



# Dat Do

## Curriculum Vitae

### Personal Informations

Office **309 Jones Laboratory, 5747 S. Ellis Avenue, Chicago, IL 60637**  
Phone **+1(734)-450-3596**  
Email [dodat@uchicago.edu](mailto:dodat@uchicago.edu)  
Website <https://dodat97.github.io>

### Education

2019–2025 **Ph.D., Statistics**, *University of Michigan*, Ann Arbor, MI, USA.  
Advised by Professor Jonathan Terhorst and XuanLong Nguyen; GPA : 4/4  
2015–2019 **B.S., Mathematics**, *Hanoi University of Science*, Hanoi, Vietnam.  
Advised by Professor Cong Son Ta and Tien Dung Nguyen; GPA : 3.88/4

### Employment

8/2025– **William H. Kruskal Instructor**, *Department of Statistics, University of Chicago*, Chicago, IL, USA.  
Advised by Professor Matthew Stephens  
5/2025– **Research Associate II**, *Department of Statistics, University of Michigan*, Ann Arbor, MI, USA.  
7/2025  
Hosted by Professor Jonathan Terhorst  
1/2023– **Graduate Student Research Assistant**, *Department of Statistics, University of Michigan*, Ann Arbor, MI, USA.  
4/2025  
Advised by Professor Jonathan Terhorst and XuanLong Nguyen  
9/2019– **Graduate Student Instructor**, *Department of Statistics, University of Michigan*, Ann Arbor, MI, USA.  
12/2022

### Publications

\* stands for equal contribution.

1. **Dat Do**, Linh Do, Scott A. McKinley, Jonathan Terhorst, and XuanLong Nguyen. "Dendrogram of mixing measures: Hierarchical clustering and model selection for finite mixture models", under review with *Biometrika*, [ArXiv:2403.01684](https://arxiv.org/abs/2403.01684).
2. **Dat Do**, Linh Do, and XuanLong Nguyen. "Strong identifiability and parameter learning in regression with heterogeneous response.", *Electronic Journal of Statistics* 19(1), 131-203, (2025). [ArXiv:2212.04091](https://arxiv.org/abs/2212.04091)

3. Jiacheng Zhu\*, Aritra Guha\*, **Dat Do\***, Mengdi Xu, XuanLong Nguyen, and Ding Zhao. "[Functional optimal transport: regularized map estimation and domain adaptation for functional data](#)", *Journal of Machine Learning Research* 25, no. 276 (2024): 1-49, [ArXiv:2102.03895](#).
4. **Dat Do\***, Huy Nguyen\*, Khai Nguyen, and Nhat Ho. "[Minimax optimal rate for parameter estimation in multivariate deviated models](#)", *Advances in Neural Information Processing Systems* 36 (NeurIPS 2023), [ArXiv:2301.11808](#).
5. **Dat Do\***, Nhat Ho\*, and XuanLong Nguyen. "[Beyond Black Box Densities: Parameter Learning for the Deviated Components](#)", *Advances in Neural Information Processing Systems* 35 (NeurIPS 2022), [ArXiv:2202.02651](#).
6. Khang Le, Dung Le, Huy Nguyen, **Dat Do**, Tung Pham, and Nhat Ho. "[Entropic gromov-wasserstein between gaussian distributions](#)" In *International Conference on Machine Learning*, pp. 12164-12203. PMLR, 2022 [ArXiv: 2108.10961](#).
7. Cong Son Ta, Van Dung Le, **Dat Do**, and Trang Ta. "[Generalized Marcinkiewicz laws for weighted dependent random vectors in Hilbert spaces.](#)" *Theory of Probability & Its Applications* 67 (3), 434-451.

---

## In preparation

1. **Dat Do**, Sunrit Chakraborty, Jonathan Terhorst, and XuanLong Nguyen. "Moment tensors of Dirichlet distributions and posterior contraction behavior of the Latent Dirichlet Allocation." To be submitted to *Annals of Statistics*.
2. **Dat Do**, XuanLong Nguyen, and Jonathan Terhorst. "Concentration behavior of parameter estimation and model selection in admixture models." To be submitted to *Proceedings of the National Academy of Sciences*.
3. Linh Do, **Dat Do**, Keisha J. Cook, Nathan Rayens, Christine K. Payne, and Scott A. McKinley. "Continuous Piecewise Linear Approximation using Stochastic Searching and applications in particle tracking." To be submitted to *Statistical Science*.

---

## Talks

- 11/2024 "Inferring and Interpreting Heterogeneous Models with Dendrogram" Computation and Applied Math Seminar, Tulane University (New Orleans, Louisiana).
- 8/2024 "Dendrogram of mixing measures: Hierarchical clustering and model selection for finite mixture models." Joint Statistical Meetings (Portland, Oregon).
- 3/2024 "Dendrogram of mixing measures: Hierarchical clustering and model selection for finite mixture models." Michigan Student Symposium for Interdisciplinary Statistical Sciences 2024 (Ann Arbor, Michigan).
- 2/2022 "Functional optimal transport: map estimation and domain adaptation for functional data." AAAI Workshop 2022 on Optimal Transport and Structured Data Modeling (virtually).
- 3/2022 "Functional optimal transport: map estimation and domain adaptation for functional data." Michigan Student Symposium for Interdisciplinary Statistical Sciences 2022 (Ann Arbor, Michigan).
- 10/2022 "Strong identifiability and parameter learning in regression with heterogeneous response." 13th International Conference on Bayesian Nonparametrics (Puerto Varas, Chile).

---

## Teaching and Advising

2019–2023 **Graduate Student Instructor**, *Department of Statistics, University of Michigan*.

- STATS 425: Introduction to Probability
- STATS 412: Introduction to Probability and Statistics
- STATS 510: Probability and Distribution theory
- STATS 425: Introduction to Probability
- STATS 621: Probability Theory 2
- STATS 430: Applied Probability
- STATS 499: Honors Seminar

---

## Computer skills

Programming Python, Julia, R, Java

Python Influential in using JAX, tskit, Numpy, Matplotlib, PyTorch

Demonstration LaTeX

---

## Academic Awards, Titles, and Scholarships

- Best Poster Award, 14th Bayesian Non-parametrics conference, Los Angeles, California 2025 to the work “Dendrogram of Mixing Measures”.
- Outstanding Doctoral Dissertation Award, Department of Statistics, University of Michigan 2025
- Junior Travel Awards, 13th Bayesian Nonparametric Conference, Chile 2022
- Outstanding First-year Ph.D. Student, Department of Statistics, University of Michigan 2020
- Valedictorian of Hanoi University of Science, Vietnam National University 2019
- Second Prize, Vietnam Mathematical Olympiad 2014
- First Prize, Vietnam Mathematical Olympiad 2013

---

## Leaves and Modified Duties

2023 First child born on January 23; 3 weeks parental leaves.

---

## References

- XuanLong Nguyen, Professor, Department of Statistics, University of Michigan.
- Jonathan Terhorst, Associate Professor, Department of Statistics, University of Michigan.
- Scott McKinley, Associate Professor, Department of Mathematics, Tulane University.