Applications of Uncertainty Quantification and Sensitivity Analysis in Smart Materials and Adaptive Structures

Ralph C. Smith
Department of Mathematics
North Carolina State University

William S. Oates
Department of Mechanical Engineering
Florida A&M/Florida State University

![Graphs showing nominal stress, PDF, and time response plots.](image-url)
Course Structure

Overview: 10:30-3:00

1. Introduction: Motivating examples
2. Overview of terminology and inverse problems
3. Bayesian inference
4. Forward uncertainty propagation
5. Global sensitivity analysis
6. Surrogate model construction
1. Introduction: Predictive Science

**Components:** All involve uncertainty

- Experiments
- Models
- Simulations

- **Experimental results are believed by everyone, except for the person who ran the experiment**, source anonymous, quoted by Max Gunzburger, Florida State University.
- **Essentially, all models are wrong, but some are useful**, George E.P. Box, Industrial Statistician.
- **Computational results are believed by no one, except the person who wrote the code**, source anonymous, quoted by Max Gunzburger, Florida State University.
- **I have always done uncertainty quantification. The difference now is that it is capitalized.** Bill Browning, Applied Mathematics Incorporated.
Modeling Strategy

General Strategy: Conservation of stuff

<table>
<thead>
<tr>
<th>Stuff</th>
<th>x</th>
<th>x + Δx</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
\frac{d\text{Stuff}}{dt} = \text{Stuff in} - \text{Stuff out} + \text{Stuff created} - \text{Stuff destroyed}
\]

Continuity Equation:

\[
\frac{\partial (\rho \Delta x)}{\partial t} = \phi(t, x) - \phi(t, x + \Delta x)
\]

\[
\Rightarrow \lim_{\Delta x \to 0} \frac{\partial \rho}{\partial t} = \lim_{\Delta x \to 0} \frac{\phi(t, x) - \phi(t, x + \Delta x)}{\Delta x}
\]

\[
\Rightarrow \frac{\partial \rho}{\partial t} + \frac{\partial \phi}{\partial x} = 0
\]

Density: \(\rho(t, x)\) - Stuff per unit length or volume

Rate of Flow: \(\phi(t, x)\) - Stuff per second

More Generally:

\[
\Rightarrow \frac{\partial \rho}{\partial t} + \frac{\partial \phi}{\partial x} = \text{Sources} - \text{Sinks}
\]
Example 1: Weather Models

Challenges:

1. Coupling between temperature, pressure gradients, precipitation, aerosol species, etc.;
2. Models and inputs contain uncertainties;
3. Numerical grids necessarily larger than many phenomena; e.g., clouds;
4. Sensors positions may be uncertain; e.g., weather balloons, ocean buoys.

Goal:

1. Assimilate data to quantify uncertain initial conditions and parameters;
2. Make predictions with quantified uncertainties.
Equations of Atmospheric Physics

Conservation Relations:

\[
\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0
\]

\[
\frac{\partial \mathbf{v}}{\partial t} = -\mathbf{v} \cdot \nabla \mathbf{v} - \frac{1}{\rho} \nabla p - g \hat{k} - 2\Omega \times \mathbf{v}
\]

\[
\rho c_v \frac{\partial T}{\partial t} + p \nabla \cdot \mathbf{v} = -\nabla \cdot \mathbf{F} + \nabla \cdot (k \nabla T) + \rho \mathbf{q}(T, p, \rho)
\]

\[p = \rho RT\]

Water

\[
\frac{\partial m_j}{\partial t} = -\mathbf{v} \cdot \nabla m_j + S_{m_j}(T, m_j, x_j, \rho), \quad j = 1, 2, 3,
\]

Aerosol

\[
\frac{\partial x_j}{\partial t} = -\mathbf{v} \cdot \nabla x_j + S_{x_j}(T, x_j, \rho), \quad j = 1, \ldots, J,
\]

Constitutive Closure Relations: e.g.,

\[
S_{m_2} = S_1 + S_2 + S_3 - S_4
\]

where

\[
S_1 = \bar{\rho}(m_2 - m_2^*)^2 \left[ 1.2 \times 10^{-4} + \left( 1.569 \times 10^{-12} \frac{n_r}{d_0(m_2 - m_2^*)} \right) \right]^{-1}
\]
Ensemble Predictions

Ensemble Predictions:

Cone of Uncertainty:

00 UTC on August 26, 2005
12 UTC on August 26, 2005

General Questions:

• What is expected snowfall in Denver on March 4?
• What are average high and low temperatures?
• Note: Quantities are statistical in nature.
Example 2: Quantum-Informed Continuum Models

Objectives:

- Employ density function theory (DFT) to construct/calibrate continuum energy relations.
- Polydomain structure – Lead titanate

\[ u(P_i, \varepsilon_{ij}, P_{i,j}) = u_M(\varepsilon_{ij}) + u_L(P_i) + u_C(P_i, \varepsilon_{ij}) + u_G(P_{i,j}) \]

where

\[ u_L(P_3) = \alpha_1 P_3^2 + \alpha_{11} P_3^4 + \alpha_{111} P_3^6 \]
\[ u_C(P_3, \varepsilon_{ii}) = -q_{11} \varepsilon_{11} P_3^2 - q_{12} (\varepsilon_{11} P_3^2 + \varepsilon_{22} P_3^2) \]
\[ u_G(P_{3,1}) = \frac{1}{2} g_{44} P_{3,1}^2 \]

Domain Wall Energy: \( E_{180}(q_{180}) = \int_{-\infty}^{\infty} (u - u_0) dx_1 \)

Question: Can parameters be uniquely determined by DFT simulations?

Broad Objective:

- Use UQ/SA to help bridge scales from quantum to system
**Example 3: Viscoelastic Material Models**

**Application:** Adaptive materials for legged robotics

- **Figure:** Billy Oates

**Material Behavior:** Significant rate dependence
Example: Viscoelastic Material Models

**Material Behavior:** Significant rate dependence

**Finite-Deformation Model:** Nonlinear, non-affine

\[ \psi(q) = \psi_\infty(G_e, G_c, \lambda_{\text{max}}) + \gamma(\eta, \beta, \gamma) \]

- Dissipative energy function \( \gamma \)
- Conserved hyperelastic energy function

\[ \psi_\infty^N = \frac{1}{6} G_c l_1 - G_c \lambda_{\text{max}}^2 \ln(3\lambda_{\text{max}}^2 - I_1) + G_e \sum_j \left( \lambda_j + \frac{1}{\lambda_j} \right) \]

**Parameters:**

\[ q = [G_e, G_c, \lambda_{\text{max}}, \eta, \beta, \gamma] \]

- \( G_c \): Crosslink network modulus
- \( G_e \): Plateau modulus
- \( \lambda_{\text{max}} \): Max stretch effective affine tube
- \([\eta, \beta, \gamma] \): Viscoelastic parameters

**Uncertainty Quantification Goals:**

- Quantify measurement errors.
- Quantify uncertainty in parameters.
- Use statistics to quantify accuracy of considered models.
Example 3: Multiscale Model Development

Example: PZT-Based Macro-Fiber Composites
Example 4: Multiscale Model Development

**Example:** PZT-Based Macro-Fiber Composites

\[
\rho \ddot{u} = \nabla \cdot \sigma + F \\
\nabla \cdot D = 0, \quad D = \varepsilon_0 E + P \\
\nabla \times E = 0, \quad E = -\nabla \phi
\]

Continuum Energy Relations

\[
P^\alpha = d_\alpha \sigma + \chi^{\alpha} E + P_R^\alpha \\
\varepsilon^\alpha = s^{E\alpha} \sigma + d_\alpha E + \varepsilon_R^{\alpha}
\]

