

for any $\boldsymbol{x} \in \mathcal{X}$.

Proof. As we assumed that the loss functions $f_1, \ldots, f_T$ and the constraint functions $g_1, \ldots, g_T$ are D-Lipschitz and L-smooth, it follows that $h_1, \ldots, h_T$ are $D(1+\lambda_q)$-Lipschitz and $L(1+\lambda_q)$-smooth. Then

$$\begin{aligned}
\mathbb{E}\left[\sum_{k=1}^{K} (h_k^q(\boldsymbol{x_k^q}) - h_k^q(\boldsymbol{x})) \mid \mathcal{F}_{(q-1)K}\right] &\leq \mathbb{E}\left[\sum_{k=1}^{K} \left(h_k^q(\boldsymbol{x_k^q}) - \min_{\boldsymbol{x} \in \mathcal{X}} h_k^q(\boldsymbol{x})\right) \mid \mathcal{F}_{(q-1)K}\right] \\
&\leq (C_0 + C_1 D(1+\lambda_q) + C_2 L(1+\lambda_q)) K^{\alpha}.
\end{aligned}$$

Here, we have $K = T^{\frac{1}{3-2\alpha}}$, as required. ☐

Next, observe that

$$\sum_{t=(q-1)K+1}^{qK} (L_t(\boldsymbol{x_t}, \lambda) - L_t(\boldsymbol{x_t}, \lambda_q)) = \sum_{k=1}^{K} (L_k^q(\boldsymbol{x_k^q}, \lambda) - L_k^q(\boldsymbol{x_k^q}, \lambda_q)). \tag{9}$$

Lemma A.2. *Let $\lambda_1, \dots, \lambda_Q$ be the dual variables chosen by Algorithm 1. Then for any $\lambda \geq 0$,*

$$\sum_{q=1}^{Q} \sum_{k=1}^{K} (L_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}, \lambda) - L_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}, \lambda_q)) \leq \frac{1}{2\mu}\lambda^2 + G^2K^2Q\mu + \theta^2\mu \sum_{q=1}^{Q} \lambda_q^2.$$

Proof. Note that

$$\lambda_{q+1} = \left[\lambda_q + \mu \sum_{k=1}^{K} \nabla_\lambda L_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}, \lambda_q)\right]_+ .$$

Then

$$\begin{aligned}
(\lambda_{q+1} - \lambda)^2 &\leq \left(\lambda_q + \mu \sum_{k=1}^{K} \nabla_\lambda L_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}, \lambda_q) - \lambda\right)^2 \\
&= (\lambda_q - \lambda)^2 + \mu^2 \left(\sum_{k=1}^{K} g_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}) - \theta\lambda_q\right)^2 + 2\mu \left(\sum_{k=1}^{K} \nabla_\lambda L_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}, \lambda_q)\right)^\top (\lambda_q - \lambda) \\
&\leq (\lambda_q - \lambda)^2 + 2\mu^2(G^2K^2 + \theta^2\lambda_q^2) + 2\mu \sum_{k=1}^{K} (L_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}, \lambda_q) - L_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}, \lambda)) .
\end{aligned}$$

Therefore, it follows that

$$\begin{aligned}
&\sum_{q=1}^{Q} \sum_{k=1}^{K} (L_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}, \lambda_q) - L_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}, \lambda)) \\
&\leq \sum_{q=1}^{Q} \frac{1}{2\mu} \left((\lambda_q - \lambda)^2 - (\lambda_{q+1} - \lambda)^2\right) + G^2K^2Q\mu + \theta^2\mu \sum_{q=1}^{Q} \lambda_q^2 \\
&= \frac{1}{2\mu}\lambda^2 - \frac{1}{2\mu}(\lambda_{Q+1} - \lambda)^2 + G^2K^2Q\mu + \theta^2\mu \sum_{q=1}^{Q} \lambda_q^2 \\
&\leq \frac{1}{2\mu}\lambda^2 + G^2K^2Q\mu + \theta^2\mu \sum_{q=1}^{Q} \lambda_q^2
\end{aligned}$$

where the last inequality holds because $(\lambda_{Q+1} - \lambda)^2 \geq 0$. ☐

Note that

$$\begin{aligned}
&\sum_{q=1}^{Q} \sum_{t=(q-1)K+1}^{qK} (L_t(\boldsymbol{x_t}, \lambda) - L_t(\boldsymbol{x}, \lambda_q)) \\
&= \sum_{q=1}^{Q} \sum_{k=1}^{K} (L_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}, \lambda) - L_k^q(\boldsymbol{x}, \lambda_q)) \qquad (10) \\
&= \sum_{q=1}^{Q} \sum_{k=1}^{K} (f_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}) - f_k^q(\boldsymbol{x})) + \lambda \sum_{q=1}^{Q} \sum_{k=1}^{K} g_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}) - \sum_{q=1}^{Q} \lambda_q \sum_{k=1}^{K} g_k^q(\boldsymbol{x}) - \frac{Q\theta}{2}\lambda^2 + \frac{\theta}{2} \sum_{q=1}^{Q} \lambda_q^2.
\end{aligned}$$

Based on (8), (9), (10), Lemmas A.1 and A.2, we deduce the following lemma.

Lemma A.3. *Let $\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}$ for $k = 1, \dots, K$ and $\lambda_1, \dots, \lambda_Q$ be the decisions and the dual variable chosen by Algorithm 1.*

Then for any $\boldsymbol{x} \in \mathcal{X}$ and $\lambda \geq 0$,

$$\mathbb{E}\left[\sum_{q=1}^{Q}\sum_{k=1}^{K}(f_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}) - f_k^q(\boldsymbol{x})) + \lambda\sum_{q=1}^{Q}\sum_{k=1}^{K} g_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}) - \sum_{q=1}^{Q}\lambda_q\sum_{k=1}^{K} g_k^q(\boldsymbol{x})\right]$$
$$\leq \frac{1}{2}\left(\frac{1}{\mu} + Q\theta\right)\lambda^2 + G^2K^2Q\mu + Q(C_0 + C_1D + C_2L)T^{\frac{\alpha}{3-2\alpha}} + \frac{Q}{2}(C_1D + C_2L)T^{\frac{\alpha}{3-2\alpha}+\beta}.$$

Proof. Note that

$$\begin{aligned}
&\sum_{q=1}^{Q}\sum_{k=1}^{K}(f_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}) - f_k^q(\boldsymbol{x})) + \lambda\sum_{q=1}^{Q}\sum_{k=1}^{K} g_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}) - \sum_{q=1}^{Q}\lambda_q\sum_{k=1}^{K} g_k^q(\boldsymbol{x})\\
&= \sum_{q=1}^{Q}\sum_{k=1}^{K}(L_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}, \lambda) - L_k^q(\boldsymbol{x}, \lambda_q)) + \frac{Q\theta}{2}\lambda^2 - \frac{\theta}{2}\sum_{q=1}^{Q}\lambda_q^2\\
&= \sum_{q=1}^{Q}\sum_{k=1}^{K}(h_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}) - h_k^q(\boldsymbol{x})) + \sum_{q=1}^{Q}\sum_{k=1}^{K}(L_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}, \lambda) - L_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}, \lambda_q)) + \frac{Q\theta}{2}\lambda^2 - \frac{\theta}{2}\sum_{q=1}^{Q}\lambda_q^2\\
&\leq \sum_{q=1}^{Q}\sum_{k=1}^{K}(h_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}) - h_k^q(\boldsymbol{x})) + \frac{1}{2}\left(\frac{1}{\mu} + Q\theta\right)\lambda^2 + G^2K^2Q\mu + \left(\theta^2\mu - \frac{\theta}{2}\right)\sum_{q=1}^{Q}\lambda_q^2
\end{aligned}$$

where the first equality is from (10), the second equality is due to (8) and (9), and the last inequality follows from Lemma A.2. By Lemma A.1,

$$\begin{aligned}
\mathbb{E}\left[\sum_{q=1}^{Q}\sum_{k=1}^{K}(h_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}) - h_k^q(\boldsymbol{x}))\right] &= \sum_{q=1}^{Q}\mathbb{E}\left[\mathbb{E}\left[\sum_{k=1}^{K}(h_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}) - h_k^q(\boldsymbol{x})) \mid \mathcal{F}_{(q-1)K}\right]\right]\\
&\leq \sum_{q=1}^{Q}\mathbb{E}\left[(C_0 + C_1D(1+\lambda_q) + C_2L(1+\lambda_q))T^{\frac{\alpha}{3-2\alpha}}\right].
\end{aligned}$$

Here, using the fact that $a + b \geq 2\sqrt{ab}$ for any $a, b \geq 0$, we obtain

$$\lambda_q T^{\frac{\alpha}{3-2\alpha}} \leq \frac{1}{2}T^{\frac{\alpha}{3-2\alpha}+\beta} + \frac{\lambda_q^2}{2}T^{\frac{\alpha}{3-2\alpha}-\beta}.$$

Therefore, it follows that

$$\begin{aligned}
&\sum_{q=1}^{Q}(C_0 + C_1D(1+\lambda_q) + C_2L(1+\lambda_q))T^{\frac{\alpha}{3-2\alpha}}\\
&\leq Q(C_0 + C_1D + C_2L)T^{\frac{\alpha}{3-2\alpha}} + \frac{Q}{2}(C_1D + C_2L)T^{\frac{\alpha}{3-2\alpha}+\beta} + \frac{1}{2}(C_1D + C_2L)T^{\frac{\alpha}{3-2\alpha}-\beta}\sum_{q=1}^{Q}\lambda_q^2\\
&= Q(C_0 + C_1D + C_2L)T^{\frac{\alpha}{3-2\alpha}} + \frac{Q}{2}(C_1D + C_2L)T^{\frac{\alpha}{3-2\alpha}+\beta} + \frac{\theta}{6}\sum_{q=1}^{Q}\lambda_q^2.
\end{aligned}$$

