# PERSONAL INFORMATION Hien Duy Nguyen

School of Computing, Engineering and Mathematical Sciences, La Trobe

University, Melbourne, Australia

- **\*\*** +61 434 253 952
- <u>h.nguyen5@latrobe.edu.au</u>
- hiendn.github.io

# ACADEMIC AND RESEARCH

## **EMPLOYMENT**

### 2023- Professor

Institute of Mathematics for Industry Kyushu University, Fukuoka, Japan

## 2023- Associate Professor

Department of Mathematical and Physical Sciences La Trobe University, Melbourne, Australia

#### 2021–2023 Senior Lecturer

School of Mathematics and Physics University of Queensland, Brisbane, Australia

#### 2019-2021 Senior Lecturer

Department of Mathematics and Statistics La Trobe University, Melbourne, Australia

### 2016-2020 DECRA Research Fellow

Australian Research Council (ARC) grant number: DE170101134. La Trobe University, Melbourne, Australia

## 2016-2019 Lecturer

Department of Mathematics and Statistics La Trobe University, Melbourne, Australia

## 2015-2016 Postdoctoral Research Fellow

School of Mathematics and Physics, and Centre for Advanced Imaging University of Queensland, Brisbane, Australia

## **Predoctoral Positions**

#### 2010 Vacation Research Scholar

Diamantina Institute, University of Queensland, Brisbane, Australia

## 2009 Research Assistant

Commonwealth Scientific and Industrial Research Organisation, Australia

# 2008 Vacation Research Scholar

Commonwealth Scientific and Industrial Research Organisation, Australia

## HONORARY, ADJUNCT, AND

# SERVICE APPOINTMENTS

## 2023- Honorary Associate Professor

School of Mathematics and Physics

University of Queensland, Brisbane, Australia

## 2021-2023 Adjunct Senior Lecturer

Department of Mathematics and Statistics La Trobe University, Melbourne, Australia

### 2020-2021 Deputy Head of Department

Department of Mathematics and Statistics La Trobe University, Melbourne, Australia

# 2018, 2019 Visiting Research Fellow

Statify (previously Mistis) Team

Inria Centre at the University Grenoble Alpes, France

### 2015, 2019 Visiting Research Fellow

SIMEXP Laboratory

Centre de Reserche de l'Institut Universitaire de Geriatrie de Montreal, Canada

### 2017, 2018 Visiting Research and Teaching Fellow

Lab of Mathematics: Nicolas Oresme University of Caen, Normandy, France

#### TERTIARY EDUCATION \_

### 2011–2015 **PhD in Statistics** (Dean's Award for Outstanding RHD Theses)

University of Queensland, Brisbane, Australia

Supervisors: Prof Geoffrey McLachlan, Dr Ian Wood, and Dr Andrew Janke

Thesis title: Finite mixture models for linear regression data

## 2006–2011 Bachelor of Science with First Class Honours in Statistics (Mathematics and Statistics majors; University Medal)

University of Queensland, Brisbane, Australia

Supervisors: Dr Ian Wood and Dr Andrew George

Thesis title: Variable selection in linear models using metaheuristic algorithms with applications in genome-wide as-

sociation studies

2010 **Study Abroad** (Scholarship; Spring Semester)

University of California, Berkeley, USA

## 2006-2010 Bachelor of Economics

University of Queensland, Brisbane, Australia

#### **TEACHING AND COURSE**

#### COORDINATION

### **Teaching**

- 2022, 2023 Mathematics for Data Science 1 (MATH7501, University of Queensland; Undergraduate)
- 2022, 2023 Problems and Applications in Modern Statistics (STAT3500, University of Queensland; Undergraduate)
- 2021–2023 Statistical Learning (STAT3006, University of Queensland; Undergraduate)
- 2020, 2021 Models for bioinformatics (STA5MB, La Trobe University; Postgraduate)
- 2016 Experimental design (STAT3003, University of Queensland; Undergraduate)

#### Coordination

- 2020, 2021 Models for bioinformatics (STA5MB, La Trobe University; Postgraduate)
- 2020 **Biostatistics** (STA2BS, La Trobe University; Undergraduate)
- 2020 Engineering probability and statistics (STM2EPS, La Trobe University; Undergraduate)

### **Predoctoral Teaching**

- 2009-2014 Tutor (School of Economics, University of Queensland)
- 2007-2014 Tutor (School of Mathematics and Physics, University of Queensland)

