

# Yuchen Zhang

---

## CONTACT

<https://zhangyuc.github.io/>  
Email: zhangyuc@cs.stanford.edu  
Phone: (+1)-510-423-1353

## ADDRESS

Gates Computer Science 254,  
Stanford University,  
Stanford, CA 94305.

## Education

<b>University of California, Berkeley</b> Doctor of Philosophy in Computer Science Advised by Michael I. Jordan and Martin J. Wainwright	<b>2011 - 2016</b>
<b>University of California, Berkeley</b> Master of Arts in Statistics	<b>2011 - 2013</b>
<b>Tsinghua University</b> Bachelor of Engineering in Computer Science Supervised by Andrew C. Yao	<b>2007 - 2011</b>

## Employment

<b>Post-doc Researcher at Stanford University</b> Hosted by Percy Liang and Moses Charikar	<b>2016 - Now</b>
<b>Intern at Baidu</b> Project: burst detection in web search	<b>Winter, 2015</b>
<b>Intern at Microsoft Research Redmond</b> Project: convex optimization	<b>Summer, 2014</b>
<b>Intern at Google Mountain View</b> Project: personalized recommender systems	<b>Summer, 2013</b>
<b>Intern at Microsoft Research Asia</b> Project: click modeling for web search and online advertising	<b>2010 - 2011</b>

## Selected Awards & Honors

<b>2017</b>	Best Paper Award, Conference on Learning Theory (COLT).
<b>2016</b>	Outstanding Reviewer Award, International Conference on Machine Learning (ICML).
<b>2015</b>	Baidu Fellowship (awards 8 PhD students every year worldwide).
<b>2013</b>	Microsoft Research PhD Fellowship Finalist.
<b>2011</b>	Outstanding Undergraduate Dissertation Award, Tsinghua University.
<b>2006</b>	Silver Medal in Asian Physics Olympiad.
<b>2006</b>	Gold Medal in Chinese Physics Olympiad (5 <sup>th</sup> among 400,000 participants).

## Research Projects

My research interest lies in the field of *artificial intelligence*. My past projects spanned machine learning, natural language processing and statistical methods. They consist of new algorithms, fundamental theory and practical software. Below is a list of projects that I have been working on (with publication references):

### Natural language processing

- Semantic parsing for question answering systems [C1].

- Efficient algorithm for parsing natural language into executable logical forms.
- Impact: 11x-16x faster and 13% more accurate than the state-of-the-art parser.

## Machine learning and optimization

- Deep learning and non-convex optimization [C2,C3,C4,C5,C6].
  - Algorithms for non-convex optimization (with applications to deep learning).
  - Impact: stronger theoretical guarantee and improved empirical performance on DNN/CNN/LSTM.
- Convex optimization [J1,C9].
  - Efficient algorithm for convex optimization.
  - Impact: orders-of-magnitude faster convergence than SGD in theory/practice.
- Crowdsourcing [J4,C10].
  - Algorithm for estimating the true labels from noisy crowdsourced data.
  - Impact: first algorithm to guarantee the best possible statistical accuracy.
- Personalized recommender systems [C12].
  - Non-parametric model for modeling very sparse user-item relations.
  - Impact: improved state-of-the-art models by 9%-12% on the quality of recommending less frequent items for online shopping.
- Web search and online advertising [C16,C17,C18,C19,C20,C21]
  - Learning from massive search & ads click logs to improve ranking quality.
  - Impact: currently in production as a part of Microsoft Bing.

## Distributed computing

- Distributed algorithms for machine learning [J5,J6,C7,C8,C14,C15].
  - Communication-efficient algorithms for distributed machine learning on large clusters.
  - Impact: guarantee the best possible accuracy and the least possible computation/communication.
- Fundamental theory of distributed computing [C11,C13,M2].
  - Understanding the fundamental trade-offs between communication, computation and statistical accuracy on any distributed system.
  - Impact: seminal papers on the study of communication complexity of distributed statistical estimation.
- Programming interface for parallelizing stochastic algorithms [M1].
  - A user-friendly framework for parallelizing online algorithms on Spark.
  - Impact: component of Berkeley Data Analytics Stack (BDAS), 80+ Github stars.

