
Yee Whye Teh

Curriculum Vitae

Department of Statistics Webpage: <http://www.stats.ox.ac.uk/~teh>
24-29 St Giles Email: y.w.teh@stats.ox.ac.uk
Oxford OX1 3LB Mobile: +44-7392100886
United Kingdom

Brief Biography

I am a Professor of Statistical Machine Learning at the Department of Statistics, University of Oxford, a Principal Research Scientist at DeepMind, an Alan Turing Institute Faculty Fellow and an ELLIS Fellow, co-director of the ELLIS Robust Machine Learning Programme and co-director of the ELLIS@Oxford ELLIS unit. I obtained my PhD at the University of Toronto, and did postdoctoral work at the University of California at Berkeley and National University of Singapore, where I was a Lee Kuan Yew Postdoctoral Fellow. I was a Lecturer and a Reader at the Gatsby Computational Neuroscience Unit, UCL and an ERC Consolidator Fellow.

My research interests are in machine learning, computational statistics and artificial intelligence, in particular probabilistic models, Bayesian nonparametrics, large scale learning and deep learning. I also have interests in using statistical and machine learning tools to solve problems in genetics, genomics, linguistics, neuroscience and artificial intelligence.

I was programme co-chair of the International Conference on Artificial Intelligence and Statistics 2010, Machine Learning Summer School 2014 (Iceland), and the International Conference on Machine Learning 2017, an editor for a IEEE TPAMI Special Issue on Bayesian nonparametrics, and is/was an associate/action editor for Bayesian Analysis, IEEE Transactions on Pattern Analysis and Machine Intelligence, Machine Learning Journal, Journal of the Royal Statistical Society Series B, Statistical Sciences and Journal of Machine Learning Research. I have been area chair or senior area chair for NIPS, ICML and AISTATS on multiple occasions.

Academic History

Sep 2016 - Present Faculty Fellow.

Alan Turing Institute.

Apr 2016 - Present RSIV Professor of Statistical Machine Learning.

Department of Statistics, University of Oxford.

May 2014 - May 2019 ERC Consolidator Fellow.

Sep 2012 - Mar 2016 Professor of Statistical Machine Learning.

Titular Professor, Department of Statistics, University of Oxford.

Tutorial Fellow, University College, Oxford.

Jan 2007 - Aug 2011 Lecturer, and

Aug 2011 - August 2012 Reader in Computational Statistics and Machine Learning.

Gatsby Computational Neuroscience Unit, University College London.

- Mar 2011 - Jun 2011** Visitor.
Wellcome Trust Centre for Human Genetics, University of Oxford.
Host: Gilean McVean and Chris Holmes.
- Aug 2005 - Dec 2006** Lee Kuan Yew Postdoctoral Fellow.
Department of Computer Science, National University of Singapore.
Host: Wee Sun Lee.
- Feb 2003 - Dec 2004** Postdoctoral Fellow.
Division of Computer Science, University of California at Berkeley.
Supervisors: Michael I. Jordan, David A. Forsyth.
- Jan 2000 - Jan 2003** Doctor of Philosophy.
Department of Computer Science, University of Toronto.
Supervisor: Geoffrey E. Hinton.
Thesis: Bethe Free Energy and Contrastive Divergence Approximations for Undirected Graphical Models.
- Jul 1999 - Aug 2001** Affiliate research student.
Gatsby Computational Neuroscience Unit, University College London.
Supervisor: Geoffrey E. Hinton
- Sep 1997 - Dec 2000** Master of Science.
Department of Computer Science, University of Toronto.
Supervisor: Geoffrey E. Hinton.
Thesis: Learning to Parse Images.
- Sep 1994 - Aug 1997** Bachelor of Mathematics.
Faculty of Mathematics, University of Waterloo.
Double honours in Computer Science and Pure Mathematics.
- Apr 1997 - Aug 1997** Research assistant.
Department of Computer Science, University of Waterloo.
Supervisor: Fahiem Bacchus.

Industrial Activities

- 2019 - Present** Principal Research Scientist. DeepMind.
- 2016 - 2019** Senior Staff Research Scientist. DeepMind.
Solve intelligence and use it to make the world a better place.
- 2014 - 2018** Consultant on Machine Learning and Predictive Analytics. Swhere Plc.
UK start-up providing predictive analytics and machine learning expertises.
- 2010 - 2018** Founder and Advisor on Machine Learning and Search. ClassDo, Inc.
Japanese/Singaporean start-up focussing on online education.

Professional Activities

Government Scientific Advice and Policy Work

- 2020** Royal Society Data Evaluation and Learning for Viral Epidemics (DELVE) Initiative, UK.
<https://rs-delve.github.io/>.
Co-lead of action team (with Neil Lawrence).

2016 Royal Society Project Working Group on Machine Learning, UK.

Community organization

2020 ELLIS Unit Oxford.

Co-director (with Mike Wooldridge and Steven Roberts).

2019 ELLIS Programme on Robust Machine Learning.

Co-director (with Chris Holmes and Sami Kaski).

Supervision

Ph.D. students

2021 - Present Eduard Oravkin (co-supervised with Patrick Rebeschini).

2021 - Present Guneet Dhillon (co-supervised with Tom Rainforth, Arnaud Doucet, George Deligiannidis).

2020 - Present Desi Ivanova (co-supervised with Tom Rainforth).

2020 - Present Ning Miao (co-supervised with Tom Rainforth).

2020 - Present M. Faaiz Taufiq (co-supervised with Arnaud Doucet).

2020 - Present Yuyang Shi (co-supervised with Tom Rainforth and Arnaud Doucet).

2020 - Present Cong Lu (co-supervised with Michael Osborne).

2020 - Present Mrinank Sharma (co-supervised with Tom Rainforth and Eric Nalisnick).

2019 - Present Michael Hutchinson (co-supervised with Max Welling).

2019 - Present Bryn Elesedy (co-supervised with Varun Kanade).

2019 - Present Bobby He (co-supervised with Arnaud Doucet, George Deligiannidis).

2019 - Present Sheheryar Zaidi (co-supervised with Arnaud Doucet).

2018 - Present Jin Xu.

2018 - Present Emilien Dupont (co-supervised with Arnaud Doucet).

2018 - Present Charline Le Lan (co-supervised with Shimon Whiteson).

2018 - Present Jean-Francois Ton (co-supervised with Dino Sejdinovic).

2018 - Present Tim Rudner (co-supervised with Yarin Gal).

2018 - Present Joost van Amersfoort (co-supervised with Yarin Gal).

2017 - Present Adam Golinski (co-supervised with Tom Rainforth and Frank Wood)

2017 - 2021 Adam Foster. Postdoctoral Fellow, Microsoft Research Cambridge.

2017 - 2021 Emile Mathieu. Postdoctoral Fellow, Statistics, Oxford.

2017 - 2021 Bradley Gram-Hansen. CTO & Co-founder, Intelligent Networks.

2017 - 2021 Dominic Richards. Research Scientist, Amazon.

2017 - 2021 Yuan Zhou.

2016 - 2020 Xenia Miscouridou. Postdoctoral Fellow, Statistics, Imperial College London.

2016 - 2020 Giuseppe Di Benedetto. Machine Learning Scientist, Amazon.

2016 - 2020 Chris Maddison. Assistant Professor, Computer Science, Toronto.

2016 - 2020 Adam Kosiorek. Research Scientist at DeepMind.

2015 - 2019 Xiaoyu Lu. Research Scientist at Amazon Cambridge
2015 - 2019 Leonard Hasenclever. Research Scientist at DeepMind.
2015 - 2019 Jovana Mitrovic. Research Scientist at DeepMind.
2015 - 2019 Valerio Peronne. Research Scientist at Amazon Berlin.
2015 - 2019 Hyunjik Kim. Research Scientist at DeepMind.
2015 - 2019 Stefan Webb. Applied Machine Learning Researcher at Twitter Cortex.
2014 - 2016 Marco Battiston. Lecturer, Statistics, Lancaster University.
2014 - Present Frauke Harms.
2013 - 2018 Thibaut Lienart. Visiting research fellow, Australian National University.
2013 - 2018 Tamara Fernandez Aguilar. Assistant Professor, Adolfo Ibanez University, Chile.
2011 - 2016 Balaji Lakshminarayanan. Research Scientist, Google Brain.
2011 - 2017 María Lomelí. Research Scientist, Facebook.
2009 - 2015 Lloyd Elliott. Assistant Professor, Biostatistics, Simon Fraser University.
2008 - 2020 Jan Gasthaus. Amazon Berlin Machine Learning Science.
2008 - 2014 Charles Blundell. Research Scientist, DeepMind.
2008 - 2012 Dorota Glowacka. Postdoc, Computer Science, University of Helsinki.
2007 - 2012 Vinayak Rao. Associate Professor, Statistics, Purdue University.

