Improve Single-Point Zeroth-Order Optimization Using High-Pass and Low-Pass Filters

Xin Chen [1] **Yujie Tang** [1] **Na Li** [1]

Abstract

Single-point zeroth-order optimization (SZO) is useful in solving online black-box optimization and control problems in time-varying environments, as it queries the function value only once at each time step. However, the vanilla SZO method is known to suffer from a large estimation variance and slow convergence, which seriously limits its practical application. In this work, we borrow the idea of high-pass and low-pass filters from extremum seeking control (continuous-time version of SZO) and develop a novel SZO method called HLF-SZO by integrating these filters. It turns out that the high-pass filter coincides with the residual feedback method, and the low-pass filter can be interpreted as the momentum method. As a result, the proposed HLF-SZO achieves a much smaller variance and much faster convergence than the vanilla SZO method, and empirically outperforms the residual-feedback SZO method, which are verified via extensive numerical experiments.

1. Introduction

This paper considers solving the generic unconstrained optimization problem:

$$\min_{\boldsymbol{x}\in\mathbb{R}^d} f(\boldsymbol{x}), \tag{1}$$

where $\boldsymbol{x} \in \mathbb{R}^d$ is the decision variable and $f : \mathbb{R}^d \to \mathbb{R}$ is the objective function. A straightforward solution scheme is to apply the gradient descent method (Ruder, 2016), e.g., $\boldsymbol{x} \leftarrow \boldsymbol{x} - \eta\nabla f(\boldsymbol{x})$ with the step size η. However, in many practical applications, the first-order information of function f, i.e., the gradient $\nabla f(\boldsymbol{x})$, may be unavailable or too expensive to procure, and one can only access the zeroth-order information, i.e., function evaluations. To this end, zeroth-order (or derivative-free) methods (Nesterov & Spokoiny, 2017) are developed to solve black-box optimization and control problems, which essentially estimate the gradients using perturbed function evaluations. Zeroth-order optimization (ZO) has attracted a great deal of recent attention and has been used for a broad spectrum of applications, such as reinforcement learning (Malik et al., 2019; Li et al., 2019), adversarial training (Chen et al., 2017), physical system control (Chen et al., 2021b; 2020), online sensor management (Liu et al., 2020), etc.

According to the number of queried function evaluations at each iteration, ZO methods can be categorized into two types: single-point and multi-point (Liu et al., 2020). As suggested by the name, single-point ZO (SZO) (Flaxman et al., 2005) only needs to query the function value once at each iteration, making it particularly suitable for online optimization and control problems. In contrast, multi-point ZO, such as two-point ZO (Shamir, 2017), requires two or more function evaluations in the same instantaneous time; this may not be practical when the environment is non-stationary or changed by the implementation of a function evaluation. For example, two-point ZO can not be applied to non-stationary reinforcement learning problems, because it needs two different policy evaluations in the same environment, which is impossible since the environment changes after each policy evaluation (Zhang et al., 2021). Nevertheless, it is known that the vanilla SZO method (Flaxman et al., 2005) suffers from a large estimation variance and slow convergence, which seriously jeopardizes its practical application. There have been multiple studies (Saha & Tewari, 2011; Dekel et al., 2015; Gasnikov et al., 2017; Zhang et al., 2021) on improving the convergence rate of SZO. In particular, the residual-feedback SZO method proposed in the recent work Zhang et al. (2021) achieves the state-of-the-art performance to date.

In the field of control, there is a continuous-time version of SZO known as *extremum seeking (ES) control* (Ariyur & Krstic, 2003; Tan et al., 2010). ES control is a classic adaptive control technique that uses only output feedback to steer a dynamical system to a state where the objective function attains an extremum. ES control can also be adopted

[1]John A. Paulson School of Engineering and Applied Sciences, Harvard University, MA, US. Correspondence to: Xin Chen <chenxin2336@gmail.com>.

Proceedings of the 39^{th} International Conference on Machine Learning, Baltimore, Maryland, USA, PMLR 162, 2022.

Table 1. The iteration complexity of ZO methods for solving Lipschitz and smooth objective functions (1).

METHOD	LITERATURE	COMPLEXITY	
		CONVEX	NON-CONVEX
VANILLA SZO	GASNIKOV ET AL. (2017)	$\mathcal{O}(d^2/\epsilon^3)$	–
RESIDUAL-FEEDBACK SZO	ZHANG ET AL. (2021)	$\mathcal{O}(d^3/\epsilon^{\frac{3}{2}})$	$\mathcal{O}(d^3/\epsilon^{\frac{3}{2}})$
HLF-SZO	THIS WORK	$\mathcal{O}(d^{\frac{3}{2}}/\epsilon^{\frac{3}{2}})$	$\mathcal{O}(d^{\frac{3}{2}}/\epsilon^{\frac{3}{2}})$
TWO-POINT ZO	NESTEROV & SPOKOINY (2017)	$\mathcal{O}(d/\epsilon)$	$\mathcal{O}(d/\epsilon)$

*In the convex setting, the accuracy is measured by $f(\overline{\boldsymbol{x}}_T)-f(\boldsymbol{x}^*)\leq\epsilon$, where $\boldsymbol{x}^*\in\arg\min_{\boldsymbol{x}\in\mathbb{R}^d}f(\boldsymbol{x})$ and $\overline{\boldsymbol{x}}_T=\frac{1}{T}\sum_{k=1}^T\boldsymbol{x}_k$. In the non-convex setting, the accuracy is measured by $\frac{1}{T}\sum_{k=1}^T\|\nabla f(\boldsymbol{x}_k)\|^2\leq\epsilon$.

as a zeroth-order algorithm to solve static-map optimization problems (Poveda & Li, 2021; Dürr et al., 2013; Ye & Hu, 2016). In addition, although not essential for the ES system operation, a *high-pass filter* and a *low-pass filter* are usually integrated into the control loop, because these filters can significantly improve the transient behavior and mitigate oscillations (see Section 2.2 for detailed explanations). Despite their close connection, ES control and SZO have been mostly studied separately in the control and optimization communities. Then, a natural question to ask is

"Can we borrow some tools from extremum seeking control, such as the high-pass and low-pass filters, to improve the performance of single-point zeroth-order optimization?"

Contributions. Motivated by this question, we develop a novel SZO method called **HLF-SZO** (**H**igh/**L**ow-pass **F**ilter **SZO**) by integrating a high-pass filter and a low-pass filter into the vanilla SZO method. The main contributions of this paper are explained below:

1) We find that the integration of a high-pass filter can be interpreted as the residual feedback scheme proposed in the recent work Zhang et al. (2021), which can greatly reduce the variance of SZO and lead to faster convergence. And the integration of a low-pass filter can be interpreted as the momentum (heavy-ball) optimization method (Polyak, 1964; Qian, 1999), which can further accelerate the convergence.

2) We prove that the proposed HLF-SZO method achieves the iteration complexity of $\mathcal{O}(d^{\frac{3}{2}}/\epsilon^{\frac{3}{2}})$ for Lipschitz and smooth objective functions in both convex and nonconvex cases. This iteration complexity is better than the complexity $\mathcal{O}(d^2/\epsilon^3)$ of the vanilla SZO method (Gasnikov et al., 2017), and has better dependency on the problem dimension d compared with the complexity $\mathcal{O}(d^3/\epsilon^{\frac{3}{2}})$ of the residual-feedback SZO method (Zhang et al., 2021), although it is inferior to the complexity $\mathcal{O}(d/\epsilon)$ of two-point methods. See Table 1 for a summary of the best known iteration complexity of ZO methods.

3) Extensive numerical experiments show that the proposed HLF-SZO method exhibits a much smaller variance and much faster convergence than the vanilla SZO; it empirically outperforms the residual-feedback SZO and has comparable performance to the two-point ZO method.

Moreover, this paper explores a new direction to improve ZO schemes by leveraging the connection between ZO and continuous-time ES control, and we exemplify the possibility that useful tools from ES control, such as high-pass and low-pass filters, can indeed help boost SZO. There is much more to study in this direction. For example, this paper only adopts the simplest high-pass and low-pass filters, while higher-order filters or compensators that are used in ES control to enhance stability (Krstić, 2000) may also be applied to further improve the performance of ZO.

Other Related Work on ZO. The work (Novitskii & Gasnikov, 2021) leverages the higher-order smoothness and proves an improved complexity bound of SZO for solving stochastic convex optimization problems. References (Jongeneel et al., 2021; Jongeneel, 2021) propose new smoothed gradient approximation schemes by using complex analysis tools from numerical differentiation, which boost the convergence of ZO and address the numerical cancellation issue. In Wang & Spall (2021); Wang et al. (2021), a gradient estimation method that uses only the system measurements is developed based on the simultaneous perturbation stochastic approximation algorithm and the complex-step gradient approximation. Berahas et al. (2022) presents a theoretical and empirical comparison of several gradient approximation methods for derivative-free optimization, including finite difference, linear interpolation, Gaussian smoothing, etc.

Notations. Let $\mathbb{B}_d\coloneqq\{\boldsymbol{x}\in\mathbb{R}^d:\|\boldsymbol{x}\|_2\leq 1\}$ denote the closed unit ball of dimension d, and let $\mathbb{S}_{d-1}\coloneqq\{\boldsymbol{x}\in\mathbb{R}^d:\|\boldsymbol{x}\|_2=1\}$ denote the unit sphere.

2. Preliminaries on ZO and ES Control

2.1. Zeroth-Order Optimization

The SZO method estimates the gradient $\nabla f(\boldsymbol{x})$ with only one query of the function value at each iteration step. Specif-

ically, upon defining

$$\mathsf{G}_f^{(1)}(\boldsymbol{x};r,\boldsymbol{u}) := \frac{d}{r} f(\boldsymbol{x}+r\boldsymbol{u})\boldsymbol{u}, \tag{2}$$

for $\boldsymbol{x},\boldsymbol{u} \in \mathbb{R}^d$ and $r > 0$, one can show that if $\boldsymbol{u}$ is uniformly sampled from the unit sphere $\mathbb{S}_{d-1}$, the term $\mathsf{G}_f^{(1)}(\boldsymbol{x};r,\boldsymbol{u})$ is an unbiased estimator for the gradient of a smoothed version of function f (Flaxman et al., 2005):[1]

$$\mathbb{E}_{\boldsymbol{u}\sim\mathrm{Unif}(\mathbb{S}_{d-1})}\big[\mathsf{G}_f^{(1)}(\boldsymbol{x};r,\boldsymbol{u})\big] = \nabla f_r(\boldsymbol{x}),$$

where $f_r(\boldsymbol{x}) := \mathbb{E}_{\bar{\boldsymbol{u}}\sim\mathrm{Unif}(\mathbb{B}_d)}\big[f(\boldsymbol{x}+r\bar{\boldsymbol{u}})\big]$. By plugging the single-point gradient estimator $\mathsf{G}_f^{(1)}(\boldsymbol{x};r,\boldsymbol{u})$ into the gradient descent iterations for solving (1), we obtain the vanilla SZO method (3):

$$(\textbf{Vanilla SZO}) : \boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \eta\cdot\frac{d}{r} f(\boldsymbol{x}_k+r\boldsymbol{u}_k)\boldsymbol{u}_k, \tag{3}$$

where $\{\boldsymbol{u}_k : k = 0,1,2,\cdots\}$ are i.i.d. random directions sampled from the uniform distribution $\mathrm{Unif}(\mathbb{S}_{d-1})$. Here, $\eta > 0$ is the step size and the parameter $r > 0$ is called smoothing radius. One can show the convergence of (3) by appropriately choosing the parameters η and r under some regular conditions (Gasnikov et al., 2017).

However, the single-point gradient estimator (2) generally suffers from a high estimation variance, which leads to slow convergence of the vanilla SZO method (3). To overcome this issue and achieve faster convergence, two-point ZO methods (Nesterov & Spokoiny, 2017) are developed as:

$$(\textbf{Two-point ZO}) :$$

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \eta\frac{d}{2r}\big(f(\boldsymbol{x}_k+r\boldsymbol{u}_k) - f(\boldsymbol{x}_k-r\boldsymbol{u}_k)\big)\boldsymbol{u}_k, \tag{4}$$

$$\text{(or)}\ \ \boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \eta\frac{d}{r}\big(f(\boldsymbol{x}_k+r\boldsymbol{u}_k) - f(\boldsymbol{x}_k)\big)\boldsymbol{u}_k, \tag{5}$$

where two function evaluations are needed in each iteration. Both of the two-point gradient estimators $\mathsf{G}_f^{(2)}(\boldsymbol{x};r,\boldsymbol{u}) := \frac{d}{2r}(f(\boldsymbol{x}+r\boldsymbol{u})-f(\boldsymbol{x}-r\boldsymbol{u}))\boldsymbol{u}$ and $\tilde{\mathsf{G}}_f^{(2)}(\boldsymbol{x};r,\boldsymbol{u}) := \frac{d}{r}(f(\boldsymbol{x}+r\boldsymbol{u}) - f(\boldsymbol{x}))\boldsymbol{u}$ share the same expectation as $\mathsf{G}_f^{(1)}(\boldsymbol{x};r,\boldsymbol{u})$, but they generally have smaller variances and thus lead to faster convergence of the two-point ZO methods (4) (5).

