AI & Statistics 2019

April 16-18, 2019
Naha, Okinawa

Contents

About AISTATS2019

Welcome to AISTATS2019!

The 22nd International Conference on Artificial Intelligence and Statistics (AISTATS 2019) is held in Naha, Okinawa, Japan from Tuesday, 16 April 2019 to Thursday, 18 April 2019 at the LOISIR Hotel Naha (3-2-1 Nishi, Naha-shi, Okinawa, 900-0036, Japan).

Since its inception in 1985, AISTATS has been an interdisciplinary gathering of researchers at the intersection of artificial intelligence, machine learning, statistics, and related areas.

Organizing committee

Program Chairs
Kamalika Chaudhuri (University of California, San Diego)
Masashi Sugiyama (RIKEN and University of Tokyo)

Local Chair
Ichiro Takeuchi (Nagoya Institute of Technology and RIKEN)

Workflow Chairs
James Foulds (University of Maryland, Baltimore County)
Makoto Yamada (Kyoto University and RIKEN)

Publication Chairs
Purna Sarkar (University of Texas at Austin)
Akiko Takeda (University of Tokyo and RIKEN)

Publicity Chair
Daniel Hsu (Columbia University)

Area Chairs

Aaditya Ramdas (Carnegie Mellon University)
Adrian Weller (University of Cambridge)
Alex Dimakis (University of Texas at Austin)
Alp Kucukelbir (Columbia University)
Amin Karbasi (Yale University)
Amir Globerson (Tel Aviv University, Google)
Andrew Dai (Google Brain)
Andrew Gordon Wilson (Cornell University)
Andriy Mnih (DeepMind)
Antti Honkela (University of Helsinki)
Arindam Banerjee (University of Minnesota)
Armand Joulin (Facebook AI Research)
Aryeh Kontorovich (Ben-Gurion University of the Negev)
Balaji Lakshminarayanan (Google DeepMind)
Barbara Engelhardt (Princeton University)
Barnabas Poczos (Carnegie Mellon University)
Been Kim (Google)
Bharath Sriperumbudur (Pennsylvania State University)
Boqing Gong (Tencent AI Lab)
Borja Balle (Amazon)
Brian Kulis (Boston University)
Byron Boots (Georgia Institute of Technology)
Chicheng Zhang (Microsoft Research)
Claire Monteleoni (University of Colorado Boulder)
Daniel Hernandez-Lobato (Universidad Autonoma de Madrid)
Daniel Hsu (Columbia University)
Daniel Sheldon (University of Massachusetts, Amherst)
Danilo Jimenez Rezende (Google DeepMind)
Dino Sejdinovic (University of Oxford)
Dustin Tran (Columbia University)
Edwin Bonilla (Data61)
Ery Arias-Castro (University of California, San Diego)
Finale Doshi-Velez (Harvard)
Francesco Orabona (Boston University)
Gang Niu (RIKEN)
Han Liu (Northwestern)
Hidetoshi Shimodaira (Kyoto University)
Hossein Mobahi (Google Research)
Ilya Tolstikhin (Google)
Isabel Valera (MPI-IS)
Ivor Tsang (University of Technology Sydney)

James Tin-Yau Kwok (The Hong Kong University of Science and Technology)
James Zou (Microsoft)
Jimmy Ba (University of Toronto)
Joan Bruna (Courant Institute of Mathematical Sciences)
John Duchi (Stanford University)
Jose Miguel Hernandez-Lobato (University of Cambridge)
Justin Solomon (MIT)
Katherine Heller (Duke University)
Kenji Fukumizu (The Institute of Statistical Mathematics)
Koby Crammer (Technion)
Laura Balzano (University of Michigan)
Le Song (Ant Financial and Georgia Institute of Technology)
Lihong Li (Google Inc.)
Liwei Wang (Peking University)
Marco Cuturi (ENSAE/CREST)
Martin Takac (Lehigh University)
Matthias Hein (University of Tuebingen)
Maxim Raginsky (University of Illinois)
Michael Mahoney (University of California, Berkeley)
Miguel Carreira-Perpinan (University of California, Merced)
Mijung Park (MPI Tuebingen)
Minmin Chen (Google)
Mladen Kolar (U Chicago)
Mohammad Emtiyaz Khan (RIKEN)
Nan Jiang (University of Illinois at Urbana-Champaign)
Nathan Kallus (Cornell Tech)
Nathan Srebro (Toyota Technical Institute of Chicago)
Negar Kiyavash (Georgia Institute of Technology)
Nihar Shah (Carnegie Mellon University)
Nika Haghtalab (Cornell University)
Peter Bartlett (University of California, Berkeley)
Ping Li (Baidu Research)
Praneeth Netrapalli (Microsoft Research)
Prateek Jain (Microsoft Research)
Purnamrita Sarkar (University of Texas at Austin)
Rafael Frongillo (CU Boulder)
Raman Arora (Johns Hopkins University)
Rebecca Willett (U Chicago)
Remi Munos (DeepMind)
Ricardo Silva (UCL)
Richard Nock (Data61, CSIRO)
Ruth Urner (York University)

Samory Kpotufe (Columbia University)
Samuel Kaski (Aalto University)
Sanjoy Dasgupta (University of California, San Diego)
Sanmi Koyejo (University of Illinois, Urbana-Champaign)
Sayan Mukherjee (Duke)
Seth Flaxman (Imperial College London)
Seungjin Choi (POSTECH)
Sewoong Oh (University of Illinois at Urbana-Champaign)
Shai Ben-David (University of Waterloo)
Silvia Chiappa (DeepMind)
Simon Lacoste-Julien (University of Montreal)
Sinead Williamson (University of Texas at Austin)
Stephen Bach (Brown University)
Sujay Sanghavi (University of Texas at Austin)
Taiji Suzuki (University of Tokyo and RIKEN)
Trevor Campbell (UBC)
Varun Kanade (Oxford)
Vikas Sindhwani (Google)
Vinayak Rao (Purdue University)
Volkan Cevher (EPFL)
Wee Sun Lee (NUS)
Xiaojin Zhu (University of Wisconsin-Madison)
Yarin Gal (Oxford)
Yingyu Liang (University of Wisconsin Madison)
Yingzhen Li (Microsoft Research Cambridge)
Yisong Yue (Caltech)
Yutian Chen (Deepmind)
Yu-Xiang Wang (University of California, Santa Barbara)
Zico Kolter (Carnegie Mellon University)
Zoltan Szabo (Ecole Polytechnique)

Sponsors

OMRON TOSHIBA

SoftBank TOYOTA

Conference Information

Basic Information

- Date: April 16 (Tue) - 18 (Thu), 2019
- Venue: Loisir HOTEL Naha (3rd Floor)
  - Hotel web-site:

    `https://www.solarehotels.com/en/hotel/okinawa/loisir-naha/`
  - Address: 3-2-1 Nishi, Naha-shi, Okinawa, Japan
  - Tel: +81-98-868-2222
- WiFi (the capacity is limited; use your own one if possible):
  - SSID: LOISIR
  - Password: 19930202
- Registration desk will open at Foyer (3rd floor) as follows:
  - Day0 (April 15, Mon): 17:00 - 20:00
  - Day1 (April 16, Tue): 07:00 - 19:00
  - Day2 (April 17, Wed): 08:30 - 19:00
  - Day3 (April 18, Thu): 08:30 - 18:00
- Banquet is held from 19:00 to 21:00 at Day2 (Apr. 17, Wed)
- Breakfast and lunch tickets from Apr.16 to 18 are provided

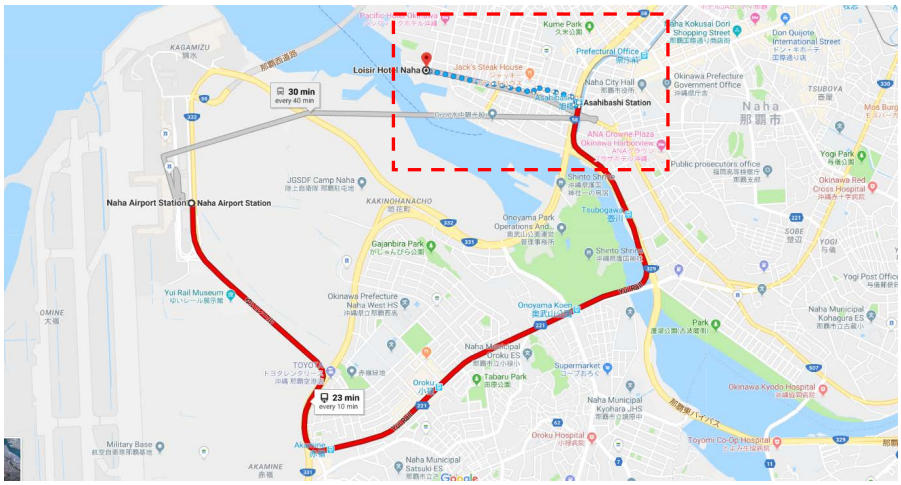
How to access to the venue

- Hotel web-site for access:

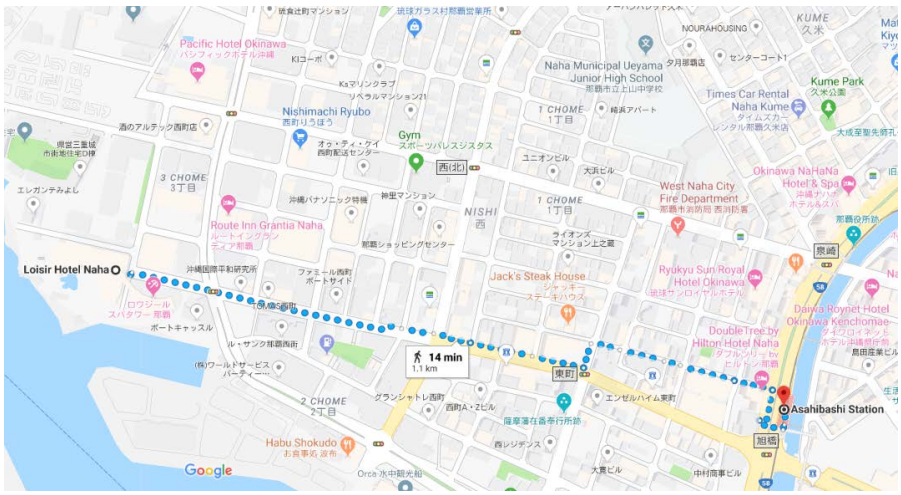
  `https://www.solarehotels.com/en/hotel/okinawa/loisir-naha/access.html`

- Fly to Naha airport (no other choices)

- From the airport to the venue, use a taxi or a monorail. If you use taxi, it takes only 10-15 mins. If you use the monorail, get off at *Asahibashi* station (5-th station from the airport), and walk 10 mins.



From the airport to the venue



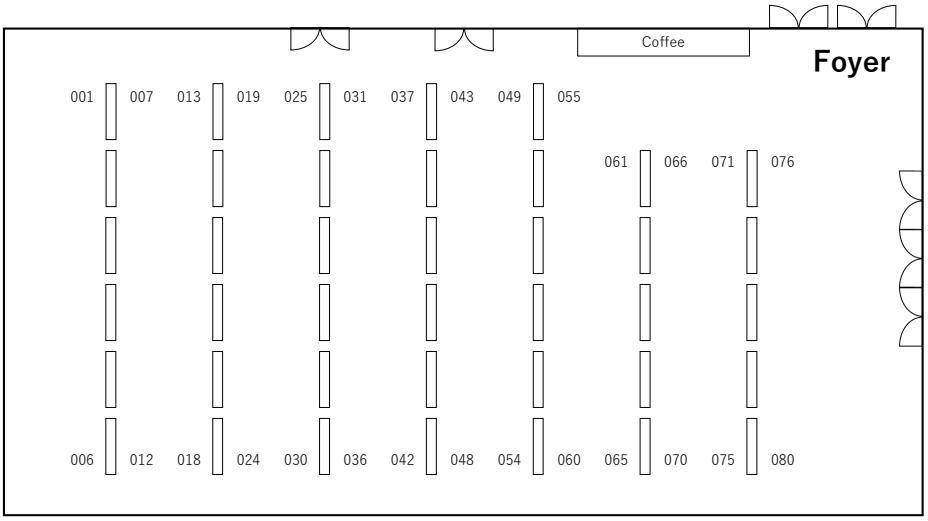
From *Asahibashi* station to the venue

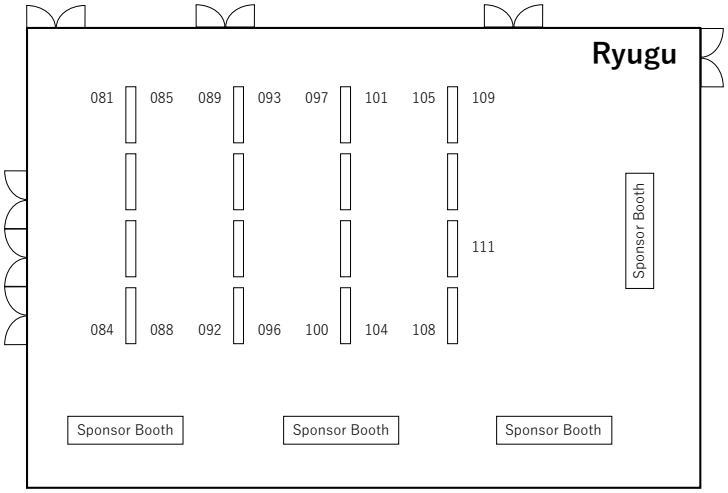
Floor map

- Registration desk is located at *Foyer* (3rd floor)
- Main session and banquet is held at *Tenpi* (3rd floor)
- Poster session, Sponsor booth, and coffee break are held at *Foyer* and *Ryugu* (3rd floor)
- Breakfast and lunch are available at *Amairo* (1st floor)

Poster session information

- Poster board size is 2100mm Height and 900mm Width
- Poster layout





- Please put the poster in place at lunch time of the presentation day, and put it away before the morning session in the next day.

