

University of Wisconsin–Madison
Department of Statistics
1300 University Ave
1220 Medical Sciences Center
Madison, WI, 53706
✉ kdlevin@wisc.edu
📄 pages.stat.wisc.edu/~kdlevin

Keith D. Levin

Research Interests

Statistical methods for network data; high-dimensional statistics; concentration inequalities; randomized linear algebra; streaming and sketching algorithms; applications to neuroscience.

Education

- 2017 **Ph.D., Computer Science**, Johns Hopkins University, Baltimore, MD.
Thesis: *Graph Inference with Applications to Low-Resource Audio Search and Indexing*, under supervision of Professors Carey E. Priebe (JHU Applied Mathematics and Statistics), Vincent Lyzinski (University of Maryland Mathematics and Statistics) and Ben Van Durme (JHU Computer Science)
- 2016 **M.S.E., Computer Science**, Johns Hopkins University, Baltimore, MD.
- 2011 **B.S., Linguistics & Psychology**, Northeastern University, Boston, MA.

Positions

- 2020–Present **Assistant Professor**, University of Wisconsin–Madison Department of Statistics, Madison, WI.
- 2017–2020 **Postdoctoral Research Fellow**, University of Michigan Department of Statistics, Ann Arbor, MI.
Statistical methods for analyzing multiple networks, supervised by Elizaveta Levina.
- 2017 **Postdoctoral Researcher**, Johns Hopkins University Department of Applied Mathematics and Statistics, Baltimore, MD.
Network inference: theory, methods and applications, supervised by Carey E. Priebe.
- 2012–2017 **Graduate Student**, Johns Hopkins University Department of Computer Science, Baltimore, MD.
Ph.D. research on applying graph inference techniques to large-scale search problems in speech recognition. Additional research on streaming clustering algorithms.
- 2010–2012 **Avoke Caller Experience Analyst**, Raytheon BBN Technologies, Cambridge, MA.
Compiled and analyzed call center data; prototyped topic models for call categorization; prototyped keyword spotting system for event detection.
- 2007–2011 **Research Assistant**, Sentence Processing Laboratory, Department of Psychology, Northeastern University, Boston, MA.
Ran experiments, designed experimental stimuli, gathered and processed experimental data.

2009–2011 **Research Assistant**, Speech Perception Laboratory, Department of Psychology, Northeastern University, Boston, MA.
Maintained laboratory equipment and records.

Peer-reviewed Publications

K. Levin, A. Lodhia, and E. Levina. Recovering shared structure from multiple networks with unknown edge distributions. *Journal of Machine Learning*, 23(3):1–48, 2022.

K. Levin, F. Roosta, M. Tang, M. W. Mahoney, and C. E. Priebe. Limit theorems for out-of-sample extensions of the adjacency and Laplacian spectral embeddings. *Journal of Machine Learning Research*, 22(194):1–59, 2021.

A. Lodhia, K. Levin, and E. Levina. Matrix means and a novel high-dimensional shrinkage phenomenon. *Bernoulli*, 28(4):2578–2605, 2022.

V. Braverman, H. Lang, K. Levin, and Y. Rudoy. Metric k -median clustering in insertion-only streams. *Discrete Applied Mathematics*, 304:164–180, 2021.

C. M. Le, K. Levin, P. J. Bickel, and E. Levina. Comment: Ridge regression and regularization of large matrices. *Technometrics*, 62(4):443–446, 2020.

J. Yoder, L. Chen, H. Pao, E. Bridgeford, K. Levin, D. E. Fishkind, C. E. Priebe, and V. Lyzinski. Vertex nomination: The canonical sampling and the extended spectral nomination schemes. *Computational Statistics and Data Analysis*, 145:106916, 2020.

J. T. Vogelstein, E. W. Bridgeford, B. D. Pedigo, J. Chung, K. Levin, B. Mensh, and C. E. Priebe. Connectal coding: discovering the structures linking cognitive phenotypes to individual histories. *Current Opinion in Neurobiology*, 55:199–212, 2019.

