

37th International Conference on Machine Learning (ICML 2020)

Proceedings of Machine Learning Research Volume 119

Online
13 – 18 July 2020

Part 1 of 15

Editors:

Hal Daumé III
Aarti Singh

ISBN: 978-1-7138-2112-0

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© (2020) by the Authors
All rights reserved.

Printed with permission by Curran Associates, Inc. (2021)

For permission requests, please contact International Machine Learning Society (IMLS)
at the address below.

International Machine Learning Society (IMLS)
c/o Priscilla Rasmussen
209 North Eighth Street
Stroudsburg PA 18360
USA

Phone: (570) 476-8006
Fax: (570) 476-0860

rasmussen@ptd.net

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
Web: www.proceedings.com

TABLE OF CONTENTS

PART 1

SELECTIVE DYNA-STYLE PLANNING UNDER LIMITED MODEL CAPACITY	1
<i>Z. Abbas, S. Sokota, E. Talvittie, M. White</i>	
EFFICIENT OPTIMISTIC EXPLORATION IN LINEAR-QUADRATIC REGULATORS VIA LAGRANGIAN RELAXATION	11
<i>Marc Abeille, Alessandro Lazaric</i>	
SUPER-EFFICIENCY OF AUTOMATIC DIFFERENTIATION FOR FUNCTIONS DEFINED AS A MINIMUM	20
<i>Pierre Ablin, Gabriel Peyré, Thomas Moreau</i>	
A GEOMETRIC APPROACH TO ARCHETYPAL ANALYSIS VIA SPARSE PROJECTIONS	30
<i>V. Abrol, P. Sharma</i>	
CONTEXT-AWARE LOCAL DIFFERENTIAL PRIVACY	40
<i>Jayadev Acharya, K. A. Bonawitz, Peter Kairouz, Daniel Ramage, Ziteng Sun</i>	
EFFICIENT INTERVENTION DESIGN FOR CAUSAL DISCOVERY WITH LATENTS	51
<i>Raghavendra Addanki, Shiva Prasad Kasiviswanathan, Andrew McGregor, Cameron Musco</i>	
THE NEURAL TANGENT KERNEL IN HIGH DIMENSIONS: TRIPLE DESCENT AND A MULTI-SCALE THEORY OF GENERALIZATION	62
<i>B. Adlam, J. Pennington</i>	
RANK AGGREGATION FROM PAIRWISE COMPARISONS IN THE PRESENCE OF ADVERSARIAL CORRUPTIONS	73
<i>Arpit Agarwal, Shivani Agarwal, Sanjeev Khanna, Prathamesh Patil</i>	
BOOSTING FOR CONTROL OF DYNAMICAL SYSTEMS	84
<i>Naman Agarwal, Nataly Brukhim, Elad Hazan, Zhou Lu</i>	
AN OPTIMISTIC PERSPECTIVE ON OFFLINE REINFORCEMENT LEARNING	92
<i>R. Agarwal, D. Schuurmans, M. Norouzi</i>	
OPTIMAL BOUNDS BETWEEN F-DIVERGENCES AND INTEGRAL PROBABILITY METRICS	103
<i>Rohit Agrawal, Thibaut Horel</i>	
LAZYITER: A FAST ALGORITHM FOR COUNTING MARKOV EQUIVALENT DAGS AND DESIGNING EXPERIMENTS	113
<i>A. Teshmizi, S. Salenkalebar, N. Kyavash</i>	
LEARNING WHAT TO DEFER FOR MAXIMUM INDEPENDENT SETS	122
<i>S. Ahn, Y. Seo, J. Shin</i>	
WHY BIGGER IS NOT ALWAYS BETTER: ON FINITE AND INFINITE NEURAL NETWORKS	133
<i>Laurence Aitchison</i>	
DISCRIMINATIVE JACKKNIFE: QUANTIFYING UNCERTAINTY IN DEEP LEARNING VIA HIGHER-ORDER INFLUENCE FUNCTIONS	142
<i>Ahmed M. Alaa, Mihaela van der Schaar</i>	
FREQUENTIST UNCERTAINTY IN RECURRENT NEURAL NETWORKS VIA BLOCKWISE INFLUENCE FUNCTIONS	152
<i>Ahmed M. Alaa, Mihaela van der Schaar</i>	
RANDOM EXTRAPOLATION FOR PRIMAL-DUAL COORDINATE DESCENT	168
<i>Ahmet Alacaoglu, Olivier Fercoq, Volkan Cevher</i>	
A NEW REGRET ANALYSIS FOR ADAM-TYPE ALGORITHMS	179
<i>Ahmet Alacaoglu, Yura Malitsky, Panayotis Mertikopoulos, Volkan Cevher</i>	
RESTARTED BAYESIAN ONLINE CHANGE-POINT DETECTOR ACHIEVES OPTIMAL DETECTION DELAY	188
<i>R. Alami, O. Maillard, R. Feraud</i>	
MAXIMUM LIKELIHOOD WITH BIAS-CORRECTED CALIBRATION IS HARD-TO-BEAT AT LABEL SHIFT ADAPTATION	199
<i>A. Alexandari, A. Kundaje, A. Shrikumar</i>	
THE IMPLICIT REGULARIZATION OF STOCHASTIC GRADIENT FLOW FOR LEAST SQUARES	210
<i>A. Ali, E. Dobriban, R. Tibshirani</i>	
STRUCTURAL LANGUAGE MODELS OF CODE	222
<i>Uri Alon, Roy Sadaka, Omer Levy, Eran Yahav</i>	
LOWFER: LOW-RANK BILINEAR POOLING FOR LINK PREDICTION	234
<i>Saadullah Amin, Stalin Varanasi, Katherine Ann Dunfield, Günter Neumann</i>	
DISCOUNT FACTOR AS A REGULARIZER IN REINFORCEMENT LEARNING	246
<i>R. Amit, R. Meir, K. Ciosek</i>	
NEURO-SYMBOLIC VISUAL REASONING: DISENTANGLING "VISUAL" FROM "REASONING"	256
<i>Saeed Amizadeh, Hamid Palangi, Oleksandr Polozov, Yichen Huang, Kazuhito Koishida</i>	
THE DIFFERENTIABLE CROSS-ENTROPY METHOD	268
<i>Brandon Amos, Denis Yarats</i>	
CUSTOMIZING ML PREDICTIONS FOR ONLINE ALGORITHMS	280
<i>Keerti Anand, Rong Ge, Debmalya Panigrahi</i>	

FAIRWASHING EXPLANATIONS WITH OFF-MANIFOLD DETERGENT	291
<i>Christopher J. Anders, Plamen Pasliev, Ann-Kathrin Dombrowski, Klaus-Robert Müller, Pan Kessel</i>	
POPULATION-BASED BLACK-BOX OPTIMIZATION FOR BIOLOGICAL SEQUENCE DESIGN	301
<i>Christof Angermueller, David Belanger, Andreea Gane, Zelda Mariet, David Dohan, Kevin Murphy, Lucy Colwell, D Sculley</i>	
LOW-LOSS CONNECTION OF WEIGHT VECTORS: DISTRIBUTION-BASED APPROACHES	312
<i>I. Anokhin, D. Yarotsky</i>	
ONLINE METRIC ALGORITHMS WITH UNTRUSTED PREDICTIONS	322
<i>Antonios Antoniadis, Christian Coester, Marek Eliáš, Adam Polak, Bertrand Simon</i>	
NADS: NEURAL ARCHITECTURE DISTRIBUTION SEARCH FOR UNCERTAINTY AWARENESS	333
<i>Randy Ardywibowo, Shahin Boluki, Xinyu Gong, Zhangyang Wang, Xiaoning Qian</i>	
PROVABLE REPRESENTATION LEARNING FOR IMITATION LEARNING VIA BI-LEVEL OPTIMIZATION	344
<i>Sanjeev Arora, Simon S. Du, Sham Kakade, Yuping Luo, Nikunj Saunshi</i>	
QUANTUM BOOSTING	354
<i>Srinivasan Arunachalam, Reevu Maity</i>	
BLACK-BOX CERTIFICATION AND LEARNING UNDER ADVERSARIAL PERTURBATIONS	365
<i>Hassan Ashtiani, Vinayak Pathak, Ruth Urner</i>	
INVERTIBLE GENERATIVE MODELS FOR INVERSE PROBLEMS: MITIGATING REPRESENTATION ERROR AND DATASET BIAS	376
<i>Muhammad Asim, Max Daniels, Oscar Leong, Ali Ahmed, Paul Hand</i>	
ON THE CONVERGENCE OF NESTEROV'S ACCELERATED GRADIENT METHOD IN STOCHASTIC SETTINGS	387
<i>M. Assran, M. Rabbat</i>	
SAFE SCREENING RULES FOR 0-REGRESSION FROM PERSPECTIVE RELAXATIONS	398
<i>Alper Atamtürk, Andrés Gómez</i>	
ADVERSARIAL LEARNING GUARANTEES FOR LINEAR HYPOTHESES AND NEURAL NETWORKS	408
<i>Pranjal Awasthi, Natalie S. Frank, Mehryar Mohri</i>	
SAMPLE AMPLIFICATION: INCREASING DATASET SIZE EVEN WHEN LEARNING IS IMPOSSIBLE	419
<i>B. Axelrod, S. Garg, V. Sharan, G. Valiant</i>	
SPARSE CONVEX OPTIMIZATION VIA ADAPTIVELY REGULARIZED HARD THRESHOLDING	429
<i>Kyriakos Axiotis, Maxim Sviridenko</i>	
MODEL-BASED REINFORCEMENT LEARNING WITH VALUE-TARGETED REGRESSION	440
<i>A. Ayoub, Z. Jia, C. Szepesvari, M. Wang, L. Yang</i>	
FORECASTING SEQUENTIAL DATA USING CONSISTENT KOOPMAN AUTOENCODERS	452
<i>Omri Azencot, N. Benjamin Erichson, Vanessa Lin, Michael W. Mahoney</i>	
CONSTANT CURVATURE GRAPH CONVOLUTIONAL NETWORKS	463
<i>G. Bachmann, G. Becigneul, O. Ganea</i>	
SCALABLE NEAREST NEIGHBOR SEARCH FOR OPTIMAL TRANSPORT	474
<i>Arturs Backurs, Yihe Dong, Piotr Indyk, Ilya Razenshteyn, Tal Wagner</i>	
AGENT57: OUTPERFORMING THE ATARI HUMAN BENCHMARK	484
<i>A. Badia, B. Piot, S. Kapturovski, P. Sprechmann, A. Vitvitskiy, D. Guo, C. Blundell</i>	
FIDUCIARY BANDITS	495
<i>Gal Bahar, Omer Ben-Porat, Kevin Leyton-Brown, Moshe Tennenholtz</i>	
LEARNING DE-BIASED REPRESENTATIONS WITH BIASED REPRESENTATIONS	505
<i>Hyojin Bahng, Sanghyuk Chun, Sangdoon Yun, Jaegul Choo, Seong Joon Oh</i>	
DEEP K-NN FOR NOISY LABELS	517
<i>D. Bahri, H. Jiang, M. Gupta</i>	
PROVABLE SELF-PLAY ALGORITHMS FOR COMPETITIVE REINFORCEMENT LEARNING	528
<i>Yu Bai, Chi Jin</i>	
SPARSE SUBSPACE CLUSTERING WITH ENTROPY-NORM	538
<i>L. Bai, J. Liang</i>	
CORESETS FOR CLUSTERING IN GRAPHS OF BOUNDED TREEWIDTH	546
<i>Daniel Baker, Vladimir Braverman, Lingxiao Huang, Shaofeng H.-C. Jiang, Robert Krauthgamer, Xuan Wu</i>	
REFINED BOUNDS FOR ALGORITHM CONFIGURATION: THE KNIFE-EDGE OF DUAL CLASS APPROXIMABILITY	557
<i>Maria-Florina Balcan, Tuomas Sandholm, Ellen Vitercik</i>	
READY POLICY ONE: WORLD BUILDING THROUGH ACTIVE LEARNING	568
<i>Philip J. Ball, Jack Parker-Holder, Aldo Pacchiano, Krzysztof Choromanski, Stephen Roberts</i>	
STOCHASTIC OPTIMIZATION FOR REGULARIZED WASSERSTEIN ESTIMATORS	579
<i>Marin Ballu, Quentin Berthet, Francis Bach</i>	
DUAL MIRROR DESCENT FOR ONLINE ALLOCATION PROBLEMS	590
<i>Santiago Balseiro, Haihao Lu, Vahab Mirrokni</i>	
INDUCTIVE-BIAS-DRIVEN REINFORCEMENT LEARNING FOR EFFICIENT SCHEDULES IN HETEROGENEOUS CLUSTERS	606
<i>Subho S. Banerjee, Saurabh Jha, Zbigniew T. Kalbarczyk, Ravishankar K. Iyer</i>	
UNILMV2: PSEUDO-MASKED LANGUAGE MODELS FOR UNIFIED LANGUAGE MODEL PRE-TRAINING	619
<i>Hangbo Bao, Li Dong, Furu Wei, Wenhui Wang, Nan Yang, Xiaodong Liu, Yu Wang, Songhao Piao, Jianfeng Gao, Ming Zhou, Hsiao-Wuen Hon</i>	

FAST OSCAR AND OWL REGRESSION VIA SAFE SCREENING RULES	630
<i>Runxue Bao, Bin Gu, Heng Huang</i>	
OPTION DISCOVERY IN THE ABSENCE OF REWARDS WITH MANIFOLD ANALYSIS	641
<i>Amitay Bar, Ronen Talmon, Ron Meir</i>	
LEARNING THE PIECE-WISE CONSTANT GRAPH STRUCTURE OF A VARYING ISING MODEL	652
<i>Batiste Le Bars, Pierre Humbert, Argyris Kalogeratos, Nicolas Vayatis</i>	
FREQUENCY BIAS IN NEURAL NETWORKS FOR INPUT OF NON-UNIFORM DENSITY	662
<i>Ronen Basri, Meirav Galun, Amnon Geifman, David Jacobs, Yoni Kasten, Shira Kritchman</i>	
PRIVATE QUERY RELEASE ASSISTED BY PUBLIC DATA	672
<i>Raef Bassily, Albert Cheu, Shay Moran, Aleksandar Nikolov, Jonathan Ullman, Zhiwei Steven Wu</i>	
ECLIPSE: AN EXTREME-SCALE LINEAR PROGRAM SOLVER FOR WEB-APPLICATIONS	681
<i>Kinjal Basu, Amol Ghoting, Rahul Mazumder, Yao Pan</i>	
ON SECOND-ORDER GROUP INFLUENCE FUNCTIONS FOR BLACK-BOX PREDICTIONS	692
<i>Samyadeep Basu, Xuchen You, Soheil Feizi</i>	
KERNEL INTERPOLATION WITH CONTINUOUS SAMPLING	702
<i>Ayoub Belhadji, Rémi Bardenet, Pierre Chainais</i>	
DECOUPLED GREEDY LEARNING OF CNNs	713
<i>Eugene Belilovsky, Michael Eickenberg, Edouard Oyallon</i>	
THE COST-FREE NATURE OF OPTIMALLY TUNING TIKHONOV REGULARIZERS AND OTHER ORDERED SMOOTHERS	723
<i>Pierre C. Bellec, Dana Yang</i>	
DEFENSE THROUGH DIVERSE DIRECTIONS	733
<i>Christopher M. Bender, Yang Li, Yifeng Shi, Michael K. Reiter, Junier B. Oliva</i>	
INTERFERENCE AND GENERALIZATION IN TEMPORAL DIFFERENCE LEARNING	744
<i>Emmanuel Bengio, Joelle Pineau, Doina Precup</i>	
PRESELECTION BANDITS	755
<i>Viktor Bengs, Eyke Hüllermeier</i>	
EFFICIENT POLICY LEARNING FROM SURROGATE-LOSS CLASSIFICATION REDUCTIONS	765
<i>Andrew Bennett, Nathan Kallus</i>	
TRAINING NEURAL NETWORKS FOR AND BY INTERPOLATION	776
<i>Leonard Berrada, Andrew Zisserman, M. Pawan Kumar</i>	

PART 2

IMPLICIT DIFFERENTIATION OF LASSO-TYPE MODELS FOR HYPERPARAMETER OPTIMIZATION	787
<i>Q. Bertrand, Q. Klopfenstein, M. Blondel, S. Vaïter, A. Gramfort, J. Salmon</i>	
ONLINE LEARNING WITH IMPERFECT HINTS	799
<i>A. Bhaskara, A. Cutkosky, R. Kumar, M. Purohit</i>	
WHEN ARE NON-PARAMETRIC METHODS ROBUST?	809
<i>R. Bhattacharjee, K. Chaudhuri</i>	
LEARNING AND SAMPLING OF ATOMIC INTERVENTIONS FROM OBSERVATIONS	819
<i>A. Bhattacharyya, S. Gayen, S. Kandasamy, A. Maran, N. Vinodchandran</i>	
NEAR-OPTIMAL SAMPLE COMPLEXITY BOUNDS FOR LEARNING LATENT K-POLYTOPES AND APPLICATIONS TO AD-MIXTURES	831
<i>Chiranjib Bhattacharyya, Ravindran Kannan</i>	
LOW-RANK BOTTLENECK IN MULTI-HEAD ATTENTION MODELS	841
<i>Srinadh Bhojanapalli, Chulhee Yun, Ankit Singh Rawat, Sashank Reddi, Sanjiv Kumar</i>	
SPECTRAL CLUSTERING WITH GRAPH NEURAL NETWORKS FOR GRAPH POOLING	851
<i>Filippo Maria Bianchi, Daniele Grattarola, Cesare Alippi</i>	
TIME SERIES DECONFOUNDER: ESTIMATING TREATMENT EFFECTS OVER TIME IN THE PRESENCE OF HIDDEN CONFOUNDERS	861
<i>Ioana Bica, Ahmed M. Alaa, Mihaela van der Schaar</i>	
ADVERSARIAL ROBUSTNESS FOR CODE	873
<i>Pavol Bielek, Martin Vechev</i>	
THE BOOMERANG SAMPLER	885
<i>Joris Bierkens, Sebastiano Grazi, Kengo Kamatani, Gareth Roberts</i>	
TIGHT BOUNDS ON MINIMAX REGRET UNDER LOGARITHMIC LOSS VIA SELF-CONCORDANCE	896
<i>Blair Bilodeau, Dylan J. Foster, Daniel M. Roy</i>	
MY FAIR BANDIT: DISTRIBUTED LEARNING OF MAX-MIN FAIRNESS WITH MULTI-PLAYER BANDITS	907
<i>I. Bistriz, T. Baharav, A. Leshem, N. Bambos</i>	
PROVABLE GUARANTEES FOR DECISION TREE INDUCTION: THE AGNOSTIC SETTING	918
<i>Guy Blanc, Jane Lange, Li-Yang Tan</i>	
FAST DIFFERENTIABLE SORTING AND RANKING	927
<i>M. Blondel, O. Teboul, Q. Berthet, J. Djolonga</i>	
BEYOND SIGNAL PROPAGATION: IS FEATURE DIVERSITY NECESSARY IN DEEP NEURAL NETWORK INITIALIZATION?	937
<i>Y. Blumenfeld, D. Gilboa, D. Soudry</i>	

MODULATING SURROGATES FOR BAYESIAN OPTIMIZATION	947
<i>E. Bodin, M. Kaiser, I. Kazlauskaitė, Z. Dai, N. Campbell, C. Ek</i>	
DEEP COORDINATION GRAPHS	957
<i>W. Bohmer, V. Kurin, S. Whiteson</i>	
LORENTZ GROUP EQUIVARIANT NEURAL NETWORK FOR PARTICLE PHYSICS	969
<i>Alexander Bogatskiy, Brandon Anderson, Jan T. Offermann, Marwah Roussi, David W. Miller, Risi Kondor</i>	
EFFICIENT ROBUSTNESS CERTIFICATES FOR DISCRETE DATA: SPARSITY-AWARE RANDOMIZED SMOOTHING FOR GRAPHS, IMAGES AND MORE	980
<i>Aleksandar Bojchevski, Johannes Klicpera, Stephan Günnemann</i>	
PROPER NETWORK INTERPRETABILITY HELPS ADVERSARIAL ROBUSTNESS IN CLASSIFICATION	991
<i>A. Boopathy, S. Liu, G. Zhang, C. Liu, P. Chen, S. Chang, L. Daniel</i>	
SPECTRUM DEPENDENT LEARNING CURVES IN KERNEL REGRESSION AND WIDE NEURAL NETWORKS	1001
<i>Blake Bordelon, Abdulkadir Canatar, Cengiz Pehlevan</i>	
SMALL DATA, BIG DECISIONS: MODEL SELECTION IN THE SMALL-DATA REGIME	1012
<i>Jörg Bornschein, Francesco Visin, Simon Osindero</i>	
LATENT VARIABLE MODELLING WITH HYPERBOLIC NORMALIZING FLOWS	1022
<i>Avishke Joey Bose, Ariella Smofsky, Renjie Liao, Prakash Panangaden, William L. Hamilton</i>	
TIGHTENING EXPLORATION IN UPPER CONFIDENCE REINFORCEMENT LEARNING	1033
<i>Hippolyte Bourel, Odalric-Ambrym Maillard, Mohammad Sadegh Talebi</i>	
PREFERENCE MODELING WITH CONTEXT-DEPENDENT SALIENT FEATURES	1044
<i>Amanda Bower, Laura Balzano</i>	
ADVERSARIAL FILTERS OF DATASET BIASES	1055
<i>Ronan Le Bras, Swabha Swayamdipta, Chandra Bhagavatula, Rowan Zellers, Matthew E. Peters, Ashish Sabharwal, Yejin Choi</i>	
CALIBRATION, ENTROPY RATES, AND MEMORY IN LANGUAGE MODELS	1066
<i>Mark Braverman, Xinyi Chen, Sham Kakade, Karthik Narasimhan, Cyril Zhang, Yi Zhang</i>	
SCHATTEN NORMS IN MATRIX STREAMS: HELLO SPARSITY, GOODBYE DIMENSION	1077
<i>Vladimir Braverman, Robert Krauthgamer, Aditya Krishnan, Roi Sinoff</i>	
ALL IN THE EXPONENTIAL FAMILY: BREGMAN DUALITY IN THERMODYNAMIC VARIATIONAL INFERENCE	1088
<i>R. Brekelmans, V. Masrani, F. Wood, G. Steeg, A. Galstyan</i>	
ESTIMATING THE NUMBER AND EFFECT SIZES OF NON-NULL HYPOTHESES	1100
<i>Jennifer Brennan, Ramya Korlakai Vinayak, Kevin Jamieson</i>	
THE FAST ALGORITHM FOR SUBMODULAR MAXIMIZATION	1111
<i>Adam Breuer, Eric Balkanski, Yaron Singer</i>	
GNN-FILM: GRAPH NEURAL NETWORKS WITH FEATURE-WISE LINEAR MODULATION	1121
<i>M. Brockschmidt</i>	
TASKNORM: RETHINKING BATCH NORMALIZATION FOR META-LEARNING	1130
<i>John Bronskill, Jonathan Gordon, James Requeima, Sebastian Nowozin, Richard E. Turner</i>	
SAFE IMITATION LEARNING VIA FAST BAYESIAN REWARD INFERENCE FROM PREFERENCES	1142
<i>Daniel S. Brown, Russell Coleman, Ravi Srinivasan, Scott Niekum</i>	
A PAIRWISE FAIR AND COMMUNITY-PRESERVING APPROACH TO K-CENTER CLUSTERING	1155
<i>Brian Brubach, Darshan Chakrabarti, John P. Dickerson, Samir Khuller, Aravind Srinivasan, Leonidas Tsepenekas</i>	
SCALABLE EXACT INFERENCE IN MULTI-OUTPUT GAUSSIAN PROCESSES	1167
<i>Wessel P. Bruinsma, Eric Perim, Will Tebbutt, J. Scott Hosking, Arno Solin, Richard E. Turner</i>	
ONLINE PRICING WITH OFFLINE DATA: PHASE TRANSITION AND INVERSE SQUARE LAW	1179
<i>Jinzhi Bu, David Simchi-Levi, Yunzong Xu</i>	
EMPIRICAL STUDY OF THE BENEFITS OF OVERPARAMETERIZATION IN LEARNING LATENT VARIABLE MODELS	1188
<i>Rares-Darius Buhai, Yoni Halpern, Yoon Kim, Andrej Risteski, David Sontag</i>	
DEBAYES: A BAYESIAN METHOD FOR DEBIASING NETWORK EMBEDDINGS	1197
<i>Maarten Buyl, Tijl De Bie</i>	
STRUCTURED PREDICTION WITH PARTIAL LABELLING THROUGH THE INFIMUM LOSS	1207
<i>Vivien Cabannes, Alessandro Rudi, Francis Bach</i>	
ONLINE LEARNED CONTINUAL COMPRESSION WITH ADAPTIVE QUANTIZATION MODULES	1217
<i>L. Caccia, E. Belilovsky, M. Caccia, J. Pineau</i>	
BOOSTED HISTOGRAM TRANSFORM FOR REGRESSION	1228
<i>Yuchao Cai, Hanyuan Hang, Hanfang Yang, Zhouchen Lin</i>	
ON VALIDATION AND PLANNING OF AN OPTIMAL DECISION RULE WITH APPLICATION IN HEALTHCARE STUDIES	1239
<i>Hengrui Cai, Wenbin Lu, Rui Song</i>	
UNCERTAINTY QUANTIFICATION FOR NONCONVEX TENSOR COMPLETION: CONFIDENCE INTERVALS, HETEROSCEDASTICITY AND OPTIMALITY	1248
<i>Changxiao Cai, H. Vincent Poor, Yuxin Chen</i>	
PROVABLY EFFICIENT EXPLORATION IN POLICY OPTIMIZATION	1260
<i>Q. Cai, Z. Yang, C. Jin, Z. Wang</i>	
NEAR-LINEAR TIME GAUSSIAN PROCESS OPTIMIZATION* WITH ADAPTIVE BATCHING AND RESPARSIFICATION	1272
<i>Daniele Calandriello, Luigi Carratino, Alessandro Lazaric, Michal Valko, Lorenzo Rosasco</i>	

POISSON LEARNING: GRAPH BASED SEMI-SUPERVISED LEARNING AT VERY LOW LABEL RATES	1283
<i>Jeff Calder, Brendan Cook, Matthew Thorpe, Dejan Slepčev</i>	
EXPLORE, DISCOVER AND LEARN: UNSUPERVISED DISCOVERY OF STATE-COVERING SKILLS	1294
<i>Victor Campos, Alex Trotti, Caiming Xiong, Richard Socher, Xavier Giro-i-Nieto, Jordi Torres</i>	
LOGARITHMIC REGRET FOR LEARNING LINEAR QUADRATIC REGULATORS EFFICIENTLY	1305
<i>A. Cassel, A. Cohen, T. Koren</i>	
FULLY PARALLEL HYPERPARAMETER SEARCH: RESHAPED SPACE-FILLING	1315
<i>M.-L. Cauwet, C. Couprie, J. Dehos, P. Luc, J. Rapin, M. Riviere, F. Teytaud, O. Teytaud, N. Usunier</i>	
DATA PREPROCESSING TO MITIGATE BIAS: A MAXIMUM ENTROPY BASED APPROACH	1326
<i>L. Elisa Celis, Vijay Keswani, Nisheeth K. Vishnoi</i>	
META-LEARNING WITH STOCHASTIC LINEAR BANDITS	1337
<i>L. Cella, A. Lazaric, M. Pontil</i>	
DESCRIPTION BASED TEXT CLASSIFICATION WITH REINFORCEMENT LEARNING	1348
<i>Duo Chai, Wei Wu, Qinghong Han, Wu Fei, Jiwei Li</i>	
CONCISE EXPLANATIONS OF NEURAL NETWORKS USING ADVERSARIAL TRAINING	1360
<i>P. Chalasani, J. Chen, A. Chowdhury, S. Jha, X. Wu</i>	
UNLABELLED DATA IMPROVES BAYESIAN UNCERTAINTY CALIBRATION UNDER COVARIATE SHIFT	1369
<i>Alex J. Chan, Ahmed M. Alaa, Zhaozhi Qian, Mihaela van der Schaar</i>	
IMPUTER: SEQUENCE MODELLING VIA IMPUTATION AND DYNAMIC PROGRAMMING	1380
<i>William Chan, Chitwan Saharia, Geoffrey Hinton, Mohammad Norouzi, Navdeep Jaitly</i>	
OPTIMIZING FOR THE FUTURE IN NON-STATIONARY MDPS	1391
<i>Y. Chandak, G. Theodorou, S. Shankar, M. White, S. Mahadevan, P. Thomas</i>	
LEARNING TO SIMULATE AND DESIGN FOR STRUCTURAL ENGINEERING	1403
<i>Kai-Hung Chang, Chin-Yi Cheng</i>	
DECENTRALIZED REINFORCEMENT LEARNING: GLOBAL DECISION-MAKING VIA LOCAL ECONOMIC TRANSACTIONS	1414
<i>M. Chang, S. Kaushik, S. Weinberg, T. Griffiths, S. Levine</i>	
INVARIANT RATIONALIZATION	1425
<i>Shiyu Chang, Yang Zhang, Mo Yu, Tommi S. Jaakkola</i>	
CIRCUIT-BASED INTRINSIC METHODS TO DETECT OVERFITTING	1436
<i>Satrajit Chatterjee, Alan Mishchenko</i>	
BETTER DEPTH-WIDTH TRADE-OFFS FOR NEURAL NETWORKS THROUGH THE LENS OF DYNAMICAL SYSTEMS	1446
<i>Vaggos Chatziafratis, Sai Ganesh Nagarajan, Ioannis Panageas</i>	
EXPLAINABLE AND DISCOURSE TOPIC-AWARE NEURAL LANGUAGE UNDERSTANDING	1456
<i>Yatin Chaudhary, Hinrich Schütze, Pankaj Gupta</i>	
UNCERTAINTY-AWARE LOOKAHEAD FACTOR MODELS FOR QUANTITATIVE INVESTING	1466
<i>Lakshay Chauhan, John Alberg, Zachary C. Lipton</i>	
DEEP REASONING NETWORKS FOR UNSUPERVISED PATTERN DE-MIXING WITH CONSTRAINT REASONING	1477
<i>Di Chen, Yiwei Bai, Wenting Zhao, Sebastian Ament, John M. Gregoire, Carla P. Gomes</i>	
SELF-PU: SELF BOOSTED AND CALIBRATED POSITIVE-UNLABELED TRAINING	1487
<i>Xuxi Chen, Wuyang Chen, Tianlong Chen, Ye Yuan, Chen Gong, Kewei Chen, Zhangyang Wang</i>	
LEARNING TO STOP WHILE LEARNING TO PREDICT	1498
<i>Xinshi Chen, Hanjun Dai, Yu Li, Xin Gao, Le Song</i>	
COMBINATORIAL PURE EXPLORATION FOR DUELING BANDITS	1509
<i>Wei Chen, Yihan Du, Longbo Huang, Haoyu Zhao</i>	
GRAPH OPTIMAL TRANSPORT FOR CROSS-DOMAIN ALIGNMENT	1520
<i>Liqun Chen, Zhe Gan, Yu Cheng, Linjie Li, Lawrence Carin, Jingjing Liu</i>	
STABILIZING DIFFERENTIABLE ARCHITECTURE SEARCH VIA PERTURBATION-BASED REGULARIZATION	1532
<i>Xiangning Chen, Cho-Jui Hsieh</i>	
MAPPING NATURAL-LANGUAGE PROBLEMS TO FORMAL-LANGUAGE SOLUTIONS USING STRUCTURED NEURAL REPRESENTATIONS	1544
<i>Kezhen Chen, Qiuyuan Huang, Hamid Palangi, Paul Smolensky, Kenneth D. Forbus, Jianfeng Gao</i>	
CONVOLUTIONAL KERNEL NETWORKS FOR GRAPH-STRUCTURED DATA	1554
<i>Dexiong Chen, Laurent Jacob, Julien Mairal</i>	