Timetable

Tuesday, 16 April	
7:30-9:00	Breakfast (90 min)
9:00-9:10	opening (10 min)
9:10 - 10:10	Invited talk 1 (1h) ***Speaker: Robert Tibshirani***
10:10-10:30	coffee break (20 min)
10:30 - 11:50	session 1a (20 min x 4) ***Learning Theory***
11:50-13:20	lunch (90 min)
13:20 - 14:20	session 1b (20 min x 3) ***Statistics and Machine Learning 1***
14:20-14:40	coffee break (20 min)
14:40 - 15:40	session 1c (20 min x 3) ***Statistics and Machine Learning 2***
15:40 - 18:40	poster session 1 (3 hours)

Wednesday, 17 April	
7:30-9:00	Breakfast (90 min)
9:00-10:00	Invited talk 2 (1h) ***Speaker: Poling Loh***
10:00-10:30	coffee break (30 min)
10:30 - 11:50	session 2a (20 min x 4) ***Bayesian Methods***
11:50-13:30	lunch (100 min)
13:30 - 14:30	session 2b (20 min x 3) ***Inference***
14:30-14:50	coffee break (20 min)
14:50 - 15:50	session 2c (20 min x 3) ***Privacy and Fairness***
15:50 - 18:50	poster session 2 (3 hours)
19:00 - 21:00	Banquet

Thursday, 18 April	
7:30-9:00	Breakfast (90 min)
9:00-10:20	session 3a (20 min x 4) ***Optimization***
10:20-10:40	coffee break (20 min)
10:40 - 12:00	session 3b (20 min x 4) ***Reinforcement Learning and Bandits***
12:00-13:30	lunch (90 min)
13:30 - 16:30	poster session 3 (3 hours)
16:30 - 17:30	Invited talk 3 (1h) ***Speaker: Zhi-Hua Zhou***
17:30 - 17:40	closing (10 min)

Day 1: April 16 (Tuesday), 2019

7:30–9:00 **Breakfast**
9:00–9:10 **Opening Remarks**
9:10–10:10 **Invited Talk 1**

- *Statistical Learning and Sparsity*
  Robert Tibshirani (Stanford University)

10:10–10:30 **Coffee Break**
10:30–11:50 **Oral Session 1a: Learning Theory**

- *Learning Rules-First Classifiers*
  Deborah Cohen (Google)*; Amit Daniely (Google); Amir Globerson (Google); Gal Elidan (Google)
- *Multitask Metric Learning: Theory and Algorithm*
  Boyu Wang (Princeton University)*; Hejia Zhang (Princeton); Peng Liu (University of Toronto); Zebang Shen (Zhejiang University; Tencent AI Lab); Joelle Pineau (McGill / Facebook)
- *Active Ranking with Subset-wise Preferences in the Plackett-Luce model*
  Aadirupa Saha (Indian Institute of Science)*; Aditya Gopalan (Indian Institute of Science (IISc), Bangalore)
- *Convergence of Gradient Descent on Separable Data*
  Mor Shpigel Nacson (Technion)*; Jason Lee (USC); Suriya Gunasekar (TTI Chicago); Pedro Henrique Pamplona Savarese (Toyota Technical Institute of Chicago); Nathan Srebro (Toyota Technical Institute of Chicago); Daniel Soudry (Technion)

11:50–13:20 **Lunch Break**
13:20–14:20 **Oral Session 1b: Statistics and Machine Learning 1**

- *Nonlinear ICA Using Auxiliary Variables and Generalized Contrastive Learning*
  Aapo Hyvarinen (UCL and U Helsinki)*; Hiroaki Sasaki (Nara Institute of Science and Technology); Richard Turner ()
- *Estimating Network Structure from Incomplete Event Data*
  Ben Mark (University of Wisconsin-Madison)*; Garvesh Raskutti (UW-Madison); Rebecca Willett (U Chicago)
- *A Higher-Order Kolmogorov-Smirnov Test*
  Veeranjaneyulu Sadhanala (Carnegie Mellon University)*; Aaditya Ramdas (Carnegie Mellon University); Yu-Xiang Wang (UC Santa Barbara); Ryan Tibshirani (CMU)

14:20–14:40 **Coffee Break**
14:40–15:40 **Oral Session 1c: Statistics and Machine Learning 2**

- *A Stein–Papangelou Goodness-of-Fit Test for Point Processes*
  Jiasen Yang (Purdue University)*; Vinayak Rao (Purdue University); Jennifer Neville (Purdue University)

- *Defending against Whitebox Adversarial Attacks via Randomized Discretization*
  Yuchen Zhang (Microsoft)*; Percy Liang ()
- *A Swiss Army Infinitesimal Jackknife*
  Ryan Giordano (UC Berkeley)*; William T Stephenson (MIT); Runjing Liu (UC Berkeley); Michael Jordan (UC Berkeley); Tamara Broderick (MIT)

15:40–18:40 **Poster Session 1**

Day 2: April 17 (Wednesday), 2019

7:30–9:00 **Breakfast**
9:00–10:00 **Invited Talk 2**

- *Data Science for Networked Data*
  Poling Loh (University of Wisconsin-Madison)

10:00–10:30 **Coffee Break**
10:30–11:50 **Oral Session 2a: Bayesian Methods**

- *A Bayesian model for sparse graphs with flexible degree distribution and overlapping community structure*
  Juho Lee (University of Oxford)*; Lancelot James (Hong Kong University of Science and Technology); Seungjin Choi (POSTECH); Francois Caron (Oxford)
- *On Multi-Cause Approaches to Causal Inference with Unobserved Counfounding: Two Cautionary Failure Cases and A Promising Alternative*
  Alexander D'Amour (Google Brain)*
- *Conditionally Independent Multiresolution Gaussian Processes*
  Jalil Taghia (Uppsala University)*; Thomas Schön (Uppsala University)
- *Deep learning with differential Gaussian process flows*
  Pashupati R Hegde (Aalto University)*; Markus Heinonen (Aalto University); Harri Lähdesmäki (Aalto University); Samuel Kaski (Aalto University)

11:50–13:30 **Lunch Break**
13:30–14:30 **Oral Session 2b: Inference**

- *Block Stability for MAP Inference*
  unter Lang (MIT)*; David Sontag (MIT); Aravindan Vijayaraghavan ()
- *Reparameterizing Distributions on Lie Groups*
  Tim R Davidson (University of Amsterdam)*; Pim de Haan (University of Amsterdam); Luca Falorsi (University of Amsterdam); Patrick Forré (University of Amsterdam)
- *Unbiased Smoothing using Particle Independent Metropolis-Hastings* Lawrece T Middleton (University of Oxford)*; Arnaud Doucet (Oxford University); Pierre Jacob (Harvard University); George Deligiannidis (Oxford)

14:30–14:50 **Coffee Break**
14:50–15:50 **Oral Session 2c: Privacy and Fairness**

- *Subsampled Renyi Differential Privacy and Analytical Moments Accountant*
  Yu-Xiang Wang (UC Santa Barbara)*; Borja Balle (Amazon); Shiva Kasiviswanathan (Amazon AWS AI)
- *Renyi Differentially Private ERM for Smooth Objectives*
  Chen Chen (University of Georgia); Jaewoo Lee (University of Georgia)*; Dan Kifer (Pennsylva State Univ., USA)
- *Attenuating Bias in Word Vectors*
  Sunipa Dev (University of Utah)*; Jeff Phillips (University of Utah)

15:50–18:50 **Poster Session 2**
19:00–21:00 **Banquet**

Day 3: April 18 (Thursday), 2019

7:30–9:00 **Breakfast**
9:00–10:20 **Oral Session 3a: Optimization**

- *Stochastic Gradient Descent with Exponential Convergence Rates of Expected Classification Errors*
  Atsushi Nitanda (The University of Tokyo / RIKEN)*; Taiji Suzuki (University of Tokyo / RIKEN)
- *Sobolev Descent*
  Youssef Mroueh (IBM Research)*; Tom Sercu (IBM); Anant Raj (Max-Planck Institute for Intelligent Systems)
- *Empirical Risk Minimization and Stochastic Gradient Descent for Relational Data*
  Victor Veitch (Columbia University)*; Morgane Austern (Columbia University); Wenda Zhou (Columbia University); Peter Orbanz (); David Blei (Columbia University)
- *Stochastic Algorithms with Descent Guarantees for ICA*
  Pierre Ablin (Inria)*; Alexandre Gramfort (Inria); Jean-François Cardoso (CNRS - Institut d'astrophysique de Paris); Francis Bach (INRIA - Ecole Normale Supérieure)

10:20–10:40 **Coffee Break**
10:40–12:00 **Oral Session 3b: Reinforcement Learning and Bandits**

- *Rotting Bandits Are No Harder Than Stochastic Bandits*
  Julien Seznec (lelivrescolaire.fr)*; Andrea Locatelli (Uni Magdeburg); Alexandra Carpentier (Otto-von-Guericke-Universität Magdeburg); Alessandro Lazaric (FAIR); Michal Valko (Inria)
- *Thompson Sampling for Cascading Bandits*
  Wang Chi Cheung (IHPC, ASTAR); Vincent Tan (NUS); Zixin Zhong (NUS)*

- *Distilling Policy Distillation*
  Wojciech M Czarnecki (DeepMind)*; Razvan Pascanu (Google Deepmind); Simon Osindero (DeepMind); Siddhant Jayakumar (DeepMind); Grzegorz M Swirszcz (DeepMind); Max Jaderberg (Google)
- *Theoretical Analysis of Efficiency and Robustness of Softmax and Gap-Increasing Operators in Reinforcement Learning*
  Tadashi Kozuno (Okinawa Institute of Science and Technology)*; Eiji Uchibe (ATR Computational Neuroscience Labs.); Kenji Doya (Okinawa Institute of Science and Technology)

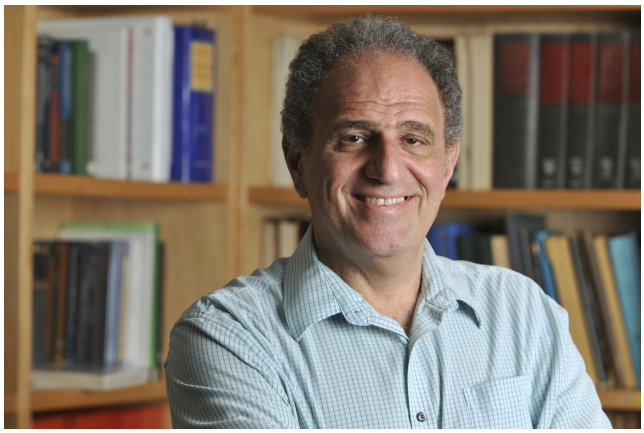
12:00–13:30 **Lunch Break**
13:30–16:30 **Poster Session 3**
16:30–17:30 **Invited Talk 3**

- *An Exploration to Non-NN Deep Models based on Non-Differentiable Modules*
  Zhi-Hua Zhou (Nanjing University)

17:30-17:40 **Closing Remarks**

Invited Talks

Profressor Robert Tibshirani, Stanford University



Statistical Learning and Sparsity

I will review the lasso method for high dimensional supervised learning and discuss some new developments in the area, including the Pliable Lasso, and post-selection inference for understanding the important features. I will also describe some applications of these methods to my own collaborative work, including prediction of platelet usage at Stanford Hospital.

Biography: Rob Tibshirani is a Professor of Statistics, and Biomedical Data Science at Stanford University. His main interests are in applied statistics, biostatistics, and data science. He is most well-known for the LASSO, which is a shrinkage and selection method for linear regression. He is the co-author of the books Generalized Additive Models (with T. Hastie), An Introduction to the Bootstrap (with B. Efron), An Introduction to Statistical Learning (with G. James, D. Witten and T. Hastie), Sparsity in Statistics (with T. Hastie and M. Wainwright) and the widely used Elements of Statistical Learning (with T. Hastie and J. Friedman). His current research focuses on problems in biology and genomics, medicine, and industry.