Postdoctoral fellows

2017 - 2019 Tom Rainforth. Florence Nightingale Fellow, Statistics, University of Oxford and JRF, Christchurch College Oxford.
2017 - 2019 Ben Bloem-Reddy. Assistant Professor, Statistics, University of British Columbia.
2016 - 2018 Marco Battiston. Lecturer, Statistics, Lancaster University.
2015 - 2017 Seth Flaxman. Associate Professor, Computer Science, Oxford.
2014 - 2017 Konstantina Palla. Research Scientist, Microsoft Research Cambridge.
2014 - 2015 Levi Boyles. Applied Scientist, Microsoft Bing.
2013 - 2015 Pierre Jacob. Professor, Statistics, ESSEC Business School, France.
2013 - 2015 Sebastian Vollmer. Associate Professor, Data Science, University of Warwick.
2013 - 2015 François Caron. Associate Professor, Statistics, University of Oxford.
2011 - 2014 Nicolas Heess. Research Scientist, DeepMind.
2010 - 2013 Andriy Mnih. Research Scientist, DeepMind.
2008 - 2011 Katherine Heller. Research Scientist, Google Brain.
2007 - 2010 Dilan Görür. Research Scientist, DeepMind.
2007 - 2009 Frank Wood. Associate Professor, Computer Science, University of British Columbia.

Other students

- 2018** Chris Chen (summer intern).
- 2017** Jin Xu (summer intern).
- 2016** Lazar Valkov (MSc).
- 2015** Matej Balog (MMath&CS).
- 2014** Xiaoyu Lu (MMath&Stats).
- 2013** Minjie Xu (summer intern).
- 2013** Xingwei Zhu (MSc; co-supervised with Phil Blunsom).
- 2013** Ho-Joon Kim (MMath&Stats).
- 2013** Ugochukwu Okoroafor (MMath&Stats).
- 2011 - 2013** Sam Patterson (MPhil).
- 2011** Sam Patterson (MSc).
- 2011** Marius Cobzarenco (MSc).
- 2010** Krikamol Muandet (MSc).
- 2008** Jan Gasthaus (MSc; co-supervised with Frank Wood).
- 2007** Daniel Roy (summer intern)
- 2006** Cai Junfu (BSc; co-supervised with Wee Sun Lee).

Ph.D. examiner/committee member

- 2020** Gabrielle Abbati, Engineering, University of Oxford.
- 2020** Rob Cornish, Statistics, University of Oxford.
- 2018** Yishu Miao, Computer Science, University of Oxford.
- 2018** Willie Neiswanger, Machine Learning, Carnegie-Mellon University.
- 2016** Adrien Todeschini, INRIA Bordeaux - Sud-Ouest, Universite de Bordeaux.
- 2016** Creighton (Koa) Heaukulani, Engineering, University of Cambridge.
- 2016** Giacomo Zanella, Statistics, University of Warwick.
- 2014** Tjun Kiat Teo, Statistics, University of Oxford.
- 2014** Karl-Moritz Hermann, Computer Science, University of Oxford.
- 2012** Liangliang Wang, Statistics, University of British Columbia.
- 2012** David Knowles, Engineering, University of Cambridge.
- 2011** Carl Scheffler, Physics, University of Cambridge.
- 2011** Daniel Roy, Electrical Engineering/Computer Science, Massachusetts Institute of Technology.
- 2010** Shahid Iqbal, Statistics, University College London.
- 2008** Hanna Wallach, Physics, University of Cambridge.
- 2007** David Stern, Physics, University of Cambridge.

Habilitation a Diriger des Recherches examiner/committee member

- 2019** Julyan Arbel, Inria Grenoble.
- 2013** Francois Caron, Inria Bordeaux.

Scholarships and Awards

- 2021** Test of Time Award (with Max Welling).
Bayesian Learning via Stochastic Gradient Langevin Dynamics.
International Conference of Machine Learning.
- 2019** IMS Medallion Lecture.
Institute of Mathematical Statistics.
- 2009** Best Paper Award (with Frank Wood).
International Conference on Artificial Intelligence and Statistics.
- 1998, 2002** Ontario Graduate Scholarship (CDN\$11859, CDN\$15000).
Ministry of Training, Colleges and Universities, Ontario, Canada.
- 1999 - 2002** University of Toronto Fellowship (Tuition and living expenses).
University of Toronto, Canada.
- 1997** University of Toronto Open Masters Fellowship (CDN\$7200).
University of Toronto, Canada.

Grants

- 2018** DeepMind Scholarship.
GBP 300K over 4 years for two studentships.
- 2018** Oxford-Tencent Collaboration in Large Scale Machine Learning.
GBP 300K over 3 years for one studentship and 1 postdoc.
- 2016** DeepMind Scholarship.
GBP 150K over 4 years.
- 2016** Learning to Infer in Graphical Models and Probabilistic Programmes.
Microsoft Research PhD Scholarship.
GBP 110K over 3 years.
- 2015** Amazon AWS in Education Grant.
USD 10K in AWS credits.
- 2014** BigBayes: Rich, Structured and Efficient Learning of Big Bayesian Models.
European Research Council Consolidator Fellowship.
EUR 1.92 million over 5 years.
- 2013** Learning Compositional Semantic Relations.
Xerox Foundation Award.
EUR 90K over 3 years. co-PI (with 1 other).
- 2013** Bayesian Inference for Big Data with Stochastic Gradient Markov Chain Monte Carlo.
EPSRC. GBP 0.75M over 3 years. co-PI (along with 2 other co-PIs).
- 2011** Composing Learning for Artificial Cognitive Systems.
European Union STREP. EUR 4.5M over 4 years. co-PI (along with 9 other co-PIs).
- 2010** Renewal of funding for the Gatsby Computational Neuroscience Unit.
Gatsby Charitable Foundation.
GBP 6.9M over 5 years. co-PI (along with 4 other co-PIs).
- 2005** Lee Kuan Yew Postdoctoral Fellowship.
Lee Kuan Yew Endowment Fund, Singapore.
Salary and SGD178K grant over three years.

Scientific and Peer Reviewing Activities

Conference organization

- 2020** World Meeting of the International Society for Bayesian Analysis, Yunnan, China.
Scientific committee member.
- 2019** International Conference on Bayesian Nonparametrics, Oxford, UK.
Organizing member.
- 2017** International Conference on Machine Learning, Sydney, Australia.
Program co-chair (with Doina Precup).
- 2014** Machine Learning Summer School, Iceland.
Program co-chair (with Samuel Kaski and Jukka Corander).
- 2010** International Conference on Artificial Intelligence and Statistics, Sardinia, Italy.
Program co-chair (with Mike Titterton).

Workshop organization

- 2018** NeurIPS Workshop on Continual Learning.
Montreal, Canada.
Co-organizers: Razvan Pascanu, Marc Pickett, Mark Ring.
- 2018** NeurIPS Workshop on Critiquing and Correcting Trends in Machine Learning.
Montreal, Canada.
Co-organizers: Tom Rainforth, Matt Kusner, Ben Bloem-Reddy, Brooks Paige, Rich Caruana.
- 2016** ATI Scoping Workshop on High-Dimensional Statistical Models & Big Data: Methodology & Applications.
British Library, London, UK.
Co-organizers: Alex Beskos, Arnaud Doucet, Simon Godsill, Konstantinos Zygalakis, Sebastian Vollmer, Dan O Crisan, Semeetpal Sidhu Singh.
- 2015** ATI Scoping Workshop on Intractable Likelihoods.
University of Warwick, UK.
Co-organizers: Gareth Roberts, David Firth, Chris Holmes, Iain Murray, Alex Beskos, John Aston.
- 2015** ATI Scoping Workshop on Deep Learning.
University of Edinburgh, UK.
Co-organizers: Amos Storkey and Krzysztof Geras, Nando De Freitas, Ben Graham, Zoubin Ghahramani, Thore Graepel, Neil Lawrence, Phil Blunsom, Iain Murray, Stephen Roberts, Andrew Zisserman, Mark Gales, Vittorio Ferrari, Andrea Vedaldi, Steve Renals, Andrew Stuart, Charles Sutton, Richard Turner, Chris Williams, Emre Ozer, Max Welling.
- 2015** ATI Scoping Workshop on Statistical and Computational Challenges in Large-Scale Data Analysis.
University of Cambridge, UK.
Co-organizers: Yi Yu, Rajen Shah, Richard Samworth, Anthony Lee, Graham Cormode.
- 2015** Scalable Monte Carlo Methods for Bayesian Analysis of Big Data.
NIPS, Montreal, Canada.
Co-organizers: Babak Shahbaba, Max Welling, Christophe Andrieu, Arnaud Doucet, Sebastian Vollmer, Pierre Jacob.

- 2015** Bayesian Inference for Big Data.
Oxford, United Kingdom.
- 2011** Gatsby Unit/CIFAR Workshop on Hierarchical Distributed Representations.
Windsor, United Kingdom.
Co-organizers: Geoffrey E. Hinton and Peter Dayan.
- 2009** Nonparametric Bayes.
NIPS, Whistler, Canada.
Co-organizers: Dilan Görür, Francois Caron, David Dunson, Zoubin Ghahramani, Michael I. Jordan.
- 2009** Grammar Induction, Representation of Language and Language Learning.
NIPS, Whistler, Canada.
Co-organizers: Alex Clark, Dorota Glowacka, John Shawe-Taylor, Chris Watkins.
- 2008** Nonparametric Bayes.
ICML/UAI/COLT, Helsinki, Finland.
Co-organizers: Romain Thibaux, Athanasios Kottas, Zoubin Ghahramani, Michael I. Jordan.
- 2008** Prior knowledge for text and language processing.
ICML/UAI/COLT, Helsinki, Finland.
Co-organizers: Marc Dymetman, Guillaume Bouchard, Hal Daume III.
- 2005** Bayesian methods for natural language processing.
NIPS, Whistler, Canada.
Co-organizer: Hal Daume III.
- 2005** Open problems and challenges for nonparametric Bayesian methods in machine learning.
NIPS, Whistler, Canada.
Co-organizer: Matthew J. Beal.
- 2003** Nonparametric Bayesian methods and infinite models.
NIPS, Whistler, Canada.
Co-organizer: Matthew J. Beal.

