SAND2016-1683C

Uncertainty Quantification in Computational Models

Habib N. Najm

hnnajm@sandia.gov

Sandia National Laboratories
Livermore, CA, USA

Computational Infrastructure for Geodynamics (CIG)
Webinar, UC Davis, CA
Feb 11, 2016

Acknowledgement

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

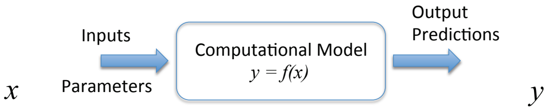
R.G. Ghanem – U. South. California, Los Angeles, CA
O.M. Knio – Duke Univ., Durham, NC
O.P. Le Maître – CNRS, Paris, France
Y.M. Marzouk – Mass. Inst. of Tech., Cambridge, MA

This work was supported by:

- DOE Office of Advanced Scientific Computing Research (ASCR), Scientific Discovery through Advanced Computing (SciDAC)
- DOE ASCR Applied Mathematics program.
- DOE Office of Basic Energy Sciences, Div. of Chem. Sci., Geosci., & Biosci.

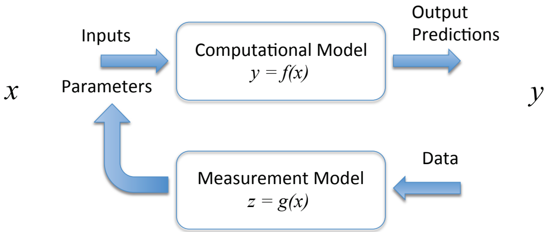
Sandia National Laboratories is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under contract DE-AC04-94-AL85000.

Uncertainty Quantification and Computational Science



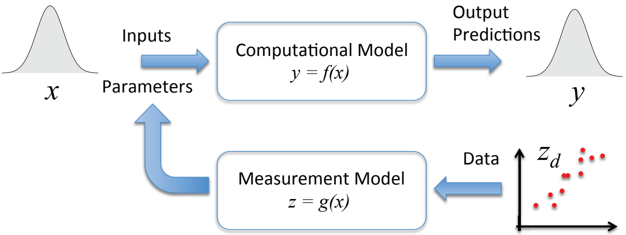
Forward problem

Uncertainty Quantification and Computational Science



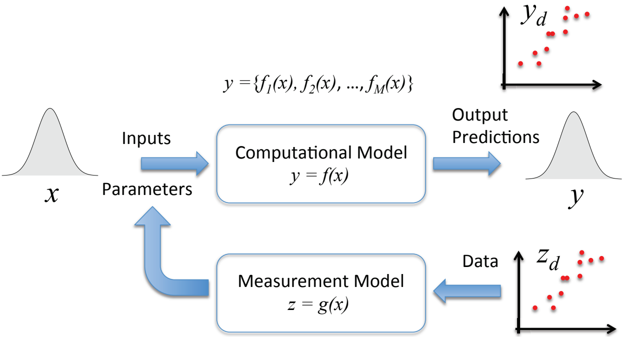
Inverse & Forward problems

Uncertainty Quantification and Computational Science



Inverse & Forward UQ

Uncertainty Quantification and Computational Science



Inverse & Forward UQ

Model validation & comparison, Hypothesis testing

Outline

Forward propagation of parametric uncertainty

Forward model: $y = f(x)$

- Local sensitivity analysis (SA) and error propagation

$$\Delta y = \frac{\mathrm{d}f}{\mathrm{d}x}\bigg|_{x_0} \Delta x$$

  This is ok for:

  - small uncertainty
  - low degree of non-linearity in $f(x)$

- Non-probabilistic methods
  - Fuzzy logic
  - Evidence theory – Dempster-Shafer theory
  - Interval math

- Probabilistic methods – this is our focus

Probabilistic Forward UQ – $y = f(x)$

Represent uncertain quantities using probability theory

- Random sampling, MC, QMC
  - Generate random samples $\{x^i\}_{i=1}^N$ from the PDF of x, $p(x)$
  - Bin the corresponding $\{y^i\}$ to construct $p(y)$
  - Not feasible for computationally expensive $f(x)$
    - slow convergence of MC/QMC methods
      $\Rightarrow$ very large N required for reliable estimates
- Build a cheap surrogate for $f(x)$, then use MC
  - Collocation – interpolants
  - Regression – fitting
- Galerkin methods
  - Polynomial Chaos (PC)
  - Intrusive and non-intrusive PC methods

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
  - as a function on a probability space
- Constraining the analysis to RVs with finite variance

  $\Rightarrow$ Represent RV as a spectral expansion in terms of orthogonal functions of standard RVs
  - Polynomial Chaos Expansion
- 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:

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

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

Orthogonality

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

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

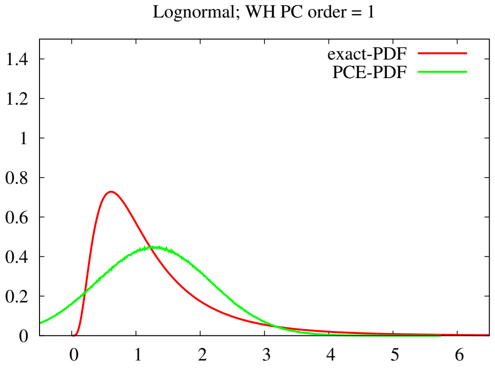
Examples:

