

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

M.S. in Electrical Engineering and Computer Science (EECS)

Advised by Prof. Tommi Jaakkola

Cambridge, MA

Aug. 2022-present

Aug. 2022-Feb. 2025

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. Noureddine El Karoui) | GPA 4.0/4.0

Berkeley, CA

Jan. 2021-Jun. 2021

RESEARCH INTERESTS

My research focuses on developing controllable and efficient generative models, via reinforcement learning, multi-modal learning, and representation learning. I work across various application domains, including language models, vision, and scientific data (e.g. biochemistry). My recent work explores:

- Reinforcement learning, reasoning and inference-time alignment for deep generative models and language models.
- Multi-modal learning, especially multi-modal representations and their interactions with generative models.
- Enhancing biochemistry discovery through generative models and agents.

PUBLICATIONS & PREPRINTS

(*: Equal Contribution)

[Google Scholar](#)

- **Chenyu Wang**, Paria Rashidinejad, DiJia Su, Song Jiang, Sid Wang, Siyan Zhao, Cai Zhou, Shannon Zejiang Shen, Feiyu Chen, Tommi Jaakkola, Yuandong Tian, Bo Liu. SPG: Sandwiched Policy Gradient for Masked Diffusion Language Models. In *arXiv preprint 2025*. [\[link\]](#)
- Sharut Gupta, Shobhita Sundaram, **Chenyu Wang**, Stefanie Jegelka, Phillip Isola. Better Together: Leveraging Unpaired Multimodal Data for Stronger Unimodal Models. In *arXiv preprint 2025*. [\[link\]](#)
- Cai Zhou, Chenxiao Yang, Yi Hu, **Chenyu Wang**, Chubin Zhang, Muhan Zhang, Lester Mackey, Tommi Jaakkola, Stephen Bates, Dinghuai Zhang. Coevolutionary Continuous Discrete Diffusion: Make Your Diffusion Language Model a Latent Reasoner. In *arXiv preprint 2025*. [\[link\]](#)
- Siyan Zhao, Mengchen Liu, Jing Huang, Miao Liu, **Chenyu Wang**, Bo Liu, Yuandong Tian, Guan Pang, Sean Bell, Aditya Grover, Feiyu Chen. Inpainting-Guided Policy Optimization for Diffusion Large Language Models. In *arXiv preprint 2025*. [\[link\]](#)
- Hanqun Cao*, Xinyi Zhou*, Zijun Gao*, **Chenyu Wang**, Xin Gao, Zhi Zhang, Chunbin Gu, Ge Liu, Pheng-Ann Heng. Lightweight MSA Design Advances Protein Folding From Evolutionary Embeddings. In *arXiv preprint 2025*. [\[link\]](#)
- **Chenyu Wang***, Cai Zhou*, Sharut Gupta, Zongyu Lin, Stefanie Jegelka, Stephen Bates, Tommi Jaakkola. Learning Diffusion Models with Flexible Representation Guidance. In *Advances in Neural Information Processing Systems, NeurIPS 2025*. (Also **Oral** at ICML 2025 FM4LS workshop.) [\[link\]](#)
- Cai Zhou*, **Chenyu Wang***, Dinghuai Zhang*, Shangyuan Tong, Yifei Wang, Stephen Bates, Tommi Jaakkola. Next Semantic Scale Prediction via Hierarchical Diffusion Language Models. In *Advances in Neural Information Processing Systems, NeurIPS 2025*. [\[link\]](#)
- Hanqun Cao*, Haosen Shi*, **Chenyu Wang**, Sinno Jialin Pan, Pheng-Ann Heng. GLID²E: A Gradient-Free Lightweight Fine-tune Approach for Discrete Sequence Design. In *Advances in Neural Information Processing Systems, NeurIPS 2025*. [\[link\]](#)
- 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 *Advances in Neural Information Processing Systems, NeurIPS 2025*. [\[link\]](#)

