

# JERRY ZHENG LI

Principal Researcher  
PhysAGI (formerly Machine Learning Foundations) Group  
Microsoft Research

jerrl@microsoft.com  
<https://jerryzli.github.io/>

## EDUCATION

### Massachusetts Institute of Technology

Ph.D., September 2018  
Thesis: Principled Approaches to Robust Machine Learning and Beyond  
Advisor: Ankur Moitra  
Electrical Engineering and Computer Science

### Massachusetts Institute of Technology

S.M., September 2014  
Thesis: The SprayList: A Scalable Relaxed Priority Queue  
Advisor: Nir Shavit  
Electrical Engineering and Computer Science

### University of Washington

B.S., magna cum laude, May 2013  
Computer Science & Engineering, Mathematics

## OTHER ACADEMIC POSITIONS

**VMWare Research Fellow, Simons Institute**, Fall 2019

**Research Intern, Google Brain**, Summer 2018  
Hosted by Vitaly Feldman

**Research Intern, Microsoft Research, Cambridge**, Spring 2016  
Hosted by Dan Alistarh and Milan Vojnovic

## PUBLICATIONS

(All authors are alphabetical unless stated otherwise)

### Journal papers

#### **The Complexity of NISQ**

Sitan Chen, Jordan Cotler, Hsin-Yuan Huang, Jerry Li  
to appear, Nature Communications  
preliminary version in Advances in Quantum Information Processing 22 (QIP 2023)

#### **Quantum Advantage in Learning from Experiments**

(by contribution) Hsin-Yuan Huang, Michael Broughton, Jordan Cotler, Sitan Chen, Jerry Li, Masoud Mohseni, Hartmut Neven, Ryan Babbush, Richard Kueng, John Preskill, Jarrod R. McClean  
Science, 376 (6598), 2022.

#### **Robustness Meets Algorithms**

Ilias Diakonikolas, Gautam Kamath, Daniel M. Kane, Jerry Li, Ankur Moitra, Alistair Stewart  
Communications of the ACM May 2021, Research Highlights

#### **Robust Estimators in High Dimensions without the Computational Intractability**

Ilias Diakonikolas, Gautam Kamath, Daniel M. Kane, Jerry Li, Ankur Moitra, Alistair Stewart  
SIAM Journal on Computing, 48(2), 2019. Special Issue for FOCS 2016.

#### **Exact Model Counting of Query Expressions: Limitations of Propositional Methods**

Paul Beame, Jerry Li, Sudeepa Roy, Dan Suciu.  
ACM Transactions on Database Systems (TODS), Vol. 42, Issue 1, pages 1:1-1:46, March 2017.

## Conference and workshop papers

### **Structured Semidefinite Programming for Recovering Structured Preconditioners**

Arun Jambulapati, Jerry Li, Christopher Musco, Kirankumar Shiragur, Aaron Sidford, Kevin Tian  
to appear, Advances in Neural Information Processing Systems 36  
preliminary version in 14th International OPT Workshop on Optimization for Machine Learning (OPT 2022)

### **Automatic Prompt Optimization with “Gradient Descent” and Beam Search**

Reid Pryzant, Dan Iter, Jerry Li, Yin Tat Lee, Chenguang Zhu, Michael Zeng  
to appear, Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP 2023)

### **The Full Landscape of Robust Mean Testing: Sharp Separations between Oblivious and Adaptive Contamination**

Clément Canonne, Samuel B. Hopkins, Jerry Li, Allen Liu, Shyam Narayanan  
Proceedings of the 64th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2023)  
**Invited to appear in special issue of SIAM Journal on Computing for FOCS 2023**

### **Matrix Completion in Almost-Verification Time**

Jon Kelner, Jerry Li, Allen Liu, Aaron Sidford, Kevin Tian  
Proceedings of the 64th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2023)

### **When Does Adaptivity Help for Quantum State Learning?**

Sitan Chen, Brice Huang, Jerry Li, Allen Liu, Mark Sellke  
Proceedings of the 64th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2023)  
preliminary version in Quantum Information Processing 22 (QIP 2023), merged with [CHLL22]

