Topics on Uncertainty Quantification

UQ challenges and techniques to address them

Khachik Sargsyan

Sandia National Laboratories, Livermore, CA
Transportation Energy Center
Reacting Flow Research Department

Reliability Seminar,
UC Berkeley, Feb 23, 2012

Puzzle

Problem

You have eight digits from 1 to 8. Drop one of them randomly. Construct a random 7-digit integer from the remaining set (no repetition of digits).

What is the probability that this number is a multiple of 9 ?

Puzzle

Problem

You have eight digits from 1 to 8. Drop one of them randomly. Construct a random 7-digit integer from the remaining set (no repetition of digits).

What is the probability that this number is a multiple of 9 ?

Puzzle

Problem

You have eight digits from 1 to 8. Drop one of them randomly. Construct a random 7-digit integer from the remaining set (no repetition of digits).
What is the probability that this number is a multiple of 9 ?

Solution

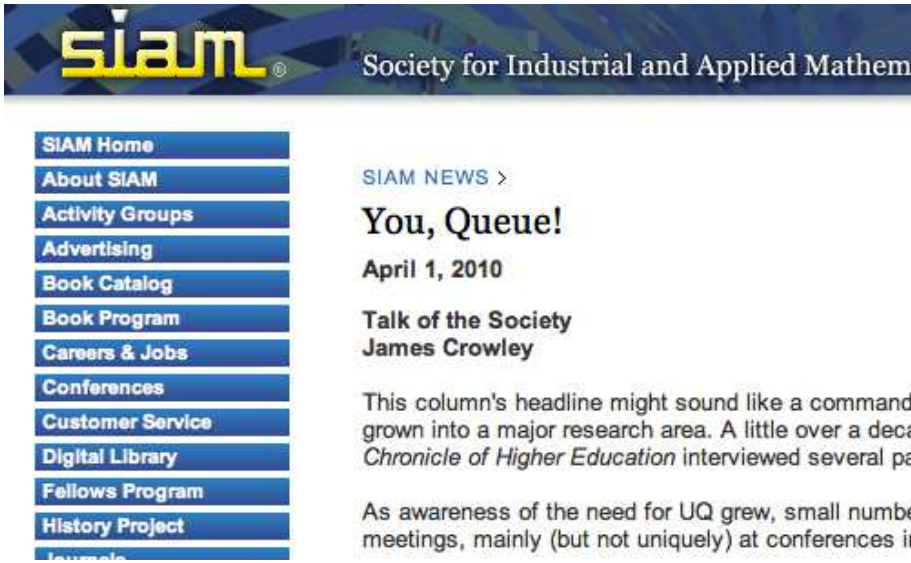
The trick is: this is not a probability question, rather a number theory one.
A number is a multiple of 9 *iff* sum of its digits is a multiple of 9.
$1 + 2 + \ldots + 8 = 36$. You can not remove any of the digits to reach a multiple of 9.
Hence, the answer is 0.

To extend,

- repeat this exercise with any number of digits.
- replace 9 with 3.

Outline

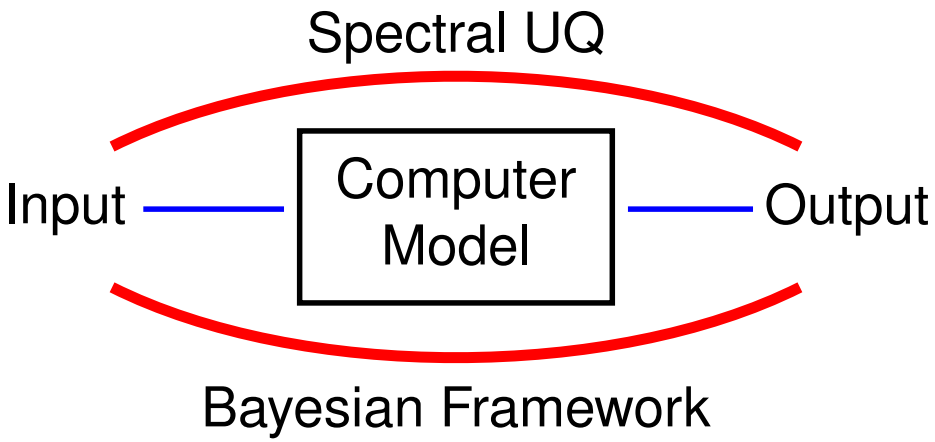
- Uncertainty quantification (UQ)
  - What, why, how?
  - Functional representation of r.v.
  - Forward UQ
- Challenges
  - Input correlations
  - Distribution tails
  - Intrinsic stochasticity
  - Limited data
  - High dimensionality
  - Nonlinear/bifurcative systems
  - ...



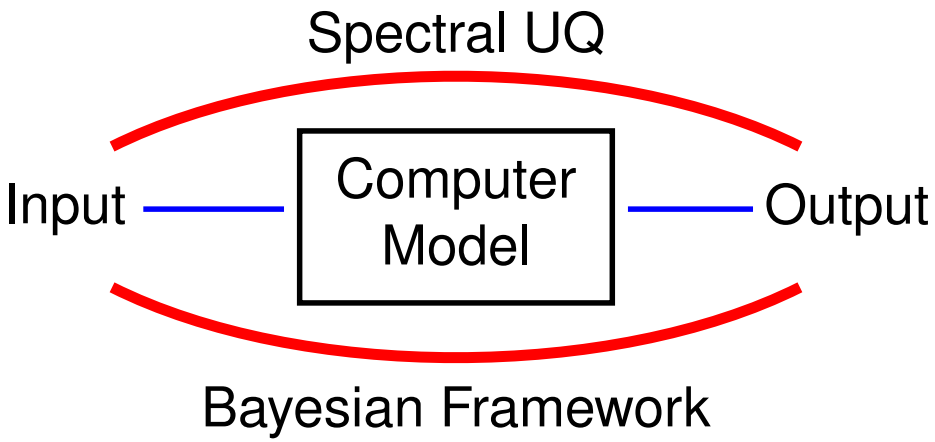
SIAM Conference on
Uncertainty Quantification

Raleigh Marriott City Center
Raleigh, North Carolina USA
April 2 - 5, 2012

Big picture


Spectral UQ
Input
Computer Model
Output
Bayesian Framework

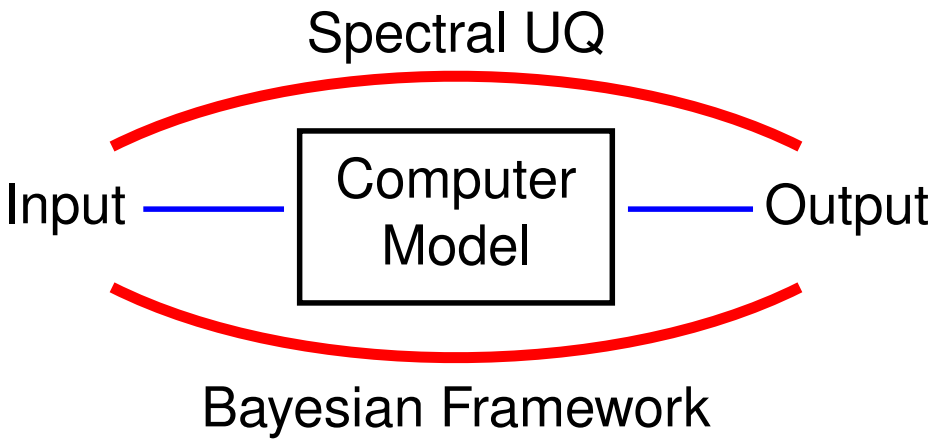
Big picture



Uncertainty sources

- Model parameters
- Initial/boundary conditions
- Model geometry/structure
- Lack of knowledge
- Data noise
- Intrinsic stochasticity

Big picture



Uncertainty sources

- Model parameters
- Initial/boundary conditions
- Model geometry/structure
- Lack of knowledge
- Data noise
- Intrinsic stochasticity

Uncertainty types

- Epistemic (reducible)
- Aleatoric (irreducible)
- Bayesian probabilistic viewpoint covers both
- Uncertainty $\sim$ Probability $\sim$ $\sim$ State of knowledge

Why UQ?

UQ explores the effect of input uncertainties on the outputs of interest.

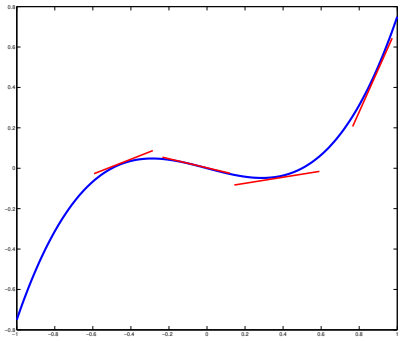
- Model predictions over range of uncertainty
- Model validation and comparison
- Confidence assessment
- Reliability analysis
- Dimensionality reduction
- Optimal design
- Decision support
- (Noisy) data assimilation
  - Combination of measurements and model predictions to obtain accurate representations.

UQ components

- Locate all sources of (manageable) uncertainties
- Parameter selection/estimation
  - Auxilliary data collection
  - Expert opinion, physical bounds, maximum entropy
  - Submodel fitting/regression
- Forward propagation of uncertainties
  - Local SA (deterministic, error propagation)
  - Interval math, evidence theory
  - Global SA (stochastic, variance-based decomposition)
- Calibration with a separate set of data (inverse UQ)
- Model (in)validation
  - No model is perfect
  - Compare model prediction with uncertainties versus data on some QoI
  - Model comparison (Bayes Factors, Model Plausibility)
- See the previous slide

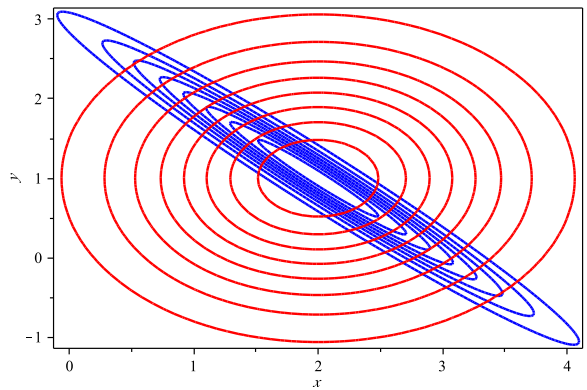
Forward UQ

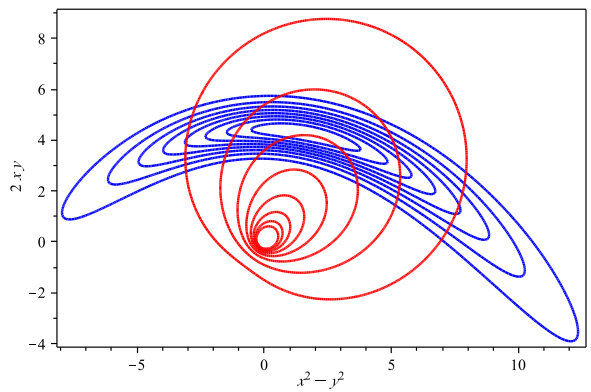
- Local sensitivity analysis
  - well, it's local
- Interval math
  - misses correlations
- Global sensitivity analysis
  - Random sampling (MC, LHS, QMC,...)
  - Analysis of variance (ANOVA) =
    = High dimensional model representation (HDMR)
  - Response surface construction



Forward UQ

- Local sensitivity analysis
  - well, it's local
- Interval math
  - misses correlations
- Global sensitivity analysis
  - Random sampling (MC, LHS, QMC,...)
  - Analysis of variance (ANOVA) =
    = High dimensional model representation (HDMR)
  - Response surface construction





Forward UQ

- Local sensitivity analysis
  - well, it's local
- Interval math
  - misses correlations
- Global sensitivity analysis
  - Random sampling (MC, LHS, QMC,...)
  - Analysis of variance (ANOVA) =
    = High dimensional model representation (HDMR)
  - Response surface construction

Spectral Methods for Forward Stochastic UQ

Polynomial Chaos represents any* L_2 random variable as a polynomial of standard random variables

- Truncated PCE: finite dimension n and order p

$$X \simeq \sum_{k=0}^{P} c_k \Psi_k(\boldsymbol{\eta})$$

with the number of terms $P+1 = \frac{(n+p)!}{n!p!}$.

- $\boldsymbol{\eta} = (\eta_1, \cdots, \eta_n)$ standard i.i.d. r.v.
  Ψ_k standard polynomials, orthogonal w.r.t. $\pi(\boldsymbol{\eta})$
  c_k spectral modes.
- Most common standard Polynomial-Variable pairs:
  (continuous) Gauss-Hermite, Legendre-Uniform,
  (discrete) Poisson-Charlier.

[Wiener, 1938; Ghanem & Spanos, 1991; Xiu & Karniadakis, 2002; Le Maître & Knio, 2010]

Polynomial Chaos represents any* L_2 random variable as a polynomial of standard random variables

- Truncated PCE: finite dimension n and order p

$$X \simeq \sum_{k=0}^{P} c_k \Psi_k(\boldsymbol{\eta})$$

  with the number of terms $P+1 = \frac{(n+p)!}{n!p!}$.

- $\boldsymbol{\eta} = (\eta_1, \cdots, \eta_n)$ standard i.i.d. r.v.
  Ψ_k standard polynomials, orthogonal w.r.t. $\pi(\boldsymbol{\eta})$
  c_k spectral modes.

- Most common standard Polynomial-Variable pairs: (continuous) Gauss-Hermite, Legendre-Uniform, (discrete) Poisson-Charlier.

[Wiener, 1938; Ghanem & Spanos, 1991; Xiu & Karniadakis, 2002; Le Maître & Knio, 2010]

Polynomial Chaos represents any* L_2 random variable as a polynomial of standard random variables

- Truncated PCE: finite dimension n and order p

$$X \simeq \sum_{k=0}^{P} c_k \Psi_k(\boldsymbol{\eta})$$

  with the number of terms $P+1 = \frac{(n+p)!}{n!p!}$.

- $\boldsymbol{\eta} = (\eta_1, \cdots, \eta_n)$ standard i.i.d. r.v.
  Ψ_k standard polynomials, orthogonal w.r.t. $\pi(\boldsymbol{\eta})$
  c_k spectral modes.

- Most common standard Polynomial-Variable pairs:
  (continuous) Gauss-Hermite, Legendre-Uniform,
  (discrete) Poisson-Charlier.

[Wiener, 1938; Ghanem & Spanos, 1991; Xiu & Karniadakis, 2002; Le Maître & Knio, 2010]

Constructing a 1D PCE for a r.v. with a given PDF

- Define a 1D PCE for r.v. X:

$$X = \sum_{k=0}^{\infty} c_k \Psi_k(\eta)$$

- Define a uniformly distributed r.v. $u \sim U(0,1)$ using *inverse CDFs*

$$X = F_X^{-1}(u), \qquad \eta = F_\eta^{-1}(u)$$

and evaluate the PC modes c_k via integrals over u :

$$c_k = \frac{\langle X \Psi_k \rangle}{\langle \Psi_k^2 \rangle} = \frac{1}{\langle \Psi_k^2 \rangle} \int_0^1 \underbrace{F_X^{-1}(u)}_{X} \Psi_k(\underbrace{F_\eta^{-1}(u)}_{\eta}) \, du$$

- E.g., for Legendre-Uniform PC, $\eta = 2u - 1$.

Constructing an nD PCE for a r.v. with a given PDF

- Define a multivariate PCE for r.v. X:

$$X = \sum_{k=0}^{P} c_k \Psi_k(\eta_1, \eta_2, \ldots, \eta_n), \qquad P + 1 = \frac{(n+p)!}{n!p!}$$

- No general procedure, but...
- Can choose $\{n, p\}$ and the mode strengths by ensuring accurate capture of
  - the PDF of X
  - select moments of X
  - some observable of interest $f(X)$
- Generalizations of the CDF transform often not feasible or not accurate enough
  - Rosenblatt transformation
  - Nataf transformation
- *Fun example:* η_1, η_2 i.i.d. gaussians, $X = \eta_1^2 + \eta_2^2$ is exponential r.v. However, no finite order 1D representation.