Professor Po-Ling Loh, University of Wisconsin-Madison

Data Science for Networked Data

We will survey a variety of problems involving mathematical analysis of network-structured data. In many scientific problems of contemporary interest, data are acquired in a very heterogeneous and non-i.i.d. fashion: Edges in a network may give rise to important correlations between node-level observations, which must be taken into account when performing data analysis. In large-scale applications, the structure of the graph may also determine the type of algorithms that may be performed. Our talk will cover topics such as influence maximization, source inference, graph hypothesis testing, immunization, and local optimization algorithms on networks.

Biography: Po-Ling Loh is an assistant professor in the ECE department at the UW-Madison, with a secondary appointment in the statistics, computer science, and industrial and systems engineering departments. From 2014-2016, Po-Ling was an assistant professor in the statistics department at the Wharton School at the University of Pennsylvania. Po-Ling received an MS in computer science and a PhD in statistics from Berkeley in 2013 and 2014, and a BS in math with a minor in English from Caltech in 2009. She was the recipient of the 2014 Erich L. Lehmann Citation from the Berkeley statistics department for an outstanding PhD dissertation in theoretical statistics, and a best paper award at the NIPS conference in 2012. Po-Ling is a recipient of an NSF CAREER award in statistics.

Professor Zhi-Hua Zhou, Nanjing University

An Exploration to Non-NN Deep Models Based on Non-differentiable Modules

The word "deep learning" is often regarded as a synonym of "deep neural networks (DNNs)". In this talk, we claim that the essential of deep learning lies in the combination of layer-by-layer processing, in-model feature transformation and sufficient model complexity, and it is not that crucial whether deep models are realized by neural networks or not. To verify the conjecture, we will show that it is possible to construct non-NN style deep models without relying on backpropagation training nor gradient-based adjustment. We advocate the exploration to non-NN deep models, because neural network based deep models have already been studied for many years while it is well-known that none model can always be the best.

Biography: Zhi-Hua Zhou is a Professor of Computer Science and Artificial Intelligence at Nanjing University. He is the founding director of the LAMDA Group and head of the department of computer science. His main research interests are in machine learning and data mining, involving ensemble methods, weakly supervised learning, multi-label learning, etc. He authored the books "Ensemble Methods: Foundations and Algorithms" and "Machine Learning (in Chinese)", and published more than 200 papers in top-tier international journals/conferences. According to Google Scholar, his publications have received more than 35,000 citations, with an H-index of 90. He has received various awards, including the National Natural Science Award of China, PAKDD Distinguished Contribution Award, Microsoft Professorship Award, etc. He served as General chair of IEEE ICDM 2016, Program chair of AAAI 2019, IJCAI 2015 Machine Learning track, SDM 2013, etc. He is a Fellow of the ACM, AAAI, AAAS, IEEE and IAPR.

Poster Sessions

Day 1: April 16 (Tuesday), 2019

Tu1 *Proximal Splitting Meets Variance Reduction*
Fabian Pedregosa (UC Berkeley)*; Kilian EJ Fatras (ENSTA ParisTech); Mattia Casotto (Kamet)

Tu2 *Optimal Noise-Adding Mechanism in Additive Differential Privacy*
Quan Geng (Google)*; Wei Ding (Google); Ruiqi Guo (Google); Sanjiv Kumar (Google Research)

Tu3 *Tossing Coins Under Monotonicity*
Matey Neykov (Carnegie Mellon University)*

Tu4 *Gaussian Regression with Convex Constraints*
Matey Neykov (Carnegie Mellon University)*

Tu5 *Risk-Averse Stochastic Convex Bandit*
Adrian Rivera Cardoso (Georgia Tech)*; Huan Xu (Georgia Inst. of Technology)

Tu6 *Error bounds for sparse classifiers in high-dimensions*
Antoine Dedieu (MIT)*

Tu7 *Boosting Survival Predictions with Auxiliary Data from Heterogeneous Domains*
Alexis Bellot (University of Oxford)*

Tu8 *Resampled Priors for Variational Autoencoders*
Matthias Bauer (MPI Intelligent Systems/University of Cambridge)*; Andriy Mnih (DeepMind)

Tu9 *Scalable Bayesian Learning for State Space Models using Variational Inference with SMC Samplers*
Marcel Hirt (University College London)*; Petros Dellaportas (UCL)

Tu10 *Scalable Thompson Sampling via Optimal Transport*
Ruiyi Zhang (Duke University)*; Zheng Wen (Adobe Research); Changyou Chen (University at Buffalo); CHEN FANG (Adobe Research, San Jose, CA)

Tu11 *Inferring Multidimensional Rates of Aging from Cross-Sectional Data*
Emma Pierson (Stanford)*; Pang Wei Koh (Stanford); Tatsunori Hashimoto (Stanford); Daphne Koller (insitro); Jure Leskovec (Stanford); Nick Eriksson (Calico); Percy Liang ()

Tu12 *Interaction Detection with Bayesian Decision Tree Ensembles*
Junliang Du (Florida State University); Antonio R Linero (Florida State University)*

Tu13 *Interaction Effects: The Lurking Problem in Machine Learning Systems*
Matt Barnes (Carnegie Mellon University)*; Artur Dubrawski (CMU)

Tu14 *Towards a Theoretical Understanding of Hashing-Based Neural Nets*
Yibo Lin (UT-Austin)*; Zhao Song (Harvard University); Lin Yang (Princeton University)

Tu15 *Faster First-Order Methods for Stochastic Non-Convex Optimization on Riemannian Manifolds*

Pan Zhou (NUS)*; Jiashi Feng (NUS); Xiaotong Yuan (Nanjing University of Information Science and Technology)

Tu16 *A Low-Level Probabilistic Programming Language for Non-Differentiable Models*
Yuan Zhou (University of Oxford); Bradley J Gram-Hansen (University of Oxford)*; Tobias Kohn (University of Oxford); Tom Rainforth (University of Oxford); Frank Wood (University of British Columbia); Hongseok Yang ()

Tu17 *Identifiability of Generalized Hypergeometric Distribution (GHD) Directed Acyclic Graphical Models*
Gunwoong Park (University of Seoul)*; Hyewon Park (University of Seoul)

Tu18 *Unbiased Implicit Variational Inference*
Michalis Titsias (Athens University)*; Francisco Ruiz ("Columbia University, University of Cambridge")

Tu19 *Efficient Linear Bandits through Matrix Sketching*
Ilja Kuzborskij (University of Milan)*; Leonardo Cella (); Nicolo A Cesa Bianchi (Milan University)

Tu20 *Orthogonal Estimation of Wasserstein Distances*
Mark Rowland (DeepMind)*; Jiri Hron (University of Cambridge); Krzysztof Choromanski (Google Brain Robotics); Tamas Sarlos (Google Research); Yunhao Tang (Columbia University); Adrian Weller (Cambridge University)

Tu21 *Linear Convergence of the Primal-Dual Gradient Method for Convex-Concave Saddle Point Problems without Strong Convexity*
Simon S Du (Carnegie Mellon University); Wei Hu (Princeton University)*

Tu22 *Greedy and IHT Algorithms for Non-convex Optimization with Monotone Costs of Non-zeros*
Shinsaku Sakaue (NTT)*

Tu23 *KAMA-NNs: low-dimensional rotation based neural networks*
Krzysztof Choromanski (Google Brain Robotics)*; Aldo Pacchiano (UC Berkeley); Jeffrey Pennington (Google Brain); Yunhao Tang (Columbia University)

Tu24 *Statistical Windows in Testing for the Initial Distribution of a Reversible Markov Chain*
Quentin Berthet (University of Cambridge); Varun Kanade (Oxford)*

Tu25 *Sketching for Latent Dirichlet-Categorical Models*
Joseph Tassarotti (Carnegie-Mellon University)*; Jean-Baptiste Tristan (Oracle Labs); Michael Wick (Oracle Labs)

Tu26 *Adaptive Activity Monitoring with Uncertainty Quantification in Switching Gaussian Process Models*
Randy Ardywibowo (Texas A&M University)*; Guang Zhao (Texas A&M University); Zhangyang Wang (TAMU); Bobak J Mortazavi (Texas A&M University); Shuai Huang (University of Washington); Xiaoning Qian (Texas A&M University)

Tu27 *Near Optimal Algorithms for Hard Submodular Programs with Discounted Cooperative Costs*
Rishabh Krishnan Iyer (Microsoft Corporation)*; Jeffrey Bilmes (University of Washington)

Tu28 *Fast Stochastic Algorithms for Low-rank and Nonsmooth Matrix Problems*
Dan Garber (Technion)*; Atara Kaplan (Technion)

Tu29 *Logarithmic Regret for Online Gradient Descent Beyond Strong Convexity*
Dan Garber (Technion)*

Tu30 *Accelerated Coordinate Descent with Arbitrary Sampling and Best Rates for Minibatches*

Filip Hanzely (KAUST)*; Peter Richtarik (KAUST)

Tu31 *Globally-convergent Iteratively Reweighted Least Squares for Robust Regression Problems*
Bhaskar Mukhoty (Indian Institute of Technology Kanpur)*; Govind Gopakumar (Goldman Sachs); Prateek Jain (Microsoft Research); Purushottam Kar (Indian Institute of Technology Kanpur)

Tu32 *Modularity-based Sparse Soft Graph Clustering*
Alexandre Hollocou (INRIA, Paris)*; Thomas Bonald (Telecom Paristech); Marc Lelarge (INRIA-ENS)

Tu33 *Pathwise Derivatives for Multivariate Distributions*
Martin Jankowiak (Uber AI Labs)*; Theofanis Karaletsos (Uber AI Labs)

Tu34 *Distributed Inexact Newton-type Pursuit for Non-convex Sparse Learning*
Bo Liu (Rutgers University)*; Xiaotong Yuan (Nanjing University of Information Science and Technology); Lezi Wang (Rutgers Uniersity); Qingshan Liu (Nanjing University of Information Science & Technology); Junzhou Huang (University of Texas at Arlington); Dimitris Metaxas (Rutgers)

Tu35 *Vine copula structure learning via Monte Carlo tree search*
Bo Chang (University of British Columbia)*; Shenyi Pan (University of British Columbia); Harry Joe (University of British Columbia)

Tu36 *Blind Demixing via Wirtinger Flow with Random Initialization*
Jialin Dong (ShanghaiTech University); Yuanming Shi (ShanghaiTech University)*

Tu37 *Performance Metric Elicitation from Pairwise Classifier Comparisons*
Gaurush Hiranandani (UNIVERSITY OF ILLINOIS, URBANA-CH)*; Shant Boodaghians (UIUC); Ruta Mehta (UIUC); Sanmi Koyejo (University of Illinois, Urbana-Champaign)

Tu38 *Analysis of Network Lasso for Semi-Supervised Regression*
Alexander Jung (Aalto University)*

Tu39 *Learning Mixtures of Smooth Product Distributions: Identifiability and Algorithm*
Nikolaos Kargas (UMN)*; Nicholas Sidiropoulos (University of Virginia)

Tu40 *Robust Low-Rank Estimation from Quantized Observations*
Jie Shen (Stevens Institute of Technology)*; Pranjal Awasthi (Rutgers University); Ping Li (Baidu Research)

Tu41 *Foundations of Sequence-to-Sequence Modeling for Time Series*
Zelda Mariet (Massachusetts Institute of Technology)*; Vitaly Kuznetsov (Google)

Tu42 *Nearly Optimal Adaptive Procedure with Change Detection for Piecewise-Stationary Bandit*
Yang Cao (Uber Technologies, Inc)*; Zheng Wen (Adobe Research); Branislav Kveton (Google Research); Yao Xie (Georgia Tech)

Tu43 *An Optimal Algorithm for Stochastic Three-Composite Optimization*
Renbo Zhao (MIT)*; William Haskell (NUS); Vincent Tan (NUS)

Tu44 *Lifelong Optimization with Low Regret*
Yi-Shan Wu (Academia Sinica)*; Po-An Wang (LIONS, EPFL); Chi-Jen Lu (Academia Sinica)

Tu45 *Sparse Multivariate Bernoulli Processes in High Dimensions*
Parthe Pandit (University of California Los Angeles)*; Arash Amini (UCLA); Mojtaba Sahraee-Ardakan (UCLA); Sundeep Rangan (NYU); Alyson K Fletcher (UCLA)

Tu46 *An Optimal Algorithm for Stochastic and Adversarial Bandits*
Julian Zimmert (University of Copenhagen)*; Yevgeny Seldin (University of Copenhagen)

Tu47 *Efficient Bayesian Experimental Design for Implicit Models*
Steven Kleinegesse (University of Edinburgh)*; Michael U. Gutmann (University of Edinburgh)

Tu48 *Local Saddle Point Optimization: A Curvature Exploitation Approach*
Leonard Adolphs (ETHZ)*; Hadi Daneshmand (ETH); Aurelien Lucchi (ETH Zurich); Thomas Hofmann (ETH Zurich)

Tu49 *Testing Conditional Independence on Discrete Data using Stochastic Complexity*
Alexander Marx (Max-Planck-Institut for Informatics)*; Jilles Vreeken (CISPA Helmholtz Center for Information Security)

Tu50 *Distributionally Robust Submodular Maximization*
Matthew Staib (MIT)*; Bryan Wilder (University of Southern California); Stefanie Jegelka (MIT)

