Online Companion: Non-stationary Stochastic Optimization

Omar Besbes — Columbia University
Yonatan Gur — Stanford University
Assaf Zeevi* — Columbia University

first version: July 2013, current version: June 17, 2015

B Proofs of additional results

Proof of Proposition 1. The proof of the proposition is established in two steps. In the first step, we limit nature to a class of function sequences $\mathcal{V}'$ where in every epoch nature is limited to one of two specific cost functions, and show that $\mathcal{V}' \subset \mathcal{V}$. In the second step, we show that whenever $\phi \in \left\{\phi^{(0)}, \phi^{(1)}\right\}$, any admissible policy must incur regret of at least order T, even when nature is limited to the set $\mathcal{V}'$.

Step 1. Let $\mathcal{X} = [0,1]$ and fix $T \geq 1$. Let $V_T \in \{1, \ldots, T\}$ and assume that C_1 is a constant such that $V_T \geq C_1 T$. Let $C = \min\left\{C_1, \left(\frac{1}{2} - \nu\right)^2\right\}$ where ν appears in (2), and we assume $\nu < 1/2$. Consider the following two quadratic functions:

$$f^1(x) = x^2 - x + \frac{3}{4}, \qquad f^2(x) = x^2 - (1+2C)\,x + \frac{3}{4} + C.$$

Denoting $x_k^* = \arg\min_{x\in[0,1]} f^k(x)$, we have $x_1^* = \frac{1}{2}$, and $x_2^* = \frac{1}{2}+C$. Define $\mathcal{V}' = \left\{f \,;\, f_t \in \left\{f^1, f^2\right\} \ \forall t \in \mathcal{T}\right\}$. Then, for any sequence in $\mathcal{V}'$ the total functional variation is:

$$\sum_{t=2}^{T} \sup_{x\in\mathcal{X}} |f_t - f_{t-1}| \ \leq\ \sum_{t=2}^{T} \sup_{x\in\mathcal{X}} |2Cx - C| \ \leq\ CT \ \leq\ C_1 T \ \leq\ V_T.$$

For any sequence in $\mathcal{V}'$ the total functional variation (3) is bounded by V_T, and therefore $\mathcal{V}' \subset \mathcal{V}$.

Step 2. Fix $\phi \in \left\{\phi^{(0)}, \phi^{(1)}\right\}$, and let $\pi \in \mathcal{P}_\phi$. Let $\tilde{f}$ to be a random sequence in which in each epoch f_t is drawn according to a discrete uniform distribution over $\left\{f^1, f^2\right\}$ ($\tilde{f}_t$ is independent of $\mathcal{H}_t$ for any $t \in \mathcal{T}$). Any realization of $\tilde{f}$ is a sequence in $\mathcal{V}'$. In particular, taking expectation over $\tilde{f}$, one has:

$$\begin{aligned}
\mathcal{R}_\phi^\pi(\mathcal{V}', T) &\geq \mathbb{E}^{\pi,\tilde{f}}\left[\sum_{t=1}^{T} \tilde{f}_t(X_t) - \sum_{t=1}^{T} \tilde{f}_t(x_t^*)\right] \\
&= \mathbb{E}^{\pi}\left[\sum_{t=1}^{T}\left(\frac{1}{2}\left(f^1(X_t) + f^2(X_t)\right) - \frac{1}{2}\left(f^1(x_1^*) + f^2(x_2^*)\right)\right)\right] \\
&\geq \sum_{t=1}^{T} \min_{x\in[0,1]}\left\{x^2 - (1+C)\,x + \frac{1}{4} + \frac{C}{2} + \frac{C^2}{2}\right\} \ = \ T \cdot \frac{C^2}{4},
\end{aligned}$$

where the minimum is obtained at $x^* = \frac{1+C}{2}$. Since $\mathcal{V}' \subseteq \mathcal{V}$, we have established that

$$\mathcal{R}_\phi^\pi(\mathcal{V}, T) \ \geq\ \mathcal{R}_\phi^\pi(\mathcal{V}', T) \ \geq\ \frac{C^2}{4} \cdot T,$$

which concludes the proof. ■

*Correspondence: ob2105@columbia.edu, ygur@stanford.edu, assaf@gsb.columbia.edu.

Proof of Proposition 3. Fix $T \geq 1$, and $1 \leq V_T \leq T$. Let π be the OGD algorithm with $\eta_{t+1} = \eta$ for any $t = 1, \ldots, T-1$. Fix $\Delta_T \in \{1, \ldots, T\}$ (to be specified below), and define a partition of $\mathcal{T}$ into batches $\mathcal{T}_1, \ldots, \mathcal{T}_m$ of size Δ_T each (except perhaps $\mathcal{T}_m$) according to (8); this partition is only for analysis purposes. Fix $f \in \mathcal{V}$. By Flaxman et al. (2005) we have that (see analysis in their Lemma 3.1):[1]

$$\mathbb{E}^{\pi}\left[\sum_{t \in \mathcal{T}_j} f_t(X_t)\right] - \inf_{x \in \mathcal{X}}\left\{\sum_{t \in \mathcal{T}_j} f_t(x)\right\} \leq \frac{4r^2}{\eta} + \Delta_T \cdot \frac{\eta G^2}{2}, \tag{B-1}$$

for any $j = 1, \ldots, m$, where r is the radius of the set $\mathcal{X}$. Following the proof of Proposition 2, we have:

$$\begin{aligned}
\mathbb{E}^{\pi}\left[\sum_{t=1}^{T} f_t(X_t)\right] - \sum_{t=1}^{T} f_t(x_t^*) &\leq \sum_{j=1}^{m}\left(\mathbb{E}^{\pi}\left[\sum_{t \in \mathcal{T}_j} f_t(X_t)\right] - \inf_{x \in \mathcal{X}}\left\{\sum_{t \in \mathcal{T}_j} f_t(x)\right\}\right) + 2 \cdot \Delta_T V_T \\
&\overset{(a)}{\leq} \frac{8T}{\Delta_T} \cdot \frac{r^2}{\eta} + T\eta G^2 + 2 \cdot \Delta_T V_T,
\end{aligned}$$

for any $1 \leq \Delta_T \leq T$, where (a) follows (B-1). Taking $\Delta_T = \left\lceil (T/V_T)^{2/3} \right\rceil$ and $\eta = \frac{r}{G}\left(V_T/T\right)^{1/3}$ we get:

$$\mathbb{E}^{\pi}\left[\sum_{t=1}^{T} f_t(X_t)\right] - \sum_{t=1}^{T} f_t(x_t^*) \leq (9rG + 4) \cdot V^{1/3} T^{2/3}.$$

Since the above holds for any $f \in \mathcal{V}$, we established:

$$\mathcal{R}^{\pi}_{\phi^{(1)}}(\mathcal{V}, T) \leq (9rG + 4) \cdot V^{1/3} T^{2/3}.$$

This concludes the proof. ■

Proofs of Lemma A-1 and Lemma A-3. We start by proving Lemma A-1. Suppose that $\phi = \phi^{(1)}$. In the proof we use the notation defined in §4 and in the proof of Theorem 3. For any $t \in \mathcal{T}$ denote $Y_t = \phi^{(1)}(X_t, \cdot)$, and denote by $y_t \in \mathbb{R}^d$ the realized feedback observation at epoch t. For convenience, for any $t \geq 1$ we further denote $y^t = (y_1, \ldots, y_t)$. Fix $\pi \in \mathcal{P}_{\phi}$. Letting $u \in \mathcal{U}$, we denote $x_1 = \pi_1(u)$, and $x_t := \pi_t\left(y^{t-1}, u\right)$ for $t \in \{2, \ldots, T\}$. For any $f \in \mathcal{F}$ and $\tau \geq 2$, one has:

$$\begin{aligned}
d\mathbb{P}_f^{\pi,\tau}\left\{y^{\tau}, u\right\} &= d\mathbb{P}_f\left\{y^{\tau-1}, u\right\} d\mathbb{P}_f^{\pi,\tau-1}\left\{y^{\tau-1}, u\right\} \\
&\overset{(a)}{=} d\mathbb{P}_f\left\{y_{\tau} \left| x_{\tau}\right.\right\} d\mathbb{P}_f^{\pi,\tau-1}\left\{y^{\tau-1}, u\right\} \\
&\overset{(b)}{=} dG\left(y_{\tau} - \nabla f(x_{\tau})\right) d\mathbb{P}_f^{\pi,\tau-1}\left\{y^{\tau-1}, u\right\},
\end{aligned} \tag{B-2}$$

where: (a) holds since by the first part of Assumption 1 the feedback at epoch τ depends on the history only through $x_{\tau} = \pi_{\tau}\left(y^{\tau-1}, u\right)$; and (b) follows from the feedback structure given in the first part of Assumption 1. Fix $f, g \in \mathcal{F}$ and $\tau \geq 2$. One has:

[1] The expression adjusts the analysis in Flaxman et al. (2005) to allow an arbitrary (and apotentially random) X_0.

