

# Nhat Pham Minh Ho

---

CONTACT INFORMATION	Department of Statistics and Data Sciences University of Texas, Austin GDC 7.504, 2317 Speedway D9800 Austin, Texas 78712-1823, USA	Phone: (734) 355-9907 E-mail: minhnhath@utexas.edu <a href="https://nhatptnk8912.github.io/">https://nhatptnk8912.github.io/</a>
ACADEMIC APPOINTMENTS	<b>University of Texas, Austin, TX, USA</b> <ul style="list-style-type: none"><li>• Assistant Professor (tenure-track), Department of Statistics and Data Sciences</li><li>• Core Member, Machine Learning Laboratory</li></ul> <b>University of California, Berkeley, CA, USA</b> <ul style="list-style-type: none"><li>• Post-doctoral Scholar in EECS</li><li>• Mentors: Professors Michael I. Jordan and Martin J. Wainwright</li></ul>	2020-present  2017-2020
EDUCATION	<b>University of Michigan, Ann Arbor, Michigan USA</b> Ph.D in Statistics, 2012 - 2017 <ul style="list-style-type: none"><li>• Advisors: Professors Long Nguyen and Ya'acov Ritov</li><li>• Thesis: Parameter estimation and multilevel clustering with mixture and hierarchical models</li></ul> Master of Arts in Statistics, 2012-2013  <b>Ho Chi Minh City University of Science, Ho Chi Minh City, Vietnam</b> B.S in Mathematics and Computer Science, 2007-2011 <ul style="list-style-type: none"><li>• Major: Partial Differential Equations (PDE) and Numerical Analysis</li><li>• Rank: 3/3200 - <i>Summa Cum Laude</i></li></ul>	
RESEARCH INTERESTS	A central theme of my research focuses on four principles of statistics and data science: heterogeneity, interpretability, stability, and scalability: <ul style="list-style-type: none"><li>• Heterogeneity (Bayesian nonparametrics, hierarchical and mixture models)</li><li>• Interpretability (model misspecification, deep generative model, convolutional neural networks)</li><li>• Stability (robust statistics, (non)-convex optimization in statistical settings)</li><li>• Scalability (statistical and computational optimal transport, approximate Bayesian inference, distributed computing)</li></ul>	
PAPERS TO BE SUBMITTED (PREPRINTS, COMPLETE MANUSCRIPTS)	(Authors of papers marked with * are equally contributed. Authors of papers marked with ** are in alphabetical order)  Chiao-Yu Yang, Lihua Lei, <b>Nhat Ho</b> , Will Fithian. <a href="#">The bag-of-null-statistics procedure: An adaptive framework for selecting a better test statistics</a> . <i>To be submitted, Journal of the Royal Statistical Society: Series B</i> .  <b>Nhat Ho**</b> , Stephen Walker**. <a href="#">On efficient estimation of cumulative distribution function and its application to sampling</a> . <i>To be submitted, Biometrika</i> .  Aritra Guha, <b>Nhat Ho</b> , Chiao-Yu Yang, Long Nguyen, Michael I. Jordan. <a href="#">Bayesian sieves for unbounded parameter space and excess mass contraction under misspecification</a> . <i>To be submitted, Annals of Statistics</i> .	

Dat Do\*, **Nhat Ho\***, XuanLong Nguyen. [Statistical efficiency of parameter estimation in generalized contaminated models](#). *To be submitted, Annals of Statistics*.

Wenlong Mou, **Nhat Ho**, Martin J. Wainwright, Peter Bartlett, Michael I. Jordan. [A Poisson equation perspective on Bernstein-Von Mises theorem](#). *To be submitted, Annals of Statistics*.

Tudor Manole, **Nhat Ho**. [Uniform convergence rates of parameter estimation under Gaussian identifiable mixtures](#). *To be submitted, Annals of Statistics*.

Tudor Manole\*, **Nhat Ho\***. [Revisiting convergence rate of penalized MLE on finite mixture models](#). *To be submitted, Journal of Machine Learning Research*.

JOURNAL  
SUBMISSIONS

**Nhat Ho\*\***, Stephen Walker\*\*. [Multivariate smoothing via the Fourier integral theorem and Fourier kernel](#). *Under review, Journal of Machine Learning Research*.

