

BOYA HOU

123 Coordinated Science Laboratory, 1308 W Main St, Urbana, IL

boyahou2@illinois.edu <https://boyahou.github.io/>

ACADEMIC APPOINTMENT

University of Illinois Urbana-Champaign

Since October 2024

Postdoctoral Research Associate at CSL and Carl R. Woese Institute for Genomic Biology

Advisors: Maxim Raginsky, Olgica Milenkovic

EDUCATION

University of Illinois Urbana-Champaign

2024

PhD in Electrical and Computer Engineering

Advisor: Subhonmesh Bose

University of Illinois Urbana-Champaign

2019

Master of Engineering in Electrical and Computer Engineering

Zhejiang University

2019

Bachelor of Engineering in Electrical Engineering

RESEARCH INTERESTS

My research focuses on developing machine learning methods for complex systems across science and engineering, spanning learning, probabilistic sampling, control, and stability analysis. Drawing on machine learning, operator theory, control and dynamical systems, I take a unified approach to these problems through learned operator representations that govern the evolution of functions or distributions. These operators are linear even when the underlying system is nonlinear. I construct them from trajectory data with finite-sample guarantees and exploit their structure to design efficient algorithms for each task, with provable guarantees that black-box models do not offer. Active research directions include¹:

- Operator learning with finite-sample guarantees: learning operator representations of nonlinear dynamical systems from trajectory data, with compression, multi-agent, and online algorithms that scale to large, distributed and streaming data, applied to data-driven decision-making [1, 2, 6, 7, 8, 10, 14].
- Operator-guided sampling and steering of diffusion processes: learning-based methods that combine operator learning with optimal stochastic control to sample from target path measures, applied to transition path sampling in molecular dynamics [12, 15], with natural extensions to problems that share mathematical foundations, such as generative modeling with diffusion models and Schrödinger bridges.
- Applications in scientific AI and engineering: molecular and cellular dynamics [11,12,15,16], analysis of drug intervention [11,16], and power grid operations [3,4,5,9].
- Dynamical systems for AI, e.g., data attribution and unlearning for LLM [13].

¹ Bracketed numbers refer to the Full List of Publications

SELECTED PUBLICATIONS

For convenience, each paper title below is a clickable link to the full text. * stands for co-first authorships. Bracketed numbers refer to paper in this section.

- [1] **B. Hou**, S. Sanjari, N. Dahlin, S. Bose, U. Vaidya. [*Sparse Learning of Dynamical System in Reproducing Kernel Hilbert Space: An Operator-Theoretic Approach*](#). In Proceedings of the Fortieth International Conference on Machine Learning (ICML), 2023.

Summary: operator representations describe nonlinear dynamics through linear operators that act on functions or distributions of states; however, these operators are infinite-dimensional. I establish the first finite-sample convergence guarantees from trajectory data for learning these operators, in both discrete and continuous time. I further introduce a compression scheme that makes the algorithm scalable and regularizes the learning problem.

- [2] **B. Hou**, S. Sanjari, N. Dahlin, A. Koppel, S. Bose. [*Nonparametric Sparse Online Learning of the Koopman Operator*](#). To appear at SIAM Journal on Control and Optimization (SICON), 2026.

Summary: the guarantees in [1] apply to data collected in advance, whereas in many applications, data arrives sequentially. I designed stochastic operator gradient descent, which is the first online algorithm for learning linear operators with provable guarantees. I establish both asymptotic and finite-time convergence and incorporate compression to control model complexity. I further applied this operator-learning framework to reinforcement learning.

- [3] **B. Hou***, S. Wang*, S. Ambekar, M. Raginsky, O. Milenkovic. [*Transition Path Sampling Using Koopman Operators and Exit-Time Optimal Control*](#). In submission, 2026. Short version accepted at NeurIPS 2026 workshop.

Summary: A core problem in machine learning is to steer a diffusion process so that it generates samples from a target distribution. I study a particularly challenging case, transition path sampling (TPS), which samples rare transitions between metastable states. These transitions take far longer than any affordable simulation. Most existing ML methods fix the duration of a transition in advance and train a neural-network sampler by repeated simulation, which can be costly. Building on [1,2], I instead first learn an operator representation of the dynamics, whose spectrum identifies conformations of interest. I then formulate TPS as an optimal control problem up to an exit time, incorporating this spectral information into a cost that penalizes the time a path lingers before it transitions and derive the optimal controller in closed form. The controller is then computed from a single convex optimization problem without additional simulation in the loop. I have validated the method on alanine dipeptide, and I am now extending it to fast-folding proteins.

- [4] **B. Hou**, M. Raginsky, A. Pandey, O. Milenkovic. [*Spatially-Coupled Network RNA Velocities: A Control-Theoretic Perspective*](#). In submission, 2026.