Constructing an nD PCE for a r.v. with a given PDF

- Define a multivariate PCE for r.v. X:

$$X = \sum_{k=0}^{P} c_k \Psi_k(\eta_1, \eta_2, \ldots, \eta_n), \qquad P + 1 = \frac{(n+p)!}{n!p!}$$

- No general procedure, but...
- Can choose $\{n, p\}$ and the mode strengths by ensuring accurate capture of
  - the PDF of X
  - select moments of X
  - some observable of interest $f(X)$
- Generalizations of the CDF transform often not feasible or not accurate enough
  - Rosenblatt transformation
  - Nataf transformation
- *Fun example:* η_1, η_2 i.i.d. gaussians, $X = \eta_1^2 + \eta_2^2$ is exponential r.v. However, no finite order 1D representation.

PC Surrogate Construction

- Build/presume PC for input parameter λ

$$\lambda(\boldsymbol{\eta}) = \sum_{k=0}^{P} \lambda_k \Psi_k(\boldsymbol{\eta})$$

- E.g., uniform on an interval, or gaussian with known moments,

$$\lambda = \lambda_0 + \lambda_1 \eta$$

- Forward function $f(\cdot)$, output u

$$u = f(\lambda(\boldsymbol{\eta})) \qquad\qquad u = \sum_{k=0}^{P} u_k \Psi_k(\boldsymbol{\eta}) \equiv g(\boldsymbol{\eta})$$

- Any forward UQ method can be used to construct the surrogate
- For optimization, inverse problems, the surrogate $g(\boldsymbol{\eta})$ can replace the expensive forward function $f(\lambda(\boldsymbol{\eta}))$
- Global sensitivity information for free (Sobol indices, variance-based decomposition)

PC Surrogate Construction

- Build/presume PC for input parameter λ

$$\lambda(\boldsymbol{\eta}) = \sum_{k=0}^{P} \lambda_k \Psi_k(\boldsymbol{\eta})$$

- E.g., uniform on an interval, or gaussian with known moments,

$$\lambda = \lambda_0 + \lambda_1 \eta$$

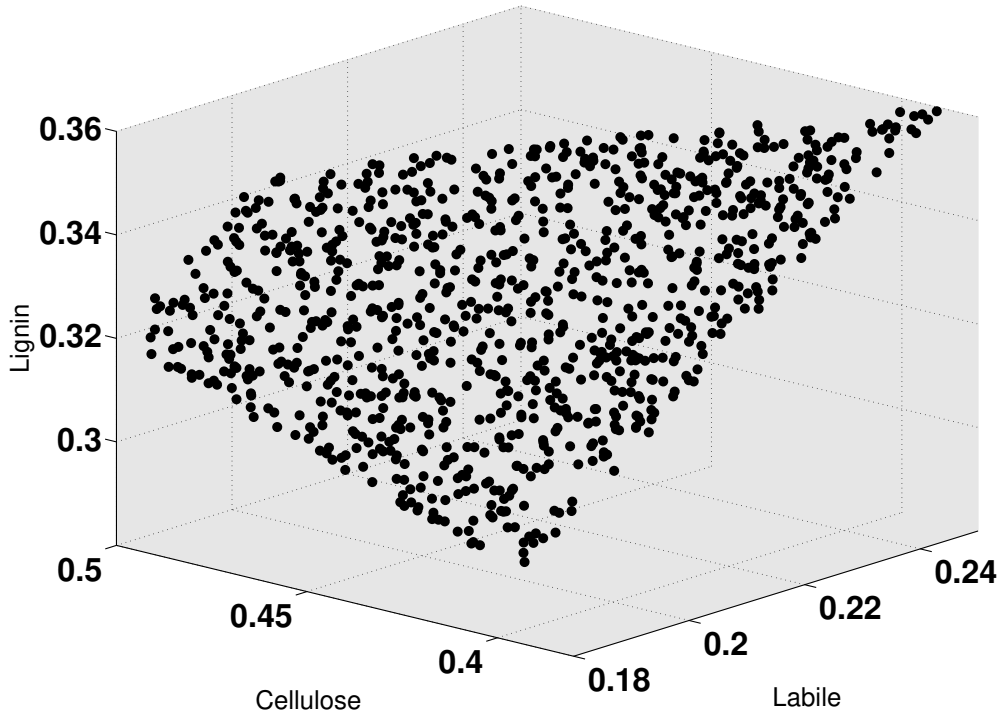
- Forward function $f(\cdot)$, output u

$$u = f(\lambda(\boldsymbol{\eta})) \qquad\qquad u = \sum_{k=0}^{P} u_k \Psi_k(\boldsymbol{\eta}) \equiv g(\boldsymbol{\eta})$$

- Any forward UQ method can be used to construct the surrogate
- For optimization, inverse problems, the surrogate $g(\boldsymbol{\eta})$ can replace the expensive forward function $f(\lambda(\boldsymbol{\eta}))$
- Global sensitivity information for free (Sobol indices, variance-based decomposition)

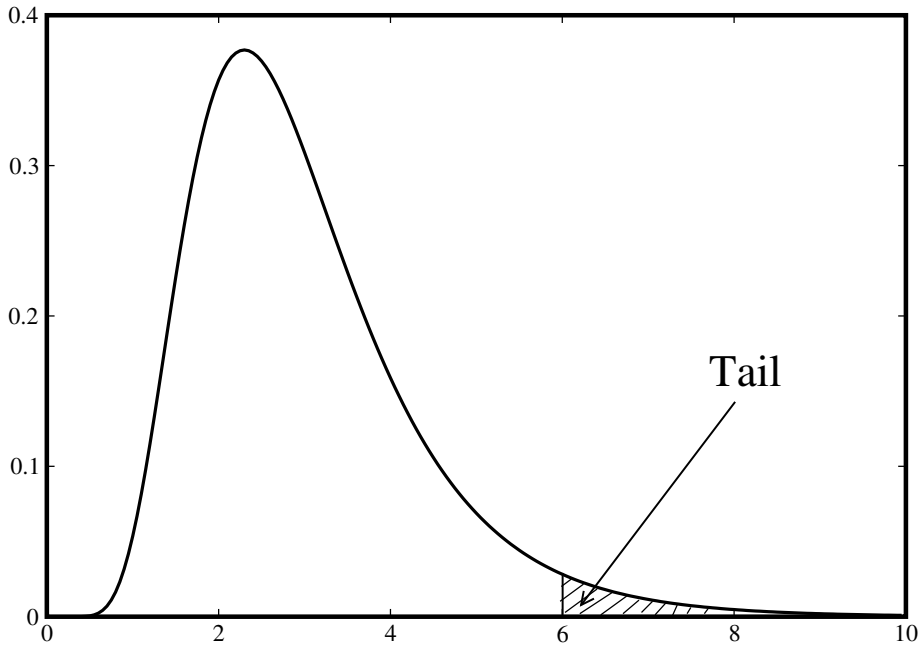
UQ methods are challenged by..

- Input correlations
- Tail regions
- Limited data
- Intrinsic stochasticity
- Curse of dimensionality
- Nonlinearities, Bifurcations, Bimodalities



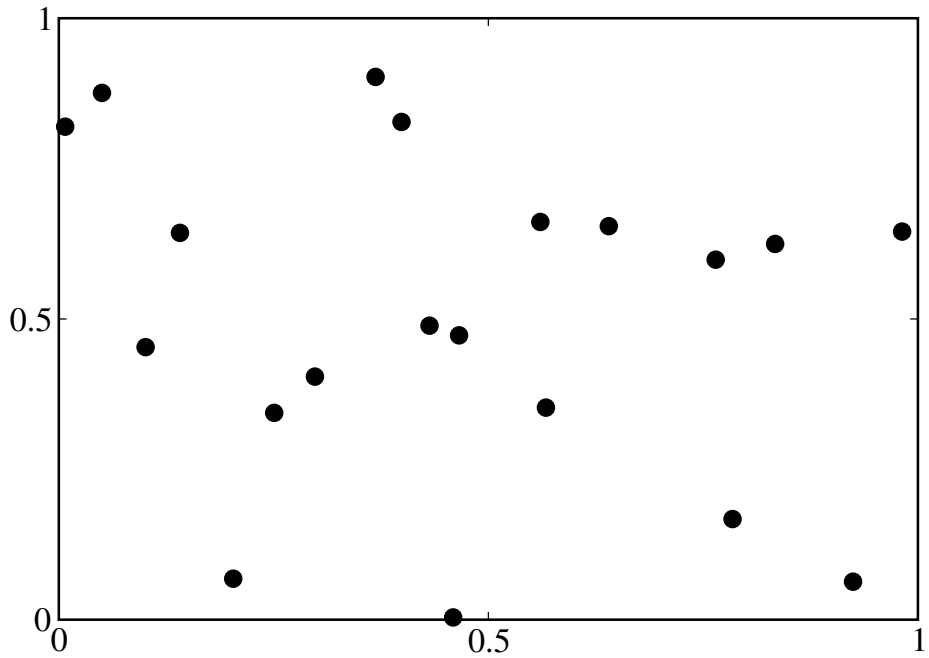
UQ methods are challenged by..

- Input correlations
- Tail regions
- Limited data
- Intrinsic stochasticity
- Curse of dimensionality
- Nonlinearities, Bifurcations, Bimodalities



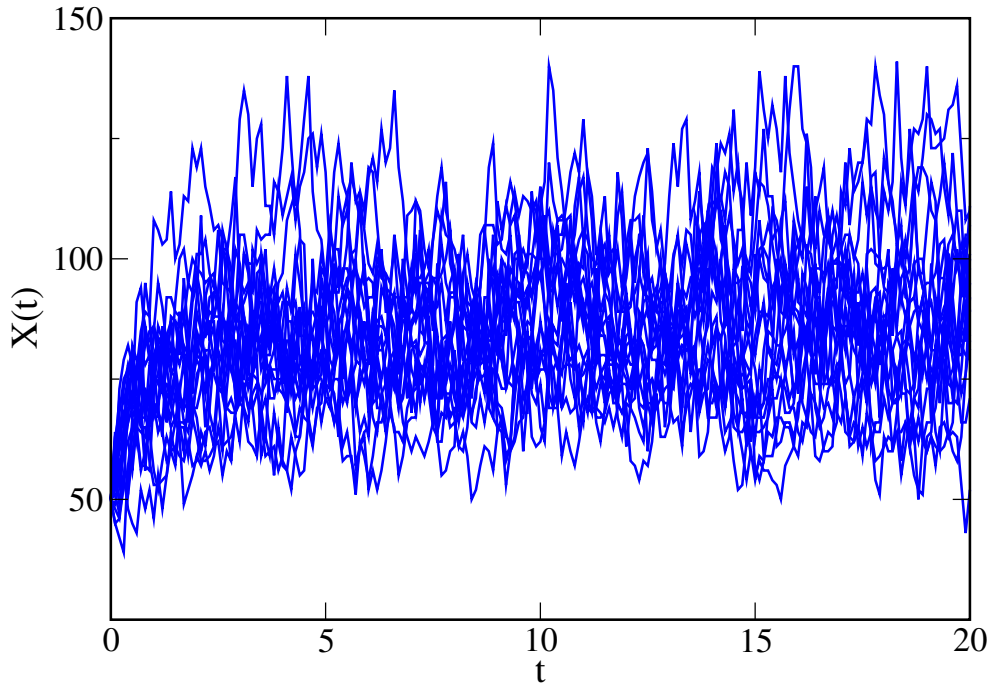
UQ methods are challenged by..

- Input correlations
- Tail regions
- Limited data
- Intrinsic stochasticity
- Curse of dimensionality
- Nonlinearities,
  Bifurcations,
  Bimodalities



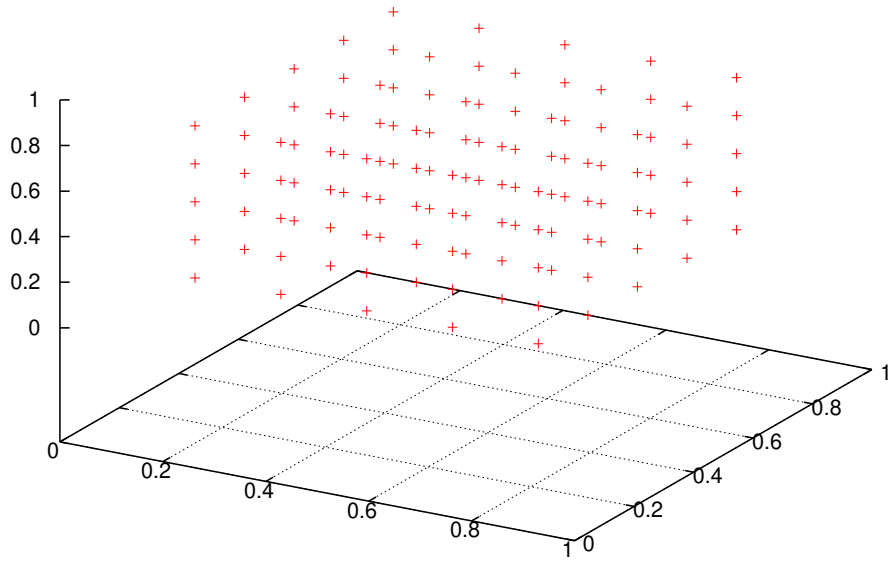
UQ methods are challenged by..

- Input correlations
- Tail regions
- Limited data
- Intrinsic stochasticity
- Curse of dimensionality
- Nonlinearities, Bifurcations, Bimodalities



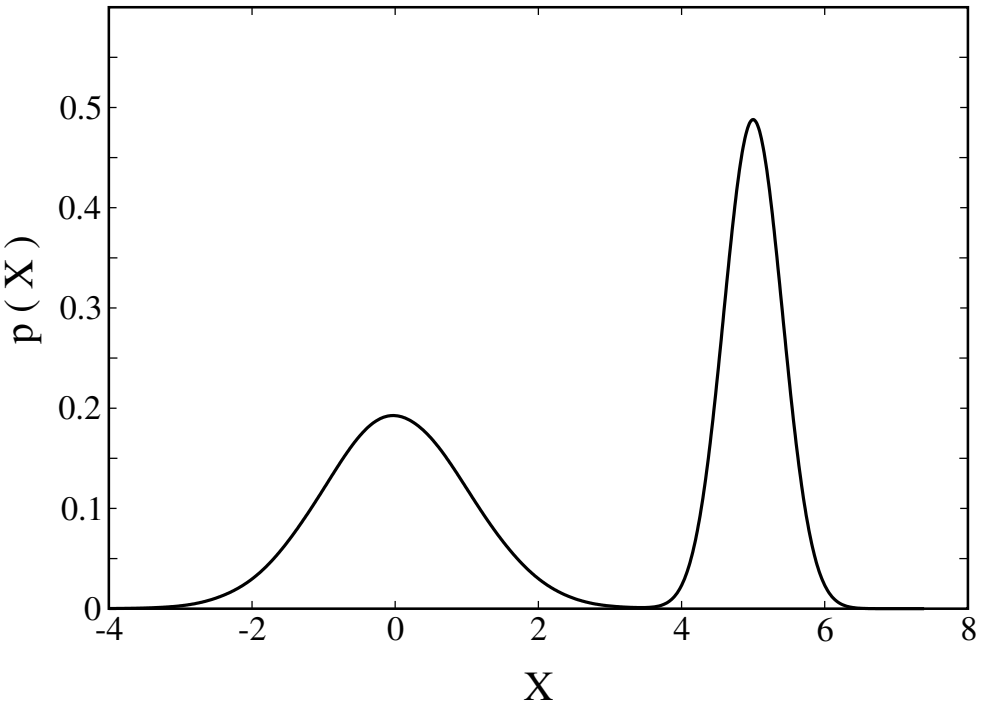
UQ methods are challenged by..

- Input correlations
- Tail regions
- Limited data
- Intrinsic stochasticity
- Curse of dimensionality
- Nonlinearities,
  Bifurcations,
  Bimodalities



UQ methods are challenged by..

