Challenges in Uncertainty Quantification in Computational Models

Habib N. Najm

hnnajm@sandia.gov

Sandia National Laboratories,
Livermore, CA

Computational and Theoretical Challenges in
Interdisciplinary Predictive Modeling Over Random Fields
12th Annual Red Raider Mini-Symposium
Department of Mathematics and Statistics
Texas Tech University, Lubbock, TX
October 26, 2012

Acknowledgement

B.J. Debusschere, R.D. Berry, K. Sargsyan, C. Safta,
K. Chowdhary — Sandia National Laboratories, CA

R.G. Ghanem — U. South. California, Los Angeles, CA
O.M. Knio — Duke Univ., Durham, NC
O.P. Le Maître — CNRS, Paris, France
Y.M. Marzouk — Mass. Inst. of Tech., Cambridge, MA

This work was supported by:

- DOE Office of Basic Energy Sciences, Div. of Chem. Sci., Geosci., & Biosci.
- DOE Office of Advanced Scientific Computing Research (ASCR), Scientific Discovery through Advanced Computing (SciDAC)
- DOE ASCR Applied Mathematics program.
- 2009 American Recovery and Reinvesment Act.
- DOE Office of Biological and Environmental Research

Sandia National Laboratories is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under contract DE-AC04-94-AL85000.

Outline

The Case for Uncertainty Quantification (UQ)

UQ is needed in:

- Assessment of confidence in computational predictions
- Validation and comparison of scientific/engineering models
- Design optimization
- Use of computational predictions for decision-support
- Assimilation of observational data and model construction
- Multiscale and multiphysics model coupling

Overview of UQ Methods

Estimation of model/parametric uncertainty

- Expert opinion, data collection
- Regression analysis, fitting, parameter estimation
- Bayesian inference of uncertain models/parameters

Forward propagation of uncertainty in models

- Local sensitivity analysis (SA) and error propagation
- Fuzzy logic; Evidence theory — interval math
- Probabilistic framework — Global SA / stochastic UQ
    - Random sampling, statistical methods
    - Galerkin methods
        - Polynomial Chaos (PC) — intrusive/non-intrusive
    - Collocation, interpolants, regression, fitting ... PC/other

Probabilistic Forward UQ & Polynomial Chaos Representation of Random Variables

- With $y = f(x)$, x a random variable, estimate the RV y
- Can describe a RV in terms of its density, moments, characteristic function, or most fundamentally as a function on a probability space
- Constraining the analysis to RVs with finite variance, enables the representation of a RV as a spectral expansion in terms of orthogonal functions of standard RVs.
  - Polynomial Chaos
- Enables the use of available functional analysis methods for forward UQ

Polynomial Chaos Expansion (PCE)

- Model uncertain quantities as random variables (RVs)
- Given a *germ* $\boldsymbol{\xi}(\omega) = \{\xi_1, \cdots, \xi_n\}$ – a set of *i.i.d.* RVs
  – where $p(\boldsymbol{\xi})$ is uniquely determined by its moments

Any RV in $L^2(\Omega, \mathfrak{S}(\boldsymbol{\xi}), P)$ can be written as a PCE:

$$u(\mathbf{x}, t, \omega) = f(\mathbf{x}, t, \boldsymbol{\xi}) \simeq \sum_{k=0}^{P} u_k(\mathbf{x}, t)\Psi_k(\boldsymbol{\xi}(\omega))$$

– $u_k(\mathbf{x}, t)$ are mode strengths
– $\Psi_k()$ are functions orthogonal w.r.t. $p(\boldsymbol{\xi})$

With dimension n and order p: $$P + 1 = \frac{(n+p)!}{n!p!}$$

Orthogonality

By construction, the functions $\Psi_k()$ are orthogonal with respect to the density of ξ

$$u_k(\boldsymbol{x}, t) = \frac{\langle u\Psi_k \rangle}{\langle \Psi_k^2 \rangle} = \frac{1}{\langle \Psi_k^2 \rangle} \int u(\boldsymbol{x}, t; \lambda(\boldsymbol{\xi}))\, \Psi_k(\boldsymbol{\xi}) p_{\boldsymbol{\xi}}(\boldsymbol{\xi}) d\boldsymbol{\xi}$$

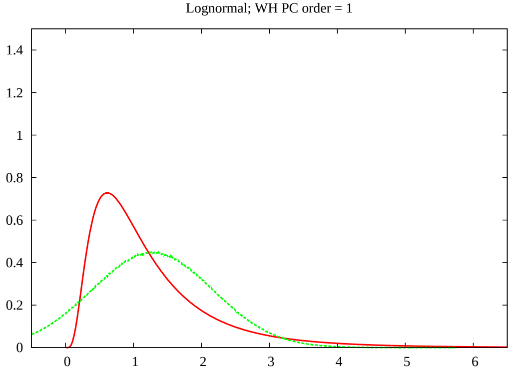
Examples:

- Hermite polynomials with Gaussian basis
- Legendre polynomials with Uniform basis, ...
- Global versus Local PC methods
    - Adaptive domain decomposition of the support of ξ

PC Illustration: WH PCE for a Lognormal RV

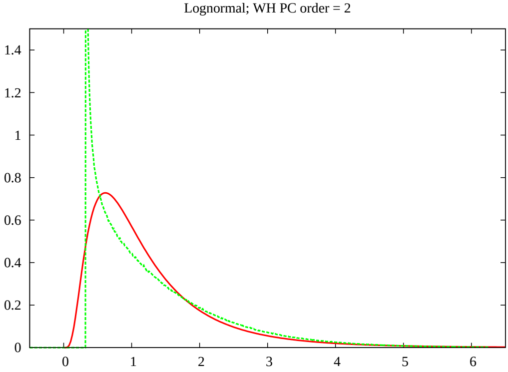
- Wiener-Hermite PCE constructed for a Lognormal RV
- PCE-sampled PDF superposed on true PDF
- Order = 1

$$\begin{aligned} u &= \sum_{k=0}^{P} u_k \Psi_k(\xi) \\ &= u_0 + u_1 \xi \end{aligned}$$



