Jiaxin Shi

jiaxins.io

ishijiaxin AT gmail.com

Education

2015/9-2020/6 Ph.D., Computer Science & Technology
Tsinghua University, Beijing. Advisor: Jun Zhu.

2011/8-2015/7 B.Eng., Computer Science & Technology
Tsinghua University, Beijing.

Employment

2025/12- Research Scientist
Meta Superintelligence Labs, TBD Lab, CA.

2023/3-2025/11 Research Scientist
Google DeepMind, New York.

2022/7-2023/3 Postdoctoral Scholar, Department of Statistics
Stanford University, CA. Advisor: Emily B. Fox.

2020/8-2022/6 Postdoctoral Researcher
Microsoft Research New England, Cambridge, MA, worked with Dr. Lester Mackey.

Oct-Dec 2019 Research Intern
Vector Institute, Toronto, short-term visit hosted by Prof. Roger Grosse.

Jun-Sep 2019 Research Scientist Intern
DeepMind, London, worked with Dr. Andriy Mnih and Dr. Michalis Titsias.

Jul-Sep 2018 Research Intern
RIKEN Center for Advanced Intelligence Project, Tokyo, worked with Dr. Emtiyaz Khan.

Nov-Jul 2015 Intern
Mobvoi Inc., Beijing, worked with Dr. Libin Shen.

Honors & Awards

2022 **NeurIPS 2022 Outstanding Paper Award**.
2022 Top Reviewer, NeurIPS 2022.
2020 Outstanding Thesis Award, Tsinghua University.
2019 Best Student Paper Runner-Up, Symposium on Advances in Approximate Bayesian Inference (AABI).
2018 Microsoft Research PhD Fellowship, Asia-Pacific Region.
2016 Honorable mention (ranked 1/80), Duke-Tsinghua Machine Learning Summer School.
2015 Excellent Graduate Award, Department of Computer Science & Technology, Tsinghua University.
2014 First prize (8 out of 77), Tsinghua Contribution Award of Laboratory Construction.

Professional Service

Area Chair

Neural Information Processing Systems (NeurIPS) 2023-2026
International Conference on Machine Learning (ICML) 2026, 2027
International Conference on Learning Representations (ICLR) 2027
International Conference on Artificial Intelligence and Statistics (AISTATS) 2023, 2025, 2026

Senior Program Committee

AAAI Conference on Artificial Intelligence 2022

Reviewer

Annals of Statistics
Journal of Machine Learning Research (JMLR)
Transactions on Machine Learning Research (TMLR)
International Conference on Artificial Intelligence and Statistics (AISTATS) 2021, 2022
International Conference on Learning Representations (ICLR) 2020-2026
Neural Information Processing Systems (NeurIPS) 2019-2022
International Conference on Machine Learning (ICML) 2019, 2021, 2023, 2025
Asian Conference on Machine Learning (ACML) 2019, 2020
Symposium on Advances in Approximate Bayesian Inference (AABI) 2021

Organizing Commitee

ICLR 2025 Workshop on World Models: Understanding, Modelling and Scaling
Workflow Chair. International Conference on Artificial Intelligence and Statistics (AISTATS) 2024

Patent

Jiaxin Shi, Kehang Han, Arnaud Doucet, and Michail Titsias. Masked diffusion models with state-dependent masking schedules, November 27 2025. US Patent App. 19/216,465.

Publications

(*) denotes equal contribution.

Preprints

Zachary Wu, Joshua Abramson, Thomas Frerix, Alexander E Chu, Ruijie K Zhang, Luca Schulz, Amy E Danson, Tristan OC Kwan, Wenliang K Li, Jacob Kelly, et al. Designing enzymes for new-to-nature chemistry and non-natural substrates with AlphaProtein Novo. *bioRxiv*, pages 2026–10, 2026.

Jiaxin Shi and Michalis K. Titsias. Demystifying diffusion objectives: Reweighted losses are better variational bounds. *arXiv preprint arXiv:2511.19664*, 2025.

Refereed Conference Publications

Yangtian Zhang, Zhe Wang, Arthur Gretton, Rex Ying, David van Dijk, Michalis K. Titsias, and **Jiaxin Shi**. Variational learning for insertion-based generation. *International Conference on Ma-*

chine Learning (ICML), 2026. **Spotlight Presentation (top 2.2%)**.

