



Ph.D. DISSERTATION DEFENSE

| | |
|---------------------------|---|
| Candidate: | Jiali Cui |
| Degree: | Doctor of Philosophy |
| School/Department: | Charles V. Schaefer Jr. School of Engineering and Science, Department of Computer Science. |
| Date: | Monday, March 9th, 2026 |
| Time/Location: | 1:00 p.m. Gateway South (GS) 024. |
| Title: | Modelling and Learning for Hierarchical Generative Model and Energy-Based Model |
| Chairperson: | Dr. Tian Han, Department of Computer Science |
| Committee Members: | Dr. Philippos Mordohai, Department of Computer Science Dr. Zhuo Feng, Department of Electrical and Computer Engineering Dr. Zining Zhu, Department of Computer Science |

ABSTRACT

Real-world data exhibit hierarchical and compositional structure, where high-level semantics interact with fine-grained details. A central goal of modern generative modeling is to build models that capture this structure, not only synthesizing high-quality samples but also learning meaningful representations across multi-level abstraction. Hierarchical generative model and energy-based model (EBM) are two fundamental generative models factorized by distinct formulations, featuring unique properties and driving active research exploration.

This dissertation first introduces the foundations of hierarchical generative models and energy-based models, including their formulations, learning methods, and practical limitations. We then present unified modeling and learning principle that focus on hierarchical latent structures and expressive energy-based models to address these limitations. The dissertation further extends these ideas from single-modality settings to multimodal settings, with the goal of learning shared abstractions and enabling coherent generation across heterogeneous data types.

Specifically, we develop: (i) A unified framework incorporating expressive EBM priors to address the limitations of non-informative Gaussian priors in hierarchical generative models [1]. (ii) A novel design leveraging both architectural hierarchical structure and EBM priors for learning effective hierarchical data representations [2]. (iii) innovative methods integrating EBMs prior within diffusion probabilistic and multi-stage contrastive learning scheme to enhance EBM learning [3, 4]. (iv) A general learning method that learns EBM and generator model in a cooperative fashion [5]. (v) probabilistic frameworks learning expressive prior for shared latent representations across multimodalities [6, 7]. (vi) a hierarchical multimodal framework that enables structured representation learning and coherent cross modal generation. [8].

References

- [1] **Jiali Cui**, Ying Nian Wu, and Tian Han. "Learning joint latent space ebm prior model for multi-layer generator."
Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2023.
- [2] **Jiali Cui**, Ying Nian Wu, and Tian Han. "Learning hierarchical features with joint latent space energy-based prior."
Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). 2023.
- [3] **Jiali Cui**, and Tian Han. "Learning latent space hierarchical EBM diffusion models."
The International Conference on Machine Learning (ICML). 2024.
- [4] **Jiali Cui**, and Tian Han. "Learning adaptive multi-stage energy-based prior for hierarchical generative model."
The Transactions on Machine Learning Research (TMLR). 2026.
- [5] **Jiali Cui**, and Tian Han. "Learning energy-based model via dual-MCMC teaching."
Advances in Neural Information Processing (NeurIPS). 2023.
- [6] Shiyu Yuan, **Jiali Cui**, Hanao Li, and Tian Han. "Learning multimodal latent generative models with energy-based prior."
The European Conference on Computer Vision (ECCV). 2024.
- [7] **Jiali Cui**, Yanying Chen, Yanxia Zhang, and Matthew Klenk. "ShaLa: multimodal shared latent generative modelling"
Association for the Advancement of Artificial Intelligence (AAAI). 2026.
- [8] **Jiali Cui**, Yanying Chen, Tingwei Liu, and Tian Han. "Learning hierarchical multimodal latent generative model."
Under review.