2.2. Extremum Seeking (ES) Control

Extremum seeking control is a type of model-free control that uses only output feedback to steer a dynamical system to the extremum of an unknown function (Ariyur & Krstic, 2003). A simple ES scheme for solving the scalar version

[1] Another choice of the distribution to sample the random direction $\boldsymbol{u}$ is the Gaussian distribution $\mathcal{N}(\boldsymbol{0},\boldsymbol{I}/d)$. We use $\mathrm{Unif}(\mathbb{S}_{d-1})$ in this paper for the simplicity of exposition, while the results can be also extended to the Gaussian distribution case.

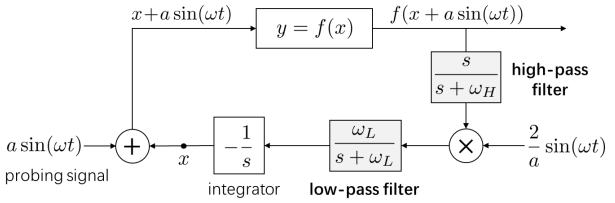

Figure 1. The block diagram of ES system with a high-pass filter $\frac{s}{s+\omega_H}$ and a low-pass filter $\frac{\omega_L}{s+\omega_L}$ for solving $\min_x f(x)$.

of problem (1) is illustrated as Figure 1. Essentially, the ES system adopts a small sinusoidal probing signal $a\sin(\omega t)$ as the perturbation to estimate the gradient $\nabla f(x)$, where positive parameters a and ω are the amplitude and frequency.

To analyze the ES system shown in Figure 1, we can first ignore the high-pass filter and the low-pass filter, as they are not essential to the system operation. Following the block diagram in Figure 1, we start from the state x and add a sinusoidal probing signal $a\sin(\omega t)$ to it, and the resultant value $x + a\sin\omega t$ is taken as the input to the static map $y = f(x)$. Then, the output $y = f(x + a\sin(\omega t))$ is multiplied by $\frac{2}{a}\sin(\omega t)$ for correlation and the loop is closed after passing through the integrator $-\frac{1}{s}$. As a consequence, the ES system dynamics can be formulated as

$$(\textbf{Extremum Seeking Dynamics}) :$$

$$\dot{x} = -\frac{2}{a} f\big(x(t) + a\sin(\omega t)\big)\sin(\omega t). \tag{6}$$

The formulation of the ES dynamics (6) is quite similar to the vanilla SZO method (3), except that a sinusoidal probing signal is used instead of random directions. Besides, the ES scheme can be easily extended to the multivariate case by using an appropriate probing signal vector with different frequencies (Chen et al., 2021a; Poveda & Li, 2021).

The rationale behind (6) is that with a small a and a large ω, the ES dynamics (6) behaves approximately like the gradient descent flow $\dot{x} = -\nabla f(x)$, which steers x to a (local) minimum $x^* \in \arg\min_x f(x)$ under regular conditions (Chill & Fašangová, 2010). Specifically, when ω is large, the ES dynamics (6) exhibits a timescale separation property, where the fast-time variation results from the high-frequency sinusoidal signal $\sin(\omega t)$, while the slow-time variation induced by the integrator $-\frac{1}{s}$ dominates the evolution of x. As a is small, we consider the Taylor expansion of the output $f(x + a\sin(\omega t))$ around x:

$$f(x+a\sin(\omega t)) = f(x) + a\sin(\omega t)\nabla f(x) + \mathcal{O}(a^2). \tag{7}$$

Then to analyze the dominant slow-time variation, one can compute the average dynamics of (6) to wash out the peri-

odic fast-time variation, which is $\dot{x} = -h_{\text{ave}}(x)$ with

$$\begin{aligned} h_{\text{ave}}(x) &:= \frac{1}{T}\int_0^T \frac{2}{a} f(x + a\sin(\omega t))\sin(\omega t)\, dt \\ &= \frac{1}{T}\int_0^T \frac{2}{a} f(x)\sin(\omega t) + 2\sin^2(\omega t)\nabla f(x) + \mathcal{O}(a)\, dt \\ &= \nabla f(x) + \mathcal{O}(a), \end{aligned} \tag{8}$$

where $T = \frac{2\pi}{\omega}$. Equation (8) reveals that the average dynamics of (6) is actually the gradient descent flow plus a small perturbation term $\mathcal{O}(a)$. See Ariyur & Krstic (2003); Tan et al. (2010) for more explanations.

Remark 2.1. (**Add High-Pass and Low-Pass Filters**). As shown in Figure 1, a high-pass filter $\frac{s}{s+\omega_H}$ and a lower-pass filter $\frac{\omega_L}{s+\omega_L}$ are integrated into the ES system to mitigate oscillations and improve the transient behavior, where ω_H and ω_L are the cut-off frequency parameters. Intuitively, the high-pass filter serves to wash out the constant and low-frequency components of the output signal $f(x+a\sin(\omega t))$. As indicated in the integral (8), the first term $f(x)$ in the Taylor expansion (7) is useless and causes large oscillations in the transient process, and it is slowly-varying and thus can be removed by a high-pass filter. In terms of the low-pass filter, it works to wash out high-frequency oscillations (induced by the sinusoidal probing signal) of the gradient estimation, because a clean and constant gradient estimator is desirable before passing through the integrator. More explanations are provided in Tan et al. (2010; 2006).

3. Algorithm Design

In this section, we first present the proposed SZO algorithms that incorporate high-pass and low-pass filters, and then describe the derivation process.

3.1. SZO Methods with High-Pass/Low-Pass Filters

3.1.1. Integrate a High-Pass Filter

We mimic how ES control applies the high-pass filter and integrate it into the vanilla SZO (3), leading to the following new SZO method (9):

$$\begin{aligned} &(\textbf{HF-SZO}): \qquad \forall k = 1, 2, \cdots \\ &\begin{cases} z_k = (1-\beta)z_{k-1} + f(\boldsymbol{x}_k + r\boldsymbol{u}_k) - f(\boldsymbol{x}_{k-1} + r\boldsymbol{u}_{k-1}) \\ \boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \eta \cdot \dfrac{d}{r} z_k \boldsymbol{u}_k, \end{cases} \end{aligned} \tag{9}$$

where $z_k \in \mathbb{R}$ is an intermediate variable and $\beta \geq 0$ is an adjustable parameter. Particularly, when setting $\beta = 1$, the HF-SZO (9) becomes the residual-feedback SZO method proposed in Zhang et al. (2021); and (9) reduces to the vanilla SZO (3) when $\beta = 0$ and $z_0 = f(\boldsymbol{x}_0 + r\boldsymbol{u}_0)$.

3.1.2. Integrate a Low-Pass Filter

Similarly, by integrating a low-pass filter into the vanilla SZO (3), we obtain a new SZO method (10):

$$\begin{aligned} &(\textbf{LF-SZO}): \qquad \forall k = 1, 2, \cdots \\ &\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \eta \frac{d}{r} f(\boldsymbol{x}_k + r\boldsymbol{u}_k)\boldsymbol{u}_k + \alpha(\boldsymbol{x}_k - \boldsymbol{x}_{k-1}), \end{aligned} \tag{10}$$

where $\alpha \in [0, 1]$ is an adjustable parameter.

Compared with (3), the LF-SZO method (10) has an additional "momentum" term $\alpha(\boldsymbol{x}_k - \boldsymbol{x}_{k-1})$ that is introduced by the low-pass filter. When setting $\alpha = 0$, the LF-SZO method (10) reduces to the vanilla SZO method (3).

3.1.3. Integrate Both a High-Pass Filter and a Low-Pass Filter

By integrating both a high-pass filter and a low-pass filter into (3), we develop the HLF-SZO method (11):

$$\begin{aligned} &(\textbf{HLF-SZO}): \qquad \forall k = 1, 2, \cdots \\ &\begin{cases} z_k = (1-\beta)z_{k-1} + f(\boldsymbol{x}_k + r\boldsymbol{u}_k) - f(\boldsymbol{x}_{k-1} + r\boldsymbol{u}_{k-1}) \\ \boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \eta \cdot \dfrac{d}{r} z_k \boldsymbol{u}_k + \alpha(\boldsymbol{x}_k - \boldsymbol{x}_{k-1}), \end{cases} \end{aligned} \tag{11}$$

which is basically the combination of (9) and (10).

Note that in the HLF-SZO (11), only one function evaluation $f(\boldsymbol{x}_k + r\boldsymbol{u}_k)$ is queried at each iteration k, while the value $f(\boldsymbol{x}_{k-1} + r\boldsymbol{u}_{k-1})$ is directly inherited from the last iteration, which is the same for the HF-SZO (9). As for the choices of adjustable parameters α and β, the theoretical analysis in Section 4 shows that $\beta = 1$ is optimal for convergence, which is also validated by the numerical experiments conducted in Section 5. Besides, a momentum parameter $\alpha = 0.9$ or a similar value is suggested (Ruder, 2016) and there have been extensive studies (Polyak, 1964; Tao et al., 2021; Qian, 1999) on the role and selection of the momentum parameter. See Section 4 for more discussions.

Remark 3.1. High-pass and low-pass filters are classic tools in the field of control and signal processing. It is interesting to find their connections to optimization approaches. Specifically, the integration of a *high-pass filter* coincides with the *residual-feedback method* proposed in Zhang et al. (2021), which can significantly reduce the estimation variance of SZO methods. This is also consistent with the function of a high-pass filter explained in Remark 2.1. In addition, the integration of a *low-pass filter* can be interpreted as the *momentum (heavy-ball) method* (Polyak, 1964), which can accelerate the convergence. These observations are further explained in Section 4 and are verified via extensive numerical experiments in Section 5.

3.2. Detailed Derivation Process

For the ES system illustrated in Figure 1, we denote f as the output of the static map, and let z be the signal after passing f through the high-pass filter $\frac{s}{s+\omega_H}$. Then we have

$$\mathcal{L}\{z\} = \frac{s}{s+\omega_H}\mathcal{L}\{f\} \iff \dot{z} + \omega_H z = \dot{f}, \tag{12}$$

where $\mathcal{L}\{\cdot\}$ denotes the Laplacian transform. We discretize the continuous-time dynamics (12) under the time gap δ:

$$\begin{aligned}&\frac{z_k - z_{k-1}}{\delta} + \omega_H z_{k-1} = \frac{f_k - f_{k-1}}{\delta}\\ &\iff \quad z_k = (1-\delta\omega_H)z_{k-1} + f_k - f_{k-1},\end{aligned} \tag{13}$$

where $f_k := f(\boldsymbol{x}_k + r\boldsymbol{u}_k)$. Essentially, z_k can be regarded as the refined value of f_k after passing it through the high-pass filter. Denote $\beta := \delta\omega_H \geq 0$. Then replacing f_k by z_k in the vanilla SZO (3) leads to the HF-SZO method (9).