PART 3

LEARNING FLAT LATENT MANIFOLDS WITH VAES	1565
<i>Nutan Chen, Alexej Klushyn, Francesco Ferroni, Justin Bayer, Patrick van der Smagt</i>	
A SIMPLE FRAMEWORK FOR CONTRASTIVE LEARNING OF VISUAL REPRESENTATIONS	1575
<i>T. Chen, S. Kornblith, M. Norouzi, G. Hinton</i>	
RETRO*: LEARNING RETROSYNTHETIC PLANNING WITH NEURAL GUIDED A* SEARCH	1586
<i>Binghong Chen, Chengtao Li, Hanjun Dai, Le Song</i>	
DIFFERENTIABLE PRODUCT QUANTIZATION FOR END-TO-END EMBEDDING COMPRESSION	1595
<i>Ting Chen, Lala Li, Yizhou Sun</i>	

ON EFFICIENT CONSTRUCTIONS OF CHECKPOINTS	1605
<i>Yu Chen, Zhenming Liu, Bin Ren, Xin Jin</i>	
ANGULAR VISUAL HARDNESS	1615
<i>Beidi Chen, Weiyang Liu, Zhiding Yu, Jan Kautz, Anshumali Shrivastava, Animesh Garg, Anima Anandkumar</i>	
ESTIMATING THE ERROR OF RANDOMIZED NEWTON METHODS: A BOOTSTRAP APPROACH	1627
<i>Jessie X.T. Chen, Miles E. Lopes</i>	
VFLOW: MORE EXPRESSIVE GENERATIVE FLOWS WITH VARIATIONAL DATA AUGMENTATION	1638
<i>J. Chen, C. Lu, B. Chenli, J. Zhu, T. Tian</i>	
MORE DATA CAN EXPAND THE GENERALIZATION GAP BETWEEN ADVERSARIALLY ROBUST AND STANDARD MODELS	1648
<i>Lin Chen, Yifei Min, Mingrui Zhang, Amin Karbasi</i>	
AN ACCELERATED DFO ALGORITHM FOR FINITE-SUM CONVEX FUNCTIONS	1659
<i>Y. Chen, A. Orvieto, A. Lucchi</i>	
GENERATIVE PRETRAINING FROM PIXELS	1669
<i>Mark Chen, Alec Radford, Rewon Child, Jeff Wu, Heewoo Jun, David Luan, Ilya Sutskever</i>	
NEGATIVE SAMPLING IN SEMI-SUPERVISED LEARNING	1682
<i>J. Chen, V. Shah, A. Kyrillidis</i>	
OPTIMIZATION FROM STRUCTURED SAMPLES FOR COVERAGE FUNCTIONS	1693
<i>Wei Chen, Xiaoming Sun, Jialin Zhang, Zhijie Zhang</i>	
SIMPLE AND DEEP GRAPH CONVOLUTIONAL NETWORKS	1703
<i>Ming Chen, Zhewei Wei, Zengfeng Huang, Bolin Ding, Yaliang Li</i>	
ON BREAKING DEEP GENERATIVE MODEL-BASED DEFENSES AND BEYOND	1714
<i>Yanzhi Chen, Renjie Xie, Zhanxing Zhu</i>	
AUTOMATED SYNTHETIC-TO-REAL GENERALIZATION	1724
<i>Wuyang Chen, Zhiding Yu, Zhangyang Wang, Anima Anandkumar</i>	
(LOCALLY) DIFFERENTIALLY PRIVATE COMBINATORIAL SEMI-BANDITS	1735
<i>X. Chen, K. Zheng, Z. Zhou, Y. Yang, W. Chen, L. Wang</i>	
HIGH-DIMENSIONAL ROBUST MEAN ESTIMATION VIA GRADIENT DESCENT	1746
<i>Y. Cheng, I. Diakonikolas, R. Ge, M. Soltanolkotabi</i>	
CLUB: A CONTRASTIVE LOG-RATIO UPPER BOUND OF MUTUAL INFORMATION	1757
<i>Pengyu Cheng, Weituo Hao, Shuyang Dai, Jiachang Liu, Zhe Gan, Lawrence Carin</i>	
LEARNING WITH BOUNDED INSTANCE- AND LABEL-DEPENDENT LABEL NOISE	1767
<i>Jiacheng Cheng, Tongliang Liu, Kotagiri Ramamohanarao, Dacheng Tao</i>	
MUTUAL TRANSFER LEARNING FOR MASSIVE DATA	1778
<i>Ching-Wei Cheng, Xingye Qiao, Guang Cheng</i>	
STOCHASTIC GRADIENT AND LANGEVIN PROCESSES	1788
<i>X. Cheng, D. Yin, P. Bartlett, M. Jordan</i>	
REPRESENTATION LEARNING VIA ADVERSARIALLY-CONTRASTIVE OPTIMAL TRANSPORT	1798
<i>Anoop Cherian, Shuchin Aeron</i>	
CONVERGENCE RATES OF VARIATIONAL INFERENCE IN SPARSE DEEP LEARNING	1809
<i>Badr-Eddine Chérif-Abdellatif</i>	
REINFORCEMENT LEARNING FOR NON-STATIONARY MARKOV DECISION PROCESSES: THE BLESSING OF (MORE) OPTIMISM	1821
<i>Wang Chi Cheung, David Simchi-Levi, Ruihao Zhu</i>	
STREAMING CORESETS FOR SYMMETRIC TENSOR FACTORIZATION	1833
<i>Rachit Chhaya, Jayesh Choudhari, Anirban Dasgupta, Supratim Shit</i>	
ON CORESETS FOR REGULARIZED REGRESSION	1844
<i>Rachit Chhaya, Anirban Dasgupta, Supratim Shit</i>	
HOW TO SOLVE FAIR K-CENTER IN MASSIVE DATA MODELS	1855
<i>A. Chiplunkar, S. Kale, S. Ramamoorthy</i>	
FAIR GENERATIVE MODELING VIA WEAK SUPERVISION	1865
<i>K. Choi, A. Grover, T. Singh, R. Shu, S. Ermon</i>	
ENCODING MUSICAL STYLE WITH TRANSFORMER AUTOENCODERS	1877
<i>K. Choi, C. Hawthorne, I. Simon, M. Dinculescu, J. Engel</i>	
K-MEANS++: FEW MORE STEPS YIELD CONSTANT APPROXIMATION	1887
<i>Davin Choo, Christoph Grunau, Julian Portmann, Václav Rozhon</i>	
STOCHASTIC FLOWS AND GEOMETRIC OPTIMIZATION ON THE ORTHOGONAL GROUP	1896
<i>K. Choromanski, D. Cheikhi, J. Davis, V. Likhoshesterov, A. Nazaret, A. Bahamou, X. Song, M. Akarte, J. Parker-Holder, J. Bergquist, Y. Gao, A. Pacchiano, T. Sarlos, A. Weller, V. Sindhvani</i>	
UNBIASED RISK ESTIMATORS CAN MISLEAD: A CASE STUDY OF LEARNING WITH COMPLEMENTARY LABELS	1907
<i>Yu-Ting Chou, Gang Niu, Hsuan-Tien Lin, Masashi Sugiyama</i>	
DATA-DEPENDENT DIFFERENTIALLY PRIVATE PARAMETER LEARNING FOR DIRECTED GRAPHICAL MODELS	1917
<i>A. Chowdhury, T. Rekatsinas, S. Jha</i>	
ONLINE CONTINUAL LEARNING FROM IMBALANCED DATA	1930
<i>Aristotelis Chrysakis, Marie-Francine Moens</i>	
DISTANCE METRIC LEARNING WITH JOINT REPRESENTATION DIVERSIFICATION	1940
<i>Xu Chu, Yang Lin, Yasha Wang, Xiting Wang, Hailong Yu, Xin Gao, Qi Tong</i>	

SEMISMOOTH NEWTON ALGORITHM FOR EFFICIENT PROJECTIONS ONTO 1, -NORM BALL	1952
<i>Dejun Chu, Changshui Zhang, Shiliang Sun, Qing Tao</i>	
ESTIMATING GENERALIZATION UNDER DISTRIBUTION SHIFTS VIA DOMAIN-INVARIANT REPRESENTATIONS	1962
<i>Ching-Yao Chuang, Antonio Torralba, Stefanie Jegelka</i>	
SCALABLE AND EFFICIENT COMPARISON-BASED SEARCH WITHOUT FEATURES	1973
<i>D. Chumalov, L. Maystre, M. Grossglauer</i>	
FEATURE-MAP-LEVEL ONLINE ADVERSARIAL KNOWLEDGE DISTILLATION	1984
<i>Inseop Chung, SeongUk Park, Jangho Kim, Nojun Kwak</i>	
TEACHING WITH LIMITED INFORMATION ON THE LEARNER'S BEHAVIOUR	1994
<i>F. Cicalese, S. Filho, E. Laber, M. Molinaro</i>	
DEEP DIVERGENCE LEARNING	2005
<i>Kubra Cilingir, Rachel Manzelli, Brian Kulis</i>	
MODEL FUSION WITH KULLBACK-LEIBLER DIVERGENCE	2016
<i>Sebastian Clatici, Mikhail Yurochkin, Soumya Ghosh, Justin Solomon</i>	
LEVERAGING PROCEDURAL GENERATION TO BENCHMARK REINFORCEMENT LEARNING	2026
<i>Karl Cobbe, Christopher Hesse, Jacob Hilton, John Schulman</i>	
COMPOSABLE SKETCHES FOR FUNCTIONS OF FREQUENCIES: BEYOND THE WORST CASE	2035
<i>Edith Cohen, Ofir Geri, Rasmus Pagh</i>	
HEALING PRODUCTS OF GAUSSIAN PROCESS EXPERTS	2046
<i>S. Cohen, R. Mubvha, T. Marwala, M. Deisenroth</i>	
ON EFFICIENT LOW DISTORTION ULTRAMETRIC EMBEDDING	2056
<i>Vincent Cohen-Addad, Karthik C. S., Guillaume Lagarde</i>	
SUB-LINEAR MEMORY SKETCHES FOR NEAR NEIGHBOR SEARCH ON STREAMING DATA	2067
<i>B. Coleman, R. Baraniuk, A. Shrivastava</i>	
WORD-LEVEL SPEECH RECOGNITION WITH A LETTER TO WORD ENCODER	2078
<i>Ronan Collobert, Awni Hannun, Gabriel Synnaeve</i>	
BOOSTING FRANK-WOLFE BY CHASING GRADIENTS	2089
<i>Cyrille W. Combettes, Sebastian Pokutta</i>	
LEARNING OPINIONS IN SOCIAL NETWORKS	2100
<i>Vincent Conitzer, Debmalya Panigrahi, Hanrui Zhang</i>	
RELAXING BIJECTIVITY CONSTRAINTS WITH CONTINUOUSLY INDEXED NORMALISING FLOWS	2111
<i>R. Cornish, A. Caterini, G. Delgiannidis, A. Doucet</i>	
ADAPTIVE REGION-BASED ACTIVE LEARNING	2122
<i>Corinna Cortes, Giulia DeSalvo, Claudio Gentile, Mehryar Mohri, Ningshan Zhang</i>	
ONLINE LEARNING WITH DEPENDENT STOCHASTIC FEEDBACK GRAPHS	2132
<i>Corinna Cortes, Giulia DeSalvo, Claudio Gentile, Mehryar Mohri, Ningshan Zhang</i>	
LEARNABLE GROUP TRANSFORM FOR TIME-SERIES	2142
<i>R. Cosentino, B. Aazhang</i>	
DINO: DISTRIBUTED NEWTON-TYPE OPTIMIZATION METHOD	2152
<i>Rixon Crane, Fred Roosta</i>	
CAUSAL MODELING FOR FAIRNESS IN DYNAMICAL SYSTEMS	2163
<i>Elliot Creager, David Madras, Toniann Pitassi, Richard Zemel</i>	
MINIMALLY DISTORTED ADVERSARIAL EXAMPLES WITH A FAST ADAPTIVE BOUNDARY ATTACK	2174
<i>Francesco Croce, Matthias Hein</i>	
RELIABLE EVALUATION OF ADVERSARIAL ROBUSTNESS WITH AN ENSEMBLE OF DIVERSE PARAMETER-FREE ATTACKS	2184
<i>Francesco Croce, Matthias Hein</i>	
REAL-TIME OPTIMISATION FOR ONLINE LEARNING IN AUCTIONS	2195
<i>L. Croissant, M. Abeille, C. Calauzenes</i>	
PRIVATELY DETECTING CHANGES IN UNKNOWN DISTRIBUTIONS	2205
<i>Rachel Cummings, Sara Krehbiel, Yuliia Lut, Wanrong Zhang</i>	
FLEXIBLE AND EFFICIENT LONG-RANGE PLANNING THROUGH CURIOUS EXPLORATION	2216
<i>A. Curtis, M. Xin, D. Arumugam, K. Feigels, D. Yamins</i>	
PARAMETER-FREE, DYNAMIC, AND STRONGLY-ADAPTIVE ONLINE LEARNING	2228
<i>Ashok Cutkosky</i>	
MOMENTUM IMPROVES NORMALIZED SGD	2238
<i>A. Cutkosky, H. Mehta</i>	
SUPERVISED QUANTILE NORMALIZATION FOR LOW-RANK MATRIX APPROXIMATION	2247
<i>Marco Cuturi, Olivier Teboul, Jonathan Niles-Weed, Jean-Philippe Vert</i>	
DOUBLE TROUBLE IN DOUBLE DESCENT: BIAS AND VARIANCE(S) IN THE LAZY REGIME	2258
<i>Stéphane d'Ascoli, Maria Refinetti, Giulio Biroli, Florent Krzakala</i>	
R2-B2: RECURSIVE REASONING-BASED BAYESIAN OPTIMIZATION FOR NO-REGRET LEARNING IN GAMES	2269
<i>Z. Dai, Y. Chen, B. Low, P. Jaillet, T. Ho</i>	
SCALABLE DEEP GENERATIVE MODELING FOR SPARSE GRAPHS	2280
<i>Hanjun Dai, Azade Nazi, Yujia Li, Bo Dai, Dale Schuurmans</i>	
THE USUAL SUSPECTS? REASSESSING BLAME FOR VAE POSTERIOR COLLAPSE	2291
<i>Bin Dai, Ziyu Wang, David Wipf</i>	

CONFIDENCE SETS AND HYPOTHESIS TESTING IN A LIKELIHOOD-FREE INFERENCE SETTING	2301
<i>Niccolò Dalmaso, Rafael Izbicki, Ann B. Lee</i>	
GOODNESS-OF-FIT TESTS FOR INHOMOGENEOUS RANDOM GRAPHS	2313
<i>Soham Dan, Bhaswar B. Bhattacharya</i>	
SHARP STATISTICAL GUARANTEES FOR ADVERSARIALLY ROBUST GAUSSIAN CLASSIFICATION	2323
<i>Chen Dan, Yuting Wei, Pradeep Ravikumar</i>	
ADVERSARIAL ATTACKS ON PROBABILISTIC AUTOREGRESSIVE FORECASTING MODELS	2334
<i>Raphaël Dang-Nhu, Gagandeep Singh, Pavol Bielik, Martin Vechev</i>	

PART 4

SUBSPACE FITTING MEETS REGRESSION: THE EFFECTS OF SUPERVISION AND ORTHONORMALITY CONSTRAINTS ON DOUBLE DESCENT OF GENERALIZATION ERRORS	2344
<i>Yehuda Dar, Paul Mayer, Lorenzo Luzzi, Richard G. Baraniuk</i>	
PROBING EMERGENT SEMANTICS IN PREDICTIVE AGENTS VIA QUESTION ANSWERING	2354
<i>Abhishek Das, Federico Carnevale, Hamza Merzic, Laura Rimell, Rosalia Schneider, Josh Abramson, Alden Hung, Arun Ahuja, Stephen Clark, Gregory Wayne, Felix Hill</i>	
LOW-VARIANCE AND ZERO-VARIANCE BASELINES FOR EXTENSIVE-FORM GAMES	2370
<i>Trevor Davis, Martin Schmid, Michael Bowling</i>	
COMBINING DIFFERENTIABLE PDE SOLVERS AND GRAPH NEURAL NETWORKS FOR FLUID FLOW PREDICTION	2380
<i>F. Belbute-Peres, T. Economou, J. Kolter</i>	
REPRESENTING UNORDERED DATA USING COMPLEX-WEIGHTED MULTISSET AUTOMATA	2390
<i>Justin DeBenedetto, David Chiang</i>	
AN END-TO-END DIFFERENTIALLY PRIVATE LATENT DIRICHLET ALLOCATION USING A SPECTRAL ALGORITHM	2399
<i>Christopher DeCarolis, Mukul Ram, Seyed Esmaili, Yu-Xiang Wang, Furong Huang</i>	
GAMIFICATION OF PURE EXPLORATION FOR LINEAR BANDITS	2410
<i>R. Degenne, P. Menard, X. Shang, M. Valko</i>	
STRUCTURE ADAPTIVE ALGORITHMS FOR STOCHASTIC BANDITS	2421
<i>R. Degenne, H. Shao, W. Koolen</i>	
RANDOMLY PROJECTED ADDITIVE GAUSSIAN PROCESSES FOR REGRESSION	2431
<i>I. Delbridge, D. Bindel, A. Wilson</i>	
INTERPRETING ROBUST OPTIMIZATION VIA ADVERSARIAL INFLUENCE FUNCTIONS	2442
<i>Zhun Deng, Cynthia Dwork, Jialiang Wang, Linjun Zhang</i>	
NON-CONVEX LEARNING VIA REPLICA EXCHANGE STOCHASTIC GRADIENT MCMC	2452
<i>Wei Deng, Qi Feng, Liyao Gao, Faming Liang, Guang Lin</i>	
TOWARDS UNDERSTANDING THE DYNAMICS OF THE FIRST-ORDER ADVERSARIES	2462
<i>Zhun Deng, Hangfeng He, Jiaoyang Huang, Weijie J. Su</i>	
ROBUST PRICING IN DYNAMIC MECHANISM DESIGN	2472
<i>Yuan Deng, Sébastien Lahaie, Vahab Mirrokni</i>	
A SWISS ARMY KNIFE FOR MINIMAX OPTIMAL TRANSPORT	2482
<i>Softien Dhoub, Ievgen Redko, Tanguy Kerdoncuff, Rémi Emonet, Marc Sebban</i>	
MARGIN-AWARE ADVERSARIAL DOMAIN ADAPTATION WITH OPTIMAL TRANSPORT	2492
<i>Softien Dhoub, Ievgen Redko, Carole Lartizien</i>	
ENHANCING SIMPLE MODELS BY EXPLOITING WHAT THEY ALREADY KNOW	2503
<i>Amit Dhurandhar, Karthikeyan Shanmugam, Ronny Luss</i>	
SPECTRAL FRANK-WOLFE ALGORITHM: STRICT COMPLEMENTARITY AND LINEAR CONVERGENCE	2513
<i>Lijun Ding, Yingjie Fei, Qiantong Xu, Chengrun Yang</i>	
GENERALIZATION GUARANTEES FOR SPARSE KERNEL APPROXIMATION WITH ENTROPIC OPTIMAL FEATURES	2523
<i>Liang Ding, Rui Tuo, Shahin Shahrampour</i>	
LAYERED SAMPLING FOR ROBUST OPTIMIZATION PROBLEMS	2534
<i>Hu Ding, Zixiu Wang</i>	
GROWING ADAPTIVE MULTI-HYPERPLANE MACHINES	2545
<i>Nemanja Djuric, Zhuang Wang, Slobodan Vucetic</i>	
INEXACT TENSOR METHODS WITH DYNAMIC ACCURACIES	2555
<i>Nikita Doikov, Yurii Nesterov</i>	
PROVABLE SMOOTHNESS GUARANTEES FOR BLACK-BOX VARIATIONAL INFERENCE	2565
<i>Justin Domke</i>	
OPTIMAL DIFFERENTIAL PRIVACY COMPOSITION FOR EXPONENTIAL MECHANISMS	2575
<i>Jinshuo Dong, David Durfee, Ryan Rogers</i>	
MULTINOMIAL LOGIT BANDIT WITH LOW SWITCHING COST	2585
<i>Kefan Dong, Yingkai Li, Qin Zhang, Yuan Zhou</i>	
TOWARDS ADAPTIVE RESIDUAL NETWORK TRAINING: A NEURAL-ODE PERSPECTIVE	2594
<i>Chengyu Dong, Liyuan Liu, Zichao Li, Jingbo Shang</i>	
ON THE EXPRESSIVITY OF NEURAL NETWORKS FOR DEEP REINFORCEMENT LEARNING	2605
<i>K. Dong, Y. Luo, T. Yu, C. Finn, T. Ma</i>	

COLLAPSED AMORTIZED VARIATIONAL INFERENCE FOR SWITCHING NONLINEAR DYNAMICAL SYSTEMS	2616
<i>Z. Dong, B. Seybold, K. Murphy, H. Bui</i>	
EXPERT LEARNING THROUGH GENERALIZED INVERSE MULTIOBJECTIVE OPTIMIZATION: MODELS, INSIGHTS, AND ALGORITHMS	2626
<i>Chaosheng Dong, Bo Zeng</i>	
THE COMPLEXITY OF FINDING STATIONARY POINTS WITH STOCHASTIC GRADIENT DESCENT	2636
<i>Y. Drori, O. Shamir</i>	
OPTIMAL NON-PARAMETRIC LEARNING IN REPEATED CONTEXTUAL AUCTIONS WITH STRATEGIC BUYER	2646
<i>Alexey Drutsa</i>	
RESERVE PRICING IN REPEATED SECOND-PRICE AUCTIONS WITH STRATEGIC BIDDERS	2658
<i>Alexey Drutsa</i>	
NGBOOST: NATURAL GRADIENT BOOSTING FOR PROBABILISTIC PREDICTION	2670
<i>Tony Duan, Anand Avati, Daisy Yi Ding, Khanh K. Thai, Sanjay Basu, Andrew Ng, Alejandro Schuler</i>	
MINIMAX-OPTIMAL OFF-POLICY EVALUATION WITH LINEAR FUNCTION APPROXIMATION	2681
<i>Y. Duan, Z. Jia, M. Wang</i>	
ONLINE BAYESIAN MOMENT MATCHING BASED SAT SOLVER HEURISTICS	2690
<i>Haonan Duan, Saeed Nejati, George Trimponias, Pascal Poupart, Vijay Ganesh</i>	
FAMILYWISE ERROR RATE CONTROL BY INTERACTIVE UNMASKING	2700
<i>Boyan Duan, Aaditya Ramdas, Larry Wasserman</i>	
COOPERATIVE MULTI-AGENT BANDITS WITH HEAVY TAILS	2710
<i>Abhimanyu Dubey, Alex Pentland</i>	
KERNEL METHODS FOR COOPERATIVE MULTI-AGENT CONTEXTUAL BANDITS	2720
<i>Abhimanyu Dubey, Alex Pentland</i>	
OPTIMIZATION THEORY FOR RELU NEURAL NETWORKS TRAINED WITH NORMALIZATION LAYERS	2731
<i>Y. Dukler, Q. Gu, G. Montufar</i>	
EQUIVARIANT NEURAL RENDERING	2741
<i>Emilien Dupont, Miguel Angel Bautista, Alex Colburn, Aditya Sankar, Joshua Susskind, Qi Shan</i>	
ON CONTRASTIVE LEARNING FOR LIKELIHOOD-FREE INFERENCE	2751
<i>C. Durkan, I. Murray, G. Papamakarios</i>	
EFFICIENT AND SCALABLE BAYESIAN NEURAL NETS WITH RANK-1 FACTORS	2762
<i>Michael W. Dusenberry, Ghassen Jerfel, Yeming Wen, Yi-An Ma, Jasper Snoek, Katherine Heller, Balaji Lakshminarayanan, Dustin Tran</i>	
SPARSE GAUSSIAN PROCESSES WITH SPHERICAL HARMONIC FEATURES	2773
<i>V. Dutoit, N. Durrande, J. Hensman</i>	
IS THERE A TRADE-OFF BETWEEN FAIRNESS AND ACCURACY? A PERSPECTIVE USING MISMATCHED HYPOTHESIS TESTING	2783
<i>Sanghamitra Dutta, Dennis Wei, Hazar Yueksel, Pin-Yu Chen, Sijia Liu, Kush R. Varshney</i>	
SELF-CONCORDANT ANALYSIS OF FRANK-WOLFE ALGORITHMS	2794
<i>Pavel Dvurechensky, Petr Ostroukhov, Kamil Safin, Shimrit Shern, Mathias Staudigl</i>	
ESTIMATING Q(S,S') WITH DEEP DETERMINISTIC DYNAMICS GRADIENTS	2805
<i>Ashley D. Edwards, Himanshu Sahni, Rosanne Liu, Jane Hung, Ankit Jain, Rui Wang, Adrien Ecoffet, Thomas Miconi, Charles Isbell, Jason Yosinski</i>	
TRAINING LINEAR NEURAL NETWORKS: NON-LOCAL CONVERGENCE AND COMPLEXITY RESULTS	2816
<i>Armin Eftekhari</i>	
STUDENT-TEACHER CURRICULUM LEARNING VIA REINFORCEMENT LEARNING: PREDICTING HOSPITAL INPATIENT ADMISSION LOCATION	2828
<i>Rasheed el-Bouri, David Eyre, Peter Watkinson, Tingting Zhu, David A. Clifton</i>	
DECISION TREES FOR DECISION-MAKING UNDER THE PREDICT-THEN-OPTIMIZE FRAMEWORK	2838
<i>A. Elmachtoub, J. Liang, R. McNellis</i>	
REVISITING SPATIAL INVARIANCE WITH LOW-RANK LOCAL CONNECTIVITY	2848
<i>Gamaleldin F. Elsayed, Prajit Ramachandran, Jonathon Shlens, Simon Kornblith</i>	
DIVIDE AND CONQUER: LEVERAGING INTERMEDIATE FEATURE REPRESENTATIONS FOR QUANTIZED TRAINING OF NEURAL NETWORKS	2860
<i>A. Elthakeb, P. Pilligundla, F. Miresghallah, A. Cloninger, H. Esmailzadeh</i>	
GENERALIZATION ERROR OF GENERALIZED LINEAR MODELS IN HIGH DIMENSIONS	2872
<i>M. Emami, M. Sahraee-Ardakan, P. Pandit, S. Rangan, A. Fletcher</i>	
PARALLEL ALGORITHM FOR NON-MONOTONE DR-SUBMODULAR MAXIMIZATION	2882
<i>A. Ene, H. Nguyen</i>	
CONTINUOUS-TIME BAYESIAN NETWORKS WITH CLOCKS	2892
<i>Nicolai Engelmann, Dominik Linzner, Heinz Koeppl</i>	
IDENTIFYING STATISTICAL BIAS IN DATASET REPLICATION	2902
<i>Logan Engstrom, Andrew Ilyas, Shibani Santurkar, Dimitris Tsipras, Jacob Steinhardt, Aleksander Madry</i>	
DISTRIBUTED ONLINE OPTIMIZATION OVER A HETEROGENEOUS NETWORK WITH ANY-BATCH MIRROR DESCENT	2913
<i>Nima Eshraghi, Ben Liang</i>	

