Appendix

Table of Contents

A Implementation Details of the Numerical Experiments

All experiments are conducted using Matlab R2019a on Ubuntu 18.04 with the AMD Ryzen 2700X 8-core processor and 16GB 2133MHz memory.

In the non-stationary LQR experiments, we select $n_x = 6$, $n_u = 6$ and $\gamma = 0.5$. The dynamical matrices A_0 and B_0 at episode 0 are randomly generated from a Gaussian distribution $\mathcal{N}(0, 0.1^2)$. Then, we generate the time-varying dynamical matrices according to $A_{t+1} = A_t + 0.01M_t$ and $B_{t+1} = B_t + 0.01N_t$, where M_t and N_t are random matrices whose entries are uniformly sampled from [0,1]. To evaluate the cost function $V_t(K_t)$ given the policy parameter K_t at episode t, we roll out a trajectory of length $H = 50$ using policy parameter K_t and sum up the collected rewards.

In the non-stationary resource allocation experiments, the policy function $\pi_{i,t}(o_i; \theta_{i,t})$ is parameterized as the following: $a_{ij} = \exp(z_{ij}) / \sum_j \exp(z_{ij})$, where $z_{ij} = \sum_{p=1}^{9} \psi_p(o_i)\theta_{ij}(p)$ and $\theta_i = [\dots, \theta_{ij}, \dots]^T$ and episode index t is omitted for notation simplicity. Specifically, the feature function $\psi_p(o_i)$ is selected as $\psi_p(o_i) = \|o_i - c_p\|^2$, where c_p is the parameter of the p-th feature function. Effectively, the agents need to make decisions on 64 actions, and each action is decided by 9 parameters. Therefore, the problem dimension is $d = 576$. The discount factor is set as $\gamma = 0.75$ and the length of horizon $H = 30$. The time-varying sensitivity parameter $\zeta_{i,t}$ is generated as follows: let $\zeta_{i,0} = 1$ and $\zeta_{i,t+1} = \zeta_{i,t} + 0.1P_t$, where P_t is a random number uniformly sampled from $[-1, 1]$.

B Proof of Lemma 2.5

By definition of the residual feedback, we have

$$\begin{aligned}
\mathbb{E}[\|\tilde{g}_t(x_t)\|^2] &= \mathbb{E}[\frac{1}{\delta^2}\big(f_t(x_t + \delta u_t) - f_{t-1}(x_{t-1} + \delta u_{t-1})\big)^2\|u_t\|^2] \\
&\leq \frac{2}{\delta^2}\mathbb{E}[\big(f_t(x_t + \delta u_t) - f_t(x_{t-1} + \delta u_{t-1})\big)^2\|u_t\|^2] \\
&\quad + \frac{2}{\delta^2}\mathbb{E}[\big(f_t(x_{t-1} + \delta u_{t-1}) - f_{t-1}(x_{t-1} + \delta u_{t-1})\big)^2\|u_t\|^2].
\end{aligned} \tag{14}$$

Since u_t is independent of x_{t-1}, u_{t-1} and the generation of functions f_{t-1} and f_t, we have that $\frac{2}{\delta^2}\mathbb{E}[\big(f_t(x_{t-1}+\delta u_{t-1}) - f_{t-1}(x_{t-1}+\delta u_{t-1})\big)^2\|u_t\|^2] \leq \frac{2d}{\delta^2}\mathbb{E}[\big(f_t(x_{t-1}+\delta u_{t-1}) - f_{t-1}(x_{t-1}+\delta u_{t-1})\big)^2]$. Moreover, adding and subtracting $f_t(x_{t-1}+\delta u_t)$ in the term $\big(f_t(x_t+\delta u_t) - f_t(x_{t-1}+\delta u_{t-1})\big)^2$ of the above inequality, we obtain that

$$\begin{aligned}\mathbb{E}[\|\tilde{g}(x_t)\|^2] \leq & \frac{4}{\delta^2}\mathbb{E}[\big(f_t(x_t+\delta u_t) - f_t(x_{t-1}+\delta u_t)\big)^2\|u_t\|^2] \\ & + \frac{4}{\delta^2}\mathbb{E}[\big(f_t(x_{t-1}+\delta u_t) - f_t(x_{t-1}+\delta u_{t-1})\big)^2\|u_t\|^2] \\ & + \frac{2d}{\delta^2}\mathbb{E}[\big(f_t(x_{t-1}+\delta u_{t-1}) - f_{t-1}(x_{t-1}+\delta u_{t-1})\big)^2].\end{aligned} \tag{15}$$

Since $f_t \in C^{0,0}$ is Lipschitz with constant L_0, we further obtain that

$$\begin{aligned}\mathbb{E}[\|\tilde{g}(x_t)\|^2] \leq & \frac{4L_0^2}{\delta^2}\mathbb{E}[\|x_t - x_{t-1}\|^2\|u_t\|^2] + 4L_0^2\mathbb{E}[\|u_t - u_{t-1}\|^2\|u_t\|^2] \\ & + \frac{2d}{\delta^2}\mathbb{E}[\big(f_t(x_{t-1}+\delta u_{t-1}) - f_{t-1}(x_{t-1}+\delta u_{t-1})\big)^2].\end{aligned} \tag{16}$$

Note that u_t is a Gaussian vector independent from $x_t - x_{t-1}$, we then obtain that $\mathbb{E}[\|x_t - x_{t-1}\|^2\|u_t\|^2] = d\mathbb{E}[\|x_t - x_{t-1}\|^2]$. Furthermore, using Lemma 1 in Nesterov & Spokoiny (2017), we know that $\mathbb{E}[\|u_t - u_{t-1}\|^2\|u_t\|^2] \leq 2\mathbb{E}[(\|u_t\|^2 + \|u_{t-1}\|^2)\|u_t\|^2] = 2\mathbb{E}[(\|u_t\|^4] + 2\mathbb{E}[\|u_{t-1}\|^2\|u_t\|^2] \leq 4(d+4)^2$. Substituting these bounds into inequality (16), we obtain that

$$\begin{aligned}\mathbb{E}[\|\tilde{g}(x_t)\|^2] \leq & \frac{4dL_0^2}{\delta^2}\mathbb{E}[\|x_t - x_{t-1}\|^2] + 16L_0^2(d+4)^2 \\ & + \frac{2d}{\delta^2}\mathbb{E}[\big(f_t(x_{t-1}+\delta u_{t-1}) - f_{t-1}(x_{t-1}+\delta u_{t-1})\big)^2].\end{aligned}$$

Since $x_t = \Pi_{\mathcal{X}}\big[x_{t-1} - \eta\tilde{g}(x_{t-1})\big]$, we get that $\|x_t - x_{t-1}\| = \|\Pi_{\mathcal{X}}\big[x_{t-1} - \eta\tilde{g}(x_{t-1})\big] - \Pi_{\mathcal{X}}\big[x_{t-1}\big]\| \leq \eta\|\tilde{g}(x_{t-1})\|$ due to the nonexpansiveness of the projection operator onto a convex set. Therefore, we have that

$$\begin{aligned}\mathbb{E}[\|\tilde{g}_t(x_t)\|^2] \leq & \frac{4dL_0^2\eta^2}{\delta^2}\mathbb{E}[\|\tilde{g}_{t-1}(x_{t-1})\|^2] + 16L_0^2(d+4)^2 \\ & + \frac{2d}{\delta^2}\mathbb{E}[\big(f_t(x_{t-1}+\delta u_{t-1}) - f_{t-1}(x_{t-1}+\delta u_{t-1})\big)^2].\end{aligned}$$

The proof is complete.

