NIPS Proceedings (/)β

search

Books (/)

Advances in Neural Information Processing Systems 31 (NIPS 2018) pre-proceedings

The papers below appear in Advances in Neural Information Processing Systems 31 edited by S. Bengio (/author/samy-bengio-350) and H. Wallach (/author/hanna-wallach-8515) and H. Larochelle (/author/hugo-larochelle-3213) and K. Grauman (/author/kristen-grauman-3364) and N. Cesa-Bianchi (/author/nicolo-cesa-bianchi-6344) and R. Garnett (/author/roman-garnett-7864).
They are proceedings from the conference, "Neural Information Processing Systems 2018."

- Efficient Algorithms for Non-convex Isotonic Regression through Submodular Optimization (/paper/7286-efficient-algorithms-for-non-convex-isotonic-regression-through-submodular-optimization) Francis Bach (/author/francis-bach-6335)
- Structure-Aware Convolutional Neural Networks (/paper/7287-structure-aware-convolutional-neural-networks) Jianlong Chang (/author/jianlong-chang-10489), Jie Gu (/author/jie-gu-10490), Lingfeng Wang (/author/lingfeng-wang-10491), GAOFENG MENG (/author/gaofeng-meng-10492), SHIMING XIANG (/author/shiming-xiang-10493), Chunhong Pan (/author/chunhong-pan-10494)
- Kalman Normalization: Normalizing Internal Representations Across Network Layers (/paper/7288-kalman-normalization-normalizing-internal-representations-across-network-layers) Guangrun Wang (/author/guangrun-wang-10495), jiefeng peng (/author/jiefeng-peng-10496), Ping Luo (/author/ping-luo-7572), Xinjiang Wang (/author/xinjiang-wang-10497), Liang Lin (/author/liang-lin-5593)
- HOGWILD!-Gibbs can be PanAccurate (/paper/7289-hogwild-gibbs-can-be-panaccurate) Constantinos Daskalakis (/author/constantinos-daskalakis-8166), Nishanth Dikkala (/author/nishanth-dikkala-9360), Siddhartha Jayanti (/author/siddhartha-jayanti-10498)
- Text-Adaptive Generative Adversarial Networks: Manipulating Images with Natural Language (/paper/7290-text-adaptive-generative-adversarial-networks-manipulating-images-with-natural-language) Seonghyeon Nam (/author/seonghyeon-nam-10499), Yunji Kim (/author/yunji-kim-10500), Seon Joo Kim (/author/seon-joo-kim-10501)
- IntroVAE: Introspective Variational Autoencoders for Photographic Image Synthesis (/paper/7291-introvae-introspective-variational-autoencoders-for-photographic-image-synthesis) Huaibo Huang (/author/huaibo-huang-10502), zhihang li (/author/zhihang-li-

10503), Ran He (/author/ran-he-9747), Zhenan Sun (/author/zhenan-sun-9746), Tieniu Tan (/author/tieniu-tan-6456)

- Doubly Robust Bayesian Inference for Non-Stationary Streaming Data with *beta*-Divergences (/paper/7292-doubly-robust-bayesian-inference-for-non-stationary-streaming-data-with-beta-divergences) Jeremias Knoblauch (/author/jeremias-knoblauch-10504), Jack E. Jewson (/author/jack-e-jewson-10505), Theodoros Damoulas (/author/theodoros-damoulas-10506)
- Adapted Deep Embeddings: A Synthesis of Methods for k-Shot Inductive Transfer Learning (/paper/7293-adapted-deep-embeddings-a-synthesis-of-methods-for-k-shot-inductive-transfer-learning) Tyler Scott (/author/tyler-scott-10507), Karl Ridgeway (/author/karl-ridgeway-10508), Michael C. Mozer (/author/michael-c-mozer-191)
- Generalized Inverse Optimization through Online Learning (/paper/7294-generalized-inverse-optimization-through-online-learning) Chaosheng Dong (/author/chaosheng-dong-10509), Yiran Chen (/author/yiran-chen-9167), Bo Zeng (/author/bo-zeng-10510)
- An Off-policy Policy Gradient Theorem Using Emphatic Weightings (/paper/7295-an-off-policy-policy-gradient-theorem-using-emphatic-weightings) Ehsan Imani (/author/ehsan-imani-10511), Eric Graves (/author/eric-graves-10512), Martha White (/author/martha-white-4533)
- Supervised autoencoders: Improving generalization performance with unsupervised regularizers (/paper/7296-supervised-autoencoders-improving-generalization-performance-with-unsupervised-regularizers) Lei Le (/author/lei-le-10513), Andrew Patterson (/author/andrew-patterson-10514), Martha White (/author/martha-white-4533)
- Visual Object Networks: Image Generation with Disentangled 3D Representations (/paper/7297-visual-object-networks-image-generation-with-disentangled-3d-representations) Jun-Yan Zhu (/author/jun-yan-zhu-9438), Zhoutong Zhang (/author/zhoutong-zhang-9570), Chengkai Zhang (/author/chengkai-zhang-8531), Jiajun Wu (/author/jiajun-wu-8096), Antonio Torralba (/author/antonio-torralba-2458), Josh Tenenbaum (/author/josh-tenenbaum-6308), Bill Freeman (/author/bill-freeman-4274)
- Understanding Weight Normalized Deep Neural Networks with Rectified Linear Units (/paper/7298-understanding-weight-normalized-deep-neural-networks-with-rectified-linear-units) Yixi Xu (/author/yixi-xu-10515), Xiao Wang (/author/xiao-wang-10516)
- Learning Pipelines with Limited Data and Domain Knowledge: A Study in Parsing Physics Problems (/paper/7299-learning-pipelines-with-limited-data-and-domain-knowledge-a-study-in-parsing-physics-problems) Mrinmaya Sachan (/author/mrinmaya-sachan-10517), Kumar Avinava Dubey (/author/kumar-avinava-dubey-7131), Tom M. Mitchell (/author/tom-m-mitchell-859), Dan Roth (/author/dan-roth-6990), Eric P. Xing (/author/eric-p-xing-6314)
- Learning long-range spatial dependencies with horizontal gated recurrent units (/paper/7300-learning-long-range-spatial-dependencies-with-horizontal-gated-recurrent-units) Drew Linsley (/author/drew-linsley-10518), Junkyung Kim (/author/junkyung-kim-10519), Vijay Veerabadran (/author/vijay-veerabadran-10520), Charles Windolf (/author/charles-windolf-10521), Thomas Serre (/author/thomas-serre-2272)

- Joint Sub-bands Learning with Clique Structures for Wavelet Domain Super-Resolution (/paper/7301-joint-sub-bands-learning-with-clique-structures-for-wavelet-domain-super-resolution) Zhisheng Zhong (/author/zhisheng-zhong-10522), Tiancheng Shen (/author/tiancheng-shen-10523), Yibo Yang (/author/yibo-yang-10524), Zhouchen Lin (/author/zhouchen-lin-3724), Chao Zhang (/author/chao-zhang-6817)
- Fast Similarity Search via Optimal Sparse Lifting (/paper/7302-fast-similarity-search-via-optimal-sparse-lifting) Wenye Li (/author/wenye-li-3336), Jingwei Mao (/author/jingwei-mao-10525), Yin Zhang (/author/yin-zhang-10526), Shuguang Cui (/author/shuguang-cui-10527)
- Learning Deep Disentangled Embeddings With the F-Statistic Loss (/paper/7303-learning-deep-disentangled-embeddings-with-the-f-statistic-loss) Karl Ridgeway (/author/karl-ridgeway-10508), Michael C. Mozer (/author/michael-c-mozer-191)
- Geometrically Coupled Monte Carlo Sampling (/paper/7304-geometrically-coupled-monte-carlo-sampling) Mark Rowland (/author/mark-rowland-9398), Krzysztof M. Choromanski (/author/krzysztof-m-choromanski-6269), François Chalus (/author/francois-chalus-10528), Aldo Pacchiano (/author/aldo-pacchiano-10529), Tamas Sarlos (/author/tamas-sarlos-10530), Richard E. Turner (/author/richard-e-turner-7449), Adrian Weller (/author/adrian-weller-7545)
- Cooperative Holistic Scene Understanding: Unifying 3D Object, Layout, and Camera Pose Estimation (/paper/7305-cooperative-holistic-scene-understanding-unifying-3d-object-layout-and-camera-pose-estimation) Siyuan Huang (/author/siyuan-huang-10531), Siyuan Qi (/author/siyuan-qi-10532), Yinxue Xiao (/author/yinxue-xiao-10533), Yixin Zhu (/author/yixin-zhu-10534), Ying Nian Wu (/author/ying-nian-wu-10535), Song-Chun Zhu (/author/song-chun-zhu-6683)
- An Efficient Pruning Algorithm for Robust Isotonic Regression (/paper/7306-an-efficient-pruning-algorithm-for-robust-isotonic-regression) Cong Han Lim (/author/cong-han-lim-7672)
- PAC-learning in the presence of adversaries (/paper/7307-pac-learning-in-the-presence-of-adversaries) Daniel Cullina (/author/daniel-cullina-10536), Arjun Nitin Bhagoji (/author/arjun-nitin-bhagoji-10537), Prateek Mittal (/author/prateek-mittal-10538)
- Sparse DNNs with Improved Adversarial Robustness (/paper/7308-sparse-dnns-with-improved-adversarial-robustness) Yiwen Guo (/author/yiwen-guo-8632), Chao Zhang (/author/chao-zhang-6817), Changshui Zhang (/author/changshui-zhang-10539), Yurong Chen (/author/yurong-chen-8634)
- Snap ML: A Hierarchical Framework for Machine Learning (/paper/7309-snap-ml-a-hierarchical-framework-for-machine-learning) Celestine Dünner (/author/celestine-dunner-10040), Thomas Parnell (/author/thomas-parnell-10041), Dimitrios Sarigiannis (/author/dimitrios-sarigiannis-10540), Nikolas Ioannou (/author/nikolas-ioannou-10541), Andreea Anghel (/author/andreea-anghel-10542), Gummadi Ravi (/author/gummadi-ravi-10543), Madhusudanan Kandasamy (/author/madhusudanan-kandasamy-10544), Haralampos Pozidis (/author/haralampos-pozidis-10545)

- See and Think: Disentangling Semantic Scene Completion (/paper/7310-see-and-think-disentangling-semantic-scene-completion) Shice Liu (/author/shice-liu-10546), YU HU (/author/yu-hu-10547), Yiming Zeng (/author/yiming-zeng-10548), Qiankun Tang (/author/qiankun-tang-10549), Beibei Jin (/author/beibei-jin-10550), Yinhe Han (/author/yinhe-han-10551), Xiaowei Li (/author/xiaowei-li-10552)
- Chain of Reasoning for Visual Question Answering (/paper/7311-chain-of-reasoning-for-visual-question-answering) Chenfei Wu (/author/chenfei-wu-10553), Jinlai Liu (/author/jinlai-liu-10554), Xiaojie Wang (/author/xiaojie-wang-10555), Xuan Dong (/author/xuan-dong-10556)
- Sigsoftmax: Reanalysis of the Softmax Bottleneck (/paper/7312-sigsoftmax-reanalysis-of-the-softmax-bottleneck) Sekitoshi Kanai (/author/sekitoshi-kanai-9432), Yasuhiro Fujiwara (/author/yasuhiro-fujiwara-9433), Yuki Yamanaka (/author/yuki-yamanaka-10557), Shuichi Adachi (/author/shuichi-adachi-10558)
- Deep Non-Blind Deconvolution via Generalized Low-Rank Approximation (/paper/7313-deep-non-blind-deconvolution-via-generalized-low-rank-approximation) Wenqi Ren (/author/wenqi-ren-10559), Jiawei Zhang (/author/jiawei-zhang-10560), Lin Ma (/author/lin-ma-10561), Jinshan Pan (/author/jinshan-pan-10562), Xiaochun Cao (/author/xiaochun-cao-10563), Wangmeng Zuo (/author/wangmeng-zuo-7575), Wei Liu (/author/wei-liu-7251), Ming-Hsuan Yang (/author/ming-hsuan-yang-1946)
- Bayesian Pose Graph Optimization via Bingham Distributions and Tempered Geodesic MCMC (/paper/7314-bayesian-pose-graph-optimization-via-bingham-distributions-and-tempered-geodesic-mcmc) Tolga Birdal (/author/tolga-birdal-10564), Umut Simsekli (/author/umut-simsekli-5225), Mustafa Onur Eken (/author/mustafa-onur-eken-10565), Slobodan Ilic (/author/slobodan-ilic-10566)
- MetaAnchor: Learning to Detect Objects with Customized Anchors (/paper/7315-metaanchor-learning-to-detect-objects-with-customized-anchors) Tong Yang (/author/tong-yang-10567), Xiangyu Zhang (/author/xiangyu-zhang-10568), Zeming Li (/author/zeming-li-10569), Wenqiang Zhang (/author/wenqiang-zhang-10570), Jian Sun (/author/jian-sun-7877)
- Image Inpainting via Generative Multi-column Convolutional Neural Networks (/paper/7316-image-inpainting-via-generative-multi-column-convolutional-neural-networks) Yi Wang (/author/yi-wang-10571), Xin Tao (/author/xin-tao-10572), Xiaojuan Qi (/author/xiaojuan-qi-10573), Xiaoyong Shen (/author/xiaoyong-shen-10574), Jiaya Jia (/author/jiaya-jia-7486)
- On Misinformation Containment in Online Social Networks (/paper/7317-on-misinformation-containment-in-online-social-networks) Amo Tong (/author/amo-tong-10575), Ding-Zhu Du (/author/ding-zhu-du-10576), Weili Wu (/author/weili-wu-10577)
- A^2-Nets: Double Attention Networks (/paper/7318-a2-nets-double-attention-networks) Yunpeng Chen (/author/yunpeng-chen-10071), Yannis Kalantidis (/author/yannis-kalantidis-10578), Jianshu Li (/author/jianshu-li-9370), Shuicheng Yan (/author/shuicheng-yan-4679), Jiashi Feng (/author/jiashi-feng-6686)

- Self-Supervised Generation of Spatial Audio for 360° Video (/paper/7319-self-supervised-generation-of-spatial-audio-for-360-video) Pedro Morgado (/author/pedro-morgado-10579), Nuno Nvasconcelos (/author/nuno-nvasconcelos-9103), Timothy Langlois (/author/timothy-langlois-10580), Oliver Wang (/author/oliver-wang-9441)
- How Many Samples are Needed to Estimate a Convolutional Neural Network? (/paper/7320-how-many-samples-are-needed-to-estimate-a-convolutional-neural-network) Simon S. Du (/author/simon-s-du-8945), Yining Wang (/author/yining-wang-7526), Xiyu Zhai (/author/xiyu-zhai-10581), Sivaraman Balakrishnan (/author/sivaraman-balakrishnan-4913), Ruslan R. Salakhutdinov (/author/ruslan-r-salakhutdinov-2816), Aarti Singh (/author/aarti-singh-4037)
- Algorithmic Regularization in Learning Deep Homogeneous Models: Layers are Automatically Balanced (/paper/7321-algorithmic-regularization-in-learning-deep-homogeneous-models-layers-are-automatically-balanced) Simon S. Du (/author/simon-s-du-8945), Wei Hu (/author/wei-hu-9180), Jason D. Lee (/author/jason-d-lee-4783)
- Optimization for Approximate Submodularity (/paper/7322-optimization-for-approximate-submodularity) Yaron Singer (/author/yaron-singer-8169), Avinatan Hassidim (/author/avinatan-hassidim-10582)
- (Probably) Concave Graph Matching (/paper/7323-probably-concave-graph-matching) Haggai Maron (/author/haggai-maron-10583), Yaron Lipman (/author/yaron-lipman-10584)
- Deep Defense: Training DNNs with Improved Adversarial Robustness (/paper/7324-deep-defense-training-dnns-with-improved-adversarial-robustness) Ziang Yan (/author/ziang-yan-10585), Yiwen Guo (/author/yiwen-guo-8632), Changshui Zhang (/author/changshui-zhang-10539)
- Rest-Katyusha: Exploiting the Solution's Structure via Scheduled Restart Schemes (/paper/7325-rest-katyusha-exploiting-the-solutions-structure-via-scheduled-restart-schemes) Junqi Tang (/author/junqi-tang-10586), Mohammad Golbabaee (/author/mohammad-golbabaee-10587), Francis Bach (/author/francis-bach-6335), Mike E. davies (/author/mike-e-davies-10588)
- Implicit Reparameterization Gradients (/paper/7326-implicit-reparameterization-gradients) Mikhail Figurnov (/author/mikhail-figurnov-9110), Shakir Mohamed (/author/shakir-mohamed-4022), Andriy Mnih (/author/andriy-mnih-2595)
- Training DNNs with Hybrid Block Floating Point (/paper/7327-training-dnns-with-hybrid-block-floating-point) Mario Drumond (/author/mario-drumond-10589), Tao LIN (/author/tao-lin-10590), Martin Jaggi (/author/martin-jaggi-7651), Babak Falsafi (/author/babak-falsafi-10591)
- A Model for Learned Bloom Filters and Optimizing by Sandwiching (/paper/7328-a-model-for-learned-bloom-filters-and-optimizing-by-sandwiching) Michael Mitzenmacher (/author/michael-mitzenmacher-9145)
- Soft-Gated Warping-GAN for Pose-Guided Person Image Synthesis (/paper/7329-soft-gated-warping-gan-for-pose-guided-person-image-synthesis) Haoye Dong (/author/haoye-dong-10592), Xiaodan Liang (/author/xiaodan-liang-9217), Ke Gong (/author/ke-gong-

10593), Hanjiang Lai (/author/hanjiang-lai-10594), Jia Zhu (/author/jia-zhu-10595), Jian Yin (/author/jian-yin-10596)

- Deep Functional Dictionaries: Learning Consistent Semantic Structures on 3D Models from Functions (/paper/7330-deep-functional-dictionaries-learning-consistent-semantic-structures-on-3d-models-from-functions) Minhyuk Sung (/author/minhyuk-sung-10597), Hao Su (/author/hao-su-4630), Ronald Yu (/author/ronald-yu-10598), Leonidas J. Guibas (/author/leonidas-j-guibas-3539)
- Nonlocal Neural Networks, Nonlocal Diffusion and Nonlocal Modeling (/paper/7331-nonlocal-neural-networks-nonlocal-diffusion-and-nonlocal-modeling) Yunzhe Tao (/author/yunzhe-tao-10599), Qi Sun (/author/qi-sun-10600), Qiang Du (/author/qiang-du-10601), Wei Liu (/author/wei-liu-7251)
- Are ResNets Provably Better than Linear Predictors? (/paper/7332-are-resnets-provably-better-than-linear-predictors) Ohad Shamir (/author/ohad-shamir-3602)
- Learning to Decompose and Disentangle Representations for Video Prediction (/paper/7333-learning-to-decompose-and-disentangle-representations-for-video-prediction) Jun-Ting Hsieh (/author/jun-ting-hsieh-10602), Bingbin Liu (/author/bingbin-liu-10603), De-An Huang (/author/de-an-huang-10604), Li F. Fei-Fei (/author/li-f-fei-fei-7198), Juan Carlos Niebles (/author/juan-carlos-niebles-10605)
- Multi-Task Learning as Multi-Objective Optimization (/paper/7334-multi-task-learning-as-multi-objective-optimization) Ozan Sener (/author/ozan-sener-8928), Vladlen Koltun (/author/vladlen-koltun-4518)
- Combinatorial Optimization with Graph Convolutional Networks and Guided Tree Search (/paper/7335-combinatorial-optimization-with-graph-convolutional-networks-and-guided-tree-search) Zhuwen Li (/author/zhuwen-li-10606), Qifeng Chen (/author/qifeng-chen-10607), Vladlen Koltun (/author/vladlen-koltun-4518)
- Self-Erasing Network for Integral Object Attention (/paper/7336-self-erasing-network-for-integral-object-attention) Qibin Hou (/author/qibin-hou-10608), PengTao Jiang (/author/pengtao-jiang-10609), Yunchao Wei (/author/yunchao-wei-10610), Ming-Ming Cheng (/author/ming-ming-cheng-10611)
- LinkNet: Relational Embedding for Scene Graph (/paper/7337-linknet-relational-embedding-for-scene-graph) Sanghyun Woo (/author/sanghyun-woo-10612), Dahun Kim (/author/dahun-kim-10613), Donghyeon Cho (/author/donghyeon-cho-10614), In So Kweon (/author/in-so-kweon-10615)
- How to Start Training: The Effect of Initialization and Architecture (/paper/7338-how-to-start-training-the-effect-of-initialization-and-architecture) Boris Hanin (/author/boris-hanin-10616), David Rolnick (/author/david-rolnick-10617)
- Which Neural Net Architectures Give Rise to Exploding and Vanishing Gradients? (/paper/7339-which-neural-net-architectures-give-rise-to-exploding-and-vanishing-gradients) Boris Hanin (/author/boris-hanin-10616)
- Explanations based on the Missing: Towards Contrastive Explanations with Pertinent Negatives (/paper/7340-explanations-based-on-the-missing-towards-contrastive-

explanations-with-pertinent-negatives) Amit Dhurandhar (/author/amit-dhurandhar-10618), Pin-Yu Chen (/author/pin-yu-chen-10619), Ronny Luss (/author/ronny-luss-3756), Chun-Chen Tu (/author/chun-chen-tu-10620), Paishun Ting (/author/paishun-ting-10621), Karthikeyan Shanmugam (/author/karthikeyan-shanmugam-7412), Payel Das (/author/payel-das-10622)

- HitNet: Hybrid Ternary Recurrent Neural Network (/paper/7341-hitnet-hybrid-ternary-recurrent-neural-network) Peiqi Wang (/author/peiqi-wang-10623), Xinfeng Xie (/author/xinfeng-xie-10624), Lei Deng (/author/lei-deng-10625), Guoqi Li (/author/guoqi-li-10626), Dongsheng Wang (/author/dongsheng-wang-10627), Yuan Xie (/author/yuan-xie-10628)
- A Unified Framework for Extensive-Form Game Abstraction with Bounds (/paper/7342-a-unified-framework-for-extensive-form-game-abstraction-with-bounds) Christian Kroer (/author/christian-kroer-10629), Tuomas Sandholm (/author/tuomas-sandholm-2336)
- Removing the Feature Correlation Effect of Multiplicative Noise (/paper/7343-removing-the-feature-correlation-effect-of-multiplicative-noise) Zijun Zhang (/author/zijun-zhang-10630), Yining Zhang (/author/yining-zhang-10631), Zongpeng Li (/author/zongpeng-li-10632)
- Maximum-Entropy Fine Grained Classification (/paper/7344-maximum-entropy-fine-grained-classification) Abhimanyu Dubey (/author/abhimanyu-dubey-10633), Otkrist Gupta (/author/otkrist-gupta-10634), Ramesh Raskar (/author/ramesh-raskar-10635), Nikhil Naik (/author/nikhil-naik-10636)
- On Learning Markov Chains (/paper/7345-on-learning-markov-chains) Yi HAO (/author/yi-hao-10637), Alon Orlitsky (/author/alon-orlitsky-521), Venkatadheeraj Pichapati (/author/venkatadheeraj-pichapati-10376)
- A Neural Compositional Paradigm for Image Captioning (/paper/7346-a-neural-compositional-paradigm-for-image-captioning) Bo Dai (/author/bo-dai-6604), Sanja Fidler (/author/sanja-fidler-4379), Dahua Lin (/author/dahua-lin-4831)
- Quantifying Learning Guarantees for Convex but Inconsistent Surrogates (/paper/7347-quantifying-learning-guarantees-for-convex-but-inconsistent-surrogates) Kirill Struminsky (/author/kirill-struminsky-10638), Simon Lacoste-Julien (/author/simon-lacoste-julien-7161), Anton Osokin (/author/anton-osokin-5553)
- Dialog-based Interactive Image Retrieval (/paper/7348-dialog-based-interactive-image-retrieval) Xiaoxiao Guo (/author/xiaoxiao-guo-6764), Hui Wu (/author/hui-wu-10639), Yu Cheng (/author/yu-cheng-8899), Steven Rennie (/author/steven-rennie-3286), Gerald Tesauro (/author/gerald-tesauro-53), Rogerio Feris (/author/rogerio-feris-10640)
- SPIDER: Near-Optimal Non-Convex Optimization via Stochastic Path-Integrated Differential Estimator (/paper/7349-spider-near-optimal-non-convex-optimization-via-stochastic-path-integrated-differential-estimator) Cong Fang (/author/cong-fang-10074), Chris Junchi Li (/author/chris-junchi-li-8836), Zhouchen Lin (/author/zhouchen-lin-3724), Tong Zhang (/author/tong-zhang-1901)

- Are GANs Created Equal? A Large-Scale Study (/paper/7350-are-gans-created-equal-a-large-scale-study) Mario Lucic (/author/mario-lucic-7416), Karol Kurach (/author/karol-kurach-7296), Marcin Michalski (/author/marcin-michalski-10641), Sylvain Gelly (/author/sylvain-gelly-10211), Olivier Bousquet (/author/olivier-bousquet-2049)
- Learning Disentangled Joint Continuous and Discrete Representations (/paper/7351-learning-disentangled-joint-continuous-and-discrete-representations) Emilien Dupont (/author/emilien-dupont-10642)
- TADAM: Task dependent adaptive metric for improved few-shot learning (/paper/7352-tadam-task-dependent-adaptive-metric-for-improved-few-shot-learning) Boris Oreshkin (/author/boris-oreshkin-10643), Pau Rodríguez López (/author/pau-rodriguez-lopez-10644), Alexandre Lacoste (/author/alexandre-lacoste-9288)
- Do Less, Get More: Streaming Submodular Maximization with Subsampling (/paper/7353-do-less-get-more-streaming-submodular-maximization-with-subsampling) Moran Feldman (/author/moran-feldman-10004), Amin Karbasi (/author/amin-karbasi-6295), Ehsan Kazemi (/author/ehsan-kazemi-10645)
- Sparse Covariance Modeling in High Dimensions with Gaussian Processes (/paper/7354-sparse-covariance-modeling-in-high-dimensions-with-gaussian-processes) Rui Li (/author/rui-li-10646), Kishan KC (/author/kishan-kc-10647), Feng Cui (/author/feng-cui-10648), Justin Domke (/author/justin-domke-4762), Anne Haake (/author/anne-haake-10649)
- Deep Neural Nets with Interpolating Function as Output Activation (/paper/7355-deep-neural-nets-with-interpolating-function-as-output-activation) Bao Wang (/author/bao-wang-10650), Xiyang Luo (/author/xiyang-luo-10651), Zhen Li (/author/zhen-li-5157), Wei Zhu (/author/wei-zhu-10652), Zuoqiang Shi (/author/zuoqiang-shi-10653), Stanley Osher (/author/stanley-osher-10654)
- FishNet: A Versatile Backbone for Image, Region, and Pixel Level Prediction (/paper/7356-fishnet-a-versatile-backbone-for-image-region-and-pixel-level-prediction) Shuyang Sun (/author/shuyang-sun-10655), Jiangmiao Pang (/author/jiangmiao-pang-10656), Jianping Shi (/author/jianping-shi-10657), Shuai Yi (/author/shuai-yi-10658), Wanli Ouyang (/author/wanli-ouyang-8795)
- Visual Memory for Robust Path Following (/paper/7357-visual-memory-for-robust-path-following) Ashish Kumar (/author/ashish-kumar-10659), Saurabh Gupta (/author/saurabh-gupta-10660), David Fouhey (/author/david-fouhey-10661), Sergey Levine (/author/sergey-levine-4516), Jitendra Malik (/author/jitendra-malik-1449)
- KDGAN: Knowledge Distillation with Generative Adversarial Networks (/paper/7358-kdgan-knowledge-distillation-with-generative-adversarial-networks) Xiaojie Wang (/author/xiaojie-wang-10555), Rui Zhang (/author/rui-zhang-10662), Yu Sun (/author/yu-sun-8604), Jianzhong Qi (/author/jianzhong-qi-10663)
- Long short-term memory and Learning-to-learn in networks of spiking neurons (/paper/7359-long-short-term-memory-and-learning-to-learn-in-networks-of-spiking-neurons) Guillaume Bellec (/author/guillaume-bellec-10664), Darjan Salaj (/author/darjan-

salaj-10665), Anand Subramoney (/author/anand-subramoney-10666), Robert Legenstein (/author/robert-legenstein-8312), Wolfgang Maass (/author/wolfgang-maass-618)

- Greedy Hash: Towards Fast Optimization for Accurate Hash Coding in CNN (/paper/7360-greedy-hash-towards-fast-optimization-for-accurate-hash-coding-in-cnn) Shupeng Su (/author/shupeng-su-10667), Chao Zhang (/author/chao-zhang-6817), Kai Han (/author/kai-han-10668), Yonghong Tian (/author/yonghong-tian-10669)
- Informative Features for Model Comparison (/paper/7361-informative-features-for-model-comparison) Wittawat Jitkrittum (/author/wittawat-jitkrittum-8358), Heishiro Kanagawa (/author/heishiro-kanagawa-9032), Patsorn Sangkloy (/author/patsorn-sangkloy-10670), James Hays (/author/james-hays-10671), Bernhard Schölkopf (/author/bernhard-scholkopf-1472), Arthur Gretton (/author/arthur-gretton-2650)
- PointCNN: Convolution On X-Transformed Points (/paper/7362-pointcnn-convolution-on-x-transformed-points) Yangyan Li (/author/yangyan-li-9027), Rui Bu (/author/rui-bu-10672), Mingchao Sun (/author/mingchao-sun-10673), Wei Wu (/author/wei-wu-4960), Xinhan Di (/author/xinhan-di-10674), Baoquan Chen (/author/baoquan-chen-10675)
- Connectionist Temporal Classification with Maximum Entropy Regularization (/paper/7363-connectionist-temporal-classification-with-maximum-entropy-regularization) Hu Liu (/author/hu-liu-10676), Sheng Jin (/author/sheng-jin-10677), Changshui Zhang (/author/changshui-zhang-10539)
- Large Margin Deep Networks for Classification (/paper/7364-large-margin-deep-networks-for-classification) Gamaleldin Elsayed (/author/gamaleldin-elsayed-10678), Dilip Krishnan (/author/dilip-krishnan-4261), Hossein Mobahi (/author/hossein-mobahi-7958), Kevin Regan (/author/kevin-regan-10679), Samy Bengio (/author/samy-bengio-350)
- Generalizing Graph Matching beyond Quadratic Assignment Model (/paper/7365-generalizing-graph-matching-beyond-quadratic-assignment-model) Tianshu Yu (/author/tianshu-yu-10680), Junchi Yan (/author/junchi-yan-9861), Yilin Wang (/author/yilin-wang-10681), Wei Liu (/author/wei-liu-7251), baoxin Li (/author/baoxin-li-10682)
- Solving Large Sequential Games with the Excessive Gap Technique (/paper/7366-solving-large-sequential-games-with-the-excessive-gap-technique) Christian Kroer (/author/christian-kroer-10629), Gabriele Farina (/author/gabriele-farina-10683), Tuomas Sandholm (/author/tuomas-sandholm-2336)
- Discrimination-aware Channel Pruning for Deep Neural Networks (/paper/7367-discrimination-aware-channel-pruning-for-deep-neural-networks) Zhuangwei Zhuang (/author/zhuangwei-zhuang-10684), Mingkui Tan (/author/mingkui-tan-10685), Bohan Zhuang (/author/bohan-zhuang-10686), Jing Liu (/author/jing-liu-10687), Yong Guo (/author/yong-guo-10688), Qingyao Wu (/author/qingyao-wu-10689), Junzhou Huang (/author/junzhou-huang-5423), Jinhui Zhu (/author/jinhui-zhu-10690)
- On the Dimensionality of Word Embedding (/paper/7368-on-the-dimensionality-of-word-embedding) Zi Yin (/author/zi-yin-10691), Yuanyuan Shen (/author/yuanyuan-shen-10692)
- Reinforced Continual Learning (/paper/7369-reinforced-continual-learning) Ju Xu (/author/ju-xu-10693), Zhanxing Zhu (/author/zhanxing-zhu-8367)

- Uncertainty-Aware Attention for Reliable Interpretation and Prediction (/paper/7370-uncertainty-aware-attention-for-reliable-interpretation-and-prediction) Jay Heo (/author/jay-heo-10694), Hae Beom Lee (/author/hae-beom-lee-10695), Saehoon Kim (/author/saehoon-kim-10696), Juho Lee (/author/juho-lee-8128), Kwang Joon Kim (/author/kwang-joon-kim-10697), Eunho Yang (/author/eunho-yang-5404), Sung Ju Hwang (/author/sung-ju-hwang-7214)
- DropMax: Adaptive Variational Softmax (/paper/7371-dropmax-adaptive-variational-softmax) Hae Beom Lee (/author/hae-beom-lee-10695), Juho Lee (/author/juho-lee-8128), Saehoon Kim (/author/saehoon-kim-10696), Eunho Yang (/author/eunho-yang-5404), Sung Ju Hwang (/author/sung-ju-hwang-7214)
- Posterior Concentration for Sparse Deep Learning (/paper/7372-posterior-concentration-for-sparse-deep-learning) Veronika Rockova (/author/veronika-rockova-9694), nicholas polson (/author/nicholas-polson-10698)
- A flexible model for training action localization with varying levels of supervision (/paper/7373-a-flexible-model-for-training-action-localization-with-varying-levels-of-supervision) Guilhem Chéron (/author/guilhem-cheron-10699), Jean-Baptiste Alayrac (/author/jean-baptiste-alayrac-10700), Ivan Laptev (/author/ivan-laptev-4922), Cordelia Schmid (/author/cordelia-schmid-7293)
- A Deep Bayesian Policy Reuse Approach Against Non-Stationary Agents (/paper/7374-a-deep-bayesian-policy-reuse-approach-against-non-stationary-agents) YAN ZHENG (/author/yan-zheng-10701), Zhaopeng Meng (/author/zhaopeng-meng-10702), Jianye Hao (/author/jianye-hao-10703), Zongzhang Zhang (/author/zongzhang-zhang-10704), Tianpei Yang (/author/tianpei-yang-10705), Changjie Fan (/author/changjie-fan-10706)
- Empirical Risk Minimization in Non-interactive Local Differential Privacy Revisited (/paper/7375-empirical-risk-minimization-in-non-interactive-local-differential-privacy-revisited) Di Wang (/author/di-wang-6556), Marco Gaboardi (/author/marco-gaboardi-10707), Jinhui Xu (/author/jinhui-xu-3682)
- Low-shot Learning via Covariance-Preserving Adversarial Augmentation Networks (/paper/7376-low-shot-learning-via-covariance-preserving-adversarial-augmentation-networks) Hang Gao (/author/hang-gao-10708), Zheng Shou (/author/zheng-shou-10709), Alireza Zareian (/author/alireza-zareian-10710), Hanwang Zhang (/author/hanwang-zhang-10711), Shih-Fu Chang (/author/shih-fu-chang-6715)
- Learning semantic similarity in a continuous space (/paper/7377-learning-semantic-similarity-in-a-continuous-space) Michel Deudon (/author/michel-deudon-10712)
- MetaReg: Towards Domain Generalization using Meta-Regularization (/paper/7378-metareg-towards-domain-generalization-using-meta-regularization) Yogesh Balaji (/author/yogesh-balaji-10713), Swami Sankaranarayanan (/author/swami-sankaranarayanan-10714), Rama Chellappa (/author/rama-chellappa-811)
- Boosted Sparse and Low-Rank Tensor Regression (/paper/7379-boosted-sparse-and-low-rank-tensor-regression) Lifang He (/author/lifang-he-10715), Kun Chen (/author/kun-chen-

10716), Wanwan Xu (/author/wanwan-xu-10717), Jiayu Zhou (/author/jiayu-zhou-5003), Fei Wang (/author/fei-wang-10718)

- Domain-Invariant Projection Learning for Zero-Shot Recognition (/paper/7380-domain-invariant-projection-learning-for-zero-shot-recognition) An Zhao (/author/an-zhao-10719), Mingyu Ding (/author/mingyu-ding-10720), Jiechao Guan (/author/jiechao-guan-10721), Zhiwu Lu (/author/zhiwu-lu-10722), Tao Xiang (/author/tao-xiang-10723), Ji-Rong Wen (/author/ji-rong-wen-10724)
- Neural-Symbolic VQA: Disentangling Reasoning from Vision and Language Understanding (/paper/7381-neural-symbolic-vqa-disentangling-reasoning-from-vision-and-language-understanding) Kexin Yi (/author/kexin-yi-10725), Jiajun Wu (/author/jiajun-wu-8096), Chuang Gan (/author/chuang-gan-10726), Antonio Torralba (/author/antonio-torralba-2458), Pushmeet Kohli (/author/pushmeet-kohli-4400), Josh Tenenbaum (/author/josh-tenenbaum-6308)
- Frequency-Domain Dynamic Pruning for Convolutional Neural Networks (/paper/7382-frequency-domain-dynamic-pruning-for-convolutional-neural-networks) Zhenhua Liu (/author/zhenhua-liu-10727), Jizheng Xu (/author/jizheng-xu-10728), Xiulian Peng (/author/xiulian-peng-10729), Ruiqin Xiong (/author/ruiqin-xiong-10730)
- Quadratic Decomposable Submodular Function Minimization (/paper/7383-quadratic-decomposable-submodular-function-minimization) Pan Li (/author/pan-li-9723), Niao He (/author/niao-he-7138), Olgica Milenkovic (/author/olgica-milenkovic-9724)
- A Block Coordinate Ascent Algorithm for Mean-Variance Optimization (/paper/7384-a-block-coordinate-ascent-algorithm-for-mean-variance-optimization) Tengyang Xie (/author/tengyang-xie-10731), Bo Liu (/author/bo-liu-4493), Yangyang Xu (/author/yangyang-xu-10732), Mohammad Ghavamzadeh (/author/mohammad-ghavamzadeh-3309), Yinlam Chow (/author/yinlam-chow-7145), Daoming Lyu (/author/daoming-lyu-10733), Daesub Yoon (/author/daesub-yoon-10734)
- *\ell_1*-regression with Heavy-tailed Distributions (/paper/7385-ell_1-regression-with-heavy-tailed-distributions) Lijun Zhang (/author/lijun-zhang-6394), Zhi-Hua Zhou (/author/zhi-hua-zhou-6485)
- Neural Nearest Neighbors Networks (/paper/7386-neural-nearest-neighbors-networks) Tobias Plötz (/author/tobias-plotz-10735), Stefan Roth (/author/stefan-roth-3063)
- Efficient nonmyopic batch active search (/paper/7387-efficient-nonmyopic-batch-active-search) Shali Jiang (/author/shali-jiang-10487), Gustavo Malkomes (/author/gustavo-malkomes-8210), Matthew Abbott (/author/matthew-abbott-10488), Benjamin Moseley (/author/benjamin-moseley-8391), Roman Garnett (/author/roman-garnett-7864)
- A Game-Theoretic Approach to Recommendation Systems with Strategic Content Providers (/paper/7388-a-game-theoretic-approach-to-recommendation-systems-with-strategic-content-providers) Omer Ben-Porat (/author/omer-ben-porat-9610), Moshe Tennenholtz (/author/moshe-tennenholtz-2308)
- Interactive Structure Learning with Structural Query-by-Committee (/paper/7389-interactive-structure-learning-with-structural-query-by-committee) Christopher Tosh

(/author/christopher-tosh-10736), Sanjoy Dasgupta (/author/sanjoy-dasgupta-2220)

- Global Geometry of Multichannel Sparse Blind Deconvolution on the Sphere (/paper/7390-global-geometry-of-multichannel-sparse-blind-deconvolution-on-the-sphere) Yanjun Li (/author/yanjun-li-10737), Yoram Bresler (/author/yoram-bresler-10738)
- Video-to-Video Synthesis (/paper/7391-video-to-video-synthesis) Ting-Chun Wang (/author/ting-chun-wang-10739), Ming-Yu Liu (/author/ming-yu-liu-7201), Jun-Yan Zhu (/author/jun-yan-zhu-9438), Nikolai Yakovenko (/author/nikolai-yakovenko-10740), Andrew Tao (/author/andrew-tao-10741), Jan Kautz (/author/jan-kautz-9470), Bryan Catanzaro (/author/bryan-catanzaro-10742)
- How To Make the Gradients Small Stochastically: Even Faster Convex and Nonconvex SGD (/paper/7392-how-to-make-the-gradients-small-stochastically-even-faster-convex-and-nonconvex-sgd) Zeyuan Allen-Zhu (/author/zeyuan-allen-zhu-8933)
- Synthesize Policies for Transfer and Adaptation across Tasks and Environments (/paper/7393-synthesize-policies-for-transfer-and-adaptation-across-tasks-and-environments) Hexiang Hu (/author/hexiang-hu-10743), Liyu Chen (/author/liyu-chen-10744), Boqing Gong (/author/boqing-gong-6799), Fei Sha (/author/fei-sha-2468)
- Adversarial vulnerability for any classifier (/paper/7394-adversarial-vulnerability-for-any-classifier) Alhussein Fawzi (/author/alhussein-fawzi-8874), Hamza Fawzi (/author/hamza-fawzi-10745), Omar Fawzi (/author/omar-fawzi-10746)
- Evolution-Guided Policy Gradient in Reinforcement Learning (/paper/7395-evolution-guided-policy-gradient-in-reinforcement-learning) Shauharda Khadka (/author/shauharda-khadka-10747), Kagan Tumer (/author/kagan-tumer-1498)
- Toddler-Inspired Visual Object Learning (/paper/7396-toddler-inspired-visual-object-learning) Sven Bambach (/author/sven-bambach-10748), David Crandall (/author/david-crandall-8786), Linda Smith (/author/linda-smith-10749), Chen Yu (/author/chen-yu-10750)
- Alternating optimization of decision trees, with application to learning sparse oblique trees (/paper/7397-alternating-optimization-of-decision-trees-with-application-to-learning-sparse-oblique-trees) Miguel A. Carreira-Perpinan (/author/miguel-a-carreira-perpinan-8356), Pooya Tavallali (/author/pooya-tavallali-10751)
- FD-GAN: Pose-guided Feature Distilling GAN for Robust Person Re-identification (/paper/7398-fd-gan-pose-guided-feature-distilling-gan-for-robust-person-re-identification) Yixiao Ge (/author/yixiao-ge-10752), Zhuowan Li (/author/zhuowan-li-10753), Haiyu Zhao (/author/haiyu-zhao-10754), Guojun Yin (/author/guojun-yin-10755), Shuai Yi (/author/shuai-yi-10658), Xiaogang Wang (/author/xiaogang-wang-3673), hongsheng Li (/author/hongsheng-li-8796)
- New Insight into Hybrid Stochastic Gradient Descent: Beyond With-Replacement Sampling and Convexity (/paper/7399-new-insight-into-hybrid-stochastic-gradient-descent-beyond-with-replacement-sampling-and-convexity) Pan Zhou (/author/pan-zhou-10756), Xiaotong Yuan (/author/xiaotong-yuan-8551), Jiashi Feng (/author/jiashi-feng-6686)
- The Lingering of Gradients: How to Reuse Gradients Over Time (/paper/7400-the-lingering-of-gradients-how-to-reuse-gradients-over-time) Zeyuan Allen-Zhu

(/author/zeyuan-allen-zhu-8933), David Simchi-Levi (/author/david-simchi-levi-10757), Xinshang Wang (/author/xinshang-wang-10758)

- Unsupervised Learning of View-invariant Action Representations (/paper/7401-unsupervised-learning-of-view-invariant-action-representations) Junnan Li (/author/junnan-li-10759), Yongkang Wong (/author/yongkang-wong-10760), Qi Zhao (/author/qi-zhao-3936), Mohan Kankanhalli (/author/mohan-kankanhalli-10761)
- Fairness Behind a Veil of Ignorance: A Welfare Analysis for Automated Decision Making (/paper/7402-fairness-behind-a-veil-of-ignorance-a-welfare-analysis-for-automated-decision-making) Hoda Heidari (/author/hoda-heidari-10762), Claudio Ferrari (/author/claudio-ferrari-10763), Krishna Gummadi (/author/krishna-gummadi-9400), Andreas Krause (/author/andreas-krause-3758)
- Global Gated Mixture of Second-order Pooling for Improving Deep Convolutional Neural Networks (/paper/7403-global-gated-mixture-of-second-order-pooling-for-improving-deep-convolutional-neural-networks) Qilong Wang (/author/qilong-wang-10764), Zilin Gao (/author/zilin-gao-10765), Jiangtao Xie (/author/jiangtao-xie-10766), Wangmeng Zuo (/author/wangmeng-zuo-7575), Peihua Li (/author/peihua-li-10767)
- Image-to-image translation for cross-domain disentanglement (/paper/7404-image-to-image-translation-for-cross-domain-disentanglement) Abel Gonzalez-Garcia (/author/abel-gonzalez-garcia-10768), Joost van de Weijer (/author/joost-van-de-weijer-10769), Yoshua Bengio (/author/yoshua-bengio-178)
- Gradient Sparsification for Communication-Efficient Distributed Optimization (/paper/7405-gradient-sparsification-for-communication-efficient-distributed-optimization) Jianqiao Wangni (/author/jianqiao-wangni-10770), Jialei Wang (/author/jialei-wang-9104), Ji Liu (/author/ji-liu-4638), Tong Zhang (/author/tong-zhang-1901)
- Revisiting Multi-Task Learning with ROCK: a Deep Residual Auxiliary Block for Visual Detection (/paper/7406-revisiting-multi-task-learning-with-rock-a-deep-residual-auxiliary-block-for-visual-detection) Taylor Mordan (/author/taylor-mordan-10771), Nicolas THOME (/author/nicolas-thome-10772), Gilles Henaff (/author/gilles-henaff-10773), Matthieu Cord (/author/matthieu-cord-6543)
- Adaptive Online Learning in Dynamic Environments (/paper/7407-adaptive-online-learning-in-dynamic-environments) Lijun Zhang (/author/lijun-zhang-6394), Shiyin Lu (/author/shiyin-lu-10774), Zhi-Hua Zhou (/author/zhi-hua-zhou-6485)
- FRAGE: Frequency-Agnostic Word Representation (/paper/7408-frage-frequency-agnostic-word-representation) Chengyue Gong (/author/chengyue-gong-10473), Di He (/author/di-he-9115), Xu Tan (/author/xu-tan-10775), Tao Qin (/author/tao-qin-3844), Liwei Wang (/author/liwei-wang-4410), Tie-Yan Liu (/author/tie-yan-liu-8972)
- Generative Neural Machine Translation (/paper/7409-generative-neural-machine-translation) Harshil Shah (/author/harshil-shah-10776), David Barber (/author/david-barber-1451)
- Found Graph Data and Planted Vertex Covers (/paper/7410-found-graph-data-and-planted-vertex-covers) Austin R. Benson (/author/austin-r-benson-7276), Jon Kleinberg