RIGGING THE LOTTERY: MAKING ALL TICKETS WINNERS	2923
<i>Utku Evcı, Trevor Gale, Jacob Menick, Pablo Samuel Castro, Erich Elsen</i>	
FASTER GRAPH EMBEDDINGS VIA COARSENING	2934
<i>Mathew Fahrback, Gramoz Goranci, Richard Peng, Sushant Sachdeva, Chi Wang</i>	
LATENT BERNOULLI AUTOENCODER	2945
<i>Jiri Fajtl, Vasileios Argyriou, Dorothy Monekosso, Paolo Remagnino</i>	
OPTIMAL SEQUENTIAL MAXIMIZATION [1MM] ONE INTERVIEW IS ENOUGH!	2956
<i>Moein Falahatgar, Alon Orlitsky, Venkatadheeraj Pichapati</i>	
SPECTRAL GRAPH MATCHING AND REGULARIZED QUADRATIC RELAXATIONS: ALGORITHM AND THEORY	2966
<i>Z. Fan, C. Mao, Y. Wu, J. Xu</i>	
ON HYPERPARAMETER TUNING IN GENERAL CLUSTERING PROBLEMS	2977
<i>X. Fan, Y. Yue, P. Sarkar, Y. Wang</i>	
ONLINE MIRROR DESCENT AND DUAL AVERAGING: KEEPING PACE IN THE DYNAMIC CASE	2989
<i>Huang Fang, Nicholas J. A. Harvey, Victor S. Portella, Michael P. Friedlander</i>	
STOCHASTIC REGRET MINIMIZATION IN EXTENSIVE-FORM GAMES	2999
<i>Gabriele Farina, Christian Kroer, Tuomas Sandholm</i>	
DO GANS ALWAYS HAVE NASH EQUILIBRIA?	3010
<i>Farzan Farnia, Asuman Ozdaglar</i>	
GROWING ACTION SPACES	3021
<i>G. Farquhar, L. Gustafson, Z. Lin, S. Whiteson, N. Usunier, G. Synnaeve</i>	
IMPROVED OPTIMISTIC ALGORITHMS FOR LOGISTIC BANDITS	3033
<i>Louis Faury, Marc Abeille, Clément Calauzènes, Olivier Fercoq</i>	
REVISITING FUNDAMENTALS OF EXPERIENCE REPLAY	3042
<i>William Fedus, Prajit Ramachandran, Rishabh Agarwal, Yoshua Bengio, Hugo Larochelle, Mark Rowland, Will Dabney</i>	
LEARNING WITH MULTIPLE COMPLEMENTARY LABELS	3053
<i>Lei Feng, Takuo Kaneko, Bo Han, Gang Niu, Bo An, Masashi Sugiyama</i>	
GLOBAL CONCAVITY AND OPTIMIZATION IN A CLASS OF DYNAMIC DISCRETE CHOICE MODELS	3063
<i>Yiding Feng, Ekaterina Khmel'nitskaya, Denis Nekipelov</i>	
THE INTRINSIC ROBUSTNESS OF STOCHASTIC BANDITS TO STRATEGIC MANIPULATION	3073
<i>Z. Feng, D. Parkes, H. Xu</i>	
ACCOUNTABLE OFF-POLICY EVALUATION WITH KERNEL BELLMAN STATISTICS	3083
<i>Y. Feng, T. Ren, Z. Tang, Q. Liu</i>	
KERNELIZED STEIN DISCREPANCY TESTS OF GOODNESS-OF-FIT FOR TIME-TO-EVENT DATA	3093
<i>T. Fernandez, N. Rivera, W. Xu, A. Gretton</i>	
WHY ARE LEARNED INDEXES SO EFFECTIVE?	3104
<i>Paolo Ferragina, Fabrizio Lillo, Giorgio Vinciguerra</i>	
IMPLICIT LEARNING DYNAMICS IN STACKELBERG GAMES: EQUILIBRIA CHARACTERIZATION, CONVERGENCE ANALYSIS, AND EMPIRICAL STUDY	3114
<i>T. Fiez, B. Chasnov, L. Ratliff</i>	

PART 5

CAN AUTONOMOUS VEHICLES IDENTIFY, RECOVER FROM, AND ADAPT TO DISTRIBUTION SHIFTS?	3126
<i>Angelos Filos, Panagiotis Tigas, Rowan McAllister, Nicholas Rhinehart, Sergey Levine, Yarin Gal</i>	
HOW TO TRAIN YOUR NEURAL ODE: THE WORLD OF JACOBIAN AND KINETIC REGULARIZATION	3135
<i>C. Finlay, J. Jacobsen, L. Nurbekyan, A. Oberman</i>	
GENERALIZING CONVOLUTIONAL NEURAL NETWORKS FOR EQUIVARIANCE TO LIE GROUPS ON ARBITRARY CONTINUOUS DATA	3146
<i>M. Finzi, S. Stanton, P. Izmailov, A. Wilson</i>	
INFORMATION PARTICLE FILTER TREE: AN ONLINE ALGORITHM FOR POMDPS WITH BELIEF-BASED REWARDS ON CONTINUOUS DOMAINS	3158
<i>Johannes Fischer, Ömer Sahin Tas</i>	
TOPIC MODELING VIA FULL DEPENDENCE MIXTURES	3169
<i>D. Fisher, M. Kozdoba, S. Mannor</i>	
P-NORM FLOW DIFFUSION FOR LOCAL GRAPH CLUSTERING	3180
<i>Kimon Fountoulakis, Di Wang, Shenghao Yang</i>	
STOCHASTIC LATENT RESIDUAL VIDEO PREDICTION	3191
<i>Jean-Yves Franceschi, Edouard Delasalles, Mickaël Chen, Sylvain Lamprier, Patrick Gallinari</i>	
LEVERAGING FREQUENCY ANALYSIS FOR DEEP FAKE IMAGE RECOGNITION	3205
<i>Joel Frank, Thorsten Eisenhofer, Lea Schönherr, Asja Fischer, Dorothea Kolossa, Thorsten Holz</i>	
LINEAR MODE CONNECTIVITY AND THE LOTTERY TICKET HYPOTHESIS	3217
<i>J. Frankle, G. Dziugaite, D. Roy, M. Carbin</i>	
NO-REGRET AND INCENTIVE-COMPATIBLE ONLINE LEARNING	3228
<i>Rupert Freeman, David M. Pennock, Chara Podimata, Jennifer Wortman Vaughan</i>	
FAST AND THREE-RIOUS: SPEEDING UP WEAK SUPERVISION WITH TRIPLET METHODS	3238
<i>Daniel Y. Fu, Mayee F. Chen, Frederic Sala, Sarah M. Hooper, Kayvon Fatahalian, Christopher Ré</i>	
AUTOGAN-DISTILLER: SEARCHING TO COMPRESS GENERATIVE ADVERSARIAL NETWORKS	3250
<i>Yonggan Fu, Wuyang Chen, Haotao Wang, Haoran Li, Yingyan Lin, Zhangyang Wang</i>	

DON'T WASTE YOUR BITS! SQUEEZE ACTIVATIONS AND GRADIENTS FOR DEEP NEURAL NETWORKS VIA TINYSRIPT	3262
<i>Fangcheng Fu, Yuzheng Hu, Yihan He, Jiawei Jiang, Yingxia Shao, Ce Zhang, Bin Cui</i>	
DESSILBI: EXPLORING STRUCTURAL SPARSITY OF DEEP NETWORKS VIA DIFFERENTIAL INCLUSION PATHS	3273
<i>Y. fu, C. Liu, D. Li, X. Sun, J. Zeng, Y. Yao</i>	
APPROXIMATION GUARANTEES OF LOCAL SEARCH ALGORITHMS VIA LOCALIZABILITY OF SET FUNCTIONS	3285
<i>Kaito Fujii</i>	
ACCELERATING THE DIFFUSION-BASED ENSEMBLE SAMPLING BY NON-REVERSIBLE DYNAMICS	3295
<i>F. Futami, I. Sato, M. Sugiyama</i>	
STOCHASTIC BANDITS WITH ARM-DEPENDENT DELAYS	3306
<i>A. Manegueu, C. Vernade, A. Carpentier, M. Valko</i>	
ABSTRACTION MECHANISMS PREDICT GENERALIZATION IN DEEP NEURAL NETWORKS	3315
<i>Alex Gain, Hava Siegelmann</i>	
A FREE-ENERGY PRINCIPLE FOR REPRESENTATION LEARNING	3325
<i>Yansong Gao, Pratik Chaudhari</i>	
CAN STOCHASTIC ZERO-TH-ORDER FRANK-WOLFE METHOD CONVERGE FASTER FOR NON-CONVEX PROBLEMS?	3335
<i>Hongchang Gao, Heng Huang</i>	
ONLINE CONVEX OPTIMIZATION IN THE RANDOM ORDER MODEL	3345
<i>Dan Garber, Gal Korcia, Kfir Y. Levy</i>	
SYMBOLIC NETWORK: GENERALIZED NEURAL POLICIES FOR RELATIONAL MDPS	3355
<i>Sankalp Garg, Aniket Bajpai, Mausam</i>	
PREDICTING DELIBERATIVE OUTCOMES	3366
<i>V. Garg, T. Jaakkola</i>	
GENERALIZATION AND REPRESENTATIONAL LIMITS OF GRAPH NEURAL NETWORKS	3377
<i>V. Garg, S. Jegelka, T. Jaakkola</i>	
DEEP PQR: SOLVING INVERSE REINFORCEMENT LEARNING USING ANCHOR ACTIONS	3389
<i>Sinong Geng, Houssam Nassif, Carlos A. Manzanares, A. Max Reppen, Ronnie Sircar</i>	
MULTILINEAR LATENT CONDITIONING FOR GENERATING UNSEEN ATTRIBUTE COMBINATIONS	3400
<i>Markos Georgopoulos, Grigorios G Chrysos, Maja Pantic, Yannis Panagakis</i>	
GENERALISATION ERROR IN LEARNING WITH RANDOM FEATURES AND THE HIDDEN MANIFOLD MODEL	3410
<i>Federica Gerace, Bruno Loureiro, Florent Krzakala, Marc Mézard, Lenka Zdeborová</i>	
BLACK-BOX METHODS FOR RESTORING MONOTONICITY	3421
<i>E. Gergatsouli, B. Lucier, C. Tzamos</i>	
ONLINE MULTI-KERNEL LEARNING WITH GRAPH-STRUCTURED FEEDBACK	3432
<i>Pouya M Ghari, Yanning Shen</i>	
TASK-ORIENTED ACTIVE PERCEPTION AND PLANNING IN ENVIRONMENTS WITH PARTIALLY KNOWN SEMANTICS	3442
<i>M. Ghasemi, E. Bulgur, U. Topcu</i>	
CHARACTERIZING DISTRIBUTION EQUIVALENCE AND STRUCTURE LEARNING FOR CYCLIC AND ACYCLIC DIRECTED GRAPHS	3452
<i>AmirEmad Ghassami, Alan Yang, Negar Kiyavash, Kun Zhang</i>	
PRIVATE COUNTING FROM ANONYMOUS MESSAGES: NEAR-OPTIMAL ACCURACY WITH VANISHING COMMUNICATION OVERHEAD	3463
<i>Badih Ghazi, Ravi Kumar, Pasin Manurangsi, Rasmus Pagh</i>	
ALIGNED CROSS ENTROPY FOR NON-AUTOREGRESSIVE MACHINE TRANSLATION	3473
<i>Marjan Ghazvininejad, Vladimir Karpukhin, Luke Zettlemoyer, Omer Levy</i>	
GRADIENT TEMPORAL-DIFFERENCE LEARNING WITH REGULARIZED CORRECTIONS	3482
<i>S. Ghassian, A. Patterson, S. Garg, D. Gupta, A. White, M. White</i>	
A DISTRIBUTIONAL FRAMEWORK FOR DATA VALUATION	3493
<i>Amirata Ghorbani, Michael P. Kim, James Zou</i>	
FRACTAL GAUSSIAN NETWORKS: A SPARSE RANDOM GRAPH MODEL BASED ON GAUSSIAN MULTIPLICATIVE CHAOS	3503
<i>S. Ghosh, K. Balasubramanian, X. Yang</i>	
REPRESENTATIONS FOR STABLE OFF-POLICY REINFORCEMENT LEARNING	3514
<i>D. Ghosh, M. Bellemare</i>	
ADAPTIVE SKETCHING FOR FAST AND CONVERGENT CANONICAL POLYADIC DECOMPOSITION	3524
<i>Kareem S. Aggour, Alex Gittens, Bülent Yener</i>	
ONE SIZE FITS ALL: CAN WE TRAIN ONE DENOISER FOR ALL NOISE LEVELS?	3534
<i>Abhiram Gnansambandam, Stanley H. Chan</i>	
SUPERPOLYNOMIAL LOWER BOUNDS FOR LEARNING ONE-LAYER NEURAL NETWORKS USING GRADIENT DESCENT	3545
<i>Surbhi Goel, Aravind Gollakota, Zhihan Jin, Sushrut Karmalkar, Adam Klivans</i>	
SIMGANS: SIMULATOR-BASED GENERATIVE ADVERSARIAL NETWORKS FOR ECG SYNTHESIS TO IMPROVE DEEP ECG CLASSIFICATION	3555
<i>Tomer Golany, Daniel Freedman, Kira Radinsky</i>	

UNRAVELING META-LEARNING: UNDERSTANDING FEATURE REPRESENTATIONS FOR FEW-SHOT TASKS	3565
<i>M. Goldblum, S. Reich, L. Fowl, R. Ni, V. Cherepanova, T. Goldstein</i>	
TOWARDS A GENERAL THEORY OF INFINITE-WIDTH LIMITS OF NEURAL CLASSIFIERS	3575
<i>Eugene A. Golikov</i>	
DIFFERENTIALLY PRIVATE SET UNION	3585
<i>Sivakanth Gopi, Pankaj Gulhane, Janardhan Kulkarni, Judy Hanwen Shen, Milad Shokouhi, Sergey Yekhanin</i>	
THE CONTINUOUS CATEGORICAL: A NOVEL SIMPLEX-VALUED EXPONENTIAL FAMILY	3595
<i>Elliott Gordon-Rodriguez, Gabriel Loaiza-Ganem, John P. Cunningham</i>	
AUTOMATIC REPARAMETERISATION OF PROBABILISTIC PROGRAMS	3606
<i>M. Gorinova, D. Moore, M. Hoffman</i>	
INTERPRETABLE OFF-POLICY EVALUATION IN REINFORCEMENT LEARNING BY HIGHLIGHTING INFLUENTIAL TRANSITIONS	3616
<i>Omer Gottesman, Joseph Futoma, Yao Liu, Sonali Parbhoo, Leo Anthony Celi, Emma Brunskill, Finale Doshi-Velez</i>	
LEARNING TO NAVIGATE THE SYNTHETICALLY ACCESSIBLE CHEMICAL SPACE USING REINFORCEMENT LEARNING	3626
<i>S. Gottipati, B. Sattarov, S. Niu, Y. Pathak, H. Wei, S. Liu, K. Thomas, S. Blackburn, C. Coley, J. Tang, S. Chandar, Y. Bengio</i>	
ORDINAL NON-NEGATIVE MATRIX FACTORIZATION FOR RECOMMENDATION	3638
<i>Olivier Gouvert, Thomas Oberlin, Cédric Févotte</i>	
POWER-BERT: ACCELERATING BERT INFERENCE VIA PROGRESSIVE WORD-VECTOR ELIMINATION	3648
<i>Saurabh Goyal, Anamitra Roy Choudhury, Saurabh M. Raje, Venkatesan T. Chakaravarthy, Yogish Sabharwal, Ashish Verma</i>	
PACKIT: A VIRTUAL ENVIRONMENT FOR GEOMETRIC PLANNING	3658
<i>Akit Goyal, Jia Deng</i>	
DROCC: DEEP ROBUST ONE-CLASS CLASSIFICATION	3669
<i>S. Goyal, A. Raghunathan, M. Jain, H. Simhadri, P. Jain</i>	
SCALABLE GAUSSIAN PROCESS SEPARATION FOR KERNELS WITH A NON-STATIONARY PHASE	3680
<i>Jan Grafshoff, Alexandra Jankowski, Philipp Rostalski</i>	
LEARNING THE STEIN DISCREPANCY FOR TRAINING AND EVALUATING ENERGY-BASED MODELS WITHOUT SAMPLING	3690
<i>Will Grathwohl, Kuan-Chieh Wang, Jörn-Henrik Jacobsen, David Duvenaud, Richard Zemel</i>	
ON THE ITERATION COMPLEXITY OF HYPERGRADIENT COMPUTATION	3706
<i>R. Grazi, L. Franceschi, M. Pontil, S. Salzo</i>	
ROBUST LEARNING WITH THE HILBERT-SCHMIDT INDEPENDENCE CRITERION	3717
<i>Daniel Greenfeld, Uri Shalit</i>	
MONTE-CARLO TREE SEARCH AS REGULARIZED POLICY OPTIMIZATION	3727
<i>J. Grill, F. Althe, Y. Tang, T. Hubert, M. Valko, I. Antonoglou, R. Munos</i>	
NEAR-TIGHT MARGIN-BASED GENERALIZATION BOUNDS FOR SUPPORT VECTOR MACHINES	3737
<i>Allan Grønlund, Lior Kamma, Kasper Green Larsen</i>	
IMPLICIT GEOMETRIC REGULARIZATION FOR LEARNING SHAPES	3747
<i>Amos Gropp, Lior Yariv, Niv Haim, Matan Atzmon, Yaron Lipman</i>	
IMPROVING THE GATING MECHANISM OF RECURRENT NEURAL NETWORKS	3758
<i>A. Gu, C. Gulcehre, T. Paine, M. Hoffman, R. Pascanu</i>	
RECURRENT HIERARCHICAL TOPIC-GUIDED RNN FOR LANGUAGE GENERATION	3768
<i>Dandan Guo, Bo Chen, Ruiying Lu, Mingyuan Zhou</i>	
BREAKING THE CURSE OF SPACE EXPLOSION: TOWARDS EFFICIENT NAS WITH CURRICULUM SEARCH	3780
<i>Yong Guo, Yaofu Chen, Yin Zheng, Peilin Zhao, Jian Chen, Junzhou Huang, Mingkui Tan</i>	
CERTIFIED DATA REMOVAL FROM MACHINE LEARNING MODELS	3790
<i>Chuan Guo, Tom Goldstein, Awni Hannun, Laurens van der Maaten</i>	
LTF: A LABEL TRANSFORMATION FRAMEWORK FOR CORRECTING TARGET SHIFT	3801
<i>Jiaxian Guo, Mingming Gong, Tongliang Liu, Kun Zhang, Dacheng Tao</i>	
LEARNING TO BRANCH FOR MULTI-TASK LEARNING	3812
<i>Pengsheng Guo, Chen-Yu Lee, Daniel Ulbricht</i>	
COMMUNICATION-EFFICIENT DISTRIBUTED STOCHASTIC AUC MAXIMIZATION WITH DEEP NEURAL NETWORKS	3822
<i>Z. Guo, M. Liu, Z. Yuan, L. Shen, W. Liu, T. Yang</i>	
BOOTSTRAP LATENT-PREDICTIVE REPRESENTATIONS FOR MULTITASK REINFORCEMENT LEARNING	3833
<i>Z. Guo, B. Pires, B. Piot, J. Grill, F. Althe, R. Munos, M. Azar</i>	
ACCELERATING LARGE-SCALE INFERENCE WITH ANISOTROPIC VECTOR QUANTIZATION	3845
<i>R. Guo, P. Sun, E. Lindgren, Q. Geng, D. Simcha, F. Chern, S. Kumar</i>	
SAFE DEEP SEMI-SUPERVISED LEARNING FOR UNSEEN-CLASS UNLABELED DATA	3855
<i>Lan-Zhe Guo, Zhen-Yu Zhang, Yuan Jiang, Yu-Feng Li, Zhi-Hua Zhou</i>	
NEURAL TOPIC MODELING WITH CONTINUAL LIFELONG LEARNING	3865
<i>Pankaj Gupta, Yatin Chaudhary, Thomas Runkler, Hinrich Schütze</i>	
MULTIDIMENSIONAL SHAPE CONSTRAINTS	3876
<i>M. Gupta, E. Loidor, O. Mangylov, N. Morioka, T. Narayan, S. Zhao</i>	

PART 6

REALM: RETRIEVAL-AUGMENTED LANGUAGE MODEL PRE-TRAINING	3887
<i>K. Guu, K. Lee, Z. Tung, P. Pasupat, M. Chang</i>	
STREAMING SUBMODULAR MAXIMIZATION UNDER A K-SET SYSTEM CONSTRAINT	3897
<i>Ran Haba, Ehsan Kazemi, Moran Feldman, Amin Karbasi</i>	
LET’S AGREE TO AGREE: NEURAL NETWORKS SHARE CLASSIFICATION ORDER ON REAL DATASETS	3908
<i>G. Hacohen, L. Choshen, D. Weinshall</i>	
OPTIMAL APPROXIMATION FOR UNCONSTRAINED NON-SUBMODULAR MINIMIZATION	3919
<i>Marwa El Halabi, Stefanie Jegelka</i>	
FEDBOOST: COMMUNICATION-EFFICIENT ALGORITHMS FOR FEDERATED LEARNING	3931
<i>J. Hamer, M. Mohri, A. Suresh</i>	
POLYNOMIAL TENSOR SKETCH FOR ELEMENT-WISE FUNCTION OF LOW-RANK MATRIX	3942
<i>Insu Han, Haim Avron, Jinwoo Shin</i>	
DRWR: A DIFFERENTIABLE RENDERER WITHOUT RENDERING FOR UNSUPERVISED 3D STRUCTURE LEARNING FROM SILHOUETTE IMAGES	3952
<i>Zhizhong Han, Chao Chen, Yu-Shen Liu, Matthias Zwicker</i>	
SIGUA: FORGETTING MAY MAKE LEARNING WITH NOISY LABELS MORE ROBUST	3964
<i>B. Han, G. Niu, X. Yu, Q. Yao, M. Xu, I. Tsang, M. Sugiyama</i>	
TRAINING BINARY NEURAL NETWORKS THROUGH LEARNING WITH NOISY SUPERVISION	3975
<i>Kai Han, Yunhe Wang, Yixing Xu, Chunjing Xu, Enhua Wu, Chang Xu</i>	
STOCHASTIC SUBSPACE CUBIC NEWTON METHOD	3985
<i>Filip Hanzely, Nikita Doikov, Yurii Nesterov, Peter Richtárik</i>	
VARIANCE REDUCED COORDINATE DESCENT WITH ACCELERATION: NEW METHOD WITH A SURPRISING APPLICATION TO FINITE-SUM PROBLEMS	3997
<i>Filip Hanzely, Dmitry Kovalev, Peter Richtárik</i>	
DATA AMPLIFICATION: INSTANCE-OPTIMAL PROPERTY ESTIMATION	4007
<i>Yi Hao, Alon Orlitsky</i>	
DYNAMIC KNAPSACK OPTIMIZATION TOWARDS EFFICIENT MULTI-CHANNEL SEQUENTIAL ADVERTISING	4018
<i>Xiaotian Hao, Zhaoqing Peng, Yi Ma, Guan Wang, Junqi Jin, Jianye Hao, Shan Chen, Rongquan Bai, Mingzhou Xie, Miao Xu, Zhenzhe Zheng, Chuan Yu, Han Li, Jian Xu, Kun Gai</i>	
IMPROVING GENERALIZATION BY CONTROLLING LABEL-NOISE INFORMATION IN NEURAL NETWORK WEIGHTS	4029
<i>Harutyunyan, Hrayr</i>	
A NATURAL LOTTERY TICKET WINNER: REINFORCEMENT LEARNING WITH ORDINARY NEURAL CIRCUITS	4040
<i>Ramin Hasani, Mathias Lechner, Alexander Amini, Daniela Rus, Radu Grosu</i>	
BAYESIAN GRAPH NEURAL NETWORKS WITH ADAPTIVE CONNECTION SAMPLING	4052
<i>Arman Hasanzadeh, Ehsan Hajiramezani, Shahin Boluki, Mingyuan Zhou, Nick Duffield, Krishna Narayanan, Xiaoning Qian</i>	
COMIC: COMPLEMENTARY TASK LEARNING & MIMICRY FOR REUSABLE SKILLS	4063
<i>Leonard Hasenclever, Fabio Pardo, Raia Hadsell, Nicolas Heess, Josh Merel</i>	
CONTRASTIVE MULTI-VIEW REPRESENTATION LEARNING ON GRAPHS	4074
<i>Kaveh Hassani, Amir Hosein Khasahmadi</i>	
NESTED SUBSPACE ARRANGEMENT FOR REPRESENTATION OF RELATIONAL DATA	4085
<i>Nozomi Hata, Shizuo Kaji, Akihiro Yoshida, Katsuki Fujisawa</i>	
THE TREE ENSEMBLE LAYER: DIFFERENTIABILITY MEETS CONDITIONAL COMPUTATION	4096
<i>Hussein Hazimeh, Natalia Ponomareva, Petros Mol, Zhenyu Tan, Rahul Mazumder</i>	
HIERARCHICALLY DECOUPLED IMITATION FOR MORPHOLOGICAL TRANSFER	4107
<i>Donald J. Hejna III, Pieter Abbeel, Lerrel Pinto</i>	
GRADIENT-FREE ONLINE LEARNING IN GAMES WITH DELAYED REWARDS	4120
<i>Amélie Héliou, Panayotis Mertikopoulos, Zhengyuan Zhou</i>	
DATA-EFFICIENT IMAGE RECOGNITION WITH CONTRASTIVE PREDICTIVE CODING	4130
<i>Olivier J. Héroux, Aravind Srinivas, Jeffrey De Fauw, Ali Razavi, Carl Doersch, S. M. Ali Eslami, Aaron van den Oord</i>	
MINIMAX RATE FOR LEARNING FROM PAIRWISE COMPARISONS IN THE BTL MODEL	4141
<i>Julien Hendrickx, Alex Olshevsky, Venkatesh Saligrama</i>	
STATISTICALLY PRECONDITIONED ACCELERATED GRADIENT METHOD FOR DISTRIBUTED OPTIMIZATION	4151
<i>Hadrien Hendrikx, Lin Xiao, Sébastien Bubeck, Francis Bach, Laurent Massoulié</i>	
COST-EFFECTIVE INTERACTIVE ATTENTION LEARNING WITH NEURAL ATTENTION PROCESSES	4176
<i>Jay Heo, Junhyeon Park, Hyewon Jeong, Kwangjoon Kim, Juho Lee, Eunho Yang, Sung Ju Hwang</i>	
LIKELIHOOD-FREE MCMC WITH AMORTIZED APPROXIMATE RATIO ESTIMATORS	4187
<i>J. Hermans, V. Begy, G. Louppe</i>	
TOWARDS NON-PARAMETRIC DRIFT DETECTION VIA DYNAMIC ADAPTING WINDOW INDEPENDENCE DRIFT DETECTION (DAWIDD)	4199
<i>Fabian Hinder, André Artelt, Barbara Hammer</i>	
OPTIMIZATION AND ANALYSIS OF THE PAP@K METRIC FOR RECOMMENDER SYSTEMS	4210
<i>Gaurush Hiranandani, Warut Vajitbenjaronk, Oluwasanmi Koyejo, Prateek Jain</i>	

