Uncertainty Quantification in Random Heterogeneous Media

Nicholas Zabaras
Materials Process Design and Control Laboratory
Sibley School of Mechanical and Aerospace Engineering
101 Rhodes Hall
Cornell University
Ithaca, NY 14853-3801
Email: nzabaras@gmail.com
URL: http://mpdc.mae.cornell.edu/

Work supported by: OSD/AFOSR MURI on Uncertainty Quantification, AFOSR Computational Mathematics, NSF Computational Mathematics & Materials Programs, DOE Office of Science, ...

School of Mechanical Engineering, Purdue University, April 12, 2012
Thermal and fluid transport in heterogeneous media are ubiquitous.

- Range from large scale systems (geothermal systems) to the small scale.
- Most critical devices/applications utilize heterogeneous/polycrystalline/functionally graded materials.
- Properties depend on the distribution of material/microstructure.

- But only possess limited information about the microstructure/property distribution (e.g. 2D images).

Incorporate limited information into stochastic analysis:
- worst case scenarios
- variations on physical properties.

Thermal transport through polycrystalline and functionally graded materials.

Hydrodynamic transport through heterogeneous permeable media.
PROBLEM DEFINITION

Observed input
\[ A = \{a(i)\}_{i=1}^{S_A} \]

Data collection

Deterministic Solver

Bayesian Training
- Tree construction
- Experimental design
- Output correlations
- HDMR terms

Surrogate Model

Reduced input space

Statistics

PDFs

Error bars

Output space
DATA-DRIVEN STOCHASTIC INPUT MODELS

Input reduction:
KPCA, Isomap, LLE, LEM, HEM, GTM, etc.

Density estimation:
Polynomial Chaos Representation

$R(\alpha) = z$

$\alpha = C(z)$

$z_i = \sum_m \beta_{im} H_m(\xi)$

Input reconstruction:
Optimization problem constrained on physical info.
Coupling data driven model generation with a Multiscale stochastic modeling framework

Seamlessly couple stochastic analysis with multiscale analysis.

Multiscale framework (large deformation/thermal evolution) + Adaptive stochastic collocation framework

Provides roadmap to efficiently link any validated multiscale framework

Coupled with a data-driven input model strategy to analyze realistic stochastic multiscale problems.

Mean statistics

Higher-order statistics

Limited data

Stochastic multiscale framework

Statistics extraction + model generation
Some input parameters, boundary/initial conditions or operational sequences have a significant impact on the stochastic solution.

Improved input representation (i.e. lesser uncertainty) significantly reduces the uncertainty in the final solution.

The stochastic solution is sensitive to which parameters or input terms? Need a robust, seamlessly interfacing methodology to estimate this.

- Which input parameters to consider as random?
- Quantify the effect of uncertainty on each parameter on system response
- Decide the length scale of experimentation and the needed fidelity of each measurement.
Sparse Grid Collocation:

- Interpolation based method
- Use hierarchical, adaptive sparse grids with basis functions of local support
- Integrate with HDMR (ANOVA) modeling

\[
f(Y) = f_0 + \sum_{i=1}^{N} f_i(Y_i) + \sum_{i_1 < i_2}^{N} f_{i_1i_2}(Y_{i_1}, Y_{i_2}) + \sum_{i_1 < i_2 < i_3}^{N} f_{i_1i_2i_3}(Y_{i_1}, Y_{i_2}, Y_{i_3}) + \ldots
\]

Multi-output Gaussian Process Modeling:

- GP non-stationary treed model
- Predictive variance biased by input PDF to define the tree
- Derive a predictive distribution from the given data
The left vertical wall (cold) is supposed to be a one dimensional Gaussian stochastic process with mean $-0.5$ and exponential covariance

\[
\text{Cov}[x_1, x_2] = s^2 \exp\left\{ -\frac{|x_1 - x_2|}{L_C} \right\},
\]

where $s^2$ is the variance of the signal and $L_C$ the correlation length.

We set $L_C = 1$, $s = 1$, $Pr = 1$, $Ra = 5000$. We test against 80000 MC samples.

The problem is solved using spectral elements (Nektar).
This is a Bayesian implementation (to be discussed later on) - Multi-output Gaussian Process Modeling.

Prediction for the standard deviation of the velocity $u$ compared to a MC estimate for $K = 8$ input dimensions.
Prediction for the standard deviation of the temperature $T$ compared to a MC estimate for $K = 8$ input dimensions.
Natural Convection ($K = 4, \quad \delta = 10^{-5}$): Comparing the prediction at a random input point with the true response.
Numerical example: Flow through random media

Basic equation for pressure and velocity in a domain

\[ \nabla \cdot \mathbf{u} = f \quad \text{in} \quad D \]
\[ \mathbf{u} = -k \nabla p \quad \text{in} \quad D \]

where \( f(x) \) denotes the deterministic source/sink term. Homogeneous boundary condition is applied. Mixed finite element method is used to solve the deterministic problem at the collocation points.

