

Qifan Song

Department of Statistics

Purdue University

250 N. University Street

West Lafayette, IN 47907-2066

Phone: 765-494-4058

Office: MATH 208

Email: qfsong@purdue.edu

URL: www.stat.purdue.edu/~qfsong

Employment

ASSOCIATE PROFESSOR, DEPARTMENT OF STATISTICS, PURDUE UNIVERSITY, 2021 - PRESENT

ASSISTANT PROFESSOR, DEPARTMENT OF STATISTICS, PURDUE UNIVERSITY, 2014 - 2021

GRADUATE FELLOW, IAMCS, TEXAS A&M UNIVERSITY, 2009-2012.

Education

PH.D. IN STATISTICS, TEXAS A&M UNIVERSITY, AUGUST 2014.

Advisor: Dr. Faming Liang

B.S. IN STATISTICS, PEKING UNIVERSITY, BEIJING, CHINA, MAY 2009.

Teaching Experience

INSTRUCTOR, STAT 526 (ADVANCED STAT. METHODS): Purdue University

INSTRUCTOR, STAT 417 (STATISTICAL THEORY): Purdue University

INSTRUCTOR, STAT 525 (INTERMEDIATE STAT. METHODS): Purdue University

INSTRUCTOR, STAT 516 (BASIC PROB. AND APPLICATIONS): Purdue University.

INSTRUCTOR, STAT 517 (STATISTICAL INFERENCE): Purdue University.

INSTRUCTOR, STAT 514 (DESIGN OF EXPERIMENT): Purdue University.

INSTRUCTOR, STAT 101 (FRESHMEN ORIENTATION SEMINAR): Purdue University

INSTRUCTOR, STAT 302 (INTRO. BIOSTATISTICS): Texas A&M University.

LAB INSTRUCTOR, STAT 302 (INTRO. BIOSTATISTICS): Texas A&M University, Fall 2011–Spring 2012.

Department and University Service

STATISTICS UNDERGRADUATE PROGRAM CHAIR: Fall 2021 - Present

UNIVERSITY FACULTY SENATE: Fall 2018 - Spring 2024

Research Interests

High dimensional data analysis: Bayesian variable selection, Bayesian shrinkage, penalized estimator, false discovery control, Bernstein von-Mises theorem.

Bayesian inferences: variational Bayes inferences, Bayesian semiparametric modeling.

Big Data analysis: parallel Bayesian computing, divide-and-combine approaches.

Bayesian Computation: ergodicity analysis for Monte Carlo Markov chain (MCMC), Adaptive MCMC algorithm, Stochastic gradient MCMC.

Modern Data Science: Bayesian sparse DNN modeling, adversarial learning, large language models, generative models.

Research Publications

Liang, J., Zhang, Q., Deng, W., **Song, Q.**, and Lin, G. (2025+), Bayesian Federated Learning with Hamiltonian Monte Carlo: Algorithm and Theory, accepted by *Journal of Computational and Graphical Statistics*

Kim, S., **Song, Q.** and Liang, F (2025). A New Paradigm of Generative Adversarial Networks based on Randomized Decision Rules, *Statistica Sinica*, **35**, 1-22

