

- [A graph similarity for deep learning](#) *Seongmin Ok*
- [An Unsupervised Information-Theoretic Perceptual Quality Metric](#) *Sangnie Bhardwaj, Ian Fischer, Johannes Ballé, Troy Chinen*
- [Self-Supervised MultiModal Versatile Networks](#) *Jean-Baptiste Alayrac, Adria Recasens, Rosalia Schneider, Relja Arandjelović, Jason Ramapuram, Jeffrey De Fauw, Lucas Smaira, Sander Dieleman, Andrew Zisserman*
- [Benchmarking Deep Inverse Models over time, and the Neural-Adjoint method](#) *Simiao Ren, Willie Padilla, Jordan Malof*
- [Off-Policy Evaluation and Learning for External Validity under a Covariate Shift](#) *Masatoshi Uehara, Masahiro Kato, Shota Yasui*
- [Neural Methods for Point-wise Dependency Estimation](#) *Yao-Hung Hubert Tsai, Han Zhao, Makoto Yamada, Louis-Philippe Morency, Russ R. Salakhutdinov*
- [Fast and Flexible Temporal Point Processes with Triangular Maps](#) *Oleksandr Shchur, Nicholas Gao, Marin Biloš, Stephan Günnemann*
- [Backpropagating Linearly Improves Transferability of Adversarial Examples](#) *Yiwen Guo, Qizhang Li, Hao Chen*
- [PyGlove: Symbolic Programming for Automated Machine Learning](#) *Daiyi Peng, Xuanyi Dong, Esteban Real, Mingxing Tan, Yifeng Lu, Gabriel Bender, Hanxiao Liu, Adam Kraft, Chen Liang, Quoc Le*
- [Fourier Sparse Leverage Scores and Approximate Kernel Learning](#) *Tamas Erdelyi, Cameron Musco, Christopher Musco*
- [Improved Algorithms for Online Submodular Maximization via First-order Regret Bounds](#) *Nicholas Harvey, Christopher Liaw, Tasuku Soma*
- [Synbols: Probing Learning Algorithms with Synthetic Datasets](#) *Alexandre Lacoste, Pau Rodríguez López, Frederic Branchaud-Charron, Parmida Atighehchian, Massimo Caccia, Issam Hadj Laradji, Alexandre Drouin, Matthew Craddock, Laurent Charlin, David Vázquez*
- [Adversarially Robust Streaming Algorithms via Differential Privacy](#) *Avinatan Hasidim, Haim Kaplan, Yishay Mansour, Yossi Matias, Uri Stemmer*
- [Trading Personalization for Accuracy: Data Debugging in Collaborative Filtering](#) *Long Chen, Yuan Yao, Feng Xu, Miao Xu, Hanghang Tong*
- [Cascaded Text Generation with Markov Transformers](#) *Yuntian Deng, Alexander Rush*
- [Improving Local Identifiability in Probabilistic Box Embeddings](#) *Shib Dasgupta, Michael Boratko, Dongxu Zhang, Luke Vilnis, Xiang Li, Andrew McCallum*
- [Permute-and-Flip: A new mechanism for differentially private selection](#) *Ryan McKenna, Daniel R. Sheldon*
- [Deep reconstruction of strange attractors from time series](#) *William Gilpin*
- [Reciprocal Adversarial Learning via Characteristic Functions](#) *Shengxi Li, Zeyang Yu, Min Xiang, Danilo Mandic*
- [Statistical Guarantees of Distributed Nearest Neighbor Classification](#) *Jiexin Duan, Xingye Qiao, Guang Cheng*
- [Stein Self-Repulsive Dynamics: Benefits From Past Samples](#) *Mao Ye, Tongzheng Ren, Qiang Liu*
- [The Statistical Complexity of Early-Stopped Mirror Descent](#) *Tomas Vaskevicius, Varun Kanade, Patrick Rebeschini*
- [Algorithmic recourse under imperfect causal knowledge: a probabilistic approach](#) *Amir-Hossein Karimi, Julius von Kügelgen, Bernhard Schölkopf, Isabel Valera*
- [Quantitative Propagation of Chaos for SGD in Wide Neural Networks](#) *Valentin De Bortoli, Alain Durmus, Xavier Fontaine, Umut Simsekli*

- [A Causal View on Robustness of Neural Networks](#) *Cheng Zhang, Kun Zhang, Yingzhen Li*
- [Minimax Classification with 0-1 Loss and Performance Guarantees](#) *Santiago Mazuelas, Andrea Zanoni, Aritz Pérez*
- [How to Learn a Useful Critic? Model-based Action-Gradient-Estimator Policy Optimization](#) *Pierluca D'Oro, Wojciech Jaśkowski*
- [Coresets for Regressions with Panel Data](#) *Lingxiao Huang, K Sudhir, Nisheeth Vishnoi*
- [Learning Composable Energy Surrogates for PDE Order Reduction](#) *Alex Beatson, Jordan Ash, Geoffrey Roeder, Tianju Xue, Ryan P. Adams*
- [Efficient Contextual Bandits with Continuous Actions](#) *Maryam Majzoubi, Chicheng Zhang, Rajan Chari, Akshay Krishnamurthy, John Langford, Aleksandrs Slivkins*
- [Achieving Equalized Odds by Resampling Sensitive Attributes](#) *Yaniv Romano, Stephen Bates, Emmanuel Candes*
- [Multi-Robot Collision Avoidance under Uncertainty with Probabilistic Safety Barrier Certificates](#) *Wenhao Luo, Wen Sun, Ashish Kapoor*
- [Hard Shape-Constrained Kernel Machines](#) *Pierre-Cyril Aubin-Frankowski, Zoltan Szabo*
- [A Closer Look at the Training Strategy for Modern Meta-Learning](#) *JIAXIN CHEN, Xiao-Ming Wu, Yanke Li, Qimai LI, Li-Ming Zhan, Fu-lai Chung*
- [On the Value of Out-of-Distribution Testing: An Example of Goodhart's Law](#) *Damien Teney, Ehsan Abbasnejad, Kushal Kafle, Robik Shrestha, Christopher Kanan, Anton van den Hengel*
- [Generalised Bayesian Filtering via Sequential Monte Carlo](#) *Ayman Boustati, Omer Deniz Akyildiz, Theodoros Damoulas, Adam Johansen*
- [Deterministic Approximation for Submodular Maximization over a Matroid in Nearly Linear Time](#) *Kai Han, zongmai Cao, Shuang Cui, Benwei Wu*
- [Flows for simultaneous manifold learning and density estimation](#) *Johann Brehmer, Kyle Cranmer*
- [Simultaneous Preference and Metric Learning from Paired Comparisons](#) *Austin Xu, Mark Davenport*
- [Efficient Variational Inference for Sparse Deep Learning with Theoretical Guarantee](#) *Jincheng Bai, Qifan Song, Guang Cheng*
- [Learning Manifold Implicitly via Explicit Heat-Kernel Learning](#) *Yufan Zhou, Changyou Chen, Jinhui Xu*
- [Deep Relational Topic Modeling via Graph Poisson Gamma Belief Network](#) *Chaojie Wang, Hao Zhang, Bo Chen, Dongsheng Wang, Zhengjue Wang, Mingyuan Zhou*
- [One-bit Supervision for Image Classification](#) *Hengtong Hu, Lingxi Xie, Zewei Du, Richang Hong, Qi Tian*
- [What is being transferred in transfer learning?](#) *Behnam Neyshabur, Hanie Sedghi, Chiyuan Zhang*
- [Submodular Maximization Through Barrier Functions](#) *Ashwinkumar Badanidiyuru, Amin Karbasi, Ehsan Kazemi, Jan Vondrak*
- [Neural Networks with Recurrent Generative Feedback](#) *Yujia Huang, James Gornet, Sihui Dai, Zhiding Yu, Tan Nguyen, Doris Tsao, Anima Anandkumar*
- [Learning to Extrapolate Knowledge: Transductive Few-shot Out-of-Graph Link Prediction](#) *Jinheon Baek, Dong Bok Lee, Sung Ju Hwang*
- [Exploiting weakly supervised visual patterns to learn from partial annotations](#) *Kaustav Kundu, Joseph Tighe*
- [Improving Inference for Neural Image Compression](#) *Yibo Yang, Robert Bamler, Stephan Mandt*
- [Neuron Merging: Compensating for Pruned Neurons](#) *Woojeong Kim, Suhyun Kim, Mincheol Park, Geunseok Jeon*
- [FixMatch: Simplifying Semi-Supervised Learning with Consistency and Confidence](#) *Kihyuk Sohn, David Berthelot, Nicholas Carlini, Zizhao Zhang, Han Zhang, Colin A. Raffel, Ekin Dogus Cubuk, Alexey Kurakin,*

Chun-Liang Li

- [Reinforcement Learning with Combinatorial Actions: An Application to Vehicle Routing](#) Arthur Delarue, Ross Anderson, Christian Tjandraatmadja
- [Towards Playing Full MOBA Games with Deep Reinforcement Learning](#) Deheng Ye, Guibin Chen, Wen Zhang, Sheng Chen, Bo Yuan, Bo Liu, Jia Chen, Zhao Liu, Fuhao Qiu, Hongsheng Yu, Yinyuting Yin, Bei Shi, Liang Wang, Tengfei Shi, Qiang Fu, Wei Yang, Lanxiao Huang, Wei Liu
- [Rankmax: An Adaptive Projection Alternative to the Softmax Function](#) Weiwei Kong, Walid Krichene, Nicolas Mayoraz, Steffen Rendle, Li Zhang
- [Online Agnostic Boosting via Regret Minimization](#) Nataly Brukhim, Xinyi Chen, Elad Hazan, Shay Moran
- [Causal Intervention for Weakly-Supervised Semantic Segmentation](#) Dong Zhang, Hanwang Zhang, Jinhui Tang, Xian-Sheng Hua, Qianru Sun
- [Belief Propagation Neural Networks](#) Jonathan Kuck, Shuvam Chakraborty, Hao Tang, Rachel Luo, Jiaming Song, Ashish Sabharwal, Stefano Ermon
- [Over-parameterized Adversarial Training: An Analysis Overcoming the Curse of Dimensionality](#) Yi Zhang, Orestis Plevrakis, Simon S. Du, Xingguo Li, Zhao Song, Sanjeev Arora
- [Post-training Iterative Hierarchical Data Augmentation for Deep Networks](#) Adil Khan, Khadija Fraz
- [Debugging Tests for Model Explanations](#) Julius Adebayo, Michael Muelly, Ilaria Lliccardi, Been Kim
- [Robust compressed sensing using generative models](#) Ajil Jalal, Liu Liu, Alexandros G. Dimakis, Constantine Caramanis
- [Fairness without Demographics through Adversarially Reweighted Learning](#) Preethi Lahoti, Alex Beutel, Jilin Chen, Kang Lee, Flavien Prost, Nithum Thain, Xuezhi Wang, Ed Chi
- [Stochastic Latent Actor-Critic: Deep Reinforcement Learning with a Latent Variable Model](#) Alex X. Lee, Anusha Nagabandi, Pieter Abbeel, Sergey Levine
- [Ridge Rider: Finding Diverse Solutions by Following Eigenvectors of the Hessian](#) Jack Parker-Holder, Luke Metz, Cinjon Resnick, Hengyuan Hu, Adam Lerer, Alistair Letcher, Alexander Peysakhovich, Aldo Pacchiano, Jakob Foerster
- [The route to chaos in routing games: When is price of anarchy too optimistic?](#) Thiparat Chotibut, Fryderyk Falniowski, Michał Misiurewicz, Georgios Piliouras
- [Online Algorithm for Unsupervised Sequential Selection with Contextual Information](#) Arun Verma, Manjesh Kumar Hanawal, Csaba Szepesvari, Venkatesh Saligrama
- [Adapting Neural Architectures Between Domains](#) Yanxi Li, Zhaohui Yang, Yunhe Wang, Chang Xu
- [What went wrong and when? Instance-wise feature importance for time-series black-box models](#) Sana Tonekaboni, Shalmali Joshi, Kieran Campbell, David K. Duvenaud, Anna Goldenberg
- [Towards Better Generalization of Adaptive Gradient Methods](#) Yingxue Zhou, Belhal Karimi, Jinxing Yu, Zhiqiang Xu, Ping Li
- [Learning Guidance Rewards with Trajectory-space Smoothing](#) Tanmay Gangwani, Yuan Zhou, Jian Peng
- [Variance Reduction via Accelerated Dual Averaging for Finite-Sum Optimization](#) Chaobing Song, Yong Jiang, Yi Ma
- [Tree! I am no Tree! I am a low dimensional Hyperbolic Embedding](#) Rishi Sonthalia, Anna Gilbert
- [Deep Structural Causal Models for Tractable Counterfactual Inference](#) Nick Pawlowski, Daniel Coelho de Castro, Ben Glocker
- [Convolutional Generation of Textured 3D Meshes](#) Dario Pavllo, Graham Spinks, Thomas Hofmann, Marie-Francine Moens, Aurelien Lucchi
- [A Statistical Framework for Low-bitwidth Training of Deep Neural Networks](#) Jianfei Chen, Yu Gai, Zhewei

Yao, Michael W. Mahoney, Joseph E. Gonzalez

- Better Set Representations For Relational Reasoning *Qian Huang, Horace He, Abhay Singh, Yan Zhang, Ser Nam Lim, Austin R. Benson*
- AutoSync: Learning to Synchronize for Data-Parallel Distributed Deep Learning *Hao Zhang, Yuan Li, Zhijie Deng, Xiaodan Liang, Lawrence Carin, Eric Xing*
- A Combinatorial Perspective on Transfer Learning *Jianan Wang, Eren Sezener, David Budden, Marcus Hutter, Joel Veness*
- Hardness of Learning Neural Networks with Natural Weights *Amit Daniely, Gal Vardi*
- Higher-Order Spectral Clustering of Directed Graphs *Steinar Laenen, He Sun*
- Primal-Dual Mesh Convolutional Neural Networks *Francesco Milano, Antonio Loquercio, Antoni Rosinol, Davide Scaramuzza, Luca Carlone*
- The Advantage of Conditional Meta-Learning for Biased Regularization and Fine Tuning *Giulia Denevi, Massimiliano Pontil, Carlo Ciliberto*
- Watch out! Motion is Blurring the Vision of Your Deep Neural Networks *Qing Guo, Felix Juefei-Xu, Xiaofei Xie, Lei Ma, Jian Wang, Bing Yu, Wei Feng, Yang Liu*
- Sinkhorn Barycenter via Functional Gradient Descent *Zebang Shen, Zhenfu Wang, Alejandro Ribeiro, Hamed Hassani*
- Coresets for Near-Convex Functions *Murad Tukan, Alaa Maalouf, Dan Feldman*
- Bayesian Deep Ensembles via the Neural Tangent Kernel *Bobby He, Balaji Lakshminarayanan, Yee Whye Teh*
- Improved Schemes for Episodic Memory-based Lifelong Learning *Yunhui Guo, Mingrui Liu, Tianbao Yang, Tajana Rosing*
- Adaptive Sampling for Stochastic Risk-Averse Learning *Sebastian Curi, Kfir Y. Levy, Stefanie Jegelka, Andreas Krause*
- Deep Wiener Deconvolution: Wiener Meets Deep Learning for Image Deblurring *Jiangxin Dong, Stefan Roth, Bernt Schiele*
- Discovering Reinforcement Learning Algorithms *Junhyuk Oh, Matteo Hessel, Wojciech M. Czarnecki, Zhongwen Xu, Hado P. van Hasselt, Satinder Singh, David Silver*
- Taming Discrete Integration via the Boon of Dimensionality *Jeffrey Dudek, Dror Fried, Kuldeep S Meel*
- Blind Video Temporal Consistency via Deep Video Prior *Chenyang Lei, Yazhou Xing, Qifeng Chen*
- Simplify and Robustify Negative Sampling for Implicit Collaborative Filtering *Jingtao Ding, Yuhan Quan, Quanming Yao, Yong Li, Depeng Jin*
- Model Selection for Production System via Automated Online Experiments *Zhenwen Dai, Praveen Chandar, Ghazal Fazelnia, Benjamin Carterette, Mounia Lalmas*
- On the Almost Sure Convergence of Stochastic Gradient Descent in Non-Convex Problems *Panayotis Mertikopoulos, Nadav Hallak, Ali Kavis, Volkan Cevher*
- Automatic Perturbation Analysis for Scalable Certified Robustness and Beyond *Kaidi Xu, Zhouxing Shi, Huan Zhang, Yihan Wang, Kai-Wei Chang, Minlie Huang, Bhavya Kailkhura, Xue Lin, Cho-Jui Hsieh*
- Adaptation Properties Allow Identification of Optimized Neural Codes *Luke Rast, Jan Drugowitsch*
- Global Convergence and Variance Reduction for a Class of Nonconvex-Nonconcave Minimax Problems *Junchi Yang, Negar Kiyavash, Niao He*
- Model-Based Multi-Agent RL in Zero-Sum Markov Games with Near-Optimal Sample Complexity *Kaiqing Zhang, Sham Kakade, Tamer Basar, Lin Yang*
- Conservative Q-Learning for Offline Reinforcement Learning *Aviral Kumar, Aurick Zhou, George Tucker, Sergey Levine*

- [Online Influence Maximization under Linear Threshold Model](#) *Shuai Li, Fang Kong, Kejie Tang, Qizhi Li, Wei Chen*
- [Ensembling geophysical models with Bayesian Neural Networks](#) *Ushnish Sengupta, Matt Amos, Scott Hosking, Carl Edward Rasmussen, Matthew Juniper, Paul Young*
- [Delving into the Cyclic Mechanism in Semi-supervised Video Object Segmentation](#) *Yuxi Li, Ning Xu, Jinlong Peng, John See, Weiyao Lin*
- [Asymmetric Shapley values: incorporating causal knowledge into model-agnostic explainability](#) *Christopher Frye, Colin Rowat, Ilya Feige*
- [Understanding Deep Architecture with Reasoning Layer](#) *Xinshi Chen, Yufei Zhang, Christoph Reisinger, Le Song*
- [Planning in Markov Decision Processes with Gap-Dependent Sample Complexity](#) *Anders Jonsson, Emilie Kaufmann, Pierre Menard, Omar Darwiche Domingues, Edouard Leurent, Michal Valko*
- [Provably Good Batch Off-Policy Reinforcement Learning Without Great Exploration](#) *Yao Liu, Adith Swaminathan, Alekh Agarwal, Emma Brunskill*
- [Detection as Regression: Certified Object Detection with Median Smoothing](#) *Ping-yeh Chiang, Michael Curry, Ahmed Abdelkader, Aounon Kumar, John Dickerson, Tom Goldstein*
- [Contextual Reserve Price Optimization in Auctions via Mixed Integer Programming](#) *Joey Huchette, Haihao Lu, Hossein Esfandiari, Vahab Mirrokni*
- [ExpandNets: Linear Over-parameterization to Train Compact Convolutional Networks](#) *Shuxuan Guo, Jose M. Alvarez, Mathieu Salzmann*
- [FleXOR: Trainable Fractional Quantization](#) *Dongsoo Lee, Se Jung Kwon, Byeongwook Kim, Yongkweon Jeon, Baeseong Park, Jeongin Yun*
- [The Implications of Local Correlation on Learning Some Deep Functions](#) *Eran Malach, Shai Shalev-Shwartz*
- [Learning to search efficiently for causally near-optimal treatments](#) *Samuel Håkansson, Viktor Lindblom, Omer Gottesman, Fredrik D. Johansson*
- [A Game Theoretic Analysis of Additive Adversarial Attacks and Defenses](#) *Ambar Pal, Rene Vidal*
- [Posterior Network: Uncertainty Estimation without OOD Samples via Density-Based Pseudo-Counts](#) *Bertrand Charpentier, Daniel Zügner, Stephan Günnemann*
- [Recurrent Quantum Neural Networks](#) *Johannes Bausch*
- [No-Regret Learning and Mixed Nash Equilibria: They Do Not Mix](#) *Emmanouil-Vasileios Vlatakis-Gkaragkounis, Lampros Flokas, Thanasis Lianeas, Panayotis Mertikopoulos, Georgios Piliouras*
- [A Unifying View of Optimism in Episodic Reinforcement Learning](#) *Gergely Neu, Ciara Pike-Burke*
- [Continuous Submodular Maximization: Beyond DR-Submodularity](#) *Moran Feldman, Amin Karbasi*
- [An Asymptotically Optimal Primal-Dual Incremental Algorithm for Contextual Linear Bandits](#) *Andrea Tirinzoni, Matteo Pirotta, Marcello Restelli, Alessandro Lazaric*
- [Assessing SATNet's Ability to Solve the Symbol Grounding Problem](#) *Oscar Chang, Lampros Flokas, Hod Lipson, Michael Spranger*
- [A Bayesian Nonparametrics View into Deep Representations](#) *Michał Jamroz, Marcin Kurdziel, Mateusz Opala*
- [On the Similarity between the Laplace and Neural Tangent Kernels](#) *Amnon Geifman, Abhay Yadav, Yoni Kasten, Meirav Galun, David Jacobs, Basri Ronen*
- [A causal view of compositional zero-shot recognition](#) *Yuval Atzmon, Felix Kreuk, Uri Shalit, Gal Chechik*
- [HiPPO: Recurrent Memory with Optimal Polynomial Projections](#) *Albert Gu, Tri Dao, Stefano Ermon, Atri Rudra, Christopher Ré*
- [Auto Learning Attention](#) *Benteng Ma, Jing Zhang, Yong Xia, Dacheng Tao*

- [CASTLE: Regularization via Auxiliary Causal Graph Discovery](#) *Trent Kyono, Yao Zhang, Mihaela van der Schaar*
- [Long-Tailed Classification by Keeping the Good and Removing the Bad Momentum Causal Effect](#) *Kaihua Tang, Jianqiang Huang, Hanwang Zhang*
- [Explainable Voting](#) *Dominik Peters, Ariel D. Procaccia, Alexandros Psomas, Zixin Zhou*
- [Deep Archimedean Copulas](#) *Chun Kai Ling, Fei Fang, J. Zico Kolter*
- [Re-Examining Linear Embeddings for High-Dimensional Bayesian Optimization](#) *Ben Letham, Roberto Calandra, Akshara Rai, Eytan Bakshy*
- [UnModNet: Learning to Unwrap a Modulo Image for High Dynamic Range Imaging](#) *Chu Zhou, Hang Zhao, Jin Han, Chang Xu, Chao Xu, Tiejun Huang, Boxin Shi*
- [Thunder: a Fast Coordinate Selection Solver for Sparse Learning](#) *Shaogang Ren, Weijie Zhao, Ping Li*
- [Neural Networks Fail to Learn Periodic Functions and How to Fix It](#) *Liu Ziyin, Tilman Hartwig, Masahito Ueda*
- [Distribution Matching for Crowd Counting](#) *Boyu Wang, Huidong Liu, Dimitris Samaras, Minh Hoai Nguyen*
- [Correspondence learning via linearly-invariant embedding](#) *Riccardo Marin, Marie-Julie Rakotosaona, Simone Melzi, Maks Ovsjanikov*
- [Learning to Dispatch for Job Shop Scheduling via Deep Reinforcement Learning](#) *Cong Zhang, Wen Song, Zhiguang Cao, Jie Zhang, Puay Siew Tan, Xu Chi*
- [On Adaptive Attacks to Adversarial Example Defenses](#) *Florian Tramer, Nicholas Carlini, Wieland Brendel, Aleksander Madry*
- [Sinkhorn Natural Gradient for Generative Models](#) *Zebang Shen, Zhenfu Wang, Alejandro Ribeiro, Hamed Hassani*
- [Online Sinkhorn: Optimal Transport distances from sample streams](#) *Arthur Mensch, Gabriel Peyré*
- [Ultrahyperbolic Representation Learning](#) *Marc Law, Jos Stam*
- [Locally-Adaptive Nonparametric Online Learning](#) *Ilja Kuzborskij, Nicolò Cesa-Bianchi*
- [Compositional Generalization via Neural-Symbolic Stack Machines](#) *Xinyun Chen, Chen Liang, Adams Wei Yu, Dawn Song, Denny Zhou*
- [Graphon Neural Networks and the Transferability of Graph Neural Networks](#) *Luana Ruiz, Luiz Chamon, Alejandro Ribeiro*
- [Unreasonable Effectiveness of Greedy Algorithms in Multi-Armed Bandit with Many Arms](#) *Mohsen Bayati, Nima Hamidi, Ramesh Johari, Khashayar Khosravi*
- [Gamma-Models: Generative Temporal Difference Learning for Infinite-Horizon Prediction](#) *Michael Janner, Igor Mordatch, Sergey Levine*
- [Deep Transformers with Latent Depth](#) *Xian Li, Asa Cooper Stickland, Yuqing Tang, Xiang Kong*
- [Neural Mesh Flow: 3D Manifold Mesh Generation via Diffeomorphic Flows](#) *Kunal Gupta, Manmohan Chandraker*
- [Statistical control for spatio-temporal MEG/EEG source imaging with desparsified mutli-task Lasso](#) *Jerome-Alexis Chevalier, Joseph Salmon, Alexandre Gramfort, Bertrand Thirion*
- [A Scalable MIP-based Method for Learning Optimal Multivariate Decision Trees](#) *Haoran Zhu, Pavankumar Murali, Dzong Phan, Lam Nguyen, Jayant Kalagnanam*
- [Efficient Exact Verification of Binarized Neural Networks](#) *Kai Jia, Martin Rinard*
- [Ultra-Low Precision 4-bit Training of Deep Neural Networks](#) *Xiao Sun, Naigang Wang, Chia-Yu Chen, Jiamin Ni, Ankur Agrawal, Xiaodong Cui, Swagath Venkataramani, Kaoutar El Maghraoui, Vijayalakshmi (Viji) Srinivasan, Kailash Gopalakrishnan*
- [Bridging the Gap between Sample-based and One-shot Neural Architecture Search with BONAS](#) *Han Shi,*

Renjie Pi, Hang Xu, Zhenguo Li, James Kwok, Tong Zhang

- On Numerosity of Deep Neural Networks *Xi Zhang, Xiaolin Wu*
- Outlier Robust Mean Estimation with Subgaussian Rates via Stability *Ilias Diakonikolas, Daniel M. Kane, Ankit Pensia*
- Self-Supervised Relationship Probing *Jiuxiang Gu, Jason Kuen, Shafiq Joty, Jianfei Cai, Vlad Morariu, Handong Zhao, Tong Sun*
- Information Theoretic Counterfactual Learning from Missing-Not-At-Random Feedback *Zifeng Wang, Xi Chen, Rui Wen, Shao-Lun Huang, Ercan Kuruoglu, Yefeng Zheng*
- Prophet Attention: Predicting Attention with Future Attention *Fenglin Liu, Xuancheng Ren, Xian Wu, Shen Ge, Wei Fan, Yuexian Zou, Xu Sun*
- Language Models are Few-Shot Learners *Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel Ziegler, Jeffrey Wu, Clemens Winter, Chris Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, Dario Amodei*
- Margins are Insufficient for Explaining Gradient Boosting *Allan Grønlund, Lior Kamma, Kasper Green Larsen*
- Fourier-transform-based attribution priors improve the interpretability and stability of deep learning models for genomics *Alex Tseng, Avanti Shrikumar, Anshul Kundaje*
- MomentumRNN: Integrating Momentum into Recurrent Neural Networks *Tan Nguyen, Richard Baraniuk, Andrea Bertozzi, Stanley Osher, Bao Wang*
- Marginal Utility for Planning in Continuous or Large Discrete Action Spaces *Zaheen Ahmad, Levi Lelis, Michael Bowling*
- Projected Stein Variational Gradient Descent *Peng Chen, Omar Ghattas*
- Minimax Lower Bounds for Transfer Learning with Linear and One-hidden Layer Neural Networks *Mohammadreza Mousavi Kalan, Zalan Fabian, Salman Avestimehr, Mahdi Soltanolkotabi*
- SE(3)-Transformers: 3D Roto-Translation Equivariant Attention Networks *Fabian Fuchs, Daniel Worrall, Volker Fischer, Max Welling*
- On the equivalence of molecular graph convolution and molecular wave function with poor basis set *Masashi Tsubaki, Teruyasu Mizoguchi*
- The Power of Predictions in Online Control *Chenkai Yu, Guanya Shi, Soon-Jo Chung, Yisong Yue, Adam Wierman*
- Learning Affordance Landscapes for Interaction Exploration in 3D Environments *Tushar Nagarajan, Kristen Grauman*
- Cooperative Multi-player Bandit Optimization *Ilai Bistritz, Nicholas Bambos*
- Tight First- and Second-Order Regret Bounds for Adversarial Linear Bandits *Shinji Ito, Shuichi Hirahara, Tasuku Soma, Yuichi Yoshida*
- Just Pick a Sign: Optimizing Deep Multitask Models with Gradient Sign Dropout *Zhao Chen, Jiquan Ngiam, Yanping Huang, Thang Luong, Henrik Kretzschmar, Yuning Chai, Dragomir Anguelov*
- A Loss Function for Generative Neural Networks Based on Watson's Perceptual Model *Steffen Czolbe, Oswin Krause, Ingemar Cox, Christian Igel*
- Dynamic Fusion of Eye Movement Data and Verbal Narrations in Knowledge-rich Domains *Ervine Zheng, Qi Yu, Rui Li, Pengcheng Shi, Anne Haake*
- Scalable Multi-Agent Reinforcement Learning for Networked Systems with Average Reward *Guannan Qu, Yiheng Lin, Adam Wierman, Na Li*

- [Optimizing Neural Networks via Koopman Operator Theory](#) Akshunna S. Dogra, William Redman
- [SVGD as a kernelized Wasserstein gradient flow of the chi-squared divergence](#) Sinho Chewi, Thibaut Le Gouic, Chen Lu, Tyler Maunu, Philippe Rigollet
- [Adversarial Robustness of Supervised Sparse Coding](#) Jeremias Sulam, Ramchandran Muthukumar, Raman Arora
- [Differentiable Meta-Learning of Bandit Policies](#) Craig Boutilier, Chih-wei Hsu, Branislav Kveton, Martin Mladenov, Csaba Szepesvari, Manzil Zaheer
- [Biologically Inspired Mechanisms for Adversarial Robustness](#) Manish Reddy Vuyyuru, Andrzej Banburski, Nishka Pant, Tomaso Poggio
- [Statistical-Query Lower Bounds via Functional Gradients](#) Surbhi Goel, Aravind Gollakota, Adam Klivans
- [Near-Optimal Reinforcement Learning with Self-Play](#) Yu Bai, Chi Jin, Tiancheng Yu
- [Network Diffusions via Neural Mean-Field Dynamics](#) Shushan He, Hongyuan Zha, Xiaojing Ye
- [Self-Distillation as Instance-Specific Label Smoothing](#) Zhilu Zhang, Mert Sabuncu
- [Towards Problem-dependent Optimal Learning Rates](#) Yunbei Xu, Assaf Zeevi
- [Cross-lingual Retrieval for Iterative Self-Supervised Training](#) Chau Tran, Yuqing Tang, Xian Li, Jiatao Gu
- [Rethinking pooling in graph neural networks](#) Diego Mesquita, Amauri Souza, Samuel Kaski
- [Pointer Graph Networks](#) Petar Veličković, Lars Buesing, Matthew Overlan, Razvan Pascanu, Oriol Vinyals, Charles Blundell
- [Gradient Regularized V-Learning for Dynamic Treatment Regimes](#) Yao Zhang, Mihaela van der Schaar
- [Faster Wasserstein Distance Estimation with the Sinkhorn Divergence](#) Lénaïc Chizat, Pierre Roussillon, Flavien Léger, François-Xavier Vialard, Gabriel Peyré
- [Forethought and Hindsight in Credit Assignment](#) Veronica Chelu, Doina Precup, Hado P. van Hasselt
- [Robust Recursive Partitioning for Heterogeneous Treatment Effects with Uncertainty Quantification](#) Hyun-Suk Lee, Yao Zhang, William Zame, Cong Shen, Jang-Won Lee, Mihaela van der Schaar
- [Rescuing neural spike train models from bad MLE](#) Diego Arribas, Yuan Zhao, Il Memming Park
- [Lower Bounds and Optimal Algorithms for Personalized Federated Learning](#) Filip Hanzely, Slavomír Hanzely, Samuel Horváth, Peter Richtarik
- [Black-Box Certification with Randomized Smoothing: A Functional Optimization Based Framework](#) Dinghui Zhang, Mao Ye, Chengyue Gong, Zhanxing Zhu, Qiang Liu
- [Deep Imitation Learning for Bimanual Robotic Manipulation](#) Fan Xie, Alexander Chowdhury, M. Clara De Paolis Kaluza, Linfeng Zhao, Lawson Wong, Rose Yu
- [Stationary Activations for Uncertainty Calibration in Deep Learning](#) Lassi Meronen, Christabella Irwanto, Arno Solin
- [Ensemble Distillation for Robust Model Fusion in Federated Learning](#) Tao Lin, Lingjing Kong, Sebastian U. Stich, Martin Jaggi
- [Falcon: Fast Spectral Inference on Encrypted Data](#) Qian Lou, Wen-jie Lu, Cheng Hong, Lei Jiang
- [On Power Laws in Deep Ensembles](#) Ekaterina Lobacheva, Nadezhda Chirkova, Maxim Kodryan, Dmitry P. Vetrov
- [Practical Quasi-Newton Methods for Training Deep Neural Networks](#) Donald Goldfarb, Yi Ren, Achraf Bahamou
- [Approximation Based Variance Reduction for Reparameterization Gradients](#) Tomas Geffner, Justin Domke
- [Inference Stage Optimization for Cross-scenario 3D Human Pose Estimation](#) Jianfeng Zhang, Xuecheng Nie, Jiashi Feng
- [Consistent feature selection for analytic deep neural networks](#) Vu C. Dinh, Lam S. Ho

- Glance and Focus: a Dynamic Approach to Reducing Spatial Redundancy in Image Classification Yulin Wang, Kangchen Lv, Rui Huang, Shiji Song, Le Yang, Gao Huang
- Information Maximization for Few-Shot Learning Malik Boudiaf, Imtiaz Ziko, Jérôme Rony, Jose Dolz, Pablo Piantanida, Ismail Ben Ayed
- Inverse Reinforcement Learning from a Gradient-based Learner Giorgia Ramponi, Gianluca Drappo, Marcello Restelli
- Bayesian Multi-type Mean Field Multi-agent Imitation Learning Fan Yang, Alina Vereshchaka, Changyou Chen, Wen Dong
- Bayesian Robust Optimization for Imitation Learning Daniel Brown, Scott Niekum, Marek Petrik
- Multiview Neural Surface Reconstruction by Disentangling Geometry and Appearance Lior Yariv, Yoni Kasten, Dror Moran, Meirav Galun, Matan Atzmon, Basri Ronen, Yaron Lipman
- Riemannian Continuous Normalizing Flows Emile Mathieu, Maximilian Nickel
- Attention-Gated Brain Propagation: How the brain can implement reward-based error backpropagation Isabella Pozzi, Sander Bohte, Pieter Roelfsema
- Asymptotic Guarantees for Generative Modeling Based on the Smooth Wasserstein Distance Ziv Goldfeld, Kristjan Greenewald, Kengo Kato
- Online Robust Regression via SGD on the l_1 loss Scott Pesme, Nicolas Flammarion
- PRANK: motion Prediction based on RANKing Yuriy Biktairov, Maxim Stebelev, Irina Rudenko, Oleh Shliazhko, Boris Yangel
- Fighting Copycat Agents in Behavioral Cloning from Observation Histories Chuan Wen, Jierui Lin, Trevor Darrell, Dinesh Jayaraman, Yang Gao
- Tight Nonparametric Convergence Rates for Stochastic Gradient Descent under the Noiseless Linear Model Raphaël Berthier, Francis Bach, Pierre Gaillard
- Structured Prediction for Conditional Meta-Learning Ruohan Wang, Yiannis Demiris, Carlo Ciliberto
- Optimal Lottery Tickets via Subset Sum: Logarithmic Over-Parameterization is Sufficient Ankit Pensia, Shashank Rajput, Alliot Nagle, Harit Vishwakarma, Dimitris Papailiopoulos
- The Hateful Memes Challenge: Detecting Hate Speech in Multimodal Memes Douwe Kiela, Hamed Firooz, Aravind Mohan, Vedanuj Goswami, Amanpreet Singh, Pratik Ringshia, Davide Testuggine
- Stochasticity of Deterministic Gradient Descent: Large Learning Rate for Multiscale Objective Function Lingkai Kong, Molei Tao
- Identifying Learning Rules From Neural Network Observables Aran Navebi, Sanjana Srivastava, Surya Ganguli, Daniel L. Yamins
- Optimal Approximation - Smoothness Tradeoffs for Soft-Max Functions Alessandro Epasto, Mohammad Mahdian, Vahab Mirrokni, Emmanouil Zampetakis
- Weakly-Supervised Reinforcement Learning for Controllable Behavior Lisa Lee, Ben Eysenbach, Russ R. Salakhutdinov, Shixiang (Shane) Gu, Chelsea Finn
- Improving Policy-Constrained Kidney Exchange via Pre-Screening Duncan McElfresh, Michael Curry, Tuomas Sandholm, John Dickerson
- Learning abstract structure for drawing by efficient motor program induction Lucas Tian, Kevin Ellis, Marta Kryven, Josh Tenenbaum
- Why Do Deep Residual Networks Generalize Better than Deep Feedforward Networks? --- A Neural Tangent Kernel Perspective Kaixuan Huang, Yuqing Wang, Molei Tao, Tuo Zhao
- Dual Instrumental Variable Regression Krikamol Muandet, Arash Mehrjou, Si Kai Lee, Anant Raj
- Stochastic Gradient Descent in Correlated Settings: A Study on Gaussian Processes Hao Chen, Lili Zheng,

