Efficient ML Algorithms: From Theory to Practice

Shaocong Ma

October 9, 2025

University of Maryland, College Park

UNIVERSITY OF MARYLAND

Today's Topic

- *Revisiting Zeroth-Order Optimization: Minimum-Variance Two-Point Estimators and Directionally Aligned Perturbations* (ICLR 2025 Spotlight)
  Shaocong Ma, Heng Huang.
  - We developed the DAP, a zeroth-order optimization method specifically designed to overcome key bottlenecks in large language model training and AI4Science tasks.
- Future Research
  - Efficient Inference-Time Training of Large Language Models
  - Fine-Tuning the LLM Agent using Zeroth-Order Optimization
  - Addressing the Theoretical Limit of Multi-Level Optimization

Motivation

- Zeroth-order optimization
  - Training LLMs: Saving GPU memory
  - Training Physics Models: Non-differentiable module
- However, existing gradient estimator fails to focus on the "important directions" .
  - Existing literature indicates that some neurons in LLMs are more important.
  - Our work[1] indicates that the mesh optimization has "important nodes".

[1] Shaocong Ma, James Diffenderfer, Bhavya Kailkhura, and Yi Zhou. "*Deep learning of PDE correction and mesh adaption without automatic differentiation.*" Machine Learning 114.3 (2025): 1-25.

DAPs: Directionally Aligned Perturbation

Free Lunch! Identifying the important direction just requires to slightly modify the random perturbation in traditional two-point estimator:

$$\hat{\nabla} f(x) = \frac{f(x+\mu v) - f(x)}{\mu} v$$

Here, μ is a small perturbation size. v is a random vector.

- Traditional Method:
  - $\mathbb{E}[vv^\top] = I$ (Unbiasedness)
  - $\|v\|$ is fixed (Minimum Variance).
- DAP:
  - $\mathbb{E}[vv^\top] = I$ (Unbiasedness)
  - $\nabla f(x)^\top v$ is fixed (Minimum Variance).

Traditional Methods Cannot Identify the Important Directions

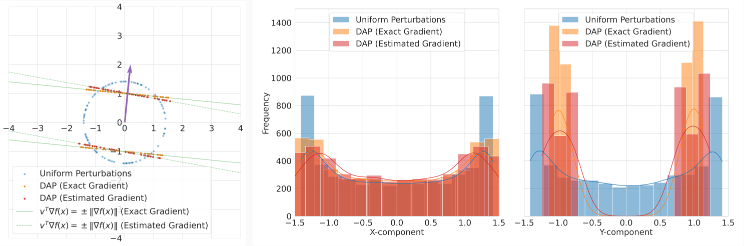

Figure 1: Illustration of the *directional alignment property* of DAP in $d = 2$ with estimating the gradient of $f(x) = x_1^2 + x_2^2$ at $x = \begin{bmatrix} 0.1 & 1 \end{bmatrix}^\top$. Traditional zeroth-order estimator is symmetric, but we really need a non-symmetric estimator.

Traditional Methods Cannot Identify the Important Directions

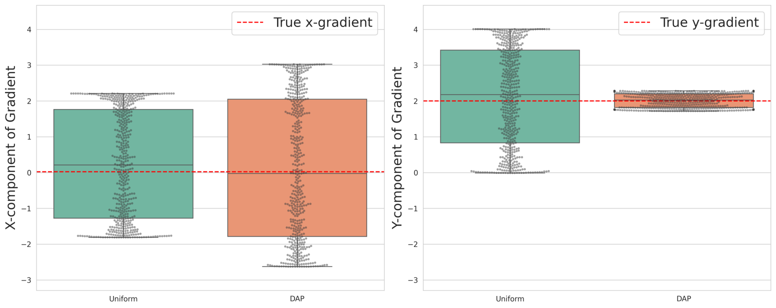

Figure 2: Comparison of gradient estimation performance with estimating the gradient of $f(x) = x_1^2 + x_2^2$ at $x = \begin{bmatrix} 0.1 & 1 \end{bmatrix}^\top$ between uniform random perturbations and DAPs. The non-symmetric estimator is more accurate in the direction with larger gradient.

