

Hanpu Shen

hanpushen@gmail.com • Irvine, CA

Reinforcement Learning · Embodied AI · Deep Generative Models · World Models

EDUCATION

University of California, Irvine

09/2022 – 06/2027 (expected)

Ph.D. in Statistics · GPA 3.94 / 4.0 · Advisors: Prof. Weining Shen, Prof. Roy Fox

Irvine, CA

Southern University of Science and Technology

09/2018 – 06/2022

B.S. in Statistics & Data Science · GPA 3.83 / 4.0 (Major 3.89, Top 5%)

Shenzhen, China

SELECTED PUBLICATIONS

*All papers first-authored. * denotes equal contribution.*

1. REG: In-Sample RL via Regularizing the Evaluation Gap

Hanpu Shen, Weining Shen, Roy Fox. [ICML 2026](#).

2. Differentially Private Non-Convex Learning: From Generalized Linear Models to Multi-Layer Neural Networks

Hanpu Shen*, Cheng-Long Wang*, Zihang Xiang, Yiming Ying, Di Wang. [AAMAS 2026](#).

RESEARCH

REG: In-Sample RL via Regularizing the Evaluation Gap

2024 – 2026

Advisor: Prof. Roy Fox, Dept. of Computer Science, UCI

[ICML 2026](#)

- Avoided querying out-of-distribution actions by bounding the off-policy evaluation gap; used Fenchel duality to turn the resulting optimization into an equivalent RL algorithm. Replaced IQL's expectile regression with a simpler, principled critic loss backed by a theoretical guarantee on the gap.
- Proposed an orthogonal policy gradient that adds on-policy information to the in-sample gradient, encouraging mode-seeking policies; matched or outperformed a diffusion baseline on D4RL MuJoCo and AntMaze.
- Used the learned critic to select high-quality policies, reaching under 5% regret relative to Top@10 trajectories.

Differentially Private Non-Convex Learning: From GLMs to Multi-Layer Neural Networks 2022 – 2023

Advisor: Prof. Di Wang, Dept. of Computer Science, KAUST

[AAMAS 2026](#)

- Constructed a convex surrogate loss upper-bounding generalization error for monotone, L-Lipschitz link functions over a constrained parameter space.
- Developed a privacy-preserving ERM algorithm with a dimension-independent excess-risk bound $O(\log^2(n)/\sqrt{n})$ for constrained non-convex DP-GLMs via objective perturbation, and used Neural Tangent Kernel analysis to derive the first population-risk bound for DP-SGD on arbitrary-depth over-parameterized ReLU networks.

EXPERIENCE

Amazon, Global Supply Chain Optimization Technologies (SCOT)

06/2025 – 09/2025

Applied Scientist Intern

Bellevue, WA

- **Built the External-Products-Aware Choice Model (EPACM)**, a hierarchical probabilistic model coupling view-propensity and utility networks to fuse Amazon behavioral logs with external market data and predict purchase probabilities for both existing and previously unseen products, **raising total sales by more than 4%** over prior models during Prime Day 2025.
- Derived a differentiable Gauss–Laguerre approximation of the intractable choice integral, curbing gradient explosion in low-utility regimes and improving long-tail convergence relative to Taylor expansions.

HONORS & AWARDS

First-Class Merit Student Scholarship, SUSTech (2019, 2020, 2021)

TECHNICAL SKILLS

Programming: Python, R, MATLAB, Java

ML & Tools: JAX, PyTorch, Claude Code, Codex, agentic coding workflows