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 wish to understand the foundations of learning and decision making towards developing intelligent systems. I am particularly engaged in building methods for probabilistic deep learning, scalable Gaussian processes, physics-inspired machine learning, kernel learning, and training of deep neural networks. I have applied my work to time series, vision, NLP, spatial statistics, counterfactual inference, public policy, medicine, and physics. I also believe in open and reproducible research, and maintain several software libraries.

CURRENT POSITION

Assistant Professor, New York University **July 2019 –**
Courant Institute of Mathematical Sciences and Center for Data Science
Computer Science Department, Mathematics Department (Affiliated)

ACADEMIC BACKGROUND

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

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

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

- **Supervisor**: Zoubin Ghahramani
- **Thesis**: *Covariance Kernels for Fast Automatic Pattern Discovery and Extrapolation with Gaussian Processes.* October 25, 2014.

BSc (Hons), University of British Columbia **May 2008**
Mathematics and Physics

- A+ Graduating Average, Highest Ranking Honours Physics Thesis.
- **Thesis**: *Position and Energy Reconstruction from Scintillation Light in a Liquid Xenon Gamma Ray Detector designed for PET.*

AWARDS

- 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
- NSERC Canadian Graduate Scholarship (Masters) (CGS-M) (Declined) 2009
- 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] A.G. Wilson, P. Izmailov. Bayesian Deep Learning and a Probabilistic Perspective of Generalization. *Advances in Neural Information Processing Systems* (NeurIPS), 2020.

[2] 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.

[3] 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.

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

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

[6] 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.

[7] 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.

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

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

[10] 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**.

[11] 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.

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

[13] 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.

[14] 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.

[15] 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.

[16] 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.

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

[18] 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.

[19] 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.

[20] 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**.

[21] 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**.

[22] 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.

[23] 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**.

[24] 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.

[25] 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.

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

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

[28] 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.

[29] 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.

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

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

[32] 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.

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

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

[35] 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.

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

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

[38] 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.

[39] 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.

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

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

[42] 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.

[43] 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**.

[44] 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.

[45] 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.

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

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

[48] 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**.

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

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

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

Pre-prints

[52] W. Maddox, G. Benton, A.G. Wilson. *Rethinking Parameter Counting in Deep Models: Effective Dimensionality Revisited*. 2020.

[53] A.G. Wilson. *The Case for Bayesian Deep Learning*. 2019.

[54] A.G. Wilson, C. Dann, and H. Nickisch. Thoughts on massively scalable Gaussian processes. arXiv pre-print 2015. *Extended version in preparation for* JMLR.

[55] S. Flaxman, A. Gelman, D.B. Neill, A.J. Smola, A. Vehtari, and A.G. Wilson. Fast hierarchical Gaussian processes. 2015.

[56] A.G. Wilson, Y. Wu, D. J. Holland, S. Nowozin, M.D. Mantle, L.F Gladden, and A. Blake. Bayesian inference for NMR spectroscopy. arXiv pre-print 2014. *In preparation for the Electronic Journal of Statistics*.

[57] A.G. Wilson*, E. Gilboa* (equal contribution), A. Nehorai, and J.P. Cunningham. GPatt: Fast multidimensional pattern extrapolation with Gaussian processes. arXiv pre-print 2013. Extended into: *Fast kernel learning for multidimensional pattern extrapolation* at NeurIPS 2014.

Refereed Workshop Papers

[58] M. Al-Shedivat, A.G. Wilson, Y. Saatchi, Z. Hu, and E.P. Xing. Scalable GP-LSTMs with Semi-Stochastic Gradients. *NeurIPS Workshop on Bayesian Deep Learning*, 2016.

[59] Y. Ma, R. Garnett, J. Schneider, and A.G. Wilson. Fast Bayesian Optimization via Conjugate Sampling. *NeurIPS Workshop on Practical Bayesian Nonparametrics*, 2016.

[60] M. Van der Wilk, A.G. Wilson, and C.E. Rasmussen. Variational inference for latent variable modelling of correlation structure. *NeurIPS Workshop on Advances in Variational Inference*, 2014.

[61] A. Shah, A.G. Wilson, and Z. Ghahramani. Student-t processes for Bayesian optimisation. *NeurIPS Workshop on Bayesian Optimization*, 2013.

Reports

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

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

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

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

[66] 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.

Code Repository Selection

- `https://cims.nyu.edu/~andrewgw/code`. Main resource page for code repositories from myself and collaborators.
- `https://github.com/cornellius-gp/gpytorch`. A popular library for massively scalable Gaussian processes. Joint work with Jake Gardner, Geoff Pleiss, and Kilian Weinberger.
- `https://github.com/andrewgordonwilson`. Provides the `bayesgan` repository. Joint work with Yunus Saatchi.
- `https://cims.nyu.edu/~andrewgw/pattern`. Provides numerous resources for scalable and flexible kernel learning. Joint work with Hannes Nickisch.
- `https://github.com/benathi/word2gm`. A popular library for probabilistic word embeddings. Joint work with Ben Athiwaratkun.

