YUN YANG

Contact Information

Department of Mathematics
University of Maryland, College Park
College Park, MD 20742, USA

Phone: (919) 699-3620
Email: yy84@umd.edu
Website: sites.google.com/site/yunyangstat/

Education And Training

Postdoctoral Fellow August 2014 – May 2016
Department of EECS, University of California Berkeley Berkeley, California, USA

- Host: Dr. Michael Jordan and Dr. Martin Wainwright

Ph.D. in Statistics August 2011 – May 2014
Department of Statistical Science, Duke University Durham, North Carolina, USA

- Thesis Topic: Nonparametric Bayes for big data
- Advisor: Dr. David Dunson and Dr. Surya Tokdar

Bachelor of Science in Mathematics August 2007 – May 2011
Department of Mathematical Sciences, Tsinghua University Beijing, China

- Specialization: mathematics and physics (first two years)
- Thesis Topic: Brownian motion with absorbing boundaries

Professional Experience

Associate Professor July 2024 – Present
Department of Mathematics, University of Maryland, College Park College Park, Maryland, USA

Associate Professor August 2022 – May 2024
Department of Statistics, University of Illinois Urbana-Champaign Champaign, Illinois, USA

Assistant Professor August 2018 – July 2022
Department of Statistics, University of Illinois Urbana-Champaign Champaign, Illinois, USA

Assistant Professor August 2016 – May 2018
Department of Statistics, Florida State University Tallahassee, Florida, USA

Research Interests

Bayesian statistics, high-dimensional statistics, machine learning, Markov chain Monte Carlo, non-parametric statistics, optimal transport, statistical learning theory, variational inference.

Published Journal Articles

* indicates mentored PhD students

(J1) Yun Yang, Qiaochu He, Xiaolin Hu. A compact neural network for training support vector machines, *Neurocomputing*, 85: 193–198, 2012.

(J2) Yun Yang, Surya Tokdar. Minimax-optimal non-parametric regression in high dimensions, *Annals of Statistics*, 43: 652–674, 2015.

(J3) Yun Yang, David Dunson. Bayesian conditional tensor factorizations for high-dimensional classification, *Journal of the American Statistical Association*, 111: 656–669, 2016.

(J4) Yun Yang, David Dunson. Bayesian manifold regression, *Annals of Statistics*, 44: 876–905, 2016.

(J5) Yun Yang, Martin Wainwright, Michael Jordan. On the computational complexity of high-dimensional Bayesian variable selection, *Annals of Statistics*, 44: 2497–2532, 2016.

(J6) Yun Yang, Mert Pilanci, Martin Wainwright. Randomized sketches for kernels: fast and optimal non-parametric regression, *Annals of Statistics*, 45: 991–1023, 2017.

(J7) Yun Yang, Surya Tokdar. Joint estimation of quantile planes over arbitrary predictor spaces, *Journal of the American Statistical Association*, 112: 1107–1120, 2017.

(J8) Antonio Linero, Yun Yang. Bayesian regression tree ensembles that adapt to smoothness and sparsity, *Journal of the Royal Statistical Society: Series B*, 80: 1087–1110, 2018.

(J9) Roumen Varbanov, Eric Chicken, Antonio Linero, Yun Yang. A Bayesian approach to sequential monitoring of nonlinear profiles using wavelets, *Quality and Reliability Engineering International*, 35:761–775, 2019.

(J10) Michael Jordan, Jason Lee, Yun Yang. Communication-efficient distributed statistical learning, *Journal of the American Statistical Association*, 114:668–681,2019.

(J11) Anirban Bhattacharya, Debdeep Pati, Yun Yang. Bayesian fractional posteriors, *Annals of Statistics*, 48:39–66, 2019.

(J12) Qiaochu He, Yun Yang, Lingquan Bai, Baoseng Zhang, Smart energy storage management via information systems design, *Energy Economics*, 85: 104542, 2020.

(J13) Wright Shamp, Roumen Varbanov, Eric Chicken, Antonio Linero, Yun Yang. Computationally efficient Bayesian sequential function monitoring, *Journal of Quality Technology*, 54:1–19, 2020.

