

Po-Ling Loh

CONTACT INFORMATION	Statistical Laboratory Centre for Mathematical Sciences Wilberforce Road, Cambridge, CB3 0WB	<i>Phone:</i> +1 (443) 968-5029 <i>E-mail:</i> pll28@cam.ac.uk <i>WWW:</i> https://www.dpmms.cam.ac.uk/~pll28
RESEARCH INTERESTS	High-dimensional statistics, errors-in-variables, graphical models, nonconvex optimization, causality, robust statistics, medical imaging, network inference, differential privacy	
EMPLOYMENT	University of Cambridge <i>Lecturer</i> Department of Pure Mathematics and Mathematical Statistics	January 2021 – present
	Columbia University <i>Visiting Associate Professor</i> Department of Statistics	2019 – 2020
	University of Wisconsin - Madison <i>Associate Professor of Statistics</i> <i>Assistant Professor of Statistics</i> <i>Assistant Professor of Electrical & Computer Engineering</i>	2019 – 2020 2018 – 2019 2016 – 2018
	Secondary appointments: Department of Computer Sciences, Department of Industrial & Systems Engineering Affiliated faculty, Wisconsin Institute for Discovery Affiliated member, UW-Madison Teaching Academy	
	University of Pennsylvania <i>Assistant Professor</i> The Wharton School, Department of Statistics Secondary appointment, Department of Computer & Information Science Affiliated faculty, Warren Center for Network & Data Sciences	2014 – 2016
EDUCATION	University of California, Berkeley Ph.D., Statistics, May 2014 <ul style="list-style-type: none">• Dissertation topic: “High-dimensional statistics with systematically corrupted data”• Adviser: Martin Wainwright M.S., Computer Science, May 2013 Designated Emphasis in Communication, Computation, and Statistics, December 2011	
	California Institute of Technology B.S., Mathematics, minor in English, June 2009	

HONORS AND
AWARDS

ARO Young Investigator Program Award, 2019–2022
IMS Tweedie New Researcher Award, 2019
Bernoulli Society New Researcher Award, 2019
NSF CAREER Award, 2018–2023
Simons Foundation Fellowship, 2018: Awarded to visit Isaac Newton Institute in Cambridge, UK
Erich L. Lehmann Citation, 2014: Berkeley statistics department award for outstanding PhD dissertation in theoretical statistics
Best student paper award, NIPS conference, 2012
Hertz Fellowship, 2009–2014
NSF Graduate Research Fellowship, 2012–2014
National Defense Science & Engineering Graduate Fellowship (NDSEG), 2009–2012
Mabel Beckman Award, 2009: Caltech award for leadership and academic excellence
Axline Merit Award, 2005–2009: Full-tuition scholarship at Caltech
William Lowell Putnam Math Competition Honorable Mention, 2005: Placed 57th in the nation, one of top 5 Caltech students

TEACHING

University of Cambridge

Instructor

2021–present

- Robust Statistics, Lent 2021

Columbia University

Instructor

2019–2020

- STAT 8201: Robustness Theory, Fall 2019
- STAT 8201: Spectral Methods for Machine Learning and Statistics, Spring 2020

University of Wisconsin, Madison

Instructor

2016–2019

- CS/ECE/ME 532: Matrix Methods in Machine Learning, Spring 2019
- STAT 992/CS 838/ECE 901: Robustness Theory, Fall 2018
- ECE 203: Signals, Information, and Computation, Spring 2018 (co-instructor)
- ECE 901: Spectral Methods for Machine Learning and Statistics, Spring 2017
- ECE 331: Probability and Stochastic Processes, Fall 2016 & 2017

University of Pennsylvania

Instructor

2015–2016

- STAT 431: Statistical Inference, Spring 2016
- STAT 430: Probability, Spring 2015

University of California, Berkeley

Graduate Student Instructor

Fall 2011

- Stat 241A/CS 281A: Statistical Learning Theory

California Institute of Technology

Undergraduate Teaching Assistant

2007–2009

- IST 4: Information and Logic, Spring 2008 & 2009
- Math 2b: Probability and Statistics, Winter 2008
- Math 1a: Calculus, Fall 2007

Winner of Associated Students of Caltech Teaching Award, 2008.

FUNDING

Current

- New directions in privacy and robustness. Isaac Newton Trust Research Grant, £24k, 2021–22.

Previous

- Something old, something new: Robust statistics in the 21st century. NSF CAREER, \$400k, 2018–2023.
- Big data analytics solutions to improve power plant efficiency: Online monitoring, visualization, prognosis, and maintenance decision making. DOE NEUP, PI: Kaibo Liu, co-PI: Po-Ling Loh, Todd Allen, \$800k, 2018–2021.
- Developing a theory for function optimization on graphs using local information. NSF EAGER, PI: Varun Jog, co-PI: Po-Ling Loh, \$175k, 2018–2020.
- Resource allocation and statistical estimation in epidemic networks: Scalable algorithms and analysis. ARO YIP, \$360k, 2019–2022.
- Can machines be trusted? Robustification of deep learning for medical imaging. NIH R01, PI: Alan McMillan, co-PI: Po-Ling Loh, Varun Jog, \$1,275k, 2020–2024.

PREPRINTS

Pensia, A., Jog, V., and Loh, P. 2021. Robust regression with covariate filtering: Heavy tails and adversarial contamination. *Submitted*.

Zhang, X., Zhu, X., and Loh, P. 2021. Provable training set debugging for linear regression. *Submitted*.

Liu, Z. and Loh, P. 2021. Robust W-GAN-based estimation under Wasserstein contamination. *Submitted*.

Pensia, A., Tovar Lopez, A., Jog, V., and Loh, P. 2020. Analyzing generalization error of learning algorithms: From information theory to optimal transport. *Submitted*.

