THE UNIVERSITY OF CHICAGO | THE COMMITTEE ON COMPUTATIONAL AND APPLIED MATHEMATICS

Dissertation Defense:

Nathan Waniorek

Computational and Applied Mathematics
University of Chicago

High-Dimensional Analysis of Ensemble Kalman Filters and Hierarchical Bayesian Inverse Problems

Wednesday, April 15, 2026 at 12:00 PM-2:00 PM

Jones 111, 5747 S. Ellis Ave. Chicago, IL 60637

.

The ensemble Kalman filter (EnKF) is a widely used algorithm for high-dimensional data assimilation tasks, such as numerical weather prediction. Despite its widespread use and empirical success, the theoretical underpinnings of the EnKF are largely not understood. The first part of this thesis aims to provide a rigorous analysis of the EnKF in high/infinite-dimensional state spaces. We first consider discretizations of Bayesian inverse problems on a Hilbert space, where we prove that the statistical complexity of the discretized ensemble Kalman update is independent of the mesh refinement. Our results hold for the important practical settings of finite-element discretizations and graph-based discretizations. Next, we establish a filter accuracy result for the EnKF applied to partially-observed nonlinear dynamical systems, addressing an important open problem in the data assimilation literature. Previous EnKF accuracy guarantees have been limited to either fully observed dynamical systems or linear, making the results inapplicable to realistic systems. We provide a natural condition coupling the nonlinear dynamics and partial observations such that the estimation error of the EnKF remains small in the long-time horizon. We also prove a filter accuracy result for the EnKF with a surrogate dynamics model, validating the use of machine-learned forecast models in ensemble data assimilation. Finally, we present learning-theoretic guarantees for learning dissipative models with operator learning methods, providing a theoretical understanding of why machine-learning forecast models can be learned with moderate sample complexity.

The second part of this thesis analyzes a family of hierarchical Bayesian models originally developed in the computational mathematics community for inverse problems in medical imaging. The main contribution of this work is to bring techniques from the high dimensional statistics literature to establish the first known reconstruction error guarantees for these models. Our analysis leverages a property of hierarchical Bayesian regularizers that we call approximate decomposability to obtain non-asymptotic bounds on the reconstruction error attained by maximum a posteriori estimators. The new theory explains how hierarchical Bayesian models that exploit sparsity, group sparsity, and sparse representations of the unknown parameter can achieve accurate reconstructions in high-dimensional settings.