FASTMath: UQ Algorithms

FASTMath UQ Team Members: T. Casey, B. Debusschere, M. Eldred, G. Geraci, R. Ghanem, J. Jakeman, Y. Marzouk, H. Najm, C. Safta, and K. Sargsyan

The SciDAC FASTMath uncertainty quantification (UQ) team works on development of robust UQ methods and software, necessary for predictive large-scale computational modeling in applications of relevance to DOE/SC. Working on SciDAC partnership projects, we focus on hardening and adapting UQ capabilities to provide effective solutions according to project needs.

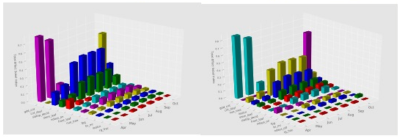
Low-Rank tensor train surrogates for high-dimensional computational models

Deploy surrogates that exploit structure in parameter to output map - seek low-rank functional tensor-train (LRFTT) representation to reveal couplings in high-dimensional models.

- Approach is analogous to low-rank tensor decompositions
- Flexibility when choosing univariate functions
  - Can combine both spectral and kernel representations
- Model fitting using Stochastic Gradient Descent

$$f(x_1, x_2, \ldots, x_d) \cong \mathcal{F}_1(x_1)\mathcal{F}_2(x_2)\cdots\mathcal{F}_d(x_d)$$

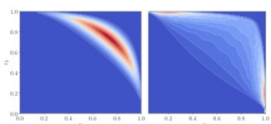
$$\mathcal{F}_k(x_k) = \begin{bmatrix} f_k^{1,1}(x_k) & \cdots & f_k^{1,r_k}(x_k) \\ \vdots & \ddots & \vdots \\ f_k^{r_{k-1},1}(x_k) & \cdots & f_k^{r_{k-1},r_k}(x_k) \end{bmatrix}$$



Total effect Sobol indices for Leaf Area Index (left) and Gross Primary Production (right) extracted from the LRFTT surrogate

Surrogates for dependent random variables

- Building accurate approximations of models with dependent random variables is challenging.
- Building approximations using samples from the probability measure is ill conditioned.
- We build polynomial chaos interpolants using Leja sequences, which are nested samples from the induced distribution that focus samples in high probability regions while maintaining stability.

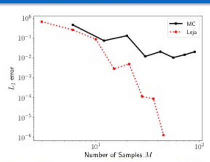


(Left) Dependent measure of random variables. (Right) Induced measure for degree 20 polynomial

- For a fixed budget, our approach can obtain errors which are orders of magnitude smaller than those obtained using existing methods

Sampling discrete random variables

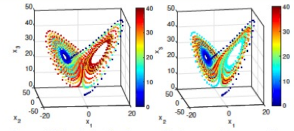
- Define polynomial chaos expansion (PCE) basis functions orthogonal *w.r.t.* standard probability masses
  - *e.g.* Charlier, Hahn, Krawtchouk
- Generate set of candidate samples
  - Nested samples in high-probability regions
- Compute Leja sequence via pivoted LU decomposition



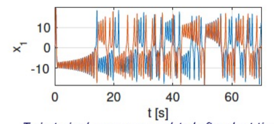
Errors between model values and PCEs constructed using Monte Carlo and Leja samples, for a standard test problem (Genz oscillatory)

Parameter inference in chaotic systems

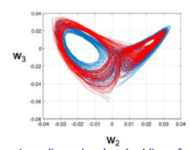
- Pointwise distances no longer useful in chaotic systems where trajectories become uncorrelated after finite time
- Classical approaches compare milestones of the trajectories to quantify shape changes
- Formalize these ideas using manifold learning, diffusion maps
- Compare trajectories in the coordinates defined by the diffusion map



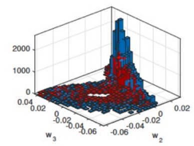
Lorenz63 trajectories for small changes in parameters, overall trajectory structure is essentially unchanged



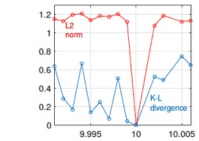
Trajectories become uncorrelated after short times



