

# **Advances in Neural Information Processing Systems 31**

32nd Conference on Neural Information  
Processing Systems (NeurIPS 2018)

Montreal, Canada  
3-8 December 2018

Volume 1 of 15

## **Editors:**

**Samy Bengio  
Hanna Wallach  
Hugo Larochelle**

**Kristen Grauman  
Nicolò Cesa-Bianchi  
Roman Garnett**

ISBN: 978-1-5108-8447-2

**Printed from e-media with permission by:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571



**Some format issues inherent in the e-media version may also appear in this print version.**

Copyright© (2018) by individual authors and NIPS  
All rights reserved.

Printed by Curran Associates, Inc. (2019)

For permission requests, please contact Neural Information Processing Systems (NIPS)  
at the address below.

Neural Information Processing Systems  
10010 North Torrey Pines Road  
La Jolla, CA 92037  
USA

Phone: (858) 453-1623

Fax: (858) 587-0417

[info@nips.cc](mailto:info@nips.cc)

**Additional copies of this publication are available from:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571 USA  
Phone: 845-758-0400  
Fax: 845-758-2633  
Email: [curran@proceedings.com](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

# Contents

Contents .....	iii
Preface .....	lix
Donors .....	lxiv
Neural Information Processing Systems Foundation .....	lxvii
Committees .....	lxix
Reviewers .....	lxxvi
<b>Efficient Algorithms for Non-convex Isotonic Regression through Submodular Optimization,</b> FRANCIS BACH .....	1
<b>Structure-Aware Convolutional Neural Networks,</b> JIANLONG CHANG, JIE GU, and LINGFENG WANG, GAOFENG MENG, SHIMING XIANG, CHUNHONG PAN .....	11
<b>Kalman Normalization: Normalizing Internal Representations Across Network Layers,</b> GUANGRUN WANG, JIEFENG PENG, and PING LUO, XINJIANG WANG, LIANG LIN .....	21
<b>HOGWILD!-Gibbs can be PanAccurate,</b> CONSTANTINOS DASKALAKIS, NISHANTH DIKKALA, and SIDDHARTHA JAYANTI .....	32
<b>Text-Adaptive Generative Adversarial Networks: Manipulating Images with Natural Language,</b> SEONGHYEON NAM, YUNJI KIM, and SEON KIM .....	42
<b>IntroVAE: Introspective Variational Autoencoders for Photographic Image Synthesis,</b> HUAIBO HUANG, ZHIHANG LI, and RAN HE, ZHENAN SUN, TIENIU TAN .....	52
<b>Doubly Robust Bayesian Inference for Non-Stationary Streaming Data with <math>\beta</math>-Divergences,</b> JEREMIAS KNOBLAUCH, JACK JEWSON, and THEODOROS DAMOULAS .....	64
<b>Adapted Deep Embeddings: A Synthesis of Methods for k-Shot Inductive Transfer Learning,</b> TYLER SCOTT, KARL RIDGEWAY, and MICHAEL MOZER .....	76
<b>Generalized Inverse Optimization through Online Learning,</b> CHAOSHENG DONG, YIRAN CHEN, and BO ZENG .....	86
<b>An Off-policy Policy Gradient Theorem Using Emphatic Weightings,</b> EHSAN IMANI, ERIC GRAVES, and MARTHA WHITE .....	96

<b>Supervised autoencoders: Improving generalization performance with unsupervised regularizers,</b> LEI LE, ANDREW PATTERSON, and MARTHA WHITE .....	107
<b>Visual Object Networks: Image Generation with Disentangled 3D Representations,</b> JUN-YAN ZHU, ZHOUTONG ZHANG, and CHENGKAI ZHANG, JIAJUN WU, ANTONIO TORRALBA, JOSHUA TENENBAUM, BILL FREEMAN .....	118
<b>Understanding Weight Normalized Deep Neural Networks with Rectified Linear Units,</b> YIXI XU, XIAO WANG .....	130
<b>Learning Pipelines with Limited Data and Domain Knowledge: A Study in Parsing Physics Problems,</b> MRINMAYA SACHAN, AVINAVA DUBEY, and TOM MITCHELL, DAN ROTH, ERIC XING .....	140
<b>Learning long-range spatial dependencies with horizontal gated recurrent units,</b> DREW LINSLEY, JUNKYUNG KIM, and VIJAY VEERABADRAN, CHARLES WINDOLF, THOMAS SERRE .....	152
<b>Joint Sub-bands Learning with Clique Structures for Wavelet Domain Super-Resolution,</b> ZHISHENG ZHONG, TIANCHENG SHEN, and YIBO YANG, ZHOUCHE LIN, CHAO ZHANG .....	165
<b>Fast Similarity Search via Optimal Sparse Lifting,</b> WENYE LI, JINGWEI MAO, and YIN ZHANG, SHUGUANG CUI .....	176
<b>Learning Deep Disentangled Embeddings With the F-Statistic Loss,</b> KARL RIDGEWAY, MICHAEL MOZER .....	185
<b>Geometrically Coupled Monte Carlo Sampling,</b> MARK ROWLAND, KRZYSZTOF CHOROMANSKI, and FRANÇOIS CHALUS, ALDO PACCHIANO, TAMAS SARLOS, RICHARD TURNER, ADRIAN WELLER .....	195
<b>Cooperative Holistic Scene Understanding: Unifying 3D Object, Layout, and Camera Pose Estimation,</b> SIYUAN HUANG, SIYUAN QI, and YINXUE XIAO, YIXIN ZHU, YING WU, SONG-CHUN ZHU .....	207
<b>An Efficient Pruning Algorithm for Robust Isotonic Regression,</b> CONG LIM .....	219
<b>PAC-learning in the presence of evasion adversaries,</b> DANIEL CULLINA, ARJUN BHAGOJI, and PRATEEK MITTAL .....	230
<b>Sparse DNNs with Improved Adversarial Robustness,</b> YIWEN GUO, CHAO ZHANG, and CHANGSHUI ZHANG, YURONG CHEN .....	242
<b>Snap ML: A Hierarchical Framework for Machine Learning,</b> CELESTINE DUENNER, THOMAS PARNELL, and DIMITRIOS SARIGIANNIS, NIKOLAS IOANNOU, ANDREEA ANGHEL, GUMMADI RAVI, MADHUSUDANAN KANDASAMY, CHARALAMPOS POZIDIS .....	252
<b>See and Think: Disentangling Semantic Scene Completion,</b> SHICE LIU, YU HU, and YIMING ZENG, QIAN KUN TANG, BEIBEI JIN, YINHE HAN, XIAOWEI LI .....	263
<b>Chain of Reasoning for Visual Question Answering,</b> CHENFEI WU, JINLAI LIU, and XIAOJIE WANG, XUAN DONG .....	275
<b>Sigsoftmax: Reanalysis of the Softmax Bottleneck,</b> SEKITOSHI KANAI, YASUHIRO FUJIWARA, and YUKI YAMANAKA, SHUICHI ADACHI .....	286

<b>Deep Non-Blind Deconvolution via Generalized Low-Rank Approximation,</b> WENQI REN, JIAWEI ZHANG, and LIN MA, JINSHAN PAN, XIAOCHUN CAO, WANGMENG ZUO, WEI LIU, MING-HSUAN YANG .....	297
<b>Bayesian Pose Graph Optimization via Bingham Distributions and Tempered Geodesic MCMC,</b> TOLGA BIRDAL, UMUT SIMSEKLI, and MUSTAFA EKEN, SLOBODAN ILIC .....	308
<b>MetaAnchor: Learning to Detect Objects with Customized Anchors,</b> TONG YANG, XIANGYU ZHANG, and ZEMING LI, WENQIANG ZHANG, JIAN SUN .....	320
<b>Image Inpainting via Generative Multi-column Convolutional Neural Networks,</b> YI WANG, XIN TAO, and XIAOJUAN QI, XIAOYONG SHEN, JIAYA JIA .....	331
<b>On Misinformation Containment in Online Social Networks,</b> GUANGMO TONG, DING-ZHU DU, and WEILI WU .....	341
<b>A<sup>2</sup>-Nets: Double Attention Networks,</b> YUNPENG CHEN, YANNIS KALANTIDIS, and JIANSU LI, YAN SHUICHENG, JIASHI FENG .....	352
<b>Self-Supervised Generation of Spatial Audio for 360° Video,</b> PEDRO MORGADO, NUNO VASCONCELOS, and TIMOTHY LANGLOIS, OLIVER WANG .....	362
<b>How Many Samples are Needed to Estimate a Convolutional Neural Network?,</b> SIMON DU, YINING WANG, and XIYU ZHAI, SIVARAMAN BALAKRISHNAN, RUSLAN SALAKHUTDINOV, AARTI SINGH .....	373
<b>Algorithmic Regularization in Learning Deep Homogeneous Models: Layers are Automatically Balanced,</b> SIMON DU, WEI HU, and JASON LEE .....	384
<b>Optimization for Approximate Submodularity,</b> YARON SINGER, AVINATAN HASSIDIM .....	396
<b>(Probably) Concave Graph Matching,</b> HAGGAI MARON, YARON LIPMAN .....	408
<b>Deep Defense: Training DNNs with Improved Adversarial Robustness,</b> ZIANG YAN, YIWEN GUO, and CHANGSHUI ZHANG .....	419
<b>Rest-Katyusha: Exploiting the Solution’s Structure via Scheduled Restart Schemes,</b> JUNQI TANG, MOHAMMAD GOLBABAEE, and FRANCIS BACH, MIKE DAVIES .....	429
<b>Implicit Reparameterization Gradients,</b> MICHAEL FIGURNOV, SHAKIR MOHAMED, and ANDRIY MNIH .....	441
<b>Training DNNs with Hybrid Block Floating Point,</b> MARIO DRUMOND, TAO LIN, and MARTIN JAGGI, BABAK FALSAFI .....	453
<b>A Model for Learned Bloom Filters and Optimizing by Sandwiching,</b> MICHAEL MITZENMACHER .....	464
<b>Soft-Gated Warping-GAN for Pose-Guided Person Image Synthesis,</b> HAQYE DONG, XIAODAN LIANG, and KE GONG, HANJIANG LAI, JIA ZHU, JIAN YIN .....	474
<b>Deep Functional Dictionaries: Learning Consistent Semantic Structures on 3D Models from Functions,</b> MINHUYK SUNG, HAO SU, and RONALD YU, LEONIDAS GUIBAS .....	485

<b>Nonlocal Neural Networks, Nonlocal Diffusion and Nonlocal Modeling,</b> YUNZHE TAO, QI SUN, and QIANG DU, WEI LIU .....	496
<b>Are ResNets Provably Better than Linear Predictors?,</b> OHAD SHAMIR .....	507
<b>Learning to Decompose and Disentangle Representations for Video Prediction,</b> JUN-TING HSIEH, BINGBIN LIU, and DE-AN HUANG, LI FEI-FEI, JUAN NIEBLES .....	517
<b>Multi-Task Learning as Multi-Objective Optimization,</b> OZAN SENER, VLADLEN KOLTUN .....	527
<b>Combinatorial Optimization with Graph Convolutional Networks and Guided Tree Search,</b> ZHUWEN LI, QIFENG CHEN, and VLADLEN KOLTUN .....	539
<b>Self-Erasing Network for Integral Object Attention,</b> QIBIN HOU, PENG-TAO JIANG, and YUNCHAO WEI, MING-MING CHENG .....	549
<b>LinkNet: Relational Embedding for Scene Graph,</b> SANGHYUN WOO, DAHUN KIM, and DONGHYEON CHO, IN KWEON .....	560
<b>How to Start Training: The Effect of Initialization and Architecture,</b> BORIS HANIN, DAVID ROLNICK .....	571
<b>Which Neural Net Architectures Give Rise to Exploding and Vanishing Gradients?,</b> BORIS HANIN .....	582
<b>Explanations based on the Missing: Towards Contrastive Explanations with Pertinent Negatives,</b> AMIT DHURANDHAR, PIN-YU CHEN, and RONNY LUSS, CHUN-CHEN TU, PAI-SHUN TING, KARTHIKEYAN SHANMUGAM, PAYEL DAS .....	592
<b>HitNet: Hybrid Ternary Recurrent Neural Network,</b> PEIQI WANG, XINFENG XIE, and LEI DENG, GUOQI LI, DONGSHENG WANG, YUAN XIE .....	604
<b>A Unified Framework for Extensive-Form Game Abstraction with Bounds,</b> CHRISTIAN KROER, TUOMAS SANDHOLM .....	615
<b>Removing the Feature Correlation Effect of Multiplicative Noise,</b> ZIJUN ZHANG, YINING ZHANG, and ZONGPENG LI .....	627
<b>Maximum-Entropy Fine Grained Classification,</b> ABHIMANYU DUBEY, OTKRIST GUPTA, and RAMESH RASKAR, NIKHIL NAIK .....	637
<b>On Learning Markov Chains,</b> YI HAO, ALON ORLITSKY, and VENKATADHEERAJ PICHAPATI .....	648
<b>A Neural Compositional Paradigm for Image Captioning,</b> BO DAI, SANJA FIDLER, and DAHUA LIN .....	658
<b>Quantifying Learning Guarantees for Convex but Inconsistent Surrogates,</b> KIRILL STRUMINSKY, SIMON LACOSTE-JULIEN, and ANTON OSOKIN .....	669
<b>Dialog-based Interactive Image Retrieval,</b> XIAOXIAO GUO, HUI WU, and YU CHENG, STEVEN RENNIE, GERALD TESAURO, ROGERIO FERIS .....	678
<b>SPIDER: Near-Optimal Non-Convex Optimization via Stochastic Path-Integrated Differential Estimator,</b> CONG FANG, CHRIS LI, and ZHOUCHE LIN, TONG ZHANG .....	689

<b>Are GANs Created Equal? A Large-Scale Study,</b> MARIO LUCIC, KAROL KURACH, and MARCIN MICHALSKI, SYLVAIN GELLY, OLIVIER BOUSQUET .....	700
<b>Learning Disentangled Joint Continuous and Discrete Representations,</b> EMILIEN DUPONT .....	710
<b>TADAM: Task dependent adaptive metric for improved few-shot learning,</b> BORIS ORESHKIN, PAU RODRIGUEZ, and ALEXANDRE LACOSTE ...	721
<b>Do Less, Get More: Streaming Submodular Maximization with Subsampling,</b> MORAN FELDMAN, AMIN KARBASI, and EHSAN KAZEMI .....	732
<b>Deep Neural Nets with Interpolating Function as Output Activation,</b> BAO WANG, XIYANG LUO, and ZHEN LI, WEI ZHU, ZUOQIANG SHI, STANLEY OSHER .....	743
<b>FishNet: A Versatile Backbone for Image, Region, and Pixel Level Prediction,</b> SHUYANG SUN, JIANGMIAO PANG, and JIANPING SHI, SHUAI YI, WANLI OUYANG .....	754
<b>Visual Memory for Robust Path Following,</b> ASHISH KUMAR, SAURABH GUPTA, and DAVID FOUHEY, SERGEY LEVINE, JITENDRA MALIK .....	765
<b>KDGAN: Knowledge Distillation with Generative Adversarial Networks,</b> XIAOJIE WANG, RUI ZHANG, and YU SUN, JIANZHONG QI .....	775
<b>Long short-term memory and Learning-to-learn in networks of spiking neurons,</b> GUILLAUME BELLEC, DARJAN SALAJ, and ANAND SUBRAMONEY, ROBERT LEGENSTEIN, WOLFGANG MAASS .....	787
<b>Greedy Hash: Towards Fast Optimization for Accurate Hash Coding in CNN,</b> SHUPENG SU, CHAO ZHANG, and KAI HAN, YONGHONG TIAN .....	798
<b>Informative Features for Model Comparison,</b> WITTAWAT JITKRITUM, HEISHIRO KANAGAWA, and PATSORN SANGKLOY, JAMES HAYS, BERNHARD SCHÖLKOPF, ARTHUR GRETTON .....	808
<b>PointCNN: Convolution On X-Transformed Points,</b> YANGYAN LI, RUI BU, and MINGCHAO SUN, WEI WU, XINHAN DI, BAOQUAN CHEN .....	820
<b>Connectionist Temporal Classification with Maximum Entropy Regularization,</b> HU LIU, SHENG JIN, and CHANGSHUI ZHANG .....	831
<b>Large Margin Deep Networks for Classification,</b> GAMALELDIN ELSAYED, DILIP KRISHNAN, and HOSSEIN MOBAHI, KEVIN REGAN, SAMY BENGIO .....	842
<b>Generalizing Graph Matching beyond Quadratic Assignment Model,</b> TIANSHU YU, JUNCHI YAN, and YILIN WANG, WEI LIU, BAOXIN LI ...	853
<b>Solving Large Sequential Games with the Excessive Gap Technique,</b> CHRISTIAN KROER, GABRIELE FARINA, and TUOMAS SANDHOLM .....	864
<b>Discrimination-aware Channel Pruning for Deep Neural Networks,</b> ZHUANGWEI ZHUANG, MINGKUI TAN, and BOHAN ZHUANG, JING LIU, YONG GUO, QINGYAO WU, JUNZHOU HUANG, JINHUI ZHU .....	875
<b>On the Dimensionality of Word Embedding,</b> ZI YIN, YUANYUAN SHEN	887
<b>Reinforced Continual Learning,</b> JU XU, ZHANXING ZHU .....	899

<b>Uncertainty-Aware Attention for Reliable Interpretation and Prediction,</b> JAY HEO, HAE LEE, and SAEHOON KIM, JUHO LEE, KWANG KIM, EUNHO YANG, SUNG HWANG .....	909
<b>DropMax: Adaptive Variational Softmax,</b> HAE LEE, JUHO LEE, and SAEHOON KIM, EUNHO YANG, SUNG HWANG .....	919
<b>Posterior Concentration for Sparse Deep Learning,</b> VERONIKA ROCKOVA, NICHOLAS POLSON .....	930
<b>A flexible model for training action localization with varying levels of supervision,</b> GUILHEM CHERON, JEAN-BAPTISTE ALAYRAC, and IVAN LAPTEV, CORDELIA SCHMID .....	942
<b>A Deep Bayesian Policy Reuse Approach Against Non-Stationary Agents,</b> YAN ZHENG, ZHAOPENG MENG, and JIANYE HAO, ZONGZHANG ZHANG, TIANPEI YANG, CHANGJIE FAN .....	954
<b>Empirical Risk Minimization in Non-interactive Local Differential Privacy Revisited,</b> DI WANG, MARCO GABOARDI, and JINHUI XU .....	965
<b>Low-shot Learning via Covariance-Preserving Adversarial Augmentation Networks,</b> HANG GAO, ZHENG SHOU, and ALIREZA ZAREIAN, HANWANG ZHANG, SHIH-FU CHANG .....	975
<b>Learning semantic similarity in a continuous space,</b> MICHEL DEUDON	986
<b>MetaReg: Towards Domain Generalization using Meta-Regularization,</b> YOGESH BALAJI, SWAMI SANKARANARAYANAN, and RAMA CHELLAPPA .....	998
<b>Boosted Sparse and Low-Rank Tensor Regression,</b> LIFANG HE, KUN CHEN, and WANWAN XU, JIAYU ZHOU, FEI WANG .....	1009
<b>Domain-Invariant Projection Learning for Zero-Shot Recognition,</b> AN ZHAO, MINGYU DING, and JIECHAO GUAN, ZHIWU LU, TAO XIANG, JI-RONG WEN .....	1019
<b>Neural-Symbolic VQA: Disentangling Reasoning from Vision and Language Understanding,</b> KEXIN YI, JIAJUN WU, and CHUANG GAN, ANTONIO TORRALBA, PUSHMEET KOHLI, JOSHUA TENENBAUM .....	1031
<b>Frequency-Domain Dynamic Pruning for Convolutional Neural Networks,</b> ZHENHUA LIU, JIZHENG XU, and XIULIAN PENG, RUIQIN XIONG	1043
<b>Quadratic Decomposable Submodular Function Minimization,</b> PAN LI, NIAO HE, and OLGICA MILENKOVIC .....	1054
<b>A Block Coordinate Ascent Algorithm for Mean-Variance Optimization,</b> TENG YANG XIE, BO LIU, and YANG YANG XU, MOHAMMAD GHAVAMZADEH, YINLAM CHOW, DAOMING LYU, DAESUB YOON .....	1065
<b><math>\ell_1</math>-regression with Heavy-tailed Distributions,</b> LIJUN ZHANG, ZHI-HUA ZHOU .....	1076
<b>Neural Nearest Neighbors Networks,</b> TOBIAS PLOETZ, STEFAN ROTH	1087
<b>Efficient nonmyopic batch active search,</b> SHALI JIANG, GUSTAVO MALKOMES, and MATTHEW ABBOTT, BENJAMIN MOSELEY, ROMAN GARNETT	1099
<b>A Game-Theoretic Approach to Recommendation Systems with Strategic Content Providers,</b> OMER BEN-PORAT, MOSHE TENNENHOLTZ	1110



<b>Interactive Structure Learning with Structural Query-by-Committee</b> , CHRISTOPHER TOSH, SANJOY DASGUPTA .....	1121
<b>Global Geometry of Multichannel Sparse Blind Deconvolution on the Sphere</b> , YANJUN LI, YORAM BRESLER .....	1132
<b>Video-to-Video Synthesis</b> , TING-CHUN WANG, MING-YU LIU, and JUN-YAN ZHU, GUILIN LIU, ANDREW TAO, JAN KAUTZ, BRYAN CATANZARO ..	1144
<b>How To Make the Gradients Small Stochastically: Even Faster Convex and Nonconvex SGD</b> , ZEYUAN ALLEN-ZHU .....	1157
<b>Synthesized Policies for Transfer and Adaptation across Tasks and Environments</b> , HEXIANG HU, LIYU CHEN, and BOQING GONG, FEI SHA ...	1168
<b>Adversarial vulnerability for any classifier</b> , ALHUSSEIN FAWZI, HAMZA FAWZI, and OMAR FAWZI .....	1178
<b>Evolution-Guided Policy Gradient in Reinforcement Learning</b> , SHAUHARDA KHADKA <sup>4</sup> , KAGAN TUMER .....	1188
<b>Toddler-Inspired Visual Object Learning</b> , SVEN BAMBACH, DAVID CRANDALL, and LINDA SMITH, CHEN YU .....	1201
<b>Alternating optimization of decision trees, with application to learning sparse oblique trees</b> , MIGUEL CARREIRA-PERPINAN, POOYA TAVALLALI .....	1211
<b>FD-GAN: Pose-guided Feature Distilling GAN for Robust Person Re-identification</b> , YIXIAO GE, ZHUOWAN LI, and HAIYU ZHAO, GUOJUN YIN, SHUAI YI, XIAOGANG WANG, HONGSHENG LI .....	1222
<b>New Insight into Hybrid Stochastic Gradient Descent: Beyond With-Replacement Sampling and Convexity</b> , PAN ZHOU, XIAOTONG YUAN, and JIASHI FENG .....	1234
<b>The Lingering of Gradients: How to Reuse Gradients over Time</b> , ZEYUAN ALLEN-ZHU, DAVID SIMCHI-LEVI, and XINSHANG WANG .....	1244
<b>Unsupervised Learning of View-invariant Action Representations</b> , JUNNAN LI, WONG YONGKANG, and QI ZHAO, MOHAN KANKANHALLI .....	1254
<b>Fairness Behind a Veil of Ignorance: A Welfare Analysis for Automated Decision Making</b> , HODA HEIDARI, CLAUDIO FERRARI, and KRISHNA GUMMADI, ANDREAS KRAUSE .....	1265
<b>Global Gated Mixture of Second-order Pooling for Improving Deep Convolutional Neural Networks</b> , QILONG WANG, ZILIN GAO, and JIANGTAO XIE, WANGMENG ZUO, PEIHUA LI .....	1277
<b>Image-to-image translation for cross-domain disentanglement</b> , ABEL GONZALEZ-GARCIA, JOOST WEIJER, and YOSHUA BENGIO .....	1287
<b>Gradient Sparsification for Communication-Efficient Distributed Optimization</b> , JIANQIAO WANGNI, JIALEI WANG, and JI LIU, TONG ZHANG	1299
<b>Revisiting Multi-Task Learning with ROCK: a Deep Residual Auxiliary Block for Visual Detection</b> , TAYLOR MORDAN, NICOLAS THOME, and GILLES HENAFF, MATTHIEU CORD .....	1310
<b>Adaptive Online Learning in Dynamic Environments</b> , LIJUN ZHANG, SHIYIN LU, and ZHI-HUA ZHOU .....	1323