As shown in Figure 1, a low-pass filter $\frac{\omega_L}{s+\omega_L}$ is added before passing the gradient estimation to the integrator. We denote $\boldsymbol{g} \in \mathbb{R}^d$ as the gradient estimator and $\boldsymbol{y} \in \mathbb{R}^d$ as the refined signal after passing $\boldsymbol{g}$ through the low-pass filter. Then we have the relation:

$$\mathcal{L}\{\boldsymbol{y}\} = \frac{\omega_L}{s+\omega_L}\mathcal{L}\{\boldsymbol{g}\} \iff \dot{\boldsymbol{y}} = \omega_L(-\boldsymbol{y} + \boldsymbol{g}). \tag{14}$$

We discretize the resultant dynamics (14) and the integrator dynamics $\dot{\boldsymbol{x}} = -\boldsymbol{y}$ under the time gap δ, and obtain

$$\begin{aligned}&\frac{\boldsymbol{y}_{k+1} - \boldsymbol{y}_k}{\delta} = \omega_L(-\boldsymbol{y}_k + \boldsymbol{g}_k)\\ &\iff \quad \boldsymbol{y}_{k+1} = (1-\delta\omega_L)\boldsymbol{y}_k + \delta\omega_L\boldsymbol{g}_k,\end{aligned} \tag{15a}$$

$$\frac{\boldsymbol{x}_{k+1} - \boldsymbol{x}_k}{\delta} = -\boldsymbol{y}_{k+1} \iff \boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \delta\boldsymbol{y}_{k+1}. \tag{15b}$$

Denote $\eta := \delta^2\omega_L$ and $\alpha := 1 - \delta\omega_L$. After eliminating the variable $\boldsymbol{y}$ from (15) and letting $\boldsymbol{g}_k = \frac{d}{r}f(\boldsymbol{x}_k + r\boldsymbol{u}_k)\boldsymbol{u}_k$ as in the vanilla SZO (3), we obtain the LF-SZO (10).

When both a high-pass filter and a low-pass filter are applied, we obtain the HLF-SZO method (11), which is basically the combination of (13) and (15).

4. Theoretic Analysis

In this section, we analyze the convergence of the proposed HLF-SZO method (11). We make the following assumption throughout our analysis.

Assumption 4.1. The function f is G-Lipschitz and L-smooth, i.e., for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$, we have

$$|f(\boldsymbol{x}) - f(\boldsymbol{y})| \leq G\|\boldsymbol{x} - \boldsymbol{y}\|, \ \|\nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y})\| \leq L\|\boldsymbol{x} - \boldsymbol{y}\|.$$

Note that Assumption 4.1 is mainly for theoretical analysis, and the proposed HLF-SZO works well for a wide range of problems that may not satisfy this assumption, as verified by the numerical experiments in Section 5.

We first consider the case when the objective function f is convex, and the following theorem states the convergence properties of the HLF-SZO (11).

Theorem 4.2. *(Convex Case). Let $\alpha \in [0, 1)$, $\beta \in (0, 2)$, and the total number of iterations T be given. Suppose that Assumption 4.1 holds, f is convex and has a finite minimizer $\boldsymbol{x}^* \in \mathbb{R}^d$. Then by choosing*

$$\eta \leq \frac{(1-\alpha)(1-|1-\beta|)^2}{20LdT^{\frac{1}{3}}}, \ \frac{4\eta dG}{(1-|1-\beta|)(1-\alpha)} \leq r \leq \frac{G}{LT^{\frac{1}{3}}}, \tag{16}$$

the HLF-SZO method (11) achieves

$$\begin{aligned}&\mathbb{E}\left[f(\overline{\boldsymbol{x}}_T)\right] - f(\boldsymbol{x}^*)\\ &\leq \frac{3(1-\alpha)\|\boldsymbol{x}_1 - \boldsymbol{x}^*\|^2}{4\eta T} + \frac{3G^2}{2LT^{2/3}} + \mathcal{O}(\frac{d}{T}),\end{aligned} \tag{17}$$

where $\overline{\boldsymbol{x}}_T := \frac{1}{T}\sum_{k=1}^T \boldsymbol{x}_k$. Moreover, by letting η achieve the equality in (16), we have

$$\begin{aligned}&\mathbb{E}\left[f(\overline{\boldsymbol{x}}_T)\right] - f(\boldsymbol{x}^*)\\ &\leq \frac{d}{T^{2/3}}\left(\frac{15L\|\boldsymbol{x}_1 - \boldsymbol{x}^*\|^2}{(1-|1-\beta|)^2} + \frac{3G^2}{2Ld}\right) + \mathcal{O}(\frac{d}{T}).\end{aligned} \tag{18}$$

The proof of Theorem 4.2 is provided in Appendix B. Some key implications of Theorem 4.2 are discussed below:

1. **Convergence rate**: The dominant term on the right-hand side of (18) is $\mathcal{O}(d/T^{\frac{2}{3}})$, which is better than the convergence rate $\mathcal{O}(d^{\frac{2}{3}}/T^{\frac{1}{3}})$ of the vanilla SZO method (Gasnikov et al., 2017) whenever $T > d$. Compared with the residual-feedback SZO method with the rate of $\mathcal{O}(d^2/T^{\frac{2}{3}})$ given in (Zhang et al., 2021), the HLF-SZO (11) has the same dependency on T but achieves better dependency on the problem dimension d, which is due to the refined analysis on the second moment of the zeroth-order gradient estimator. On the other hand, all these SZO methods are inferior to the two-point ZO methods (Nesterov & Spokoiny, 2017) that achieve the convergence rate of $\mathcal{O}(d/T)$. See Table 1 for the iteration complexity of these ZO methods.

2. **Choice of β**: It is seen that the optimal β that minimizes the dominant term in (18) is given by $\beta = 1$. This choice of β implied by the theoretical analysis is consistent with our empirical results in Section 5. As mentioned above, the HF-SZO method (9) with $\beta = 1$ is equivalent to the residual-feedback SZO method (Zhang et al., 2021).

3. **Choice of α**: It is seen that the parameter α does not appear in the dominant term on the right-hand side of (18).

This suggests that, theoretically, the HLF-SZO (11) converges at least as fast as the HF-SZO (9) (or the residual-feedback SZO). Moreover, the numerical results in Section 5 show that the HLF-SZO (11) with $\alpha > 0$ empirically achieves faster convergence than the residual-feedback SZO (Zhang et al., 2021).

Next, we establish the convergence of HLF-SZO (11) for a nonconvex objective function f with Theorem 4.3.

Theorem 4.3. *(Nonconvex Case). Let $\alpha \in [0,1)$, $\beta \in (0,2)$, and the total number of iterations T be sufficiently large. Suppose that Assumption 4.1 holds and $f^* := \inf_{x\in\mathbb{R}^d} f(x) > -\infty$. Then by choosing (16), the HLF-SZO method (11) achieves*

$$\frac{1}{T}\sum_{k=1}^{T}\mathbb{E}\big[\|\nabla f(\boldsymbol{x}_k)\|^2\big] \le \frac{4(1-\alpha)(f(\boldsymbol{x}_1)-f^*)}{\eta T} + \frac{8G^2}{T^{2/3}} + \mathcal{O}(\frac{d}{T}). \quad (19)$$

Moreover, by letting η achieve the equality in (16), we have

$$\frac{1}{T}\sum_{k=1}^{T}\mathbb{E}\big[\|\nabla f(\boldsymbol{x}_k)\|^2\big] \le \frac{d}{T^{2/3}}\left(\frac{80L(f(\boldsymbol{x}_1)-f^*)}{(1-|1-\beta|)^2} + \frac{8G^2}{d}\right) + \mathcal{O}(\frac{d}{T}). \quad (20)$$

The proof of Theorem 4.3 is provided in Appendix C. Note that in (19) and (20), we use the ergodic rate of the squared norm of the gradient $\|\nabla f(\boldsymbol{x}_k)\|^2$ to characterize the convergence behavior. This is common for local methods of unconstrained smooth nonconvex problems, where one does not aim for global optimal solutions (Ghadimi & Lan, 2013; Nesterov & Spokoiny, 2017). Besides, it is observed that the bound (20) is very similar to (18), and the discussions on Theorem 4.2 can also be applied to the smooth nonconvex case with minor modifications.

5. Numerical Experiments

In this section, we first study the properties of the proposed SZO methods that incorporate high-pass and low-pass filters. Then the HLF-SZO (11) is tested on logistic regression, ridge regression, an artificial test function, and the linear quadratic regulator (LQR) problem, in the comparison with the residual-feedback SZO and the two-point ZO method.

5.1. Properties of SZO Methods with Filters

This subsection compares the performance of vanilla SZO (3), HF-SZO (9), LF-SZO (10), and HLF-SZO (11), and studies the impact of the parameter β on HLF-SZO (11), via the numerical tests on logistic regression.

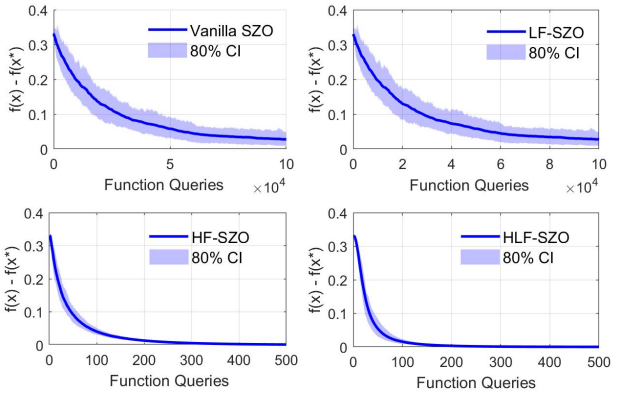

Figure 2. The convergence results of applying vanilla SZO (3), HF-SZO (9), LF-SZO (10), and HLF-SZO (11) to solve the logistic regression (21) in Case 1a).

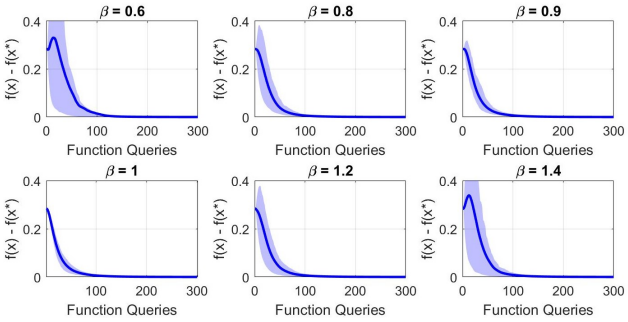

Figure 3. The convergence results of HLF-SZO (11) under different β for solving the logistic regression (21) in Case 1b).

Consider solving the logistic regression problem (21) (Uribe et al., 2020):

$$\min_{\boldsymbol{x}\in\mathbb{R}^d} f(\boldsymbol{x}) = \frac{1}{N}\sum_{i=1}^{N}\log\big(1+\exp(-y_i\cdot A_i^\top \boldsymbol{x})\big), \quad (21)$$

where $A_i \in \mathbb{R}^d$ is one of the data samples, $y_i \in \{-1,1\}$ is the corresponding class label, and N is the total sample size. In our experiments, each element of a data sample A_i is randomly and independently generated from the uniform distribution $\mathrm{Unif}([-1,1])$, and the label is computed by $y_i = \mathrm{sign}(A_i^\top \boldsymbol{x}_\star + \epsilon_i)$ with $\boldsymbol{x}_\star = 0.5\mathbf{1}_d$ and $\epsilon_i \sim \mathrm{Unif}([-0.5, 0.5])$.

Case 1a). Let $d = 2$ and $N = 200$. We set $r = 0.1$, $\beta = 1$, $\alpha = 0.9$ and the initial condition $\boldsymbol{x}_0 = \mathbf{0}_d$ for the SZO methods. Then we manually optimize the stepsize η of each SZO method to achieve its fastest convergence[2]. The selected stepsizes are 5×10^{-4}, 0.3, 5×10^{-5}, and 0.05 for

[2]We gradually tune the stepsize for each SZO method and select the one that has the fastest convergence without divergence.

the vanilla SZO (3), HF-SZO (9), LF-SZO (10), and HLF-SZO (11), respectively. We run each SZO method 200 times and calculate the mean and 80%-confidence interval (CI) of the distance to optimality, i.e., $f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*)$. Figure 2 illustrates the experimental results of solving the logistic regression (21) with these SZO methods.

Case 1b). Next, we tune the parameter β from 0.6 to 1.4 and run experiments for each case to study the impact of β. Other settings are the same as the above Case 1a). The convergence results of HLF-SZO (11) with different β are shown in Figure 3.

The key observations are summarized as follows:

1) From Figure 2, it is seen that HF-SZO and HLF-SZO have much smaller variances and converge much faster than the vanilla SZO. It indicates that the integration of a high-pass filter can greatly reduce the estimation variance of SZO, and thus leads to faster convergence.

2) Figure 2 also shows that HLF-SZO converges even faster than HF-SZO due to the integration of a low-pass filter. It implies that the momentum term introduced by the low-pass filter can further accelerate convergence, which is verified by more numerical tests in the next subsections.

3) From Figure 3, it is observed that the case with $\beta = 1$ achieves the least variance and fastest convergence. This is consistent with the theoretical analysis in Section 4 that $\beta = 1$ is the optimal choice.

