Forgetting-Factor Regrets for Online Convex Optimization

Yuhang Liu, Wenxiao Zhao, *Member, IEEE*, and George Yin, *Life Fellow, IEEE*

***Abstract*—This article develops a class of novel algorithms for online convex optimization. The key construct is a forgetting-factor regret. It introduces weights to the objective functions at each time instant t and allows the weights of the past objective functions decaying to zero. We establish the forgetting-factor regret bounds of classical algorithms including online gradient descent algorithms, online gradient-free algorithms, and online Frank–Wolfe algorithms. In addition, the article introduces online gradient descent algorithm with a forgetting factor, and analyze its performance under the new regret. Sufficient conditions are obtained to guarantee the bounds of the forgetting-factor regret of the above algorithms being of the order $\mathrm{o}(1)$, which guarantees the tracking performance for minimizers of time-varying objective functions. Finally, our results are tested through numerical demonstration.**

***Index Terms*—Forgetting-factor regret, iterative optimization algorithm, online convex optimization.**

I. Introduction

ONLINE convex optimization (OCO) has received much attention lately because many applications such as online routing, advertising selection for search engines, and spam filtering, etc., all fall into this category [1]. In such a scenario, optimization algorithms need to be performed for a sequence of convex objective functions $\{f_t(\cdot), t = 1, \ldots, T\}$. At each time instance t, the algorithm generates a prediction x_{t+1} based on available information including $\{x_i, i = 1, \ldots, t\}$ as well as $\{f_i(\cdot), i = 1, \ldots, t\}$, and then the loss of prediction $f_{t+1}(x_{t+1})$ is obtained for the next round of optimization process. Note that the OCO can also be considered for a continuous-time variable $t \in \mathbb{R}^+$; see, e.g., [2]. In this article, we mainly focus on the discrete time setting. Note that the optimal solutions of OCO are time varying. As a result, the traditional convex optimization algorithms for the time-invariant objective functions may not be feasible. A widely applied index for describing the performance of algorithms for time-varying objective functions is the so-called static regret function, denoted by regret_T^S in this article, which measures the cumulative difference between the loss of the estimates and the best-fixed points:

$$\text{regret}_T^S = \sum_{t=1}^{T} f_t(x_t) - \min_{u \in X} \sum_{t=1}^{T} f_t(u) \tag{1}$$

where X denotes the feasible set. An optimization algorithm for OCO is said to be acceptable if regret_T^S is sublinear with respect to T, that is,

$$\lim_{T \to \infty} \text{regret}_T^S / T = 0.$$

The definition of the regret function first appeared in [3], and has been widely applied ever since in areas including online learning [4], [5], [6], information theory [7], game theory [8], etc. Under this framework, performance of many classes of algorithms for OCO have been evaluated. For example, the projected online gradient descent (OGD) algorithm with $O(1/\sqrt{t})$ step size and the regularized follow-the-leader algorithm with $O(1/\sqrt{T})$ step size are studied in [9] and [10] and $O(\sqrt{T})$ regret bounds are given. The OGD algorithm with $O(1/t)$ step size is further investigated in [10] with a logarithmic regret bound $O(\log T)$ established. There are also many variants of the above algorithms with sublinear regrets, see the Frank–Wolfe approach-based projection-free algorithm [11], [12], the δ-smoothing-based algorithm for OCO without gradient information (named the bandit optimization in literature [13]), etc., and also [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], and references therein. However, as indicated in [27], algorithms with sublinear regrets do not necessarily guarantee a satisfactory tracking performance for minimizers of OCO.

Example (Hazan and Seshadhri [27]): With an even integer $T > 0$, define the loss functions of OCO by

$$f_t(x) = \begin{cases} (x-1)^2, & t = 1, \ldots, \frac{T}{2} \\ (x+1)^2, & t = \frac{T}{2}+1, \ldots, T. \end{cases} \tag{2}$$

Manuscript received 12 June 2023; accepted 2 December 2023. Date of publication 6 December 2023; date of current version 30 July 2024. The work of Yuhang Liu and Wenxiao Zhao was supported in part by National Key Research and Development Program of China under Grant 2018YFA0703800, in part by the Chinese Academy of Sciences (CAS) Project for Young Scientists in Basic Research under Grant YSBR-008, and in part by the Strategic Priority Research Program of Chinese Academy of Sciences under Grant XDA27000000. The work of George Yin was supported in part by the National Science Foundation under Grant DMS-2204240. Recommended by Associate Editor A. Olshevsky. *(Corresponding author: Wenxiao Zhao.)*

Yuhang Liu and Wenxiao Zhao are with the Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China, and also with the School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: liuyuhang@amss.ac.cn; wxzhao@amss.ac.cn).

George Yin is with the Department of Mathematics, University of Connecticut, Storrs, CT 06269 USA (e-mail: gyin@wayne.edu).

Digital Object Identifier 10.1109/TAC.2023.3340120

By using the follow-the-leader algorithm (see, e.g., [6]):

$$x_{t+1} = \operatorname*{argmin}_{x\in X} \sum_{s=1}^{t} f_s(x) \tag{3}$$

it follows that

$$x_t = \begin{cases} 1, & t = 1, \ldots, \frac{T}{2} \\ \frac{T}{t} - 1, & t = \frac{T}{2}+1, \ldots, T. \end{cases} \tag{4}$$

As shown in [6], for the above algorithm, the static regret function is bounded by $O(\log T)$. On the other hand, it is direct to check that the estimate x_T equals 0, rather than the minimizer -1 of $f_T(\cdot)$.

Based on regret_T^S defined in (1), some other regrets are given in the literature for evaluating performance of algorithms; see the adaptive regret regret_T^A in [27] and the dynamic regret regret_T^D in [28] defined by

$$\text{regret}_T^A = \sup_{[r,s]\subset[T]} \left\{ \sum_{t=r}^{s} f_t(x_t) - \min_{u\in X} \sum_{t=r}^{s} f_t(u) \right\} \tag{5}$$

and

$$\text{regret}_T^D = \sum_{t=1}^{T} f_t(x_t) - \sum_{t=1}^{T} f_t(x_t^*) \tag{6}$$

respectively, where $x_t^* \in \operatorname{argmin}_{x\in X} f_t(x)$, $t \geq 1$, and $[T]$ and $[r, s]$ denote the sets $\{1, \ldots, T\}$ and $\{r, \ldots, s\}$. The studies for the adaptive regret and the dynamic regret can be found in [29], [30], [9], [31], [32], [33], [34], [35], [36], [37], [38], and [39], respectively, which aim at connecting different classes of regrets and establishing the corresponding regret bounds of optimization algorithms as tight as possible. For a given algorithm $\{x_t\}_{t=1}^T$, it is readily checked that

$$\text{regret}_T^D \geq \text{regret}_T^A \geq \text{regret}_T^S. \tag{7}$$

However, for an algorithm with a sub-linear dynamic regret, there is no guarantee on $f_t(x_t) - f_t(x_t^*) \to 0$ as $t \to \infty$. For example, assuming $f_t(x_t) - f_t(x_t^*) = a_t = 1$ if $t = 2^m$, $m \geq 1$ and $f_t(x_t) - f_t(x_t^*) = a_t = 0$ otherwise, then $\lim_{T\to\infty} \sum_{t=1}^{T} a_t/T \leq \lim_{T\to\infty} \log_2 T/T = 0$, while $f_t(x_t) - f_t(x_t^*)$ does not converge to zero.

From the definitions of regret_T^S, regret_T^A, and regret_T^D, the time-varying objective functions $\{f_t(\cdot), t = 1, \ldots, T\}$ are treated equally for each time instant t. In many applications such as target tracking in systems and control, portfolio management, and property price forecast, etc., in addition to guarantee a sublinear regret, it is much desirable to ensure that at the terminal instant T, the bound of $0 \leq f_T(x_T) - f_T(x_T^*)$ is small and furthermore, $f_T(x_T) - f_T(x_T^*) \to 0$ as $T \to \infty$, which generally cannot be guaranteed by the sublinear regret_T^S, regret_T^A, and regret_T^D.

In [7], the prediction with expert advice is considered and the *discounted regrets* of the following form are introduced:

$$\text{regret}_{E,T}^{\text{dis}} = \sum_{t=1}^{T} \beta_{T-t} \left(l(p_t, y_t) - l(f_{E,t}, y_t) \right) \tag{8}$$

where $p_t \in \mathcal{D}$ with $\mathcal{D}$ being the *prediction space*, $y_t \in \mathcal{Y}$ with $\mathcal{Y}$ being the *outcome space*, $\{f_{E,t}, E \in \mathcal{E}\}$ being a set of references called *experts*, $l : \mathcal{D} \times \mathcal{Y} \to \mathbb{R}$ being the loss function, and $\{\beta_t\}_{t\geq 1}$ being a sequence of positive numbers satisfying $\beta_t \to 0$ as $t \to \infty$. Under the assumption that $\sum_{t=1}^{\infty} \beta_t < \infty$, [7, Th. 2.7] establishes a positive lower bound for the discounted regrets, i.e., there exists some $C > 0$ such that for any forecasting strategy, there is a sequence of outcomes such that

$$\max_{E\in\mathcal{E}} \{\text{regret}_{E,T}^{\text{dis}}\} \geq C \quad \forall\, T \geq 1. \tag{9}$$

From (9) it can be observed that if the $\{\beta_t\}_{t\geq 1}$ decreases too quickly, then, except for trivial cases, there is no hope to prove that the discounted regrets converge to zero. Discussions for the case $\sum_{t=1}^{\infty} \beta_t = \infty$ can be found in [7, Th. 2.8] and [51, Sec. VI].

Note that to achieve good tracking performance of $f_T(x_T) - f_T(x_T^*)$, $f_t(\cdot)$ and x_t at time t near T is more informative compared with those at earlier time instants. Motivated by this and similar to the above discounted regret, and aiming at characterizing the tracking performance of iterative algorithms for general OCO, in this article, we propose a regret for the convex loss functions, namely, regret with a forgetting factor, for which the weighting coefficients are introduced to the objective functions at each time instant t and the weights for the past objective functions are allowed to decay to zero asymptotically. To be specific, with a fixed $\rho \in (0, 1)$ the regret, namely, forgetting-factor regret denoted by regret_T^F in this article, is defined by

$$\text{regret}_T^F \triangleq \sum_{t=1}^{T} \rho^{T-t} \left(f_t(x_t) - f_t(x_t^*) \right) \tag{10}$$

where $\{x_t, t = 1, \ldots, T\}$ is generated from an algorithm for OCO with $x_t^* \in \operatorname{argmin}_{x\in X} f_t(x)$, $t = 1, \ldots, T$, and $\rho \in (0, 1)$ is the forgetting factor. Note that in regret_T^F, the effect of $\rho^{T-t}(f_t(x_t) - f_t(x_t^*))$ decays as the magnitude of $T - t$ increases.

Under the framework of the forgetting-factor regret and for an optimization algorithm of OCO, in this article we

1) analyze the upper bound of regret_T^F, and
2) introduce sufficient conditions such that $\lim_{T\to\infty} \text{regret}_T^F = 0$.

This establishes the tracking performance of the algorithm since $f_T(x_T) - f_T(x_T^*) \leq \text{regret}_T^F$. For a given algorithm $\{x_t\}_{t=1}^T$, by (7) and (10) it is easy to check that if $\text{regret}_T^F = o(1)$ as $T \to \infty$, then $\text{regret}_T^D = o(T)$, $\text{regret}_T^A \leq o(T)$, and $\text{regret}_T^S \leq o(T)$.

The contributions of the article are as follows.

1) We are able to analyze the performance of the classical algorithms for OCO under regret_T^F.
   We first analyze regret_T^F of the online gradient descent algorithm (OGDA). Then, for the bandit information model where the gradients of the objective functions are unavailable, we establish bounds of regret_T^F for the online gradient-free algorithms (OGFA) with δ-smoothing and deterministic difference techniques.

By introducing a line-search procedure into the online Frank–Wolfe algorithm (OFWA), we investigate regret$_T^F$ of the corresponding algorithm, which reduces the computation complexity compared with OGDA and OGFA.

2) Under the framework of regret$_T^F$, we introduce a new class of algorithm–online gradient descent algorithm with a forgetting factor (OGDA-F) and analyze the corresponding regret$_T^F$.
3) We establish bounds of regret$_T^F$'s for OGDA, OGFA, OFWA, as well as OGDA-F and derive sufficient conditions to ensure regret$_T^F$'s of these algorithms tending to 0 as $T \to \infty$. Our knowledge was not known in literature in the past. In addition, the regret bound of OGDA-F extends the well-established estimation error bound for time-varying linear stochastic systems; see, e.g., [40] for the nonlinear OCO problems.

The rest of the article is organized as follows. In Section II, we give the detailed problem formulation of OCO and the definition of regret$_T^F$, analyze the performance of OGDA, OGFA, and OFWA, and propose a new algorithm OGDA-F together with detailed analysis. In Section III, we provide some simulation examples to demonstrate our algorithms. Then, we make a number of concluding remarks in Section IV. Finally, in Appendix, we list some definitions and results for convex optimization.

Notation: Denote by $(\Omega, \mathcal{F}, \mathbb{P})$ the probability space and $\mathbb{E}(\cdot)$ the mathematical expectation operator. Denote by $M[i,j]$ the (i,j)th entry of the matrix M and by $M[i,:]$ and $M[:,j]$ its ith row and jth column, respectively. Denote by e^k the vector in $\mathbb{R}^d$ with the kth component being 1 and the others being 0. Denote by $\|\cdot\|$ the Euclidean norm on $\mathbb{R}^d$ and by $\langle\cdot,\cdot\rangle$ the inner product on $\mathbb{R}^d$. The projection operator $\mathcal{P}_X(\cdot)$ onto the set X is defined by $\mathcal{P}_X(y) \triangleq \arg\min_{x\in X}\|x-y\|$. The gradient of the function $f(\cdot):\mathbb{R}^d \to \mathbb{R}$ at a given x is denoted by $\nabla f(x)$ if it exists. For a set X, denote its interior by X^o.