$$
\begin{aligned}
\mathcal{K}\left(\mathbb{P}_f^{\pi,\tau}\|\mathbb{P}_g^{\pi,\tau}\right) &= \int_{u,y^\tau}\log\left(\frac{d\mathbb{P}_f^{\pi,\tau}\left\{y^\tau,u\right\}}{d\mathbb{P}_g^{\pi,\tau}\left\{y^\tau,u\right\}}\right)d\mathbb{P}_f^{\pi,\tau}\left\{y^\tau,u\right\}\\
&\overset{(a)}{=} \int_{u,y^\tau}\log\left(\frac{dG\left(y_\tau-\nabla f(x_\tau)\right)d\mathbb{P}_f^{\pi,\tau-1}\left\{y^{\tau-1},u\right\}}{dG\left(y_\tau-\nabla g(x_\tau)\right)d\mathbb{P}_g^{\pi,\tau-1}\left\{y^{\tau-1},u\right\}}\right)dG\left(y_\tau-\nabla f(x_\tau)\right)d\mathbb{P}_f^{\pi,\tau-1}\left\{y^{\tau-1},u\right\}
\end{aligned}
$$

where (a) holds by (B-2). We have that $\mathcal{K}\left(\mathbb{P}_f^{\pi,\tau}\|\mathbb{P}_g^{\pi,\tau}\right)=A_\tau+B_\tau$, where:

$$
\begin{aligned}
A_\tau &:= \int_{u,y^\tau}\log\left(\frac{d\mathbb{P}_f^{\pi,\tau-1}\left\{y^{\tau-1},u\right\}}{d\mathbb{P}_g^{\pi,\tau-1}\left\{y^{\tau-1},u\right\}}\right)dG\left(y_\tau-\nabla f(x_\tau)\right)d\mathbb{P}_f^{\pi,\tau-1}\left\{y^{\tau-1},u\right\}\\
&= \int_{u,y^{\tau-1}}\log\left(\frac{d\mathbb{P}_f^{\pi,\tau-1}\left\{y^{\tau-1},u\right\}}{d\mathbb{P}_g^{\pi,\tau-1}\left\{y^{\tau-1},u\right\}}\right)\left[\int_{y_\tau}dG\left(y_\tau-\nabla f(x_\tau)\right)\right]d\mathbb{P}_f^{\pi,\tau-1}\left\{y^{\tau-1},u\right\}\\
&= \int_{u,y^{\tau-1}}\log\left(\frac{d\mathbb{P}_f^{\pi,\tau-1}\left\{y^{\tau-1},u\right\}}{d\mathbb{P}_g^{\pi,\tau-1}\left\{y^{\tau-1},u\right\}}\right)d\mathbb{P}_f^{\pi,\tau-1}\left\{y^{\tau-1},u\right\} \;=\; \mathcal{K}\left(\mathbb{P}_f^{\pi,\tau-1}\|\mathbb{P}_g^{\pi,\tau-1}\right),
\end{aligned}
$$

and

$$
\begin{aligned}
B_\tau &:= \int_{u,y^\tau}\log\left(\frac{dG\left(y_\tau-\nabla f(x_\tau)\right)}{dG\left(y_\tau-\nabla g(x_\tau)\right)}\right)dG\left(y_\tau-\nabla f(x_\tau)\right)d\mathbb{P}_f^{\pi,\tau-1}\left\{y^{\tau-1},u\right\}\\
&= \int_{u,y^{\tau-1}}\int_{y_\tau}\left[\log\left(\frac{dG\left(y_\tau-\nabla f(x_\tau)\right)}{dG\left(y_\tau-\nabla g(x_\tau)\right)}\right)dG\left(y_\tau-\nabla f(x_\tau)\right)\right]d\mathbb{P}_f^{\pi,\tau-1}\left\{y^{\tau-1},u\right\}\\
&\overset{(b)}{\leq} \tilde{C}\int_{u,y^{\tau-1}}\left\|\nabla f_\tau(x_\tau)-g_\tau(x_\tau)\right\|^2 d\mathbb{P}_f^{\pi,\tau-1}\left\{y^{\tau-1},u\right\} \;=\; \tilde{C}\mathbb{E}_f^\pi\left\|\nabla f_\tau(x_\tau)-g_\tau(x_\tau)\right\|^2,
\end{aligned}
$$

where (b) follows the second part of Assumption 1. Repeating the above arguments, one has:

$$
\mathcal{K}\left(\mathbb{P}_f^{\pi,\tau}\|\mathbb{P}_g^{\pi,\tau}\right) \;\leq\; \mathcal{K}\left(\mathbb{P}_f^{\pi,1}\|\mathbb{P}_g^{\pi,1}\right)+\tilde{C}\mathbb{E}_f^\pi\left[\sum_{t=2}^{\tau}\left\|\nabla f_t(x_t)-g_t(x_t)\right\|^2\right].
$$

From the above it is also clear that:

$$
\begin{aligned}
\mathcal{K}\left(\mathbb{P}_f^{\pi,1}\|\mathbb{P}_g^{\pi,1}\right) &= \int_{u,y_1}\log\left(\frac{d\mathbb{P}_f^{\pi,1}\left\{y_1,u\right\}}{d\mathbb{P}_g^{\pi,1}\left\{y_1,u\right\}}\right)d\mathbb{P}_f^{\pi,1}\left\{y_1,u\right\}\\
&= \int_u\left[\int_{y_1}\log\left(\frac{dG\left(y_1-\nabla f(x_1)\right)}{dG\left(y_1-\nabla g(x_1)\right)}\right)dG\left(y_1-\nabla f(x_1)\right)\right]d\mathbf{P}_u\left\{u\right\}\\
&\leq \tilde{C}\int_u\left\|\nabla f_1(x_1)-\nabla g_1(x_1)\right\|^2 d\mathbf{P}_u\left\{u\right\} \;=\; \tilde{C}\mathbb{E}_f^\pi\left\|\nabla f_1(x_1)-\nabla g_1(x_1)\right\|^2.
\end{aligned}
$$

Hence, we have established that for any $\tau\geq 1$:

$$
\mathcal{K}\left(\mathbb{P}_f^{\pi,\tau}\|\mathbb{P}_g^{\pi,\tau}\right) \;\leq\; \tilde{C}\sum_{t=1}^{\tau}\mathbb{E}_f^\pi\left\|\nabla f_t(x_t)-g_t(x_t)\right\|^2.
$$

Finally, following the steps above, the proof of Lemma A-3 (for the feedback structure $\phi = \phi^{(0)}$) is immediate, using the notation introduced in the proof of Theorem 5 for cost feedback structure, along with Assumption 2. This concludes the proof. ■

Performance analysis of OGD algorithm without restarting. We consider the performance of the OGD algorithm *without restarting*, relative to the dynamic benchmark. The following illustrates that this algorithm will yield linear regret for a broad set of variation budgets.

Example 1 (Failure of OGD without restarting) Consider a partition of the horizon $\mathcal{T}$ into batches $\mathcal{T}_1, \dots \mathcal{T}_m$ according to (8), with each batch of size Δ_T. Consider the following cost functions:

$$g_1(x) = (x-\alpha)^2, \qquad g_2(x) = x^2; \qquad x \in [-1,3].$$

Assume that nature selects the cost function to be $g_1(\cdot)$ in the even batches and $g_2(\cdot)$ in the odd batches. Assume that at every epoch t, after selecting an action $x_t \in \mathcal{X}$, a *noiseless* access to the gradient of the cost function at point x_t is granted, that is, $\phi_t^{(1)}(x, f_t) = f_t'(x)$ for all $x \in \mathcal{X}$ and $t \in \mathcal{T}$. Assume that the decision maker is applying the OGD algorithm with a sequence of step sizes $\{\eta_t\}_{t=2}^T$, and $x_1 = 1$. We consider two classes of step size sequences that have been shown to be rate optimal in two instances of OCO settings (see Flaxman et al. (2005), and Hazan et al. (2007)).

1. Suppose $\eta_t = \eta = C/\sqrt{T}$. Then, selecting a batch size Δ_T of order $\sqrt{T}$, and $\alpha = 1 + (1+2\eta)^{\Delta_T}$, the variation budget V_T is at most of order $\sqrt{T}$, and there is a constant C_1 such that $\mathcal{R}_\phi^\pi(\mathcal{V}, T) \ge C_1 T$.
2. Suppose that $\eta_t = C/t$. Then, selecting a batch size Δ_T of order T, and $\alpha = 1$, the variation budget V_T is a fixed constant, and there is a constant C_2 such that $\mathcal{R}_\phi^\pi(\mathcal{V}, T) \ge C_2 T$. ■

Proof of claims made in Example 1. Fix $T \ge 1$. Let $\mathcal{X} = [-1,3]$ (we assume that ν, appearing in (2), is smaller than 1) and consider the following two functions: $g^1(x) = (x-\alpha)^2$, and $g^2(x) = x^2$. We assume that in each epoch t, after selecting an action x_t, there is a noiseless access to the gradient of the cost function, evaluated at point x_t. The the deterministic actions are generated by an OGD algorithm:

$$x_{t+1} = P_{\mathcal{X}}\left(x_t - \eta_{t+1} \cdot f_t'(x_t)\right), \qquad \text{for all } t \ge 1,$$

with the initial selection $x_1 = 1$. In the first part we consider the case of $\eta_t = \eta = C/\sqrt{T}$, and in a second part we consider the case of $\eta_t = C/t$. The structure of both parts is similar: first we analyze the variation of the instance, showing it is sublinear. Then, by analyzing the sequence of decisions $\{x_t\}_{t=1}^T$ that is generated by the Online Gradient Descent policy, we show that in a linear portion of the horizon there is a constant C_2 such that $|x_t - x_t^*| > C_2$, and therefore a linear regret is incurred.

Part 1. Assume that $\eta_t = \eta = C/\sqrt{T} \le 1/2$. Select $\Delta_T = \left\lfloor 1 + \frac{1}{2\eta} \right\rfloor$, and set $\alpha = 1 + (1-2\eta)^{\Delta_T}$ (note that $1 \le \alpha \le 2$). We assume that nature selects the cost function to be $g_1(\cdot)$ in the even batches and $g_2(\cdot)$ in the odd batches. We start by analyzing the variation along the horizon:

$$
\begin{aligned}
\sum_{t=2}^{T} \sup_{x\in\mathcal{X}} |f_t(x) - f_{t-1}(x)| &\le \left(\left\lceil \frac{T}{\Delta_T} \right\rceil - 1 \right) \cdot \sup_{x\in\mathcal{X}} |g_2(x) - g_1(x)| \\
&\le \frac{T}{\Delta_T} \cdot \sup_{x\in\mathcal{X}} \left|\alpha^2 - 2\alpha x\right| \\
&\overset{(a)}{\le} \frac{8T}{\Delta_T} = \frac{8T}{\left\lfloor 2 + \frac{1}{2\eta} \right\rfloor} \\
&\le 16T\eta = 16C \cdot \sqrt{T},
\end{aligned}
$$

where (a) follows from $1 \le \alpha \le 2$ and $-1 \le x \le 3$. Next, we analyze the incurred regret. We start by analyzing decisions generated by the OGD algorithm throughout the first two batches. Recalling that $x_1 = 1$ and that $g_2(\cdot)$ is the cost function throughout the first batch, one has for any $2 \le t \le \Delta_T + 1$:

$$
\begin{aligned}
x_t &= x_{t-1} - \eta \cdot f'(x_{t-1}) \\
&= x_{t-1} - \eta \cdot 2x_{t-1} = x_{t-1}(1 - 2\eta) \\
&= x_1 (1 - 2\eta)^{t-1} = (1 - 2\eta)^{t-1} \\
&= \exp\{(t-1)\ln(1 - 2\eta)\} \\
&\overset{(a)}{\ge} \exp\left\{(t-1)(-2\eta - 2\eta^2)\right\} \\
&\overset{(b)}{\ge} \exp\{-1 - \eta\} \\
&\overset{(c)}{>} \frac{1}{e^2},
\end{aligned}
$$

where: (a) follows since for any $-1 < x \le 1$ one has $\ln(1+x) \ge x - \frac{x^2}{2}$; (b) follows from $t \le \Delta_T \le 1 + \frac{1}{2\eta}$; and (c) follows from $\eta \le \frac{1}{2} < 1$. Since $x_t^* = 0$ for any $1 \le t \le \Delta_T$, one has:

$$
x_t - x_t^* > \frac{1}{e^2},
$$

for any $1 \le t \le \Delta_T$. At the end of the first batch the cost function changes from $f(\cdot)$ to $g(\cdot)$. Note that the first action of the second batch is $x_{\Delta_T+1} = (1 - 2\eta)^{\Delta_T}$. Since $g_1(\cdot)$ is the cost function throughout the second batch, for any $\Delta_T + 2 \le t \le 2\Delta_T + 1$ one has:

$$
\begin{aligned}
x_t &= x_{t-1} - \eta \cdot g'(x_{t-1}) \\
&= x_{t-1} - \eta \cdot 2\left(x_{t-1} - \alpha\right).
\end{aligned}
$$

Using the transformation $y_t = x_t - \alpha$ for all t, one has:

$$
\begin{aligned}
y_t &= y_{t-1} - \eta \cdot 2y_{t-1} \;=\; y_{t-1}(1-2\eta) \\
&= y_{\Delta_T+1}(1-2\eta)^{t-\Delta_T-1} \\
&= x_{\Delta_T+1}(1-2\eta)^{t-\Delta_T-1} - \alpha(1-2\eta)^{t-\Delta_T-1} \\
&= (1-2\eta)^{t-1} - (1-2\eta)^{t-\Delta_T-1} - (1-2\eta)^{t-1} \\
&= -(1-2\eta)^{t-\Delta_T-1} \\
&= -\exp\{(t-\Delta_T-1)\ln(1-2\eta)\} \\
&\overset{(a)}{\leq} -\exp\left\{(t-\Delta_T-1)(-2\eta-2\eta^2)\right\} \\
&\overset{(b)}{\leq} -\exp\{-1-\eta\} \\
&\overset{(c)}{<} -\frac{1}{e^2},
\end{aligned}
$$

where: (a) holds since for any $-1 < x \leq 1$ one has $\ln(1+x) \geq x - \frac{x^2}{2}$; (b) follows from $t \leq 2\Delta_T \leq 1 + \frac{1}{2\eta} + \Delta_T$; and (c) follows from $\eta \leq \frac{1}{2} < 1$. Finally, recalling that $x_t^* = \alpha$ and using the transformation $y_t = x_t - \alpha$, one has for any $\Delta_T + 1 \leq t \leq 2\Delta_T$:

$$
x_t^* - x_t \;=\; y_t \;<\; -\frac{1}{e^2}.
$$

In the beginning of the third batch $g_2(\cdot)$ becomes the cost function once again. We note that the first action of the third batch is the same as the first action of the first batch:

$$
x_{2\Delta_T+1} \;=\; \alpha + y_{2\Delta_T+1} \;=\; \alpha - (1-2\eta)^{2\Delta_T+1-\Delta_T-1} \;=\; \alpha - (1-2\eta)^{\Delta_T} \;=\; 1 \;=\; x_1,
$$

and therefore the actions taken in the first two batches are repeated throughout the horizon. We conclude that for any $1 \leq t \leq T$,

$$
|x_t - x_t^*| \;>\; \frac{1}{e^2}.
$$

Finally, we calculate the regret incurred throughout the horizon. Using Taylor expansion, one has

$$
\sum_{t=1}^{T}(f_t(x_t) - f_t(x_t^*)) \;=\; \sum_{t=1}^{T}(x_t - x_t^*)^2 \;>\; \sum_{t=1}^{T}\frac{1}{e^4} \;=\; \frac{T}{e^4}.
$$

Part 2. For concreteness we assume in this part that T is even and larger than 2. We show that linear regret can be incurred when $\eta_t = \frac{C}{t}$. Set $\alpha = 1$ and $\Delta_T = T/2$ (therefore we have two batches). Assume that nature selects $g_1(\cdot)$ to be the cost function in the first batch, $g_2(\cdot)$ to be the cost function in the second batch. We start by analyzing the variation along the horizon. Recalling that there is only one change in the cost function, one has:

$$
\begin{aligned}
\sum_{t=2}^{T}\sup_{x\in\mathcal{X}}|f_t(x) - f_{t-1}(x)| &= \sup_{x\in\mathcal{X}}|g_2(x) - g_1(x)| \\
&= \sup_{x\in\mathcal{X}}\left|\alpha^2 - 2\alpha x\right| \;=\; \sup_{x\in\mathcal{X}}|1-2x| \;\overset{(a)}{=}\; 5,
\end{aligned}
$$

where (a) holds because $-1 \leq x \leq 3$. Since $x_1 = 1$, and $g_1'(1) = 0$, one obtains $x_t = 1$ for all

$1 \le t \le \lceil \frac{T}{2} \rceil + 1$. After $\lceil T/2 \rceil$ epochs, the cost function changes from $g_1(\cdot)$ to $g_2(\cdot)$, and for all $\lceil \frac{T}{2} \rceil + 2 \le t \le T$ one has:

$$\begin{aligned}
x_t &= x_{t-1} - \eta_t \cdot g_2'(x_{t-1}) \\
&= x_{t-1} - \eta_t \cdot 2x_{t-1} \; = \; x_{t-1}\left(1 - 2\eta_t\right) \\
&= x_{\frac{T}{2}+1} \prod_{t'=\frac{T}{2}+1}^{t} \left(1 - 2\eta_{t'}\right) \; = \; \prod_{t'=\frac{T}{2}+1}^{t} \left(1 - 2\eta_{t'}\right) \\
&\overset{(a)}{\ge} \left(1 - 2\eta_{\frac{T}{2}+2}\right)^{t-\frac{T}{2}-1} \; = \; \left(1 - \frac{4C}{T+4}\right)^{t-\frac{T}{2}-1} \\
&= \exp\left\{\left(t - \frac{T}{2} - 1\right) \ln\left(1 - \frac{4C}{T+4}\right)\right\} \\
&\overset{(b)}{\ge} \exp\left\{\left(t - \frac{T}{2} - 1\right)\left(-\frac{4C}{T+4} - \frac{8C}{(T+4)^2}\right)\right\} \\
&\overset{(c)}{\ge} \exp\left\{-4C - \frac{8C^2}{T+4}\right\} \; > \; \exp\left\{-4C - 2C^2\right\},
\end{aligned}$$

where: (a) holds since $\{\eta_t\}$ is a decreasing sequence; (b) holds since $\ln(1+x) \ge x - \frac{x^2}{2}$ for any $-1 < x \le 1$; and (c) is obtained using $t < T + \frac{T}{2} + 5$. Since $x_t^* = 0$ for any $\frac{T}{2} + 1 \le t \le T$, one has:

$$x_t - x_t^* \; > \; \frac{1}{e^{2C(2+C)}},$$

for all $\frac{T}{2} + 1 \le t \le T$. Finally, we calculate the regret incurred throughout the horizon. Recalling that throughout the first batch no regret is incurred, and using Taylor expansion, one has:

$$\sum_{t=1}^{T} \left(f_t(x_t) - f_t(x_t^*)\right) \; = \; \sum_{t=\frac{T}{2}+1}^{T} \left(f(x_t) - f(x_t^*)\right) \; = \; \sum_{t=\frac{T}{2}+1}^{T} \left(x_t - x_t^*\right)^2 \; \ge \; \sum_{t=\frac{T}{2}+1}^{T} \frac{1}{e^{4C(2+C)}} \; = \; \frac{T}{2e^{4C(2+C)}}.$$

This concludes the proof. ■

C Auxiliary results for OCO settings

C.1 Preliminaries

In this section we develop auxiliary results that provide bounds on the regret with respect to the single best action in the adversarial setting. As discussed in §1, the OCO literature most often considers few different feedback structures; typical examples include full access to the cost/gradient after the action X_t is selected, as well as a noiseless access to the cost/gradient evaluated at X_t. However, in this section we consider the feedback structures $\phi^{(0)}$ and $\phi^{(1)}$, where noisy access to the cost/gradient is granted.

We define admissible online algorithms exactly as admissible policies are defined in §2.[2] More precisely, letting U be a random variable defined over a probability space $(\mathbb{U}, \mathcal{U}, \mathbf{P}_u)$, we let $\mathcal{A}_1 : \mathbb{U} \to \mathbb{R}^d$

[2]We use the different terminology and notation only to highlight the different objectives: a policy π is designed to minimize regret with respect to the dynamic oracle, while an online algorithm $\mathcal{A}$ is designed to minimize regret compared to the static single best action benchmark.

and $\mathcal{A}_t : \mathbb{R}^{(t-1)k} \times \mathbb{U} \to \mathbb{R}^d$ for $t = 2, 3, \ldots$ be measurable functions, such that X_t, the action at time t, is given by

$$X_t = \begin{cases} \mathcal{A}_1(U) & t = 1, \\ \mathcal{A}_t(\phi_{t-1}(X_{t-1}, f_{t-1}), \ldots, \phi_1(X_1, f_1), U) & t = 2, 3, \ldots, \end{cases}$$

where $k = 1$ if $\phi = \phi^{(0)}$, and $k = d$ if $\phi = \phi^{(1)}$. The mappings $\{\mathcal{A}_t : t = 1, \ldots, T\}$ together with the distribution $\mathbf{P}_u$ define the class of admissible online algorithms with respect to feedback ϕ, which is exactly the class $\mathcal{P}_\phi$. The filtration $\{\mathcal{H}_t,\ t = 1, \ldots, T\}$ is defined exactly as in §2. Given a feedback structure $\phi \in \{\phi^{(0)}, \phi^{(1)}\}$, the objective is to minimize the regret compared to the single best action:

$$\mathcal{G}^{\mathcal{A}}_\phi(\mathcal{F}, T) = \sup_{f \in \mathcal{F}} \left\{ \mathbb{E}^{\mathcal{A}} \left[\sum_{t=1}^T f_t(X_t) \right] - \min_{x \in \mathcal{X}} \left\{ \sum_{t=1}^T f_t(x) \right\} \right\}.$$

We note that while most results in the OCO literature allow sequences that can adjust the cost function adversarially at each epoch, we consider the above setting where nature commits to a sequence of functions in advance. This, along with the setting of noisy cost/gradient observations, is done for the sake of consistency with the non-stationary stochastic framework we propose in this paper.