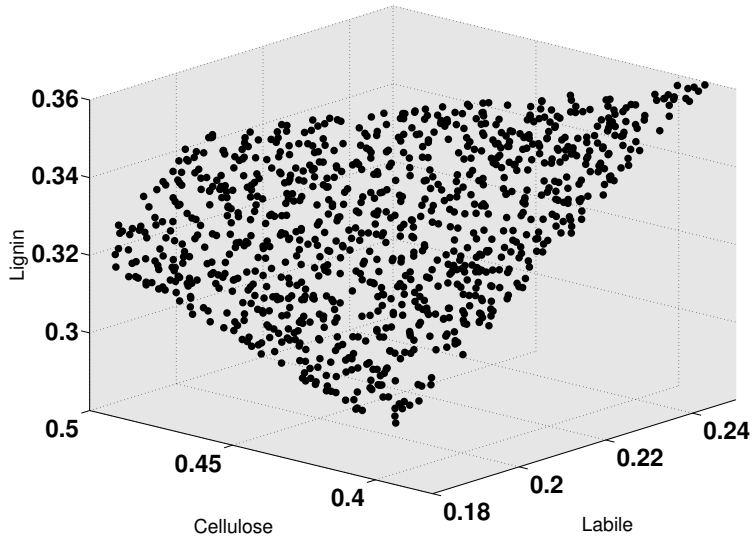
- Input correlations
- Tail regions
- Limited data
- Intrinsic stochasticity
- Curse of dimensionality
- Nonlinearities, Bifurcations, Bimodalities



Input correlations: Rosenblatt transformation

- Rosenblatt transformation maps any (not necessarily independent) set of random variables $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$ to uniform i.i.d.'s $\{\eta_i\}_{i=1}^n$ [Rosenblatt, 1952].

$$
\begin{aligned}
\eta_1 &= F_1(\xi_1) \\
\eta_2 &= F_{2|1}(\xi_2|\xi_1) \\
\eta_3 &= F_{3|2,1}(\xi_3|\xi_2, \xi_1) \\
&\vdots \\
\eta_n &= F_{n|n-1,\ldots,1}(\xi_n|\xi_{n-1}, \ldots, \xi_1)
\end{aligned}
$$



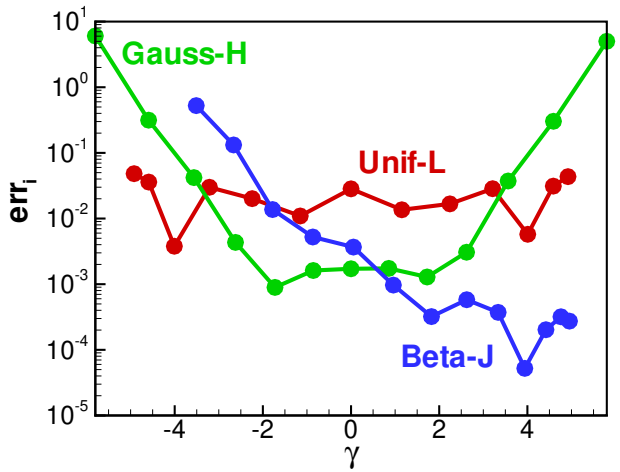
- Inverse Rosenblatt transformation $\boldsymbol{\xi} = R^{-1}(\boldsymbol{\eta})$ ensures a well-defined quadrature integration to build PC [Sargsyan *et al.*, 2010]

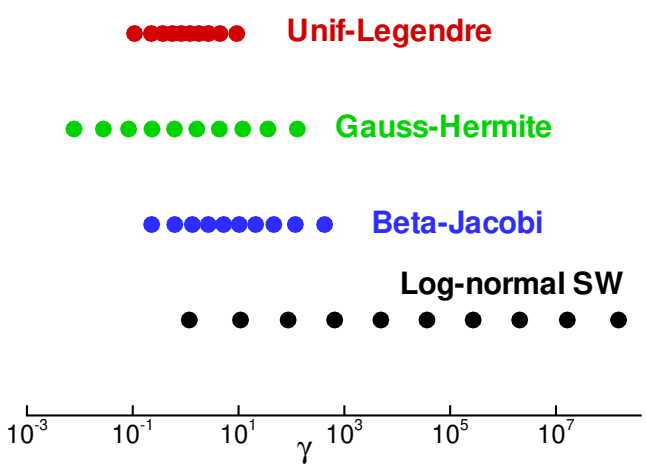
$$c_k = \langle \boldsymbol{\xi} \Psi_k(\boldsymbol{\eta}) \rangle = \int R^{-1}(\boldsymbol{\eta}) \Psi_k(\boldsymbol{\eta}) d\boldsymbol{\eta}$$

- Caveat: if only samples of $\boldsymbol{\xi}$ are available, the conditional distributions are hard to evaluate accurately.

Tail regions: Custom-based polynomial chaos

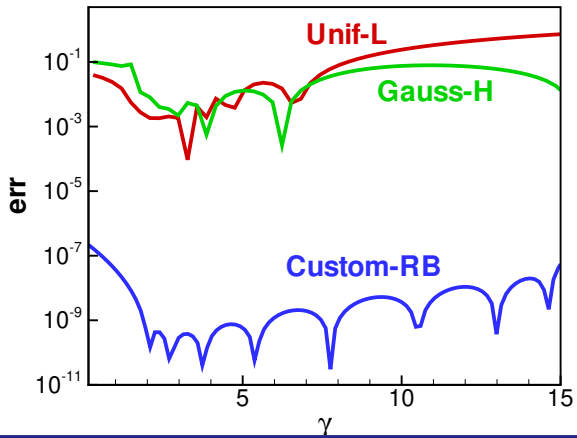
- A significant number of simulations are required to accurately represent tail regions
- Pointwise error of PC representation is large at low-probability regions: $\int \left| f(\eta) - \sum_{k=0}^{P} c_k \Psi_k(\eta) \right|^2 \pi(\eta) d\eta$
- Construct spectral expansions based on...
  - Non-classical bases that cluster points in the tail region
    - Jacobi for Beta distributions, Stieltjes-Wigert for log-normal distributions
  - Bases tailored to the expected behavior of the output
    - Build a custom PC basis according to the input



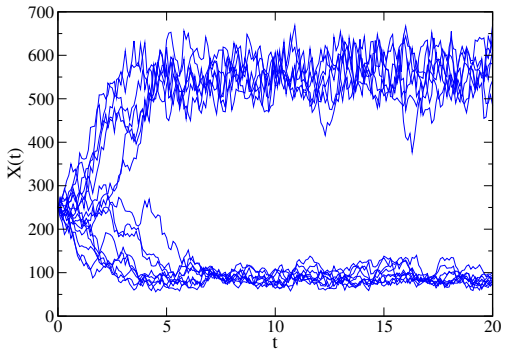


Tail regions: Custom-based polynomial chaos

- A significant number of simulations are required to accurately represent tail regions
- Pointwise error of PC representation is large at low-probability regions: $\int \left| f(\eta) - \sum_{k=0}^{P} c_k \Psi_k(\eta) \right|^2 \pi(\eta) d\eta$
- Construct spectral expansions based on...
  - Non-classical bases that cluster points in the tail region
    - Jacobi for Beta distributions, Stieltjes-Wigert for log-normal distributions
  - Bases tailored to the expected behavior of the output
    - Build a custom PC basis according to the input distribution



Stochastic systems: quadrature methods choke



- Stochastic chemical kinetics
  [Gillespie, 1977]
- Climate buzzword: stochastic physics
  [Palmer & WIlliams, 2009]

- Quadrature formulae presume a degree of smoothness

$$u_k = \frac{1}{\left\langle \Psi_k^2 \right\rangle} \int u(\lambda(\boldsymbol{\eta}))\, \Psi_k(\boldsymbol{\eta}) \pi(\boldsymbol{\eta}) d\boldsymbol{\eta} \approx \sum_* u(\boldsymbol{\lambda}(\boldsymbol{\eta}_*)) \Psi_k(\boldsymbol{\eta}_*) w_*$$

- Sparse-Quadrature formulae are *ill-conditioned* and highly-sensitive to noise
  - No convergence with order
  - Error grows with increased dimensionality
- Options in the presence of noise:
  - RMS fitting for PC coefficients
  - Bayesian inference of PC coefficients

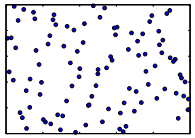
Alternative methods to obtain PC coefficients

$$X \simeq \sum_{k=0}^{P} c_k \Psi_k(\boldsymbol{\eta}) \qquad\qquad c_k = \frac{\langle X(\boldsymbol{\eta})\Psi_k(\boldsymbol{\eta})\rangle}{\langle \Psi_k^2(\boldsymbol{\eta})\rangle}$$

The integral $\langle X(\boldsymbol{\eta})\Psi_k(\boldsymbol{\eta})\rangle = \int X(\boldsymbol{\eta})\Psi_k(\boldsymbol{\eta})\pi(\boldsymbol{\eta})d\boldsymbol{\eta}$ can be estimated by

- Monte-Carlo

$$\frac{1}{K}\sum_{j=1}^{K} X(\boldsymbol{\eta}_j)\Psi_k(\boldsymbol{\eta}_j)$$



many samples from $\pi(\boldsymbol{\eta})$

- Quadrature

$$\sum_{j=1}^{Q} X(\boldsymbol{\eta}_j)\Psi_k(\boldsymbol{\eta}_j)w_j$$

samples at quadrature

- *Bayesian inference*

any (number of) samples

$$P(c_k|X(\boldsymbol{\eta}_j)) \propto P(X(\boldsymbol{\eta}_j)|c_k)P(c_k)$$

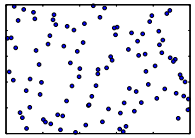
Alternative methods to obtain PC coefficients

$$X \simeq \sum_{k=0}^{P} c_k \Psi_k(\boldsymbol{\eta}) \qquad\qquad c_k = \frac{\langle X(\boldsymbol{\eta})\Psi_k(\boldsymbol{\eta})\rangle}{\langle \Psi_k^2(\boldsymbol{\eta})\rangle}$$

The integral $\langle X(\boldsymbol{\eta})\Psi_k(\boldsymbol{\eta})\rangle = \int X(\boldsymbol{\eta})\Psi_k(\boldsymbol{\eta})\pi(\boldsymbol{\eta})d\boldsymbol{\eta}$ can be estimated by

- Monte-Carlo

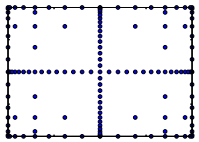
$$\frac{1}{K}\sum_{j=1}^{K} X(\boldsymbol{\eta}_j)\Psi_k(\boldsymbol{\eta}_j)$$



many samples from $\pi(\boldsymbol{\eta})$

- Quadrature

$$\sum_{j=1}^{Q} X(\boldsymbol{\eta}_j)\Psi_k(\boldsymbol{\eta}_j)w_j$$



samples at quadrature

- *Bayesian inference*

any (number of) samples

$$P(c_k|X(\boldsymbol{\eta}_j)) \propto P(X(\boldsymbol{\eta}_j)|c_k)P(c_k)$$

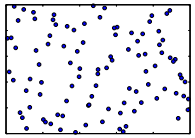
Alternative methods to obtain PC coefficients

$$X \simeq \sum_{k=0}^{P} c_k \Psi_k(\boldsymbol{\eta}) \qquad\qquad c_k = \frac{\langle X(\boldsymbol{\eta})\Psi_k(\boldsymbol{\eta})\rangle}{\langle \Psi_k^2(\boldsymbol{\eta})\rangle}$$

The integral $\langle X(\boldsymbol{\eta})\Psi_k(\boldsymbol{\eta})\rangle = \int X(\boldsymbol{\eta})\Psi_k(\boldsymbol{\eta})\pi(\boldsymbol{\eta})d\boldsymbol{\eta}$ can be estimated by

- Monte-Carlo

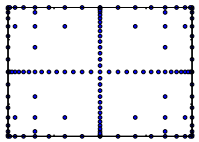
$$\frac{1}{K}\sum_{j=1}^{K} X(\boldsymbol{\eta}_j)\Psi_k(\boldsymbol{\eta}_j)$$



many samples from $\pi(\boldsymbol{\eta})$

- Quadrature

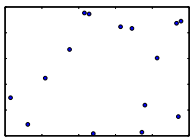
$$\sum_{j=1}^{Q} X(\boldsymbol{\eta}_j)\Psi_k(\boldsymbol{\eta}_j)w_j$$



samples at quadrature

- *Bayesian inference*

$$P(c_k|X(\boldsymbol{\eta}_j)) \propto P(X(\boldsymbol{\eta}_j)|c_k)P(c_k)$$



any (number of) samples

Bayesian inference handles the intrinsic stochasticity well in presence of limited data

$$X \simeq \sum_{k=0}^{P} c_k \Psi_k(\boldsymbol{\eta}) \equiv g_{\boldsymbol{c}}(\boldsymbol{\eta}) \qquad \overbrace{P(\boldsymbol{c}|\mathcal{D})}^{\text{Posterior}} \propto \overbrace{P(\mathcal{D}|\boldsymbol{c})}^{\text{Likelihood}} \overbrace{P(\boldsymbol{c})}^{\text{Prior}}$$

$$\text{Data} \qquad \mathcal{D} \equiv \{(\tilde{\boldsymbol{\eta}}_1, X_1), \ldots, (\tilde{\boldsymbol{\eta}}_N, X_N)\} = (\tilde{\boldsymbol{\eta}}, \boldsymbol{X})$$

$$L(\boldsymbol{c}) = P(\mathcal{D}|\boldsymbol{c}) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^N \prod_{i=1}^{N} \exp\left(-\frac{(X_i - g_{\boldsymbol{c}}(\boldsymbol{\eta}))^2}{2\sigma^2}\right)$$

- Noise model is assumed gaussian with σ^2 fixed or inferred
- Conjugate prior on $\boldsymbol{c}$, uniform or gaussian
- The whole posterior distribution is accessible, *i.e.* uncertain response surface
- Input parameters can have arbitrary values

Bayesian inference handles the intrinsic stochasticity well in presence of limited data

$$X \simeq \sum_{k=0}^{P} c_k \Psi_k(\boldsymbol{\eta}) \equiv g_{\boldsymbol{c}}(\boldsymbol{\eta}) \qquad \overbrace{P(\boldsymbol{c}|\mathcal{D})}^{\text{Posterior}} \propto \overbrace{P(\mathcal{D}|\boldsymbol{c})}^{\text{Likelihood}} \overbrace{P(\boldsymbol{c})}^{\text{Prior}}$$

$$\text{Data} \qquad \mathcal{D} \equiv \{(\tilde{\boldsymbol{\eta}}_1, X_1), \ldots, (\tilde{\boldsymbol{\eta}}_N, X_N)\} = (\tilde{\boldsymbol{\eta}}, \boldsymbol{X})$$

$$L(\boldsymbol{c}) = P(\mathcal{D}|\boldsymbol{c}) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^N \prod_{i=1}^{N} \exp\left(-\frac{(X_i - g_{\boldsymbol{c}}(\boldsymbol{\eta}))^2}{2\sigma^2}\right)$$