(/author/jon-kleinberg-10248)

- Joint Active Feature Acquisition and Classification with Variable-Size Set Encoding (/paper/7411-joint-active-feature-acquisition-and-classification-with-variable-size-set-encoding) Hajin Shim (/author/hajin-shim-10777), Sung Ju Hwang (/author/sung-ju-hwang-7214), Eunho Yang (/author/eunho-yang-5404)
- Regularization Learning Networks: Deep Learning for Tabular Datasets (/paper/7412-regularization-learning-networks-deep-learning-for-tabular-datasets) Ira Shavitt (/author/ira-shavitt-10778), Eran Segal (/author/eran-segal-2233)
- Multitask Boosting for Survival Analysis with Competing Risks (/paper/7413-multitask-boosting-for-survival-analysis-with-competing-risks) Alexis Bellot (/author/alexis-bellot-10779), Mihaela van der Schaar (/author/mihaela-van-der-schaar-9351)
- Geometry Based Data Generation (/paper/7414-geometry-based-data-generation) Ofir Lindenbaum (/author/ofir-lindenbaum-10780), Jay Stanley (/author/jay-stanley-10781), Guy Wolf (/author/guy-wolf-10782), Smita Krishnaswamy (/author/smita-krishnaswamy-10783)
- SLAYER: Spike Layer Error Reassignment in Time (/paper/7415-slayer-spike-layer-error-reassignment-in-time) Sumit Bam Shrestha (/author/sumit-bam-shrestha-10784), Garrick Orchard (/author/garrick-orchard-10785)
- On Oracle-Efficient PAC RL with Rich Observations (/paper/7416-on-oracle-efficient-pac-rl-with-rich-observations) Christoph Dann (/author/christoph-dann-8074), Nan Jiang (/author/nan-jiang-9663), Akshay Krishnamurthy (/author/akshay-krishnamurthy-5059), Alekh Agarwal (/author/alekh-agarwal-3606), John Langford (/author/john-langford-2139), Robert E. Schapire (/author/robert-e-schapire-710)
- Gradient Descent for Spiking Neural Networks (/paper/7417-gradient-descent-for-spiking-neural-networks) Dongsung Huh (/author/dongsung-huh-10786), Terrence J. Sejnowski (/author/terrence-j-sejnowski-54)
- Generalizing Tree Probability Estimation via Bayesian Networks (/paper/7418-generalizing-tree-probability-estimation-via-bayesian-networks) Cheng Zhang (/author/cheng-zhang-10164), Frederick A Matsen IV (/author/frederick-a-matsen-iv-10787)
- Where Do You Think You're Going?: Inferring Beliefs about Dynamics from Behavior (/paper/7419-where-do-you-think-youre-going-inferring-beliefs-about-dynamics-from-behavior) Sid Reddy (/author/sid-reddy-10788), Anca Dragan (/author/anca-dragan-9037), Sergey Levine (/author/sergey-levine-4516)
- Designing by Training: Acceleration Neural Network for Fast High-Dimensional Convolution (/paper/7420-designing-by-training-acceleration-neural-network-for-fast-high-dimensional-convolution) Longquan Dai (/author/longquan-dai-10789), Liang Tang (/author/liang-tang-10790), Yuan Xie (/author/yuan-xie-10628), Jinhui Tang (/author/jinhui-tang-10791)
- Understanding the Role of Adaptivity in Machine Teaching: The Case of Version Space Learners (/paper/7421-understanding-the-role-of-adaptivity-in-machine-teaching-the-case-

of-version-space-learners) Yuxin Chen (/author/yuxin-chen-8045), Adish Singla (/author/adish-singla-10792), Oisin Mac Aodha (/author/oisin-mac-aodha-10793), Pietro Perona (/author/pietro-perona-1183), Yisong Yue (/author/yisong-yue-5193)

- A loss framework for calibrated anomaly detection (/paper/7422-a-loss-framework-for-calibrated-anomaly-detection)
- PacGAN: The power of two samples in generative adversarial networks (/paper/7423-pacgan-the-power-of-two-samples-in-generative-adversarial-networks) Zinan Lin (/author/zinan-lin-10794), Ashish Khetan (/author/ashish-khetan-8572), Giulia Fanti (/author/giulia-fanti-9587), Sewoong Oh (/author/sewoong-oh-4356)
- Variational Memory Encoder-Decoder (/paper/7424-variational-memory-encoder-decoder) Hung Le (/author/hung-le-10795), Truyen Tran (/author/truyen-tran-10796), Thin Nguyen (/author/thin-nguyen-10797), Svetha Venkatesh (/author/svetha-venkatesh-3855)
- Stochastic Composite Mirror Descent: Optimal Bounds with High Probabilities (/paper/7425-stochastic-composite-mirror-descent-optimal-bounds-with-high-probabilities) Yunwen Lei (/author/yunwen-lei-8408), Ke Tang (/author/ke-tang-9661)
- Hybrid Retrieval-Generation Reinforced Agent for Medical Image Report Generation (/paper/7426-hybrid-retrieval-generation-reinforced-agent-for-medical-image-report-generation) Yuan Li (/author/yuan-li-10798), Xiaodan Liang (/author/xiaodan-liang-9217), Zhiting Hu (/author/zhiting-hu-9046), Eric P. Xing (/author/eric-p-xing-6314)
- Overcoming Language Priors in Visual Question Answering with Adversarial Regularization (/paper/7427-overcoming-language-priors-in-visual-question-answering-with-adversarial-regularization) Sainandan Ramakrishnan (/author/sainandan-ramakrishnan-10799), Aishwarya Agrawal (/author/aishwarya-agrawal-10800), Stefan Lee (/author/stefan-lee-8782)
- Hybrid Knowledge Routed Modules for Large-scale Object Detection (/paper/7428-hybrid-knowledge-routed-modules-for-large-scale-object-detection) ChenHan Jiang (/author/chenhan-jiang-10801), Hang Xu (/author/hang-xu-10802), Xiaodan Liang (/author/xiaodan-liang-9217), Liang Lin (/author/liang-lin-5593)
- Bilinear Attention Networks (/paper/7429-bilinear-attention-networks) Jin-Hwa Kim (/author/jin-hwa-kim-9074), Jaehyun Jun (/author/jaehyun-jun-10105), Byoung-Tak Zhang (/author/byoung-tak-zhang-9080)
- Parsimonious Quantile Regression of Financial Asset Tail Dynamics via Sequential Learning (/paper/7430-parsimonious-quantile-regression-of-financial-asset-tail-dynamics-via-sequential-learning) Xing Yan (/author/xing-yan-10803), Weizhong Zhang (/author/weizhong-zhang-10804), Lin Ma (/author/lin-ma-10561), Wei Liu (/author/wei-liu-7251), Qi Wu (/author/qi-wu-10805)
- Multi-Class Learning: From Theory to Algorithm (/paper/7431-multi-class-learning-from-theory-to-algorithm) Jian Li (/author/jian-li-3569), Yong Liu (/author/yong-liu-950), Rong Yin (/author/rong-yin-10806), Hua Zhang (/author/hua-zhang-10807), Lizhong Ding (/author/lizhong-ding-10808), Weiping Wang (/author/weiping-wang-10809)

- Multivariate Time Series Imputation with Generative Adversarial Networks (/paper/7432-multivariate-time-series-imputation-with-generative-adversarial-networks) Yonghong Luo (/author/yonghong-luo-10810), Xiangrui Cai (/author/xiangrui-cai-10811), Ying ZHANG (/author/ying-zhang-10812), Jun Xu (/author/jun-xu-10813), Yuan xiaojie (/author/yuan-xiaojie-10814)
- Learning Versatile Filters for Efficient Convolutional Neural Networks (/paper/7433-learning-versatile-filters-for-efficient-convolutional-neural-networks) Yunhe Wang (/author/yunhe-wang-8989), Chang Xu (/author/chang-xu-8990), Chunjing XU (/author/chunjing-xu-10815), Chao Xu (/author/chao-xu-8992), Dacheng Tao (/author/dacheng-tao-4229)
- Accelerated Stochastic Matrix Inversion: General Theory and Speeding up BFGS Rules for Faster Second-Order Optimization (/paper/7434-accelerated-stochastic-matrix-inversion-general-theory-and-speeding-up-bfgs-rules-for-faster-second-order-optimization) Robert Gower (/author/robert-gower-10816), Filip Hanzely (/author/filip-hanzely-10817), Peter Richtarik (/author/peter-richtarik-8292), Sebastian U. Stich (/author/sebastian-u-stich-10062)
- DifNet: Semantic Segmentation by Diffusion Networks (/paper/7435-difnet-semantic-segmentation-by-diffusion-networks) Peng Jiang (/author/peng-jiang-10818), Fanglin Gu (/author/fanglin-gu-10819), Yunhai Wang (/author/yunhai-wang-10820), Changhe Tu (/author/changhe-tu-10821), Baoquan Chen (/author/baoquan-chen-10675)
- Conditional Adversarial Domain Adaptation (/paper/7436-conditional-adversarial-domain-adaptation) Mingsheng Long (/author/mingsheng-long-8558), ZHANGJIE CAO (/author/zhangjie-cao-10474), Jianmin Wang (/author/jianmin-wang-8560), Michael I. Jordan (/author/michael-i-jordan-6338)
- Neighbourhood Consensus Networks (/paper/7437-neighbourhood-consensus-networks) Ignacio Rocco (/author/ignacio-rocco-10822), Mircea Cimpoi (/author/mircea-cimpoi-10823), Relja Arandjelović (/author/relja-arandjelovic-10824), Akihiko Torii (/author/akihiko-torii-10825), Tomas Pajdla (/author/tomas-pajdla-10826), Josef Sivic (/author/josef-sivic-4273)
- Relating Leverage Scores and Density using Regularized Christoffel Functions (/paper/7438-relating-leverage-scores-and-density-using-regularized-christoffel-functions) Edouard Pauwels (/author/edouard-pauwels-9149), Francis Bach (/author/francis-bach-6335), Jean-Philippe Vert (/author/jean-philippe-vert-7375)
- Non-Local Recurrent Network for Image Restoration (/paper/7439-non-local-recurrent-network-for-image-restoration) Ding Liu (/author/ding-liu-10827), Bihan Wen (/author/bihan-wen-10828), Yuchen Fan (/author/yuchen-fan-10829), Chen Change Loy (/author/chen-change-loy-8944), Thomas S. Huang (/author/thomas-s-huang-2118)
- Bayesian Semi-supervised Learning with Graph Gaussian Processes (/paper/7440-bayesian-semi-supervised-learning-with-graph-gaussian-processes) Yin Cheng Ng (/author/yin-cheng-ng-9220), Nicolò Colombo (/author/nicolo-colombo-9271), Ricardo Silva (/author/ricardo-silva-3670)

- Foreground Clustering for Joint Segmentation and Localization in Videos and Images (/paper/7441-foreground-clustering-for-joint-segmentation-and-localization-in-videos-and-images) Abhishek Sharma (/author/abhishek-sharma-7199)
- Video Prediction via Selective Sampling (/paper/7442-video-prediction-via-selective-sampling) Jingwei Xu (/author/jingwei-xu-10830), Bingbing Ni (/author/bingbing-ni-10831), Xiaokang Yang (/author/xiaokang-yang-10463)
- Distilled Wasserstein Learning for Word Embedding and Topic Modeling (/paper/7443-distilled-wasserstein-learning-for-word-embedding-and-topic-modeling) Hongteng Xu (/author/hongteng-xu-9586), Wenlin Wang (/author/wenlin-wang-10832), Wei Liu (/author/wei-liu-7251), Lawrence Carin (/author/lawrence-carin-3001)
- Learning to Exploit Stability for 3D Scene Parsing (/paper/7444-learning-to-exploit-stability-for-3d-scene-parsing) Yilun Du (/author/yilun-du-10833), Zhijian Liu (/author/zhijian-liu-10834), Hector Basevi (/author/hector-basevi-10835), Ales Leonardis (/author/ales-leonardis-4381), Bill Freeman (/author/bill-freeman-4274), Josh Tenenbaum (/author/josh-tenenbaum-6308), Jiajun Wu (/author/jiajun-wu-8096)
- Neural Guided Constraint Logic Programming for Program Synthesis (/paper/7445-neural-guided-constraint-logic-programming-for-program-synthesis) Lisa Zhang (/author/lisa-zhang-10836), Gregory Rosenblatt (/author/gregory-rosenblatt-10837), Ethan Fetaya (/author/ethan-fetaya-10838), Renjie Liao (/author/renjie-liao-9343), William Byrd (/author/william-byrd-10839), Matthew Might (/author/matthew-might-10840), Raquel Urtasun (/author/raquel-urtasun-4509), Richard Zemel (/author/richard-zemel-6522)
- Genetic-Gated Networks for Deep Reinforcement Learning (/paper/7446-genetic-gated-networks-for-deep-reinforcement-learning) Simyung Chang (/author/simyung-chang-10841), John Yang (/author/john-yang-10842), Jaeseok Choi (/author/jaeseok-choi-10843), Nojun Kwak (/author/nojun-kwak-10844)
- Fighting Boredom in Recommender Systems with Linear Reinforcement Learning (/paper/7447-fighting-boredom-in-recommender-systems-with-linear-reinforcement-learning) Romain WARLOP (/author/romain-warlop-10845), Alessandro Lazaric (/author/alessandro-lazaric-3735), Jérémie Mary (/author/jeremie-mary-10846)
- Enhancing the Accuracy and Fairness of Human Decision Making (/paper/7448-enhancing-the-accuracy-and-fairness-of-human-decision-making) Isabel Valera (/author/isabel-valera-5657), Adish Singla (/author/adish-singla-10792), Manuel Gomez Rodriguez (/author/manuel-gomez-rodriguez-6268)
- Temporal Regularization for Markov Decision Process (/paper/7449-temporal-regularization-for-markov-decision-process) Pierre Thodoroff (/author/pierre-thodoroff-10847), Audrey Durand (/author/audrey-durand-10848), Joelle Pineau (/author/joelle-pineau-2636), Doina Precup (/author/doina-precup-1573)
- The Pessimistic Limits and Possibilities of Margin-based Losses in Semi-supervised Learning (/paper/7450-the-pessimistic-limits-and-possibilities-of-margin-based-losses-in-semi-supervised-learning) Jesse Krijthe (/author/jesse-krijthe-10849), Marco Loog (/author/marco-loog-10850)

- Simple random search of static linear policies is competitive for reinforcement learning (/paper/7451-simple-random-search-of-static-linear-policies-is-competitive-for-reinforcement-learning) Horia Mania (/author/horia-mania-10851), Aurelia Guy (/author/aurelia-guy-10852), Benjamin Recht (/author/benjamin-recht-3537)
- Generating Informative and Diverse Conversational Responses via Adversarial Information Maximization (/paper/7452-generating-informative-and-diverse-conversational-responses-via-adversarial-information-maximization) Yizhe Zhang (/author/yizhe-zhang-8612), Michel Galley (/author/michel-galley-10853), Jianfeng Gao (/author/jianfeng-gao-8345), Zhe Gan (/author/zhe-gan-7916), Xiujun Li (/author/xiujun-li-10854), Chris Brockett (/author/chris-brockett-10855), Bill Dolan (/author/bill-dolan-10856)
- Entropy and mutual information in models of deep neural networks (/paper/7453-entropy-and-mutual-information-in-models-of-deep-neural-networks) Marylou Gabrié (/author/marylou-gabrie-10857), Andre Manoel (/author/andre-manoel-10858), Clément Luneau (/author/clement-luneau-10859), jean barbier (/author/jean-barbier-8963), Nicolas Macris (/author/nicolas-macris-8965), Florent Krzakala (/author/florent-krzakala-6406), Lenka Zdeborová (/author/lenka-zdeborova-6407)
- Collaborative Learning for Deep Neural Networks (/paper/7454-collaborative-learning-for-deep-neural-networks) Guocong Song (/author/guocong-song-10860), Wei Chai (/author/wei-chai-10861)
- High Dimensional Linear Regression using Lattice Basis Reduction (/paper/7455-high-dimensional-linear-regression-using-lattice-basis-reduction) Ilias Zadik (/author/ilias-zadik-10862), David Gamarnik (/author/david-gamarnik-7259)
- Symbolic Graph Reasoning Meets Convolutions (/paper/7456-symbolic-graph-reasoning-meets-convolutions) Xiaodan Liang (/author/xiaodan-liang-9217), Zhiting Hu (/author/zhiting-hu-9046), Hao Zhang (/author/hao-zhang-9976), Liang Lin (/author/liang-lin-5593), Eric P. Xing (/author/eric-p-xing-6314)
- DVAE#: Discrete Variational Autoencoders with Relaxed Boltzmann Priors (/paper/7457-dvae-discrete-variational-autoencoders-with-relaxed-boltzmann-priors) Arash Vahdat (/author/arash-vahdat-5369), Evgeny Andriyash (/author/evgeny-andriyash-10863), William Macready (/author/william-macready-10864)
- Partially-Supervised Image Captioning (/paper/7458-partially-supervised-image-captioning) Peter Anderson (/author/peter-anderson-10865), Stephen Gould (/author/stephen-gould-3941), Mark Johnson (/author/mark-johnson-2520)
- 3D-Aware Scene Manipulation via Inverse Graphics (/paper/7459-3d-aware-scene-manipulation-via-inverse-graphics) Shunyu Yao (/author/shunyu-yao-10866), Tzu Ming Hsu (/author/tzu-ming-hsu-10867), Jun-Yan Zhu (/author/jun-yan-zhu-9438), Jiajun Wu (/author/jiajun-wu-8096), Antonio Torralba (/author/antonio-torralba-2458), Bill Freeman (/author/bill-freeman-4274), Josh Tenenbaum (/author/josh-tenenbaum-6308)
- Random Feature Stein Discrepancies (/paper/7460-random-feature-stein-discrepancies) Jonathan Huggins (/author/jonathan-huggins-9129), Lester Mackey (/author/lester-mackey-8079)

- Distributed Stochastic Optimization via Adaptive SGD (/paper/7461-distributed-stochastic-optimization-via-adaptive-sgd) Ashok Cutkosky (/author/ashok-cutkosky-8949), Róbert Busa-Fekete (/author/robert-busa-fekete-8268)
- Precision and Recall for Time Series (/paper/7462-precision-and-recall-for-time-series) Nesime Tatbul (/author/nesime-tatbul-10868), Tae Jun Lee (/author/tae-jun-lee-10869), Stan Zdonik (/author/stan-zdonik-10870), Mejbah Alam (/author/mejbah-alam-10871), Justin Gottschlich (/author/justin-gottschlich-10872)
- Deep Attentive Tracking via Reciprocative Learning (/paper/7463-deep-attentive-tracking-via-reciprocative-learning) Shi Pu (/author/shi-pu-10873), Yibing Song (/author/yibing-song-10874), Chao Ma (/author/chao-ma-10875), Honggang Zhang (/author/honggang-zhang-10876), Ming-Hsuan Yang (/author/ming-hsuan-yang-1946)
- Virtual Class Enhanced Discriminative Embedding Learning (/paper/7464-virtual-class-enhanced-discriminative-embedding-learning) Binghui Chen (/author/binghui-chen-10877), Weihong Deng (/author/weihong-deng-10878), Haifeng Shen (/author/haifeng-shen-10879)
- Attention in Convolutional LSTM for Gesture Recognition (/paper/7465-attention-in-convolutional-lstm-for-gesture-recognition) Liang Zhang (/author/liang-zhang-3868), Guangming Zhu (/author/guangming-zhu-10880), Lin Mei (/author/lin-mei-10881), Peiyi Shen (/author/peiyi-shen-10882), Syed Afaq Ali Shah (/author/syed-afaq-ali-shah-10883), Mohammed Bennamoun (/author/mohammed-bennamoun-10884)
- Pelee: A Real-Time Object Detection System on Mobile Devices (/paper/7466-pelee-a-real-time-object-detection-system-on-mobile-devices) Jun Wang (/author/jun-wang-5143), Tanner Bohn (/author/tanner-bohn-10885), Charles Ling (/author/charles-ling-10886)
- Universal Growth in Production Economies (/paper/7467-universal-growth-in-production-economies) Simina Branzei (/author/simina-branzei-10887), Ruta Mehta (/author/ruta-mehta-10888), Noam Nisan (/author/noam-nisan-10889)
- Bayesian Model Selection Approach to Boundary Detection with Non-Local Priors (/paper/7468-bayesian-model-selection-approach-to-boundary-detection-with-non-local-priors) Fei Jiang (/author/fei-jiang-10890), Guosheng Yin (/author/guosheng-yin-10891), Francesca Dominici (/author/francesca-dominici-10892)
- Efficient Stochastic Gradient Hard Thresholding (/paper/7469-efficient-stochastic-gradient-hard-thresholding) Pan Zhou (/author/pan-zhou-10756), Xiaotong Yuan (/author/xiaotong-yuan-8551), Jiashi Feng (/author/jiashi-feng-6686)
- SplineNets: Continuous Neural Decision Graphs (/paper/7470-splinenets-continuous-neural-decision-graphs) Cem Keskin (/author/cem-keskin-10893), Shahram Izadi (/author/shahram-izadi-10894)
- Generalized Zero-Shot Learning with Deep Calibration Network (/paper/7471-generalized-zero-shot-learning-with-deep-calibration-network) Shichen Liu (/author/shichen-liu-10895), Mingsheng Long (/author/mingsheng-long-8558), Jianmin Wang (/author/jianmin-wang-8560), Michael I. Jordan (/author/michael-i-jordan-6338)

- Neural Architecture Search with Bayesian Optimisation and Optimal Transport (/paper/7472-neural-architecture-search-with-bayesian-optimisation-and-optimal-transport) Kirthevasan Kandasamy (/author/kirthevasan-kandasamy-8276), Willie Neiswanger (/author/willie-neiswanger-10896), Jeff Schneider (/author/jeff-schneider-6415), Barnabas Poczos (/author/barnabas-poczos-7509), Eric P. Xing (/author/eric-p-xing-6314)
- Embedding Logical Queries on Knowledge Graphs (/paper/7473-embedding-logical-queries-on-knowledge-graphs) Will Hamilton (/author/will-hamilton-9530), Payal Bajaj (/author/payal-bajaj-9402), Marinka Zitnik (/author/marinka-zitnik-10897), Dan Jurafsky (/author/dan-jurafsky-10898), Jure Leskovec (/author/jure-leskovec-4767)
- Learning Optimal Reserve Price against Non-myopic Bidders (/paper/7474-learning-optimal-reserve-price-against-non-myopic-bidders) Jinyan Liu (/author/jinyan-liu-10899), Zhiyi Huang (/author/zhiyi-huang-10900), Xiangning Wang (/author/xiangning-wang-10901)
- Sequential Context Encoding for Duplicate Removal (/paper/7475-sequential-context-encoding-for-duplicate-removal) Lu Qi (/author/lu-qi-10902), Shu Liu (/author/shu-liu-10903), Jianping Shi (/author/jianping-shi-10657), Jiaya Jia (/author/jiaya-jia-7486)
- Discovery of Latent 3D Keypoints via End-to-end Geometric Reasoning (/paper/7476-discovery-of-latent-3d-keypoints-via-end-to-end-geometric-reasoning) Supasorn Suwajanakorn (/author/supasorn-suwajanakorn-10904), Noah Snavely (/author/noah-snavely-10905), Jonathan J. Tompson (/author/jonathan-j-tompson-7614), Mohammad Norouzi (/author/mohammad-norouzi-5636)
- Nonparametric learning from Bayesian models with randomized objective functions (/paper/7477-nonparametric-learning-from-bayesian-models-with-randomized-objective-functions) Simon Lyddon (/author/simon-lyddon-10906), Stephen Walker (/author/stephen-walker-10907), Chris C. Holmes (/author/chris-c-holmes-10908)
- SEGA: Variance Reduction via Gradient Sketching (/paper/7478-sega-variance-reduction-via-gradient-sketching) Filip Hanzely (/author/filip-hanzely-10817), Konstantin Mishchenko (/author/konstantin-mishchenko-10909), Peter Richtarik (/author/peter-richtarik-8292)
- Automatic Program Synthesis of Long Programs with a Learned Garbage Collector (/paper/7479-automatic-program-synthesis-of-long-programs-with-a-learned-garbage-collector) Amit Zohar (/author/amit-zohar-10910), Lior Wolf (/author/lior-wolf-9480)
- One-Shot Unsupervised Cross Domain Translation (/paper/7480-one-shot-unsupervised-cross-domain-translation) Sagie Benaim (/author/sagie-benaim-9479), Lior Wolf (/author/lior-wolf-9480)
- Regularizing by the Variance of the Activations' Sample-Variances (/paper/7481-regularizing-by-the-variance-of-the-activations-sample-variances) Etai Littwin (/author/etai-littwin-10911), Lior Wolf (/author/lior-wolf-9480)
- Overlapping Clustering Models, and One (class) SVM to Bind Them All (/paper/7482-overlapping-clustering-models-and-one-class-svm-to-bind-them-all) Xueyu Mao (/author/xueyu-mao-10912), Purnamrita Sarkar (/author/purnamrita-sarkar-6868), Deepayan Chakrabarti (/author/deepayan-chakrabarti-4089)

- Algorithmic Linearly Constrained Gaussian Processes (/paper/7483-algorithmic-linearly-constrained-gaussian-processes) Markus Lange-Hegermann (/author/markus-lange-hegermann-10913)
- DeepExposure: Learning to Expose Photos with Asynchronously Reinforced Adversarial Learning (/paper/7484-deepexposure-learning-to-expose-photos-with-asynchronously-reinforced-adversarial-learning) Runsheng Yu (/author/runsheng-yu-10914), Wenyu Liu (/author/wenyu-liu-5206), Yasen Zhang (/author/yasen-zhang-10915), Zhi Qu (/author/zhi-qu-10916), Deli Zhao (/author/deli-zhao-3865), Bo Zhang (/author/bo-zhang-3841)
- Norm matters: efficient and accurate normalization schemes in deep networks (/paper/7485-norm-matters-efficient-and-accurate-normalization-schemes-in-deep-networks) Elad Hoffer (/author/elad-hoffer-9639), Ron Banner (/author/ron-banner-10917), Itay Golan (/author/itay-golan-10918), Daniel Soudry (/author/daniel-soudry-6992)
- Dual Principal Component Pursuit: Improved Analysis and Efficient Algorithms (/paper/7486-dual-principal-component-pursuit-improved-analysis-and-efficient-algorithms) Zhihui Zhu (/author/zhihui-zhu-10919), Yifan Wang (/author/yifan-wang-9451), Daniel Robinson (/author/daniel-robinson-10920), Daniel Naiman (/author/daniel-naiman-10921), Rene Vidal (/author/rene-vidal-10922), Manolis Tsakiris (/author/manolis-tsakiris-10923)
- MULAN: A Blind and Off-Grid Method for Multichannel Echo Retrieval (/paper/7487-mulan-a-blind-and-off-grid-method-for-multichannel-echo-retrieval) Helena Peic Tukuljac (/author/helena-peic-tukuljac-10924), Antoine Deleforge (/author/antoine-deleforge-10925), Remi Gribonval (/author/remi-gribonval-6284)
- Mixture Matrix Completion (/paper/7488-mixture-matrix-completion) Daniel Pimentel-Alarcon (/author/daniel-pimentel-alarcon-10926)
- Trajectory Convolution for Action Recognition (/paper/7489-trajectory-convolution-for-action-recognition) Yue Zhao (/author/yue-zhao-10927), Yuanjun Xiong (/author/yuanjun-xiong-7250), Dahua Lin (/author/dahua-lin-4831)
- The Description Length of Deep Learning models (/paper/7490-the-description-length-of-deep-learning-models) Léonard Blier (/author/leonard-blier-10928), Yann Ollivier (/author/yann-ollivier-10929)
- A Smoothed Analysis of the Greedy Algorithm for the Linear Contextual Bandit Problem (/paper/7491-a-smoothed-analysis-of-the-greedy-algorithm-for-the-linear-contextual-bandit-problem) Sampath Kannan (/author/sampath-kannan-10930), Jamie H. Morgenstern (/author/jamie-h-morgenstern-8075), Aaron Roth (/author/aaron-roth-8384), Bo Waggoner (/author/bo-waggoner-8386), Zhiwei Steven Wu (/author/zhiwei-steven-wu-10931)
- Revisiting Decomposable Submodular Function Minimization with Incidence Relations (/paper/7492-revisiting-decomposable-submodular-function-minimization-with-incidence-relations) Pan Li (/author/pan-li-9723), Olgica Milenkovic (/author/olgica-milenkovic-9724)
- A Practical Algorithm for Distributed Clustering and Outlier Detection (/paper/7493-a-practical-algorithm-for-distributed-clustering-and-outlier-detection) Jiecao Chen

(/author/jiecao-chen-9268), Erfan Sadeqi Azer (/author/erfan-sadeqi-azer-10932), Qin Zhang (/author/qin-zhang-9270)

- Learning to Reconstruct Shapes from Unseen Classes (/paper/7494-learning-to-reconstruct-shapes-from-unseen-classes) Xiuming Zhang (/author/xiuming-zhang-10933), Zhoutong Zhang (/author/zhoutong-zhang-9570), Chengkai Zhang (/author/chengkai-zhang-8531), Josh Tenenbaum (/author/josh-tenenbaum-6308), Bill Freeman (/author/bill-freeman-4274), Jiajun Wu (/author/jiajun-wu-8096)
- BourGAN: Generative Networks with Metric Embeddings (/paper/7495-bourgan-generative-networks-with-metric-embeddings) Chang Xiao (/author/chang-xiao-10934), Peilin Zhong (/author/peilin-zhong-10935), Changxi Zheng (/author/changxi-zheng-10936)
- Smoothed analysis of the low-rank approach for smooth semidefinite programs (/paper/7496-smoothed-analysis-of-the-low-rank-approach-for-smooth-semidefinite-programs) Thomas Pumir (/author/thomas-pumir-10937), Samy Jelassi (/author/samy-jelassi-10938), Nicolas Boumal (/author/nicolas-boumal-5149)
- Zero-Shot Transfer with Deictic Object-Oriented Representation in Reinforcement Learning (/paper/7497-zero-shot-transfer-with-deictic-object-oriented-representation-in-reinforcement-learning) Ofir Marom (/author/ofir-marom-10939), Benjamin Rosman (/author/benjamin-rosman-10940)
- Overfitting or perfect fitting? Risk bounds for classification and regression rules that interpolate (/paper/7498-overfitting-or-perfect-fitting-risk-bounds-for-classification-and-regression-rules-that-interpolate) Mikhail Belkin (/author/mikhail-belkin-2150), Daniel J. Hsu (/author/daniel-j-hsu-6464), Partha Mitra (/author/partha-mitra-10941)
- Breaking the Span Assumption Yields Fast Finite-Sum Minimization (/paper/7499-breaking-the-span-assumption-yields-fast-finite-sum-minimization) Robert Hannah (/author/robert-hannah-10313), Yanli Liu (/author/yanli-liu-10942), Daniel O'Connor (/author/daniel-oconnor-10943), Wotao Yin (/author/wotao-yin-10215)
- Structured Local Minima in Sparse Blind Deconvolution (/paper/7500-structured-local-minima-in-sparse-blind-deconvolution) Yuqian Zhang (/author/yuqian-zhang-10305), Han-wen Kuo (/author/han-wen-kuo-10944), John Wright (/author/john-wright-3722)
- GIANT: Globally Improved Approximate Newton Method for Distributed Optimization (/paper/7501-giant-globally-improved-approximate-newton-method-for-distributed-optimization) Shusen Wang (/author/shusen-wang-5570), Farbod Roosta-Khorasani (/author/farbod-roosta-khorasani-8452), Peng Xu (/author/peng-xu-2927), Michael W. Mahoney (/author/michael-w-mahoney-4286)
- Modelling sparsity, heterogeneity, reciprocity and community structure in temporal interaction data (/paper/7502-modelling-sparsity-heterogeneity-reciprocity-and-community-structure-in-temporal-interaction-data) Xenia Miscouridou (/author/xenia-miscouridou-10945), Francois Caron (/author/francois-caron-4438), Yee Whye Teh (/author/yee-whye-teh-1925)
- Non-monotone Submodular Maximization in Exponentially Fewer Iterations (/paper/7503-non-monotone-submodular-maximization-in-exponentially-fewer-iterations) Eric Balkanski

(/author/eric-balkanski-9081), Adam Breuer (/author/adam-breuer-10946), Yaron Singer (/author/yaron-singer-8169)

- MetaGAN: An Adversarial Approach to Few-Shot Learning (/paper/7504-metagan-an-adversarial-approach-to-few-shot-learning) Ruixiang ZHANG (/author/ruixiang-zhang-10947), Tong Che (/author/tong-che-8830), Zoubin Ghahramani (/author/zoubin-ghahramani-1020), Yoshua Bengio (/author/yoshua-bengio-178), Yangqiu Song (/author/yangqiu-song-10948)
- Local Differential Privacy for Evolving Data (/paper/7505-local-differential-privacy-for-evolving-data) Matthew Joseph (/author/matthew-joseph-8923), Aaron Roth (/author/aaron-roth-8384), Jonathan Ullman (/author/jonathan-ullman-8645), Bo Waggoner (/author/bo-waggoner-8386)
- Gaussian Process Conditional Density Estimation (/paper/7506-gaussian-process-conditional-density-estimation) Vincent Dutordoir (/author/vincent-dutordoir-10949), Hugh Salimbeni (/author/hugh-salimbeni-10095), James Hensman (/author/james-hensman-5588), Marc Deisenroth (/author/marc-deisenroth-5510)
- Meta-Gradient Reinforcement Learning (/paper/7507-meta-gradient-reinforcement-learning) Zhongwen Xu (/author/zhongwen-xu-9699), Hado P. van Hasselt (/author/hado-p-van-hasselt-7150), David Silver (/author/david-silver-4280)
- Modular Networks: Learning to Decompose Neural Computation (/paper/7508-modular-networks-learning-to-decompose-neural-computation) Louis Kirsch (/author/louis-kirsch-10950), Julius Kunze (/author/julius-kunze-10951), David Barber (/author/david-barber-1451)
- Learning to Navigate in Cities Without a Map (/paper/7509-learning-to-navigate-in-cities-without-a-map) Piotr Mirowski (/author/piotr-mirowski-10952), Matt Grimes (/author/matt-grimes-10953), Mateusz Malinowski (/author/mateusz-malinowski-7378), Karl Moritz Hermann (/author/karl-moritz-hermann-7898), Keith Anderson (/author/keith-anderson-10954), Denis Teplyashin (/author/denis-teplyashin-10955), Karen Simonyan (/author/karen-simonyan-6376), koray kavukcuoglu (/author/koray-kavukcuoglu-7564), Andrew Zisserman (/author/andrew-zisserman-2581), Raia Hadsell (/author/raia-hadsell-10081)
- Query Complexity of Bayesian Private Learning (/paper/7510-query-complexity-of-bayesian-private-learning) Kuang Xu (/author/kuang-xu-10956)
- A theory on the absence of spurious solutions for nonconvex and nonsmooth optimization (/paper/7511-a-theory-on-the-absence-of-spurious-solutions-for-nonconvex-and-nonsmooth-optimization) Cedric Josz (/author/cedric-josz-10957), Yi Ouyang (/author/yi-ouyang-9581), Richard Zhang (/author/richard-zhang-9439), Javad Lavaei (/author/javad-lavaei-10958), Somayeh Sojoudi (/author/somayeh-sojoudi-10959)
- Recurrent World Models Facilitate Policy Evolution (/paper/7512-recurrent-world-models-facilitate-policy-evolution) David Ha (/author/david-ha-10960), Jürgen Schmidhuber (/author/jurgen-schmidhuber-583)

- Ridge Regression and Provable Deterministic Ridge Leverage Score Sampling (/paper/7513-ridge-regression-and-provable-deterministic-ridge-leverage-score-sampling) Shannon McCurdy (/author/shannon-mccurdy-10961)
- Wasserstein Variational Inference (/paper/7514-wasserstein-variational-inference) Luca Ambrogioni (/author/luca-ambrogioni-9516), Umut Güçlü (/author/umut-guclu-8714), Yağmur Güçlütürk (/author/yagmur-gucluturk-10036), Max Hinne (/author/max-hinne-9517), Marcel A. J. van Gerven (/author/marcel-a-j-van-gerven-9172), Eric Maris (/author/eric-maris-9519)
- How Does Batch Normalization Help Optimization? (/paper/7515-how-does-batch-normalization-help-optimization) Shibani Santurkar (/author/shibani-santurkar-10962), Dimitris Tsipras (/author/dimitris-tsipras-10963), Andrew Ilyas (/author/andrew-ilyas-10964), Aleksander Madry (/author/aleksander-madry-10965)
- Verifiable Reinforcement Learning via Policy Extraction (/paper/7516-verifiable-reinforcement-learning-via-policy-extraction) Osbert Bastani (/author/osbert-bastani-8893), Yewen Pu (/author/yewen-pu-10966), Armando Solar-Lezama (/author/armando-solar-lezama-8106)
- Leveraged volume sampling for linear regression (/paper/7517-leveraged-volume-sampling-for-linear-regression) Michal Derezinski (/author/michal-derezinski-7428), Manfred K. Warmuth (/author/manfred-k-warmuth-1253), Daniel J. Hsu (/author/daniel-j-hsu-6464)
- Model Agnostic Supervised Local Explanations (/paper/7518-model-agnostic-supervised-local-explanations) Gregory Plumb (/author/gregory-plumb-10967), Denali Molitor (/author/denali-molitor-10968), Ameet S. Talwalkar (/author/ameet-s-talwalkar-8939)
- A Linear Speedup Analysis of Distributed Deep Learning with Sparse and Quantized Communication (/paper/7519-a-linear-speedup-analysis-of-distributed-deep-learning-with-sparse-and-quantized-communication) Peng Jiang (/author/peng-jiang-10818), Gagan Agrawal (/author/gagan-agrawal-10969)
- Active Learning for Non-Parametric Regression Using Purely Random Trees (/paper/7520-active-learning-for-non-parametric-regression-using-purely-random-trees) Jack Goetz (/author/jack-goetz-9513), Ambuj Tewari (/author/ambuj-tewari-3304), Paul Zimmerman (/author/paul-zimmerman-10970)
- Tree-to-tree Neural Networks for Program Translation (/paper/7521-tree-to-tree-neural-networks-for-program-translation) Xinyun Chen (/author/xinyun-chen-8805), Chang Liu (/author/chang-liu-8802), Dawn Song (/author/dawn-song-4395)
- Batch-Instance Normalization for Adaptively Style-Invariant Neural Networks (/paper/7522-batch-instance-normalization-for-adaptively-style-invariant-neural-networks) Hyeonseob Nam (/author/hyeonseob-nam-10971), Hyo-Eun Kim (/author/hyo-eun-kim-10972)
- Structural Causal Bandits: Where to Intervene? (/paper/7523-structural-causal-bandits-where-to-intervene) Sanghack Lee (/author/sanghack-lee-6584), Elias Bareinboim (/author/elias-bareinboim-6583)

- Answerer in Questioner's Mind: Information Theoretic Approach to Goal-Oriented Visual Dialog (/paper/7524-answerer-in-questioners-mind-information-theoretic-approach-to-goal-oriented-visual-dialog) Sang-Woo Lee (/author/sang-woo-lee-9075), Yu-Jung Heo (/author/yu-jung-heo-10973), Byoung-Tak Zhang (/author/byoung-tak-zhang-9080)
- A Unified Feature Disentangler for Multi-Domain Image Translation and Manipulation (/paper/7525-a-unified-feature-disentangler-for-multi-domain-image-translation-and-manipulation) Alexander H. Liu (/author/alexander-h-liu-10974), Yen-Cheng Liu (/author/yen-cheng-liu-10975), Yu-Ying Yeh (/author/yu-ying-yeh-10976), Yu-Chiang Frank Wang (/author/yu-chiang-frank-wang-10977)
- Online Learning with an Unknown Fairness Metric (/paper/7526-online-learning-with-an-unknown-fairness-metric) Stephen Gillen (/author/stephen-gillen-10978), Christopher Jung (/author/christopher-jung-10979), Michael Kearns (/author/michael-kearns-2449), Aaron Roth (/author/aaron-roth-8384)
- Isolating Sources of Disentanglement in Variational Autoencoders (/paper/7527-isolating-sources-of-disentanglement-in-variational-autoencoders) Tian Qi Chen (/author/tian-qi-chen-10980), Xuechen Li (/author/xuechen-li-10981), Roger B. Grosse (/author/roger-b-grosse-6301), David K. Duvenaud (/author/david-k-duvenaud-4917)
- Contextual bandits with surrogate losses: Margin bounds and efficient algorithms (/paper/7528-contextual-bandits-with-surrogate-losses-margin-bounds-and-efficient-algorithms) Dylan J. Foster (/author/dylan-j-foster-8174), Akshay Krishnamurthy (/author/akshay-krishnamurthy-5059)
- Representation Learning for Treatment Effect Estimation from Observational Data (/paper/7529-representation-learning-for-treatment-effect-estimation-from-observational-data) Liuyi Yao (/author/liuyi-yao-10982), Sheng Li (/author/sheng-li-9514), Yaliang Li (/author/yaliang-li-10983), Mengdi Huai (/author/mengdi-huai-10984), Jing Gao (/author/jing-gao-4441), Aidong Zhang (/author/aidong-zhang-10985)
- Representation Balancing MDPs for Off-policy Policy Evaluation (/paper/7530-representation-balancing-mdps-for-off-policy-policy-evaluation) Yao Liu (/author/yao-liu-10986), Omer Gottesman (/author/omer-gottesman-10987), Aniruddh Raghu (/author/aniruddh-raghu-10988), Matthieu Komorowski (/author/matthieu-komorowski-10989), Aldo A. Faisal (/author/aldo-a-faisal-10990), Finale Doshi-Velez (/author/finale-doshi-velez-8323), Emma Brunskill (/author/emma-brunskill-6649)
- Out of the Box: Reasoning with Graph Convolution Nets for Factual Visual Question Answering (/paper/7531-out-of-the-box-reasoning-with-graph-convolution-nets-for-factual-visual-question-answering) Medhini Narasimhan (/author/medhini-narasimhan-10991), Svetlana Lazebnik (/author/svetlana-lazebnik-3247), Alexander Schwing (/author/alexander-schwing-9411)
- Causal Discovery from Discrete Data using Hidden Compact Representation (/paper/7532-causal-discovery-from-discrete-data-using-hidden-compact-representation) Ruichu Cai (/author/ruichu-cai-10992), Jie Qiao (/author/jie-qiao-10993), Kun Zhang

(/author/kun-zhang-6824), Zhenjie Zhang (/author/zhenjie-zhang-10994), Zhifeng Hao (/author/zhifeng-hao-10995)

- Natasha 2: Faster Non-Convex Optimization Than SGD (/paper/7533-natasha-2-faster-non-convex-optimization-than-sgd) Zeyuan Allen-Zhu (/author/zeyuan-allen-zhu-8933)
- Minimax Statistical Learning with Wasserstein distances (/paper/7534-minimax-statistical-learning-with-wasserstein-distances) Jaeho Lee (/author/jaeho-lee-10996), Maxim Raginsky (/author/maxim-raginsky-3246)
- Provable Variational Inference for Constrained Log-Submodular Models (/paper/7535-provable-variational-inference-for-constrained-log-submodular-models) Josip Djolonga (/author/josip-djolonga-6707), Stefanie Jegelka (/author/stefanie-jegelka-3773), Andreas Krause (/author/andreas-krause-3758)
- Learning Hierarchical Semantic Image Manipulation through Structured Representations (/paper/7536-learning-hierarchical-semantic-image-manipulation-through-structured-representations) Seunghoon Hong (/author/seunghoon-hong-8189), Xinchen Yan (/author/xinchen-yan-8089), Thomas S. Huang (/author/thomas-s-huang-2118), Honglak Lee (/author/honglak-lee-3292)
- Processing of missing data by neural networks (/paper/7537-processing-of-missing-data-by-neural-networks) Marek Śmieja (/author/marek-smieja-10997), Łukasz Struski (/author/ukasz-struski-10998), Jacek Tabor (/author/jacek-tabor-10999), Bartosz Zieliński (/author/bartosz-zielinski-11000), Przemysław Spurek (/author/przemysaw-spurek-11001)
- Safe Active Learning for Time-Series Modeling with Gaussian Processes (/paper/7538-safe-active-learning-for-time-series-modeling-with-gaussian-processes) Christoph Zimmer (/author/christoph-zimmer-11002), Mona Meister (/author/mona-meister-11003), Duy Nguyen-Tuong (/author/duy-nguyen-tuong-11004)
- Optimal Algorithms for Non-Smooth Distributed Optimization in Networks (/paper/7539-optimal-algorithms-for-non-smooth-distributed-optimization-in-networks) Kevin Scaman (/author/kevin-scaman-7317), Francis Bach (/author/francis-bach-6335), Sebastien Bubeck (/author/sebastien-bubeck-6650), Laurent Massoulié (/author/laurent-massoulie-11005), Yin Tat Lee (/author/yin-tat-lee-11006)
- Computing Higher Order Derivatives of Matrix and Tensor Expressions (/paper/7540-computing-higher-order-derivatives-of-matrix-and-tensor-expressions) Soeren Laue (/author/soeren-laue-5352), Matthias Mitterreiter (/author/matthias-mitterreiter-11007), Joachim Giesen (/author/joachim-giesen-3017)
- Paraphrasing Complex Network: Network Compression via Factor Transfer (/paper/7541-paraphrasing-complex-network-network-compression-via-factor-transfer) Jangho Kim (/author/jangho-kim-11008), Seonguk Park (/author/seonguk-park-11009), Nojun Kwak (/author/nojun-kwak-10844)
- Analytic solution and stationary phase approximation for the Bayesian lasso and elastic net (/paper/7542-analytic-solution-and-stationary-phase-approximation-for-the-bayesian-lasso-and-elastic-net) Tom Michoel (/author/tom-michoel-11010)