Note that

$$\left(\theta^2\mu - \frac{\theta}{2}\right)\sum_{q=1}^{Q}\lambda_q^2 + \frac{\theta}{6}\sum_{q=1}^{Q}\lambda_q^2 = \left(\theta^2\mu - \frac{\theta}{3}\right)\sum_{q=1}^{Q}\lambda_q^2 = \left(\frac{\theta}{Q+1} - \frac{\theta}{3}\right)\sum_{q=1}^{Q}\lambda_q^2.$$

Since $\lambda_1 = 0$ and $Q+1 \geq 3$ for $Q \geq 2$, we have

$$\left(\frac{\theta}{Q+1} - \frac{\theta}{3}\right)\sum_{q=1}^{Q} \lambda_q^2 \leq 0,$$

as required. $\square$

Proof of Theorem 1. By plugging in (3) to the inequality given in Lemma A.3, it follows that

$$\begin{aligned}
&\mathbb{E}\left[\sum_{q=1}^{Q}\sum_{k=1}^{K}(f_k^q(\boldsymbol{x_k^q}) - f_k^q(\boldsymbol{x})) + \lambda\sum_{q=1}^{Q}\sum_{k=1}^{K} g_k^q(\boldsymbol{x_k^q}) - \sum_{q=1}^{Q}\lambda_q\sum_{k=1}^{K} g_k^q(\boldsymbol{x})\right] \\
&\leq \frac{1}{2}\left(\frac{1}{\mu} + Q\theta\right)\lambda^2 + G^2K^2Q\mu + Q(C_0 + C_1D + C_2L)T^{\frac{\alpha}{3-2\alpha}} + \frac{Q}{2}(C_1D + C_2L)T^{\frac{\alpha}{3-2\alpha}+\beta} \\
&= \frac{1}{2}(2Q+1)\theta\lambda^2 + \frac{G^2K^2Q}{\theta(Q+1)} + Q(C_0 + C_1D + C_2L)T^{\frac{\alpha}{3-2\alpha}} + \frac{Q}{2}(C_1D + C_2L)T^{\frac{\alpha}{3-2\alpha}+\beta} \\
&\leq 2Q\theta\lambda^2 + \frac{G^2K^2}{\theta} + Q(C_0 + C_1D + C_2L)T^{\frac{\alpha}{3-2\alpha}} + \frac{Q}{2}(C_1D + C_2L)T^{\frac{\alpha}{3-2\alpha}+\beta} \\
&= 6(C_1D + C_2L)T^{\frac{2-\alpha}{3-2\alpha}-\beta}\lambda^2 + \frac{G^2}{3(C_1D + C_2L)}T^{\frac{2-\alpha}{3-2\alpha}+\beta} \\
&\quad + (C_0 + C_1D + C_2L)T^{\frac{2-\alpha}{3-2\alpha}} + \frac{1}{2}(C_1D + C_2L)T^{\frac{2-\alpha}{3-2\alpha}+\beta} \\
&\leq 6(C_1D + C_2L)T^{\frac{2-\alpha}{3-2\alpha}-\beta}\lambda^2 + \left(\frac{G^2}{3(C_1D + C_2L)} + \frac{3(C_0 + C_1D + C_2L)}{2}\right)T^{\frac{2-\alpha}{3-2\alpha}+\beta}.
\end{aligned}$$

Moreover, as $\bar{g}(\boldsymbol{x}^*) \leq 0$,

$$\begin{aligned}
\mathbb{E}\left[\sum_{q=1}^{Q}\lambda_q\sum_{k=1}^{K} g_k^q(\boldsymbol{x}^*)\right] &= \sum_{q=1}^{Q}\mathbb{E}\left[\mathbb{E}\left[\lambda_q\sum_{k=1}^{K} g_k^q(\boldsymbol{x}^*) \mid \mathcal{F}_{(q-1)K}\right]\right] \\
&= \sum_{q=1}^{Q}\mathbb{E}\left[\lambda_q\sum_{k=1}^{K}\mathbb{E}\left[g_k^q(\boldsymbol{x}^*) \mid \mathcal{F}_{(q-1)K}\right]\right] \\
&= \sum_{q=1}^{Q}\mathbb{E}\left[\lambda_q\sum_{k=1}^{K}\bar{g}(\boldsymbol{x}^*)\right] \\
&\leq 0.
\end{aligned}$$

This implies that

$$\begin{aligned}
&\mathbb{E}\left[\sum_{q=1}^{Q}\sum_{k=1}^{K}(f_k^q(\boldsymbol{x_k^q}) - f_k^q(\boldsymbol{x}^*)) + \lambda\sum_{q=1}^{Q}\sum_{k=1}^{K} g_k^q(\boldsymbol{x_k^q})\right] \\
&\leq 6(C_1D + C_2L)T^{\frac{2-\alpha}{3-2\alpha}-\beta}\lambda^2 + \left(\frac{G^2}{3(C_1D + C_2L)} + \frac{3(C_0 + C_1D + C_2L)}{2}\right)T^{\frac{2-\alpha}{3-2\alpha}+\beta}.
\end{aligned}$$

Next, we set

$$\lambda = \frac{1}{12(C_1D + C_2L)T^{\frac{2-\alpha}{3-2\alpha}-\beta}}\left[\mathbb{E}\left[\sum_{q=1}^{Q}\sum_{k=1}^{K} g_k^q(\boldsymbol{x_k^q})\right]\right]_+ .$$

Then we obtain

$$\mathbb{E}\left[\sum_{q=1}^{Q}\sum_{k=1}^{K}(f_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}) - f_k^q(\boldsymbol{x}^*))\right] \leq -\frac{1}{24(C_1D+C_2L)T^{\frac{2-\alpha}{3-2\alpha}-\beta}}\left[\mathbb{E}\left[\sum_{q=1}^{Q}\sum_{k=1}^{K} g_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}})\right]\right]_+^2 + \left(\frac{G^2}{3(C_1D+C_2L)} + \frac{3(C_0+C_1D+C_2L)}{2}\right)T^{\frac{2-\alpha}{3-2\alpha}+\beta}.$$

As a result,

$$\mathbb{E}\left[\sum_{q=1}^{Q}\sum_{k=1}^{K}(f_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}) - f_k^q(\boldsymbol{x}^*))\right] \leq \left(\frac{G^2}{3(C_1D+C_2L)} + \frac{3(C_0+C_1D+C_2L)}{2}\right)T^{\frac{2-\alpha}{3-2\alpha}+\beta}.$$

Moreover,

$$\begin{aligned}
&\frac{1}{24(C_1D+C_2L)T^{\frac{2-\alpha}{3-2\alpha}-\beta}}\left[\mathbb{E}\left[\sum_{q=1}^{Q}\sum_{k=1}^{K} g_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}})\right]\right]_+^2 \\
&\leq -\mathbb{E}\left[\sum_{q=1}^{Q}\sum_{k=1}^{K}(f_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}) - f_k^q(\boldsymbol{x}^*))\right] + \left(\frac{G^2}{3(C_1D+C_2L)} + \frac{3(C_0+C_1D+C_2L)}{2}\right)T^{\frac{2-\alpha}{3-2\alpha}+\beta} \\
&\leq TDR + \left(\frac{G^2}{3(C_1D+C_2L)} + \frac{3(C_0+C_1D+C_2L)}{2}\right)T^{\frac{2-\alpha}{3-2\alpha}+\beta}
\end{aligned}$$

where the second inequality holds because $f_1, \ldots, f_T$ are D-Lipschitz. Then

$$\begin{aligned}
&\mathbb{E}\left[\sum_{q=1}^{Q}\sum_{k=1}^{K} g_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}})\right] \\
&\leq \left[\mathbb{E}\left[\sum_{q=1}^{Q}\sum_{k=1}^{K} g_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}})\right]\right]_+ \\
&\leq \sqrt{24(C_1D+C_2L)T^{\frac{2-\alpha}{3-2\alpha}-\beta}\left(TDR + \left(\frac{G^2}{3(C_1D+C_2L)} + \frac{3(C_0+C_1D+C_2L)}{2}\right)T^{\frac{2-\alpha}{3-2\alpha}+\beta}\right)} \\
&= O\left(T^{\frac{5-3\alpha}{6-4\alpha}-\frac{\beta}{2}}\right).
\end{aligned}$$

Then the result follows as

$$\mathbb{E}\left[\sum_{t=1}^{T}(f_t(\boldsymbol{x_t}) - f_t(\boldsymbol{x}^*))\right] = \mathbb{E}\left[\sum_{q=1}^{Q}\sum_{k=1}^{K}(f_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}}) - f_k^q(\boldsymbol{x}^*))\right],$$

$$\mathbb{E}\left[\sum_{t=1}^{T} g_t(\boldsymbol{x_t})\right] = \mathbb{E}\left[\sum_{q=1}^{Q}\sum_{k=1}^{K} g_k^q(\boldsymbol{x}_{\boldsymbol{k}}^{\boldsymbol{q}})\right],$$

as required. □

B Performance Analysis of PMFWG (Algorithm 2)

We provide the proof of Lemma 5.1 in Appendix B.1. Then, in Appendix B.2 we show Theorem 2.

B.1 Regret of FTPL: Proof of Lemma 5.1

Recall that our adaptive variant of Follow-The-Perturbed-Leader proceeds as the following setup:

- Choose a random perturbation vector $\boldsymbol{p}$ from $[0, 1/\delta]^d$ for some $\delta > 0$ uniformly at random.
- For $t = 1, \ldots, T$:
  - Choose
$$\boldsymbol{x_t} \in \operatorname*{argmin}_{\boldsymbol{x} \in \mathcal{X}} \left(\boldsymbol{p} + \sum_{s=1}^{t-1} \boldsymbol{w_s} \right)^\top \boldsymbol{x}.$$
  - Observe $\boldsymbol{w_t}$.

Define $M(\boldsymbol{w}) = \operatorname{argmin}_{\boldsymbol{x} \in \mathcal{X}} \boldsymbol{w}^\top \boldsymbol{x}$ for $\boldsymbol{w} \in \mathbb{R}^d$.

