Topic E: Scalable Uncertainty Quantification Algorithms for Truly Predictive Integrated Simulations in Magnetic Fusion Energy Sciences

Tan Bui-Thanh
The University of Texas at Austin

Inverse problems and uncertainty quantification (UQ) are pervasive in engineering and science, especially in scientific discovery and decision-making for complex, natural, engineered, and societal systems. For simulations that serve as a basis for design, control, discovery, and decision-making for complex problems, we argue, their solutions must be equipped with the degree of confidence. In other words, we have to *quantify the uncertainty* in the solution due to observation noise and numerical discretization errors, among other error sources. Though the past decades have seen tremendous advances in both theories and computational algorithms for quantifying the uncertainty in the solution of forward/inverse problems for many scientific disciplines, little work has been done on UQ for integrated simulations in fusion energy sciences. Thus, *there is a critical need to develop effective UQ strategies for integrated simulations in magnetic fusion energy sciences to go beyond interpretive simulations in order to enable scientific discovery of new plasma phenomena*.

The *long-term goal* of our research program is to establish a research program that addresses grand challenges in forward/inverse UQ for large-scale simulations in high dimensional parameter spaces: *one of the keys to reliable and predictive computational sciences*. We have a track record of developing scalable parallel algorithms for large-scale UQ, [7, 16], scalable UQ algorithms [6, 13–15, 17], large-scale PDE-constrained optimization, [16], and large-scale model and data reduction [4, 8, 18, 19]. Our *objective in this white paper* is to focus on developing scalable UQ algorithms that permits scientists to go beyond interpretive simulations in fusion energy sciences. The following specific aims are proposed:

Aim 1: New dimensional reduction techniques with error bounds. One of the major obstacles in dealing with UQ in high dimensional parameter spaces is the curse of dimensionality. To mitigate this effect, we propose to develop a quantity of interest (QoI) informed model reduction technique with provable probabilistic error bounds. The method uncovers the dominant low dimensional parameter subspaces using a detailed sensitivity of the QoI via efficient adjoint technique together with a randomized singular value decomposition approach. This will then followed by an application of effective stochastic collocation approaches on sparse grids in low dimensional reduced parameter spaces to tackle the problem of forward propagation of uncertainty.

Aim 2: Scalable Bayesian Inversion with Optimal Transport theory. While Bayesian inference is a systematic approach to account for most, if not all, uncertainties, it is prohibitively expensive for inverse problems governed by large-scale forward partial differential equation (PDE) model in extremely high dimensional parameter spaces. In particular, contemporary Markov chain Monte Carlo approaches are impractical as an inference tool for large-scale complex systems. Our strategy here is radically different. Specifically, we represent probability distribution with particles, which are then transformed, via an optimal transport theory, to particles representing the probability distribution, and hence the statistics, of the forward or inverse solutions. Since each particle requires one forward PDE solve, the method naturally accommodates computing resource constraints by limiting the number of particles. Moreover, the method simultaneously tackles both inverse and forward propagation of uncertainty in a unified framework with rigorous mathematical foundation.

Impact. We will have developed scalable and reliable methods for quantifying the uncertainty in forward/inverse fusion energy simulations. Our *advanced UQ development* will have positive impacts on fusion energy sciences that require uncertainty quantification. In particular, its potential impact on truly predictive simulations of multiphysics plasma systems will be demonstrated through its ability to efficiently and accurately quantify the uncertainties. This will, in turn, provides a means to enable scientific discovery of new plasma phenomena and decision-making in magnetic fusion energy sciences. Below are a few additional details on the proposed aims.

Aim 1. Various reduced-order modeling techniques, see, e.g., [2, 3, 20, 23, 26, 30, 32] exist, including our previous work on the proper orthogonal decomposition and model-constrained model order reduction approach for large-scale systems in high dimensional parameter spaces [4, 8, 18, 19]. We have also developed structure-exploiting model order reduction techniques in [5, 19] to efficiently propagate the uncertainty through large-scale computational fluid dynamic model in turbomachines. Nevertheless, most reduction approaches do not discover inherent low dimensional structure through detailed sensitivity analysis, and hence are not scalable for extremely high dimensional parameter spaces. Our idea is to look for a reduced subspace along which the QoI is most sensitive. Parameters outside this dominant subspace contribute insignificantly to the outcome of the QoI, and hence can be ignored. Our approach is to pose this problem as a stochastic optimization problem which can be solved efficiently in parallel using existing stochastic algorithms. The novelty of this approach is that it allows us to quantify the error for each step (and hence for total error bound) and to exploit the detailed sensitivity of the QoI. We shall use adjoint method to compute the sensitivity efficiently. To build the explicit and fast reduced model on the dominant subspace, we shall conduct both stochastic collocation on sparse grid [1] and our Hessian-informed response surface method [15].