- 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

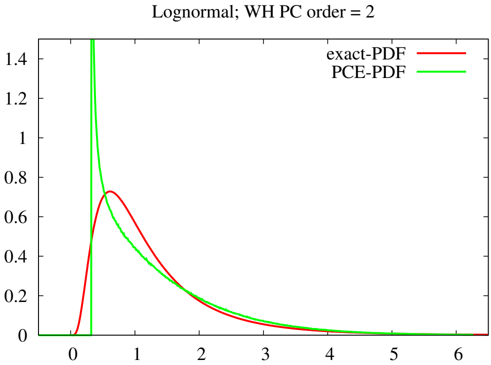
$$
\begin{aligned}
u &= \sum_{k=0}^{P} u_k \Psi_k(\xi) \\
&= u_0 + u_1 \xi
\end{aligned}
$$



PC Illustration: WH PCE for a Lognormal RV

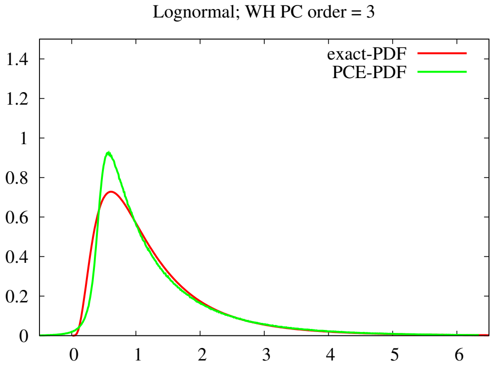
- Wiener-Hermite PCE constructed for a Lognormal RV
- PCE-sampled PDF superposed on true PDF
- Order = 2

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



PC Illustration: WH PCE for a Lognormal RV

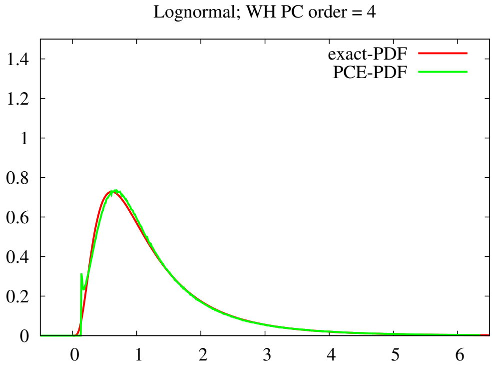
- Wiener-Hermite PCE constructed for a Lognormal RV
- PCE-sampled PDF superposed on true PDF
- Order = 3



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

PC Illustration: WH PCE for a Lognormal RV

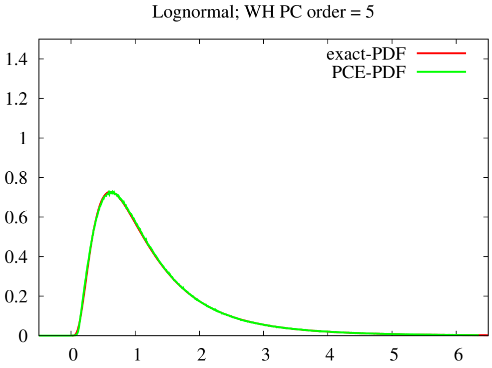
- Wiener-Hermite PCE constructed for a Lognormal RV
- PCE-sampled PDF superposed on true PDF
- Order = 4



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

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



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

Random Fields

- A random variable is a function on an event space Ω
  - No dependence on other coordinates –*e.g.* space or time
- A random field is a function on a product space $\Omega \times D$
  - *e.g.* sea surface temperature $T_{\mathsf{ss}}(z, \omega)$, $z \equiv (\boldsymbol{x}, t)$
- It is a more complex object than a random variable
  - A combination of an infinite number of random variables
- In many physical systems, uncertain field quantities, described by random fields:
  - are smooth, *i.e.*
  - they have an underlying *low dimensional structure*

  due to large correlation length-scales

Random Fields – KLE

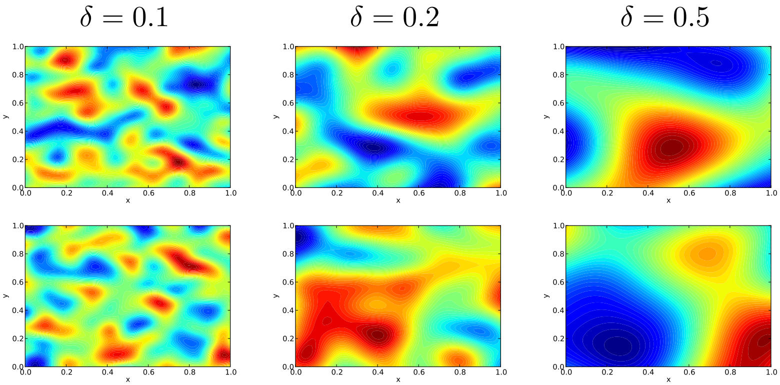
- Smooth random fields can be represented with a small no. of stochastic degrees of freedom
- A random field $M(x,\omega)$ with
  - a mean function: $\mu(x)$
  - a continuous covariance function: $C(x_1,x_2) = \langle [M(x_1,\omega)-\mu(x_1)][M(x_2,\omega)-\mu(x_2)]\rangle$

  can be represented with the Karhunen-Loeve Expansion (KLE)

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

  where
  - λ_i and $\phi_i(x)$ are the eigenvalues and eigenfunctions of the covariance function $C(\cdot,\cdot)$
  - η_i are uncorrelated zero-mean unit-variance RVs
- KLE $\Rightarrow$ representation of random fields using PC

