Predictive Modeling and Uncertainty Quantification in Polycrystalline Materials

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, ...

Department of Aerospace Engineering Sciences, Mechanics and Materials (MaM) Colloquium, May 4th, 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
Problem Definition

Observed input
\[ A = \{\alpha^{(s)}\}_{s=1}^{S_A} \]

Data collection

Deterministic Solver

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

Surrogate Model

Reduced input space

Reduced density estimation

Reconstruction

Statistics

PDFs

Error bars

Output space
**Data-Driven Stochastic Input Models**

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

\[ \mathcal{R}(\alpha) = z \]

\[ \alpha = C(z) \]

**Density estimation:**
Polynomial Chaos Representation

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

**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
Design in the Presence of Uncertainties

Design the boundary heat flux variability that results in desired PDFs of the temperature at sensor locations.

Limited information is provided for the underlying heterogeneous medium.

Lesser input uncertainty results in reduced variability in the design 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

Observed inputs
The left vertical wall (cold) is a 1D 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.

(c) $K=8$, MGP ($N = 20, \delta = 10^{-3}$): std. of $T$

(d) $K=8$, MC: std. of $T$
Natural Convection Multioutput GPs

(a) Mean $T$ prediction  (b) Absolute error  (c) 2 std.’s of $T$  (d) Input point

Natural Convection ($K = 4$, $\delta = 10^{-5}$): 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.

- We treat the permeability as a log random field obtained from the K-L expansion

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

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 $N = 500$

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

$\sigma^2 = 0.01$

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

$\sigma^2 = 0.25$

$\sigma^2 = 2.0$

Standard deviation of the $v$ velocity-component along the cross section $y = 0.5$ for different $\sigma^2$
PDF of the velocity-component at point $(0, 0.5)$ for different $\sigma^2$ with $N = 500$

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.

$L = 0.25, N = 500$
Convergence of the normalized errors of the standard deviation of the velocity-component for different $\sigma^2$.

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

\(-\nabla \cdot (a_K(\omega, \cdot)\nabla u(\omega, \cdot)) = f(\cdot), \text{ in } D = [0, 1]^2, \quad f(x, y) = 100 \cos(x) \sin(y),\)

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

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

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

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

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

\[ \omega_k = U\left(\left[\sqrt{3}, \sqrt{3}\right]\right) \]

\[ \text{Cov}[\log(a_K - 0.5)]((x_1, y_1), (x_2, y_2)) = \exp \left(-\frac{(x_1 - x_2)^2}{L_c^2}\right) \]
Elliptic Benchmark Problem

The $L_2$ norm of the error in variance of the elliptic problem with $K = 40$ inputs as a function of the observed samples for MGP and ASGC.

- Sparse grids work using points from an a-priori defined grid.
- Bayesian GPs use active learning to select points and defining the tree.

The correlation length is set to $L_c = 0.6$.
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 & Zabaras (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:**
  \[ k(\alpha^{(s)}, \alpha^{(s')}) = \exp \left( - \frac{\| \alpha^{(s)} - \alpha^{(s')} \|^2}{2\sigma^2} \right) \]

- **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.
- 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 $z$, reconstruct a field $a = C(z)$.

- Assuming locally linear reduced manifold:
  - Find $L$ observed nearest neighbors $z^{(s_i)}$.
  - Find the corresponding observed high-dimensional inputs $\alpha^{(s_i)}$.
- Assume the following form for the reconstruction:

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

where

$$\{d_i^*\}_{i=1}^{L} = \arg \min_{\sum_{i=1}^{L} d_i = 1} \| R \left( \sum_{i=1}^{L} d_i \alpha^{(s_i)} \right) - z \|^2_2$$
Reconstruct a test sample:

- Reconstruction from Kernel PCA with k=30
- Reconstructed with Linear PCA: k = 30

10 nearest neighbor are used in reconstruction

Reconstruction from Kernel PCA with k=30
**Density Estimation**

**Problem Definition:** Given the reduced observations \( \{ z^{(s)} \}_{s=1}^{S_A} \), deduce their probability density.