V. Lyzinski, K. Levin, and C. E. Priebe. On consistent vertex nomination schemes. *Journal of Machine Learning Research*, 20(69):1–39, 2019.

C. M. Le, K. Levin, and E. Levina. Estimating a network from multiple noisy realizations. *Electronic Journal of Statistics*, 12(2):4697–4740, 2018.

K. Levin, F. Roosta, M. W. Mahoney, and C. E. Priebe. Out-of-sample extension of graph adjacency spectral embedding. In *Proceedings of the International Conference on Machine Learning*, Stockholm, Sweden, 2018.

A. Athreya, D. E. Fishkind, K. Levin, V. Lyzinski, Y. Park, Y. Qin, D. L. Sussman, M. Tang, J. T. Vogelstein, and C. E. Priebe. Statistical inference on random dot product graphs: a survey. *Journal of Machine Learning Research*, 18(226):1–92, 2018.

K. Levin and V. Lyzinski. Laplacian eigenmaps from sparse, noisy similarity measurements. *IEEE Transactions on Signal Processing*, 65(8):1988–2003, 2017.

S. Settle, K. Levin, H. Kamper, and K. Livescu. Query-by-example search with discriminative neural acoustic word embeddings. In *Proceedings of INTERSPEECH*, Stockholm, Sweden, 2017.

V. Lyzinski, K. Levin, D. E. Fishkind, and C. E. Priebe. On the consistency of the likelihood maximization vertex nomination scheme: Bridging the gap between maximum likelihood estimation and graph matching. *Journal of Machine Learning Research*, 17(179):1–34, 2016.

V. Braverman, H. Lang, K. Levin, and M. Monemizadeh. Clustering problems on sliding windows. In *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms*, Arlington, VA, USA, 2016.

V. Braverman, H. Lang, K. Levin, and M. Monemizadeh. Clustering on sliding windows in polylogarithmic space. In *Proceedings of the 35th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science*, Bangalore, India, 2015.

K. Levin, A. Jansen, and B. van Durme. Segmental acoustic indexing for zero resource keyword search. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Brisbane, Australia, 2015.

K. Levin, K. Henry, A. Jansen, and K. Livescu. Fixed-dimensional acoustic embeddings of variable-length segments in low-resource settings. In *Proceedings of the IEEE Automatic Speech Recognition and Understanding Workshop*, Olomouc, Czech Republic, 2013.

A. Jansen, E. Dupoux, S. Goldwater, M. Johnson, S. Khudanpur, K. Church, N. Feldman, H. Hermansky, F. Metze, R. Rose, M. Seltzer, P. Clark, I. McGraw, B. Varadarajan, E. Bennett, B. Borschinger, J. Chiu, E. Dunbar, A. Fourtassi, D. Harwath, C. Lee, K. Levin, A. Norouzi, V. Peddinti, R. Richardson, T. Schatz, and S. Thomas. A summary of the 2012 CLSP workshop on zero resource speech technologies and models of early language acquisition. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Vancouver, Canada, 2013.

Preprints and Manuscripts

K. Levin and B. Betancourt. Fast sampling of exchangeable partitions. *In preparation*, 2022+.

K. Levin and M. Fredrickson. Testing node-level treatment effects in networks. *In preparation*, 2022+.

A. Hayes and K. Levin. Mediation in low-dimensional network models. *In preparation*, 2022+.

K. Levin and Y. Zhang. Asymptotic behavior of average shortest path length under the graphon. *In preparation*, 2022+.

C. E. Priebe, Y. Park, K. Levin, E. Levina, and Y. Qin. Bootstrapping network data: Conditional versus marginal approaches. *In preparation*, 2022+.

C. E. Priebe K. Levin and V. Lyzinski. On the role of features in vertex nomination: Content and context together are better (sometimes). *Submitted, arXiv:2005.02151*, 2020.