Lemma B.1. *The sequence $\{\boldsymbol{w_t}\}_{t \in [T]}$ satisfies the following.*

$$\sum_{t=1}^{T} M\left(\sum_{s=1}^{t} \boldsymbol{w_s} \right)^\top \boldsymbol{w_t} \leq \sum_{t=1}^{T} M\left(\sum_{s=1}^{T} \boldsymbol{w_s} \right)^\top \boldsymbol{w_t}.$$

Proof. We proceed by induction. The case $T = 1$ is obvious. Now assume it holds for $T \geq 1$, i.e.,

$$\sum_{t=1}^{T} M\left(\sum_{s=1}^{t} \boldsymbol{w_s} \right)^\top \boldsymbol{w_t} \leq \sum_{t=1}^{T} M\left(\sum_{s=1}^{T} \boldsymbol{w_s} \right)^\top \boldsymbol{w_t}.$$

Consider

$$\begin{aligned}
\sum_{t=1}^{T+1} M\left(\sum_{s=1}^{t} \boldsymbol{w_s} \right)^\top \boldsymbol{w_t} &= \sum_{t=1}^{T} M\left(\sum_{s=1}^{t} \boldsymbol{w_s} \right)^\top \boldsymbol{w_t} + M\left(\sum_{s=1}^{T+1} \boldsymbol{w_s} \right)^\top \boldsymbol{w_{T+1}} \\
&\leq M\left(\sum_{s=1}^{T+1} \boldsymbol{w_s} \right)^\top \boldsymbol{w_{T+1}} + \sum_{t=1}^{T} M\left(\sum_{s=1}^{T} \boldsymbol{w_s} \right)^\top \boldsymbol{w_t} \\
&= M\left(\sum_{s=1}^{T+1} \boldsymbol{w_s} \right)^\top \boldsymbol{w_{T+1}} + M\left(\sum_{s=1}^{T} \boldsymbol{w_s} \right)^\top \left(\sum_{t=1}^{T} \boldsymbol{w_t} \right) \\
&\leq M\left(\sum_{s=1}^{T+1} \boldsymbol{w_s} \right)^\top \boldsymbol{w_{T+1}} + M\left(\sum_{s=1}^{T} \boldsymbol{w_s} + \boldsymbol{w_{T+1}} \right)^\top \left(\sum_{t=1}^{T} \boldsymbol{w_t} \right) \\
&= \sum_{t=1}^{T+1} M\left(\sum_{s=1}^{T+1} \boldsymbol{w_s} \right)^\top \boldsymbol{w_t}
\end{aligned}$$

where the first inequality comes from the induction hypothesis and the second inequality is because of the definition of $M(\cdot)$, as required. □

Lemma B.2. *Define $\boldsymbol{p_0} := \boldsymbol{0}$. Then for any sequence $\boldsymbol{p_1}, \ldots, \boldsymbol{p_T}$ we have*

$$\sum_{t=1}^{T} M\left(\boldsymbol{p_t} + \sum_{s=1}^{t} \boldsymbol{w_s} \right)^\top \boldsymbol{w_t} \leq \sum_{t=1}^{T} M\left(\sum_{s=1}^{T} \boldsymbol{w_s} \right)^\top \boldsymbol{w_t} + R \sum_{t=1}^{T} \|\boldsymbol{p_t} - \boldsymbol{p_{t-1}}\|_\infty.$$

Consequently, if $\boldsymbol{p_t} = \boldsymbol{p}$ for all $t \in [T]$, we have

$$\sum_{t=1}^{T} M\left(\boldsymbol{p} + \sum_{s=1}^{t} \boldsymbol{w_s}\right)^{\top} \boldsymbol{w_t} \leq \sum_{t=1}^{T} M\left(\sum_{s=1}^{T} \boldsymbol{w_s}\right)^{\top} \boldsymbol{w_t} + R\|\boldsymbol{p}\|_{\infty}.$$

Proof. We invoke Lemma B.1 with $\boldsymbol{w_t'} = \boldsymbol{w_t} + \boldsymbol{p_t} - \boldsymbol{p_{t-1}}$ to obtain

$$\begin{aligned}
&\sum_{t=1}^{T} M\left(\boldsymbol{p_t} + \sum_{s=1}^{t} \boldsymbol{w_s}\right)^{\top} (\boldsymbol{w_t} + \boldsymbol{p_t} - \boldsymbol{p_{t-1}}) \\
&\leq \sum_{t=1}^{T} M\left(\boldsymbol{p_T} + \sum_{s=1}^{T} \boldsymbol{w_s}\right)^{\top} (\boldsymbol{w_t} + \boldsymbol{p_t} - \boldsymbol{p_{t-1}}) \\
&= M\left(\boldsymbol{p_T} + \sum_{s=1}^{T} \boldsymbol{w_s}\right)^{\top} \left(\boldsymbol{p_T} + \sum_{t=1}^{T} \boldsymbol{w_t}\right) \\
&\leq M\left(\sum_{s=1}^{T} \boldsymbol{w_s}\right)^{\top} \left(\boldsymbol{p_T} + \sum_{t=1}^{T} \boldsymbol{w_t}\right) \\
&= M\left(\sum_{s=1}^{T} \boldsymbol{w_s}\right)^{\top} \left(\sum_{t=1}^{T} \boldsymbol{w_t}\right) + M\left(\sum_{s=1}^{T} \boldsymbol{w_s}\right)^{\top} \left(\sum_{t=1}^{T} (\boldsymbol{p_t} - \boldsymbol{p_{t-1}})\right)
\end{aligned}$$

where the second inequality follows from the definition of $M(\cdot)$. This inequality implies that

$$\begin{aligned}
&\sum_{t=1}^{T} M\left(\boldsymbol{p_t} + \sum_{s=1}^{t} \boldsymbol{w_s}\right)^{\top} \boldsymbol{w_t} \\
&\leq M\left(\sum_{s=1}^{T} \boldsymbol{w_s}\right)^{\top} \left(\sum_{t=1}^{T} \boldsymbol{w_t}\right) + \sum_{t=1}^{T} \left(M\left(\sum_{s=1}^{T} \boldsymbol{w_s}\right) - M\left(\boldsymbol{p_t} + \sum_{s=1}^{t} \boldsymbol{w_s}\right)\right)^{\top} (\boldsymbol{p_t} - \boldsymbol{p_{t-1}}) \\
&\leq M\left(\sum_{s=1}^{T} \boldsymbol{w_s}\right)^{\top} \left(\sum_{t=1}^{T} \boldsymbol{w_t}\right) + R\sum_{t=1}^{T} \|\boldsymbol{p_t} - \boldsymbol{p_{t-1}}\|_{\infty},
\end{aligned}$$

as required. $\square$

Lemma B.3. *Let $\mathbf{p}$ be a vector sampled from $[0, 1/\delta]^n$ uniformly at random, and $\boldsymbol{w}, \boldsymbol{w'}, \boldsymbol{w''}$ be three arbitrary fixed vectors. Then*

$$\mathbb{E}\left[(M(\boldsymbol{w} + \boldsymbol{p}) - M(\boldsymbol{w} + \boldsymbol{w'} + \boldsymbol{p}))^{\top} \boldsymbol{w''}\right] \leq R\|\boldsymbol{w''}\|_{\infty} \sum_{i=1}^{d} \min\left\{1, \delta|w_i'|\right\}.$$

Proof. We consider the hypercubes $C = \boldsymbol{w} + [0, 1/\delta]^d$, $C' = \boldsymbol{w} + \boldsymbol{w'} + [0, 1/\delta]^d$. Note that $\tilde{\boldsymbol{p}} := \boldsymbol{w} + \boldsymbol{p}$ is uniformly distributed over C and $\tilde{\boldsymbol{p}}' := \boldsymbol{w} + \boldsymbol{w'} + \boldsymbol{p}$ is uniformly distributed over C'. Let $C_\cap = C \cap C'$ and $C_\circ = (C \cup C') \setminus (C \cap C')$. Now

$$\begin{aligned}
&\mathbb{E}\left[(M(\tilde{\boldsymbol{p}}) - M(\tilde{\boldsymbol{p}}'))^{\top} \boldsymbol{w''}\right] \\
&= \mathbb{E}\left[M(\tilde{\boldsymbol{p}})^{\top} \boldsymbol{w''} \mid \tilde{\boldsymbol{p}} \in C_\cap\right] \mathbb{P}[\tilde{\boldsymbol{p}} \in C_\cap] + \mathbb{E}\left[M(\tilde{\boldsymbol{p}})^{\top} \boldsymbol{w''} \mid \tilde{\boldsymbol{p}} \in C_\circ\right] \mathbb{P}[\tilde{\boldsymbol{p}} \in C_\circ] \\
&\quad - \mathbb{E}\left[M(\tilde{\boldsymbol{p}}')^{\top} \boldsymbol{w''} \mid \tilde{\boldsymbol{p}}' \in C_\cap\right] \mathbb{P}[\tilde{\boldsymbol{p}}' \in C_\cap] - \mathbb{E}\left[M(\tilde{\boldsymbol{p}}')^{\top} \boldsymbol{w''} \mid \tilde{\boldsymbol{p}}' \in C_\circ\right] \mathbb{P}[\tilde{\boldsymbol{p}}' \in C_\circ].
\end{aligned}$$

Since $\boldsymbol{p}$ is uniform on the hypercube, the distribution of $\tilde{\boldsymbol{p}} \mid \tilde{\boldsymbol{p}} \in C_\cap$ is equal to the distribution of $\tilde{\boldsymbol{p}}' \mid \tilde{\boldsymbol{p}}' \in C_\cap$, and also the probability that each vector is in $C_\cap$ is equal. Therefore,

$$\begin{aligned}
&\mathbb{E}\left[\left(M(\tilde{\boldsymbol{p}}) - M(\tilde{\boldsymbol{p}}')\right)^\top \boldsymbol{w}''\right] \\
&= \mathbb{E}\left[M(\tilde{\boldsymbol{p}})^\top \boldsymbol{w}'' \mid \tilde{\boldsymbol{p}} \in C_\circ\right] \mathbb{P}[\tilde{\boldsymbol{p}} \in C_\circ] - \mathbb{E}\left[M(\tilde{\boldsymbol{p}}')^\top \boldsymbol{w}'' \mid \tilde{\boldsymbol{p}}' \in C_\circ\right] \mathbb{P}[\tilde{\boldsymbol{p}}' \in C_\circ].
\end{aligned}$$

