An Information Theoretic Approach Towards Predictive Materials Modeling

Professor Nicholas Zabaras
Materials Process Design and Control Laboratory and
Center for Applied Mathematics
101 Frank H T Rhodes Hall,
Sibley School of Mechanical and Aerospace Engineering,
Cornell University, Ithaca, NY 14853, USA
nzabaras@gmail.com
http://mpdc.mae.cornell.edu/

Scientific Computing and Imaging Institute, University of Utah, Salt Lake City, Utah, April 18, 2014
Outline

- Introduction and Core Issues in Predictive Science & Engineering
- Common Themes in Predictive Modeling
- A Bayesian Approach to Uncertainty Quantification
- Predictive Materials Science: An information Theoretic Approach to Coarse Graining
- Modeling of Random Polycrystalline Materials
- Graph Theoretic Approach to UQ for Multiscale Problems
- Closing Remarks
Simulations of physical systems accounting explicitly for uncertainty in mission/operating conditions, material properties, and manufacturing imperfections, by using state-of-the-art stochastic modeling & computing approaches.

New paradigm: *Integration of computational mathematics, computational statistics, scientific computing with engineering science & design.*

Data-Driven Science: *A transformative integration of theory, modeling and experiments.*

Design in the presence of uncertainties: Uncertainty is not simply viewed as a unquantifiable risk but as an opportunity to endow greater flexibility and robustness in the system.
Numerical solution of stochastic multiscale/multiphysics PDEs with emphasis on high-dimensional problems

Certified stochastic reduced-order (surrogate) models

Stochastic multiscale science (information theoretic approaches to coarse graining, information propagation across scales, etc.)

Multiscale Inverse Problems

Probabilistic approach to systems synthesis, design-optimization and control under uncertainty

Model error and model selection

Modeling of rare events

Uncertainty visualization, ... and many more
Predictive Science & Engineering

Computational Mathematics

Computational Statistics

Computer Science & Scientific Computing

Uncertainty Quantification

Data-Driven Science

Stochastic Multiscaling

Graph Theoretic Inference

Design with Uncertainty

Exascale Computing

Complex Interconnected Systems

Materials Genome

Materials Science

Nanoscale Science & Device Design

Energy Systems

Fluid – Structure Interaction
Problem 1: Stochastic Input Model Construction

**Rock Permeability**

**Problem 2: Uncertainty Propagation**

How can you construct a probabilistic input model \( p(x) \)?

How can you quantify the uncertainty in the response \( p(y) \)?
Applications: Common UQ Themes

Secondary Oil Recovery  ~1-2 hours

Rock Permeability

Simulation

Water Saturation

Disc Forging  ~0.5-1 day

Mechanical Properties

Microstructure

Challenges:
- Stochastic input high-dimensional
- Input/output Multi-scale
- Response Multi-output/space/time
- Expensive Solver
- Limited simulations

Component Life Prediction

Microstructure
Surface finish
Lubricant properties
... Fatigue Life

~3-10 days
Stochastic Input Model Reduction

- High-dimensionality does not allow direct density estimation.

\[ p(x) = ? \text{(HARD)} \]

- Data live in low dimensional manifolds embedded in high-dimensional spaces.

\[ R^{1000 \ldots} + \text{Physical Restrictions} \rightarrow \text{Lower-D Manifold} \]

- The manifolds are in general non-linear.

Data-Driven Stochastic Input Model

**High-dimensional Space**

**Non-linear Dimensionality Reduction**

**Reconstruction**

**Reduced Space**

**Reduction:**
- Kernel Principal Component Analysis
- Isomap, Diffusion Maps, GTM,…

**Reconstruction:**
- Exploit locally linear nature of data manifolds
- Optimization problem

**Density Estimation**
- Polynomial Chaos Exp. to known variables
- Graph Structure Learning (Wan & Zabaras, 2013)
Example: Channelized Permeability

Data:
- Channelized permeability
- 0 and 1
- 2025 dimensions
- 1000 observations

Stochastic Input Model:
- 30 dimensions
- Uniform random variables

Construction Framework

Samples

Uncertainty Propagation

Variance of water saturation

Stochastic Input: Key Ideas
- Reduction
- Reconstruction
- Density Estimation

Uncertainty Propagation

\[ p(x) \quad \rightarrow \quad p(y) \]

Problem 2: Uncertainty Propagation

• Easiest way: Monte Carlo

• Monte Carlo requires millions of observations to converge.

• However, with \textbf{expensive code} we have \textbf{limited samples}.

• Limited samples means \textbf{epistemic uncertainty in predicted statistics}.

• Capturing epistemic uncertainty, requires \textbf{Bayesian} methodology.

• Moving away from gPCE, Sparse Grid Collocation, etc.

• We build a \textbf{Bayesian Surrogate} to replace the expensive code.
A Gaussian Process (GP) is a generalization of a multivariate Gaussian distribution to infinitely many variables.

\[ p \left( \hat{f}(\cdot) \right) = \text{GP} \]

Mean function instead of mean vector.

Covariance function instead of covariance matrix.

The most important choice:
- Regularity
- Periodicity
- Length scales
**Gaussian Process Regression**

Data: Input-Output $\mathcal{D}$

Prior GP $\hat{f}(\cdot) \sim p(\hat{f}(\cdot))$

Bayes Rule

Posterior GP $\hat{f}(\cdot)|\mathcal{D} \sim p(\hat{f}(\cdot)|\mathcal{D})$

95% error bars

Possible surrogates

We use every possible surrogate compatible with $\mathcal{D}$. Epistemic uncertainty induced by limited data.
**Example: Stochastic Elliptic PDE**

10 Inputs

Bayesian surrogate + Active Learning

Adaptive Sparse Grid Collocation.

Significantly outperforms competitive methods!

Error in mean predicted variance of response.

Number of observations.