<b>FRAGE: Frequency-Agnostic Word Representation,</b> CHENGYUE GONG, DI HE, and XU TAN, TAO QIN, LIWEI WANG, TIEYAN LIU .....	1334
<b>Generative Neural Machine Translation,</b> HARSHIL SHAH, DAVID BARBER .....	1346
<b>Found Graph Data and Planted Vertex Covers,</b> AUSTIN BENSON, JON KLEINBERG .....	1356
<b>Joint Active Feature Acquisition and Classification with Variable-Size Set Encoding,</b> HAJIN SHIM, SUNG HWANG, and EUNHO YANG .....	1368
<b>Regularization Learning Networks: Deep Learning for Tabular Datasets,</b> IRA SHAVITT, ERAN SEGAL .....	1379
<b>Multitask Boosting for Survival Analysis with Competing Risks,</b> ALEXIS BELLOT, MIHAELA SCHAAR .....	1390
<b>Geometry Based Data Generation,</b> OFIR LINDENBAUM, JAY STANLEY, and GUY WOLF, SMITA KRISHNASWAMY .....	1400
<b>SLAYER: Spike Layer Error Reassignment in Time,</b> SUMIT SHRESTHA, GARRICK ORCHARD .....	1412
<b>On Oracle-Efficient PAC RL with Rich Observations,</b> CHRISTOPH DANN, NAN JIANG, and AKSHAY KRISHNAMURTHY, ALEKH AGARWAL, JOHN LANGFORD, ROBERT SCHAPIRE .....	1422
<b>Gradient Descent for Spiking Neural Networks,</b> DONGSUNG HUH, TERRY SEJNOWSKI .....	1433
<b>Generalizing Tree Probability Estimation via Bayesian Networks,</b> CHENG ZHANG, FREDERICK IV .....	1444
<b>Where Do You Think You're Going?: Inferring Beliefs about Dynamics from Behavior,</b> SIDDHARTH REDDY, ANCA DRAGAN, and SERGEY LEVINE .....	1454
<b>Designing by Training: Acceleration Neural Network for Fast High-Dimensional Convolution,</b> LONGQUAN DAI, LIANG TANG, and YUAN XIE, JINHUI TANG .....	1466
<b>Understanding the Role of Adaptivity in Machine Teaching: The Case of Version Space Learners,</b> YUXIN CHEN, ADISH SINGLA, and OISIN AODHA, PIETRO PERONA, YISONG YUE .....	1476
<b>A loss framework for calibrated anomaly detection,</b> ADITYA MENON, ROBERT WILLIAMSON .....	1487
<b>PacGAN: The power of two samples in generative adversarial networks,</b> ZINAN LIN, ASHISH KHETAN, and GIULIA FANTI, SEWOONG OH .	1498
<b>Variational Memory Encoder-Decoder,</b> HUNG LE, TRUYEN TRAN, and THIN NGUYEN, SVETHA VENKATESH .....	1508
<b>Stochastic Composite Mirror Descent: Optimal Bounds with High Probabilities,</b> YUNWEN LEI, KE TANG .....	1519
<b>Hybrid Retrieval-Generation Reinforced Agent for Medical Image Report Generation,</b> YUAN LI, XIAODAN LIANG, and ZHITING HU, ERIC XING .....	1530

<b>Overcoming Language Priors in Visual Question Answering with Adversarial Regularization,</b> SAINANDAN RAMAKRISHNAN, AISHWARYA AGRAWAL, and STEFAN LEE .....	1541
<b>Hybrid Knowledge Routed Modules for Large-scale Object Detection,</b> CHENHAN JIANG, HANG XU, and XIAODAN LIANG, LIANG LIN ..	1552
<b>Bilinear Attention Networks,</b> JIN-HWA KIM, JAEHYUN JUN, and BYOUNG-TAK ZHANG .....	1564
<b>Parsimonious Quantile Regression of Financial Asset Tail Dynamics via Sequential Learning,</b> XING YAN, WEIZHONG ZHANG, and LIN MA, WEI LIU, QI WU .....	1575
<b>Multi-Class Learning: From Theory to Algorithm,</b> JIAN LI, YONG LIU, and RONG YIN, HUA ZHANG, LIZHONG DING, WEIPING WANG .....	1586
<b>Multivariate Time Series Imputation with Generative Adversarial Networks,</b> YONGHONG LUO, XIANGRUI CAI, and YING ZHANG, JUN XU, YUAN XIAOJIE .....	1596
<b>Learning Versatile Filters for Efficient Convolutional Neural Networks,</b> YUNHE WANG, CHANG XU, and CHUNJING XU, CHAO XU, DACHENG TAO .....	1608
<b>Accelerated Stochastic Matrix Inversion: General Theory and Speeding up BFGS Rules for Faster Second-Order Optimization,</b> ROBERT GOWER, FILIP HANZELY, and PETER RICHTARIK, SEBASTIAN STICH	1619
<b>DifNet: Semantic Segmentation by Diffusion Networks,</b> PENG JIANG, FANGLIN GU, and YUNHAI WANG, CHANGHE TU, BAOQUAN CHEN ....	1630
<b>Conditional Adversarial Domain Adaptation,</b> MINGSHENG LONG, ZHANGJIE CAO, and JIANMIN WANG, MICHAEL JORDAN .....	1640
<b>Neighbourhood Consensus Networks,</b> IGNACIO ROCCO, MIRCEA CIMPOI, and RELJA ARANDJELOVIĆ, AKIHIKO TORII, TOMAS PAJDLA, JOSEF SIVIC .....	1651
<b>Relating Leverage Scores and Density using Regularized Christoffel Functions,</b> EDOUARD PAUWELS, FRANCIS BACH, and JEAN-PHILIPPE VERT .....	1663
<b>Non-Local Recurrent Network for Image Restoration,</b> DING LIU, BIHAN WEN, and YUCHEN FAN, CHEN LOY, THOMAS HUANG .....	1673
<b>Bayesian Semi-supervised Learning with Graph Gaussian Processes,</b> YIN NG, NICOLO COLOMBO, and RICARDO SILVA .....	1683
<b>Foreground Clustering for Joint Segmentation and Localization in Videos and Images,</b> ABHISHEK SHARMA .....	1695
<b>Video Prediction via Selective Sampling,</b> JINGWEI XU, BINGBING NI, and XIAOKANG YANG .....	1705
<b>Distilled Wasserstein Learning for Word Embedding and Topic Modeling,</b> HONGTENG XU, WENLIN WANG, and WEI LIU, LAWRENCE DUKE	1716
<b>Learning to Exploit Stability for 3D Scene Parsing,</b> YILUN DU, ZHIJIAN LIU, and HECTOR BASEVI, ALES LEONARDIS, BILL FREEMAN, JOSHUA TENENBAUM, JIAJUN WU .....	1726

<b>Neural Guided Constraint Logic Programming for Program Synthesis,</b> LISA ZHANG, GREGORY ROSENBLATT, and ETHAN FETAYA, RENJIE LIAO, WILLIAM BYRD, MATTHEW MIGHT, RAQUEL URTASUN, RICHARD ZEMEL .....	1737
<b>Genetic-Gated Networks for Deep Reinforcement Learning,</b> SIMYUNG CHANG, JOHN YANG, and JAESEOK CHOI, NOJUN KWAK .....	1747
<b>Fighting Boredom in Recommender Systems with Linear Reinforcement Learning,</b> ROMAIN WARLOP, ALESSANDRO LAZARIC, and JÉRÉMIE MARY .....	1757
<b>Enhancing the Accuracy and Fairness of Human Decision Making,</b> ISABEL VALERA, ADISH SINGLA, and MANUEL RODRIGUEZ .....	1769
<b>Temporal Regularization for Markov Decision Process,</b> PIERRE THODOROFF, AUDREY DURAND, and JOELLE PINEAU, DOINA PRECUP .....	1779
<b>The Pessimistic Limits and Possibilities of Margin-based Losses in Semi-supervised Learning,</b> JESSE KRIJTJE, MARCO LOOG .....	1790
<b>Simple random search of static linear policies is competitive for reinforcement learning,</b> HORIA MANIA, AURELIA GUY, and BENJAMIN RECHT .....	1800
<b>Generating Informative and Diverse Conversational Responses via Adversarial Information Maximization,</b> YIZHE ZHANG, MICHEL GALLEY, and JIANFENG GAO, ZHE GAN, XIUJUN LI, CHRIS BROCKETT, BILL DOLAN .....	1810
<b>Entropy and mutual information in models of deep neural networks,</b> MARYLOU GABRIÉ, ANDRE MANOEL, and CLÉMENT LUNEAU, JEAN BARBIER, NICOLAS MACRIS, FLORENT KRZAKALA, LENKA ZDEBOROVA .....	1821
<b>Collaborative Learning for Deep Neural Networks,</b> GUOCONG SONG, WEI CHAI .....	1832
<b>High Dimensional Linear Regression using Lattice Basis Reduction,</b> ILIAS ZADIK, DAVID GAMARNIK .....	1842
<b>Symbolic Graph Reasoning Meets Convolutions,</b> XIAODAN LIANG, ZHITING HU, and HAO ZHANG, LIANG LIN, ERIC XING .....	1853
<b>DVAE#: Discrete Variational Autoencoders with Relaxed Boltzmann Priors,</b> ARASH VAHDAT, EVGENY ANDRIYASH, and WILLIAM MACREADY .....	1864
<b>Partially-Supervised Image Captioning,</b> PETER ANDERSON, STEPHEN GOULD, and MARK JOHNSON .....	1875
<b>3D-Aware Scene Manipulation via Inverse Graphics,</b> SHUNYU YAO, TZU-MING HSU, and JUN-YAN ZHU, JIAJUN WU, ANTONIO TORRALBA, BILL FREEMAN, JOSHUA TENENBAUM .....	1887
<b>Random Feature Stein Discrepancies,</b> JONATHAN HUGGINS, LESTER MACKAY .....	1899
<b>Distributed Stochastic Optimization via Adaptive SGD,</b> ASHOK CUTKOSKY, ROBERT BUSA-FEKETE .....	1910
<b>Precision and Recall for Time Series,</b> NESIME TATBUL, TAE LEE, and STAN ZDONIK, MEJBAH ALAM, JUSTIN GOTTSCHLICH .....	1920

<b>Deep Attentive Tracking via Reciprocative Learning,</b> SHI PU, YIBING SONG, and CHAO MA, HONGGANG ZHANG, MING-HSUAN YANG .....	1931
<b>Virtual Class Enhanced Discriminative Embedding Learning,</b> BINGHUI CHEN, WEIHONG DENG, and HAIFENG SHEN .....	1942
<b>Attention in Convolutional LSTM for Gesture Recognition,</b> LIANG ZHANG, GUANGMING ZHU, and LIN MEI, PEIYI SHEN, SYED SHAH, MOHAMMED BENNAMOUN .....	1953
<b>Pelee: A Real-Time Object Detection System on Mobile Devices,</b> ROBERT WANG, XIANG LI, and CHARLES LING .....	1963
<b>Universal Growth in Production Economies,</b> SIMINA BRANZEI, RUTA MEHTA, and NOAM NISAN .....	1973
<b>Bayesian Model Selection Approach to Boundary Detection with Non-Local Priors,</b> FEI JIANG, GUOSHENG YIN, and FRANCESCA DOMINICI	1974
<b>Efficient Stochastic Gradient Hard Thresholding,</b> PAN ZHOU, XIAOTONG YUAN, and JIASHI FENG .....	1984
<b>SplineNets: Continuous Neural Decision Graphs,</b> CEM KESKIN, SHAHRAM IZADI .....	1994
<b>Generalized Zero-Shot Learning with Deep Calibration Network,</b> SHICHEN LIU, MINGSHENG LONG, and JIANMIN WANG, MICHAEL JORDAN ...	2005
<b>Neural Architecture Search with Bayesian Optimisation and Optimal Transport,</b> KIRTHEVASAN KANDASAMY, WILLIE NEISWANGER, and JEFF SCHNEIDER, BARNABAS POZOS, ERIC XING .....	2016
<b>Embedding Logical Queries on Knowledge Graphs,</b> WILLIAM HAMILTON, PAYAL BAJAJ, and MARINKA ZITNIK, DAN JURAFSKY, JURE LESKOVEC .....	2026
<b>Learning Optimal Reserve Price against Non-myopic Bidders,</b> JINYAN LIU, ZHIYI HUANG, and XIANGNING WANG .....	2038
<b>Sequential Context Encoding for Duplicate Removal,</b> LU QI, SHU LIU, and JIANPING SHI, JIAYA JIA .....	2049
<b>Discovery of Latent 3D Keypoints via End-to-end Geometric Reasoning,</b> SUPASORN SUWAJANAKORN, NOAH SNAVELY, and JONATHAN TOMPSON, MOHAMMAD NOROUZI .....	2059
<b>Nonparametric learning from Bayesian models with randomized objective functions,</b> SIMON LYDDON, STEPHEN WALKER, and CHRIS HOLMES .....	2071
<b>SEGA: Variance Reduction via Gradient Sketching,</b> FILIP HANZELY, KONSTANTIN MISHCHENKO, and PETER RICHTARIK .....	2082
<b>Automatic Program Synthesis of Long Programs with a Learned Garbage Collector,</b> AMIT ZOHAR, LIOR WOLF .....	2094
<b>One-Shot Unsupervised Cross Domain Translation,</b> SAGIE BENAIM, LIOR WOLF .....	2104
<b>Regularizing by the Variance of the Activations' Sample-Variiances,</b> ETAI LITWIN, LIOR WOLF .....	2115

<b>Overlapping Clustering Models, and One (class) SVM to Bind Them All</b> , XUEYU MAO, PURNAMRITA SARKAR, and DEEPAYAN CHAKRABARTI .....	2126
<b>Algorithmic Linearly Constrained Gaussian Processes</b> , MARKUS LANGE-HEGERMANN .....	2137
<b>DeepExposure: Learning to Expose Photos with Asynchronously Reinforced Adversarial Learning</b> , RUNSHENG YU, WENYU LIU, and YASEN ZHANG, ZHI QU, DELI ZHAO, BO ZHANG .....	2149
<b>Norm matters: efficient and accurate normalization schemes in deep networks</b> , ELAD HOFFER, RON BANNER, and ITAY GOLAN, DANIEL SOUDRY .....	2160
<b>Dual Principal Component Pursuit: Improved Analysis and Efficient Algorithms</b> , ZHIHUI ZHU, YIFAN WANG, and DANIEL ROBINSON, DANIEL NAIMAN, RENE VIDAL, MANOLIS TSAKIRIS .....	2171
<b>MULAN: A Blind and Off-Grid Method for Multichannel Echo Retrieval</b> , HELENA TUKULJAC, ANTOINE DELEFORGE, and REMI GRIBONVAL .....	2182
<b>Mixture Matrix Completion</b> , DANIEL PIMENTEL-ALARCON .....	2193
<b>Trajectory Convolution for Action Recognition</b> , YUE ZHAO, YUANJUN XIONG, and DAHUA LIN .....	2204
<b>The Description Length of Deep Learning models</b> , LÉONARD BLIER, YANN OLLIVIER .....	2216
<b>A Smoothed Analysis of the Greedy Algorithm for the Linear Contextual Bandit Problem</b> , SAMPATH KANNAN, JAMIE MORGENSTERN, and AARON ROTH, BO WAGGONER, ZHIWEI WU .....	2227
<b>Revisiting Decomposable Submodular Function Minimization with Incidence Relations</b> , PAN LI, OLGICA MILENKOVIC .....	2237
<b>A Practical Algorithm for Distributed Clustering and Outlier Detection</b> , JIECAO CHEN, ERFAN AZER, and QIN ZHANG .....	2248
<b>Learning to Reconstruct Shapes from Unseen Classes</b> , XIUMING ZHANG, ZHOUTONG ZHANG, and CHENGKAI ZHANG, JOSHUA TENENBAUM, BILL FREEMAN, JIAJUN WU .....	2257
<b>BourGAN: Generative Networks with Metric Embeddings</b> , CHANG XIAO, PEILIN ZHONG, and CHANGXI ZHENG .....	2269
<b>Smoothed analysis of the low-rank approach for smooth semidefinite programs</b> , THOMAS PUMIR, SAMY JELASSI, and NICOLAS BOUMAL .....	2281
<b>Zero-Shot Transfer with Deictic Object-Oriented Representation in Reinforcement Learning</b> , OFIR MAROM, BENJAMIN ROSMAN .....	2291
<b>Overfitting or perfect fitting? Risk bounds for classification and regression rules that interpolate</b> , MIKHAIL BELKIN, DANIEL HSU, and PARTHA MITRA .....	2300
<b>Breaking the Span Assumption Yields Fast Finite-Sum Minimization</b> , ROBERT HANNAH, YANLI LIU, and DANIEL O'CONNOR, WOTAO YIN .....	2312

<b>Structured Local Minima in Sparse Blind Deconvolution,</b> YUQIAN ZHANG, HAN-WEN KUO, and JOHN WRIGHT .....	2322
<b>GIANT: Globally Improved Approximate Newton Method for Distributed Optimization,</b> SHUSEN WANG, FARBOD ROOSTA-KHORASANI, and PENG XU, MICHAEL MAHONEY .....	2332
<b>Modelling sparsity, heterogeneity, reciprocity and community structure in temporal interaction data,</b> XENIA MISCOURIDOU, FRANCOIS CARON, and YEE TEH .....	2343
<b>Non-monotone Submodular Maximization in Exponentially Fewer Iterations,</b> ERIC BALKANSKI, ADAM BREUER, and YARON SINGER .....	2353
<b>MetaGAN: An Adversarial Approach to Few-Shot Learning,</b> RUIXIANG ZHANG, TONG CHE, and ZOUBIN GHAHRAMANI, YOSHUA BENGIO, YANGQIU SONG .....	2365
<b>Local Differential Privacy for Evolving Data,</b> MATTHEW JOSEPH, AARON ROTH, and JONATHAN ULLMAN, BO WAGGONER .....	2375
<b>Gaussian Process Conditional Density Estimation,</b> VINCENT DUTORDOIR, HUGH SALIMBENI, and JAMES HENSMAN, MARC DEISENROTH ..	2385
<b>Meta-Gradient Reinforcement Learning,</b> ZHONGWEN XU, HADO HASSELT, and DAVID SILVER .....	2396
<b>Modular Networks: Learning to Decompose Neural Computation,</b> LOUIS KIRSCH, JULIUS KUNZE, and DAVID BARBER .....	2408
<b>Learning to Navigate in Cities Without a Map,</b> PIOTR MIROWSKI, MATT GRIMES, and MATEUSZ MALINOWSKI, KARL HERMANN, KEITH ANDERSON, DENIS TEPLYASHIN, KAREN SIMONYAN, KORAY KAVUKCUOGLU, ANDREW ZISSERMAN, RAIA HADSELL .....	2419
<b>Query Complexity of Bayesian Private Learning,</b> KUANG XU .....	2431
<b>A theory on the absence of spurious solutions for nonconvex and nonsmooth optimization,</b> CEDRIC JOSZ, YI OUYANG, and RICHARD ZHANG, JAVAD LAVAEI, SOMAYEH SOJOUDI .....	2441
<b>Recurrent World Models Facilitate Policy Evolution,</b> DAVID HA, JÜRGEN SCHMIDHUBER .....	2450
<b>Ridge Regression and Provable Deterministic Ridge Leverage Score Sampling,</b> SHANNON MCCURDY .....	2463
<b>Wasserstein Variational Inference,</b> LUCA AMBROGIONI, UMUT GÜÇLÜ, and YAĞMUR GÜÇLÜTÜRK, MAX HINNE, MARCEL GERVEN, ERIC MARIS .....	2473
<b>How Does Batch Normalization Help Optimization?,</b> SHIBANI SANTURKAR, DIMITRIS TSIPRAS, and ANDREW ILYAS, ALEKSANDER MADRY ..	2483
<b>Verifiable Reinforcement Learning via Policy Extraction,</b> OSBERT BASTANI, YEWEN PU, and ARMANDO SOLAR-LEZAMA .....	2494
<b>Leveraged volume sampling for linear regression,</b> MICHAL DEREZINSKI, MANFRED WARMUTH, and DANIEL HSU .....	2505
<b>Model Agnostic Supervised Local Explanations,</b> GREGORY PLUMB, DENALI MOLITOR, and AMEET TALWALKAR .....	2515

<b>A Linear Speedup Analysis of Distributed Deep Learning with Sparse and Quantized Communication,</b> PENG JIANG, GAGAN AGRAWAL .....	2525
<b>Active Learning for Non-Parametric Regression Using Purely Random Trees,</b> JONATHAN GOETZ, AMBUJ TEWARI, and PAUL ZIMMERMAN .....	2537
<b>Tree-to-tree Neural Networks for Program Translation,</b> XINYUN CHEN, CHANG LIU, and DAWN SONG .....	2547
<b>Batch-Instance Normalization for Adaptively Style-Invariant Neural Networks,</b> HYEONSEOB NAM, HYO-EUN KIM .....	2558
<b>Structural Causal Bandits: Where to Intervene?,</b> SANGHACK LEE, ELIAS BAREINBOIM .....	2568
<b>Answerer in Questioner’s Mind: Information Theoretic Approach to Goal-Oriented Visual Dialog,</b> SANG-WOO LEE, YU-JUNG HEO, and BYOUNG-TAK ZHANG .....	2579
<b>A Unified Feature Disentangler for Multi-Domain Image Translation and Manipulation,</b> ALEXANDER LIU, YEN-CHENG LIU, and YU-YING YEH, YU-CHIANG WANG .....	2590
<b>Online Learning with an Unknown Fairness Metric,</b> STEPHEN GILLEN, CHRISTOPHER JUNG, and MICHAEL KEARNS, AARON ROTH .....	2600
<b>Isolating Sources of Disentanglement in Variational Autoencoders,</b> RICKY CHEN, XUECHEN LI, and ROGER GROSSE, DAVID DUVENAUD .....	2610
<b>Contextual bandits with surrogate losses: Margin bounds and efficient algorithms,</b> DYLAN FOSTER, AKSHAY KRISHNAMURTHY .....	2621
<b>Representation Learning for Treatment Effect Estimation from Observational Data,</b> LIUYI YAO, SHENG LI, and YALIANG LI, MENGDI HUAI, JING GAO, AIDONG ZHANG .....	2633
<b>Representation Balancing MDPs for Off-policy Policy Evaluation,</b> YAO LIU, OMER GOTTESMAN, and ANIRUDDH RAGHU, MATTHIEU KOMOROWSKI, ALDO FAISAL, FINALE DOSHI-VELEZ, EMMA BRUNSKILL .....	2644
<b>Out of the Box: Reasoning with Graph Convolution Nets for Factual Visual Question Answering,</b> MEDHINI NARASIMHAN, SVETLANA LAZEBNIK, and ALEXANDER SCHWING .....	2654
<b>Causal Discovery from Discrete Data using Hidden Compact Representation,</b> RUICHU CAI, JIE QIAO, and KUN ZHANG, ZHENJIE ZHANG, ZHIFENG HAO .....	2666
<b>Natasha 2: Faster Non-Convex Optimization Than SGD,</b> ZEYUAN ALLEN-ZHU .....	2675
<b>Minimax Statistical Learning with Wasserstein distances,</b> JAEHO LEE, MAXIM RAGINSKY .....	2687
<b>Provable Variational Inference for Constrained Log-Submodular Models,</b> JOSIP DJOLONGA, STEFANIE JEGELKA, and ANDREAS KRAUSE ...	2697
<b>Learning Hierarchical Semantic Image Manipulation through Structured Representations,</b> SEUNGHOO HONG, XINCHEN YAN, and HONGLAK LEE, THOMAS HUANG .....	2708