Homogenized Energy Model (HEM)

\[
P = d(E, \sigma) \sigma + \chi^{E} E + P_{irr}(E, \sigma) \\
\varepsilon = s^{E} \sigma + d(E, \sigma) E + \varepsilon_{irr}(E, \sigma)
\]
Example: PZT-Based MFC and Robobee

Beam Model: 20 parameters -- quantify uncertainties

\[ \rho \frac{\partial^2 w}{\partial t^2} + \gamma \frac{\partial w}{\partial t} - \frac{\partial^2 M}{\partial x^2} = 0 \]

\[ M = -c^E I \frac{\partial^2 w}{\partial x^2} - c_D I \frac{\partial^3 w}{\partial x^2 \partial t} \]
\[ - [k_1 e(E, \sigma_0) E + k_2 \varepsilon_{\text{irr}}(E, \sigma_0)] \chi_{\text{MFC}}(X) \]

Homogenized Energy Model (HEM)

2nd Example: Robobee Drive Mechanism
2. Challenge: Terminology and Notation

**Terminology:**

- **Inputs:** Parameters, initial conditions, boundary conditions, exogenous forces; e.g., parameters in HIV models, initial conditions in weather models.
- **Outputs or Responses:** Quantities that we experimentally or numerically measure; e.g., viral load, outlet temperature in reactor.
- **Quantities of Interest (QoI):** Statistical quantity that we want to compute; e.g., average CRUD buildup, expected profit for a given design.

**Input Notation:** Can vary even within disciplines!

- Math Control Community: \( q = [q_1, \ldots, q_p] \)
- Math Reduced-Order Community: \( p = [p_1, \ldots, p_q] \)
- Statistics: \( \theta = [\theta_1, \ldots, \theta_d] \)
- Nuclear Engineering: \( \alpha = [\alpha_1, \ldots, \alpha_k] \)
- Active subspace community: \( x = [x_1, \ldots, x_p] \)

**Note:** Same variability in notation for outputs and quantities of interest
First Challenge: Terminology and Notation

Terminology:

- Linearly parameterized problems: e.g., portfolio model $y = c_1 q_1 + c_2 q_2$
  - Rare in applications except constitutive relations and image processing
- Nonlinearly parameterized problems: typical case
  - Differs from linear or nonlinear in state; e.g., spring model

\[
\frac{d^2 y(t)}{dt^2} + ky(t) = 0
\]
\[
y(0) = y_0, \quad \frac{dy}{dt}(0) = 0
\]

Inputs: $q = [k, y_0]$

Response: Displacement $y(t) = y_0 \cos(\sqrt{k} \cdot t)$

\[
\dot{y}(t) + ky(t) = 0
\]
\[
y(0) = y_0, \quad \frac{dy}{dt}(0) = 0
\]

Note:

- Linear state dependence
- Nonlinear parameter dependence
Uncertainty Quantification

*I have always done uncertainty quantification. The difference now is that it is capitalized.* Bill Browning, Applied Mathematics Incorporated.

**Note:** The field of “Uncertainty Quantification” has grown rapidly over the last 20 years. How is “Capital UQ” different from what statisticians do extremely well every day?

- E.g., When I proposed a course on “Uncertainty Quantification” in Mathematics, I had to carefully justify its existence to Statistics.
- Statistics students are now starting to take the course.
Uncertainty Quantification

I have always done uncertainty quantification. The difference now is that it is capitalized. Bill Browning, Applied Mathematics Incorporated.

**Note:** The field of “Uncertainty Quantification” has grown rapidly over the last 20 years. How is “Capital UQ” different from what statisticians do extremely well every day?

- E.g., When I proposed a course on “Uncertainty Quantification” in Mathematics, I had to carefully justify its existence to Statistics.
- Statistics students are now starting to take the course.

**My Definition of “Capital UQ”:** The synergy between statistics, applied mathematics and domain sciences required to quantify uncertainties in inputs and QoI when models are too computationally complex to permit sole reliance on sampling-based methods.”

- Involves orthogonal polynomial techniques, sparse grids, high-D (infinite-D) approximation theory, randomized linear algebra … and a lot of statistics!

*No one trusts a model except the man who wrote it; everyone trusts an observation except the man who made it,* Harlow Shapely.
Steps in Uncertainty Quantification

**Note:** Uncertainty quantification requires synergy between statistics, mathematics and application area.

- Input Representation
- Local Sensitivity Analysis
- Global Sensitivity Analysis
- Parameter Selection
- Model Discrepancy
- Surrogate Models
- Model Calibration
- Uncertainty Propagation
- Stochastic Spectral Methods
- Sparse Grids
Deterministic Model Calibration

Example: MFC

\[ \frac{\partial^2 w}{\partial t^2} + \gamma \frac{\partial w}{\partial t} - \frac{\partial^2 M}{\partial x^2} = 0 \]

\[
M = -c^E I \frac{\partial^2 w}{\partial x^2} - c_D I \frac{\partial^3 w}{\partial x^2 \partial t} - \left[ k_1 e(E, \sigma_0) E + k_2 \varepsilon_{irr}(E, \sigma_0) \right] \chi_{MFC}(x)
\]

Homogenized Energy Model (HEM)

Note: 20 parameters

Point Estimates: Ordinary least squares

\[ q^0 = \arg \min_q \frac{1}{2} \sum_{j=1}^N \left[ w_j - w^N(t_j, \bar{x}, q) \right]^2 \]
Deterministic Model Calibration

Representative Parameter Values:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_+$ (m/V)</td>
<td>$478.10 \times 10^{-12}$</td>
</tr>
<tr>
<td>$\sigma_I$ (V/m)</td>
<td>$6.47 \times 10^6$</td>
</tr>
<tr>
<td>$\tau_{180}$ (s)</td>
<td>$2.80 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

1 Hz Input
(400, 800 VDC)

30 Hz Input

146 Hz Input

Phase Space

Frequency Sweep

Note: Point estimates but no quantification of uncertainty in:
- Model
- Parameters
- Data
Objectives for Uncertainty Quantification

Goal: Replace point estimates with distributions or credible intervals

E.g., Parameter Densities

E.g., Response Intervals
Objectives for Uncertainty Quantification

**Example:** Helmholtz energy $\psi(P, q) = \alpha_1 P^2 + \alpha_{11} P^4$, $q = [\alpha_1, \alpha_{11}]$

**Statistical Model:** Describes observation process

$$\upsilon_i = \psi(P_i, q) + \varepsilon_i \quad , \quad i = 1, \ldots, n$$

**Common Assumption:** $\varepsilon_i \sim N(0, \sigma^2)$

**UQ Goals:** Quantify parameter and response uncertainties
Strategy 1: Perform Experiments or High-Fidelity Simulations

Example: Helmholtz energy $\psi(P, q) = \alpha_1 P^2 + \alpha_{11} P^4$, $q = [\alpha_1, \alpha_{11}]$

Statistical Model: Describes observation process

$\nu_i = \psi(P_i, q) + \varepsilon_i$, $i = 1, \ldots, n$

Common Assumption: $\varepsilon_i \sim N(0, \sigma^2)$

UQ Goals: Quantify parameter and response uncertainties

Strategy 1: Perform experiments; e.g., 1
Strategy 1: Perform Experiments or High-Fidelity Simulations

Example: Helmholtz energy $\psi(P, q) = \alpha_1 P^2 + \alpha_{11} P^4$, $q = [\alpha_1, \alpha_{11}]$

