

# Andrew Gordon Wilson

---

CONTACT INFORMATION      Operations Research and Information Engineering      andrew@cornell.edu  
Cornell University      <https://people.orie.cornell.edu/andrew>  
235 Rhodes Hall, 136 Hoy Road, Ithaca, NY, 14853, USA      <https://twitter.com/andrewgwils>

RESEARCH INTERESTS      I wish to understand the foundations of learning and decision making towards developing intelligent systems. I often (1) take a probabilistic approach; (2) care about automatically discovering scientifically interpretable structure in data; and (3) am excited about bridging connections with scientific computing. I am particularly active in building methods for probabilistic deep learning, scalable Gaussian processes, kernel learning, and training of deep neural networks. My work has been applied to time series, vision, NLP, spatial statistics, public policy, medicine, and physics. I believe in open and reproducible research.

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

ACADEMIC BACKGROUND      **Research Fellow, Carnegie Mellon University**      March 2014 – August 2016  
*Machine Learning Department, School of Computer Science*  
• **Mentors:** Eric P. Xing and Alexander J. Smola

**PhD, Trinity College, University of Cambridge**      October 2009 – March 2014  
*Machine Learning, Department of Engineering*  
• **Supervisor:** Zoubin Ghahramani  
• **Thesis:** *Covariance Kernels for Fast Automatic Pattern Discovery and Extrapolation with Gaussian Processes*. January 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

- Best Poster Award (with Ben Athiwaratkun), NIPS ML Train Workshop      2017
- Outstanding PhD Dissertation (£10,000), G-Research      2014
- Outstanding Reviewer Award, Neural Information Processing Systems (NIPS)      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] P. Izmailov\*, D. Podoprikin\*, T. Garipov\*, D. Vetrov, A.G. Wilson. Averaging Weights Leads to Wider Optima and Better Generalization, *Uncertainty in Artificial Intelligence (UAI)*, 2018. **Oral presentation.**
- [2] 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.
- [3] 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.

- [4] B. Athiwaratkun, A.G. Wilson, and A. Anandkumar. Probabilistic FastText. *Association for Computational Linguistics (ACL)*, 2018. **Oral presentation.**
- [5] B. Athiwaratkun and A.G. Wilson. Hierarchical Density Order Embeddings. *International Conference on Learning Representations (ICLR)*, 2018.
- [6] 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.
- [7] 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.
- [8] Y. Saatchi and A.G. Wilson. Bayesian GAN. *Neural Information Processing Systems (NIPS)*, 2017. **Spotlight.**
- [9] J. Wu, M. Poloczek, A.G. Wilson, and P. Frazier. Bayesian optimization with gradients. *Neural Information Processing Systems (NIPS)*, 2017. **Oral presentation.**
- [10] K. Dong, D. Eriksson, H. Nickisch, D. Bindel, and A.G. Wilson. Scalable log determinants for Gaussian process kernel learning. *Neural Information Processing Systems (NIPS)*, 2017.
- [11] A. Loeb, P. Jang, M. Davidow, and A.G. Wilson. Scalable Lévy process kernel learning. *Neural Information Processing Systems (NIPS)*, 2017.
- [12] B. Athiwaratkun and A.G. Wilson. Multimodal Word Distributions. *Association for Computational Linguistics (ACL)*, 2017.
- [13] 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.
- [14] A.G. Wilson\*, Z. Hu\* (equal contribution), R. Salakhutdinov, and E.P. Xing. Stochastic Variational Deep Kernel Learning. *Neural Information Processing Systems (NIPS)*, 2016.
- [15] A.G. Wilson\*, Z. Hu\* (equal contribution), R. Salakhutdinov, and E.P. Xing. Deep kernel learning. *Artificial Intelligence and Statistics (AISTATS)*, 2016.
- [16] 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.
- [17] 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.
- [18] A.G. Wilson, C. Dann, C.G. Lucas, and E.P. Xing. The human kernel. In *Neural Information Processing Systems (NIPS)*, 2015. **Spotlight.**
- [19] A.G. Wilson and H. Nickisch. Kernel interpolation for scalable structured Gaussian processes (KISS-GP). *International Conference on Machine Learning (ICML)*, 2015.
- [20] 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.
- [21] 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.**
- [22] A.G. Wilson\*, E. Gilboa\* (equal contribution), A. Nehorai, and J.P. Cunningham. Fast kernel learning for multidimensional pattern extrapolation. *Neural Information Processing Systems (NIPS)*, 2014.
- [23] 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.
- [24] A. Shah, A.G. Wilson, and Z. Ghahramani. Student- $t$  processes as alternatives to Gaussian processes. *Artificial Intelligence and Statistics (AISTATS)*, 2014.
- [25] A.G. Wilson and R.P. Adams. Gaussian process kernels for pattern discovery and extrapolation. *International Conference on Machine Learning (ICML)*, 2013. **Oral presentation.**
- [26] 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.**
- [27] A.G. Wilson, D.A. Knowles, and Z. Ghahramani. Gaussian process regression networks. *International Conference on Machine Learning (ICML)*, 2012. **Oral presentation.**