RF Illustration: KL of 2D Gaussian Process

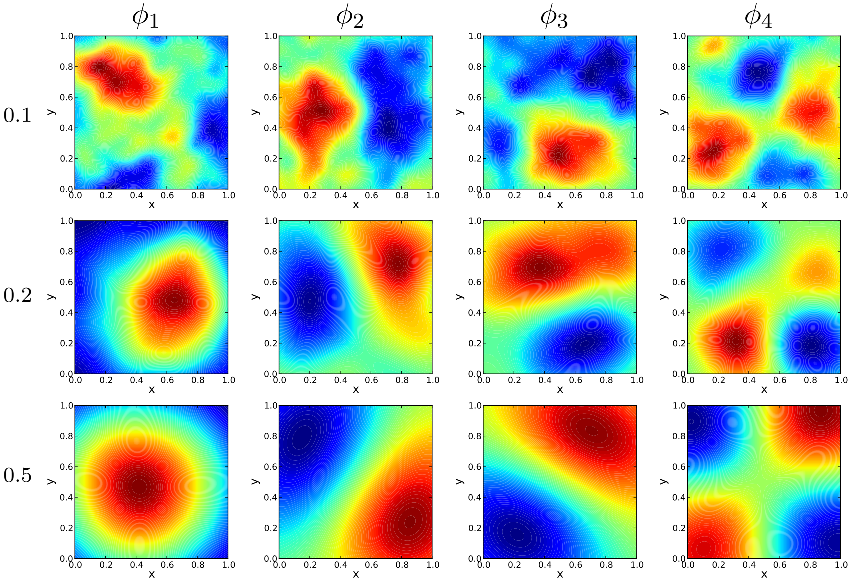


- 2D Gaussian Process with covariance:

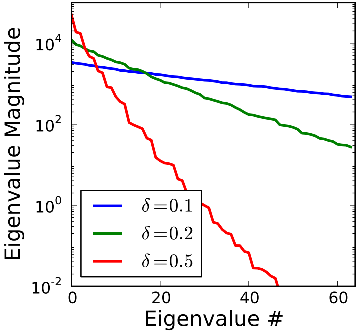
$$C(x_1, x_2) = \text{exp}(-||x_1 - x_2||^2/\delta^2)$$

- Realizations smoother as covariance length δ increases

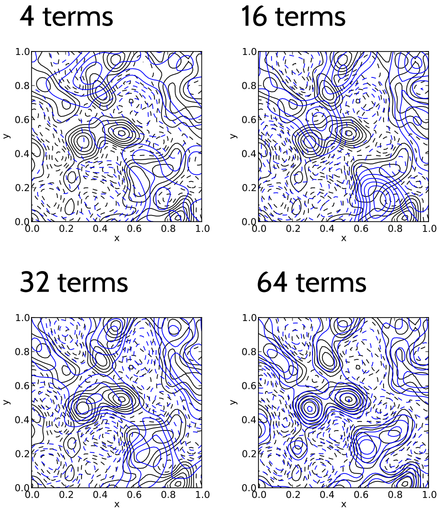
RF Illustration: 2D KL - Modes for $\delta = 0.1 - 0.5$



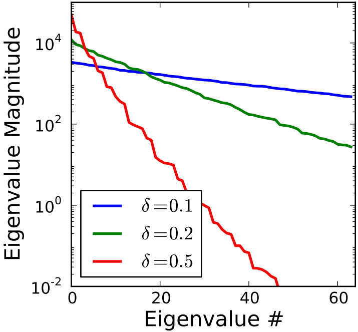
RF Illustration: 2D KL - eigenvalue spectrum



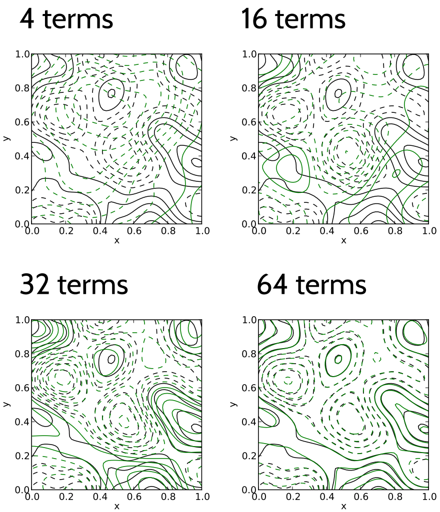
$\delta = 0.1$



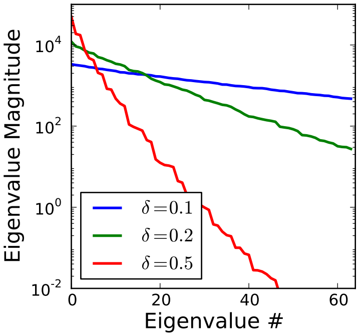
RF Illustration: 2D KL - eigenvalue spectrum

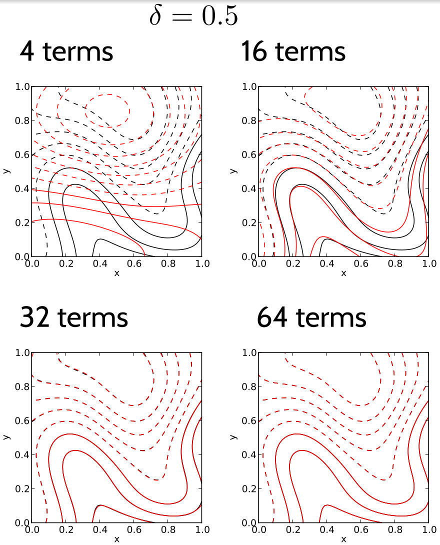


$\delta = 0.2$



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
- Utility
  - Moments: $\mathsf{E}(u) = u_0$, $\mathsf{var}(u) = \sum_{k=1}^{P} u_k^2 \langle \Psi_k^2 \rangle$, $\ldots$
  - Global Sensitivities – fractional variances, Sobol' indices
  - Surrogate for forward model