Jiacheng Zhu, Jeongyeol Kwon, **Nhat Ho**, Constantine Caramanis. [On the computational and statistical complexity of over-parameterized matrix sensing](#). *Under review, Journal of Machine Learning Research*.

Mingzhang Yin, **Nhat Ho**, Bowei Yan, Mingyuan Zhou. [Probabilistic best subset selection by gradient-based optimization](#). *Under review, Journal of the American Statistical Association*.

**Nhat Ho**, Chiao-Yu Yang, Michael I. Jordan. [Convergence rates for Gaussian mixtures of experts](#). *Under review, Journal of Machine Learning Research*.

Chiao-Yu Yang, Eric Xia, **Nhat Ho**, Michael I. Jordan. [Posterior distribution for the number of clusters in Dirichlet process mixture models](#). *Under review, Journal of Machine Learning Research*.

CONFERENCE  
SUBMISSIONS

Khai Nguyen, Dinh Quoc, **Nhat Ho**, Tung Pham, Hung Bui, Dinh Phung, Trung Le. [BoMb-OT: On batch of mini-batches optimal transport](#). *Under review, ICML*.

Khang Le\*, Huy Nguyen\*, Quang Nguyen, **Nhat Ho**, Tung Pham, Hung Bui. [On robust optimal transport: computational complexity, low-rank approximation, and barycenter computation](#). *Under review, ICML*.

Son Nguyen, Khai Nguyen, **Nhat Ho**, Tung Pham, Hung Bui. [Improving Bayesian inference in deep neural networks with variational structured dropout](#). *Under review, ICML*.

Tam Le\*, Viet Huynh\*, **Nhat Ho\***, Makoto Yamada, Dinh Phung. [On scalable variant of Wasserstein barycenter](#). *Under review, UAI*.

Dang Nguyen, Khai Nguyen, **Nhat Ho**, Dinh Phung, Hung Bui. [On efficient model fusion of neural networks](#). *To be submitted, NeurIPS*.

Huy Nguyen, Dat Do, **Nhat Ho**, Dinh Phung, Hung Bui, Trung Le. [A unified framework for label shift via optimal transport](#). *To be submitted, NeurIPS*.

Dat Do, Khanh Pham, **Nhat Ho**, Dinh Phung, Hung Bui, Trung Le. [Conformalized ensemble learning](#). *To be submitted, NeurIPS*.

Trung Nguyen, Hieu Pham, Tam Le, **Nhat Ho**, Tung Pham, Son Hua. [Point-set distances for learning representations of 3D point clouds](#). *Under review, ICCV*.

**Nhat Ho\***, Raaz Dwivedi\*, Koulik Khamaru\*, Martin J. Wainwright, Michael I. Jordan, Bin Yu. [Instability, computational efficiency, and statistical accuracy](#). *Under revision*.

Wenlong Mou\*, **Nhat Ho\***, Martin J. Wainwright, Peter Bartlett, Michael I. Jordan. [A diffusion process perspective on posterior contraction rates for parameters](#). *arXiv preprint arXiv: 1909.00966*, *Under revision*.

Tan Nguyen\*, **Nhat Ho\***, Ankit Patel, Anima Anankumar, Michael I. Jordan, Richard Baraniuk. [A Bayesian perspective of convolutional neural networks through a deconvolutional generative model](#). *Under revision*.

Tianyi Lin, **Nhat Ho**, Michael I. Jordan. [On the acceleration of the Sinkhorn and Greenkhorn algorithms for optimal transport](#). *Under revision*.

Wenlong Mou, **Nhat Ho**, Martin J. Wainwright, Peter Bartlett, Michael I. Jordan. [Sampling for Bayesian mixture models: MCMC with polynomial-time mixing](#). *Under revision*.

**Nhat Ho\***, Tianyi Lin\*, Michael I. Jordan. [On structured filtering-clustering: global error bound and optimal first-order algorithms](#). *Under revision*.

Tianyi Lin, **Nhat Ho**, Marco Cuturi, Michael I. Jordan. [On the complexity of approximating multi-marginal optimal transport](#). *Under revision*.

Avi Feller\*\*, Evan Greif\*\*, **Nhat Ho\*\***, Luke W. Miratrix\*\*, Natesh S. Pillai\*\*. [Weakly separation in mixture models and implications for principal stratification](#). *Under revision*.