Conference and workshop advisory board

- 2017 - 2019** International Machine Learning Society.
- 2016 - 2017** StatScale EPSRC Programme.
- 2015** NIPS Workshop on Bayesian Nonparametrics.
- 2015** NIPS Workshop on Advances in Approximate Bayesian Inference.
- 2011** NIPS Workshop on Bayesian Nonparametrics: Hope or Hype?
- 2015, 2011, 2009** Conference on Bayesian Nonparametrics.

Conference senior programme committee and session organization

- 2012** Session Organizer for International Conference of the ERCIM Working Group on Computing and Statistics.
- 2009** International Joint Conference on Artificial Intelligence.
- 2007 - Present** I am actively involved as area chair or senior area chair for Neural Information Processing Systems, International Conference on Machine Learning and International Conference on Artificial Intelligence and Statistics over the years.

Conference reviewer

Artificial Intelligence and Statistics
Computer Vision and Pattern Recognition
Empirical Methods on Natural Language Processing
International Conference on Machine Learning
International Joint Conference on Artificial Intelligence
Neural Information Processing Systems
Uncertainty in Artificial Intelligence

Associate/action editor

2016 Electronic Journal of Statistics Special Issue on Bayesian Nonparametrics
2015 IEEE TPAMI Special Issue on Bayesian Nonparametrics
2014 - 2016 Statistical Sciences
2013 - 2017 Journal of Machine Learning Research
2011 - 2016 Journal of the Royal Statistical Society B
2011 - 2013 IEEE Transactions on Pattern Analysis and Machine Intelligence
2011 - 2013 Machine Learning Journal
2009 - 2012 Bayesian Analysis

Journal reviewer

Artificial Intelligence Journal
Annals of the Institute of Statistical Mathematics
Bayesian Analysis
Biometrika
Genetics
IEEE Signal Processing Magazine
IEEE Transactions on Image Processing
IEEE Transactions on Knowledge and Data Engineering
IEEE Transactions on Neural Networks
IEEE Transactions on Pattern Analysis and Machine Intelligence
IEEE Transactions on Speech, Audio, and Language Processing
International Journal of Neural Systems
Journal of Computational and Graphical Statistics
Journal of Machine Learning Research
Journal of the American Statistical Association
Machine Learning Journal
Neural Computation
Statistics and Computing

Grant reviewer

2012 - 2016 Peer review college, Engineering and Physical Sciences Research Council, UK
Biotechnology and Biological Sciences Research Council, UK
Engineering and Physical Sciences Research Council, UK

Leverhulme Trust, UK
National Research Foundation, Singapore
National Science Foundation, USA
Netherlands Organisation for Scientific Research, The Netherlands

Teaching

Qualifications

2011 Postgraduate Certificate in Learning and Teaching in Higher Education

Courses

- 2020** Advanced Topics in Machine Learning.
Department of Statistics, University of Oxford.
Cross-listed Part C (fourth year undergraduate) and MSc Applied Statistics course.
- 2012 - 2014** Statistical Machine Learning and Data Mining.
Department of Statistics, University of Oxford.
Cross-listed Part C (fourth year undergraduate) and MSc Applied Statistics course.
- 2012 - 2013** Simulation.
Department of Statistics, University of Oxford.
Part A (second year undergraduate) course.
- 2008 - 2012** Machine Learning II.
Gatsby Computational Neuroscience Unit, University College London.
Ph.D. level course.
Course on a number of advanced topics in machine learning.
- 2007 - 2013** Unsupervised and Probabilistic Learning (co-taught with Maneesh Sahani).
Gatsby Computational Neuroscience Unit, University College London.
M.Sc. and Ph.D. level course.
Introduction to probabilistic and graphical models for unsupervised learning.

Tutorials and short courses

- 2020** Tutorial on meta-learning.
Machine Learning Summer School, Tuebingen.
- 2019-2020** Two week module on Machine Learning.
StatML.IO Centre for Doctoral Training, University of Oxford.
- 2016** Short Course on Big Data and Bayesian Learning.
Greek Stochastics θ , Greece.
- 2014 - 2016, 2018** Two week module on Machine Learning.
OxWASP Centre for Doctoral Training, University of Oxford.
- 2014** Day Courses on Markov Chains and Monte Carlo Markov Chain Methods.
Doctoral Training Centre, University of Oxford.
- 2013** Day Course on Monte Carlo Markov Chain Methods.
Doctoral Training Centre, University of Oxford.

- 2013** Tutorial on Bayesian nonparametrics.
Machine Learning Summer School, Tuebingen.
- 2012** Tutorial on Bayesian nonparametrics.
Abdus Salam International Centre for Theoretical Physics Summer School, Italy.
- 2011** Tutorial on modern Bayesian nonparametrics.
Neural Information Processing Systems, Granada, Spain.
- 2011** Tutorial on Bayesian nonparametrics.
Machine Learning Summer Schools,
Singapore (July) and Bordeaux, France (September).
- 2011** Bayesian nonparametrics.
PhD Course at Section for Cognitive Systems,
Denmark Technical University, Informatics, Denmark.
- 2010** Probabilistic and Bayesian Modelling.
Signal Theory and Communications Department, University Carlos III in Madrid.
Week-long graduate-level intensive course.
- 2010** Tutorial on Bayesian nonparametrics and applications.
Korean Advanced Institute of Science and Technology, South Korea.
- 2010** Tutorial on Bayesian nonparametrics and applications.
Workshop on Probabilistic Modeling for Computer Vision Applications.
Center for Mathematical Investigations, Mexico.
- 2009** Tutorial on Bayesian nonparametrics.
Machine Learning Summer School, University of Cambridge, UK.
Microsoft Research Cambridge, UK.
University of Toronto, Canada.
- 2008 - 2010** Clustering: tutorial and practical course.
Adaptive Modelling of Complex Data. University College London.
- 2008** Tutorial on hierarchical clustering.
EPSRC Winter School on Mathematics for Data Modelling, University of Sheffield, UK.
- 2007** Tutorial on Dirichlet processes and hierarchical Dirichlet processes.
Machine Learning Advanced Tutorial Series, University of Cambridge, UK.
- 2007** Bayesian nonparametric models: applications in machine learning and beyond.
Issac Newton Institute Workshop on Bayesian Nonparametric Regression, UK.
- 2007** Dirichlet processes: tutorial and practical course.
Machine Learning Summer School.
Max Planck Institute for Biological Cybernetics, Tübingen, Germany.
- 2006** Dirichlet processes.
Max Planck Institute for Biological Cybernetics Retreat, Tübingen, Germany.

Invited Conference and Workshop Talks

- 2020** NordStat, Sweden (**keynote**).
- 2019** Conference on Uncertainty in Artificial Intelligence (UAI), Tel Aviv, Israel (**keynote**)
- 2019** Institute of Mathematical Statistics (IMS) Meeting, Denver, USA. (**medallion lecture**)