Requirement:

- RVs in L^2, *i.e.* with finite variance, on $(\Omega, \mathfrak{S}(\boldsymbol{\xi}), P)$

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

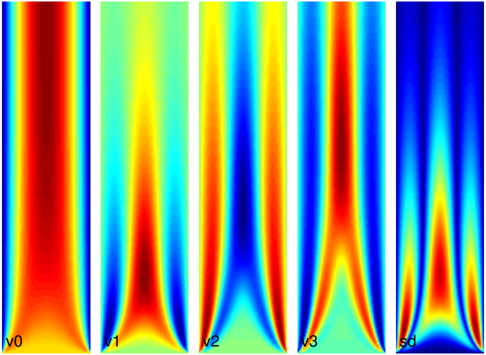
- Given model equations: $\mathcal{M}(u(\boldsymbol{x},t);\lambda) = 0$
- Express uncertain parameters/variables using PCEs

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

- Substitute in model equations; apply Galerkin projection
- New set of equations: $\mathcal{G}(U(\boldsymbol{x},t),\Lambda) = 0$
  - with $U = [u_0,\ldots,u_P]^T$, $\Lambda = [\lambda_0,\ldots,\lambda_P]^T$
- Solving this deterministic 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
  - $\mathsf{v} = \sum_{i=0}^{P} \mathsf{v}_i \Psi_i$
  - v_0: mean
  - v_i: i-th order mode
  - $\sigma^2 = \sum_{i=1}^{P} \mathsf{v}_i^2 \langle \Psi_i^2 \rangle$



v_0 v_1 v_2 v_3 σ

(Le Maître *et al.*, J. Comput. Phys., 2001)

Intrusive PC UQ Pros/Cons

Cons:

- Reformulation of governing equations
- New discretizations
- New numerical solution method
  - Consistency, Convergence, Stability
  - Global vs. multi-element local PC constructions
- New solvers and model codes
  - Opportunities for automated code transformation
- New preconditioners

Pros:

- Tailored solvers <u>can</u> deliver superior performance

Non-intrusive 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

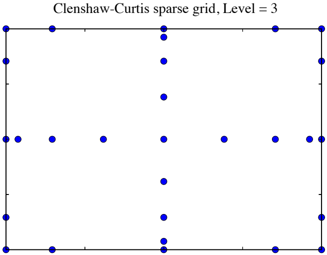
PC and High-Dimensionality

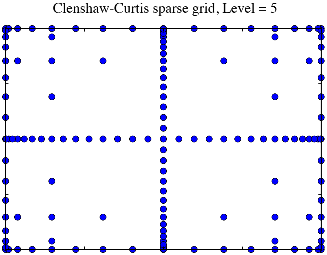
Dimensionality n of the PC basis: $\boldsymbol{\xi} = \{\xi_1, \ldots, \xi_n\}$

- $n \approx$ number of uncertain parameters
- $P + 1 = (n+p)!/n!p!$ grows fast with n

Impacts:

- Size of intrusive PC system
- Hi-D projection integrals $\Rightarrow$ large # non-intrusive samples
  - Sparse quadrature methods

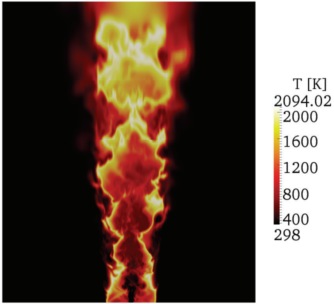


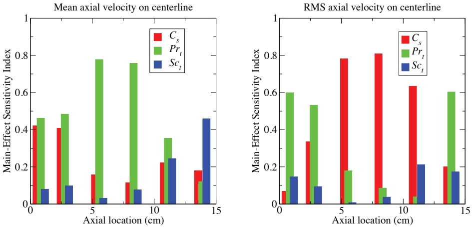


UQ in LES computations: turbulent bluff-body flame

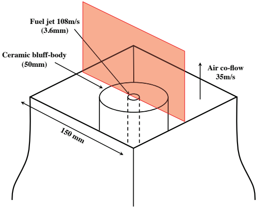
with M. Khalil, G. Lacaze, & J. Oefelein, Sandia Nat. Labs

- CH_4-H_2 jet, air coflow, 3D flow
- Re=9500, LES subgrid modeling
- 12×10^6 mesh cells, 1024 cores
- 3 days run time, 2×10^5 time steps
- 3 uncertain parameters (C_s, Pr_t, Sc_t)
- 2^{nd}-order PC, 25 sparse-quad. pts





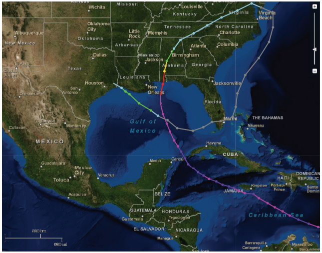
Main-Effect Sensitivity Indices

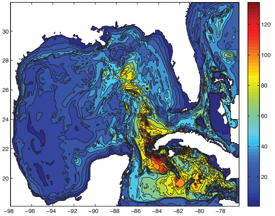


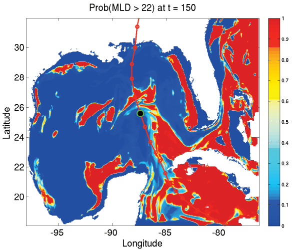
J. Oefelein & G. Lacaze, SNL

UQ in Ocean Modeling – Gulf of Mexico

A. Alexanderian, J. Winokur, I. Sraj, O.M. Knio, Duke Univ.
A. Srinivasan, M. Iskandarani, Univ. Miami; W.C. Thacker, NOAA







