

Yizhou Zhang

Pasadena, CA, 91125 • (626) 316-2381 • yzhang8@caltech.edu

[Personal Webpage](#) • [LinkedIn](#) • [Google Scholar](#)

Research Interests

I am interested in solving **decision-making** problems in the **multi-agent** world with modern **machine learning** techniques, both theoretically and practically. In particular, my research focuses on the following (and related) fields:

- ◇ Strategic learning and decision making through **game theory**;
- ◇ Robust decision making against uncertainty through **RL** and **control**;
- ◇ Leveraging **generative AI** towards data-efficient decision making;
- ◇ Social aspects of decision making: **economics, privacy, and safety**.

Education

California Institute of Technology

Sep. 2023 – Present

PhD in Computing and Mathematical Sciences

Department of Computing and Mathematical Sciences

Advisors: Prof. Adam Wierman and Prof. Eric Mazumdar.

Tsinghua University

Aug. 2019 – Jun. 2023

B.Eng in Computer Science (Yao Class)

Institute for Interdisciplinary Information Sciences

GPA: 3.85/4.0

Experiences

California Institute of Technology

Feb. 2022 – Aug. 2022

Department of Computing and Mathematical Sciences

Visiting student, attended the Visiting Undergraduate Research Program (VURP).

Worked with Prof. Adam Wierman.

Shanghai Qizhi Institute

Jun. 2021 – Aug. 2021

Research Intern on Platform Economics: Design, Optimization and Regulation.

Advisor: Prof. Zhixuan Fang

Publications and Manuscripts

- [1] **Provably Convergent Actor-Critic in Risk-averse MARL.**

Yizhou Zhang and Eric Mazumdar.

In submission.

- [2] **Training Generalizable Collaborative Agents via Strategic Risk Aversion.**

Chengrui Qu, **Yizhou Zhang**, Nicola Lanzetti and Eric Mazumdar.

In submission.

- [3] **KL-regularization Itself is Differentially Private in Bandits and RLHF** ([link](#)).
Yizhou Zhang, Kishan Panaganti, Laixi Shi, Juba Ziani and Adam Wierman.
In submission.
- [4] **Convergent Q-Learning for Infinite-Horizon General-Sum Markov Games through Behavioral Economics** ([link](#)).
Yizhou Zhang and Eric Mazumdar.
Proceedings of the 64th IEEE Conference on Decision and Control (CDC), 2025.
- [5] **Learning to Steer Learners in Games** ([link](#)).
Yizhou Zhang, Yian Ma and Eric Mazumdar.
Proceedings of the 42nd International Conference on Machine Learning (ICML), 2025.
- [6] **Global Convergence of Localized Policy Iteration in Networked Multi-Agent Reinforcement Learning** ([link](#)).
Yizhou Zhang*, Guannan Qu*, Pan Xu*, Yiheng Lin, Zaiwei Chen and Adam Wierman.
Proceedings of the ACM on Measurement and Analysis of Computing Systems, 2023.

Selected Research Projects

Multi-Agent RL through Behavioral Economics

Nov. 2024 – Jan. 2026

with Eric Mazumdar, Caltech

also with Chengrui Qu and Nicolas Lanzetti, Caltech

Incorporated risk-aversion and bounded rationality from human decision-making to Q-learning in infinite-horizon general-sum Markov games. We proved that under the monotonicity assumption of the underlying game, the solution concept of risk-averse quantal-response equilibrium (RQE) is unique, and can be computed through provably convergent Q-learning or Actor-Critic algorithms in discounted infinite-horizon Markov games. Based on this, we developed a scalable risk-averse actor-critic for general-sum MARL that empirically provides better convergence.

Inherent differential privacy of KL-regularization

Nov. 2024 – Present

with Kishan Panaganti, Tencent; Laixi Shi, Johns Hopkins;

Juba Ziani, Georgia Tech and Adam Wierman, Caltech.

While guaranteeing DP generally requires explicitly injecting noise either to the algorithm itself or to its outputs, the intrinsic randomness of existing algorithms presents an opportunity to achieve DP “for free”. We explored regularization in achieving DP across three different offline decision-making problems: multi-armed bandits, linear contextual bandits, and reinforcement learning from human feedback (RLHF). We show that adding KL-regularization to the learning objective (a common approach in optimization algorithms) makes the action sampled from the resulting stochastic policy itself differentially private.

Learning to Steer Learners in Games

Dec. 2023 – Jan. 2025

with Eric Mazumdar, Caltech and Yian Ma, UC San Diego

Studied whether a no-regret learner can be exploited in a repeated general-sum bimatrix game such that its opponent can achieve Stackelberg value even without the knowledge of the payoff matrix of the learner. We first showed that this is impossible in general if the optimizer only knows that the learner is using a no-regret algorithm. We then reduced the problem for the optimizer to that of recovering the learner’s payoff structure. We demonstrated the effectiveness of this approach by analyzing two examples: one where the learner uses an ascent algorithm, and another where the learner uses stochastic mirror ascent.

Networked MARL with Local Information

Feb. 2022 – Sept. 2022

Supervisor: Adam Wierman, Caltech

Studied a multi-agent reinforcement learning problem on a network of agents where each agent

cooperatively maximizes the average of their entropy-regularized long-term rewards. We showed that even if the policy of each agent is restricted to its κ -hop neighborhood, there is a bounded performance gap between that policy and the globally optimal policy. We proposed a novel policy iteration algorithm for each agent which only uses the information inside its κ -hop neighborhood and proved its convergence to a near-globally-optimal policy.

Talks

Convergent Q-learning in Discounted Markov Games.

64th IEEE Conference on Decision and Control (CDC), Rio de Janeiro, Brazil.

Dec. 2025

Convergent Q-learning in Discounted Markov Games.

64th IEEE CDC Workshop: Game Theory Meets MPC: Advances in Multi-Agent Control, Rio de Janeiro, Brazil.

Dec. 2025

Learning to Steer Learners in Games.

RSRG/FALCON weekly seminar, Caltech, Pasadena, CA.

Jul. 2025

Learning to Steer Learners in Games.

AIX: Academia-Industry X Workshop, Caltech, Pasadena, CA.

Jul. 2025

Service

Reviewer: *Transactions on Machine Learning Research (TMLR)*; *International Conference on Machine Learning (ICML)* (2026); *IEEE Conference on Decision and Control (CDC)* (2025); *International Conference on Artificial Intelligence and Statistics (AISTATS)* (2026); *ACM SIGMETRICS* (2026); *Learning for Dynamics and Control (L4DC)* (2026); and *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*.

Organizer: 2025-2026 RSRG/FALCON weekly seminar, Caltech.

Skills

Programming: Python, C/C++, MATLAB, \LaTeX

Languages: Chinese (Native); English (TOEFL iBT 111; GRE 330).