C Proof of Theorem 3.2

Note that $f_{\delta,t}(x)$ is convex for all t, we then conclude that

$$f_{\delta,t}(x_t) - f_{\delta,t}(x) \leq \langle \nabla f_{\delta,t}(x_t), x_t - x\rangle, \text{ for all } x \in \mathcal{X}, \tag{17}$$

Adding and subtracting $\tilde{g}_t(x_t)$ after $\nabla f_{\delta,t}(x_t)$ in above inequality, and taking expectation over u_t on both sides, we obtain that

$$\mathbb{E}\big[f_{\delta,t}(x_t) - f_{\delta,t}(x)\big] \leq \mathbb{E}\big[\langle \tilde{g}_t(x_t), x_t - x\rangle\big]. \tag{18}$$

Since $x_{t+1} = \Pi_{\mathcal{X}}\big[x_t - \eta\tilde{g}(x_t)\big]$, for any $x \in \mathcal{X}$ we have that

$$\begin{aligned}\|x_{t+1} - x\|^2 &= \|\Pi_{\mathcal{X}}\big[x_t - \eta\tilde{g}(x_t)\big] - \Pi_{\mathcal{X}}\big[x\big]\|^2 \\ &\leq \|x_t - \eta\tilde{g}(x_t) - x\|^2 \\ &= \|x_t - x\|^2 - 2\eta\langle\tilde{g}_t(x_t), x_t - x\rangle + \eta^2\|\tilde{g}_t(x_t)\|^2.\end{aligned} \tag{19}$$

Rearranging the above inequality yields that

$$\langle\tilde{g}_t(x_t), x_t - x\rangle = \frac{1}{2\eta}\big(\|x_t - x\|^2 - \|x_{t+1} - x\|^2\big) + \frac{\eta}{2}\|\tilde{g}_t(x_t)\|^2. \tag{20}$$

Taking expectation on both sides of the above inequality over u_t and substituting the resulting bound into (18), we obtain that

$$\mathbb{E}\Big[\sum_{t=0}^{T} f_{\delta,t}(x_t) - \sum_{t=0}^{T} f_{\delta,t}(x)\Big] \leq \frac{1}{2\eta}\|x_0 - x\|^2 + \frac{\eta}{2}\mathbb{E}\Big[\sum_{t=0}^{T}\|\tilde{g}_t(x_t)\|^2\Big]. \tag{21}$$

Since $f_t(x) \in C^{0,0}$, we know that $|f_{\delta,t}(x) - f_t(x)| \leq \delta L_0\sqrt{d}$. Therefore, we obtain from the above inequality that

$$\begin{aligned}\mathbb{E}\big[\sum_{t=0}^{T} f_t(x_t) - \sum_{t=0}^{T} f_t(x)\big] &= \mathbb{E}\big[\sum_{t=0}^{T} f_{\delta,t}(x_t) - \sum_{t=0}^{T} f_{\delta,t}(x)\big] + \mathbb{E}\big[\sum_{t=0}^{T}\big(f_t(x_t) - f_{\delta,t}(x_t)\big) - \sum_{t=0}^{T}\big(f_t(x) - f_{\delta,t}(x)\big)\big]\\ &\leq \frac{1}{2\eta}\|x_0 - x\|^2 + \frac{\eta}{2}\mathbb{E}\big[\sum_{t=0}^{T}\|\tilde{g}_t(x_t)\|^2\big] + 2\sqrt{d}L_0\delta T. \end{aligned} \tag{22}$$

Telescoping the bound in (5) over $t = 1, 2, ..., T$, adding $\mathbb{E}\big[\|\tilde{g}_0(x_0)\|^2\big]$ on both sides, adding $\frac{4dL_0^2\eta^2}{\delta^2}\mathbb{E}[\|\tilde{g}_T(x_T)\|^2]$ to the right hand side and using Assumption 3.1, we obtain that

$$\mathbb{E}\big[\sum_{t=0}^{T}\|\tilde{g}_t(x_t)\|^2\big] \leq \frac{1}{1-\alpha}\mathbb{E}\big[\|\tilde{g}_0(x_0)\|^2\big] + \frac{16}{1-\alpha}L_0^2(d+4)^2T + \frac{2dV_f^2}{1-\alpha}\frac{1}{\delta^2}T, \tag{23}$$

where $\alpha = \frac{4dL_0^2\eta^2}{\delta^2}$. Substituting the above bound into (22) yields that

$$\begin{aligned}\mathbb{E}\big[\sum_{t=0}^{T} f_t(x_t) - \sum_{t=0}^{T} f_t(x)\big] \leq{}& \frac{1}{2\eta}\|x_0 - x\|^2 + \frac{\eta}{2(1-\alpha)}\mathbb{E}\big[\|\tilde{g}_0(x_0)\|^2\big] + \frac{16}{1-\alpha}L_0^2(d+4)^2\eta T\\ &+ 2\sqrt{d}L_0\delta T + \frac{2dV_f^2}{1-\alpha}\frac{\eta}{\delta^2}T. \end{aligned} \tag{24}$$

Since above inequality holds for all $x \in \mathcal{X}$, we can replace x with x^*. When the upper bound on $\|x_0 - x^*\| \leq R$ is known, let $\eta = \frac{R^{\frac{3}{2}}}{2\sqrt{2}L_0\sqrt{d}T^{\frac{3}{4}}}$ and $\delta = \frac{\sqrt{R}}{T^{\frac{1}{4}}}$, so that $\alpha = \frac{4dL_0^2\eta^2}{\delta^2} = \frac{R^2}{2T} \leq \frac{1}{2}$, when $T \geq R^2$. Then, we obtain that

$$\begin{aligned}\mathbb{E}\big[\sum_{t=0}^{T} f_t(x_t) - \sum_{t=0}^{T} f_t(x^*)\big] \leq{}& \sqrt{2}L_0\sqrt{d}RT^{\frac{3}{4}} + \frac{\mathbb{E}\big[\|\tilde{g}_0(x_0)\|^2\big]R^{\frac{3}{2}}}{2\sqrt{2d}L_0T^{\frac{3}{4}}}\\ &+ 8\sqrt{2}\frac{(d+4)^2}{\sqrt{d}}L_0R^{\frac{3}{2}}T^{\frac{1}{4}} + 2L_0\sqrt{d}RT^{\frac{3}{4}} + \frac{\sqrt{2d}RV_f^2}{L_0}T^{\frac{3}{4}}. \end{aligned} \tag{25}$$

When R is unknown, let $\eta = \frac{1}{2\sqrt{2}L_0\sqrt{d}T^{\frac{3}{4}}}$ and $\delta = \frac{1}{T^{\frac{1}{4}}}$, so that $\alpha = \frac{4dL_0^2\eta^2}{\delta^2} = \frac{1}{2T} \leq \frac{1}{2}$. Then, we obtain that

$$\begin{aligned}\mathbb{E}\big[\sum_{t=0}^{T} f_t(x_t) - \sum_{t=0}^{T} f_t(x^*)\big] \leq{}& \sqrt{2}L_0\sqrt{d}R^2T^{\frac{3}{4}} + \frac{\mathbb{E}\big[\|\tilde{g}_0(x_0)\|^2\big]}{2\sqrt{2d}L_0T^{\frac{3}{4}}} + 8\sqrt{2}\frac{(d+4)^2}{\sqrt{d}}L_0T^{\frac{1}{4}}\\ &+ 2\sqrt{d}L_0T^{\frac{3}{4}} + \frac{\sqrt{2d}V_f^2}{L_0}T^{\frac{3}{4}}. \end{aligned} \tag{26}$$

On the other hand, we can let $\eta = \frac{R^{\frac{3}{2}}}{2\sqrt{2}L_0\sqrt{d}T^{\frac{3}{4}}}$ and $\delta = \frac{\sqrt{R}}{L_0^qT^{\frac{1}{4}}}$, where $q \in \mathbb{R}$ is a user-specific parameter. With this choice of parameters, we get $\alpha = \frac{4dL_0^2\eta^2}{\delta^2} = \frac{L_0^{2q}R^2}{2T} \leq \frac{1}{2}$ when $T \geq L_0^{2q}R^2$ and, as a result, we obtain that

$$\begin{aligned}\mathbb{E}\big[\sum_{t=0}^{T} f_t(x_t) - \sum_{t=0}^{T} f_t(x^*)\big] \leq{}& \sqrt{2}L_0\sqrt{d}RT^{\frac{3}{4}} + \frac{\mathbb{E}\big[\|\tilde{g}_0(x_0)\|^2\big]R^{\frac{3}{2}}}{2\sqrt{2d}L_0T^{\frac{3}{4}}} + 8\sqrt{2}\frac{(d+4)^2}{\sqrt{d}}L_0R^{\frac{3}{2}}T^{\frac{1}{4}}\\ &+ 2L_0^{1-q}\sqrt{d}RT^{\frac{3}{4}} + \sqrt{2dR}L_0^{2q-1}V_f^2T^{\frac{3}{4}}. \end{aligned} \tag{27}$$

D PROOF OF THEOREM 3.4

Since $f_t(x) \in C^{1,1}$, we know that $|f_{\delta,t}(x) - f_t(x)| \le \delta^2 L_1 d$. Following the same proof logic as that for proving (22), we obtain that

$$\mathbb{E}\Big[\sum_{t=0}^{T} f_t(x_t) - \sum_{t=0}^{T} f_t(x)\Big] \le \frac{1}{2\eta}\|x_0 - x\|^2 + \frac{\eta}{2}\mathbb{E}\Big[\sum_{t=0}^{T}\|\tilde{g}_t(x_t)\|^2\Big] + 2dL_1\delta^2 T. \tag{28}$$