- Demystifying excessively volatile human learning: A Bayesian persistent prior and a neural approximation (/paper/7543-demystifying-excessively-volatile-human-learning-a-bayesian-persistent-prior-and-a-neural-approximation) Chaitanya Ryali (/author/chaitanya-ryali-11011), Gautam Reddy (/author/gautam-reddy-11012), Angela J. Yu (/author/angela-j-yu-2266)
- Empirical Risk Minimization Under Fairness Constraints (/paper/7544-empirical-risk-minimization-under-fairness-constraints) Michele Donini (/author/michele-donini-11013), Luca Oneto (/author/luca-oneto-4934), Shai Ben-David (/author/shai-ben-david-2010), John S. Shawe-Taylor (/author/john-s-shawe-taylor-7341), Massimiliano Pontil (/author/massimiliano-pontil-2047)
- Unsupervised Learning of Shape and Pose with Differentiable Point Clouds (/paper/7545-unsupervised-learning-of-shape-and-pose-with-differentiable-point-clouds) Eldar Insafutdinov (/author/eldar-insafutdinov-11014), Alexey Dosovitskiy (/author/alexey-dosovitskiy-7576)
- Continuous-time Value Function Approximation in Reproducing Kernel Hilbert Spaces (/paper/7546-continuous-time-value-function-approximation-in-reproducing-kernel-hilbert-spaces) Motoya Ohnishi (/author/motoya-ohnishi-11015), Masahiro Yukawa (/author/masahiro-yukawa-11016), Mikael Johansson (/author/mikael-johansson-11017), Masashi Sugiyama (/author/masashi-sugiyama-1885)
- Gradient Descent Meets Shift-and-Invert Preconditioning for Eigenvector Computation (/paper/7547-gradient-descent-meets-shift-and-invert-preconditioning-for-eigenvector-computation) Zhiqiang Xu (/author/zhiqiang-xu-11018)
- Factored Bandits (/paper/7548-factored-bandits) Julian Zimmert (/author/julian-zimmert-11019), Yevgeny Seldin (/author/yevgeny-seldin-3401)
- Delta-encoder: an effective sample synthesis method for few-shot object recognition (/paper/7549-delta-encoder-an-effective-sample-synthesis-method-for-few-shot-object-recognition) Eli Schwartz (/author/eli-schwartz-11020), Leonid Karlinsky (/author/leonid-karlinsky-4636), Joseph Shtok (/author/joseph-shtok-11021), Sivan Harary (/author/sivan-harary-11022), Mattias Marder (/author/mattias-marder-11023), Abhishek Kumar (/author/abhishek-kumar-4631), Rogerio Feris (/author/rogerio-feris-10640), Raja Giryes (/author/raja-giryes-11024), Alex Bronstein (/author/alex-bronstein-11025)
- Metric on Nonlinear Dynamical Systems with Perron-Frobenius Operators (/paper/7550-metric-on-nonlinear-dynamical-systems-with-perron-frobenius-operators) Isao Ishikawa (/author/isao-ishikawa-11026), Keisuke Fujii (/author/keisuke-fujii-11027), Masahiro Ikeda (/author/masahiro-ikeda-11028), Yuka Hashimoto (/author/yuka-hashimoto-11029), Yoshinobu Kawahara (/author/yoshinobu-kawahara-3437)
- Learning a High Fidelity Pose Invariant Model for High-resolution Face Frontalization (/paper/7551-learning-a-high-fidelity-pose-invariant-model-for-high-resolution-face-frontalization) Jie Cao (/author/jie-cao-11030), Yibo Hu (/author/yibo-hu-11031), Hongwen Zhang (/author/hongwen-zhang-11032), Ran He (/author/ran-he-9747), Zhenan Sun (/author/zhenan-sun-9746)

- Mirrored Langevin Dynamics (/paper/7552-mirrored-langevin-dynamics) Ya-Ping Hsieh (/author/ya-ping-hsieh-8122), Ali Kavis (/author/ali-kavis-11033), Paul Rolland (/author/paul-rolland-11034), Volkan Cevher (/author/volkan-cevher-3964)
- Moonshine: Distilling with Cheap Convolutions (/paper/7553-moonshine-distilling-with-cheap-convolutions) Elliot J. Crowley (/author/elliot-j-crowley-11035), Gavin Gray (/author/gavin-gray-11036), Amos J. Storkey (/author/amos-j-storkey-2459)
- Stochastic Cubic Regularization for Fast Nonconvex Optimization (/paper/7554-stochastic-cubic-regularization-for-fast-nonconvex-optimization) Nilesh Tripuraneni (/author/nilesh-tripuraneni-8347), Mitchell Stern (/author/mitchell-stern-10020), Chi Jin (/author/chi-jin-5251), Jeffrey Regier (/author/jeffrey-regier-9737), Michael I. Jordan (/author/michael-i-jordan-6338)
- Adaptation to Easy Data in Prediction with Limited Advice (/paper/7555-adaptation-to-easy-data-in-prediction-with-limited-advice) Tobias Thune (/author/tobias-thune-11037), Yevgeny Seldin (/author/yevgeny-seldin-3401)
- Differentially Private Bayesian Inference for Exponential Families (/paper/7556-differentially-private-bayesian-inference-for-exponential-families) Garrett Bernstein (/author/garrett-bernstein-11038), Daniel R. Sheldon (/author/daniel-r-sheldon-3728)
- Playing hard exploration games by watching YouTube (/paper/7557-playing-hard-exploration-games-by-watching-youtube) Yusuf Aytar (/author/yusuf-aytar-8611), Tobias Pfaff (/author/tobias-pfaff-11039), David Budden (/author/david-budden-11040), Thomas Paine (/author/thomas-paine-11041), Ziyu Wang (/author/ziyu-wang-10190), Nando de Freitas (/author/nando-de-freitas-2019)
- Dialog-to-Action: Conversational Question Answering Over a Large-Scale Knowledge Base (/paper/7558-dialog-to-action-conversational-question-answering-over-a-large-scale-knowledge-base) Daya Guo (/author/daya-guo-11042), Duyu Tang (/author/duyu-tang-11043), Nan Duan (/author/nan-duan-11044), Ming Zhou (/author/ming-zhou-11045), Jian Yin (/author/jian-yin-10596)
- Norm-Ranging LSH for Maximum Inner Product Search (/paper/7559-norm-ranging-lsh-for-maximum-inner-product-search) Xiao Yan (/author/xiao-yan-11046), Jinfeng Li (/author/jinfeng-li-11047), Xinyan Dai (/author/xinyan-dai-11048), Hongzhi Chen (/author/hongzhi-chen-11049), James Cheng (/author/james-cheng-7470)
- Optimization over Continuous and Multi-dimensional Decisions with Observational Data (/paper/7560-optimization-over-continuous-and-multi-dimensional-decisions-with-observational-data) Dimitris Bertsimas (/author/dimitris-bertsimas-6821), Christopher McCord (/author/christopher-mccord-11050)
- Fast Estimation of Causal Interactions using Wold Processes (/paper/7561-fast-estimation-of-causal-interactions-using-wold-processes) Flavio Figueiredo (/author/flavio-figueiredo-11051), Guilherme Resende Borges (/author/guilherme-resende-borges-11052), Pedro O.S. Vaz de Melo (/author/pedro-os-vaz-de-melo-11053), Renato Assunção (/author/renato-assuncao-11054)

- When do random forests fail? (/paper/7562-when-do-random-forests-fail) Cheng Tang (/author/cheng-tang-11055), Damien Garreau (/author/damien-garreau-7343), Ulrike von Luxburg (/author/ulrike-von-luxburg-10397)
- Near Optimal Exploration-Exploitation in Non-Communicating Markov Decision Processes (/paper/7563-near-optimal-exploration-exploitation-in-non-communicating-markov-decision-processes) Ronan Fruit (/author/ronan-fruit-9845), Matteo Pirotta (/author/matteo-pirotta-6757), Alessandro Lazaric (/author/alessandro-lazaric-3735)
- Optimistic optimization of a Brownian (/paper/7564-optimistic-optimization-of-a-brownian) Jean-Bastien Grill (/author/jean-bastien-grill-8019), Michal Valko (/author/michal-valko-7335), Remi Munos (/author/remi-munos-6446)
- Practical Methods for Graph Two-Sample Testing (/paper/7565-practical-methods-for-graph-two-sample-testing) Debarghya Ghoshdastidar (/author/debarghya-ghoshdastidar-7254), Ulrike von Luxburg (/author/ulrike-von-luxburg-10397)
- NAIS-Net: Stable Deep Networks from Non-Autonomous Differential Equations (/paper/7566-nais-net-stable-deep-networks-from-non-autonomous-differential-equations) Marco Ciccone (/author/marco-ciccone-11056), Marco Gallieri (/author/marco-gallieri-11057), Jonathan Masci (/author/jonathan-masci-6578), Christian Osendorfer (/author/christian-osendorfer-11058), Faustino Gomez (/author/faustino-gomez-6580)
- On the Global Convergence of Gradient Descent for Over-parameterized Models using Optimal Transport (/paper/7567-on-the-global-convergence-of-gradient-descent-for-over-parameterized-models-using-optimal-transport) Lénaïc Chizat (/author/lenaic-chizat-11059), Francis Bach (/author/francis-bach-6335)
- Constructing Deep Neural Networks by Bayesian Network Structure Learning (/paper/7568-constructing-deep-neural-networks-by-bayesian-network-structure-learning) Raanan Y. Rohekar (/author/raanan-y-rohekar-11060), Shami Nisimov (/author/shami-nisimov-11061), Yaniv Gurwicz (/author/yaniv-gurwicz-11062), Guy Koren (/author/guy-koren-11063), Gal Novik (/author/gal-novik-11064)
- Weakly Supervised Dense Event Captioning in Videos (/paper/7569-weakly-supervised-dense-event-captioning-in-videos) Xuguang Duan (/author/xuguang-duan-11065), Wenbing Huang (/author/wenbing-huang-9901), Chuang Gan (/author/chuang-gan-10726), Jingdong Wang (/author/jingdong-wang-11066), Wenwu Zhu (/author/wenwu-zhu-11067), Junzhou Huang (/author/junzhou-huang-5423)
- Faithful Inversion of Generative Models for Effective Amortized Inference (/paper/7570-faithful-inversion-of-generative-models-for-effective-amortized-inference) Stefan Webb (/author/stefan-webb-11068), Adam Golinski (/author/adam-golinski-11069), Rob Zinkov (/author/rob-zinkov-11070), Siddharth Narayanaswamy (/author/siddharth-narayanaswamy-10287), Tom Rainforth (/author/tom-rainforth-9038), Yee Whye Teh (/author/yee-whye-teh-1925), Frank Wood (/author/frank-wood-3062)
- From Stochastic Planning to Marginal MAP (/paper/7571-from-stochastic-planning-to-marginal-map) Hao Cui (/author/hao-cui-11071), Radu Marinescu (/author/radu-marinescu-8371), Roni Khardon (/author/roni-khardon-2261)

- On Binary Classification in Extreme Regions (/paper/7572-on-binary-classification-in-extreme-regions) Hamid JALALZAI (/author/hamid-jalalzai-11072), Stephan Clémençon (/author/stephan-clemencon-9313), Anne Sabourin (/author/anne-sabourin-11073)
- Near-Optimal Policies for Dynamic Multinomial Logit Assortment Selection Models (/paper/7573-near-optimal-policies-for-dynamic-multinomial-logit-assortment-selection-models) Yining Wang (/author/yining-wang-7526), Xi Chen (/author/xi-chen-4483), Yuan Zhou (/author/yuan-zhou-11074)
- Q-learning with Nearest Neighbors (/paper/7574-q-learning-with-nearest-neighbors) Devavrat Shah (/author/devavrat-shah-3543), Qiaomin Xie (/author/qiaomin-xie-11075)
- Global Convergence of Langevin Dynamics Based Algorithms for Nonconvex Optimization (/paper/7575-global-convergence-of-langevin-dynamics-based-algorithms-for-nonconvex-optimization) Pan Xu (/author/pan-xu-9210), Jinghui Chen (/author/jinghui-chen-11076), Difan Zou (/author/difan-zou-11077), Quanquan Gu (/author/quanquan-gu-5348)
- Asymptotic optimality of adaptive importance sampling (/paper/7576-asymptotic-optimality-of-adaptive-importance-sampling) François Portier (/author/francois-portier-11078), Bernard Delyon (/author/bernard-delyon-11079)
- Learning latent variable structured prediction models with Gaussian perturbations (/paper/7577-learning-latent-variable-structured-prediction-models-with-gaussian-perturbations) Kevin Bello (/author/kevin-bello-11080), Jean Honorio (/author/jean-honorio-4384)
- The Nearest Neighbor Information Estimator is Adaptively Near Minimax Rate-Optimal (/paper/7578-the-nearest-neighbor-information-estimator-is-adaptively-near-minimax-rate-optimal) Jiantao Jiao (/author/jiantao-jiao-11081), Weihao Gao (/author/weihao-gao-8827), Yanjun Han (/author/yanjun-han-4536)
- Deep Reinforcement Learning of Marked Temporal Point Processes (/paper/7579-deep-reinforcement-learning-of-marked-temporal-point-processes) Utkarsh Upadhyay (/author/utkarsh-upadhyay-11082), Abir De (/author/abir-de-8678), Manuel Gomez Rodriguez (/author/manuel-gomez-rodriguez-6268)
- Evidential Deep Learning to Quantify Classification Uncertainty (/paper/7580-evidential-deep-learning-to-quantify-classification-uncertainty) Murat Sensoy (/author/murat-sensoy-11083), Lance Kaplan (/author/lance-kaplan-11084), Melih Kandemir (/author/melih-kandemir-11085)
- Parsimonious Bayesian deep networks (/paper/7581-parsimonious-bayesian-deep-networks) Mingyuan Zhou (/author/mingyuan-zhou-4435)
- Single-Agent Policy Tree Search With Guarantees (/paper/7582-single-agent-policy-tree-search-with-guarantees) Laurent Orseau (/author/laurent-orseau-11086), Levi Lelis (/author/levi-lelis-11087), Tor Lattimore (/author/tor-lattimore-7451), Theophane Weber (/author/theophane-weber-8262)
- Semi-crowdsourced Clustering with Deep Generative Models (/paper/7583-semi-crowdsourced-clustering-with-deep-generative-models) Yucen Luo (/author/yucen-luo-

8756), TIAN TIAN (/author/tian-tian-8363), Jiaxin Shi (/author/jiaxin-shi-11088), Jun Zhu (/author/jun-zhu-3840), Bo Zhang (/author/bo-zhang-3841)

- The committee machine: Computational to statistical gaps in learning a two-layers neural network (/paper/7584-the-committee-machine-computational-to-statistical-gaps-in-learning-a-two-layers-neural-network) Benjamin Aubin (/author/benjamin-aubin-11089), Antoine Maillard (/author/antoine-maillard-11090), jean barbier (/author/jean-barbier-8963), Florent Krzakala (/author/florent-krzakala-6406), Nicolas Macris (/author/nicolas-macris-8965), Lenka Zdeborová (/author/lenka-zdeborova-6407)
- Realistic Evaluation of Deep Semi-Supervised Learning Algorithms (/paper/7585-realistic-evaluation-of-deep-semi-supervised-learning-algorithms) Avital Oliver (/author/avital-oliver-11091), Augustus Odena (/author/augustus-odena-11092), Colin A. Raffel (/author/colin-a-raffel-11093), Ekin Dogus Cubuk (/author/ekin-dogus-cubuk-11094), Ian Goodfellow (/author/ian-goodfellow-4369)
- Contextual Combinatorial Multi-armed Bandits with Volatile Arms and Submodular Reward (/paper/7586-contextual-combinatorial-multi-armed-bandits-with-volatile-arms-and-submodular-reward) Lixing Chen (/author/lixing-chen-11095), Jie Xu (/author/jie-xu-11096), Zhuo Lu (/author/zhuo-lu-11097)
- Training deep learning based denoisers without ground truth data (/paper/7587-training-deep-learning-based-denoisers-without-ground-truth-data) Shakarim Soltanayev (/author/shakarim-soltanayev-11098), Se Young Chun (/author/se-young-chun-11099)
- Re-evaluating evaluation (/paper/7588-re-evaluating-evaluation) David Balduzzi (/author/david-balduzzi-5288), Karl Tuyls (/author/karl-tuyls-10027), Julien Perolat (/author/julien-perolat-10028), Thore Graepel (/author/thore-graepel-1666)
- Deep, complex, invertible networks for inversion of transmission effects in multimode optical fibres (/paper/7589-deep-complex-invertible-networks-for-inversion-of-transmission-effects-in-multimode-optical-fibres) Oisín Moran (/author/oisin-moran-11100), Piergiorgio Caramazza (/author/piergiorgio-caramazza-11101), Daniele Faccio (/author/daniele-faccio-11102), Roderick Murray-Smith (/author/roderick-murray-smith-1768)
- Multivariate Convolutional Sparse Coding for Electromagnetic Brain Signals (/paper/7590-multivariate-convolutional-sparse-coding-for-electromagnetic-brain-signals) Tom Dupré la Tour (/author/tom-dupre-la-tour-9545), Thomas Moreau (/author/thomas-moreau-11103), Mainak Jas (/author/mainak-jas-9544), Alexandre Gramfort (/author/alexandre-gramfort-4730)
- Data-Efficient Hierarchical Reinforcement Learning (/paper/7591-data-efficient-hierarchical-reinforcement-learning) Ofir Nachum (/author/ofir-nachum-9795), Shixiang (Shane) Gu (/author/shixiang-shane-gu-8332), Honglak Lee (/author/honglak-lee-3292), Sergey Levine (/author/sergey-levine-4516)
- Speaker-Follower Models for Vision-and-Language Navigation (/paper/7592-speaker-follower-models-for-vision-and-language-navigation) Daniel Fried (/author/daniel-fried-11104), Ronghang Hu (/author/ronghang-hu-7396), Volkan Cirik (/author/volkan-cirik-

11105), Anna Rohrbach (/author/anna-rohrbach-11106), Jacob Andreas (/author/jacob-andreas-7421), Louis-Philippe Morency (/author/louis-philippe-morency-11107), Taylor Berg-Kirkpatrick (/author/taylor-berg-kirkpatrick-7420), Kate Saenko (/author/kate-saenko-3825), Dan Klein (/author/dan-klein-2132), Trevor Darrell (/author/trevor-darrell-678)

- Inequity aversion improves cooperation in intertemporal social dilemmas (/paper/7593-inequity-aversion-improves-cooperation-in-intertemporal-social-dilemmas) Edward Hughes (/author/edward-hughes-11108), Joel Z. Leibo (/author/joel-z-leibo-5026), Matthew Phillips (/author/matthew-phillips-11109), Karl Tuyls (/author/karl-tuyls-10027), Edgar Dueñez-Guzman (/author/edgar-duenez-guzman-11110), Antonio García Castañeda (/author/antonio-garcia-castaneda-11111), Iain Dunning (/author/iain-dunning-11112), Tina Zhu (/author/tina-zhu-11113), Kevin McKee (/author/kevin-mckee-11114), Raphael Koster (/author/raphael-koster-11115), Heather Roff (/author/heather-roff-11116), Thore Graepel (/author/thore-graepel-1666)
- Learning Gaussian Processes by Minimizing PAC-Bayesian Generalization Bounds (/paper/7594-learning-gaussian-processes-by-minimizing-pac-bayesian-generalization-bounds) David Reeb (/author/david-reeb-11117), Andreas Doerr (/author/andreas-doerr-11118), Sebastian Gerwinn (/author/sebastian-gerwinn-3704), Barbara Rakitsch (/author/barbara-rakitsch-6623)
- Probabilistic Matrix Factorization for Automated Machine Learning (/paper/7595-probabilistic-matrix-factorization-for-automated-machine-learning) Nicolo Fusi (/author/nicolo-fusi-11119), Rishit Sheth (/author/rishit-sheth-10175), Melih Elibol (/author/melih-elibol-11120)
- Stochastic Spectral and Conjugate Descent Methods (/paper/7596-stochastic-spectral-and-conjugate-descent-methods) Dmitry Kovalev (/author/dmitry-kovalev-11121), Peter Richtarik (/author/peter-richtarik-8292), Eduard Gorbunov (/author/eduard-gorbunov-11122), Elnur Gasanov (/author/elnur-gasanov-11123)
- Recurrent Relational Networks (/paper/7597-recurrent-relational-networks) Rasmus Palm (/author/rasmus-palm-11124), Ulrich Paquet (/author/ulrich-paquet-4136), Ole Winther (/author/ole-winther-1474)
- But How Does It Work in Theory? Linear SVM with Random Features (/paper/7598-but-how-does-it-work-in-theory-linear-svm-with-random-features) Yitong Sun (/author/yitong-sun-11125), Anna Gilbert (/author/anna-gilbert-11126), Ambuj Tewari (/author/ambuj-tewari-3304)
- Learning to Optimize Tensor Programs (/paper/7599-learning-to-optimize-tensor-programs) Tianqi Chen (/author/tianqi-chen-8249), Lianmin Zheng (/author/lianmin-zheng-11127), Eddie Yan (/author/eddie-yan-11128), Ziheng Jiang (/author/ziheng-jiang-11129), Thierry Moreau (/author/thierry-moreau-11130), Luis Ceze (/author/luis-ceze-9877), Carlos Guestrin (/author/carlos-guestrin-2592), Arvind Krishnamurthy (/author/arvind-krishnamurthy-11131)
- Boosting Black Box Variational Inference (/paper/7600-boosting-black-box-variational-inference) Francesco Locatello (/author/francesco-locatello-9485), Gideon Dresdner

(/author/gideon-dresdner-11132), Rajiv Khanna (/author/rajiv-khanna-7681), Isabel Valera (/author/isabel-valera-5657), Gunnar Raetsch (/author/gunnar-raetsch-9487)

- Nearly tight sample complexity bounds for learning mixtures of Gaussians via sample compression schemes (/paper/7601-nearly-tight-sample-complexity-bounds-for-learning-mixtures-of-gaussians-via-sample-compression-schemes) Hassan Ashtiani (/author/hassan-ashtiani-9087), Shai Ben-David (/author/shai-ben-david-2010), Nicholas Harvey (/author/nicholas-harvey-11133), Christopher Liaw (/author/christopher-liaw-11134), Abbas Mehrabian (/author/abbas-mehrabian-11135), Yaniv Plan (/author/yaniv-plan-8586)
- Actor-Critic Policy Optimization in Partially Observable Multiagent Environments (/paper/7602-actor-critic-policy-optimization-in-partially-observable-multiagent-environments) Sriram Srinivasan (/author/sriram-srinivasan-8974), Marc Lanctot (/author/marc-lanctot-4270), Vinicius Zambaldi (/author/vinicius-zambaldi-9928), Julien Perolat (/author/julien-perolat-10028), Karl Tuyls (/author/karl-tuyls-10027), Remi Munos (/author/remi-munos-6446), Michael Bowling (/author/michael-bowling-2940)
- Step Size Matters in Deep Learning (/paper/7603-step-size-matters-in-deep-learning) Kamil Nar (/author/kamil-nar-11136), Shankar Sastry (/author/shankar-sastry-2659)
- Derivative Estimation in Random Design (/paper/7604-derivative-estimation-in-random-design) Yu Liu (/author/yu-liu-9182), Kris De Brabanter (/author/kris-de-brabanter-11137)
- Zeroth-order (Non)-Convex Stochastic Optimization via Conditional Gradient and Gradient Updates (/paper/7605-zeroth-order-non-convex-stochastic-optimization-via-conditional-gradient-and-gradient-updates) Krishnakumar Balasubramanian (/author/krishnakumar-balasubramanian-10431), Saeed Ghadimi (/author/saeed-ghadimi-11138)
- Latent Gaussian Activity Propagation: Using Smoothness and Structure to Separate and Localize Sounds in Large Noisy Environments (/paper/7606-latent-gaussian-activity-propagation-using-smoothness-and-structure-to-separate-and-localize-sounds-in-large-noisy-environments) Daniel Johnson (/author/daniel-johnson-11139), Daniel Gorelik (/author/daniel-gorelik-11140), Ross E. Mawhorter (/author/ross-e-mawhorter-11141), Kyle Suver (/author/kyle-suver-11142), Weiqing Gu (/author/weiqing-gu-11143), Steven Xing (/author/steven-xing-11144), Cody Gabriel (/author/cody-gabriel-11145), Peter Sankhagowit (/author/peter-sankhagowit-11146)
- Hybrid-MST: A Hybrid Active Sampling Strategy for Pairwise Preference Aggregation (/paper/7607-hybrid-mst-a-hybrid-active-sampling-strategy-for-pairwise-preference-aggregation) JING LI (/author/jing-li-11147), Rafal Mantiuk (/author/rafal-mantiuk-9463), Junle Wang (/author/junle-wang-11148), Suiyi Ling (/author/suiyi-ling-11149), Patrick Le Callet (/author/patrick-le-callet-11150)
- Infinite-Horizon Gaussian Processes (/paper/7608-infinite-horizon-gaussian-processes) Arno Solin (/author/arno-solin-11151), James Hensman (/author/james-hensman-5588), Richard E. Turner (/author/richard-e-turner-7449)
- Dimensionality Reduction for Stationary Time Series via Stochastic Nonconvex Optimization (/paper/7609-dimensionality-reduction-for-stationary-time-series-via-

stochastic-nonconvex-optimization) Minshuo Chen (/author/minshuo-chen-11152), Lin Yang (/author/lin-yang-9084), Mengdi Wang (/author/mengdi-wang-9065), Tuo Zhao (/author/tuo-zhao-5638)

- Sequence-to-Segment Networks for Segment Detection (/paper/7610-sequence-to-segment-networks-for-segment-detection) Zijun Wei (/author/zijun-wei-9089), Boyu Wang (/author/boyu-wang-11153), Minh Hoai Nguyen (/author/minh-hoai-nguyen-9091), Jianming Zhang (/author/jianming-zhang-11154), Zhe Lin (/author/zhe-lin-10439), Xiaohui Shen (/author/xiaohui-shen-9160), Radomir Mech (/author/radomir-mech-11155), Dimitris Samaras (/author/dimitris-samaras-3108)
- Scaling the Poisson GLM to massive neural datasets through polynomial approximations (/paper/7611-scaling-the-poisson-glm-to-massive-neural-datasets-through-polynomial-approximations) David Zoltowski (/author/david-zoltowski-11156), Jonathan W. Pillow (/author/jonathan-w-pillow-5713)
- Multiplicative Weights Updates with Constant Step-Size in Graphical Constant-Sum Games (/paper/7612-multiplicative-weights-updates-with-constant-step-size-in-graphical-constant-sum-games) Yun Kuen Cheung (/author/yun-kuen-cheung-11157)
- Why Is My Classifier Discriminatory? (/paper/7613-why-is-my-classifier-discriminatory) Irene Chen (/author/irene-chen-11158), Fredrik D. Johansson (/author/fredrik-d-johansson-8266), David Sontag (/author/david-sontag-3668)
- Multi-Layered Gradient Boosting Decision Trees (/paper/7614-multi-layered-gradient-boosting-decision-trees) Ji Feng (/author/ji-feng-11159), Yang Yu (/author/yang-yu-8150), Zhi-Hua Zhou (/author/zhi-hua-zhou-6485)
- Learn What Not to Learn: Action Elimination with Deep Reinforcement Learning (/paper/7615-learn-what-not-to-learn-action-elimination-with-deep-reinforcement-learning) Tom Zahavy (/author/tom-zahavy-9839), Matan Haroush (/author/matan-haroush-11160), Nadav Merlis (/author/nadav-merlis-11161), Daniel J. Mankowitz (/author/daniel-j-mankowitz-8913), Shie Mannor (/author/shie-mannor-2093)
- Communication Efficient Parallel Algorithms for Optimization on Manifolds (/paper/7616-communication-efficient-parallel-algorithms-for-optimization-on-manifolds) Bayan Saparbayeva (/author/bayan-saparbayeva-11162), Michael Zhang (/author/michael-zhang-11163), Lizhen Lin (/author/lizhen-lin-10428)
- Neural Code Comprehension: A Learnable Representation of Code Semantics (/paper/7617-neural-code-comprehension-a-learnable-representation-of-code-semantics) Tal Ben-Nun (/author/tal-ben-nun-11164), Alice Shoshana Jakobovits (/author/alice-shoshana-jakobovits-11165), Torsten Hoefler (/author/torsten-hoefler-11166)
- Tight Bounds for Collaborative PAC Learning via Multiplicative Weights (/paper/7618-tight-bounds-for-collaborative-pac-learning-via-multiplicative-weights) Jiecao Chen (/author/jiecao-chen-9268), Qin Zhang (/author/qin-zhang-9270), Yuan Zhou (/author/yuan-zhou-11074)
- BinGAN: Learning Compact Binary Descriptors with a Regularized GAN (/paper/7619-bingan-learning-compact-binary-descriptors-with-a-regularized-gan) Maciej Zieba

(/author/maciej-zieba-11167), Piotr Semberecki (/author/piotr-semberecki-11168), Tarek El-Gaaly (/author/tarek-el-gaaly-11169), Tomasz Trzcinski (/author/tomasz-trzcinski-6818)

- Modern Neural Networks Generalize on Small Data Sets (/paper/7620-modern-neural-networks-generalize-on-small-data-sets) Matthew Olson (/author/matthew-olson-11170), Abraham Wyner (/author/abraham-wyner-11171), Richard Berk (/author/richard-berk-11172)
- Escaping Saddle Points in Constrained Optimization (/paper/7621-escaping-saddle-points-in-constrained-optimization) Aryan Mokhtari (/author/aryan-mokhtari-8769), Asuman Ozdaglar (/author/asuman-ozdaglar-6501), Ali Jadbabaie (/author/ali-jadbabaie-6451)
- Adversarial Attacks on Stochastic Bandits (/paper/7622-adversarial-attacks-on-stochastic-bandits) Kwang-Sung Jun (/author/kwang-sung-jun-8224), Lihong Li (/author/lihong-li-4101), Yuzhe Ma (/author/yuzhe-ma-11173), Jerry Zhu (/author/jerry-zhu-2978)
- Optimal Subsampling with Influence Functions (/paper/7623-optimal-subsampling-with-influence-functions) Daniel Ting (/author/daniel-ting-11174), Eric Brochu (/author/eric-brochu-11175)
- A Bandit Approach to Sequential Experimental Design with False Discovery Control (/paper/7624-a-bandit-approach-to-sequential-experimental-design-with-false-discovery-control) Kevin G. Jamieson (/author/kevin-g-jamieson-5174), Lalit Jain (/author/lalit-jain-8204)
- Equality of Opportunity in Classification: A Causal Approach (/paper/7625-equality-of-opportunity-in-classification-a-causal-approach) Junzhe Zhang (/author/junzhe-zhang-11176), Elias Bareinboim (/author/elias-bareinboim-6583)
- Towards Understanding Acceleration Tradeoff between Momentum and Asynchrony in Nonconvex Stochastic Optimization (/paper/7626-towards-understanding-acceleration-tradeoff-between-momentum-and-asynchrony-in-nonconvex-stochastic-optimization) Tianyi Liu (/author/tianyi-liu-11177), Shiyang Li (/author/shiyang-li-11178), Jianping Shi (/author/jianping-shi-10657), Enlu Zhou (/author/enlu-zhou-11179), Tuo Zhao (/author/tuo-zhao-5638)
- Unsupervised Attention-guided Image-to-Image Translation (/paper/7627-unsupervised-attention-guided-image-to-image-translation) Youssef Alami Mejjati (/author/youssef-alami-mejjati-11180), Christian Richardt (/author/christian-richardt-11181), James Tompkin (/author/james-tompkin-11182), Darren Cosker (/author/darren-cosker-11183), Kwang In Kim (/author/kwang-in-kim-11184)
- Inferring Networks From Random Walk-Based Node Similarities (/paper/7628-inferring-networks-from-random-walk-based-node-similarities) Jeremy Hoskins (/author/jeremy-hoskins-11185), Cameron Musco (/author/cameron-musco-8032), Christopher Musco (/author/christopher-musco-8033), Babis Tsourakakis (/author/babis-tsourakakis-11186)
- NEON2: Finding Local Minima via First-Order Oracles (/paper/7629-neon2-finding-local-minima-via-first-order-oracles) Zeyuan Allen-Zhu (/author/zeyuan-allen-zhu-8933), Yuanzhi Li (/author/yuanzhi-li-8643)

- Zeroth-Order Stochastic Variance Reduction for Nonconvex Optimization (/paper/7630-zeroth-order-stochastic-variance-reduction-for-nonconvex-optimization) Sijia Liu (/author/sijia-liu-11187), Bhavya Kailkhura (/author/bhavya-kailkhura-11188), Pin-Yu Chen (/author/pin-yu-chen-10619), Paishun Ting (/author/paishun-ting-10621), Shiyu Chang (/author/shiyu-chang-9374), Lisa Amini (/author/lisa-amini-11189)
- Online Structured Laplace Approximations for Overcoming Catastrophic Forgetting (/paper/7631-online-structured-laplace-approximations-for-overcoming-catastrophic-forgetting) Hippolyt Ritter (/author/hippolyt-ritter-11190), Aleksandar Botev (/author/aleksandar-botev-11191), David Barber (/author/david-barber-1451)
- DeepProbLog: Neural Probabilistic Logic Programming (/paper/7632-deepproblog-neural-probabilistic-logic-programming) Robin Manhaeve (/author/robin-manhaeve-11192), Sebastijan Dumancic (/author/sebastijan-dumancic-11193), Angelika Kimmig (/author/angelika-kimmig-9340), Thomas Demeester (/author/thomas-demeester-11194), Luc De Raedt (/author/luc-de-raedt-11195)
- Convergence of Cubic Regularization for Nonconvex Optimization under KL Property (/paper/7633-convergence-of-cubic-regularization-for-nonconvex-optimization-under-kl-property) Yi Zhou (/author/yi-zhou-8417), Zhe Wang (/author/zhe-wang-11196), Yingbin Liang (/author/yingbin-liang-8418)
- Direct Estimation of Differences in Causal Graphs (/paper/7634-direct-estimation-of-differences-in-causal-graphs) Yuhao Wang (/author/yuhao-wang-10267), Chandler Squires (/author/chandler-squires-11197), Anastasiya Belyaeva (/author/anastasiya-belyaeva-11198), Caroline Uhler (/author/caroline-uhler-10270)
- Sublinear Time Low-Rank Approximation of Distance Matrices (/paper/7635-sublinear-time-low-rank-approximation-of-distance-matrices) Ainesh Bakshi (/author/ainesh-bakshi-11199), David Woodruff (/author/david-woodruff-6661)
- Variational PDEs for Acceleration on Manifolds and Application to Diffeomorphisms (/paper/7636-variational-pdes-for-acceleration-on-manifolds-and-application-to-diffeomorphisms) Ganesh Sundaramoorthi (/author/ganesh-sundaramoorthi-11200), Anthony Yezzi (/author/anthony-yezzi-11201)
- Bayesian Inference of Temporal Task Specifications from Demonstrations (/paper/7637-bayesian-inference-of-temporal-task-specifications-from-demonstrations) Ankit Shah (/author/ankit-shah-11202), Pritish Kamath (/author/pritish-kamath-11203), Julie A. Shah (/author/julie-a-shah-7253), Shen Li (/author/shen-li-11204)
- Data center cooling using model-predictive control (/paper/7638-data-center-cooling-using-model-predictive-control) Nevena Lazic (/author/nevena-lazic-11205), Craig Boutilier (/author/craig-boutilier-2040), Tyler Lu (/author/tyler-lu-11206), Eehern Wong (/author/eehern-wong-11207), Binz Roy (/author/binz-roy-11208), MK Ryu (/author/mk-ryu-11209), Greg Imwalle (/author/greg-imwalle-11210)
- Acceleration through Optimistic No-Regret Dynamics (/paper/7639-acceleration-through-optimistic-no-regret-dynamics) Jun-Kun Wang (/author/jun-kun-wang-10363), Jacob D. Abernethy (/author/jacob-d-abernethy-4782)

- Lipschitz regularity of deep neural networks: analysis and efficient estimation (/paper/7640-lipschitz-regularity-of-deep-neural-networks-analysis-and-efficient-estimation) Aladin Virmaux (/author/aladin-virmaux-11211), Kevin Scaman (/author/kevin-scaman-7317)
- Minimax Estimation of Neural Net Distance (/paper/7641-minimax-estimation-of-neural-net-distance) Kaiyi Ji (/author/kaiyi-ji-11212), Yingbin Liang (/author/yingbin-liang-8418)
- Leveraging the Exact Likelihood of Deep Latent Variable Models (/paper/7642-leveraging-the-exact-likelihood-of-deep-latent-variable-models) Pierre-Alexandre Mattei (/author/pierre-alexandre-mattei-11213), Jes Frellsen (/author/jes-frellsen-9900)
- Bipartite Stochastic Block Models with Tiny Clusters (/paper/7643-bipartite-stochastic-block-models-with-tiny-clusters) Stefan Neumann (/author/stefan-neumann-11214)
- Learning sparse neural networks via sensitivity-driven regularization (/paper/7644-learning-sparse-neural-networks-via-sensitivity-driven-regularization) Enzo Tartaglione (/author/enzo-tartaglione-11215), Skjalg Lepsøy (/author/skjalg-lepsy-11216), Attilio Fiandrotti (/author/attilio-fiandrotti-11217), Gianluca Francini (/author/gianluca-francini-11218)
- Faster Online Learning of Optimal Threshold for Consistent F-measure Optimization (/paper/7645-faster-online-learning-of-optimal-threshold-for-consistent-f-measure-optimization) Xiaoxuan Zhang (/author/xiaoxuan-zhang-11219), Mingrui Liu (/author/mingrui-liu-9569), Xun Zhou (/author/xun-zhou-11220), Tianbao Yang (/author/tianbao-yang-5489)
- Direct Runge-Kutta Discretization Achieves Acceleration (/paper/7646-direct-runge-kutta-discretization-achieves-acceleration) Jingzhao Zhang (/author/jingzhao-zhang-11221), Aryan Mokhtari (/author/aryan-mokhtari-8769), Suvrit Sra (/author/suvrit-sra-2850), Ali Jadbabaie (/author/ali-jadbabaie-6451)
- Adversarial Examples that Fool both Computer Vision and Time-Limited Humans (/paper/7647-adversarial-examples-that-fool-both-computer-vision-and-time-limited-humans) Gamaleldin Elsayed (/author/gamaleldin-elsayed-10678), Shreya Shankar (/author/shreya-shankar-11222), Brian Cheung (/author/brian-cheung-11223), Nicolas Papernot (/author/nicolas-papernot-11224), Alexey Kurakin (/author/alexey-kurakin-11225), Ian Goodfellow (/author/ian-goodfellow-4369), Jascha Sohl-Dickstein (/author/jascha-sohl-dickstein-7915)
- Stochastic Nested Variance Reduced Gradient Descent for Nonconvex Optimization (/paper/7648-stochastic-nested-variance-reduced-gradient-descent-for-nonconvex-optimization) Dongruo Zhou (/author/dongruo-zhou-11226), Pan Xu (/author/pan-xu-9210), Quanquan Gu (/author/quanquan-gu-5348)
- Faster Neural Networks Straight from JPEG (/paper/7649-faster-neural-networks-straight-from-jpeg) Lionel Gueguen (/author/lionel-gueguen-11227), Alex Sergeev (/author/alex-sergeev-11228), Ben Kadlec (/author/ben-kadlec-11229), Rosanne Liu (/author/rosanne-liu-11230), Jason Yosinski (/author/jason-yosinski-7288)

- TopRank: A practical algorithm for online stochastic ranking (/paper/7650-toprank-a-practical-algorithm-for-online-stochastic-ranking) Tor Lattimore (/author/tor-lattimore-7451), Branislav Kveton (/author/branislav-kveton-2628), Shuai Li (/author/shuai-li-11231), Csaba Szepesvari (/author/csaba-szepesvari-6449)
- Learning from discriminative feature feedback (/paper/7651-learning-from-discriminative-feature-feedback) Sanjoy Dasgupta (/author/sanjoy-dasgupta-2220), Akansha Dey (/author/akansha-dey-11233), Nicholas Roberts (/author/nicholas-roberts-11232), Sivan Sabato (/author/sivan-sabato-4663)
- RetGK: Graph Kernels based on Return Probabilities of Random Walks (/paper/7652-retgk-graph-kernels-based-on-return-probabilities-of-random-walks) Zhen Zhang (/author/zhen-zhang-11234), Mianzhi Wang (/author/mianzhi-wang-11235), Yijian Xiang (/author/yijian-xiang-11236), Yan Huang (/author/yan-huang-8094), Arye Nehorai (/author/arye-nehorai-7327)
- Deep Generative Markov State Models (/paper/7653-deep-generative-markov-state-models) Hao Wu (/author/hao-wu-8672), Andreas Mardt (/author/andreas-mardt-11237), Luca Pasquali (/author/luca-pasquali-11238), Frank Noe (/author/frank-noe-8673)
- Early Stopping for Nonparametric Testing (/paper/7654-early-stopping-for-nonparametric-testing) Meimei Liu (/author/meimei-liu-11239), Guang Cheng (/author/guang-cheng-8287)
- Solving Non-smooth Constrained Programs with Lower Complexity than $\mathcal{O}(1/\varepsilon)$: A Primal-Dual Homotopy Smoothing Approach (/paper/7655-solving-non-smooth-constrained-programs-with-lower-complexity-than-mathcalo1varepsilon-a-primal-dual-homotopy-smoothing-approach) Xiaohan Wei (/author/xiaohan-wei-9599), Hao Yu (/author/hao-yu-9597), Qing Ling (/author/qing-ling-11240), Michael Neely (/author/michael-neely-9598)
- Heterogeneous Bitwidth Binarization in Convolutional Neural Networks (/paper/7656-heterogeneous-bitwidth-binarization-in-convolutional-neural-networks) Joshua Fromm (/author/joshua-fromm-11241), Shwetak Patel (/author/shwetak-patel-11242), Matthai Philipose (/author/matthai-philipose-11243)
- Unsupervised Learning of Object Landmarks through Conditional Image Generation (/paper/7657-unsupervised-learning-of-object-landmarks-through-conditional-image-generation) Tomas Jakab (/author/tomas-jakab-11244), Ankush Gupta (/author/ankush-gupta-11245), Hakan Bilen (/author/hakan-bilen-8995), Andrea Vedaldi (/author/andrea-vedaldi-3408)
- Probabilistic Neural Programmed Networks for Scene Generation (/paper/7658-probabilistic-neural-programmed-networks-for-scene-generation) Zhiwei Deng (/author/zhiwei-deng-11246), Jiacheng Chen (/author/jiacheng-chen-11247), YIFANG FU (/author/yifang-fu-11248), Greg Mori (/author/greg-mori-4070)
- The streaming rollout of deep networks - towards fully model-parallel execution (/paper/7659-the-streaming-rollout-of-deep-networks-towards-fully-model-parallel-execution) Volker Fischer (/author/volker-fischer-11249), Jan Koehler (/author/jan-koehler-11250), Thomas Pfeil (/author/thomas-pfeil-11251)

- KONG: Kernels for ordered-neighborhood graphs (/paper/7660-kong-kernels-for-ordered-neighborhood-graphs) Moez Draief (/author/moez-draief-11252), Konstantin Kutzkov (/author/konstantin-kutzkov-11253), Kevin Scaman (/author/kevin-scaman-7317), Milan Vojnovic (/author/milan-vojnovic-7125)
- GumBolt: Extending Gumbel trick to Boltzmann priors (/paper/7661-gumbolt-extending-gumbel-trick-to-boltzmann-priors) Amir H. Khoshaman (/author/amir-h-khoshaman-11254), Mohammad Amin (/author/mohammad-amin-11255)
- Neural Networks Trained to Solve Differential Equations Learn General Representations (/paper/7662-neural-networks-trained-to-solve-differential-equations-learn-general-representations) Martin Magill (/author/martin-magill-11256), Faisal Qureshi (/author/faisal-qureshi-11257), Hendrick de Haan (/author/hendrick-de-haan-11258)
- Beauty-in-averageness and its contextual modulations: A Bayesian statistical account (/paper/7663-beauty-in-averageness-and-its-contextual-modulations-a-bayesian-statistical-account) Chaitanya Ryali (/author/chaitanya-ryali-11011), Angela J. Yu (/author/angela-j-yu-2266)
- Distributed Weight Consolidation: A Brain Segmentation Case Study (/paper/7664-distributed-weight-consolidation-a-brain-segmentation-case-study) Patrick McClure (/author/patrick-mcclure-11259), Charles Y. Zheng (/author/charles-y-zheng-7510), Jakub Kaczmarzyk (/author/jakub-kaczmarzyk-11260), John Rogers-Lee (/author/john-rogers-lee-11261), Satra Ghosh (/author/satra-ghosh-11262), Dylan Nielson (/author/dylan-nielson-11263), Peter A. Bandettini (/author/peter-a-bandettini-11264), Francisco Pereira (/author/francisco-pereira-2842)
- Efficient Projection onto the Perfect Phylogeny Model (/paper/7665-efficient-projection-onto-the-perfect-phylogeny-model) Bei Jia (/author/bei-jia-11265), Surjyendu Ray (/author/surjyendu-ray-11266), Sam Safavi (/author/sam-safavi-11267), José Bento (/author/jose-bento-6511)
- TETRIS: TilE-matching the TRemendous Irregular Sparsity (/paper/7666-tetris-tile-matching-the-tremendous-irregular-sparsity) Yu Ji (/author/yu-ji-11268), Ling Liang (/author/ling-liang-11269), Lei Deng (/author/lei-deng-10625), Youyang Zhang (/author/youyang-zhang-11270), Youhui Zhang (/author/youhui-zhang-11271), Yuan Xie (/author/yuan-xie-10628)
- Cooperative neural networks (CoNN): Exploiting prior independence structure for improved classification (/paper/7667-cooperative-neural-networks-conn-exploiting-prior-independence-structure-for-improved-classification) Harsh Shrivastava (/author/harsh-shrivastava-11272), Eugene Bart (/author/eugene-bart-11273), Bob Price (/author/bob-price-11274), Hanjun Dai (/author/hanjun-dai-7969), Bo Dai (/author/bo-dai-6604), Srinivas Aluru (/author/srinivas-aluru-11275)
- Differentially Private Robust Low-Rank Approximation (/paper/7668-differentially-private-robust-low-rank-approximation) Raman Arora (/author/raman-arora-4325), Vladimir braverman (/author/vladimir-braverman-11276), Jalaj Upadhyay (/author/jalaj-upadhyay-11277)