- Hanchen Wang*, Yichun He*, Paula P Coelho*, Matthew Bucci*, ..., **Chenyu Wang**, ..., Aviv Regev. SpatialAgent: An Autonomous AI Agent for Spatial Biology. In *bioRxiv preprint 2025*. [\[link\]](#)
- Yuhui Zhang*, Yuchang Su*, **Chenyu Wang**, Tianhong Li, Zoe Wefers, Jeffrey Nirschl, James Burgess, Daisy Ding, Alejandro Lozano, Emma Lundberg, Serena Yeung-Levy. CellFlux: Simulating Cellular Morphology Changes via Flow Matching. In *International Conference on Machine Learning, ICML 2025*. [\[link\]](#)
- Masatoshi Uehara, Yulai Zhao, **Chenyu Wang**, Xiner Li, Aviv Regev, Sergey Levine, Tommaso Biancalani. Inference-Time Alignment in Diffusion Models with Reward-Guided Generation: Tutorial and Review. In *arXiv preprint 2025*. [\[link\]](#)
- Yuhui Zhang*, Yuchang Su*, Yiming Liu, Xiaohan Wang, James Burgess, Elaine Sui, **Chenyu Wang**, Josiah Aklilu, Alejandro Lozano, Anjiang Wei, Ludwig Schmidt, Serena Yeung-Levy. Automated Generation of Challenging Multiple-Choice Questions for Vision Language Model Evaluation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2025*. [\[link\]](#)
- **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 *International Conference on Learning Representations, ICLR 2025*. [\[link\]](#)
- **Chenyu Wang***, Sharut Gupta*, Xinyi Zhang, Sana Tonekaboni, Stefanie Jegelka, Tommi Jaakkola, Caroline Uhler. An Information Criterion for Controlled Disentanglement of Multimodal Data. In *International Conference on Learning Representations, ICLR 2025*. (Also **Oral** and **Honorable Mention Award** at NeurIPS 2024 UniReps workshop.) [\[link\]](#)
- 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*. (Also **Oral** at NeurIPS 2024 SSL workshop.) [\[link\]](#) [\[MIT CSAIL News\]](#)
- 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*. (Also **Oral** at ICLR 2024 MLGenX workshop.) [\[link\]](#)
- **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

Research Intern at Meta FAIR

Advised by Yuandong Tian and Bo Liu

Menlo Park, CA

May 2025-Aug. 2025

Research Intern at Genentech

Advised by Aviv Regev, Hanchen Wang, and Masatoshi Uehara

South San Francisco, CA

May 2024-Aug. 2024

Research Intern at Princeton University

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

Princeton, NJ

Jun. 2021-Dec. 2021

Research Intern at USC

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

Los Angeles, CA

Jan. 2021-Jun. 2021

Research Assistant at Tsinghua University

Beijing, China

HONORS & AWARDS

- Citadel GQS PhD Fellowship (the only recipient in EECS), 2025
- D. E. Shaw Research Doctoral Fellowship, 2025
- Honorable Mention Award in NeurIPS 2024 UniReps Workshop, 2024
- 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

Meta	Menlo Park
<i>Research Intern in Meta FAIR</i>	May 2025-Aug. 2025
Genentech	South San Francisco
<i>Research Intern in Dr. Aviv Regev’s Lab</i>	May 2024-Aug. 2024
Jane Street Asia Limited	Hong Kong
<i>Quantitative Trading Intern (Return offer extended)</i>	Jun. 2021-Sept. 2021
WizardQuant Capital Management	Zhuhai, China
<i>Quantitative Research Intern, Quantitative Research Department</i>	Jun. 2020-Aug. 2020
Techsharpe Quant Capital Management	Beijing, China
<i>Data Analyst Intern, Trading Department</i>	Jan. 2020-Feb. 2020

SERVICES

- Reviewer: ICML 2025, ICLR 2025/2026, NeurIPS 2024/2025, PLOS Computational Biology
- Teaching Assistant: MIT [6.8300](#) Advances in Computer Vision
- Volunteer: [WiDS](#) Cambridge Datathon student volunteer

INVITED TALKS

• MIT CSB Seminar	2025
• MIT Machine Learning in Biology Working Group (MLBWG)	2025
• NeurIPS 2024 UniReps Workshop	2024
• Broad Institute of MIT and Harvard Models, Inference & Algorithms (MIA) Post Doc Lightning Talks	2024
• MIT ML Tea Talk	2024
• MIT LIDS Student Conference	2024
• Valence Labs Molecular Modeling & Drug Discovery (M2D2) Reading Group	2024
• Temporal Graph Reading Group	2023

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)