

38th Conference on Uncertainty in Artificial Intelligence (UAI 2022)

Proceedings of Machine Learning Research Volume 180

Eindhoven, The Netherlands
1-5 August 2022

Part 1 of 3

ISBN: 978-1-7138-6329-8

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© (2022) by Association For Uncertainty in Artificial Intelligence (AUAI)
All rights reserved.

Printed by Curran Associates, Inc. (2023)

For permission requests, please contact Association For Uncertainty in Artificial Intelligence (AUAI)
at the address below.

Association For Uncertainty in Artificial Intelligence (AUAI)
Department of Computer Science
University of British Columbia
Vancouver BC Canada V6T 1Z4

Phone: (703) 787-8044
Fax: (703) 478-0829

www.auai.org

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

VOLUME 1

| | |
|--|-----|
| Mutation-Driven Follow the Regularized Leader for Last-iterate Convergence in Zero-sum Games..... | 1 |
| <i>Kenshi Abe, Mitsuki Sakamoto, Atsushi Iwasaki</i> | |
| NeuroBE: Escalating Neural Network Approximations of Bucket Elimination..... | 11 |
| <i>Sakshi Agarwal, Kalev Kask, Alex Ihler, Rina Dechter</i> | |
| Regret Guarantees for Model-Based Reinforcement Learning with Long-term Average Constraints | 22 |
| <i>Mridul Agarwal, Qinbo Bai, Vaneet Aggarwal</i> | |
| GNN2GNN: Graph Neural Networks to Generate Neural Networks..... | 32 |
| <i>Andrea Agiollo, Andrea Omicini</i> | |
| Neuro-Symbolic Entropy Regularization | 43 |
| <i>Kareem Ahmed, Eric Wang, Kai-Wei Chang, Guy Van Den Broeck</i> | |
| Non-Parametric Inference of Relational Dependence | 54 |
| <i>Ragib Ahsan, Zahra Fatemi, David Arbour, Elena Zheleva</i> | |
| Data Dependent Randomized Smoothing..... | 64 |
| <i>Motasem Alfarra, Adel Bibi, Philip H. S. Torr, Bernard Ghanem</i> | |
| Multi-Winner Approval Voting Goes Epistemic..... | 75 |
| <i>Tahar Allouche, Jérôme Lang, Florian Yger</i> | |
| Inductive Synthesis of Finite-State Controllers for POMDPs | 85 |
| <i>Roman Andriushchenko, Milan Češka, Sebastian Junges, Joost-Pieter Katoen</i> | |
| Discovery of Extended Summary Graphs in Time Series | 96 |
| <i>Charles K. Assaad, Emilie Devijver, Eric Gaussier</i> | |
| Asymmetric DQN for Partially Observable Reinforcement Learning..... | 107 |
| <i>Andrea Baisero, Brett Daley, Christopher Amato</i> | |
| Physics Guided Neural Networks for Spatio-Temporal Super-resolution of Turbulent Flows | 118 |
| <i>Tianshu Bao, Shengyu Chen, Taylor T Johnson, Peyman Givi, Shervin Sammak, Xiaowei Jia</i> | |
| Byzantine-Tolerant Distributed Multiclass Sparse Linear Discriminant Analysis | 129 |
| <i>Yajie Bao, Weidong Liu, Xiaojun Mao, Weijia Xiong</i> | |
| Equilibrium Aggregation: Encoding Sets Via Optimization | 139 |
| <i>Sergey Bartunov, Fabian B. Fuchs, Timothy P. Lillicrap</i> | |
| Empirical Bayes Approach to Truth Discovery Problems..... | 150 |
| <i>Tsviel Ben Shabat, Reshef Meir, David Azriel</i> | |
| On Early Extinction and the Effect of Travelling in the SIR Model | 159 |
| <i>Petra Berenbrink, Colin Cooper, Cristina Gava, David Kohan Marzagão, Frederik Mallmann-Trenn, Tomasz Radzik</i> | |
| Learning Soft Interventions in Complex Equilibrium Systems | 170 |
| <i>Michel Besserve, Bernhard Schölkopf</i> | |

| | |
|---|-----|
| Identifying Near-Optimal Decisions in Linear-in-parameter Bandit Models with Continuous Decision Sets | 181 |
| <i>Sanjay P. Bhat, Chaitanya Amballa</i> | |
| Offline Change Detection Under Contamination | 191 |
| <i>Sujay Bhatt, Guanhua Fang, Ping Li</i> | |
| On Testability of the Front-Door Model Via Verma Constraints | 202 |
| <i>Rohit Bhattacharya, Razieh Nabi</i> | |
| Robustness of Model Predictions Under Extension | 213 |
| <i>Tineke Blom, Joris M. Mooij</i> | |
| Information Theoretic Approach to Detect Collusion in Multi-Agent Games | 223 |
| <i>Trevor Bonjour, Vaneet Aggarwal, Bharat Bhargava</i> | |
| Lifting in Multi-Agent Systems Under Uncertainty | 233 |
| <i>Tanya Braun, Marcel Gehrke, Florian Lau, Ralf Möller</i> | |
| On-The-fly Adaptation of Patrolling Strategies in Changing Environments | 244 |
| <i>Tomáš Brázdil, David Klačka, Antonín Kucera, Vi't Musil, Petr Novotný, Vojtech Reháč</i> | |
| On the Inductive Bias of Neural Networks for Learning Read-Once DNFs | 255 |
| <i>Ido Bronstein, Alon Brutzkus, Amir Globerson</i> | |
| AUTM Flow: Atomic Unrestricted Time Machine for Monotonic Normalizing Flows | 266 |
| <i>Difeng Cai, Yuliang Ji, Huan He, Qiang Ye, Yuanzhe Xi</i> | |
| Active Approximately Metric-Fair Learning | 275 |
| <i>Yiting Cao, Chao Lan</i> | |
| Capturing Actionable Dynamics with Structured Latent Ordinary Differential Equations | 286 |
| <i>Paidamoyo Chapfuwa, Sherri Rose, Lawrence Carin, Edward Meeds, Ricardo Henao</i> | |
| Privacy-Aware Compression for Federated Data Analysis | 296 |
| <i>Kamalika Chaudhuri, Chuan Guo, Mike Rabbat</i> | |
| The Optimal Noise in Noise-Contrastive Learning is Not What You Think | 307 |
| <i>Omar Chehab, Alexandre Gramfort, Aapo Hyvärinen</i> | |
| A Competitive Analysis of Online Failure-Aware Assignment | 317 |
| <i>Mengjing Chen, Pingzhong Tang, Zihong Wang, Shenke Xiao, Xiwang Yang</i> | |
| Stackmix: A Complementary Mix Algorithm | 326 |
| <i>John Chen, Samarth Sinha, Anastasios Kyrillidis</i> | |
| Knowledge Representation Combining Quaternion Path Integration and Depth-Wise Atrous Circular Convolution | 336 |
| <i>Xinyuan Chen, Zhongmei Zhou, Meichun Gao, Daya Shi, Mohd N. Husen</i> | |
| Sublinear Time Algorithms for Greedy Selection in High Dimensions | 346 |
| <i>Qi Chen, Kai Liu, Ruilong Yao, Hu Ding</i> | |
| Shoring Up the Foundations: Fusing Model Embeddings and Weak Supervision | 357 |
| <i>Mayee F. Chen, Daniel Y. Fu, Dyah Adila, Michael Zhang, Frederic Sala, Kayvon Fatahalian, Christopher Ré</i> | |