- Meta-Learning MCMC Proposals (/paper/7669-meta-learning-mcmc-proposals) Tongzhou Wang (/author/tongzhou-wang-11278), YI WU (/author/yi-wu-8464), Dave Moore (/author/dave-moore-11279), Stuart J. Russell (/author/stuart-j-russell-2312)
- An Information-Theoretic Analysis for Thompson Sampling with Many Actions (/paper/7670-an-information-theoretic-analysis-for-thompson-sampling-with-many-actions) Shi Dong (/author/shi-dong-11280), Benjamin Van Roy (/author/benjamin-van-roy-1265)
- Flexible and accurate inference and learning for deep generative models (/paper/7671-flexible-and-accurate-inference-and-learning-for-deep-generative-models) Eszter Vértes (/author/eszter-vertes-11281), Maneesh Sahani (/author/maneesh-sahani-1670)
- The Price of Privacy for Low-rank Factorization (/paper/7672-the-price-of-privacy-for-low-rank-factorization) Jalaj Upadhyay (/author/jalaj-upadhyay-11277)
- Regret Bounds for Robust Adaptive Control of the Linear Quadratic Regulator (/paper/7673-regret-bounds-for-robust-adaptive-control-of-the-linear-quadratic-regulator) Sarah Dean (/author/sarah-dean-11282), Horia Mania (/author/horia-mania-10851), Nikolai Matni (/author/nikolai-matni-11283), Benjamin Recht (/author/benjamin-recht-3537), Stephen Tu (/author/stephen-tu-9347)
- Bilevel Distance Metric Learning for Robust Image Recognition (/paper/7674-bilevel-distance-metric-learning-for-robust-image-recognition) Jie Xu (/author/jie-xu-11096), Lei Luo (/author/lei-luo-11284), Cheng Deng (/author/cheng-deng-10457), Heng Huang (/author/heng-huang-4609)
- Differentially Private Uniformly Most Powerful Tests for Binomial Data (/paper/7675-differentially-private-uniformly-most-powerful-tests-for-binomial-data) Jordan Awan (/author/jordan-awan-11285), Aleksandra Slavković (/author/aleksandra-slavkovic-11286)
- Scalable Coordinated Exploration in Concurrent Reinforcement Learning (/paper/7676-scalable-coordinated-exploration-in-concurrent-reinforcement-learning) Maria Dimakopoulou (/author/maria-dimakopoulou-11287), Ian Osband (/author/ian-osband-6756), Benjamin Van Roy (/author/benjamin-van-roy-1265)
- Integrated accounts of behavioral and neuroimaging data using flexible recurrent neural network models (/paper/7677-integrated-accounts-of-behavioral-and-neuroimaging-data-using-flexible-recurrent-neural-network-models) Amir Dezfouli (/author/amir-dezfouli-7941), Richard Morris (/author/richard-morris-11288), Fabio T. Ramos (/author/fabio-t-ramos-7238), Peter Dayan (/author/peter-dayan-632), Bernard Balleine (/author/bernard-balleine-11289)
- BML: A High-performance, Low-cost Gradient Synchronization Algorithm for DML Training (/paper/7678-bml-a-high-performance-low-cost-gradient-synchronization-algorithm-for-dml-training) Songtao Wang (/author/songtao-wang-11290), Dan Li (/author/dan-li-11291), Yang Cheng (/author/yang-cheng-11292), Jinkun Geng (/author/jinkun-geng-11293), Yanshu Wang (/author/yanshu-wang-11294), Shuai Wang (/author/shuai-wang-11295), Shu-Tao Xia (/author/shu-tao-xia-10131), Jianping Wu (/author/jianping-wu-11296)
- Inexact trust-region algorithms on Riemannian manifolds (/paper/7679-inexact-trust-region-algorithms-on-riemannian-manifolds) Hiroyuki Kasai (/author/hiroyuki-kasai-11297),

Bamdev Mishra (/author/bamdev-mishra-11298)

- Can We Gain More from Orthogonality Regularizations in Training Deep Networks? (/paper/7680-can-we-gain-more-from-orthogonality-regularizations-in-training-deep-networks) Nitin Bansal (/author/nitin-bansal-11299), Xiaohan Chen (/author/xiaohan-chen-11300), Zhangyang Wang (/author/zhangyang-wang-7234)
- Binary Rating Estimation with Graph Side Information (/paper/7681-binary-rating-estimation-with-graph-side-information) Kwangjun Ahn (/author/kwangjun-ahn-11301), Kangwook Lee (/author/kangwook-lee-11302), Hyunseung Cha (/author/hyunseung-cha-11303), Changho Suh (/author/changho-suh-9636)
- SimplE Embedding for Link Prediction in Knowledge Graphs (/paper/7682-simple-embedding-for-link-prediction-in-knowledge-graphs) Seyed Mehran Kazemi (/author/seyed-mehran-kazemi-9339), David Poole (/author/david-poole-2324)
- Differentially Private Contextual Linear Bandits (/paper/7683-differentially-private-contextual-linear-bandits) Roshan Shariff (/author/roshan-shariff-11304), Or Sheffet (/author/or-sheffet-7425)
- Submodular Field Grammars: Representation, Inference, and Application to Image Parsing (/paper/7684-submodular-field-grammars-representation-inference-and-application-to-image-parsing) Abram L. Friesen (/author/abram-l-friesen-5073), Pedro M. Domingos (/author/pedro-m-domingos-7410)
- A Bridging Framework for Model Optimization and Deep Propagation (/paper/7685-a-bridging-framework-for-model-optimization-and-deep-propagation) Risheng Liu (/author/risheng-liu-5180), Shichao Cheng (/author/shichao-cheng-11305), xiaokun liu (/author/xiaokun-liu-11306), Long Ma (/author/long-ma-11307), Xin Fan (/author/xin-fan-11308), Zhongxuan Luo (/author/zhongxuan-luo-11309)
- Completing State Representations using Spectral Learning (/paper/7686-completing-state-representations-using-spectral-learning) Nan Jiang (/author/nan-jiang-9663), Alex Kulesza (/author/alex-kulesza-3510), Satinder Singh (/author/satinder-singh-6765)
- Optimization of Smooth Functions with Noisy Observations: Local Minimax Rates (/paper/7687-optimization-of-smooth-functions-with-noisy-observations-local-minimax-rates) Yining Wang (/author/yining-wang-7526), Sivaraman Balakrishnan (/author/sivaraman-balakrishnan-4913), Aarti Singh (/author/aarti-singh-4037)
- Adding One Neuron Can Eliminate All Bad Local Minima (/paper/7688-adding-one-neuron-can-eliminate-all-bad-local-minima) SHIYU LIANG (/author/shiyu-liang-11310), Ruoyu Sun (/author/ruoyu-sun-8432), Jason D. Lee (/author/jason-d-lee-4783), R. Srikant (/author/r-srikant-8403)
- Mean-field theory of graph neural networks in graph partitioning (/paper/7689-mean-field-theory-of-graph-neural-networks-in-graph-partitioning) Tatsuro Kawamoto (/author/tatsuro-kawamoto-11311), Masashi Tsubaki (/author/masashi-tsubaki-11312), Tomoyuki Obuchi (/author/tomoyuki-obuchi-11313)
- The Physical Systems Behind Optimization Algorithms (/paper/7690-the-physical-systems-behind-optimization-algorithms) Lin Yang (/author/lin-yang-9084), Raman Arora

(/author/raman-arora-4325), Vladimir braverman (/author/vladimir-braverman-11276), Tuo Zhao (/author/tuo-zhao-5638)

- Mallows Models for Top-k Lists (/paper/7691-mallows-models-for-top-k-lists) Flavio Chierichetti (/author/flavio-chierichetti-5099), Anirban Dasgupta (/author/anirban-dasgupta-5491), Shahrzad Haddadan (/author/shahrzad-haddadan-11314), Ravi Kumar (/author/ravi-kumar-4090), Silvio Lattanzi (/author/silvio-lattanzi-7248)
- Amortized Inference Regularization (/paper/7692-amortized-inference-regularization) Rui Shu (/author/rui-shu-11315), Hung H. Bui (/author/hung-h-bui-8377), Shengjia Zhao (/author/shengjia-zhao-9146), Mykel J. Kochenderfer (/author/mykel-j-kochenderfer-11316), Stefano Ermon (/author/stefano-ermon-5196)
- Maximum Causal Tsallis Entropy Imitation Learning (/paper/7693-maximum-causal-tsallis-entropy-imitation-learning) Kyungjae Lee (/author/kyungjae-lee-11317), Sungjoon Choi (/author/sungjoon-choi-11318), Songhwai Oh (/author/songhwai-oh-11319)
- Limited Memory Kelley's Method Converges for Composite Convex and Submodular Objectives (/paper/7694-limited-memory-kelleys-method-converges-for-composite-convex-and-submodular-objectives) Song Zhou (/author/song-zhou-11320), Swati Gupta (/author/swati-gupta-11321), Madeleine Udell (/author/madeleine-udell-9049)
- Semi-Supervised Learning with Declaratively Specified Entropy Constraints (/paper/7695-semi-supervised-learning-with-declaratively-specified-entropy-constraints) Haitian Sun (/author/haitian-sun-11322), William W. Cohen (/author/william-w-cohen-1648), Lidong Bing (/author/lidong-bing-11323)
- End-to-end Symmetry Preserving Inter-atomic Potential Energy Model for Finite and Extended Systems (/paper/7696-end-to-end-symmetry-preserving-inter-atomic-potential-energy-model-for-finite-and-extended-systems) Linfeng Zhang (/author/linfeng-zhang-11324), Jiequn Han (/author/jiequn-han-11325), Han Wang (/author/han-wang-11326), Wissam Saidi (/author/wissam-saidi-11327), Roberto Car (/author/roberto-car-11328), Weinan E (/author/weinan-e-11329)
- Sparsified SGD with Memory (/paper/7697-sparsified-sgd-with-memory) Sebastian U. Stich (/author/sebastian-u-stich-10062), Jean-Baptiste Cordonnier (/author/jean-baptiste-cordonnier-11330), Martin Jaggi (/author/martin-jaggi-7651)
- Exponentiated Strongly Rayleigh Distributions (/paper/7698-exponentiated-strongly-rayleigh-distributions) Zelda E. Mariet (/author/zelda-e-mariet-8823), Suvrit Sra (/author/suvrit-sra-2850), Stefanie Jegelka (/author/stefanie-jegelka-3773)
- Importance Weighting and Variational Inference (/paper/7699-importance-weighting-and-variational-inference) Justin Domke (/author/justin-domke-4762), Daniel R. Sheldon (/author/daniel-r-sheldon-3728)
- Transfer Learning from Speaker Verification to Multispeaker Text-To-Speech Synthesis (/paper/7700-transfer-learning-from-speaker-verification-to-multispeaker-text-to-speech-synthesis) Ye Jia (/author/ye-jia-11331), Yu Zhang (/author/yu-zhang-4618), Ron Weiss (/author/ron-weiss-11332), Quan Wang (/author/quan-wang-11333), Jonathan Shen (/author/jonathan-shen-11334), Fei Ren (/author/fei-ren-11335), zhifeng Chen

(/author/zhifeng-chen-9242), Patrick Nguyen (/author/patrick-nguyen-11336), Ruoming Pang (/author/ruoming-pang-11337), Ignacio Lopez Moreno (/author/ignacio-lopez-moreno-11338), Yonghui Wu (/author/yonghui-wu-9243)

- Expanding Holographic Embeddings for Knowledge Completion (/paper/7701-expanding-holographic-embeddings-for-knowledge-completion) Yexiang Xue (/author/yexiang-xue-9109), Yang Yuan (/author/yang-yuan-9004), Zhitian Xu (/author/zhitian-xu-11339), Ashish Sabharwal (/author/ashish-sabharwal-3334)
- Lifelong Inverse Reinforcement Learning (/paper/7702-lifelong-inverse-reinforcement-learning) Jorge Armando Mendez Mendez (/author/jorge-armando-mendez-mendez-11340), Shashank Shivkumar (/author/shashank-shivkumar-11341), Eric Eaton (/author/eric-eaton-11342)
- Explaining Deep Learning Models -- A Bayesian Non-parametric Approach (/paper/7703-explaining-deep-learning-models-a-bayesian-non-parametric-approach) Wenbo Guo (/author/wenbo-guo-11343), Sui Huang (/author/sui-huang-11344), Yunzhe Tao (/author/yunzhe-tao-10599), Xinyu Xing (/author/xinyu-xing-11345), Lin Lin (/author/lin-lin-11346)
- Third-order Smoothness Helps: Faster Stochastic Optimization Algorithms for Finding Local Minima (/paper/7704-third-order-smoothness-helps-faster-stochastic-optimization-algorithms-for-finding-local-minima) Yaodong Yu (/author/yaodong-yu-11347), Pan Xu (/author/pan-xu-9210), Quanquan Gu (/author/quanquan-gu-5348)
- COLA: Decentralized Linear Learning (/paper/7705-cola-decentralized-linear-learning) Lie He (/author/lie-he-11348), An Bian (/author/an-bian-9447), Martin Jaggi (/author/martin-jaggi-7651)
- MiME: Multilevel Medical Embedding of Electronic Health Records for Predictive Healthcare (/paper/7706-mime-multilevel-medical-embedding-of-electronic-health-records-for-predictive-healthcare) Edward Choi (/author/edward-choi-8854), Cao Xiao (/author/cao-xiao-11349), Walter Stewart (/author/walter-stewart-9328), Jimeng Sun (/author/jimeng-sun-8369)
- Adaptive Sampling Towards Fast Graph Representation Learning (/paper/7707-adaptive-sampling-towards-fast-graph-representation-learning) Wenbing Huang (/author/wenbing-huang-9901), Tong Zhang (/author/tong-zhang-1901), Yu Rong (/author/yu-rong-11350), Junzhou Huang (/author/junzhou-huang-5423)
- Hunting for Discriminatory Proxies in Linear Regression Models (/paper/7708-hunting-for-discriminatory-proxies-in-linear-regression-models) Samuel Yeom (/author/samuel-yeom-11351), Anupam Datta (/author/anupam-datta-11352), Matt Fredrikson (/author/matt-fredrikson-11353)
- Towards Robust Detection of Adversarial Examples (/paper/7709-towards-robust-detection-of-adversarial-examples) Tianyu Pang (/author/tianyu-pang-11354), Chao Du (/author/chao-du-11355), Yinpeng Dong (/author/yinpeng-dong-11356), Jun Zhu (/author/jun-zhu-6461)

- Active Matting (/paper/7710-active-matting) Xin Yang (/author/xin-yang-11357), Ke Xu (/author/ke-xu-8661), Shaozhe Chen (/author/shaozhe-chen-11358), Shengfeng He (/author/shengfeng-he-11359), Baocai Yin Yin (/author/baocai-yin-yin-11360), Rynson Lau (/author/rynson-lau-11361)
- Learning filter widths of spectral decompositions with wavelets (/paper/7711-learning-filter-widths-of-spectral-decompositions-with-wavelets) Haidar Khan (/author/haidar-khan-11362), Bulent Yener (/author/bulent-yener-11363)
- Byzantine Stochastic Gradient Descent (/paper/7712-byzantine-stochastic-gradient-descent) Dan Alistarh (/author/dan-alistarh-8258), Zeyuan Allen-Zhu (/author/zeyuan-allen-zhu-8933), Jerry Li (/author/jerry-li-9638)
- PG-TS: Improved Thompson Sampling for Logistic Contextual Bandits (/paper/7713-pg-ts-improved-thompson-sampling-for-logistic-contextual-bandits) Bianca Dumitrascu (/author/bianca-dumitrascu-11364), Karen Feng (/author/karen-feng-11365), Barbara Engelhardt (/author/barbara-engelhardt-11366)
- Spectral Filtering for General Linear Dynamical Systems (/paper/7714-spectral-filtering-for-general-linear-dynamical-systems) Elad Hazan (/author/elad-hazan-3630), HOLDEN LEE (/author/holden-lee-11367), Karan Singh (/author/karan-singh-10382), Cyril Zhang (/author/cyril-zhang-10383), Yi Zhang (/author/yi-zhang-3967)
- On Learning Intrinsic Rewards for Policy Gradient Methods (/paper/7715-on-learning-intrinsic-rewards-for-policy-gradient-methods) Zeyu Zheng (/author/zeyu-zheng-11368), Junhyuk Oh (/author/junhyuk-oh-8191), Satinder Singh (/author/satinder-singh-6765)
- Boolean Decision Rules via Column Generation (/paper/7716-boolean-decision-rules-via-column-generation) Sanjeeb Dash (/author/sanjeeb-dash-11369), Oktay Gunluk (/author/oktay-gunluk-11370), Dennis Wei (/author/dennis-wei-8840)
- Adversarial Text Generation via Feature-Mover's Distance (/paper/7717-adversarial-text-generation-via-feature-movers-distance) Liqun Chen (/author/liqun-chen-10055), Shuyang Dai (/author/shuyang-dai-11371), Chenyang Tao (/author/chenyang-tao-11372), Haichao Zhang (/author/haichao-zhang-6277), Zhe Gan (/author/zhe-gan-7916), Dinghan Shen (/author/dinghan-shen-10024), Yizhe Zhang (/author/yizhe-zhang-8612), Guoyin Wang (/author/guoyin-wang-10025), Ruiyi Zhang (/author/ruiyi-zhang-11373), Lawrence Carin (/author/lawrence-carin-3001)
- Fast Rates of ERM and Stochastic Approximation: Adaptive to Error Bound Conditions (/paper/7718-fast-rates-of-erm-and-stochastic-approximation-adaptive-to-error-bound-conditions) Mingrui Liu (/author/mingrui-liu-9569), Xiaoxuan Zhang (/author/xiaoxuan-zhang-11219), Lijun Zhang (/author/lijun-zhang-6394), Jing Rong (/author/jing-rong-10478), Tianbao Yang (/author/tianbao-yang-5489)
- Learning Bounds for Greedy Approximation with Explicit Feature Maps from Multiple Kernels (/paper/7719-learning-bounds-for-greedy-approximation-with-explicit-feature-maps-from-multiple-kernels) Shahin Shahrampour (/author/shahin-shahrampour-6450), Vahid Tarokh (/author/vahid-tarokh-9905)

- A Mathematical Model For Optimal Decisions In A Representative Democracy (/paper/7720-a-mathematical-model-for-optimal-decisions-in-a-representative-democracy) Malik Magdon-Ismail (/author/malik-magdon-ismail-1568), Lirong Xia (/author/lirong-xia-5550)
- Negotiable Reinforcement Learning for Pareto Optimal Sequential Decision-Making (/paper/7721-negotiable-reinforcement-learning-for-pareto-optimal-sequential-decision-making) Nishant Desai (/author/nishant-desai-11374), Andrew Critch (/author/andrew-critch-11375), Stuart J. Russell (/author/stuart-j-russell-2312)
- Non-metric Similarity Graphs for Maximum Inner Product Search (/paper/7722-non-metric-similarity-graphs-for-maximum-inner-product-search) Stanislav Morozov (/author/stanislav-morozov-11376), Artem Babenko (/author/artem-babenko-11377)
- Recurrently Controlled Recurrent Networks (/paper/7723-recurrently-controlled-recurrent-networks) Yi Tay (/author/yi-tay-11378), Anh Tuan Luu (/author/anh-tuan-luu-11379), Siu Cheung Hui (/author/siu-cheung-hui-11380)
- Fast greedy algorithms for dictionary selection with generalized sparsity constraints (/paper/7724-fast-greedy-algorithms-for-dictionary-selection-with-generalized-sparsity-constraints) Kaito Fujii (/author/kaito-fujii-8453), Tasuku Soma (/author/tasuku-soma-8293)
- Deep Reinforcement Learning in a Handful of Trials using Probabilistic Dynamics Models (/paper/7725-deep-reinforcement-learning-in-a-handful-of-trials-using-probabilistic-dynamics-models) Kurtland Chua (/author/kurtland-chua-11381), Roberto Calandra (/author/roberto-calandra-11382), Rowan McAllister (/author/rowan-mcallister-9688), Sergey Levine (/author/sergey-levine-4516)
- A Smoother Way to Train Structured Prediction Models (/paper/7726-a-smoother-way-to-train-structured-prediction-models) Venkata Krishna Pillutla (/author/venkata-krishna-pillutla-11383), Vincent Roulet (/author/vincent-roulet-9546), Sham M. Kakade (/author/sham-m-kakade-2572), Zaid Harchaoui (/author/zaid-harchaoui-7292)
- Context-dependent upper-confidence bounds for directed exploration (/paper/7727-context-dependent-upper-confidence-bounds-for-directed-exploration) Raksha Kumaraswamy (/author/raksha-kumaraswamy-11384), Matthew Schlegel (/author/matthew-schlegel-11385), Adam White (/author/adam-white-4743), Martha White (/author/martha-white-4533)
- A Unified View of Piecewise Linear Neural Network Verification (/paper/7728-a-unified-view-of-piecewise-linear-neural-network-verification) Rudy R. Bunel (/author/rudy-r-bunel-9015), Ilker Turkaslan (/author/ilker-turkaslan-11386), Philip Torr (/author/philip-torr-3805), Pushmeet Kohli (/author/pushmeet-kohli-4400), Pawan K. Mudigonda (/author/pawan-k-mudigonda-8606)
- Hierarchical Graph Representation Learning with Differentiable Pooling (/paper/7729-hierarchical-graph-representation-learning-with-differentiable-pooling) Zhitao Ying (/author/zhitao-ying-9531), Jiaxuan You (/author/jiaxuan-you-11387), Christopher Morris (/author/christopher-morris-11388), Xiang Ren (/author/xiang-ren-11389), Will Hamilton (/author/will-hamilton-9530), Jure Leskovec (/author/jure-leskovec-4767)

- Non-Ergodic Alternating Proximal Augmented Lagrangian Algorithms with Optimal Rates (/paper/7730-non-ergodic-alternating-proximal-augmented-lagrangian-algorithms-with-optimal-rates) Quoc Tran Dinh (/author/quoc-tran-dinh-8155)
- Information-based Adaptive Stimulus Selection to Optimize Communication Efficiency in Brain-Computer Interfaces (/paper/7731-information-based-adaptive-stimulus-selection-to-optimize-communication-efficiency-in-brain-computer-interfaces) Boyla Mainsah (/author/boyla-mainsah-11390), Dmitry Kalika (/author/dmitry-kalika-11391), Leslie Collins (/author/leslie-collins-11392), Siyuan Liu (/author/siyuan-liu-11393), Chandra Throckmorton (/author/chandra-throckmorton-11394)
- Porcupine Neural Networks: Approximating Neural Network Landscapes (/paper/7732-porcupine-neural-networks-approximating-neural-network-landscapes) Soheil Feizi (/author/soheil-feizi-7662), Hamid Javadi (/author/hamid-javadi-9578), Jesse Zhang (/author/jesse-zhang-11395), David Tse (/author/david-tse-7988)
- Fairness Through Computationally-Bounded Awareness (/paper/7733-fairness-through-computationally-bounded-awareness) Michael Kim (/author/michael-kim-11396), Omer Reingold (/author/omer-reingold-8383), Guy Rothblum (/author/guy-rothblum-11397)
- Adaptive Negative Curvature Descent with Applications in Non-convex Optimization (/paper/7734-adaptive-negative-curvature-descent-with-applications-in-non-convex-optimization) Mingrui Liu (/author/mingrui-liu-9569), Zhe Li (/author/zhe-li-9267), Xiaoyu Wang (/author/xiaoyu-wang-11398), Jinfeng Yi (/author/jinfeng-yi-5487), Tianbao Yang (/author/tianbao-yang-5489)
- Is Q-Learning Provably Efficient? (/paper/7735-is-q-learning-provably-efficient) Chi Jin (/author/chi-jin-5251), Zeyuan Allen-Zhu (/author/zeyuan-allen-zhu-8933), Sebastien Bubeck (/author/sebastien-bubeck-6650), Michael I. Jordan (/author/michael-i-jordan-6338)
- Interpreting Neural Network Judgments via Minimal, Stable, and Symbolic Corrections (/paper/7736-interpreting-neural-network-judgments-via-minimal-stable-and-symbolic-corrections) Xin Zhang (/author/xin-zhang-11399), Armando Solar-Lezama (/author/armando-solar-lezama-8106), Rishabh Singh (/author/rishabh-singh-9691)
- Measures of distortion for machine learning (/paper/7737-measures-of-distortion-for-machine-learning) Leena Chennuru Vankadara (/author/leena-chennuru-vankadara-11400), Ulrike von Luxburg (/author/ulrike-von-luxburg-10397)
- On the Local Minima of the Empirical Risk (/paper/7738-on-the-local-minima-of-the-empirical-risk) Chi Jin (/author/chi-jin-5251), Lydia T. Liu (/author/lydia-t-liu-11401), Rong Ge (/author/rong-ge-6993), Michael I. Jordan (/author/michael-i-jordan-6338)
- Densely Connected Attention Propagation for Reading Comprehension (/paper/7739-densely-connected-attention-propagation-for-reading-comprehension) Yi Tay (/author/yi-tay-11378), Anh Tuan Luu (/author/anh-tuan-luu-11379), Siu Cheung Hui (/author/siu-cheung-hui-11380), Jian Su (/author/jian-su-11402)
- Bandit Learning with Positive Externalities (/paper/7740-bandit-learning-with-positive-externalities) Virag Shah (/author/virag-shah-11403), Jose Blanchet (/author/jose-blanchet-

11404), Ramesh Johari (/author/ramesh-johari-4920)

- Learning Confidence Sets using Support Vector Machines (/paper/7741-learning-confidence-sets-using-support-vector-machines) Wenbo Wang (/author/wenbo-wang-11405), Xingye Qiao (/author/xingye-qiao-11406)
- Efficient Neural Network Robustness Certification with General Activation Functions (/paper/7742-efficient-neural-network-robustness-certification-with-general-activation-functions) Huan Zhang (/author/huan-zhang-9151), Tsui-Wei Weng (/author/tsui-wei-weng-11407), Pin-Yu Chen (/author/pin-yu-chen-10619), Cho-Jui Hsieh (/author/cho-jui-hsieh-6366), Luca Daniel (/author/luca-daniel-11408)
- Hessian-based Analysis of Large Batch Training and Robustness to Adversaries (/paper/7743-hessian-based-analysis-of-large-batch-training-and-robustness-to-adversaries) Zhewei Yao (/author/zhewei-yao-11409), Amir Gholami (/author/amir-gholami-11410), Qi Lei (/author/qi-lei-8545), Kurt Keutzer (/author/kurt-keutzer-11411), Michael W. Mahoney (/author/michael-w-mahoney-4286)
- Neural Edit Operations for Biological Sequences (/paper/7744-neural-edit-operations-for-biological-sequences) Satoshi Koide (/author/satoshi-koide-11412), Keisuke Kawano (/author/keisuke-kawano-11413), Takuro Kutsuna (/author/takuro-kutsuna-11414)
- Objective and efficient inference for couplings in neuronal networks (/paper/7745-objective-and-efficient-inference-for-couplings-in-neuronal-networks) Yu Terada (/author/yu-terada-11415), Tomoyuki Obuchi (/author/tomoyuki-obuchi-11313), Takuya Isomura (/author/takuya-isomura-11416), Yoshiyuki Kabashima (/author/yoshiyuki-kabashima-1795)
- Learning from Group Comparisons: Exploiting Higher Order Interactions (/paper/7746-learning-from-group-comparisons-exploiting-higher-order-interactions) Yao Li (/author/yao-li-7191), Minhao Cheng (/author/minhao-cheng-11417), Kevin Fujii (/author/kevin-fujii-11418), Fushing Hsieh (/author/fushing-hsieh-11419), Cho-Jui Hsieh (/author/cho-jui-hsieh-6366)
- Supervising Unsupervised Learning (/paper/7747-supervising-unsupervised-learning) Vikas Garg (/author/vikas-garg-6645)
- Nonparametric Bayesian Lomax delegate racing for survival analysis with competing risks (/paper/7748-nonparametric-bayesian-lomax-delegate-racing-for-survival-analysis-with-competing-risks) Quan Zhang (/author/quan-zhang-11420), Mingyuan Zhou (/author/mingyuan-zhou-4435)
- Adversarially Robust Generalization Requires More Data (/paper/7749-adversarially-robust-generalization-requires-more-data) Ludwig Schmidt (/author/ludwig-schmidt-8011), Shibani Santurkar (/author/shibani-santurkar-10962), Dimitris Tsipras (/author/dimitris-tsipras-10963), Kunal Talwar (/author/kunal-talwar-8026), Aleksander Madry (/author/aleksander-madry-10965)
- Improving Exploration in Evolution Strategies for Deep Reinforcement Learning via a Population of Novelty-Seeking Agents (/paper/7750-improving-exploration-in-evolution-strategies-for-deep-reinforcement-learning-via-a-population-of-novelty-seeking-agents)

Edoardo Conti (/author/edoardo-conti-11421), Vashisht Madhavan (/author/vashisht-madhavan-11422), Felipe Petroski Such (/author/felipe-petroski-such-11423), Joel Lehman (/author/joel-lehman-11424), Kenneth Stanley (/author/kenneth-stanley-11425), Jeff Clune (/author/jeff-clune-7289)

- Practical exact algorithm for trembling-hand equilibrium refinements in games (/paper/7751-practical-exact-algorithm-for-trembling-hand-equilibrium-refinements-in-games) Gabriele Farina (/author/gabriele-farina-10683), Nicola Gatti (/author/nicola-gatti-11426), Tuomas Sandholm (/author/tuomas-sandholm-2336)
- LAG: Lazily Aggregated Gradient for Communication-Efficient Distributed Learning (/paper/7752-lag-lazily-aggregated-gradient-for-communication-efficient-distributed-learning) Tianyi Chen (/author/tianyi-chen-11427), Georgios Giannakis (/author/georgios-giannakis-8486), Tao Sun (/author/tao-sun-10312), Wotao Yin (/author/wotao-yin-10215)
- Scalable Robust Matrix Factorization with Nonconvex Loss (/paper/7753-scalable-robust-matrix-factorization-with-nonconvex-loss) Quanming Yao (/author/quanming-yao-11428), James Kwok (/author/james-kwok-7945)
- Power-law efficient neural codes provide general link between perceptual bias and discriminability (/paper/7754-power-law-efficient-neural-codes-provide-general-link-between-perceptual-bias-and-discriminability) Michael Morais (/author/michael-morais-11429), Jonathan W. Pillow (/author/jonathan-w-pillow-5713)
- Geometry-Aware Recurrent Neural Networks for Active Visual Recognition (/paper/7755-geometry-aware-recurrent-neural-networks-for-active-visual-recognition) Ricson Cheng (/author/ricson-cheng-11430), Ziyan Wang (/author/ziyan-wang-11431), Katerina Fragkiadaki (/author/katerina-fragkiadaki-7383)
- Unsupervised Adversarial Invariance (/paper/7756-unsupervised-adversarial-invariance) Ayush Jaiswal (/author/ayush-jaiswal-11432), Rex Yue Wu (/author/rex-yue-wu-11433), Wael Abd-Almageed (/author/wael-abd-almageed-11434), Prem Natarajan (/author/prem-natarajan-11435)
- Content preserving text generation with attribute controls (/paper/7757-content-preserving-text-generation-with-attribute-controls) Lajanugen Logeswaran (/author/lajanugen-logeswaran-11436), Honglak Lee (/author/honglak-lee-3292), Samy Bengio (/author/samy-bengio-350)
- Multi-armed Bandits with Compensation (/paper/7758-multi-armed-bandits-with-compensation) Siwei Wang (/author/siwei-wang-11437), Longbo Huang (/author/longbo-huang-11438)
- GradiVeQ: Vector Quantization for Bandwidth-Efficient Gradient Aggregation in Distributed CNN Training (/paper/7759-gradiveq-vector-quantization-for-bandwidth-efficient-gradient-aggregation-in-distributed-cnn-training) Mingchao Yu (/author/mingchao-yu-11439), Zhifeng Lin (/author/zhifeng-lin-11440), Krishna Narra (/author/krishna-narra-11441), Songze Li (/author/songze-li-11442), Youjie Li (/author/youjie-li-11443), Nam Sung Kim (/author/nam-sung-kim-11444), Alexander Schwing (/author/alexander-schwing-9411),

Murali Annavaram (/author/murali-annavaram-11445), Salman Avestimehr (/author/salman-avestimehr-10066)

- Learning in Games with Lossy Feedback (/paper/7760-learning-in-games-with-lossy-feedback) Zhengyuan Zhou (/author/zhengyuan-zhou-6557), Panayotis Mertikopoulos (/author/panayotis-mertikopoulos-10310), Susan Athey (/author/susan-athey-10453), Nicholas Bambos (/author/nicholas-bambos-10311), Peter W. Glynn (/author/peter-w-glynn-2046), Yinyu Ye (/author/yinyu-ye-11446)
- Scalable methods for 8-bit training of neural networks (/paper/7761-scalable-methods-for-8-bit-training-of-neural-networks) Ron Banner (/author/ron-banner-10917), Itay Hubara (/author/itay-hubara-7178), Elad Hoffer (/author/elad-hoffer-9639), Daniel Soudry (/author/daniel-soudry-6992)
- Dropping Symmetry for Fast Symmetric Nonnegative Matrix Factorization (/paper/7762-dropping-symmetry-for-fast-symmetric-nonnegative-matrix-factorization) Zhihui Zhu (/author/zhihui-zhu-10919), Xiao Li (/author/xiao-li-7985), Kai Liu (/author/kai-liu-11447), Qiuwei Li (/author/qiuwei-li-11448)
- Link Prediction Based on Graph Neural Networks (/paper/7763-link-prediction-based-on-graph-neural-networks) Muhan Zhang (/author/muhan-zhang-11449), Yixin Chen (/author/yixin-chen-3208)
- Why so gloomy? A Bayesian explanation of human pessimism bias in the multi-armed bandit task (/paper/7764-why-so-gloomy-a-bayesian-explanation-of-human-pessimism-bias-in-the-multi-armed-bandit-task) Dalin Guo (/author/dalin-guo-11450), Angela J. Yu (/author/angela-j-yu-2266)
- Near-Optimal Time and Sample Complexities for Solving Markov Decision Processes with a Generative Model (/paper/7765-near-optimal-time-and-sample-complexities-for-solving-markov-decision-processes-with-a-generative-model) Aaron Sidford (/author/aaron-sidford-11451), Mengdi Wang (/author/mengdi-wang-9065), Xian Wu (/author/xian-wu-11452), Lin Yang (/author/lin-yang-9084), Yinyu Ye (/author/yinyu-ye-11446)
- ChannelNets: Compact and Efficient Convolutional Neural Networks via Channel-Wise Convolutions (/paper/7766-channelnets-compact-and-efficient-convolutional-neural-networks-via-channel-wise-convolutions) Hongyang Gao (/author/hongyang-gao-11453), Zhengyang Wang (/author/zhengyang-wang-11454), Shuiwang Ji (/author/shuiwang-ji-4078)
- Causal Inference and Mechanism Clustering of A Mixture of Additive Noise Models (/paper/7767-causal-inference-and-mechanism-clustering-of-a-mixture-of-additive-noise-models) Shoubo Hu (/author/shoubo-hu-11455), Zhitang Chen (/author/zhitang-chen-6823), Vahid Partovi Nia (/author/vahid-partovi-nia-11456), Laiwan CHAN (/author/laiwan-chan-11457), Yanhui Geng (/author/yanhui-geng-11458)
- Contour location via entropy reduction leveraging multiple information sources (/paper/7768-contour-location-via-entropy-reduction-leveraging-multiple-information-sources) Alexandre Marques (/author/alexandre-marques-11459), Remi Lam (/author/remi-lam-8665), Karen Willcox (/author/karen-willcox-8666)

- Assessing Generative Models via Precision and Recall (/paper/7769-assessing-generative-models-via-precision-and-recall) Mehdi S. M. Sajjadi (/author/mehdi-s-m-sajjadi-11460), Olivier Bachem (/author/olivier-bachem-9124), Mario Lucic (/author/mario-lucic-7416), Olivier Bousquet (/author/olivier-bousquet-2049), Sylvain Gelly (/author/sylvain-gelly-10211)
- Multiple-Step Greedy Policies in Approximate and Online Reinforcement Learning (/paper/7770-multiple-step-greedy-policies-in-approximate-and-online-reinforcement-learning) Yonathan Efroni (/author/yonathan-efroni-11461), Gal Dalal (/author/gal-dalal-11462), Bruno Scherrer (/author/bruno-scherrer-4011), Shie Mannor (/author/shie-mannor-2093)
- A Convex Duality Framework for GANs (/paper/7771-a-convex-duality-framework-for-gans) Farzan Farnia (/author/farzan-farnia-7987), David Tse (/author/david-tse-7988)
- Horizon-Independent Minimax Linear Regression (/paper/7772-horizon-independent-minimax-linear-regression) Alan Malek (/author/alan-malek-7143), Peter L. Bartlett (/author/peter-l-bartlett-6348)
- Exploiting Numerical Sparsity for Efficient Learning : Faster Eigenvector Computation and Regression (/paper/7773-exploiting-numerical-sparsity-for-efficient-learning-faster-eigenvector-computation-and-regression) Neha Gupta (/author/neha-gupta-11463), Aaron Sidford (/author/aaron-sidford-11451)
- Experimental Design for Cost-Aware Learning of Causal Graphs (/paper/7774-experimental-design-for-cost-aware-learning-of-causal-graphs) Erik Lindgren (/author/erik-lindgren-8973), Murat Kocaoglu (/author/murat-kocaoglu-7411), Alexandros G. Dimakis (/author/alexandros-g-dimakis-7413), Sriram Vishwanath (/author/sriram-vishwanath-8274)
- Task-Driven Convolutional Recurrent Models of the Visual System (/paper/7775-task-driven-convolutional-recurrent-models-of-the-visual-system) Aran Nayebi (/author/aran-nayebi-8986), Daniel Bear (/author/daniel-bear-11464), Jonas Kubilius (/author/jonas-kubilius-9759), Kohitij Kar (/author/kohitij-kar-11465), Surya Ganguli (/author/surya-ganguli-4596), David Sussillo (/author/david-sussillo-9342), James J. DiCarlo (/author/james-j-dicarlo-6474), Daniel L. Yamins (/author/daniel-l-yamins-6472)
- Meta-Reinforcement Learning of Structured Exploration Strategies (/paper/7776-meta-reinforcement-learning-of-structured-exploration-strategies) Abhishek Gupta (/author/abhishek-gupta-11466), Russell Mendonca (/author/russell-mendonca-11467), YuXuan Liu (/author/yuxuan-liu-11468), Pieter Abbeel (/author/pieter-abbeel-2686), Sergey Levine (/author/sergey-levine-4516)
- Sample Efficient Stochastic Gradient Iterative Hard Thresholding Method for Stochastic Sparse Linear Regression with Limited Attribute Observation (/paper/7777-sample-efficient-stochastic-gradient-iterative-hard-thresholding-method-for-stochastic-sparse-linear-regression-with-limited-attribute-observation) Tomoya Murata (/author/tomoya-murata-9461), Taiji Suzuki (/author/taiji-suzuki-4892)
- Semi-supervised Deep Kernel Learning: Regression with Unlabeled Data by Minimizing Predictive Variance (/paper/7778-semi-supervised-deep-kernel-learning-regression-with-

unlabeled-data-by-minimizing-predictive-variance) Neal Jean (/author/neal-jean-11469), Sang Michael Xie (/author/sang-michael-xie-11470), Stefano Ermon (/author/stefano-ermon-5196)

- Generalizing to Unseen Domains via Adversarial Data Augmentation (/paper/7779-generalizing-to-unseen-domains-via-adversarial-data-augmentation) Riccardo Volpi (/author/riccardo-volpi-11471), Hongseok Namkoong (/author/hongseok-namkoong-8456), Ozan Sener (/author/ozan-sener-8928), John C. Duchi (/author/john-c-duchi-3453), Vittorio Murino (/author/vittorio-murino-3325), Silvio Savarese (/author/silvio-savarese-8929)
- Hyperbolic Neural Networks (/paper/7780-hyperbolic-neural-networks) Octavian Ganea (/author/octavian-ganea-11472), Gary Becigneul (/author/gary-becigneul-11473), Thomas Hofmann (/author/thomas-hofmann-969)
- Breaking the Curse of Horizon: Infinite-Horizon Off-Policy Estimation (/paper/7781-breaking-the-curse-of-horizon-infinite-horizon-off-policy-estimation) Qiang Liu (/author/qiang-liu-5417), Lihong Li (/author/lihong-li-4101), Ziyang Tang (/author/ziyang-tang-11474), Dengyong Zhou (/author/dengyong-zhou-2649)
- Learning Task Specifications from Demonstrations (/paper/7782-learning-task-specifications-from-demonstrations) Marcell Vazquez-Chanlatte (/author/marcell-vazquez-chanlatte-11475), Susmit Jha (/author/susmit-jha-11476), Ashish Tiwari (/author/ashish-tiwari-11477), Mark K. Ho (/author/mark-k-ho-9018), Sanjit Seshia (/author/sanjit-seshia-11478)
- Learning a latent manifold of odor representations from neural responses in piriform cortex (/paper/7783-learning-a-latent-manifold-of-odor-representations-from-neural-responses-in-piriform-cortex) Anqi Wu (/author/anqi-wu-7132), Stan Pashkovski (/author/stan-pashkovski-11479), Sandeep R. Datta (/author/sandeep-r-datta-8962), Jonathan W. Pillow (/author/jonathan-w-pillow-5713)
- Fully Understanding The Hashing Trick (/paper/7784-fully-understanding-the-hashing-trick) Lior Kamma (/author/lior-kamma-11480), Casper B. Freksen (/author/casper-b-freksen-11481), Kasper Green Larsen (/author/kasper-green-larsen-11482)
- Evolved Policy Gradients (/paper/7785-evolved-policy-gradients) Rein Houthooft (/author/rein-houthooft-9000), Yuhua Chen (/author/yuhua-chen-11483), Phillip Isola (/author/phillip-isola-5198), Bradly Stadie (/author/bradly-stadie-9541), Filip Wolski (/author/filip-wolski-10157), OpenAI Jonathan Ho (/author/openai-jonathan-ho-9542), Pieter Abbeel (/author/pieter-abbeel-2686)
- The Spectrum of the Fisher Information Matrix of a Single-Hidden-Layer Neural Network (/paper/7786-the-spectrum-of-the-fisher-information-matrix-of-a-single-hidden-layer-neural-network) Jeffrey Pennington (/author/jeffrey-pennington-8301), Pratik Worah (/author/pratik-worah-9770)
- Learning Concave Conditional Likelihood Models for Improved Analysis of Tandem Mass Spectra (/paper/7787-learning-concave-conditional-likelihood-models-for-improved-analysis-of-tandem-mass-spectra) John T. Halloran (/author/john-t-halloran-10252), David M. Rocke (/author/david-m-rocke-10253)