PC Illustration: WH PCE for a Lognormal RV

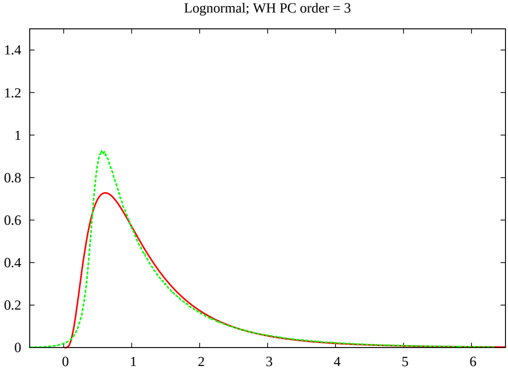
- Wiener-Hermite PCE constructed for a Lognormal RV
- PCE-sampled PDF superposed on true PDF
- Order = 2



$$
\begin{aligned}
u &= \sum_{k=0}^{P} u_k \Psi_k(\xi) \\
&= u_0 + u_1 \xi + u_2(\xi^2 - 1)
\end{aligned}
$$

PC Illustration: WH PCE for a Lognormal RV

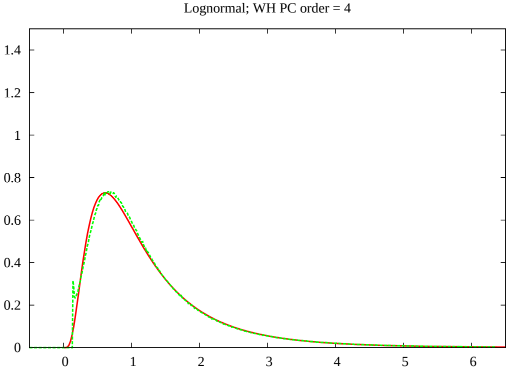
- Wiener-Hermite PCE constructed for a Lognormal RV
- PCE-sampled PDF superposed on true PDF
- Order = 3



$$
\begin{aligned}
u &= \sum_{k=0}^{P} u_k \Psi_k(\xi) \\
&= u_0 + u_1\xi + u_2(\xi^2 - 1) + u_3(\xi^3 - 3\xi)
\end{aligned}
$$

PC Illustration: WH PCE for a Lognormal RV

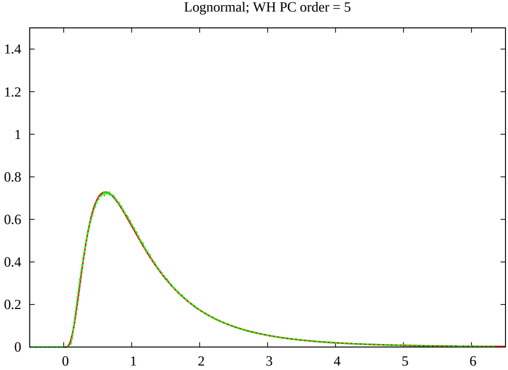
- Wiener-Hermite PCE constructed for a Lognormal RV
- PCE-sampled PDF superposed on true PDF
- Order = 4



$$
\begin{aligned}
u &= \sum_{k=0}^{P} u_k \Psi_k(\xi) \\
&= u_0 + u_1\xi + u_2(\xi^2 - 1) + u_3(\xi^3 - 3\xi) + u_4(\xi^4 - 6\xi^2 + 3)
\end{aligned}
$$

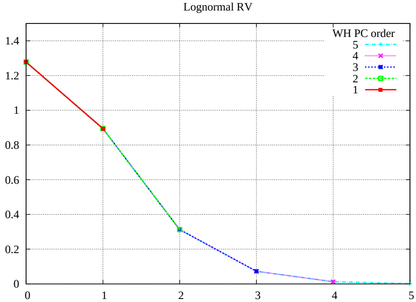
PC Illustration: WH PCE for a Lognormal RV

- Wiener-Hermite PCE constructed for a Lognormal RV
- PCE-sampled PDF superposed on true PDF
- Order = 5



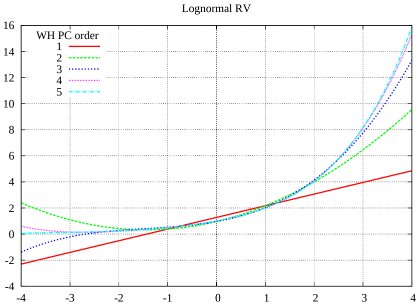
$$
\begin{aligned}
u &= \sum_{k=0}^{P} u_k \Psi_k(\xi) \\
&= u_0 + u_1\xi + u_2(\xi^2 - 1) + u_3(\xi^3 - 3\xi) + u_4(\xi^4 - 6\xi^2 + 3) \\
&\quad + u_5(\xi^5 - 10\xi^3 + 15\xi)
\end{aligned}
$$

PC Illustration: WH PCE for a Lognormal RV



PC mode amplitudes

u_0–u_5



PC function $u(\xi)$

Order 1–5

- Fifth-order Wiener-Hermite PCE represents the given Lognormal well
- Higher order terms are negligible

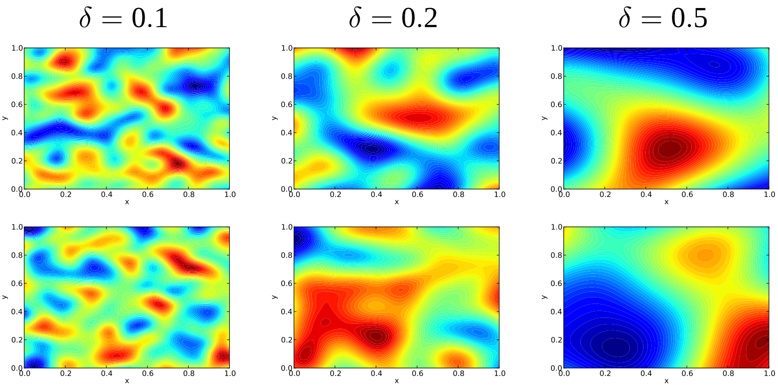
Random Fields – KLE

- Smooth random fields can be repesented with a small no. of stochastic degrees of freedom
- Karhunen-Loeve Expansion (KLE) for a RF with a continuous covariance function

$$M(x,\omega) = \mu(x) + \sum_{i=1}^{\infty} \sqrt{\lambda_i}\eta_i(\omega)\phi_i(x)$$