5.2. Comparison with Other ZO Algorithms

This subsection compares the performance of the proposed HLF-SZO method (11) with the residual-feedback SZO method (Zhang et al., 2021) and the two-point ZO method (4), through the tests on logistic regression, ridge regression, and Beale function. Here, the vanilla SZO method (3) is not considered for comparison, because it has a much larger variance and much slower convergence.

In the following cases, we run each ZO method 200 times and calculate the mean and 80%-confidence interval (CI) of the distance to optimality, i.e., $f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*)$. Figure 4 illustrates the convergence results of these ZO methods.

Case 2a). Consider solving the **logistic regression** problem (21) with $d = 50$ and $N = 1000$. We manually optimize the stepsize η of each ZO method to achieve its fastest convergence. The selected stepsizes are 1.5×10^{-2}, 4.5×10^{-2}, and 0.7 for the HLF-SZO (11), the residual-feedback SZO, and the two-point ZO (4). Other settings are the same as Case 1a) in Section 5.1.

Case 2b). Consider solving the **ridge regression** problem

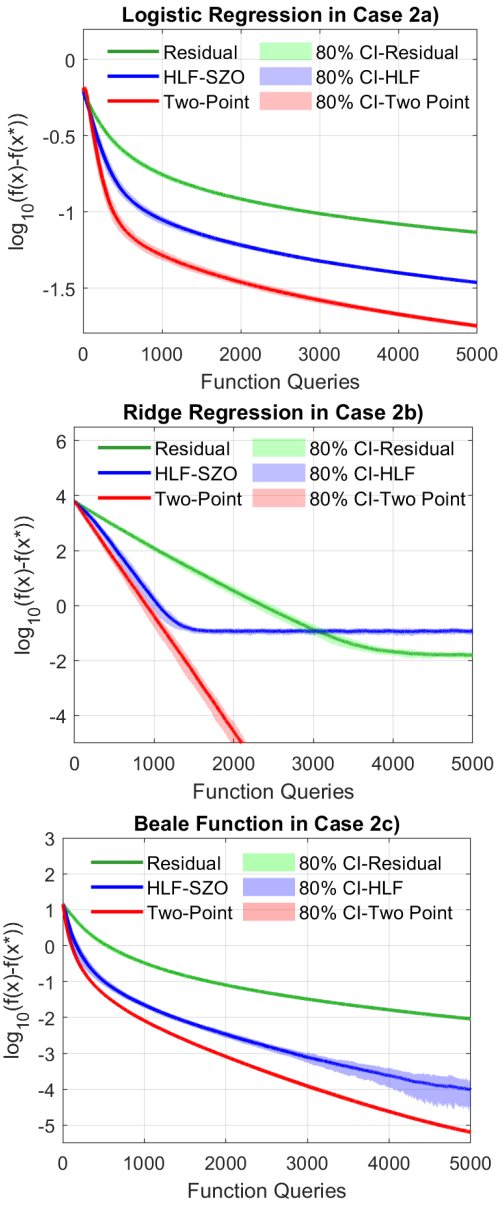

Figure 4. The convergence results of residual-feedback SZO, HLF-SZO (11), and two-point ZO (4) for solving logistic regression (21) in Case 2a), ridge regression (22) in Case 2b), and Beale function (23) in Case 2c).

(22) (Uribe et al., 2020):

$$\min_{\boldsymbol{x}\in\mathbb{R}^d} f(\boldsymbol{x}) = \frac{1}{2}||\boldsymbol{b} - H\boldsymbol{x}||_2^2 + \frac{1}{2}c||\boldsymbol{x}||_2^2. \tag{22}$$

In our experiments, each entry of matrix $H \in \mathbb{R}^{N\times d}$ is generated randomly and independently from Gaussian distribution $\mathcal{N}(0,1)$. The vector $\boldsymbol{b} \in \mathbb{R}^N$ is constructed by letting $\boldsymbol{b} = H\boldsymbol{x}_\star + \boldsymbol{\epsilon}$ for a certain predefined $\boldsymbol{x}_\star \in \mathbb{R}^d$ and $\boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{0}, 0.1\boldsymbol{I})$. The regularization constant is set to be $c = 0.1$. Let $d = 50$, $N = 1000$ and $\boldsymbol{x}_\star = 0.5 \times \boldsymbol{1}_d$. For the ZO methods, we set $r = 0.1$, $\alpha = 0.9$, $\beta = 1$, and manually optimize the stepsize η to achieve the fastest convergence. The selected stepsizes are 1×10^{-6}, 2.4×10^{-6}, and 2×10^{-5} for the HLF-SZO (11), the residual-feedback SZO, and the two-point ZO (4), respectively.

Case 2c). Consider minimizing the nonconvex **Beale function** given by (23), which is an artificial test function and has a global minimum at $\boldsymbol{x}^* = [3; 0.5]$ with $f(\boldsymbol{x}^*) = 0$.

$$f(\boldsymbol{x}) = (1.5 - x_1 + x_1x_2)^2 + (2.25 - x_1 + x_1x_2^2)^2 + (2.625 - x_1 + x_1x_2^3)^2. \tag{23}$$

For the ZO methods, we set $r = 0.01$, $\beta = 1$, $\alpha = 0.9$, $\boldsymbol{x}_0 = [0; 0]$, and manually optimize the stepsize η to achieve the fastest convergence. The selected stepsizes are 2×10^{-4}, 5.8×10^{-4}, and 6×10^{-3} for the HLF-SZO (11), the residual-feedback SZO, and the two-point ZO (4), respectively.

From Figure 4, it is seen that the proposed HLF-SZO (11) converges faster than the residual-feedback SZO method in all cases, and its performance is comparable to that of the two-point ZO method (4). Since the residual-feedback SZO is equivalent to the HF-SZO (9) with $\beta = 1$, these results indicate that the integration of a low-pass filter can further accelerate the convergence of SZO via the momentum term.

5.3. Control Policy Optimization for Linear Quadratic Regulator (LQR) Problem

We consider solving the linear quadratic regulator (LQR) problem (Anderson & Moore, 2007; Fazel et al., 2018) (24) with the linear feedback control policy $\boldsymbol{u}_t = -K\boldsymbol{x}_t$.

$$\min_K \quad J(K) := \mathbb{E}\Big[\sum_{t=0}^{\infty} \gamma^t \big(\boldsymbol{x}_t^\top Q \boldsymbol{x}_t + \boldsymbol{u}_t^\top R \boldsymbol{u}_t\big)\Big] \tag{24a}$$

$$\text{s.t.} \quad \boldsymbol{x}_{t+1} = A\boldsymbol{x}_t + B\boldsymbol{u}_t + \boldsymbol{w}_t. \tag{24b}$$

Here, $\boldsymbol{x}_t \in \mathbb{R}^{n_x}$ and $\boldsymbol{u}_t \in \mathbb{R}^{n_u}$ are the state and the control input at time step t, respectively. $\boldsymbol{w}_t \sim \mathcal{N}(0, \delta^2 I)$ denotes the random system noise. In the simulations, we set $n_x = 20$, $n_u = 15$, and thus the problem dimension is $d = 300$. Let the discount factor γ be 0.9. Each element in the transition matrices $A \in \mathbb{R}^{n_x\times n_x}$ and $B \in \mathbb{R}^{n_x\times n_u}$ are randomly generated from Unif$([-0.1, 0])$ and Unif$([0, 0.02])$, respectively. Let Q and R be the identity matrix. To evaluate the total cost $J(K)$, we run the episode with a finite horizon $H = 50$ to approximate it, which has been checked to be sufficiently long for the system to converge. The initial state is set as $\boldsymbol{x}_0 = \boldsymbol{1}_{n_x}$.

We apply the HLF-SZO (11), residual-feedback SZO, and two-point ZO (4) to solve the LQR problem (24). The initial K_0 is constructed by $K_0 = K^* + \tilde{K}$, where K^* is the optimal solution and each element in $\tilde{K}$ is randomly generated from Unif$([-1, 1])$. We set $r = 0.1$, $\alpha = 0.9$, $\beta = 1$, and tune the standard deviation δ of $\boldsymbol{w}_t$ to simulate different levels of system noise, as in the following cases. In each case, we manually optimize the stepsize η of each ZO method to achieve its fastest convergence.

Case 3a). Let $\delta = 0$. The selected stepsizes are 7.5×10^{-6}, 2.25×10^{-5}, and 7.5×10^{-4} for HLF-SZO (11), residual-feedback SZO, and two-point ZO (4), respectively.

Case 3b). Let $\delta = 0.005$. The selected stepsizes are 4×10^{-6}, 2.4×10^{-5}, and 1.2×10^{-4} for HLF-SZO (11), residual-feedback SZO, and two-point ZO (4), respectively.

Case 3c). Let $\delta = 0.01$. The selected stepsizes are 3×10^{-6}, 2.1×10^{-5}, and 9×10^{-5} for HLF-SZO (11), residual-feedback SZO, and two-point ZO (4), respectively.

The convergence results of these ZO methods are illustrated in Figure 5. It is observed that larger noise leads to slower convergence and lower convergence accuracy of ZO methods. In all the cases, the HLF-SZO (11) can converge faster than the residual-feedback SZO and achieve comparable performance to the two-point ZO method (4).

6. Conclusion

In this work, by leveraging the idea of high-pass and low-pass filters from extremum seeking control, we develop a novel single-point zeroth-order optimization (SZO) method called HLF-SZO. Interestingly, we find that the integration of a high-pass filter coincides with the residual feedback scheme proposed in Zhang et al. (2021), and the integration of a low-pass filter can be interpreted as the momentum method. We prove that the HLF-SZO method achieves an iteration complexity of $\mathcal{O}(d^{\frac{3}{2}}/\epsilon^{\frac{3}{2}})$ for both convex and nonconvex objective functions under the Lipschitz and smoothness conditions. This iteration complexity is improved over the vanilla SZO method (Gasnikov et al., 2017) and the residual-feedback SZO method (Zhang et al., 2021), although it is inferior to the two-point ZO methods. Extensive numerical experiments demonstrate that the HLF-SZO method achieves a much smaller variance and much faster convergence than the vanilla SZO method; it empirically outperforms the residual-feedback SZO method and has comparable performance to the two-point ZO method.

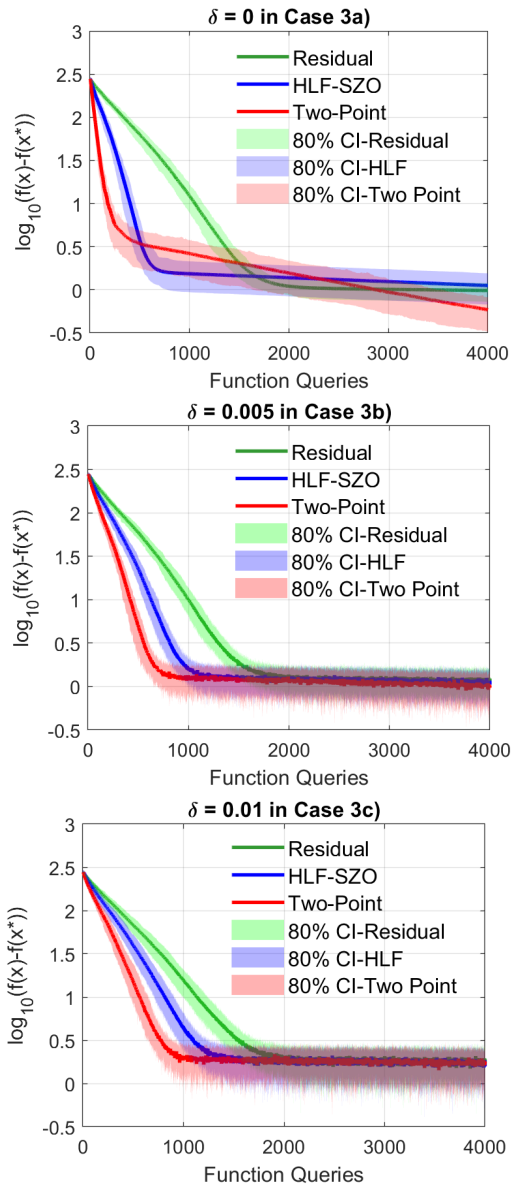

Figure 5. The convergence results of residual-feedback SZO, HLF-SZO (11), and two-point ZO (4) for solving the LQR problem (24) under different levels of system noise δ.

