Marianne Arriola

New York, NY | marriola@cs.cornell.edu | m-arriola.com | GitHub

Interests: Deep Learning, Generative AI for text, Discrete Diffusion

Education

Ph.D. student in Computer Science, Cornell University — Sept 2023–Jun 2027
Advisor: Volodymyr Kuleshov

B.S., Computer Science, University of California, Santa Barbara — Aug 2019–Jun 2023

Current Research

Diffusion Language Models
Committee: Volodymyr Kuleshov (Chair), Mohamed Abdelfattah, Fei Wang

- I explore diffusion LMs which enable faster & more controllable generation with better benchmark performance for math & planning compared to traditional LLMs
- I design novel parameterizations and architectures to improve model quality [1,2,3,4], training efficiency [1,3,4], and inference speed [1,3]

Publications

Selected Papers

[1] Marianne Arriola, Aaron Gokalsan, Justin Chiu, Zhihan Yang, Zhixuan Qi, Jiaqi Han, Subham Sahoo, Volodymyr Kuleshov. "Block Diffusion: Interpolating between Autoregressive and Diffusion Language Models." *ICLR 2025.* **Oral presentation (Top 1.77%)**.

[2] Subham Sahoo, Marianne Arriola, Yair Schiff, Aaron Gokaslan, Edgar Marroquin, Justin Chiu, Alexander Rush, Volodymyr Kuleshov "Simple and Effective Masked Diffusion Language Models." *NeurIPS 2024.*

[3] Marianne Arriola*, Yair Schiff*, Hao Phung, Aaron Gokaslan, Volodymyr Kuleshov "Encoder-Decoder Block Diffusion Language Models for Efficient Training and Inference." *NeurIPS 2025.*

---

[4] Marianne Arriola, Naveen Venkat, Jon Granskog, Anastasis Germanidis. "Adapting Autoregressive Vision Language Models for Parallel Diffusion Decoding." *Runway Research Blog.*

[5] Yair Schiff, Omer Belhasin, Roy Uziel, Guanghan Wang, Marianne Arriola, Gilad Turok, Michael Elad, Volodymyr Kuleshov. "Learn from Your Mistakes: Self-Correcting Masked Diffusion Models." *arXiv preprint* (2026).

[6] Guanghan Wang, Gilad Turok, Yair Schiff, Marianne Arriola, Volodymyr Kuleshov. "d2: Improved Techniques for Training Reasoning Diffusion Language Models." *arXiv preprint* (2025).

[7] Marianne Arriola, Weishen Pan, Manqi Zhou, Qiannan Zhang, Chang Su, Fei Wang. "Joint Analysis of Single-Cell Data across Cohorts with Missing Modalities." *arXiv* (Feb 2024).

[8] Marianne Arriola & Kadina Johnston "Identifying Optimal Proteins by Their Structure Using Graph Neural Networks." *Caltech URJ* (Jun 2022).

Employment

NVIDIA, Research Intern. Santa Clara, CA — Feb 2026–Aug 2026
Runway AI, Research Intern. New York, NY — Jun 2025–Sept 2025
MIT CSAIL, Research Intern. Cambridge, MA — Jun 2022–Nov 2022
Caltech, Research Intern. Caltech, Pasadena, CA — Jun 2021–Aug 2021

Open-Source Contributions

Block Diffusion Language Models (GitHub). 960 stars.

- Led development of a diffusion LM with arbitrary-length generation and KV caching

Encoder-Decoder Diffusion Language Models (GitHub). 35 stars.

- Co-led development of a diffusion LM with an encoder-decoder architecture for faster inference

Projects

Runway AI, New York, NY — Jun 2025 – Sept 2025

- Developed a distillation framework to adapt a 7B autoregressive vision-language model for parallel diffusion decoding while maintaining benchmark performance
- Built an end-to-end pipeline for data creation, model adaptation, and benchmarking, and shared the work in a technical blog post

MIT CSAIL, Cambridge, MA — Jun 2022 – Nov 2022
PI: Justin Solomon

- Designed a memory-efficient point cloud representation using geometric primitives
- Developed a graph neural network for hybrid point clouds that matches state-of-the-art segmentation performance while using 50% less memory

Caltech, Pasadena, CA — Jun 2021 – Aug 2021
PI: Frances Arnold

- Proposed a data-driven method to iteratively refine existing protein structures toward desired properties (e.g., substrate specificity)
- Built a graph neural network to predict protein function from structural graph representations

Patents

Xiangru Huang, Marianne Arriola, Yue Wang, Vitor Campagnolo Guizilini, Rares Andrei Ambrus, Justin Solomon. "Hybrid Geometric Primitive Representations for Point Clouds." U.S. patent pending.

Selected Talks

Meta Fundamental AI Research (FAIR) — Feb 2026
International Conference on Learning Representations (ICLR) — Apr 2025
Amazon Artificial General Intelligence (AGI) — Apr 2025

Awards

Top Reviewer, NeurIPS 2025	Oct 2025
NSF Graduate Research Fellowship	Mar 2023–Jun 2028
Bowers CIS Dean's Excellence Fellowship, Cornell University	Mar 2023–Jun 2029

Skills: Python, PyTorch, C++, MATLAB, Bash