- Li, W., Li, H., Honorio, J. and **Song, Q.**, (2024) PyXAB - A Python Library for X-Armed Bandit and Online Blackbox Optimization Algorithms, *Journal of Open Source Software*, doi.org/10.21105/joss.06507
- Lee, H., **Song, Q.** and Honorio, J. (2024), Support Recovery in Sparse PCA with General Missing Data (2024), *Proceedings of the Fortieth Conference on Uncertainty in Artificial Intelligence*, **PMLR 244**:2163-2187,
- Wang, Z., Chen, Y., **Song, Q.**, Zhang, R. (2024), Enhancing Low-Precision Sampling via Stochastic Gradient Hamiltonian Monte Carlo, *Transactions on Machine Learning Research* (04/24)
- Zhang, P., **Song, Q.** and Liang, F.(2024), A Langevinized Ensemble Kalman Filter for Large-Scale Dynamic Learning, *Statistica Sinica* **34**, 1071-1091
- Sohn, J., **Song, Q.** and Lin, G. (2024), Fair Supervised Learning with A Simple Random Sampler of Sensitive Attributes, *Proceedings of The 27th International Conference on Artificial Intelligence and Statistics*, PMLR **238**:1594-1602
- Xing, Y., Lin, Xi., **Song, Q.**, Xu, Y., Zeng, B. and Cheng, G. (2024), Better Representations via Adversarial Training in Pre-Training: A Theoretical Perspective, *Proceedings of The 27th International Conference on Artificial Intelligence and Statistics*, PMLR **238**:199-207
- Li, W., **Song, Q.** and Honorio, J. (2024), Personalized Federated X-armed Bandit, *Proceedings of The 27th International Conference on Artificial Intelligence and Statistics*, PMLR **238**:37-45
- Haldar, R. and Xing, Y. and **Song, Q.** (2024), Effect of Ambient-Intrinsic Dimension Gap on Adversarial Vulnerability, *Proceedings of The 27th International Conference on Artificial Intelligence and Statistics*, PMLR **238**:1090-1098.
- Li, W., **Song, Q.**, Honorio, J. and Lin, G.(2024), Federated X-Armed Bandit, *Proceedings of the AAAI Conference on Artificial Intelligence*, **38**(12), 13628-13636. <https://doi.org/10.1609/aaai.v38i12.29267>
- Li, W., Wang, C.-H., Cheng, G. and **Song, Q.** (2023) Optimum-statistical Collaboration Towards General and Efficient Black-box Optimization, *Transactions on Machine Learning Research* (05/2023)
- Song, Q.** and Liang, F.(2023), Nearly optimal Bayesian Shrinkage for High Dimensional Regression, *Science China Mathematics* **66**:409-442.
- Song, Q.** and Cheng, G. (2023), Optimal False Discovery Control of Minimax Estimator, *Bernoulli* **29**: 1959-1982.
- Chen, Y., **Song, Q.**, and Su, J.(2023), On a Class of Multivariate Mixtures of Gamma Distributions: Actuarial Applications and Estimation Via Stochastic Gradient Methods, *Variance* **16**(1).
- Lee, H., **Song, Q.**, and Honorio, J.(2022), Support Recovery in Sparse PCA with Incomplete Data, accepted by *Advances in Neural Information Processing Systems* **35**: 27321-27332.
- Xing, Y., **Song, Q.**, and Cheng, G.(2022), Why Do Artificially Generated Data Help Adversarial Robustness, accepted by *Advances in Neural Information Processing Systems* **35**: 954-966..
- Xing, Y., **Song, Q.**, and Cheng, G.(2022), Phase Transition from Clean Training to Adversarial Training, accepted by *Advances in Neural Information Processing Systems* **35**: 9330-9343.
- Chen, Z., Wang, Z., **Song, Q.**, and Xie, J. (2022), Data-guided Treatment Recommendation with Feature Scores, *Statistica Sinica* **32**:2497-2519.
- Xing, Y., **Song, Q.**, and Cheng, G.(2022), Benefit of Interpolation in Nearest Neighbor Algorithms, *SIAM Journal on Mathematics of Data Science*, **4**:935-956.