| | |
|---|-----|
| On the Definition and Computation of Causal Treewidth..... | 368 |
| <i>Yizuo Chen, Adnan Darwiche</i> | |
| Offline Reinforcement Learning Under Value and Density-Ratio Realizability: The Power of Gaps..... | 378 |
| <i>Jinglin Chen, Nan Jiang</i> | |
| Greedy Modality Selection Via Approximate Submodular Maximization..... | 389 |
| <i>Runxiang Cheng, Gargi Balasubramaniam, Yifei He, Yao-Hung Hubert Tsai, Han Zhao</i> | |
| Feature Selection for Discovering Distributional Treatment Effect Modifiers..... | 400 |
| <i>Yoichi Chikahara, Makoto Yamada, Hisashi Kashima</i> | |
| Combating the Instability of Mutual Information-Based Losses Via Regularization..... | 411 |
| <i>Kwanghee Choi, Siyeong Lee</i> | |
| A Geometric Method for Improved Uncertainty Estimation in Real-Time..... | 422 |
| <i>Gabriella Chouraqi, Liron Cohen, Gil Einziger, Liel Leman</i> | |
| Cyclic Test Time Augmentation with Entropy Weight Method..... | 433 |
| <i>Sewhan Chun, Jae Young Lee, Junmo Kim</i> | |
| Greedy Equivalence Search in the Presence of Latent Confounders..... | 443 |
| <i>Tom Claassen, Ioan G. Bucur</i> | |
| Counterfactual Inference of Second Opinions..... | 453 |
| <i>Nina L. Corvelo Benz, Manuel Gomez Rodriguez</i> | |
| Variational Message Passing Neural Network for Maximum-A-Posteriori (MAP) Inference..... | 464 |
| <i>Zijun Cui, Hanjing Wang, Tian Gao, Kartik Talamadupula, Qiang Ji</i> | |
| On Provably Robust meta-Bayesian Optimization..... | 475 |
| <i>Zhongxiang Dai, Yizhou Chen, Haibin Yu, Bryan Kian Hsiang Low, Patrick Jaillet</i> | |
| Individual Fairness in Feature-Based Pricing for Monopoly Markets..... | 486 |
| <i>Shantanu Das, Swapnil Dhamal, Ganesh Ghalme, Shweta Jain, Sujit Gujar</i> | |
| Faster Non-Convex Federated Learning Via Global and Local Momentum..... | 496 |
| <i>Rudrajit Das, Anish Acharya, Abolfazl Hashemi, Sujay Sanghavi, Inderjit S. Dhillon, Ufuk Topcu</i> | |
| Multi-Objective Bayesian Optimization Over High-dimensional Search Spaces..... | 507 |
| <i>Samuel Daulton, David Eriksson, Maximilian Balandat, Eytan Bakshy</i> | |
| Bayesian Structure Learning with Generative Flow Networks..... | 518 |
| <i>Tristan Deleu, António Góis, Chris Emezue, Mansi Rankawat, Simon Lacoste-Julien, Stefan Bauer, Yoshua Bengio</i> | |
| Bayesian Spillover Graphs for Dynamic Networks..... | 529 |
| <i>Grace Deng, David S. Matteson</i> | |
| Multiclass Classification for Hawkes Processes..... | 539 |
| <i>Christophe Denis, Charlotte Dion-Blanc, Laure Sansonnet</i> | |
| Quantification of Credal Uncertainty in Machine Learning: A Critical Analysis and Empirical Comparison..... | 548 |
| <i>Eyke Hüllermeier, Sébastien Destercke, Mohammad Hossein Shaker</i> | |