Statistical Model: Describes observation process

$$v_i = \psi(P_i, q) + \varepsilon_i, \quad i = 1, \ldots, n$$

Common Assumption: $\varepsilon_i \sim N(0, \sigma^2)$

UQ Goals: Quantify parameter and response uncertainties

Strategy 1: Perform experiments; e.g., 2

![Graph of Helmholtz Energy vs Polarization P](image)
Strategy 1: Perform Experiments or High-Fidelity Simulations

Example: Helmholtz energy $\psi(P, q) = \alpha_1 P^2 + \alpha_{11} P^4$, $q = [\alpha_1, \alpha_{11}]$

Statistical Model: Describes observation process

$$v_i = \psi(P_i, q) + \varepsilon_i, \quad i = 1, \ldots, n$$

Common Assumption: $\varepsilon_i \sim N(0, \sigma^2)$

UQ Goals: Quantify parameter and response uncertainties

Strategy 1: Perform experiments; e.g., 3
Strategy 1: Perform Experiments or High-Fidelity Simulations

Example: Helmholtz energy $\psi(P, q) = \alpha_1 P^2 + \alpha_{11} P^4$, $q = [\alpha_1, \alpha_{11}]$

Statistical Model: Describes observation process

$V_i = \psi(P_i, q) + \epsilon_i$, $i = 1, \ldots, n$

Common Assumption: $\epsilon_i \sim N(0, \sigma^2)$

UQ Goals: Quantify parameter and response uncertainties

Strategy 1: Perform many experiments; e.g., 1000
Strategy 1: Perform Experiments or High-Fidelity Simulations

**Example:** Helmholtz energy $\psi(P, q) = \alpha_1 P^2 + \alpha_{11} P^4$, $q = [\alpha_1, \alpha_{11}]$

**Statistical Model:** Describes observation process

$$v_i = \psi(P_i, q) + \epsilon_i \quad , \quad i = 1, \ldots, n$$

**Common Assumption:** $\epsilon_i \sim N(0, \sigma^2)$

**UQ Goals:** Quantify parameter and response uncertainties

**Strategy 1:** Perform many experiments; e.g., 1000
Strategy 1: Perform Experiments or High-Fidelity Simulations

Example: Helmholtz energy $\psi(P, q) = \alpha_1 P^2 + \alpha_{11} P^4$, $q = [\alpha_1, \alpha_{11}]$

Statistical Model: Describes observation process

$\nu_i = \psi(P_i, q) + \epsilon_i$, $i = 1, \ldots, n$

Common Assumption: $\epsilon_i \sim N(0, \sigma^2)$

UQ Goals: Quantify parameter and response uncertainties

Strategy 1: Perform many experiments; e.g., 1000

Problem: Often cannot perform required number of experiments or high-fidelity simulations.

Solution: Statistical inference
3. Statistical Inference

**Goal:** The goal in statistical inference is to make conclusions about a phenomenon based on observed data.

**Frequentist:** Observations made in the past are analyzed with a specified model. Result is regarded as confidence about state of real world.

- Probabilities defined as frequencies with which an event occurs if experiment is repeated several times.
- Parameter Estimation:
  - Relies on *estimators* derived from different data sets and a specific sampling distribution.
  - Parameters may be unknown but are fixed and deterministic.

**Bayesian:** Interpretation of probability is subjective and can be updated with new data.

- Parameter Estimation: *Parameters are considered to be random variables having associated densities.*
Linear Regression

Statistical Model:
\[ Y = Xq_0 + \varepsilon \]

Assumptions:
(i) \( \mathbb{E}(\varepsilon_i) = 0 \)

(ii) \( \varepsilon_i \) iid (independent and identically distributed)
\[ \Rightarrow \text{var}(\varepsilon_i) = \sigma^2 \]
\[ \mathbb{E}[(\varepsilon_i - \mathbb{E}(\varepsilon_i))(\varepsilon_j - \mathbb{E}(\varepsilon_j))] = \text{cov}(\varepsilon_i, \varepsilon_j) = 0 \text{ for } i \neq j \]

Examples:
- iid errors
- Not identically distributed
- Not independent
Polarization Example

**Statistical Model:** For $i = 1, \ldots, n$

\[ \nu_i = \psi(P_i, q) + \varepsilon_i \quad \varepsilon_i \sim N(0, \sigma^2) \]

\[ = \alpha_1 P_i^2 + \alpha_{11} P_i^4 + \varepsilon_i \]

\[ \Rightarrow \begin{bmatrix} \nu_i \end{bmatrix} = \begin{bmatrix} P_i^2 & P_i^4 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_{11} \end{bmatrix} + \begin{bmatrix} \varepsilon_i \end{bmatrix} \]

\[ \Rightarrow \nu = Xq + \varepsilon \]

**Statistical Quantities:**

\[ q = (X^T X)^{-1} X^T \nu \]
Polarization Example

Statistical Model: For \( i = 1, \ldots, n \)

\[
\nu_i = \psi(P_i, q) + \varepsilon_i \quad \varepsilon_i \sim N(0, \sigma^2)
\]

\[
= \alpha_1 P_i^2 + \alpha_{11} P_i^4 + \varepsilon_i
\]

\[
\Rightarrow \left[ \begin{array}{c} \nu_i \\ \varepsilon_i \end{array} \right] = \left[ \begin{array}{cc} P_i^2 & P_i^4 \end{array} \right] \left[ \begin{array}{c} \alpha_1 \\ \alpha_{11} \end{array} \right] + \left[ \begin{array}{c} \varepsilon_i \end{array} \right]
\]

\[
\Rightarrow \nu = Xq + \varepsilon
\]

Statistical Quantities:

\[
q = (X^T X)^{-1} X^T \nu
\]

Note: \( \mathbb{E}(q) = \mathbb{E}[(X^T X)^{-1} X^T \nu] \)

\[
= (X^T X)^{-1} X^T \mathbb{E}(\nu)
\]

\[
= q_0 \quad \nu = Xq_0 + \varepsilon
\]
Polarization Example

Statistical Model: For $i = 1, \ldots, n$

$$\nu_i = \psi(P_i, q) + \epsilon_i \quad \epsilon_i \sim N(0, \sigma^2)$$

$$= \alpha_1 P_i^2 + \alpha_{11} P_i^4 + \epsilon_i$$

$$\Rightarrow \begin{bmatrix} \nu_i \end{bmatrix} = \begin{bmatrix} P_i^2 & P_i^4 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_{11} \end{bmatrix} + \begin{bmatrix} \epsilon_i \end{bmatrix}$$

$$\Rightarrow \nu = Xq + \epsilon$$

Statistical Quantities:

$$q = (X^T X)^{-1} X^T \nu$$

And: Let $A = (X^T X)^{-1} X^T$

$$V(q) = \mathbb{E}[(q - q_0)(q - q_0)^T]$$

$$= \mathbb{E}[(q_0 + A\epsilon - q_0)(q_0 + A\epsilon - q_0)^T]$$

since $q = A\gamma = A(Xq_0 + \epsilon)$

$$= A\mathbb{E}(\epsilon \epsilon^T) A^T$$

$$= \sigma^2 (X^T X)^{-1}$$

Note: $\mathbb{E}(q) = \mathbb{E}[(X^T X)^{-1} X^T \nu]$

$$= (X^T X)^{-1} X^T \mathbb{E}(\nu)$$

$$= q_0$$

$$\nu = Xq_0 + \epsilon$$
Polarization Example

Statistical Model: For $i = 1, \ldots, n$

$$
\nu_i = \psi(P_i, q) + \varepsilon_i \quad \varepsilon_i \sim N(0, \sigma^2) \\
= \alpha_1 P_i^2 + \alpha_{11} P_i^4 + \varepsilon_i \\
\Rightarrow \begin{bmatrix} \nu_i \end{bmatrix} = \begin{bmatrix} P_i^2 & P_i^4 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_{11} \end{bmatrix} + \begin{bmatrix} \varepsilon_i \end{bmatrix} \\
\Rightarrow \nu = Xq + \varepsilon
$$

Statistical Quantities:

$$
q = (X^T X)^{-1} X^T \nu \\
V = \sigma^2 (X^T X)^{-1} = \begin{bmatrix} 8.8 & -17.4 \\ -17.4 & 37.6 \end{bmatrix}
$$

Note: Covariance matrix incorporates “geometry”

Goal: Employ Bayesian inference for UQ
**Statistical Inference**

**Goal:** The goal in statistical inference is to make conclusions about a phenomenon based on observed data.

**Frequentist:** Observations made in the past are analyzed with a specified model. Result is regarded as confidence about state of real world.

