

Cheng Zhang

CONTACT INFORMATION	Department of Probability and Statistics Peking University Beijing, 100871, China	Tel: (+86) 13621390837 Email: chengzhang@math.pku.edu.cn Webpage: zcrabbit.github.io
RESEARCH INTERESTS	<ul style="list-style-type: none">• Statistics: Scalable Bayesian Inference (e.g., Markov Chain Monte Carlo, Variational Inference), Bayesian Nonparametric Models (e.g., Gaussian Processes), Sparse Modelling• Machine Learning: Probabilistic Graphical Models, Deep Bayesian Learning• Computational Biology: Bayesian Phylogenetic Inference• Scientific Computing: Bayesian Inverse Problems	
EDUCATION	University of California, Irvine, Irvine, CA Ph.D., Computational Mathematics, 2011–2016 <ul style="list-style-type: none">• Dissertation: <i>Scalable Hamiltonian Monte Carlo via Surrogate Methods</i>• Advisors:<ul style="list-style-type: none">* Hongkai Zhao (Computational Mathematics)* Babak Shahbaba (Statistics/Machine Learning) Peking University, Beijing, China M.S., Computational Mathematics, 2008–2011 B.S., Mathematics and Applied Mathematics, 2004–2008	
PROFESSIONAL POSITIONS	Assistant Professor Department of Probability and Statistics, School of Mathematical Sciences, Peking University	Aug 2019 to present
	Postdoctoral Research Fellow Computational Biology Program, Fred Hutchinson Cancer Research Center, Advisor: Frederick A. Matsen IV	Jan 2017 to July 2019
PUBLICATIONS	<ol style="list-style-type: none">1. Functional Gradient Flows for Constrained Sampling. Zhang, S.*, Yu, L.*, Cheng, Z.*, and Zhang, C. In <i>Advances in Neural Information Processing Systems</i>, 2024.2. Improved Tree Probability Estimation with Stochastic Optimization and Variance Reduction. Xie, T. Y.*, Yuan, M. S.*, Deng, M. H. and Zhang, C. <i>Statistics and Computing</i>, 34(186), 2024.3. Variational Supertrees for Bayesian Phylogenetics. Karcher, M., Zhang, C., Matsen F. A. <i>Bulletin of Mathematical Biology</i>, 86(114), 2024.4. A Variational Approach to Bayesian Phylogenetic Inference. Zhang, C. and Matsen F. A. <i>Journal of Machine Learning Research</i>, 25(145), pp. 1–56, 2024.	

5. Kernel Semi-Implicit Variational Inference.
Cheng, Z.*, Yu, L.*, Xie, T., Zhang, S., and **Zhang, C.**
In *Proceedings of the 41st International Conference on Machine Learning*, 2024.
6. Reflected Flow Matching.
Xie, T.*, Zhu, Y.*, Yu, L.*, Yang, T., Cheng, Z., Zhang, S., Zhang, X. and **Zhang, C.**
In *Proceedings of the 41st International Conference on Machine Learning*, 2024.
7. Reduced-order Model-based Variational Inference with Normalizing Flows for Bayesian Elliptic Inverse Problems.
Wu, Z. Z., **Zhang, C.** and Zhang, Z. W.
Journal of Computational and Applied Mathematics, 441, 2024.
8. ARTree: A Deep Autoregressive Model for Phylogenetic Inference.
Xie, T. and **Zhang, C.**
In *Advances in Neural Information Processing Systems*, **spotlight**, 2023.
9. Hierarchical Semi-Implicit Variational Inference with Application to Diffusion Model Acceleration.
Yu, L.*, Xie, T.*, Zhu, Y.*, Yang, T., Zhang, X. and **Zhang, C.**
In *Advances in Neural Information Processing Systems*, 2023.
10. Particle-based Variational Inference with Generalized Wasserstein Gradient Flow.
Cheng, Z.*, Zhang, S.*, Yu, L. and **Zhang, C.**
In *Advances in Neural Information Processing Systems*, 2023.
11. A Topology-marginal Composite Likelihood via a Generalized Phylogenetic Pruning Algorithm.
Jun, S. H., Nasif, H., Jennings-Shaffer, C., Rich, D. H., Kooperberg, A., Fourment, M., **Zhang, C.**, Suchard, M. A., and Matsen, F. A.
Algorithms for Molecular Biology, **18**(10), 2023.
12. A Data-driven and Model-based Accelerated Hamiltonian Monte Carlo method for Bayesian elliptic inverse problems.
Li, S., **Zhang, C.**, Zhang, Z. and Zhao, H.
Statistics and Computing, **33**(90), 2023.
13. Learnable Topological Features for Phylogenetic Inference via Graph Neural Networks.
Zhang, C.
In *Proceedings of the 11th International Conference on Learning Representations*, 2023.
14. Semi-Implicit Variational Inference via Score Matching.
Yu, L. and **Zhang, C.**
In *Proceedings of the 11th International Conference on Learning Representations*, **notable top 25% (spotlight)**, 2023.
15. Non-bifurcating Phylogenetic Tree Inference via The Adaptive LASSO.
Zhang, C.*, Dinh, V.* and Matsen F. A.
Journal of the American Statistical Association, **116**(534), pages 858-873, 2021.
16. Improved Variational Bayesian Phylogenetic Inference with Normalizing Flows
Zhang, C.
In *Advances in Neural Information Processing Systems*, **oral**(1.1%), 2020.
17. Learning, Using, and Extending Variational Distributions of Phylogenetic Trees.
Matsen F. A., Fourment, M., Karcher M., Magee A., Swanepoel, C. and **Zhang,**