### **Query lower bounds for log-concave sampling**

Sinho Chewi, Jaume de Dios Pont, Jerry Li, Chen Lu, Shyam Narayanan  
Proceedings of the 64th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2023)

### **Sampling Is as Easy as Learning the Score: Theory for Diffusion Models With Minimal Data Assumptions**

Sitan Chen, Sinho Chewi, Jerry Li, Yuanzhi Li, Adil Salim, Anru R. Zhang  
Proceedings of the 11th International Conference on Learning Representations (ICLR 2023), **Notable top 5%**

### **Learning Polynomial Transformations**

Sitan Chen, Jerry Li, Yuanzhi Li, Anru R. Zhang  
Proceedings of the 55th ACM Symposium on Theory of Computing (STOC 2023)

### **REAP: A Large-Scale Realistic Adversarial Patch Benchmark**

Nabeel Hingun, Chawin Sitawarin, Jerry Li, David Wagner  
Proceedings of the 36th International Conference on Computer Vision (ICCV 2023)

### **Learning (Very) Simple Generative Models Is Hard**

Sitan Chen, Jerry Li, Yuanzhi Li  
Advances in Neural Information Processing Systems 35 (NeurIPS 2022)

### **Robust Model Selection and Nearly-Proper Learning for GMMs**

Jerry Li, Allen Liu, Ankur Moitra  
Advances in Neural Information Processing Systems 35 (NeurIPS 2022)

### **Tight Bounds for Quantum State Certification with Incoherent Measurements**

Sitan Chen, Brice Huang, Jerry Li, Allen Liu  
Proceedings of the 63th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2022)  
Quantum Information Processing 22 (QIP 2023), merged with [CHLLS23]

### **The Price of Tolerance in Distribution Testing**

Clément Canonne, Gautam Kamath, Ayush Jain, Jerry Li  
Proceedings of the 35th Annual Conference on Learning Theory (COLT 2022)

### **Clustering Mixtures with Almost Optimal Separation in Polynomial Time**

Jerry Li, Allen Liu

Proceedings of the 54th ACM Symposium on Theory of Computing (STOC 2022)

**Invited to appear in special issue of SIAM Journal on Computing for STOC 2022**

**Clustering Mixture Models in Almost-Linear Time via List-Decodable Mean Estimation**

Ilias Diakonikolas, Daniel M. Kane, Daniel Kongsgaard, Jerry Li, Kevin Tian

Proceedings of the 54th ACM Symposium on Theory of Computing (STOC 2022)

**Minimax Optimality (Probably) Doesn't Imply Distribution Learning for GANs**

Sitan Chen, Jerry Li, Yuanzhi Li, Raghu Meka

Proceedings of the 10th International Conference on Learning Representations (ICLR 2022)

**Toward Instance-Optimal State Certification With Incoherent Measurements**

Sitan Chen, Jerry Li, Ryan O'Donnell

Quantum Information Processing 21 (QIP 2022)

Proceedings of the 35th Annual Conference on Learning Theory (COLT 2022)

**Robust Regression Revisited: Acceleration and Improved Estimation Rates**

Arun Jambulapati, Jerry Li, Tselil Schramm, Kevin Tian

Advances in Neural Information Processing Systems 34 (NeurIPS 2021)

**List-Decodable Mean Estimation in Nearly-PCA Time**

Ilias Diakonikolas, Daniel M. Kane, Daniel Kongsgaard, Jerry Li, Kevin Tian

Advances in Neural Information Processing Systems 34 (NeurIPS 2021), **Spotlight Presentation**

**A Hierarchy for Replica Quantum Advantage**

Sitan Chen, Jordan Cotler, Hsin-Yuan Huang, Jerry Li

Quantum Information Processing 21 (QIP 2022), merged with [CCHL21] below

**Exponential Separations between Learning With and Without Quantum Memory**

Sitan Chen, Jordan Cotler, Hsin-Yuan Huang, Jerry Li

Proceedings of the 62th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2021)