Liu, Z., Zhang, J., Jog, V., Loh, P., and McMillan, A. B. 2020. Robustifying deep networks for image segmentation. *Submitted*.

Wang, D. and Loh, P. 2020. Adaptive estimation and statistical inference for high-dimensional graph-based linear models. *Submitted*.

Wang, D., Fu, H., and Loh, P. 2020. Boosting algorithms for estimating optimal individualized treatment rules. *Submitted*.

Loh, P. 2018. Scale calibration for high-dimensional robust regression. *Submitted*.

Khim, J. and Loh, P. 2018. Adversarial risk bounds for binary classification via function transformation.

Khim, J. and Loh, P. 2018. A theory of maximum likelihood for weighted infection graphs.

PUBLICATIONS

Pensia, A., Jog, V., and Loh, P. 2021. Estimating location parameters in sample-heterogeneous distributions. To appear in *Information and Inference*.

- Loh, P. 2021. Searching for structure in complex data: A modern statistical quest. *Snapshots of Modern Mathematics from Oberwolfach*.
- Jog, V. and Loh, P. 2021. Teaching and learning in uncertainty. *IEEE Transactions on Information Theory* 67(1): 598–615.
- Kim, M., Ou, E., Loh, P., Allen, T., and Agasie, R. 2020. RNN-based online anomaly detection in nuclear reactors for highly imbalanced datasets with uncertainty. *Nuclear Engineering and Design* 364: 110699.
- Pensia, A., Jog, V., and Loh, P. 2020. Extracting robust and accurate features via a robust information bottleneck. *IEEE Journal on Selected Areas in Information Theory* 1–14.
- Khim, J. and Loh, P. 2020. Permutation tests for infection graphs. *Journal of the American Statistical Association* 1–13.
- Xu, M., Jog, V., and Loh, P. 2020. Optimal rates for community estimation in the weighted stochastic block model. *Annals of Statistics* 48(1): 183–204.
- Loh, P. 2020. Review of “High-dimensional statistics: A non-asymptotic viewpoint.” *Bulletin of the American Mathematical Society*.
- Yan, W., Loh, P., Li, C., Huang, Y., and Yang, L. 2019. Conquering the worst case of infections in networks. *IEEE Access* 8(1): 2835–2846.
- Charles, Z., Feng, Z., Rajput, S., Loh, P., and Papailiopoulos, D. 2019. Does data augmentation lead to positive margin? *Proceedings of the International Conference on Machine Learning (ICML)*.
- Pensia, A., Jog, V., and Loh, P. 2019. Mean estimation for entangled single-sample distributions. *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*.
- Khim, J., Jog, V., and Loh, P. 2019. Adversarial influence maximization. *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*.
- Mazrooyisebdani, M., Nair, V., Loh, P., Remsik, A., Young, B., Dodd, K., Kang, T., Williams, J., and Prabhakaran, V. 2018. Evaluation of changes in motor network following BCI therapy based on graph theory analysis. *Frontiers in Neuroscience* 12: 861.
- Pydi, M., Jog, V., and Loh, P. 2018. Graph-based ascent algorithms for function maximization. *Proceedings of the 56th Annual Allerton Conference on Communication, Control, and Computing*.
- Feng, Z. and Loh, P. 2018. Online learning with graph-structured feedback against adaptive adversaries. *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*.
- Pensia, A., Jog, V., and Loh, P. 2018. Generalization error bounds for noisy, iterative algorithms. *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*.
- Loh, P. and Tan, X. 2018. High-dimensional robust precision matrix estimation: Cellwise corruption under ϵ -contamination. *Electronic Journal of Statistics* 12(1): 1429–1467.
- Ko, J., Baldassano, S., Loh, P., Kording, K., Litt, B., and Issadore, D. 2018. Machine learning to detect signatures of disease in liquid biopsies – A user’s guide. *Lab on a Chip* 18(3): 395–405.
- Jog, V. and Loh, P. 2018. Persistence of centrality in random growing trees. *Random Structures*

and *Algorithms* 52(1): 136–157.

Loh, P. and Wainwright, M. J. 2017. Support recovery without incoherence: A case for nonconvex regularization. *Annals of Statistics* 45(6): 1–28.

Loh, P. 2017. On lower bounds for statistical learning theory. *Entropy* 19(11), 617. Feature paper, special issue on information theory in machine learning and data science.

Wibisono, A., Jog, V., and Loh, P. 2017. Information and estimation in Fokker-Planck channels. *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*.

Loh, P. 2017. Statistical consistency and asymptotic normality for high-dimensional robust M -estimators. *Annals of Statistics* 45(2): 866–896.

Jog, V. and Loh, P. 2016. Analysis of centrality in sublinear preferential attachment trees via the CMJ branching process. *IEEE Transactions on Network Science and Engineering* 4(1):1–12.

Khim, J. and Loh, P. 2016. Confidence sets for the source of a diffusion in regular trees. *IEEE Transactions on Network Science and Engineering* 4(1): 27–40.

Khim, J., Jog, V., and Loh, P. 2016. Computing and maximizing influence in linear threshold and triggering models. *Proceedings of the Neural Information Processing Systems (NIPS) Conference*.

Cheng, M., Sriramulu, A., Muralidhar, S., Loo, B., Huang, L., and Loh, P. 2016. Collection, exploration and analysis of crowdfunding social networks. *Proceedings of the Third International Workshop on Exploratory Search in Databases and the Web (ExploreDB)*.

Jog, V. and Loh, P. 2015. Recovering communities in weighted stochastic block models. *Proceedings of the 53rd Annual Allerton Conference on Communication, Control, and Computing*.

Jog, V. and Loh, P. 2015. On model misspecification and KL separation for Gaussian graphical models. *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*.