<b>Processing of missing data by neural networks</b> , MAREK ŚMIEJA, LUKASZ STRUSKI, and JACEK TABOR, BARTOSZ ZIELIŃSKI, PRZEMYSŁAW SPUREK .....	2719
<b>Safe Active Learning for Time-Series Modeling with Gaussian Processes</b> , CHRISTOPH ZIMMER, MONA MEISTER, and NGUYEN-TUONG DUY .....	2730
<b>Optimal Algorithms for Non-Smooth Distributed Optimization in Networks</b> , KEVIN SCAMAN, FRANCIS BACH, and SEBASTIEN BUBECK, LAURENT MASSOULIE, YIN LEE .....	2740
<b>Computing Higher Order Derivatives of Matrix and Tensor Expressions</b> , SÖREN LAUE, MATTHIAS MITTERREITER, and JOACHIM GIESEN .....	2750
<b>Paraphrasing Complex Network: Network Compression via Factor Transfer</b> , JANGHO KIM, SEONGUK PARK, and NOJUN KWAK .....	2760
<b>Analytic solution and stationary phase approximation for the Bayesian lasso and elastic net</b> , TOM MICHOEL .....	2770
<b>Demystifying excessively volatile human learning: A Bayesian persistent prior and a neural approximation</b> , CHAITANYA RYALI, GAUTAM REDDY, and ANGELA YU .....	2781
<b>Empirical Risk Minimization Under Fairness Constraints</b> , MICHELE DONINI, LUCA ONETO, and SHAI BEN-DAVID, JOHN SHAWE-TAYLOR, MASSIMILIANO PONTIL .....	2791
<b>Unsupervised Learning of Shape and Pose with Differentiable Point Clouds</b> , ELДАР INSAFUTDINOV, ALEXEY DOSOVITSKIY .....	2802
<b>Continuous-time Value Function Approximation in Reproducing Kernel Hilbert Spaces</b> , MOTOYA OHNISHI, MASAHIRO YUKAWA, and MIKAEL JOHANSSON, MASASHI SUGIYAMA .....	2813
<b>Gradient Descent Meets Shift-and-Invert Preconditioning for Eigenvector Computation</b> , ZHIQIANG XU .....	2825
<b>Factored Bandits</b> , JULIAN ZIMMERT, YEVGENY SELDIN .....	2835
<b>Delta-encoder: an effective sample synthesis method for few-shot object recognition</b> , ELI SCHWARTZ, LEONID KARLINSKY, and JOSEPH SHTOK, SIVAN HARARY, MATTIAS MARDER, ABHISHEK KUMAR, ROGERIO FERIS, RAJA GIRYES, ALEX BRONSTEIN .....	2845
<b>Metric on Nonlinear Dynamical Systems with Perron-Frobenius Operators</b> , ISAO ISHIKAWA, KEISUKE FUJII, and MASAHIRO IKEDA, YUKA HASHIMOTO, YOSHINOBU KAWAHARA .....	2856
<b>Learning a High Fidelity Pose Invariant Model for High-resolution Face Frontalization</b> , JIE CAO, YIBO HU, and HONGWEN ZHANG, RAN HE, ZHENAN SUN .....	2867
<b>Mirrored Langevin Dynamics</b> , YA-PING HSIEH, ALI KAVIS, and PAUL ROLLAND, VOLKAN CEVHER .....	2878
<b>Moonshine: Distilling with Cheap Convolutions</b> , ELLIOT CROWLEY, GAVIN GRAY, and AMOS STORKEY .....	2888

<b>Stochastic Cubic Regularization for Fast Nonconvex Optimization,</b> NILESH TRIPURANENI, MITCHELL STERN, and CHI JIN, JEFFREY REGIER, MICHAEL JORDAN .....	2899
<b>Adaptation to Easy Data in Prediction with Limited Advice,</b> TOBIAS THUNE, YEVGENY SELDIN .....	2909
<b>Differentially Private Bayesian Inference for Exponential Families,</b> GARRETT BERNSTEIN, DANIEL SHELDON .....	2919
<b>Playing hard exploration games by watching YouTube,</b> YUSUF AYTAR, TOBIAS PFAFF, and DAVID BUDDEN, THOMAS PAINE, ZIYU WANG, NANDO FREITAS .....	2930
<b>Dialog-to-Action: Conversational Question Answering Over a Large-Scale Knowledge Base,</b> DAYA GUO, DUYU TANG, and NAN DUAN, MING ZHOU, JIAN YIN .....	2942
<b>Norm-Ranging LSH for Maximum Inner Product Search,</b> XIAO YAN, JINFENG LI, and XINYAN DAI, HONGZHI CHEN, JAMES CHENG .....	2952
<b>Optimization over Continuous and Multi-dimensional Decisions with Observational Data,</b> DIMITRIS BERTSIMAS, CHRISTOPHER MCCORD	2962
<b>Fast Estimation of Causal Interactions using Wold Processes,</b> FLAVIO FIGUEIREDO, GUILHERME BORGES, and PEDRO MELO, RENATO ASSUNÇÃO .....	2971
<b>When do random forests fail?,</b> CHENG TANG, DAMIEN GARREAU, and ULRIKE LUXBURG .....	2983
<b>Near Optimal Exploration-Exploitation in Non-Communicating Markov Decision Processes,</b> RONAN FRUIT, MATTEO PIROTTA, and ALESSANDRO LAZARIC .....	2994
<b>Optimistic optimization of a Brownian,</b> JEAN-BASTIEN GRILL, MICHAL VALKO, and REMI MUNOS .....	3005
<b>Practical Methods for Graph Two-Sample Testing,</b> DEBARGHYA GHOSHDASTIDAR, ULRIKE LUXBURG .....	3015
<b>NAIS-Net: Stable Deep Networks from Non-Autonomous Differential Equations,</b> MARCO CICCONE, MARCO GALLIERI, and JONATHAN MASCI, CHRISTIAN OSENDORFER, FAUSTINO GOMEZ .....	3025
<b>On the Global Convergence of Gradient Descent for Over- parameterized Models using Optimal Transport,</b> LÉNAÏC CHIZAT, FRANCIS BACH .....	3036
<b>Constructing Deep Neural Networks by Bayesian Network Structure Learning,</b> RAANAN ROHEKAR, SHAMI NISIMOV, and YANIV GURWICZ, GUY KOREN, GAL NOVIK .....	3047
<b>Weakly Supervised Dense Event Captioning in Videos,</b> XUGUANG DUAN, WENBING HUANG, and CHUANG GAN, JINGDONG WANG, WENWU ZHU, JUNZHOU HUANG .....	3059
<b>Faithful Inversion of Generative Models for Effective Amortized Inference,</b> STEFAN WEBB, ADAM GOLINSKI, and ROB ZINKOV, N SIDDHARTH, TOM RAINFORTH, YEE TEH, FRANK WOOD .....	3070
<b>From Stochastic Planning to Marginal MAP,</b> HAO CUI, RADU MARINESCU, and RONI KHARDON .....	3081

<b>On Binary Classification in Extreme Regions,</b> HAMID JALALZAI, STEPHAN CLEMENCON, and ANNE SABOURIN .....	3092
<b>Near-Optimal Policies for Dynamic Multinomial Logit Assortment Selection Models,</b> YINING WANG, XI CHEN, and YUAN ZHOU .....	3101
<b>Q-learning with Nearest Neighbors,</b> DEVAVRAT SHAH, QIAOMIN XIE ..	3111
<b>Global Convergence of Langevin Dynamics Based Algorithms for Nonconvex Optimization,</b> PAN XU, JINGHUI CHEN, and DIFAN ZOU, QUANQUAN GU .....	3122
<b>Asymptotic optimality of adaptive importance sampling,</b> FRANÇOIS PORTIER, BERNARD DELYON .....	3134
<b>Learning latent variable structured prediction models with Gaussian perturbations,</b> KEVIN BELLO, JEAN HONORIO .....	3145
<b>The Nearest Neighbor Information Estimator is Adaptively Near Minimax Rate-Optimal,</b> JIANTAO JIAO, WEIHAO GAO, and YANJUN HAN	3156
<b>Deep Reinforcement Learning of Marked Temporal Point Processes,</b> UTKARSH UPADHYAY, ABIR DE, and MANUEL RODRIGUEZ .....	3168
<b>Evidential Deep Learning to Quantify Classification Uncertainty,</b> MURAT SENSOY, LANCE KAPLAN, and MELIH KANDEMIR .....	3179
<b>Parsimonious Bayesian deep networks,</b> MINGYUAN ZHOU .....	3190
<b>Single-Agent Policy Tree Search With Guarantees,</b> LAURENT ORSEAU, LEVI LELIS, and TOR LATTIMORE, THEOPHANE WEBER .....	3201
<b>Semi-crowdsourced Clustering with Deep Generative Models,</b> YUCEN LUO, TIAN TIAN, and JIAXIN SHI, JUN ZHU, BO ZHANG .....	3212
<b>The committee machine: Computational to statistical gaps in learning a two-layers neural network,</b> BENJAMIN AUBIN, ANTOINE MAILLARD, and JEAN BARBIER, FLORENT KRZAKALA, NICOLAS MACRIS, LENKA ZDEBOROVA .....	3223
<b>Realistic Evaluation of Deep Semi-Supervised Learning Algorithms,</b> AVITAL OLIVER, AUGUSTUS ODENA, and COLIN RAFFEL, EKIN CUBUK, IAN GOODFELLOW .....	3235
<b>Contextual Combinatorial Multi-armed Bandits with Volatile Arms and Submodular Reward,</b> LIXING CHEN, JIE XU, and ZHUO LU ..	3247
<b>Training deep learning based denoisers without ground truth data,</b> SHAKARIM SOLTANAYEV, SE CHUN .....	3257
<b>Re-evaluating evaluation,</b> DAVID BALDUZZI, KARL TUYLS, and JULIEN PEROLAT, THORE GRAEPEL .....	3268
<b>Deep, complex, invertible networks for inversion of transmission effects in multimode optical fibres,</b> OISÍN MORAN, PIERGIORGIO CARAMAZZA, and DANIELE FACCIO, RODERICK MURRAY-SMITH .....	3280
<b>Multivariate Convolutional Sparse Coding for Electromagnetic Brain Signals,</b> TOM TOUR, THOMAS MOREAU, and MAINAK JAS, ALEXANDRE GRAMFORT .....	3292
<b>Data-Efficient Hierarchical Reinforcement Learning,</b> OFIR NACHUM, SHIXIANG GU, and HONGLAK LEE, SERGEY LEVINE .....	3303

<b>Speaker-Follower Models for Vision-and-Language Navigation,</b> DANIEL FRIED, RONGHANG HU, and VOLKAN CIRIK, ANNA ROHRBACH, JACOB ANDREAS, LOUIS-PHILIPPE MORENCY, TAYLOR BERG-KIRKPATRICK, KATE SAENKO, DAN KLEIN, TREVOR DARRELL .....	3314
<b>Inequity aversion improves cooperation in intertemporal social dilemmas,</b> EDWARD HUGHES, JOEL LEIBO, and MATTHEW PHILLIPS, KARL TUYLS, EDGAR DUEÑEZ-GUZMAN, ANTONIO CASTAÑEDA, IAIN DUNNING, TINA ZHU, KEVIN MCKEE, RAPHAEL KOSTER, HEATHER ROFF, THORE GRAEPEL .....	3326
<b>Learning Gaussian Processes by Minimizing PAC-Bayesian Generalization Bounds,</b> DAVID REEB, ANDREAS DOERR, and SEBASTIAN GERWINN, BARBARA RAKITSCH .....	3337
<b>Probabilistic Matrix Factorization for Automated Machine Learning,</b> NICOLO FUSI, RISHIT SHETH, and MELIH ELIBOL .....	3348
<b>Stochastic Spectral and Conjugate Descent Methods,</b> DMITRY KOVALEV, PETER RICHTARIK, and EDUARD GORBUNOV, ELNUR GASANOV ..	3358
<b>Recurrent Relational Networks,</b> RASMUS PALM, ULRICH PAQUET, and OLE WINTHER .....	3368
<b>But How Does It Work in Theory? Linear SVM with Random Features,</b> YITONG SUN, ANNA GILBERT, and AMBUJ TEWARI .....	3379
<b>Learning to Optimize Tensor Programs,</b> TIANQI CHEN, LIANMIN ZHENG, and EDDIE YAN, ZIHENG JIANG, THIERRY MOREAU, LUIS CEZE, CARLOS GUESTRIN, ARVIND KRISHNAMURTHY .....	3389
<b>Boosting Black Box Variational Inference,</b> FRANCESCO LOCATELLO, GIDEON DRESDNER, and RAJIV KHANNA, ISABEL VALERA, GUNNAR RAETSCH	3401
<b>Nearly tight sample complexity bounds for learning mixtures of Gaussians via sample compression schemes,</b> HASSAN ASHTIANI, SHAI BEN-DAVID, and NICK HARVEY, CHRISTOPHER LIAW, ABBAS MEHRABIAN, YANIV PLAN .....	3412
<b>Actor-Critic Policy Optimization in Partially Observable Multiagent Environments,</b> SRINIVASAN SRIRAM, MARC LANCTOT, and VINICIUS ZAMBALDI, JULIEN PEROLAT, KARL TUYLS, REMI MUNOS, MICHAEL BOWLING .....	3422
<b>Step Size Matters in Deep Learning,</b> KAMIL NAR, SHANKAR SASTRY .	3436
<b>Derivative Estimation in Random Design,</b> YU LIU, KRIS BRABANTER	3445
<b>Zeroth-order (Non)-Convex Stochastic Optimization via Conditional Gradient and Gradient Updates,</b> KRISHNA BALASUBRAMANIAN, SAEED GHADIMI .....	3455
<b>Latent Gaussian Activity Propagation: Using Smoothness and Structure to Separate and Localize Sounds in Large Noisy Environments,</b> DANIEL JOHNSON, DANIEL GORELIK, and ROSS MAWHORTER, KYLE SUVER, WEIQING GU, STEVEN XING, CODY GABRIEL, PETER SANKHAGOWIT .....	3465
<b>Hybrid-MST: A Hybrid Active Sampling Strategy for Pairwise Preference Aggregation,</b> JING LI, RAFAL MANTIUK, and JUNLE WANG, SUIYI LING, PATRICK CALLET .....	3475

<b>Infinite-Horizon Gaussian Processes,</b> ARNO SOLIN, JAMES HENSMAN, and RICHARD TURNER .....	3486
<b>Dimensionality Reduction for Stationary Time Series via Stochastic Nonconvex Optimization,</b> MINSHUO CHEN, LIN YANG, and MENGDI WANG, TUO ZHAO .....	3496
<b>Sequence-to-Segment Networks for Segment Detection,</b> ZIJUN WEI, BOYU WANG, and MINH NGUYEN, JIANMING ZHANG, ZHE LIN, XIAOHUI SHEN, RADOMIR MECH, DIMITRIS SAMARAS .....	3507
<b>Scaling the Poisson GLM to massive neural datasets through polynomial approximations,</b> DAVID ZOLTOWSKI, JONATHAN PILLOW ...	3517
<b>Multiplicative Weights Updates with Constant Step-Size in Graphical Constant-Sum Games,</b> YUN CHEUNG .....	3528
<b>Why Is My Classifier Discriminatory?,</b> IRENE CHEN, FREDRIK JOHANSSON, and DAVID SONTAG .....	3539
<b>Multi-Layered Gradient Boosting Decision Trees,</b> JI FENG, YANG YU, and ZHI-HUA ZHOU .....	3551
<b>Learn What Not to Learn: Action Elimination with Deep Reinforcement Learning,</b> TOM ZAHAVY, MATAN HARUSH, and NADAV MERLIS, DANIEL MANKOWITZ, SHIE MANNOR .....	3562
<b>Communication Efficient Parallel Algorithms for Optimization on Manifolds,</b> BAYAN SAPARBAYEVA, MICHAEL ZHANG, and LIZHEN LIN .....	3574
<b>Neural Code Comprehension: A Learnable Representation of Code Semantics,</b> TAL BEN-NUN, ALICE JAKOBOVITS, and TORSTEN HOEFLER .....	3585
<b>Tight Bounds for Collaborative PAC Learning via Multiplicative Weights,</b> JIECAO CHEN, QIN ZHANG, and YUAN ZHOU .....	3598
<b>BinGAN: Learning Compact Binary Descriptors with a Regularized GAN,</b> MACIEJ ZIEBA, PIOTR SEMBERECKI, and TAREK EL-GAALY, TOMASZ TRZCINSKI .....	3608
<b>Modern Neural Networks Generalize on Small Data Sets,</b> MATTHEW OLSON, ABRAHAM WYNER, and RICHARD BERK .....	3619
<b>Escaping Saddle Points in Constrained Optimization,</b> ARYAN MOKHTARI, ASUMAN OZDAGLAR, and ALI JADBABAIE .....	3629
<b>Adversarial Attacks on Stochastic Bandits,</b> KWANG-SUNG JUN, LIHONG LI, and YUZHE MA, XIAOJIN ZHU .....	3640
<b>Optimal Subsampling with Influence Functions,</b> DANIEL TING, ERIC BROCHU .....	3650
<b>A Bandit Approach to Sequential Experimental Design with False Discovery Control,</b> KEVIN JAMIESON, LALIT JAIN .....	3660
<b>Equality of Opportunity in Classification: A Causal Approach,</b> JUNZHE ZHANG, ELIAS BAREINBOIM .....	3671
<b>Towards Understanding Acceleration Tradeoff between Momentum and Asynchrony in Nonconvex Stochastic Optimization,</b> TIANYI LIU, SHIYANG LI, and JIANPING SHI, ENLU ZHOU, TUO ZHAO .....	3682

<b>Unsupervised Attention-guided Image-to-Image Translation,</b> YOUSSEF MEJJATI, CHRISTIAN RICHARDT, and JAMES TOMPKIN, DARREN COSKER, KWANG KIM .....	3693
<b>Inferring Networks From Random Walk-Based Node Similarities,</b> JEREMY HOSKINS, CAMERON MUSCO, and CHRISTOPHER MUSCO, CHARALAMPOS TSOURAKAKIS .....	3704
<b>NEON2: Finding Local Minima via First-Order Oracles,</b> ZEYUAN ALLEN-ZHU, YUANZHI LI .....	3716
<b>Zeroth-Order Stochastic Variance Reduction for Nonconvex Optimization,</b> SIJIA LIU, BHAVYA KAILKHURA, and PIN-YU CHEN, PAI-SHUN TING, SHIYU CHANG, LISA AMINI .....	3727
<b>Online Structured Laplace Approximations for Overcoming Catastrophic Forgetting,</b> HIPPOLYT RITTER, ALEKSANDAR BOTEV, and DAVID BARBER .....	3738
<b>DeepProbLog: Neural Probabilistic Logic Programming,</b> ROBIN MANHAEVE, SEBASTIJAN DUMANCIC, and ANGELIKA KIMMIG, THOMAS DEMEESTER, LUC RAEDT .....	3749
<b>Convergence of Cubic Regularization for Nonconvex Optimization under KL Property,</b> YI ZHOU, ZHE WANG, and YINGBIN LIANG .....	3760
<b>Direct Estimation of Differences in Causal Graphs,</b> YUHAO WANG, CHANDLER SQUIRES, and ANASTASIYA BELYAEVA, CAROLINE UHLER .....	3770
<b>Sublinear Time Low-Rank Approximation of Distance Matrices,</b> AINESH BAKSHI, DAVID WOODRUFF .....	3782
<b>Variational PDEs for Acceleration on Manifolds and Application to Diffeomorphisms,</b> GANESH SUNDARAMOORTHY, ANTHONY YEZZI .....	3793
<b>Bayesian Inference of Temporal Task Specifications from Demonstrations,</b> ANKIT SHAH, PRITISH KAMATH, and JULIE SHAH, SHEN LI .....	3804
<b>Data center cooling using model-predictive control,</b> NEVENA LAZIC, CRAIG BOUTILIER, and TYLER LU, EEHERN WONG, BINZ ROY, MOONKYUNG RYU, GREG IMWALLE .....	3814
<b>Acceleration through Optimistic No-Regret Dynamics,</b> JUN-KUN WANG, JACOB ABERNETHY .....	3824
<b>Lipschitz regularity of deep neural networks: analysis and efficient estimation,</b> ALADIN VIRMAUX, KEVIN SCAMAN .....	3835
<b>Minimax Estimation of Neural Net Distance,</b> KAIYI JI, YINGBIN LIANG .....	3845
<b>Leveraging the Exact Likelihood of Deep Latent Variable Models,</b> PIERRE-ALEXANDRE MATTEI, JES FRELLSEN .....	3855
<b>Bipartite Stochastic Block Models with Tiny Clusters,</b> STEFAN NEUMANN .....	3867
<b>Learning sparse neural networks via sensitivity-driven regularization,</b> ENZO TARTAGLIONE, SKJALG LEPSØY, and ATTILIO FIANDROTTI, GIANLUCA FRANCIANI .....	3878

<b>Faster Online Learning of Optimal Threshold for Consistent F-measure Optimization,</b> XIAOXUAN ZHANG, MINGRUI LIU, and XUN ZHOU, TIANBAO YANG .....	3889
<b>Direct Runge-Kutta Discretization Achieves Acceleration,</b> JINGZHAO ZHANG, ARYAN MOKHTARI, and SUVRIT SRA, ALI JADBABAIE ....	3900
<b>Adversarial Examples that Fool both Computer Vision and Time-Limited Humans,</b> GAMALELDIN ELSAYED, SHREYA SHANKAR, and BRIAN CHEUNG, NICOLAS PAPERNOT, ALEXEY KURAKIN, IAN GOODFELLOW, JASCHA SOHL-DICKSTEIN .....	3910
<b>Stochastic Nested Variance Reduction for Nonconvex Optimization,</b> DONGRUO ZHOU, PAN XU, and QUANQUAN GU .....	3921
<b>Faster Neural Networks Straight from JPEG,</b> LIONEL GUEGUEN, ALEX SERGEEV, and BEN KADLEC, ROSANNE LIU, JASON YOSINSKI .....	3933
<b>TopRank: A practical algorithm for online stochastic ranking,</b> TOR LATTIMORE, BRANISLAV KVETON, and SHUAI LI, CSABA SZEPESVARI ...	3945
<b>Learning from discriminative feature feedback,</b> SANJOY DASGUPTA, AKANSHA DEY, and NICHOLAS ROBERTS, SIVAN SABATO .....	3955
<b>RetGK: Graph Kernels based on Return Probabilities of Random Walks,</b> ZHEN ZHANG, MIANZHI WANG, and YIJIAN XIANG, YAN HUANG, ARYE NEHORAI .....	3964
<b>Deep Generative Markov State Models,</b> HAO WU, ANDREAS MARDT, and LUCA PASQUALI, FRANK NOE .....	3975
<b>Early Stopping for Nonparametric Testing ,</b> MEIMEI LIU, GUANG CHENG .....	3985
<b>Solving Non-smooth Constrained Programs with Lower Complexity than <math>\mathcal{O}(1/\varepsilon)</math>: A Primal-Dual Homotopy Smoothing Approach,</b> XIAOHAN WEI, HAO YU, and QING LING, MICHAEL NEELY .....	3995
<b>Heterogeneous Bitwidth Binarization in Convolutional Neural Networks,</b> JOSHUA FROMM, SHWETAK PATEL, and MATTHAI PHILIPOSE ...	4006
<b>Unsupervised Learning of Object Landmarks through Conditional Image Generation,</b> TOMAS JAKAB, ANKUSH GUPTA, and HAKAN BILEN, ANDREA VEDALDI .....	4016
<b>Probabilistic Neural Programmed Networks for Scene Generation,</b> ZHIWEI DENG, JIACHENG CHEN, and YIFANG FU, GREG MORI .....	4028
<b>The streaming rollout of deep networks - towards fully model-parallel execution,</b> VOLKER FISCHER, JAN KOEHLER, and THOMAS PFEIL .....	4039
<b>KONG: Kernels for ordered-neighborhood graphs,</b> MOEZ DRAIEF, KONSTANTIN KUTZKOV, and KEVIN SCAMAN, MILAN VOJNOVIC .....	4051
<b>GumBolt: Extending Gumbel trick to Boltzmann priors,</b> AMIR KHOSHAMAN, MOHAMMAD AMIN .....	4061
<b>Neural Networks Trained to Solve Differential Equations Learn General Representations,</b> MARTIN MAGILL, FAISAL QURESHI, and HENDRICK HAAN .....	4071