Aritra Guha, **Nhat Ho**, XuanLong Nguyen. [On posterior contraction of parameters and interpretability in Bayesian mixture modeling](#). *To appear, Bernoulli*.

Viet Huynh\*, **Nhat Ho\***, Nhan Dam, XuanLong Nguyen, Mikhail Yurochkin, Hung Bui, Dinh Phung. [On efficient multilevel clustering via Wasserstein distances](#). *To appear, Journal of Machine Learning Research*.

Raaz Dwivedi\*, **Nhat Ho\***, Koulik Khamaru\*, Martin J. Wainwright, Michael I. Jordan, Bin Yu. [Singularity, misspecification, and the convergence rate of EM](#). *Annals of Statistics*, 48(6), 3161-3182, 2020.

**Nhat Ho**, XuanLong Nguyen, Ya'acov Ritov. [Robust estimation of mixing measures in finite mixture models](#). *Bernoulli*, 26(2), 828-857, 2020.

**Nhat Ho** and XuanLong Nguyen. [Singularity structures and impacts on parameter estimation in finite mixtures of distributions](#). *SIAM Journal on Mathematics of Data Science*, 1(4), 730-758, 2019.

**Nhat Ho** and XuanLong Nguyen. [Convergence rates of parameter estimation for some weakly identifiable finite mixtures](#). *Annals of Statistics*, 44(6), 2726-2755, 2016.

**Nhat Ho** and XuanLong Nguyen. [On strong identifiability and convergence rates of parameter estimation in finite mixtures](#). *Electronic Journal of Statistics*, 10(1), 271-307, 2016.

Trung Le, Tuan Nguyen, **Nhat Ho**, Hung Bui, Dinh Phung. [LAMDA: label matching deep domain adaptation](#). *ICML*, 2021.

Jeong Y. Kwon, **Nhat Ho**, Constantine Caramanis. [On the minimax optimality of the EM algorithm](#)

for learning two-component mixed linear regression. *AISTATS, 2021*.

Tam Le\*, **Nhat Ho\***, Makoto Yamada. Flow-based alignment approaches for probability measures in different spaces. *AISTATS, 2021*.

Khai Nguyen, **Nhat Ho**, Tung Pham, Hung Bui. Distributional sliced-Wasserstein and applications to generative modeling. *ICLR, 2021 (Spotlight)*.

Khai Nguyen, Son Nguyen, **Nhat Ho**, Tung Pham, Hung Bui. Improving relational regularized autoencoders with spherical sliced fused Gromov-Wasserstein. *ICLR, 2021*.

Tianyi Lin\*, Chenyou Fan\*, **Nhat Ho**, Marco Cuturi, Michael I. Jordan. Projection robust Wasserstein distance and Riemannian optimization. *Advances in NeurIPS, 2020 (Spotlight)*.

Tianyi Lin, **Nhat Ho**, Xi Chen, Marco Cuturi, Michael I. Jordan. Revisiting fixed support Wasserstein barycenter: Computational hardness and efficient algorithms. *Advances in NeurIPS, 2020*.

Khiem Pham\*, Khang Le\*, **Nhat Ho**, Tung Pham, Hung Bui. On unbalanced optimal transport: an analysis of Sinkhorn algorithm. *Proceedings of the ICML, 2020*.

Raaz Dwivedi\*, **Nhat Ho\***, Koulik Khamaru\*, Martin J. Wainwright, Michael I. Jordan, Bin Yu. Sharp analysis of Expectation-Maximization for weakly identifiable models. *AISTATS, 2020*.

Wenshuo Guo, **Nhat Ho**, Michael I. Jordan. Fast algorithms for computational optimal transport and Wasserstein barycenters. *AISTATS, 2020*.

Tianyi Lin\*, **Nhat Ho\***, Michael I. Jordan. On efficient optimal transport: an analysis of greedy and accelerated mirror descent algorithms. *Proceedings of the ICML, 2019*.

**Nhat Ho\***, Viet Huynh\*, Dinh Phung, Michael I. Jordan. A probabilistic approach to multilevel clustering via composite transportation distance. *AISTATS, 2019*.

