EC 504, Engineering Center
115 Library Drive
Rochester, MI, 48309
jiehu@oakland.edu
jiehu-research.github.io
jiehu2017
Google Scholar

Jie Hu

Curriculum Vitae

Research Interests

My research centers on advancing machine learning theory by developing efficient and scalable distributed optimization methods and improving sampling techniques in data mining. I investigate how enhanced non-Markovian sampling can theoretically accelerate the training process in distributed machine learning tasks. I seek advancements in MCMC for wider applications, such as policy evaluation in reinforcement learning, distributed learning across IoT, mobile, vehicular, and robotic networks, control of virus spreading in epidemics, representation of graph structures for graph neural networks, and fast Bayesian inference.

Experience

2025–Current **Assistant Professor in Computer Science and Engineering**, *Oakland University*, Rochester, MI
- Research: Exploring efficient data sampling algorithms in very large state space with applications in Generative AI, molecular dynamics, and statistical physics.
- Teaching: Introduction to Computer Networks (Fall 2025)

Education

2019–2025 **Ph.D. – Electrical Engineering**, *North Carolina State University*, Raleigh, NC
- **Proposed Dissertation Title**: *Statistical Efficiency in a Non-Markovian Era: Improving Machine Learning and Distributed Optimization*
- Advisor: Prof. Do Young Eun

2017–2019 **M.S. – Electrical Engineering**, *Northwestern University*, Evanston, IL
- **M.S. Thesis**: *Performance Analysis of DSRC Vehicle-to-Vehicle Safety Communication*
- Advisor: Prof. Randall Berry

2012—2016 **B.E. – Telecommunication Engineering**, *Wuhan University of Technology*, Wuhan, China

Honors and Awards

2023 ○ **Outstanding Paper Award**, International Conference on Machine Learning (ICML). Link: https://icml.cc/Conferences/2023/Awards
One of 6 *winners out of* 6538 *submitted papers.*

2022 ○ Scholar Award, Conference on Neural Information Processing Systems (NeurIPS).

2013 ○ First-Class Scholarship of Wuhan University of Technology (Top 3%).

Publications

* stands for the co-first author

1. **Jie Hu**, Yi-Ting Ma, Do Young Eun, "Beyond Self-Repellent Kernels: History-Driven Target Towards Efficient Nonlinear MCMC on General Graphs," ICML (**Oral, Top 1% out of 12107 submitted papers**),

2025.

2. **Jie Hu**, Yi-Ting Ma, Do Young Eun, "Does Worst-Performing Agent Lead the Pack? Analyzing Agent Dynamics in Unified Distributed SGD," NeurIPS, 2024.
3. **Jie Hu**, Vishwaraj Doshi, Do Young Eun, "Central Limit Theorem for Two-Timescale Stochastic Approximation with Markovian Noise: Theory and Applications," AISTATS, 2024.
4. **Jie Hu***, Vishwaraj Doshi*, Do Young Eun, "Accelerating Distributed Stochastic Optimization via Self-Repellent Random Walks," ICLR (**Oral, Top 1.18% out of 7304 submitted papers**), 2024.
5. Vishwaraj Doshi, **Jie Hu**, Do Young Eun, "Self-Repellent Random Walks on General Graphs–Achieving Minimal Sampling Variance via Nonlinear Markov Chains," ICML (**Outstanding Paper Award**), 2023.
6. Vishwaraj Doshi*, **Jie Hu***, Do Young Eun, "Bi-SIS Epidemics on Graphs-Quantitative Analysis of Coexistence Equilibria," IEEE CDC, 2022.
7. **Jie Hu***, Vishwaraj Doshi*, Do Young Eun, "Efficiency ordering of stochastic gradient descent," NeurIPS, 2022.
8. **Jie Hu**, Vishwaraj Doshi, Do Young Eun, "Minimizing File Transfer Time in Opportunistic Spectrum Access Model," IEEE Transactions on Mobile Computing, 2022.
9. **Jie Hu**, Vishwaraj Doshi, Do Young Eun, "Opportunistic spectrum access: Does maximizing throughput minimize file transfer time?," International Symposium on Modeling and Optimization in Mobile, Ad hoc, and Wireless Networks (WiOpt), 2021.
10. Liu Cao, Hao Yin, **Jie Hu**, Lyutianyang Zhang, "Performance analysis and improvement on DSRC application for V2V communication," IEEE Vehicular Technology Conference (VTC), 2020.

