

A. Academic History:

- University of California, Berkeley: PhD in Computer Science with a minor in Statistics, 2012.
- Harvard University: Bachelor of Arts in Mathematics, Magna Cum Laude with Highest Honors in Mathematics, 2006.
- Harvard University: Master of Science in Computer Science, 2006.

B. Employment Record:

- University of California, Berkeley, 2007-2012: Graduate student researcher and teaching assistant.
- Microsoft Research, New England, Summer 2009: Internship mentored by Adam Kalai.
- IBM Research, Almaden, Summer 2010: Internship under the supervision of Ken Clarkson and the mentorship of Vitaly Feldman.
- Microsoft Research, New England, 2012-2013: Postdoctoral researcher.
- Stanford University, 2013-present: Assistant Professor, Department of Computer Science.

C. Professional Activities

- Co-organized (with Dan Boneh and Percy Liang) 3-week Summer School for high-school students on *Computer Security and Machine Learning*, with an emphasis on current research trends in robust machine learning.
- Co-organized (with Ankur Moitra and David Donoho) a summer school at UCSD on the research frontiers of Robust Learning and Estimation (>100 registered attendees), coordinated by Shachar Lovett.
- ACM Doctoral Dissertation Award committee, 2019, 2020.
- Program committee member for the following conferences: Highlights of Algorithms (HALG 2019), Innovations in Theoretical Computer Science (ITCS 2016, 2018, 2020), Conference on Learning Theory (COLT 2016, 2019), Foundations of Computer Science (FOCS 2018), Symposium on Theory of Computing (STOC, 2014).

D. University and Departmental Service

- Departmental committees: Curriculum Committee (2014-2015), PhD admissions (2013-2014, 2014-2015, 2015-2016, 2017-2018), Masters Admissions (2016-2017), Broad Area Faculty Search (2018-2019, 2019-2020), CS “Vision” committee (2017-2018).
- Co-organize weekly Statistical Machine Learning lunches (attendance ~60 PhD students and faculty), 2016-present.
- Organized weekly Theory Seminar talk series, and weekly Theory Lunches (attendance ~50 undergraduates, PhD students, and faculty), 2013-2016.
- Faculty Fellow for Men’s Varsity Waterpolo Team (2017-present), and Women’s Varsity Waterpolo Team, (2019-present).
- Co-organized periodic university-wide happy hours for junior faculty, 2013-2016.

E. Awards and Honors

- Young Investigator Award, Office of Naval Research, 2018.
- Hellman Faculty Scholar, 2017.
- Sloan Foundation Research Fellowship, 2016.
- Okawa Foundation Research Grant, 2015.
- Inclusion in the 2014 ACM Computing Review;s “Best of Computing” list for the paper *An automatic inequality prover and instance optimal identity testing*.
- NSF CAREER Award, 2014.
- ACM Doctoral Dissertation Award, Honorable Mention, 2012.
- ACM Symposium on Principles of Database Systems (PODS) Best Paper Award, 2009.

## F. BIBLIOGRAPHICAL INFORMATION:

### PUBLICATIONS

The publications in theoretical computer science venues list author names in alphabetical order. Several of the publications in machine learning venues list students first, in order of contribution.

Legend: My PhD students, and supervised undergraduate and M.S. students are indicated in **boldface**. Collaborators who were Stanford PhD students *not* advised by me are underlined.

