www.quest-scidac.org

Advanced Algorithms for Uncertainty Quantification and Their Application to Large Scale, High-dimensional, Computational Problems

Kenny Chowdhary, Mike Eldred, John Jakeman, Bert Debusschere, Cosmin Safta, Khachik Sargsyan, Prashant Rai, and Habib Najm
Sandia National Laboratories

Support for this work was provided through the Scientific Discovery through Advanced Computing (SciDAC) project funded by the U.S. Department of Energy, Office of Science, Advanced Scientific Computing Research.
Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia National Laboratories
www.sandia.gov

Model Error

Motivation

- Models rely on physical assumptions and are not perfect
- Model error (discrepancy from truth) is often non-negligible
- Model error estimation is useful for model validation, comparison and reliable predictions
- Explicit statistical modeling of model error has shortcomings

$$y_i = \underbrace{f(x_i;\lambda) + \delta(x_i)}_{\text{truth}} + \epsilon_i^{\mathrm{d}}$$

Embedded Model Error Approach

- Instead, we embed the model error directly into the model
- Cast some inputs as random variables and infer the parameters describing their PDF
- Density estimation problem for Λ
- Can be formulated in a Bayesian framework

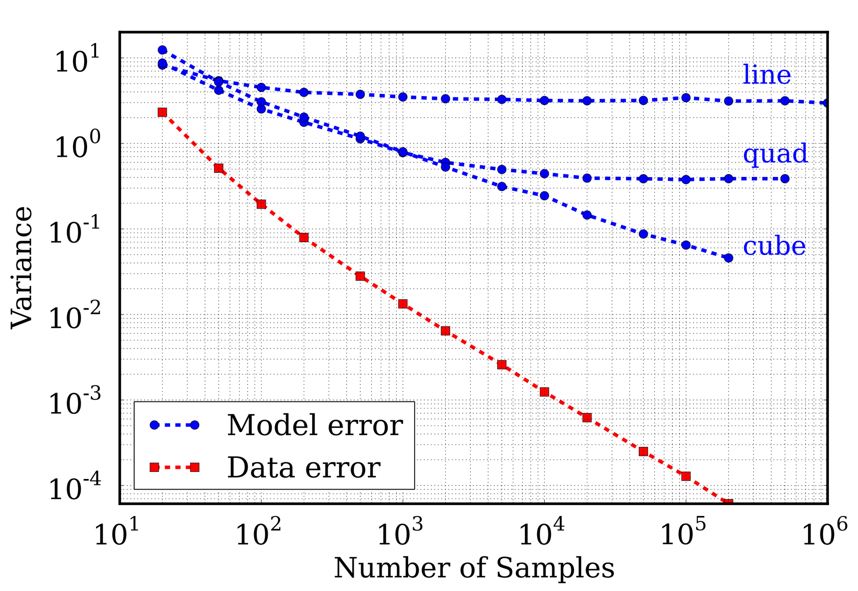
$$y_i = f(x_i;\Lambda) + \epsilon_i^{\mathrm{d}}$$

