

Shixun Wu

Email: swu264@ucr.edu ◊  shixun404

EDUCATION

University of California, Riverside

Ph.D Student in Computer Science

Sep. 2022 - Present

Columbia University

M.S. in Science

Sep. 2020 - May 2022

Peking University

B.S. in Computer Science





B.S. in Economics(Double Major)

Sep. 2016 - Jul. 2020

WORK EXPERIENCE

1. Graduate Researcher@UCR&Lawrence Berkeley National Laboratory *Sep. 2022 - Present*
 - DECODE: Data-driven Exascale Control of Optically Driven Excitations in Chemical and Material Systems
 - Implement GEMM&FFT from scratch, achieving comparable or faster performance with cuBLAS/cuFFT.
 - Explore fault tolerant GEMM&FFT with a low overhead of 5% ~ 10% compared to cuBLAS/cuFFT.
2. Main Contributor@AI4Finance Foundation&Columbia University *Aug. 2021 - Jul. 2022*
 - Develop multi-agent reinforcement learning algorithms in ElegantRL, a RL library with 3.1k starts on GitHub.
 - Co-leader of ElegantRL_Solver, a high-performance solver, outperforming Gurobi for dense cases in MaxCut Problem.
3. Machine Learning Engineer Intern@Noah Ark's Lab *Jul. 2020 - May 2021*
 - *MineRL, NeurIPS 2020 Competition*, Rank top 1 among 90+ teams.
 - Research on multi-agent reinforcement learning, worked on agent cooperation.
4. Algorithm Engineer Intern @ Megvii *Aug. 2019 - Jan. 2020*
 - Reproduced a real-time 3D reconstruction paper Double-Fusion.
 - Accelerate the reconstruction algorithm with cuBLAS.

SELECTED GITHUB REPO

1.  Fault-Tolerant-GEMM-on-NVIDIA-GPUs. [code][paper]
2.  High-Performance FFT implementation on GPUs. (*Released soon.*)
3.  ElegantRL, a popular reinforcement learning library (received over 3.1k GitHub stars). [code]
4.  ElegantRL_Solver, a high-performance RL Solvers.[code]

TECHNICAL STRENGTHS

Programming Languages: C/C++, CUDA, SIMD, OpenMP, Python, PyTorch.

SELECTED PUBLICATIONS

1. [ICS'23] **Shixun Wu***, Yujia Zhai*, Jinyang Liu, Jiajun Huang, Zizhe Jian, Bryan Wong, Zizhong Chen. "Anatomy of High-Performance GEMM with Online Fault Tolerance on GPUs." *The 37th ACM International Conference on Supercomputing*, Orlando, FL, USA, June 21–23, 2023. DOI: 10.1145/3577193.3593715.[paper]
2. [HPDC'23] **Shixun Wu***, Yujia Zhai*, Jiajun Huang, Zizhe Jian, Zizhong Chen. "FT-GEMM: A Fault Tolerant High Performance GEMM Implementation on x86 CPUs." *The 32nd ACM International Symposium on High-Performance Parallel and Distributed Computing*, Orlando, FL, USA, June 21–23, 2023. DOI: 10.1145/3588195.3595947.[paper]
3. [ICLR'23] Xiaoyang Liu, Zechu Li, **Shixun Wu**, Xiaodong Wang. "Stationary Deep Reinforcement Learning With Quantum K-Spin Hamiltonian Regularization." *ICLR 2023 Workshop on Physics for Machine Learning. 2023*
4. [TSP'23] Jeremy Johnston, Xiaoyang Liu, **Shixun Wu**, Xiaodong Wang. "A Curriculum Learning Approach to Optimization with Application to Downlink Beamforming." *IEEE Transactions on Signal Processing (2023)*, Major Revision.

5. **Shixun Wu**, Yujia Zhai, Jinyang Liu, Jiajun Huang, Zizhe Jian, Yiliu Li, Zizhong Chen. "TurboFFT: A High-Performance Fast Fourier Transform with Fault Tolerance on GPUs", *2023*, Under Submission.
6. **Shixun Wu**, Yujia Zhai, Jinyang Liu, Jiajun Huang, Zizhe Jian, Zizhong Chen. "High-performance ZGEMM on GPUs using Tensor Cores" *2023*, in progress.

HONORS AND AWARDS

- Third Prize at UCRPC, University of California, Riverside, 2023
- Distinguished Dean's Fellowship, University of California, Riverside, 2022
- Second Prize in PKU ACM in 2017, 2018.
- PKU May 4th Scholarship.