#### Refereed Conference/Symposia Proceedings

1. **V. Sharan**, A. Sidford and G. Valiant, *Memory-Sample Tradeoffs for Linear Regression with Small Error*, Symposium on Theory of Computing (STOC), 2019.
2. **M. Guan** and G. Valiant, *A Surprising Density of Illusionable Natural Speech*, Annual Meeting of the Cognitive Science Society (CogSci), 2019.
3. **K.S. Tai**, P. Bailis and G. Valiant, *Equivariant Transformer Networks*, International Conference on Machine Learning (ICML), 2019.
4. **V. Sharan**, **K.S. Tai**, P. Bailis and G. Valiant, *Compressed Factorization: Fast and Accurate Low-Rank Factorization of Compressively-Sensed Data*, International Conference on Machine Learning (ICML), 2019.
5. R. Vinayak, **W. Kong**, G. Valiant, and S. Kakade, *Maximum Likelihood Estimation for Learning Populations of Parameters*, International Conference on Machine Learning (ICML), 2019.
6. **M. Qiao** and G. Valiant, *A Theory of Selective Prediction*, Conference on Learning Theory (COLT), 2019.
7. **W. Kong** and G. Valiant, *Estimating Learnability in the Sublinear Data Regime*, Neural Information Processing Systems (NeurIPS), 2018.
8. **S. Garg**, **V. Sharan**, **B.H. Zhang** and G. Valiant, *A Spectral View of Adversarially Robust Features*, Neural Information Processing Systems (NeurIPS), 2018.
9. **V. Sharan**, S. Kakade, P. Liang and G. Valiant, *Prediction with a Short Memory*, Symposium on Theory of Computing (STOC), 2018.
10. **M. Meister** and G. Valiant, *A Data Prism: Semi-Verified Learning in the Small-alpha Regime*, Conference on Learning Theory (COLT), 2018.

11. **M. Qiao** and G. Valiant, *Learning Discrete Distributions from Untrusted Batches*, Innovations in Theoretical Computer Science (ITCS), 2018.
12. Q. Huang, S. Kakade, **W. Kong** and G. Valiant, *Recovering Structured Probability Matrices*, Innovations in Theoretical Computer Science (ITCS), 2018.
13. **K.S. Tai**, **V. Sharan**, P. Bailis and G. Valiant, *Finding Heavily-Weighted Features in Data Streams*, SIGMOD, 2018.
14. D. Cohen-Steiner, **W. Kong**, C. Sohler and G. Valiant, *Approximating the Spectrum of a Graph*, SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2018.
15. K. Tian, **W. Kong** and G. Valiant, *Learning Populations of Parameters*, Neural Information Processing Systems (NeurIPS), 2017.
16. **V. Sharan**, S. Kakade, P. Liang and G. Valiant, *Learning Overcomplete HMMs*, Neural Information Processing Systems (NeurIPS), 2017.
17. **V. Sharan** and G. Valiant, *Orthogonalized ALS: A Theoretically Principled Tensor Decomposition Algorithm for Practical Use*, International Conference on Machine Learning (ICML), 2017.
18. A. Raghunathan, G. Valiant and J. Zou, *Estimating the Unseen from Multiple Populations*, International Conference on Machine Learning (ICML), 2017.
19. M. Charikar, J. Steinhardt and G. Valiant, *Learning from Untrusted Data*, Symposium on Theory of Computing (STOC), 2017.
20. J. Steinhardt, G. Valiant and M. Charikar, *Avoiding Imposters and Delinquents: Adversarial Crowdsourcing and Peer Prediction*, Neural Information Processing Systems (NeurIPS), 2016.
21. J. Steinhardt, G. Valiant and S. Wager, *Memory, Communication, and Statistical Queries*, Conference on Learning Theory (COLT), 2016.
22. G. Valiant and P. Valiant, *Instance Optimal Learning of Discrete Distributions*, Symposium on Theory of Computing (STOC), 2016.
23. B. Bhattacharya and G. Valiant, *Testing closeness with unequal sized samples*, Neural Information Processing Systems (NeurIPS), 2015.
24. G. Valiant and P. Valiant, *An automatic Inequality Prover and Instance Optimal Identity Testing*, Foundations of Computer Science (FOCS), 2014.
25. A. Livnat, C. Papadimitriou, A. Rubinfeld, A. Wan and G. Valiant, *Satisfiability and Evolution*, Foundations of Computer Science (FOCS), 2014.
26. A. Andoni, R. Panigrahy, G. Valiant and L. Zhang, *Learning Polynomials with Neural Networks*, International Conference on Machine Learning (ICML), 2014.
27. A. Agarwal, S. Kakade, N. Karampatziakis, L. Song and G. Valiant, *Learning Least Squares Revisited: Scalable Approaches for Multi-class Prediction*, International Conference on Machine Learning (ICML), 2014.
28. S. Chan, I. Diakonikolas, G. Valiant and P. Valiant, *Optimal Algorithms for Testing Closeness of Discrete Distributions*, Symposium on Discrete Algorithms (SODA), 2014.
29. A. Andoni, R. Panigrahy, G. Valiant and L. Zhang, *Learning Sparse Polynomial Functions*, Symposium on Discrete Algorithms (SODA), 2014.
30. G. Valiant and P. Valiant, *Estimating the Unseen: Improved Estimators for Entropy and Other Properties*, Neural Information Processing Systems (NeurIPS), 2013.
31. G. Valiant, *Finding Correlations in Subquadratic Time, with Applications to Learning Parities and Juntas*, Foundations of Computer Science (FOCS), 2012.

