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: minhnhat@utexas.edu
http://people.eecs.berkeley.edu/ minhnhat/

ACADEMIC
APPOINTMENTS

Assistant Professor (tenure-track)   2020-present

- Department of Statistics and Data Sciences
- **University of Texas, Austin**, TX, USA

Post-doctoral Scholar in EECS   2017-2020

- **University of California, Berkeley**, CA, USA
- Mentors: Professors Michael I. Jordan and Martin J. Wainwright

EDUCATION

University of Michigan, Ann Arbor, Michigan USA

Ph.D in Statistics, 2012 - 2017

- Advisors: Professors Long Nguyen and Ya'acov Ritov
- Thesis: Parameter estimation and multilevel clustering with mixture and hierarchical models

Master of Arts in Statistics, 2012-2013

Ho Chi Minh City University of Science, Ho Chi Minh City, Vietnam

B.S in Mathematics and Computer Science, 2007-2011
- Major: Partial Differential Equations (PDE) and Numerical Analysis
- Rank: 3/3200 - *Summa Cum Laude*

RESEARCH
INTERESTS

A central theme of my research focuses on four principles of statistics and data science: heterogeneity, interpretability, stability, and scalability:

- Heterogeneity (Bayesian nonparametrics, hierarchical and mixture models)

- Interpretability (model misspecification, deep generative model, convolutional neural networks)

- Stability (robust statistics, (non)-convex optimization in statistical settings)

- Scalability (statistical and computational optimal transport, approximate Bayesian inference, distributed computing)

PAPERS
TO BE SUBMITTED
(PREPRINTS,
COMPLETE
MANUSCRIPTS)

(Authors of papers marked with * are equally contributed. Authors of papers marked with ** are in alphabetical order)

Nhat Ho, XuanLong Nguyen. Statistical efficiency of parameter estimation in generalized contaminated models. *To be submitted, Annals of Statistics.*

Chiao-Yu Yang, Lihua Lei, **Nhat Ho**, Will Fithian. The bag-of-null-statistics procedure: An adaptive framework for selecting a better test statistics. *To be submitted, Journal of the Royal Statistical Society: Series B.*

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

Aritra Guha, Chiao-Yu Yang, **Nhat Ho**, Long Nguyen, Michael I. Jordan. On the growing parameter space under Dirichlet process mixture models. *To be submitted, Annals of Statistics.*

Tudor Manole, Sivaraman Balakrishnan, Larry Wasserman, **Nhat Ho**. Uniform convergence rates of parameter estimation under weakly identifiable mixtures. *To be submitted, Annals of Statistics.*

JOURNAL
SUBMISSIONS

Nhat Ho*, Raaz Dwivedi*, Koulik Khamaru*, Martin J. Wainwright, Michael I. Jordan, Bin Yu. Instability, computational efficiency, and statistical accuracy. *Under review, Annals of Statistics.*

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

Tan Nguyen*, **Nhat Ho***, Ankit Patel, Anima Anankumar, Michael I. Jordan, Richard Baraniuk. Neural rendering model: Joint generation and prediction for semi-supervised learning. *arXiv preprint arXiv: 1811.02657, Under review, Journal of Machine Learning Research.*

Nhat Ho, Chiao-Yu Yang, Michael I. Jordan. Convergence rates for Gaussian mixtures of experts. *arXiv preprint arXiv: 1907.04377, 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, Annals of Statistics.*

CONFERENCE
SUBMISSIONS

Jeong Y. Kwon, **Nhat Ho**, Constantine Caramanis. On the minimax optimality of the EM algorithm for learning two-component mixed linear regression. *Under review, AISTATS.*

Khai Nguyen, **Nhat Ho**, Tung Pham, Hung Bui. Distributional sliced-Wasserstein and applications to generative modeling. *Under review, ICLR.*

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

Trung Le, Khanh Nguyen, **Nhat Ho**, Hung Bui, Dinh Phung. LAMDA: label matching deep domain adaptation with theoretical guarantee. *arXiv preprint arXiv:1811.06199, Under review, AISTATS.*

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

Tam Le*, Viet Huynh*, **Nhat Ho***, Makoto Yamada, Dinh Phung. On scalable variant of Wasserstein barycenter. *Under review, AISTATS.*

Trung Nguyen, **Nhat Ho**, Tam Le, Hieu Pham, Son Hua. Point-set distances for learning representations of 3D point clouds. *Under review, CVPR.*