Loh, P. and Wainwright, M. J. 2015. Regularized M -estimators with nonconvexity: Statistical and algorithmic theory for local optima. *Journal of Machine Learning Research* 16 (2015) 559–616.

Loh, P. and Wibisono, A. 2014. Concavity of reweighted Kikuchi approximation. *Proceedings of the Neural Information Processing Systems (NIPS) Conference*.

Loh, P. and Bühlmann, P. 2014. High-dimensional learning of linear causal networks via inverse covariance estimation. *Journal of Machine Learning Research* 15 (2014) 3065–3105.

Loh, P. and Nowozin, S. 2013. Faster Hoeffding racing: Bernstein races via jackknife estimates. *Proceedings of the Algorithmic Learning Theory (ALT) Conference*.

Loh, P. and Wainwright, M. J. 2013. Structure estimation for discrete graphical models: Generalized covariance matrices and their inverses. *Annals of Statistics* 41(6): 3022–3049.

Loh, P. and Wainwright, M. J. 2012. Corrupted and missing predictors: Minimax bounds for high-dimensional linear regression. *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*.

Loh, P. and Wainwright, M. J. 2011. High-dimensional regression with noisy and missing data: Provable guarantees with nonconvexity. *Annals of Statistics* 40(3): 1637–1664.

- Loh, P., Zhou, H., and Bruck, J. 2009. The robustness of stochastic switching networks. *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*.
- BOOK CHAPTERS
- Loh, P. 2020. Modern perspectives on robust statistics. To appear in *IISA Series on Statistics and Data Science*, Eds. Ghosh, S., Datta, S., and Kundu, S.
- Loh, P. 2019. Neighborhood selection methods. In *Handbook of Graphical Models*, Eds. Maathuis, M., Drton, M., Lauritzen, S., and Wainwright, M.
- CONFERENCE ABSTRACTS
- Shen, M., Kabra, P., Periyasamy, S., Ozturk, M., Jog, V., Loh, P., Laeseke, P., Morgan, D., and Kleedehn, M. 2020. Identifying active extravasation on arteriograms using artificial intelligence. *Presented at the Society of Interventional Radiology (SIR) annual meeting*.
- Hou, A. and Loh, P. 2019. Robust budget allocation. *Presented at the SIAM workshop on network science*.
- Shen, M., Ozturk, M., Loh, P., Jog, V., Laeseke, P., Morgan, D., and Kleedehn, M. 2019. Identifying active extravasation on arteriograms using artificial intelligence. *Presented at the Cardiovascular and Interventional Radiological Society of Europe (CIRSE)*.
- Park, C., Chen, W., Loh, P., Perlman, S., Kim, D., Robbins, J., and McMillan, A. 2019. Low dose PET/MR imaging in Crohn's disease. *Presented at the International Society for Magnetic Resonance in Medicine (ISMRM)*.
- DEPARTMENT SEMINARS
- A modern take on Huber regression. Biostatistics colloquium, Columbia University, 2021.
- Estimating location parameters in entangled single-sample distributions. Statistics seminar, Harvard University, 2020.
- Estimating location parameters in entangled single-sample distributions. DPMMS seminar, University of Cambridge, 2020.
- Estimating location parameters in entangled single-sample distributions. Statistics colloquium, University of Virginia, 2019.
- Estimating location parameters in entangled single-sample distributions. Math and data seminar, New York University, 2019.
- Mean estimation for entangled single-sample distributions. Statistics seminar, USC Marshall, 2019.
- Mean estimation for entangled single-sample distributions. Statistics seminar, Penn State University, 2019.
- Statistical inference for infectious disease modeling. Statistics seminar, UC Berkeley, 2019.
- Statistical inference for infectious disease modeling. Statistics seminar, Cornell University, 2019.
- Statistical inference for infectious disease modeling. Statistics seminar, Duke University, 2019.
- Statistical inference for infectious disease modeling. Statistics seminar, Columbia University, 2018.
- Two inference problems for network contagion. ORFE colloquium, Princeton University, 2018.

Statistical inference for infectious disease modeling. Statistics seminar, University of Michigan, 2018.

Scale calibration for high-dimensional robust regression. Statistics seminar, University of Wisconsin - Madison, 2018.

Statistical inference for infectious disease modeling. ORIE colloquium, Cornell University, 2018.

Statistical inference for infectious disease modeling. Statistics and biostatistics seminar, Rutgers University, 2018.

Estimation and inference for network contagion. Probability seminar, McGill University math department, Montreal, Canada, 2018.

Statistical inference for infectious disease modeling. Statistics colloquium, University of Iowa statistics department, 2018.

Influence maximization in stochastic and adversarial settings. London School of Economics statistics department, UK, 2018.

Statistical inference for infectious disease modeling. Statistical science seminar, University College London math department, UK, 2018.

Theory for local optima of nonconvex high-dimensional M -estimators. Cantab Capital Institute for Mathematics of Information seminar, Cambridge University math department, UK, 2018.

Statistical inference for infectious disease modeling. Statistics seminar, University of York math department, UK, 2018.

Modeling disease propagation in networks: Source-finding and influence maximization. Machine learning seminar, University of Bristol, UK, 2018.

Statistical inference for infectious disease modeling. Microsoft Distinguished Seminar, University of Washington statistics department, 2017.

Statistical inference for infectious disease modeling. Stochastics seminar, Georgia Tech math department, 2017.

Theory for local optima of nonconvex high-dimensional M -estimators. Department of Electrical and Computer Engineering seminar, Instituto Superior Técnico, Lisbon, Portugal, 2017.

Modeling disease propagation in networks: Source-finding and influence maximization. Department of Statistics and Applied Probability seminar, National University of Singapore, 2017.

