

Mohsin Hasan

Tel: 647-571-8737

Email: hasanmohsin19@gmail.com

GitHub: <https://github.com/hasanmohsin>

LinkedIn: <https://www.linkedin.com/in/mohsin-hasan-421891254/>

Website: <https://hasanmohsin.github.io/>

EDUCATION

University of Montreal, Mila

(Sep 2023 – Expected May 2027)

PhD in Computer Science – Supervisor: Yoshua Bengio, Year 3

Topics: RL Fine-tuning, Diffusion and Flow Models, Sampling, Bayesian Inference

University of Waterloo

(Sep 2021 – Aug 2023)

Research Master's in Computer Science – Supervisor: Pascal Poupart

Topics: Bayesian Federated Learning, Calibration

CGPA: 4.0/4.0 (Cumulative Percentage: 100%)

University of Toronto, St. George Campus

(Sep 2016 – May 2021)

Bachelor's in Electrical and Computer Engineering (U of T Scholar)

CGPA: 3.99/4.0 (Cumulative Percentage: 95.5%)

- Rank 2 in Computer Engineering 2nd and 3rd Year

- Rank 1 in Computer Engineering 1st Year

SKILLS

- Technical languages: Python, MATLAB, LaTeX, C++, C, Perl, Java
- Software/Platforms: PyTorch, TensorFlow, Slurm Compute Cluster, Linux, Simulink, Microsoft Office

RESEARCH EXPERIENCE

PhD Research Student – University of Montreal, Mila

(Sep 2023 – Present)

- Develop scalable methods for reward fine-tuning of generative models (particularly diffusion/flow models), through both RL training and inference-time methods
- Design scalable methods to train samplers for energy functions (both discrete and continuous), with applications to problems in biology and physics
- Led and collaborated on multiple projects published at NeurIPS, ICLR and ICML

Master's Research Student – University of Waterloo, Vector Institute

(Oct 2021 – Aug 2023)

- Developed methods for efficient Bayesian inference in a federated setting
- Invented and evaluated a method for aggregating Bayesian posteriors in the predictive space, enabling federated learning in a single communication round with accurate calibration
- Outperformed state-of-the-art methods on heterogeneous data
- Published as first author at AAAI 2024

Research Student – University of Toronto (with Will Grathwohl)

(May 2021 – Aug 2021)

- Investigated sampling from an intractable density using Learned Stein Discrepancy (LSD)
- Implemented the sampler and compared it to baselines on simple distributions and RBMs
- Formulated an alternative version of the sampler that implicitly minimized KL divergence instead of LSD
- Outperformed Stein Variational Gradient Descent in higher dimensions, but underperformed the Langevin sampler

Research Student – University of Toronto, Vector Institute (CS PAIR Lab)

(Aug 2020 – Aug 2021)

- Led a research project designing a deep learning algorithm to construct an action graph from an instructional video

- The model took video and transcript text as input and generated a “visually grounded action graph”, with nodes containing critical action steps (as text) and graph edges representing action interdependencies. Entities in action steps were also linked to bounding boxes and tracked over the video.
- Adapted a Transformer-based architecture to find joint visual-linguistic representations over the input

PROFESSIONAL EXPERIENCE

Engineering Intern – Qualcomm Markham (Display Team)

(May 2019 – Aug 2020)

- Designed image-processing algorithms that enhance image quality (e.g. lighting and saturation) using deep learning and general statistical techniques
- Developed a learning-based algorithm for simulating high-dynamic-range images by manipulating lighting content in standard-dynamic-range images

PUBLICATIONS

Adaptive Order Policies for Masked Diffusion

(May 2026)

Mohsin Hasan*, Jama Hussein Mohamud*, Mirco Ravanelli, Yoshua Bengio

- Under review, preprint at <https://arxiv.org/abs/2606.00295>

Discrete Feynman-Kac Correctors

(Jan 2026)

Mohsin Hasan*, Viktor Ohanesian*, Artem Gazizov, Yoshua Bengio, Alán Aspuru-Guzik, Roberto Bondesan, Marta Skreta, Kirill Neklyudov

- Under review, preprint at <https://arxiv.org/abs/2601.10403>

Outsourced diffusion sampling: Efficient posterior inference in latent spaces of generative models

(Feb 2025)

Siddarth Venkatraman*, **Mohsin Hasan***, Minsu Kim, Luca Scimeca, Marcin Sendera, Yoshua Bengio, Glen Berseth, Nikolay Malkin

- ICML 2025, <https://arxiv.org/abs/2502.06999>

Steering masked discrete diffusion models via discrete denoising posterior prediction

(Oct 2024)

Jarrid Rector-Brooks, **Mohsin Hasan**, Zhangzhi Peng, Zachary Quinn, Chenghao Liu, Sarthak Mittal, Nouha Dziri, Michael Bronstein, Yoshua Bengio, Pranam Chatterjee, Alexander Tong, Avishek Joey Bose

- ICLR 2025, <https://arxiv.org/abs/2410.08134>

Amortizing intractable inference in diffusion models for vision, language, and control

(May 2024)

Siddarth Venkatraman*, Moksh Jain*, Luca Scimeca*, Minsu Kim*, Marcin Sendera*, **Mohsin Hasan**, Luke Rowe, Sarthak Mittal, Pablo Lemos, Emmanuel Bengio, Alexandre Adam, Jarrid Rector-Brooks, Yoshua Bengio, Glen Berseth, Nikolay Malkin

- NeurIPS 2024, <https://arxiv.org/abs/2405.20971>

Estimating Expectations without Sampling: Neural Stein Estimation

(May 2024)

Mohsin Hasan, Dinghui Zhang, Cheikh Ahmed, Awa Khouna, Yoshua Bengio

- Appeared at the Approximate Inference Workshop <https://openreview.net/pdf?id=by9cKVIs0F>

Calibrated One Round Federated Learning with Bayesian Inference in the Predictive Space

(Feb 2024)

Mohsin Hasan, Guojun Zhang, Kaiyang Guo, Xi Chen, Pascal Poupart

- AAAI 2024, <https://ojs.aaai.org/index.php/AAAI/article/view/29122>

TALKS

International Workshop on Trustworthy Federated Learning in conjunction with IJCAI 2022

(July 2022)

15-minute presentation (<https://federated-learning.org/fl-ijcai-2022/>)

ACHIEVEMENTS

- Received NSERC USRA for undergraduate research
- Received over **\$56,000** in scholarships and awards
- One of eight students to receive the First Year Research Fellowship (**\$5,000**)

(May 2018)
(Apr 2017 – Present)
(May 2017)

RELEVANT COURSEWORK

- Graduate: Dynamical Systems - MAT6215 (**A+**)
- Graduate: Probabilistic Graphical Models - IFT6269 (**A+**)
- Graduate: Stochastic Processes - AMATH777 (**100%**)
- Graduate: Optimization - CS794 (**100%**)
- Graduate: Reinforcement Learning - CS885 (**100%**)
- Undergraduate: Deep Learning - CSC413 (**96%**)

(University of Montreal)
(University of Montreal)
(University of Waterloo)
(University of Waterloo)
(University of Waterloo)
(University of Toronto)

TEACHING ASSISTANT

- Graduate: Representation Learning - IFT6135
- Graduate: Reinforcement Learning - CS885

(University of Montreal)
(University of Waterloo)