C.2 Upper bounds

The first two results of this section, Lemma 1 and Lemma 2, analyze the performance of the EGS algorithm (given in §5) under structure $(\mathcal{F}_s, \phi^{(0)})$ and the OGD algorithm (given in §4) under structure $(\mathcal{F}_s, \phi^{(1)})$, respectively. To the best of our knowledge, the upper bound in Lemma 1 is not documented in the Online Convex Optimization literature[3]. Lemma 2 adapts Theorem 1 in Hazan et al. (2007) (that considered noiseless access to the gradient) to the feedback structure $\phi^{(1)}$.

Lemma 1 (Performance of EGS in the adversarial setting) *Consider the feedback structure $\phi = \phi^{(0)}$. Let $\mathcal{A}$ be the EGS algorithm given in §5.2, with $a_t = 2d/Ht$ and $\delta_t = h_t = a_t^{1/4}$ for all $t \in \{1, \ldots, T-1\}$. Then, there exists a constant $\bar{C}$, independent of T such that for any $T \geq 1$,*

$$\mathcal{G}^{\mathcal{A}}_\phi(\mathcal{F}_s, T) \ \leq\ \bar{C}\sqrt{T}.$$

Proof. Let $\phi = \phi^{(0)}$. Fix $T \geq 1$ and $f \in \mathcal{F}_s$. Let $\mathcal{A}$ be the EGS algorithm, with the selection $a_t = 2d/Ht$ and $\delta_t = h_t = a_t^{1/4}$ for all $t \in \{1, \ldots, T-1\}$. We assume that $\delta_t \leq \nu$ for all $t \in \mathcal{T}$; in the end of the proof we discuss the case in which the former does not hold. For the sequence $\{\delta_t\}_{t=1}^T$, we denote by $\mathcal{X}_{\delta_t}$ the δ_t-interior of the action set $\mathcal{X}$: $\mathcal{X}_{\delta_t} = \{x \in \mathcal{X} \mid \mathbf{B}_{\delta_t}(x) \subseteq \mathcal{X}\}$. We have for all $f_t \in \mathcal{F}_s$:

$$\mathbb{E}\left[\phi_t^{(0)}(X_t, f_t)\ |X_t = x\right] \ =\ f_t(x) \quad \text{and} \quad \sup_{x \in \mathcal{X}} \left\{ \mathbb{E}\left[\left(\phi_t^{(0)}(x, f_t) \right)^2 \right] \right\} \ \leq\ G^2 + \sigma^2, \tag{C-3}$$

for some $\sigma \geq 0$. At any $t \in \mathcal{T}$ the gradient estimator is:

$$\hat{\nabla}_{h_t} f_t(X_t) \ =\ \frac{\phi_t^{(0)}(X_t + h_t\psi_t, f_t)\,\psi_t}{h_t},$$

[3]The feasibility of an upper bound of order $\sqrt{T}$ on the regret in an adversarial setting with noisy access to the cost and with strictly convex cost functions was suggested by Agarwal et al. (2010) without further details or proof.

for a fixed $h_t > 0$, and where $\{\psi_t\}$ is a sequence of iid random variables, drawn uniformly over the set $\left\{\pm e^{(1)}, \ldots, \pm e^{(d)}\right\}$, where $e^{(k)}$ denotes the unit vector with 1 at the k^{th} coordinate. In particular, we denote $\psi_t = Y_t W_t$, where Y_t and W_t are independent random variables, $\mathbb{P}\{y_t = 1\} = \mathbb{P}\{y_t = -1\} = 1/2$, and $W_t = e^{(k)}$ with probability $1/d$ for all $k \in \{1, \ldots, d\}$. The estimated gradient step is

$$Z_{t+1} \;=\; P_{\mathcal{X}_{\delta_t}}\left(Z_t - a_t \hat{\nabla}_{h_t} f_t(Z_t)\right), \qquad X_{t+1} \;=\; Z_{t+1} + h_{t+1}\psi_t,$$

where $P_{\mathcal{X}_{\delta_t}}$ denotes the Euclidean projection operator over the set $\mathcal{X}_{\delta_t}$. Note that $Z_t \in \mathcal{X}$, $X_t \in \mathcal{X}$, and $X_t + h_t\psi_t \in \mathcal{X}$ for all $t \in \mathcal{T}$. Since $\|\psi_t\| = 1$ for all $t \in \mathcal{T}$, one has:

$$\mathbb{E}\left[\left\|\hat{\nabla}_{h_t} f_t(Z_t)\right\|^2 \; |Z_t = z\right] \;=\; \frac{\mathbb{E}\left[\left(\phi_t^{(0)}\left(z + h_t\psi_t, f_t\right)\right)^2\right]}{h_t^2} \;\leq\; \frac{G^2 + \sigma^2}{h_t^2} \qquad \text{for all } z \in \mathcal{X}, \tag{C-4}$$

using (C-3). Then,

$$\mathbb{E}\left[\hat{\nabla}_{h_t} f_t(Z_t) \; |Z_t = z, \psi_t = \psi\right] \;=\; \frac{\mathbb{E}\left[\phi_t^{(0)}\left(Z_t + h_t\psi_t, f_t\right)\psi_t \; |Z_t = z, \psi_t = \psi\right]}{h_t} \;=\; \frac{f_t\left(z + h_t\psi\right)\psi}{h_t}.$$

Therefore, taking expectation with respect to ψ, one has

$$\begin{aligned}
\mathbb{E}\left[\hat{\nabla}_{h_t} f_t(Z_t) \; |Z_t = z\right] \;&=\; \mathbb{E}_{Y,W}\left[\frac{f_t\left(z + h_t\psi\right)\psi}{h_t}\right] \;=\; \frac{1}{d}\sum_{k=1}^{d} \frac{\left(f_t(z + h_t e^{(k)}) - f_t(z - h_t e^{(k)})\right)e^{(k)}}{2h_t} \\
&\overset{(a)}{\geq}\; \frac{1}{d}\sum_{k=1}^{d}\left(\nabla f_t(z - h_t e^{(k)}) \cdot e^{(k)}\right)e^{(k)} \\
&\overset{(b)}{\geq}\; \frac{1}{d}\sum_{k=1}^{d}\left(\nabla f_t(z) \cdot e^{(k)} - G h_t\right)e^{(k)} \;=\; \frac{1}{d}\nabla f_t(z) - \frac{G h_t}{d} \cdot \bar{e},
\end{aligned}$$

where $\bar{e}$ denotes a vector of ones. The equalities and inequalities above hold componentwise, where (a) follows from a Taylor expansion and the convexity of f_t: $f_t(z + h_t e^{(k)}) - f_t(z - h_t e^{(k)}) \geq \nabla f_t(z - h_t e^{(k)}) \cdot \left(2h_t e^{(k)}\right)$, for any $1 \leq k \leq d$, and (b) follows from a Taylor expansion, the convexity of f_t, and (10):

$$\nabla f_t(z - h_t e^{(k)}) \cdot e^{(k)} \;\geq\; \nabla f_t(z) \cdot e^{(k)} - \left(h e^{(k)}\right) \cdot \left(\nabla^2 f_t\right) e^{(k)} \;\geq\; \nabla f_t(z) \cdot e^{(k)} - G h_t,$$

for any $1 \leq k \leq d$. Therefore, for all $z \in \mathcal{X}$ and for all $t \in \mathcal{T}$:

$$\left\|\frac{1}{d}\nabla f_t(z) - \mathbb{E}\left[\hat{\nabla}_{h_t} f_t(Z_t) \; |Z_t = z\right]\right\| \;\leq\; \frac{G h_t}{\sqrt{d}}. \tag{C-5}$$

Define x^* as the single best action: $x^* = \arg\min_{x \in \mathcal{X}}\left\{\sum_{t=1}^{T} f_t(x)\right\}$. Then, for any $t \in \mathcal{T}$, one has

$$f_t(x^*) \;\geq\; f_t(Z_t) + \nabla f_t(Z_t) \cdot (x^* - Z_t) + \frac{1}{2}H\left\|x^* - Z_t\right\|^2,$$

and hence:

$$f_t(Z_t) - f_t(x^*) \;\leq\; \nabla f_t(Z_t) \cdot (Z_t - x^*) - \frac{1}{2}H\left\|Z_t - x^*\right\|^2. \tag{C-6}$$