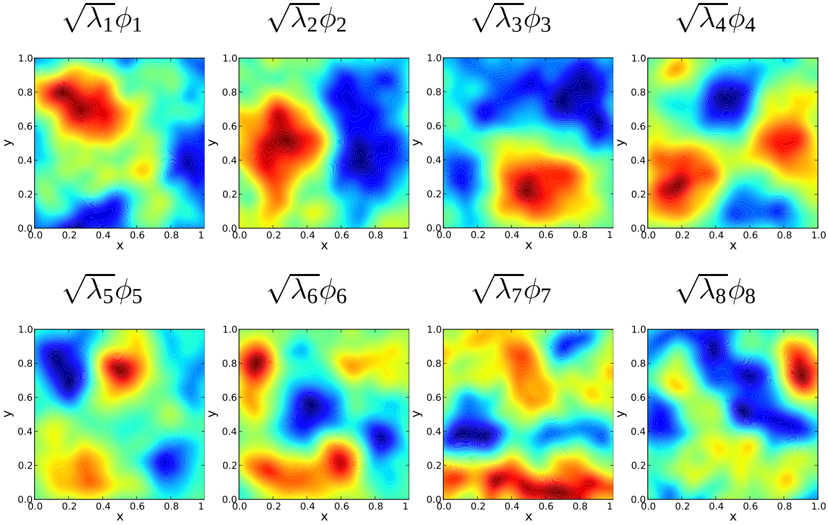
- $\mu(x)$ is the mean of $M(x,\omega)$ at x
- λ_i and $\phi_i(x)$ are the eigenvalues and eigenfunctions of the covariance $C(x_1,x_2) = \langle [M(x_1,\omega) - \mu(x_1)][M(x_2,\omega) - \mu(x_2)] \rangle$
- The η_i are uncorrelated zero-mean unit-variance RVs
- KLE $\Rightarrow$ representation of random fields using PC

RF Illustration: KL of 2D Gaussian Process

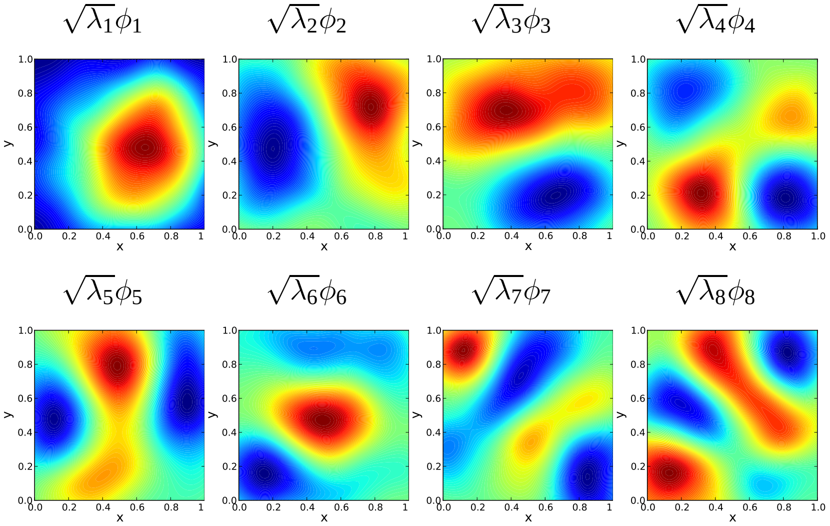


- 2D Gaussian Process with covariance:
  $Cov(x_1, x_2) = exp(-||x_1 - x_2||^2/\delta^2)$
- Realizations smoother as covariance length δ increases

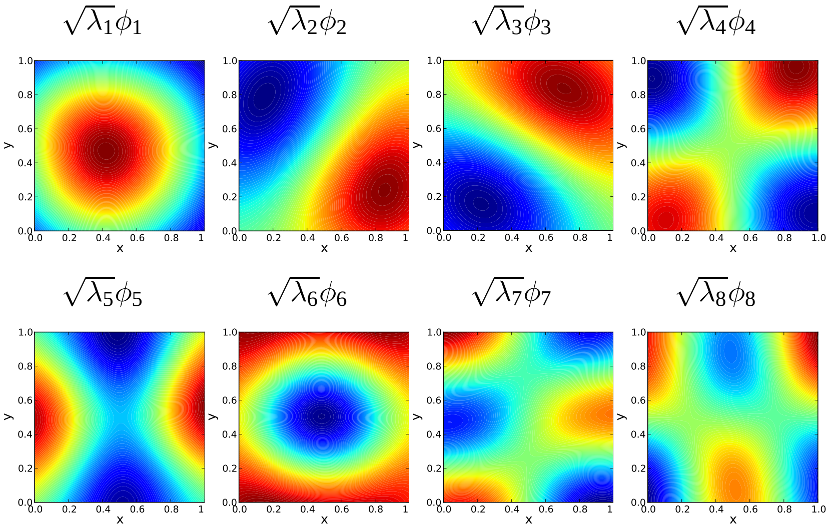
RF Illustration: 2D KL - Modes for $\delta = 0.1$



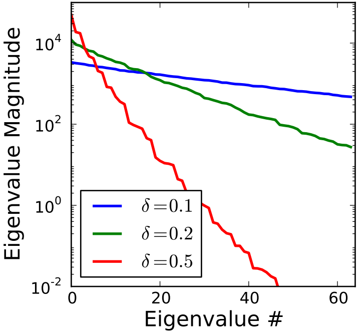
RF Illustration: 2D KL - Modes for $\delta = 0.2$