Substituting the bound in (23) into the above inequality, we obtain that

$$\mathbb{E}\Big[\sum_{t=0}^{T} f_t(x_t) - \sum_{t=0}^{T} f_t(x)\Big] \le \frac{1}{2\eta}\|x_0 - x\|^2 + \frac{\eta}{2(1-\alpha)}\mathbb{E}\big[\|\tilde{g}_0(x_0)\|^2\big] + \frac{16}{1-\alpha}L_0^2(d+4)^2\eta T$$
$$+ 2dL_1\delta^2 T + \frac{2dV_f^2}{1-\alpha}\frac{\eta}{\delta^2}T. \tag{29}$$

Since above inequality holds for all $x \in \mathcal{X}$, we can replace x with x^*. Assuming the bound $\|x_0 - x^*\| \le R$ is known, let $\eta = \frac{R^{\frac{4}{3}}}{2\sqrt{2}L_0 d^{\frac{2}{3}} T^{\frac{2}{3}}}$ and $\delta = \frac{R^{\frac{1}{3}}}{d^{\frac{1}{6}} T^{\frac{1}{6}}}$ so that $\alpha = \frac{4dL_0^2\eta^2}{\delta^2} = \frac{R^2}{2T} \le \frac{1}{2}$ when $T \ge R^2$. Plugging these parameters into above inequality, we finally obtain that

$$\mathbb{E}\Big[\sum_{t=0}^{T} f_t(x_t) - \sum_{t=0}^{T} f_t(x)\Big] \le \sqrt{2}L_0 d^{\frac{2}{3}} R^{\frac{2}{3}} T^{\frac{2}{3}} + \frac{\mathbb{E}\big[\|\tilde{g}_0(x_0)\|^2\big]R^{\frac{4}{3}}}{2\sqrt{2}L_0 d^{\frac{2}{3}} T^{\frac{2}{3}}} + 8\sqrt{2}L_0\frac{(d+4)^2}{d^{\frac{2}{3}}}R^{\frac{4}{3}}T^{\frac{1}{3}}$$
$$+ 2L_1 d^{\frac{2}{3}} R^{\frac{2}{3}} T^{\frac{2}{3}} + \frac{\sqrt{2}}{L_0} d^{\frac{2}{3}} R^{\frac{2}{3}} V_f^2 T^{\frac{2}{3}}. \tag{30}$$

When the bound $\|x_0 - x^*\| \le R$ is unknown. Choose $\eta = \frac{1}{2\sqrt{2}L_0 d^{\frac{2}{3}} T^{\frac{2}{3}}}$ and $\delta = \frac{1}{d^{\frac{1}{6}} T^{\frac{1}{6}}}$ so that $\alpha = \frac{4dL_0^2\eta^2}{\delta^2} = \frac{1}{2T} \le \frac{1}{2}$. Plugging these parameters into above inequality, we finally obtain that

$$\mathbb{E}\Big[\sum_{t=0}^{T} f_t(x_t) - \sum_{t=0}^{T} f_t(x)\Big] \le \sqrt{2}L_0 d^{\frac{2}{3}}\|x_0 - x\|^2 T^{\frac{2}{3}} + \frac{\mathbb{E}\big[\|\tilde{g}_0(x_0)\|^2\big]}{2\sqrt{2}L_0 d^{\frac{2}{3}} T^{\frac{2}{3}}} + 8\sqrt{2}L_0\frac{(d+4)^2}{d^{\frac{2}{3}}}T^{\frac{1}{3}}$$
$$+ 2d^{\frac{2}{3}} L_1 T^{\frac{2}{3}} + \frac{\sqrt{2}}{L_0} d^{\frac{2}{3}} V_f^2 T^{\frac{2}{3}}. \tag{31}$$

The proof is complete.

E PROOF OF THEOREM 4.2

We first consider the case where Assumption 4.1.1 holds. Note that $f_t(x) \in C^{0,0}$. According to Lemma 2.2, $f_{\delta,t}(x)$ has $L_{1,\delta}$-Lipschitz continuous gradient with $L_{1,\delta} = \frac{\sqrt{d}}{\delta}L_0$. Furthermore, according to Lemma 1.2.3 in Nesterov (2013), we have the following inequality

$$\begin{aligned} f_{\delta,t}(x_{t+1}) &\le f_{\delta,t}(x_t) + \langle \nabla f_{\delta,t}(x_t), x_{t+1} - x_t\rangle + \frac{L_{1,\delta}}{2}\|x_{t+1} - x_t\|^2 \\ &= f_{\delta,t}(x_t) - \eta\langle \nabla f_{\delta,t}(x_t), \tilde{g}_t(x_t)\rangle + \frac{L_{1,\delta}\eta^2}{2}\|\tilde{g}_t(x_t)\|^2 \\ &= f_{\delta,t}(x_t) - \eta\langle \nabla f_{\delta,t}(x_t), \Delta_t\rangle - \eta\|\nabla f_{\delta,t}(x_t)\|^2 + \frac{L_{1,\delta}\eta^2}{2}\|\tilde{g}_t(x_t)\|^2, \end{aligned} \tag{32}$$

where $\Delta_t = \tilde{g}_t(x_t) - \nabla f_{\delta,t}(x_t)$. According to Lemma 2.4, we know that $\mathbb{E}_{u_t}[\tilde{g}_t(x_t)] = \nabla f_{\delta,t}(x_t)$. Therefore, taking expectation over u_t conditional on x_t on both sides of inequality (32) and rearranging terms, we obtain that

$$\eta\mathbb{E}[\|\nabla f_{\delta,t}(x_t)\|^2] \le \mathbb{E}[f_{\delta,t}(x_t)] - \mathbb{E}[f_{\delta,t}(x_{t+1})] + \frac{L_{1,\delta}\eta^2}{2}\mathbb{E}[\|\tilde{g}_t(x_t)\|^2]$$
$$\le \mathbb{E}[f_{\delta,t}(x_t)] - \mathbb{E}[f_{\delta,t+1}(x_{t+1})] + \frac{L_{1,\delta}\eta^2}{2}\mathbb{E}[\|\tilde{g}_t(x_t)\|^2] + \mathbb{E}[f_{\delta,t+1}(x_{t+1})] - \mathbb{E}[f_{\delta,t}(x_{t+1})], \tag{33}$$

where the expectation is conditional on x_t. Then, we can further condition both sides of (33) on x_0 without changing the sign of inequality, and then apply the tower rule of conditional expectation to make the expectation in (33) become full expectation. Telescoping the above inequality over $t = 0, ..., T-1$ and dividing both sides by η, we obtain that

$$\begin{aligned}\sum_{t=0}^{T-1} \mathbb{E}[\|\nabla f_{\delta,t}(x_t)\|^2] &\le \frac{\mathbb{E}[f_{\delta,0}(x_0)] - \mathbb{E}[f_{\delta,T}(x_T)]}{\eta} + \frac{L_{1,\delta}\eta}{2}\sum_{t=0}^{T-1}\mathbb{E}[\|\tilde{g}_t(x_t)\|^2] + \frac{W_T}{\eta} \\ &\le \frac{\mathbb{E}[f_{\delta,0}(x_0)] - f^*_{\delta,T}}{\eta} + \frac{L_{1,\delta}\eta}{2}\sum_{t=0}^{T-1}\mathbb{E}[\|\tilde{g}_t(x_t)\|^2] + \frac{W_T}{\eta},\end{aligned} \tag{34}$$

where $f^*_{\delta,T}$ is the lower bound of the smoothed function $f_{\delta,T}(x)$. $f^*_{\delta,T}$ must exist because we assume the orignal function $f_t(x)$ is lower bounded and the smoothed function has a bounded distance from $f_t(x)$ due to Lemma 2.2 for all t.