Tu51 *A Robust Zero-Sum Game Framework for Pool-based Active Learning*
Dixian Zhu (University of Iowa)*; Zhe Li (The University of Iowa); Xiaoyu Wang (-); Boqing Gong (Tencent AI Lab); Tianbao Yang (University of Iowa)

Tu52 *Support and Invertibility in Domain-Invariant Representations*
Fredrik D Johansson (MIT)*; David Sontag (MIT); Rajesh Ranganath (New York University)

Tu53 *Efficient Inference in Multi-task Cox Process Models*
Virginia Aglietti (University of Warwick)*; Theodoros Damoulas (University of Warwick); Edwin V Bonilla (Data61)

Tu54 *Optimization of Inf-Convolution Regularized Nonconvex Composite Problems*
Emanuel Laude (TU Munich)*; Tao Wu (TU Munich); Daniel Cremers (TUM)

Tu55 *On Connecting Stochastic Gradient MCMC and Differential Privacy*
Bai Li (Duke University)*; Changyou Chen (University at Buffalo); Hao Liu (Nanjing University); Lawrence Carin Duke (CS)

Tu56 *What made you do this? Understanding black-box decisions with sufficient input subsets*
Brandon Carter (MIT CSAIL)*; Jonas Mueller (MIT); Siddhartha Jain (MIT CSAIL); David Gifford (MIT CSAIL)

Tu57 *Computation Efficient Coded Linear Transform*
Sinong Wang (The Ohio State University)*; Jiashang Liu (The Ohio State University); Ness Shroff (The Ohio State University); Pengyu Yang (The Ohio State University)

Tu58 *Mixing of Hamiltonian Monte Carlo on strongly log-concave distributions 2: Numerical integrators*
Oren Mangoubi (EPFL)*; Aaron Smith (University of Ottawa)

Tu59 *Temporal Quilting for Survival Analysis*
Changhee Lee (UCLA)*; William Zame (UCLA); Ahmed M. Alaa (University of California, Los Angeles); Mihaela van der Schaar ()

Tu60 *Learning Classifiers with Fenchel-Young Losses: Generalized Entropies, Margins, and Algorithms*
Mathieu Blondel (NTT)*; Andre Martins (Unbabel); Vlad Niculae (Instituto de Telecomunicações)

Tu61 *On Target Shift in Adversarial Domain Adaptation*
Yitong Li (Duke University)*; David Carlson (Duke)

Tu62 *Optimal testing in the experiment-rich regime*
Sven Schmit (Stitch Fix, Inc)*; Ramesh Johari (Stanford University); Virag Shah (Stanford

University)

Tu63 *Reversible Jump Probabilistic Programming*
David A Roberts (University of Queensland)*; Marcus Gallagher (The University of Queensland); Thomas Taimre (University of Queensland)

Tu64 *Graph Embedding with Shifted Inner Product Similarity and Its Improved Approximation Capability*
Akifumi Okuno (Kyoto University and RIKEN AIP)*; Geewook Kim (Kyoto University / RIKEN AIP); Hidetoshi Shimodaira (Kyoto University)

Tu65 *Mixed Graphical Model with Ordinal Data: Parameter Estimation and Statistical Inference*
Huijie Feng (Cornell University)*; Yang Ning (Cornell University)

Tu66 *Robust Graph Embedding with Noisy Link Weights*
Akifumi Okuno (Kyoto University and RIKEN AIP)*; Hidetoshi Shimodaira (Kyoto University)

Tu67 *Exploring Fast and Communication-Efficient Algorithms in Large-scale Distributed Networks*
Yue Yu (Tsinghua University)*; Jiaxiang Wu (Tencent AI Lab); Junzhou Huang (University of Texas at Arlington)

Tu68 *Fisher Information and Natural Gradient Learning in Random Deep Networks*
Shun-ichi Amari (RIKEN BSI)*

Tu69 *Robust descent using smoothed multiplicative noise*
Matthew J Holland (Osaka University)*

Tu70 *Classification using margin pursuit*
Matthew J Holland (Osaka University)*

Tu71 *Linear Queries Estimation with Local Differential Privacy*
Raef Bassily (The Ohio State University)*

Tu72 *Bayesian Learning of Neural Network Architectures*
Georgi Dikov (Technical University of Munich)*; Justin Bayer (Volkswagen Group)

Tu73 *Nonlinear Acceleration of Primal-Dual Algorithms*
Raghu Bollapragada (Northwestern); Damien Scieur (Princeton University); Alexandre d'Aspremont (Ecole Normale Superieure)*

Tu74 *Gaussian Process Latent Variable Alignment Learning*
Ieva Kazlauskaite (University of Bath)*; Carl Henrik Ek (Bristol University); Neill Campbell (University of Bath)

Tu75 *Pseudo-Bayesian Learning with Kernel Fourier Transform as Prior*
Gaël Letarte (Université Laval)*; Emilie Morvant (University Jean Monnet, St-Etienne); Pascal Germain (Inria)

Tu76 *Forward Amortized Inference for Likelihood-Free Variational Marginalization*
Luca Ambrogioni (Donders Institute)*; Julia Berezutskaya (University of Utrecht); Eva van den Borne (Radboud University); Umut Güçlü (Radboud University, Donders Institute for Brain, Cognition and Behaviour); Yağmur Güçlütürk (Radboud University, Donders Institute for Brain, Cognition and Behaviour); Max Hinne (University of Amsterdam); Eric Maris (Donders Institute); Marcel van Gerven (Radboud University, Donders Institute for Brain, Cognition and Behaviour)

Tu77 *SpikeCaKe: Semi-Analytic Nonparametric Bayesian Inference forSpike-Spike Neuronal Connectivity*
Luca Ambrogioni (Donders Institute)*; Patrick Ebel (Radboud University); Max Hinne (University of Amsterdam); Umut Güçlü (Radboud University, Donders Institute for Brain, Cognition and Behaviour); Marcel van Gerven (Radboud University, Donders Institute for

Brain, Cognition and Behaviour); Eric Maris (Donders Institute)

Tu78 *Scalable Gaussian Process Inference with Finite-data Mean and Variance Guarantees*
Jonathan H Huggins (Harvard)*; Trevor Campbell (UBC); Mikolaj Kasprzak (Oxford); Tamara Broderick (MIT)

Tu79 *Exponential convergence rates for Batch Normalization: The power of length-direction decoupling in non-convex optimization*
Jonas Kohler (ETHZ)*; Hadi Daneshmand (ETH); Aurelien Lucchi (ETH Zurich); Klaus Neymeyr (University of Rostock); Ming Zhou (University of Rostock)

Tu80 *A new evaluation framework for topic modeling algorithms based on synthetic corpora*
Hanyu Shi (Northwestern university); Martin Gerlach (Northwestern University); Isabel Diersen (Northwestern University); Doug Downey (Northwestern University); Luis Amaral (Northwestern University)*

Tu81 *On Kernel Derivative Approximation with Random Fourier Features*
Zoltan Szabo (Ecole Polytechnique)*; Bharath Sriperumbudur (Penn State)

Tu82 *Sequential Neural Likelihood: Fast Likelihood-free Inference with Autoregressive Flows*
George Papamakarios (University of Edinburgh)*; David Sterratt (University of Edinburgh); Iain Murray (University of Edinburgh)

Tu83 *Optimal Transport for Multi-source Domain Adaptation under Target Shift*
Ievgen Redko (Laboratoire Hubert Curien)*; Rémi Flamary (Université Côte d'Azur); Nicolas Courty (UBS); Devis Tuia (Wageningen University and Research)

Tu84 *Deep Neural Networks Learn Non-Smooth Functions Effectively*
Masaaki Imaizumi (Institute of Statistical Mathematics)*; Kenji Fukumizu (The Institute of Statistical Mathematics)

Tu85 *Fisher-Rao Metric, Geometry, and Complexity of Neural Networks*
Tengyuan Liang (University of Chicago); Tomaso Poggio (MIT); Alexander Rakhlin (MIT); James Stokes ()*

Tu86 *Accelerated Decentralized Optimization with Local Updates for Smooth and Strongly Convex Objectives*
Hadrien Hendrikx (INRIA - Ecole Normale Supérieure)*; Laurent Massoulie (Microsoft-Inria Joint Center); Francis Bach (INRIA - Ecole Normale Supérieure)

Tu87 *Interaction Matters: A Note on Non-asymptotic Local Convergence of Generative Adversarial Networks*
Tengyuan Liang (University of Chicago)*; James Stokes ()

Tu88 *On Constrained Nonconvex Stochastic Optimization: A Case Study for Generalized Eigenvalue Decomposition*
Zhehui Chen (Georgia Tech); Xingguo Li (University of Minnesota); Lin Yang (Princeton University); Jarvis Haupt (UMN); Tuo Zhao (Georgia Tech)*

Tu89 *Generalized Boltzmann Machine with Deep Neural Structure*
Yingru Liu (Stony Brook University); Dongliang Xie (Beijing University of Posts and Telecommunications)*; xin wang (Department of Electrical and Computer Engineering, Stony Brook University)

Tu90 *Extreme Stochastic Variational Inference: Distributed Inference for Large Scale Mixture Models*
Parameswaran Raman (UC Santa Cruz)*; Jiong Zhang (UT-Austin); Shihao Ji (Georgia State University); Hsiang-Fu Yu (Amazon); S.V.N. Vishwanathan (UCSD / Amazon); Inderjit Dhillon (University of Texas at Austin)

Tu91 *Correcting the bias in least squares regression with volume-rescaled sampling*
Michal Derezinski (UC Berkeley)*; Manfred K. Warmuth (UCSC); Daniel Hsu (Columbia University)

Tu92 *Conservative Exploration using Interleaving*
Sumeet Katariya (Univ of Wisconsin-Madison)*; Branislav Kveton (Google Research); Zheng Wen (Adobe Research); Vamsi K Potluru (Comcast Cable)

Tu93 *Active Exploration in Markov Decision Processes*
Jean Tarbouriech (FAIR); Alessandro Lazaric (FAIR)*

Tu94 *On the Convergence of Stochastic Gradient Descent with Adaptive Stepsizes*
Xiaoyu Li (Boston University); Francesco Orabona (Boston University)*

Tu95 *Bandit Online Learning with Unknown Delays*
Bingcong Li (University of Minnesota)*; Tianyi Chen (University of Minnesota); Georgios B. Giannakis (University of Minnesota)

Tu96 *Learning Invariant Representations by Kernel Warping*
Yingyi Ma (UIC); Vignesh Ganapathiraman (UIC); Xinhua Zhang (UI Chicago)*

Tu97 *β^3-IRT: A New Item Response Model and its Applications*
YU CHEN (UNIVERSITY OF BRISTOL)*; Telmo M Silva Filho (Universidade Federal de Pernambuco); Ricardo B Prudencio (UFPE); Tom Diethe (Amazon); Peter Flach (University of Bristol)

Tu98 *Auditing Model Prediction Reliability After-the-Fact with Resampling Uncertainty Estimation*
Peter Schulam (Johns Hopkins University)*; Suchi Saria (Johns Hopkins University)

Tu99 *Universal Statistics of Fisher Information in Deep Neural Networks: Mean Field Approach*
Ryo Karakida (National Institute of Advanced Industrial Science and Technology)*; Shotaro Akaho (AIST); Shun-ichi Amari (RIKEN BSI)

Tu100 *Conditional Sparse L_p-norm Regression With Optimal Probability*
Brendan Juba (WASHINGTON UNIVERSITY ST LOUIS); David Woodruff (Carnegie Mellon University); Hai S Le (WASHINGTON UNIVERSITY ST LOUIS)*; John Hainline (Washington University in St. Louis)

Tu101 *On the Connection Between Learning Two-Layer Neural Networks and Tensor Decomposition*
Marco Mondelli (Stanford University)*; Andrea Montanari (Stanford University)

Tu102 *Autoencoding any Data through Kernel Autoencoders*
Pierre Laforgue (Telecom ParisTech)*; Stéphan Clémençon (Télécom ParisTech); Florence d'Alche-Buc (Télécom ParisTech)

Tu103 *Towards Understanding the Generalization Bias of Two Layer Convolutional Linear Classifiers with Gradient Descent*
Yifan Wu (Carnegie Mellon University)*; Barnabas Poczos (Carnegie Mellon University); Aarti Singh (Carnegie Mellon University)

Tu104 *Learning to Optimize under Non-Stationarity*
Wang Chi Cheung (IHPC, ASTAR); David Simchi-Levi (MIT); Ruihao Zhu (MIT)*

Tu105 *SPONGE: A generalized eigenproblem for clustering signed networks*
Mihai Cucuringu (University of Oxford and the Alan Turing Institute)*; Peter Davies (University of Warwick); Aldo Glielmo (Department of Physics, King's College London, Strand, London WC2R 2LS, United Kingdom); Hemant Tyagi (Alan Turing Institute)

Tu106 *Deep Neural Networks with Multi-Branch Architectures Are Intrinsically Less Non-Convex*
Hongyang Zhang (Carnegie Mellon University)*; Junru Shao (Carnegie Mellon University); Ruslan Salakhutdinov (Carnegie Mellon University)