| | |
|---|-----|
| Balancing Adaptability and Non-Exploitability in Repeated Games | 559 |
| <i>Anthony Digiovanni, Ambuj Tewari</i> | |
| Variational- And Metric-based Deep Latent Space for Out-of-distribution Detection..... | 569 |
| <i>Or Dinari, Oren Freifeld</i> | |
| Revisiting DP-Means: Fast Scalable Algorithms Via Parallelism and Delayed Cluster Creation..... | 579 |
| <i>Or Dinari, Oren Freifeld</i> | |
| X-MEN: Guaranteed XOR-maximum Entropy Constrained Inverse Reinforcement Learning | 589 |
| <i>Fan Ding, Yexiang Xue</i> | |
| Improving Sign-Random-projection Via Count Sketch | 599 |
| <i>Punit Pankaj Dubey, Bhisham Dev Verma, Rameshwar Pratap, Keegan Kang</i> | |
| ResIST: Layer-Wise Decomposition of ResNets for Distributed Training | 610 |
| <i>Chen Dun, Cameron R. Wolfe, Christopher M. Jermaine, Anastasios Kyrillidis</i> | |
| Learning Explainable Templated Graphical Models | 621 |
| <i>Varun Embar, Sriram Srinivasa, Lise Getoor</i> | |
| SENTINEL: Taming Uncertainty with Ensemble Based Distributional Reinforcement Learning | 631 |
| <i>Hannes Eriksson, Debabrota Basu, Mina Alibeigi, Christos Dimitrakakis</i> | |
| Temporal Abstractions-Augmented Temporally Contrastive Learning: An Alternative to the Laplacian in RL | 641 |
| <i>Akram Erraqabi, Marlos C. Machado, Mingde Zhao, Sainbayar Sukhbaatar, Alessandro Lazaric, Denoyer Ludovic, Yoshua Bengio</i> | |
| Implicit Kernel Meta-Learning Using Kernel Integral Forms | 652 |
| <i>John Isak Texas Falk, Carlo Cilibert, Massimiliano Pontil</i> | |
| Self-Distribution Distillation: Efficient Uncertainty Estimation | 663 |
| <i>Yassir Fathullah, Mark J. F. Gales</i> | |
| Sequential Algorithmic Modification with Test Data Reuse | 674 |
| <i>Jean Feng, Gene Pennllo, Nicholas Petrick, Berkman Sahiner, Romain Pirracchio, Alexej Gossmann</i> | |
| Estimating Transfer Entropy Under Long Ranged Dependencies..... | 685 |
| <i>Sahil Garg, Umang Gupta, Yu Chen, Syamantak Datta Gupta, Yeshaya Adler, Anderson Schneider, Yuriy Nevmyvaka</i> | |
| Mitigating Statistical Bias Within Differentially Private Synthetic Data | 696 |
| <i>Sahra Ghalebikesabi, Harry Wilde, Jack Jewson, Arnaud Doucet, Sebastian Vollmer, Chris Holmes</i> | |
| Neural-Progressive Hedging: Enforcing Constraints in Reinforcement Learning with Stochastic Programming | 707 |
| <i>Supriyo Ghosh, Laura Wynter, Shiau Hong Lim, Duc Thien Nguyen</i> | |
| Do Bayesian Variational Autoencoders Know What They Don't Know? | 718 |
| <i>Misha Glazunov, Apostolis Zarras</i> | |
| Robust Expected Information Gain for Optimal Bayesian Experimental Design Using Ambiguity Sets | 728 |
| <i>Jinwoo Go, Tobin Isaac</i> | |

| | |
|--|-----|
| Efficient and Transferable Adversarial Examples from Bayesian Neural Networks..... | 738 |
| <i>Martin Gubri, Maxime Cordy, Mike Papadakis, Yves Le Traon, Koushik Sen</i> | |
| Learning a Neural Pareto Manifold Extractor with Constraints | 749 |
| <i>Soumyajit Gupta, Gurpreet Singh, Raghu Bollapragada, \matthew Lease</i> | |
| Modeling Extremes with D -Max-decreasing Neural Networks | 759 |
| <i>Ali Hasan, Khalil Elkhail, Yuting Ng, João M. Pereira, Sina Farsiu, Jose Blanchet, Vahid Tarokh</i> | |
| Generalizing Off-Policy Learning Under Sample Selection Bias | 769 |
| <i>Tobias Hatt, Daniel Tschernutter, Stefan Feuerriegel</i> | |
| Reinforcement Learning in Many-Agent Settings Under Partial Observability | 780 |
| <i>Keyang He, Prashant Doshi, Bikramjit Banerjee</i> | |
| Variational Multiple Shooting for Bayesian ODEs with Gaussian Processes | 790 |
| <i>Pashupati Hegde, Çagatay Yildiz, Harri Lähdesmäki, Samuel Kaski, Markus Heinonen</i> | |
| Learning Sparse Representations of Preferences Within Choquet Expected Utility Theory | 800 |
| <i>Margot Herin, Patrice Perny, Nataliya Sokolovska</i> | |

VOLUME 2

| | |
|---|-----|
| Quadratic Metric Elicitation for Fairness and Beyond | 811 |
| <i>Gaurush Hiranandani, Jatin Mathur, Harikrishna Narasimhan, Oluwasanmi Koyejo</i> | |
| Fast Predictive Uncertainty for Classification with Bayesian Deep Networks..... | 822 |
| <i>Marius Hobbhahn, Agustinus Kristiadi, Philipp Hennig</i> | |
| CIGMO: Categorical Invariant Representations in a Deep Generative Framework..... | 833 |
| <i>Haruo Hosoya</i> | |
| Near-Optimal Thompson Sampling-based Algorithms for Differentially Private Stochastic Bandits | 844 |
| <i>Bingshan Hu, Nidhi Hegde</i> | |
| Uncertainty-Aware Pseudo-labeling for Quantum Calculations | 853 |
| <i>Kexin Huang, Vishnu Sresht, Brajesh Rai, Mykola Bordyuh</i> | |
| A Mutually Exciting Latent Space Hawkes Process Model for Continuous-Time Networks | 863 |
| <i>Zhipeng Huang, Hadeel Soliman, Subhadeep Paul, Kevin S. Xu</i> | |
| Binary Independent Component Analysis: A Non-Stationarity-based Approach | 874 |
| <i>Antti Hyttinen, Vitória Barin Pacela, Aapo Hyvärinen</i> | |
| Balancing Utility and Scalability in Metric Differential Privacy | 885 |
| <i>Jacob Imola, Shiva Kasiviswanathan, Stephen White, Abhinav Aggarwal, Nathanael Teissier</i> | |
| Towards Painless Policy Optimization for Constrained MDPs | 895 |
| <i>Arushi Jain, Sharan Vaswani, Reza Babanezhad, Csaba Szepesvári, Doina Precup</i> | |
| Fedvarp: Tackling the Variance Due to Partial Client Participation in Federated Learning | 906 |
| <i>Divyansh Jhunjhunwala, Pranay Sharma, Aushim Nagarkatti, Gauri Joshi</i> | |
| Orthogonal Gromov-Wasserstein Discrepancy with Efficient Lower Bound | 917 |
| <i>Hongwei Jin, Zishun Yu, Xinhua Zhang</i> | |

