Andrew Gordon Wilson

CONTACT INFORMATION

Courant Institute of Mathematical Sciences
New York University
60 5th Ave, New York, NY, 10011, USA

andrewgw@cims.nyu.edu
https://cims.nyu.edu/~andrewgw
https://twitter.com/andrewgwils
Google Scholar Profile

RESEARCH INTERESTS

I aim to develop a prescriptive understanding of how to build intelligent systems. My works spans generalization theory, architecture design, optimization, uncertainty representation, time series forecasting, scaling laws, synthetic data, out-of-distribution generalization, scientific discovery, and beyond. I also believe in reproducible research, and my work is the basis for several open-source libraries.

CURRENT POSITION

Research Lead, Perplexity **2026 – Present**
Leading the frontier research efforts at Perplexity AI

Professor, New York University **2024 – Present**
Associate Professor (with tenure) **2022 – 2024**
Assistant Professor **2019 – 2022**
Courant Institute of Mathematical Sciences and Center for Data Science
Computer Science Department, Mathematics Department (Affiliated)

ACADEMIC BACKGROUND

Assistant Professor, Cornell University **2016 – 2019**
Operations Research and Information Engineering
Field member of ORIE, Computer Science, Statistics, and Applied Mathematics

Research Fellow, Carnegie Mellon University **2014 – 2016**
Machine Learning Department, School of Computer Science

PhD, Trinity College, University of Cambridge **2009 – 2014**
Machine Learning, Department of Engineering

BSc (Hons), University of British Columbia **May 2008**
Mathematics and Physics
A+ Graduating Average, Highest Ranking Honours Physics Thesis.

AWARDS

- Test of Time Award, Honourable Mention (for *Deep Kernel Learning*), AISTATS 2026
- Heilbronn Distinguished Fellow, Isaac Newton Institute, University of Cambridge 2025
- Best Paper Award, Molecular Machine Learning Conference 2025
- Best Paper Award, ICML Exploration in AI Workshop 2025
- Best Paper, Runner Up, ICLR Genetics Workshop (AI4NA) 2025
- Best Paper Award, ICML Theoretical Foundations Workshop 2024
- Outstanding Paper Award, NeurIPS AI DrugX Workshop 2024
- Outstanding Paper Award (for work on *Bayesian Model Selection*), ICML 2022
- NSF Career Award 2022
- Outstanding Area Chair, ICLR 2022
- Amazon Machine Learning Research Award ($60,000 + $100,000 AWS Credits) 2020
- Amazon Research Award ($100,000) 2019
- Best Paper Award, NeurIPS Time Series Workshop 2019
- Facebook Research Award ($130,000) 2018
- Best Poster Award (with Ben Athiwaratkun), NeurIPS ML Train Workshop 2017
- Outstanding PhD Dissertation (£10,000), G-Research 2014
- Outstanding Reviewer Award, Neural Information Processing Systems (NeurIPS) 2013

- Best Student Paper Award, Uncertainty in Artificial Intelligence (UAI) 2011
- Schiff Foundation Studentship 2009-2014
- NSERC Postgraduate Scholarship (Doctoral) (PGS-D) 2010-2013
- Trinity College Overseas Bursary 2009-2013
- Cambridge Commonwealth Trust 2009-2013
- John Collison Memorial Scholarship in Mathematics 2007-2008
- Dean's Honour List, Science Scholar, Undergraduate Program Scholarship, TRIUMF Research Scholarship, NSERC Undergraduate Research Scholarship (USRA)

Refereed Publications

[1] E. Su, A. Potapczynski, S. Qiu, E. Hughes, A. G. Wilson. *Epiplexity Guided Data Selection and Generation for Out-of-Distribution Generalization.* Advances in Neural Information Processing Systems (NeurIPS), 2026.

[2] E. Baron, A. N. Amin, A. G. Wilson. *Deep Generative Models for Phylogenetic Inference with Complex Evolutionary Processes.* Advances in Neural Information Processing Systems (NeurIPS), 2026.

[3] S. Qiu, Z. Chen, A. Potapczynski, M. Marek, A. G. Wilson. *Scaling Arbitrary Architectures and Optimizers with Automatic Parameterization.* Advances in Neural Information Processing Systems (NeurIPS), 2026.

[4] S. Qiu, M. Finzi, Y. Zheng, K. Zhang, A. G. Wilson. *Requential Coding: Pushing the Limits of Model Compression with Self-Generated Training Data.* Advances in Neural Information Processing Systems (NeurIPS), 2026.

[5] V. Baherwani, Z. Chen, S. Qiu, A. G. Wilson, P. Izmailov. *Emergent Capabilities Arise Randomly from Learning Sparse Attention Patterns.* Advances in Neural Information Processing Systems (NeurIPS), 2026.

[6] M. L. Wiemann, L. M. Smith, P. Melchior, S. Mishra-Sharma, A. G. Wilson, P. Izmailov, C. Cuesta-Lázaro. *DiscoverPhysics: Benchmarking LLMs for Out-of-the-Box Scientific Thinking.* Advances in Neural Information Processing Systems (NeurIPS), 2026.

[7] T. Papamarkou *et al. Agentic AI Orchestration Should Be Bayes-Consistent.* International Conference on Machine Learning (ICML), 2026.

[8] X. Wang, W. Cao, A. G. Wilson, Z. Zeng. *Automatic Layer Selection for Hallucination Detection.* International Conference on Machine Learning (ICML), 2026.

[9] B. N. Oreshkin, M. Jauhari, R. K. Selvam, M. Wolff, S. Ramasubramanian, K. G. Olivares, T. Konstantinova, A. Potapczynski, W. Pan, M. Cao, D. Efimov, M. W. Mahoney, A. G. Wilson. *Zero-shot Forecasting by Simulation Alone.* International Conference on Learning Representations (ICLR), 2026.

[10] A. Yu, D. C. Maddix, B. Han, X. Zhang, A. F. Ansari, O. Shchur, C. Faloutsos, A. G. Wilson, M. W. Mahoney, B. Wang. *Understanding Transformers for Time Series: Rank Structure, Flow-of-ranks, and Compressibility.* International Conference on Learning Representations (ICLR), 2026.

[11] A. Yu, D. C. Maddix, B. Han, X. Zhang, A. F. Ansari, O. Shchur, C. Faloutsos, A. G. Wilson, M. W. Mahoney, B. Wang. *Understanding the Implicit Biases of Design Choices for Time Series Foundation Models.* International Conference on Learning Representations (ICLR), 2026.

[12] N. A. Chandra, Y. L. Li, A. N. Amin, A. Ali, J. Rollins, S. W. Ober, A. Raghu, A. G. Wilson. *A Unification of Discrete, Gaussian, and Simplicial Diffusion.* International Conference on Learning Representations (ICLR), 2026.

[13] E. Baron, A. Amin, R. Weitzman, S. d'Oelsnitz, D. S. Marks, A. G. Wilson. *Shrinking Proteins with Diffusion.* International Conference on Learning Representations (ICLR), 2026.

[14] A. Potapczynski, A. Ali, A. G. Wilson. *Understanding and Relaxing the Limitations of Transformers for Linear Algebra.* International Conference on Learning Representations (ICLR), 2026.

[15] Y. Zheng, Z. Li, S. Fan, A.G. Wilson, K Zhang. *Diverse Dictionary Learning.* International Conference on Learning Representations (ICLR), 2026.

[16] A. Amin, N. Gruver, A.G. Wilson. *Why Masking Diffusion Works: Condition on the Jump Schedule for Improved Discrete Diffusion.* Advances in Neural Information Processing Systems (NeurIPS), 2025.

[17] M. Marek, S. Lotfi, A. Somasundaram, A. G. Wilson, M. Goldblum *Small Batch Size Training for Language Models: When Vanilla SGD Works, and Why Gradient Accumulation is Wasteful.* Advances in Neural Information Processing Systems (NeurIPS), 2025.

[18] S. Qiu, Z. Chen, H. Phan, Q. Lei, A.G. Wilson. *How to Scale Second-Order Optimization.* Advances in Neural Information Processing Systems (NeurIPS), 2025.

[19] A.G. Wilson. *Deep Learning is Not So Mysterious or Different.* International Conference on Machine Learning (ICML), 2025. **Spotlight**.

[20] Y. L. Li, D. Lu, P. Kirichenko, S. Qiu, T. G. J. Rudner, C. B. Bruss, A. G. Wilson. *Out-of-Distribution Detection Methods Answer the Wrong Questions.* International Conference on Machine Learning (ICML), 2025.