- 2018** Knowledge Discovery and Data Mining (KDD), London, UK. **(keynote)**
- 2018** Bayesian Young Statisticians Meeting (BAYSM), Warwick, UK. **(keynote)**
- 2018** Bayesian Nonparametrics Workshop, Bordeaux, France. **(keynote)**
- 2018** Bayesian Computation, Barcelona, Spain.
- 2017** Neural Information Processing Systems (NIPS), Long Beach, USA. **(keynote)**
- 2017** Machine Learning and Signal Processing (MLSP), Tokyo, Japan. **(keynote)**
- 2017** OxWaSP Machine Learning Symposium, Warwick, UK.
- 2016** Advanced Computational methods for Complex Models in Biology Workshop, London, UK.
- 2016** EPSRC Network for Computational Statistics and Machine Learning Workshop: Data Science: pulling together computational statistics and machine learning? Edinburgh, UK.
- 2016** Newton Institute Workshop on Bayesian Methods for Networks, Cambridge, UK.
- 2016** Intractable Likelihoods Workshop, Lancaster, UK.
- 2016** World Meeting of the International Society for Bayesian Analysis, Italy.
- 2015** NIPS Workshop on Nonparametric Methods for Large Scale Representation Learning, Canada.
- 2015** NIPS Workshop on Advances in Approximate Bayesian Inference (Panel), Canada.
- 2015** NIPS Workshop on Bayesian Nonparametrics: The Next Generation (Panel), Canada.
- 2015** ATI Open Workshop on Deep Learning, UK.
- 2015** ATI Scoping Workshop on Deep Learning, UK.
- 2015** ATI Scoping Workshop on Intractable Likelihoods, UK.
- 2015** ATI Scoping Workshop on Distributed Machine Learning and Optimization, UK.
- 2015** Conference on Bayesian Nonparametrics, USA.
- 2015** Objective Bayes Workshop, Spain.
- 2015** Symposium on Big Data and Predictive Computational Modeling, Germany.
- 2014** UCL-Duke University Workshop on Sensing and Analysis of High-Dimensional Data, UK.
- 2014** Isaac Newton Institute Workshop on Monte Carlo Inference for High-Dimensional Statistical Models, UK.
- 2014** MCMSki, France.
- 2013** Banff Workshop on Random Measures and Measure-Valued Processes.
Banff International Research Station for Mathematical Innovation and Discovery.
- 2013** Royal Statistical Society Meeting celebrating 250 years since the reading of Bayes' paper, UK.
- 2013** Institut des Hautes Etudes Scientifiques (IHES) Workshop, France.
- 2012** Institute of Mathematical Sciences Programme on Meeting the Challenges of High Dimension—Statistical Methodology, Theory and Applications, Singapore.
- 2012** Joint Statistical Meetings, San Diego, USA.
- 2012** 2nd Institute of Mathematical Statistics Asian Pacific Rim Meeting, Tsukuba, Japan.
- 2012** International Society for Bayesian Analysis World Meeting, Kyoto, Japan.
- 2011** Efficient MCMC for Continuous Time Discrete State Systems.
Bayes-250 Workshop, Edinburgh, UK.
- 2011** Bayesian Tools for Natural Language Learning.
Conference on Natural Language Learning, Oregon, USA. **(keynote)**

- 2011** Modelling Genetic Variations using Fragmentation-Coagulation Processes.
Bayesian Nonparametrics Workshop, Veracruz, Mexico.
- 2010** MCMC for Bayesian Nonparametric Models: A Bag of Tricks.
NIPS Workshop on Monte Carlo Methods, Canada.
- 2010** Hierarchical Bayesian Nonparametric Models for Language and Text (**keynote**).
ACM-SIGIR Workshop on Feature Generation and Selection for Information Retrieval, Switzerland.
- 2010** Hierarchical Bayesian Nonparametric Models for Language and Text.
International Society for Bayesian Analysis World Meeting, Valencia, Spain.
- 2010** Bayesian Rose Trees.
CRiSM Workshop on Model Uncertainty, Warwick, UK.
- 2010** Bayesian Rose Trees.
Workshop in Mixture Estimation and Applications, ICMS, Edinburgh, UK.
- 2009** Nonparametric Surrogate Priors.
Gatsby Deep Learning Workshop, London, UK.
- 2009** Bayesian Nonparametrics in Document and Language Modelling.
European Workshop on Challenges in Modern Massive Data Sets, Copenhagen, Denmark.
- 2009** The Mondrian Process: Bayesian Nonparametrics for Relational Data.
Bayesian Nonparametrics Workshop, Turin, Italy.
- 2008** Collapsed Variational Inference for Infinite State Bayesian Networks.
London Mathematical Society Durham Symposium on Mathematical Aspects of Graphical Models, UK.
- 2008** Bayesian Language Models.
CRiSM/INI Workshop on Bayesian Analysis of High Dimensional Data, Warwick, UK.
- 2007** Setting the Stage: Complementary Priors and Variational Bounds.
NIPS Deep Learning Symposium, Vancouver, Canada.
- 2007** Bayesian Agglomerative Clustering with Coalescents.
PASCAL Entente-Cordiale Workshop, Paris, France.
- 2007** A Hierarchical Bayesian Language Model based on Pitman-Yor Processes.
PASCAL Entente-Cordiale Workshop, London, UK.
- 2006** Some Applications of Hierarchical Dirichlet Processes (**keynote**).
Bayesian Nonparametrics Workshop, Jeju, Korea.
- 2006** Beam Sampling for the Infinite Hidden Markov Model (**keynote**).
ICML Workshop on Bayesian Nonparametric Models, Pittsburgh, USA.
- 2004** Hierarchical Dirichlet Processes.
Snowbird Learning Workshop, Utah, USA.
- 2002** Automatic Alignment of Local Representations.
NIPS Workshop on Spectral Methods, Vancouver, Canada.
- 2001** Passing and Bouncing Messages for Generalized Inference.
Gatsby Approximate Inference Workshop, London, UK.

Invited Seminars

- 2017** Vector Institute, Toronto, Canada.
Statistics, Lancaster University, UK.
- 2016** Computational Statistics and Machine Learning, UCL, UK.
- 2015** Intractable likelihoods (i-like) programme seminar, UK.
Statistics and Mathematics, Vienna University of Economics and Business, Austria.
Statistics, National University of Singapore, Singapore.
Statistics, University of Bristol.
- 2014** Computer Science, University of Toronto, Canada.
Statistics, University of Toronto, Canada.
Statistics, Columbia University, USA.
Statistics, University of Texas at Austin, USA.
Statistics, University of Cambridge, UK.
SNN Adaptive Intelligence, Radboud University Nijmegen, The Netherlands.
- 2013** Statistics, Duke University, USA.
Statistics, University of Newcastle, UK.
Information Engineering, University of Oxford, UK.
Computer Science, University of Oxford, UK.
Statistics, University of Glasgow, UK.
- 2012** Computing, Macquarie University, Australia.
Computer Science, University of Toronto, Canada.
Mathematical Sciences, Queen Mary University of London, UK.
Computing Lab, Cambridge University, UK.
Statistics, Oxford University, UK.
Statistics, Warwick University, UK.
- 2011** Statistics and Applied Mathematics, University of Turin, Italy.
Xerox Research Centre Europe, France.
Statistics, University of Bristol, UK.
Computer Science and Information Systems, Birkbeck College, London, UK.
Decision Sciences, Bocconi University, Italy.
Wellcome Trust Centre for Human Genetics, University of Oxford, UK.
Engineering Science, University of Oxford, UK.
Computing, University of Oxford, UK.
CSAIL, Massachusetts Institute of Technology, USA.
Engineering, University of Cambridge, UK.
Institute of Statistical Mathematics, Japan.
Tokyo Institute of Technology, Japan.
NTT Research Labs, Japan.
- 2010** Defence Science Organization (DSO), Singapore.
Informatics, University of Edinburgh, UK.
Google Research and Computer Science, University of Waterloo, Canada.
Computer Science, University of Southern California, USA.
Salk Institute, University of Southern California, USA.
Information and Computer Science, University of California at Irvine, USA.

- Computer Science, University of Glasgow, UK.
 Statistics, University of Leeds, UK.
 Statistics, UCL, UK.
 Computer Science, University of Sheffield, UK.
- 2009** Hong Kong University of Science and Technology, Hong Kong.
 Computer Science, University of Sheffield, UK.
 Engineering, University of Cambridge, UK.
 Computer Science, University of Birmingham, UK.
- 2008** Atomic Energy and Alternative Energies Commission, Saclay, France.
 Yahoo! Research Silicon Valley, USA.
 Google Research, USA.
 NEC Laboratories America, USA.
 Information and Computer Science, University of California at Irvine, USA.
 Computer Science, University of California at Berkeley, USA.
 Engineering, University of Cambridge, UK.
 Computer Science, Royal Holloway University of London, UK.
 Computer Science, University of Bristol, UK.
 School of Computing, National University of Singapore, Singapore.
 Nanyang Technological University, Singapore.
- 2007** SNN Adaptive Intelligence, Radboud University Nijmegen, The Netherlands.
 Computer Science, University of Toronto, Canada.
 Computer Science, University of Manchester, UK.
 Mathematics, Statistics and Actuarial Science, University of Kent, UK.
 Informatics, University of Edinburgh, UK.
- 2006** Information and Computer Science, University of California at Irvine, USA.
 Computer Science and Engineering, State University of New York at Buffalo, USA.
- 2005** Computer Science, University of Toronto, Canada.
- 2004** Microsoft Research Redmond, USA.
 Intel Research, USA.
 Computer Science, University of Pennsylvania, USA.
 Computer Science, University of Massachusetts Amherst, USA.
 CSAIL, Massachusetts Institute of Technology, USA.
 Computer Science, New York University, USA.
 Computer Science, Columbia University, USA.
 Computer Science, University of Toronto, Canada.
 School of Computing, National University of Singapore, Singapore.
- 2003** Honda Research Institute, USA. Computer Science, University of Toronto, Canada.
 Information and Computer Science, University of California at Irvine, USA.
- 2002** Redwood Neuroscience Institute, Menlo Park, USA.
 Computer Science, University of Toronto, Canada.
 Gatsby Computational Neuroscience Unit, UCL, UK.
 Computer Science, Carnegie Mellon University, USA.
- 2001** Computer Science, University of Waterloo, Canada.

Publications

Up-to-date publication list at Google Scholar ([Profile](#)). Citation count: > 37000, h-index: 63.

Invited Journal Papers

1. F. Wood, J. Gasthaus, C. Archambeau, L. James, and Y. W. Teh. The sequence memoizer. *Communications of the Association for Computing Machines*, 54(2):91–98, 2011. [url](#). [pdf](#).