Since $\mathbb{P}[\tilde{\boldsymbol{p}} \in C_\cap] = \mathbb{P}[\tilde{\boldsymbol{p}}' \in C_\cap]$ we have $\mathbb{P}[\tilde{\boldsymbol{p}} \in C_\circ] = \mathbb{P}[\tilde{\boldsymbol{p}}' \in C_\circ]$, hence

$$\begin{aligned}
&\mathbb{E}\left[\left(M(\tilde{\boldsymbol{p}}) - M(\tilde{\boldsymbol{p}}')\right)^\top \boldsymbol{w}''\right] \\
&= \mathbb{E}\left[M(\tilde{\boldsymbol{p}})^\top \boldsymbol{w}'' \mid \tilde{\boldsymbol{p}} \in C_\circ\right] \mathbb{P}[\tilde{\boldsymbol{p}} \in C_\circ] - \mathbb{E}\left[M(\tilde{\boldsymbol{p}}')^\top \boldsymbol{w}'' \mid \tilde{\boldsymbol{p}} \in C_\circ\right] \mathbb{P}[\tilde{\boldsymbol{p}}' \in C_\circ] \\
&= \mathbb{P}[\tilde{\boldsymbol{p}} \in C_\circ] \left(\mathbb{E}\left[M(\tilde{\boldsymbol{p}}) \mid \tilde{\boldsymbol{p}} \in C_\circ\right] - \mathbb{E}\left[M(\tilde{\boldsymbol{p}}') \mid \tilde{\boldsymbol{p}}' \in C_\circ\right]\right)^\top \boldsymbol{w}'' \\
&\le R \|\boldsymbol{w}''\|_\infty \mathbb{P}[\tilde{\boldsymbol{p}} \in C_\circ].
\end{aligned}$$

We now estimate $\mathbb{P}[\tilde{\boldsymbol{p}} \in C_\circ]$. If $\tilde{\boldsymbol{p}} \in C \setminus C'$ then there exists some $i \in [d]$ such that $\tilde{p}_i \in [w_i, w_i + 1/\delta]$ but $\tilde{p}_i \notin [w_i + w_i', w_i + w_i' + 1/\delta]$. We can compute that

$$\mathbb{P}[\tilde{p}_i \in [w_i, w_i + 1/\delta] \setminus [w_i + w_i', w_i + w_i' + 1/\delta]] = \begin{cases} 1, & |w_i'| \ge 1/\delta \\ \delta |w_i'|, & |w_i'| < 1/\delta \end{cases} = \min\{1, \delta |w_i'|\}.$$

Now observe that

$$\begin{aligned}
\mathbb{P}\left[\tilde{\boldsymbol{p}} \in C_\circ\right] &= \mathbb{P}\left[\tilde{\boldsymbol{p}} \in C \setminus C'\right] \\
&= \mathbb{P}\left[\exists i \in [d] \text{ s.t. } \tilde{p}_i \in [w_i, w_i + 1/\delta] \setminus [w_i + w_i', w_i + w_i' + 1/\delta]\right] \\
&= 1 - \mathbb{P}\left[\forall i \in [d],\ \tilde{p}_i \in [w_i, w_i + 1/\delta] \cap [w_i + w_i', w_i + w_i' + 1/\delta]\right] \\
&= 1 - \prod_{i=1}^d \left(1 - \mathbb{P}[\tilde{p}_i \in [w_i, w_i + 1/\delta] \setminus [w_i + w_i', w_i + w_i' + 1/\delta]]\right) \\
&= 1 - \prod_{i=1}^d \left(1 - \min\{1, \delta |w_i'|\}\right) \\
&= 1 - \prod_{i=1}^d \max\{0, 1 - \delta |w_i'|\} \\
&= \begin{cases} 1 - \prod_{i=1}^d \left(1 - \delta |w_i'|\right), & \|\boldsymbol{w}'\|_\infty \le 1/\delta \\ 1, & \|\boldsymbol{w}'\|_\infty > 1/\delta, \end{cases}
\end{aligned}$$

where the fourth equality follows since each $\tilde{p}_i$ is independent. A union bound then gives

$$\begin{aligned}
\mathbb{P}\left[\tilde{\boldsymbol{p}} \in C_\circ\right] &= \mathbb{P}\left[\tilde{\boldsymbol{p}} \in C \setminus C'\right] \\
&= \mathbb{P}\left[\exists i \in [d] \text{ s.t. } \tilde{p}_i \in [w_i, w_i + 1/\delta] \setminus [w_i + w_i', w_i + w_i' + 1/\delta]\right] \\
&\le \sum_{i=1}^d \mathbb{P}\left[\tilde{p}_i \in [w_i, w_i + 1/\delta] \setminus [w_i + w_i', w_i + w_i' + 1/\delta]\right] \\
&= \sum_{i=1}^d \min\{1, \delta |w_i'|\},
\end{aligned}$$

as required. □

Theorem 5 (Follow-the-Perturbed-Leader regret bound)**.** *Let $\boldsymbol{x_1}, \ldots, \boldsymbol{x_T}$ be the sequence of decisions chosen by FTPL for the sequence of linear vectors $\boldsymbol{w_1}, \ldots, \boldsymbol{w_T}$. Then*

$$\mathbb{E}\left[\sum_{t=1}^{T} \boldsymbol{w_t}^\top \boldsymbol{x_t}\right] - \min_{\boldsymbol{x}\in\mathcal{X}} \sum_{t=1}^{T} \boldsymbol{w_t}^\top \boldsymbol{x} \leq \frac{R}{\delta} + \delta d R \sum_{t=1}^{T} \|\boldsymbol{w_t}\|_\infty^2$$

where the expectation is taken with respect to the randomness of the random perturbation vectors.

Proof. According to Lemma B.2 we have

$$\begin{aligned}\sum_{t=1}^{T} M\left(\boldsymbol{p} + \sum_{s=1}^{t} \boldsymbol{w_s}\right)^\top \boldsymbol{w_t} &\leq \sum_{t=1}^{T} M\left(\sum_{s=1}^{T} \boldsymbol{w_s}\right)^\top \boldsymbol{w_t} + R\|\boldsymbol{p}\|_\infty \\ &= \min_{\boldsymbol{x}\in\mathcal{X}} \sum_{t=1}^{T} \boldsymbol{w_t}^\top \boldsymbol{x} + R\|\boldsymbol{p}\|_\infty \\ &\leq \min_{\boldsymbol{x}\in\mathcal{X}} \sum_{t=1}^{T} \boldsymbol{w_t}^\top \boldsymbol{x} + \frac{R}{\delta}\end{aligned}$$

where the final inequality follows since $\boldsymbol{p} \in [0, 1/\delta]^d$. Remember that $\boldsymbol{x_t} = M\left(\boldsymbol{p} + \sum_{s=1}^{t-1} \boldsymbol{w_s}\right)$. Therefore,

$$\sum_{t=1}^{T} \boldsymbol{w_t}^\top \boldsymbol{x_t} \leq \min_{\boldsymbol{x}\in\mathcal{X}} \sum_{t=1}^{T} \boldsymbol{w_t}^\top \boldsymbol{x} + \frac{R}{\delta} + \sum_{t=1}^{T}\left(M\left(\boldsymbol{p} + \sum_{s=1}^{t-1} \boldsymbol{w_s}\right) - M\left(\boldsymbol{p} + \sum_{s=1}^{t} \boldsymbol{w_s}\right)\right)^\top \boldsymbol{w_t}.$$

Take expectations of both sides to get

$$\begin{aligned}&\mathbb{E}_{\boldsymbol{p}}\left[\sum_{t=1}^{T} \boldsymbol{w_t}^\top \boldsymbol{x_t}\right] \\ &\leq \min_{\boldsymbol{x}\in\mathcal{X}} \sum_{t=1}^{T} \boldsymbol{w_t}^\top \boldsymbol{x} + \frac{R}{\delta} + \sum_{t=1}^{T} \mathbb{E}_{\boldsymbol{p}}\left[M\left(\boldsymbol{p} + \sum_{s=1}^{t-1} \boldsymbol{w_s}\right) - M\left(\boldsymbol{p} + \sum_{s=1}^{t} \boldsymbol{w_s}\right)\right]^\top \boldsymbol{w_t}.\end{aligned}$$

We now apply Lemma B.3 to each term in the last sum to get

$$\begin{aligned}\mathbb{E}_{\boldsymbol{p}}\left[\sum_{t=1}^{T} \boldsymbol{w_t}^\top \boldsymbol{x_t}\right] - \min_{\boldsymbol{x}\in\mathcal{X}} \sum_{t=1}^{T} \boldsymbol{w_t}^\top \boldsymbol{x} &\leq \frac{R}{\delta} + \sum_{t=1}^{T} R\|\boldsymbol{w_t}\|_\infty \sum_{i=1}^{d} \min\{1, \delta|w_{t,i}|\} \\ &\leq \frac{R}{\delta} + \sum_{t=1}^{T} R\|\boldsymbol{w_t}\|_\infty \sum_{i=1}^{d} \delta|w_{t,i}| \\ &\leq \frac{R}{\delta} + \delta d R \sum_{t=1}^{T} \|\boldsymbol{w_t}\|_\infty^2,\end{aligned}$$

as required. □

Having proved Theorem 5, we can prove Lemma 5.1. Let us define filtration $\{\mathcal{F}_t : t \geq 0\}$ where $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_t = \sigma(\boldsymbol{\omega_1}, \ldots, \boldsymbol{\omega_t})$ being the σ-algebra generated by the set of random samples $\{\boldsymbol{\omega_1}, \ldots, \boldsymbol{\omega_t}\}$. Note that $\boldsymbol{x_t}$ and λ_t are $\mathcal{F}_{t-1}$-measurable for all $t \geq 1$.

