Siwei Chen

Beijing, China 2300017751@stu.pku.edu.cn 18810613228 pkuchensw

Personal Statement

I am currently a senior at Yuanpei College, Peking University, majoring in Data Science and Big Data Technology. My research focuses on efficient and self-improving large language models, spanning reinforcement-learning-based post-training, recursive self-improvement, and diffusion language models for parallel generation. I am particularly interested in co-designing learning algorithms, model architectures, and systems to improve the capability, efficiency, and deployability of LLMs.

Education

Yuanpei College, Peking University *Sept 2023 - Present*
B.S. in Data Science and Big Data Technology;

- GPA: **3.8/4.0**; Major Rank: 1/13; Overall Rank: 43/355
- **Basic Courses:** Machine Learning (99), Practice of Programming in C&C++ (98), Data Structure and Algorithm (A) (95), Probability theory and statistics in information science (95), Advanced Algebra (93)
- **Advanced Courses:** Programming in Artificial Intelligence (99), Principles, Challenges, and Industrial Applications of Large Language Models (99), Directed Research in AI System (95)
- **Teaching Assistant:** Probability theory and statistics in information science, Outdoor Exploration

Tsinghua University High School *Sept 2017 – Jun 2023*
High School Certificate;

- College entrance examination: 693/750

Research Experience

Efficient Reinforcement Learning for Large Language Models *May 2025 - Present*
Advisor: Prof. Bin Cui and Prof. Xupeng Miao, Peking University

- Investigated efficiency bottlenecks in LLM reinforcement learning, including the straggler effect caused by long-tail rollout lengths and the length explosion problem in LLM RL training.
- Proposed DARTS (Distribution-Aware Active Rollout Trajectory Shaping), a system-level optimization paradigm that actively reshapes the model's output distribution to reduce intra-prompt redundancy. Designed Dual-End Sampling, Adaptive Redundant Allocation, and a token-level streaming mechanism to accelerate RL training while preserving reasoning capability.
- Proposed QLPO, a resampling-based variant of GRPO that introduces implicit length control by favoring short-correct and long-incorrect trajectories without modifying the reward function.

Efficient Diffusion Language Models *Feb 2026 - Present*
Advisor: Prof. Fan Lai, University of Illinois Urbana-Champaign

- Identified prefix-dependence and early-rejection bottlenecks in proposal–verification decoding, and formulated verified progress per model evaluation as a key objective for efficient parallel generation.
- Developed BRISK-DLM, combining risk–reward proposal learning with a lightweight prefix-conditioned corrector distilled from verifier preferences, without additional backbone evaluations.
- Improved end-to-end throughput by 22.7% at 8B and 29.8% at 32B while preserving task quality.

Context Management Benchmark for Long-Horizon LLM Agents *Feb 2026 - May 2026*
Advisor: Prof. Bin Cui and Prof. Xupeng Miao, Peking University

- Worked on MemClaw-Bench, a capability-centered benchmark for evaluating context management in long-horizon tool-using agents under compressed-context settings.
- Designed diagnostic tasks and evaluation protocols covering memory capabilities such as constraint tracking, version updating, procedure transfer, source conflict resolution, task resumption, and staleness judgment.

Selected Publications

DARTS: Distribution-Aware Active Rollout Trajectory Shaping for Accelerating LLM Reinforcement Learning (ICML 2026) May 2025 - Jan 2026

Yujie Wang†, ***Siwei Chen†***, Longzan Luo†, Xinyi Liu, Xupeng Miao, Fangcheng Fu, Bin Cui (Co-first author)

arxiv.org/abs/2605.30859

QLPO: Quadrant-weighted sampling for Length-aware Policy Optimization (COLM 2026) Dec 2025 - March 2026

Siwei Chen, Siqi Chen, Xupeng Miao, Bin Cui (First Author)

arxiv.org/abs/2607.21793

From Position Risks to Block Survival: Faster Generation for Diffusion Language Models (Under Review) Feb 2026 - Sep 2026

Siwei Chen†, Yuxiang Wan†, Yifan Yu, Fan Lai (Co-first author)

arxiv.org/abs/2609.33390

COEVO: Co-Evolving Context and Parameters for Recursive Self-Improvement (Under Review) May 2026 - Sep 2026

Siwei Chen, Xinping Bao, Xinyu Cai, Yuan Cao, Wan Jiang, Shaohong Chen (First Author)

arxiv.org/abs/2609.33398

MemClaw-Bench: A Capability-Centered Benchmark for Context Management in Long-Horizon Agents (Under Review) Feb 2026 - May 2026

Siwei Chen, Xupeng Miao, Bin Cui (First Author)

Projects

Usagi: ICS course helper *Usagi project*

- Developed a Retrieval-Augmented Generation enhanced LLM question-answering framework. Backed by a comprehensive corpus of learning materials, the system supports RAG-based source tracing, predictive question generation, and learner cognitive state diagnosis. Over 100 students have used our system and given positive feedback.

MyTorch: Custom PyTorch for CNN training *MyTorch*

- Implemented a custom PyTorch based on CUDA programming. My implementation includes basic kernels, encapsulation between Python and C, and the construction of a simple computational graph. The project can ultimately train a CNN for image recognition.

Honors and Awards

- 2025-2026 National Scholarship, China's highest-level undergraduate scholarship
- 2025-2026 Peking University Outstanding Student Model, Academic Innovation Award
- 2026 SenseTime Scholarship, selected only 30 recipients nationwide
- 2024-2025 Outstanding Research Award, Peking University Third-Class Scholarship
- 2023-2024 Peking University Outstanding Student, Peking University Second-Class Scholarship
- 2023 Beijing Outstanding Student

Additional Information

Programming Languages and Tools: Python, C/C++, Rust, PyTorch, CUDA

Language: New TOEFL iBT: 5.5/6 (Reading 6.0, Listening 5.0, Writing 5.0, Speaking 5.0)

I enjoy all kinds of sports ranging from basketball, running to Ultimate Frisbee and mountain climbing. I am a member of the Shanying Club (Mountaineering Association of Peking University) and I won the second prize at the PKU Ultimate Frisbee Cup (2024,2025).