II. Performance Analysis of Algorithms Under the Forgetting Factor Regret

Definition 1: Let $\{f_t(\cdot), t=1,\ldots,T\}$ be a sequence of convex objective functions over the constraint set $X \subset \mathbb{R}^d$. Let $\mathcal{A}$ be an algorithm for the above OCO problem. Denote by $\{x_t\}_{1\le t\le T}$ the estimates generated by $\mathcal{A}$. With a given $\rho \in (0,1)$, the forgetting-factor regret of algorithm $\mathcal{A}$ is defined by

$$\text{regret}_T^F = \sum_{t=1}^{T} \rho^{T-t}\left(f_t(x_t) - f_t(x_t^*)\right) \tag{11}$$

where $x_t^* \in \arg\min_{x\in X} f_t(x)$, $t=1,\ldots,T$.

With the definition, an algorithm for OCO possesses good tracking performance provided that $\lim_{T\to\infty}\text{regret}_T^F = 0$. We first impose a set of conditions for the analysis to follow:

1) The feasible set $X \subset \mathbb{R}^d$ is compact and convex.
2) For any $t \ge 0$, the objective function $f_t(\cdot)$ is convex and differentiable in X and there exists $L > 0$ such that $\|\nabla f_t(x)\| \le L$, $x \in X$, $t = 1,\ldots,T$.
3) $\{f_t(\cdot), 0 \le t \le T\}$ are L_S-smooth in X.

The definition of L_S-smooth functions is given in Appendix. By A1), we can define a positive constant $M \triangleq \sup\{\|x\| : x \in X\} < \infty$. As indicated in [32] and [34], it is impossible to exactly track the optimizer defined by an arbitrarily varying optimization problem. Thus, we introduce the following notation to characterize the loss functions:

$$\begin{cases} \theta_t \triangleq \|x_{t+1}^* - x_t^*\| \\ F_{t,t+1}^{\sup} \triangleq \sup_{x\in X}|f_t(x) - f_{t+1}(x)| \end{cases} \tag{12}$$

which will be used throughout the article.

A. Forgetting-Factor Regret of OGDA

The OGAD algorithm is one of the most widely applied algorithms for online optimization, which incorporated with a projection operator can be formulated as follows.

Algorithm 1: Online Gradient Descent Algorithm (OGDA).

Initialization: An initial estimate x_0, a step size $\alpha > 0$ and the maximal number T of iterations.
For $t = 0,\ldots,T$,
update the estimate as

$$\widehat{x}_{t+1} = x_t - \alpha\nabla f_t(x_t) \tag{13}$$

$$x_{t+1} = \mathcal{P}_X\left(\widehat{x}_{t+1}\right) \tag{14}$$

end;

Theorem 1: Let $\{x_t\}_{t\ge1}$ be a sequence generated by Algorithm 1. Assume A1)–A3) hold and for each $t = 1,\ldots,T$, the objective function $f_t(\cdot)$ is σ-strongly convex. Then for any $\rho \in (\max\{1/2, 1-\sigma/(4L_S)\}, 1)$,

$$\text{regret}_T^F \le \frac{1}{\beta}\left\{\rho^T\left(f_1(x_1) - f_1(x_1^*)\right) + 2\sum_{t=1}^{T}\rho^{T-t}F_{t,t+1}^{\sup} + L\sum_{t=1}^{T}\rho^{T-t}\theta_t\right\} \tag{15}$$

provided that $\frac{4(1-\rho)}{\sigma} < \alpha \le \min\{1/L_S, 2/\sigma\}$, where $\beta = \rho + \frac{\sigma\alpha}{4} - 1 > 0$. Moreover, if $F_{t,t+1}^{\sup} = o(1)$ and $\theta_t = o(1)$ as $t \to \infty$, then $\lim_{T\to\infty}\text{regret}_T^F = 0$.

Theorem 1 further requires that $F_{t,t+1}^{\sup} = o(1)$ and $\theta_t = o(1)$ as $t \to \infty$. It means that the varying minimizers of the OCO problem should not change too fast with respect to the time instants $t \ge 1$. Let us consider the following illustrative example. Choose $f_t(x) \triangleq \frac{1}{2}\|x - z_t\|^2$, $t \ge 1$ with a compact and convex feasible set X and $z_t \in X$, $t \ge 1$. Assume that the changing of the minimizers is slow, i.e., $\|z_t - z_{t+1}\| = o(1)$ as $t \to \infty$. Set $M \triangleq \sup\{\|x\| : x \in X\} < \infty$. From the definitions of $F_{t,t+1}^{\sup}$ and θ_t, it is direct to check that $F_{t,t+1}^{\sup} \le 2M\|z_t - z_{t+1}\| = 2M\theta_t$ and $F_{t,t+1}^{\sup} = o(1)$ and $\theta_t = o(1)$ follows.

Theorem 1 establishes the upper bound of the forgetting-factor regret of OGDA and gives the sufficient conditions to guarantee regret$_T^F = o(1)$ as $T \to \infty$. From Theorem 1, for OGDA with the appropriately selected forgetting factor ρ and the step size

constant α, the tracking of the time-varying minimizers of OCO can be guaranteed, i.e., $f_T(x_T) - f_T(x_T^*) \to 0$ as $T \to \infty$. A key step toward establishing the upper bound of regret$_T^F$ is to find a recursive formula for $f_t(x_t) - f_t(x_t^*)$, $t \geq 0$ by using the properties of the projection operator and the convexity of the loss functions. The details are given as follows.

Proof: By the definition of $F_{t,t+1}^{\sup}$, it directly follows that

$$\begin{aligned} f_{t+1}(x_{t+1}) - f_{t+1}(x_{t+1}^*) &= f_{t+1}(x_{t+1}) - f_t(x_{t+1}) + f_t(x_{t+1}) \\ &\quad - f_t(x_t) + f_t(x_t) - f_t(x_t^*) + f_t(x_t^*) - f_{t+1}(x_{t+1}^*) \\ &\leq F_{t,t+1}^{\sup} + f_t(x_{t+1}) - f_t(x_t) + f_t(x_t) - f_t(x_t^*) \\ &\quad + f_t(x_t^*) - f_{t+1}(x_{t+1}^*). \end{aligned} \tag{16}$$

For $f_t(x_t^*) - f_{t+1}(x_{t+1}^*)$, the convexity of $f_{t+1}(\cdot)$ yields that

$$\begin{aligned} &f_t(x_t^*) - f_{t+1}(x_{t+1}^*) \\ &= f_t(x_t^*) - f_{t+1}(x_t^*) + f_{t+1}(x_t^*) - f_{t+1}(x_{t+1}^*) \\ &\leq F_{t,t+1}^{\sup} + \left\langle \nabla f_{t+1}(x_t^*), x_t^* - x_{t+1}^* \right\rangle \leq F_{t,t+1}^{\sup} + L\theta_t \end{aligned} \tag{17}$$

where the last inequality follows from the Schwartz inequality and Assumption A2).

Next we consider $f_t(x_{t+1}) - f_t(x_t)$. By A3),

$$\begin{aligned} &f_t(x_{t+1}) - f_t(x_t) \\ &\leq \langle \nabla f_t(x_t), x_{t+1} - x_t \rangle + \frac{L_S}{2} \|x_{t+1} - x_t\|^2 \end{aligned} \tag{18}$$

which together with the stepsize condition $\alpha \leq 1/L_S$ yields

$$\begin{aligned} &f_t(x_{t+1}) - f_t(x_t) \\ &\leq \langle \nabla f_t(x_t), x_{t+1} - x_t \rangle + \frac{1}{2\alpha} \|x_{t+1} - x_t\|^2 . \end{aligned} \tag{19}$$

On the other hand, from (13)–(14), Algorithm 1 can be rewritten as

$$\begin{aligned} x_{t+1} &= \arg\min_{x \in X} \|x - x_t + \alpha \nabla f_t(x_t)\|^2 \\ &= \arg\min_{x \in X} \left\{ \langle \nabla f_t(x_t), x - x_t \rangle + \frac{1}{2\alpha} \|x - x_t\|^2 \right\}. \end{aligned} \tag{20}$$

Combining (19) and (20) and by using the convexity of $f_t(\cdot)$, we have

$$\begin{aligned} f_t(x_{t+1}) - f_t(x_t) &\leq \min_{x \in X} \left\{ \langle \nabla f_t(x_t), x - x_t \rangle + \frac{1}{2\alpha} \|x - x_t\|^2 \right\} \\ &\leq \min_{x \in X} \left\{ f_t(x) - f_t(x_t) + \frac{1}{2\alpha} \|x - x_t\|^2 \right\}. \end{aligned} \tag{21}$$

Note that X is a convex set and $f_t(\cdot)$ is convex in X. For any fixed $\eta \in [0,1]$, by setting $x = (1-\eta)x_t + \eta x_t^*$ in (21), we have

$$\begin{aligned} &f_t(x_{t+1}) - f_t(x_t) \\ &\leq (1-\eta) f_t(x_t) + \eta f_t(x_t^*) - f_t(x_t) + \frac{\eta^2}{2\alpha} \|x_t - x_t^*\|^2 \\ &= -\eta \left(f_t(x_t) - f_t(x_t^*) \right) + \frac{\eta^2}{2\alpha} \|x_t - x_t^*\|^2 . \end{aligned} \tag{22}$$

By using the σ-strong convexity of $f_t(\cdot)$ and Lemma A1, we obtain

$$\frac{\sigma}{2} \|x_t - x_t^*\|^2 \leq f_t(x_t) - f_t(x_t^*). \tag{23}$$

Set $\eta = \frac{\sigma\alpha}{2}$. By noting $\alpha \leq \frac{2}{\sigma}$, it is direct to check that $\eta \leq 1$. Then from (22) and (23), we obtain

$$\begin{aligned} f_t(x_{t+1}) - f_t(x_t) &\leq \left(\frac{\eta^2}{\sigma\alpha} - \eta \right) \left(f_t(x_t) - f_t(x_t^*) \right) \\ &= -\frac{\sigma\alpha}{4} \left(f_t(x_t) - f_t(x_t^*) \right) \end{aligned} \tag{24}$$

which incorporated with (16) and (17) gives

$$\begin{aligned} &f_{t+1}(x_{t+1}) - f_{t+1}(x_{t+1}^*) \\ &\leq \left(1 - \frac{\sigma\alpha}{4}\right) \left(f_t(x_t) - f_t(x_t^*) \right) + 2F_{t,t+1}^{\sup} + L\theta_t \\ &= (\rho - \beta) \left(f_t(x_t) - f_t(x_t^*) \right) + 2F_{t,t+1}^{\sup} + L\theta_t \end{aligned} \tag{25}$$

where $\beta = \rho + \frac{\sigma\alpha}{4} - 1$ being positive since $\frac{4(1-\rho)}{\sigma} < \alpha \leq \min\{1/L_S, 2/\sigma\}$.

By multiplying ρ^{T-t} to both sides of (25) and then summing up the terms for $t = 1, \ldots, T$, we finally obtain

$$\begin{aligned} \beta \sum_{t=1}^{T} \rho^{T-t} \left(f_t(x_t) - f_t(x_t^*) \right) &\leq \sum_{t=1}^{T} \left(\rho^{T-t+1} \left(f_t(x_t) - f_t(x_t^*) \right) \right. \\ &\quad \left. - \rho^{T-t} \left(f_{t+1}(x_{t+1}) - f_{t+1}(x_{t+1}^*) \right) \right) + 2 \sum_{t=1}^{T} \rho^{T-t} \left(F_{t,t+1}^{\sup} + L\theta_t \right) \\ &\leq \rho^T \left(f_1(x_1) - f_1(x_1^*) \right) + 2 \sum_{t=1}^{T} \rho^{T-t} F_{t,t+1}^{\sup} + L \sum_{t=1}^{T} \rho^{T-t} \theta_t. \end{aligned} \tag{26}$$

Hence, (15) is proved.

If $F_{t,t+1}^{\sup} = o(1)$ and $\theta_t = o(1)$ as $t \to \infty$, then $\lim_{T\to\infty} \text{regret}_T^F = 0$ follows directly from the fact that for any positive sequence $\{\gamma_t\}_{t \geq 0}$ tending to zero, $\lim_{T\to\infty} \sum_{t=1}^{T} \rho^{T-t} \gamma_t = 0$; (see, e.g., [49, Lemma 3.1]). ■

B. Forgetting-Factor Regret of Online Gradient-Free Algorithms

In applications, to obtain the gradient/subgradient information is sometimes computationally expensive, and even impracticable in some cases, such as online source localization, online routing of data networks, online placement of advertisement [10], as well as the graphical model inference [41], and the structured-prediction [42] in statistics, etc. For OCO, this is called the partial information model or the bandit model and in such a scenario, at time instant t, the value of $f_t(\cdot)$ at x_t can be observed but the gradient $\nabla f_t(x_t)$ is unavailable.

The so-called δ-smoothing technique is widely applied for designing gradient-free algorithms for OCO [13]. In fact, such an approach is rooted to a variation of the random direction methods in stochastic approximation [43]; see also some recent progress in stochastic approximation [44] and stochastic convex

optimization [45]. In this section, we first analyze regret_T^F of OGFA with the δ-smoothing technique. Then we introduce a deterministic difference based OGFA algorithm and analyze the corresponding regret_T^F.

