Weihao Kong

✉ kweihao@gmail.com
🏠 homes.cs.washington.edu/~whkong

Research Interests

○ Learning from a large number of data sources: A common modern machine learning scenario involves a large amount of data contributed by a large number of heterogeneous data sources, with each data source providing a modest amount of data. How well can we learn in this setting? To what extent can a large number of sources compensate for the lack of data from each source? What is the fundamental limit of learning? This problem has been studied under multi-task learning, meta-learning, federated learning, few-shot learning, empirical bayesian by different communities.
Meta-learning for linear models: [1, 2]; Learning populations of binomial parameters: [3, 5, 10]

○ Estimating learnability: Without enough data to learn a good model for prediction, is it possible to tell whether a good model exists? This is surprisingly possible under linear model assumptions [7]. An analogs question can be asked in the contextual bandits setting: Without enough rounds to learn a good policy, is it possible to estimate the value of the optimal policy [4]?

○ Robust machine learning: How to design learning algorithms that are provably robust against a small fraction of malicious data/users? The problem becomes more and more pressing with the advancement of federated learning.
Robust meta-learning and robust PCA: [1]; Robust linear regression: [6].

Education

Sep 2013 - Sep 2019	**Ph.D.**, *Computer Science*, advised by Gregory Valiant. Stanford University
Sep 2009 - Jun 2013	**B.Eng.**, *ACM-Honored class of Computer Science*. Shanghai Jiao Tong University

Professional Experience

Sep 2019 - present	**Postdoc researcher**, working with Sham M. Kakade. University of Washington
July 2018 - Aug 2018	**Visiting researcher**, *research on learning populations of parameters*, hosted by Sham M. Kakade. University of Washington
Jun 2017 - Aug 2017	**Visiting researcher**, *research on robust statistics*, hosted by Ilias Diakonikolas. University of Southern California
Jun 2014 - Sep 2014	**Research intern**, *research on nonparametric Bayesian approaches to clustering*, supervised by Bo-June(Paul) Hsu. Microsoft Research, Redmond
Aug 2012 - April 2013	**Research intern**, *research on an allocation algorithm for display advertising*, supervised by by Tao Qin. Microsoft Research Asia, IECA group

Teaching Experience

Teaching Assistant

Spring 2017	**CS168 Modern Algorithmic Toolbox**, *Instructor: Tim Roughgarden and Gregory Valiant*.
Fall 2015	**CS265 Randomized Algorithms and Probabilistic Analysis**, *Instructor: Gregory Valiant*.
Spring 2015	**CS261 Optimization and Algorithmic Paradigms**, *Instructor: Tim Roughgarden*.
Fall 2014	**CS265 Randomized Algorithms and Probabilistic Analysis**, *Instructor: Gregory Valiant*.

Services

Conference reviewer, *COLT17, SODA18, ITCS18, AAAI 18, ICML 19, FOCS 19, NeurIPS 19, ICML 20, FOCS 20, RANDOM 20, NeurIPS 20*.

Journel reviewer, *Biometrika, Electronic Journal of Statistics, Journal of Machine Learning Research (JMLR), Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*.

Organizer, *Stanford Theory Seminar 16-17*.

Publications

[1] Weihao Kong, Raghav Somani, Sham Kakade, and Sewoong Oh. Robust meta-learning for mixed linear regression with small batches. *arXiv preprint arXiv:2006.09702*, 2020.

[2] Weihao Kong, Somani Raghav, Zhao Song, Sham M Kakade, and Sewoong Oh. Meta-learning for mixed linear

regression. In *International Conference on Machine Learning (**ICML**)*, 2020.

[3] Ramya Korlakai Vinayak, Weihao Kong, and Sham M Kakade. Optimal estimation of change in a population of parameters. *arXiv preprint arXiv:1911.12568*, 2019.

[4] Weihao Kong, Gregory Valiant, and Emma Brunskill. Sublinear optimal policy value estimation in contextual bandits. *The 23rd International Conference on Artificial Intelligence and Statistics (**AISTATS**)*, 2020.

[5] Ramya Korlakai Vinayak, Weihao Kong, Gregory Valiant, and Sham Kakade. Maximum likelihood estimation for learning populations of parameters. In *International Conference on Machine Learning (**ICML**)*, 2019.

[6] Ilias Diakonikolas, Weihao Kong, and Alistair Stewart. Efficient algorithms and lower bounds for robust linear regression. *In ACM-SIAM Symposium on Discrete Algorithms (**SODA**)*, 2019.

[7] Weihao Kong and Gregory Valiant. Estimating learnability in the sublinear data regime. *In Advances in Neural Information Processing Systems (**NeurIPS**)*, 2018.

[8] David Cohen-Steiner, Weihao Kong, Christian Sohler, and Gregory Valiant. Approximating the spectrum of a graph. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (**KDD**)*. ACM, 2018.

[9] Qingqing Huang, Sham M. Kakade, Weihao Kong, and Gregory Valiant. Recovering Structured Probability Matrices. In *9th Innovations in Theoretical Computer Science Conference (**ITCS**)*, 2018.

[10] Kevin Tian, Weihao Kong, and Gregory Valiant. Learning populations of parameters. In *Advances in Neural Information Processing Systems (**NIPS**)*, 2017.

[11] Weihao Kong, Gregory Valiant, et al. Spectrum estimation from samples. *The Annals of Statistics*, 45(5):2218–2247, 2017.

[12] Weihao Kong, Jian Li, Tie-Yan Liu, and Tao Qin. Optimal allocation for chunked-reward advertising. In *Web and Internet Economics (**WINE**)*. Springer, 2013.

[13] Weihao Kong and Wu-Jun Li. Isotropic hashing. In *Advances in Neural Information Processing Systems 25 (**NIPS**)*, 2012.

[14] Weihao Kong, Wu-Jun Li, and Minyi Guo. Manhattan hashing for large-scale image retrieval. In *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval (**SIGIR**)*, 2012.

[15] Weihao Kong and Wu-Jun Li. Double-bit quantization for hashing. In *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence (**AAAI**)*, 2012.