- Differentially Private k-Means with Constant Multiplicative Error (/paper/7788-differentially-private-k-means-with-constant-multiplicative-error) Uri Stemmer (/author/uri-stemmer-9720), Haim Kaplan (/author/haim-kaplan-11484)
- Policy Optimization via Importance Sampling (/paper/7789-policy-optimization-via-importance-sampling) Alberto Maria Metelli (/author/alberto-maria-metelli-9689), Matteo Papini (/author/matteo-papini-9922), Francesco Faccio (/author/francesco-faccio-11485), Marcello Restelli (/author/marcello-restelli-3736)
- Estimating Learnability in the Sublinear Data Regime (/paper/7790-estimating-learnability-in-the-sublinear-data-regime) Weihao Kong (/author/weihao-kong-5677), Gregory Valiant (/author/gregory-valiant-6736)
- Algorithmic Assurance: An Active Approach to Algorithmic Testing using Bayesian Optimisation (/paper/7791-algorithmic-assurance-an-active-approach-to-algorithmic-testing-using-bayesian-optimisation) Shivapratap Gopakumar (/author/shivapratap-gopakumar-11486), Sunil Gupta (/author/sunil-gupta-9893), Santu Rana (/author/santu-rana-9892), Vu Nguyen (/author/vu-nguyen-9266), Svetha Venkatesh (/author/svetha-venkatesh-3855)
- Community Exploration: From Offline Optimization to Online Learning (/paper/7792-community-exploration-from-offline-optimization-to-online-learning) Xiaowei Chen (/author/xiaowei-chen-11487), Weiran Huang (/author/weiran-huang-11488), Wei Chen (/author/wei-chen-4262), John C. S. Lui (/author/john-c-s-lui-11489)
- A Dual Framework for Low-rank Tensor Completion (/paper/7793-a-dual-framework-for-low-rank-tensor-completion) Madhav Nimishakavi (/author/madhav-nimishakavi-11490), Pratik Kumar Jawanpuria (/author/pratik-kumar-jawanpuria-8260), Bamdev Mishra (/author/bamdev-mishra-11298)
- Low-rank Interaction with Sparse Additive Effects Model for Large Data Frames (/paper/7794-low-rank-interaction-with-sparse-additive-effects-model-for-large-data-frames) Geneviève Robin (/author/genevieve-robin-11491), Hoi-To Wai (/author/hoi-to-wai-11492), Julie Josse (/author/julie-josse-11493), Olga Klopp (/author/olga-klopp-7306), Eric Moulines (/author/eric-moulines-4051)
- Inference Aided Reinforcement Learning for Incentive Mechanism Design in Crowdsourcing (/paper/7795-inference-aided-reinforcement-learning-for-incentive-mechanism-design-in-crowdsourcing) Zehong Hu (/author/zehong-hu-11494), Yitao Liang (/author/yitao-liang-11495), Jie Zhang (/author/jie-zhang-11496), Zhao Li (/author/zhao-li-11497), Yang Liu (/author/yang-liu-8670)
- Middle-Out Decoding (/paper/7796-middle-out-decoding) Shikib Mehri (/author/shikib-mehri-11498), Leonid Sigal (/author/leonid-sigal-2726)
- First-order Stochastic Algorithms for Escaping From Saddle Points in Almost Linear Time (/paper/7797-first-order-stochastic-algorithms-for-escaping-from-saddle-points-in-almost-linear-time) Yi Xu (/author/yi-xu-9010), Jing Rong (/author/jing-rong-10478), Tianbao Yang (/author/tianbao-yang-5489)

- To Trust Or Not To Trust A Classifier (/paper/7798-to-trust-or-not-to-trust-a-classifier) Heinrich Jiang (/author/heinrich-jiang-9364), Been Kim (/author/been-kim-7252), Melody Guan (/author/melody-guan-11499), Maya Gupta (/author/maya-gupta-4237)
- Reparameterization Gradient for Non-differentiable Models (/paper/7799-reparameterization-gradient-for-non-differentiable-models) Wonyeol Lee (/author/wonyeol-lee-11500), Hangyeol Yu (/author/hangyeol-yu-11501), Hongseok Yang (/author/hongseok-yang-11502)
- A Simple Proximal Stochastic Gradient Method for Nonsmooth Nonconvex Optimization (/paper/7800-a-simple-proximal-stochastic-gradient-method-for-nonsmooth-nonconvex-optimization) Zhize Li (/author/zhize-li-8431), Jian Li (/author/jian-li-3569)
- Multimodal Generative Models for Scalable Weakly-Supervised Learning (/paper/7801-multimodal-generative-models-for-scalable-weakly-supervised-learning) Mike Wu (/author/mike-wu-11503), Noah Goodman (/author/noah-goodman-3514)
- How Much Restricted Isometry is Needed In Nonconvex Matrix Recovery? (/paper/7802-how-much-restricted-isometry-is-needed-in-nonconvex-matrix-recovery) Richard Zhang (/author/richard-zhang-9439), Cedric Josz (/author/cedric-josz-10957), Somayeh Sojoudi (/author/somayeh-sojoudi-10959), Javad Lavaei (/author/javad-lavaei-10958)
- Occam's razor is insufficient to infer the preferences of irrational agents (/paper/7803-occams-razor-is-insufficient-to-infer-the-preferences-of-irrational-agents) Stuart Armstrong (/author/stuart-armstrong-11504), Sören Mindermann (/author/soren-mindermann-11505)
- Manifold Structured Prediction (/paper/7804-manifold-structured-prediction) Alessandro Rudi (/author/alessandro-rudi-6400), Carlo Ciliberto (/author/carlo-ciliberto-8528), GianMaria Marconi (/author/gianmaria-marconi-11506), Lorenzo Rosasco (/author/lorenzo-rosasco-3011)
- Fast Greedy MAP Inference for Determinantal Point Process to Improve Recommendation Diversity (/paper/7805-fast-greedy-map-inference-for-determinantal-point-process-to-improve-recommendation-diversity) Laming Chen (/author/laming-chen-11507), Guoxin Zhang (/author/guoxin-zhang-11508), Eric Zhou (/author/eric-zhou-11509)
- Learning Others' Intentional Models in Multi-Agent Settings Using Interactive POMDPs (/paper/7806-learning-others-intentional-models-in-multi-agent-settings-using-interactive-pomdps) Yanlin Han (/author/yanlin-han-11510), Piotr Gmytrasiewicz (/author/piotr-gmytrasiewicz-11511)
- Contextual Pricing for Lipschitz Buyers (/paper/7807-contextual-pricing-for-lipschitz-buyers) Jieming Mao (/author/jieming-mao-11512), Renato Leme (/author/renato-leme-9780), Jon Schneider (/author/jon-schneider-11513)
- Online Improper Learning with an Approximation Oracle (/paper/7808-online-improper-learning-with-an-approximation-oracle) Elad Hazan (/author/elad-hazan-3630), Wei Hu (/author/wei-hu-9180), Yuanzhi Li (/author/yuanzhi-li-8643), zhiyuan li (/author/zhiyuan-li-11514)
- Bandit Learning in Concave N-Person Games (/paper/7809-bandit-learning-in-concave-n-person-games) Mario Bravo (/author/mario-bravo-11515), David Leslie (/author/david-

leslie-11516), Panayotis Mertikopoulos (/author/panayotis-mertikopoulos-10310)

- On Fast Leverage Score Sampling and Optimal Learning (/paper/7810-on-fast-leverage-score-sampling-and-optimal-learning) Alessandro Rudi (/author/alessandro-rudi-6400), Daniele Calandriello (/author/daniele-calandriello-7146), Luigi Carratino (/author/luigi-carratino-9975), Lorenzo Rosasco (/author/lorenzo-rosasco-3011)
- Unsupervised Video Object Segmentation for Deep Reinforcement Learning (/paper/7811-unsupervised-video-object-segmentation-for-deep-reinforcement-learning) Vikash Goel (/author/vikash-goel-11517), Jameson Weng (/author/jameson-weng-11518), Pascal Poupart (/author/pascal-poupart-2378)
- Efficient inference for time-varying behavior during learning (/paper/7812-efficient-inference-for-time-varying-behavior-during-learning) Nicholas G. Roy (/author/nicholas-g-roy-7325), Ji Hyun Bak (/author/ji-hyun-bak-8902), Athena Akrami (/author/athena-akrami-9287), Carlos Brody (/author/carlos-brody-7765), Jonathan W. Pillow (/author/jonathan-w-pillow-5713)
- Learning convex polytopes with margin (/paper/7813-learning-convex-polytopes-with-margin) Lee-Ad Gottlieb (/author/lee-ad-gottlieb-7543), Eran Kaufman (/author/eran-kaufman-11519), Aryeh Kontorovich (/author/aryeh-kontorovich-6644), Gabriel Nivasch (/author/gabriel-nivasch-11520)
- Critical initialisation for deep signal propagation in noisy rectifier neural networks (/paper/7814-critical-initialisation-for-deep-signal-propagation-in-noisy-rectifier-neural-networks) Arnu Pretorius (/author/arnu-pretorius-11521), Elan van Biljon (/author/elan-van-biljon-11522), Steve Kroon (/author/steve-kroon-11523), Herman Kamper (/author/herman-kamper-11524)
- Insights on representational similarity in neural networks with canonical correlation (/paper/7815-insights-on-representational-similarity-in-neural-networks-with-canonical-correlation) Ari Morcos (/author/ari-morcos-11525), Maithra Raghu (/author/maithra-raghu-8856), Samy Bengio (/author/samy-bengio-350)
- Variational Inference with Tail-adaptive f-Divergence (/paper/7816-variational-inference-with-tail-adaptive-f-divergence) Dilin Wang (/author/dilin-wang-8892), Hao Liu (/author/hao-liu-10228), Qiang Liu (/author/qiang-liu-5417)
- Mental Sampling in Multimodal Representations (/paper/7817-mental-sampling-in-multimodal-representations) Jianqiao Zhu (/author/jianqiao-zhu-11526), Adam Sanborn (/author/adam-sanborn-3588), Nick Chater (/author/nick-chater-4433)
- Adversarially Robust Optimization with Gaussian Processes (/paper/7818-adversarially-robust-optimization-with-gaussian-processes) Ilija Bogunovic (/author/ilija-bogunovic-8510), Jonathan Scarlett (/author/jonathan-scarlett-8511), Stefanie Jegelka (/author/stefanie-jegelka-3773), Volkan Cevher (/author/volkan-cevher-3964)
- Learning to Multitask (/paper/7819-learning-to-multitask) Yu Zhang (/author/yu-zhang-4618), Ying Wei (/author/ying-wei-11527), Qiang Yang (/author/qiang-yang-3973)
- Loss Functions for Multiset Prediction (/paper/7820-loss-functions-for-multiset-prediction) Sean Welleck (/author/sean-welleck-10177), Zixin Yao (/author/zixin-yao-11528), Yu Gai

(/author/yu-gai-11529), Jialin Mao (/author/jialin-mao-10452), Zheng Zhang (/author/zheng-zhang-7177), Kyunghyun Cho (/author/kyunghyun-cho-7192)

- Computing Kantorovich-Wasserstein Distances on *d*-dimensional histograms using *(d+1)*-partite graphs (/paper/7821-computing-kantorovich-wasserstein-distances-on-d-dimensional-histograms-using-d1-partite-graphs) Gennaro Auricchio (/author/gennaro-auricchio-11530), Federico Bassetti (/author/federico-bassetti-11531), Stefano Gualandi (/author/stefano-gualandi-11532), Marco Veneroni (/author/marco-veneroni-11533)
- Neural Interaction Transparency (NIT): Disentangling Learned Interactions for Improved Interpretability (/paper/7822-neural-interaction-transparency-nit-disentangling-learned-interactions-for-improved-interpretability) Michael Tsang (/author/michael-tsang-11534), Hanpeng Liu (/author/hanpeng-liu-11535), Sanjay Purushotham (/author/sanjay-purushotham-11536), Pavankumar Murali (/author/pavankumar-murali-11537), Yan Liu (/author/yan-liu-5231)
- CapProNet: Deep Feature Learning via Orthogonal Projections onto Capsule Subspaces (/paper/7823-cappronet-deep-feature-learning-via-orthogonal-projections-onto-capsule-subspaces) Liheng Zhang (/author/liheng-zhang-11538), Marzieh Edraki (/author/marzieh-edraki-11539), Guo-Jun Qi (/author/guo-jun-qi-11540)
- Gamma-Poisson Dynamic Matrix Factorization Embedded with Metadata Influence (/paper/7824-gamma-poisson-dynamic-matrix-factorization-embedded-with-metadata-influence) Trong Dinh Thac Do (/author/trong-dinh-thac-do-11541), Longbing Cao (/author/longbing-cao-11542)
- Masking: A New Perspective of Noisy Supervision (/paper/7825-masking-a-new-perspective-of-noisy-supervision) Bo Han (/author/bo-han-11543), Jiangchao Yao (/author/jiangchao-yao-11544), Gang Niu (/author/gang-niu-4971), Mingyuan Zhou (/author/mingyuan-zhou-4435), Ivor Tsang (/author/ivor-tsang-8394), Ya Zhang (/author/ya-zhang-3209), Masashi Sugiyama (/author/masashi-sugiyama-1885)
- On GANs and GMMs (/paper/7826-on-gans-and-gmms) Eitan Richardson (/author/eitan-richardson-11545), Yair Weiss (/author/yair-weiss-1523)
- Differential Properties of Sinkhorn Approximation for Learning with Wasserstein Distance (/paper/7827-differential-properties-of-sinkhorn-approximation-for-learning-with-wasserstein-distance) Giulia Luise (/author/giulia-luise-11546), Alessandro Rudi (/author/alessandro-rudi-6400), Massimiliano Pontil (/author/massimiliano-pontil-2047), Carlo Ciliberto (/author/carlo-ciliberto-8528)
- Practical Deep Stereo (PDS): Toward applications-friendly deep stereo matching (/paper/7828-practical-deep-stereo-pds-toward-applications-friendly-deep-stereo-matching) Stepan Tulyakov (/author/stepan-tulyakov-11547), Anton Ivanov (/author/anton-ivanov-11548), François Fleuret (/author/francois-fleuret-6804)
- A Bayes-Sard Cubature Method (/paper/7829-a-bayes-sard-cubature-method) Toni Karvonen (/author/toni-karvonen-11549), Chris J. Oates (/author/chris-j-oates-11550), Simo Sarkka (/author/simo-sarkka-11551)

- Dual Swap Disentangling (/paper/7830-dual-swap-disentangling) Zunlei Feng (/author/zunlei-feng-11552), Xinchao Wang (/author/xinchao-wang-11553), Chenglong Ke (/author/chenglong-ke-11554), An-Xiang Zeng (/author/an-xiang-zeng-11555), Dacheng Tao (/author/dacheng-tao-4229), Mingli Song (/author/mingli-song-11556)
- Diverse Ensemble Evolution: Curriculum Data-Model Marriage (/paper/7831-diverse-ensemble-evolution-curriculum-data-model-marriage) Tianyi Zhou (/author/tianyi-zhou-7616), Shengjie Wang (/author/shengjie-wang-7998), Jeff A. Bilmes (/author/jeff-a-bilmes-2747)
- Binary Classification from Positive-Confidence Data (/paper/7832-binary-classification-from-positive-confidence-data) Takashi Ishida (/author/takashi-ishida-10241), Gang Niu (/author/gang-niu-4971), Masashi Sugiyama (/author/masashi-sugiyama-1885)
- Deep Generative Models for Distribution-Preserving Lossy Compression (/paper/7833-deep-generative-models-for-distribution-preserving-lossy-compression) Michael Tschannen (/author/michael-tschannen-9486), Eirikur Agustsson (/author/eirikur-agustsson-9548), Mario Lucic (/author/mario-lucic-7416)
- Exact natural gradient in deep linear networks and its application to the nonlinear case (/paper/7834-exact-natural-gradient-in-deep-linear-networks-and-its-application-to-the-nonlinear-case) Alberto Bernacchia (/author/alberto-bernacchia-11557), Mate Lengyel (/author/mate-lengyel-6286), Guillaume Hennequin (/author/guillaume-hennequin-4300)
- Constructing Fast Network through Deconstruction of Convolution (/paper/7835-constructing-fast-network-through-deconstruction-of-convolution) Yunho Jeon (/author/yunho-jeon-11558), Junmo Kim (/author/junmo-kim-11559)
- Memory Replay GANs: Learning to Generate New Categories without Forgetting (/paper/7836-memory-replay-gans-learning-to-generate-new-categories-without-forgetting) Chenshen Wu (/author/chenshen-wu-11560), Luis Herranz (/author/luis-herranz-11561), Xialei Liu (/author/xialei-liu-11562), yaxing wang (/author/yaxing-wang-11563), Joost van de Weijer (/author/joost-van-de-weijer-10769), Bogdan Raducanu (/author/bogdan-raducanu-11564)
- The Convergence of Sparsified Gradient Methods (/paper/7837-the-convergence-of-sparsified-gradient-methods) Dan Alistarh (/author/dan-alistarh-8258), Torsten Hoefler (/author/torsten-hoefler-11166), Mikael Johansson (/author/mikael-johansson-11017), Nikola Konstantinov (/author/nikola-konstantinov-11565), Sarit Khirirat (/author/sarit-khirirat-11566), Cedric Renggli (/author/cedric-renggli-11567)
- Automating Bayesian optimization with Bayesian optimization (/paper/7838-automating-bayesian-optimization-with-bayesian-optimization) Gustavo Malkomes (/author/gustavo-malkomes-8210), Roman Garnett (/author/roman-garnett-7864)
- Stacked Semantics-Guided Attention Model for Fine-Grained Zero-Shot Learning (/paper/7839-stacked-semantics-guided-attention-model-for-fine-grained-zero-shot-learning) yunlong yu (/author/yunlong-yu-11568), Zhong Ji (/author/zhong-ji-11569), Yanwei Fu (/author/yanwei-fu-11570), Jichang Guo (/author/jichang-guo-11571), Yanwei

Pang (/author/yanwei-pang-11572), Zhongfei (Mark) Zhang (/author/zhongfei-mark-zhang-8900)

- Dirichlet-based Gaussian Processes for Large-scale Calibrated Classification (/paper/7840-dirichlet-based-gaussian-processes-for-large-scale-calibrated-classification) Dimitrios Milios (/author/dimitrios-milios-11573), Raffaello Camoriano (/author/raffaello-camoriano-8296), Pietro Michiardi (/author/pietro-michiardi-11574), Lorenzo Rosasco (/author/lorenzo-rosasco-3011), Maurizio Filippone (/author/maurizio-filippone-8222)
- Multi-Task Zipping via Layer-wise Neuron Sharing (/paper/7841-multi-task-zipping-via-layer-wise-neuron-sharing) Xiaoxi He (/author/xiaoxi-he-11575), Zimu Zhou (/author/zimu-zhou-11576), Lothar Thiele (/author/lothar-thiele-11577)
- Dimensionally Tight Bounds for Second-Order Hamiltonian Monte Carlo (/paper/7842-dimensionally-tight-bounds-for-second-order-hamiltonian-monte-carlo) Oren Mangoubi (/author/oren-mangoubi-11578), Nisheeth Vishnoi (/author/nisheeth-vishnoi-11579)
- Approximation algorithms for stochastic clustering (/paper/7843-approximation-algorithms-for-stochastic-clustering) David Harris (/author/david-harris-11580), Shi Li (/author/shi-li-11581), Aravind Srinivasan (/author/aravind-srinivasan-11582), Khoa Trinh (/author/khoa-trinh-11583), Thomas Pensyl (/author/thomas-pensyl-11584)
- Evolutionary Stochastic Gradient Descent for Optimization of Deep Neural Networks (/paper/7844-evolutionary-stochastic-gradient-descent-for-optimization-of-deep-neural-networks) Xiaodong Cui (/author/xiaodong-cui-9378), Wei Zhang (/author/wei-zhang-7010), Zoltán Tüske (/author/zoltan-tuske-11585), Michael Picheny (/author/michael-picheny-11586)
- Learning to Infer Graphics Programs from Hand-Drawn Images (/paper/7845-learning-to-infer-graphics-programs-from-hand-drawn-images) Kevin Ellis (/author/kevin-ellis-8105), Daniel Ritchie (/author/daniel-ritchie-8917), Armando Solar-Lezama (/author/armando-solar-lezama-8106), Josh Tenenbaum (/author/josh-tenenbaum-6308)
- Graphical Generative Adversarial Networks (/paper/7846-graphical-generative-adversarial-networks) Chongxuan LI (/author/chongxuan-li-10008), Max Welling (/author/max-welling-2187), Jun Zhu (/author/jun-zhu-6461), Bo Zhang (/author/bo-zhang-3841)
- Variational Learning on Aggregate Outputs with Gaussian Processes (/paper/7847-variational-learning-on-aggregate-outputs-with-gaussian-processes) Ho Chung Law (/author/ho-chung-law-9585), Dino Sejdinovic (/author/dino-sejdinovic-5539), Ewan Cameron (/author/ewan-cameron-11587), Tim Lucas (/author/tim-lucas-11588), Seth Flaxman (/author/seth-flaxman-11589), Katherine Battle (/author/katherine-battle-11590), Kenji Fukumizu (/author/kenji-fukumizu-1361)
- MacNet: Transferring Knowledge from Machine Comprehension to Sequence-to-Sequence Models (/paper/7848-macnet-transferring-knowledge-from-machine-comprehension-to-sequence-to-sequence-models) Boyuan Pan (/author/boyuan-pan-11591), Yazheng Yang (/author/yazheng-yang-11592), Hao Li (/author/hao-li-10262), Zhou Zhao (/author/zhou-zhao-11593), Yueting Zhuang (/author/yueting-zhuang-11594), Deng Cai (/author/deng-cai-3206), Xiaofei He (/author/xiaofei-he-2553)

- Poison Frogs! Targeted Clean-Label Poisoning Attacks on Neural Networks (/paper/7849-poison-frogs-targeted-clean-label-poisoning-attacks-on-neural-networks) Ali Shafahi (/author/ali-shafahi-11595), W. Ronny Huang (/author/w-ronny-huang-11596), Mahyar Najibi (/author/mahyar-najibi-11597), Octavian Suciu (/author/octavian-suciu-11598), Christoph Studer (/author/christoph-studer-10265), Tudor Dumitras (/author/tudor-dumitras-11599), Tom Goldstein (/author/tom-goldstein-8021)
- Information Constraints on Auto-Encoding Variational Bayes (/paper/7850-information-constraints-on-auto-encoding-variational-bayes) Romain Lopez (/author/romain-lopez-11600), Jeffrey Regier (/author/jeffrey-regier-9737), Michael I. Jordan (/author/michael-i-jordan-6338), Nir Yosef (/author/nir-yosef-11601)
- Recurrent Transformer Networks for Semantic Correspondence (/paper/7851-recurrent-transformer-networks-for-semantic-correspondence) Seungryong Kim (/author/seungryong-kim-11602), Stephen Lin (/author/stephen-lin-11603), SANG RYUL JEON (/author/sang-ryul-jeon-11604), Dongbo Min (/author/dongbo-min-11605), Kwanghoon Sohn (/author/kwanghoon-sohn-11606)
- Online convex optimization for cumulative constraints (/paper/7852-online-convex-optimization-for-cumulative-constraints) Jianjun Yuan (/author/jianjun-yuan-11607), Andrew Lamperski (/author/andrew-lamperski-11608)
- Predict Responsibly: Improving Fairness and Accuracy by Learning to Defer (/paper/7853-predict-responsibly-improving-fairness-and-accuracy-by-learning-to-defer) David Madras (/author/david-madras-11609), Toni Pitassi (/author/toni-pitassi-6521), Richard Zemel (/author/richard-zemel-6522)
- Deep State Space Models for Unconditional Word Generation (/paper/7854-deep-state-space-models-for-unconditional-word-generation) Florian Schmidt (/author/florian-schmidt-11610), Thomas Hofmann (/author/thomas-hofmann-969)
- ResNet with one-neuron hidden layers is a Universal Approximator (/paper/7855-resnet-with-one-neuron-hidden-layers-is-a-universal-approximator) Hongzhou Lin (/author/hongzhou-lin-8294), Stefanie Jegelka (/author/stefanie-jegelka-3773)
- Transfer of Value Functions via Variational Methods (/paper/7856-transfer-of-value-functions-via-variational-methods) Andrea Tirinzoni (/author/andrea-tirinzoni-11611), Rafael Rodriguez Sanchez (/author/rafael-rodriguez-sanchez-11612), Marcello Restelli (/author/marcello-restelli-3736)
- The Cluster Description Problem - Complexity Results, Formulations and Approximations (/paper/7857-the-cluster-description-problem-complexity-results-formulations-and-approximations) Ian Davidson (/author/ian-davidson-11613), Antoine Gourru (/author/antoine-gourru-11614), S Ravi (/author/s-ravi-11615)
- Sharp Bounds for Generalized Uniformity Testing (/paper/7858-sharp-bounds-for-generalized-uniformity-testing) Ilias Diakonikolas (/author/ilias-diakonikolas-7127), Daniel M. Kane (/author/daniel-m-kane-11616), Alistair Stewart (/author/alistair-stewart-11617)
- Deep Neural Networks with Box Convolutions (/paper/7859-deep-neural-networks-with-box-convolutions) Egor Burkov (/author/egor-burkov-11618), Victor Lempitsky

(/author/victor-lempitsky-4675)

- Learning towards Minimum Hyperspherical Energy (/paper/7860-learning-towards-minimum-hyperspherical-energy) Weiyang Liu (/author/weiyang-liu-9985), Rongmei Lin (/author/rongmei-lin-11619), Zhen Liu (/author/zhen-liu-11620), Lixin Liu (/author/lixin-liu-11621), Zhiding Yu (/author/zhiding-yu-9987), Bo Dai (/author/bo-dai-6604), Le Song (/author/le-song-3161)
- LF-Net: Learning Local Features from Images (/paper/7861-lf-net-learning-local-features-from-images) Yuki Ono (/author/yuki-ono-11622), Eduard Trulls (/author/eduard-trulls-11623), Pascal Fua (/author/pascal-fua-6780), Kwang Moo Yi (/author/kwang-moo-yi-11624)
- SLANG: Fast Structured Covariance Approximations for Bayesian Deep Learning with Natural Gradient (/paper/7862-slang-fast-structured-covariance-approximations-for-bayesian-deep-learning-with-natural-gradient) Aaron Mishkin (/author/aaron-mishkin-11625), Frederik Kunstner (/author/frederik-kunstner-11626), Didrik Nielsen (/author/didrik-nielsen-11627), Mark Schmidt (/author/mark-schmidt-4137), Mohammad Emtiyaz Khan (/author/mohammad-emtiyaz-khan-11628)
- Tangent: Automatic differentiation using source-code transformation for dynamically typed array programming (/paper/7863-tangent-automatic-differentiation-using-source-code-transformation-for-dynamically-typed-array-programming) Bart van Merrienboer (/author/bart-van-merrienboer-11629), Dan Moldovan (/author/dan-moldovan-11630), Alexander Wiltschko (/author/alexander-wiltschko-11631)
- Multi-domain Causal Structure Learning in Linear Systems (/paper/7864-multi-domain-causal-structure-learning-in-linear-systems) AmirEmad Ghassami (/author/amiremad-ghassami-9829), Negar Kiyavash (/author/negar-kiyavash-9831), Biwei Huang (/author/biwei-huang-11632), Kun Zhang (/author/kun-zhang-6824)
- Privacy Amplification by Subsampling: Tight Analyses via Couplings and Divergences (/paper/7865-privacy-amplification-by-subsampling-tight-analyses-via-couplings-and-divergences) Borja Balle (/author/borja-balle-5500), Gilles Barthe (/author/gilles-barthe-11633), Marco Gaboardi (/author/marco-gaboardi-10707)
- Exponentially Weighted Imitation Learning for Batched Historical Data (/paper/7866-exponentially-weighted-imitation-learning-for-batched-historical-data) Qing Wang (/author/qing-wang-11634), Jiechao Xiong (/author/jiechao-xiong-8814), Lei Han (/author/lei-han-11635), peng sun (/author/peng-sun-11636), Han Liu (/author/han-liu-3555), Tong Zhang (/author/tong-zhang-1901)
- Algebraic tests of general Gaussian latent tree models (/paper/7867-algebraic-tests-of-general-gaussian-latent-tree-models) Dennis Leung (/author/dennis-leung-11637), Mathias Drton (/author/mathias-drton-4742)
- Navigating with Graph Representations for Fast and Scalable Decoding of Neural Language Models (/paper/7868-navigating-with-graph-representations-for-fast-and-scalable-decoding-of-neural-language-models) Minjia Zhang (/author/minjia-zhang-11638),

Wenhan Wang (/author/wenhan-wang-11639), Xiaodong Liu (/author/xiaodong-liu-11640), Jianfeng Gao (/author/jianfeng-gao-8345), Yuxiong He (/author/yuxiong-he-11641)

- Deep Structured Prediction with Nonlinear Output Transformations (/paper/7869-deep-structured-prediction-with-nonlinear-output-transformations) Colin Graber (/author/colin-graber-11642), Ofer Meshi (/author/ofer-meshi-4519), Alexander Schwing (/author/alexander-schwing-9411)
- Sequential Test for the Lowest Mean: From Thompson to Murphy Sampling (/paper/7870-sequential-test-for-the-lowest-mean-from-thompson-to-murphy-sampling) Emilie Kaufmann (/author/emilie-kaufmann-6655), Wouter M. Koolen (/author/wouter-m-koolen-4876), Aurélien Garivier (/author/aurelien-garivier-4848)
- Distributed Learning without Distress: Privacy-Preserving Empirical Risk Minimization (/paper/7871-distributed-learning-without-distress-privacy-preserving-empirical-risk-minimization) Bargav Jayaraman (/author/bargav-jayaraman-11643), Lingxiao Wang (/author/lingxiao-wang-11644), David Evans (/author/david-evans-11645), Quanquan Gu (/author/quanquan-gu-5348)
- A no-regret generalization of hierarchical softmax to extreme multi-label classification (/paper/7872-a-no-regret-generalization-of-hierarchical-softmax-to-extreme-multi-label-classification) Marek Wydmuch (/author/marek-wydmuch-11646), Kalina Jasinska (/author/kalina-jasinska-11647), Mikhail Kuznetsov (/author/mikhail-kuznetsov-11648), Róbert Busa-Fekete (/author/robert-busa-fekete-8268), Krzysztof Dembczynski (/author/krzysztof-dembczynski-8385)
- Efficient Formal Safety Analysis of Neural Networks (/paper/7873-efficient-formal-safety-analysis-of-neural-networks) Shiqi Wang (/author/shiqi-wang-11649), Kexin Pei (/author/kexin-pei-11650), Justin Whitehouse (/author/justin-whitehouse-11651), Junfeng Yang (/author/junfeng-yang-11652), Suman Jana (/author/suman-jana-11653)
- Bayesian Distributed Stochastic Gradient Descent (/paper/7874-bayesian-distributed-stochastic-gradient-descent) Michael Teng (/author/michael-teng-11654), Frank Wood (/author/frank-wood-3062)
- Visualizing the Loss Landscape of Neural Nets (/paper/7875-visualizing-the-loss-landscape-of-neural-nets) Hao Li (/author/hao-li-10262), Zheng Xu (/author/zheng-xu-10264), Gavin Taylor (/author/gavin-taylor-11655), Christoph Studer (/author/christoph-studer-10265), Tom Goldstein (/author/tom-goldstein-8021)
- The Limits of Post-Selection Generalization (/paper/7876-the-limits-of-post-selection-generalization) Jonathan Ullman (/author/jonathan-ullman-8645), Adam Smith (/author/adam-smith-6507), Kobbi Nissim (/author/kobbi-nissim-9718), Uri Stemmer (/author/uri-stemmer-9720), Thomas Steinke (/author/thomas-steinke-11656)
- Graph Convolutional Policy Network for Goal-Directed Molecular Graph Generation (/paper/7877-graph-convolutional-policy-network-for-goal-directed-molecular-graph-generation) Jiaxuan You (/author/jiaxuan-you-11387), Bowen Liu (/author/bowen-liu-11657), Zhitao Ying (/author/zhitao-ying-9531), Vijay Pande (/author/vijay-pande-11658), Jure Leskovec (/author/jure-leskovec-4767)

- On Controllable Sparse Alternatives to Softmax (/paper/7878-on-controllable-sparse-alternatives-to-softmax) Anirban Laha (/author/anirban-laha-11659), Saneem Ahmed Chemmengath (/author/saneem-ahmed-chemmengath-11660), Priyanka Agrawal (/author/priyanka-agrawal-11661), Mitesh Khapra (/author/mitesh-khapra-7180), Karthik Sankaranarayanan (/author/karthik-sankaranarayanan-11662), Harish G. Ramaswamy (/author/harish-g-ramaswamy-5280)
- L4: Practical loss-based stepsize adaptation for deep learning (/paper/7879-l4-practical-loss-based-stepsize-adaptation-for-deep-learning) Michal Rolinek (/author/michal-rolinek-11663), Georg Martius (/author/georg-martius-11664)
- Learning Latent Subspaces in Variational Autoencoders (/paper/7880-learning-latent-subspaces-in-variational-autoencoders) Jack Klys (/author/jack-klys-11665), Jake Snell (/author/jake-snell-10007), Richard Zemel (/author/richard-zemel-6522)
- Turbo Learning for CaptionBot and DrawingBot (/paper/7881-turbo-learning-for-captionbot-and-drawingbot) Qiuyuan Huang (/author/qiuyuan-huang-11666), Pengchuan Zhang (/author/pengchuan-zhang-11667), Dapeng Wu (/author/dapeng-wu-11668), Lei Zhang (/author/lei-zhang-3167)
- Learning to Teach with Dynamic Loss Functions (/paper/7882-learning-to-teach-with-dynamic-loss-functions) Lijun Wu (/author/lijun-wu-9656), Fei Tian (/author/fei-tian-9658), Yingce Xia (/author/yingce-xia-9116), Yang Fan (/author/yang-fan-11669), Tao Qin (/author/tao-qin-3844), Lai Jian-Huang (/author/lai-jian-huang-11670), Tie-Yan Liu (/author/tie-yan-liu-8972)
- Multi-View Silhouette and Depth Decomposition for High Resolution 3D Object Representation (/paper/7883-multi-view-silhouette-and-depth-decomposition-for-high-resolution-3d-object-representation) Edward Smith (/author/edward-smith-11671), Scott Fujimoto (/author/scott-fujimoto-11672), David Meger (/author/david-meger-11673)
- Size-Noise Tradeoffs in Generative Networks (/paper/7884-size-noise-tradeoffs-in-generative-networks) Bolton Bailey (/author/bolton-bailey-11674), Matus J. Telgarsky (/author/matus-j-telgarsky-6642)
- Online Adaptive Methods, Universality and Acceleration (/paper/7885-online-adaptive-methods-universality-and-acceleration) Yehuda Kfir Levy (/author/yehuda-kfir-levy-11675), Alp Yurtsever (/author/alp-yurtsever-8154), Volkan Cevher (/author/volkan-cevher-3964)
- Compact Generalized Non-local Network (/paper/7886-compact-generalized-non-local-network) Kaiyu Yue (/author/kaiyu-yue-11676), Ming Sun (/author/ming-sun-11677), Yuchen Yuan (/author/yuchen-yuan-11678), Feng Zhou (/author/feng-zhou-4291), Errui Ding (/author/errui-ding-11679), Fuxin Xu (/author/fuxin-xu-11680)
- On the Local Hessian in Back-propagation (/paper/7887-on-the-local-hessian-in-back-propagation) Huishuai Zhang (/author/huishuai-zhang-8416), Wei Chen (/author/wei-chen-4262), Tie-Yan Liu (/author/tie-yan-liu-8972)
- The Everlasting Database: Statistical Validity at a Fair Price (/paper/7888-the-everlasting-database-statistical-validity-at-a-fair-price) Blake E. Woodworth (/author/blake-e-

woodworth-8482), Vitaly Feldman (/author/vitaly-feldman-6643), Saharon Rosset (/author/saharon-rosset-2490), Nati Srebro (/author/nati-srebro-5014)

- Lipschitz-Margin Training: Scalable Certification of Perturbation Invariance for Deep Neural Networks (/paper/7889-lipschitz-margin-training-scalable-certification-of-perturbation-invariance-for-deep-neural-networks) Yusuke Tsuzuku (/author/yusuke-tsuzuku-11681), Issei Sato (/author/issei-sato-4567), Masashi Sugiyama (/author/masashi-sugiyama-1885)
- Proximal SCOPE for Distributed Sparse Learning (/paper/7890-proximal-scope-for-distributed-sparse-learning) Shenyi Zhao (/author/shenyi-zhao-11682), Gong-Duo Zhang (/author/gong-duo-zhang-11683), Ming-Wei Li (/author/ming-wei-li-11684), Wu-Jun Li (/author/wu-jun-li-7360)
- On Coresets for Logistic Regression (/paper/7891-on-coresets-for-logistic-regression) Alexander Munteanu (/author/alexander-munteanu-11685), Chris Schwiegelshohn (/author/chris-schwiegelshohn-11686), Christian Sohler (/author/christian-sohler-11687), David Woodruff (/author/david-woodruff-6661)
- Neural Ordinary Differential Equations (/paper/7892-neural-ordinary-differential-equations) Tian Qi Chen (/author/tian-qi-chen-10980), Yulia Rubanova (/author/yulia-rubanova-11688), Jesse Bettencourt (/author/jesse-bettencourt-11689), David K. Duvenaud (/author/david-k-duvenaud-4917)
- Unsupervised Learning of Artistic Styles with Archetypal Style Analysis (/paper/7893-unsupervised-learning-of-artistic-styles-with-archetypal-style-analysis) Daan Wynen (/author/daan-wynen-11690), Cordelia Schmid (/author/cordelia-schmid-7293), Julien Mairal (/author/julien-mairal-3911)
- Approximating Real-Time Recurrent Learning with Random Kronecker Factors (/paper/7894-approximating-real-time-recurrent-learning-with-random-kronecker-factors) Asier Mujika (/author/asier-mujika-10284), Florian Meier (/author/florian-meier-10285), Angelika Steger (/author/angelika-steger-10286)
- Contamination Attacks and Mitigation in Multi-Party Machine Learning (/paper/7895-contamination-attacks-and-mitigation-in-multi-party-machine-learning) Jamie Hayes (/author/jamie-hayes-9677), Olga Ohrimenko (/author/olga-ohrimenko-11691)
- An Improved Analysis of Alternating Minimization for Structured Multi-Response Regression (/paper/7896-an-improved-analysis-of-alternating-minimization-for-structured-multi-response-regression) Sheng Chen (/author/sheng-chen-7218), Arindam Banerjee (/author/arindam-banerjee-5303)
- Incorporating Context into Language Encoding Models for fMRI (/paper/7897-incorporating-context-into-language-encoding-models-for-fmri) Shailee Jain (/author/shailee-jain-11692), Alexander Huth (/author/alexander-huth-11693)
- CatBoost: unbiased boosting with categorical features (/paper/7898-catboost-unbiased-boosting-with-categorical-features) Liudmila Prokhorenkova (/author/liudmila-prokhorenkova-11694), Gleb Gusev (/author/gleb-gusev-9278), Aleksandr Vorobev

(/author/aleksandr-vorobev-11695), Anna Veronika Dorogush (/author/anna-veronika-dorogush-11696), Andrey Gulin (/author/andrey-gulin-11697)

- Query K-means Clustering and the Double Dixie Cup Problem (/paper/7899-query-k-means-clustering-and-the-double-dixie-cup-problem) I Chien (/author/i-chien-11698), Chao Pan (/author/chao-pan-10140), Olgica Milenkovic (/author/olgica-milenkovic-9724)
- Training Neural Networks Using Features Replay (/paper/7900-training-neural-networks-using-features-replay) Zhouyuan Huo (/author/zhouyuan-huo-11699), Bin Gu (/author/bin-gu-11700), Heng Huang (/author/heng-huang-4609)
- Modeling Dynamic Missingness of Implicit Feedback for Recommendation (/paper/7901-modeling-dynamic-missingness-of-implicit-feedback-for-recommendation) Menghan Wang (/author/menghan-wang-11701), Mingming Gong (/author/mingming-gong-11702), Xiaolin Zheng (/author/xiaolin-zheng-11703), Kun Zhang (/author/kun-zhang-6824)
- Representation Learning of Compositional Data (/paper/7902-representation-learning-of-compositional-data) Marta Avalos (/author/marta-avalos-11704), Richard Nock (/author/richard-nock-3961), Cheng Soon Ong (/author/cheng-soon-ong-8657), Julien Rouar (/author/julien-rouar-11705), Ke Sun (/author/ke-sun-8354)
- Model-based targeted dimensionality reduction for neuronal population data (/paper/7903-model-based-targeted-dimensionality-reduction-for-neuronal-population-data) Mikio Aoi (/author/mikio-aoi-11706), Jonathan W. Pillow (/author/jonathan-w-pillow-5713)
- On gradient regularizers for MMD GANs (/paper/7904-on-gradient-regularizers-for-mmd-gans) Michael Arbel (/author/michael-arbel-11707), Dougal Sutherland (/author/dougal-sutherland-11708), Mikołaj Bińkowski (/author/mikoaj-binkowski-11709), Arthur Gretton (/author/arthur-gretton-2650)
- Heterogeneous Multi-output Gaussian Process Prediction (/paper/7905-heterogeneous-multi-output-gaussian-process-prediction) Pablo Moreno-Muñoz (/author/pablo-moreno-munoz-11710), Antonio Artés (/author/antonio-artes-11711), Mauricio Álvarez (/author/mauricio-alvarez-10173)
- Large-Scale Stochastic Sampling from the Probability Simplex (/paper/7906-large-scale-stochastic-sampling-from-the-probability-simplex) Jack Baker (/author/jack-baker-11712), Paul Fearnhead (/author/paul-fearnhead-11713), Emily Fox (/author/emily-fox-4032), Christopher Nemeth (/author/christopher-nemeth-11714)
- Policy Regret in Repeated Games (/paper/7907-policy-regret-in-repeated-games) Raman Arora (/author/raman-arora-4325), Michael Dinitz (/author/michael-dinitz-11715), Teodor Vanislavov Marinov (/author/teodor-vanislavov-marinov-10120), Mehryar Mohri (/author/mehryar-mohri-2319)
- A Theory-Based Evaluation of Nearest Neighbor Models Put Into Practice (/paper/7908-a-theory-based-evaluation-of-nearest-neighbor-models-put-into-practice) Hendrik Fichtenberger (/author/hendrik-fichtenberger-11716), Dennis Rohde (/author/dennis-rohde-11717)
- Banach Wasserstein GAN (/paper/7909-banach-wasserstein-gan) Jonas Adler (/author/jonas-adler-11718), Sebastian Lunz (/author/sebastian-lunz-11719)

- Provable Gaussian Embedding with One Observation (/paper/7910-provable-gaussian-embedding-with-one-observation) Ming Yu (/author/ming-yu-9211), Zhuoran Yang (/author/zhuoran-yang-8225), Tuo Zhao (/author/tuo-zhao-5638), Mladen Kolar (/author/mladen-kolar-8400), Princeton Zhaoran Wang (/author/princeton-zhaoran-wang-10432)
- BRITS: Bidirectional Recurrent Imputation for Time Series (/paper/7911-brits-bidirectional-recurrent-imputation-for-time-series) Wei Cao (/author/wei-cao-8429), Dong Wang (/author/dong-wang-11720), Jian Li (/author/jian-li-3569), Hao Zhou (/author/hao-zhou-8692), Lei Li (/author/lei-li-6667), Yitan Li (/author/yitan-li-11721)
- M-Walk: Learning to Walk over Graphs using Monte Carlo Tree Search (/paper/7912-m-walk-learning-to-walk-over-graphs-using-monte-carlo-tree-search) Yelong Shen (/author/yelong-shen-8343), Jianshu Chen (/author/jianshu-chen-8341), Po-Sen Huang (/author/po-sen-huang-11722), Yuqing Guo (/author/yuqing-guo-11723), Jianfeng Gao (/author/jianfeng-gao-8345)
- Extracting Relationships by Multi-Domain Matching (/paper/7913-extracting-relationships-by-multi-domain-matching) Yitong Li (/author/yitong-li-10098), michael Murias (/author/michael-murias-10454), geraldine Dawson (/author/geraldine-dawson-10456), David E. Carlson (/author/david-e-carlson-6582)
- Efficient Gradient Computation for Structured Output Learning with Rational and Tropical Losses (/paper/7914-efficient-gradient-computation-for-structured-output-learning-with-rational-and-tropical-losses) Corinna Cortes (/author/corinna-cortes-1046), Vitaly Kuznetsov (/author/vitaly-kuznetsov-7524), Mehryar Mohri (/author/mehryar-mohri-2319), Dmitry Storcheus (/author/dmitry-storcheus-11724), Scott Yang (/author/scott-yang-7320)
- Generative Probabilistic Novelty Detection with Adversarial Autoencoders (/paper/7915-generative-probabilistic-novelty-detection-with-adversarial-autoencoders) Stanislav Pidhorskyi (/author/stanislav-pidhorskyi-11725), Ranya Almohsen (/author/ranya-almohsen-11726), Gianfranco Doretto (/author/gianfranco-doretto-10379)
- Diminishing Returns Shape Constraints for Interpretability and Regularization (/paper/7916-diminishing-returns-shape-constraints-for-interpretability-and-regularization) Maya Gupta (/author/maya-gupta-4237), Dara Bahri (/author/dara-bahri-11727), Andrew Cotter (/author/andrew-cotter-5178), Kevin Canini (/author/kevin-canini-8474)
- Scalable Hyperparameter Transfer Learning (/paper/7917-scalable-hyperparameter-transfer-learning) Valerio Perrone (/author/valerio-perrone-11728), Rodolphe Jenatton (/author/rodolphe-jenatton-4574), Matthias W. Seeger (/author/matthias-w-seeger-8845), Cedric Archambeau (/author/cedric-archambeau-11729)
- Stochastic Nonparametric Event-Tensor Decomposition (/paper/7918-stochastic-nonparametric-event-tensor-decomposition) Shandian Zhe (/author/shandian-zhe-9309), Yishuai Du (/author/yishuai-du-11730)
- Scaling Gaussian Process Regression with Derivatives (/paper/7919-scaling-gaussian-process-regression-with-derivatives) David Eriksson (/author/david-eriksson-10332), Kun

Dong (/author/kun-dong-10333), Eric Lee (/author/eric-lee-11731), David Bindel (/author/david-bindel-8352), Andrew G. Wilson (/author/andrew-g-wilson-8073)