RF Illustration: 2D KL - Modes for $\delta = 0.5$

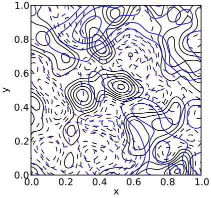


RF Illustration: 2D KL - eigenvalue spectrum

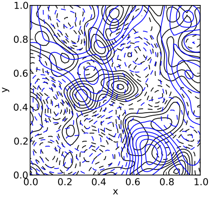


$\delta = 0.1$

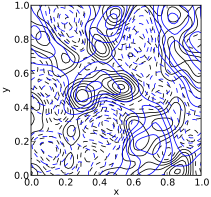
4 terms



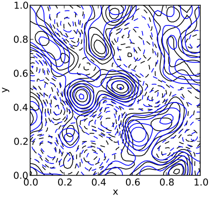
16 terms



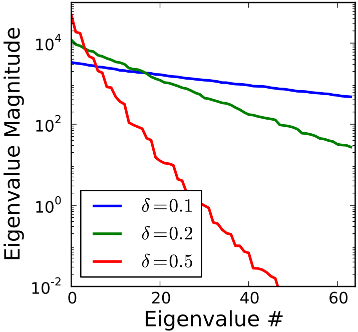
32 terms

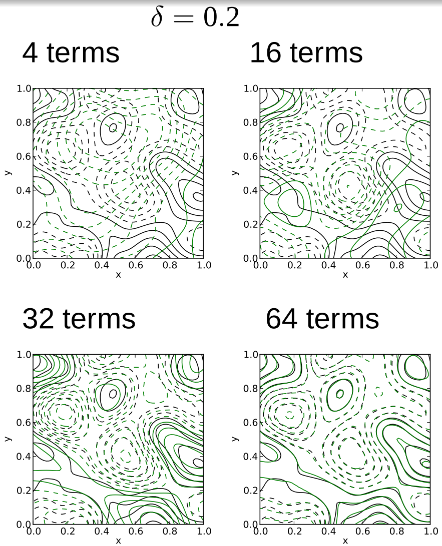


64 terms

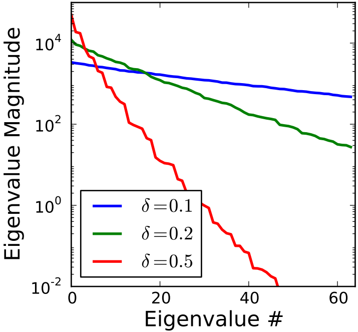


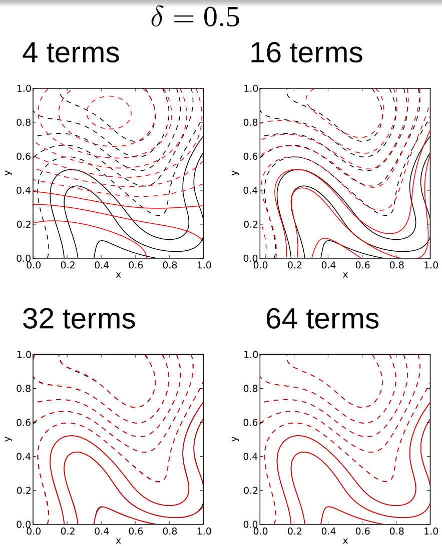
RF Illustration: 2D KL - eigenvalue spectrum





RF Illustration: 2D KL - eigenvalue spectrum





Essential Use of PC in UQ

Strategy:

- Represent model parameters/solution as random variables
- Construct PCEs for uncertain parameters
- Evaluate PCEs for model outputs

Advantages:

- Computational efficiency
- Sensitivity information

Requirement:

- Random variables in L^2, i.e. with finite variance

Intrusive PC UQ: A direct *non-sampling* method

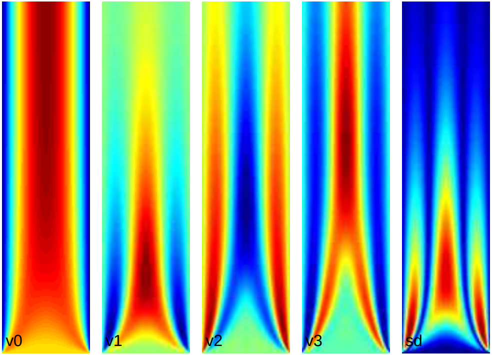
- Given model equations: $\mathcal{M}(u(\mathbf{x},t);\lambda) = 0$
- Express uncertain parameters/variables using PCEs

$$u = \sum_{k=0}^{P} u_k \Psi_k; \quad \lambda = \sum_{k=0}^{P} \lambda_k \Psi_k$$

- Substitute in model equations; apply Galerkin projection
- New set of equations: $\mathcal{G}(U(\mathbf{x},t),\Lambda) = 0$
  - with $U = [u_0, \dots, u_P]^T$, $\Lambda = [\lambda_0, \dots, \lambda_P]^T$
- Solving this system *once* provides the full specification of uncertain model ouputs

Laminar 2D Channel Flow with Uncertain Viscosity

- Incompressible flow
- Viscosity PCE
  - $\nu = \nu_0 + \nu_1 \xi$
- Streamwise velocity
  - $\mathrm{v} = \sum_{i=0}^{P} \mathrm{v}_i \Psi_i$
  - v_0: mean
  - v_i: i-th order mode
  - $\sigma^2 = \sum_{i=1}^{P} \mathrm{v}_i^2 \langle \Psi_i^2 \rangle$



v_0 v_1 v_2 v_3 σ

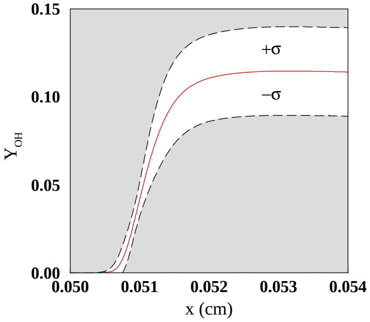
Non-intrusive Spectral Projection (NISP) PC UQ