Bridging ZO and extremum seeking control to benefit each other is a promising direction of research. This paper is a preliminary attempt to throw light on this direction, and much remains to be studied. Our future work includes 1) extending HLF-SZO to stochastic optimization problems, 2) designing better filters and time-discretization schemes to further improve convergence, etc.

Acknowledgements

This work was supported by the funding programs: NSF CAREER ECCS-1553407, NSF AI Institute 2112085, NSF CNS 2003111, and ONR YIP N00014-19-1-2217.

References

Anderson, B. D. and Moore, J. B. *Optimal control: linear quadratic methods*. Courier Corporation, 2007.

Ariyur, K. B. and Krstic, M. *Real time optimization by extremum-seeking control*. John Wiley & Sons, 2003.

Berahas, A. S., Cao, L., Choromanski, K., and Scheinberg, K. A theoretical and empirical comparison of gradient approximations in derivative-free optimization. *Foundations of Computational Mathematics*, 22(2):507–560, 2022.

Chen, P.-Y., Zhang, H., Sharma, Y., Yi, J., and Hsieh, C.-J. Zoo: Zeroth order optimization based black-box attacks to deep neural networks without training substitute models. In *Proceedings of the 10th ACM workshop on artificial intelligence and security*, pp. 15–26, 2017.

Chen, X., Poveda, J. I., and Li, N. Model-free optimal voltage control via continuous-time zeroth-order methods. *arXiv preprint arXiv:2103.14703*, 2021a.

Chen, X., Poveda, J. I., and Li, N. Safe model-free optimal voltage control via continuous-time zeroth-order methods. In *2021 60th IEEE Conference on Decision and Control (CDC)*, pp. 4064–4070. IEEE, 2021b.

Chen, Y., Bernstein, A., Devraj, A., and Meyn, S. Model-free primal-dual methods for network optimization with application to real-time optimal power flow. In *2020 American Control Conference (ACC)*, pp. 3140–3147. IEEE, 2020.

Chill, R. and Fašangová, E. Gradient systems. *MatFyzPress, Prague*, 2010.

Dekel, O., Eldan, R., and Koren, T. Bandit smooth convex optimization: Improving the bias-variance tradeoff. In *NIPS*, pp. 2926–2934, 2015.

Dürr, H.-B., Zeng, C., and Ebenbauer, C. Saddle point seeking for convex optimization problems. *IFAC Proceedings Volumes*, 46(23):540–545, 2013.

Fazel, M., Ge, R., Kakade, S., and Mesbahi, M. Global convergence of policy gradient methods for the linear quadratic regulator. In *International Conference on Machine Learning*, pp. 1467–1476. PMLR, 2018.

Flaxman, A. D., Kalai, A. T., and McMahan, H. B. Online convex optimization in the bandit setting: Gradient descent without a gradient. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 385–394, 2005.

Gasnikov, A. V., Krymova, E. A., Lagunovskaya, A. A., Usmanova, I. N., and Fedorenko, F. A. Stochastic online optimization. single-point and multi-point non-linear multi-armed bandits. convex and strongly-convex case. *Automation and Remote Control*, 78(2):224–234, 2017.

Ghadimi, S. and Lan, G. Stochastic first-and zeroth-order methods for nonconvex stochastic programming. *SIAM Journal on Optimization*, 23(4):2341–2368, 2013.

Jongeneel, W. Imaginary zeroth-order optimization. *arXiv preprint arXiv:2112.07488*, 2021.

Jongeneel, W., Yue, M.-C., and Kuhn, D. Small errors in random zeroth order optimization are imaginary. *arXiv preprint arXiv:2103.05478*, 2021.

Krstić, M. Performance improvement and limitations in extremum seeking control. *Systems & Control Letters*, 39 (5):313–326, 2000.

Li, Y., Tang, Y., Zhang, R., and Li, N. Distributed reinforcement learning for decentralized linear quadratic control: A derivative-free policy optimization approach. *arXiv preprint arXiv:1912.09135*, 2019.

Liu, S., Chen, P.-Y., Kailkhura, B., Zhang, G., Hero III, A. O., and Varshney, P. K. A primer on zeroth-order optimization in signal processing and machine learning: Principals, recent advances, and applications. *IEEE Signal Processing Magazine*, 37(5):43–54, 2020.

Malik, D., Pananjady, A., Bhatia, K., Khamaru, K., Bartlett, P., and Wainwright, M. Derivative-free methods for policy optimization: Guarantees for linear quadratic systems. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pp. 2916–2925. PMLR, 2019.

Nesterov, Y. *Introductory Lectures on Convex Optimization: A Basic Course*. Springer Science+Business Media, LLC, 2004.

Nesterov, Y. and Spokoiny, V. Random gradient-free minimization of convex functions. *Foundations of Computational Mathematics*, 17(2):527–566, 2017.

Novitskii, V. and Gasnikov, A. Improved exploiting higher order smoothness in derivative-free optimization and continuous bandit. *arXiv preprint arXiv:2101.03821*, 2021.

Polyak, B. T. Some methods of speeding up the convergence of iteration methods. *Ussr computational mathematics and mathematical physics*, 4(5):1–17, 1964.

Poveda, J. I. and Li, N. Robust hybrid zero-order optimization algorithms with acceleration via averaging in time. *Automatica*, 123:109361, 2021.

Qian, N. On the momentum term in gradient descent learning algorithms. *Neural networks*, 12(1):145–151, 1999.

Ruder, S. An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747*, 2016.

Saha, A. and Tewari, A. Improved regret guarantees for online smooth convex optimization with bandit feedback. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, pp. 636–642. JMLR Workshop and Conference Proceedings, 2011.

Shamir, O. An optimal algorithm for bandit and zero-order convex optimization with two-point feedback. *The Journal of Machine Learning Research*, 18(1):1703–1713, 2017.

Tan, Y., Nešić, D., and Mareels, I. On non-local stability properties of extremum seeking control. *Automatica*, 42 (6):889–903, 2006.

Tan, Y., Moase, W. H., Manzie, C., Nešić, D., and Mareels, I. M. Extremum seeking from 1922 to 2010. In *Proceedings of the 29th Chinese control conference*, pp. 14–26. IEEE, 2010.

Tao, W., Long, S., Wu, G., and Tao, Q. The role of momentum parameters in the optimal convergence of adaptive polyak's heavy-ball methods. *arXiv preprint arXiv:2102.07314*, 2021.

Uribe, C. A., Lee, S., Gasnikov, A., and Nedić, A. A dual approach for optimal algorithms in distributed optimization over networks. In *2020 Information Theory and Applications Workshop (ITA)*, pp. 1–37. IEEE, 2020.

Wang, L. and Spall, J. C. Improved spsa using complex variables with applications in optimal control problems. In *2021 American Control Conference (ACC)*, pp. 3519–3524. IEEE, 2021.

Wang, L., Zhu, J., and Spall, J. C. Model-free optimal control using spsa with complex variables. In *2021 55th Annual Conference on Information Sciences and Systems (CISS)*, pp. 1–5. IEEE, 2021.

Ye, M. and Hu, G. Distributed extremum seeking for constrained networked optimization and its application to energy consumption control in smart grid. *IEEE Transactions on Control Systems Technology*, 24(6):2048–2058, 2016.

Zhang, Y., Zhou, Y., Ji, K., and Zavlanos, M. M. A new one-point residual-feedback oracle for black-box learning and control. *Automatica*, pp. 110006, 2021.

A. Notations and Preliminary Results

It can be checked that the iterations can be equivalently written as

$$\begin{aligned}\boldsymbol{g}_k &= \frac{d}{r} z_k \boldsymbol{u}_k, \quad z_k = (1-\beta) z_{k-1} + f(\boldsymbol{x}_k + r\boldsymbol{u}_k) - f(\boldsymbol{x}_{k-1} + r\boldsymbol{u}_{k-1}),\\ \boldsymbol{p}_k &= \alpha \boldsymbol{p}_{k-1} + \eta \boldsymbol{g}_k,\\ \boldsymbol{x}_{k+1} &= \boldsymbol{x}_k - \boldsymbol{p}_k,\end{aligned}$$

for $k \geq 1$, and we set $\boldsymbol{p}_0 = \boldsymbol{0}$, $\boldsymbol{x}_1 = \boldsymbol{x}_0$ and $z_0 = 0$. We denote $\tilde{\beta} \coloneqq 1 - |1-\beta|$ for notational simplicity. We treat $\boldsymbol{x}_1 = \boldsymbol{x}_0$ as deterministic quantities and let $\mathcal{F}_k$ denote the filtration generated by $\boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ and $\boldsymbol{u}_0, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_{k-1}$ for $k \geq 1$.

The following lemma deals with the expectation of $\boldsymbol{g}_k$.

Lemma A.1. *Denote $\mathbb{B}_d \coloneqq \{\boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x}\| \leq 1\}$. For all $k \geq 1$, we have*

$$\mathbb{E}[\boldsymbol{g}_k \,|\, \mathcal{F}_k] = \nabla f_r(\boldsymbol{x}_k),$$

where $f_r : \mathbb{R}^d \to \mathbb{R}$ is defined by $f_r(\boldsymbol{x}) \coloneqq \mathbb{E}_{\boldsymbol{y} \sim \mathrm{Unif}(\mathbb{B}_d)}[f(\boldsymbol{x} + r\boldsymbol{y})]$. Moreover, for all $\boldsymbol{x} \in \mathbb{R}^d$,

$$|f_r(\boldsymbol{x}) - f(\boldsymbol{x})| \leq \frac{1}{2} L r^2, \qquad \|\nabla f_r(\boldsymbol{x}) - \nabla f(\boldsymbol{x})\| \leq Lr,$$

and if f is convex, then $f_r(\boldsymbol{x}) \geq f(\boldsymbol{x})$.

Proof. The result $\mathbb{E}[\boldsymbol{g}_k \,|\, \mathcal{F}_k] = \nabla f_r(\boldsymbol{x}_k)$ is standard in zeroth-order optimization (Flaxman et al., 2005). By the L-smoothness of f, we have

$$r\langle \boldsymbol{y}, \nabla f(\boldsymbol{x})\rangle - \frac{L}{2} r^2 \|\boldsymbol{y}\|^2 \leq f(\boldsymbol{x} + r\boldsymbol{y}) - f(\boldsymbol{x}) \leq r\langle \boldsymbol{y}, \nabla f(\boldsymbol{x})\rangle + \frac{L}{2} r^2 \|\boldsymbol{y}\|^2,$$

and by taking the expectation with $\boldsymbol{y} \sim \mathrm{Unif}(\mathbb{B}_d)$, we get

$$-\frac{L}{2} r^2 \leq f_r(\boldsymbol{x}) - f(\boldsymbol{x}) \leq \frac{L}{2} r^2$$

for any $\boldsymbol{x} \in \mathbb{R}^d$. If f is convex, then $f(\boldsymbol{x}+r\boldsymbol{y}) - f(\boldsymbol{x}) \geq r\langle \boldsymbol{y}, \nabla f(\boldsymbol{x})\rangle$, and by by taking the expectation with $\boldsymbol{y} \sim \mathrm{Unif}(\mathbb{B}_d)$, we get $f_r(\boldsymbol{x}) \geq f(\boldsymbol{x})$.

Finally, regarding the gradient $\nabla f_r(\boldsymbol{x})$, we have

$$\begin{aligned}\|\nabla f_r(\boldsymbol{x}) - \nabla f(\boldsymbol{x})\| &= \|\nabla \mathbb{E}_{\boldsymbol{y} \sim \mathrm{Unif}(\mathbb{B}_d)}[f(\boldsymbol{x} + r\boldsymbol{y})] - \nabla f(\boldsymbol{x})\|\\ &= \|\mathbb{E}_{\boldsymbol{y} \sim \mathrm{Unif}(\mathbb{B}_d)}[\nabla f(\boldsymbol{x} + r\boldsymbol{y}) - \nabla f(\boldsymbol{x})]\|\\ &\leq \mathbb{E}_{\boldsymbol{y} \sim \mathrm{Unif}(\mathbb{B}_d)}[Lr\|\boldsymbol{y}\|] \leq Lr,\end{aligned}$$

where we interchange integration and differentiation in the second step, and it can be justified by the dominated convergence theorem. □

The following lemma deals with the accumulated second moment of z_k.