Raaz Dwivedi\*, **Nhat Ho\***, Koulik Khamaru\*, Martin J. Wainwright, Michael I. Jordan. Theoretical guarantees for EM under misspecified Gaussian mixture models. *Advances in NeurIPS, 2018*.

**Nhat Ho**, XuanLong Nguyen, Mikhail Yurochkin, Hung Bui, Viet Huynh, Dinh Phung. Multi-level clustering via Wasserstein means. *Proceedings of the ICML, 2017*.

REFEREED  
WORKSHOP PAPERS Yujia Huang\*, Sihua Dai\*, Tan Nguyen\*, **Nhat Ho\***, Ankit Patel, Anima Anankumar, Michael Jordan, Richard Baraniuk. Neural rendering model: rethinking neural networks from the joint generation and prediction perspective. *Conference on Mathematical Theory of Deep Neural Networks, 2019*.

**Nhat Ho\***, Tan Nguyen\*, Ankit Patel, Anima Anankumar, Michael Jordan, Richard Baraniuk. Neural rendering model: Joint generation and prediction for semi-supervised learning. *NIPS workshop on Integration of Deep Learning Theories, 2018*.

**Nhat Ho\***, Tan Nguyen\*, Ankit Patel, Anima Anankumar, Michael Jordan, Richard Baraniuk. Latent-dependent deep rendering model. *ICML workshop on Theoretical Foundations and Applications of Deep Generative Models, 2018*.

TEACHING

Spring 2021, SDS 323 Statistical Learning and Inference

WORKSHOP  
ORGANIZATION

*Integration of Deep Learning Theories* at NeurIPS 2018, Palais des Congrès de Montréal, Canada.

Co-organize with Professors Richard Baraniuk, Stephane Mallat, Anima Anandkumar, and Ankit Patel

INVITED SEMINAR  
PRESENTATIONS

Instability, statistical accuracy, and computational efficiency. *BLISS Seminar, Department of Electrical Engineering and Computer Sciences, April, 2021, UC Berkeley.* (Invited talk)

Statistical and computational perspectives on latent variable models. *Department of Decision Sciences at Bocconi University, Italy, November 2020.* (Invited talk)

Statistical and computational perspectives on latent variable models. *Young Data Science Researcher Seminar, ETH Zurich, July 2020.* (Invited talk)

Statistical and computational perspectives on latent variable models. *Statistics Seminar, Department of Biostatistics, March 2020, University of California, Berkeley.* (Invited talk - Canceled)

Statistical and computational perspectives on latent variable models. *Statistics Seminar, Department of Statistics and Data Science, February 2020, CMU.* (Invited talk)

Statistical and computational perspectives on latent variable models. *Statistics Seminar, Department of Statistics, February 2020, University of California, Los Angeles.* (Invited talk)

Statistical and computational perspectives on latent variable models. *Statistics Seminar, Department of Statistics and Data Science, February 2020, Cornell University.* (Invited talk)

Statistical and computational perspectives on latent variable models. *Statistics Seminar, Department of Statistics, February 2020, Rutgers University.* (Invited talk - Canceled)

Statistical and computational perspectives on latent variable models. *Statistics Seminar, Department of Statistics, February 2020, Purdue University.* (Invited talk - Canceled)

Statistical and computational perspectives on latent variable models. *Statistics Seminar, Department of Data Science, February 2020, University of California, San Diego.* (Invited talk)

Statistical and computational perspectives on latent variable models. *Statistics Seminar, Department of Data Science, February 2020, John Hopkins University.* (Invited talk - Canceled)

Statistical and computational perspectives on latent variable models. *Statistics Seminar, Department of Statistics, February 2020, University of Wisconsin, Madison.* (Invited talk - Canceled)

Statistical and computational perspectives on latent variable models. *Statistics Seminar, Department of Statistics, March 2020, University of Toronto.* (Invited talk - Canceled)

Statistical and computational perspectives on latent variable models. *Statistics Seminar, Department of Statistics, February 2020, Duke University.* (Invited talk)

Statistical and computational perspectives on latent variable models. *Statistics Seminar, Department of Statistics and Data Sciences, January 2020, University of Texas, Austin.* (Invited talk)

Statistical and computational perspectives on latent variable models. *Statistics Seminar, Department of Statistics, January 2020, University of Illinois, Urbana-Champaign.* (Invited talk)

