

CHENYU (MONICA) WANG

<https://chenyuwang-monica.github.io> | wangchy@mit.edu | (+1)617-902-8630

MIT Stata Center, 32 Vassar St. G418, Cambridge MA 02139

EDUCATION BACKGROUND

Massachusetts Institute of Technology

Ph.D. Student in Electrical Engineering and Computer Science (EECS) | GPA 5.0/5.0

Advised by Prof. Tommi Jaakkola

Cambridge, MA

Aug. 2022-present

Tsinghua University

Bachelor of Economics, Minor in Data Science and Technology | GPA 3.99/4.0 (Ranking 1/192)

Advised by Prof. Mingsheng Long, Prof. Mengdi Wang, and Prof. Cyrus Shahabi

Beijing, China

Sep. 2018-Jun. 2022

University of California, Berkeley

Exchange Student, Department of Statistics (Instructed by Prof. Nouredine El Karoui) | GPA 4.0/4.0

Berkeley, CA

Jan. 2021-Jun. 2021

RESEARCH INTERESTS

My research interests lie broadly in machine learning, representation learning, generative models, and AI for science. Recently my research focuses on multi-modal representation learning, diffusion generative models, and controlled generation, with applications to biology and drug discovery.

PUBLICATIONS & PREPRINTS

(*: Equal Contribution)

[Google Scholar](#)

- **Chenyu Wang***, Masatoshi Uehara*, Yichun He, Amy Wang, Tommaso Biancalani, Avantika Lal, Tommi Jaakkola, Sergey Levine, Hanchen Wang, Aviv Regev. Fine-Tuning Discrete Diffusion Models via Reward Optimization with Applications to DNA and Protein Design. In *NeurIPS 2024 Workshop on Machine Learning in Structural Biology*. [\[link\]](#)
- **Chenyu Wang***, Sharut Gupta*, Xinyi Zhang, Sana Tonekaboni, Stefanie Jegelka, Tommi Jaakkola, Caroline Uhler. An Information Criterion for Controlled Disentanglement of Multimodal Data. In *NeurIPS 2024 Workshop on Unifying Representations in Neural Models*. (**Honorable Mention Award**) [\[link\]](#)
- Xiner Li, Yulai Zhao, **Chenyu Wang**, Gabriele Scalia, Gokcen Eraslan, Surag Nair, Tommaso Biancalani, Aviv Regev, Sergey Levine, Masatoshi Uehara. Derivative-Free Guidance in Continuous and Discrete Diffusion Models with Soft Value-Based Decoding. In *NeurIPS 2024 Workshop on AI for New Drug Modalities*. [\[link\]](#)
- Sharut Gupta*, **Chenyu Wang***, Yifei Wang*, Tommi Jaakkola, Stefanie Jegelka. In-Context Symmetries: Self-Supervised Learning through Contextual World Models. In *Advances in Neural Information Processing Systems, NeurIPS 2024*. [\[link\]](#)
- Hannes Stark*, Bowen Jing*, **Chenyu Wang**, Gabriele Corso, Bonnie Berger, Regina Barzilay, Tommi Jaakkola. Dirichlet Flow Matching with Applications to DNA Sequence Design. In *International Conference on Machine Learning, ICML 2024*. [\[link\]](#)
- **Chenyu Wang**, Sharut Gupta, Caroline Uhler, Tommi S. Jaakkola. Removing Biases from Molecular Representations via Information Maximization. In *International Conference on Learning Representations, ICLR 2024*. [\[link\]](#)
- **Chenyu Wang***, Joseph Kim*, Le Cong, Mengdi Wang. Neural Bandits for Protein Sequence Optimization. In *56th Annual Conference on Information Sciences and Systems, CISS 2022*. [\[link\]](#)
- **Chenyu Wang***, Zongyu Lin*, Xiaochen Yang, Jiao Sun, Mingxuan Yue, Cyrus Shahabi. HAGEN: Homophily-Aware Graph Convolutional Recurrent Network for Crime Forecasting. In *AAAI Conference on Artificial Intelligence, AAAI 2022*. (**Oral Presentation.**) [\[link\]](#)
- Yang Shu*, Zhangjie Cao*, **Chenyu Wang**, Jianmin Wang, Mingsheng Long. Open Domain Generalization with Domain-augmented Meta-learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2021*. [\[link\]](#)