Multiple Spatio-Temporal Outputs

- The response is a vector of $q$-dimensions:
  \[
  f(x) = (f_1(x), \ldots, f_q(x)).
  \]

- It might also be a function of space/time:
  \[
  x = (\xi, x_s, t).
  \]

- The prior is a $q$-dimensional GP:
  \[
  \hat{f}(\cdot) \sim p\left(\hat{f}(\cdot)\right) = \text{GP}_q\left(\hat{f}(\cdot) \mid m(\cdot), k(\cdot, \cdot)\Sigma\right).
  \]

\[
\Sigma \in \mathbb{R}^{q \times q}: \text{Captures linear correlations between outputs.}
\]

\[
\begin{align*}
  m &: \mathcal{X} \rightarrow \mathbb{R}^q \\
  k(x, x') &= k_\xi(\xi, \xi')k_s(x_s, x_s')k_t(t, t')
\end{align*}
\]

Reference:
Example: Flow Through Porous Media

Input: $\xi$ = Random permeability (50 dimensions).

Response:
- $u(x_s, t; \xi)$: x-velocity
- $v(x_s, t; \xi)$: y-velocity
- $p(x_s, t; \xi)$: pressure

Summary of Key ideas:
- Build Bayesian Surrogate and replace code.
- Effect of epistemic uncertainty quantified.
- Informative selection of simulations
- Multiple spatio/temporal outputs and correlations
- Discontinuities in stochastic space like h-adaptivity
- High-dimensions through HDMR
- Sparse Gaussian Processes
- Many open issues....

Reference:

120 observations + GP $\approx$ 100,000 MC simulations!
Electronic Structure

Nanoscale molecular

Research Areas III and IV: Entropic Coarse Graining and Stochastic Multiscaling

Macroscale Forging Problem

Mesoscale crystalline structure

Research Area II: Quantifying uncertainty in interatomic potentials

Research Area I: Structure Representation

Research Area V: Modeling Rare Events

Research Area VI: Materials Design – Location Specific Microstructures
Towards Materials by Design

Can you find the material that has this property (or is close) doing only computer experiments?

Guide experimental efforts.

Save money and time.

Can you quantify how uncertainty propagates across scales?
Can you infuse these tasks with experimental data collected on all scales?
Towards Materials by Design

Stochastic Coarse Graining:

Stochastic Coarse Graining

Fine Scale Model
• $n \gg 1$ degrees of freedom ($r$)
• probability density of $r$

Coarse Scale Model
• $N \ll n$ degrees of freedom ($R$)
• probability density of $R$ given $c$
• $c$ missing parameters

Relative Entropy Principle
• Find $c$ by minimizing $\text{Rel Entr.}$
• Noisy gradients and Hessian

Stochastic Optimization
• Robbins-Monro
• Stochastic Meta-descent
• Pairwise Conjugation
Stochastic Coarse Graining

Statistics

Fine:

\[ p(q) \propto e^{-\beta V(q)} \]

\[ q \rightarrow Q = \xi(q) \]

Coarse:

\[ p(Q) \propto e^{-\beta V_{CG}(Q)} \]

Example: Coarse Graining Water

Applications: Protein folding, DNA, Complex fluids, etc.

From:
Run Quantum Mechanical Codes on a sample set of structures:
- Construct formation energies
- Construct “design matrix”

Map of Thermodynamic Quantities
- Sample states according to Boltzmann factor
- Average over quantity evaluated for these states

Fast Surrogate Models
  - Cluster Expansion
    - Polynomial expansion on a model for property based on what atom sits where on a fixed Bravais lattice

Match Boltzmann Distribution
  - Relative Entropy
    - Information Theoretic arguments minimizing “distance” to true distribution:
    - $\exp \left( -\frac{E}{kT} \right)$
    - Answers the question: What cluster expansion gives best match?

Thermodynamic Monte Carlo Simulations

We show better agreement with experiments!
Optimal Cluster Expansion for Thermodynamic Modeling

Methodology

1) Which clusters needed to best describe ground states?

1) What should we choose for $\gamma_i$? (the Effective Cluster Interactions or ECI)

2) Traditionally: Use least squares fitting with leave-one-out cross validation

3) **New Idea** In computing thermodynamic quantities: Use relative entropy, matching Boltzmann distribution!

Determines $\phi_i$ (and $M$) in the polynomial (cluster expansion):

$$E(\sigma | y) = \sum_{i=1}^{M} \gamma_i \phi_i(\sigma),$$

Toy model reveals: Better predictive capabilities!

Relative Entropy

$$S_{rel}[y] := \sum_{\sigma} p(\sigma | \beta) \ln \left[ \frac{p(\sigma | \beta)}{p(\sigma | y, \beta)} \right].$$

J. Kristensen & N. Zabaras, PRB, 2013a, 2013b
Phase Transitions in Magnesium Lithium

Results – Magnesium Lithium --

Difference in ECI between least squares and relative entropy large around transitions (see right figure)

Relative entropy: Agrees better with experiments!

Experiments:
- 33 % Mg in Mg-Li
  Difference = 56 K
- 50 % Mg in Mg-Li
  Difference = 90 K
- 66 % Mg in Mg-Li
  Difference = 33 K

Difference in ECI 140-200 K
Surrogate Models: Induced Uncertainty

• **Interest**
  – Thermodynamic computations with quantum mechanical energies:
    • Compute phase diagrams
  – Optimization of properties:
    • Minimize thermal conductivity
  – Etc.

• **Requirement**
  – Traverse configuration space with Monte Carlo methods
  – Requires million+ energy evaluations
  – We need fast way of obtaining free energy

• **Problem**
  – Accurate (e.g., ab initio) method not feasible
    • Too expensive (takes too long!)

• **Solution**
  – Sacrifice accuracy for speed:
    • Build surrogate model based on few ab initio simulations
  – Obtain energies in required time
Surrogate Models: Induced Uncertainty

- **Example of expensive computer codes**
  - Quantum mechanical energies obtained via VASP[1]
  - Heat conductivity from LAMMPS[2]

- **Building a surrogate model**
  - Approximate computer code with parametrized surface
  - E.g., expand response surface in a basis

- **Predictions**
  - Predictions (such as phase transitions) are made based on surrogates

- **Important problem not addressed so far:**
  - What is the induced uncertainty in predicted quantities from using particular parametrizations of surrogates?
    - E.g.: How does the choice of basis functions and expansion coefficients affect accuracy in predictions
  - Can help answer whether surrogate can even be used to capture response surface!

