

# HENRY MOSS

henry.moss@secondmind.ai

## RESEARCH INTERESTS

---

**Bayesian optimisation & active learning:** information-theory; batch design; multi-fidelity.

**Gaussian processes:** scalable models; multi-fidelity; structural kernels.

**Experimental design:** gene design; molecular search.

**Natural language processing:** text-to-speech, hyper-parameter optimisation, model selection.

## CURRENT ROLE

---

**Research Scientist**, Secondmind, Cambridge.

2022 -

- Co-lead of a team of 4 engineers, designing an ML system for electric vehicle engine calibration.

**Senior Machine Learning Researcher**, Secondmind, Cambridge.

2021 - 2022

## EDUCATION

---

**PhD in Machine Learning and Statistics**, STOR-i CDT, Lancaster University.

2017 - 2021

- *General-purpose Information-theoretical Bayesian Optimisation*
  - Extended information-theoretical Bayesian optimisation for batch and multi-fidelity designs.
  - Developed Bayesian optimisation methods for high-cost string design problems.
- Supervision by Prof. David Leslie (Statistics) and Prof. Paul Rayson (Computer Science).

**MRes in Statistics and Operational Research**, STOR-i CDT, Lancaster University.

2016 - 2017

**MA in Mathematics**, Emmanuel College, University of Cambridge.

2013 - 2016

## PUBLICATIONS

---

### Published:

Picheny V., **Moss H. B.**, Durrande N. & Torossian L. Bayesian Quantile and Expectile Optimisation. In *Association of Uncertainty in Artificial Intelligence (UAI)*, 2022.

**Moss H. B.**, Ober S. W. & Picheny V. Information-theoretic Inducing Point Placement for High-throughput Bayesian Optimisation. In *The International Conference on Machine Learning: Workshop on Real World Experimental Design and Active Learning (ICML: Real-ML Workshop)*, 2022.

Payeles A., **Moss H. B.** & Picheny V. A Penalisation Method for Batch Multi-objective Bayesian Optimisation with Application in Heat-exchanger Design. In *The International Conference on Machine Learning: Workshop on Real World Experimental Design and Active Learning (ICML: Real-ML Workshop)*, 2022

Griffiths R., Klarner L., **Moss H. B.**, Ravuri A., Truong S. T., Ranković B., Schwaller P., Du Y., Jamasb A. R., Schwartz J., Tripp A., Kell G., Bourached A., Chan A., Moss J. Guo G., Lee A. & Jiang Tang. GAUCHE: A Library for GAussian Processes and Bayesian Optimisation in CHEmistry. In *The International Conference on Machine Learning: AI for Science Workshop (ICML: AI4Science Workshop)*, 2022

Vakili S., **Moss H. B.**, Artmev A., Dutordoir V. & Picheny V. Scalable Thompson Sampling using Sparse Gaussian Process Models. In *The Conference on Neural Information Processing Systems (NeurIPS)*, 2021.

**Moss H. B.**, Leslie D. S., Gonzalez J. & Rayson P. General-purpose Information-based Bayesian Optimisation. In *The Journal for Machine Learning Research (JMLR)*, 2021.

**Moss H. B.**, Beck D., Leslie D. S., Gonzalez J. & Rayson P. Bayesian Optimisation over String spaces. In *The Conference on Neural Information Processing Systems (NeurIPS)*, 2020 (**spotlight**).

**Moss H. B.** & Griffiths R. Gaussian Process Molecule Property Prediction With FlowMO. In *The Conference on Neural Information Processing Systems: Machine Learning for Molecules Workshop (NeurIPS: ML4Molecules Workshop)*, 2020 (**selected talk**).

**Moss H. B.**, Leslie D. S. & Rayson P. BOSH: Bayesian Optimisation by Sampling Hierarchically. In *The International Conference on Machine Learning: Workshop on Real World Experimental Design and Active Learning* (ICML: Real-ML Workshop), 2020.

**Moss H. B.**, Leslie D. S. & Rayson P. MUMBO: Multi-task Max-value Bayesian Optimisation. In *The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases* (ECML), 2020.