- To impose the non-negativity of the permeability, we will treat the permeability as a log random field obtained from the K-L expansion

\[ Y(\omega) = \log(K(\omega)) = \sum_{i=1}^{N} \sqrt{\lambda_i} \phi_i(\mathbf{x}) Y_i \]

where \( Y \) is a zero mean Gaussian random field with covariance function

\[ \text{Cov}(x, y) = \sigma^2 \exp \left( -\frac{|x_1 - y_1|}{L} - \frac{|x_2 - y_2|}{L} \right) \]

where \( L \) is the correlation length and \( \sigma \) is the standard deviation.
Standard deviations for different $\sigma^2$ with $N = 500$

$e = 8.08 \times 10^{-4}$

$\sigma^2 = 0.01$

$e = 7.37 \times 10^{-4}$

$\sigma^2 = 0.25$

Adaptive Sparse Grids + Adaptive HDMR

Standard deviation of the $v$ velocity-component along the cross section $y = 0.5$ for different $\sigma^2$
For low input variability, even 1st-order expansion is accurate.

For high input variability, the 1st-order expansion deviates from MC. More component terms are needed to improve accuracy.

For moderate input variability, 1st order does not deviate significantly from MC. However, a few 2nd-order terms are still needed.

PDF of the $v$ velocity-component at point (0,0.5) for different $\sigma^2$ with $N = 500$.
Convergence of the normalized errors with $N = 500$

Direct solution of the 500 dimensional problem using ASGC is impractical due to the huge computational cost. Convergence rate deteriorates with increasing input variability. However, it is still better than that of MC.
ELLIPTIC BENCHMARK PROBLEM

\[-\nabla \cdot (a_K(\omega, \cdot) \nabla u(\omega, \cdot)) = f(\cdot), \text{ in } D = [0, 1]^2,\]

\[
u(\omega, \cdot) = 0, \text{ on } \partial D,
\]

\[
f(x, y) = 100 \cos(x) \sin(y).
\]

\[
\log(a_K(\omega, x, y) - 0.5) = 1 + \omega_1 \left(\frac{\sqrt{\pi L}}{2}\right)^{1/2} + \sum_{k=2}^K \xi_k \phi_k(x) \omega_k,
\]

\[
\xi_k := \left(\sqrt{\pi L}\right)^{1/2} \exp \left(-\frac{\left[\frac{k}{2}\right] \pi L}{8}\right), \text{ for } k \geq 2,
\]

\[
\phi_k(x) := \begin{cases} 
\sin \left(\frac{\left[\frac{k}{2}\right] \pi x}{L_p}\right), & \text{if } k \text{ is even,} \\
\cos \left(\frac{\left[\frac{k}{2}\right] \pi x}{L_p}\right), & \text{if } k \text{ is odd.}
\end{cases}
\]

\[L_p = \max\{1, 2L_c\} \text{ and } L = \frac{L_c}{L_p},\]

\[\omega_K = \mathcal{U} \left([-\sqrt{3}, \sqrt{3}]\right)\]

\[
\text{Cov}[\log(a_K - 0.5)]((x_1, y_1), (x_2, y_2)) = \exp \left(-\frac{(x_1 - x_2)^2}{L_c^2}\right).
\]
INPUT REDUCTION: Kernel PCA

Use a non-linear map

\[ \Phi : \mathbb{R}^M \rightarrow F \subset \mathbb{R}^M, \quad \alpha \rightarrow \phi(\alpha) = A, \]

to unwrap the observed data

\[ \{\alpha^{(s)}\}_{s=1}^{S_A} \rightarrow \{\Phi(\alpha^{(s)})\}_{s=1}^{S_A} = \{A^{(s)}\}_{s=1}^{S_A} \]

on \( F \) (feature space). Then, do PCA (K-L) on the feature space \( F \).

Reduction achieved by keeping just a few terms of PCA.

References: [Schlkopf et al. (1998)], [Ma & Zabarás (2011)].
Fortunately only dot products are needed:

\[ \Phi(\alpha^{(s)}) \cdot \Phi(\alpha^{(s')}) := k(\alpha^{(s)}, \alpha^{(s')}) \]

Choices of kernels:

- **Standard PCA:**
  \[ k(\alpha^{(s)}, \alpha^{(s')}) = (\alpha^{(s)} \cdot \alpha^{(s')}) \]

- **Gaussian Kernel:**
  \[ k(\alpha^{(s)}, \alpha^{(s')}) = \exp \left( -\frac{\|\alpha^{(s)} - \alpha^{(s')}\|^2}{2\sigma^2} \right) \cdot \]

- In this work, we use:
  \[ \sigma^2 = c \frac{1}{S_A} \sum_{s=1}^{S_A} \min_{j \neq i} \|\alpha^{(s)} - \alpha^{(s')}\|^2, \quad s = 1, \ldots, S_A, \]
  \[ c = \text{constant}. \]
INPUT DATA

- SGeMS (Stanford Geostatistical Modeling Software)
- 0-1 image large scale image of channelized permeability field
- Cut it in 1000 45x45 pieces to generate training set
INPUT REDUCTION: Kernel PCA

- Training set consists of 1000 samples.
- Each sample has 2025 dimensions.
INPUT REDUCTION: Input Data

- Performing KPCA reduction.
- We keep 30 eigenvalues of feature space.
- 75% of the field energy.
- Residual variance (in feature space) is 0.003.
Problem Definition: Given a reduced input $\mathbf{z}$, reconstruct a field $\mathbf{a} = C(\mathbf{z})$.

- Assuming locally linear reduced manifold:
  - Find $L$ observed nearest neighbors $\mathbf{z}^{(s)}$
  - Find the corresponding observed high-dimensional inputs $\mathbf{a}^{(s)}$.