32. C. Daskalakis, I. Diakonikolas, R. Servedio, G. Valiant and P. Valiant, *Testing  $k$ -modal Distributions: Optimal Algorithms via Reductions*, Symposium on Discrete Algorithms (SODA), 2013.
33. G. Valiant and P. Valiant, *The Power of Linear Estimators*, Foundations of Computer Science (FOCS), 2011.
34. G. Valiant and P. Valiant, *Estimating the Unseen: An  $n/\log(n)$  sample Estimator for Entropy and Support Size, Shown Optimal via New CLTs*, Symposium on Theory of Computing (STOC), 2011.
35. N. Nisan, M. Schapira, G. Valiant and A. Zohar, *Best-Response Auctions*, Conference on Electronic Commerce (EC), 2011.
36. C. Papadimitriou and G. Valiant, *A New Look at Selfish Routing*, Innovations in Theoretical Computer Science (ITCS), 2010.
37. A. Moitra and G. Valiant, *Settling the Polynomial Learnability of Mixtures of Gaussians*, Foundations of Computer Science (FOCS), 2010.
38. A. Kalai, A. Moitra and G. Valiant, *Efficiently Learning Mixtures of Two Gaussians*, Symposium on Theory of Computing (STOC), 2010.
39. G. Gottlob, S.T. Lee and G. Valiant, *Size and Treewidth Bounds for Conjunctive Queries*, Symposium on Principles of Database Systems (PODS), 2009.
40. C. Daskalakis, G. Shoenebeck, G. Valiant and P. Valiant, *On the Complexity of Nash Equilibria of Action-Graph Games*, Symposium on Discrete Algorithms (SODA), 2009.
41. H.L. Chen, T. Roughgarden and G. Valiant, *Designing Network Protocols for Good Equilibria*, Symposium on Discrete Algorithms (SODA), 2008.
42. G. Valiant and T. Roughgarden, *Braess's Paradox in Large Random Graphs*, Conference on Electronic Commerce (EC), 2006.

#### Refereed Conference/Symposia Proceedings in Press/Accepted

43. **T. Ginart**, **M. Guan**, G. Valiant and J. Zou, *Making AI Forget You: Data Deletion in Machine Learning*, Neural Information Processing Systems (NeurIPS), 2019 (to appear).
44. **B. Axelrod**, I. Diakonikolas, A. Stewart, A. Sidiropoulos and G. Valiant, *A Polynomial Time Algorithm for Log-Concave Maximum Likelihood via Locally Exponential Families*, Neural Information Processing Systems (NeurIPS), 2019 (to appear).

#### Refereed Journal Publications

1. **W. Kong** and G. Valiant, *Spectrum Estimation from Samples*. Annals of Statistics, 45(5), 2218-2247, 2017.
2. G. Valiant and P. Valiant, *An automatic inequality prover and instance optimal identity testing*. SIAM Journal on Computing, 46(1): 429-455, 2017.
3. G. Valiant and P. Valiant, *Estimating the unseen: improved estimators for entropy and other properties*. Journal of the ACM, 64(6), 1-41, 2017.
4. J. Zou, G. Valiant, P. Valiant, K. Karczewski, S.O. Chan, K. Samocha, M. Lek, S. Sunyaev, M. Daly and D.G. MacArthur, *Quantifying the unobserved protein-coding variants in*

