Yuxuan Song

Tsinghua University – Beijing, China

+8615121167187 • yxsong0816@gmail.com

Homepage: `https://yuxuansong.com/`

Education

Tsinghua University **Beijing, China**
Sep. 2022-Jan. 2026(expected)

- Ph.D. in Computer Science and Technology
- Advised by Prof. Wei-Ying Ma.

Shanghai Jiao Tong University **Shanghai, China**
Sep. 2013-Mar. 2020

- Bachelor and Master of Science in Computer Science and Technology
- Research Assistant, Apex Data and Knowledge Management Lab
- Advised by Prof. Yong Yu and Prof. Weinan Zhang.

Professional

Bytedance Seed, LLM team Research Intern — May 2024 - current
Microsoft Research Asia Research Intern — Sep. 2019 - Mar. 2020

Research Interests

My ultimate goal is to integrate generalizable intelligence (LLMs) with advanced AI tools (e.g., AF3) to solve challenging real-world scientific problems and boost scientific discovery. My Ph.D. research has focused on geometric generative models (bio/chemical/material structures), discrete diffusion models (biological sequences, Diffusion LLMs), and reasoning in LLMs.

Publications (* for Equal Contribution)

- **Y. Song***, Z. Zhang*, Y. Pei*, J. Gong, Q. Yu, Z. Zhang, M. Wang, H. Zhou, J. Liu, W.-Y. Ma. ShortListing Model: A Streamlined SimplexDiffusion for Discrete Variable Generation. **NeurIPS 2025**.
- **Y. Song***, Z. Zhang*, C. Luo*, P. Gao, F. Xia, H. Luo, Z. Li, Y. Yang, H. Yu, X. Qu, Y. Fu, J. Su, G. Zhang, W. Huang, M. Wang, L. Yan, X. Jia, J. Liu, W.-Y. Ma, Y.-Q. Zhang, Y. Wu, H. Zhou. Seed Diffusion: A Large-Scale Diffusion Language Model with High-Speed Inference. **Technical Report**.
- **Y. Song***, J. Gong*, Y. Qu, M. Zheng, H. Zhou, J. Liu, W. Ma . Unified Generative Modeling of 3D Molecules with Bayesian Flow Networks. **ICLR 2024** (**Oral 85/7262**)
- H. Wu*, **Y. Song***, J. Gong, Z. Cao, Y. Ouyang, J. Zhang, H. Zhou, W. Ma, J. Liu. A Periodic Bayesian Flow for Material Generation. **ICLR 2025** (**Spotlight**)
- Z. Zhang*, **Y. Song***, Y. Wang, J. Gong, H. Wu, D. Zhou, H. Zhou, W.-Y. Ma. Accelerating 3D Molecule Generative Models with Trajectory Diagnosis. **NeurIPS 2025**.
- J. Gong*, Y. Pei*, S. Long*, **Y. Song***, Z. Zhang, W. Huang, Z. Cao, S. Zhang, H. Zhou, W. Ma. Steering Protein Family Design through Profile Bayesian Flow. **ICLR 2025** (**Oral 207/11672**)
- **Y. Song***, J. Gong*, M. Xu, Z. Cao, Y. Lan, S. Ermon, H. Zhou, W. Ma . Equivariant Flow Matching with Hybrid Probability Transport for 3D Molecule Generation. **NeurIPS 2023**
- **Y. Song***, J. Shi*, J. Gong*, M. Xu, S. Ermon, H. Zhou, W.-Y. Ma. Smooth Interpolation for Improved Discrete Graph Generative Models. **ICML 2025**.