Proof of Lemma 5.1. Note that the expected regret of FTPL under the sequence of linear functions $h_1^k(\boldsymbol{v}), \ldots, h_T^k(\boldsymbol{v})$ can be bounded using Theorem 5, as follows.

$$
\begin{aligned}
\mathbb{E}\left[\sum_{t=1}^T h_t^k(\boldsymbol{v_t^k}) - \min_{\boldsymbol{v}\in\mathcal{X}} \sum_{t=1}^T h_t^k(\boldsymbol{v})\right] &= \mathbb{E}\left[\mathbb{E}\left[\sum_{t=1}^T h_t^k(\boldsymbol{v_t^k}) \mid \mathcal{F}_T\right] - \min_{\boldsymbol{v}\in\mathcal{X}} \sum_{t=1}^T h_t^k(\boldsymbol{v})\right] \\
&\leq \mathbb{E}\left[\frac{R}{\delta} + \delta d R \sum_{t=1}^T \|\nabla f_t(\boldsymbol{x_t^k}) + \lambda_t \nabla g_t(\boldsymbol{x_t^k})\|_\infty^2\right] \\
&\leq \mathbb{E}\left[\frac{R}{\delta} + 2\delta d R \sum_{t=1}^T \left(\|\nabla f_t(\boldsymbol{x_t^k})\|_\infty^2 + \lambda_t^2 \|\nabla g_t(\boldsymbol{x_t^k})\|_\infty^2\right)\right]
\end{aligned} \tag{11}
$$

where the first equality holds due to the tower rule, the first inequality follows from Theorem 5, and the second inequality holds because $\|\boldsymbol{x}+\boldsymbol{y}\|_\infty^2 \leq (\|\boldsymbol{x}\|_\infty + \|\boldsymbol{y}\|_\infty)^2 \leq 2\|\boldsymbol{x}\|_\infty^2 + 2\|\boldsymbol{y}\|_\infty^2$ for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$. Note that

$$
\begin{aligned}
&\mathbb{E}\left[\frac{R}{\delta} + 2\delta d R \sum_{t=1}^T \left(\|\nabla f_t(\boldsymbol{x_t^k})\|_\infty^2 + \lambda_t^2 \|\nabla g_t(\boldsymbol{x_t^k})\|_\infty^2\right)\right] \\
&= \frac{R}{\delta} + 2\delta d R \sum_{t=1}^T \left(\|\nabla f_t(\boldsymbol{x_t^k})\|_\infty^2 + \mathbb{E}\left[\lambda_t^2 \|\nabla g_t(\boldsymbol{x_t^k})\|_\infty^2\right]\right) \\
&\leq \frac{R}{\delta} + 2\delta d R \sum_{t=1}^T \left(D^2 + D^2 \mathbb{E}\left[\lambda_t^2\right]\right)
\end{aligned} \tag{12}
$$

where the first inequality follows from Assumption 2. Since

$$
\delta = \frac{1}{2D\sqrt{d}T^{\frac{1}{2}+\beta}},
$$

it follows from (11) that

$$
\begin{aligned}
\mathbb{E}\left[\sum_{t=1}^T h_t^k(\boldsymbol{v_t^k}) - \min_{\boldsymbol{v}\in\mathcal{X}} \sum_{t=1}^T h_t^k(\boldsymbol{v})\right] &\leq \frac{R}{\delta} + 2R\delta d \sum_{t=1}^T (D^2 + D^2 \mathbb{E}\left[\lambda_t^2\right]) \\
&= RD\sqrt{d}\left(3T^{\frac{1}{2}+\beta} + \frac{1}{T^{\frac{1}{2}+\beta}} \mathbb{E}\left[\sum_{t=1}^T \lambda_t^2\right]\right),
\end{aligned}
$$

as required. □

B.2 Regret: Proof of Theorem 2

Lemma B.4. *Let $\gamma_k = 2/(k+1)$ for $k \geq 1$. Then for any $\ell \leq K+1$,*

$$
\prod_{k=\ell}^K (1-\gamma_k) \leq \left(\frac{\ell+1}{K+2}\right)^2.
$$

Proof. If $\ell = K+1$, then both sides are equal to 1, so the inequality is satisfied. Assume that $\ell \leq K$. Then

$$
\prod_{k=\ell}^K (1-\gamma_k) \leq \exp\left(-\sum_{k=\ell}^K \frac{2}{k+1}\right) \leq \exp\left(-\int_{x=\ell+1}^{K+2} \frac{2}{x} dx\right) = \left(\frac{\ell+1}{K+2}\right)^2,
$$

as required. □

First, we prove the following lemma, which holds because f_t and g_t are smooth (Assumption 3). We closely follow the proof of Chen et al. [3, Theorem 1]. We define a constant C_3 as

$$C_3 = 4RD + 5R^2L + \frac{R(4D+5RL)^2}{4D\sqrt{d}} + 3RD\sqrt{d}.$$

Lemma B.5. *Let $\gamma_k = 2/(k+1)$. Then for any $\boldsymbol{x} \in \mathcal{X}$,*

$$\mathbb{E}\left[\sum_{t=1}^{T} L_t(\boldsymbol{x_t}, \lambda_t) - \sum_{t=1}^{T} L_t(\boldsymbol{x}, \lambda_t)\right] \le C_3 T^{\frac{1}{2}+\beta} + \frac{2RD\sqrt{d}}{T^{\frac{1}{2}+\beta}} \sum_{t=1}^{T} \mathbb{E}\left[\lambda_t^2\right].$$

Proof. Observe first that

$$\sum_{t=1}^{T} L_t(\boldsymbol{x_t}, \lambda_t) - \sum_{t=1}^{T} L_t(\boldsymbol{x}, \lambda_t) = \sum_{t=1}^{T} \left(f_t(\boldsymbol{x_t}) - f_t(\boldsymbol{x})\right) + \sum_{t=1}^{T} \lambda_t \left(g_t(\boldsymbol{x_t}) - g_t(\boldsymbol{x})\right). \tag{13}$$

Let $t \ge 1$. Note that for any $1 \le k \le K$,

$$\begin{aligned}
& f_t(\boldsymbol{x_t^{k+1}}) - f_t(\boldsymbol{x}) \\
&= f_t(\boldsymbol{x_t^k} + \gamma_k(\boldsymbol{v_t^k} - \boldsymbol{x_t^k})) - f_t(\boldsymbol{x}) \\
&\le f_t(\boldsymbol{x_t^k}) - f_t(\boldsymbol{x}) + \gamma_k \nabla f_t(\boldsymbol{x_t^k})^\top (\boldsymbol{v_t^k} - \boldsymbol{x_t^k}) + \frac{L\gamma_k^2 R^2}{2} \\
&= f_t(\boldsymbol{x_t^k}) - f_t(\boldsymbol{x}) + \gamma_k \nabla f_t(\boldsymbol{x_t^k})^\top (\boldsymbol{x} - \boldsymbol{x_t^k}) + \gamma_k \nabla f_t(\boldsymbol{x_t^k})^\top (\boldsymbol{v_t^k} - \boldsymbol{x}) + \frac{L\gamma_k^2 R^2}{2} \\
&\le (1-\gamma_k)\big(f_t(\boldsymbol{x_t^k}) - f_t(\boldsymbol{x})\big) + \gamma_k \nabla f_t(\boldsymbol{x_t^k})^\top (\boldsymbol{v_t^k} - \boldsymbol{x}) + \frac{L\gamma_k^2 R^2}{2}
\end{aligned} \tag{14}$$

where the first inequality holds because f_t is L-smooth and $\|\boldsymbol{v_t^k} - \boldsymbol{x_t^k}\|_1 \le R$ whereas the second inequality follows from the convexity of f_t. Then it follows from (14) that

$$\begin{aligned}
f_t(\boldsymbol{x_t}) - f_t(\boldsymbol{x}) &= f_t(\boldsymbol{x_t^{K+1}}) - f_t(\boldsymbol{x}) \\
&\le \prod_{k=1}^{K}(1-\gamma_k)\big(f_t(\boldsymbol{x_t^1}) - f_t(\boldsymbol{x})\big) \\
&\quad + \sum_{k=1}^{K} \gamma_k \left(\nabla f_t(\boldsymbol{x_t^k})^\top (\boldsymbol{v_t^k} - \boldsymbol{x}) + \frac{L\gamma_k R^2}{2}\right) \prod_{j=k+1}^{K}(1-\gamma_j) \\
&= \prod_{k=1}^{K}(1-\gamma_k)\left(f_t(\boldsymbol{x_{t-1}}) - f_t(\boldsymbol{x})\right) \\
&\quad + \sum_{k=1}^{K} \gamma_k \left(\nabla f_t(\boldsymbol{x_t^k})^\top (\boldsymbol{v_t^k} - \boldsymbol{x}) + \frac{L\gamma_k R^2}{2}\right) \prod_{j=k+1}^{K}(1-\gamma_j) \\
&\le \prod_{k=1}^{K}(1-\gamma_k) DR \\
&\quad + \sum_{k=1}^{K} \gamma_k \left(\nabla f_t(\boldsymbol{x_t^k})^\top (\boldsymbol{v_t^k} - \boldsymbol{x}) + \frac{L\gamma_k R^2}{2}\right) \prod_{j=k+1}^{K}(1-\gamma_j)
\end{aligned} \tag{15}$$

where the last inequality holds because f_t is D-Lipschitz. Similarly, we deduce that

$$\begin{aligned}
&\mathbb{E}\left[\lambda_t\left(g_t(\boldsymbol{x_t})-g_t(\boldsymbol{x})\right)\right] \\
&\leq \prod_{k=1}^{K}(1-\gamma_k)DR\mathbb{E}\left[\lambda_t\right] \\
&\quad+\sum_{k=1}^{K}\gamma_k\left(\mathbb{E}\left[\lambda_t\nabla g_t(\boldsymbol{x_t^k})^\top(\boldsymbol{v_t^k}-\boldsymbol{x})\right]+\frac{L\gamma_k R^2}{2}\mathbb{E}\left[\lambda_t\right]\right)\prod_{j=k+1}^{K}(1-\gamma_j).
\end{aligned} \tag{16}$$