**Moss H. B.**, Aggarwal V., Prateek N., Gonzalez J. & Barra-Chicote R. BOFFIN TTS: Few-shot Speaker Adaptation By Bayesian Optimisation. In *The International Conference on Acoustics, Speech and Signal Processing* (ICASSP), 2020.

**Moss H. B.**, Moore A., Leslie D. S. & Rayson P. FIESTA: Fast IdEntification of State-of-The-Art Models Using Adaptive Bandit Algorithms. In *The Annual Meeting of the Association of Computational Linguists* (ACL), 2019.

**Moss H. B.**, Leslie D. S. & Rayson P. Using  $J$ - $K$ -fold Cross Validation to Reduce Variance when Tuning Natural Language Processing Models. In *The International Conference on Computational Linguistics* (COLING), 2018 (area chair favourite).

### Under Review:

**Moss H. B.**, Ober S. W. & Picheny V. Inducing Point Allocation for Sparse Gaussian Processes in High-throughput Bayesian Optimisation Loops.

Ranković B., Griffiths R., **Moss H. B.** & Schwaller P. Bayesian Optimisation-accelerated Additives Screening and Yield Improvements in Chemical Reactions.

Qing J., **Moss H. B.** & Couckuyt I.  $\{PF\}^2$ ES: Parallel Feasible Pareto Frontier Entropy Search for Multi-Objective Bayesian Optimization under Unknown Constraints.

Payeles A., **Moss H. B.** & Picheny V. HIPPO: Highly Parallelisable Pareto Optimisation for Multi-Objective Bayesian Optimisation.

Griffiths R., Thawani A., Jamasb A., **Moss H. B.**, Bourached A., Jones P., McCorkindale W. & Aldrick A. A Case for Domain Expert Dataset Curation in Machine-Learning Enabled Chemistry.

Chang P., Verma P., John ST., **Moss H. B.**, Picheny V. & Solin A. Fantasizing with Dual GPs in Bayesian Optimization and Active Learning.

## PRIZES

---

<b>Extra Mile award:</b> Embodying Secondmind’s “Delight Every Customer” value.	2022
<b>NeurIPS spotlight paper:</b> Top 3% of submissions at NeurIPS.	2020
<b>ML4Molecules contributed talk:</b> Top 5% of submissions at the ML4Molecules Workshop at NeurIPS.	2020
<b>Nick Smith prize:</b> Best second-year Statistics PhD student at Lancaster University.	2019
<b>Area chair favourite:</b> Nominated for overall best paper at COLING 2018.	2018

## PRESENTATIONS

---

<b>Lancaster Alumni Day:</b> <i>Job Hunting and Building an Early Career</i> (panel Session).	2022
<b>Secondmind Seminar:</b> <i>Inducing Point Allocation for Sparse Gaussian Processes ...</i> (talk).	2022
<b>Real-ML@ICML:</b> <i>Information-theoretic Inducing Point ...</i> and <i>A penalisation method for ...</i> (posters).	2022
<b>UAI:</b> <i>Quantile and Expectile Optimisation</i> (poster).	2022
<b>Secondmind:</b> <i>Introduction to Multi-Fidelity Modelling</i> and <i>Introduction to Profile Optimisation</i> (talks).	2022
<b>Univ. Cambridge ML@Computer-Lab</b> and <b>Univ. Bern:</b> <i>GIBBON</i> (talk).	2022
<b>Gaussian Process Summer School:</b> <i>Overview of Secondmind’s Toolboxes</i> (talk).	2021
<b>Secondmind:</b> <i>Introduction to Information Theory</i> (talk).	2021
<b>NeurIPS:</b> <i>BOSS: Bayesian Optimisation over String Spaces</i> (talk).	2020
<b>ML4Molecules@NeurIPS:</b> <i>Gaussian Process Molecule Property Prediction with FlowMo.</i> (talk).	2020
<b>ECML:</b> <i>MUMBO: Multi-task Max-value Bayesian Optimisation</i> (talk).	2020
<b>Microsoft Research Cambridge:</b> <i>Bayesian Optimisation in Gene Design Loops</i> (talk).	2020
<b>Mathematics of Data Science Conference</b> and <b>ICML: RealML Workshop:</b> <i>BOSH</i> (talk).	2020