K. Levin and E. Levina. Bootstrapping networks with latent space structure. *Submitted, arXiv:1907.10821*, 2019.

K. Levin, A. Athreya, M. Tang, C. E. Priebe, and V. Lyzinski. A central limit theorem for the omnibus embedding. *Submitted, arXiv:1705.09355*, 2017.

Teaching Experience

- Fall 2022 **Instructor**, STAT340: Introduction to Data Modeling II, University of Wisconsin, Department of Statistics.
Undergraduate introduction to statistical modeling and analysis.
- Spring 2022 **Instructor**, STAT679: Computing for Data Science and Statistics, University of Wisconsin, Department of Statistics.
Graduate-level survey of Python for data science.
- Fall 2021 **Instructor**, STAT340: Introduction to Data Modeling II, University of Wisconsin, Department of Statistics.
Undergraduate introduction to statistical modeling and analysis.
- Spring 2021 **Instructor**, STAT679: Computing for Data Science and Statistics, University of Wisconsin, Department of Statistics.
Graduate-level survey of Python for data science.
- Fall 2020 **Instructor**, STAT605: Data Science Computing Project, University of Wisconsin, Department of Statistics.
Graduate-level introduction to computing for data analysis.
- Fall 2019 **Instructor**, STATS426: Introduction to Theoretical Statistics, University of Michigan, Department of Statistics.
Upper-level undergraduate course on mathematical statistics.
- 2017-2019 (4 times) **Instructor**, STATS507: Data Analysis using Python, University of Michigan, Department of Statistics.
Masters/Ph.D.-level survey of tools for large-scale data analysis in the Python ecosystem.
- 2016 (twice) **Teaching Assistant**, EN 550.430: Introduction to Statistics, Johns Hopkins University, Department of Applied Mathematics and Statistics.
Upper-level undergraduate/first-year graduate course on mathematical statistics.
- Spring 2014 **Teaching Assistant**, EN 600.363/463: Introduction to Algorithms, Johns Hopkins University, Department of Computer Science.
Upper-level undergraduate/first-year graduate course on analysis of algorithms.

Presentations

- August, 2022 **Joint Statistical Meetings, Contributed Talk**, Bootstrapping Network Data: Conditional versus Marginal Distributions.

- July, 2022 **Levina/Zhu Research Group, University of Michigan, Invited Talk**, Bootstrapping Network Data: Conditional versus Marginal Distributions.
- July, 2022 **International Chinese Statistical Association China Conference, Invited Talk**, Bootstrapping Network Data: Conditional versus Marginal Distributions.
- June, 2022 **Annual Meeting of the Western North American Region of the International Biometric Society, Invited Talk**, Bootstrapping Network Data: Conditional versus Marginal Distributions.
- June, 2022 **International Conference on Econometrics and Statistics (EcoSta), Invited Talk**, Limit theorems for out-of-sample extensions of spectral graph embeddings.
- May, 2022 **New England Statistics Symposium, Invited Talk**, Limit theorems for out-of-sample extensions of spectral graph embeddings.
- April, 2022 **University of Delaware Department of Applied Economics and Statistics, Statistics Seminar**, Averaging Connectomes: Beyond the Arithmetic Mean.
- March, 2022 **University of Florida Department of Statistics, Department Seminar**, Averaging Connectomes: Beyond the Arithmetic Mean.
- Dec, 2021 **University of Pittsburgh Department of Statistics, Department Seminar**, Averaging Connectomes: Beyond the Arithmetic Mean.
- Nov, 2021 **University of Wisconsin-Madison SILO Seminar**, Averaging Connectomes: Beyond the Arithmetic Mean.
- April, 2021 **University of Wisconsin-Madison Probability Seminar**, Network Embeddings and Latent Space Models.
- April, 2021 **Ohio State University Department of Statistics Seminar Series**, Averaging Connectomes: Beyond the Arithmetic Mean.
- March, 2021 **International Biometric Society Eastern North American Region Spring Meeting, Invited Talk**, Averaging Connectomes: Beyond the Arithmetic Mean.
- Dec, 2020 **International Conference of the Computational and Financial Economics Working Group on Computational and Mathematical Statistics (CMStatistics), Invited Talk**, Estimation and bootstrapping for collections of low-rank networks.
- July, 2020 **University of Maryland College Park Department of Mathematics, Networks Seminar**, Bootstrapping networks with latent space structure.
- May, 2020 **SIAM Conference on Mathematics of Data Science Minisymposium on Machine Learning on Data with Low Dimensional Structures, Invited Talk (canceled due to COVID-19)**, Bootstrapping networks with latent space structure.
- Jan, 2020 **University of Wisconsin-Madison Department of Statistics, Department Seminar**, Bootstrapping networks with latent space structure.
- Jan, 2020 **Boston University Department of Mathematics & Statistics, Department Seminar**, Bootstrapping networks with latent space structure.
- Jan, 2020 **McGill University Department of Mathematics and Statistics, Department Seminar**, Bootstrapping networks with latent space structure.
- Jan, 2020 **Cornell University Department of Statistics and Data Science, Department Seminar**, Bootstrapping networks with latent space structure.