- **Y. Song***, Z. Zhang*, Y. Pei*, J. Gong, M. Wang, H. Zhou, J. Liu, W.-Y. Ma. ShortListing Model: A Streamlined Simplex Diffusion for Biological Sequence Generation. **ICLR 2025 MLGenX Workshop**.
- K. Qiu*, **Y. Song***, Z. Fan, P. Liu, Z. Zhang, M. Zheng, H. Zhou, W.-Y. Ma. Piloting Structure-Based Drug Design via Modality-Specific Optimal Schedule. **ICML 2025**.
- Y. Qu*, K. Qiu*, **Y. Song***, J. Gong, J. Han, M. Zheng, H. Zhou, W. Ma. MolCRAFT: Structure-Based Drug Design in Continuous Parameter Space. **ICML 2024**.
- B. Qiang*, **Y. Song***, M. Xu, J. Gong, B. Gao, H. Zhou, W. Ma, Y. Lan. Coarse-to-Fine: a Hierarchical Diffusion Model for Molecule Generation in 3D. **ICML 2023**.
- W. Shi*, **Y. Song***, H. Zhou, L. Li. Follow Your Path: a Progressive Method for Knowledge Distillation. **ECML/PKDD 2021**.
- **Y. Song**, L. Yu, Z. Cao, Z. Zhou, J. Shen, S. Shao, W. Zhang and Y. Yu. Improving Domain Adaptation with Variational Information Bottleneck. **ECAI 2020**.
- **Y. Song**, N. Miao, H. Zhou, L. Yu and L. Li. Improving Maximum Likelihood Training for Text Generation with Density Ratio Estimation. **AISTATS 2020**.
- **Y. Song**, M. Xu, L. Yu, H. Zhou, S. Shao and Y. Yu. Infomax Neural Joint Source-Channel Coding via Adversarial Bit Flip. **AAAI 2020**.
- **Y. Song**, H. Cai, K. Ren, W. Zhang and Y. Yu. Volume Ranking and Sequential Selection in Programmatic Display Advertising. **CIKM 2017**
- H. Wu, **Y. Song**, Z. Zhang, Z. Zhang, H. Zhou, W.-Y. Ma, J. Liu. Rationalized All-Atom Protein Design with Unified Multi-modal Bayesian Flow. **NeurIPS 2025**.
- R. Jiao, H. Wu, W. Huang, **Y. Song**, Y. Ouyang, Y. Rong, T. Xu, P. Wang, H. Zhou, W.-Y. Ma, J. Liu, Y. Liu. MOF-BFN: Metal-Organic Frameworks Structure Prediction via Bayesian Flow Networks. **NeurIPS 2025**.
- W. Liu, J. Feng, H. Yu, **Y. Song**, Y. Li, S. Zhang, L. Bai, W.-Y. Ma, H. Zhou. Retro-R1: LLM-based Agentic Retrosynthesis. **NeurIPS 2025**.
- J. Han*, M. Jiang*, **Y. Song**, J. Leskovec, S. Ermon, M. Xu*. f-PO: Generalized Preference Optimization with f-divergence Minimization. **AISTATS 2025**
- Y. Wang, **Y. Song**, M. Xu, R. Wang, H. Zhou, W. Ma. RetroDiff: Retrosynthesis as Multi-stage Distribution Interpolation. **AISTATS 2025**
- B. Qiang, W. Shi, **Y. Song** and M. Wu. PROflow: An iterative refinement model for PROTAC-induced structure prediction. **ICLR 2024 GEM workshop, Oral**
- N. Miao, **Y. Song**, H. Zhou and L. Li. Do You Have the Right Scissors? Tailoring Pre-trained Language Models via Monte-Carlo Methods. **ACL 2020**
- Z. Zhou, J. Liang, **Y. Song**, L. Yu, H. Wang, Z. Zhang, W. Zhang and Y. Yu. Lipschitz Generative Adversarial Nets. **ICML 2019**
- G. Lu, Z. Zhou, **Y. Song**, K. Ren and Y. Yu.Guiding the One-to-one Mapping in CycleGAN via Optimal Transport. **AAAI 2019**
- Z. Zhou, H. Cai, S. Rong, **Y. Song**, K. Ren, W. Zhang, Y. Yu and J. Wang. Activation Maximization Generative Adversarial Nets. **ICLR 2018**

Honors and Awards

- **Bytedance Scholarship Awards**.(15 Ph.D. students from Singapore and China) 2024
- **Dean Prize of AIR, Tsinghua Univeristy**.(Top 3 Ph.D. students of the institute) 2025&2024
- **Shanghai Jiao Tong University Outstanding Graduate**. 2017

- **Shanghai Jiao Tong University Excellent Scholarship**. 2014&2015&2016
- **The First Prize in National High-School Mathematical Olympiad**. 2012

Academic Services

- **Conference Reviewer**: NeurIPS (2020-2025),ICML(2021-2025), ICLR(2024-2025), AISTATS(2025).