Aim 2. There are many ways to explore the posterior distribution to estimate the inverse solution and its uncertainty. However, we argue that, Markov chain Monte Carlo (MCMC) approaches [21, 22, 24, 25, 27–29], or more generally Monte Carlo methods, are perhaps the most general UQ tool. The problem is that standard MCMC methods for posterior exploration often require millions of samples to converge, especially in high dimensions. As a result, millions of expensive forward simulations are necessary—an intractable proposition. More importantly, they discard costly work during the Metropolization if a sample is rejected. *Thus, tackling the UQ challenges for fusion energy systems requires radically different approaches*. We propose to develop an ensemble-based approach in which the prior probability distribution are represented by empirical distribution with particles (or samples), which are then transformed, via an optimization problem, to particles representing the posterior probability distribution, and hence the statistics, of the inverse solution.

Our proposed approach induces *several advantages over the contemporary counterparts*. *First*, it is a consistent Monte Carlo approach that provably converges to the exact posterior distribution. Consequently, it does not require Metropolization and hence avoiding discarding useful expensive simulations. *Second*, since each particle requires one forward PDE solve, the method naturally accommodates computing resource constraints by matching the number of particles to that of allowable forward solves. *Third*, unlike existing MCMC approaches that couple the Markov chain and forward PDE solves, the proposed method decouples the process of constructing the posterior particles and PDE simulations. As such, it is well suited for current and future supercomputer infrastructures. *Fourth*, the beauty of this approach is that it casts the sampling challenge into a large-scale optimization problem that capitalizes on our decade of work on parallel large-scale optimization algorithms [5, 7, 9–12, 14, 16, 17, 31].

References

[1] I. Babuška, F. Nobile, and R. Tempone. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 45:1005–1034, 2007.

[2] Gal Berkooz, Philip Holmes, and John L.Lumley. The proper orthogonal decomposition in the analysis of turbulent flows. *Annual Reviews of Fluid Mechanics*, 25:539–575, 1993.

[3] S. Boyaval, C. Le Bris, T. Lelievre, Y. Maday, N.C. Nguyen, and A. T. Patera. Reduced basis techniques for stochastic problems. 2009. Submitted to Archives of Computational Methods in Engineering.

[4] T. Bui-Thanh, K. Willcox, O. Ghattas, and B. van Bloemen Waanders. Goal-oriented model-constrained optimization for reduction of large-scale systems. *Journal of Computational Physics*, 224:880–896, 2007.

[5] Tan Bui-Thanh. *Model-Constrained Optimization Methods for Reduction of Parameterized Large-Scale Systems*. PhD thesis, Department of Aeronautics and Astronautics, MIT, 2007.

[6] Tan Bui-Thanh. Discretization-invariant MCMC methods for PDE-constrained Bayesian inverse problems in infinite dimensional parameter spaces. *Submitted*, 2014.

[7] Tan Bui-Thanh, Carsten Burstedde, Omar Ghattas, James Martin, Georg Stadler, and Lucas C. Wilcox. Extreme-scale UQ for Bayesian inverse problems governed by PDEs. In *SC12: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, 2012. Gordon Bell Prize finalist.

[8] Tan Bui-Thanh, Murali Damodaran, and Karen Willcox. Aerodynamic data reconstruction and inverse design using proper orthogonal decomposition. *AIAA Journal*, 42(8):1505–1516, 2004.

[9] Tan Bui-Thanh and Omar Ghattas. Analysis of the Hessian for inverse scattering problems. Part I: Inverse shape scattering of acoustic waves. *Inverse Problems*, 28(5):055001, 2012.

[10] Tan Bui-Thanh and Omar Ghattas. Analysis of the Hessian for inverse scattering problems. Part II: Inverse medium scattering of acoustic waves. *Inverse Problems*, 28(5):055002, 2012.

[11] Tan Bui-Thanh and Omar Ghattas. Non-stationary Gaussian process response for posterior interpolation: a comparative study. 2012. In preparation.