[21] S. Qiu, L. Xiao, A. G. Wilson, J. Pennington, A. Agarwala. *Scaling Collapse Reveals Universal Dynamics in Compute-Optimally Trained Neural Networks.* International Conference on Machine Learning (ICML), 2025. **Oral presentation.**

[22] L. Masserano, A. F. Ansari, B. Han, X. Zhang, C. Faloutsos, M. W. Mahoney, A. G. Wilson, Y. Park, S. Rangapuram, D. C. Maddix, B. Wang. *Enhancing Foundation Models for Time Series Forecasting via Wavelet-based Tokenization.* International Conference on Machine Learning (ICML), 2025.

[23] A. Amin, A. Potapczynski, A. G. Wilson. *Training Flexible Models of Genetic Variant Effects from Functional Annotations using Accelerated Linear Algebra.* International Conference on Machine Learning (ICML), 2025. **Paper Award at ICLR Genetics Workshop**.

[24] Y. Kuang, N. Amsel, S. Lotfi, S. Qiu, A. Potapczynski, A. G. Wilson. *Customizing the Inductive Biases of Softmax Attention using Structured Matrices.* International Conference on Machine Learning (ICML), 2025.

[25] Y. Liu, M. Jovanovic, K. Mallayya, W. J. Maddox, A. G. Wilson, S. Klemenz, L. M. Schoop, E. A. Kim. *Materials expert-artificial intelligence for materials discovery.* Nature Communications Materials, 2025.

[26] B. Burkhart, T. Dharmawardena, et al. *A nearby dark molecular cloud in the Local Bubble revealed via H2 fluorescence.* Nature Astronomy, 2025.

[27] M. Finzi, S. Kapoor, D. Granziol, A. Gu, C. De Sa, Z. Kolter, A. G. Wilson. *Compute-Optimal LLMs Provably Generalize Better with Scale.* International Conference on Learning Representations (ICLR), 2025.

[28] A. Amin, N. Gruver, Y. Kuang, Y. L. Li, H. Elliott, C. McCarter, A. Raghu, P. Greenside, A. G. Wilson. *Bayesian Optimization of Antibodies Informed by a Generative Model of Evolving Sequences.* International Conference on Learning Representations (ICLR), 2025.

[29] T. G. J. Rudner, X. Pan, Y. L. Li, R. Shwartz-Ziv, A. G. Wilson. *Fine-Tuning with Uncertainty-Aware Priors Makes Vision and Language Foundation Models More Reliable.* Artificial Intelligence and Statistics (AISTATS), 2025.

[30] A. Potapczynski, S. Qiu, M. A. Finzi, C. Ferri, Z. Chen, M. Goldblum, C. B. Bruss, C. De Sa, A. G. Wilson. *Searching for Efficient Linear Layers over a Continuous Space of Structured Matrices.* Advances in Neural Information Processing Systems (NeurIPS), 2024.

[31] S. Lotfi, Y. Kuang, M. A. Finzi, B. Amos, M. Goldblum, A. G. Wilson. *Unlocking Tokens as Data Points for Generalization Bounds on Larger Language Models.* Advances in Neural Information Processing Systems (NeurIPS), 2024. Best Paper Award, ICML Theoretical Foundations Workshop).

[32] S. Kapoor, N. Gruver, M. Roberts, A. Pal, S. Dooley, K. M. Collins, U. Bhatt, A. Weller, M. Goldblum, A. G. Wilson. *Large Language Models Must Be Taught to Know What They Don't Know.* Advances in Neural Information Processing Systems (NeurIPS), 2024.

[33] A. Amin, A.G. Wilson. *Scalable and Flexible Causal Discovery with an Efficient Test for Adjacency.* International Conference on Machine Learning (ICML), 2024.

[34] S. Lotfi, M. Finzi, Y. Kuang, T. Rudner, M. Goldblum, A. G. Wilson. *Non-Vacuous Generalization Bounds for Large Language Models.* International Conference on Machine Learning (ICML), 2024.

[35] M. Goldblum, M. Finzi, K. Rowan, A. G. Wilson. *The No Free Lunch Theorem, Kolmogorov Complexity, and the Role of Inductive Biases in Machine Learning.* International Conference on Machine Learning (ICML), 2024. **Spotlight.**

[36] T. Papamarkou *et. al. Bayesian Deep Learning in the Age of Large-Scale AI.* International Conference on Machine Learning (ICML), 2024.

[37] H. Phan, A.G. Wilson, Q. Lei. *Controllable Prompt Tuning For Balancing Group Distributional Robustness.* International Conference on Machine Learning (ICML), 2024.

[38] S. Qiu, B. Han, D. Maddix, S. Zhang, B. Wang, A.G. Wilson. *Transferring Knowledge From Large Foundation Models to Small Downstream Models.* International Conference on Machine Learning (ICML), 2024.

[39] S. Qiu, A. Potapczynski, M. Finzi, A.G. Wilson. *Compute Better Spent: Replacing Dense Layers with Structured Matrices.* International Conference on Machine Learning (ICML), 2024.

[40] S. Lavoie, P. Kirichenko, M. Ibrahim, M. Assran, A. G. Wilson, A. Courville, N. Ballas. *Modeling Caption Diversity in Contrastive Visual Language Pretraining.* International Conference on Machine Learning (ICML), 2024.

[41] Y. Li, T. Rudner, A. G. Wilson. *A Study of Bayesian Neural Network Surrogates for Bayesian Optimization.* International Conference on Learning Representations (ICLR), 2024.

[42] N. Gruver, A. Sriram, A. Madotto, A. G. Wilson, C. L. Zitnick, Z. W. Ulissi. *Fine-Tuned Language Models Generate Stable Inorganic Materials as Text.* International Conference on Learning Representations (ICLR), 2024.

[43] Y. Zhang, T. Rudner, A.G. Wilson, J. Kempe. *Mind the GAP: Improving Robustness to Subpopulation Shifts with Group-Aware Priors.* Artificial Intelligence and Statistics (AISTATS), 2024. **Oral Presentation**.

[44] G. Detomasso, A. Gasparin, M. Donini, M. Seeger, A.G. Wilson, C. Archambeau. *Fortuna: A Library for Uncertainty Quantification in Deep Learning.* Journal of Machine Learning Research (JMLR), 2024.

[45] A. Potapczynski, M. Finzi, G. Pleiss, A. G. Wilson. *Exploiting Compositional Structure for Automatic and Efficient Numerical Linear Algebra.* Advances in Neural Information Processing Systems (NeurIPS), 2023.

[46] N. Gruver, M. Finzi, S. Qiu, A.G. Wilson. *Large Language Models Are Zero Shot Time Series Forecasters.* Advances in Neural Information Processing Systems (NeurIPS), 2023.

[47] P. Kirichenko, M. Ibrahim, R. Balestriero, D. Bouchacourt, S. R. Vedantam, H. Firooz, A.G. Wilson. *Understanding the detrimental class-level effects of data augmentation.* Advances in Neural Information Processing Systems (NeurIPS), 2023.

[48] R. Shwartz-Ziv, M. Goldblum, Y. L. Li, C. B. Bruss, A. G. Wilson. *Simplifying Neural Network Training Under Class Imbalance.* Advances in Neural Information Processing Systems (NeurIPS), 2023.

[49] T. G. J. Rudner, S. Kapoor, S. Qiu, A. G. Wilson. *Should We Learn Most Likely Functions or Parameters?* Advances in Neural Information Processing Systems (NeurIPS), 2023.

[50] N. Gruver, S. D. Stanton, N. C. Frey, T. G. J. Rudner, I. Hotzel, J. Lafrance-Vanasse, A. Rajpal, K. Cho, A. G. Wilson. *Protein Design with Guided Discrete Diffusion.* Advances in Neural Information Processing Systems (NeurIPS), 2023.

[51] M. Goldblum, H. Souri, R. Ni, M. Shu, V. U. Prabhu, G. Somepalli, P. Chattopadhyay, A. Bardes, M. Ibrahim, J. Hoffman, R. Chellappa, A. G. Wilson, T. Goldstein. *Battle of the Backbones: A Large-Scale Comparison of Pretrained Models across Computer Vision Tasks.* Advances in Neural Information Processing Systems (NeurIPS), 2023.

[52] V. Cherepanova, G. Somepalli, J. Geiping, C. B. Bruss, A. G. Wilson, T. Goldstein, M. Goldblum. *A Performance-Driven Benchmark for Feature Selection in Tabular Deep Learning.* Advances in Neural Information Processing Systems (NeurIPS), 2023.

[53] Y. Wang, T. G. J. Rudner, A.G. Wilson. *Visual Explanations of Image-Text Representations via Multi-Modal Information Bottleneck Attribution.* Advances in Neural Information Processing Systems (NeurIPS), 2023.

[54] R. Ali, X. Shi, K. Lin, A. Zhang, A.G. Wilson. *Automated Few-Shot Classification with Instruction-Finetuned Language Models.* Findings of Empirical Methods in Natural Language Processing (EMNLP), 2023.

