

# Tan Minh Nguyen

---

## CONTACT INFORMATION

Department of Mathematics  
National University of Singapore  
10 Lower Kent Ridge Road  
Singapore, 119076

Phone: +65 8308 2743  
E-mail: [tanmn@nus.edu.sg](mailto:tanmn@nus.edu.sg)  
<https://tanmnguyen89.github.io/>

## ACADEMIC APPOINTMENTS

**National University of Singapore**, Singapore 2023–present

- Assistant Professor (tenure-track), Department of Mathematics

**University of California, Los Angeles**, CA, USA 2020–present

- Post-doctoral Scholar in Mathematics
- Mentor: Professor Stanley J. Osher

## EDUCATION

**Rice University**, Houston, Texas, USA

Ph.D. in Electrical and Computer Engineering, 2014–2020

- Advisor: Professor Richard G. Baraniuk
- Thesis: Momentum-based Methods for Training and Designing Deep Neural Networks

M.S. in Electrical and Computer Engineering, 2014–2016

- Advisor: Professor Richard G. Baraniuk
- Thesis: Bridging Theory and Practice in Deep Learning with the Deep Rendering Model

B.S. in Electrical and Computer Engineering, 2011–2014

## RESEARCH INTERESTS

My research focuses on the interplay of the interpretability, robustness, and efficiency of machine learning models from three principled approaches:

- Optimization (primal-dual frameworks for deep learning models, momentum-based neural networks, fast multipole transformers)
- Differential equation (Nesterov neural ordinary differential equations, graph neural diffusion)
- Statistical modeling (mixture and nonparametric kernel regression frameworks for transformers, deep generative models)

## JOURNAL PUBLICATIONS

**Tan M. Nguyen\***, Nhat Ho\*, Ankit B. Patel, Anima Anandkumar, Michael I. Jordan, Richard G. Baraniuk. “A Bayesian Perspective of Convolutional Neural Networks through a Deconvolutional Generative Model”. *Journal of Machine Learning Research*, 2023.

Bao Wang\*, **Tan M. Nguyen\***, Andrea L. Bertozzi, Richard G. Baraniuk, Stanley J. Osher. “Scheduled Restart Momentum for Accelerated Stochastic Gradient Descent”. *SIAM Journal on Imaging Sciences*, 2022.

Bao Wang, Hedi Xia, **Tan M. Nguyen**, Stanley J. Osher. “How Does Momentum Benefit Deep Neural Networks Architecture Design? A Few Case Studies”. *Research in the Mathematical Sciences*, 2022.

Yue Wang, Jianghao Shen, Ting-Kuei Hu, Pengfei Xu, **Tan M. Nguyen**, Richard Baraniuk, Zhangyang Wang, Yingyan Lin. “Dual Dynamic Inference: Enabling More Efficient, Adaptive, and Controllable Deep Inference”. *IEEE Journal of Selected Topics in Signal Processing*, 2020.

---

\*: co-first author

Khang Nguyen, Hieu Nong, Vinh Nguyen, Nhat Ho, Stanley J. Osher, **Tan M. Nguyen**. “Revisiting Over-smoothing and Over-squashing using Ollivier-Ricci Curvature”. *International Conference on Machine Learning (ICML), 2023 (journal-quality)*.

Hien Dang, Tho Tran Huu, Stanley Osher, Hung Tran-The, Nhat Ho, **Tan M. Nguyen**. “Neural Collapse in Deep Linear Networks: From Balanced to Imbalanced Data”. *International Conference on Machine Learning (ICML), 2023 (journal-quality)*.

**Tan M. Nguyen\***, Tam Nguyen\*, Nhat Ho, Andrea Bertozzi, Richard G. Baraniuk, Stanley J. Osher. “A Primal-Dual Framework for Transformers and Neural Networks”. *International Conference on Learning Representations (ICLR), 2023 (journal-quality, notable-top-25%)*.

Khai Nguyen, Tongzheng Ren, Huy Nguyen, Litu Rout, **Tan M. Nguyen**, Nhat Ho. “Hierarchical Sliced Wasserstein Distance”. *International Conference on Learning Representations (ICLR), 2023 (journal-quality)*.

**Tan M. Nguyen\***, Tam Nguyen\*, Long Bui, Hai Do, Dung Le, Hung Tran-The, Khuong Nguyen, Richard G. Baraniuk, Nhat Ho, Stanley J. Osher. “A Probabilistic Framework for Pruning Transformers via a Finite Admixture of Keys”. *International Conference on Acoustics, Speech, and Signal Processing (ICASSP, notable-top-3%), 2023*.

