

Neural Information Processing Systems

# Advances in Neural Information Processing Systems 24

25th Annual Conference on Neural Information  
Processing Systems 2011

December 12-15, 2011  
Granada, Spain

Volume 1 of 3

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

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571  
[www.proceedings.com](http://www.proceedings.com)

**ISBN: 978-1-61839-599-3**

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

**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© (2011) by Neural Information Processing Systems  
All rights reserved.

Printed by Curran Associates, Inc. (2012)

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

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

Phone: (858) 453-4100  
Fax: (858) 453-8534

info@nips.cc

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

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571 USA  
Phone: 845-758-0400  
Fax: 845-758-2634  
Email: curran@proceedings.com  
Web: www.proceedings.com

# Contents

Contents .....	iii
Preface .....	xxv
Donors .....	xxviii
NIPS foundation .....	xxix
Committees .....	xxx
Reviewers .....	xxxiii
<b>Maximum Margin Multi-Instance Learning</b> , HUA WANG, Univ. of Texas at Arlington, HENG HUANG, University of Texas Arlington, FARHAD KAMANGAR, University of Texas at Arlington, FEIPING NIE, University of Texas Arlington, and CHRIS DING, University of Texas at Arlington .....	1
<b>Shaping Level Sets with Submodular Functions</b> , , and FRANCIS BACH, INRIA - Ecole Normale Superieure .....	10
<b>Nonlinear Inverse Reinforcement Learning with Gaussian Processes</b> , SERGEY LEVINE, Stanford University, ZORAN POPOVIC, University of Washington, and VLADLEN KOLTUN, Stanford University .....	19
<b>Video Annotation and Tracking with Active Learning</b> , CARL VONDRICK, Massachusetts Institute of Technology, and DEVA RAMANAN, .....	28
<b>On U-processes and clustering performance</b> , , and STÉPHAN CLÉMENTÇON, Telecom ParisTech .....	37
<b>Penalty Decomposition Methods for Rank Minimization</b> , YONG ZHANG, and ZHAOSONG LU, Simon Fraser University .....	46
<b>Sparse Manifold Clustering and Embedding</b> , EHSAN ELHAMIFAR, Johns Hopkins University, and RENE VIDAL, .....	55
<b>Unifying Non-Maximum Likelihood Learning Objectives with Minimum KL Contraction</b> , , and SIWEI LYU, University at Albany SUNY ..	64
<b>Image Parsing with Stochastic Scene Grammar</b> , YIBIAO ZHAO, University of California, Los Angeles, and SONG-CHUN ZHU, UCLA .....	73
<b>A Reinforcement Learning Theory for Homeostatic Regulation</b> , MEHDI KERAMATI, Group for Neural Theory, and BORIS GUTKIN, Group for Neural Theory, DEC, ENS, Paris, France .....	82
<b>Learning large-margin halfspaces with more malicious noise</b> , PHIL LONG, Google, and ROCCO SERVEDIO, Columbia University .....	91
<b>On Strategy Stitching in Large Extensive Form Multiplayer Games</b> , RICHARD GIBSON, and DUANE SZAFRON, University of Alberta ....	100

<b>Efficient Inference in Fully Connected CRFs with Gaussian Edge Potentials,</b> PHILIPP KRÄHENBÜHL, and VLADLEN KOLTUN, Stanford University .....	109
<b>Transfer Learning by Borrowing Examples for Multiclass Object Detection,</b> JOSEPH LIM, Massachusetts Institute of Technology, RUSLAN SALAKHUTDINOV, University of Toronto, and ANTONIO TORRALBA, Massachusetts Institute of Technology .....	118
<b>Environmental statistics and the trade-off between model-based and TD learning in humans,</b> DYLAN SIMON, and NATHANIEL DAW, New York University .....	127
<b>Variational Learning for Recurrent Spiking Networks,</b> DANILO REZENDE, EPFL, DAAN WIERSTRA, , and WULFRAM GERSTNER, Ecole Polytechnique Federale de Lausanne .....	136
<b>Multiple Instance Learning on Structured Data,</b> DAN ZHANG, Purdue University, YAN LIU, University of Southern California, LUO SI, JIAN ZHANG, Purdue University, and RICHARD LAWRENCE, IBM TJ Watson Research Center .....	145
<b>Manifold Precip: An Annealing Technique for Diverse Sampling of Manifolds,</b> NITESH SHROFF, PAVAN TURAGA, University of Maryland, and RAMA CHELLAPPA, University of Maryland College Park .....	154
<b>A Global Structural EM Algorithm for a Model of Cancer Progression,</b> ALI TOFIGH, McGill University, ERIK SJHLUND, Stockholm University, MATTIAS HHGLUND, Lund University, and JENS LAGERGREN, KTH .....	163
<b>Action-Gap Phenomenon in Reinforcement Learning,</b> , and AMIR-MASSOUD FARAHMAND, McGill University .....	172
<b>Generalized Lasso based Approximation of Sparse Coding for Visual Recognition,</b> NOBUYUKI MORIOKA, University of New South Wales, and SHIN'ICHI SATOH, NII .....	181
<b>Matrix Completion for Multi-label Image Classification,</b> RICARDO CABRAL, Instituto Superior Tecnico VAT 501507930, FERNANDO DE LA TORRE, Carnegie Mellon University, JOAO COSTEIRA, Instituto Superior Tecnico VAT- 501507930, and ALEXANDRE BERNARDINO, ISR/ Instituto Superior Tecnico .....	190
<b>Multi-View Learning of Word Embeddings via CCA,</b> PARAMVEER DHILLON, DEAN FOSTER, and LYLE UNGAR, University of Pennsylvania .....	199
<b>Global Solution of Fully-Observed Variational Bayesian Matrix Factorization is Column-Wise Independent,</b> SHINICHI NAKAJIMA, Nikon Corporation, MASASHI SUGIYAMA, Tokyo Institute of Technology, and S. DERIN BABACAN, University of Illinois at Urbana-Champaign .....	208
<b>Estimating time-varying input signals and ion channel states from a single voltage trace of a neuron,</b> RYOTA KOBAYASHI, Ritsumeikan University, YASUHIRO TSUBO, RIKEN, PETR LANSKY, Academy of Sciences, and SHIGERU SHINOMOTO, Kyoto University .....	217
<b>Additive Gaussian Processes,</b> DAVID DUVENAUD, University of Cambridge, HANNES NICKISCH, Philips Research, and CARL EDWARD RASMUSSEN, University of Cambridge .....	226

<b>Inferring Interaction Networks using the IBP applied to microRNA Target Prediction,</b> HAI-SON LE, and ZIV BAR-JOSEPH, Carnegie Mellon University	235
<b>Semantic Labeling of 3D Point Clouds for Indoor Scenes,</b> HEMA KOPPULA, ABHISHEK ANAND, THORSTEN JOACHIMS, and ASHUTOSH SAXENA, Cornell University	244
<b>Learning Higher-Order Graph Structure with Features by Structure Penalty,</b> SHILIN DING, University of Wisconsin-Madison, GRACE WAHBA, University of Wisconsin, and XIAOJIN (JERRY) ZHU, U. Wisconsin-Madison	253
<b>Analysis and Improvement of Policy Gradient Estimation,</b> TINGTING ZHAO, HIROTAKA HACHIYA, GANG NIU, and MASASHI SUGIYAMA, Tokyo Institute of Technology	262
<b>Dimensionality Reduction Using the Sparse Linear Model,</b> IOANNIS GKIOULEKAS, and TODD ZICKLER, Harvard University	271
<b>Robust Multi-Class Gaussian Process Classification,</b> DANIEL HERNÁNDEZ-LOBATO, Université catholique de Louvain, JOSE MIGUEL HERNÁNDEZ-LOBATO, Cambridge University, and PIERRE DUPONT, Louvain University	280
<b>Maximum Margin Multi-Label Structured Prediction,</b> , and CHRISTOPH LAMPERT, IST Austria	289
<b>Extracting Speaker-Specific Information with a Regularized Siamese Deep Network,</b> KE CHEN, The University of Manchester, and AHMAD SALMAN, University of Manchester	298
<b>Thinning Measurement Models and Questionnaire Design,</b> , and RICARDO SILVA, University College London	307
<b>Inductive reasoning about chimeric creatures,</b> , and CHARLES KEMP, Carnegie Mellon University	316
<b>Optimal Reinforcement Learning for Gaussian Systems,</b> , and PHILIPP HENNIG, Max Planck Institute for Intelligent Systems	325
<b>A Denoising View of Matrix Completion,</b> WEIRAN WANG, MIGUEL CARREIRA-PERPINAN, University of California, Merced, and ZHENG DONG LU, MSRA	334
<b>Efficient Online Learning via Randomized Rounding,</b> NICOLÒ CESA-BIANCHI, Università degli Studi di Milano, and OHAD SHAMIR, Microsoft Research	343
<b>Efficient Methods for Overlapping Group Lasso,</b> LEI YUAN, Arizona State University, JUN LIU, Siemens Corporate Research, and JIEPING YE, Arizona State University	352
<b>Differentially Private M-Estimators,</b> , and JING LEI, Carnegie Mellon University	361
<b>Multiple Instance Filtering,</b> KAMIL WNUK, University of California, Los Angeles, and STEFANO SOATTO, UCLA	370
<b>Phase transition in the family of p-resistances,</b> MORTEZA ALAMGIR, Max Planck Institute for Intelligent Systems, and ULRIKE VON LUXBURG, Max Plank Institute for Intelligent Systems	379