Summing up (15) and (16) for $t=1,\dots,T$, we obtain

$$\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^{T}L_t(\boldsymbol{x_t},\lambda_t)-\sum_{t=1}^{T}L_t(\boldsymbol{x},\lambda_t)\right] \\
&\leq \sum_{t=1}^{T}\prod_{k=1}^{K}(1-\gamma_k)DR\left(1+\mathbb{E}\left[\lambda_t\right]\right) \\
&\quad+\sum_{t=1}^{T}\sum_{k=1}^{K}\gamma_k\mathbb{E}\left[\left(\nabla f_t(\boldsymbol{x_t^k})+\lambda_t\nabla g_t(\boldsymbol{x_t^k})\right)^\top(\boldsymbol{v_t^k}-\boldsymbol{x})\right]\prod_{j=k+1}^{K}(1-\gamma_j) \\
&\quad+\sum_{t=1}^{T}\sum_{k=1}^{K}\frac{L\gamma_k^2R^2}{2}\left(1+\mathbb{E}\left[\lambda_t\right]\right)\prod_{j=k+1}^{K}(1-\gamma_j).
\end{aligned} \tag{17}$$

Here,

$$\begin{aligned}
&\sum_{t=1}^{T}\prod_{k=1}^{K}(1-\gamma_k)DR\left(1+\mathbb{E}\left[\lambda_t\right]\right)+\sum_{t=1}^{T}\sum_{k=1}^{K}\frac{L\gamma_k^2R^2}{2}\left(1+\mathbb{E}\left[\lambda_t\right]\right)\prod_{j=k+1}^{K}(1-\gamma_j) \\
&\leq \sum_{t=1}^{T}\frac{4RD}{(K+2)^2}\left(1+\mathbb{E}\left[\lambda_t\right]\right)+\sum_{t=1}^{T}\sum_{k=1}^{K}\frac{2R^2L}{(k+1)^2}\left(\frac{k+2}{K+2}\right)^2\left(1+\mathbb{E}\left[\lambda_t\right]\right) \\
&\leq \sum_{t=1}^{T}\frac{4RD}{K+2}\left(1+\mathbb{E}\left[\lambda_t\right]\right)+\sum_{t=1}^{T}\frac{9R^2LK}{2\left(K+2\right)^2}\left(1+\mathbb{E}\left[\lambda_t\right]\right) \\
&\leq \sum_{t=1}^{T}\frac{4RD}{T^{\frac{1}{2}+\beta}}\left(1+\mathbb{E}\left[\lambda_t\right]\right)+\sum_{t=1}^{T}\frac{5R^2L}{T^{\frac{1}{2}+\beta}}\left(1+\mathbb{E}\left[\lambda_t\right]\right)
\end{aligned} \tag{18}$$

where the first inequality follows from Lemma B.4, the second inequality holds because $K+2\geq 1$ and $2(k+2)\leq$

$3(k+1)$, and the third inequality is obtained using $K+2 \geq T^{\frac{1}{2}+\beta}$ and $K \leq T^{\frac{1}{2}+\beta}$. Furthermore,

$$\begin{aligned}
&\sum_{t=1}^{T}\sum_{k=1}^{K}\gamma_k \prod_{j=k+1}^{K}(1-\gamma_j)\mathbb{E}\left[\left(\nabla f_t(\boldsymbol{x_t^k}) + \lambda_t \nabla g_t(\boldsymbol{x_t^k})\right)^\top (\boldsymbol{v_t^k} - \boldsymbol{x})\right] \\
&= \sum_{k=1}^{K}\gamma_k \prod_{j=k+1}^{K}(1-\gamma_j)\sum_{t=1}^{T}\mathbb{E}\left[\left(\nabla f_t(\boldsymbol{x_t^k}) + \lambda_t \nabla g_t(\boldsymbol{x_t^k})\right)^\top (\boldsymbol{v_t^k} - \boldsymbol{x})\right] \\
&= \sum_{k=1}^{K}\gamma_k \prod_{j=k+1}^{K}(1-\gamma_j)\mathbb{E}\left[\sum_{t=1}^{T} h_t^k(\boldsymbol{v_t^k}) - \sum_{t=1}^{T} h_t^k(\boldsymbol{x})\right] \\
&\leq \sum_{k=1}^{K}\gamma_k \prod_{j=k+1}^{K}(1-\gamma_j)\mathcal{R}^{\mathcal{E}}(T) \\
&\leq \sum_{k=1}^{K}\frac{2}{(k+1)}\left(\frac{k+2}{K+2}\right)^2\mathcal{R}^{\mathcal{E}}(T) \\
&\leq \sum_{k=1}^{K}\frac{3(k+2)}{(K+2)^2}\mathcal{R}^{\mathcal{E}}(T) \\
&\leq \frac{3}{2}\mathcal{R}^{\mathcal{E}}(T)
\end{aligned} \tag{19}$$

where the second inequality is from Lemma B.4. By Lemma 5.1, (17), (18), and (19),

$$\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^{T} L_t(\boldsymbol{x_t}, \lambda_t) - \sum_{t=1}^{T} L_t(\boldsymbol{x}, \lambda_t)\right] \\
&\leq \sum_{t=1}^{T}\frac{4RD + 5R^2L}{T^{\frac{1}{2}+\beta}}\left(1 + \mathbb{E}\left[\lambda_t\right]\right) + RD\sqrt{d}\left(3T^{\frac{1}{2}+\beta} + \frac{1}{T^{\frac{1}{2}+\beta}}\sum_{t=1}^{T}\mathbb{E}[\lambda_t^2]\right).
\end{aligned} \tag{20}$$

Observe that

$$\frac{4RD + 5R^2L}{T^{\frac{1}{2}+\beta}}\lambda_t \leq \frac{RD\sqrt{d}}{T^{\frac{1}{2}+\beta}}\lambda_t^2 + \frac{(4RD + 5R^2L)^2}{4RD\sqrt{d}T^{\frac{1}{2}+\beta}} = \frac{RD\sqrt{d}}{T^{\frac{1}{2}+\beta}}\lambda_t^2 + \frac{R(4D + 5RL)^2}{4D\sqrt{d}T^{\frac{1}{2}+\beta}}. \tag{21}$$

as $2\sqrt{pq} \leq p + q$ for any $p, q \geq 0$. Based on (20) and (21),

$$\begin{aligned}
\mathbb{E}\left[\sum_{t=1}^{T} L_t(\boldsymbol{x_t}, \lambda_t) - \sum_{t=1}^{T} L_t(\boldsymbol{x}, \lambda_t)\right] &\leq \sum_{t=1}^{T}\left(4RD + 5R^2L + \frac{R(4D + 5RL)^2}{4D\sqrt{d}}\right)\frac{1}{T^{\frac{1}{2}+\beta}} \\
&\quad + 3RD\sqrt{d}T^{\frac{1}{2}+\beta} + \frac{2RD\sqrt{d}}{T^{\frac{1}{2}+\beta}}\sum_{t=1}^{T}\mathbb{E}\left[\lambda_t^2\right] \\
&\leq C_3 T^{\frac{1}{2}+\beta} + \frac{2RD\sqrt{d}}{T^{\frac{1}{2}+\beta}}\sum_{t=1}^{T}\mathbb{E}\left[\lambda_t^2\right],
\end{aligned}$$

as required. □

Next, based on the concavity of $L_t(\boldsymbol{x_t}, \lambda)$ with respect to λ, we show the following lemma.

Lemma B.6. *For any $\lambda \geq 0$,*

$$\mathbb{E}\left[\sum_{t=1}^{T} L_t(\boldsymbol{x_t}, \lambda) - \sum_{t=1}^{T} L_t(\boldsymbol{x_t}, \lambda_t)\right] \leq \frac{1}{2\mu}\lambda^2 + G^2T\mu + \mu\theta^2\sum_{t=1}^{T}\mathbb{E}\left[\lambda_t^2\right].$$

Proof. Note that

$$
\begin{aligned}
&(\lambda_{t+1}-\lambda)^2 \\
&\le (\lambda_t + \mu\nabla_\lambda L_t(\boldsymbol{x_t},\lambda_t)-\lambda)^2 \\
&= (\lambda_t-\lambda)^2 + \mu^2(g_t(\boldsymbol{x_t})-\theta\lambda_t)^2 + 2\mu\nabla_\lambda L_t(\boldsymbol{x_t},\lambda_t)^\top(\lambda_t-\lambda) \\
&\le (\lambda_t-\lambda)^2 + \mu^2(g_t(\boldsymbol{x_t})-\theta\lambda_t)^2 - 2\mu\left(L_t(\boldsymbol{x_t},\lambda)-L_t(\boldsymbol{x_t},\lambda_t)\right) - \mu\theta(\lambda_t-\lambda)^2
\end{aligned} \tag{22}
$$

where the first inequality is because λ_{t+1} is the projection of $\lambda_t + \mu\nabla_\lambda L_t(\boldsymbol{x_t},\lambda_t)$, the equality is because $\nabla_\lambda L_t(\boldsymbol{x_t},\lambda_t) = g_t(\boldsymbol{x_t}) - \theta\lambda_t$, the second inequality holds because L_t is θ-strongly concave with respect to λ. Note that

$$
\begin{aligned}
\mathbb{E}\left[(g_t(\boldsymbol{x_t})-\theta\lambda_t)^2\right] &\le \mathbb{E}\left[2(g_t(\boldsymbol{x_t}))^2 + 2\theta^2\lambda_t^2\right] \\
&= 2\mathbb{E}\left[\mathbb{E}\left[(g_t(\boldsymbol{x_t}))^2 + 2\theta^2\lambda_t^2 \mid \mathcal{F}_{t-1}\right]\right] \\
&\le 2G^2 + 2\theta^2\mathbb{E}\left[\lambda_t^2\right].
\end{aligned} \tag{23}
$$