Quantum Information Processing 21 (QIP 2022)

**Invited to appear in special issue of SIAM Journal on Computing for FOCS 2021**

**Finding the Mode of a Kernel Density Estimate**

Jasper C.H. Lee, Jerry Li, Christopher Musco, Jeff M. Phillips, Wai Ming Tai

Proceedings of the 29th European Symposium on Algorithms (ESA 2021)

**Statistical Query Algorithms and Low-Degree Tests Are Almost Equivalent**

Matthew Brennan, Guy Bresler, Samuel B. Hopkins, Jerry Li, Tselil Schramm

Proceedings of the 34th Annual Conference on Learning Theory (COLT 2021), **Best Paper Runner Up**

**Aligning AI With Shared Human Values**

Dan Hendrycks, Collin Burns, Steven Basart, Andrew Critch, Jerry Li, Dawn Song, Jacob Steinhardt

Proceedings of the 9th International Conference on Learning Representations (ICLR 2021)

**Byzantine-Resilient Non-Convex Stochastic Gradient Descent**

Dan Alistarh, Zeyuan Allen-Zhu, Faeze Ebrahimiaghazani, Jerry Li

Proceedings of the 9th International Conference on Learning Representations (ICLR 2021)

**Robust and Heavy-Tailed Mean Estimation Made Simple, via Regret Minimization**

Samuel B. Hopkins, Jerry Li, Fred Zhang

Advances in Neural Information Processing Systems 33 (NeurIPS 2020)

**Robust Sub-Gaussian Principal Component Analysis and Width-Independent Schatten Packing**

Arun Jambulapati, Jerry Li, Kevin Tian

Advances in Neural Information Processing Systems 33 (NeurIPS 2020), **Spotlight Presentation**

**Learning Structured Distributions From Untrusted Batches: Faster and Simpler**

Sitan Chen, Jerry Li, Ankur Moitra

Advances in Neural Information Processing Systems 33 (NeurIPS 2020)

**Robust Covariance Estimation in Nearly-Matrix Multiplication Time**

Jerry Li, Guanghao Ye

Advances in Neural Information Processing Systems 33 (NeurIPS 2020)

**Entanglement is Necessary for Optimal Quantum Property Testing**

Sébastien Bubeck, Sitan Chen, Jerry Li

Proceedings of the 61th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2020)

**Randomized Smoothing of All Shapes and Sizes**

(by contribution) Greg Yang, Tony Duan, Edward Hu, Hadi Salman, Ilya Razenshteyn, Jerry Li

Proceedings of the 37th International Conference on Machine Learning (ICML 2020)

**Positive Semidefinite Programming: Mixed, Parallel, and Width-Independent**

Arun Jambulapati, Yin Tat Lee, Jerry Li, Swati Padmanabhan, Kevin Tian

Proceedings of the 52th ACM Symposium on Theory of Computing (STOC 2020)

**Efficiently Learning Structured Distributions from Untrusted Batches**

Sitan Chen, Jerry Li, Ankur Moitra

Proceedings of the 52th ACM Symposium on Theory of Computing (STOC 2020)

**Learning Mixtures of Linear Regressions in Subexponential Time via Fourier Moments**

Sitan Chen, Jerry Li, Zhao Song

Proceedings of the 52th ACM Symposium on Theory of Computing (STOC 2020)

**Low-rank Toeplitz Matrix Estimation via Random Ultra-Sparse Rulers**

Hannah Lawrence, Jerry Li, Cameron Musco, Christopher Musco

Proceedings of the 45th International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2020)

**The Sample Complexity of Toeplitz Covariance Estimation**

Yonina Eldar, Jerry Li, Cameron Musco, Christopher Musco

Proceedings of the 31st Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2020)