PEER-REVIEWED

### **Journal Articles**

- S Grot, S Smine, S Potvin, M Darcey, V Pavlov, S Genon, H Nguyen, and P Orban. Label-based meta-analysis of functional brain dysconnectivity across mood and psychotic disorders. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 131(110950), 2024.
- [2] P K T Goh, A Pulemotov, H Nguyen, N Pinto, and R Olive. Treatment duration by morphology and location of impacted maxillary canines: a CBCT investigation. *American Journal of Orthodontics and Dentofacial Orthopedics*, (to appear), 2024.
- [3] F Almasi, M J Stear, M Khansefid, H Nguyen, A Desai, and J E Pryce. Innovative use of sensor technology to study grazing behaviour and its associations with parasitic resistance in sheep. *Small Ruminant Research*, 232(107223), 2024.
- [4] T P Wallis, A Jiang, K Young, H Hou, K Kudo, A J McCann, N Durisic, M Joensuu, D Oelz, H Nguyen, R S Gormal, and F A Meunier. Super-resolved trajectory-derived nanoclustering analysis using spatiotemporal indexing. *Nature Communications*, 14:3353, 2023.
- [5] T T Nguyen, F Chamroukhi, H D Nguyen, and G J McLachlan. Approximation of probability density functions via location-scale finite mixtures in Lesbegue spaces. *Communications in Statistics Theory and Methods*, 52:5048–5059, 2023.
- [6] H D Nguyen and M Gupta. Finite sample inference for empirical Bayesian methods. *Scandinavian Journal of Statistics*, 50:1616–1640, 2023.
- [7] H D Nguyen, D Fryer, and G J McLachlan. Order selection with confidence for finite mixture models. *Journal of the Korean Statistical Society*, 52:154–184, 2023.
- [8] R Navarathna, D T Le, A R Hamann, H D Nguyen, T M Stace, and A Fedorov. Passive superconducting circulator on a chip. *Physical Review Letters*, 130:037001, 2023.
- [9] J Arbel, S Girard, H D Nguyen, and A Usseglio-Carleve. Multiple expectile-based distribution: properties, Bayesian inference and applications. *Journal of Statistical Planning and Inference*, 225:146–170, 2023.
- [10] F Almasi, M Stear, M Khansefid, H Nguyen, A Desai, and J E Pryce. The repeatability and heritability of traits derived from accelerometer sensors associated with grazing and rumination time in an extensive sheep farming system. Frontiers in Animal Science, 4(1154797), 2023.
- [11] S Urchs, A Tam, P Orban, C Moreau, Y Benhajali, H D Nguyen, A C Evans, and P Bellec. Subtypes of functional connectivity associate robustly with ASD diagnosis. *eLife*, 11:e56257, 2022.
- [12] R Sohi, A Carroll, H Nguyen, Z Almasi, J Miller, J Trompf, A Bervan, B I Godoy, M Stear, A Desai, and M Jois. Determination of ewe behaviour around lambing time and prediction of parturition seven days prior to lambing by tri-axial accelerometer sensors in an extensive farming system. *Animal Production Science*, 62:1729–1738, 2022.
- [13] D C Phan, L T Truong, H D Nguyen, and R Tay. Modelling the safety effects of train commuters' access modes. *Journal of Advanced Transportation*, 2022:3473397, 2022.
- [14] T T Nguyen, H D Nguyen, F Chamroukhi, and F Forbes. A non-asymptotic approach for model selection via penalization in high-dimensional mixture of experts. *Electronic Journal of Statistics*, 16:4742–4822, 2022.
- [15] H D Nguyen and F Forbes. Global implicit function theorems and the online expectation—maximisation algorithm. *Australian and New Zealand Journal of Statistics*, 64:255–281, 2022.
- [16] J Gao, D A Burgard, B J Tscharke, F Y Lai, J W O'Brien, H D Nguyen, Q Zheng, J Li, P Du, X Li, D Wang, S Castiglioni, C Cruz-Cruz, J A Baz-Lomba, V Yargeau, E Emke, K V Thomas, J F Mueller, and P K Thai. Refining the estimation of amphetamine consumption by wastewater-based epidemiology. Water Research, 225:119182, 2022.
- [17] F Forbes, H D Nguyen, T T Nguyen, and J Arbel. Approximate Bayesian computation with surrogate posteriors. *Statistics and Computing*, 32:85, 2022.
- [18] J-B Durand, F Forbes, C D Phan, L Truong, H D Nguyen, and F Dama. Bayesian nonparametric spatial prior for traffic crash risk mapping: a case study of Victoria, Australia. *Australian and New Zealand Journal of Statistics*, 64:171–204, 2022.
- [19] F Almasi, H Nguyen, D Heydarian, R Sohi, S Nikbin, C J Jenvey, E Halliwell, E N Ponnampalam, A Desai, M Jois, and M J Stear. Quantification of behavioural variation among sheep grazing on pasture using accelerometer sensors. *Animal Production Science*, 62:1527–1538, 2022.
- [20] H D Nguyen, T T Nguyen, F Chamroukhi, and G McLachlan. Approximations of conditional probability density functions in Lesbegue spaces via mixture of experts models. *Journal of Statistical Distributions and Applications*, 8(13), 2021.
- [21] D V Fryer, I Strümke, and H Nguyen. Explaining the data or explaining a model? Shapley values that uncover non-linear dependencies. PeerJ Computer Science, 7(e582), 2021.
- [22] D Fryer, I Strümke, and H Nguyen. Shapley values for feature selection: the good, the bad, and the axioms. IEEE Access, 9:144352 – 144360, 2021.
- [23] M Vladimirova, S Girard, H Nguyen, and J Arbel. Sub-Weibull distributions: generalizing sub-Gaussian and sub-Exponential properties to heavier-tailed distributions. *Stat*, 9:e318, 2020.
- [24] E Redivo, H Nguyen, and M Gupta. Bayesian clustering of skewed and multimodal data using geometric skew normal distributions. *Computational Statistics and Data Analysis*, 152:107044, 2020.
- [25] T T Nguyen, H D Nguyen, F Chamroukhi, and G J McLachlan. Approximation by finite mixtures of continuous density functions that vanish at infinity. Cogent Mathematics and Statistics, 7:1750861, 2020.