- [28] A.G. Wilson and Z. Ghahramani. Generalised Wishart processes. *Uncertainty in Artificial Intelligence (UAI)*, 2011. **Best Student Paper Award**.
- [29] A.G. Wilson and Z. Ghahramani. Copula processes. *Neural Information Processing Systems (NIPS)*, 2010. **Spotlight**.

## PRE-PRINTS

- [30] B. Athiwaratkun, M. Finzi, P. Izmailov, A.G. Wilson. Improving Consistency-Based Semi-Supervised Learning, 2018.
- [31] W. Herlands, D.B. Neill, H. Nickisch, and A.G. Wilson. Change Surfaces for Expressive Multidimensional Changepoints and Counterfactual Prediction, 2018.
- [32] T. Garipov\*, P. Izmailov\*, D. Podoprikin\*, D. Vetrov, A.G. Wilson. Loss Surfaces, Mode Connectivity, and Fast Ensembling of DNNs, 2018.
- [33] A.G. Wilson, C. Dann, and H. Nickisch. Thoughts on massively scalable Gaussian processes. arXiv pre-print 2015. *Extended version in preparation for JMLR*.
- [34] S. Flaxman, A. Gelman, D.B. Neill, A.J. Smola, A. Vehtari, and A.G. Wilson. Fast hierarchical Gaussian processes. 2015.
- [35] 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*.
- [36] 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 NIPS 2014.

REFEREED  
WORKSHOP  
PAPERS

- [37] M. Al-Shedivat, A.G. Wilson, Y. Saatchi, Z. Hu, and E.P. Xing. Scalable GP-LSTMs with Semi-Stochastic Gradients. *NIPS Workshop on Bayesian Deep Learning*, 2016.
- [38] Y. Ma, R. Garnett, J. Schneider, and A.G. Wilson. Fast Bayesian Optimization via Conjugate Sampling. *NIPS Workshop on Practical Bayesian Nonparametrics*, 2016.
- [39] M. Van der Wilk, A.G. Wilson, and C.E. Rasmussen. Variational inference for latent variable modelling of correlation structure. *NIPS Workshop on Advances in Variational Inference*, 2014.
- [40] A. Shah, A.G. Wilson, and Z. Ghahramani. Student- $t$  processes for Bayesian optimisation. *NIPS Workshop on Bayesian Optimization*, 2013.

## REPORTS

- [41] A.G. Wilson. Covariance kernels for fast automatic pattern discovery and extrapolation with Gaussian processes. PhD Thesis, University of Cambridge. January 2014.
- [42] A.G. Wilson. The change point kernel. Technical report, University of Cambridge. Nov 2013.
- [43] A.G. Wilson. A process over all stationary covariance kernels. Technical report, University of Cambridge. June 2012.
- [44] A.G. Wilson. Latent Gaussian process models. First year report, University of Cambridge. August 2010.
- [45] 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  
REPOSITORIES

- <http://people.orie.cornell.edu/andrew/code>. Main resource page for code repositories from myself and collaborators.
- <https://github.com/andrewgordonwilson>. Provides the `bayesgan` repository (> 880 stars, 135 forks). Joint work with Yunus Saatchi.
- <http://people.orie.cornell.edu/andrew/pattern>. Provides numerous resources for scalable and flexible kernel learning. Joint work with Hannes Nickisch.
- <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/benathi/word2gm>. A popular library for probabilistic word embeddings. Joint work with Ben Athiwaratkun.

## SELECTED TALKS

- 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 (NIPS), International Conference on Machine Learning (ICML), Artificial Intelligence and Statistics (AISTATS), Uncertainty in Artificial Intelligence (UAI), International Conference on Learning Representations (ICLR), International Joint Conference on Artificial Intelligence (IJCAI).