Raed AL Kontar, Garvesh Raskutti

- Interventional Few-Shot Learning *Zhongqi Yue, Hanwang Zhang, Qianru Sun, Xian-Sheng Hua*
- Minimax Value Interval for Off-Policy Evaluation and Policy Optimization *Nan Jiang, Jiawei Huang*
- Biased Stochastic First-Order Methods for Conditional Stochastic Optimization and Applications in Meta Learning *Yifan Hu, Siqi Zhang, Xin Chen, Niao He*
- ShiftAddNet: A Hardware-Inspired Deep Network *Haoran You, Xiaohan Chen, Yongan Zhang, Chaojian Li, Sicheng Li, Zihao Liu, Zhangyang Wang, Yingyan Lin*
- Network-to-Network Translation with Conditional Invertible Neural Networks *Robin Rombach, Patrick Esser, Bjorn Ommer*
- Intra-Processing Methods for Debiasing Neural Networks *Yash Savani, Colin White, Naveen Sundar Govindarajulu*
- Finding Second-Order Stationary Points Efficiently in Smooth Nonconvex Linearly Constrained Optimization Problems *Songtao Lu, Meisam Razaviyayn, Bo Yang, Kejun Huang, Mingyi Hong*
- Model-based Policy Optimization with Unsupervised Model Adaptation *Jian Shen, Han Zhao, Weinan Zhang, Yong Yu*
- Implicit Regularization and Convergence for Weight Normalization *Xiaoxia Wu, Edgar Dobriban, Tongzheng Ren, Shanshan Wu, Zhiyuan Li, Suriya Gunasekar, Rachel Ward, Qiang Liu*
- Geometric All-way Boolean Tensor Decomposition *Changlin Wan, Wennan Chang, Tong Zhao, Sha Cao, Chi Zhang*
- Modular Meta-Learning with Shrinkage *Yutian Chen, Abram L. Friesen, Feryal Behbahani, Arnaud Doucet, David Budden, Matthew Hoffman, Nando de Freitas*
- A/B Testing in Dense Large-Scale Networks: Design and Inference *Preetam Nandy, Kinjal Basu, Shaunak Chatterjee, Ye Tu*
- What Neural Networks Memorize and Why: Discovering the Long Tail via Influence Estimation *Vitaly Feldman, Chiyuan Zhang*
- Partially View-aligned Clustering *Zhenyu Huang, Peng Hu, Joey Tianyi Zhou, Jiancheng Lv, Xi Peng*
- Partial Optimal Transport with applications on Positive-Unlabeled Learning *Laetitia Chapel, Mokhtar Z. Alaya, Gilles Gasso*
- Toward the Fundamental Limits of Imitation Learning *Nived Rajaraman, Lin Yang, Jiantao Jiao, Kannan Ramchandran*
- Logarithmic Pruning is All You Need *Laurent Orseau, Marcus Hutter, Omar Rivasplata*
- Hold me tight! Influence of discriminative features on deep network boundaries *Guillermo Ortiz-Jimenez, Apostolos Modas, Seyed-Mohsen Moosavi, Pascal Frossard*
- Learning from Mixtures of Private and Public Populations *Raef Bassily, Shay Moran, Anupama Nandi*
- Adversarial Weight Perturbation Helps Robust Generalization *Dongxian Wu, Shu-Tao Xia, Yisen Wang*
- Stateful Posted Pricing with Vanishing Regret via Dynamic Deterministic Markov Decision Processes *Yuval Emek, Ron Lavi, Rad Niazadeh, Yangguang Shi*
- Adversarial Self-Supervised Contrastive Learning *Minseon Kim, Jihoon Tack, Sung Ju Hwang*
- Normalizing Kalman Filters for Multivariate Time Series Analysis *Emmanuel de Bézenac, Syama Sundar Rangapuram, Konstantinos Benidis, Michael Bohlke-Schneider, Richard Kurle, Lorenzo Stella, Hilaf Hasson, Patrick Gallinari, Tim Januschowski*
- Learning to summarize with human feedback *Nisan Stiennon, Long Ouyang, Jeffrey Wu, Daniel Ziegler, Ryan Lowe, Chelsea Voss, Alec Radford, Dario Amodei, Paul F. Christiano*
- Fourier Spectrum Discrepancies in Deep Network Generated Images *Tarik Dzanic, Karan Shah, Freddie*

Witherden

- [Lamina-specific neuronal properties promote robust, stable signal propagation in feedforward networks](#) Dongqi Han, Erik De Schutter, Sungho Hong
- [Learning Dynamic Belief Graphs to Generalize on Text-Based Games](#) Ashutosh Adhikari, Xingdi Yuan, Marc-Alexandre Côté, Mikuláš Zelinka, Marc-Antoine Rondeau, Romain Laroche, Pascal Poupart, Jian Tang, Adam Trischler, Will Hamilton
- [Triple descent and the two kinds of overfitting: where & why do they appear?](#) Stéphane d'Ascoli, Levent Sagun, Giulio Biroli
- [Multimodal Graph Networks for Compositional Generalization in Visual Question Answering](#) Raeid Saqr, Karthik Narasimhan
- [Learning Graph Structure With A Finite-State Automaton Layer](#) Daniel Johnson, Hugo Larochelle, Daniel Tarlow
- [A Universal Approximation Theorem of Deep Neural Networks for Expressing Probability Distributions](#) Yulong Lu, Jianfeng Lu
- [Unsupervised object-centric video generation and decomposition in 3D](#) Paul Henderson, Christoph H. Lampert
- [Domain Generalization for Medical Imaging Classification with Linear-Dependency Regularization](#) Haoliang Li, Yufei Wang, Renjie Wan, Shiqi Wang, Tie-Qiang Li, Alex Kot
- [Multi-label classification: do Hamming loss and subset accuracy really conflict with each other?](#) Guoqiang Wu, Jun Zhu
- [A Novel Automated Curriculum Strategy to Solve Hard Sokoban Planning Instances](#) Dieqiao Feng, Carla P. Gomes, Bart Selman
- [Causal analysis of Covid-19 Spread in Germany](#) Atalanti Mastakouri, Bernhard Schölkopf
- [Locally private non-asymptotic testing of discrete distributions is faster using interactive mechanisms](#) Thomas Berrett, Cristina Butucea
- [Adaptive Gradient Quantization for Data-Parallel SGD](#) Fartash Faghri, Iman Tabrizian, Ilia Markov, Dan Alistarh, Daniel M. Roy, Ali Ramezani-Kebrya
- [Finite Continuum-Armed Bandits](#) Solenne Gaucher
- [Removing Bias in Multi-modal Classifiers: Regularization by Maximizing Functional Entropies](#) Itai Gat, Idan Schwartz, Alexander Schwing, Tamir Hazan
- [Compact task representations as a normative model for higher-order brain activity](#) Severin Berger, Christian K. Machens
- [Robust-Adaptive Control of Linear Systems: beyond Quadratic Costs](#) Edouard Leurent, Odalric-Ambrym Maillard, Denis Efimov
- [Co-exposure Maximization in Online Social Networks](#) Sijing Tu, Cigdem Aslay, Aristides Gionis
- [UCLID-Net: Single View Reconstruction in Object Space](#) Benoit Guillard, Edoardo Remelli, Pascal Fua
- [Reinforcement Learning for Control with Multiple Frequencies](#) Jongmin Lee, Byung-Jun Lee, Kee-Eung Kim
- [Complex Dynamics in Simple Neural Networks: Understanding Gradient Flow in Phase Retrieval](#) Stefano Sarao Mannelli, Giulio Biroli, Chiara Cammarota, Florent Krzakala, Pierfrancesco Urbani, Lenka Zdeborová
- [Neural Message Passing for Multi-Relational Ordered and Recursive Hypergraphs](#) Naganand Yadati
- [A Unified View of Label Shift Estimation](#) Saurabh Garg, Yifan Wu, Sivaraman Balakrishnan, Zachary Lipton
- [Optimal Private Median Estimation under Minimal Distributional Assumptions](#) Christos Tzamos, Emmanouil-Vasileios Vlatakis-Gkaragkounis, Ilias Zadik
- [Breaking the Communication-Privacy-Accuracy Trilemma](#) Wei-Ning Chen, Peter Kairouz, Ayfer Ozgur

- [Audeo: Audio Generation for a Silent Performance Video](#) Kun Su, Xiulong Liu, Eli Shlizerman
- [Ode to an ODE](#) Krzysztof M. Choromanski, Jared Quincy Davis, Valerii Likhoshesterov, Xingyou Song, Jean-Jacques Slotine, Jacob Varley, Honglak Lee, Adrian Weller, Vikas Sindhwani
- [Self-Distillation Amplifies Regularization in Hilbert Space](#) Hossein Mobahi, Mehrdad Farajtabar, Peter Bartlett
- [Coupling-based Invertible Neural Networks Are Universal Diffeomorphism Approximators](#) Takeshi Teshima, Isao Ishikawa, Koichi Tojo, Kenta Oono, Masahiro Ikeda, Masashi Sugiyama
- [Community detection using fast low-cardinality semidefinite programming](#) Po-Wei Wang, J. Zico Kolter
- [Modeling Noisy Annotations for Crowd Counting](#) Jia Wan, Antoni Chan
- [An operator view of policy gradient methods](#) Dibya Ghosh, Marlos C. Machado, Nicolas Le Roux
- [Demystifying Contrastive Self-Supervised Learning: Invariances, Augmentations and Dataset Biases](#) Senthil Purushwalkam, Abhinav Gupta
- [Online MAP Inference of Determinantal Point Processes](#) Aditya Bhaskara, Amin Karbasi, Silvio Lattanzi, Morteza Zadimoghaddam
- [Video Object Segmentation with Adaptive Feature Bank and Uncertain-Region Refinement](#) Yongqing Liang, Xin Li, Navid Jafari, Jim Chen
- [Inferring learning rules from animal decision-making](#) Zoe Ashwood, Nicholas A. Roy, Ji Hyun Bak, Jonathan W. Pillow
- [Input-Aware Dynamic Backdoor Attack](#) Tuan Anh Nguyen, Anh Tran
- [How hard is to distinguish graphs with graph neural networks?](#) Andreas Loukas
- [Minimax Regret of Switching-Constrained Online Convex Optimization: No Phase Transition](#) Lin Chen, Qian Yu, Hannah Lawrence, Amin Karbasi
- [Dual Manifold Adversarial Robustness: Defense against \$L_p\$ and non- \$L_p\$ Adversarial Attacks](#) Wei-An Lin, Chun Pong Lau, Alexander Levine, Rama Chellappa, Soheil Feizi
- [Cross-Scale Internal Graph Neural Network for Image Super-Resolution](#) Shangchen Zhou, Jiawei Zhang, Wangmeng Zuo, Chen Change Loy
- [Unsupervised Representation Learning by Invariance Propagation](#) Feng Wang, Huaping Liu, Di Guo, Sun Fuchun
- [Restoring Negative Information in Few-Shot Object Detection](#) Yukuan Yang, Fangyun Wei, Miaoqing Shi, Guoqi Li
- [Do Adversarially Robust ImageNet Models Transfer Better?](#) Hadi Salman, Andrew Ilyas, Logan Engstrom, Ashish Kapoor, Aleksander Madry
- [Robust Correction of Sampling Bias using Cumulative Distribution Functions](#) Bijan Mazaheri, Siddharth Jain, Jehoshua Bruck
- [Personalized Federated Learning with Theoretical Guarantees: A Model-Agnostic Meta-Learning Approach](#) Alireza Fallah, Aryan Mokhtari, Asuman Ozdaglar
- [Pixel-Level Cycle Association: A New Perspective for Domain Adaptive Semantic Segmentation](#) Guoliang Kang, Yunchao Wei, Yi Yang, Yueting Zhuang, Alexander Hauptmann
- [Classification with Valid and Adaptive Coverage](#) Yaniv Romano, Matteo Sesia, Emmanuel Candes
- [Learning Global Transparent Models consistent with Local Contrastive Explanations](#) Tejaswini Pedapati, Avinash Balakrishnan, Karthikeyan Shanmugam, Amit Dhurandhar
- [Learning to Approximate a Bregman Divergence](#) Ali Siahkamari, XIDE XIA, Venkatesh Saligrama, David Castañón, Brian Kulis
- [Diverse Image Captioning with Context-Object Split Latent Spaces](#) Shweta Mahajan, Stefan Roth
- [Learning Disentangled Representations of Videos with Missing Data](#) Armand Comas, Chi Zhang, Zlatan

Feric, Octavia Camps, Rose Yu

- Natural Graph Networks *Pim de Haan, Taco S. Cohen, Max Welling*
- Continual Learning with Node-Importance based Adaptive Group Sparse Regularization *Sangwon Jung, Hongjoon Ahn, Sungmin Cha, Taesup Moon*
- Towards Crowdsourced Training of Large Neural Networks using Decentralized Mixture-of-Experts *Max Ryabinin, Anton Gusev*
- Bidirectional Convolutional Poisson Gamma Dynamical Systems *wenchao chen, Chaojie Wang, Bo Chen, Yicheng Liu, Hao Zhang, Mingyuan Zhou*
- Deep Reinforcement and InfoMax Learning *Bogdan Mazouze, Remi Tachet des Combes, Thang Long Doan, Philip Bachman, R Devon Hjelm*
- On ranking via sorting by estimated expected utility *Clement Calauzenes, Nicolas Usunier*
- Distribution-free binary classification: prediction sets, confidence intervals and calibration *Chirag Gupta, Aleksandr Podkopaev, Aaditya Ramdas*
- Closing the Dequantization Gap: PixelCNN as a Single-Layer Flow *Didrik Nielsen, Ole Winther*
- Sequence to Multi-Sequence Learning via Conditional Chain Mapping for Mixture Signals *Jing Shi, Xuankai Chang, Pengcheng Guo, Shinji Watanabe, Yusuke Fujita, Jiaming Xu, Bo Xu, Lei Xie*
- Variance reduction for Random Coordinate Descent-Langevin Monte Carlo *ZHIYAN DING, Qin Li*
- Language as a Cognitive Tool to Imagine Goals in Curiosity Driven Exploration *Cédric Colas, Tristan Karch, Nicolas Lair, Jean-Michel Dussoux, Clément Moulin-Frier, Peter Dominey, Pierre-Yves Oudeyer*
- All Word Embeddings from One Embedding *Sho Takase, Sosuke Kobayashi*
- Primal Dual Interpretation of the Proximal Stochastic Gradient Langevin Algorithm *Adil Salim, Peter Richtarik*
- How to Characterize The Landscape of Overparameterized Convolutional Neural Networks *Yihong Gu, Weizhong Zhang, Cong Fang, Jason D. Lee, Tong Zhang*
- On the Tightness of Semidefinite Relaxations for Certifying Robustness to Adversarial Examples *Richard Zhang*
- Submodular Meta-Learning *Arman Adibi, Aryan Mokhtari, Hamed Hassani*
- Rethinking Pre-training and Self-training *Barret Zoph, Golnaz Ghiasi, Tsung-Yi Lin, Yin Cui, Hanxiao Liu, Ekin Dogus Cubuk, Quoc Le*
- Unsupervised Sound Separation Using Mixture Invariant Training *Scott Wisdom, Efthymios Tzinis, Hakan Erdogan, Ron Weiss, Kevin Wilson, John Hershey*
- Adaptive Discretization for Model-Based Reinforcement Learning *Sean Sinclair, Tianyu Wang, Gauri Jain, Siddhartha Banerjee, Christina Yu*
- CodeCMR: Cross-Modal Retrieval For Function-Level Binary Source Code Matching *Zeping Yu, Wenxin Zheng, Jiaqi Wang, Qiyi Tang, Sen Nie, Shi Wu*
- On Warm-Starting Neural Network Training *Jordan Ash, Ryan P. Adams*
- DAGs with No Fears: A Closer Look at Continuous Optimization for Learning Bayesian Networks *Dennis Wei, Tian Gao, Yue Yu*
- OOD-MAML: Meta-Learning for Few-Shot Out-of-Distribution Detection and Classification *Taewon Jeong, Heeyoung Kim*
- An Imitation from Observation Approach to Transfer Learning with Dynamics Mismatch *Siddharth Desai, Ishan Durugkar, Haresh Karnan, Garrett Warnell, Josiah Hanna, Peter Stone*
- Learning About Objects by Learning to Interact with Them *Martin Lohmann, Jordi Salvador, Aniruddha Kembhavi, Roozbeh Mottaghi*
- Learning discrete distributions with infinite support *Doron Cohen, Aryeh Kontorovich, Geoffrey Wolfer*

- [Dissecting Neural ODEs](#) *Stefano Massaroli, Michael Poli, Jinkyoo Park, Atsushi Yamashita, Hajime Asama*
- [Teaching a GAN What Not to Learn](#) *Siddarth Asokan, Chandra Seelamantula*
- [Counterfactual Data Augmentation using Locally Factored Dynamics](#) *Silviu Pitis, Elliot Creager, Animesh Garg*
- [Rethinking Learnable Tree Filter for Generic Feature Transform](#) *Lin Song, Yanwei Li, Zhengkai Jiang, Zeming Li, Xiangyu Zhang, Hongbin Sun, Jian Sun, Nanning Zheng*
- [Self-Supervised Relational Reasoning for Representation Learning](#) *Massimiliano Patacchiola, Amos J. Storkey*
- [Sufficient dimension reduction for classification using principal optimal transport direction](#) *Cheng Meng, Jun Yu, Jingyi Zhang, Ping Ma, Wenxuan Zhong*
- [Fast Epigraphical Projection-based Incremental Algorithms for Wasserstein Distributionally Robust Support Vector Machine](#) *Jiajin Li, Caihua Chen, Anthony Man-Cho So*
- [Differentially Private Clustering: Tight Approximation Ratios](#) *Badi Ghazi, Ravi Kumar, Pasin Manurangsi*
- [On the Power of Louvain in the Stochastic Block Model](#) *Vincent Cohen-Addad, Adrian Kosowski, Frederik Mallmann-Trenn, David Saulpic*
- [Fairness with Overlapping Groups; a Probabilistic Perspective](#) *Forest Yang, Mouhamadou Cisse, Sanmi Koyejo*
- [AttendLight: Universal Attention-Based Reinforcement Learning Model for Traffic Signal Control](#) *Afshin Oroojlooy, Mohammadreza Nazari, Davood Hajinezhad, Jorge Silva*
- [Searching for Low-Bit Weights in Quantized Neural Networks](#) *Zhaohui Yang, Yunhe Wang, Kai Han, Chunjing XU, Chao Xu, Dacheng Tao, Chang Xu*
- [Adaptive Reduced Rank Regression](#) *Qiong Wu, Felix MF Wong, Yanhua Li, Zhenming Liu, Varun Kanade*
- [From Predictions to Decisions: Using Lookahead Regularization](#) *Nir Rosenfeld, Anna Hilgard, Sai Srivatsa Ravindranath, David C. Parkes*
- [Sequential Bayesian Experimental Design with Variable Cost Structure](#) *Sue Zheng, David Hayden, Jason Pacheco, John W. Fisher III*
- [Predictive inference is free with the jackknife+-after-bootstrap](#) *Byol Kim, Chen Xu, Rina Barber*
- [Counterfactual Predictions under Runtime Confounding](#) *Amanda Coston, Edward Kennedy, Alexandra Chouldechova*
- [Learning Loss for Test-Time Augmentation](#) *Ildoo Kim, Younghoon Kim, Sungwoong Kim*
- [Balanced Meta-Softmax for Long-Tailed Visual Recognition](#) *Jiawei Ren, Cunjun Yu, shunan sheng, Xiao Ma, Haiyu Zhao, Shuai Yi, hongsheng Li*
- [Efficient Exploration of Reward Functions in Inverse Reinforcement Learning via Bayesian Optimization](#) *Sreejith Balakrishnan, Quoc Phong Nguyen, Bryan Kian Hsiang Low, Harold Soh*
- [MDP Homomorphic Networks: Group Symmetries in Reinforcement Learning](#) *Elise van der Pol, Daniel Worrall, Herke van Hoof, Frans Oliehoek, Max Welling*
- [How Can I Explain This to You? An Empirical Study of Deep Neural Network Explanation Methods](#) *Jeya Vikranth Jeyakumar, Joseph Noor, Yu-Hsi Cheng, Luis Garcia, Mani Srivastava*
- [On the Error Resistance of Hinge-Loss Minimization](#) *Kunal Talwar*
- [Munchausen Reinforcement Learning](#) *Nino Vieillard, Olivier Pietquin, Matthieu Geist*
- [Object Goal Navigation using Goal-Oriented Semantic Exploration](#) *Devendra Singh Chaplot, Dhiraj Prakashchand Gandhi, Abhinav Gupta, Russ R. Salakhutdinov*
- [Efficient semidefinite-programming-based inference for binary and multi-class MRFs](#) *Chirag Pabbaraju, Po-Wei Wang, J. Zico Kolter*

- [Funnel-Transformer: Filtering out Sequential Redundancy for Efficient Language Processing](#) *Zihang Dai, Guokun Lai, Yiming Yang, Quoc Le*
- [Semantic Visual Navigation by Watching YouTube Videos](#) *Matthew Chang, Arjun Gupta, Saurabh Gupta*
- [Heavy-tailed Representations, Text Polarity Classification & Data Augmentation](#) *Hamid Jalalzai, Pierre Colombo, Chloé Clavel, Eric Gaussier, Giovanna Varni, Emmanuel Vignon, Anne Sabourin*
- [SuperLoss: A Generic Loss for Robust Curriculum Learning](#) *Thibault Castells, Philippe Weinzaepfel, Jerome Revaud*
- [CogMol: Target-Specific and Selective Drug Design for COVID-19 Using Deep Generative Models](#) *Vijil Chenthamarakshan, Payel Das, Samuel Hoffman, Hendrik Strobelt, Inkit Padhi, Kar Wai Lim, Benjamin Hoover, Matteo Manica, Jannis Born, Teodoro Laino, Aleksandra Mojsilovic*
- [Memory Based Trajectory-conditioned Policies for Learning from Sparse Rewards](#) *Yijie Guo, Jongwook Choi, Marcin Moczulski, Shengyu Feng, Samy Bengio, Mohammad Norouzi, Honglak Lee*
- [Liberty or Depth: Deep Bayesian Neural Nets Do Not Need Complex Weight Posterior Approximations](#) *Sebastian Farquhar, Lewis Smith, Yarin Gal*
- [Improving Sample Complexity Bounds for \(Natural\) Actor-Critic Algorithms](#) *Tengyu Xu, Zhe Wang, Yingbin Liang*
- [Learning Differential Equations that are Easy to Solve](#) *Jacob Kelly, Jesse Bettencourt, Matthew J. Johnson, David K. Duvenaud*
- [Stability of Stochastic Gradient Descent on Nonsmooth Convex Losses](#) *Raef Bassily, Vitaly Feldman, Cristóbal Guzmán, Kunal Talwar*
- [Influence-Augmented Online Planning for Complex Environments](#) *Jinke He, Miguel Suau de Castro, Frans Oliehoek*
- [PAC-Bayes Learning Bounds for Sample-Dependent Priors](#) *Pranjal Awasthi, Satyen Kale, Stefani Karp, Mehryar Mohri*
- [Reward-rational \(implicit\) choice: A unifying formalism for reward learning](#) *Hong Jun Jeon, Smitha Milli, Anca Dragan*
- [Probabilistic Time Series Forecasting with Shape and Temporal Diversity](#) *Vincent LE GUEN, Nicolas THOME*
- [Low Distortion Block-Resampling with Spatially Stochastic Networks](#) *Sarah Hong, Martin Arjovsky, Darryl Barnhart, Ian Thompson*
- [Continual Deep Learning by Functional Regularisation of Memorable Past](#) *Pingbo Pan, Siddharth Swaroop, Alexander Immer, Runa Eschenhagen, Richard Turner, Mohammad Emtiyaz E. Khan*
- [Distance Encoding: Design Provably More Powerful Neural Networks for Graph Representation Learning](#) *Pan Li, Yanbang Wang, Hongwei Wang, Jure Leskovec*
- [Fast Fourier Convolution](#) *Lu Chi, Borui Jiang, Yadong Mu*
- [Unsupervised Learning of Dense Visual Representations](#) *Pedro O. O. Pinheiro, Amjad Almahairi, Ryan Benmalek, Florian Golemo, Aaron C. Courville*
- [Higher-Order Certification For Randomized Smoothing](#) *Jeet Mohapatra, Ching-Yun Ko, Tsui-Wei Weng, Pin-Yu Chen, Sijia Liu, Luca Daniel*
- [Learning Structured Distributions From Untrusted Batches: Faster and Simpler](#) *Sitan Chen, Jerry Li, Ankur Moitra*
- [Hierarchical Quantized Autoencoders](#) *Will Williams, Sam Ringer, Tom Ash, David MacLeod, Jamie Dougherty, John Hughes*
- [Diversity can be Transferred: Output Diversification for White- and Black-box Attacks](#) *Yusuke Tashiro, Yang Song, Stefano Ermon*

- [POLY-HOOT: Monte-Carlo Planning in Continuous Space MDPs with Non-Asymptotic Analysis](#) *Weichao Mao, Kaiqing Zhang, Qiaomin Xie, Tamer Basar*
- [AvE: Assistance via Empowerment](#) *Yuqing Du, Stas Tiomkin, Emre Kiciman, Daniel Polani, Pieter Abbeel, Anca Dragan*
- [Variational Policy Gradient Method for Reinforcement Learning with General Utilities](#) *Junyu Zhang, Alec Koppel, Amrit Singh Bedi, Csaba Szepesvari, Mengdi Wang*
- [Reverse-engineering recurrent neural network solutions to a hierarchical inference task for mice](#) *Rylan Schaeffer, Mikail Khona, Leenoy Meshulam, Brain Laboratory International, Ila Fiete*
- [Temporal Positive-unlabeled Learning for Biomedical Hypothesis Generation via Risk Estimation](#) *Uchenna Akujuobi, Jun Chen, Mohamed Elhoseiny, Michael Spranger, Xiangliang Zhang*
- [Efficient Low Rank Gaussian Variational Inference for Neural Networks](#) *Marcin Tomczak, Siddharth Swaroop, Richard Turner*
- [Privacy Amplification via Random Check-Ins](#) *Borja Balle, Peter Kairouz, Brendan McMahan, Om Thakkar, Abhradeep Guha Thakurta*
- [Probabilistic Circuits for Variational Inference in Discrete Graphical Models](#) *Andy Shih, Stefano Ermon*
- [Your Classifier can Secretly Suffice Multi-Source Domain Adaptation](#) *Naveen Venkat, Jogendra Nath Kundu, Durgesh Singh, Ambareesh Revanur, Venkatesh Babu R*
- [Labelling unlabelled videos from scratch with multi-modal self-supervision](#) *Yuki Asano, Mandela Patrick, Christian Rupprecht, Andrea Vedaldi*
- [A Non-Asymptotic Analysis for Stein Variational Gradient Descent](#) *Anna Korba, Adil Salim, Michael Arbel, Giulia Luise, Arthur Gretton*
- [Robust Meta-learning for Mixed Linear Regression with Small Batches](#) *Weihaio Kong, Raghav Somani, Sham Kakade, Sewoong Oh*
- [Bayesian Deep Learning and a Probabilistic Perspective of Generalization](#) *Andrew G. Wilson, Pavel Izmailov*
- [Unsupervised Learning of Object Landmarks via Self-Training Correspondence](#) *Dimitrios Mallis, Enrique Sanchez, Matthew Bell, Georgios Tzimiropoulos*
- [Randomized tests for high-dimensional regression: A more efficient and powerful solution](#) *Yue Li, Ilmun Kim, Yuting Wei*
- [Learning Representations from Audio-Visual Spatial Alignment](#) *Pedro Morgado, Yi Li, Nuno Nvasconcelos*
- [Generative View Synthesis: From Single-view Semantics to Novel-view Images](#) *Tewodros Amberbir Habtegebrail, Varun Jampani, Orazio Gallo, Didier Stricker*
- [Towards More Practical Adversarial Attacks on Graph Neural Networks](#) *Jiaqi Ma, Shuangrui Ding, Qiaozhu Mei*
- [Multi-Task Reinforcement Learning with Soft Modularization](#) *Ruihan Yang, Huazhe Xu, YI WU, Xiaolong Wang*
- [Causal Shapley Values: Exploiting Causal Knowledge to Explain Individual Predictions of Complex Models](#) *Tom Heskes, Evi Sijben, Ioan Gabriel Bucur, Tom Claassen*
- [On the training dynamics of deep networks with L2L2 regularization](#) *Aitor Lewkowycz, Guy Gur-Ari*
- [Improved Algorithms for Convex-Concave Minimax Optimization](#) *Yuanhao Wang, Jian Li*
- [Deep Variational Instance Segmentation](#) *Jialin Yuan, Chao Chen, Fuxin Li*
- [Learning Implicit Functions for Topology-Varying Dense 3D Shape Correspondence](#) *Feng Liu, Xiaoming Liu*
- [Deep Multimodal Fusion by Channel Exchanging](#) *Yikai Wang, Wenbing Huang, Fuchun Sun, Tingyang Xu, Yu Rong, Junzhou Huang*
- [Hierarchically Organized Latent Modules for Exploratory Search in Morphogenetic Systems](#) *Mayalen Etcheverry, Clément Moulin-Frier, Pierre-Yves Oudeyer*

- [AI Feynman 2.0: Pareto-optimal symbolic regression exploiting graph modularity](#) *Silviu-Marian Udrescu, Andrew Tan, Jiahai Feng, Orisvaldo Neto, Tailin Wu, Max Tegmark*
- [Delay and Cooperation in Nonstochastic Linear Bandits](#) *Shinji Ito, Daisuke Hatano, Hanna Sumita, Kei Takemura, Takuro Fukunaga, Naonori Kakimura, Ken-Ichi Kawarabayashi*
- [Probabilistic Orientation Estimation with Matrix Fisher Distributions](#) *David Mohlin, Josephine Sullivan, Gérald Bianchi*
- [Minimax Dynamics of Optimally Balanced Spiking Networks of Excitatory and Inhibitory Neurons](#) *Qianyi Li, Cengiz Pehlevan*
- [Telescoping Density-Ratio Estimation](#) *Benjamin Rhodes, Kai Xu, Michael U. Gutmann*
- [Towards Deeper Graph Neural Networks with Differentiable Group Normalization](#) *Kaixiong Zhou, Xiao Huang, Yuening Li, Daochen Zha, Rui Chen, Xia Hu*
- [Stochastic Optimization for Performative Prediction](#) *Celestine Mender-Dünner, Juan Perdomo, Tijana Zrnic, Moritz Hardt*
- [Learning Differentiable Programs with Admissible Neural Heuristics](#) *Ameesh Shah, Eric Zhan, Jennifer Sun, Abhinav Verma, Yisong Yue, Swarat Chaudhuri*
- [Improved guarantees and a multiple-descent curve for Column Subset Selection and the Nystrom method](#) *Michal Dereziński, Rajiv Khanna, Michael W. Mahoney*
- [Domain Adaptation as a Problem of Inference on Graphical Models](#) *Kun Zhang, Mingming Gong, Petar Stojanov, Biwei Huang, QINGSONG LIU, Clark Glymour*
- [Network size and size of the weights in memorization with two-layers neural networks](#) *Sebastien Bubeck, Ronen Eldan, Yin Tat Lee, Dan Mikulincer*
- [Certifying Strategyproof Auction Networks](#) *Michael Curry, Ping-yeh Chiang, Tom Goldstein, John Dickerson*
- [Continual Learning of Control Primitives : Skill Discovery via Reset-Games](#) *Kelvin Xu, Siddharth Verma, Chelsea Finn, Sergey Levine*
- [HOI Analysis: Integrating and Decomposing Human-Object Interaction](#) *Yong-Lu Li, Xinpeng Liu, Xiaoqian Wu, Yizhuo Li, Cewu Lu*
- [Strongly local p-norm-cut algorithms for semi-supervised learning and local graph clustering](#) *Meng Liu, David F. Gleich*
- [Deep Direct Likelihood Knockoffs](#) *Mukund Sudarshan, Wesley Tansey, Rajesh Ranganath*
- [Meta-Neighborhoods](#) *Siyuan Shan, Yang Li, Junier B. Oliva*
- [Neural Dynamic Policies for End-to-End Sensorimotor Learning](#) *Shikhar Bahl, Mustafa Mukadam, Abhinav Gupta, Deepak Pathak*
- [A new inference approach for training shallow and deep generalized linear models of noisy interacting neurons](#) *Gabriel Mahuvas, Giulio Isacchini, Olivier Marre, Ulisse Ferrari, Thierry Mora*
- [Decision-Making with Auto-Encoding Variational Bayes](#) *Romain Lopez, Pierre Boyeau, Nir Yosef, Michael Jordan, Jeffrey Regier*
- [Attribution Preservation in Network Compression for Reliable Network Interpretation](#) *Geondo Park, June Yong Yang, Sung Ju Hwang, Eunho Yang*
- [Feature Importance Ranking for Deep Learning](#) *Maksymilian Wojtas, Ke Chen*
- [Causal Estimation with Functional Confounders](#) *Aahlad Puli, Adler Perotte, Rajesh Ranganath*
- [Model Inversion Networks for Model-Based Optimization](#) *Aviral Kumar, Sergey Levine*
- [Hausdorff Dimension, Heavy Tails, and Generalization in Neural Networks](#) *Umut Simsekli, Ozan Sener, George Deligiannidis, Murat A. Erdogdu*
- [Exact expressions for double descent and implicit regularization via surrogate random design](#) *Michal*

Derezinski, Feynman T. Liang, Michael W. Mahoney

- Certifying Confidence via Randomized Smoothing *Aounon Kumar, Alexander Levine, Soheil Feizi, Tom Goldstein*
- Learning Physical Constraints with Neural Projections *Shuqi Yang, Xingzhe He, Bo Zhu*
- Robust Optimization for Fairness with Noisy Protected Groups *Serena Wang, Wenshuo Guo, Harikrishna Narasimhan, Andrew Cotter, Maya Gupta, Michael Jordan*
- Noise-Contrastive Estimation for Multivariate Point Processes *Hongyuan Mei, Tom Wan, Jason Eisner*
- A Game-Theoretic Analysis of the Empirical Revenue Maximization Algorithm with Endogenous Sampling *Xiaotie Deng, Ron Lavi, Tao Lin, Qi Qi, Wenwei WANG, Xiang Yan*
- Neural Path Features and Neural Path Kernel : Understanding the role of gates in deep learning *Chandrashekar Lakshminarayanan, Amit Vikram Singh*
- Multiscale Deep Equilibrium Models *Shaojie Bai, Vladlen Koltun, J. Zico Kolter*
- Sparse Graphical Memory for Robust Planning *Scott Emmons, Ajay Jain, Misha Laskin, Thanard Kurutach, Pieter Abbeel, Deepak Pathak*
- Second Order PAC-Bayesian Bounds for the Weighted Majority Vote *Andres Masegosa, Stephan Lorenzen, Christian Igel, Yevgeny Seldin*
- Dirichlet Graph Variational Autoencoder *Jia Li, Jianwei Yu, Jiajin Li, Honglei Zhang, Kangfei Zhao, Yu Rong, Hong Cheng, Junzhou Huang*
- Modeling Task Effects on Meaning Representation in the Brain via Zero-Shot MEG Prediction *Mariya Toneva, Otilia Stretcu, Barnabas Póczos, Leila Wehbe, Tom M. Mitchell*
- Counterfactual Vision-and-Language Navigation: Unravelling the Unseen *Amin Parvaneh, Ehsan Abbasnejad, Damien Teney, Javen Qinfeng Shi, Anton van den Hengel*
- Robust Quantization: One Model to Rule Them All *moran shkolnik, Brian Chmiel, Ron Banner, Gil Shomron, Yury Nahshan, Alex Bronstein, Uri Weiser*
- Enabling certification of verification-agnostic networks via memory-efficient semidefinite programming *Sumanth Dathathri, Krishnamurthy Dvijotham, Alexey Kurakin, Aditi Raghunathan, Jonathan Uesato, Rudy R. Bunel, Shreya Shankar, Jacob Steinhardt, Ian Goodfellow, Percy S. Liang, Pushmeet Kohli*
- Federated Accelerated Stochastic Gradient Descent *Honglin Yuan, Tengyu Ma*
- Robust Density Estimation under Besov IPM Losses *Ananya Uppal, Shashank Singh, Barnabas Póczos*
- An analytic theory of shallow networks dynamics for hinge loss classification *Franco Pellegrini, Giulio Biroli*
- Fixed-Support Wasserstein Barycenters: Computational Hardness and Fast Algorithm *Tianyi Lin, Nhat Ho, Xi Chen, Marco Cuturi, Michael Jordan*
- Learning to Orient Surfaces by Self-supervised Spherical CNNs *Riccardo Spezialetti, Federico Stella, Marlon Marcon, Luciano Silva, Samuele Salti, Luigi Di Stefano*
- Adam with Bandit Sampling for Deep Learning *Rui Liu, Tianyi Wu, Barzan Mozafari*
- Parabolic Approximation Line Search for DNNs *Maximus Mutschler, Andreas Zell*
- Agnostic Learning of a Single Neuron with Gradient Descent *Spencer Frei, Yuan Cao, Quanquan Gu*
- Statistical Efficiency of Thompson Sampling for Combinatorial Semi-Bandits *Pierre Perrault, Etienne Boursier, Michal Valko, Vianney Perchet*
- Analytic Characterization of the Hessian in Shallow ReLU Models: A Tale of Symmetry *Yossi Arjevani, Michael Field*
- Generative causal explanations of black-box classifiers *Matthew O'Shaughnessy, Gregory Canal, Marissa Connor, Christopher Rozell, Mark Davenport*
- Sub-sampling for Efficient Non-Parametric Bandit Exploration *Dorian Baudry, Emilie Kaufmann, Odalric-*

Ambrym Maillard

- [Learning under Model Misspecification: Applications to Variational and Ensemble methods](#) *Andres Masegosa*
- [Language Through a Prism: A Spectral Approach for Multiscale Language Representations](#) *Alex Tamkin, Dan Jurafsky, Noah Goodman*
- [DVERGE: Diversifying Vulnerabilities for Enhanced Robust Generation of Ensembles](#) *Huanrui Yang, Jingyang Zhang, Hongliang Dong, Nathan Inkawhich, Andrew Gardner, Andrew Touchet, Wesley Wilkes, Heath Berry, Hai Li*
- [Towards practical differentially private causal graph discovery](#) *Lun Wang, Qi Pang, Dawn Song*
- [Independent Policy Gradient Methods for Competitive Reinforcement Learning](#) *Constantinos Daskalakis, Dylan J. Foster, Noah Golowich*
- [The Value Equivalence Principle for Model-Based Reinforcement Learning](#) *Christopher Grimm, Andre Barreto, Satinder Singh, David Silver*
- [Structured Convolutions for Efficient Neural Network Design](#) *Yash Bhalgat, Yizhe Zhang, Jamie Menjay Lin, Fatih Porikli*
- [Latent World Models For Intrinsically Motivated Exploration](#) *Aleksandr Ermolov, Nicu Sebe*
- [Estimating Rank-One Spikes from Heavy-Tailed Noise via Self-Avoiding Walks](#) *Jingqiu Ding, Samuel Hopkins, David Steurer*
- [Policy Improvement via Imitation of Multiple Oracles](#) *Ching-An Cheng, Andrey Kolobov, Alekh Agarwal*
- [Training Generative Adversarial Networks by Solving Ordinary Differential Equations](#) *Chongli Qin, Yan Wu, Jost Tobias Springenberg, Andy Brock, Jeff Donahue, Timothy Lillicrap, Pushmeet Kohli*
- [Learning of Discrete Graphical Models with Neural Networks](#) *Abhijith Jayakumar, Andrey Lokhov, Sidhant Misra, Marc Vuffray*
- [RepPoints v2: Verification Meets Regression for Object Detection](#) *Yihong Chen, Zheng Zhang, Yue Cao, Liwei Wang, Stephen Lin, Han Hu*
- [Unfolding the Alternating Optimization for Blind Super Resolution](#) *zhengxiong luo, Yan Huang, Shang Li, Liang Wang, Tieniu Tan*
- [Entrywise convergence of iterative methods for eigenproblems](#) *Vasileios Charisopoulos, Austin R. Benson, Anil Damle*
- [Learning Object-Centric Representations of Multi-Object Scenes from Multiple Views](#) *Nanbo Li, Cian Eastwood, Robert Fisher*
- [A Catalyst Framework for Minimax Optimization](#) *Junchi Yang, Siqi Zhang, Negar Kiyavash, Niao He*
- [Self-supervised Co-Training for Video Representation Learning](#) *Tengda Han, Weidi Xie, Andrew Zisserman*
- [Gradient Estimation with Stochastic Softmax Tricks](#) *Max Paulus, Dami Choi, Daniel Tarlow, Andreas Krause, Chris J. Maddison*
- [Meta-Learning Requires Meta-Augmentation](#) *Janarthanan Rajendran, Alexander Irpan, Eric Jang*
- [SLIP: Learning to predict in unknown dynamical systems with long-term memory](#) *Paria Rashidinejad, Jiantao Jiao, Stuart Russell*
- [Improving GAN Training with Probability Ratio Clipping and Sample Reweighting](#) *Yue Wu, Pan Zhou, Andrew G. Wilson, Eric Xing, Zhiting Hu*
- [Bayesian Bits: Unifying Quantization and Pruning](#) *Mart van Baalen, Christos Louizos, Markus Nagel, Rana Ali Amjad, Ying Wang, Tijmen Blankevoort, Max Welling*
- [On Testing of Samplers](#) *Kuldeep S Meel, Yash Pralhad Pote, Sourav Chakraborty*
- [Gaussian Process Bandit Optimization of the Thermodynamic Variational Objective](#) *Vu Nguyen, Vaden*