- [26] H D Nguyen, F Forbes, and G J McLachlan. Mini-batch learning of exponential family finite mixture models. Statistics and Computing, 30:731–748, 2020.
- [27] H D Nguyen, J Arbel, H Lü, and F Forbes. Approximate Bayesian computation via the energy statistic. IEEE Access, 8:131683–131698, 2020.
- [28] D V Fryer, I Strümke, and H D Nguyen. Shapley value confidence intervals for attributing variance explained. *Frontiers in Applied Mathematics and Statistics*, 6:587199, 2020.
- [29] J Bagnall, A Jones, N Karavarsamis, and H Nguyen. The fully-visible Boltzmann machine and the Senate of the 45th Australian Parliament in 2016. *Journal of Computational Social Science*, 3:55–81, 2020.
- [30] J Arbel, O Marchal, and H D Nguyen. On strict sub-Gaussianity, optimal proxy variance and symmetry for bounded random variables. *ESAIM: Probability and Statistics*, 24:39–55, 2020.
- [31] L Truong, H Nguyen, H Nguyen, and H Vu. Pedestrian overpass use and its relationship with digital and social distractions, and overpass characteristics. *Accident Analysis and Prevention*, 131:234–238, 2019.
- [32] H D Nguyen, Y Yee, G J McLachlan, and J P Lerch. False discovery rate control for grouped or discretely supported p-values with application to a neuroimaging study. SORT, 43:1–22, 2019.
- [33] H D Nguyen and G J McLachlan. On approximation via convolution-defined mixture models. *Communications in Statistics Theory and Methods*, 48:3945–3955, 2019.
- [34] H D Nguyen, F Chamroukhi, and F Forbes. Approximation results regarding the multiple-output mixture of linear experts model. *Neurocomputing*, 366:208–214, 2019.
- [35] H D Nguyen. Asymptotic normality of the time-domain generalized least squares estimator for linear regression models. Stat, 8(e248), 2019.
- [36] A T Jones, J J Bagnall, and H D Nguyen. BoltzMM: an R package for maximum pseudolikelihood estimation of fully-visible Boltzmann machines. *Journal of Open Source Software*, 4:1193, 2019. https://doi.org/10.21105/joss.01193.
- [37] D Fryer, H Nguyen, and P Orban. studentlife: tidy handling and navigation of a valuable mobile-health dataset. *Journal of Open Source Software*, 4(40), 2019.
- [38] F Chamroukhi and H D Nguyen. Model-based clustering and classification of functional data. WIREs Data Mining and Knowledge Discovery, e1298, 2019.
- [39] P Orban, C Dansereau, L Desbois, V Mongeau-Perusse, C-E Giguere, H Nguyen, A Mendrek, E Stip, and P Bellec. Multisite generalizability of schizophrenia diagnosis classification based on functional brain connectivity. *Schizophrenia Research*, 192:167–171, 2018.
- [40] H D Nguyen, D H Wang, and G J McLachlan. Randomized mixture models for probability density approximation and estimation. *Information Sciences*, 467:135–148, 2018.
- [41] H D Nguyen, G J McLachlan, J F P Ullmann, V Voleti, W Li, E M C Hillman, D C Reutens, and A L Janke. Whole-volume clustering of time series data from zebrafish brain calcium images via mxiture modeling. *Statistical Analysis and Data Mining*, 11:5–16, 2018.
- [42] H D Nguyen and G J McLachlan. Some theoretical results regarding the polygonal distribution. *Communications in Statistics Theory and Methods*, 47:5083–5095, 2018.
- [43] H D Nguyen and G J McLachlan. Chunked-and-averaged estimators for vector parameters. Statistics and Probability Letters, 137:336–342, 2018.
- [44] H D Nguyen, A T Jones, and G J McLachlan. Stream-suitable optimization algorithms for some soft-margin support vector machine variants. *Japanese Journal of Statistics and Data Science*, 1:81–108, 2018.
- [45] H D Nguyen and F Chamroukhi. An introduction to the practical and theoretical aspects of mixture-of-experts modeling. WIREs Data Mining and Knowledge Discovery, e1246, 2018.
- [46] H D Nguyen. Near universal consistency of the maximum pseudolikelihood estimator for discrete models. *Journal of the Korean Statistical Society*, 47:90–98, 2018.
- [47] L R Lloyd-Jones, H D Nguyen, and G J McLachlan. A globally convergent algorithm for lasso-penalized mixture of linear regression models. *Computational Statistics and Data Analysis*, 119:19–38, 2018.
- [48] A T Jones, H D Nguyen, and G J McLachlan. logKDE: log-transformed kernel density estimation. Journal of Open Source Software, 3:870, 2018.
- [49] C Oyarzun, A Sanjurjo, and H Nguyen. Response functions. European Economic Review, 98:1–31, 2017.
- [50] H D Nguyen, G J McLachlan, P Orban, P Bellec, and A L Janke. Maximum pseudolikelihood estimation for a model-based clustering of time series data. *Neural Computation*, 29:990–1020, 2017.
- [51] H D Nguyen and G J McLachlan. Progress on a conjecture regarding the triangular distribution. *Communications in Statistics Theory and Methods*, 46:11261–11271, 2017.
- [52] H D Nguyen. An introduction to MM algorithms for machine learning and statistical estimation. WIREs Data Mining and Knowledge Discovery, 7(e1198), 2017.
- [53] H D Nguyen and I A Wood. Asymptotic normality of the maximum pseudolikelihood estimator for fully visible Boltzmann machines. IEEE Transactions on Neural Networks and Learning Systems, 27:897–902, 2016.

- [54] H D Nguyen and I A Wood. A block successive lower-bound maximization algorithm for the maximum pseudolikelihood estimation of fully visible Boltzmann machines. *Neural Computation*, 28:485–492, 2016.
- [55] H D Nguyen, G J McLachlan, and I A Wood. Mixtures of spatial spline regressions for clustering and classification. *Computational Statistics and Data Analysis*, 93:76–85, 2016.
- [56] H D Nguyen, G J McLachlan, J F P Ullmann, and A L Janke. Spatial clustering of time-series via mixture of autoregressions models and Markov Random Fields. Statistica Neerlandica, 70:414–439, 2016.
- [57] H D Nguyen, G J McLachlan, J F P Ullmann, and A L Janke. Laplace mixture autoregressive models. Statistics and Probability Letters, 110:18–24, 2016.
- [58] H D Nguyen and G J McLachlan. Maximum likelihood estimation of triangular and polygonal distributions. *Computational Statistics and Data Analysis*, 106:23–36, 2016.
- [59] H D Nguyen and G J McLachlan. Linear mixed models with marginally symmetric nonparametric random-effects. *Computational Statistics and Data Analysis*, 106:151–169, 2016.
- [60] H D Nguyen and G J McLachlan. Laplace mixture of linear experts. Computational Statistics and Data Analysis, 93:177–191, 2016.
- [61] H D Nguyen, L R Lloyd-Jones, and G J McLachlan. A universal approximation theorem for mixture-of-experts models. Neural Computation, 28:2585–2593, 2016.
- [62] H D Nguyen, L R Lloyd-Jones, and G J McLachlan. A block minorization-maximization algorithm for heteroscedastic regression. *IEEE Signal Processing Letters*, 23:1031–1135, 2016.
- [63] L R Lloyd-Jones, H D Nguyen, G J McLachlan, W Sumpton, and Y-G Wang. Mixture of time dependent growth models with an application to blue swimmer crab length-frequency data. *Biometrics*, 72:1255–1265, 2016.
- [64] H D Nguyen and G J McLachlan. Maximum likelihood estimation of Gaussian mixture models without matrix operations. Advances in Data Analysis and Classification, 9:371–394, 2015.
- [65] H D Nguyen, G J McLachlan, N Cherbuin, and A L Janke. False discovery rate control in magnetic resonance imaging studies via Markov random fields. *IEEE Transactions on Medical Imaging*, 33:1735–1748, 2014.
- [66] L R Lloyd-Jones, H D Nguyen, Y-G Wang, and M F O'Neill. Improved estimation of size-transition matrices using tagrecapture data. Canadian Journal of Fisheries and Aquatic Sciences, 71:1385–1394, 2014.
- [67] D Chen, A Shah, H Nguyen, D Loo, K Inder, and M Hill. Online quantitative proteomics p-value calculator for permutation-based statistical testing of peptide ratios. *Journal of Proteomics Research*, 13:4184–4191, 2014.
- [68] H D Nguyen, M M Hill, and I A Wood. A robust permutation test for quantitative SILAC proteomics experiments. Journal of Integrated OMICS, 2(80-93), 2012.
- [69] K L Inder, Y Z Zheng, M J Davis, H Moon, D Loo, H Nguyen, J A Clements, R G Parton, L J Foster, and M M Hill. Expression of PRTF in PC-3 cells modulated cholesterol dynamics and actin cytoskeleton impacting secretion pathways. *Molecular and Cellular Proteomics*, 11(M111.012245), 2012.