- **Possible solution:**
  - Model parametrized surrogate in fully Bayesian way

---

What is the parametrization uncertainty in predicted quantities?

Model surrogate via Bayesian paradigm

- Put prior belief on parametrization
  \[ p(\theta) \]
  - This represents our belief about the parametrization before seeing any data
  - Uncertainty in \( \theta \) propagates to prediction \( I \) as follows:

\[
p(I) = \int d\theta \delta(I[f(\cdot; \theta)] - I) p(\theta)
\]

Quantify Predictive Uncertainty

- Then, update our belief about the quantity of interest to:

\[ p(I|\mathcal{D}) = \int d\theta \delta(I[f(\cdot; \theta)] - I)p(\theta|\mathcal{D}) \]

- Quantify uncertainty in predictions:
  - In other words: Draw samples from \( p(I|D) \)

```
comment: Collect parametrizations of surrogate: \( \mathcal{T} = \{\theta_i\}_{i=1}^M \) from (some) Bayesian posterior

comment: Collect (expensive) data set \( \mathcal{D} = \{x_j\}_{j=1}^N \)

for i ← 1 to M
    for j ← 1 to N
        do {        do compute \( f(x_j; \theta_i) \)
           Compute and store quantity \( I_i[\{f(x_j; \theta_i)\}_{j=1}^N] \)

comment: Compute 1st, 50th, and 99th percentiles of \( \{I_i\}_{i=1}^M \)
```
• Replace VASP response surface with surrogate model
  – What should the surrogate model be?

• Quantum Mechanical energy invariant under space group operations of the lattice
  – Surrogate must account for this

• Possible (and popular) surrogate in alloy modeling:
  – The Cluster Expansion

\[ E^{(i)} = \sum_{\mathcal{F}} J_{\mathcal{F}} \langle \Pi_{\mathcal{F}} \rangle_{\sigma^{(i)}} \]

Accounts for symmetries

The Cluster Expansion

+ Corresponds to: Multidimensional discrete Fourier transform[1]
+ Generalized Ising model

\[ E^{(t)} = J_0 + J_1 \sum_i \sigma_i^{(t)} + J_{2,1} \sum_{i,j} \sigma_i^{(t)} \sigma_j^{(t)} + \cdots + J_{3,1} \sum_{i,j,k} \sigma_i^{(t)} \sigma_j^{(t)} \sigma_k^{(t)} + \cdots \]

Bayesian Parametrization of the Surrogate

- Infinite series: True response surface is part of parametrization class
- Truncation required:
  - How do we choose truncated parametrization?

\[ E^{(t)} = J_0 + J_1 \sum_i \sigma_i^{(t)} + J_{2,1} \sum_{i,j} \sigma_i^{(t)} \sigma_j^{(t)} + \cdots + J_{3,1} \sum_{i,j,k} \sigma_i^{(t)} \sigma_j^{(t)} \sigma_k^{(t)} + \cdots \]

\[ \gamma = (1, 1, 0, \cdots, 1, \cdots) \]

\[ \beta_\gamma = (J_0, J_1, 0, \cdots, J_{3,1}, \cdots) \]

\[ \theta = (\gamma, \beta_\gamma) \]

\[ E^{(i)} = \sum_F J_F \langle \Pi_F \rangle_{\sigma^{(i)}} \quad \rightarrow \quad y = X \beta_\gamma \]
RJMCMC Chain: Algorithm

Input: The number of iterations \( T \). Random walk step size \( \varepsilon \).
Data: \( X \) and \( y \).
Output: \( \{\theta^{(t)} = (\beta^{(t)}, \gamma^{(t)})|t \in \{0, \ldots, T\}\} \).

begin
\[\text{Initialization: set } \theta^{(0)} = (\beta^{(0)}, \gamma^{(0)}) \text{ and } t = 1.\]
repeat
  if \( k^{(t-1)} = 1 \) then
    \[k^{(t)} \leftarrow k^{(t-1)} + U([0,1]).\]
  else if \( k^{(t-1)} = p \) then
    \[k^{(t)} \leftarrow k^{(t-1)} - U([0,1]).\]
  else
    \[k^{(t)} \leftarrow k^{(t-1)} + U([-1,0,1]).\]
end
Sample \( S \sim N(0,\varepsilon^2). \)
\[K \leftarrow \gamma^{(t-1)} \text{ and } K^c \leftarrow \{1, \ldots, p\} \setminus K.\]
if \( k^{(t)} = k^{(t-1)} \) then
  Sample \( j \sim U(K). \)
  Update \( \beta_j^{(t)} \leftarrow \beta_j^{(t-1)} + s \) with an MH step, details in Section 3.2.
else if \( k^{(t)} = k^{(t-1)} + 1 \) then
  Sample \( j \sim U(K^c). \)
  Perform a “birth” move and update \( \beta_j^{(t-1)} \), details in Section 3.3.
else
  Sample \( j \sim U(K). \)
  Perform a “death” move and update \( \beta_j^{(t-1)} \), details in Section 3.3.
end
\[t \leftarrow t + 1.\]
until \( t = T.\)
end

\[\min \left\{ \left( \frac{\|y - X\beta^{(1)}\|_2}{\|\beta^{(1)}\|_1} \right)^{-k} \times \left( \frac{\|y - X\beta^{(n-1)}\|_2}{\|\beta^{(n-1)}\|_1} \right)^{-(n-1)}, 1 \right\}. \]

(10)

\[\min \left\{ \frac{k^2}{p-k} \times \frac{\|\beta^{(k+1)}\|_1}{\|\beta^{(1)}\|_1} \times \frac{\|y - X\beta^{(n-1)}\|_2}{\|\gamma - y\|_2} \times \frac{p(\gamma \to \gamma)}{p(\gamma' \to \gamma')} \times N(u; 0, \varepsilon^2)^{-1}, 1 \right\}. \]

(11)
Quantify Uncertainty in Ground State

Data set $D$

LASSO with 10-fold CV
Least squares

Approximate ground state line

99 % confidence region

Form energy (meV)

% Mg in Mg-Li
Phase Transition in Si-Ge
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
- Expensive deterministic simulation

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

**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
Different microstructure realizations satisfying some experimental correlations

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.
**An Intuitive Picture of the Problem**

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 $d << n$.