OPTIMIZING DYNAMIC STRUCTURES WITH BAYESIAN GENERATIVE SEARCH	4221
<i>Minh Hoang, Carl Kingsford</i>	
LEARNING TASK-AGNOSTIC EMBEDDING OF MULTIPLE BLACK-BOX EXPERTS FOR MULTI-TASK MODEL FUSION	4232
<i>T. Hoang, C. Lam, B. Low, P. Jaillet</i>	
PARAMETERIZED RATE-DISTORTION STOCHASTIC ENCODER	4243
<i>Quan Hoang, Trung Le, Dinh Phung</i>	
TOPOLOGICALLY DENSIFIED DISTRIBUTIONS	4254
<i>C. Hofer, F. Graf, M. Niethammer, R. Kwitt</i>	
GRAPH FILTRATION LEARNING	4264
<i>Christoph D. Hofer, Florian Graf, Bastian Rieck, Marc Niethammer, Roland Kwitt</i>	
BLACK-BOX VARIATIONAL INFERENCE AS DISTILLED LANGEVIN DYNAMICS	4274
<i>Matthew Hoffman, Yi-An Ma</i>	
LEARNING MIXTURES OF GRAPHS FROM EPIDEMIC CASCADES	4292
<i>Jessica Hoffmann, Soumya Basu, Surbhi Goel, Constantine Caramanis</i>	
SET FUNCTIONS FOR TIME SERIES	4303
<i>M. Horn, M. Moor, C. Bock, B. Rieck, K. Borgwardt</i>	
LIFTED DISJOINT PATHS WITH APPLICATION IN MULTIPLE OBJECT TRACKING	4314
<i>Andrea Hornakova, Roberto Henschel, Bodo Rosenhahn, Paul Swoboda</i>	
INFINITE ATTENTION: NNGP AND NTK FOR DEEP ATTENTION NETWORKS	4326
<i>J. Hron, Y. Bahri, J. Sohl-Diskstein, R. Novak</i>	
THE NON-IID DATA QUAGMIRE OF DECENTRALIZED MACHINE LEARNING	4337
<i>K. Hsieh, A. Phanishayee, O. Mutlu, P. Gibbons</i>	
"OTHER-PLAY" FOR ZERO-SHOT COORDINATION	4349
<i>Hengyuan Hu, Adam Lerer, Alex Peysakhovich, Jakob Foerster</i>	
XTREME: A MASSIVELY MULTILINGUAL MULTI-TASK BENCHMARK FOR EVALUATING CROSS-LINGUAL GENERALIZATION	4361
<i>Junjie Hu, Sebastian Ruder, Aditya Siddhant, Graham Neubig, Orhan Firat, Melvin Johnson</i>	
MOMENTUM-BASED POLICY GRADIENT METHODS	4372
<i>Feihu Huang, Shangqian Gao, Jian Pei, Heng Huang</i>	
FROM IMPORTANCE SAMPLING TO DOUBLY ROBUST POLICY GRADIENT	4384
<i>J. Huang, N. Jiang</i>	
EVALUATING LOSSY COMPRESSION RATES OF DEEP GENERATIVE MODELS	4394
<i>S. Huang, A. Makhzani, Y. Cao, R. Grosse</i>	
ONE POLICY TO CONTROL THEM ALL: SHARED MODULAR POLICIES FOR AGENT-AGNOSTIC CONTROL	4405
<i>Wenlong Huang, Igor Mordatch, Deepak Pathak</i>	
COMMUNICATION-EFFICIENT DISTRIBUTED PCA BY RIEMANNIAN OPTIMIZATION	4415
<i>Long-Kai Huang, Sinno Jialin Pan</i>	
IMPROVING TRANSFORMER OPTIMIZATION THROUGH BETTER INITIALIZATION	4425
<i>Xiao Shi Huang, Felipe Pérez, Jimmy Ba, Maksims Volkovs</i>	
MORE INFORMATION SUPERVISED PROBABILISTIC DEEP FACE EMBEDDING LEARNING	4434
<i>Ying Huang, Shangfeng Qiu, Wenwei Zhang, Xianghui Luo, Jinzhuo Wang</i>	
GENERATING PROGRAMMATIC REFERRING EXPRESSIONS VIA PROGRAM SYNTHESIS	4445
<i>Jiani Huang, Calvin Smith, Osbert Bastani, Rishabh Singh, Aws Albarghouthi, Mayur Naik</i>	
INSTAHLIDE: INSTANCE-HIDING SCHEMES FOR PRIVATE DISTRIBUTED LEARNING	4457
<i>Yangsibo Huang, Zhao Song, Kai Li, Sanjeev Arora</i>	
ACCELERATED STOCHASTIC GRADIENT-FREE AND PROJECTION-FREE METHODS	4469
<i>Feihu Huang, Lue Tao, Songcan Chen</i>	
DEEP GRAPH RANDOM PROCESS FOR RELATIONAL-THINKING-BASED SPEECH RECOGNITION	4481
<i>Hengguan Huang, Fuzhao Xue, Hao Wang, Ye Wang</i>	
DYNAMICS OF DEEP NEURAL NETWORKS AND NEURAL TANGENT HIERARCHY	4492
<i>Jiaoyang Huang, Horng-Tzer Yau</i>	
CURVATURE-CORRECTED LEARNING DYNAMICS IN DEEP NEURAL NETWORKS	4502
<i>Dongsung Huh</i>	
MULTIGRID NEURAL MEMORY	4511
<i>Tri Huynh, Michael Maire, Matthew R. Walter</i>	
META-LEARNING WITH SHARED AMORTIZED VARIATIONAL INFERENCE	4522
<i>Ekaterina Iakovleva, Jakob Verbeek, Karteek Alahari</i>	
LINEAR LOWER BOUNDS AND CONDITIONING OF DIFFERENTIABLE GAMES	4533
<i>Adam Ibrahim, Waiss Azizian, Gauthier Gidel, Ioannis Mitliagkas</i>	
FAST DETERMINISTIC CUR MATRIX DECOMPOSITION WITH ACCURACY ASSURANCE	4544
<i>Yasutoshi Ida, Sekitoshi Kanai, Yasuhiro Fujiwara, Tomoharu Iwata, Koh Takeuchi, Hisashi Kashima</i>	
DO WE NEED ZERO TRAINING LOSS AFTER ACHIEVING ZERO TRAINING ERROR?	4554
<i>T. Ishida, I. Yamane, T. Sakai, G. Niu, M. Sugiyama</i>	
SEMI-SUPERVISED LEARNING WITH NORMALIZING FLOWS	4565
<i>Pavel Izmailov, Polina Kirichenko, Marc Finzi, Andrew Gordon Wilson</i>	
IMPLICIT REGULARIZATION OF RANDOM FEATURE MODELS	4581
<i>Arthur Jacot, Berfin Simsek, Francesco Spadaro, Clément Hongler, Franck Gabriel</i>	

CORRELATION CLUSTERING WITH ASYMMETRIC CLASSIFICATION ERRORS	4591
<i>Jafar Jafarov, Sanchit Kalhan, Konstantin Makarychev, Yury Makarychev</i>	
OPTIMAL ROBUST LEARNING OF DISCRETE DISTRIBUTIONS FROM BATCHES	4601
<i>Ayush Jain, Alon Orlitsky</i>	
GENERALIZATION TO NEW ACTIONS IN REINFORCEMENT LEARNING	4611
<i>Ayush Jain, Andrew Szot, Joseph J. Lim</i>	
TAILS OF LIPSCHITZ TRIANGULAR FLOWS	4623
<i>P. Jaini, I. Kobyzev, Y. Yu, M. Brubaker</i>	
LEARNING PORTABLE REPRESENTATIONS FOR HIGH-LEVEL PLANNING	4632
<i>Steven James, Benjamin Rosman, George Konidaris</i>	
DEBIASED SINKHORN BARYCENTERS	4642
<i>Hicham Janati, Marco Cuturi, Alexandre Gramfort</i>	
PARAMETRIC GAUSSIAN PROCESS REGRESSORS	4652
<i>Martin Jankowiak, Geoff Pleiss, Jacob R. Gardner</i>	
INVERSE ACTIVE SENSING: 0.95 MODELING AND UNDERSTANDING TIMELY DECISION-MAKING	4663
<i>Daniel Jarrett, Mihaela van der Schaar</i>	

PART 7

SOURCE SEPARATION WITH DEEP GENERATIVE PRIORS	4674
<i>Vivek Jayaram, John Thickstun</i>	
EXTRAGRADIENT WITH PLAYER SAMPLING FOR FASTER NASH EQUILIBRIUM FINDING	4686
<i>S. Jelassi, C. Domingo-Enrich, D. Scieur, A. Mensch, J. Bruna</i>	
T-GD: TRANSFERABLE GAN-GENERATED IMAGES DETECTION FRAMEWORK	4696
<i>Hyeonseong Jeon, Youngoh Bang, Junyaup Kim, Simon S. Woo</i>	
HISTORY-GRADIENT AIDED BATCH SIZE ADAPTATION FOR VARIANCE REDUCED ALGORITHMS	4712
<i>K. Ji, Z. Wang, B. Weng, Y. Zhou, W. Zhang, Y. Liang</i>	
INFORMATION-THEORETIC LOCAL MINIMA CHARACTERIZATION AND REGULARIZATION	4723
<i>Zhiwei Jia, Hao Su</i>	
OPTIMIZING BLACK-BOX METRICS WITH ADAPTIVE SURROGATES	4734
<i>Qijia Jiang, Olaoluwa Adigun, Harikrishna Narasimhan, Mahdi Milani Fard, Maya Gupta</i>	
BINOCULARS FOR EFFICIENT, NONMYOPIC SEQUENTIAL EXPERIMENTAL DESIGN	4744
<i>Shali Jiang, Henry Chai, Javier González, Roman Garnett</i>	
BEYOND SYNTHETIC NOISE: DEEP LEARNING ON CONTROLLED NOISY LABELS	4754
<i>Lu Jiang, Di Huang, Mason Liu, Weilong Yang</i>	
IMPLICIT CLASS-CONDITIONED DOMAIN ALIGNMENT FOR UNSUPERVISED DOMAIN ADAPTATION	4766
<i>Xiang Jiang, Qicheng Lao, Stan Matwin, Mohammad Havaei</i>	
ASSOCIATIVE MEMORY IN ITERATED OVERPARAMETERIZED SIGMOID AUTOENCODERS	4778
<i>Y. Jiang, C. Pehlevan</i>	
HIERARCHICAL GENERATION OF MOLECULAR GRAPHS USING STRUCTURAL MOTIFS	4789
<i>Wengong Jin, Regina Barzilay, Tommi Jaakkola</i>	
MULTI-OBJECTIVE MOLECULE GENERATION USING INTERPRETABLE SUBSTRUCTURES	4799
<i>Wengong Jin, Regina Barzilay, Tommi Jaakkola</i>	
LEARNING ADVERSARIAL MARKOV DECISION PROCESSES WITH BANDIT FEEDBACK AND UNKNOWN TRANSITION	4810
<i>Chi Jin, Tiancheng Jin, Haipeng Luo, Suvrit Sra, Tiancheng Yu</i>	
REWARD-FREE EXPLORATION FOR REINFORCEMENT LEARNING	4820
<i>Chi Jin, Akshay Krishnamurthy, Max Simchowitz, Tiancheng Yu</i>	
WHAT IS LOCAL OPTIMALITY IN NONCONVEX-NONCONCAVE MINIMAX OPTIMIZATION?	4830
<i>C. Jin, P. Netrapalli, M. Jordan</i>	
EFFICIENTLY SOLVING MDPS WITH STOCHASTIC MIRROR DESCENT	4840
<i>Y. Jin, A. Sidford</i>	
COMPUTATIONAL-STATISTICAL TRADEOFFS IN INFERRING COMBINATORIAL STRUCTURES OF ISING MODEL	4851
<i>Ying Jin, Zhaoran Wang, Junwei Lu</i>	
ADASCALE SGD: A USER-FRIENDLY ALGORITHM FOR DISTRIBUTED TRAINING	4860
<i>Tyler B. Johnson, Pulkit Agrawal, Haijie Gu, Carlos Guestrin</i>	
GUIDED LEARNING OF NONCONVEX MODELS THROUGH SUCCESSIVE FUNCTIONAL GRADIENT OPTIMIZATION	4870
<i>R. Johnson, T. Zhang</i>	
ON RELATIVISTIC F-DIVERGENCES	4880
<i>A. Jolicoeur-Martineau</i>	
FAIR K-CENTERS VIA MAXIMUM MATCHING	4889
<i>Matthew Jones, Huy Lê Nguyễn, Thy Nguyen</i>	
BEING BAYESIAN ABOUT CATEGORICAL PROBABILITY	4899
<i>Taejong Joo, Uijung Chung, Min-Gwan Seo</i>	
EVALUATING THE PERFORMANCE OF REINFORCEMENT LEARNING ALGORITHMS	4911
<i>Scott M. Jordan, Yash Chandak, Daniel Cohen, Mengxue Zhang, Philip S. Thomas</i>	

EVALUATING THE PERFORMANCE OF REINFORCEMENT LEARNING ALGORITHMS	4923
<i>Scott M. Jordan, Yash Chandak, Daniel Cohen, Mengxue Zhang, Philip S. Thomas</i>	
STOCHASTIC DIFFERENTIAL EQUATIONS WITH VARIATIONAL WISHART DIFFUSIONS	4941
<i>Martin Jørgensen, Marc Peter Deisenroth, Hugh Salimbeni</i>	
A SIMPLER APPROACH TO ACCELERATED STOCHASTIC OPTIMIZATION: ITERATIVE AVERAGING MEETS OPTIMISM	4951
<i>P. Joulani, A. Raj, A. Gyorgy, C. Szepesvari</i>	
SETS CLUSTERING	4961
<i>I. Jubran, M. Tukan, A. Maalouf, D. Feldman</i>	
DISTRIBUTION AUGMENTATION FOR GENERATIVE MODELING	4973
<i>Heewoo Jun, Rewon Child, Mark Chen, John Schulman, Aditya Ramesh, Alec Radford, Ilya Sutskever</i>	
SUB-GOAL TREES – A FRAMEWORK FOR GOAL-BASED REINFORCEMENT LEARNING	4987
<i>T. Jurgenson, O. Avner, E. Groshev, A. Tamar</i>	
PARTIAL TRACE REGRESSION AND LOW-RANK KRAUS DECOMPOSITION	4998
<i>Hachem Kadri, Stéphane Ayache, Riikka Huusari, Alain Rakotomamonjy, Liva Ralaivola</i>	
STRATEGYPROOF MEAN ESTIMATION FROM MULTIPLE-CHOICE QUESTIONS	5009
<i>Anson Kahng, Gregory Kehne, Ariel D. Procaccia</i>	
VARIATIONAL AUTOENCODERS WITH RIEMANNIAN BROWNIAN MOTION PRIORS	5020
<i>Dimitris Kalatzis, David Eklund, Georgios Arvanitidis, Soren Hauberg</i>	
DOUBLE REINFORCEMENT LEARNING FOR EFFICIENT AND ROBUST OFF-POLICY EVALUATION	5034
<i>N. Kallus, M. Uehara</i>	
STATISTICALLY EFFICIENT OFF-POLICY POLICY GRADIENTS	5045
<i>N. Kallus, M. Uehara</i>	
ON THE POWER OF COMPRESSED SENSING WITH GENERATIVE MODELS	5057
<i>Akshay Kamath, Sushrut Karmalkar, Eric Price</i>	
LEARNING AND EVALUATING CONTEXTUAL EMBEDDING OF SOURCE CODE	5066
<i>Aditya Kanade, Petros Maniatis, Gogul Balakrishnan, Kensen Shi</i>	
OPERATION-AWARE SOFT CHANNEL PRUNING USING DIFFERENTIABLE MASKS	5078
<i>Minsoo Kang, Bohyung Han</i>	
SCAFFOLD: STOCHASTIC CONTROLLED AVERAGING FOR FEDERATED LEARNING	5088
<i>S. Karimireddy, S. Kale, M. Mohri, S. Reddi, S. Stich, A. Suresh</i>	
NON-AUTOREGRESSIVE MACHINE TRANSLATION WITH DISENTANGLED CONTEXT TRANSFORMER	5100
<i>Jungo Kasai, James Cross, Marjan Ghazvininejad, Jiatao Gu</i>	
TRANSFORMERS ARE RNNs: FAST AUTOREGRESSIVE TRANSFORMERS WITH LINEAR ATTENTION	5112
<i>Angelos Katharopoulos, Apoorv Vyas, Nikolaos Pappas, François Fleuret</i>	
RATE-DISTORTION OPTIMIZATION GUIDED AUTOENCODER FOR ISOMETRIC EMBEDDING IN EUCLIDEAN LATENT SPACE	5122
<i>K. Kato, J. Zhou, T. Sasaki, A. Nakagawa</i>	
EFFICIENT NON-CONJUGATE GAUSSIAN PROCESS FACTOR MODELS FOR SPIKE COUNT DATA USING POLYNOMIAL APPROXIMATIONS	5133
<i>Stephen L. Keeley, David M. Zoltowski, Yiyi Yu, Jacob L. Yates, Spencer L. Smith, Jonathan W. Pillow</i>	
QUANTUM EXPECTATION-MAXIMIZATION FOR GAUSSIAN MIXTURE MODELS	5143
<i>Iordanis Kerenidis, Alessandro Luongo, Anupam Prakash</i>	
DIFFERENTIABLE LIKELIHOODS FOR FAST INVERSION OF 'LIKELIHOOD-FREE' DYNAMICAL SYSTEMS	5154
<i>Hans Kersting, Nicholas Krämer, Martin Schiegg, Christian Daniel, Michael Tiemann, Philipp Hennig</i>	
FEATURE NOISE INDUCES LOSS DISCREPANCY ACROSS GROUPS	5165
<i>F. Khani, P. Liang</i>	
ENTROPY MINIMIZATION IN EMERGENT LANGUAGES	5176
<i>Eugene Kharitonov, Rahma Chaabouni, Diane Bouchacourt, Marco Baroni</i>	
PRIVATE OUTSOURCED BAYESIAN OPTIMIZATION	5187
<i>D. Kharkovskii, Z. Dai, B. Low</i>	
WHAT CAN I DO HERE? A THEORY OF AFFORDANCES IN REINFORCEMENT LEARNING	5199
<i>K. Khetarpal, Z. Ahmed, G. Comanici, D. Abel, D. Precup</i>	
UNIFORM CONVERGENCE OF RANK-WEIGHTED LEARNING	5210
<i>Justin Khim, Liu Leqi, Adarsh Prasad, Pradeep Ravikumar</i>	
FACT: A DIAGNOSTIC FOR GROUP FAIRNESS TRADE-OFFS	5220
<i>J. Kim, J. Chen, A. Talwalkar</i>	
PUZZLE MIX: EXPLOITING SALIENCY AND LOCAL STATISTICS FOR OPTIMAL MIXUP	5231
<i>Jang-Hyun Kim, Wonho Choo, Hyun Oh Song</i>	
DOMAIN ADAPTIVE IMITATION LEARNING	5242
<i>Kuno Kim, Yihong Gu, Jiaming Song, Shengjia Zhao, Stefano Ermon</i>	
VARIATIONAL INFERENCE FOR SEQUENTIAL DATA WITH FUTURE LIKELIHOOD ESTIMATES	5252
<i>Geon-Hyeong Kim, Youngsoo Jang, Hongseok Yang, Kee-Eung Kim</i>	
ACTIVE WORLD MODEL LEARNING WITH PROGRESS CURIOSITY	5262
<i>Kuno Kim, Megumi Sano, Julian De Freitas, Nick Haber, Daniel Yamins</i>	
BAYESIAN EXPERIMENTAL DESIGN FOR IMPLICIT MODELS BY MUTUAL INFORMATION NEURAL ESTIMATION	5272
<i>Steven Kleinegesse, Michael U. Gutmann</i>	

OPTIMAL CONTINUAL LEARNING HAS PERFECT MEMORY AND IS NP-HARD	5283
<i>Jeremias Knoblauch, Hisham Husain, Tom Diethe</i>	
CONCEPT BOTTLENECK MODELS	5294
<i>P. Koh, T. Nguyen, Y. Tang, S. Mussmann, E. Pierson, B. Kim, P. Liang</i>	
LEARNING SIMILARITY METRICS FOR NUMERICAL SIMULATIONS	5305
<i>Georg Kohl, Kiwon Um, Nils Thuerey</i>	
EQUIVARIANT FLOWS: EXACT LIKELIHOOD GENERATIVE LEARNING FOR SYMMETRIC DENSITIES	5317
<i>J. Kohler, L. Klein, F. Noe</i>	
ONLINE LEARNING FOR ACTIVE CACHE SYNCHRONIZATION	5327
<i>Andrey Kolobov, Sébastien Bubeck, Julian Zimmert</i>	
A UNIFIED THEORY OF DECENTRALIZED SGD WITH CHANGING TOPOLOGY AND LOCAL UPDATES	5337
<i>Anastasia Koloskova, Nicolas Loizou, Sadra Boreiri, Martin Jaggi, Sebastian U. Stich</i>	
META-LEARNING FOR MIXED LINEAR REGRESSION	5350
<i>W. Kong, R. Somani, Z. Song, S. Kakade, S. Oh</i>	
SDE-NET: EQUIPPING DEEP NEURAL NETWORKS WITH UNCERTAINTY ESTIMATES	5361
<i>Lingkai Kong, Jimeng Sun, Chao Zhang</i>	
ON THE SAMPLE COMPLEXITY OF ADVERSARIAL MULTI-SOURCE PAC LEARNING	5372
<i>N. Konstantinov, E. Frantar, D. Alistarh, C. Lampert</i>	
ASYNCHRONOUS COAGENT NETWORKS	5382
<i>James E. Kostas, Chris Nota, Philip S. Thomas</i>	
BEING BAYESIAN, EVEN JUST A BIT, FIXES OVERCONFIDENCE IN RELU NETWORKS	5392
<i>A. Kristiadi, M. Hein, P. Hennig</i>	
A SEQUENTIAL SELF TEACHING APPROACH FOR IMPROVING GENERALIZATION IN SOUND EVENT RECOGNITION	5403
<i>Anurag Kumar, Vamsi Krishna Ithapu</i>	
CURSE OF DIMENSIONALITY ON RANDOMIZED SMOOTHING FOR CERTIFIABLE ROBUSTNESS	5414
<i>Aounon Kumar, Alexander Levine, Tom Goldstein, Soheil Feizi</i>	
UNDERSTANDING SELF-TRAINING FOR GRADUAL DOMAIN ADAPTATION	5424
<i>A. Kumar, T. Ma, P. Liang</i>	

PART 8

ON IMPLICIT REGULARIZATION IN -VAES	5436
<i>Abhishek Kumar, Ben Poole</i>	
PROBLEMS WITH SHAPLEY-VALUE-BASED EXPLANATIONS AS FEATURE IMPORTANCE MEASURES	5447
<i>I. Elizabeth Kumar, Suresh Venkatasubramanian, Carlos Scheidegger, Sorelle A. Friedler</i>	
EFFICIENT IDENTIFICATION IN LINEAR STRUCTURAL CAUSAL MODELS WITH AUXILIARY CUTSETS	5457
<i>D. Kumor, C. Cinelli, E. Bareinboim</i>	
TWO ROUTES TO SCALABLE CREDIT ASSIGNMENT WITHOUT WEIGHT SYMMETRY	5467
<i>Daniel Kumin, Aran Nayebi, Javier Sagastuy-Brena, Surya Ganguli, Jonathan M. Bloom, Daniel L. K. Yamins</i>	
ONLINE DENSE SUBGRAPH DISCOVERY VIA BLURRED-GRAPH FEEDBACK	5478
<i>Y. Kuroki, A. Miyauchi, J. Honda, M. Sugiyama</i>	
INDUCING AND EXPLOITING ACTIVATION SPARSITY FOR FAST NEURAL NETWORK INFERENCE	5489
<i>Mark Kurtz, Justin Kopinsky, Rati Gelashvili, Alexander Matveev, John Carr, Michael Goin, William Leiserson, Sage Moore, Bill Nell, Nir Shavit, Dan Alistarh</i>	
SOFT THRESHOLDWEIGHT REPARAMETERIZATION FOR LEARNABLE SPARSITY	5500
<i>A. Kusupati, V. Ramantujan, R. Somani, M. Wortsman, P. Jain, S. Kakade, A. Farhadi</i>	
CONTROLLING OVERESTIMATION BIAS WITH TRUNCATED MIXTURE OF CONTINUOUS DISTRIBUTIONAL QUANTILE CRITICS	5512
<i>A. Kuznetsov, P. Shvechikov, A. Grishin, D. Vetrov</i>	
PRINCIPLED LEARNING METHOD FOR WASSERSTEIN DISTRIBUTIONALLY ROBUST OPTIMIZATION WITH LOCAL PERTURBATIONS	5523
<i>Yongchan Kwon, Wonyoung Kim, Joong-Ho Won, Myunghee Cho Paik</i>	
CONCENTRATION BOUNDS FOR CVAR ESTIMATION:[0.5EX] THE CASES OF LIGHT-TAILED AND HEAVY-TAILED DISTRIBUTIONS	5533
<i>Prashanth L.A., Krishna Jagannathan, Ravi Kumar Kolla</i>	
OPTIMAL RANDOMIZED FIRST-ORDER METHODS FOR LEAST-SQUARES PROBLEMS	5543
<i>J. Lacotte, M. Pilanci</i>	
DUALITY IN RKHSS WITH INFINITE DIMENSIONAL OUTPUTS: APPLICATION TO ROBUST LOSSES	5554
<i>P. Laforgue, A. Lambert, L. Brogat-Motte, F. Alche-Buc</i>	
RECHT-RÉ NONCOMMUTATIVE ARITHMETIC-GEOMETRIC MEAN CONJECTURE IS FALSE	5564
<i>Zehua Lai, Lek-Heng Lim</i>	
BIDIRECTIONAL MODEL-BASED POLICY OPTIMIZATION	5574
<i>H. Lai, J. Shen, W. Zhang, Y. Yu</i>	
ROBUST AND STABLE BLACK BOX EXPLANATIONS	5584
<i>Himabindu Lakkaraju, Nino Arsov, Osbert Bastani</i>	

CURL: CONTRASTIVE UNSUPERVISED REPRESENTATIONS FOR REINFORCEMENT LEARNING	5595
<i>M. Laskin, A. Srinivas, P. Abbeel</i>	
EFFICIENT PROXIMAL MAPPING OF THE 1-PATH-NORM OF SHALLOW NETWORKS	5607
<i>F. Latorre, P. Rolland, N. Hallak, V. Cevher</i>	
LEARNING WITH GOOD FEATURE REPRESENTATIONS IN BANDITS AND IN RL WITH A GENERATIVE MODEL	5618
<i>Tor Lattimore, Csaba Szepesvári, Gellért Weisz</i>	
INERTIAL BLOCK PROXIMAL METHODS FOR NON-CONVEX NON-SMOOTH OPTIMIZATION	5627
<i>Le Thi Khanh Hien, Nicolas Gillis, Panagiotis Patrinos</i>	
SELF-ATTENTIVE ASSOCIATIVE MEMORY	5638
<i>Hung Le, Truyen Tran, Svetha Venkatesh</i>	
CAUSAL EFFECT IDENTIFIABILITY UNDER PARTIAL-OBSERVABILITY	5648
<i>Sanghack Lee, Elias Bareinboim</i>	
ESTIMATING MODEL UNCERTAINTY OF NEURAL NETWORKS IN SPARSE INFORMATION FORM	5658
<i>Jongseok Lee, Matthias Humt, Jianxiang Feng, Rudolph Triebel</i>	
SELF-SUPERVISED LABEL AUGMENTATION VIA INPUT TRANSFORMATIONS	5670
<i>Hankook Lee, Sung Ju Hwang, Jinwoo Shin</i>	
BATCH REINFORCEMENT LEARNING WITH HYPERPARAMETER GRADIENTS	5681
<i>Byung-Jun Lee, Jongmin Lee, Peter Vrancx, Dongho Kim, Kee-Eung Kim</i>	
ACCELERATED MESSAGE PASSING FOR ENTROPY-REGULARIZED MAP INFERENCE	5692
<i>Jonathan N. Lee, Aldo Pacchiano, Peter Bartlett, Michael I. Jordan</i>	
LEARNING COMPOUND TASKS WITHOUT TASK-SPECIFIC KNOWLEDGE VIA IMITATION AND SELF-SUPERVISED LEARNING	5703
<i>Sang-Hyun Lee, Seung-Woo Seo</i>	
CONTEXT-AWARE DYNAMICS MODEL FOR GENERALIZATION IN MODEL-BASED REINFORCEMENT LEARNING	5713
<i>Kimin Lee, Younggyo Seo, Seunghyun Lee, Honglak Lee, Jinwoo Shin</i>	
TEMPORAL PHENOTYPING USING DEEP PREDICTIVE CLUSTERING OF DISEASE PROGRESSION	5723
<i>Changhee Lee, Mihaela van der Schaar</i>	
TENSOR DENOISING AND COMPLETION BASED ON ORDINAL OBSERVATIONS	5734
<i>Chanwoo Lee, Miaoyan Wang</i>	
ANALYTIC MARCHING: AN ANALYTIC MESHING SOLUTION FROM DEEP IMPLICIT SURFACE NETWORKS	5745
<i>Jiabao Lei, Kui Jia</i>	
SGD LEARNS ONE-LAYER NETWORKS IN WGANs	5755
<i>Q. Lei, J. Lee, A. Dimakis, C. Daskalakis</i>	
FINE-GRAINED ANALYSIS OF STABILITY AND GENERALIZATION FOR STOCHASTIC GRADIENT DESCENT	5765
<i>Yunwen Lei, Yiming Ying</i>	
LEARNING QUADRATIC GAMES ON NETWORKS	5776
<i>Yan Leng, Xiaowen Dong, Junfeng Wu, Alex Pentland</i>	
ACFLOW: FLOW MODELS FOR ARBITRARY CONDITIONAL LIKELIHOODS	5787
<i>Y. Li, S. Akbar, J. Oliva</i>	
MANIFOLD IDENTIFICATION FOR ULTIMATELY COMMUNICATION-EFFICIENT DISTRIBUTED OPTIMIZATION	5798
<i>Y. Li, W. Chiang, C. Lee</i>	
NEURAL ARCHITECTURE SEARCH IN A PROXY VALIDATION LOSS LANDSCAPE	5809
<i>Yanxi Li, Minjing Dong, Yunhe Wang, Chang Xu</i>	
PENNI: PRUNED KERNEL SHARING FOR EFFICIENT CNN INFERENCE	5819
<i>Shiyu Li, Edward Hanson, Hai Li, Yiran Chen</i>	
IMPLICIT EULER SKIP CONNECTIONS: ENHANCING ADVERSARIAL ROBUSTNESS VIA NUMERICAL STABILITY	5830
<i>Mingjie Li, Lingshen He, Zhouchen Lin</i>	
CLOSED LOOP NEURAL-SYMBOLIC LEARNING VIA INTEGRATING NEURAL PERCEPTION, GRAMMAR PARSING, AND SYMBOLIC REASONING	5840
<i>Qing Li, Siyuan Huang, Yining Hong, Yixin Chen, Ying Nian Wu, Song-Chun Zhu</i>	
ACCELERATION FOR COMPRESSED GRADIENT DESCENT IN DISTRIBUTED AND FEDERATED OPTIMIZATION	5851
<i>Zhize Li, Dmitry Kovalev, Xun Qian, Peter Richtárik</i>	
ON THE RELATION BETWEEN QUALITY-DIVERSITY EVALUATION AND DISTRIBUTION-FITTING GOAL IN TEXT GENERATION	5861
<i>Jianing Li, Yanyan Lan, Jiafeng Guo, Xueqi Cheng</i>	
LATENT SPACE FACTORISATION AND MANIPULATION VIA MATRIX SUBSPACE PROJECTION	5872
<i>Xiao Li, Chenghua Lin, Ruizhe Li, Chaozheng Wang, Frank Guerin</i>	
VISUAL GROUNDING OF LEARNED PHYSICAL MODELS	5883
<i>Yunzhu Li, Toru Lin, Kexin Yi, Daniel M. Bear, Daniel L. K. Yamins, Jiajun Wu, Joshua B. Tenenbaum, Antonio Torralba</i>	
LEARNING FROM IRREGULARLY-SAMPLED TIME SERIES: A MISSING DATA PERSPECTIVE	5893
<i>S. Li, B. Marlin</i>	
EVOLUTIONARY TOPOLOGY SEARCH FOR TENSOR NETWORK DECOMPOSITION	5903
<i>Chao Li, Zhun Sun</i>	