- Xing, Y., **Song, Q.**, and Cheng, G.(2022), Unlabelled Data Helps: Minimax Analysis and Adversarial Robustness, *Proceedings of The 25th International Conference on Artificial Intelligence and Statistics*, PMLR **151**:136-168.
- Sun, Y., **Song, Q.** and Liang, F.(2022), Consistent Sparse Deep Learning: Theory and Computation, *J. Amer. Statist. Assoc.*, **117**:1981-1995.
- Sun, Y., **Song, Q.** and Liang, F.(2022), Learning Sparse Deep Neural Networks with Spike-and-Slab Priors, *Statistics and Probability Letters*, **180**, doi.org:10.1016/j.spl.2021.109246.
- Kim, S., **Song, Q.** and Liang, F.(2022) Stochastic Gradient Langevin Dynamics with Adaptive Drifts, *Journal of Statistical Computation and Simulation*, **93**: 318-336.
- Pae, B. J., Hirai, D., Monroe, J., Motaganahalli, R., **Song, Q.**, and Roseguini, B. (2022), Assessment of Skeletal Muscle Oxygenation during Exercise in Patients with Peripheral Artery Disease Using Near Infrared Spectroscopy: Impact of Skin Color and Adipose Tissue Thickness. *The FASEB Journal*, **36**, doi.org:10.1096/fasebj.2022.36.S1.R5392.
- Ro, B., Deren, A., **Song, Q.**, Kim, K.H., Kuang, S., Roseguini, B. (2022) Impact of Heat Therapy on Skeletal Muscle Mass and Function in A Mouse Model of Duchenne Muscular Dystrophy, *The FASEB Journal*, **36**, doi.org/10.1096/fasebj.2022.36.S1.R5393.
- Xing, Y., **Song, Q.**, and Cheng, G.(2021) On the Algorithmic Stability of Adversarial Training, *Advances in Neural Information Processing Systems* **34**:26523–26535.
- Kim, K., Ro, B., Damen, F., Gramling, D., Lehr, T., **Song, Q.**, Goergen, C. and Roseguini, B.(2021), Heat therapy improves body composition and muscle function, but does not affect capillary or collateral growth in a model of obesity and hindlimb ischemia, *Journal of Applied Physiology*, **130**(2): 355-368.
- Kim, K., Ro, B., Damen, F., Gramling, D., Lehr, T., **Song, Q.** and Roseguini, B.(2021), Neither peristaltic pulse dynamic compressions nor heat therapy accelerate glycogen resynthesis following shuttle running, *Medicine & Science in Sports & Exercise*, **53**(11): 2425-2435.
- Li, H., **Song, Q.** and Su, J.(2021) Robust Estimates of Insurance Misrepresentation through Kernel Quantile Regression Mixtures, *Journal of Risk and Insurance*, **88**: 625-663.
- Xing, Y., **Song, Q.**, and Cheng, G.(2021) Predictive Power of Nearest Neighbors Algorithm under Random Perturbation, *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics*, PMLR **130**: 496-504.
- Xing, Y., **Song, Q.**, and Cheng, G.(2021) On the Generalization Properties of Adversarial Training, *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics*, PMLR **130**: 505-513.
- Monroe, J., **Song, Q.**, Emery, M., Hirai, D. Motaganahalli, R., Roseguini, B. (2021), Acute effects of leg heat therapy on walking performance and cardiovascular and inflammatory responses to exercise in patients with peripheral artery disease, *Physiological Reports*, **8**:e14650
- Bai, J., **Song, Q.**, and Cheng, G.(2020) Efficient Variational Inference for Sparse Deep Learning with Theoretical Guarantee, *Advances in Neural Information Processing Systems* **33**:466-476.
- Song, Q.**(2020) Bayesian Shrinkage towards Sharp Minimaxity, *Electronic Journal of Statistics*, **14**(2): 2714-2741;
- Ro, B., Deren, A., **Song, Q.**, Kim, K.H., Kuang, S., Roseguini, B. (2020) Impact of Heat Therapy on Skeletal Muscle Mass and Function in A Mouse Model of Duchenne Muscular Dystrophy, *FASEB Journal*, **36**, <https://doi.org/10.1096/fasebj.2022.36.S1.R5393>
- Kim, K., Reid, B., Casey, C., Bender, B., Ro, B., **Song, Q.**, Trewin, A., Petersen, A., Kuang, S., Gavin, T., and Roseguini, B. (2020) Effects of repeated local heat therapy on skeletal muscle structure and function in humans, *Journal of Applied Physiology* **128**(3): 483-492.

