

# Yi Sun

CONTACT INFORMATION	Address: Department of Statistics, The University of Chicago, Chicago, IL 60637. Email: <a href="mailto:ysisun@statistics.uchicago.edu">ysisun@statistics.uchicago.edu</a> Webpage: <a href="http://ysisun.io">ysisun.io</a>
RESEARCH	Probability and applications to machine learning and high-dimensional statistics.
EMPLOYMENT	<b>The University of Chicago</b> Chicago, IL Assistant Professor (tenure-track), 2020–present. <b>Columbia University</b> New York, NY Joseph F. Ritt Assistant Professor, 2019–2020; Simons Fellow, 2016–2019.
EDUCATION	<b>Massachusetts Institute of Technology</b> Cambridge, MA Ph.D., Mathematics, advised by Pavel Etingof, 2011–2016. <b>University of Cambridge</b> Cambridge, UK M.A.St., Mathematics, with distinction, 2010–2011. <b>Harvard University</b> Cambridge, MA A.M., Mathematics, 2006–2010. A.B., Mathematics, <i>magna cum laude</i> , with secondary field in Economics, 2006–2010. Phi Beta Kappa (one of 24 juniors inducted)
GRANTS FELLOWSHIPS AND AWARDS	NSF Grant DMS-2054838, 2021–2024. (\$238,603, Highly Recommended) NSF Grant DMS-1701654/2039183, 2017–2021. (\$141,999, Highly Recommended) Simons Junior Fellowship, 2016–2019. (\$364,214) Open Philanthropy Project Grant, 2019. (\$10,000, co-PI) NSF Mathematical Sciences Postdoctoral Research Fellowship, 2016–2019 (declined). Johnson Prize for best research paper by MIT graduate student in mathematics, 2016. NSF Graduate Research Fellowship, 2012–2015. Churchill Scholarship, 2010–2011. / MIT Praecis Presidential Fellowship, 2011–2012. COMAP Math Contest in Modeling, Outstanding Winner, SIAM Prize, 2008 and 2009. Intel Science Talent Search, 2 <sup>nd</sup> Place, 2006. / Putnam Competition, 10 <sup>th</sup> Place, 2009. Int'l Math Olym., Silver Medal, 2006. / Asian Pacific Math Olym., Gold Medal, 2005. Int'l Physics Olympiad, Gold Medal, 2004. / USA Computing Olympiad, Finalist, 2005.
MATHEMATICS AND STATISTICS RESEARCH	25. <i>Maximum likelihood for high-noise group orbit estimation and single-particle cryo-EM</i> (with Z. Fan, R. Lederman, T. Wang, and S. Xu), Ann. Stat. <b>52</b> (2024), 52–77. <a href="https://arxiv.org/abs/2107.01305">arXiv:2107.01305</a> 24. <i>Likelihood landscape and maximum likelihood estimation for the discrete orbit recovery model</i> (with Z. Fan, T. Wang, and Y. Wu), Comm. Pure Appl. Math. <b>76</b> (2023), 1208–1302. <a href="https://arxiv.org/abs/2004.00041">arXiv:2004.00041</a> 23. <i>Probabilistic conformal blocks for Liouville CFT on the torus</i> (with P. Ghosal, G. Remy, and X. Sun), Duke Math. J., to appear. <a href="https://arxiv.org/abs/2003.03802">arXiv:2003.03802</a> 22. <i>Principal components in linear mixed models with general bulk</i> (with Z. Fan and Z. Wang), Ann. Stat. <b>49</b> (2021), 1489–1513. <a href="https://arxiv.org/abs/1903.09592">arXiv:1903.09592</a> 21. <i>Gaussian fluctuations for products of random matrices</i> (with V. Gorin), Amer. J. Math. <b>144</b> (2022), 287–393. <a href="https://arxiv.org/abs/1812.06532">arXiv:1812.06532</a> 20. <i>Spiked covariances and principal components analysis in high-dimensional random effects models</i> (with Z. Fan and I. Johnstone), preprint, 2018. <a href="https://arxiv.org/abs/1806.09529">arXiv:1806.09529</a> 19. <i>Affine Macdonald conjectures and special values of Felder-Varchenko functions</i> (with E. Rains and A. Varchenko), Sel. Math. N. S. <b>24</b> (2018), 1549–1591. <a href="https://arxiv.org/abs/1610.01917">arXiv:1610.01917</a> 18. <i>Laguerre and Jacobi analogues of the Warren process</i> (single author, with an appendix by A. Sarantsev), submitted, 2017. <a href="https://arxiv.org/abs/1610.01635">arXiv:1610.01635</a> 17. <i>Traces of intertwiners for quantum affine algebras and difference equations (after Etingof-Schiffmann-Varchenko)</i> (single author), Transform. Groups <b>23</b> (2018), 1167–1215. <a href="https://arxiv.org/abs/1609.09038">arXiv:1609.09038</a>