**Tan M. Nguyen\***, Minh Pham\*, Tam Nguyen, Khai Nguyen, Stanley J. Osher, Nhat Ho. “FourierFormer: Transformer Meets Generalized Fourier Integral Theorem”. *Conference on Neural Information Processing Systems (NeurIPS), 2022 (journal-quality, acceptance rate: 25.6%)*.

**Tan M. Nguyen\***, Tam Nguyen\*, Hai Do, Khai Nguyen, Vishwanath Saragadam, Minh Pham, Khuong Nguyen, Nhat Ho, Stanley J. Osher. “Improving Transformer with an Admixture of Attention Heads”. *Conference on Neural Information Processing Systems (NeurIPS), 2022 (journal-quality, acceptance rate: 25.6%)*.

Nghia Nguyen\*, **Tan M. Nguyen\***, Huyen Vo, Stanley J. Osher, Thieu Vo. “Improving Neural Ordinary Differential Equations with Nesterov’s Accelerated Gradient Method”. *Conference on Neural Information Processing Systems (NeurIPS), 2022 (journal-quality, acceptance rate: 25.6%)*.

Tam Nguyen\*, **Tan M. Nguyen\***, Dung Le, Khuong Nguyen, Anh Tran, Richard G. Baraniuk, Nhat Ho, Stanley J. Osher. “Improving Transformers with Probabilistic Attention Keys”. *International Conference on Machine Learning (ICML), 2022 (journal-quality, acceptance rate: 21.9%)*.

Matthew Thorpe\*, **Tan M. Nguyen\***, Hedi Xia\*, Thomas Strohmmer, Andrea Bertozzi, Stanley J. Osher, Bao Wang. “GRAND++: Graph Neural Diffusion with a Source Term”. *International Conference on Learning Representations (ICLR), 2022 (journal-quality, acceptance rate: 32.9%)*.

**Tan M. Nguyen**, Richard G. Baraniuk, Mike Kirby, Stanley J. Osher, Bao Wang. “Momentum Transformer: Closing the Performance Gap Between Self-attention and Its Linearization”. *Mathematical and Scientific Machine Learning (MSML), 2022*.

**Tan M. Nguyen**, Animesh Garg, Richard G Baraniuk, Anima Anandkumar. “InfoCNF: An Efficient Conditional Continuous Normalizing Flow with Adaptive Solvers”. *Asilomar Conference, 2022*.

**Tan M. Nguyen**, Vai Suliafu, Stanley J. Osher, Long Chen, Bao Wang. “FMMformer: Efficient and Flexible Transformer via Decomposed Near-field and Far-field Attention”. *Conference on Neural Information Processing Systems (NeurIPS), 2021 (journal-quality, acceptance rate: 25.7%)*.

Hedi Xia, Vai Suliafu, Hangjie Ji, **Tan M. Nguyen**, Andrea Bertozzi, Stanley J. Osher, Bao Wang. “Heavy Ball Neural Ordinary Differential Equations”. *Conference on Neural Information Processing Systems (NeurIPS)*, 2021 (journal-quality, acceptance rate: 25.7%).

**Tan M. Nguyen**, Richard G. Baraniuk, Andrea Bertozzi, Stanley J. Osher, Bao Wang. “MomentumRNN: Integrating Momentum into Recurrent Neural Networks”. *Conference on Neural Information Processing Systems (NeurIPS)*, 2020 (journal-quality, acceptance rate: 20.1%).

Yujia Huang, James Gornet, Sihui Dai, Zhiding Yu, **Tan M. Nguyen**, Doris Tsao, Anima Anandkumar. “Neural Networks with Recurrent Generative Feedback”. *Conference on Neural Information Processing Systems (NeurIPS)*, 2020 (journal-quality, acceptance rate: 20.1%).

**Tan M. Nguyen\***, Nhat Ho\*, Ankit B. Patel, Anima Anandkumar, Michael I. Jordan, Richard G. Baraniuk. “Neural Rendering Model: Joint Generation and Prediction for Semi-Supervised Learning”. *Deep Math Conference (DeepMath)*, 2019. (Oral presentation)

**Tan M. Nguyen**, Wanjia Liu, Fabian Sinz, Richard G. Baraniuk, Andreas S. Tolias, Xaq Pitkow, Ankit B. Patel. “Towards a Cortically Inspired Deep Learning Model: Semi-Supervised Learning, Divisive Normalization, and Synaptic Pruning”. *Conference on Cognitive Computational Neuroscience (CCN)*, 2017.