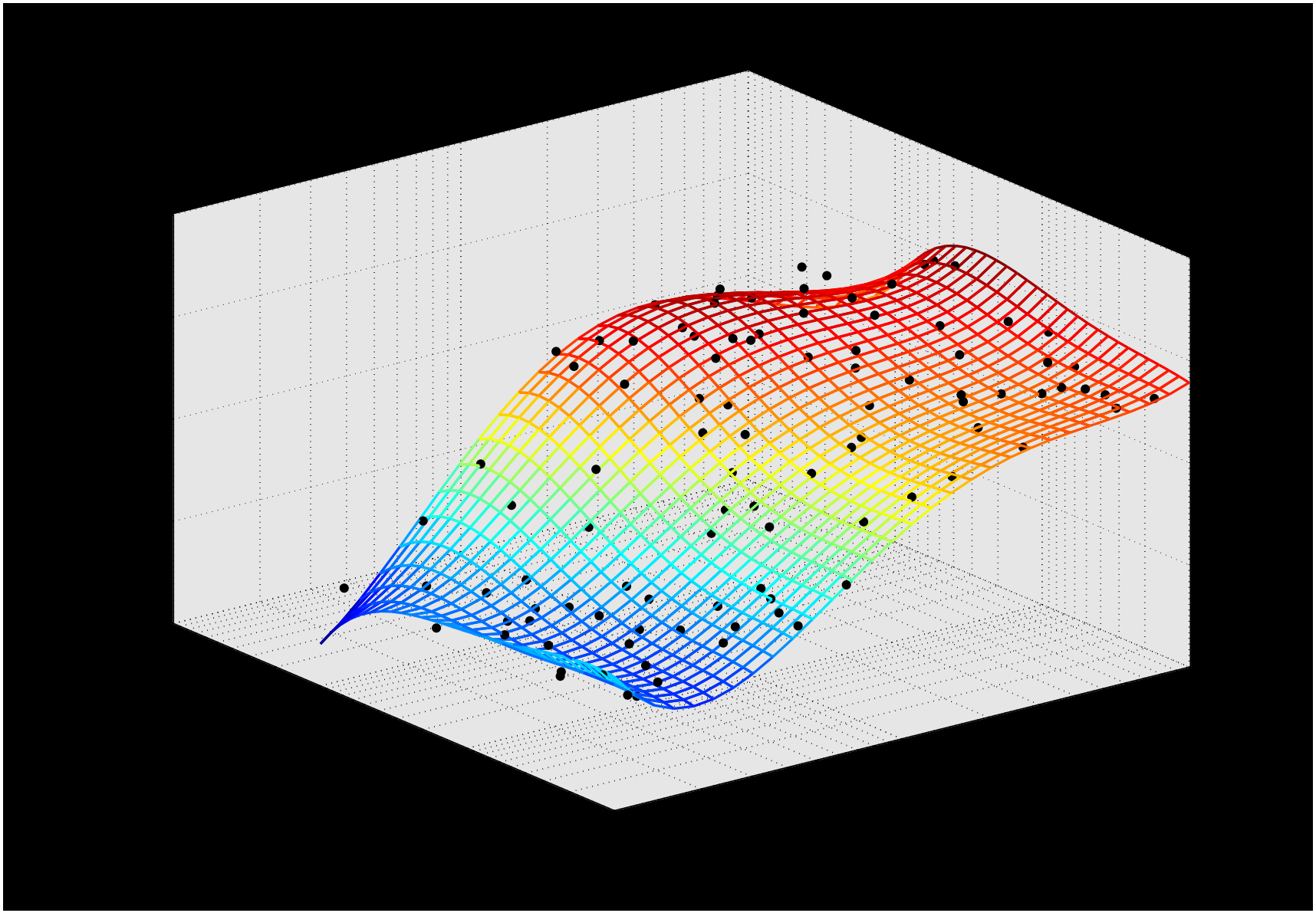
- Posterior is *a multivariate normal*

$$\boldsymbol{c} \quad \in \quad \mathcal{N}(\underbrace{(\boldsymbol{P}^T\boldsymbol{P})^{-1}\boldsymbol{P}^T\boldsymbol{X}}_{\boldsymbol{\mu}}, \underbrace{\sigma^2(\boldsymbol{P}^T\boldsymbol{P})^{-1}}_{\boldsymbol{\Sigma}}), \qquad \text{where } \boldsymbol{P}_{ik} = \Psi_k(\tilde{\boldsymbol{\eta}}_i)$$

- The response surface is *a gaussian process*

$$\sum_{k=0}^{P} c_k \Psi_k(\boldsymbol{\eta}) = \boldsymbol{\Psi}(\boldsymbol{\eta})^T \boldsymbol{c} \quad \in \quad \mathcal{GP}(\boldsymbol{\Psi}(\boldsymbol{\eta})^T\boldsymbol{\mu}, \boldsymbol{\Psi}(\boldsymbol{\eta})\boldsymbol{\Sigma}\boldsymbol{\Psi}(\boldsymbol{\eta}')^T)$$

Uncertain response surface in two-parameter case, 4th order PC



Curse of Dimensionality

- (Dim-adaptive) Sparse quadrature integration [Gerstner, 2003]
- High Dimensional Model Representation [Rabitz & Alis, 1999]
  - would not handle strong nonlinearities
  - tried cut-HDMR in a chemical kinetics context: fails!
- Proper Generalized Decomposition [Nuoy, 2010]

- Turn it into the *blessing of dimensionality* [Donoho, 2000]
- Compressive sensing in spectral methods [Doostan *et al.*, 2009]
- Bayesian compressive sensing [Ji *et al.*, 2008]

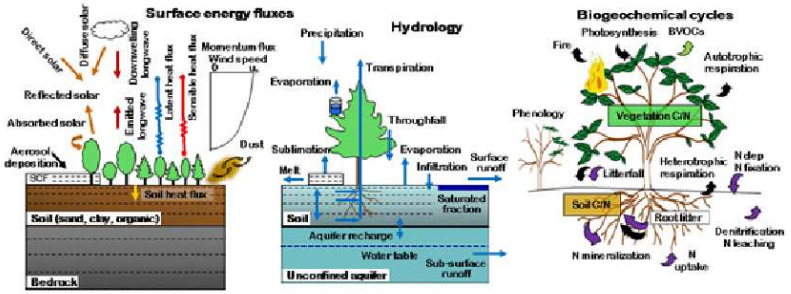
Curse of Dimensionality

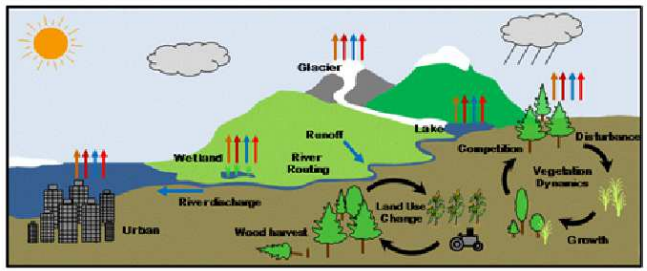
- (Dim-adaptive) Sparse quadrature integration [Gerstner, 2003]
- High Dimensional Model Representation [Rabitz & Alis, 1999]
  - would not handle strong nonlinearities
  - tried cut-HDMR in a chemical kinetics context: fails!
- Proper Generalized Decomposition [Nuoy, 2010]

- Turn it into the *blessing of dimensionality* [Donoho, 2000]
- Compressive sensing in spectral methods [Doostan *et al.*, 2009]
- Bayesian compressive sensing [Ji *et al.*, 2008]

short answer: no free lunch

Application of Interest: Community Land Model





http://www.cesm.ucar.edu/models/clm/

- Nested computational grid hierarchy
- Represents spatial heterogeneity of the land surface
- A single-site, 1000-yr simulation takes ~ 10 hrs on 1 CPU
- Involves ~ 80 input parameters

Polynomial Chaos Surrogate for CLM

Input parameters are represented via their cumulative distribution function (CDF) $F(\cdot)$, such that, with $\eta_i \sim \text{Uniform}[-1,1]$, we have:

$$\lambda_i = F_{\lambda_i}^{-1}\left(\frac{\eta_i+1}{2}\right), \qquad \text{for } i = 1,2,\ldots,d.$$

If input parameters are uniform $\lambda_i \sim \text{Uniform}[a_i, b_i]$, then

$$\lambda_i = \frac{a_i+b_i}{2} + \frac{b_i-a_i}{2}\,\eta_i.$$

Output is represented with respect to Legendre polynomials

$$f(\boldsymbol{\lambda}(\boldsymbol{\eta})) \approx y_{\boldsymbol{c}}(\boldsymbol{\eta}) \equiv \sum_{k=0}^{K} c_k \Psi_k(\boldsymbol{\eta}).$$

Iterative Bayesian Compressive Sensing (iBCS)

- The number of polynomial basis terms grows fast; a p-th order, d-dimensional basis has a total of $(p+d)!/(p!d!)$ terms.
- Dimensionality reduction by using Gaussian *sparsity* priors.

$$p(\boldsymbol{c}) \propto \prod_{k=0}^{K} \exp\left(-\frac{c_k^2}{2s_k^2}\right).$$

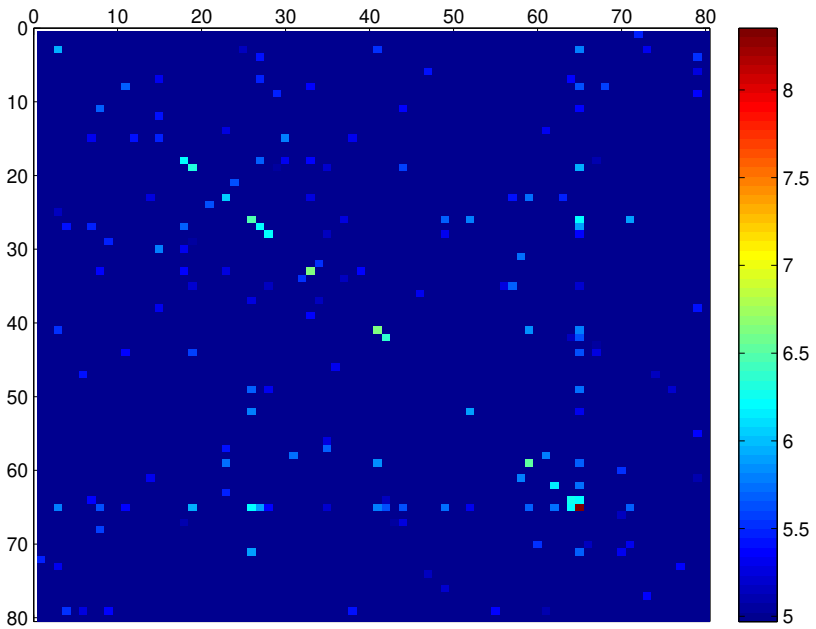
The parameters $(\sigma^2, s_0^2, \ldots, s_K^2)$ are fixed by evidence maximization, and bases corresponding to small s_i^2 are discarded [Ji *et al.* 2008].

- *Iterative BCS*: We implement an iterative procedure that allows increasing the order for the relevant basis terms while maintaining the dimensionality reduction [Sargsyan *et al.* 2012].

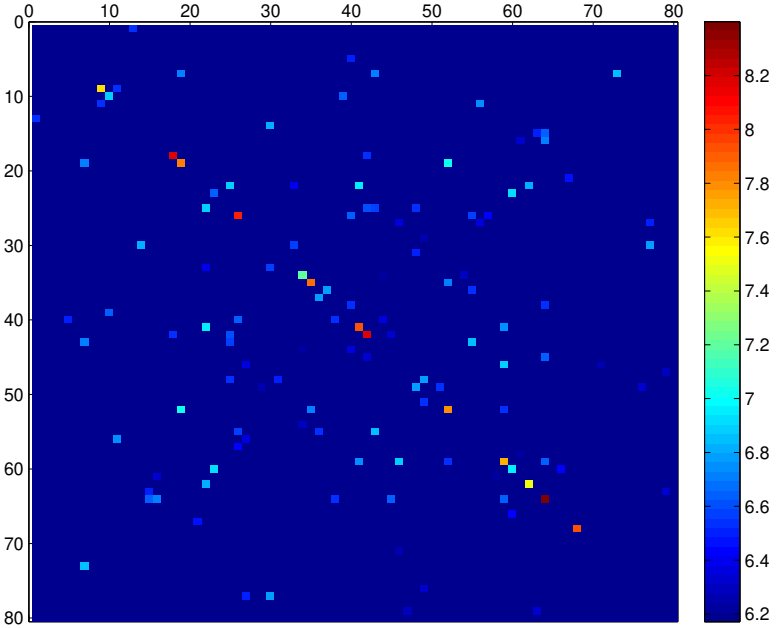
Community Land Model - 2nd order BCS

- Most influential input parameter couplings for each output
- About 100 out of 3000 terms retained

TOTVEGC



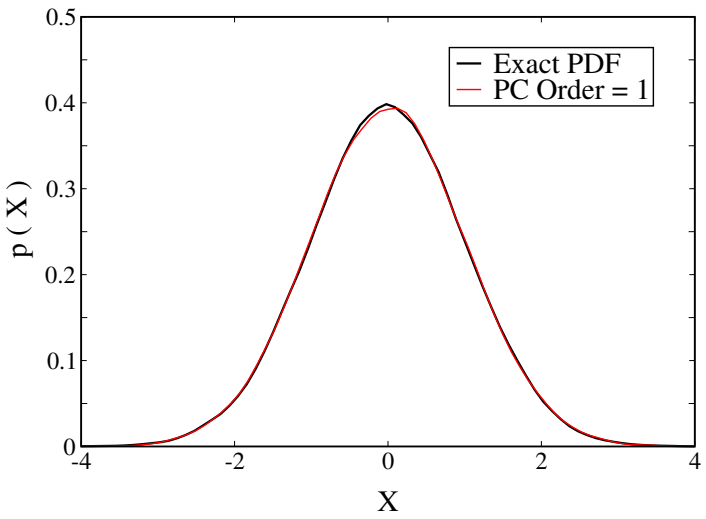
TOTSOMC



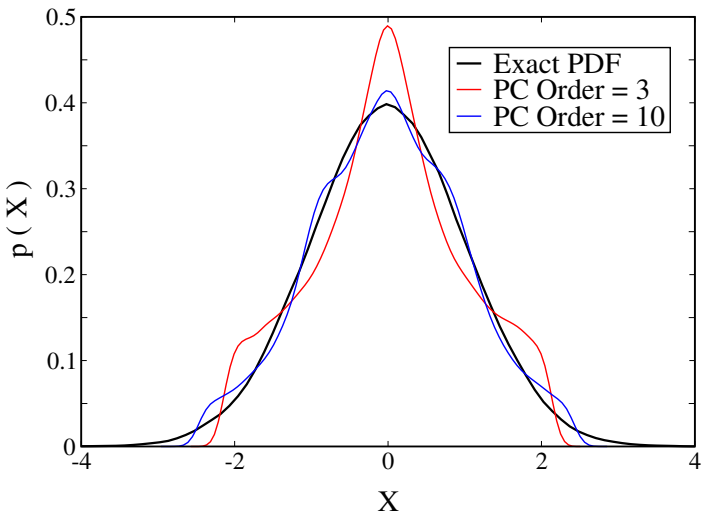
Nonlinear/bimodal r.v.

Normal Random Variable

Gauss-Hermite PC



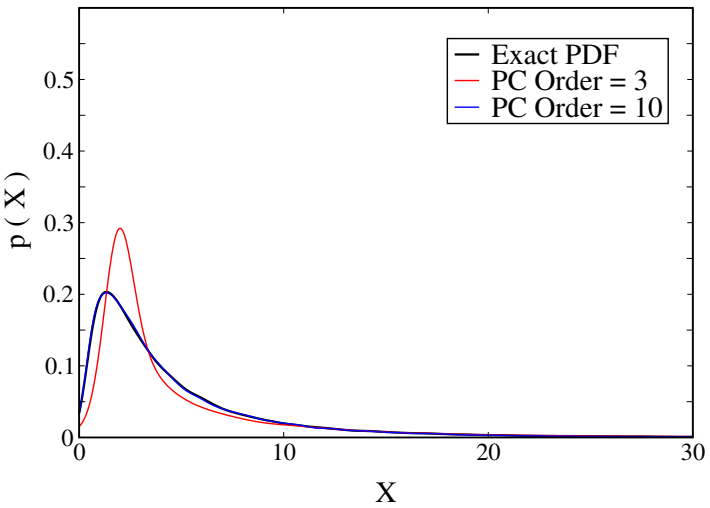
Legendre-Uniform PC



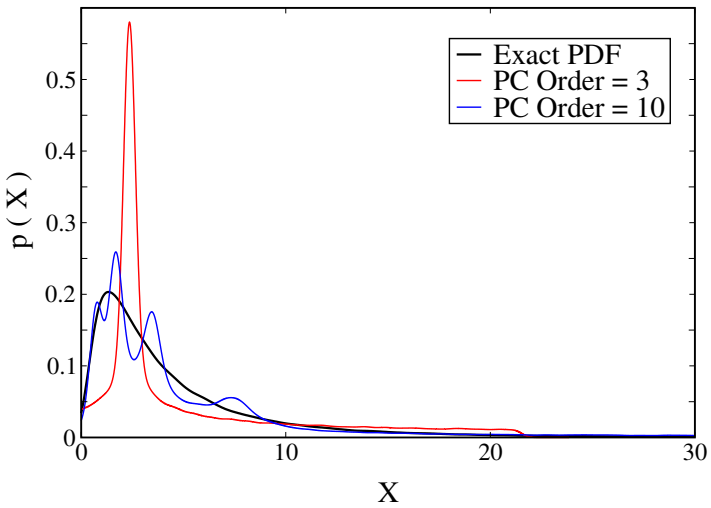
Nonlinear/bimodal r.v.