Tu107 *Are we there yet? Manifold identification of gradient-related proximal methods*
Yifan Sun (University of British Columbia)*; Halyun Jeong (UBC); Julie Nutini (University of British Columbia); Mark Schmidt (University of British Columbia)

Tu108 *Hadamard Response: Estimating Distributions Privately, Efficiently, and with Little Communication*
Jayadev Acharya (Cornell University)*; Ziteng Sun (Cornell University); Huanyu Zhang (Cornell University)

Tu109 *Accelerated Bayesian Additive Regression Trees*
Jingyu He (University of Chicago)*; Saar Yalov (Arizona State University); P. Richard Hahn (Arizona State University)

Tu110 *Online Multiclass Boosting with Bandit Feedback*
Daniel T. Zhang (University of Michigan); Young Hun Jung (University of Michigan)*; Ambuj Tewari (University of Michigan)

Tu111 *Auto-Encoding Total Correlation Explanation*
Shuyang Gao (ISI USC)*; Rob Brekelmans (USC / ISI); Greg Ver Steeg (University of Southern California); Aram Galstyan (USC Information Sciences Institute)

Day 2: April 17 (Wednesday), 2019

We1 *Towards Efficient Data Valuation Based on the Shapley Value*
Ruoxi Jia (UC Berkeley)*; David Dao (ETH); Boxin Wang (Zhejiang University); Frances Ann Hubis (ETH Zurich); Nick Hynes (UC Berkeley); Bo Li (University of Illinois at Urbana–Champaign); Ce Zhang (ETH); Dawn Song (UC Berkeley); Costas J. Spanos (University of California at Berkeley)

We2 *Bayesian optimisation under uncertain inputs*
Rafael Oliveira (The University of Sydney)*; Lionel Ott (The University of Sydney); Fabio Ramos (U Sydney)

We3 *Optimal Minimization of the Sum of Three Convex Functions with a Linear Operator*
Seyoon Ko (Seoul National University); Donghyeon Yu (Inha University); Joong-Ho Won (Seoul National University)*

We4 *Fast and Faster Convergence of SGD for Over-Parameterized Models and an Accelerated Perceptron*
Sharan Vaswani (University of British Columbia)*; Francis Bach (INRIA - Ecole Normale Supérieure); Mark Schmidt (University of British Columbia)

We5 *No-regret algorithms for online k-submodular maximization*
Tasuku Soma (University of Tokyo)*

We6 *Lagrange Coded Computing: Optimal Design for Resiliency, Security, and Privacy*
Qian Yu (University of Southern California)*; Netanel Raviv (Caltech); Songze Li (University

of Southern California); Seyed Mohammadreza Mousavi Kalan (University of Southern California); Mahdi Soltanolkotabi (USC); Salman Avestimehr (University of Southern California)

We7 *Model Consistency for Learning with Mirror-Stratifiable Regularizers*
Jalal Fadili (GREYC, CNRS, ENSICAEN, Université de Caen); Guillaume Garrigos (École Normale Supérieure de Paris)*; Jérôme Malick (CNRS and LJK); Gabriel Peyré (CNRS and ENS)

We8 *From Cost-Sensitive to Tight F-measure Bounds*
Kevin Bascol (Université Saint-Etienne)*; Emonet Rémi (Laboratoire Hubert Curien); Elisa Fromont (IRISA, INRIA, FR); Amaury Habrard (Université Saint-Etienne); Guillaume METZLER (Université Saint Etienne); Marc Sebban (Université Saint-Etienne)

We9 *Feature subset selection for multinomial logit model via mixed-integer optimization*
Shunsuke Kamiya (Tokyo University of Agriculture and Technology)*; Ryuhei Miyashiro (Tokyo University of Agriculture and Technology); Yuichi Takano (University of Tsukuba)

We10 *Low-precision Random Fourier Features for Memory-constrained Kernel Approximation*
Jian Zhang (Stanford)*; Avner May (Stanford University); Tri Dao (Stanford University); Christopher Re (Stanford University)

We11 *Restarting Frank-Wolfe*
Thomas Kerdreux (INRIA/ ENS)*; Alexandre d'Aspremont (Ecole Normale Superieure); Sebastian Pokutta (Gatech)

We12 *Fast and Accurate Inference with Adaptive Ensemble Prediction for Deep Neural Networks based on Confidence Level*
Hiroshi Inoue (IBM Research - Tokyo)*

We13 *Infinite Task Learning in RKHSs*
Romain R Brault (Telecom ParisTech); Alex Lambert (Télécom ParisTech)*; Zoltan Szabo (Ecole Polytechnique); Florence d'Alche-Buc (Télécom ParisTech); Maxime Sangnier (Sorbonne University)

We14 *Detection of Planted Solutions for Flat Satisfiability Problems*
Quentin Berthet (University of Cambridge)*; Jordan Ellenberg (University of Wisconsin-Madison)

We15 *Markov Properties of Discrete Determinantal Point Processes*
Kayvan Sadeghi (University College London)*; Alessandro Rinaldo (Carnegie Mellon University)

We16 *Analysis of Thompson Sampling for Combinatorial Multi-armed Bandit with Probabilistically Triggered Arms*
Alihan Huyuk (Bilkent University)*; Cem Tekin (Bilkent University)

We17 *Support Localization and the Fisher Metric for off-the-grid Sparse Regularization*
Clarice Poon (DAMTP, University of Cambridge); Nicolas Keriven (Ecole Normale Supérieure)*; Gabriel Peyré (CNRS and ENS)

We18 *Fast Gaussian process based gradient matching for parameter identification in systems of nonlinear ODEs*
Philippe Wenk (ETH Zurich)*; Alkis Gotovos (ETH); Stefan Bauer (MPI IS); Nico S Gorbach (); Andreas Krause (ETH Zürich); Joachim Buhmann (ETH Zurich)

We19 *Semi-Generative Modelling: Covariate-Shift Adaptation with Cause and Effect Features*
Julius von Kügelgen (University of Cambridge)*; Marco Loog (TU Delft); Alexander Mey (TU Delft)

We20 *A Continuous-Time View of Early Stopping for Least Squares Regression*
Alnur Ali ()*; Ryan Tibshirani (Carnegie Mellon University); Zico Kolter (Carnegie Mellon

University)

We21 *Towards Clustering High-dimensional Gaussian Mixture Clouds in Linear Running Time*
Dan Kushnir (Nokia Bell Labs)*; Shirin Jalali (Bell Labs); Iraj Saniee (Nokia Bell Labs)

We22 *Classifying Signals on Irregular Domains via Convolutional Cluster Pooling*
Angelo Porrello (University of Modena and Reggio Emilia)*; Davide Abati (University of Modena and Reggio Emilia); SIMONE CALDERARA (University of Modena and Reggio Emilia, Italy); Rita Cucchiara (Universita Di Modena E Reggio Emilia)

We23 *Wasserstein regularization for sparse multi-task regression*
Hicham Janati (INRIA)*; Marco Cuturi (ENSAE/CREST); Alexandre Gramfort (Inria)

We24 *Black Box Quantiles for Kernel Learning*
Anthony P Tompkins (The University of Sydney)*; Ransalu Senanayake (University of Sydney); Philippe Morere (The University of Sydney); Fabio Ramos (U Sydney)

We25 *Adversarial Variational Optimization of Non-Differentiable Simulators*
Gilles Louppe (University of Liège)*; Joeri Hermans (University of Liège); Kyle Cranmer (New York University)

We26 *Active Probabilistic Inference on Matrices for Pre-Conditioning in Stochastic Optimization*
Filip de Roos (Max Planck Institute for Intelligent Systems)*; Philipp Hennig (University of Tübingen)

We27 *Projection Free Online Learning over Smooth Sets*
Kfir Yehuda Levy (ETH)*; Andreas Krause (ETH Zürich)

We28 *Confidence Scoring Using Whitebox Meta-models with Linear Classifier Probes*
Tongfei Chen (Johns Hopkins University)*; Jiri Navratil (IBM Thomas J. Watson Research Center); Vijay Iyengar (IBM Research); Karthikeyan Shanmugam (IBM Research)

We29 *Learning Influence-Receptivity Network Structure with Guarantee*
Ming Yu (University of Chicago)*; Varun Gupta (University of Chicago Booth School of Business); Mladen Kolar (University of Chicago Booth School of Business)

We30 *Iterative Bayesian Learning for Crowdsourced Regression*
Jungseul Ok (UIUC); Sewoong Oh (UIUC); Jinwoo Shin (KAIST); Yung Yi (KAIST); Yunhun Jang (KAIST)*

We31 *Nonconvex Matrix Factorization from Rank-One Measurements*
Yuanxin Li (Carnegie Mellon University); Cong Ma (Princeton University); Yuxin Chen (Princeton University); Yuejie Chi (CMU)*

We32 *Fast and Robust Shortest Paths on Manifolds Learned from Data*
Georgios Arvanitidis (Technical University of Denmark)*; Soren Hauberg (Technical University of Denmark, Denmark); Philipp Hennig (University of Tübingen); Michael Schober (Bosch Center for Artificial Intelligence)

We33 *Training a Spiking Neural Network with Equilibrium Propagation*
Peter E.D. O'Connor (University of Amsterdam)*

We34 *Learning One-hidden-layer ReLU Networks via Gradient Descent*
Xiao Zhang (University of Virginia)*; Yaodong Yu (University of Virginia); Lingxiao Wang (University of California, Los Angeles); Quanquan Gu (University of California, Los Angeles)

We35 *Gain estimation of linear dynamical systems using Thompson Sampling*
Matias I Müller (KTH Royal Institute of Technology)*; Cristian R Rojas (KTH Royal Institute of Technology)

We36 *Universal Hypothesis Testing with Kernels: Asymptotically Optimal Tests for Goodness of Fit*

Shengyu Zhu (Huawei Noah's Ark Lab)*; Biao Chen (Syracuse University); Pengfei Yang (Cubist Systematic Strategies); Zhitang Chen (Huawei Noah's Ark Lab)

We37 *Calibrating Deep Convolutional Gaussian Processes*
Gia-Lac Tran (EURECOM); Edwin V Bonilla (Data61); John Cunningham (); Pietro Michiardi (EURECOM); Maurizio Filippone (EURECOM)*

We38 *Sample Complexity of Sinkhorn Divergences*
Aude Genevay (U Paris Dauphine)*; Marco Cuturi (ENSAE/CREST); Gabriel Peyré (CNRS and ENS); Francis Bach (INRIA - Ecole Normale Supérieure); Lénaïc Chizat (INRIA)

We39 *Adaptive Gaussian Copula ABC*
Yanzhi Chen (University of Edinburgh)*; Michael U. Gutmann (University of Edinburgh)

We40 *Top Feasible Arm Identification*
Julian Katz-Samuels (University of Michigan)*; Clay Scott ()

We41 *Direct Acceleration of SAGA using Sampled Negative Momentum*
Kaiwen Zhou (The Chinese University of Hong Kong)*; Qinghua Ding (Tsinghua University); Fanhua Shang (Xidian University); James Cheng (CUHK); Danli Li (The Chinese University of Hong Kong); Zhiquan Luo (The Chinese University of Hong Kong)

We42 *Does data interpolation contradict statistical optimality?*
Mikhail Belkin (Ohio State University); Alexander Rakhlin (MIT)*; Alexandre Tsybakov (CREST, ENSAE)

We43 *Inverting Supervised Representations with Autoregressive Neural Density Models*
Charlie Nash (The University of Edinburgh)*; Nate Kushman (Microsoft Research); Chris Williams (Edinburgh)

We44 *Connecting Weighted Automata and Recurrent Neural Networks through Spectral Learning*
guillaume rabusseau (McGill)*; Tianyu Li (McGill University); Doina Precup (McGill University)

We45 *A Family of Exact, Distribution-Free Goodness-of-Fit Tests for High-Dimensional Discrete Distributions*
Feras Saad (Massachusetts Institute of Technology)*; Cameron Freer (MIT); Nate Ackerman (Harvard University); Vikash Mansinghka (Massachusetts Institute of Technology)

We46 *Differentially Private Online Submodular Minimization*
Rachel Cummings (Georgia Tech)*; Adrian Rivera Cardoso (Georgia Tech)

We47 *Semi-supervised clustering for de-duplication*
Shrinu Kushagra (University of Waterloo)*; Shai Ben-David (University of Waterloo); Ihab F Ilyas (U. of Waterloo)

We48 *Finding the Bandit in a Graph: Sequential Search-and-Stop*
Pierre Perrault (Inria Lille - Nord Europe)*; Vianney Perchet (ENS Paris-Saclay & Criteo); Michal Valko (Inria)

We49 *Statistical Learning under Nonstationary Mixing Processes*
Steve Hanneke (Toyota Technological Institute at Chicago)*; Liu Yang (Independent)

We50 *On Structure Priors for Learning Bayesian Networks*
Ralf Eggeling (University of Tübingen)*; Jussi Viinikka (); Aleksis Vuoksenmaa (University of Helsinki); Mikko Koivisto ()

We51 *Partial Optimality of Dual Decomposition for MAP Inference in Pairwise MRFs*
Alex Bauer (TU Berlin)*; Shinichi Nakajima (Technische Universität Berlin); Nico Goernitz (TU Berlin); Klaus-Robert Müller (Technische Universität Berlin)