- Probabilities defined as frequencies with which an event occurs if experiment is repeated several times.

- Parameter Estimation:
  - Relies on *estimators* derived from different data sets and a specific sampling distribution.
  - Parameters may be unknown but are fixed and deterministic.

**Bayesian:** Interpretation of probability is subjective and can be updated with new data.

- Parameter Estimation: *Parameters are considered to be random variables having associated densities.*
Bayesian Inference: Simpler Model

Example: Displacement-force relation (Hooke’s Law)

\[ s_i = E e_i + \varepsilon_i, \quad i = 1, \ldots, N \]

\[ \varepsilon_i \sim N(0, \sigma^2) \]

Parameter: Stiffness \( E \)

Strategy: Use model fit to data to update prior information

Non-normalized Bayes’ Relation:

\[ \pi(E|s) = e^{-\frac{\sum_{i=1}^{N} [s_i - E e_i]^2}{2\sigma^2}} \pi_0(E) \]
Bayesian Inference

**Bayes’ Relation:** Specifies posterior in terms of likelihood and prior

\[
\pi(q|\nu) = \frac{\pi(\nu|q)\pi_0(q)}{\int_{\mathbb{R}^p} \pi(\nu|q)\pi_0(q) dq}
\]

- **Likelihood:** \( e^{-\sum_{i=1}^{N} [s_i - Ee_i]^2 / 2\sigma^2} \)
- \( q = E \)
- \( \nu = [s_1, \ldots, s_N] \)

**Problem:** Can require high-dimensional integration

- e.g., MFC Model: \( p = 20! \)

**Solution:** Sampling-based Markov Chain Monte Carlo (MCMC) algorithms.

- Metropolis algorithms first used by nuclear physicists during Manhattan Project in 1940’s to understand particle movement underlying first atomic bomb.
Bayesian Inference: Motivation

Bayes’ Relation for Sets:

\[
P(B_i | A) = \frac{P(B_i \cap A)}{P(A)} = \frac{P(A | B_i) P(B_i)}{\sum_i P(A | B_i) P(B_i)}
\]

Bayes’ Relation for Functions: Specifies posterior in terms of likelihood, prior, and normalization constant.

Note:

- **Prior Distribution**: Quantifies prior knowledge of parameter values.
- **Likelihood**: Probability of observing a data if we have a certain set of parameter values.
- **Posterior Distribution**: Conditional probability distribution of unknown parameters given observed data (Updated distribution based on how model fits new data).
Bayesian Model Calibration

Bayesian Model Calibration:

• Parameters assumed to be random variables

\[ \pi(q|\nu) = \frac{\pi(\nu|q)\pi_0(q)}{\int_{\mathbb{R}_p} \pi(\nu|q)\pi_0(q) dq} \]

Example: Coin Flip

\[ \gamma_i(\omega) = \begin{cases} 0 & , \quad \omega = T \\ 1 & , \quad \omega = H \end{cases} \]

Likelihood:

\[ \pi(\nu|q) = \prod_{i=1}^{N} q^{\nu_i}(1 - q)^{1-\nu_i} \]

\[ = q^{N_1}(1 - q)^{N_0} \]

Posterior with Noninformative Prior: \( \pi_0(q) = 1 \)

\[ \pi(q|\nu) = \frac{q^{N_1}(1 - q)^{N_0}}{\int_0^1 q^{N_1}(1 - q)^{N_0} dq} = \frac{(N + 1)!}{N_0!N_1!} q^{N_1}(1 - q)^{N_0} \]
Bayesian Model Calibration:

• Parameters considered to be random variables with associated densities.

\[ \pi(q|\nu) = \frac{\pi(\nu|q)\pi_0(q)}{\int_{\mathbb{R}^p} \pi(\nu|q)\pi_0(q)\,dq} \]

Problem:

• Often requires high dimensional integration;
  - \( p = 20 \) for MFC example

Strategies:

• Sampling methods
• Sparse grid quadrature techniques
Markov Chain Monte Carlo Methods

Strategy:
- Sample values from proposal distribution $J(q^*|q^{k-1})$ that reflects geometry of posterior distribution
- Compute $r(q^*|q^{k-1}) = \frac{\pi(v|q^*)\pi_0(q^*)}{\pi(v|q^{k-1})\pi_0(q^{k-1})}$
  * If $r \geq 1$, accept with probability $\alpha = 1$
  * If $r < 1$, accept with probability $\alpha = r$

Intuition: Consider flat prior $\pi_0(q) = 1$ and Gaussian observation model

$$\pi(v|q) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-SS_q/2\sigma^2} \quad SS_q = \sum_{i=1}^{N} [v_i - f(t_i, q)]^2$$
Delayed Rejection Adaptive Metropolis (DRAM)

**Algorithm:** [Haario et al., 2006] – MATLAB, Python, R

1. Determine $q^0 = \arg \min_q \sum_{i=1}^{N} [u_i - \psi(P_i, q)]^2$

**Example:** Helmholtz energy

\[ u_i = \psi(P_i, q) + \varepsilon_i \quad \varepsilon_i \sim N(0, \sigma^2) \]
\[ = \alpha_1 P_i^2 + \alpha_{11} P_i^4 + \varepsilon_i \]
Delayed Rejection Adaptive Metropolis (DRAM)

**Algorithm:** [Haario et al., 2006] – MATLAB, Python, R

1. Determine \( q^0 = \arg \min_q \sum_{i=1}^{N} [v_i - \psi(P_i, q)]^2 \)

2. For \( k = 1, \ldots, M \)
   (a) Construct candidate \( q^* \sim N(q^{k-1}, V) \)

**Example:** Helmholtz energy

\[
v_i = \psi(P_i, q) + \varepsilon_i \quad \varepsilon_i \sim N(0, \sigma^2)
\]

\[
= \alpha_1 P_i^2 + \alpha_{11} P_i^4 + \varepsilon_i
\]

**Recall:** Covariance \( V \) incorporates geometry
Delayed Rejection Adaptive Metropolis (DRAM)

**Algorithm:** [Haario et al., 2006] – MATLAB, Python, R

1. Determine \( q^0 = \arg \min_q \sum_{i=1}^{N} [u_i - \psi(P_i, q)]^2 \)

2. For \( k = 1, \ldots, M \)
   (a) Construct candidate \( q^* \sim N(q^{k-1}, V) \)
   (b) Compute likelihood
   \[
   SS_{q^*} = \sum_{i=1}^{N} u_i - \psi(P_i, q^*) \]
   \[
   \pi(u|q) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-SS_q/2\sigma^2}
   \]
   (c) Accept \( q^* \) with probability dictated by likelihood
Delayed Rejection Adaptive Metropolis (DRAM)