| | |
|---|------|
| If You've Trained One You've Trained Them All: Inter-Architecture Similarity Increases with Robustness..... | 928 |
| <i>Haydn T. Jones, Jacob M. Springer, Garrett T. Kenyon, Juston S. Moore</i> | |
| Decision-Theoretic Planning with Communication in Open Multiagent Systems | 938 |
| <i>Anirudh Kakarlapudi, Gayathri Anil, Adam Eck, Prashant Doshi, Leen-Kiat Soh</i> | |
| Optimal Control of Partially Observable Markov Decision Processes with Finite Linear Temporal Logic Constraints | 949 |
| <i>Krishna C. Kalagarla, Kartik Dhruva, Dongming Shen, Rahul Jain, Ashutosh Nayyar, Pierluigi Nuzzo</i> | |
| Test for Non-Negligible Adverse Shifts | 959 |
| <i>Vathy M Kamulete</i> | |
| Improved Feature Importance Computation for Tree Models Based on the Banzhaf Value..... | 969 |
| <i>Adam Karczmarz, Tomasz Michalak, Anish Mukherjee, Piotr Sankowski, Piotr Wygocki</i> | |
| Dynamic Relocation in Ridesharing Via Fixpoint Construction | 980 |
| <i>Ian A. Kash, Zhongkai Wen, Lenore D. Zuck</i> | |
| Restless and Uncertain: Robust Policies for Restless Bandits Via Deep Multi-Agent Reinforcement Learning | 990 |
| <i>Jackson A. Killian, Lily Xu, Arpita Biswas, Milind Tambe</i> | |
| Combinatorial Bayesian Optimization with Random Mapping Functions to Convex Polytopes | 1001 |
| <i>Jungtaek Kim, Seungjin Choi, Minsu Cho</i> | |
| On the Effectiveness of Adversarial Training Against Common Corruptions | 1012 |
| <i>Klim Kireev, Maksym Andriushchenko, Nicolas Flammarion</i> | |
| Revisiting the General Identifiability Problem..... | 1022 |
| <i>Yaroslav Kivva, Ehsan Mokhtarian, Jalal Etesami, Negar Kiyavash</i> | |
| Hitting Times for Continuous-Time imprecise-Markov Chains..... | 1031 |
| <i>Thomas Krak</i> | |
| Calibrated Ensembles Can Mitigate Accuracy Tradeoffs Under Distribution Shift..... | 1041 |
| <i>Ananya Kumar, Tengyu Ma, Percy Liang, Aditi Raghunathan</i> | |
| Greedy Relaxations of the Sparsest Permutation Algorithm | 1052 |
| <i>Wai-Yin Lam, Bryan Andrews, Joseph Ramsey</i> | |
| Interpolating Between Sampling and Variational Inference with Infinite Stochastic Mixtures | 1063 |
| <i>Richard D. Lange, Ari S. Benjamin, Ralf M. Haefner, Xaq Pitkow</i> | |
| Systematized Event-Aware Learning for Multi-object Tracking | 1074 |
| <i>Hyemin Lee, Daijin Kim</i> | |
| Fixing the Bethe Approximation: How Structural Modifications in a Graph Improve Belief Propagation..... | 1085 |
| <i>Harald Leisenberger, Franz Pernkopf, Christian Knoll</i> | |
| Recursive Monte Carlo and Variational Inference with Auxiliary Variables..... | 1096 |
| <i>Alexander K. Lew, Marco Cusumano-Towner, Vikash K. Mansinghka</i> | |

| | |
|---|------|
| Solving Structured Hierarchical Games Using Differential Backward Induction | 1107 |
| <i>Zun Li, Feiran Jia, Aditya Mate, Shahin Jabbari, Mithun Chakraborty, Milind Tambe, Yevgeniy Vorobeychik</i> | |
| PDQ-Net: Deep Probabilistic Dual Quaternion Network for Absolute Pose Regression on SE(3)..... | 1118 |
| <i>Wenjie Li, Wasif Naeem, Jia Liu, Dequan Zheng, Wei Hao, Lijun Chen</i> | |
| Accelerating Training of Batch Normalization: A Manifold Perspective | 1128 |
| <i>Mingyang Yi</i> | |
| Deep Dirichlet Process Mixture Models..... | 1138 |
| <i>Naiqi Li, Wenjie Li, Yong Jiang, Shu-Tao Xia</i> | |
| Proportional Allocation of Indivisible Resources Under Ordinal and Uncertain Preferences..... | 1148 |
| <i>Zihao Li, Xiaohui Bei, Zhenzhen Yan</i> | |
| Efficient Resource Allocation with Fairness Constraints in Restless Multi-Armed Bandits..... | 1158 |
| <i>Dexun. Li, Pradeep Varakantham</i> | |
| A Label Efficient Two-Sample Test..... | 1168 |
| <i>Weizhi Li, Gautam Dasarathy, Karthikeyan Natesan Ramamurthy, Visar Berisha</i> | |
| ℓ_∞ -Bounds of the MLE in the BTL Model Under General Comparison Graphs | 1178 |
| <i>Wanshan Li, Shamindra Shrotriya, Alessandro Rinaldo</i> | |
| AdaCat: Adaptive Categorical Discretization for Autoregressive Models | 1188 |
| <i>Qiyang Li, Ajay Jain, Pieter Abbeel</i> | |
| Laplace Approximated Gaussian Process State-Space Models | 1199 |
| <i>Jakob Lindinger, Barbara Rakitsch, Christoph Lippert</i> | |
| Dimension Reduction for High-Dimensional Small Counts with KL Divergence..... | 1210 |
| <i>Yurong Ling, Jing-Hao Xue</i> | |
| Federated Online Clustering of Bandits | 1221 |
| <i>Xutong Liu, Haoru Zhao, Tong Yu, Shuai Li, John C. S. Lui</i> | |
| PathFlow: A Normalizing Flow Generator that Finds Transition Paths..... | 1232 |
| <i>Tianyi Liu, Weihao Gao, Zhirui Wang, Chong Wang</i> | |
| SASH: Efficient Secure Aggregation Based on SHPRG for Federated Learning | 1243 |
| <i>Zizhen Liu, Si Chen, Jing Ye, Junfeng Fan, Huawei Li, Xiaowei Li</i> | |
| Offline Policy Optimization with Eligible Actions | 1253 |
| <i>Yao Liu, Yannis Flet-Berliac, Emma Brunskill</i> | |
| Data Poisoning Attacks on Off-Policy Policy Evaluation Methods | 1264 |
| <i>Elita Lobo, Harvineet Singh, Marek Petrik, Cynthia Rudin, Himabindu Lakkaraju</i> | |
| Nonparametric Exponential Family Graph Embeddings for Multiple Representation Learning | 1275 |
| <i>Chien Lu, Jaakko Peltonen, Timo Nummenmaa, Jyrki Nummenmaa</i> | |
| Local Calibration: Metrics and Recalibration..... | 1286 |
| <i>Rachel Luo, Aadyot Bhatnagar, Yu Bai, Shengjia Zhao, Huan Wang, Caiming Xiong, Silvio Savarese, Stefano Ermon, Edward Schmerling, Marco Pavone</i> | |