Lemma A.2. *We have*

$$\sum_{k=1}^{T} \mathbb{E}\big[|z_k|^2\big] \leq \frac{5r^2}{\tilde{\beta}^2 d} \sum_{k=1}^{T-1} \mathbb{E}\big[\|\nabla f(\boldsymbol{x}_k)\|^2\big] + \frac{2G^2}{\tilde{\beta}^2} \sum_{k=1}^{T-1} \mathbb{E}\big[\|\boldsymbol{p}_k\|^2\big] + \frac{10 T r^4 L^2}{\tilde{\beta}^2} + \frac{5r^2}{2\tilde{\beta} d} \|\nabla f(\boldsymbol{x}_1)\|^2.$$

Proof. For $k = 1$, by the initial conditions imposed on $\boldsymbol{x}_0$ and z_0, we have

$$\begin{aligned}
\mathbb{E}\big[|z_1|^2\big] &= \mathbb{E}\big[|f(\boldsymbol{x}_1+r\boldsymbol{u}_1) - f(\boldsymbol{x}_1+r\boldsymbol{u}_0)|^2\big] = \mathbb{E}\Big[\big|f(\boldsymbol{x}_1+r\boldsymbol{u}_1) - f(\boldsymbol{x}_1) - (f(\boldsymbol{x}_1+r\boldsymbol{u}_0) - f(\boldsymbol{x}_1))\big|^2\Big] \\
&= \mathbb{E}\left[\left|\int_0^r \langle \nabla f(\boldsymbol{x}_1+s\boldsymbol{u}_1), \boldsymbol{u}_1\rangle \, ds - \int_0^r \langle \nabla f(\boldsymbol{x}_1+s\boldsymbol{u}_0), \boldsymbol{u}_0\rangle \, ds\right|^2\right] \\
&= \mathbb{E}\left[\left|r\langle \nabla f(\boldsymbol{x}_1), \boldsymbol{u}_1-\boldsymbol{u}_0\rangle + \int_0^r \langle \nabla f(\boldsymbol{x}_1+s\boldsymbol{u}_1)-\nabla f(\boldsymbol{x}_1), \boldsymbol{u}_1\rangle \, ds - \int_0^r \langle \nabla f(\boldsymbol{x}_1+s\boldsymbol{u}_0)-\nabla f(\boldsymbol{x}_1), \boldsymbol{u}_0\rangle \, ds\right|^2\right] \\
&\le \frac{5}{4} r^2 \nabla f(\boldsymbol{x}_1)^\top \mathbb{E}\big[(\boldsymbol{u}_1 - \boldsymbol{u}_0)(\boldsymbol{u}_1 - \boldsymbol{u}_0)^\top\big] \, \nabla f(\boldsymbol{x}_1) \\
&\quad + 10\,\mathbb{E}\left[\left|\int_0^r \|\nabla f(\boldsymbol{x}_1+s\boldsymbol{u}_1) - \nabla f(\boldsymbol{x}_1)\|\|\boldsymbol{u}_1\| \, ds\right|^2\right] + 10\,\mathbb{E}\left[\left|\int_0^r \|\nabla f(\boldsymbol{x}_1+s\boldsymbol{u}_0) - \nabla f(\boldsymbol{x}_1)\|\|\boldsymbol{u}_0\| \, ds\right|^2\right] \\
&\le \frac{5r^2}{2d}\|\nabla f(\boldsymbol{x}_1)\|^2 + 10\mathbb{E}\left[\left|\int_0^r L|s|\|\boldsymbol{u}_1\|^2 \, ds\right|^2\right] + 10\mathbb{E}\left[\left|\int_0^r L|s|\|\boldsymbol{u}_0\|^2 \, ds\right|^2\right] \\
&\le \frac{5r^2}{2d}\|\nabla f(\boldsymbol{x}_1)\|^2 + 5L^2 r^4.
\end{aligned}$$

The first inequality above is the application of the inequality $(a+b+c)^2 \le \frac{5}{4}a^2 + 10b^2 + 10c^2$, as $\frac{5}{4}a^2 + 10b^2 + 10c^2 - (a+b+c)^2 = (\frac{1}{\sqrt{8}}a - \sqrt{8}b)^2 + (\frac{1}{\sqrt{8}}a - \sqrt{8}c)^2 + (b-c)^2 \ge 0$. The second inequality above is due to $\mathbb{E}\big[(\boldsymbol{u}_1 - \boldsymbol{u}_0)(\boldsymbol{u}_1 - \boldsymbol{u}_0)^\top\big] = \frac{2}{d} I_d$.

Then for $k > 1$, by the inequality $(a+b)^2 \le (1+\gamma)a^2 + (1+\frac{1}{\gamma})b^2$ for any $\gamma > 0$ (frequently used below to bound the square of a sum), we let $\gamma = \frac{1-|1-\beta|}{|1-\beta|}$ and get

$$\begin{aligned}
\mathbb{E}\big[|z_k|^2\big] &\le \frac{1}{|1-\beta|} \cdot (1-\beta)^2 \, \mathbb{E}\big[|z_{k-1}|^2\big] + \frac{1}{1-|1-\beta|} \, \mathbb{E}\Big[\big|f(\boldsymbol{x}_k+r\boldsymbol{u}_k) - f(\boldsymbol{x}_{k-1}+r\boldsymbol{u}_{k-1})\big|^2\Big] \\
&= (1-\tilde{\beta}) \, \mathbb{E}\big[|z_{k-1}|^2\big] + \tilde{\beta}^{-1} \mathbb{E}\Big[\big|f(\boldsymbol{x}_k+r\boldsymbol{u}_k) - f(\boldsymbol{x}_{k-1}+r\boldsymbol{u}_{k-1})\big|^2\Big]
\end{aligned}$$

when $\beta \neq 1$. And obviously this inequality also extends to the $\beta = 1$ case, because $z_k = f(\boldsymbol{x}_k+r\boldsymbol{u}_k) - f(\boldsymbol{x}_{k-1}+r\boldsymbol{u}_{k-1})$ and $\tilde{\beta} = 1$ when $\beta = 1$. Furthermore,

$$\begin{aligned}
&\mathbb{E}\Big[\big|f(\boldsymbol{x}_k+r\boldsymbol{u}_k) - f(\boldsymbol{x}_{k-1}-r\boldsymbol{u}_{k-1})\big|^2\Big] \\
=&\mathbb{E}\Big[\big|f(\boldsymbol{x}_k+r\boldsymbol{u}_k) - f(\boldsymbol{x}_{k-1}+r\boldsymbol{u}_k) + f(\boldsymbol{x}_{k-1}+r\boldsymbol{u}_k) - f(\boldsymbol{x}_{k-1}-r\boldsymbol{u}_{k-1})\big|^2\Big] \\
\le& \, 2\Big(\mathbb{E}\Big[\big|f(\boldsymbol{x}_{k-1}+r\boldsymbol{u}_k) - f(\boldsymbol{x}_{k-1}+r\boldsymbol{u}_{k-1})\big|^2\Big] + \mathbb{E}\Big[\big|f(\boldsymbol{x}_k+r\boldsymbol{u}_k) - f(\boldsymbol{x}_{k-1}+r\boldsymbol{u}_k)\big|^2\Big]\Big).
\end{aligned}$$

For the first term, we have

$$\begin{aligned}
&\mathbb{E}\big[|f(\boldsymbol{x}_{k-1}+r\boldsymbol{u}_k) - f(\boldsymbol{x}_{k-1}+r\boldsymbol{u}_{k-1})|^2\big] \\
&= \mathbb{E}\Bigg[\bigg|r\langle\nabla f(\boldsymbol{x}_{k-1}), \boldsymbol{u}_k-\boldsymbol{u}_{k-1}\rangle + \int_0^r \langle\nabla f(\boldsymbol{x}_{k-1}+s\boldsymbol{u}_k)-\nabla f(\boldsymbol{x}_{k-1}), \boldsymbol{u}_k\rangle\, ds \\
&\qquad - \int_0^r \langle\nabla f(\boldsymbol{x}_{k-1}+s\boldsymbol{u}_{k-1})-\nabla f(\boldsymbol{x}_{k-1}), \boldsymbol{u}_{k-1}\rangle\, ds\bigg|^2\Bigg] \\
&\le 5\,\mathbb{E}\left[\left(\int_0^r (\|\nabla f(\boldsymbol{x}_{k-1}+s\boldsymbol{u}_k)-\nabla f(\boldsymbol{x}_{k-1})\| + \|\nabla f(\boldsymbol{x}_{k-1}+s\boldsymbol{u}_{k-1})-\nabla f(\boldsymbol{x}_{k-1})\|)\, ds\right)^2\right] \\
&\quad + \frac{5}{4}r^2\,\mathbb{E}\big[|\langle\nabla f(\boldsymbol{x}_{k-1}), \boldsymbol{u}_k-\boldsymbol{u}_{k-1}\rangle|^2\big] \\
&\le 5\,\mathbb{E}\left[\left(\int_0^r (Ls\|\boldsymbol{u}_k\| + Ls\|\boldsymbol{u}_{k-1}\|)\, ds\right)^2\right] \\
&\quad + \frac{5}{4}r^2\,\mathbb{E}\big[\nabla f(\boldsymbol{x}_{k-1})^\top \mathbb{E}\big[(\boldsymbol{u}_k-\boldsymbol{u}_{k-1})(\boldsymbol{u}_k-\boldsymbol{u}_{k-1})^\top \,\big|\, \mathcal{F}_{k-1}\big]\,\nabla f(\boldsymbol{x}_{k-1})\big] \\
&= 5r^4L^2 + \frac{5r^2}{2d}\mathbb{E}\big[\|\nabla f(\boldsymbol{x}_{k-1})\|^2\big]
\end{aligned}$$

where we used $\|\boldsymbol{u}_k\| = \|\boldsymbol{u}_{k-1}\| = 1$, the law of total expectation, and

$$\mathbb{E}\big[(\boldsymbol{u}_k-\boldsymbol{u}_{k-1})(\boldsymbol{u}_k-\boldsymbol{u}_{k-1})^\top \,\big|\, \mathcal{F}_{k-1}\big] = \mathbb{E}\big[\boldsymbol{u}_k\boldsymbol{u}_k^\top + \boldsymbol{u}_{k-1}\boldsymbol{u}_{k-1}^\top|\mathcal{F}_{k-1}\big] = \frac{2}{d}I_d,$$

and for the second term, we have

$$\mathbb{E}\big[|f(\boldsymbol{x}_k+r\boldsymbol{u}_k) - f(\boldsymbol{x}_{k-1}+r\boldsymbol{u}_k)|^2\big] \le G^2\,\mathbb{E}\left[\|\boldsymbol{x}_k - \boldsymbol{x}_{k-1}\|^2\right] = G^2\,\mathbb{E}\left[\|\boldsymbol{p}_{k-1}\|^2\right].$$

Therefore for $k > 1$,

$$\mathbb{E}\big[|z_k|^2\big] \le (1-\tilde{\beta})\,\mathbb{E}\big[|z_{k-1}|^2\big] + \frac{5r^2}{\tilde{\beta}d}\mathbb{E}\big[\|\nabla f(\boldsymbol{x}_{k-1})\|^2\big] + \frac{10r^4L^2}{\tilde{\beta}} + \frac{2G^2}{\tilde{\beta}}\mathbb{E}\big[\|\boldsymbol{p}_{k-1}\|^2\big].$$

By summing over $k = 1, \ldots, T$, we get

$$\begin{aligned}
\sum_{k=1}^T \mathbb{E}\big[|z_k|^2\big] &\le (1-\tilde{\beta})\sum_{k=1}^{T-1}\mathbb{E}\big[|z_k|^2\big] + \frac{5r^2}{\tilde{\beta}d}\sum_{k=1}^{T-1}\mathbb{E}\big[\|\nabla f(\boldsymbol{x}_k)\|^2\big] + \frac{10(T-1)r^4L^2}{\tilde{\beta}} \\
&\quad + \frac{2G^2}{\tilde{\beta}}\sum_{k=1}^{T-1}\mathbb{E}\big[\|\boldsymbol{p}_k\|^2\big] + \frac{5r^2}{2d}\mathbb{E}\big[\|\nabla f(\boldsymbol{x}_1)\|^2\big] + 5r^4L^2 \\
&\le (1-\tilde{\beta})\sum_{k=1}^{T}\mathbb{E}\big[|z_k|^2\big] + \frac{5r^2}{\tilde{\beta}d}\sum_{k=1}^{T-1}\mathbb{E}\big[\|\nabla f(\boldsymbol{x}_k)\|^2\big] + \frac{10Tr^4L^2}{\tilde{\beta}} + \frac{2G^2}{\tilde{\beta}}\sum_{k=1}^{T-1}\mathbb{E}\big[\|\boldsymbol{p}_k\|^2\big] + \frac{5r^2}{2d}\|\nabla f(\boldsymbol{x}_1)\|^2,
\end{aligned}$$

which then implies the desired result. □

The following lemma then provides bounds for the accumulated second moment of $\boldsymbol{g}_k$ and $\boldsymbol{p}_k$.