- human populations provides a roadmap for large-scale sequencing projects*, Nature Communications (7), 2016.
5. G. Valiant, *Finding correlations in subquadratic time, with applications to learning parities and the closest pair problem*, Journal of the ACM, 62(2), 1:45, 2015.
  6. A. Kalai, A. Moitra and G. Valiant, *Disentangling Gaussians*. Communications of the ACM, 55(2), 113-120, 2012.
  7. G. Gottlob, S.T. Lee, G. Valiant and P. Valiant, *Size and Treewidth Bounds for Conjunctive Queries*, Journal of the ACM, 59(3), 1-35, 2012.
  8. H. Chen, T. Roughgarden, and G. Valiant, *Designing Network Protocols for Good Equilibria*, SIAM Journal on Computing, 39(5), 1799-1832, 2010.
  9. T. Roughgarden and G. Valiant, *Braess's Paradox in Large Random Graphs*. Random Structures and Algorithms, 37(4), 495-515, 2010.
  10. S. Stewart and G. Valiant, *Martian Subsurface Properties and Crater Formation Processes Inferred from Fresh Impact Crater Geometries*, Meteoritics & Planetary Sciences, 41(10), 2006.
- Refereed Journal Publications Submitted:
    1. **W. Kong** and G. Valiant, *Estimating Learnability in the Sublinear Data Regime*. Submitted to the Journal of the ACM, March, 2019, under review as of September 20, 2019.

#### Non-refereed Publications

- Non-refereed pre-print servers
  1. **B. Axelrod, S. Garg, V. Sharan** and G. Valiant, *Sample Amplification: Increasing Dataset Size Even When Learning is Impossible*. Arxiv.org/abs/1904.12053, April, 2019.
  2. **G. Blanc, N. Gupta**, G. Valiant and P. Valiant, *Implicit Regularization for Deep Neural Networks Driven by an Ornstein-Uhlenbeck like Process*. Arxiv.org/abs/1904.09080, April, 2019.
- Book Chapters Submitted
  1. G. Valiant and P. Valiant, *Instance Optimal Learning and Estimation*, in Beyond Worst-Case Analysis, edited by Tim Roughgarden. Draft submitted August, 2019.

#### PRESENTATIONS (January 2016 - October, 2019)

- Invited Presentations
  1. University of California, San Diego, Robust Learning and Statistics Summer School, 8/12/19-8/15/19.
  2. Google TechTalk, 12/17/18.
  3. University of California, Berkeley, Simons Center Lecture, 11/4/18.
  4. IEEE Silicon Valley Artificial Intelligence Chapter, Invited Lecture, 8/23/18.
  5. University of California, Berkeley, Simons Center Workshop, 6/6/18.
  6. University of Warwick, Data Summarization Workshop, 3/13/18.
  7. Information Theory and Applications (ITA), 2/15/18.
  8. Los Alamos, Physics Informed Machine Learning Workshop, 1/24/18.
  9. Allerton, 10/6/17.
  10. Dagstuhl Workshop on Learning and Games, 6/20/17.