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.
Data-Driven Uncertainty Quantification

**Microstructure space**

START: Extract microstructural features

- Property variability
- Property convex hull
- Property distribution

**Low-dimensional space**

Model reduction

- Low-D points
- Feature reconstruction
- Low-D representation reconstruction
- Sample from hypercube

- PCE
- Low-D space
- Map to hypercube

**Microstructure 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

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

Variability of Fatigue Properties of Ni-based Superalloys

Study the variability of fatigue indicator parameters (FIPs) of nickel-based superalloy microstructures.

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

**Model Reduction**
Linear and kernel principal component to simplify polycrystalline microstructures.

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

**Objective**
Study the variability of mechanical properties of large scale forging process due to microstructure variation.

**Challenge**
Resolve the “curse of dimensionality” induced by microstructure location dependence.

Grain structures of a nickel-base superalloy turbine disk

**Solution**
Bi-orthogonal Karhunen-Loeve expansion.

- **Bi-orthogonal KLE**
  \[
  \mathbf{A}(x,s,\omega) = \mathbf{\bar{A}}(x,s) + \sum_{j=1}^{\infty} \sqrt{\lambda_j} \mathbf{\Psi}_j(s) \Phi_j(x,\omega)
  \]

- **Second-level KLE**
  \[
  \mathbf{\Phi}(x,\omega) = \mathbf{\Phi}(x) + \sum_{j=1}^{\infty} \sqrt{\gamma_j} \mathbf{\psi}_j(x) \phi_j(\omega)
  \]

- **Forcing**
  \[
  \nabla_x \cdot \{ \mathbf{p} \} + \mathbf{f} = \mathbf{0} \\
  \hat{\mathbf{f}} - \hat{\mathbf{f}}_0 = \alpha \mu \beta \sqrt{\rho}
  \]

- **PCE**
  \[
  \mathbf{\psi}(\mathbf{\alpha'}) = \mathbf{\psi}(\mathbf{\alpha}) = \sum_{j=1}^{\infty} \gamma_j \mathbf{\psi}_j(\mathbf{\alpha'})
  \]

- **(e) Properties distribution**
  - Stress mean
  - Stress std
  - Strength mean
  - Strength std

- **(b) Macro-random coupled modes**
  - **(c) Reduced surrogate space**
  - **(d) Uniform distribution**

- **(a) Multiscale microstructure sample**
  - Stochastic Multiscale Simulation

- **(f) Forging**

- **(g) Simulation**

- **(h) Multiscale**

- **(i) Karhunen**

- **(j) Loeve**

- **(k) Second**

- **(l) Level KLE**

- **(m) PCE**

- **(n) Compute**

- **(o) Variation**

- **(p) Stochastic**

- **(q) Induced**

- **(r) Microstructure**

- **(s) Location**

- **(t) Dependence**

- **(u) Simulate**

- **(v) Coupled**

- **(w) Modes**

- **(x) Series**

- **(y) Expansion**

- **(z) Multiscale**

- **(aa) Forging**

- **(bb) Process**

- **(cc) Induced**

- **(dd) Microstructure**

- **(ee) Location**

- **(ff) Dependence**

- **(gg) Grain**

- **(hh) Structures**

- **(ii) Nickel-base**

- **(jj) Superalloy**

- **(kk) Turbine disk**

Stochastic Multiscale Models: A Graph Theoretic Approach

- Pose multiscale SPDEs in graphs
- Factorize conditional PDF of responses given the stochastic input on a graph using ‘clique’ potentials.
- Introduce hidden variables to account for coarse graining in constitutive models.
- Non-parametric approach
- All parameters are learned with local inference (EM, SMC, Variational,..)
- Conditional and marginal PDFs are computed with belief propagation – potentially linear scaling!
- The probabilistic graphical model can be used as a surrogate model (provides response for any realization of the input)
- Data and models become one and the same!

- J. Wan and N. Zabaras, JCP, 2013
- P. Chen and N. Zabaras, JCP, 2013
Consider flow in porous media as an example of a multiscale SPDE. A probabilistic model for \( p(Y|a) \) (response given permeability) is constructed with the help of graphical model.

We assume a constant pressure, \( h_k \), on each coarse element. Besides, the flow in coarse elements interact with each other through edges. Thus the flux on middle points of edges of coarse elements are also of interest.

Finally, the target physical responses include pressure and flux, i.e. \( Y = (u, h) \) where \( h = \{h_1, ..., h_{N_c}\} \) denotes the pressure and \( u = \{u_1, ..., u_{M_c}\} \) the flux.
As there is little prior information on the relationship between random variables \( \{u, h\} \) and \( a \), all nodes in each coarse element are assumed correlated and thus are linked in the graph.

The spatial correlations between physical responses are considered.

**Assumption**
- Each response \( u_i \), \( h_k \) is only correlated to its neighboring nodes (including local feature \( a_k \)) within the same coarse element.
- Long distance interactions among variables are ignored.
Probabilistic Modeling of SPDEs

Maximal clique: potential
\[ q_k \left( u_{I_k}, h_k; a_k \right) \]

Joint probability
\[ p(u, h | a) \propto \prod_k q_k \left( u_{I_k}, h_k; a_k \right) \]
In this way, the nodes on a coarse element $E_k : \{ u_{i \in I_k}, h_k, a_k \}$, form a maximal clique in an undirected graphical model. $I_k$ is the index set of responses $u_i$ in element $E_k$.

$$p(u, h \mid a) \propto \prod_{k} q_k (u_{I_k}, h_k ; a_k)$$

where $q_k()$ is the potential function of the maximal clique on $E_k$. Thus the global joint distribution is factorized into local potential functions on coarse elements.