| | |
|--|------|
| Data Sampling Affects the Complexity of Online SGD Over Dependent Data | 1296 |
| <i>Shaocong Ma, Ziyi Chen, Yi Zhou, Kaiyi Ji, Yingbin Liang</i> | |
| Low-Precision Arithmetic for Fast Gaussian Processes | 1306 |
| <i>Wesley J. Maddox, Andres Potapczynski, Andrew Gordon Wilson</i> | |
| Perturbation Type Categorization for Multiple Adversarial Perturbation Robustness | 1317 |
| <i>Pratyush Maini, Xinyun Chen, Bo Li, Dawn Song</i> | |
| A Causal Bandit Approach to Learning Good Atomic Interventions in Presence of Unobserved Confounders | 1328 |
| <i>Aurghya Maiti, Vineet Nair, Gaurav Sinha</i> | |
| Case-Based Off-policy Evaluation Using Prototype Learning | 1339 |
| <i>Anton Matsson, Fredrik D. Johansson</i> | |
| Multistate Analysis with Infinite Mixtures of Markov Chains | 1350 |
| <i>Lucas Maystre, Tiffany Wu, Roberto Sanchis-Ojeda, Tony Jebara</i> | |
| Forget-Me-not! Contrastive Critics for Mitigating Posterior Collapse | 1360 |
| <i>Sachit Menon, David Blei, Carl Vondrick</i> | |
| Can Mean Field Control (mfc) Approximate Cooperative Multi Agent Reinforcement Learning (marl) with Non-Uniform Interaction? | 1371 |
| <i>Washim Uddin Mondal, Vaneet Aggarwal, Satish V. Ukkusuri</i> | |
| Monotonicity Regularization: Improved Penalties and Novel Applications to Disentangled Representation Learning and Robust Classification | 1381 |
| <i>João Monteiro, Mohamed Osama Ahmed, Hosein Hajimirsadeghi, Greg Mori</i> | |
| Set-Valued Prediction in Hierarchical Classification with Constrained Representation Complexity | 1392 |
| <i>Thomas Mortier, Eyke Hüllermeier, Krzysztof Dembczynski, Willem Waegeman</i> | |
| Safety Aware Changepoint Detection for Piecewise I.i.d. Bandits | 1402 |
| <i>Subhojyoti Mukherjee</i> | |
| ReVar: Strengthening Policy Evaluation Via Reduced Variance Sampling | 1413 |
| <i>Subhojyoti Mukherjee, Josiah P. Hanna, Robert D Nowak</i> | |
| Probabilistic Surrogate Networks for Simulators with Unbounded Randomness | 1423 |
| <i>Andreas Munk, Berend Zwartsenberg, Adam Scibior, Atılım Günes G. Baydin, Andrew Stewart, Goran Fernlund, Anoush Poursartip, Frank Wood</i> | |
| Data Augmentation in Bayesian Neural Networks and the Cold Posterior Effect | 1434 |
| <i>Seth Nabarro, Stoil Ganev, Adrià Garriga-Alonso, Vincent Fortuin, Mark Van Der Wilk, Laurence Aitchison</i> | |
| Semiparametric Causal Sufficient Dimension Reduction of Multidimensional Treatments | 1445 |
| <i>Razieh Nabi, Todd McNutt, Ilya Shpitser</i> | |
| Partially Adaptive Regularized Multiple Regression Analysis for Estimating Linear Causal Effects | 1456 |
| <i>Hisayoshi Nanmo, Manabu Kuroki</i> | |
| Efficient Learning of Sparse and Decomposable PDEs Using Random Projection | 1466 |
| <i>Md Nasim, Xinghang Zhang, Anter El-Azab, Yexiang Xue</i> | |