High-dimensional robust regression. Econometrics and statistics workshop, University of Chicago Booth School of Business, 2017.

Confidence sets for the source of a diffusion in regular trees. Probability seminar, UW-Madison math department, 2017.

Influence maximization in stochastic and adversarial settings. Systems science seminar, ECE department, University of Michigan, 2017.

Influence maximization in stochastic and adversarial settings. Research seminar, Seminar für Statistik, ETH Zürich, 2017.

From information to estimation in stochastic differential equations. Statistics seminar, EPFL, 2017.

Influence maximization in stochastic and adversarial settings. Stochastics and statistics seminar, MIT, 2016.

Modeling disease propagation in networks: Source-finding and influence maximization. Statistics seminar, Yale University, 2016.

Confidence sets for the source of a diffusion in regular trees. Probability seminar, University of Delaware math department, 2016.

New perspectives for robust/high-dimensional estimation. Statistics seminar, USC Marshall School, 2016.

New perspectives in robust, high-dimensional statistics. Systems, Information, Learning, and Optimization (SILO) seminar, UW-Madison, 2016.

On centrality in random growing trees: Confidence sets for source estimators and persistence. Wilks statistics seminar, Princeton University, 2016.

New perspectives for robust/high-dimensional estimation. Research seminar, Seminar für Statistik, ETH Zürich, 2015.

New perspectives for robust/high-dimensional estimation. Research seminar, Max Planck Institute for Intelligent Systems, Tübingen, Germany, 2015.

High-dimensional robust regression. Statistics department seminar, Harvard University, 2015.

High-dimensional robust regression. Department of Statistics and Applied Probability seminar, National University of Singapore, 2015.

PDW methods for support recovery in nonconvex high-dimensional problems. Statistics department seminar, University of Washington, 2015.

PDW methods for support recovery in nonconvex high-dimensional problems. ORFE optimization seminar, Princeton University, 2015.

What can be learned with linear regression. PennSIVE/Biostats seminar, University of Pennsylvania, 2014.

Nonconvex methods for high-dimensional regression with noisy and missing data. Electrical engineering seminar, Caltech, 2014.

Nonconvex methods for high-dimensional regression with noisy and missing data. WID seminar, UW-Madison, 2014.

Nonconvex methods for high-dimensional regression with noisy and missing data. Statistics seminar, Yale University, 2014.

Nonconvex methods for high-dimensional regression with noisy and missing data. Statistics seminar, Carnegie Mellon University, 2014.

Nonconvex methods for high-dimensional regression with noisy and missing data. ORFE seminar, Princeton University, 2014.

Nonconvex methods for high-dimensional regression with noisy and missing data. ORC seminar, MIT, 2014.

Nonconvex methods for high-dimensional regression with noisy and missing data. Statistics seminar, University of Chicago, 2014.

Nonconvex methods for high-dimensional regression with noisy and missing data. ORIE/statistics seminar, Cornell University, 2014.

Nonconvex methods for high-dimensional regression with noisy and missing data. Statistics seminar, Columbia University, 2014.

Nonconvex methods for high-dimensional regression with noisy and missing data. Statistics seminar, University of Michigan, 2014.

Nonconvex methods for high-dimensional regression with noisy and missing data. Statistics and scientific computation seminar, UT Austin, 2014.

Nonconvex methods for high-dimensional regression with noisy and missing data. Statistics seminar, University of Pennsylvania, 2014.

Nonconvex methods for high-dimensional regression with noisy and missing data. Statistics seminar, Stanford University, 2014.

Nonconvex methods for high-dimensional regression with noisy and missing data. Statistics seminar, UC Davis, 2014.

Beyond the graphical Lasso: Structure learning via inverse covariance estimation. Research seminar, Department of Mathematics and Computer Science, Universität Basel, 2013.

Local optima of nonconvex regularized M -estimators. Research seminar, Seminar für Statistik, ETH Zürich, 2013.

Beyond the graphical Lasso: Structure learning via inverse covariance estimation. Research seminar, Max Planck Institute for Intelligent Systems, Tübingen, Germany, 2013.

Local optima of nonconvex regularized M -estimators. Department of Statistics and Applied Probability seminar, National University of Singapore, 2013.

Learning with systematic corruptions: Regression-based methods with applications to MRI and graph estimation. Systems, Information, Learning, and Optimization (SILO) seminar, UW-Madison, 2012.

INVITED TALKS

Robust regression with covariate filtering: Heavy tails and adversarial contamination. Conference on Information Sciences and Systems (CISS), 2021.

Teaching and learning in uncertainty. Special session on graphs in data science, American Mathematical Society (AMS) Spring Southeastern Virtual Sectional Meeting, 2021.

Estimating location parameters in entangled single-sample distributions. Algorithmic Learning Theory (ALT) mentoring workshop, 2021.

Statistical inference for infectious disease modeling. Joint probability seminar, Universities of Frankfurt/Mainz/Darmstadt, 2021.

A tutorial on robust statistics. Machine Learning in Science and Engineering Conference, Data Science Institute, Columbia University, 2020.

Robustifying machine learning algorithms. Computing in Engineering Forum, UW-Madison, Madison, WI, 2020.

Extracting robust and accurate features via a robust information bottleneck. Simons Institute deep learning reunion workshop, Berkeley, CA, 2020.

Discussant, JASA Theory & Methods invited session, Joint Statistical Meetings (JSM), Philadelphia, PA, 2020.

Extracting robust and accurate features via a robust information bottleneck. Statistics meets machine learning, MFO workshop, Oberwolfach, Germany, 2020.

Robust information bottleneck. NeurIPS workshop on information theory and machine learning, Vancouver, Canada, 2019.