16. *Matrix models for multilevel Heckman-Opdam and multivariate Bessel measures* (single author), Ann. inst. Henri Poincaré (B) Probab. Stat., to appear. [arXiv:1609.09096](#)
  15. *Traces of intertwiners for quantum affine  $\mathfrak{sl}_2$  and Felder-Varchenko functions* (single author), Commun. Math. Phys. **347** (2016), 573-653. [arXiv:1508.03918](#)
  14. *The polynomial representation of the type  $A_{n-1}$  rational Cherednik algebra in characteristic  $p \mid n$*  (with S. Devadas), Commun. Algebra **45** (2016), 1926-1934. [arXiv:1505.07891](#)
  13. *A representation-theoretic proof of the branching rule for Macdonald polynomials* (single author), Math. Res. Lett. **23** (2016), 887-927. [arXiv:1412.0714](#)
  12. *A new integral formula for Heckman-Opdam hypergeometric functions* (single author), Adv. Math. **289** (2016), 1157-1204. [arXiv:1406.3772](#)
  11. *Finite dimensional representations of the rational Cherednik algebra for  $G_4$*  (single author), J. Algebra **323** (2010), 2864-2887. [arXiv:0910.5527](#)
- COMPUTER  
SCIENCE  
RESEARCH
10. D. Kang, T. Hashimoto, I. Stoica, and Y. Sun, *ZK-IMG: Attested Images via Zero-Knowledge Proofs to Fight Disinformation*, preprint, 2022. [arXiv:2211.04775](#)
  9. D. Kang, T. Hashimoto, I. Stoica, and Y. Sun, *Scaling up trustless DNN inference with zero-knowledge proofs*, preprint, 2022. [arXiv:2210.08674](#)
  8. B. Hanin\* and Y. Sun\*, *How data augmentation affects optimization for linear regression*, NeurIPS 2021. DeepMath 2020, OPT 2020. [arXiv:2010.11171](#)
  7. D. Kang, A. Derhacopian, K. Tsuji, T. Hebert, P. Bailis, T. Fukami, T. Hashimoto, Y. Sun, M. Zaharia, *Exploiting proximity search and easy examples to select rare events*, NeurIPS DCAI workshop 2021.
  6. D. Kang\*, J. Guibas\*, P. Bailis, T. Hashimoto, Y. Sun, and M. Zaharia, *Accelerating Approximate Aggregation Queries with Expensive Predicates*, VLDB 2021. [arXiv:2108.06313](#)
  5. D. Kang\*, Y. Sun\*, D. Hendrycks, T. Brown, and J. Steinhardt, *Testing robustness against unforeseen adversaries*, submitted, 2019. [arXiv:1908.08016](#)
  4. T. Hashimoto, Y. Sun, and T. Jaakkola, *From random walks to distances on unweighted graphs*, NIPS 2015. [arXiv:1511.00573](#)
  3. T. Hashimoto, Y. Sun, and T. Jaakkola, *Metric recovery from directed unweighted graphs*, NIPS 2014 workshop (Best Student Paper), AISTATS 2015. [arXiv:1411.5720](#)
  2. Y. Sun and M. Sundararajan, *Axiomatic attribution for multilinear functions*, ACM Conf. on Electronic Commerce 2011. [arXiv:1102.0989](#)
- OTHER  
RESEARCH
1. P. Y. Wang, Y. Sun, R. Axel, LF Abbott, and R. G. Yang, *Evolving the olfactory system with machine learning*, Neuron, in press, 2021. CCN 2019, NeurIPS 2019 Neuro+AI Workshop.
- RESEARCH  
PRESENTATIONS
56. NeurIPS 2021 December 2021  
How data augmentation affects optimization for linear regression (poster)
  55. Princeton: Wilks Statistics Seminar October 2021  
Maximum likelihood for high-noise group orbit estimation and single-particle cryo-EM
  54. Luminy: Modern analysis related to root systems with applications October 2021  
Gaussian fluctuations for products of random matrices
  53. Simons Society of Fellows Alumni Symposium October 2021  
Maximum likelihood for high-noise group orbit estimation and single particle cryo-electron microscopy
  52. Online conference on Integrability in Conformal Probability October 2021  
Probabilistic construction of conformal blocks for Liouville CFT on the torus
  51. UChicago: Statistics Consulting Seminar February 2021  
Learning under a group action and the orbit recovery problem
  50. UChicago: Probability Seminar February 2021

