Shansan Gong

Google Scholar · Homepage · The University of Hong Kong · hisansas@gmail.com

---

RESEARCH SUMMARY

My research focuses on **text diffusion models** for conditional generation, reasoning, and scaling. I am passionate about developing the next generation of diffusion reasoning models and exploring their capabilities. My interests also include long-context modeling and building personalized, tool-augmented systems.

EDUCATION

2023.09 – Present **The University of Hong Kong (HKU)** Hong Kong SAR
PhD in Computer Science

2019.06 – 2022.03 **Shanghai Jiao Tong University (SJTU)** Shanghai, CN
M.Sc. in Computer Science

2015.09 – 2019.06 **Shanghai Jiao Tong University (SJTU)** Shanghai, CN
B.Eng. in Information Engineering

RESEARCH EXPERIENCE

2025.01 – 2025.07 **Research Intern, Apple** Seattle, US
Code Generation with Masked Diffusion Models; Diffusion-native RL Algorithms

2023.11 – 2024.10 **Research Intern, Tencent AI Lab** Shenzhen, CN
Reasoning & Scaling of Diffusion Language Models

2022.05 – 2023.08 **Researcher, Shanghai AI Lab** Shanghai, CN
Diffusion Text for Conditional Generation; Long-Context Modeling

---

PUBLICATIONS

* denotes equal contribution.

Selected Publications

Jiacheng Ye, **Shansan Gong**, Jiahui Gao, Junming Fan, Shuang Wu, Wei Bi, Haoli Bai, Lifeng Shang, Lingpeng Kong. Dream-VL & Dream-VLA: Open Vision-Language and Vision-Language-Action Models with Diffusion Language Model Backbone. **Technical Report, 2025**

Huangjie Zheng, **Shansan Gong**, Ruixiang Zhang, Tianrong Chen, Jiatao Gu, Mingyuan Zhou, Navdeep Jaitly, Yizhe Zhang. Continuously Augmented Discrete Diffusion Model for Categorical Generative Modeling (CADD). **ICLR, 2026**

Zirui Wu, Lin Zheng, Zhihui Xie, Jiacheng Ye, Jiahui Gao, **Shansan Gong**, Yansong Feng, Zhenguo Li, Wei Bi, Guorui Zhou, Lingpeng Kong. DreamOn: Diffusion Language Models For Code Infilling Beyond Fixed-size Canvas. [code] **ICLR, 2026**

Shansan Gong, Ruixiang Zhang, Huangjie Zheng, Jiatao Gu, Navdeep Jaitly, Lingpeng Kong, Yizhe Zhang. DiffuCoder: Understanding and Improving Masked Diffusion Models for Code Generation. [code] **ICLR, 2026**

Shansan Gong*, Shivam Agarwal*, Yizhe Zhang, Jiacheng Ye, Lin Zheng, Mukai Li, Chenxin An, Peilin Zhao, Wei Bi, Jiawei Han, Hao Peng, Lingpeng Kong. Scaling Diffusion Language Models via Adaptation from Autoregressive Models (DiffuLLaMA). [code] **ICLR, 2025**

Jiacheng Ye, Jiahui Gao, **Shansan Gong**, Lin Zheng, Xin Jiang, Zhenguo Li, Lingpeng Kong. Beyond Autoregression: Discrete Diffusion for Complex Reasoning and Planning. [code] **ICLR, 2025**

Jiacheng Ye*, **Shansan Gong***, Liheng Chen, Lin Zheng, Jiahui Gao, Han Shi, Chuan Wu, Zhenguo Li, Wei Bi, Lingpeng Kong. Diffusion of Thoughts: Chain-of-Thought Reasoning in Diffusion Language Models (DoT). [code] **NeurIPS, 2024**

Shansan Gong, Mukai Li, Jiangtao Feng, Zhiyong Wu, Lingpeng Kong. DiffuSeq-v2: Bridging Discrete and Continuous Text Spaces for Accelerated Seq2Seq Diffusion Models. [code] **EMNLP Findings, 2023**

Shansan Gong, Mukai Li, Jiangtao Feng, Zhiyong Wu, Lingpeng Kong. DiffuSeq: Sequence to Sequence Text Generation With Diffusion Models. [code] **ICLR, 2023**