## Other projects

- Theoretical statistics [J2,J3].
- Theoretical computer science [J7,C22].

## Manuscripts

- [M1] **Y. Zhang** and MI. Jordan. Splash: User-friendly Programming Interface for Parallelizing Stochastic Algorithms. *arXiv:1506.07552*, 2015.
- [M2] J. Duchi, MI. Jordan, M. Wainwright and **Y. Zhang** (alpha-beta order). Optimality Guarantees for Distributed Statistical Estimation. *arXiv:1405.0782*, 2014.

## Journal Publications

- [J1] **Y. Zhang** and L. Xiao. Stochastic Primal-Dual Coordinate Method for Regularized Empirical Risk Minimization. *Journal of Machine Learning Research*.

- [J2] **Y. Zhang**, M. Wainwright and M.I. Jordan. Optimal prediction for sparse linear models? Lower bounds for coordinate-separable M-estimators. *Electronic Journal of Statistics*.
- [J3] X. Chen, A. Guntuboyina and **Y. Zhang** (alpha-beta order). On Bayes Risk Lower Bounds. *Journal of Machine Learning Research*.
- [J4] **Y. Zhang**, X. Chen, D. Zhou and M.I. Jordan. Spectral Methods meet EM: A Provably Optimal Algorithm for Crowdsourcing. *Journal of Machine Learning Research*.
- [J5] **Y. Zhang**, J. Duchi and M. Wainwright. Divide and Conquer Kernel Ridge Regression: A Distributed Algorithm with Minimax Optimal Rates. *Journal of Machine Learning Research*.
- [J6] **Y. Zhang**, J. Duchi and M. Wainwright. Communication-Efficient Algorithms for Statistical Optimization. *Journal of Machine Learning Research*.
- [J7] **Y. Zhang** and X. Sun. The Antimagicalness of the Cartesian Product of Graphs. *Theoretical Computer Science*.

## Conference Publications

- [C1] **Y. Zhang**, P. Pasupat, P. Liang. Macro Grammars and Holistic Triggering for Efficient Semantic Parsing. *Empirical Methods on Natural Language Processing (EMNLP)*, 2017.
- [C2] **Y. Zhang**, P. Liang, M. Wainwright. Convexified Convolutional Neural Networks. *International Conference on Machine Learning (ICML)*, 2017.
- [C3] **Y. Zhang**, P. Liang, M. Charikar. A Hitting Time Analysis of Stochastic Gradient Langevin Dynamics. *Conference on Learning Theory (COLT)*, 2017 (**Best paper award**).
- [C4] **Y. Zhang**, J.D. Lee, M. Wainwright and M.I. Jordan. On the Learnability of Fully-connected Neural Networks. *Artificial Intelligence and Statistics (AISTATS)*, 2017.
- [C5] C. Jin, **Y. Zhang**, S. Balakrishnan, M. Wainwright, M.I. Jordan. Local Maxima in the Likelihood of Gaussian Mixture Models: Structural Results and Algorithmic Consequences. *Neural Information Processing Systems (NIPS)*, 2016.
- [C6] **Y. Zhang**, J.D. Lee, M.I. Jordan.  $\ell_1$ -regularized Neural Networks are Improperly Learnable in Polynomial Time. *International Conference on Machine Learning (ICML)*, 2016.
- [C7] **Y. Zhang**, M. Wainwright and M.I. Jordan. Distributed Estimation of Generalized Matrix Rank: Efficient Algorithms and Lower Bounds. *International Conference on Machine Learning (ICML)*, 2015.
- [C8] **Y. Zhang** and L. Xiao. DiSCO: Communication-Efficient Distributed Optimization of Self-Concordant Loss. *International Conference on Machine Learning (ICML)*, 2015.
- [C9] **Y. Zhang** and L. Xiao. Stochastic Primal-Dual Coordinate Method for Regularized Empirical Risk Minimization. *International Conference on Machine Learning (ICML)*, 2015.
- [C10] **Y. Zhang**, X. Chen, D. Zhou and M.I. Jordan. Spectral Methods meet EM: A Provably Optimal Algorithm for Crowdsourcing. *Neural Information Processing Systems (NIPS)*, 2014. (**Spotlight presentation, 4.8% acceptance rate**)
- [C11] **Y. Zhang**, M. Wainwright and M.I. Jordan. Lower Bounds on the Performance of Polynomial-time Algorithms for Sparse Linear Regression. *Conference on Learning Theory (COLT)*, 2014.
- [C12] **Y. Zhang**, A. Ahmed, V. Josifovski and A. Smola. Taxonomy Discovery for Personalized Recommendation. *ACM International Conference on Web Search and Data Mining (WSDM)*, 2014.
- [C13] **Y. Zhang**, J. Duchi, M. Wainwright and M.I. Jordan. Information-theoretic Lower Bounds for Distributed Statistical Estimation with Communication Constraints. *Neural Information Processing Systems (NIPS)*, 2013. (**Oral presentation, 1.4% acceptance rate**)