<b>Convergent Bounds on the Euclidean Distance</b> , YOONHO HWANG, and HEE-KAP AHN, POSTECH .....	388
<b>Heavy-tailed Distances for Gradient Based Image Descriptors</b> , YANGQING JIA, University of California, Berkeley, and TREVOR DARRELL, UC Berkeley .....	397
<b>RTRMC: A Riemannian trust-region method for low-rank matrix completion</b> , NICOLAS BOUMAL, UCLouvain, and PIERRE-ANTOINE ABSIL, U.C.Louvain .....	406
<b>Expressive Power and Approximation Errors of Restricted Boltzmann Machines</b> , GUIDO MONTUFAR, Max Planck Institute for Mathematics in the Sciences, JOHANNES RAUH, and NIHAT AY, MPI for Mathematics in the Sciences .....	415
<b>History distribution matching method for predicting effectiveness of HIV combination therapies</b> , , and JASMINA BOGOJESKA, Max-Planck Institute for Informatics .....	424
<b>Semi-supervised Regression via Parallel Field Regularization</b> , BINBIN LIN, CHIYUAN ZHANG, and XIAOFEI HE, Zhejiang University .....	433
<b>Object Detection with Grammar Models</b> , ROSS GIRSHICK, University of Chicago, PEDRO FELZENSZWALB, Brown University, and DAVID MCALLESTER, TTI-Chicago .....	442
<b>Non-Asymptotic Analysis of Stochastic Approximation Algorithms for Machine Learning</b> , FRANCIS BACH, INRIA - Ecole Normale Superieure, and ERIC MOULINES, Telecom Paristech .....	451
<b>On fast approximate submodular minimization</b> , STEFANIE JEGELKA, Max Planck Institute for Intelligent Systems, HUI LIN, and JEFF BILMES, University of Washington .....	460
<b>Emergence of Multiplication in a Biophysical Model of a Wide-Field Visual Neuron for Computing Object Approaches: Dynamics, Peaks, &amp; Fits</b> , , and MATTHIAS KEIL, University of Barcelona	469
<b>Efficient anomaly detection using bipartite k-NN graphs</b> , KUMAR SRICHARAN, University of Michigan Ann Arbor, and ALFRED HERO, University of Michigan .....	478
<b>Projection onto A Nonnegative Max-Heap</b> , JUN LIU, Siemens Corporate Research, LIANG SUN, and JIEPING YE, Arizona State University ....	487
<b>Improving Topic Coherence with Regularized Topic Models</b> , DAVID NEWMAN, University of California, Irvine, EDWIN BONILLA, NICTA, and WRAY BUNTINE, .....	496
<b>A Two-Stage Weighting Framework for Multi-Source Domain Adaptation</b> , QIAN SUN, RITA CHATTOPADHYAY, SETHURAMAN PANCHANATHAN, and JIEPING YE, Arizona State University .....	505
<b>An ideal observer model for identifying the reference frame of objects</b> , JOSEPH AUSTERWEIL, UC Berkeley, ABRAM FRIESEN, University of Washington, and TOM GRIFFITHS, University of California, Berkeley .....	514
<b>Generalized Beta Mixtures of Gaussians</b> , ARTIN ARMAGAN, Duke University, DAVID DUNSON, , and MERLISE CLYDE, Duke University .....	523

<b>Large-Scale Sparse Principal Component Analysis with Application to Text Data</b> , YOUWEI ZHANG, UC Berkeley, and LAURENT GHAOUI, .....	532
<b>Simultaneous Sampling and Multi-Structure Fitting with Adaptive Reversible Jump MCMC</b> , TRUNG PHAM, The University of Adelaide, TAT-JUN CHIN, the University of Adelaide, JIN YU, and DAVID SUTER, The University of Adelaide .....	540
<b><math>\theta</math>-MRF: Capturing Spatial and Semantic Structure in the Parameters for Scene Understanding</b> , CONGCONG LI, ASHUTOSH SAXENA, and TSUHAN CHEN, Cornell University .....	549
<b>Crowdclustering</b> , RYAN GOMES, PETER WELINDER, Caltech, ANDREAS KRAUSE, ETH Zurich, and PIETRO PERONA, Caltech .....	558
<b>Fast and Balanced: Efficient Label Tree Learning for Large Scale Object Recognition</b> , JIA DENG, Princeton University, SANJEEV SATHEESH, Stanford University, ALEXANDER BERG, Stony Brook, and FEI FEI LI, Stanford University .....	567
<b>Target Neighbor Consistent Feature Weighting for Nearest Neighbor Classification</b> , ICHIRO TAKEUCHI, Nagoya Institute of Technology, and MASASHI SUGIYAMA, Tokyo Institute of Technology .....	576
<b>The Impact of Unlabeled Patterns in Rademacher Complexity Theory for Kernel Classifiers</b> , LUCA ONETO, DAVIDE ANGUIA, University of Genoa, ALESSANDRO GHIO, University of Genoa, and SANDRO RIDELLA, University of Genoa, Italy .....	585
<b>Relative Density-Ratio Estimation for Robust Distribution Comparison</b> , MAKOTO YAMADA, Tokyo Institute of Technology, TAIJI SUZUKI, University of Tokyo, TAKAFUMI KANAMORI, Nagoya university, HIROTAKA HACHIYA, and MASASHI SUGIYAMA, Tokyo Institute of Technology .....	594
<b>Solving Decision Problems with Limited Information</b> , DENIS MAUA, and CASSIO DE CAMPOS, Dalle Molle Institute for Artificial Intelligence .....	603
<b>Linearized Alternating Direction Method with Adaptive Penalty for Low-Rank Representation</b> , ZHOUCHE LIN, Microsoft Research Asia, RISHENG LIU, and ZHIXUN SU, Dalian University of Technology .....	612
<b>Learning a Tree of Metrics with Disjoint Visual Features</b> , SUNG JU HWANG, KRISTEN GRAUMAN, University of Texas at Austin, and FEI SHA, University of Southern California .....	621
<b>Efficient inference in matrix-variate Gaussian models with i.i.d observation noise</b> , OLIVER STEGLE, Max Planck Institute Biological Cybernetics, CHRISTOPH LIPPERT, MPI for Developmental Biology and MPI for Biological Cybernetics, JORIS MOOIJ, Radboud University Nijmegen, NEIL LAWRENCE, University of Sheffield, and KARSTEN BORGWARDT, MPIs Tuebingen .....	630
<b>On Causal Discovery with Cyclic Additive Noise Models</b> , JORIS MOOIJ, Radboud University Nijmegen, DOMINIK JANZING, MPI for Biological Cybernetics, TOM HESKES, Radboud University Nijmegen, and BERNHARD SCHÖLKOPF, Max Planck Institute for Intelligent Systems .....	639
<b>Learning to Agglomerate Superpixel Hierarchies</b> , VIREN JAIN, Howard Hughes Medical Institute, SRINIVAS TURAGA, University College	