<b>Beauty-in-averageness and its contextual modulations: A Bayesian statistical account</b> , CHAITANYA RYALI, ANGELA YU .....	4082
<b>Distributed Weight Consolidation: A Brain Segmentation Case Study</b> , PATRICK MCCLURE, CHARLES ZHENG, and JAKUB KACZMARZYK, JOHN ROGERS-LEE, SATRAJIT GHOSH, DYLAN NIELSON, PETER BANDETTINI, FRANCISCO PEREIRA .....	4093
<b>Efficient Projection onto the Perfect Phylogeny Model</b> , BEI JIA, SURJYENDU RAY, and SAM SAFAVI, JOSE BENTO .....	4104
<b>TETRIS: Tile-matching the TRemendous Irregular Sparsity</b> , YU JI, LING LIANG, and LEI DENG, YOUYANG ZHANG, YOUHUI ZHANG, YUAN XIE .....	4115
<b>Cooperative neural networks (CoNN): Exploiting prior independence structure for improved classification</b> , HARSH SHRIVASTAVA, EUGENE BART, and BOB PRICE, HANJUN DAI, BO DAI, SRINIVAS ALURU .....	4126
<b>Differentially Private Robust Low-Rank Approximation</b> , RAMAN ARORA, VLADIMIR BRAVERMAN, and JALAJ UPADHYAY .....	4137
<b>Meta-Learning MCMC Proposals</b> , TONGZHOU WANG, YI WU, and DAVID MOORE, STUART RUSSELL .....	4146
<b>An Information-Theoretic Analysis for Thompson Sampling with Many Actions</b> , SHI DONG, BENJAMIN ROY .....	4157
<b>Flexible and accurate inference and learning for deep generative models</b> , ESZTER VERTES, MANEESH SAHANI .....	4166
<b>The Price of Privacy for Low-rank Factorization</b> , JALAJ UPADHYAY .....	4176
<b>Regret Bounds for Robust Adaptive Control of the Linear Quadratic Regulator</b> , SARAH DEAN, HORIA MANIA, and NIKOLAI MATNI, BENJAMIN RECHT, STEPHEN TU .....	4188
<b>Bilevel Distance Metric Learning for Robust Image Recognition</b> , JIE XU, LEI LUO, and CHENG DENG, HENG HUANG .....	4198
<b>Differentially Private Uniformly Most Powerful Tests for Binomial Data</b> , JORDAN AWAN, ALEKSANDRA SLAVKOVIĆ .....	4208
<b>Scalable Coordinated Exploration in Concurrent Reinforcement Learning</b> , MARIA DIMAKOPOULOU, IAN OSBAND, and BENJAMIN ROY .....	4219
<b>Integrated accounts of behavioral and neuroimaging data using flexible recurrent neural network models</b> , AMIR DEZFOULI, RICHARD MORRIS, and FABIO RAMOS, PETER DAYAN, BERNARD BALLEINE .....	4228
<b>BML: A High-performance, Low-cost Gradient Synchronization Algorithm for DML Training</b> , SONGTAO WANG, DAN LI, and YANG CHENG, JINKUN GENG, YANSHU WANG, SHUAI WANG, SHUTAO XIA, JIANPING WU .....	4238
<b>Inexact trust-region algorithms on Riemannian manifolds</b> , HIROYUKI KASAI, BAMDEV MISHRA .....	4249
<b>Can We Gain More from Orthogonality Regularizations in Training Deep Networks?</b> , NITIN BANSAL, XIAOHAN CHEN, and ZHANGYANG WANG .....	4261



<b>Binary Rating Estimation with Graph Side Information,</b> KWANGJUN AHN, KANGWOOK LEE, and HYUNSEUNG CHA, CHANGHO SUH ...	4272
<b>Simple Embedding for Link Prediction in Knowledge Graphs,</b> SEYED KAZEMI, DAVID POOLE .....	4284
<b>Differentially Private Contextual Linear Bandits,</b> ROSHAN SHARIFF, OR SHEFFET .....	4296
<b>Submodular Field Grammars: Representation, Inference, and Application to Image Parsing,</b> ABRAM FRIESEN, PEDRO DOMINGOS ...	4307
<b>A Bridging Framework for Model Optimization and Deep Propagation,</b> RISHENG LIU, SHICHAO CHENG, and XIAOKUN LIU, LONG MA, XIN FAN, ZHONGXUAN LUO .....	4318
<b>Completing State Representations using Spectral Learning,</b> NAN JIANG, ALEX KULESZA, and SATINDER SINGH .....	4328
<b>Optimization of Smooth Functions with Noisy Observations: Local Minimax Rates,</b> YINING WANG, SIVARAMAN BALAKRISHNAN, and AARTI SINGH .....	4338
<b>Adding One Neuron Can Eliminate All Bad Local Minima,</b> SHIYU LIANG, RUOYU SUN, and JASON LEE, RAYADURGAM SRIKANT .....	4350
<b>Mean-field theory of graph neural networks in graph partitioning,</b> TATSURO KAWAMOTO, MASASHI TSUBAKI, and TOMOYUKI OBUCHI .....	4361
<b>The Physical Systems Behind Optimization Algorithms,</b> LIN YANG, RAMAN ARORA, and VLADIMIR BRAVERMAN, TUO ZHAO .....	4372
<b>Mallows Models for Top-k Lists,</b> FLAVIO CHIERICHETTI, ANIRBAN DASGUPTA, and SHAHRZAD HADDADAN, RAVI KUMAR, SILVIO LATTANZI .....	4382
<b>Amortized Inference Regularization,</b> RUI SHU, HUNG BUI, and SHENGJIA ZHAO, MYKEL KOCHENDERFER, STEFANO ERMON .....	4393
<b>Maximum Causal Tsallis Entropy Imitation Learning,</b> KYUNGJAE LEE, SUNGJOON CHOI, and SONGHWAI OH .....	4403
<b>Limited Memory Kelley’s Method Converges for Composite Convex and Submodular Objectives,</b> SONG ZHOU, SWATI GUPTA, and MADELEINE UDELL .....	4414
<b>Semi-Supervised Learning with Declaratively Specified Entropy Constraints,</b> HAITIAN SUN, WILLIAM COHEN, and LIDONG BING .....	4425
<b>End-to-end Symmetry Preserving Inter-atomic Potential Energy Model for Finite and Extended Systems,</b> LINFENG ZHANG, JIEQUN HAN, and HAN WANG, WISSAM SAIDI, ROBERTO CAR, WEINAN E .....	4436
<b>Sparsified SGD with Memory,</b> SEBASTIAN STICH, JEAN-BAPTISTE CORDONNIER, and MARTIN JAGGI .....	4447
<b>Exponentiated Strongly Rayleigh Distributions,</b> ZELDA MARIET, SUVRIT SRA, and STEFANIE JEGELKA .....	4459
<b>Importance Weighting and Variational Inference,</b> JUSTIN DOMKE, DANIEL SHELDON .....	4470
<b>Transfer Learning from Speaker Verification to Multispeaker Text-To-Speech Synthesis,</b> YE JIA, YU ZHANG, and RON WEISS, QUAN	

WANG, JONATHAN SHEN, FEI REN, ZHIFENG CHEN, PATRICK NGUYEN, RUOMING PANG, IGNACIO MORENO, YONGHUI WU .....	4480
<b>Expanding Holographic Embeddings for Knowledge Completion,</b> YEXIANG XUE, YANG YUAN, and ZHITIAN XU, ASHISH SABHARWAL .....	4491
<b>Lifelong Inverse Reinforcement Learning,</b> JORGE MENDEZ, SHASHANK SHIVKUMAR, and ERIC EATON .....	4502
<b>Explaining Deep Learning Models – A Bayesian Non-parametric Approach,</b> WENBO GUO, SUI HUANG, and YUNZHE TAO, XINYU XING, LIN LIN .....	4514
<b>Third-order Smoothness Helps: Faster Stochastic Optimization Algorithms for Finding Local Minima,</b> YAODONG YU, PAN XU, and QUANQUAN GU .....	4525
<b>COLA: Decentralized Linear Learning,</b> LIE HE, YATAO BIAN, and MARTIN JAGGI .....	4536
<b>MiME: Multilevel Medical Embedding of Electronic Health Records for Predictive Healthcare,</b> EDWARD CHOI, CAO XIAO, and WALTER STEWART, JIMENG SUN .....	4547
<b>Adaptive Sampling Towards Fast Graph Representation Learning,</b> WENBING HUANG, TONG ZHANG, and YU RONG, JUNZHOU HUANG .....	4558
<b>Hunting for Discriminatory Proxies in Linear Regression Models,</b> SAMUEL YEOM, ANUPAM DATTA, and MATT FREDRIKSON .....	4568
<b>Towards Robust Detection of Adversarial Examples,</b> TIANYU PANG, CHAO DU, and YINPENG DONG, JUN ZHU .....	4579
<b>Active Matting,</b> XIN YANG, KE XU, and CITY KONG), SHAOZHE CHEN, SHENGFENG HE, BAOCAI YIN .....	4590
<b>Learning filter widths of spectral decompositions with wavelets,</b> HAIDAR KHAN, BÜLENT YENER .....	4601
<b>Byzantine Stochastic Gradient Descent,</b> DAN ALISTARH, ZEYUAN ALLEN-ZHU, and JERRY LI .....	4613
<b>PG-TS: Improved Thompson Sampling for Logistic Contextual Bandits,</b> BIANCA DUMITRASCU, KAREN FENG, and BARBARA ENGELHARDT	4624
<b>Spectral Filtering for General Linear Dynamical Systems,</b> ELAD HAZAN, HOLDEN LEE, and KARAN SINGH, CYRIL ZHANG, YI ZHANG .....	4634
<b>On Learning Intrinsic Rewards for Policy Gradient Methods,</b> ZEYU ZHENG, JUNHYUK OH, and SATINDER SINGH .....	4644
<b>Boolean Decision Rules via Column Generation,</b> SANJEEB DASH, OKTAY GUNLUK, and DENNIS WEI .....	4655
<b>Adversarial Text Generation via Feature-Mover’s Distance,</b> LIQUN CHEN, SHUYANG DAI, and CHENYANG TAO, HAICHAO ZHANG, ZHE GAN, DINGHAN SHEN, YIZHE ZHANG, GUOYIN WANG, RUIYI ZHANG, LAWRENCE DUKE .....	4666
<b>Fast Rates of ERM and Stochastic Approximation: Adaptive to Error Bound Conditions,</b> MINGRUI LIU, XIAOXUAN ZHANG, and LIJUN ZHANG, RONG JIN, TIANBAO YANG .....	4678

<b>Learning Bounds for Greedy Approximation with Explicit Feature Maps from Multiple Kernels,</b> SHAHIN SHAHRAMPOUR, VAHID TAROKH	4690
<b>A Mathematical Model For Optimal Decisions In A Representative Democracy ,</b> MALIK MAGDON-ISMAIL, LIRONG XIA	4702
<b>Negotiable Reinforcement Learning for Pareto Optimal Sequential Decision-Making,</b> NISHANT DESAI, ANDREW CRITCH, and STUART RUSSELL	4712
<b>Non-metric Similarity Graphs for Maximum Inner Product Search,</b> STANISLAV MOROZOV, ARTEM BABENKO	4721
<b>Recurrently Controlled Recurrent Networks,</b> YI TAY, ANH LUU, and SIU HUI	4731
<b>Fast greedy algorithms for dictionary selection with generalized sparsity constraints,</b> KAITO FUJII, TASUKU SOMA	4744
<b>Deep Reinforcement Learning in a Handful of Trials using Probabilistic Dynamics Models,</b> KURLAND CHUA, ROBERTO CALANDRA, and ROWAN MCALLISTER, SERGEY LEVINE	4754
<b>A Smoother Way to Train Structured Prediction Models,</b> KRISHNA PILLUTLA, VINCENT ROULET, and SHAM KAKADE, ZAID HARCHAOUI	4766
<b>Context-dependent upper-confidence bounds for directed exploration,</b> RAKSHA KUMARASWAMY, MATTHEW SCHLEGEL, and ADAM WHITE, MARTHA WHITE	4779
<b>A Unified View of Piecewise Linear Neural Network Verification,</b> RUDY BUNEL, ILKER TURKASLAN, and PHILIP TORR, PUSHMEET KOHLI, M. KUMAR	4790
<b>Hierarchical Graph Representation Learning with Differentiable Pooling,</b> REX YING, JIAXUAN YOU, and CHRISTOPHER MORRIS, XIANG REN, WILLIAM HAMILTON, JURE LESKOVEC	4800
<b>Non-Ergodic Alternating Proximal Augmented Lagrangian Algorithms with Optimal Rates,</b> DINH TRAN	4811
<b>Information-based Adaptive Stimulus Selection to Optimize Communication Efficiency in Brain-Computer Interfaces,</b> BOYLA MAINSAH, DMITRY KALIKA, and LESLIE COLLINS, SIYUAN LIU, CHANDRA THROCKMORTON	4820
<b>Porcupine Neural Networks: Approximating Neural Network Landscapes,</b> SOHEIL FEIZI, HAMID JAVADI, and JESSE ZHANG, DAVID TSE	4831
<b>Fairness Through Computationally-Bounded Awareness,</b> MICHAEL KIM, OMER REINGOLD, and GUY ROTHBLUM	4842
<b>Adaptive Negative Curvature Descent with Applications in Non-convex Optimization,</b> MINGRUI LIU, ZHE LI, and XIAOYU WANG, JINFENG YI, TIANBAO YANG	4853
<b>Is Q-Learning Provably Efficient?,</b> CHI JIN, ZEYUAN ALLEN-ZHU, and SEBASTIEN BUBECK, MICHAEL JORDAN	4863
<b>Interpreting Neural Network Judgments via Minimal, Stable, and Symbolic Corrections,</b> XIN ZHANG, ARMANDO SOLAR-LEZAMA, and RISHABH SINGH	4874

<b>Measures of distortion for machine learning</b> , LEENA VANKADARA, ULRIKE LUXBURG .....	4886
<b>On the Local Minima of the Empirical Risk</b> , CHI JIN, LYDIA LIU, and RONG GE, MICHAEL JORDAN .....	4896
<b>Densely Connected Attention Propagation for Reading Comprehension</b> , YI TAY, ANH LUU, and SIU HUI, JIAN SU .....	4906
<b>Bandit Learning with Positive Externalities</b> , VIRAG SHAH, JOSE BLANCHET, and RAMESH JOHARI .....	4918
<b>Learning Confidence Sets using Support Vector Machines</b> , WENBO WANG, XINGYE QIAO .....	4929
<b>Efficient Neural Network Robustness Certification with General Activation Functions</b> , HUAN ZHANG, LILY WENG, and PIN-YU CHEN, CHO-JUI HSIEH, LUCA DANIEL .....	4939
<b>Hessian-based Analysis of Large Batch Training and Robustness to Adversaries</b> , ZHEWEI YAO, AMIR GHOLAMI, and QI LEI, KURT KEUTZER, MICHAEL MAHONEY .....	4949
<b>Neural Edit Operations for Biological Sequences</b> , SATOSHI KOIDE, KEISUKE KAWANO, and TAKURO KUTSUNA .....	4960
<b>Objective and efficient inference for couplings in neuronal networks</b> , YU TERADA, TOMOYUKI OBUCHI, and TAKUYA ISOMURA, YOSHIYUKI KABASHIMA .....	4971
<b>Learning from Group Comparisons: Exploiting Higher Order Interactions</b> , YAO LI, MINHAO CHENG, and KEVIN FUJII, FUSHING HSIEH, CHO-JUI HSIEH .....	4981
<b>Supervising Unsupervised Learning</b> , VIKAS GARG, ADAM KALAI .....	4991
<b>Nonparametric Bayesian Lomax delegate racing for survival analysis with competing risks</b> , QUAN ZHANG, MINGYUAN ZHOU .....	5002
<b>Adversarially Robust Generalization Requires More Data</b> , LUDWIG SCHMIDT, SHIBANI SANTURKAR, and DIMITRIS TSIPRAS, KUNAL TALWAR, ALEKSANDER MADRY .....	5014
<b>Improving Exploration in Evolution Strategies for Deep Reinforcement Learning via a Population of Novelty-Seeking Agents</b> , EDOARDO CONTI, VASHISHT MADHAVAN, and FELIPE SUCH, JOEL LEHMAN, KENNETH STANLEY, JEFF CLUNE .....	5027
<b>Practical exact algorithm for trembling-hand equilibrium refinements in games</b> , GABRIELE FARINA, NICOLA GATTI, and TUOMAS SANDHOLM .....	5039
<b>LAG: Lazily Aggregated Gradient for Communication-Efficient Distributed Learning</b> , TIANYI CHEN, GEORGIOS GIANNAKIS, and TAO SUN, WOTAO YIN .....	5050
<b>Scalable Robust Matrix Factorization with Nonconvex Loss</b> , QUANMING YAO, JAMES KWOK .....	5061
<b>Power-law efficient neural codes provide general link between perceptual bias and discriminability</b> , MICHAEL MORAIS, JONATHAN PILLOW .....	5071

<b>Geometry-Aware Recurrent Neural Networks for Active Visual Recognition,</b> RICSON CHENG, ZIYAN WANG, and KATERINA FRAGKIADAKI	5081
<b>Unsupervised Adversarial Invariance,</b> AYUSH JAISWAL, YUE WU, and WAEL ABD-ALMAGEED, PREM NATARAJAN	5092
<b>Content preserving text generation with attribute controls,</b> LAJANUGEN LOGESWARAN, HONGLAK LEE, and SAMY BENGIO	5103
<b>Multi-armed Bandits with Compensation,</b> SIWEI WANG, LONGBO HUANG	5114
<b>GradiVeQ: Vector Quantization for Bandwidth-Efficient Gradient Aggregation in Distributed CNN Training,</b> MINGCHAO YU, ZHIFENG LIN, and KRISHNA NARRA, SONGZE LI, YOUJIE LI, NAM KIM, ALEXANDER SCHWING, MURALI ANNAVARAM, SALMAN AVESTIMEHR	5123
<b>Learning in Games with Lossy Feedback,</b> ZHENGYUAN ZHOU, PANAYOTIS MERTIKOPOULOS, and SUSAN ATHEY, NICHOLAS BAMBOS, PETER GLYNN, YINYU YE	5134
<b>Scalable methods for 8-bit training of neural networks,</b> RON BANNER, ITAY HUBARA, and ELAD HOFFER, DANIEL SOUDRY	5145
<b>Dropping Symmetry for Fast Symmetric Nonnegative Matrix Factorization,</b> ZHIHUI ZHU, XIAO LI, and KAI LIU, QIUWEI LI	5154
<b>Link Prediction Based on Graph Neural Networks,</b> MUHAN ZHANG, YIXIN CHEN	5165
<b>Why so gloomy? A Bayesian explanation of human pessimism bias in the multi-armed bandit task,</b> DALIN GUO, ANGELA YU	5176
<b>Near-Optimal Time and Sample Complexities for Solving Markov Decision Processes with a Generative Model,</b> AARON SIDFORD, MENGDI WANG, and XIAN WU, LIN YANG, YINYU YE	5186
<b>ChannelNets: Compact and Efficient Convolutional Neural Networks via Channel-Wise Convolutions,</b> HONGYANG GAO, ZHENGYANG WANG, and SHUIWANG JI	5197
<b>Causal Inference and Mechanism Clustering of A Mixture of Additive Noise Models,</b> SHOUBO HU, ZHITANG CHEN, and VAHID NIA, LAIWAN CHAN, YANHUI GENG	5206
<b>Contour location via entropy reduction leveraging multiple information sources,</b> ALEXANDRE MARQUES, RÉMI LAM, and KAREN WILLCOX	5217
<b>Assessing Generative Models via Precision and Recall,</b> MEHDI SAJJADI, OLIVIER BACHEM, and MARIO LUCIC, OLIVIER BOUSQUET, SYLVAIN GELLY	5228
<b>Multiple-Step Greedy Policies in Approximate and Online Reinforcement Learning,</b> YONATHAN EFRONI, GAL DALAL, and BRUNO SCHERRER, SHIE MANNOR	5238
<b>A Convex Duality Framework for GANs,</b> FARZAN FARNIA, DAVID TSE	5248
<b>Horizon-Independent Minimax Linear Regression,</b> ALAN MALEK, PETER BARTLETT	5259

<b>Exploiting Numerical Sparsity for Efficient Learning : Faster Eigenvector Computation and Regression,</b> NEHA GUPTA, AARON SIDFORD .....	5269
<b>Experimental Design for Cost-Aware Learning of Causal Graphs,</b> ERIK LINDGREN, MURAT KOCAOGLU, and ALEX DIMAKIS, SRIRAM VISHWANATH .....	5279
<b>Task-Driven Convolutional Recurrent Models of the Visual System,</b> ARAN NAYEBI, DANIEL BEAR, and JONAS KUBILIUS, KOHITIJ KAR, SURYA GANGULI, DAVID SUSSILLO, JAMES DICARLO, DANIEL YAMINS .....	5290
<b>Meta-Reinforcement Learning of Structured Exploration Strategies,</b> ABHISHEK GUPTA, RUSSELL MENDONCA, and YUXUAN LIU, PIETER ABBEEL, SERGEY LEVINE .....	5302
<b>Sample Efficient Stochastic Gradient Iterative Hard Thresholding Method for Stochastic Sparse Linear Regression with Limited Attribute Observation,</b> TOMOYA MURATA, TAJI SUZUKI .....	5312
<b>Semi-supervised Deep Kernel Learning: Regression with Unlabeled Data by Minimizing Predictive Variance,</b> NEAL JEAN, SANG XIE, and STEFANO ERMON .....	5322
<b>Generalizing to Unseen Domains via Adversarial Data Augmentation,</b> RICCARDO VOLPI, HONGSEOK NAMKOONG, and OZAN SENER, JOHN DUCHI, VITTORIO MURINO, SILVIO SAVARESE .....	5334
<b>Hyperbolic Neural Networks,</b> OCTAVIAN GANEA, GARY BECIGNEUL, and THOMAS HOFMANN .....	5345
<b>Breaking the Curse of Horizon: Infinite-Horizon Off-Policy Estimation,</b> QIANG LIU, LIHONG LI, and ZIYANG TANG, DENNY ZHOU .....	5356
<b>Learning Task Specifications from Demonstrations,</b> MARCELL VAZQUEZ-CHANLATTE, SUSMIT JHA, and ASHISH TIWARI, MARK HO, SANJIT SESHIA .....	5367
<b>Learning a latent manifold of odor representations from neural responses in piriform cortex,</b> ANQI WU, STAN PASHKOVSKI, and SANDEEP DATTA, JONATHAN PILLOW .....	5378
<b>Fully Understanding The Hashing Trick,</b> LIOR KAMMA, CASPER FREKSEN, and KASPER LARSEN .....	5389
<b>Evolved Policy Gradients,</b> REIN HOUTHOOFT, RICHARD CHEN, and PHILLIP ISOLA, BRADLY STADIE, FILIP WOLSKI, JONATHAN HO, PIETER ABBEEL .....	5400
<b>The Spectrum of the Fisher Information Matrix of a Single-Hidden-Layer Neural Network,</b> JEFFREY PENNINGTON, PRATIK WORAH .....	5410
<b>Learning Concave Conditional Likelihood Models for Improved Analysis of Tandem Mass Spectra,</b> JOHN HALLORAN, DAVID ROCKE ..	5420
<b>Differentially Private k-Means with Constant Multiplicative Error,</b> URI STEMMER, HAIM KAPLAN .....	5431
<b>Policy Optimization via Importance Sampling,</b> ALBERTO METELLI, MATTEO PAPINI, and FRANCESCO FACCIO, MARCELLO RESTELLI .....	5442