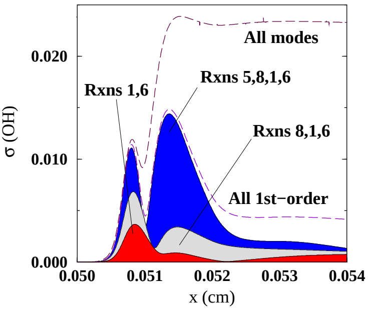
- *Sampling*-based
- Relies on black-box utilization of the computational model
- Evaluate projection integrals *numerically*
- For any quantity of interest $\phi(\boldsymbol{x}, t; \lambda) = \sum_{k=0}^{P} \phi_k(\boldsymbol{x}, t)\Psi_k(\boldsymbol{\xi})$

$$\phi_k(\boldsymbol{x}, t) = \frac{1}{\langle \Psi_k^2 \rangle} \int \phi(\boldsymbol{x}, t; \lambda(\boldsymbol{\xi}))\, \Psi_k(\boldsymbol{\xi}) p_{\boldsymbol{\xi}}(\boldsymbol{\xi}) d\boldsymbol{\xi}, \quad k = 0, \ldots, P$$

- Integrals can be evaluated using
  - A variety of (Quasi) Monte Carlo methods
    - Slow convergence; $\sim$ indep. of dimensionality
  - Quadrature/Sparse-Quadrature methods
    - Fast convergence; depends on dimensionality

1D H_2-O_2 SCWO Flame NISP UQ/Chemkin-Premix





- Fast growth in OH uncertainty in the primary reaction zone
- Constant uncertainty and mean of OH in post-flame region
- Uncertainty in pre-exponential of Rxn.5 (H_2O_2+OH=H_2O+HO_2) has largest contribution to uncertainty in predicted OH

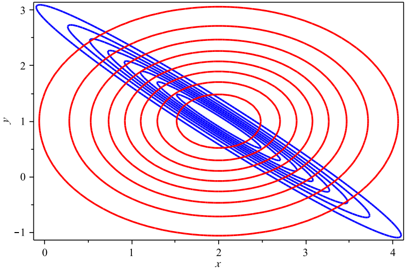
Other non-intrusive methods

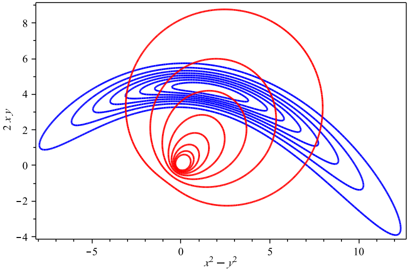
- Response surface employing PC or other functional basis
- Collocation: Fit interpolant to samples
  - Oscillation concern
- Regression: Estimate best-fit response surface
  - Least-squares
  - Bayesian inference
- Useful when quadrature methods are infeasible, e.g. when
  - Can’t choose sample locations; samples given *a priori*
  - Can’t take enough samples
  - Forward model is noisy

Inverse UQ – Estimation of Uncertain Inputs

- Forward UQ requires specification of uncertain inputs
- Probabilistic setting
  - Require joint PDF on input space
- Bayesian setting
  - PDF on uncertain inputs can be found given data
  - Probabilistic inference – an inverse problem
- Uncertainty in computational predictions can depend strongly on detailed structure of the parametric PDF

The strong role of detailed input PDF structure





- Simple nonlinear algebraic model $(u, v) = (x^2 - y^2, 2xy)$
- Two input PDFs, $p(x, y)$
  - same nominals/bounds
  - different correlation structure
- Drastically different output PDFs
  - different nominals and bounds

Bayes formula for Parameter Inference

- Data Model (fit model + noise model): $y = f(\lambda) * g(\epsilon)$
- Bayes Formula:

$$p(\lambda, y) = p(\lambda|y)p(y) = p(y|\lambda)p(\lambda)$$

$$\underbrace{p(\lambda|y)}_{\text{Posterior}} = \frac{\overbrace{p(y|\lambda)}^{\text{Likelihood}} \; \overbrace{p(\lambda)}^{\text{Prior}}}{\underbrace{p(y)}_{\text{Evidence}}}$$

- Prior: knowledge of λ prior to data
- Likelihood: forward model and measurement noise
- Posterior: combines information from prior and data
- Evidence: normalizing constant for present context

Exploring the Posterior

- Given any sample λ, the un-normalized posterior probability can be easily computed

$$p(\lambda|y) \propto p(y|\lambda)p(\lambda)$$

- Explore posterior w/ Markov Chain Monte Carlo (MCMC)
  - Metropolis-Hastings algorithm:
    - Random walk with proposal PDF & rejection rules
  - Computationally intensive, $\mathcal{O}(10^5)$ samples
  - Each sample: evaluation of the forward model
    - Surrogate models
- Evaluate moments/marginals from the MCMC statistics

UQ Challenges - Characterization of Uncertain Inputs

- Computational Model $\mathcal{M}(u, \lambda) = 0$
  - Uncertain input parameter λ
  - Experimental Measurement $F(y, \lambda) = 0$
- Uncertain model inputs can be estimated from data on y
  - Regression
  - Bayesian inference
- Quite frequently, we have partial data/information
  - Partial missing data, *e.g.* failed measurements
  - Full data loss – No data, but have summary information, *e.g.* moments and/or quantiles on
    - data – processed data products
    - fitted parameters

Dealing with Partial Data – Imputation

Bayesian Multiple Imputation (Rubin, 1987)

- Use the posterior predictive conditioned on *observed* data to generate replicates of the missing data z^t_{mis}, $t = 1, \ldots, m$
- For each full data set, $(z_{\mathrm{obs}}, z^t_{\mathrm{mis}})$ apply Bayesian inference to get a posterior density on the model parameters
- Marginalize over the missing data to get the observed data posterior $p(\beta|z_{\mathrm{obs}})$

$$p(\beta|z_{\mathrm{obs}}) = \int q(\beta|z_{\mathrm{mis}}, z_{\mathrm{obs}}) f(z_{\mathrm{mis}}|z_{\mathrm{obs}}) dz_{\mathrm{mis}}$$

Parameter Estimation in the Absence of Data

- Frequently:
  - we know summary statistics about data or parameters from previous work
  - the raw data used to arrive at these statistics is not available
- How can we construct a joint PDF on the parameters?
- In the absence of data, the structure of the fit model, combined with the summary statistics, implicitly inform the joint PDF on the parameters
- Goal: Make available information *explicit* in the joint PDF

