

Weixuan Yuan

Zürich, Switzerland (open to relocation) · weixuan.yuan@inf.ethz.ch · [Homepage](#)

Research interests: LLM reasoning, post-training (RLVR, synthetic data), AI for math and formal verification (Lean 4), agent harnesses, theoretical computer science (graph algorithms).

Seeking research internships: post-training, AI for math, agent harnesses (6 months).

Education

ETH Zürich — Ph.D. in Computer Science Advisor: Prof. Rasmus Kyng. Joined the group as a Research Assistant in October 2024.	May 2025–May 2029 (expected)
ETH Zürich — M.Sc. in Data Science GPA 5.92 / 6.00 (ranked 1st; Willi Studer Prize 2025).	September 2022–October 2024
Xi'an Jiaotong University — B.Sc. in Mathematics GPA 93.43 / 100 · Honors Youth Program.	September 2018–July 2022

Selected Research

AuXSpec: Learning to Autoformalize Program Specifications via Executable Postconditions

Co-first author · under review at ICLR 2027. Training LLMs to translate natural-language programming problems into formal specifications in Lean.

- Proposed *executable postconditions*: an executable counterpart of each declarative Lean specification that can be checked on any test case via `#guard`, turning specification correctness into a cheap automatic signal for data filtering and a core component of the RL reward.
- Co-designed the Lean specification format and co-built a multistage LLM pipeline for specification synthesis, verification, and repair, producing 3,037 formalized programming problems and ~19.5K SFT samples.
- Reviewed 300 benchmark problems and audited 60 model-generated specifications.
- With SFT + RL, a 27B model improved SpecGen pass@4 from 0.7% to 60.3% on AuXSpec-Test and from 21.9% to 84.7% on Verina.

Deterministic Gomory–Hu Trees — FOCS 2025

Co-authored the first deterministic almost-linear-time algorithm for Gomory–Hu trees (all-pairs minimum cuts), running in $m^{1+o(1)}$ time on graphs with m edges. Presented the paper with Wuwei Yuan at FOCS 2025. [\[slides\]](#)

Publications

Deterministic Almost-Linear-Time Gomory-Hu Trees

Amir Abboud, Rasmus Kyng, Jason Li, Debmalya Panigrahi, Maximilian Probst Gutenberg, Thatchaphol Saranurak, **Weixuan Yuan**, Wuwei Yuan.

IEEE Symposium on Foundations of Computer Science (FOCS 2025). Authors listed alphabetically. [\[paper\]](#)

Statistics (undergraduate work, before 2023; [†] equal contribution)

Xuehu Zhu[†], **Weixuan Yuan**[†], Zongben Xu, Wenlin Dai. Optimal subsampling for principal component analysis.

Science China Mathematics 68, 2993–3016, 2025. [\[paper\]](#)

Xinru Jia[†], **Weixuan Yuan**[†], Xingqiu Zhao, Xuehu Zhu. Optimal subsampling for estimation of dimension reduction directions. *Scandinavian Journal of Statistics* 53(1), 575–611, 2026. [\[paper\]](#)

Technical Skills

Programming: Python, Lean 4.

Post-training: verl (GRPO and other RLVR methods).

Agentic tools: Codex, Claude Code (daily use; custom skills and AGENTS.md-based workflows).