TRAIN LARGE, THEN COMPRESS: RETHINKING MODEL SIZE FOR EFFICIENT TRAINING AND INFERENCE OF TRANSFORMERS	5914
<i>Z. Li, E. Wallace, S. Shen, K. Lin, K. Keutzer, D. Klein, J. Gonzalez</i>	
ALMOST TUNE-FREE VARIANCE REDUCTION	5925
<i>B. Li, L. Wang, G. Giannakis</i>	
NEARLY LINEAR ROW SAMPLING ALGORITHM FOR QUANTILE REGRESSION	5935
<i>Y. Li, R. Wang, L. Yang, H. Zhang</i>	
TEMPORAL LOGIC POINT PROCESSES	5946
<i>S. Li, L. Wang, R. Zhang, X. Chang, X. Liu, Y. Xie, Y. Qi, L. Song</i>	
INPUT-SPARSITY LOW RANK APPROXIMATION IN SCHATTEN NORM	5957
<i>Yi Li, David P. Woodruff</i>	
RIFLE: BACKPROPAGATION IN DEPTH FOR DEEP TRANSFER LEARNING THROUGH RE-INITIALIZING THE FULLY-CONNECTED LAYER	5966
<i>Xingjian Li, Haoyi Xiong, Haozhe An, Chengzhong Xu, Dejing Dou</i>	
ON A PROJECTIVE ENSEMBLE APPROACH TO TWO SAMPLE TEST FOR EQUALITY OF DISTRIBUTIONS	5976
<i>Zhimei Li, Yaowu Zhang</i>	
DO WE REALLY NEED TO ACCESS THE SOURCE DATA? SOURCE HYPOTHESIS TRANSFER FOR UNSUPERVISED DOMAIN ADAPTATION	5984
<i>Jian Liang, Dapeng Hu, Jiashi Feng</i>	
VARIABLE SKIPPING FOR AUTOREGRESSIVE RANGE DENSITY ESTIMATION	5996
<i>Eric Liang, Zongheng Yang, Ion Stoica, Pieter Abbeel, Yan Duan, Xi Chen</i>	
ADAPTIVE DROPLET ROUTING IN DIGITAL MICROFLUIDIC BIOCHIPS USING DEEP REINFORCEMENT LEARNING	6006
<i>Tung-Che Liang, Zhanwei Zhong, Yaas Bigdeli, Tsung-Yi Ho, Krishnendu Chakrabarty, Richard Fair</i>	
AR-DAE: TOWARDS UNBIASED NEURAL ENTROPY GRADIENT ESTIMATION	6017
<i>Jae Hyun Lim, Aaron Courville, Christopher Pal, Chin-Wei Huang</i>	
HIERARCHICAL VERIFICATION FOR ADVERSARIAL ROBUSTNESS	6028
<i>Cong Han Lim, Raquel Urtasun, Ersin Yumer</i>	
ON GRADIENT DESCENT ASCENT FOR NONCONVEX-CONCAVE MINIMAX PROBLEMS	6039
<i>T. Lin, C. Jin, M. Jordan</i>	
EXTRAPOLATION FOR LARGE-BATCH TRAINING IN DEEP LEARNING	6050
<i>Tao Lin, Lingjing Kong, Sebastian U. Stich, Martin Jaggi</i>	
ON THE THEORETICAL PROPERTIES OF THE NETWORK JACKKNIFE	6061
<i>Qiaohui Lin, Robert Lunde, Purnamrita Sarkar</i>	
HANDLING THE POSITIVE-DEFINITE CONSTRAINT IN THE BAYESIAN LEARNING RULE	6072
<i>W. Lin, M. Schmidt, M. Khan</i>	
INFOGAN-CR AND MODELCENTRALITY: SELF-SUPERVISED MODEL TRAINING AND SELECTION FOR DISENTANGLING GANS	6083
<i>Z. Lin, K. Thekumparampil, G. Fanti, S. Oh</i>	
IMPROVING GENERATIVE IMAGINATION IN OBJECT-CENTRIC WORLD MODELS	6096
<i>Zhixuan Lin, Yi-Fu Wu, Skand Peri, Bofeng Fu, Jindong Jiang, Sungjin Ahn</i>	
GENERALIZED AND SCALABLE OPTIMAL SPARSE DECISION TREES	6106
<i>Jimmy Lin, Chudi Zhong, Diane Hu, Cynthia Rudin, Margo Seltzer</i>	
FINITE-TIME LAST-ITERATE CONVERGENCE FOR MULTI-AGENT LEARNING IN GAMES	6117
<i>T. Lin, Z. Zhou, P. Mertikopoulos, M. Jordan</i>	
TIME-AWARE LARGE KERNEL CONVOLUTIONS	6128
<i>Vasileios Lioutas, Yuhong Guo</i>	
UNDERSTANDING THE CURSE OF HORIZON IN OFF-POLICY EVALUATION VIA CONDITIONAL IMPORTANCE SAMPLING	6140
<i>Yao Liu, Pierre-Luc Bacon, Emma Brunskill</i>	
SPARSE SHRUNK ADDITIVE MODELS	6150
<i>Guodong Liu, Hong Chen, Heng Huang</i>	
BOOSTING DEEP NEURAL NETWORK EFFICIENCY WITH DUAL-MODULE INFERENCE	6161
<i>Liu Liu, Lei Deng, Zhaodong Chen, Yuke Wang, Shuangchen Li, Jingwei Zhang, Yihua Yang, Zhenyu Gu, Yufei Ding, Yuan Xie</i>	
SAMPLE COMPLEXITY BOUNDS FOR 1-BIT COMPRESSIVE SENSING AND BINARY STABLE EMBEDDINGS WITH GENERATIVE PRIORS	6172
<i>Zhaoqiang Liu, Selwyn Gomes, Avtansh Tiwari, Jonathan Scarlett</i>	
PEER LOSS FUNCTIONS: LEARNING FROM NOISY LABELS WITHOUT KNOWING NOISE RATES	6182
<i>Yang Liu, Hongyi Guo</i>	
AN IMITATION LEARNING APPROACH FOR CACHE REPLACEMENT	6193
<i>E. Liu, M. Hashemi, K. Swersky, P. Ranganathan, J. Ahn</i>	
EXPLORATION THROUGH REWARD BIASING: REWARD-BIASED MAXIMUM LIKELIHOOD ESTIMATION FOR STOCHASTIC MULTI-ARMED BANDITS	6204
<i>Xi Liu, Ping-Chun Hsieh, Yu-Heng Hung, Anirban Bhattacharya, P. R. Kumar</i>	

PART 9

HALLUCINATIVE TOPOLOGICAL MEMORY FOR ZERO-SHOT VISUAL PLANNING	6215
<i>Kara Liu, Thanard Kurutach, Christine Tung, Pieter Abbeel, Aviv Tamar</i>	
A CHANCE-CONSTRAINED GENERATIVE FRAMEWORK FOR SEQUENCE OPTIMIZATION	6227
<i>Xianggen Liu, Qiang Liu, Sen Song, Jian Peng</i>	
MIN-MAX OPTIMIZATION WITHOUT GRADIENTS: CONVERGENCE AND APPLICATIONS TO BLACK-BOX EVATION AND POISONING ATTACKS	6238
<i>S. Liu, S. Lu, X. Chen, Y. Feng, K. Xu, A. Al-Dujaili, M. Hong, U. O'Reilly</i>	
MEDIAN MATRIX COMPLETION: FROM EMBARRASSMENT TO OPTIMALITY	6250
<i>Weidong Liu, Xiaojun Mao, Raymond K. W. Wong</i>	
A GENERIC FIRST-ORDER ALGORITHMIC FRAMEWORK FOR BI-LEVEL PROGRAMMING BEYOND LOWER-LEVEL SINGLETON	6261
<i>R. Liu, P. Mu, X. Yuan, S. Zeng, J. Zhang</i>	
LEARNING DEEP KERNELS FOR NON-PARAMETRIC TWO-SAMPLE TESTS	6272
<i>Feng Liu, Wenkai Xu, Jie Lu, Guangquan Zhang, Arthur Gretton, D. J. Sutherland</i>	
LEARNING TO ENCODE POSITION FOR TRANSFORMER WITH CONTINUOUS DYNAMICAL MODEL	6283
<i>Xuanqing Liu, Hsiang-Fu Yu, Inderjit S. Dhillon, Cho-Jui Hsieh</i>	
FINDING TRAINABLE SPARSE NETWORKS THROUGH NEURAL TANGENT TRANSFER	6292
<i>Tianlin Liu, Friedemann Zenke</i>	
WEAKLY-SUPERVISED DISENTANGLEMENT WITHOUT COMPROMISES	6304
<i>Francesco Locatello, Ben Poole, Gunnar Rätsch, Bernhard Schölkopf, Olivier Bachem, Michael Tschannen</i>	
TOO RELAXED TO BE FAIR	6316
<i>Michael Lohaus, Michaël Perrot, Ulrike von Luxburg</i>	
STOCHASTIC HAMILTONIAN GRADIENT METHODS FOR SMOOTH GAMES	6326
<i>Nicolas Loizou, Hugo Berard, Alexia Jolicoeur-Martineau, Pascal Vincent, Simon Lacoste-Julien, Ioannis Mitliagkas</i>	
ERROR ESTIMATION FOR SKETCHED SVD VIA THE BOOTSTRAP	6338
<i>Miles E. Lopes, N. Benjamin Erichson, Michael W. Mahoney</i>	
DIFFERENTIATING THROUGH THE FRETCHET MEAN	6349
<i>A. Lou, I. Katsman, Q. Jiang, S. Belongie, S. Lim, C. Sa</i>	
WORKING MEMORY GRAPHS	6360
<i>R. Loynd, R. Fernandez, A. Celikyilmaz, A. Swaminathan, M. Hausknecht</i>	
MONQUA: MODULO QUANTIZED COMMUNICATION IN DECENTRALIZED SGD	6371
<i>Yucheng Lu, Christopher De Sa</i>	
A MEAN-FIELD ANALYSIS OF DEEP RESNET AND BEYOND: TOWARDS PROVABLE OPTIMIZATION VIA OVERPARAMETERIZATION FROM DEPTH	6382
<i>Yiping Lu, Chao Ma, Yulong Lu, Jianfeng Lu, Lexing Ying</i>	
COUNTERING LANGUAGE DRIFT WITH SEEDED ITERATED LEARNING	6393
<i>Y. Lu, S. Singhal, F. Strub, O. Pietquin, A. Courville</i>	
DOES LABEL SMOOTHING MITIGATE LABEL NOISE?	6404
<i>M. Lukasik, S. Bhojanapalli, A. Menon, S. Kumar</i>	
IMPROVED COMMUNICATION COST IN DISTRIBUTED PAGERANK COMPUTATION – A THEORETICAL STUDY	6415
<i>Siqiang Luo</i>	
PROGRESSIVE GRAPH LEARNING FOR OPEN-SET DOMAIN ADAPTATION	6424
<i>Yadan Luo, Zijian Wang, Zi Huang, Mahsa Baktashmotlagh</i>	
ADVERSARIAL NONNEGATIVE MATRIX FACTORIZATION	6435
<i>Lei Luo, Yanfu Zhang, Heng Huang</i>	
LEARNING ALGEBRAIC MULTIGRID USING GRAPH NEURAL NETWORKS	6445
<i>Ilay Luz, Meirav Galun, Haggai Maron, Ronen Basri, Irad Yavneh</i>	
PROGRESSIVE IDENTIFICATION OF TRUE LABELS FOR PARTIAL-LABEL LEARNING	6456
<i>Jiaqi Lv, Miao Xu, Lei Feng, Gang Niu, Xin Geng, Masashi Sugiyama</i>	
BANDITS WITH ADVERSARIAL SCALING	6467
<i>Thodoris Lykouris, Vahab Mirrokni, Renato Paes Leme</i>	
EFFICIENT CONTINUOUS PARETO EXPLORATION IN MULTI-TASK LEARNING	6478
<i>Pingchuan Ma, Tao Du, Wojciech Matusik</i>	
CONVEX REPRESENTATION LEARNING FOR GENERALIZED INVARIANCE IN SEMI-INNER-PRODUCT SPACE	6488
<i>Yingyi Ma, Vignesh Ganapathiraman, Yaoliang Yu, Xinhua Zhang</i>	
NORMALIZED LOSS FUNCTIONS FOR DEEP LEARNING WITH NOISY LABELS	6499
<i>Xingjun Ma, Hanxun Huang, Yisen Wang, Simone Romano, Sarah Erfani, James Bailey</i>	
QUADRATICALLY REGULARIZED SUBGRADIENT METHODS FOR WEAKLY CONVEX OPTIMIZATION WITH WEAKLY CONVEX CONSTRAINTS	6510
<i>Runchao Ma, Qihang Lin, Tianbao Yang</i>	
UNDERSTANDING THE IMPACT OF MODEL INCOHERENCE ON CONVERGENCE OF INCREMENTAL SGD WITH RANDOM RESHUFFLE	6521
<i>S. Ma, Y. Zhou</i>	

ADVERSARIAL NEURAL PRUNING WITH LATENT VULNERABILITY SUPPRESSION	6531
<i>Divyam Madaan, Jinwoo Shin, Sung Ju Hwang</i>	
INDIVIDUAL FAIRNESS FOR K-CLUSTERING	6542
<i>Sepideh Mahabadi, Ali Vakilian</i>	
MULTI-TASK LEARNING WITH USER PREFERENCES: GRADIENT DESCENT WITH CONTROLLED ASCENT IN PARETO OPTIMIZATION	6553
<i>D. Mahapatra, V. Rajan</i>	
HOW RECURRENT NETWORKS IMPLEMENT CONTEXTUAL PROCESSING IN SENTIMENT ANALYSIS	6564
<i>Niru Maheswaranathan, David Sussillo</i>	
CONVERGENCE OF A STOCHASTIC GRADIENT METHOD WITH MOMENTUM FOR NON-SMOOTH NON-CONVEX OPTIMIZATION	6576
<i>Vien V. Mai, Mikael Johansson</i>	
CONVERGENCE OF A STOCHASTIC GRADIENT METHOD WITH MOMENTUM FOR NON-SMOOTH NON-CONVEX OPTIMIZATION	6586
<i>Vien V. Mai, Mikael Johansson</i>	
ADVERSARIAL ROBUSTNESS AGAINST THE UNION OF MULTIPLE PERTURBATION MODELS	6596
<i>Pratyush Maini, Eric Wong, J. Zico Kolter</i>	
EVOLUTIONARY REINFORCEMENT LEARNING FOR SAMPLE-EFFICIENT MULTIAGENT COORDINATION	6607
<i>Shauharda Khadka, Somdeb Majumdar, Santiago Miret, Stephen McAleer, Kagan Tumer</i>	
ESTIMATION OF BOUNDS ON POTENTIAL OUTCOMES FOR DECISION MAKING	6617
<i>Maggie Makar, Fredrik Johansson, John Guttag, David Sontag</i>	
OPTIMAL TRANSPORT MAPPING VIA INPUT CONVEX NEURAL NETWORKS	6628
<i>Ashok Vardhan Makkuva, Amirhossein Taghvaei, Jason D. Lee, Sewoong Oh</i>	
PROVING THE LOTTERY TICKET HYPOTHESIS: PRUNING IS ALL YOU NEED	6638
<i>E. Malach, G. Yehudai, S. Shalev-schwartz, O. Shamir</i>	
FROM LOCAL SGD TO LOCAL FIXED-POINT METHODS FOR FEDERATED LEARNING	6648
<i>Grigory Malinovsky, Dmitry Kovalev, Elnur Gasanov, Laurent Condat, Peter Richtárik</i>	
ADAPTIVE GRADIENT DESCENT WITHOUT DESCENT	6658
<i>Yura Malitsky, Konstantin Mishchenko</i>	
EMERGENCE OF SEPARABLE MANIFOLDS IN DEEP LANGUAGE REPRESENTATIONS	6669
<i>Jonathan Mamou, Hang Le, Miguel A Del Rio, Cory Stephenson, Hanlin Tang, Yoon Kim, SueYeon Chung</i>	
ADAPTIVE ADVERSARIAL MULTI-TASK REPRESENTATION LEARNING	6680
<i>Yuren Mao, Weiwei Liu, Xuemin Lin</i>	
ON LEARNING SETS OF SYMMETRIC ELEMENTS	6690
<i>H. Maron, O. Litany, G. Chechik, E. Fetaya</i>	
STOCHASTICALLY DOMINANT DISTRIBUTIONAL REINFORCEMENT LEARNING	6701
<i>John D. Martin, Michal Lyskawinski, Xiaohu Li, Brendan Englot</i>	
MINIMAX PARETO FAIRNESS: A MULTI OBJECTIVE PERSPECTIVE	6711
<i>N. Martinez, M. Bertran, G. Sapiro</i>	
PREDICTIVE MULTIPLICITY IN CLASSIFICATION	6721
<i>Charles T. Marx, Flavio du Pin Calmon, Berk Ustun</i>	
ADDING SEEMINGLY UNINFORMATIVE LABELS HELPS IN LOW DATA REGIMES	6731
<i>Christos Matusoukas, Albert Bou I Hernandez, Yue Liu, Karin Dembrower, Gisele Miranda, Emir Konuk, Johan Fredin Haslum, Athanasios Zouzos, Peter Lindholm, Fredrik Strand, Kevin Smith</i>	
FAST AND CONSISTENT LEARNING OF HIDDEN MARKOV MODELS BY INCORPORATING NON-CONSECUTIVE CORRELATIONS	6741
<i>Robert Mattila, Cristian R. Rojas, Eric Moulines, Vikram Krishnamurthy, Bo Wahlberg</i>	
ON APPROXIMATE THOMPSON SAMPLING WITH LANGEVIN ALGORITHMS	6753
<i>Eric Mazumdar, Aldo Pacchiano, Yi-An Ma, Peter L. Bartlett, Michael I. Jordan</i>	
NEURAL DATALOG THROUGH TIME: INFORMED TEMPORAL MODELING VIA LOGICAL SPECIFICATION	6764
<i>H. Mei, G. Qin, M. Xu, J. Eisner</i>	
ON THE GLOBAL CONVERGENCE RATES OF SOFTMAX POLICY GRADIENT METHODS	6776
<i>Jincheng Mei, Chenjun Xiao, Csaba Szepesvári, Dale Schuurmans</i>	
SCALABLE IDENTIFICATION OF PARTIALLY OBSERVED SYSTEMS WITH CERTAINTY-EQUIVALENT EM	6786
<i>Kunal Menda, Jean de Beedelièvre, Jayesh K. Gupta, Ilan Kroo, Mykel J. Kochenderfer, Zachary Manchester</i>	
RANDOMIZED BLOCK-DIAGONAL PRECONDITIONING FOR PARALLEL LEARNING	6797
<i>Celestine Mender-Dünner, Aurelien Lucchi</i>	
TRAINING BINARY NEURAL NETWORKS USING THE BAYESIAN LEARNING RULE	6808
<i>Xiangming Meng, Roman Bachmann, Mohammad Emtiyaz Khan</i>	
CONTROL FREQUENCY ADAPTATION VIA ACTION PERSISTENCE IN BATCH REINFORCEMENT LEARNING	6818
<i>Alberto Maria Metelli, Flavio Mazzolini, Lorenzo Bisi, Luca Sabbioni, Marcello Restelli</i>	
THE ROLE OF REGULARIZATION IN CLASSIFICATION OF HIGH-DIMENSIONAL NOISY GAUSSIAN MIXTURE	6830
<i>Francesca Mignacco, Florent Krzakala, Yue M. Lu, Pierfrancesco Urbani, Lenka Zdeborová</i>	
PROJECTIVE PREFERENTIAL BAYESIAN OPTIMIZATION	6840
<i>Petrus Mikkola, Milica Todorovic, Jari Järvi, Patrick Rinke, Samuel Kaski</i>	

VIDEOONENET: BIDIRECTIONAL CONVOLUTIONAL RECURRENT ONENET WITH TRAINABLE DATA STEPS FOR VIDEO PROCESSING	6849
<i>Zoltán Á. Milacski, Barnabás Póczos, András Lorincz</i>	
THE EFFECT OF NATURAL DISTRIBUTION SHIFT ON QUESTION ANSWERING MODELS	6861
<i>John Miller, Karl Krauth, Benjamin Recht, Ludwig Schmidt</i>	
STRATEGIC CLASSIFICATION IS CAUSAL MODELING IN DISGUISE	6873
<i>John Miller, Smiha Milli, Moritz Hardt</i>	
AUTOMATIC SHORTCUT REMOVAL FOR SELF-SUPERVISED REPRESENTATION LEARNING	6883
<i>Mathias Minderer, Olivier Bachem, Neil Houlsby, Michael Tschannen</i>	
LEARNING REASONING STRATEGIES IN END-TO-END DIFFERENTIABLE PROVING	6894
<i>Pasquale Minervini, Sebastian Riedel, Pontus Stenetorp, Edward Grefenstette, Tim Rocktäschel</i>	
CORESETS FOR DATA-EFFICIENT TRAINING OF MACHINE LEARNING MODELS	6906
<i>B. Mirzasoleiman, J. Bilmes, J. Leskovec</i>	
KINEMATIC STATE ABSTRACTION AND PROVABLY EFFICIENT RICH-OBSERVATION REINFORCEMENT LEARNING	6917
<i>D. Misra, M. Henaff, A. Krishnamurthy, J. Langford</i>	
LEARNING TO COMBINE TOP-DOWN AND BOTTOM-UP SIGNALS IN RECURRENT NEURAL NETWORKS WITH ATTENTION OVER MODULES	6928
<i>Sarthak Mittal, Alex Lamb, Anirudh Goyal, Vikram Voleti, Murray Shanahan, Guillaume Lajoie, Michael Mozer, Yoshua Bengio</i>	
OPTIMIZING LONG-TERM SOCIAL WELFARE IN RECOMMENDER SYSTEMS: A CONSTRAINED MATCHING APPROACH	6943
<i>M. Mladenov, E. Creager, O. Ben-Porat, K. Swersky, R. Zemel, C. Boutilier</i>	
TRANSFORMATION OF RELU-BASED RECURRENT NEURAL NETWORKS FROM DISCRETE-TIME TO CONTINUOUS-TIME	6955
<i>Zahra Monfared, Daniel Durstewitz</i>	
EFFICIENTLY LEARNING ADVERSARIAL ROBUST HALFSPACES WITH NOISE	6966
<i>Omar Montasser, Surbhi Goel, Ilias Diakonikolas, Nathan Srebro</i>	
AN END-TO-END APPROACH FOR THE VERIFICATION PROBLEM: LEARNING THE RIGHT DISTANCE	6978
<i>J. Monteiro, I. Albuquerque, J. Alam, R. Hjelm, T. Falk</i>	
CONFIDENCE-AWARE LEARNING FOR DEEP NEURAL NETWORKS	6990
<i>Jooyoung Moon, Jihyo Kim, Younghak Shin, Sangheum Hwang</i>	

PART 10

TOPOLOGICAL AUTOENCODERS	7001
<i>M. Moor, M. Horn, B. Rieck, K. Borgwardt</i>	
EXPLAINABLE K-MEANS AND K-MEDIANS CLUSTERING	7012
<i>Sanjoy Dasgupta, Nave Frost, Michal Moshkovitz, Cyrus Rashtchian</i>	
FAIR LEARNING WITH PRIVATE DEMOGRAPHIC DATA	7023
<i>Hussein Mozannar, Mesrob I. Ohannessian, Nathan Srebro</i>	
CONSISTENT ESTIMATORS FOR LEARNING TO DEFER TO AN EXPERT	7033
<i>Hussein Mozannar, David Sontag</i>	
CONTINUOUS-TIME LOWER BOUNDS FOR GRADIENT-BASED ALGORITHMS	7045
<i>Michael Muehlebach, Michael I. Jordan</i>	
TWO SIMPLE WAYS TO LEARN INDIVIDUAL FAIRNESS METRICS FROM DATA	7054
<i>Debarghya Mukherjee, Mikhail Yurochkin, Moulinath Banerjee, Yuekai Sun</i>	
UNIQUE PROPERTIES OF FLAT MINIMA IN DEEP NETWORKS	7065
<i>Rotem Mulayoff, Tomer Michaeli</i>	
FAST COMPUTATION OF NASH EQUILIBRIA IN IMPERFECT INFORMATION GAMES	7076
<i>Remi Munos, Julien Perolat, Jean-Baptiste Lespiau, Mark Rowland, Bart De Vylder, Marc Lanctot, Finbarr Timbers, Daniel Hennes, Shayegan Omidshaftei, Audrunas Gruslys, Mohammad Gheshlaghi Azar, Edward Lockhart, Karl Tuyls</i>	
MISSING DATA IMPUTATION USING OPTIMAL TRANSPORT	7087
<i>Boris Muzellec, Julie Josse, Claire Boyer, Marco Cuturi</i>	
SEMI-PARAMETRIC NONLINEAR BIPARTITE GRAPH REPRESENTATION LEARNING WITH PROVABLE GUARANTEES	7098
<i>S. Na, Y. Luo, Z. Yang, Z. Wang, M. Kolar</i>	
FULL LAW IDENTIFICATION IN GRAPHICAL MODELS OF MISSING DATA: COMPLETENESS RESULTS	7110
<i>Razieh Nabi, Rohit Bhattacharya, Ilya Shpitser</i>	
VOICE SEPARATION WITH AN UNKNOWN NUMBER OF MULTIPLE SPEAKERS	7121
<i>Eliya Nachmani, Yossi Adi, Lior Wolf</i>	
RELIABLE FIDELITY AND DIVERSITY METRICS FOR GENERATIVE MODELS	7133
<i>Muhammad Ferjad Naeem, Seong Joon Oh, Youngjung Uh, Yunjey Choi, Jaejun Yoo</i>	
FROM CHAOS TO ORDER: SYMMETRY AND CONSERVATION LAWS IN GAME DYNAMICS	7143
<i>Sai Ganesh Nagarajan, David Balduzzi, Georgios Piliouras</i>	
UP OR DOWN? ADAPTIVE ROUNDING FOR POST-TRAINING QUANTIZATION	7154
<i>Markus Nagel, Rana Ali Amjad, Mart van Baalen, Christos Louizos, Tijmen Blankevoort</i>	
GOAL-AWARE PREDICTION: LEARNING TO MODEL WHAT MATTERS	7164
<i>Suraj Nair, Silvio Savarese, Chelsea Finn</i>	