- Differentially Private Testing of Identity and Closeness of Discrete Distributions (/paper/7920-differentially-private-testing-of-identity-and-closeness-of-discrete-distributions) Jayadev Acharya (/author/jayadev-acharya-5296), Ziteng Sun (/author/ziteng-sun-11732), Huanyu Zhang (/author/huanyu-zhang-11733)
- Bayesian Adversarial Learning (/paper/7921-bayesian-adversarial-learning) Nanyang Ye (/author/nanyang-ye-9462), Zhanxing Zhu (/author/zhanxing-zhu-8367)
- Efficient Convex Completion of Coupled Tensors using Coupled Nuclear Norms (/paper/7922-efficient-convex-completion-of-coupled-tensors-using-coupled-nuclear-norms) Kishan Wimalawarne (/author/kishan-wimalawarne-7693), Hiroshi Mamitsuka (/author/hiroshi-mamitsuka-6488)
- Maximizing Induced Cardinality Under a Determinantal Point Process (/paper/7923-maximizing-induced-cardinality-under-a-determinantal-point-process) Jennifer A. Gillenwater (/author/jennifer-a-gillenwater-7601), Alex Kulesza (/author/alex-kulesza-3510), Sergei Vassilvitskii (/author/sergei-vassilvitskii-8508), Zelda E. Mariet (/author/zelda-e-mariet-8823)
- Causal Inference with Noisy and Missing Covariates via Matrix Factorization (/paper/7924-causal-inference-with-noisy-and-missing-covariates-via-matrix-factorization) Nathan Kallus (/author/nathan-kallus-11734), Xiaojie Mao (/author/xiaojie-mao-11735), Madeleine Udell (/author/madeleine-udell-9049)
- rho-POMDPs have Lipschitz-Continuous epsilon-Optimal Value Functions (/paper/7925-rho-pomdps-have-lipschitz-continuous-epsilon-optimal-value-functions) Mathieu Fehr (/author/mathieu-fehr-11736), Olivier Buffet (/author/olivier-buffet-4581), Vincent Thomas (/author/vincent-thomas-4582), Jilles Dibangoye (/author/jilles-dibangoye-11737)
- Online Structure Learning for Feed-Forward and Recurrent Sum-Product Networks (/paper/7926-online-structure-learning-for-feed-forward-and-recurrent-sum-product-networks) Agastya Kalra (/author/agastya-kalra-11738), Abdullah Rashwan (/author/abdullah-rashwan-11739), Wei-Shou Hsu (/author/wei-shou-hsu-8506), Pascal Poupart (/author/pascal-poupart-2378), Prashant Doshi (/author/prashant-doshi-8031), Georgios Trimponias (/author/georgios-trimponias-11740)
- Uncertainty Sampling is Preconditioned Stochastic Gradient Descent on Zero-One Loss (/paper/7927-uncertainty-sampling-is-preconditioned-stochastic-gradient-descent-on-zero-one-loss) Stephen Mussmann (/author/stephen-mussmann-11741), Percy S. Liang (/author/percy-s-liang-3559)
- A Probabilistic U-Net for Segmentation of Ambiguous Images (/paper/7928-a-probabilistic-u-net-for-segmentation-of-ambiguous-images) Simon Kohl (/author/simon-kohl-11742), Bernardino Romera-Paredes (/author/bernardino-romera-paredes-6602), Clemens Meyer (/author/clemens-meyer-11743), Jeffrey De Fauw (/author/jeffrey-de-fauw-11744), Joseph R. Ledsam (/author/joseph-r-ledsam-11745), Klaus Maier-Hein (/author/klaus-maier-hein-11746), S. M. Ali Eslami (/author/s-m-ali-eslami-7648), Danilo Jimenez Rezende

(/author/danilo-jimenez-rezende-7298), Olaf Ronneberger (/author/olaf-ronneberger-11747)

- Unorganized Malicious Attacks Detection (/paper/7929-unorganized-malicious-attacks-detection) Ming Pang (/author/ming-pang-11748), Wei Gao (/author/wei-gao-11749), Min Tao (/author/min-tao-11750), Zhi-Hua Zhou (/author/zhi-hua-zhou-6485)
- Causal Inference via Kernel Deviance Measures (/paper/7930-causal-inference-via-kernel-deviance-measures) Jovana Mitrovic (/author/jovana-mitrovic-11751), Dino Sejdinovic (/author/dino-sejdinovic-5539), Yee Whye Teh (/author/yee-whye-teh-1925)
- Bayesian Alignments of Warped Multi-Output Gaussian Processes (/paper/7931-bayesian-alignments-of-warped-multi-output-gaussian-processes) Markus Kaiser (/author/markus-kaiser-11752), Clemens Otte (/author/clemens-otte-11753), Thomas Runkler (/author/thomas-runkler-11754), Carl Henrik Ek (/author/carl-henrik-ek-11755)
- Hybrid Macro/Micro Level Backpropagation for Training Deep Spiking Neural Networks (/paper/7932-hybrid-macromicro-level-backpropagation-for-training-deep-spiking-neural-networks) Yingyezhe Jin (/author/yingyezhe-jin-11756), Wenrui Zhang (/author/wenrui-zhang-11757), Peng Li (/author/peng-li-11758)
- Gen-Oja: Simple & Efficient Algorithm for Streaming Generalized Eigenvector Computation (/paper/7933-gen-oja-simple-efficient-algorithm-for-streaming-generalized-eigenvector-computation) Kush Bhatia (/author/kush-bhatia-8349), Aldo Pacchiano (/author/aldo-pacchiano-10529), Nicolas Flammarion (/author/nicolas-flammarion-11759), Peter L. Bartlett (/author/peter-l-bartlett-6348), Michael I. Jordan (/author/michael-i-jordan-6338)
- Efficient online algorithms for fast-rate regret bounds under sparsity (/paper/7934-efficient-online-algorithms-for-fast-rate-regret-bounds-under-sparsity) Pierre Gaillard (/author/pierre-gaillard-5458), Olivier Wintenberger (/author/olivier-wintenberger-11760)
- GILBO: One Metric to Measure Them All (/paper/7935-gilbo-one-metric-to-measure-them-all) Alexander A. Alemi (/author/alexander-a-alemi-8798), Ian Fischer (/author/ian-fischer-11761)
- Predictive Uncertainty Estimation via Prior Networks (/paper/7936-predictive-uncertainty-estimation-via-prior-networks) Andrey Malinin (/author/andrey-malinin-11762), Mark Gales (/author/mark-gales-2225)
- Dual Policy Iteration (/paper/7937-dual-policy-iteration) Wen Sun (/author/wen-sun-9557), Geoffrey J. Gordon (/author/geoffrey-j-gordon-1357), Byron Boots (/author/byron-boots-3782), J. Bagnell (/author/j-bagnell-9560)
- A probabilistic population code based on neural samples (/paper/7938-a-probabilistic-population-code-based-on-neural-samples) Sabyasachi Shivkumar (/author/sabyasachi-shivkumar-11763), Richard Lange (/author/richard-lange-11764), Ankani Chattoraj (/author/ankani-chattoraj-8267), Ralf Haefner (/author/ralf-haefner-11765)
- Manifold-tiling Localized Receptive Fields are Optimal in Similarity-preserving Neural Networks (/paper/7939-manifold-tiling-localized-receptive-fields-are-optimal-in-similarity-preserving-neural-networks) Anirvan Sengupta (/author/anirvan-sengupta-11766), Cengiz

Pehlevan (/author/cengiz-pehlevan-8235), Mariano Tepper (/author/mariano-tepper-11767), Alexander Genkin (/author/alexander-genkin-11768), Dmitri Chklovskii (/author/dmitri-chklovskii-8236)

- On the Convergence and Robustness of Training GANs with Regularized Optimal Transport (/paper/7940-on-the-convergence-and-robustness-of-training-gans-with-regularized-optimal-transport) Maziar Sanjabi (/author/maziar-sanjabi-10068), Jimmy Ba (/author/jimmy-ba-6545), Meisam Razaviyayn (/author/meisam-razaviyayn-7668), Jason D. Lee (/author/jason-d-lee-4783)
- Model-Agnostic Private Learning (/paper/7941-model-agnostic-private-learning) Raef Bassily (/author/raef-bassily-9719), Abhradeep Guha Thakurta (/author/abhradeep-guha-thakurta-8027), Om Dipakbhai Thakkar (/author/om-dipakbhai-thakkar-11769)
- Constrained Generation of Semantically Valid Graphs via Regularizing Variational Autoencoders (/paper/7942-constrained-generation-of-semantically-valid-graphs-via-regularizing-variational-autoencoders) Tengfei Ma (/author/tengfei-ma-11770), Jie Chen (/author/jie-chen-9669), Cao Xiao (/author/cao-xiao-11349)
- Provably Correct Automatic Sub-Differentiation for Qualified Programs (/paper/7943-provably-correct-automatic-sub-differentiation-for-qualified-programs) Sham M. Kakade (/author/sham-m-kakade-2572), Jason D. Lee (/author/jason-d-lee-4783)
- Deep Homogeneous Mixture Models: Representation, Separation, and Approximation (/paper/7944-deep-homogeneous-mixture-models-representation-separation-and-approximation) Priyank Jaini (/author/priyank-jaini-11771), Pascal Poupart (/author/pascal-poupart-2378), Yaoliang Yu (/author/yaoliang-yu-7539)
- Parameters as interacting particles: long time convergence and asymptotic error scaling of neural networks (/paper/7945-parameters-as-interacting-particles-long-time-convergence-and-asymptotic-error-scaling-of-neural-networks) Grant Rotskoff (/author/grant-rotskoff-11772), Eric Vanden-Eijnden (/author/eric-vanden-eijnden-11773)
- Hierarchical Reinforcement Learning for Zero-shot Generalization with Subtask Dependencies (/paper/7946-hierarchical-reinforcement-learning-for-zero-shot-generalization-with-subtask-dependencies) Sungryull Sohn (/author/sungryull-sohn-11774), Junhyuk Oh (/author/junhyuk-oh-8191), Honglak Lee (/author/honglak-lee-3292)
- A Simple Unified Framework for Detecting Out-of-Distribution Samples and Adversarial Attacks (/paper/7947-a-simple-unified-framework-for-detecting-out-of-distribution-samples-and-adversarial-attacks) Kimin Lee (/author/kimin-lee-11775), Kibok Lee (/author/kibok-lee-11776), Honglak Lee (/author/honglak-lee-3292), Jinwoo Shin (/author/jinwoo-shin-6409)
- End-to-End Differentiable Physics for Learning and Control (/paper/7948-end-to-end-differentiable-physics-for-learning-and-control) Filipe de Avila Belbute-Peres (/author/filipe-de-avila-belbute-peres-11777), Kevin Smith (/author/kevin-smith-11778), Kelsey Allen (/author/kelsey-allen-11779), Josh Tenenbaum (/author/josh-tenenbaum-6308), J. Zico Kolter (/author/j-zico-kolter-8526)

- BRUNO: A Deep Recurrent Model for Exchangeable Data (/paper/7949-bruno-a-deep-recurrent-model-for-exchangeable-data) Iryna Korshunova (/author/iryna-korshunova-11780), Jonas Degrave (/author/jonas-degrave-11781), Ferenc Huszar (/author/ferenc-huszar-5504), Yarin Gal (/author/yarin-gal-7645), Arthur Gretton (/author/arthur-gretton-2650), Joni Dambre (/author/joni-dambre-11782)
- Stimulus domain transfer in recurrent models for large scale cortical population prediction on video (/paper/7950-stimulus-domain-transfer-in-recurrent-models-for-large-scale-cortical-population-prediction-on-video) Fabian Sinz (/author/fabian-sinz-6402), Alexander S. Ecker (/author/alexander-s-ecker-7866), Paul Fahey (/author/paul-fahey-11783), Edgar Walker (/author/edgar-walker-11784), Erick Cobos (/author/erick-cobos-11785), Emmanouil Froudarakis (/author/emmanouil-froudarakis-11786), Dimitri Yatsenko (/author/dimitri-yatsenko-11787), Zachary Pitkow (/author/zachary-pitkow-5127), Jacob Reimer (/author/jacob-reimer-11788), Andreas Tolias (/author/andreas-tolias-2548)
- Mapping Images to Scene Graphs with Permutation-Invariant Structured Prediction (/paper/7951-mapping-images-to-scene-graphs-with-permutation-invariant-structured-prediction) Roei Herzig (/author/roei-herzig-11789), Moshiko Raboh (/author/moshiko-raboh-11790), Gal Chechik (/author/gal-chechik-1785), Jonathan Berant (/author/jonathan-berant-11791), Amir Globerson (/author/amir-globerson-2199)
- Distributed Multi-Player Bandits - a Game of Thrones Approach (/paper/7952-distributed-multi-player-bandits-a-game-of-thrones-approach) Ilai Bistritz (/author/ilai-bistritz-11792), Amir Leshem (/author/amir-leshem-4422)
- Efficient Loss-Based Decoding on Graphs for Extreme Classification (/paper/7953-efficient-loss-based-decoding-on-graphs-for-extreme-classification) Itay Evron (/author/itay-evron-11793), Edward Moroshko (/author/edward-moroshko-11794), Koby Crammer (/author/koby-crammer-1989)
- Chaining Mutual Information and Tightening Generalization Bounds (/paper/7954-chaining-mutual-information-and-tightening-generalization-bounds) Amir Asadi (/author/amir-asadi-11795), Emmanuel Abbe (/author/emmanuel-abbe-8271), Sergio Verdu (/author/sergio-verdu-11796)
- Implicit Probabilistic Integrators for ODEs (/paper/7955-implicit-probabilistic-integrators-for-odes) Onur Teymur (/author/onur-teymur-8924), Han Cheng Lie (/author/han-cheng-lie-11797), Tim Sullivan (/author/tim-sullivan-11798), Ben Calderhead (/author/ben-calderhead-3978)
- Learning Attentional Communication for Multi-Agent Cooperation (/paper/7956-learning-attentional-communication-for-multi-agent-cooperation) Jiechuan Jiang (/author/jiechuan-jiang-11799), Zongqing Lu (/author/zongqing-lu-11800)
- Training Deep Models Faster with Robust, Approximate Importance Sampling (/paper/7957-training-deep-models-faster-with-robust-approximate-importance-sampling) Tyler B. Johnson (/author/tyler-b-johnson-8460), Carlos Guestrin (/author/carlos-guestrin-2592)

- Bandit Learning with Implicit Feedback (/paper/7958-bandit-learning-with-implicit-feedback) Yi Qi (/author/yi-qi-11801), Qingyun Wu (/author/qingyun-wu-11802), Hongning Wang (/author/hongning-wang-11803), Jie Tang (/author/jie-tang-11804), Maosong Sun (/author/maosong-sun-5401)
- Unsupervised Text Style Transfer using Language Models as Discriminators (/paper/7959-unsupervised-text-style-transfer-using-language-models-as-discriminators) Zichao Yang (/author/zichao-yang-11805), Zhiting Hu (/author/zhiting-hu-9046), Chris Dyer (/author/chris-dyer-7286), Eric P. Xing (/author/eric-p-xing-6314), Taylor Berg-Kirkpatrick (/author/taylor-berg-kirkpatrick-7420)
- Relational recurrent neural networks (/paper/7960-relational-recurrent-neural-networks) Adam Santoro (/author/adam-santoro-10152), Ryan Faulkner (/author/ryan-faulkner-11806), David Raposo (/author/david-raposo-10153), Jack Rae (/author/jack-rae-8824), Mike Chrzanowski (/author/mike-chrzanowski-11807), Theophane Weber (/author/theophane-weber-8262), Daan Wierstra (/author/daan-wierstra-5118), Oriol Vinyals (/author/oriol-vinyals-5571), Razvan Pascanu (/author/razvan-pascanu-7405), Timothy Lillicrap (/author/timothy-lillicrap-8124)
- Streaming Kernel PCA with *\tilde{O}(\sqrt{n})* Random Features (/paper/7961-streaming-kernel-pca-with-tildeosqrtn-random-features) Md Enayat Ullah (/author/md-enayat-ullah-11808), Poorya Mianjy (/author/poorya-mianjy-10121), Teodor Vanislavov Marinov (/author/teodor-vanislavov-marinov-10120), Raman Arora (/author/raman-arora-4325)
- REFUEL: Exploring Sparse Features in Deep Reinforcement Learning for Fast Disease Diagnosis (/paper/7962-refuel-exploring-sparse-features-in-deep-reinforcement-learning-for-fast-disease-diagnosis) Yu-Shao Peng (/author/yu-shao-peng-11809), Kai-Fu Tang (/author/kai-fu-tang-11810), Hsuan-Tien Lin (/author/hsuan-tien-lin-11811), Edward Chang (/author/edward-chang-9539)
- Bayesian Model-Agnostic Meta-Learning (/paper/7963-bayesian-model-agnostic-meta-learning) Jaesik Yoon (/author/jaesik-yoon-11812), Taesup Kim (/author/taesup-kim-11813), Ousmane Dia (/author/ousmane-dia-11814), Sungwoong Kim (/author/sungwoong-kim-5154), Yoshua Bengio (/author/yoshua-bengio-178), Sungjin Ahn (/author/sungjin-ahn-11815)
- Disconnected Manifold Learning for Generative Adversarial Networks (/paper/7964-disconnected-manifold-learning-for-generative-adversarial-networks) Mahyar Khayatkhoei (/author/mahyar-khayatkhoei-11816), Maneesh K. Singh (/author/maneesh-k-singh-11817), Ahmed Elgammal (/author/ahmed-elgammal-11818)
- Unsupervised Cross-Modal Alignment of Speech and Text Embedding Spaces (/paper/7965-unsupervised-cross-modal-alignment-of-speech-and-text-embedding-spaces) Yu-An Chung (/author/yu-an-chung-11819), Wei-Hung Weng (/author/wei-hung-weng-11820), Schrasing Tong (/author/schrasing-tong-11821), James Glass (/author/james-glass-7706)
- Learning Signed Determinantal Point Processes through the Principal Minor Assignment Problem (/paper/7966-learning-signed-determinantal-point-processes-through-the-

principal-minor-assignment-problem) Victor-Emmanuel Brunel (/author/victor-emmanuel-brunel-11822)

- Out-of-Distribution Detection using Multiple Semantic Label Representations (/paper/7967-out-of-distribution-detection-using-multiple-semantic-label-representations) Gabi Shalev (/author/gabi-shalev-11823), Yossi Adi (/author/yossi-adi-10420), Joseph Keshet (/author/joseph-keshet-2387)
- Stochastic Chebyshev Gradient Descent for Spectral Optimization (/paper/7968-stochastic-chebyshev-gradient-descent-for-spectral-optimization) Insu Han (/author/insu-han-11824), Haim Avron (/author/haim-avron-6660), Jinwoo Shin (/author/jinwoo-shin-6409)
- Revisiting *(\epsilon, \gamma, \tau)*-similarity learning for domain adaptation (/paper/7969-revisiting-epsilon-gamma-tau-similarity-learning-for-domain-adaptation) Sofiane Dhouib (/author/sofiane-dhouib-11825), Ievgen Redko (/author/ievgen-redko-11826)
- How to tell when a clustering is (approximately) correct using convex relaxations (/paper/7970-how-to-tell-when-a-clustering-is-approximately-correct-using-convex-relaxations) Marina Meila (/author/marina-meila-1285)
- Constant Regret, Generalized Mixability, and Mirror Descent (/paper/7971-constant-regret-generalized-mixability-and-mirror-descent) Zakaria Mhammedi (/author/zakaria-mhammedi-11827), Robert C. Williamson (/author/robert-c-williamson-575)
- A Bayesian Approach to Generative Adversarial Imitation Learning (/paper/7972-a-bayesian-approach-to-generative-adversarial-imitation-learning) Wonseok Jeon (/author/wonseok-jeon-11828), Seokin Seo (/author/seokin-seo-11829), Kee-Eung Kim (/author/kee-eung-kim-9743)
- Plug-in Estimation in High-Dimensional Linear Inverse Problems: A Rigorous Analysis (/paper/7973-plug-in-estimation-in-high-dimensional-linear-inverse-problems-a-rigorous-analysis) Alyson K. Fletcher (/author/alyson-k-fletcher-3910), Parthe Pandit (/author/parthe-pandit-11830), Sundeep Rangan (/author/sundeep-rangan-3908), Subrata Sarkar (/author/subrata-sarkar-11831), Philip Schniter (/author/philip-schniter-9757)
- Constrained Cross-Entropy Method for Safe Reinforcement Learning (/paper/7974-constrained-cross-entropy-method-for-safe-reinforcement-learning) Min Wen (/author/min-wen-11832), Ufuk Topcu (/author/ufuk-topcu-11833)
- Multi-Agent Generative Adversarial Imitation Learning (/paper/7975-multi-agent-generative-adversarial-imitation-learning) Jiaming Song (/author/jiaming-song-9966), Hongyu Ren (/author/hongyu-ren-11834), Dorsa Sadigh (/author/dorsa-sadigh-11835), Stefano Ermon (/author/stefano-ermon-5196)
- Adaptive Learning with Unknown Information Flows (/paper/7976-adaptive-learning-with-unknown-information-flows) Yonatan Gur (/author/yonatan-gur-7332), Ahmadreza Momeni (/author/ahmadreza-momeni-11836)
- Forecasting Treatment Responses Over Time Using Recurrent Marginal Structural Networks (/paper/7977-forecasting-treatment-responses-over-time-using-recurrent-marginal-structural-networks) Bryan Lim (/author/bryan-lim-11837)

- Generative modeling for protein structures (/paper/7978-generative-modeling-for-protein-structures) Namrata Anand (/author/namrata-anand-11838), Possu Huang (/author/possu-huang-11839)
- Inference in Deep Gaussian Processes using Stochastic Gradient Hamiltonian Monte Carlo (/paper/7979-inference-in-deep-gaussian-processes-using-stochastic-gradient-hamiltonian-monte-carlo) Marton Havasi (/author/marton-havasi-11840), José Miguel Hernández-Lobato (/author/jose-miguel-hernandez-lobato-6612), Juan José Murillo-Fuentes (/author/juan-jose-murillo-fuentes-11841)
- Knowledge Distillation by On-the-Fly Native Ensemble (/paper/7980-knowledge-distillation-by-on-the-fly-native-ensemble) xu lan (/author/xu-lan-11842), Xiatian Zhu (/author/xiatian-zhu-11843), Shaogang Gong (/author/shaogang-gong-11844)
- Non-Adversarial Mapping with VAEs (/paper/7981-non-adversarial-mapping-with-vaes) Yedid Hoshen (/author/yedid-hoshen-9782)
- Generalisation in humans and deep neural networks (/paper/7982-generalisation-in-humans-and-deep-neural-networks) Robert Geirhos (/author/robert-geirhos-11845), Carlos R. M. Temme (/author/carlos-r-m-temme-11846), Jonas Rauber (/author/jonas-rauber-11847), Heiko H. Schütt (/author/heiko-h-schutt-11848), Matthias Bethge (/author/matthias-bethge-2385), Felix A. Wichmann (/author/felix-a-wichmann-2706)
- Towards Text Generation with Adversarially Learned Neural Outlines (/paper/7983-towards-text-generation-with-adversarially-learned-neural-outlines) Sandeep Subramanian (/author/sandeep-subramanian-11849), Sai Rajeswar Mudumba (/author/sai-rajeswar-mudumba-11850), Alessandro Sordoni (/author/alessandro-sordoni-10384), Adam Trischler (/author/adam-trischler-10225), Aaron C. Courville (/author/aaron-c-courville-2205), Chris Pal (/author/chris-pal-9890)
- cpSGD: Communication-efficient and differentially-private distributed SGD (/paper/7984-cpsgd-communication-efficient-and-differentially-private-distributed-sgd) Naman Agarwal (/author/naman-agarwal-11851), Ananda Theertha Suresh (/author/ananda-theertha-suresh-7152), Felix Xinnan X. Yu (/author/felix-xinnan-x-yu-8749), Sanjiv Kumar (/author/sanjiv-kumar-2539), Brendan McMahan (/author/brendan-mcmahan-6393)
- GPyTorch: Blackbox Matrix-Matrix Gaussian Process Inference with GPU Acceleration (/paper/7985-gpytorch-blackbox-matrix-matrix-gaussian-process-inference-with-gpu-acceleration) Jacob Gardner (/author/jacob-gardner-8209), Geoff Pleiss (/author/geoff-pleiss-10245), Kilian Q. Weinberger (/author/kilian-q-weinberger-2966), David Bindel (/author/david-bindel-8352), Andrew G. Wilson (/author/andrew-g-wilson-8073)
- Diffusion Maps for Textual Network Embedding (/paper/7986-diffusion-maps-for-textual-network-embedding) Xinyuan Zhang (/author/xinyuan-zhang-11852), Yitong Li (/author/yitong-li-10098), Dinghan Shen (/author/dinghan-shen-10024), Lawrence Carin (/author/lawrence-carin-3001)
- Simple, Distributed, and Accelerated Probabilistic Programming (/paper/7987-simple-distributed-and-accelerated-probabilistic-programming) Dustin Tran (/author/dustin-tran-7944), Matthew W. Hoffman (/author/matthew-w-hoffman-7266), Dave Moore

(/author/dave-moore-11279), Christopher Suter (/author/christopher-suter-11853), Srinivas Vasudevan (/author/srinivas-vasudevan-11854), Alexey Radul (/author/alexey-radul-11855)

- VideoCapsuleNet: A Simplified Network for Action Detection (/paper/7988-videocapsulenet-a-simplified-network-for-action-detection) Kevin Duarte (/author/kevin-duarte-11856), Yogesh Rawat (/author/yogesh-rawat-11857), Mubarak Shah (/author/mubarak-shah-11858)
- Rectangular Bounding Process (/paper/7989-rectangular-bounding-process) Xuhui Fan (/author/xuhui-fan-11859), Bin Li (/author/bin-li-11860), Scott SIsson (/author/scott-sisson-11861)
- Improved Algorithms for Collaborative PAC Learning (/paper/7990-improved-algorithms-for-collaborative-pac-learning) Huy Nguyen (/author/huy-nguyen-7141), Lydia Zakynthinou (/author/lydia-zakynthinou-11862)
- Sparse Attentive Backtracking: Temporal Credit Assignment Through Reminding (/paper/7991-sparse-attentive-backtracking-temporal-credit-assignment-through-reminding) Nan Rosemary Ke (/author/nan-rosemary-ke-10063), Anirudh Goyal ALIAS PARTH GOYAL (/author/anirudh-goyal-alias-parth-goyal-8537), Olexa Bilaniuk (/author/olexa-bilaniuk-11863), Jonathan Binas (/author/jonathan-binas-11864), Michael C. Mozer (/author/michael-c-mozer-191), Chris Pal (/author/chris-pal-9890), Yoshua Bengio (/author/yoshua-bengio-178)
- Communication Compression for Decentralized Training (/paper/7992-communication-compression-for-decentralized-training) Hanlin Tang (/author/hanlin-tang-11865), Shaoduo Gan (/author/shaoduo-gan-11866), Ce Zhang (/author/ce-zhang-6471), Tong Zhang (/author/tong-zhang-1901), Ji Liu (/author/ji-liu-4638)
- Depth-Limited Solving for Imperfect-Information Games (/paper/7993-depth-limited-solving-for-imperfect-information-games) Noam Brown (/author/noam-brown-8275), Tuomas Sandholm (/author/tuomas-sandholm-2336), Brandon Amos (/author/brandon-amos-10227)
- Training Deep Neural Networks with 8-bit Floating Point Numbers (/paper/7994-training-deep-neural-networks-with-8-bit-floating-point-numbers) Naigang Wang (/author/naigang-wang-11867), Jungwook Choi (/author/jungwook-choi-11868), Daniel Brand (/author/daniel-brand-11869), Chia-Yu Chen (/author/chia-yu-chen-11870), Kailash Gopalakrishnan (/author/kailash-gopalakrishnan-11871)
- Scalar Posterior Sampling with Applications (/paper/7995-scalar-posterior-sampling-with-applications) Georgios Theocharous (/author/georgios-theocharous-2708), Zheng Wen (/author/zheng-wen-6418), Yasin Abbasi (/author/yasin-abbasi-6448), Nikos Vlassis (/author/nikos-vlassis-9045)
- Understanding Batch Normalization (/paper/7996-understanding-batch-normalization) Nils Bjorck (/author/nils-bjorck-11872), Carla P. Gomes (/author/carla-p-gomes-3333), Bart Selman (/author/bart-selman-3335), Kilian Q. Weinberger (/author/kilian-q-weinberger-2966)

- Adversarial Scene Editing: Automatic Object Removal from Weak Supervision (/paper/7997-adversarial-scene-editing-automatic-object-removal-from-weak-supervision) Rakshith R. Shetty (/author/rakshith-r-shetty-11873), Mario Fritz (/author/mario-fritz-4398), Bernt Schiele (/author/bernt-schiele-6798)
- Attacks Meet Interpretability: Attribute-steered Detection of Adversarial Samples (/paper/7998-attacks-meet-interpretability-attribute-steered-detection-of-adversarial-samples) Guanhong Tao (/author/guanhong-tao-11874), Shiqing Ma (/author/shiqing-ma-11875), Yingqi Liu (/author/yingqi-liu-11876), Xiangyu Zhang (/author/xiangyu-zhang-10568)
- On Neuronal Capacity (/paper/7999-on-neuronal-capacity) Pierre Baldi (/author/pierre-baldi-99), Roman Vershynin (/author/roman-vershynin-11877)
- Breaking the Activation Function Bottleneck through Adaptive Parameterization (/paper/8000-breaking-the-activation-function-bottleneck-through-adaptive-parameterization) Sebastian Flennerhag (/author/sebastian-flennerhag-11878), Hujun Yin (/author/hujun-yin-11879), John Keane (/author/john-keane-11880), Mark Elliot (/author/mark-elliot-11881)
- Learning Loop Invariants for Program Verification (/paper/8001-learning-loop-invariants-for-program-verification) Xujie Si (/author/xujie-si-11882), Hanjun Dai (/author/hanjun-dai-7969), Mukund Raghothaman (/author/mukund-raghothaman-11883), Mayur Naik (/author/mayur-naik-4826), Le Song (/author/le-song-3161)
- Cooperative Learning of Audio and Video Models from Self-Supervised Synchronization (/paper/8002-cooperative-learning-of-audio-and-video-models-from-self-supervised-synchronization) Bruno Korbar (/author/bruno-korbar-11884), Du Tran (/author/du-tran-5615), Lorenzo Torresani (/author/lorenzo-torresani-2736)
- Towards Robust Interpretability with Self-Explaining Neural Networks (/paper/8003-towards-robust-interpretability-with-self-explaining-neural-networks) David Alvarez Melis (/author/david-alvarez-melis-11885), Tommi Jaakkola (/author/tommi-jaakkola-1019)
- Deep State Space Models for Time Series Forecasting (/paper/8004-deep-state-space-models-for-time-series-forecasting) Syama Sundar Rangapuram (/author/syama-sundar-rangapuram-6352), Matthias W. Seeger (/author/matthias-w-seeger-8845), Jan Gasthaus (/author/jan-gasthaus-4126), Lorenzo Stella (/author/lorenzo-stella-11886), Yuyang Wang (/author/yuyang-wang-11887), Tim Januschowski (/author/tim-januschowski-11888)
- Constrained Graph Variational Autoencoders for Molecule Design (/paper/8005-constrained-graph-variational-autoencoders-for-molecule-design) Qi Liu (/author/qi-liu-11889), Miltiadis Allamanis (/author/miltiadis-allamanis-11890), Marc Brockschmidt (/author/marc-brockschmidt-11891), Alexander Gaunt (/author/alexander-gaunt-11892)
- Learning Libraries of Subroutines for Neurally–Guided Bayesian Program Induction (/paper/8006-learning-libraries-of-subroutines-for-neurallyguided-bayesian-program-induction) Kevin Ellis (/author/kevin-ellis-8105), Lucas Morales (/author/lucas-morales-11893), Mathias Sablé-Meyer (/author/mathias-sable-meyer-11894), Armando Solar-

Lezama (/author/armando-solar-lezama-8106), Josh Tenenbaum (/author/josh-tenenbaum-6308)

- Neural Architecture Optimization (/paper/8007-neural-architecture-optimization) Renqian Luo (/author/renqian-luo-11895), Fei Tian (/author/fei-tian-9658), Tao Qin (/author/tao-qin-3844), Enhong Chen (/author/enhong-chen-5478), Tie-Yan Liu (/author/tie-yan-liu-8972)
- Preference Based Adaptation for Learning Objectives (/paper/8008-preference-based-adaptation-for-learning-objectives) Yao-Xiang Ding (/author/yao-xiang-ding-11896), Zhi-Hua Zhou (/author/zhi-hua-zhou-6485)
- Distributed *k*-Clustering for Data with Heavy Noise (/paper/8009-distributed-k-clustering-for-data-with-heavy-noise) Shi Li (/author/shi-li-11581), Xiangyu Guo (/author/xiangyu-guo-11897)
- Beyond Log-concavity: Provable Guarantees for Sampling Multi-modal Distributions using Simulated Tempering Langevin Monte Carlo (/paper/8010-beyond-log-concavity-provable-guarantees-for-sampling-multi-modal-distributions-using-simulated-tempering-langevin-monte-carlo) HOLDEN LEE (/author/holden-lee-11367), Andrej Risteski (/author/andrej-risteski-8048), Rong Ge (/author/rong-ge-6993)
- A General Method for Amortizing Variational Filtering (/paper/8011-a-general-method-for-amortizing-variational-filtering) Joseph Marino (/author/joseph-marino-11898), Milan Cvitkovic (/author/milan-cvitkovic-11899), Yisong Yue (/author/yisong-yue-5193)
- A Reduction for Efficient LDA Topic Reconstruction (/paper/8012-a-reduction-for-efficient-lda-topic-reconstruction) Matteo Almanza (/author/matteo-almanza-11900), Flavio Chierichetti (/author/flavio-chierichetti-5099), Alessandro Panconesi (/author/alessandro-panconesi-11901), Andrea Vattani (/author/andrea-vattani-11902)
- Cluster Variational Approximations for Structure Learning of Continuous-Time Bayesian Networks from Incomplete Data (/paper/8013-cluster-variational-approximations-for-structure-learning-of-continuous-time-bayesian-networks-from-incomplete-data) Dominik Linzner (/author/dominik-linzner-11903), Heinz Koeppl (/author/heinz-koeppl-11904)
- RenderNet: A deep convolutional network for differentiable rendering from 3D shapes (/paper/8014-rendernet-a-deep-convolutional-network-for-differentiable-rendering-from-3d-shapes) Thu H. Nguyen-Phuoc (/author/thu-h-nguyen-phuoc-11905), Chuan Li (/author/chuan-li-11906), Stephen Balaban (/author/stephen-balaban-11907), Yongliang Yang (/author/yongliang-yang-11908)
- Robust Hypothesis Testing Using Wasserstein Uncertainty Sets (/paper/8015-robust-hypothesis-testing-using-wasserstein-uncertainty-sets) RUI GAO (/author/rui-gao-11909), Liyan Xie (/author/liyan-xie-11910), Yao Xie (/author/yao-xie-7968), Huan Xu (/author/huan-xu-3387)
- Robust Detection of Adversarial Attacks by Modeling the Intrinsic Properties of Deep Neural Networks (/paper/8016-robust-detection-of-adversarial-attacks-by-modeling-the-intrinsic-properties-of-deep-neural-networks) Zhihao Zheng (/author/zhihao-zheng-11911), Pengyu Hong (/author/pengyu-hong-11912)

- Monte-Carlo Tree Search for Constrained POMDPs (/paper/8017-monte-carlo-tree-search-for-constrained-pomdps) Jongmin Lee (/author/jongmin-lee-11913), Geon-hyeong Kim (/author/geon-hyeong-kim-11914), Pascal Poupart (/author/pascal-poupart-2378), Kee-Eung Kim (/author/kee-eung-kim-9743)
- Learning to Repair Software Vulnerabilities with Generative Adversarial Networks (/paper/8018-learning-to-repair-software-vulnerabilities-with-generative-adversarial-networks) Jacob Harer (/author/jacob-harer-11915), Onur Ozdemir (/author/onur-ozdemir-11916), Tomo Lazovich (/author/tomo-lazovich-11917), Christopher Reale (/author/christopher-reale-11918), Rebecca Russell (/author/rebecca-russell-11919), Louis Kim (/author/louis-kim-11920), peter chin (/author/peter-chin-11921)
- Layer-Wise Coordination between Encoder and Decoder for Neural Machine Translation (/paper/8019-layer-wise-coordination-between-encoder-and-decoder-for-neural-machine-translation) Tianyu He (/author/tianyu-he-11922), Xu Tan (/author/xu-tan-10775), Yingce Xia (/author/yingce-xia-9116), Di He (/author/di-he-9115), Tao Qin (/author/tao-qin-3844), Zhibo Chen (/author/zhibo-chen-11923), Tie-Yan Liu (/author/tie-yan-liu-8972)
- Dirichlet belief networks for topic structure learning (/paper/8020-dirichlet-belief-networks-for-topic-structure-learning) He Zhao (/author/he-zhao-11924), Lan Du (/author/lan-du-4183), Wray Buntine (/author/wray-buntine-2349), Mingyuan Zhou (/author/mingyuan-zhou-4435)
- Stochastic Expectation Maximization with Variance Reduction (/paper/8021-stochastic-expectation-maximization-with-variance-reduction) Jianfei Chen (/author/jianfei-chen-6460), Jun Zhu (/author/jun-zhu-3840), Yee Whye Teh (/author/yee-whye-teh-1925), Tong Zhang (/author/tong-zhang-1901)
- Submodular Maximization via Gradient Ascent: The Case of Deep Submodular Functions (/paper/8022-submodular-maximization-via-gradient-ascent-the-case-of-deep-submodular-functions) Wenruo Bai (/author/wenruo-bai-7999), William Stafford Noble (/author/william-stafford-noble-11925), Jeff A. Bilmes (/author/jeff-a-bilmes-2747)
- The challenge of realistic music generation: modelling raw audio at scale (/paper/8023-the-challenge-of-realistic-music-generation-modelling-raw-audio-at-scale) Sander Dieleman (/author/sander-dieleman-6494), Aaron van den Oord (/author/aaron-van-den-oord-6493), Karen Simonyan (/author/karen-simonyan-6376)
- Spectral Signatures in Backdoor Attacks (/paper/8024-spectral-signatures-in-backdoor-attacks) Brandon Tran (/author/brandon-tran-11926), Jerry Li (/author/jerry-li-9638), Aleksander Madry (/author/aleksander-madry-10965)
- Reward learning from human preferences and demonstrations in Atari (/paper/8025-reward-learning-from-human-preferences-and-demonstrations-in-atari) Borja Ibarz (/author/borja-ibarz-11927), Jan Leike (/author/jan-leike-10047), Tobias Pohlen (/author/tobias-pohlen-12413), Geoffrey Irving (/author/geoffrey-irving-8797), Shane Legg (/author/shane-legg-3690), Dario Amodei (/author/dario-amodei-10050)
- Approximate Knowledge Compilation by Online Collapsed Importance Sampling (/paper/8026-approximate-knowledge-compilation-by-online-collapsed-importance-

sampling) Tal Friedman (/author/tal-friedman-11928), Guy Van den Broeck (/author/guy-van-den-broeck-6274)

- Neural Arithmetic Logic Units (/paper/8027-neural-arithmetic-logic-units) Andrew Trask (/author/andrew-trask-11929), Felix Hill (/author/felix-hill-11930), Scott E. Reed (/author/scott-e-reed-7879), Jack Rae (/author/jack-rae-8824), Chris Dyer (/author/chris-dyer-7286), Phil Blunsom (/author/phil-blunsom-3916)
- Pipe-SGD: A Decentralized Pipelined SGD Framework for Distributed Deep Net Training (/paper/8028-pipe-sgd-a-decentralized-pipelined-sgd-framework-for-distributed-deep-net-training) Youjie Li (/author/youjie-li-11443), Mingchao Yu (/author/mingchao-yu-11439), Songze Li (/author/songze-li-11442), Salman Avestimehr (/author/salman-avestimehr-10066), Nam Sung Kim (/author/nam-sung-kim-11444), Alexander Schwing (/author/alexander-schwing-9411)
- Improved Expressivity Through Dendritic Neural Networks (/paper/8029-improved-expressivity-through-dendritic-neural-networks) Xundong Wu (/author/xundong-wu-11931), Xiangwen Liu (/author/xiangwen-liu-11932), wei li (/author/wei-li-11933), qing wu (/author/qing-wu-11934)
- Efficient Anomaly Detection via Matrix Sketching (/paper/8030-efficient-anomaly-detection-via-matrix-sketching) Vatsal Sharan (/author/vatsal-sharan-9515), Parikshit Gopalan (/author/parikshit-gopalan-11935), Udi Wieder (/author/udi-wieder-11936)
- Learning to Specialize with Knowledge Distillation for Visual Question Answering (/paper/8031-learning-to-specialize-with-knowledge-distillation-for-visual-question-answering) Jonghwan Mun (/author/jonghwan-mun-10171), Kimin Lee (/author/kimin-lee-11775), Jinwoo Shin (/author/jinwoo-shin-6409), Bohyung Han (/author/bohyung-han-7598)
- A Lyapunov-based Approach to Safe Reinforcement Learning (/paper/8032-a-lyapunov-based-approach-to-safe-reinforcement-learning) Yinlam Chow (/author/yinlam-chow-7145), Ofir Nachum (/author/ofir-nachum-9795), Edgar Duenez-Guzman (/author/edgar-duenez-guzman-11937), Mohammad Ghavamzadeh (/author/mohammad-ghavamzadeh-3309)
- Credit Assignment For Collective Multiagent RL With Global Rewards (/paper/8033-credit-assignment-for-collective-multiagent-rl-with-global-rewards) Duc Thien Nguyen (/author/duc-thien-nguyen-10052), Akshat Kumar (/author/akshat-kumar-4845), Hoong Chuin Lau (/author/hoong-chuin-lau-10053)
- Statistical Optimality of Stochastic Gradient Descent on Hard Learning Problems through Multiple Passes (/paper/8034-statistical-optimality-of-stochastic-gradient-descent-on-hard-learning-problems-through-multiple-passes) Loucas Pillaud-Vivien (/author/loucas-pillaud-vivien-11938), Alessandro Rudi (/author/alessandro-rudi-6400), Francis Bach (/author/francis-bach-6335)
- Does mitigating ML's impact disparity require treatment disparity? (/paper/8035-does-mitigating-mls-impact-disparity-require-treatment-disparity) Zachary Lipton

(/author/zachary-lipton-11939), Julian McAuley (/author/julian-mcauley-11940), Alexandra Chouldechova (/author/alexandra-chouldechova-11941)

- Proximal Graphical Event Models (/paper/8036-proximal-graphical-event-models) Debarun Bhattacharjya (/author/debarun-bhattacharjya-11942), Dharmashankar Subramanian (/author/dharmashankar-subramanian-7434), Tian Gao (/author/tian-gao-8360)
- Bayesian Control of Large MDPs with Unknown Dynamics in Data-Poor Environments (/paper/8037-bayesian-control-of-large-mdps-with-unknown-dynamics-in-data-poor-environments) Mahdi Imani (/author/mahdi-imani-11943), Seyede Fatemeh Ghoreishi (/author/seyede-fatemeh-ghoreishi-11944), Ulisses M. Braga-Neto (/author/ulisses-m-braga-neto-11945)
- Learning Overparameterized Neural Networks via Stochastic Gradient Descent on Structured Data (/paper/8038-learning-overparameterized-neural-networks-via-stochastic-gradient-descent-on-structured-data) Yuanzhi Li (/author/yuanzhi-li-8643), Yingyu Liang (/author/yingyu-liang-6637)
- Hamiltonian Variational Auto-Encoder (/paper/8039-hamiltonian-variational-auto-encoder) Anthony L. Caterini (/author/anthony-l-caterini-11946), Arnaud Doucet (/author/arnaud-doucet-1821), Dino Sejdinovic (/author/dino-sejdinovic-5539)
- Modelling and unsupervised learning of symmetric deformable object categories (/paper/8040-modelling-and-unsupervised-learning-of-symmetric-deformable-object-categories) James Thewlis (/author/james-thewlis-9500), Hakan Bilen (/author/hakan-bilen-8995), Andrea Vedaldi (/author/andrea-vedaldi-3408)
- Graphical model inference: Sequential Monte Carlo meets deterministic approximations (/paper/8041-graphical-model-inference-sequential-monte-carlo-meets-deterministic-approximations) Fredrik Lindsten (/author/fredrik-lindsten-5468), Jouni Helske (/author/jouni-helske-11947), Matti Vihola (/author/matti-vihola-11948)
- Statistical mechanics of low-rank tensor decomposition (/paper/8042-statistical-mechanics-of-low-rank-tensor-decomposition) Jonathan Kadmon (/author/jonathan-kadmon-8873), Surya Ganguli (/author/surya-ganguli-4596)
- Variational Bayesian Monte Carlo (/paper/8043-variational-bayesian-monte-carlo) Luigi Acerbi (/author/luigi-acerbi-7588)
- Sample-Efficient Reinforcement Learning with Stochastic Ensemble Value Expansion (/paper/8044-sample-efficient-reinforcement-learning-with-stochastic-ensemble-value-expansion) Jacob Buckman (/author/jacob-buckman-11949), Danijar Hafner (/author/danijar-hafner-10386), George Tucker (/author/george-tucker-9768), Eugene Brevdo (/author/eugene-brevdo-11950), Honglak Lee (/author/honglak-lee-3292)
- Efficient Online Portfolio with Logarithmic Regret (/paper/8045-efficient-online-portfolio-with-logarithmic-regret) Haipeng Luo (/author/haipeng-luo-7459), Chen-Yu Wei (/author/chen-yu-wei-9224), Kai Zheng (/author/kai-zheng-11951)
- Algorithms and Theory for Multiple-Source Adaptation (/paper/8046-algorithms-and-theory-for-multiple-source-adaptation) Judy Hoffman (/author/judy-hoffman-7393), Mehryar Mohri (/author/mehryar-mohri-2319), Ningshan Zhang (/author/ningshan-zhang-11952)