**Algorithm:** [Haario et al., 2006] – MATLAB, Python, R

1. Determine \( q^0 = \arg \min_q \sum_{i=1}^N [v_i - \psi(P_i, q)]^2 \)

2. For \( k = 1, \ldots, M \)
   (a) Construct candidate \( q^* \sim N(q^{k-1}, V) \)
   (b) Compute likelihood

   \[
   SS_{q^*} = \sum_{i=1}^N v_i - \psi(P_i, q^*)]^2
   \]

   \[
   \pi(v|q) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-SS_q/2\sigma^2}
   \]

   (c) Accept \( q^* \) with probability dictated by likelihood

\[\frac{1}{(2\pi\sigma^2)^{n/2}} e^{-SS_q/2\sigma^2} \]
Delayed Rejection Adaptive Metropolis (DRAM)

**Algorithm:** [Haario et al., 2006] – MATLAB, Python, R

1. Determine $q^0 = \arg \min_q \sum_{i=1}^{N} [\nu_i - \psi(P_i, q)]^2$

2. For $k = 1, \ldots, M$
   (a) Construct candidate $q^* \sim N(q^{k-1}, V)$
   (b) Compute likelihood
   \[ SS_{q^*} = \sum_{i=1}^{N} \nu_i - \psi(P_i, q^*) ]^2 \]
   \[ \pi(\nu|q) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-SS_q/2\sigma^2} \]
   (c) Accept $q^*$ with probability dictated by likelihood
Delayed Rejection Adaptive Metropolis (DRAM)

**Algorithm:** [Haario et al., 2006] – MATLAB, Python, R

1. Determine $q^0 = \arg\min_q \sum_{i=1}^{N} [u_i - \psi(P_i, q)]^2$

2. For $k = 1, \ldots, M$
   (a) Construct candidate $q^* \sim N(q^{k-1}, V)$
   (b) Compute likelihood
      \[
      SS_{q^*} = \sum_{i=1}^{N} u_i - \psi(P_i, q^*)]^2
      \]
      \[
      \pi(u|q) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-SS_q/2\sigma^2}
      \]
   (c) Accept $q^*$ with probability dictated by likelihood

\[ q^* \]
\[ q^{k-1} \]
Delayed Rejection Adaptive Metropolis (DRAM)

Algorithm: [Haario et al., 2006] – MATLAB, Python, R

1. Determine $q^0 = \arg \min_q \sum_{i=1}^N [v_i - \psi(P_i, q)]^2$

2. For $k = 1, \ldots, M$
   (a) Construct candidate $q^* \sim N(q^{k-1}, V)$
   (b) Compute likelihood
      
      \[
      SS_{q^*} = \sum_{i=1}^N v_i - \psi(P_i, q^*)]^2
      \]
      
      \[
      \pi(v|q) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-SS_q/2\sigma^2}
      \]
   (c) Accept $q^*$ with probability dictated by likelihood
**Delayed Rejection Adaptive Metropolis (DRAM)**

**Algorithm:** [Haario et al., 2006] – MATLAB, Python, R

1. Determine $q^0 = \arg \min_q \sum_{i=1}^{N} [v_i - \psi(P_i, q)]^2$

2. For $k = 1, \ldots, M$
   (a) Construct candidate $q^* \sim N(q^{k-1}, V)$
   (b) Compute likelihood

   $$SS_{q^*} = \sum_{i=1}^{N} v_i - \psi(P_i, q^*)^2$$

   $$\pi(v|q) = \frac{1}{(2\pi\sigma^2)^n/2} e^{-SS_q/2\sigma^2}$$

   (c) Accept $q^*$ with probability dictated by likelihood

**Note:**
- Delayed Rejection: Shrink proposal: $\gamma V$
- Adaptive Metropolis: Update proposal as samples are accepted
Delayed Rejection Adaptive Metropolis (DRAM)

**Example:** Helmholtz energy with 3 parameters

\[ \psi(P, q) = \alpha_1 P^2 + \alpha_{11} P^4 + \alpha_{111} P^6 \]

**Note:** Similar results for \( \alpha_{11} \) and \( \alpha_{111} \)

**Pairwise Plots:** Quantify correlation

![Pairwise plots](image)
Example: Viscoelastic Material Models

**Material Behavior:** Significant rate dependence

**Finite-Deformation Model:** Nonlinear, non-affine

\[ \psi(q) = \psi_\infty(G_e, G_c, \lambda_{\text{max}}) + \gamma(\eta, \beta, \gamma) \]

- Dissipative energy function \( \gamma \)
- Conserved hyperelastic energy function

\[ \psi_\infty^N = \frac{1}{6} G_c I_1 - G_c \lambda_{\text{max}}^2 \ln(3 \lambda_{\text{max}}^2 - I_1) + G_e \sum_j \left( \lambda_j + \frac{1}{\lambda_j} \right) \]

**Parameters:**

\[ q = [G_e, G_c, \lambda_{\text{max}}, \eta, \beta, \gamma] \]

- \( G_c \): Crosslink network modulus
- \( G_e \): Plateau modulus
- \( \lambda_{\text{max}} \): Max stretch effective affine tube

\[ [\eta, \beta, \gamma] \]: Viscoelastic parameters

**UQ Goals:**

- Quantify uncertainty in parameters.
- Use UQ for model selection
  - E.g., linear versus nonlinear.
- Quantify models’ predictive capabilities for range of stretch rates.
Viscoelastic Model

Reduced Parameter Set:

\[ q = [\gamma, \eta, \beta] \], Fixed hyperelastic parameters

**Note:** Fastest stretch rate (0.67 Hz)

**Question:** How do we quantify uncertainty in response (stress)?

**Solution:** Propagate parameter and measurement uncertainties through model … in a few slides!
Bayesian Inference: Exercise

Example: Consider the spring model

\[ \ddot{z} + C\dot{z} + Kz = 0 \]
\[ z(0) = 2, \quad \dot{z}(0) = -C \]

and synthetic data generated with errors

\[ \varepsilon \sim iid N(0, \sigma^2) \text{ where } \sigma = 0.1. \]

Website: [https://rsmith.math.ncsu.edu/UQ_TIA/](https://rsmith.math.ncsu.edu/UQ_TIA/)

Exercise:

Download the code `spring_mcmc_C_K_sigma.m` from Chapter 8 of the website, which is a basic Metropolis algorithm for inferring C, K and the measurements. Run the code and familiarize yourself with the algorithm. We will modify the code to compute uncertainties associated with the displacement y.
Bayesian Inference: Advantages and Disadvantages

Advantages:

• Advantageous over frequentist inference when data is limited.
• Directly provides parameter densities, which can subsequently be propagated to construct response uncertainties.
• Can be used to infer non-identifiable parameters if priors are tight.
• Provides natural framework for experimental design.