(J14) Yun Yang, Anirban Bhattacharya, Debdeep Pati. α-variational inference with statistical guarantees, *Annals of Statistics*, 48: 886–905, 2020.

(J15) Qiaochu He, Tiantian Nie, Yun Yang, Zuojun Shen. Beyond rebalancing: crowd-sourcing and geo-fencing for shared-mobility systems, *Production and Operations Management*, 30: 3448–3466, 2021.

(J16) Xiaohui Chen, Yun Yang. Hanson-Wright inequality in Hilbert spaces with application to K-means clustering for non-Euclidean data, *Bernoulli*, 29: 586–614, 2021.

(J17) Xiaohui Chen, Yun Yang. Diffusion K-means clustering on manifolds: Provable exact recovery via semidefinite relaxations, *Applied and Computational Harmonic Analysis*, 52: 303–347, 2021.

(J18) Xi Chen, Jason Lee, He Li, Yun Yang. Distributed estimation for principal component analysis: an enlarged eigenspace analysis, *Journal of the American Statistical Association*, 2021.

(J19) Xiaohui Chen, Yun Yang. Cutoff for exact recovery of Gaussian mixture models, *IEEE Transactions on Information Theory*, 67: 4223 – 4238, 2021.

(J20) Ke Li*, Yun Yang, Naveen Narisetty. Regret lower bound and optimal algorithm for high-dimensional contextual linear bandit, *Electronic Journal of Statistics*, 2022.

(J21) Meimei Liu, Zuofeng Shang, Yun Yang, Guang Cheng. Nonparametric testing under randomized sketching, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44:4280 – 4290, 2022.

(J22) Peng Zhao*, Qiaoche He, Yun Yang. High-dimensional linear regression via implicit regularization, *Biometrika*, 109: 1033 — 1046, 2022.

(J23) Rong Tang*, Yun Yang. Bayesian inference for risk minimization via exponentially tilted empirical likelihood, *Journal of the Royal Statistical Society: Series B*, 84: 1257 – 1286.2022.

(J24) Shishuang He*, Yinyin Chen*, Yun Yang, Feng Liang. Learning topic models: identifiability and finite-sample analysis, *Journal of the American Statistical Association*, 1 – 16, 2022.

(J25) Weilong Zhao, Zishen Xu, Yue Mu, Yun Yang, Wei Wu. Model-based statistical depth with applications to functional data, *Journal of Nonparametric Statistics*, 1 – 44, 2023.

(J26) Yifan Chen*, Tianning Xu, Dilek Hakkani-Tur, Di Jin, Yun Yang, Ruoqing Zhu. Calibrate and debias layer-wise sampling for graph convolutional networks, *Transactions on Machine Learning Research*, 2023.

(J27) Shuang Zhou*, Debdeep Pati, Tianying Wang, Yun Yang, Raymond J. Carroll. Gaussian processes with errors in variables: theory and computation, *Journal of Machine Learning Research*, 24: 1 – 53, 2023.

(J28) Teng Wu*, Naveen N. Narisetty, Yun Yang. Statistical inference via conditional Bayesian posteriors in high-dimensional linear regression, *Electronic Journal of Statistics*, 17: 769 – 797, 2023.

(J29) Rong Tang*, Yun Yang. Minimax rate of distribution estimation on unknown submanifold under adversarial losses, *Annals of Statistics*, 51: 1282 – 1308, 2023.

(J30) *Yangfan Zhang, Yun Yang. Bayesian model selection via mean-field variational approximation, *Journal of the Royal Statistical Society: Series B*, 2024.

(J31) *Rong Tang, Yun Yang. On the computational complexity of Metropolis-adjusted Langevin algorithms for Bayesian posterior sampling,*Journal of Machine Learning Research*, 2024.

PEER REVIEWED CONFERENCE PUBLICATIONS

(C1) Bruno Cornelis, Yun Yang, Joshua Vogelstein, Ann Dooms, Ingrid Daubechies, David Dunson. Bayesian crack detection in ultra high resolution multimodal images of paintings, *18th International Conference on Digital Signal Processing (ICDSP)*, 2013.

(C2) Roumen Varbanov, Eric Chicken, Antonio Linero, Yun Yang. Wavelet-based Bayesian profile monitoring, *Industrial and Systems Engineering Research Conference (ISERC)*, 2017.