Patrick Pynadath, **Jiaxin Shi**, and Ruqi Zhang. CANDI: Hybrid discrete-continuous diffusion models. *International Conference on Machine Learning (ICML)*, 2026.

Celine Lee, Jing Nathan Yan, Chen Liang, **Jiaxin Shi**, Yin Zhang, Jeremiah Liu, Pengcheng Yin, Fernando Pereira, Ed Chi, Derek Cheng, et al. The efficiency gap in byte modeling. *International Conference on Machine Learning (ICML)*, 2026.

Andrew Campbell, Valentin De Bortoli, **Jiaxin Shi**, and Arnaud Doucet. Self-speculative masked diffusions. *International Conference on Representation Learning (ICLR)*, 2026.

Yixiu Zhao, **Jiaxin Shi**, Feng Chen, Shaul Druckmann, Lester Mackey, and Scott Linderman. Informed correctors for discrete diffusion models. *Advances in Neural Information Processing Systems (NeurIPS)*, 2025.

Xidong Feng, Vivek Veeriah, Marcus Chiam, Michael D Dennis, Federico Barbero, Johan Obando-Ceron, **Jiaxin Shi**, Satinder Singh, Shaobo Hou, Nenad Tomasev, and Tom Zahavy. Generating Creative Chess Puzzles. *Advances in Neural Information Processing Systems (NeurIPS)*, 2025.

Zhe Wang, **Jiaxin Shi**, Nicolas Heess, Arthur Gretton, and Michalis K. Titsias. Learning-order autoregressive models with application to molecular graph generation. In *International Conference on Machine Learning (ICML)*, 2025.

Jiaxin Shi*, Kehang Han*, Zhe Wang, Arnaud Doucet, and Michalis K. Titsias. Simplified and generalized masked diffusion for discrete data. *Advances in Neural Information Processing Systems (NeurIPS)*, 37:103131–103167, 2024.

Jiaxin Shi and Lester Mackey. A finite-particle convergence rate for stein variational gradient descent. *Advances in Neural Information Processing Systems (NeurIPS)*, 36, 2023.

Jiaxin Shi, Ke Alexander Wang, and Emily B. Fox. Sequence modeling with multiresolution convolutional memory. *International Conference on Machine Learning (ICML)*, 2023.

Jiaxin Shi, Yuhao Zhou, Jessica Hwang, Michalis K. Titsias, and Lester Mackey. Gradient estimation with discrete Stein operators. *Advances in Neural Information Processing Systems (NeurIPS)*, 2022. **NeurIPS 2022 Outstanding Paper Award**.

Zhijie Deng, **Jiaxin Shi**, and Jun Zhu. NeuralEF: Deconstructing kernels by deep neural networks. *International Conference on Machine Learning (ICML)*, 2022.

Michalis K. Titsias and **Jiaxin Shi**. Double control variates for gradient estimation in discrete latent variable models. *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2022.

Jiaxin Shi, Chang Liu, and Lester Mackey. Sampling with mirrored Stein operators. *International Conference on Learning Representations (ICLR)*, 2022. **Spotlight Presentation (top 5.2%)**.

Shengyang Sun, **Jiaxin Shi**, Andrew Gordon Wilson, and Roger B Grosse. Scalable variational Gaussian processes via harmonic kernel decomposition. In *International Conference on Machine Learning (ICML)*, pages 9955–9965, 2021.

Yuhao Zhou, **Jiaxin Shi**, and Jun Zhu. Nonparametric score estimators. In *International Conference on Machine Learning (ICML)*, pages 11513–11522, 2020.

Jiaxin Shi, Michalis K. Titsias, and Andriy Mnih. Sparse orthogonal variational inference for Gaus-

sian processes. *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2020. **Best Student Paper Runner-Up at AABI 2019**.

Yang Song, Sahaj Garg, **Jiaxin Shi**, and Stefano Ermon. Sliced score matching: A scalable approach to density and score estimation. *The 35th Conference on Uncertainty in Artificial Intelligence (UAI)*, 2019. **Oral Presentation (top 8.7%)**.

Jiaxin Shi, Mohammad Emtiyaz Khan, and Jun Zhu. Scalable training of inference networks for Gaussian-process models. In *International Conference on Machine Learning (ICML)*, pages 5758–5768, 2019.

Shengyang Sun*, Guodong Zhang*, **Jiaxin Shi***, and Roger Grosse. Functional variational Bayesian neural networks. In *International Conference on Learning Representations (ICLR)*, 2019.