### **Conference Papers**

- [70] F Forbes, H D Nguyen, and T T Nguyen. Bayesian likelihood free inference using mixtures of experts. In *Proceedings of the International Joint Conference on Neural Networks*, 2024.
- [71] S Roohi, R Skarbez, and H Nguyen. Reliable emotion recognition in conversation: quantifying and communicating uncertainty. In *IJCNN Workshop on Trustworthy and Responsible AI: theory, applications and challanges*, 2023.
- [72] TT Nguyen, D N Nguyen, H D Nguyen, and F Chamroukhi. A non-asymptotic risk bound for model selection in high-dimensional mixture of experts via joint rank and variable selection. In *Proceedings of the Australasian Jount Conference on Artificial Intelligence (AJCAI)*. Springer, 2023.
- [73] D Fryer, D Lowing, I Strümke, and H Nguyen. Multi-choice Explanations: a new cooperative game structure for XAI. In IJCNN Workshop on Trustworthy and Responsible AI: theory, applications and challanges, 2023.
- [74] T T Nguyen, F Chamroukhi, H D Nguyen, and F Forbes. Model selection by penalization in mixture of experts models with a non-asymptotic approach. In 53emes Journees de Statistique de la Societe Française de Statistique (SFdS), 2022.
- [75] H Nguyen, F Forbes, G Fort, and O Cappe. An online Minorization–Maximization Algorithm. In *Proceedings of the International Federation of Classification Societies*, 2022.
- [76] F Almasi, M Khansefid, H Nguyen, A Desai, J E Pryce, and M Stear. Repeatability estimates of grazing and rumination activity of Merino sheep measured using wearable sensors. In *Proceedings of the World Congress on Genetics Applied to Livestock Production*, 2022.
- [77] D C Phan, L T Truong, H D Nguyen, and R Tay. Can walking and cycling for train access improve road safety? In *Australian Road Safety Conference (ARSC2021)*, 2021.
- [78] D Fryer, H Nguyen, and P Castellazzi. k-means on positive definite matrices, and an application to clustering in radar image sequences. In *Proceedings of the IEEE Symposium Series on Computational Intelligence*, 2020.
- [79] H D Nguyen. An introduction to approximate Bayesian computation. In Proceedings of the Research School on Statistics and Data Science (RSSDS), 2019.

- [80] F Chamroukhi, F Lecocq, and H D Nguyen. Regularized estimation and feature selection in mixtures of Gaussian-gated experts models. In *Proceedings of the Research School on Statistics and Data Science (RSSDS)*, 2019.
- [81] H D Nguyen, A T Jones, and G J McLachlan. Positive data kernel density estimation via the logKDE package for R. In *Proceedings of the Sixteenth Australasian Data Mining Conference*, 2018.
- [82] H D Nguyen and G J McLachlan. Iteratively-reweighted least-squares fitting of support vector machines: a majorization-minimization algorithm approach. In *Proceedings of the 2017 Future Technologies Conference (FTC)*, 2017.
- [83] H D Nguyen. A two-sample Kolmogorov-Smirnov-like test for Big Data. In *Proceedings of the Fifteenth Australasian Data Mining Conference*, 2017.
- [84] H D Nguyen and G J McLachlan. Asymptotic inference for hidden process regression models. In Proceedings of the IEEE Statistical Signal Processing Workshop, 2014.
- [85] H D Nguyen, A L Janke, N Cherbuin, G J McLachlan, P Sachdev, and K J Anstey. Spatial false discovery rate control for magnetic resonance imaging studies. In *Proceedings of the 2013 Digital Imaging: Techniques and Applications (DICTA) Conference*, 2013.
- [86] H D Nguyen and I A Wood. Variable selection in statistical models using population-based incremental learning with applications to genome-wide association studies. In *Proceedings of the 2012 IEEE Congress on Evolutionary Computation* (CEC), 2012.

#### **Book Chapters**

- [87] G J McLachlan, S K Ng, and H D Nguyen. EM Algorithm. In Wiley StatsRef: Statistics Reference Online. Wiley, Chichester, 2021.
- [88] H D Nguyen and A T Jones. Big data-appropriate clustering via stochastic approximation and Gaussian mixture models. In *Data Analytics: Concepts, Techniques, and Applications*. CRC Press, 2018.
- [89] H D Nguyen, G J McLachlan, and M M Hill. Permutation tests with false discovery corrections for comparative-profiling proteomics experiments. In *Methods in Molecular Biology: Proteomics Bioinformatics*. Springer, 2017.