- C.
In *Proceedings of the 14th Machine Learning in Computational Biology*, 2019.
18. Variational Bayesian Phylogenetic Inference.
Zhang, C. and Matsen F. A.
In *Proceedings of the 7th International Conference on Learning Representations*, 2019.
 19. Generalizing Tree Probability Estimation via Bayesian Networks.
Zhang, C. and Matsen F. A.
In *Advances in Neural Information Processing Systems*, **spotlight**(3.5%), 2018.
 20. Variational Hamiltonian Monte Carlo via Score Matching.
Zhang, C., Shahbaba, B., and Zhao, H.
Bayesian Analysis, **13**(2), pages 486–506, 2018.
 21. Probabilistic Path Hamiltonian Monte Carlo.
Dinh, V.*, Bilge, A.*, **Zhang, C.***, and Matsen F. A.
In *Proceedings of the 34th International Conference on Machine Learning*, pp. 1009–1018, 2017.
 22. Hamiltonian Monte Carlo Acceleration Using Surrogate Functions with Random Bases.
Zhang, C., Shahbaba, B., and Zhao, H.
Statistics and Computing, **27**(6), pp. 1473–1490, 2017.
 23. Precomputing Strategy for Hamiltonian Monte Carlo Method Based on Regularity in Parameter Space.
Zhang, C., Shahbaba, B., and Zhao, H.
Computational Statistics, **32**(1), pp. 253–279, 2017.

- SELECTED TALKS
- **Invited** The 3rd Joint Congress on Evolutionary Biology, Montreal, Canada. *ARTree: A Deep Autoregressive Model for Phylogenetic Inference*, July, 2024
 - **Invited** The 7th International Conference on Econometrics and Statistics, Beijing, China. *Semi-Implicit Variational Inference via Score Matching*, July, 2024
 - **Invited** International Conference on Frontiers of Data Science, Hangzhou, China. *Semi-Implicit Variational Inference via Score Matching*. July, 2024
 - **Invited** The International Chinese Statistical Association (ICSA) 2023 China Conference, Chengdu, China. *Learnable Topological Features for Phylogenetic Inference*. July, 2023
 - **Invited** The 34th Conference on Neural Information Processing Systems, Vancouver, Canada. *Improved Variational Bayesian Phylogenetic Inference with Normalizing Flows*. Dec, 2020
 - **Invited** The 17th Annual Meeting of the Chinese Society for Industrial and Applied Mathematics (CSIAM 2019), Foshan, China. *Modern Bayesian Approaches and Applications in Deep Learning*. Sep, 2019
 - **Invited** The Annual Meeting of the Canadian Society of Applied and Industrial Mathematics (CAIMS 2019), Whistler, BC. *Variational Bayesian Phylogenetic Inference*. Jun, 2019
 - **Invited** SIAM Conference on Computational Science and Engineering (CSE19), Spokane, USA. *Scalable Bayesian Inference for Inverse Problems*. Feb, 2019

- **Invited** The 32nd Conference on Neural Information Processing Systems, Montreal, Canada. *Generalizing Tree Probability Estimation via Bayesian Networks*. Dec, 2018
- **Invited** Joint Statistical Meeting 2018, Vancouver, BC. *Variational Hamiltonian Monte Carlo via Score Matching*. Aug, 2018
- **Invited** The 34th International Conference on Machine Learning, Sydney, Australia. *Probabilistic Path Hamiltonian Monte Carlo*. Aug, 2017
- **Seminar Talk** AI/ML Seminar, Department of Computer Science, UC Irvine. *Variational Hamiltonian Monte Carlo via Score Matching*. Nov, 2016

TEACHING EXPERIENCE

Instructor at Peking University

- Statistical Models and Computing Methods Fall 2020, 2021, 2022, 2023
- Bayesian Theory and Computation Spring 2021, 2022, 2024
- Modern Computational Statistics Fall 2019

Teaching Assistant at University of California, Irvine

- Math 2D - Multivariable Calculus Spring 2016
- Math 130B - Probability and Stochastic Process Winter 2016
- Math 105B - Numerical Analysis Winter 2016
- Math 2E - Multivariable Calculus Spring 2015
- Math 6G - Linear Algebra Spring 2015
- Math 2B - Single Variable Calculus Fall 2013 – Spring 2014

REVIEWER

Journals

- *Journal of Machine Learning Research*
- *Statistics and Computing*
- *Bayesian Analysis*
- *Bioinformatics*
- *IEEE Transactions on Neural Networks and Learning Systems*
- *Transactions on Machine Learning Research*

Conferences

- *ICML 2020, 2021, 2022, 2023, 2024*
- *ICLR 2021, 2022, 2023, 2024, 2025*
- *NeurIPS 2022, 2023, 2024*