<b>Estimating Learnability in the Sublinear Data Regime,</b> WEIHAO KONG, GREGORY VALIANT .....	5455
<b>Algorithmic Assurance: An Active Approach to Algorithmic Testing using Bayesian Optimisation,</b> SHIVAPRATAP GOPAKUMAR, SUNIL GUPTA, and SANTU RANA, VU NGUYEN, SVETHA VENKATESH .....	5465
<b>Community Exploration: From Offline Optimization to Online Learning,</b> XIAOWEI CHEN, WEIRAN HUANG, and WEI CHEN, JOHN LUI ....	5474
<b>A Dual Framework for Low-rank Tensor Completion,</b> MADHAV NIMISHAKAVI, PRATIK JAWANPURIA, and BAMDEV MISHRA .....	5484
<b>Low-rank Interaction with Sparse Additive Effects Model for Large Data Frames,</b> GENEVIÈVE ROBIN, HOI-TO WAI, and JULIE JOSSE, OLGA KLOPP, ERIC MOULINES .....	5496
<b>Inference Aided Reinforcement Learning for Incentive Mechanism Design in Crowdsourcing,</b> ZEHONG HU, YITAO LIANG, and JIE ZHANG, ZHAO LI, YANG LIU .....	5507
<b>Middle-Out Decoding,</b> SHIKIB MEHRI, LEONID SIGAL .....	5518
<b>First-order Stochastic Algorithms for Escaping From Saddle Points in Almost Linear Time,</b> YI XU, RONG JIN, and TIANBAO YANG .	5530
<b>To Trust Or Not To Trust A Classifier,</b> HEINRICH JIANG, BEEN KIM, and MELODY GUAN, MAYA GUPTA .....	5541
<b>Reparameterization Gradient for Non-differentiable Models,</b> WONYEOL LEE, HANGYEOL YU, and HONGSEOK YANG .....	5553
<b>A Simple Proximal Stochastic Gradient Method for Nonsmooth Nonconvex Optimization,</b> ZHIZE LI, JIAN LI .....	5564
<b>Multimodal Generative Models for Scalable Weakly-Supervised Learning,</b> MIKE WU, NOAH GOODMAN .....	5575
<b>How Much Restricted Isometry is Needed In Nonconvex Matrix Recovery?,</b> RICHARD ZHANG, CEDRIC JOSZ, and SOMAYEH SOJOUDI, JAVAD LAVAEI .....	5586
<b>Occam’s razor is insufficient to infer the preferences of irrational agents,</b> STUART ARMSTRONG, SÖREN MINDERMANN .....	5598
<b>Manifold Structured Prediction,</b> ALESSANDRO RUDI, CARLO CILIBERTO, and GIAN MARCONI, LORENZO ROSASCO .....	5610
<b>Fast Greedy MAP Inference for Determinantal Point Process to Improve Recommendation Diversity,</b> LAMING CHEN, GUOXIN ZHANG, and HANNING ZHOU .....	5622
<b>Learning Others’ Intentional Models in Multi-Agent Settings Using Interactive POMDPs,</b> YANLIN HAN, PIOTR GMYTRASIEWICZ ....	5634
<b>Contextual Pricing for Lipschitz Buyers,</b> JIEMING MAO, RENATO LEME, and JON SCHNEIDER .....	5643
<b>Online Improper Learning with an Approximation Oracle,</b> ELAD HAZAN, WEI HU, and YUANZHI LI, ZHIYUAN LI .....	5652
<b>Bandit Learning in Concave N-Person Games,</b> MARIO BRAVO, DAVID LESLIE, and PANAYOTIS MERTIKOPOULOS .....	5661

<b>On Fast Leverage Score Sampling and Optimal Learning,</b> ALESSANDRO RUDI, DANIELE CALANDRIELLO, and LUIGI CARRATINO, LORENZO ROSASCO .....	5672
<b>Unsupervised Video Object Segmentation for Deep Reinforcement Learning,</b> VIKASH GOEL, JAMESON WENG, and PASCAL POUPART .....	5683
<b>Efficient inference for time-varying behavior during learning,</b> NICHOLAS ROY, JI BAK, and ATHENA AKRAMI, CARLOS BRODY, JONATHAN PILLOW .....	5695
<b>Learning convex polytopes with margin,</b> LEE-AD GOTTLIEB, ERAN KAUFMAN, and ARYEH KONTOROVICH, GABRIEL NIVASCH .....	5706
<b>Critical initialisation for deep signal propagation in noisy rectifier neural networks,</b> ARNU PRETORIUS, ELAN BILJON, and STEVE KROON, HERMAN KAMPER .....	5717
<b>Insights on representational similarity in neural networks with canonical correlation,</b> ARI MORCOS, MAITHRA RAGHU, and SAMY BENGIO .....	5727
<b>Variational Inference with Tail-adaptive f-Divergence,</b> DILIN WANG, HAO LIU, and QIANG LIU .....	5737
<b>Mental Sampling in Multimodal Representations,</b> JIANQIAO ZHU, ADAM SANBORN, and NICK CHATER .....	5748
<b>Adversarially Robust Optimization with Gaussian Processes,</b> ILIJA BOGUNOVIC, JONATHAN SCARLETT, and STEFANIE JEGELKA, VOLKAN CEVHER .....	5760
<b>Learning to Multitask,</b> YU ZHANG, YING WEI, and QIANG YANG .....	5771
<b>Loss Functions for Multiset Prediction,</b> SEAN WELLECK, ZIXIN YAO, and YU GAI, JIALIN MAO, ZHENG ZHANG, KYUNGHYUN CHO .....	5783
<b>Computing Kantorovich-Wasserstein Distances on <math>d</math>-dimensional histograms using <math>(d+1)</math>-partite graphs,</b> GENNARO AURICCHIO, FEDERICO BASSETTI, and STEFANO GUALANDI, MARCO VENERONI .....	5793
<b>Neural Interaction Transparency (NIT): Disentangling Learned Interactions for Improved Interpretability,</b> MICHAEL TSANG, HANPENG LIU, and SANJAY PURUSHOTHAM, PAVANKUMAR MURALI, YAN LIU .....	5804
<b>CapProNet: Deep Feature Learning via Orthogonal Projections onto Capsule Subspaces,</b> LIHENG ZHANG, MARZIEH EDRAKI, and GUO-JUN QI .....	5814
<b>Gamma-Poisson Dynamic Matrix Factorization Embedded with Metadata Influence,</b> TRONG DO, LONGBIN CAO .....	5824
<b>Masking: A New Perspective of Noisy Supervision,</b> BO HAN, JIANGCHAO YAO, and GANG NIU, MINGYUAN ZHOU, IVOR TSANG, YA ZHANG, MASASHI SUGIYAMA .....	5836
<b>On GANs and GMMs,</b> EITAN RICHARDSON, YAIR WEISS .....	5847
<b>Differential Properties of Sinkhorn Approximation for Learning with Wasserstein Distance,</b> GIULIA LUISE, ALESSANDRO RUDI, and MASSIMILIANO PONTIL, CARLO CILIBERTO .....	5859
<b>Practical Deep Stereo (PDS): Toward applications-friendly deep stereo matching,</b> STEPAN TULYAKOV, ANTON IVANOV, and FRANCOIS FLEURET .....	5871



<b>A Bayes-Sard Cubature Method,</b> TONI KARVONEN, CHRIS OATES, and SIMO SÄRKKÄ .....	5882
<b>Dual Swap Disentangling,</b> ZUNLEI FENG, XINCHAO WANG, and CHENGLONG KE, AN-XIANG ZENG, DACHENG TAO, MINGLI SONG .....	5894
<b>Diverse Ensemble Evolution: Curriculum Data-Model Marriage,</b> TIANYI ZHOU, SHENGJIE WANG, and JEFF BILMES .....	5905
<b>Binary Classification from Positive-Confidence Data,</b> TAKASHI ISHIDA, GANG NIU, and MASASHI SUGIYAMA .....	5917
<b>Deep Generative Models for Distribution-Preserving Lossy Compression,</b> MICHAEL TSCHANNEN, EIRIKUR AGUSTSSON, and MARIO LUCIC .....	5929
<b>Exact natural gradient in deep linear networks and its application to the nonlinear case,</b> ALBERTO BERNACCHIA, MÁTÉ LENGYEL, and GUILLAUME HENNEQUIN .....	5941
<b>Constructing Fast Network through Deconstruction of Convolution,</b> YUNHO JEON, JUNMO KIM .....	5951
<b>Memory Replay GANs: learning to generate images from new categories without forgetting,</b> CHENSHEN WU, LUIS HERRANZ, and XIALEI LIU, YAXING WANG, JOOST WEIJER, BOGDAN RADUCANU .....	5962
<b>The Convergence of Sparsified Gradient Methods,</b> DAN ALISTARH, TORSTEN HOEFLER, and MIKAEL JOHANSSON, NIKOLA KONSTANTINOV, SARIT KHIRIRAT, CEDRIC RENGGLI .....	5973
<b>Automating Bayesian optimization with Bayesian optimization,</b> GUSTAVO MALKOMES, ROMAN GARNETT .....	5984
<b>Stacked Semantics-Guided Attention Model for Fine-Grained Zero-Shot Learning,</b> YUNLONG YU, ZHONG JI, and YANWEI FU, JICHANG GUO, YANWEI PANG, ZHONGFEI ZHANG .....	5995
<b>Dirichlet-based Gaussian Processes for Large-scale Calibrated Classification,</b> DIMITRIOS MILIOS, RAFFAELLO CAMORIANO, and PIETRO MICHIARDI, LORENZO ROSASCO, MAURIZIO FILIPPONE .....	6005
<b>Multi-Task Zipping via Layer-wise Neuron Sharing,</b> XIAOXI HE, ZIMU ZHOU, and LOTHAR THIELE .....	6016
<b>Dimensionally Tight Bounds for Second-Order Hamiltonian Monte Carlo,</b> OREN MANGOUBI, NISHEETH VISHNOI .....	6027
<b>Approximation algorithms for stochastic clustering,</b> DAVID HARRIS, SHI LI, and ARAVIND SRINIVASAN, KHOA TRINH, THOMAS PENSYL .....	6038
<b>Evolutionary Stochastic Gradient Descent for Optimization of Deep Neural Networks,</b> XIAODONG CUI, WEI ZHANG, and ZOLTAN TUSKE, MICHAEL PICHENY .....	6048
<b>Learning to Infer Graphics Programs from Hand-Drawn Images,</b> KEVIN ELLIS, DANIEL RITCHIE, and ARMANDO SOLAR-LEZAMA, JOSHUA TENENBAUM .....	6059
<b>Graphical Generative Adversarial Networks,</b> CHONGXUAN LI, MAX WELLING, and JUN ZHU, BO ZHANG .....	6069

<b>Variational Learning on Aggregate Outputs with Gaussian Processes,</b> HO LAW, DINO SEJDINOVIC, and EWAN CAMERON, TIM LUCAS, SETH FLAXMAN, KATHERINE BATTLE, KENJI FUKUMIZU .....	6081
<b>MacNet: Transferring Knowledge from Machine Comprehension to Sequence-to-Sequence Models,</b> BOYUAN PAN, YAZHENG YANG, and HAO LI, ZHOU ZHAO, YUETING ZHUANG, DENG CAI, XIAOFEI HE .....	6092
<b>Poison Frogs! Targeted Clean-Label Poisoning Attacks on Neural Networks,</b> ALI SHAFABI, W. HUANG, and MAHYAR NAJIBI, OCTAVIAN SUCIU, CHRISTOPH STUDER, TUDOR DUMITRAS, TOM GOLDSTEIN .....	6103
<b>Information Constraints on Auto-Encoding Variational Bayes,</b> ROMAIN LOPEZ, JEFFREY REGIER, and MICHAEL JORDAN, NIR YOSEF .....	6114
<b>Recurrent Transformer Networks for Semantic Correspondence,</b> SEUNGRYONG KIM, STEPHEN LIN, and SANGRYUL JEON, DONGBO MIN, KWANGHOON SOHN .....	6126
<b>Online convex optimization for cumulative constraints,</b> JIANJUN YUAN, ANDREW LAMPERSKI .....	6137
<b>Predict Responsibly: Improving Fairness and Accuracy by Learning to Defer,</b> DAVID MADRAS, TONIANN PITASSI, and RICHARD ZEMEL .....	6147
<b>Deep State Space Models for Unconditional Word Generation,</b> FLORIAN SCHMIDT, THOMAS HOFMANN .....	6158
<b>ResNet with one-neuron hidden layers is a Universal Approximator,</b> LIN HONGZHOU, STEFANIE JEGELKA .....	6169
<b>Transfer of Value Functions via Variational Methods,</b> ANDREA TIRINZONI, RAFAEL RODRIGUEZ, and MARCELLO RESTELLI .....	6179
<b>The Cluster Description Problem - Complexity Results, Formulations and Approximations,</b> IAN DAVIDSON, ANTOINE GOURRU, and S RAVI .....	6190
<b>Sharp Bounds for Generalized Uniformity Testing,</b> ILIAS DIAKONIKOLAS, DANIEL KANE, and ALISTAIR STEWART .....	6201
<b>Deep Neural Networks with Box Convolutions,</b> EGOR BURKOV, VICTOR LEMPITSKY .....	6211
<b>Learning towards Minimum Hyperspherical Energy,</b> WEIYANG LIU, RONGMEI LIN, and ZHEN LIU, LIXIN LIU, ZHIDING YU, BO DAI, LE SONG .....	6222
<b>LF-Net: Learning Local Features from Images,</b> YUKI ONO, EDUARD TRULLS, and PASCAL FUA, KWANG YI .....	6234
<b>SLANG: Fast Structured Covariance Approximations for Bayesian Deep Learning with Natural Gradient,</b> AARON MISHKIN, FREDERIK KUNSTNER, and DIDRIK NIELSEN, MARK SCHMIDT, MOHAMMAD KHAN .....	6245
<b>Tangent: Automatic differentiation using source-code transformation for dynamically typed array programming,</b> BART MERRIENBOER, DAN MOLDOVAN, and ALEXANDER WILTSCHKO .....	6256
<b>Multi-domain Causal Structure Learning in Linear Systems,</b> AMIREMAD GHASSAMI, NEGAR KIYAVASH, and BIWEI HUANG, KUN ZHANG ..	6266

<b>Privacy Amplification by Subsampling: Tight Analyses via Couplings and Divergences,</b> BORJA BALLE, GILLES BARTHE, and MARCO GABOARDI .....	6277
<b>Exponentially Weighted Imitation Learning for Batched Historical Data,</b> QING WANG, JIECHAO XIONG, and LEI HAN, PENG SUN, HAN LIU, TONG ZHANG .....	6288
<b>Algebraic tests of general Gaussian latent tree models,</b> DENNIS LEUNG, MATHIAS DRTON .....	6298
<b>Navigating with Graph Representations for Fast and Scalable Decoding of Neural Language Models,</b> MINJIA ZHANG, WENHAN WANG, and XIAODONG LIU, JIANFENG GAO, YUXIONG HE .....	6308
<b>Deep Structured Prediction with Nonlinear Output Transformations,</b> COLIN GRABER, OFER MESHI, and ALEXANDER SCHWING .....	6320
<b>Sequential Test for the Lowest Mean: From Thompson to Murphy Sampling,</b> EMILIE KAUFMANN, WOUTER KOOLEN, and AURÉLIEN GARIVIER .....	6332
<b>Distributed Learning without Distress: Privacy-Preserving Empirical Risk Minimization,</b> BARGAV JAYARAMAN, LINGXIAO WANG, and DAVID EVANS, QUANQUAN GU .....	6343
<b>A no-regret generalization of hierarchical softmax to extreme multi-label classification,</b> MAREK WYDMUCH, KALINA JASINSKA, and MIKHAIL KUZNETSOV, ROBERT BUSA-FEKETE, KRZYSZTOF DEMBCZYŃSKI ..	6355
<b>Efficient Formal Safety Analysis of Neural Networks,</b> SHIQI WANG, KEXIN PEI, and JUSTIN WHITEHOUSE, JUNFENG YANG, SUMAN JANA .....	6367
<b>Bayesian Distributed Stochastic Gradient Descent,</b> MICHAEL TENG, FRANK WOOD .....	6378
<b>Visualizing the Loss Landscape of Neural Nets,</b> HAO LI, ZHENG XU, and GAVIN TAYLOR, CHRISTOPH STUDER, TOM GOLDSTEIN .....	6389
<b>The Limits of Post-Selection Generalization,</b> JONATHAN ULLMAN, ADAM SMITH, and KOBBI NISSIM, URI STEMMER, THOMAS STEINKE .....	6400
<b>Graph Convolutional Policy Network for Goal-Directed Molecular Graph Generation,</b> JIAXUAN YOU, BOWEN LIU, and REX YING, VIJAY PANDE, JURE LESKOVEC .....	6410
<b>On Controllable Sparse Alternatives to Softmax,</b> ANIRBAN LAHA, SANEEM CHEMMENGATH, and PRIYANKA AGRAWAL, MITESH KHAPRA, KARTHIK SANKARANARAYANAN, HARISH RAMASWAMY .....	6422
<b>L4: Practical loss-based stepsize adaptation for deep learning,</b> MICHAL ROLINEK, GEORG MARTIUS .....	6433
<b>Learning Latent Subspaces in Variational Autoencoders,</b> JACK KLYS, JAKE SNELL, and RICHARD ZEMEL .....	6444
<b>Turbo Learning for CaptionBot and DrawingBot,</b> QIUYUAN HUANG, PENGCHUAN ZHANG, and DAPENG WU, LEI ZHANG .....	6455
<b>Learning to Teach with Dynamic Loss Functions,</b> LIJUN WU, FEI TIAN, and YINGCE XIA, YANG FAN, TAO QIN, JIAN-HUANG LAI, TIEYAN LIU ..	6466
<b>Multi-View Silhouette and Depth Decomposition for High Resolution 3D Object Representation,</b> EDWARD SMITH, SCOTT FUJIMOTO, and DAVID MEGER .....	6478

<b>Size-Noise Tradeoffs in Generative Networks,</b> BOLTON BAILEY, MATUS TELGARSKY .....	6489
<b>Online Adaptive Methods, Universality and Acceleration,</b> Kfir LEVY, ALP YURTSEVER, and VOLKAN CEVHER .....	6500
<b>Compact Generalized Non-local Network,</b> KAIYU YUE, MING SUN, and YUCHEN YUAN, FENG ZHOU, ERRUI DING, FUXIN XU .....	6510
<b>On the Local Hessian in Back-propagation,</b> HUIHUAI ZHANG, WEI CHEN, and TIEYAN LIU .....	6520
<b>The Everlasting Database: Statistical Validity at a Fair Price,</b> BLAKE WOODWORTH, VITALY FELDMAN, and SAHARON ROSSET, NATHAN SREBRO .....	6531
<b>Lipschitz-Margin Training: Scalable Certification of Perturbation Invariance for Deep Neural Networks,</b> YUSUKE TSUZUKU, ISSEI SATO, and MASASHI SUGIYAMA .....	6541
<b>Proximal SCOPE for Distributed Sparse Learning,</b> SHEN-YI ZHAO, GONG-DUO ZHANG, and MING-WEI LI, WU-JUN LI .....	6551
<b>On Coresets for Logistic Regression,</b> ALEXANDER MUNTEANU, CHRIS SCHWIEGELSHOHN, and CHRISTIAN SOHLER, DAVID WOODRUFF .....	6561
<b>Neural Ordinary Differential Equations,</b> RICKY CHEN, YULIA RUBANOVA, and JESSE BETTENCOURT, DAVID DUVENAUD .....	6571
<b>Unsupervised Learning of Artistic Styles with Archetypal Style Analysis,</b> DAAN WYNEN, CORDELIA SCHMID, and JULIEN MAIRAL .....	6584
<b>Approximating Real-Time Recurrent Learning with Random Kronecker Factors,</b> ASIER MUJIKA, FLORIAN MEIER, and ANGELIKA STEGER .....	6594
<b>Contamination Attacks and Mitigation in Multi-Party Machine Learning,</b> JAMIE HAYES, OLGA OHRIMENKO .....	6604
<b>An Improved Analysis of Alternating Minimization for Structured Multi-Response Regression,</b> SHENG CHEN, ARINDAM BANERJEE .....	6616
<b>Incorporating Context into Language Encoding Models for fMRI,</b> SHAILEE JAIN, ALEXANDER HUTH .....	6628
<b>CatBoost: unbiased boosting with categorical features,</b> LIUDMILA PROKHORENKOVA, GLEB GUSEV, and ALEKSANDR VOROBEV, ANNA DOROGUSH, ANDREY GULIN .....	6638
<b>Query K-means Clustering and the Double Dixie Cup Problem,</b> I CHIEN, CHAO PAN, and OLGICA MILENKOVIC .....	6649
<b>Training Neural Networks Using Features Replay,</b> ZHOUYUAN HUO, BIN GU, and HENG HUANG .....	6659
<b>Modeling Dynamic Missingness of Implicit Feedback for Recommendation,</b> MENGHAN WANG, MINGMING GONG, and XIAOLIN ZHENG, KUN ZHANG .....	6669
<b>Representation Learning of Compositional Data,</b> MARTA AVALOS-FERNANDEZ, RICHARD NOCK, and CHENG ONG, JULIEN ROUAR, KE SUN .....	6679

<b>Model-based targeted dimensionality reduction for neuronal population data</b> , MIKIO AOI, JONATHAN PILLOW .....	6690
<b>On gradient regularizers for MMD GANs</b> , MICHAEL ARBEL, DOUGAL SUTHERLAND, and MIKOLAJ BINKOWSKI, ARTHUR GRETTON .....	6700
<b>Heterogeneous Multi-output Gaussian Process Prediction</b> , PABLO MORENO-MUNOZ, ANTONIO RODRÍGUEZ, and MAURICIO ALVAREZ .....	6711
<b>Large-Scale Stochastic Sampling from the Probability Simplex</b> , JACK BAKER, PAUL FEARNHEAD, and EMILY FOX, CHRISTOPHER NEMETH ..	6721
<b>Policy Regret in Repeated Games</b> , RAMAN ARORA, MICHAEL DINITZ, and TEODOR MARINOV, MEHRYAR MOHRI .....	6732
<b>A Theory-Based Evaluation of Nearest Neighbor Models Put Into Practice</b> , HENDRIK FICHTENBERGER, DENNIS ROHDE .....	6742
<b>Banach Wasserstein GAN</b> , JONAS ADLER, SEBASTIAN LUNZ .....	6754
<b>Provable Gaussian Embedding with One Observation</b> , MING YU, ZHUORAN YANG, and TUO ZHAO, MLADEN KOLAR, ZHAORAN WANG .....	6764
<b>BRITS: Bidirectional Recurrent Imputation for Time Series</b> , WEI CAO, DONG WANG, and JIAN LI, HAO ZHOU, LEI LI, YITAN LI .....	6775
<b>M-Walk: Learning to Walk over Graphs using Monte Carlo Tree Search</b> , YELONG SHEN, JIANSHU CHEN, and PO-SEN HUANG, YUQING GUO, JIANFENG GAO .....	6786
<b>Extracting Relationships by Multi-Domain Matching</b> , YITONG LI, MICHAEL MURIAS, and GERALDINE DAWSON, DAVID CARLSON .....	6798
<b>Efficient Gradient Computation for Structured Output Learning with Rational and Tropical Losses</b> , CORINNA CORTES, VITALY KUZNETSOV, and MEHRYAR MOHRI, DMITRY STORCHEUS, SCOTT YANG .....	6810
<b>Generative Probabilistic Novelty Detection with Adversarial Autoencoders</b> , STANISLAV PIDHORSKYI, RANYA ALMOHSEN, and GIANFRANCO DORETTO .....	6822
<b>Diminishing Returns Shape Constraints for Interpretability and Regularization</b> , MAYA GUPTA, DARA BAHRI, and ANDREW COTTER, KEVIN CANINI .....	6834
<b>Scalable Hyperparameter Transfer Learning</b> , VALERIO PERRONE, RODOLPHE JENATTON, and MATTHIAS SEEGER, CEDRIC ARCHAMBEAU .....	6845
<b>Stochastic Nonparametric Event-Tensor Decomposition</b> , SHANDIAN ZHE, YISHUAI DU .....	6856
<b>Scaling Gaussian Process Regression with Derivatives</b> , DAVID ERIKSSON, KUN DONG, and ERIC LEE, DAVID BINDEL, ANDREW WILSON .....	6867
<b>Differentially Private Testing of Identity and Closeness of Discrete Distributions</b> , JAYADEV ACHARYA, ZITENG SUN, and HUANYU ZHANG .....	6878
<b>Bayesian Adversarial Learning</b> , NANYANG YE, ZHANXING ZHU .....	6892
<b>Efficient Convex Completion of Coupled Tensors using Coupled Nuclear Norms</b> , KISHAN WIMALAWARNE, HIROSHI MAMITSUKA .....	6902
<b>Maximizing Induced Cardinality Under a Determinantal Point Process</b> , JENNIFER GILLENWATER, ALEX KULESZA, and ZELDA MARIET, SERGEI VASSILVTSKII .....	6911