Masrani, Rob Brekelmans, Michael Osborne, Frank Wood

- MiniLM: Deep Self-Attention Distillation for Task-Agnostic Compression of Pre-Trained Transformers *Wenhui Wang, Furu Wei, Li Dong, Hangbo Bao, Nan Yang, Ming Zhou*
- Optimal Epoch Stochastic Gradient Descent Ascent Methods for Min-Max Optimization *Yan Yan, Yi Xu, Qihang Lin, Wei Liu, Tianbao Yang*
- Woodbury Transformations for Deep Generative Flows *You Lu, Bert Huang*
- Graph Contrastive Learning with Augmentations *Yuning You, Tianlong Chen, Yongduo Sui, Ting Chen, Zhangyang Wang, Yang Shen*
- Gradient Surgery for Multi-Task Learning *Tianhe Yu, Saurabh Kumar, Abhishek Gupta, Sergey Levine, Karol Hausman, Chelsea Finn*
- Bayesian Probabilistic Numerical Integration with Tree-Based Models *Harrison Zhu, Xing Liu, Ruya Kang, Zhichao Shen, Seth Flaxman, Francois-Xavier Briol*
- Deep learning versus kernel learning: an empirical study of loss landscape geometry and the time evolution of the Neural Tangent Kernel *Stanislav Fort, Gintare Karolina Dziugaite, Mansheej Paul, Sepideh Kharaghani, Daniel M. Roy, Surya Ganguli*
- Graph Meta Learning via Local Subgraphs *Kexin Huang, Marinka Zitnik*
- Stochastic Deep Gaussian Processes over Graphs *Naiqi Li, Wenjie Li, Jifeng Sun, Yinghua Gao, Yong Jiang, Shu-Tao Xia*
- Bayesian Causal Structural Learning with Zero-Inflated Poisson Bayesian Networks *Junsouk Choi, Robert Chapkin, Yang Ni*
- Evaluating Attribution for Graph Neural Networks *Benjamin Sanchez-Lengeling, Jennifer Wei, Brian Lee, Emily Reif, Peter Wang, Wesley Qian, Kevin McCloskey, Lucy Colwell, Alexander Wiltchko*
- On Second Order Behaviour in Augmented Neural ODEs *Alexander Norcliffe, Cristian Bodnar, Ben Day, Nikola Simidjievski, Pietro Lió*
- Neuron Shapley: Discovering the Responsible Neurons *Amirata Ghorbani, James Y. Zou*
- Stochastic Normalizing Flows *Hao Wu, Jonas Köhler, Frank Noe*
- GPU-Accelerated Primal Learning for Extremely Fast Large-Scale Classification *John T. Halloran, David M. Rocke*
- Random Reshuffling is Not Always Better *Christopher M. De Sa*
- Model Agnostic Multilevel Explanations *Karthikeyan Natesan Ramamurthy, Bhanukiran Vinzamuri, Yunfeng Zhang, Amit Dhurandhar*
- NeuMiss networks: differentiable programming for supervised learning with missing values. *Marine Le Morvan, Julie Josse, Thomas Moreau, Erwan Scornet, Gael Varoquaux*
- Revisiting Parameter Sharing for Automatic Neural Channel Number Search *Jiaxing Wang, Haoli Bai, Jiaxiang Wu, Xupeng Shi, Junzhou Huang, Irwin King, Michael Lyu, Jian Cheng*
- Differentially-Private Federated Linear Bandits *Abhimanyu Dubey, Alex 'Sandy' Pentland*
- Is Plug-in Solver Sample-Efficient for Feature-based Reinforcement Learning? *Qiwen Cui, Lin Yang*
- Learning Physical Graph Representations from Visual Scenes *Daniel Bear, Chaofei Fan, Damian Mrowca, Yunzhu Li, Seth Alter, Aran Nayebi, Jeremy Schwartz, Li F. Fei-Fei, Jiajun Wu, Josh Tenenbaum, Daniel L. Yamins*
- Deep Graph Pose: a semi-supervised deep graphical model for improved animal pose tracking *Anqi Wu, Estefany Kelly Buchanan, Matthew Whiteway, Michael Schartner, Guido Meijer, Jean-Paul Noel, Erica Rodriguez, Claire Everett, Amy Norovich, Evan Schaffer, Neeli Mishra, C. Daniel Salzman, Dora Angelaki, Andrés Bendesky, The International Brain Laboratory The International Brain Laboratory, John P.*

Cunningham, Liam Paninski

- Meta-learning from Tasks with Heterogeneous Attribute Spaces Tomoharu Iwata, Atsutoshi Kumagai
- Estimating decision tree learnability with polylogarithmic sample complexity Guy Blanc, Neha Gupta, Jane Lange, Li-Yang Tan
- Sparse Symplectically Integrated Neural Networks Daniel DiPietro, Shiyong Xiong, Bo Zhu
- Continuous Object Representation Networks: Novel View Synthesis without Target View Supervision Nicolai Hani, Selim Engin, Jun-Jee Chao, Volkan Isler
- Multimodal Generative Learning Utilizing Jensen-Shannon-Divergence Thomas Sutter, Imant Daunhawer, Julia Vogt
- Solver-in-the-Loop: Learning from Differentiable Physics to Interact with Iterative PDE-Solvers Kiwon Um, Robert Brand, Yun (Raymond) Fei, Philipp Holl, Nils Thuerey
- Reinforcement Learning with General Value Function Approximation: Provably Efficient Approach via Bounded Eluder Dimension Ruosong Wang, Russ R. Salakhutdinov, Lin Yang
- Predicting Training Time Without Training Luca Zancato, Alessandro Achille, Avinash Ravichandran, Rahul Bhotika, Stefano Soatto
- How does This Interaction Affect Me? Interpretable Attribution for Feature Interactions Michael Tsang, Sirisha Rambhatla, Yan Liu
- Optimal Adaptive Electrode Selection to Maximize Simultaneously Recorded Neuron Yield John Choi, Krishan Kumar, Mohammad Khazali, Katie Wingel, Mahdi Choudhury, Adam S. Charles, Bijan Pesaran
- Neurosymbolic Reinforcement Learning with Formally Verified Exploration Greg Anderson, Abhinav Verma, Isil Dillig, Swarat Chaudhuri
- Wavelet Flow: Fast Training of High Resolution Normalizing Flows Jason J. Yu, Konstantinos G. Derpanis, Marcus A. Brubaker
- Multi-task Batch Reinforcement Learning with Metric Learning Jiachen Li, Quan Vuong, Shuang Liu, Minghua Liu, Kamil Ciosek, Henrik Christensen, Hao Su
- On $1/n$ neural representation and robustness Josue Nassar, Piotr Sokol, Sueyeon Chung, Kenneth D. Harris, Il Memming Park
- Boundary thickness and robustness in learning models Yaoqing Yang, Rajiv Khanna, Yaodong Yu, Amir Gholami, Kurt Keutzer, Joseph E. Gonzalez, Kannan Ramchandran, Michael W. Mahoney
- Demixed shared component analysis of neural population data from multiple brain areas Yu Takagi, Steven Kennerley, Jun-ichiro Hirayama, Laurence Hunt
- Learning Kernel Tests Without Data Splitting Jonas Kübler, Wittawat Jitkrittum, Bernhard Schölkopf, Krikamol Muandet
- Unsupervised Data Augmentation for Consistency Training Qizhe Xie, Zihang Dai, Eduard Hovy, Thang Luong, Quoc Le
- Subgroup-based Rank-1 Lattice Quasi-Monte Carlo Yueming LYU, Yuan Yuan, Ivor Tsang
- Minibatch vs Local SGD for Heterogeneous Distributed Learning Blake E. Woodworth, Kumar Kshitij Patel, Nati Srebro
- Multi-task Causal Learning with Gaussian Processes Virginia Aglietti, Theodoros Damoulas, Mauricio Álvarez, Javier González
- Proximity Operator of the Matrix Perspective Function and its Applications Joong-Ho (Johann) Won
- Generative 3D Part Assembly via Dynamic Graph Learning jialei huang, Guanqi Zhan, Qingnan Fan, Kaichun Mo, Lin Shao, Baoquan Chen, Leonidas J. Guibas, Hao Dong
- Improving Natural Language Processing Tasks with Human Gaze-Guided Neural Attention Ekta Sood, Simon

Tannert, Philipp Mueller, Andreas Bulling

- The Power of Comparisons for Actively Learning Linear Classifiers *Max Hopkins, Daniel Kane, Shachar Lovett*
- From Boltzmann Machines to Neural Networks and Back Again *Surbhi Goel, Adam Klivans, Frederic Koehler*
- Crush Optimism with Pessimism: Structured Bandits Beyond Asymptotic Optimality *Kwang-Sung Jun, Chicheng Zhang*
- Pruning neural networks without any data by iteratively conserving synaptic flow *Hidenori Tanaka, Daniel Kunin, Daniel L. Yamins, Surya Ganguli*
- Detecting Interactions from Neural Networks via Topological Analysis *Zirui Liu, Qingquan Song, Kaixiong Zhou, Ting-Hsiang Wang, Ying Shan, Xia Hu*
- Neural Bridge Sampling for Evaluating Safety-Critical Autonomous Systems *Aman Sinha, Matthew O'Kelly, Russ Tedrake, John C. Duchi*
- Interpretable and Personalized Apprenticeship Scheduling: Learning Interpretable Scheduling Policies from Heterogeneous User Demonstrations *Rohan Paleja, Andrew Silva, Letian Chen, Matthew Gombolay*
- Task-Agnostic Online Reinforcement Learning with an Infinite Mixture of Gaussian Processes *Mengdi Xu, Wenhao Ding, Jiacheng Zhu, ZUXIN LIU, Baiming Chen, Ding Zhao*
- Benchmarking Deep Learning Interpretability in Time Series Predictions *Aya Abdelsalam Ismail, Mohamed Gunady, Hector Corrada Bravo, Soheil Feizi*
- Federated Principal Component Analysis *Andreas Grammenos, Rodrigo Mendoza Smith, Jon Crowcroft, Cecilia Mascolo*
- (De)Randomized Smoothing for Certifiable Defense against Patch Attacks *Alexander Levine, Soheil Feizi*
- SMYRF - Efficient Attention using Asymmetric Clustering *Giannis Daras, Nikita Kitaev, Augustus Odena, Alexandros G. Dimakis*
- Introducing Routing Uncertainty in Capsule Networks *Fabio De Sousa Ribeiro, Georgios Leontidis, Stefanos Kollias*
- A Simple and Efficient Smoothing Method for Faster Optimization and Local Exploration *Kevin Scaman, Ludovic DOS SANTOS, Merwan Barlier, Igor Colin*
- Hyperparameter Ensembles for Robustness and Uncertainty Quantification *Florian Wenzel, Jasper Snoek, Dustin Tran, Rodolphe Jenatton*
- Neutralizing Self-Selection Bias in Sampling for Sortition *Bailey Flanigan, Paul Gözl, Anupam Gupta, Ariel D. Procaccia*
- On the Convergence of Smooth Regularized Approximate Value Iteration Schemes *Elena Smirnova, Elvis Dohmatob*
- Off-Policy Evaluation via the Regularized Lagrangian *Mengjiao Yang, Ofir Nachum, Bo Dai, Lihong Li, Dale Schuurmans*
- The LoCA Regret: A Consistent Metric to Evaluate Model-Based Behavior in Reinforcement Learning *Harm Van Seijen, Hadi Nekoei, Evan Racah, Sarath Chandar*
- Neural Power Units *Niklas Heim, Tomas Pevny, Vasek Smidl*
- Towards Scalable Bayesian Learning of Causal DAGs *Jussi Viinikka, Antti Hyttinen, Johan Pensar, Mikko Koivisto*
- A Dictionary Approach to Domain-Invariant Learning in Deep Networks *Ze Wang, Xiuyuan Cheng, Guillermo Sapiro, Qiang Qiu*
- Bootstrapping neural processes *Juho Lee, Yoonho Lee, Jungtaek Kim, Eunho Yang, Sung Ju Hwang, Yee Whye Teh*

- [Large-Scale Adversarial Training for Vision-and-Language Representation Learning](#) Zhe Gan, Yen-Chun Chen, Linjie Li, Chen Zhu, Yu Cheng, Jingjing Liu
- [Most ReLU Networks Suffer from \$\ell_2\$ Adversarial Perturbations](#) Amit Daniely, Hadas Shacham
- [Compositional Visual Generation with Energy Based Models](#) Yilun Du, Shuang Li, Igor Mordatch
- [Factor Graph Grammars](#) David Chiang, Darcey Riley
- [Erdos Goes Neural: an Unsupervised Learning Framework for Combinatorial Optimization on Graphs](#) Nikolaos Karalias, Andreas Loukas
- [Autoregressive Score Matching](#) Chenlin Meng, Lantao Yu, Yang Song, Jiaming Song, Stefano Ermon
- [Debiasing Distributed Second Order Optimization with Surrogate Sketching and Scaled Regularization](#) Michal Derezhinski, Burak Bartan, Mert Pilanci, Michael W. Mahoney
- [Neural Controlled Differential Equations for Irregular Time Series](#) Patrick Kidger, James Morrill, James Foster, Terry Lyons
- [On Efficiency in Hierarchical Reinforcement Learning](#) Zheng Wen, Doina Precup, Morteza Ibrahimi, Andre Barreto, Benjamin Van Roy, Satinder Singh
- [On Correctness of Automatic Differentiation for Non-Differentiable Functions](#) Wonyeol Lee, Hangeol Yu, Xavier Rival, Hongseok Yang
- [Probabilistic Linear Solvers for Machine Learning](#) Jonathan Wenger, Philipp Hennig
- [Dynamic Regret of Policy Optimization in Non-Stationary Environments](#) Yingjie Fei, Zhuoran Yang, Zhaoran Wang, Qiaomin Xie
- [Multipole Graph Neural Operator for Parametric Partial Differential Equations](#) Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Andrew Stuart, Kaushik Bhattacharya, Anima Anandkumar
- [BlockGAN: Learning 3D Object-aware Scene Representations from Unlabelled Images](#) Thu H. Nguyen-Phuoc, Christian Richardt, Long Mai, Yongliang Yang, Niloy Mitra
- [Online Structured Meta-learning](#) Huaxiu Yao, Yingbo Zhou, Mehrdad Mahdavi, Zhenhui (Jessie) Li, Richard Socher, Caiming Xiong
- [Learning Strategic Network Emergence Games](#) Rakshit Trivedi, Hongyuan Zha
- [Towards Interpretable Natural Language Understanding with Explanations as Latent Variables](#) Wangchunshu Zhou, Jinyi Hu, Hanlin Zhang, Xiaodan Liang, Maosong Sun, Chenyan Xiong, Jian Tang
- [The Mean-Squared Error of Double Q-Learning](#) Wentao Weng, Harsh Gupta, Niao He, Lei Ying, R. Srikant
- [What Makes for Good Views for Contrastive Learning?](#) Yonglong Tian, Chen Sun, Ben Poole, Dilip Krishnan, Cordelia Schmid, Phillip Isola
- [Denoising Diffusion Probabilistic Models](#) Jonathan Ho, Ajay Jain, Pieter Abbeel
- [Barking up the right tree: an approach to search over molecule synthesis DAGs](#) John Bradshaw, Brooks Paige, Matt J. Kusner, Marwin Segler, José Miguel Hernández-Lobato
- [On Uniform Convergence and Low-Norm Interpolation Learning](#) Lijia Zhou, Danica J. Sutherland, Nati Srebro
- [Bandit Samplers for Training Graph Neural Networks](#) Ziqi Liu, Zhengwei Wu, Zhiqiang Zhang, Jun Zhou, Shuang Yang, Le Song, Yuan Qi
- [Sampling from a k-DPP without looking at all items](#) Daniele Calandriello, Michal Derezhinski, Michal Valko
- [Uncovering the Topology of Time-Varying fMRI Data using Cubical Persistence](#) Bastian Rieck, Tristan Yates, Christian Bock, Karsten Borgwardt, Guy Wolf, Nicholas Turk-Browne, Smita Krishnaswamy
- [Hierarchical Poset Decoding for Compositional Generalization in Language](#) Yinuo Guo, Zeqi Lin, Jian-Guang Lou, Dongmei Zhang
- [Evaluating and Rewarding Teamwork Using Cooperative Game Abstractions](#) Tom Yan, Christian Kroer, Alexander Peysakhovich

- [Exchangeable Neural ODE for Set Modeling](#) Yang Li, Haidong Yi, Christopher Bender, Siyuan Shan, Junier B. Oliva
- [Profile Entropy: A Fundamental Measure for the Learnability and Compressibility of Distributions](#) Yi Hao, Alon Orlitsky
- [CoADNet: Collaborative Aggregation-and-Distribution Networks for Co-Salient Object Detection](#) Qijian Zhang, Runmin Cong, Junhui Hou, Chongyi Li, Yao Zhao
- [Regularized linear autoencoders recover the principal components, eventually](#) Xuchan Bao, James Lucas, Sushant Sachdeva, Roger B. Grosse
- [Semi-Supervised Partial Label Learning via Confidence-Rated Margin Maximization](#) Wei Wang, Min-Ling Zhang
- [GramGAN: Deep 3D Texture Synthesis From 2D Exemplars](#) Tiziano Portenier, Siavash Arjomand Bigdeli, Orcun Goksel
- [UWSOD: Toward Fully-Supervised-Level Capacity Weakly Supervised Object Detection](#) Yunhang Shen, Rongrong Ji, Zhiwei Chen, Yongjian Wu, Feiyue Huang
- [Learning Restricted Boltzmann Machines with Sparse Latent Variables](#) Guy Bresler, Rares-Darius Buhai
- [Sample Complexity of Asynchronous Q-Learning: Sharper Analysis and Variance Reduction](#) Gen Li, Yuting Wei, Yuejie Chi, Yuantao Gu, Yuxin Chen
- [Curriculum learning for multilevel budgeted combinatorial problems](#) Adel Nabli, Margarida Carvalho
- [FedSplit: an algorithmic framework for fast federated optimization](#) Reese Pathak, Martin J. Wainwright
- [Estimation and Imputation in Probabilistic Principal Component Analysis with Missing Not At Random Data](#) Aude Sportisse, Claire Boyer, Julie Josse
- [Correlation Robust Influence Maximization](#) Louis Chen, Divya Padmanabhan, Chee Chin Lim, Karthik Natarajan
- [Neuronal Gaussian Process Regression](#) Johannes Friedrich
- [Nonconvex Sparse Graph Learning under Laplacian Constrained Graphical Model](#) Jiaxi Ying, José Vinícius de Miranda Cardoso, Daniel Palomar
- [Synthetic Data Generators -- Sequential and Private](#) Olivier Bousquet, Roi Livni, Shay Moran
- [Uncertainty Quantification for Inferring Hawkes Networks](#) Haoyun Wang, Liyan Xie, Alex Cuzzo, Simon Mak, Yao Xie
- [Implicit Distributional Reinforcement Learning](#) Yuguang Yue, Zhendong Wang, Mingyuan Zhou
- [Auxiliary Task Reweighting for Minimum-data Learning](#) Baifeng Shi, Judy Hoffman, Kate Saenko, Trevor Darrell, Huijuan Xu
- [Small Nash Equilibrium Certificates in Very Large Games](#) Brian Zhang, Tuomas Sandholm
- [Training Linear Finite-State Machines](#) Arash Ardakani, Amir Ardakani, Warren Gross
- [Efficient active learning of sparse halfspaces with arbitrary bounded noise](#) Chicheng Zhang, Jie Shen, Pranjal Awasthi
- [Swapping Autoencoder for Deep Image Manipulation](#) Taesung Park, Jun-Yan Zhu, Oliver Wang, Jingwan Lu, Eli Shechtman, Alexei Efros, Richard Zhang
- [Self-Supervised Few-Shot Learning on Point Clouds](#) Charu Sharma, Manohar Kaul
- [Faster Differentially Private Samplers via Rényi Divergence Analysis of Discretized Langevin MCMC](#) Arun Ganesh, Kunal Talwar
- [Learning identifiable and interpretable latent models of high-dimensional neural activity using pi-VAE](#) Ding Zhou, Xue-Xin Wei
- [RL Unplugged: A Collection of Benchmarks for Offline Reinforcement Learning](#) Caglar Gulcehre, Ziyu Wang,

Alexander Novikov, Thomas Paine, Sergio Gómez, Konrad Zolna, Rishabh Agarwal, Josh S. Merel, Daniel J. Mankowitz, Cosmin Paduraru, Gabriel Dulac-Arnold, Jerry Li, Mohammad Norouzi, Matthew Hoffman, Nicolas Heess, Nando de Freitas

- [Dual T: Reducing Estimation Error for Transition Matrix in Label-noise Learning](#) Yu Yao, Tongliang Liu, Bo Han, Mingming Gong, Jiankang Deng, Gang Niu, Masashi Sugiyama
- [Interior Point Solving for LP-based prediction+optimisation](#) Jayanta Mandi, Tias Guns
- [A simple normative network approximates local non-Hebbian learning in the cortex](#) Siavash Golkar, David Lipshutz, Yanis Bahroun, Anirvan Sengupta, Dmitri Chklovskii
- [Kernelized information bottleneck leads to biologically plausible 3-factor Hebbian learning in deep networks](#) Roman Pogodin, Peter Latham
- [Understanding the Role of Training Regimes in Continual Learning](#) Seyed Iman Mirzadeh, Mehrdad Farajtabar, Razvan Pascanu, Hassan Ghasemzadeh
- [Fair regression with Wasserstein barycenters](#) Evgenii Chzhen, Christophe Denis, Mohamed Hebiri, Luca Oneto, Massimiliano Pontil
- [Training Stronger Baselines for Learning to Optimize](#) Tianlong Chen, Weiyi Zhang, Zhou Jingyang, Shiyu Chang, Sijia Liu, Lisa Amini, Zhangyang Wang
- [Exactly Computing the Local Lipschitz Constant of ReLU Networks](#) Matt Jordan, Alexandros G. Dimakis
- [Strictly Batch Imitation Learning by Energy-based Distribution Matching](#) Daniel Jarrett, Ioana Bica, Mihaela van der Schaar
- [On the Ergodicity, Bias and Asymptotic Normality of Randomized Midpoint Sampling Method](#) Ye He, Krishnakumar Balasubramanian, Murat A. Erdogdu
- [A Single-Loop Smoothed Gradient Descent-Ascent Algorithm for Nonconvex-Concave Min-Max Problems](#) Jiawei Zhang, Peijun Xiao, Ruoyu Sun, Zhiquan Luo
- [Generating Correct Answers for Progressive Matrices Intelligence Tests](#) Niv Pekar, Yaniv Benny, Lior Wolf
- [HyNet: Learning Local Descriptor with Hybrid Similarity Measure and Triplet Loss](#) Yurun Tian, Axel Barroso Laguna, Tony Ng, Vassileios Balntas, Krystian Mikolajczyk
- [Preference learning along multiple criteria: A game-theoretic perspective](#) Kush Bhatia, Ashwin Pananjady, Peter Bartlett, Anca Dragan, Martin J. Wainwright
- [Multi-Plane Program Induction with 3D Box Priors](#) Yikai Li, Jiayuan Mao, Xiuming Zhang, Bill Freeman, Josh Tenenbaum, Noah Snavely, Jiajun Wu
- [Online Neural Connectivity Estimation with Noisy Group Testing](#) Anne Draelos, John Pearson
- [Once-for-All Adversarial Training: In-Situ Tradeoff between Robustness and Accuracy for Free](#) Haotao N. Wang, Tianlong Chen, Shupeng Gui, Tingkuei Hu, Ji Liu, Zhangyang Wang
- [Implicit Neural Representations with Periodic Activation Functions](#) Vincent Sitzmann, Julien Martel, Alexander Bergman, David Lindell, Gordon Wetzstein
- [Rotated Binary Neural Network](#) Mingbao Lin, Rongrong Ji, Zihan Xu, Baochang Zhang, Yan Wang, Yongjian Wu, Feiyue Huang, Chia-Wen Lin
- [Community detection in sparse time-evolving graphs with a dynamical Bethe-Hessian](#) Lorenzo Dall'Amico, Romain Couillet, Nicolas Tremblay
- [Simple and Principled Uncertainty Estimation with Deterministic Deep Learning via Distance Awareness](#) Jeremiah Liu, Zi Lin, Shreyas Padhy, Dustin Tran, Tania Bedrax Weiss, Balaji Lakshminarayanan
- [Adaptive Learning of Rank-One Models for Efficient Pairwise Sequence Alignment](#) Govinda Kamath, Tavor Baharav, Ilan Shomorony
- [Hierarchical nucleation in deep neural networks](#) Diego Doimo, Aldo Glielmo, Alessio Ansuini, Alessandro

Laio

- [Fourier Features Let Networks Learn High Frequency Functions in Low Dimensional Domains](#) *Matthew Tancik, Pratul Srinivasan, Ben Mildenhall, Sara Fridovich-Keil, Nithin Raghavan, Utkarsh Singhal, Ravi Ramamoorthi, Jonathan Barron, Ren Ng*
- [Graph Geometry Interaction Learning](#) *Shichao Zhu, Shirui Pan, Chuan Zhou, Jia Wu, Yanan Cao, Bin Wang*
- [Differentiable Augmentation for Data-Efficient GAN Training](#) *Shengyu Zhao, Zhijian Liu, Ji Lin, Jun-Yan Zhu, Song Han*
- [Heuristic Domain Adaptation](#) *Shuhao Cui, Xuan Jin, Shuhui Wang, Yuan He, Qingming Huang*
- [Learning Certified Individually Fair Representations](#) *Anian Ruoss, Mislav Balunovic, Marc Fischer, Martin Vechev*
- [Part-dependent Label Noise: Towards Instance-dependent Label Noise](#) *Xiaobo Xia, Tongliang Liu, Bo Han, Nannan Wang, Mingming Gong, Haifeng Liu, Gang Niu, Dacheng Tao, Masashi Sugiyama*
- [Tackling the Objective Inconsistency Problem in Heterogeneous Federated Optimization](#) *Jianyu Wang, Qinghua Liu, Hao Liang, Gauri Joshi, H. Vincent Poor*
- [An Improved Analysis of \(Variance-Reduced\) Policy Gradient and Natural Policy Gradient Methods](#) *Yanli Liu, Kaiqing Zhang, Tamer Basar, Wotao Yin*
- [Geometric Exploration for Online Control](#) *Orestis Plevrakis, Elad Hazan*
- [Automatic Curriculum Learning through Value Disagreement](#) *Yunzhi Zhang, Pieter Abbeel, Lerrel Pinto*
- [MRI Banding Removal via Adversarial Training](#) *Aaron Defazio, Tullie Murrell, Michael Recht*
- [The NetHack Learning Environment](#) *Heinrich Küttler, Nantas Nardelli, Alexander Miller, Roberta Raileanu, Marco Selvatici, Edward Grefenstette, Tim Rocktäschel*
- [Language and Visual Entity Relationship Graph for Agent Navigation](#) *Yicong Hong, Cristian Rodriguez, Yuankai Qi, Qi Wu, Stephen Gould*
- [ICAM: Interpretable Classification via Disentangled Representations and Feature Attribution Mapping](#) *Cher Bass, Mariana da Silva, Carole Sudre, Petru-Daniel Tudosi, Stephen Smith, Emma Robinson*
- [Spectra of the Conjugate Kernel and Neural Tangent Kernel for linear-width neural networks](#) *Zhou Fan, Zhichao Wang*
- [No-Regret Learning Dynamics for Extensive-Form Correlated Equilibrium](#) *Andrea Celli, Alberto Marchesi, Gabriele Farina, Nicola Gatti*
- [Estimating weighted areas under the ROC curve](#) *Andreas Maurer, Massimiliano Pontil*
- [Can Implicit Bias Explain Generalization? Stochastic Convex Optimization as a Case Study](#) *Assaf Dauber, Meir Feder, Tomer Koren, Roi Livni*
- [Generalized Hindsight for Reinforcement Learning](#) *Alexander Li, Lerrel Pinto, Pieter Abbeel*
- [Critic Regularized Regression](#) *Ziyu Wang, Alexander Novikov, Konrad Zolna, Josh S. Merel, Jost Tobias Springenberg, Scott E. Reed, Bobak Shahriari, Noah Siegel, Caglar Gulcehre, Nicolas Heess, Nando de Freitas*
- [Boosting Adversarial Training with Hypersphere Embedding](#) *Tianyu Pang, Xiao Yang, Yinpeng Dong, Kun Xu, Jun Zhu, Hang Su*
- [Beyond Homophily in Graph Neural Networks: Current Limitations and Effective Designs](#) *Jiong Zhu, Yujun Yan, Lingxiao Zhao, Mark Heimann, Leman Akoglu, Danai Koutra*
- [Modeling Continuous Stochastic Processes with Dynamic Normalizing Flows](#) *Ruizhi Deng, Bo Chang, Marcus A. Brubaker, Greg Mori, Andreas Lehrmann*
- [Efficient Online Learning of Optimal Rankings: Dimensionality Reduction via Gradient Descent](#) *Dimitris Fotakis, Thanasis Lianas, Georgios Piliouras, Stratis Skoulakis*

- [Training Normalizing Flows with the Information Bottleneck for Competitive Generative Classification](#) *Lynton Ardizzone, Radek Mackowiak, Carsten Rother, Ullrich Köthe*
- [Detecting Hands and Recognizing Physical Contact in the Wild](#) *Supreeth Narasimhaswamy, Trung Nguyen, Minh Hoai Nguyen*
- [On the Theory of Transfer Learning: The Importance of Task Diversity](#) *Nilesh Tripuraneni, Michael Jordan, Chi Jin*
- [Finite-Time Analysis of Round-Robin Kullback-Leibler Upper Confidence Bounds for Optimal Adaptive Allocation with Multiple Plays and Markovian Rewards](#) *Vrettos Moulos*
- [Neural Star Domain as Primitive Representation](#) *Yuki Kawana, Yusuke Mukuta, Tatsuya Harada*
- [Off-Policy Interval Estimation with Lipschitz Value Iteration](#) *Ziyang Tang, Yihao Feng, Na Zhang, Jian Peng, Qiang Liu*
- [Inverse Rational Control with Partially Observable Continuous Nonlinear Dynamics](#) *Minhae Kwon, Saurabh Daptardar, Paul R. Schrater, Xaq Pitkow*
- [Deep Statistical Solvers](#) *Balthazar Donon, Zhengying Liu, Wenzhuo LIU, Isabelle Guyon, Antoine Marot, Marc Schoenauer*
- [Distributionally Robust Parametric Maximum Likelihood Estimation](#) *Viet Anh Nguyen, Xuhui Zhang, Jose Blanchet, Angelos Georgioui*
- [Secretary and Online Matching Problems with Machine Learned Advice](#) *Antonios Antoniadis, Themis Gouleakis, Pieter Kleer, Pavel Kolev*
- [Deep Transformation-Invariant Clustering](#) *Tom Monnier, Thibault Groueix, Mathieu Aubry*
- [Overfitting Can Be Harmless for Basis Pursuit, But Only to a Degree](#) *Peizhong Ju, Xiaojun Lin, Jia Liu*
- [Improving Generalization in Reinforcement Learning with Mixture Regularization](#) *KAIXIN WANG, Bingyi Kang, Jie Shao, Jiashi Feng*
- [Pontryagin Differentiable Programming: An End-to-End Learning and Control Framework](#) *Wanxin Jin, Zhaoran Wang, Zhuoran Yang, Shaoshuai Mou*
- [Learning from Aggregate Observations](#) *Yivan Zhang, Nontawat Charoenphakdee, Zhenguo Wu, Masashi Sugiyama*
- [The Devil is in the Detail: A Framework for Macroscopic Prediction via Microscopic Models](#) *Yingxiang Yang, Negar Kiyavash, Le Song, Niao He*
- [Subgraph Neural Networks](#) *Emily Alsentzer, Samuel Finlayson, Michelle Li, Marinka Zitnik*
- [Demystifying Orthogonal Monte Carlo and Beyond](#) *Han Lin, Haoxian Chen, Krzysztof M. Choromanski, Tianyi Zhang, Clement Laroche*
- [Optimal Robustness-Consistency Trade-offs for Learning-Augmented Online Algorithms](#) *Alexander Wei, Fred Zhang*
- [A Scalable Approach for Privacy-Preserving Collaborative Machine Learning](#) *Jinhyun So, Basak Guler, Salman Avestimehr*
- [Glow-TTS: A Generative Flow for Text-to-Speech via Monotonic Alignment Search](#) *Jaehyeon Kim, Sungwon Kim, Jungil Kong, Sungroh Yoon*
- [Towards Learning Convolutions from Scratch](#) *Behnam Neyshabur*
- [Cycle-Contrast for Self-Supervised Video Representation Learning](#) *Quan Kong, Wenpeng Wei, Ziwei Deng, Tomoaki Yoshinaga, Tomokazu Murakami*
- [Posterior Re-calibration for Imbalanced Datasets](#) *Junjiao Tian, Yen-Cheng Liu, Nathaniel Glaser, Yen-Chang Hsu, Zsolt Kira*
- [Novelty Search in Representational Space for Sample Efficient Exploration](#) *Ruo Yu Tao, Vincent Francois-*

Lavet, Joelle Pineau

- Robust Reinforcement Learning via Adversarial training with Langevin Dynamics Parameswaran Kamalaruban, Yu-Ting Huang, Ya-Ping Hsieh, Paul Rolland, Cheng Shi, Volkan Cevher
- Adversarial Blocking Bandits Nicholas Bishop, Hau Chan, Debmalya Mandal, Long Tran-Thanh
- Online Algorithms for Multi-shop Ski Rental with Machine Learned Advice Shufan Wang, Jian Li, Shiqiang Wang
- Multi-label Contrastive Predictive Coding Jiaming Song, Stefano Ermon
- Rotation-Invariant Local-to-Global Representation Learning for 3D Point Cloud SEOHYUN KIM, JaeYoo Park, Bohyung Han
- Learning Invariants through Soft Unification Nuri Cingillioglu, Alessandra Russo
- One Solution is Not All You Need: Few-Shot Extrapolation via Structured MaxEnt RL Saurabh Kumar, Aviral Kumar, Sergey Levine, Chelsea Finn
- Variational Bayesian Monte Carlo with Noisy Likelihoods Luigi Acerbi
- Finite-Sample Analysis of Contractive Stochastic Approximation Using Smooth Convex Envelopes Zaiwei Chen, Siva Theja Maguluri, Sanjay Shakkottai, Karthikeyan Shanmugam
- Self-Supervised Generative Adversarial Compression Chong Yu, Jeff Pool
- An efficient nonconvex reformulation of stagewise convex optimization problems Rudy R. Bunel, Oliver Hinder, Srinadh Bhojanapalli, Krishnamurthy Dvijotham
- From Finite to Countable-Armed Bandits Anand Kalvit, Assaf Zeevi
- Adversarial Distributional Training for Robust Deep Learning Yinpeng Dong, Zhijie Deng, Tianyu Pang, Jun Zhu, Hang Su
- Meta-Learning Stationary Stochastic Process Prediction with Convolutional Neural Processes Andrew Foong, Wessel Bruinsma, Jonathan Gordon, Yann Dubois, James Requeima, Richard Turner
- Theory-Inspired Path-Regularized Differential Network Architecture Search Pan Zhou, Caiming Xiong, Richard Socher, Steven Chu Hong Hoi
- Conic Descent and its Application to Memory-efficient Optimization over Positive Semidefinite Matrices John C. Duchi, Oliver Hinder, Andrew Naber, Yinyu Ye
- Learning the Geometry of Wave-Based Imaging Konik Kothari, Maarten de Hoop, Ivan Dokmanić
- Greedy inference with structure-exploiting lazy maps Michael Brennan, Daniele Bigoni, Olivier Zahm, Alessio Spantini, Youssef Marzouk
- Nimble: Lightweight and Parallel GPU Task Scheduling for Deep Learning Woosuk Kwon, Gyeong-In Yu, Eunji Jeong, Byung-Gon Chun
- Finding the Homology of Decision Boundaries with Active Learning Weizhi Li, Gautam Dasarathy, Karthikeyan Natesan Ramamurthy, Visar Berisha
- Reinforced Molecular Optimization with Neighborhood-Controlled Grammars Chencheng Xu, Qiao Liu, Minlie Huang, Tao Jiang
- Natural Policy Gradient Primal-Dual Method for Constrained Markov Decision Processes Dongsheng Ding, Kaiqing Zhang, Tamer Basar, Mihailo Jovanovic
- Classification Under Misspecification: Halfspaces, Generalized Linear Models, and Evolvability Sitan Chen, Frederic Koehler, Ankur Moitra, Morris Yau
- Certified Defense to Image Transformations via Randomized Smoothing Marc Fischer, Maximilian Baader, Martin Vechev
- Estimation of Skill Distribution from a Tournament Ali Jadbabaie, Anuran Makur, Devavrat Shah
- Reparameterizing Mirror Descent as Gradient Descent Ehsan Amid, Manfred K. K. Warmuth

- [General Control Functions for Causal Effect Estimation from IVs](#) *Aahlad Puli, Rajesh Ranganath*
- [Optimal Algorithms for Stochastic Multi-Armed Bandits with Heavy Tailed Rewards](#) *Kyungjae Lee, Hongjun Yang, Sungbin Lim, Songhwai Oh*
- [Certified Robustness of Graph Convolution Networks for Graph Classification under Topological Attacks](#) *Hongwei Jin, Zhan Shi, Venkata Jaya Shankar Ashish Peruri, Xinhua Zhang*
- [Zero-Resource Knowledge-Grounded Dialogue Generation](#) *Linxiao Li, Can Xu, Wei Wu, YUFAN ZHAO, Xueliang Zhao, Chongyang Tao*
- [Targeted Adversarial Perturbations for Monocular Depth Prediction](#) *Alex Wong, Safa Cicek, Stefano Soatto*
- [Beyond the Mean-Field: Structured Deep Gaussian Processes Improve the Predictive Uncertainties](#) *Jakob Lindinger, David Reeb, Christoph Lippert, Barbara Rakitsch*
- [Offline Imitation Learning with a Misspecified Simulator](#) *Shengyi Jiang, Jingcheng Pang, Yang Yu*
- [Multi-Fidelity Bayesian Optimization via Deep Neural Networks](#) *Shibo Li, Wei Xing, Robert Kirby, Shandian Zhe*
- [PlanGAN: Model-based Planning With Sparse Rewards and Multiple Goals](#) *Henry Charlesworth, Giovanni Montana*
- [Bad Global Minima Exist and SGD Can Reach Them](#) *Shengchao Liu, Dimitris Papailiopoulos, Dimitris Achlioptas*
- [Optimal Prediction of the Number of Unseen Species with Multiplicity](#) *Yi Hao, Ping Li*
- [Characterizing Optimal Mixed Policies: Where to Intervene and What to Observe](#) *Sanghack Lee, Elias Bareinboim*
- [Factor Graph Neural Networks](#) *Zhen Zhang, Fan Wu, Wee Sun Lee*
- [A Closer Look at Accuracy vs. Robustness](#) *Yao-Yuan Yang, Cyrus Rashtchian, Hongyang Zhang, Russ R. Salakhutdinov, Kamalika Chaudhuri*
- [Curriculum Learning by Dynamic Instance Hardness](#) *Tianyi Zhou, Shengjie Wang, Jeffrey Bilmes*
- [Spin-Weighted Spherical CNNs](#) *Carlos Esteves, Ameesh Makadia, Kostas Daniilidis*
- [Learning to Execute Programs with Instruction Pointer Attention Graph Neural Networks](#) *David Bieber, Charles Sutton, Hugo Larochelle, Daniel Tarlow*
- [AutoPrivacy: Automated Layer-wise Parameter Selection for Secure Neural Network Inference](#) *Qian Lou, Song Bian, Lei Jiang*
- [Baxter Permutation Process](#) *Masahiro Nakano, Akisato Kimura, Takeshi Yamada, Naonori Ueda*
- [Characterizing emergent representations in a space of candidate learning rules for deep networks](#) *Yinan Cao, Christopher Summerfield, Andrew Saxe*
- [Fast, Accurate, and Simple Models for Tabular Data via Augmented Distillation](#) *Rasool Fakoor, Jonas W. Mueller, Nick Erickson, Pratik Chaudhari, Alexander J. Smola*
- [Adaptive Probing Policies for Shortest Path Routing](#) *Aditya Bhaskara, Sreenivas Gollapudi, Kostas Kollias, Kamesh Munagala*
- [Approximate Heavily-Constrained Learning with Lagrange Multiplier Models](#) *Harikrishna Narasimhan, Andrew Cotter, Yichen Zhou, Serena Wang, Wenshuo Guo*
- [Faster Randomized Infeasible Interior Point Methods for Tall/Wide Linear Programs](#) *Agniva Chowdhury, Palma London, Haim Avron, Petros Drineas*
- [Sliding Window Algorithms for k-Clustering Problems](#) *Michele Borassi, Alessandro Epasto, Silvio Lattanzi, Sergei Vassilvitskii, Morteza Zadimoghaddam*
- [AdaShare: Learning What To Share For Efficient Deep Multi-Task Learning](#) *Ximeng Sun, Rameswar Panda, Rogerio Feris, Kate Saenko*