Statistical and computational perspectives on latent variable models. *Statistics Seminar, Booth School of Management, January 2020, University of Chicago.* (Invited talk)

Statistical and computational perspectives on latent variable models. *Statistics Seminar, Department of Data Science and Operation Research, January 2020, University of Southern California.* (Invited talk)

Statistical and computational perspectives on latent variable models. *Statistics Seminar, Department of Statistics, January 2020, North Carolina State University.* (Invited talk)

Statistical and computational perspectives on latent variable models. *Statistics Seminar, Department of Statistics, January 2020, University of Minnesota, Twin Cities.* (Invited talk)

Statistical and computational perspectives on latent variable models. *Statistics Seminar, Krannert School of Management, January 2020, Purdue University.* (Invited talk)

Statistical and computational perspectives on latent variable models. *Statistics Seminar, Department of Statistics, January 2020, Pennsylvania State University.* (Invited talk)

On multi-layer latent variable models: computational and statistical perspective. *Mathematics of Data and Decisions Seminar, Department of Mathematics, December, 2019, UC Davis.* (Invited talk)

On optimal transport in machine learning and data science: computational, modeling, and theoretical perspective. *Research Seminar, VinAI Research, October, 2019, Ha Noi, Viet Nam.* (Invited talk)

Statistical and computational perspective of mixture and hierarchical models. *BLISS Seminar, Department of Electrical Engineering and Computer Sciences, October, 2019, UC Berkeley.* (Invited talk-Canceled)

Singularity structures of mixture models: Statistical and computational perspective. *Department Seminar, Department of Electrical Engineering and Computer Sciences, Rice, November, 2018, Houston, Texas.* (Invited talk)

CONFERENCE,  
WORKSHOP  
PRESENTATIONS

Statistical efficiency of parameter estimation in generalized contaminated models. *International Indian Statistical Association (IISA) Conference, 2021.* (Invited talk)

Statistical efficiency of parameter estimation in generalized contaminated models. *International Chinese Statistical Association (ICSA), Xi'an University, Xi'an, China, 2021.* (Invited talk)

Statistical efficiency of parameter estimation in generalized contaminated models. *International Conference on Econometrics and Statistics (EcoSta), Yonsei University, Seoul, South Korea, 2021.* (Invited talk)

Convergence rates for Gaussian mixtures of experts. *International Indian Statistical Association (IISA) Conference, University of Illinois at Chicago, 2020.* (Invited talk - Canceled)

On efficient optimal transport: an analysis of greedy and accelerated mirror descent algorithms. *International Conference on Machine Learning (ICML), Long Beach, CA, 2019.* (Oral presentation)

Singularity structures of mixture models: Statistical and computational perspective. *Joint Statistical Meetings (JSM), August, 2019, Denver, Colorado.* (Invited talk)

Singularity structures of parameter estimation in finite mixtures of distributions. *Joint Stanford/Berkeley Applied Math Event, November 2018, University of California, Berkeley.* (Invited talk)

Singularity structure of parameter space and posterior contraction in finite mixture models. *Joint Statistical Meetings (JSM), August, 2017, Baltimore, Maryland.* (Invited talk)

Singularity structures and parameter estimation behavior in finite mixtures of distributions. *Non-parametric Statistics Workshop: Integration of Theory, Methods, and Applications, October, 2016, Ann Arbor, Michigan.* (Poster)

Singularity structures and impacts on parameter estimation in finite mixtures of distributions. *Shannon Centennial Symposium, September, 2016, Ann Arbor, Michigan.* (Poster)

Singularity structures and parameter estimation behavior in finite mixtures of distributions. *Joint Statistical Meetings (JSM), August, 2016, Chicago, Illinois.* (Contributed talk)

Singularity structures and parameter estimation behavior in finite mixtures of distributions. *Conference on Statistical Learning and Data Science, June, 2016, University of North Carolina at Chapel Hill.* (Poster)

Singularity structures and parameter estimation behavior in finite mixtures of distributions. *Statistical Machine Learning Student Workshop, June, 2016, University of Michigan, Ann Arbor.* (Contributed talk)