Next, we consider the case where Assumption 4.1.2 holds. Summing the bound in (5) from $t = 1, ..., T$, adding $\mathbb{E}\big[\|\tilde{g}_0(x_0)\|^2\big]$ on both sides, and adding $\frac{4dL_0^2\eta^2}{\delta^2}\mathbb{E}[\|\tilde{g}_T(x_T)\|^2]$ to the right hand side, we obtain that

$$\mathbb{E}\big[\sum_{t=0}^{T}\|\tilde{g}_t(x_t)\|^2\big] \le \frac{1}{1-\alpha}\mathbb{E}\big[\|\tilde{g}_0(x_0)\|^2\big] + \frac{16}{1-\alpha}L_0^2(d+4)^2T + \frac{2d}{1-\alpha}\frac{\widetilde{W}_T}{\delta^2}, \tag{35}$$

Substituting this bound into the inequality (34), we obtain that

$$\begin{aligned}\sum_{t=0}^{T-1} \mathbb{E}[\|\nabla f_{\delta,t}(x_t)\|^2] &\le \frac{\mathbb{E}[f_{\delta,0}(x_0)] - f^*_{\delta,T}}{\eta} + \frac{W_T}{\eta} + \frac{\sqrt{d}L_0\eta}{2\delta}\frac{1}{1-\alpha}\mathbb{E}\big[\|\tilde{g}_0(x_0)\|^2\big] \\ &\quad + \frac{\sqrt{d}L_0\eta}{2\delta}\frac{16}{1-\alpha}L_0^2(d+4)^2T + \frac{\sqrt{d}L_0\eta}{2\delta}\frac{2d}{1-\alpha}\frac{\widetilde{W}_T}{\delta^2}.\end{aligned} \tag{36}$$

To fullfill the requirement that $|f_t(x) - f_{\delta,t}(x)| \le \epsilon_f$, we set the exporation parameter $\delta = \frac{\epsilon_f}{d^{\frac{1}{2}}L_0}$. In addition, let the stepsize be $\eta = \frac{\epsilon_f^{1.5}}{2\sqrt{2}L_0^2 d^{1.5} T^{\frac{1}{2}}}$. Then, we have that $\alpha = \frac{4dL_0^2\eta^2}{\delta^2} = \frac{\epsilon_f}{2dT} \le \frac{1}{2}$ when $T \ge \frac{1}{d\epsilon_f}$. Therefore, we have that $\frac{1}{1-\alpha} \le 2$. Substituting this bound and the choices of η and δ into the bound (**??**), we finally obtain that

$$\begin{aligned}\sum_{t=0}^{T-1} \mathbb{E}[\|\nabla f_{\delta,t}(x_t)\|^2] &\le 2\sqrt{2}L_0^2\big(\mathbb{E}[f_{\delta,0}(x_0)] - f^*_{\delta,T} + W_T\big)\frac{d^{1.5}}{\epsilon_f^{1.5}}T^{\frac{1}{2}} + \frac{\epsilon_f^{\frac{1}{2}}\mathbb{E}\big[\|\tilde{g}_0(x_0)\|^2\big]}{2\sqrt{2dT}} \\ &\quad + 4\sqrt{2}L_0\epsilon_f^{\frac{1}{2}}\frac{(d+4)^2}{d^{\frac{1}{2}}}T^{\frac{1}{2}} + \frac{L_0^2}{\sqrt{2}}\frac{d^{1.5}\widetilde{W}_T}{\epsilon_f^{1.5}T^{\frac{1}{2}}}.\end{aligned} \tag{37}$$

The proof is complete.

F Proof of Theorem 4.3

We first consider the case where Assumption 4.1.1 holds. Note that when $f_t \in C^{1,1}$ with Lipschitz constant L_1, the smoothed function $f_{\delta,t} \in C^{1,1}$ with Lipschitz constant L_1. Therefore, following the proof of Theorem 4.2 but replacing $L_{1,\delta}$ with L_1, we obtain that

$$\sum_{t=0}^{T-1} \mathbb{E}[\|\nabla f_{\delta,t}(x_t)\|^2] \le \frac{\mathbb{E}[f_{\delta,0}(x_0)] - f^*_{\delta,T}}{\eta} + \frac{L_1\eta}{2}\sum_{t=0}^{T-1}\mathbb{E}[\|\tilde{g}_t(x_t)\|^2] + \frac{W_T}{\eta}. \tag{38}$$

Since $f_t \in C^{1,1}$, according to Lemma 2.2, we have that $\|\nabla f_{\delta,t}(x) - \nabla f_t(x)\| \le \delta L_1(d+3)^{3/2}$.

Furthermore, we have that

$$\begin{aligned}&\sum_{t=0}^{T-1}\mathbb{E}[\|\nabla f(x_t)\|^2] = \sum_{t=0}^{T-1}\mathbb{E}[\|\nabla f(x_t) - \nabla f_{\delta,t}(x_t) + \nabla f_{\delta,t}(x_t)\|^2] \\ &\le 2\mathbb{E}[\|\nabla f(x_t) - \nabla f_{\delta,t}(x_t)\|^2] + 2\mathbb{E}[\|\nabla f_{\delta,t}(x_t)\|^2].\end{aligned} \tag{39}$$

Next, we consider the case where Assumption 4.1.2 holds. Substituting the bound in (35) into (38) and using the bound in (39), we obtain that

$$\sum_{t=0}^{T-1}\mathbb{E}[\|\nabla f(x_t)\|^2] \le 2\frac{\mathbb{E}[f_{\delta,0}(x_0)]-f^*_{\delta,T}}{\eta}+2\frac{W_T}{\eta}+\frac{L_1}{1-\alpha}\mathbb{E}\big[\|\tilde{g}_0(x_0)\|^2\big]\eta+\frac{16L_1}{1-\alpha}L_0^2(d+4)^2\eta T$$
$$+\frac{2dL_1\widetilde{W}_T}{1-\alpha}\frac{\eta}{\delta^2}+2L_1^2(d+3)^3\delta^2T, \tag{40}$$

Choose $\eta=\frac{1}{2\sqrt{2}L_0d^{\frac{4}{3}}T^{\frac{1}{2}}}$ and $\delta=\frac{1}{d^{\frac{5}{6}}T^{\frac{1}{4}}}$. Then, $\alpha=\frac{4dL_0^2\eta^2}{\delta^2}=\frac{1}{2\sqrt{T}}\le\frac{1}{2}$ and $\frac{1}{1-\alpha}\le 2$. Substituting these results into the above inequality, we finally obtain that

$$\sum_{t=0}^{T-1}\mathbb{E}[\|\nabla f(x_t)\|^2] \le 4\sqrt{2}L_0\big(\mathbb{E}[f_{\delta,0}(x_0)]-f^*_{\delta,T}+W_T\big)d^{\frac{4}{3}}T^{\frac{1}{2}}+\frac{L_1\mathbb{E}\big[\|\tilde{g}_0(x_0)\|^2\big]}{\sqrt{2}L_0d^{\frac{4}{3}}T^{\frac{1}{2}}}$$
$$+8\sqrt{2}L_1L_0\frac{(d+4)^2}{d^{\frac{4}{3}}}T^{\frac{1}{2}}+\frac{\sqrt{2}L_1}{L_0}d^{\frac{4}{3}}\widetilde{W}_T+2L_1^2\frac{(d+3)^3}{d^{\frac{5}{3}}}T^{\frac{1}{2}}. \tag{41}$$

The proof is complete.

G Proof of Lemma 5.2

Consider the case when $F_t(x,\xi)\in C^{0,0}$ with $L_0(\xi)$. According to (13), we have that

$$\mathbb{E}[\|\tilde{g}_t(x_t)\|^2]=\mathbb{E}[\frac{1}{\delta^2}\big(F_t(x_t+\delta u_t,\xi_t)-F_{t-1}(x_{t-1}+\delta u_{t-1},\xi_{t-1})\big)^2\|u_t\|^2]$$
$$\le\frac{2}{\delta^2}\mathbb{E}[\big(F_t(x_t+\delta u_t,\xi_t)-F_t(x_{t-1}+\delta u_{t-1},\xi_t)\big)^2\|u_t\|^2]$$
$$+\frac{2}{\delta^2}\mathbb{E}[\big(F_t(x_{t-1}+\delta u_{t-1},\xi_t)-F_{t-1}(x_{t-1}+\delta u_{t-1},\xi_{t-1})\big)^2\|u_t\|^2]. \tag{42}$$