Yucen Luo, Tian Tian, **Jiaxin Shi**, Jun Zhu, and Bo Zhang. Semi-crowdsourced clustering with deep generative models. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 3216–3226. 2018.

Jiaxin Shi, Shengyang Sun, and Jun Zhu. A spectral approach to gradient estimation for implicit distributions. *International Conference on Machine Learning (ICML)*, pages 4644–4653, 2018.

Jingwei Zhuo, Chang Liu, **Jiaxin Shi**, Jun Zhu, Ning Chen, and Bo Zhang. Message passing Stein variational gradient descent. In *International Conference on Machine Learning (ICML)*, pages 6013–6022, 2018.

Jiaxin Shi*, Shengyang Sun*, and Jun Zhu. Kernel implicit variational inference. *International Conference on Learning Representations (ICLR)*, 2018.

Journal Publications

Ke Alexander Wang, **Jiaxin Shi**, and Emily B Fox. Test-time regression: a unifying framework for designing sequence models with associative memory. *Journal of Machine Learning Research*, 27(169):1–41, 2026.

Zhijie Deng*, **Jiaxin Shi***, Hao Zhang, Peng Cui, Cewu Lu, and Jun Zhu. Neural eigenfunctions are structured representation learners. *IEEE Transaction on Pattern Analysis and Machine Intelligence (TPAMI)*, 2025.

Mengchen Liu, **Jiaxin Shi**, Kelei Cao, Jun Zhu, and Shixia Liu. Analyzing the training processes of deep generative models. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):77–87, 2018.

Mengchen Liu, **Jiaxin Shi**, Zhen Li, Chongxuan Li, Jun Zhu, and Shixia Liu. Towards better analysis of deep convolutional neural networks. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):91–100, 2017. **Most cited paper of TVCG 2017**.

Fanglue Zhang, Jue Wang, Eli Shechtman, Ziye Zhou, **Jiaxin Shi**, and Shimin Hu. Plenopatch: Patch-based plenoptic image manipulation. *IEEE Transactions on Visualization and Computer Graphics*, 23(5):1561–1573, 2017.

Workshop Papers

Zhe Wang, **Jiaxin Shi**, Nicolas Heess, Michalis K. Titsias, Arthur Gretton, and Yee Whye Teh. Modeling molecular sequences with learning-order autoregressive models. In *ICML 2025 Generative AI*

and Biology (GenBio) Workshop, 2025.

Jiaxin Shi and Lester Mackey. A finite-particle convergence rate for Stein variational gradient descent. In *OPT 2022: Optimization for Machine Learning (NeurIPS 2022 Workshop)*, 2022.

Alex Wang, Matthew E. Levine, **Jiaxin Shi**, and Emily Fox. Learning absorption rates in glucose-insulin dynamics from meal covariates. In *NeurIPS 2022 Workshop on Learning from Time Series for Health*, 2022.

Michalis K. Titsias and **Jiaxin Shi**. Double control variates for gradient estimation in discrete latent variable models. *4th Symposium on Advances in Approximate Bayesian Inference*, 2022.

Jiaxin Shi, Chang Liu, and Lester Mackey. Sampling with mirrored Stein operators. *4th Symposium on Advances in Approximate Bayesian Inference*, 2022.

Shengyang Sun*, **Jiaxin Shi***, and Roger Grosse. Neural networks as inter-domain inducing points. *3rd Symposium on Advances in Approximate Bayesian Inference*, 2020.

Jiaxin Shi, Michalis Titsias, and Andriy Mnih. Sparse orthogonal variational inference for Gaussian processes. *2nd Symposium on Advances in Approximate Bayesian Inference*, Vancouver, 2019.

Yuhao Zhou, **Jiaxin Shi**, and Jun Zhu. Spectral estimators for gradient fields of log-densities. *ICML Workshop on Stein's Method*, Long Beach, CA, 2019.

Shengyang Sun*, Guodong Zhang*, **Jiaxin Shi***, and Roger Grosse. Functional variational Bayesian neural networks. *NeurIPS Bayesian Deep Learning Workshop*, Montréal, 2018.

Yucen Luo, Tian Tian, **Jiaxin Shi**, Jun Zhu, and Bo Zhang. Semi-crowdsourced clustering with deep generative models. *ICML Workshop on Theoretical Foundations and Applications of Deep Generative Models*, Stockholm, 2018.