Summary: A cell's fate is the outcome of a networked dynamical process, yet RNA velocity models typically treat genes and cells in isolation. In collaboration with AbbVie, I designed a spatially coupled network RNA velocity model that integrates intracellular gene-regulatory structure with intercellular interactions to infer cell dynamics. The control-theoretic formulation enables analysis of steady states, stability, and consensus-like behavior across cell populations, and provides a principled basis for studying how targeted perturbations propagate through regulatory networks and spatial cell neighborhoods.

FULL LIST OF PUBLICATIONS AND PREPRINTS

Each paper title below is a link to the full paper. * stands for co-first authorships.

Published Journal Articles

- [1] **B. Hou**, S. Sanjari, N. Dahlin, A. Koppel, S. Bose. [Nonparametric Sparse Online Learning of the Koopman Operator](#). To appear at SIAM Journal on Control and Optimization (SICON), 2026.
- [2] J. Lee, B. Hamzi, **B. Hou**, H. Owhadi, G. Santin, U. Vaidya. [Kernel Methods for the Approximation of the Eigenfunctions of the Koopman Operator](#). Physica D: Nonlinear Phenomena, 2025.
- [3] A. Reddy Ramapuram Matavalam, **B. Hou**, H. Choi, S. Bose, U. Vaidya. [Data-Driven Transient Stability Analysis Using the Koopman Operator](#). International Journal of Electrical Power and Energy Systems, 2024.
- [4] **B. Hou**, A. Reddy Ramapuram Matavalam, S. Bose, U. Vaidya. [Propagating Uncertainty Through System Dynamics in Reproducing Kernel Hilbert Space](#). Physica D: Nonlinear Phenomena. Special issue: Topics at the Interface of Machine Learning and Dynamical Systems, 2024.
- [5] **B. Hou**, S. Bose, L. Marla and K. Haran. [Impact of Aviation Electrification on Airports: Flight Scheduling and Charging](#). IEEE Transactions on Intelligent Transportation Systems, 2023.

Published Conference Proceedings

- [6] **B. Hou**, S. Sanjari, N. Dahlin, S. Bose, U. Vaidya. [Sparse Learning of Dynamical System in Reproducing Kernel Hilbert Space: An Operator-Theoretic Approach](#). In Proceedings of the Fortieth International Conference on Machine Learning (ICML), 2023.
- [7] **B. Hou**, S. Sanjari, N. Dahlin, S. Bose. [Compressed Decentralized Learning of Conditional Mean Embedding Operators in Reproducing Kernel Hilbert Space](#). In Proceedings of the 37th Association for the Advancement of Artificial Intelligence (AAAI) Conference on Artificial Intelligence, 2023.
- [8] **B. Hou**, S. Bose and U. Vaidya. [Sparse Learning of Kernel Transfer Operators](#). In Proceedings of Asilomar Conference on Signals, Systems, and Computers, 2021.
- [9] **B. Hou**, S. Bose, and K. Haran. [Powering Electric Aircraft at O'Hare Airport: A Case Study](#). In Proceedings of IEEE Power and Energy Society General Meeting, 2020.

Preprints

- [10] **B. Hou**, M. Raginsky. [A Behavioral Framework for Data-Driven Modeling of Nonlinear Systems in Vector-Valued Reproducing Kernel Hilbert Spaces](#). Under revision at IEEE Transactions on Automatic Control (TAC), 2026.
- [11] **B. Hou**, M. Raginsky, A. Pandey, O. Milenkovic. [Spatially-Coupled Network RNA Velocities: A Control-Theoretic Perspective](#). Under review at IEEE Transactions on Control of Network Systems (TCNS), 2026.
- [12] **B. Hou***, S. Wang*, S. Ambekar, M. Raginsky, O. Milenkovic. [Transition Path Sampling Using](#)

Koopman Operators and Exit-Time Optimal Control. In submission, 2026. Short version accepted at a NeurIPS 2026 workshop.

- [13] P. Niu*, E. Ma*, **B. Hou***, H. Zhou, D. Zhou, H. Zhang, R. Etesami, O. Milenkovic. GUARD: Guided Unlearning and Retention via Data Attribution for Large Language Models. In submission, 2026.

Monograph (in preparation)

- [14] **B. Hou**, S. Bose, N. Dahlin, U. Vaidya. Nonparametric Method for Learning Dynamical Systems. Sample chapters are available on my website, full monograph in preparation.

Workshops & Non-Archival Venues

- [15] **B. Hou***, S. Wang*, S. Ambekar, M. Raginsky, O. Milenkovic. *Transition Path Sampling Using Koopman Operators and Exit-Time Optimal Control*. Accepted at NeurIPS workshop, 2026.
- [16] **B. Hou**, M. Raginsky, A. Pandey, O. Milenkovic. *Multi Network RNA Velocity*. National Institute for Theory and Mathematics in Biology (NITMB) MathBio Convergence Conference, 2025.
- [17] **B. Hou**, A. Reddy Ramapuram Matavalam, S. Bose, U. Vaidya. *Propagating Uncertainty Through System Dynamics in Reproducing Kernel Hilbert Space*. Late breaking poster paper at American Control Conference (ACC), 2023.