Denote by $\mathbb{B}$ the d-dimensional unit ball and by $\mathbb{S}$ the d-dimensional unit sphere.

Definition 2 (see [13]): At time instant t, for the loss function $f_t(x)$ with $x \in \mathbb{R}^d$, its δ-smoothing function $\widehat{f}_{t,\delta}(x)$ is given by

$$\widehat{f}_{t,\delta}(x) \triangleq \mathbb{E}[f_t(x+\delta v)] \tag{27}$$

where $\delta > 0$ is a scalar constant and v is a random vector uniformly distributed over the d-dimensional unit ball $\mathbb{B}$.

Before proceeding further, we strengthen Assumption A1) as follows.

A1') The constraint set $X \subset \mathbb{R}^d$ is compact and convex and satisfies $r\mathbb{B} \subset X \subset R\mathbb{B}$ for some $R > r > 0$. Moreover, there exists an constant $L_1 > 0$ such that $\sup_{1\le t\le T, x\in X} |f_t(x)| \le L_1$.

Choose $\delta \in (0, r)$ and set $X_\delta \triangleq (1-\delta/r)X$. It can be proved that (see, e.g., [13]),

$$X_\delta + \delta\mathbb{B} \subset X \tag{28}$$

and

$$\nabla \widehat{f}_{t,\delta}(x) = \frac{d}{\delta}\mathbb{E}[f_t(x+\delta u)u] \tag{29}$$

where u is a random vector uniformly distributed over the d-dimensional unit sphere $\mathbb{S}$. OGFA with the δ-smoothing technique is given as follows:

Algorithm 2: Online Gradient-Free Algorithm (OGFA) with δ-Smoothing Technique.

Initialization: An initial estimate x_0, a step size α, a smoothing parameter δ, and the maximal number T of iterations.

For $t = 0, \ldots, T$**,**

choose u_t a random vector uniformly distributed over $\mathbb{S}$ and independent of $\{u_1, \ldots, u_{t-1}\}$,

update the estimate as

$$g_t = \frac{d}{\delta}\left(f_t(x_t+\delta u_t) - f_t(x_t)\right)u_t \tag{30}$$

$$x_{t+1} = P_{X_\delta}\left(x_t - \alpha g_t\right) \tag{31}$$

end;

Theorem 2: Let $\{x_t\}_{t\ge 1}$ be a sequence generated by Algorithm 2. Suppose that A1'), A2), and A3) hold and $f_t(\cdot), t = 1, \ldots, T$ are σ-strongly convex. Then for any $\rho \in (\max\{1/2, 1-\sigma/(4L_S)\}, 1)$,

$$\begin{aligned}\mathbb{E}[\text{regret}_T^F] &= \sum_{t=1}^{T}\rho^{T-t}\mathbb{E}\left[f_t(x_t) - f_t(x_t^*)\right] \\ &\le \frac{1}{\beta}\left\{\rho^T\mathbb{E}(f_1(x_1) - f_1(x_1^*)) + 2\sum_{t=1}^{T}\rho^{T-t}F_{t,t+1}^{\sup}\right. \\ &\quad \left. + L\sum_{t=1}^{T}\rho^{T-t}\theta_t + \frac{C_{\delta,\alpha}}{1-\rho}\right\}\end{aligned} \tag{32}$$

provided that $\frac{4(1-\rho)}{\sigma} < \alpha \le \min\{1/L_S, 2/\sigma\}$, where $\beta = \rho + \frac{\sigma\alpha}{4} - 1 > 0$ and $C_{\delta,\alpha} = (2L + \frac{\sigma L}{2}\alpha + \frac{\sigma L_1}{4r}\alpha)\delta + 2d^2L^2\alpha$. Moreover, if $F_{t,t+1}^{\sup} = o(1)$ and $\theta_t = o(1)$ as $t \to \infty$, then $\limsup_{T\to\infty}\mathbb{E}[\text{regret}_T^F] \le \frac{C_{\delta,\alpha}}{\beta(1-\rho)}$.

Theorem 2 establishes the upper bound of the forgetting-factor regret of Algorithm 2. Note that for OGFA with δ-smoothing technique, the upper bound of $\mathbb{E}[\text{regret}_T^F]$ is related by a positive constant $C_{\delta,\alpha}$. This is because the δ-smoothing based difference (30) serves as an estimate for $\nabla\widehat{f}_{t,\delta}(x)$ given by (29), not for the gradient of the optimization function $\nabla f_t(x)$ itself. The proof of Theorem 2 generally follows the similar lines of Theorem 1, by paying attention to the difference between the δ-smoothing function $\widehat{f}_{t,\delta}(x)$ and the optimization function $f_t(x)$, i.e., $\widehat{f}_{t,\delta}(x) - f_t(x)$.

Proof: A similar analysis as (16), (17) leads to

$$\begin{aligned}&f_{t+1}(x_{t+1}) - f_{t+1}(x_{t+1}^*) \\ &\le 2F_{t,t+1}^{\sup} + f_t(x_{t+1}) - f_t(x_t) + f_t(x_t) - f_t(x_t^*) + L\theta_t.\end{aligned} \tag{33}$$

By (27), (28), and Assumption A2), for any $x \in X_\delta$ we have the following estimate:

$$\begin{aligned}\left|\widehat{f}_{t,\delta}(x) - f_t(x)\right| &= \left|\mathbb{E}[f_t(x+\delta v)] - f_t(x)\right| \\ &\le \mathbb{E}\left|f_t(x+\delta v) - f_t(x)\right| \\ &\le \mathbb{E}[L\delta\|v\|] \le L\delta\end{aligned} \tag{34}$$

where the last inequality holds because v is uniformly distributed over the unit ball $\mathbb{B}$ and thus

$$f_t(x_{t+1}) - f_t(x_t) \le 2L\delta + \widehat{f}_{t,\delta}(x_{t+1}) - \widehat{f}_{t,\delta}(x_t). \tag{35}$$

Next, we consider the upper bound of $\widehat{f}_{t,\delta}(x_{t+1}) - \widehat{f}_{t,\delta}(x_t)$.

Noting A3) and (28), by the assumption that X is compact and the Lebesgue's dominated convergence theorem, it can be proven that $\nabla\widehat{f}_{t,\delta}(x) = \mathbb{E}[\nabla f_t(x+\delta v)]$ and $\widehat{f}_{t,\delta}(x)$ is also L_S-smooth over X_δ, i.e.,

$$\begin{aligned}&\widehat{f}_{t,\delta}(x_{t+1}) - \widehat{f}_{t,\delta}(x_t) \\ &\le \left\langle \nabla\widehat{f}_{t,\delta}(x_t), x_{t+1} - x_t\right\rangle + \frac{L_S}{2}\|x_{t+1} - x_t\|^2\end{aligned} \tag{36}$$

since by (31), $x_t \in X_\delta$. From (36) and noting $\alpha \le 1/L_S$ and (30),

$$\begin{aligned}&\widehat{f}_{t,\delta}(x_{t+1}) - \widehat{f}_{t,\delta}(x_t) \\ &\le \left\langle \nabla\widehat{f}_{t,\delta}(x_t), x_{t+1} - x_t\right\rangle + \frac{1}{2\alpha}\|x_{t+1} - x_t\|^2 \\ &= \langle g_t, x_{t+1} - x_t\rangle + \frac{1}{2\alpha}\|x_{t+1} - x_t\|^2 \\ &\quad + \left\langle \nabla\widehat{f}_{t,\delta}(x_t) - g_t, x_{t+1} - x_t\right\rangle.\end{aligned} \tag{37}$$

By (29) and (30) and noting A1'), A2), we get $\|g_t\| \le \frac{d}{\delta}\cdot L\|\delta u_t\| = dL$ and $\|\nabla\widehat{f}_{t,\delta}(x)\| = \|\frac{d}{\delta}\mathbb{E}[(f_t(x+\delta u))u]\| =$

$\|\frac{d}{\delta}\mathbb{E}[(f_t(x+\delta u)-f_t(x))u]\| \le dL$ for any $x \in X_\delta$. Then by Lemma A2 in Appendix and (31) it leads to

$$\begin{aligned}&\left\langle \nabla \widehat{f}_{t,\delta}(x_t) - g_t, x_{t+1}-x_t\right\rangle\\ &\le \left(\|\nabla \widehat{f}_{t,\delta}(x_t)\| + \|g_t\|\right)\|x_{t+1}-x_t\|\\ &\le 2dL\cdot(\alpha\|g_t\|) \le 2d^2L^2\alpha \end{aligned} \tag{38}$$

which combining with (37) yields

$$\begin{aligned}&\widehat{f}_{t,\delta}(x_{t+1}) - \widehat{f}_{t,\delta}(x_t)\\ &\le \langle g_t, x_{t+1}-x_t\rangle + \frac{1}{2\alpha}\|x_{t+1}-x_t\|^2 + 2d^2L^2\alpha. \end{aligned} \tag{39}$$

On the other hand, noting that $P_{X_\delta}(\cdot)$ is a projection operator, for Algorithm 2 we have

$$x_{t+1} = \arg\min_{x\in X_\delta}\left\{\langle g_t, x-x_t\rangle + \frac{1}{2\alpha}\|x-x_t\|^2\right\}. \tag{40}$$

Combining (39) and (40), we obtain

$$\begin{aligned}&\widehat{f}_{t,\delta}(x_{t+1}) - \widehat{f}_{t,\delta}(x_t)\\ &\le \min_{x\in X_\delta}\left\{\langle g_t, x-x_t\rangle + \frac{1}{2\alpha}\|x-x_t\|^2\right\} + 2d^2L^2\alpha\\ &= \min_{x\in X_\delta}\left\{\langle \nabla \widehat{f}_{t,\delta}(x_t), x-x_t\rangle + \frac{1}{2\alpha}\|x-x_t\|^2\right.\\ &\quad \left. + \langle g_t - \nabla \widehat{f}_{t,\delta}(x_t), x-x_t\rangle\right\} + 2d^2L^2\alpha\\ &\le \min_{x\in X_\delta}\left\{\widehat{f}_{t,\delta}(x) - \widehat{f}_{t,\delta}(x_t) + \frac{1}{2\alpha}\|x-x_t\|^2\right.\\ &\quad \left. + \langle g_t - \nabla \widehat{f}_{t,\delta}(x_t), x-x_t\rangle\right\} + 2d^2L^2\alpha \end{aligned} \tag{41}$$

where the last inequality holds because of the convexity of $\widehat{f}_{t,\delta}(\cdot)$ over X_δ.

Denote $\tilde{x}_t^* = \arg\min_{x\in X_\delta} f_t(x)$. By setting $x = (1-\frac{\sigma\alpha}{2})x_t + \frac{\sigma\alpha}{2}\tilde{x}_t^* \in X_\delta$ in (41) and carrying out a similar analysis as (22)–(24), we have

$$\begin{aligned}&\widehat{f}_{t,\delta}(x_{t+1}) - \widehat{f}_{t,\delta}(x_t) \le -\frac{\sigma\alpha}{4}\left(\widehat{f}_{t,\delta}(x_t) - \widehat{f}_{t,\delta}(\tilde{x}_t^*)\right)\\ &\quad + \frac{\sigma\alpha}{2}\langle g_t - \nabla \widehat{f}_{t,\delta}(x_t), \tilde{x}_t^* - x_t\rangle + 2d^2L^2\alpha. \end{aligned} \tag{42}$$

Note that for the conditional expectation $\mathbb{E}[g_t|x_t] = \nabla \widehat{f}_{t,\delta}(x_t)$ and thus

$$\begin{aligned}&\mathbb{E}\left\langle g_t - \nabla \widehat{f}_{t,\delta}(x_t), \tilde{x}_t^* - x_t\right\rangle\\ &= \mathbb{E}\left\langle \mathbb{E}\left[g_t - \nabla \widehat{f}_{t,\delta}(x_t)|x_t\right], \tilde{x}_t^* - x_t\right\rangle = 0 \end{aligned} \tag{43}$$

from which and by taking the mathematical expectation to both sides of (42), we have

$$\begin{aligned}&\mathbb{E}\left(\widehat{f}_{t,\delta}(x_{t+1}) - \widehat{f}_{t,\delta}(x_t)\right)\\ &\le -\frac{\sigma\alpha}{4}\mathbb{E}\left(\widehat{f}_{t,\delta}(x_t) - \widehat{f}_{t,\delta}(\tilde{x}_t^*)\right) + 2d^2L^2\alpha. \end{aligned} \tag{44}$$

Next we consider the term $\widehat{f}_{t,\delta}(x_t) - \widehat{f}_{t,\delta}(\tilde{x}_t^*)$. By A1'), we know that $0 \in X$. Then, by the convexity of $f_t(\cdot)$ we have

$$\begin{aligned}f_t(\tilde{x}_t^*) &= \min_{x\in X_\delta} f_t(x) = \min_{x\in X} f_t\left((1-\delta/r)\,x\right)\\ &\le \min_{x\in X}\{(\delta/r f_t(0) + (1-\delta/r)\,f_t(x)\} \le \frac{\delta}{r}L_1 + f_t(x_t^*). \end{aligned} \tag{45}$$

Using $|f_t(x) - \widehat{f}_{t,\delta}(x)| \le L\delta$ for all $x \in X_\delta$, we obtain

$$\begin{aligned}&\widehat{f}_{t,\delta}(x_t) - \widehat{f}_{t,\delta}(\tilde{x}_t^*)\\ &= \widehat{f}_{t,\delta}(x_t) - f_t(x_t) + f_t(\tilde{x}_t^*) - \widehat{f}_{t,\delta}(\tilde{x}_t^*) + f_t(x_t) - f_t(\tilde{x}_t^*)\\ &\ge -2L\delta + f_t(x_t) - f_t(\tilde{x}_t^*) \ge -2L\delta + f_t(x_t) - f_t(x_t^*) - \frac{\delta}{r}L_1 \end{aligned} \tag{46}$$

where for the last inequality (45) is applied.