- Hurricane Ivan, Sep. 2004
- HYCOM ocean model (hycom.org)
- Predicted Mixed Layer Depth (MLD)
- Four uncertain parameters, *i.i.d.* U
  - subgrid mixing & wind drag params
- 385 sparse quadrature samples

(Alexanderian *et al.*, Winokur *et. al.*, *Comput. Geosci.*, 2012, 2013)

Inverse UQ – Estimation of Uncertain Parameters

Forward UQ requires specification of uncertain inputs

Probabilistic setting

- Require joint PDF on input space
- Statistical inference – an inverse problem

Bayesian setting

- Given Data: PDF on uncertain inputs can be estimated using Bayes formula
  - Bayesian Inference
- Given Constraints: PDF on uncertain inputs can be estimated using the Maximum Entropy principle
  - MaxEnt Methods

Bayes formula for Parameter Inference

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

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

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

- 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

The Prior

- Prior $p(\lambda)$ comes from
  - Physical constraints
  - Prior data
  - Prior knowledge
- The prior can be **uninformative**
- It can be chosen to impose **regularization**
- Unknown aspects of the prior can be added to the rest of the parameters as hyperparameters
- The choice of prior can be crucial when there is little information in the data relative to the number of degrees of freedom in the inference problem
- When there is sufficient information in the data, the data can overrule the prior

Construction of the Likelihood $p(y|\lambda)$

- Where does probability enter the mapping $\lambda \to y$ in $p(y|\lambda)$?
- Through a presumed error model:
- Example:
    - Model:

$$y_m = g(\lambda)$$

    - Data: y
    - Error between data and model prediction: ϵ

$$y = g(\lambda) + \epsilon$$

- Model this error as a random variable
- Example
    - Error is due to instrument measurement noise
    - Instrument has Gaussian errors, with no bias

$$\epsilon \sim N(0, \sigma^2)$$

Construction of the Likelihood $p(y|\lambda)$ - cont'd

For any given λ, this implies

$$y|\lambda, \sigma \sim N(g(\lambda), \sigma^2)$$

or

$$p(y|\lambda, \sigma) = \frac{1}{\sqrt{2\pi}\,\sigma} \text{exp} \left(-\frac{(y - g(\lambda))^2}{2\sigma^2} \right)$$

Given N measurements $(y_1, \ldots, y_N)$, and presuming independent identically distributed (*iid*) noise

$$\begin{aligned} y_i &= g(\lambda) + \epsilon_i \\ \epsilon_i &\sim N(0, \sigma^2) \\ L(\lambda) = p(y_1, \ldots, y_N|\lambda, \sigma) &= \prod_{i=1}^{N} p(y_i|\lambda, \sigma) \end{aligned}$$

Likelihood Modeling

- This is frequently the *core* modeling challenge
  - Error model: a statistical model for the discrepancy between the forward model and the data
  - composition of the error model with the forward model
- Error model composed of discrepancy between
  - data and the truth – (data error)
  - model prediction and the truth – (model error)
- Mean bias and correlated/uncorrelated noise structure
- Hierarchical Bayes modeling, and dependence trees

$$p(\phi, \theta | D) = p(\phi | \theta, D) p(\theta | D)$$

- Choice of observable – constraint on Quantity of Interest?

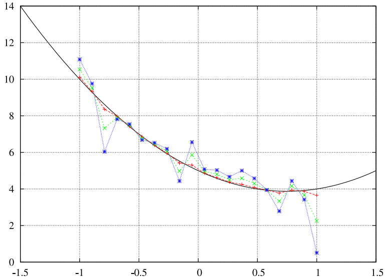
Exploring the Posterior

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

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

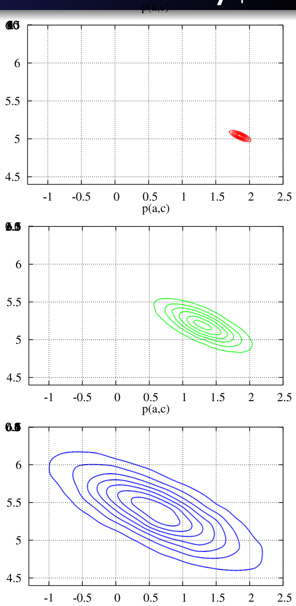
- Explore posterior w/ Markov Chain Monte Carlo (MCMC)
  - Metropolis-Hastings algorithm:
    - Random walk with proposal PDF & rejection rules
  - Computationally intensive, $\mathcal{O}(10^5)$ samples
  - Each sample: evaluation of the forward model
    - Surrogate models
- Evaluate moments/marginals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

Bayesian inference illustration: noise$\uparrow$ $\Rightarrow$ uncertainty$\uparrow$



- data: $y = 2x^2 - 3x + 5 + \epsilon$
- $\epsilon \sim \mathcal{N}(0, \sigma^2)$, $\sigma = \{0.1, 0.5, 1.0\}$
- Fit model $y = ax^2 + bx + c$

Marginal posterior density $p(a, c)$:



Bayesian inference – High Dimensionality Challenge

- Judgement on local/global posterior peaks is difficult
  - Multiple chains; Tempering
- Choosing a good starting point is very important
  - An initial optimization strategy is useful, albeit not trivial
- Choosing good MCMC proposals, and attaining good mixing
  - Likelihood-informed
    - Markov jump in those dimensions informed by data
    - Sample from prior in complement of dimensions
    - Adaptive proposal learning from MCMC samples
    - Log-Posterior Hessian $\Rightarrow$ local Gaussian approx.
    - Adaptive, Geometric, Langevin MCMC
  - Dimension independent
    - Proposal design: good MCMC performance in hiD
  - Literature: A. Stuart, M. Girolami, K. Law, T. Cui, Y. Marzouk