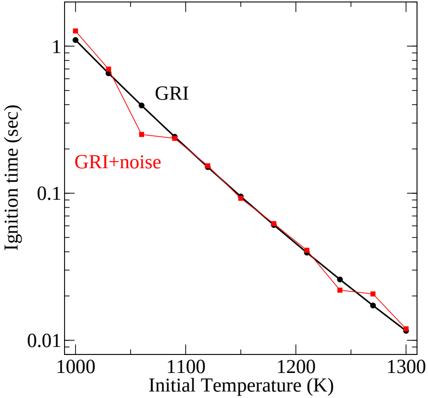
Data Free Inference (DFI)

- Discover a consensus joint PDF on the parameters consistent with given information in the absence of data

Generate ignition "data" using a detailed model+noise

- Ignition using a detailed chemical model for methane-air chemistry
- Ignition time versus Initial Temperature
- Multiplicative noise error model
- 11 data points:

$$
\begin{aligned}
d_i &= t_{\mathrm{ig},i}^{\mathsf{GRI}}(1+\sigma\epsilon_i) \\
\epsilon &\sim N(0,1)
\end{aligned}
$$



Fitting with a simple chemical model

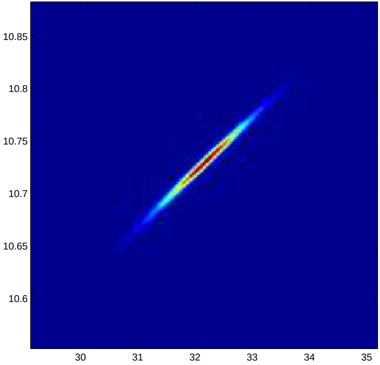
- Fit a global single-step irreversible chemical model

$$CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O$$

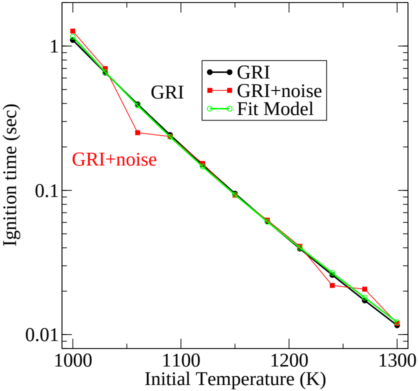
$$\begin{aligned} \Re &= [CH_4][O_2]k_f \\ k_f &= A\exp(-E/R^oT) \end{aligned}$$

- Infer 3-D parameter vector $(\ln A, \ln E, \ln\sigma)$
- Use an adaptive MCMC procedure
- Start at the maximum likelihood estimate

Bayesian Inference Posterior and Nominal Prediction



Marginal joint posterior on $(\ln A, \ln E)$ exhibits strong correlation



Nominal fit model is consistent with the true model

DFI Challenge — Chemical Model Problem

Discarding initial data, reconstruct marginal $(\ln A, \ln E)$ posterior using the following information

- Form of fit model
- Range of initial temperature
- Nominal fit parameter values of $\ln A$ and $\ln E$
- Marginal 5% and 95% quantiles on $\ln A$ and $\ln E$

Further, for now, presume

- Multiplicative Gaussian errors
- $N = 8$ data points

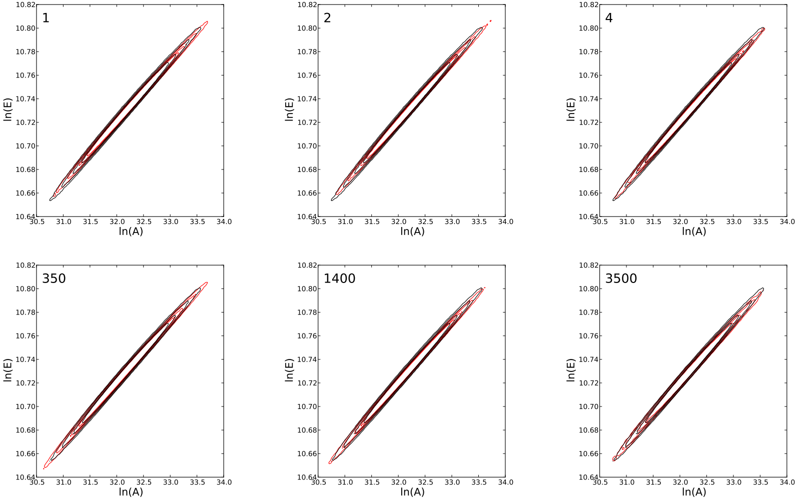
DFI Algorithm Structure

Basic idea:

- Explore the space of hypothetical data sets
  - MCMC chain on the data
  - Each state defines a data set
- For each data set:
  - MCMC chain on the parameters
  - Evaluate statistics on resulting posterior
  - Accept data set if posterior is consistent with given information
- Evaluate pooled posterior from all acceptable posteriors
  Logarithmic pooling:

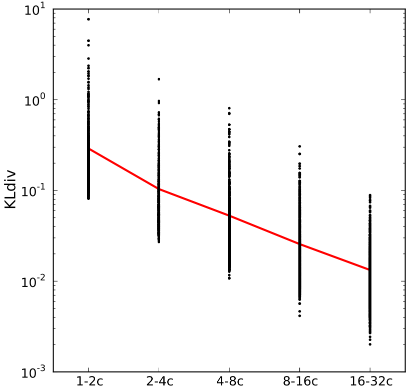
$$p(\lambda|y) = \left[\prod_{i=1}^{K} p(\lambda|y_i)\right]^{1/K}$$

2D Marginal Pooled Posteriors vs. Reference Posterior



Empirical Convergence of 3D Pooled Posteriors

- Kullback-Leibler divergence between posteriors of successively increased data volumes
- Combinatorial choices of chains pooled at each stage
  - statistical scatter of KLdiv
- Overall convergence evident $\propto 1/N$