(C3) Debdeep Pati, Anirban Bhattacharya, Yun Yang. On the Statistical Optimality of Variational Bayes, *Artificial Intelligence and Statistics Conference (AISTATS)*, 2018.

(C4) Yun Yang, Zuofeng Shang, Guang Cheng. Non-asymptotic analysis for nonparametric testing, *Conference on Learning Theory (COLT)*, 2020.

(C5) Yifan Chen*, Yun Yang. Fast statistical leverage score approximation in kernel ridge regression, *Artificial Intelligence and Statistics Conference (AISTATS)*, 2021.

(C6) Yifan Chen*, Yun Yang. Accumulations of projections—a unified framework for random sketches in kernel ridge regression, *Artificial Intelligence and Statistics Conference (AISTATS)*, 2021.

(C7) Rong Tang*, Yun Yang. On empirical Bayes variational autoencoder: an excess risk bound, *Conference on Learning Theory (COLT)*, 2021.

(C8) Yifan Chen*, Qi Zeng, Heng Ji, Yun Yang. Skyformer: remodel self-attention with Gaussian kernel and Nyström method, *Conference on Neural Information Processing Systems (NeurIPS)*, 2021.

(C9) Yifan Chen*, Qi Zeng, Dilek Hakkani-Tur, Di Jin, Heng Ji, Yun Yang. Sketching as a tool for understanding and accelerating self-attention for long sequences, *North American Chapter of the Association for Computational Linguistics (NAACL)*, 2022.

(C10) Honggang Wang, Anirban Bhattacharya, Debdeep Pati, Yun Yang. Structured variational inference in Bayesian state-space models, *Artificial Intelligence and Statistics Conference (AISTATS)*, 2022.

(C11) Yubo Zhuang*, Xiaohui Chen, Yun Yang. Sketch-and-lift: scalable subsampled semidefinite program for K-means clustering, *Artificial Intelligence and Statistics Conference (AISTATS)*, 2022.

(C12) Rentian Yao*, Xiaohui Chen, Yun Yang. Mean-field nonparametric estimation of interacting particle systems, *Conference on Learning Theory (COLT)*, 2022.

(C13) Yubo Zhuang*, Xiaohui Chen, Yun Yang. Wasserstein K-means for clustering probability distributions, *Conference on Neural Information Processing Systems (NeurIPS)*, 2022.

(C14) Rong Tang*, Yun Yang. Minimax nonparametric two-sample test under adversarial losses, *Artificial Intelligence and Statistics Conference (AISTATS)*, 2023.

(C15) Yifan Chen*, Rentian Yao*, Yun Yang, Jie Chen. A Gromov–Wasserstein geometric view of spectrum-preserving graph coarsening, *International Conference on Machine Learning (ICML)*, 2023.

(C16) Yubo Zhuang*, Xiaohui Chen, Yun Yang. Likelihood adjusted semidefinite programs for clustering heterogeneous data, *International Conference on Machine Learning (ICML)*, 2023.

(C17) *Rong Tang, Yun Yang. Adaptivity of diffusion models to manifold structures, *Artificial Intelligence and Statistics Conference (AISTATS)*, 2024.

(C18) *Rentian Yao, *Linjun Huang, Yun Yang. Minimizing convex functionals over space of probability measures via KL divergence gradient flow, *Artificial Intelligence and Statistics Conference (AISTATS)*, 2024.

(C19) *Yubo Zhuang, Xiaohui Chen, Yun Yang, Richard Y Zhang. Statistically optimal K-means clustering via nonnegative low-rank semidefinite programming, *International Conference on Learning Representations (ICLR)*, Selected for Oral, 2024.

SUBMITTED MANUSCRIPTS AND TECHNICAL REPORTS

(T1) Yun Yang, David Dunson. Sequential Markov chain Monte Carlo, arXiv:1308.3861.

(T2) Yun Yang, David Dunson. Minimax optimal Bayesian aggregation, arXiv:1403.1345.

(T3) Yun Yang, Guang Cheng, David Dunson. Semiparametric Bernstein-von Mises theorem: second order studies, arXiv:1503.04493.