Combining (22) and (23),

$$
\begin{aligned}
&\mathbb{E}\left[(\lambda_{t+1}-\lambda)^2\right] \\
&\le \mathbb{E}\left[(\lambda_t-\lambda)^2\right] + \mu^2(2G^2+2\theta^2\lambda_t^2) - \mathbb{E}\left[2\mu\left(L_t(\boldsymbol{x_t},\lambda)-L_t(\boldsymbol{x_t},\lambda_t)\right)\right] - \mu\theta(\lambda_t-\lambda)^2.
\end{aligned} \tag{24}
$$

Then it follows from (24) that

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^T L_t(\boldsymbol{x_t},\lambda) - \sum_{t=1}^T L_t(\boldsymbol{x_t},\lambda_t)\right] \\
&\le \mathbb{E}\left[\sum_{t=1}^T \frac{1}{2\mu}\left((\lambda_t-\lambda)^2-(\lambda_{t+1}-\lambda)^2\right) + \sum_{t=1}^T \mu(G^2+\theta^2\lambda_t^2) - \sum_{t=1}^T \frac{\theta}{2}(\lambda_t-\lambda)^2\right] \\
&\le \frac{1}{2\mu}(\lambda_1-\lambda)^2 + G^2T\mu + \sum_{t=1}^T \mu\theta^2\mathbb{E}\left[\lambda_t^2\right] \\
&= \frac{1}{2\mu}\lambda^2 + G^2T\mu + \sum_{t=1}^T \mu\theta^2\mathbb{E}\left[\lambda_t^2\right]
\end{aligned}
$$

where the last inequality holds due to $\lambda_1 = 0$. Therefore, the assertion of the lemma holds, as required. $\square$

Based on Lemmas B.5 and B.6, we show the following lemma. Recall that

$$
C_1 = 4RD + 5R^2L + \frac{R(4D+5RL)^2}{4D\sqrt{d}} + 3RD\sqrt{d} + \frac{G^2}{12RD\sqrt{d}} = C_3 + \frac{G^2}{12RD\sqrt{d}}.
$$

Lemma B.7. *Let $\gamma_k = 2/(k+1)$. Then for any $\boldsymbol{x}\in\mathcal{X}$ and $\lambda \ge 0$,*

$$
\begin{aligned}
&\mathbb{E}\left[\left(\sum_{t=1}^T f_t(\boldsymbol{x_t}) - \sum_{t=1}^T f_t(\boldsymbol{x})\right) + \left(\sum_{t=1}^T \lambda g_t(\boldsymbol{x_t}) - \sum_{t=1}^T \lambda_t g_t(\boldsymbol{x})\right)\right] \\
&\le C_1 T^{\frac{1}{2}+\beta} + 24RD\sqrt{d}T^{\frac{1}{2}-\beta}\lambda^2.
\end{aligned}
$$

Proof. Adding the two inequalities proved in Lemmas B.5 and B.6, we obtain

$$\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^{T} L_t(\boldsymbol{x_t}, \lambda) - \sum_{t=1}^{T} L_t(\boldsymbol{x}, \lambda_t)\right] \\
&\leq C_3 T^{\frac{1}{2}+\beta} + \frac{2RD\sqrt{d}}{T^{\frac{1}{2}+\beta}} \sum_{t=1}^{T} \mathbb{E}\left[\lambda_t^2\right] + \frac{1}{2\mu}\lambda^2 + G^2 T\mu + \sum_{t=1}^{T} \mu\theta^2 \mathbb{E}\left[\lambda_t^2\right].
\end{aligned} \tag{25}$$

By (25),

$$\begin{aligned}
&\mathbb{E}\left[\left(\sum_{t=1}^{T} f_t(\boldsymbol{x_t}) - \sum_{t=1}^{T} f_t(\boldsymbol{x})\right) + \left(\sum_{t=1}^{T} \lambda g_t(\boldsymbol{x_t}) - \sum_{t=1}^{T} \lambda_t g_t(\boldsymbol{x})\right)\right] \\
&\leq C_3 T^{\frac{1}{2}+\beta} + G^2 T\mu + \left(\frac{1}{2}T\theta + \frac{1}{2\mu}\right)\lambda^2 + \sum_{t=1}^{T}\left(\frac{2RD\sqrt{d}}{T^{\beta}} + \mu\theta^2 - \frac{\theta}{2}\right)\mathbb{E}\left[\lambda_t^2\right] \\
&\leq C_3 T^{\frac{1}{2}+\beta} + G^2 T\mu + \left(\frac{1}{2}T\theta + \frac{1}{2\mu}\right)\lambda^2 \\
&\leq C_3 T^{\frac{1}{2}+\beta} + \frac{G^2}{12RD\sqrt{d}} T^{\frac{1}{2}+\beta} + 24RD\sqrt{d}T^{\frac{1}{2}-\beta}\lambda^2 \\
&= C_1 T^{\frac{1}{2}+\beta} + 24RD\sqrt{d}T^{\frac{1}{2}-\beta}\lambda^2
\end{aligned}$$

where the second inequality holds because $\lambda_1 = 0$ and for $T \geq 1$,

$$\frac{2RD\sqrt{d}}{T^{\frac{1}{2}+\beta}} + \mu\theta^2 - \frac{\theta}{2} = \frac{\theta}{6} + \frac{\theta}{T+2} - \frac{\theta}{2} \leq \frac{\theta}{6} + \frac{\theta}{3} - \frac{\theta}{2} = 0,$$

as required. □

Lastly, we need the following lemma.

Lemma B.8. *For any $\{\lambda_t\}_{t=1}^{T} \subseteq \mathbb{R}_+$ such that λ_t is $\mathcal{F}_{t-1}$-measurable for $t \geq 1$, then*

$$\mathbb{E}\left[\sum_{t=1}^{T} \lambda_t g_t(\boldsymbol{x}^*)\right] \leq 0$$

Proof. Note that

$$\begin{aligned}
\mathbb{E}[\lambda_t g_t(\boldsymbol{x}^*)] &= \mathbb{E}\left[\mathbb{E}[\lambda_t g_t(\boldsymbol{x}^*) \mid \mathcal{F}_{t-1}]\right] \\
&= \mathbb{E}\left[\lambda_t \mathbb{E}[g_t(\boldsymbol{x}^*) \mid \mathcal{F}_{t-1}]\right] \\
&= \mathbb{E}\left[\lambda_t \bar{g}(\boldsymbol{x}^*)\right] \\
&= \mathbb{E}\left[\lambda_t\right] \bar{g}(\boldsymbol{x}^*) \\
&\leq 0
\end{aligned}$$

where the first equality comes from the tower rule, the second equality holds because λ_t is $\mathcal{F}_{t-1}$-measurable, the third equality follows as g_t is independent of $\{\boldsymbol{\omega_1}, \ldots, \boldsymbol{\omega_{t-1}}\}$, the fourth equality holds because $\bar{g}(\boldsymbol{x}^*)$ is a constant, and the last inequality follows from $\lambda_t \geq 0$ and $\bar{g}(\boldsymbol{x}^*) \leq 0$. Then

$$\mathbb{E}\left[\sum_{t=1}^{T} \lambda_t g_t(\boldsymbol{x}^*)\right] = \sum_{t=1}^{T} \mathbb{E}[\lambda_t g_t(\boldsymbol{x}^*)] \leq 0,$$

as required. □

Now we are ready to prove Theorem 2.

Proof of Theorem 2. By Lemma B.7,

$$\begin{aligned}&\mathbb{E}\left[\left(\sum_{t=1}^{T} f_t(\boldsymbol{x_t}) - \sum_{t=1}^{T} f_t(\boldsymbol{x})\right) + \left(\sum_{t=1}^{T} \lambda g_t(\boldsymbol{x_t}) - \sum_{t=1}^{T} \lambda_t g_t(\boldsymbol{x})\right)\right] \\ &\leq C_1 T^{\frac{1}{2}+\beta} + 24RD\sqrt{d}T^{\frac{1}{2}-\beta}\lambda^2.\end{aligned} \tag{26}$$

Here, we set

$$\lambda = \left(48RD\sqrt{d}T^{\frac{1}{2}-\beta}\right)^{-1}\left[\mathbb{E}\left[\sum_{t=1}^{T} g_t(\boldsymbol{x_t})\right]\right]_+$$

where $(p)_+ = \max\{p, 0\}$. Then

$$\mathbb{E}\left[\sum_{t=1}^{T} \lambda g_t(\boldsymbol{x_t})\right] = \left(48RD\sqrt{d}T^{\frac{1}{2}-\beta}\right)^{-1}\left[\mathbb{E}\left[\sum_{t=1}^{T} g_t(\boldsymbol{x_t})\right]\right]_+^2 = 48RD\sqrt{d}T^{\frac{1}{2}-\beta}\lambda^2.$$

Together with (26), this implies

$$\begin{aligned}&\mathbb{E}\left[\sum_{t=1}^{T} f_t(\boldsymbol{x_t}) - \sum_{t=1}^{T} f_t(\boldsymbol{x})\right] \\ &\leq C_1 T^{\frac{1}{2}+\beta} + \mathbb{E}\left[\sum_{t=1}^{T} \lambda_t g_t(\boldsymbol{x})\right] - \frac{1}{96RD\sqrt{d}T^{\frac{1}{2}-\beta}}\left[\mathbb{E}\left[\sum_{t=1}^{T} g_t(\boldsymbol{x_t})\right]\right]_+^2.\end{aligned} \tag{27}$$