[55] S. Qiu, A. Potapczynski, P. Izmailov, A.G. Wilson. *Simple and Fast Group Robustness by Automatic Feature Reweighting*. International Conference on Machine Learning (ICML), 2023.

[56] S. Lotfi, P. Izmailov, G. Benton, M. Goldblum, A.G. Wilson. *Bayesian Model Selection, the Marginal Likelihood, and Generalization*. Extended version. Award winning papers track. Journal of Machine Learning Research (JMLR), 2023.

[57] M.A. Finzi, A. Boral, L. Zepeda-Nunez, A.G. Wilson, F. Sha. *User-defined Event Sampling and Uncertainty Quantification in Diffusion Models for Physical Dynamical Systems*. International Conference on Machine Learning (ICML), 2023.

[58] T.G.J. Rudner, S. Kapoor, S. Qiu, A.G. Wilson. *Function-Space Regularization in Neural Networks*. International Conference on Machine Learning (ICML), 2023.

[59] S. Stanton, W. Maddox, A.G. Wilson. *Bayesian Optimization with Conformal Prediction Sets*. Artificial Intelligence and Statistics (AISTATS), 2023.

[60] N. Gruver, M. Finzi, M. Goldblum, A.G. Wilson. *The Lie Derivative for Measuring Learned Equivariance*. International Conference on Learning Representations (ICLR), 2023. **Top 5% notable paper.**

[61] M. Finzi, A. Potapczynski, M. Choptuik, A.G. Wilson. *A Stable and Scalable Method for Solving Initial Value PDEs with Neural Networks*. International Conference on Learning Representations (ICLR), 2023.

[62] J. Geiping, M. Goldblum, G. Somepalli, R. Shwartz-Ziv, T. Goldstein, A.G. Wilson. *How Much Data Are Augmentations Worth? An Investigation into Scaling Laws, Invariance, and Implicit Regularization*. International Conference on Learning Representations (ICLR), 2023.

[63] P. Kirichenko, P. Izmailov, A.G. Wilson. *Last Layer Re-Training is Sufficient for Robustness to Spurious Correlations*. International Conference on Learning Representations (ICLR), 2023. **Top 25% notable paper.**

[64] R. Levin, V. Cherepanova, A. Schwarzschild, A. Bansal, C. B. Bruss, T. Goldstein, A.G. Wilson, M. Goldblum. *Transfer Learning with Deep Tabular Models*. International Conference on Learning Representations (ICLR), 2023.

[65] Z. Liu, Z. Tang, X. Shi, A. Zhang, M. Li, A. Shrivastava, A.G. Wilson. *Learning Multimodal Data Augmentation in Feature Space*. International Conference on Learning Representations (ICLR), 2023.

[66] S. Lotfi, M. Finzi, S. Kapoor, A. Potapczynski, M. Goldblum, A.G. Wilson. *PAC-Bayes Compression Bounds So Tight That They Can Explain Generalization*. Advances in Neural Information Processing Systems (NeurIPS), 2022.

[67] R. Shwartz-Ziv, M. Goldblum, H. Souri, S. Kapoor, C. Zhu, Y. LeCun, A.G. Wilson. *Pre-Train Your Loss: Easy Bayesian Transfer Learning with Informative Priors*. Advances in Neural Information Processing Systems (NeurIPS), 2022.

[68] P. Izmailov, P. Kirichenko, N. Gruver, A.G. Wilson. *On Feature Learning in the Presence of Spurious Correlations*. Advances in Neural Information Processing Systems (NeurIPS), 2022.

[69] S. Kapoor, W. Maddox, P. Izmailov, A.G. Wilson *On Uncertainty, Tempering, and Data Augmentation in Bayesian Classification*. Advances in Neural Information Processing Systems (NeurIPS), 2022.

[70] W. Yang, P. Kirichenko, M. Goldblum, A.G. Wilson. *Chroma-VAE: Mitigating Shortcut Learning with Generative Classifiers*. Neural Information Processing Systems (NeurIPS), 2022.

[71] S. Lotfi, P. Izmailov, G. Benton, M. Goldblum, A.G. Wilson. *Bayesian Model Selection, the Marginal Likelihood, and Generalization*. International Conference on Machine Learning (ICML), 2022. **Outstanding Paper Award**.

[72] S. Stanton, W. Maddox, N. Gruver, P. Maffettone, E. Delaney, P. Greenside, A.G. Wilson. *Accelerating Bayesian Optimization for Biological Sequence Design with Denoising Autoencoders*. International Conference on Machine Learning (ICML), 2022.

[73] R. Zhang, C. De Sa, A.G. Wilson. *Low-Precision Stochastic Gradient Langevin Dynamics*. International Conference on Machine Learning (ICML), 2022.

[74] G. Benton, W. Maddox, A.G. Wilson. *Volatility Based Kernels and Moving Average Means for Accurate Forecasting with Gaussian Processes*. International Conference on Machine Learning (ICML), 2022.

[75] W. Maddox, A. Potapczynski, A.G. Wilson. *Low Precision Arithmetic for Fast Gaussian Processes*. Uncertainty in Artificial Intelligence (UAI), 2022.

[76] J. Venderley et. al. *Harnessing Interpretable and Unsupervised Machine Learning to Address Big Data from Modern X-ray Diffraction*. Proceedings of the National Academy of Sciences (PNAS), 2022.

[77] N. Gruver, M. Finzi, A.G. Wilson. *Deconstructing the Inductive Biases of Hamiltonian Neural Networks*. International Conference on Learning Representations (ICLR), 2022.

[78] A.G. Wilson *et. al. Evaluating Approximate Inference in Bayesian Deep Learning*. Journal of Machine Learning Research (JMLR), 2022.

[79] P. Izmailov, P. Nicholson, S. Lotfi, A.G. Wilson. *Dangers of Bayesian Model Averaging under Covariate Shift*. Advances in Neural Information Processing Systems (NeurIPS), 2021.

[80] S. Stanton, P. Izmailov, P. Kirichenko, A. Alemi, A.G. Wilson. *Does Knowledge Distillation Really Work?* Advances in Neural Information Processing Systems (NeurIPS), 2021.

[81] M. Finzi, G. Benton, A.G. Wilson. *Residual Pathway Priors for Soft Equivariance Constraints*. Advances in Neural Information Processing Systems (NeurIPS), 2021.

[82] W. Maddox, M. Balandat, A.G. Wilson, E. Bakshy. *Bayesian Optimization with High-Dimensional Outputs*. Advances in Neural Information Processing Systems (NeurIPS), 2021.

[83] W. Maddox, S. Stanton, A.G. Wilson. *Conditioning Sparse Variational Gaussian Processes for Online Decision-making*. Advances in Neural Information Processing Systems (NeurIPS), 2021.

[84] P. Izmailov, S. Vikram, M. Hoffman, A.G. Wilson. What Are Bayesian Neural Network Posteriors Really Like? *International Conference on Machine Learning* (ICML), 2021. **Long oral presentation** (≈1% accept rate).

[85] M. Finzi, M. Welling, A.G. Wilson. A Practical Method for Constructing Equivariant Multilayer Perceptrons for Arbitrary Matrix Groups. *International Conference on Machine Learning* (ICML), 2021. **Long oral presentation** (≈1% accept rate).

[86] S. Kapoor, M. Finzi, A. Wang, A.G. Wilson. SKIing on Simplices: Kernel Interpolation on the Permutohedral Lattice for Scalable Gaussian Processes. *International Conference on Machine Learning* (ICML), 2021. **Long oral presentation** (≈1% accept rate).

[87] G. Benton, W. Maddox, S. Lotfi, A.G. Wilson. Loss Surface Simplexes for Mode Connecting Volumes and Fast Ensembling. *International Conference on Machine Learning* (ICML), 2021.

[88] S. Sun, J. Shi, A.G. Wilson, R. Grosse. Scalable Variational Gaussian Processes via Harmonic Kernel Decomposition. *International Conference on Machine Learning* (ICML), 2021.

[89] B. Amos, S. Stanton, D. Yarats, A.G. Wilson. On the model-based stochastic value gradient for continuous reinforcement learning. *Learning for Dynamics and Control* (L4DC), 2021. **Oral presentation** (≈1% accept rate).

[90] S. Stanton, W. Maddox, I. Delbridge, A.G. Wilson. Kernel Interpolation for Scalable Online Gaussian Processes. *Artificial Intelligence and Statistics* (AISTATS), 2021.

[91] W. Maddox, S. Tang, P. Moreno, A. Damianou, A.G. Wilson. Fast Adaptation with Linearized Neural Networks. *Artificial Intelligence and Statistics* (AISTATS), 2021.

[92] A.G. Wilson, P. Izmailov. Bayesian Deep Learning and a Probabilistic Perspective of Generalization. *Advances in Neural Information Processing Systems* (NeurIPS), 2020.