(T4) Yun Yang, Debdeep Pati. Bayesian model selection consistency and oracle inequality with intractable marginal likelihood, arXiv:1701.00311.

(T5) Yun Yang. Statistical inference for high dimensional regression via Constrained Lasso, arXiv:1704.05098.

(T6) Yun Yang, Anirban Bhattacharya, Debdeep Pati. Frequentist coverage and sup-norm convergence rate in Gaussian process regression, arXiv:1708.04753.

(T7) *Wei Han, Yun Yang. Statistical inference in mean-field variational Bayes, arXiv:1911.01525.

(T8) *Yangfan Zhang, Yun Yang. Efficient inference for stochastic gradient descent.

(T9) *Ke Li, Yun Yang, Naveen Narisetty. High-dimensional linear bandits with variable selection.

(T10) *Rentian Yao, Yun Yang. Mean-field variational inference via Wasserstein gradient flow, arXiv:2207.08074.

(T11) *Rong Tang, Yun Yang. Estimating distributions with low-dimensional structures using mixtures of generative models, arXiv:2301.00890.

(T12) Anirban Bhattacharya, Debdeep Pati, Yun Yang. On the convergence of coordinate ascent variational inference, arXiv:2306.01122.

(T13) *Yifan Chen, Yun Yang. Accumulative sub-sampling sketching for approximate matrix multiplication.

(T14) *Rentian Yao, Xiaohui Chen, Yun Yang. Wasserstein proximal coordinate gradient algorithms.

(T15) *Rong Tang, Anirban Bhattacharya, Debdeep Pati, Yun Yang. Robust Bayesian inference on Riemannian submanifold.

(T16) Meimei Liu, Zuofeng Shang, Yun Yang. Scalable statistical inference in non-parametric least squares, arXiv:2310.00881.

GRANTS AND AWARDS

- **Grant:** Collaborative research: theoretical and algorithmic foundations of variational Bayesian inference (PI), funded by the *National Science Foundation* (NSF-DMS # 2210717). Total amount: $134,186, June 2022 – May 2025.
- **Grant:** Fast and robust Gaussian process inference for Bayesian nonparametric learning (PI), funded by the *National Science Foundation* (NSF-DMS # 1810831). Total amount: $120,000, June 2018 – May 2021.
- **Grant:** Bayesian inference via fractional posteriors (PI), funded by *Florida State University* (FYAP). Total amount: $20,000, June 2017 – August 2017.

- **University of Illinois List of Teachers Ranked as Excellent by Their Students**: Fall 2021 , Spring 2023.
- **PhD Fellowship**, Department of Statistical Science, Duke University, 2011.
- **Tsinghua First-Class Scholarship**, 2007 to 2011.
- **Gold Medal and the Best Score Special Award**, 8th Asian Physics Olympiad (APhO), 2007.

INVITED TALKS

- Bayesian inference for risk minimization via exponentially tilted empirical likelihood, *EAC ISBA Conference*, Qingdao, China, June 2023.
- Learning topic models: identifiability and finite-sample analysis, *ICSA Applied Statistics Symposium*, Ann Arbor, Michigan, June 2023.
- Implicit estimation of high-dimensional distributions using generative models, *ACMS Colloquium*, University of Notre Dame, April 2023.
- Implicit estimation of high-dimensional distributions using generative models, *Seminar talk*, Georgia Institute of Technology, March 2023.
- Implicit estimation of high-dimensional distributions using generative models, *Seminar talk*, University of Wisconsin Madison, February 2023.
- Implicit estimation of high-dimensional distributions using generative models, *Seminar talk*, Columbia University, February 2023.
- Mean-field variational inference via Wasserstein gradient flow, *Conference on Advances in Data Science,* College Station, Texas, October, 2022.
- Bayesian inference for risk minimization via exponentially tilted empirical likelihood, *ICSA conference (virtual)*, Xi'an, China, July, 2022.
- Fast and Accurate Computation for Large-Scale Kernel Ridge Regression, *International Indian Statistical Association Conference,* virtual conference, May 2021.
- Approximate Bayesian Computation via Variational Approximation, *Seminar talk (virtual)*, Department of Decision Sciences, Bocconi University, Milan, Italy, May 2021.
- Non-asymptotic analysis for nonparametric testing, *Conference on Learning Theory (COLT)*, virtual conference, July 2020.
- Using equivalent kernel to understand kernel ridge regression, *BayesComp*, Gainesville, Florida, Jan 2020.
- Fast and optimal Bayesian inference via variational approximations, *Seminar talk*, Department of Statistics and Actuarial Science, University of Iowa, Iowa City, USA, November 2019.
- Smoothness and sparsity adaptive Bayesian tree ensemble method for high-dimensional nonparametric regression, *ICSA conference*, Nankai, China, July 2019.
- Approximate Bayesian inference via variational approximation, *Conference on Mixture/Non-regular Models*, Guilin, China, August 2018.