11. Highlights of Algorithms (HALG), 6/10/17.
  12. University of California, Berkeley, Simons Center Workshop on Approximation Algorithms, 4/27/17.
  13. Google Research Talk, 4/14/17.
  14. Stanford Center for Professional Development, Webinar, 3/17.
  15. Information Theory and Applications (ITA), 2/16/17.
  16. Neural Information Processing Systems (NeurIPS) Workshop on Tensor Methods, 12/10/16.
  17. Workshop on Local Algorithms (MSR/MIT), 10/15/16.
  18. DIMACS Cryptography and Security Workshop, 7/11/16.
  19. The Mathematics Behind Big Data Analysis, invited talk at SIAM conference on Discrete Math, 6/7/16.
  20. Information Theory Society, Public Lecture, 3/23/16.
  21. Henri Poincare Institute, as part of the Nexus of Information and Computation Theories Program, 3/11/16
  22. Joint UCSF/UC Berkeley Simons Center Seminar, 2/11/16
  23. Symposium on Discrete Algorithms (SODA), Sublinear Algorithms Workshop, 1/7/2016.
- Contributed Conference Presentations
    1. Innovations in Theoretical Computer Science (ITCS), 1/14/18.
    2. Symposium on Theory of Computing (STOC), 6/20/16.
  - Department Seminars
    1. UCLA Electrical and Computer Engineering Seminar, 10/28/19.
    2. Princeton Theory Lunch, 10/18/19.
    3. MIT Theoretical Computer Science Seminar, 5/16/19.
    4. Columbia Theory Seminar, 5/8/19.
    5. University of Chicago, Statistics Departmental Seminar, 3/11/19.
    6. University of Washington Statistics Colloquium, 11/30/18.
    7. Vienna IST Colloquium, 11/18/18.
    8. University of California, Berkeley, Theory Seminar, 4/10/18.
    9. Carnegie Mellon University, Dept. Statistics and Data Science Seminar, 9/18/17.
    10. Carnegie Mellon University, Joint Machine Learning and Computer Science Dept. Seminar, 9/19/17.
    11. University of Washington, Theory Seminar, 5/30/17.
    12. Caltech Computer Science Seminar, 11/17/16.
    13. ETH Zurich, Departmental Colloquium, 11/9/16.
    14. Ecole Polytechnique Federale de Lausanne, Departmental Colloquium, 11/8/16
    15. New York University, Theory Seminar, 4/28/16.
    16. Columbia University, Departmental Seminar, 4/29/16.
    17. Stanford Statistics Seminar, 4/26/16.
    18. USC Electrical Engineering Seminar, 3/2/16
    19. MIT LIDS/TOC Colloquium, 2/23/16.
    20. Stanford ISL Colloquium, 2/23/16.

## STUDENTS

- Former PhD Students
  1. Weihao Kong, Thesis: *The Surprising Power of Little Data*. Graduation Date: September, 2019. (Currently post-doc at University of Washington.)
  2. Hongyang Zhang (co-advised with Ashish Goel), Thesis: *Algorithms and Generalization for Large-Scale Matrices and Tensors*. Graduation Date: September, 2019. (Will be joining Northeastern as an assistant professor in Fall, 2020. Currently post-doc at UPenn.)
- Current PhD Students
  1. Vatsal Sharan, Thesis: *Perspectives on Memory in Learning*. Anticipated Graduation Date: June, 2020.
  2. Kai Sheng Tai (co-advised with Peter Bailis). Anticipated Graduation Date: June, 2021.
  3. Brian Axelrod (co-advised with Omer Reingold). Thesis: *Understanding and Harnessing Randomness*. Anticipated Graduation Date: June, 2022.
  4. Shivam Garg. Anticipated Graduation Date: June, 2022.
  5. Neha (co-advised with Moses Charikar). Anticipated Graduation Date: June, 2022.
  6. Jay Mardia (co-advised with Tsachy Weissman). Anticipated Graduation Date: June, 2022.
  7. Annie Marsden (co-advised with John Duchi). Anticipated Graduation Date: June, 2022.
  8. Melody Guan. Anticipated Graduation Date: June, 2022.
  9. Mingda Qiao. Anticipated Graduation Date: June, 2023.
- Supervised or Supported Post-Docs
  1. Clement Canonne, supported by the Motwani Fellowship, and mentored by the Theory group. Appointed 2018.
  2. Yuanzhi Li, jointly supported by me, the Motwani Fellowship and Tengyu Ma. Appointed 2018.
- Supervised Undergraduate Researchers
  1. Collin Schlager. Anticipated Date of Graduation: June, 2021.
  2. Justin Chen. Anticipated Date of Graduation: June, 2020.
  3. Brian Hu Zhang. Date of Graduation: June, 2019.
  4. Guy Blanc. Date of Graduation: June, 2019.
  5. Michela Meister. Date of Graduation: December, 2018.
  6. Yongxing Deng. Date of Graduation: June, 2015.
- Supervised M.S. Researchers
  1. Michela Meister. Date of Graduation: June, 2017.