Probabilistic conformal blocks for Liouville CFT on the torus	
49. NeurIPS 2020 Workshop: OPT 2020 Data augmentation as stochastic optimization (poster)	December 2020
48. DeepMath 2020 Data augmentation as stochastic optimization	November 2020
47. Bernoulli-IMS One World Symposium Likelihood landscape and maximum likelihood estimation for the discrete orbit recovery model	August 2020
46. Google X Testing robustness against unforeseen adversaries	March 2020
45. UW Madison: Mathematics Colloquium Fluctuations for products of random matrices	February 2020
44. UChicago: Statistics Colloquium Fluctuations for products of random matrices	January 2020
43. AMS Fall Western Sectional Meeting Fluctuations for products of random matrices	November 2019
42. ICML 2019 Workshop: Uncertainty and Robustness in DL (poster) Transfer of robustness against adversarial and stochastic distortions	June 2019
41. OpenAI Transfer of robustness against adversarial and stochastic distortions	June 2019
40. Virginia: Integrable Probability Summer School Fluctuations for products of random matrices	June 2019
39. UCSD: Probability Seminar Fluctuations for products of random matrices	January 2019
38. Yale: Geometry, Symmetry, and Physics Seminar Affine Macdonald conjectures and special values of Felder-Varchenko functions	April 2018
37. Simons Society of Fellows Retreat A probabilistic view on random covariance matrices	February 2018
36. PCMI: Research Program on Random Matrices Algebraic structures for multilevel eigenvalue densities	July 2017
35. Rochester: Probability Seminar Laguerre and Jacobi analogues of the Warren process	April 2017
34. Perimeter Institute: Mathematical Physics Seminar Affine Macdonald conjectures and special values of Felder-Varchenko functions	April 2017
33. Rutgers: Lie Group / Quantum Mathematics Seminar Affine Macdonald conjectures and special values of Felder-Varchenko functions	April 2017
32. Columbia-Princeton Probability Day Laguerre and Jacobi analogues of the Warren process	March 2017
31. ESI: Workshop on Elliptic Hypergeometric Functions Affine Macdonald conjectures and special values of Felder-Varchenko functions	March 2017
30. Columbia: Probability Seminar Laguerre and Jacobi analogues of the Warren process	November 2016
29. Columbia: Mathematical Physics Seminar Affine Macdonald conjectures and special values of Felder-Varchenko functions	October 2016
28. IESC: QIS's, CFT's, and Stochastic Processes (poster) Laguerre and Jacobi analogues of the Warren process	September 2016
27. MIT: Infinite-Dimensional Algebra Seminar Traces of intertwiners for quantum affine $\mathfrak{sl}_2$ and Felder-Varchenko functions	March 2016
26. MIT: Integrable Probability Seminar Laguerre and Jacobi analogues of the Warren process	February 2016

