

**Bayesian Learning for
High-Dimensional Nonlinear Dynamical Systems:
Methodologies, Numerics and Applications to Fluid Flows**

by

Jing Lin

B. Eng., Tsinghua University (2012)

Submitted to the Department of Mechanical Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Mechanical Engineering and Computation

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2020

© Massachusetts Institute of Technology 2020. All rights reserved.

Author
Department of Mechanical Engineering
August 20th, 2020

Certified by
Pierre F.J. Lermusiaux
Professor, Department of Mechanical Engineering
Thesis Supervisor

Accepted by
Nicolas Hadjiconstantinou
Professor, Department of Mechanical Engineering
Chairman, Department Committee on Graduate Theses

Accepted by
Youssef Marzouk
Professor, Department of Aeronautics and Astronautics
Co-Director, Center for Computational Science and Engineering

**Bayesian Learning for
High-Dimensional Nonlinear Dynamical Systems:
Methodologies, Numerics and Applications to Fluid Flows**

by
Jing Lin

Submitted to the Department of Mechanical Engineering
on August 20th, 2020, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Mechanical Engineering and Computation

Abstract

The rapidly-growing computational power and the increasing capability of uncertainty quantification, statistical inference, and machine learning have opened up new opportunities for utilizing data to assist, identify and refine physical models. In this thesis, we focus on Bayesian learning for a particular class of models: high-dimensional nonlinear dynamical systems, which have been commonly used to predict a wide range of transient phenomena including fluid flows, heat transfer, biogeochemical dynamics, and other advection-diffusion-reaction-based transport processes. Even though such models often express the differential form of fundamental laws, they commonly contain uncertainty in their initial and boundary values, parameters, forcing and even formulation. Learning such components from sparse observation data by principled Bayesian inference is very challenging due to the systems' high-dimensionality and nonlinearity.

We systematically study the theoretical and algorithmic properties of a Bayesian learning methodology built upon previous efforts in our group to address this challenge. Our systematic study breaks down into the three hierarchical components of the Bayesian learning and we develop new numerical schemes for each. The first component is on uncertainty quantification for stochastic dynamical systems and fluid flows. We study dynamic low-rank approximations using the dynamically orthogonal (DO) equations including accuracy and computational costs, and develop new numerical schemes for re-orthonormalization, adaptive subspace augmentation, residual-driven closure, and stochastic Navier-Stokes integration. The second part is on Bayesian data assimilation, where we study the properties of and connections among the different families of nonlinear and non-Gaussian filters. We derive an ensemble square-root filter based on minimal-correction second-moment matching that works especially well under the adversity of small ensemble size, sparse observations and chaotic dynamics. We also obtain a localization technique for filtering with high-dimensional systems that can be applied to nonlinear non-Gaussian inference with both brute force Monte Carlo (MC) and reduced subspace modeling in a unified way. Furthermore, we develop a mutual-information-based

adaptive sampling strategy for filtering to identify the most informative observations with respect to the state variables and/or parameters, utilizing the sub-modularity of mutual information due to the conditional independence of observation noise. The third part is on active Bayesian model learning, where we have a discrete set of candidate dynamical models and we infer the model formulation that best explains the data using principled Bayesian learning. To predict the observations that are most useful to learn the model formulation, we further extend the above adaptive sampling strategy to identify the data that are expected to be most informative with respect to both state variables and the uncertain model identity.

To investigate and showcase the effectiveness and efficiency of our theoretical and numerical advances for uncertainty quantification, Bayesian data assimilation, and active Bayesian learning with stochastic nonlinear high-dimensional dynamical systems, we apply our dynamic data-driven reduced subspace approach to several dynamical systems and compare our results against those of brute force MC and other existing methods. Specifically, we analyze our advances using several drastically different dynamical regimes modeled by the nonlinear Lorenz-96 ordinary differential equations as well as turbulent bottom gravity current dynamics modeled by the 2-D unsteady incompressible Reynolds-averaged Navier-Stokes (RANS) partial differential equations. We compare the accuracy, efficiency, and robustness of different methodologies and algorithms. With the Lorenz-96 system, we show how the performance differs under periodic, weakly chaotic, and very chaotic dynamics and under different observation layouts. With the bottom gravity current dynamics, we show how model parameters, domain geometries, initial fields, and boundary forcing formulations can be identified and how the Bayesian methodology performs when the candidate model space does not contain the true model. The results indicate that our active Bayesian learning framework can better infer the state variables and dynamical model identity with fewer observations than many alternative approaches in the literature.