Using the bound in Assumption 5.1 and the fact that the generation of random objective functions $F_{t-1}(\cdot,\xi_{t-1})$ and $F_t(\cdot,\xi_t)$ are independent of u_t, we get that $\frac{2}{\delta^2}\mathbb{E}[\big(F_t(x_{t-1}+\delta u_{t-1},\xi_t)-F_{t-1}(x_{t-1}+\delta u_{t-1},\xi_{t-1})\big)^2\|u_t\|^2]\le\frac{2d}{\delta^2}V_{f,\xi}^2$. In addition, adding and subtracting $F_t(x_{t-1}+\delta u_t,\xi_t)$ in $\big(F_t(x_t+\delta u_t,\xi_t)-F_t(x_{t-1}+\delta u_{t-1},\xi_t)\big)^2$ in above inequality, we obtain that

$$\begin{aligned}\mathbb{E}[\|\tilde{g}_t(x_t)\|^2]\le&\frac{4}{\delta^2}\mathbb{E}[\big(F_t(x_t+\delta u_t,\xi_t)-F_t(x_{t-1}+\delta u_t,\xi_t)\big)^2\|u_t\|^2]\\&+\frac{4}{\delta^2}\mathbb{E}[\big(F_t(x_{t-1}+\delta u_t,\xi_t)-F_t(x_{t-1}+\delta u_{t-1},\xi_t)\big)^2\|u_t\|^2]\\&+\frac{2d}{\delta^2}\mathbb{E}[\big(F_t(x_{t-1}+\delta u_{t-1},\xi_t)-F_{t-1}(x_{t-1}+\delta u_{t-1},\xi_{t-1})\big)^2].\end{aligned} \tag{43}$$

By Lipschitz continuity of $F_t(\cdot;\xi_t)$, we can bound the first two items on the right hand side of above inequality following the same procedure after inequality (16) and get that

$$\begin{aligned}\mathbb{E}[\|\tilde{g}_t(x_t)\|^2]\le&\frac{4dL_0^2\eta^2}{\delta^2}\mathbb{E}[\|\tilde{g}_t(x_{t-1})\|^2]+16L_0^2(d+4)^2\\&+\frac{2d}{\delta^2}\mathbb{E}[\big(F_t(x_{t-1}+\delta u_{t-1},\xi_t)-F_{t-1}(x_{t-1}+\delta u_{t-1},\xi_{t-1})\big)^2].\end{aligned}$$

The proof is complete.

H Residual-Feedback Convex Optimization with Unit Sphere Sampling

Consider the online convex zeroth-order optimization problem (P) with a compact constraint set $\mathcal{X}$. In this section, we assume that the objective function $f(x)$ cannot be queried outside the constraint set $\mathcal{X}$. To satisfy this requirement, we estimate the gradient as

$$\widetilde{g}_t(x_t) := \frac{d}{\delta}\big(f_t(x_t + \delta u_t) - f_{t-1}(x_{t-1} + \delta u_{t-1})\big)u_t, \tag{44}$$

where u_{t-1} and u_t are independently and uniformly sampled from the unit sphere $\mathbb{S} := \{x \in \mathbb{R}^d : \|x\| = 1\}$. Consider the smoothed function $f_\delta(x) = \mathbb{E}_{v\in\mathbb{B}}\big[f(x + \delta v)\big]$, where the random vector v is uniformly sampled from the unit ball $\mathbb{B} = \{x \in \mathbb{R}^d : \|x\| \le 1\}$. Then, we have the following lemma

Lemma H.1. *The function $\widetilde{g}_t(x_t)$ is an unbiased estimate of the gradient $\nabla f_\delta(x_t)$, i.e., $\mathbb{E}\big[\widetilde{g}_t(x_t)\big] = \nabla f_\delta(x_t)$.*

Proof. Since u_t is sampled independently from x_{t-1} and u_{t-1}, and u_t has zero mean, it is straightforward to complete the proof by applying Lemma 2.1 in Flaxman et al. (2005). □

To ensure that the iterates are confined within the constraint set $\mathcal{X}$, we consider the update

$$x_{t+1} = \Pi_{(1-\xi)\mathcal{X}}\big(x_t - \eta\tilde{g}_t(x_t)\big), \tag{45}$$

where the set $(1-\xi)\mathcal{X} := \{(1-\xi)x : \forall x \in \mathcal{X}\}$ is a shrinked version of the original constraint set $\mathcal{X}$. The goal is to select a parameter ξ so that for every $x_\xi \in (1-\xi)\mathcal{X}$, $x_\xi + \delta u \in \mathcal{X}$ for every $u \in \mathbb{S}$. To achieve this, we first make the following assumption that is inspired by Flaxman et al. (2005); Bubeck et al. (2012).

Assumption H.2. *There exist contants r and $\bar{r}$ such that $r\mathbb{B} \subset \mathcal{X} \subset \bar{r}\mathbb{B}$.*

Then, we have the following lemma.

Lemma H.3. *If the parameter ξ satisfies $1 \ge \xi \ge \frac{\delta}{r}$, then for every iterate x_t obtained using (45), we have that $x_t + \delta u_t \in \mathcal{X}$ for all $u_t \in \mathbb{S}$.*

Proof. When $1 \ge \xi \ge \frac{\delta}{r}$, we get that $\|\delta u\| \le \xi r$. Therefore, there exists $x' \in r\mathbb{B} \subset \mathcal{X}$ such that the vector $\delta u = \xi x'$. Since $x_t \in (1-\xi)\mathcal{X}$, there exists $x \in \mathcal{X}$ such that $x_t = (1-\xi)x$, and there exists $x' \in \mathcal{X}$ such that $\delta u = \xi x'$. As a result, we have that $x_t + \delta u = (1-\xi)x + \xi x' \in \mathcal{X}$. This is because set $\mathcal{X}$ is convex. □

Next, we study the regret $R_T := \mathbb{E}\Big[\sum_{t=0}^{T-1} f_t(x_t) - \min_{x\in\mathcal{X}} \sum_{t=0}^{T-1} f_t(x)\Big]$ achieved by executing the online update (45) We do so in the following two steps. First, in Lemma H.4, we provide an upper bound on the difference between the optimal solution that lies in the set $(1-\xi)\mathcal{X}$ and the one that lies in the set $\mathcal{X}$, i.e., $\min_{x\in(1-\xi)\mathcal{X}} \sum_{t=0}^{T-1} f_t(x) - \min_{x\in\mathcal{X}} \sum_{t=0}^{T-1} f_t(x)$; Then, in Theorem H.7, we bound the regret defined by the expected difference between the function values achieved by running the update (45) and the term $\min_{x\in(1-\xi)\mathcal{X}} \sum_{t=0}^{T-1} f_t(x)$, i.e., $\mathbb{E}\Big[\sum_{t=0}^{T-1} f_t(x_t) - \min_{x\in(1-\xi)\mathcal{X}} \sum_{t=0}^{T-1} f_t(x)\Big]$. Adding the two bounds above, we can complete the proof.

In the following lemma we provide a bound on $\min_{x\in(1-\xi)\mathcal{X}} \sum_{t=0}^{T-1} f_t(x) - \min_{x\in\mathcal{X}} \sum_{t=0}^{T-1} f_t(x)$.

Lemma H.4. *If the function f_t is convex and $f_t \in C^{0,0}$ with Lipschitz constant L_0 for all time t, we have that*

$$\sum_{t=0}^{T-1} f_t(x_\xi^*) - \sum_{t=0}^{T-1} f_t(x^*) \le \bar{r}L_0\xi T, \tag{46}$$

where $x_\xi^* = \arg\min_{x\in(1-\xi)\mathcal{X}} \sum_{t=0}^{T-1} f_t(x)$ and $x^* = \arg\min_{x\in\mathcal{X}} \sum_{t=0}^{T-1} f_t(x)$.

Proof. Since $x^* \in \mathcal{X}$, we have that $(1-\xi)x^* \in (1-\xi)\mathcal{X}$. Moreover, since x_ξ^* is the minimizer in the set $(1-\xi)\mathcal{X}$, we get that

$$\sum_{t=0}^{T-1} f_t(x_\xi^*) \le \sum_{t=0}^{T-1} f_t((1-\xi)x^*). \tag{47}$$

Also, since f_t is convex and $(1-\xi)x^* = (1-\xi)x^* + \xi 0$, we have that

$$\begin{aligned} f_t((1-\xi)x^*) &\le (1-\xi)f(x^*) + \xi f_t(0) \\ &\le (1-\xi)f_t(x^*) + \xi f_t(x^*) - \xi f_t(x^*) + \xi f_t(0) \\ &\le f_t(x^*) + \xi L_0 \|x^*\| \le f_t(x^*) + \bar{r} L_0 \xi, \end{aligned} \tag{48}$$

where the last inequality is due to the fact that $x^* \in \mathcal{X} \subset \bar{r}\mathbb{B}$. Summing the inequality (48) over time, we obtain that

$$\sum_{t=0}^{T-1} f_t((1-\xi)x^*) - \sum_{t=0}^{T-1} f_t(x^*) \le \bar{r} L_0 \xi T. \tag{49}$$

Adding up the inequalities (47) and (49) and rearranging terms completes the proof. □

Next, we study the regret $\mathbb{E}\Big[\sum_{t=0}^{T-1} f_t(x_t) - \min_{x\in(1-\xi)\mathcal{X}} \sum_{t=0}^{T-1} f_t(x)\Big]$ following similar steps as in Section 3. First, we can bound the difference between the smoothed objective function $f_{\delta,t}$ and f_t for every time step t as follows.