Practical Experiments

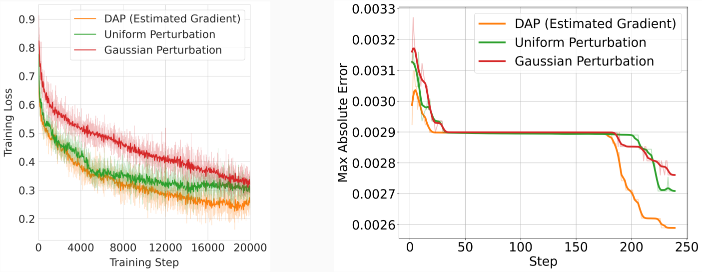

Figure 3: Comparison of training loss curves among different random perturbations on Large Language Model Fine Tuning and Mesh Optimization for the Physical Numerical Solver .

Future Research

Efficient Inference-Time Training of Large Language Models

- The generation process of LLMs is a Markov decision process.
  Policy:
  $$p_\theta(\text{Next Token} \mid \text{Previous Token})$$
  Transition:
  $$\mathbb{P}(s' = \text{Next Token} \mid s = \text{Previous Token}, a = \text{Next Token}) = 1$$
- Maximize the Reward Given by the Process Reward Model:
  $$\theta \leftarrow \theta + \eta \times \text{Reward} \times \nabla_\theta \log p_\theta$$
- Efficient Algorithms (Connections to Previous Research):
  - (NIPS 2020) Variance-reduced off-policy TDC learning: Non-asymptotic convergence analysis
    Shaocong Ma, Yi Zhou, Shaofeng Zou.
  - (ICLR 2021) Greedy-GQ with Variance Reduction: Finite-time Analysis and Improved Complexity
    Shaocong Ma, Ziyi Chen, Yi Zhou, Shaofeng Zou.

Fine-Tuning the LLM Agent using Zeroth-Order Optimization

- A LLM agent contains many non-differentiable modules
  - Black-Box LLM parameters (temperature, LoRA weights, ...)
  - Prompts
  - RAG, Search, Tool Use, ...
- Efficient fine-tuning requires advanced techniques:
  - (Submitted) Hybrid Fine-Tuning of LLMs: Theoretical Insights on Generalized Smoothness and Convergence
    Shaocong Ma, Peiran Yu, Heng Huangs.
  - (ICLR 2025 Spotlight) Revisiting Zeroth-Order Optimization: Minimum-Variance Two-Point Estimators and Directionally Aligned Perturbations
    Shaocong Ma, Heng Huang.

Addressing the Theoretical Limit of Multi-Level Optimization

- Min-Max Optimization:

$$\min_{\theta} \max_{\omega} f(\theta, \omega)$$

The solution (θ^*, ω^*) is a Nash equilibrium of the min-max game.

- Multi-Level Optimization:

$$\begin{aligned} \min_{\theta} \quad & f(\theta, \omega_1, \omega_2, \ldots, \omega_N) \\ \text{s.t.} \quad & \omega_1 \in \operatorname{arg\,min} g_1(\theta, \omega_1, \omega_2, \ldots, \omega_N) \\ & \omega_2 \in \operatorname{arg\,min} g_2(\theta, \omega_1, \omega_2, \ldots, \omega_N) \\ & \vdots \end{aligned}$$

- The solution belongs to Correlated Equilibria:
  - (JMLR 2023) Decentralized Robust V-Learning for Solving Markov Games with Model Uncertainty
    Shaocong Ma, Ziyi Chen, Shaofeng Zou, Yi Zhou.
  - (NIPS 2022) Finding Correlated Equilibrium of Constrained Markov Game: A Primal-Dual Approach
    Ziyi Chen, **Shaocong Ma**, Yi Zhou.