Disadvantages:

• More computationally intense than frequentist inference.
• Can be difficult to confirm that chains have burned-in or converged.
Delayed Rejection Adaptive Metropolis (DRAM)

Websites:

- http://helios.fmi.fi/~lainema/mcmc/
Delayed Rejection Adaptive Metropolis (DRAM)

We fit the Monod model

\[ y = \theta_1 \frac{1}{\theta_2 + 1} + \epsilon, \quad \epsilon \sim N(0, \sigma^2) \]

to observations

- **x (mg / L COD):** 28 55 83 110 138 225 375
- **y (1 / h):** 0.053 0.060 0.112 0.105 0.099 0.122 0.125

First clear some variables from possible previous runs.

```matlab
clear data model options
```

Next, create a data structure for the observations and control variables. Typically one could make a structure data that contains fields `xdata` and `ydata`.

```matlab
data.xdata = [28 55 83 110 138 225 375]'; % x (mg / L COD)
data.ydata = [0.053 0.060 0.112 0.105 0.099 0.122 0.125]'; % y (1 / h)
```

Construct model

```matlab
modelfun = @(x,theta) theta(1)*x./(theta(2)+x);
ssfun = @(theta,data) sum((data.ydata-modelfun(data.xdata,theta)).^2);
model.ssfun = ssfun;
model.sigma2 = 0.01^2;
```
Delayed Rejection Adaptive Metropolis (DRAM)

Input parameters

\[
\text{params} = \{ \\
\quad \{\text{'theta1}', \text{tmin(1)}, 0\} \\
\quad \{\text{'theta2}', \text{tmin(2)}, 0\} \}; \\
\]

and set options

\[
\text{options.nsimu} = 4000; \\
\text{options.updatesigma} = 1; \\
\text{options.qcov} = \text{tcov}; \\
\]

Run code

\[
\text{[res,chain,s2chain] = mcmcrun(model,data,params,options);} \\
\]
Delayed Rejection Adaptive Metropolis (DRAM)

Plot results
figure(2); clf
mcmcplot(chain,[],res,'chainpanel');
figure(3); clf
mcmcplot(chain,[],res,'pairs');

Examples:
• Several available in MCMC_EXAMPLES
• ODE solver illustrated in algae example
Delayed Rejection Adaptive Metropolis (DRAM)

Construct credible and prediction intervals

```matlab
figure(5); clf
out = mcmcpred(res,chain,[],x,modelfun);
mcmcpredplot(out);
hold on
plot(data.xdata,data.ydata,'s'); % add data points to the plot
xlabel('x [mg/L COD]');
ylabel('y [1/h]');
hold off
title('Predictive envelopes of the model')
```
Bayesian Inference: Exercise

Example: Helmholtz energy with 3 parameters

\[ \psi(P, q) = \alpha_1 P^2 + \alpha_{11} P^4 + \alpha_{111} P^6 \]

Exercise:

1. Download the code Helmholtz_DRAM.m, and associated functions and data Helmholtz.txt, from the website

   https://rsmith.math.ncsu.edu/SPIE_SHORT_COURSE19/

   Also download and unzip MCMC_Stat.zip, which contains the DRAM software. You will need to set your paths. Run the code and generate the chains, marginal distributions and pairwise plots. The final plot are credible and prediction intervals, which we will discuss later.

2. Now modify the code to infer just the first two parameters. You may need to additionally modify your initial values.
Note: Uncertainty quantification requires synergy between statistics, mathematics and application area.
4. Forward Uncertainty Propagation: Linear Models

**Note:** Analytic mean and variance relations

**Example:** Helmholtz energy

\[ \gamma_i = \alpha_1 P_i^2 + \alpha_{11} P_i^4 + \varepsilon_i, \quad \text{var}[\varepsilon_i] = \sigma^2 \]

**Model Statistics:**

Let \( \bar{\alpha}_1, \bar{\alpha}_{11} \) and \( \text{var}(\alpha_1), \text{var}(\alpha_{11}) \) denote parameter means and variance. Then

\[
\mathbb{E}[\alpha_1 P_i^2 + \alpha_{11} P_i^4] = \bar{\alpha}_1 P_i^2 + \bar{\alpha}_{11} P_i^4 \\
\text{var}[\alpha_1 P_i^2 + \alpha_{11} P_i^4] = P_i^4 \text{var}[\alpha_1] + P_i^8 \text{var}[\alpha_{11}] + 2P_i^6 \text{cov}[\alpha_1, \alpha_{11}] 
\]

**Response Statistics:** Assume measurement errors uncorrelated from model response.

\[
\mathbb{E}[\gamma] = \bar{\alpha}_1 P_i^2 + \bar{\alpha}_{11} P_i^4 \\
\text{var}[\gamma] = P_i^4 \text{var}[\alpha_1] + P_i^8 \text{var}[\alpha_{11}] + 2P_i^6 \text{cov}[\alpha_1, \alpha_{11}] + \sigma^2 
\]

**Problem:** Models almost always nonlinearly parameterized
Forward Uncertainty Propagation: Sampling Methods

Strategy 1: Randomly sample from parameter and measurement error distributions and propagate through model to quantify response uncertainty.

Advantages:

• Applicable to nonlinear models.
• Parameters can be correlated and non-Gaussian.
• Straight-forward to apply and convergence rate is independent of number of parameters.
• Can directly incorporate both parameter and measurement uncertainties.

Disadvantages:

• Very slow convergence rate: $\mathcal{O}(1/\sqrt{M})$ where M is the number of samples.
• 100-fold more evaluations required to gain additional place of accuracy.
• This motivates numerical analysis techniques.

Strategy 2: Employ numerical surrogate representations to analytically propagate uncertainties.
Confidence, Credible and Prediction Intervals

Note:
- We now know how to compute the mean response for the QoI.
- How do we compute appropriate intervals?

Example: Helmholtz energy $\psi(P, q) = \alpha_1 P^2 + \alpha_{11} P^4$, $q = [\alpha_1, \alpha_{11}]$
Confidence, Credible and Prediction Intervals

Data: $\mathcal{Y} = [\mathcal{Y}_1, \cdots, \mathcal{Y}_n]$ of iid random observations

Confidence Interval (Frequentist): A $100 \times (1 - \alpha)\%$ confidence interval for a fixed, unknown parameter $q_0$ is a random interval $[L_c(\mathcal{Y}), U_c(\mathcal{Y})]$, having probability at least $1 - \alpha$ of covering $q_0$ under the joint distribution of $\mathcal{Y}$.

Credible Interval (Bayesian): A $100 \times (1 - \alpha)\%$ credible interval is that having probability at least $1 - \alpha$ of containing $q$.

Strategy: Sample out of parameter density $\rho_Q(q)$
Confidence, Credible and Prediction Intervals

**Data:** $\mathbf{Y} = [Y_1, \cdots, Y_n]$ of iid random observations

**Prediction Interval:** A $100 \times (1 - \alpha)$% prediction interval for a future observable $Y_{n+1}$ is a random interval $[L_c(Y), U_c(Y)]$ having probability at least $1 - \alpha$ of containing $Y_{n+1}$ under the joint distribution of $(Y, Y_{n+1})$.

**Example:** Consider linear model

$$Y_i = q_0 + q_1 x_i + \varepsilon_i, \quad i = 1, \cdots, n$$
Prediction Intervals for the Viscoelastic Model

**Linear Non-Affine Model:** Not accurate for predicting higher stretch rates

\[
\frac{d\lambda}{dt} = 6.7 \times 10^{-5} \text{ Hz}
\]

\[
\frac{d\lambda}{dt} = 0.335 \text{ Hz}
\]

\[
\frac{d\lambda}{dt} = 0.67 \text{ Hz}
\]
Prediction Intervals for the Viscoelastic Model

**Linear Non-Affine Model:**

\[ \frac{d\lambda}{dt} = 0.335 \text{ Hz} \]

\[ \frac{d\lambda}{dt} = 6.7 \times 10^{-5} \text{ Hz} \]

\[ \frac{d\lambda}{dt} = 0.67 \text{ Hz} \]

**Nonlinear Non-Affine Model:** Significantly more accurate over range of stretch rates!

\[ \frac{d\lambda}{dt} = 6.7 \times 10^{-5} \text{ Hz} \]

\[ \frac{d\lambda}{dt} = 0.335 \text{ Hz} \]

\[ \frac{d\lambda}{dt} = 0.67 \text{ Hz} \]
Prediction Intervals: Exercise

Example: Consider the spring model

$$\ddot{z} + C\dot{z} + Kz = 0$$

$$z(0) = 2, \quad \dot{z}(0) = -C$$

and synthetic data generated with errors

$$\varepsilon \overset{iid}{\sim} N(0, \sigma^2) \text{ where } \sigma = 0.1.$$ 

Exercise:

1. Use the code spring_mcmc_C_K_sigma.m, which you downloaded from Chapter 8 of the website [https://rsmith.math.ncsu.edu/UQ_TIA/](https://rsmith.math.ncsu.edu/UQ_TIA/) to compute the uncertainty in the displacement $z(2)$ by sampling out of the densities for $K$, $C$ and the measurement error.