We52 *Sparse Feature Selection in Kernel Discriminant Analysis via Optimal Scoring*
Alexander F Lapanowski (Texas A&M University)*; Irina Gaynanova (Texas A&M University)

We53 *Learning Natural Programs from a Few Examples in Real-Time*
Nagarajan Natarajan (Microsoft Research)*; Danny Simmons (Microsoft); Naren Datha (Microsoft Research); Prateek Jain (Microsoft Research); Sumit Gulwani (Microsoft Research)

We54 *Truncated Back-propagation for Bilevel Optimization*
Amirreza Shaban (Georgia Institute of Technology)*; Ching-An Cheng (Georgia Institute of Technology); Nathan Hatch (Georgia Institute of Technology); Byron Boots (Georgia Institute of Technology)

We55 *Variable selection for Gaussian processes via sensitivity analysis of the posterior predictive distribution*
Topi Paananen (Aalto University)*; Juho Piironen (Aalto University); Michael R Andersen (Aalto University); Aki Vehtari (Aalto University)

We56 *Lifted Weight Learning of Markov Logic Networks Revisited*
Ondrej Kuzelka (University of Leuven)*; Vyacheslav Kungurtsev (Czech Technical University)

We57 *Causal discovery in the presence of missing data*
Ruibo Tu (KTH Royal Institute of Technology)*; Cheng Zhang (Microsoft); Paul Ackermann (Karolinska Institutet); Karthika Mohan (U C Berkeley); Hedvig Kjellström (KTH Royal Institute of Technology); Kun Zhang (Carnegie Mellon University)

We58 *Learning Tree Structures from Noisy Data*
Konstantinos Nikolakakis (Rutgers University)*; Dionysios Kalogerias (Princeton University); Anand D Sarwate (Rutgers University)

We59 *Active multiple matrix completion with adaptive confidence sets*
Andrea Locatelli (Uni Magdeburg)*; Alexandra Carpentier (Otto-von-Guericke-Universität Magdeburg); Michal Valko (Inria)

We60 *Confidence-based Graph Convolutional Networks for Semi-Supervised Learning*
Shikhar Vashishth (Indian Institute of Science)*; Prateek Yadav (Indian Institute of Science); Manik Bhandari (Indian Institute of Science); Partha Talukdar (Indian Institute of Science)

We61 *Negative Momentum for Improved Game Dynamics*
Gauthier Gidel (MILA)*; Reyhane Askari Hemmat (MILA); Mohammad Pezeshki (MILA); Gabriel Huang (MILA); Rémi Le Priol (MILA); Simon Lacoste-Julien (University of Montreal); Ioannis Mitliagkas (University of Montreal)

We62 *Data-dependent compression of random features for large-scale kernel approximation*
Raj Agrawal (MIT)*; Jonathan H Huggins (Harvard); Trevor Campbell (UBC); Tamara Broderick (MIT)

We63 *Large-Margin Classification in Hyperbolic Space*
Hyunghoon Cho (MIT)*; Benjamin DeMeo (Harvard University); Jian Peng (UIUC); Bonnie Berger (MIT)

We64 *Generalizing the theory of cooperative inference*
Pei Wang (Rutgers University-Newark)*; Pushpi Paranamana (Rutgers University-Newark); Patrick Shafto (Rutgers University)

We65 *MaxHedge: Maximizing a Maximum Online with Theoretical Performance Guarantees*
Stephen U Pasteris (University College London)*; Fabio Vitale (University of Lille); Kevin Chan (US army); Shiqiang Wang (IBM Research)

We66 *The Gaussian Process Autoregressive Regression Model (GPAR)*
James R Requeima (University of Cambridge)*; Wessel P Bruinsma (University of Cambridge); William Tebbutt (University of Cambridge); Richard Turner ()

We67 *Towards Optimal Transport with Global Invariances*
David Alvarez-Melis (MIT)*; Stefanie Jegelka (MIT); Tommi Jaakkola (MIT)

We68 *Unsupervised Alignment of Embeddings with Wasserstein Procrustes*
Edouard Grave (Facebook AI Research)*; Armand Joulin (Facebook AI Research); Quentin Berthet (University of Cambridge)

We69 *Sequential Patient Recruitment and Allocation for Adaptive Clinical Trials*
Onur Atan (UCLA)*; William Zame (UCLA); Mihaela van der Schaar ()

We70 *Probabilistic Forecasting with Spline Quantile Function RNNs*
Konstantinos Benidis (Amazon); Jan Gasthaus (Amazon Research)*; Bernie Wang (Amazon); Syama Sundar Rangapuram (Amazon); David Salinas (Amazon); Valentin Flunkert (Amazon); Tim Januschowski (Amazon Research)

We71 *Exponential Weights on the Hypercube in Polynomial Time*
Sudeep Raja Putta (University of Massachusetts Amherst)*; Abhishek Shetty (Microsoft Research)

We72 *Sharp Analysis of Learning with Discrete Losses*
Alex Nowak (INRIA, Ecole Normale Supérieure); Francis Bach (INRIA - Ecole Normale Supérieure); Alessandro Rudi (INRIA, Ecole Normale Superieure)*

We73 *Designing Optimal Binary Rating Systems*
Nikhil Garg (Stanford University)*; Ramesh Johari (Stanford University)

We74 *Stochastic Negative Mining for Learning with Large Output Spaces*
Sashank Reddi (Google)*; Satyen Kale (Google); Felix Yu (Google); Daniel Holtmann-Rice (Google); Jiecao Chen (Indiana University Bloomington)

We75 *Learning One-hidden-layer Neural Networks under General Input Distributions*
Weihao Gao (UIUC); Ashok V Makkuva (University of Illinois at Urba); Sewoong Oh (UIUC)*; Pramod Viswanath (UIUC)

We76 *A Geometric Perspective on the Transferability of Adversarial Directions*
Zachary B Charles (University of Wisconsin - Madison)*; Harrison Rosenberg (University of Wisconsin-Madison); Dimitris Papailiopoulos (University of Wisconsin-Madison)

We77 *Non-linear process convolutions for multi-output Gaussian processes*
Mauricio A Alvarez (University of Sheffield)*; Wil Ward (University of Sheffield)

We78 *Lovász Convolutional Networks*
Prateek Yadav (Indian Institute of Science)*; Madhav R Nimishakavi (Indian Institute of Science); Naganand Yadati (Indian Institute of Science); Shikhar Vashishth (Indian Institute of Science); Arun Rajkumar (Conduent Labs); Partha Talukdar (Indian Institute of Science)

We79 *Bridging the gap between regret minimization and best arm identification, with application to A/B tests*
Rémy Degenne (CWI); Thomas Nedelec (Criteo)*; Clement Calauzenes (Criteo); Vianney Perchet (ENS Paris-Saclay)

We80 *Gaussian process modulated Cox processes under linear inequality constraints*
Andrés F LOPEZ-LOPERA (Mines Saint-Etienne)*; ST John (PROWLER.io); Nicolas Durrande (PROWLER.io)

We81 *Implicit Kernel Learning*
Chun-Liang Li (Carnegie Mellon University)*; Wei-Cheng Chang (Carnegie Mellon University); Youssef Mroueh (IBM Research); Yiming Yang (Carnegie Mellon University); Barnabas Poczos (Carnegie Mellon University)

We82 *Bounding Inefficiency of Equilibria in Continuous Actions Games using Submodularity and Curvature*
Pier Giuseppe PGS Sessa (ETH Zürich)*; Maryam Kamgarpour (ETH Zürich); Andreas Krause (ETH Zürich)

We83 *Variational Information Planning for Sequential Decision Making*
Jason Pacheco (Brown University)*; John Fisher (MIT)

We84 *Projection-Free Bandit Convex Optimization*
Lin Chen (Yale University)*; Mingrui Zhang (Yale University); Amin Karbasi (Yale)

We85 *Provable Robustness of ReLU networks via Maximization of Linear Regions*
Francesco Croce (Saarland University); Maksym Andriushchenko (Saarland University); Matthias Hein (University of Tuebingen)*

We86 *Test without Trust: Optimal Locally Private Distribution Testing*
Jayadev Acharya (Cornell University); Clement Canonne (Stanford University)*; Cody Freitag (Cornell University); Himanshu Tyagi (IISC)

We87 *Distributed Maximization of Submodular plus Diversity Functions for Multi-label Feature Selection on Huge Datasets*
Mehrdad Ghadiri (University of British Columbia)*; Mark Schmidt (University of British Columbia)

We88 *On Euclidean k-Means Clustering with alpha-Center Proximity*
Amit Deshpande (Microsoft Research); Anand Louis (Indian Institute of Science, Bangalore, India); Apoorv V Singh (Indian Institute of Science)*

We89 *Noisy Blackbox Optimization using Multi-fidelity Queries: A Tree Search Approach*
Rajat Sen (University of Texas at Austin)*; Kirthevasan Kandasamy (Carnegie Mellon University); Sanjay Shakkottai (University of Texas at Austin)

We90 *Safe Convex Learning under Uncertain Constraints*
Ilnura Usmanova (ETH Zurich)*; Andreas Krause (ETH Zürich); Maryam Kamgarpour ()

We91 *The non-parametric bootstrap and spectral analysis in moderate and high-dimension*
Noureddine El Karoui (UC Berkeley)*; Elizabeth Purdom (UC Berkeley)

We92 *Knockoffs for the Mass: New Feature Importance Statistics with False Discovery Guarantees*
Jaime Roquero Gimenez (Stanford University); Amirata Ghorbani (Stanford University); James Zou (Stanford University)*

We93 *Training Variational Autoencoders with Buffered Stochastic Variational Inference*
Rui Shu (Stanford University); Hung Bui (Google)*; Jay Whang (Stanford University); Stefano Ermon (Stanford University)

We94 *Regularized Contextual Bandits*
Xavier Fontaine (ENS Paris-Saclay)*; Vianney Perchet (Ecole Normale Supérieure Paris-Saclay, Université Paris Saclay); Quentin Berthet (University of Cambridge)

We95 *Risk-Sensitive Generative Adversarial Imitation Learning*
Jonathan Lacotte (Stanford University)*; Mohammad Ghavamzadeh (FAIR); Yinlam Chow (DeepMind); Marco Pavone (Stanford University)

We96 *Learning Controllable Fair Representations*
Jiaming Song (Stanford)*; Pratyusha Kalluri (Stanford University); Aditya Grover (Stanford

University); Shengjia Zhao (Stanford University); Stefano Ermon (Stanford University)

We97 *Multi-Task Time Series Analysis applied to Drug Response Modelling*
Alex Bird (Alan Turing Institute)*; Chris Williams (Edinburgh)

We98 *Improving the Stability of the Knockoff Procedure: Multiple Simultaneous Knockoffs and Entropy Maximization*
Jaime Roquero Gimenez (Stanford University); James Zou (Stanford University)*

We99 *Know Your Boundaries: Constraining Gaussian Processes by Variational Harmonic Features*
Arno Solin (Aalto University)*; Manon Kok (Delft University of Technology)

We100 *Distributional reinforcement learning with linear function approximation*
Marc G. Bellemare (Google Brain)*; Nicolas Le Roux (Google); Pablo Samuel Castro (Google); Subhodeep Moitra (Google, Inc.)

