HARVARD
School of Engineering and Applied Sciences

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

Xin Chen, Yujie Tang, Na Li

John A. Paulson School of Engineering and Applied Sciences, Harvard University
Email: xchen2@g.harvard.edu

Zeroth-Order Optimization (ZO)

Single-Point ZO (SZO)

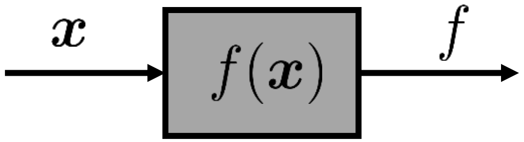
$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \eta \frac{d}{r} \underbrace{f(\boldsymbol{x}_k + r\boldsymbol{u}_k)\boldsymbol{u}_k}$$

single-point gradient estimator

- ✓ one function evaluation → online problems
- ✗ large variance and slow convergence.

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

unknown



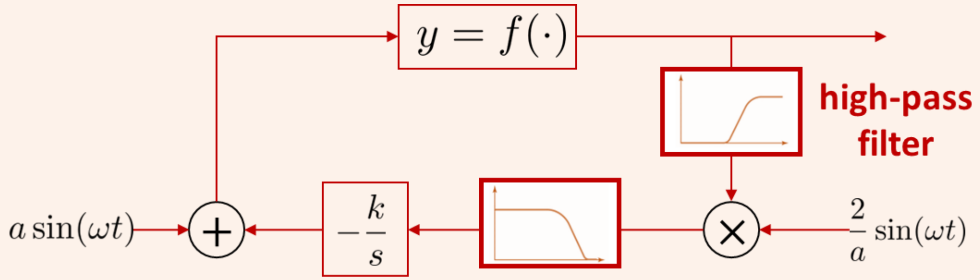
only function evaluation is available

“Can we borrow the idea of high-pass and low-pass filters from ES control to improve SZO?”

Extremum Seeking (ES) Control

$$\dot{\boldsymbol{x}} = -k \cdot \frac{2}{a} f(\boldsymbol{x} + a\sin(\boldsymbol{\omega}t))\sin(\boldsymbol{\omega}t)$$

continuous-time dynamics



“Can we borrow the idea of high-pass and low-pass filters to improve SZO?”

YES!

Vanilla **SZO**

$$x_{k+1} = x_k - \eta \frac{d}{r} f(x_k + r u_k) u_k$$

$+$

High-pass **F**ilter

$$\frac{s}{s+\omega_H}$$

&

Low-pass **F**ilter

$$\frac{\omega_L}{s+\omega_L}$$

recycled
from last iteration

Our proposed **HLF-SZO**

$=$

$$x_{k+1} = x_k - \eta \cdot \frac{d}{r}\Big(f(x_k + r u_k) - f(x_{k-1} + r u_{k-1})\Big) u_k + \alpha(x_k - x_{k-1})$$

residual feedback [1]

“momentum”

[1] Yan Zhang, et al. A new one-point residual-feedback oracle for black-box learning and control. Automatica, 2021.

Performance Comparison

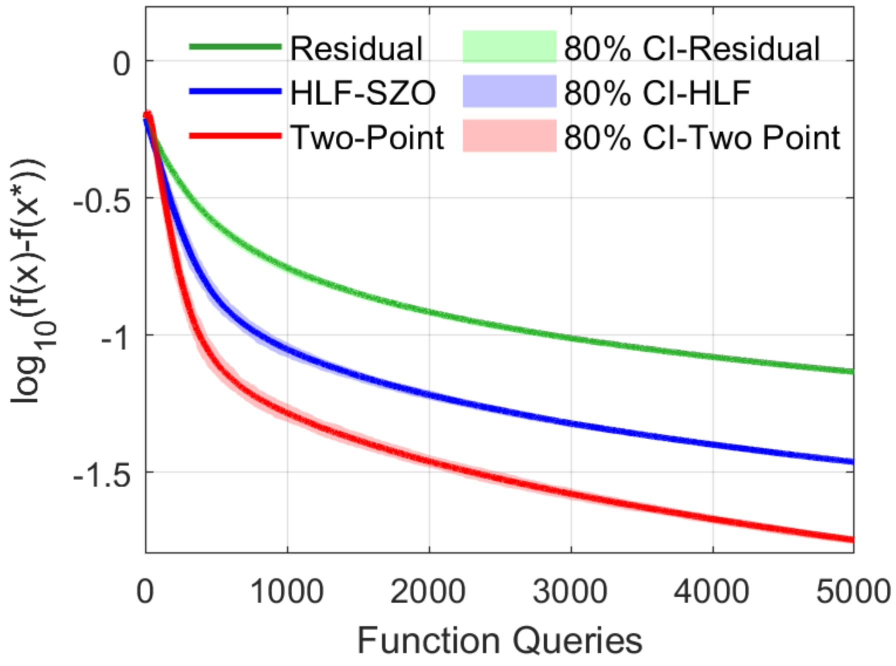
$\min_{\boldsymbol{x}\in\mathbb{R}^d} f(\boldsymbol{x})$	**Convex**	**Nonconvex**

Vanilla SZO < Residual SZO < **HLF-SZO** < Two-Point

Vanilla SZO	Residual SZO	HLF-SZO	Two-Point
$\mathcal{O}(d^2/\epsilon^3)$	$\mathcal{O}(d^2/\epsilon^{\frac{3}{2}})$	$\mathcal{O}(d^{\frac{3}{2}}/\epsilon^{\frac{3}{2}})$	$\mathcal{O}(d/\epsilon)$

- Logistic Regression (d=50, N=1000)

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



X. Chen, Y. Tang, N. Li, “Improve Single-Point Zeroth-Order Optimization Using High-Pass and Low-Pass Filters”, arXiv:2111.01701, 2022.

Takeaway

- We borrow high-pass and low-pass filters from ES control and develop **HLF-SZO** method with

  ✓ **smaller variance** ✓ **faster convergence**

-

Control		**Optimization**
high-pass filter	⟷	residual feedback
low-pass filer	⟷	momentum

- We explore a new direction to improve ZO schemes by leveraging the close connection between ZO and continuous-time ES control.