#### OTHER PUBLICATIONS \_

- [90] H Nguyen, editor. Statistics and Data Science: Proceedings of the 2019 Research School on Statistics and Data Science (RSSDS), Singapore, 2019. Springer.
- [91] H Nguyen, S Lee, and F Forbes. A Festschrift for Geoff McLachlan. Australian and New Zealand Journal of Statistics, 64:111–116, 2022.
- [92] H D Nguyen, J Bagnall-Guerreiro, and A T Jones. Universal inference with composite likelihoods. In *Proceedings of the 63rd ISI World Statistics Congress*, 2021.
- [93] H D Nguyen. Finite sample inference for generic autoregressive models. In *Proceedings of FMfl 2021*, Math-for-Industry Lecture Notes. Institute of Mathematics for Industry, Kyushu University, 2021.
- [94] R Gray, H Nguyen, D Bressington, M Jones, and D Thompson. Comment on Mothers' voices and white noise on premature infants' psychological reactions in a neonatal intensive care unit: A multi-arm randomised controlled trial. *International Journal of Nursing Studies*, (104050), 2021.
- [95] A T Jones, H D Nguyen, and J J Bagnall. BoltzMM: Boltzmann Machines with MM Algorithms. Software published in the Comprehensive R Archive Network, 2019. https://CRAN.R- project.org/package=BoltzMM.
- [96] H D Nguyen, A T Jones, and G J McLachlan. logKDE: Computing log-transformed kernel density estmates for positive data. Software published in the Comprehensive R Archive Network, 2018. https://CRAN.R-project.org/package=logKDE.
- [97] J Bagnall, A T Jones, and H Nguyen. Analysing the voting patterns of the Senate of the 45th Australian Parliament via fully-visible Boltzmann machines. Poster presented at UseR! 2018, 2018, https://hal.archives-ouvertes.fr/hal-01838443v1.
- [98] H D Nguyen. A novel algorithm for clustering of data on the unit sphere via mixture models. In *JSM Proceedings: Statistical Computing Section*, 2017.
- [99] G J McLachlan and H D Nguyen. Contribution to the discussion of paper by M. Drton and M. Plummer. Journal of the Royal Statistical Society B, 79:365, 2017.
- [100] A T Jones and H D Nguyen. lowmemtkmeans: Low memory use trimmed k-means. Software published in the Comprehensive R Archive Network, 2016. https://CRAN.R-project.org/package=lowmemtkmeans.
- [101] H D Nguyen. NostalgiR: Advanced text-based plots. Software published in the Comprehensive R Archive Network, 2015. https://CRAN.R-project.org/package=NostalgiR.
- [102] H D Nguyen. Finite mixture models for regression problems. PhD thesis, University of Queensland, 2015. https://doi.org/10.14264/uql.2015.584.

SUBMITTED MANUSCRIPTS,

- [103] H D Nguyen, TT Nguyen, J Arbel, and F Forbes. Concentration results for approximate Bayesian computation without identifiability. 2023.
- [104] H D Nguyen. PanIC: consistent information criteria for general model selection problems. 2023.
- [105] T T Nguyen, F Chamroukhi, H D Nguyen, and F Forbes. A non-asymptotic model selection in block-diagonal mixture of polynomial experts models. 2021.
- [106] S Urchs, H D Nguyen, C Moreau, C Dansereau, A Tam, A C Evans, and P Bellec. Reportducible functional connectivity endophenotype confers high risk of ASD diagnosis in subset of individuals. 2020.
- [107] M Terrett, D Fryer, T Doody, H Nguyen, and P Castellazzi. SARGDV: Efficient identification of groundwatter-dependent vegetation using synthetic aperture radar. 2020.
- [108] T T Nguyen, H D Nguyen, F Chamroukhi, and G J McLachlan. An  $l_1$ -oracle inequality for the lasso in mixture-of-experts regression models. 2020.
- [109] H D Nguyen and D V Fryer. A binary-response regression model based on support vector machines. 2020.
- [110] H D Nguyen. A note on the convergence of the Gaussian mean shift algorithm. https://arxiv.org/abs/1703.02337, 2017.
- [111] H D Nguyen, G J McLachlan, J F P Ullmann, and A L Janke. Faster functional clustering via Gaussian mixture models. https://arxiv.org/abs/1608.05481, 2016.

#### GRANTS AND RESEARCH

#### **FUNDING**

#### 2023-2026 Inria Associate Teams (48,000 EURO)

Co-investigator: Dr Florence Forbes (Inria Centre at the University Grenoble Alpes)

Project: Variance-reduced Optimization Methods and Bayesian Approximation Techniques for scalable inference (WOMBAT)

#### 2023–2025 **Discovery Project** (360,000 AUD; DP230100905)

Australian Research Council

Co-investigators: Dr Xin Guo (University of Queensland), Dr Florence Forbes (Inria Centre at the University Grenoble Alpes), and A Prof Gersende Fort (CNRS)

Project: Stochastic majorization-minimization algorithms for data science

# 2021–2023 Australian Government Road Safety Innovation Fund (101,248 AUD)

Co-investigator: Dr Long Truong (La Trobe University)

Project: Development of planning-stage crash prediction tools for use at the planning stages

#### 2019–2021 Inria Associate Teams (36,000 EURO)

Co-investigator: Dr Florence Forbes (Inria Centre at the University Grenoble Alpes)

Project: Latent analysis, adversarial networks, and DimEnsionality Reduction (LANDER)

## 2018–2021 Discovery Project (242,194 AUD; DP180101192)

Australian Research Council

Co-investigators: Prof Geoffrey McLachlan (University of Queensland) and Dr Sharon Lee (University of Queensland)

Project: Classification methods for providing personalised and class decisions

## 2017–2020 Discovery Early Career Research Award (360,000 AUD; DE170101134)

Australian Research Council

Project: Feasible algorithms for big inference

#### 2018 AFRAN Call for Initiatives (3,000 AUD)

Australian-French Association for Research and Innovation

Project: Research School on Statistics and Data Science

## 2018 FASIC RESEARCHERS Program (2,545 EURO)

Ministry of Europe and Foreign Affairs, and Ministry of Higher Education, Research and Innovation (France)

Co-awardee: Dr Florence Forbes (Inria Centre at the University Grenoble Alpes)