Jiaxin Shi*, Shengyang Sun*, and Jun Zhu. Implicit variational inference with kernel density ratio fitting. *ICML Workshop on Implicit Models*, Sydney, 2017.

Open Source Projects

I created and lead the development of **ZhuSuan** (GitHub, Documentation), an open-source differentiable probabilistic programming library based on Tensorflow. The project had **2K stars** on GitHub by Aug 2021.

Jiaxin Shi, Jianfei Chen, Jun Zhu, Shengyang Sun, Yucen Luo, Yihong Gu, and Yuhao Zhou. ZhuSuan: A library for Bayesian deep learning. *arXiv preprint arXiv:1709.05870*, 2017.

Invited Talks

Oct 2026 **COLM 2026 Workshop: Non-Autoregressive Language Models for Fast & Flexible Text Generation**, San Francisco, CA.

Aug 2026 **Simons Institute for the Theory of Computing**, Diffusion Generative Modeling: Progress and Next Steps Workshop, Berkeley, CA.

Feb 2026 **The University of Texas at Austin**, NSF AI Institute for Foundations of Machine Learning, IFML Seminar Series.

Dec 2025 **NeurIPS 2025 Workshop on Structured Probabilistic Inference & Generative Modeling**, San Diego, CA.

Jul 2025 **Microsoft Research New England.** Generative Modeling & Sampling Seminar, virtual.

Jul 2025 **INFORMS Applied Probability Society Conference**. Denoising Generative Models: Theory and Applications session, Atlanta, GA.

Feb 2025 **Toyota Technological Institute at Chicago**.

Nov 2024 **Learning on Graphs Conference (LoG) New York meetup**.

Oct 2024 **Imperial College London**, Department of Computing.

Sep 2024 **Generative Models and Uncertainty Quantification (GenU 2024)**, Copenhagen, Denmark.

May 2024 **Institute of Computing Technology, Chinese Academy of Sciences**, Beijing.

May 2024 **Tsinghua University**, TSAIL group.

May 2024 **University of Warwick**, Foundations of AI Seminar Series.

May 2024 **University College London**, Fundamentals of Statistical Machine Learning Group.

Mar 2024 **Workshop on Functional Inference and Machine Intelligence (FIMI 2024)**, Bristol, UK.

Mar 2024 **Workshop on Optimal Transport: From Theory To Applications (OT Berlin 2024)**, Berlin.

Dec 2023 **FAI-Seminar (International Seminar on Foundational Artificial Intelligence)**, virtual.

May 2023 **Westlake University, City University of Hong Kong & Peking University**, Deep Learning Theory Seminar Series, virtual.

Mar 2023 **University of California, Berkeley**, Biostats Seminar.

Jan 2023 **Beijing Academy of Artificial Intelligence**, Qingyuan Talk.

Dec 2022 **North Carolina State University**, Efficient & Intelligent Computing Lab.

Apr 2022 **Banff International Research Station**, Advances in Stein's method and its applications in Machine Learning and Optimization, Banff, Alberta.

Feb 2022 **Imperial College London & Oxford**, Centre for Doctoral Training in Statistics & Machine Learning.

Jan 2022 **Max Planck Institute for Intelligent Systems & Cyber Valley**, Scientific Symposium.

Aug 2021 **RIKEN Center for Advanced Intelligence Project** (RIKEN-AIP), Tokyo.

Sep 2019 **University of Bristol**, School of Mathematics.

May 2019 **PaperWeekly local meetup**, Beijing.

Apr 2019 **RealAI Inc.**, Beijing.

Nov 2018 **Nanjing University**, Symposium on Machine Learning and Applications (MLA).

Jul 2018 **Renmin University**, International Forum on Statistics, Beijing.

Mar 2018 **GPU Technology Conference**, San Jose, CA.

Jun 2017 **Jiangmen TechBeat**, Beijing.

Teaching

Oct 2025 **Guest Lecture**
Deep Learning (StatML.io), Oxford University.

2019-2020 **Instructor of the lectures on probabilistic programming**
70240413: Statistical Machine Learning, Tsinghua University.

Spring 2018 **Guest Lecture**
70240033: Artificial Intelligence, Tsinghua University.

Spring 2018 **Teaching Assistant**
70240413: Statistical Machine Learning, Tsinghua University.

Jul-Aug 2017 **Teaching Assistant**
Duke-Tsinghua Machine Learning Summer School 2017.

Spring 2017 **Teaching Assistant**
70240413: Statistical Machine Learning, Tsinghua University.