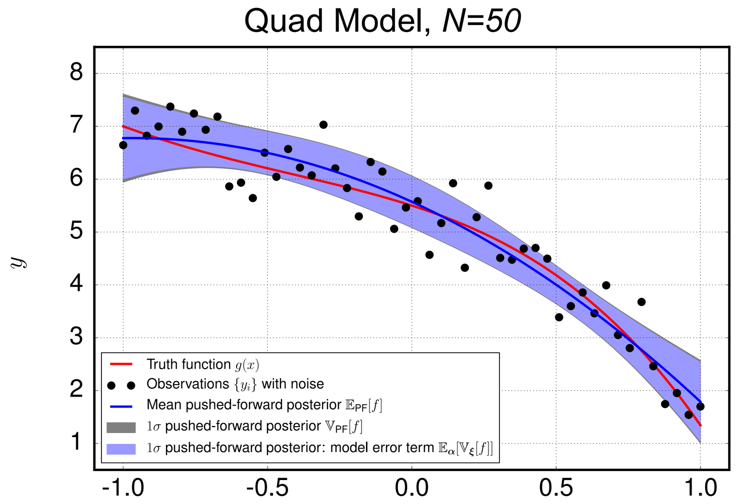
Advantages

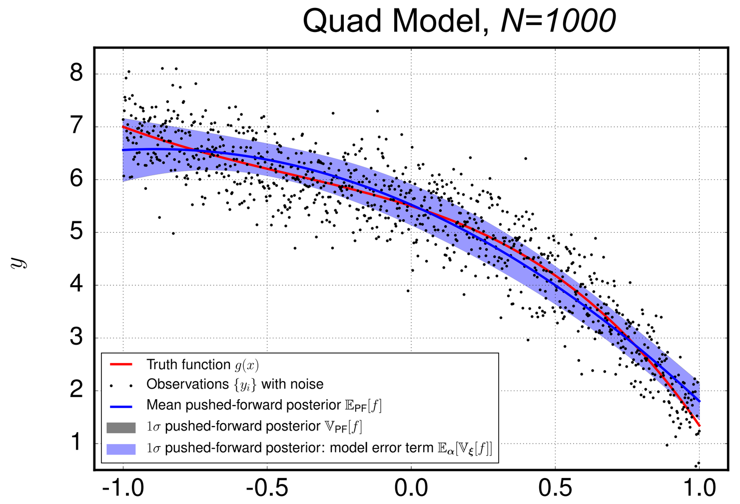
- Allows well-defined model-to-model calibration
- Discrepancy correlations are driven by the model
- Respects physical constraints
- Allows calibrated predictions of multiple QoIs
- Disambiguates model and data error
- Predictive uncertainty is broken into parts due to model error and data noise

Toy example

- Consider "truth" $g(x)=6+x^2+0.5(x+1)^{3.5}$, measured with noise $\sigma=0.5$
- Calibrate three approximate models
  - Line : $f(x,\lambda)=\lambda_0+\lambda_1 x$
  - Quad : $f(x,\lambda)=\lambda_0+\lambda_1 x+\lambda_2 x^2$
  - Cube : $f(x,\lambda)=\lambda_0+\lambda_1 x+\lambda_2 x^2+\lambda_3 x^3$
- Data noise component of predictive uncertainty reduces with more data
- Model error converges to a limiting value, because all three models are wrong!







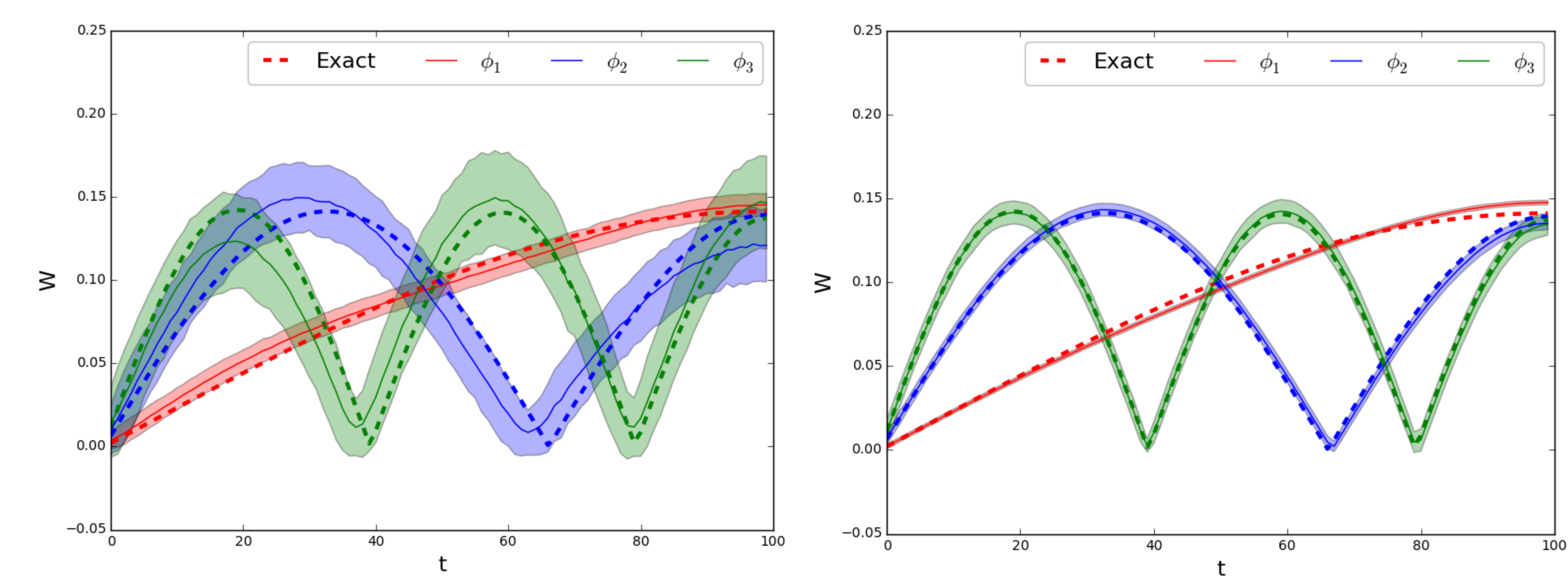
Random Fields

Relevance

- Many applications involve uncertain inputs/outputs that have spatial or time dependence
- Such an uncertain function, represented probabilistically, is a random field/process.
  - It is a random variable at each space/time location
  - Generally with some correlation structure in space/time
  - An infinite-dimensional object
- The Karhunen Loeve expansion (KLE) provides an optimal representation of random fields, employing a (small) number of eigenmodes of its covariance function

