

Shiyang (Lilian) Pan

@ sp7207@nyu.edu |  github.com/WaAaaAterfall |  linkedin.com/in/shiyangpan

RESEARCH INTERESTS

Generative models for interpretable, identifiable latent dynamics (flow matching, variational inference, Gaussian processes), causal representation learning, and inverse optimal control, with applications to neural and behavioral time series.

EDUCATION

New York University, Center for Data Science

Ph.D. in Data Science. Advisor: John Pearson

New York, NY

Aug 2026 – May 2029 (expected)

Duke University

Ph.D. in Electrical and Computer Engineering (transferred to NYU with advisor)

Durham, NC

Aug 2024 – Aug 2026

M.S. in Electrical and Computer Engineering, GPA: 4.00/4.00

Aug 2022 – May 2024

Xi'an Jiaotong-Liverpool University (XJTLU)

Suzhou, China

University of Liverpool

Liverpool, UK

Dual B.Sc. in Mathematics, First Class Honours, GPA: 3.96/4.00

Sep 2018 – Jun 2022

PUBLICATIONS

- [1] Ganchao Wei, Daniela de Albuquerque, Miles Martinez, **Shiyang Pan**, and John Pearson. “Dynamic compression flows for neuroscience data.” *ICML*, 2026.
- [2] Yi Li, X. Hermione Xu, **Shiyang Pan**, Xu An, Patrick J. Mulcahey, Zixuan Qiu, Shengli Zhao, Nuo Li, John Pearson, and Z. Josh Huang. “Motor cortical output integrates distorted proprioceptive feedback.” *bioRxiv*, 2025. Under review at *Nature Neuroscience*.
- [3] Jianhua Liu, Shujing Zhang, Jing Cai, Cheng Xue, Zhiyong Jin, **Shiyang Pan**, Ruoxuan Yang, Shenwei Zhang, and Mingfeng Cao. “Pupil dynamics in macaque recognition memory tasks: Investigating physiological mechanisms.” *Physiology and Behavior*, 2025.
- [4] **Shiyang Pan**, Yuxin Zhang, Zhen Wei, Jia Meng, and Daiyun Huang. “Prediction and motif analysis of 2'-O-methylation using a hybrid deep learning model from RNA primary sequence and nanopore signals.” *Current Bioinformatics*, 2022.

SELECTED RESEARCH EXPERIENCE

Motor Control in Mouse Reaching

Advisor: John Pearson

Nov 2025 – Present

NYU and Duke University

Developing latent variable models to study how mouse motor cortex controls reaching behavior, using mechanical and optogenetic perturbations as causal interventions.

- Developing a generative model that explains trial-to-trial variability via latent control signals, combining a GP-VAE over controls with flow matching to capture nonlinear, multimodal reach dynamics.
- Developing causal representation learning methods to identify latent causal structure of reaching from perturbation-paired data.
- Modeled how motor cortical output integrates distorted proprioceptive feedback using GLM and Poisson LDS models.

Dynamic Compression Flows

Advisor: John Pearson

Oct 2025 – Jan 2026

Duke University

Contributed to a dual flow-matching model that learns identifiable low-dimensional latent dynamics. Ran baseline comparisons against existing methods across neural and behavioral video datasets, and produced the comparison figures.

Online Latent Models for Calcium Imaging

Feb 2024 – Present

Advisor: John Pearson

Duke University

Developed a state-space model to discover low-dimensional population dynamics from streaming large-scale calcium imaging data, fitted with particle-filter EM strategy and implemented in JAX.

Real-Time Analysis for Live Neuroscience Experiments

May 2023 – Aug 2023

Advisor: John Pearson

Duke University

Built a Python system to classify neurons from streaming calcium imaging during live experiments, with an interface for monitoring results.

Deep Learning for RNA Modification

Apr 2021 – Mar 2022

Advisor: Jia Meng

XJTLU

Predicted RNA 2'-O-methylation sites from sequence and nanopore signals with a hybrid CNN-RNN model, and applied integrated gradients to identify the sequence motifs driving predictions.

INDUSTRY EXPERIENCE

Bosch China

Feb 2022 – May 2022

Data Engineer Intern

Suzhou, China

- Developed an unsupervised pipeline to classify vehicle operating states from multivariate sensor data, using PCA, GMM, and k-means to identify distinct operating regimes without ground-truth labels.
- Designed and implemented Python and MATLAB workflows to process large-scale sensor data across five vehicle types.

TEACHING

Theory and Algorithms for Machine Learning (CS 671)

Fall 2025

Teaching Assistant

Duke University

Advanced Algebra Seminar

Sep 2020 – Aug 2021

Undergraduate Instructor, Dept. of Mathematical Sciences

XJTLU

SKILLS

Programming: Python (JAX, PyTorch, NumPy, SciPy), MATLAB, R, C/C++

Modeling: flow matching, variational autoencoders, Gaussian processes, causal representation learning, state-space models, generalized linear models

Inference: variational inference, EM, sequential Monte Carlo

Control: optimal control, inverse optimal control, reinforcement learning, Koopman operator methods

AWARDS

University Academic Excellence Scholarship (top 2%), XJTLU, 2019–2021

Summer Undergraduate Research Fellowship, XJTLU, 2021