[12] Tan Bui-Thanh and Omar Ghattas. Analysis of the Hessian for inverse scattering problems. Part III: Inverse medium scattering of electromagnetic waves. Inverse Problems and Imaging, 2013.

[13] Tan Bui-Thanh and Omar Ghattas. An analysis of infinite dimensional Bayesian inverse shape acoustic scattering and its numerical approximation. *SIAM Journal of Uncertainty Quantification*, 2:203–222, 2014. To Appear.

[14] Tan Bui-Thanh and Omar Ghattas. A scalable MAP solver for Bayesian inverse problems with Besov priors. *Inverse Problems and Imaging*, 9(1):27–53, 2015.

[15] Tan Bui-Thanh, Omar Ghattas, and David Higdon. Adaptive Hessian-based nonstationary Gaussian process response surface method for probability density approximation with application to Bayesian solution of large-scale inverse problems. *SIAM Journal on Scientific Computing*, 34(6):A2837–A2871, 2012.

[16] Tan Bui-Thanh, Omar Ghattas, James Martin, and Georg Stadler. A computational framework for infinite-dimensional Bayesian inverse problems Part I: The linearized case, with application to global seismic inversion. *SIAM Journal on Scientific Computing*, 35(6):A2494–A2523, 2013.

[17] Tan Bui-Thanh and Mark Andrew Girolami. Solving large-scale PDE-constrained Bayesian inverse problems with Riemann manifold Hamiltonian Monte Carlo. *Inverse Problems*, Special Issue:114014, 2014.

[18] Tan Bui-Thanh, Karen Willcox, and Omar Ghattas. Model reduction for large-scale systems with high-dimensional parametric input space. *SIAM Journal on Scientific Computing*, 30:3270–3288, 2008.

[19] Tan Bui-Thanh, Karen Willcox, and Omar Ghattas. Parametric reduced-order models for probabilistic analysis of unsteady aerodynamic applications. *AIAA Journal*, 46:2520–2529, 2008.

[20] John Burkhardt, Max Gunzburger, and Hyung-Chun Lee. Centroidal voronoi tessellation-based reduced-order modeling of complex systems. Technical report, Florida State University, 2003.

[21] S. Duane, A. D. Kennedy, B. Pendleton, and D. Roweth. Hybrid Monte Carlo. *Phys. Lett. B*, 195:216–222, 1987.

[22] Mark Girolami and Ben Calderhead. Riemann manifold Langevin and Hamiltonian Monte Carlo methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(2):123–214, 2011.

[23] Serkan Gugercin and Athanasios C. Antoulas. A survey of model reduction by balanced truncation and some new results. *International Journal of Control*, 77(8):748–766, 2004.

[24] W. Keith Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.

[25] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953.

[26] N.C. Nguyen, G. Rozza, D.B.P. Huynh, and A.T. Patera. Reduced basis approximation and a posteriori error estimation for parametrized parabolic PDEs; Application to real-time Bayesian parameter estimation. In L. Biegler, G. Biros, O. Ghattas, M. Heinkenschloss, D. Keyes, B. Mallick, L. Tenorio, B. van Bloemen Waanders, and K. Willcox, editors, *Large Scale Inverse Problems and Quantification of Uncertainty*. John Wiley & Sons, 2011.

[27] Christian P. Robert and George Casella. *Monte Carlo Statistical Methods (Springer Texts in Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2005.

[28] Gareth O. Roberts and Jeffrey S. Rosenthal. Optimal scaling of discrete approximations to Langevin diffusions. *J. R. Statist. Soc. B*, 60:255–268, 1997.

[29] Gareth O. Roberts and Richard L. Tweedie. Exponential convergence of Langevin distributions and their discrete approximations. *Bernoulli*, 2(4):341–363, 1996.

[30] L. Sirovich. Turbulence and the dynamics of coherent structures, i-iii. *Quart. Appl. Math.*, 45:561–590, 1987.

[31] Lucas C. Wilcox, Georg Stadler, Tan Bui-Thanh, and Omar Ghattas. Discretely exact derivatives for hyperbolic PDE-constrained optimization problems discretized by the discontinuous Galerkin method. *Journal of Scientific Computing*, 2014. To Appear.

[32] K. Willcox and A. Megretski. Fourier series for accurate, stable, reduced-order models in large-scale linear applications. *SIAM Journal on Scientific Computing*, 26(3):944–962 (electronic), 2005.