# Aditya **Ravuri**

PhD Student · Statistician

(+44 / 0) 7774 528 671 | aditya.ravuri@gmail.com | infprobscix

## Education

### University of Cambridge — *Cambridge, UK*

PhD in Computer Science, **Probabilistic Machine Learning** — *Oct. 2021 - Present*

- Focusing on unifying probabilistic interpretations of unsupervised dimensionality reduction methods in single cell genomics, with a focus on interpretability and identifiability. Concurrently also working on a wide range of case studies of science-constrained unsupervised learning across climate science and biology. Supervised by Prof. Neil Lawrence and funded by the Accelerate Programme for Scientific Discovery.
- Current work involves working with latent Gaussian processes on manifolds and graphs, diffusions (SDEs), Bayesian optimisation over discrete spaces, extensions of hidden Markov models and probabilistic programming.
- Responsibilities include group roles, undergraduate and masters project supervisions.

### University of Cambridge — *Cambridge, UK*

MPhil in Management — *Oct. 2016 - Jun. 2017*

- Commendation (70%). **Focus:** Strategic Valuation (top prize), Economics. **Audits:** Measure Theory, Philosophy of Science, Physics.
- Jaguar Land Rover consulting project: Modeled the contribution of JLR to the UK economy in a highly collaborative setting. Conducted original research on improving the accuracy of input-output multipliers. Highly commended for the work and presentation.
- Coloplast Internship (Aug. '17): Aided implementation of a time series forecasting model (involving splines, linear filters and ARIMA error processes) and automated model fitting.

### Heriot-Watt University — *Edinburgh, UK*

BSc Hons in Actuarial Science, **Statistics Major** — *Sep. 2013 - Jun. 2016*

- **Awards:** Distinction (83%) (top 3%), CT1-8, Volunteering - Bronze (ChessSoc President, Student Union Exec, Mentor). Obtained BSc at age 18.
- **Focus:** Statistics, Quantitative Risk Management, Mathematical Finance. Project Areas: GARCH models, copulas, extreme value theory, stochastic calculus, liquidity risk, economic scenario generators and advanced statistical inference. Published in the actuarial magazine.

## Skills

| | |
|---|---|
| **Programming** | R, Python, Stan, PyTorch, Tensorflow |
| **Languages** | English, Japanese (basic), French (basic), Hindi, Telugu |
| **Other** | ggplot, data.table, Rcpp, sparklyr, git, LaTeX, astrophotography, other PPLs |

## Contributions

**Web** — falmity.com: Personal projects (e.g. speech synthesis using Gaussian processes) and minimal examples (e.g. GAMs as sparse GPs, the Griffin-Lim algorithm, state-space models). Stats Stack Exchange: top 2% contributor in 2018.

**Code** — Contributed to scipy (an efficient Toeplitz matrix-vector product function), geometric-kernels (kernels on graphs), gauche (kernels on spaces of graphs), sympy, torch, gpy and gpytorch. Wrote code for the papers below, including (torch, tf and stan based) packages for working with scalable GPLVMs, GPLVMs with normalizing flows and a probabilistic framework for ice core dating for the British Antarctic Survey. Currently also working on code for Bayesian optimisation over proteins.

## Papers

1. **Ravuri, A.** and Lawrence, N. D. (preprint available on request). *Selected Dimensionality Reduction methods as Probabilistic Inference algorithms with links to Gaussian Processes on manifolds and graphs*. The work interprets UMAP, (t-)SNE, Diffusion Maps, MEU (and thus LLE and LE) as graph based probabilistic models with ties to graph Gaussian processes, which are approximations of Gaussian processes on data-defined manifolds.
2. Lalchand V.*, **Ravuri A.***, Dann E. M.* et. al. (2022). *Modelling Technical and Biological Effects in scRNA-seq data with Scalable GPLVMs*, accessible here (MLCB Oral 2022).
3. Griffiths R. R.*, Klarner L.*, Moss H. B.*, **Ravuri A.***, Truong S.* et. al. *GAUCHE: A Library for Gaussian Processes and Bayesian Optimisation in Chemistry*, accessible here (AI4Science & ReALML ICML Workshops 2022).
4. Lalchand, V., **Ravuri, A.** and Lawrence, N. D. (2022). *Generalised GPLVMs with SVI, AISTATS 2022*, accessible here.
5. **Ravuri, A.** et al. (2022). *Ice Core Dating using Probabilistic Programming*, accessible here (NeurIPS GPs & Spatiotemporal Modelling Workshop 2022).

# Employment

## Barclays

*London, UK*

Quant Analyst + Developer

*Dec. 2018 - Sep. 2021*

- Designed and productionized large-scale statistical models for balance sheet simulation of term deposits and loans, accounting for customer behavior and economic trends. Modeling mainly involved Markovian models, GAMs and time series models.
- In addition to this, I assisted with and reviewed other model implementations (e.g. for current accounts, savings and mortgages). I also piloted new tools, created knowledge-bases, worked on automation and performed exploratory work to identify areas of efficiency (e.g. with Spark, Rcpp, Docker). In some cases, I reduced execution times from days to seconds.

## Sciemus

*London, UK*

Data Scientist + Statistician

*Sep. 2017 - Dec. 2018*

- Was involved with building and maintaining end-to-end stats/tech related solutions, particularly in the space, weather and power business areas. This involved data cleaning, analysis, modeling, documentation, web-app development and deployment (using Shiny, Dash, Flask), basic linux server and database maintenance (using postgres), research and development of infrastructure (e.g. aiding development of a distributed computing cluster on AWS).
- On the modeling side, I've worked with GLMs for assessing risk probabilities, Hidden Markov models & sparse Gaussian Processes to model rates based on large-scale weather data, importance sampling & subset simulation to accelerate simulations and other ideas in Bayesian statistics.