- Song, Q.**, Sun, Y., Ye, M., Liang, F. (2020) Extended Stochastic Gradient MCMC for Large-Scale Bayesian Variable Selection, *Biometrika*, **107**(4): 997-1004,
- Song, Q.**, Cheng, Guang. (2019) Bayesian Fusion Estimation via t Shrinkage, *Sankhya series A*, doi.org: 10.1007/s13171-019-00177-0
- Shi, R., Liang, F., **Song, Q.**, Luo, Y. and Ghosh, M. (2019) A Blockwise Consistency Method for Parameter Estimation of Complex Models, *Sankhya series B*, doi.org: 10.1007/s13571-018-0183-0.
- Kim, K., Reid, B., Ro, B., Casey, C., **Song, Q.**, Kuang, S. and Roseguini, B. (2019) Heat therapy improves soleus muscle force in a model of ischemia-induced muscle damage, *Journal of Applied Physiology* **127**(1):215-228.
- Kim, K., Kuang, S., **Song, Q.**, Gavin, T., and Roseguini, B.(2019) Impact of heat therapy on recovery after eccentric exercise in humans, *Journal of Applied Physiology* **126**(4):965-976.
- Song, Q.**, (2018) An overview of Reciporical Lasso for high dimensional data analysis, WIREs Computational statistics, **10** doi: 10.1002/wics.1416.
- Kim, K., Reid, B., Ro, B., Hester, B., Caesy, C. **Song, Q.** and Roseguini, B. (2018) Impact of heat therapy on skeletal muscle structure and function in a mouse model of peripheral arterial disease, *The FASEB Journal*, **32**, 853.12
- Kuhlenhoelter, A., Kim, K., Neff, D., Nie, Y., Blaize, A.N., Wong, B., Kuang, S., Stout, J., **Song, Q.**, Gavin, T. and Roseguini, B. (2016) Heat therapy promotes the expression of angiogenic regulators in human skeletal muscle. *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology*, **311**, R377-R391
- Liang, F., Kim, J. and **Song, Q.** (2016) A Bootstrap Metropolis-Hastings Algorithm for Bayesian Analysis of Big Data. *Technometrics*, **58**, 304-318.
- Liang, F., Jin, I-H., **Song, Q.**, and Liu, J.S. (2016) An Adaptive Exchange Algorithm for Sampling from Distribution with Intractable Normalized Constants. *J. Amer. Statist. Assoc.*, **111**, 377-393.
- Liang, F., **Song, Q.** and Qiu, P.(2015) An Equivanlent measure of Partial Correlation Coefficients for High Dimensional Gaussian Graphic Models. *J. Amer. Statist. Assoc.*, **110**, 1248-1265.
- Song, Q.** and Liang, F.(2015) High Dimensional Variable Selection with Reciprocal L_1 -Regularization. *J. Amer. Statist. Assoc.*, **110**, 1607-1620.
- Song, Q.** and Liang, F.(2015) A split-and-Merge Bayesian Variable Selection Approach for Ultra-high dimensional Regression. *J. R. Statist. Soc. B*, **77**, 947-972 .
- Song, Q.**, Wu, M. and Liang, F.(2014) Weak Convergence Rates of Population versus Single-Chain Stochastic Approximation MCMC algorithms. *Adv. Appl. Prob.*, **46**, 1059-1083.
- Liang, F., **Song, Q.**, and Yu. K. (2013) Bayesian Subset Modeling for High Dimensional Generalized Linear Models. *J. Amer. Statist. Assoc.*, **108**, 589-606.
- Liang, F., Cheng, Y., **Song, Q.**, Park, J., and Yang, P. (2013) A Resampling-Based Stochastic Approximation Method for Analysis of Large Geostatistical Data. *J. Amer. Statist. Assoc.*, **108**, 325-339.