- [Approximate Cross-Validation for Structured Models](#) *Soumya Ghosh, Will Stephenson, Tin D. Nguyen, Sameer Deshpande, Tamara Broderick*
- [Exemplar VAE: Linking Generative Models, Nearest Neighbor Retrieval, and Data Augmentation](#) *Sajad Norouzi, David J. Fleet, Mohammad Norouzi*
- [Debiased Contrastive Learning](#) *Ching-Yao Chuang, Joshua Robinson, Yen-Chen Lin, Antonio Torralba, Stefanie Jegelka*
- [UCSG-NET- Unsupervised Discovering of Constructive Solid Geometry Tree](#) *Kacper Kania, Maciej Zieba, Tomasz Kajdanowicz*
- [Generalized Boosting](#) *Arun Suggala, Bingbin Liu, Pradeep Ravikumar*
- [COT-GAN: Generating Sequential Data via Causal Optimal Transport](#) *Tianlin Xu, Li Kevin Wenliang, Michael Munn, Beatrice Acciaio*
- [Impossibility Results for Grammar-Compressed Linear Algebra](#) *Amir Abboud, Arturs Backurs, Karl Bringmann, Marvin Künnemann*
- [Understanding spiking networks through convex optimization](#) *Allan Mancoo, Sander Keemink, Christian K. Machens*
- [Better Full-Matrix Regret via Parameter-Free Online Learning](#) *Ashok Cutkosky*
- [Large-Scale Methods for Distributionally Robust Optimization](#) *Daniel Levy, Yair Carmon, John C. Duchi, Aaron Sidford*
- [Analysis and Design of Thompson Sampling for Stochastic Partial Monitoring](#) *Taira Tsuchiya, Junya Honda, Masashi Sugiyama*
- [Bandit Linear Control](#) *Asaf Cassel, Tomer Koren*
- [Refactoring Policy for Compositional Generalizability using Self-Supervised Object Proposals](#) *Tongzhou Mu, Jiayuan Gu, Zhiwei Jia, Hao Tang, Hao Su*
- [PEP: Parameter Ensembling by Perturbation](#) *Alireza Mehrtash, Purang Abolmaesumi, Polina Golland, Tina Kapur, Demian Wassermann, William Wells*
- [Theoretical Insights Into Multiclass Classification: A High-dimensional Asymptotic View](#) *Christos Thrampoulidis, Samet Oymak, Mahdi Soltanolkotabi*
- [Adversarial Example Games](#) *Joey Bose, Gauthier Gidel, Hugo Berard, Andre Cianflone, Pascal Vincent, Simon Lacoste-Julien, Will Hamilton*
- [Residual Distillation: Towards Portable Deep Neural Networks without Shortcuts](#) *Guilin Li, Junlei Zhang, Yunhe Wang, Chuanjian Liu, Matthias Tan, Yunfeng Lin, Wei Zhang, Jiashi Feng, Tong Zhang*
- [Provably Efficient Neural Estimation of Structural Equation Models: An Adversarial Approach](#) *Luofeng Liao, You-Lin Chen, Zhuoran Yang, Bo Dai, Mladen Kolar, Zhaoran Wang*
- [Security Analysis of Safe and Seldonian Reinforcement Learning Algorithms](#) *Pinar Ozisik, Philip S. Thomas*
- [Learning to Play Sequential Games versus Unknown Opponents](#) *Pier Giuseppe Sessa, Ilija Bogunovic, Maryam Kamgarpour, Andreas Krause*
- [Further Analysis of Outlier Detection with Deep Generative Models](#) *Ziyu Wang, Bin Dai, David Wipf, Jun Zhu*
- [Bridging Imagination and Reality for Model-Based Deep Reinforcement Learning](#) *Guangxiang Zhu, Minghao Zhang, Honglak Lee, Chongjie Zhang*
- [Neural Networks Learning and Memorization with \(almost\) no Over-Parameterization](#) *Amit Daniely*
- [Exploiting Higher Order Smoothness in Derivative-free Optimization and Continuous Bandits](#) *Arya Akhavan, Massimiliano Pontil, Alexandre Tsybakov*
- [Towards a Combinatorial Characterization of Bounded-Memory Learning](#) *Alon Gonen, Shachar Lovett, Michal Moshkovitz*

- [Chaos, Extremism and Optimism: Volume Analysis of Learning in Games](#) Yun Kuen Cheung, Georgios Piliouras
- [On Regret with Multiple Best Arms](#) Yinglun Zhu, Robert Nowak
- [Matrix Completion with Hierarchical Graph Side Information](#) Adel Elmahdy, Junhyung Ahn, Changho Suh, Soheil Mohajer
- [Is Long Horizon RL More Difficult Than Short Horizon RL?](#) Ruosong Wang, Simon S. Du, Lin Yang, Sham Kakade
- [Hamiltonian Monte Carlo using an adjoint-differentiated Laplace approximation: Bayesian inference for latent Gaussian models and beyond](#) Charles Margossian, Aki Vehtari, Daniel Simpson, Raj Agrawal
- [Adversarial Learning for Robust Deep Clustering](#) Xu Yang, Cheng Deng, Kun Wei, Junchi Yan, Wei Liu
- [Learning Mutational Semantics](#) Brian Hie, Ellen Zhong, Bryan Bryson, Bonnie Berger
- [Learning to Learn Variational Semantic Memory](#) Xiantong Zhen, Yingjun Du, Huan Xiong, Qiang Qiu, Cees Snoek, Ling Shao
- [Myersonian Regression](#) Allen Liu, Renato Leme, Jon Schneider
- [Learnability with Indirect Supervision Signals](#) Kaifu Wang, Qiang Ning, Dan Roth
- [Towards Safe Policy Improvement for Non-Stationary MDPs](#) Yash Chandak, Scott Jordan, Georgios Theodorou, Martha White, Philip S. Thomas
- [Finer Metagenomic Reconstruction via Biodiversity Optimization](#) Simon Foucart, David Koslicki
- [Causal Discovery in Physical Systems from Videos](#) Yunzhu Li, Antonio Torralba, Anima Anandkumar, Dieter Fox, Animesh Garg
- [Glyph: Fast and Accurately Training Deep Neural Networks on Encrypted Data](#) Qian Lou, Bo Feng, Geoffrey Charles Fox, Lei Jiang
- [Smoothed Analysis of Online and Differentially Private Learning](#) Nika Haghtalab, Tim Roughgarden, Abhishek Shetty
- [Self-Paced Deep Reinforcement Learning](#) Pascal Klink, Carlo D'Eramo, Jan R. Peters, Joni Pajarinen
- [Kalman Filtering Attention for User Behavior Modeling in CTR Prediction](#) Hu Liu, Jing LU, Xiwei Zhao, Sulong Xu, Hao Peng, Yutong Liu, Zehua Zhang, Jian Li, Junsheng Jin, Yongjun Bao, Weipeng Yan
- [Towards Maximizing the Representation Gap between In-Domain & Out-of-Distribution Examples](#) Jay Nandy, Wynne Hsu, Mong Li Lee
- [Fully Convolutional Mesh Autoencoder using Efficient Spatially Varying Kernels](#) Yi Zhou, Chenglei Wu, Zimo Li, Chen Cao, Yuting Ye, Jason Saragih, Hao Li, Yaser Sheikh
- [GNNGuard: Defending Graph Neural Networks against Adversarial Attacks](#) Xiang Zhang, Marinka Zitnik
- [Geo-PIFu: Geometry and Pixel Aligned Implicit Functions for Single-view Human Reconstruction](#) Tong He, John Collomosse, Hailin Jin, Stefano Soatto
- [Optimal visual search based on a model of target detectability in natural images](#) Shima Rashidi, Krista Ehinger, Andrew Turpin, Lars Kulik
- [Towards Convergence Rate Analysis of Random Forests for Classification](#) Wei Gao, Zhi-Hua Zhou
- [List-Decodable Mean Estimation via Iterative Multi-Filtering](#) Ilias Diakonikolas, Daniel Kane, Daniel Kongsgaard
- [Exact Recovery of Mangled Clusters with Same-Cluster Queries](#) Marco Bressan, Nicolò Cesa-Bianchi, Silvio Lattanzi, Andrea Paudice
- [Steady State Analysis of Episodic Reinforcement Learning](#) Huang Bojun
- [Direct Feedback Alignment Scales to Modern Deep Learning Tasks and Architectures](#) Julien Launay, Iacopo Poli, François Boniface, Florent Krzakala

- [Bayesian Optimization for Iterative Learning](#) *Vu Nguyen, Sebastian Schulze, Michael Osborne*
- [Minimax Bounds for Generalized Linear Models](#) *Kuan-Yun Lee, Thomas Courtade*
- [Projection Robust Wasserstein Distance and Riemannian Optimization](#) *Tianyi Lin, Chenyou Fan, Nhat Ho, Marco Cuturi, Michael Jordan*
- [CoinDICE: Off-Policy Confidence Interval Estimation](#) *Bo Dai, Ofir Nachum, Yinlam Chow, Lihong Li, Csaba Szepesvari, Dale Schuurmans*
- [Simple and Fast Algorithm for Binary Integer and Online Linear Programming](#) *Xiaocheng Li, Chunlin Sun, Yinyu Ye*
- [Learning Diverse and Discriminative Representations via the Principle of Maximal Coding Rate Reduction](#) *Yaodong Yu, Kwan Ho Ryan Chan, Chong You, Chaobing Song, Yi Ma*
- [Learning Rich Rankings](#) *Arjun Seshadri, Stephen Ragain, Johan Ugander*
- [Color Visual Illusions: A Statistics-based Computational Model](#) *Elad Hirsch, Ayellet Tal*
- [Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks](#) *Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, Sebastian Riedel, Douwe Kiela*
- [Universal guarantees for decision tree induction via a higher-order splitting criterion](#) *Guy Blanc, Neha Gupta, Jane Lange, Li-Yang Tan*
- [Trade-offs and Guarantees of Adversarial Representation Learning for Information Obfuscation](#) *Han Zhao, Jianfeng Chi, Yuan Tian, Geoffrey J. Gordon*
- [A Boolean Task Algebra for Reinforcement Learning](#) *Geraud Nangue Tasse, Steven James, Benjamin Rosman*
- [Learning with Differentiable Perturbed Optimizers](#) *Quentin Berthet, Mathieu Blondel, Olivier Teboul, Marco Cuturi, Jean-Philippe Vert, Francis Bach*
- [Optimal Learning from Verified Training Data](#) *Nicholas Bishop, Long Tran-Thanh, Enrico Gerding*
- [Online Linear Optimization with Many Hints](#) *Aditya Bhaskara, Ashok Cutkosky, Ravi Kumar, Manish Purohit*
- [Dynamical mean-field theory for stochastic gradient descent in Gaussian mixture classification](#) *Francesca Mignacco, Florent Krzakala, Pierfrancesco Urbani, Lenka Zdeborová*
- [Causal Discovery from Soft Interventions with Unknown Targets: Characterization and Learning](#) *Amin Jaber, Murat Kocaoglu, Karthikeyan Shanmugam, Elias Bareinboim*
- [Exploiting the Surrogate Gap in Online Multiclass Classification](#) *Dirk van der Hoeven*
- [The Pitfalls of Simplicity Bias in Neural Networks](#) *Harshay Shah, Kaustav Tamuly, Aditi Raghunathan, Prateek Jain, Praneeth Netrapalli*
- [Automatically Learning Compact Quality-aware Surrogates for Optimization Problems](#) *Kai Wang, Bryan Wilder, Andrew Perrault, Milind Tambe*
- [Empirical Likelihood for Contextual Bandits](#) *Nikos Karampatziakis, John Langford, Paul Mineiro*
- [Can Q-Learning with Graph Networks Learn a Generalizable Branching Heuristic for a SAT Solver?](#) *Vitaly Kurin, Saad Godil, Shimon Whiteson, Bryan Catanzaro*
- [Non-reversible Gaussian processes for identifying latent dynamical structure in neural data](#) *Virginia Rutten, Alberto Bernacchia, Maneesh Sahani, Guillaume Hennequin*
- [Listening to Sounds of Silence for Speech Denoising](#) *Ruilin Xu, Rundi Wu, Yuko Ishiwaka, Carl Vondrick, Changxi Zheng*
- [BoxE: A Box Embedding Model for Knowledge Base Completion](#) *Ralph Abboud, Ismail Ceylan, Thomas Lukasiewicz, Tommaso Salvatori*
- [Coherent Hierarchical Multi-Label Classification Networks](#) *Eleonora Giunchiglia, Thomas Lukasiewicz*

- [Walsh-Hadamard Variational Inference for Bayesian Deep Learning](#) *Simone Rossi, Sebastien Marmin, Maurizio Filippone*
- [Federated Bayesian Optimization via Thompson Sampling](#) *Zhongxiang Dai, Bryan Kian Hsiang Low, Patrick Jaillet*
- [MultiON: Benchmarking Semantic Map Memory using Multi-Object Navigation](#) *Saim Wani, Shivansh Patel, Unnat Jain, Angel Chang, Manolis Savva*
- [Neural Complexity Measures](#) *Yoonho Lee, Juho Lee, Sung Ju Hwang, Eunho Yang, Seungjin Choi*
- [Optimal Iterative Sketching Methods with the Subsampled Randomized Hadamard Transform](#) *Jonathan Lacotte, Sifan Liu, Edgar Dobriban, Mert Pilanci*
- [Provably adaptive reinforcement learning in metric spaces](#) *Tongyi Cao, Akshay Krishnamurthy*
- [ShapeFlow: Learnable Deformation Flows Among 3D Shapes](#) *Chiyu Jiang, Jingwei Huang, Andrea Tagliasacchi, Leonidas J. Guibas*
- [Self-Supervised Learning by Cross-Modal Audio-Video Clustering](#) *Humam Alwassel, Dhruv Mahajan, Bruno Korbar, Lorenzo Torresani, Bernard Ghanem, Du Tran*
- [Optimal Query Complexity of Secure Stochastic Convex Optimization](#) *Wei Tang, Chien-Ju Ho, Yang Liu*
- [DynaBERT: Dynamic BERT with Adaptive Width and Depth](#) *Lu Hou, Zhiqi Huang, Lifeng Shang, Xin Jiang, Xiao Chen, Qun Liu*
- [Generalization Bound of Gradient Descent for Non-Convex Metric Learning](#) *MINGZHI DONG, Xiaochen Yang, Rui Zhu, Yujiang Wang, Jing-Hao Xue*
- [Dynamic Submodular Maximization](#) *Morteza Monemizadeh*
- [Inference for Batched Bandits](#) *Kelly Zhang, Lucas Janson, Susan Murphy*
- [Approximate Cross-Validation with Low-Rank Data in High Dimensions](#) *Will Stephenson, Madeleine Udell, Tamara Broderick*
- [GANSpace: Discovering Interpretable GAN Controls](#) *Erik Härkönen, Aaron Hertzmann, Jaakko Lehtinen, Sylvain Paris*
- [Differentiable Expected Hypervolume Improvement for Parallel Multi-Objective Bayesian Optimization](#) *Samuel Daulton, Maximilian Balandat, Eytan Bakshy*
- [Neuron-level Structured Pruning using Polarization Regularizer](#) *Tao Zhuang, Zhixuan Zhang, Yuheng Huang, Xiaoyi Zeng, Kai Shuang, Xiang Li*
- [Limits on Testing Structural Changes in Ising Models](#) *Aditya Gangrade, Bobak Nazer, Venkatesh Saligrama*
- [Field-wise Learning for Multi-field Categorical Data](#) *Zhibin Li, Jian Zhang, Yongshun Gong, Yazhou Yao, Qiang Wu*
- [Continual Learning in Low-rank Orthogonal Subspaces](#) *Arslan Chaudhry, Naeemullah Khan, Puneet Dokania, Philip Torr*
- [Unsupervised Learning of Visual Features by Contrasting Cluster Assignments](#) *Mathilde Caron, Ishan Misra, Julien Mairal, Priya Goyal, Piotr Bojanowski, Armand Joulin*
- [Sharpened Generalization Bounds based on Conditional Mutual Information and an Application to Noisy, Iterative Algorithms](#) *Mahdi Haghifam, Jeffrey Negrea, Ashish Khisti, Daniel M. Roy, Gintare Karolina Dziugaite*
- [Learning Deformable Tetrahedral Meshes for 3D Reconstruction](#) *Jun Gao, Wenzheng Chen, Tommy Xiang, Alec Jacobson, Morgan McGuire, Sanja Fidler*
- [Information theoretic limits of learning a sparse rule](#) *Clément Luneau, Jean Barbier, Nicolas Macris*
- [Self-supervised learning through the eyes of a child](#) *Emin Orhan, Vaibhav Gupta, Brenden M. Lake*
- [Unsupervised Semantic Aggregation and Deformable Template Matching for Semi-Supervised Learning](#) *Tao Han, Junyu Gao, Yuan Yuan, Qi Wang*

- [A game-theoretic analysis of networked system control for common-pool resource management using multi-agent reinforcement learning](#) *Arnu Pretorius, Scott Cameron, Elan van Biljon, Thomas Makkink, Shahil Mawjee, Jeremy du Plessis, Jonathan Shock, Alexandre Laterre, Karim Beguir*
- [What shapes feature representations? Exploring datasets, architectures, and training](#) *Katherine Hermann, Andrew Lampinen*
- [Optimal Best-arm Identification in Linear Bandits](#) *Yassir Jedra, Alexandre Proutiere*
- [Data Diversification: A Simple Strategy For Neural Machine Translation](#) *Xuan-Phi Nguyen, Shafiq Joty, Kui Wu, Ai Ti Aw*
- [Interstellar: Searching Recurrent Architecture for Knowledge Graph Embedding](#) *Yongqi Zhang, Quanming Yao, Lei Chen*
- [CoSE: Compositional Stroke Embeddings](#) *Emre Aksan, Thomas Deselaers, Andrea Tagliasacchi, Otmar Hilliges*
- [Learning Multi-Agent Coordination for Enhancing Target Coverage in Directional Sensor Networks](#) *Jing Xu, Fangwei Zhong, Yizhou Wang*
- [Biological credit assignment through dynamic inversion of feedforward networks](#) *Bill Podlaski, Christian K. Machens*
- [Discriminative Sounding Objects Localization via Self-supervised Audiovisual Matching](#) *Di Hu, Rui Qian, Minyue Jiang, Xiao Tan, Shilei Wen, Errui Ding, Weiyao Lin, Dejing Dou*
- [Learning Multi-Agent Communication through Structured Attentive Reasoning](#) *Murtaza Rangwala, Ryan Williams*
- [Private Identity Testing for High-Dimensional Distributions](#) *Clément L. Canonne, Gautam Kamath, Audra McMillan, Jonathan Ullman, Lydia Zakynthinou*
- [On the Optimal Weighted \$\ell_2\ell_2\$ Regularization in Overparameterized Linear Regression](#) *Denny Wu, Ji Xu*
- [An Efficient Asynchronous Method for Integrating Evolutionary and Gradient-based Policy Search](#) *Kyunghyun Lee, Byeong-Uk Lee, Ukcheol Shin, In So Kweon*
- [MetaSDF: Meta-Learning Signed Distance Functions](#) *Vincent Sitzmann, Eric Chan, Richard Tucker, Noah Snively, Gordon Wetzstein*
- [Simple and Scalable Sparse k-means Clustering via Feature Ranking](#) *Zhiyue Zhang, Kenneth Lange, Jason Xu*
- [Model-based Adversarial Meta-Reinforcement Learning](#) *Zichuan Lin, Garrett Thomas, Guangwen Yang, Tengyu Ma*
- [Graph Policy Network for Transferable Active Learning on Graphs](#) *Shengding Hu, Zheng Xiong, Meng Qu, Xingdi Yuan, Marc-Alexandre Côté, Zhiyuan Liu, Jian Tang*
- [Towards a Better Global Loss Landscape of GANs](#) *Ruoyu Sun, Tiantian Fang, Alexander Schwing*
- [Weighted QMIX: Expanding Monotonic Value Function Factorisation for Deep Multi-Agent Reinforcement Learning](#) *Tabish Rashid, Gregory Farquhar, Bei Peng, Shimon Whiteson*
- [BanditPAM: Almost Linear Time k-Medoids Clustering via Multi-Armed Bandits](#) *Mo Tiwari, Martin J. Zhang, James Mayclin, Sebastian Thrun, Chris Piech, Ilan Shomorony*
- [UDH: Universal Deep Hiding for Steganography, Watermarking, and Light Field Messaging](#) *Chaoning Zhang, Philipp Benz, Adil Karjauv, Geng Sun, In So Kweon*
- [Evidential Sparsification of Multimodal Latent Spaces in Conditional Variational Autoencoders](#) *Masha Itkina, Boris Ivanovic, Ransalu Senanayake, Mykel J. Kochenderfer, Marco Pavone*
- [An Unbiased Risk Estimator for Learning with Augmented Classes](#) *Yu-Jie Zhang, Peng Zhao, Lanjihong Ma, Zhi-Hua Zhou*

- AutoBSS: An Efficient Algorithm for Block Stacking Style Search *Yikang Zhang, Jian Zhang, Zhao Zhong*
- Pushing the Limits of Narrow Precision Inferencing at Cloud Scale with Microsoft Floating Point *Bitá Darvish Rouhani, Daniel Lo, Ritchie Zhao, Ming Liu, Jeremy Fowers, Kalin Ovtcharov, Anna Vinogradsky, Sarah Massengill, Lita Yang, Ray Bittner, Alessandro Forin, Haishan Zhu, Taesik Na, Prerak Patel, Shuai Che, Lok Chand Koppaka, XIA SONG, Subhojit Som, Kaustav Das, Saurabh T, Steve Reinhardt, Sitaram Lanka, Eric Chung, Doug Burger*
- Stochastic Optimization with Laggard Data Pipelines *Naman Agarwal, Rohan Anil, Tomer Koren, Kunal Talwar, Cyril Zhang*
- Self-supervised Auxiliary Learning with Meta-paths for Heterogeneous Graphs *Dasol Hwang, Jinyoung Park, Sunyoung Kwon, KyungMin Kim, Jung-Woo Ha, Hyunwoo J. Kim*
- GPS-Net: Graph-based Photometric Stereo Network *Zhuokun Yao, Kun Li, Ying Fu, Haofeng Hu, Boxin Shi*
- Consistent Structural Relation Learning for Zero-Shot Segmentation *Peike Li, Yunchao Wei, Yi Yang*
- Model Selection in Contextual Stochastic Bandit Problems *Aldo Pacchiano, My Phan, Yasin Abbasi Yadkori, Anup Rao, Julian Zimmert, Tor Lattimore, Csaba Szepesvari*
- Truncated Linear Regression in High Dimensions *Constantinos Daskalakis, Dhruv Rohatgi, Emmanouil Zampetakis*
- Incorporating Pragmatic Reasoning Communication into Emergent Language *Yipeng Kang, Tonghan Wang, Gerard de Melo*
- Deep Subspace Clustering with Data Augmentation *Mahdi Abavisani, Alireza Naghizadeh, Dimitris Metaxas, Vishal Patel*
- An Empirical Process Approach to the Union Bound: Practical Algorithms for Combinatorial and Linear Bandits *Julian Katz-Samuels, Lalit Jain, zohar karnin, Kevin G. Jamieson*
- Can Graph Neural Networks Count Substructures? *Zhengdao Chen, Lei Chen, Soledad Villar, Joan Bruna*
- A Bayesian Perspective on Training Speed and Model Selection *Clare Lyle, Lisa Schut, Robin Ru, Yarin Gal, Mark van der Wilk*
- On the Modularity of Hypernetworks *Tomer Galanti, Lior Wolf*
- Doubly Robust Off-Policy Value and Gradient Estimation for Deterministic Policies *Nathan Kallus, Masatoshi Uehara*
- Provably Efficient Neural GTD for Off-Policy Learning *Hoi-To Wai, Zhuoran Yang, Zhaoran Wang, Mingyi Hong*
- Learning Discrete Energy-based Models via Auxiliary-variable Local Exploration *HanJun Dai, Rishabh Singh, Bo Dai, Charles Sutton, Dale Schuurmans*
- Stable and expressive recurrent vision models *Drew Linsley, Alekh Karkada Ashok, Lakshmi Narasimhan Govindarajan, Rex Liu, Thomas Serre*
- Entropic Optimal Transport between Unbalanced Gaussian Measures has a Closed Form *Hicham Janati, Boris Muzellec, Gabriel Peyré, Marco Cuturi*
- BRP-NAS: Prediction-based NAS using GCNs *Lukasz Dudziak, Thomas Chau, Mohamed Abdelfattah, Royson Lee, Hyeji Kim, Nicholas Lane*
- Deep Shells: Unsupervised Shape Correspondence with Optimal Transport *Marvin Eisenberger, Aysim Toker, Laura Leal-Taixé, Daniel Cremers*
- ISTA-NAS: Efficient and Consistent Neural Architecture Search by Sparse Coding *Yibo Yang, Hongyang Li, Shan You, Fei Wang, Chen Qian, Zhouchen Lin*
- Rel3D: A Minimally Contrastive Benchmark for Grounding Spatial Relations in 3D *Ankit Goyal, Kaiyu Yang, Dawei Yang, Jia Deng*

- [Regularizing Black-box Models for Improved Interpretability](#) Gregory Plumb, Maruan Al-Shedivat, Ángel Alexander Cabrera, Adam Perer, Eric Xing, Ameet Talwalkar
- [Trust the Model When It Is Confident: Masked Model-based Actor-Critic](#) Feiyang Pan, Jia He, Dandan Tu, Qing He
- [Semi-Supervised Neural Architecture Search](#) Renqian Luo, Xu Tan, Rui Wang, Tao Qin, Enhong Chen, Tie-Yan Liu
- [Consistency Regularization for Certified Robustness of Smoothed Classifiers](#) Jongheon Jeong, Jinwoo Shin
- [Robust Multi-Agent Reinforcement Learning with Model Uncertainty](#) Kaiqing Zhang, TAO SUN, Yunzhe Tao, Sahika Genc, Sunil Mallya, Tamer Basar
- [SIRI: Spatial Relation Induced Network For Spatial Description Resolution](#) peiyao wang, Weixin Luo, Yanyu Xu, Haojie Li, Shugong Xu, Jianyu Yang, Shenghua Gao
- [Adaptive Shrinkage Estimation for Streaming Graphs](#) Nesreen Ahmed, Nick Duffield
- [Make One-Shot Video Object Segmentation Efficient Again](#) Tim Meinhardt, Laura Leal-Taixé
- [Depth Uncertainty in Neural Networks](#) Javier Antoran, James Allingham, José Miguel Hernández-Lobato
- [Non-Euclidean Universal Approximation](#) Anastasis Kratsios, Ievgen Bilokopytov
- [Constraining Variational Inference with Geometric Jensen-Shannon Divergence](#) Jacob Deasy, Nikola Simidjievski, Pietro Lió
- [Gibbs Sampling with People](#) Peter Harrison, Raja Marjeh, Federico Adolfi, Pol van Rijn, Manuel Anglada-Tort, Ofer Tchernichovski, Pauline Larrouy-Maestri, Nori Jacoby
- [HM-ANN: Efficient Billion-Point Nearest Neighbor Search on Heterogeneous Memory](#) Jie Ren, Minjia Zhang, Dong Li
- [FrugalML: How to use ML Prediction APIs more accurately and cheaply](#) Lingjiao Chen, Matei Zaharia, James Y. Zou
- [Sharp Representation Theorems for ReLU Networks with Precise Dependence on Depth](#) Guy Bresler, Dheeraj Nagaraj
- [Shared Experience Actor-Critic for Multi-Agent Reinforcement Learning](#) Filippos Christianos, Lukas Schäfer, Stefano Albrecht
- [Monotone operator equilibrium networks](#) Ezra Winston, J. Zico Kolter
- [When and How to Lift the Lockdown? Global COVID-19 Scenario Analysis and Policy Assessment using Compartmental Gaussian Processes](#) Zhaozhi Qian, Ahmed M. Alaa, Mihaela van der Schaar
- [Unsupervised Learning of Lagrangian Dynamics from Images for Prediction and Control](#) Yaofeng Desmond Zhong, Naomi Leonard
- [High-Dimensional Sparse Linear Bandits](#) Botao Hao, Tor Lattimore, Mengdi Wang
- [Non-Stochastic Control with Bandit Feedback](#) Paula Gradu, John Hallman, Elad Hazan
- [Generalized Leverage Score Sampling for Neural Networks](#) Jason D. Lee, Ruoqi Shen, Zhao Song, Mengdi Wang, Zheng Yu
- [An Optimal Elimination Algorithm for Learning a Best Arm](#) Avinatan Hassidim, Ron Kupfer, Yaron Singer
- [Efficient Projection-free Algorithms for Saddle Point Problems](#) Cheng Chen, Luo Luo, Weinan Zhang, Yong Yu
- [A mathematical model for automatic differentiation in machine learning](#) Jérôme Bolte, Edouard Pauwels
- [Unsupervised Text Generation by Learning from Search](#) Jingjing Li, Zichao Li, Lili Mou, Xin Jiang, Michael Lyu, Irwin King
- [Learning Compositional Rules via Neural Program Synthesis](#) Maxwell Nye, Armando Solar-Lezama, Josh Tenenbaum, Brenden M. Lake

- [Incorporating BERT into Parallel Sequence Decoding with Adapters](#) Junliang Guo, Zhirui Zhang, Linli Xu, Hao-Ran Wei, Boxing Chen, Enhong Chen
- [Estimating Fluctuations in Neural Representations of Uncertain Environments](#) Sahand Farhoodi, Mark Plitt, Lisa Giocomo, Uri Eden
- [Discover, Hallucinate, and Adapt: Open Compound Domain Adaptation for Semantic Segmentation](#) KwanYong Park, Sanghyun Woo, Inkyu Shin, In So Kweon
- [SURF: A Simple, Universal, Robust, Fast Distribution Learning Algorithm](#) Yi Hao, Ayush Jain, Alon Orlitsky, Vaishakh Ravindrakumar
- [Understanding Approximate Fisher Information for Fast Convergence of Natural Gradient Descent in Wide Neural Networks](#) Ryo Karakida, Kazuki Osawa
- [General Transportability of Soft Interventions: Completeness Results](#) Juan Correa, Elias Bareinboim
- [GALT-prop: A biologically plausible learning rule derived from backpropagation of error](#) Nasir Ahmad, Marcel A. J. van Gerven, Luca Ambrogioni
- [Lipschitz Bounds and Provably Robust Training by Laplacian Smoothing](#) Vishaal Krishnan, Abed AlRahman Al Makdah, Fabio Pasqualetti
- [SCOP: Scientific Control for Reliable Neural Network Pruning](#) Yehui Tang, Yunhe Wang, Yixing Xu, Dacheng Tao, Chunjing XU, Chao Xu, Chang Xu
- [Provably Consistent Partial-Label Learning](#) Lei Feng, Jiaqi Lv, Bo Han, Miao Xu, Gang Niu, Xin Geng, Bo An, Masashi Sugiyama
- [Robust, Accurate Stochastic Optimization for Variational Inference](#) Akash Kumar Dhaka, Alejandro Catalina, Michael R. Andersen, Måns Magnusson, Jonathan Huggins, Aki Vehtari
- [Discovering conflicting groups in signed networks](#) Ruo-Chun Tzeng, Bruno Ordozgoiti, Aristides Gionis
- [Learning Some Popular Gaussian Graphical Models without Condition Number Bounds](#) Jonathan Kelner, Frederic Koehler, Raghu Meka, Ankur Moitra
- [Sense and Sensitivity Analysis: Simple Post-Hoc Analysis of Bias Due to Unobserved Confounding](#) Victor Veitch, Anisha Zaveri
- [Mix and Match: An Optimistic Tree-Search Approach for Learning Models from Mixture Distributions](#) Matthew Faw, Rajat Sen, Karthikeyan Shanmugam, Constantine Caramanis, Sanjay Shakkottai
- [Understanding Double Descent Requires A Fine-Grained Bias-Variance Decomposition](#) Ben Adlam, Jeffrey Pennington
- [VIME: Extending the Success of Self- and Semi-supervised Learning to Tabular Domain](#) Jinsung Yoon, Yao Zhang, James Jordon, Mihaela van der Schaar
- [The Smoothed Possibility of Social Choice](#) Lirong Xia
- [A Decentralized Parallel Algorithm for Training Generative Adversarial Nets](#) Mingrui Liu, Wei Zhang, Youssef Mroueh, Xiaodong Cui, Jarret Ross, Tianbao Yang, Payel Das
- [Phase retrieval in high dimensions: Statistical and computational phase transitions](#) Antoine Maillard, Bruno Loureiro, Florent Krzakala, Lenka Zdeborová
- [Fair Performance Metric Elicitation](#) Gaurush Hiranandani, Harikrishna Narasimhan, Sanmi Koyejo
- [Hybrid Variance-Reduced SGD Algorithms For Minimax Problems with Nonconvex-Linear Function](#) Quoc Tran Dinh, Deyi Liu, Lam Nguyen
- [Belief-Dependent Macro-Action Discovery in POMDPs using the Value of Information](#) Genevieve Flaspohler, Nicholas A. Roy, John W. Fisher III
- [Soft Contrastive Learning for Visual Localization](#) Janine Thoma, Danda Pani Paudel, Luc V. Gool
- [Fine-Grained Dynamic Head for Object Detection](#) Lin Song, Yanwei Li, Zhengkai Jiang, Zeming Li, Hongbin

Sun, Jian Sun, Nanning Zheng

- [LoCo: Local Contrastive Representation Learning](#) *Yuwen Xiong, Mengye Ren, Raquel Urtasun*
- [Modeling and Optimization Trade-off in Meta-learning](#) *Katelyn Gao, Ozan Sener*
- [SnapBoost: A Heterogeneous Boosting Machine](#) *Thomas Parnell, Andreea Anghel, Małgorzata Łazuka, Nikolas Ioannou, Sebastian Kurella, Peshal Agarwal, Nikolaos Papandreou, Haralampos Pozidis*
- [On Adaptive Distance Estimation](#) *Yeshwanth Cherapanamjeri, Jelani Nelson*
- [Stage-wise Conservative Linear Bandits](#) *Ahmadreza Moradipari, Christos Thrampoulidis, Mahnoosh Alizadeh*
- [RELATE: Physically Plausible Multi-Object Scene Synthesis Using Structured Latent Spaces](#) *Sebastien Ehrhardt, Oliver Groth, Aron Monszpart, Martin Engelcke, Ingmar Posner, Niloy Mitra, Andrea Vedaldi*
- [Metric-Free Individual Fairness in Online Learning](#) *Yahav Bechavod, Christopher Jung, Steven Z. Wu*
- [GreedyFool: Distortion-Aware Sparse Adversarial Attack](#) *Xiaoyi Dong, Dongdong Chen, Jianmin Bao, Chuan Qin, Lu Yuan, Weiming Zhang, Nenghai Yu, Dong Chen*
- [VAEM: a Deep Generative Model for Heterogeneous Mixed Type Data](#) *Chao Ma, Sebastian Tschitschek, Richard Turner, José Miguel Hernández-Lobato, Cheng Zhang*
- [RetroXpert: Decompose Retrosynthesis Prediction Like A Chemist](#) *Chaochao Yan, Qianggang Ding, Peilin Zhao, Shuangjia Zheng, JINYU YANG, Yang Yu, Junzhou Huang*
- [Sample-Efficient Optimization in the Latent Space of Deep Generative Models via Weighted Retraining](#) *Austin Tripp, Erik Daxberger, José Miguel Hernández-Lobato*
- [Improved Sample Complexity for Incremental Autonomous Exploration in MDPs](#) *Jean Tarbouriech, Matteo Pirota, Michal Valko, Alessandro Lazaric*
- [TinyTL: Reduce Memory, Not Parameters for Efficient On-Device Learning](#) *Han Cai, Chuang Gan, Ligeng Zhu, Song Han*
- [RD22: Reward Decomposition with Representation Decomposition](#) *Zichuan Lin, Derek Yang, Li Zhao, Tao Qin, Guangwen Yang, Tie-Yan Liu*
- [Self-paced Contrastive Learning with Hybrid Memory for Domain Adaptive Object Re-ID](#) *Yixiao Ge, Feng Zhu, Dapeng Chen, Rui Zhao, hongsheng Li*
- [Fairness constraints can help exact inference in structured prediction](#) *Kevin Bello, Jean Honorio*
- [Instance-based Generalization in Reinforcement Learning](#) *Martin Bertran, Natalia Martinez, Mariano Phielipp, Guillermo Sapiro*
- [Smooth And Consistent Probabilistic Regression Trees](#) *Sami Alkhoury, Emilie Devijver, Marianne Clausel, Myriam Tami, Eric Gaussier, georges Oppenheim*
- [Computing Valid p-value for Optimal Changepoint by Selective Inference using Dynamic Programming](#) *Vo Nguyen Le Duy, Hiroki Toda, Ryota Sugiyama, Ichiro Takeuchi*
- [Factorized Neural Processes for Neural Processes: K-Shot Prediction of Neural Responses](#) *Ronald (James) Cotton, Fabian Sinz, Andreas Tolias*
- [Winning the Lottery with Continuous Sparsification](#) *Pedro Savarese, Hugo Silva, Michael Maire*
- [Adversarial robustness via robust low rank representations](#) *Pranjal Awasthi, Himanshu Jain, Ankit Singh Rawat, Aravindan Vijayaraghavan*
- [Joints in Random Forests](#) *Alvaro Correia, Robert Peharz, Cassio P. de Campos*
- [Compositional Generalization by Learning Analytical Expressions](#) *Qian Liu, Shengnan An, Jian-Guang Lou, Bei Chen, Zeqi Lin, Yan Gao, Bin Zhou, Nanning Zheng, Dongmei Zhang*
- [JAX MD: A Framework for Differentiable Physics](#) *Samuel Schoenholz, Ekin Dogus Cubuk*
- [An implicit function learning approach for parametric modal regression](#) *Yangchen Pan, Ehsan Imani, Amir-massoud Farahmand, Martha White*