London, K BRIGGMAN, MPI, MORITZ HELMSTAEDTER, WINFRIED DENK, Max Planck Institute for Medical Research, and H. SEBASTIAN SEUNG, Massachusetts Institute of Technology .....	648
<b>A Convergence Analysis of Log-Linear Training</b> , SIMON WIESLER, and HERMANN NEY, RWTH Aachen .....	657
<b>Shallow vs. Deep Sum-Product Networks</b> , OLIVIER DELALLEAU, and YOSHUA BENGIO, University of Montreal .....	666
<b>Signal Estimation Under Random Time-Warpings and Nonlinear Signal Alignment</b> , SEBASTIAN KURTEK, ANUJ SRIVASTAVA, and WEI WU, Florida State University .....	675
<b>From Bandits to Experts: On the Value of Side-Observations</b> , SHIE MANNOR, Technion, and OHAD SHAMIR, Microsoft Research .....	684
<b>Hogwild: A Lock-Free Approach to Parallelizing Stochastic Gradient Descent</b> , BENJAMIN RECHT, UW-Madison, CHRISTOPHER RE, University of Wisconsin, STEPHEN WRIGHT, and FENG NIU, University of Wisconsin-Madison .....	693
<b>Clustered Multi-Task Learning Via Alternating Structure Optimization</b> , JIAYU ZHOU, JIANHUI CHEN, and JIEPING YE, Arizona State University .....	702
<b>Why The Brain Separates Face Recognition From Object Recognition</b> , JOEL LEIBO, JIM MUTCH, and TOMASO POGGIO, Massachusetts Institute of Technology .....	711
<b>Reinforcement Learning using Kernel-Based Stochastic Factorization</b> , ANDRE BARRETO, DOINA PRECUP, and JOELLE PINEAU, McGill University .....	720
<b>k-NN Regression Adapts to Local Intrinsic Dimension</b> , , and SAMORY KPOTUFE, Max Planck Institute .....	729
<b>Learning unbelievable probabilities</b> , XAQ PITKOW, University of Rochester, YASHAR AHMADIAN, and KEN MILLER, Columbia University .....	738
<b>A Machine Learning Approach to Predict Chemical Reactions</b> , MATTHEW KAYALA, and PIERRE BALDI, University of California, Irvine .....	747
<b>Dynamical segmentation of single trials from population neural data</b> , BILJANA PETRESKA, Gatsby Unit, UCL, BYRON YU, Carnegie Mellon University, JOHN CUNNINGHAM, Stanford University, GOPAL SANTHANAM, , STEPHEN RYU, Dept Electrical Eng., Stanford University, KRISHNA SHENOY, Stanford University, and MANEESH SAHANI, Gatsby Unit, UCL .....	756
<b>Recovering Intrinsic Images with a Global Sparsity Prior on Reflectance</b> , PETER GEHLER, Max Planck Institute Informatik, CARSTEN ROTHER, Microsoft, MARTIN KIEFEL, MPI for Intelligent Systems, LUMIN ZHANG, MPI, and BERNHARD SCHÖLKOPF, Max Planck Institute for Intelligent Systems .....	765
<b>Probabilistic Modeling of Dependencies Among Visual Short- Term Memory Representations</b> , EMIN ORHAN, and ROBERT JACOBS, University of Rochester .....	774
<b>Optimistic Optimization of a Deterministic Function without the Knowledge of its Smoothness</b> , , and REMI MUNOS, INRIA Lille - Nord Europe .....	783



<b>Reconstructing Patterns of Information Diffusion from Incomplete Observations,</b> FLAVIO CHIERICHETTI, Cornell, JON KLEINBERG, Cornell University, and DAVID LIBEN-NOWELL, Carleton College .....	792
<b>Dynamic Pooling and Unfolding Recursive Autoencoders for Paraphrase Detection,</b> RICHARD SOCHER, ERIC HUANG, JEFFREY PENNIN, ANDREW NG, Stanford University, and CHRISTOPHER MANNING, Stanford .....	801
<b>Active Learning Ranking from Pairwise Preferences with Almost Optimal Query Complexity,</b> , and NIR AILON, Technion .....	810
<b>Modelling Genetic Variations using Fragmentation-Coagulation Processes,</b> YEE WHYE TEH, Gatsby Computational Neuroscience Unit, UCL, CHARLES BLUNDELL, Gatsby Unit, UCL, and LLOYD ELLIOTT, University College London .....	819
<b>Prediction strategies without loss,</b> MICHAEL KAPRALOV, Stanford University, and RINA PANIGRAHY, Microsoft Research .....	828
<b>Data Skeletonization via Reeb Graphs,</b> XIAOYIN GE, Ohio State University, ISSAM SAFA, MIKHAIL BELKIN, and YUSU WANG, The Ohio State University .....	837
<b>Information Rates and Optimal Decoding in Large Neural Populations,</b> KAMIAR RAHNAMA RAD, Columbia University, and LIAM PANINSKI, .....	846
<b>Selective Prediction of Financial Trends with Hidden Markov Models,</b> DMITRY PIDAN, and RAN EL-YANIV, Technion .....	855
<b>Maximal Cliques that Satisfy Hard Constraints with Application to Deformable Object Model Learning,</b> XINGGANG WANG, XIANG BAI, Huazhong University of Science and Technology, XINGWEI YANG, , WENYU LIU, Huazhong University of Science and Technology, and LONGJIAN LATECKI, Temple University .....	864
<b>Distributed Delayed Stochastic Optimization,</b> ALEKH AGARWAL, and JOHN DUCHI, University of California Berkeley .....	873
<b>Greedy Algorithms for Structurally Constrained High Dimensional Problems,</b> AMBUJ TEWARI, UT Austin, PRADEEP RAVIKUMAR, University of Texas, Austin, and INDERJIT DHILLON, University of Texas at Austin .....	882
<b>Newtron: an Efficient Bandit algorithm for Online Multiclass Prediction,</b> ELAD HAZAN, Technion, and SATYEN KALE, IBM .....	891
<b>Learning Sparse Representations of High Dimensional Data on Large Scale Dictionaries,</b> ZHEN JAMES XIANG, HAO XU, and PETER RAMADGE, Princeton University .....	900
<b>Minimax Localization of Structural Information in Large Noisy Matrices,</b> MLADEN KOLAR, SIVARAMAN BALAKRISHNAN, Carnegie Mellon University, ALESSANDRO RINALDO, , and AARTI SINGH, Carnegie Mellon University .....	909
<b>Maximum Covariance Unfolding : Manifold Learning for Bimodal Data,</b> VIJAY MAHADEVAN, University of California, San Diego, CHI WAH WONG, University of California San Diego, JOSE COSTA PEREIRA, UC San	

Diego, TOM LIU, University of California, San Diego, NUNO VASCONCELOS, and LAWRENCE SAUL, UC San Diego .....	918
<b>Efficient Learning of Generalized Linear and Single Index Models with Isotonic Regression,</b> SHAM KAKADE, Microsoft Research, ADAM KALAI, Microsoft Research New England, VARUN KANADE, Harvard University, and OHAD SHAMIR, Microsoft Research .....	927
<b>On the Analysis of Multi-Channel Neural Spike Data,</b> BO CHEN, DAVID CARLSON, and LAWRENCE CARIN, Duke University .....	936
<b>Learning Eigenvectors for Free,</b> WOUTER KOOLEN, Royal Holloway, University of London, WOJCIECH KOTLOWSKI, Centrum Wiskunde en Informatica, and MANFRED WARMUTH, Univ. of Calif. at Santa Cruz .....	945
<b>Noise Thresholds for Spectral Clustering,</b> SIVARAMAN BALAKRISHNAN, MIN XU, AKSHAY KRISHNAMURTHY, and AARTI SINGH, Carnegie Mellon University .....	954
<b>The Kernel Beta Process,</b> LU REN, YINGJIAN WANG, Duke University, DAVID DUNSON, , and LAWRENCE CARIN, Duke University .....	963
<b>Statistical Performance of Convex Tensor Decomposition,</b> RYOTA TOMIOKA, TAJI SUZUKI, University of Tokyo, KOHEI HAYASHI, Nara Institute of Science and Technology, and HISASHI KASHIMA, .....	972
<b>Probabilistic amplitude and frequency demodulation,</b> RICHARD TURNER, , and MANEESH SAHANI, Gatsby Unit, UCL .....	981
<b>Directed Graph Embedding: an Algorithm based on Continuous Limits of Laplacian-type Operators,</b> DOMINIQUE PERRAULT-JONCAS, and MARINA MEILA, University of Washington .....	990
<b>Efficient coding of natural images with a population of noisy Linear-Nonlinear neurons,</b> YAN KARKLIN, New York University, and EERO SIMONCELLI, HHMI / New York University .....	999
<b>Complexity of Inference in Latent Dirichlet Allocation,</b> DAVID SONTAG, New York University, and DAN ROY, University of Cambridge .....	1008
<b>ICA with Reconstruction Cost for Efficient Overcomplete Feature Learning,</b> QUOC V. LE, ALEXANDRE KARPENKO, JIQUAN NGIAM, and ANDREW NG, Stanford University .....	1017
<b>Lower Bounds for Passive and Active Learning,</b> MAXIM RAGINSKY, University of Illinois at Urbana-Champaign, and ALEXANDER RAKHLIN, University of Pennsylvania .....	1026
<b>Stochastic convex optimization with bandit feedback,</b> ALEKH AGARWAL, University of California Berkeley, DEAN FOSTER, University of Pennsylvania, DANIEL HSU, SHAM KAKADE, Microsoft Research, and ALEXANDER RAKHLIN, University of Pennsylvania .....	1035
<b>Structure Learning for Optimization,</b> SHULIN YANG, University of Washington, and ALI RAHIMI, Red Bow Labs .....	1044
<b>Inverting Grice’s Maxims to Learn Rules from Natural Language Extractions,</b> MOHAMMAD SHAHED SOROWER, THOMAS DIETTERICH, JANARDHAN RAO DOPPA, WALKER ORR, PRASAD TADEPALLI, and XIAOLI FERN, Oregon State University .....	1053