Lemma A.3. *Suppose the quantity*

$$\theta := 1 - \frac{4\eta^2d^2G^2(1+\alpha^2)}{\tilde{\beta}^2r^2(1-\alpha^2)^2}$$

is positive. Then we have

$$\sum_{k=1}^{T}\mathbb{E}\big[\|\boldsymbol{g}_k\|^2\big] \le \frac{5d}{\theta\tilde{\beta}^2}\sum_{k=1}^{T}\mathbb{E}\big[\|\nabla f(\boldsymbol{x}_k)\|^2\big] + \frac{10Tr^2L^2d^2}{\theta\tilde{\beta}^2} + \frac{5d}{2\theta\tilde{\beta}}\|\nabla f(\boldsymbol{x}_1)\|^2, \tag{25}$$

$$\sum_{k=1}^{T}\mathbb{E}\big[\|\boldsymbol{p}_k\|^2\big] \le \frac{2\eta^2(1+\alpha^2)}{(1-\alpha^2)^2}\sum_{k=1}^{T}\mathbb{E}\big[\|\boldsymbol{g}_k\|^2\big]. \tag{26}$$

Proof. By the definition of $\boldsymbol{p}_k$, we have

$$\begin{aligned}\mathbb{E}\big[\|\boldsymbol{p}_k\|^2\big] &\le \left(1+\frac{1-\alpha^2}{2\alpha^2}\right)\alpha^2\mathbb{E}\big[\|\boldsymbol{p}_{k-1}\|^2\big] + \left(1+\frac{2\alpha^2}{1-\alpha^2}\right)\eta^2\mathbb{E}\big[\|\boldsymbol{g}_k\|^2\big]\\ &= \frac{1+\alpha^2}{2}\mathbb{E}\big[\|\boldsymbol{p}_{k-1}\|^2\big] + \frac{1+\alpha^2}{1-\alpha^2}\eta^2\mathbb{E}\big[\|\boldsymbol{g}_k\|^2\big].\end{aligned}$$

By summing over $k = 1, \dots, T$, we get

$$\sum_{k=1}^{T}\mathbb{E}\big[\|\boldsymbol{p}_k\|^2\big] \le \frac{1+\alpha^2}{2}\sum_{k=1}^{T-1}\mathbb{E}\big[\|\boldsymbol{p}_k\|^2\big] + \frac{1+\alpha^2}{1-\alpha^2}\eta^2\sum_{k=1}^{T}\mathbb{E}\big[\|\boldsymbol{g}_k\|^2\big],$$

which then implies (26).

Now, since $\boldsymbol{g}_k = \frac{d}{r}z_k\boldsymbol{u}_k$ and $\|\boldsymbol{u}_k\| = 1$, we see that $\mathbb{E}\big[\|\boldsymbol{g}_k\|^2\big] = \frac{d^2}{r^2}\mathbb{E}\big[|z_k|^2\big]$. By using the bound in Lemma A.2, we get

$$\sum_{k=1}^{T}\mathbb{E}\big[\|\boldsymbol{g}_k\|^2\big] \le \sum_{k=1}^{T-1}\frac{2d^2G^2}{\tilde{\beta}^2r^2}\mathbb{E}\big[\|\boldsymbol{p}_k\|^2\big] + \frac{5d}{\tilde{\beta}^2}\sum_{k=1}^{T-1}\mathbb{E}\big[\|\nabla f(\boldsymbol{x}_k)\|^2\big] + \frac{10Tr^2L^2d^2}{\tilde{\beta}^2} + \frac{5d}{2\tilde{\beta}}\|\nabla f(\boldsymbol{x}_1)\|^2,$$

After plugging in the bound (26), we can show that

$$\left(1 - \frac{4\eta^2d^2G^2(1+\alpha^2)}{\tilde{\beta}^2r^2(1-\alpha^2)^2}\right)\sum_{k=1}^{T}\mathbb{E}\big[\|\boldsymbol{g}_k\|^2\big] \le \frac{5d}{\tilde{\beta}^2}\sum_{k=1}^{T}\mathbb{E}\big[\|\nabla f(\boldsymbol{x}_k)\|^2\big] + \frac{10Tr^2L^2d^2}{\tilde{\beta}^2} + \frac{5d}{2\tilde{\beta}}\|\nabla f(\boldsymbol{x}_1)\|^2,$$

which gives (25). □

B. Proof of Theorem 4.2

Denote $\boldsymbol{w}_k \coloneqq \boldsymbol{x}_k - \alpha\boldsymbol{p}_{k-1}/(1-\alpha)$. We have

$$\boldsymbol{w}_{k+1} = \boldsymbol{x}_{k+1} - \frac{\alpha}{1-\alpha}\boldsymbol{p}_k = \boldsymbol{x}_k - \frac{1}{1-\alpha}\boldsymbol{p}_k = \boldsymbol{x}_k - \frac{\alpha}{1-\alpha}\boldsymbol{p}_{k-1} - \frac{\eta}{1-\alpha}\boldsymbol{g}_k = \boldsymbol{w}_k - \frac{\eta}{1-\alpha}\boldsymbol{g}_k,$$

and thus

$$\|\boldsymbol{w}_{k+1} - \boldsymbol{x}^*\|^2 = \|\boldsymbol{w}_k - \boldsymbol{x}^*\|^2 - \frac{2\eta}{1-\alpha}\langle\boldsymbol{w}_k - \boldsymbol{x}^*, \boldsymbol{g}_k\rangle + \frac{\eta^2}{(1-\alpha)^2}\|\boldsymbol{g}_k\|^2.$$

By taking the expectation conditioned on $\mathcal{F}_k$, we get

$$\begin{aligned}\mathbb{E}\big[\|\boldsymbol{w}_{k+1} - \boldsymbol{x}^*\|^2 \,\big|\, \mathcal{F}_k\big] &= \|\boldsymbol{w}_k - \boldsymbol{x}^*\|^2 - \frac{2\eta}{1-\alpha}\langle\boldsymbol{w}_k - \boldsymbol{x}^*, \nabla f_r(\boldsymbol{x}_k)\rangle + \frac{\eta^2}{(1-\alpha)^2}\mathbb{E}\big[\|\boldsymbol{g}_k\|^2 \,\big|\, \mathcal{F}_k\big]\\ &= \|\boldsymbol{w}_k - \boldsymbol{x}^*\|^2 - \frac{2\eta}{1-\alpha}\langle\boldsymbol{x}_k - \boldsymbol{x}^*, \nabla f_r(\boldsymbol{x}_k)\rangle\\ &\quad - \frac{2\eta\alpha}{(1-\alpha)^2}\langle\boldsymbol{x}_k - \boldsymbol{x}_{k-1}, \nabla f_r(\boldsymbol{x}_k)\rangle + \frac{\eta^2}{(1-\alpha)^2}\mathbb{E}\big[\|\boldsymbol{g}_k\|^2 \,\big|\, \mathcal{F}_k\big].\end{aligned}$$

Noticing that the convexity of f implies

$$-\langle\boldsymbol{x}_k - \boldsymbol{x}_{k-1}, \nabla f_r(\boldsymbol{x}_k)\rangle \le -(f_r(\boldsymbol{x}_k) - f_r(\boldsymbol{x}_{k-1})),$$

$$-\langle \boldsymbol{x}_k - \boldsymbol{x}^*, \nabla f_r(\boldsymbol{x}_k)\rangle \le -(f_r(\boldsymbol{x}_k) - f_r(\boldsymbol{x}^*)),$$

we get

$$\mathbb{E}\left[\|\boldsymbol{w}_{k+1} - \boldsymbol{x}^*\|^2\right] \le \mathbb{E}\left[\|\boldsymbol{w}_k - \boldsymbol{x}^*\|^2\right] - \frac{2\eta}{1-\alpha}\mathbb{E}[f_r(\boldsymbol{x}_k) - f_r(\boldsymbol{x}^*)] - \frac{2\eta\alpha}{(1-\alpha)^2}\mathbb{E}[f_r(\boldsymbol{x}_k) - f_r(\boldsymbol{x}_{k-1})] + \frac{\eta^2}{(1-\alpha)^2}\mathbb{E}\left[\|\boldsymbol{g}_k\|^2\right].$$

Now let us assume $\theta := 1 - 4\eta^2 d^2 G^2(1+\alpha^2)/[\tilde{\beta}^2 r^2(1-\alpha^2)^2] > 0$. By taking the telescoping sum over $k = 1, \ldots, T$ and plugging in the bound on $\sum_{k=1}^T \mathbb{E}\left[\|\boldsymbol{g}_k\|^2\right]$ in Lemma A.3, we get

$$\mathbb{E}\left[\|\boldsymbol{w}_{T+1} - \boldsymbol{x}^*\|^2\right] \le \|\boldsymbol{w}_1 - \boldsymbol{x}^*\|^2 - \frac{2\eta}{1-\alpha}\sum_{k=1}^T \mathbb{E}[f_r(\boldsymbol{x}_k) - f_r(\boldsymbol{x}^*)] - \frac{2\eta\alpha}{(1-\alpha)^2}\mathbb{E}[f_r(\boldsymbol{x}_T) - f_r(\boldsymbol{x}_1)] + \frac{10\eta^2 Ld}{(1-\alpha)^2\theta\tilde{\beta}^2}\sum_{k=1}^T \mathbb{E}[f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*)] + \frac{10T\eta^2 r^2 L^2 d^2}{(1-\alpha)^2\theta\tilde{\beta}^2} + \frac{5\eta^2 Ld}{(1-\alpha)^2\theta\tilde{\beta}}(f(\boldsymbol{x}_1) - f(\boldsymbol{x}^*)),$$

where we also used the inequality $\|\nabla f(\boldsymbol{x}_k)\|^2 \le 2L(f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*))$ (see equation (2.1.7) in (Nesterov, 2004)). By Lemma A.1, we get

$$\mathbb{E}\left[\|\boldsymbol{w}_{T+1} - \boldsymbol{x}^*\|^2\right] \le \|\boldsymbol{w}_1 - \boldsymbol{x}^*\|^2 - \frac{2\eta}{1-\alpha}\sum_{k=1}^T \mathbb{E}[f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*)] - \frac{2\eta\alpha}{(1-\alpha)^2}\mathbb{E}[f(\boldsymbol{x}_T) - f(\boldsymbol{x}_1)] + \frac{2\eta T r^2 L}{1-\alpha} + \frac{2\eta\alpha r^2 L}{(1-\alpha)^2} + \frac{10\eta^2 Ld}{(1-\alpha)^2\theta\tilde{\beta}^2}\sum_{k=1}^T \mathbb{E}[f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*)] + \frac{10T\eta^2 r^2 L^2 d^2}{(1-\alpha)^2\theta\tilde{\beta}^2} + \frac{5\eta^2 Ld}{(1-\alpha)^2\theta\tilde{\beta}}(f(\boldsymbol{x}_1) - f(\boldsymbol{x}^*)),$$

which further leads to

$$\begin{aligned}
&\left(1 - \frac{5\eta Ld}{(1-\alpha)\theta\tilde{\beta}^2}\right)\frac{1}{T}\sum_{k=1}^T \mathbb{E}[f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*)] \\
&\le \left(1 - \frac{5\eta Ld}{(1-\alpha)\theta\tilde{\beta}^2}\right)\frac{1}{T}\sum_{k=1}^T \mathbb{E}[f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*)] + \frac{\alpha}{1-\alpha}\frac{\mathbb{E}[f(\boldsymbol{x}_T) - f(\boldsymbol{x}^*)]}{T} \\
&\le \frac{(1-\alpha)\|\boldsymbol{x}_1 - \boldsymbol{x}^*\|^2}{2\eta T} + \frac{5\eta r^2 L^2 d^2}{(1-\alpha)\theta\tilde{\beta}^2} + r^2 L + \frac{\alpha r^2 L}{(1-\alpha)T} + \left(\frac{\alpha}{1-\alpha} + \frac{5\eta Ld}{2(1-\alpha)\theta\tilde{\beta}}\right)\frac{f(\boldsymbol{x}_1) - f(\boldsymbol{x}^*)}{T}.
\end{aligned}$$

Now, by letting

$$\eta \le \frac{(1-\alpha)\tilde{\beta}^2}{20LdT^{1/3}}, \qquad \frac{4\eta dG}{\tilde{\beta}(1-\alpha)} \le r \le \frac{G}{LT^{1/3}},$$

we see that

$$\theta \ge 1 - \frac{1}{4}\cdot\frac{(1+\alpha^2)}{(1+\alpha)^2} \ge \frac{3}{4}, \qquad 1 - \frac{5\eta Ld}{(1-\alpha)\theta\tilde{\beta}^2} \ge 1 - \frac{1}{3T^{1/3}} \ge \frac{2}{3},$$

and then we obtain

$$\begin{aligned}
&\frac{1}{T}\sum_{k=1}^T \mathbb{E}[f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*)] \\
&\le \frac{3(1-\alpha)\|\boldsymbol{x}_1 - \boldsymbol{x}^*\|^2}{4\eta T} + \frac{G^2 d}{2LT} + \frac{3G^2}{2LT^{2/3}} + \frac{3\alpha G^2}{2(1-\alpha)LT^{5/3}} + \frac{3}{2}\left(\frac{\alpha}{1-\alpha} + \frac{10\eta Ld}{3(1-\alpha)\tilde{\beta}}\right)\frac{f(\boldsymbol{x}_1) - f(\boldsymbol{x}^*)}{T}, \\
&\le \frac{3(1-\alpha)\|\boldsymbol{x}_1 - \boldsymbol{x}^*\|^2}{4\eta T} + \frac{3G^2}{2LT^{2/3}} + \mathcal{O}(d/T),
\end{aligned}$$

which implies the desired bound as f is convex.