| | |
|--|------|
| Linearizing Contextual Bandits with Latent State Dynamics..... | 1477 |
| <i>Elliot Nelson, Debarun Bhattacharjya, Tian Gao, Miao Liu, Djallel Bouneffouf, Pascal Poupart</i> | |
| CounterGAN: Generating Counterfactuals for Real-Time Recourse and Interpretability Using Residual GANs..... | 1488 |
| <i>Daniel Nemirovsky, Nicolas Thiebaut, Ye Xu, Abhishek Gupta</i> | |
| Robust Bayesian Recourse | 1498 |
| <i>Tuan-Duy H. Nguyen, Ngoc Bui, Duy Nguyen, Man-Chung Yue, Viet Anh Nguyen</i> | |
| Efficient and Accurate Top-K Recovery from Choice Data..... | 1509 |
| <i>Duc Nguyen</i> | |
| Cycle Class Consistency with Distributional Optimal Transport and Knowledge Distillation for Unsupervised Domain Adaptation..... | 1519 |
| <i>Tuan Nguyen, Van Nguyen, Trung Le, He Zhao, Quan Hung Tran, Dinh Phung</i> | |
| Ordinal Causal Discovery..... | 1530 |
| <i>Yang Ni, Bani Mallick</i> | |
| An Explore-Then-commit Algorithm for Submodular Maximization Under Full-bandit Feedback..... | 1541 |
| <i>Guanyu Nie, Mridul Agarwal, Abhishek Kumar Umrawal, Vaneet Aggarwal, Christopher John Quinn</i> | |
| Evaluating High-Order Predictive Distributions in Deep Learning..... | 1552 |
| <i>Ian Osband, Zheng Wen, Seyed Mohammad Asghari, Vikranth Dwaracherla, Xiuyuan Lu, Benjamin Van Roy</i> | |
| Understanding and Mitigating the Limitations of Prioritized Experience Replay..... | 1561 |
| <i>Yangchen Pan, Jincheng Mei, Amir-Massoud Farahmand, Martha White, Hengshuai Yao, Mohsen Rohani, Jun Luo</i> | |
| Robust Learning of Tractable Probabilistic Models..... | 1572 |
| <i>Rohith Peddi, Tahrira Rahman, Vibhav Gogate</i> | |
| Attribution of Predictive Uncertainties in Classification Models..... | 1582 |
| <i>Iker Perez, Piotr Skalski, Alec Barns-Graham, Jason Wong, David Sutton</i> | |
| Learning Large Bayesian Networks with Expert Constraints..... | 1592 |
| <i>Vaidyanathan Peruvemba Ramaswamy, Stefan Szeider</i> | |
| AND/OR Branch-And-bound for Computational Protein Design Optimizing K^* | 1602 |
| <i>Bobak Pezeshki, Radu Marinescu, Alexander Ihler, Rina Dechter</i> | |
| Identifiability of Sparse Causal Effects Using Instrumental Variables | 1613 |
| <i>Niklas Pfister, Jonas Peters</i> | |

VOLUME 3

| | |
|--|------|
| Bayesian Quantile and Expectile Optimisation | 1623 |
| <i>Victor Picheny, Henry Moss, Léonard Torossian, Nicolas Durrande</i> | |
| Using Hierarchies to Efficiently Combine Evidence with Dempster’s Rule of Combination..... | 1634 |
| <i>Daira Pinto Prieto, Ronald De Haan</i> | |

| | |
|--|------|
| Voronoi Density Estimator for High-Dimensional Data: Computation, Compactification and Convergence..... | 1644 |
| <i>Vladislav Polianskii, Giovanni Luca Marchetti, Alexander Kravberg, Anastasiia Varava, Florian T. Pokorny, Danica Kragic</i> | |
| Clustering a Union of Linear Subspaces Via Matrix Factorization and Innovation Search | 1654 |
| <i>Mostafa Rahmani</i> | |
| Learning in Markov Games: Can We Exploit a General-Sum Opponent? | 1665 |
| <i>Giorgia Ramponi, Marcello Restelli</i> | |
| Expectation Programming: Adapting Probabilistic Programming Systems to Estimate Expectations Efficiently | 1676 |
| <i>Tim Reichelt, Adam Golinski, Luke Ong, Tom Rainforth</i> | |
| A Free Lunch from the Noise: Provable and Practical Exploration for Representation Learning..... | 1686 |
| <i>Tongzheng Ren, Tianjun Zhang, Csaba Szepesvári, Bo Dai</i> | |
| Quantum Perceptron Revisited: Computational-Statistical Tradeoffs..... | 1697 |
| <i>Mathieu Roget, Giuseppe Di Molfetta, Hachem Kadri</i> | |
| Resolving Label Uncertainty with Implicit Posterior Models | 1707 |
| <i>Esther Rolf, Nikolay Malkin, Alexandros Graikos, Ana Jojic, Caleb Robinson, Nebojsa Jojic</i> | |
| Feature Learning and Random Features in Standard Finite-Width Convolutional Neural Networks: An Empirical Study | 1718 |
| <i>Maxim Samarin, Volker Roth, David Belius</i> | |
| Robust Identifiability in Linear Structural Equation Models of Causal Inference | 1728 |
| <i>Karthik A. Sankararaman, Anand Louis, Navin Goyal</i> | |
| How Unfair is Private Learning?..... | 1738 |
| <i>Amartya Sanyal, Yaxi Hu, Fanny Yang</i> | |
| Probabilistic Spatial Transformer Networks | 1749 |
| <i>Pola Schwöbel, Frederik Rahbæk Warburg, Martin Jørgensen, Kristoffer Hougaard Madsen, Søren Hauberg</i> | |
| Learning Functions on Multiple Sets Using Multi-Set Transformers..... | 1760 |
| <i>Kira A. Selby, Ahmad Rashid, Ivan Kobyzev, Mehdi Rezagholizadeh, Pascal Poupart</i> | |
| SymNet 2.0: Effectively Handling Non-Fluents and Actions in Generalized Neural Policies for RDDDL Relational MDPs..... | 1771 |
| <i>Vishal Sharma, Daman Arora, Florian Geißer, Mausam, Parag Singla</i> | |
| Reframed GES with a Neural Conditional Dependence Measure | 1782 |
| <i>Xinwei Shen, Shengyu Zhu, Jiji Zhang, Shoubo Hu, Zhitang Chen</i> | |
| Conditional Simulation Using Diffusion Schrödinger Bridges | 1792 |
| <i>Yuyang Shi, Valentin De Bortoli, George Deligiannidis, Arnaud Doucet</i> | |
| Neural Ensemble Search Via Bayesian Sampling | 1803 |
| <i>Yao Shu, Yizhou Chen, Zhongxiang Dai, Bryan Kian Hsiang Low</i> | |
| Shifted Compression Framework: Generalizations and Improvements | 1813 |
| <i>Egor Shulgin, Peter Richtárik</i> | |

