# Jingqi Fan

[jingqifanfanfan@gmail.com](mailto:jingqifanfanfan@gmail.com) | [jingqi-fan.github.io](https://jingqi-fan.github.io) | [linkedin.com/in/jingqi-fan](https://linkedin.com/in/jingqi-fan)

## EDUCATION

- **Northeastern University** *Sep. 2022 – Jul. 2026*
  *B.Eng. in Software Engineering* Shenyang, China
  ◦ GPA: 3.67/4.00     TOEFL: 96/120
  ◦ Selected courses: Algorithm Design and Analysis (4.0/4.0), Numerical Analysis (4.0/4.0), Information Retrieval Theory (4.0/4.0), Data Structures and Algorithms (4.0/4.0), Computer Networks (4.0/4.0), Principles of Computer Organization (3.7/4.0), Operating System (4.0/4.0), Python Data Analysis (4.0/4.0), C++ Programming (4.0/4.0).

## RESEARCH INTERESTS

Reinforcement Learning Theory, Multi-Armed Bandits, Online Learning, Operations Management.

## PUBLICATIONS                                                            † = CORRESPONDING AUTHOR

[1] **Jingqi Fan**, Zilong Wang, Shuai Li†, and Linghe Kong. **Multi-player Multi-armed Bandits with Delayed Feedback**. In The 34th International Joint Conference on Artificial Intelligence (**IJCAI 2025**). [PDF] [OpenReview] .

[2] **Jingqi Fan**, Canzhe Zhao, Shuai Li, and Siwei Wang†. **Decentralized Asynchronous Multi-player Bandits**. Under review. [Arxiv].

## RESEARCH EXPERIENCE

- **Differentiable Simulation of Large-Scale Queueing Networks for RL** *Jul. 2025 – Present*
  *Research Assistant, NYU Stern School of Business* New York, USA
  Advisor: Prof. Zhengyuan Zhou
  ◦ Built a fully tensorized, differentiable discrete-event simulation of large-scale queueing networks with TorchRL, removing the bottleneck of NumPy–Torch conversions in Stable-Baselines.
  ◦ Implemented and evaluated both classical scheduling policies (c$\mu$-rule, MaxWeight, MaxPressure, Fluid) and RL algorithms (PPO, Conservative PPO, PPO with behavior cloning).

- **Chapter 3 of RL Monograph: Online learning and Exploration-Exploitation Trade-off** *Jan. 2025 – Mar. 2025*
  *Research Assistant, Shanghai Jiao Tong University* Shanghai, China
  Advisor: Prof. Weinan Zhang
  ◦ Contributed to a nationally funded reinforcement learning monograph under the 101 Reform Project; authored the chapter "Online Learning and Exploration–Exploitation Trade-off".
  ◦ Provided theoretical analyses and code implementations of multiple online learning algorithms ($\epsilon$-greedy, ETC, UCB, elimination, Thompson Sampling). Open-sourced the [code] and the [PDF].

- **Asynchronous Feedback in Competitive Multi-Player Multi-Armed Bandits** *Aug. 2023 – Sep. 2024*
  *Research Assistant, Shanghai Jiao Tong University* Shanghai, China
  Advisor: Prof. Shuai Li
  ◦ Designed algorithms for competitive multi-player multi-armed bandits with delayed feedback and collisions. Established regret upper and lower bounds showing near-optimality.
  ◦ Validated the algorithms through numerical simulations and real-world experiments, demonstrating significant improvements over standard baselines. The related paper has been published in **IJCAI 2025**.

## INTERNSHIP

- **Decentralized Asynchronous Multi-player Multi-armed Bandits** *Dec. 2024 – Apr. 2025*
  *Research Intern, Theory Center, Microsoft Research Asia* Beijing, China
  Mentor: Dr. Siwei Wang
  ◦ Proposed a decentralized algorithm for competitive multi-player bandits when players join and leave at arbitrary times, while also being robust to heterogeneous reward structures.
  ◦ Validated the algorithm through large-scale simulations up to 100 arms. Simulated IoT network environments to further evaluate its effectiveness. The related paper is under review [Arxiv].

- **Collaborative LLM Agents for Personalized Stock Alerts and Recommendations** *Mar. 2025 – Jul. 2025*
  *Research Intern, Financial Algorithm Industry Research Institute, Shenyang Linglong Technology Co., Ltd.* Shenyang, China
  Mentor: Dr. Jun Na
  ◦ Designed a collaborative multi-agent LLM system that generates dynamic stock reports (morning, midday, and closing) by retrieving data, performing retrieval-augmented generation, and integrating stock analysis tools.
  ◦ Implemented PPO-based algorithms to optimize the timing of report delivery, ensuring relevance and timeliness.
  ◦ Developed user modeling methods to infer interests and financial literacy from historical dialogues, enabling personalized recommendations and adaptive report styles.