Taking the mathematical expectation on both sides of (46) and then substituting it into (44) yield

$$\begin{aligned}&\mathbb{E}\left(\widehat{f}_{t,\delta}(x_{t+1}) - \widehat{f}_{t,\delta}(x_t)\right)\\ &\le -\frac{\sigma\alpha}{4}\mathbb{E}(f_t(x_t) - f_t(x_t^*)) + \frac{\sigma\alpha}{4}\left(2L\delta + \frac{\delta}{r}L_1\right) + 2d^2L^2\alpha. \end{aligned} \tag{47}$$

Combining (33), (35), and (47), we obtain

$$\begin{aligned}&\mathbb{E}\left(f_{t+1}(x_{t+1}) - f_{t+1}(x_{t+1}^*)\right)\\ &\le \left(1-\frac{\sigma\alpha}{4}\right)\mathbb{E}\left(f_t(x_t) - f_t(x_t^*)\right) + 2F_{t,t+1}^{\sup} + L\theta_t\\ &\quad + \left(2L\delta + \frac{\sigma\alpha}{4}\left(2L\delta + \frac{\delta}{r}L_1\right) + 2d^2L^2\alpha\right) \end{aligned} \tag{48}$$

from which together with a similar discussion as (25), (26), we can prove (32). This completes the proof. ■

Motivated by the Kiefer–Wolfowitz algorithm for stochastic optimization [43], [46], in what follows, we introduce another class of gradient-free algorithm, which, to the authors' knowledge, has not been discussed in the literature. Let $\{c_t\}_{t\ge 0}$ be a sequence of positive numbers. At time t, assume that the values of $f_t(\cdot)$ at the points $x = x_t + c_t e^k$ and $x = x_t - c_t e^k$ with $k = 1, \ldots, d$ can be observed, where e^k is the vector in $\mathbb{R}^d$ with the kth component being 1 and the others being 0. Denote the observed function values by $[z_t^k]^+ = f_t(x_t + c_t e^k)$ and $[z_t^k]^- = f_t(x_t - c_t e^k)$ and define the differences

$$h_t[k] = \frac{[z_t^k]^+ - [z_t^k]^-}{2c_t}, \quad k = 1, \ldots, d \tag{49}$$

and

$$h_t = [h_t[1], \ldots, h_t[d]]^\top \tag{50}$$

which serve as estimates for the gradients of objective functions.

Algorithm 3: Online Gradient-Free Algorithm (OGFA) with Deterministic Difference Technique.

Initialization: An initial estimate x_0, a step size α, the maximal number T of iterations, and a positive decreasing sequence $\{c_t\}_{t=1}^{T}$.

For $t = 0, \ldots, T$,

update the estimate as

$$h_t = \left[\frac{[z_t^1]^+ - [z_t^1]^-}{2c_t}, \ldots, \frac{[z_t^d]^+ - [z_t^d]^-}{2c_t}\right]^T \tag{51}$$

$$\widehat{x}_{t+1} = x_t - \alpha h_t, \tag{52}$$

$$x_{t+1} = \mathcal{P}_X\left(\widehat{x}_{t+1}\right), \tag{53}$$

end;

Before proceeding further, we need the following assumptions.

A4) $\{f_t(\cdot), t = 1, \ldots, T\}$ are second-order differentiable and there exists a positive constant L_H such that

$$\left|\nabla^2 f_t[k,k]\right| \le L_H, \quad t = 1, \ldots, T, \quad k = 1, \ldots, d \tag{54}$$

where $\nabla^2 f_t(\cdot)$ represents the Hessian matrix of $f_t(\cdot)$.

A5) $\{c_t\}_{t\ge 0}$ is a positive sequence decreasing to zero.

Theorem 3: Let $\{x_t\}_{t\ge 1}$ be the sequence generated by (51)–(53) with initial value $x_0 \in X$. Suppose that $f_t(\cdot)$, $t = 1, \ldots, T$ are σ-strongly convex and A1)–A5) are satisfied. Further, assume that $\alpha \in (\frac{4(1-\rho)}{\sigma}, \min\{1/L_S, 2/\sigma\}]$ and $1 > \rho > \max\{1/2, 1 - \sigma/(4L_S)\}$. Then,

$$\begin{aligned}\text{regret}_T^F &= \sum_{t=1}^{T} \rho^{T-t}\left(f_t(x_t) - f_t(x_t^*)\right)\\ &\le \frac{1}{\beta}\left\{\rho^T\left(f_1(x_1) - f_1(x_1^*)\right) + 2\sum_{t=1}^{T}\rho^{T-t}F_{t,t+1}^{\sup}\right.\\ &\quad \left. + L\sum_{t=1}^{T}\rho^{T-t}\theta_t + 4M\sqrt{d}L_H\sum_{t=1}^{T}\rho^{T-t}c_t\right\}\end{aligned} \tag{55}$$

where $\beta = \rho + \frac{\sigma\alpha}{4} - 1 > 0$. Moreover, if $c_t = o(1)$, $F_{t,t+1}^{\sup} = o(1)$, and $\theta_t = o(1)$ as $t \to \infty$, then $\lim_{T\to\infty} \text{regret}_T^F = 0$.

Remark 1: Theorem 3 establishes the upper bound of the forgetting-factor regret of Algorithm 3. Here we make some comparisons between Algorithms 2 and 3. Noting (29) and (30), $g_t = \frac{d}{\delta}(f_t(x_t + \delta u_t) - f_t(x_t))u_t$ applied to Algorithm 2 serves as an unbiased estimate for $\nabla \widehat{f}_{t,\delta}(x)$, while the deterministic difference h_t applied in (51) serves as an estimate for $\nabla f_t(x)$ itself. From Theorem 2 we find that if $F_{t,t+1}^{\sup} = o(1)$ and $\theta_t = o(1)$, for OGFA with the δ-smoothing technique, $\limsup_{T\to\infty} \mathbb{E}[\text{regret}_T^F] = O(\delta) + O(\alpha)$, which depends on the choice of the smoothing parameter δ and the step size α. On the other hand, for OGFA with the deterministic difference technique, Theorem 3 ensures $\limsup_{T\to\infty} \text{regret}_T^F = 0$ provided that $F_{t,t+1}^{\sup} = o(1)$ and $\theta_t = o(1)$ as $t \to \infty$. Hence, Algorithm 3 guarantees better performance than Algorithm 2 in regard of the tracking error of $f_T(x_T) - f_T(x_T^*)$. The proof of Theorem 3 generally follows the similar lines of Theorem 1, by properly analyzing the deterministic difference h_t given by (51).

Proof: Since the proof is similar to that of Theorem 1, here we only give a sketch.

By the definitions in (49) and (50), for any $t \ge 0$ and $k \in \{1, \ldots, d\}$, for h_t it follows that

$$\begin{aligned}(e^k)^\top h_t = h_t[k] &= \frac{[z_t^k]^+ - [z_t^k]^-}{2c_t}\\ &= \frac{f_t(x_t + c_t e^k) - f_t(x_t - c_t e^k)}{2c_t}.\end{aligned} \tag{56}$$

By Taylor's expansion and noting that $f_t(\cdot)$ is second-order differentiable, we obtain that

$$h_t = \nabla f_t(x_t) + c_t p_t \tag{57}$$

where

$$\begin{aligned}p_t[k] = \langle e^k, p_t\rangle = \frac{1}{2}(e^k)^\top &\left[\nabla^2 f_t\left(x_t + [\beta_t^k]^+[\theta_t^k]^+ c_t e^k\right)[\theta_t^k]^+\right.\\ &\left. + \nabla^2 f_t\left(x_t + [\beta_t^k]^-[\theta_t^k]^- c_t e^k\right)[\theta_t^k]^-\right] e^k\end{aligned} \tag{58}$$

with $[\theta_t^k]^+, [\theta_t^k]^-, [\beta_t^k]^+, [\beta_t^k]^- \in [-1, 1]$ for $k \in \{1, \ldots, d\}$.

Formula (52) and (53) are equivalent to

$$x_{t+1} = \arg\min_{x\in X}\left\{\langle h_t, x - x_t\rangle + \frac{1}{2\alpha}\|x - x_t\|^2\right\}. \tag{59}$$

Similar to (18)–(19) and noting (57), we obtain that

$$\begin{aligned}f_t(x_{t+1}) &\le f_t(x_t) + \langle \nabla f_t(x_t), x_{t+1} - x_t\rangle + \frac{1}{2\alpha}\|x_{t+1} - x_t\|^2\\ &= f_t(x_t) + \langle h_t, x_{t+1} - x_t\rangle + \frac{1}{2\alpha}\|x_{t+1} - x_t\|^2\\ &\quad - \langle c_t p_t, x_{t+1} - x_t\rangle.\end{aligned} \tag{60}$$

Combining (59) and (60) leads to

$$\begin{aligned}f_t(x_{t+1}) - f_t(x_t) \le \min_{x\in X}&\left\{\langle h_t, x - x_t\rangle + \frac{1}{2\alpha}\|x - x_t\|^2\right\}\\ &- \langle c_t p_t, x_{t+1} - x_t\rangle.\end{aligned} \tag{61}$$

By A1), A4), and (58), for $\langle c_t p_t, x_{t+1} - x_t\rangle$, we have

$$\begin{aligned}&\langle c_t p_t, x_{t+1} - x_t\rangle \le c_t\|p_t\|\|x_{t+1} - x_t\|\|\\ &\le c_t\sqrt{d}L_H(\|x_{t+1}\| + \|x_t\|) \le 2M\sqrt{d}L_H c_t.\end{aligned} \tag{62}$$

Again by using (57), we obtain

$$\begin{aligned}f_t(x_{t+1}) - f_t(x_t) \le \min_{x\in X}&\left\{\langle \nabla f_t(x_t), x - x_t\rangle + \frac{1}{2\alpha}\|x - x_t\|^2\right.\\ &\left. + \langle c_t p_t, x - x_t\rangle\right\} - \langle c_t p_t, x_{t+1} - x_t\rangle \le \min_{x\in X}\{f_t(x) - f_t(x_t)\\ &\left. + \frac{1}{2\alpha}\|x - x_t\|^2 + 2M\sqrt{d}L_H c_t\right\} + 2M\sqrt{d}L_H c_t.\end{aligned} \tag{63}$$

TABLE I
SELECTION OF STEP SIZE

	Step Size α
Algorithms 1, 2, 3	$\alpha \in \left(\frac{4(1-\rho)}{\sigma}, \min\{1/L_S, 2/\sigma\}\right]$

Algorithm 4: Online Frank–Wolfe Algorithm (OFWA) with Line-Search for Step Size.

Initialization: An initial estimate x_1 and the maximal number T of iterations.

For $t = 1, \ldots, T$,

perform the linear optimization:

$$v_t = \arg\min_{v \in X} \langle \nabla f_t(x_t), v \rangle \tag{64}$$

select the step size by line-search:

$$\alpha_t = \arg\min_{\alpha \in [0,1]} f_t(x_t + \alpha(v_t - x_t)) \tag{65}$$

update the estimates as:

$$x_{t+1} = (1 - \alpha_t)x_t + \alpha_t v_t \tag{66}$$

end;

By (57)–(63) and following the proofs of Theorem 1, we can obtain the result. ■

Remark 2: Under the framework of the forgetting-factor regret, Theorems 1–3 establish a unified interval for selection of the step size for Algorithms 1–3; see, e.g., Table I. It indicates that as far as the tracking of the time-varying minimizers of OCO is of concern, the step size should be chosen neither too small nor too large. To the authors' knowledge, such interval for selection of the step size has not been reported in the literature.

C. Forgetting-Factor Regret of Online Frank–Wolfe Algorithm

In the design of Algorithms 1–3, there is a projection operator, which usually results in high computational complexity for high-dimensional problems. To deal with this difficulty, in literature the online conditional gradient algorithm (OCG) [11], [35], the one-shot Frank–Wolfe algorithm [12], [17], [34], are introduced. The key idea of these algorithms lies in replacing the projection operation with a linear optimization, which can be performed more efficiently.

In what follows, we propose an online projection-free algorithm and analyze the corresponding forgetting-factor regret. The algorithm is motivated by the one-shot Frank–Wolfe algorithm [34] as well as the line-search procedure for selection of the step size, see, e.g. [35]. To the author's knowledge, combination of the one-shot Frank-Wolfe algorithm with the line-search procedure for the step size is new and has not been reported in the existing literature.

Theorem 4: For $\{x_t\}_{t\geq 1}$ generated by Algorithm 4, if A1)–A3) are satisfied, then for any fixed $\rho \in (0,1)$ and $\alpha_0 \in (0,1]$, it holds that

$$\begin{aligned}\text{regret}_T^F &= \sum_{t=1}^{T} \rho^{T-t} \left(f_t(x_t) - f_t(x_t^*)\right) \\ &\leq \sum_{t=1}^{T} \rho^{T-t}(1-\alpha_0)^{t-1} \left(f_1(x_1) - f_1(x_1^*)\right) \\ &\quad + 2\sum_{t=1}^{T} \rho^{T-t} \widetilde{F}_t^{(\alpha_0)} + L \sum_{t=1}^{T} \rho^{T-t} \widetilde{\theta}_t^{(\alpha_0)} + \frac{2M^2 L_S}{1-\rho}\alpha_0\end{aligned} \tag{67}$$

where $\widetilde{F}_t^{(\alpha_0)} = \sum_{s=1}^{t-1}(1-\alpha_0)^{t-s-1} F_{s,s+1}^{\sup}$ and $\widetilde{\theta}_t^{(\alpha_0)} = \sum_{s=1}^{t-1}(1-\alpha_0)^{t-s-1}\theta_s$. Moreover, if $F_{t,t+1}^{\sup} = o(1)$ and $\theta_t = o(1)$ as $t \to \infty$, then $\lim_{T\to\infty} \text{regret}_T^F = 0$.