- [SDF-SRN: Learning Signed Distance 3D Object Reconstruction from Static Images](#) *Chen-Hsuan Lin, Chaoyang Wang, Simon Lucey*
- [Coresets for Robust Training of Deep Neural Networks against Noisy Labels](#) *Baharan Mirzasoleiman, Kaidi Cao, Jure Leskovec*
- [Adapting to Misspecification in Contextual Bandits](#) *Dylan J. Foster, Claudio Gentile, Mehryar Mohri, Julian Zimmert*
- [Convergence of Meta-Learning with Task-Specific Adaptation over Partial Parameters](#) *Kaiyi Ji, Jason D. Lee, Yingbin Liang, H. Vincent Poor*
- [MetaPerturb: Transferable Regularizer for Heterogeneous Tasks and Architectures](#) *Jeong Un Ryu, JaeWoong Shin, Hae Beom Lee, Sung Ju Hwang*
- [Learning to solve TV regularised problems with unrolled algorithms](#) *Hamza Cherkaoui, Jeremias Sulam, Thomas Moreau*
- [Object-Centric Learning with Slot Attention](#) *Francesco Locatello, Dirk Weissenborn, Thomas Unterthiner, Aravindh Mahendran, Georg Heigold, Jakob Uszkoreit, Alexey Dosovitskiy, Thomas Kipf*
- [Improving robustness against common corruptions by covariate shift adaptation](#) *Steffen Schneider, Evgenia Rusak, Luisa Eck, Oliver Bringmann, Wieland Brendel, Matthias Bethge*
- [Deep Smoothing of the Implied Volatility Surface](#) *Damien Ackerer, Natasa Tagasovska, Thibault Vatter*
- [Probabilistic Inference with Algebraic Constraints: Theoretical Limits and Practical Approximations](#) *Zhe Zeng, Paolo Morettin, Fanqi Yan, Antonio Vergari, Guy Van den Broeck*
- [Provable Online CP/PARAFAC Decomposition of a Structured Tensor via Dictionary Learning](#) *Sirisha Rambhatla, Xingguo Li, Jarvis Haupt*
- [Look-ahead Meta Learning for Continual Learning](#) *Gunshi Gupta, Karmesh Yadav, Liam Paull*
- [A polynomial-time algorithm for learning nonparametric causal graphs](#) *Ming Gao, Yi Ding, Bryon Aragam*
- [Sparse Learning with CART](#) *Jason Klusowski*
- [Proximal Mapping for Deep Regularization](#) *Mao Li, Yingyi Ma, Xinhua Zhang*
- [Identifying Causal-Effect Inference Failure with Uncertainty-Aware Models](#) *Andrew Jesson, Sören Mindermann, Uri Shalit, Yarin Gal*
- [Hierarchical Granularity Transfer Learning](#) *Shaobo Min, Hongtao Xie, Hantao Yao, Xuran Deng, Zheng-Jun Zha, Yongdong Zhang*
- [Deep active inference agents using Monte-Carlo methods](#) *Zafeirios Fountas, Noor Sajid, Pedro Mediano, Karl Friston*
- [Consistent Estimation of Identifiable Nonparametric Mixture Models from Grouped Observations](#) *Alexander Ritchie, Robert A. Vandermeulen, Clayton Scott*
- [Manifold structure in graph embeddings](#) *Patrick Rubin-Delanchy*
- [Adaptive Learned Bloom Filter \(Ada-BF\): Efficient Utilization of the Classifier with Application to Real-Time Information Filtering on the Web](#) *Zhenwei Dai, Anshumali Shrivastava*
- [MCUNet: Tiny Deep Learning on IoT Devices](#) *Ji Lin, Wei-Ming Chen, Yujun Lin, John Cohn, Chuang Gan, Song Han*
- [In search of robust measures of generalization](#) *Gintare Karolina Dziugaite, Alexandre Drouin, Brady Neal, Nitish Rajkumar, Ethan Caballero, Linbo Wang, Ioannis Mitliagkas, Daniel M. Roy*
- [Task-agnostic Exploration in Reinforcement Learning](#) *Xuezhou Zhang, Yuzhe Ma, Adish Singla*
- [Multi-task Additive Models for Robust Estimation and Automatic Structure Discovery](#) *Yingjie Wang, Hong Chen, Feng Zheng, Chen Xu, Tieliang Gong, Yanhong Chen*
- [Provably Efficient Reward-Agnostic Navigation with Linear Value Iteration](#) *Andrea Zanette, Alessandro*

Lazaric, Mykel J. Kochenderfer, Emma Brunskill

- Softmax Deep Double Deterministic Policy Gradients *Ling Pan, Qingpeng Cai, Longbo Huang*
- Online Decision Based Visual Tracking via Reinforcement Learning *ke Song, Wei Zhang, Ran Song, Yibin Li*
- Efficient Marginalization of Discrete and Structured Latent Variables via Sparsity *Gonçalo Correia, Vlad Niculae, Wilker Aziz, André Martins*
- DeepI2I: Enabling Deep Hierarchical Image-to-Image Translation by Transferring from GANs *yaxing wang, Lu Yu, Joost van de Weijer*
- Distributional Robustness with IPMs and links to Regularization and GANs *Hisham Husain*
- A shooting formulation of deep learning *François-Xavier Vialard, Roland Kwitt, Susan Wei, Marc Niethammer*
- CSI: Novelty Detection via Contrastive Learning on Distributionally Shifted Instances *Jihoon Tack, Sangwoo Mo, Jongheon Jeong, Jinwoo Shin*
- Learning Implicit Credit Assignment for Cooperative Multi-Agent Reinforcement Learning *Meng Zhou, Ziyu Liu, Pengwei Sui, Yixuan Li, Yuk Ying Chung*
- MATE: Plugging in Model Awareness to Task Embedding for Meta Learning *Xiaohan Chen, Zhangyang Wang, Siyu Tang, Krikamol Muandet*
- Restless-UCB, an Efficient and Low-complexity Algorithm for Online Restless Bandits *Siwei Wang, Longbo Huang, John C. S. Lui*
- Predictive Information Accelerates Learning in RL *Kuang-Huei Lee, Ian Fischer, Anthony Liu, Yijie Guo, Honglak Lee, John Canny, Sergio Guadarrama*
- Robust and Heavy-Tailed Mean Estimation Made Simple, via Regret Minimization *Sam Hopkins, Jerry Li, Fred Zhang*
- High-Fidelity Generative Image Compression *Fabian Mentzer, George D. Toderici, Michael Tschannen, Eirikur Agustsson*
- A Statistical Mechanics Framework for Task-Agnostic Sample Design in Machine Learning *Bhavya Kailkhura, Jayaraman Thiagarajan, Qunwei Li, Jize Zhang, Yi Zhou, Timo Bremer*
- Counterexample-Guided Learning of Monotonic Neural Networks *Aishwarya Sivaraman, Golnoosh Farnadi, Todd Millstein, Guy Van den Broeck*
- A Novel Approach for Constrained Optimization in Graphical Models *Sara Rouhani, Tahrira Rahman, Vibhav Gogate*
- Global Convergence of Deep Networks with One Wide Layer Followed by Pyramidal Topology *Quynh N. Nguyen, Marco Mondelli*
- On the Trade-off between Adversarial and Backdoor Robustness *Cheng-Hsin Weng, Yan-Ting Lee, Shan-Hung (Brandon) Wu*
- Implicit Graph Neural Networks *Fangda Gu, Heng Chang, Wenwu Zhu, Somayeh Sojoudi, Laurent El Ghaoui*
- Rethinking Importance Weighting for Deep Learning under Distribution Shift *Tongtong Fang, Nan Lu, Gang Niu, Masashi Sugiyama*
- Guiding Deep Molecular Optimization with Genetic Exploration *Sungsoo Ahn, Junsu Kim, Hankook Lee, Jinwoo Shin*
- Temporal Spike Sequence Learning via Backpropagation for Deep Spiking Neural Networks *Wenrui Zhang, Peng Li*
- TSPNet: Hierarchical Feature Learning via Temporal Semantic Pyramid for Sign Language Translation *DONGXU LI, Chenchen Xu, Xin Yu, Kaihao Zhang, Benjamin Swift, Hanna Suominen, Hongdong Li*
- Neural Topographic Factor Analysis for fMRI Data *Eli Sennesh, Zulqarnain Khan, Yiyu Wang, J Benjamin Hutchinson, Ajay Satpute, Jennifer Dy, Jan-Willem van de Meent*

- [Neural Architecture Generator Optimization](#) *Robin Ru, Pedro Esperança, Fabio Maria Carlucci*
- [A Bandit Learning Algorithm and Applications to Auction Design](#) *Kim Thang Nguyen*
- [MetaPoison: Practical General-purpose Clean-label Data Poisoning](#) *W. Ronny Huang, Jonas Geiping, Liam Fowl, Gavin Taylor, Tom Goldstein*
- [Sample Efficient Reinforcement Learning via Low-Rank Matrix Estimation](#) *Devavrat Shah, Dogyoon Song, Zhi Xu, Yuzhe Yang*
- [Training Generative Adversarial Networks with Limited Data](#) *Tero Karras, Miika Aittala, Janne Hellsten, Samuli Laine, Jaakko Lehtinen, Timo Aila*
- [Deeply Learned Spectral Total Variation Decomposition](#) *Tamara G. Grossmann, Yury Korolev, Guy Gilboa, Carola Schoenlieb*
- [FracTrain: Fractionally Squeezing Bit Savings Both Temporally and Spatially for Efficient DNN Training](#) *Yonggan Fu, Haoran You, Yang Zhao, Yue Wang, Chaojian Li, Kailash Gopalakrishnan, Zhangyang Wang, Yingyan Lin*
- [Improving Neural Network Training in Low Dimensional Random Bases](#) *Frithjof Gressmann, Zach Eaton-Rosen, Carlo Luschi*
- [Safe Reinforcement Learning via Curriculum Induction](#) *Matteo Turchetta, Andrey Kolobov, Shital Shah, Andreas Krause, Alekh Agarwal*
- [Leverage the Average: an Analysis of KL Regularization in Reinforcement Learning](#) *Nino Vieillard, Tadashi Kozuno, Bruno Scherrer, Olivier Pietquin, Remi Munos, Matthieu Geist*
- [How Robust are the Estimated Effects of Nonpharmaceutical Interventions against COVID-19?](#) *Mrinank Sharma, Sören Mindermann, Jan Brauner, Gavin Leech, Anna Stephenson, Tomáš Gavenčiak, Jan Kulveit, Yee Whye Teh, Leonid Chindelevitch, Yarin Gal*
- [Beyond Individualized Recourse: Interpretable and Interactive Summaries of Actionable Recourses](#) *Kaivalya Rawal, Himabindu Lakkaraju*
- [Generalization error in high-dimensional perceptrons: Approaching Bayes error with convex optimization](#) *Benjamin Aubin, Florent Krzakala, Yue Lu, Lenka Zdeborová*
- [Projection Efficient Subgradient Method and Optimal Nonsmooth Frank-Wolfe Method](#) *Kiran K. Thekumparampil, Prateek Jain, Praneeth Netrapalli, Sewoong Oh*
- [PGM-Explainer: Probabilistic Graphical Model Explanations for Graph Neural Networks](#) *Minh Vu, My T. Thai*
- [Few-Cost Salient Object Detection with Adversarial-Paced Learning](#) *Dingwen Zhang, HaiBin Tian, Jungong Han*
- [Minimax Estimation of Conditional Moment Models](#) *Nishanth Dikkala, Greg Lewis, Lester Mackey, Vasilis Syrgkanis*
- [Causal Imitation Learning With Unobserved Confounders](#) *Junzhe Zhang, Daniel Kumor, Elias Bareinboim*
- [Your GAN is Secretly an Energy-based Model and You Should Use Discriminator Driven Latent Sampling](#) *Tong Che, Ruixiang ZHANG, Jascha Sohl-Dickstein, Hugo Larochelle, Liam Paull, Yuan Cao, Yoshua Bengio*
- [Learning Black-Box Attackers with Transferable Priors and Query Feedback](#) *Jiancheng YANG, Yangzhou Jiang, Xiaoyang Huang, Bingbing Ni, Chenglong Zhao*
- [Locally Differentially Private \(Contextual\) Bandits Learning](#) *Kai Zheng, Tianle Cai, Weiran Huang, Zhenguo Li, Liwei Wang*
- [Invertible Gaussian Reparameterization: Revisiting the Gumbel-Softmax](#) *Andres Potapczynski, Gabriel Loaiza-Ganem, John P. Cunningham*
- [Kernel Based Progressive Distillation for Adder Neural Networks](#) *Yixing Xu, Chang Xu, Xinghao Chen, Wei Zhang, Chunjing XU, Yunhe Wang*

- [Adversarial Soft Advantage Fitting: Imitation Learning without Policy Optimization](#) Paul Barde, Julien Roy, Wonseok Jeon, Joelle Pineau, Chris Pal, Derek Nowrouzezahrai
- [Agree to Disagree: Adaptive Ensemble Knowledge Distillation in Gradient Space](#) Shangchen Du, Shan You, Xiaojie Li, Jianlong Wu, Fei Wang, Chen Qian, Changshui Zhang
- [The Wasserstein Proximal Gradient Algorithm](#) Adil Salim, Anna Korba, Giulia Luise
- [Universally Quantized Neural Compression](#) Eirikur Agustsson, Lucas Theis
- [Temporal Variability in Implicit Online Learning](#) Nicolò Campolongo, Francesco Orabona
- [Investigating Gender Bias in Language Models Using Causal Mediation Analysis](#) Jesse Vig, Sebastian Gehrmann, Yonatan Belinkov, Sharon Qian, Daniel Nevo, Yaron Singer, Stuart Shieber
- [Off-Policy Imitation Learning from Observations](#) Zhuangdi Zhu, Kaixiang Lin, Bo Dai, Jiayu Zhou
- [Escaping Saddle-Point Faster under Interpolation-like Conditions](#) Abhishek Roy, Krishnakumar Balasubramanian, Saeed Ghadimi, Prasant Mohapatra
- [Matérn Gaussian Processes on Riemannian Manifolds](#) Viacheslav Borovitskiy, Alexander Terenin, Peter Mostowsky, Marc Deisenroth (he/him)
- [Improved Techniques for Training Score-Based Generative Models](#) Yang Song, Stefano Ermon
- [wav2vec 2.0: A Framework for Self-Supervised Learning of Speech Representations](#) Alexei Baevski, Yuhao Zhou, Abdelrahman Mohamed, Michael Auli
- [A Maximum-Entropy Approach to Off-Policy Evaluation in Average-Reward MDPs](#) Nevena Lazic, Dong Yin, Mehrdad Farajtabar, Nir Levine, Dilan Gorur, Chris Harris, Dale Schuurmans
- [Instead of Rewriting Foreign Code for Machine Learning, Automatically Synthesize Fast Gradients](#) William Moses, Valentin Churavy
- [Does Unsupervised Architecture Representation Learning Help Neural Architecture Search?](#) Shen Yan, Yu Zheng, Wei Ao, Xiao Zeng, Mi Zhang
- [Value-driven Hindsight Modelling](#) Arthur Guez, Fabio Viola, Theophane Weber, Lars Buesing, Steven Kapturowski, Doina Precup, David Silver, Nicolas Heess
- [Dynamic Regret of Convex and Smooth Functions](#) Peng Zhao, Yu-Jie Zhang, Lijun Zhang, Zhi-Hua Zhou
- [On Convergence of Nearest Neighbor Classifiers over Feature Transformations](#) Luka Rimanic, Cedric Renggli, Bo Li, Ce Zhang
- [Mitigating Manipulation in Peer Review via Randomized Reviewer Assignments](#) Steven Jecmen, Hanrui Zhang, Ryan Liu, Nihar Shah, Vincent Conitzer, Fei Fang
- [Contrastive learning of global and local features for medical image segmentation with limited annotations](#) Krishna Chaitanya, Ertunc Erdil, Neerav Karani, Ender Konukoglu
- [Self-Supervised Graph Transformer on Large-Scale Molecular Data](#) Yu Rong, Yatao Bian, Tingyang Xu, Weiyang Xie, Ying WEI, Wenbing Huang, Junzhou Huang
- [Generative Neurosymbolic Machines](#) Jindong Jiang, Sungjin Ahn
- [How many samples is a good initial point worth in Low-rank Matrix Recovery?](#) Jialun Zhang, Richard Zhang
- [CSER: Communication-efficient SGD with Error Reset](#) Cong Xie, Shuai Zheng, Sanmi Koyejo, Indranil Gupta, Mu Li, Haibin Lin
- [Efficient estimation of neural tuning during naturalistic behavior](#) Edoardo Balzani, Kaushik Lakshminarasimhan, Dora Angelaki, Cristina Savin
- [High-recall causal discovery for autocorrelated time series with latent confounders](#) Andreas Gerhardus, Jakob Runge
- [Forget About the LiDAR: Self-Supervised Depth Estimators with MED Probability Volumes](#) Juan Luis GonzalezBello, Munchurl Kim

- [Joint Contrastive Learning with Infinite Possibilities](#) *Qi Cai, Yu Wang, Yingwei Pan, Ting Yao, Tao Mei*
- [Robust Gaussian Covariance Estimation in Nearly-Matrix Multiplication Time](#) *Jerry Li, Guanghao Ye*
- [Adversarially-learned Inference via an Ensemble of Discrete Undirected Graphical Models](#) *Adarsh Keshav Jeewajee, Leslie Kaelbling*
- [GS-WGAN: A Gradient-Sanitized Approach for Learning Differentially Private Generators](#) *Dingfan Chen, Tribhuvanesh Orekondy, Mario Fritz*
- [SurVAE Flows: Surjections to Bridge the Gap between VAEs and Flows](#) *Didrik Nielsen, Priyank Jaini, Emiel Hoogeboom, Ole Winther, Max Welling*
- [Learning Causal Effects via Weighted Empirical Risk Minimization](#) *Yonghan Jung, Jin Tian, Elias Bareinboim*
- [Revisiting the Sample Complexity of Sparse Spectrum Approximation of Gaussian Processes](#) *Minh Hoang, Nghia Hoang, Hai Pham, David Woodruff*
- [Incorporating Interpretable Output Constraints in Bayesian Neural Networks](#) *Wanqian Yang, Lars Lorch, Moritz Graule, Himabindu Lakkaraju, Finale Doshi-Velez*
- [Multi-Stage Influence Function](#) *Hongge Chen, Si Si, Yang Li, Ciprian Chelba, Sanjiv Kumar, Duane Boning, Cho-Jui Hsieh*
- [Probabilistic Fair Clustering](#) *Seyed Esmaeili, Brian Brubach, Leonidas Tsepenekas, John Dickerson*
- [Stochastic Segmentation Networks: Modelling Spatially Correlated Aleatoric Uncertainty](#) *Miguel Monteiro, Loic Le Folgoc, Daniel Coelho de Castro, Nick Pawlowski, Bernardo Marques, Konstantinos Kamnitsas, Mark van der Wilk, Ben Glocker*
- [ICE-BeeM: Identifiable Conditional Energy-Based Deep Models Based on Nonlinear ICA](#) *Ilyes Khemakhem, Ricardo Monti, Diederik Kingma, Aapo Hyvarinen*
- [Testing Determinantal Point Processes](#) *Khashayar Gatmiry, Maryam Aliakbarpour, Stefanie Jegelka*
- [CogLTX: Applying BERT to Long Texts](#) *Ming Ding, Chang Zhou, Hongxia Yang, Jie Tang*
- [f-GAIL: Learning f-Divergence for Generative Adversarial Imitation Learning](#) *Xin Zhang, Yanhua Li, Ziming Zhang, Zhi-Li Zhang*
- [Non-parametric Models for Non-negative Functions](#) *Ulysse Marteau-Ferey, Francis Bach, Alessandro Rudi*
- [Uncertainty Aware Semi-Supervised Learning on Graph Data](#) *Xujiang Zhao, Feng Chen, Shu Hu, Jin-Hee Cho*
- [ConvBERT: Improving BERT with Span-based Dynamic Convolution](#) *Zi-Hang Jiang, Weihao Yu, Daquan Zhou, Yunpeng Chen, Jiashi Feng, Shuicheng Yan*
- [Practical No-box Adversarial Attacks against DNNs](#) *Qizhang Li, Yiwen Guo, Hao Chen*
- [Breaking the Sample Size Barrier in Model-Based Reinforcement Learning with a Generative Model](#) *Gen Li, Yuting Wei, Yuejie Chi, Yuantao Gu, Yuxin Chen*
- [Walking in the Shadow: A New Perspective on Descent Directions for Constrained Minimization](#) *Hassan Mortagy, Swati Gupta, Sebastian Pokutta*
- [Path Sample-Analytic Gradient Estimators for Stochastic Binary Networks](#) *Alexander Shekhovtsov, Viktor Yanush, Boris Flach*
- [Reward Propagation Using Graph Convolutional Networks](#) *Martin Klissarov, Doina Precup*
- [LoopReg: Self-supervised Learning of Implicit Surface Correspondences, Pose and Shape for 3D Human Mesh Registration](#) *Bharat Lal Bhatnagar, Cristian Sminchisescu, Christian Theobalt, Gerard Pons-Moll*
- [Fully Dynamic Algorithm for Constrained Submodular Optimization](#) *Silvio Lattanzi, Slobodan Mitrović, Ashkan Norouzi-Fard, Jakub M. Tarnawski, Morteza Zadimoghaddam*
- [Robust Optimal Transport with Applications in Generative Modeling and Domain Adaptation](#) *Yogesh Balaji, Rama Chellappa, Soheil Feizi*

- [Autofocused oracles for model-based design](#) *Clara Fannjiang, Jennifer Listgarten*
- [Debiasing Averaged Stochastic Gradient Descent to handle missing values](#) *Aude Sportisse, Claire Boyer, Aymeric Dieuleveut, Julie Josse*
- [Trajectory-wise Multiple Choice Learning for Dynamics Generalization in Reinforcement Learning](#) *Younggyo Seo, Kimin Lee, Ignasi Clavera Gilaberte, Thanard Kurutach, Jinwoo Shin, Pieter Abbeel*
- [CompPress: Self-Supervised Learning by Compressing Representations](#) *Soroush Abbasi Koochpayegani, Ajinkya Tejankar, Hamed Pirsiavash*
- [Sample complexity and effective dimension for regression on manifolds](#) *Andrew McRae, Justin Romberg, Mark Davenport*
- [The phase diagram of approximation rates for deep neural networks](#) *Dmitry Yarotsky, Anton Zhevnerchuk*
- [Timeseries Anomaly Detection using Temporal Hierarchical One-Class Network](#) *Lifeng Shen, Zhuocong Li, James Kwok*
- [EcoLight: Intersection Control in Developing Regions Under Extreme Budget and Network Constraints](#) *Sachin Chauhan, Kashish Bansal, Rijurekha Sen*
- [Reconstructing Perceptive Images from Brain Activity by Shape-Semantic GAN](#) *Tao Fang, Yu Qi, Gang Pan*
- [Emergent Complexity and Zero-shot Transfer via Unsupervised Environment Design](#) *Michael Dennis, Natasha Jaques, Eugene Vinitzky, Alexandre Bayen, Stuart Russell, Andrew Critch, Sergey Levine*
- [A Spectral Energy Distance for Parallel Speech Synthesis](#) *Alexey Gritsenko, Tim Salimans, Rianne van den Berg, Jasper Snoek, Nal Kalchbrenner*
- [Simulating a Primary Visual Cortex at the Front of CNNs Improves Robustness to Image Perturbations](#) *Joel Dapello, Tiago Marques, Martin Schrimpf, Franziska Geiger, David Cox, James J. DiCarlo*
- [Learning from Positive and Unlabeled Data with Arbitrary Positive Shift](#) *Zayd Hammoudeh, Daniel Lowd*
- [Deep Energy-based Modeling of Discrete-Time Physics](#) *Takashi Matsubara, Ai Ishikawa, Takaharu Yaguchi*
- [Quantifying Learnability and Describability of Visual Concepts Emerging in Representation Learning](#) *Iro Laina, Ruth Fong, Andrea Vedaldi*
- [Self-Learning Transformations for Improving Gaze and Head Redirection](#) *Yufeng Zheng, Seonwook Park, Xucong Zhang, Shalini De Mello, Otmar Hilliges*
- [Language-Conditioned Imitation Learning for Robot Manipulation Tasks](#) *Simon Stepputtis, Joseph Campbell, Mariano Phielipp, Stefan Lee, Chitta Baral, Heni Ben Amor*
- [POMDPs in Continuous Time and Discrete Spaces](#) *Bastian Alt, Matthias Schultheis, Heinz Koepl*
- [Exemplar Guided Active Learning](#) *Jason S. Hartford, Kevin Leyton-Brown, Hadas Raviv, Dan Padnos, Shahar Lev, Barak Lenz*
- [Grasp Proposal Networks: An End-to-End Solution for Visual Learning of Robotic Grasps](#) *Chaozheng Wu, Jian Chen, Qiaoyu Cao, Jianchi Zhang, Yunxin Tai, Lin Sun, Kui Jia*
- [Node Embeddings and Exact Low-Rank Representations of Complex Networks](#) *Sudhanshu Chanpuriya, Cameron Musco, Konstantinos Sotiropoulos, Charalampos Tsourakakis*
- [Fictitious Play for Mean Field Games: Continuous Time Analysis and Applications](#) *Sarah Perrin, Julien Perolat, Mathieu Lauriere, Matthieu Geist, Romuald Elie, Olivier Pietquin*
- [Steering Distortions to Preserve Classes and Neighbors in Supervised Dimensionality Reduction](#) *Benoît Colange, Jaakko Peltonen, Michael Aupetit, Denys Dutykh, Sylvain Lespinats*
- [On Infinite-Width Hypernetworks](#) *Etai Littwin, Tomer Galanti, Lior Wolf, Greg Yang*
- [Interferobot: aligning an optical interferometer by a reinforcement learning agent](#) *Dmitry Sorokin, Alexander Ulanov, Ekaterina Sazhina, Alexander Lvovsky*
- [Program Synthesis with Pragmatic Communication](#) *Yewen Pu, Kevin Ellis, Marta Kryven, Josh Tenenbaum,*

Armando Solar-Lezama

- Principal Neighbourhood Aggregation for Graph Nets *Gabriele Corso, Luca Cavalleri, Dominique Beaini, Pietro Liò, Petar Veličković*
- Reliable Graph Neural Networks via Robust Aggregation *Simon Geisler, Daniel Zügner, Stephan Günnemann*
- Instance Selection for GANs *Terrance DeVries, Michal Drozdal, Graham W. Taylor*
- Linear Disentangled Representations and Unsupervised Action Estimation *Matthew Painter, Adam Prugel-Bennett, Jonathon Hare*
- Video Frame Interpolation without Temporal Priors *Youjian Zhang, Chaoyue Wang, Dacheng Tao*
- Learning compositional functions via multiplicative weight updates *Jeremy Bernstein, Jiawei Zhao, Markus Meister, Ming-Yu Liu, Anima Anandkumar, Yisong Yue*
- Sample Complexity of Uniform Convergence for Multicalibration *Eliran Shabat, Lee Cohen, Yishay Mansour*
- Differentiable Neural Architecture Search in Equivalent Space with Exploration Enhancement *Miao Zhang, Huiqi Li, Shirui Pan, Xiaojun Chang, Zongyuan Ge, Steven Su*
- The interplay between randomness and structure during learning in RNNs *Friedrich Schuessler, Francesca Mastrogiuseppe, Alexis Dubreuil, Srdjan Ostojic, Omri Barak*
- A Generalized Neural Tangent Kernel Analysis for Two-layer Neural Networks *Zixiang Chen, Yuan Cao, Quanquan Gu, Tong Zhang*
- Instance-wise Feature Grouping *Aria Masoomi, Chieh Wu, Tingting Zhao, Zifeng Wang, Peter Castaldi, Jennifer Dy*
- Robust Disentanglement of a Few Factors at a Time using rPU-VAE *Benjamin Estermann, Markus Marks, Mehmet Fatih Yanik*
- PC-PG: Policy Cover Directed Exploration for Provable Policy Gradient Learning *Alekh Agarwal, Mikael Henaff, Sham Kakade, Wen Sun*
- Group Contextual Encoding for 3D Point Clouds *Xu Liu, Chengtao Li, Jian Wang, Jingbo Wang, Boxin Shi, Xiaodong He*
- Latent Bandits Revisited *Joey Hong, Branislav Kveton, Manzil Zaheer, Yinlam Chow, Amr Ahmed, Craig Boutilier*
- Is normalization indispensable for training deep neural network? *Jie Shao, Kai Hu, Changhu Wang, Xiangyang Xue, Bhiksha Raj*
- Optimization and Generalization of Shallow Neural Networks with Quadratic Activation Functions *Stefano Sarao Mannelli, Eric Vanden-Eijnden, Lenka Zdeborová*
- Intra Order-preserving Functions for Calibration of Multi-Class Neural Networks *Amir Rahimi, Amirreza Shaban, Ching-An Cheng, Richard Hartley, Byron Boots*
- Linear Time Sinkhorn Divergences using Positive Features *Meyer Scetbon, Marco Cuturi*
- VarGrad: A Low-Variance Gradient Estimator for Variational Inference *Lorenz Richter, Ayman Boustati, Nikolas Nüsken, Francisco Ruiz, Omer Deniz Akyildiz*
- A Convolutional Auto-Encoder for Haplotype Assembly and Viral Quasispecies Reconstruction *Ziqi Ke, Haris Vikaló*
- Promoting Stochasticity for Expressive Policies via a Simple and Efficient Regularization Method *Qi Zhou, Yufei Kuang, Zherui Qiu, Houqiang Li, Jie Wang*
- Adversarial Counterfactual Learning and Evaluation for Recommender System *Da Xu, Chuanwei Ruan, Evren Korpeoglu, Sushant Kumar, Kannan Achan*
- Memory-Efficient Learning of Stable Linear Dynamical Systems for Prediction and Control *Giorgos ("Yorgos") Mamakoukas, Orest Xherija, Todd Murphey*

- [Evolving Normalization-Activation Layers](#) *Hanxiao Liu, Andy Brock, Karen Simonyan, Quoc Le*
- [ScaleCom: Scalable Sparsified Gradient Compression for Communication-Efficient Distributed Training](#) *Chia-Yu Chen, Jiamin Ni, Songtao Lu, Xiaodong Cui, Pin-Yu Chen, Xiao Sun, Naigang Wang, Swagath Venkataramani, Vijayalakshmi (Viji) Srinivasan, Wei Zhang, Kailash Gopalakrishnan*
- [RelationNet++: Bridging Visual Representations for Object Detection via Transformer Decoder](#) *Cheng Chi, Fangyun Wei, Han Hu*
- [Efficient Learning of Discrete Graphical Models](#) *Marc Vuffray, Sidhant Misra, Andrey Lokhov*
- [Near-Optimal SQ Lower Bounds for Agnostically Learning Halfspaces and ReLUs under Gaussian Marginals](#) *Ilias Diakonikolas, Daniel Kane, Nikos Zarifis*
- [Neurosymbolic Transformers for Multi-Agent Communication](#) *Jeevana Priya Inala, Yichen Yang, James Paulos, Yewen Pu, Osbert Bastani, Vijay Kumar, Martin Rinard, Armando Solar-Lezama*
- [Fairness in Streaming Submodular Maximization: Algorithms and Hardness](#) *Marwa El Halabi, Slobodan Mitrović, Ashkan Norouzi-Fard, Jakab Tardos, Jakub M. Tarnawski*
- [Smoothed Geometry for Robust Attribution](#) *Zifan Wang, Haofan Wang, Shakul Ramkumar, Piotr Mardziel, Matt Fredrikson, Anupam Datta*
- [Fast Adversarial Robustness Certification of Nearest Prototype Classifiers for Arbitrary Seminorms](#) *Sascha Saralajew, Lars Holdijk, Thomas Villmann*
- [Multi-agent active perception with prediction rewards](#) *Mikko Lauri, Frans Oliehoek*
- [A Local Temporal Difference Code for Distributional Reinforcement Learning](#) *Pablo Tano, Peter Dayan, Alexandre Pouget*
- [Learning with Optimized Random Features: Exponential Speedup by Quantum Machine Learning without Sparsity and Low-Rank Assumptions](#) *Hayata Yamasaki, Sathyawageeswar Subramanian, Sho Sonoda, Masato Koashi*
- [CaSPR: Learning Canonical Spatiotemporal Point Cloud Representations](#) *Davis Rempe, Tolga Birdal, Yongheng Zhao, Zan Gojcic, Srinath Sridhar, Leonidas J. Guibas*
- [Deep Automodulators](#) *Ari Heljakka, Yuxin Hou, Juho Kannala, Arno Solin*
- [Convolutional Tensor-Train LSTM for Spatio-Temporal Learning](#) *Jiahao Su, Wonmin Byeon, Jean Kossaifi, Furong Huang, Jan Kautz, Anima Anandkumar*
- [The Potts-Ising model for discrete multivariate data](#) *Zahra Razaee, Arash Amini*
- [Interpretable multi-timescale models for predicting fMRI responses to continuous natural speech](#) *Shailee Jain, Vy Vo, Shivangi Mahto, Amanda LeBel, Javier S. Turek, Alexander Huth*
- [Group-Fair Online Allocation in Continuous Time](#) *Semih Cayci, Swati Gupta, Atilla Eryilmaz*
- [Decentralized TD Tracking with Linear Function Approximation and its Finite-Time Analysis](#) *Gang Wang, Songtao Lu, Georgios Giannakis, Gerald Tesauro, Jian Sun*
- [Understanding Gradient Clipping in Private SGD: A Geometric Perspective](#) *Xiangyi Chen, Steven Z. Wu, Mingyi Hong*
- [\$O\(n\)\$ Connections are Expressive Enough: Universal Approximability of Sparse Transformers](#) *Chulhee Yun, Yin-Wen Chang, Srinadh Bhojanapalli, Ankit Singh Rawat, Sashank Reddi, Sanjiv Kumar*
- [Identifying signal and noise structure in neural population activity with Gaussian process factor models](#) *Stephen Keeley, Mikio Aoi, Yiyi Yu, Spencer Smith, Jonathan W. Pillow*
- [Equivariant Networks for Hierarchical Structures](#) *Renhao Wang, Marjan Albooyeh, Siamak Ravanbakhsh*
- [MinMax Methods for Optimal Transport and Beyond: Regularization, Approximation and Numerics](#) *Luca De Gennaro Aquino, Stephan Eckstein*
- [A Discrete Variational Recurrent Topic Model without the Reparametrization Trick](#) *Mehdi Rezaee, Francis*

Ferraro

- [Transferable Graph Optimizers for ML Compilers](#) Yanqi Zhou, Sudip Roy, Amirali Abdolrashidi, Daniel Wong, Peter Ma, Qiumin Xu, Hanxiao Liu, Phitchaya Phothilimtha, Shen Wang, Anna Goldie, Azalia Mirhoseini, James Laudon
- [Learning with Operator-valued Kernels in Reproducing Kernel Krein Spaces](#) Akash Saha, Balamurugan Palaniappan
- [Learning Bounds for Risk-sensitive Learning](#) Jaeho Lee, Sejun Park, Jinwoo Shin
- [Simplifying Hamiltonian and Lagrangian Neural Networks via Explicit Constraints](#) Marc Finzi, Ke Alexander Wang, Andrew G. Wilson
- [Beyond accuracy: quantifying trial-by-trial behaviour of CNNs and humans by measuring error consistency](#) Robert Geirhos, Kristof Meding, Felix A. Wichmann
- [Provably Efficient Reinforcement Learning with Kernel and Neural Function Approximations](#) Zhuoran Yang, Chi Jin, Zhaoran Wang, Mengdi Wang, Michael Jordan
- [Constant-Expansion Suffices for Compressed Sensing with Generative Priors](#) Constantinos Daskalakis, Dhruv Rohatgi, Emmanouil Zampetakis
- [RANet: Region Attention Network for Semantic Segmentation](#) Dingguo Shen, Yuanfeng Ji, Ping Li, Yi Wang, Di Lin
- [A random matrix analysis of random Fourier features: beyond the Gaussian kernel, a precise phase transition, and the corresponding double descent](#) Zhenyu Liao, Romain Couillet, Michael W. Mahoney
- [Learning sparse codes from compressed representations with biologically plausible local wiring constraints](#) Kion Fallah, Adam Willats, Ninghao Liu, Christopher Rozell
- [Self-Imitation Learning via Generalized Lower Bound Q-learning](#) Yunhao Tang
- [Private Learning of Halfspaces: Simplifying the Construction and Reducing the Sample Complexity](#) Haim Kaplan, Yishay Mansour, Uri Stemmer, Eliad Tsfadia
- [Directional Pruning of Deep Neural Networks](#) Shih-Kang Chao, Zhanyu Wang, Yue Xing, Guang Cheng
- [Smoothly Bounding User Contributions in Differential Privacy](#) Alessandro Epasto, Mohammad Mahdian, Jieming Mao, Vahab Mirrokni, Lijie Ren
- [Accelerating Training of Transformer-Based Language Models with Progressive Layer Dropping](#) Minjia Zhang, Yuxiong He
- [Online Planning with Lookahead Policies](#) Yonathan Efroni, Mohammad Ghavamzadeh, Shie Mannor
- [Learning Deep Attribution Priors Based On Prior Knowledge](#) Ethan Weinberger, Joseph Janizek, Su-In Lee
- [Using noise to probe recurrent neural network structure and prune synapses](#) Eli Moore, Rishidev Chaudhuri
- [NanoFlow: Scalable Normalizing Flows with Sublinear Parameter Complexity](#) Sang-gil Lee, Sungwon Kim, Sungroh Yoon
- [Group Knowledge Transfer: Federated Learning of Large CNNs at the Edge](#) Chaoyang He, Murali Annavaram, Salman Avestimehr
- [Neural FFTs for Universal Texture Image Synthesis](#) Morteza Mardani, Guilin Liu, Aysegul Dundar, Shiqiu Liu, Andrew Tao, Bryan Catanzaro
- [Graph Cross Networks with Vertex Infomax Pooling](#) Maosen Li, Siheng Chen, Ya Zhang, Ivor Tsang
- [Instance-optimality in differential privacy via approximate inverse sensitivity mechanisms](#) Hilal Asi, John C. Duchi
- [Calibration of Shared Equilibria in General Sum Partially Observable Markov Games](#) Nelson Vadori, Sumitra Ganesh, Prashant Reddy, Manuela Veloso
- [MOPO: Model-based Offline Policy Optimization](#) Tianhe Yu, Garrett Thomas, Lantao Yu, Stefano Ermon,

