

# Yuhao Huang

✉ Address: 72 S Central Campus Dr, Salt Lake City, UT 84112  
☎ Phone: +1 (812) 508 1849 / +86 15298359912 (WeChat)  
✉ Email: u1430219@utah.edu / m15298359912@gmail.com  
🔗 LinkedIn Profile



I am a Ph.D. student in Machine Learning @University of Utah. My research focuses on **generative models, graph learning, and large language models**, with publications in top-tier conferences including **ICLR, ICML and CVPR**. I also have research and industry experience in AI and quantitative analysis for finance.

## Education




- 
- |   |                     |
|---|---------------------|
| 🏠 <b>University of Utah, UT, USA</b>  | Aug 2022 – Present  |
| Ph.D. in Applied Mathematics (with a specialization in Machine Learning & Data Science)<br>GPA: 3.93/4.00 |                     |
| 🏠 <b>Northwestern University, IL, USA</b>   | Sep 2019 – Jan 2021 |
| Master in Applied Mathematics (Machine Learning & Data Science Track)<br>GPA: 3.78/4.00                   |                     |
| 🏠 <b>Hohai University, Jiangsu, China</b>   | Sep 2015 – Jun 2019 |
| Bachelor in Computing Science<br>GPA: 87/100  |                     |

## Research Interest



(The **red square** indicates one of my referenced publications or preprints.)



### Deep Generative Model

- Improving diffusion (score-based) and flow-based models, with a focus on  theoretical analysis [1] from differential equations point of view,  model- and algorithm-level efficiency optimization [4, 7] and  applications to image [4, 5], spatiotemporal data and video [1], and scientific data [7] sampling.

### Numerical and Stochastic Method with Applications to Deep Generative Models

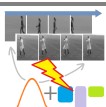
- Improving stochastic algorithms for  nonconvex optimization [2]  more efficient training for diffusion (score-based) models [4], image restoration tasks incorporating flow/diffusion models [5].



### Graph Learning & Large Language Model

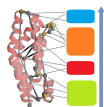
- Geometric learning integrated with LLM / State Space Model (Mamba)[3, 6].

## Selected Research Projects



### Efficient Diffusion/Flow Matching with Guidance for Spatiotemporal Data (e.g. Video, Dynamical System) Generation

- Developed on-step sampling diffusion or flow model for sampling the spatiotemporal Data, e.g. video, dynamical system data with previous states or specified information as guidance. Significantly reduced the time complexity from  $\mathcal{O}(LN)$  to  $\mathcal{O}(L)$ .
- Designed a U-ViT-based architecture to parameterize trajectory integration over vector fields, tailored for spatiotemporal data such as videos and latent variables applications. Designed a deep equilibrium-based model for mean vector field regressor, significantly enhancing memory efficiency.
- Developed an auto-regressive sampling method that uses the previous state as guidance for simulating dynamical system data, enhancing memory efficiency.
- Developed Jacobian regularization framework for robustness and training acceleration.



## Representation Learning for Protein Sequential Data with LLM, State Space Model & Local Geometric Learning

- Developed GNN-based local geometric learning models and applied a protein fragment algorithm to tokenize protein sequences into a hidden sequential tensor representation.
- Applied hidden sequential tensor representations derived from GNN-based tokenization to fine-tune large-scale protein language models such as ESM-2 or Mamba for protein classification and property prediction.

## Industry Experience

---



### Machine Learning & AI Research

- **Argonne National Laboratory**, Chicago Area, IL May 2024 – Aug 2024  
*Research Intern*,  
✍️ Worked on flow matching/normalizing generative models in scientific applications.
- **Snips Media**, Chicago, IL Jun 2020 – Aug 2020  
*Machine Learning Engineer Intern*,  
✍️ Worked on GANs for data augmentation and YOLO for object detection.



### Quantitative Research for the Financial Industrial

- **AQUMON Digital Wealth Management**, Hong Kong SAR Nov 2021 – Jul 2022  
*Quantitative Researcher Intern*,  
✍️ Developed ML models (ARIMA, LSTM, CNN, Attention, Transformer) for time series forecasting.
- **SGD Asset Management**, Shenzhen, China Feb 2021 – Jul 2021  
*Quantitative Researcher Intern*,  
✍️ Developed ETF arbitrage models integrated with machine learning techniques for tick-level time series data forecasting.

## Publications and Preprints

---

(\*: equal contribution)

1. **ICML2025** **Yuhao Huang**, Taos Transue, Shih-Hsin Wang, William M Feldman, Hong Zhang, Bao Wang. “Improving Flow Matching by Aligning Flow Divergence”. Proceedings of the 42nd International Conference on Machine Learning, 2025.
2. **CVPR2025** Tao Sun, **Yuhao Huang**, Li Shen, Kele Xu, Bao Wang. “Investigating the Role of Weight Decay in Enhancing Nonconvex SGD”. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2025.
3. **ICLR2025, Oral 1.8%** **Yuhao Huang\***, Shih-Hsin Wang\*, Justin M. Baker, Yuan-En Sun, Qi Tang, and Bao Wang. “A Theoretically-Principled Sparse, Connected, and Rigid Graph Representation of Molecules”. The 13th International Conference on Learning Representations, 2025.
4. **ICLR2024** **Yuhao Huang**, Qingsong Wang, Akwum Onwunta, Bao Wang. ”Efficient Score Matching with Deep Equilibrium Layers”. The 12th International Conference on Learning Representations, 2024.
5. Under review Fan Jia, **Yuhao Huang**, Shih-Hsin Wang, Bao Wang. “Plug-and-Play Image Restoration with Flow Matching: A Continuous Viewpoint”. Under review at NeurIPS, 2025
6. Under review Shih-Hsin Wang, **Yuhao Huang**, Taos Transue, Justin M. Baker, Jonathan Forstater, Thomas Strohmmer, Bao Wang. “Towards Multiscale Graph-based Protein Learning with Geometric Secondary Structural Motifs”. Under review at NeurIPS, 2025.
7. Under review **Yuhao Huang**, Justin Baker, Shih-Hsin Wang, Massimiliano Lupo Pasini, Andrea L. Bertozzi, Bao Wang. “A Regularized Training of E(n)-Equivariant Graph Neural Network-assisted Generative Models”.
8. Preprint **Yuhao Huang**, David Chopp. “Fast Iterative Algorithm for Eikonal Equation and Applications”. arxiv.2106.15869

## Skills

---

- **Programming Languages:** Python, C, CUDA Programming, Linux
- **Scientific Computing Tools:** Matlab, PETSc
- **Frameworks:** PyTorch, Jax, Tensorflow
- **Libraries:** Pytorch-diffeq, Pytorch-geometric, Matplotlib, Pandas, Numpy

## Teaching & Service Experience

---

- **Reviewer:** ICLR 2025, ICML 2025
- **Volunteer:** ICLR 2025
- **Teaching:** MATH 2210, University of Utah

## Invited Conference & Presentation

---

- ICLR 2025 Oral Presentaion: “A Theoretically-Principled Sparse, Connected, and Rigid Graph Representation of Molecules”.
- Inaugural CAMDA (Center for Approximation and Mathematical Data Analytics at Texas A&M University) Conference 2023.