[93] M. Balandat, B. Karrer, D. Jiang, S. Daulton, B. Letham, A.G. Wilson, E. Bakshy. BoTorch: An Efficient Differentiable Monte-Carlo Framework for Bayesian Optimization. *Advances in Neural Information Processing Systems* (NeurIPS), 2020.

[94] P. Kirichenko, P. Izmailov, A.G. Wilson. Why Normalizing Flows Fail to Detect Out-of-Distribution Data. *Advances in Neural Information Processing Systems* (NeurIPS), 2020.

[95] G. Benton, M. Finzi, P. Izmailov, A.G. Wilson. Learning Invariances in Neural Networks. *Advances in Neural Information Processing Systems* (NeurIPS), 2020.

[96] M. Finzi, A. Wang, A.G. Wilson. Simplifying Hamiltonian and Lagrangian Neural Networks via Explicit Constraints. *Advances in Neural Information Processing Systems* (NeurIPS), 2020.

[97] Y. Wu, P. Zhou, A.G. Wilson, E.P. Xing, Z. Hu. Improving GAN Training with Probability Ratio Clipping and Sample Reweighting. *Advances in Neural Information Processing Systems* (NeurIPS), 2020.

[98] M. Finzi, S. Stanton, P. Izmailov, A.G. Wilson. Generalizing Convolutional Networks for Equivariance to Lie Groups on Arbitrary Continuous Data. *International Conference on Machine Learning* (ICML), 2020.

[99] I. Delbridge, D. Bindel, A.G. Wilson. Randomly Projected Additive Gaussian Processes for Regression. *International Conference on Machine Learning* (ICML), 2020.

[100] P. Izmailov, P. Kirichenko, M. Finzi, A.G. Wilson. Semi-Supervised learning with Normalizing Flows. *International Conference on Machine Learning* (ICML), 2020.

[101] R. Zhang, C. Li, C. Chen, A.G. Wilson. Cyclical Stochastic Gradient MCMC for Bayesian Deep Learning. *International Conference on Learning Representations* (ICLR), 2020. **Oral presentation**.

[102] W. Maddox, P. Izmailov, T. Garipov, D. Vetrov, A.G. Wilson. A Simple Baseline for Bayesian Uncertainty in Deep Learning. *Advances in Neural Information Processing Systems* (NeurIPS), 2019.

[103] G. Benton, J. Salkey, W. Maddox, J. Albinati, A.G. Wilson. Function-Space Distributions over Kernels. *Advances in Neural Information Processing Systems* (NeurIPS), 2019.

[104] K. A. Wang, G. Pleiss, J. Gardner, S. Tyree, K. Weinberger, A.G. Wilson. Exact Gaussian Processes on a Million Data Points. *Advances in Neural Information Processing Systems* (NeurIPS), 2019.

[105] P. Izmailov, W. Maddox, P. Kirichenko, T. Garipov, D. Vetrov, A.G. Wilson. Subspace Inference for Bayesian Deep Learning. *Uncertainty In Artificial Intelligence* (UAI), 2019.

[106] J. Wu, S. Toscano-Palmerin, P. I. Frazier, A. G. Wilson. Practical Multi-fidelity Bayesian Optimization for Hyperparameter Tuning. *Uncertainty in Artificial Intelligence* (UAI), 2019.

[107] G. Yang, T. Chen, P. Kirichenko, J. Bai, A.G. Wilson, C. de Sa. SWALP: Stochastic Weight Averaging in Low Precision Training. *International Conference on Machine Learning* (ICML), 2019.

[108] C. Guo, J. Gardner, Y. You, A.G. Wilson, K.Q. Weinberger. Simple Black-box Adversarial Attacks. *International Conference on Machine Learning* (ICML), 2019.

[109] W. Herlands, D.B. Neill, H. Nickisch, A.G. Wilson. Change Surfaces for Expressive Multidimensional Changepoints and Counterfactual Prediction. *Journal of Machine Learning Research* (JMLR), 2019.

[110] B. Athiwaratkun, M. Finzi, P. Izmailov, A.G. Wilson. There are Many Consistent Explanations of Unlabeled Data: Why You Should Average. *International Conference on Learning Representations* (ICLR), 2019.

[111] T. Garipov*, P. Izmailov*, D. Podoprikhin*, D. Vetrov, A.G. Wilson. Loss Surfaces, Mode Connectivity, and Fast Ensembling of DNNs. *Advances in Neural Information Processing Systems* (NeurIPS), 2018. **Spotlight**.

[112] J. Gardner, G. Pleiss, D. Bindel, K. Weinberger, A.G. Wilson. GPyTorch: Blackbox Matrix-Matrix Gaussian Process Inference with GPU Acceleration. *Advances in Neural Information Processing Systems* (NeurIPS), 2018. **Spotlight**.

[113] D. Eriksson, K. Dong, E. Lee, D. Bindel, A.G. Wilson. Scaling Gaussian Process Regression with Derivatives. *Advances in Neural Information Processing Systems* (NeurIPS), 2018.

[114] P. Izmailov*, D. Podoprikhin*, T. Garipov*, D. Vetrov, A.G. Wilson. Averaging Weights Leads to Wider Optima and Better Generalization, *Uncertainty in Artificial Intelligence* (UAI), 2018. **Oral presentation**.

[115] G. Pleiss, J. Gardner, K.Q. Weinberger, and A.G. Wilson. Constant time predictive distributions for Gaussian processes. *International Conference on Machine Learning* (ICML), 2018.

[116] W. Herlands, E. McFowland III, A.G. Wilson, and D.B. Neill. Automated Local Regression Discontinuity Design Discovery. *Knowledge Discovery and Data Mining* (KDD), 2018.

[117] B. Athiwaratkun, A.G. Wilson, and A. Anandkumar. Probabilistic FastText. *Association for Computational Linguistics* (ACL), 2018. **Oral presentation**.

[118] B. Athiwaratkun and A.G. Wilson. Hierarchical Density Order Embeddings. *International Conference on Learning Representations* (ICLR), 2018.

[119] J. Gardner, G. Pleiss, R. Wu, K.Q. Weinberger, and A.G. Wilson. Product Kernel Interpolation for Scalable Gaussian Processes. *Artificial Intelligence and Statistics* (AISTATS), 2018.

[120] W. Herlands, E. McFowland, A.G. Wilson, and D.B. Neill. Gaussian Process Subset Scanning for Anomalous Pattern Detection in Non-iid Data. *Artificial Intelligence and Statistics* (AISTATS), 2018.

[121] Y. Saatchi and A.G. Wilson. Bayesian GAN. *Neural Information Processing Systems* (NeurIPS), 2017. **Spotlight**.

[122] J. Wu, M. Poloczek, A.G. Wilson, and P. Frazier. Bayesian optimization with gradients. *Neural Information Processing Systems* (NeurIPS), 2017. **Oral presentation**.

[123] K. Dong, D. Eriksson, H. Nickisch, D. Bindel, and A.G. Wilson. Scalable log determinants for Gaussian process kernel learning. *Neural Information Processing Systems* (NeurIPS), 2017.

[124] A. Loeb, P. Jang, M. Davidow, and A.G. Wilson. Scalable Lévy process kernel learning. *Neural Information Processing Systems* (NeurIPS), 2017.

[125] B. Athiwaratkun and A.G. Wilson. Multimodal Word Distributions. *Association for Computational Linguistics* (ACL), 2017.

[126] M. Al-Shedivat, A.G. Wilson, Y. Saatchi, Z. Hu, and E.P. Xing. Learning Scalable Deep Kernels with Recurrent Structure. *Journal of Machine Learning Research* (JMLR), 2017.

[127] A.G. Wilson*, Z. Hu* (equal contribution), R. Salakhutdinov, and E.P. Xing. Stochastic Variational Deep Kernel Learning. *Neural Information Processing Systems* (NeurIPS), 2016.

[128] A.G. Wilson*, Z. Hu* (equal contribution), R. Salakhutdinov, and E.P. Xing. Deep kernel learning. *Artificial Intelligence and Statistics* (AISTATS), 2016.

[129] W. Herlands, A.G. Wilson, S. Flaxman, H. Nickisch, D.B. Neill, and E.P. Xing. Scalable Gaussian processes for characterizing multidimensional change surfaces. *Artificial Intelligence and Statistics* (AISTATS), 2016.

[130] J. Oliva*, A. Dubey* (equal contribution), A.G. Wilson, B. Poczos, J. Schneider, and E.P. Xing. Bayesian nonparametric kernel learning. *Artificial Intelligence and Statistics* (AISTATS), 2016.

[131] A.G. Wilson, C. Dann, C.G. Lucas, and E.P. Xing. The human kernel. In *Neural Information Processing Systems* (NeurIPS), 2015. **Spotlight**.