**Provably Robust Deep Learning via Adversarially Trained Smoothed Classifiers**

Hadi Salman, Greg Yang, Jerry Li, Pengchuan Zhang\*, Huan Zhang\*, Ilya Razenshteyn\*, Sébastien Bubeck\*

\* reverse alphabetical order

Advances in Neural Information Processing Systems 32 (NeurIPS 2019), **Spotlight Presentation**

**Quantum Entropy Scoring for Fast Robust Mean Estimation and Improved Outlier Detection**

Yihe Dong, Samuel B. Hopkins, Jerry Li

Advances in Neural Information Processing Systems 32 (NeurIPS 2019), **Spotlight Presentation**

**Sever: A Robust Meta-Algorithm for Stochastic Optimization**

Ilias Diakonikolas, Daniel Kane, Gautam Kamath, Jerry Li, Jacob Steinhardt, Alistair Stewart

Preliminary version in preliminary version in Workshop on Secure ML (SecML 2018), **Oral Presentation**

Proceedings of the 36th International Conference on Machine Learning (ICML 2019)

**How hard is Robust Mean Estimation?**

Samuel B. Hopkins and Jerry Li

Proceedings of the 32nd Annual Conference on Learning Theory (COLT 2019)

**On Mean Estimation For General Norms with Statistical Queries**

Alexandr Andoni, Jerry Li, Ilya Razenshteyn, Aleksandar Nikolov, Erik Waingarten

Proceedings of the 32nd Annual Conference on Learning Theory (COLT 2019)

**Differentially Private Estimation in High Dimensions**

Gautam Kamath, Jerry Li, Vikrant Singhal, Jonathan Ullman

Preliminary version in Theory and Practice of Differential Privacy (TPDP 2018)

Proceedings of the 32nd Annual Conference on Learning Theory (COLT 2019)

**Spectral Signatures for Backdoor Attacks**

(by contribution) Brandon Tran, Jerry Li, Aleksander Mądry  
Advances in Neural Information Processing Systems 31 (NIPS 2018)

**Byzantine Stochastic Gradient Descent**

Dan Alistarh, Zeyuan Allen-Zhu, Jerry Li  
Advances in Neural Information Processing Systems 31 (NIPS 2018)

**On the Limitations of First Order Approximation in GAN dynamics**

Jerry Li, Aleksander Mądry, John Peebles, Ludwig Schmidt  
Preliminary version in Workshop on Principled Approaches to Deep Learning (PADL 2017) as *Towards Understanding the Dynamics of Generative Adversarial Networks*  
Proceedings of the 35th International Conference on Machine Learning (ICML 2018)

**Fast and Sample-Efficient Algorithms for Learning Multidimensional Histograms**

Ilias Diakonikolas, Jerry Li, Ludwig Schmidt  
Proceedings of the 31st Conference on Learning Theory (COLT 2018)

**Asynchronous Balanced Allocations with Applications to Approximate Counting**

Dan Alistarh, Justin Kopinsky, Jerry Li, Giorgi Nadiradze  
Proceedings of the 30th ACM Symposium on Parallelism in Algorithms and Architectures (SPAA 2018)

**Mixture Models, Robustness, and Sum of Squares Proofs**

Samuel B. Hopkins and Jerry Li  
Proceedings of the 50th ACM Symposium on Theory of Computing (STOC 2018)

**Robustly Learning a Gaussian in High Dimensions: Getting Optimal Error, Efficiently**

Ilias Diakonikolas, Gautam Kamath, Daniel M. Kane, Jerry Li, Ankur Moitra, Alistair Stewart  
Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2018)

**Communication-Efficient Distributed Learning of Discrete Distributions**

Ilias Diakonikolas, Elena Grigorescu, Jerry Li, Abhiram Natarajan, Krzysztof Onak, Ludwig Schmidt  
Advances in Neural Information Processing Systems 30 (NIPS 2017), **Oral Presentation**