Thesis Supervisor: Pierre F.J. Lermusiaux
Title: Professor, Department of Mechanical Engineering

Acknowledgments

This eight-year journey is a truly humbling experience to me. Despite all the difficulties, I still consider myself lucky, and this adventure will not be possible without many people's help.

First, I want to express my gratitude to my advisor Professor Pierre Lermusiaux. Thank you for inviting me to your MSEAS group and for introducing me to the field of uncertainty quantification, data assimilation and Bayesian inference, which I had little clue of when I first joined the group. Thank you for giving me the opportunity to work on the Bayesian model learning project. The countless tread stones demanded by the end goal caused quite a few detours, but also broadened my horizon and benefited me greatly. This is truly an eye-opening journey. Thank you also for financially supporting me all these years under grants N00014-11-1-0337 (Active Transfer Learning), N00014-14-1-0476 (Science of Autonomy - LEARNS), N00014-14-1-0725 (Bayesian Nonlinear Assimilation of Eulerian and Lagrangian Coastal Flow Data), and N00014-19-1-2693 (Interdisciplinary Nonlinear Bayesian Data Assimilation) all through the Office of Naval Research (ONR). I am also thankful for the opportunity to be the teaching assistant in your course 2.29 and 2.005. Pierre, I wish you all the best for your future career.

I also want to thank Professor Ahmed Ghoniem and Professor Anthony Patera for serving in my thesis committee. Thank you for the inspiring questions and helpful advice throughout the committee meetings and defense, and I deeply appreciate your kind encouragement and compliment. Thank you also for being tolerant about my several failed attempt to defend.

I would like to thank Dr. Patrick Haley, the research scientist in our group, for his dedication to maintaining the cluster over the years. Pat, you are always swift in resolving cluster issues and answering my requests. Your timely and clear instructions have made life much easier for a Linux idiot like me. All the heavy-duty computation will be much more of a pain without you, the cluster guardian.

I am also thankful for Leslie Regan (MechE), Kate Nelson (CCSE), Marcia Munger (MSEAS), and Lisa Mayer (MSEAS), the administrative staff who have offered me great help over the years, especially Leslie. Not only are they professional and efficient in navigating students through the miscellaneous administrative procedures, but their commitment to caring for students is also heart-

warming.

Of course, I am deeply grateful for the accompany of all the other MSEAS buddies throughout these eight years. People came and left, and it is amazing to recall all who I have come across. Let me list all the previous and current MSEAS members (excluding Pierre and Pat) who had an overlap with me and whom I have known personally, ordered by the time they joined MSEAS: Wayne Leslie, Matt Ueckermann, Tapovan Lolla, Pete Lu, Akash Phadnis, Chris Mirabito, Deepak Subramani, Doojoon Jang, Jen Landry, John Aoussou, Sydney Sroka, Abhinav Gupta, Arkopal Dutt, Florian Feppon, Corbin Foucart, Chinmay Kulkarni, Johnathan Vo, Wael Ali, Sudip Jana, Yulin Pan, Jade Coulin, Stefano Deluca, Manmeet Bhabra, Manan Doshi, Akis Gkirgkis, Mike Humara, Aaron Charous, Zach Duguid, Aditya Ghodgaonkar, Jacob Heuss, and Tony Ryu. Besides, I also consider Saviz Mowlavi and Rohit Supekar as quasi-MSEAS members due to their close personal connections with many of us. I really appreciate all the friendship and kindness all of you generously offer me. I cannot overemphasize how much your presence enriches my life, and how much I have learned from you. It is truly my privilege to share a portion of this journey of life with you.

Regarding research, I have also benefited greatly from these great buddies, especially some of those more senior ones, and I want to specifically thank them for their contributions to this work.

Matt, thank you for introducing me to your beloved projection methods for the incompressible Navier-Stokes equations, and thank you for initially developing the MATLAB finite volume framework, which becomes the workhorse of my work on bottom gravity currents. By improving, extending, and testing it over the years, I gained a lot of concrete understanding of the nuts and bolts of an incompressible Navier-Stokes solver.

Tapovan, thank you for your work on mutual-information-based adaptive sampling, from which I absorbed a lot of prior knowledge on this topic. That part of your thesis established a rich context for my subsequent advancement. Also I appreciate you, Deepak and Pat's generous help during my preparation for the qualifying exam.