[132] A.G. Wilson and H. Nickisch. Kernel interpolation for scalable structured Gaussian processes (KISS-GP). *International Conference on Machine Learning* (ICML), 2015.

[133] S. Flaxman, A.G. Wilson, D.B. Neill, H. Nickisch, and A.J. Smola. Fast kronecker inference in Gaussian processes with non-Gaussian likelihoods. *International Conference on Machine Learning* (ICML), 2015.

[134] Z. Yang, A.J. Smola, L. Song, and A.G. Wilson. À la carte – learning fast kernels. *Artificial Intelligence and Statistics* (AISTATS), 2015. **Oral presentation**.

[135] A.G. Wilson*, E. Gilboa* (equal contribution), A. Nehorai, and J.P. Cunningham. Fast kernel learning for multidimensional pattern extrapolation. *Neural Information Processing Systems* (NeurIPS), 2014.

[136] Y. Wu, D.J. Holland, M.D., Mantle, A.G. Wilson, S. Nowozin, A. Blake, and L.F. Gladden. A Bayesian method to quantifying chemical composition using NMR: application to porous media systems. *European Signal Processing Conference* (EUSIPCO), 2014.

[137] A. Shah, A.G. Wilson, and Z. Ghahramani. Student-t processes as alternatives to Gaussian processes. *Artificial Intelligence and Statistics* (AISTATS), 2014.

[138] A.G. Wilson and R.P. Adams. Gaussian process kernels for pattern discovery and extrapolation. *International Conference on Machine Learning* (ICML), 2013. **Oral presentation**.

[139] A.G. Wilson and Z. Ghahramani. Modelling input dependent correlations between multiple responses. *European Conference on Machine Learning* (ECML), 2012. **Nectar Track** for "*significant machine learning results*". **Oral presentation**.

[140] A.G. Wilson, D.A. Knowles, and Z. Ghahramani. Gaussian process regression networks. *International Conference on Machine Learning* (ICML), 2012. **Oral presentation**.

[141] A.G. Wilson and Z. Ghahramani. Generalised Wishart processes. *Uncertainty in Artificial Intelligence* (UAI), 2011. **Best Student Paper Award**.

[142] A.G. Wilson and Z. Ghahramani. Copula processes. *Neural Information Processing Systems* (NeurIPS), 2010. **Spotlight**.

Reports

[143] A.G. Wilson. The Case for Bayesian Deep Learning. Technical report, NYU, 2019.

[144] A.G. Wilson. Covariance kernels for fast automatic pattern discovery and extrapolation with Gaussian processes. PhD Thesis, University of Cambridge. October 2014.

[145] A.G. Wilson. The change point kernel. Technical report, University of Cambridge. Nov 2013.

[146] A.G. Wilson. A process over all stationary covariance kernels. Technical report, University of Cambridge. June 2012.

[147] A.G. Wilson and D. Duvenaud. Learning Intrinsic Dimension with Probabilistic Methods. University of Cambridge, 2012.

[148] A.G. Wilson. Latent Gaussian process models. First year report, University of Cambridge. August 2010.

[149] A.G. Wilson. Position and energy reconstruction from scintillation light in a liquid xenon gamma ray detector designed for PET. Honours undergraduate thesis, UBC. May 2008.

Books

[150] *Probabilistic Machine Learning: Advanced Topics*. MIT Press, 2023. Co-authored chapters on Gaussian processes and Bayesian deep learning with Kevin P. Murphy.

[151] *Dive into Deep Learning*. `d2l.ai`. Cambridge University Press, 2023. Authored a chapter on Gaussian processes.

Service

- WiML Mentor, NeurIPS 2025
- WiML Mentor, ICML 2025
- WiML Mentor, NeurIPS 2024
- AI Faculty Search Committee, NYU, 2024-2025
- AI Tenure Committee, NYU, 2024-2025
- ICLR Mentor, 2024
- AI Faculty Search Committee, NYU, 2023-2024
- AI Tenure Committee, NYU, 2023-2024
- AI Director Search Committee, NYU, 2023-2024
- AI Faculty Search Committee, NYU, 2022-2023
- WiML PhD Admissions Panelist, 2023
- PhD Admissions Committee, Computer Science, NYU, 2022-2023
- PhD Admissions Committee, Data Science, NYU, 2022-2023
- Fellowship Committee, Computer Science, NYU, 2022-2023
- Tenure Committee, Computer Science, NYU, 2022-2023
- Fellowship Committee, Computer Science, NYU, 2021-2022
- AI Faculty Search Committee, NYU, 2021-2022
- Pathways to AI Mentor, NYU, 2022.
- CURP Mentor, NYU, 2021.
- PhD Admissions Committee, Computer Science, NYU, 2021-2022
- PhD Admissions Committee, Data Science, NYU, 2021-2022
- CDS Faculty Fellow Committee, NYU, 2021-2022
- PhD Admissions Committee, Computer Science, NYU, 2020-2021
- PhD Admissions Committee, Data Science, NYU, 2020-2021
- Fellowship Committee, Computer Science, NYU, 2020-2021
- AI Faculty Search Committee, NYU, 2020-2021
- CS Faculty Fellow Committee, NYU, 2020-2021
- CDS Faculty Fellow Committee, NYU, 2020-2021
- Mentor for CDS Undergraduate Research Program for URM, NYU, 2020-2021

- PhD Admissions Committee, Computer Science, NYU, 2019-2020
- PhD Admissions Committee, Data Science, NYU, 2019-2020
- Fellowship Committee, Computer Science, NYU, 2019-2020
- Faculty Search Committee, Computer Science, NYU, 2019-2020
- Faculty Search Committee (Biostatistics), CDS, NYU, 2019-2020
- CDS Faculty Fellow Committee, CDS, NYU, 2019-2020
- ORIE PhD Admissions Committee, Cornell, 2018-2019
- ORIE PhD Admissions Committee, Cornell, 2017-2018

Code Repositories

- `https://cims.nyu.edu/~andrewgw/code`. Main resource page for code repositories from myself and collaborators. Includes the CoLA (fast matrix operations), GPyTorch (Gaussian processes), BoTorch (Bayesian optimization), Fortuna (uncertainty representation), EMLP (equivariance symmetries), Bayesian GAN, and probabilistic word embedding libraries, as well as resources for kernel learning, Hamiltonian neural networks, and invariance learning.

Selected Talks

- Berkeley AI for Science — July 2026
- Riken AIP — June 2026
- MLSS — June 2026
- Chai Workshop, Berkeley — May 2026
- Cornell Seminar — April 2026
- Tower Capital — March 2026
- Renaissance Technologies — March 2026
- MIT, Institute for AI and Fundamental Interactions — February 2026
- Riken AIP — February 2026
- Keynote Speaker, MLDS and DEDS (Kyoto, Japan) — February 2026
- Oxford Mathematics Department — January 2026
- AI Security Institute (AISI) — January 2026
- Keynote Speaker, SPIGM Workshop (NeurIPS) — December 2025
- MBZUAI Workshop — October 2025
- FLaNN Seminar — September 2025
- Keynote Speaker, AutoML Conference — September 2025
- Keynote Speaker, ICML Foundation Models for Structured Data — July 2025
- Isaac Newton Institute, Uncertainty Representation — June 2025
- Cold Spring Harbor Laboratory (CSHL) Seminar — June 2025
- AI Forum, Boston — June 2025
- Princeton University, CSML Seminar — May 2025
- Columbia University, Foundations Seminar — April 2025
- Flatiron Learning on Graphs (LoG) Workshop — April 2025
- UAI ColorAI Workshop — March 2025
- Georgia Tech, ML Seminar Series — February 2025
- ELLIS Seminar Speaker, Cambridge UK — February 2025
- NeurIPS Time Series Workshop — December 2024
- CCM Flatiron Institute Colloquium — October 2024
- Dagstuhl on Bayesian Deep Learning — October 2024
- Simons Workshop on Transformers as a Computational Model — September 2024
- Inference and Theory for Bayesian Neural Networks, JSM Meeting — August 2024