JOURNAL
PUBLICATIONS

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.*

Tianyi Lin, **Nhat Ho**, Michael I. Jordan. On the acceleration of the Sinkhorn and Greenkhorn algorithms for optimal transport. *arXiv preprint arXiv:1906.01437, Under revision.*

Wenlong Mou, **Nhat Ho**, Martin J. Wainwright, Peter Bartlett, Michael I. Jordan. Sampling for Bayesian mixture models: MCMC with polynomial-time mixing. *arXiv preprint arXiv: 1912.05153, Under revision.*

Nhat Ho*, Tianyi Lin*, Michael I. Jordan. On structured filtering-clustering: global error bound and optimal first-order algorithms. *arXiv preprint arXiv:1904.07462, Under revision.*

Tianyi Lin, **Nhat Ho**, Marco Cuturi, Michael I. Jordan. On the complexity of approximating multi-marginal optimal transport. *arXiv preprint arXiv: 1910.00152, Under revision.*

Avi Feller**, Evan Greif**, **Nhat Ho****, Luke W. Miratrix**, Natesh S. Pillai**. Weakly separation in mixture models and implications for principal stratification. *arXiv preprint arXiv:1602.06595, Under revision.*

Aritra Guha, **Nhat Ho**, XuanLong Nguyen. On posterior contraction of parameters and interpretability in Bayesian mixture modeling. *arXiv preprint arXiv:1901.05078, 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.*

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.*

CONFERENCE
PUBLICATIONS

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. The-oretical 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 gen-eration 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 work-shop 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 Applica-tions of Deep Generative Models, 2018.*

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

SELECTED HONORS
AND AWARDS

Conference on Statistical Learning and Data Science Travel Award, UNC Chapel Hill, 2016

Best Poster Award, Michigan Student Symposium for Interdisciplinary Statistical Sciences, 2016

NSF Conference for Complex Systems Travel Award, 2015

Rackham School of Graduate Studies Conference Travel Grant, 2015, 2016

Departmental Fellowship, Department of Statistics, University of Michigan, Fall 2012, Winter 2013

Highest Distinction Graduation Award, Ho Chi Minh University of Science, 2011

Outstanding Student Award, Deparment of Mathematics and Computer Science, Ho Chi Minh University of Science, 2008 - 2011

Odon Vallet Fellowship, Ho Chi Minh University of Science, 2008 - 2011

TEACHING
EXPERIENCE

University of Michigan
- Summer 2016: Instructor (GSI), Big Data Summer Institute, School of Public Health, Univer-sity of Michigan. This workshop introduces undergraduate students to the intersection of big data and human health. Instructor meets with the students, answer their questions, and help with their projects on analyzing the Panama paper data and Twitter data.
- Winter 2016: Graduate Student Instructor (GSI), STATS 414, Introduction to Bayesian Data Analysis. This course covers basic computational tools (Markov Chain Monte Carlo) needed for Bayesian data analysis. Instructor reviews new topics, discusses examples, answers homework, and holds office hours.
- Fall 2015: Graduate Student Instructor (GSI), STATS 250, Introduction to Statistics and Data Analysis. This course covers basic topics of statistics. Lab instructor reviews new topics, discusses examples, answers questions, and guides lab projects.

- Fall 2012- Fall 2013: Graduate Student Instructor (GSI), STATS 412, Probability and Statistics for Engineering.

Ho Chi Minh University of Science
- Winter 2012: Teaching Assistant, Advanced Calculus, Advanced Computer Science Program, Ho Chi Minh University of Science
- Fall 2011: Teaching Assistant, Advanced Calculus, Department of Mathematics and Computer Science, Ho Chi Minh University of Science

INVITED SEMINAR PRESENTATIONS

On optimal transport in machine learning and data sciences: computational, modeling, and theoretical perspective. *BLISS Seminar, Department of Electrical Engineering and Computer Sciences, November, 2020, UC Berkeley.* (Invited talk)

On optimal transport in machine learning and data sciences: computational, modeling, and theoretical perspective. *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 Chinese Statistical Association (ICSA), Xi'an University, Xi'an, China, 2021.* (Invited talk)

Convergence rates for Gaussian mixtures of experts. *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. *Nonparametric 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)

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 (5 papers)
- SIAM Journal on Mathematics of Data Science (1 paper)
- Mathematical Statistics and Learning (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 (1 paper)
- Stat (1 paper)

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

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

MEMBERSHIP

American Statistical Association

International Society for Bayesian Analysis