Theorem 4 establishes the upper bound of the forgetting-factor regret of Algorithm 4 and gives the sufficient conditions to guarantee $\text{regret}_T^F = o(1)$ as $T \to \infty$. The key step toward establishing the upper bound of regret_T^F is to find a recursive formula for $f_t(x_t) - f_t(x_t^*)$, $t \geq 0$ by using the linear optimization (64) and the line-search procedure (65) for selection of the step size.

Proof: Noting that α_t is selected by the line-search procedure (65), for any fixed $\alpha_0 \in [0,1]$, it follows that

$$f_t(x_{t+1}) = f_t\left(x_t + \alpha_t(v_t - x_t)\right) \leq f_t\left(x_t + \alpha_0(v_t - x_t)\right) \tag{68}$$

from which together with the assumption that $f_t(x)$ is L_S-smooth over X,

$$\begin{aligned}f_t(x_{t+1}) &\leq f_t\left(x_t + \alpha_0(v_t - x_t)\right) \\ &\leq f_t(x_t) + \langle \nabla f_t(x_t), \alpha_0(v_t - x_t)\rangle \\ &\quad + \frac{L_S}{2}\alpha_0^2 \|v_t - x_t\|^2.\end{aligned} \tag{69}$$

By (64) we know that $\langle \nabla f_t(x_t), v_t\rangle \leq \langle \nabla f_t(x_t), x_t^*\rangle$, which combining with (69) gives

$$f_t(x_{t+1}) \leq f_t(x_t) + \alpha_0 \langle \nabla f_t(x_t), x_t^* - x_t\rangle + \frac{L_S}{2}\alpha_0^2 \|v_t - x_t\|^2. \tag{70}$$

By the convexity of $f_t(\cdot)$ in X, we have

$$\langle \nabla f_t(x_t), x_t^* - x_t\rangle \leq f_t(x_t^*) - f_t(x_t). \tag{71}$$

By (70) and (71), we obtain

$$f_t(x_{t+1}) \leq f_t(x_t) - \alpha_0 \left(f_t(x_t) - f_t(x_t^*)\right) + \frac{L_S}{2}\alpha_0^2 \|v_t - x_t\|^2. \tag{72}$$

On the other hand, a similar analysis as (16), (17) leads to

$$\begin{aligned}&f_{t+1}(x_{t+1}) - f_{t+1}(x_{t+1}^*) \\ &\leq 2F_{t,t+1}^{\sup} + f_t(x_{t+1}) - f_t(x_t) + f_t(x_t) - f_t(x_t^*) + L\theta_t.\end{aligned} \tag{73}$$

Combining (72) and (73) and noting that A1) ensures $\|v_t - x_t\|^2 \leq 4M^2$, we obtain

$$\begin{aligned}f_{t+1}(x_{t+1}) - f_{t+1}(x_{t+1}^*) &\leq (1-\alpha_0)\left(f_t(x_t) - f_t(x_t^*)\right) \\ &\quad + 2F_{t,t+1}^{\sup} + L\theta_t + 2M^2 L_S \alpha_0^2.\end{aligned} \tag{74}$$

and furthermore,

$$\begin{aligned} & f_t(x_t) - f_t(x_t^*) \\ \le & (1-\alpha_0)^{t-1}\left(f_1(x_1) - f_1(x_1^*)\right) + 2\sum_{s=1}^{t-1}(1-\alpha_0)^{t-s-1}F_{s,s+1}^{\sup} \\ & + L\sum_{s=1}^{t-1}(1-\alpha_0)^{t-s-1}\theta_s + 2\,M^2 L_S \sum_{s=1}^{t-1}(1-\alpha_0)^{t-s-1}\alpha_0^2 \\ \le & (1-\alpha_0)^{t-1}\left(f_1(x_1) - f_1(x_1^*)\right) + 2\sum_{s=1}^{t-1}(1-\alpha_0)^{t-s-1}F_{s,s+1}^{\sup} \\ & + L\sum_{s=1}^{t-1}(1-\alpha_0)^{t-s-1}\theta_s + 2\,M^2 L_S \alpha_0. \end{aligned} \tag{75}$$

Multiplying ρ^{T-t} to both sides of (75) and then summing up from $t=1$ to T leads to (67).

From (67) it follows that if $F_{t,t+1}^{\sup} = o(1)$ and $\theta_t = o(1)$ as $t \to \infty$, then

$$0 \le \limsup_{T\to\infty} \text{regret}_T^F \le \frac{2M^2 L_S}{1-\rho}\alpha_0.$$

Noting that α_0 is arbitrary in the interval (0, 1], we obtain $\lim_{T\to\infty} \text{regret}_T^F = 0$. The proof is thus concluded. ■

D. OGDA With Forgetting Factor

In the above sections, we analyze the forgetting-factor regrets of OGDA, OGFA, and OFWA. Note that the forgetting factor ρ does not appear in these algorithms. On the other hand, since $\{f_t(\cdot), t=1,\dots,T\}$ are time varying, to predict the minimizer x_{t+1}^* of $f_{t+1}(x), x \in X$, it is reasonable to use the information in $\{x_s, f_s(x_s), \nabla f_s(x_s)\}$ at the recent time instants s near to t, while those at time instants away from t may be negligible. Based on the consideration, in what follows, we introduce an OGDA with a forgetting factor and analyze the corresponding regret_T^F.

We first reformulate the OCO problem as follows. Let $\{f_t(\cdot), t=1,\dots,T\}$ be a sequence of convex functions over the constraint set X. Denote $x_t^* = \arg\min_{x\in X} f_t(x), t=1,\dots,T$ and assume there exists an unknown matrix $A \in \mathbb{R}^{d\times d}$ with $\|A\| \le 1$ such that

$$x_{t+1}^* = Ax_t^* + \xi_t \tag{76}$$

where ξ_t is a random noise.

Here we consider the full information model for OCO, i.e., the gradient $\nabla f_t(\cdot), t=1,\dots,T$ are available for the algorithm design. We first consider the case of $X = \mathbb{R}^d$. The idea arises from the following observation. Recall that the classical regularized approach for OCO, see, e.g., [10], can be formulated by $x_{t+1} = \arg\min_{x\in\mathbb{R}^d}\{\sum_{s=1}^t \langle \nabla f_s(x_s), x\rangle + \frac{1}{2\alpha}\|x\|^2\}$. By introducing a forgetting factor $\rho_0 \in [0,1]$ into the regularized algorithm, i.e.,

$$x_{t+1} = \arg\min_{x\in\mathbb{R}^d}\left\{\sum_{s=1}^t \rho_0^{t-s}\langle d_s, x\rangle + \frac{1}{2\alpha}\|x\|^2\right\} \tag{77}$$

the predict x_{t+1} for x_{t+1}^* is then generated by

$$x_{t+1} = \rho_0 x_t - \alpha d_t \tag{78}$$

where α is a positive step size, $d_t = \nabla f_t(x_t) + \epsilon_t$ and ϵ_t is the observation noise for the gradient. From (77), (78) and taking the constraint X into consideration, we propose the following algorithm.

Algorithm 5: OGDA with Forgetting Factor.

Initialization: An initial estimate x_1, a forgetting factor $\rho_0 \in [0,1]$, the maximal number T of iterations, and a sequence of step size $\{\alpha_t\}_{t=1}^T$.

For $t = 1,\dots,T$**,** update the estimate as:

$$\widehat{x}_{t+1} = \rho_0 x_t - \alpha_t d_t \tag{79}$$

$$x_{t+1} = \mathcal{P}_X\left(\widehat{x}_{t+1}\right) \tag{80}$$

end;

Remark 3: The OCO model (76) is similar to the model considered in [47]. The difference lies in that the matrix A is known in [47] as a priori information, while in this paper we assume A is a unknown matrix. Examples of the model (76) include the prediction of random signals in signal processing, the identification of time-varying parameter vectors in systems and control, etc.

With the forgetting factor ρ_0, (76) can be rewritten as

$$x_{t+1}^* = \rho_0 x_t^* + \omega_t \tag{81}$$

where

$$\omega_t = (A - \rho_0 I)x_t^* + \xi_t. \tag{82}$$

Before proceeding further, we pose the following assumption for the theoretical analysis.

A6) For the random noises $\{\epsilon_t\}_{t=1}^T$ and $\{\xi_t\}_{t=1}^T$,

a) $\mathbb{E}[\epsilon_t] = 0$, $\mathbb{E}\|\epsilon_t\|^2 \le v^2$ for some positive constant v and any $t = 1,\dots,T$;

b) ϵ_s and ϵ_t are mutually independent for any $1 \le s < t \le T$;

c) $\sup_{t=1,\dots,T}\mathbb{E}\|\xi_t\| < \infty$;

d) ξ_s and ξ_t are mutually independent for any $1 \le s < t \le T$.

Set the sequence of σ-algebras $\mathcal{F}_t \triangleq \sigma\{\epsilon_s, \xi_s : 1 \le s \le t\}$, $t = 1,\dots,T$. From the definition of $\mathcal{F}_t$ and noting Algorithm 5 and the assumption $x_{t+1}^* = Ax_t^* + \xi_t$, it is direct to check that both $f_{t+1}(x_{t+1})$ and $f_{t+1}(x_{t+1}^*)$ are random variables measurable with respect to $\mathcal{F}_t$ and in the following the forgetting factor regret of Algorithm 5 in mathematical expectation will be considered.

The key step for analyzing Algorithm 5 is to find the recursive formula for the sequence $\{\mathbb{E}\|x_t - x_t^*\|^2\}_{t\ge 0}$, which is given by the following technical lemma.

Lemma 1: Let $\{x_t\}_{t=1}^T$ be generated by Algorithm 5. Assume that A1), A2), and A6) hold and $f_t(\cdot), t=1,\dots,T$ are σ-strongly convex. Furthermore, assume the constraint set X

satisfies $ax \in X$ for $\forall\, a \in (0,1)$ if $x \in X$. Then,

$$\mathbb{E}\|x_{t+1}-x_{t+1}^*\|^2 \le \rho_0^2\mathbb{E}\|x_t-x_t^*\|^2-2\rho_0\alpha_t\mathbb{E}\left(f_t(x_t)-f_t(x_t^*)\right) - \rho_0\sigma\alpha_t\mathbb{E}\|x_t-x_t^*\|^2+(L+v)^2\alpha_t^2+4M\mathbb{E}\|\omega_t\|. \tag{83}$$

Proof: By A1) and the Schwartz inequality, we have the following chain of equality and inequality:

$$\begin{aligned}
&\mathbb{E}\left\|x_{t+1}-x_{t+1}^*\right\|^2-\mathbb{E}\left\|\rho_0x_t-\rho_0x_t^*\right\|^2\\
=\;&-\mathbb{E}\left\|\rho_0x_t-x_{t+1}\right\|^2+\mathbb{E}\left\langle\rho_0x_t-x_{t+1},\rho_0x_t^*-x_{t+1}^*\right\rangle\\
&+\mathbb{E}\left\langle\rho_0x_t-x_{t+1},\rho_0x_t^*+x_{t+1}^*-2x_{t+1}\right\rangle\\
&+\mathbb{E}\left\langle2x_{t+1}-x_{t+1}^*-\rho_0x_t^*,\rho_0x_t^*-x_{t+1}^*\right\rangle\\
\le\;&-\mathbb{E}\left\|\rho_0x_t-x_{t+1}\right\|^2+\mathbb{E}\left\langle\rho_0x_t-x_{t+1},2\rho_0x_t^*-2x_{t+1}\right\rangle\\
&+4M\mathbb{E}\|\omega_t\|.
\end{aligned}\tag{84}$$

By (79), (80), and the definition of the projection operator $\mathcal{P}_X(\cdot)$, we have that for Algorithm 5

$$x_{t+1}=\arg\min_{x\in X}\left\{\langle2\alpha_td_t-2\rho_0x_t,x\rangle+\|x\|^2\right\}. \tag{85}$$

By Lemma A3 given in Appendix with $x^*=x_{t+1}$ and $x=\rho_0x_t^*$, we obtain

$$2\left\langle x_{t+1}+\alpha_td_t-\rho_0x_t,\rho_0x_t^*-x_{t+1}\right\rangle\ge0 \tag{86}$$

and thus

$$\begin{aligned}
&\mathbb{E}\left\langle\rho_0x_t-x_{t+1},\rho_0x_t^*-x_{t+1}\right\rangle\le\alpha_t\mathbb{E}\left\langle d_t,\rho_0x_t^*-x_{t+1}\right\rangle\\
&=\alpha_t\mathbb{E}\left\langle d_t,\rho_0x_t^*-\rho_0x_t\right\rangle+\alpha_t\mathbb{E}\left\langle d_t,\rho_0x_t-x_{t+1}\right\rangle\\
&=E_t(1)+E_t(2)+E_t(3)
\end{aligned}\tag{87}$$

where

$$E_t(1)=\rho_0\alpha_t\mathbb{E}\left[\langle\nabla f_t(x_t),x_t^*-x_t\rangle\right] \tag{88}$$

$$E_t(2)=\rho_0\alpha_t\mathbb{E}\left[\langle\epsilon_t,x_t^*-x_t\rangle\right] \tag{89}$$

and

$$E_t(3)=\alpha_t\mathbb{E}\left[\langle d_t,\rho_0x_t-x_{t+1}\rangle\right]. \tag{90}$$

By the fact that $f_t(\cdot)$ is σ-strongly convex, we have

$$-E_t(1)\ge\rho_0\alpha_t\mathbb{E}\left(f_t(x_t)-f_t(x_t^*)\right)+\rho_0\alpha_t\frac{\sigma}{2}\mathbb{E}\|x_t-x_t^*\|^2. \tag{91}$$

By (79), (80), and assumption A6) b), we know that $x_t-x_t^*$ is measurable with respect to $\mathcal{F}_{t-1}$ and thus independent of ϵ_t. Hence

$$E_t(2)=\rho_0\alpha_t\mathbb{E}\langle\epsilon_t,x_t-x_t^*\rangle=\rho_0\alpha_t\langle\mathbb{E}\epsilon_t,\mathbb{E}\left(x_t^*-x_t\right)\rangle=0 \tag{92}$$

where in the last equation assumption A6) c) is used.