Papers submitted or under review

- Sohn, J., **Song, Q.** and Lin, G. Task-tailored Pre-processing: Fair Downstream Supervised Learning, submitted to *IEEE Transactions on Pattern Analysis and Machine Intelligence*
- Ro, B, Castanhas, L., Crenshaw, F., Janczy, A., Swift, P. **Song, Q.**, Zhou, X., Parker, J. and Roseguini, B. Feasibility of unsupervised, home-based leg heat therapy in older adults: a pilot, sham-controlled trial, submitted to *Journal of Applied Physiology*

Haldar, R., Wang, Z., **Song, Q.**, Lin, G., and Xing, Y., LLM Safety Alignment is Divergence Estimation in Disguise, submitted to *ICML 2025*

Haldar, R. and **Song, Q.**, On Neural Network approximation of ideal adversarial attack and convergence of adversarial training, under revision *SIAM Journal on Mathematics of Data Science*

Xu, C., Cheng, Y., **Song, Q.**, Zhao, Y., Wang, X., Bayesian variable selection based on empirical likelihood for ultra-high dimensional data, submitted to *The Annals of Applied Statistics*

Xing, Y., Lin, X., Suh, N., **Song, Q.**, and Cheng, G., Benefits of Transformer: In-Context Learning in Linear Regression Tasks with Unstructured Data, submitted to *ICML 2025*

Haldar, R., Xing, Y., **Song, Q.** and Lin, G., Adversarial Vulnerability as a Consequence of On-Manifold Inseparability, submitted to *ICML 2025*

Sohn, J. and **Song, Q.**, Parallely Tempered Generative Adversarial Networks, under revision for *J. Amer. Statist. Assoc.*

Li, W., Honorio, J. and **Song, Q.**, On the Sharp Analysis of Online Blackbox Optimization, submitted

Lee, H., Mazumder, R., **Song, Q.** and Honorio, J., Matrix Completion from General Deterministic Sampling Patterns, submitted

Liang, J., Lin, G. and **Song, Q.**, Bridging Differential Privacy to Consistent Applications: A Tight Characterization of the Worst-Case Risks under Knowledgeable Attacks, submitted

Students

Graduated PhD students (5): Jincheng Bai, Yue Xing, Jiajun Liang, Wenjie Li, Hanbyul Lee

Current PhD students (7): Ziyi Wang, Jinwon Sohn, Rajdeep Haldar, Heekyung Ahn, Shanyun Gao, Kyle Conrad, Qin Shi

Past postdoc researcher (1): Ashwini Maurya

Funds

JMHC Breakthrough Research Award (Purdue Honors College), *Home-based Heat Therapy in Older Adults with Type II Diabetes*, \$52,411.5, 01/2025-12/2025, Role: co-PI

Indiana Center for Diabetes and Metabolic Diseases (CDMD) Pilot and Feasibility (P&F) program *Home-based Heat Therapy in Older Adults with Type II Diabetes*, \$50,000, 09/2024-08/2025, Role: co-PI

SOA 2021 Individual Grants Competition *Interpretable Random Forests for Risk-informed Modeling in Healthcare Management*, \$13,000, 07/01/2021-06/30/2022, Role: co-PI

CAS-20278375 (Casualty Actuarial Society) *Multivariate Mixtures of Gamma Distributions: A Joyful Toolkit for Modeling Dependent Insurance Data*, 07/01/2020-06/30/2021, \$14,000, Role: co-PI

NSF-DMS-1811812 (National Science Foundation), *High Dimensional Semiparametric Estimation and Inferences*, 08/01/2018-07/31/2021, \$159,985, Role: PI.

DARPA-BAA-16-4 (DARPA), *Diagnostic Methods & Software for Fitted-Model Checking*, 08/12/2016-08/12/2020, \$800,000, Role: Co-PI.

AHA-16SDG27600003 (American Heart Association), *Thermotherapy for Intermittent Claudication*, 01/01/2016-12/31/2019, \$308,000, Role: co-investigator.

CTSI-UL1TR001108 (Indiana Clinical Translational Sciences Institute), *Heat therapy to reduce leg pain and improve walking tolerance in patients with symptomatic peripheral artery disease*, 9/1/2017 - 8/31/2019, \$75,000, Role: co-investigator.