- Assume the following form for the reconstruction:

$$C^L(\mathbf{z}) = \sum_{l=1}^{L} d_i^* \alpha^{(s_l)}.$$ 

where

$$\{d_i^*\}_{i=1}^{L} = \arg \min \sum_{l=1}^{L} d_i \alpha^{(s_l)} \left\| \mathcal{R} \left( \sum_{l=1}^{L} d_i \alpha^{(s_l)} \right) - \mathbf{z} \right\|_2^2$$
Reconstruction

- Reconstruct a test sample:

- Reconstruction from Kernel PCA with $k=30$
**Problem Definition:** Given the reduced observations \( \{z^{(s)}\}_{s=1}^{SA} \), deduce their probability density.

- Expand using a GPC representation:
  \[
  z_i(\xi) = \sum_{m} \beta_{im} H_m(\xi), \quad i = 1, \ldots, K
  \]

- Coefficients may be found via:
  - Maximum likelihood: [Desceler et al. (2006)], [Stefanou et al. (2009)].
  - Rosenblatt transformation: [Rosenblatt (1952)], [Das et al. (2009)].

- Here, we use Rosenblatt assuming **independence**.
SAMPLES FROM THE INPUT MODEL

- Sampling from the 30-dimensional space and reconstructing ...
Probability Model of Reduced Variables

- Construct stochastic input model in a reduced space while preserving the correlation between random variables.

Manifold learning on high-dimensional input data.

Graphical model learning of joint PDF of reduced random variables \( p(\eta_1 \cdots \eta_m) \)

Polynomial chaos representation of reduced random variables
\[
\eta_i = \sum_{j=1}^{p} c_{ij} \psi_j(\xi) \quad \text{where} \quad \xi_i \sim N(0,1)
\]

Low dimensional decomposition of joint PDF via Bayesian network
\[
p(\eta) = \prod_{i=1}^{t} p(\eta_i | \Pi_{\eta_i})
\]
where \( \Pi_{\eta_i} \) indicates the parent nodes that depend on \( \eta_i \)

\[\text{Wan & Zabaras, 2012}\]
Two Phase Flow through Porous Medium

Water and oil, ignore gravity effects and capillary forces, assume porosity is a constant.

\[ \nabla \cdot \mathbf{u} = 0 \]

\[ \mathbf{u} = -K(x, \omega) \nabla p, \forall x \in D, \]

\[ \frac{\partial S(x, t, \omega)}{\partial t} + \mathbf{u} \cdot \nabla S(x, t, \omega) = 0, \forall x \in D, t \in [0, T], \]

\[ p = \bar{p}, \text{ on } \partial D_p, \mathbf{u} \cdot \mathbf{n} = 0, \text{ on } \partial D_u. \]

- Deterministic solver: Mixed FEM on a 45x45 grid.
- Permeability is defined as constant on each element.
- We look at the response at the saturation at 0.2 PVI.
- We don’t need HDMR for 30D input.
Multi-output Gaussian Processes are used here for uncertainty quantification (to be discussed later on).

Comparison of standard deviations of the saturation $S$ using MGP with approx. 800 (a) and 6,500 (b) samples with a MC simulation using $10^6$ samples at 0.2 PVI.
Goal: quantifying property variability of materials due to microstructure uncertainty
- Microstructures of engineering materials are random
- Properties of the material/product are microstructure dependent
- Uncertainty quantification is important (error bars on properties)

Challenge:
- High dimensional stochastic input
- Time consuming stochastic/deterministic simulation

Strategy:
- Model reduction techniques
- Efficient SPDE solver
- Robust physical solver
Uncertainty Sources in Polycrystals

**Macroscale uncertainties**
- Initial and boundary conditions (e.g. Die shape)
- Dual microstructures
- Process parameters: 
  - Temperature
  - Strain rate
  - etc.

**Mesoscale uncertainties**
- Topology
- Two-phase features
- Orientation
- Model Parameters: CRSS, etc.

**Microscale uncertainties**
- Particle size/volume fraction
- Particle shape
- Dislocation configuration
- Parameters: APB energy, etc.

Uncertainty propagation
Given some experimental correlation that the microstructure/property variation satisfies.

Construct several plausible ‘images’ of the microstructure/property.

Each of these ‘images’ consists of, say, $n$ pixels.

Each image is a point in n-dimensional space.

But each and every ‘image’ is related.

That is, all these images lie on a unique curve (manifold) in $\mathbb{R}^n$.

Can a low-dimensional parameterization of this curve be computed?

**Strategy:** based on a variant of the ‘manifold learning’ problem.
Manifold Learning: An Intuitive Approach

Given a set of \( N \) unordered points belonging to a manifold \( \mathcal{M} \) embedded in a high-dimensional space \( \mathbb{R}^n \), find a low-dimensional region \( \mathcal{A} \subset \mathbb{R}^d \) that parameterizes \( \mathcal{M} \), where \( d \ll n \).

Ensure that nearby points on the manifold map to nearby points in the low-dimensional space and faraway points map to faraway points in the low-dimensional space.