By A2), A6) a), and the Lyapunov inequality,

$$\mathbb{E}\left\|d_t\right\|=\mathbb{E}\left\|\nabla f_t(x_t)+\epsilon_t\right\|\le\mathbb{E}\left\|\nabla f_t(x_t)\right\|+\mathbb{E}\left\|\epsilon_t\right\|\le L+v \tag{93}$$

and

$$\begin{aligned}
\mathbb{E}\left\|d_t\right\|^2&=\mathbb{E}\left\|\nabla f_t(x_t)+\epsilon_t\right\|^2\\
&\le\mathbb{E}\left\|\nabla f_t(x_t)\right\|^2+\mathbb{E}\left\|\epsilon_t\right\|^2+2\mathbb{E}\left\|\nabla f_t(x_t)\right\|\left\|\epsilon_t\right\|\\
&\le L^2+v^2+2Lv=(L+v)^2.
\end{aligned}\tag{94}$$

By using the Cauchy inequality and (94), for $E_t(3)$ we have the following estimate:

$$\begin{aligned}
E_t(3)=\alpha_t\mathbb{E}\left[\langle d_t,\rho_0x_t-x_{t+1}\rangle\right]&\le\mathbb{E}\left(\alpha_t\left\|d_t\right\|\left\|\rho_0x_t-x_{t+1}\right\|\right)\\
&\le\frac{1}{2}\alpha_t^2\mathbb{E}\left\|d_t\right\|^2+\frac{1}{2}\mathbb{E}\left\|\rho_0x_t-x_{t+1}\right\|^2\\
&\le\frac{(L+v)^2}{2}\alpha_t^2+\frac{1}{2}\mathbb{E}\left\|\rho_0x_t-x_{t+1}\right\|^2.
\end{aligned}\tag{95}$$

Combining (87)–(95) leads to

$$\begin{aligned}
&\mathbb{E}\left\langle\rho_0x_t-x_{t+1},2\rho_0x_t^*-2x_{t+1}\right\rangle\\
&\le-2\rho_0\alpha_t\mathbb{E}\left(f_t(x_t)-f_t(x_t^*)\right)-\rho_0\sigma\alpha_t\mathbb{E}\left\|x_t-x_t^*\right\|^2\\
&\quad+(L+v)^2\alpha_t^2+\mathbb{E}\left\|\rho_0x_t-x_{t+1}\right\|^2
\end{aligned}\tag{96}$$

which incorporated with (84) yields (83). ■

Theorem 5: Let $\{x_t\}_{t=1}^T$ be generated by Algorithm 5. Assume that A1), A2), and A6) hold and the constraint set X satisfies $ax\in X$ for $\forall\, a\in(0,1)$ if $x\in X$. Then, for any $0\le\rho\le\rho_0^2$,

a) If $f_t(\cdot)$, $t=1,\ldots,T$ are convex and the step size $\alpha_t\equiv\alpha>0$, then

$$\begin{aligned}
\mathbb{E}[\text{regret}_T^F]&=\sum_{t=1}^T\rho^{T-t}\mathbb{E}\left(f_t(x_t)-f_t(x_t^*)\right)\\
&\le\frac{\rho_0^{2T-1}}{2\alpha}\mathbb{E}\left\|x_1-x_1^*\right\|^2+\frac{(L+v)^2\alpha}{2\rho_0(1-\rho_0^2)}\\
&\quad+\frac{2\,M}{\rho_0\alpha}\sum_{t=1}^T\rho_0^{2T-2t}\mathbb{E}\|\omega_t\|
\end{aligned}\tag{97}$$

b) If $f_t(\cdot)$, $t=1,\ldots,T$ are σ-strongly convex and the step size $\{\alpha_t\}_{t=1}^T$ is a decreasing positive sequence, then

$$\begin{aligned}
\mathbb{E}[\text{regret}_T^F]&=\sum_{t=1}^T\rho^{T-t}\mathbb{E}\left(f_t(x_t)-f_t(x_t^*)\right)\\
&\le\frac{\rho_0^{2T-1}}{2\alpha_1}\mathbb{E}\|x_1-x_1^*\|^2+\sum_{t=1}^T\rho_0^{2T-2t-1}\\
&\quad\times\left(\frac{(L+v)^2}{2}\alpha_t+2M\frac{\mathbb{E}\|\omega_t\|}{\alpha_t}\right)\\
&\quad+\frac{1}{2}\sum_{t=2}^T\rho_0^{2T-2t+1}\mathbb{E}\|x_t-x_t^*\|^2\\
&\quad\times\left(\frac{1}{\alpha_t}-\frac{1}{\alpha_{t-1}}-\frac{\sigma}{\rho_0}\right).
\end{aligned}\tag{98}$$

Moreover, if $\alpha_0 = 1$, $\alpha_t = c/t$ with $0 < c \le \frac{\rho_0}{\sigma}$, and $\mathbb{E}\|\omega_t\| = o(\alpha_t)$ as $t \to \infty$, then $\mathbb{E}[\text{regret}_T^F] = o(1)$ as $T \to \infty$.

Proof: Here we only prove Theorem 5b), while a) can be proved by setting $\sigma = 0$ and $\alpha_t \equiv \alpha$.

By Lemma 1, it follows that

$$\begin{aligned}&\rho_0\mathbb{E}\left(f_t(x_t) - f_t(x_t^*)\right)\\ &\le \frac{\rho_0^2\mathbb{E}\left\|x_t - x_t^*\right\|^2 - \mathbb{E}\left\|x_{t+1} - x_{t+1}^*\right\|^2}{2\alpha_t} - \frac{\sigma}{2}\rho_0\mathbb{E}\left\|x_t - x_t^*\right\|^2\\ &\quad + \frac{(L+v)^2}{2}\alpha_t + 2M\frac{\mathbb{E}\|\omega_t\|}{\alpha_t}. \end{aligned} \tag{99}$$

Set $\rho_1 = \rho_0^2$. Multiplying ρ_1^{T-t-1} to both sides of (99) and summing up $t = 1, \ldots, T$ leads to

$$\begin{aligned}&\sum_{t=1}^{T}\rho_1^{T-t-1/2}\mathbb{E}\left(f_t(x_t) - f_t(x_t^*)\right)\\ &\le \frac{1}{2}\sum_{t=1}^{T}\frac{\rho_1^{T-t}\mathbb{E}\left\|x_t - x_t^*\right\|^2 - \rho_1^{T-t-1}\mathbb{E}\left\|x_{t+1} - x_{t+1}^*\right\|^2}{\alpha_t}\\ &\quad - \frac{\sigma}{2}\sum_{t=1}^{T}\rho_1^{T-t-1/2}\mathbb{E}\left\|x_t - x_t^*\right\|^2 + \frac{(L+v)^2}{2}\sum_{t=1}^{T}\rho_1^{T-t-1}\alpha_t\\ &\quad + 2M\sum_{t=1}^{T}\rho_1^{T-t-1}\frac{\mathbb{E}\|\omega_t\|}{\alpha_t}. \end{aligned} \tag{100}$$

We have the following chain of equalities and inequalities:

$$\begin{aligned}&\sum_{t=1}^{T}\frac{\rho_1^{T-t}\mathbb{E}\left\|x_t - x_t^*\right\|^2 - \rho_1^{T-t-1}\mathbb{E}\left\|x_{t+1} - x_{t+1}^*\right\|^2}{\alpha_t}\\ &= \sum_{t=1}^{T}\frac{\rho_1^{T-t}\mathbb{E}\left\|x_t - x_t^*\right\|^2}{\alpha_t} - \sum_{t=1}^{T}\frac{\rho_1^{T-t-1}\mathbb{E}\left\|x_{t+1} - x_{t+1}^*\right\|^2}{\alpha_t}\\ &= \sum_{t=1}^{T}\frac{\rho_1^{T-t}\mathbb{E}\left\|x_t - x_t^*\right\|^2}{\alpha_t} - \sum_{t=2}^{T+1}\frac{\rho_1^{T-t}\mathbb{E}\left\|x_t - x_t^*\right\|^2}{\alpha_{t-1}}\\ &\le \frac{\rho_1^{T-1}}{\alpha_1}\mathbb{E}\left\|x_1 - x_1^*\right\|^2 + \sum_{t=2}^{T}\mathbb{E}\left\|x_t - x_t^*\right\|^2\left(\frac{\rho_1^{T-t}}{\alpha_t} - \frac{\rho_1^{T-t}}{\alpha_{t-1}}\right) \end{aligned} \tag{101}$$

from which and noting (100)

$$\begin{aligned}&\sum_{t=1}^{T}\rho_1^{T-t-1/2}\mathbb{E}\left(f_t(x_t) - f_t(x_t^*)\right) \le \frac{\rho_1^{T-1}}{2\alpha_1}\mathbb{E}\left\|x_1 - x_1^*\right\|^2\\ &+ \frac{1}{2}\sum_{t=2}^{T}\mathbb{E}\left\|x_t - x_t^*\right\|^2\left(\frac{\rho_1^{T-t}}{\alpha_t} - \frac{\rho_1^{T-t}}{\alpha_{t-1}} - \sigma\rho_1^{T-t-1/2}\right)\\ &+ \frac{(L+v)^2}{2}\sum_{t=1}^{T}\rho_1^{T-t-1}\alpha_t + 2M\sum_{t=1}^{T}\rho_1^{T-t-1}\frac{\mathbb{E}\|\omega_t\|}{\alpha_t}. \end{aligned} \tag{102}$$

Noting $0 \le \rho \le \rho_0^2 = \rho_1$, we arrive that

$$\begin{aligned}&\sum_{t=1}^{T}\rho^{T-t}\mathbb{E}\left(f_t(x_t) - f_t(x_t^*)\right) \le \sum_{t=1}^{T}\rho_1^{T-t}\mathbb{E}\left(f_t(x_t) - f_t(x_t^*)\right)\\ &\le \frac{\rho_0^{2T-1}}{2\alpha_1}\mathbb{E}\left\|x_1 - x_1^*\right\|^2\\ &+ \frac{\rho_0}{2}\sum_{t=2}^{T}\mathbb{E}\left\|x_t - x_t^*\right\|^2\left(\frac{\rho_0^{2T-2t}}{\alpha_t} - \frac{\rho_0^{2T-2t}}{\alpha_{t-1}} - \sigma\rho_0^{2T-2t-1}\right)\\ &+ \frac{(L+v)^2}{2}\sum_{t=1}^{T}\rho_0^{2T-2t-1}\alpha_t + 2M\sum_{t=1}^{T}\rho_0^{2T-2t-1}\frac{\mathbb{E}\|\omega_t\|}{\alpha_t}. \end{aligned} \tag{103}$$

Thus, the conclusion of (98) is proved.

If $\alpha_0 = 1$, $\alpha_t = c/t$ with $0 < c \le \frac{\rho_0}{\sigma}$ and $\mathbb{E}\|\omega_t\| = o(\alpha_t)$ as $t \to \infty$, then it is direct to check that $\mathbb{E}[\text{regret}_T^F] = o(1)$ as $T \to \infty$. ■

Remark 4: Equation (97) in Theorem 5 indicates that to obtain a small bound of regret_T^F, an optimal step size α for Algorithm 5 should be chosen as a tradeoff between the magnitude of the gradients of objective functions and the observation noise variance, i.e., $L + v$, as well as the magnitude of changing of the time-varying minimizers of $f_t(\cdot)$, i.e., $\sup_t \mathbb{E}\|\omega_t\|$. In [40], some classical results for optimal tracking of time-varying parameters in linear stochastic systems are obtained. To be specific, let the linear systems be given by

$$y_{t+1} = \theta_t^\top \varphi_t + \omega_{t+1}$$

where $\{\theta_t\}_{t\ge1}$ are the unknown time-varying parameters, $\{y_{t+1}, \varphi_t\}_{t\ge1}$ are the observations, and $\{\omega_t\}_{t\ge1}$ are the system noises. Denote the estimates for $\{\theta_t\}_{t\ge1}$ by $\{\widehat{\theta}_t\}_{t\ge1}$. In [40, Th. 3.1] it is established that for the classical algorithms for estimating $\{\theta_t\}_{t\ge1}$, including LMS, FFLS, and the KF-based algorithm, the estimation error is bounded by

$$\mathbb{E}[\|(\theta_t - \widehat{\theta}_t)(\theta_t - \widehat{\theta}_t)^\top\|] = O\left(\mu + \frac{1}{\mu}\right) \tag{104}$$

where $\mu > 0$ is the step size adopted in LMS, FFLS, the KF-based algorithm. Note that (97) characterizes the tracking performance by $\mathbb{E}(f_T(x_T) - f_T(x_T^*)) = O(\alpha + \frac{1}{\alpha})$. So in this sense (97) can be regarded as a parallel extension of results in [40] to the nonlinear OCO problem.