Bayesian Estimation of KLE

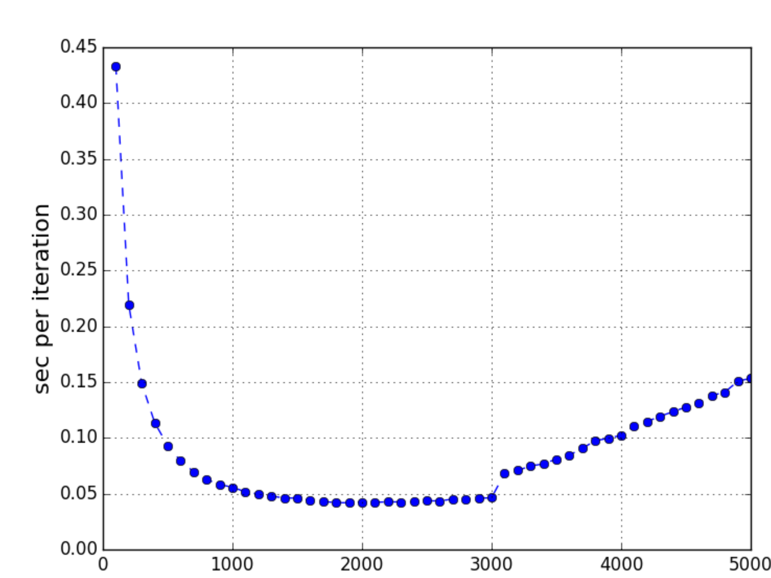
- Uncertain KLE given limited # of samples; Bayesian framework
- Compute principal directions of maximum variance
- Produce error bounds on the principal modes themselves
- The figure shows PCA modes in solid colored lines overlaid with the uncertainty: shaded regions.



More samples, less uncertainty. Uncertainty is shown as the shaded regions.

Large-Scale Parallel SVD Algorithms

- KLE requires eigenvectors of sample covariance matrix.
- We utilize Trilinos' Anasazi package for large-scale eigenvector/eigenvalue parallel algorithms.
- Using John's Hopkins Turbulence database, we analysis time-dependent particle velocities for isotropic turbulence.
- In this example, the random field is the fluid velocity profile at 512^3 data points (~134 million data points).



- This figure shows the seconds per iteration for Anasazi's block Krylov Schur Solver on a 512^3-dimensional data set.
- Note that after about 3000 processors, communication bottlenecks cause slow down.

Hessian informed MCMC

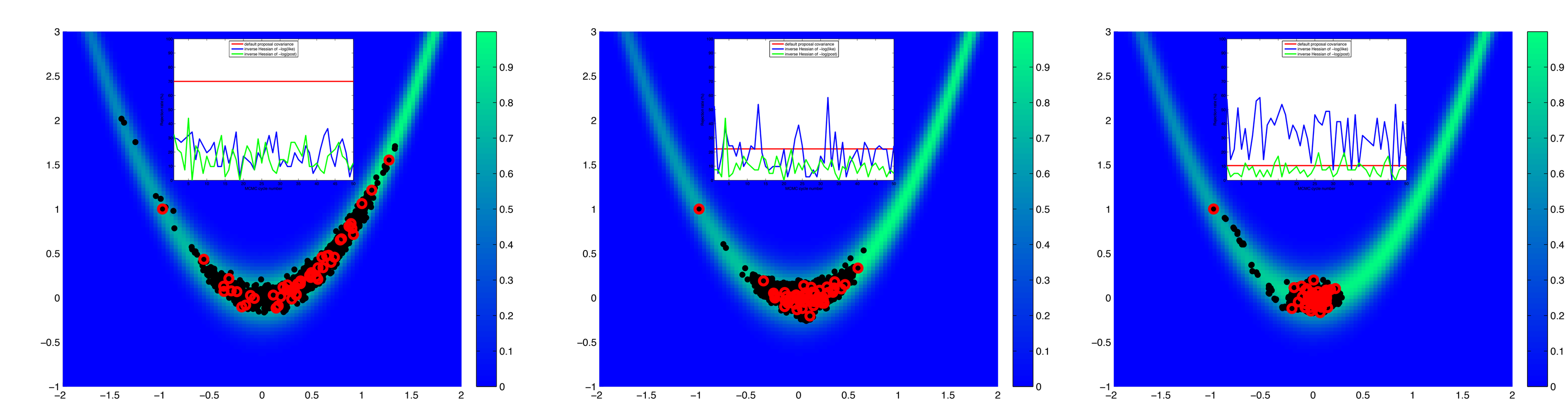
Bayesian Inference

- Bayesian inference is an approach for updating subjective beliefs (i.e., prior assumptions about uncertainty) based on evidence (i.e., observational data).
- Bayes' rule expresses the posterior probability density of ξ conditioned on d as the product of a likelihood model $p(\mathrm{d}|\xi)$, that quantifies the probability that a random variable ξ could produce the observational data d, and $p(\xi)$, the prior density of ξ.