**QSGD: Communication-Efficient SGD via Gradient Quantization and Encoding**

Dan Alistarh, Demjan Grubić, Jerry Li, Ryota Tomioka, Milan Vojnovic  
Preliminary version appeared in Optimization for Machine Learning 2016  
Advances in Neural Information Processing Systems 30 (NIPS 2017)  
**Spotlight Presentation**

**Being Robust (in High Dimensions) Can Be Practical**

Ilias Diakonikolas, Gautam Kamath, Daniel M. Kane, Jerry Li, Ankur Moitra, Alistair Stewart  
Proceedings of the 34th International Conference on Machine Learning (ICML 2017)

**ZipML: An End-to-end Bitwise Framework for Dense Generalized Linear Models**

(by contribution) Hantian Zhang\*, Jerry Li\*, Kaan Kara, Dan Alistarh, Ji Liu, Ce Zhang  
\*equal contribution  
Proceedings of the 34th International Conference on Machine Learning (ICML 2017)

**The Power of Choice in Priority Scheduling**

Dan Alistarh, Justin Kopinsky, Jerry Li, Giorgi Nadiradze  
Proceedings of the 25th ACM Symposium on Principles of Distributed Computing (PODC 2017)

**Robust Sparse Estimation Tasks in High Dimensions**

Jerry Li  
Proceedings of the 30th Annual Conference on Learning Theory (COLT 2017)  
Merged with *Computationally Efficient Robust Estimation of Sparse Functionals*

**Robust Proper Learning for Mixtures of Gaussians via Systems of Polynomial Inequalities**

Jerry Li, Ludwig Schmidt  
Proceedings of the 30th Annual Conference on Learning Theory (COLT 2017)

**Sample Optimal Density Estimation in Nearly-Linear Time**

Jayadev Acharya, Ilias Diakonikolas, Jerry Li, Ludwig Schmidt

Proceedings of the 28th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2017)

**Robust Estimators in High Dimensions without the Computational Intractability**

Ilias Diakonikolas, Gautam Kamath, Daniel M. Kane, Jerry Li, Ankur Moitra, Alistair Stewart

Proceedings of the 57th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2016)

**Invited to the SIAM Journal on Computing Special Issue for FOCS 2016**

**Invited to Highlights of Algorithms 2017 (HALG 2017)**

**Fast Algorithms for Segmented Regression**

Jayadev Acharya, Ilias Diakonikolas, Jerry Li, Ludwig Schmidt

Proceedings of the 33th International Conference on Machine Learning (ICML 2016)

**Replacing Mark Bits with Randomness in Fibonacci Heaps**

Jerry Li, John Peebles

Proceedings of the 42nd International Colloquium on Automata, Languages, and Programming (ICALP 2015)

**Fast and Near-Optimal Algorithms for Approximating Distributions by Histograms**

Jayadev Acharya, Ilias Diakonikolas, Chinmay Hegde, Jerry Li, Ludwig Schmidt.

Proceedings of the 23rd ACM Symposium on Principles of Database Systems (PODS 2015)

**The SprayList: A Scalable Relaxed Priority Queue**

Dan Alistarh, Justin Kopinsky, Jerry Li, Nir Shavit

Proceedings of 20th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming (PPoPP 2015)

**Best Artifact Award**

(see also my master's thesis)

**On the Importance of Registers for Computability**

Rati Gelashvili, Mohsen Ghaffari, Jerry Li, Nir Shavit

Proceedings of the 18th International Conference on Principles of Distributed Systems (OPODIS 2014)

The following two papers are subsumed by the journal paper *Exact Model Counting of Query Expressions: Limitations of Propositional Methods*:

**Model Counting of Query Expressions: Limitations of Propositional Methods**

Paul Beame, Jerry Li, Sudeepa Roy, Dan Suciu.

Proceedings of the 17th International Conference on Database Theory (ICDT 2014)

**Lower bounds for exact model counting and applications in probabilistic databases**

Paul Beame, Jerry Li, Sudeepa Roy, Dan Suciu.