Remark 5: By setting $\rho = \rho_0 = 1$, Theorem 5 also generates a bound of the dynamic regret of OGDA.

Remark 6: The OCO model (76) can also be seen as a way of adding a priori knowledge about the sequence of convex optimization functions $f_t(x), t = 1, \ldots, T$. Similar idea, named the online learning with a predictable sequence, was considered in [52], etc. To be specific, the online linear optimization $f_t(x) = \langle x, y_t\rangle$, $t \ge 1$ with a predictable sequence $y_t = M_t(y_1, \ldots, y_{t-1}) + \epsilon_t$, where ϵ_t is the noise is considered in [52] and the static regrets are analyzed for both the accurate and inaccurate prediction model. Compared with [52], this article considers the general OCO functions $f_t(x), t = 1, \ldots, T$ and

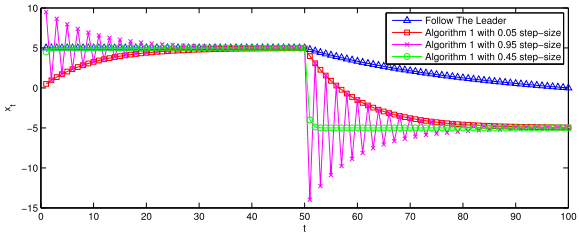


Fig. 1. Estimates of follow the leader algorithm and Algorithm 1 with $\alpha = 0.05,\ 0.45$, and 0.95.

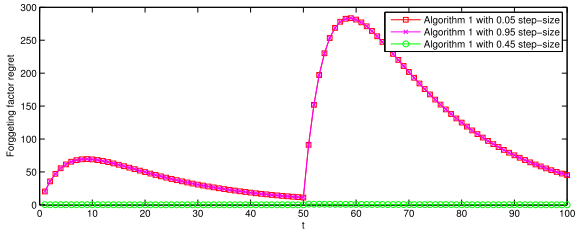


Fig. 2. Forgetting factor regret of Algorithm 1 with $\alpha = 0.05,\ 0.45$, and 0.95.

analyze the performance of Algorithm 5 under the new regret function regret_T^F.

III. Simulation Studies

Example 1: (Hazan and Seshadhri's example, [27]) Define the loss functions of OCO by

$$f_t(x) = \begin{cases} (x-5)^2, & t = 1, \ldots, \frac{T}{2} \\ (x+5)^2, & t = \frac{T}{2}+1, \ldots, T. \end{cases} \tag{105}$$

Set $T = 100$ and the constraint set $X = [-10, 10]$. We test the performance of the leader algorithm and Algorithm 1 through this example.

By the leader algorithm, it follows that

$$x_{t+1} = \arg\min_{x \in X} \sum_{s=1}^{t} f_s(x) \tag{106}$$

and

$$x_t = \begin{cases} 5, & t = 1, \ldots, \frac{T}{2} \\ \frac{5T}{t} - 5, & t = \frac{T}{2}+1, \ldots, T. \end{cases} \tag{107}$$

As shown in [6], for the above algorithm, the static regret function is bounded by $\text{regret}_T^S = O(\log T)$. On the other hand, the estimate x_T equals 0, rather than the minimizers -5 of $f_T(\cdot)$.

Set the forgetting factor $\rho = 0.95$. For this example, it is readily checked that the interval for selection of the step size, see, e.g., Theorem 1, $(\frac{4(1-\rho)}{\sigma}, \min\{1/L_S, 2/\sigma\}] = (0.1, 0.5]$. We testify the performance of Algorithm 1 with $\alpha = 0.05,\ 0.45$, and 0.95. It is clear that only the step size $\alpha = 0.45$ belongs to the above interval.

Fig. 1 shows trajectories of the estimates generated from the leader algorithm and Algorithm 1 with $\alpha = 0.05,\ 0.45$, and 0.95. Fig. 2 shows the values of $\text{regret}_t^F, t = 1, \ldots, T$ of Algorithm 1 with $\alpha = 0.05,\ 0.45$, and 0.95. Simulation results indicate that for OCO, as far as the tracking of the time-varying minimizers is concerned, performance of Algorithm 1 can be well evaluated by the forgetting-factor regret.

Example 2: We consider the following system (see, e.g., [48]), which is often used for the modeling of aircraft tracking, spacecraft intersection, etc.

Denote by z_s and $\tilde{z}_s$ the positions of the tracker and the target at the time interval $s \in [t, t+1)$,

$$z_s = \sum_{k=1}^{d} x_t[k] c_{k,t}(s)$$

$$\tilde{z}_s = \sum_{k=1}^{d} \xi_t[k] c_{k,t}(s)$$

where $x_t \in \mathbb{R}^d$ and $\xi_t \in \mathbb{R}^d$ are the coordinate vectors of the tracker and the target at time t, respectively, and $\{c_{k,t}(s), k = 1, \ldots, d\}$ with $s \in [t, t+1)$ are normalized orthogonal functions on the interval $[t, t+1)$, i.e.,

$$\int_t^{t+1} \langle c_{k,t}(s), c_{l,t}(s) \rangle \, \mathrm{d}s = \begin{cases} 1, & \text{if } k = l \\ 0, & \text{otherwise.} \end{cases}$$

The values of the loss function at time t are given by

$$\begin{aligned} f_t(x_t) &= \zeta_1 \langle \pi_t, x_t \rangle + \zeta_2 \int_t^{t+1} \|z_s - \tilde{z}_s\|^2 \, ds \\ &= \zeta_1 \langle \pi_t, x_t \rangle + \zeta_2 \|x_t - \xi_t\|^2. \end{aligned}$$

Here, we consider the 1-D system, i.e., $d = 1$ with a constraint set $X = [-2, 2]$. Set $T = 1000$, $\zeta_1 = 0$, $\zeta_2 = 1$, and the target coordinate $\xi_t = \frac{100}{t^2}$. We test the performance of Algorithms 1–5. To be specific, we set the forgetting factor $\rho = 0.8$. The step size is chosen as $\alpha = 0.5$ for Algorithms 1–3; for Algorithm 2 the smoothing parameter is set as $\delta = 0.01$; for Algorithm 3 $\{c_t = \frac{1}{t}\}_{t=1}^T$. For Algorithm 5, choosing the step size $\alpha_t = \sqrt{\rho}/2t$, based on which we can prove that $|\omega_t| = |(1-\rho^{\frac{1}{2}})\xi_t + (\xi_{t+1} - \xi_t)| = O(\frac{1}{t^2})$, $|\omega_t|/\alpha_t = o(1)$, and conditions of Theorem 5b) are satisfied. For comparison, we also test the performance of Algorithm 5 with a constant step size $\alpha = 0.1$. In fact, by setting $\rho_0 = \rho = 1$ in Theorem 5, Algorithm 5 becomes OGDA. By noting $|\omega_t| = |\xi_{t+1} - \xi_t| = O(\frac{1}{t^3})$ and $|\omega_t|/\alpha_t = O(|\frac{1}{t^2} - \frac{1}{(t+1)^2}|) = O(\frac{1}{t^2})$, Theorem 5 ensures that the dynamic regret of OGDA with $\alpha_t = 1/(2t)$ is bounded by $\text{regret}_T^D = O(\log T)$.

Fig. 3 shows the tracking errors $\{|x_t - \xi_t|\}_{t=1}^T$ of the above algorithms. It can be seen that Algorithms 1–5 generally perform better than OGDA, even if OGDA ensures a dynamic regret bound $\text{regret}_T^D = O(\log T)$. Fig. 4 shows the forgetting-factor regrets of Algorithms 1–5. To further test the performance of algorithms, we carry out ten Monte Carlo simulations. In Fig. 5 , the vertical axis shows the averaged regret_t^Fs of the ten simulations and the horizontal axis plots the averaged computation time, which shows that Algorithm 4 equipped with the Frank–Wolfe technique requires less computational complexity as expected.

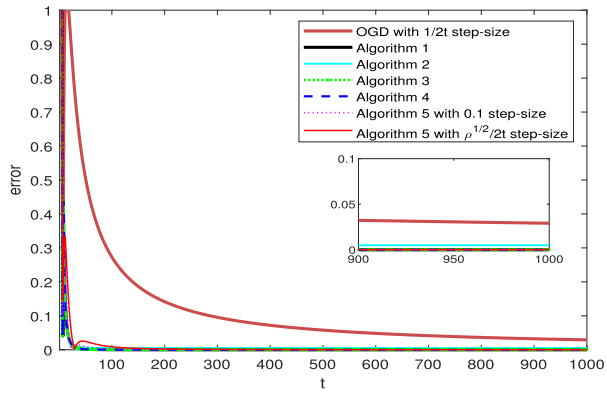


Fig. 3. Tracking errors $|x_t - \xi_t|$ of Algorithms 1–5 and OGDA with step size $\{\alpha_t = 1/(2t)\}$.

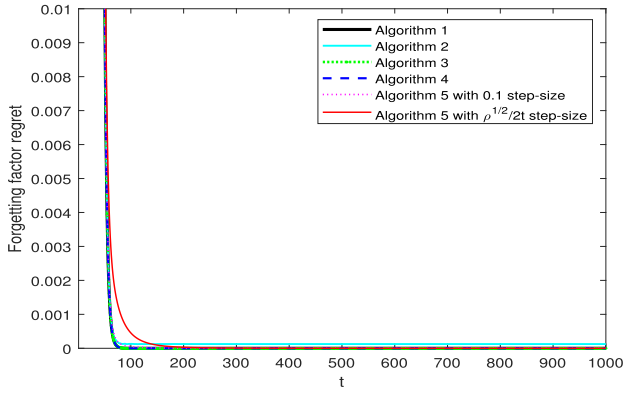


Fig. 4. Forgetting factor regrets of Algorithms 1–5.

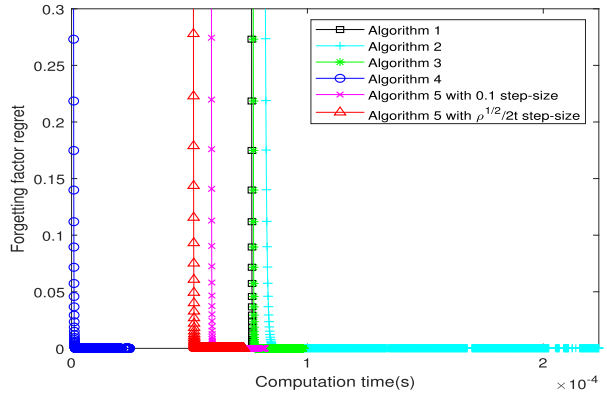


Fig. 5. Computation time of Algorithms 1–5 (unit: second).

IV. Concluding Remarks

In this article, aiming at further improving the tracking performance of the time-varying minimizers of OCO, we propose a new regret, namely, regret with a forgetting factor. We establish the forgetting-factor regret bounds of classical algorithms including OGDA, OGFA, and OFWA as well as a new algorithm, namely, online gradient descent algorithm with a forgetting factor.

The OGDA (Algorithm 1) is one of the most classical algorithms for OCO and it is suitable for the full information model being available, i.e., the gradients of the cost functions can be observed at each time instant t. On the other hand, when only the bandit model is available, the gradients of the cost functions cannot be directly observed and in such cases the OGFA with the δ-smoothing technique and the deterministic difference technique (Algorithms 2 and 3) can be applied. In the design of Algorithms 1–3, there is a projection operator, which usually results in high computational complexity for high-dimensional problems. So for the high-dimensional OCO problem, the OFWA (Algorithm 4) can be considered. When a priori knowledge about the sequence of convex optimization functions $f_t(x), t = 1, \ldots, T$ is available, such knowledge can be used in the algorithm design like the OGDA with a forgetting factor (Algorithm 5).

For future research, it is of interest to relax the conditions on the objective functions such as strong convexity, smoothness, etc., and to obtain a more precise upper bound of regret$_T^F$. It is also of interest to apply the theoretical results in this article to practical scenarios. Theoretical results in the article indicate that the bounds of the forgetting factor regret regret$_T^F$ will be different for different sequences of $\{f_t(\cdot)\}_{t=1}^T$. Another interesting topic is to investigate, within the worst-case framework, a priori bounds on regret$_T^F$ that do not involve individual optimization functions but depend only on some gross geometric features of the problem such as the diameter of the decision set, the Lipschitz constant, the strong convexity, and the smoothness constant, etc.

Appendix

Definition A1: Denote by $\mathrm{dom}(f) \subset \mathbb{R}^d$ the domain of $f(\cdot)$. A vector-valued function $\partial f(x) \in \mathbb{R}^d$ is the subgradient of a nonsmooth convex function $f(x) : \mathrm{dom}(f) \to \mathbb{R}$ if for any $x, y \in \mathrm{dom}(f)$, $f(y) - f(x) \geq \partial f(x)^\top (y - x)$.

Definition A2: Let $f(\cdot)$ be a convex function on a convex set S. Denoting a subgradient of $f(x)$ by $\partial f(x), x \in S$, $f(\cdot)$ is said to be σ-strongly convex on S, if for any $x, y \in S$,

$$f(y) - f(x) \geq \partial f(x)^\top (y - x) + \frac{\sigma}{2}\|y - x\|^2$$

and is said to be L_S-smooth on S, if for any $x, y \in S$,

$$f(y) - f(x) \leq \partial f(x)^\top (y - x) + \frac{L_S}{2}\|y - x\|^2.$$

Lemma A1 (see [11]): Assume that $f(x)$ is σ-strongly convex and differentiable over $\mathrm{dom}(f)$. For any $x \in \mathrm{dom}(f)$,

$$f(x) - f(x^*) \geq \frac{\sigma}{2}\|x - x^*\|^2 \tag{108}$$

with $x^* \in \arg\min_{x \in \mathrm{dom}(f)} f(x)$.