PCA  3D data

Linear approach  Non-linear approach: unraveling the curve
Data-Driven Uncertainty Quantification

Microstructure space

START: Extract microstructural features

Low-dimensional space

Model reduction

Feature reconstruction

Low-D points

PCE

Low-D representation reconstruction

Sample from hypercube

Map to hypercube

Unit hypercube

Property variability

Property convex hull

Property distribution

END

Microstructure space

Low-dimensional space

DATABASE

Sample from hypercube

Map to hypercube

Unit hypercube

Model reduction

Feature reconstruction

Low-D points

PCE

Low-D representation reconstruction

Sample from hypercube

Map to hypercube

Unit hypercube

Property distribution

Property convex hull

Microstructure space

START: Extract microstructural features

Low-dimensional space

Model reduction

Feature reconstruction

Low-D points

PCE

Low-D representation reconstruction

Sample from hypercube

Map to hypercube

Unit hypercube

Property distribution

Property convex hull

Microstructure space

START: Extract microstructural features

Low-dimensional space

Model reduction

Feature reconstruction

Low-D points

PCE

Low-D representation reconstruction

Sample from hypercube

Map to hypercube

Unit hypercube

Property distribution

Property convex hull
Problem set up:
1. Deterministic (multiscale) solver
2. Limited data.
4. Multiple outputs

Bayesian Methods (GP, RVM, etc.)

Builds surrogate surface
Informative data collection
Error bars for the statistics

Hierarchical binary trees
Capturing correlations

Observed inputs

Bilionis & Zabaras, 2011a,b, 2012a,b
Collocation vs Bayesian

Here is a take-home message:

- Collocation methods produce a single surrogate for the code

- Bayesian methods produce a probability distribution over the possible surrogates (Bayesian surrogate)

This is what saves the day!
1. A complex solver for the deterministic problem is already given and we must use it.
2. The solver is very expensive to run.
3. We wish to quantify our confidence about the results.

Only a few samples can be observed!!!
- Which observations are the most informative?
- What about 3?
Problem:
• Given a Bayesian surrogate based, which one is the most informative input point to observe?

The one that maximizes the predictive variance \((D. J. C. MacKay, Bayesian interpolation, Neural Computation 4 (3) (1992) 415–447.) biased by the input distribution.

\[ x_{\text{new}, n+1} = \arg \max_{x \in \mathcal{X}} \sigma_F^2(x; \theta^*; D_{i,1,n}) p(x) \]

Initial Data → Update Bayesian Surrogate → Find most informative observation

I. Bilionis and N. Zabaras, 2011a
All Bayesian methods proceed as follows:

- Collect data.
- Construct a **probability measure over the possible surrogates**.
- Sample a surrogate and interrogate it (analytically or via MC) to get a sample of the statistics.
- Repeat the previous step to get confidence intervals for the statistics.

**Epistemic uncertainty introduced from an inaccurate surrogate.**
Non-stationary Response

Non-stationarity is emulated by using trees of stationary models (GP’s, RVM’s) with each leaf (stochastic element) covering a different region of the stochastic domain.

The tree construction is biased by:
- The variability of the function (local bumps, discontinuities, etc.).
- The underlying input probability density (unimportant regions are not refined).

I. Bilionis and N. Zabaras, 2011a
Assume that $f$ is a $q$-dimensional Gaussian Process

$$f(\cdot) \mid B, \Sigma, r \sim \mathcal{N}_q (m(\cdot; B), c(\cdot, \cdot; r) \Sigma)$$

with mean:

$$m(x; B) = B^T h(x)$$

and covariance:

$$\text{Cov} [f(x_1, x_2) \mid B, \Sigma, r] = c(x_1, x_2) \Sigma.$$ 

over all variables (input, space, time).

$$x = (\xi, x_s, t)$$

Specifies our prior belief about the response surface of the code.
The Covariance Function

- Specifies the assumed regularity of the response.

- Here we use the **Squared Exponential Covariance** function:

\[
c(x_1, x_2; r) = \exp \left\{ -\frac{1}{2} \sum_{s=1}^{k} \frac{(x_{1s} - x_{2s})^2}{r_s^2} \right\}
\]

- The parameters can be interpreted as the length scale of each dimension.

- Essentially we are a priori assuming that the response is infinitely differentiable.
The Likelihood Function

- Observations:

\[ X = (x_1, \ldots, x_n)^T \quad Y = (y_1, \ldots, y_n)^T \]

- The likelihood function is:

\[ Y|B, \Sigma, r \sim \mathcal{N}_{n\times q}(HB, \Sigma, A) \]

where H is the design matrix:

\[ H^T = (h(x_1) \ldots h(x_n))^T \]

and A the covariance matrix:

\[ A_{ij} = c(x_i, x_j; r) \]
By using Bayes Theorem, the **predictive distribution** at a new point \( x^* \) is:

\[
y^*|Y, B, \Sigma, \theta \sim \mathcal{N}_q \left( m^* + (Y - HB)^T A^{-1} a^*, \Sigma (a^* - (a^*)^T A^{-1} a^*) \right)
\]

where

\[
m^* := m(x^*; b)
\]

\[
a^* := (c(x^*, x_1; \theta) \ldots c(x^*, x_n; \theta))^T
\]

and

\[
a^* = c(x^*, x^*; \theta).
\]
The Full Model

To complete the model, we need to assign a prior on the hyper-parameters:

\[ \pi(B, \Sigma, \theta) = \pi(B)\pi(\Sigma)\pi(r) \]

We choose **non-informative** improper priors for the weights and the correlation matrix:

\[ p(B) \propto 1 \]

\[ \pi(\Sigma) \propto |\Sigma|^{-\frac{q+1}{2}} \]

and a log-logistic prior for the length scales:

\[ p(r_i) \propto (1 + r_i^2)^{-1} \]
The posterior of the hyper-parameters is:

\[ p(B, \Sigma, r|Y) \propto p(B)p(\Sigma)p(r)p(Y|B, \Sigma, r) \]

