# Zelin Tan

tanzl@mail.ustc.edu.cn · Homepage · Google Scholar

## Education

**University of Science and Technology of China**, Artificial Intelligence, *Ph.D. Student* 2025.6 - Now
**University of Science and Technology of China**, Computer Technology, *M.Eng.* 2022.9 - 2025.6
**University of Science and Technology of China**, Computer Science *B.Eng.* 2018.9 - 2022.6

## Internship

**Shanghai AI Lab: Agent&LLM Post Training** Mentors: Chen Zhang and Shuyue Hu 2025.7 - Now

- **Enhancing LLM Skill-Use Capability**: Designed a data pipeline synthesizing long-horizon skill-dependent tasks and their multi-turn solution trajectories to improve the LLM's Skill-use ability.
- **Reasoning-Paradigm Routing for LLM Agents**: Proposed *Select-then-Solve*, a lightweight router that picks a paradigm(ReAct, Plan-Execute etc.) before solving, improving both accuracy and inference efficiency.
- **Rubric-based RL Post-Training for LLMs**: Via decoupled normalization of the advantage function, addressed reward-signal exhaustion in conventional RLVR and reward hacking in Rubric RLVR.
- **RL Scaling Laws for LLMs**: Studied scaling laws of GRPO-based RL post-training for LLM mathematical agents across model sizes and hyperparameter settings.
- **Agent-Human discussion forum**: We present AgentPanel, a multi-agent forum for human–AI collaboration in scientific exploration.
- **Survey on LLM-based Agents**: Surveyed frontier progress of agentic reinforcement learning across games, interaction, and tool use, producing a survey report.

**Meituan: LLM Inference Optimization** Mentors: Yineng Zhang and Ke Bao 2024.5 - 2024.7

- **Block-granularity automatic prefix caching**: Learning to manage GPU memory via a block-based prefix tree to enable automatic prefix caching.
- Operator fusion for Generative Recommendation(GR) models based on Tensor-RT

**Tencent: LLM Inference Optimization** Mentor: Hande Dong 2023.12 - 2024.5

- **Chunked prefill** with dynamic split-fuse on internal serving engine (based on vLLM v0.2.7)
- **SM-Aware Operator Parallelism Framework**: Leveraged intra-GPU compute-resource isolation to run prefill and decode attention kernels in parallel.
- **GPU-memory-aware dynamic chunk-size adjustment**, mitigating the recomputation overhead introduced by preemption under high load.

## Publications

- Scaling Behaviors of LLM Reinforcement Learning Post-Training: An Empirical Study in Mathematical Reasoning **(First author, ACL 2026 Main)**
- PAPO: Stabilizing Rubric Integration Training via Decoupled Advantage Normalization **(First author, EMNLP 2026 Under review)**
- SKT: Skill-Use Training at Scale via Verified Synthetic Data Generation**(First author, AAAI 2027 Under Review)**
- Select-then-Solve: Paradigm Routing as Inference-Time Optimization for LLM Agents **(EMNLP 2026 Under review)**
- The Landscape of Agentic Reinforcement Learning for LLMs: A Survey **(TMLR 2025)**
- SkillsBench: Benchmarking How Well Agent Skills Work Across Diverse Tasks **(NeurIPS 2026 Under review)**
- AgentAsk: Multi-Agent Systems Need to Ask **(ACL 2026 Main)**

## Open Source

- **RL Infra ROLL** – Contributor
- AgentPanel – Contributor
- **SkillsBench** – Contributor
- SM-Aware Operator Parallelism Framework – Designer