<b>Active Classification based on Value of Classifier,</b> TIANSHI GAO, and DAPHNE KOLLER, Stanford University .....	1062
<b>Group Anomaly Detection using Flexible Genre Models,</b> LIANG XIONG, BARNABAS POZOS, and JEFF SCHNEIDER, Carnegie Mellon University .....	1071
<b>Approximating Semidefinite Programs in Sublinear Time,</b> DAN GARBER, and ELAD HAZAN, Technion .....	1080
<b>SpaRCS: Recovering low-rank and sparse matrices from compressive measurements,</b> ANDREW WATERS, ASWIN SANKARANARAYANAN, and RICHARD BARANIUK, Rice University .....	1089
<b>Budgeted Optimization with Concurrent Stochastic-Duration Experiments,</b> JAVAD AZIMI, Oregon State University, ALAN FERN, afern@eecs.oregonstate.edu, and XIAOLI FERN, Oregon State University .....	1098
<b>Online Submodular Set Cover, Ranking, and Repeated Active Learning,</b> ANDREW GUILLORY, and JEFF BILMES, University of Washington	1107
<b>Structured sparse coding via lateral inhibition,</b> ARTHUR SZLAM, KAROL GREGOR, NYU, and YANN LECUN, New York U .....	1116
<b>Sparse Filtering,</b> JIQUAN NGIAM, PANG WEI KOH, ZHENGHAO CHEN, SONIA BHASKAR, and ANDREW NG, Stanford University .....	1125
<b>Divide-and-Conquer Matrix Factorization,</b> LESTER MACKEY, University of California, Berkeley, AMEET TALWALKAR, UC Berkeley, and MICHAEL JORDAN, University of California .....	1134
<b>Im2Text: Describing Images Using 1 Million Captioned Photographs,</b> VICENTE ORDONEZ, GIRISH KULKARNI, and TAMARA BERG, Stony Brook University .....	1143
<b>Nonstandard Interpretations of Probabilistic Programs for Efficient Inference,</b> DAVID WINGATE, Massachusetts Institute of Technology, NOAH GOODMAN, , ANDREAS STUHLMUELLER, Massachusetts Institute of Technology, and JEFFREY SISKIND, Purdue University .....	1152
<b>Collective Graphical Models,</b> DANIEL SHELDON, and THOMAS DIETTERICH, Oregon State University .....	1161
<b>Metric Learning with Multiple Kernels,</b> JUN WANG, HUYEN DO, ADAM WOZNICA, and ALEXANDROS KALOUSIS, University of Geneva .....	1170
<b>ShareBoost: Efficient multiclass learning with feature sharing,</b> SHAI SHALEV-SHWARTZ, Hebrew University, YONATAN WEXLER, , and AMNON SHASHUA, Hebrew University of Jerusalem .....	1179
<b>Active dendrites: adaptation to spike-based communication,</b> BALAZS UJFALUSSY, and MATE LENGYEL, University of Cambridge .....	1188
<b>Message-Passing for Approximate MAP Inference with Latent Variables,</b> JIARONG JIANG, University of Maryland College Park, PIYUSH RAI, University of Utah, and HAL DAUME III, University of Maryland .....	1197
<b>A More Powerful Two-Sample Test in High Dimensions using Random Projection,</b> MILES LOPES, LAURENT JACOB, and MARTIN WAINWRIGHT, UC Berkeley .....	1206

<b>Orthogonal Matching Pursuit with Replacement,</b> PRATEEK JAIN, , AMBUJ TEWARI, UT Austin, and INDERJIT DHILLON, University of Texas at Austin .....	1215
<b>Composite Multiclass Losses,</b> ELODIE VERNET, ENS Cachan, ROBERT WILLIAMSON, Australian National University, and MARK REID, The Australian National University .....	1224
<b>Beating SGD: Learning SVMs in Sublinear Time,</b> ELAD HAZAN, TOMER KOREN, Technion, and NATI SREBRO, TTI-Chicago .....	1233
<b>Greedy Model Averaging,</b> DONG DAI, and TONG ZHANG, Rutgers University .....	1242
<b>Large-Scale Category Structure Aware Image Categorization,</b> BIN ZHAO, Carnegie Mellon University, FEI FEI LI, Stanford University, and ERIC XING, Carnegie Mellon University .....	1251
<b>On the accuracy of l1-filtering of signals with block-sparse structure,</b> ANATOLI IOUDITSKI, UJF, FATMA KILINC KARZAN, Carnegie Mellon University, ARKADI NEMIROVSKI, Georgia Institute of Technology, and BORIS POLYAK, Institute for Control Sciences, RAS Moscow .....	1260
<b>Multilinear Subspace Regression: An Orthogonal Tensor Decomposition Approach,</b> QIBIN ZHAO, RIKEN Brain Science Institute, CESAR CAIAFA, DANILO MANDIC, , LIQING ZHANG, Shanghai Jiao Tong University, TONIO BALL, ANDREAS SCHULZE-BONHAGE, Albert-Ludwigs-University, and ANDRZEJ CICHOCKI, RIKEB Brain Science Institute .....	1269
<b>Finite Time Analysis of Stratified Sampling for Monte Carlo,</b> ALEXANDRA CARPENTIER, INRIA Lille Nord Europe, and REMI MUNOS, INRIA Lille - Nord Europe .....	1278
<b>Monte Carlo Value Iteration with Macro-Actions,</b> ZHAN WEI LIM, DAVID HSU, and LEE SUN, National University of Singapore .....	1287
<b>Structured Learning for Cell Tracking,</b> XINGHUA LOU, University of Heidelberg, and FRED HAMPRECHT, Univ. of Heidelberg .....	1296
<b>Two is better than one: distinct roles for familiarity and recollection in retrieving palimpsest memories,</b> CRISTINA SAVIN, University of Cambridge, PETER DAYAN, Gatsby Computational Neuroscience Unit, and MATE LENGYEL, University of Cambridge .....	1305
<b>Algorithms and hardness results for parallel large margin learning,</b> ROCCO SERVEDIO, Columbia University, and PHIL LONG, Google .....	1314
<b>Portmanteau Vocabularies for Multi-Cue Image Representation,</b> FAHAD KHAN, Computer Vision Center, JOOST VAN DE WEIJER, Centre de Visió per Computador, ANDREW BAGDANOV, University of Florence., and MARIA VANRELL, Computer Vision Center .....	1323
<b>Boosting with Maximum Adaptive Sampling,</b> CHARLES DUBOUT, IDIAP, and FRANCOIS FLEURET, Idiap Research Institute .....	1332
<b>Gaussian Process Training with Input Noise,</b> ANDREW MCHUTCHON, Cambridge University, and CARL EDWARD RASMUSSEN, University of Cambridge .....	1341
<b>Empirical models of spiking in neural populations,</b> JAKOB MACKE, University College London, LARS BUESING, Gatsby Computational	