Proceedings of the 29th Conference on Uncertainty in Artificial Intelligence (UAI 2013)

**Plenary Presentation**

**THESES**

**Principled Approaches to Robust Machine Learning and Beyond**

Ph.D thesis, 2018

**George M. Sprowls Award for outstanding Ph.D. theses in EECS at MIT**

**The SprayList: A Scalable Relaxed Priority Queue**

Masters thesis, 2015

**PATENTS**

**Efficient training of neural networks**

Dan Alistarh, Jerry Li, Ryota Tomioka, Milan Vojnovic

in submission

**TEACHING EXPERIENCE**

## Past Interns / Residents

Jaume de Dios Pont (Research Intern, Summer 2023), co-advised with Adil Salim.  
Jane Lange (Research Intern, Summer 2023)  
Sidhanth Mohanty (Research Intern, Summer 2022)  
Allen Liu (Research Intern, Summer 2021, Summer 2022)  
Huiying Li (Research Intern, Summer 2020), co-advised with Ece Kamar and Emre Kıcıman.  
Kai Xiao (Research Intern, Summer 2020), co-advised with Sébastien Bubeck.  
Kevin Tian (Research Intern, Summer 2020).  
Ivan Evtimov (Research Intern, Spring 2020), co-advised with Weidong Cui, Ece Kamar, and Emre Kıcıman.  
Tony Duan (MSR AI Resident, 2019—2020).  
Hadi Salman (MSR AI Resident, 2018—2019).  
Sitan Chen (Research Intern, Summer 2019)

## Past Classes

<b>Instructor</b>	University of Washington	Fall 2019
CSE 499M: Principled Approaches to Robust Machine Learning		Fall 2019
<b>Teaching Assistant</b>	Massachusetts Institute of Technology	Fall 2014
6.852: Distributed Algorithms		Fall 2014
<b>Teaching Assistant</b>	University of Washington	Spring 2012 - Spring 2013
Mathematics REU		Summer 2013
MATH 334/5/6: Advanced Accelerated Second Year Honors Calculus		Fall 2012-Spring 2013
CSE 373: Algorithms and Data Structures		Spring 2012
CSE 344: Databases		Winter 2012

## RELEVANT COURSEWORK

MIT: Graph Theory and Combinatorics, Topics in Combinatorics, Machine Learning, Distributed Systems, Advanced Algorithms, Distributed Algorithms, Multicore Programming, Randomness and Computation, Algorithmist's Toolkit.

UW: Randomized Algorithms, Databases, Advanced Real Analysis, Advanced Complex Analysis, Algebraic and Smooth Topology, Continuous and Discrete Probability, Group Theory.

## HONORS AND AWARDS

George M. Sprowls Award for outstanding Ph.D. theses in EECS at MIT	2018
Simons Fellowship	2018
NSF Graduate Research Fellowship	2014-2017
NDSEG Graduate Research Fellowship (declined)	2014
MIT Akamai Presidential Graduate Fellowship	September 2013 - May 2014
University of Washington Outstanding Graduating Senior in Mathematics	May 2013
Microsoft Endowed Scholarship	2013-2014
University of Washington Gullickson Award for Outstanding Junior in Mathematics	May 2012
Jerre Noe Endowed Scholarship	2012-2013
Meritorious Winner Mathematical Contest in Modeling	Fall 2012
University of Washington Outstanding Second Year Student in Mathematics	May 2010
Outstanding Winner Mathematical Contest in Modeling	Fall 2010
National Merit Scholarship Finalist	Fall 2008
Early Entrance Program Scholarship	Fall 2008
Dean's list at University of Washington	Fall 2009 - Spring 2013

## SERVICE

PC member for: STOC 22, SODA 2022, SODA 2024, ITCS 2024  
Organizer of MIT Theory Lunch (Fall 2013 - Summer 2014)  
External reviewer for: STOC, FOCS, SODA, COLT, ITCS, RANDOM, ICALP, NIPS, ICML, ICDT, DISC, JMLR, AISTATS, IEEE Transactions on Information Theory

## REFERENCES

available upon request