Lognormal Random Variable

Gauss-Hermite PC



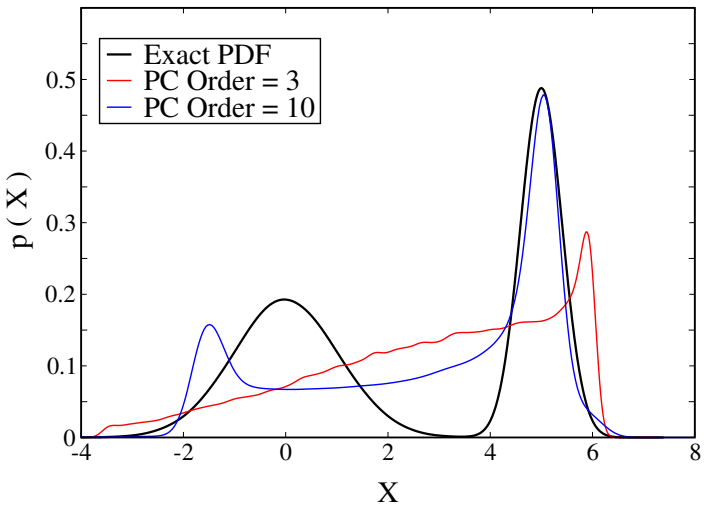
Legendre-Uniform PC



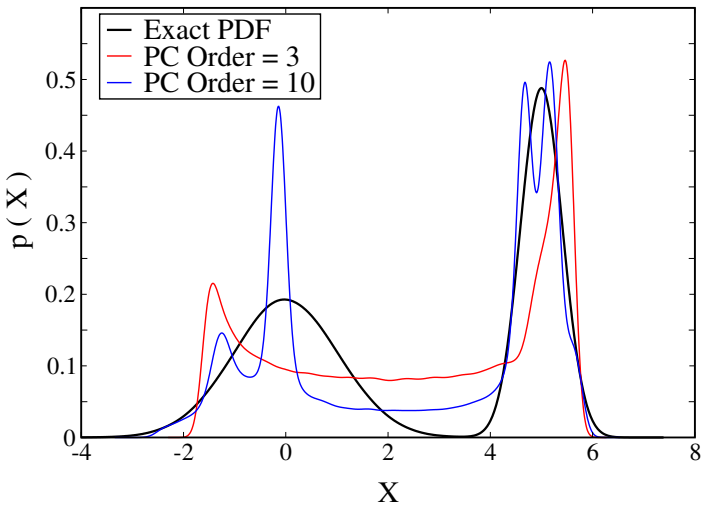
Nonlinear/bimodal r.v.

Binormal Random Variable

Gauss-Hermite PC



Legendre-Uniform PC



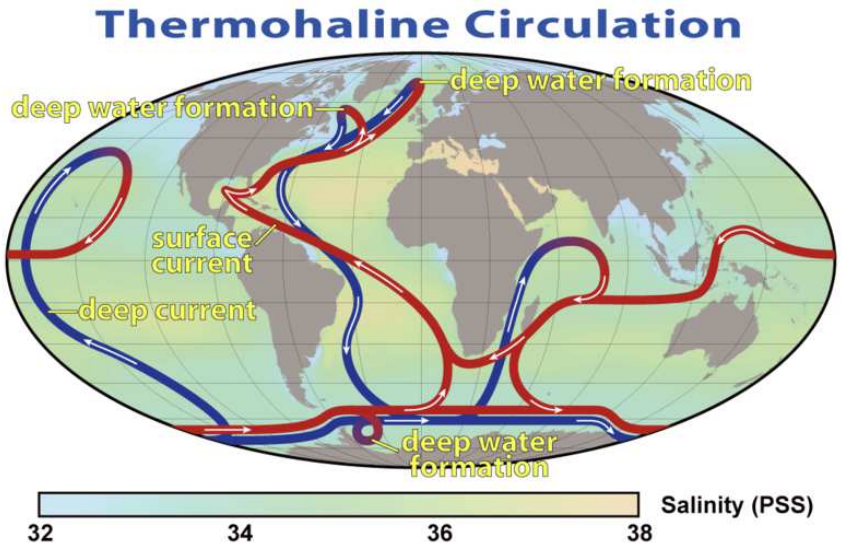
Discontinuities/Nonlinearities/Bifurcations

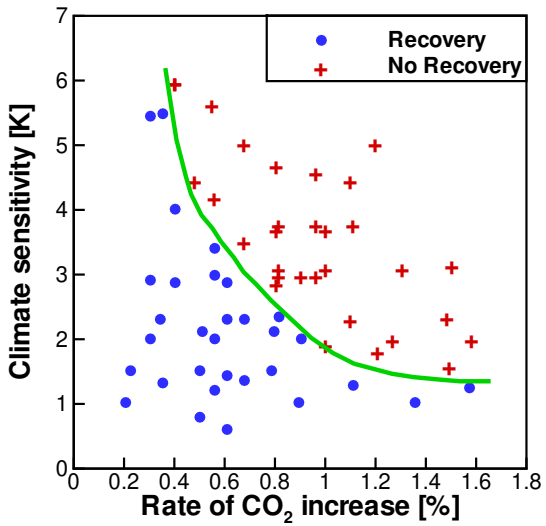
- Stochastic domain decomposition
  - Wiener-Haar expansions,
    Multiblock expansions,
    Multiwavelets, [Le Maître *et al*, 2004,2007]
  - also known as Multielement PC [Wan & Karniadakis, 2009]
- Data domain decomposition [Sargsyan *et al*, 2009,2010]
  - Data clustering, classification
  - Mixture PC expansions

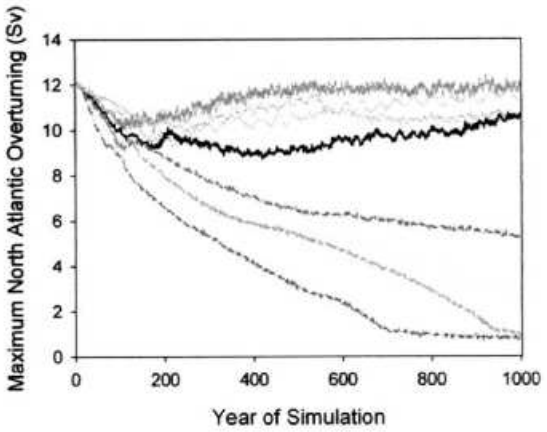
- Adaptive setting helps
- Does not scale with dimensionality
- For expensive models, can not split much
- *Need a 'smart' domain decomposition*

Webster et al - J. Environ. Syst. 31: 39-59, 2007

- Computational model - EMIC
- Input parameters
  - Rate of CO_2 increase (r)
  - Climate sensitivity (λ)
- Output observable
  - Overturning streamfunction (Z)

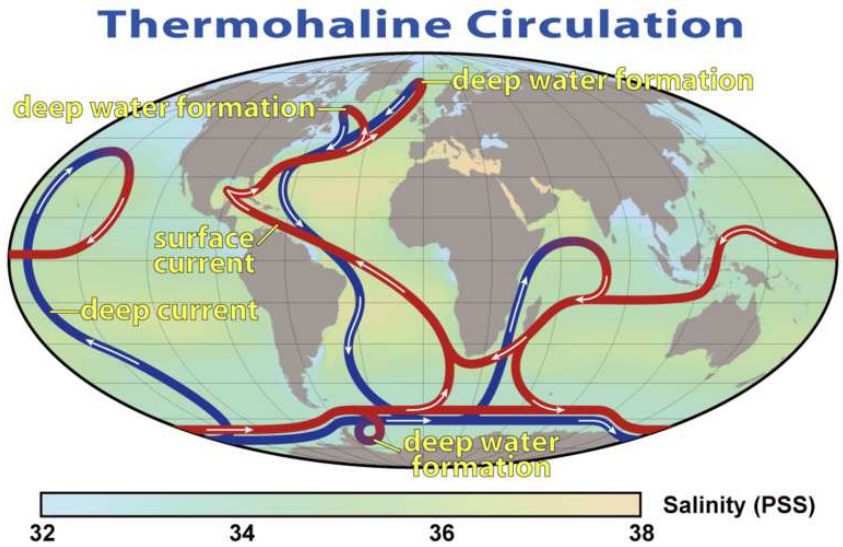


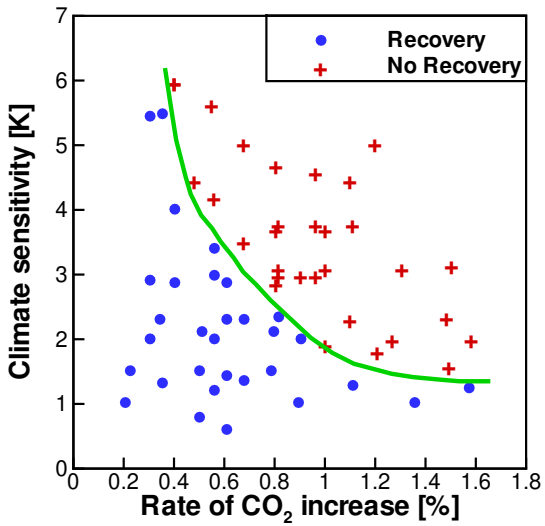


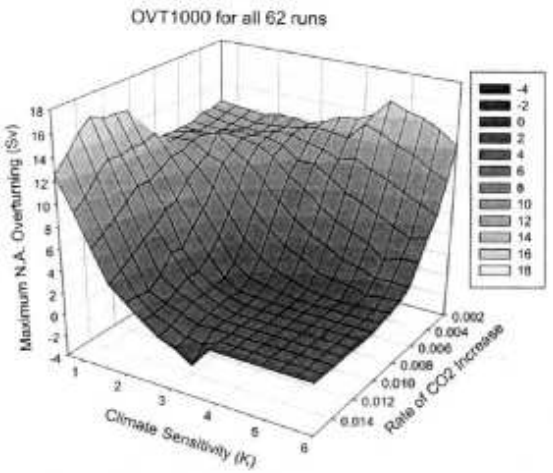


Global representations fail to capture discontinuities

- Computational model - EMIC
- Input parameters
  - Rate of CO_2 increase (r)
  - Climate sensitivity (λ)
- Output observable
  - Overturning streamfunction (Z)







UQ & Discontinuities - Proposed Methodology

Our approach locates the discontinuity first so the domain can be subdivided into regions with smooth model response where spectral uncertainty quantification methods can be used.

Two-step approach:

- **Bayesian inference of the location of the discontinuity**
- **Polynomial chaos representation via parameter domain mapping at each side of the discontinuity**

Bayesian Inference of the Location of Discontinuity

- Parameterize the discontinuity: $r \approx p_{\boldsymbol{c}}(\lambda) = \sum_{k=0}^{K} c_k \lambda^k$

- Approximation model:

$$\mathcal{M}_{\boldsymbol{c}} \equiv g(\lambda, r) = m_L + (m_R - m_L)\frac{1 + \tanh\left(\alpha(r - p_{\boldsymbol{c}}(\lambda))\right)}{2}$$

- Noise model postulated: $\sigma(\lambda, r)$

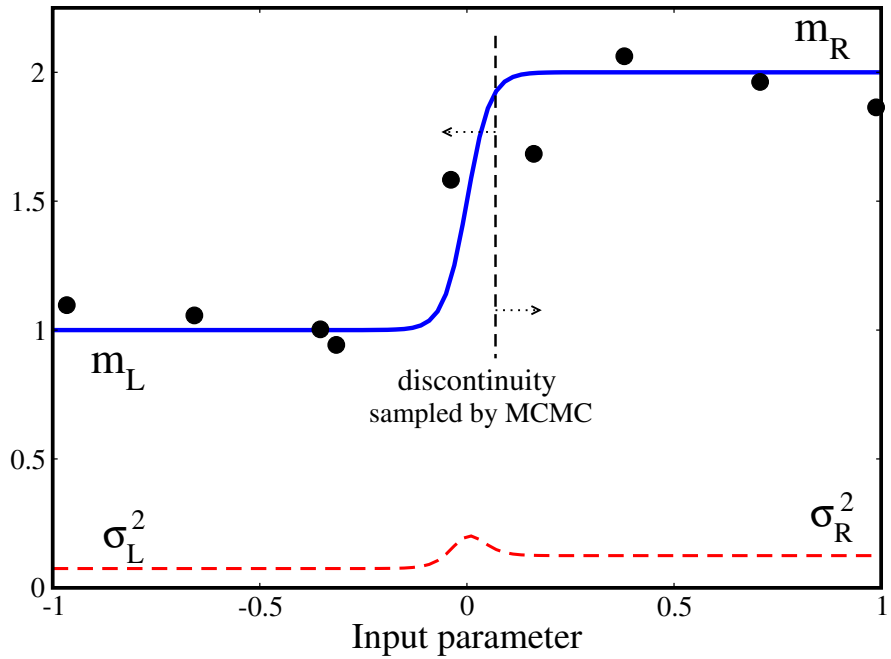
- Likelihood function:

$$\log P(\mathcal{D}|\mathcal{M}_{\boldsymbol{c}}) = \sum_{i=1}^{N} \log\left(P(z_i|\mathcal{M}_{\boldsymbol{c}})\right) = -\sum_{i=1}^{N} \frac{\left(z_i - g(\lambda, r)\right)^2}{2\sigma(\lambda, r)^2}.$$

Bayesian Inference of the Location of Discontinuity

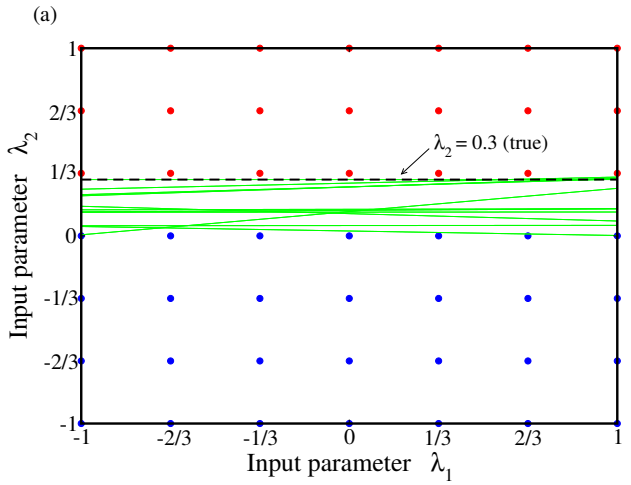
- Parameterize the discontinuity: $r \approx p_{\boldsymbol{c}}(\lambda) = \sum_{k=0}^{K} c_k \lambda^k$

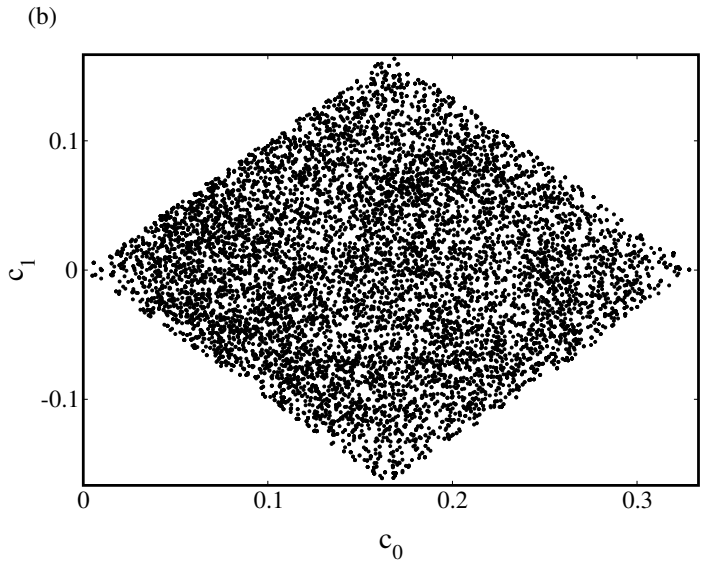
- Bayes' formula: $P(\mathcal{M}|\mathcal{D}) = \frac{P(\mathcal{D}|\mathcal{M})P(\mathcal{M})}{P(\mathcal{D})}$