Neuroscience Unit, University College London, JOHN CUNNINGHAM, University of Cambridge, BYRON YU, Carnegie Mellon University, KRISHNA SHENOY, Stanford University, and MANEESH SAHANI, Gatsby Unit, UCL .....	1350
<b>Learning Probabilistic Non-Linear Latent Variable Models for Tracking Complex Activities,</b> ANGELA YAO, JUERGEN GALL, ETH Zurich, LUC GOOL, , and RAQUEL URTASUN, TTI-Chicago .....	1359
<b>Bayesian Partitioning of Large-Scale Distance Data,</b> DAVID ADAMETZ, and VOLKER ROTH, University of Basel .....	1368
<b>From Stochastic Nonlinear Integrate-and-Fire to Generalized Linear Models,</b> SKANDER MENSI, Brain-Mind Institute, RICHARD NAUD, and WULFRAM GERSTNER, Ecole Polytechnique Federale de Lausanne .....	1377
<b>On the Completeness of First-Order Knowledge Compilation for Lifted Probabilistic Inference,</b> , and GUY VAN DEN BROECK, Katholieke Universiteit Leuven .....	1386
<b>Hierarchical Topic Modeling for Analysis of Time-Evolving Personal Choices,</b> XIANXING ZHANG, Duke University, DAVID DUNSON, , and LAWRENCE CARIN, Duke University .....	1395
<b>An Exact Algorithm for F-Measure Maximization,</b> KRZYSZTOF DEMBCZYNSKI, Poznan University of Technology, WILLEM WAEGEMAN, Ghent University, WEIWEI CHENG, University of Marburg, and EYKE HULLERMEIER, Philipps-Universität Marburg .....	1404
<b>Co-regularized Multi-view Spectral Clustering,</b> ABHISHEK KUMAR, University of Maryland, College Park, PIYUSH RAI, University of Utah, and HAL DAUME III, University of Maryland .....	1413
<b>Sequence learning with hidden units in spiking neural networks,</b> JOHANNI BREA, Universität Bern, WALTER SENN, , and JEAN-PASCAL PFISTER, Cambridge University .....	1422
<b>Identifying Alzheimer’s Disease-Related Brain Regions from Multi-Modality Neuroimaging Data using Sparse Composite Linear Discrimination Analysis,</b> SHUAI HUANG, JING LI, JIEPING YE, TERESA WU, Arizona State University, KEWEI CHEN, ADAM FLEISHER, and ERIC REIMAN, Banner Alzheimer’s Institute .....	1431
<b>A blind sparse deconvolution method for neural spike identification,</b> CHAITANYA EKANADHAM, Courant Institute, New York University, DANIEL TRANCHINA, Courant Institute, NYU, and EERO SIMONCELLI, HHMI / New York University .....	1440
<b>How Do Humans Teach: On Curriculum Learning and Teaching Dimension,</b> FAISAL KHAN, University of Wisconsin-Madison, XIAOJIN (JERRY) ZHU, U. Wisconsin-Madison, and BILGE MUTLU, University of Wisconsin-Madison .....	1449
<b>Convergence Rates of Inexact Proximal-Gradient Methods for Convex Optimization,</b> MARK SCHMIDT, INRIA - SIERRA Project Team, NICOLAS LE ROUX, INRIA, and FRANCIS BACH, INRIA - Ecole Normale Supérieure .....	1458
<b>Joint 3D Estimation of Objects and Scene Layout,</b> ANDREAS GEIGER, KIT, CHRISTIAN WOJEK, , and RAQUEL URTASUN, TTI-Chicago ..	1467

<b>Spatial distance dependent Chinese restaurant processes for image segmentation,</b> SOUMYA GHOSH, Brown University, ANDREI UNGUREANU, Morgan Stanley, ERIK SUDDERTH, Brown University, and DAVID BLEI, Princeton University	1476
<b>Pylon Model for Semantic Segmentation,</b> VICTOR LEMPITSKY, Yandex, ANDREA VEDALDI, and ANDREW ZISSERMAN, University of Oxford	1485
<b>t-divergence Based Approximate Inference,</b> NAN DING, S.V.N. VISHWANATHAN, Purdue University, and YUAN (ALAN) QI, Purdue university	1494
<b>Learning person-object interactions for action recognition in still images,</b> VINCENT DELAITRE, Ecole Normale Supérieure, JOSEF SIVIC, INRIA / Ecole Normale Supérieure, and IVAN LAPTEV, INRIA	1503
<b>Submodular Multi-Label Learning,</b> JAMES PETTERSON, NICTA, and TIBERIO CAETANO, NICTA/Australian National University	1512
<b>Uniqueness of Belief Propagation on Signed Graphs,</b> , and YUSUKE WATANABE, The Institute of Statistical Mathematics	1521
<b>Higher-Order Correlation Clustering for Image Segmentation,</b> SUNGWOONG KIM, KAIST, SEBASTIAN NOWOZIN, Microsoft Research Cambridge, PUSHMEET KOHLI, Microsoft Research, and CHANG D. YOO, KAIST	1530
<b>Optimal learning rates for least squares SVMs using Gaussian kernels,</b> MONA EBERTS, and INGO STEINWART, University of Stuttgart	1539
<b>Learning Auto-regressive Models from Sequence and Non-sequence Data,</b> TZU-KUO HUANG, and JEFF SCHNEIDER, Carnegie Mellon University	1548
<b>Committing Bandits,</b> LOC BUI, Technion, RAMESH JOHARI, Stanford University, and SHIE MANNOR, Technion	1557
<b>Energetically Optimal Action Potentials,</b> MARTIN STEMMLER, LMU Munich, BISWA SENGUPTA, University of Cambridge, SIMON LAUGHLIN, Cambridge University, and JEREMY NIVEN, University of Sussex	1566
<b>Unifying Framework for Fast Learning Rate of Non-Sparse Multiple Kernel Learning,</b> , and TAJI SUZUKI, University of Tokyo	1575
<b>See the Tree Through the Lines: The Shazoo Algorithm,</b> FABIO VITALE, NICOLÒ CESA-BIANCHI, Università degli Studi di Milano, CLAUDIO GENTILE, Università dell'Insubria, and GIOVANNI ZAPPELLA, Università degli studi di Milano	1584
<b>The Fast Convergence of Boosting,</b> , and MATUS TELGARSKY, University of California, San Diego	1593
<b>Multi-armed bandits on implicit metric spaces,</b> , and ALEKSANDRS SLIVKINS, Microsoft Research	1602
<b>Learning Anchor Planes for Classification,</b> ZIMING ZHANG, Oxford Brookes University, LUBOR LADICKY, University of Oxford, PHILIP TORR, Oxford Brookes, and AMIR SAFFARI, Sony Computer Entertainment Europe	1611
<b>Infinite Latent SVM for Classification and Multi-task Learning,</b> JUN ZHU, Carnegie Mellon University, NING CHEN, Tsinghua University, and ERIC XING, Carnegie Mellon University	1620

<b>Facial Expression Transfer with Input-Output Temporal Restricted Boltzmann Machines</b> , MATTHEW ZEILER, GRAHAM TAYLOR, New York University, LEONID SIGAL, Disney Research, Pittsburgh, IAIN MATTHEWS, Disney Research Pittsburgh, and ROB FERGUS, NYU	1629
<b>Universal low-rank matrix recovery from Pauli measurements</b> , and YI-KAI LIU, National Institute of Standards and Technology	1638
<b>Better Mini-Batch Algorithms via Accelerated Gradient Methods</b> , ANDREW COTTER, TTI Chicago, OHAD SHAMIR, Microsoft Research, NATI SREBRO, TTI-Chicago, and KARTHIK SRIDHARAN, Toyota Technological Institute at Chicago	1647
<b>Adaptive Hedge</b> , TIM VAN ERVEN, VU University, PETER GRUNWALD, CWI, WOUTER KOOLEN, Royal Holloway, University of London, and STEVEN ROOIJ, CWI	1656
<b>Agnostic Selective Classification</b> , YAIR WIENER, and RAN EL-YANIV, Technion	1665
<b>Comparative Analysis of Viterbi Training and Maximum Likelihood Estimation for HMMs</b> , ARMEN ALLAHVERDYAN, Yerevan Physics Institute, and ARAM GALSTYAN, University of Southern California	1674
<b>PAC-Bayesian Analysis of Contextual Bandits</b> , YEVGENY SELDIN, Max Planck Institute for Intelligent Systems, PETER AUER, University of Leoben, FRANCOIS LAVIOLETTE, Université Laval, JOHN SHAWE-TAYLOR, University College London, and RONALD ORTNER, Montanuniversitaet Leoben	1683
<b>Bayesian Spike-Triggered Covariance Analysis</b> , IL MEMMING PARK, and JONATHAN PILLOW, University of Texas at Austin	1692
<b>Non-conjugate Variational Message Passing for Multinomial and Binary Regression</b> , DAVID KNOWLES, University of Cambridge, and TOM MINKA, Microsoft Research Ltd	1701
<b>Learning to Search Efficiently in High Dimensions</b> , ZHEN LI, UIUC, HUAZHONG NING, Google Inc, LIANGLIANG CAO, IBM T. J. Watson Research Center, TONG ZHANG, Rutgers University, YIHONG GONG, , and THOMAS HUANG, UIUC	1710
<b>A Non-Parametric Approach to Dynamic Programming</b> , OLIVER KROEMER, Technische Universität Darmstadt, and JAN PETERS, Technische Universitaet Darmstadt	1719
<b>Advice Refinement in Knowledge-Based SVMs</b> , GAUTAM KUNAPULI, University of Wisconsin-Madison, RICHARD MACLIN, University of Minnesota Duluth, and JUDE SHAVLIK, University of Wisconsin at Madison	1728
<b>Kernel Bayes' Rule</b> , KENJI FUKUMIZU, Institute of Statistical Mathematics, LE SONG, Carnegie Mellon University, and ARTHUR GRETTON, Gatsby Unit, UCL	1737
<b>Transfer from Multiple MDPs</b> , ALESSANDRO LAZARIC, INRIA Lille-Nord Europe, and MARCELLO RESTELLI, Politecnico di Milano	1746
<b>Sparse Bayesian Multi-Task Learning</b> , CEDRIC ARCHAMBEAU, SHENGBO GUO, and ONNO ZOETER, Xerox Research Centre Europe	1755
<b>Online Learning: Stochastic, Constrained, and Smoothed Adversaries</b> , ALEXANDER RAKHLIN, University of Pennsylvania, KARTHIK	