Challenges in PC UQ – High-Dimensionality

- Dimensionality n of the PC basis: $\boldsymbol{\xi} = \{\xi_1, \ldots, \xi_n\}$
  - number of degrees of freedom
  - $P+1 = (n+p)!/n!p!$ grows fast with n
- Impacts:
  - Size of intrusive system
  - # non-intrusive (sparse) quadrature samples
- Generally $n \approx$ number of uncertain parameters
- Reduction of n:
  - Sensitivity analysis
  - Dependencies/correlations among parameters
  - Dominant eigenmodes of random fields
  - Manifold learning: Isomap, Diffusion maps
  - Sparsification: Compressed Sensing, LASSO

PC Quadrature in hiD

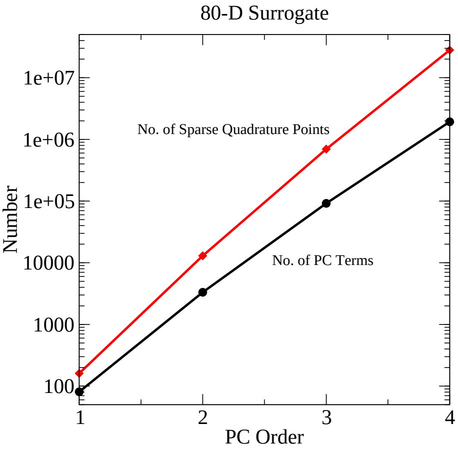
Full quadrature: $N = (N_{1D})^n$

Sparse Quadrature

- Wide range of methods
- Nested & hierarchical
- Clenshaw-Curtis: $N = \mathcal{O}(n^p)$
- Adaptive – greedy algorithms

Number of points can still be excessive in hi-D

- Large no. of terms
- Reduction/sparsity



PC coefficients via sparse regression

PCE:

$$y = f(x) \simeq \sum_{k=0}^{K-1} c_k \Psi_k(x)$$

with $x \in \mathbb{R}^n$, Ψ_k max order p, and $K = (p+n)!/p!/n!$

- N samples $(x_1, y_1), \ldots, (x_N, y_N)$
- Estimate K terms $c_0, \ldots, c_{K-1}$, s.t.

$$\min \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{c}\|_2^2$$

where $\boldsymbol{y} \in \mathbb{R}^N$, $\boldsymbol{c} \in \mathbb{R}^K$, $\boldsymbol{A}_{ik} = \Psi_k(x_i)$, $\boldsymbol{A} \in \mathbb{R}^{N\times K}$

With $N << K \Rightarrow$ under-determined

- Need some form of regularization

Regularization – Compressive Sensing (CS)

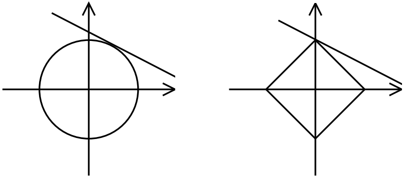
- ℓ_2-norm — Tikhonov regularization; Ridge regression:

$$\min \{\|y - Ac\|_2^2 + \|c\|_2^2\}$$

- ℓ_1-norm — Compressive Sensing; LASSO; basis pursuit

$$\begin{aligned} \min \{\|y - Ac\|_2^2 + \|c\|_1\} & \\ \min \{\|y - Ac\|_2^2\} & \quad \text{subject to } \|c\|_1 \le \epsilon \\ \min \{\|c\|_1\} & \quad \text{subject to } \|y - Ac\|_2^2 \le \epsilon \end{aligned}$$

$\Rightarrow$ discovery of sparse signals



Bayesian Regression

- Bayes formula

$$p(\boldsymbol{c}|D) \propto p(D|\boldsymbol{c})\pi(\boldsymbol{c})$$

- Bayesian regression: prior as a regularizer, *e.g.*
  - Log Likelihood $\Leftrightarrow$ $\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{c}\|_2^2$
  - Log Prior $\Leftrightarrow$ $\|\boldsymbol{c}\|_p^p$
- Laplace sparsity priors $\pi(c_k|\alpha) = \frac{1}{2\alpha} e^{-|c_k|/\alpha}$
- LASSO (Tibshirani 1996) ... formally:

$$\min \{\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{c}\|_2^2 + \lambda\|\boldsymbol{c}\|_1\}$$

Solution $\sim$ the posterior mode of $\boldsymbol{c}$ in the Bayesian model

$$y \sim \mathcal{N}(\boldsymbol{A}\boldsymbol{c}, I_N), \qquad c_k \sim \frac{1}{2\alpha} e^{-|c_k|/\alpha}$$

- Bayesian LASSO (Park & Casella 2008)

Bayesian Compressive Sensing (BCS)

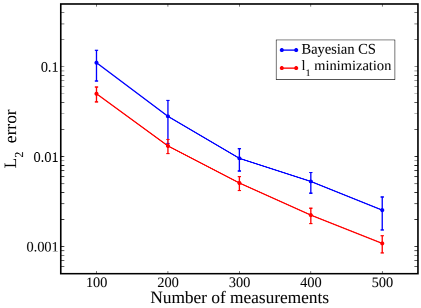
- BCS (Ji 2008; Babacan 2010)— hierarchical priors:
  - Gaussian priors $\mathcal{N}(0, \sigma_k^2)$ on the c_k
  - Gamma priors on the σ_k^2

  $\Rightarrow$ Laplace sparsity priors on the c_k
- Evidence maximization establishes ML estimates of the σ_k
  - many of which are found $\approx 0 \quad \Rightarrow \quad c_k \approx 0$
  - iteratively include terms that lead to the largest increase in the evidence
- iterative BCS (iBCS) (Sargsyan 2012):
  - adaptive iterative order growth
  - BCS on order-p Legendre-Uniform PC
  - repeat with order-$p+1$ terms added to surviving p-th order terms

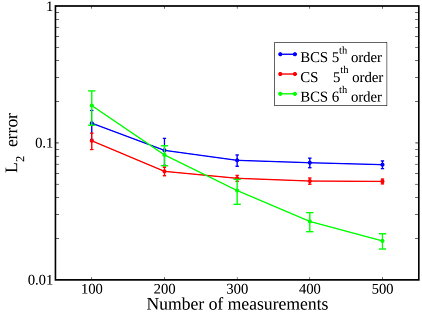
CS and BCS

Oscillatory Genz function

- $f(x) = \cos(2\pi r + \sum_{i=1}^{n} a_i x_i); \quad a_i \propto 1/i^2; \quad r = 0$
- Legendre-Uniform PC, 10^{th}-order in 5d; $(5,6)^{th}$-order in 10d

