

Shirley Xiaoqi Liu

Department of Statistics, University of Oxford, UK | shirley.liu@stats.ox.ac.uk | shirleyliuxq.github.io

My research develops statistical and information-theoretic foundations for machine learning, with a focus on learning under data heterogeneity. I characterise fundamental statistical limits and develop principled algorithms for sequential decision-making and high-dimensional inference; I analyse training dynamics with provable guarantees.

RESEARCH

Postdoctoral Researcher, University of Oxford

Oct 2024 – present

Host: Prof. Patrick Rebeschini

- Characterise training dynamics of canonical statistical and machine learning models, with a focus on early stopping, benign overfitting, and feature learning
- Sequential decision-making with rigorous guarantees, including bandits and adaptive inference
- Uncertainty quantification using assumption-lean methods, including e-values, conformal prediction, and prediction-powered inference

PhD, University of Cambridge

Oct 2019 – Jun 2024

Advisor: Prof. Ramji Venkataramanan

Thesis: Message passing algorithms for high-dimensional statistical estimation and communication

- Efficient offline changepoint localisation in high-dimensional regression, in the asymptotic proportional regime
- Near-optimal sketching schemes for sparse, low-rank matrices
- Efficient encoding and decoding schemes for communication in large user networks, in the asymptotic proportional regime

SELECTED PREPRINTS & PUBLICATIONS

- [1] A. Buna, **S.X. Liu**, P. Rebeschini, “Minimax Optimal Early-Stopped Gradient Descent for Gaussian Mixture Classification”, in submission, 2026. ([paper](#), [talk](#))
- [2] J. Allison, ..., **X. Liu**, ... (alphabetical order), “Laser writing in glass for dense, fast and efficient archival data storage”, *Nature*, 2026. ([paper](#))
- [3] **X. Liu**, D. Baudry, J. Zimmert, P. Rebeschini, A. Akhavan, “Non-stationary Bandit Convex Optimization: A Comprehensive Study”, *Annual Conference on Neural Information Processing Systems*, 2025. ([paper](#), [talk](#), [poster](#))
- [4] **X. Liu**, P. Pascual Cobo, R. Venkataramanan, “Many-User Multiple Access with Random User Activity: Achievability Bounds and Efficient Schemes”, *IEEE Transactions on Information Theory*, 2025. ([paper](#), [conference version](#), [talk](#), [poster](#))
- [5] **X. Liu**, K. Hsieh, R. Venkataramanan, “Coded many-user multiple access via Approximate Message Passing”, *Information Theory, Probability and Statistical Learning: A Festschrift in Honor of Andrew Barron*, 2025. ([preprint](#), [conference version](#), [talk](#))
- [6] G. Arpino, **X. Liu**, J. Gontarek, R. Venkataramanan, “Inferring Change Points in High-Dimensional Regression via Approximate Message Passing”, *Journal of Machine Learning Research*, 2025. ([paper](#))
- [7] G. Arpino, **X. Liu**, R. Venkataramanan, “Inferring Change Points in High-Dimensional Linear Regression via Approximate Message Passing”, *International Conference on Machine Learning*, 2024. ([paper](#), [talk](#), [poster](#), [code](#))
- [8] **X. Liu**, R. Venkataramanan, “Sketching Sparse Low-Rank Matrices With Near-Optimal Sample- and Time-Complexity Using Message Passing”, *IEEE Transactions on Information Theory*, 2023. ([paper](#), [conference version](#))

EDUCATION

BA and MEng, University of Cambridge

Oct 2015 – Jun 2019

Information and Computer Engineering (Honours with Distinction): ranked top 3–7% each year in a cohort of 300+

Overseas Family School, Singapore

Aug 2013 – Jun 2015

IB Diploma (Bilingual): 45/45, top 0.3% globally

SELECTED TALKS

- [1] “Bayes-Assisted Confidence Sequences for CDFs”, Safe Anytime-Valid Inference (SAVI) Conference, University of Twente, July 2026.
- [2] “Mind the U Curve: Why We Stop Gradient Descent Early”, Meet Our Group Seminar Series, Department of Statistics, University of Oxford, May 2026.
- [3] “Learning Under Non-Stationarity: Statistical Inference & Bandit Convex Optimization”, Algorithms and Computationally Intensive Inference Seminar, Department of Statistics, University of Warwick, March 2026. ([slides](#))
- [4] “A Friendly Talk on Bandit Convex Optimization in Changing Environments”, Oxford Young Statisticians Seminar,

Department of Statistics, University of Oxford, March 2026.

- [5] “Tight Confidence Sequences From Universal Portfolio”, Hypothesis Testing with E-values Reading Group, Department of Statistics, University of Oxford, March 2026. ([slides](#))
- [6] “Tutorial on Sequential Anytime-Valid Inference Using E-processes”, Statistics in AI CDT Module, Department of Statistics, University of Oxford, November 2025. ([slides](#))
- [7] “Sequential Anytime-Valid Inference Using E-processes”, Hypothesis Testing with E-values Reading Group, Department of Statistics, University of Oxford, October 2025. ([slides](#))
- [8] “Loss landscapes and optimization in over-parameterized neural networks”, Learning Theory and Statistical Optimization Reading Group, Department of Statistics, University of Oxford, November 2024. ([slides](#))
- [9] “Communication over many-user channels via Approximate Message Passing”, Information Theory Seminar, Department of Mathematics, University of Cambridge, May 2024. ([slides](#))

SERVICE & OUTREACH

- **Reviewer:** NeurIPS (top reviewer 2024), ICLR, TMLR, ISIT
- **Lecturer & speaker,** Oxford Algorithmic Statistics CDT module (Nov 2025): lecture on uncertainty quantification via e-values and e-processes; research talk on non-stationary bandit optimisation
- **Research project supervision** (2022–present): MSc projects on in-context learning in transformers, e-values, and prediction-powered inference (2025–present); UNIQplus (Oxford widening-access research internship programme) project on sequential uncertainty quantification (2026); MEng projects on high-dimensional statistical estimation for heterogeneous data (2023–2024)
- **Organiser & speaker,** Cambridge Information Engineering Divisional Conference (Mar 2022)
- **Teaching:** Supervisor for 3F7 Information Theory & Coding and 3F4 Data Transmission (2019–2024), and Cambridge AI+ Programme (2022, 2023)

SCHOLARSHIPS & AWARDS

- Cambridge Trust International PhD Scholar (£49,000/year)
- 2020 British Education Award (Nov 2019)
- Best Thesis Presentation Award (Jun 2019)
- Newnham College Scholar, Cambridge (2016–2019)

WORK EXPERIENCE

Research Intern, Microsoft Research, Cambridge

Apr – Jun 2023

- Designed a novel error correction coding scheme for Project Silica (glass-based cloud data storage system), increasing transmission rate by 5–10%
- Clarified a key information-theoretic performance metric and used it to build a unifying evaluation framework for selecting the optimal coding scheme
- Enhanced image classification decoder in PyTorch
- Co-authored the corresponding section of a resulting *Nature* publication

Software Development Intern, MediaTek Inc, Cambourne

Jun – Sep 2017

REFEREES

Prof. Patrick Rebeschini

patrick.rebeschini@stats.ox.ac.uk

Department of Statistics, Oxford

Prof. Ramji Venkataramanan

rv285@cam.ac.uk

Department of Engineering, Cambridge

Prof. Ioannis Kontoyiannis

ik355@cam.ac.uk

Department of Pure Mathematics and Mathematical Statistics, Cambridge