It can be sampled via a Gibb’s procedure:

- Initialize \( B_0, \Sigma_0, r_0 \)
- Sample
  \[ B_{i+1} \sim p(B|Y, \Sigma_i, r_i). \]
- Sample
  \[ \Sigma_{i+1} \sim p(\Sigma|Y, B_{i+1}, r_i). \]
- Sample
  \[ r_{i+1} \sim p(r|Y, B_{i+1}, \Sigma_{i+1}). \]
The Conditionals Posteriors

- **B** follows a matrix-normal:
  \[
  p(B|Y, \Sigma, r) = \mathcal{N}_{m \times q}(B; B_0, \Sigma, H^T A^{-1} H).
  \]

- **\Sigma** follows and inverse-Wishart distribution:
  \[
  p(\Sigma|Y, B, r) = W^{-1}(\Sigma; Y^T G Y, n - m).
  \]

- The length-scales follows:
  \[
  p(r|Y, B, \Sigma) \propto p(r|A)^{-q/2} \exp \left\{ -\frac{1}{2} \text{tr} \left[ \Sigma^{-1} (Y - HB)^T A^{-1} (Y - HB) \right] \right\}.
  \]
  A Metropolis-Hastings procedure is used for this.
Take Home Message

Each sample of the hyper parameters gives a sample surrogate model for the code identified as the mean of the predictive distribution.
We consider covariance functions of the form:

\[ c(x_1, x_2; r) := c_{\xi}(\xi_1, \xi_2; r_\xi)c_s(x_{s,1}, x_{s,2}; r_s)c_t(t_1, t_2; r_t) \]

Under the assumption that the spatial and time inputs are fixed, then:

\[ A = A_{\xi} \otimes A_s \otimes A_t \]

A great deal of this research (not shown here) is concerned with the problem of carrying out inference without ever forming this huge matrix! It can be done!
Given a leaf of the tree with GP fitted on it, we must decide:

- Refine the leaf or not?
- What is the most important dimension?

Technical details can be found here:

Refine the element $\Xi_i$, if:

$$\int_0^T \int_{\Omega} \int_{\Xi_i} \text{tr} \left[ \Sigma(a^* - a(\xi^*, x^*, t^*))^T A^{-1} a(\xi^*, x^*, t^*)) \right] p(\xi^*) dx^* dt^* \frac{T|\Omega|q}{P(\Xi_i) > \delta},$$

The covariance of the predictive distribution.

Remark: The integral can be expressed in terms of error functions!
What is the most important dimension?

Split the element perpendicular to:

\[ j^* = \arg \max_j \frac{p^i_j}{r^i_j}. \]

The splitting point is the median of the marginal conditional.
Consider the dynamical system:

\[
\frac{dy_1}{dt} = y_1 y_3,
\]
\[
\frac{dy_2}{dt} = -y_2 y_3,
\]
\[
\frac{dy_3}{dt} = -y_1^2 + y_2^2
\]

We will consider two different stochastic variants:

- **KO-1:**
  \[
y_1(0) = 1, \quad y_2(0) = \frac{\xi}{10}, \quad y_3(0) = 0, \quad \xi \sim U[-1, 1]
\]

- **KO-2**
  \[
y_1(0) = 1, \quad y_2(0) = \frac{\xi_1}{10}, \quad y_3(0) = \xi_2, \quad \xi_i \sim U[-1, 1]
\]
For each input point, we observe the response at \( n = 20 \) time instants:

\[
Y_\xi = \begin{pmatrix}
y_1(\xi, t_1) & y_2(\xi, t_1) & \cdots & y_q(\xi, t_1) \\
y_1(\xi, t_2) & y_2(\xi, t_2) & \cdots & y_q(\xi, t_2) \\
\vdots & \vdots & \ddots & \vdots \\
y_1(\xi, t_n) & y_2(\xi, t_n) & \cdots & y_q(\xi, t_n)
\end{pmatrix}
\]

Gathering everything together build the observed-data matrix:

\[
Y = \begin{pmatrix}
Y_{\xi_1} \\
Y_{\xi_2} \\
\vdots \\
Y_{\xi_n}\end{pmatrix}
\]
KO-2 has a line discontinuity:

The figure shows $y_3(t=10)$ as a function of the initial conditions.

The regions shown are the final nodes of the tree for a tolerance of $1e-7$.

The statistics can be recovered accurately with much smaller tolerance. The discontinuity does not contribute to the integrals!
KO-2 Observations for Tolerance 1e-3 (80)
KO-2 Observations for Tolerance $1e^{-4}$ (260)
KO-2 Observations for Tolerance 1e-5 (660)
KO-2 Observations for Tolerance $1e^{-6}$ (2310)
KO-2 Mean

1e-1, 20 observations

1e-2, 40 observations
KO-2 Variance

1e-1, 20 observations

1e-2, 40 observations
KO-2 Variance

1e-3, 80 observations

1e-4, 260 observations
True PDF

1e-1, 20 observations
True PDF

![Graph showing probability density for different conditions.](image)

1e-2, 40 observations

![Graph showing probability density with 40 observations.](image)
KO-2 PDF

True PDF

1e-3, 80 observations
True PDF

1e-4, 260 observations
• Moving to hierarchical trees and coupling with HDMR. Bayesian model selection for the choice of the HDMR terms

• Exploiting spatial locality of the stochastic input in order to break the UQ problem into smaller manageable ones (natural convection, porous flow etc.)
“Designing the Next Generation Exascale Statistical Tools.”