Next, using the estimated gradient step, one has

$$
\begin{aligned}
\|Z_{t+1}-x^*\|^2 &= \left\| P_{\mathcal{X}_{\delta_t}}\left(Z_t - a_t\hat{\nabla}_{h_t} f_t(Z_t)\right) - x^* \right\|^2 \\
&\overset{(a)}{\leq} \left\| Z_t - a_t\hat{\nabla}_{h_t} f_t(Z_t) - x^* \right\|^2 \\
&= \|Z_t - x^*\|^2 - 2a_t\left(Z_t - x^*\right)\cdot\hat{\nabla}_{h_t} f_t(Z_t) + a_t^2\left\|\hat{\nabla}_{h_t} f_t(Z_t)\right\|^2 \\
&= \|Z_t - x^*\|^2 - \frac{2a_t}{d}\cdot\left(Z_t - x^*\right)\cdot\nabla f_t(Z_t) + a_t^2\left\|\hat{\nabla}_{h_t} f_t(Z_t)\right\|^2 \\
&\quad + 2a_t\left(Z_t - x^*\right)\cdot\left(\frac{1}{d}\nabla f_t(Z_t) - \hat{\nabla}_{h_t} f_t(Z_t)\right) \\
&\leq \|Z_t - x^*\|^2 - \frac{2a_t}{d}\cdot\left(Z_t - x^*\right)\cdot\nabla f_t(Z_t) + a_t^2\left\|\hat{\nabla}_{h_t} f_t(Z_t)\right\|^2 \\
&\quad + 2a_t\|Z_t - x^*\|\cdot\left\|\frac{1}{d}\nabla f_t(Z_t) - \hat{\nabla}_{h_t} f_t(Z_t)\right\|,
\end{aligned}
$$

where (a) follows from a standard contraction property of the Euclidean projection operator. Taking expectation with respect to ψ_t and conditioning on Z_t, we follow (C-4) and (C-5) to obtain

$$
\mathbb{E}\left[\|Z_{t+1}-x^*\|^2\ |Z_t\right] \leq \|Z_t - x^*\|^2 - \frac{2a_t}{d}\cdot\left(Z_t - x^*\right)\cdot\nabla f_t(Z_t) + \frac{a_t^2\left(G^2+\sigma^2\right)}{h_t^2} + \frac{2Ga_th_t}{\sqrt{d}}\cdot\|Z_t - x^*\|.
$$

Taking another expectation, with respect to Z_t, we get

$$
\mathbb{E}\left[\|Z_{t+1}-x^*\|^2\right] \leq \mathbb{E}\left[\|Z_t - x^*\|^2\right] - \frac{2a_t}{d}\cdot\mathbb{E}\left[\left(Z_t - x^*\right)\cdot\nabla f_t(Z_t)\right] + \frac{a_t^2\left(G^2+\sigma^2\right)}{h_t^2} + \frac{2Ga_th_t}{\sqrt{d}}\cdot\mathbb{E}\|Z_t - x^*\|,
$$

and therefore, fixing some $\gamma > 0$, we have for all $t \in \{1, \ldots, T-1\}$:

$$
\begin{aligned}
\mathbb{E}\left[\left(Z_t - x^*\right)\cdot\nabla f_t(Z_t)\right] &\leq \frac{d}{2a_t}\left(\mathbb{E}\left[\|Z_t - x^*\|^2\right] - \mathbb{E}\left[\|Z_{t+1}-x^*\|^2\right]\right) + \frac{\left(G^2+\sigma^2\right)a_td}{2h_t^2} \\
&\quad + \gamma\cdot\frac{1}{\gamma}\cdot Gh_t\sqrt{d}\cdot\mathbb{E}\|Z_t - x^*\| \\
&\overset{(a)}{\leq} \frac{d}{2a_t}\left(\mathbb{E}\left[\|Z_t - x^*\|^2\right] - \mathbb{E}\left[\|Z_{t+1}-x^*\|^2\right]\right) + \frac{\left(G^2+\sigma^2\right)a_td}{2h_t^2} \\
&\quad + \frac{\gamma^2}{2}\cdot\mathbb{E}\left[\|Z_t - x^*\|^2\right] + \frac{G^2h_t^2d}{2\gamma^2},
\end{aligned}
\tag{C-7}
$$

where (a) holds by $ab \leq \left(a^2+b^2\right)/2$, and by Jensen's inequality. In addition, one has for any $t \in \mathcal{T}$:

$$
\begin{aligned}
\mathbb{E}\left[f_t(X_t)\right] &= \mathbb{E}\left[\mathbb{E}\left[f_t(X_t)|Z_t\right]\right] = \mathbb{E}\left[\frac{1}{2}\left(f_t(Z_t+h_t) + f_t(Z_t-h_t)\right)\right] \\
&\leq \frac{1}{2}\mathbb{E}\left[2f_t(Z_t) + h_t\left(\nabla f_t(Z_t+h_t) - \nabla f_t(Z_t-h_t)\right) - Hh_t^2\right] \\
&\leq \mathbb{E}\left[f_t(Z_t) + \frac{1}{2}Hh_t^2\right].
\end{aligned}
\tag{C-8}
$$

The regret with respect to the single best action is:

$$\begin{aligned}
\sum_{t=1}^{T} \mathbb{E}^{\mathcal{A}}\left[f_t(X_t) - f_t(x^*)\right] &\le 2G + \sum_{t=1}^{T-1} \mathbb{E}^{\pi}\left[f_t(Z_t) - f_t(x^*) + \frac{1}{2} H h_t^2\right] \\
&\overset{(a)}{\le} 2G + \sum_{t=1}^{T-1} \mathbb{E}\left[\nabla f_t(Z_t)\cdot(Z_t - x^*) - \frac{1}{2} H \left\|Z_t - x^*\right\|^2 + \frac{1}{2} H h_t^2\right] \\
&\overset{(b)}{\le} \mathbb{E}\left[\sum_{t=1}^{T-1}\left(\frac{d}{2a_t}\left(\left\|Z_t - x^*\right\|^2 - \left\|Z_{t+1} - x^*\right\|^2\right) + \frac{\left(\gamma^2 - H\right)}{2}\cdot\left\|Z_t - x^*\right\|^2\right)\right] \\
&\quad + 2G + \frac{\left(G^2+\sigma^2\right)}{2}\sum_{t=1}^{T-1}\left(\frac{a_t d}{h_t^2} + \frac{h_t^2 d}{\gamma^2}\right) + \frac{H}{2}\sum_{t=1}^{T-1} h_t^2 \\
&\overset{(c)}{=} \frac{1}{2}\sum_{t=2}^{T}\mathbb{E}\left[\left\|Z_t - x^*\right\|^2\right]\underbrace{\left(\frac{d}{a_t} - \frac{d}{a_{t-1}} + \left(\gamma^2 - H\right)\right)}_{I_t} + \mathbb{E}\left[\left\|Z_1 - x^*\right\|^2\right]\underbrace{\left(\frac{d}{2a_1} + \frac{\gamma^2 - H}{2}\right)}_{I_1} \\
&\quad - \mathbb{E}\left[\left\|Z_T - x^*\right\|^2\right]\frac{d}{2a_{T-1}} + 2G + \frac{\left(G^2+\sigma^2\right)}{2}\sum_{t=1}^{T-1}\left(\frac{a_t d}{h_t^2} + \frac{h_t^2 d}{\gamma^2}\right) + \frac{H}{2}\sum_{t=1}^{T-1} h_t^2,
\end{aligned}$$

where (a) holds by (C-6), (b) holds by (C-7), and (c) holds by rearranging the summation. By selecting $\gamma^2 = \frac{H}{2}$, $a_t = \frac{d}{(H-\gamma^2)t}$, and $h_t = \delta_t = a_t^{1/4}$, we have $I_t = 0$ for all $t \in \mathcal{T}$, and:

$$\mathbb{E}^{\mathcal{A}}\left[\sum_{t=1}^{T} f_t(X_t)\right] - \inf_{x\in\mathcal{X}}\left\{\sum_{t=1}^{T} f_t(x)\right\} \le 2G + \frac{\left(G^2+\sigma^2+H\right)d^{3/2}}{\sqrt{2H}}\cdot\sqrt{T}.$$

Since the above holds for any $f \in \mathcal{F}_s$, we conclude that

$$\mathcal{G}^{\mathcal{A}}_{\phi^{(0)}}\left(\mathcal{F}_s, T\right) \le 2G + \frac{\left(G^2+\sigma^2+H\right)d^{3/2}}{\sqrt{2H}}\cdot\sqrt{T}.$$

Finally, we consider the case in which there exists at least one time epoch t such that $\delta_t > \nu$. Then, for any such time epoch we select $h_t' = \delta_t' = \min\{\nu, \delta_t\}$. We note that the sequence $\{\delta_t\}$ is converging to 0, and therefore for any number ν there is some epoch t_ν, independent of T, such that $\delta_t \le \nu$ for any $t \ge t_\nu$. Therefore there can be no more than t_ν such epochs. In particular, it follows that such a case could add to the regret above no more than a constant (independent of T), that depends solely on ν, the dimension d, and the second derivative bound H. This concludes the proof. ■

Lemma 2 (Performance of OGD in the adversarial setting) *Consider the feedback structure $\phi = \phi^{(1)}$. Let $\mathcal{A}$ be the OGD algorithm given in §4, with the selection $\eta_{t+1} = 1/Ht$ for $t = 1, \dots T-1$. Then, there exists a constant $\bar{C}$, independent of T such that for any $T \ge 1$,*

$$\mathcal{G}^{\mathcal{A}}_{\phi}\left(\mathcal{F}_s, T\right) \le \bar{C}\log T.$$

Proof. We adapt the proof of Theorem 1 in Hazan et al. (2007) to the feedback $\phi^{(1)}$. Fix $\phi = \phi^{(1)}$, $T \ge 1$, and $f \in \mathcal{F}_s$. Selecting $\eta_t = 1/Ht$ for any $t = 2, \dots T$, one has that for any $x \in \mathcal{X}$ and f_t,

$$\mathbb{E}\left[\phi_t^{(1)}\left(X_t, f_t\right)\,\middle|X_t = x\right] = \nabla f_t(x), \quad \text{and} \quad \mathbb{E}\left[\left\|\phi_t^{(1)}\left(x, f_t\right)\right\|^2\right] \le G^2 + \sigma^2, \tag{C-9}$$

for some $\sigma \geq 0$. Define x^* as the single best action in hindsight: $x^* = \arg\min_{x\in\mathcal{X}} \left\{\sum_{t=1}^{T} f_t(x)\right\}$. Then, by a Taylor expansion, for any $x \in \mathcal{X}$ there is a point $\tilde{x}$ on the segment between x and x^* such that:

$$\begin{aligned} f_t(x^*) &= f_t(x) + \nabla f_t(x) \cdot (x^* - x) + \frac{1}{2}(x^* - x) \cdot \nabla^2 f_t(\tilde{x})(x^* - x) \\ &\overset{(a)}{\geq} f_t(x) + \nabla f_t(x) \cdot (x^* - x) + \frac{H}{2} \left\| x^* - x \right\|^2, \end{aligned}$$