Refereed Journal Papers

2. B. He, S. Zaidi, B. Elesedy, M. Hutchinson, A. Paleyes, G. Harling, A. M. Johnson, Y. Whye Teh, and n. null. Effectiveness and resource requirements of test, trace and isolate strategies for COVID in the UK. *Royal Society Open Science*, 8(3):201491, 2021. [doi](#). [url](#).
3. J. M. Brauner, S. Mindermann, M. Sharma, D. Johnston, J. Salvatier, T. Gavenčiak, A. B. Stephenson, G. Leech, G. Altman, V. Mikulik, A. J. Norman, J. T. Monrad, T. Besiroglu, H. Ge, M. A. Hartwick, Y. W. Teh, L. Chindelevitch, Y. Gal, and J. Kulveit. Inferring the effectiveness of government interventions against COVID-19. *Science*, 2020. ISSN 0036-8075. [doi](#). [url](#).
4. N. Tomašev, J. Cornebise, F. Hutter, S. Mohamed, A. Picciariello, B. Connelly, D. C. Belgrave, D. Ezer, F. C. van der Haert, F. Mugisha, et al. AI for social good: unlocking the opportunity for positive impact. *Nature Communications*, 11(1):1–6, 2020.
5. R. Schwessinger, M. Gosden, D. Downes, R. C. Brown, A. M. Oudelaar, J. Telenius, Y. W. Teh, G. Lunter, and J. R. Hughes. DeepC: predicting 3D genome folding using megabase-scale transfer learning. *Nature Methods*, 2020. [doi](#). [url](#).
6. Q. Wang, V. Rao, and Y. W. Teh. An exact auxiliary variable gibbs sampler for a class of diffusions. *Journal of Computational and Graphical Statistics*, 0(0):1–15, 2020. [doi](#). [url](#). [pdf](#).
7. B. Bloem-Reddy and Y. Teh. Probabilistic symmetries and invariant neural networks. *Journal of Machine Learning Research*, 21(90):1–61, 2020. [url](#). [pdf](#).
8. L. T. Elliott, M. De Iorio, S. Favaro, K. Adhikari, and Y. W. Teh. Modeling population structure under hierarchical dirichlet processes. *Bayesian Analysis*, 2018. [doi](#). [url](#).
9. M. Battiston, S. Favaro, D. M. Roy, and Y. W. Teh. A characterization of product-form exchangeable feature probability functions. *Annals of Applied Probability*, 28(3):1423–1448, 06 2018. [doi](#). [url](#).
10. M. Lomeli, S. Favaro, and Y. W. Teh. A marginal sampler for σ -stable Poisson-Kingman mixture models. *Journal of Computational and Graphical Statistics*, 2017. [doi](#). [url](#). [pdf](#).
11. L. Hasenclever, S. Webb, T. Lienart, S. Vollmer, B. Lakshminarayanan, C. Blundell, and Y. W. Teh. Distributed Bayesian learning with stochastic natural-gradient expectation propagation and the posterior server. *Journal of Machine Learning Research (JMLR)*, 2017. [url](#). [pdf](#).
12. V. Perrone, P. A. Jenkins, D. Spano, and Y. W. Teh. Poisson random fields for dynamic feature models. *Journal of Machine Learning Research (JMLR)*, 2017. [url](#). [pdf](#).
13. S. Flaxman, Y. W. Teh, and D. Sejdinovic. Poisson intensity estimation with reproducing kernels. *Electronic Journal of Statistics*, 11(2):5081–5104, 2017a. [doi](#). [url](#).

14. M. Battiston, S. Favaro, and Y. W. Teh. Multi-armed bandit for species discovery: A bayesian nonparametric approach. *Journal of the American Statistical Association*, 2016. doi: . url: . pdf: .
15. J. Arbel, S. Favaro, B. Nipoti, and Y. W. Teh. Bayesian nonparametric inference for discovery probabilities: credible intervals and large sample asymptotics. *Statistica Sinica*, 2017. doi: . url: . pdf: .
16. S. Favaro, A. Lijoi, C. Nava, B. Nipoti, I. Prüenster, and Y. W. Teh. On the stick-breaking representation for homogeneous NRMI. *Bayesian Analysis*, 11:697–724, 2016. doi: . url: . pdf: .
17. S. J. Vollmer, K. C. Zygalakis, and Y. W. Teh. Exploration of the (non-)asymptotic bias and variance of stochastic gradient Langevin dynamics. *Journal of Machine Learning Research (JMLR)*, 2016. url: . pdf: .
18. Y. W. Teh, A. H. Thiéry, and S. J. Vollmer. Consistency and fluctuations for stochastic gradient Langevin dynamics. *Journal of Machine Learning Research (JMLR)*, 2016. url: . pdf: .
19. L. T. Elliott and Y. W. Teh. A nonparametric HMM for genetic imputation and coalescent inference. *Electronic Journal of Statistics*, 2016. url: . pdf: .
20. S. Favaro, B. Nipoti, and Y. W. Teh. Rediscovery of Good-Turing estimators via Bayesian nonparametrics. *Biometrics*, 2015b. doi: . url: . pdf: .
21. S. Favaro, B. Nipoti, and Y. W. Teh. Random variate generation for Laguerre-type exponentially tilted α -stable distributions. *Electronic Journal of Statistics*, 9:1230–1242, 2015a. doi: . url: . pdf: .
22. P. G. Moreno, A. Artés-Rodríguez, Y. W. Teh, and F. Perez-Cruz. Bayesian nonparametric crowdsourcing. *Journal of Machine Learning Research (JMLR)*, 2015. url: . pdf: .
23. S. Favaro, M. Lomeli, B. Nipoti, and Y. W. Teh. On the stick-breaking representation of σ -stable Poisson-Kingman models. *Electronic Journal of Statistics*, 8:1063–1085, 2014a. doi: . url: . pdf: .
24. S. Favaro, M. Lomeli, and Y. W. Teh. On a class of σ -stable Poisson-Kingman models and an effective marginalized sampler. *Statistics and Computing*, 2014b. doi: . url: .
25. F. Caron, Y. W. Teh, and B. T. Murphy. Bayesian nonparametric Plackett-Luce models for the analysis of preferences for college degree programmes. *Annals of Applied Statistics*, 8(2): 1145–1181, 2014. doi: . url: . pdf: .
26. V. Rao and Y. W. Teh. Fast MCMC sampling for Markov jump processes and extensions. *Journal of Machine Learning Research (JMLR)*, 14:3295–3320, 2013. url: . pdf: .
27. S. Favaro and Y. W. Teh. MCMC for normalized random measure mixture models. *Statistical Science*, 28(3):335–359, 2013. url: . pdf: .
28. D. Görür and Y. W. Teh. Concave-convex adaptive rejection sampling. *Journal of Computational and Graphical Statistics*, 2011. doi: . url: .
29. Y. W. Teh, M. I. Jordan, M. J. Beal, and D. M. Blei. Hierarchical Dirichlet processes. *Journal of the American Statistical Association*, 101(476):1566–1581, 2006. url: . pdf: .
30. G. E. Hinton, S. Osindero, and Y. W. Teh. A fast learning algorithm for deep belief networks. *Neural Computation*, 18(7):1527–1554, 2006a. url: . pdf: .

31. G. E. Hinton, S. Osindero, M. Welling, and Y. W. Teh. Unsupervised discovery of non-linear structure using contrastive backpropagation. *Cognitive Science*, 30(4):725–731, 2006b. [url](#). [pdf](#).
32. M. Welling and Y. W. Teh. Linear response algorithms for approximate inference in graphical models. *Neural Computation*, 16:197–221, 2004. [url](#). [pdf](#).
33. Y. W. Teh, M. Welling, S. Osindero, and G. E. Hinton. Energy-based models for sparse overcomplete representations. *Journal of Machine Learning Research (JMLR)*, 4:1235–1260, 2003. [url](#). [pdf](#).
34. M. Welling and Y. W. Teh. Approximate inference in Boltzmann machines. *Artificial Intelligence*, 143(1):19–50, 2003. [url](#). [pdf](#).

Short Discussions

35. Y. W. Teh. On statistical thinking in deep learning. In *IMS Bulletin*, August 2019. [url](#). [pdf](#).
36. Y. W. Teh. Bayesian nonparametric modelling and the ubiquitous Ewens sampling formula. *Statistical Science*, 31(1):34–36, 2016. [url](#). [pdf](#).
37. M. Welling, Y. W. Teh, C. Andrieu, J. Kominiarczuk, T. Meeds, B. Shahbaba, and S. Vollmer. Bayesian inference and big data: A snapshot from a workshop. *ISBA Bulletin*, 2014. [pdf](#).

Review Articles

38. P. Orbanz and Y. W. Teh. Bayesian nonparametric models. In *Encyclopedia of Machine Learning*. Springer, 2010. [pdf](#).
39. Y. W. Teh. Dirichlet processes. In *Encyclopedia of Machine Learning*. Springer, 2010. [pdf](#).
40. Y. W. Teh and M. I. Jordan. Hierarchical Bayesian nonparametric models with applications. In N. Hjort, C. Holmes, P. Müller, and S. Walker, editors, *Bayesian Nonparametrics*. Cambridge University Press, 2010. [pdf](#).