- Fast and optimal Bayesian inference via variational approximations, *SLDS conference*, Columbia University, New York, USA, June 2018.
- Communication-efficient distributed statistical inference, *ICSA conference*, Rutgers University, New Jersey, USA, June 2018.
- Smoothness and sparsity adaptive Bayesian tree ensemble method for high-dimensional nonparametric regression, *ENAR conference*, Atlanta, GA, March 2018.
- Frequentist coverage and sup-norm convergence rate in Gaussian process regression, *CMstatistics conference*, Senate House, University of London, UK, December 2017.
- Computationally efficient high-dimensional variable selection via Bayesian procedures, *Seminar talk*, Department of Statistics, Texas A&M University, Texas, USA, November 2017.
- Bayesian model selection consistency and oracle inequality with intractable marginal likelihood, *BNP conference*, Universiteé Paris Dauphine, Paris, France, June 2017.
- Computationally efficient high-dimensional variable selection via Bayesian procedures, *Seminar talk*, Department of Statistics, University of Florida, Florida, USA, March 2017.
- Computationally efficient high-dimensional variable selection via Bayesian procedures, *Seminar talk*, Department of Statistics, Purdue University, Indiana, USA, November, 2016.
- Computationally efficient high-dimensional variable selection via Bayesian procedures, *Seminar talk*, Department of Statistics, Florida State University, Florida, USA, January 2016.
- Computationally efficient high-dimensional variable selection via Bayesian procedures, *Seminar talk*, Department of Statistics, University of Illinois Urbana-Champaign, Illinois, USA, January 2016.
- Computationally efficient high-dimensional variable selection via Bayesian procedures, *Seminar talk*, Department of Statistics, Rutgers University, New Jersey, USA, January 2016.
- Computationally efficient high-dimensional variable selection via Bayesian procedures, *Seminar talk*, Department of Statistics, University of California, Davis, California, USA, January 2016.
- Bayesian conditional tensor factorizations for high-dimensional classification, *ERCIM conference*, Oviedo, Spain, December 2012.

PROFESSIONAL EXPERIENCE

- Editorial service:
  - Associate Editor, *Journal of Computational and Graphical Statistics*, 2023 – present.
- Refereeing for Journals: Annals of Statistics, Journal of the American Statistical Association, Journal of the Royal Statistical Society, Biometrika, Journal of Multivariate Analysis, Electronic Journal of Statistics, Journal of Nonparametric Statistics, Statistica Sinica, Journal of Statistical Planning and Inference, Journal of Machine Learning Research, IEEE Transactions on Signal Processing, Journal of Selected Topics in Signal Processing, IEEE Transactions on Network Science and Engineering, Indian Journal of Statistics.

- Regular reviewers for conferences: NeurIPS, 2016 – 2020, ICML 2016 – 2020, AISTATS 2016 – 2020.
- Area chair for AISTATS 2023 – 2024.
- Judge for ASA student paper competition in Section on Bayesian Statistical Science (SBSS), 2019 – 2021.
- NESS Student Research Award committee, 2022 – 2023.
- Savage award committee member, 2022.
- Ad hoc reviewer for Economics Program of National Science Foundation (NSF), 2019.
- Chair of session "Recent development of high-dimensional modeling, inference and computation" in CMStatistics 2017.
- Departmental committee:
  - Colloquium committee, 2018 – 2019
  - MS program committee, 2018 – 2019
  - Undergraduate research committee, 2018 – 2019
  - PhD qualifying exam committee, 2019 – Present
  - PhD program committee, 2019 – Present
  - Tenure track faculty search committee, 2018 – 2019, 2022 – Present
- Membership in Professional Organizations: International Society for Bayesian Analysis, Institute of Mathematical Statistics.