- Online Reciprocal Recommendation with Theoretical Performance Guarantees (/paper/8047-online-reciprocal-recommendation-with-theoretical-performance-guarantees) Claudio Gentile (/author/claudio-gentile-1834), Nikos Parotsidis (/author/nikos-parotsidis-10108), Fabio Vitale (/author/fabio-vitale-3581)
- The promises and pitfalls of Stochastic Gradient Langevin Dynamics (/paper/8048-the-promises-and-pitfalls-of-stochastic-gradient-langevin-dynamics) Nicolas Brosse (/author/nicolas-brosse-11953), Alain Durmus (/author/alain-durmus-9187), Eric Moulines (/author/eric-moulines-4051)
- How SGD Selects the Global Minima in Over-parameterized Learning: A Dynamical Stability Perspective (/paper/8049-how-sgd-selects-the-global-minima-in-over-parameterized-learning-a-dynamical-stability-perspective) Lei Wu (/author/lei-wu-4222), Chao Ma (/author/chao-ma-10875), Weinan E (/author/weinan-e-11329)
- Differentiable MPC for End-to-end Planning and Control (/paper/8050-differentiable-mpc-for-end-to-end-planning-and-control) Brandon Amos (/author/brandon-amos-10227), Ivan Jimenez (/author/ivan-jimenez-11954), Jacob Sacks (/author/jacob-sacks-11955), Byron Boots (/author/byron-boots-3782), J. Zico Kolter (/author/j-zico-kolter-8526)
- Bilevel learning of the Group Lasso structure (/paper/8051-bilevel-learning-of-the-group-lasso-structure) Jordan Frecon (/author/jordan-frecon-11956), Saverio Salzo (/author/saverio-salzo-11957), Massimiliano Pontil (/author/massimiliano-pontil-2047)
- Constructing Unrestricted Adversarial Examples with Generative Models (/paper/8052-constructing-unrestricted-adversarial-examples-with-generative-models) Yang Song (/author/yang-song-6980), Rui Shu (/author/rui-shu-11315), Nate Kushman (/author/nate-kushman-11958), Stefano Ermon (/author/stefano-ermon-5196)
- Information-theoretic Limits for Community Detection in Network Models (/paper/8053-information-theoretic-limits-for-community-detection-in-network-models) Chuyang Ke (/author/chuyang-ke-11959), Jean Honorio (/author/jean-honorio-4384)
- Learning Conditioned Graph Structures for Interpretable Visual Question Answering (/paper/8054-learning-conditioned-graph-structures-for-interpretable-visual-question-answering) Will Norcliffe-Brown (/author/will-norcliffe-brown-11960), Stathis Vafeias (/author/stathis-vafeias-11961), Sarah Parisot (/author/sarah-parisot-11962)
- Distributionally Robust Graphical Models (/paper/8055-distributionally-robust-graphical-models) Rizal Fathony (/author/rizal-fathony-8520), Ashkan Rezaei (/author/ashkan-rezaei-11963), Mohammad Ali Bashiri (/author/mohammad-ali-bashiri-9454), Xinhua Zhang (/author/xinhua-zhang-3378), Brian Ziebart (/author/brian-ziebart-7447)
- Transfer Learning with Neural AutoML (/paper/8056-transfer-learning-with-neural-automl) Catherine Wong (/author/catherine-wong-11964), Neil Houlsby (/author/neil-houlsby-5503), Yifeng Lu (/author/yifeng-lu-11965), Andrea Gesmundo (/author/andrea-gesmundo-11966)
- Stochastic Primal-Dual Method for Empirical Risk Minimization with O(1) Per-Iteration Complexity (/paper/8057-stochastic-primal-dual-method-for-empirical-risk-minimization-with-o1-per-iteration-complexity) Conghui Tan (/author/conghui-tan-8808), Tong Zhang

(/author/tong-zhang-1901), Shiqian Ma (/author/shiqian-ma-4754), Ji Liu (/author/ji-liu-4638)

- On preserving non-discrimination when combining expert advice (/paper/8058-on-preserving-non-discrimination-when-combining-expert-advice) Avrim Blum (/author/avrim-blum-205), Suriya Gunasekar (/author/suriya-gunasekar-8422), Thodoris Lykouris (/author/thodoris-lykouris-8588), Nati Srebro (/author/nati-srebro-5014)
- Learning to Play With Intrinsically-Motivated, Self-Aware Agents (/paper/8059-learning-to-play-with-intrinsically-motivated-self-aware-agents) Nick Haber (/author/nick-haber-11967), Damian Mrowca (/author/damian-mrowca-11968), Stephanie Wang (/author/stephanie-wang-11969), Li F. Fei-Fei (/author/li-f-fei-fei-7198), Daniel L. Yamins (/author/daniel-l-yamins-6472)
- Scaling provable adversarial defenses (/paper/8060-scaling-provable-adversarial-defenses) Eric Wong (/author/eric-wong-11970), Frank Schmidt (/author/frank-schmidt-11971), Jan Hendrik Metzen (/author/jan-hendrik-metzen-11972), J. Zico Kolter (/author/j-zico-kolter-8526)
- Deep Network for the Integrated 3D Sensing of Multiple People in Natural Images (/paper/8061-deep-network-for-the-integrated-3d-sensing-of-multiple-people-in-natural-images) Andrei Zanfir (/author/andrei-zanfir-11973), Elisabeta Marinoiu (/author/elisabeta-marinoiu-11974), Mihai Zanfir (/author/mihai-zanfir-11975), Alin-Ionut Popa (/author/alin-ionut-popa-11976), Cristian Sminchisescu (/author/cristian-sminchisescu-3601)
- Almost Optimal Algorithms for Linear Stochastic Bandits with Heavy-Tailed Payoffs (/paper/8062-almost-optimal-algorithms-for-linear-stochastic-bandits-with-heavy-tailed-payoffs) Han Shao (/author/han-shao-11977), Xiaotian Yu (/author/xiaotian-yu-11978), Irwin King (/author/irwin-king-3779), Michael R. Lyu (/author/michael-r-lyu-6272)
- Data-dependent PAC-Bayes priors via differential privacy (/paper/8063-data-dependent-pac-bayes-priors-via-differential-privacy) Gintare Karolina Dziugaite (/author/gintare-karolina-dziugaite-11979), Daniel M. Roy (/author/daniel-m-roy-3409)
- Deep Poisson gamma dynamical systems (/paper/8064-deep-poisson-gamma-dynamical-systems) Dandan Guo (/author/dandan-guo-11980), Bo Chen (/author/bo-chen-5132), Hao Zhang (/author/hao-zhang-9976), Mingyuan Zhou (/author/mingyuan-zhou-4435)
- Dimensionality Reduction has Quantifiable Imperfections: Two Geometric Bounds (/paper/8065-dimensionality-reduction-has-quantifiable-imperfections-two-geometric-bounds) Kry Lui (/author/kry-lui-11981), Gavin Weiguang Ding (/author/gavin-weiguang-ding-11982), Ruitong Huang (/author/ruitong-huang-9098), Robert McCann (/author/robert-mccann-11983)
- Teaching Inverse Reinforcement Learners via Features and Demonstrations (/paper/8066-teaching-inverse-reinforcement-learners-via-features-and-demonstrations) Luis Haug (/author/luis-haug-11984), Sebastian Tschiatschek (/author/sebastian-tschiatschek-7386), Adish Singla (/author/adish-singla-10792)
- Wasserstein Distributionally Robust Kalman Filtering (/paper/8067-wasserstein-distributionally-robust-kalman-filtering) Soroosh Shafieezadeh Abadeh (/author/soroosh-

shafieezadeh-abadeh-8439), Viet Anh Nguyen (/author/viet-anh-nguyen-11985), Daniel Kuhn (/author/daniel-kuhn-6677), Peyman Mohajerin Mohajerin Esfahani (/author/peyman-mohajerin-mohajerin-esfahani-8440)

- Generalisation of structural knowledge in the hippocampal-entorhinal system (/paper/8068-generalisation-of-structural-knowledge-in-the-hippocampal-entorhinal-system) James Whittington (/author/james-whittington-11986), Timothy Muller (/author/timothy-muller-11987), Shirely Mark (/author/shirely-mark-11988), Caswell Barry (/author/caswell-barry-11989), Tim Behrens (/author/tim-behrens-11990)
- Graph Oracle Models, Lower Bounds, and Gaps for Parallel Stochastic Optimization (/paper/8069-graph-oracle-models-lower-bounds-and-gaps-for-parallel-stochastic-optimization) Blake E. Woodworth (/author/blake-e-woodworth-8482), Jialei Wang (/author/jialei-wang-9104), Adam Smith (/author/adam-smith-6507), Brendan McMahan (/author/brendan-mcmahan-6393), Nati Srebro (/author/nati-srebro-5014)
- Adversarial Regularizers in Inverse Problems (/paper/8070-adversarial-regularizers-in-inverse-problems) Sebastian Lunz (/author/sebastian-lunz-11719), Carola Schoenlieb (/author/carola-schoenlieb-11991), Ozan Öktem (/author/ozan-oktem-11992)
- Clustering Redemption–Beyond the Impossibility of Kleinberg's Axioms (/paper/8071-clustering-redemptionbeyond-the-impossibility-of-kleinbergs-axioms) Vincent Cohen-Addad (/author/vincent-cohen-addad-10314), Varun Kanade (/author/varun-kanade-4219), Frederik Mallmann-Trenn (/author/frederik-mallmann-trenn-10315)
- Co-teaching: Robust training of deep neural networks with extremely noisy labels (/paper/8072-co-teaching-robust-training-of-deep-neural-networks-with-extremely-noisy-labels) Bo Han (/author/bo-han-11543), Quanming Yao (/author/quanming-yao-11428), Xingrui Yu (/author/xingrui-yu-11993), Gang Niu (/author/gang-niu-4971), Miao Xu (/author/miao-xu-6484), Weihua Hu (/author/weihua-hu-10242), Ivor Tsang (/author/ivor-tsang-8394), Masashi Sugiyama (/author/masashi-sugiyama-1885)
- Variational Inverse Control with Events: A General Framework for Data-Driven Reward Definition (/paper/8073-variational-inverse-control-with-events-a-general-framework-for-data-driven-reward-definition) Justin Fu (/author/justin-fu-9762), Avi Singh (/author/avi-singh-11994), Dibya Ghosh (/author/dibya-ghosh-11995), Larry Yang (/author/larry-yang-11996), Sergey Levine (/author/sergey-levine-4516)
- A convex program for bilinear inversion of sparse vectors (/paper/8074-a-convex-program-for-bilinear-inversion-of-sparse-vectors) Alireza Aghasi (/author/alireza-aghasi-9846), Ali Ahmed (/author/ali-ahmed-11997), Paul Hand (/author/paul-hand-11998), Babhru Joshi (/author/babhru-joshi-11999)
- Adversarial Multiple Source Domain Adaptation (/paper/8075-adversarial-multiple-source-domain-adaptation) Han Zhao (/author/han-zhao-9043), Shanghang Zhang (/author/shanghang-zhang-12000), Guanhang Wu (/author/guanhang-wu-12001), José M. F. Moura (/author/jose-m-f-moura-12002), Joao P. Costeira (/author/joao-p-costeira-5169), Geoffrey J. Gordon (/author/geoffrey-j-gordon-1357)

- Neural Tangent Kernel: Convergence and Generalization in Neural Networks (/paper/8076-neural-tangent-kernel-convergence-and-generalization-in-neural-networks) Arthur Jacot-Guillarmod (/author/arthur-jacot-guillarmod-12003), Clement Hongler (/author/clement-hongler-12004), Franck Gabriel (/author/franck-gabriel-12005)
- Contextual Stochastic Block Models (/paper/8077-contextual-stochastic-block-models) Yash Deshpande (/author/yash-deshpande-7374), Subhabrata Sen (/author/subhabrata-sen-4396), Andrea Montanari (/author/andrea-montanari-4355), Elchanan Mossel (/author/elchanan-mossel-12006)
- A Likelihood-Free Inference Framework for Population Genetic Data using Exchangeable Neural Networks (/paper/8078-a-likelihood-free-inference-framework-for-population-genetic-data-using-exchangeable-neural-networks) Jeffrey Chan (/author/jeffrey-chan-12007), Valerio Perrone (/author/valerio-perrone-11728), Jeffrey Spence (/author/jeffrey-spence-12008), Paul Jenkins (/author/paul-jenkins-12009), Sara Mathieson (/author/sara-mathieson-12010), Yun Song (/author/yun-song-12011)
- Sequential Attend, Infer, Repeat: Generative Modelling of Moving Objects (/paper/8079-sequential-attend-infer-repeat-generative-modelling-of-moving-objects) Adam Kosiorek (/author/adam-kosiorek-9833), Hyunjik Kim (/author/hyunjik-kim-12012), Yee Whye Teh (/author/yee-whye-teh-1925), Ingmar Posner (/author/ingmar-posner-9835)
- Randomized Prior Functions for Deep Reinforcement Learning (/paper/8080-randomized-prior-functions-for-deep-reinforcement-learning) Ian Osband (/author/ian-osband-6756), John Aslanides (/author/john-aslanides-12013), Albin Cassirer (/author/albin-cassirer-12014)
- Compact Representation of Uncertainty in Clustering (/paper/8081-compact-representation-of-uncertainty-in-clustering) Craig Greenberg (/author/craig-greenberg-12015), Nicholas Monath (/author/nicholas-monath-12016), Ari Kobren (/author/ari-kobren-12017), Patrick Flaherty (/author/patrick-flaherty-3224), Andrew McGregor (/author/andrew-mcgregor-12018), Andrew McCallum (/author/andrew-mccallum-1167)
- Learning without the Phase: Regularized PhaseMax Achieves Optimal Sample Complexity (/paper/8082-learning-without-the-phase-regularized-phasemax-achieves-optimal-sample-complexity) Fariborz Salehi (/author/fariborz-salehi-12019), Ehsan Abbasi (/author/ehsan-abbasi-8040), Babak Hassibi (/author/babak-hassibi-893)
- Multilingual Anchoring: Interactive Topic Modeling and Alignment Across Languages (/paper/8083-multilingual-anchoring-interactive-topic-modeling-and-alignment-across-languages) Michelle Yuan (/author/michelle-yuan-12020), Benjamin Van Durme (/author/benjamin-van-durme-12021), Jordan L. Ying (/author/jordan-l-ying-6725)
- Estimators for Multivariate Information Measures in General Probability Spaces (/paper/8084-estimators-for-multivariate-information-measures-in-general-probability-spaces) Arman Rahimzamani (/author/arman-rahimzamani-12022), Himanshu Asnani (/author/himanshu-asnani-12023), Pramod Viswanath (/author/pramod-viswanath-7351), Sreeram Kannan (/author/sreeram-kannan-10093)

- DeepPINK: reproducible feature selection in deep neural networks (/paper/8085-deeppink-reproducible-feature-selection-in-deep-neural-networks) Yang Lu (/author/yang-lu-12024), Yingying Fan (/author/yingying-fan-12025), Jinchi Lv (/author/jinchi-lv-12026), William Stafford Noble (/author/william-stafford-noble-11925)
- HOUDINI: Lifelong Learning as Program Synthesis (/paper/8086-houdini-lifelong-learning-as-program-synthesis) Lazar Valkov (/author/lazar-valkov-9868), Dipak Chaudhari (/author/dipak-chaudhari-12027), Akash Srivastava (/author/akash-srivastava-9867), Charles Sutton (/author/charles-sutton-7542), Swarat Chaudhuri (/author/swarat-chaudhuri-12028)
- Searching for Efficient Multi-Scale Architectures for Dense Image Prediction (/paper/8087-searching-for-efficient-multi-scale-architectures-for-dense-image-prediction) Liang-Chieh Chen (/author/liang-chieh-chen-12029), Maxwell Collins (/author/maxwell-collins-8237), Yukun Zhu (/author/yukun-zhu-7888), George Papandreou (/author/george-papandreou-4535), Barret Zoph (/author/barret-zoph-12030), Florian Schroff (/author/florian-schroff-12031), Hartwig Adam (/author/hartwig-adam-12032), Jon Shlens (/author/jon-shlens-6787)
- Orthogonally Decoupled Variational Gaussian Processes (/paper/8088-orthogonally-decoupled-variational-gaussian-processes) Hugh Salimbeni (/author/hugh-salimbeni-10095), Ching-An Cheng (/author/ching-an-cheng-9120), Byron Boots (/author/byron-boots-3782), Marc Deisenroth (/author/marc-deisenroth-5510)
- Dendritic cortical microcircuits approximate the backpropagation algorithm (/paper/8089-dendritic-cortical-microcircuits-approximate-the-backpropagation-algorithm) João Sacramento (/author/joao-sacramento-12033), Rui Ponte Costa (/author/rui-ponte-costa-12034), Yoshua Bengio (/author/yoshua-bengio-178), Walter Senn (/author/walter-senn-5110)
- Learning Plannable Representations with Causal InfoGAN (/paper/8090-learning-plannable-representations-with-causal-infogan) Thanard Kurutach (/author/thanard-kurutach-12035), Aviv Tamar (/author/aviv-tamar-8289), Ge Yang (/author/ge-yang-9809), Stuart J. Russell (/author/stuart-j-russell-2312), Pieter Abbeel (/author/pieter-abbeel-2686)
- Uniform Convergence of Gradients for Non-Convex Learning and Optimization (/paper/8091-uniform-convergence-of-gradients-for-non-convex-learning-and-optimization) Dylan J. Foster (/author/dylan-j-foster-8174), Ayush Sekhari (/author/ayush-sekhari-12036), Karthik Sridharan (/author/karthik-sridharan-3842)
- Automatic differentiation in ML: Where we are and where we should be going (/paper/8092-automatic-differentiation-in-ml-where-we-are-and-where-we-should-be-going) Bart van Merrienboer (/author/bart-van-merrienboer-11629), Olivier Breuleux (/author/olivier-breuleux-12037), Arnaud Bergeron (/author/arnaud-bergeron-12038), Pascal Lamblin (/author/pascal-lamblin-3382)
- A Bayesian Nonparametric View on Count-Min Sketch (/paper/8093-a-bayesian-nonparametric-view-on-count-min-sketch) Diana Cai (/author/diana-cai-9311), Michael

Mitzenmacher (/author/michael-mitzenmacher-9145), Ryan P. Adams (/author/ryan-p-adams-3862)

- Generalized Cross Entropy Loss for Training Deep Neural Networks with Noisy Labels (/paper/8094-generalized-cross-entropy-loss-for-training-deep-neural-networks-with-noisy-labels) Zhilu Zhang (/author/zhilu-zhang-12039), Mert Sabuncu (/author/mert-sabuncu-12040)
- Loss Surfaces, Mode Connectivity, and Fast Ensembling of DNNs (/paper/8095-loss-surfaces-mode-connectivity-and-fast-ensembling-of-dnns) Timur Garipov (/author/timur-garipov-12041), Pavel Izmailov (/author/pavel-izmailov-12042), Dmitrii Podoprikhin (/author/dmitrii-podoprikhin-8109), Dmitry P. Vetrov (/author/dmitry-p-vetrov-8110), Andrew G. Wilson (/author/andrew-g-wilson-8073)
- Flexible neural representation for physics prediction (/paper/8096-flexible-neural-representation-for-physics-prediction) Damian Mrowca (/author/damian-mrowca-11968), Chengxu Zhuang (/author/chengxu-zhuang-9758), Elias Wang (/author/elias-wang-12043), Nick Haber (/author/nick-haber-11967), Li F. Fei-Fei (/author/li-f-fei-fei-7198), Josh Tenenbaum (/author/josh-tenenbaum-6308), Daniel L. Yamins (/author/daniel-l-yamins-6472)
- Legendre Decomposition for Tensors (/paper/8097-legendre-decomposition-for-tensors) Mahito Sugiyama (/author/mahito-sugiyama-6684), Hiroyuki Nakahara (/author/hiroyuki-nakahara-1258), Koji Tsuda (/author/koji-tsuda-2192)
- Reinforcement Learning of Theorem Proving (/paper/8098-reinforcement-learning-of-theorem-proving) Cezary Kaliszyk (/author/cezary-kaliszyk-12044), Josef Urban (/author/josef-urban-8800), Henryk Michalewski (/author/henryk-michalewski-12045), Miroslav Olšák (/author/miroslav-olsak-12046)
- Data Amplification: A Unified and Competitive Approach to Property Estimation (/paper/8099-data-amplification-a-unified-and-competitive-approach-to-property-estimation) Yi HAO (/author/yi-hao-10637), Alon Orlitsky (/author/alon-orlitsky-521), Ananda Theertha Suresh (/author/ananda-theertha-suresh-7152), Yihong Wu (/author/yihong-wu-12047)
- Group Equivariant Capsule Networks (/paper/8100-group-equivariant-capsule-networks) Jan Eric Lenssen (/author/jan-eric-lenssen-12048), Matthias Fey (/author/matthias-fey-12049), Pascal Libuschewski (/author/pascal-libuschewski-12050)
- Stein Variational Gradient Descent as Moment Matching (/paper/8101-stein-variational-gradient-descent-as-moment-matching) Qiang Liu (/author/qiang-liu-5417), Dilin Wang (/author/dilin-wang-8892)
- Differential Privacy for Growing Databases (/paper/8102-differential-privacy-for-growing-databases) Rachel Cummings (/author/rachel-cummings-12051), Sara Krehbiel (/author/sara-krehbiel-12052), Kevin A. Lai (/author/kevin-a-lai-12053), Uthaipon Tantipongpipat (/author/uthaipon-tantipongpipat-12054)
- Exploration in Structured Reinforcement Learning (/paper/8103-exploration-in-structured-reinforcement-learning) Jungseul Ok (/author/jungseul-ok-12055), Alexandre Proutiere

(/author/alexandre-proutiere-6653), Damianos Tranos (/author/damianos-tranos-12056)

- A Statistical Recurrent Model on the Manifold of Symmetric Positive Definite Matrices (/paper/8104-a-statistical-recurrent-model-on-the-manifold-of-symmetric-positive-definite-matrices) Rudrasis Chakraborty (/author/rudrasis-chakraborty-12057), Chun-Hao Yang (/author/chun-hao-yang-12058), Xingjian Zhen (/author/xingjian-zhen-12059), Monami Banerjee (/author/monami-banerjee-12060), Derek Archer (/author/derek-archer-12061), David Vaillancourt (/author/david-vaillancourt-12062), Vikas Singh (/author/vikas-singh-3679), Baba Vemuri (/author/baba-vemuri-12063)
- Balanced Policy Evaluation and Learning (/paper/8105-balanced-policy-evaluation-and-learning) Nathan Kallus (/author/nathan-kallus-11734)
- Distributed Multitask Reinforcement Learning with Quadratic Convergence (/paper/8106-distributed-multitask-reinforcement-learning-with-quadratic-convergence) Rasul Tutunov (/author/rasul-tutunov-12064), Dongho Kim (/author/dongho-kim-5629), Haitham Bou Ammar (/author/haitham-bou-ammar-12065)
- Improving Neural Program Synthesis with Inferred Execution Traces (/paper/8107-improving-neural-program-synthesis-with-inferred-execution-traces) Richard Shin (/author/richard-shin-8806), Illia Polosukhin (/author/illia-polosukhin-10447), Dawn Song (/author/dawn-song-4395)
- Adaptive Path-Integral Autoencoders: Representation Learning and Planning for Dynamical Systems (/paper/8108-adaptive-path-integral-autoencoders-representation-learning-and-planning-for-dynamical-systems) Jung-Su Ha (/author/jung-su-ha-12066), Young-Jin Park (/author/young-jin-park-12067), Hyeok-Joo Chae (/author/hyeok-joo-chae-12068), Soon-Seo Park (/author/soon-seo-park-12069), Han-Lim Choi (/author/han-lim-choi-12070)
- Policy-Conditioned Uncertainty Sets for Robust Markov Decision Processes (/paper/8109-policy-conditioned-uncertainty-sets-for-robust-markov-decision-processes) Andrea Tirinzoni (/author/andrea-tirinzoni-11611), Marek Petrik (/author/marek-petrik-4010), Xiangli Chen (/author/xiangli-chen-12071), Brian Ziebart (/author/brian-ziebart-7447)
- GLoMo: Unsupervised Learning of Transferable Relational Graphs (/paper/8110-glomo-unsupervised-learning-of-transferable-relational-graphs) Zhilin Yang (/author/zhilin-yang-8638), Jake Zhao (/author/jake-zhao-12072), Bhuwan Dhingra (/author/bhuwan-dhingra-12073), Kaiming He (/author/kaiming-he-7876), William W. Cohen (/author/william-w-cohen-1648), Ruslan R. Salakhutdinov (/author/ruslan-r-salakhutdinov-2816), Yann LeCun (/author/yann-lecun-200)
- Online Learning of Quantum States (/paper/8111-online-learning-of-quantum-states) Scott Aaronson (/author/scott-aaronson-12074), Xinyi Chen (/author/xinyi-chen-12075), Elad Hazan (/author/elad-hazan-3630), Satyen Kale (/author/satyen-kale-3631), Ashwin Nayak (/author/ashwin-nayak-12076)
- Wavelet regression and additive models for irregularly spaced data (/paper/8112-wavelet-regression-and-additive-models-for-irregularly-spaced-data) Asad Haris (/author/asad-

haris-12077), Ali Shojaie (/author/ali-shojaie-4564), Noah Simon (/author/noah-simon-12078)

- Inferring Latent Velocities from Weather Radar Data using Gaussian Processes (/paper/8113-inferring-latent-velocities-from-weather-radar-data-using-gaussian-processes) Rico Angell (/author/rico-angell-12079), Daniel R. Sheldon (/author/daniel-r-sheldon-3728)
- A Structured Prediction Approach for Label Ranking (/paper/8114-a-structured-prediction-approach-for-label-ranking) Anna Korba (/author/anna-korba-12080), Alexandre Garcia (/author/alexandre-garcia-12081), Florence d'Alché-Buc (/author/florence-dalche-buc-8739)
- Efficient High Dimensional Bayesian Optimization with Additivity and Quadrature Fourier Features (/paper/8115-efficient-high-dimensional-bayesian-optimization-with-additivity-and-quadrature-fourier-features) Mojmir Mutny (/author/mojmir-mutny-12082), Andreas Krause (/author/andreas-krause-3758)
- FastGRNN: A Fast, Accurate, Stable and Tiny Kilobyte Sized Gated Recurrent Neural Network (/paper/8116-fastgrnn-a-fast-accurate-stable-and-tiny-kilobyte-sized-gated-recurrent-neural-network) Aditya Kusupati (/author/aditya-kusupati-12083), Manish Singh (/author/manish-singh-4703), Kush Bhatia (/author/kush-bhatia-8349), Ashish Kumar (/author/ashish-kumar-10659), Prateek Jain (/author/prateek-jain-3906), Manik Varma (/author/manik-varma-4601)
- Reversible Recurrent Neural Networks (/paper/8117-reversible-recurrent-neural-networks) Matthew MacKay (/author/matthew-mackay-12084), Paul Vicol (/author/paul-vicol-12085), Jimmy Ba (/author/jimmy-ba-6545), Roger B. Grosse (/author/roger-b-grosse-6301)
- SING: Symbol-to-Instrument Neural Generator (/paper/8118-sing-symbol-to-instrument-neural-generator) Alexandre Defossez (/author/alexandre-defossez-12086), Neil Zeghidour (/author/neil-zeghidour-10291), Nicolas Usunier (/author/nicolas-usunier-3174), Leon Bottou (/author/leon-bottou-12087), Francis Bach (/author/francis-bach-6335)
- Learning Compressed Transforms with Low Displacement Rank (/paper/8119-learning-compressed-transforms-with-low-displacement-rank) Anna Thomas (/author/anna-thomas-8918), Albert Gu (/author/albert-gu-12088), Tri Dao (/author/tri-dao-10306), Atri Rudra (/author/atri-rudra-12089), Christopher Ré (/author/christopher-re-7222)
- Theoretical Linear Convergence of Unfolded ISTA and Its Practical Weights and Thresholds (/paper/8120-theoretical-linear-convergence-of-unfolded-ista-and-its-practical-weights-and-thresholds) Xiaohan Chen (/author/xiaohan-chen-11300), Jialin Liu (/author/jialin-liu-12090), Zhangyang Wang (/author/zhangyang-wang-7234), Wotao Yin (/author/wotao-yin-10215)
- Iterative Value-Aware Model Learning (/paper/8121-iterative-value-aware-model-learning) Amir-massoud Farahmand (/author/amir-massoud-farahmand-4865)
- Invariant Representations without Adversarial Training (/paper/8122-invariant-representations-without-adversarial-training) Daniel Moyer (/author/daniel-moyer-12091), Shuyang Gao (/author/shuyang-gao-9073), Rob Brekelmans (/author/rob-brekelmans-

12092), Aram Galstyan (/author/aram-galstyan-5050), Greg Ver Steeg (/author/greg-ver-steeg-7623)

- Robot Learning in Homes: Improving Generalization and Reducing Dataset Bias (/paper/8123-robot-learning-in-homes-improving-generalization-and-reducing-dataset-bias) Abhinav Gupta (/author/abhinav-gupta-4024), Adithyavairavan Murali (/author/adithyavairavan-murali-12093), Dhiraj Prakashchand Gandhi (/author/dhiraj-prakashchand-gandhi-12094), Lerrel Pinto (/author/lerrel-pinto-9558)
- Learning Safe Policies with Expert Guidance (/paper/8124-learning-safe-policies-with-expert-guidance) Jessie Huang (/author/jessie-huang-12095), Fa Wu (/author/fa-wu-12096), Doina Precup (/author/doina-precup-1573), Yang Cai (/author/yang-cai-12097)
- Bayesian multi-domain learning for cancer subtype discovery from next-generation sequencing count data (/paper/8125-bayesian-multi-domain-learning-for-cancer-subtype-discovery-from-next-generation-sequencing-count-data) Ehsan Hajiramezanali (/author/ehsan-hajiramezanali-12098), Siamak Zamani Dadaneh (/author/siamak-zamani-dadaneh-12099), Alireza Karbalayghareh (/author/alireza-karbalayghareh-12100), Mingyuan Zhou (/author/mingyuan-zhou-4435), Xiaoning Qian (/author/xiaoning-qian-12101)
- Learning SMaLL Predictors (/paper/8126-learning-small-predictors) Vikas Garg (/author/vikas-garg-6645), Ofer Dekel (/author/ofer-dekel-2467), Lin Xiao (/author/lin-xiao-4475)
- Phase Retrieval Under a Generative Prior (/paper/8127-phase-retrieval-under-a-generative-prior) Paul Hand (/author/paul-hand-11998), Oscar Leong (/author/oscar-leong-12102), Vlad Voroninski (/author/vlad-voroninski-9194)
- Quadrature-based features for kernel approximation (/paper/8128-quadrature-based-features-for-kernel-approximation) Marina Munkhoeva (/author/marina-munkhoeva-12103), Yermek Kapushev (/author/yermek-kapushev-12104), Evgeny Burnaev (/author/evgeny-burnaev-12105), Ivan Oseledets (/author/ivan-oseledets-12106)
- Reducing Network Agnostophobia (/paper/8129-reducing-network-agnostophobia) Akshay Raj Dhamija (/author/akshay-raj-dhamija-12107), Manuel Günther (/author/manuel-gunther-12108), Terrance Boult (/author/terrance-boult-12109)
- A Stein variational Newton method (/paper/8130-a-stein-variational-newton-method) Gianluca Detommaso (/author/gianluca-detommaso-12110), Tiangang Cui (/author/tiangang-cui-12111), Youssef Marzouk (/author/youssef-marzouk-9731), Alessio Spantini (/author/alessio-spantini-12112), Robert Scheichl (/author/robert-scheichl-12113)
- Watch Your Step: Learning Node Embeddings via Graph Attention (/paper/8131-watch-your-step-learning-node-embeddings-via-graph-attention) Sami Abu-El-Haija (/author/sami-abu-el-haija-12114), Bryan Perozzi (/author/bryan-perozzi-12115), Rami Al-Rfou (/author/rami-al-rfou-12116), Alexander A. Alemi (/author/alexander-a-alemi-8798)
- Visual Reinforcement Learning with Imagined Goals (/paper/8132-visual-reinforcement-learning-with-imagined-goals) Ashvin V. Nair (/author/ashvin-v-nair-9349), Vitchyr Pong (/author/vitchyr-pong-12117), Murtaza Dalal (/author/murtaza-dalal-12118), Shikhar Bahl

(/author/shikhar-bahl-12119), Steven Lin (/author/steven-lin-12120), Sergey Levine (/author/sergey-levine-4516)

- Deep Predictive Coding Network with Local Recurrent Processing for Object Recognition (/paper/8133-deep-predictive-coding-network-with-local-recurrent-processing-for-object-recognition) Kuan Han (/author/kuan-han-12121), Haiguang Wen (/author/haiguang-wen-12122), Yizhen Zhang (/author/yizhen-zhang-12123), Di Fu (/author/di-fu-12124), Eugenio Culurciello (/author/eugenio-culurciello-2970), Zhongming Liu (/author/zhongming-liu-12125)
- PAC-Bayes bounds for stable algorithms with instance-dependent priors (/paper/8134-pac-bayes-bounds-for-stable-algorithms-with-instance-dependent-priors) Omar Rivasplata (/author/omar-rivasplata-12126), Csaba Szepesvari (/author/csaba-szepesvari-6449), John S. Shawe-Taylor (/author/john-s-shawe-taylor-7341), Emilio Parrado-Hernandez (/author/emilio-parrado-hernandez-12127), Shiliang Sun (/author/shiliang-sun-10445)
- Beyond Grids: Learning Graph Representations for Visual Recognition (/paper/8135-beyond-grids-learning-graph-representations-for-visual-recognition) Yin Li (/author/yin-li-12128), Abhinav Gupta (/author/abhinav-gupta-4024)
- The Limit Points of (Optimistic) Gradient Descent in Min-Max Optimization (/paper/8136-the-limit-points-of-optimistic-gradient-descent-in-min-max-optimization) Constantinos Daskalakis (/author/constantinos-daskalakis-8166), Ioannis Panageas (/author/ioannis-panageas-10277)
- Coordinate Descent with Bandit Sampling (/paper/8137-coordinate-descent-with-bandit-sampling) Farnood Salehi (/author/farnood-salehi-12129), Patrick Thiran (/author/patrick-thiran-12130), Elisa Celis (/author/elisa-celis-12131)
- Deep Dynamical Modeling and Control of Unsteady Fluid Flows (/paper/8138-deep-dynamical-modeling-and-control-of-unsteady-fluid-flows) Jeremy Morton (/author/jeremy-morton-12132), Antony Jameson (/author/antony-jameson-12133), Mykel J. Kochenderfer (/author/mykel-j-kochenderfer-11316), Freddie Witherden (/author/freddie-witherden-12134)
- Confounding-Robust Policy Improvement (/paper/8139-confounding-robust-policy-improvement) Nathan Kallus (/author/nathan-kallus-11734), Angela Zhou (/author/angela-zhou-12135)
- The Importance of Sampling inMeta-Reinforcement Learning (/paper/8140-the-importance-of-sampling-inmeta-reinforcement-learning) Bradly Stadie (/author/bradly-stadie-9541), Ge Yang (/author/ge-yang-9809), Rein Houthooft (/author/rein-houthooft-9000), Peter Chen (/author/peter-chen-12136), Yan Duan (/author/yan-duan-8999), Yuhuai Wu (/author/yuhuai-wu-8701), Pieter Abbeel (/author/pieter-abbeel-2686), Ilya Sutskever (/author/ilya-sutskever-3959)
- Representer Point Selection for Explaining Deep Neural Networks (/paper/8141-representer-point-selection-for-explaining-deep-neural-networks) Chih-Kuan Yeh (/author/chih-kuan-yeh-12137), Joon Kim (/author/joon-kim-12138), Ian En-Hsu Yen (/author/ian-en-hsu-yen-7160), Pradeep K. Ravikumar (/author/pradeep-k-ravikumar-6369)

- The Effect of Network Width on the Performance of Large-batch Training (/paper/8142-the-effect-of-network-width-on-the-performance-of-large-batch-training) Lingjiao Chen (/author/lingjiao-chen-12139), Hongyi Wang (/author/hongyi-wang-12140), Jinman Zhao (/author/jinman-zhao-12141), Dimitris Papailiopoulos (/author/dimitris-papailiopoulos-8142), Paraschos Koutris (/author/paraschos-koutris-12142)
- SNIPER: Efficient Multi-Scale Training (/paper/8143-sniper-efficient-multi-scale-training) Bharat Singh (/author/bharat-singh-12143), Mahyar Najibi (/author/mahyar-najibi-11597), Larry S. Davis (/author/larry-s-davis-2789)
- The Sample Complexity of Semi-Supervised Learning with Nonparametric Mixture Models (/paper/8144-the-sample-complexity-of-semi-supervised-learning-with-nonparametric-mixture-models) Chen Dan (/author/chen-dan-12144), Liu Leqi (/author/liu-leqi-12145), Bryon Aragam (/author/bryon-aragam-12146), Pradeep K. Ravikumar (/author/pradeep-k-ravikumar-6369), Eric P. Xing (/author/eric-p-xing-6314)
- Hardware Conditioned Policies for Multi-Robot Transfer Learning (/paper/8145-hardware-conditioned-policies-for-multi-robot-transfer-learning) Tao Chen (/author/tao-chen-12147), Adithyavairavan Murali (/author/adithyavairavan-murali-12093), Abhinav Gupta (/author/abhinav-gupta-4024)
- Co-regularized Alignment for Unsupervised Domain Adaptation (/paper/8146-co-regularized-alignment-for-unsupervised-domain-adaptation) Abhishek Kumar (/author/abhishek-kumar-4631), Prasanna Sattigeri (/author/prasanna-sattigeri-10230), Kahini Wadhawan (/author/kahini-wadhawan-12148), Leonid Karlinsky (/author/leonid-karlinsky-4636), Rogerio Feris (/author/rogerio-feris-10640), Bill Freeman (/author/bill-freeman-4274), Gregory Wornell (/author/gregory-wornell-12149)
- Statistical and Computational Trade-Offs in Kernel K-Means (/paper/8147-statistical-and-computational-trade-offs-in-kernel-k-means) Daniele Calandriello (/author/daniele-calandriello-7146), Lorenzo Rosasco (/author/lorenzo-rosasco-3011)
- Assessing the Scalability of Biologically-Motivated Deep Learning Algorithms and Architectures (/paper/8148-assessing-the-scalability-of-biologically-motivated-deep-learning-algorithms-and-architectures) Sergey Bartunov (/author/sergey-bartunov-12150), Adam Santoro (/author/adam-santoro-10152), Blake Richards (/author/blake-richards-12151), Luke Marris (/author/luke-marris-12152), Geoffrey E. Hinton (/author/geoffrey-e-hinton-121), Timothy Lillicrap (/author/timothy-lillicrap-8124)
- Learning Attractor Dynamics for Generative Memory (/paper/8149-learning-attractor-dynamics-for-generative-memory) Yan Wu (/author/yan-wu-12153), Gregory Wayne (/author/gregory-wayne-8123), Karol Gregor (/author/karol-gregor-4805), Timothy Lillicrap (/author/timothy-lillicrap-8124)
- The emergence of multiple retinal cell types through efficient coding of natural movies (/paper/8150-the-emergence-of-multiple-retinal-cell-types-through-efficient-coding-of-natural-movies) Samuel Ocko (/author/samuel-ocko-12154), Jack Lindsey (/author/jack-lindsey-12155), Surya Ganguli (/author/surya-ganguli-4596), Stephane Deny (/author/stephane-deny-12156)

- Gather-Excite: Exploiting Feature Context in Convolutional Neural Networks (/paper/8151-gather-excite-exploiting-feature-context-in-convolutional-neural-networks) Jie Hu (/author/jie-hu-12157), Li Shen (/author/li-shen-5256), Samuel Albanie (/author/samuel-albanie-12158), Gang Sun (/author/gang-sun-12159), Andrea Vedaldi (/author/andrea-vedaldi-3408)
- The Global Anchor Method for Quantifying Linguistic Shifts and Domain Adaptation (/paper/8152-the-global-anchor-method-for-quantifying-linguistic-shifts-and-domain-adaptation) Zi Yin (/author/zi-yin-10691), Vin Sachidananda (/author/vin-sachidananda-12160), Balaji Prabhakar (/author/balaji-prabhakar-12161)
- Identification and Estimation of Causal Effects from Dependent Data (/paper/8153-identification-and-estimation-of-causal-effects-from-dependent-data) Eli Sherman (/author/eli-sherman-12162), Ilya Shpitser (/author/ilya-shpitser-2313)
- Deepcode: Feedback Codes via Deep Learning (/paper/8154-deepcode-feedback-codes-via-deep-learning) Hyeji Kim (/author/hyeji-kim-10094), Yihan Jiang (/author/yihan-jiang-12163), Sreeram Kannan (/author/sreeram-kannan-10093), Sewoong Oh (/author/sewoong-oh-4356), Pramod Viswanath (/author/pramod-viswanath-7351)
- Learning and Testing Causal Models with Interventions (/paper/8155-learning-and-testing-causal-models-with-interventions) Jayadev Acharya (/author/jayadev-acharya-5296), Arnab Bhattacharyya (/author/arnab-bhattacharyya-12164), Constantinos Daskalakis (/author/constantinos-daskalakis-8166), Saravanan Kandasamy (/author/saravanan-kandasamy-12165)
- Implicit Bias of Gradient Descent on Linear Convolutional Networks (/paper/8156-implicit-bias-of-gradient-descent-on-linear-convolutional-networks) Suriya Gunasekar (/author/suriya-gunasekar-8422), Jason D. Lee (/author/jason-d-lee-4783), Daniel Soudry (/author/daniel-soudry-6992), Nati Srebro (/author/nati-srebro-5014)
- DAGs with NO TEARS: Continuous Optimization for Structure Learning (/paper/8157-dags-with-no-tears-continuous-optimization-for-structure-learning) Xun Zheng (/author/xun-zheng-6463), Bryon Aragam (/author/bryon-aragam-12146), Pradeep K. Ravikumar (/author/pradeep-k-ravikumar-6369), Eric P. Xing (/author/eric-p-xing-6314)
- PAC-Bayes Tree: weighted subtrees with guarantees (/paper/8158-pac-bayes-tree-weighted-subtrees-with-guarantees) Tin D. Nguyen (/author/tin-d-nguyen-12166), Samory Kpotufe (/author/samory-kpotufe-4345)
- Multi-objective Maximization of Monotone Submodular Functions with Cardinality Constraint (/paper/8159-multi-objective-maximization-of-monotone-submodular-functions-with-cardinality-constraint) Rajan Udwani (/author/rajan-udwani-12167)
- Sanity Checks for Saliency Maps (/paper/8160-sanity-checks-for-saliency-maps) Julius Adebayo (/author/julius-adebayo-12168), Justin Gilmer (/author/justin-gilmer-10304), Michael Muelly (/author/michael-muelly-12169), Ian Goodfellow (/author/ian-goodfellow-4369), Moritz Hardt (/author/moritz-hardt-5305), Been Kim (/author/been-kim-7252)
- Probabilistic Model-Agnostic Meta-Learning (/paper/8161-probabilistic-model-agnostic-meta-learning) Chelsea Finn (/author/chelsea-finn-8627), Kelvin Xu (/author/kelvin-xu-

9796), Sergey Levine (/author/sergey-levine-4516)

- Reinforcement Learning with Multiple Experts: A Bayesian Model Combination Approach (/paper/8162-reinforcement-learning-with-multiple-experts-a-bayesian-model-combination-approach) Michael Gimelfarb (/author/michael-gimelfarb-12170), Scott Sanner (/author/scott-sanner-4817), Chi-Guhn Lee (/author/chi-guhn-lee-12171)
- e-SNLI: Natural Language Inference with Natural Language Explanations (/paper/8163-e-snli-natural-language-inference-with-natural-language-explanations) Oana-Maria Camburu (/author/oana-maria-camburu-12172), Tim Rocktäschel (/author/tim-rocktaschel-9961), Thomas Lukasiewicz (/author/thomas-lukasiewicz-12173), Phil Blunsom (/author/phil-blunsom-3916)
- Fast Approximate Natural Gradient Descent in a Kronecker Factored Eigenbasis (/paper/8164-fast-approximate-natural-gradient-descent-in-a-kronecker-factored-eigenbasis) Thomas George (/author/thomas-george-12174), César Laurent (/author/cesar-laurent-12175), Xavier Bouthillier (/author/xavier-bouthillier-8199), Nicolas Ballas (/author/nicolas-ballas-12176), Pascal Vincent (/author/pascal-vincent-2037)
- Learning convex bounds for linear quadratic control policy synthesis (/paper/8165-learning-convex-bounds-for-linear-quadratic-control-policy-synthesis) Jack Umenberger (/author/jack-umenberger-12177), Thomas B. Schön (/author/thomas-b-schon-6616)
- Neural Proximal Gradient Descent for Compressive Imaging (/paper/8166-neural-proximal-gradient-descent-for-compressive-imaging) Morteza Mardani (/author/morteza-mardani-12178), Qingyun Sun (/author/qingyun-sun-12179), David Donoho (/author/david-donoho-2681), Vardan Papyan (/author/vardan-papyan-12180), Hatef Monajemi (/author/hatef-monajemi-12181), Shreyas Vasanawala (/author/shreyas-vasanawala-12182), John Pauly (/author/john-pauly-12183)
- Towards Understanding Learning Representations: To What Extent Do Different Neural Networks Learn the Same Representation (/paper/8167-towards-understanding-learning-representations-to-what-extent-do-different-neural-networks-learn-the-same-representation) Liwei Wang (/author/liwei-wang-4410), Lunjia Hu (/author/lunjia-hu-12184), Jiayuan Gu (/author/jiayuan-gu-12185), Zhiqiang Hu (/author/zhiqiang-hu-10319), Yue Wu (/author/yue-wu-7330), Kun He (/author/kun-he-9285), John Hopcroft (/author/john-hopcroft-6600)
- Optimal Algorithms for Continuous Non-monotone Submodular and DR-Submodular Maximization (/paper/8168-optimal-algorithms-for-continuous-non-monotone-submodular-and-dr-submodular-maximization) Rad Niazadeh (/author/rad-niazadeh-12186), Tim Roughgarden (/author/tim-roughgarden-6426), Joshua Wang (/author/joshua-wang-9838)
- An intriguing failing of convolutional neural networks and the CoordConv solution (/paper/8169-an-intriguing-failing-of-convolutional-neural-networks-and-the-coordconv-solution) Rosanne Liu (/author/rosanne-liu-11230), Joel Lehman (/author/joel-lehman-11424), Piero Molino (/author/piero-molino-12187), Felipe Petroski Such (/author/felipe-petroski-such-11423), Eric Frank (/author/eric-frank-12188), Alex Sergeev (/author/alex-sergeev-11228), Jason Yosinski (/author/jason-yosinski-7288)