	25. HCM: Asymptotic Analysis in Strongly Coupled Systems (poster) Laguerre and Jacobi analogues of the Warren process	January 2016
	24. NIPS 2015 (poster) From random walks to distances on unweighted graphs	December 2015
	23. ETH Zurich: ITS Talks in Theoretical Sciences 2015 Random matrices and representation theory	November 2015
	22. UC Berkeley: RTGC Seminar Traces of intertwiners for quantum affine $\mathfrak{sl}_2$ and Felder-Varchenko functions	November 2015
	21. ETH Zurich: Mathematical Physics Seminar Traces of intertwiners for quantum affine $\mathfrak{sl}_2$ and Felder-Varchenko functions	October 2015
	20. NEU: Geometry, Physics and Representation Theory Seminar Traces of intertwiners for quantum affine $\mathfrak{sl}_2$ and Felder-Varchenko functions	October 2015
	19. Columbia: Mathematical Physics Seminar Traces of intertwiners for quantum affine $\mathfrak{sl}_2$ and Felder-Varchenko functions	October 2015
	18. Yale: Geometry, Symmetry, and Physics Seminar Traces of intertwiners for quantum affine $\mathfrak{sl}_2$ and Felder-Varchenko functions	September 2015
	17. FPSAC 2015 (poster) A representation-theoretic proof of the branching rule for Macdonald polynomials	July 2015
	16. Clay Math Inst.: Random Polymers and Algebraic Combinatorics A representation-theoretic proof of the branching rule for Macdonald polynomials	May 2015
	15. AISTATS 2015 (poster) Metric recovery from directed unweighted graphs	May 2015
	14. ICERM: Workshop on Limit Shapes (poster) A representation-theoretic proof of the branching rule for Macdonald polynomials	April 2015
	13. NIPS 2014: Workshop on Networks (poster) Metric recovery from directed unweighted graphs	December 2014
	12. UC Berkeley: GRASP Seminar A representation-theoretic proof of the branching rule for Macdonald polynomials	November 2014
	11. IHP: Workshop on Macdonald Processes and Hecke Algebras A new integral formula for Heckman-Opdam hypergeometric functions	May 2014
	10. MIT: Integrable Probability Seminar A new integral formula for Heckman-Opdam hypergeometric functions	April 2014
OUTREACH PRESENTATIONS	9. Math Olympiad Program 2018 Threshold signatures	June 2018
	8. MIT “Meta-Math” Meetup 2017 How to do a Literature Search	May 2017
	7. Summer Program in Applied Rationality and Cognition 2016 Problem Solving: Contests vs. Real Life	August 2016
	6. Math Olympiad Summer Program 2016 Distribution Testing: Is this die fair?	June 2016
	5. MIT Open House 2016 Universality: Mathematics in the real world	April 2016
	4. Math Olympiad Summer Program 2015 Fair coin flips from unfair coins	June 2015
	3. Math Olympiad Summer Program 2014 The Ising model	June 2014
	2. Math Olympiad Summer Program 2013 Random matrices	June 2013
	1. Math Olympiad Summer Program 2012 Random partitions and Fock space	June 2012