for any $t \in \mathcal{T}$, where (a) holds by (10). Substituting X_t in the above and taking expectation with respect to X_t, one has:

$$\mathbb{E}\left[f_t(X_t)\right] - f_t(x^*) \;\leq\; \mathbb{E}\left[\nabla f_t(X_t) \cdot (X_t - x^*)\right] - \frac{H}{2}\mathbb{E}\left\| x^* - X_t \right\|^2, \tag{C-10}$$

for any $t \in \mathcal{T}$. By the OGD step,

$$\left\| X_{t+1} - x^* \right\|^2 \;=\; \left\| P_{\mathcal{X}}\left(X_t - \eta_{t+1}\phi_t^{(1)}\left(X_t, f_t\right)\right) - x^* \right\|^2 \;\overset{(a)}{\leq}\; \left\| X_t - \eta_{t+1}\phi_t^{(1)}\left(X_t, f_t\right) - x^* \right\|^2,$$

where (a) follows from a standard contraction property of the Euclidean projection operator. Taking expectation with respect to X_t, one has:

$$\begin{aligned} \mathbb{E}\left\| X_{t+1} - x^* \right\|^2 &\leq \mathbb{E}\left\| X_t - x^* \right\|^2 + \eta_{t+1}^2 \mathbb{E}\left\| \phi_t^{(1)}\left(X_t, f_t\right) \right\|^2 - 2\eta_{t+1}\mathbb{E}\left[\left(\phi_t^{(1)}\left(X_t, f_t\right)\right) \cdot (X_t - x^*)\right] \\ &\overset{(a)}{\leq} \mathbb{E}\left\| X_t - x^* \right\|^2 + \eta_{t+1}^2\left(G^2 + \sigma^2\right) - 2\eta_{t+1}\mathbb{E}\left[\left(\nabla f_t(X_t)\right) \cdot (X_t - x^*)\right], \end{aligned}$$

where (a) follows from (C-9). Therefore, for any $t \in \mathcal{T}$, we get:

$$\mathbb{E}\left[\nabla f_t(X_t) \cdot (X_t - x^*)\right] \;\leq\; \frac{\mathbb{E}\left\| X_t - x^* \right\|^2 - \mathbb{E}\left\| X_{t+1} - x^* \right\|^2}{2\eta_{t+1}} + \frac{\eta_{t+1}}{2}\left(G^2 + \sigma^2\right). \tag{C-11}$$

Summing (C-10) over the horizon and using (C-11), one has:

$$\begin{aligned} \sum_{t=1}^{T}\left(\mathbb{E}\left[f_t(X_t)\right] - f_t(x^*)\right) &\leq \frac{1}{2}\sum_{t=2}^{T}\mathbb{E}\left\| X_t - x^* \right\|^2\left(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t} - H\right) \\ &\quad + \frac{1}{2}\mathbb{E}\left\| X_1 - x^* \right\|^2\left(\frac{1}{\eta_2} - \frac{H}{2}\right) - \frac{1}{2}\mathbb{E}\left\| X_{T+1} - x^* \right\|^2\left(\frac{1}{\eta_{T+1}} + \frac{H}{2}\right) \\ &\quad + \frac{\left(G^2 + \sigma^2\right)}{2}\sum_{t=1}^{T}\eta_{t+1} \qquad \text{(C-12)} \\ &\overset{(a)}{\leq} \frac{\left(G^2 + \sigma^2\right)}{2}\sum_{t=1}^{T}\frac{1}{Ht} \;\leq\; \frac{\left(G^2 + \sigma^2\right)}{2H}\left(1 + \log T\right), \end{aligned}$$

where (a) holds using $\eta_t = 1/Ht$. Since the above holds for any sequence of functions in $\mathcal{F}_s$ we have that

$$\mathcal{G}_{\phi^{(1)}}^{\mathcal{A}}\left(\mathcal{F}_s, T\right) \;\leq\; \frac{\left(G^2 + \sigma^2\right)}{2H}\left(1 + \log T\right),$$

which concludes the proof. ■

C.3 Lower bounds

The last two results of this section, Lemma 3 and Lemma 4, establish lower bounds on the best achievable performance in the adversarial setting, under the structures $(\mathcal{F}_s, \phi^{(0)})$, and $(\mathcal{F}, \phi^{(1)})$, respectively. Lemma 3 provides a lower bound that (together with the upper bound in Lemma 1) establishes that the EGS algorithm is rate optimal in a setting with strongly convex cost functions and noisy cost observations. Lemma 4 provides a lower bound that matches the upper bound in Lemma 3.1 in Flaxman et al. (2005), establishing that the OGD algorithm (with a careful selection of step-sizes), is rate optimal in a setting with general convex cost functions and noisy gradient observations.

Lemma 3 *Let Assumption 2 hold. Then, there exists a constant C, independent of T such that for any online algorithm $\mathcal{A} \in \mathcal{P}_{\phi^{(0)}}$ and for all $T \geq 1$:*

$$\mathcal{G}^{\mathcal{A}}_{\phi^{(0)}}(\mathcal{F}_s, T) \;\geq\; C\sqrt{T}.$$

Proof. Let $\mathcal{X} = [0,1]$. Consider the quadratic functions f^1 and f^2 in (A-12), used in the proof of Theorems 4 and 5. (note that δ will be selected differently). Fix some algorithm $\mathcal{A} \in \mathcal{P}_{\phi^{(0)}}$. Let $\tilde{f}$ be a random sequence where in the beginning of the horizon nature draws (according to a uniform discrete distribution) a cost function from $\{f^1, f^2\}$, and applies it throughout the horizon. Taking expectation over the random sequence $\tilde{f}$ one has

$$\mathcal{G}^{\mathcal{A}}_{\phi^{(0)}}(\mathcal{F}_s, T) \;\geq\; \frac{1}{2}\mathbb{E}^{\mathcal{A}}_{f^1}\left[\sum_{t=1}^{T}\left(f^1(X_t) - f^1(x_1^*)\right)\right] + \frac{1}{2}\mathbb{E}^{\mathcal{A}}_{f^2}\left[\sum_{t=1}^{T}\left(f^2(X_t) - f^2(x_2^*)\right)\right],$$

where the inequality follows as in step 3 of the proof of theorem 2. In the following we use notation described at the proof of Theorem 5, for the online algorithm $\mathcal{A}$. We start by bounding the Kullback-Leibler divergence between $\mathbb{P}^{\mathcal{A},\tau}_{f^1}$ and $\mathbb{P}^{\mathcal{A},\tau}_{f^2}$ for all $\tau \in \mathcal{T}$:

$$\begin{aligned}
\mathcal{K}\left(\mathbb{P}^{\mathcal{A},T}_{f^1} \| \mathbb{P}^{\mathcal{A},T}_{f^2}\right) \;&\overset{(a)}{\leq}\; \tilde{C}\mathbb{E}^{\mathcal{A}}_{f^1}\left[\sum_{t=1}^{T}\left(f^1(X_t) - f^2(X_t)\right)^2\right] \;=\; \tilde{C}\mathbb{E}^{\mathcal{A}}_{f^1}\left[\sum_{t=1}^{T}\left(\delta X_t - \frac{\delta}{2}\right)^2\right] \\
&=\; \tilde{C}\mathbb{E}^{\mathcal{A}}_{f^1}\left[\delta^2\sum_{t=1}^{T}\left(X_t - x_1^*\right)^2\right] \\
&\overset{(b)}{=}\; \tilde{C}\mathbb{E}^{\mathcal{A}}_{f^1}\left[2\delta^2\sum_{t=1}^{T}\left(f^1(X_t) - f^1(x_1^*)\right)\right] \;\overset{(c)}{\leq}\; 4\tilde{C}\delta^2\mathcal{G}^{\mathcal{A}}_{\phi^{(0)}}(\mathcal{F}_s, T),
\end{aligned} \tag{C-13}$$

where: (a) follows from Lemma A-3; (b) holds since

$$f^1(x) - f^1(x_1^*) \;=\; \nabla f^1(x_1^*)\cdot(x - x_1^*) + \frac{1}{2}\cdot\nabla f^1(x_1^*)\cdot(x - x_1^*)^2 \;=\; \frac{1}{2}(x - x_1^*)^2$$

for any $x \in \mathcal{X}$; and (c) holds by

$$\begin{aligned}\mathcal{G}_{\phi(0)}^{\mathcal{A}}(\mathcal{F}_s,T) &\geq \frac{1}{2}\mathbb{E}_{f^1}^{\mathcal{A}}\left[\sum_{t=1}^{T}\left(f^1(X_t)-f^1(x_1^*)\right)\right]+\frac{1}{2}\mathbb{E}_{f^2}^{\mathcal{A}}\left[\sum_{t=1}^{T}\left(f^2(X_t)-f^2(x_2^*)\right)\right]\\ &\geq \frac{1}{2}\mathbb{E}_{f^1}^{\mathcal{A}}\left[\sum_{t=1}^{T}\left(f^1(X_t)-f^1(x_1^*)\right)\right]. \end{aligned} \tag{C-14}$$

Therefore, for any $x_0 \in \mathcal{X}$, by Lemma A-2 with $\varphi_t = \mathbb{1}\{X_t > x_0\}$, we have:

$$\max\left\{\mathbb{P}_{f^1}^{\mathcal{A}}\left\{X_\tau > x_0\right\},\mathbb{P}_{f^2}^{\mathcal{A}}\left\{X_\tau \leq x_0\right\}\right\} \geq \frac{1}{4}\exp\left\{-4\tilde{C}\delta^2\mathcal{G}_{\phi(0)}^{\mathcal{A}}(\mathcal{F}_s,T)\right\} \qquad \text{for all } \tau\in\mathcal{T}. \tag{C-15}$$