Selected Talks

• Machine Learning in Science and Engineering (Columbia University)	Online, December 2020
• NYU Winter School (for URM)	Online, December 2020
• Gaussian Process Summer School	Online, September 2020
• Smiles Skoltech Summer School	Online, August 2020
• ICML Bayesian Deep Learning Tutorial	Online, July 2020
• Google Brain Research Seminar	Online, May 2020
• BIRS Workshop	Banff, Canada, February 2020
• NeurIPS 2019 Bayesian Deep Learning Workshop	Vancouver, Canada, December 2019
• Binghampon University (SUNY) Dean's Speaker Series	Binghamton, NY, November 2019
• MIT Broad Institute	Cambridge, USA, October 2019
• Data for Good	New York, NY, September 2019
• Los Alamos National Laboratory	Los Alamos, USA, April 2019
• New York University	NY, USA, February 2019
• University of Maryland	College Park, USA, February 2019
• University of Michigan	Ann Arbor, USA, February 2019
• UNC Chapel Hill	Chapel Hill, USA, February 2019
• MIT Seminar	Cambridge, USA, November 2018
• Boston University Seminar	Cambridge, USA, November 2018
• Allerton Conference	Allerton, IL, October 2018
• PyTorch DevCon	San Francisco, USA, October 2018
• Precision Medicine and Machine Learning	Durham, NC, August 2018
• Deep Learning Summer School	Toronto, July 2018
• SIAM ALA (Applied Linear Algebra)	Hong Kong, May 2018
• DALI 2018	Canary Islands, April 2018
• BIRS Workshop (Stats & ML)	Banff, Canada, January 2018
• UCL Gatsby	London, UK, December 2017
• University of Cambridge	Cambridge, UK, December 2017
• Microsoft Research	Cambridge, UK, December 2017
• CMStatistics	London, UK, December 2017
• AI Seminar, Cornell	Ithaca, NY, October 2017
• Statistics Seminar, Cornell	Ithaca, NY, September 2017
• Linköping University	Linköping, Sweden, April 2017
• UCLA	Los Angeles, USA, January 2017
• University of British Columbia	Vancouver, Canada, March 2016
• University of Edinburgh	Edinburgh, UK, March 2016

- University of Southern California — Los Angeles, USA, March 2016
- University of California, Irvine — Irvine, USA, March 2016
- UCLA — Los Angeles, USA, March 2016
- University of Massachusetts — Amherst, USA, March 2016
- Cornell University — Ithaca, USA, March 2016
- University of Toronto — Toronto, Canada, February 2016
- Dartmouth College — Hanover, USA, February 2016
- EPFL — Lausanne, Switzerland, February 2016
- University of Waterloo — Waterloo, Canada, January 2016
- University of Cambridge — Cambridge, UK, August 2015
- International Conference on Machine Learning — Lille, France, July 2015
- New York University — NYC, USA, June 2015
- Neural Information Processing Systems Workshop — Montreal, Canada, December 2014
- Oxford University — Oxford, UK, November 2014
- University College London — London, UK, November 2014
- Machine Learning Summer School (MLSS) — Pittsburgh, USA, July 2014
- International Conference on Machine Learning — Atlanta, USA, June 2013
- Xerox Research Seminar — Grenoble, France, November 2012
- ECML Nectar Track — Bristol, UK, September 2012
- Microsoft Research — Cambridge, UK, September 2012
- International Conference on Machine Learning — Edinburgh, UK, June 2012
- University of California, Berkeley — Berkeley, USA, May 2012
- Harvard University — Cambridge, USA, April 2012
- International Joint Conference on Artificial Intelligence — Barcelona, Spain, July 2011
- Uncertainty in Artificial Intelligence — Barcelona, Spain, July 2011
- Bayesian Econometrics Workshop — Rimini, Italy, June 2011
- ETH — Zurich, Switzerland, February 2011
- Latent Gaussian Models Workshop — Zurich, Switzerland, February 2011
- University College London — London, UK, October 2010

Reviewing and Service (Clickable Links)

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).

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.

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

NSF Panelist, 2018, 2020.

Symposia/Workshops:

- 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.

Teaching

- 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.

Research Students (Primary Advisor)

- Ben Athiwaratkun (PhD, Statistics), Year 4, April 2017 – May 2019
  (Graduated; Now Research Scientist at Amazon AI)
- Pavel Izmailov (PhD, Computer Science), Year 3, August 2017 – Present
- Polina Kirichenko (PhD, Data Science), Year 2, August 2018 – Present
- Marc Finzi (PhD, Computer Science), Year 3, August 2017 – Present
- Samuel Stanton (PhD, Data Science), Year 3, March 2017 – Present
- Wesley Maddox (PhD, Data Science), Year 3, August 2017 – Present
- Gregory Benton (PhD, Computer Sicence), Year 2, January 2018 – Present
- Ian Delbridge (Masters, Computer Science), January 2019 – Present
- Alex Wang (Masters, Computer Science), September 2018 – Present
- Patrick Nicholson (Masters, Computer Science), August 2017 – May 2018 (Graduated)

Research Students (Committee Member)

- William Herlands (PhD, Machine Learning, CMU), Year 4, August 2016 – Present
- Kun Dong (PhD, CAM), January 2017 – July 2019 (Graduated)
- Skyler Seto (PhD, Statistics), January 2017 – July 2019 (Graduated)
- Matthew Davidow (PhD, CAM), Year 4, January 2017 – Present
- Chris Browne (PhD, CAM), Year 4, January 2017 – Present
- Jiajun Gu (PhD), Year 4, March 2018 – Present
- Daniel Gilbert (PhD, Statistics), September 2017 – May 2019 (Graduated)
- Aman Agarwal (PhD, CS), Year 4, December 2017 – Present
- Zhilu Zhang (PhD, ECE), Year 3, December 2017 – Present
- Geoff Pleiss (PhD, CS), Year 4, March 2018 – Present
- Raul Astudillo Marban (PhD, ORIE), Year 3, April 2018 – Present

Grants

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.
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 2021.
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.

Example Video Lectures (Clickable Links)

2020: Loss Geometry for Practical Bayesian Deep Learning
2019: Bayesian Deep Learning
2018: Bayesian GAN
2017: Bayesian optimization with gradients (with Peter I. Frazier)
2016: Scalable Gaussian processes for scientific discovery
2015: Kernel interpolation for scalable structured Gaussian processes
2014: Kernel methods for large scale representation learning

Employment

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 anatomatical 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.