RESEARCH EXPERIENCE

Representation Learning and Generative Models with Applications to Biology

Cambridge, MA

Advised by Prof. Tommi Jaakkola, MIT EECS

Tree-Based Neural Bandits for High-Value Protein Design

Advised by Prof. Mengdi Wang, Department of Electrical Engineering, Princeton University

Homophily-Aware Graph Convolutional Recurrent Network for Crime Forecasting

Advised by Prof. Cyrus Shahabi, Department of Computer Science, USC

Open Domain Generalization with Domain-Augmented Meta-Learning

Advised by Prof. Mingsheng Long, School of Software, Tsinghua University

Understanding Chinese Bond Yield Curve: Excess Return Prediction

Advised by Prof. Hao Wang, SEM, Tsinghua

Aug. 2022-present

Princeton, NJ

Jun. 2021-Dec. 2021

Los Angeles, CA

Jan. 2021-Jun. 2021

Beijing, China

Sept. 2020-Nov. 2020

Beijing, China

Jun. 2020-Aug. 2020

HONORS & AWARDS

- MIT EECS Great Educators Fellowship, 2022
- Outstanding Undergraduate in Tsinghua (2% in Tsinghua), 2022
- Outstanding Undergraduate in Beijing, 2022
- Chen Daisun Scholarship (3 in Tsinghua SEM), 2022
- Undergraduate Commencement Student Speaker of Tsinghua SEM, 2022
- Meritorious Winner in MCM/ICM Mathematical Contest in Modelling, 2021
- Chen Xiaoyue Scholarship, 2021
- Tang Lixin Scholarship (50 in Tsinghua), 2020
- National Scholarship (0.2% in China), 2019
- Athletics Excellence Scholarship of Tsinghua, 2019
- First Class Scholarship for Freshmen of Tsinghua, 2018
- Gold medalist of 50th International Chemistry Olympiad (4 in China, 6th place in the world), 2018
- Silver medalist of 15th China Girl's Mathematical Olympiad (50 in China), 2016

WORK EXPERIENCE

Genentech

Research Intern in Dr. Aviv Regev's Lab

South San Francisco

May 2024-Aug. 2024

Jane Street Asia Limited

Quantitative Trading Intern (Return offer extended)

Hong Kong

Jun. 2021-Sept. 2021

WizardQuant Capital Management

Quantitative Research Intern, Quantitative Research Department

Zhuhai, China

Jun. 2020-Aug. 2020

Techsharpe Quant Capital Management

Data Analyst Intern, Trading Department

Beijing, China

Jan. 2020-Feb. 2020

SERVICES

- Reviewer: ICLR 2025, NeurIPS 2024, PLOS Computational Biology, ICML 2024/ICLR 2024/NeurIPS 2024 workshops

LEADERSHIP & ACTIVITIES

- Team Leader, Meritorious Winner in 2021 MCM/ICM Mathematical Contest in Modelling. Feb. 2021
- Co-president, Banking & Investment Mentor Program (A 10-year global non-profit organization). Feb. 2021-Feb. 2022
- Director of Department of Sports, Student Union of Tsinghua SEM. Mar. 2019-Sept. 2020

SKILLS & INTERESTS

- **Languages:** English (Proficient; TOEFL: 110/120); Mandarin (Native)
- **Technical Skills:** Python/C++/Matlab; Deep learning framework: PyTorch, Tensorflow; Basic knowledge of SQL and Linux.
- **Interests:** Sports (1st place in 4*400m; member of SEM basketball and soccer team), Chinese Zither (Amateur Certificate 9)