CTSI-208597 (Indiana Clinical Translational Sciences Institute), *Big Data Inference for Survival Data*, 05/01/2015 – 09/30/2016, \$12,500, Role: PI.

PRF Summer Faculty Grant (Purdue), *Theoretical foundation of High dimensional Bayesian shrinkage*, 05/01/2016 – 08/30/2016, \$8,000, Role: PI.

Professional Association

American Statistical Association, Member.

Institute of Mathematical Statistics, Member.

International Chinese Statistical Association, Member.

Honors, Awards & Fellowships

Ross-Lynn Research Scholar Fund, Purdue University, 2024.

Regina and Norman F. Carroll (Col. USAF) Research Award, Purdue University, 2024.

Ross-Lynn Research Scholar Fund, Purdue University, 2023.

College of Science Undergraduate Advising Award, Purdue University, 2023.

Regina and Norman F. Carroll (Col. USAF) Research Award, Purdue University, 2020.

The Jack Youden Prize, Technometrics, 2017.

Early Career Scholarship, Spring Research Conference, Cincinnati, OH, 2015.

Emanuel Parzen Graduate Research Fellowship Award, Dept. of Statistics, Texas A&M University, 2013.

Anant M. Kshirsagar Endowed Fellowship, Dept. of Statistics, Texas A&M University, 2012.

William S. Connor Award, Dept. of Statistics, Texas A&M University, 2010.

Outstanding Graduate Award, Peking University, 2009.

May Fourth Scholarship, Peking University, 2008.

Past Research Talks

Center for Data Science, Zhejiang University, China, Jun 2024

EcoStat 2023, Tokyo, Japan, Aug 2023

ICSA 2023, Ann Arbor, MI, Jun 2023

ICOSDA 2022, Marshall University, WV, Oct 2022.

ISBA 2022 World Meeting, Montreal, Canada, Jun 2022.

CFE-CMStatistics 2021, King's College, London, UK, Dec 2021.

Department of Department of Statistics and Probability, Michigan State University, MI, Apr 2021.

ICSA 2020 Applied Statistics Symposium, virtual, Dec 2020.

Joint Statistical Meeting, virtual, Aug 2020.

Department of Mathematics, IUPUI, Indianapolis, Feb 2020.

CFE-CMStatistics 2019, University of London, UK, Dec 2019.

11th Computational and Methodological Statistics (CMStatistics), University of Pisa, Dec 2018,

Computational strategies for large-scale statistical data analysis workshop, International Center for Mathematical Sciences, Edinburgh, UK, July 2018

The 2nd International Conference on Econometrics and Statistics (EcoSta 2018), City University of Hong Kong, Jun 2018

Hangzhou International Conference on Frontiers of Data Science, Zhejiang University, May 2018

Department of Bioinformatics and Biostatistics, the University of Louisville, Sep 2017.

Department of statistics, University of Notre Dame, Oct 2017.

10th ICSA international conference, Shanghai, Dec 2016.

Department of Mathematics, Zhejiang University, Dec 2016.

Central of Statistical Science, Tsinghua University, Dec 2016.

AISC-2016 international conference on Advances in interdisciplinary statistics and combinatorics, Greensboro, NC, Sep 2016.

ICSA 2016 symposium, Atlanta, Jun 2016.

Department of Statistics, Univeristy of Wisconsin, Madison, Apr 2016.

Joint Statistical Meeting, Seattle, Aug 2015.

Spring Research Conference, Cincinnati, Apr 2015.

Department of Biostatistics, IUPUI, Jan 2015.

Joint Statistical Meeting, Boston, Aug 2014.

Department of Statistics, Purdue University. Apr 2014.

Biostatistics Branch, National Cancer Institute, NIH. Mar 2014.

Department of Mathematics and Statistics, Boston University. Jan 2014.

International workshop on Statistical Machine Learning and Biosciences, Beijing Institute of Genomics, Chinese Academy of Sciences, May 2012.

IAMCS-KAUST Workshop on Computational Biomedicine and Geophysics, Salt Lake City. Apr 2012.