OTHER PUBLICATIONS

Zhihui Xie, Jiacheng Ye, Lin Zheng, Jiahui Gao, Jingwei Dong, Zirui Wu, Xueliang Zhao, **Shansan Gong**, Xin Jiang, Zhenguo Li, Lingpeng Kong. Dream-Coder 7B. [code] **Technical Report, 2025**

Chenxin An, Zhihui Xie, Xiaonan Li, Lei Li, Jun Zhang, **Shansan Gong**, Ming Zhong, Jingjing Xu, Xipeng Qiu, Mingxuan Wang, Lingpeng Kong. Polaris: A Post-Training Recipe for Scaling Reinforcement Learning on Advanced Reasoning Models. [code] **Technical Report, 2025**

Mukai Li, Lei Li, **Shansan Gong**, Qi Liu. GIRAFFE: Design Choices for Extending the Context Length of Visual Language Models. [code] **ACL, 2025**

Chenxin An, Jun Zhang, Ming Zhong, Lei Li, **Shansan Gong**, Yao Luo, Jingjing Xu, Lingpeng Kong. Why Does the Effective Context Length of LLMs Fall Short?. [code] **ICLR, 2025**

Chenxin An, **Shansan Gong**, Ming Zhong, Mukai Li, Jun Zhang, Lingpeng Kong, Xipeng Qiu. L-Eval: Instituting Standardized Evaluation for Long Context Language Models. [code] **ACL (Outstanding Paper), 2024**

Chenxin An, Fei Huang, Jun Zhang, **Shansan Gong**, Xipeng Qiu, Chang Zhou, Lingpeng Kong. Training-Free Long-Context Scaling of Large Language Models. [code] **ICML, 2024**

Xueliang Zhao, Xinting Huang, Tingchen Fu, Qintong Li, **Shansan Gong**, Lemao Liu, Wei Bi, Lingpeng Kong. BBA: Bi-Modal Behavioral Alignment for Reasoning with Large Vision-Language Models. **ACL Findings, 2024**

Mukai Li, **Shansan Gong**, Jiangtao Feng, Yiheng Xu, Jun Zhang, Zhiyong Wu, Lingpeng Kong. In-Context Learning with Many Demonstration Examples. [code] **arXiv, 2023**

Shansan Gong*, Zelin Zhou*, Shuo Wang, Fengjiao Chen, Xiujie Song, Xuezhi Cao, Yunsen Xian, Kenny Zhu. Transferable and Efficient: Unifying Dynamic Multi-Domain Product Categorization. **ACL Industry, 2023**

Shansan Gong, Kenny Q. Zhu. Positive, Negative and Neutral: Modeling Implicit Feedback in Session-based News Recommendation. [code] **SIGIR, 2022**

INVITED TALKS

2025.12	"DiffuCoder," Discrete Diffusion Reading Group, Cornell (d-llms.com)
2025.11	"Diffusion Language Models: From Fundamentals to Better Coding with RL," Jia Li's Group, Tsinghua University
2025.06	"Introducing Diffusion Large Language Models (dLLMs): DiffuLLaMA and DiffuCoder," Google Efficiency Research
2025.06	"The Journey of Text Diffusion Models," Noah's Group, UW NLP
2023.06	"DiffuSeq," Youth PhD Talk ICLR 2023 AI Time
2023.05	"Incorporate Diffusion Models into Conditional Text Generation," Global Lunch Seminar, SJTU CS Department

HONORS & AWARDS

ongoing	HKU Presidential PhD Scholarship (HKU-PS)
2024	ACL 2024 Outstanding Paper

2024	Tencent Rhino-bird Research Elite Program, Outstanding Student
2022	SIGIR 2022 Student Travel Award
2022	Outstanding Graduate in Shanghai Municipality

SERVICE

Reviewer	ACL, EMNLP, NeurIPS, ICLR, ICML, COLM, COLING; ACM Computing Surveys, IEEE Journals, TPAMI
Teaching Assistant	COMP2121 Discrete Mathematics, COMP7104 Advanced Database Systems
Community	Co-organizer of the HKU NLP Seminar