<b>Causal Inference with Noisy and Missing Covariates via Matrix Factorization,</b> NATHAN KALLUS, XIAOJIE MAO, and MADELEINE UDELL ..	6921
<b>rho-POMDPs have Lipschitz-Continuous epsilon-Optimal Value Functions,</b> MATHIEU FEHR, OLIVIER BUFFET, and VINCENT THOMAS, JILLES DIBANGOYE .....	6933
<b>Online Structure Learning for Feed-Forward and Recurrent Sum-Product Networks,</b> AGASTYA KALRA, ABDULLAH RASHWAN, and WILSON HSU, PASCAL POUPART, PRASHANT DOSHI, GEORGE TRIMPONIAS ..	6944
<b>Uncertainty Sampling is Preconditioned Stochastic Gradient Descent on Zero-One Loss,</b> STEPHEN MUSSMANN, PERCY LIANG .....	6955
<b>A Probabilistic U-Net for Segmentation of Ambiguous Images,</b> SIMON KOHL, BERNARDINO ROMERA-PAREDES, and CLEMENS MEYER, JEFFREY FAUW, JOSEPH LEDSAM, KLAUS MAIER-HEIN, S. ESLAMI, DANILO REZENDE, OLAF RONNEBERGER .....	6965
<b>Unorganized Malicious Attacks Detection,</b> MING PANG, WEI GAO, and MIN TAO, ZHI-HUA ZHOU .....	6976
<b>Causal Inference via Kernel Deviance Measures,</b> JOVANA MITROVIC, DINO SEJDINOVIC, and YEE TEH .....	6986
<b>Bayesian Alignments of Warped Multi-Output Gaussian Processes,</b> MARKUS KAISER, CLEMENS OTTE, and THOMAS RUNKLER, CARL EK .....	6995
<b>Hybrid Macro/Micro Level Backpropagation for Training Deep Spiking Neural Networks,</b> YINGYEZHE JIN, WENRUI ZHANG, and PENG LI .....	7005
<b>Gen-Oja: Simple &amp; Efficient Algorithm for Streaming Generalized Eigenvector Computation,</b> KUSH BHATIA, ALDO PACCHIANO, and NICOLAS FLAMMARION, PETER BARTLETT, MICHAEL JORDAN .....	7016
<b>Efficient online algorithms for fast-rate regret bounds under sparsity,</b> PIERRE GAILLARD, OLIVIER WINTENBERGER .....	7026
<b>GILBO: One Metric to Measure Them All,</b> ALEX ALEMI, IAN FISCHER .....	7037
<b>Predictive Uncertainty Estimation via Prior Networks,</b> ANDREY MALININ, MARK GALES .....	7047
<b>Dual Policy Iteration,</b> WEN SUN, GEOFFREY GORDON, and BYRON BOOTS, J. BAGNELL .....	7059
<b>A probabilistic population code based on neural samples,</b> SABYASACHI SHIVKUMAR, RICHARD LANGE, and ANKANI CHATTORAJ, RALF HAEFNER .....	7070
<b>Manifold-tiling Localized Receptive Fields are Optimal in Similarity-preserving Neural Networks,</b> ANIRVAN SENGUPTA, CENGIZ PEHLEVAN, and MARIANO TEPPER, ALEX GENKIN, DMITRI CHKLOVSKII .....	7080
<b>On the Convergence and Robustness of Training GANs with Regularized Optimal Transport,</b> MAZIAR SANJABI, JIMMY BA, and MEISAM RAZAVIYAYN, JASON LEE .....	7091
<b>Model-Agnostic Private Learning,</b> RAEF BASSILY, ABHRADEEP THAKURTA, and OM THAKKAR .....	7102

<b>Constrained Generation of Semantically Valid Graphs via Regularizing Variational Autoencoders,</b> TENGFEI MA, JIE CHEN, and CAO XIAO .....	7113
<b>Provably Correct Automatic Sub-Differentiation for Qualified Programs,</b> SHAM KAKADE, JASON LEE .....	7125
<b>Deep Homogeneous Mixture Models: Representation, Separation, and Approximation,</b> PRIYANK JAINI, PASCAL POUPART, and YAOLIANG YU .....	7136
<b>Parameters as interacting particles: long time convergence and asymptotic error scaling of neural networks,</b> GRANT ROTSKOFF, ERIC VANDEN-EIJNDEN .....	7146
<b>Hierarchical Reinforcement Learning for Zero-shot Generalization with Subtask Dependencies,</b> SUNGRYULL SOHN, JUNHYUK OH, and HONGLAK LEE .....	7156
<b>A Simple Unified Framework for Detecting Out-of-Distribution Samples and Adversarial Attacks,</b> KIMIN LEE, KIBOK LEE, and HONGLAK LEE, JINWOO SHIN .....	7167
<b>End-to-End Differentiable Physics for Learning and Control,</b> FILIPE BELBUTE-PERES, KEVIN SMITH, and KELSEY ALLEN, JOSHUA TENENBAUM, ZICO KOLTER .....	7178
<b>BRUNO: A Deep Recurrent Model for Exchangeable Data,</b> IRYNA KORSHUNOVA, JONAS DEGRAVE, and FERENC HUSZAR, YARIN GAL, ARTHUR GRETTON, JONI DAMBRE .....	7190
<b>Stimulus domain transfer in recurrent models for large scale cortical population prediction on video,</b> FABIAN SINZ, ALEXANDER ECKER, and PAUL FAHEY, EDGAR WALKER, ERICK COBOS, EMMANOUIL FROUDARAKIS, DIMITRI YATSENKO, XAQ PITKOW, JACOB REIMER, ANDREAS TOLIAS .....	7199
<b>Distributed Multi-Player Bandits - a Game of Thrones Approach,</b> ILAI BISTRITZ, AMIR LESHEM .....	7211
<b>Efficient Loss-Based Decoding on Graphs for Extreme Classification,</b> ITAY EVRON, EDWARD MOROSHKO, and KOBY CRAMMER ..	7222
<b>Chaining Mutual Information and Tightening Generalization Bounds,</b> AMIR ASADI, EMMANUEL ABBE, and SERGIO VERDÚ .....	7234
<b>Implicit Probabilistic Integrators for ODEs,</b> ONUR TEYMUR, HAN LIE, and TIM SULLIVAN, BEN CALDERHEAD .....	7244
<b>Learning Attentional Communication for Multi-Agent Cooperation,</b> JIECHUAN JIANG, ZONGQING LU .....	7254
<b>Training Deep Models Faster with Robust, Approximate Importance Sampling,</b> TYLER JOHNSON, CARLOS GUESTRIN .....	7265
<b>Bandit Learning with Implicit Feedback,</b> YI QI, QINGYUN WU, and HONGNING WANG, JIE TANG, MAOSONG SUN .....	7276
<b>Unsupervised Text Style Transfer using Language Models as Discriminators,</b> ZICHAO YANG, ZHITING HU, and CHRIS DYER, ERIC XING, TAYLOR BERG-KIRKPATRICK .....	7287

<b>Relational recurrent neural networks</b> , ADAM SANTORO, RYAN FAULKNER, and DAVID RAPOSO, JACK RAE, MIKE CHRZANOWSKI, THEOPHANE WEBER, DAAN WIERSTRA, ORIOL VINYALS, RAZVAN PASCANU, TIMOTHY LILICRAP .....	7299
<b>Streaming Kernel PCA with <math>\tilde{O}(\sqrt{n})</math> Random Features</b> , ENAYAT ULLAH, POORYA MIANJY, and TEODOR MARINOV, RAMAN ARORA .....	7311
<b>REFUEL: Exploring Sparse Features in Deep Reinforcement Learning for Fast Disease Diagnosis</b> , YU-SHAO PENG, KAI-FU TANG, and HSUAN-TIEN LIN, EDWARD CHANG .....	7322
<b>Bayesian Model-Agnostic Meta-Learning</b> , JAESIK YOON, TAESUP KIM, and OUSMANE DIA, SUNGWOONG KIM, YOSHUA BENGIO, SUNGJIN AHN .....	7332
<b>Disconnected Manifold Learning for Generative Adversarial Networks</b> , MAHYAR KHAYATKHOEI, MANEESH SINGH, and AHMED ELGAMMAL .....	7343
<b>Unsupervised Cross-Modal Alignment of Speech and Text Embedding Spaces</b> , YU-AN CHUNG, WEI-HUNG WENG, and SCHRASING TONG, JAMES GLASS .....	7354
<b>Learning Signed Determinantal Point Processes through the Principal Minor Assignment Problem</b> , VICTOR-EMMANUEL BRUNEL ..	7365
<b>Out-of-Distribution Detection using Multiple Semantic Label Representations</b> , GABI SHALEV, YOSSI ADI, and JOSEPH KESHET .....	7375
<b>Stochastic Chebyshev Gradient Descent for Spectral Optimization</b> , INSU HAN, HAIM AVRON, and JINWOO SHIN .....	7386
<b>Revisiting <math>(\epsilon, \gamma, \tau)</math>-similarity learning for domain adaptation</b> , SOFIEN DHOUIB, IEVGEN REDKO .....	7397
<b>How to tell when a clustering is (approximately) correct using convex relaxations</b> , MARINA MEILA .....	7407
<b>Constant Regret, Generalized Mixability, and Mirror Descent</b> , ZAKARIA MHAMMEDI, ROBERT WILLIAMSON .....	7419
<b>A Bayesian Approach to Generative Adversarial Imitation Learning</b> , WONSEOK JEON, SEOKIN SEO, and KEE-EUNG KIM .....	7429
<b>Plug-in Estimation in High-Dimensional Linear Inverse Problems: A Rigorous Analysis</b> , ALYSON FLETCHER, PARTHE PANDIT, and SUNDEEP RANGAN, SUBRATA SARKAR, PHILLIP SCHNITER .....	7440
<b>Constrained Cross-Entropy Method for Safe Reinforcement Learning</b> , MIN WEN, UFUK TOPCU .....	7450
<b>Multi-Agent Generative Adversarial Imitation Learning</b> , JIAMING SONG, HONGYU REN, and DORSA SADIGH, STEFANO ERMON .....	7461
<b>Adaptive Learning with Unknown Information Flows</b> , YONATAN GUR, AHMADREZA MOMENI .....	7473
<b>Forecasting Treatment Responses Over Time Using Recurrent Marginal Structural Networks</b> , BRYAN LIM, AHMED ALAA, and MIHAELA SCHAAR .....	7483
<b>Generative modeling for protein structures</b> , NAMRATA ANAND, PO-SSU HUANG .....	7494



<b>Inference in Deep Gaussian Processes using Stochastic Gradient Hamiltonian Monte Carlo</b> , MARTON HAVASI, JOSE HERNANDEZ-LOBATO, and JUAN FUENTES .....	7506
<b>Knowledge Distillation by On-the-Fly Native Ensemble</b> , XU LAN, XIATIAN ZHU, and SHAOGANG GONG .....	7517
<b>Non-Adversarial Mapping with VAEs</b> , YEDID HOSHEN .....	7528
<b>Generalisation in humans and deep neural networks</b> , ROBERT GEIRHOS, CARLOS TEMME, and JONAS RAUBER, HEIKO SCHÜTT, MATTHIAS BETHGE, FELIX WICHMANN .....	7538
<b>Towards Text Generation with Adversarially Learned Neural Outlines</b> , SANDEEP SUBRAMANIAN, SAI RAJESHWAR, and ALESSANDRO SORDONI, ADAM TRISCHLER, AARON COURVILLE, CHRISTOPHER PAL .....	7551
<b>cpSGD: Communication-efficient and differentially-private distributed SGD</b> , NAMAN AGARWAL, ANANDA SURESH, and FELIX YU, SANJIV KUMAR, BRENDAN MCMAHAN .....	7564
<b>GPyTorch: Blackbox Matrix-Matrix Gaussian Process Inference with GPU Acceleration</b> , JACOB GARDNER, GEOFF PLEISS, and KILIAN WEINBERGER, DAVID BINDEL, ANDREW WILSON .....	7576
<b>Diffusion Maps for Textual Network Embedding</b> , XINYUAN ZHANG, YITONG LI, and DINGHAN SHEN, LAWRENCE DUKE .....	7587
<b>Simple, Distributed, and Accelerated Probabilistic Programming</b> , DUSTIN TRAN, MATTHEW HOFFMAN, and DAVE MOORE, CHRISTOPHER SUTER, SRINIVAS VASUDEVAN, ALEXEY RADUL, MATTHEW JOHNSON, RIF SAUROUS .....	7598
<b>VideoCapsuleNet: A Simplified Network for Action Detection</b> , KEVIN DUARTE, YOGESH RAWAT, and MUBARAK SHAH .....	7610
<b>Rectangular Bounding Process</b> , XUHUI FAN, BIN LI, and SCOTT SISSON .....	7620
<b>Improved Algorithms for Collaborative PAC Learning</b> , HUY NGUYEN, LYDIA ZAKYNTHINO .....	7631
<b>Sparse Attentive Backtracking: Temporal Credit Assignment Through Reminding</b> , NAN KE, ANIRUDH GOYAL, and OLEXA BILANIUK, JONATHAN BINAS, MICHAEL MOZER, CHRISTOPHER PAL, YOSHUA BENGIO ..	7640
<b>Communication Compression for Decentralized Training</b> , HANLIN TANG, SHAODUO GAN, and CE ZHANG, TONG ZHANG, JI LIU .....	7652
<b>Depth-Limited Solving for Imperfect-Information Games</b> , NOAM BROWN, TUOMAS SANDHOLM, and BRANDON AMOS .....	7663
<b>Training Deep Neural Networks with 8-bit Floating Point Numbers</b> , NAIGANG WANG, JUNGWOOK CHOI, and DANIEL BRAND, CHIA-YU CHEN, KAILASH GOPALAKRISHNAN .....	7675
<b>Scalar Posterior Sampling with Applications</b> , GEORGIOS THEOCHAROUS, ZHENG WEN, and YASIN ABBASI-YADKORI, NIKOS VLASSIS ..	7685
<b>Understanding Batch Normalization</b> , JOHAN BJÖRCK, CARLA GOMES, and BART SELMAN, KILIAN WEINBERGER .....	7694
<b>Adversarial Scene Editing: Automatic Object Removal from Weak Supervision</b> , RAKSHITH SHETTY, MARIO FRITZ, and BERNT SCHIELE .....	7706

<b>Attacks Meet Interpretability: Attribute-steered Detection of Adversarial Samples,</b> GUANHONG TAO, SHIQING MA, and YINGQI LIU, XIANGYU ZHANG .....	7717
<b>On Neuronal Capacity,</b> PIERRE BALDI, ROMAN VERSHYNIN .....	7729
<b>Breaking the Activation Function Bottleneck through Adaptive Parameterization,</b> SEBASTIAN FLENNERHAG, HUNYUN YIN, and JOHN KEANE, MARK ELLIOT .....	7739
<b>Learning Loop Invariants for Program Verification,</b> XUJIE SI, HANJUN DAI, and MUKUND RAGHOTHAMAN, MAYUR NAIK, LE SONG .....	7751
<b>Cooperative Learning of Audio and Video Models from Self-Supervised Synchronization,</b> BRUNO KORBAR, DU TRAN, and LORENZO TORRESANI .....	7763
<b>Towards Robust Interpretability with Self-Explaining Neural Networks,</b> DAVID ALVAREZ-MELIS, TOMMI JAAKKOLA .....	7775
<b>Deep State Space Models for Time Series Forecasting,</b> SYAMA RANGAPURAM, MATTHIAS SEEGER, and JAN GASTHAUS, LORENZO STELLA, BERNIE WANG, TIM JANUSCHOWSKI .....	7785
<b>Constrained Graph Variational Autoencoders for Molecule Design,</b> QI LIU, MILTIADIS ALLAMANIS, and MARC BROCKSCHMIDT, ALEXANDER GAUNT .....	7795
<b>Learning Libraries of Subroutines for Neurally-Guided Bayesian Program Induction,</b> KEVIN ELLIS, LUCAS MORALES, and MATHIAS SABLÉ-MEYER, ARMANDO SOLAR-LEZAMA, JOSHUA TENENBAUM .....	7805
<b>Neural Architecture Optimization,</b> RENQIAN LUO, FEI TIAN, and TAO QIN, ENHONG CHEN, TIEYAN LIU .....	7816
<b>Preference Based Adaptation for Learning Objectives,</b> YAO-XIANG DING, ZHI-HUA ZHOU .....	7828
<b>Distributed <math>k</math>-Clustering for Data with Heavy Noise,</b> SHI LI, XIANGYU GUO .....	7838
<b>Beyond Log-concavity: Provable Guarantees for Sampling Multi-modal Distributions using Simulated Tempering Langevin Monte Carlo,</b> HOLDEN LEE, ANDREJ RISTESKI, and RONG GE .....	7847
<b>A General Method for Amortizing Variational Filtering,</b> JOE MARINO, MILAN CVITKOVIC, and YISONG YUE .....	7857
<b>A Reduction for Efficient LDA Topic Reconstruction,</b> MATTEO ALMANZA, FLAVIO CHERICHETTI, and ALESSANDRO PANCONESI, ANDREA VATTANI .....	7869
<b>Cluster Variational Approximations for Structure Learning of Continuous-Time Bayesian Networks from Incomplete Data,</b> DOMINIK LINZNER, HEINZ KOEPPL .....	7880
<b>RenderNet: A deep convolutional network for differentiable rendering from 3D shapes,</b> THU NGUYEN-PHUOC, CHUAN LI, and STEPHEN BALABAN, YONGLIANG YANG .....	7891
<b>Robust Hypothesis Testing Using Wasserstein Uncertainty Sets,</b> RUI GAO, LIYAN XIE, and YAO XIE, HUAN XU .....	7902

<b>Robust Detection of Adversarial Attacks by Modeling the Intrinsic Properties of Deep Neural Networks,</b> ZHIHAO ZHENG, PENGYU HONG .....	7913
<b>Monte-Carlo Tree Search for Constrained POMDPs,</b> JONGMIN LEE, GEON-HYEONG KIM, and PASCAL POUPART, KEE-EUNG KIM .....	7923
<b>Learning to Repair Software Vulnerabilities with Generative Adversarial Networks,</b> JACOB HARER, ONUR OZDEMIR, and TOMO LAZOVICH, CHRISTOPHER REALE, REBECCA RUSSELL, LOUIS KIM, SANG CHIN .....	7933
<b>Layer-Wise Coordination between Encoder and Decoder for Neural Machine Translation,</b> TIANYU HE, XU TAN, and YINGCE XIA, DI HE, TAO QIN, ZHIBO CHEN, TIEYAN LIU .....	7944
<b>Dirichlet belief networks for topic structure learning,</b> HE ZHAO, LAN DU, and WRAY BUNTINE, MINGYUAN ZHOU .....	7955
<b>Stochastic Expectation Maximization with Variance Reduction,</b> JIANFEI CHEN, JUN ZHU, and YEE TEH, TONG ZHANG .....	7967
<b>Submodular Maximization via Gradient Ascent: The Case of Deep Submodular Functions,</b> WENRUO BAI, WILLIAM NOBLE, and JEFF BILMES .....	7978
<b>The challenge of realistic music generation: modelling raw audio at scale,</b> SANDER DIELEMAN, AARON OORD, and KAREN SIMONYAN .....	7989
<b>Spectral Signatures in Backdoor Attacks,</b> BRANDON TRAN, JERRY LI, and ALEKSANDER MADRY .....	8000
<b>Reward learning from human preferences and demonstrations in Atari,</b> JAN LEIKE, BORJA IBARZ, and DARIO AMODEI, GEOFFREY IRVING, SHANE LEGG .....	8011
<b>Approximate Knowledge Compilation by Online Collapsed Importance Sampling,</b> TAL FRIEDMAN, GUY BROECK .....	8024
<b>Neural Arithmetic Logic Units,</b> ANDREW TRASK, FELIX HILL, and SCOTT REED, JACK RAE, CHRIS DYER, PHIL BLUNSOM .....	8035
<b>Pipe-SGD: A Decentralized Pipelined SGD Framework for Distributed Deep Net Training,</b> YOUJIE LI, MINGCHAO YU, and SONGZE LI, SALMAN AVESTIMEHR, NAM KIM, ALEXANDER SCHWING .....	8045
<b>Improved Expressivity Through Dendritic Neural Networks,</b> XUNDONG WU, XIANGWEN LIU, and WEI LI, QING WU .....	8057
<b>Efficient Anomaly Detection via Matrix Sketching,</b> VATSAL SHARAN, PARIKSHIT GOPALAN, and UDI WIEDER .....	8069
<b>Learning to Specialize with Knowledge Distillation for Visual Question Answering,</b> JONGHWAN MUN, KIMIN LEE, and JINWOO SHIN, BOHYUNG HAN .....	8081
<b>A Lyapunov-based Approach to Safe Reinforcement Learning,</b> YINLAM CHOW, OFIR NACHUM, and EDGAR DUENEZ-GUZMAN, MOHAMMAD GHAVAMZADEH .....	8092
<b>Credit Assignment For Collective Multiagent RL With Global Rewards,</b> DUC NGUYEN, AKSHAT KUMAR, and HOONG LAU .....	8102

<b>Statistical Optimality of Stochastic Gradient Descent on Hard Learning Problems through Multiple Passes,</b> LOUCAS PILLAUD-VIVIEN, ALESSANDRO RUDI, and FRANCIS BACH .....	8114
<b>Does mitigating ML’s impact disparity require treatment disparity?,</b> ZACHARY LIPTON, JULIAN MCAULEY, and ALEXANDRA CHOULDECHOVA .....	8125
<b>Proximal Graphical Event Models,</b> DEBARUN BHATTACHARJYA, DHARMASHANKAR SUBRAMANIAN, and TIAN GAO .....	8136
<b>Bayesian Control of Large MDPs with Unknown Dynamics in Data-Poor Environments,</b> MAHDI IMANI, SEYEDE GHOREISHI, and ULISSES BRAGA-NETO .....	8146
<b>Learning Overparameterized Neural Networks via Stochastic Gradient Descent on Structured Data,</b> YUANZHI LI, YINGYU LIANG ..	8157
<b>Hamiltonian Variational Auto-Encoder,</b> ANTHONY CATERINI, ARNAUD DOUCET, and DINO SEJDINOVIC .....	8167
<b>Modelling and unsupervised learning of symmetric deformable object categories,</b> JAMES THEWLIS, HAKAN BILEN, and ANDREA VEDALDI	8178
<b>Graphical model inference: Sequential Monte Carlo meets deterministic approximations,</b> FREDRIK LINDSTEN, JOUNI HELSKE, and MATTI VIHOLA .....	8190
<b>Statistical mechanics of low-rank tensor decomposition,</b> JONATHAN KADMON, SURYA GANGULI .....	8201
<b>Variational Bayesian Monte Carlo,</b> LUIGI ACERBI .....	8213
<b>Sample-Efficient Reinforcement Learning with Stochastic Ensemble Value Expansion,</b> JACOB BUCKMAN, DANIJAR HAFNER, and GEORGE TUCKER, EUGENE BREVDO, HONGLAK LEE .....	8224
<b>Efficient Online Portfolio with Logarithmic Regret,</b> HAIPENG LUO, CHEN-YU WEI, and KAI ZHENG .....	8235
<b>Algorithms and Theory for Multiple-Source Adaptation,</b> JUDY HOFFMAN, MEHRYAR MOHRI, and NINGSHAN ZHANG .....	8246
<b>Online Reciprocal Recommendation with Theoretical Performance Guarantees,</b> CLAUDIO GENTILE, NIKOS PAROTSIDIS, and FABIO VITALE ...	8257
<b>The promises and pitfalls of Stochastic Gradient Langevin Dynamics,</b> NICOLAS BROSSE, ALAIN DURMUS, and ERIC MOULINES .....	8268
<b>How SGD Selects the Global Minima in Over-parameterized Learning: A Dynamical Stability Perspective,</b> LEI WU, CHAO MA, and WEINAN E .....	8279
<b>Differentiable MPC for End-to-end Planning and Control,</b> BRANDON AMOS, IVAN JIMENEZ, and JACOB SACKS, BYRON BOOTS, ZICO KOLTER .....	8289
<b>Bilevel learning of the Group Lasso structure,</b> JORDAN FRECON, SAVERIO SALZO, and MASSIMILIANO PONTIL .....	8301
<b>Constructing Unrestricted Adversarial Examples with Generative Models,</b> YANG SONG, RUI SHU, and NATE KUSHMAN, STEFANO ERMON ....	8312