Pete, although I did not have much chance to talk to you, thank you for your initial work on model learning that becomes the starting point of my project. Thank you as well for your sophisticated plotting script for a model learning demo. Plots like fig. 7-81 are certainly your legacy.

Akash, thank you for your initial work on UQ for quadratic ODEs and for developing the MC and DO solver for quadratic ODEs in MATLAB, which becomes another workhorse in my work after some improvements and extensions. The Lorenz-96 system implemented under this framework serves as a simple and yet representative example for me to quickly prototype any algorithmic idea and obtain almost instant feedback. Also, it was a pleasure to read your thesis, ask you questions, and discuss with you in the summer of 2013 before you graduated.

Deepak, as a cohort of 2012, you have the longest overlap with me in MSEAS. Thank you for all the research discussions and life sharing in these years as we sat in the same office. I still remember the night we both sensed a mild earthquake from our office in our first semester.

John, it was a rewarding experience to collaborate with you on the paper Aoussou et al. (2018). I learned a lot from your contributions and this endeavor also gave me a break from the model learning work which got rather stuck at the time. Also, I enjoyed all the research discussions and private chats with you. I always liked your sarcastic comments on things that don't make sense.

Sydney, it was a great time to work with you on refactoring the finite volume code. The log file functionality you added has been really helpful. It is always fun to catch up with you. I don't know about you, but I will remember very well the vent in our office covered by many patchy random plastic bags.

Abhinav, somehow our MSEAS destinies were entangled from the very beginning when you came as an undergraduate intern in the summer of 2014 and worked on biological model learning. It was really a pleasure to work with you all these years. The countless discussions with you also brought me plenty of inspirations. I will remember the only miserable time I had to work in the office overnight, with your accompany.

Arkopal, it is really fortunate for me to have you as a friend. It is always a pleasure to talk to you, whether for work or for fun. Thank you and Corbin for being so kind and supportive when I once wanted to quit. Your encouragement and understanding are truly heart-warming. I will miss the time when we go swimming together.

Florian, although I didn't have much chance to talk to you personally, I really appreciate the expertise in modern mathematics you brought to the group. Thank you for popularizing the low-

rank manifold perspective on DO, which was very eye-opening to me. Thank you also for pointing me to the polar decomposition when you saw my crude preliminary work that eventually crystallizes as Lin and Lermusiaux (2020). Your comments led to a giant leap of the work at the time.

Corbin, the time and conversations we had together are really unforgettable. Your passion for Python, coding and computers in general is admirable, and I certainly learned a lot about coding from you. Although all my work has been done in MATLAB, I do agree that Python sounds a lot sexier and I have made up my mind to be converted after graduation. Thank you and Johnathan also for promoting GitHub in the group. It is infinitely better than subversion and has improved my code development experience dramatically. Thank you and Abhinav also for figuring out a convenient video-making procedure using Blender, which I have found really handy and has been a huge relief to me. I will miss the time we did weight lifting together. Thank you for teaching me to do that properly as well.

Chinmay, thank you for generously sharing your committee meeting slides with me. They look really nice and I have learned many things from them. Thank you also for gathering and organizing the color map information, which I have found very useful.

Johnathan, thank you for your compassion and encouragement when things were not going smoothly with me. Thank you for leaving me kind wishes every time you came to visit us.

Wael, thank you for always offering to help. Thank you, as well as Chinmay, Corbin and Abhinav, for being my audience in the defense rehearsal and providing me with a lot of valuable feedback. You are always calm, gentle and pleasant to work with.

Saviz, thank you for all the academic and casual conversations over the years. Thank you for all your compassion and kindness. I really miss the time when we go for a volunteer shift together.

MIT is truly an amazing place. I am grateful that I have met so many brilliant and interesting minds here, who have taught me a lot in one way or another.

Lastly, needless to say, I am extremely grateful to my family members for their love and care. I want to express my gratitude to my girlfriend Aiyu, who has come across half the earth to visit me four times in the last two years. Thank you for being so supportive and for encouraging me not

to lose heart when things didn't go well. Thank you for enduring all the pain and disappointment together with me until we see the light at the end of the tunnel.

Above all, I am forever indebted to my parents. Words are pale in front of the weight of their love and sacrifice. Thank you for always being there and accepting me as who I am. With love, joy and humility, I dedicate this thesis to you.