$$p(u, h \mid a) \propto \exp \left( - \sum_{k} \mathcal{E}_k (u_{I_k}, h_k ; a_k) \right)$$

such that

$$q_k (u_{I_k}, h_k ; a_k) = \exp \left( - \mathcal{E}_k (u_{I_k}, h_k ; a_k) \right)$$
Apply the definition of energy function in these sub-problems, the local energy functions are expressed by

\[ \mathcal{E}_k (u_{i_k}, h_k; a_k) \approx \sum_{i \in I_k} \phi_{k,i} (u_i, a_k) + \sum_{(i,j) \in I_k \times I_k, i \neq j} \phi_{k,ij} (u_i, u_j, a_k) + \phi_{k,0} (h_k, a_k) + \sum_{i \in I_k} \phi_{k,i0} (u_i, h_k, a_k) \]

where

\[ \phi_{k,i} (u_i, a_k) = f_{k,i} (a_k) u_i + f_{k,ii} (a_k) u_i^2, \quad \phi_{k,ij} (u_i, u_j, a_k) = f_{k,ij} (a_k) u_i u_j \]

\[ \phi_{k,0} (h_k, a_k) = f_{k,0} (a_k) h_k + f_{k,00} (a_k) h_k^2, \quad \phi_{k,i0} (u_i, h_k, a_k) = f_{k,i0} (a_k) u_i h_k \]

Since the functions of local features \( a_k \) in the energy functions are unknown, a nonparametric model is adopted

\[ f_{k,\cdot} (a_k) \equiv f_{k,\cdot} (a_k; \theta_k) = \theta_{k,1} + \sum_{t=2}^{r} \theta_{k,t} \zeta_t (a_k) \]

with unnormalized Gaussian kernels

\[ \zeta_t (a_k) = \exp \left( -\frac{\|a_k - \bar{a}_t\|^2}{\sigma_{\zeta}^2} \right) \]
Probabilistic Modeling of SPDEs

Maximal clique: potential

\[ q_k (u_{I_k}, h_k; a_k) = \exp \left( -\mathcal{E}_k (u_{I_k}, h_k; a_k) \right) \]

\[ \mathcal{E}_k (u_{I_k}, h_k; a_k) \approx \sum_{i \in I_k} \phi_{k,i} (u_i, a_k) + \sum_{(i,j) \in I_k \times I_k, i \neq j} \phi_{k,ij} (u_i, u_j, a_k) + \phi_{k,0} (h_k, a_k) + \sum_{i \in I_k} \phi_{k,0} (u_i, h_k, a_k) \]
Given a set of samples of input \( \{a_k^{(n)}\}_{n=1}^{N} \) and specifying the number of kernels, the centers of Gaussian kernels are determined using K-means clustering.

A typical choice of the kernel width \( \sigma_\zeta \) is the average minimum distance between two realizations in the input space, i.e.

\[
\sigma^2_\zeta = \frac{1}{N} \sum_{i=1}^{N} \min_{i \neq j} \| a_k^{(i)} - a_k^{(j)} \|^2
\]

The function \( f_{k,\cdot} \) is a mapping from local features \( a_k \) to a scalar variable \( \zeta_k \). In other words, \( f_{k,\cdot} \) projects the high-dimensional input into a low-dimensional space.

Since these variables \( \{\zeta_k\} \) are not directly observable, we call them hidden variables in the probabilistic model.

These hidden variables capture non-parametrically the effects of the fine futures on the coarse variables (constitutive equations).
Conditional probability distribution:

\[ p(u, h | a, \Theta) \propto \prod_k \prod_{i \in I_k} \exp \left( -f_{k,i} (a_k; \theta_k) u_i - f_{k,ii} (a_k; \theta_k) u_i^2 \right) \cdot \prod_{(i,j) \in I_k \times I_k, i \neq j} \exp \left( -f_{k,ij} (a_k; \theta_k) u_i u_j \right) \cdot \prod_{i \in I_k} \exp \left( -f_{k,i0} (a_k; \theta_k) u_i h_k \right) \exp \left( -f_{k,00} (a_k; \theta_k) h_k \right) \]

The relationships between hidden variables and local features are

\[ \xi_{k,i} = f_{k,i} (a_k; \theta_k), \xi_{k,ii} = f_{k,ii} (a_k; \theta_k), \xi_{k,ij} = f_{k,ij} (a_k; \theta_k) \]

\[ \xi_{k,i0} = f_{k,i0} (a_k; \theta_k), \xi_{k,0} = f_{k,0} (a_k; \theta_k), \xi_{k,00} = f_{k,00} (a_k; \theta_k) \]

Then the conditional probability can be formulated in terms of hidden variables \( \xi \)

\[ p(u, h | \xi) \propto \prod_k \prod_{i \in I_k} \exp \left( -\xi_{k,i} u_i - \xi_{k,ii} u_i^2 \right) \cdot \prod_{(i,j) \in I_k \times I_k, i \neq j} \exp \left( -\xi_{k,ij} u_i u_j \right) \cdot \prod_{i \in I_k} \exp \left( -\xi_{k,i0} u_i h_k \right) \exp \left( -\xi_{k,00} h_k \right) \]
According to the definition of hidden variables, each hidden variable is completely fixed given corresponding local features \( a_k \) and hyperparameters, i.e.

\[
p(\xi_{k,i} | \xi_{k,j}, a_k, \theta_k) = p(\xi_{k,i} | a_k, \theta_k)
\]

In other words, the hidden variables are conditionally independent on local features, thus

\[
p(\xi | a, \Theta) \propto \prod_k p(\xi_{k,0} | a_k, \theta_k) \cdot \prod_{(i,j) \in I_k \times I_k} p(\xi_{k,ij} | a_k, \theta_k) \\
\cdot \prod_{i \in I_k} p(\xi_{k,i} | a_k, \theta_k) \cdot p(\xi_{k,00} | a_k, \theta_k)
\]

where \( \Theta = \bigcup_k \theta_k \).