<b>Information-theoretic Limits for Community Detection in Network Models</b> , CHUYANG KE, JEAN HONORIO .....	8324
<b>Learning Conditioned Graph Structures for Interpretable Visual Question Answering</b> , WILL NORCLIFFE-BROWN, EFSTATHIOS VAFEIAS, and SARAH PARISOT .....	8334
<b>Distributionally Robust Graphical Models</b> , RIZAL FATHONY, ASHKAN REZAEI, and MOHAMMAD BASHIRI, XINHUA ZHANG, BRIAN ZIEBART .....	8344
<b>Transfer Learning with Neural AutoML</b> , CATHERINE WONG, NEIL HOULSBY, and YIFENG LU, ANDREA GESMUNDO .....	8356
<b>Stochastic Primal-Dual Method for Empirical Risk Minimization with <math>O(1)</math> Per-Iteration Complexity</b> , CONGHUI TAN, TONG ZHANG, and SHIQIAN MA, JI LIU .....	8366
<b>On preserving non-discrimination when combining expert advice</b> , AVRIM BLUM, SURIYA GUNASEKAR, and THODORIS LYKOURIS, NATHAN SREBRO .....	8376
<b>Learning to Play With Intrinsically-Motivated, Self-Aware Agents</b> , NICK HABER, DAMIAN MROWCA, and STEPHANIE WANG, LI FEI-FEI, DANIEL YAMINS .....	8388
<b>Scaling provable adversarial defenses</b> , ERIC WONG, FRANK SCHMIDT, and JAN METZEN, ZICO KOLTER .....	8400
<b>Deep Network for the Integrated 3D Sensing of Multiple People in Natural Images</b> , ANDREI ZANFIR, ELISABETA MARINOIU, and MIHAI ZANFIR, ALIN-IONUT POPA, CRISTIAN SMINCHISESCU .....	8410
<b>Almost Optimal Algorithms for Linear Stochastic Bandits with Heavy-Tailed Payoffs</b> , HAN SHAO, XIAOTIAN YU, and IRWIN KING, MICHAEL LYU .....	8420
<b>Data-dependent PAC-Bayes priors via differential privacy</b> , GINTARE DZIUGAITE, DANIEL ROY .....	8430
<b>Deep Poisson gamma dynamical systems</b> , DANDAN GUO, BO CHEN, and HAO ZHANG, MINGYUAN ZHOU .....	8442
<b>Dimensionality Reduction has Quantifiable Imperfections: Two Geometric Bounds</b> , YIK LUI, GAVIN DING, and RUITONG HUANG, ROBERT MCCANN .....	8453
<b>Teaching Inverse Reinforcement Learners via Features and Demonstrations</b> , LUIS HAUG, SEBASTIAN TSCHIATSCHEK, and ADISH SINGLA .....	8464
<b>Wasserstein Distributionally Robust Kalman Filtering</b> , SOROOSH SHAFIEE, VIET NGUYEN, and DANIEL KUHN, PEYMAN ESFAHANI .....	8474
<b>Generalisation of structural knowledge in the hippocampal-entorhinal system</b> , JAMES WHITTINGTON, TIMOTHY MULLER, and SHIRLEY MARK, CASWELL BARRY, TIM BEHRENS .....	8484
<b>Graph Oracle Models, Lower Bounds, and Gaps for Parallel Stochastic Optimization</b> , BLAKE WOODWORTH, JIALEI WANG, and ADAM SMITH, BRENDAN MCMAHAN, NATHAN SREBRO .....	8496
<b>Adversarial Regularizers in Inverse Problems</b> , SEBASTIAN LUNZ, CAROLA SCHOENLIEB, and OZAN ÖKTEM .....	8507

<b>Clustering Redemption-Beyond the Impossibility of Kleinberg’s Axioms,</b> VINCENT COHEN-ADDAD, VARUN KANADE, and FREDERIK MALLMANN-TRENN .....	8517
<b>Co-teaching: Robust training of deep neural networks with extremely noisy labels,</b> BO HAN, QUANMING YAO, and XINGRUI YU, GANG NIU, MIAO XU, WEIHUA HU, IVOR TSANG, MASASHI SUGIYAMA .....	8527
<b>Variational Inverse Control with Events: A General Framework for Data-Driven Reward Definition,</b> JUSTIN FU, AVI SINGH, and DIBYA GHOSH, LARRY YANG, SERGEY LEVINE .....	8538
<b>A convex program for bilinear inversion of sparse vectors,</b> ALIREZA AGHASI, ALI AHMED, and PAUL HAND, BABHRU JOSHI .....	8548
<b>Adversarial Multiple Source Domain Adaptation,</b> HAN ZHAO, SHANGHANG ZHANG, and GUANHANG WU, JOSÉ MOURA, JOAO COSTEIRA, GEOFFREY GORDON .....	8559
<b>Neural Tangent Kernel: Convergence and Generalization in Neural Networks,</b> ARTHUR JACOT, FRANCK GABRIEL, and CLEMENT HONGLER .....	8571
<b>Contextual Stochastic Block Models,</b> YASH DESHPANDE, SUBHABRATA SEN, and ANDREA MONTANARI, ELCHANAN MOSSEL .....	8581
<b>A Likelihood-Free Inference Framework for Population Genetic Data using Exchangeable Neural Networks,</b> JEFFREY CHAN, VALERIO PERRONE, and JEFFREY SPENCE, PAUL JENKINS, SARA MATHIESON, YUN SONG .....	8594
<b>Sequential Attend, Infer, Repeat: Generative Modelling of Moving Objects,</b> ADAM KOSIOREK, HYUNJIK KIM, and YEE TEH, INGMAR POSNER .....	8606
<b>Randomized Prior Functions for Deep Reinforcement Learning,</b> IAN OSBAND, JOHN ASLANIDES, and ALBIN CASSIRER .....	8617
<b>Compact Representation of Uncertainty in Clustering,</b> CRAIG GREENBERG, NICHOLAS MONATH, and ARI KOBREN, PATRICK FLAHERTY, ANDREW MCGREGOR, ANDREW MCCALLUM .....	8630
<b>Learning without the Phase: Regularized PhaseMax Achieves Optimal Sample Complexity,</b> FARIBORZ SALEHI, EHSAN ABBASI, and BABAK HASSIBI .....	8641
<b>Multilingual Anchoring: Interactive Topic Modeling and Alignment Across Languages,</b> MICHELLE YUAN, BENJAMIN DURME, and JORDAN BOYD-GRABER .....	8653
<b>Estimators for Multivariate Information Measures in General Probability Spaces,</b> ARMAN RAHIMZAMANI, HIMANSHU ASNANI, and PRAMOD VISWANATH, SREERAM KANNAN .....	8664
<b>DeepPINK: reproducible feature selection in deep neural networks,</b> YANG LU, YINGYING FAN, and JINCHI LV, WILLIAM NOBLE .....	8676
<b>HOUDINI: Lifelong Learning as Program Synthesis,</b> LAZAR VALKOV, DIPAK CHAUDHARI, and AKASH SRIVASTAVA, CHARLES SUTTON, SWARAT CHAUDHURI .....	8687

<b>Searching for Efficient Multi-Scale Architectures for Dense Image Prediction,</b> LIANG-CHIEH CHEN, MAXWELL COLLINS, and YUKUN ZHU, GEORGE PAPANDEOU, BARRET ZOPH, FLORIAN SCHROFF, HARTWIG ADAM, JON SHLENS .....	8699
<b>Orthogonally Decoupled Variational Gaussian Processes,</b> HUGH SALIMBENI, CHING-AN CHENG, and BYRON BOOTS, MARC DEISENROTH .....	8711
<b>Dendritic cortical microcircuits approximate the backpropagation algorithm,</b> JOAO SACRAMENTO, RUI COSTA, and YOSHUA BENGIO, WALTER SENN .....	8721
<b>Learning Plannable Representations with Causal InfoGAN,</b> THANARD KURUTACH, AVIV TAMAR, and GE YANG, STUART RUSSELL, PIETER ABBEEL .....	8733
<b>Uniform Convergence of Gradients for Non-Convex Learning and Optimization,</b> DYLAN FOSTER, AYUSH SEKHARI, and KARTHIK SRIDHARAN .....	8745
<b>Automatic differentiation in ML: Where we are and where we should be going,</b> BART MERRIENBOER, OLIVIER BREULEUX, and ARNAUD BERGERON, PASCAL LAMBLIN .....	8757
<b>A Bayesian Nonparametric View on Count-Min Sketch,</b> DIANA CAI, MICHAEL MITZENMACHER, and RYAN ADAMS .....	8768
<b>Generalized Cross Entropy Loss for Training Deep Neural Networks with Noisy Labels,</b> ZHILU ZHANG, MERT SABUNCU .....	8778
<b>Loss Surfaces, Mode Connectivity, and Fast Ensembling of DNNs,</b> TIMUR GARİPOV, PAVEL IZMAILOV, and DMITRII PODOPRIKHIN, DMITRY VETROV, ANDREW WILSON .....	8789
<b>Flexible neural representation for physics prediction,</b> DAMIAN MROWCA, CHENGXU ZHUANG, and ELIAS WANG, NICK HABER, LI FEI-FEI, JOSHUA TENENBAUM, DANIEL YAMINS .....	8799
<b>Legendre Decomposition for Tensors,</b> MAHITO SUGIYAMA, HIROYUKI NAKAHARA, and KOJI TSUDA .....	8811
<b>Reinforcement Learning of Theorem Proving,</b> CEZARY KALISZYK, JOSEF URBAN, and HENRYK MICHALEWSKI, MIREK OLŠÁK .....	8822
<b>Data Amplification: A Unified and Competitive Approach to Property Estimation,</b> YI HAO, ALON ORLITSKY, and ANANDA SURESH, YIHONG WU .....	8834
<b>Group Equivariant Capsule Networks,</b> JAN LENSSSEN, MATTHIAS FEY, and PASCAL LIBUSCHEWSKI .....	8844
<b>Stein Variational Gradient Descent as Moment Matching,</b> QIANG LIU, DILIN WANG .....	8854
<b>Differential Privacy for Growing Databases,</b> RACHEL CUMMINGS, SARA KREHIEL, and KEVIN LAI, UTHAIPON TANTIPONGPIPAT .....	8864
<b>Exploration in Structured Reinforcement Learning,</b> JUNGSEUL OK, ALEXANDRE PROUTIERE, and DAMIANOS TRANOS .....	8874
<b>A Statistical Recurrent Model on the Manifold of Symmetric Positive Definite Matrices,</b> RUDRASHIS CHAKRABORTY, CHUN-HAO YANG, and XINGJIAN ZHEN, MONAMI BANERJEE, DEREK ARCHER, DAVID VAILLANCOURT, VIKAS SINGH, BABA VEMURI .....	8883

<b>Balanced Policy Evaluation and Learning,</b> NATHAN KALLUS	8895
<b>Distributed Multitask Reinforcement Learning with Quadratic Convergence,</b> RASUL TUTUNOV, DONGHO KIM, and HAITHAM ANMAR	8907
<b>Improving Neural Program Synthesis with Inferred Execution Traces,</b> RICHARD SHIN, ILLIA POLOSUKHIN, and DAWN SONG	8917
<b>Adaptive Path-Integral Autoencoders: Representation Learning and Planning for Dynamical Systems,</b> JUNG-SU HA, YOUNG-JIN PARK, and HYEOK-JOO CHAE, SOON-SEO PARK, HAN-LIM CHOI	8927
<b>Policy-Conditioned Uncertainty Sets for Robust Markov Decision Processes,</b> ANDREA TIRINZONI, MAREK PETRIK, and XIANGLI CHEN, BRIAN ZIEBART	8939
<b>GLoMo: Unsupervised Learning of Transferable Relational Graphs,</b> ZHILIN YANG, JUNBO ZHAO, and BHUWAN DHINGRA, KAIMING HE, WILLIAM COHEN, RUSLAN SALAKHUTDINOV, YANN LECUN	8950
<b>Online Learning of Quantum States,</b> SCOTT AARONSON, XINYI CHEN, and ELAD HAZAN, SATYEN KALE, ASHWIN NAYAK	8962
<b>Wavelet regression and additive models for irregularly spaced data,</b> ASAD HARIS, ALI SHOJAIE, and NOAH SIMON	8973
<b>Inferring Latent Velocities from Weather Radar Data using Gaussian Processes,</b> RICO ANGELL, DANIEL SHELDON	8984
<b>A Structured Prediction Approach for Label Ranking,</b> ANNA KORBA, ALEXANDRE GARCIA, and FLORENCE D'ALCHE-BUC	8994
<b>Efficient High Dimensional Bayesian Optimization with Additivity and Quadrature Fourier Features,</b> MOJMIR MUTNY, ANDREAS KRAUSE	9005
<b>FastGRNN: A Fast, Accurate, Stable and Tiny Kilobyte Sized Gated Recurrent Neural Network,</b> ADITYA KUSUPATI, MANISH SINGH, and KUSH BHATIA, ASHISH KUMAR, PRATEEK JAIN, MANIK VARMA	9017
<b>Reversible Recurrent Neural Networks,</b> MATTHEW MACKAY, PAUL VICOL, and JIMMY BA, ROGER GROSSE	9029
<b>SING: Symbol-to-Instrument Neural Generator,</b> ALEXANDRE DEFOSSEZ, NEIL ZEGHIDOUR, and NICOLAS USUNIER, LEON BOTTOU, FRANCIS BACH	9041
<b>Learning Compressed Transforms with Low Displacement Rank,</b> ANNA THOMAS, ALBERT GU, and TRI DAO, ATRI RUDRA, CHRISTOPHER RE	9052
<b>Theoretical Linear Convergence of Unfolded ISTA and Its Practical Weights and Thresholds,</b> XIAOHAN CHEN, JIALIN LIU, and ZHANGYANG WANG, WOTAO YIN	9061
<b>Iterative Value-Aware Model Learning,</b> AMIR-MASSOUD FARAHMAND	9072
<b>Invariant Representations without Adversarial Training,</b> DANIEL MOYER, SHUYANG GAO, and ROB BREKELMANS, ARAM GALSTYAN, GREG STEEG	9084
<b>Robot Learning in Homes: Improving Generalization and Reducing Dataset Bias,</b> ABHINAV GUPTA, ADITHYAVAIRAVAN MURALI, and DHIRAJ GANDHI, LERREL PINTO	9094



<b>Learning Safe Policies with Expert Guidance,</b> JESSIE HUANG, FA WU, and DOINA PRECUP, YANG CAI .....	9105
<b>Bayesian multi-domain learning for cancer subtype discovery from next-generation sequencing count data,</b> EHSAN HAJIRAMEZANALI, SIAMAK DADANEH, and ALIREZA KARBALAYGHAREH, MINGYUAN ZHOU, XIAONING QIAN .....	9115
<b>Learning SMA<sub>LL</sub> Predictors,</b> VIKAS GARG, OFER DEKEL, and LIN XIAO .....	9125
<b>Phase Retrieval Under a Generative Prior,</b> PAUL HAND, OSCAR LEONG, and VLADISLAV VORONINSKI .....	9136
<b>Quadrature-based features for kernel approximation,</b> MARINA MUNKHOEVA, YERMEK KAPUSHEV, and EVGENY BURNAEV, IVAN OSELEDETS .....	9147
<b>Reducing Network Agnostophobia,</b> AKSHAY DHAMIJA, MANUEL GÜNTHER, and TERRANCE BOULT .....	9157
<b>A Stein variational Newton method,</b> GIANLUCA DETOMMASO, TIANGANG CUI, and YOUSSEF MARZOUK, ALESSIO SPANTINI, ROBERT SCHEICHL .....	9169
<b>Watch Your Step: Learning Node Embeddings via Graph Attention,</b> SAMI ABU-EL-HAJJA, BRYAN PEROZZI, and RAMI AL-RFOU, ALEX ALEMI .....	9180
<b>Visual Reinforcement Learning with Imagined Goals,</b> ASHVIN NAIR, VITCHYR PONG, and MURTAZA DALAL, SHIKHAR BAHL, STEVEN LIN, SERGEY LEVINE .....	9191
<b>Deep Predictive Coding Network with Local Recurrent Processing for Object Recognition,</b> KUAN HAN, HAIGUANG WEN, and YIZHEN ZHANG, DI FU, EUGENIO CULURCIELLO, ZHONGMING LIU .....	9201
<b>PAC-Bayes bounds for stable algorithms with instance-dependent priors,</b> OMAR RIVASPLATA, CSABA SZEPESVARI, and JOHN SHAWE-TAYLOR, EMILIO PARRADO-HERNÁNDEZ, SHILIANG SUN .....	9214
<b>Beyond Grids: Learning Graph Representations for Visual Recognition,</b> YIN LI, ABHINAV GUPTA .....	9225
<b>The Limit Points of (Optimistic) Gradient Descent in Min-Max Optimization,</b> CONSTANTINOS DASKALAKIS, IOANNIS PANAGEAS .....	9236
<b>Coordinate Descent with Bandit Sampling,</b> FARNOOD SALEHI, PATRICK THIRAN, and ELISA CELIS .....	9247
<b>Deep Dynamical Modeling and Control of Unsteady Fluid Flows,</b> JEREMY MORTON, ANTONY JAMESON, and MYKEL KOCHENDERFER, FREDDIE WITHERDEN .....	9258
<b>Confounding-Robust Policy Improvement,</b> NATHAN KALLUS, ANGELA ZHOU .....	9269
<b>Some Considerations on Learning to Explore via Meta-Reinforcement Learning,</b> BRADLY STADIE, GE YANG, and REIN HOUTHOOFT, XI CHEN, YAN DUAN, YUHUAI WU, PIETER ABBEEL, ILYA SUTSKEVER .....	9280
<b>Representer Point Selection for Explaining Deep Neural Networks,</b> CHIH-KUAN YEH, JOON KIM, and IAN YEN, PRADEEP RAVIKUMAR .....	9291

<b>The Effect of Network Width on the Performance of Large-batch Training</b> , LINGJIAO CHEN, HONGYI WANG, and JINMAN ZHAO, DIMITRIS PAPAILIOPOULOS, PARASCHOS KOUTRIS .....	9302
<b>SNIPER: Efficient Multi-Scale Training</b> , BHARAT SINGH, MAHYAR NAJIBI, and LARRY DAVIS .....	9310
<b>The Sample Complexity of Semi-Supervised Learning with Nonparametric Mixture Models</b> , CHEN DAN, LIU LEQI, and BRYON ARAGAM, PRADEEP RAVIKUMAR, ERIC XING .....	9321
<b>Hardware Conditioned Policies for Multi-Robot Transfer Learning</b> , TAO CHEN, ADITHYAVAIRAVAN MURALI, and ABHINAV GUPTA ...	9333
<b>Co-regularized Alignment for Unsupervised Domain Adaptation</b> , ABHISHEK KUMAR, PRASANNA SATTIGERI, and KAHINI WADHAWAN, LEONID KARLINSKY, ROGERIO FERIS, BILL FREEMAN, GREGORY WORNELL .....	9345
<b>Statistical and Computational Trade-Offs in Kernel K-Means</b> , DANIELE CALANDRIELLO, LORENZO ROSASCO .....	9357
<b>Assessing the Scalability of Biologically-Motivated Deep Learning Algorithms and Architectures</b> , SERGEY BARTUNOV, ADAM SANTORO, and BLAKE RICHARDS, LUKE MARRIS, GEOFFREY HINTON, TIMOTHY LILLICRAP .....	9368
<b>Learning Attractor Dynamics for Generative Memory</b> , YAN WU, GREG WAYNE, and KAROL GREGOR, TIMOTHY LILLICRAP .....	9379
<b>The emergence of multiple retinal cell types through efficient coding of natural movies</b> , SAMUEL OCKO, JACK LINDSEY, and SURYA GANGULI, STEPHANE DENY .....	9389
<b>Gather-Excite: Exploiting Feature Context in Convolutional Neural Networks</b> , JIE HU, LI SHEN, and SAMUEL ALBANIE, GANG SUN, ANDREA VEDALDI .....	9401
<b>The Global Anchor Method for Quantifying Linguistic Shifts and Domain Adaptation</b> , ZI YIN, VIN SACHIDANANDA, and BALAJI PRABHAKAR .....	9412
<b>Identification and Estimation of Causal Effects from Dependent Data</b> , ELI SHERMAN, ILYA SHPITSER .....	9424
<b>Deepcode: Feedback Codes via Deep Learning</b> , HYEJI KIM, YIHAN JIANG, and SREERAM KANNAN, SEWOONG OH, PRAMOD VISWANATH .....	9436
<b>Learning and Testing Causal Models with Interventions</b> , JAYADEV ACHARYA, ARNAB BHATTACHARYYA, and CONSTANTINOS DASKALAKIS, SARAVANAN KANDASAMY .....	9447
<b>Implicit Bias of Gradient Descent on Linear Convolutional Networks</b> , SURIYA GUNASEKAR, JASON LEE, and DANIEL SOUDRY, NATHAN SREBRO .....	9461
<b>DAGs with NO TEARS: Continuous Optimization for Structure Learning</b> , XUN ZHENG, BRYON ARAGAM, and PRADEEP RAVIKUMAR, ERIC XING .....	9472
<b>PAC-Bayes Tree: weighted subtrees with guarantees</b> , TIN NGUYEN, SAMORY KPOTUFE .....	9484

<b>Multi-objective Maximization of Monotone Submodular Functions with Cardinality Constraint,</b> RAJAN UDWANI .....	9493
<b>Sanity Checks for Saliency Maps,</b> JULIUS ADEBAYO, JUSTIN GILMER, and MICHAEL MUELLY, IAN GOODFELLOW, MORITZ HARDT, BEEN KIM .....	9505
<b>Probabilistic Model-Agnostic Meta-Learning,</b> CHELSEA FINN, KELVIN XU, and SERGEY LEVINE .....	9516
<b>Reinforcement Learning with Multiple Experts: A Bayesian Model Combination Approach,</b> MIKE GIMELFARB, SCOTT SANNER, and CHI-GUHN LEE .....	9528
<b>e-SNLI: Natural Language Inference with Natural Language Explanations,</b> OANA-MARIA CAMBURU, TIM ROCKTASCHEL, and THOMAS LUKASIEWICZ, PHIL BLUNSOM .....	9539
<b>Fast Approximate Natural Gradient Descent in a Kronecker Factored Eigenbasis,</b> THOMAS GEORGE, CÉSAR LAURENT, and XAVIER BOUTHILLIER, NICOLAS BALLAS, PASCAL VINCENT .....	9550
<b>Learning convex bounds for linear quadratic control policy synthesis,</b> JACK UMENBERGER, THOMAS SCHÖN .....	9561
<b>Neural Proximal Gradient Descent for Compressive Imaging,</b> MORTEZA MARDANI, QINGYUN SUN, and DAVID DONOHO, VARDAN PAPYAN, HATEF MONAJEMI, SHREYAS VASANAWALA, JOHN PAULY .....	9573
<b>Towards Understanding Learning Representations: To What Extent Do Different Neural Networks Learn the Same Representation,</b> LIWEI WANG, LUNJIA HU, and JIAYUAN GU, ZHIQIANG HU, YUE WU, KUN HE, JOHN HOPCROFT .....	9584
<b>Optimal Algorithms for Continuous Non-monotone Submodular and DR-Submodular Maximization,</b> RAD NIAZADEH, TIM ROUGHGARDEN, and JOSHUA WANG .....	9594
<b>An intriguing failing of convolutional neural networks and the CoordConv solution,</b> ROSANNE LIU, JOEL LEHMAN, and PIERO MOLINO, FELIPE SUCH, ERIC FRANK, ALEX SERGEEV, JASON YOSINSKI .....	9605
<b>Trading robust representations for sample complexity through self-supervised visual experience,</b> ANDREA TACCHETTI, STEPHEN VOINEA, and GEORGIOS EVANGELOPOULOS .....	9617
<b>Invertibility of Convolutional Generative Networks from Partial Measurements,</b> FANGCHANG MA, ULAS AYAZ, and SERTAC KARAMAN ....	9628
<b>Ex ante coordination and collusion in zero-sum multi-player extensive-form games,</b> GABRIELE FARINA, ANDREA CELLI, and NICOLA GATTI, TUOMAS SANDHOLM .....	9638
<b>Multi-Agent Reinforcement Learning via Double Averaging Primal-Dual Optimization,</b> HOI-TO WAI, ZHUORAN YANG, and ZHAORAN WANG, MINGYI HONG .....	9649
<b>Improving Online Algorithms via ML Predictions,</b> MANISH PUROHIT, ZOYA SVITKINA, and RAVI KUMAR .....	9661
<b>Global Non-convex Optimization with Discretized Diffusions,</b> MURAT ERDOGDU, LESTER MACKEY, and OHAD SHAMIR .....	9671

