

# Szilvia Ujváry

sru23@cam.ac.uk | szilviaujvary.github.io | linkedin.com/in/szilviaujvary | github.com/szilviaujvary | Google Scholar

## SUMMARY

I am a final-year PhD student at the University of Cambridge studying **language model generalisation and inductive biases**, with a focus on out-of-distribution compositional generalisation, implicit Bayesian inference, and how training data shapes model behaviour. Alongside my research, I have hands-on industry experience in LLM **pre-training** at Apple MLR and **post-training** at Reasonable, where I led fine-tuning projects in collaboration with NVIDIA.

## EDUCATION

### University of Cambridge

Sept 2023 – Apr 2027 (expected)

*PhD in Machine Learning, supervised by Ferenc Huszár*

*Cambridge, UK*

- Investigated what determines language model behaviour beyond next-token prediction loss, establishing limitations of Bayesian explanations of out-of-distribution generalisation and showing how training order introduces inductive biases into what models learn.
- Visiting student in the Computational and Biological Learning Lab (CBL), with José Miguel Hernández-Lobato
- TA and Course Developer, *Deep Neural Networks* (3rd-year module), 4th year research project supervisor

### University of Cambridge

Sept 2022 – Aug 2023

*MPhil in Machine Learning and Machine Intelligence*

*Cambridge, UK*

- Distinction, ranked 2nd of 44
- Dissertation: tighter PAC-Bayes generalisation bounds via Hamiltonian Monte Carlo (best dissertation in cohort)

### Imperial College London

Sept 2019 – July 2022

*BSc in Mathematics*

*London, UK*

- First Class Honours (87%), Dean's List in all three years (top 10%)

## WORK EXPERIENCE

### Student Researcher, LLM post-training

July 2026 – Oct 2026

*Reasonable (AI startup for formal verification)*

*Budapest, Hungary*

- Led post-training in two NVIDIA Nemotron early-access projects, fine-tuning 30B–120B parameter models to write formally verified Rust proofs, improving pass@1 by 30pp to 73.3%
- Contributed to training-data generation from frontier-model agentic traces
- Presented post-training findings to NVIDIA's Nemotron team, represented Reasonable at GTC Berlin

### Research Scientist Intern

Apr 2025 – Feb 2026

*Apple Machine Learning Research, hosted by Marco Cuturi*

*Paris, France*

- Led a project on pre-training data selection to reduce hallucinations in small language models, showing that prediction loss alone cannot distinguish factual errors from acceptable alternatives. Developed LaCy, a method that teaches models when to delegate predictions to a larger model (TMLR 2026)
- Pre-trained models of up to 1.3B parameters on ~40B tokens

### Deep Tech Fellow (part-time)

Apr 2024 – June 2024

*Lakestar Advisors LLP*

*London, UK*

- Assessed technologies, markets and companies in next-generation deep learning architectures for a venture capital fund, gave technical due diligence and an investment recommendation on an applicant company

### Research Assistant

July 2022 – Oct 2022

*University of Cambridge, supervised by Ferenc Huszár*

*Cambridge, UK*

- Connected Sharpness-Aware Minimisation with mean-field variational inference and developed VariationalSAM, combining the advantages of both (OPT Workshop, NeurIPS 2022)

### Undergraduate Research Assistant

July 2020 – Sept 2021

*Imperial College London*

*London, UK*

- Clustering and dimensionality reduction for single-cell RNA-seq data (Dr Vahid Shahrezaei, 2020)
- Feature selection for Canonical Correlation Analysis with evolutionary algorithms (Dr Marina Evangelou, 2021)

## RESEARCH

---

### Papers

- **Ujváry S.\***, Mladozeniec B.K., Hernández-Lobato J.M., Huszár F., Krasheninnikov D.\*, Is your LLM a sequence model on the training history? The origins and consequences of anticipation. Under review at ICLR 2027; **Spotlight**, 4th Workshop on High-dimensional Learning Dynamics, ICML 2026, \*equal contribution.
- **Ujváry S.**, Béthune L., Ablin P., Monteiro J., Cuturi M., Kirchhof M., LaCy: What Small Language Models Can and Should Learn is Not Just a Question of Loss. Transactions on Machine Learning Research (TMLR), 2026.
- **Ujváry S.**, Mészáros A., Brendel W., Reizinger P., Huszár F., Implicit Bayesian Inference is an Insufficient Explanation of Language Model Behaviour in Compositional Tasks. Preprint, 2025.
- Mészáros A., **Ujváry S.**, Brendel W., Reizinger P., Huszár F., Rule Extrapolation in Language Modeling: A Study of Compositional Generalisation on OOD Prompts. 38th Conference on Neural Information Processing Systems (NeurIPS), 2024, **Spotlight**.
- Reizinger P.\*, **Ujváry S.\***, Mészáros A.\*, Kerekes A.\*, Brendel W., Huszár F., Position: Understanding LLMs Requires More Than Statistical Generalization. 41st International Conference on Machine Learning (ICML), 2024, **Spotlight**, \*equal contribution.
- **Ujváry S.**, Flamich G., Fortuin V., Hernández-Lobato J.M., Estimating optimal PAC-Bayes bounds with Hamiltonian Monte Carlo. Mathematics of Modern Machine Learning Workshop, NeurIPS 2023.
- **Ujváry S.\***, Telek Z.\*, Kerekes A.\*, Mészáros A.\*, Huszár F., Rethinking Sharpness-Aware Minimization as Variational Inference. OPT Workshop on Optimization for Machine Learning, NeurIPS 2022, \*equal contribution.

### Technical Reports

- **Ujváry S.**, Mészáros A., Szilágyi B., Czanik P., Benkő B., Hughes C., Huszár F., Post-training Nemotron 3.5 Super VL for Verified Software Development. Reasonable, with NVIDIA, 2026.
- **Ujváry S.**, Benkő B., Horváth C., Verification at Lightning Speed. Reasonable, with NVIDIA, 2026.

### Service

**Reviewer:** ICML, NeurIPS, ICLR

## SELECTED TALKS

---

- **LaCy: What Small Language Models Can and Should Learn is Not Just a Question of Loss.** Invited talk, Hungarian Machine Learning Days, 2026 (~100 attendees)
- **LLMs, Implicit Bayesian Inference and Compositional Generalisation.** Invited talk, NLIP Seminar Series, Department of Computer Science and Technology, University of Cambridge, 2025
- **Position: Understanding LLMs Requires More Than Statistical Generalization.** Invited talk, Hungarian Machine Learning Days, 2023 (~100 attendees)

## TECHNICAL SKILLS

---

**Languages:** Python, Bash, R, C++

**Deep Learning:** PyTorch, JAX, HuggingFace (Transformers, PEFT), NeMo AutoModel, vLLM

**Tools:** NumPy, Pandas, Weights & Biases, Git, Docker, AWS

## SCHOLARSHIPS AND AWARDS

---

- **Citadel PhD Summit 2026:** Third Prize and a £4,000 research grant
- **ICML 2026:** Silver Reviewer award
- **Emmanuel College, University of Cambridge:** full scholarship for the MPhil (2022)
- **G-Research Prize**, Imperial College London: for academic excellence (2021)
- **Marjorie McDermott Scholarship**, Imperial College London: academic excellence and good citizenship (2020 – 2022)
- **Individual Research Project Prize**, Imperial College London: for a project on Covid-19 epidemic modelling (2020)