TrungTin Nguyen, Ph.D.
Phone: (+33) 769 978 975
Email: trung-tin.nguyen@inria.fr
trungtinnguyends@gmail.com
Homepage: trung-tinnguyen.github.io

Statify team, Inria Grenoble Rhone-Alpes (G208)
655, avenue de l'Europe, 38334 Montbonnot Saint-Martin, France.

Cover Letter and Responses to the Selection Criteria

Dear Committee,

In this document, I am writing to address the selection criteria for Position Number 23809, Lecturer (Level B) in Statistical Data Science in the School of Mathematical Sciences at Queensland University of Technology (QUT). Each point of the criteria shall be repeated (in boldface) and each response will follow its respective point. I thank the appointment committee in advanced regarding their consideration of my application.

1. **Completion of a doctoral qualification in statistics, mathematics or a closely related discipline.**
   I am currently a Postdoctoral Research Fellow in Statistics and Data Science with the topic "Bayesian model selection and simulated-based inference for complex and high-dimensional models" in the Statify team (formerly Mistis) at Inria Centre at the University of Grenoble Alpes, where I am fortunate to be mentored by Research Director Florence Forbes and Associate Researcher Julyan Arbel. Going further back in time, from 2018-2021, I was conferred a Ph.D. degree in Statistics and Data Science granted by Ministère de l'Enseignement Supérieur et de la Recherche (The Ministry of Higher Education, Research and Innovation) under the supervision of Professor Faïcel Chamroukhi. This work was carried out at Laboratoire de Mathématiques Nicolas Oresme, Université de Caen Normandie, within Doctoral School Mathématiques, Information, Ingénierie des Systèmes of Normandie Université. The thesis was written on "Model selection and approximation in high-dimensional mixture of experts models: from theory to practice" and successfully defended under the top experts in the field. The members of the Committee include two Rapporteurs: Professor Sylvain Arlot (Laboratoire de Mathématiques, Université Paris-Saclay) and Professor Judith Rousseau (University of Oxford) and Professor Christophe Biernacki (President of the Jury, University of Lille 1, Scientific Director of the Modal Team at INRIA Lille Nord-Europe). During my Ph.D. research, I am grateful to collaborate with Associate Professor Hien Duy Nguyen (La Trobe University), and Professor Geoff McLachlan (University of Queensland).
2. **Level B - Demonstrated ability to coordinate and teach undergraduate and postgraduate units in Statistics, Mathematics, and their applications to Data Science. Potential to develop units and curriculum in these fields. Level C - Demonstrated ability and experience across a wider range of units, including large introductory classes, and teaching statistics mathematics effectively to students in related disciplines such as engineering or the natural and health sciences.**
3. **Ability to develop and implement innovative teaching practices including flexible and online learning and authentic assessment.**
   Response to Questions 2 and 3:
   Having spent the majority of my career in research-only roles, my opportunities to build up my teaching record has been relatively limited. However, I have recently recommenced teaching duties as part of my role in Semester 1, 2023, where I coordinated the master course Statistical

analysis and document mining (Complementary Course, 17 hours) with responsible Professor Pedro Rodrigues, from Master 1 of Applied Mathematics (M1AM), Université Grenoble Alpes, France. In this course, I am responsible for teaching and designing the core material of the lectures, exercises and exams for the complementary courses of the Masters students on both theoretical and practical problems. This course is aimed at second year engineering students and M1AM students of the University of Grenoble Alpes who can choose one of the two following tracks of the Grenoble Master 2 in Modelling, Scientific Computing and Image Analysis and Data Science.

In addition to collaborating on course design, I have a good track record as a teaching assistant in other disciplines. I was a tutor (23 hours) for the Statistical methods for biology during the 2nd Semester, 2022, for Bachelor of Science in Science and Technology, particularly in Biology, Chemistry and Life Sciences at Université Grenoble Alpes. Furthermore, in autumn 2018, I was a teaching assistant (24 hours) for the Master 2 course Mathematical and numerical foundations of modelling and simulation using partial differential equations with the responsible Professor Jing-Rebecca Li (IDEFIX team, Inria) under the French-Vietnam Master 2 in Applied Mathematics. Going further back in time, in autumn 2017, I was a tutor (30 hours) for an undergraduate course Principles of Mathematical Analysis with the responsible Professor Duong Minh Duc, of the Bachelor of Mathematics and Computer Science, Vietnam National University Ho Chi Minh City, Vietnam.

In particular, I have accomplished several popular online courses in data science, including Machine Learning Specialization and Deep Learning Specialization, from Stanford University, USA, taught by Professor Andrew Ng et al.. During these courses, I learned a lot about developing and implementing innovative teaching practices from the content, authentic assessment and online learning tools. This gives me a lot of confidence in the design of the online course for my future career.

4. **Level B - Ability to develop and maintain an excellent research program that supports or complements existing research strengths in the School of Mathematical Sciences. Level C - Candidates would additionally be expected to demonstrate successful significant research outputs and grant outcomes.**

5. **Level B - Ability to supervise higher degree research students. Level C - Candidates should have demonstrated primary supervisory experience.**

Response to Questions 4 and 5:

Relative to my career stage, I consider my research record to be strong, having produced 17 research publications. The publications consist of 5 peer-reviewed journal articles, 7 peer-reviewed conference papers and 5 preprints under review.