Since there exist deterministic relationships between any \( \xi_{k,i} \) and \( a_k \), \( \xi_{k,i} \) takes value at \( f_{k,i}(a_k) \) with probability 1 given \( a_k \). Then the conditional probability of \( \xi_{k,i} \) on \( a_k \) is a Delta function

\[
p(\xi_{k,i} | a_k, \theta_k) = \delta(\xi_{k,i} - f_{k,i}(a_k; \theta_k))
\]
The hidden variables $\xi$ capture fine-scale effects on the coarse-scale, i.e. the influence of high-dimensional stochastic input on responses is represented by $\xi$. The undirected graphical model and the corresponding factor graph with hidden variables are depicted as
The complete probabilistic model for $p(u,h|a)$ is then factorized as a product of potential functions measuring the interactions between random variables.

$$\mu_{k,0}(h_k, \xi_{k,0}, \xi_{k,00}) = \exp(-\xi_{k,0}h_k - \xi_{k,00}h_k^2)$$

$$\mu_{k,i}(h_k, u_i, \xi_{k,l}) = \exp(-\xi_{k,l0}u_i h_k)$$

$$p(\xi_{k,ii}|a_k) = \delta_{\xi}(f_{k,ii}(a_k))$$

$$\mu_{k,i}(u_i, \xi_{k,i}, \xi_{k,ii}) = \exp(-\xi_{k,i}u_i - \xi_{k,ii}u_i^2)$$

$$\mu_{k,ij}(u_i, u_j, \xi_{k,ij}) = \exp(-\xi_{k,ij}u_i u_j)$$
We now need to learn the various parameters that define the probabilistic graphical model. Suppose we have a training set \( D = \left\{ a^{(t)}, u^{(t)}, h^{(t)} \right\}_{t=1}^{N} \). The likelihood function of training data is formulated as

\[
p(D | \Theta) \equiv \prod_{t=1}^{N} p \left( u^{(t)}, h^{(t)} | a^{(t)}, \Theta \right) p \left( a^{(t)} \right)
\]

\[
\propto \prod_{t=1}^{N} \exp \left( -\sum_{k} \mathcal{E}_{k} \left( u_{I_{k}}^{(t)}, h_{k}^{(t)}; a_{k}^{(t)}, \theta_{k} \right) \right)
\]

\[
\propto \prod_{k} \exp \left( -\sum_{t=1}^{N} \mathcal{E}_{k} \left( u_{I_{k}}^{(t)}, h_{k}^{(t)}; a_{k}^{(t)}, \theta_{k} \right) \right)
\]

Set the prior distribution for \( \Theta \), the Bayesian posterior is

\[
p(\Theta | D) \propto p(D | \Theta) p(\Theta)
\]

where \( p(\Theta) \) is chosen to be a multivariate Gaussian with mean zero and an identity covariance matrix.
As the prior leads to \( p(\Theta) = \prod_k p(\theta_k) \), the posterior distribution can be decomposed as

\[
p(\Theta | D) \equiv \prod_k p(\theta_k | D_k) \propto \prod_k \exp \left( -\sum_{t=1}^N \mathcal{E}_k \left( u_i^{(t)}, h_i^{(t)}; a_i^{(t)}, \theta_k \right) \right) p(\theta_k)
\]

As a result, \( \Theta \) in the probabilistic graphical model can be estimated locally!!

We will use a special Monte Carlo method --- **Sequential Monte Carlo** (SMC) to estimate the hyperparameters \( \theta_k \) on each coarse element through local posterior \( p(\theta_k | D_k) \). The samples that maximize the posterior will be taken as fixed parameter values in the probabilistic model for inference [de Moral etc. 2006].
Belief Propagation

Belief propagation (BP) algorithm is the most commonly used techniques for the inference problem in graphical model.

Two type of messages (update rules)

variable node to factor node

\[ m_{x_i \rightarrow \mu}^{(n)} (x_i) \leftarrow \mu_i (x_i) \prod_{\mu_{pi} \in \Gamma(x_i) \setminus \mu} m_{\mu_{pi} \rightarrow x_i}^{(n-1)} (x_i) \]

factor node to variable node

\[ m_{\mu \rightarrow x_i}^{(n)} (x_i) \leftarrow \int \mu(x_{\mu \setminus x_i}) \prod_{x_i \in x_{\mu \setminus x_i}} m_{x_i \rightarrow \mu}^{(n)} (x_i) \, dx_{\mu \setminus x_i} \]

Marginal distribution

\[ p^{(n)} (x_i) \propto \mu_i (x_i) \prod_{\mu \in \Gamma(x_i)} m_{\mu \rightarrow x_i}^{(n)} (x_i) \]

\[ p^{(n)} (x) \propto \mu(x) \prod_{x_j \in x \mu \in \Gamma(x_j)} m_{\mu \rightarrow x_j}^{(n)} (x) \]
In the BP algorithm used in this work, the messages are updated in parallel. At each iteration, we calculate the messages from each factor node to its neighboring variable nodes as well as the messages from each variable node to its neighboring factor nodes.

The messages are considered as converged if their change is less than a threshold in two successive iterations. There exists a unique message between any factor node and one of its arguments.
For all messages except those between hidden variables, since there is no prior information, they are represented non parametrically (as weighted Gaussian mixtures)

Without loss of generality, consider the message from factor node $\mu_{k,ij}(u_i, u_j, \xi_{k,ij})$ to variable node $u_i$

$$m_{\mu_{k,ij} \rightarrow u_i}(u_i) \approx \sum_{t=1}^{T} l_t \mathcal{N}(u_i; \bar{u}_i^t, \sigma_i^2)$$

At iteration $n$ of the BP algorithm, the messages between factor nodes and variables are updated by

$$m_{\mu_{k,ij} \rightarrow u_i}^{(n)}(u_i) \leftarrow \int \mu_{k,ij}(u_i, u_j, \xi_{k,ij}) m_{u_j \rightarrow \mu_{k,ij}}^{(n)}(u_j) m_{\xi_{k,ij} \rightarrow \mu_{k,ij}}^{(n)}(\xi_{k,ij}) d\xi_{k,ij} du_j$$
Another challenge lies in the update of messages between hidden variables. Although analytic expressions of \( p(a) \) and \( p(\zeta|a) \) are explicit, the joint distribution of hidden variables \( \zeta \) could be complicated such that the links between them are implicit when stochastic input has been removed from the graph.