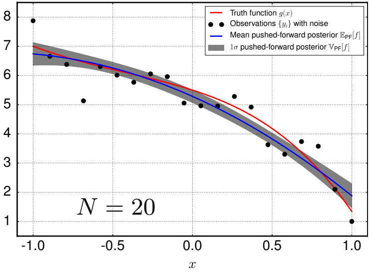
(Law 2014; Cui *et al.*, 2014,2015; Cotter *et al.*, 2013)

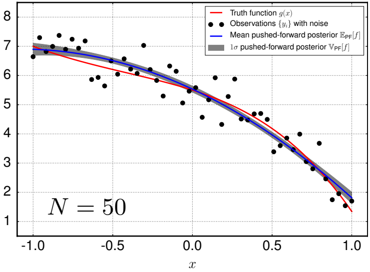
Bayesian inference – Model Error Challenge

- Quantifying model error, as distinct from data noise, is important for assessing confidence in model validity
- Conventional statistical methods for representation of model error have shortcomings when applied to physical models
- New methods are under-development for model error:
  - physical constraints are satisfied
  - feasible disambiguation of model-error/data-noise
  - calibrated model error terms adequately impact all model outputs of interest
  - uncertainties in predictions from calibrated model reflect the range of discrepancy from the truth
- Embed model error in submodel components where approximations exist

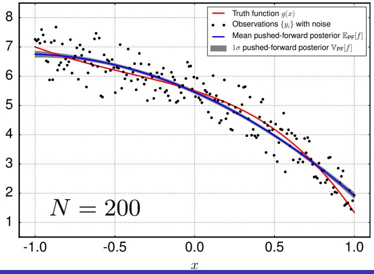
(K. Sargsyan *et al.*, 2015)

Quadratic-fit – Classical Bayesian likelihood

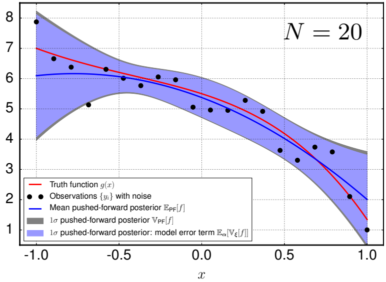


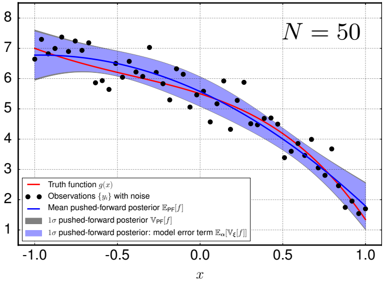


- With additional data, predictive uncertainty around the wrong model is indefinitely reducible
- Predictive uncertainty not indicative of discrepancy from truth

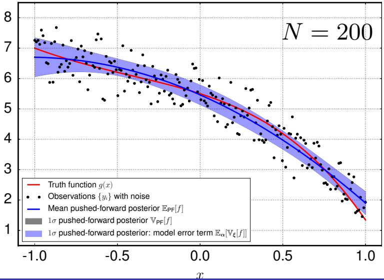


Quadratic-fit – ModErr – MargGauss

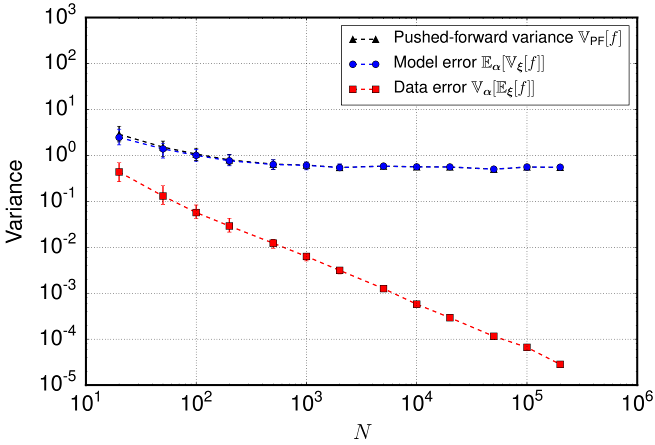




- With additional data, predictive uncertainty due to data noise is reducible
- Predictive uncertainty due to model error is not reducible



Quadratic-fit – ModErr – MargGauss



Calibrating a quadratic $f(x)$ w.r.t. $g(x) = 6 + x^2 + 0.5(x+1)^{3.5}$

Model Evidence and Complexity

Let $\mathcal{M} = \{M_1, M_2, \ldots\}$ be a set of models of interest

- Parameter estimation from data is conditioned on the model

$$p(\theta|D, M_k) = \frac{p(D|\theta, M_k)\pi(\theta|M_k)}{p(D|M_k)}$$

Evidence (marginal likelihood) for M_k:

$$p(D|M_k) = \int p(D|\theta, M_k)\pi(\theta|M_k)\mathrm{d}\theta$$

Model evidence is useful for model selection

- Choose model with maximum evidence
- Compromise between fitting data and model complexity
  - Optimal complexity – Occam's razor principle
  - Avoid overfitting

