Private Zeroth-Order Optimization with Public Data

Xuchen Gong and Tian Li (University of Chicago)

Challenges

- First-order differentially private (DP) algorithms has high computation and memory cost
- Zeroth-order DP methods are efficient to privatize as they leverage function evaluations to approximate gradients
- However, zeroth-order approaches suffer from low utilities and limited application scenarios

Our Insights

We propose to **leverage public information** (batch gradient, which is efficient to compute) to improve private zeroth-order gradient estimation; we introduce the first set of **public-data-assisted zeroth-order optimizers (PAZO)**

PAZO achieves

- stronger privacy/utility tradeoffs across **vision and language tasks** in both **pre-training and fine-tuning settings**
- outperform first-order methods (with public gradients) under tight privacy
- 16× runtime speedup

Background: DPZero

1. Sample perturbation u uniformly from sphere $\sqrt{d}\mathbb{S}^{d-1}$

2. $g \leftarrow \left(\frac{1}{|B|}\sum_{\xi\in B}\mathrm{clip}_C\left(\frac{f(x+\lambda u;\xi)-f(x-\lambda u;\xi)}{2\lambda}\right)+z\right)u \quad z\sim\frac{1}{|B|}\mathcal{N}(0,qC^2\sigma^2)$

Function queries to privatize
by clipping and adding noise

Zhang, Liang, et al. "Dpzero: Private fine-tuning of language models without backpropagation." ICML 2024.

PAZO-M: Mixing Zeroth and First-Order Gradients

$$x \leftarrow \eta(\alpha g_{\text{pub}} + (1-\alpha) g_{\text{pri}})$$

First-order grad
on public data

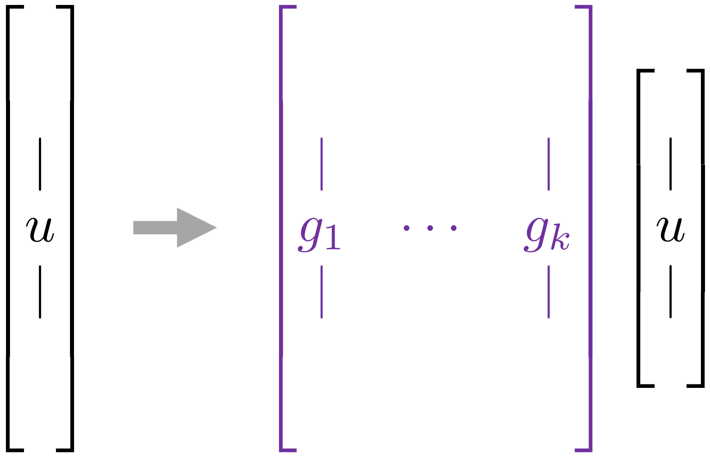
Zeroth-order grad
on private data

Sample u uniformly from sphere $d^{\frac{1}{4}}\mathbb{S}^{d-1}$

$$g_{\text{pri}} \leftarrow \left(\frac{1}{|B|} \sum_{\xi \in B} \text{clip}_C \left(\frac{f(x+\lambda u;\xi) - f(x-\lambda u;\xi)}{2\lambda} \right) + z \right) u$$

PAZO-P: Sampling in Public Gradient Subspace

(ortho)normalized gradients
on k batches of public data



$\sqrt{d}\mathbb{S}^{d-1}$

Prior work

$G(\sqrt{k}\mathbb{S}^{k-1})$

PAZO-P

Sample u uniformly from sphere $\sqrt{k}\mathbb{S}^{k-1}$

$$g \leftarrow \left(\frac{1}{|B|}\sum_{\xi \in B} \mathrm{clip}_C\left(\frac{f(x+\lambda Gu;\xi)-f(x-\lambda Gu;\xi)}{2\lambda}\right)+z\right)Gu$$

PAZO-S: Select the Best Public Gradient

1. Find the best public descent direction

$$\left\{ \begin{matrix} | & & | \\ g_1, & \cdots, & g_k \\ | & & | \end{matrix} \right\} \rightarrow \{ f(x-\eta g_1), \quad \cdots, \quad f(x-\eta g_k) \} \quad \text{with privatization}$$

$$\rightarrow j^\star \leftarrow \arg\min_{j\in[k]} \mathrm{priv}(f(x-\eta g_j))$$

2. Perturb the best candidate and compare

$$\mathrm{priv}(f(x-\eta(g_{j^\star}+z'))) \overset{?}{<} \mathrm{priv}(f(x-\eta g_{j^\star})) \qquad z' \sim \mathcal{N}(0, \epsilon^2 I_d)$$

If yes, use $g_{j^\star} + z'$
Else, use $g_{j^\star}$

Convergence

[**γ-similar**] Public B' and private data B are γ-similar if $\|\nabla f(x;B) - \nabla f(x;B')\| \le \gamma, \forall x$

Our assumptions: L-smooth, M-lipschitz, γ-similar, and optionally $|f(x;B)| \le S, \forall x$