<b>Theoretical guarantees for the EM algorithm when applied to mis-specified Gaussian mixture models,</b> RAAZ DWIVEDI, NHAT HO, and KOULIK KHAMARU, MARTIN WAINWRIGHT, MICHAEL JORDAN .....	9681
<b>Coupled Variational Bayes via Optimization Embedding,</b> BO DAI, HANJUN DAI, and NIAO HE, WEIYANG LIU, ZHEN LIU, JIANSHU CHEN, LIN XIAO, LE SONG .....	9690
<b>Improving Explorability in Variational Inference with Annealed Variational Objectives,</b> CHIN-WEI HUANG, SHAWN TAN, and ALEXANDRE LACOSTE, AARON COURVILLE .....	9701
<b>Latent Alignment and Variational Attention,</b> YUNTIAN DENG, YOON KIM, and JUSTIN CHIU, DEMI GUO, ALEXANDER RUSH .....	9712
<b>Towards Deep Conversational Recommendations,</b> LI RAYMOND, SAMIRA KAHOU, and HANNES SCHULZ, VINCENT MICHALSKI, LAURENT CHARLIN, CHRISTOPHER PAL .....	9725
<b>Unsupervised Depth Estimation, 3D Face Rotation and Replacement,</b> JOEL MONIZ, CHRISTOPHER BECKHAM, and SIMON RAJOTTE, SINA HONARI, CHRISTOPHER PAL .....	9736
<b>Generalization Bounds for Uniformly Stable Algorithms,</b> VITALY FELDMAN, JAN VONDRAK .....	9747
<b>Deep Anomaly Detection Using Geometric Transformations,</b> IZHAK GOLAN, RAN EL-YANIV .....	9758
<b>Large Scale computation of Means and Clusters for Persistence Diagrams using Optimal Transport,</b> THEO LACOMBE, MARCO CUTURI, and STEVE OUDOT .....	9770
<b>Entropy Rate Estimation for Markov Chains with Large State Space,</b> YANJUN HAN, JIANTAO JIAO, and CHUAN-ZHENG LEE, TSACHY WEISSMAN, YIHONG WU, TIANCHENG YU .....	9781
<b>Adaptive Methods for Nonconvex Optimization,</b> MANZIL ZAHEER, SASHANK REDDI, and DEVENDRA SACHAN, SATYEN KALE, SANJIV KUMAR ..	9793
<b>Object-Oriented Dynamics Predictor,</b> GUANGXIANG ZHU, ZHIAO HUANG, and CHONGJIE ZHANG .....	9804
<b>Adaptive Skip Intervals: Temporal Abstraction for Recurrent Dynamical Models,</b> ALEXANDER NEITZ, GIAMBATTISTA PARASCANDOLO, and STEFAN BAUER, BERNHARD SCHÖLKOPF .....	9816
<b>Scalable End-to-End Autonomous Vehicle Testing via Rare-event Simulation,</b> MATTHEW O'KELLY, AMAN SINHA, and HONGSEOK NAMKOONG, RUSS TEDRAKE, JOHN DUCHI .....	9827
<b>Reinforcement Learning for Solving the Vehicle Routing Problem,</b> MOHAMMADREZA NAZARI, AFSHIN OROOJLOOY, and LAWRENCE SNYDER, MARTIN TAKAC .....	9839
<b>ATOMO: Communication-efficient Learning via Atomic Sparsification,</b> HONGYI WANG, SCOTT SIEVERT, and SHENGCHAO LIU, ZACHARY CHARLES, DIMITRIS PAPALIOPOULOS, STEPHEN WRIGHT .....	9850
<b>Dynamic Network Model from Partial Observations,</b> ELAHE GHALEBI, BAHARAN MIRZASOLEIMAN, and RADU GROSU, JURE LESKOVEC ..	9862

<b>Life-Long Disentangled Representation Learning with Cross-Domain Latent Homologies</b> , ALESSANDRO ACHILLE, TOM ECCLES, and LOIC MATTHEY, CHRISTOPHER BURGESS, NICK WATTERS, ALEXANDER LERCHNER, IRINA HIGGINS .....	9873
<b>Maximizing acquisition functions for Bayesian optimization</b> , JAMES WILSON, FRANK HUTTER, and MARC DEISENROTH .....	9884
<b>On Markov Chain Gradient Descent</b> , TAO SUN, YUEJIAO SUN, and WOTAO YIN .....	9896
<b>Variance-Reduced Stochastic Gradient Descent on Streaming Data</b> , ELLANGO JOTHIMURUGESAN, ASHRAF TAHMASBI, and PHILLIP GIBBONS, SRIKANTA TIRTHAPURA .....	9906
<b>Online Robust Policy Learning in the Presence of Unknown Adversaries</b> , AARON HAVENS, ZHANHONG JIANG, and SOUMIK SARKAR ...	9916
<b>Uplift Modeling from Separate Labels</b> , IKKO YAMANE, FLORIAN YGER, and JAMAL ATIF, MASASHI SUGIYAMA .....	9927
<b>Learning Invariances using the Marginal Likelihood</b> , MARK WILK, MATTHIAS BAUER, and ST JOHN, JAMES HENSMAN .....	9938
<b>Non-delusional Q-learning and Value Iteration</b> , TYLER LU, DALE SCHUURMANS, and CRAIG BOUTILIER .....	9949
<b>Using Large Ensembles of Control Variates for Variational Inference</b> , TOMAS GEFFNER, JUSTIN DOMKE .....	9960
<b>Post: Device Placement with Cross-Entropy Minimization and Proximal Policy Optimization</b> , YUANXIANG GAO, LI CHEN, and BAOCHUN LI .....	9971
<b>Learning to Reason with Third Order Tensor Products</b> , IMANOL SCHLAG, JÜRGEN SCHMIDHUBER .....	9981
<b>Memory Augmented Policy Optimization for Program Synthesis and Semantic Parsing</b> , CHEN LIANG, MOHAMMAD NOROUZI, and JONATHAN BERANT, QUOC LE, NI LAO .....	9994
<b>Persistence Fisher Kernel: A Riemannian Manifold Kernel for Persistence Diagrams</b> , TAM LE, MAKOTO YAMADA .....	10007
<b>Neural Voice Cloning with a Few Samples</b> , SERCAN ARIK, JITONG CHEN, and KAINAN PENG, WEI PING, YANQI ZHOU .....	10019
<b>Blind Deconvolutional Phase Retrieval via Convex Programming</b> , ALI AHMED, ALIREZA AGHASI, and PAUL HAND .....	10030
<b>Scalable Laplacian K-modes</b> , IMTIAZ ZIKO, ERIC GRANGER, and ISMAIL AYED .....	10041
<b>A Retrieve-and-Edit Framework for Predicting Structured Outputs</b> , TATSUNORI HASHIMOTO, KELVIN GUU, and YONATAN OREN, PERCY LIANG .....	10052
<b>Testing for Families of Distributions via the Fourier Transform</b> , ALISTAIR STEWART, ILIAS DIAKONIKOLAS, and CLEMENT CANONNE .....	10063
<b>Thwarting Adversarial Examples: An <math>L_0</math>-Robust Sparse Fourier Transform</b> , MITALI BAFNA, JACK MURTAGH, and NIKHIL VYAS .....	10075

<b>Blockwise Parallel Decoding for Deep Autoregressive Models,</b> MITCHELL STERN, NOAM SHAZEER, and JAKOB USZKOREIT .....	10086
<b>Low-Rank Tucker Decomposition of Large Tensors Using TensorSketch,</b> OSMAN MALIK, STEPHEN BECKER .....	10096
<b>A Simple Cache Model for Image Recognition,</b> EMIN ORHAN .....	10107
<b>Clebsch-Gordan Nets: a Fully Fourier Space Spherical Convolutional Neural Network,</b> RISI KONDOR, ZHEN LIN, and SHUBHENDU TRIVEDI .....	10117
<b>Bayesian Nonparametric Spectral Estimation,</b> FELIPE TOBAR .....	10127
<b>A Spectral View of Adversarially Robust Features,</b> SHIVAM GARG, VATSAL SHARAN, and BRIAN ZHANG, GREGORY VALIANT .....	10138
<b>Synaptic Strength For Convolutional Neural Network,</b> CHEN LIN, ZHAO ZHONG, and WEI WU, JUNJIE YAN .....	10149
<b>Human-in-the-Loop Interpretability Prior,</b> ISAAC LAGE, ANDREW ROSS, and SAMUEL GERSHMAN, BEEN KIM, FINALE DOSHI-VELEZ .....	10159
<b>Learning To Learn Around A Common Mean,</b> GIULIA DENEVI, CARLO CILIBERTO, and DIMITRIOS STAMOS, MASSIMILIANO PONTIL .....	10169
<b>Backpropagation with Callbacks: Foundations for Efficient and Expressive Differentiable Programming,</b> FEI WANG, JAMES DECKER, and XILUN WU, GREGORY ESSERTEL, TIARK ROMPF .....	10180
<b>Learning with SGD and Random Features,</b> LUIGI CARRATINO, ALESSANDRO RUDI, and LORENZO ROSASCO .....	10192
<b>Total stochastic gradient algorithms and applications in reinforcement learning,</b> PAAVO PARMAS .....	10204
<b>Glow: Generative Flow with Invertible 1x1 Convolutions,</b> DIEDERIK KINGMA, PRAFULLA DHARIWAL .....	10215
<b>Nonparametric Density Estimation under Adversarial Losses,</b> SHASHANK SINGH, ANANYA UPPAL, and BOYUE LI, CHUN-LIANG LI, MANZIL ZAHEER, BARNABAS POCZOS .....	10225
<b>Generalizing Point Embeddings using the Wasserstein Space of Elliptical Distributions,</b> BORIS MUZELLEC, MARCO CUTURI .....	10237
<b>Learning to Share and Hide Intentions using Information Regularization,</b> DJ STROUSE, MAX KLEIMAN-WEINER, and JOSHUA TENENBAUM, MATTHEW BOTVINICK, DAVID SCHWAB .....	10249
<b>Predictive Approximate Bayesian Computation via Saddle Points,</b> YINGXIANG YANG, BO DAI, and NEGAR KIYAVASH, NIAO HE .....	10260
<b>Robustness of conditional GANs to noisy labels,</b> KIRAN THEKUMPARAMPIL, ASHISH KHETAN, and ZINAN LIN, SEWOONG OH .....	10271
<b>Robust Learning of Fixed-Structure Bayesian Networks,</b> YU CHENG, ILIAS DIAKONIKOLAS, and DANIEL KANE, ALISTAIR STEWART .....	10283
<b>Improving Simple Models with Confidence Profiles,</b> AMIT DHURANDHAR, KARTHIKEYAN SHANMUGAM, and RONNY LUSS, PEDER OLSEN .....	10296

<b>PCA of high dimensional random walks with comparison to neural network training,</b> JOSEPH ANTOGNINI, JASCHA SOHL-DICKSTEIN . . . . .	10307
<b>Learning to Solve SMT Formulas,</b> MISLAV BALUNOVIC, PAVOL BIELIK, and MARTIN VECHEV . . . . .	10317
<b>Lifted Weighted Mini-Bucket,</b> NICHOLAS GALLO, ALEXANDER IHLER . . . . .	10329
<b>Learning and Inference in Hilbert Space with Quantum Graphical Models,</b> SIDDARTH SRINIVASAN, CARLTON DOWNEY, and BYRON BOOTS . . . . .	10338
<b>Unsupervised Image-to-Image Translation Using Domain-Specific Variational Information Bound,</b> HADI KAZEMI, SOBHAN SOLEYMANI, and FARIBORZ TAHERKHANI, SEYED IRANMANESH, NASSER NASRABADI . . . . .	10348
<b>Adversarial Risk and Robustness: General Definitions and Implications for the Uniform Distribution,</b> DIMITRIOS DIOCHNOS, SAEED MAHLOUJIFAR, and MOHAMMAD MAHMOODY . . . . .	10359
<b>Gaussian Process Prior Variational Autoencoders,</b> FRANCESCO CASALE, ADRIAN DALCA, and LUCA SAGLIETTI, JENNIFER LISTGARTEN, NICOLO FUSI . . . . .	10369
<b>3D Steerable CNNs: Learning Rotationally Equivariant Features in Volumetric Data,</b> MAURICE WEILER, WOUTER BOOMSMA, and MARIO GEIGER, MAX WELLING, TACO COHEN . . . . .	10381
<b>Context-aware Synthesis and Placement of Object Instances,</b> DONGHOON LEE, SIFEI LIU, and JINWEI GU, MING-YU LIU, MING-HSUAN YANG, JAN KAUTZ . . . . .	10393
<b>Convex Elicitation of Continuous Properties,</b> JESSICA FINOCCHIARO, RAFAEL FRONGILLO . . . . .	10404
<b>Mesh-TensorFlow: Deep Learning for Supercomputers,</b> NOAM SHAZEER, YOULONG CHENG, and NIKI PARMAR, DUSTIN TRAN, ASHISH VASWANI, PENPORN KOANANTAKOOL, PETER HAWKINS, HYOUKJOONG LEE, MINGSHENG HONG, CLIFF YOUNG, RYAN SEPASSI, BLAKE HECHTMAN . . . . .	10414
<b>Learning Abstract Options,</b> MATTHEW RIEMER, MIAO LIU, and GERALD TESAURO . . . . .	10424
<b>Bounded-Loss Private Prediction Markets,</b> RAFAEL FRONGILLO, BO WAGGONER . . . . .	10435
<b>Temporal alignment and latent Gaussian process factor inference in population spike trains,</b> LEA DUNCKER, MANEESH SAHANI . . . . .	10445
<b>Using Trusted Data to Train Deep Networks on Labels Corrupted by Severe Noise,</b> DAN HENDRYCKS, MANTAS MAZEIKA, and DUNCAN WILSON, KEVIN GIMPEL . . . . .	10456
<b>Discretely Relaxing Continuous Variables for tractable Variational Inference,</b> TREFOR EVANS, PRASANTH NAIR . . . . .	10466
<b>Regret bounds for meta Bayesian optimization with an unknown Gaussian process prior,</b> ZI WANG, BEOMJOON KIM, and LESLIE KAEHLING . . . . .	10477
<b>Diversity-Driven Exploration Strategy for Deep Reinforcement Learning,</b> ZHANG-WEI HONG, TZU-YUN SHANN, and SHIH-YANG SU, YI-HSIANG CHANG, TSU-JUI FU, CHUN-YI LEE . . . . .	10489

<b>Deep Generative Models with Learnable Knowledge Constraints,</b> ZHITING HU, ZICHAO YANG, and RUSLAN SALAKHUTDINOV, LIANHUI QIN, XIAODAN LIANG, HAOYE DONG, ERIC XING .....	10501
<b>The Sparse Manifold Transform,</b> YUBEI CHEN, DYLAN PAITON, and BRUNO OLSHAUSEN .....	10513
<b>Bayesian Structure Learning by Recursive Bootstrap,</b> RAANAN ROHEKAR, YANIV GURWICZ, and SHAMI NISIMOV, GUY KOREN, GAL NOVIK .....	10525
<b>Complex Gated Recurrent Neural Networks,</b> MORITZ WOLTER, ANGELA YAO .....	10536
<b>Learning a Warping Distance from Unlabeled Time Series Using Sequence Autoencoders,</b> ABUBAKAR ABID, JAMES ZOU .....	10547
<b>Streamlining Variational Inference for Constraint Satisfaction Problems,</b> ADITYA GROVER, TUDOR ACHIM, and STEFANO ERMON .....	10556
<b>Fast deep reinforcement learning using online adjustments from the past,</b> STEVEN HANSEN, ALEXANDER PRITZEL, and PABLO SPRECHMANN, ANDRE BARRETO, CHARLES BLUNDELL .....	10567
<b>Improved Network Robustness with Adversary Critic,</b> ALEXANDER MATYASKO, LAP-PUI CHAU .....	10578
<b>Regret Bounds for Online Portfolio Selection with a Cardinality Constraint,</b> SHINJI ITO, DAISUKE HATANO, and HANNA SUMITA, AKIHIRO YABE, TAKURO FUKUNAGA, NAONORI KAKIMURA, KEN-ICHI KAWARABAYASHI .....	10588
<b>Sketching Method for Large Scale Combinatorial Inference,</b> WILL SUN, JUNWEI LU, and HAN LIU .....	10598
<b>Connecting Optimization and Regularization Paths,</b> ARUN SUGGALA, ADARSH PRASAD, and PRADEEP RAVIKUMAR .....	10608
<b>Fully Neural Network Based Speech Recognition on Mobile and Embedded Devices,</b> JINHWAN PARK, YOONHO BOO, and IKSOO CHOI, SUNGHO SHIN, WONYONG SUNG .....	10620
<b>Understanding Regularized Spectral Clustering via Graph Conductance,</b> YILIN ZHANG, KARL ROHE .....	10631
<b>Data-Driven Clustering via Parameterized Lloyd’s Families,</b> MARIA-FLORINA BALCAN, TRAVIS DICK, and COLIN WHITE .....	10641
<b>Learning Beam Search Policies via Imitation Learning,</b> RENATO NEGRINHO, MATTHEW GORMLEY, and GEOFFREY GORDON .....	10652
<b>Benefits of over-parameterization with EM,</b> JI XU, DANIEL HSU, and ARIAN MALEKI .....	10662
<b>Thermostat-assisted continuously-tempered Hamiltonian Monte Carlo for Bayesian learning,</b> RUI LUO, JIANHONG WANG, and YAODONG YANG, JUN WANG, ZHANXING ZHU .....	10673
<b>Robust Subspace Approximation in a Stream,</b> ROIE LEVIN, ANISH SEVEKARI, and DAVID WOODRUFF .....	10683
<b>Mean Field for the Stochastic Blockmodel: Optimization Landscape and Convergence Issues,</b> SOUMENDU MUKHERJEE, PURNAMRITA SARKAR, and Y. WANG, BOWEI YAN .....	10694



<b>Analysis of Krylov Subspace Solutions of Regularized Non-Convex Quadratic Problems,</b> YAIR CARMON, JOHN DUCHI .....	10705
<b>Autoconj: Recognizing and Exploiting Conjugacy Without a Domain-Specific Language,</b> MATTHEW HOFFMAN, MATTHEW JOHNSON, and DUSTIN TRAN .....	10716
<b>DropBlock: A regularization method for convolutional networks,</b> GOLNAZ GHIASI, TSUNG-YI LIN, and QUOC LE .....	10727
<b>Forward Modeling for Partial Observation Strategy Games - A StarCraft Defogger,</b> GABRIEL SYNNAEVE, ZEMING LIN, and JONAS GEHRING, DAN GANT, VEGARD MELLA, VASIL KHALIDOV, NICOLAS CARION, NICOLAS USUNIER .....	10738
<b>With Friends Like These, Who Needs Adversaries?,</b> SAUMYA JETLEY, NICHOLAS LORD, and PHILIP TORR .....	10749
<b>Decentralize and Randomize: Faster Algorithm for Wasserstein Barycenters,</b> PAVEL DVURECHENSKY, DARINA DVINSKIKH, and ALEXANDER GASNIKOV, CESAR URIBE, ANGELIA NEDICH .....	10760
<b>Joint Autoregressive and Hierarchical Priors for Learned Image Compression,</b> DAVID MINNEN, JOHANNES BALLÉ, and GEORGE TODERICI .....	10771
<b>Learning Temporal Point Processes via Reinforcement Learning,</b> SHUANG LI, SHUAI XIAO, and SHIXIANG ZHU, NAN DU, YAO XIE, LE SONG ..	10781
<b>Bias and Generalization in Deep Generative Models: An Empirical Study,</b> SHENGJIA ZHAO, HONGYU REN, and ARIANNA YUAN, JIAMING SONG, NOAH GOODMAN, STEFANO ERMON .....	10792
<b>Fast and Effective Robustness Certification,</b> GAGANDEEP SINGH, TIMON GEHR, and MATTHEW MIRMAN, MARKUS PÜSCHEL, MARTIN VECHEV ..	10802
<b>Support Recovery for Orthogonal Matching Pursuit: Upper and Lower bounds,</b> RAGHAV SOMANI, CHIRAG GUPTA, and PRATEEK JAIN, PRANEETH NETRAPALLI .....	10814
<b>Differentially Private Change-Point Detection,</b> SARA KREHIEL, RACHEL CUMMINGS, and WANRONG ZHANG, YAJUN MEI, RUI TUO .....	10825
<b>Multi-value Rule Sets for Interpretable Classification with Feature-Efficient Representations,</b> TONG WANG .....	10835
<b>Domain Adaptation by Using Causal Inference to Predict Invariant Conditional Distributions,</b> SARA MAGLIACANE, THIJS OMMEN, and TOM CLAASSEN, STEPHAN BONGERS, PHILIP VERSTEEG, JORIS MOOIJ .....	10846
<b>Smoothed Analysis of Discrete Tensor Decomposition and Assemblies of Neurons,</b> NIMA ANARI, CONSTANTINOS DASKALAKIS, and WOLFGANG MAASS, CHRISTOS PAPADIMITRIOU, AMIN SABERI, SANTOSH VEMPALA .....	10857
<b>MixLasso: Generalized Mixed Regression via Convex Atomic-Norm Regularization,</b> IAN YEN, WEI-CHENG LEE, and KAI ZHONG, SUNG-EN CHANG, PRADEEP RAVIKUMAR, SHOU-DE LIN .....	10868
<b>Semidefinite relaxations for certifying robustness to adversarial examples,</b> ADITI RAGHUNATHAN, JACOB STEINHARDT, and PERCY LIANG ..	10877

<b>Removing Hidden Confounding by Experimental Grounding,</b> NATHAN KALLUS, AAHLAD PULI, and URI SHALIT .....	10888
<b>Topkapi: Parallel and Fast Sketches for Finding Top-K Frequent Elements,</b> ANKUSH MANDAL, HE JIANG, and ANSHUMALI SHRIVASTAVA, VIVEK SARKAR .....	10898
<b>Contrastive Learning from Pairwise Measurements,</b> YI CHEN, ZHUORAN YANG, and YUCHEN XIE, ZHAORAN WANG .....	10909
<b>Point process latent variable models of larval zebrafish behavior,</b> ANUJ SHARMA, SCOTT LINDERMAN, and ROBERT JOHNSON, FLORIAN ENGERT .....	10919
<b>Computationally and statistically efficient learning of causal Bayes nets using path queries,</b> KEVIN BELLO, JEAN HONORIO .....	10931
<b>Sparse PCA from Sparse Linear Regression,</b> GUY BRESLER, SUNG PARK, and MADALINA PERSU .....	10942
<b>Multiple Instance Learning for Efficient Sequential Data Classification on Resource-constrained Devices,</b> DON DENNIS, CHIRAG PABBARAJU, and HARSHA SIMHADRI, PRATEEK JAIN .....	10953
<b>Transfer of Deep Reactive Policies for MDP Planning,</b> ANIKET BAJPAI, SANKALP GARG, and MAUSAM MAUSAM .....	10965
<b>The Price of Fair PCA: One Extra dimension,</b> SAMIRA SAMADI, UTHAIPON TANTIPONGPIPAT, and JAMIE MORGENSTERN, MOHIT SINGH, SANTOSH VEMPALA .....	10976
<b>GroupReduce: Block-Wise Low-Rank Approximation for Neural Language Model Shrinking,</b> PATRICK CHEN, SI SI, and YANG LI, CIPRIAN CHELBA, CHO-JUI HSIEH .....	10988
<b>Author Index</b>	