- Meta Adapative Experimentation Workshop June 2024
- Keynote for Workshop on Distribution Shifts, ICLR May 2024
- Keynote for How Far Are We from AGI Workshop, ICLR May 2024
- Distinguished Speaker Series, Boston University March 2024
- Flatiron Institute Seminar February 2024
- Distinguished Seminar Speaker, Michigan State University February 2024
- Generative Models and Uncertainty Quantification Copenhagen, September 2023
- Prescriptive Foundations for Model Construction Copenhagen, September 2023
- AI for Materials Engineering Arlington, VA, June 2023
- Meta Adaptive Experimentation Workshop NYC, April 2023
  Bayesian Neural Network Surrogates for Bayesian Optimization
- UQSay Seminar Paris, France, March 2023
  Bayesian Model Selection
- NeurIPS "I can't believe it's not better!" Workshop New Orleans, December 2022
  When Bayesian Orthodoxy Can Go Wrong: Model Selection and Out-of-Distribution Generalization
- ASM Workshop on Uncertainty in Deep Learning Washington DC, August 2022
  Promises and Pitfalls of Bayesian Deep Learning.
- Summer School on Optimization, Big Data, and Analysis. Veroli, Italy, July 2022
  The Foundations of Bayesian Model Construction
- ISBA World Meeting Montreal, Canada, July 2022
  Promises and Challenges of Bayesian Inference in Deep Learning
- Amazon Web Services Research Seminar Berlin, Germany, April 2022
  Myths and Reality in Bayesian Deep Learning
- Dagstuhl Seminar on Tractable Probabilistic Inference Schloss Dagstuhl, April 2022
  Exploiting Algebraic Structure for Scalable Inference
- Memorial Sloan Kettering Cancer Center. NYC, March 2022
  How do we build models that learn and generalize?
- SIAM Uncertainty Quantification Atlanta, April 2022
  How do we build models that learn and generalize?
- Workshop on Scalable Gaussian Processes Berne, Switzerland, January 2022
  Numerical Methods for Scalable Gaussian Processes
- Quantum Materials and Machine Learning Summer School Online, June 2021
  Probabilistic Machine Learning
- ML×Physics Panel Debate on Bayesian Neural Networks Online, May 2021
  We Should All Be Using Bayesian Inference with Neural Networks
- Flatiron Institute, Computational Methods and Data Science Seminar Online, April 2021
  Prescriptions for Probabilistic Model Construction
- University of Cambridge Online, April 2021
  Examining Critiques in Bayesian Deep Learning
- G-Research Seminar Online, February 2021
  Bayesian Deep Learning
- SecondMind Online, January 2021
  Gaussian Priors are Good and Deep Ensembles are Bayesian Model Averaging
- MLSE Online, December 2020
  Learning Symmetries
- Machine Learning in Science and Engineering (Columbia University) Online, December 2020
  How Do We Build Models that Learn and Generalize?
- NYU AI Winter School (for URM) Online, December 2020
  Introduction to Bayesian Machine Learning
- University of Oxford Online, November 2020
  Challenging Conventional Wisdom in Bayesian Deep Learning

- Toyota Technology Institute — Online, November 2020
  Bayesian Deep Learning and Prescriptions for Good Generalization
- Gaussian Process Summer School — Online, September 2020
  Representation Learning with Gaussian Processes
- Smiles Skoltech Summer School — Online, August 2020
  Bayesian Deep Learning
- ICML Bayesian Deep Learning Tutorial — Online, July 2020
  Bayesian Deep Learning and Probabilistic Model Construction
- Google Brain Research Seminar — Online, May 2020
  Loss Surface Geometry for Practical Bayesian Deep Learning
- BIRS Workshop — Banff, Canada, February 2020
  How Do We Build Models that Learn and Generalize?
- Flatiron Institute, Computational Quantum Physics Seminar — NYC, January 2020
  Understanding Generalization with Probability
- NeurIPS 2019 Bayesian Deep Learning Workshop — Vancouver, Canada, December 2019
  Exploiting Loss Surface Geometry for Practical Bayesian Deep Learning
- Binghamton University (SUNY) Dean's Speaker Series — Binghamton, NY, November 2019
  How Do We Build Models that Learn and Generalize?
- MIT Broad Institute — Cambridge, USA, October 2019
  Understanding Loss Landscapes for Bayesian Deep Learning
- Facebook Data for Good Seminar — New York, NY, September 2019
  Active Learning for Malaria Modelling
- Los Alamos National Laboratory — Los Alamos, USA, April 2019
  Exploiting Hardware Design for Scalable Gaussian Processes
- New York University — NY, USA, February 2019
  Uncertainty, Loss Valleys, and Generalization in Deep Learning
- University of Maryland — College Park, USA, February 2019
  Uncertainty, Loss Valleys, and Generalization in Deep Learning
- University of Michigan — Ann Arbor, USA, February 2019
  Bayesian Deep Learning and Probabilistic Model Construction
- UNC Chapel Hill — Chapel Hill, USA, February 2019
  Scalable Inference for High Dimensional Models
- MIT Seminar — Cambridge, USA, November 2018
  Uncertainty, Loss Valleys, and Generalization in Deep Learning
- Boston University Seminar — Cambridge, USA, November 2018
  Bayesian Deep Learning and Probabilistic Perspectives of Generalization
- Allerton Conference — Allerton, IL, October 2018
  Loss Valleys and Generalization in Deep Learning
- PyTorch DevCon — San Francisco, USA, October 2018
  GPyTorch: Scalable Numerical Linear Algebra in PyTorch
- Precision Medicine and Machine Learning — Durham, NC, August 2018
  Bayesian Deep Generative Models
- Deep Learning Summer School — Toronto, July 2018
  Bayesian Neural Networks
- SIAM ALA (Applied Linear Algebra) — Hong Kong, May 2018
  Krylov Subspace Methods for Scalable Gaussian Processes
- DALI 2018 — Canary Islands, April 2018
  Loss Landscapes and Optimization in Deep Learning
- BIRS Workshop (Stats & ML) — Banff, Canada, January 2018
  Bayesian Generative Adversarial Networks
- UCL Gatsby — London, UK, December 2017
  Bayesian Deep Generative Models

- University of Cambridge — Cambridge, UK, December 2017
  Bayesian Deep Generative Models
- Microsoft Research — Cambridge, UK, December 2017
  Bayesian Generative Adversarial Networks
- CMStatistics — London, UK, December 2017
  Stochastic MCMC in Bayesian Deep Learning
- AI Seminar, Cornell — Ithaca, NY, October 2017
  Loss Landscapes and Generalization in Deep Learning
- Statistics Seminar, Cornell — Ithaca, NY, September 2017
  Scalable Gaussian Processes for Scientific Discovery
- Linköping University — Linköping, Sweden, April 2017
  Scalable Deep Kernel Learning
- UCLA — Los Angeles, USA, January 2017
  Deep Kernel Learning
- University of British Columbia — Vancouver, Canada, March 2016
  Scalable Gaussian Processes for Scientific Discovery
- University of Edinburgh — Edinburgh, UK, March 2016
  Scalable Gaussian Processes for Scientific Discovery
- University of Southern California — Los Angeles, USA, March 2016
  Scalable Gaussian Processes for Scientific Discovery
- University of California, Irvine — Irvine, USA, March 2016
  Scalable Gaussian Processes for Scientific Discovery
- UCLA — Los Angeles, USA, March 2016
  Scalable Gaussian Processes for Scientific Discovery
- University of Massachusetts — Amherst, USA, March 2016
  Scalable Gaussian Processes for Scientific Discovery
- Cornell University — Ithaca, USA, March 2016
  Scalable Gaussian Processes for Scientific Discovery
- University of Toronto — Toronto, Canada, February 2016
  Scalable Gaussian Processes for Scientific Discovery
- Dartmouth College — Hanover, USA, February 2016
  Scalable Gaussian Processes for Scientific Discovery
- EPFL — Lausanne, Switzerland, February 2016
  Scalable Gaussian Processes for Scientific Discovery
- University of Waterloo — Waterloo, Canada, January 2016
  Scalable Gaussian Processes for Scientific Discovery
- University of Cambridge — Cambridge, UK, August 2015
  Kernel Interpolation for Scalable Gaussian Processes
- International Conference on Machine Learning — Lille, France, July 2015
  Kernel Interpolation for Scalable Structured Gaussian Processes
- New York University — NYC, USA, June 2015
  New Directions for Scalable Gaussian Processes using Numerical Methods
- Neural Information Processing Systems Workshop — Montreal, Canada, December 2014
  Kernel Methods for Representation Learning
- Oxford University — Oxford, UK, November 2014
  Kernel Methods for Representation Learning
- University College London — London, UK, November 2014
  Kernel Methods for Representation Learning
- Machine Learning Summer School (MLSS) — Pittsburgh, USA, July 2014
  Gaussian processes, Bayesian model selection, and kernel methods
- International Conference on Machine Learning — Atlanta, USA, June 2013
  Gaussian process kernels for pattern discovery and extrapolation