Ankit B Patel, **Tan M. Nguyen**, Richard Baraniuk. “A Probabilistic Framework for Deep Learning”. *Conference on Neural Information Processing Systems (NeurIPS)*, 2016 (journal-quality, acceptance rate: 23.6%).

#### WORKSHOP PAPERS

Gavin D. Portwood, Peetak P. Mitra, Mateus Dias Ribeiro, **Tan M. Nguyen**, Balasubramanya T. Nadiga, Juan A. Saenz, Michael Chertkov, Animesh Garg, Anima Anandkumar, Andreas Dengel, Richard G. Baraniuk, David P. Schmidt. “Turbulence Forecasting via Neural ODE”. *NeurIPS Workshop on Machine Learning and the Physical Sciences*, 2019.

#### WORKSHOP ORGANIZATION

*Integration of Deep Neural Models and Differential Equations* at the International Conference on Learning Representations (ICLR), 2020.

*Integration of Deep Learning Theories* at the Conference on Neural Information Processing Systems (NeurIPS), 2018.

#### AWARDED GRANTS

PI, Toyota Research Institute Grant, *A Unified Framework for Building Large-Scale Language Models in Developing Countries: From Model and System to Data*, \$100,000, November, 2022 - November, 2023.

#### SUBMITTED GRANT PROPOSALS

Co-PI, Singapore-Israel Scientific Research Program, *Mitigating Security and Privacy Risks of Large Language Models*, SGD 250,000, October, 2023 - March, 2025.

#### HONORS AND AWARDS

*Computing Innovation Postdoctoral Fellowship (CIFellows)*, 2020–2023.

*AWS Cloud Credits for Research*, \$50,000, September, 2018.

*NSF Graduate Research Fellowship*, 2016–2020.

*Ford Foundation Fellowship Honorable Mention*, 2016.

*Neuroengineering IGERT: From Cells to Systems Fellowship*, 2015 – 2017.

*Texas Instruments Fellowship*, 2014–2016. (Awarded to top incoming graduate students in the ECE department at Rice University)

*Louis J. Walsh Scholarship in Engineering*, 2013–2014. (Awarded to top undergraduate students in the ECE department at Rice University)

*Gold Scholar on Coca-Cola’s Community College Academic Team*, 2011.

TEACHING

Spring 2023 and Summer 2022, Math 156 Machine Learning, UCLA.

Fall 2021, Probability & Statistics Course, FPT Software AI Residency, Vietnam.

INVITED SEMINAR  
PRESENTATIONS

A Statistical Treatment of the Attention Mechanism in Transformers. *Analytics Center of Excellence (ACOE), IQVIA, Italy, 2022.*

Principled Models for Machine Learning. *Math Machine Learning Seminar, the Max Planck Institute for Mathematics in the Sciences and UCLA, USA, 2022.*

Principled Models for Machine Learning. *Applied Math Colloquium at UCLA, USA, 2022.*

Momentum-based Methods for Training and Designing Deep Neural Networks. *VinAI Research, Vietnam, 2020.* (Invited talk)

Scheduled Restart Momentum for Accelerated Stochastic Gradient Descent. *Machine Learning Seminar, Rice University, USA, 2020.* (Invited talk)

Deep Generative Models for Geophysical Signal Disentanglement. *Geo-Mathematical Imaging Group (GMIG) Project Review Meeting, Rice University, USA, 2018.* (Invited talk)

CONFERENCE,  
WORKSHOP  
PRESENTATIONS

A Primal-Dual Framework for Transformers and Neural Networks. *Special Session on “Mathematics of Machine Learning”, Canadian Mathematical Society Winter Meeting, 2022.* (Invited talk)

Transformer with Fourier Integral Attentions. *Deep Learning for Sequence Modeling Minisymposium, the SIAM Conference on Computational Science and Engineering (CSE), 2023.* (Invited talk)

Transformer with a Mixture of Gaussian Keys. *Geometry of Machine Learning Minisymposium, the 4th Annual Meeting of the SIAM Texas-Louisiana Section, 2021.* (Invited talk)

Momentum-Based and Fast Multipole Methods for Designing Deep Learning Models. *Mathematical Foundation of Deep Learning with the Applications to PDE Minisymposium, the 4th Annual Meeting of the SIAM Texas-Louisiana Section, 2021.* (Invited talk)