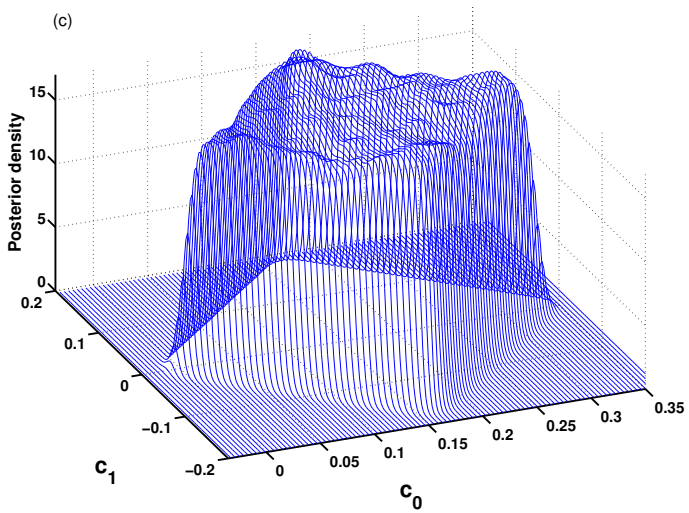


Highlights

- Any distribution of input points
- Generalizes to multiple dimensions
- Probabilistic representation







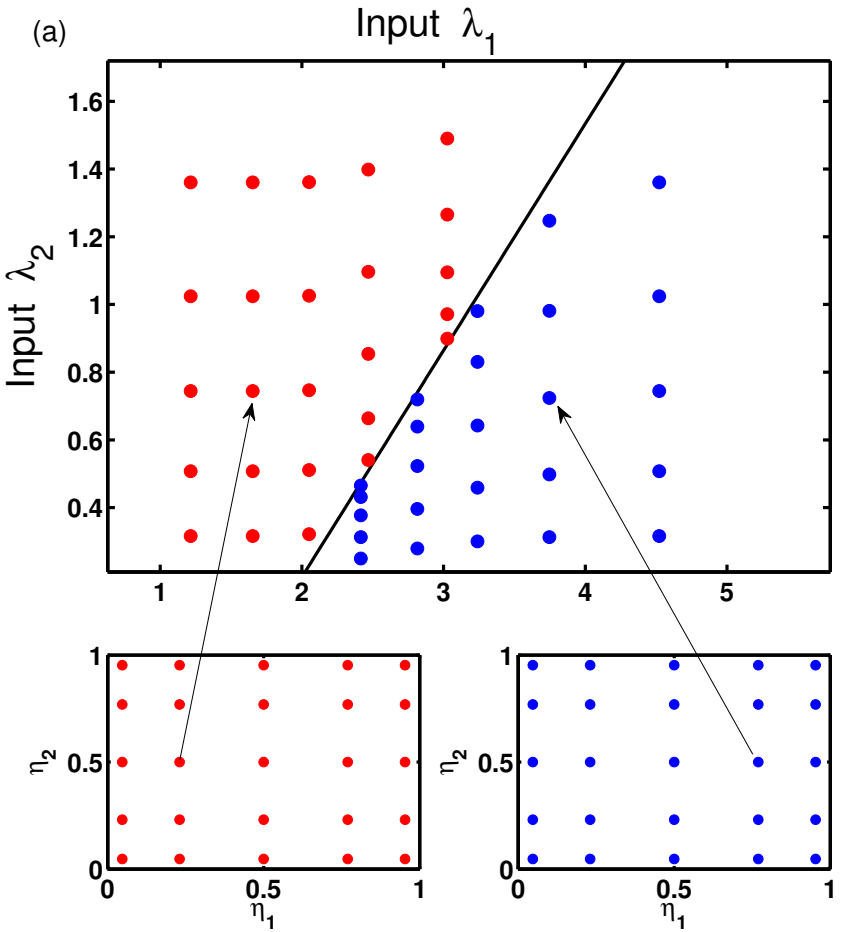
Discontinuity curve samples and their pdf

Parameter Domain Mapping

- Assume linear discontinuity
- Use Rosenblatt Transformation (RT) to map the pair of uncertain parameters (λ,r) to i.i.d. uniform random variables η_1 and η_2:

$$\begin{aligned} \lambda &= F_{\lambda}^{-1}(\eta_1), \\ r &= F_{r|\lambda}^{-1}(\eta_2|\eta_1) \end{aligned}$$

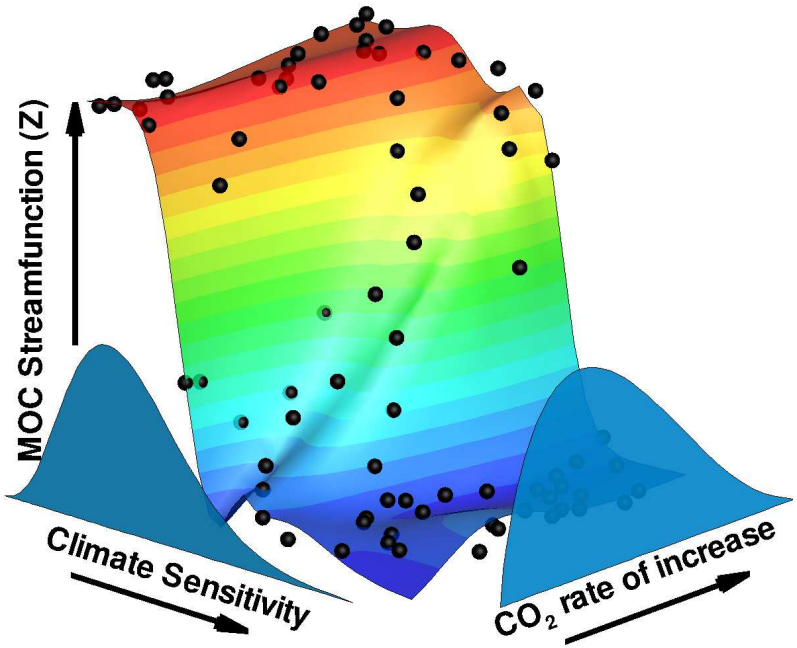
- Apply the RT mapping to both sides of the discontinuity



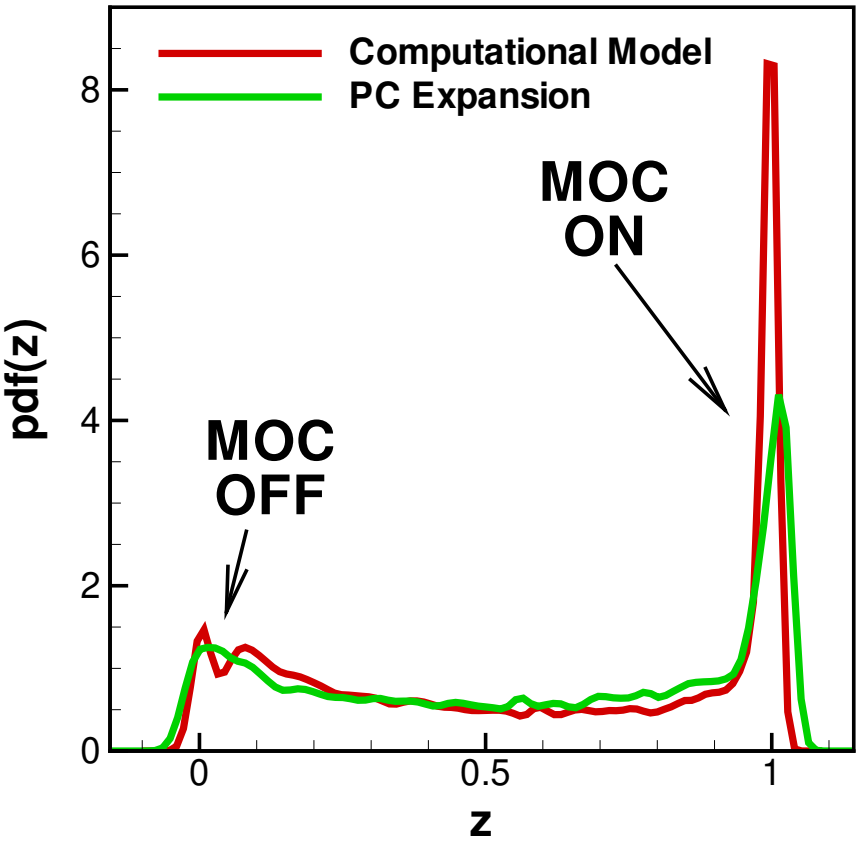
ROSENBLATT TRANSFORMATION: $(\lambda, r) \rightarrow (\eta_1, \eta_2)$

Discontinuous data represented well with the averaged PC

PCE IN (η_1, η_2) DOMAIN



OUTPUT PDF



Discontinuous data represented well with the averaged PC.

Resulting output PDF given input parameter joint PDF.

Closure

- Probabilistic UQ framework
  - Probabilistic characterization of uncertain inputs
  - Forward UQ
  - PC surrogate
  - Still many challenges
- Items not covered
  - Time/space-resolved processes
    - Karhunen-Loeve expansion
  - Inverse UQ, calibration
    - Bayesian inference accelerated by PC surrogates
  - Model comparison, validation
    - Bayes Factors, Model plausibility
  - Relationship with Gaussian processes (GP)
    - GP regression is usually non-parametric
    - GP good for PDFs, PC good for sampling
  - Discrete parameters
    - Haar wavelets, mixed Haar-PC expansions

Postdoctoral Appointee

Job Details

The Reacting Flow Research Department seeks postdoctoral researchers who will pursue research in applied math, including theoretical and algorithmic developments, as well as computational investigations, with a focus on uncertainty quantification in computational modeling of physical systems. Current projects focus on applications that include, but are not limited to, climate modeling, stochastic dynamical systems, shock hydrodynamics, complex networks, and high performance computing. Postdoctoral fellows work under the supervision of one or more research staff, and frequently in collaboration with other researchers at Sandia, other National Laboratories, or in academia. Research programs are expected to result in major impact and recognition within the academic research community, as well as relevant contributions to the long-term energy and environmental mission objectives of the U.S. Department of Energy.

Required

This position requires a PhD in engineering, computational sciences, physical sciences, or mathematics. Other required qualifications include (1) expertise in uncertainty quantification in computational models, (2) a demonstrated track record in scientific research, (3) strong verbal- and written-communication skills, and (4) the ability to work effectively as part of an interdisciplinary research team. Good scientific programming skills are a plus.

Acknowledgements

Cosmin Safta, SNL
Bert Debusschere, SNL
Habib Najm, SNL
Daniel Ricciuto, ORNL
Robert Berry, Climate Corporation
Helgi Adalsteinnson, Google
Youssef Marzouk, MIT
Olivier Le Maître, LIMSI, France

- Sandia's internal Laboratory Directed Research & Development,
- DOE, Biological and Environmental Research,
- DOE, Advanced Scientific Computing Research.

Thank You!

Sandia National Laboratories is a multi-program laboratory 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.

Literature

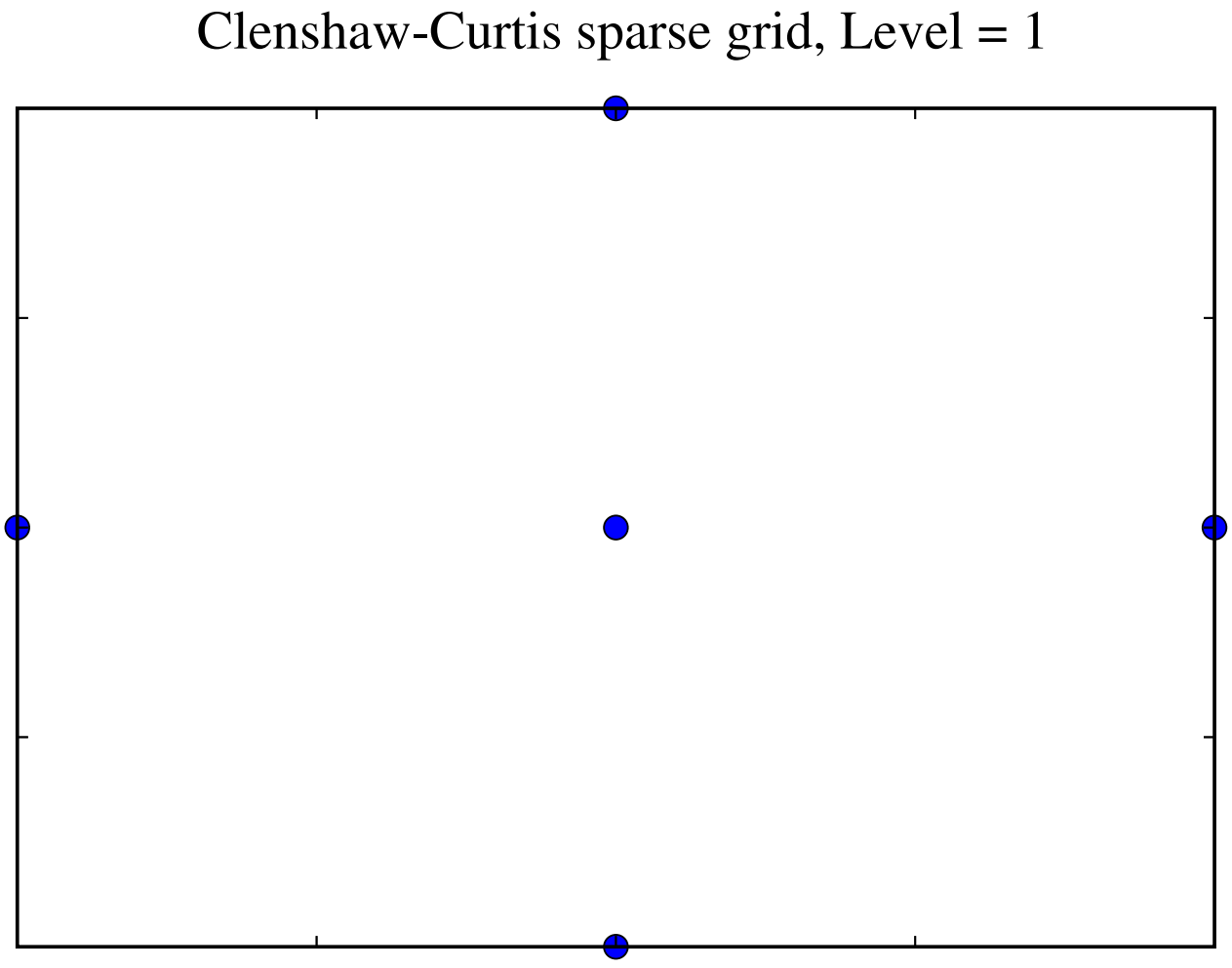
- N. Wiener, “Homogeneous Chaos”, *American Journal of Mathematics*, 60:4, pp. 897-936, 1938.
- M. Rosenblatt, “Remarks on a Multivariate Transformation”, *Ann. Math. Statist.*, 23:3, pp. 470-472, 1952.

- R. Ghanem and P. Spanos, “Stochastic Finite Elements: a Spectral Approach”, Springer, 1991.
- O. Le Maître and O. Knio, “Spectral Methods for Uncertainty Quantification with Applications to Computational Fluid Dynamics”, Springer, 2010.
- D. Xiu, “Numerical Methods for Stochastic Computations: A Spectral Method Approach”, Princeton U. Press, 2010.