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

- **Coefficients may be found via:**
  - Maximum likelihood: [Descelier et al. (2006)], [Stefanou et al. (2009)], etc.
  - 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 ...
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 \( p(\eta_1 \cdots \eta_m) \)

Initial fully connected graphical model

Graphical structure learning via conditional independence tests from data

Probabilistic model of joint PDF using kernel density estimation

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 \( \eta_i \)

\[\eta_i = \sum_{j=1}^{p} c_{ij} \psi_j(\xi) \text{ where } \xi_i \sim N(0,1)\]

Polynomial chaos representation of reduced random variables

Wan & Zabaras, 2012
Two phase flow through porous media

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

Property variability

- Property convex hull

Low-dimensional space

- Model reduction
- Feature reconstruction
- Low-D representation reconstruction
- Sample from hypercube
- Map to hypercube
- Low-D space

START: Extract microstructural features

- Property convex hull
- Property distribution
- Database

Obtain Properties

- Solve SPEDs

END

Reconstruct microstructures

Low-D points

PCE

Unit hypercube
Data-Driven Uncertainty Quantification

Microstructure space

START: Extract microstructural features

Low-dimensional space

Model reduction

Low-D points

Feature reconstruction

PCE

Low-D space

Sample from hypercube

Map to hypercube

Reconstruct microstructures

Database

Property variability

Property convex hull

Property distribution

END

END

START:

Extract microstructural features

Low-D points

Feature reconstruction

PCE

Low-D space

Sample from hypercube

Map to hypercube

Reconstruct microstructures

Database
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

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

**FCC copper**

Continuum representation of texture in Rodrigues space

Effective strain

Effective stress (MPa)

Variation of stress-strain response

Energy captured

Number of Eigenvalues

Property convex hull

\( \times 10^6 \) MPa

0.4 0.5 0.6

0.4 0.5 0.6

1 1.2 1.3

0.4 0.5 0.6

Shear Modulus

Young Modulus

Bulk Modulus

Approach

- Karhunen-Loeve expansion
- Maximum Entropy Principle

\[
A_0(s, \omega) = \bar{A}_0(s) + \sum_{i=1}^{\infty} \sqrt{\lambda_i} f_i(s, t) Y_i(\omega)
\]
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) Pcy
(b) Pr
(c) PFS
(d) Pmps

FEM Simulation

Realistic microstructure
Mean Contour plot of \( \Delta \gamma_P \)
Std Contour plot of \( \Delta \gamma_P \)
Sampled microstructure
Mean Contour plot of \( \Delta \gamma_P \)
Std Contour plot of \( \Delta \gamma_P \)
**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
  
  $$A(x,s,\omega) = \hat{A}(x,s) + \sum_{i=1}^{\infty} \sqrt{\rho_i} \psi_i(s) \Phi_i(x,\omega)$$  
  
  $$\hat{A}(x,s) := \int_\Omega A(x,s,\omega) p(\omega) d\omega$$

  $$\approx \hat{A}(x,s) + \sum_{i=1}^{r} \sqrt{\rho_i} \psi_i(s) \Phi_i(x,\omega)$$

- By minimizing the distance between the Karhunen-Loeve expansion and the random field, the microscale basis can be computed by
  
  $$\psi_i(s) = \frac{1}{\sqrt{\rho_i}} \left\{ \hat{A}, \Phi_i \right\} \quad \left\{ \Phi_i, \Phi_j \right\} := \int_X \left( \Phi_i(x,\omega) \cdot \Phi_j(x,\omega) \right) dx$$
  
  $$\left( \psi_i, \psi_j \right) := \int_S \psi_i(s) \psi_j(s) ds$$

- The macroscale basis is obtained through orthogonality condition
  
  $$\Phi_i(x,\omega) = \frac{1}{\sqrt{\rho_i}} \int_S \hat{A}(x,s,\omega) \psi_i(s) ds$$