- **Objectives:**
  - High-dimensional, Data-driven exploration of PDE-based engineering problems
  - Fast implementation and testing of novel statistical algorithms

- **Key ideas:**
  - Modularity:
  - Efficiency/Scalability:

- **Features:**
  - Optimized templated versions of core statistical models
  - Data-based input modeling
  - Uncertainty Propagation
  - Inverse problems.

**Supported by:** AFOSR-Mathematics & DOE Office of Science, ASCR
Data-Driven Uncertainty Quantification

Microstructure space

START: Extract microstructural features

Low-dimensional space

Model reduction

Feature reconstruction

Low-D points

PCE

Low-D representation reconstruction

Unit hypercube

Property variability

Property convex hull

Property distribution

Obtain Properties

Solve SPEDs

Database

Reconstruct microstructures

Low-D representation

Sample from hypercube

Unit hypercube

END

Property convex hull

Low-D space

Map to hypercube

Unit hypercube

Low-D points

PCE

Low-D representation reconstruction

Sample from hypercube

Unit hypercube

Low-D space

Map to hypercube

Unit hypercube

START: Extract microstructural features
Elastic Properties - Tensile Strength

An example take orientation distribution function (ODF) in the Rodrigues space as the stochastic input.

**Approach**
- Karhunen-Loeve expansion
- Maximum Entropy Principle

\[
A_0(s, \omega) = \bar{A}_0(s) + \sum_{i=1}^{\infty} \sqrt{\lambda_i} f_i(s, t) Y_i(\omega)
\]

Energy captured

![KLE](image1)

MaxEnt

![MaxEnt](image2)

Property convex hull

![Property convex hull](image3)

Continuum representation of texture in Rodrigues space

Variation of stress-strain response

![Variation of stress-strain response](image4)
Problem Definition

- **Problem definition**

  **Given:**
  - Grain size snapshots constrained by the mean size.
  - Texture snapshots from random process.
  - Volume fraction of $\gamma'$ precipitates.

  **Goal:**
  - The variability of fatigue resistance due to microstructure uncertainties.
  - Find the feature(s) dominate the variability of fatigue resistance

- **Methodologies**

  - PCA and KPCA model reduction to reduce the complexity of stochastic input.
  - PCE to map reduced coordinates to a known distribution.
  - Adaptive sparse grid and Monte Carlo collocation to solve stochastic partial differential equations

---

Strain-based FIPs are employed to measure the fatigue properties of nickel-based superalloy microstructures.

The FIPs are evaluated throughout the third loop of cyclic deformation.

Cumulative plastic strain per cycle

\[
P_{\text{cyc}} = \int_{\text{cyc}} \frac{2}{3} \dot{\varepsilon} dt = \int_{\text{cyc}} \frac{2}{3} \mathbf{D}^p : \mathbf{D}^p dt
\]

\[
\Delta \gamma_{\max}^p = \max_{\theta} \left( \max_{\text{cyc}} \left( \gamma_{\theta}^p \right) - \min_{\text{cyc}} \left( \gamma_{\theta}^p \right) \right)
\]

Cumulative net plastic shear strain

\[
P_r = \max \left( \int_{\text{cyc}} \dot{\gamma}_{ij} n_i t_j dt \right)
\]

The Fatemi-Socie parameter

\[
P_{FS} = \frac{\Delta \gamma_{\max}^p}{2} \left[ 1 + k^* \frac{\sigma_{n_{\max}}}{\sigma_y} \right]
\]

Maximum range of cyclic plastic shear strain

\[
P_{mps} = \frac{\Delta \gamma_{\max}^p}{2}
\]

Wen and Zabaras, 2012
Variability of Fatigue Properties of Ni-based Superalloys

Physical Model

Homogenized model and explicit two-phase structure. Solver: Taylor model FEM simulations.

Explicit structure
Homogenized model

Convex Hulls of FIPs

FIPs Distributions

(a) Pcyc
(b) Pr
(c) PFS
(d) Pmps

FEM Simulation

Realistic microstructure
Mean Contour plot of $\Delta \gamma_P^{\max}$
Std Contour plot of $\Delta \gamma_P^{\max}$

Sampled microstructure
Mean Contour plot of $\Delta \gamma_P^{\max}$
Std Contour plot of $\Delta \gamma_P^{\max}$
Multiscale Model Reduction

- **Motivation:** Microstructure features are location specific. Multiscale model reduction is needed to exploit the correlation between spatial points to resolve the “curse of dimensionality”.

- **Solution:** Bi-orthogonal decomposition.
  - Start from realizations of the microstructure random field \( (A) \) varying in both micro- (s) and macro-scales (x):
    \[
    A(x,s,\omega) : X \times S \times \Omega \to \mathbb{R}
    \]
  - Project A to a set of bi-orthogonal bases
    \[
    A(x,s,\omega) = \tilde{A}(x,s) + \sum_{i=1}^{\infty} \sqrt{\rho_i} \psi_i(s) \Phi_i(x,\omega)
    \]
  - By minimizing the distance between the Karhunen-Loeve expansion and the random field, the microscale basis can be computed by
    \[
    \psi_i(s) = \frac{1}{\sqrt{\rho_i}} \left\{ \tilde{A}, \Phi_i \right\}
    \]
    \[
    \{\Phi_i, \Phi_j\} := \int_X \Phi_i(x,\omega) \cdot \Phi_j(x,\omega) \, dx
    \]
    \[
    (\psi_i, \psi_j) := \int_S \psi_i(s) \psi_j(s) \, ds
    \]
  - The macroscale basis is obtained through orthogonality condition
    \[
    \Phi_i(x,\omega) = \frac{1}{\sqrt{\rho_i}} \int_S \tilde{A}(x,s,\omega) \psi_i(s) \, ds
    \]
  - The eigenvalue problem in K-L expansion can be defined as
    \[
    \rho_i \psi_i(s) = \int_S C(s,s') \psi_i(s') \, ds'
    \]

Wen & Zabaras, 2012
Training: establish the mapping between microstructure space and reduced-order space.