James Y. Zou, Sergey Levine, Chelsea Finn, Tengyu Ma

- Building powerful and equivariant graph neural networks with structural message-passing *Clément Vignac, Andreas Loukas, Pascal Frossard*
- Efficient Model-Based Reinforcement Learning through Optimistic Policy Search and Planning *Sebastian Curi, Felix Berkenkamp, Andreas Krause*
- Practical Low-Rank Communication Compression in Decentralized Deep Learning *Thijs Vogels, Sai Praneeth Karimireddy, Martin Jaggi*
- Mutual exclusivity as a challenge for deep neural networks *Kanishk Gandhi, Brenden M. Lake*
- 3D Shape Reconstruction from Vision and Touch *Edward Smith, Roberto Calandra, Adriana Romero, Georgia Gkioxari, David Meger, Jitendra Malik, Michal Drozdal*
- GradAug: A New Regularization Method for Deep Neural Networks *Taojiannan Yang, Sijie Zhu, Chen Chen*
- An Equivalence between Loss Functions and Non-Uniform Sampling in Experience Replay *Scott Fujimoto, David Meger, Doina Precup*
- Learning Utilities and Equilibria in Non-Truthful Auctions *Hu Fu, Tao Lin*
- Rational neural networks *Nicolas Boulle, Yuji Nakatsukasa, Alex Townsend*
- DISK: Learning local features with policy gradient *Michał Tyszkiewicz, Pascal Fua, Eduard Trulls*
- Transfer Learning via $\ell_1\ell_1$ Regularization *Masaaki Takada, Hironori Fujisawa*
- GOCor: Bringing Globally Optimized Correspondence Volumes into Your Neural Network *Prune Truong, Martin Danelljan, Luc V. Gool, Radu Timofte*
- Deep Inverse Q-learning with Constraints *Gabriel Kalweit, Maria Huegle, Moritz Werling, Joschka Boedecker*
- Optimistic Dual Extrapolation for Coherent Non-monotone Variational Inequalities *Chaobing Song, Zhengyuan Zhou, Yichao Zhou, Yong Jiang, Yi Ma*
- Prediction with Corrupted Expert Advice *Idan Amir, Idan Attias, Tomer Koren, Yishay Mansour, Roi Livni*
- Human Parsing Based Texture Transfer from Single Image to 3D Human via Cross-View Consistency *Fang Zhao, Shengcai Liao, Kaihao Zhang, Ling Shao*
- Knowledge Augmented Deep Neural Networks for Joint Facial Expression and Action Unit Recognition *Zijun Cui, Tengfei Song, Yuru Wang, Qiang Ji*
- Point process models for sequence detection in high-dimensional neural spike trains *Alex Williams, Anthony Degleris, Yixin Wang, Scott Linderman*
- Adversarial Attacks on Linear Contextual Bandits *Evrard Garcelon, Baptiste Roziere, Laurent Meunier, Jean Tarbouriech, Olivier Teytaud, Alessandro Lazaric, Matteo Pirotta*
- Meta-Consolidation for Continual Learning *Joseph K J, Vineeth N Balasubramanian*
- Organizing recurrent network dynamics by task-computation to enable continual learning *Lea Duncker, Laura Driscoll, Krishna V. Shenoy, Maneesh Sahani, David Sussillo*
- Lifelong Policy Gradient Learning of Factored Policies for Faster Training Without Forgetting *Jorge Mendez, Boyu Wang, Eric Eaton*
- Kernel Methods Through the Roof: Handling Billions of Points Efficiently *Giacomo Meanti, Luigi Carratino, Lorenzo Rosasco, Alessandro Rudi*
- Spike and slab variational Bayes for high dimensional logistic regression *Kolyan Ray, Botond Szabo, Gabriel Clara*
- Maximum-Entropy Adversarial Data Augmentation for Improved Generalization and Robustness *Long Zhao, Ting Liu, Xi Peng, Dimitris Metaxas*
- Fast geometric learning with symbolic matrices *Jean Feydy, Alexis Glaunès, Benjamin Charlier, Michael Bronstein*

- [MESA: Boost Ensemble Imbalanced Learning with MEta-SAMpler](#) *Zhining Liu, Pengfei Wei, Jing Jiang, Wei Cao, Jiang Bian, Yi Chang*
- [CoinPress: Practical Private Mean and Covariance Estimation](#) *Sourav Biswas, Yihe Dong, Gautam Kamath, Jonathan Ullman*
- [Planning with General Objective Functions: Going Beyond Total Rewards](#) *Ruosong Wang, Peilin Zhong, Simon S. Du, Russ R. Salakhutdinov, Lin Yang*
- [Scattering GCN: Overcoming Oversmoothness in Graph Convolutional Networks](#) *Yimeng Min, Frederik Wenkel, Guy Wolf*
- [KFC: A Scalable Approximation Algorithm for \$k\$ -center Fair Clustering](#) *Elfarouk Harb, Ho Shan Lam*
- [Leveraging Predictions in Smoothed Online Convex Optimization via Gradient-based Algorithms](#) *Yingying Li, Na Li*
- [Learning the Linear Quadratic Regulator from Nonlinear Observations](#) *Zakaria Mhammedi, Dylan J. Foster, Max Simchowitz, Dipendra Misra, Wen Sun, Akshay Krishnamurthy, Alexander Rakhlin, John Langford*
- [Reconciling Modern Deep Learning with Traditional Optimization Analyses: The Intrinsic Learning Rate](#) *Zhiyuan Li, Kaifeng Lyu, Sanjeev Arora*
- [Scalable Graph Neural Networks via Bidirectional Propagation](#) *Ming Chen, Zhewei Wei, Bolin Ding, Yaliang Li, Ye Yuan, Xiaoyong Du, Ji-Rong Wen*
- [Distribution Aligning Refinery of Pseudo-label for Imbalanced Semi-supervised Learning](#) *Jaehyung Kim, Youngbum Hur, Sejun Park, Eunho Yang, Sung Ju Hwang, Jinwoo Shin*
- [Assisted Learning: A Framework for Multi-Organization Learning](#) *Xun Xian, Xinran Wang, Jie Ding, Reza Ghanadan*
- [The Strong Screening Rule for SLOPE](#) *Johan Larsson, Malgorzata Bogdan, Jonas Wallin*
- [STLnet: Signal Temporal Logic Enforced Multivariate Recurrent Neural Networks](#) *Meiyi Ma, Ji Gao, Lu Feng, John Stankovic*
- [Election Coding for Distributed Learning: Protecting SignSGD against Byzantine Attacks](#) *Jy-yong Sohn, Dong-Jun Han, Beongjun Choi, Jaekyun Moon*
- [Reducing Adversarially Robust Learning to Non-Robust PAC Learning](#) *Omar Montasser, Steve Hanneke, Nati Srebro*
- [Top-k Training of GANs: Improving GAN Performance by Throwing Away Bad Samples](#) *Samarth Sinha, Zhengli Zhao, Anirudh Goyal ALIAS PARTH GOYAL, Colin A. Raffel, Augustus Odena*
- [Black-Box Optimization with Local Generative Surrogates](#) *Sergey Shirobokov, Vladislav Belavin, Michael Kagan, Andrei Ustyuzhanin, Atılım Gunes Baydin*
- [Efficient Generation of Structured Objects with Constrained Adversarial Networks](#) *Luca Di Liello, Pierfrancesco Ardino, Jacopo Gobbi, Paolo Morettin, Stefano Teso, Andrea Passerini*
- [Hard Example Generation by Texture Synthesis for Cross-domain Shape Similarity Learning](#) *Huan Fu, Shunming Li, Rongfei Jia, Mingming Gong, Binqiang Zhao, Dacheng Tao*
- [Recovery of sparse linear classifiers from mixture of responses](#) *Venkata Gandikota, Arya Mazumdar, Soumyabrata Pal*
- [Efficient Distance Approximation for Structured High-Dimensional Distributions via Learning](#) *Arnab Bhattacharyya, Sutanu Gayen, Kuldeep S Meel, N. V. Vinodchandran*
- [A Single Recipe for Online Submodular Maximization with Adversarial or Stochastic Constraints](#) *Omid Sadeghi, Prasanna Raut, Maryam Fazel*
- [Learning Sparse Prototypes for Text Generation](#) *Junxian He, Taylor Berg-Kirkpatrick, Graham Neubig*
- [Implicit Rank-Minimizing Autoencoder](#) *Li Jing, Jure Zbontar, yann lecun*

- Storage Efficient and Dynamic Flexible Runtime Channel Pruning via Deep Reinforcement Learning Jianda Chen, Shangyu Chen, Sinno Jialin Pan
- Task-Oriented Feature Distillation Linfeng Zhang, Yukang Shi, Zuoqiang Shi, Kaisheng Ma, Chenglong Bao
- Entropic Causal Inference: Identifiability and Finite Sample Results Spencer Compton, Murat Kocaoglu, Kristjan Greenewald, Dmitry Katz
- Rewriting History with Inverse RL: Hindsight Inference for Policy Improvement Ben Eysenbach, XINYANG GENG, Sergey Levine, Russ R. Salakhutdinov
- Variance-Reduced Off-Policy TDC Learning: Non-Asymptotic Convergence Analysis Shaocong Ma, Yi Zhou, Shaofeng Zou
- AdaTune: Adaptive Tensor Program Compilation Made Efficient Menghao Li, Minjia Zhang, Chi Wang, Mingqin Li
- When Do Neural Networks Outperform Kernel Methods? Behrooz Ghorbani, Song Mei, Theodor Misiakiewicz, Andrea Montanari
- STEER : Simple Temporal Regularization For Neural ODE Arnab Ghosh, Harkirat Behl, Emilien Dupont, Philip Torr, Vinay Namboodiri
- A Variational Approach for Learning from Positive and Unlabeled Data Hui Chen, Fangqing Liu, Yin Wang, Liyue Zhao, Hao Wu
- Efficient Clustering Based On A Unified View Of KK-means And Ratio-cut Shenfei Pei, Feiping Nie, Rong Wang, Xuelong Li
- Recurrent Switching Dynamical Systems Models for Multiple Interacting Neural Populations Joshua Glaser, Matthew Whiteway, John P. Cunningham, Liam Paninski, Scott Linderman
- Coresets via Bilevel Optimization for Continual Learning and Streaming Zalán Borsos, Mojmir Mutny, Andreas Krause
- Generalized Independent Noise Condition for Estimating Latent Variable Causal Graphs Feng Xie, Ruichu Cai, Biwei Huang, Clark Glymour, Zhifeng Hao, Kun Zhang
- Understanding and Exploring the Network with Stochastic Architectures Zhijie Deng, Yinpeng Dong, Shifeng Zhang, Jun Zhu
- All-or-nothing statistical and computational phase transitions in sparse spiked matrix estimation jean barbier, Nicolas Macris, Cynthia Rush
- Deep Evidential Regression Alexander Amini, Wilko Schwarting, Ava Soleimany, Daniela Rus
- Analytical Probability Distributions and Exact Expectation-Maximization for Deep Generative Networks Randall Balestriero, Sebastien PARIS, Richard Baraniuk
- Bayesian Pseudocoresets Dionysis Manousakas, Zuheng Xu, Cecilia Mascolo, Trevor Campbell
- See, Hear, Explore: Curiosity via Audio-Visual Association Victoria Dean, Shubham Tulsiani, Abhinav Gupta
- Adversarial Training is a Form of Data-dependent Operator Norm Regularization Kevin Roth, Yannic Kilcher, Thomas Hofmann
- A Biologically Plausible Neural Network for Slow Feature Analysis David Lipshutz, Charles Windolf, Siavash Golkar, Dmitri Chklovskii
- Learning Feature Sparse Principal Subspace Lai Tian, Feiping Nie, Rong Wang, Xuelong Li
- Online Adaptation for Consistent Mesh Reconstruction in the Wild Xueting Li, Sifei Liu, Shalini De Mello, Kihwan Kim, Xiaolong Wang, Ming-Hsuan Yang, Jan Kautz
- Online learning with dynamics: A minimax perspective Kush Bhatia, Karthik Sridharan
- Learning to Select Best Forecast Tasks for Clinical Outcome Prediction Yuan Xue, Nan Du, Anne Mottram, Martin Seneviratne, Andrew M. Dai

- Stochastic Optimization with Heavy-Tailed Noise via Accelerated Gradient Clipping *Eduard Gorbunov, Marina Danilova, Alexander Gasnikov*
- Adaptive Experimental Design with Temporal Interference: A Maximum Likelihood Approach *Peter W. Glynn, Ramesh Johari, Mohammad Rasouli*
- From Trees to Continuous Embeddings and Back: Hyperbolic Hierarchical Clustering *Ines Chami, Albert Gu, Vaggos Chatziafratis, Christopher Ré*
- The Autoencoding Variational Autoencoder *Taylan Cemgil, Sumedh Ghaisas, Krishnamurthy Dvijotham, Sven Gowal, Pushmeet Kohli*
- A Fair Classifier Using Kernel Density Estimation *Jaewoong Cho, Gyeongjo Hwang, Changho Suh*
- A Randomized Algorithm to Reduce the Support of Discrete Measures *Francesco Cosentino, Harald Oberhauser, Alessandro Abate*
- Distributionally Robust Federated Averaging *Yuyang Deng, Mohammad Mahdi Kamani, Mehrdad Mahdavi*
- Sharp uniform convergence bounds through empirical centralization *Cyrus Cousins, Matteo Riondato*
- COBE: Contextualized Object Embeddings from Narrated Instructional Video *Gedas Bertasius, Lorenzo Torresani*
- Knowledge Transfer in Multi-Task Deep Reinforcement Learning for Continuous Control *Zhiyuan Xu, Kun Wu, Zhengping Che, Jian Tang, Jieping Ye*
- Finite Versus Infinite Neural Networks: an Empirical Study *Jaehoon Lee, Samuel Schoenholz, Jeffrey Pennington, Ben Adlam, Lechao Xiao, Roman Novak, Jascha Sohl-Dickstein*
- Supermasks in Superposition *Mitchell Wortsman, Vivek Ramanujan, Rosanne Liu, Aniruddha Kembhavi, Mohammad Rastegari, Jason Yosinski, Ali Farhadi*
- Nonasymptotic Guarantees for Spiked Matrix Recovery with Generative Priors *Jorio Cocola, Paul Hand, Vlad Voroninski*
- Almost Optimal Model-Free Reinforcement Learning via Reference-Advantage Decomposition *Zihan Zhang, Yuan Zhou, Xiangyang Ji*
- Learning to Incentivize Other Learning Agents *Jiachen Yang, Ang Li, Mehrdad Farajtabar, Peter Sunehag, Edward Hughes, Hongyuan Zha*
- Displacement-Invariant Matching Cost Learning for Accurate Optical Flow Estimation *Jianyuan Wang, Yiran Zhong, Yuchao Dai, Kaihao Zhang, Pan Ji, Hongdong Li*
- Distributionally Robust Local Non-parametric Conditional Estimation *Viet Anh Nguyen, Fan Zhang, Jose Blanchet, Erick Delage, Yinyu Ye*
- Robust Multi-Object Matching via Iterative Reweighting of the Graph Connection Laplacian *Yunpeng Shi, Shaohan Li, Gilad Lerman*
- Meta-Gradient Reinforcement Learning with an Objective Discovered Online *Zhongwen Xu, Hado P. van Hasselt, Matteo Hessel, Junhyuk Oh, Satinder Singh, David Silver*
- Learning Strategy-Aware Linear Classifiers *Yiling Chen, Yang Liu, Chara Podimata*
- Upper Confidence Primal-Dual Reinforcement Learning for CMDP with Adversarial Loss *Shuang Qiu, Xiaohan Wei, Zhuoran Yang, Jieping Ye, Zhaoran Wang*
- Calibrating Deep Neural Networks using Focal Loss *Jishnu Mukhoti, Viveka Kulharia, Amartya Sanyal, Stuart Golodetz, Philip Torr, Puneet Dokania*
- Optimizing Mode Connectivity via Neuron Alignment *Norman Tatro, Pin-Yu Chen, Payel Das, Igor Melnyk, Prasanna Sattigeri, Rongjie Lai*
- Information Theoretic Regret Bounds for Online Nonlinear Control *Sham Kakade, Akshay Krishnamurthy, Kendall Lowrey, Motoya Ohnishi, Wen Sun*

- [A kernel test for quasi-independence](#) *Tamara Fernandez, Wenkai Xu, Marc Ditzhaus, Arthur Gretton*
- [First Order Constrained Optimization in Policy Space](#) *Yiming Zhang, Quan Vuong, Keith Ross*
- [Learning Augmented Energy Minimization via Speed Scaling](#) *Etienne Bamas, Andreas Maggiori, Lars Rohwedder, Ola Svensson*
- [Exploiting MMD and Sinkhorn Divergences for Fair and Transferable Representation Learning](#) *Luca Oneto, Michele Donini, Giulia Luise, Carlo Ciliberto, Andreas Maurer, Massimiliano Pontil*
- [Deep Rao-Blackwellised Particle Filters for Time Series Forecasting](#) *Richard Kurle, Syama Sundar Rangapuram, Emmanuel de Bézenac, Stephan Günnemann, Jan Gasthaus*
- [Why are Adaptive Methods Good for Attention Models?](#) *Jingzhao Zhang, Sai Praneeth Karimireddy, Andreas Veit, Seungyeon Kim, Sashank Reddi, Sanjiv Kumar, Suvrit Sra*
- [Neural Sparse Representation for Image Restoration](#) *Yuchen Fan, Jiahui Yu, Yiqun Mei, Yulun Zhang, Yun Fu, Ding Liu, Thomas S. Huang*
- [Boosting First-Order Methods by Shifting Objective: New Schemes with Faster Worst-Case Rates](#) *Kaiwen Zhou, Anthony Man-Cho So, James Cheng*
- [Robust Sequence Submodular Maximization](#) *Gamal Sallam, Zizhan Zheng, Jie Wu, Bo Ji*
- [Certified Monotonic Neural Networks](#) *Xingchao Liu, Xing Han, Na Zhang, Qiang Liu*
- [System Identification with Biophysical Constraints: A Circuit Model of the Inner Retina](#) *Cornelius Schröder, David Klindt, Sarah Strauss, Katrin Franke, Matthias Bethge, Thomas Euler, Philipp Berens*
- [Efficient Algorithms for Device Placement of DNN Graph Operators](#) *Jakub M. Tarnawski, Amar Phanishayee, Nikhil Devanur, Divya Mahajan, Fanny Nina Paravecino*
- [Active Invariant Causal Prediction: Experiment Selection through Stability](#) *Juan L. Gamella, Christina Heinze-Deml*
- [BOSS: Bayesian Optimization over String Spaces](#) *Henry Moss, David Leslie, Daniel Beck, Javier González, Paul Rayson*
- [Model Interpretability through the lens of Computational Complexity](#) *Pablo Barceló, Mikaël Monet, Jorge Pérez, Bernardo Subercaseaux*
- [Markovian Score Climbing: Variational Inference with \$KL\(p||q\)\$](#) *Christian Naeseth, Fredrik Lindsten, David Blei*
- [Improved Analysis of Clipping Algorithms for Non-convex Optimization](#) *Bohang Zhang, Jikai Jin, Cong Fang, Liwei Wang*
- [Bias no more: high-probability data-dependent regret bounds for adversarial bandits and MDPs](#) *Chung-Wei Lee, Haipeng Luo, Chen-Yu Wei, Mengxiao Zhang*
- [A Ranking-based, Balanced Loss Function Unifying Classification and Localisation in Object Detection](#) *Kemal Oksuz, Baris Can Cam, Emre Akbas, Sinan Kalkan*
- [StratLearner: Learning a Strategy for Misinformation Prevention in Social Networks](#) *Guangmo Tong*
- [A Unified Switching System Perspective and Convergence Analysis of Q-Learning Algorithms](#) *Donghwan Lee, Niao He*
- [Kernel Alignment Risk Estimator: Risk Prediction from Training Data](#) *Arthur Jacot, Berfin Simsek, Francesco Spadaro, Clement Hongler, Franck Gabriel*
- [Calibrating CNNs for Lifelong Learning](#) *Pravendra Singh, Vinay Kumar Verma, Pratik Mazumder, Lawrence Carin, Piyush Rai*
- [Online Convex Optimization Over Erdos-Renyi Random Networks](#) *Jinlong Lei, Peng Yi, Yiguang Hong, Jie Chen, Guodong Shi*
- [Robustness of Bayesian Neural Networks to Gradient-Based Attacks](#) *Ginevra Carbone, Matthew Wicker,*

Luca Laurenti, Andrea Patane', Luca Bortolussi, Guido Sanguinetti

- Parametric Instance Classification for Unsupervised Visual Feature learning *Yue Cao, Zhenda Xie, Bin Liu, Yutong Lin, Zheng Zhang, Han Hu*
- Sparse Weight Activation Training *Md Aamir Raihan, Tor Aamodt*
- Collapsing Bandits and Their Application to Public Health Intervention *Aditya Mate, Jackson Killian, Haifeng Xu, Andrew Perrault, Milind Tambe*
- Neural Sparse Voxel Fields *Lingjie Liu, Jiatao Gu, Kyaw Zaw Lin, Tat-Seng Chua, Christian Theobalt*
- A Flexible Framework for Designing Trainable Priors with Adaptive Smoothing and Game Encoding *Bruno Lecouat, Jean Ponce, Julien Mairal*
- The Discrete Gaussian for Differential Privacy *Clément L. Canonne, Gautam Kamath, Thomas Steinke*
- Robust Sub-Gaussian Principal Component Analysis and Width-Independent Schatten Packing *Arun Jambulapati, Jerry Li, Kevin Tian*
- Adaptive Importance Sampling for Finite-Sum Optimization and Sampling with Decreasing Step-Sizes *Ayoub El Hanchi, David Stephens*
- Learning efficient task-dependent representations with synaptic plasticity *Colin Bredenberg, Eero Simoncelli, Cristina Savin*
- A Contour Stochastic Gradient Langevin Dynamics Algorithm for Simulations of Multi-modal Distributions *Wei Deng, Guang Lin, Faming Liang*
- Error Bounds of Imitating Policies and Environments *Tian Xu, Ziniu Li, Yang Yu*
- Disentangling Human Error from Ground Truth in Segmentation of Medical Images *Le Zhang, Ryutaro Tanno, Mou-Cheng Xu, Chen Jin, Joseph Jacob, Olga Ciccarrelli, Frederik Barkhof, Daniel Alexander*
- Consequences of Misaligned AI *Simon Zhuang, Dylan Hadfield-Menell*
- Promoting Coordination through Policy Regularization in Multi-Agent Deep Reinforcement Learning *Julien Roy, Paul Barde, Félix Harvey, Derek Nowrouzezahrai, Chris Pal*
- Emergent Reciprocity and Team Formation from Randomized Uncertain Social Preferences *Bowen Baker*
- Hitting the High Notes: Subset Selection for Maximizing Expected Order Statistics *Aranyak Mehta, Uri Nadav, Alexandros Psomas, Aviad Rubinstein*
- Towards Scale-Invariant Graph-related Problem Solving by Iterative Homogeneous GNNs *Hao Tang, Zhiao Huang, Jiayuan Gu, Bao-Liang Lu, Hao Su*
- Regret Bounds without Lipschitz Continuity: Online Learning with Relative-Lipschitz Losses *Yihan Zhou, Victor Sanches Portella, Mark Schmidt, Nicholas Harvey*
- The Lottery Ticket Hypothesis for Pre-trained BERT Networks *Tianlong Chen, Jonathan Frankle, Shiyu Chang, Sijia Liu, Yang Zhang, Zhangyang Wang, Michael Carbin*
- Label-Aware Neural Tangent Kernel: Toward Better Generalization and Local Elasticity *Shuxiao Chen, Hangfeng He, Weijie Su*
- Beyond Perturbations: Learning Guarantees with Arbitrary Adversarial Test Examples *Shafi Goldwasser, Adam Tauman Kalai, Yael Kalai, Omar Montasser*
- AdvFlow: Inconspicuous Black-box Adversarial Attacks using Normalizing Flows *Hadi Mohaghegh Dolatabadi, Sarah Erfani, Christopher Leckie*
- Few-shot Image Generation with Elastic Weight Consolidation *Yijun Li, Richard Zhang, Jingwan (Cynthia) Lu, Eli Shechtman*
- On the Expressiveness of Approximate Inference in Bayesian Neural Networks *Andrew Foong, David Burt, Yingzhen Li, Richard Turner*
- Non-Crossing Quantile Regression for Distributional Reinforcement Learning *Fan Zhou, Jianing Wang,*

Xingdong Feng

- [Dark Experience for General Continual Learning: a Strong, Simple Baseline](#) *Pietro Buzzega, Matteo Boschini, Angelo Porrello, Davide Abati, SIMONE CALDERARA*
- [Learning to Utilize Shaping Rewards: A New Approach of Reward Shaping](#) *Yujing Hu, Weixun Wang, Hangtian Jia, Yixiang Wang, Yingfeng Chen, Jianye Hao, Feng Wu, Changjie Fan*
- [Neural encoding with visual attention](#) *Meenakshi Khosla, Gia Ngo, Keith Jamison, Amy Kuceyeski, Mert Sabuncu*
- [On the linearity of large non-linear models: when and why the tangent kernel is constant](#) *Chaoyue Liu, Libin Zhu, Misha Belkin*
- [PLLay: Efficient Topological Layer based on Persistent Landscapes](#) *Kwangho Kim, Jisu Kim, Manzil Zaheer, Joon Kim, Frederic Chazal, Larry Wasserman*
- [Decentralized Langevin Dynamics for Bayesian Learning](#) *Anjaly Parayil, He Bai, Jemin George, Prudhvi Gurram*
- [Shared Space Transfer Learning for analyzing multi-site fMRI data](#) *Tony Muhammad Yousefnezhad, Alessandro Selvitella, Daoqiang Zhang, Andrew Greenshaw, Russell Greiner*
- [The Diversified Ensemble Neural Network](#) *Shaofeng Zhang, Meng Liu, Junchi Yan*
- [Inductive Quantum Embedding](#) *Santosh Kumar Srivastava, Dinesh Khandelwal, Dhiraj Madan, Dinesh Garg, Hima Karanam, L Venkata Subramaniam*
- [Variational Bayesian Unlearning](#) *Quoc Phong Nguyen, Bryan Kian Hsiang Low, Patrick Jaillet*
- [Batched Coarse Ranking in Multi-Armed Bandits](#) *Nikolai Karpov, Qin Zhang*
- [Understanding and Improving Fast Adversarial Training](#) *Maksym Andriushchenko, Nicolas Flammarion*
- [Coded Sequential Matrix Multiplication For Straggler Mitigation](#) *Nikhil Krishnan Muralee Krishnan, Seyederfan Hosseini, Ashish Khisti*
- [Attack of the Tails: Yes, You Really Can Backdoor Federated Learning](#) *Hongyi Wang, Kartik Sreenivasan, Shashank Rajput, Harit Vishwakarma, Saurabh Agarwal, Jy-yong Sohn, Kangwook Lee, Dimitris Papailiopoulos*
- [Certiably Adversarially Robust Detection of Out-of-Distribution Data](#) *Julian Bitterwolf, Alexander Meinke, Matthias Hein*
- [Domain Generalization via Entropy Regularization](#) *Shanshan Zhao, Mingming Gong, Tongliang Liu, Huan Fu, Dacheng Tao*
- [Bayesian Meta-Learning for the Few-Shot Setting via Deep Kernels](#) *Massimiliano Patacchiola, Jack Turner, Elliot J. Crowley, Michael O'Boyle, Amos J. Storkey*
- [Skeleton-bridged Point Completion: From Global Inference to Local Adjustment](#) *Yinyu Nie, Yiqun Lin, Xiaoguang Han, Shihui Guo, Jian Chang, Shuguang Cui, Jian.J Zhang*
- [Compressing Images by Encoding Their Latent Representations with Relative Entropy Coding](#) *Gergely Flamich, Marton Havasi, José Miguel Hernández-Lobato*
- [Improved Guarantees for k-means++ and k-means++ Parallel](#) *Konstantin Makarychev, Aravind Reddy, Liren Shan*
- [Sparse Spectrum Warped Input Measures for Nonstationary Kernel Learning](#) *Anthony Tompkins, Rafael Oliveira, Fabio T. Ramos*
- [An Efficient Adversarial Attack for Tree Ensembles](#) *Chong Zhang, Huan Zhang, Cho-Jui Hsieh*
- [Learning Continuous System Dynamics from Irregularly-Sampled Partial Observations](#) *Zijie Huang, Yizhou Sun, Wei Wang*
- [Online Bayesian Persuasion](#) *Matteo Castiglioni, Andrea Celli, Alberto Marchesi, Nicola Gatti*

- Robust Pre-Training by Adversarial Contrastive Learning Ziyu Jiang, Tianlong Chen, Ting Chen, Zhangyang Wang
- Random Walk Graph Neural Networks Giannis Nikolentzos, Michalis Vazirgiannis
- Explore Aggressively, Update Conservatively: Stochastic Extragradient Methods with Variable Stepsize Scaling Yu-Guan Hsieh, Franck Iutzeler, Jérôme Malick, Panayotis Mertikopoulos
- Fast and Accurate k-means++ via Rejection Sampling Vincent Cohen-Addad, Silvio Lattanzi, Ashkan Norouzi-Fard, Christian Sohler, Ola Svensson
- Variational Amodal Object Completion Huan Ling, David Acuna, Karsten Kreis, Seung Wook Kim, Sanja Fidler
- When Counterpoint Meets Chinese Folk Melodies Nan Jiang, Sheng Jin, Zhiyao Duan, Changshui Zhang
- Sub-linear Regret Bounds for Bayesian Optimisation in Unknown Search Spaces Hung Tran-The, Sunil Gupta, Santu Rana, Huong Ha, Svetha Venkatesh
- Universal Domain Adaptation through Self Supervision Kuniaki Saito, Donghyun Kim, Stan Sclaroff, Kate Saenko
- Patch2Self: Denoising Diffusion MRI with Self-Supervised Learning Shreyas Fadnavis, Joshua Batson, Eleftherios Garyfallidis
- Stochastic Normalization Zhi Kou, Kaichao You, Mingsheng Long, Jianmin Wang
- Constrained episodic reinforcement learning in concave-convex and knapsack settings Kianté Brantley, Miro Dudik, Thodoris Lykouris, Sobhan Miryoosefi, Max Simchowitz, Aleksandrs Slivkins, Wen Sun
- On Learning Ising Models under Huber's Contamination Model Adarsh Prasad, Vishwak Srinivasan, Sivaraman Balakrishnan, Pradeep Ravikumar
- Cross-validation Confidence Intervals for Test Error Pierre Bayle, Alexandre Bayle, Lucas Janson, Lester Mackey
- DeepSVG: A Hierarchical Generative Network for Vector Graphics Animation Alexandre Carlier, Martin Danelljan, Alexandre Alahi, Radu Timofte
- Bayesian Attention Modules Xinjie Fan, Shujian Zhang, Bo Chen, Mingyuan Zhou
- Robustness Analysis of Non-Convex Stochastic Gradient Descent using Biased Expectations Kevin Scaman, Cedric Malherbe
- SoftFlow: Probabilistic Framework for Normalizing Flow on Manifolds Hyeongju Kim, Hyeonseung Lee, Woo Hyun Kang, Joun Yeop Lee, Nam Soo Kim
- A meta-learning approach to (re)discover plasticity rules that carve a desired function into a neural network Basile Confavreux, Friedemann Zenke, Everton Agnes, Timothy Lillicrap, Tim Vogels
- Greedy Optimization Provably Wins the Lottery: Logarithmic Number of Winning Tickets is Enough Mao Ye, Lemeng Wu, Qiang Liu
- Path Integral Based Convolution and Pooling for Graph Neural Networks Zheng Ma, Junyu Xuan, Yu Guang Wang, Ming Li, Pietro Liò
- Estimating the Effects of Continuous-valued Interventions using Generative Adversarial Networks Ioana Bica, James Jordon, Mihaela van der Schaar
- Latent Dynamic Factor Analysis of High-Dimensional Neural Recordings Heejong Bong, Zongge Liu, Zhao Ren, Matthew Smith, Valerie Ventura, Kass E. Robert
- Conditioning and Processing: Techniques to Improve Information-Theoretic Generalization Bounds Hassan Hafez-Kolahi, Zeinab Golgooni, Shohreh Kasaei, Mahdieh Soleymani
- Bongard-LOGO: A New Benchmark for Human-Level Concept Learning and Reasoning Weili Nie, Zhiding Yu, Lei Mao, Ankit B. Patel, Yuke Zhu, Anima Anandkumar

- [GAN Memory with No Forgetting](#) *Yulai Cong, Miaoyun Zhao, Jianqiao Li, Sijia Wang, Lawrence Carin*
- [Deep Reinforcement Learning with Stacked Hierarchical Attention for Text-based Games](#) *Yunqiu Xu, Meng Fang, Ling Chen, Yali Du, Joey Tianyi Zhou, Chengqi Zhang*
- [Gaussian Gated Linear Networks](#) *David Budden, Adam Marblestone, Eren Sezener, Tor Lattimore, Gregory Wayne, Joel Veness*
- [Node Classification on Graphs with Few-Shot Novel Labels via Meta Transformed Network Embedding](#) *Lin Lan, Pinghui Wang, Xuefeng Du, Kaikai Song, Jing Tao, Xiaohong Guan*
- [Online Fast Adaptation and Knowledge Accumulation \(OSAKA\): a New Approach to Continual Learning](#) *Massimo Caccia, Pau Rodriguez, Oleksiy Ostapenko, Fabrice Normandin, Min Lin, Lucas Page-Caccia, Issam Hadj Laradji, Irina Rish, Alexandre Lacoste, David Vázquez, Laurent Charlin*
- [Convex optimization based on global lower second-order models](#) *Nikita Doikov, Yurii Nesterov*
- [Simultaneously Learning Stochastic and Adversarial Episodic MDPs with Known Transition](#) *Tiancheng Jin, Haipeng Luo*
- [Relative gradient optimization of the Jacobian term in unsupervised deep learning](#) *Luigi Gresele, Giancarlo Fissore, Adrián Javaloy, Bernhard Schölkopf, Aapo Hyvarinen*
- [Self-Supervised Visual Representation Learning from Hierarchical Grouping](#) *Xiao Zhang, Michael Maire*
- [Optimal Variance Control of the Score-Function Gradient Estimator for Importance-Weighted Bounds](#) *Valentin Liévin, Andrea Dittadi, Anders Christensen, Ole Winther*
- [Explicit Regularisation in Gaussian Noise Injections](#) *Alexander Camuto, Matthew Willetts, Umut Simsekli, Stephen J. Roberts, Chris C. Holmes*
- [Numerically Solving Parametric Families of High-Dimensional Kolmogorov Partial Differential Equations via Deep Learning](#) *Julius Berner, Markus Dablander, Philipp Grohs*
- [Finite-Time Analysis for Double Q-learning](#) *Huaqing Xiong, Lin Zhao, Yingbin Liang, Wei Zhang*
- [Learning to Detect Objects with a 1 Megapixel Event Camera](#) *Etienne Perot, Pierre de Tournemire, Davide Nitti, Jonathan Masci, Amos Sironi*
- [End-to-End Learning and Intervention in Games](#) *Jiayang Li, Jing Yu, Yu Nie, Zhaoran Wang*
- [Least Squares Regression with Markovian Data: Fundamental Limits and Algorithms](#) *Dheeraj Nagaraj, Xian Wu, Guy Bresler, Prateek Jain, Praneeth Netrapalli*
- [Predictive coding in balanced neural networks with noise, chaos and delays](#) *Jonathan Kadmon, Jonathan Timcheck, Surya Ganguli*
- [Interpolation Technique to Speed Up Gradients Propagation in Neural ODEs](#) *Talgat Daulbaev, Alexandr Katrutsa, Larisa Markeeva, Julia Gusak, Andrzej Cichocki, Ivan Oseledets*
- [On the Equivalence between Online and Private Learnability beyond Binary Classification](#) *Young Jung, Baekjin Kim, Ambuj Tewari*
- [AViD Dataset: Anonymized Videos from Diverse Countries](#) *AJ Piergiovanni, Michael Ryoo*
- [Probably Approximately Correct Constrained Learning](#) *Luiz Chamon, Alejandro Ribeiro*
- [RATT: Recurrent Attention to Transient Tasks for Continual Image Captioning](#) *Riccardo Del Chiaro, Bartłomiej Twardowski, Andrew Bagdanov, Joost van de Weijer*
- [Decisions, Counterfactual Explanations and Strategic Behavior](#) *Stratis Tsirtsis, Manuel Gomez Rodriguez*
- [Hierarchical Patch VAE-GAN: Generating Diverse Videos from a Single Sample](#) *Shir Gur, Sagie Benaim, Lior Wolf*
- [A Feasible Level Proximal Point Method for Nonconvex Sparse Constrained Optimization](#) *Digvijay Boob, Qi Deng, Guanghui Lan, Yilin Wang*
- [Reservoir Computing meets Recurrent Kernels and Structured Transforms](#) *Jonathan Dong, Ruben Ohana,*

Mushegh Rafayelyan, Florent Krzakala

- Comprehensive Attention Self-Distillation for Weakly-Supervised Object Detection *Zeyi Huang, Yang Zou, B. V. K. Vijaya Kumar, Dong Huang*
- Linear Dynamical Systems as a Core Computational Primitive *Shiva Kaul*
- Ratio Trace Formulation of Wasserstein Discriminant Analysis *Hexuan Liu, Yunfeng Cai, You-Lin Chen, Ping Li*
- PAC-Bayes Analysis Beyond the Usual Bounds *Omar Rivasplata, Ilja Kuzborskij, Csaba Szepesvari, John Shawe-Taylor*
- Few-shot Visual Reasoning with Meta-Analogical Contrastive Learning *Youngsung Kim, Jinwoo Shin, Eunho Yang, Sung Ju Hwang*
- MPNet: Masked and Permuted Pre-training for Language Understanding *Kaitao Song, Xu Tan, Tao Qin, Jianfeng Lu, Tie-Yan Liu*
- Reinforcement Learning with Feedback Graphs *Christoph Dann, Yishay Mansour, Mehryar Mohri, Ayush Sekhari, Karthik Sridharan*
- Zap Q-Learning With Nonlinear Function Approximation *Shuhang Chen, Adithya M Devraj, Fan Lu, Ana Busic, Sean Meyn*
- Lipschitz-Certifiable Training with a Tight Outer Bound *Sungyoon Lee, Jaewook Lee, Saerom Park*
- Fast Adaptive Non-Monotone Submodular Maximization Subject to a Knapsack Constraint *Georgios Amanatidis, Federico Fusco, Philip Lazos, Stefano Leonardi, Rebecca Reiffenhäuser*
- Conformal Symplectic and Relativistic Optimization *Guilherme Franca, Jeremias Sulam, Daniel Robinson, Rene Vidal*
- Bayes Consistency vs. H-Consistency: The Interplay between Surrogate Loss Functions and the Scoring Function Class *Mingyuan Zhang, Shivani Agarwal*
- Inverting Gradients - How easy is it to break privacy in federated learning? *Jonas Geiping, Hartmut Bauermeister, Hannah Dröge, Michael Moeller*
- Dynamic allocation of limited memory resources in reinforcement learning *Nisheet Patel, Luigi Acerbi, Alexandre Pouget*
- CryptoNAS: Private Inference on a ReLU Budget *Zahra Ghodsi, Akshaj Kumar Veldanda, Brandon Reagan, Siddharth Garg*
- A Stochastic Path Integral Differential Estimator *Expectation Maximization Algorithm Gersende Fort, Eric Moulines, Hoi-To Wai*
- CHIP: A Hawkes Process Model for Continuous-time Networks with Scalable and Consistent Estimation *Makan Arastuie, Subhadeep Paul, Kevin Xu*
- SAC: Accelerating and Structuring Self-Attention via Sparse Adaptive Connection *Xiaoya Li, Yuxian Meng, Mingxin Zhou, Qinghong Han, Fei Wu, Jiwei Li*
- Design Space for Graph Neural Networks *Jiaxuan You, Zhitao Ying, Jure Leskovec*
- HiFi-GAN: Generative Adversarial Networks for Efficient and High Fidelity Speech Synthesis *Jungil Kong, Jaehyeon Kim, Jaekyoung Bae*
- Unbalanced Sobolev Descent *Youssef Mroueh, Mattia Rigotti*
- Identifying Mislabeled Data using the Area Under the Margin Ranking *Geoff Pleiss, Tianyi Zhang, Ethan Elenberg, Kilian Q. Weinberger*
- Combining Deep Reinforcement Learning and Search for Imperfect-Information Games *Noam Brown, Anton Bakhtin, Adam Lerer, Qucheng Gong*
- High-Throughput Synchronous Deep RL *Iou-Jen Liu, Raymond Yeh, Alexander Schwing*