- Trading robust representations for sample complexity through self-supervised visual experience (/paper/8170-trading-robust-representations-for-sample-complexity-through-self-supervised-visual-experience) Andrea Tacchetti (/author/andrea-tacchetti-10088), Stephen Voinea (/author/stephen-voinea-8125), Georgios Evangelopoulos (/author/georgios-evangelopoulos-12189)
- Invertibility of Convolutional Generative Networks from Partial Measurements (/paper/8171-invertibility-of-convolutional-generative-networks-from-partial-measurements) Fangchang Ma (/author/fangchang-ma-12190), Ulas Ayaz (/author/ulas-ayaz-12191), Sertac Karaman (/author/sertac-karaman-12192)
- Ex ante coordination and collusion in zero-sum multi-player extensive-form games (/paper/8172-ex-ante-coordination-and-collusion-in-zero-sum-multi-player-extensive-form-games) Gabriele Farina (/author/gabriele-farina-10683), Andrea Celli (/author/andrea-celli-12193), Nicola Gatti (/author/nicola-gatti-11426), Tuomas Sandholm (/author/tuomas-sandholm-2336)
- Multi-Agent Reinforcement Learning via Double Averaging Primal-Dual Optimization (/paper/8173-multi-agent-reinforcement-learning-via-double-averaging-primal-dual-optimization) Hoi-To Wai (/author/hoi-to-wai-11492), Zhuoran Yang (/author/zhuoran-yang-8225), Princeton Zhaoran Wang (/author/princeton-zhaoran-wang-10432), Mingyi Hong (/author/mingyi-hong-7669)
- Improving Online Algorithms via ML Predictions (/paper/8174-improving-online-algorithms-via-ml-predictions) Manish Purohit (/author/manish-purohit-12194), Zoya Svitkina (/author/zoya-svitkina-12195), Ravi Kumar (/author/ravi-kumar-4090)
- Global Non-convex Optimization with Discretized Diffusions (/paper/8175-global-non-convex-optimization-with-discretized-diffusions) Murat A. Erdogdu (/author/murat-a-erdogdu-9429), Lester Mackey (/author/lester-mackey-8079), Ohad Shamir (/author/ohad-shamir-3602)
- Theoretical guarantees for EM under misspecified Gaussian mixture models (/paper/8176-theoretical-guarantees-for-em-under-misspecified-gaussian-mixture-models) Raaz Dwivedi (/author/raaz-dwivedi-12196), nhật Hồ (/author/nhat-ho-12197), Koulik Khamaru (/author/koulik-khamaru-12198), Martin J. Wainwright (/author/martin-j-wainwright-1952), Michael I. Jordan (/author/michael-i-jordan-6338)
- Coupled Variational Bayes via Optimization Embedding (/paper/8177-coupled-variational-bayes-via-optimization-embedding) Bo Dai (/author/bo-dai-6604), Hanjun Dai (/author/hanjun-dai-7969), Niao He (/author/niao-he-7138), Weiyang Liu (/author/weiyang-liu-9985), Zhen Liu (/author/zhen-liu-11620), Jianshu Chen (/author/jianshu-chen-8341), Lin Xiao (/author/lin-xiao-4475), Le Song (/author/le-song-3161)
- Improving Explorability in Variational Inference with Annealed Variational Objectives (/paper/8178-improving-explorability-in-variational-inference-with-annealed-variational-objectives) Chin-Wei Huang (/author/chin-wei-huang-12199), Shawn Tan (/author/shawn-tan-12200), Alexandre Lacoste (/author/alexandre-lacoste-9288), Aaron C. Courville (/author/aaron-c-courville-2205)

- Latent Alignment and Variational Attention (/paper/8179-latent-alignment-and-variational-attention) Yuntian Deng (/author/yuntian-deng-12201), Yoon Kim (/author/yoon-kim-12202), Justin Chiu (/author/justin-chiu-12203), Demi Guo (/author/demi-guo-12204), Alexander Rush (/author/alexander-rush-12205)
- Towards Deep Conversational Recommendations (/paper/8180-towards-deep-conversational-recommendations) Raymond Li (/author/raymond-li-12206), Samira Ebrahimi Kahou (/author/samira-ebrahimi-kahou-10462), Hannes Schulz (/author/hannes-schulz-12207), Vincent Michalski (/author/vincent-michalski-7579), Laurent Charlin (/author/laurent-charlin-3275), Chris Pal (/author/chris-pal-9890)
- Unsupervised Depth Estimation, 3D Face Rotation and Replacement (/paper/8181-unsupervised-depth-estimation-3d-face-rotation-and-replacement) Joel Ruben Antony Moniz (/author/joel-ruben-antony-moniz-12208), Christopher Beckham (/author/christopher-beckham-9888), Simon Rajotte (/author/simon-rajotte-12209), Sina Honari (/author/sina-honari-12210), Chris Pal (/author/chris-pal-9890)
- Generalization Bounds for Uniformly Stable Algorithms (/paper/8182-generalization-bounds-for-uniformly-stable-algorithms) Vitaly Feldman (/author/vitaly-feldman-6643), Jan Vondrak (/author/jan-vondrak-8170)
- Deep Anomaly Detection Using Geometric Transformations (/paper/8183-deep-anomaly-detection-using-geometric-transformations) Izhak Golan (/author/izhak-golan-12211), Ran El-Yaniv (/author/ran-el-yaniv-1591)
- Large Scale computation of Means and Clusters for Persistence Diagrams using Optimal Transport (/paper/8184-large-scale-computation-of-means-and-clusters-for-persistence-diagrams-using-optimal-transport) Theo Lacombe (/author/theo-lacombe-12212), Marco Cuturi (/author/marco-cuturi-2992), Steve OUDOT (/author/steve-oudot-12213)
- Entropy Rate Estimation for Markov Chains with Large State Space (/paper/8185-entropy-rate-estimation-for-markov-chains-with-large-state-space) Yanjun Han (/author/yanjun-han-4536), Jiantao Jiao (/author/jiantao-jiao-11081), Chuan-Zheng Lee (/author/chuan-zheng-lee-12214), Tsachy Weissman (/author/tsachy-weissman-12215), Yihong Wu (/author/yihong-wu-12047), Tiancheng Yu (/author/tiancheng-yu-12216)
- Adaptive Methods for Nonconvex Optimization (/paper/8186-adaptive-methods-for-nonconvex-optimization) Manzil Zaheer (/author/manzil-zaheer-9885), Sashank Reddi (/author/sashank-reddi-12217), Devendra Sachan (/author/devendra-sachan-12218), Satyen Kale (/author/satyen-kale-3631), Sanjiv Kumar (/author/sanjiv-kumar-2539)
- Object-Oriented Dynamics Predictor (/paper/8187-object-oriented-dynamics-predictor) Guangxiang Zhu (/author/guangxiang-zhu-12219), Zhiao Huang (/author/zhiao-huang-9717), Chongjie Zhang (/author/chongjie-zhang-7638)
- Adaptive Skip Intervals: Temporal Abstraction for Recurrent Dynamical Models (/paper/8188-adaptive-skip-intervals-temporal-abstraction-for-recurrent-dynamical-models) Alexander Neitz (/author/alexander-neitz-12220), Giambattista Parascandolo (/author/giambattista-parascandolo-9467), Stefan Bauer (/author/stefan-bauer-10123), Bernhard Schölkopf (/author/bernhard-scholkopf-1472)

- Scalable End-to-End Autonomous Vehicle Testing via Rare-event Simulation (/paper/8189-scalable-end-to-end-autonomous-vehicle-testing-via-rare-event-simulation) Matthew O'Kelly (/author/matthew-okelly-12221), Aman Sinha (/author/aman-sinha-8659), Hongseok Namkoong (/author/hongseok-namkoong-8456), Russ Tedrake (/author/russ-tedrake-3995), John C. Duchi (/author/john-c-duchi-3453)
- Reinforcement Learning for Solving the Vehicle Routing Problem (/paper/8190-reinforcement-learning-for-solving-the-vehicle-routing-problem) MohammadReza Nazari (/author/mohammadreza-nazari-12222), Afshin Oroojlooy (/author/afshin-oroojlooy-12223), Lawrence Snyder (/author/lawrence-snyder-12224), Martin Takac (/author/martin-takac-7653)
- ATOMO: Communication-efficient Learning via Atomic Sparsification (/paper/8191-atomo-communication-efficient-learning-via-atomic-sparsification) Hongyi Wang (/author/hongyi-wang-12140), Scott Sievert (/author/scott-sievert-12225), Shengchao Liu (/author/shengchao-liu-12226), Zachary Charles (/author/zachary-charles-12227), Dimitris Papailiopoulos (/author/dimitris-papailiopoulos-8142), Stephen Wright (/author/stephen-wright-5125)
- Dynamic Network Model from Partial Observations (/paper/8192-dynamic-network-model-from-partial-observations) Elahe Ghalebi (/author/elahe-ghalebi-12228), Baharan Mirzasoleiman (/author/baharan-mirzasoleiman-6553), Radu Grosu (/author/radu-grosu-12229), Jure Leskovec (/author/jure-leskovec-4767)
- Life-Long Disentangled Representation Learning with Cross-Domain Latent Homologies (/paper/8193-life-long-disentangled-representation-learning-with-cross-domain-latent-homologies) Alessandro Achille (/author/alessandro-achille-12230), Tom Eccles (/author/tom-eccles-12231), Loic Matthey (/author/loic-matthey-12232), Chris Burgess (/author/chris-burgess-12233), Nicholas Watters (/author/nicholas-watters-10087), Alexander Lerchner (/author/alexander-lerchner-12234), Irina Higgins (/author/irina-higgins-12235)
- Maximizing acquisition functions for Bayesian optimization (/paper/8194-maximizing-acquisition-functions-for-bayesian-optimization) James Wilson (/author/james-wilson-12236), Frank Hutter (/author/frank-hutter-8217), Marc Deisenroth (/author/marc-deisenroth-5510)
- On Markov Chain Gradient Descent (/paper/8195-on-markov-chain-gradient-descent) Tao Sun (/author/tao-sun-10312), Yuejiao Sun (/author/yuejiao-sun-12237), Wotao Yin (/author/wotao-yin-10215)
- Variance-Reduced Stochastic Gradient Descent on Streaming Data (/paper/8196-variance-reduced-stochastic-gradient-descent-on-streaming-data) Ellango Jothimurugesan (/author/ellango-jothimurugesan-12238), Ashraf Tahmasbi (/author/ashraf-tahmasbi-12239), Phillip Gibbons (/author/phillip-gibbons-12240), Srikanta Tirthapura (/author/srikanta-tirthapura-12241)
- Online Robust Policy Learning in the Presence of Unknown Adversaries (/paper/8197-online-robust-policy-learning-in-the-presence-of-unknown-adversaries) Aaron Havens

(/author/aaron-havens-12242), Zhanhong Jiang (/author/zhanhong-jiang-10281), Soumik Sarkar (/author/soumik-sarkar-10283)

- Uplift Modeling from Separate Labels (/paper/8198-uplift-modeling-from-separate-labels) Ikko Yamane (/author/ikko-yamane-12243), Florian Yger (/author/florian-yger-12244), Jamal Atif (/author/jamal-atif-12245), Masashi Sugiyama (/author/masashi-sugiyama-1885)
- Learning Invariances using the Marginal Likelihood (/paper/8199-learning-invariances-using-the-marginal-likelihood) Mark van der Wilk (/author/mark-van-der-wilk-7646), Matthias Bauer (/author/matthias-bauer-9123), ST John (/author/st-john-12246), James Hensman (/author/james-hensman-5588)
- Non-delusional Q-learning and value-iteration (/paper/8200-non-delusional-q-learning-and-value-iteration) Tyler Lu (/author/tyler-lu-11206), Dale Schuurmans (/author/dale-schuurmans-1904), Craig Boutilier (/author/craig-boutilier-2040)
- Using Large Ensembles of Control Variates for Variational Inference (/paper/8201-using-large-ensembles-of-control-variates-for-variational-inference) Tomas Geffner (/author/tomas-geffner-12247), Justin Domke (/author/justin-domke-4762)
- Post: Device Placement with Cross-Entropy Minimization and Proximal Policy Optimization (/paper/8202-post-device-placement-with-cross-entropy-minimization-and-proximal-policy-optimization) Yuanxiang Gao (/author/yuanxiang-gao-12248), Li Chen (/author/li-chen-12249), Baochun Li (/author/baochun-li-12250)
- Learning to Reason with Third Order Tensor Products (/paper/8203-learning-to-reason-with-third-order-tensor-products) Imanol Schlag (/author/imanol-schlag-12251), Jürgen Schmidhuber (/author/jurgen-schmidhuber-583)
- Memory Augmented Policy Optimization for Program Synthesis and Semantic Parsing (/paper/8204-memory-augmented-policy-optimization-for-program-synthesis-and-semantic-parsing) Chen Liang (/author/chen-liang-12252), Mohammad Norouzi (/author/mohammad-norouzi-5636), Jonathan Berant (/author/jonathan-berant-11791), Quoc V. Le (/author/quoc-v-le-7287), Ni Lao (/author/ni-lao-4860)
- Persistence Fisher Kernel: A Riemannian Manifold Kernel for Persistence Diagrams (/paper/8205-persistence-fisher-kernel-a-riemannian-manifold-kernel-for-persistence-diagrams) Tam Le (/author/tam-le-12253), Makoto Yamada (/author/makoto-yamada-4961)
- Neural Voice Cloning with a Few Samples (/paper/8206-neural-voice-cloning-with-a-few-samples) Sercan Arik (/author/sercan-arik-10466), Jitong Chen (/author/jitong-chen-12254), Kainan Peng (/author/kainan-peng-10469), Wei Ping (/author/wei-ping-7991), Yanqi Zhou (/author/yanqi-zhou-10471)
- Blind Deconvolutional Phase Retrieval via Convex Programming (/paper/8207-blind-deconvolutional-phase-retrieval-via-convex-programming) Ali Ahmed (/author/ali-ahmed-11997), Alireza Aghasi (/author/alireza-aghasi-9846), Paul Hand (/author/paul-hand-11998)
- Scalable Laplacian K-modes (/paper/8208-scalable-laplacian-k-modes) Imtiaz Ziko (/author/imtiaz-ziko-12255), Eric Granger (/author/eric-granger-1779), Ismail Ben Ayed

(/author/ismail-ben-ayed-12256)

- A Retrieve-and-Edit Framework for Predicting Structured Outputs (/paper/8209-a-retrieve-and-edit-framework-for-predicting-structured-outputs) Tatsunori B. Hashimoto (/author/tatsunori-b-hashimoto-10408), Kelvin Guu (/author/kelvin-guu-12257), Yonatan Oren (/author/yonatan-oren-12258), Percy S. Liang (/author/percy-s-liang-3559)
- Testing for Families of Distributions via the Fourier Transform (/paper/8210-testing-for-families-of-distributions-via-the-fourier-transform) Alistair Stewart (/author/alistair-stewart-11617), Ilias Diakonikolas (/author/ilias-diakonikolas-7127), Clement Canonne (/author/clement-canonne-12259)
- Thwarting Adversarial Examples: An *L_0*-Robust Sparse Fourier Transform (/paper/8211-thwarting-adversarial-examples-an-l_0-robust-sparse-fourier-transform) Mitali Bafna (/author/mitali-bafna-12260), Jack Murtagh (/author/jack-murtagh-12261), Nikhil Vyas (/author/nikhil-vyas-12262)
- Blockwise Parallel Decoding for Deep Autoregressive Models (/paper/8212-blockwise-parallel-decoding-for-deep-autoregressive-models) Mitchell Stern (/author/mitchell-stern-10020), Noam Shazeer (/author/noam-shazeer-8322), Jakob Uszkoreit (/author/jakob-uszkoreit-10298)
- Low-Rank Tucker Decomposition of Large Tensors Using TensorSketch (/paper/8213-low-rank-tucker-decomposition-of-large-tensors-using-tensorsketch) Osman Asif Malik (/author/osman-asif-malik-12263), Stephen Becker (/author/stephen-becker-5275)
- A Simple Cache Model for Image Recognition (/paper/8214-a-simple-cache-model-for-image-recognition) Emin Orhan (/author/emin-orhan-5029)
- Clebsch–Gordan Nets: a Fully Fourier Space Spherical Convolutional Neural Network (/paper/8215-clebschgordan-nets-a-fully-fourier-space-spherical-convolutional-neural-network) Risi Kondor (/author/risi-kondor-3436), Zhen Lin (/author/zhen-lin-12264), Shubhendu Trivedi (/author/shubhendu-trivedi-7345)
- Bayesian Nonparametric Spectral Estimation (/paper/8216-bayesian-nonparametric-spectral-estimation) Felipe Tobar (/author/felipe-tobar-8084)
- A Spectral View of Adversarially Robust Features (/paper/8217-a-spectral-view-of-adversarially-robust-features) Shivam Garg (/author/shivam-garg-12265), Vatsal Sharan (/author/vatsal-sharan-9515), Brian Zhang (/author/brian-zhang-12266), Gregory Valiant (/author/gregory-valiant-6736)
- Synaptic Strength For Convolutional Neural Network (/paper/8218-synaptic-strength-for-convolutional-neural-network) CHEN LIN (/author/chen-lin-12267), Zhao Zhong (/author/zhao-zhong-12268), Wu Wei (/author/wu-wei-12269), Junjie Yan (/author/junjie-yan-12270)
- Human-in-the-Loop Interpretability Prior (/paper/8219-human-in-the-loop-interpretability-prior) Isaac Lage (/author/isaac-lage-12271), Andrew Ross (/author/andrew-ross-12272), Samuel J. Gershman (/author/samuel-j-gershman-7284), Been Kim (/author/been-kim-7252), Finale Doshi-Velez (/author/finale-doshi-velez-8323)

- Learning To Learn Around A Common Mean (/paper/8220-learning-to-learn-around-a-common-mean) Giulia Denevi (/author/giulia-denevi-12273), Carlo Ciliberto (/author/carlo-ciliberto-8528), Dimitris Stamos (/author/dimitris-stamos-7225), Massimiliano Pontil (/author/massimiliano-pontil-2047)
- Backpropagation with Callbacks: Foundations for Efficient and Expressive Differentiable Programming (/paper/8221-backpropagation-with-callbacks-foundations-for-efficient-and-expressive-differentiable-programming) Fei Wang (/author/fei-wang-10718), James Decker (/author/james-decker-12274), Xilun Wu (/author/xilun-wu-12275), Gregory Essertel (/author/gregory-essertel-12276), Tiark Rompf (/author/tiark-rompf-12277)
- Learning with SGD and Random Features (/paper/8222-learning-with-sgd-and-random-features) Luigi Carratino (/author/luigi-carratino-9975), Alessandro Rudi (/author/alessandro-rudi-6400), Lorenzo Rosasco (/author/lorenzo-rosasco-3011)
- Total stochastic gradient algorithms and applications in reinforcement learning (/paper/8223-total-stochastic-gradient-algorithms-and-applications-in-reinforcement-learning) Paavo Parmas (/author/paavo-parmas-12278)
- Glow: Generative Flow with Invertible 1x1 Convolutions (/paper/8224-glow-generative-flow-with-invertible-1x1-convolutions) Durk P. Kingma (/author/durk-p-kingma-4708), Prafulla Dhariwal (/author/prafulla-dhariwal-12279)
- Nonparametric Density Estimation under Adversarial Losses (/paper/8225-nonparametric-density-estimation-under-adversarial-losses) Shashank Singh (/author/shashank-singh-7508), Ananya Uppal (/author/ananya-uppal-12280), Boyue Li (/author/boyue-li-10302), Chun-Liang Li (/author/chun-liang-li-9709), Manzil Zaheer (/author/manzil-zaheer-9885), Barnabas Poczos (/author/barnabas-poczos-7509)
- Generalizing Point Embeddings using the Wasserstein Space of Elliptical Distributions (/paper/8226-generalizing-point-embeddings-using-the-wasserstein-space-of-elliptical-distributions) Boris Muzellec (/author/boris-muzellec-12281), Marco Cuturi (/author/marco-cuturi-2992)
- Learning to Share and Hide Intentions using Information Regularization (/paper/8227-learning-to-share-and-hide-intentions-using-information-regularization) Daniel Strouse (/author/daniel-strouse-12282), Max Kleiman-Weiner (/author/max-kleiman-weiner-12283), Josh Tenenbaum (/author/josh-tenenbaum-6308), Matt Botvinick (/author/matt-botvinick-12284), David J. Schwab (/author/david-j-schwab-8698)
- Predictive Approximate Bayesian Computation via Saddle Points (/paper/8228-predictive-approximate-bayesian-computation-via-saddle-points) Yingxiang Yang (/author/yingxiang-yang-10145), Bo Dai (/author/bo-dai-6604), Negar Kiyavash (/author/negar-kiyavash-9831), Niao He (/author/niao-he-7138)
- Robustness of conditional GANs to noisy labels (/paper/8229-robustness-of-conditional-gans-to-noisy-labels) Kiran K. Thekumparampil (/author/kiran-k-thekumparampil-8147), Ashish Khetan (/author/ashish-khetan-8572), Zinan Lin (/author/zinan-lin-10794), Sewoong Oh (/author/sewoong-oh-4356)

- Robust Learning of Fixed-Structure Bayesian Networks (/paper/8230-robust-learning-of-fixed-structure-bayesian-networks) Yu Cheng (/author/yu-cheng-8899), Ilias Diakonikolas (/author/ilias-diakonikolas-7127), Daniel Kane (/author/daniel-kane-12285), Alistair Stewart (/author/alistair-stewart-11617)
- Improving Simple Models with Confidence Profiles (/paper/8231-improving-simple-models-with-confidence-profiles) Amit Dhurandhar (/author/amit-dhurandhar-10618), Karthikeyan Shanmugam (/author/karthikeyan-shanmugam-7412), Ronny Luss (/author/ronny-luss-3756), Peder A. Olsen (/author/peder-a-olsen-3287)
- PCA of high dimensional random walks with comparison to neural network training (/paper/8232-pca-of-high-dimensional-random-walks-with-comparison-to-neural-network-training) Joseph Antognini (/author/joseph-antognini-12286), Jascha Sohl-Dickstein (/author/jascha-sohl-dickstein-7915)
- Learning to Solve SMT Formulas (/paper/8233-learning-to-solve-smt-formulas) Mislav Balunovic (/author/mislav-balunovic-12287), Pavol Bielik (/author/pavol-bielik-12288), Martin Vechev (/author/martin-vechev-12289)
- Lifted Weighted Mini-Bucket (/paper/8234-lifted-weighted-mini-bucket) Nicholas Gallo (/author/nicholas-gallo-12290), Alexander T. Ihler (/author/alexander-t-ihler-5418)
- Learning and Inference in Hilbert Space with Quantum Graphical Models (/paper/8235-learning-and-inference-in-hilbert-space-with-quantum-graphical-models) Siddarth Srinivasan (/author/siddarth-srinivasan-12291), Carlton Downey (/author/carlton-downey-7952), Byron Boots (/author/byron-boots-3782)
- Unsupervised Image-to-Image Translation Using Domain-Specific Variational Information Bound (/paper/8236-unsupervised-image-to-image-translation-using-domain-specific-variational-information-bound) Hadi Kazemi (/author/hadi-kazemi-12292), Sobhan Soleymani (/author/sobhan-soleymani-12293), Fariborz Taherkhani (/author/fariborz-taherkhani-12294), Seyed Iranmanesh (/author/seyed-iranmanesh-10443), Nasser Nasrabadi (/author/nasser-nasrabadi-12295)
- Adversarial Risk and Robustness: General Definitions and Implications for the Uniform Distribution (/paper/8237-adversarial-risk-and-robustness-general-definitions-and-implications-for-the-uniform-distribution) Dimitrios Diochnos (/author/dimitrios-diochnos-12296), Saeed Mahloujifar (/author/saeed-mahloujifar-12297), Mohammad Mahmoody (/author/mohammad-mahmoody-12298)
- Gaussian Process Prior Variational Autoencoders (/paper/8238-gaussian-process-prior-variational-autoencoders) Francesco Paolo Casale (/author/francesco-paolo-casale-12299), Adrian Dalca (/author/adrian-dalca-12300), Luca Saglietti (/author/luca-saglietti-12301), Jennifer Listgarten (/author/jennifer-listgarten-3009), Nicolo Fusi (/author/nicolo-fusi-11119)
- 3D Steerable CNNs: Learning Rotationally Equivariant Features in Volumetric Data (/paper/8239-3d-steerable-cnns-learning-rotationally-equivariant-features-in-volumetric-data) Maurice Weiler (/author/maurice-weiler-12302), Wouter Boomsma (/author/wouter-

boomsma-9899), Mario Geiger (/author/mario-geiger-12303), Max Welling (/author/max-welling-2187), Taco Cohen (/author/taco-cohen-12304)

- Context-aware Synthesis and Placement of Object Instances (/paper/8240-context-aware-synthesis-and-placement-of-object-instances) Donghoon Lee (/author/donghoon-lee-12305), Ming-Yu Liu (/author/ming-yu-liu-7201), Ming-Hsuan Yang (/author/ming-hsuan-yang-1946), Sifei Liu (/author/sifei-liu-9612), Jinwei Gu (/author/jinwei-gu-9615), Jan Kautz (/author/jan-kautz-9470)
- Convex Elicitation of Continuous Properties (/paper/8241-convex-elicitation-of-continuous-properties) Jessica Finocchiaro (/author/jessica-finocchiaro-12306), Rafael Frongillo (/author/rafael-frongillo-6349)
- Mesh-TensorFlow: Deep Learning for Supercomputers (/paper/8242-mesh-tensorflow-deep-learning-for-supercomputers) Noam Shazeer (/author/noam-shazeer-8322), Youlong Cheng (/author/youlong-cheng-12307), Niki Parmar (/author/niki-parmar-10296), Dustin Tran (/author/dustin-tran-7944), Ashish Vaswani (/author/ashish-vaswani-10295), Penporn Koanantakool (/author/penporn-koanantakool-12308), Peter Hawkins (/author/peter-hawkins-12309), HyoukJoong Lee (/author/hyoukjoong-lee-12310), Mingsheng Hong (/author/mingsheng-hong-12311), Cliff Young (/author/cliff-young-12312), Ryan Sepassi (/author/ryan-sepassi-12313), Blake Hechtman (/author/blake-hechtman-12314)
- Learning Abstract Options (/paper/8243-learning-abstract-options) Matthew Riemer (/author/matthew-riemer-12315), Miao Liu (/author/miao-liu-6632), Gerald Tesauro (/author/gerald-tesauro-53)
- Bounded-Loss Private Prediction Markets (/paper/8244-bounded-loss-private-prediction-markets) Rafael Frongillo (/author/rafael-frongillo-6349), Bo Waggoner (/author/bo-waggoner-8386)
- Temporal alignment and latent Gaussian process factor inference in population spike trains (/paper/8245-temporal-alignment-and-latent-gaussian-process-factor-inference-in-population-spike-trains) Lea Duncker (/author/lea-duncker-12316), Maneesh Sahani (/author/maneesh-sahani-1670)
- Using Trusted Data to Train Deep Networks on Labels Corrupted by Severe Noise (/paper/8246-using-trusted-data-to-train-deep-networks-on-labels-corrupted-by-severe-noise) Dan Hendrycks (/author/dan-hendrycks-12317), Mantas Mazeika (/author/mantas-mazeika-12318), Duncan Wilson (/author/duncan-wilson-12319), Kevin Gimpel (/author/kevin-gimpel-4056)
- Discretely Relaxing Continuous Variables for tractable Variational Inference (/paper/8247-discretely-relaxing-continuous-variables-for-tractable-variational-inference) Trefor Evans (/author/trefor-evans-12320), Prasanth Nair (/author/prasanth-nair-12321)
- Regret bounds for meta Bayesian optimization with an unknown Gaussian process prior (/paper/8248-regret-bounds-for-meta-bayesian-optimization-with-an-unknown-gaussian-process-prior) Zi Wang (/author/zi-wang-6462), Beomjoon Kim (/author/beomjoon-kim-6357), Leslie Pack Kaelbling (/author/leslie-pack-kaelbling-6967)

- Diversity-Driven Exploration Strategy for Deep Reinforcement Learning (/paper/8249-diversity-driven-exploration-strategy-for-deep-reinforcement-learning) Zhang-Wei Hong (/author/zhang-wei-hong-12322), Tzu-Yun Shann (/author/tzu-yun-shann-12323), Shih-Yang Su (/author/shih-yang-su-12324), Yi-Hsiang Chang (/author/yi-hsiang-chang-12325), Tsu-Jui Fu (/author/tsu-jui-fu-12326), Chun-Yi Lee (/author/chun-yi-lee-12327)
- Deep Generative Models with Learnable Knowledge Constraints (/paper/8250-deep-generative-models-with-learnable-knowledge-constraints) Zhiting Hu (/author/zhiting-hu-9046), Zichao Yang (/author/zichao-yang-11805), Ruslan R. Salakhutdinov (/author/ruslan-r-salakhutdinov-2816), LIANHUI Qin (/author/lianhui-qin-12328), Xiaodan Liang (/author/xiaodan-liang-9217), Haoye Dong (/author/haoye-dong-10592), Eric P. Xing (/author/eric-p-xing-6314)
- The Sparse Manifold Transform (/paper/8251-the-sparse-manifold-transform) Yubei Chen (/author/yubei-chen-12329), Dylan Paiton (/author/dylan-paiton-12330), Bruno Olshausen (/author/bruno-olshausen-12331)
- Bayesian Structure Learning by Recursive Bootstrap (/paper/8252-bayesian-structure-learning-by-recursive-bootstrap) Raanan Y. Rohekar (/author/raanan-y-rohekar-11060), Yaniv Gurwicz (/author/yaniv-gurwicz-11062), Shami Nisimov (/author/shami-nisimov-11061), Guy Koren (/author/guy-koren-11063), Gal Novik (/author/gal-novik-11064)
- Complex Gated Recurrent Neural Networks (/paper/8253-complex-gated-recurrent-neural-networks) Moritz Wolter (/author/moritz-wolter-12332), Angela Yao (/author/angela-yao-5068)
- Learning a Warping Distance from Unlabeled Time Series Using Sequence Autoencoders (/paper/8254-learning-a-warping-distance-from-unlabeled-time-series-using-sequence-autoencoders) Abubakar Abid (/author/abubakar-abid-12333), James Y. Zou (/author/james-y-zou-6498)
- Streamlining Variational Inference for Constraint Satisfaction Problems (/paper/8255-streamlining-variational-inference-for-constraint-satisfaction-problems) Aditya Grover (/author/aditya-grover-8762), Tudor Achim (/author/tudor-achim-12334), Stefano Ermon (/author/stefano-ermon-5196)
- Fast deep reinforcement learning using online adjustments from the past (/paper/8256-fast-deep-reinforcement-learning-using-online-adjustments-from-the-past) Steven Hansen (/author/steven-hansen-12335), Alexander Pritzel (/author/alexander-pritzel-9158), Pablo Sprechmann (/author/pablo-sprechmann-6374), Andre Barreto (/author/andre-barreto-9700), Charles Blundell (/author/charles-blundell-4903)
- Improved Network Robustness with Adversary Critic (/paper/8257-improved-network-robustness-with-adversary-critic) Alexander Matyasko (/author/alexander-matyasko-12336), Lap-Pui Chau (/author/lap-pui-chau-12337)
- Regret Bounds for Online Portfolio Selection with a Cardinality Constraint (/paper/8258-regret-bounds-for-online-portfolio-selection-with-a-cardinality-constraint) Shinji Ito (/author/shinji-ito-8828), Daisuke Hatano (/author/daisuke-hatano-10014), Sumita Hanna (/author/sumita-hanna-12338), Akihiro Yabe (/author/akihiro-yabe-10010), Takuro

Fukunaga (/author/takuro-fukunaga-10013), Naonori Kakimura (/author/naonori-kakimura-10012), Ken-Ichi Kawarabayashi (/author/ken-ichi-kawarabayashi-10011)

- Sketching Method for Large Scale Combinatorial Inference (/paper/8259-sketching-method-for-large-scale-combinatorial-inference) Wei Sun (/author/wei-sun-8286), Junwei Lu (/author/junwei-lu-12339), Han Liu (/author/han-liu-3555)
- Connecting Optimization and Regularization Paths (/paper/8260-connecting-optimization-and-regularization-paths) Arun Suggala (/author/arun-suggala-10070), Adarsh Prasad (/author/adarsh-prasad-7267), Pradeep K. Ravikumar (/author/pradeep-k-ravikumar-6369)
- Fully Neural Network Based Speech Recognition on Mobile and Embedded Devices (/paper/8261-fully-neural-network-based-speech-recognition-on-mobile-and-embedded-devices) Jinhwan Park (/author/jinhwan-park-12340), Yoonho Boo (/author/yoonho-boo-10222), Iksoo Choi (/author/iksoo-choi-10221), Sungho Shin (/author/sungho-shin-12341), Wonyong Sung (/author/wonyong-sung-10223)
- Understanding Regularized Spectral Clustering via Graph Conductance (/paper/8262-understanding-regularized-spectral-clustering-via-graph-conductance) Yilin Zhang (/author/yilin-zhang-12342), Karl Rohe (/author/karl-rohe-6641)
- Data-Driven Clustering via Parameterized Lloyd's Families (/paper/8263-data-driven-clustering-via-parameterized-lloyds-families) Maria-Florina F. Balcan (/author/maria-florina-f-balcan-6635), Travis Dick (/author/travis-dick-12343), Colin White (/author/colin-white-12344)
- Learning Beam Search Policies via Imitation Learning (/paper/8264-learning-beam-search-policies-via-imitation-learning) Renato Negrinho (/author/renato-negrinho-7361), Matthew Gormley (/author/matthew-gormley-12345), Geoffrey J. Gordon (/author/geoffrey-j-gordon-1357)
- Benefits of over-parameterization with EM (/paper/8265-benefits-of-over-parameterization-with-em) Ji Xu (/author/ji-xu-8466), Daniel J. Hsu (/author/daniel-j-hsu-6464), Arian Maleki (/author/arian-maleki-8467)
- Thermostat-assisted continuously-tempered Hamiltonian Monte Carlo for Bayesian learning (/paper/8266-thermostat-assisted-continuously-tempered-hamiltonian-monte-carlo-for-bayesian-learning) Rui Luo (/author/rui-luo-12346), Jianhong Wang (/author/jianhong-wang-12347), Yaodong Yang (/author/yaodong-yang-12348), Jun WANG (/author/jun-wang-12349), Zhanxing Zhu (/author/zhanxing-zhu-8367)
- Robust Subspace Approximation in a Stream (/paper/8267-robust-subspace-approximation-in-a-stream) Roie Levin (/author/roie-levin-12350), Anish Prasad Sevekari (/author/anish-prasad-sevekari-12351), David Woodruff (/author/david-woodruff-6661)
- Mean Field for the Stochastic Blockmodel: Optimization Landscape and Convergence Issues (/paper/8268-mean-field-for-the-stochastic-blockmodel-optimization-landscape-and-convergence-issues) Soumendu Sundar Mukherjee (/author/soumendu-sundar-mukherjee-10427), Purnamrita Sarkar (/author/purnamrita-sarkar-6868), Y. X. Rachel Wang (/author/y-x-rachel-wang-12352), Bowei Yan (/author/bowei-yan-8988)

- Analysis of Krylov Subspace Solutions of Regularized Non-Convex Quadratic Problems (/paper/8269-analysis-of-krylov-subspace-solutions-of-regularized-non-convex-quadratic-problems) Yair Carmon (/author/yair-carmon-12353), John C. Duchi (/author/john-c-duchi-3453)
- Autoconj: Recognizing and Exploiting Conjugacy Without a Domain-Specific Language (/paper/8270-autoconj-recognizing-and-exploiting-conjugacy-without-a-domain-specific-language) Matthew D. Hoffman (/author/matthew-d-hoffman-12354)
- DropBlock: A regularization method for convolutional networks (/paper/8271-dropblock-a-regularization-method-for-convolutional-networks) Golnaz Ghiasi (/author/golnaz-ghiasi-12355), Tsung-Yi Lin (/author/tsung-yi-lin-12356), Quoc V. Le (/author/quoc-v-le-7287)
- Forward Modeling for Partial Observation Strategy Games - A StarCraft Defogger (/paper/8272-forward-modeling-for-partial-observation-strategy-games-a-starcraft-defogger) Gabriel Synnaeve (/author/gabriel-synnaeve-12357), Zeming Lin (/author/zeming-lin-12358), Jonas Gehring (/author/jonas-gehring-12359), Dan Gant (/author/dan-gant-12360), Vegard Mella (/author/vegard-mella-12361), Vasil Khalidov (/author/vasil-khalidov-12362), Nicolas Carion (/author/nicolas-carion-12363), Nicolas Usunier (/author/nicolas-usunier-3174)
- With Friends Like These, Who Needs Adversaries? (/paper/8273-with-friends-like-these-who-needs-adversaries) Saumya Jetley (/author/saumya-jetley-12364), Nicholas Lord (/author/nicholas-lord-12365), Philip Torr (/author/philip-torr-3805)
- Decentralize and Randomize: Faster Algorithm for Wasserstein Barycenters (/paper/8274-decentralize-and-randomize-faster-algorithm-for-wasserstein-barycenters) Pavel Dvurechenskii (/author/pavel-dvurechenskii-9276), Darina Dvinskikh (/author/darina-dvinskikh-12366), Alexander Gasnikov (/author/alexander-gasnikov-9277), Cesar Uribe (/author/cesar-uribe-12367), Angelia Nedich (/author/angelia-nedich-12368)
- Joint Autoregressive and Hierarchical Priors for Learned Image Compression (/paper/8275-joint-autoregressive-and-hierarchical-priors-for-learned-image-compression) David Minnen (/author/david-minnen-12369), Johannes Ballé (/author/johannes-balle-9915), George D. Toderici (/author/george-d-toderici-12370)
- Learning Temporal Point Processes via Reinforcement Learning (/paper/8276-learning-temporal-point-processes-via-reinforcement-learning) Shuang Li (/author/shuang-li-7967), Shuai Xiao (/author/shuai-xiao-9860), Shixiang Zhu (/author/shixiang-zhu-12371), Nan Du (/author/nan-du-5356), Yao Xie (/author/yao-xie-7968), Le Song (/author/le-song-3161)
- Bias and Generalization in Deep Generative Models: An Empirical Study (/paper/8277-bias-and-generalization-in-deep-generative-models-an-empirical-study) Shengjia Zhao (/author/shengjia-zhao-9146), Hongyu Ren (/author/hongyu-ren-11834), Arianna Yuan (/author/arianna-yuan-12372), Jiaming Song (/author/jiaming-song-9966), Noah Goodman (/author/noah-goodman-3514), Stefano Ermon (/author/stefano-ermon-5196)
- Fast and Effective Robustness Certification (/paper/8278-fast-and-effective-robustness-certification) Gagandeep Singh (/author/gagandeep-singh-12373), Timon Gehr

(/author/timon-gehr-12374), Matthew Mirman (/author/matthew-mirman-12375), Markus Püschel (/author/markus-puschel-12376), Martin Vechev (/author/martin-vechev-12289)

- Support Recovery for Orthogonal Matching Pursuit: Upper and Lower bounds (/paper/8279-support-recovery-for-orthogonal-matching-pursuit-upper-and-lower-bounds) Raghav Somani (/author/raghav-somani-12377), Chirag Gupta (/author/chirag-gupta-12378), Prateek Jain (/author/prateek-jain-3906), Praneeth Netrapalli (/author/praneeth-netrapalli-6558)
- Differentially Private Change-Point Detection (/paper/8280-differentially-private-change-point-detection) Rachel Cummings (/author/rachel-cummings-12051), Sara Krehbiel (/author/sara-krehbiel-12052), Yajun Mei (/author/yajun-mei-12380), Rui Tuo (/author/rui-tuo-12381), Wanrong Zhang (/author/wanrong-zhang-12379)
- Multi-value Rule Sets for Interpretable Classification with Feature-Efficient Representations (/paper/8281-multi-value-rule-sets-for-interpretable-classification-with-feature-efficient-representations) Tong Wang (/author/tong-wang-12382)
- Domain Adaptation by Using Causal Inference to Predict Invariant Conditional Distributions (/paper/8282-domain-adaptation-by-using-causal-inference-to-predict-invariant-conditional-distributions) Sara Magliacane (/author/sara-magliacane-8775), Thijs van Ommen (/author/thijs-van-ommen-12383), Tom Claassen (/author/tom-claassen-4570), Stephan Bongers (/author/stephan-bongers-12384), Philip Versteeg (/author/philip-versteeg-12385), Joris M. Mooij (/author/joris-m-mooij-3032)
- Smoothed Analysis of Discrete Tensor Decomposition and Assemblies of Neurons (/paper/8283-smoothed-analysis-of-discrete-tensor-decomposition-and-assemblies-of-neurons) Nima Anari (/author/nima-anari-12386), Constantinos Daskalakis (/author/constantinos-daskalakis-8166), Wolfgang Maass (/author/wolfgang-maass-618), Christos Papadimitriou (/author/christos-papadimitriou-12387), Amin Saberi (/author/amin-saberi-12388), Santosh Vempala (/author/santosh-vempala-8164)
- MixLasso: Generalized Mixed Regression via Convex Atomic-Norm Regularization (/paper/8284-mixlasso-generalized-mixed-regression-via-convex-atomic-norm-regularization) Ian En-Hsu Yen (/author/ian-en-hsu-yen-7160), Wei-Cheng Lee (/author/wei-cheng-lee-12389), Kai Zhong (/author/kai-zhong-7344), Sung-En Chang (/author/sung-en-chang-12390), Pradeep K. Ravikumar (/author/pradeep-k-ravikumar-6369), Shou-De Lin (/author/shou-de-lin-7476)
- Semidefinite relaxations for certifying robustness to adversarial examples (/paper/8285-semidefinite-relaxations-for-certifying-robustness-to-adversarial-examples) Aditi Raghunathan (/author/aditi-raghunathan-10365), Jacob Steinhardt (/author/jacob-steinhardt-7966), Percy S. Liang (/author/percy-s-liang-3559)
- Removing Hidden Confounding by Experimental Grounding (/paper/8286-removing-hidden-confounding-by-experimental-grounding) Nathan Kallus (/author/nathan-kallus-11734), Aahlad Manas Puli (/author/aahlad-manas-puli-12391), Uri Shalit (/author/uri-shalit-4258)

- Topkapi: Parallel and Fast Sketches for Finding Top-K Frequent Elements (/paper/8287-topkapi-parallel-and-fast-sketches-for-finding-top-k-frequent-elements) Ankush Mandal (/author/ankush-mandal-12392), He Jiang (/author/he-jiang-12393), Anshumali Shrivastava (/author/anshumali-shrivastava-5151), Vivek Sarkar (/author/vivek-sarkar-12394)
- Contrastive Learning from Pairwise Measurements (/paper/8288-contrastive-learning-from-pairwise-measurements) Yi Chen (/author/yi-chen-12395), Zhuoran Yang (/author/zhuoran-yang-8225), Yuchen Xie (/author/yuchen-xie-12396), Princeton Zhaoran Wang (/author/princeton-zhaoran-wang-10432)
- Point process latent variable models of larval zebrafish behavior (/paper/8289-point-process-latent-variable-models-of-larval-zebrafish-behavior) Anuj Sharma (/author/anuj-sharma-12397), Robert Johnson (/author/robert-johnson-12398), Florian Engert (/author/florian-engert-12399), Scott Linderman (/author/scott-linderman-7186)
- Computationally and statistically efficient learning of causal Bayes nets using path queries (/paper/8290-computationally-and-statistically-efficient-learning-of-causal-bayes-nets-using-path-queries) Kevin Bello (/author/kevin-bello-11080), Jean Honorio (/author/jean-honorio-4384)
- Sparse PCA from Sparse Linear Regression (/paper/8291-sparse-pca-from-sparse-linear-regression) Guy Bresler (/author/guy-bresler-7258), Sung Min Park (/author/sung-min-park-12400), Madalina Persu (/author/madalina-persu-12401)
- Multiple Instance Learning for Efficient Sequential Data Classification on Resource-constrained Devices (/paper/8292-multiple-instance-learning-for-efficient-sequential-data-classification-on-resource-constrained-devices) Don Dennis (/author/don-dennis-12402), Chirag Pabbaraju (/author/chirag-pabbaraju-12403), Harsha Vardhan Simhadri (/author/harsha-vardhan-simhadri-12404), Prateek Jain (/author/prateek-jain-3906)
- Transfer of Deep Reactive Policies for MDP Planning (/paper/8293-transfer-of-deep-reactive-policies-for-mdp-planning) Aniket (Nick) Bajpai (/author/aniket-nick-bajpai-12405), Sankalp Garg (/author/sankalp-garg-12406), None (/author/none-12407)
- The Price of Fair PCA: One Extra dimension (/paper/8294-the-price-of-fair-pca-one-extra-dimension) Samira Samadi (/author/samira-samadi-12408), Uthaipon Tantipongpipat (/author/uthaipon-tantipongpipat-12054), Jamie H. Morgenstern (/author/jamie-h-morgenstern-8075), Mohit Singh (/author/mohit-singh-12409), Santosh Vempala (/author/santosh-vempala-8164)
- GroupReduce: Block-Wise Low-Rank Approximation for Neural Language Model Shrinking (/paper/8295-groupreduce-block-wise-low-rank-approximation-for-neural-language-model-shrinking) Patrick Chen (/author/patrick-chen-12410), Si Si (/author/si-si-7479), Yang Li (/author/yang-li-12411), Ciprian Chelba (/author/ciprian-chelba-12412), Cho-Jui Hsieh (/author/cho-jui-hsieh-6366)

© 1987 – 2018 Neural Information Processing Systems Foundation, Inc.