- Determine the dimensionality of reduced space (d):

\[ P_{\text{Energy}}(d) = \frac{\sum_{i=1}^{d} \langle E_i \rangle}{\sum_{j=1}^{M} \langle E_j \rangle} \]

Energy:
\[ E_i(\omega) := \int_{x} \rho_i(x, \omega) \Phi_i(x, \omega) dx \]

- Decompose macromodes through PCE:
\[ \Phi_i(x, \omega) = \sum_{j} \phi_{ij}(x) \gamma_j(\omega) \]

Uniform-Legendre PCE
\[ \phi_{ij}(x) = \frac{2j+1}{2} \int_{-1}^{1} \Phi_i(x, \xi) \gamma_j(\xi) d\xi, \]
\[ i = 1, \ldots, r, \ j = 0, \ldots, p \]

Testing: sample in the low-D space and reconstruct real microstructures
\[ e = \frac{1}{\dim} \sum_{i=1}^{\dim} \left| \frac{R_{\text{Original}}^i - R_{\text{Restored}}^i}{R_{\text{Original}}^i} \right| \]
Stochastic Multiscale Forging

Mean effective strength

Standard deviation effective strength

Initial samples
MC with 16 dim
MC with 30 dim
ASGC with 3 dim
Full Probabilistic Response of Properties in the Disk

(a) 3D representation of the disk with color-coded regions indicating different properties.
(b) PDF plot for Equiv Strain ranging from 0.6 to 0.7.
(c) PDF plot for Equiv Stress ranging from 700 to 1100.
(d) PDF plot for Equiv Strength ranging from 265 to 268.
Goal: Accurately and efficiently investigate effective and local mechanical properties/response of polycrystalline materials based on realistic microstructure images.

Solution strategy: Green’s function method in combination with fast Fourier transform for solving the equilibrium equations.

Merit: no complex FEM meshing, no “inversion” of big matrices, can consider both intergranular and intragranular interactions, can use experimental or simulation based microstructure images as input.

Any constitutive model can be considered: elasto-viscoplasticity, single phase or multi-phase alloys, fatigue properties, etc.

Wen & Zabaras, 2012
Consider the fluid flow problem with multiple scales

\[
\mathbf{u}(\mathbf{x}, \omega) = -a(\mathbf{x}, \omega) \nabla p(\mathbf{x}, \omega) \quad \forall \mathbf{x} \in D
\]

\[
\nabla \cdot \mathbf{u}(\mathbf{x}, \omega) = f(\mathbf{x}) \quad \text{with} \quad p = \bar{p} \quad \text{on} \quad \partial D_p, \quad \mathbf{u} \cdot \mathbf{n} = \bar{u} \quad \text{on} \quad \partial D_u \quad \partial D = \partial D_p \cup \partial D_u
\]

- fine grid \( T_h = \bigcup_{i=1}^{N_h} e_i \) : material properties \( a(\mathbf{x}, \omega) = \{ a_i(\omega) \}_{i=1}^{N_h} \)
- coarse grid \( T_c = \bigcup_{k=1}^{N_c} E_k \) : physical responses \( y(\mathbf{x}, \omega) := (\mathbf{u}(\mathbf{x}, \omega), p(\mathbf{x}, \omega))^T \)

**Objective**

Construct a regression model for the solution to Multiscale SPDEs

\[
y_K(\mathbf{x}, \omega) = \sum_i W_i(\omega) \psi_i^K(a_K) \quad \forall \mathbf{x} \in E_K
\]

where \( \psi_i(a(x, \omega)) \) are vector basis functions and \( W_i(\omega) \) are scalar weights
**Problem definition**
Given stochastic input $a_K$ on the $K$-th coarse element, predict multiscale basis function values $\psi^K_i$ at any fixed locations $x \in E_K$ with a regression model.

**Multiscale basis functions**
$$\psi^K_i = -a_K(x) \nabla \varphi^K_i$$

**Local regression models**
$$\psi^K_i \approx c^K_m \delta^K_m(a_K) \text{ with kernel functions } g^K_m(a_K) = \exp\left(-\frac{\|a_K - a^K_m\|^2}{L^2}\right)$$

Given a set of $N$ training data $\left\{a_k^{(n)}, \psi^K_i^{(n)}\right\}_{n=1}^N$, we get the constant coefficients $c^K_m$ by minimizing the squared error function
$$\mathbb{E} = \frac{1}{2} \sum_{n=1}^N \left(\psi^K_i(a_k^{(n)}) - \psi^K_i^{(n)}\right)^2$$

**Steps**
1. **Step 1**: Local regression model for multiscale basis functions
2. **Step 2**: Multiscale basis functions
3. **Step 3**: Local regression models
4. **Step 4**: Approximation of $\psi^K_i$
Problem definition
Given material properties \( a(x, \omega) = \{a_i(\omega)\}_{i=1}^{N_b} \) defined on the finest scale, predict the weights, \( W_i(\omega) \), of basis functions in the global regression model for physical responses on the coarse scale.