I have been one of the lead authors on 13 of the 17 research articles, demonstrating my research independence. In addition, my research has received over 147 citations and I currently have an h-index of 7 on Google Scholar (since 2020), which is significant for my career stage. Of the remaining 4 publications, 2 were first authored by one of the Master and PhD students, which is indicative of the quality of supervision and research guidance I am able to provide, despite not having the opportunity to be their main supervisor. I believe that training future academic leaders is one of the core responsibilities of any research-oriented academic. I hope to have more opportunities to do this in future work.

My publications have appeared in high level statistics journals such as Electronic Journal of Statistics, Statistics and Computing, and Communications in Statistics - Theory and Methods, as well as data science and computer science conferences such as *a spotlight acceptance* to Conference on Neural Information Processing Systems (Top 1 conference on Artificial Intelligence based on h5-index and h5-median: Google scholar), 2022 IMS International Conference on Statistics and Data Science, and BNP13 - 13th International Conference on Bayesian Non-

parametrics. Furthermore, one of my applied research has been published in the top conference Conference on Empirical Methods in Natural Language Processing (Top 2 conference in Computational Linguistics in Computer Science and Engineering based on h5-index and h5-median: Google scholar). The quality and status of these research outlets allude towards the quality and impact of my research output.

Moreover, my interests in mathematical statistics are well aligned with those of the School of Mathematical Sciences of the Faculty of Science at the QUT. In particular, Distinguished Professor Kerrie Mengersen, whose research on Bayesian computational methods and complex systems modelling using mixture models is closely related to my research on asymptotic and non-asymptotic Bayesian mixture of experts models. My research regarding approximate Bayesian computation methods also closely aligns with the work of Professor Chris Drovandi, and Dr. Leah South, who both are active researchers in simulation-based inference, especially in approximate Bayesian computation and Bayesian synthetic likelihood (BSL). More specifically, together with Research Director Florence Forbes, Associate Researcher Julyan Arbel and Associate Professor Hien Duy Nguyen, we recently had some discussions with Professor Chris Drovandi, and Dr. Leah South about a combination between a mixture of expert model, namely Gaussian Locally Linear Mapping (GLLiM), and BSL proposal. The use of GLLiM within a BSL framework provides a double parametric approximation that allows to drastically reduce the number of simulations required and offers an alternative to overcome the curse of dimensionality in BSL approaches. At the International Conference on Monte Carlo Methods and Applications in Paris in June 2023, I have also had an interesting discussion with Dr. David Warne about the potential improvement of multifidelity multilevel Monte Carlo to accelerate approximate Bayesian parameter inference for partially observed stochastic processes using mixture of experts models.

I am also able to complement the School of Mathematical Sciences in a number of ways. I am currently pursuing a research program in the *approximation, computational and convergence rate theory of sparse mixture of experts models in deep neural networks* as it relates to statistical and machine learning methods. This research fits well with the School's current research strengths in applied, numerical and computational mathematics, statistics and operations research, and has the potential to provide synergistic research outputs towards both of these strengths. Another way in which my research can complement the department is through my work on the *application of deep neural networks mixture of experts to complex and high-dimensional models*, particularly in natural language processing (large language model), biostatistics (genomics, transcriptomics, proteomics) and computer vision (image segmentation, geospatial analysis and remote sensing), which fits well with the research strengths in machine learning and statistical computation. They all share the common goal of improving the analysis and interpretation of scalable Bayesian model selection and simulation-based inference for high-dimensional heterogeneous statistics. In particular, for high-dimensional, multimodal, heterogeneous, streaming, lifelong and large-scale data, it is worth noting that a simple mixture of experts and its recent generalization, deep neural network mixtures of experts, are extremely well suited. This research has the potential to add value to the School's current work, yielding higher profile and more impactful publications.

6. **Ability to work effectively in teams to support the School, Faculty, and University, and the ability to contribute to and foster a collegial and inclusive workplace within the school and across the University.**
I have been successful in building professional collaborative relationships at both the national and international levels. Internationally, I have produced publications and am in continual co-operation with researchers at The University of Queensland (Australia), La Trobe University (Australia), The University of Texas at Austin (USA), Institute for Infocomm Research, A*Star (Singapore), Singapore Management University (Singapore), University of Padova (Italy), Max Planck Research School for Intelligent Systems (Germany), and Vietnam National University Ho Chi Minh City (Vietnam). Nationally, in France, I have collaborative relationships with researchers at Inria, technological research institute IRT SystemX, Normandy University, and

Université Grenoble Alpes, among other institutions.

Although I believe I can work independently, I am keen to work with faculty from your institution to learn from their expertise. I therefore hope that you will give serious consideration to my application as I am extremely excited about the prospect of joining the School of Mathematical Sciences of the Faculty of Science at the QUT as a Lecturer (Level B).

I believe that my academic and professional profile is a good fit for the School of Mathematical Sciences of the Faculty of Science at the QUT as a Lecturer (Level B). I am a hardworking and productive statistician with a broad interest across the spectrum of subdisciplines and cross-disciplines. I am also very able in building collaborative relationships, both nationally and internationally. Furthermore, I am an experienced supervisor with an enthusiasm for teaching, and I am keen and able to perform service duties. I hope that you will give my application serious consideration, as I am very excited about the prospect of joining the the School of Mathematical Sciences of the Faculty of Science at the QUT.

Last but not least, thank you very much for your time and consideration. I am looking forward to hearing news from you.

Yours sincerely,

TrungTin Nguyen,
Postdoctoral Research Fellow in
Statistics and Data Science.