Low-dimensional embedding of attractor manifold structure



Trajectories represented as densities in manifold coordinates

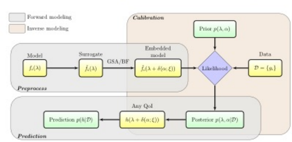


Distances between data sets using L_2 distance vs. KL-divergence

Model error quantification and propagation

Bayesian framework for *embedded* model structural error representation, propagation and attribution

- Polynomial Chaos representation of embedded structural error corrections
- Simultaneous estimation of physical parameters and model error
- Predictive variance attribution to model error, data noise and surrogate error
- Workflow implemented in UQ toolkit (*www.sandia.gov/uqtoolkit*) UQTk



Applications

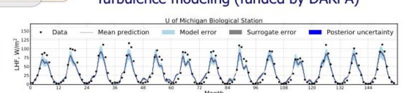
E3SM Land Model (BER partnership)

Plasma Surface Interactions (FES partnership)

Chemical kinetic modeling (funded by BES)

Turbulence modeling (funded by DARPA)

Predictive attribution of model error, surrogate error and posterior uncertainty (nearly negligible) for a calibration of single-site E3SM Land Model given latent heat flux observations at U of Michigan Bio Station

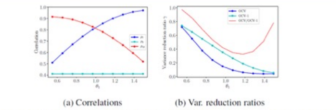


Method: Sargsyan, Najm, Ghanem, IJCK, 2015; Sargsyan, Huan, Najm, IJUQ, 2019

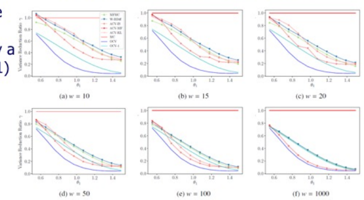
Applications: Huan et. al, AIAA J, 2018; Hakim et. al, CTM, 2018; Cekmer et. al, IJUQ, 2018; Rizzi et. al, CMAME, 2019

Multifidelity strategies for forward UQ

- At the **state-of-the-art** all the multilevel/multifidelity strategies assume ***a priori* a hierarchy** among models
- From a practical standpoint this corresponds to **recursive sampling strategy**
- **A recursive sampling** approach (*e.g.* MLMC and MFMC) **limits** the maximum achievable **variance reduction** to the one obtained by a single known LF model (i.e. OCV-1)
- We **designed** several **Approximate Control Variates (ACV) schemes** that **overcome this issue** and attain a larger variance reduction for a fixed computational budget. They converge to the Optimal Control Variate (OCV) estimator



***Three models tunable test case**: correlations among the three models are controlled by a single parameter. ACV estimators have the potential to converge to the variance reduction of OCV, whereas recursive sampling approaches (MLMC and MFMC) converge only to OCV-1*

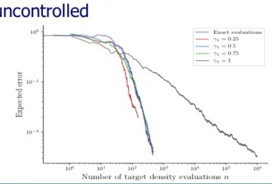


***Variance reduction** obtained for the ACV and recursive estimators under several cost scenarios. The cost of an evaluation is 1 for HF, and $(1/w, 1/w^2)$ for the LF models*

Rate-optimal local approximation MCMC

Continual refinement of surrogate models within MCMC

- Useful for Bayesian inference with computationally intensive models
- Yet previous methods, while asymptotically exact, left balance between surrogate error and Monte Carlo error uncontrolled
- **New approach:**
  - Rate-optimal refinement strategy balances bias and variance of sample estimator at finite time
  - Lyapunov function correction ensures convergence for heavier tails, improves efficiency



FASTMath-UQ partnerships

Partnership project title	Funding offices
Plasma Surface Interactions: Predicting the Performance and Impact of Dynamic Plasma Facing Component Surfaces (PSI2)	SC-FES/ASCR
Optimization of Sensor Networks for Improving Climate Model Predictions (OSCM)	SC-BER/ASCR
Probabilistic Sea Level Projections from Ice Sheet and Earth System Model (ProSPect)	SC-BER/ASCR
Simulation of Fission Gas in Uranium Oxide Nuclear Fuel	NE/ASCR

More Information: http://www.fastmath-scidac.org or contact Habib Najm, Sandia National Labs, hnnajm@sandia.gov, 925-294-2054

Argonne National Laboratory · Berkeley Lab · Lawrence Livermore National Laboratory · Sandia National Laboratories · Oak Ridge National Laboratory · Massachusetts Institute of Technology · Rensselaer · SMU · USC University of Southern California