We101 *Matroids, Matchings, and Fairness*
Flavio Chierichetti (Sapienza University); Ravi Kumar (Google)*; Silvio Lattanzi (Google); Sergei Vassilvtiskii (Google)

We102 *Dynamical Isometry is Achieved in Residual Networks in a Universal Way for any Activation Function*
Wojciech Tarnowski (Jagiellonian University in Kraków)*; Piotr Warchoł (Jagiellonian University in Kraków); Stanisław Jastrzębski (Jagiellonian University); Jacek Tabor (Jagiellonian University in Kraków); Maciej Nowak (Jagiellonian University in Kraków)

We103 *The Termination Critic*
Anna Harutyunyan (DeepMind)*; Will Dabney (DeepMind); Diana Borsa (DeepMind); Nicolas Heess (DeepMind); Remi Munos (DeepMind); Doina Precup (McGill University)

We104 *Consistent Online Optimization: Convex and Submodular*
Mohammad Reza Karimi Jaghargh (ETH Zurich)*; Andreas Krause (ETH Zürich); Silvio Lattanzi (Google); Sergei Vassilvtiskii (Google)

We105 *Learning Determinantal Point Processes by Corrective Negative Sampling*
Zelda Mariet (Massachusetts Institute of Technology)*; Mike Gartrell (Criteo AI Lab); Suvrit Sra (Massachusetts Institute of Technology, USA)

We106 *Probabilistic Semantic Inpainting with Pixel Constrained CNNs*
Emilien Dupont (Schlumberger)*

We107 *Least Squares Estimation of Weakly Convex Functions*
Sun Sun (University of Waterloo); Yaoliang Yu (University of Waterloo)*

We108 *Interval Estimation of Individual-Level Causal Effects Under Unobserved Confounding*
Nathan Kallus (Cornell Tech)*; Xiaojie Mao (Cornell University); Angela Zhou (Cornell University)

We109 *Amortized Variational Inference with Graph Convolutional Networks for Gaussian Processes*
Linfeng Liu (Tufts University)*; Liping Liu (Tufts University)

We110 *Online Decentralized Leverage Score Sampling for Streaming Multidimensional Time Series*
Rui Xie (University of Georgia); Zengyan Wang (University of Georgia); Shuyang Bai (University of Georgia); Ping Ma (University of Georgia); Wenxuan Zhong ()*

We111 *Interpretable Cascade Classifiers with Abstention*
Matthieu Clertant (University Paris 6)*; Nataliya Sokolovska (University Paris 6); Yann Chevaleyre (Université Paris Dauphine); Blaise HANCZAR (Université d Evry)

Day 3: April 18 (Thursday), 2019

Th1 *Kernel Exponential Family Estimation via Doubly Dual Embedding*
Bo Dai (Google Brain)*; Hanjun Dai (Georgia Tech); Arthur Gretton (Gatsby Computational Neuroscience Unit); Dale E Schuurmans (Google Inc.); Le Song (Ant Financial & Georgia Institute of Technology); Niao He (University of Illinois at Urbana-Champaign)

Th2 *Revisiting Adversarial Risk*
Arun Sai Suggala (Carnegie Mellon University)*; Adarsh Prasad (Carnegie Mellon University); Pradeep Ravikumar (Carnegie Mellon University)

Th3 *A Memoization Framework for Scaling Submodular Optimization to Large Scale Problems*
Rishabh Krishnan Iyer (Microsoft Corporation)*; Jeffrey Bilmes (University of Washington)

Th4 *Bernoulli Race Particle Filters*
Sebastian M Schmon (University of Oxford)*; Arnaud Doucet (Oxford University); George Deligiannidis (Oxford)

Th5 *Augmented Ensemble MCMC sampling in Factorial Hidden Markov Models*
Kaspar Martens (University of Oxford); Michalis Titsias (Athens University); Christopher Yau (University of Birmingham)*

Th6 *Probabilistic Riemannian submanifold learning with wrapped Gaussian process latent variable models*
Anton Mallasto (University of Copenhagen)*; Soren Hauberg (Technical University of Denmark, Denmark); Aasa Feragen (University of Copenhagen, Denmark)

Th7 *Two-temperature logistic regression based on the Tsallis divergence*
Ehsan Amid (UCSC)*; Manfred K. Warmuth (UCSC); Sriram Srinivasan (UC Santa Cruz)

Th8 *Avoiding Latent Variable Collapse with Generative Skip Models*
Adji Bousso Dieng (Columbia University)*; Yoon Kim (Harvard University); Alexander Rush (Harvard); David Blei (Columbia University)

Th9 *SMOGS: Social Network Metrics of Game Success*
Fan Bu (Duke University)*; Sonia Xu (Duke University); Katherine Heller (Duke University); Alexander Volfovsky (Duke University)

Th10 *Fast Algorithms for Sparse Reduced-Rank Regression*
Benjamin Dubois (Ecole des Ponts ParisTech)*; Guillaume Obozinski (Ecole des Ponts ParisTech); Jean-François Delmas (Ecole des Ponts ParisTech)

Th11 *Stay Positive: The Benefits of Better Models in Stochastic Optimization*
Hilal Asi (Stanford University)*; John Duchi (Stanford University)

Th12 *Online learning with feedback graphs and switching costs*
Anshuka Rangi (University of California San Diego)*; Massimo Franceschetti (UC San Diego)

Th13 *Interpretable Almost-Exact Matching for Causal Inference*
Awa Dieng (Duke University, USA); Yameng Liu (Duke University, USA); Sudeepa Roy (Duke University, USA); Cynthia Rudin (Duke)*; Alexander Volfovsky (Duke University)

Th14 *Statistical Optimal Transport via Factored Couplings*
Aden Forrow (MIT)*; Jan-Christian Hütter (MIT); Mor Nitzan (Broad Institute); Philippe

Rigollet (MIT); Geoffrey Schiebinger (MIT, Broad Institute); Jonathan Weed (MIT)

Th15 *HS^2: Active Learning over Hypergraphs*
I Chien (UIUC)*; Huozhi Zhou (UIUC); Pan Li (UIUC)

Th16 *Clustering Time Series with Nonlinear Dynamics: A Bayesian Non-Parametric and Particle-Based Approach*
Alexander Lin (Harvard University)*; Yingzhuo Zhang (Harvard University); Jeremy Heng (Harvard University); Stephen Allsop (Massachusetts Institute of Technology); Kay Tye (Salk Institute for Biological Sciences); Pierre Jacob (Harvard University); Demba Ba (Harvard)

Th17 *Efficient Nonconvex Empirical Risk Minimization via Adaptive Sample Size Methods*
Aryan Mokhtari (MIT)*; Asuman Ozdaglar (MIT); Ali Jadbabaie ()

Th18 *An Optimal Control Approach to Sequential Machine Teaching*
Laurent Lessard (University of Wisconsin-Madison); Xuezhou Zhang (University of Wisconsin-Madison)*; Xiaojin Zhu (University of Wisconsin-Madison)

Th19 *Smoothed Online Optimization for Regression and Control*
Gautam Goel (California Institute of Technology)*; Adam Wierman (California Institute of Technology)

Th20 *Variational Compressive Sensing using Uncertainty Autoencoders*
Aditya Grover (Stanford University)*; Stefano Ermon (Stanford University)

Th21 *Structured Disentangled Representations*
Babak Esmaeili (Northeastern University)*; Hao Wu (Northeastern University); Sarthak Jain (Northeastern University); Alican Bozkurt (Northeastern University); N Siddharth (Unversity of Oxford); Brooks Paige (Alan Turing Institute); Dana Brooks (Northeastern University); Jennifer Dy (Northeastern); Jan-Willem van de Meent (Northeastern)

Th22 *Locally Private Mean Estimation: Z-test and Tight Confidence Intervals*
Marco Gaboardi (Univeristy at Buffalo); Ryan Rogers (); Or Sheffet (University of Alberta)*

Th23 *Estimation of Non-Normalized Mixture Models*
Takeru Matsuda (U Tokyo)*; Aapo Hyvarinen (UCL & U Helsinki)

Th24 *A Topological Regularizer for Classifiers via Persistent Homology*
Chao Chen (Stony Brook University)*; Xiuyan Ni (City University of New York); Qinxun Bai (Boston University); Yusu Wang (Ohio State University)

Th25 *Overcomplete Independent Component Analysis via SDP*
Anastasia Podosinnikova (MIT)*; Amelia Perry (MIT); Alex Wein (MIT); Alex Wein (NYU); Francis Bach (INRIA - Ecole Normale Supérieure); Alexandre d'Aspremont (Ecole Normale Superieure); David Sontag (MIT)

Th26 *Doubly Semi-Implicit Variational Inference*
Dmitry Molchanov (National Research University Higher School of Economics, Samsung)*; Valery Kharitonov (National Research University Higher School of Economics); Artem Sobolev (Samsung); Dmitry P Vetrov (Higher School of Economics)

Th27 *LocalNysation: A bottom up approach to efficient localized kernel regression*
Nicole Muecke (University of Stuttgart)*

Th28 *Scalable High-Order Gaussian Process Regression*
Shandian Zhe (University of Utah)*; Wei Xing (University of Utah); Robert Kirby (University of Utah)

Th29 *Bayesian Learning of Conditional Kernel Mean Embeddings for Automatic Likelihood-Free Inference*
Kelvin Hsu (University of Sydney)*; Fabio Ramos (U Sydney)

Th30 *Parallel Asynchronous Stochastic Coordinate Descent with Auxiliary Variables*
Hsiang-Fu Yu (Amazon)*; Cho-Jui Hsieh (UCLA, Google Research); Inderjit Dhillon (University of Texas at Austin)

Th31 *Credit Assignment Techniques in Stochastic Computation Graphs*
Theophane Weber (DeepMind)*; Nicolas Heess (DeepMind); Lars Buesing (DeepMind); David Silver (-)

Th32 *Efficient Bayesian Optimization for Target Vector Estimation*
Anders Kirk Uhrenholt (University of Glasgow)*; Bjoern Sand Jensen (University of Glasgow)

Th33 *Correspondence Analysis Using Neural Networks*
Hsiang Hsu (Harvard University)*; Salman Salamatian (MIT); Flavio Calmon (Harvard University)

Th34 *Interpolating between Optimal Transport and MMD using Sinkhorn Divergences*
Jean Feydy (École Normale Supérieure)*; Thibault Séjourné (ENS); Alain TROUVE (Ecole Normale Superieure de Cachan); François-Xavier Vialard (Université de Marne-la-Vallée); Gabriel Peyré (CNRS and ENS)

Th35 *Multi-Observation Regression*
Rafael Frongillo (CU Boulder); Nishant Mehta (University of Victoria)*; Tom Morgan (Harvard University); Bo Waggoner (University of Colorado)

Th36 *Adaptive MCMC via Combining Local Samplers*
Kiarash Shaloudegi (Imperial College London); Andras Gyorgy (DeepMind)*

Th37 *Variance reduction properties of the reparameterization trick*
Ming Xu (University of New South Wales); Matias Quiroz (University of New South Wales)*; Robert Kohn (University of New South Wales); Scott SIsson ()

Th38 *Hierarchical Clustering for Euclidean Data*
Vaggos Chatziafratis (Stanford University, California)*; Moses Charikar (Stanford University, California); Rad Niazadeh (Stanford University, California); Grigory Yaroslavtsev (Indiana University, Bloomington)

Th39 *Stochastic Variance-Reduced Cubic Regularization for Nonconvex Optimization*
Zhe Wang (Ohio State University)*; Yi Zhou (Ohio State University); Yingbin Liang (The Ohio State University); Guanghui Lan (Georgia Tech)

Th40 *Variational Noise-Contrastive Estimation*
Benjamin J Rhodes (University of Edinburgh)*; Michael U. Gutmann (University of Edinburgh)

Th41 *Improving Quadrature for Constrained Integrands*
Henry R Chai (Washington University in St. Louis)*; Roman Garnett (-)

Th42 *High Dimensional Inference in Partially Linear Models*
Ying Zhu (Purdue University)*; Zhuqing Yu (AbbVie Inc); Guang Cheng (Purdue University)

Th43 *Cost aware Inference for IoT Devices*
Pengkai Zhu (Boston University)*; Nan Feng (); Durmus Alp Emre Acar (Boston University); Prateek Jain (Microsoft Research); Venkatesh Saligrama (Boston University)

Th44 *Banded Matrix Operators for Gaussian Markov Models in the Automatic Differentiation Era*
Nicolas Durrande (PROWLER.io)*; Vincent Adam (PROWLER.io); Lucas Bordeaux (PROWLER.io); Stefanos Eleftheriadis (Prowler.io); James Hensman (PROWLER.io)

Th45 *A Unified Weight Learning Paradigm for Multi-view Learning*
Lai Tian (Northwestern Polytechnical University)*; Feiping Nie (Northwestern Polytechnical University); Xuelong Li (Northwestern Polytechnical University, China; Chinese Academy of Science, China)

Th46 *Region-Based Active Learning*
Corinna Cortes (Google); Giulia DeSalvo (Google); Claudio Gentile (Google Research); Mehryar Mohri (NYU); Ningshan Zhang (NYU)*

Th47 *Precision Matrix Estimation with Noisy and Missing Data*
Roger Fan (University of Michigan)*; Byoungwook Jang (University of Michigan); Yuekai Sun (University of Michigan); Shuheng Zhou (University of California, Riverside)

Th48 *Exploring k out of Top ρ Fraction of Arms in Stochastic Bandits*
Wenbo Ren (Ohio State University)*; Jia Liu (Iowa State University); Ness Shroff (The Ohio State University)

Th49 *AutoML from Service Provider's Perspective: Multi-device, Multi-tenant Model Selection with GP-EI*
Chen Yu (University of Rochester)*; Bojan Karlaš (); Jie Zhong (Cal State LA); Ce Zhang (ETH); Ji Liu (University of Rochester)

Th50 *On Theory for BART*
Veronika Rockova (University of Chicago)*; Enakshi Saha (University of Chicago)

Th51 *Deep Conditioned Poisson Factor Model for Multi-label Learning*
Rajat Panda (IIT Kanpur); Ankit Pensia (Indian Institute of Technology Kanpur); Mingyuan Zhou (University of Texas at Austin); Piyush Rai (IIT Kanpur)*

Th52 *On the Dynamics of Gradient Descent for Autoencoders*
Thanh V Nguyen (Iowa State University)*; Chinmay Hegde (Iowa State University); Raymond K. W. Wong (Texas A&M University)

Th53 *Complexities in Projection-Free Stochastic Non-convex Minimization*
Zebang Shen (Zhejiang University; Tencent AI Lab)*; Hui Qian (Zhejiang University); Cong Fang (Peking University); Peilin Zhao (Tencent AI Lab); Junzhou Huang (University of Texas at Arlington)

Th54 *Differentiable Antithetic Sampling for Variance Reduction in Stochastic Variational Inference*
Mike H Wu (Stanford University)*; Stefano Ermon (Stanford University); Noah Goodman (Stanford University)