AWARDS/HONORS

- The NSF-Simons National Institute for Theory and Mathematics in Biology (NITMB) Early Career Travel Award, 2025.
- Best Poster Award, Power Systems Engineering Research Center (PSERC), Industry Advisory Board (IAB) Meeting, 2024.
- Rising Stars in EECS, Georgia Tech, 2023
- Mavis Future Faculty Fellows (MF3), UIUC, 2023-2024.
- M.A. Pai Scholarship, UIUC, 2023.
- Association for the Advancement of Artificial Intelligence (AAAI) Student Scholarship, 2023.
- Second-place winner, United States Association for Energy Economics (USAEE) Case Competition, 2019.
- UCLA Cross-disciplinary Scholars in Science and Technology (CSST) Scholarship, 2017.

INVITED TALKS AND PRESENTATIONS

AI&ML for Complex Systems

- January 2027, Joint Mathematics Meetings (JMM), AMS Special Session on Recent Advances in Scientific Machine Learning: Theory, Algorithms, and Applications. (Talk)
- March 2026, Autonomy Talks Series. (Talk). Video can be found [here](#).
- Invited speaker at the Operator Learning in Control workshop at the 64th IEEE Conference on

Decision and Control, Rio de Janeiro, Brazil. (Unable to attend due to visa concerns)

- April 2025, UIUC, the 11th Midwest Workshop on Control and Game Theory. (Poster)
- December 2024, Power Systems Engineering Research Center (PSERC), Industry Advisory Board (IAB) Meeting. (Poster)
- October 2024, Cornell ORIE, Young Researcher's Workshop. (Poster)
- September 2024, Sandia National Laboratories, Sandia Machine Learning and Deep Learning (MLDL) Workshop. (Talk)
- July 2024, Fields Institute, the 4th Symposium on Machine Learning and Dynamical Systems. (Poster)
- May 2024, Alan Turing Institute, Machine Learning and Dynamical Systems Seminar. (Talk)
- April 2024, Northwestern, the 10th Midwest Workshop on Control and Game Theory. (Poster)
- April 2024, UIUC, group meeting of Prof. R. Srikant and Prof. Carolyn Beck. (Talk)
- March 2024, MIT, group meeting of Prof. Navid Azizan. (Talk)
- Feb. 2024, Georgia Tech, Algorithms and Randomness Center (ARC) colloquium. (Talk)
- Nov. 2023, Georgia Tech, EECS Rising Stars Workshop. (Talk)
- Nov. 2023, University of Illinois, Urbana-Champaign, Grad Seminar in Special Topics. (Talk)

AI&ML For Cellular Dynamics and Drug Intervention

- May 2026, IGB Symposium, UIUC. (Talk)
- August 2025, the NSF-Simons National Institute for Theory and Mathematics in Biology (NITMB). MathBio Convergence Conference. (Talk)

TEACHING & MENTORING

- Teaching Assistant, ECE 365 Data Science and Engineering, UIUC. Fall 2021
- Organized and taught weekly reading group on Learning in Games and Mean Field Games, with participants from Prof. Subhonmesh Bose's and Prof. Tamer Basar's groups. Fall 2023
- Organized and taught weekly reading group on Functional and Operator-Theoretic Foundation of Learning, joined by Prof. Umesh Vaidya (Clemson), Prof. Amarsagar Reddy Ramapuram Matavalam (AUS), Prof. Subhonmesh Bose and their students. Summer & Fall 2021
- Mentoring: Peizhi Niu, Shane Wang, Siddharth Ambekar, Jonghyeon Lee.

ACADEMIC ACTIVITIES

- Journals Reviewed: Journal of Machine Learning Research (JMLR), SIAM Journal on Control and Optimization (SICON), IEEE Transactions on Automatic Control (TAC), Applied Energy, Nonlinear Dynamics, Frontiers in Applied Mathematics and Statistics
- Conferences Reviewed: ICLR, NeurIPS, AACL, CDC, ACC, L4DC.
- Session organizer, Operator Learning and Behavioral Modeling, Allerton Conference, 2026
- Visiting undergrad scholar, Henry Samueli School of Engineering, UCLA, 2017.

REFERENCES

- Subhonmesh Bose, Associate Professor
Department of Electrical and Computer Engineering
University of Illinois Urbana-Champaign
(217) 244-2101, bores@illinois.edu
- Olgica Milenkovic, Professor
Department of Electrical and Computer Engineering
University of Illinois Urbana-Champaign
(217) 244-7358, milenkov@illinois.edu
- Maxim Raginsky, Professor
Department of Electrical and Computer Engineering
University of Illinois Urbana-Champaign
(217) 244-1782, maxim@illinois.edu
- Umesh Vaidya, Professor
Department of Mechanical Engineering
Clemson University
(864) 656-7471, uvaidya@clemson.edu