| | |
|---|------|
| PAC-Bayesian Domain Adaptation Bounds for Multiclass Learners..... | 1824 |
| <i>Anthony Sicilia, Katherine Atwell, Malihe Alikhani, Seong Jae Hwang</i> | |
| VQ-Flows: Vector Quantized Local Normalizing Flows | 1835 |
| <i>Sahil Sidheekh, Chris B. Dock, Tushar Jain, Radu Balan, Maneesh K. Singh</i> | |
| Enhanced Adaptive Optics Control with Image to Image Translation | 1846 |
| <i>Jeffrey Smith, Jesse Cranney, Charles Gretton, Damien Gratadour</i> | |
| Fast Inference and Transfer of Compositional Task Structures for Few-Shot Task Generalization..... | 1857 |
| <i>Sungryull Sohn, Hyunjae Woo, Jongwook Choi, Lyubing Qiang, Izzeddin Gur, Aleksandra Faust, Honglak Lee</i> | |
| Mutual Information Based Bayesian Graph Neural Network for Few-Shot Learning | 1866 |
| <i>Kaiyi Song, Kun Yue, Liang Duan, Mingze Yang, Angsheng Li</i> | |
| SMT-Based Weighted Model Integration with Structure Awareness | 1876 |
| <i>Giuseppe Spallitta, Gabriele Masina, Paolo Morettin, Andrea Passerini, Roberto Sebastiani</i> | |
| A Robustness Test for Estimating Total Effects with Covariate Adjustment..... | 1886 |
| <i>Zehao Su, Leonard Henckel</i> | |
| Simplified and Unified Analysis of Various Learning Problems by Reduction to Multiple-Instance Learning | 1896 |
| <i>Daiki Suehiro, Eiji Takimoto</i> | |
| Marginal MAP Estimation for Inverse RL Under Occlusion with Observer Noise | 1907 |
| <i>Prasanth Sengadu Suresh, Prashant Doshi</i> | |
| High-Probability Bounds for Robust Stochastic Frank-Wolfe Algorithm..... | 1917 |
| <i>Tongyi Tang, Krishna Balasubramanian, Thomas Chun Man Lee</i> | |
| Contrastive Latent Variable Models for Neural Text Generation..... | 1928 |
| <i>Zhiyang Teng, Chenhua Chen, Yan Zhang, Yue Zhang</i> | |
| Semi-Supervised Novelty Detection Using Ensembles with Regularized Disagreement..... | 1939 |
| <i>Alexandru Tifrea, Eric Stavarache, Fanny Yang</i> | |
| Efficient Inference for Dynamic Topic Modeling with Large Vocabularies..... | 1950 |
| <i>Federico Tomasi, Mounia Lalmas, Zhenwen Dai</i> | |
| Learning Linear non-Gaussian Polytree Models | 1960 |
| <i>Daniele Tramontano, Anthea Monod, Mathias Drton</i> | |
| Multi-Source Domain Adaptation Via Weighted Joint Distributions Optimal Transport | 1970 |
| <i>Rosanna Turrisi, Rémi Flamary, Alain Rakotomamonjy, Massimiliano Pontil</i> | |
| Towards Unsupervised Open World Semantic Segmentation..... | 1981 |
| <i>Svenja Uhlemeyer, Matthias Rottmann, Hanno Gottschalk</i> | |
| Learning Invariant Weights in Neural Networks..... | 1992 |
| <i>Tycho F. A. Van Der Ouderaa, Mark Van Der Wilk</i> | |
| Causal Forecasting: Generalization Bounds for Autoregressive Models | 2002 |
| <i>Leena Chennuru Vankadara, Philipp Michael Faller, Michaela Hardt, Lenon Minorics, Debarghya Ghoshdastidar, Dominik Janzing</i> | |

| | |
|---|------|
| Intervention Target Estimation in the Presence of Latent Variables | 2013 |
| <i>Burak Varici, Karthikeyan Shanmugam, Prasanna Sattigeri, Ali Tajer</i> | |
| Bayesian Federated Estimation of Causal Effects from Observational Data..... | 2024 |
| <i>Thanh Vinh Vo, Young Lee, Trong Nghia Hoang, Tze-Yun Leong</i> | |
| Bias Aware Probabilistic Boolean Matrix Factorization | 2035 |
| <i>Changlin Wan, Pengtao Dang, Tong Zhao, Yong Zang, Chi Zhang, Sha Cao</i> | |
| Meta-Learning Without Data Via Wasserstein Distributionally-robust Model Fusion | 2045 |
| <i>Zhenyi Wang, Xiaoyang Wang, Li Shen, Qiuling Suo, Kaiqiang Song, Dong Yu, Yan Shen, Mingchen Gao</i> | |
| Detecting Textual Adversarial Examples Through Randomized Substitution and Vote..... | 2056 |
| <i>Xiaosen Wang, Xiong Yifeng, Kun He</i> | |
| ST-MAML : A Stochastic-task Based Method for Task-heterogeneous Meta-learning | 2066 |
| <i>Zhe Wang, Jake Grigsby, Arshdeep Sekhon, Yanjun Qi</i> | |
| Toward Learning Human-Aligned Cross-domain Robust Models by Countering Misaligned Features | 2075 |
| <i>Haohan Wang, Zeyi Huang, Hanlin Zhang, Yong Jae Lee, Eric P. Xing</i> | |
| Generalized Bayesian Quadrature with Spectral Kernels..... | 2085 |
| <i>Houston Warren, Rafael Oliveira, Fabio Ramos</i> | |
| Causal Discovery Under a Confounder Blanket..... | 2096 |
| <i>David S. Watson, Ricardo Silva</i> | |
| A New Constructive Criterion for Markov Equivalence of MAGs..... | 2107 |
| <i>Marcel Wienöbst, Max Bannach, Maciej Liskiewicz</i> | |
| Residual Bootstrap Exploration for Stochastic Linear Bandit..... | 2117 |
| <i>Shuang Wu, Chi-Hua Wang, Yuantong Li, Guang Cheng</i> | |
| Differentially Private Multi-Party Data Release for Linear Regression..... | 2128 |
| <i>Ruihan Wu, Xin Yang, Yuanshun Yao, Jiankai Sun, Tianyi Liu, Q. Kilian Weinberger, Chong Wang</i> | |
| Partial Likelihood Thompson Sampling..... | 2138 |
| <i>Han Wu, Stefan Wager</i> | |
| Fine-Grained Matching with Multi-perspective Similarity Modeling for Cross-modal Retrieval | 2148 |
| <i>Xiumin Xie, Chuanwen Hou, Zhixin Li</i> | |
| Deterministic Policy Gradient: Convergence Analysis | 2159 |
| <i>Huaqing. Xiong, Tengyu Xu, Lin Zhao, Yingbin Liang, Wei Zhang</i> | |
| Finite-Horizon Equilibria for Neuro-symbolic Concurrent Stochastic Games..... | 2170 |
| <i>Rui Yan, Gabriel Santos, Xiaoming Duan, David Parker, Marta Kwiatkowska</i> | |
| Addressing Token Uniformity in Transformers Via Singular Value Transformation | 2181 |
| <i>Hanqi Yan, Lin Gui, Wenjie Li, Yulan He</i> | |
| Differentially Private SGDA for Minimax Problems | 2192 |
| <i>Zhenhuan Yang, Shu Hu, Yunwen Lei, Kush R Vashney, Siwei Lyu, Yiming Ying</i> | |