Research Experience

Adaptive Sampling and Its Application to Distributed Optimization

- tackled a challenging set of open problems, Markov Chain Monte Carlo (MCMC) with self-repellent random walk (SRRW) and history-driven target (HDT) scheme. It goes beyond traditional non-backtracking approaches and paves the way for new research directions in MCMC sampling. By leveraging temporal information, we theoretically prove that SRRW achieves minimal sampling variance and accelerate distributed learning process.
- Experiments show that graph information estimation and model training error decreased nearly threefold over standard algorithms.
- Earned recognition at ICML 2023 (***Outstanding Paper Award***), ICLR 2024 (***Oral***), and ICML 2025 (***Oral***).

Central Limit Theorem on Two-Timescale Stochastic Approximation

- Established first-of-its-kind central limit theorem to demonstrate the coupled dynamics of two iterations influenced by the underlying controlled Markov chain, which evolves according to both iterates.
- Theoretically and numerically studied its application to (i) accelerated distributed mini-max / bi-level optimization using efficient data sampling strategies; (ii) statistical properties of gradient-based TD learning algorithms (GTD2, TDC).
- Work appeared in AISTATS 2024.

Impact of Data Sampling in Distributed Learning

- Demonstrated the long-ignored effect of data sampling strategy in distributed optimization for a single agent where datasets are spread over a graph, e.g., environmental features in wireless sensor networks, robot exploration of unknown environment. Theoretically proved accelerated training processes through efficient data sampling methods, bridging the gap between MCMC and optimization.
- Extension to multi-agent systems and emphasized the equal contribution of each agent to the overall system performance, instead of "worst-performing agent" in the current multi-agent literature. Identified that improving sampling strategies of a fraction of agents will accelerate the training process of the whole multi-agent system.
- Verified using Python that a more efficient random walk strategy improves convergence speed over a baseline in logistic regression. In distributed image classification with edge computing, this reduced ResNet-18 training time by 20% on CIFAR-10 dataset.
- Works published in NeurIPS 2022 and NeurIPS 2024.

Quantitative Analysis of Epidemics on Graphs

- Developed theoretical analysis and simulations to quantify and visualize the Bi-SIS epidemic on complex networks, demonstrating the practical implications of how two viruses compete over a general graph and how tight the quantitative bound of the infection ratio is to predict virus spreads.
- Work presented at IEEE CDC 2022.

File Transfer Time in Dynamic Spectrum Access

- Led the research of opportunistic spectrum access models in cognitive radio and designed algorithms to minimize file transfer time where every channel's availability follows either Bernoulli distribution or Markov process. Saved on average over 10% transmission time in the simulation using Python and pyOpt to solve mixed integer programming, compared to the baseline throughput-optimal algorithm.
- Work published in WiOpt 2021 and IEEE TMC 2022.

Teaching & Mentoring

Teaching Assistant

Spring 2024 **ECE 570 Computer Networks**, *North Carolina State University*
Fall 2020
- Support students with programming tasks, exam and project assessments.
- Host office hours to address inquiries about homework, networking architecture, protocols taught in class.

Spring 2022 **ECE 220 Analytical Foundations of ECE**, *North Carolina State University*
Fall 2021
Spring 2021
- Conduct laboratory sessions for implementation of linear algebra, differential equations, and Fourier analysis using MATLAB.
- Grade homework assignments, and provide additional support to students struggling with complex mathematical concepts.

Grader

Winter 2019 **EECS 422 Random Processes in Communications and Control 1**, *Northwestern University*

Fall 2018 **EECS 308 Communications Systems**, *Northwestern University*

Mentoring Experience

Aug 2023–Current Yi-Ting Ma, Ph.D. student at North Carolina State University

Presentations

Invited Talks
- **"Statistical Efficiency in Data Sampling: Enhancing Machine Learning and Distributed Optimization Beyond the Markovian Regime,"** CCSL seminar series talk, North Carolina State University, October 2024.

Contributed Presentations

- **"Beyond Self-Repellent Kernels: History-Driven Target Towards Efficient Nonlinear MCMC on General Graphs,"** ICML (Oral), Vancouver, Canada, July 2025.
- **"Bi-SIS Epidemics on Graphs-Quantitative Analysis of Coexistence Equilibria,"** IEEE CDC, Cancun, Mexico, December 2022.

Service

Reviewer

- IEEE Transactions on Network Science and Engineering
- Transactions on Machine Learning Research
- NeurIPS 2024 (Top Reviewer), 2025
- ICLR 2025 - 2026
- AISTATS 2024 - 2026
- ICML 2025
- AAAI 2026