To bypass the difficulties in passing messages between hidden variables, the graphical model is transformed as follows:
Step 0: Initialization

- initialize the input messages as the marginal distributions of hidden variables

\[ m^{(0)}(\xi_{k,ij}) = p(\xi_{k,ij}), m^{(0)}(\xi_{k,i}) = p(\xi_{k,i}) \]

- initialize all messages between \( u \) and factor node as

\[ m^{(0)}_{\mu_{k,ij} \rightarrow u_i}(u_i) = \sum_{t=1}^{n} l_t N(u_i; \bar{u}_i^{(t)}, \sigma_i^2) \]

with \( l_t = 1/n, \bar{u}_i^{(t)} = t/n, \sigma_i^2 = 1 \)
Generalized Belief Propagation: Iteration

- At step $t$

  - Input messages $m^{(t)}(\xi_{k,i})$ have been fixed in step $t-1$

  - Update the message from $\mu_{k,i}$ to $u_i$

  $m^{(t)}_{\mu_{k,i} \rightarrow u_i} \leftarrow \int \mu_{k,i} (u_i, \xi_{k,i}, \xi_{k,ii}) m^{(t)}(\xi_{k,i}) m^{(t)}(\xi_{k,ii}) d\xi_{k,i} d\xi_{k,ii}$

  This is obtained by making samples $\left(u_i, \xi_{k,i}, \xi_{k,ii}\right)$ of function $\mu_{k,i}(u_i, \xi_{k,i}, \xi_{k,ii}) m^{(t)}(\xi_{k,i}) m^{(t)}(\xi_{k,ii})$ using standard Metropolis-Hastings sampler

  $m^{(t)}_{\mu_{k,j} \rightarrow u_i}$ are obtained by Gaussian mixture approximation of samples $(u_i)$, i.e.

  $m^{(t)}_{\mu_{k,j} \rightarrow u_i}(u_i) \approx \sum_{t=1}^{n} l_i N(u_i; \bar{u}_i^{(t)}, \sigma_i^2)$

  - Update the message from $\mu_{k,ij}$ to $u_i$ and the message from $\mu_{k,ij}$ to $u_j$ in the same way

  $m^{(t)}_{\mu_{k,ij} \rightarrow u_i} \leftarrow \int \mu_{k,ij} (u_i, u_i, \xi_{k,ij}) m^{(t)}(\xi_{k,ij}) m^{(t-1)}_{u_j \rightarrow \mu_{k,ij}}(u_j) d\xi_{k,ij} du_j$

  $m^{(t)}_{\mu_{k,ij} \rightarrow u_j} \leftarrow \int \mu_{k,ij} (u_i, u_i, \xi_{k,ij}) m^{(t)}(\xi_{k,ij}) m^{(t-1)}_{u_i \rightarrow \mu_{k,ij}}(u_i) d\xi_{k,ij} du_i$
Generalized Belief propagation: Iteration

- At step $t$

  \[
  m^{(t)}(\xi_{k,ii})
  \]

  • Update the input message at step $t + 1$

  \[
  m^{(t)}_{\mu_{k,ij} \rightarrow \xi_{k,ij}}(\xi_{k,ij}) \leftarrow \int m^{(t)}_{\mu_{i} \rightarrow u_{i}}(u_{i}, u_{i}, \xi_{k,ij}) m^{(t)}_{\mu_{j} \rightarrow u_{j}}(u_{j}) du_{i} du_{j}
  \]

  \[
  m^{(t+1)}(\xi_{k,ij}) = p(\xi_{k,ij}) / m^{(t)}_{\mu_{k,ij} \rightarrow \xi_{k,ij}}(\xi_{k,ij})
  \]
Generalized Belief Propagation: Convergence

- At step $t$, compute the marginal distributions of each physical responses by multiplying all incoming messages from neighboring factor nodes

\[ p(u_i) \propto \prod_{\mu \in \Gamma(u_i)} m_{\mu \rightarrow i}(u_i) \approx \sum_{i=1}^{n} l_i N(u_i; \bar{u}_i^{(t)}, \sigma_i^2) \]

if \( \max_t |\bar{u}_i^{(t)} - \bar{u}_i^{(t-1)}| < \varepsilon \), the marginal distribution of $u_i$ converges. The same rule is applied to pressure $h_k$.

- Stop iteration until all marginal distributions converge.
Given a realization of stochastic input, $a^{(n)}$, the values of hidden variables can be directly obtained through functions $f(.)$.

As the hidden variables are observed in this case, there is no message between them. Then the factor graph corresponds to the conditional distribution $p(u,h|\xi^{(n)})$. The unobserved variables are $(u,h)$.

When belief propagation is performed, we get the marginals of physical responses conditioned on the input. Let the expectation $E(u_i|a^{(n)})$ and $E(h_k|a^{(n)})$ be predicted values of physical responses, we can get a surrogate model by running the belief propagation algorithm on a factor graph given a realization of stochastic input.
Stochastic Multiscale Models: A Graph Theoretic Approach

\[ \text{Cov}(x,y) = \sigma^2 \exp \left( -\frac{|x_1 - y_1|}{L_1} \right) \]
\[ \text{Cov}(x,y) = \sigma^2 \exp \left( -\frac{|x_2 - y_2|}{L_2} \right) \]
\[ \sigma = 1.0, L_1 = 0.1, L_2 = 0.1 \]

Predicted physical responses given a realization of stochastic input (a)-(c) x-velocity, y-velocity and pressure obtained from direct simulation, and (d)-(f) x-velocity, y-velocity and pressure predicted by the probabilistic graphical model (trained with 60 data points)
Predicted mean of x-velocity (a) MC simulation with $10^6$ samples, and probabilistic model trained by (b) 20, (c) 40, (d) 60 data.
Predicted variance of x-velocity (a) MC simulation with $10^6$ samples, and probabilistic model trained by (b) 20, (c) 40, (d) 60 data.
Predicted marginal PDFs

Predicted marginal PDF of x-velocity at point (0.5,0.4375) with (a) 2 and (b) 4 Gaussian components in nonparametric messages

Predicted marginal PDF of y-velocity at point (0.4375,0.5) with (a) 2 and (b) 4 Gaussian components in nonparametric messages
The joint PDF of the x-velocity $u_1$ at (0.5, 0.4375) and $u_2$ at (0.375, 0.4375): (a) direct simulation (b) probabilistic graphical model; the joint PDF of y-velocity $v_1$ at (0.4375, 0.5) and $v_2$ at (0.4375, 0.375): (c) direct simulation (d) probabilistic graphical model.
Multiscale Disk Forging: Microstructure Uncertainty