POLYGEN: AN AUTOREGRESSIVE GENERATIVE MODEL OF 3D MESHES	7177
<i>C. Nash, Y. Ganin, S. Eslami, P. Battaglia</i>	
BAYESIAN SPARSIFICATION OF DEEP C-VALUED NETWORKS	7187
<i>Ivan Nazarov, Evgeny Burnaev</i>	
ORACLE EFFICIENT PRIVATE NON-CONVEX OPTIMIZATION	7200
<i>Seth Neel, Aaron Roth, Giuseppe Vietri, Zhiwei Steven Wu</i>	
STOCHASTIC FRANK-WOLFE FOR CONSTRAINED FINITE-SUM MINIMIZATION	7209
<i>Geoffrey Néglar, Gideon Dresdner, Alicia Yi-Ting Tsai, Laurent El Ghaoui, Francesco Locatello, Robert M. Freund, Fabian Pedregosa</i>	
IN DEFENSE OF UNIFORM CONVERGENCE: GENERALIZATION VIA DERANDOMIZATION WITH AN APPLICATION TO INTERPOLATING PREDICTORS	7219
<i>J. Negrea, G. Dziugaite, D. Roy</i>	
INVOLUTIVE MCMC: A UNIFYING FRAMEWORK	7229
<i>K. Neklyudov, M. Welling, E. Egorov, D. Vetrov</i>	
AGGREGATION OF MULTIPLE KNOCKOFFS	7239
<i>T. Nguyen, J. Chevalier, B. Thirion, S. Arlot</i>	
LEEP: A NEW MEASURE TO EVALUATE TRANSFERABILITY OF LEARNED REPRESENTATIONS	7250
<i>Cuong V. Nguyen, Tal Hassner, Matthias Seeger, Cedric Archambeau</i>	
GRAPH HOMOMORPHISM CONVOLUTION	7262
<i>Hoang NT, Takanori Maehara</i>	
KNOWING THE WHAT BUT NOT THE WHERE IN BAYESIAN OPTIMIZATION	7273
<i>Vu Nguyen, Michael A Osborne</i>	
ROBUST BAYESIAN CLASSIFICATION USING AN OPTIMISTIC SCORE RATIO	7283
<i>Viet Anh Nguyen, Nian Si, Jose Blanchet</i>	
STREAMING K-SUBMODULAR MAXIMIZATION UNDER NOISE SUBJECT TO SIZE CONSTRAINT	7294
<i>L. Nguyen, M. Thai</i>	
LP-SPARSEMAP: DIFFERENTIABLE RELAXED OPTIMIZATION FOR SPARSE STRUCTURED PREDICTION	7304
<i>Vlad Niculae, André F. T. Martins</i>	
SEMI-SUPERVISED STYLEGAN FOR DISENTANGLEMENT LEARNING	7316
<i>Weili Nie, Tero Karras, Animesh Garg, Shoubhik Debnath, Anjul Patney, Ankit B. Patel, Anima Anandkumar</i>	
SUPERVISED LEARNING: NO LOSS NO CRY	7326
<i>Richard Nock, Aditya Krishna Menon</i>	
CONSISTENT STRUCTURED PREDICTION WITH MAX-MIN MARGIN MARKOV NETWORKS	7337
<i>A. Nowak-Vila, F. Bach, A. Rudi</i>	
T-BASIS: A COMPACT REPRESENTATION FOR NEURAL NETWORKS	7348
<i>Anton Obukhov, Maxim Rakhuba, Stamatios Georgoulis, Menelaos Kanakis, Dengxin Dai, Luc Van Gool</i>	
ELIMINATING THE INVARIANCE ON THE LOSS LANDSCAPE OF LINEAR AUTOENCODERS	7361
<i>Reza Ofiadeh, Jiayi Shen, Zhangyang Wang, Dylan Shell</i>	
ON THE (IN)TRACTABILITY OF COMPUTING NORMALIZING CONSTANTS FOR THE PRODUCT OF DETERMINANTAL POINT PROCESSES	7370
<i>N. Ohsaka, T. Matsuoka</i>	
CAN INCREASING INPUT DIMENSIONALITY IMPROVE DEEP REINFORCEMENT LEARNING?	7380
<i>Kei Ota, Tomoaki Oiki, Devsh K. Jha, Toshisada Mariyama, Daniel Nikovski</i>	
INTERFEROMETRIC GRAPH TRANSFORM: A DEEP UNSUPERVISED GRAPH REPRESENTATION	7390
<i>Edouard Oyallon</i>	
LEARNING TO SCORE BEHAVIORS FOR GUIDED POLICY OPTIMIZATION	7401
<i>Aldo Pacchiano, Jack Parker-Holder, Yunhao Tang, Anna Choromanska, Krzysztof Choromanski, Michael I. Jordan</i>	
NEURAL CLUSTERING PROCESSES	7411
<i>A. Pakman, Y. Wang, C. Mitelut, J. Lee, L. Paninski</i>	
RECOVERY OF SPARSE SIGNALS FROM A MIXTURE OF LINEAR SAMPLES	7422
<i>Arya Mazumdar, Soumyabrata Pal</i>	
ADVERSARIAL MUTUAL INFORMATION FOR TEXT GENERATION	7432
<i>B. Pan, Y. Yang, K. Liang, B. Kailkhura, Z. Jin, X. Hua, D. Cai, B. Li</i>	
STABILIZING TRANSFORMERS FOR REINFORCEMENT LEARNING	7443
<i>E. Parisotto, H. Song, J. Rae, R. Pascanu, C. Gulcehre, S. Jayakumar, M. Jaderberg, R. Kaufman, A. Clark, S. Noury, M. Botvinick, N. Heess, R. Hadsell</i>	
MULTIRESOLUTION TENSOR LEARNING FOR EFFICIENT AND INTERPRETABLE SPATIAL ANALYSIS	7455
<i>J. Park, K. Carr, S. Zheng, Y. Yue, R. Yu</i>	
META VARIANCE TRANSFER: LEARNING TO AUGMENT FROM THE OTHERS	7466
<i>Seong-Jin Park, Seungju Han, Jiwon Baek, Insoo Kim, Juhwan Song, Hae Beom Lee, Jae-Joon Han, Sung Ju Hwang</i>	
STRUCTURED POLICY ITERATION FOR LINEAR QUADRATIC REGULATOR	7477
<i>Youngsuk Park, Ryan A. Rossi, Zheng Wen, Gang Wu, Handong Zhao</i>	
REGULARIZED OPTIMAL TRANSPORT IS GROUND COST ADVERSARIAL	7488
<i>F. Paty, M. Cuturi</i>	
REDUCING SAMPLING ERROR IN BATCH TEMPORAL DIFFERENCE LEARNING	7499
<i>Brahma S. Pavse, Ishan Durugkar, Josiah P. Hanna, Peter Stone</i>	
AVERAGE-CASE ACCELERATION THROUGH SPECTRAL DENSITY ESTIMATION	7509
<i>F. Pedregosa, D. Scieur</i>	

EINSUM NETWORKS: FAST AND SCALABLE LEARNING OF TRACTABLE PROBABILISTIC CIRCUITS	7519
<i>Robert Peharz, Steven Lang, Antonio Vergari, Karl Stelzner, Alejandro Molina, Martin Trapp, Guy Van den Broeck, Kristian Kersting, Zoubin Ghahramani</i>	
LEARNING SELECTION STRATEGIES IN BUCHBERGER’S ALGORITHM	7531
<i>Dylan Peifer, Michael Stillman, Daniel Halpern-Leistner</i>	
NON-AUTOREGRESSIVE NEURAL TEXT-TO-SPEECH	7542
<i>Kainan Peng, Wei Ping, Zhao Song, Kexin Zhao</i>	
PERFORMATIVE PREDICTION	7555
<i>Juan C. Perdomo, Tijana Zrnic, Celestine Mender-Dünner, Moritz Hardt</i>	
CONSTRUCTIVE UNIVERSAL HIGH-DIMENSIONAL DISTRIBUTION GENERATION THROUGH DEEP RELU NETWORKS	7566
<i>Dmytro Perekrstenko, Stephan Müller, Helmut Bölcskei</i>	
BUDGETED ONLINE INFLUENCE MAXIMIZATION	7576
<i>P. Perrault, J. Healey, Z. Wen, M. Valko</i>	
LOW BIAS LOW VARIANCE GRADIENT ESTIMATES FOR BOOLEAN STOCHASTIC NETWORKS	7588
<i>Adeel Pervez, Taco Cohen, Efstathios Gavves</i>	
ON CONVERGENCE-DIAGNOSTIC BASED STEP SIZES FOR STOCHASTIC GRADIENT DESCENT	7597
<i>Scott Pesme, Aymeric Dieuleveut, Nicolas Flammarion</i>	
SAMPLE FACTORY: EGOCENTRIC 3D CONTROL FROM PIXELS AT 100000 FPS WITH ASYNCHRONOUS REINFORCEMENT LEARNING	7608
<i>Aleksei Petrenko, Zhehui Huang, Tushar Kumar, Gaurav Sukhatme, Vladlen Koltun</i>	
IPBOOST – NON-CONVEX BOOSTING VIA INTEGER PROGRAMMING	7619
<i>Marc Pfetsch, Sebastian Pokutta</i>	
ON UNBALANCED OPTIMAL TRANSPORT: AN ANALYSIS OF SINKHORN ALGORITHM	7629
<i>Khiem Pham, Khang Le, Nhat Ho, Tung Pham, Hung Bui</i>	
SCALABLE DIFFERENTIAL PRIVACY WITH CERTIFIED ROBUSTNESS IN ADVERSARIAL LEARNING	7639
<i>NhatHai Phan, My T. Thai, Han Hu, Ruoming Jin, Tong Sun, Dejing Dou</i>	
NEURAL NETWORKS ARE CONVEX REGULARIZERS: EXACT POLYNOMIAL-TIME CONVEX OPTIMIZATION FORMULATIONS FOR TWO-LAYER NETWORKS	7651
<i>Mert Pilanci, Tolga Ergen</i>	
WAVEFLOW: A COMPACT FLOW-BASED MODEL FOR RAW AUDIO	7662
<i>Wei Ping, Kainan Peng, Kexin Zhao, Zhao Song</i>	
RANDOMIZATION MATTERS HOW TO DEFEND AGAINST STRONG ADVERSARIAL ATTACKS	7673
<i>Rafael Pinot, Raphael Ettetdgui, Geovani Rizk, Yann Chevaleyre, Jamal Atif</i>	
EFFICIENT DOMAIN GENERALIZATION VIA COMMON-SPECIFIC LOW-RANK DECOMPOSITION	7684
<i>Vihari Piratla, Praneeth Netrapalli, Sunita Sarawagi</i>	
DISSECTING NON-VACUOUS GENERALIZATION BOUNDS BASED ON THE MEAN-FIELD APPROXIMATION	7695
<i>Konstantinos Pitas</i>	
MAXIMUM ENTROPY GAIN EXPLORATION FOR LONG HORIZON MULTI-GOAL REINFORCEMENT LEARNING	7706
<i>S. Pitis, H. Chan, S. Zhao, B. Stadie, J. Ba</i>	
EXPLAINING GROUPS OF POINTS IN LOW-DIMENSIONAL REPRESENTATIONS	7718
<i>G. Plumb, J. Terhorst, S. Sankaramana, A. Talwalkar</i>	
ON THE UNREASONABLE EFFECTIVENESS OF THE GREEDY ALGORITHM: GREEDY ADAPTS TO SHARPNESS	7728
<i>Sebastian Pokutta, Mohit Singh, Alfredo Torrico</i>	
SKEW-FIT: STATE-COVERING SELF-SUPERVISED REINFORCEMENT LEARNING	7739
<i>V. Pong, M. Dalal, S. Lin, A. Nair, S. Bahl, S. Levine</i>	
SOFTSORT: A CONTINUOUS RELAXATION FOR THE ARGSORT OPERATOR	7749
<i>Sebastian Prillo, Julian Martin Eisenschlos</i>	
GRAPH-BASED NEAREST NEIGHBOR SEARCH: FROM PRACTICE TO THEORY	7759
<i>Liudmila Prokhorenkova, Aleksandr Shekhovtsov</i>	
ADVERSARIAL RISK VIA OPTIMAL TRANSPORT AND OPTIMAL COUPLINGS	7770
<i>Muni Sreenivas Pydi, Varun Jog</i>	

PART 11

DEEP ISOMETRIC LEARNING FOR VISUAL RECOGNITION	7780
<i>Haozhi Qi, Chong You, Xiaolong Wang, Yi Ma, Jitendra Malik</i>	
UNSUPERVISED SPEECH DECOMPOSITION VIA TRIPLE INFORMATION BOTTLENECK	7792
<i>Kaizhi Qian, Yang Zhang, Shiyu Chang, David Cox, Mark Hasegawa-Johnson</i>	
SCALABLE DIFFERENTIABLE PHYSICS FOR LEARNING AND CONTROL	7803
<i>Yi-Ling Qiao, Junbang Liang, Vladlen Koltun, Ming C. Lin</i>	
ROBUST ONE-BIT RECOVERY VIA RELU GENERATIVE NETWORKS: NEAR-OPTIMAL STATISTICAL RATE AND GLOBAL LANDSCAPE ANALYSIS	7813
<i>Shuang Qiu, Xiaohan Wei, Zhuoran Yang</i>	
FEW-SHOT RELATION EXTRACTION VIA BAYESIAN META-LEARNING ON RELATION GRAPHS	7823
<i>Meng Qu, Tianyu Gao, Louis-Pascal A. C. Xhonneux, Jian Tang</i>	

DEEPCODA: PERSONALIZED INTERPRETABILITY FOR COMPOSITIONAL HEALTH DATA	7833
<i>Thomas P. Quinn, Dang Nguyen, Santu Rana, Sunil Gupta, Svetha Venkatesh</i>	
FAST AND PRIVATE SUBMODULAR AND K-SUBMODULAR FUNCTIONS MAXIMIZATION WITH MATROID CONSTRAINTS	7843
<i>Akbar Rafiey, Yuichi Yoshida</i>	
TRANSPARENCY PROMOTION WITH MODEL-AGNOSTIC LINEAR COMPETITORS	7854
<i>Hassan Rafique, Tong Wang, Qihang Lin, Arshia Sighani</i>	
UNDERSTANDING AND MITIGATING THE TRADEOFF BETWEEN ROBUSTNESS AND ACCURACY	7865
<i>A. Raghunathan, S. Xie, F. Yang, J. Duchi, P. Liang</i>	
FAST ADAPTATION TO NEW ENVIRONMENTS VIA POLICY-DYNAMICS VALUE FUNCTIONS	7876
<i>Roberta Raileanu, Max Goldstein, Arthur Szlam, Rob Fergus</i>	
IMPROVING ROBUSTNESS OF DEEP-LEARNING-BASED IMAGE RECONSTRUCTION	7888
<i>Ankit Raj, Yoram Bresler, Bo Li</i>	
MULTI-PRECISION POLICY ENFORCED TRAINING (MUPPET): A PRECISION-SWITCHING STRATEGY FOR QUANTISED FIXED-POINT TRAINING OF CNNs	7899
<i>A. Rajagopal, D. Vink, S. Venieris, C. Bouganis</i>	
A GAME THEORETIC FRAMEWORK FOR MODEL BASED REINFORCEMENT LEARNING	7909
<i>Aravind Rajeswaran, Igor Mordatch, Vikash Kumar</i>	
CLOSING THE CONVERGENCE GAP OF SGD WITHOUT REPLACEMENT	7920
<i>Shashank Rajput, Anant Gupta, Dimitris Papailiopoulos</i>	
POLICY TEACHING VIA ENVIRONMENT POISONING: TRAINING-TIME ADVERSARIAL ATTACKS AGAINST REINFORCEMENT LEARNING	7930
<i>Amin Rakhsha, Goran Radanovic, Rati Devidze, Xiaojin Zhu, Adish Singla</i>	
IMPLICIT GENERATIVE MODELING FOR EFFICIENT EXPLORATION	7941
<i>Neale Ratzlaff, Qinxun Bai, Li Fuxin, Wei Xu</i>	
UNIVERSAL EQUIVARIANT MULTILAYER PERCEPTRONS	7952
<i>Stamak Ravanbakhsh</i>	
AUTOML-ZERO: EVOLVING MACHINE LEARNING ALGORITHMS FROM SCRATCH	7963
<i>Esteban Real, Chen Liang, David R. So, Quoc V. Le</i>	
LEARNING HUMAN OBJECTIVES BY EVALUATING HYPOTHETICAL BEHAVIOR	7976
<i>Siddharth Reddy, Anca D. Dragan, Sergey Levine, Shane Legg, Jan Leike</i>	
OPTIMISTIC BOUNDS FOR MULTI-OUTPUT PREDICTION	7986
<i>Henry W.J. Reeve, Ata Kabán</i>	
ACTIVE LEARNING ON ATTRIBUTED GRAPHS VIA GRAPH COGNIZANT LOGISTIC REGRESSION AND PREEMPTIVE QUERY GENERATION	7997
<i>Florence Regol, Soumyasundar Pal, Yingxue Zhang, Mark Coates</i>	
THE SAMPLE COMPLEXITY OF BEST-K ITEMS SELECTION FROM PAIRWISE COMPARISONS	8007
<i>Wenbo Ren, Jia Liu, Ness B. Shroff</i>	
NETGAN WITHOUT GAN: FROM RANDOM WALKS TO LOW-RANK APPROXIMATIONS	8029
<i>L. Rendsburg, H. Heidrich, U. Luxburg</i>	
NORMALIZING FLOWS ON TORI AND SPHERES	8039
<i>D. Rezende, G. Papamakarios, S. Racaniere, M. Albergo, G. Kanwar, P. Shanahan, K. Cranmer</i>	
OVERFITTING IN ADVERSARIALLY ROBUST DEEP LEARNING	8049
<i>Leslie Rice, Eric Wong, J. Zico Kolter</i>	
OVERFITTING IN ADVERSARIALLY ROBUST DEEP LEARNING	8061
<i>Leslie Rice, Eric Wong, J. Zico Kolter</i>	
DECENTRALISED LEARNING WITH DISTRIBUTED GRADIENT DESCENT AND RANDOM FEATURES	8075
<i>Dominic Richards, Patrick Rebeschini, Lorenzo Rosasco</i>	
INTERPRETATIONS ARE USEFUL: PENALIZING EXPLANATIONS TO ALIGN NEURAL NETWORKS WITH PRIOR KNOWLEDGE	8086
<i>Laura Rieger, Chandan Singh, W. James Murdoch, Bin Yu</i>	
STRENGTH FROM WEAKNESS: FAST LEARNING USING WEAK SUPERVISION	8097
<i>J. Robinson, S. Jegelka, S. Sra</i>	
ON SEMI-PARAMETRIC INFERENCE FOR BART	8107
<i>Veronika Rocková</i>	
FR-TRAIN: A MUTUAL INFORMATION-BASED APPROACH TO FAIR AND ROBUST TRAINING	8117
<i>Yuji Roh, Kangwook Lee, Steven Euijong Whang, Changho Suh</i>	
BALANCING COMPETING OBJECTIVES WITH NOISY DATA: SCORE-BASED CLASSIFIERS FOR WELFARE-AWARE MACHINE LEARNING	8128
<i>Esther Rolf, Max Simchowitz, Sarah Dean, Lydia T. Liu, Daniel Björckegren, Moritz Hardt, Joshua Blumensack</i>	
DOUBLE-LOOP UNADJUSTED LANGEVIN ALGORITHM	8139
<i>Paul Rolland, Armin Eftekhar, Ali Kavis, Volkan Cevher</i>	
REVERSE-ENGINEERING DEEP RELU NETWORKS	8148
<i>David Rolnick, Konrad P. Körding</i>	
ATTENTIVE GROUP EQUIVARIANT CONVOLUTIONAL NETWORKS	8158
<i>David W. Romero, Erik J. Bekkers, Jakub M. Tomczak, Mark Hoogendoorn</i>	
FINITE-TIME CONVERGENCE IN CONTINUOUS-TIME OPTIMIZATION	8170
<i>Orlando Romero, Mouhacine Benosman</i>	
NEAR-OPTIMAL REGRET BOUNDS FOR STOCHASTIC SHORTEST PATH	8180
<i>A. Cohen, H. Kaplan, Y. Mansour, A. Rosenberg</i>	

PREDICTING CHOICE WITH SET-DEPENDENT AGGREGATION	8190
<i>N. Rosenfeld, K. Oshiba, Y. Singer</i>	
CERTIFIED ROBUSTNESS TO LABEL-FLIPPING ATTACKS VIA RANDOMIZED SMOOTHING	8200
<i>Elan Rosenfeld, Ezra Winston, Pradeep Ravikumar, J. Zico Kolter</i>	
REVISITING TRAINING STRATEGIES AND GENERALIZATION PERFORMANCE IN DEEP METRIC LEARNING	8212
<i>K. Roth, T. Mülbich, S. Sinha, P. Gupta, B. Ommer, J. Cohen</i>	
FETCHSGD: COMMUNICATION-EFFICIENT FEDERATED LEARNING WITH SKETCHING	8223
<i>D. Rothchild, A. Panda, E. Ullah, N. Iykin, I. Stoica, V. Braverman, J. Gonzalez, R. Arora</i>	
SIMPLE AND SHARP ANALYSIS OF K-MEANS 	8236
<i>Václav Rozhon</i>	
BAYESIAN OPTIMISATION OVER MULTIPLE CONTINUOUS AND CATEGORICAL INPUTS	8246
<i>Binxin Ru, Ahsan S. Alvi, Vu Nguyen, Michael A. Osborne, Stephen J Roberts</i>	
INTER-DOMAIN DEEP GAUSSIAN PROCESSES	8256
<i>T. Rudner, D. Sejdinovic, Y. Gal</i>	
BIO-INSPIRED HASHING FOR UNSUPERVISED SIMILARITY SEARCH	8265
<i>C. Ryali, J. Hopfield, L. Grinberg, D. Krotov</i>	
ADVERSARIAL ATTACKS ON COPYRIGHT DETECTION SYSTEMS	8277
<i>Parsa Saadatpanah, Ali Shafahi, Tom Goldstein</i>	
BOUNDING THE FAIRNESS AND ACCURACY OF CLASSIFIERS FROM POPULATION STATISTICS	8286
<i>S. Sabato, E. Yom-Tov</i>	
RADIOACTIVE DATA: TRACING THROUGH TRAINING	8296
<i>A. Sablayrolles, M. Douze, C. Schmid, H. Jegou</i>	
CAUSAL STRUCTURE DISCOVERY FROM DISTRIBUTIONS ARISING FROM MIXTURES OF DAGS	8306
<i>B. Saeed, S. Panigrahi, C. Uhler</i>	
AN INVESTIGATION OF WHY OVERPARAMETERIZATION EXACERBATES SPURIOUS CORRELATIONS	8316
<i>S. Sagawa, A. Raghunathan, P. Koh, P. Liang</i>	
IMPROVED SLEEPING BANDITS WITH STOCHASTIC ACTIONS SETS AND ADVERSARIAL REWARDS	8327
<i>A. Saha, P. Gaillard, M. Valko</i>	
FROM PAC TO INSTANCE-OPTIMAL SAMPLE COMPLEXITY IN THE PLACKETT-LUCE MODEL	8336
<i>A. Saha, A. Gopalan</i>	
MEASURING NON-EXPERT COMPREHENSION OF MACHINE LEARNING FAIRNESS METRICS	8346
<i>Debjani Saha, Candice Schumann, Duncan C. McElfresh, John P. Dickerson, Michelle L. Mazurek, Michael Carl Tschantz</i>	
FROM SETS TO MULTISSETS: PROVABLE VARIATIONAL INFERENCE FOR PROBABILISTIC INTEGER SUBMODULAR MODELS	8357
<i>Aytunc Sahin, Yatao Bian, Joachim M. Buhmann, Andreas Krause</i>	
COUNTERFACTUAL CROSS-VALIDATION: STABLE MODEL SELECTION PROCEDURE FOR CAUSAL INFERENCE MODELS	8367
<i>Yuta Saito, Shota Yasui</i>	
INFERRING DQN STRUCTURE FOR HIGH-DIMENSIONAL CONTINUOUS CONTROL	8377
<i>Andrey Sakryukin, Chedy Raïssi, Mohan S. Kankanhalli</i>	
THE PERFORMANCE ANALYSIS OF GENERALIZED MARGIN MAXIMIZER (GMM) ON SEPARABLE DATA	8386
<i>F. Salehi, E. Abbasi, B. Hassibi</i>	
STOCHASTIC COORDINATE MINIMIZATION WITH PROGRESSIVE PRECISION FOR STOCHASTIC CONVEX OPTIMIZATION	8396
<i>Sudeep Salgia, Qing Zhao, Sattar Vakili</i>	
A QUANTILE-BASED APPROACH FOR HYPERPARAMETER TRANSFER LEARNING	8407
<i>David Salinas, Huibin Shen, Valerio Perrone</i>	
SPECTRAL SUBSAMPLING MCMC FOR STATIONARY TIME SERIES	8418
<i>Robert Salomone, Matias Quiroz, Robert Kohn, Mattias Villani, Minh-Ngoc Tran</i>	
LEARNING TO SIMULATE COMPLEX PHYSICS WITH GRAPH NETWORKS	8428
<i>Alvaro Sanchez-Gonzalez, Jonathan Godwin, Tobias Pfaff, Rex Ying, Jure Leskovec, Peter W. Battaglia</i>	
EXPLICIT GRADIENT LEARNING FOR BLACK-BOX OPTIMIZATION	8438
<i>Elad Sarafian, Mor Sinay, Yoram Louzoun, Noa Agmon, Sarit Kraus</i>	
DETECTING OUT-OF-DISTRIBUTION EXAMPLES WITH GRAM MATRICES	8449
<i>C. Sastry, S. Oore</i>	
CONSTRAINED MARKOV DECISION PROCESSES VIA BACKWARD VALUE FUNCTIONS	8460
<i>Harsh Satija, Philip Amortila, Joelle Pineau</i>	
A SAMPLE COMPLEXITY SEPARATION BETWEEN NON-CONVEX AND CONVEX META-LEARNING	8470
<i>Nikunj Saunshi, Yi Zhang, Mikhail Khodak, Sanjeev Arora</i>	
HARMONIC DECOMPOSITIONS OF CONVOLUTIONAL NETWORKS	8480
<i>Meyer Scetbon, Zaid Harchaoui</i>	
IMPLICIT COMPETITIVE REGULARIZATION IN GANS	8491
<i>Florian Schäfer, Hongkai Zheng, Anima Anandkumar</i>	
OFF-POLICY ACTOR-CRITIC WITH SHARED EXPERIENCE REPLAY	8503
<i>Simon Schmitt, Matteo Hessel, Karen Simonyan</i>	
DISCRIMINATIVE ADVERSARIAL SEARCH FOR ABSTRACTIVE SUMMARIZATION	8513
<i>Thomas Scialom, Paul-Alexis Dray, Sylvain Lamprier, Benjamin Piwowarski, Jacopo Staiano</i>	

UNIVERSAL AVERAGE-CASE OPTIMALITY OF POLYAK MOMENTUM	8523
<i>Damien Scieur, Fabian Pedregosa</i>	
RANDOM MATRIX THEORY PROVES THAT DEEP LEARNING REPRESENTATIONS OF GAN-DATA BEHAVE AS GAUSSIAN MIXTURES	8531
<i>Mohamed El Amine Seddik, Cosme Louart, Mohamed Tamaazousti, Romain Couillet</i>	
PLANNING TO EXPLORE VIA SELF-SUPERVISED WORLD MODELS	8541
<i>Ramanan Sekar, Oleh Rybkin, Kostas Daniilidis, Pieter Abbeel, Danijar Hafner, Deepak Pathak</i>	

PART 12

AN EXPLICITLY RELATIONAL NEURAL NETWORK ARCHITECTURE	8551
<i>M. Shanahan, K. Nikiforou, A. Creswell, C. Kaplanis, D. Barrett, M. Garnele</i>	
OPTIMISTIC POLICY OPTIMIZATION WITH BANDIT FEEDBACK	8562
<i>Y. Efroni, L. Shani, A. Rosenberg, S. Mannor</i>	
NEURAL KERNELS WITHOUT TANGENTS	8572
<i>V. Shankar, A. Fang, W. Guo, S. Fridovich-Kell, L. Schmidt, J. Ragan-Kelley, B. Recht</i>	
LEARNING ROBOT SKILLS WITH TEMPORAL VARIATIONAL INFERENCE	8582
<i>Tanmay Shankar, Abhinav Gupta</i>	
EVALUATING MACHINE ACCURACY ON IMAGENET	8592
<i>V. Shankar, R. Roelofs, H. Mania, A. Fang, B. Recht, L. Schmidt</i>	
CHANNEL EQUILIBRIUM NETWORKS FOR LEARNING DEEP REPRESENTATION	8603
<i>Wenqi Shao, Shitao Tang, Xingang Pan, Ping Tan, Xiaogang Wang, Ping Luo</i>	
CONTROLVAE: CONTROLLABLE VARIATIONAL AUTOENCODER	8614
<i>H. Shao, S. Yao, D. Sun, A. Zhang, S. Liu, D. Liu, J. Wang, T. Abdelzaher</i>	
LOOKAHEAD-BOUNDED Q-LEARNING	8624
<i>Ibrahim El Shar, Daniel R. Jiang</i>	
CAUSAL STRATEGIC LINEAR REGRESSION	8635
<i>Yonadav Shavit, Benjamin L. Edelman, Brian Axelrod</i>	
ADAPTIVE SAMPLING FOR ESTIMATING PROBABILITY DISTRIBUTIONS	8646
<i>S. Shekhar, T. Javidi, M. Ghavamzadeh</i>	
PDO-ECONVS: PARTIAL DIFFERENTIAL OPERATOR BASED EQUIVARIANT CONVOLUTIONS	8656
<i>Zhengyang Shen, Lingshen He, Zhouchen Lin, Jinwen Ma</i>	
DEEP REINFORCEMENT LEARNING WITH ROBUST AND SMOOTH POLICY	8666
<i>Qianli Shen, Yan Li, Haoming Jiang, Zhaoran Wang, Tuo Zhao</i>	
EDUCATING TEXT AUTOENCODERS: LATENT REPRESENTATION GUIDANCE VIA DENOISING	8678
<i>T. Shen, J. Mueller, R. Barzilay, T. Jaakkola</i>	
LEARNING FOR DOSE ALLOCATION IN ADAPTIVE CLINICAL TRIALS WITH SAFETY CONSTRAINTS	8689
<i>Cong Shen, Zhiyang Wang, Sofia S. Villar, Mihaela van der Schaar</i>	
POWERNORM: RETHINKING BATCH NORMALIZATION IN TRANSFORMERS	8700
<i>Sheng Shen, Zhewei Yao, Amir Gholami, Michael W. Mahoney, Kurt Keutzer</i>	
EXTREME MULTI-LABEL CLASSIFICATION FROM AGGREGATED LABELS	8711
<i>Y. Shen, H. Yu, S. Sanghavi, I. Dhillon</i>	
ONE-SHOT DISTRIBUTED RIDGE REGRESSION IN HIGH DIMENSIONS	8722
<i>Edgar Dobriban, Yue Sheng</i>	
LANDSCAPE CONNECTIVITY AND DROPOUT STABILITY OF SGD SOLUTIONS FOR OVER-PARAMETERIZED NEURAL NETWORK	8732
<i>A. Shevchenko, M. Mondelli</i>	
INCREMENTAL SAMPLING WITHOUT REPLACEMENT FOR SEQUENCE MODELS	8744
<i>K. Shi, D. Bieber, C. Sutton</i>	
MESSAGE PASSING LEAST SQUARES FRAMEWORK AND ITS APPLICATION TO ROTATION SYNCHRONIZATION	8755
<i>Yunpeng Shi, Gilad Lerman</i>	
DOES THE MARKOV DECISION PROCESS FIT THE DATA: TESTING FOR THE MARKOV PROPERTY IN SEQUENTIAL DECISION MAKING	8766
<i>Chengchun Shi, Runzhe Wan, Rui Song, Wenbin Lu, Ling Leng</i>	
A GRAPH TO GRAPHS FRAMEWORK FOR RETROSYNTHESIS PREDICTION	8777
<i>Chence Shi, Minkai Xu, Hongyu Guo, Ming Zhang, Jian Tang</i>	
INFORMATIVE DROPOUT FOR ROBUST REPRESENTATION LEARNING: A SHAPE-BIAS PERSPECTIVE	8787
<i>B. Shi, D. Zhang, Q. Dai, Z. Zhu, Y. Mu, J. Wang</i>	
DISPERSED EXPONENTIAL FAMILY MIXTURE VAES FOR INTERPRETABLE TEXT GENERATION	8799
<i>Wenxian Shi, Hao Zhou, Ning Miao, Lei Li</i>	
ON CONDITIONAL VERSUS MARGINAL BIAS IN MULTI-ARMED BANDITS	8811
<i>Jaehyeok Shin, Alessandro Rinaldo, Aaditya Ramdas</i>	
PREDICTIVE CODING FOR LOCALLY-LINEAR CONTROL	8821
<i>Rui Shu, Tung Nguyen, Yinlam Chow, Tuan Pham, Khoat Than, Mohammad Ghavamzadeh, Stefano Ermon, Hung Bui</i>	
A MARKOV DECISION PROCESS MODEL FOR SOCIO-ECONOMIC SYSTEMS IMPACTED BY CLIMATE CHANGE	8831
<i>Salman Sadiq Shuvo, Yasin Yilmaz, Alan Bush, Mark Hafen</i>	