Estimating location parameters in entangled single-sample distributions. Conference on big data and machine learning in econometrics, finance, and statistics. Stevanovich Center for Financial Mathematics, University of Chicago, Chicago, IL, 2019.

Robust information bottleneck. Workshop on recent trends in clustering and classification. Toyota Technological Institute, Chicago, IL, 2019.

Teaching and learning in uncertainty. Workshop on estimation of entropy and other functionals: Statistics meets information theory. University of Cambridge, Cambridge, UK, 2019.

Optimization challenges from robust statistics. Bernoulli Society new researchers' session. World Statistics Congress, Kuala Lumpur, Malaysia, 2019.

Robust information bottleneck. Program on deep learning opening workshop. Statistics and Applied Mathematical Sciences Institute (SAMSI), Durham, NC, 2019.

Mean estimation for entangled single-sample distributions. Joint Statistical Meetings (JSM), Denver, CO, 2019.

Statistical inference for infectious disease modeling. Tweedie lecture, IMS New Researcher's Conference, Fort Collins, CO, 2019.

Mean estimation for entangled single-sample distributions. Workshop on recent themes in resource tradeoffs: Privacy, fairness, and robustness, Institute for Mathematics and Its Applications (IMA), University of Minnesota, Minneapolis, MN, 2019.

Teaching and learning in uncertainty. London Symposium on Information Theory, Kings College London, UK, 2019.

Mean estimation for entangled single-sample distributions. Conference on Information Sciences and Systems (CISS), Johns Hopkins University, Baltimore, MD, 2019.

The not-so-rough landscape of nonconvex M -estimators, Workshop on rough landscapes: From physics to algorithms, Kavli Institute for Theoretical Physics (KITP), Santa Barbara, CA, 2019.

Theory of local optima of nonconvex high-dimensional M -estimators (short course), Tata Institute

for Fundamental Research (TIFR), Mumbai, India, 2019.

Data science for networked data. Women in Machine Learning workshop, Montreal, Canada, 2018.

Robust M -estimation for high-dimensional regression. Larry Brown memorial workshop, University of Pennsylvania, Philadelphia, PA, 2018.

Scale calibration for high-dimensional robust regression. Workshop on robust and high-dimensional statistics, Simons Institute, Berkeley, CA, 2018.

Two results on generalization error. Workshop on theoretical foundations of deep learning, Georgia Tech, Atlanta, GA, 2018.

Scale calibration for high-dimensional robust regression. Workshop on higher-order asymptotics and post-selection inference, St. Louis, MO, 2018.

Robust estimation via (non)convex M -estimation. Workshop on computational efficiency and high-dimensional robust statistics, Toyota Technological Institute, Chicago, IL, 2018.

Robust estimation, efficiency, and Lasso debiasing. Joint Statistical Meetings (JSM), Vancouver, BC, 2018.

A theory of maximum likelihood for weighted infection graphs. IMS Vilnius meeting, Vilnius, Lithuania, 2018.

Two inference problems for network contagion. IMS Asia Pacific Rim meeting, Singapore, 2018.

Robust estimation, efficiency, and Lasso debiasing. International Chinese Statistical Association (ICSA) conference, New Brunswick, NJ, 2018.

New algorithms for resource allocation in networks. Midwest statistical machine learning colloquium, Iowa State University, Ames, IA, 2018.

Two inference problems for network contagion. CRM workshop on combinatorial statistics, Montreal, Canada, 2018.

New algorithms for resource allocation in networks. BRANDS workshop, USC Marshall, Los Angeles, CA, 2018.

Statistical inference for infectious disease modeling. Seminar for coding theory for inference, learning, and optimization, Dagstuhl, Germany, 2018.

Data science for networked data. Workshop on big data and the role of statistical scalability, Isaac Newton Institute, Cambridge, UK, 2018.

New perspectives for robust/high-dimensional estimation. Econometric Institute workshop on robust statistics, Erasmus University, Rotterdam, Netherlands, 2018.

Influence maximization in stochastic and adversarial settings. STOR-i annual conference, Lancaster University, UK, 2018.

Statistical inference for infectious disease modeling. Workshop on modern statistics for complex data structures, King Abdullah University of Science and Technology, Saudi Arabia, 2017.

Optimization challenges from robust statistics. CMO-BIRS workshop on emerging challenges in data science, Oaxaca, Mexico, 2017.

Permutation tests for infection graphs. Allerton conference on communication, control, and computing, Monticello, IL, 2017.

Optimal rates for community estimation in the weighted stochastic block model. Eurandom workshop on community detection and network reconstruction, Eindhoven, Netherlands, 2017.

Nonconvex optimization for statistical estimation. OpenAI, San Francisco, CA, 2017.

Statistical inference for disease modeling. Institute for Disease Modeling, Seattle, WA, 2017.

Robust estimation, efficiency, and Lasso debiasing. Workshop on higher-order asymptotics and post-selection inference, St. Louis, MO, 2017.

Modeling disease propagation in networks. Joint Statistical Meetings (JSM), Baltimore, MD, 2017.

Source finding for disease propagation in networks. ISI World Statistics Congress, Marrakech, Morocco, 2017.

Community recovery in weighted stochastic block models. Workshop on scalable statistical inference, Isaac Newton Institute, Cambridge, UK, 2017.

High-dimensional robust regression. IMS China international conference on stochastics and probability, Nanning, China, 2017.

Robust high-dimensional linear regression: A statistical perspective. STOC workshop on robustness and nonconvexity, Montreal, Canada, 2017.

High-dimensional robust regression. Conference on nonconvex statistical learning, University of Southern California, 2017.

Theory for local optima in nonconvex regression. Minisymposium on nonconvex optimization in data analysis, SIAM OP17 conference, Vancouver, BC, 2017.