SRIDHARAN, Toyota Technological Institute at Chicago, and AMBUJ TEWARI, UT Austin .....	1764
<b>Learning in Hilbert vs. Banach Spaces: A Measure Embedding Viewpoint,</b> BHARATH SRIPERUMBUDUR, UCL, KENJI FUKUMIZU, Institute of Statistical Mathematics, and GERT LANCKRIET, U.C. San Diego ...	1773
<b>Sparse Recovery with Brownian Sensing,</b> ALEXANDRA CARPENTIER, INRIA Lille Nord Europe, ODALRIC-AMBRYM MAILLARD, Montanuniversität Leoben, and REMI MUNOS, INRIA Lille - Nord Europe .....	1782
<b>An Unsupervised Decontamination Procedure For Improving The Reliability Of Human Judgments,</b> MICHAEL MOZER, BENJAMIN LINK, University of Colorado, and HAROLD PASHLER, University of California, San Diego .....	1791
<b>Bayesian Bias Mitigation for Crowdsourcing,</b> FABIAN WAUTHIER, UC Berkeley, and MICHAEL JORDAN, University of California .....	1800
<b>Ranking annotators for crowdsourced labeling tasks,</b> VIKAS RAYKAR, Siemens Healthcare, and SHIPENG YU, Siemens Medical Solutions ...	1809
<b>Clustering via Dirichlet Process Mixture Models for Portable Skill Discovery,</b> SCOTT NIEKUM, University of Massachusetts Amherst, and ANDREW BARTO, University of Massachusetts .....	1818
<b>Probabilistic Joint Image Segmentation and Labeling,</b> ADRIAN ION, TU Wien & IST Austria, JOAO CARREIRA, and CRISTIAN SMINCHISESCU, University of Bonn .....	1827
<b>Variance Reduction in Monte-Carlo Tree Search,</b> JOEL VENESS, MARC LANCTOT, and MICHAEL BOWLING, University of Alberta .....	1836
<b>Learning Patient-Specific Cancer Survival Distributions as a Sequence of Dependent Regressors,</b> CHUN-NAM YU, RUSS GREINER, University of Alberta, HSIU-CHIN LIN, University of Alberta /University of Edinburgh, and VICKIE BARACOS, .....	1845
<b>An Application of Tree-Structured Expectation Propagation for Channel Decoding,</b> PABLO OLMOS, Universidad de Sevilla, LUIS SALAMANCA, University of Sevilla, JUAN JOSÉ MURILLO FUENTES, Universidad de Sevilla, and FERNANDO PEREZ-CRUZ, University Carlos III in Madrid .....	1854
<b>High-Dimensional Graphical Model Selection: Tractable Graph Families and Necessary Conditions,</b> ANIMASHREE ANANDKUMAR, UC Irvine, VINCENT TAN, University of Wisconsin-Madison, and ALAN WILLSKY, Massachusetts Institute of Technology .....	1863
<b>Structural equations and divisive normalization for energy- dependent component analysis,</b> JUN-ICHIRO HIRAYAMA, Kyoto University, and AAPO HYVARINEN, .....	1872
<b>Robust Lasso with missing and grossly corrupted observations,</b> NAM NGUYEN, Johns Hopkins University, NASSER NASRABADI, Army Research Lab, and TRAC TRAN, Johns Hopkins University .....	1881
<b>A concave regularization technique for sparse mixture models,</b> MARTIN LARSSON, Operations Research and Information Engineering, and JOHAN UGANDER, Cornell University .....	1890



<b>Learning a Distance Metric from a Network</b> , BLAKE SHAW, Foursquare, BERT HUANG, University of Maryland, and TONY JEBARA, Columbia University .....	1899
<b>Variance Penalizing AdaBoost</b> , PANNAGADATTA SHIVASWAMY, Cornell University, and TONY JEBARA, Columbia University .....	1908
<b>Efficient Offline Communication Policies for Factored Multiagent POMDPs</b> , JOÃO MESSIAS, Instituto Superior Técnico, 501 507 930 TU Lisbon, MATTHIJS SPAAN, Delft University of Technology, and PEDRO LIMA, Instituto Superior Técnico, 501 507 930 TU Lisbon .....	1917
<b>Sparse recovery by thresholded non-negative least squares</b> , MARTIN SLAWSKI, and MATTHIAS HEIN, Saarland University .....	1926
<b>On Learning Discrete Graphical Models using Greedy Methods</b> , ALI JALALI, CHRISTOPHER JOHNSON, University of Texas at Austin, and PRADEEP RAVIKUMAR, University of Texas, Austin .....	1935
<b>Policy Gradient Coagent Networks</b> , , and PHILIP THOMAS, University of Massachusetts Amherst .....	1944
<b>Iterative Learning for Reliable Crowdsourcing Systems</b> , DAVID KARGER, MIT, SEWOONG OH, and DEVAVRAT SHAH, Massachusetts Institute of Technology .....	1953
<b>A Model for Temporal Dependencies in Event Streams</b> , ASELA GUNAWARDANA, CHRISTOPHER MEEK, Microsoft Research, and PUYANG XU, Johns Hopkins University .....	1962
<b>Unsupervised learning models of primary cortical receptive fields and receptive field plasticity</b> , ANDREW SAXE, MANEESH BHAND, RITVIK MUDUR, BIPIN SURESH, and ANDREW NG, Stanford University .....	1971
<b>The Doubly Correlated Nonparametric Topic Model</b> , DAE IL KIM, and ERIK SUDDERTH, Brown University .....	1980
<b>MAP Inference for Bayesian Inverse Reinforcement Learning</b> , JAEDEUG CHOI, and KEE-EUNG KIM, KAIST .....	1989
<b>Similarity-based Learning via Data Driven Embeddings</b> , PURUSHOTTAM KAR, Indian Institute of Technology Kanpur, and PRATEEK JAIN, .....	1998
<b>Predicting Dynamic Difficulty</b> , OLANA MISSURA, and THOMAS GAERTNER, University of Bonn .....	2007
<b>Sparse Estimation with Structured Dictionaries</b> , , and DAVID WIPF, Microsoft Research Asia .....	2016
<b>Spectral Methods for Learning Multivariate Latent Tree Structure</b> , ANIMASHREE ANANDKUMAR, U.C.Irvine, KAMALIKA CHAUDHURI, UC San Diego, DANIEL HSU, SHAM KAKADE, Microsoft Research, LE SONG, Carnegie Mellon University, and TONG ZHANG, Rutgers University .....	2025
<b>How biased are maximum entropy models?</b> , JAKOB MACKE, University College London, IAIN MURRAY, University of Edinburgh, and PETER LATHAM, University College London .....	2034
<b>Active learning of neural response functions with Gaussian processes</b> , MIJUNG PARK, The University of Texas at Austin, GREG HORWITZ, , and JONATHAN PILLOW, University of Texas at Austin .....	2043