Lemma H.5. *Consider a function f and its smoothed version f_δ. It holds that*

$$|f_\delta(x) - f(x)| \le \begin{cases} \delta L_0, \ \textit{if } f \in C^{0,0}, \\ \delta^2 L_1, \ \textit{if } f \in C^{1,1}. \end{cases}$$

Proof. Recall that $f_\delta(x) = \mathbb{E}_{v\in\mathbb{B}}\big[f(x+\delta v)\big]$. Then, we have that

$$\begin{aligned} |f_\delta(x) - f(x)| &= |\mathbb{E}_{v\in\mathbb{B}}\big[f(x+\delta v) - f(x)\big]| \\ &\le \mathbb{E}_{v\in\mathbb{B}}\big[|f(x+\delta v) - f(x)|\big] \\ &\le \mathbb{E}_{v\in\mathbb{B}}\big[L_0\|\delta v\|\big]. \end{aligned} \tag{50}$$

Furthermore, since $v \in \mathbb{B}$, we have that $\|\delta v\| \le \delta$. Combining this inequality with (50), we have that $|f_\delta(x) - f(x)| \le \mathbb{E}_{v\in\mathbb{B}}\big[\delta L_0\big] = L_0\delta$. When the function $f \in C^{1,1}$ with Lipschitz constant L_1, we have that

$$\langle \nabla f(x), \delta v\rangle - \frac{L_1}{2}\|\delta v\|^2 \le f(x+\delta v) - f(x) \le \langle \nabla f(x), \delta v\rangle + \frac{L_1}{2}\|\delta v\|^2, \tag{51}$$

for all $v \in \mathbb{B}$. Taking the expectation of (51) over v sampled uniformly from the unit ball $\mathbb{B}$ and recalling that v is sampled independently from x and has zero mean, we get that

$$-L_1\delta^2 \le -\frac{L_1}{2}\mathbb{E}_{v\in\mathbb{B}}\big[\|\delta v\|^2\big] \le \mathbb{E}_{v\in\mathbb{B}}\big[f(x+\delta v) - f(x)\big] \le \frac{L_1}{2}\mathbb{E}_{v\in\mathbb{B}}\big[\|\delta v\|^2\big] \le L_1\delta^2. \tag{52}$$

In addition, because $|f_\delta(x) - f(x)| = |\mathbb{E}_{v\in\mathbb{B}}\big[f(x+\delta v) - f(x)\big]|$, we obtain that $|f_\delta(x) - f(x)| \le L_1\delta^2$. The proof is complete. □

The next lemma provides a bound on the second moment of the gradient estimate (44) under update (45).

Lemma H.6 (Second moment)**.** *Assume that $f_t \in C^{0,0}$ with Lipschitz constant L_0 for all time t. Then, under the ZO update rule in* (45)*, the second moment of the residual feedback* (44) *satisfies:*

$$\mathbb{E}[\|\widetilde{g}_t(x_t)\|^2] \le \frac{4d^2L_0^2\eta^2}{\delta^2}\mathbb{E}[\|\widetilde{g}_{t-1}(x_{t-1})\|^2] + D_t, \tag{53}$$

where $D_t := 16d^2L_0^2 + \frac{2d^2}{\delta^2}\mathbb{E}\big[\big(f_t(x_{t-1}+\delta u_{t-1}) - f_{t-1}(x_{t-1}+\delta u_{t-1})\big)^2\big]$.

Proof. By definition of the residual feedback (44), we have that

$$\begin{aligned}
\mathbb{E}[\|\tilde{g}_t(x_t)\|^2] &= \mathbb{E}[\frac{d^2}{\delta^2}\big(f_t(x_t+\delta u_t) - f_{t-1}(x_{t-1}+\delta u_{t-1})\big)^2\|u_t\|^2] \\
&\leq \frac{2d^2}{\delta^2}\mathbb{E}[\big(f_t(x_t+\delta u_t) - f_t(x_{t-1}+\delta u_{t-1})\big)^2\|u_t\|^2] \\
&\qquad + \frac{2d^2}{\delta^2}\mathbb{E}[\big(f_t(x_{t-1}+\delta u_{t-1}) - f_{t-1}(x_{t-1}+\delta u_{t-1})\big)^2\|u_t\|^2] \qquad (54) \\
&\leq \frac{2d^2}{\delta^2}\mathbb{E}[\big(f_t(x_t+\delta u_t) - f_t(x_{t-1}+\delta u_{t-1})\big)^2] \\
&\qquad + \frac{2d^2}{\delta^2}\mathbb{E}[\big(f_t(x_{t-1}+\delta u_{t-1}) - f_{t-1}(x_{t-1}+\delta u_{t-1})\big)^2],
\end{aligned}$$

where the last inequality is because $u_t \in \mathbb{S}$. Moreover, adding and subtracting $f_t(x_{t-1}+\delta u_t)$ to the term $\big(f_t(x_t+\delta u_t) - f_t(x_{t-1}+\delta u_{t-1})\big)^2$ in the inequality (54), we obtain

$$\begin{aligned}
\mathbb{E}[\|\tilde{g}(x_t)\|^2] \leq &\frac{4d^2}{\delta^2}\mathbb{E}[\big(f_t(x_t+\delta u_t) - f_t(x_{t-1}+\delta u_t)\big)^2] \\
&+ \frac{4d^2}{\delta^2}\mathbb{E}[\big(f_t(x_{t-1}+\delta u_t) - f_t(x_{t-1}+\delta u_{t-1})\big)^2] \qquad (55) \\
&+ \frac{2d^2}{\delta^2}\mathbb{E}[\big(f_t(x_{t-1}+\delta u_{t-1}) - f_{t-1}(x_{t-1}+\delta u_{t-1})\big)^2].
\end{aligned}$$

Since $f_t \in C^{0,0}$ is Lipschitz with constant L_0, we further obtain that

$$\begin{aligned}
\mathbb{E}[\|\tilde{g}(x_t)\|^2] \leq &\frac{4d^2L_0^2}{\delta^2}\mathbb{E}[\|x_t - x_{t-1}\|^2] + 4d^2L_0^2\mathbb{E}[\|u_t - u_{t-1}\|^2] \\
&+ \frac{2d^2}{\delta^2}\mathbb{E}[\big(f_t(x_{t-1}+\delta u_{t-1}) - f_{t-1}(x_{t-1}+\delta u_{t-1})\big)^2]. \qquad (56)
\end{aligned}$$

Since $u_t \in \mathbb{S}$, we get that $\mathbb{E}[\|u_t - u_{t-1}\|^2] \leq 4$. Substituting this bound into inequality (56), we obtain that

$$\begin{aligned}
\mathbb{E}[\|\tilde{g}(x_t)\|^2] \leq &\frac{4d^2L_0^2}{\delta^2}\mathbb{E}[\|x_t - x_{t-1}\|^2] + 16d^2L_0^2 \\
&+ \frac{2d^2}{\delta^2}\mathbb{E}[\big(f_t(x_{t-1}+\delta u_{t-1}) - f_{t-1}(x_{t-1}+\delta u_{t-1})\big)^2]. \qquad (57)
\end{aligned}$$

Since $x_t = \Pi_{(1-\xi)\mathcal{X}}\big[x_{t-1} - \eta\tilde{g}(x_{t-1})\big]$, we get that $\|x_t - x_{t-1}\| = \|\Pi_{(1-\xi)\mathcal{X}}\big[x_{t-1} - \eta\tilde{g}(x_{t-1})\big] - \Pi_{(1-\xi)\mathcal{X}}\big[x_{t-1}\big]\| \leq \eta\|\tilde{g}(x_{t-1})\|$ due to the nonexpansiveness of the projection operator onto a convex set. Therefore, we have that

$$\begin{aligned}
\mathbb{E}[\|\tilde{g}_t(x_t)\|^2] \leq &\frac{4d^2L_0^2\eta^2}{\delta^2}\mathbb{E}[\|\tilde{g}_{t-1}(x_{t-1})\|^2] + 16d^2L_0^2 \\
&+ \frac{2d^2}{\delta^2}\mathbb{E}[\big(f_t(x_{t-1}+\delta u_{t-1}) - f_{t-1}(x_{t-1}+\delta u_{t-1})\big)^2]. \qquad (58)
\end{aligned}$$

The proof is complete. □

Using Lemmas H.1-H.6, we can obtain the main theorem for online convex optimization using (45).