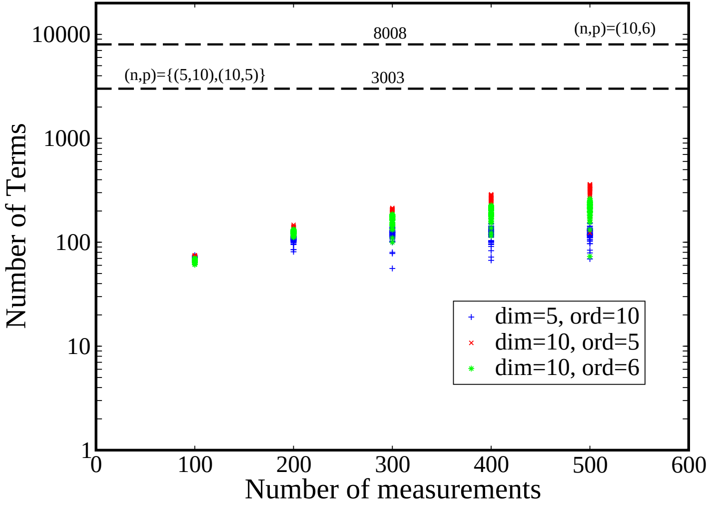


5d



10d

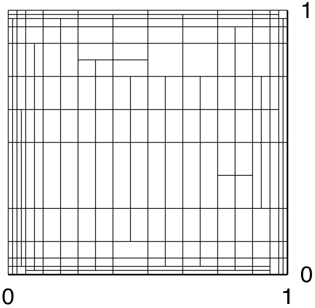
Oscillatory function – BCS number of terms

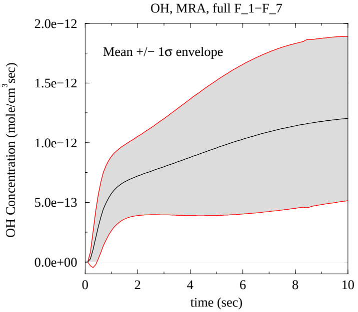


Challenges in PC UQ – Non-Linearity

- Bifurcative response at critical parameter values
  - Rayleigh-Bénard convection
  - Transition to turbulence
  - Chemical ignition
- Discontinuous $u(\lambda(\xi))$
  - Failure of global PCEs in terms of smooth $\Psi_k()$
  - $\Leftrightarrow$ failure of Fourier series in representing a step function
- Local PC methods
  - Subdivide support of $\lambda(\xi)$ into regions of smooth $u \circ \lambda(\xi)$
  - Employ PC with compact support basis on each region
  - A spectral-element vs. spectral construction
  - Domain mapping

Multi-Block Multiwavelet PC UQ in Ignition



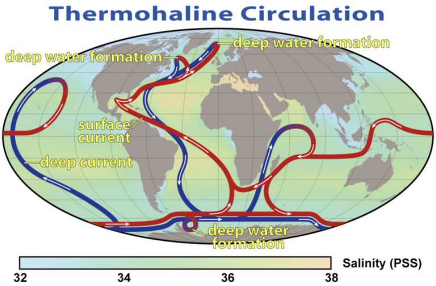


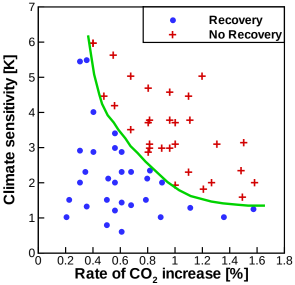
- H_2-O_2 supercritical water oxidation model
- Empirically-based uncertainty in all 7 reactions
- Adaptive refinement of MW block decomposition

(Le Maître, 2004, 2007)

- Challenges in hi-dimensional contexts

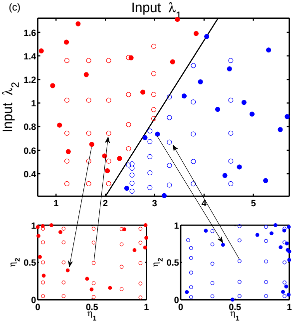
Uncertainty in Discontinuous Climate Response

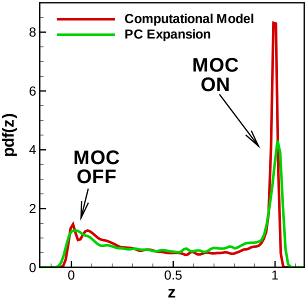




- Atlantic meridional ocean circulation (AMOC)
- Predicted response to increasing CO_2 (Webster, 2007)
- Circulation ON/OFF response over parameter space
  - Rate of CO_2 increase
  - Climate sensitivity

Domain Mapping for Discontinuous Response





- Initial set of computational samples
- Discover uncertain discontinuity with Bayesian inference
- Map sub-domains to unit hypercubes; Rosenblatt transform
- PC quadrature in mapped domains; map back
- Marginalize over uncertain curve (Sargsyan, 2012)

Challenges in PC UQ – Time Dynamics

- Systems with limit-cycle or chaotic dynamics
- Large amplification of phase errors over long time horizon
- PC order needs to be increased in time to retain accuracy
- Time shifting/scaling remedies

- Futile to attempt representation of detailed turbulent velocity field $\boldsymbol{v}(\boldsymbol{x}, t; \lambda(\boldsymbol{\xi}))$ as a PCE
  - Fast loss of correlation due to energy cascade
  - Problem studied in 60's and 70's
- Focus on flow statistics, *e.g.* Mean/RMS quantities
  - Well behaved
  - Argues for non-intrusive methods with DNS/LES of turbulent flow

Closure

- Probabilistic UQ framework
- Forward UQ
  - Polynomial Chaos representation of random variables
  - Intrusive and non-intrusive forward PC UQ methods
- Inverse UQ
  - Bayesian methods
- Highlighted UQ Challenges
  - Missing data – Imputation / DFI
  - High dimensionality
  - Non-linearity
  - Long term oscillatory dynamics