Book Chapters

41. M. De Iorio, S. Favaro, and Y. W. Teh. Bayesian inference on population structure: From parametric to nonparametric modeling. In *Nonparametric Bayesian Inference in Biostatistics*. Springer, 2015. [doi](#). [url](#).
42. C. Blundell, Y. W. Teh, and K. A. Heller. Discovering non-binary hierarchical structures with Bayesian rose trees. In C. P. Robert, K. Mengersen, and M. Titterton, editors, *Mixture Estimation and Applications*. John Wiley & Sons, 2011. [url](#). [pdf](#).

Invited Conference Papers

43. G. E. Hinton, M. Welling, Y. W. Teh, and S. Osindero. A new view of ICA. In *Proceedings of the International Conference on Independent Component Analysis and Blind Signal Separation*, volume 3, 2001. [pdf](#).
44. G. E. Hinton and Y. W. Teh. Discovering multiple constraints that are frequently approximately satisfied. In *International Conference on Uncertainty in Artificial Intelligence (UAI)*, volume 17, pages 227–234, 2001. [pdf](#).

Peer-Reviewed Conference Papers

45. S. L. Chau, J.-F. Ton, J. González, Y. W. Teh, and D. Sejdinovic. Bayesimp: Uncertainty quantification for causal data fusion. In *Advances in Neural Information Processing Systems (NeurIPS)*, December 2021. [url](#). [pdf](#).
46. J. Xu, H. Kim, T. Rainforth, and Y. W. Teh. Group equivariant subsampling. In *Advances in Neural Information Processing Systems (NeurIPS)*, December 2021. [url](#). [pdf](#).
47. S. Zaidi, A. Zela, T. Elsken, C. C. Holmes, F. Hutter, and Y. W. Teh. Neural ensemble search for uncertainty estimation and dataset shift. In *Advances in Neural Information Processing Systems (NeurIPS)*, December 2021. [url](#). [pdf](#).
48. M. Hutchinson, A. Terenin, V. Borovitskiy, S. Takao, Y. Teh, and M. Deisenroth. Vector-valued gaussian processes on riemannian manifolds via gauge independent projected kernels. In *Advances in Neural Information Processing Systems (NeurIPS)*, December 2021a. [url](#). [pdf](#).
49. T. G. J. Rudner, C. Lu, M. A. Osborne, Y. Gal, and Y. W. Teh. On pathologies in kl-regularized reinforcement learning from expert demonstrations. In *Advances in Neural Information Processing Systems (NeurIPS)*, December 2021. [url](#). [pdf](#).
50. E. Mathieu, A. Foster, and Y. W. Teh. On contrastive representations of stochastic processes. In *Advances in Neural Information Processing Systems (NeurIPS)*, December 2021. [url](#). [pdf](#).
51. J. Schwarz, S. Jayakumar, R. Pascanu, P. E. Latham, and Y. W. Teh. Powerpropagation: A sparsity inducing weight reparameterisation. In *Advances in Neural Information Processing Systems (NeurIPS)*, December 2021. [url](#). [pdf](#).
52. M. J. Hutchinson, C. Le Lan, S. Zaidi, E. Dupont, Y. W. Teh, and H. Kim. LieTransformer: Equivariant self-attention for lie groups. In *International Conference on Machine Learning (ICML)*, 18–24 Jul 2021b. [url](#). [pdf](#).
53. P. Holderrieth, M. J. Hutchinson, and Y. W. Teh. Equivariant learning of stochastic fields: Gaussian processes and steerable conditional neural processes. In *International Conference on Machine Learning (ICML)*, 18–24 Jul 2021. [url](#). [pdf](#).
54. J.-F. Ton, L. CHAN, Y. W. Teh, and D. Sejdinovic. Noise contrastive meta-learning for conditional density estimation using kernel mean embeddings. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*. PMLR, 13–15 Apr 2021. [url](#). [pdf](#).
55. S. Hayou, J.-F. Ton, A. Doucet, and Y. W. Teh. Robust pruning at initialization. In *International Conference on Learning Representations (ICLR)*, 2021. [url](#).
56. B. He, B. Lakshminarayanan, and Y. W. Teh. Bayesian deep ensembles via the neural tangent kernel. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2020.
57. J. Lee, Y. Lee, J. Kim, E. Yang, S. J. Hwang, and Y. W. Teh. Bootstrapping neural processes. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2020.
58. M. Sharma, S. Mindermann, J. Brauner, G. Leech, A. Stephenson, T. Gavenčiak, J. Kulveit, Y. W. Teh, L. Chindelevitch, and Y. Gal. How robust are the estimated effects of nonpharmaceutical interventions against COVID-19? In *Advances in Neural Information Processing Systems (NeurIPS)*, 2020.
59. M. K. Titsias, J. Schwarz, A. G. de G. Matthews, R. Pascanu, and Y. W. Teh. Functional regularisation for continual learning with gaussian processes. In *International Conference on Learning Representations*, 2020. [url](#).

60. S. M. Jayakumar, W. M. Czarnecki, J. Menick, J. Schwarz, J. Rae, S. Osindero, Y. W. Teh, T. Harley, and R. Pascanu. Multiplicative interactions and where to find them. In *International Conference on Learning Representations*, 2020. [url](#). [pdf](#).
61. A. Foster, M. Jankowiak, M. O’Meara, Y. W. Teh, and T. Rainforth. A unified stochastic gradient approach to designing bayesian-optimal experiments. In S. Chiappa and R. Calandra, editors, *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pages 2959–2969. PMLR, 26–28 Aug 2020. [url](#). [pdf](#).
62. G. D. Benedetto, F. Caron, and Y. W. Teh. Non-exchangeable feature allocation models with sublinear growth of the feature sizes. In S. Chiappa and R. Calandra, editors, *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pages 3208–3218. PMLR, 26–28 Aug 2020. [url](#). [pdf](#).
63. U. Simsekli, L. Zhu, Y. W. Teh, and M. Gurbuzbalaban. Fractional underdamped Langevin dynamics: Retargeting SGD with momentum under heavy-tailed gradient noise. In H. D. III and A. Singh, editors, *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 8970–8980. PMLR, 13–18 Jul 2020. [url](#). [pdf](#).
64. J. Van Amersfoort, L. Smith, Y. W. Teh, and Y. Gal. Uncertainty estimation using a single deep deterministic neural network. In H. D. III and A. Singh, editors, *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 9690–9700. PMLR, 13–18 Jul 2020. [url](#). [pdf](#).
65. J. Xu, J.-F. Ton, H. Kim, A. Kosiorek, and Y. W. Teh. MetaFun: Meta-learning with iterative functional updates. In H. D. III and A. Singh, editors, *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 10617–10627. PMLR, 13–18 Jul 2020. [url](#). [pdf](#).
66. Y. Zhou, H. Yang, Y. W. Teh, and T. Rainforth. Divide, conquer, and combine: a new inference strategy for probabilistic programs with stochastic support. In H. D. III and A. Singh, editors, *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 11534–11545. PMLR, 13–18 Jul 2020. [url](#). [pdf](#).
67. E. Dupont, A. Doucet, and Y. W. Teh. Augmented neural ODEs. In *Advances in Neural Information Processing Systems (NeurIPS)*. 8–14 Dec 2019. [url](#). [pdf](#).
68. D. Rao, F. Visin, A. Rusu, R. Pascanu, Y. W. Teh, and R. Hadsell. Continual unsupervised representation learning. In *Advances in Neural Information Processing Systems (NeurIPS)*. 8–14 Dec 2019. [url](#). [pdf](#).
69. S. Ge, S. Wang, Y. W. Teh, L. Wang, and L. Elliott. Random tessellation forests. In *Advances in Neural Information Processing Systems (NeurIPS)*. 8–14 Dec 2019. [url](#). [pdf](#).
70. E. Mathieu, C. Le Lan, C. J. Maddison, R. Tomioka, and Y. W. Teh. Continuous hierarchical representations with Poincaré variational auto-encoders. In *Advances in Neural Information Processing Systems (NeurIPS)*. 8–14 Dec 2019a. [url](#). [pdf](#).
71. A. Foster, M. Jankowiak, E. Bingham, P. Horsfall, Y. W. Teh, T. Rainforth, and N. Goodman. Variational bayesian optimal experimental design. In *Advances in Neural Information Processing Systems (NeurIPS)*. 8–14 Dec 2019. [url](#). [pdf](#).