Theorem H.7 (Regret for Convex Lipschitz f_t)**.** *Let Assumption 3.1 hold. Assume that $f_t \in C^{0,0}$ is convex with Lipschitz constant L_0 for all t. Run ZO with residual feedback for $T > \bar{r}^2L_0^{2q}$ iterations with $\eta = \frac{\bar{r}^{\frac{3}{2}}}{2\sqrt{2}L_0\sqrt{d}T^{\frac{3}{4}}}$ and $\delta = \frac{\sqrt{\bar{r}d}}{L_0^qT^{\frac{1}{4}}}$, where $q \in \mathbb{R}$ is a user-specified parameter. Then, we have that*

$$\begin{aligned}
R_T \leq\ &4\sqrt{2\bar{r}d}L_0T^{\frac{3}{4}} + \frac{\mathbb{E}\big[\|\tilde{g}_0(x_0)\|^2\big]\bar{r}^{\frac{3}{2}}}{2\sqrt{2d}L_0T^{\frac{3}{4}}} + 8\sqrt{2}d^{\frac{3}{2}}L_0\bar{r}^{\frac{3}{2}}T^{\frac{1}{4}} \\
&+ (2+\frac{\bar{r}}{r})L_0^{1-q}\sqrt{d\bar{r}}T^{\frac{3}{4}} + \frac{\sqrt{2d\bar{r}}V_f^2}{L_0^{1-2q}}T^{\frac{3}{4}}. \qquad (59)
\end{aligned}$$

Asymptotically, we have $R_T = \mathcal{O}\big((L_0 + L_0^{1-q} + {L_0}^{2q-1}V_f^2)\sqrt{d\bar{r}}T^{\frac{3}{4}}\big)$.

Proof. First, we provide a bound on the regret that compares the sum of the function values obtained using (45) to that obtained for the optimizer x_ξ^* in the shrinked constraint set $(1-\xi)\mathcal{X}$, i.e., $\mathbb{E}\Big[\sum_{t=0}^{T-1} f_t(x_t) - \min_{x\in(1-\xi)\mathcal{X}} \sum_{t=0}^{T-1} f_t(x)\Big]$. Since $f_{\delta,t}(x)$ is convex for all t, we conclude that

$$f_{\delta,t}(x_t) - f_{\delta,t}(x) \leq \langle \nabla f_{\delta,t}(x_t), x_t - x\rangle, \text{ for all } x \in (1-\xi)\mathcal{X}. \tag{60}$$

Adding and subtracting $\tilde{g}_t(x_t)$ to $\nabla f_{\delta,t}(x_t)$ in inequality (60), and taking the expectation of both sides with respect to u_t, we obtain that

$$\mathbb{E}\big[f_{\delta,t}(x_t) - f_{\delta,t}(x)\big] \leq \mathbb{E}\big[\langle \tilde{g}_t(x_t), x_t - x\rangle\big]. \tag{61}$$

Since $x_{t+1} = \Pi_{(1-\xi)\mathcal{X}}\big[x_t - \eta\tilde{g}(x_t)\big]$, for any $x \in (1-\xi)\mathcal{X}$ we have that

$$\begin{aligned}\|x_{t+1} - x\|^2 &= \|\Pi_{(1-\xi)\mathcal{X}}\big[x_t - \eta\tilde{g}(x_t)\big] - \Pi_{(1-\xi)\mathcal{X}}\big[x\big]\|^2 \\ &\leq \|x_t - \eta\tilde{g}(x_t) - x\|^2 \\ &= \|x_t - x\|^2 - 2\eta\langle \tilde{g}_t(x_t), x_t - x\rangle + \eta^2\|\tilde{g}_t(x_t)\|^2.\end{aligned} \tag{62}$$

Rearranging the terms in inequality (62) yields

$$\langle \tilde{g}_t(x_t), x_t - x\rangle \leq \frac{1}{2\eta}\big(\|x_t - x\|^2 - \|x_{t+1} - x\|^2\big) + \frac{\eta}{2}\|\tilde{g}_t(x_t)\|^2. \tag{63}$$

Taking the expectation of both sides of inequality (63) with respect to u_t and substituting the resulting bound into (61), we obtain that

$$\mathbb{E}\Big[\sum_{t=0}^{T-1} f_{\delta,t}(x_t) - \sum_{t=0}^{T-1} f_{\delta,t}(x)\Big] \leq \frac{1}{2\eta}\|x_0 - x\|^2 + \frac{\eta}{2}\mathbb{E}\Big[\sum_{t=0}^{T-1}\|\tilde{g}_t(x_t)\|^2\Big]. \tag{64}$$

Since $f_t(x) \in C^{0,0}$, we know that $|f_{\delta,t}(x) - f_t(x)| \leq \delta L_0$. Therefore, we obtain

$$\begin{aligned}\mathbb{E}\big[\sum_{t=0}^{T-1} f_t(x_t) - \sum_{t=0}^{T-1} f_t(x)\big] &= \mathbb{E}\big[\sum_{t=0}^{T-1} f_{\delta,t}(x_t) - \sum_{t=0}^{T-1} f_{\delta,t}(x)\big] \\ &\quad + \mathbb{E}\big[\sum_{t=0}^{T-1}\big(f_t(x_t) - f_{\delta,t}(x_t)\big) - \sum_{t=0}^{T-1}\big(f_t(x) - f_{\delta,t}(x)\big)\big] \\ &\leq \frac{1}{2\eta}\|x_0 - x\|^2 + \frac{\eta}{2}\mathbb{E}\big[\sum_{t=0}^{T-1}\|\tilde{g}_t(x_t)\|^2\big] + 2L_0\delta T,\end{aligned} \tag{65}$$

where we have made use of the bound in (64). Telescoping the bound in (53) over $t = 1, 2, ..., T-1$, adding $\mathbb{E}\big[\|\tilde{g}_0(x_0)\|^2\big]$ to both sides, and adding $\frac{4d^2L_0^2\eta^2}{\delta^2}\mathbb{E}[\|\tilde{g}_{T-1}(x_{T-1})\|^2]$ to the right hand side, we obtain that

$$\mathbb{E}\big[\sum_{t=0}^{T-1}\|\tilde{g}_t(x_t)\|^2\big] \leq \frac{1}{1-\alpha}\mathbb{E}\big[\|\tilde{g}_0(x_0)\|^2\big] + \frac{16}{1-\alpha}d^2L_0^2T + \frac{2d^2V_f^2}{1-\alpha}\frac{1}{\delta^2}T, \tag{66}$$

where $\alpha = \frac{4d^2L_0^2\eta^2}{\delta^2}$. Substituting the bound in (66) into (65) yields

$$\begin{aligned}\mathbb{E}\big[\sum_{t=0}^{T-1} f_t(x_t) - \sum_{t=0}^{T-1} f_t(x)\big] &\leq \frac{1}{2\eta}\|x_0 - x\|^2 + \frac{\eta}{2(1-\alpha)}\mathbb{E}\big[\|\tilde{g}_0(x_0)\|^2\big] + \frac{16}{1-\alpha}d^2L_0^2\eta T \\ &\quad + 2L_0\delta T + \frac{2d^2V_f^2}{1-\alpha}\frac{\eta}{\delta^2}T.\end{aligned} \tag{67}$$

Since inequality (67) holds for all $x \in (1-\xi)\mathcal{X}$, we can replace x in (67) with x_ξ^*. Furthermore, using Lemma H.4, we have that

$$\sum_{t=0}^{T-1} f_t(x_\xi^*) - \sum_{t=0}^{T-1} f_t(x^*) \leq \bar{r}L_0\xi T. \tag{68}$$