- [Contrastive Learning with Adversarial Examples](#) *Chih-Hui Ho, Nuno Nvasconcelos*
- [Mixed Hamiltonian Monte Carlo for Mixed Discrete and Continuous Variables](#) *Guangyao Zhou*
- [Adversarial Sparse Transformer for Time Series Forecasting](#) *Sifan Wu, Xi Xiao, Qianggang Ding, Peilin Zhao, Ying Wei, Junzhou Huang*
- [The Surprising Simplicity of the Early-Time Learning Dynamics of Neural Networks](#) *Wei Hu, Lechao Xiao, Ben Adlam, Jeffrey Pennington*
- [CLEARER: Multi-Scale Neural Architecture Search for Image Restoration](#) *Yuanbiao Gou, Boyun Li, Zitao Liu, Songfan Yang, Xi Peng*
- [Hierarchical Gaussian Process Priors for Bayesian Neural Network Weights](#) *Theofanis Karaletsos, Thang D. Bui*
- [Compositional Explanations of Neurons](#) *Jesse Mu, Jacob Andreas*
- [Calibrated Reliable Regression using Maximum Mean Discrepancy](#) *Peng Cui, Wenbo Hu, Jun Zhu*
- [Directional convergence and alignment in deep learning](#) *Ziwei Ji, Matus Telgarsky*
- [Functional Regularization for Representation Learning: A Unified Theoretical Perspective](#) *Siddhant Garg, Yingyu Liang*
- [Provably Efficient Online Hyperparameter Optimization with Population-Based Bandits](#) *Jack Parker-Holder, Vu Nguyen, Stephen J. Roberts*
- [Understanding Global Feature Contributions With Additive Importance Measures](#) *Ian Covert, Scott M. Lundberg, Su-In Lee*
- [Online Non-Convex Optimization with Imperfect Feedback](#) *Amélie Héliou, Matthieu Martin, Panayotis Mertikopoulos, Thibaud Rahier*
- [Co-Tuning for Transfer Learning](#) *Kaichao You, Zhi Kou, Mingsheng Long, Jianmin Wang*
- [Multifaceted Uncertainty Estimation for Label-Efficient Deep Learning](#) *Weishi Shi, Xujiang Zhao, Feng Chen, Qi Yu*
- [Continuous Surface Embeddings](#) *Natalia Neverova, David Novotny, Marc Szafraniec, Vasil Khalidov, Patrick Labatut, Andrea Vedaldi*
- [Succinct and Robust Multi-Agent Communication With Temporal Message Control](#) *Sai Qian Zhang, Qi Zhang, Jieyu Lin*
- [Big Bird: Transformers for Longer Sequences](#) *Manzil Zaheer, Guru Guruganesh, Kumar Avinava Dubey, Joshua Ainslie, Chris Alberti, Santiago Ontanon, Philip Pham, Anirudh Ravula, Qifan Wang, Li Yang, Amr Ahmed*
- [Neural Execution Engines: Learning to Execute Subroutines](#) *Yujun Yan, Kevin Swersky, Danai Koutra, Parthasarathy Ranganathan, Milad Hashemi*
- [Random Reshuffling: Simple Analysis with Vast Improvements](#) *Konstantin Mishchenko, Ahmed Khaled Ragab Bayoumi, Peter Richtarik*
- [Long-Horizon Visual Planning with Goal-Conditioned Hierarchical Predictors](#) *Karl Pertsch, Oleh Rybkin, Frederik Ebert, Shenghao Zhou, Dinesh Jayaraman, Chelsea Finn, Sergey Levine*
- [Statistical Optimal Transport posed as Learning Kernel Embedding](#) *Saketha Nath Jagarlapudi, Pratik Kumar Jawanpuria*
- [Dual-Resolution Correspondence Networks](#) *Xinghui Li, Kai Han, Shuda Li, Victor Prisacariu*
- [Advances in Black-Box VI: Normalizing Flows, Importance Weighting, and Optimization](#) *Abhinav Agrawal, Daniel R. Sheldon, Justin Domke*
- [f-Divergence Variational Inference](#) *Neng Wan, Dapeng Li, NAIRA HOVAKIMYAN*
- [Unfolding recurrence by Green's functions for optimized reservoir computing](#) *Sandra Nestler, Christian Keup,*

David Dahmen, Matthieu Gilson, Holger Rauhut, Moritz Helias

- [The Dilemma of TriHard Loss and an Element-Weighted TriHard Loss for Person Re-Identification](#) *Yihao Lv, Youzhi Gu, Liu Xinggao*
- [Disentangling by Subspace Diffusion](#) *David Pfau, Irina Higgins, Alex Botev, Sébastien Racanière*
- [Towards Neural Programming Interfaces](#) *Zachary Brown, Nathaniel Robinson, David Wingate, Nancy Fulda*
- [Discovering Symbolic Models from Deep Learning with Inductive Biases](#) *Miles Cranmer, Alvaro Sanchez Gonzalez, Peter Battaglia, Rui Xu, Kyle Cranmer, David Spergel, Shirley Ho*
- [Real World Games Look Like Spinning Tops](#) *Wojciech M. Czarnecki, Gauthier Gidel, Brendan Tracey, Karl Tuyls, Shayegan Omidshafiei, David Balduzzi, Max Jaderberg*
- [Cooperative Heterogeneous Deep Reinforcement Learning](#) *Han Zheng, Pengfei Wei, Jing Jiang, Guodong Long, Qinghua Lu, Chengqi Zhang*
- [Mitigating Forgetting in Online Continual Learning via Instance-Aware Parameterization](#) *Hung-Jen Chen, An-Chieh Cheng, Da-Cheng Juan, Wei Wei, Min Sun*
- [ImpatientCapsAndRuns: Approximately Optimal Algorithm Configuration from an Infinite Pool](#) *Gellert Weisz, András György, Wei-I Lin, Devon Graham, Kevin Leyton-Brown, Csaba Szepesvari, Brendan Lucier*
- [Dense Correspondences between Human Bodies via Learning Transformation Synchronization on Graphs](#) *Xiangru Huang, Haitao Yang, Etienne Vouga, Qixing Huang*
- [Reasoning about Uncertainties in Discrete-Time Dynamical Systems using Polynomial Forms](#) *Sriram Sankaranarayanan, Yi Chou, Eric Goubault, Sylvie Putot*
- [Applications of Common Entropy for Causal Inference](#) *Murat Kocaoglu, Sanjay Shakkottai, Alexandros G. Dimakis, Constantine Caramanis, Sriram Vishwanath*
- [SGD with shuffling: optimal rates without component convexity and large epoch requirements](#) *Kwangjun Ahn, Chulhee Yun, Suvrit Sra*
- [Unsupervised Joint k-node Graph Representations with Compositional Energy-Based Models](#) *Leonardo Cotta, Carlos H. C. Teixeira, Ananthram Swami, Bruno Ribeiro*
- [Neural Manifold Ordinary Differential Equations](#) *Aaron Lou, Derek Lim, Isay Katsman, Leo Huang, Qingxuan Jiang, Ser Nam Lim, Christopher M. De Sa*
- [CO-Optimal Transport](#) *Vayer Titouan, Ievgen Redko, Rémi Flamary, Nicolas Courty*
- [Continuous Meta-Learning without Tasks](#) *James Harrison, Apoorva Sharma, Chelsea Finn, Marco Pavone*
- [A mathematical theory of cooperative communication](#) *Pei Wang, Junqi Wang, Pushpi Paranamana, Patrick Shafto*
- [Penalized Langevin dynamics with vanishing penalty for smooth and log-concave targets](#) *Avetik Karagulyan, Arnak Dalalyan*
- [Learning Invariances in Neural Networks from Training Data](#) *Gregory Benton, Marc Finzi, Pavel Izmailov, Andrew G. Wilson*
- [A Finite-Time Analysis of Two Time-Scale Actor-Critic Methods](#) *Yue Frank Wu, Weitong ZHANG, Pan Xu, Quanquan Gu*
- [Pruning Filter in Filter](#) *Fanxu Meng, Hao Cheng, Ke Li, Huixiang Luo, Xiaowei Guo, Guangming Lu, Xing Sun*
- [Learning to Mutate with Hypergradient Guided Population](#) *Zhiqiang Tao, Yaliang Li, Bolin Ding, Ce Zhang, Jingren Zhou, Yun Fu*
- [A convex optimization formulation for multivariate regression](#) *Yunzhang Zhu*
- [Online Meta-Critic Learning for Off-Policy Actor-Critic Methods](#) *Wei Zhou, Yiyang Li, Yongxin Yang, Huaimin Wang, Timothy Hospedales*
- [The All-or-Nothing Phenomenon in Sparse Tensor PCA](#) *Jonathan Niles-Weed, Ilias Zadik*

- [Synthesize, Execute and Debug: Learning to Repair for Neural Program Synthesis](#) Kavi Gupta, Peter Ebert Christensen, Xinyun Chen, Dawn Song
- [ARMA Nets: Expanding Receptive Field for Dense Prediction](#) Jiahao Su, Shiqi Wang, Furong Huang
- [Diversity-Guided Multi-Objective Bayesian Optimization With Batch Evaluations](#) Mina Konakovic Lukovic, Yunsheng Tian, Wojciech Matusik
- [SOLOv2: Dynamic and Fast Instance Segmentation](#) Xinlong Wang, Rufeng Zhang, Tao Kong, Lei Li, Chunhua Shen
- [Robust Recovery via Implicit Bias of Discrepant Learning Rates for Double Over-parameterization](#) Chong You, Zhihui Zhu, Qing Qu, Yi Ma
- [Axioms for Learning from Pairwise Comparisons](#) Ritesh Noothigattu, Dominik Peters, Ariel D. Procaccia
- [Continuous Regularized Wasserstein Barycenters](#) Lingxiao Li, Aude Genevay, Mikhail Yurochkin, Justin M. Solomon
- [Spectral Temporal Graph Neural Network for Multivariate Time-series Forecasting](#) Defu Cao, Yujing Wang, Juanyong Duan, Ce Zhang, Xia Zhu, Congrui Huang, Yunhai Tong, Bixiong Xu, Jing Bai, Jie Tong, Qi Zhang
- [Online Multitask Learning with Long-Term Memory](#) Mark Herbster, Stephen Pasteris, Lisa Tse
- [Fewer is More: A Deep Graph Metric Learning Perspective Using Fewer Proxies](#) Yuehua Zhu, Muli Yang, Cheng Deng, Wei Liu
- [Adaptive Graph Convolutional Recurrent Network for Traffic Forecasting](#) LEI BAI, Lina Yao, Can Li, Xianzhi Wang, Can Wang
- [On Reward-Free Reinforcement Learning with Linear Function Approximation](#) Ruosong Wang, Simon S. Du, Lin Yang, Russ R. Salakhutdinov
- [Robustness of Community Detection to Random Geometric Perturbations](#) Sandrine Peche, Vianney Perchet
- [Learning outside the Black-Box: The pursuit of interpretable models](#) Jonathan Crabbe, Yao Zhang, William Zame, Mihaela van der Schaar
- [Breaking Reversibility Accelerates Langevin Dynamics for Non-Convex Optimization](#) Xuefeng GAO, Mert Gurbuzbalaban, Lingjiong Zhu
- [Robust large-margin learning in hyperbolic space](#) Melanie Weber, Manzil Zaheer, Ankit Singh Rawat, Aditya K. Menon, Sanjiv Kumar
- [Replica-Exchange Nos'e-Hoover Dynamics for Bayesian Learning on Large Datasets](#) Rui Luo, Qiang Zhang, Yaodong Yang, Jun Wang
- [Adversarially Robust Few-Shot Learning: A Meta-Learning Approach](#) Micah Goldblum, Liam Fowl, Tom Goldstein
- [Neural Anisotropy Directions](#) Guillermo Ortiz-Jimenez, Apostolos Modas, Seyed-Mohsen Moosavi, Pascal Frossard
- [Digraph Inception Convolutional Networks](#) Zekun Tong, Yuxuan Liang, Changsheng Sun, Xinke Li, David Rosenblum, Andrew Lim
- [PAC-Bayesian Bound for the Conditional Value at Risk](#) Zakaria Mhammedi, Benjamin Guedj, Robert C. Williamson
- [Stochastic Stein Discrepancies](#) Jackson Gorham, Anant Raj, Lester Mackey
- [On the Role of Sparsity and DAG Constraints for Learning Linear DAGs](#) Ignavier Ng, AmirEmad Ghassami, Kun Zhang
- [Cream of the Crop: Distilling Prioritized Paths For One-Shot Neural Architecture Search](#) Houwen Peng, Hao Du, Hongyuan Yu, Qi Li, Jing Liao, Jianlong Fu
- [Fair Multiple Decision Making Through Soft Interventions](#) Yaowei Hu, Yongkai Wu, Lu Zhang, Xintao Wu

- [Representation Learning for Integrating Multi-domain Outcomes to Optimize Individualized Treatment](#) *Yuan Chen, Donglin Zeng, Tianchen Xu, Yuanjia Wang*
- [Learning to Play No-Press Diplomacy with Best Response Policy Iteration](#) *Thomas Anthony, Tom Eccles, Andrea Tacchetti, János Kramár, Ian Gemp, Thomas Hudson, Nicolas Porcel, Marc Lanctot, Julien Perolat, Richard Everett, Satinder Singh, Thore Graepel, Yoram Bachrach*
- [Inverse Learning of Symmetries](#) *Mario Wieser, Sonali Parbhoo, Aleksander Wieczorek, Volker Roth*
- [DiffGCN: Graph Convolutional Networks via Differential Operators and Algebraic Multigrid Pooling](#) *Moshe Eliasof, Eran Treister*
- [Distributed Newton Can Communicate Less and Resist Byzantine Workers](#) *Avishek Ghosh, Raj Kumar Maity, Arya Mazumdar*
- [Efficient Nonmyopic Bayesian Optimization via One-Shot Multi-Step Trees](#) *Shali Jiang, Daniel Jiang, Maximilian Balandat, Brian Karrer, Jacob Gardner, Roman Garnett*
- [Effective Diversity in Population Based Reinforcement Learning](#) *Jack Parker-Holder, Aldo Pacchiano, Krzysztof M. Choromanski, Stephen J. Roberts*
- [Elastic-InfoGAN: Unsupervised Disentangled Representation Learning in Class-Imbalanced Data](#) *Utkarsh Ojha, Krishna Kumar Singh, Cho-Jui Hsieh, Yong Jae Lee*
- [Direct Policy Gradients: Direct Optimization of Policies in Discrete Action Spaces](#) *Guy Lorberbom, Chris J. Maddison, Nicolas Heess, Tamir Hazan, Daniel Tarlow*
- [Hybrid Models for Learning to Branch](#) *Prateek Gupta, Maxime Gasse, Elias Khalil, Pawan Mudigonda, Andrea Lodi, Yoshua Bengio*
- [WoodFisher: Efficient Second-Order Approximation for Neural Network Compression](#) *Sidak Pal Singh, Dan Alistarh*
- [Bi-level Score Matching for Learning Energy-based Latent Variable Models](#) *Fan Bao, Chongxuan LI, Kun Xu, Hang Su, Jun Zhu, Bo Zhang*
- [Counterfactual Contrastive Learning for Weakly-Supervised Vision-Language Grounding](#) *Zhu Zhang, Zhou Zhao, Zhijie Lin, Jieming zhu, Xiuqiang He*
- [Decision trees as partitioning machines to characterize their generalization properties](#) *Jean-Samuel Leboeuf, Frédéric LeBlanc, Mario Marchand*
- [Learning to Prove Theorems by Learning to Generate Theorems](#) *Mingzhe Wang, Jia Deng*
- [3D Self-Supervised Methods for Medical Imaging](#) *Aiham Taleb, Winfried Loetzsch, Noel Danz, Julius Severin, Thomas Gaertner, Benjamin Bergner, Christoph Lippert*
- [Bayesian filtering unifies adaptive and non-adaptive neural network optimization methods](#) *Laurence Aitchison*
- [Worst-Case Analysis for Randomly Collected Data](#) *Justin Chen, Gregory Valiant, Paul Valiant*
- [Truthful Data Acquisition via Peer Prediction](#) *Yiling Chen, Yiheng Shen, Shuran Zheng*
- [Learning Robust Decision Policies from Observational Data](#) *Muhammad Osama, Dave Zachariah, Peter Stoica*
- [Byzantine Resilient Distributed Multi-Task Learning](#) *Jiani Li, Waseem Abbas, Xenofon Koutsoukos*
- [Reinforcement Learning in Factored MDPs: Oracle-Efficient Algorithms and Tighter Regret Bounds for the Non-Episodic Setting](#) *Ziping Xu, Ambuj Tewari*
- [Improving model calibration with accuracy versus uncertainty optimization](#) *Ranganath Krishnan, Omesh Tickoo*
- [The Convolution Exponential and Generalized Sylvester Flows](#) *Emiel Hoogeboom, Victor Garcia Satorras, Jakub Tomczak, Max Welling*

- [An Improved Analysis of Stochastic Gradient Descent with Momentum](#) Yanli Liu, Yuan Gao, Wotao Yin
- [Precise expressions for random projections: Low-rank approximation and randomized Newton](#) Michal Dereziński, Feynman T. Liang, Zhenyu Liao, Michael W. Mahoney
- [The MAGICAL Benchmark for Robust Imitation](#) Sam Toyer, Rohin Shah, Andrew Critch, Stuart Russell
- [X-CAL: Explicit Calibration for Survival Analysis](#) Mark Goldstein, Xintian Han, Aahlad Puli, Adler Perotte, Rajesh Ranganath
- [Decentralized Accelerated Proximal Gradient Descent](#) Haishan Ye, Ziang Zhou, Luo Luo, Tong Zhang
- [Making Non-Stochastic Control \(Almost\) as Easy as Stochastic](#) Max Simchowitz
- [BERT Loses Patience: Fast and Robust Inference with Early Exit](#) Wangchunshu Zhou, Canwen Xu, Tao Ge, Julian McAuley, Ke Xu, Furu Wei
- [Optimal and Practical Algorithms for Smooth and Strongly Convex Decentralized Optimization](#) Dmitry Kovalev, Adil Salim, Peter Richtarik
- [BAIL: Best-Action Imitation Learning for Batch Deep Reinforcement Learning](#) Xinyue Chen, Zijian Zhou, Zheng Wang, Che Wang, Yanqiu Wu, Keith Ross
- [Regularizing Towards Permutation Invariance In Recurrent Models](#) Edo Cohen-Karlik, Avichai Ben David, Amir Globerson
- [What Did You Think Would Happen? Explaining Agent Behaviour through Intended Outcomes](#) Herman Yau, Chris Russell, Simon Hadfield
- [Batch normalization provably avoids ranks collapse for randomly initialised deep networks](#) Hadi Daneshmand, Jonas Kohler, Francis Bach, Thomas Hofmann, Aurelien Lucchi
- [Choice Bandits](#) Arpit Agarwal, Nicholas Johnson, Shivani Agarwal
- [What if Neural Networks had SVDs?](#) Alexander Mathiasen, Frederik Hvilshøj, Jakob Rødsgaard Jørgensen, Anshul Nasery, Davide Mottin
- [A Matrix Chernoff Bound for Markov Chains and Its Application to Co-occurrence Matrices](#) Jiezhong Qiu, Chi Wang, Ben Liao, Richard Peng, Jie Tang
- [CoMIR: Contrastive Multimodal Image Representation for Registration](#) Nicolas Pielawski, Elisabeth Wetzer, Johan Öfverstedt, Jiahao Lu, Carolina Wählby, Joakim Lindblad, Natasa Sladoje
- [Ensuring Fairness Beyond the Training Data](#) Debmalaya Mandal, Samuel Deng, Suman Jana, Jeannette Wing, Daniel J. Hsu
- [How do fair decisions fare in long-term qualification?](#) Xueru Zhang, Ruibo Tu, Yang Liu, Mingyan Liu, Hedvig Kjellström, Kun Zhang, Cheng Zhang
- [Pre-training via Paraphrasing](#) Mike Lewis, Marjan Ghazvininejad, Gargi Ghosh, Armen Aghajanyan, Sida Wang, Luke Zettlemoyer
- [GCN meets GPU: Decoupling “When to Sample” from “How to Sample”](#) Morteza Ramezani, Weilin Cong, Mehrdad Mahdavi, Anand Sivasubramaniam, Mahmut Kandemir
- [Continual Learning of a Mixed Sequence of Similar and Dissimilar Tasks](#) Zixuan Ke, Bing Liu, Xingchang Huang
- [All your loss are belong to Bayes](#) Christian Walder, Richard Nock
- [HAWQ-V2: Hessian Aware trace-Weighted Quantization of Neural Networks](#) Zhen Dong, Zhewei Yao, Daiyaan Arfeen, Amir Gholami, Michael W. Mahoney, Kurt Keutzer
- [Sample-Efficient Reinforcement Learning of Undercomplete POMDPs](#) Chi Jin, Sham Kakade, Akshay Krishnamurthy, Qinghua Liu
- [Non-Convex SGD Learns Halfspaces with Adversarial Label Noise](#) Ilias Diakonikolas, Vasilis Kontonis, Christos Tzamos, Nikos Zarifis

- [A Tight Lower Bound and Efficient Reduction for Swap Regret](#) *Shinji Ito*
- [DisCor: Corrective Feedback in Reinforcement Learning via Distribution Correction](#) *Aviral Kumar, Abhishek Gupta, Sergey Levine*
- [OTLDA: A Geometry-aware Optimal Transport Approach for Topic Modeling](#) *Viet Huynh, He Zhao, Dinh Phung*
- [Measuring Robustness to Natural Distribution Shifts in Image Classification](#) *Rohan Taori, Achal Dave, Vaishaal Shankar, Nicholas Carlini, Benjamin Recht, Ludwig Schmidt*
- [Can I Trust My Fairness Metric? Assessing Fairness with Unlabeled Data and Bayesian Inference](#) *Disi Ji, Padhraic Smyth, Mark Steyvers*
- [RandAugment: Practical Automated Data Augmentation with a Reduced Search Space](#) *Ekin Dogus Cubuk, Barret Zoph, Jon Shlens, Quoc Le*
- [Asymptotic normality and confidence intervals for derivatives of 2-layers neural network in the random features model](#) *Yiwei Shen, Pierre C Bellec*
- [DisARM: An Antithetic Gradient Estimator for Binary Latent Variables](#) *Zhe Dong, Andriy Mnih, George Tucker*
- [Variational Inference for Graph Convolutional Networks in the Absence of Graph Data and Adversarial Settings](#) *Pantelis Elinas, Edwin V. Bonilla, Louis Tiao*
- [Supervised Contrastive Learning](#) *Prannay Khosla, Piotr Teterwak, Chen Wang, Aaron Sarna, Yonglong Tian, Phillip Isola, Aaron Maschiot, Ce Liu, Dilip Krishnan*
- [Learning Optimal Representations with the Decodable Information Bottleneck](#) *Yann Dubois, Douwe Kiela, David J. Schwab, Ramakrishna Vedantam*
- [Meta-trained agents implement Bayes-optimal agents](#) *Vladimir Mikulik, Grégoire Delétang, Tom McGrath, Tim Genewein, Miljan Martic, Shane Legg, Pedro Ortega*
- [Learning Agent Representations for Ice Hockey](#) *Guiliang Liu, Oliver Schulte, Pascal Poupart, Mike Rudd, Mehrsan Javan*
- [Weak Form Generalized Hamiltonian Learning](#) *Kevin Course, Trefor Evans, Prasanth Nair*
- [Neural Non-Rigid Tracking](#) *Aljaz Bozic, Pablo Palafox, Michael Zollhöfer, Angela Dai, Justus Thies, Matthias Niessner*
- [Collegial Ensembles](#) *Etai Littwin, Ben Myara, Sima Sabah, Joshua Susskind, Shuangfei Zhai, Oren Golan*
- [ICNet: Intra-saliency Correlation Network for Co-Saliency Detection](#) *Wen-Da Jin, Jun Xu, Ming-Ming Cheng, Yi Zhang, Wei Guo*
- [Improved Variational Bayesian Phylogenetic Inference with Normalizing Flows](#) *Cheng Zhang*
- [Deep Metric Learning with Spherical Embedding](#) *Dingyi Zhang, Yingming Li, Zhongfei Zhang*
- [Preference-based Reinforcement Learning with Finite-Time Guarantees](#) *Yichong Xu, Ruosong Wang, Lin Yang, Aarti Singh, Artur Dubrawski*
- [AdaBelief Optimizer: Adapting Stepsizes by the Belief in Observed Gradients](#) *Juntang Zhuang, Tommy Tang, Yifan Ding, Sekhar C. Tatikonda, Nicha Dvornek, Xenophon Papademetris, James Duncan*
- [Interpretable Sequence Learning for Covid-19 Forecasting](#) *Sercan Arik, Chun-Liang Li, Jinsung Yoon, Rajarishi Sinha, Arkady Epshteyn, Long Le, Vikas Menon, Shashank Singh, Leyou Zhang, Martin Nikoltchev, Yash Sonthalia, Hootan Nakhost, Elli Kanal, Tomas Pfister*
- [Off-policy Policy Evaluation For Sequential Decisions Under Unobserved Confounding](#) *Hongseok Namkoong, Ramtin Keramati, Steve Yadlowsky, Emma Brunskill*
- [Modern Hopfield Networks and Attention for Immune Repertoire Classification](#) *Michael Widrich, Bernhard Schäfl, Milena Pavlović, Hubert Ramsauer, Lukas Gruber, Markus Holzleitner, Johannes Brandstetter, Geir Kjetil Sandve, Victor Greiff, Sepp Hochreiter, Günter Klambauer*

- [One Ring to Rule Them All: Certifiably Robust Geometric Perception with Outliers](#) *Heng Yang, Luca Carlone*
- [Task-Robust Model-Agnostic Meta-Learning](#) *Liam Collins, Aryan Mokhtari, Sanjay Shakkottai*
- [R-learning in actor-critic model offers a biologically relevant mechanism for sequential decision-making](#) *Sergey Shuvaev, Sarah Starosta, Duda Kvitsiani, Adam Kepecs, Alexei Koulakov*
- [Revisiting Frank-Wolfe for Polytopes: Strict Complementarity and Sparsity](#) *Dan Garber*
- [Fast Convergence of Langevin Dynamics on Manifold: Geodesics meet Log-Sobolev](#) *Xiao Wang, Qi Lei, Ioannis Panageas*
- [Tensor Completion Made Practical](#) *Allen Liu, Ankur Moitra*
- [Optimization and Generalization Analysis of Transduction through Gradient Boosting and Application to Multi-scale Graph Neural Networks](#) *Kenta Oono, Taiji Suzuki*
- [Content Provider Dynamics and Coordination in Recommendation Ecosystems](#) *Omer Ben-Porat, Itay Rosenberg, Moshe Tennenholtz*
- [Almost Surely Stable Deep Dynamics](#) *Nathan Lawrence, Philip Loewen, Michael Forbes, Johan Backstrom, Bhushan Gopaluni*
- [Experimental design for MRI by greedy policy search](#) *Tim Bakker, Herke van Hoof, Max Welling*
- [Expert-Supervised Reinforcement Learning for Offline Policy Learning and Evaluation](#) *Aaron Sonabend, Junwei Lu, Leo Anthony Celi, Tianxi Cai, Peter Szolovits*
- [ColdGANs: Taming Language GANs with Cautious Sampling Strategies](#) *Thomas Scialom, Paul-Alexis Dray, Sylvain Lamprier, Benjamin Piwowarski, Jacopo Staiano*
- [Hedging in games: Faster convergence of external and swap regrets](#) *Xi Chen, Binghui Peng*
- [The Origins and Prevalence of Texture Bias in Convolutional Neural Networks](#) *Katherine Hermann, Ting Chen, Simon Kornblith*
- [Time-Reversal Symmetric ODE Network](#) *In Huh, Eunho Yang, Sung Ju Hwang, Jinwoo Shin*
- [Provable Overlapping Community Detection in Weighted Graphs](#) *Jimit Majmudar, Stephen Vavasis*
- [Fast Unbalanced Optimal Transport on a Tree](#) *Ryoma Sato, Makoto Yamada, Hisashi Kashima*
- [Acceleration with a Ball Optimization Oracle](#) *Yair Carmon, Arun Jambulapati, Qijia Jiang, Yujia Jin, Yin Tat Lee, Aaron Sidford, Kevin Tian*
- [Avoiding Side Effects By Considering Future Tasks](#) *Victoria Krakovna, Laurent Orseau, Richard Ngo, Miljan Martic, Shane Legg*
- [Handling Missing Data with Graph Representation Learning](#) *Jiaxuan You, Xiaobai Ma, Yi Ding, Mykel J. Kochenderfer, Jure Leskovec*
- [Improving Auto-Augment via Augmentation-Wise Weight Sharing](#) *Keyu Tian, Chen Lin, Ming Sun, Luping Zhou, Junjie Yan, Wanli Ouyang*
- [MMA Regularization: Decorrelating Weights of Neural Networks by Maximizing the Minimal Angles](#) *Zhennan Wang, Canqun Xiang, Wenbin Zou, Chen Xu*
- [HRN: A Holistic Approach to One Class Learning](#) *Wenpeng Hu, Mengyu Wang, Qi Qin, Jinwen Ma, Bing Liu*
- [The Generalized Lasso with Nonlinear Observations and Generative Priors](#) *Zhaoqiang Liu, Jonathan Scarlett*
- [Fair regression via plug-in estimator and recalibration with statistical guarantees](#) *Evgenii Chzhen, Christophe Denis, Mohamed Hebiri, Luca Oneto, Massimiliano Pontil*
- [Modeling Shared responses in Neuroimaging Studies through MultiView ICA](#) *Hugo Richard, Luigi Gresele, Aapo Hyvarinen, Bertrand Thirion, Alexandre Gramfort, Pierre Ablin*
- [Efficient Planning in Large MDPs with Weak Linear Function Approximation](#) *Roshan Shariff, Csaba Szepesvari*
- [Efficient Learning of Generative Models via Finite-Difference Score Matching](#) *Tianyu Pang, Kun Xu,*

Chongxuan LI, Yang Song, Stefano Ermon, Jun Zhu

- [Semialgebraic Optimization for Lipschitz Constants of ReLU Networks](#) Tong Chen, Jean B. Lasserre, Victor Magron, Edouard Pauwels
- [Linear-Sample Learning of Low-Rank Distributions](#) Ayush Jain, Alon Orlitsky
- [Transferable Calibration with Lower Bias and Variance in Domain Adaptation](#) Ximei Wang, Mingsheng Long, Jianmin Wang, Michael Jordan
- [Generalization bound of globally optimal non-convex neural network training: Transportation map estimation by infinite dimensional Langevin dynamics](#) Taiji Suzuki
- [Online Bayesian Goal Inference for Boundedly Rational Planning Agents](#) Tan Zhi-Xuan, Jordyn Mann, Tom Silver, Josh Tenenbaum, Vikash Mansinghka
- [BayReL: Bayesian Relational Learning for Multi-omics Data Integration](#) Ehsan Hajiramezanali, Arman Hasanzadeh, Nick Duffield, Krishna Narayanan, Xiaoning Qian
- [Weakly Supervised Deep Functional Maps for Shape Matching](#) Abhishek Sharma, Maks Ovsjanikov
- [Domain Adaptation with Conditional Distribution Matching and Generalized Label Shift](#) Remi Tachet des Combes, Han Zhao, Yu-Xiang Wang, Geoffrey J. Gordon
- [Rethinking the Value of Labels for Improving Class-Imbalanced Learning](#) Yuzhe Yang, Zhi Xu
- [Provably Robust Metric Learning](#) Lu Wang, Xuanqing Liu, Jinfeng Yi, Yuan Jiang, Cho-Jui Hsieh
- [Iterative Deep Graph Learning for Graph Neural Networks: Better and Robust Node Embeddings](#) Yu Chen, Lingfei Wu, Mohammed Zaki
- [COPT: Coordinated Optimal Transport on Graphs](#) Yihe Dong, Will Sawin
- [No Subclass Left Behind: Fine-Grained Robustness in Coarse-Grained Classification Problems](#) Nimit Sohoni, Jared Dunnmon, Geoffrey Angus, Albert Gu, Christopher Ré
- [Model Rubik's Cube: Twisting Resolution, Depth and Width for TinyNets](#) Kai Han, Yunhe Wang, Qiulin Zhang, Wei Zhang, Chunjing XU, Tong Zhang
- [Self-Adaptive Training: beyond Empirical Risk Minimization](#) Lang Huang, Chao Zhang, Hongyang Zhang
- [Effective Dimension Adaptive Sketching Methods for Faster Regularized Least-Squares Optimization](#) Jonathan Lacotte, Mert Pilanci
- [Near-Optimal Comparison Based Clustering](#) Michaël Perrot, Pascal Esser, Debarghya Ghoshdastidar
- [Multi-Task Temporal Shift Attention Networks for On-Device Contactless Vitals Measurement](#) Xin Liu, Josh Fromm, Shwetak Patel, Daniel McDuff
- [A new convergent variant of Q-learning with linear function approximation](#) Diogo Carvalho, Francisco S. Melo, Pedro Santos
- [TaylorGAN: Neighbor-Augmented Policy Update Towards Sample-Efficient Natural Language Generation](#) Chun-Hsing Lin, Siang-Ruei Wu, Hung-yi Lee, Yun-Nung Chen
- [Neural Networks with Small Weights and Depth-Separation Barriers](#) Gal Vardi, Ohad Shamir
- [Untangling tradeoffs between recurrence and self-attention in artificial neural networks](#) Giancarlo Kerg, Bhargav Kanuparthi, Anirudh Goyal ALIAS PARTH GOYAL, Kyle Goyette, Yoshua Bengio, Guillaume Lajoie
- [Dual-Free Stochastic Decentralized Optimization with Variance Reduction](#) Hadrien Hendrikx, Francis Bach, Laurent Massoulié
- [Online Learning in Contextual Bandits using Gated Linear Networks](#) Eren Sezener, Marcus Hutter, David Budden, Jianan Wang, Joel Veness
- [Throughput-Optimal Topology Design for Cross-Silo Federated Learning](#) Othmane MARFOQ, CHUAN XU, Giovanni Neglia, Richard Vidal
- [Quantized Variational Inference](#) Amir Dib

- [Asymptotically Optimal Exact Minibatch Metropolis-Hastings](#) *Ruqi Zhang, A. Feder Cooper, Christopher M. De Sa*
- [Learning Search Space Partition for Black-box Optimization using Monte Carlo Tree Search](#) *Linnan Wang, Rodrigo Fonseca, Yuandong Tian*
- [Feature Shift Detection: Localizing Which Features Have Shifted via Conditional Distribution Tests](#) *Sean Kulinski, Saurabh Bagchi, David I. Inouye*
- [Unifying Activation- and Timing-based Learning Rules for Spiking Neural Networks](#) *Jinseok Kim, Kyungsu Kim, Jae-Joon Kim*
- [Space-Time Correspondence as a Contrastive Random Walk](#) *Allan Jabri, Andrew Owens, Alexei Efros*
- [The Flajolet-Martin Sketch Itself Preserves Differential Privacy: Private Counting with Minimal Space](#) *Adam Smith, Shuang Song, Abhradeep Guha Thakurta*
- [Exponential ergodicity of mirror-Langevin diffusions](#) *Sinho Chewi, Thibaut Le Gouic, Chen Lu, Tyler Maunu, Philippe Rigollet, Austin Stromme*
- [An Efficient Framework for Clustered Federated Learning](#) *Avishek Ghosh, Jichan Chung, Dong Yin, Kannan Ramchandran*
- [Autoencoders that don't overfit towards the Identity](#) *Harald Steck*
- [Polynomial-Time Computation of Optimal Correlated Equilibria in Two-Player Extensive-Form Games with Public Chance Moves and Beyond](#) *Gabriele Farina, Tuomas Sandholm*
- [Parameterized Explainer for Graph Neural Network](#) *Dongsheng Luo, Wei Cheng, Dongkuan Xu, Wenchao Yu, Bo Zong, Haifeng Chen, Xiang Zhang*
- [Recursive Inference for Variational Autoencoders](#) *Minyoung Kim, Vladimir Pavlovic*
- [Flexible mean field variational inference using mixtures of non-overlapping exponential families](#) *Jeffrey Spence*
- [HYDRA: Pruning Adversarially Robust Neural Networks](#) *Vikash Sehwal, Shiqi Wang, Prateek Mittal, Suman Jana*
- [NVAE: A Deep Hierarchical Variational Autoencoder](#) *Arash Vahdat, Jan Kautz*
- [Can Temporal-Difference and Q-Learning Learn Representation? A Mean-Field Theory](#) *Yufeng Zhang, Qi Cai, Zhuoran Yang, Yongxin Chen, Zhaoran Wang*
- [What Do Neural Networks Learn When Trained With Random Labels?](#) *Hartmut Maennel, Ibrahim M. Alabdulmohsin, Ilya O. Tolstikhin, Robert Baldock, Olivier Bousquet, Sylvain Gelly, Daniel Keysers*
- [Counterfactual Prediction for Bundle Treatment](#) *Hao Zou, Peng Cui, Bo Li, Zheyang Shen, Jianxin Ma, Hongxia Yang, Yue He*
- [Beta Embeddings for Multi-Hop Logical Reasoning in Knowledge Graphs](#) *Hongyu Ren, Jure Leskovec*
- [Learning Disentangled Representations and Group Structure of Dynamical Environments](#) *Robin Quessard, Thomas Barrett, William Clements*
- [Learning Linear Programs from Optimal Decisions](#) *Yingcong Tan, Daria Terekhov, Andrew Delong*
- [Wisdom of the Ensemble: Improving Consistency of Deep Learning Models](#) *Lijing Wang, Dipanjan Ghosh, Maria Gonzalez Diaz, Ahmed Farahat, Mahbubul Alam, Chetan Gupta, Jiangzhuo Chen, Madhav Marathe*
- [Universal Function Approximation on Graphs](#) *Rickard Br  el Gabrielsson*
- [Accelerating Reinforcement Learning through GPU Atari Emulation](#) *Steven Dalton, iuri frosio*
- [EvolveGraph: Multi-Agent Trajectory Prediction with Dynamic Relational Reasoning](#) *Jiachen Li, Fan Yang, Masayoshi Tomizuka, Chiho Choi*
- [Comparator-Adaptive Convex Bandits](#) *Dirk van der Hoeven, Ashok Cutkosky, Haipeng Luo*
- [Model-based Reinforcement Learning for Semi-Markov Decision Processes with Neural ODEs](#) *Jianzhun Du,*