TEACHING EXPERIENCE

Instructor, University of Illinois Urbana-Champaign

- STAT 408: Actuarial Statistics I, Spring 2023
- STAT 410: Statistics and probability II, Spring 2022, Spring 2021, Spring 2019
- STAT 424: Analysis of variance, Spring 2019
- STAT 510: Mathematical Statistics I, Fall 2019, Spring 2020, Fall 2020
- STAT 511: Advanced Mathematical Statistics, Fall 2021, Fall 2022, Fall 2023
- STAT 578: Modern Statistical Inference, Spring 2023

Instructor, Florida State University

- STA4321/5323: Introduction to mathematical statistics I, Fall 2016, Spring 2017, Fall 2017
- STA6448: Advanced probability and inference II, Spring 2018

Teaching Assistant, Duke University

- STA104: Probability, Spring 2012
- STA961: Statistical stochastic processes, Spring 2014

PhD Student Advising

- Chenyang Wang, University of Illinois Urbana-Champaign (Role: Doctoral thesis advisor, 2023 – Present)
- Linjun Huang, University of Illinois Urbana-Champaign (Role: Doctoral thesis advisor, 2022 – Present)
- Yubo Zhuang, University of Illinois Urbana-Champaign (Role: Doctoral thesis co-advisor with Dr. Xiaohui Chen, 2021 – Present)
- Rentian Yao, University of Illinois Urbana-Champaign (Role: Doctoral thesis co-advisor with Dr. Xiaohui Chen, 2020 – Present)
- Shishuang He, University of Illinois Urbana-Champaign (Role: Doctoral thesis co-advisor with Dr. Feng Liang, 2020 – Present),
- Rong Tang, University of Illinois Urbana-Champaign (Role: Doctoral thesis advisor, 2019 – 2023), now Assistant Professor in Mathematics at Hong Kong University of Science and Technology
- Yifan Chen, University of Illinois Urbana-Champaign (Role: Doctoral thesis advisor, 2019 – 2023), now Assistant Professor in Computer Science at Hong Kong Baptist University
- Yangfan Zhang, University of Illinois Urbana-Champaign (Role: Doctoral thesis co-advisor with Dr. Xiaofeng Shao, 2020 – 2022), now at Two-Sigma
- Ke Li, University of Illinois Urbana-Champaign (Role: Doctoral thesis co-advisor with Dr. Naveen Narisetty, 2019 – 2022), now at Facebook
- Wei Han, University of Illinois Urbana-Champaign (Role: Doctoral thesis advisor, 2018 – 2022)
- Peng Zhao, Florida State University (Role: Doctoral thesis advisor, 2017 – 2018), now Assistant Professor in Applied Economics and Statistics at University of Delaware
- Shuang Zhou, Florida State University (Role: Doctoral thesis advisor, 2017 – 2018), now Assistant Professor in Statistics at Arizona State University

Doctoral Dissertation committee

- Abhishek Ojha, University of Illinois Urbana-Champaign (2022 – Present)
- Zihe Liu, University of Illinois Urbana-Champaign (2021 – Present)
- Tianning Xu, University of Illinois Urbana-Champaign (2021 – Present)
- Teng Wu, University of Illinois Urbana-Champaign (2020 – 2022)
- Yan Liu, University of Illinois Urbana-Champaign (2019 – 2021)
- Yubai Yuan, University of Illinois Urbana-Champaign (2018 – 2020)
- Yujia Deng, University of Illinois Urbana-Champaign (2018 – 2020)
- Yinyin Chen, University of Illinois Urbana-Champaign (2018 – 2020)
- Roumen Varbanov, Florida State University (2017 – 2018)
- Lizhe Sun, Florida State University (2017 – 2018)
- Libo Wang, Florida State University (2015 – 2017)

TECHNICAL SKILLS

- Computer programming: C/C++, Python.
- Statistical software: R, Matlab.
- Applications: Latex, Microsoft Office.