DISTRIBUTIONALLY ROBUST POLICY EVALUATION AND LEARNING IN OFFLINE CONTEXTUAL BANDITS	8843
<i>Nian Si, Fan Zhang, Zhengyuan Zhou, Jose Blanchet</i>	
PIECEWISE LINEAR REGRESSION VIA A DIFFERENCE OF CONVEX FUNCTIONS	8854
<i>A. Siahkamari, A. Gangrade, B. Kulis, V. Saligrama</i>	
LEARNING FAIR POLICIES IN MULTIOBJECTIVE (DEEP) REINFORCEMENT LEARNING WITH AVERAGE AND DISCOUNTED REWARDS	8864
<i>Umer Siddique, Paul Weng, Matthieu Zimmer</i>	
DEEP GAUSSIAN MARKOV RANDOM FIELDS	8875
<i>P. Siden, F. Lindsten</i>	
COLLABORATIVE MACHINE LEARNING WITH INCENTIVE-AWARE MODEL REWARDS	8886
<i>Rachael Hwee Ling Sim, Yehong Zhang, Mun Choon Chan, Bryan Kian Hsiang Low</i>	
A GENERATIVE MODEL FOR MOLECULAR DISTANCE GEOMETRY	8896
<i>Gregor N. C. Simm, José Miguel Hernández-Lobato</i>	
REINFORCEMENT LEARNING FOR MOLECULAR DESIGN GUIDED BY QUANTUM MECHANICS	8906
<i>Gregor N. C. Simm, Robert Pinsler, José Miguel Hernández-Lobato</i>	
FRACTIONAL UNDERDAMPED LANGEVIN DYNAMICS: RETARGETING SGD WITH MOMENTUM UNDER HEAVY-TAILED GRADIENT NOISE	8917
<i>Umut Simsekli, Lingjiong Zhu, Yee Whye Teh, Mert Gürbüzbalaban</i>	
SECOND-ORDER PROVABLE DEFENSES AGAINST ADVERSARIAL ATTACKS	8928
<i>Sahil Singla, Soheil Feizi</i>	
FORMULAZERO: DISTRIBUTIONALLY ROBUST ONLINE ADAPTATION VIA OFFLINE POPULATION SYNTHESIS	8939
<i>Aman Sinha</i>	
SMALL-GAN: SPEEDING UP GAN TRAINING USING CORE-SETS	8952
<i>Samarth Sinha, Han Zhang, Anirudh Goyal, Yoshua Bengio, Hugo Larochelle, Augustus Odena</i>	
INTERPRETABLE, MULTIDIMENSIONAL, MULTIMODAL ANOMALY DETECTION WITH NEGATIVE SAMPLING FOR DETECTION OF DEVICE FAILURE	8963
<i>John Sipple</i>	
STRUCTURED LINEAR CONTEXTUAL BANDITS: A SHARP AND GEOMETRIC SMOOTHED ANALYSIS	8973
<i>Vidyashankar Sivakumar, Zhiwei Steven Wu, Arindam Banerjee</i>	
OPTIMIZER BENCHMARKING NEEDS TO ACCOUNT FOR HYPERPARAMETER TUNING	8983
<i>P. Teja, F. Mai, T. Vogels, M. Jaggi, F. Fleuret</i>	
WHEN EXPLANATIONS LIE: WHY MANY MODIFIED BP ATTRIBUTIONS FAIL	8993
<i>L. Sixt, M. Granz, T. Landgraf</i>	
MULTICLASS NEURAL NETWORK MINIMIZATION VIA TROPICAL NEWTON POLYTOPE APPROXIMATION	9005
<i>Georgios Smyrnis, Petros Maragos</i>	
BRIDGING THE GAP BETWEEN F-GANS AND WASSERSTEIN GANS	9015
<i>Jiaming Song, Stefano Ermon</i>	
PROVABLY EFFICIENT MODEL-BASED POLICY ADAPTATION	9025
<i>Y. Song, A. Mavalankar, W. Sun, S. Gao</i>	
HYPERNETWORK APPROACH TO GENERATING POINT CLOUDS	9036
<i>Przemysław Spurek, Sebastian Winczowski, Jacek Tabor, Maciej Zamorski, Maciej Zieba, Tomasz Trzcinski</i>	
ROBUSTNESS TO SPURIOUS CORRELATIONS VIA HUMAN ANNOTATIONS	9046
<i>M. Srivastava, T. Hashimoto, P. Liang</i>	
WHICH TASKS SHOULD BE LEARNED TOGETHER IN MULTI-TASK LEARNING?	9057
<i>Trevor Standley, Amir Zamir, Dawn Chen, Leonidas Guibas, Jitendra Malik, Silvio Savarese</i>	
RESPONSIVE SAFETY IN REINFORCEMENT LEARNING BY PID LAGRANGIAN METHODS	9070
<i>Adam Stooke, Joshua Achiam, Pieter Abbeel</i>	
LEARNING DISCRETE STRUCTURED REPRESENTATIONS BY ADVERSARIAL MAXIMIZING MUTUAL INFORMATION	9081
<i>Karl Stratos, Sam Wiseman</i>	
CONFIDENCE-CALIBRATED ADVERSARIAL TRAINING: GENERALIZING TO UNSEEN ATTACKS	9092
<i>David Stutz, Matthias Hein, Bernt Schiele</i>	
DOUBLY ROBUST OFF-POLICY EVALUATION WITH SHRINKAGE	9104
<i>Yi Su, Maria Dimakopoulou, Akshay Krishnamurthy, Miroslav Dudik</i>	
TASK UNDERSTANDING FROM CONFUSING MULTI-TASK DATA	9114
<i>Xin Su, Yizhou Jiang, Shangqi Guo, Feng Chen</i>	
CONQR: MITIGATING DELUSIONAL BIAS IN DEEP Q-LEARNING	9124
<i>DiJia (Andy) Su, Jayden Ooi, Tyler Lu, Dale Schuurmans, Craig Boutilier</i>	
ADAPTIVE ESTIMATOR SELECTION FOR OFF-POLICY EVALUATION	9133
<i>Akshay Krishnamurthy</i>	
GENERATIVE TEACHING NETWORKS: ACCELERATING NEURAL ARCHITECTURE SEARCH BY LEARNING TO GENERATE SYNTHETIC TRAINING DATA	9143
<i>F. Such, A. Rawal, J. Lehman, K. Stanley, J. Clune</i>	
IMPROVING THE SAMPLE AND COMMUNICATION COMPLEXITY FOR DECENTRALIZED NON-CONVEX OPTIMIZATION: JOINT GRADIENT ESTIMATION AND TRACKING	9154
<i>Haoran Sun, Songtao Lu, Mingyi Hong</i>	

TEST-TIME TRAINING WITH SELF-SUPERVISION FOR GENERALIZATION UNDER DISTRIBUTION SHIFTS	9166
<i>Yu Sun, Xiaolong Wang, Zhuang Liu, John Miller, Alexei A. Efros, Moritz Hardt</i>	
AN EM APPROACH TO NON-AUTOREGRESSIVE CONDITIONAL SEQUENCE GENERATION	9186
<i>Zhiqing Sun, Yiming Yang</i>	
THE SHAPLEY TAYLOR INTERACTION INDEX	9196
<i>Kedar Dhamdhere, Ashish Agarwal, Mukund Sundararajan</i>	
THE MANY SHAPLEY VALUES FOR MODEL EXPLANATION	9210
<i>Mukund Sundararajan, Amir Najmi</i>	
MULTI-OBJECTIVE BAYESIAN OPTIMIZATION USING PARETO-FRONTIER ENTROPY	9221
<i>S. Suzuki, S. Takeno, T. Tamura, K. Shitara, M. Karasuyama</i>	
THE K-TIED NORMAL DISTRIBUTION: A COMPACT PARAMETERIZATION OF GAUSSIAN MEAN FIELD POSTERIORES IN BAYESIAN NEURAL NETWORKS	9231
<i>Jakub Swiatkowski, Kevin Roth, Bastiaan S. Veeling, Linh Tran, Joshua V. Dillon, Jasper Snoek, Stephan Mandt, Tim Salimans, Rodolphe Jenatton, Sebastian Nowozin</i>	
MULTI-AGENT ROUTING VALUE ITERATION NETWORK	9242
<i>Quinlan Sykora, Mengye Ren, Raquel Urtasun</i>	
DISTINGUISHING CAUSE FROM EFFECT USING QUANTILES: BIVARIATE QUANTILE CAUSAL DISCOVERY	9253
<i>Natasa Tagasovska, Valérie Chavez-Demoulin, Thibault Vatter</i>	
QUANTIZED DECENTRALIZED STOCHASTIC LEARNING OVER DIRECTED GRAPHS	9266
<i>Hossein Taheri, Aryan Mokhtari, Hamed Hassani, Ramtin Pedarsani</i>	
MULTI-FIDELITY BAYESIAN OPTIMIZATION WITH MAX-VALUE ENTROPY SEARCH AND ITS PARALLELIZATION	9276
<i>S. Takeno, H. Fukuoka, Y. Tsukada, T. Koyama, M. Shiga, I. Takeuchi, M. Karasuyama</i>	
FIEDLER REGULARIZATION: LEARNING NEURAL NETWORKS WITH GRAPH SPARSITY	9288
<i>Edric Tam, David Dunson</i>	
DROPNET: REDUCING NEURAL NETWORK COMPLEXITY VIA ITERATIVE PRUNING	9298
<i>John Tan Chong Min, Mehul Motani</i>	

PART 13

REINFORCEMENT LEARNING FOR INTEGER PROGRAMMING: LEARNING TO CUT	9309
<i>Y. Tang, S. Agrawal, Y. Faenza</i>	
THE BUCKLEY-OSTHUS MODEL AND THE BLOCK PREFERENTIAL ATTACHMENT MODEL: STATISTICAL ANALYSIS AND APPLICATION	9319
<i>X. Guo, F. Tang, W. Tang</i>	
CLINICIAN-IN-THE-LOOP DECISION MAKING: REINFORCEMENT LEARNING WITH NEAR-OPTIMAL SET-VALUED POLICIES	9329
<i>Shengpu Tang, Aditya Modi, Michael W. Sjoding, Jenna Wiens</i>	
TAYLOR EXPANSION POLICY OPTIMIZATION	9339
<i>Y. Tang, M. Valko, R. Munos</i>	
VARIATIONAL IMITATION LEARNING WITH DIVERSE-QUALITY DEMONSTRATIONS	9349
<i>Voot Tangkaratt, Bo Han, Mohammad Emtyaz Khan, Masashi Sugiyama</i>	
LEARNING DISCONNECTED MANIFOLDS: A NO GAN'S LAND	9360
<i>Ugo Tanielian, Thibaut Issenhuth, Elvis Dohmatob, Jérémie Mary</i>	
NO-REGRET EXPLORATION IN GOAL-ORIENTED REINFORCEMENT LEARNING	9370
<i>Jean Tarbouriech, Evrard Garcelon, Michal Valko, Matteo Pirota, Alessandro Lazaric</i>	
SPARSE SINKHORN ATTENTION	9380
<i>Yi Tay, Dara Bahri, Liu Yang, Donald Metzler, Da-Cheng Juan</i>	
INDUCTIVE RELATION PREDICTION BY SUBGRAPH REASONING	9390
<i>Komal K. Teru, Etienne G. Denis, William L. Hamilton</i>	
FEW-SHOT DOMAIN ADAPTATION BY CAUSAL MECHANISM TRANSFER	9400
<i>Takeshi Teshima, Issei Sato, Masashi Sugiyama</i>	
STUDENT SPECIALIZATION IN DEEP RECTIFIED NETWORKS WITH FINITE WIDTH AND INPUT DIMENSION	9412
<i>Yuandong Tian</i>	
SEQUENTIAL TRANSFER IN REINFORCEMENT LEARNING WITH A GENERATIVE MODEL	9423
<i>A. Tirinzoni, R. Poiani, M. Restelli</i>	
CONVOLUTIONAL DICTIONARY LEARNING BASED AUTO-ENCODERS FOR NATURAL EXPONENTIAL-FAMILY DISTRIBUTIONS	9435
<i>Bahareh Tolooshams, Andrew H. Song, Simona Temereanca, Demba Ba</i>	
MULTI-STEP GREEDY REINFORCEMENT LEARNING ALGORITHMS	9446
<i>M. Tomar, Y. Efroni, M. Ghavamzadeh</i>	
CHOICE SET OPTIMIZATION UNDER DISCRETE CHOICE MODELS OF GROUP DECISIONS	9456
<i>Kiran Tomlinson, Austin R. Benson</i>	
TRAJECTORYNET: A DYNAMIC OPTIMAL TRANSPORT NETWORK FOR MODELING CELLULAR DYNAMICS	9468
<i>Alexander Tong, Jessie Huang, Guy Wolf, David van Dijk, Smita Krishnaswamy</i>	

ALLEVIATING PRIVACY ATTACKS VIA CAUSAL LEARNING	9479
<i>S. Tople, A. Sharma, A. Nori</i>	
BAYESIAN LEARNING FROM SEQUENTIAL DATA USING GAUSSIAN PROCESSES WITH SIGNATURE COVARIANCES.....	9490
<i>Csaba Toth, Harald Oberhauser</i>	
FUNDAMENTAL TRADEOFFS BETWEEN INVARIANCE AND SENSITIVITY TO ADVERSARIAL PERTURBATIONS	9503
<i>F. Tramer, J. Behrmann, N. Carlini, N. Papernot, J. Jacobsen</i>	
STOCHASTIC GAUSS-NEWTON ALGORITHMS FOR NONCONVEX COMPOSITIONAL OPTIMIZATION	9514
<i>Quoc Tran-Dinh, Nhan H. Pham, Lam M. Nguyen</i>	
BAYESIAN DIFFERENTIAL PRIVACY FOR MACHINE LEARNING	9525
<i>Aleksei Triastcyn, Boi Faltings</i>	
SINGLE POINT TRANSDUCTIVE PREDICTION	9535
<i>N. Tripuraneni, L. Mackey</i>	
GRAPHOPT: LEARNING OPTIMIZATION MODELS OF GRAPH FORMATION	9545
<i>R. Trivedi, J. Yang, H. Zha</i>	
TRANSFER LEARNING WITHOUT KNOWING: REPROGRAMMING BLACK-BOX MACHINE LEARNING MODELS WITH SCARCE DATA AND LIMITED RESOURCES	9556
<i>Yun-Yun Tsai, Pin-Yu Chen, Tsung-Yi Ho</i>	
FROM IMAGENET TO IMAGE CLASSIFICATION: CONTEXTUALIZING PROGRESS ON BENCHMARKS	9567
<i>D. Tsipras, S. Santurkar, L. Engstrom, A. Ilyas, A. Madry</i>	
NORMALIZED FLAT MINIMA: EXPLORING SCALE INVARIANT DEFINITION OF FLAT MINIMA FOR NEURAL NETWORKS USING PAC-BAYESIAN ANALYSIS	9578
<i>Yusuke Tsuzuku, Issei Sato, Masashi Sugiyama</i>	
APPROXIMATING STACKED AND BIDIRECTIONAL RECURRENT ARCHITECTURES WITH THE DELAYED RECURRENT NEURAL NETWORK.....	9590
<i>Javier S. Turek, Shailee Jain, Vy A. Vo, Mihai Capota, Alexander G. Huth, Theodore L. Willke</i>	
MINIMAXWEIGHT AND Q-FUNCTION LEARNING FOR OFF-POLICY EVALUATION	9601
<i>M. Uehara, J. Huang, N. Jiang</i>	
STOCHASTICRANK: GLOBAL OPTIMIZATION OF SCALE-FREE DISCRETE FUNCTIONS.....	9611
<i>Aleksei Ustimenko, Liudmila Prokhorenkova</i>	
UNDIRECTED GRAPHICAL MODELS AS APPROXIMATE POSTERiors	9622
<i>Arash Vahdat, Evgeny Andriyash, William G. Macready</i>	
UNCERTAINTY ESTIMATION USING A SINGLE DEEP DETERMINISTIC NEURAL NETWORK.....	9632
<i>Joost van Amersfoort, Lewis Smith, Yee Whye Teh, Yarin Gal</i>	
DEEP MOLECULAR PROGRAMMING: A NATURAL IMPLEMENTATION OF BINARY-WEIGHT RELU NEURAL NETWORKS.....	9643
<i>Marko Vasic, Cameron Chalk, Sarfraz Khurshid, David Soloveichik</i>	
LINEAR BANDITS WITH STOCHASTIC DELAYED FEEDBACK.....	9654
<i>C. Vernade, A. Carpentier, T. Lattimore, G. Zappella, B. Ermi, M. Brueckner</i>	
NON-STATIONARY DELAYED BANDITS WITH INTERMEDIATE OBSERVATIONS.....	9664
<i>Claire Vernade, András György, Timothy A. Mann</i>	
OPTIONS AS RESPONSES: GROUNDING BEHAVIOURAL HIERARCHIES IN MULTI-AGENT REINFORCEMENT LEARNING	9675
<i>Alexander Sasha Vezhnevets, Yuhuai Tony Wu, Maria Eckstein, Rémi Leblond, Joel Z. Leibo</i>	
BORN-AGAIN TREE ENSEMBLES	9685
<i>Thibaut Vidal, Maximilian Schiffer</i>	
PRIVATE REINFORCEMENT LEARNING WITH PAC AND REGRET GUARANTEES.....	9696
<i>Giuseppe Vietri, Borja Balle, Akshay Krishnamurthy, Zhiwei Steven Wu</i>	
NEW ORACLE-EFFICIENT ALGORITHMS FOR PRIVATE SYNTHETIC DATA RELEASE	9707
<i>Giuseppe Vietri, Grace Tian, Mark Bun, Thomas Steinke, Zhiwei Steven Wu</i>	
CONDITIONAL GRADIENT METHODS FOR STOCHASTICALLY CONSTRAINED CONVEX MINIMIZATION	9717
<i>Maria-Luiza Vladarean, Ahmet Alacaoglu, Ya-Ping Hsieh, Volkan Cevher</i>	
UNSUPERVISED DISCOVERY OF INTERPRETABLE DIRECTIONS IN THE GAN LATENT SPACE	9728
<i>Andrey Voynov, Artem Babenko</i>	
SAFE REINFORCEMENT LEARNING IN CONSTRAINED MARKOV DECISION PROCESSES	9739
<i>Akifumi Wachi, Yanan Sui</i>	
ORTHOGONALIZED SGD AND NESTED ARCHITECTURES FOR ANYTIME NEURAL NETWORKS	9749
<i>Chengcheng Wan, Henry Hoffmann, Shan Lu, Michael Maire</i>	
PROJECTION-FREE DISTRIBUTED ONLINE CONVEX OPTIMIZATION WITH O(T) COMMUNICATION COMPLEXITY.....	9760
<i>Yuanyu Wan, Wei-Wei Tu, Lijun Zhang</i>	
LOGISTIC REGRESSION FOR MASSIVE DATA WITH RARE EVENTS.....	9771
<i>HaiYing Wang</i>	
ON THE GLOBAL OPTIMALITY OF MODEL-AGNOSTIC META-LEARNING: REINFORCEMENT LEARNING AND SUPERVISED LEARNING.....	9779
<i>L. Wang, Q. Cai, Z. Yang, Z. Wang</i>	
TOWARDS ACCURATE POST-TRAINING NETWORK QUANTIZATION VIA BIT-SPLIT AND STITCHING	9789
<i>Peisong Wang, Qiang Chen, Xiangyu He, Jian Cheng</i>	

SELF-MODULATING NONPARAMETRIC EVENT-TENSOR FACTORIZATION	9799
<i>Zheng Wang, Xinqi Chu, Shandian Zhe</i>	
UPPER BOUNDS FOR MODEL-FREE ROW-SPARSE PRINCIPAL COMPONENT ANALYSIS	9810
<i>G. Wang, S. Dey</i>	
ROMA: MULTI-AGENT REINFORCEMENT LEARNING WITH EMERGENT ROLES	9818
<i>Tonghan Wang, Heng Dong, Victor Lesser, Chongjie Zhang</i>	
NONSTATIONARY NONSEPARABLE RANDOM FIELDS	9829
<i>Kangrui Wang, Oliver Hamelijnck, Theodoros Damoulas, Mark Steel</i>	
CONTINUOUSLY INDEXED DOMAIN ADAPTATION	9840
<i>Hao Wang, Hao He, Dina Katabi</i>	
LEARNING EFFICIENT MULTI-AGENT COMMUNICATION: AN INFORMATION BOTTLENECK APPROACH	9850
<i>Rundong Wang, Xu He, Runsheng Yu, Wei qiu, Bo An, Zinovi Rabinovich</i>	
FRUSTRATINGLY SIMPLE FEW-SHOT OBJECT DETECTION	9861
<i>Xin Wang, Thomas E. Huang, Trevor Darrell, Joseph E. Gonzalez, Fisher Yu</i>	
UNDERSTANDING CONTRASTIVE REPRESENTATION LEARNING THROUGH ALIGNMENT AND UNIFORMITY ON THE HYPERSPHERE	9871
<i>Tongzhou Wang, Phillip Isola</i>	
ENHANCED POET: OPEN-ENDED REINFORCEMENT LEARNING THROUGH UNBOUNDED INVENTION OF LEARNING CHALLENGES AND THEIR SOLUTIONS	9882
<i>R. Wang, J. Leman, A. Rawal, J. Zhi, Y. Li, J. Clune, K. Stanley</i>	
HAAR GRAPH POOLING	9894
<i>Yu Guang Wang, Ming Li, Zheng Ma, Guido Montúfar, Xiaosheng Zhuang, Yanan Fan</i>	
DEEP STREAMING LABEL LEARNING	9905
<i>Z. Wang, L. Liu, D. Tao</i>	
BOXHED: BOOSTED EXACT HAZARD ESTIMATOR WITH DYNAMIC COVARIATES	9915
<i>Xiaochen Wang, Arash Pakbin, Bobak J. Mortazavi, Hongyu Zhao, Donald K.K. Lee</i>	
OPTIMIZING DATA USAGE VIA DIFFERENTIABLE REWARDS	9925
<i>Xinyi Wang, Hieu Pham, Paul Michel, Antonios Anastasopoulos, Jaime Carbonell, Graham Neubig</i>	
BANDITS FOR BMO FUNCTIONS	9938
<i>Tianyu Wang, Cynthia Rudin</i>	
WHEN DEEP DENOISING MEETS ITERATIVE PHASE RETRIEVAL	9949
<i>Yaotian Wang, Xiaohang Sun, Jason W. Fleischer</i>	
DOUBLY STOCHASTIC VARIATIONAL INFERENCE FOR NEURAL PROCESSES WITH HIERARCHICAL LATENT VARIABLES	9960
<i>Qi Wang, Herke van Hoof</i>	
LOSS FUNCTION SEARCH FOR FACE RECOGNITION	9971
<i>Xiaobo Wang, Shuo Wang, Cheng Chi, Shijeng Zhang, Tao Mei</i>	
SEQUENTIAL COOPERATIVE BAYESIAN INFERENCE	9981
<i>Junqi Wang, Pei Wang, Patrick Shafto</i>	
NEURAL NETWORK CONTROL POLICY VERIFICATION WITH PERSISTENT ADVERSARIAL PERTURBATIONS	9992
<i>Yuh-Shyang Wang, Tsui-Wei Weng, Luca Daniel</i>	
COST-EFFECTIVELY IDENTIFYING CAUSAL EFFECTS WHEN ONLY RESPONSE VARIABLE IS OBSERVABLE	10002
<i>Tian-Zuo Wang, Xi-Zhu Wu, Sheng-Jun Huang, Zhi-Hua Zhou</i>	
STRIVING FOR SIMPLICITY AND PERFORMANCE IN OFF-POLICY DRL: OUTPUT NORMALIZATION AND NON-UNIFORM SAMPLING	10012
<i>Che Wang, Yanqiu Wu, Quan Vuong, Keith Ross</i>	
ON DIFFERENTIALLY PRIVATE STOCHASTIC CONVEX OPTIMIZATION WITH HEAVY-TAILED DATA	10023
<i>Di Wang, Hanshen Xiao, Srinu Devadas, Jinhui Xu</i>	
BREAKING THE CURSE OF MANY AGENTS: PROVABLE MEAN EMBEDDING Q-ITERATION FOR MEAN-FIELD REINFORCEMENT LEARNING	10034
<i>L. Wang, Z. Yang, Z. Wang</i>	
THOMPSON SAMPLING VIA LOCAL UNCERTAINTY	10046
<i>Zhendong Wang, Mingyuan Zhou</i>	
A NEARLY-LINEAR TIME ALGORITHM FOR EXACT COMMUNITY RECOVERY IN STOCHASTIC BLOCK MODEL	10057
<i>Peng Wang, Zirui Zhou, Anthony Man-Cho So</i>	
LEARNING REPRESENTATIONS THAT SUPPORT EXTRAPOLATION	10067
<i>Taylor W. Webb, Zachary Dulberg, Steven M. Frankland, Alexander A. Petrov, Randall C. O'Reilly, Jonathan D. Cohen</i>	

PART 14

POKED: A SEMI-SUPERVISED SYSTEM FOR WORD SENSE DISAMBIGUATION	10078
<i>Feng Wei</i>	
TUNING-FREE PLUG-AND-PLAY PROXIMAL ALGORITHM FOR INVERSE IMAGING PROBLEMS	10089
<i>Kaixuan Wei, Angelica Aviles-Rivero, Jingwei Liang, Ying Fu, Carola-Bibiane Schönlieb, Hua Huang</i>	