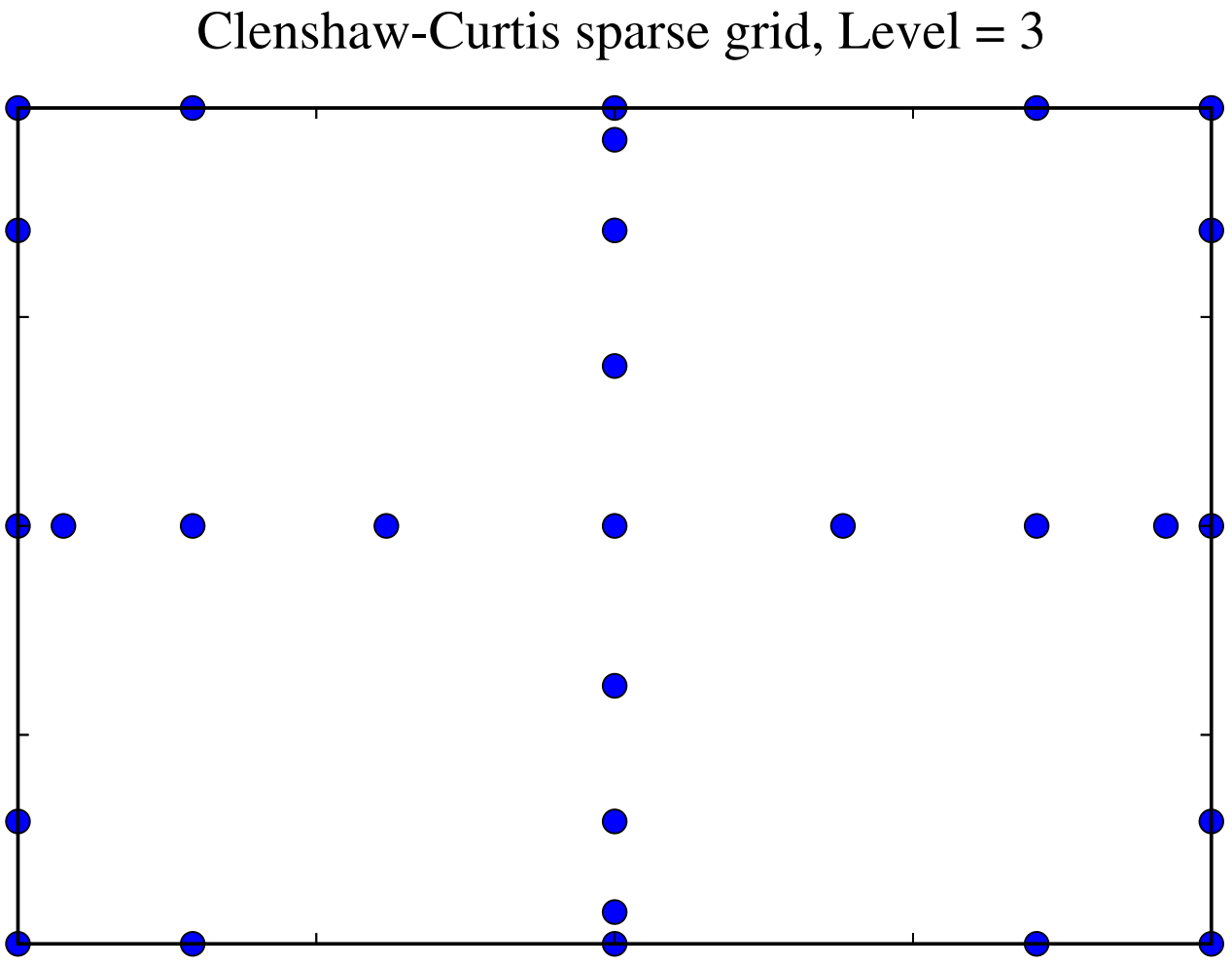
- D. Xiu and G.E. Karniadakis, “The Wiener-Askey Polynomial Chaos for Stochastic Differential Equations”, *SIAM J. Sci. Comput.*, 24:2, 2002.
- B. Debusschere, H. Najm, P. Pébay, O. Knio, R. Ghanem and O. Le Maître, “Numerical Challenges in the Use of Polynomial Chaos Representations for Stochastic Proecesses”, *SIAM J. Sci. Comp.*, 26:2, 2004.
- S. Ji, Y. Xue and L. Carin, “Bayesian Compressive Sensing”, *IEEE Trans. Signal Proc.*, 56:6, 2008.

- K. Sargsyan, C. Safta, B. Debusschere and H. Najm, “Multiparameter spectral representation of competence dynamics in Bacillus Subtilis”. *Submitted to IEEE Trans. Comp. Biol. and Bioinformatics*, 2012.
- K. Sargsyan, C. Safta, B. Debusschere and H. Najm, “Uncertainty quantification given discontinuous model response and a limited number of model runs”. *SIAM J. Sci. Comp.*, 34:1, 2012.
- K. Sargsyan, C. Safta, R. Berry, J. Ray, B. Debusschere and H. Najm, “Efficient uncertainty quantification methodologies for high-dimensional climate land models”, Sandia Report, SAND2011-8757, Nov. 2011.
- K. Sargsyan, B. Debusschere, H. Najm and O. Le Maître, “Spectral representation and reduced order modeling of the dynamics of stochastic reaction networks via adaptive data partitioning”. *SIAM J. Sci. Comp.*, 31:6, 2010.
- K. Sargsyan, B. Debusschere, H. Najm and Y. Marzouk, “Bayesian inference of spectral expansions for predictability assessment in stochastic reaction networks”. *J. Comp. Theor. Nanosc.*, 6:10, 2009.

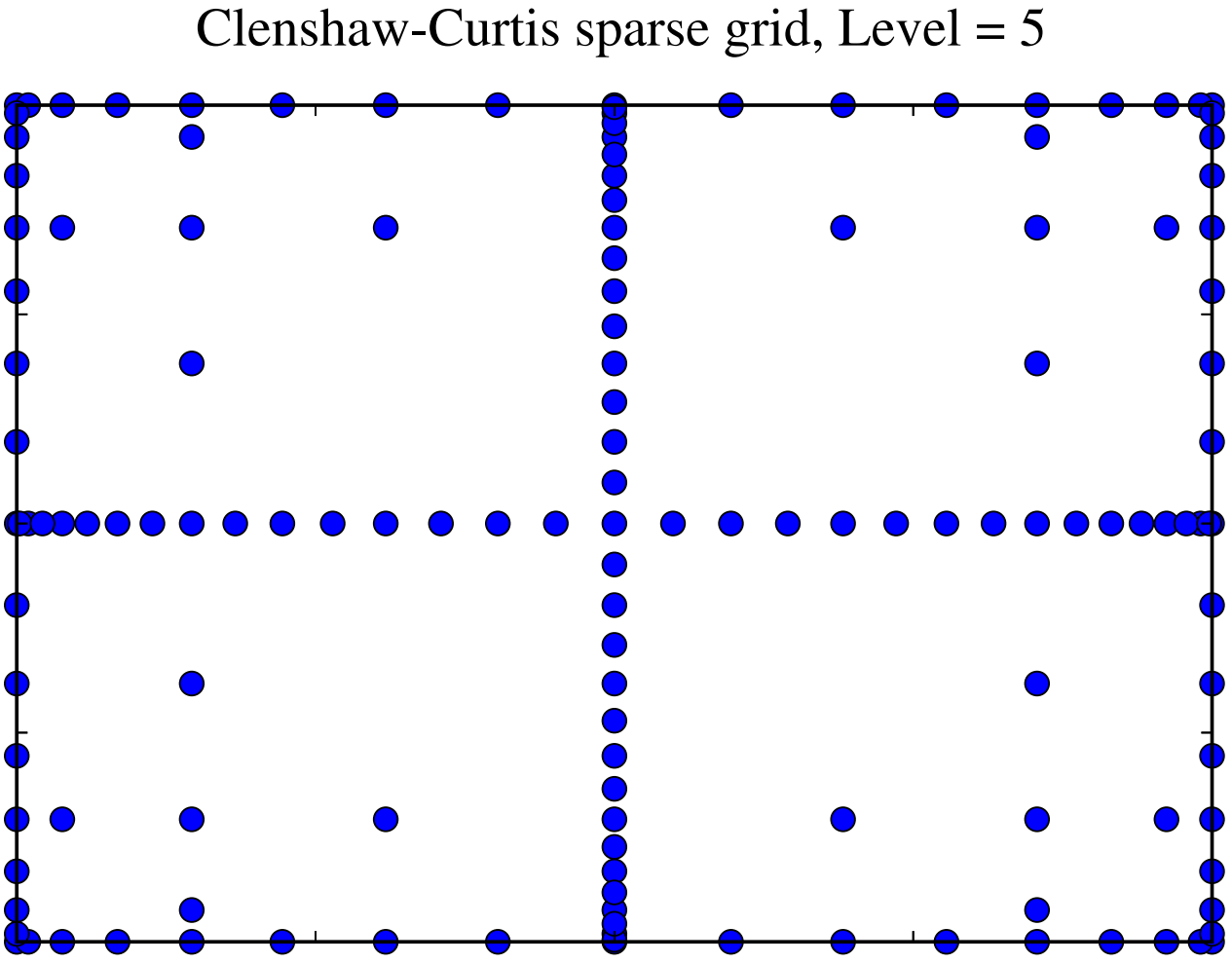
Sparse quadrature integration well-suited for high-dimensional *smooth* integrands



Sparse quadrature integration well-suited for high-dimensional *smooth* integrands



Sparse quadrature integration well-suited for high-dimensional *smooth* integrands

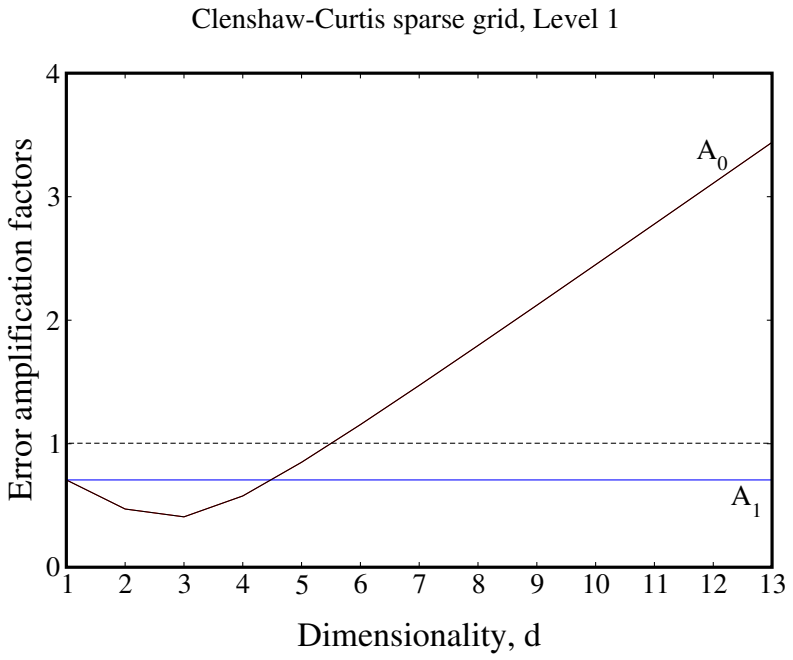


Sparse quadrature integration fails for noisy integrands

$$Y \simeq \sum_{k=0}^{P} c_k \Psi_k(\boldsymbol{\eta})$$

$$c_k = \frac{\langle Y(\boldsymbol{\eta})\Psi_k(\boldsymbol{\eta})\rangle}{\langle \Psi_k^2(\boldsymbol{\eta})\rangle}$$

$$c_k \approx \frac{1}{\langle \Psi_k^2(\boldsymbol{\eta})\rangle} \sum_{j=1}^{Q} Y(\boldsymbol{\eta}_j)\Psi_k(\boldsymbol{\eta}_j) w_j$$



$\text{Noise}_Y \sim \sigma \Longrightarrow \text{Error}_{c_k} \sim A_k \sigma$

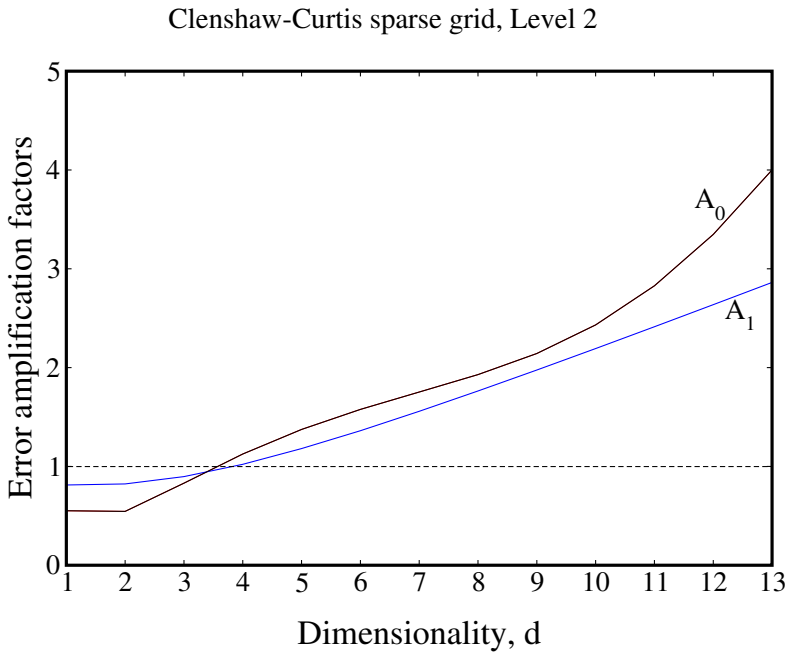
- amplification factor A_k grows with dimensionality
  - CC, level 1: $A_0 = \frac{1}{3}\sqrt{(d-3)^2 + \frac{d}{2}}, \qquad A_1 = \frac{1}{\sqrt{2}}.$
- blame the negative weights.
- for full quadrature, $\frac{1}{n^{d/2}} \leq A_0 \leq 1$, no amplification!

Sparse quadrature integration fails for noisy integrands

$$Y \simeq \sum_{k=0}^{P} c_k \Psi_k(\boldsymbol{\eta})$$

$$c_k = \frac{\langle Y(\boldsymbol{\eta})\Psi_k(\boldsymbol{\eta})\rangle}{\langle \Psi_k^2(\boldsymbol{\eta})\rangle}$$

$$c_k \approx \frac{1}{\langle \Psi_k^2(\boldsymbol{\eta})\rangle} \sum_{j=1}^{Q} Y(\boldsymbol{\eta}_j)\Psi_k(\boldsymbol{\eta}_j) w_j$$



$\text{Noise}_Y \sim \sigma \Longrightarrow \text{Error}_{c_k} \sim A_k \sigma$

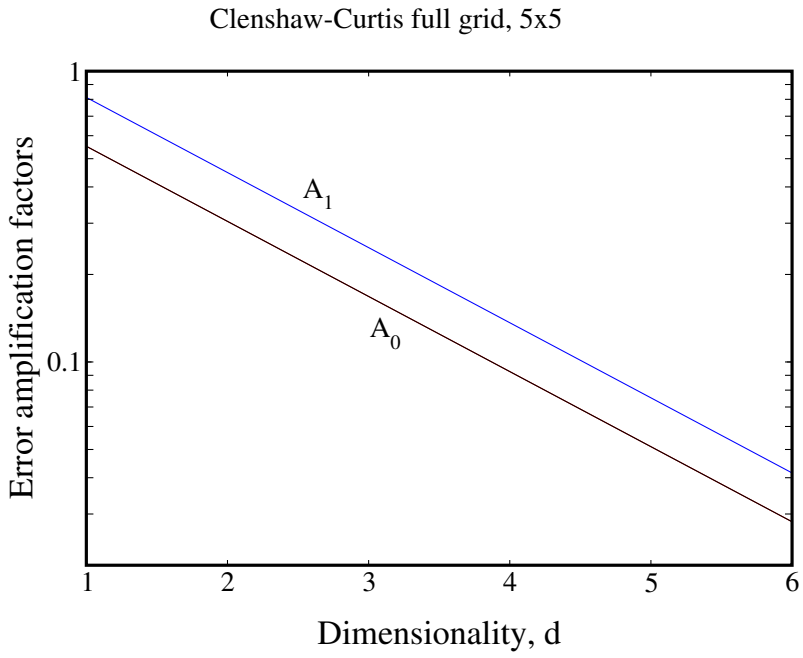
- amplification factor A_k grows with dimensionality
  - CC, level 1: $A_0 = \frac{1}{3}\sqrt{(d-3)^2 + \frac{d}{2}}, \qquad A_1 = \frac{1}{\sqrt{2}}.$
- blame the negative weights.
- for full quadrature, $\frac{1}{n^{d/2}} \leq A_0 \leq 1$, no amplification!