- [C14] **Y. Zhang**, J. Duchi and M. Wainwright. Divide and Conquer Kernel Ridge Regression. *Conference on Learning Theory (COLT)*, 2013.
- [C15] **Y. Zhang**, J. Duchi and M. Wainwright. Communication-Efficient Algorithms for Statistical Optimization. *Neural Information Processing Systems (NIPS)*, 2012.
- [C16] W. Chen, D. Wang, **Y. Zhang** and Q. Yang. Understanding Click Noise: A Noise-aware Click Model for Web Search. *ACM International Conference on Web Search and Data Mining (WSDM)*, 2012.
- [C17] **Y. Zhang**, W. Chen and D. Wang, Q. Yang. User-click Modeling for Understanding and Predicting Search-behavior. *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, 2011.
- [C18] B. Hu, **Y. Zhang**, G. Wang, Q. Yang, W. Chen. Characterize Search Intent Diversity into Click Models. *International World Wide Web Conference (WWW)*, 2011.
- [C19] **Y. Zhang**, D. Wang, G. Wang, W. Chen, Z. Zhang, B. Hu and L. Zhang. Learning Click Model via Probit Bayesian Inference. *ACM International Conference on Information and Knowledge Management (CIKM)*, 2010.
- [C20] D. Wang, W. Chen, G. Wang, **Y. Zhang** and B. Hu. Explore Click Models for Search Ranking. *ACM International Conference on Information and Knowledge Management (CIKM)*, short paper, 2010.
- [C21] F. Zhong, D. Wang, G. Wang, W. Chen, **Y. Zhang**, Z. Chen and H. Wang. Incorporating Post-Click Behaviors Into a Click Model. *Annual International ACM SIGIR Conference (SIGIR)*, 2010.
- [C22] **Y. Zhang** and L. Zhang. Extracting Independent Rules: a New Perspective of Boosting. *International Symposium on Artificial Intelligence and Mathematics (ISAIM)*, 2010.

## Teaching

Graduate Student Instructor, Introduction to machine learning, UC Berkeley

**Spring, 2015**

Graduate Student Instructor, Randomized algorithms for matrices and data, UC Berkeley

**Fall, 2013**

## Service

**Journal Reviewer:** Journal of Machine Learning Research, Annals of Statistics, Mathematical Programming, ACM Transactions on the Web.

**Conference Reviewer:** ICML (2013 - ), NIPS (2013 - ), AISTAT (2015 - ), IJCAI (2015 - ).

## Programming

Capable of Python, C/C++, C#, Java, Scala, MATLAB.

## References

**Michael I. Jordan**

Pehong Chen Distinguished Professor  
EECS and Statistics, UC Berkeley  
jordan@cs.berkeley.edu

**Martin J. Wainwright**

Professor  
EECS and Statistics, UC Berkeley  
wainwrig@eecs.berkeley.edu

**Lin Xiao**

Principle Researcher  
Machine Learning Department  
Microsoft Research Redmond  
lin.xiao@microsoft.com

**Percy Liang**

Assistant Professor  
Computer Science Department  
Stanford University  
pliang@cs.stanford.edu