OTHER PUBLICATIONS	6. 54 <sup>th</sup> <i>International Mathematical Olympiad</i> (with J. Berman and Z. Feng), <i>Mathematics Magazine</i> <b>86</b> (2013), 309–313.	
	5. 53 <sup>rd</sup> <i>International Mathematical Olympiad</i> (with Z. Feng), <i>Mathematics Magazine</i> <b>85</b> (2012), 312–317.	
	4. 52 <sup>nd</sup> <i>International Mathematical Olympiad</i> (with Z. Feng), <i>Mathematics Magazine</i> <b>84</b> (2011), 316–319.	
	3. 51 <sup>st</sup> <i>International Mathematical Olympiad</i> (with Z. Feng and P. Loh), <i>Mathematics Magazine</i> <b>83</b> (2010), 320–323.	
	2. <i>A simulation based model of traffic circles</i> (with C. Chang and Z. Fan), <i>The UMAP Journal</i> <b>30</b> (2009), 225–244.	
	1. <i>hsolve: A difficulty metric and puzzle generator for Sudoku</i> (with C. Chang and Z. Fan), <i>The UMAP Journal</i> <b>29</b> (2008), 303–324.	
TEACHING	<b>University of Chicago</b>	2020–present
	Instructor for: Introduction to Mathematical Probability (2020, 2022), Topics in Deep Learning: Discriminative Models (2021, 2022), Statistical Theory and Methods I (2021), Distribution Theory (2022).	
	<b>Columbia University</b>	2017–2020
	Instructor for: Calculus II (2017), Graduate reading course on representation theory (2019), Calculus II (2019), Calculus II (2020).	
	<b>Cyberspace Mathematical Competition</b>	Summer 2020
	Problem Captain. Manage grading team for one of 8 problems for first year of international online math competition.	
	<b>US National Math Olympiad Summer Program</b>	Summers 2007–2018
	Instructor (2010, 2012–2018); Assistant (2007–2009). Design curriculum, give lectures, and personally coach US team to International Mathematical Olympiad.	
	<b>MIT MathROOTS</b>	Summers 2015–2016
	Academic Coordinator. Design curriculum, give lectures, and manage academic team, guest lectures, and website for first two years of outreach program in problem solving for underrepresented minority students. Covered on MIT homepage and Notices of the AMS.	
PROFESSIONAL ACTIVITIES	<b>MIT Undergraduate Research Opportunities Program</b>	Fall 2012–2015
	Mentor two undergraduate research projects, leading to published research paper.	
	• Sheela Devadas (rational Cherednik algebras in char $p \leq n$ ), 2014–2015.	
	• Ryan Yoo (characters of rational Cherednik algebras in char $p > n$ ), 2012–2014.	
	<b>Massachusetts Institute of Technology</b>	Spring 2015
	Teaching Assistant for Differential Equations. Evaluations: 6.2 (7.0)	
	<b>MIT Directed Reading Program</b>	January 2011
	Mentor reading project on representation theory of the symmetric group.	
	<b>Harvard University</b>	Spring 2009
	Course Assistant for Probability Theory. Evaluations: 4.3 (5.0)	
SERVICE	<b>Columbia Probability Seminar</b>	2016–2020
	Co-organize weekly probability seminar.	
	<b>Summer School in Probability</b>	Summer 2017
	Co-organize graduate summer school “Dyson-Schwinger equations, topological expansions, and random matrices” at Columbia.	
	<b>MIT Interacting Particle Systems Learning Seminar</b>	2012–2013
	Organize learning seminar on recent developments in interacting particle systems.	
	<b>Google Research</b>	Summer 2010
	Research intern. Research attribution and cost-sharing methods, leading to paper published in EC 2010. Mentor: Mukund Sundararajan	
SERVICE	Reviewing: Communications in Mathematical Physics, Probability Theory and Related	

Fields, Selecta Mathematica (N.S.), SIGMA, Journal of Theoretical Probability, Euro-  
physics Letters, Information and Inference, Algebraic Combinatorics, OPT 2021.

Qualifying Exam Committee: Ivan Danilenko (Columbia), Maithreya Sitaraman (Columbia)

Dissertation Committee: Qing Yan (UChicago)

Hiring Committees: Kruskal Instructor (UChicago, 2021)

LANGUAGES

Python, PyTorch, C++,  $\text{\LaTeX}$ , Magma, Mathematica / Mandarin (native), French