Sparse quadrature integration fails for noisy integrands

$$Y \simeq \sum_{k=0}^{P} c_k \Psi_k(\boldsymbol{\eta})$$

$$c_k = \frac{\langle Y(\boldsymbol{\eta})\Psi_k(\boldsymbol{\eta})\rangle}{\langle \Psi_k^2(\boldsymbol{\eta})\rangle}$$

$$c_k \approx \frac{1}{\langle \Psi_k^2(\boldsymbol{\eta})\rangle} \sum_{j=1}^{Q} Y(\boldsymbol{\eta}_j)\Psi_k(\boldsymbol{\eta}_j) w_j$$



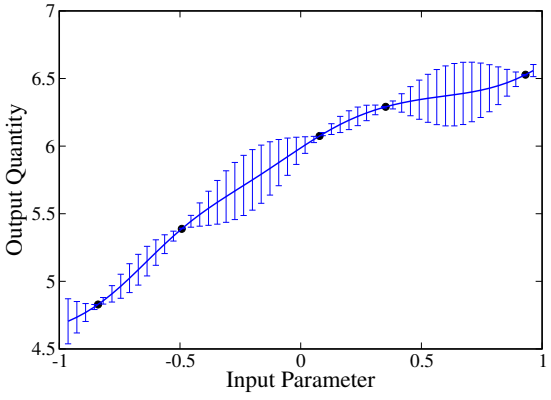
$\text{Noise}_Y \sim \sigma \Longrightarrow \text{Error}_{c_k} \sim A_k \sigma$

- amplification factor A_k grows with dimensionality
  - CC, level 1: $A_0 = \frac{1}{3}\sqrt{(d-3)^2 + \frac{d}{2}}, \qquad A_1 = \frac{1}{\sqrt{2}}.$
- blame the negative weights.
- for full quadrature, $\frac{1}{n^{d/2}} \leq A_0 \leq 1$, no amplification!

Limited Data

Both observational experiments and computer model simulations are expensive.

- Need to infer functional representation based on limited number of model runs/experiments.
  - Interpolation (kriging)
  - Gaussian Process emulation to assess the lack-of-knowledge [O'Hagan]
  - Extended to stochastic model setting



- *Bayesian* experimental design
  - What are the best locations to take observations?
  - At which parameter sets to run climate models to gain maximal information?

High Dimensional Model Representation (HDMR)

breaks the function into group-wise contributions of input variables

$$f(\boldsymbol{\lambda}) = f(\lambda_1, \ldots, \lambda_d) = f_0 + \sum_i f_i(\lambda_i) + \sum_{i<j} f_{ij}(\lambda_i, \lambda_j) + \sum_{i<j<k} f_{ijk}(\lambda_i, \lambda_j, \lambda_k) + \cdots$$

Component functions are found by

$$f_0 = \int_{R^d} f(\boldsymbol{\lambda}) d\boldsymbol{\lambda}, \qquad f_i(\lambda_i) = \int_{R^{d-1}} f(\lambda_i, \boldsymbol{\lambda}_{\bar{i}}) d\boldsymbol{\lambda}_{\bar{i}} - f_0$$

$$f_{ij}(\lambda_i, \lambda_j) = \int_{R^{d-2}} f(\lambda_i, \lambda_j, \boldsymbol{\lambda}_{\bar{ij}}) d\boldsymbol{\lambda}_{\bar{ij}} - f_i(\lambda_i) - f_j(\lambda_j) - f_0$$

- Component function $f_{i_1 \ldots i_s}(\lambda_{i_1}, \ldots, \lambda_{i_s})$ is found by a $(d-s)$-dimensional integral. Still too high-dimensional.
- Otherwise called ANOVA decomposition (analysis of variance)
- Exact in the limit, but not unique.

High Dimensional Model Representation (HDMR)

breaks the function into group-wise contributions of input variables

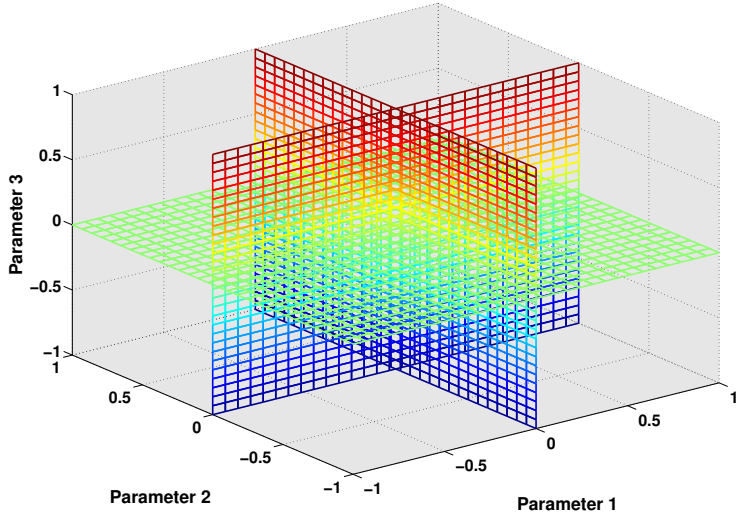
$$f(\boldsymbol{\lambda}) = f(\lambda_1, \ldots, \lambda_d) = f_0 + \sum_i f_i(\lambda_i) + \sum_{i<j} f_{ij}(\lambda_i, \lambda_j) + \sum_{i<j<k} f_{ijk}(\lambda_i, \lambda_j, \lambda_k) + \cdots$$

Component functions are found by

$$f_0 = \int_{R^d} f(\boldsymbol{\lambda}) d\boldsymbol{\lambda}, \qquad f_i(\lambda_i) = \int_{R^{d-1}} f(\lambda_i, \boldsymbol{\lambda}_{\bar{i}}) d\boldsymbol{\lambda}_{\bar{i}} - f_0$$

$$f_{ij}(\lambda_i, \lambda_j) = \int_{R^{d-2}} f(\lambda_i, \lambda_j, \boldsymbol{\lambda}_{\bar{ij}}) d\boldsymbol{\lambda}_{\bar{ij}} - f_i(\lambda_i) - f_j(\lambda_j) - f_0$$

- Two major variants:
  - Cut-HDMR
  - RS-HDMR



High Dimensional Model Representation (HDMR)

breaks the function into group-wise contributions of input variables

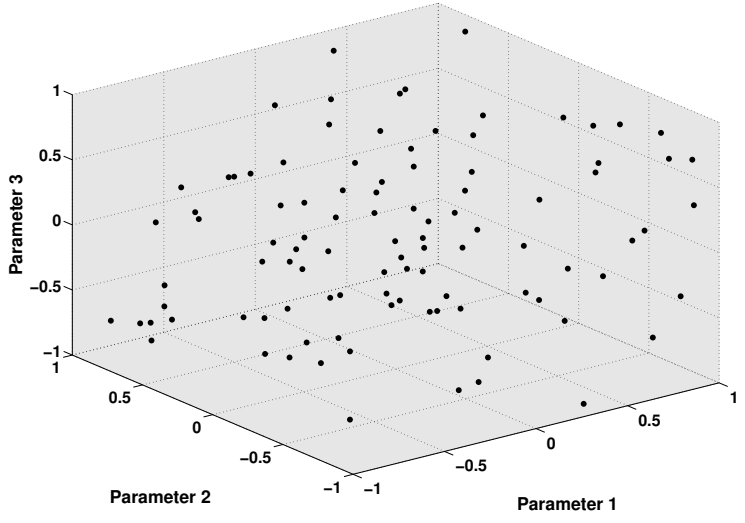
$$f(\boldsymbol{\lambda}) = f(\lambda_1, \ldots, \lambda_d) = f_0 + \sum_i f_i(\lambda_i) + \sum_{i<j} f_{ij}(\lambda_i, \lambda_j) + \sum_{i<j<k} f_{ijk}(\lambda_i, \lambda_j, \lambda_k) + \cdots$$

Component functions are found by

$$f_0 = \int_{R^d} f(\boldsymbol{\lambda}) d\boldsymbol{\lambda}, \qquad f_i(\lambda_i) = \int_{R^{d-1}} f(\lambda_i, \boldsymbol{\lambda}_{\bar{i}}) d\boldsymbol{\lambda}_{\bar{i}} - f_0$$

$$f_{ij}(\lambda_i, \lambda_j) = \int_{R^{d-2}} f(\lambda_i, \lambda_j, \boldsymbol{\lambda}_{\bar{ij}}) d\boldsymbol{\lambda}_{\bar{ij}} - f_i(\lambda_i) - f_j(\lambda_j) - f_0$$

- Two major variants:
  - Cut-HDMR
  - RS-HDMR

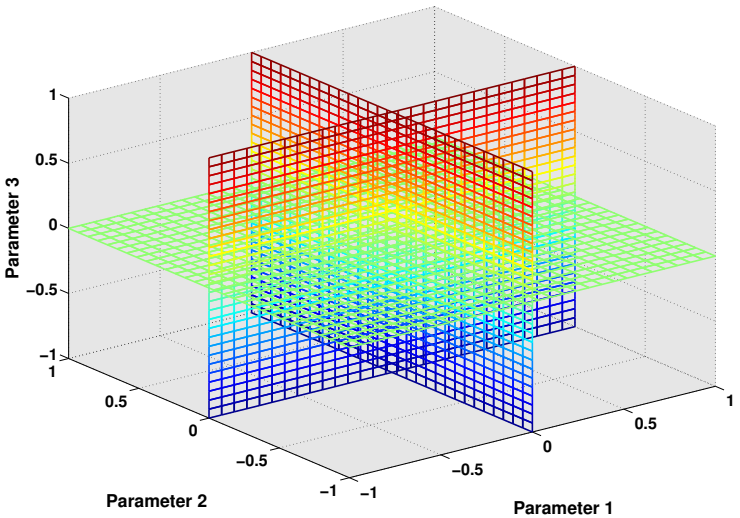


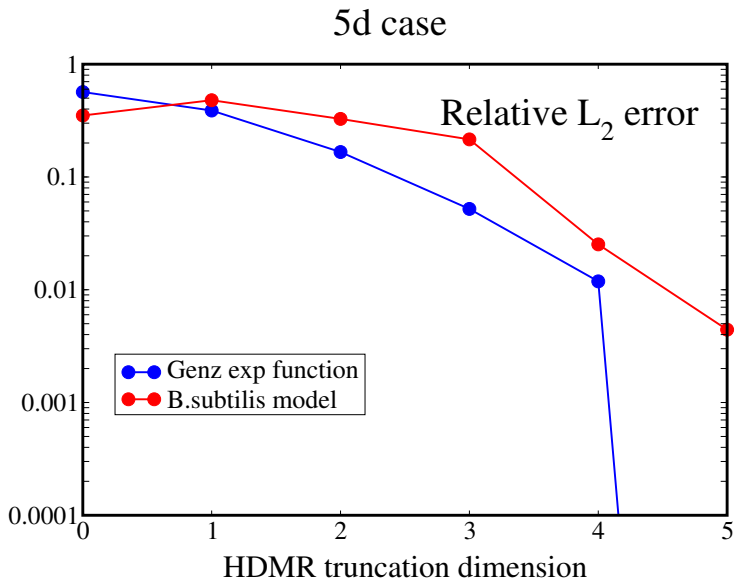
cut-HDMR disregards corners in the parameter space and does not guarantee accuracy in general

- Component functions are

$$f_0 = f(\boldsymbol{\lambda}^a), \qquad f_i(\lambda_i) = f(\lambda_i, \boldsymbol{\lambda}_{\bar{i}}^a) - f_0$$
$$f_{ij}(\lambda_i, \lambda_j) = f(\lambda_i, \lambda_j, \boldsymbol{\lambda}_{\bar{ij}}^a) - f_i(\lambda_i) - f_j(\lambda_j) - f_0$$

- Relies on values at lower-dimensional hyperplanes
- Depends on the anchor point $\boldsymbol{\lambda}^a$
- Does not account for ‘corners’



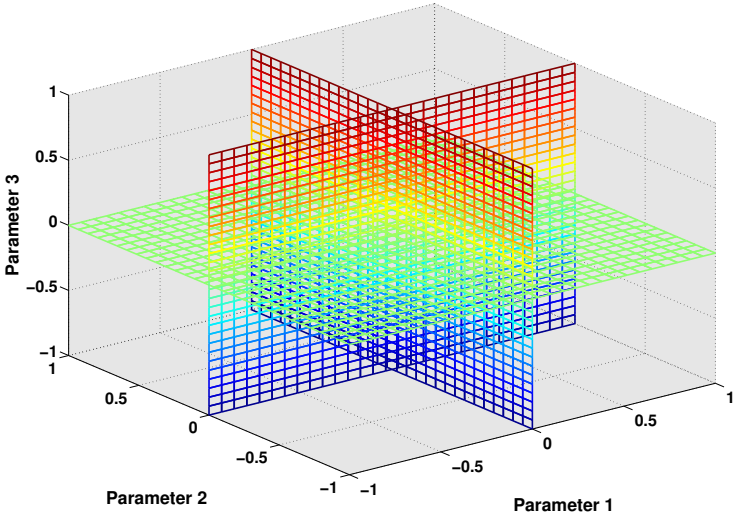


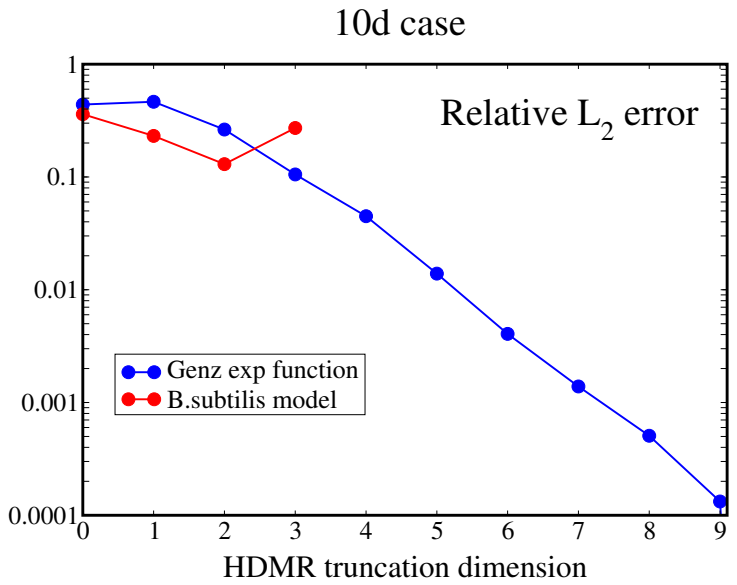
cut-HDMR disregards corners in the parameter space and does not guarantee accuracy in general

- Component functions are

$$f_0 = f(\boldsymbol{\lambda}^a), \qquad f_i(\lambda_i) = f(\lambda_i, \boldsymbol{\lambda}_{\bar{i}}^a) - f_0$$
$$f_{ij}(\lambda_i, \lambda_j) = f(\lambda_i, \lambda_j, \boldsymbol{\lambda}_{\bar{ij}}^a) - f_i(\lambda_i) - f_j(\lambda_j) - f_0$$

- Relies on values at lower-dimensional hyperplanes
- Depends on the anchor point $\boldsymbol{\lambda}^a$
- Does not account for ‘corners’





Random Sampling (RS) HDMR in principle equivalent to PC expansion with Monte-Carlo integration

$$f(\boldsymbol{\lambda}) = f(\lambda_1, \ldots, \lambda_d) = f_0 + \sum_i f_i(\lambda_i) + \sum_{i<j} f_{ij}(\lambda_i, \lambda_j) + \sum_{i<j<k} f_{ijk}(\lambda_i, \lambda_j, \lambda_k) + \cdots$$

- Component functions are found by

$$f_0 = \int_{R^d} f(\boldsymbol{\lambda}) d\boldsymbol{\lambda}, \qquad f_i(\lambda_i) = \int_{R^{d-1}} f(\lambda_i, \boldsymbol{\lambda}_{\bar{i}}) d\boldsymbol{\lambda}_{\bar{i}} - f_0$$

$$f_{ij}(\lambda_i, \lambda_j) = \int_{R^{d-2}} f(\lambda_i, \lambda_j, \boldsymbol{\lambda}_{\bar{ij}}) d\boldsymbol{\lambda}_{\bar{ij}} - f_i(\lambda_i) - f_j(\lambda_j) - f_0$$

- MC integrals still too expensive (new random samples needed for each hyperplane)
- Represent component functions with a polynomial expansion and use same set of samples
- Equivalent to Monte-Carlo PC with reordered multiindices!
  - PC (total order):
    $f(\xi_1, \xi_2) = 1 + [\xi_1 + \xi_2] + [(\xi_1^2 - 1) + \xi_1\xi_2 + (\xi_2^2 - 1)] + \ldots$
  - RS-HDMR:
    $f(\xi_1, \xi_2) = 1 + [\xi_1 + (\xi_1^2 - 1)] + [\xi_2 + (\xi_2^2 - 1)] + \xi_1\xi_2 \ldots$
- In future: employ Bayesian inference on component functions.