Source estimation and centrality in random growing trees. Probability and combinatorics: Twelfth annual workshop, Holetown, Barbados, 2017.

Provable bounds for influence maximization. International Chinese Statistical Association (ICSA) conference, Shanghai, China, 2016.

Confidence sets for the source of a diffusion in regular trees. International Conference on Data Science, Fudan University, 2016.

Source estimation and centrality in random growing trees. NIPS workshop on learning in high dimensions with structure, Barcelona, Spain, 2016.

Computing and maximizing influence in linear threshold and triggering models. Institute for Operations Research and the Management Sciences (INFORMS) conference, Nashville, TN, 2016.

Adversarial influence maximization. Workshop on theoretical foundations for learning from easy data, Lorentz Center, Leiden, Netherlands, 2016.

Computing and maximizing influence in linear threshold and triggering models. Penn Institute for Computational Science (PICS) symposium on emerging paradigms in scientific discovery, University of Pennsylvania, 2016.

Computing and maximizing influence in linear threshold and triggering models. Allerton conference on communication, control, and computing, Monticello, IL, 2016.

Computing and maximizing influence in linear threshold and triggering models. International Indian Statistical Association (IISA) conference, Corvallis, OR, 2016.

What can be learned with the graphical Lasso: Generalizations of the Gaussian assumption. Joint Statistical Meetings (JSM), Chicago, IL, 2016.

High-dimensional learning of linear causal networks via inverse covariance estimation. CRM workshop on statistical causal inference and its applications to genetics, Montreal, Canada, 2016.

Robust techniques in high-dimensional statistics. IMS World Congress on Probability and Statistics, Toronto, Canada, 2016.

High-dimensional consistency of robust precision matrix estimators. International Conference on Robust Statistics (ICORS), Geneva, Switzerland, 2016.

High-dimensional robust regression. ICML workshop on advances in nonconvex analysis and optimization, New York, NY, 2016.

Community estimation in weighted SBMs. New England Statistics Symposium, Yale University, 2016.

What can be learned with the graphical Lasso: Generalizations of the Gaussian assumption. Probabilistic graphical model workshop, Institute of Statistical Mathematics (ISM), Tokyo, Japan, 2016.

PDW methods for support recovery in nonconvex high-dimensional problems. Conference on Information Sciences and Systems (CISS), Princeton, NJ, 2016.

On centrality in random growing trees: Confidence sets for source estimators and persistence. Computationally and statistically efficient inference for complex large-scale data, MFO workshop, Oberwolfach, Germany, 2016.

Joint estimation of location and scale for robust high-dimensional linear regression. Research seminar, Research Center for Statistics, Université de Genève, 2016.

Confidence sets for the source of a diffusion in regular trees. Physics informed machine learning workshop, Santa Fe, NM, 2016.

High-dimensional consistency of robust precision matrix estimators. International Indian Statistical Association (IISA) conference, Pune, India, 2015.

Recent advances in high-dimensional robust regression. Computational and Methodological Statistics (CMStatistics) conference, London, UK, 2015.

High-dimensional precision matrix estimation: Cellwise corruption under epsilon-contamination. Workshop on current and future challenges in robust statistics, Banff, Canada, 2015.

Robust methods for high-dimensional regression. Institute for Operations Research and the Man-

agement Sciences (INFORMS) conference, Philadelphia, PA, 2015.

Community detection in weighted stochastic block models. Allerton conference on communication, control, and computing, Monticello, IL, 2015.

On model misspecification and KL separation for Gaussian graphical models. SIAM conference on applied algebraic geometry, Daejeon, Korea, 2015.

Concavity of reweighted Kikuchi approximation. International Symposium on Mathematical Programming (ISMP) conference, Pittsburgh, PA, 2015.

On model misspecification and KL separation for Gaussian graphical models. International Symposium on Information Theory (ISIT) conference, Hong Kong, 2015.

PDW methods for support recovery in nonconvex high-dimensional problems. Workshop on methodological advances in statistics related to big data, Cantabria, Spain, 2015.

High-dimensional robust regression. Probabilistic techniques in modern statistics, MFO workshop, Oberwolfach, Germany, 2015.

High-dimensional robust M -estimators. Workshop on statistical inference for large-scale data, Simon Fraser University, Vancouver, Canada, 2015.

PDW methods for support recovery in nonconvex high-dimensional problems. Workshop on optimization and matrix methods in big data, Fields Institute, Toronto, Canada, 2015.

Local optima of nonconvex M -estimators. International Conference on Robust Statistics (ICORS), Kolkata, India, 2015.

Local optima of nonconvex M -estimators. CIRM meeting in mathematical statistics, Luminy, France, 2014.

Local optima of nonconvex M -estimators. NIPS workshop on robustness, Montreal, Canada, 2014.

Beyond the graphical Lasso: Structure learning via inverse covariance estimation. International Indian Statistical Association (IISA) conference, Riverside, CA, 2014.

Beyond the graphical Lasso: Structure learning via inverse covariance estimation. ICML workshop on covariance selection and graphical model structure learning, Beijing, China, 2014.

High-dimensional learning of linear causal networks via inverse covariance estimation. ICML workshop on causal modeling and machine learning, Beijing, China, 2014.

Regularized M -estimators with nonconvexity: Statistical and algorithmic theory for local optima. Spotlight presentation at Neural Information Processing Systems (NIPS) conference, Lake Tahoe, NV, 2013.

Faster Hoeffding racing: Bernstein races via jackknife estimates. Algorithmic Learning Theory (ALT) conference, Singapore, 2013.

Recent results on undirected graph estimation: Corruption and non-Gaussianity. Seminar für Statistik lunch seminar, ETH Zürich, 2013.