Probabilistic model of weights
\[
p(w | \xi) \propto \prod_i \phi_i(w_i | \xi) \prod_{i,j} \phi_{i,j}(w_i, w_j | \xi)
\]

edge potential \( \phi_i(w_i) := \exp\left( b_i w_i \frac{1}{2} \left( \sum_{r=1}^{4} \sum_{K} d_i a_{r,K}(\xi_{r,K})^2 w_i^2 \right) \psi_{K,r} = \psi_i^K(x_i) \in \xi_K \) \]

self potential \( \phi_{i,j}(w_i, w_j) = \exp\left( -w_i w_j \sum_{r=1}^{4} \sum_{K} d_i a_{r,K}(\xi_{r,K})^2 \psi_{K,r} = a_K(x_r) \right) \)

Gaussian belief propagation (GaBP) prediction of marginal distribution of unobserved variables --- weights
\[
\phi_i(w_i) \propto N(\mu_i, P_{w_i}^{-1})
\]
The joint distribution:

\[ p(X, Y) = \frac{1}{Z} \prod_{(i,j) \in \mathcal{E}} \psi_{i,j}(y_i, y_j) \prod_{i \in \mathcal{V}} \varphi_i(x_i, y_i) \]

Nonparametric Representation:

\[ \psi_{i,j}(y_i, y_j) = \sum_{m=1}^{M} \omega^{(m)} \mathcal{N}(|y_i - y_j|; \mu^{(m)}, \Sigma^{(m)}) \]
\[ \varphi_i(x_i, y_i) = \sum_{m=1}^{M} \omega^{(m)} \mathcal{N}([x_i, y_i]; \mu_{xy}^{(m)}, \Lambda^{(m)}) \]

**Gaussian Model:** In the Gaussian model, all the potential functions are modeled by Gaussian functions \( \mathcal{N}(x; \mu, \Sigma) \), the unknown parameters are \( \mu \) and \( \Sigma \). Learned by MLE.

**Nonparametric Model:** In the nonparametric model, all the potential functions are modeled as Gaussian mixtures, \( \sum_{k=1}^{K} \omega_k \mathcal{N}(x_n; \mu_k, \Sigma_k) \), the unknown parameters are \( \omega_k, \mu_k \) and \( \Sigma_k \). Learned by the EM method.

---

**Graphical Model Learning**

Chen and Zabaras, 2011
Inference Problem

General Belief Propagation Algorithm:

The marginal for $p(y_{ij})$ can be calculated by

$$p(y_{ij}) \propto \varphi_{i,j}(x_{i,j}, y_{i,j}) \prod_{y_{k,l} \in \Gamma(y_{i,j})} m_{y_{k,l}}(y_{i,j})$$

Nonparametric Belief Propagation Algorithm (Main steps):

- Determine the marginal influence $\zeta(y_{k,l}) = \int \psi(y_{k,l}, y_{i,j}) dy_{i,j}$
- Sampling from $y_{k,l}^{(t)} \sim \zeta(y_{k,l}) \psi_{k,l}(x_{k,l}, f_{k,l}) \prod_{(p,q) \in \Gamma(y_{k,l}) \setminus y_{i,j}} m_{p,q}(y_{k,l})$
- Obtain the belief update by Sampling from $y_{i,j}^{(t)} \sim \psi(y_{i,j}, y_{k,l} = y_{k,l}^{(t)})$

Loopy Belief Propagation Algorithm:
**Main process:**

- **Start with a set of training data (initial microstructures, final properties, textures)**
- **Graph learning**
  - PCA, MLE, EM
- **Create a Gaussian model or a Nonparametric model**
- **Inference**
  - Nonparametric belief propagation, Gaussian mixture reduction
- **Obtain the prediction for the output given a new input**
Final Property Prediction

Stress field prediction given a new initial texture:

Each component of the nonparametric model:

- **mean**
- **std**
- **weight**
Final Texture Prediction

Final Texture Prediction Given a new initial Texture:

Orientation density function prediction at a random point:

Pole figure for the Initial texture at a random point

Pole figure for the final texture at that point

Pole figure for the reconstructed final texture by Gaussian model

Pole figure for the reconstructed final texture by the Nonparametric model
Inverse Problem

(A) Find Initial Texture Field for Desired Forged Disk Properties (here equivalent stress)

Desired

Gaussian Model

Nonparametric Model

(B) Find Initial Texture Field for Desired Texture on the Forged Disk

True Initial texture

True Final texture

Gaussian Model

Nonparametric Model
Conclusions

- Information theoretic approaches to predictive science
- Need for managing complexity in stochastic multiscale models
- The Bayesian approach offers:
  - Probability measure over possible surrogates (finite number of samples).
  - Experimental design.
  - Non-stationary responses can be modeled by trees of surrogates.
  - Samples of the surrogate distribution can be used for: Uncertainty Quantification tasks, Sensitivity analysis, Model calibration.
  - Future research: Corregionalization models, fully Bayesian trees.
  - Dealing with high-dimensional inputs.
- Use of probabilistic graphical models allow stochastic multiscale problems to be addressed as inference problems in graphs.