Too much model complexity leads to overfitting

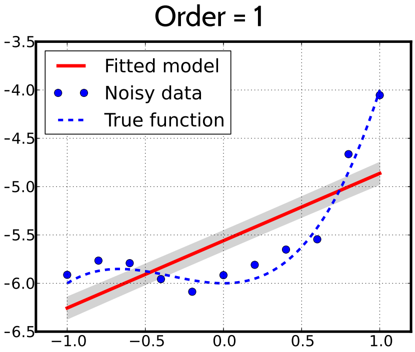
Data model: $i = 1, \ldots, N$

$$
\begin{aligned}
y_i &= x_i^3 + x_i^2 - 6 + \epsilon_i \\
\epsilon_i &\sim N(0, s)
\end{aligned}
$$

Bayesian regression with Legendre PCE fit models, order 1-10

$$y_m = \sum_{k=0}^{P} c_k \psi_k(x)$$

Uniform priors $\pi(c_k)$, $k = 0, \ldots, P$



Fitted model pushed-forward posterior versus the data

Too much model complexity leads to overfitting

Data model: $i = 1, \ldots, N$

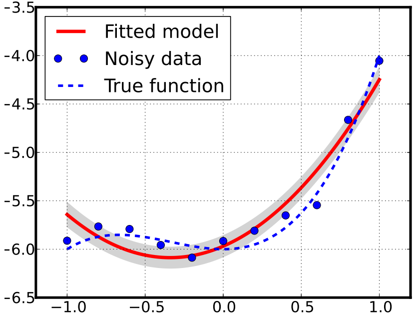
$$\begin{aligned} y_i &= x_i^3 + x_i^2 - 6 + \epsilon_i \\ \epsilon_i &\sim N(0, s) \end{aligned}$$

Bayesian regression with Legendre PCE fit models, order 1-10

$$y_m = \sum_{k=0}^{P} c_k \psi_k(x)$$

Uniform priors $\pi(c_k)$, $k = 0, \ldots, P$

Order = 2



Fitted model pushed-forward posterior versus the data

Too much model complexity leads to overfitting

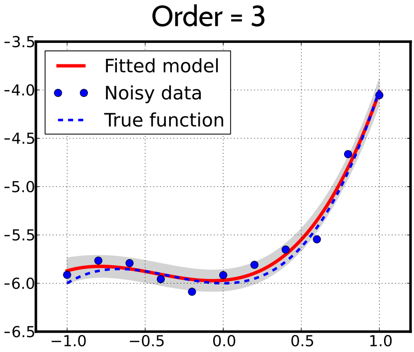
Data model: $i = 1, \ldots, N$

$$\begin{aligned} y_i &= x_i^3 + x_i^2 - 6 + \epsilon_i \\ \epsilon_i &\sim N(0, s) \end{aligned}$$

Bayesian regression with Legendre PCE fit models, order 1-10

$$y_m = \sum_{k=0}^{P} c_k \psi_k(x)$$

Uniform priors $\pi(c_k)$, $k = 0, \ldots, P$



Fitted model pushed-forward posterior versus the data

Too much model complexity leads to overfitting

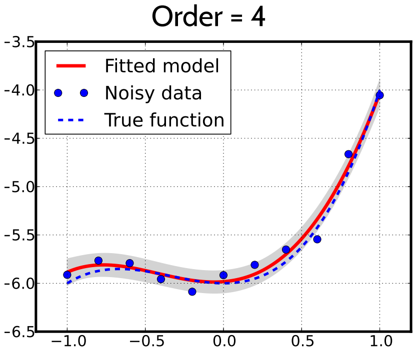
Data model: $i = 1, \ldots, N$

$$\begin{aligned} y_i &= x_i^3 + x_i^2 - 6 + \epsilon_i \\ \epsilon_i &\sim N(0, s) \end{aligned}$$

Bayesian regression with Legendre PCE fit models, order 1-10

$$y_m = \sum_{k=0}^{P} c_k \psi_k(x)$$

Uniform priors $\pi(c_k)$, $k = 0, \ldots, P$



Fitted model pushed-forward posterior versus the data

Too much model complexity leads to overfitting

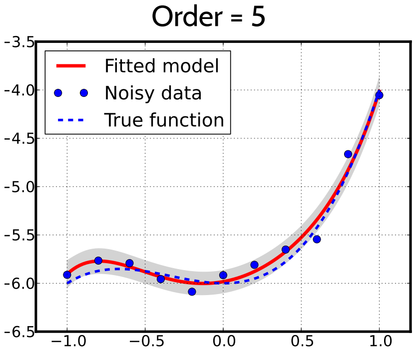
Data model: $i = 1, \ldots, N$

$$\begin{aligned} y_i &= x_i^3 + x_i^2 - 6 + \epsilon_i \\ \epsilon_i &\sim N(0, s) \end{aligned}$$

Bayesian regression with Legendre PCE fit models, order 1-10

$$y_m = \sum_{k=0}^{P} c_k \psi_k(x)$$

Uniform priors $\pi(c_k)$, $k = 0, \ldots, P$



Fitted model pushed-forward posterior versus the data

Too much model complexity leads to overfitting

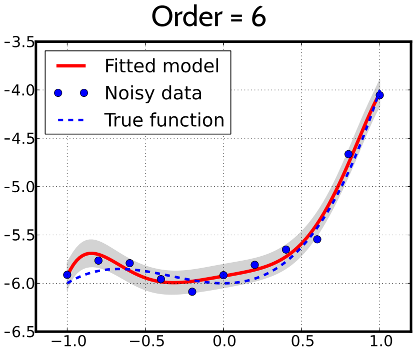
Data model: $i = 1, \ldots, N$

$$\begin{aligned} y_i &= x_i^3 + x_i^2 - 6 + \epsilon_i \\ \epsilon_i &\sim N(0, s) \end{aligned}$$