MODEL-FREE REINFORCEMENT LEARNING IN INFINITE-HORIZON AVERAGE-REWARD MARKOV DECISION PROCESSES	10101
<i>C. Wei, M. Jafarinia-Jahromi, H. Luo, H. Sharma, R. Jain</i>	
THE IMPLICIT AND EXPLICIT REGULARIZATION EFFECTS OF DROPOUT	10112
<i>C. Wei, S. Kakade, T. Ma</i>	
ONLINE CONTROL OF THE FALSE COVERAGE RATE AND FALSE SIGN RATE	10124
<i>Asaf Weinstein, Aaditya Ramdas</i>	
BATCH STATIONARY DISTRIBUTION ESTIMATION	10134
<i>Junfeng Wen, Bo Dai, Lihong Li, Dale Schuurmans</i>	
DOMAIN AGGREGATION NETWORKS FOR MULTI-SOURCE DOMAIN ADAPTATION	10145
<i>Junfeng Wen, Russ Greiner, Dale Schuurmans</i>	
TOWARDS UNDERSTANDING THE REGULARIZATION OF ADVERSARIAL ROBUSTNESS ON NEURAL NETWORKS	10156
<i>Yuxin Wen, Shuai Li, Kui Jia</i>	
AMORTISED LEARNING BY WAKE-SLEEP	10167
<i>L. Wenliang, T. Moskovitz, H. Kanagawa, M. Sahani</i>	
HOW GOOD IS THE BAYES POSTERIOR IN DEEP NEURAL NETWORKS REALLY?	10179
<i>Florian Wenzel, Kevin Roth, Bastiaan S. Veeling, Jakub Swiatkowski, Linh Tran, Stephan Mandt, Jasper Snoek, Tim Salimans, Rodolphe Jenatton, Sebastian Nowozin</i>	
PREDICTIVE SAMPLING WITH FORECASTING AUTOREGRESSIVE MODELS	10191
<i>Auke Wiggers, Emiel Hooeboom</i>	
STATE SPACE EXPECTATION PROPAGATION: EFFICIENT INFERENCE SCHEMES FOR TEMPORAL GAUSSIAN PROCESSES	10201
<i>William J. Wilkinson, Paul E. Chang, Michael Riis Andersen, Arno Solin</i>	
EFFICIENT NONPARAMETRIC STATISTICAL INFERENCE ON POPULATION FEATURE IMPORTANCE USING SHAPLEY VALUES	10213
<i>Brian D. Williamson, Jean Feng</i>	
EFFICIENTLY SAMPLING FUNCTIONS FROM GAUSSIAN PROCESS POSTERIORIS	10223
<i>J. Wilson, V. Borovitskiy, A. Terenin, P. Mostowsky, M. Deisenroth</i>	
LEARNING TO RANK LEARNING CURVES	10234
<i>Martin Wistuba, Tejaswini Pedapati</i>	
CAUSAL INFERENCE USING GAUSSIAN PROCESSES WITH STRUCTURED LATENT CONFOUNDERS	10244
<i>S. Witty, K. Takatsu, D. Jensen, V. Mansinghka</i>	
NEAR INPUT SPARSITY TIME KERNEL EMBEDDINGS VIA ADAPTIVE SAMPLING	10255
<i>David P. Woodruff, Amir Zandieh</i>	
IS LOCAL SGD BETTER THAN MINIBATCH SGD?	10265
<i>B. Woodworth, K. Patel, S. Stich, Z. Dai, B. Bullins, H. McMahan, O. Shamir, N. Srebro</i>	
OBTAINING ADJUSTABLE REGULARIZATION FOR FREE VIA ITERATE AVERAGING	10275
<i>J. Wu, V. Braverman, L. Yang</i>	
DELTA GRAD: RAPID RETRAINING OF MACHINE LEARNING MODELS	10286
<i>Yinjun Wu, Edgar Dobriban, Susan B. Davidson</i>	
ON THE NOISY GRADIENT DESCENT THAT GENERALIZES AS SGD	10298
<i>J. Wu, W. Hu, H. Xiong, J. Huan, V. Braverman, Z. Zhu</i>	
STRONGER AND FASTER WASSERSTEIN ADVERSARIAL ATTACKS	10308
<i>Kaiwen Wu, Allen Houze Wang, Yaoliang Yu</i>	
SEQUENCE GENERATION WITH MIXED REPRESENTATIONS	10319
<i>Lijun Wu, Shufang Xie, Yingce Xia, Yang Fan, Tao Qin, Jianhuang Lai, Tie-Yan Liu</i>	
ADVERSARIAL ROBUSTNESS VIA RUNTIME MASKING AND CLEANSING	10330
<i>Y. Wu, C. Yuan, S. Wu</i>	
ON THE GENERALIZATION EFFECTS OF LINEAR TRANSFORMATIONS IN DATA AUGMENTATION	10341
<i>S. Wu, H. Zhang, G. Valiant, C. Re</i>	
AMORTIZED POPULATION GIBBS SAMPLERS WITH NEURAL SUFFICIENT STATISTICS	10352
<i>Hao Wu, Heiko Zimmermann, Eli Sennesh, Tuan Anh Le, Jan-Willem van de Meent</i>	
CONTINUOUS GRAPH NEURAL NETWORKS	10363
<i>Louis-Pascal A. C. Xhonneux, Meng Qu, Jian Tang</i>	
A FLEXIBLE FRAMEWORK FOR NONPARAMETRIC GRAPHICAL MODELING THAT ACCOMMODATES MACHINE LEARNING	10373
<i>Yunhua Xiang, Noah Simon</i>	
GENERATIVE FLOWS WITH MATRIX EXPONENTIAL	10383
<i>Changyi Xiao, Ligang Liu</i>	
TRAINABILITY_GENERALIZATION_ICML2020 (6)	10393
<i>Lechao Xiao</i>	
OPTIMALLY SOLVING TWO-AGENT DECENTRALIZED POMDPs UNDER ONE-SIDED INFORMATION SHARING	10404
<i>Y. Xie, J. Dibangoye, O. Buffet</i>	
MAXIMUM-AND-CONCATENATION NETWORKS	10414
<i>Xingyu Xie, Hao Kong, Jianlong Wu, Wayne Zhang, Guangcan Liu, Zhouchen Lin</i>	
ZENO++: ROBUST FULLY ASYNCHRONOUS SGD	10426
<i>Cong Xie, Oluwasanmi Koyejo, Indranil Gupta</i>	

LOWER COMPLEXITY BOUNDS FOR FINITE-SUM CONVEX-CONCAVE MINIMAX OPTIMIZATION PROBLEMS	10435
<i>Guangzeng Xie, Luo Luo, Yijiang Lian, Zhihua Zhang</i>	
ON THE NUMBER OF LINEAR REGIONS OF CONVOLUTIONAL NEURAL NETWORKS	10445
<i>Huan Xiong, Lei Huang, Mengyang Yu, Li Liu, Fan Zhu, Ling Shao</i>	
ON LAYER NORMALIZATION IN THE TRANSFORMER ARCHITECTURE	10455
<i>Ruibin Xiong†, Yunchang Yang, Di He, Kai Zheng, Shuxin Zheng, Chen Xing, Huishuai Zhang, Yanyan Lan, Liwei Wang, Tie-Yan Liu</i>	
ON VARIATIONAL LEARNING OF CONTROLLABLE REPRESENTATIONS FOR TEXT WITHOUT SUPERVISION	10465
<i>P. Xu, J. Cheung, Y. Cao</i>	
CLASS-WEIGHTED CLASSIFICATION: TRADE-OFFS AND ROBUST APPROACHES	10475
<i>Ziyu Xu, Chen Dan, Justin Khim, Pradeep Ravikumar</i>	
A FINITE-TIME ANALYSIS OF Q-LEARNING WITH NEURAL NETWORK FUNCTION APPROXIMATION	10486
<i>Pan Xu, Quanquan Gu</i>	
UNDERSTANDING AND STABILIZING GANS' TRAINING DYNAMICS USING CONTROL THEORY	10497
<i>Kun Xu, Chongxuan Li, Jun Zhu, Bo Zhang</i>	
LEARNING AUTOENCODERS WITH RELATIONAL REGULARIZATION	10507
<i>Hongteng Xu, Dixin Luo, Ricardo Henao, Svati Shah, Lawrence Carin</i>	
LEARNING FACTORIZED WEIGHT MATRIX FOR JOINT FILTERING	10518
<i>Xiangyu Xu, Yongrui Ma, Wenxiu Sun</i>	
VARIATIONAL LABEL ENHANCEMENT	10528
<i>Ning Xu, Jun Shu, Yun-Peng Liu, Xin Geng</i>	
PREDICTION-GUIDED MULTI-OBJECTIVE REINFORCEMENT LEARNING FOR CONTINUOUS ROBOT CONTROL	10538
<i>Jie Xu, Yunsheng Tian, Pingchuan Ma, Daniela Rus, Shinjiro Sueda, Wojciech Matusik</i>	
METAFUN: META-LEARNING WITH ITERATIVE FUNCTIONAL UPDATES	10548
<i>Jin Xu, Jean-Francois Ton, Hyunjik Kim, Adam R. Kosiorek, Yee Whye Teh</i>	
VIDEO PREDICTION VIA EXAMPLE GUIDANCE	10559
<i>Jingwei Xu, Huazhe Xu, Bingbing Ni, Xiaokang Yang, Trevor Darrell</i>	
AMORTIZED FINITE ELEMENT ANALYSIS FOR FAST PDE-CONSTRAINED OPTIMIZATION	10569
<i>Tianju Xue, Alex Beaton, Sigrid Adriaenssens, Ryan P. Adams</i>	
FEATURE SELECTION USING STOCHASTIC GATES	10579
<i>Yutaro Yamada, Ofir Lindenbaum, Sahand Negahban, Yuval Kluger</i>	
STOCHASTIC OPTIMIZATION FOR NON-CONVEX INF-PROJECTION PROBLEMS	10591
<i>Y. Yan, Y. Xu, L. Zhang, X. Wang, T. Yang</i>	
VARIATIONAL BAYESIAN QUANTIZATION	10601
<i>Yibo Yang, Robert Bamler, Stephan Mandt</i>	
ENERGY-BASED PROCESSES FOR EXCHANGEABLE DATA	10612
<i>Sherry Yang</i>	
RANDOMIZED SMOOTHING OF ALL SHAPES AND SIZES	10624
<i>G. Yang, T. Duan, J. Hu, H. Salman, H. Razenshteyn, J. Li</i>	
Q-VALUE PATH DECOMPOSITION FOR DEEP MULTIAGENT REINFORCEMENT LEARNING	10637
<i>Yaodong Yang, Jianye Hao, Guangyong Chen, Hongyao Tang, Yingfeng Chen, Yujing Hu, Changjie Fan, Zhongyu Wei</i>	
IMPROVING MOLECULAR DESIGN BY STOCHASTIC ITERATIVE TARGET AUGMENTATION	10647
<i>K. Yang, W. Jin, K. Swanson, R. Barzilay, T. Jaakkola</i>	
ON THE CONSISTENCY OF TOP-K SURROGATE LOSSES	10658
<i>Forest Yang, Sanmi Koyejo</i>	
INTERPOLATION BETWEEN RESIDUAL AND NON-RESIDUAL NETWORKS	10667
<i>Zonghan Yang, Yang Liu, Chenglong Bao, Zuoqiang Shi</i>	
REINFORCEMENT LEARNING IN FEATURE SPACE: MATRIX BANDIT, KERNELS, AND REGRET BOUND	10677
<i>L. Yang, M. Wang</i>	
MULTI-AGENT DETERMINANTAL Q-LEARNING	10688
<i>Yaodong Yang, Ying Wen, Liheng Chen, Jun Wang, Kun Shao, David Mguni, Weinan Zhang</i>	
RETHINKING BIAS-VARIANCE TRADE-OFF FOR GENERALIZATION OF NEURAL NETWORKS	10698
<i>Zitong Yang, Yaodong Yu, Chong You, Jacob Steinhardt, Yi Ma</i>	
UNSUPERVISED TRANSFER LEARNING FOR SPATIOTEMPORAL PREDICTIVE NETWORKS	10709
<i>Zhiyu Yao, Yunbo Wang, Mingsheng Long, Jianmin Wang</i>	
SEARCHING TO EXPLOIT MEMORIZATION EFFECT IN LEARNING WITH NOISY LABELS	10720
<i>Q. Yao, H. Yang, B. Han, G. Niu, J. Kwok</i>	
GRAPH-BASED, SELF-SUPERVISED PROGRAM REPAIR FROM DIAGNOSTIC FEEDBACK	10730
<i>Michihiro Yasunaga, Percy Liang</i>	
PRETRAINED GENERALIZED AUTOREGRESSIVE MODEL WITH ADAPTIVE PROBABILISTIC LABEL CLUSTERS FOR EXTREME MULTI-LABEL TEXT CLASSIFICATION	10740
<i>Hui Ye, Zhiyu Chen, Da-Han Wang, Brian D. Davison</i>	
GOOD SUBNETWORKS PROVABLY EXIST: PRUNING VIA GREEDY FORWARD SELECTION	10751
<i>Mao Ye, Chengyue Gong, Lizhen Nie, Denny Zhou, Adam Klivans, Qiang Liu</i>	
IT'S NOT WHAT MACHINES CAN LEARN, IT'S WHAT WE CANNOT TEACH	10762
<i>Gal Yehuda, Moshe Gabel, Assaf Schuster</i>	

DATA VALUATION USING REINFORCEMENT LEARNING	10773
<i>Jinsung Yoon, Sercan Ö. Arık, Tomas Pfister</i>	
XTARNET: LEARNING TO EXTRACT TASK-ADAPTIVE REPRESENTATION FOR INCREMENTAL FEW-SHOT LEARNING	10783
<i>Sung Whan Yoon, Do-Yeon Kim, Jun Seo, Jaekyun Moon</i>	
ROBUSTIFYING SEQUENTIAL NEURAL PROCESSES	10792
<i>J. Yoon, G. Singh, S. Ahn</i>	
WHEN DOES SELF-SUPERVISION HELP GRAPH CONVOLUTIONAL NETWORKS?	10802
<i>Yuning You, Tianlong Chen, Zhangyang Wang, Yang Shen</i>	
GRAPH STRUCTURE OF NEURAL NETWORKS	10812
<i>Jiaxuan You, Jure Leskovec, Kaiming He, Saining Xie</i>	
SIMULTANEOUS INFERENCE FOR MASSIVE DATA: DISTRIBUTED BOOTSTRAP	10823
<i>Yang Yu, Shih-Kang Chao, Guang Cheng</i>	
GRAPHICAL MODELS MEET BANDITS: A VARIATIONAL THOMPSON SAMPLING APPROACH	10833
<i>Tong Yu, Branislav Kveton, Zheng Wen, Ruiyi Zhang, Ole J. Mengshoel</i>	
LABEL-NOISE ROBUST DOMAIN ADAPTATION	10844
<i>Xiyu Yu, Tongliang Liu, Mingming Gong, Kun Zhang, Kayhan Batmanghelich, Dacheng Tao</i>	
INTRINSIC REWARD DRIVEN IMITATION LEARNING VIA GENERATIVE MODEL	10856
<i>X. Yu, Y. Lyu, I. Tsang</i>	

PART 15

GRAPH CONVOLUTIONAL NETWORK FOR RECOMMENDATION WITH LOW-PASS COLLABORATIVE FILTERS	10867
<i>Wenhui Yu, Zheng Qin</i>	
FEDERATED LEARNING WITH ONLY POSITIVE LABELS	10878
<i>Felix X. Yu, Ankit Singh Rawat, Aditya Krishna Menon, Sanjiv Kumar</i>	
TRAINING DEEP ENERGY-BASED MODELS WITH F-DIVERGENCE MINIMIZATION	10889
<i>L. Yu, Y. Song, J. Song, S. Ermon</i>	
GRAPH RANDOM NEURAL FEATURES FOR DISTANCE-PRESERVING GRAPH REPRESENTATIONS	10900
<i>Daniele Zambon, Cesare Alippi, Lorenzo Livi</i>	
LEARNING NEAR OPTIMAL POLICIES WITH LOW INHERENT BELLMAN ERROR	10910
<i>A. Zanette, A. Lazaric, M. Kochenderfer, E. Brunskill</i>	
SCALING UP HYBRID PROBABILISTIC INFERENCE WITH LOGICAL AND ARITHMETIC CONSTRAINTS VIA MESSAGE PASSING	10922
<i>Z. Zeng, P. Morettin, F. Yan, A. Vergari, G. Broeck</i>	
LEARNING CALIBRATABLE POLICIES USING PROGRAMMATIC STYLE-CONSISTENCY	10933
<i>Eric Zhan, Albert Tseng, Yisong Yue, Adith Swaminathan, Matthew Hausknecht</i>	
DESIGNING OPTIMAL DYNAMIC TREATMENT REGIMES: A CAUSAL REINFORCEMENT LEARNING APPROACH	10944
<i>Junzhe Zhang, Elias Bareinboim</i>	
ROBUSTNESS TO PROGRAMMABLE STRING TRANSFORMATIONS VIA AUGMENTED ABSTRACT TRAINING	10955
<i>Y. Zhang, A. Albarghouti, L. D'Antoni</i>	
CONVERGING TO TEAM-MAXMIN EQUILIBRIA IN ZERO-SUM MULTIPLAYER GAMES	10967
<i>Y. Zhang, B. An</i>	
GENERATIVE ADVERSARIAL IMITATION LEARNING WITH NEURAL NETWORK PARAMETERIZATION: GLOBAL OPTIMALITY AND CONVERGENCE RATE	10978
<i>Y. Zhang, Q. Cai, Z. Yang, Z. Wang</i>	
CAUTIOUS ADAPTATION FOR REINFORCEMENT LEARNING IN SAFETY-CRITICAL SETTINGS	10989
<i>J. Zhang, B. Cheung, C. Finn, S. Levine, D. Jayaraman</i>	
LEARNING THE VALUATIONS OF A K-DEMAND AGENT	11000
<i>Hanrui Zhang, Vincent Conitzer</i>	
A TREE-STRUCTURED DECODER FOR IMAGE-TO-MARKUP GENERATION	11010
<i>Jianshu Zhang, Jun Du, Yongxin Yang, Yi-Zhe Song, Si Wei, Lirong Dai</i>	
APPROXIMATION CAPABILITIES OF NEURAL ODES AND INVERTIBLE RESIDUAL NETWORKS	11020
<i>Han Zhang, Xi Gao, Jacob Unterman, Tom Arodz</i>	
RANDOM HYPERVOLUME SCALARIZATIONS FOR PROVABLE MULTI-OBJECTIVE BLACK BOX OPTIMIZATION	11030
<i>D. Golovin, Q. Zhang</i>	
SPREAD DIVERGENCE	11040
<i>M. Zhang, P. Hayes, T. Bird, R. Habib, D. Barber</i>	
MIX-N-MATCH: ENSEMBLE AND COMPOSITIONAL METHODS FOR UNCERTAINTY CALIBRATION IN DEEP LEARNING	11051
<i>Jize Zhang, Bhavya Kailkhura, T. Yong-Jin Han</i>	
PRIVATELY LEARNING MARKOV RANDOM FIELDS	11063
<i>Huanyu Zhang, Gautam Kamath, Janardhan Kulkarni, Zhiwei Steven Wu</i>	

LEARNING STRUCTURED LATENT FACTORS FROM DEPENDENT DATA:A GENERATIVE MODEL FRAMEWORK FROM INFORMATION-THEORETIC PERSPECTIVE	11075
<i>Ruixiang Zhang, Masanori Koyama, Katsuhiko Ishiguro</i>	
OPTIMAL ESTIMATOR FOR UNLABELED LINEAR REGRESSION	11087
<i>H. Zhang, P. Li</i>	
DUAL-PATH DISTILLATION: A UNIFIED FRAMEWORK TO IMPROVE BLACK-BOX ATTACKS	11097
<i>Yonggang Zhang, Ya Li, Tongliang Liu, Xinmei Tian</i>	
COMPLEXITY OF FINDING STATIONARY POINTS OF NONSMOOTH NONCONVEX FUNCTIONS	11107
<i>J. Zhang, H. Lin, S. Jegelka, S. Sra, A. Jadbabaie</i>	
SELF-ATTENTIVE HAWKES PROCESS	11117
<i>Qiang Zhang, Aldo Lipani, Omer Kirnap, Emine Yilmaz</i>	
GRADIENTDICE: RETHINKING GENERALIZED OFFLINE ESTIMATION OF STATIONARY VALUES	11128
<i>Shangdong Zhang, Bo Liu, Shimon Whiteson</i>	
PROVABLY CONVERGENT TWO-TIMESCALE OFF-POLICY ACTOR-CRITIC WITH FUNCTION APPROXIMATION	11139
<i>Shangdong Zhang, Bo Liu, Hengshuai Yao, Shimon Whiteson</i>	
INVARIANT CAUSAL PREDICTION FOR BLOCK MDPS	11150
<i>Amy Zhang, Clare Lyle, Shagun Sodhani, Angelos Filos, Marta Kwiatkowska, Joelle Pineau, Yarin Gal, Doina Precup</i>	
ADAPTIVE REWARD-POISONING ATTACKS AGAINST REINFORCEMENT LEARNING	11161
<i>Xuezhou Zhang, Yuzhe Ma, Adish Singla, Xiaojin Zhu</i>	
CAUSE: LEARNING GRANGER CAUSALITY FROM EVENT SEQUENCES USING ATTRIBUTION METHODS	11171
<i>W. Zhang, T. Panum, S. Jha, P. Chalasani, D. Page</i>	
CONVEX CALIBRATED SURROGATES FOR THE MULTI-LABEL F-MEASURE	11182
<i>Mingyuan Zhang, Harish G. Ramaswamy, Shivani Agarwal</i>	
SPARSIFIED LINEAR PROGRAMMING FOR ZERO-SUM EQUILIBRIUM FINDING	11192
<i>Brian Hu Zhang, Tuomas Sandholm</i>	
FAST LEARNING OF GRAPH NEURAL NETWORKS WITH GUARANTEED GENERALIZABILITY: ONE-HIDDEN-LAYER CASE	11204
<i>Shuai Zhang, Meng Wang, Sijia Liu, Pin-Yu Chen, Jinjun Xiong</i>	
ATTACKS WHICH DO NOT KILL TRAINING MAKE ADVERSARIAL LEARNING STRONGER	11214
<i>Jingfeng Zhang, Xilie Xu, Bo Han, Gang Niu, Lizhen Cui, Masashi Sugiyama, Mohan Kankanhalli</i>	
A FLEXIBLE LATENT SPACE MODEL FOR MULTILAYER NETWORKS	11225
<i>Xuefei Zhang, Songkai Xue, Ji Zhu</i>	
PERCEPTUAL GENERATIVE AUTOENCODERS	11235
<i>Zijun Zhang, Ruixiang Zhang, Zongpeng Li, Yoshua Bengio, Liam Paull</i>	
VARIANCE REDUCTION IN STOCHASTIC PARTICLE-OPTIMIZATION SAMPLING	11244
<i>Jianyi Zhang, Yang Zhao, Ruiyi Zhang, Lawrence Carin, Changyou Chen</i>	
LEARNING WITH FEATURE AND DISTRIBUTION EVOLVABLE STREAMS	11254
<i>Zhen-Yu Zhang, Peng Zhao, Yuan Jiang, Zhi-Hua Zhou</i>	
PEGASUS: PRE-TRAINING WITH EXTRACTED GAP-SENTENCES FOR ABSTRACTIVE SUMMARIZATION	11265
<i>Jingqing Zhang, Yao Zhao, Mohammad Saleh, Peter J. Liu</i>	
ON LEVERAGING PRETRAINED GANS FOR GENERATION WITH LIMITED DATA	11277
<i>Miaoyun Zhao, Yulai Cong, Lawrence Carin</i>	
ON LEARNING LANGUAGE-INVARIANT REPRESENTATIONS FOR UNIVERSAL MACHINE TRANSLATION	11289
<i>Han Zhao, Junjie Hu, Andrej Risteski</i>	
DO RNN AND LSTM HAVE LONG MEMORY?	11302
<i>Jingyu Zhao, Feiqing Huang, Jia Lv, Yanjie Duan, Zhen Qin, Guodong Li, Guangjian Tian</i>	
FEATURE QUANTIZATION IMPROVES GAN TRAINING	11313
<i>Yang Zhao, Chunyuan Li, Ping Yu, Jianfeng Gao, Changyou Chen</i>	
INDIVIDUAL CALIBRATION WITH RANDOMIZED FORECASTING	11324
<i>Shengjia Zhao, Tengyu Ma, Stefano Ermon</i>	
SMALLER, MORE ACCURATE REGRESSION FORESTS USING TREE ALTERNATING OPTIMIZATION	11335
<i>Arman Zharmagambetov, Miguel Á. Carreira-Perpiñán</i>	
LEARNING TO LEARN KERNELS WITH VARIATIONAL RANDOM FEATURES	11346
<i>Xiantong Zhen, Haoliang Sun, Yingjun Du, Jun Xu, Yilong Yin, Ling Shao, Cees Snoek</i>	
SHARP COMPOSITION BOUNDS FOR GAUSSIAN DIFFERENTIAL PRIVACY VIA EDGEWORTH EXPANSION	11357
<i>Qinqing Zheng, Jinshuo Dong, Qi Long, Weijie Su</i>	
WHAT CAN LEARNED INTRINSIC REWARDS CAPTURE?	11373
<i>Zeyu Zheng, Junhyuk Oh, Matteo Hessel, Zhongwen Xu, Manuel Kroiss, Hado van Hasselt, David Silver, Satinder Singh</i>	
ERROR-BOUNDED CORRECTION OF NOISY LABELS	11384
<i>Songzhu Zheng, Pengxiang Wu, Aman Goswami, Mayank Goswami, Dimitris Metaxas, Chao Chen</i>	
ROBUST GRAPH REPRESENTATION LEARNING VIA NEURAL SPARSIFICATION	11395
<i>Cheng Zheng, Bo Zong, Wei Cheng, Dongjin Song, Jingchao Ni, Wenchao Yu, Haifeng Chen, Wei Wang</i>	
BISECTION-BASED PRICING FOR REPEATED CONTEXTUAL AUCTIONS AGAINST STRATEGIC BUYER	11406
<i>Anton Zhiyanov, Alexey Drutsa</i>	

BEST ARM IDENTIFICATION FOR CASCADING BANDITS IN THE FIXED CONFIDENCE SETTING	11418
<i>Z. Zhong, W. Cheung, V. Tan</i>	
NEURAL CONTEXTUAL BANDITS WITH UCB-BASED EXPLORATION	11429
<i>D. Zhou, L. Li, Q. Gu</i>	
MONET3D: TOWARDS ACCURATE MONOCULAR 3D OBJECT LOCALIZATION IN REAL TIME	11440
<i>Xichuan Zhou, Yicong Peng, Chungqiao Long, Fengbo Ren, Cong Shi</i>	
NONPARAMETRIC SCORE ESTIMATORS	11450
<i>Y. Zhou, J. Shi, J. Zhu</i>	
TIME-CONSISTENT SELF-SUPERVISION FOR SEMI-SUPERVISED LEARNING	11460
<i>T. Zhou, S. Wang, J. Blimes</i>	
DIVIDE, CONQUER, AND COMBINE: A NEW INFERENCE STRATEGY FOR PROBABILISTIC PROGRAMS WITH STOCHASTIC SUPPORT	11471
<i>Yuan Zhou, Hongseok Yang, Yee Whye Teh, Tom Rainforth</i>	
GO WIDE, THEN NARROW: EFFICIENT TRAINING OF DEEP THIN NETWORKS	11483
<i>Denny Zhou, Mao Ye, Chen Chen, Tianjian Meng, Mingxing Tan, Xiaodan Song, Quoc Le, Qiang Liu, Dale Schuurmans</i>	
HYBRID STOCHASTIC-DETERMINISTIC MINIBATCH PROXIMAL GRADIENT: LESS-THAN-SINGLE-PASS OPTIMIZATION WITH NEARLY OPTIMAL GENERALIZATION	11493
<i>Pan Zhou, Xiao-Tong Yuan</i>	
ROBUST OUTLIER ARM IDENTIFICATION	11503
<i>Y. Zhu, S. Katariya, R. Nowak</i>	
VARIANCE REDUCTION AND QUASI-NEWTON FOR PARTICLE-BASED VARIATIONAL INFERENCE	11513
<i>Michael H. Zhu, Chang Liu, Jun Zhu</i>	
CAUSAL EFFECT ESTIMATION AND OPTIMAL DOSE SUGGESTIONS IN MOBILE HEALTH INTERNATIONAL CONFERENCE ON MACHINE LEARNING (ICML 2020)	11525
<i>Liangyu Zhu, Wenbin Lu, Rui Song</i>	
THOMPSON SAMPLING ALGORITHMS FOR MEAN-VARIANCE BANDITS	11536
<i>Q. Zhu, V. Tan</i>	
LEARNING ADVERSARIALY ROBUST REPRESENTATIONS VIA WORST-CASE MUTUAL INFORMATION MAXIMIZATION	11546
<i>Sicheng Zhu, Xiao Zhang, David Evans</i>	
LINEAR CONVERGENCE OF RANDOMIZED PRIMAL-DUAL COORDINATE METHOD FOR LARGE-SCALE LINEAR CONSTRAINED CONVEX PROGRAMMING	11556
<i>Daoli Zhu, Lei Zhao</i>	
WHEN DEMANDS EVOLVE LARGER AND NOISIER: LEARNING AND EARNING IN A GROWING ENVIRONMENT	11565
<i>Feng Zhu, Zeyu Zheng</i>	
ADAPTIVE CHECKPOINT ADJOINT METHOD FOR GRADIENT ESTIMATION IN NEURAL ODE	11575
<i>Juntang Zhuang, Nicha Dvornek, Xiaoxiao Li, Sekhar Tatikonda, Xenophon Papademetris, James Duncan</i>	
LEARNING OPTIMAL TREE MODELS UNDER BEAM SEARCH	11586
<i>Jingwei Zhuo, Zirui Xu, Wei Dai, Han Zhu, Han Li, Jian Xu, Kun Gai</i>	
LAPLACIAN REGULARIZED FEW-SHOT LEARNING	11596
<i>Imtiaz Masud Ziko, Jose Dolz, Eric Granger, Ismail Ben Ayed</i>	
INFLUENZA FORECASTING FRAMEWORK BASED ON GAUSSIAN PROCESSES	11607
<i>C. Zimmer, R. Yaesoubi</i>	
A GENERAL RECURRENT STATE SPACE FRAMEWORK FOR MODELING NEURAL DYNAMICS DURING DECISION-MAKING	11616
<i>David M. Zoltowski, Jonathan W. Pillow, Scott W. Linderman</i>	
TRANSFORMER HAWKES PROCESS	11628
<i>S. Zao, H. Jiang, Z. Li, T. Zhao, H. Zha</i>	
Author Index	