Set $x_0 = \frac{1}{2}\left(x_1^*+x_2^*\right) = 1/2+\delta/4$. Then, following step 3 in the proof of Theorem 5, one has:

$$\begin{aligned}\mathcal{G}_{\phi(0)}^{\mathcal{A}}(\mathcal{F}_s,T) &\geq \frac{1}{2}\sum_{t=1}^{T}\left(f^1(x_0)-f^1(x_1^*)\right)\mathbb{P}_{f^1}^{\mathcal{A}}\left\{X_t > x_0\right\}+\frac{1}{2}\sum_{t=1}^{T}\left(f^2(x_0)-f^2(x_2^*)\right)\mathbb{P}_{f^2}^{\mathcal{A}}\left\{X_t \leq x_0\right\}\\ &\geq \frac{\delta^2}{16}\sum_{t=1}^{T}\left(\mathbb{P}_{f^1}^{\mathcal{A}}\left\{X_t > x_0\right\}+\mathbb{P}_{f^2}^{\mathcal{A}}\left\{X_t \leq x_0\right\}\right)\\ &\geq \frac{\delta^2}{16}\sum_{t=1}^{T}\max\left\{\mathbb{P}_{f^1}^{\mathcal{A}}\left\{X_t > x_0\right\},\mathbb{P}_{f^2}^{\mathcal{A}}\left\{X_t \leq x_0\right\}\right\}\\ &\overset{(a)}{\geq} \frac{\delta^2}{16}\sum_{t=1}^{T}\frac{1}{4}\exp\left\{-4\tilde{C}\delta^2\mathcal{G}_{\phi(0)}^{\mathcal{A}}(\mathcal{F}_s,T)\right\} = \frac{\delta^2 T}{16}\exp\left\{-4\tilde{C}\delta^2\mathcal{G}_{\phi(0)}^{\mathcal{A}}(\mathcal{F}_s,T)\right\}\end{aligned}$$

where (a) holds by (C-15). Set $\delta = \left(\frac{4}{\tilde{C}T}\right)^{1/4}$. Then, one has for $\beta = 8\sqrt{\tilde{C}/T}$:

$$\beta\mathcal{G}_{\phi(0)}^{\mathcal{A}}(\mathcal{F}_s,T) \geq \exp\left\{-\beta\mathcal{G}_{\phi(0)}^{\mathcal{A}}(\mathcal{F}_s,T)\right\}. \tag{C-16}$$

Let y_0 be the unique solution to the equation $y = \exp\{-y\}$. Then, (C-16) implies $\beta\mathcal{G}_{\phi(0)}^{\mathcal{A}}(\mathcal{S},T) \geq y_0$. In particular, since $y_0 > 1/2$ this implies

$$\mathcal{G}_{\phi(0)}^{\mathcal{A}}(\mathcal{F}_s,T) \geq 1/(2\beta) = \frac{1}{16\sqrt{\tilde{C}}}\cdot\sqrt{T}.$$

This concludes the proof. ■

Lemma 4 *Let Assumption 1 hold. Then, there exists a constant C, independent of T, such that for any online algorithm $\mathcal{A}\in\mathcal{P}_{\phi(1)}$ and for all $T \geq 1$:*

$$\mathcal{G}_{\phi(1)}^{\mathcal{A}}(\mathcal{F},T) \geq C\sqrt{T}.$$

Proof. Fix $T \geq 1$. Let $\mathcal{X} = [0,1]$, and consider functions f^1 and f^2 that are given in (A-7), and used in the proof of Theorem 2 (note that δ will be selected differently). Let $\tilde{f}$ be a random sequence of cost functions, where in the beginning of the time horizon nature draws (from a uniform discrete distribution) a function from $\left\{f^1,f^2\right\}$, and applies it throughout the horizon.

Fix $\mathcal{A}\in\mathcal{P}_{\phi(1)}$. In the following we use notation described in the proof of Theorem 2, as well as in Lemma 3. Set $\delta = 1/\sqrt{16\tilde{C}T}$, where $\tilde{C}$ is the constant that appears in Assumption 1. Then:

$$\begin{aligned}\mathcal{K}\left(\mathbb{P}_{f^1}^{\mathcal{A},T}\|\mathbb{P}_{f^2}^{\mathcal{A},T}\right) &\overset{(a)}{\leq} \tilde{C}\mathbb{E}_{f^1}^{\mathcal{A}}\left[\sum_{t=1}^{T}\left(\nabla f^1(X_t)-\nabla f^2(X_t)\right)^2\right] \\ &= \tilde{C}\mathbb{E}_{f^1}^{\mathcal{A}}\left[\sum_{t=1}^{T}16\delta^2 X_t^2\right] \leq 16\tilde{C}T\delta^2 \overset{(b)}{\leq} 1,\end{aligned} \tag{C-17}$$

where (a) follows from Lemma A-1, and (b) holds by $\delta = 1/\sqrt{16\tilde{C}T}$. Since $\mathcal{K}(\mathbb{P}_1^{\mathcal{A},\tau}\|\mathbb{P}_2^{\mathcal{A},\tau})$ is non-decreasing in τ throughout the horizon, we deduce that the Kullback-Leibler divergence is bounded by 1 throughout the horizon. Therefore, for any $x_0 \in \mathcal{X}$, by Lemma A-2 with $\varphi_\tau = \mathbb{1}\{X_\tau \leq x_0\}$ and $\beta = 1$, one has:

$$\max\left\{\mathbb{P}_{f^1}^{\mathcal{A}}\left\{X_\tau \leq x_0\right\}, \mathbb{P}_{f^2}^{\mathcal{A}}\left\{X_t > x_0\right\}\right\} \geq \frac{1}{4e} \qquad \text{for all } \tau \in \mathcal{T}. \tag{C-18}$$

Set $x_0 = \frac{1}{2}\left(x_1^* + x_2^*\right) = \frac{1}{2}$. Taking expectation over $\tilde{f}$ and following step 3 in the proof of Theorem 2, one has:

$$\begin{aligned}\mathcal{G}_{\phi^{(1)}}^{\mathcal{A}}\left(\mathcal{F},T\right) &\geq \frac{1}{2}\mathbb{E}_{f^1}^{\mathcal{A}}\left[\sum_{t=1}^{T}\left(f^1(X_t)-f^1(x_1^*)\right)\right] + \frac{1}{2}\mathbb{E}_{f^2}^{\mathcal{A}}\left[\sum_{t=1}^{T}\left(f^2(X_t)-f^2(x_2^*)\right)\right] \\ &\geq \frac{1}{2}\sum_{t=1}^{T}\left(f^1(x_0)-f^1(x_1^*)\right)\mathbb{P}_{f^1}^{\mathcal{A}}\left\{X_t > x_0\right\} + \frac{1}{2}\sum_{t=1}^{T}\left(f^2(x_0)-f^2(x_2^*)\right)\mathbb{P}_{f^2}^{\mathcal{A}}\left\{X_t \leq x_0\right\} \\ &\geq \left(\frac{\delta}{4}+\frac{\delta^2}{2}\right)\sum_{t=1}^{T}\left(\mathbb{P}_{f^1}^{\mathcal{A}}\left\{X_t > x_0\right\} + \mathbb{P}_{f^2}^{\mathcal{A}}\left\{X_t \leq x_0\right\}\right) \\ &\geq \left(\frac{\delta}{4}+\frac{\delta^2}{2}\right)\sum_{t=1}^{T}\max\left\{\mathbb{P}_{f^1}^{\pi}\left\{X_t > x_0\right\}, \mathbb{P}_{f^2}^{\mathcal{A}}\left\{X_t \leq x_0\right\}\right\} \\ &\overset{(a)}{\geq} \left(\frac{\delta}{4}+\frac{\delta^2}{2}\right)\sum_{t=1}^{T}\frac{1}{4}\exp\left\{-1\right\} \geq \frac{\delta T}{16e} \overset{(b)}{=} \frac{1}{64e\sqrt{\tilde{C}}}\cdot\sqrt{T},\end{aligned}$$

where (a) holds by (C-18), and (b) holds by $\delta = 1/\sqrt{16\tilde{C}T}$. This concludes the proof. ■

D Numerical Results

We illustrate the upper bounds on the regret by numerical experiments measuring the average regret that is incurred in the presence of various patterns of changing costs, and under different feedback structures and noise. We compare the performance of the restarted OGD and restarted EGS against the performance achieved by applying the respective subroutine without restarting, and with fixed step sizes. We note that policies of fixed step sizes, while having no performance guarantees relative the dynamic oracle, are considered in the SA literature as practical approach to sequential stochastic optimization of general cost functions when these may change; see, e.g., chapter 4 of Benveniste et al. (1990).

Variation and feedback. We fix $\mathcal{X} = [-2,3]$ and consider the sequence of quadratic cost functions $f_t(x_t) = \frac{x_t^2}{2} - b_t x_t + 1$, where the coefficient b_t is time-varying. In particular, for a given horizon length $T \geq 1000$, we let τ be the random time in which the cost begins to change, drawn from a discrete uniform distribution over $\{1, 2, \ldots, \lfloor T/4 \rfloor\}$. Then, we consider the following variation patterns:

$$b_t^{shock} = \begin{cases} 1 & \text{if } t \leq \tau \\ 0 & \text{otherwise} \end{cases} \qquad b_t^{decay} = \begin{cases} 1 & \text{if } t \leq \tau \\ e^{-10(t-\tau)/T} & \text{otherwise} \end{cases} \qquad b_t^{linear} = \begin{cases} 1 & \text{if } t \leq \tau \\ \frac{T-t}{T-\tau} & \text{otherwise} \end{cases}$$

for all $t = 1, \ldots, T$, where α is some decay parameter in $(0, 1)$. One may observe that the variation may be bounded by the budget $V_T = 1$ in all the considered patterns. Let $\{\varepsilon_t\}_t^T$ be a sequence of independent normal random variables with zero mean and standard deviation σ. We consider the case of noisy access to the cost, where $\phi_t^{(0)}(x_t, f_t) = f_t(x_t) + \varepsilon_t$ for each $t \in \mathcal{T}$, and the case of noisy access to the gradient, where $\phi_t^{(1)}(x_t, f_t) = \nabla f_t(x_t) + \varepsilon_t$ for each $t \in \mathcal{T}$. At each epoch $t \in \mathcal{T}$ an action $X_t \in \mathcal{X}$ is selected, and then the expected cost $f_t(X_t)$ is incurred, and a feedback $\phi_t(x_t, f_t)$ is observed.