<b>Priors over Recurrent Continuous Time Processes,</b> ARDAVAN SAEEDI, University of British Columbia, and ALEXANDRE BOUCHARD-CÔTÈ, UBC .....	2052
<b>Learning to Learn with Compound HD Models,</b> RUSLAN SALAKHUTDINOV, University of Toronto, JOSH TENENBAUM, and ANTONIO TORRALBA, Massachusetts Institute of Technology .....	2061
<b>Anatomically Constrained Decoding of Finger Flexion from Electrocorticographic Signals,</b> ZUOGUAN WANG, Rensselaer Polytechnic Institute, GERWIN SCHALK, Wadsworth Center, and QIANG JI, RPI .....	2070
<b>Active Learning with a Drifting Distribution,</b> , and LIU YANG, Carnegie Mellon University .....	2079
<b>PiCoDes: Learning a Compact Code for Novel-Category Recognition,</b> ALESSANDRO BERGAMO, Dartmouth College, LORENZO TORRESANI, Dartmouth, and ANDREW FITZGIBBON, Microsoft Research .....	2088
<b>Confidence Sets for Network Structure,</b> DAVID CHOI, PATRICK WOLFE, and EDOARDO AIROLDI, Harvard University .....	2097
<b>Prismatic Algorithm for Discrete D.C. Programming Problem,</b> YOSHINOBU KAWAHARA, and TAKASHI WASHIO, Osaka University .....	2106
<b>Hierarchical Matching Pursuit for Image Classification: Architecture and Fast Algorithms,</b> LIEFENG BO, University of Washington, XIAOFENG REN, Intel Labs Seattle, and DIETER FOX, Intel Labs Seattle and University of Washington .....	2115
<b>Multiclass Boosting: Theory and Algorithms,</b> MOHAMMAD SABERIAN, and NUNO VASCONCELOS, UC San Diego .....	2124
<b>Learning with the weighted trace-norm under arbitrary sampling distributions,</b> RINA FOYGEL, University of Chicago, RUSLAN SALAKHUTDINOV, University of Toronto, OHAD SHAMIR, Microsoft Research, and NATI SREBRO, TTI-Chicago .....	2133
<b>Scalable Training of Mixture Models via Coresets,</b> DAN FELDMAN, California Institute of Technology, MATTHEW FAULKNER, Caltech, and ANDREAS KRAUSE, ETH Zurich .....	2142
<b>Generalised Coupled Tensor Factorisation,</b> KENAN YILMAZ, Boğaziçi University, ALI TAYLAN CEMGIL, and UMUT SIMSEKLI, Bogazici University ..	2151
<b>Nearest Neighbor based Greedy Coordinate Descent,</b> INDERJIT DHILLON, University of Texas at Austin, PRADEEP RAVIKUMAR, University of Texas, Austin, and AMBUJ TEWARI, UT Austin .....	2160
<b>The Fixed Points of Off-Policy TD,</b> , and J. ZICO KOLTER, Massachusetts Institute of Technology .....	2169
<b>Generalizing from Several Related Classification Tasks to a New Unlabeled Sample,</b> GILLES BLANCHARD, Universität Potsdam (DE), GYEMIN LEE, and CLAY SCOTT, University of Michigan .....	2178
<b>Trace Lasso: a trace norm regularization for correlated designs,</b> EDOUARD GRAVE, INRIA - Ecole normale supérieure, GUILLAUME OBOZINSKI, INRIA / ENS, and FRANCIS BACH, INRIA - Ecole Normale Supérieure .....	2187

<b>Statistical Tests for Optimization Efficiency,</b> LEVI BOYLES, UC Irvine, ANOOP KORATTIKARA, University of California, Irvine, DEVA RAMANAN, , and MAX WELLING, University of California Irvine .....	2196
<b>Generalization Bounds and Consistency for Latent Structural Probit and Ramp Loss,</b> DAVID MCALLESTER, and JOSEPH KESHET, TTI-Chicago .....	2205
<b>A Brain-Machine Interface Operating with a Real-Time Spiking Neural Network Control Algorithm,</b> JULIE DETHIER, University of Liege, PAUL NUYUJUKIAN, Stanford University, CHRIS ELIASMITH, U of Waterloo, TERRENCE STEWART, University of Waterloo, SHAUKI ELASAAD, , KRISHNA SHENOY, and KWABENA BOAHEN, Stanford University .....	2213
<b>Multi-Bandit Best Arm Identification,</b> VICTOR GABILLON, MOHAMMAD GHAVAMZADEH, INRIA Lille - Nord Europe, ALESSANDRO LAZARIC, INRIA Lille-Nord Europe, and SEBASTIEN BUBECK, Princeton University .....	2222
<b>Randomized Algorithms for Comparison-based Search,</b> DOMINIQUE TSCHOPP, , SUHAS DIGGAVI, UCLA, PAYAM DELGOSHA, Sharif University of Technology, and SOHEIL MOHAJER, Princeton .....	2231
<b>Active Ranking using Pairwise Comparisons,</b> KEVIN JAMIESON, University of Wisconsin - Madison, and ROB NOWAK, University of Wisconsin-Madison .....	2240
<b>An Empirical Evaluation of Thompson Sampling,</b> OLIVIER CHAPELLE, Yahoo!, and LIHONG LI, Yahoo! Research .....	2249
<b>Blending Autonomous Exploration and Apprenticeship Learning,</b> THOMAS WALSH, University of Kansas, DANIEL HEWLETT, Google, and CLAYTON MORRISON, University of Arizona .....	2258
<b>Nonnegative dictionary learning in the exponential noise model for adaptive music signal representation,</b> ONUR DIKMEN, Telecom ParisTech, and CÉDRIC FÉVOTTE, CNRS LTCI; Télécom ParisTech .....	2267
<b>Evaluating the inverse decision-making approach to preference learning,</b> ALAN JERN, CHRISTOPHER LUCAS, and CHARLES KEMP, Carnegie Mellon University .....	2276
<b>Sparse Features for PCA-Like Linear Regression,</b> CHRISTOS BOUSIDIS, IBM T.J. Watson Research, PETROS DRINEAS, RPI & NSF, and MALIK MAGDON-ISMAIL, RPI .....	2285
<b>The Manifold Tangent Classifier,</b> SALAH RIFAI, Université de montreal, YANN DAUPHIN, Université de Montréal, PASCAL VINCENT, YOSHUA BENGIO, University of Montreal, and XAVIER MULLER, Université of Montreal	2294
<b>Analytical Results for the Error in Filtering of Gaussian Processes,</b> ALEX SUSEMIHL, Berlin Institute of Technology, RON MEIR, Technion, and MANFRED OPPER, .....	2303
<b>Improved Algorithms for Linear Stochastic Bandits,</b> YASIN ABBASI-YADKORI, University of Alberta, DAVID PAL, Google, and CSABA SZEPESVARI, University of Alberta .....	2312
<b>Testing a Bayesian Measure of Representativeness Using a Large Image Database,</b> JOSHUA ABBOTT, University of California, Berkeley, KATHERINE HELLER, Massachusetts Institute of Technology, ZOUBIN	

GHAHRAMANI, University of Cambridge & CMU, and TOM GRIFFITHS, University of California, Berkeley .....	2321
<b>Sparse Inverse Covariance Matrix Estimation Using Quadratic Approximation,</b> CHO-JUI HSIEH, MATYAS SUSTIK, INDERJIT DHILLON, University of Texas at Austin, and PRADEEP RAVIKUMAR, University of Texas, Austin .....	2330
<b>Spike and Slab Variational Inference for Multi-Task and Multiple Kernel Learning,</b> MICHALIS TITSIAS, University of Manchester, and MIGUEL LÁZARO-GREDILLA, Universidad Carlos III de Madrid .....	2339
<b>Practical Variational Inference for Neural Networks,</b> , and ALEX GRAVES, University of Toronto .....	2348
<b>Neuronal Adaptation for Sampling-Based Probabilistic Inference in Perceptual Bistability,</b> DAVID REICHERT, PEGGY SERIES, and AMOS STORKEY, University of Edinburgh .....	2357
<b>Beyond Spectral Clustering - Tight Relaxations of Balanced Graph Cuts,</b> MATTHIAS HEIN, and SIMON SETZER, Saarland University ..	2366
<b>Fast and Accurate k-means For Large Datasets,</b> MICHAEL SHINDLER, Oregon State University, ALEX WONG, University of California, Los Angeles, and ADAM MEYERSON, Google, Inc. ....	2375
<b>A rational model of causal inference with continuous causes,</b> MICHAEL PACER, UC, Berkeley, and TOM GRIFFITHS, University of California, Berkeley .....	2384
<b>Quasi-Newton Methods for Markov Chain Monte Carlo,</b> YICHUAN ZHANG, and CHARLES SUTTON, University of Edinburgh .....	2393
<b>TD-Gamma: Re-evaluating Complex Backups in Temporal Difference Learning,</b> GEORGE KONIDARIS, Massachusetts Institute of Technology, SCOTT NIEKUM, and PHILIP THOMAS, University of Massachusetts Amherst .....	2402
<b>Speedy Q-Learning,</b> MOHAMMAD GHESHLAGHI AZAR, Radboud University of Nijmegen, REMI MUNOS, MOHAMMAD GHAVAMZADEH, INRIA Lille - Nord Europe, and HILBERT KAPPEN, Radboud University .....	2411
<b>Regularized Laplacian Estimation and Fast Eigenvector Approximation,</b> PATRICK PERRY, NYU, and MICHAEL MAHONEY, Stanford .....	2420
<b>Understanding the Intrinsic Memorability of Images,</b> PHILLIP ISOLA, Massachusetts Institute of Technology, DEVI PARIKH, TTIC, ANTONIO TORRALBA, and AUDE OLIVA, Massachusetts Institute of Technology .....	2429
<b>The Local Rademacher Complexity of Lp-Norm Multiple Kernel Learning,</b> MARIUS KLOFT, TU Berlin, and GILLES BLANCHARD, Universität Potsdam (DE) .....	2438
<b>Contextual Gaussian Process Bandit Optimization,</b> ANDREAS KRAUSE, and CHENG SOON ONG, ETH Zurich .....	2447
<b>Co-Training for Domain Adaptation,</b> MINMIN CHEN, KILIAN WEINBERGER, Washington University, and JOHN BLITZER, Google Research ..	2456
<b>Autonomous Learning of Action Models for Planning,</b> NEVILLE MEHTA, PRASAD TADEPALLI, Oregon State University, and ALAN FERN, afern@eecs.oregonstate.edu .....	2465