Project: Latent analysis, adversarial networks, and dimensionality reduction

# 2017-2019 Start Up Grant (15,000 AUD)

La Trobe University

2015 AK Head Travelling Scholarship (13,000 AUD) Australian Academy of Science 2011–2014 ASPREE-ENVISION Scholarship (21,000 AUD) Australian National University 2011–2014 Advantage RHD Scholarship (17,500 AUD) University of Queensland 2011-2014 Australian Postgraduate Award (87,000 AUD) University of Queensland 2010 Honours Research Scholarship (5,000 AUD) Commonwealth Scientific and Industrial Research Organization 2010 Summer Research Scholarship (3,000 AUD) University of Queensland 2009 Economics Jubilee Scholarship (8,000 AUD) University of Queensland 2009 National Priority Scholarship (2,000 AUD) University of Queensland 2008 Vacation Research Scholarship (8,500 AUD) Commonwealth Scientific and Industrial Research Organization 2008 Elizabeth Norsworthy Power Industry Bursary (7,500 AUD) Stanwell Corporation Limited 2008 **Prize for Third Year Statistics** (500 AUD) Department of Education, Employment and Workplace Relations 2007, 2008 Undergraduate Scholarship (8,000 AUD) Australian Bureau of Statistics 2006, 2007 Commonwealth Educational Cost Scholarship (4,000 AUD) University of Queensland 2006 Kenneth Swanwick Memorial Prize (100 AUD) University of Queensland 2006 John Black Prize (100 AUD) University of Queensland **TEACHING AWARDS** 2013, 2014 Distinguished Teaching Award School of Economics, University of Queensland 2012 **Excellence in Tutoring Award** School of Mathematics and Physics, University of Queensland POST-DOCTORAL SUPERVISION 2023-Trung Tin Nguyen (University of Queensland) Co-supervised with Dr Xin Guo, University of Queensland Project: Stochastic optimization algorithms for mixtures of experts 2021–2023 (Inria Centre at the University of Grenoble Alpes) Co-supervised with Dr Florence Forbes, Inria Centre at the University Grenoble Alpes Project: Non-asymptotic methods for Bayesian computation STUDENT SUPERVISION 2024-Jiading Liu (PhD, University of Queensland)

Co-supervised with Dr Xin Guo, University of Queensland

Project: Kernel learning methods

2024- Jacob Westerhout (PhD, University of Queensland)

Co-supervised with Dr Xin Guo, University of Queensland

Project: Weak convergence of empirical processes and the delta method

2023 (Honours, University of Queensland)

Project: Asymptotic distributions of optimal values arising from sample averaging processes

2019–2024 Daniel Fryer (PhD, University of Queensland)

Co-supervised with Prof Geoffrey McLachlan and Dr Ian Wood, University of Queensland

Project: Interpretable methods in AI and statistics

2019–2024 Fazel Almasi (PhD, La Trobe University)

Co-supervised with Profs Michael Stear, Jennie Pryce, Aniruddha Desai, and Drs Markandeya Jois and Eric Pon-

nampalam, La Trobe University

Project: Parasitological and behavioral study of lambs using sensor technology

2023— Samad Roohi (PhD, La Trobe University)

Co-supervised with Dr Richard Skarbez, La Trobe University

Project: Natural language models for affective conversation systems

2023 **Qingyuan Zhang** (Honours, University of Queensland)

Project: Determination of optimal constraints for regularization problems in the Ivanov form

2022-2023 Mark Chiu Chong (Honours, University of Queensland)

Project: Risk bounds for mixture density estimation on compact sets via Bregman divergences

2022 **Callum Macfarlane** (Honours, University of Queensland)

Co-supervised with Em Prof Alan Mark and Dr Martin Stroet, University of Queensland

Project: Optimisation methods of refined point charge fitting for molecular modelling forcefields

2021–2022 Nicholas Williams (Honours, University of Queensland)

Co-supervised with Dr Ramiro Lafuente, University of Queensland

Project: Links between maximum likelihood estimation and geometric invariant theory

2019–2022 Cong Duc Phan (M Engineering, La Trobe University)

Co-supervised with Dr Long Truong, La Trobe University

Project: Exploring the Network-wide Road Safety Impacts of Public Transport Usage and Congestion Levels

2018–2021 Trung Tin Nguyen (PhD, University of Caen, Normandy)

Co-supervised with Prof Faicel Chamroukhi, University of Caen, Normandy

Project: Learning and approximation theory of mixture models and mixture of experts

2020 **Nauvoo Perez** (AMSI Vacation Scholar, La Trobe University)

Project: Statistical learning for time-to-relapse of colorectal cancer patients

## **CONFERENCE PRESENTATIONS**

#### **Talks**

2023 PanIC: an information criterion approach for generic model selection problems

Conference: Australian Statistical Conference

Location: Wollongong, Australia

2023 A non-asymptotic risk bound for model selection in a high-dimensional mixture of experts via joint rank and

variable selection

Conference: Australasian Joint Conference for Artificial Intelligence

Location: Brisbane, Australia

2023 Convergence of Bayesian posterior statistics and related quantities

Conference: Sequential Monte Carlo Down Under

Location: Brisbane, Australia

2022 (SAM)<sup>2</sup>: a family of sequential sample averaging algorithms via majorization-minimization

Conference: CSDA and EcoSta Workshop on Statistical Data Science

Location: Bologna, Italy (Delivered online)

2022 An online MM algorithm

Conference: 17th conference of the International Federation of Classification Societies

Location: Porto, Portugal

## 2021 Finite sample inference for nonlinear autoregressive models

Conference: 15th International Conference of the ERCIM WG on Computational and Methodological Statistics (CM-

Statistics)

Location: London, UK (Delivered online)

## 2021 Universal inference with composite likelihoods

Conference: ISI World Statistics Congress

Location: Online

## 2021 Distance-based ABC procedures

Conference: ABC in Svalbard

Location: Melbourne, Australia (Delivered onsite and online)

#### 2020 Global implicit function theorems and the online EM algorithm

Conference: 14th International Conference of the ERCIM WG on Computational and Methodological Statistics (CM-

Statistics)

Location: London, UK (Delivered online)

## 2019 Mixture of autoregressive moving average models

Conference: 13th International Conference of the ERCIM WG on Computational and Methodological Statistics (CM-

Statistics)

Location: London, UK

## 2019 Approximate Bayesian computation with discrepancy measurements

Conference: Research School on Statistics and Data Science (RSSDS2019)

Location: Melbourne, Australia

## 2019 Mixtures of local logistic regressions for nonlinear classification when data are heterogeneous

Conference: 3rd International Conference on Econometrics and Statistics (EcoSta 2019)

Location: Taichung, Taiwan

#### 2018 Fast Gaussian mixture model estimation using online EM algorithms

Conference: 11th International Conference of the ERCIM WG on Computational and Methodological Statistics (CM-

Statistics)

Location: Pisa, Italy

# 2018 On approximations via convolution-defined mixture models

Conference: 2nd Italian-French Statistical Seminar

Location: Grenoble, France

## 2017 Stream-suitable optimization algorithms for some soft-margin support vector machines

Conference: 10th International Conference of the ERCIM WG on Computational and Methodological Statistics (CM-

Statistics)

Location: London, UK

# 2017 Iteratively-reweighted least-squares fitting of support vector machines: a majorization-minimization algorithm approach (Best Paper Award)

(2001: app. 7...a.a.)