Policies and performance. In the case of noisy gradient access we use a version of the restarted OGD policy that is considered in §4 and §5.1, where the first action in batch $j \geq 2$ is obtained by taking a gradient step from the last action of batch $j-1$; in other words, only the sequence of gradient steps $\{\eta_t\}$ is restarted. Similarly, when the feedback consists of noisy cost observations we use a variation of the restarted EGS policy that is considered in §5.2, where only the sequences $\{a_t\}$, $\{h_t\}$, and $\{\delta_t\}$ are restarted. Given the actions $\{X_t\}$ generated throughout T epoch by a policy π under feedback ϕ and a sequence f of cost functions, we measure the regret relative to the dynamic oracle $R_\phi^\pi(f, T) = \sum_{t=1}^T (f_t(X_t) - f_t(x_t^*))$. We denote the relative loss (in percentage) relative to the dynamic oracle by $L_\phi^\pi(f, T) = 100 \cdot R_\phi^\pi(f, T) / \left(\sum_{t=1}^T f_t(x_t^*)\right)$. We refer to policy (OGD or EGS) as "non-restarted" when applied without restarting, and as "fixed step size of a" when these apply a fixed (non-updating) step size a ($\eta_t = a$ in the OGD; $a_t = a$ and $\delta_t = h_t = a^{1/4}$ in the EGS).

For each of the considered variation patterns, and for each value of $\sigma \in \{0.1, 0.3, 1\}$, we simulated the action paths of the policy (restated EGS for feedback $\phi^{(0)}$, restarted OGD for feedback $\phi^{(1)}$) for various values of $T \in \{1000, 5000, \ldots, 37000\}$, replicating each instance 10^3 times and calculating the average regret and relative loss. Assuming the structure $R_\phi^\pi(f, T) = cT^\alpha$, we estimate the coefficients c and α from fitting the log-regret as a function of log-time.

Results and discussion. Table 1 details the estimated coefficients c and α under the considered variation patterns and σ values, for the restarted OGD and feedback $\phi^{(1)}$. Table 1 also includes the average loss (in %) relative to the dynamic oracle for two representative values of T. Table 2 includes the respective results for the restarted EGS and feedback $\phi^{(0)}$. In all the linear fits we observed $R^2 > 0.98$, and the standard error of the percentage loss was always below 5% of the policy's average performance.

Figure 1 depicts the averaged regret the restarted policies incur at each epoch, for one representative instance (decay-type variation, $T = 1000$, $\sigma = 0.3$). For illustration purposes, Figure 1 also includes the regret incurred at each epoch by the subroutine policies (OGD/EGS) when those applied without restarting. The estimated values of α (capturing the regret rates) are consistent with the theoretical bounds (0.5 for restarted OGD under $\phi^{(1)}$, 0.67 for restarted EGS under $\phi^{(0)}$). Together with the

Variation pattern	b_t^{shock}			b_t^{decay}			b_t^{linear}		
σ	0.1	0.3	1	0.1	0.3	1	0.1	0.3	1
α	0.54	0.54	0.54	0.47	0.47	0.52	0.47	0.51	0.54
c	0.26	0.32	1.02	0.14	0.16	0.89	0.05	0.09	0.94
$L_\phi^\pi(f,T)$, $T = 5000$:									
Restarted	0.56	0.68	2.02	0.05	0.17	1.56	0.03	0.17	1.78
Non-restarted	3.02	3.02	3.08	4.94	4.95	5.01	5.81	5.81	5.86
Fixed step size of 0.1	0.09	0.31	2.80	0.05	0.27	2.82	0.05	0.31	3.26
Fixed step size of 0.01	0.64	0.66	0.89	0.22	0.25	0.49	0.18	0.21	0.49
Fixed step size of 0.001	2.87	2.87	2.89	2.41	2.41	2.43	2.44	2.45	2.47
$L_\phi^\pi(f,T)$, $T = 25000$:									
Restarted	0.26	0.32	0.94	0.02	0.07	0.71	0.01	0.08	0.82
Non-restarted	1.49	1.49	1.50	3.49	3.50	3.51	5.41	5.41	5.41
Fixed step size of 0.1	0.04	0.25	2.70	0.03	0.25	2.75	0.04	0.29	3.21
Fixed step size of 0.01	0.13	0.15	0.38	0.03	0.06	0.29	0.03	0.06	0.34
Fixed step size of 0.001	0.85	0.85	0.88	0.43	0.44	0.46	0.37	0.38	0.41

Table 1: **Performance of restarted OGD under noisy gradient observations.**

Variation pattern	b_t^{shock}			b_t^{decay}			b_t^{linear}		
σ	0.1	0.3	1	0.1	0.3	1	0.1	0.3	1
α	0.68	0.68	0.68	0.67	0.67	0.68	0.67	0.67	0.68
c	2.22	2.28	2.84	2.13	2.18	2.88	2.09	2.16	2.83
$L_\phi^\pi(f,T)$, $T = 5000$:									
Restarted	14.45	14.82	19.02	14.42	14.89	19.01	15.52	16.06	21.20
Non-restarted	37.71	38.03	41.25	29.96	31.30	33.35	28.49	29.46	39.35
Fixed step size of 0.1	26.58	27.33	34.99	26.78	27.53	35.62	28.49	29.46	39.35
Fixed step size of 0.01	8.48	8.70	11.22	8.10	8.33	10.87	8.48	8.86	11.82
Fixed step size of 0.001	5.22	5.30	5.95	4.94	4.81	5.53	4.96	5.03	5.80
$L_\phi^\pi(f,T)$, $T = 25000$:									
Restarted	8.44	8.67	11.19	8.35	8.58	11.18	8.92	9.19	12.27
Non-restarted	31.27	31.31	32.54	32.42	31.22	33.14	19.10	19.42	23.56
Fixed step size of 0.1	26.29	27.00	34.51	26.57	27.30	35.12	28.32	29.25	38.94
Fixed step size of 0.01	7.89	8.10	10.59	7.86	8.08	10.61	8.41	8.66	11.59
Fixed step size of 0.001	3.26	3.32	4.07	2.84	2.92	3.69	2.96	3.04	3.91

Table 2: **Performance of restarted EGS under noisy cost observations.**

estimation of the coefficient c (ranges in $[0.05, 0.94]$ for restarted OGD, in $[2.09, 2.88]$ for restarted EGS), these demonstrate the actual performance of the policies in a variety of cost-varying instances. One may observe that when σ is larger (observations are more noisy) the multiplying constant typically increases. The estimated loss values indicate the extant at which each policy's performance is "close" the one of the dynamic oracle (the restarted policies as well as the subroutine themselves, when applied without restarting, gets "closer" to the dynamic oracle when T grows).

One may observe that, not surprisingly, the restarted policies consistently outperform the OCO policies when these are not restarted. Considering policies with a fixed step size, we observe that different setting are characterized by different step sizes are considered to be the (ex-post) "best" in different settings; the "right" step size is effected by the variation pattern, the feedback structure, and the noisiness of the observations. Indeed, comparing the performance of the restarting procedure to

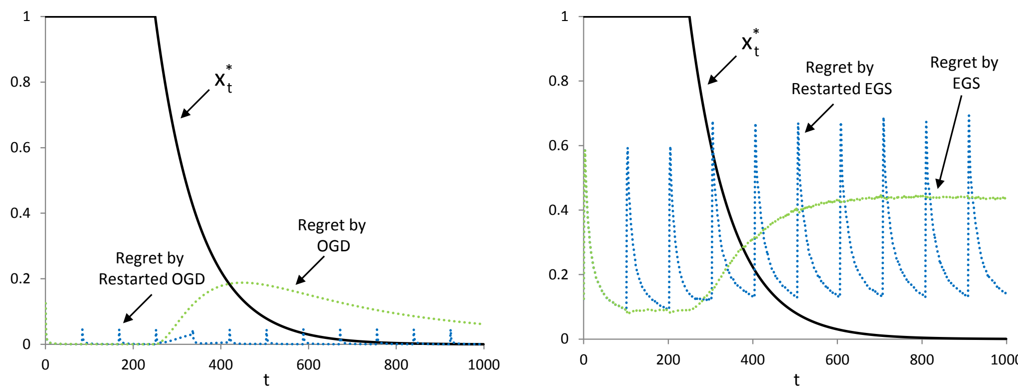

Figure 1: **Regret in the presence of changing cost.** The cost $f_t(x_t) = \frac{x_t^2}{2} - b_t^{decay} x_t + 1$ is reflected by $x_t^* = b_t^{decay}$, with $T = 1,000$, $\tau = 250$, and $\sigma = 0.3$. *(Left)* The average regret incurred at each epoch by the restarting procedure with OGD as a subroutine, and the one incurred at each epoch by OGD (without restarting), under feedback $\phi_t^{(1)}$. *(Right)* The average regret incurred at each epoch by the restarting procedure with EGS as a subroutine, and the one incurred in each epoch by EGS (without restarting), under feedback $\phi_t^{(0)}$.

that of policies with fixed step size, one may observe that in various settings the restarting procedure is outperformed by a certain fixed step size; this occur more often for small values of T. While there are various heuristics to set a-priory a fixed step size, non of those have any performance guarantee relative to the dynamic oracle for arbitrary variation (even when the variation is known to be fixed). We note that policies with a fixed step-size may perform well (and even better than known rate-optimal policies) over finite horizons even when the environment is stationary.[4] While the restarting policies were not designed and tuned in this paper to optimize practical performance, in most of the instances that are considered here they perform at least "on par" with policies with the considered fixed step-sizes.

References

Agarwal, A., O. Dekel, and L. Xiao (2010). Optimal algorithms for online convex optimization with multi-point bandit feedback. *In Proceedings of the 23rd Annual Conference on Learning Theory (COLT)*, 28–40.

Benveniste, A., P. Priouret, and M. Metivier (1990). *Adaptive algorithms and stochastic approximations*. Springer-Verlag, New York.

Flaxman, A. D., A. T. Kalai, and H. B. McMahan (2005). Online convex optimization in the bandit setting: Gradient descent without gradient. *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, 385–394.

Hazan, E., A. Agarwal, and S. Kale (2007). Logarithmic regret algorithms for online convex optimization. *Machine Learning 69*, 169–192.

[4]Repeating the numerical analysis for various *stationary* settings we observed that policies with fixed step sizes may incur relative loss of less than 0.02 percent under noisy gradient access, and less than 1 percent under noisy cost access; for various (small enough) values of T these policies outperformed rate optimal SA policies.