Bayesian regression with Legendre PCE fit models, order 1-10

$$y_m = \sum_{k=0}^{P} c_k \psi_k(x)$$

Uniform priors $\pi(c_k)$, $k = 0, \ldots, P$



Fitted model pushed-forward posterior versus the data

Too much model complexity leads to overfitting

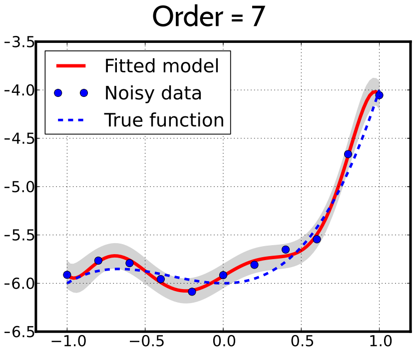
Data model: $i = 1, \ldots, N$

$$y_i = x_i^3 + x_i^2 - 6 + \epsilon_i$$
$$\epsilon_i \sim N(0, s)$$

Bayesian regression with Legendre PCE fit models, order 1-10

$$y_m = \sum_{k=0}^{P} c_k \psi_k(x)$$

Uniform priors $\pi(c_k)$, $k = 0, \ldots, P$



Fitted model pushed-forward posterior versus the data

Too much model complexity leads to overfitting

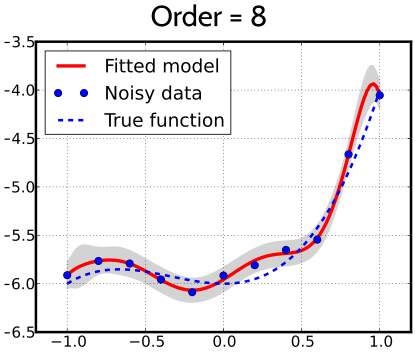
Data model: $i = 1, \ldots, N$

$$\begin{aligned} y_i &= x_i^3 + x_i^2 - 6 + \epsilon_i \\ \epsilon_i &\sim N(0, s) \end{aligned}$$

Bayesian regression with Legendre PCE fit models, order 1-10

$$y_m = \sum_{k=0}^{P} c_k \psi_k(x)$$

Uniform priors $\pi(c_k)$, $k = 0, \ldots, P$



Fitted model pushed-forward posterior versus the data

Too much model complexity leads to overfitting

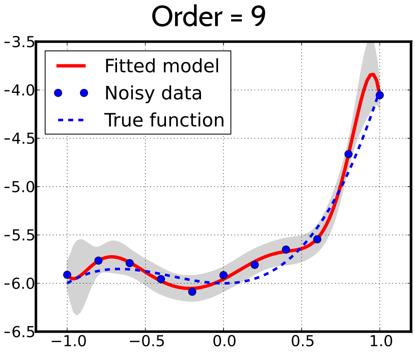
Data model: $i = 1, \ldots, N$

$$\begin{aligned} y_i &= x_i^3 + x_i^2 - 6 + \epsilon_i \\ \epsilon_i &\sim N(0, s) \end{aligned}$$

Bayesian regression with Legendre PCE fit models, order 1-10

$$y_m = \sum_{k=0}^{P} c_k \psi_k(x)$$

Uniform priors $\pi(c_k)$, $k = 0, \ldots, P$



Fitted model pushed-forward posterior versus the data

Too much model complexity leads to overfitting

Data model: $i = 1, \ldots, N$

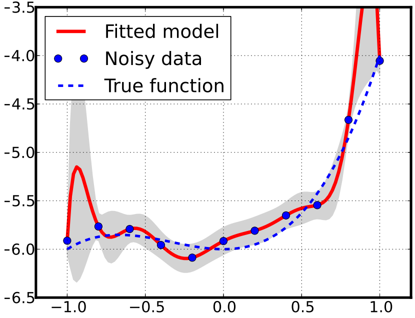
$$\begin{aligned} y_i &= x_i^3 + x_i^2 - 6 + \epsilon_i \\ \epsilon_i &\sim N(0, s) \end{aligned}$$

Bayesian regression with Legendre PCE fit models, order 1-10

$$y_m = \sum_{k=0}^{P} c_k \psi_k(x)$$

Uniform priors $\pi(c_k)$, $k = 0, \ldots, P$

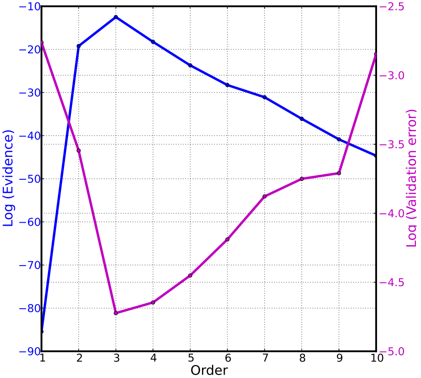
Order = 10



Fitted model pushed-forward posterior versus the data

Evidence and Cross-Validation Error

- Model evidence peaks at the true polynomial order of 3
- Cross validation error is equally minimal at order 3
- Models with optimal complexity are robust to cross validation



Cross validation error and model evidence versus order

Closure

- Probabilistic UQ framework
    - Polynomial Chaos representation of random variables
- Forward UQ
    - Intrusive and non-intrusive forward PC UQ methods
- Inverse UQ
    - Parameter estimation via Bayesian inference
    - Model error
    - Model complexity
- Challenges
    - High dimensionality
    - Intrusive Galerkin stability
    - Nonlinearity
    - Time dynamics
    - Model error