- The eigenvalue problem in K-L expansion can be defined as
  
  $$\rho_i \psi_i(s) = \int_S C(s,s') \psi_i(s') ds'$$

**Orthogonality:**

$$\left( \psi_i, \psi_j \right) = \delta_{ij}$$

$$\left\{ \Phi_i, \Phi_j \right\} = \delta_{ij}$$

---

Wen & Zabaras, 2012
Multiscale Model Reduction

- Training: establish the mapping between microstructure space and reduced-order space.
  - Determine the dimensionality of reduced space (d):
    \[ P_{\text{Energy}}(d) = \frac{\sum_{i=1}^{d} \left\langle E_i \right\rangle}{\sum_{j=1}^{M} \left\langle E_j \right\rangle} \]
    
    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
  \[ \phi_{ij}(x) = \frac{2j+1}{2} \int_{-1}^{1} \Phi_i(x,\zeta_j) \gamma_j(\zeta_i) d\zeta_i, \]
  \[ i = 1,\ldots,r, \quad j = 0,\ldots,p \]

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

Initial samples  MC with 16 dim  MC with 30 dim  ASGC with 3 dim

Standard deviation effective strength
Full Probabilistic Response of Properties in the Disk

(a)  
(b)  
(c)  
(d)
Bayesian Framework for UQ

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

Bayesian Methods (GP, RVM, etc.)

Error bars for the statistics

Hierarchical binary trees

Capturing correlations

Informative data collection

Builds surrogate surface

Material Process Design and Control Laboratory

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)
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.

\[ x^{\text{new},n+1} = \arg \max_{x \in X_i^{1,1}} \sigma_f^2(x; \theta^*; D_i^{1,1,n}) p(x) \]

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

\[
f(\cdot) \mid B, \Sigma, r \sim \mathcal{N}_q \left( m(\cdot; B), c(\cdot, \cdot; r)\Sigma \right)
\]

with mean:

\[
E[f(x) \mid B, \Sigma, r] = m(x; B)
\]

and covariance:

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

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:

\[
\begin{align*}
    c(x_1, x_2; r) & \equiv c_\xi(\xi_1, \xi_2; r_\xi)c_s(x_{s,1}, x_{s,2}; r_s)c_t(t_1, t_2; r_t) \\
    c_\xi(\xi_1, \xi_2; r_\xi) & = \exp \left\{ -\frac{1}{2} \sum_{s=1}^{k} \frac{(x_{1s} - x_{2s})^2}{r_{\xi,s}^2} \right\}
\end{align*}
\]

and analogously for \( c_s(\cdot, \cdot; r_s) \) and \( c_t(\cdot, \cdot; r_t) \).

- 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:
  \[ X_s = (x_{s,1}, \ldots, x_{s,n_s})^T \in \mathbb{R}^{n_s \times k_s} \]
  \[ X_t = (t_1, \ldots, t_{n_t}) \in \mathbb{R}^{n_t \times 1} \]
  \[ X_\xi = (\xi_1, \ldots, \xi_{n_\xi})^T \in \mathbb{R}^{n_\xi \times k_\xi} \]
  \[ Y_{\xi,i} = (y_{\xi,i,1}, \ldots y_{\xi,i,n_t})^T \in \mathbb{R}^{n_t \times q} \]
  \[ Y_{\xi,i,j} = (y_{\xi,i,j,1}, \ldots y_{\xi,i,j,q})^T \in \mathbb{R}^{q \times 1} \]
  \[ Y = (Y_{\xi_1}^T \ldots Y_{\xi_{n_\xi}}^T)^T \in \mathbb{R}^{n \times q} \]

- The likelihood function is:
  \[ Y|B, \Sigma, r \sim \mathcal{N}_{n \times q}(HB, \Sigma, A) \]

where \( H \) is the design matrix:

\[ H = (h(x_1)^T \ldots h(x_n)^T)^T \in \mathbb{R}^{n \times m} \]

and \( A \) the covariance matrix:

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

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

where

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

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

and

$$a^* = c(x^*, x^*; \theta)$$
To complete the model, we need to assign a prior on the hyperparameters:

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

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

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

and a log-logistic prior for the length scales:

\[ \pi(r) = \prod_{s=1}^{k} \pi(r_s), \quad \pi(r_s) \propto (1 + r_s^2)^{-1} \]
The posterior of the hyper-parameters is:

$$p(B, \Sigma, r|Y) \propto \pi(B)\pi(\Sigma)\pi(r)p(Y|B, \Sigma, r)$$

It can be sampled via a Gibb’s procedure:

- Initialize $B_0$, $\Sigma_0$ and $r_0$
- 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:
  \[ \rho(B|Y, \Sigma, r) = \mathcal{N}_{m \times q}(B; B_0, \Sigma, H^T A^{-1} H) \]

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

- The length-scales follows:
  \[ \rho(r|Y, B, \Sigma) \propto \pi(r)|A|^{-q/2} \exp \left\{ -\frac{1}{2} \text{tr} \left[ \Sigma^{-1} (Y - HB)^T A^{-1} (Y - HB) \right] \right\} \]

A Metropolis-Hastings procedure is used for this.
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:

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

\[ H = H_\xi \otimes H_s \otimes H_t \]

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:

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

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:

\[ j^* = \arg \max_j \frac{p_j^i}{r_j^i} \]

The splitting point is the median of the marginal conditional.

Probability that the j-th dimension of a random input point falls in element i.
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:**
  \[ y_1(0) = 1, \ y_2(0) = \xi / 10, \ y_3(0) = 0, \ \xi \sim U[-1, 1] \]

- **KO-2**
  \[ y_1(0) = 1, \ y_2(0) = \xi_1 / 10, \ y_3(0) = \xi_2, \xi_i \sim U[-1, 1] \]
Observations

For each input point, we observe the response at n=20 time instants:

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

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 $1\times10^{-5}$ (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
KO-2 PDF

True PDF

1e-1, 20 observations
True PDF

1e-2, 40 observations
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

*I. Bilionis, N. Zabaras, in preparation*
"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
Working in Stochastic Modeling does NOT imply that we are done with physical modeling...

A typical stochastic simulation to compute property variability at a material point may require thousands of calls to a deterministic 3D simulator!

How do you even mesh these microstructures?
Goal: Accurately and efficiently investigate effective and local mechanical properties/response of polycrystalline materials based on realistic microstructure images

Solution strategy: Greens function method in combination with fast Fourier transform for solving 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.
Equilibrium equation: \( \sigma_{ij,k} = \mathcal{L}_{ijkl} \dot{e}_{kl,j} + \varphi_{ij,k} = \mathcal{L}_{ijkl} v_{k,j} + \varphi_{ij,j} = 0 \quad \varphi(x) = \sigma(x) - \mathcal{L}_0 : \dot{e}(x) \)

Elasto-plastic constitutive: \( \dot{e}(x) = \dot{e}^e(x) + \dot{e}^p(x) = M : \dot{\sigma} + M^p(\sigma) \)

Equilibrium equation in Greens function form: \( -\mathcal{L}_{ijkl} G_{km,lj}(x-x') + \delta_{im} \delta(x-x') = 0 \)

Equilibrium equation in Fourier space: \( \xi_i \xi_j \mathcal{L}_{ijkl} \hat{G}_{km}(\xi) = -\delta_{im} \)

Solution: \( \text{FFT}(v_i) = \hat{v}_i = i \xi_j \hat{G}_{im}(\xi) \hat{\varphi}_{mj}(\xi) \quad \text{FFT}(v_{i,k}) = \hat{v}_{i,k} = -\xi_k \xi_j \hat{G}_{im}(\xi) \hat{\varphi}_{mj}(\xi) \)

For each time step, the following iterative algorithm is considered:

1. At the beginning of the 1\textsuperscript{st} iteration, give an initial guess to the strain rate: \( n+1 \dot{e}^0(x) = n \dot{e}(x) \). Then compute initial stress, \( n+1 \sigma^0(x) \), using local constitutive relations.
2. Compute the polarization field, \( i \varphi(x) \), for the \( i \text{th} \) iteration: \( \varphi(x) = \hat{\sigma}(x) - \mathcal{L}_0 : \hat{e}(x) \)
3. Transform the polarization field to Fourier space via FFT: \( i \hat{\varphi}(\xi) = \text{FFT}(i \varphi(x)) \)
4. Compute strain rate in the Fourier space at the \( (i+1) \text{th} \) iteration:
   \[
   i+1 \dot{e}(\xi) = \text{sym} \left( \hat{\Gamma}(\xi) : i \hat{\varphi}(\xi) \right); \quad \forall \xi \neq 0, \quad \text{and} \quad i+1 \dot{e}(0) = \hat{E} \quad \hat{\Gamma}_{ijkl}(\xi) = -\xi_j \xi_l \hat{G}_{ik}(\xi) \]
5. Transform strain rate back to the real space through inverse FFT: \( i+1 \dot{e}^n(x) = \text{FFT}^{-1}(i+1 \dot{e}(\xi)) \)
6. Compute stress (and/or stress rate) field using updated strain rate according to constitutive model.
7. Check for convergence
   \[
   \delta = \left| \frac{i+1 e - i e}{i+1 e} \right| \quad e = \left( \frac{\| \text{div}(i+1 \sigma) \|^2}{\| i+1 \sigma \|^2} \right)^{1/2} = \left( \frac{\| \xi \times i+1 \sigma(\xi) \|^2}{\| i+1 \sigma \|^2} \right)^{1/2} \quad \text{Equilibrium error}
   \]
   If not converged, repeat steps (2) to (7).
Crystal Plasticity using Fast Fourier Transform

- Plane strain deformation:
  
- Effective stress – strain curve of the microstructure.
Strain, plastic strain, and stress fields:

CVPFFT

CEPFFT

CPFEM

Strain

Plastic strain

Stress

Micromechanical Response
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:

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

Graphical Model Learning

- 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.
The message update function is

\[
m_{i-1,j}(y_{i,j}) \propto \int \psi(y_{i,j}, y_{i-1,j}) \varphi_{i-1,j}(x_{i-1,j}, y_{i-1,j}) \prod_{(p,q)\in \Gamma(y_{i-1,j})/y_i} m_{p,q}(y_{i-1,j}) dy_{i-1,j}
\]

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 \( y_{k,l}^{(t)} \sim \zeta(y_{k,l}) \psi_{k,l}(x_{k,l}, y_{k,l}) \prod_{(p,q)\in \Gamma(y_{k,l})/y_i} m_{p,q}(y_{k,l}) \)
- 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:
Application in the Disk Forging Problem

- **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

- Probabilistic Graphical model representation
- Multiscale Forging simulation
- Non-Gaussian distribution
- Use Belief propagation to do prediction
- Reduce to 2D
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)

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

- Data-driven modeling/design of heterogeneous media provides many open mathematical challenges
- Information theoretic approaches to predictive science
- Need for managing complexity in stochastic multiscale models
- The Bayesian approach offers:
  - Probability measure over surrogates (finite number of samples)
  - Active learning (experimental design)
  - Non-stationary responses (trees of surrogates)
  - Samples of the surrogate distribution can be used for: UQ, Sensitivity Analysis, Model Calibration, ...
  - Current research on Corregionalization models, fully Bayesian trees
- Use of probabilistic graphical models allows stochastic multiscale problems to be addressed as inference problems in graphs - possibly linear scaling algorithms!
Predictive Materials Science: Open Problems

- Surrogate reduced-order stochastic microstructure models
- Bayesian approach to model error: interatomic potentials, ..
- Information theoretic approaches to coarse graining in the presence of uncertainties
- Rare events: Fracture, Nucleation, ...
- Stochastic Multiscale Modeling: Curse of Dimensionality
- Exploring Process/Structure/Property Relations: The Materials Genome Project
- Other...