**Area Chair/SPC:** AAAI 2018, AISTATS 2018, UAI 2018, NIPS 2018.

**NSF Panelist,** 2018.

**Symposia/Workshops:**

- 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 NIPS 2017 symposium (~ **5000 in attendance**)  
*Interpretable Machine Learning.*  
<http://interpretable.ml>  
arXiv index: <https://arxiv.org/abs/1711.09889>
- Co-organiser of the NIPS 2017 workshop  
*Bayesian Deep Learning.*
- Lead organiser of NIPS 2016 workshop  
*Interpretable Machine Learning for Complex Systems.*  
arXiv index: <https://arxiv.org/abs/1611.09139>
- Lead organiser of the NIPS 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 NIPS 2014 workshop  
*Modern Nonparametrics 3: Automating the Learning Pipeline.*

## TEACHING

- ORIE 4742: *Information Theory, Probabilistic Modeling, and Deep Learning.* Spring 2018.
- ORIE 6741: *Bayesian Machine Learning.* 100 graduate students enrolled. Fall 2017.
- Designed the new undergraduate course ORIE 4742: *Information Theory, Probabilistic Modeling, and Deep Learning* at Cornell University. Spring 2017.  
Calendar description: <https://classes.cornell.edu/browse/roster/SP17/class/ORIE/4742>
- Designed the new PhD course ORIE 6741: *Bayesian Machine Learning* at Cornell University. 43 graduate students enrolled.  
Course website: <https://people.orie.cornell.edu/andrew/orie6741>. Fall 2016.
- 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 2014 Machine Learning Summer School.

RESEARCH  
STUDENTS  
(PRIMARY  
ADVISED)

- Ben Athiwaratkun (PhD, Statistics), Year 3, April 2017 – Present
- William Herlands (PhD, Machine Learning, CMU), Year 3, co-advised by Daniel Neill, August 2016 – Present
- Pavel Izmailov (PhD, ORIE), Year 1, August 2017 – Present
- Marc Finzi (PhD, ORIE), Year 1, August 2017 – Present
- Samuel Stanton (PhD, ORIE), Year 1, March 2017 – Present

- Wesley Maddox (PhD, Statistics), Year 1, August 2017 – Present
- Ruqi Zhang (PhD, Statistics), Year 2, August 2017 – Present
- Cheng Perng Phoo (PhD, CS), Year, 1, co-advised by Bharath Hariharan, August 2017 – Present
- Michael Luo (Masters, CS), August 2017 – Present
- Patrick Nicholson (Masters, CS), August 2017 – Present

RESEARCH  
STUDENTS  
(COMMITTEE  
MEMBER)

- Kun Dong (PhD, CAM), Year 3, January 2017 – Present
- Skyler Seto (PhD, Statistics), Year 3, January 2017 – Present
- Chris Browne (PhD, CAM), Year 2, January 2017 – Present
- Daniel Gilbert (PhD, Statistics), Year 3, September 2017 – Present
- Aman Agarwal (PhD, CS), Year 3, December 2017 – Present
- Zhilu Zhang (PhD, ECE), Year 2, December 2017 – Present
- Geoff Pleiss (PhD, CS), Year 3, March 2018 – Present
- Raul Astudillo Marban (PhD, ORIE), Year 1, April 2018 – Present

GRANTS

NSF IIS-1563887. *Scaling Machine Learning for Automating Scientific Discovery in Astrophysics*. August 2016 – July 2019. Co-PI. **\$1,100,000**.

VIDEO LECTURES  
(CLICKABLE  
LINKS)

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  
 2013: Kernels for automatic pattern discovery and extrapolation  
 2012: Gaussian process regression networks  
 2010: Copula processes

EMPLOYMENT

**Microsoft Research**, Cambridge, UK **07/2012 – 09/2012**  
*Research Intern, Supervisor: Sebastian Nowozin*

- 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, Supervisor: Douglas Bryman*

- 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.

REFERENCES  
(PLEASE DO NOT  
CONTACT WITHOUT  
MY PERMISSION)

Alexander J. Smola Director of AI Research Amazon Web Services Professor Machine Learning Department Carnegie Mellon University Pittsburgh, PA, 15213, USA	Eric P. Xing CEO, Petuum Inc. Professor Machine Learning Department Carnegie Mellon University Pittsburgh, PA, 15213, USA	Zoubin Ghahramani Chief Scientist, Uber Professor Department of Engineering University of Cambridge Cambridge, CB2 1PZ, UK
--	--	---

Ruslan Salakhutdinov Director of Apple AI Research Associate Professor Machine Learning Department Carnegie Mellon University Pittsburgh, PA, 15213, USA	Sebastian Nowozin Principal Researcher Microsoft Research Cambridge Cambridge, CB1 2FB, UK
---	---