$$p(\xi|\mathbf{d}) \propto p(\mathbf{d}|\xi)\,p(\xi)$$

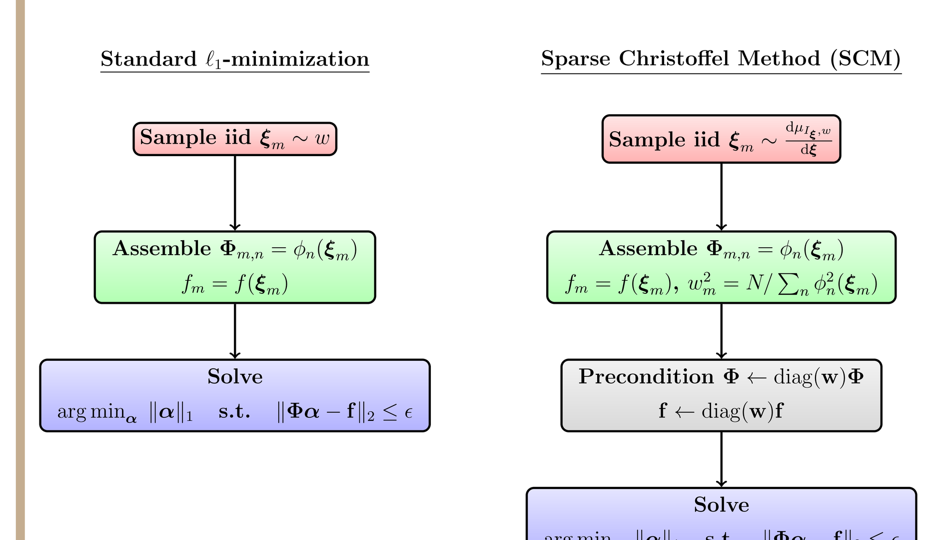
Improving MCMC with adaptive proposals

- MCMC provides a sequential method to sample from the posterior density.
- MCMC needs many samples, mixing time can increase with dimension, and choosing a good proposal can be difficult.
- We use a Gauss-Newton approximation to the Hessian of the posterior (with a PCE emulator in lieu of the true forward model) as a proposal density.

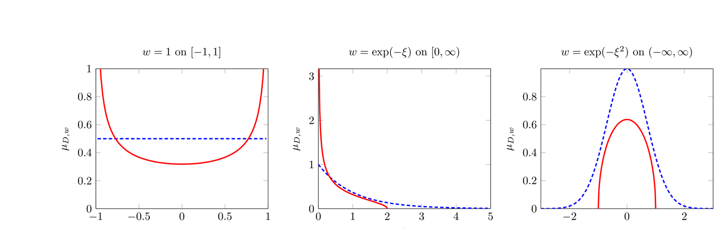


These three figures illustrate MCMC sampling with exact Hessians of the full posterior. We also show rejection rates comparisons between different choice of proposals. Proposals based on Hessians including prior information perform better with lower rejection rates, especially for more compact/local priors.

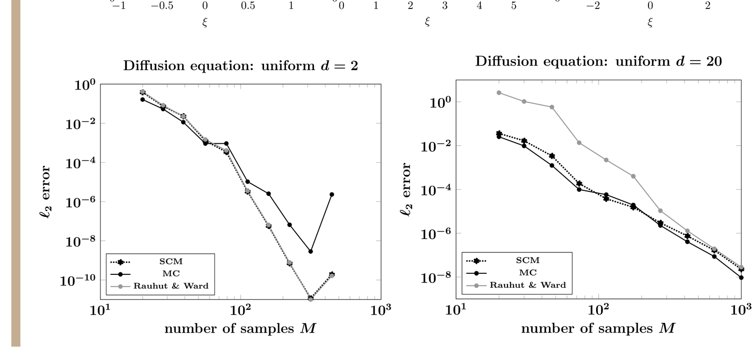
Preconditioned l-1 Sampling for PCE



- We design a sampling scheme with better reconstruction error for a given number of random samples.
- The reconstructed function is a sparse PCE.
- Sampling scheme is based on a change of measure, similar to importance sampling.
- Change of measure reduces the condition number of the Vandermonde matrix and reduces the max norm bound of each column.



- Example change of measure functions (in red) for standard densities. Note that new measures are compact.



- Performs just as good as the state-of-the-art algorithms in low-dimensions, but better in higher dimensions.
- In higher dimensions, Monte Carlo techniques are comparable.

U.S. DEPARTMENT OF ENERGY | Office of Science