Hoang-Son Nguyen

sonngh.00@gmail.com | [hsnguyen24.github.io] | [Google Scholar]

SUMMARY

Equipped with research experience in deep representation learning and graph machine learning, I aim to develop generative models that can capture real-world dynamics so as to robustly generalize to unseen real scenarios. These world models serve as reliable and interpretable backbones for agentic decision making and generating novel artifacts to advance scientific frontiers. I also aim to explore how principled generative modeling can contribute to computer vision and biomedicine.

EDUCATION

Master of Science in Artificial Intelligence — *Sep. 2024 - (Expected) June 2026*
Oregon State University — Current GPA : 3.95/4.0

Bachelor of Engineering in Artificial Intelligence — *Sep. 2019 - Mar. 2024*
The Chinese University of Hong Kong (CUHK) — First Class Honours

PUBLICATIONS & PREPRINTS

1. Diverse Influence Component Analysis: A Geometric Approach to Nonlinear Mixture Identifiability,
Hoang-Son Nguyen, Xiao Fu,
Advanced in Neural Information Processing Systems (NeuRIPS), 2025. [PDF]

2. Learning Graphs from Smooth Signals under Partial Observations: A Robustness Analysis, [PDF]
Hoang-Son Nguyen, Hoi-To Wai,
(Under Review) *International Conference on Acoustics, Speech and Signal Processing* (ICASSP), 2026.

3. On Detecting Low-Pass Graph Signals under Partial Observations,
Hoang-Son Nguyen, Hoi-To Wai,
IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM), 2024. [PDF]

4. On the Stability of Low Pass Graph Filter with a Large Number of Edge Rewires,
Hoang-Son Nguyen, Hoi-To Wai,
International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2022. [PDF]

HONORS & AWARDS (SELECTED)

Best Student Paper Award, IEEE SAM — *July 2024*
For the best student works at IEEE Sensor Array and Multichannel Signal Processing Workshop.

Charles K. Kao Research Scholarship — *Mar. 2023*
For outstanding achievements in undergraduate research at CUHK.

Undergraduate Admission Scholarship — *Sep. 2019*
Full-ride scholarship for the entire undergraduate study at CUHK.

RESEARCH EXPERIENCE

Oregon State University — *Sep. 2024 – Present*
Graduate Research Assistant (Advisor: Xiao Fu)

- Proposed Diverse Influence Component Analysis, a representation learning scheme for disentangling factors of variation; proved rigorous identifiability results; applied in **single-cell data analysis** and **unsupervised concept discovery** from images; published at **NeuRIPS 2025** [1].
- Constructed a **noise-robust identifiability analysis** of volume-minimizing NMF under anchor features; ongoing effort to generalize the result to sufficiently scattered condition-based NMF.
- (Ongoing) Developing a principle for structured generative model for content-style disentanglement from multidomain paired/unpaired data; with applications in long-tail generation.

The Chinese University of Hong Kong — *Sep. 2021 – Sep. 2024*
Undergraduate Research (Advisor: Hoi-To Wai)

- Proved that graph topology learning methods are robust against missing node data, if the node features are smooth with graph; applied to **learn subgraphs of social networks** and **sensor networks** from partially observed signals; submitted to **ICASSP 2026** [2] and presented at **Graph Signal Processing Workshop 2025**.
- Proposed a method to detect smoothness level of signals across graphs by only observing a subset of nodes; applied in **robustification of community detection**; published at **SAM 2024**, with **best student paper award** [3].
- Analyzed **stability of graph filters (building blocks of GNN)**, under large graph topology perturbation; proved that stability follows if underlying graph model is not significantly perturbed; published in **ICASSP 2022** [4].

WORKSHOPS & PRESENTATIONS

Graph Topology Learning with Smooth Signals under Partial Observations *May 2025*
Graph Signal Processing Workshop, Montreal, Canada.

Graph Learning with Low-pass Graph Signal Processing *Sep. 2024*
Faculty of Data Science & AI at National Economics University, Hanoi, Vietnam.

MISCELLANEOUS

Research Keywords: Generative Modeling, Deep Representation Learning, Self-supervised & Unsupervised Learning, Graph Machine Learning, Graph Signal Processing, Optimization, Matrix/Tensor Decomposition.

Coursework (selected): Optimization, Big Data Analytics, Tensor Methods, Online Learning, Machine Learning Theory, Information Theory, Simulation, Approximation Theory, Functional Analysis, Computer Systems, Stochastic Models, Algorithms, Linear Systems and Control, Time Series.

Programming: Python, C/C++, MATLAB, PyTorch, Git, Linux, Hadoop/Spark, LaTeX.

Languages: Vietnamese (native), English (fluent).

Reviewer: Causality and Large Models @ NeuRIPS (2024), IEEE ICASSP (2025), IEEE Transactions on Signal Processing (2025).

Volunteer: Delivery and Warehouse Assistant, Feeding Hong Kong (2023).

SIDE PROJECTS (AND SOME LESSONS LEARNED)

Cluster-then-Bandit [GitHub]

- Designed a new linear bandit algorithm that utilizes the hidden graph structure underlying rewards of the arms, by estimating a graph-based regularizer for subsequent LinUCB procedure.
- Empirically demonstrated superior regret performance over vanilla linear bandit algorithms; received an A from my instructor (Dr. Huazheng Wang at Oregon State University).

K-Means with MapReduce [GitHub]

- Implemented from scratch a distributed K-means clustering pipeline for parallelly processing and clustering MNIST dataset, using MapReduce approach on a Apache Hadoop cluster hosted on Amazon AWS.
- Achieved 44% accuracy while merely using K-means algorithm running on distributed machines with only CPU. Learned how to stay calm and debug a small distributed data processing system.

Atomization Energy Prediction [GitHub]

- Benchmarked classical machine learning and graph neural network in predicting atomization energy on QM7 molecular dataset in quantum chemistry.
- Learned the hard lesson that simple classical algorithm like Gradient Boosting sometimes works better (while being faster) than neural architectures with complicated inductive biases like CNN or GNN; and a data-centric approach focusing on careful pre-processing is sometimes more important than algorithmic development.

Pacman with SDL [GitHub]

- Designed and implemented from-scratch a solo project on Pacman game using Simple DirectMedia Layer, using various classical search algorithms such as A-star and UCS.
- Learned how to structure a small software engineering project and how to write non-trivial Git commit message.
- My friends said they really enjoyed the game.