C. Proof of Theorem 4.3

We let $\boldsymbol{w}_k \coloneqq \boldsymbol{x}_k - \alpha \boldsymbol{p}_{k-1}/(1-\alpha)$, and it can be checked that $\boldsymbol{w}_{k+1} = \boldsymbol{w}_k - \eta \boldsymbol{g}_k/(1-\alpha)$. By the L-smoothness of f, we have

$$\begin{aligned} f(\boldsymbol{w}_{k+1}) &\le f(\boldsymbol{w}_k) - \frac{\eta}{1-\alpha}\langle \boldsymbol{g}_k, \nabla f(\boldsymbol{w}_k)\rangle + \frac{\eta^2 L}{2(1-\alpha)^2}\|\boldsymbol{g}_k\|^2 \\ &= f(\boldsymbol{w}_k) - \frac{\eta}{1-\alpha}\langle \boldsymbol{g}_k, \nabla f(\boldsymbol{w}_k) - \nabla f(\boldsymbol{x}_k)\rangle - \frac{\eta}{1-\alpha}\langle \boldsymbol{g}_k, \nabla f(\boldsymbol{x}_k)\rangle + \frac{\eta^2 L}{2(1-\alpha)^2}\|\boldsymbol{g}_k\|^2. \end{aligned}$$

By taking the expectation conditioned on $\mathcal{F}_k$, we get

$$\begin{aligned} \mathbb{E}[f(\boldsymbol{w}_{k+1})|\mathcal{F}_k] \le f(\boldsymbol{w}_k) &- \frac{\eta}{1-\alpha}\langle \nabla f_r(\boldsymbol{x}_k), \nabla f(\boldsymbol{w}_k) - \nabla f(\boldsymbol{x}_k)\rangle - \frac{\eta}{1-\alpha}\langle \nabla f_r(\boldsymbol{x}_k), \nabla f(\boldsymbol{x}_k)\rangle \\ &+ \frac{\eta^2 L}{2(1-\alpha)^2}\mathbb{E}[\|\boldsymbol{g}_k\|^2|\mathcal{F}_k]. \end{aligned}$$

Note that

$$\begin{aligned} & -\langle \nabla f_r(\boldsymbol{x}_k), \nabla f(\boldsymbol{w}_k) - \nabla f(\boldsymbol{x}_k)\rangle \\ =\ & -\langle \nabla f(\boldsymbol{x}_k), \nabla f(\boldsymbol{w}_k) - \nabla f(\boldsymbol{x}_k)\rangle - \langle \nabla f_r(\boldsymbol{x}_k) - \nabla f(\boldsymbol{x}_k), \nabla f(\boldsymbol{w}_k) - \nabla f(\boldsymbol{x}_k)\rangle \\ \le\ & \frac{1}{4}\|\nabla f(\boldsymbol{x}_k)\|^2 + \|\nabla f(\boldsymbol{w}_k) - \nabla f(\boldsymbol{x}_k)\|^2 + \|\nabla f_r(\boldsymbol{x}_k) - \nabla f(\boldsymbol{x}_k)\|^2 + \frac{1}{4}\|\nabla f(\boldsymbol{w}_k) - \nabla f(\boldsymbol{x}_k)\|^2 \\ =\ & \frac{1}{4}\|\nabla f(\boldsymbol{x}_k)\|^2 + \|\nabla f_r(\boldsymbol{x}_k) - \nabla f(\boldsymbol{x}_k)\|^2 + \frac{5}{4}\|\nabla f(\boldsymbol{w}_k) - \nabla f(\boldsymbol{x}_k)\|^2 \\ \le\ & \frac{1}{4}\|\nabla f(\boldsymbol{x}_k)\|^2 + L^2 r^2 + \frac{5}{4}L^2\|\boldsymbol{w}_k - \boldsymbol{x}_k\|^2 \qquad \text{(by Lemma A.1 and } L\text{-Smoothness of } f) \\ =\ & \frac{1}{4}\|\nabla f(\boldsymbol{x}_k)\|^2 + L^2 r^2 + \frac{5L^2\alpha^2}{4(1-\alpha)^2}\|\boldsymbol{p}_{k-1}\|^2, \end{aligned}$$

and

$$\begin{aligned} -\langle \nabla f_r(\boldsymbol{x}_k), \nabla f(\boldsymbol{x}_k)\rangle =\ & -\|\nabla f(\boldsymbol{x}_k)\|^2 - \langle \nabla f_r(\boldsymbol{x}_k) - \nabla f(\boldsymbol{x}_k), \nabla f(\boldsymbol{x}_k)\rangle \\ \le\ & -\|\nabla f(\boldsymbol{x}_k)\|^2 + \|\nabla f_r(\boldsymbol{x}_k) - \nabla f(\boldsymbol{x}_k)\|^2 + \frac{1}{4}\|\nabla f(\boldsymbol{x}_k)\|^2 \\ =\ & -\frac{3}{4}\|\nabla f(\boldsymbol{x}_k)\|^2 + \|\nabla f_r(\boldsymbol{x}_k) - \nabla f(\boldsymbol{x}_k)\|^2 \\ \le\ & -\frac{3}{4}\|\nabla f(\boldsymbol{x}_k)\|^2 + L^2 r^2. \end{aligned}$$

Therefore

$$\begin{aligned} \mathbb{E}[f(\boldsymbol{w}_{k+1})|\mathcal{F}_k] \le f(\boldsymbol{w}_k) &- \frac{\eta}{2(1-\alpha)}\|\nabla f(\boldsymbol{x}_k)\|^2 + \frac{2\eta r^2 L^2}{1-\alpha} \\ &+ \frac{5\eta L^2\alpha^2}{4(1-\alpha)^3}\|\boldsymbol{p}_{k-1}\|^2 + \frac{\eta^2 L}{2(1-\alpha)^2}\mathbb{E}[\|\boldsymbol{g}_k\|^2|\mathcal{F}_k]. \end{aligned}$$

Now, by taking the total expectation and summing over $k = 1, \ldots, T$, we get

$$\begin{aligned} \mathbb{E}[f(\boldsymbol{w}_{T+1})] \le f(\boldsymbol{w}_1) &- \frac{\eta}{2(1-\alpha)}\sum_{k=1}^{T}\mathbb{E}[\|\nabla f(\boldsymbol{x}_k)\|^2] + \frac{2\eta r^2 L^2 T}{1-\alpha} \\ &+ \frac{5\eta L^2\alpha^2}{4(1-\alpha)^3}\sum_{k=1}^{T-1}\mathbb{E}[\|\boldsymbol{p}_k\|^2] + \frac{\eta^2 L}{2(1-\alpha)^2}\sum_{k=1}^{T}\mathbb{E}[\|\boldsymbol{g}_k\|^2]. \end{aligned}$$

Assuming $\theta \coloneqq 1 - 4\eta^2 d^2 G^2 (1+\alpha^2)/[\tilde{\beta}^2 r^2 (1-\alpha^2)^2] > 0$ and plugging in the bounds in Lemma A.3, we get

$$\mathbb{E}[f(\boldsymbol{w}_{T+1})] \leq f(\boldsymbol{w}_1) - \frac{\eta}{2(1-\alpha)} \sum_{k=1}^{T} \mathbb{E}[\|\nabla f(\boldsymbol{x}_k)\|^2] + \frac{2\eta r^2 L^2 T}{1-\alpha} + \frac{\eta^2 L}{2(1-\alpha)^2} \left(1 + \frac{5\eta L \alpha^2 (1+\alpha^2)}{(1-\alpha)^3 (1+\alpha)^2}\right) \left(\frac{5d}{\theta \tilde{\beta}^2} \sum_{k=1}^{T} \mathbb{E}\big[\|\nabla f(\boldsymbol{x}_k)\|^2\big] + \frac{10 T r^2 L^2 d^2}{\theta \tilde{\beta}^2} + \frac{5d}{2\theta \tilde{\beta}} \|\nabla f(\boldsymbol{x}_1)\|^2 \right).$$

Now let T be sufficiently large such that

$$T^{1/3} \geq \frac{\alpha^2 \tilde{\beta}^2}{2d(1-\alpha)^2},$$

and let

$$\eta \leq \frac{(1-\alpha)\tilde{\beta}^2}{20 L d T^{1/3}}, \qquad \frac{4\eta d G}{\tilde{\beta}(1-\alpha)} \leq r \leq \frac{G}{L T^{1/3}},$$

We then have $\eta \leq (1-\alpha)^3/(10\alpha^2 L)$, and thus

$$\theta \geq 1 - \frac{1}{4} \cdot \frac{(1+\alpha^2)}{(1+\alpha)^2} \geq \frac{3}{4}, \qquad \frac{5\eta L \alpha^2 (1+\alpha^2)}{(1-\alpha)^3 (1+\alpha)^2} \leq \frac{1+\alpha^2}{2(1+\alpha)^2} \leq \frac{1}{2}.$$

Therefore

$$\begin{aligned}
\mathbb{E}[f(\boldsymbol{w}_{T+1})] &\leq f(\boldsymbol{w}_1) - \frac{\eta}{2(1-\alpha)} \left(1 - \frac{10\eta L d}{(1-\alpha)\tilde{\beta}^2}\right) \sum_{k=1}^{T} \mathbb{E}[\|\nabla f(\boldsymbol{x}_k)\|^2] \\
&\quad + \frac{2\eta r^2 L^2 T}{1-\alpha} + \frac{\eta^2 L}{(1-\alpha)^2 \tilde{\beta}^2} \cdot 10 T r^2 L^2 d^2 + \frac{5\eta^2 L d}{2(1-\alpha)^2 \tilde{\beta}} \|\nabla f(\boldsymbol{x}_1)\|^2 \\
&\leq f(\boldsymbol{x}_1) - \frac{\eta}{4(1-\alpha)} \sum_{k=1}^{T} \mathbb{E}[\|\nabla f(\boldsymbol{x}_k)\|^2] + \frac{\eta}{1-\alpha} \cdot 2G^2 T^{1/3} + \frac{\eta}{1-\alpha} \cdot \frac{G^2 d}{2} + \frac{\eta}{1-\alpha} \frac{\tilde{\beta}}{8T^{1/3}} \|\nabla f(\boldsymbol{x}_1)\|^2.
\end{aligned}$$

Since $f^* \leq \mathbb{E}[f(\boldsymbol{w}_{T+1})]$, it leads to

$$\frac{1}{T} \sum_{k=1}^{T} \mathbb{E}\big[\|\nabla f(\boldsymbol{x}_k)\|^2\big] \leq \frac{4(1-\alpha)(f(\boldsymbol{x}_1) - f^*)}{\eta T} + \frac{8G^2}{T^{2/3}} + \frac{2G^2 d}{T} + \frac{\tilde{\beta} \|\nabla f(\boldsymbol{x}_1)\|^2}{2T^{4/3}}.$$