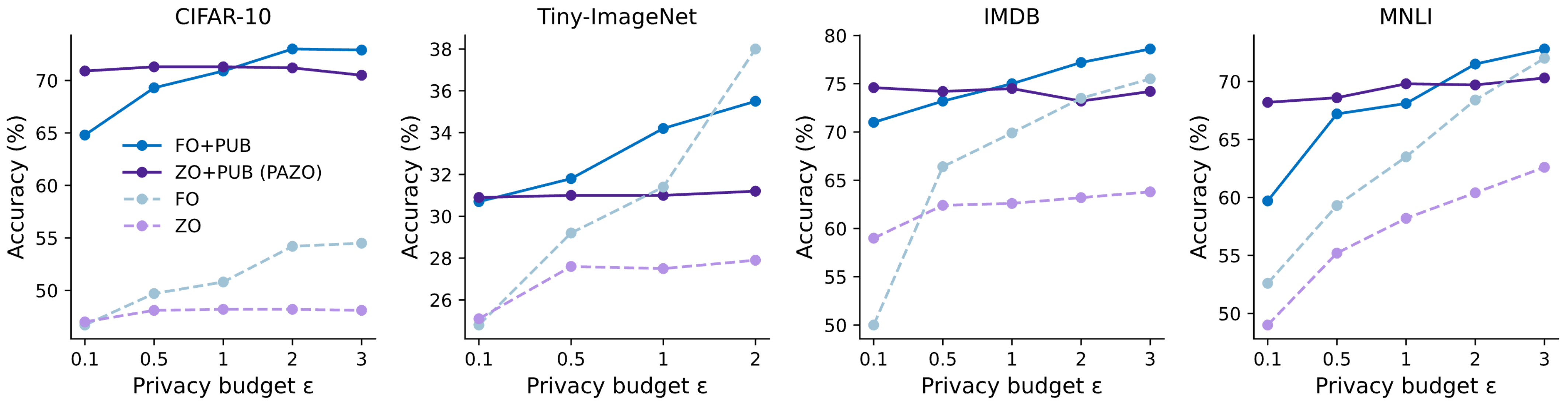
Method	wo. $\lvert f(x)\rvert \le S$	w. $\lvert f(x)\rvert \le S$
DP-SGD	$O(\sqrt{d})$	/
DPZero	/	$O(\sqrt{d}\log d)$
PAZO-M	$O(\frac{1-\alpha}{\alpha}\sqrt{d})$	$O(\frac{1-\alpha}{\alpha}d^{\frac{1}{4}})$
PAZO-P	$O(k)$	$O(\sqrt{k}\log k)$
PAZO-S	$O(c)$	

c is constant independent of k and d.

PAZO-M improves prior work by $d^{\frac{1}{4}}\log d$

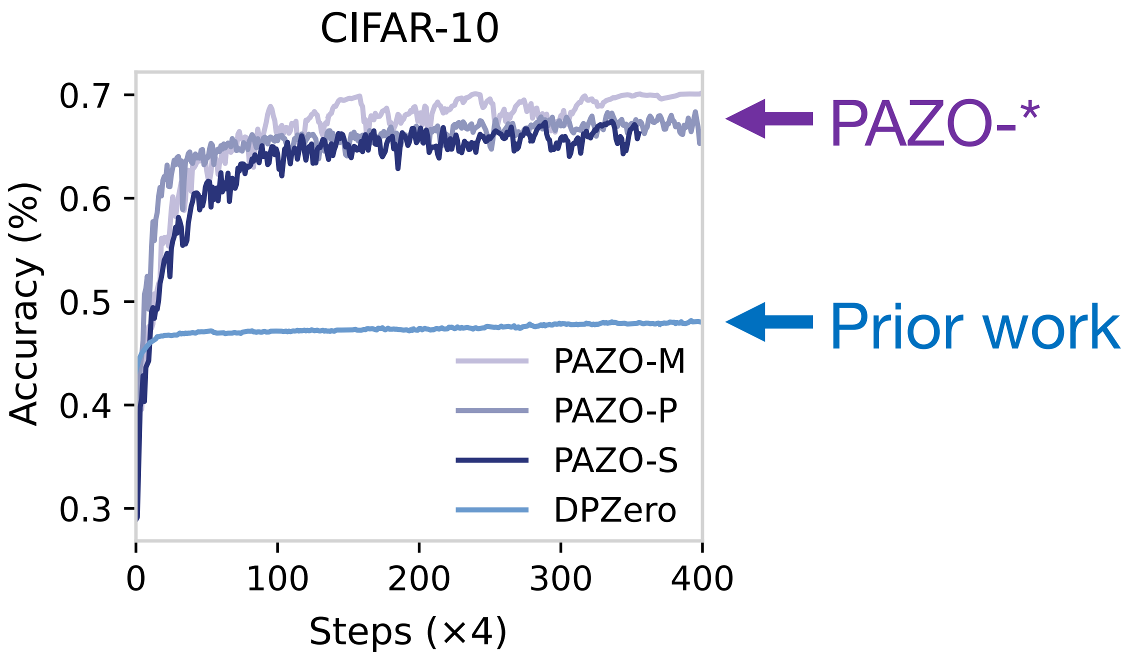
PAZO-{P,S} achieve d-independent rates

Improved Privacy/Utility Tradeoffs



- **Without public data**, vanilla **zeroth-order (ZO)** underperforms **first-order methods (FO)**
- **With public data**, **PAZO** outperforms the best **first-order methods with public data (FO+PUB)**

Time Efficiency



Slow convergence is a known disadvantage of zeroth-order methods; PAZO-* **converges faster** than vanilla ZO (DPZero)

Future Work

- Sharpen the convergence bounds by considering other data similarity metrics
- Explore a broader set of (public, private) dataset pairs in practical DP applications

Paper: openreview.net/pdf?id=zytlTzY4IW

Code: github.com/xuchengong/pazo

Check out our poster at **Exhibit Hall C,D,E** during **4 Dec 11-14:00 PST**