Brain-inspired Robust Vision Using Convolutional Neural Networks with Feedback. *NeurIPS NeuroAI Workshop, 2019.* (Poster)

Conditional Continuous Normalizing Flows for Physics-Inspired Learning. *NVIDIA Onsite Research Event, 2019.* (Lightning Talk)

Neural Rendering Model: Rethinking Neural Networks from the Joint Generation and Prediction Perspective. *NeurIPS Workshop on Integration of Deep Learning Theories, 2018.* (Contributed talk)

EnergyNet: Energy-Efficient Dynamic Inference. *NeurIPS Workshop on Compact Deep Neural Network Representation with Industrial Applications, 2018.* (Poster)

Tremor Generative Adversarial Network (TremorGAN): Deep Generative Model Approach for Geophysical Signal Generation. *NeurIPS Workshop on Machine Learning for Geophysical and Geochemical Signals, 2018*. (Poster)

The Latent-Dependent Deep Rendering Model. *ICML Workshop on Theoretical Foundations and Applications of Deep Generative Models, 2018*. (Poster)

Mixed Reality Generative Adversarial Networks: Closing the Visual Gap between Synthetic and Real Images. *Amazon Graduate Research Symposium, 2017*. (Poster)

A Probabilistic Framework for Deep Learning. *Computational and System Neuroscience Conference (COSYNE), 2016*. (Poster)

#### MENTORING

Hedi Xia, Ph.D. student in the Mathematics Department at UCLA

Tam Nguyen, Resident at FPT Software AI. (Next: Ph.D. at Rice University)

Nghia Nguyen, Resident at FPT Software AI. (Next: Ph.D. at the University of Pennsylvania)

Yujia Huang, Ph.D. student in the Computing + Mathematical Sciences Department at Caltech

Wanjia Liu, Master student in the Computer Science Department at Rice University. (Next: Google)

Si Hui Dai, Undergraduate student in the Computing + Mathematical Sciences Department at Caltech. (Next: Ph.D. at Princeton)

Ethan Perez, Undergraduate student in the Computer Science Department at Rice University. (Next: Ph.D. at NYU)

#### PROFESSIONAL SERVICES

##### **Journal reviewing**

- Journal of Machine Learning Research
- Transactions on Pattern Analysis and Machine Intelligence
- Transactions on Machine Learning Research
- Machine Learning with Applications
- Information and Inference
- IEEE Journal on Selected Areas in Information Theory
- Statistical Applications in Genetics and Molecular Biology

##### **Conference reviewing**

- International Conference on Machine Learning (ICML)
- Conference on Neural Information Processing Systems (NeurIPS)
- International Conference on Learning Representations (ICLR)
- AAAI Conference on Artificial Intelligence (AAAI)
- Asilomar Conference on Signals, Systems, and Computers

##### **Other service**

- Summer Undergraduate Research Fellowship Program at Caltech, Pasadena, California, *Project Mentor*
- Machine Learning Lunch at Rice University, Houston, Texas, *Organizer*

- Deep Learning Meeting at Rice University, Houston, Texas, *Organizer*

#### MEMBERSHIP

Institute of Electrical and Electronics Engineers (IEEE)

#### REFERENCES

##### **Professor Stanley J. Osher**

Professor of Mathematics, Computer Science, Electrical Engineering, and Chemical and Biomolecular Engineering  
 Director of Special Projects, Institute for Pure and Applied Mathematics (IPAM)  
 University of California, Los Angeles  
 Email: [sjo@math.ucla.edu](mailto:sjo@math.ucla.edu)

##### **Professor Andrea L. Bertozzi**

Distinguished Professor of Mathematics and Mechanical and Aerospace Engineering  
 Betsy Wood Knapp Chair for Innovation and Creativity  
 Director of Applied Mathematics  
 University of California, Los Angeles  
 Email: [bertozzi@math.ucla.edu](mailto:bertozzi@math.ucla.edu)

##### **Professor Richard G. Baraniuk**

C. Sidney Burris Professor of Electrical and Computer Engineering  
 Founder & Director, OpenStax  
 Rice University, Houston, Texas  
 Email: [richb@rice.edu](mailto:richb@rice.edu)

##### **Professor Don Blasius**

Professor of Mathematics  
 University of California, Los Angeles  
 Email: [blasius@math.ucla.edu](mailto:blasius@math.ucla.edu)