Summing inequalities (67) and (68), we obtain

$$\mathbb{E}\big[\sum_{t=0}^{T-1} f_t(x_t) - \sum_{t=0}^{T-1} f_t(x^*)\big] \leq \frac{1}{2\eta}\|x_0 - x_\xi^*\|^2 + \frac{\eta}{2(1-\alpha)}\mathbb{E}\big[\|\tilde{g}_0(x_0)\|^2\big] + \frac{16}{1-\alpha}d^2L_0^2\eta T + 2L_0\delta T + \frac{2d^2V_f^2}{1-\alpha}\frac{\eta}{\delta^2}T + \bar{r}L_0\xi T, \tag{69}$$

where $\|x_0 - x_\xi^*\|^2 \leq 4\bar{r}^2$. According to Lemma H.3, we can select $\xi = \frac{\delta}{r}$ to guarantee that all iterates $x_t + \delta u_t \in \mathcal{X}$ for all $u_t \in \mathbb{S}$. Furthermore, let $\eta = \frac{\bar{r}^{\frac{3}{2}}}{2\sqrt{2}L_0\sqrt{d}T^{\frac{3}{4}}}$ and $\delta = \frac{\sqrt{\bar{r}d}}{L_0^qT^{\frac{1}{4}}}$, where $q \in \mathbb{R}$ is a user-specified parameter. Then, $\alpha = \frac{4d^2L_0^2\eta^2}{\delta^2} = \frac{1}{2T}\bar{r}^2L_0^{2q} \leq \frac{1}{2}$ when $T \geq \bar{r}^2L_0^{2q}$. Substituting these parameter values into (69), we obtain that

$$\mathbb{E}\big[\sum_{t=0}^{T-1} f_t(x_t) - \sum_{t=0}^{T-1} f_t(x^*)\big] \leq 4\sqrt{2\bar{r}d}L_0T^{\frac{3}{4}} + \frac{\mathbb{E}\big[\|\tilde{g}_0(x_0)\|^2\big]\bar{r}^{\frac{3}{2}}}{2\sqrt{2d}L_0T^{\frac{3}{4}}} + 8\sqrt{2}d^{\frac{3}{2}}L_0\bar{r}^{\frac{3}{2}}T^{\frac{1}{4}} + (2+\frac{\bar{r}}{r})L_0^{1-q}\sqrt{d\bar{r}}T^{\frac{3}{4}} + L_0^{2q-1}\sqrt{2d\bar{r}}V_f^2T^{\frac{3}{4}}. \tag{70}$$

The proof is complete. □

I Discussion on the Online Optimization with Adversaries

In Section 2, we consider online optimization problems where the sequence of the objective functions $\{f_t\}_t$ is randomly generated and is independent of the agent's decisions. This assumption is satisfied when the non-stationarity of the environment is caused by the nature. In this section, we consider a different scenario where the objective function is selected by an opponent. Specifically, at time t, the agent selects a decision $x_t + \delta u_t$, then the opponent selects a objective function f_t according to the history information $H_t = \{x_0 + \delta u_0, f_0, \ldots, x_{t-1} + \delta u_{t-1}, f_{t-1}, x_t + \delta u_t\}$ to maximize the agent's regret.

When the gradient estimator (3) is applied, where the searching direction u_t is sampled from Gaussian distribution $\mathcal{N}(0, I)$, we have the following Lemma in adversarial scenario.

Lemma I.1 (Second moment)**.** *Assume that $f_t \in C^{0,0}$ with Lipschitz constant L_0 for all time t. Then, under the ZO update rule in (4), the second moment of the residual feedback satisfies: for all t,*

$$\mathbb{E}[\|\widetilde{g}_t(x_t)\|^2] \leq \frac{4dL_0^2\eta^2}{\delta^2}\mathbb{E}[\|\widetilde{g}_{t-1}(x_{t-1})\|^2] + D_t, \tag{71}$$

where $D_t := 16L_0^2(d+4)^2 + \frac{2}{\delta^2}\mathbb{E}\big[\big(f_t(x_{t-1}+\delta u_{t-1}) - f_{t-1}(x_{t-1}+\delta u_{t-1})\big)^2\|u_t\|^2\big].$

Proof. The proof is essentially the same as the proof of Lemma 2.5, except that the bound $\frac{2}{\delta^2}\mathbb{E}\big[\big(f_t(x_{t-1}+\delta u_{t-1}) - f_{t-1}(x_{t-1}+\delta u_{t-1})\big)^2\|u_t\|^2\big] \leq \frac{2d}{\delta^2}\mathbb{E}\big[\big(f_t(x_{t-1}+\delta u_{t-1}) - f_{t-1}(x_{t-1}+\delta u_{t-1})\big)^2\big]$ used under (14) does not apply in the adversary case, because the selection of the function f_t depends on u_t. Since the other derivations in the proof of Lemma 2.5 does not rely on the independence between u_t and f_t, they still hold. It is straightforward to obtain the bound in (71). □

Next, we present the assumptions on the adversary agent for online convex optimization problems.

Assumption I.2 (Bounded Adversary)**.** *Given the history H_t, the adversary agent selects a function f_t such that for all time t there exists a constant V_f^2 that satisfies*

$$|f_t(x_{t-1}+\delta u_{t-1}) - f_{t-1}(x_{t-1}+\delta u_{t-1})|^2 \leq V_f^2. \tag{72}$$

Then, within the expectation term in D_t in the bound (71), for any realization of the random vector u_t, the bound $\big(f_t(x_{t-1}+\delta u_{t-1}) - f_{t-1}(x_{t-1}+\delta u_{t-1})\big)^2 \leq V_f^2$ holds according to Assumption I.2. Therefore, we have that

$$\frac{2}{\delta^2}\mathbb{E}\big[\big(f_t(x_{t-1}+\delta u_{t-1}) - f_{t-1}(x_{t-1}+\delta u_{t-1})\big)^2\|u_t\|^2\big] \leq \frac{2}{\delta^2}\mathbb{E}\big[V_f^2\|u_t\|^2\big] \leq \frac{2d}{\delta^2}V_f^2. \tag{73}$$

Therefore, after combining Lemma I.1 and Assumption I.2, we can achieve the bound on the second moment $\mathbb{E}\big[\|\tilde{g}_t(x_t)\|^2\big]$

$$\mathbb{E}\big[\|\tilde{g}_t(x_t)\|^2\big] \leq \frac{4dL_0^2\eta^2}{\delta^2}\mathbb{E}[\|\widetilde{g}_{t-1}(x_{t-1})\|^2] + 16L_0^2(d+4)^2 + \frac{2d}{\delta^2}V_f^2. \tag{74}$$

This is the same bound we obtained by combining Lemma 2.5 and Assumption 3.1. And it can be used to obtain (23) in the proofs of Theorems 3.2, which is also used in 3.4. Then, it is straightforward to follow the same proofs of Theorems 3.2 and 3.4 to get the same regret bounds in online convex optimization problems under adversarial environment.

Finally, we present the assumptions on the adversary agent for non-stationary non-convex optimization problems.

Assumption I.3. *From time $t = 0$ to T, the adversary agent selects a sequence of objective functions $\{f_t\}$ such that*

1. *There exists a constant W_T that satisfies $\sum_{t=1}^{T} \mathbb{E}[f_{\delta,t}(x_t) - f_{\delta,t-1}(x_t)] \leq W_T$, where the expectation is taken with respect to x_t.*
2. *At time $t \geq 1$, given the history H_t, the adversary agent selects a function f_t such that there exists a constant $V_{f,t}^2$ that satisfies*

$$|f_t(x_{t-1} + \delta u_{t-1}) - f_{t-1}(x_{t-1} + \delta u_{t-1})|^2 \leq V_{f,t}^2. \tag{75}$$

Furthermore, we have that

$$\sum_{t=1}^{T} V_{f,t}^2 \leq \widetilde{W}_T. \tag{76}$$

Different from Assumption I.2, where at each time t, the adversary should select a function f_t according to a uniform function variation bound V_f^2, Assumption I.3.2 allows the adversary to select f_t according to a varying function variation bound $V_{f,t}^2$. However, there also exists a budget $\widetilde{W}_T$ for the adversary, which represents the total variation on the functions that the adversary is allowed to make from time $t = 0$ to T.

Then, similar to the discussion under Assumption I.2, within the expectation term in D_t in the bound (71), for any realization of the random vector u_t, the bound $\big(f_t(x_{t-1} + \delta u_{t-1}) - f_{t-1}(x_{t-1} + \delta u_{t-1})\big)^2 \leq V_{f,t}^2$ holds at time t according to Assumption I.3. Therefore, we can combine Lemma I.1 and Assumption I.3 and use similar derivation in (73) to achieve the same bounds in (34), (35) and (38), which are used in the proof of Theorems 4.2 and 4.3. The other part of the proofs remains the same. Therefore, by combining Lemma I.1 and Assumption I.3, we achieve the same regret bounds in Theorems 4.2 and 4.3 in online non-stationary non-convex optimization problems under adversarial environment.