<b>ICASSP, Amazon Research Cambridge and University of Melbourne:</b> <i>BOFFIN TTS</i> (talk).	2020
<b>Prowler.io, Sheffield, Lancaster and Manchester Universities;</b> <i>MUMBO</i> (talk).	2019
<b>Amazon Intern Colloquium:</b> <i>Rapid Speaker Adaptation with Bayesian Optimisation</i> (poster).	2019
<b>ACL:</b> <i>FIESTA: Fast Identification of SOTA</i> (talk).	2019
<b>Google NLP Summit:</b> <i>Reliable and Efficient Hyper-parameter Tuning for NLP</i> (poster).	2019
<b>STOR-i Forum and Lancaster Data Science Group:</b> <i>A Crash Course in Bayes Opt</i> (talk).	2018
<b>Microsoft AI Summer School and COLING:</b> <i>Using J-K-fold Cross Validation ...</i> (poster)	2018
<b>UCREL Summer School in Corpus-based NLP:</b> <i>Instabilities in NLP models</i> (poster).	2017

## INTERNSHIPS AND SUMMER SCHOOLS

<b>Visiting researcher</b> , School of Computing and Information Systems, University of Melbourne.	2020
• Derived Bayesian optimisation for sequence design under syntactic constraints.	
<b>Amazon PhD internship</b> , Text-To-Speech Team, Amazon Alexa, Cambridge.	2019
• Used Bayesian optimisation to fine-tune neural systems to synthesise new voices with limited data.	
<b>Amazon Intern Colloquium</b> , Amazon Research, Cambridge.	2019
<b>Google NLP Summit</b> , Google Research, Zurich.	2019
<b>Microsoft AI Summer School</b> , Microsoft Research, Cambridge.	2018
<b>UCREL NLP Summer School</b> , Lancaster University.	2017
<b>Wellcome Sanger internship</b> , University of Cambridge.	2016
• Designed a system to automatically flag promising compounds during image-based drug screening.	
<b>Summer research internship</b> , STOR-i CDT, Lancaster University.	2016
<b>Equity analyst:</b> Oil & Gas researcher at Redburn International.	2015

## RESEARCH SUPPORT

<b>Visiting researcher grant:</b> Collaboration sponsorship from the University of Melbourne.	2020
<b>STOR-i research fund:</b> Support for visit to the University of Melbourne from Lancaster University.	2020
<b>Workshop sponsorship:</b> Support for Bayesian optimisation workshop from Amazon Research.	2019
<b>Faculty of Science and Technology travel grant:</b> Travel support from Lancaster University.	2018
<b>STOR-i PhD scholarship:</b> Full funding for MRes and PhD programme.	2016

## OTHER RELEVANT EXPERIENCE

<b>Workshop Organiser:</b> Organised interdisciplinary workshop on Gaussian processes in Astronomy.	2022
<b>Seminar Coordinator:</b> Ran Secondmind's external speaker seminar series.	2022
<b>Internship Coordinator:</b> Ran Secondmind's summer internship programme.	2022
<b>Statistical Consultant:</b> UK Environment Agency: Flood Hydrology Road Map.	2022
<b>Reviewer:</b> Reviewed manuscripts for Neural Computation and NeurIPS.	2022
<b>Reading Group Coordinator:</b> Ran Secondmind's summer internal reading group.	2021
<b>Internship Supervisor:</b> Supervisor for two summer placement students.	2021
<b>Program Committee:</b> Committee member for the NeurIPS ML4Molecules Workshop.	2020 - 2021
<b>Reviewer:</b> Reviewed manuscripts for AISTATS and ECML.	2020
<b>MSc Masterclass:</b> Designed and taught short course on Bayesian optimisation.	2019 - 2021
<b>Statistical Consultant:</b> Natural Language Identification Project, Lancaster University.	2018
<b>STOR-i Computing Team:</b> Assist peers with coding and distributed computing.	2017 - 2019
<b>Tutor:</b> Undergraduate Mathematics and MSc Data Science.	2017 - 2019
<b>Lancaster University Outreach:</b> Interactive sessions with local secondary schools.	2016 - 2018

## COMPUTING SKILLS

Python, Tensorflow, PyTorch, GPflow, Git, Maintainer of Trieste, Contributor to Emukit, GPy, GPyOpt, Cython, MXNet, R, C.