- Jan, 2020 **McMaster University Department of Mathematics and Statistics, Department Seminar**, Bootstrapping networks with latent space structure.
- Jan, 2020 **University of Notre Dame Department of Applied and Computational Mathematics and Statistics, Department Seminar**, Bootstrapping networks with latent space structure.
- Dec, 2019 **University of New Hampshire Department of Mathematics Department Seminar**, Bootstrapping networks with latent space structure.
- Dec, 2019 **University of Massachusetts Amherst, Department of Mathematics and Statistics, Department Seminar**, Bootstrapping networks with latent space structure.
- Sep, 2019 **George Mason University Department of Statistics, Department Seminar**, Limit theorems for out-of-sample extensions of spectral graph embeddings.
- Sep, 2019 **University of Maryland College Park Department of Mathematics, Statistics Seminar**, Limit theorems for out-of-sample extensions of spectral graph embeddings.
- July, 2019 **Joint Statistical Meetings, Invited Talk**, Matrix means for network estimation with applications to fMRI data.
- July, 2019 **New Researchers Conference**, Bootstrapping networks with latent space structure.
- May, 2019 **NetSci Statistical Inference in Network Models**, Testing treatment effects on network topology.
- May, 2019 **New England Statistics Symposium, Invited Talk**, Bootstrapping networks with latent space structure.
- Mar, 2019 **University of Michigan Complex Systems Seminar**, Recovering low-rank structure from multiple networks with unknown edge distributions.
- Sep, 2018 **Mahoney Lab, University of California, Berkeley, Invited Talk**, Out-of-sample extensions for spectral embeddings of graphs.
- July, 2018 **Joint Statistical Meetings, Contributed Talk**, Inferring low-rank population structure from multiple network samples.
- Feb, 2018 **Boston University Department of Mathematics and Statistics, Statistics and Probability Seminar**, Joint embedding of multiple random dot product graphs.
- Nov, 2017 **IEEE International Conference on Data Mining, Data-Driven Discovery of Models Workshop**, A central limit theorem for an omnibus embedding of random dot product graphs.
- June, 2017 **MIT Lincoln Laboratory, Artificial Intelligence Technology and Systems Group, Invited Talk**, Recent advances in vertex nomination.
- May, 2017 **Mahoney Lab, University of California, Berkeley, Invited Talk**, Recent advances in vertex nomination.
- May, 2017 **Levina/Zhu Research Group, University of Michigan, Invited Talk**, Laplacian eigenmaps from sparse, noisy similarity measurements.
- Nov, 2015 **Johns Hopkins University Department of Applied Mathematics and Statistics, Student Seminar**, Laplacian eigenmaps with noisy kernels.