Conference: 2nd Future Technologies Conference

Location: Vancouver, Canada

# 2017 A two-sample Kolmogorov-Smirnov-like test for Big Data

Conference: 15th Australasian Data Mining Conference

Location: Melbourne, Australia

## 2017 Novel algorithm for clustering of data on the unit sphere via mixture models

Conference: Joint Statistics Meeting

Location: Baltimore, USA

# 2016 Fast model-based clustering of functional data via Gaussian mixture models

Conference: 9th International Conference of the ERCIM WG on Computational and Methodological Statistics (CM-

Statistics)

Location: Seville, Spain

# 2015 Spatial clustering of time-series via mixtures of autoregressive models and Markov random fields

Conference: 8th International Conference of the ERCIM WG on Computational and Methodological Statistics (CM-

Statistics)

Location: London, UK

# 2013 Spatial false discovery rate control for magnetic resonance imaging studies

Conference: International Conference on Digital Image Computing: Techniques and Applications (DICTA)

Location: Hobart, Australia

# Variable selection in statistical models using population-based incremental learning with applications to genome-wide association studies

Conference: IEEE Congress on Evolutionary Computation

Location: Brisbane, Australia

**Posters** 

### 2014 Asymptotic inference for hidden process regression models

Conference: IEEE Statistical Signal Processing Workshop

Location: Gold Coast, Australia

## INVITED PRESENTATIONS AND

#### **WORKSHOPS**

2024 The limits of some Bayesian model evaluation statistics (Kyushu University, Statistical Science Seminar)

ocation: Kyushu University, Japan

2024 MM Algorithms for Optimization of Statistical Models (4-day Intensive Course for the Joint Graduate School of

Mathematics for Innovation)

Location: Kyushu University, Japan

2023 PanIC at the Boathouse

Conference: ANU RSFAS Summer Research Camp

Location: Canberra, Australia

2022 Asymptotic concentration of the ABC posterior estimator without identifiability (ICASM)

Location: University of Economics, Ho Chi Minh City, Vietnam (Delivered online)

2021 Finite sample inference for generic autoregressive models (Math for Industry Forum)

Location: VIASM, Hanoi, Vietnam (Delivered online)

2021 Approximate Bayesian computation (Intellectual Climate Fund Seminar)

Location: La Trobe University, Melbourne (Delivered online)

2021 Universal inference with composite likelihoods (Research School of Finance, Actuarial Studies and Statistics

Seminar)

Location: Australian National University, Canberra (Delivered online)

2021 Order selection with confidence for mixture models (Department of Mathematics and Statistics Colloquium)

Location: Macquarie University, Sydney (Delivered online)

2020 Machine learning for statistical analysis – it's just a game! (AgriBio, Intellectual Climate Fund talk)

Location: La Trobe University, Melbourne (Delivered online)

2020 Multivariate analyses—A normal approach (AgriBio, Intellectual Climate Fund talk)

Location: La Trobe University, Melbourne (Delivered online)

2020 **The multiple testing problem in high-throughput biology** (AgriBio Lectures in Mathematical Biology)

Location: La Trobe University, Melbourne (Delivered online)

2020 Approximate Bayesian computation via the energy statistic (One World Approximate Bayesian Computation

(ABC) Seminar)

Location: University of Warwick, UK (Delivered online)

#### 2020 Shapley values for linear regression models and its application to explainable Al

Location: University of Sydney, Australia

2019 Approximation by finite mixtures of continuous density functions that vanish at infinity Location: University of Glasgow, Scotland 2019 Approximate Bayesian computation using the energy statistic (CDAC Seminar) La Trobe University, Australia 2018 Optimization theory for statistics and machine learning (4-week course) Location: University of Caen, France 2018 Minibatch and incremental learning of exponential family mixtures Location: RMIT. Australia 2018 Minibatch and incremental learning of exponential family mixtures, and the soft k-means clustering problem Location: Inria Centre at the University Grenoble Alpes, France 2018 The fully-visible Boltzmann machine, maximum pseudolikelihood estimation, and the Senate of the 45th Australian Parliament (Monash Econometrics and Business Statistics Seminar) Location: Monash University, Australia 2018 A maximum likelihood oddity (Statistics Seminar) La Trobe University, Australia 2018 MM algorithms for statistical inference and machine learning problems (S4D) Location: University of Caen, France 2018 Theory of statistical inference: a lazy approach to obtaining asymptotic results in parametric models (S4D) Location: University of Caen, France 2017 False discovery rate control under rounding of p-values Location: Centre de Reserche de l'Institut Universitaire de Geriatrie de Montreal, Canada 2017 The Stone-Weierstrass theorem and neural networks (Kyushu-Latrobe Joint Seminar) Location: La Trobe University, Australia 2017 False discovery rate control under rounding of p-values Location: Inria Centre at the University Grenoble Alpes, France 2017 Majorization-minimization (MM) algorithms for statistical inference and machine learning problems (Research Summer School in Statistics and Big Data Science) University of Caen, France 2017 Chunked-and-averaged estimators for statistically embarrassingly parallel computation and online learning (Workshop on Big Data Analysis) Location: La Trobe University, Australia 2016 Whole-volume clustering of calcium imaged zebrafish brains via model-based functional data analysis Location: La Trobe University, Australia 2016 A novel approach to clustering time series data in large spatial arrays (Centre for Advanced Imaging Seminar Series) Location: University of Queensland, Australia 2016 Model-based methods for clustering of spatial time series data (3-day Invited Workshop for the School of Computer Science and Software Engineering) University of Western Australia, Australia 2015 Finite mixture models and false discovery rate control in MRI studies (Mouse Imaging Research Centre) Toronto Centre for Phenogenomics, Canada 2015 Finite mixture models and false discovery rate control in MRI studies (Feindel Brain Imaging Lecture Series at the McConnell Brain Imaging Centre) McGill University, Canada Location: 2013 Spatial false discovery rate control for magnetic resonance imaging studies (Neuroimaging and Brain Lab) Location: Australian National University, Australia