- Xerox Research Seminar — Grenoble, France, November 2012
  Gaussian process neural networks
- ECML Nectar Track — Bristol, UK, September 2012
  Input dependent correlations between many responses
- Microsoft Research — Cambridge, UK, September 2012
  Spectral Mixture Kernels for Extrapolation
- International Conference on Machine Learning — Edinburgh, UK, June 2012
  Gaussian Process Regression Networks
- University of California, Berkeley — Berkeley, USA, May 2012
  Generalised Wishart Processes
- Harvard University — Cambridge, USA, April 2012
  Gaussian Process Regression Networks
- International Joint Conference on Artificial Intelligence — Barcelona, Spain, July 2011
  Generalised Wishart Processes
- Uncertainty in Artificial Intelligence — Barcelona, Spain, July 2011
  Generalised Wishart Processes
- Bayesian Econometrics Workshop — Rimini, Italy, June 2011
  Bayesian Nonparametric Volatility Modelling
- ETH — Zurich, Switzerland, February 2011
  New Stochastic Processes for Input Dependent Correlations
- Latent Gaussian Models Workshop — Zurich, Switzerland, February 2011
  Generalised Wishart Processes
- University College London — London, UK, October 2010
  Copula Processes

Reviewing and Outreach

Nature, Biometrika, Neural Computation, Neurocomputing, Journal of Machine Learning Research (JMLR), Electronic Journal of Statistics, Journal of Artificial Intelligence Research (JAIR), IEEE Transactions on Neural Networks, IEEE Transactions on Pattern Analysis and Machine Intelligence, Advances in Neural Information Processing Systems (NeurIPS), International Conference on Machine Learning (ICML), Artificial Intelligence and Statistics (AISTATS), Uncertainty in Artificial Intelligence (UAI), International Conference on Learning Representations (ICLR), Systems and Machine Learning (SysML), International Joint Conference on Artificial Intelligence (IJCAI).

Program Chair for ICML 2027 (primarily responsible for organizing the conference)

Workshop Chair for ICML 2024, ICML 2025 (responsible for running workshops)

Tutorial Chair for NeurIPS 2022, NeurIPS 2023 (responsible for running tutorials)

EXPO Chair for ICML 2019, ICML 2020 (responsible for engaging with corporate research and selecting corporate demonstrations, workshops, and panels).

Senior Area Chair: ICML 2022, NeurIPS 2022, NeurIPS 2023, ICML 2024, ICML 2025, NeurIPS 2024

Editorial Board: Editorial Board Member of New RSS Data Science and Artificial Intelligence Journal, 2014

Co-chair of the IMS Committee on Machine Learning and AI, 2024

Area Chair/SPC: AAAI 2018, AISTATS 2018, UAI 2018, NeurIPS 2018, AISTATS 2019, ICML 2019, IJCAI 2019, UAI 2019, NeurIPS 2019, AAAI 2020, ICLR 2020, IJCAI 2020, UAI 2020, NeurIPS 2020, ICLR 2021, UAI 2021, ICML 2021, NeurIPS 2021, ICLR 2022

ICML 2020 Tutorial Speaker on Bayesian Deep Learning:
`https://www.youtube.com/watch?v=E1qhGw8QxqY`

NSF Panelist, 2018, 2020, 2022

NeurIPS 2021 Competition on Approximate Bayesian Inference.
Website link here.

Founding Member and Mentor for CDS Undergraduate Research Program (CURP).
`https://cds.nyu.edu/curp/`
Mentored URM students with regular meetings for Spring 2021, in collaboration with Prof. Laure Zanna, on probabilistic models for studying climate change.

Symposia/Workshops:

- Co-organiser of ICLR 2024 workshop
  AI for Differential Equations.
- Co-organiser of NeurIPS 2021 workshop
  Bayesian Deep Learning.
- Co-organiser of ELLIS 2020 workshop
  Bayesian Deep Learning.
- Co-organiser of NeurIPS 2019 workshop
  Learning with All Experience: Integrating Learning Paradigms.
- Co-organiser of NeurIPS 2018 workshop
  Bayesian Deep Learning.
- Co-organiser of UAI 2018 workshop
  Uncertainty in Deep Learning.
- Co-organiser of ICML 2018 workshop
  Theoretical Foundations and Applications of Deep Generative Models.
  `https://sites.google.com/view/tadgm`
- Lead organiser of NeurIPS 2017 symposium ($\sim$ 5000 in attendance)
  Interpretable Machine Learning.
  `http://interpretable.ml`
  arXiv index: `https://arxiv.org/abs/1711.09889`
- Co-organiser of the NeurIPS 2017 workshop
  Bayesian Deep Learning.
- Lead organiser of NeurIPS 2016 workshop
  Interpretable Machine Learning for Complex Systems.
  arXiv index: `https://arxiv.org/abs/1611.09139`
- Lead organiser of the NeurIPS 2015 workshop
  Nonparametric Methods for Large Scale Representation Learning.
- Co-organiser of the ICML 2015 workshop
  Large Scale Kernel Learning: Challenges and New Opportunities.
- Co-organiser of the NeurIPS 2014 workshop
  Modern Nonparametrics 3: Automating the Learning Pipeline.

Selection of Teaching

Teaching page, with evaluations: `https://cims.nyu.edu/~andrewgw/teaching/`.

- Bayesian Machine Learning. Fall 2020-2025
- Honors Introduction to Data Science for PhD Students. Fall 2023-2025.
- Honors Introduction to Data Science for PhD Students. Fall 2021.
- Quantum Materials and Machine Learning Summer School Lecturer. Presented lectures on probabilistic machine learning, k-means, PCA, density estimation, Gaussian processes, Bayesian neural networks. June 2021.
- CS 473: *Introduction to Machine Learning* at NYU Courant. Spring 2021.
- ICML 2020 Tutorial on Bayesian Deep Learning.
- NYU AI Winter School Lecturer on Probabilistic Machine Learning. December 2020.
- PhD course on *Bayesian Machine Learning* NYU CDS & CS. Fall 2020.

- Gaussian Process Summer School Lecturer. September 2020.
- SMILES Summer School Lecturer on Bayesian neural networks. July 2020.
- PhD course on *Bayesian Machine Learning* (CSCI-GA.3033-027) at NYU Courant. Fall 2019.
- CS/ORIE/STSCI 1380: *Data Science for All.* Freshman undergraduate course. Spring 2019. Cornell University.
- Lecturer on Bayesian Neural Networks at DLRL 2018.
- Designed the new undergraduate course ORIE 4742: *Information Theory, Probabilistic Modeling, and Deep Learning* at Cornell University.
  Calendar description: `https://classes.cornell.edu/browse/roster/SP17/class/ORIE/4742`.
  Spring 2017, 2018.
- Designed the new PhD course CS/ORIE 6741: *Bayesian Machine Learning* at Cornell University. Course website: `https://people.orie.cornell.edu/andrew/orie6741`.
  Fall 2016, 2017, 2018.
- Lecturer on Markov chain Monte Carlo, Model Selection, and Advanced Gaussian Processes in Probabilistic Graphical Models (10-708), CMU.
- Lecturer on Kernel Methods at the MLSS 2014.

Current Research Students (Primary Advisor)

- Andres Potapczynski (PhD, Data Science), September 2021 – Present
- Shikai Qiu (PhD, Computer Science), September 2022 – Present
- Yucen (Lily) Li (PhD, Computer Science), September 2022 – Present
- Ethan Baron (PhD, Computer Science), September 2024 – Present
- Martin Marek (PhD, Computer Science), September 2024 – Present
- Yixi (Charlie) Chen (PhD, Computer Science), September 2024 – Present
- Vatsal Baherwani (PhD, Data Science), September 2025 – Present

Current Postdoctoral Fellows

- Alan Amin, September 2023 – Present

Alumni (* denotes primary advisor)

- Sanae Lotfi* (PhD, Data Science), September 2020 – May 2025.
  Now Research Scientist at Meta FAIR.
- Sanyam Kapoor* (PhD, Data Science), September 2020 – May 2025.
  Now Research Scientist at NVIDIA.
- Micah Goldblum* (Postdoc, Data Science), September 2021 – June 2024.
  Now Assistant Professor at Columbia University.
- Pavel Izmailov* (PhD, Computer Science), August 2017 – May 2023.
  Now at Anthropic, then Assistant Professor at NYU.
- Marc Finzi* (PhD, Computer Science), August 2017 – May 2023.
  Now Postdoctoral Fellow at CMU.
- Polina Kirichenko* (PhD, Data Science), August 2018 – May 2024.
  Now Research Scientist at Meta FAIR.
- Ben Athiwaratkun* (PhD, Statistics, Cornell), April 2017 – May 2019.
  Now Senior Research Scientist at Amazon AI.
- Jacob Gardner* (Postdoc, Cornell), September 2018 – May 2019.
  Now Assistant Professor at the University of Pennsylvania.
- Ruqi Zhang (PhD, Statistics, Cornell), September 2017 – May 2018.
  Now Assistant Professor at Purdue University.
- Greg Benton* (PhD, Computer Science), January 2019 – January 2023.
  Now Engineer at Celonis.
- Samuel Stanton* (PhD, Data Science), March 2017 – August 2022.
  Now Research Scientist at Prescient Design.