2. Now download the code spring_dram.m and functions from the website [https://rsmith.math.ncsu.edu/SPIE_SHORT_COURSE19/](https://rsmith.math.ncsu.edu/SPIE_SHORT_COURSE19/) and run it to construct 95% prediction intervals for the spring model.
What UQ Can and Cannot Do (Not Comprehensive)

Can Do:

• Quantify uncertainty in model parameters or inputs based on experimental data or high-fidelity model simulations.

• Quantify correlation between model inputs.

• Quantify uncertainties in statistical quantities-of-interest. This is critical when specifying model predictions with quantified uncertainty.

Cannot Do:

• Accommodate or replace missing physics in models.
  
  o However, when combined with validation, it can indicate missing physics.

  o Research topic: quantifying model discrepancy.

• Guarantee optimal parameter values. However, it can be more robust than gradient-based optimization.

• Rank parameter sensitivity. This is addressed next!
Uncertainty Quantification Challenges

E.g., Viscoelastic model

\[ \psi(q) = \psi_\infty(G_e, G_c, \lambda_{\text{max}}) + \gamma(\eta, \beta, \gamma) \]

- Dissipative energy function \( \gamma \)
- Conserved hyperelastic energy function \( \psi_\infty \)

Parameters:

\[ q = [G_e, G_c, \lambda_{\text{max}}, \eta, \beta, \gamma] \]

Challenge 1:

- How do we isolate set of parameters that are identifiable in the sense that they can be uniquely inferred from data?

Challenge 2:

- How do we do uncertainty propagation for computationally intensive models? E.g., we have computational budget of 5000 but UQ requires 120,000 evaluations.
Parameter Selection Techniques

First Issue: Parameters often not identifiable in the sense that they are not uniquely determined by the data.

Example 1: Spring model

\[ m \frac{dy}{dt^2} + ky = 0 \]

\[ y(0) = y_0, \quad \frac{dy}{dt}(0) = 0 \]

Solution: \( y(t, q) = y_0 \cos \left( \sqrt{\frac{k}{m}} \cdot t \right) \)

Note: \( q = [k, m] \) not jointly identifiable
Parameter Selection Techniques

First Issue: Parameters often not identifiable in the sense that they are not uniquely determined by the data.

Example 1: Spring model

\[
\frac{m}{dt^2} \frac{dy}{dt} + ky = 0
\]

\[
y(0) = y_0, \quad \frac{dy}{dt}(0) = 0
\]

Solution: \(y(t, q) = y_0 \cos \left( \sqrt{\frac{k}{m}} \cdot t \right)\)

Note: \(q = [k, m]\) not jointly identifiable

Example 2: Polydomain structure – Lead Titanate

\[
u(P_i, \varepsilon_{ij}, P_{i,j}) = u_M(\varepsilon_{ij}) + u_L(P_i) + u_C(P_i, \varepsilon_{ij}) + u_G(P_{i,j})
\]

where

\[
u_L(P_3) = \alpha_1 P_3^2 + \alpha_{11} P_3^4 + \alpha_{111} P_3^6
\]

\[
u_C(P_3, \varepsilon_{ii}) = -q_{11} \varepsilon_{11} P_3^2 - q_{12} \left( \varepsilon_{11} P_3^2 + \varepsilon_{22} P_3^2 \right)
\]

\[
u_G(P_{3,1}) = \frac{1}{2} g_{44} P_{3,1}^2
\]

Domain Wall Energy: \(E_{180}(q_{180}) = \int_{-\infty}^{\infty} (u - u_0) dx_1\)

Question: Can parameters be uniquely determined by DFT simulations?
Parameter Selection: Required for models with unidentifiable or noninfluential inputs

- e.g., Nuclear neutron transport codes can have 100,000 inputs
Sensitivity Analysis: Motivation

Example: Linear constitutive relation

\[ \sigma = E e + c \frac{de}{dt} \]

Nominal Values: \( E = 100 \), \( c = 0.1 \)

\( e = 0.001 \), \( \frac{de}{dt} = 0.1 \)

Question: To which parameter \( E \) or \( c \) is stress most sensitive?

Local Sensitivity Analysis:

\[ \frac{\partial \sigma}{\partial E} = e = 0.001 \]

\[ \frac{\partial \sigma}{\partial c} = \frac{de}{dt} = 0.1 \]

Conclusion: Model most sensitive to damping parameter \( c \)

Limitations:

- Does not accommodate potential uncertainty in parameters.
- Does not accommodate potential correlation between parameters.
- Sensitive to units and magnitudes of parameters.
Global Sensitivity Analysis

Example: Linear elastic constitute relation

\[ \sigma = Ee + c \frac{de}{dt} \]

Nominal Values: \( E = 100, \ c = 0.1 \)

Uncertainty: 10% of nominal values

\[ E \sim \mathcal{U}(90, 110), \ c \sim \mathcal{U}(0.09, 0.11) \]

Local Sensitivities:

\[ \frac{\partial \sigma}{\partial E} = e = 0.001 \]
\[ \frac{\partial \sigma}{\partial c} = \frac{de}{dt} = 0.1 \]

Global Sensitivity: \( E \) is more influential
Global Sensitivity Analysis

Example: Linear elastic constitutive relation

\[ \sigma = E e + c \frac{de}{dt} \]

Nominal Values: \( E = 100, \ c = 0.1 \)

Uncertainty: 10% of nominal values

\( E \sim \mathcal{U}(90, 110), \ c \sim \mathcal{U}(0.09, 0.11) \)

Statistical Motivation: Consider variability of expected values \( D_i = \text{var} [\mathbb{E}(Y|q_i)] \)
Variance-Based Methods

**Sobol Representation:** For now, take \( Q_i \sim \mathcal{U}(0,1) \) and \( \Gamma = [0,1]^p \)

Take

\[
f(q) = f_0 + \sum_{i=1}^{p} f_i(q_i) + \sum_{1 \leq i < j \leq p} f_{ij}(q_i, q_j)
\]

**Analogy:** Taylor or Fourier series

With appropriate assumptions,

\[
f_0 = \int_{\Gamma} f(q) \, dq
\]

\[
f_i(q_i) = \int_{\Gamma^{p-1}} f(q) \, dq_{\sim i} - f_0
\]
Variance-Based Methods

Sobol Representation: For now, take $Q_i \sim \mathcal{U}(0, 1)$ and $\Gamma = [0, 1]^p$

Take

$$f(q) = f_0 + \sum_{i=1}^{p} f_i(q_i) + \sum_{1 \leq i < j \leq p} f_{ij}(q_i, q_j)$$

With appropriate assumptions,

$$f_0 = \int_{\Gamma} f(q) \, dq$$
$$f_i(q_i) = \int_{\Gamma^{p-1}} f(q) \, dq_{\sim i} - f_0$$

Variances:

$$D_i = \int_{0}^{1} f_i^2(q_i) \, dq_i$$
$$D = \text{var}(Y)$$

Sobol Indices: $S_i = \frac{D_i}{D}$

Analogy: Taylor or Fourier series

Assumption: Mutually independent parameters

Statistical Interpretation:

$$D_i = \text{var}[\mathbb{E}(Y|q_i)] \Rightarrow S_i = \frac{\text{var}[\mathbb{E}(Y|q_i)]}{\text{var}(Y)}$$
Global Sensitivity Analysis: Morris Screening