- Oct, 2014 **Johns Hopkins University Center for Language and Speech Processing, Student Seminar**, Low-resource fixed-dimensional embeddings of variable-length speech audio segments.
- Oct, 2013 **Mid-Atlantic Student Colloquium on Speech, Language and Learning**, Toward faster audio search using context-dependent hashing.

██████████ **Awards and Memberships**

- 2021-2024 **NSF Grant**, FRG: Collaborative Research: Flexible Network Inference, NSF DMS 2052632, \$200,000.
- 2021-2022 **MTLE Fellow**, Madison Teaching and Learning Excellence.
- 2022 **Domestic Travel Award**, UW-Madison Office of the Vice Chancellor for Research and Graduate Education.
- 2013 **Best Student Paper**, IEEE Automatic Speech Recognition and Understanding Workshop.
- Member**, American Statistical Association.
- Member**, Institute of Mathematical Statistics.
- Member**, Bernoulli Society.
- Member**, Association for Computing Machinery.
- Member**, Institute of Electrical and Electronics Engineers.

██████████ **Conferences and Workshops Organized**

- 2021 **Organizer**, Low-rank models in multiple-network analysis, Satellite Workshop of Networks 2021.

██████████ **Professional Service**

- 2016-present Reviewer: the Annals of Statistics (2018,2021), the Annals of Applied Statistics (2020), Biometrika (2020,2021), NeurIPS (2020), NeurIPS Workshop: Your Model is Wrong (2021), International Conference on Machine Learning (2019), Journal of Machine Learning Research (2020,2022), Journal of the Royal Statistical Society Series B (2020,2020,2021), IEEE Transactions on Pattern Analysis and Machine Intelligence (2016,2021,2021,2022), Electronic Journal of Statistics (2020,2021,2022), Journal of Computational and Graphical Statistics (2018,2021), ACM-SIAM Symposium on Discrete Algorithms (2018), Machine Learning (2017), Pattern Recognition Letters (2017), Complex Networks (2018), the R Journal (2019), Multivariate Behavioral Research (2016), Royal Society Proceedings A (2020)
- 2021-present University of Wisconsin Department of Statistics Computing Committee
- 2020-present University of Wisconsin Department of Statistics MS Exam Committee
- 2020-present University of Wisconsin Department of Statistics Seminar Committee
- 2019 Organizer: University of Michigan Department of Statistics Fall Preview Day
- 2015–2017 Johns Hopkins University Center for Language and Speech Processing Faculty Liason
- 2014–2016 Johns Hopkins University Computer Science Ph.D. Admissions Committee

■ Mentoring

- 2021-present Mentoring PhD student project on network mediation in causal inference.
- 2020-present Mentoring early-career graduate student in project on hierarchical stochastic block-models.
- 2022-present Mentoring masters-level student project on degree distribution of random spanning trees.
 - 2021-2022 Mentored masters-level student project on extreme values of degree distributions in network data.
 - 2021-2022 Mentor for WISCERS, mentoring undergraduate in project on networks and basic reproductive number.
 - 2019-2020 Oversaw two undergraduate statistics students in analyzing neuroscientific data from a study on addiction in a rat model.
 - 2019-2020 Oversaw undergraduate data science project collecting and analyzing spatio-temporal data on locations and ownership structure of payday loan companies.
 - 2018-2019 Oversaw two undergraduate statistics students in analyzing neuroscientific data from a study on addiction in a rat model.
 - 2017-2019 Mentored early-career domestic graduate student in project on network regression with application to schizophrenia fMRI data.
 - 2018-2019 Oversaw undergraduate statistics student class project to predict schizophrenia diagnoses from fMRI data.

■ Computing Skills

Programming languages: C/C++, Python, R, MATLAB, LISP, Julia, Java.

High-performance computing: HTCondor, Slurm, Hadoop MapReduce, Google TensorFlow, Keras, Apache Spark, OpenMP, Message Passing Interface.

■ Languages

Native fluency in English; Limited working proficiency in Spanish and French; Elementary proficiency in Hebrew, Russian, Yiddish and German

■ Personal

United States citizen.