- Wesley Maddox* (PhD, Data Science), August 2017 – May 2022.
  Now Quantitative Researcher at Jump Trading.
- Geoff Pleiss (PhD, CS, Cornell), March 2018 – July 2020.
  Now Assistant Professor at the University of British Columbia.
- Alex Wang* (Advised as both Undergraduate and Masters, Computer Science, Cornell), September 2018 – May 2020. Now PhD student in Computer Science at Stanford University.
- Ian Delbridge* (Masters, Computer Science, Cornell), September 2018 – May 2020.
  Now Data Scientist at Klaviyo
- Patrick Nicholson* (Advised as both Undergraduate and Masters, Computer Science, Cornell), August 2017 – May 2018. Now Research Scientist at Covera Health.

External PhD Thesis Examiner

- Joost van Amersfoort (PhD, Computer Science, Oxford). February 2023. Co-examiner. Prof. Atilim Baydin.
- Sebastian Farquhar (PhD, Computer Science, Oxford). April 2022. Co-examiner. Prof. Stephen Roberts.
- Marina Munkhoeva (PhD, Computer Science, Skoltech). April 2021. Co-examiners: Prof. Michael Bronstein, Prof. Andrzej Cichocki, Prof. Victor Lempitsky, Prof. Maxim Panov
- Yermek Kapushev (PhD, Computer Science, Skoltech). February 2021. Co-examiners: Prof. Maurizio Filippone, Prof. Maxim Fedorov, Prof. Alexey Zaytsev
- Guillermo Valle Perez (PhD, Physics, Oxford). February 2021. Co-examiner: Prof. Ard Louis.
- Konstantinos Pitas (PhD, Computer Science, EPFL). September 2020. Co-examiners: Prof. Mario Figueiredo, Prof. Martin Jaggi

Research Students (Committee Member)

- Zahra Kadkhodaie (PhD, Neuroscience, NYU), April 2023 – Present
- Benjamin Lyo (PhD, Neuroscience, NYU), March 2022 – Present
- Yunzhen Feng (PhD, CDS, NYU), March 2023 – Present
- Ben Jakubowski (PhD, CS, NYU), February 2021 – Present
- Irina Espejo Morales (PhD, CDS, NYU), September 2021 – May 2023 (Graduated)
- Ilya Kulikov (PhD, CS, NYU), September 2021 – May 2021 (Graduated)
- Jason Lee (PhD, CS, NYU), September 2019 – May 2021 (Graduated)
- William Herlands (PhD, Machine Learning, CMU), August 2016 – May 2020 (Graduated)
- Kun Dong (PhD, CAM, Cornell), January 2017 – July 2019 (Graduated)
- Skyler Seto (PhD, Statistics, Cornell), January 2017 – July 2019 (Graduated)
- Matthew Davidow (PhD, CAM, Cornell), January 2017 – July 2020 (Graduated)
- Daniel Gilbert (PhD, Statistics, Cornell), September 2017 – May 2019 (Graduated)
- Aman Agarwal (PhD, CS, Cornell), December 2017 – July 2020 (Graduated)
- Geoff Pleiss (PhD, CS, Cornell), March 2018 – July 2020 (Graduated)

Grants

Schmidt Sciences. *Safety Pre-Training for Generalization under Distribution Shifts and Weak Supervision.* $3,000,000. Lead PI. 2026–2028. With P. Izmailov, M. Goldblum.
DARPA-AIQ. *Consistency, Robustness and Generalization Guarantees for Understanding Language Models.* $1,800,000. Co-PI. 2025–2028. With S. Feizi, F. Huang.
NSF HDR-2118310. *Quantum Integration of Data and Emergence at Atomic Scales.* Co-PI. $2,400,000. August 2022 - August 2025. With E. Kim, K. Weinberger, M. Greiner, L. Schoop.
NSF CAREER IIS-2145492. *New Frontiers in Bayesian Deep Learning.* $485,000. Sole PI. May 2022 - May 2027.
Capital One Research Grant. $150,000. December 2021.
Amazon Research Gift. $70,000. December 2021.
Google Research Gift. $10,000. November 2021.
NSF CDS&E-MSS 2134216. *Statistical and Computational Foundations of Deep Generative Models.* $1,200,000. Co-PI. August 2021 – August 2026. With E. Vanden-Eijnden, J. Bruna, J. Niles-Weed, G. Ben-Arous, M. Gabre.

Facebook Research Gift. $75,000. August 2021.
Amazon Research Gift. $100,000. May 2021.
BigHat Biosciences Research Gift. $50,000. May 2021.
Amazon Research Award. *Scalable Numerical Methods for Probabilistic Deep Learning.* September 2020. Sole PI. $60,000 + $100,000 AWS Credits.
NIH R01 DA048764-01A1. *Analyzing Sequential, Multiple Assignment, Randomized Trials in the Presence of Partial Compliance.* Co-I. $1,585,000. September 2019 – September 2023. With A. Ertefaie, B Johnson, M. Kosorok, J. McKay.
NSF IIS-1910266. *Scalable Online Gaussian Processes.* Sole PI. $400,000. August 2019 – August 2022.
NSF I-DISRE 1934714. *Understanding Subatomic-Scale Quantum Matter Data Using Machine Learning Tools.* Co-PI. $1,300,000. August 2019 – August 2022. With E. Kim, K. Weinberger
Amazon Research Award. *New Directions for Non-Convex Optimization in Deep Learning.* January 2019. Sole PI. $80,000 + $20,000 AWS Credits.
Google Cloud Award. Hundreds of TPUs on Google Cloud. December 2018. Sole PI.
Facebook Research Award. *Scalable Gaussian Processes.* November 2018. Sole PI. $130,000.
NSF IIS-1563887. *Scaling Machine Learning for Automating Scientific Discovery in Astrophysics.* August 2016 – July 2019. Co-PI. $1,200,000. With B. Poczos, R. Mandelbaum, E.P. Xing.

Example Video Lectures (Clickable Links)

2025: Deep Learning is Not So Mysterious or Different
2024: How Do We Build a General Intelligence?
2024: Machine Learning is Linear Algebra
2024: Simons Institute Polylogue Interview
2020: Bayesian Deep Learning and Probabilistic Model Construction
2020: Introduction to Bayesian Machine Learning
2017: Bayesian Optimization with Gradients (with Peter I. Frazier)
2016: Scalable Gaussian Processes for Scientific Discovery

Employment

Perplexity AI, NYC **2026 – Present**
Research Lead
- I lead Perplexity Frontier Research. This work spans pre-training, continual learning, RL post-training, multi-agent orchestration, and evaluation. We are building a system for automatic scientific research.

Amazon Supply Chain Optimization Technologies (SCOT), NYC **2020 – 2026**
Amazon Scholar
- In this 1 day/week position, I develop foundation models for time-series forecasting.

Microsoft Research, Cambridge, UK **07/2012 – 09/2012**
Research Intern
- I developed Bayesian inference techniques, and new Bayesian nonparametric models, for NMR spectroscopy. These new machine learning techniques can be used to make predictions about chemical concentrations and the progress of chemical reactions, and are markedly different from conventional NMR spectroscopy techniques.

TRIUMF, Vancouver, Canada **09/2007 – 08/2008**
Researcher
- Positron Emission Tomography (PET) is used to visualise functional activity, as opposed to anatomical structure; for example, it can be used to trace thought processes. At TRIUMF, the world's largest cyclotron laboratory, I independently devised image reconstruction algorithms necessary for the operation of a groundbreaking new PET device.

University of British Columbia, Vancouver, Canada **05/2007 – 08/2007**
Teaching Assistant, Mathematics Department
- I was the teaching assistant for a third year class in partial differential equations. I graded approximately 70 assignments weekly, and gave tutorial lectures twice weekly, where I derived theorems and explained concepts. I also tutored individuals and groups, and helped students with test preparation.

University of British Columbia, Vancouver, Canada **05/2006 – 08/2006**

Researcher, Physics Department, Supervisor: Matthew Choptuik

- I worked on developing a scientific programming language. I wrote a grammar and a parser to interpret the rules of the language. The language numerically solves partial differential equations, given the equations and the boundary conditions. The language also generates `C` and `Fortran` solution templates, and animated visualizations of the solution. I used `C`, `Fortran`, `Perl`, `Flex` (`Lex`), `Bison` (`Yacc`), `tcsh` and `bash`. The project consisted of 182 sources written in these languages. I also worked on a code-driver, using `Perl`, to generate fully functioning `C` and `Fortran` programs from a small number of declarations in an input file. This work was motivated to assist in using general relativity to model physical problems.

MISC

I am a classically trained pianist. I particularly like Glenn Gould's playing of Bach. I also enjoy reading about modern physics, and writing essays.