72. A. Kosiorek, S. Sabour, Y. W. Teh, and G. E. Hinton. Stacked capsule autoencoders. In *Advances in Neural Information Processing Systems (NeurIPS)*. 8–14 Dec 2019. [url](#). [pdf](#).
73. T. A. Le, A. R. Kosiorek, N. Siddharth, Y. W. Teh, and F. Wood. Revisiting reweighted wake-sleep for models with stochastic control flow. In *International Conference on Uncertainty in Artificial Intelligence (UAI)*, 22–25 July 2019. [pdf](#). [nick](#)
74. J. Lee, Y. Lee, J. Kim, A. Kosiorek, S. Choi, and Y. W. Teh. Set transformer: A framework for attention-based permutation-invariant neural networks. In *International Conference on Machine Learning (ICML)*, volume 97 of *Proceedings of Machine Learning Research*. PMLR, 09–15 Jun 2019. [url](#). [pdf](#).
75. E. Mathieu, T. Rainforth, N. Siddharth, and Y. W. Teh. Disentangling disentanglement in variational autoencoders. In *International Conference on Machine Learning (ICML)*, volume 97. PMLR, 09–15 Jun 2019b. [url](#). [pdf](#).
76. J. Merel, L. Hasenclever, A. Galashov, A. Ahuja, V. Pham, G. Wayne, Y. W. Teh, and N. Heess. Neural probabilistic motor primitives for humanoid control. In *International Conference on Learning Representations (ICLR)*, 6–9 May 2019. [url](#). [pdf](#).
77. E. Nalisnick, A. Matsukawa, Y. W. Teh, D. Gorur, and B. Lakshminarayanan. Do deep generative models know what they don’t know? In *International Conference on Learning Representations (ICLR)*, 6–9 May 2019. [url](#). [pdf](#).
78. A. Galashov, S. Jayakumar, L. Hasenclever, D. Tirumala, J. Schwarz, G. Desjardins, W. M. Czarnecki, Y. W. Teh, R. Pascanu, and N. Heess. Information asymmetry in KL-regularized RL. In *International Conference on Learning Representations (ICLR)*, 6–9 May 2019. [url](#). [pdf](#).
79. S. Webb, T. Rainforth, Y. W. Teh, and M. P. Kumar. A statistical approach to assessing neural network robustness. In *International Conference on Learning Representations (ICLR)*, 6–9 May 2019. [url](#). [pdf](#).
80. H. Kim, A. Mnih, J. Schwarz, M. Garnelo, A. Eslami, D. Rosenbaum, O. Vinyals, and Y. W. Teh. Attentive neural processes. In *International Conference on Learning Representations (ICLR)*, 6–9 May 2019. [url](#). [pdf](#).
81. X. Miscouridou, F. Caron, and Y. W. Teh. Modelling sparsity, heterogeneity, reciprocity and community structure in temporal interaction data. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 2343–2352. December 2018. [url](#).
82. S. Webb, A. Golinski, R. Zinkov, S. N, T. Rainforth, Y. W. Teh, and F. Wood. Faithful inversion of generative models for effective amortized inference. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 3070–3080. December 2018. [url](#).
83. J. Mitrovic, D. Sejdinovic, and Y. W. Teh. Causal inference via kernel deviance measures. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 6986–6994. December 2018. [url](#).
84. A. Kosiorek, H. Kim, Y. W. Teh, and I. Posner. Sequential attend, infer, repeat: Generative modelling of moving objects. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 8606–8616. December 2018. [url](#).
85. W. M. Czarnecki, S. M. Jayakumar, M. Jaderberg, L. Hasenclever, Y. W. Teh, S. Osindero, N. Heess, and R. Pascanu. Mix & match – agent curricula for reinforcement learning. In *International Conference on Machine Learning (ICML)*, 2018.

86. J. Schwarz, J. Luketina, W. M. Czarnecki, A. Grabska-Barwinska, Y. W. Teh, R. Pascanu, and R. Hadsell. Progress & compress: A scalable framework for continual learning. In *International Conference on Machine Learning (ICML)*, 2018.
87. T. Rainforth, A. R. Kosiorek, T. A. Le, C. J. Maddison, M. Igl, F. Wood, and Y. W. Teh. Tighter variational bounds are not necessarily better. In *International Conference on Machine Learning (ICML)*, 2018.
88. M. Garnelo, D. Rosenbaum, C. Maddison, T. Ramalho, D. Saxton, M. Shanahan, Y. W. Teh, D. J. Rezende, and S. M. A. Eslami. Conditional neural processes. In *International Conference on Machine Learning (ICML)*, 2018.
89. B. Bloem-Reddy, A. Foster, E. Mathieu, and Y. W. Teh. Sampling and inference for beta neutral-to-the-left models of sparse networks. In *International Conference on Uncertainty in Artificial Intelligence (UAI)*, 2018.
90. M. Rowland, M. Bellemare, W. Dabney, R. Munos, and Y. W. Teh. An analysis of categorical distributional reinforcement learning. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2018. [url](#). [pdf](#).
91. H. Kim and Y. W. Teh. Scaling up the automatic statistician: Scalable structure discovery using Gaussian processes. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2018. [url](#). [pdf](#).
92. Y. W. Teh, V. Bapst, W. M. Czarnecki, J. Quan, J. Kirkpatrick, R. Hadsell, N. Heess, and R. Pascanu. Distral: Robust multitask reinforcement learning. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2017. [url](#). [pdf](#).
93. C. J. Maddison, J. Lawson, G. Tucker, N. Heess, M. Norouzi, A. Mnih, A. Doucet, and Y. W. Teh. Filtering variational objectives. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2017a.
94. S. Flaxman, Y. W. Teh, and D. Sejdinovic. Poisson intensity estimation with reproducing kernels. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2017b. [url](#). [pdf](#).
95. X. Lu, V. Perrone, L. Hasenclever, Y. W. Teh, and S. Vollmer. Relativistic Monte Carlo . In A. Singh and J. Zhu, editors, *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2017. [url](#). [pdf](#).
96. C. J. Maddison, A. Mnih, and Y. W. Teh. The concrete distribution: A continuous relaxation of discrete random variables. In *International Conference on Learning Representations (ICLR)*, 2017b. [url](#). [pdf](#).
97. J. Mitrovic, D. Sejdinovic, and Y. W. Teh. DR-ABC: Approximate Bayesian computation with kernel-based distribution regression. In *International Conference on Machine Learning (ICML)*, 2016. [url](#). [pdf](#).
98. M. Balog, B. Lakshminarayanan, Z. Ghahramani, D. M. Roy, and Y. W. Teh. The Mondrian kernel. In *International Conference on Uncertainty in Artificial Intelligence (UAI)*, 2016. [pdf](#).
99. B. Lakshminarayanan, D. M. Roy, and Y. W. Teh. Mondrian forests for large-scale regression when uncertainty matters. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2016. [url](#). [pdf](#).
100. T. Fernandez, N. Rivera, and Y. W. Teh. Gaussian processes for survival analysis. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2016. [url](#). [pdf](#).

101. M. Lomeli, S. Favaro, and Y. W. Teh. A hybrid sampler for Poisson-Kingman mixture models. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2015. [url](#). [pdf](#).
102. T. Lienart, Y. W. Teh, and A. Doucet. Expectation particle belief propagation. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2015. [url](#). [pdf](#).
103. B. Lakshminarayanan, D. M. Roy, and Y. W. Teh. Particle Gibbs for Bayesian additive regression trees. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2015. [url](#). [pdf](#).
104. B. Paige, F. Wood, A. Doucet, and Y. W. Teh. Asynchronous anytime sequential Monte Carlo. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2014. [url](#). [pdf](#).
105. B. Lakshminarayanan, D. Roy, and Y. W. Teh. Mondrian forests: Efficient online random forests. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2014. [url](#). [pdf](#).
106. M. Xu, B. Lakshminarayanan, Y. W. Teh, J. Zhu, and B. Zhang. Distributed Bayesian posterior sampling via moment sharing. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2014. [url](#). [pdf](#).
107. S. Patterson and Y. W. Teh. Stochastic gradient Riemannian Langevin dynamics on the probability simplex. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2013. [url](#). [pdf](#).
108. C. Blundell and Y. W. Teh. Bayesian hierarchical community discovery. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2013. [url](#). [pdf](#).
109. X. Zhang, W. S. Lee, and Y. W. Teh. Learning with invariances via linear functionals on reproducing kernel Hilbert space. In *Advances in Neural Information Processing Systems*, 2013. [url](#). [pdf](#).
110. B. Lakshminarayanan, D. Roy, and Y. W. Teh. Top-down particle filtering for Bayesian decision trees. In *International Conference on Machine Learning (ICML)*, 2013. [pdf](#).
111. C. Chen, V. A. Rao, W. Buntine, and Y. W. Teh. Dependent normalized random measures. In *International Conference on Machine Learning (ICML)*, 2013. [url](#). [pdf](#).
112. A. Mnih and Y. W. Teh. A fast and simple algorithm for training neural probabilistic language models. In *International Conference on Machine Learning (ICML)*, 2012a. [pdf](#).
113. B. Alexe, N. Heess, Y. W. Teh, and V. Ferrari. Searching for objects driven by context. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2012. [url](#). [pdf](#).
114. L. Elliott and Y. W. Teh. Scalable imputation of genetic data with a discrete fragmentation-coagulation process. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2012. [url](#). [pdf](#).
115. F. Caron and Y. W. Teh. Bayesian nonparametric models for ranked data. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2012. [url](#). [pdf](#).
116. V. Rao and Y. W. Teh. MCMC for continuous-time discrete-state systems. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2012. [url](#). [pdf](#).
117. A. Mnih and Y. W. Teh. Learning label trees for probabilistic modelling of implicit feedback. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2012b. [url](#). [pdf](#).
118. N. Heess, D. Silver, and Y. W. Teh. Actor-critic reinforcement learning with energy-based policies. In *JMLR Workshop and Conference Proceedings: EWRL 2012*, 2012. [url](#). [pdf](#).
119. Y. W. Teh, C. Blundell, and L. T. Elliott. Modelling genetic variations with fragmentation-coagulation processes. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2011. [url](#). [pdf](#).