Location-specific microstructures

Mesoscale uncertainties

Finite discretization of microstructure

Linking Mesoscale and Macroscale via the integration point

Address high dimensionality issue

Stochastic Homogenization

Macro FE mesh

Macroscale uncertainties

Uncertainty propagation problem addressed by solving an inference problem on the graph

Due to symmetry, 2D forging problem is actually solved

[P. Chen & N. Zabaras, JCP, 2013]
[J. Wan & N. Zabaras, JCP, 2013]
The structure of the graph

\[
p(Y, S, \xi | \mathcal{D}) \approx \prod_{i=1}^{N_G} p\left(y_i | s_i, \mathcal{D}\right) \left( \prod_{j \in \Gamma(i), j > i} p\left(y_i, y_j | s_i, s_j, \mathcal{D}\right) \right) p(s_i | \xi) p(\xi)
\]

\[
\propto \prod_{i=1}^{N_G} \left( \prod_{k \in V(y)} \varphi_k \left(y_k, s_k\right) \prod_{(i,j) \in E(y)} \psi_{i,j} \left(y_i, y_j, s_i, s_j\right) \prod_{k \in V(s)} \zeta_k \left(s_k, \xi\right) \right) \rho(\xi)
\]

[The structure of the graph diagram]

[Construction of the probabilistic graphical model (factor graph)]

[p. Chen & N. Zabaras, JCP, 2013]

[D. Koller & N. Frideman, Probabilistic Graphical Model, 2009]
**Pole figure reconstruction**

Test the localized model reduction

- Test
- Reconstructed
- Test
- Reconstructed

(a) Located at the bottom of the preform  
(b) Located at the top of the preform
Equivalent Strain

(a) MC: N=8000

(b) NGM: N=100

(c) NGM: N=100

(d) NGM: N=400

(e) Mean

(f) std
Slip System Resistance

(a) MC: N=8000

(b) NGM: N=100

(c) NGM: N=400

Mean

std
Comparison of pdf at a single point located around the center of the deformed body.
Surrogate Prediction

Equiv. strain

(a)

Equiv. stress

(b)

(c)

(d)

Slip system resistance

(e)

(f)

True response

Predicted response with N=400
Issues to Address

- The graph considered is based on a *discretization of the spatial domain*. This allows an obvious modeling of correlations in space.
  - This is not a feasible approach if data are sparse and not on a grid.

- *Data are fully prescribed on a grid* – this needs to be relaxed as data are sparsely defined in the spatial domain.
  - Need for inference on a graph with incomplete data.

- A decomposition of the stochastic space is performed that in general may not be possible
  - Curse of Dimensionality
  - Insufficient data for parts of the stochastic support space.
  - *DATA SHOULD SPEAK ON ITSELF AS HOW THE STOCHASTIC SPACE SHOULD BE PARTITIONED*
Probabilistic graphical model representation

- **Infinite mixture of GPs** (Dirichlet Prior) – Data cluster in finite number of clusters.
- Computing all parameters using **variational inference** (rather than MCMC)
- Predictions are done via mixture of infinite number of GPs (model average, over fitting not an issue)

Peng & Zabaras, JCP (submitted)
Numerical Example: KO-1d problem

- Kraichnan-Orszag (KO) problem

Governing equation:
\[
\begin{align*}
\frac{dy_1}{dt} &= y_1y_3, \\
\frac{dy_2}{dt} &= -y_2y_3, \\
\frac{dy_3}{dt} &= -y_1^2 + y_2^2,
\end{align*}
\]

Initial condition:
\[
\begin{align*}
y_1 &= 1, \\
y_2 &= 0.1x, \\
y_3 &= 0, \\
x &\sim U([-1, 1])
\end{align*}
\]

Predicted variance \((n_\xi = 91)\) v.s MC variance \((n=1e4)\)

Data-selection process
KO-2d problem

Predicted variance \((n_\xi = 200, 400)\) v.s MC variance \((n=1e4)\)
KO-2d pdf prediction

\[
\begin{array}{ccc}
\text{pdf} & \text{pdf} & \text{pdf} \\
\text{95% CI} & \text{95% CI} & \text{95% CI} \\
\text{mean} & \text{mean} & \text{mean} \\
\text{MC} & \text{MC} & \text{MC} \\
\end{array}
\]

\[
\begin{array}{ccc}
y_2(t=4) & y_2(t=4) & y_3(t=4) \\
0 & 2 & 0 \\
0 & 1 & 0 \\
0 & 1 & 0 \\
\end{array}
\]

\[
\begin{array}{ccc}
y_2(t=6) & y_2(t=6) & y_3(t=6) \\
0 & 2 & 0 \\
0 & 1 & 0 \\
0 & 1 & 0 \\
\end{array}
\]

\[
\begin{array}{ccc}
y_2(t=8) & y_2(t=8) & y_3(t=8) \\
0 & 2 & 0 \\
0 & 1 & 0 \\
0 & 1 & 0 \\
\end{array}
\]

\[
\begin{array}{ccc}
y_2(t=10) & y_2(t=10) & y_3(t=10) \\
0 & 2 & 0 \\
0 & 1 & 0 \\
0 & 1 & 0 \\
\end{array}
\]

\[
\begin{array}{ccc}
n_\xi = 100 & n_\xi = 200 & n_\xi = 400 \\
0 & 2 & 0 \\
0 & 1 & 0 \\
0 & 1 & 0 \\
\end{array}
\]
KO-3d problem

Predicted variance ($n_\xi = 1000, 4000$) v.s MC variance ($n=1e4$)

$n_\xi = 1000$

$n_\xi = 4000$
Predicted pdf ($n_\xi = 4000$) v.s MC estimates ($n=1e4$)
Simulations of physical systems accounting explicitly for uncertainty in mission/operating conditions, material properties, and manufacturing imperfections.

New paradigm: Integration of computational mathematics, computational statistics, scientific computing with engineering science & design.

Data-Driven Science: A transformative integration of theory, modeling and experiments.

Design in the presence of uncertainties: Uncertainty is viewed as an opportunity to endow greater flexibility and robustness in the system.