The beginner's guide to genetic algorithms (General Interest/Lay Audience Seminar at the School of Mathematics and Physics)

and Physics)

Location: University of Queensland, Australia

#### PROFESSIONAL SERVICE

# 2024 Member of the Graduate Program of Mathematics for Innovation

Joint Graduate School of Mathematics for Innovation

Institution: Kyushu University, Japan

#### 2024- La Trobe-Kyushu Joint Seminar Administrator

Institute of Mathematics for Industry

Institution: Kyushu University, Japsan

### 2022-2023 Chair of Engagement, Advancement and Internationalization

School of Mathematics and Physics

Institution: University of Queensland, Australia

## 2022-2023 UQ School of Mathematics and Physics Representative

Institution: Australian Data Science Network

#### 2022-2023 Faculty and School Student Misconduct Committee Member

School of Mathematics and Physics

Institution: University of Queensland, Australia

# 2022-2023 Faculty EMCA Reference Group Representative and School EMCA Committee Member

School of Mathematics and Physics

Institution: University of Queensland, Australia

## 2022-2023 Mentoring Chair

Research Committee, School of Mathematics and Physics

Institution: University of Queensland, Australia

### 2021-2023 Statistics, Modelling and Operations Research Seminar Administrator

School of Mathematics and Physics

Institution: University of Queensland, Australia

## 2020-2021 Mathematics, Statistics, and Data Science Majors Course Advisor

Bachelor of Science Program, College of Science, Health and Engineering

Institution: La Trobe University, Australia

# 2020-2021 Elected Member of the Academic Board

Levels C/D Representative

Institution: La Trobe University, Australia

#### 2017, 2018 Statistics Seminar Administrator

Department of Mathematics and Statistics

Institution: La Trobe University, Australia

#### 2012, 2013 Treasurer

Mathematics Students Society

Institution: University of Queensland, Australia

## 2011–2013 First-Year Learning Centre Tutor

School of Mathematics and Physics

Institution: University of Queensland, Australia

# ACADEMIC SERVICE

# **Grant Assessor**

2018 Free University of Bozen-Bolzano, Italy

## 2017- Australian Research Council, Australia

Discovery Early Career Research Awards; Discovery Projects; Linkage Projects; Future Fellowships; Laureate Fellowships

#### **Editorial Positions**

#### 2024- Associate Editor

Japanese Journal of Statistics and Data Science (Springer)

### 2023- Book Review Editor

Australian and New Zealand Journal of Statistics (Wiley)

### 2019- Associate Editor (Statistics)

Frontiers in Applied Mathematics and Statistics (Frontiers)

#### 2019-2023 Handling Editor (Statistical Computing)

Australian and New Zealand Journal of Statistics (Wiley)

#### 2018- Technical Editor

Australian and New Zealand Journal of Statistics (Wiley)

## **Conference Commitments**

#### 2023 Organization Committee Member (Chair of Sponsorships)

67th Annual Meeting of the Australian Mathematical Society (AustMS 2023), Brisbane Australia

#### 2020 Program Committee Member

18th Australasian Data Mining Conference (AusDM), Canberra Australia

## 2019 Workshop Co-chair

Research School in Statistics and Data Science (RSSDS), Melbourne Australia

# 2019 **Program Committee Member**

17th Australasian Data Mining Conference (AusDM), Adelaide Australia

### 2018 Program Committee Member and Tract Chair (Statistics in Data Science)

16th Australasian Data Mining Conference (AusDM), Bathurst Australia

# 2018 Program Committee Member

Research Summer School on Statistics for Data Science (S4D)

## 2018 Program Committee Member (Applications and Technologies in Big Data)

2nd International Conference on Smart Grid Assisted Internet of Things (SGIoT)

#### **Peer Review**

Advances in Data Analysis and Classification (Springer); Applied Mathematical Modelling (Elsevier); Annals of Applied Statistics (Project Euclid); Australasian Journal of Information Systems (Australian Computer Society); Australian and New Zealand Journal of Statistics (Wiley); BMC Bioinformatics (BioMed Central); Computational Statistics (Springer); Computational Statistics and Data Analysis (Elsevier); Communications in Statistics – Simulation and Computation (Taylor Francis); Electronic Journal of Statistics (Project Euclid); IEEE Transactions on Fuzzy Systems (IEEE); IEEE Transactions on Image Processing (IEEE); IEEE Transactions on Medical Imaging (IEEE); Information Sciences (Elsevier); International Journal of Computers and Applications (Taylor Francis); International Journal of Machine Intelligence and Sensory Signal Processing (Inderscience); Journal of the American Statistical Association (Taylor Francis); Journal of Statistical Statistics (Taylor Francis); Journal of Open Source Software (Open Source Initiative); Journal of Statistical Computation and Simulation (Taylor Francis); Journal of Statistical Planning and Inference (Elsevier); Knowledge-Based Systems (Elsevier); Royal Society Open Science (Royal Society); Statistical Modelling: An International Journal (SAGE); Statistical Analysis and Data Mining (Wiley); Statistics Surveys (Project Euclid); Statistics and Computing (Springer); Statistics in Medicine (Wiley); WIREs: Data Mining and Knowledge Discovery (Wiley)

## PROFESSIONAL MEMBERSHIPS

2022– Australian Mathematical Society (Fellowship Committee Member 2022–)

2019– Statistical Society of Australia (Victoria Branch, Council Member 2020–2021; Queensland Branch, President 2023–)

2018– Australian-French Association for Research and Innovation (AFRAN)