In particular, we consider $\boldsymbol{x} = \boldsymbol{x}^*$. Note that λ_t is $\mathcal{F}_{t-1}$-measurable for all $t \geq 1$. Then by Lemma B.8,

$$\mathbb{E}\left[\sum_{t=1}^{T} \lambda_t g_t(\boldsymbol{x}^*)\right] \leq 0 \tag{28}$$

Then it follows from (27) and (28) that

$$\mathbb{E}\left[\sum_{t=1}^{T} f_t(\boldsymbol{x_t}) - \sum_{t=1}^{T} f_t(\boldsymbol{x}^*)\right] \leq C_1 T^{\frac{1}{2}+\beta}.$$

as required. Moreover, Since $f_1, \ldots, f_T$ are D-Lipschitz, it follows from (27) that

$$\left[\mathbb{E}\left[\sum_{t=1}^{T} g_t(\boldsymbol{x_t})\right]\right]_+^2 \leq 96RD\sqrt{d}T^{\frac{1}{2}-\beta}\left(RDT + C_1 T^{\frac{1}{2}+\beta}\right).$$

Therefore,

$$\mathbb{E}\left[\sum_{t=1}^{T} g_t(\boldsymbol{x_t})\right] \leq C_2 T^{\frac{3}{4}-\frac{\beta}{2}}$$

where

$$C_2 = \sqrt{96RD\sqrt{d}\,(RD + C_1)},$$

as required. $\square$

C Analysis for Stochastic Loss Functions under Strong Duality (Section 6)

We first prove Lemmas 6.1 and 6.2. Then based on these lemmas, we prove Theorems 3 and 4.

Proof of Lemma 6.1. Note that

$$\begin{aligned}
\mathbb{E}\left[\sum_{t=1}^{T} g_t(\boldsymbol{x_t})\right] &= \mathbb{E}\left[\sum_{q=1}^{Q}\sum_{k=1}^{K} g_{(q-1)K+k}(\boldsymbol{x_{(q-1)K+k}})\right] \\
&= \sum_{q=1}^{Q}\sum_{k=1}^{K} \mathbb{E}\left[\mathbb{E}\left[g_{(q-1)K+k}(\boldsymbol{x_{(q-1)K+k}}) \mid \mathcal{F}_{(q-1)K+k-1}\right]\right] \\
&= \mathbb{E}\left[\sum_{q=1}^{Q}\sum_{k=1}^{K} \bar{g}(\boldsymbol{x_{(q-1)K+k}})\right] \\
&= \mathbb{E}\left[\sum_{q=1}^{Q}\sum_{k=1}^{K} \bar{L}(\boldsymbol{x_{(q-1)K+k}}, \lambda^* + 1) - \sum_{q=1}^{Q}\sum_{k=1}^{K} \bar{L}(\boldsymbol{x_{(q-1)K+k}}, \lambda^*)\right] \\
&\leq \mathbb{E}\left[\sum_{q=1}^{Q}\sum_{k=1}^{K} \bar{L}(\boldsymbol{x_{(q-1)K+k}}, \lambda^* + 1) - \sum_{q=1}^{Q}\sum_{k=1}^{K} \bar{L}(\boldsymbol{x^*}, \lambda_q)\right]
\end{aligned}$$

where the last inequality follows from (7). Moreover, note that

$$\begin{aligned}
&\mathbb{E}\left[\sum_{q=1}^{Q}\sum_{k=1}^{K} \left(f_{(q-1)K+k}(\boldsymbol{x_{(q-1)K+k}}) + (\lambda^* + 1) g_{(q-1)K+k}(\boldsymbol{x_{(q-1)K+k}})\right)\right] \\
&= \sum_{q=1}^{Q}\sum_{k=1}^{K} \mathbb{E}\left[\mathbb{E}\left[f_{(q-1)K+k}(\boldsymbol{x_{(q-1)K+k}}) + (\lambda^* + 1) g_{(q-1)K+k}(\boldsymbol{x_{(q-1)K+k}}) \mid \mathcal{F}_{(q-1)K+k-1}\right]\right] \\
&= \mathbb{E}\left[\sum_{q=1}^{Q}\sum_{k=1}^{K} \left(\bar{f}(\boldsymbol{x_{(q-1)K+k}}) + (\lambda^* + 1)\bar{g}(\boldsymbol{x_{(q-1)K+k}})\right)\right] \\
&= \mathbb{E}\left[\sum_{q=1}^{Q}\sum_{k=1}^{K} \bar{L}(\boldsymbol{x_{(q-1)K+k}}, \lambda^* + 1)\right].
\end{aligned}$$

Similarly,

$$\mathbb{E}\left[\sum_{q=1}^{Q}\sum_{k=1}^{K} \left(f_{(q-1)K+k}(\boldsymbol{x^*}) + \lambda_q g_{(q-1)K+k}(\boldsymbol{x^*})\right)\right] = \mathbb{E}\left[\sum_{t=1}^{T} \bar{L}(\boldsymbol{x^*}, \lambda_q)\right].$$

Therefore, the assertion follows, as required. $\square$

Proof of Lemma 6.2. Note that

$$\begin{aligned}
\mathbb{E}\left[\sum_{t=1}^{T} g_t(\boldsymbol{x_t})\right] &= \sum_{t=1}^{T} \mathbb{E}\left[\mathbb{E}\left[g_t(\boldsymbol{x_t}) \mid \mathcal{F}_{t-1}\right]\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{T} \bar{g}(\boldsymbol{x_t})\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{T} \bar{L}(\boldsymbol{x_t}, \lambda^* + 1) - \sum_{t=1}^{T} \bar{L}(\boldsymbol{x_t}, \lambda^*)\right] \\
&\le \mathbb{E}\left[\sum_{t=1}^{T} \bar{L}(\boldsymbol{x_t}, \lambda^* + 1) - \sum_{t=1}^{T} \bar{L}(\boldsymbol{x^*}, \lambda_t)\right]
\end{aligned}$$

where the last inequality follows from (7). Moreover, note that

$$\begin{aligned}
\mathbb{E}\left[\sum_{t=1}^{T} \left(f_t(\boldsymbol{x_t}) + (\lambda^* + 1) g_t(\boldsymbol{x_t})\right)\right] &= \sum_{t=1}^{T} \mathbb{E}\left[\mathbb{E}\left[f_t(\boldsymbol{x_t}) + (\lambda^* + 1) g_t(\boldsymbol{x_t}) \mid \mathcal{F}_{t-1}\right]\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{T} \left(\bar{f}(\boldsymbol{x_t}) + (\lambda^* + 1)\bar{g}(\boldsymbol{x_t})\right)\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{T} \bar{L}(\boldsymbol{x_t}, \lambda^* + 1)\right].
\end{aligned}$$

Similarly,

$$\mathbb{E}\left[\sum_{t=1}^{T} \left(f_t(\boldsymbol{x^*}) + \lambda_t g_t(\boldsymbol{x^*})\right)\right] = \mathbb{E}\left[\sum_{t=1}^{T} \bar{L}(\boldsymbol{x^*}, \lambda_t)\right].$$

Therefore, the assertion follows, as required. □

As we have shown Lemmas 6.1 and 6.2, we are ready to prove Theorems 3 and 4.

Proof of Theorem 3. By Theorem 1, we know that

$$\mathbb{E}\left[\sum_{t=1}^{T} f_t(\boldsymbol{x_t}) - \sum_{t=1}^{T} f_t(\boldsymbol{x^*})\right] = O\left(T^{\frac{2-\alpha}{3-2\alpha}}\right).$$

Therefore, it suffices to prove the upper bound on the constraint violation. Note that

$$\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^{T} g_t(\boldsymbol{x_t})\right] \\
&\le \mathbb{E}\left[\sum_{q=1}^{Q}\sum_{k=1}^{K} \left(f_{(q-1)K+k}(\boldsymbol{x_{(q-1)K+k}}) + (\lambda^* + 1) g_{(q-1)K+k}(\boldsymbol{x_{(q-1)K+k}})\right)\right] \\
&\quad - \mathbb{E}\left[\sum_{q=1}^{Q}\sum_{k=1}^{K} \left(f_{(q-1)K+k}(\boldsymbol{x^*}) + \lambda_q g_{(q-1)K+k}(\boldsymbol{x^*})\right)\right] \\
&\le \frac{1}{2}\left(\frac{1}{\mu} + Q\theta\right)\lambda^{*2} + G^2K^2Q\mu + Q(C_0 + C_1D + C_2L)T^{\frac{\alpha}{3-2\alpha}} + \frac{Q}{2}(C_1D + C_2L)T^{\frac{\alpha}{3-2\alpha}} \\
&\le O\left(T^{\frac{2-\alpha}{3-2\alpha}}\right)
\end{aligned}$$

where the first inequality is due to Lemma 6.1, the second inequality is from Lemma A.3, and the last inequality holds because of (3). □

Proof of Theorem 4. By Theorem 2, we know that

$$\mathbb{E}\left[\sum_{t=1}^{T} f_t(\boldsymbol{x_t}) - \sum_{t=1}^{T} f_t(\boldsymbol{x}^*)\right] = O\left(\sqrt{T}\right).$$

Therefore, it suffices to prove the upper bound on the constraint violation. Note that

$$\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^{T} g_t(\boldsymbol{x_t})\right] \\
&\leq \mathbb{E}\left[\sum_{t=1}^{T} (f_t(\boldsymbol{x_t}) + (\lambda^* + 1) g_t(\boldsymbol{x_t})) - \sum_{t=1}^{T} (f_t(\boldsymbol{x}^*) + \lambda_t g_t(\boldsymbol{x}^*))\right] \\
&\leq C_1 T^{\frac{1}{2}+\beta} + 24RD\sqrt{d}T^{\frac{1}{2}-\beta}\lambda^{*2} \\
&\leq O(\sqrt{T})
\end{aligned}$$

where the first inequality is due to Lemma 6.2, the second inequality is from Lemma B.7, and the last inequality holds because of (5). □