Singularity structures and parameter estimation in mixtures of skew normal distributions. *Michigan Student Symposium for Interdisciplinary Statistical Sciences (MSSISS), March, 2016, Ann Arbor, MI.* (Poster)

Weak Identifiability and convergence rates of parameter estimation in over-fitted Gaussian mixture models. *Student Seminar, January, 2016, Department of Statistics, University of Michigan, Ann Arbor.* (Invited talk)

Intrinsic difficulties for the inference of mixing measures in finite mixtures of univariate skew normal distributions. *From Industrial Statistics to Data Science, October, 2015, Ann Arbor, Michigan.* (Poster)

Posterior concentration of mixing parameters in some weakly identifiable finite mixture models. *10th Conference on Bayesian Nonparametrics, June, 2015, Raleigh, North Carolina.* (Poster)

Weak identifiability and optimal rate of convergence of mixing measures in over-fitted Gaussian mixture model. *NSF Conference - Statistics for Complex Systems, June, 2015, Madison, Wisconsin.* (Poster)

Weak identifiability and optimal rate of convergence of mixing measures in over-fitted Gaussian mixture model. *Statistical Machine Learning Student Workshop, June, 2015, University of Michigan, Ann Arbor.* (Contributed talk)

Optimal convergence rate of parameter estimation in over-fitted finite Gaussian mixture model. *Michigan Student Symposium for Interdisciplinary Statistical Sciences (MSSISS), 2015, Ann Arbor, MI.* (Poster)

Identifiability and convergence rate of parameter estimators in exact-fitted finite mixture models. *Statistical Machine Learning Student Workshop, June, 2014, University of Michigan, Ann Arbor.* (Contributed talk)

- Fanqi Yan, Phd Student in Department of Computer Science
- Khang Nguyen, Incoming Phd Student in Department of Electrical and Computer Engineering (Co-mentored with Constantine Caramanis)
- Khai Nguyen, Incoming Phd Student in Department of Statistics and Data Sciences
- Xizewen Han, Phd Student in Department of Statistics and Data Sciences (Co-mentored with Mingyuan Zhou)
- Fuheng Cui, Phd Student in Department of Statistics and Data Sciences

**Phd committee member**

- Jiacheng Zhuo, Phd Student in the Department of Computer Science
- Xizewen Han, Phd Student in the Department of Statistics and Data Sciences
- Korawat Tanwisuth, Phd Student in Department of Statistics and Data Sciences

**PROFESSIONAL  
SERVICES**

**Journal reviewing:**

- Annals of Statistics (5 papers)
- Journal of the American Statistical Association (1 paper)
- Bernoulli (1 paper)
- Electronic Journal of Statistics (2 papers)
- Journal of Machine Learning Research (7 papers)
- SIAM Journal on Mathematics of Data Science (1 paper)
- Statistical Science (1 paper)
- Mathematical Statistics and Learning (1 paper)
- Mathematical Programming (1 paper)
- Annals of Applied Probability (1 paper)
- IEEE Transactions on Information Theory (1 paper)
- IEEE Transactions on Pattern Analysis and Machine Intelligence (1 paper)
- IEEE Transactions on Signal Processing (1 paper)
- Journal of Combinatorial Optimization (1 paper)
- IEEE Letters to Control System Society (1 paper)
- Pattern Recognition (1 paper)
- Entropy (2 papers)
- Stat (1 paper)
- Algorithms (1 paper)

**Conference reviewing/ Program Committee:**

- International Conference on Machine Learning (ICML), 2015, 2017, 2018, 2019, 2021
- Neural Information Processing Systems (NIPS), 2015, 2017, 2020, 2021
- International Conference on Learning Representations (ICLR), 2021
- International Conference on Artificial Intelligence and Statistics (AISTATS), 2021
- Conference on Learning Theory (COLT), 2021
- Conference on Uncertainty in Artificial Intelligence (UAI), 2021

**Professional Committees Member:**



- ASA SLDS Student Paper Award Committee, 2021

**Other services:**

- Nonparametric Statistics Workshop: Integration of Theory, Methods, and Applications, October 2016, Ann Arbor, Michigan, *Student Assistant*
- Extreme Value Analysis (EVA) conference, June 2015, Ann Arbor, Michigan, *Student Assistant*

MEMBERSHIP

American Statistical Association

International Society for Bayesian Analysis