Zhewei Yao | Curriculum Vitae

Soda 465, Berkeley, CA 94704

+1 707-742-9179 • ✉ zheweiy@berkeley.edu • 🌐 yaozhewei.github.io
in zhewei-yao • ⓞ yaozhewei

I am a Ph.D. student in the RISELab (former AMPLab), BDD and Math Department at University of California at Berkeley. I am advised by Michael Mahoney. My research interest lies in computing statistics, optimization and machine learning. Currently, I am interested in leveraging tools from randomized linear algebra to provide efficient and scalable solutions for large-scale optimization and learning problems. I am also working on the theory and application of deep learning.

Education

University of California at Berkeley   **CA, USA**
Ph.D. in Applied Mathematics, Department of Mathematics   *Sep. 2016–Present*

Shanghai Jiao Tong University   **Shanghai China**
B.S. in Applied Mathematics, Zhiyuan Honor College   *Sep. 2012–Jun. 2016*

Publications

Shallow Learning for Fluid Flow Reconstruction with Limited Sensors and Limited Data
NB Erichson, L Mathelin, Z Yao, SL Brunton, MW Mahoney, JN Kutz
arxiv preprint 1902.07358

JumpReLU: A Retrofit Defense Strategy for Adversarial Attacks
B Erichson, **Z Yao***, *MW Mahoney*
Under review

Trust Region Based Adversarial Attack on Neural Networks
***Z Yao**, A Gholami, P Xu, K Keutzer, MW Mahoney*
arxiv preprint 1812.06371
Proc. CVPR 2019

Parameter Re-Initialization through Cyclical Batch Scheduling
N Mu, **Z Yao***, *A Gholami, K Keutzer, MW Mahoney*
arxiv preprint 1812.01216
Proc. MLSYS Workshop at NeurIPS 2018

On the Computational Inefficiency of Large Batch Sizes for Stochastic Gradient Descent
*N Golmant, N Vemuri, **Z Yao**, V Feinberg, A Gholami, K Rothauge, MW Mahoney, J Gonzalez*
arxiv preprint 1811.12941
Under Review

Large batch size training of neural networks with adversarial training and second-order information

Z Yao*, A Gholami*, K Keutzer, MW Mahoney

arxiv preprint [1810.01021](#)

Under Review

Hessian-based Analysis of Large Batch Training and Robustness to Adversaries

Z Yao*, A Gholami*, Q Lei K Keutzer, MW Mahoney

arxiv preprint [1802.08241](#)

Proc. NeurIPS 2018

Inexact non-convex Newton-type methods

Z Yao, P Xu, F Roosta-Khorasani, MW Mahoney

arxiv preprint [1802.06925](#)

Under review

A hybrid adaptive MCMC algorithm in function spaces

Q Zhou, Z Hu, Z Yao, J Li

arxiv preprint [1607.01458](#)

SIAM/ASA Journal on Uncertainty Quantification 5 (1), 621-639

On an adaptive preconditioned Crank–Nicolson MCMC algorithm for infinite dimensional Bayesian inference

Z Hu*, Z Yao*, J Li

arxiv preprint [1511.05838](#)

Journal of Computational Physics 332, 492-503

A TV-Gaussian prior for infinite-dimensional Bayesian inverse problems and its numerical implementation

Z Yao*, Z Hu*, J Li

arxiv preprint [1510.05239](#)

Inverse Problems 32 (7), 075006 (*Highlight Paper*)

Research Experiences

University of California at Berkeley   **CA, USA**

Ph.D. Researcher at RiseLab and BDD   *Sep. 2016–Present*

- Use Hessian information to train neural networks for large batch training

- Investigate the scaling behavior of stochastic gradient descent with large batch sizes for neural networks

- Propose stochastic variants of 2nd-order methods for non-convex optimization problem and establish theories

- Apply deep learning to other fields, e.g. scientific datasets and fluid dynamics

- Analyzed large batch training and robustness of neural networks with Hessian information

Lawrence Berkeley Notional Laboratory   **CA, USA**

Researcher intern at NERSC   *May. 2018–Aug. 2018*

- Implemented CPU Parallelization of PyTorch to train large climate dataset (over 400 Gb)

- Tested robustness on models trained with scientific datasets

○ **Shanghai Jiao Tong University** **Shanghai, China**
Undergraduate Researcher *Sep. 2014–Jun. 2016*
  - Considered MCMC algorithm in infinite-dimensional space

  - Designed a TG-prior with better edge-preserving property and two new adaptive algorithms

Others

- **Programming Languages:** C++, Matlab, Python, Pytorch, Tensorflow

- **Conference Reviewer:** NeurIPS 2018, ICLR 2019

- **Teaching:**
  Stat 89A: Linear Algebra for Data Science **UC Berkeley**
  Graduate Student Instructor *Spring 2018*

  Math 16A: Analytic Geometry and Calculus **UC Berkeley**
  Graduate Student Instructor *Spring 2017 & Fall 2016*