Joseph Futoma, Finale Doshi-Velez

- The Adaptive Complexity of Maximizing a Gross Substitutes Valuation *Ron Kupfer, Sharon Qian, Eric Balkanski, Yaron Singer*
- A Robust Functional EM Algorithm for Incomplete Panel Count Data *Alexander Moreno, Zhenke Wu, Jamie Roslyn Yap, Cho Lam, David Wetter, Inbal Nahum-Shani, Walter Dempsey, James M. Rehg*
- Graph Stochastic Neural Networks for Semi-supervised Learning *Haibo Wang, Chuan Zhou, Xin Chen, Jia Wu, Shirui Pan, Jilong Wang*
- Compositional Zero-Shot Learning via Fine-Grained Dense Feature Composition *Dat Huynh, Ehsan Elhamifar*
- A Benchmark for Systematic Generalization in Grounded Language Understanding *Laura Ruis, Jacob Andreas, Marco Baroni, Diane Bouchacourt, Brenden M. Lake*
- Weston-Watkins Hinge Loss and Ordered Partitions *Yutong Wang, Clayton Scott*
- Reinforcement Learning with Augmented Data *Misha Laskin, Kimin Lee, Adam Stooke, Lerrel Pinto, Pieter Abbeel, Aravind Srinivas*
- Towards Minimax Optimal Reinforcement Learning in Factored Markov Decision Processes *Yi Tian, Jian Qian, Suvrit Sra*
- Graduated Assignment for Joint Multi-Graph Matching and Clustering with Application to Unsupervised Graph Matching Network Learning *Runzhong Wang, Junchi Yan, Xiaokang Yang*
- Estimating Training Data Influence by Tracing Gradient Descent *Garima Pruthi, Frederick Liu, Satyen Kale, Mukund Sundararajan*
- Joint Policy Search for Multi-agent Collaboration with Imperfect Information *Yuandong Tian, Qucheng Gong, Yu Jiang*
- Adversarial Bandits with Corruptions: Regret Lower Bound and No-regret Algorithm *lin yang, Mohammad Hajiesmaili, Mohammad Sadegh Talebi, John C. S. Lui, Wing Shing Wong*
- Beta R-CNN: Looking into Pedestrian Detection from Another Perspective *Zixuan Xu, Banghuai Li, Ye Yuan, Anhong Dang*
- Batch Normalization Biases Residual Blocks Towards the Identity Function in Deep Networks *Soham De, Sam Smith*
- Learning Retrospective Knowledge with Reverse Reinforcement Learning *Shangdong Zhang, Vivek Veeriah, Shimon Whiteson*
- Dialog without Dialog Data: Learning Visual Dialog Agents from VQA Data *Michael Cogswell, Jiasen Lu, Rishabh Jain, Stefan Lee, Devi Parikh, Dhruv Batra*
- GCOMB: Learning Budget-constrained Combinatorial Algorithms over Billion-sized Graphs *Sahil Manchanda, AKASH MITTAL, Anuj Dhawan, Sourav Medya, Sayan Ranu, Ambuj Singh*
- A General Large Neighborhood Search Framework for Solving Integer Linear Programs *Jialin Song, ravi lanka, Yisong Yue, Bistra Dilkina*
- A Theoretical Framework for Target Propagation *Alexander Meulemans, Francesco Carzaniga, Johan Suykens, João Sacramento, Benjamin F. Grewe*
- OrganITE: Optimal transplant donor organ offering using an individual treatment effect *Jeroen Berrevoets, James Jordon, Ioana Bica, alexander gimson, Mihaela van der Schaar*
- The Complete Lasso Tradeoff Diagram *Hua Wang, Yachong Yang, Zhiqi Bu, Weijie Su*
- On the universality of deep learning *Emmanuel Abbe, Colin Sandon*
- Regression with reject option and application to kNN *Ahmed Zaoui, Christophe Denis, Mohamed Hebiri*
- The Primal-Dual method for Learning Augmented Algorithms *Etienne Bamas, Andreas Maggiori, Ola*

Svensson

- [FLAMBE: Structural Complexity and Representation Learning of Low Rank MDPs](#) Alekh Agarwal, Sham Kakade, Akshay Krishnamurthy, Wen Sun
- [A Class of Algorithms for General Instrumental Variable Models](#) Niki Kilbertus, Matt J. Kusner, Ricardo Silva
- [Black-Box Ripper: Copying black-box models using generative evolutionary algorithms](#) Antonio Barbalau, Adrian Cosma, Radu Tudor Ionescu, Marius Popescu
- [Bayesian Optimization of Risk Measures](#) Sait Cakmak, Raul Astudillo Marban, Peter Frazier, Enlu Zhou
- [TorsionNet: A Reinforcement Learning Approach to Sequential Conformer Search](#) Tarun Gogineni, Ziping Xu, Exequiel Punzalan, Runxuan Jiang, Joshua Kammeraad, Ambuj Tewari, Paul Zimmerman
- [GRAF: Generative Radiance Fields for 3D-Aware Image Synthesis](#) Katja Schwarz, Yiyi Liao, Michael Niemeyer, Andreas Geiger
- [PIE-NET: Parametric Inference of Point Cloud Edges](#) Xiaogang Wang, Yuelang Xu, Kai Xu, Andrea Tagliasacchi, Bin Zhou, Ali Mahdavi-Amiri, Hao Zhang
- [A Simple Language Model for Task-Oriented Dialogue](#) Ehsan Hosseini-Asl, Bryan McCann, Chien-Sheng Wu, Semih Yavuz, Richard Socher
- [A Continuous-Time Mirror Descent Approach to Sparse Phase Retrieval](#) Fan Wu, Patrick Rebeschini
- [Confidence sequences for sampling without replacement](#) Ian Waudby-Smith, Aaditya Ramdas
- [A mean-field analysis of two-player zero-sum games](#) Carles Domingo-Enrich, Samy Jelassi, Arthur Mensch, Grant Rotskoff, Joan Bruna
- [Leap-Of-Thought: Teaching Pre-Trained Models to Systematically Reason Over Implicit Knowledge](#) Alon Talmor, Oyvind Tafjord, Peter Clark, Yoav Goldberg, Jonathan Berant
- [Pipeline PSRO: A Scalable Approach for Finding Approximate Nash Equilibria in Large Games](#) Stephen Mcaleer, JB Lanier, Roy Fox, Pierre Baldi
- [Improving Sparse Vector Technique with Renyi Differential Privacy](#) Yuqing Zhu, Yu-Xiang Wang
- [Latent Template Induction with Gumbel-CRFs](#) Yao Fu, Chuanqi Tan, Bin Bi, Mosha Chen, Yansong Feng, Alexander Rush
- [Instance Based Approximations to Profile Maximum Likelihood](#) Nima Anari, Moses Charikar, Kirankumar Shiragur, Aaron Sidford
- [Factorizable Graph Convolutional Networks](#) Yiding Yang, Zunlei Feng, Mingli Song, Xinchao Wang
- [Guided Adversarial Attack for Evaluating and Enhancing Adversarial Defenses](#) Gaurang Sriramanan, Sravanti Addepalli, Arya Baburaj, Venkatesh Babu R
- [A Study on Encodings for Neural Architecture Search](#) Colin White, Willie Neiswanger, Sam Nolen, Yash Savani
- [Noise2Same: Optimizing A Self-Supervised Bound for Image Denoising](#) Yaochen Xie, Zhengyang Wang, Shuiwang Ji
- [Early-Learning Regularization Prevents Memorization of Noisy Labels](#) Sheng Liu, Jonathan Niles-Weed, Narges Razavian, Carlos Fernandez-Granda
- [LAPAR: Linearly-Assembled Pixel-Adaptive Regression Network for Single Image Super-resolution and Beyond](#) Wenbo Li, Kun Zhou, Lu Qi, Nianjuan Jiang, Jiangbo Lu, Jiaya Jia
- [Learning Parities with Neural Networks](#) Amit Daniely, Eran Malach
- [Consistent Plug-in Classifiers for Complex Objectives and Constraints](#) Shiv Kumar Tavker, Harish Guruprasad Ramaswamy, Harikrishna Narasimhan
- [Movement Pruning: Adaptive Sparsity by Fine-Tuning](#) Victor Sanh, Thomas Wolf, Alexander Rush
- [Sanity-Checking Pruning Methods: Random Tickets can Win the Jackpot](#) Jingdong Su, Yihang Chen, Tianle

Cai, Tianhao Wu, Ruiqi Gao, Liwei Wang, Jason D. Lee

- [Online Matrix Completion with Side Information](#) *Mark Herbster, Stephen Pasteris, Lisa Tse*
- [Position-based Scaled Gradient for Model Quantization and Pruning](#) *Jangho Kim, KiYoon Yoo, Nojun Kwak*
- [Online Learning with Primary and Secondary Losses](#) *Avrim Blum, Han Shao*
- [Graph Information Bottleneck](#) *Tailin Wu, Hongyu Ren, Pan Li, Jure Leskovec*
- [The Complexity of Adversarially Robust Proper Learning of Halfspaces with Agnostic Noise](#) *Ilias Diakonikolas, Daniel M. Kane, Pasin Manurangsi*
- [Adaptive Online Estimation of Piecewise Polynomial Trends](#) *Dheeraj Baby, Yu-Xiang Wang*
- [RNNPool: Efficient Non-linear Pooling for RAM Constrained Inference](#) *Oindrila Saha, Aditya Kusupati, Harsha Vardhan Simhadri, Manik Varma, Prateek Jain*
- [Agnostic Learning with Multiple Objectives](#) *Corinna Cortes, Mehryar Mohri, Javier Gonzalvo, Dmitry Storcheus*
- [3D Multi-bodies: Fitting Sets of Plausible 3D Human Models to Ambiguous Image Data](#) *Benjamin Biggs, David Novotny, Sebastien Ehrhardt, Hanbyul Joo, Ben Graham, Andrea Vedaldi*
- [Auto-Panoptic: Cooperative Multi-Component Architecture Search for Panoptic Segmentation](#) *Yangxin Wu, Gengwei Zhang, Hang Xu, Xiaodan Liang, Liang Lin*
- [Differentiable Top-k with Optimal Transport](#) *Yujia Xie, Hanjun Dai, Minshuo Chen, Bo Dai, Tuo Zhao, Hongyuan Zha, Wei Wei, Tomas Pfister*
- [Information-theoretic Task Selection for Meta-Reinforcement Learning](#) *Ricardo Luna Gutierrez, Matteo Leonetti*
- [A Limitation of the PAC-Bayes Framework](#) *Roi Livni, Shay Moran*
- [On Completeness-aware Concept-Based Explanations in Deep Neural Networks](#) *Chih-Kuan Yeh, Been Kim, Sercan Arik, Chun-Liang Li, Tomas Pfister, Pradeep Ravikumar*
- [Stochastic Recursive Gradient Descent Ascent for Stochastic Nonconvex-Strongly-Concave Minimax Problems](#) *Luo Luo, Haishan Ye, Zhichao Huang, Tong Zhang*
- [Why Normalizing Flows Fail to Detect Out-of-Distribution Data](#) *Polina Kirichenko, Pavel Izmailov, Andrew G. Wilson*
- [Explaining Naive Bayes and Other Linear Classifiers with Polynomial Time and Delay](#) *Joao Marques-Silva, Thomas Gerspacher, Martin Cooper, Alexey Ignatiev, Nina Narodytska*
- [Unsupervised Translation of Programming Languages](#) *Baptiste Roziere, Marie-Anne Lachaux, Lowik Chanussot, Guillaume Lample*
- [Adversarial Style Mining for One-Shot Unsupervised Domain Adaptation](#) *Yawei Luo, Ping Liu, Tao Guan, Junqing Yu, Yi Yang*
- [Optimally Deceiving a Learning Leader in Stackelberg Games](#) *Georgios Birmpas, Jiarui Gan, Alexandros Hollender, Francisco Marmolejo, Ninad Rajgopal, Alexandros Voudouris*
- [Online Optimization with Memory and Competitive Control](#) *Guanya Shi, Yiheng Lin, Soon-Jo Chung, Yisong Yue, Adam Wierman*
- [IDEAL: Inexact DEcentralized Accelerated Augmented Lagrangian Method](#) *Yossi Arjevani, Joan Bruna, Bugra Can, Mert Gurbuzbalaban, Stefanie Jegelka, Hongzhou Lin*
- [Evolving Graphical Planner: Contextual Global Planning for Vision-and-Language Navigation](#) *Zhiwei Deng, Karthik Narasimhan, Olga Russakovsky*
- [Learning from Failure: De-biasing Classifier from Biased Classifier](#) *Junhyun Nam, Hyuntak Cha, Sungsoo Ahn, Jaeho Lee, Jinwoo Shin*
- [Likelihood Regret: An Out-of-Distribution Detection Score For Variational Auto-encoder](#) *Zhisheng Xiao, Qing*

Yan, Yali Amit

- [Deep Diffusion-Invariant Wasserstein Distributional Classification](#) Sung Woo Park, Dong Wook Shu, Junseok Kwon
- [Finding All \$\epsilon\$ -Good Arms in Stochastic Bandits](#) Blake Mason, Lalit Jain, Ardhendu Tripathy, Robert Nowak
- [Meta-Learning through Hebbian Plasticity in Random Networks](#) Elias Najarro, Sebastian Risi
- [A Computational Separation between Private Learning and Online Learning](#) Mark Bun
- [Top-KAST: Top-K Always Sparse Training](#) Siddhant Jayakumar, Razvan Pascanu, Jack Rae, Simon Osindero, Erich Elsen
- [Meta-Learning with Adaptive Hyperparameters](#) Sungyong Baik, Myungsub Choi, Janghoon Choi, Heewon Kim, Kyoung Mu Lee
- [Tight last-iterate convergence rates for no-regret learning in multi-player games](#) Noah Golowich, Sarath Pattathil, Constantinos Daskalakis
- [Curvature Regularization to Prevent Distortion in Graph Embedding](#) Hongbin Pei, Bingzhe Wei, Kevin Chang, Chunxu Zhang, Bo Yang
- [Perturbing Across the Feature Hierarchy to Improve Standard and Strict Blackbox Attack Transferability](#) Nathan Inkawhich, Kevin Liang, Binghui Wang, Matthew Inkawhich, Lawrence Carin, Yiran Chen
- [Statistical and Topological Properties of Sliced Probability Divergences](#) Kimia Nadjahi, Alain Durmus, L  na  c Chizat, Soheil Kolouri, Shahin Shahrampour, Umut Simsekli
- [Probabilistic Active Meta-Learning](#) Jean Kaddour, Steindor Saemundsson, Marc Deisenroth (he/him)
- [Knowledge Distillation in Wide Neural Networks: Risk Bound, Data Efficiency and Imperfect Teacher](#) Guangda Ji, Zhanxing Zhu
- [Adversarial Attacks on Deep Graph Matching](#) Zijie Zhang, Zeru Zhang, Yang Zhou, Yelong Shen, Ruoming Jin, Dejing Dou
- [The Generalization-Stability Tradeoff In Neural Network Pruning](#) Brian Bartoldson, Ari Morcos, Adrian Barbu, Gordon Erlebacher
- [Gradient-EM Bayesian Meta-Learning](#) Yayi Zou, Xiaoqi Lu
- [Logarithmic Regret Bound in Partially Observable Linear Dynamical Systems](#) Sahin Lale, Kamyar Azizzadenesheli, Babak Hassibi, Anima Anandkumar
- [Linearly Converging Error Compensated SGD](#) Eduard Gorbunov, Dmitry Kovalev, Dmitry Makarenko, Peter Richtarik
- [Canonical 3D Deformer Maps: Unifying parametric and non-parametric methods for dense weakly-supervised category reconstruction](#) David Novotny, Roman Shapovalov, Andrea Vedaldi
- [A Self-Tuning Actor-Critic Algorithm](#) Tom Zahavy, Zhongwen Xu, Vivek Veeriah, Matteo Hessel, Junhyuk Oh, Hado P. van Hasselt, David Silver, Satinder Singh
- [The Cone of Silence: Speech Separation by Localization](#) Teerapat Jenrungrot, Vivek Jayaram, Steve Seitz, Ira Kemelmacher-Shlizerman
- [High-Dimensional Bayesian Optimization via Nested Riemannian Manifolds](#) No  mie Jaquier, Leonel Rozo
- [Train-by-Reconnect: Decoupling Locations of Weights from Their Values](#) Yushi Qiu, Reiji Suda
- [Learning discrete distributions: user vs item-level privacy](#) Yuhan Liu, Ananda Theertha Suresh, Felix Xinnan X. Yu, Sanjiv Kumar, Michael Riley
- [Matrix Completion with Quantified Uncertainty through Low Rank Gaussian Copula](#) Yuxuan Zhao, Madeleine Udell
- [Sparse and Continuous Attention Mechanisms](#) Andr   Martins, Ant  nio Farinhas, Marcos Treviso, Vlad Niculae, Pedro Aguiar, Mario Figueiredo

- [Generalized Focal Loss: Learning Qualified and Distributed Bounding Boxes for Dense Object Detection](#) *Xiang Li, Wenhai Wang, Lijun Wu, Shuo Chen, Xiaolin Hu, Jun Li, Jinhui Tang, Jian Yang*
- [Learning by Minimizing the Sum of Ranked Range](#) *Shu Hu, Yiming Ying, xin wang, Siwei Lyu*
- [Robust Deep Reinforcement Learning against Adversarial Perturbations on State Observations](#) *Huan Zhang, Hongge Chen, Chaowei Xiao, Bo Li, Mingyan Liu, Duane Boning, Cho-Jui Hsieh*
- [Understanding Anomaly Detection with Deep Invertible Networks through Hierarchies of Distributions and Features](#) *Robin Schirrmeyer, Yuxuan Zhou, Tonio Ball, Dan Zhang*
- [Fair Hierarchical Clustering](#) *Sara Ahmadian, Alessandro Epasto, Marina Knittel, Ravi Kumar, Mohammad Mahdian, Benjamin Moseley, Philip Pham, Sergei Vassilvitskii, Yuyan Wang*
- [Self-training Avoids Using Spurious Features Under Domain Shift](#) *Yining Chen, Colin Wei, Ananya Kumar, Tengyu Ma*
- [Improving Online Rent-or-Buy Algorithms with Sequential Decision Making and ML Predictions](#) *Shom Banerjee*
- [CircleGAN: Generative Adversarial Learning across Spherical Circles](#) *Woohyeon Shim, Minsu Cho*
- [WOR and pp's: Sketches for \$\ell_p\ell_p\$ -Sampling Without Replacement](#) *Edith Cohen, Rasmus Pagh, David Woodruff*
- [Hypersolvers: Toward Fast Continuous-Depth Models](#) *Michael Poli, Stefano Massaroli, Atsushi Yamashita, Hajime Asama, Jinkyoo Park*
- [Log-Likelihood Ratio Minimizing Flows: Towards Robust and Quantifiable Neural Distribution Alignment](#) *Ben Usman, Avneesh Sud, Nick Dufour, Kate Saenko*
- [Escaping the Gravitational Pull of Softmax](#) *Jincheng Mei, Chenjun Xiao, Bo Dai, Lihong Li, Csaba Szepesvari, Dale Schuurmans*
- [Regret in Online Recommendation Systems](#) *Kaito Ariu, Narae Ryu, Se-Young Yun, Alexandre Proutiere*
- [On Convergence and Generalization of Dropout Training](#) *Poorya Mianjy, Raman Arora*
- [Second Order Optimality in Decentralized Non-Convex Optimization via Perturbed Gradient Tracking](#) *Isidoros Tziotis, Constantine Caramanis, Aryan Mokhtari*
- [Implicit Regularization in Deep Learning May Not Be Explainable by Norms](#) *Noam Razin, Nadav Cohen*
- [POMO: Policy Optimization with Multiple Optima for Reinforcement Learning](#) *Yeong-Dae Kwon, Jinho Choo, Byoungjip Kim, Iljoo Yoon, Youngjune Gwon, Seungjai Min*
- [Uncertainty-aware Self-training for Few-shot Text Classification](#) *Subhabrata Mukherjee, Ahmed Awadallah*
- [Learning to Learn with Feedback and Local Plasticity](#) *Jack Lindsey, Ashok Litwin-Kumar*
- [Every View Counts: Cross-View Consistency in 3D Object Detection with Hybrid-Cylindrical-Spherical Voxelization](#) *Qi Chen, Lin Sun, Ernest Cheung, Alan L. Yuille*
- [Sharper Generalization Bounds for Pairwise Learning](#) *Yunwen Lei, Antoine Ledent, Marius Kloft*
- [A Measure-Theoretic Approach to Kernel Conditional Mean Embeddings](#) *Junhyung Park, Krikamol Muandet*
- [Quantifying the Empirical Wasserstein Distance to a Set of Measures: Beating the Curse of Dimensionality](#) *Nian Si, Jose Blanchet, Soumyadip Ghosh, Mark Squillante*
- [Bootstrap Your Own Latent - A New Approach to Self-Supervised Learning](#) *Jean-Bastien Grill, Florian Strub, Florent Althé, Corentin Tallec, Pierre Richemond, Elena Buchatskaya, Carl Doersch, Bernardo Avila Pires, Zhaohan Guo, Mohammad Gheshlaghi Azar, Bilal Piot, koray kavukcuoglu, Remi Munos, Michal Valko*
- [Towards Theoretically Understanding Why Sgd Generalizes Better Than Adam in Deep Learning](#) *Pan Zhou, Jiashi Feng, Chao Ma, Caiming Xiong, Steven Chu Hong Hoi, Weinan E*
- [RSKDD-Net: Random Sample-based Keypoint Detector and Descriptor](#) *Fan Lu, Guang Chen, Yinlong Liu, Zhongnan Qu, Alois Knoll*

- [Efficient Clustering for Stretched Mixtures: Landscape and Optimality](#) *Kaizheng Wang, Yuling Yan, Mateo Diaz*
- [A Group-Theoretic Framework for Data Augmentation](#) *Shuxiao Chen, Edgar Dobriban, Jane Lee*
- [The Statistical Cost of Robust Kernel Hyperparameter Turning](#) *Raphael Meyer, Christopher Musco*
- [How does Weight Correlation Affect Generalisation Ability of Deep Neural Networks?](#) *Gaojie Jin, Xinping Yi, Liang Zhang, Lijun Zhang, Sven Schewe, Xiaowei Huang*
- [ContraGAN: Contrastive Learning for Conditional Image Generation](#) *Minguk Kang, Jaesik Park*
- [On the distance between two neural networks and the stability of learning](#) *Jeremy Bernstein, Arash Vahdat, Yisong Yue, Ming-Yu Liu*
- [A Topological Filter for Learning with Label Noise](#) *Pengxiang Wu, Songzhu Zheng, Mayank Goswami, Dimitris Metaxas, Chao Chen*
- [Personalized Federated Learning with Moreau Envelopes](#) *Canh T. Dinh, Nguyen Tran, Josh Nguyen*
- [Avoiding Side Effects in Complex Environments](#) *Alex Turner, Neale Ratzlaff, Prasad Tadepalli*
- [No-regret Learning in Price Competitions under Consumer Reference Effects](#) *Negin Golrezaei, Patrick Jaillet, Jason Cheuk Nam Liang*
- [Geometric Dataset Distances via Optimal Transport](#) *David Alvarez-Melis, Nicolo Fusi*
- [Task-Agnostic Amortized Inference of Gaussian Process Hyperparameters](#) *Sulin Liu, Xingyuan Sun, Peter J. Ramadge, Ryan P. Adams*
- [A novel variational form of the Schatten-pp quasi-norm](#) *Paris Giampouras, Rene Vidal, Athanasios Rontogiannis, Benjamin Haeffele*
- [Energy-based Out-of-distribution Detection](#) *Weitang Liu, Xiaoyun Wang, John Owens, Yixuan Li*
- [On the Loss Landscape of Adversarial Training: Identifying Challenges and How to Overcome Them](#) *Chen Liu, Mathieu Salzmann, Tao Lin, Ryota Tomioka, Sabine Ssstrunk*
- [User-Dependent Neural Sequence Models for Continuous-Time Event Data](#) *Alex Boyd, Robert Bamler, Stephan Mandt, Padhraic Smyth*
- [Active Structure Learning of Causal DAGs via Directed Clique Trees](#) *Chandler Squires, Sara Magliacane, Kristjan Greenewald, Dmitriy Katz, Murat Kocaoglu, Karthikeyan Shanmugam*
- [Convergence and Stability of Graph Convolutional Networks on Large Random Graphs](#) *Nicolas Keriven, Alberto Bietti, Samuel Vaiter*
- [BoTorch: A Framework for Efficient Monte-Carlo Bayesian Optimization](#) *Maximilian Balandat, Brian Karrer, Daniel Jiang, Samuel Daulton, Ben Letham, Andrew G. Wilson, Eytan Bakshy*
- [Reconsidering Generative Objectives For Counterfactual Reasoning](#) *Danni Lu, Chenyang Tao, Junya Chen, Fan Li, Feng Guo, Lawrence Carin*
- [Robust Federated Learning: The Case of Affine Distribution Shifts](#) *Amirhossein Reisizadeh, Farzan Farnia, Ramtin Pedarsani, Ali Jadbabaie*
- [Quantile Propagation for Wasserstein-Approximate Gaussian Processes](#) *Rui Zhang, Christian Walder, Edwin V. Bonilla, Marian-Andrei Rizoiu, Lexing Xie*
- [Generating Adjacency-Constrained Subgoals in Hierarchical Reinforcement Learning](#) *Tianren Zhang, Shangqi Guo, Tian Tan, Xiaolin Hu, Feng Chen*
- [High-contrast “gaudy” images improve the training of deep neural network models of visual cortex](#) *Benjamin Cowley, Jonathan W. Pillow*
- [Duality-Induced Regularizer for Tensor Factorization Based Knowledge Graph Completion](#) *Zhanqiu Zhang, Jianyu Cai, Jie Wang*
- [Distributed Training with Heterogeneous Data: Bridging Median- and Mean-Based Algorithms](#) *Xiangyi Chen,*

Tiancong Chen, Haoran Sun, Steven Z. Wu, Mingyi Hong

- H-Mem: Harnessing synaptic plasticity with Hebbian Memory Networks *Thomas Limbacher, Robert Legenstein*
- Neural Unsigned Distance Fields for Implicit Function Learning *Julian Chibane, Mohamad Aymen mir, Gerard Pons-Moll*
- Curriculum By Smoothing *Samarth Sinha, Animesh Garg, Hugo Larochelle*
- Fast Transformers with Clustered Attention *Apoorv Vyas, Angelos Katharopoulos, François Fleuret*
- The Convex Relaxation Barrier, Revisited: Tightened Single-Neuron Relaxations for Neural Network Verification *Christian Tjandraatmadja, Ross Anderson, Joey Huchette, Will Ma, KRUNAL KISHOR PATEL, Juan Pablo Vielma*
- Strongly Incremental Constituency Parsing with Graph Neural Networks *Kaiyu Yang, Jia Deng*
- AOT: Appearance Optimal Transport Based Identity Swapping for Forgery Detection *Hao Zhu, Chaoyou Fu, Qianyi Wu, Wayne Wu, Chen Qian, Ran He*
- Uncertainty-Aware Learning for Zero-Shot Semantic Segmentation *Ping Hu, Stan Sclaroff, Kate Saenko*
- Delta-STN: Efficient Bilevel Optimization for Neural Networks using Structured Response Jacobians *Juhan Bae, Roger B. Grosse*
- First-Order Methods for Large-Scale Market Equilibrium Computation *Yuan Gao, Christian Kroer*
- Minimax Optimal Nonparametric Estimation of Heterogeneous Treatment Effects *Zijun Gao, Yanjun Han*
- Residual Force Control for Agile Human Behavior Imitation and Extended Motion Synthesis *Ye Yuan, Kris Kitani*
- A General Method for Robust Learning from Batches *Ayush Jain, Alon Orlitsky*
- Not All Unlabeled Data are Equal: Learning to Weight Data in Semi-supervised Learning *Zhongzheng Ren, Raymond Yeh, Alexander Schwing*
- Hard Negative Mixing for Contrastive Learning *Yannis Kalantidis, Mert Bulent Sariyildiz, Noe Pion, Philippe Weinzaepfel, Diane Larlus*
- MOREL: Model-Based Offline Reinforcement Learning *Rahul Kidambi, Aravind Rajeswaran, Praneeth Netrapalli, Thorsten Joachims*
- Weisfeiler and Leman go sparse: Towards scalable higher-order graph embeddings *Christopher Morris, Gaurav Rattan, Petra Mutzel*
- Adversarial Crowdsourcing Through Robust Rank-One Matrix Completion *Qianqian Ma, Alex Olshevsky*
- Learning Semantic-aware Normalization for Generative Adversarial Networks *Heliang Zheng, Jianlong Fu, Yanhong Zeng, Jiebo Luo, Zheng-Jun Zha*
- Differentiable Causal Discovery from Interventional Data *Philippe Brouillard, Sébastien Lachapelle, Alexandre Lacoste, Simon Lacoste-Julien, Alexandre Drouin*
- One-sample Guided Object Representation Disassembling *Zunlei Feng, Yongming He, Xinchao Wang, Xin Gao, Jie Lei, Cheng Jin, Mingli Song*
- Extrapolation Towards Imaginary 0-Nearest Neighbour and Its Improved Convergence Rate *Akifumi Okuno, Hidetoshi Shimodaira*
- Robust Persistence Diagrams using Reproducing Kernels *Siddharth Vishwanath, Kenji Fukumizu, Satoshi Kuriki, Bharath K. Sriperumbudur*
- Contextual Games: Multi-Agent Learning with Side Information *Pier Giuseppe Sessa, Ilija Bogunovic, Andreas Krause, Maryam Kamgarpour*
- Goal-directed Generation of Discrete Structures with Conditional Generative Models *Amina Mollaysa, Brooks Paige, Alexandros Kalousis*

- [Beyond Lazy Training for Over-parameterized Tensor Decomposition](#) *Xiang Wang, Chenwei Wu, Jason D. Lee, Tengyu Ma, Rong Ge*
- [Denoised Smoothing: A Provable Defense for Pretrained Classifiers](#) *Hadi Salman, Mingjie Sun, Greg Yang, Ashish Kapoor, J. Zico Kolter*
- [Minibatch Stochastic Approximate Proximal Point Methods](#) *Hilal Asi, Karan Chadha, Gary Cheng, John C. Duchi*
- [Attribute Prototype Network for Zero-Shot Learning](#) *Wenjia Xu, Yongqin Xian, Jiuniu Wang, Bernt Schiele, Zeynep Akata*
- [CrossTransformers: spatially-aware few-shot transfer](#) *Carl Doersch, Ankush Gupta, Andrew Zisserman*
- [Learning Latent Space Energy-Based Prior Model](#) *Bo Pang, Tian Han, Erik Nijkamp, Song-Chun Zhu, Ying Nian Wu*
- [SEVIR : A Storm Event Imagery Dataset for Deep Learning Applications in Radar and Satellite Meteorology](#) *Mark Veillette, Siddharth Samsi, Chris Mattioli*
- [Lightweight Generative Adversarial Networks for Text-Guided Image Manipulation](#) *Bowen Li, Xiaojuan Qi, Philip Torr, Thomas Lukasiewicz*
- [High-Dimensional Contextual Policy Search with Unknown Context Rewards using Bayesian Optimization](#) *Qing Feng , Ben Letham, Hongzi Mao, Eytan Bakshy*
- [Model Fusion via Optimal Transport](#) *Sidak Pal Singh, Martin Jaggi*
- [On the Stability and Convergence of Robust Adversarial Reinforcement Learning: A Case Study on Linear Quadratic Systems](#) *Kaiqing Zhang, Bin Hu, Tamer Basar*
- [Learning Individually Inferred Communication for Multi-Agent Cooperation](#) *gang Ding, Tiejun Huang, Zongqing Lu*
- [Set2Graph: Learning Graphs From Sets](#) *Hadar Serviansky, Nimrod Segol, Jonathan Shlomi, Kyle Cranmer, Eilam Gross, Haggai Maron, Yaron Lipman*
- [Graph Random Neural Networks for Semi-Supervised Learning on Graphs](#) *Wenzheng Feng, Jie Zhang, Yuxiao Dong, Yu Han, Huanbo Luan, Qian Xu, Qiang Yang, Evgeny Kharlamov, Jie Tang*
- [Gradient Boosted Normalizing Flows](#) *Robert Giaquinto, Arindam Banerjee*
- [Open Graph Benchmark: Datasets for Machine Learning on Graphs](#) *Weihua Hu, Matthias Fey, Marinka Zitnik, Yuxiao Dong, Hongyu Ren, Bowen Liu, Michele Catasta, Jure Leskovec*
- [Towards Understanding Hierarchical Learning: Benefits of Neural Representations](#) *Minshuo Chen, Yu Bai, Jason D. Lee, Tuo Zhao, Huan Wang, Caiming Xiong, Richard Socher*
- [Texture Interpolation for Probing Visual Perception](#) *Jonathan Vacher, Aida Davila, Adam Kohn, Ruben Coen-Cagli*
- [Hierarchical Neural Architecture Search for Deep Stereo Matching](#) *Xuelian Cheng, Yiran Zhong, Mehrtash Harandi, Yuchao Dai, Xiaojun Chang, Hongdong Li, Tom Drummond, Zongyuan Ge*
- [MuSCLE: Multi Sweep Compression of LiDAR using Deep Entropy Models](#) *Sourav Biswas, Jerry Liu, Kelvin Wong, Shenlong Wang, Raquel Urtasun*
- [Implicit Bias in Deep Linear Classification: Initialization Scale vs Training Accuracy](#) *Edward Moroshko, Blake E. Woodworth, Suriya Gunasekar, Jason D. Lee, Nati Srebro, Daniel Soudry*
- [Focus of Attention Improves Information Transfer in Visual Features](#) *Matteo Tiezzi, Stefano Melacci, Alessandro Betti, Marco Maggini, Marco Gori*
- [Auditing Differentially Private Machine Learning: How Private is Private SGD?](#) *Matthew Jagielski, Jonathan Ullman, Alina Oprea*
- [A Dynamical Central Limit Theorem for Shallow Neural Networks](#) *Zhengdao Chen, Grant Rotskoff, Joan*

Bruna, Eric Vanden-Eijnden

- Measuring Systematic Generalization in Neural Proof Generation with Transformers *Nicolas Gontier, Koustuv Sinha, Siva Reddy, Chris Pal*
- Big Self-Supervised Models are Strong Semi-Supervised Learners *Ting Chen, Simon Kornblith, Kevin Swersky, Mohammad Norouzi, Geoffrey E. Hinton*
- Learning from Label Proportions: A Mutual Contamination Framework *Clayton Scott, Jianxin Zhang*
- Fast Matrix Square Roots with Applications to Gaussian Processes and Bayesian Optimization *Geoff Pleiss, Martin Jankowiak, David Eriksson, Anil Damle, Jacob Gardner*
- Self-Adaptively Learning to Demoiré from Focused and Defocused Image Pairs *Lin Liu, Shanxin Yuan, Jianzhuang Liu, Liping Bao, Gregory Slabaugh, Qi Tian*
- Confounding-Robust Policy Evaluation in Infinite-Horizon Reinforcement Learning *Nathan Kallus, Angela Zhou*
- Model Class Reliance for Random Forests *Gavin Smith, Roberto Mansilla, James Goulding*
- Follow the Perturbed Leader: Optimism and Fast Parallel Algorithms for Smooth Minimax Games *Arun Suggala, Praneeth Netrapalli*
- Agnostic QQ-learning with Function Approximation in Deterministic Systems: Near-Optimal Bounds on Approximation Error and Sample Complexity *Simon S. Du, Jason D. Lee, Gaurav Mahajan, Ruosong Wang*
- Learning to Adapt to Evolving Domains *Hong Liu, Mingsheng Long, Jianmin Wang, Yu Wang*
- Synthesizing Tasks for Block-based Programming *Umair Ahmed, Maria Christakis, Aleksandr Efremov, Nigel Fernandez, Ahana Ghosh, Abhik Roychoudhury, Adish Singla*
- Scalable Belief Propagation via Relaxed Scheduling *Vitalii Aksenov, Dan Alistarh, Janne H. Korhonen*
- Firefly Neural Architecture Descent: a General Approach for Growing Neural Networks *Lemeng Wu, Bo Liu, Peter Stone, Qiang Liu*
- Risk-Sensitive Reinforcement Learning: Near-Optimal Risk-Sample Tradeoff in Regret *Yingjie Fei, Zhuoran Yang, Yudong Chen, Zhaoran Wang, Qiaomin Xie*
- Learning to Decode: Reinforcement Learning for Decoding of Sparse Graph-Based Channel Codes *Salman Habib, Allison Beemer, Joerg Kliewer*
- Faster DBSCAN via subsampled similarity queries *Heinrich Jiang, Jennifer Jang, Jakub Lacki*
- De-Anonymizing Text by Fingerprinting Language Generation *Zhen Sun, Roei Schuster, Vitaly Shmatikov*
- Multiparameter Persistence Image for Topological Machine Learning *Mathieu Carrière, Andrew Blumberg*
- PLANS: Neuro-Symbolic Program Learning from Videos *Raphaël Dang-Nhu*
- Matrix Inference and Estimation in Multi-Layer Models *Parthe Pandit, Mojtaba Sahraee Ardakan, Sundeep Rangan, Philip Schniter, Alyson K. Fletcher*
- MeshSDF: Differentiable Iso-Surface Extraction *Edoardo Remelli, Artem Lukoianov, Stephan Richter, Benoit Guillard, Timur Bagautdinov, Pierre Baque, Pascal Fua*
- Variational Interaction Information Maximization for Cross-domain Disentanglement *HyeongJoo Hwang, Geon-Hyeong Kim, Seunghoon Hong, Kee-Eung Kim*
- Provably Efficient Exploration for Reinforcement Learning Using Unsupervised Learning *Fei Feng, Ruosong Wang, Wotao Yin, Simon S. Du, Lin Yang*
- Faithful Embeddings for Knowledge Base Queries *Haitian Sun, Andrew Arnold, Tania Bedrax Weiss, Fernando Pereira, William W. Cohen*
- Wasserstein Distances for Stereo Disparity Estimation *Divyansh Garg, Yan Wang, Bharath Hariharan, Mark Campbell, Kilian Q. Weinberger, Wei-Lun Chao*
- Multi-agent Trajectory Prediction with Fuzzy Query Attention *Nitin Kamra, Hao Zhu, Dweep Kumarbhai*

Trivedi, Ming Zhang, Yan Liu

- Multilabel Classification by Hierarchical Partitioning and Data-dependent Grouping *Shashanka Ubaru, Sanjeeb Dash, Arya Mazumdar, Oktay Gunluk*
- An Analysis of SVD for Deep Rotation Estimation *Jake Levinson, Carlos Esteves, Kefan Chen, Noah Snively, Angjoo Kanazawa, Afshin Rostamizadeh, Ameesh Makadia*
- Can the Brain Do Backpropagation? --- Exact Implementation of Backpropagation in Predictive Coding Networks *Yuhang Song, Thomas Lukasiewicz, Zhenghua Xu, Rafal Bogacz*
- Manifold GPLVMs for discovering non-Euclidean latent structure in neural data *Kristopher Jensen, Ta-Chu Kao, Marco Tripodi, Guillaume Hennequin*
- Distributed Distillation for On-Device Learning *Ilai Bistriz, Ariana Mann, Nicholas Bambos*
- COOT: Cooperative Hierarchical Transformer for Video-Text Representation Learning *Simon Ging, Mohammadreza Zolfaghari, Hamed Pirsiavash, Thomas Brox*
- Passport-aware Normalization for Deep Model Protection *Jie Zhang, Dongdong Chen, Jing Liao, Weiming Zhang, Gang Hua, Nenghai Yu*
- Sampling-Decomposable Generative Adversarial Recommender *Binbin Jin, Defu Lian, Zheng Liu, Qi Liu, Jianhui Ma, Xing Xie, Enhong Chen*
- Limits to Depth Efficiencies of Self-Attention *Yoav Levine, Noam Wies, Or Sharir, Hofit Bata, Amnon Shashua*