No voodoo here! Learning discrete graphical models via inverse covariance estimation. Oral pre-

resentation at Neural Information Processing Systems (NIPS) conference, Lake Tahoe, NV, 2012. **Winner of best student paper award.**

Corrupted and missing predictors: Minimax bounds for high-dimensional linear regression. International Symposium on Information Theory (ISIT) conference, Cambridge, MA, 2012.

High-dimensional regression with noisy and missing data. Oral presentation at Neural Information Processing Systems (NIPS) conference, Granada, Spain, 2011.

OTHER INVITED
WORKSHOPS

Semester program on computational complexity of statistical inference (long-term participant). Simons Institute, Berkeley, CA, 2021.

Summer program on foundations of deep learning (long-term participant). Simons Institute, Berkeley, CA, 2019.

Workshop on statistical and computational aspects of learning with complex structure. Oberwolfach, Germany, 2019.

Probability and combinatorics: Thirteenth and fourteenth annual workshops. Holetown, Barbados, 2018 & 2019.

Sackler Forum on machine learning. National Academy of Sciences, Washington, DC, 2017.

Workshop on inference in high dimensional regression. American Institute of Mathematics, San Jose, CA, 2015.

OTHER
PRESENTATIONS

Data science for networked data. Big Data Meetup, Madison, WI, 2019.

Estimation and inference in engineering systems. UW Advance: Big Data seminar, 2018 & 2019.

Illuminating medical imaging. SoundWaves event, Wisconsin Institute for Discovery, 2018.

Statistical inference for infectious disease modeling. Shannon Channel seminar, 2018.

Data science for networked data. Women in Data Science event, UW-Madison, 2018.

Networking and networks in big data. Women in Big Data Midwest Chapter kickoff, Madison, WI, 2017.

Mathematical models for network science. Grainger Institute for Engineering early career faculty seminar series, UW-Madison, 2017.

The mathematics of network science. Madison Area Technical College (MATC) math club lecture, Madison, WI, 2017.

Computationally efficient influence maximization in stochastic and adversarial models, Welling group meeting presentation, University of Amsterdam, 2016.

Learning with systematically corrupted data. Rising Stars in EECS Workshop, MIT, Cambridge, MA, 2013.

Compressed sensing with systematically corrupted data. Sense and Sense-abilities seminar, Institute for Infocomm Research, Singapore, 2013.

Graph learning with corruptions. Berkeley Statistics Annual Research Symposium (BSTARS), 2013.

Inverse covariance matrix estimation for Gaussian graphical models in the presence of noise. Women and Math (WaM) program at Institute for Advanced Study, Princeton, NJ, 2011.

PROFESSIONAL
SERVICE

Institute of Mathematical Statistics (IMS):

- Committee on special lectures, 2020–2023
- Committee on nominations, 2017–2018
- Committee on new researchers, 2014–2017

American Statistical Association (ASA):

- Student paper award committee, Section on Learning and Data Science, 2019–2020
- Publications officer, Section on Nonparametric Statistics (elected position), 2017–2020

Associate editor:

- Statistica Sinica, 2017–present
- Journal of the American Statistical Association (JASA), 2019–present
- Foundations of Data Science, 2019–present
- Annals of the Institute of Statistical Mathematics (AISM), 2020–2024
- New England Journal of Statistics in Data Science, 2021–present
- Book Reviews of the American Mathematical Society, 2021–present
- Journal on Selected Areas in Information Theory (JSAIT) special issue, “Statistical inference and estimation,” 2020 (guest editor)

Advisory board member:

- Learning Theory Mentorship Programs, 2021–present
- Women in Machine Learning, 2020–2023
- STOR-i Centre for Doctoral Training, Lancaster University, 2019–present
- Midwest Machine Learning Symposium, 2018–present
- StatScale program, Engineering and Physical Sciences Research Council (EPSRC), 2016–2021

Grant reviewer:

- Natural Sciences and Engineering Research Council of Canada (one instance)
- Swiss National Science Foundation (two instances)
- European Science Foundation (one instance)
- Dutch Research Council (one instance)
- National Institutes of Health (one instance)
- National Science Foundation (five instances)

Program committee: INFORMS Optimization Society (IOS) conference, 2016; Neural Information Processing Systems (NIPS) conference, 2016; Artificial Intelligence and Statistics (AISTATS) conference, 2017; Women in Machine Learning (WiML) workshop, 2017; Conference on Learning Theory (COLT), 2018; International Indian Statistical Association (IISA) conference, 2018; Algorithmic Learning Theory (ALT), 2019; International Conference on Machine Learning (ICML), 2017, 2018, 2019, 2021; Foundations of Data Science (FODS) conference, 2020; NeurIPS workshops, 2020; International Chinese Statistical Association (ICSA) conference, 2021.

Journal reviewer: Annals of Statistics; Journal of Machine Learning Research; Journal of the Royal Statistical Society: Series B; Statistica Sinica; Journal of the American Statistical Association; Biometrika; Electronic Journal of Statistics; Machine Learning Journal; Annals of the Institute of Statistical Mathematics; Bernoulli; Journal of Statistical Planning and Inference; Mathematical Programming; IEEE Transactions on Information Theory; Transactions on Pattern Analysis and Machine Intelligence; Probability Theory and Related Fields; IEEE Transactions on Networking; Annals of Applied Probability; Behaviormetrika; Algorithms; Operations Research

Conference reviewer: Conference on Learning Theory (COLT), 2013; International Conference on Machine Learning (ICML), 2013, 2014, 2015, 2016; Neural Information Processing Systems (NIPS), 2013, 2014, 2015; Uncertainty in Artificial Intelligence (UAI), 2015, 2016; Artificial Intelligence and Statistics (AISTATS), 2015, 2016; International Symposium on Information Theory (ISIT), 2016

CONFERENCE AND
WORKSHOP
ORGANIZATION

Invited session on theoretical advances in deep learning, Joint Statistical Meetings (JSM) conference, Philadelphia, PA, 2020.