| | |
|---|------|
| Self-Supervised Representations for Multi-view Reinforcement Learning..... | 2203 |
| <i>Huanhuan Yang, Dianxi Shi, Guojun Xie, Yingxuan Peng, Yi Zhang, Yantai Yang, Shaowu Yang</i> | |
| Robust Textual Embedding Against Word-Level Adversarial Attacks..... | 2214 |
| <i>Yichen Yang, Xiaosen Wang, Kun He</i> | |
| CoSPA: An Improved Masked Language Model with Copy Mechanism for Chinese Spelling Correction..... | 2225 |
| <i>Shoujian Yang, Lian Yu</i> | |
| Noisy ℓ^0 -Sparse Subspace Clustering on Dimensionality Reduced Data | 2235 |
| <i>Yingzhen Yang, Ping Li</i> | |
| Pareto Navigation Gradient Descent: A First-Order Algorithm for Optimization in Pareto Set | 2246 |
| <i>Mao Ye, Qiang Liu</i> | |
| Future Gradient Descent for Adapting the Temporal Shifting Data Distribution in Online Recommendation Systems..... | 2256 |
| <i>Mao Ye, Ruichen Jiang, Haoxiang Wang, Dhruv Choudhary, Xiaocong Du, Bhargav Bhushanam, Aryan Mokhtari, Arun Kejariwal, Qiang Liu</i> | |
| Superposing Many Tickets into One: A Performance Booster for Sparse Neural Network Training | 2267 |
| <i>Lu Yin, Vlado Menkovski, Meng Fang, Tianjin Huang, Yulong Pei, Mykola Pechenizkiy</i> | |
| Offline Stochastic Shortest Path: Learning, Evaluation and Towards Optimality..... | 2278 |
| <i>Ming Yin, Wenjing Chen, Mengdi Wang, Yu-Xiang Wang</i> | |
| Active Learning with Label Comparisons..... | 2289 |
| <i>Gal Yona, Shay Moran, Gal Elidan, Amir Globerson</i> | |
| Cross-Domain Adaptive Transfer Reinforcement Learning Based on State-action Correspondence | 2299 |
| <i>Heng You, Tianpei Yang, Yan Zheng, Jianye Hao, E. Taylor Matthew</i> | |
| Learning Binary Multi-Scale Games on Networks..... | 2310 |
| <i>Sixie Yu, P. Jeffrey Brantingham, Matthew Valasik, Yevgeniy Vorobeychik</i> | |
| Predictive Whittle Networks for Time Series..... | 2320 |
| <i>Zhongjie Yu, Fabrizio Ventola, Nils Thoma, Devendra Singh Dhami, Martin Mundt, Kristian Kersting</i> | |
| Principle of Relevant Information for Graph Sparsification..... | 2331 |
| <i>Shujian Yu, Francesco Alesiani, Wenzhe Yin, Robert Jenssen, Jose C. Principe</i> | |
| Asymptotic Optimality for Active Learning Processes | 2342 |
| <i>Xueying Zhan, Yaowei Wang, Antoni b. Chan</i> | |
| Distributed Adversarial Training to Robustify Deep Neural Networks at Scale | 2353 |
| <i>Gaoyuan Zhang, Songtao Lu, Yihua Zhang, Xiangyi Chen, Pin-Yu Chen, Quanfu Fan, Lee Martie, Lior Horesh, Mingyi Hong, Sijia Liu</i> | |
| Stability of SGD: Tightness Analysis and Improved Bounds | 2364 |
| <i>Yikai Zhang, Wenjia Zhang, Sammy Bald, Vamsi Pingali, Chao Chen, Mayank Goswami</i> | |
| Research on Video Adversarial Attack with Long Living Cycle | 2374 |
| <i>Zeyu Zhao, Ke Xu, Xinghao Jiang, Tanfeng Sun</i> | |

| | |
|---|------|
| Causal Discovery with Heterogeneous Observational Data | 2383 |
| <i>Fangting Zhou, Kejun He, Yang Ni</i> | |
| Convergence Analysis of Linear Coupling with Inexact Proximal Operator | 2394 |
| <i>Qiang Zhou, Sinno Jialin Pan</i> | |
| Information Design for Multiple Independent and Self-Interested Defenders: Work Less, Pay off More | 2404 |
| <i>Chenghan Zhou, Andrew Spivey, Haifeng Xu, Thanh Hong Nguyen</i> | |
| Causal Inference with Treatment Measurement Error: A Nonparametric Instrumental Variable Approach | 2414 |
| <i>Yuchen Zhu, Limor Gultchin, Arthur Gretton, Matt J. Kusner, Ricardo Silva</i> | |

Author Index