Lemma A2 (see [49]): For the projection operator $\mathcal{P}_X(\cdot) : \mathbb{R}^d \to X$,

$$\|\mathcal{P}_X(x) - \mathcal{P}_X(y)\| \leq \|x - y\| \ \forall x, y \in \mathbb{R}^d.$$

Lemma A3 (see [50]): If a function $f(x)$ is convex and differentiable in X, then for any $x \in X$,

$$\langle \nabla f(x^*), x - x^* \rangle \geq 0$$

with $x^* \in \arg\min_{x \in X} f(x)$.

REFERENCES

[1] L. Zhang, S. Lu, and Z. Zhou, "Adaptive online learning in dynamic environments," in *Proc. Adv. Neural Inf. Process. Syst.*, 2018, pp. 1330–1340.

[2] A. Simonetto, A. Mokhtari, A. Koppel, G. Leus, and A. Ribeiro, "A class of prediction-correction methods for time-varying convex optimization," *IEEE Trans. Signal Process.*, vol. 64, no. 17, pp. 4576–4591, Sep. 2016.

[3] V. N. Vapnik and S. Kotz, *Estimation of Dependences Based on Empirical Data*. Berlin, Germany: Springer, 2006.

[4] N. Cesa-Bianchi, Y. Freund, D. P. Haussler, D. Helmbold, R. E. Schapire, and M. K. Warmuth, "How to use expert advice," *J. ACM*, vol. 44, no. 3, pp. 427–485, 1997.

[5] J. Gordon, "Regret bounds for prediction problems," in *Proc. 12th Annu. Conf. Learn. Theory*, 1999, pp. 29–40.

[6] S. Shalev-Shwartz, "Online learning and online convex optimization," *Found. Trends Mach. Learn.*, vol. 4, no. 2, pp. 107–194, 2012.

[7] N. Cesa-Bianchi and G. Lugosi, *Prediction, Learning, and Games*. New York, NY, USA: Cambridge Univ. Press, 2006.

[8] S. Shalev-Shwartz and Y. Singer, "Convex repeated games and Fenchel duality," in *Proc. Adv. Neural Inf. Process. Syst.*, 2006, pp. 1265–1272.

[9] M. Zinkevich, "Online convex programming and generalized infinitesimal gradient ascent," in *Proc. 20th Int. Conf. Mach. Learn.*, 2003, pp. 928–936.

[10] E. Hazan, "Introduction to online convex optimization," *Found. Trends Optim.*, vol. 2, no. 3/4, pp. 157–325, 2016.

[11] E. Hazan and S. Kale, "Projection-free online learning," in *Proc. 29th Int. Conf. Mach. Learn.*, 2012, pp. 1843–1850.

[12] L. Chen, C. Harshaw, H. Hassani, and A. Karbasi, "Projection-free online optimization with stochastic gradient: From convexity to submodularity," in *Proc. 35th Int. Conf. Mach. Learn.*, 2018, pp. 814–823.

[13] A. D. Flaxman, A. T. Kalai, and H. B. Mcmahan, "Online convex optimization in the bandit setting: Gradient descent without a gradient," in *Proc. 16th Annu. ACM-SIAM Symp. Discrete Algorithms*, 2005, pp. 385–394.

[14] S. Shalev-Shwartz and Y. Singer, "A primal-dual perspective of online learning algorithms," *Mach. Learn.*, vol. 69, no. 2/3, pp. 115–142, 2007.

[15] E. Hazan, A. Agarwal, and S. Kale, "Logarithmic regret algorithms for online convex optimization," *Mach. Learn.*, vol. 69, no. 2/3, pp. 169–192, 2007.

[16] S. Shalev-Shwartz, Y. Singer, N. Srebro, and A. Cotter, "Pegasos: Primal estimated sub-gradient solver for SVM," in *Proc. 24th Int. Conf. Mach. Learn.*, 2007, pp. 807–814.

[17] J. Xie, Z. Shen, C. Zhang, B. Wang, and H. Qian, "Efficient projection-free online methods with stochastic recursive gradient," in *Proc. 34th AAAI Conf. Artif. Intell.*, 2020, pp. 6446–6453.

[18] D. J. Foster and M. Simchowitz, "Logarithmic regret for adversarial online control," in *Proc. 37th Int. Conf. Mach. Learn.*, 2020, pp. 3211–3221.

[19] N. Agarwal, E. Hazan, and K. Singh, "Logarithmic regret for online control," in *Proc. Adv. Neural Inf. Process. Syst.*, 2019, pp. 10175–10184.

[20] R. Jenatton, J. Huang, and C. Archambeau, "Adaptive algorithms for online convex optimization with long-term constraints," in *Proc. 33rd Int. Conf. Mach. Learn.*, 2016, pp. 402–411.

[21] E. Hazan and S. Kale, "Online submodular minimization," *J. Mach. Learn. Res.*, vol. 13, no. 1, pp. 2903–2922, 2012.

[22] A. Agarwal, O. Dekel, and L. Xiao, "Optimal algorithms for online convex optimization with multi-point bandit feedback," in *Proc. 23rd Conf. Learn. Theory*, 2010, pp. 28–40.

[23] E. Hazan and K. Y. Levy, "Bandit convex optimization: Towards tight bounds," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 784–792.

[24] S. Bubeck, R. Eldan, and Y. T. Lee, "Kernel-based methods for bandit convex optimization," *J. ACM*, vol. 68, no. 25, pp. 1–35, 2021.

[25] L. Chen, M. Zhang, and A. Karbasi, "Projection-free bandit convex optimization," in *Proc. 22nd Int. Conf. Artif. Intell. Statist.*, 2019, pp. 2047–2056.

[26] D. Garber and B. Kretzu, "Improved regret bounds for projection-free bandit convex optimization," in *Proc. 23rd Int. Conf. Artif. Intell. Statist.*, 2020, pp. 2196–2206.

[27] E. Hazan and C. Seshadhri, "Adaptive algorithms for online decision problems," *Electron. Colloq. Comput. Complexity*, Rep. 88, 2007.

[28] S. Shahrampour and A. Jadbabaie, "An online optimization approach for multi-agent tracking of dynamic parameters in the presence of adversarial noise," in *Proc. Amer. Control Conf.*, 2017, pp. 3306–3311.

[29] D. J. Foster, A. Rakhlin, and K. Sridharan, "Strongly adaptive online learning," in *Proc. 32nd Int. Conf. Mach. Learn.*, 2015, pp. 1405–1411.

[30] L. Zhang, T. Yang, and Z. Zhou, "Dynamic regret of strongly adaptive methods," in *Proc. 35th Int. Conf. Mach. Learn.*, 2018, pp. 5882–5891.

[31] E. C. Hall and R. M. Willett, "Dynamical models and tracking regret in online convex programming," in *Proc. 30th Int. Conf. Mach. Learn.*, 2013, pp. 579–587.

[32] T. Yang, L. Zhang, R. Jin, and J. Yi, "Tracking slowly moving clairvoyant: Optimal dynamic regret of online learning with true and noisy gradient," in *Proc. 33rd Int. Conf. Mach. Learn.*, 2016, pp. 449–457.

[33] L. Zhang, T. Yang, J. Yi, R. Jin, and Z. Zhou, "Improved dynamic regret for non-degenerate functions," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 732–741.

[34] D. S. Kalhan, A. S. Bedi, A. Koppel, K. Rajawat, and A. Banerjee, "Dynamic online learning via Frank–Wolfe algorithm," *IEEE Trans. Signal Process.*, vol. 69, pp. 932–947, 2021.

[35] Y. Wan, B. Xue, and L. Zhang, "Projection-free online learning in dynamic environments," in *Proc. 35th AAAI Conf. Artif. Intell.*, 2021, pp. 10067–10075.

[36] Y. Gur, A. J. Zeevi, and O. Besbes, "Stochastic multi-armed-bandit problem with non-stationary rewards," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 199–207.

[37] H. Luo, C. Y. Wei, A. Agarwal, and J. Langford, "Efficient contextual bandits in non-stationary worlds," in *Proc. 31st Conf. Learn. Theory*, 2018, pp. 1739–1776.

[38] W. C. Cheung, D. Simchi-Levi, and R. Zhu, "Learning to optimize under non-stationarity," in *Proc. 22nd Int. Conf. Artif. Intell. Statist.*, 2019, pp. 1079–1087.

[39] P. Zhao, G. Wang, L. Zhang, and Z. Zhou, "Bandit convex optimization in non-stationary environments," in *Proc. 23rd Int. Conf. Artif. Intell. Statist.*, 2020, pp. 1508–1518.

[40] L. Guo and L. Ljung, "Performance analysis of general tracking algorithms," *IEEE Trans. Autom. Control*, vol. 40, no. 8, pp. 1388–1402, Aug. 1995.

[41] M. J. Wainwright and M. I. Jordan, "Graphical models, exponential families, and variational inference," *Found. Trends Mach. Learn.*, vol. 1, no. 1/2, pp. 1–305, 2008.

[42] B. Taskar, V. Chatalbashev, D. Koller, and C. Guestrin, "Learning structured prediction models: A large margin approach," in *Proc. 22nd Int. Conf. Mach. Learn.*, 2005, pp. 896–903.

[43] H. J. Kushner and G. Yin, *Stochastic Approximation and Recursive Algorithms and Applications*, 2nd ed. Berlin, Germany: Springer, 2003.

[44] N. Nguyen and G. Yin, "Stochastic approximation with discontinuous dynamics, differential inclusions, and applications," *Ann. Appl. Probability*, vol. 33, no. 1, pp. 780–823, 2022.

[45] W. Zhao and G. Yin, "A class of constrained stochastic convex optimization algorithms with sparsity," *Automatica*, 2023.

[46] J. Kiefer and J. Wolfowitz, "Stochastic estimation of the maximum of a regression function," *Ann. Math. Statist.*, vol. 23, no. 3, pp. 462–466, 1952.

[47] S. Shahrampour and A. Jadbabaie, "Distributed online optimization in dynamic environments using mirror descent," *IEEE Trans. Autom. Control*, vol. 63, no. 3, pp. 714–725, Mar. 2018.

[48] X. Yi, X. Li, L. Xie, and K. H. Johansson, "Distributed online convex optimization with time-varying coupled inequality constraints," *IEEE Trans. Signal Process.*, vol. 68, pp. 731–746, 2020.

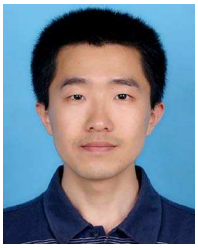
[49] S. Sundhar Ram, A. Nedić, and V. V. Veeravalli, "Distributed stochastic subgradient projection algorithms for convex optimization," *J. Optim. Theory Appl.*, vol. 147, no. 3, pp. 516–545, 2010.

[50] P. T. Harker and J. S. Pang, "Finite-dimensional variational inequality and nonlinear complementarity problems: A survey of theory, algorithms and applications," *Math. Program.*, vol. 48, no. 1/3, pp. 161–220, 1990.

[51] N. Cesa-bianchi, P. Gaillard, G. Lugosi, and G. Stoltz, "Mirror descent meets fixed share (and feels no regret)," in *Proc. 25th Adv. Neural Inf. Process. Syst.*, 2012, pp. 980–988.
[52] A. Rakhlin and K. Sridharan, "Online learning with predictable sequences," in *Proc. 26th Conf. Learn. Theory*, 2013, pp. 993–1019.

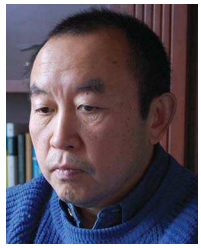
Yuhang Liu received the B.Sc. degree in mathematics from Shandong University, China, in 2018 and the Ph.D. degree in operation research and cybernetics from the Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, China, in 2023.

Her research interests are mainly in distributed/centralized approaches for online optimization and games.



Wenxiao Zhao (Member, IEEE) received the Ph.D. degree in operation research and cybernetics from the Institute of Systems Science, Academy of Mathematics and Systems Science (AMSS), Chinese Academy of Sciences (CAS), China, in 2008.

He is currently a Professor with AMSS, CAS. His research interests are mainly in system identification and adaptive control, variable and feature selection, and distributed stochastic optimization.



George Yin (Life Fellow, IEEE) received the B.S. degree in mathematics from the University of Delaware, Newark, DE, USA, in 1983, and the M.S. degree in electrical engineering and the Ph.D. degree in applied mathematics both from Brown University, Providence, RI, USA, in 1987.

He joined Wayne State University in 1987, became a Professor in 1996, and the University Distinguished Professor in 2017. He moved to the University of Connecticut in 2020. His research interests include stochastic processes, and stochastic systems theory and applications. He served as the Co-chair for a number of conferences and served on many committees for IEEE, IFAC, and SIAM. He was the Chair of the SIAM Activity Group on Control and Systems Theory, and was on the Board of Directors of the American Automatic Control Council. He is the Editor-in-Chief of *SIAM Journal on Control and Optimization,* an Associate Editor of *Applied Mathematics and Optimization*, ESAIM: Control, Optimisation and Calculus of Variations, and on the editorial boards of many other journals. He was an Associate Editor of *Automatica* 1995–2011, IEEE TRANSACTIONS ON AUTOMATIC CONTROL 1994–1998, and a Senior Editor of IEEE CONTROL SYSTEMS LETTERS 2017–2019.

Dr. Yin is a fellow of IFAC and a fellow of SIAM.