International Conference on Machine Learning (workshop co-chair), Vienna, Austria, 2020.

Invited session on robust statistics, Joint Statistical Meetings (JSM) conference, Denver, CO, 2019.

European Conference on Machine Learning & Principles and Practice of Knowledge Discovery in Databases (ECML PKDD) journal track chair, Würzburg, Germany, 2019.

First machine learning for medical imaging (ML4MI) workshop (co-organizer), UW-Madison, 2018

Midwest Machine Learning Symposium (co-founder and lead organizer), Chicago, IL, 2017 & 2018

Invited session on information theory and statistics, IMS Asia-Pacific Rim meeting, Singapore, 2018

Invited session on estimation and inference in high dimensions, International Indian Statistical Association (IISA) conference, Gainesville, FL, 2018

Invited session on statistical tools for medical imaging, International Indian Statistical Association (IISA) conference, Gainesville, FL, 2018

Isaac Newton Institute workshop on big data challenges: Heterogeneity, model misspecification, and changepoints, Windermere, UK, 2018

Dagstuhl Seminar on coding theory for inference, learning, and optimization, Dagstuhl, Germany, 2018

Statistical Scalability Programme, Isaac Newton Institute, Cambridge, UK, 2018

Invited session on machine learning, optimization, and statistics, International Indian Statistical Association (IISA) conference, Hyderabad, India, 2017

Invited session on scalable methods for complex data, Computational and Methodological Statistics (CMStatistics) conference, London, UK, 2017

(On behalf of Bernoulli Society) invited session on visualization and analysis of modern data, Conference on Data Science, Statistics & Visualization, Lisbon, Portugal, 2017

ADVISING

PhD students: Justin Khim (University of Pennsylvania, statistics), 2014–2019, Michael Liou

(UW-Madison, statistics), 2018–present, Zheng Liu (UW-Madison, statistics), 2017–present, Ankit Pensia (UW-Madison, CS), 2017–present, Duzhe Wang (UW-Madison, statistics), 2017–2020, Xiaomin Zhang (UW-Madison, CS), 2017–present

Visiting scholars: Wen Yan (Southeast University, information and communication engineering), 2017–2018

Postdoctoral researchers: Min Xu (University of Pennsylvania, statistics), 2015–2018, Andre Wibisono (UW-Madison, ECE), 2016–2018, Deepanshu Vasal (UW-Madison, ECE), 2016–2017

Masters students: Zhili Feng (UW-Madison, CS), 2017–2019, Ashley Hou (UW-Madison, ECE), 2017–2019, Elisa Ou (UW-Madison, ECE), 2017–2019, Muni Pydi (UW-Madison, ECE), 2017–2019

Thesis committee member: Kam Chung Wong (University of Michigan, statistics), Lohith Kini (University of Pennsylvania, MD/PhD), Xin Lu Tan (University of Pennsylvania, statistics), Steven Baldassano (University of Pennsylvania, MD/PhD), Soovin Yoon (UW-Madison, ISyE), Yilin Zhang (UW-Madison, statistics), Ching-pei Lee (UW-Madison, CS), Hyebin Song (UW-Madison, statistics), Changyue Song (UW-Madison, ISyE), Tan Nguyen (Queensland University of Technology, CS), Yuchen Zhou (UW-Madison, statistics), Aubrey Barnard (UW-Madison, CS), Timothée Mathieu (ENSAE), Fan Chen (UW-Madison, statistics)

Undergraduate students: Nicole Berkman (University of Pennsylvania, math), Laura Luo (University of Pennsylvania, finance), Lavi Ben Dor (University of Pennsylvania, math & business), Linden Yuan (University of Maryland, math)

High school students: Mark Yuen (Pioneer High School, San Jose, CA)

OUTREACH

Faculty mentor, UW-Madison Women’s Faculty Mentoring Program, 2020.

Member, UW-Madison Research Committee for Physical Sciences Division, 2019–2020.

Co-organizer, UW-Madison Machine Learning for Medical Imaging initiative, 2018–2020.

Committee member, Wisconsin Mathematics, Science & Engineering Talent Search, 2017–2021.

Organizer, Grainger Institute for Engineering early career faculty seminar series, 2017.

Interviewer, Hertz Foundation fellowship, 2016–present.

Guest lecturer, Madison Area Technical College (MATC) math club, 2017

Guest lecturer, Program in Algorithmic and Combinatorial Thinking (PACT) summer program, Princeton University, 2015

INTERNSHIPS

Institute for Mathematical Research (FIM), Zürich, Switzerland

Academic guest

August 2013 - November 2013

Short-term visitor, hosted by Peter Bühlmann and Sara van de Geer at ETH Zürich.

Microsoft Research, Cambridge, UK

Graduate student intern

May 2012 - August 2012

Research intern in Machine Learning & Perception group. Developed methods for speeding up decision tree learning for the Microsoft Kinect.

Sandia National Labs, Livermore, CA, USA

Graduate student intern

August 2011

Investigated techniques for MCMC sampling along low-dimensional manifolds for uncertainty quantification in chemical systems.

OTHER ACTIVITIES AIDS/LifeCycle participant, 2013–2016: Cyclist in 2013 and 2014, volunteer roadie in 2015 and 2016.

Berkeley Statistics Graduate Student Association (SGSA) co-president, 2012–2013.

Intern at Fresno Rescue Mission, Summer 2009: Worked with homeless and formerly incarcerated adults in drug rehab program.

Upper-class counselor at Caltech, 2006–2009: Hall counselor for undergraduates in Avery House.