120. V. Rao and Y. W. Teh. Gaussian process modulated renewal processes. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2011b. [url](#). [pdf](#).
121. V. Rao and Y. W. Teh. Fast MCMC sampling for Markov jump processes and continuous time Bayesian networks. In *International Conference on Uncertainty in Artificial Intelligence (UAI)*, 2011a. [pdf](#).
122. M. Welling and Y. W. Teh. Bayesian learning via stochastic gradient Langevin dynamics. In *International Conference on Machine Learning (ICML)*, 2011. [pdf](#).
123. R. Silva, C. Blundell, and Y. W. Teh. Mixed cumulative distribution networks. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2011. [pdf](#).
124. J. Gasthaus and Y. W. Teh. Improvements to the sequence memoizer. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2010. [url](#). [pdf](#).
125. C. Blundell, Y. W. Teh, and K. A. Heller. Bayesian rose trees. In *International Conference on Uncertainty in Artificial Intelligence (UAI)*, 2010. [pdf](#).
126. J. Gasthaus, F. Wood, and Y. W. Teh. Lossless compression based on the sequence memoizer. In *Data Compression Conference*, 2010. [pdf](#).
127. Y. W. Teh and D. Görür. Indian buffet processes with power-law behavior. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 22, 2009. [url](#). [pdf](#).
128. V. Rao and Y. W. Teh. Spatial normalized gamma processes. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 22, 2009. [url](#). [pdf](#).
129. A. Asuncion, M. Welling, P. Smyth, and Y. W. Teh. On smoothing and inference for topic models. In *International Conference on Uncertainty in Artificial Intelligence (UAI)*, 2009. [pdf](#).
130. F. Wood, C. Archambeau, J. Gasthaus, L. F. James, and Y. W. Teh. A stochastic memoizer for sequence data. In *International Conference on Machine Learning (ICML)*, volume 26, pages 1129–1136, 2009. [pdf](#).
131. F. Doshi, K. T. Miller, J. Van Gael, and Y. W. Teh. Variational inference for the Indian buffet process. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 5, 2009. [url](#). [pdf](#).
132. K. A. Heller, Y. W. Teh, and D. Görür. Infinite hierarchical hidden Markov models. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 5, 2009. [url](#). [pdf](#).
133. F. Wood and Y. W. Teh. A hierarchical nonparametric Bayesian approach to statistical language model domain adaptation. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2009. [url](#). [pdf](#).
134. G. R. Haffari and Y. W. Teh. Hierarchical Dirichlet trees for information retrieval. In *Proceedings of the Annual Meeting of the North American Association for Computational Linguistics and the Human Language Technology Conference*, 2009. [pdf](#).
135. D. M. Roy and Y. W. Teh. The Mondrian process. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 21, 2009. [url](#). [pdf](#).
136. D. Görür and Y. W. Teh. An efficient sequential Monte-Carlo algorithm for coalescent clustering. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 21, 2009. [url](#). [pdf](#).

137. J. Van Gael, Y. W. Teh, and Z. Ghahramani. The infinite factorial hidden Markov model. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 21, 2009. [url](#). [pdf](#).
138. J. Gasthaus, F. Wood, D. Görür, and Y. W. Teh. Dependent Dirichlet process spike sorting. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 21, pages 497–504, 2009. [url](#). [pdf](#).
139. G. Quon, Y. W. Teh, E. Chan, T. Hughes, M. Brudno, and Q. Morris. A mixture model for the evolution of gene expression in non-homogeneous datasets. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 21, 2009. [url](#). [pdf](#).
140. M. Welling, Y. W. Teh, and H. J. Kappen. Hybrid Variational/Gibbs collapsed inference in topic models. In *International Conference on Uncertainty in Artificial Intelligence (UAI)*, volume 24, 2008. [pdf](#).
141. J. Van Gael, Y. Saatchi, Y. W. Teh, and Z. Ghahramani. Beam sampling for the infinite hidden Markov model. In *International Conference on Machine Learning (ICML)*, volume 25, 2008. [pdf](#).
142. Y. W. Teh, H. Daume III, and D. M. Roy. Bayesian agglomerative clustering with coalescents. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 20, 2008a. [url](#). [pdf](#).
143. Y. W. Teh, K. Kurihara, and M. Welling. Collapsed variational inference for HDP. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 20, 2008b. [url](#). [pdf](#).
144. H. L. Chieu, W. S. Lee, and Y. W. Teh. Cooled and relaxed survey propagation for MRFs. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 20, 2008. [url](#). [pdf](#).
145. Y. J. Lim and Y. W. Teh. Variational Bayesian approach to movie rating prediction. In *Proceedings of KDD Cup and Workshop*, 2007. [pdf](#).
146. J. F. Cai, W. S. Lee, and Y. W. Teh. NUS-ML: Improving word sense disambiguation using topic features. In *Proceedings of the International Workshop on Semantic Evaluations*, volume 4, 2007a. [pdf](#).
147. J. F. Cai, W. S. Lee, and Y. W. Teh. Improving word sense disambiguation using topic features. In *Proceedings of the Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-coNLL)*, 2007b. [url](#). [pdf](#).
148. Y. W. Teh, D. Görür, and Z. Ghahramani. Stick-breaking construction for the Indian buffet process. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 11, 2007a. [pdf](#).
149. K. Kurihara, M. Welling, and Y. W. Teh. Collapsed variational Dirichlet process mixture models. In *Proceedings of the International Joint Conference on Artificial Intelligence*, volume 20, 2007. [pdf](#).
150. Y. W. Teh, D. Newman, and M. Welling. A collapsed variational Bayesian inference algorithm for latent Dirichlet allocation. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 19, pages 1353–1360, 2007b. [url](#). [pdf](#).
151. Y. W. Teh. A hierarchical Bayesian language model based on Pitman-Yor processes. In *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics*, pages 985–992, 2006. [url](#). [pdf](#).

152. E. P. Xing, K. Sohn, M. I. Jordan, and Y. W. Teh. Bayesian multi-population haplotype inference via a hierarchical Dirichlet process mixture. In *International Conference on Machine Learning (ICML)*, volume 23, 2006. [pdf](#).
153. M. Welling, T. Minka, and Y. W. Teh. Structured region graphs: Morphing EP into GBP. In *International Conference on Uncertainty in Artificial Intelligence (UAI)*, volume 21, 2005. [pdf](#).
154. Y. W. Teh, M. Seeger, and M. I. Jordan. Semiparametric latent factor models. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 10, 2005b. [pdf](#).
155. Y. W. Teh, M. I. Jordan, M. J. Beal, and D. M. Blei. Sharing clusters among related groups: Hierarchical Dirichlet processes. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 17, 2005a. [url](#). [pdf](#).
156. J. Edwards, Y. W. Teh, D. A. Forsyth, R. Bock, M. Maire, and G. Vesom. Making latin manuscripts searchable using gHMM's. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 17, 2005. [url](#). [pdf](#).
157. T. Miller, A. C. Berg, J. Edwards, M. Maire, R. White, Y. W. Teh, E. Learned-Miller, and D. A. Forsyth. Faces and names in the news. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, 2004. [pdf](#).
158. M. Welling, M. Rosen-Zvi, and Y. W. Teh. Approximate inference by markov chains on union spaces. In *International Conference on Machine Learning (ICML)*, volume 21, 2004. [pdf](#).
159. M. Welling and Y. W. Teh. Linear response algorithms for approximate inference in graphical models. *Neural Computation*, 16:197–221, 2004. [url](#). [pdf](#).
160. Y. W. Teh and M. Welling. On improving the efficiency of the iterative proportional fitting procedure. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 9, 2003. [pdf](#).
161. Y. W. Teh and S. Roweis. Automatic alignment of local representations. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 15, 2003. [url](#). [pdf](#).
162. S. Kakade, Y. W. Teh, and S. Roweis. An alternate objective function for Markovian fields. In *International Conference on Machine Learning (ICML)*, volume 19, 2002. [pdf](#).
163. Y. W. Teh and M. Welling. The unified propagation and scaling algorithm. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 14, 2002. [url](#). [pdf](#).
164. M. Welling and Y. W. Teh. Belief optimization for binary networks : A stable alternative to loopy belief propagation. In *International Conference on Uncertainty in Artificial Intelligence (UAI)*, volume 17, 2001. [pdf](#).
165. Y. W. Teh and G. E. Hinton. Rate-coded restricted Boltzmann machines for face recognition. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 13, 2001. [url](#). [pdf](#).
166. G. E. Hinton, Z. Ghahramani, and Y. W. Teh. Learning to parse images. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 12, 2000. [url](#). [pdf](#).
167. F. Bacchus and Y. W. Teh. Making forward chaining relevant. In *Proceedings of the International Conference on Artificial Intelligence Planning Systems*, 1998. [pdf](#).