<b>Gaussian process modulated renewal processes</b> , VINAYAK RAO, Gatsby Unit, UCL, and YEE WHYE TEH, Gatsby Computational Neuroscience Unit, UCL .....	2474
<b>Linear Submodular Bandits and their Application to Diversified Retrieval</b> , YISONG YUE, and CARLOS GUESTRIN, Carnegie Mellon University .....	2483
<b>Continuous-Time Regression Models for Longitudinal Networks</b> , DUY VU, Pennsylvania State University, ARTHUR ASUNCION, University of California, Irvine, DAVID HUNTER, Penn State University, and PADHRAIC SMYTH, University of California, Irvine .....	2492
<b>On Tracking The Partition Function</b> , GUILLAUME DESJARDINS, Universite de Montreal, AARON COURVILLE, , and YOSHUA BENGIO, University of Montreal .....	2501
<b>Variational Gaussian Process Dynamical Systems</b> , ANDREAS DAMIANOU, University of Sheffield, MICHALIS TITSIAS, University of Manchester, and NEIL LAWRENCE, University of Sheffield .....	2510
<b>Non-parametric Group Orthogonal Matching Pursuit for Sparse Learning with Multiple Kernels</b> , VIKAS SINDHWANI, IBM Research, and AURELIE LOZANO, IBM T.J. Watson Research Center .....	2519
<b>Selecting Receptive Fields in Deep Networks</b> , ADAM COATES, and ANDREW NG, Stanford University .....	2528
<b>Convergent Fitted Value Iteration with Linear Function Approximation</b> , , and DAN LIZOTTE, University of Waterloo .....	2537
<b>Algorithms for Hyper-Parameter Optimization</b> , JAMES BERGSTRA, Harvard, RÈMI BARDENET, Univ. Paris-Sud XI, YOSHUA BENGIO, University of Montreal, and BALÁZS KÈGL, University of Paris-Sud/CNRS .....	2546
<b>Neural Reconstruction with Approximate Message Passing (NeuRAMP)</b> , ALYSON FLETCHER, University of California, Berkeley, SUNDEEP RANGAN, Polytechnic Institute of New York University, LAV VARSHNEY, IBM, and ANIRUDDHA BHARGAVA, University of Wisconsin-Madison .....	2555
<b>Query-Aware MCMC</b> , MICHAEL WICK, University of Massachusetts, and ANDREW MCCALLUM, UMass Amherst .....	2564
<b>A reinterpretation of the policy oscillation phenomenon in approximate policy iteration</b> , , and PAUL WAGNER, Aalto University School of Science .....	2573
<b>Inferring spike-timing-dependent plasticity from spike train data</b> , IAN STEVENSON, University of California, Berkeley, and KONRAD KOERDING, .....	2582
<b>Automated Refinement of Bayes Networks' Parameters based on Test Ordering Constraints</b> , OMAR KHAN, PASCAL POUPART, University of Waterloo, and JOHN-MARK AGOSTA, Intel Corporation .....	2591
<b>A Collaborative Mechanism for Crowdsourcing Prediction Problems</b> , JACOB ABERNETHY, and RAFAEL FRONGILLO, UC Berkeley ...	2600
<b>Hierarchically Supervised Latent Dirichlet Allocation</b> , ADLER PEROTTE, FRANK WOOD, NOEMIE ELHADAD, Columbia University, and NICHOLAS BARTLETT, Columbia .....	2609

<b>Select and Sample - A Model of Efficient Neural Inference and Learning,</b> JACQUELYN SHELTON, Frankfurt Institute for Advanced Studies, JORG BORNSCHEIN, Goethe-University Frankfurt, ABDUL SABOOR SHEIKH, Frankfurt Institute for Advanced Studies, PIETRO BERKES, Brandeis University, and JORG LUCKE, Goethe-University Frankfurt .....	2618
<b>Selecting the State-Representation in Reinforcement Learning,</b> ODALRIC-AMBRYM MAILLARD, Montanuniversität Leoben, REMI MUNOS, INRIA Lille - Nord Europe, and DANIIL RYABKO, INRIA .....	2627
<b>Periodic Finite State Controllers for Efficient POMDP and DEC-POMDP Planning,</b> JONI PAJARINEN, and JAAKKO PELTONEN, Aalto University .....	2636
<b>On the Universality of Online Mirror Descent,</b> NATI SREBRO, TTI-Chicago, KARTHIK SRIDHARAN, Toyota Technological Institute at Chicago, and AMBUJ TEWARI, UT Austin .....	2645
<b>Demixed Principal Component Analysis,</b> WIELAND BRENDEL, Ecole Normale Supérieure, RANULFO ROMO, Universidad Autónoma de México, and CHRISTIAN MACHENS, .....	2654
<b>EigenNet: A Bayesian hybrid of generative and conditional models for sparse learning,</b> YUAN (ALAN) QI, Purdue University, and FENG YAN, Purdue University .....	2663
<b>Hashing Algorithms for Large-Scale Learning,</b> PING LI, Cornell, ANSHUMALI SHRIVASTAVA, JOSHUA MOORE, Cornell University, and ARND KÖNIG, Microsoft Research .....	2672
<b>Rapid Deformable Object Detection using Dual-Tree Branch-and-Bound,</b> , and IASONAS KOKKINOS, Ecole Centrale Paris / INRIA Saclay .....	2681
<b>Hierarchical Multitask Structured Output Learning for Large-scale Sequence Segmentation,</b> NICO GOERNITZ, Technical University Berlin, CHRISTIAN WIDMER, Max Planck Society, GEORG ZELLER, EMBL, ANDRE KAHLES, FML of the MPS, SOEREN SONNENBURG, TomTom, and GUNNAR RAETSCH, Max Planck Society .....	2690
<b>Predicting response time and error rates in visual search,</b> BO CHEN, VIDHYA NAVALPAKKAM, and PIETRO PERONA, Caltech .....	2699
<b>Kernel Embeddings of Latent Tree Graphical Models,</b> LE SONG, Carnegie Mellon University, ANKUR PARIKH, Carnegie Mellon, and ERIC XING, Carnegie Mellon University .....	2708
<b>Inference in continuous-time change-point models,</b> FLORIAN STIMBERG, TU Berlin, MANFRED OPPER, Technische Universität Berlin, GUIDO SANGUINETTI, University of Edinburgh, and ANDREAS RUTTOR, TU Berlin .....	2717
<b>High-dimensional regression with noisy and missing data: Provable guarantees with non-convexity,</b> PO-LING LOH, and MARTIN WAINWRIGHT, UC Berkeley .....	2726
<b>Exploiting spatial overlap to efficiently compute appearance distances between image windows,</b> BOGDAN ALEXE, ETH ZÜRICH, VIVIANA PETRESCU, , and VITTORIO FERRARI, University of Edinburgh .....	2735

**Accelerated Adaptive Markov Chain for Partition Function  
Computation**, STEFANO ERMON, CARLA P. GOMES, Cornell University,  
ASHISH SABHARWAL, IBM Watson Research Center, and BART SELMAN,  
Cornell University ..... 2744

**Subject Index**

**Author Index**