THE UNIVERSITY OF GEORGIA
DEPARTMENT OF STATISTICS

Colloquium Series

Thursday, September 22, 2022
4:00 PM, Room 204, Caldwell Building

Dr. Simon Mak

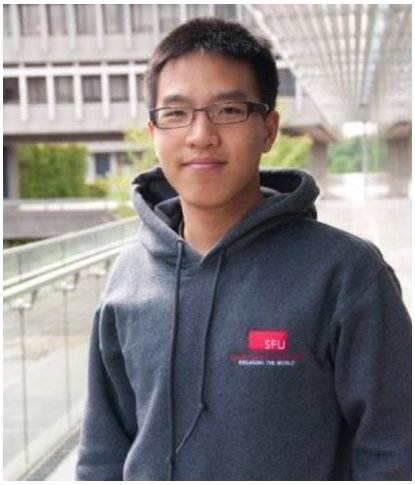
Department of Statistical Science, Duke University

Graphical multi-fidelity Gaussian process modeling, with applications to emulation of heavy-ion collisions

With advances in scientific computing, complex phenomena can now be reliably simulated via computer code. Such simulations can be very time-intensive, requiring millions of CPU hours to perform. One solution is multi-fidelity emulation, which uses data of varying accuracies (or fidelities) to train an efficient predictive model for the expensive simulator. However, for complex applications, multi-fidelity data are highly structured and embed important scientific information, which existing models cannot capture. In particular, for our high energy physics application (as well as for many problems in the physical sciences), the scientific dependencies between simulation models can be linked via a directed acyclic graph (DAG). We thus propose a new Graphical Multi-fidelity Gaussian process (GMGP) model, which embeds this DAG structure (elicited from prior scientific knowledge) within a Gaussian process framework. We show that the GMGP has desirable modeling traits and admits a scalable recursive formulation for computing the posterior predictive distribution along sub-graphs. We also present a design framework for allocating experimental runs over the DAG given a computational budget. The effectiveness of this model is then explored via a suite of numerical experiments and an application to emulating heavy-ion collisions, which sheds light on the origins of the Universe shortly after the Big Bang.

About the Speaker

Dr. Simon Mak is an Assistant Professor in the Department of Statistical Science at Duke University. Prior to Duke, he was a Postdoctoral Fellow at the Stewart School of Industrial & Systems Engineering at Georgia Tech. His research involves integrating domain knowledge (e.g., scientific theories, mechanistic models, financial principles) as prior information for statistical inference and prediction. This gives a holistic framework for interpretable statistical learning, providing a principled way for scientists to validate theories from data, and for statisticians to integrate scientific knowledge. His research tackles methodological, theoretical, and algorithmic challenges in this integration. This involves building probabilistic models on complex objects (e.g., functions, manifolds, networks), and developing efficient algorithms and data collection methods for model training. Current research is motivated from ongoing projects in nuclear physics, engineering, and finance.



For more information, please contact us at:
Phone: 706.542.5232 E-Mail: stat@uga.edu
Parking is available in the North Campus Parking Deck.
For a UGA Campus map, please see: http://aviary.camplan.uga.edu/CampusMap/Default.aspx