Th55 *Efficient Greedy Coordinate Descent for Composite Problems*
Sai Praneeth Karimireddy (EPFL)*; Anastasia Koloskova (EPFL); Sebastian Stich (EPFL); Martin Jaggi (EPFL)

Th56 *Decentralized Continuous Submodular Maximization*
Jiahao Xie (Zhejiang University)*; Chao Zhang (Zhejiang University); Zebang Shen (Zhejiang University; Tencent AI Lab); Chao Mi (Zhejiang University); Hui Qian (Zhejiang University)

Th57 *Adaptive Rao-Blackwellisation in Gibbs Sampling for Probabilistic Graphical Models*
Craig Kelly (University of Memphis)*; Somdeb Sarkhel (Adobe); Deepak Venugopal (University of Memphis)

Th58 *Derivative-Free Methods for Policy Optimization: Guarantees for Linear Quadratic Systems*
Dhruv Malik (UC Berkeley); Ashwin Pananjady (UC Berkeley)*; Kush Bhatia (UC Berkeley); Koulik Khamaru (University of California Berkeley); Peter Bartlett ("University of California, Berkeley"); Martin Wainwright (University of California at Berkeley)

Th59 *Contrasting Exploration in Parameter and Action Space: A Zeroth-Order Optimization Perspective*
Anirudh Vemula (Carnegie Mellon University)*; Wen Sun (Carnegie Mellon University); J. Andrew Bagnell (Carnegie Mellon University, USA)

Th60 *Sampling from Non-Log-Concave Distributions via Variance-Reduced Gradient Langevin Dynamics*
Difan Zou (University of California, Los Angeles); Pan Xu (UCLA); Quanquan Gu (University of California, Los Angeles)*

Th61 *Graph to Graph: a Topology Aware Approach for Graph Structures Learning and Generation*
Mingming Sun (Baidu Research)*; Ping Li (Baidu Research)

Th62 *Imitation-Regularized Offline Learning*
Yifei Ma (Amazon)*; Yu-Xiang Wang (UC Santa Barbara); Balakrishnan Narayanaswamy (Amazon)

Th63 *A maximum-mean-discrepancy goodness-of-fit test for censored data*
Tamara Fernandez (UCL)*; Arthur Gretton (Gatsby Computational Neuroscience Unit)

Th64 *Learning the structure of a nonstationary vector autoregression*
Daniel Malinsky (Johns Hopkins University)*; Peter Spirtes (Carnegie Mellon University)

Th65 *A Fast Sampling Algorithm for Maximum Inner Product Search*
QIN DING (University of California, Davis)*; Cho-Jui Hsieh (UC Davis); Hsiang-Fu Yu ()

Th66 *Minimum Volume Topic Modeling*
Byoungwook Jang (University of Michigan)*; Alfred Hero (University of Michigan)

Th67 *Binary Space Partitioning Forests*
Xuhui Fan (UNSW)*; Bin Li (Fudan University); Scott SIsson ()

Th68 *Improved Graph based Semi-Supervised Learning*
Krishnamurthy Viswanathan (Google)*; Sushant Sachdeva (University of Toronto); Andrew Tomkins (Google); Sujith Ravi ()

Th69 *Optimizing over a Restricted Policy Class in MDPs*
Ershad Banijamali (UNIVERSITY OF WATERLOO)*; Yasin Abbasi-Yadkori (Adobe Research); Mohammad Ghavamzadeh (FAIR); Nikos Vlassis (Netflix)

Th70 *Stochastic Gradient Descent on Separable Data: Exact Convergence with a Fixed Learning Rate*
Mor Shpigel Nacson (Technion)*; Nathan Srebro (Toyota Technical Institute of Chicago); Daniel Soudry (Technion)

Th71 *Multilayer Switch Networks for Generating Discrete Data*
Payam Delgosha (UC Berkeley)*; Naveen Goela (Tanium Data Science)

Th72 *A recurrent Markov state-space generative model for sequences*
Anand Ramachandran (University of Illinois at Urbana-Champaign)*; Steve Lumetta (University of Illinois at Urbana-Champaign); Eric Klee (Mayo Clinic); Deming Chen (University of Illinois at Urbana-Champaign)

Th73 *A Potential Outcomes Calculus for Identifying Conditional Path-Specific Effects*
Daniel Malinsky (Johns Hopkins University)*; Ilya Shpitser (Johns Hopkins University); Thomas Richardson (University of Washington)

Th74 *Adversarial Discrete Sequence Generation without Explicit NeuralNetworks as Discriminators*
Zhongliang Li (Google)*; Tian Xia (Pingan Technology); Kaihe Xu (Pingan Technology); Xingyu Lou (Pingan Technology); Shaojun Wang (); Jing Xiao (Ping An Technology (Shenzhen) Co., Ltd)

Th75 *Adaptive Estimation for Approximate k-Nearest-Neighbor Computations*
Daniel LeJeune (Rice University)*; Reinhard Heckel (Rice University); richard baraniuk (Rice

University)

Th76 *Model-Free Control via Reduction to Expert Prediction*
Yasin Abbasi-Yadkori (Adobe Research)*; Nevena Lazic (Google); Csaba Szepesvari (DeepMind/University of Alberta)

Th77 *Learning Predictive Models That Transport*
Adarsh Subbaswamy (Johns Hopkins University)*; Peter Schulam (Johns Hopkins University); Suchi Saria (Johns Hopkins University)

Th78 *Structured Robust Submodular Maximization: Offline and Online Algorithms*
Alfredo Torrico (Georgia Tech)*; Nika Haghtalab (Microsoft); Nima Anari (Stanford); Seffi Naor (Technion); Sebastian Pokutta (Gatech); Mohit Singh (Georgia Tech)

Th79 *Sample-Effiient Imitation Learning via Generative Adversarial Nets*
Lionel Blondé (Hesso/UniGe)*

Th80 *Probabilistic Multilevel Clustering via Composite Transportation Distance*
Nhat Ho (University of California, Berkeley)*; Viet Huynh (Monash University); Dinh Phung (Monash University); Michael Jordan (UC Berkeley)

Th81 *A General Framework for Multi-fidelity Bayesian Optimization with Gaussian Processes*
Jialin Song (Caltech)*; Yuxin Chen (Caltech); Yisong Yue (Caltech)

Th82 *Online Algorithm for Unsupervised Sensor Selection*
Arun Verma (IIT Bombay); Manjesh Kumar Hanawal (IIT Bombay); Csaba Szepesvari (DeepMind/University of Alberta)*; Venkatesh Saligrama (Boston University)

Th83 *Best of many worlds: Robust model selection for online supervised learning*
Vidya Muthukumar (UC Berkeley)*; Mitas Ray (UC Berkeley); Anant Sahai (UC Berkeley); Peter Bartlett ("University of California, Berkeley")

Th84 *Accelerating Imitation Learning with Predictive Models*
Ching-An Cheng (Georgia Institute of Technology)*; Xinyan Yan (Georgia Institute of Technology); Evangelos Theodorou (Georgia Institute of Technology); Byron Boots (Georgia Institute of Technology)

Th85 *Online Learning in Kernelized Markov Decision Processes*
Sayak Ray Chowdhury (Indian Institute of Science)*; Aditya Gopalan (Indian Institute of Science (IISc), Bangalore)

Th86 *Lifting high-dimensional non-linear models with Gaussian regressors*
Christos Thrampoulidis (University of California, Santa Barbara)*; Ankit Singh Rawat (Google)

Th87 *Domain-size Aware Markov Logic Networks*
Happy Mittal (IIT delhi)*; Ayush Bhardwaj (IIT delhi); Vibhav G Gogate (The University of Texas at Dallas); Parag Singla (IIT Delhi)

Th88 *Database Alignment with Gaussian Features*
Osman E Dai (Georgia Institute of Technology)*; Daniel Cullina (Princeton University); Negar Kiyavash (University of Illinois at Urbana-Champaign)

Th89 *Size of Interventional Markov Equivalence Classes in random DAG models*
Dmitriy Katz (IBM Research); Karthikeyan Shanmugam (IBM Research)*; Chandler Squires (Massachusetts Institute of Technology); Caroline Uhler (MIT)

Th90 *Revisit Batch Normalization: New Understanding and Refinement via Composition Optimization*
Xiangru Lian (University of Rochester)*; Ji Liu (University of Rochester)

Th91 *Multi-Order Information for Working Set Selection of Sequential Minimal Optimization*
Qimao Yang (East China Normal University); Changrong Li (CFETS Information Technology); Jun Guo (East China Normal University)*

Th92 *Harmonizable mixture kernels with variational Fourier features*
Zheyang Shen (Aalto University)*; Markus Heinonen (Aalto University); Samuel Kaski (Aalto University)

Th93 *Multiscale Gaussian Process Level Set Estimation*
Shubhanshu Shekhar (University of California, San Diego)*; Tara Javidi (University of California San Diego)

Th94 *Bayesian Nonparametric Coalescent-Tree Priors for VAEs*
Sharad Vikram (UCSD)*; Matthew D Hoffman (Google); Matthew J Johnson (Google Brain)

Th95 *Adversarial Learning of a Sampler Based on an Unnormalized Distribution*
Chunyuan Li (Microsoft Research)*; Ke Bai (Duke University); Jianqiao Li (Duke University); Guoyin Wang (Duke University); Changyou Chen (University at Buffalo); Lawrence Carin Duke (CS)

Th96 *Recovery Guarantees For Quadratic Tensors With Sparse Observations*
Hongyang Zhang (Stanford University)*; Vatsal Sharan (Stanford University); Moses Charikar (Stanford University, California); Yingyu Liang (University of Wisconsin Madison)

Th97 *Sample Efficient Graph-Based Optimization with Noisy Observations*
Thanh Tan Nguyen (Queensland University of Technology)*; Ali Shameli (Stanford University); Yasin Abbasi-Yadkori (Adobe Research); Anup Rao (Adobe Research); Branislav Kveton (Google Research)

Th98 *Robustness Guarantees for Density Clustering*
Heinrich Jiang (Google)*; Jennifer Jang (Uber); Ofir Nachum (Google)

Th99 *Fixing Mini-batch Sequences with Hierarchical Robust Partitioning*
Shengjie Wang ("University of Washington, Seattle")*; Wenruo Bai ("University of Washington, Seattle"); Chandrashekhar Lavania (University Of Washington); Jeff Bilmes (UW)

Th100 *Efficient Bayes Risk Estimation for Cost-Sensitive Classification*
Daniel Andrade (NEC)*; Yuzuru Okajima (NEC)

Th101 *Interpreting Black Box Predictions using Fisher Kernels*
Rajiv Khanna (University of California at Berkeley)*; Been Kim (Google); Joydeep Ghosh (UT Austin); Sanmi Koyejo (University of Illinois, Urbana-Champaign)

Th102 *Representation Learning on Graphs: A Reinforcement Learning Application*
Sephora Madjiheurem (University College London)*; Laura Toni (UCL)

Th103 *ABC-Strategy: Budgeted Experimental Design for Targeted Causal Structure Discovery*
Raj Agrawal (MIT)*; Caroline Uhler (MIT); Chandler Squires (Massachusetts Institute of Technology); Karren D Yang (MIT); Karthikeyan Shanmugam (IBM Research)

Th104 *Batched Stochastic Bayesian Optimization via Combinatorial Constraints Design*
Kevin K Yang (Caltech)*; Yuxin Chen (Caltech); Yisong Yue (Caltech)

Th105 *Structured Representations for Reviews: Aspect-Based Variational Hidden Factor Models*
Babak Esmaeili (Northeastern University)*; Byron Wallace (Northeastern); Jan-Willem van de Meent (Northeastern)

Th106 *Adaptive Minimax Regret against Smooth Logarithmic Losses over High-Dimensional l1-Balls via Envelope Complexity*

Kohei Miyaguchi (The University of Tokyo)*; Kenji Yamanishi (The University of Tokyo)

Th107 *Low-Dimensional Density Ratio Estimation for Covariate Shift Correction*
Petar Stojanov (Carnegie Mellon University)*; Kun Zhang (Carnegie Mellon University); Mingming Gong (University of Pittsburgh); Jaime Carbonell (CMU, LTI)

Th108 *Evaluating model calibration in classification*
Juozas vaicenavicius (Uppsala University)*; David Widmann (Uppsala University); Carl Andersson (Uppsala University); Fredrik Lindsten (Uppsala University); Jacob Roll (Veoneer Inc.); Thomas Schön (Uppsala University)

Th109 *Towards Gradient Free and Projection Free Stochastic Optimization*
Anit Kumar Sahu (Carnegie Mellon University)*; Manzil Zaheer (Carnegie Mellon University); Soummya Kar ()

Th110 *Data-Driven Approach to Multiple-Source Domain Adaptation*
Petar Stojanov (Carnegie Mellon University)*; Kun Zhang (Carnegie Mellon University); Mingming Gong (University of Pittsburgh); Jaime Carbonell (CMU, LTI)

AISTATS

AIP

TOYOTA

SoftBank

TOSHIBA

OMRON