Example: Consider independent uniformly distributed parameters on $\Gamma = [0, 1]^p$

Elementary Effect:

$$d_i^j = \frac{f(q_i^j + \Delta e_i) - f(q_i^j)}{\Delta}$$, $i^{th}$ parameter, $j^{th}$ sample

Global Sensitivity Measures: $r$ samples

$$\mu_i^* = \frac{1}{r} \sum_{j=1}^{r} |d_i^j(q)|$$

$$\sigma_i^2 = \frac{1}{r-1} \sum_{j=1}^{r} (d_i^j(q) - \mu_i)^2$$, $\mu_i = \frac{1}{r} \sum_{j=1}^{r} d_i^j(q)$
Parameter Subset Selection: Materials

Example: Polydomain structure – Lead Titanate

\[ u(P_i, \varepsilon_{ij}, P_{i,j}) = u_M(\varepsilon_{ij}) + u_L(P_i) + u_C(P_i, \varepsilon_{ij}) + u_G(P_{i,j}) \]

where

\[ u_L(P_3) = \alpha_1 P_3^2 + \alpha_{11} P_3^4 + \alpha_{111} P_3^6 \]

\[ u_C(P_3, \varepsilon_{ii}) = -q_{11} \varepsilon_{11} P_3^2 - q_{12} (\varepsilon_{11} P_3^2 + \varepsilon_{22} P_3^2) \]

\[ u_G(P_{3,1}) = \frac{1}{2} g_{44} P_{3,1}^2 \]

Domain Wall Energy:

\[ E_{180}(q_{180}) = \int_{-\infty}^{\infty} (u - u_0) \, dx \]

\[ q_{180} = [\alpha_1, \alpha_{11}, \alpha_{111}, q_{11}, q_{12}, g_{44}] \]

Result: Linear-algebra based techniques

- Only \( \alpha_{11}, q_{11}, g_{44} \) influential and can be inferred
- Prior distributions for \( \alpha_1, \alpha_{111}, q_{12} \) not informed by data
Global Sensitivity Analysis

**Example:** Quantum-informed continuum model

Question: Do we use 4\textsuperscript{th} or 6\textsuperscript{th}-order Landau energy?

\[ \psi(P, q) = \alpha_1 P^2 + \alpha_{11} P^4 + \alpha_{111} P^6 \]

Parameters:

\[ q = [\alpha_1, \alpha_{11}, \alpha_{111}] \]

Global Sensitivity Analysis:

<table>
<thead>
<tr>
<th></th>
<th>( \alpha_1 )</th>
<th>( \alpha_{11} )</th>
<th>( \alpha_{111} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S_j )</td>
<td>0.62</td>
<td>0.39</td>
<td>0.01</td>
</tr>
<tr>
<td>( S_{Ti} )</td>
<td>0.66</td>
<td>0.38</td>
<td>0.06</td>
</tr>
<tr>
<td>( \mu^*_i )</td>
<td>0.17</td>
<td>0.07</td>
<td>0.03</td>
</tr>
</tbody>
</table>

**Conclusion:** \( \alpha_{111} \) insignificant and can be fixed
Global Sensitivity Analysis

Example: Quantum-informed continuum model

Question: Do we use 4\textsuperscript{th} or 6\textsuperscript{th}-order Landau energy?

\[ \psi(P, q) = \alpha_1 P^2 + \alpha_{11} P^4 + \alpha_{111} P^6 \]

Parameters:

\[ q = [\alpha_1, \alpha_{11}, \alpha_{111}] \]

Global Sensitivity Analysis:

<table>
<thead>
<tr>
<th></th>
<th>( \alpha_1 )</th>
<th>( \alpha_{11} )</th>
<th>( \alpha_{111} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S_j )</td>
<td>0.62</td>
<td>0.39</td>
<td>0.01</td>
</tr>
<tr>
<td>( S_{Ti} )</td>
<td>0.66</td>
<td>0.38</td>
<td>0.06</td>
</tr>
<tr>
<td>( \mu_i^x )</td>
<td>0.17</td>
<td>0.07</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Conclusion:

\( \alpha_{111} \) insignificant and can be fixed

Problem: We obtain different distributions when we perform Bayesian inference with fixed non-influential parameters.
Global Sensitivity Analysis

Example: Quantum-informed continuum model

Question: Do we use 4\textsuperscript{th} or 6\textsuperscript{th}-order Landau energy?

\[ \psi(P, q) = \alpha_1 P^2 + \alpha_{11} P^4 + \alpha_{111} P^6 \]

Parameters:

\[ q = [\alpha_1, \alpha_{11}, \alpha_{111}] \]

Global Sensitivity Analysis:

<table>
<thead>
<tr>
<th></th>
<th>( \alpha_1 )</th>
<th>( \alpha_{11} )</th>
<th>( \alpha_{111} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S_k )</td>
<td>0.62</td>
<td>0.39</td>
<td>0.01</td>
</tr>
<tr>
<td>( T_k )</td>
<td>0.66</td>
<td>0.38</td>
<td>0.06</td>
</tr>
<tr>
<td>( \mu_k^* )</td>
<td>0.17</td>
<td>0.07</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Note: Must accommodate correlation

Problem:

- Parameters correlated
- Cannot fix \( \alpha_{111} \)
One Solution: Parameter Subset Selection

Consider

\[ \psi(P_i, q) \approx \psi(P_i, q^*) + \nabla_q \psi(P_i, q^*) \Delta q \]

where

\[ \nabla_q \psi(P_i, q^*) = \left[ \frac{\partial \psi}{\partial \alpha_1}(P_i, q^*), \frac{\partial \psi}{\partial \alpha_11}(P_i, q^*), \frac{\partial \psi}{\partial \alpha_{111}}(P_i, q^*) \right] \]

**Functional:** Since \( \nu_i \approx \psi(P_i, q^*) \)

\[
J(q) = \frac{1}{n} \sum_{i=1}^{n} \left[ \nu_i - \psi(P_i, q) \right]^2 \\
\approx \frac{1}{n} \sum_{i=1}^{n} \left[ \nabla_q \psi(P_i, q^*) \cdot \Delta q \right]^2 \\
= \frac{1}{n} (\chi \Delta q)^T (\chi \Delta q)
\]

**Sensitivity Matrix:**

\[
\chi(q^*) = \begin{bmatrix}
\frac{\partial \psi}{\partial \alpha_1}(P_1, q^*) & \frac{\partial \psi}{\partial \alpha_{111}}(P_1, q^*) \\
\vdots & \ddots & \vdots \\
\frac{\partial \psi}{\partial \alpha_1}(P_n, q^*) & \frac{\partial \psi}{\partial \alpha_{111}}(P_n, q^*)
\end{bmatrix}
\]

**Note:**

\[ J(q^* + \Delta q) \approx \frac{1}{n} \Delta q^T \chi^T \chi \Delta q \]
One Solution: Parameter Subset Selection

Note:

\[ J(q^* + \Delta q) \approx \frac{1}{n} \Delta q^T \chi^T \chi \Delta q \]

Strategy: Take \( \Delta q \) to be eigenvector of \( \chi^T \chi \) Fisher Information

\[ \Rightarrow \chi^T \chi \Delta q = \lambda \Delta q \]
\[ \Rightarrow J(q^* + \Delta q) \approx \frac{\lambda}{n} \| \Delta q \|^2 \]

Note: \( \lambda \approx 0 \Rightarrow \) Perturbations \( J(q^* + \Delta q) \approx 0 \)
\[ \Rightarrow \text{Nonidentifiable} \]

Note: Estimator for covariance matrix

\[ V = s^2 [\chi^T \chi]^{-1} = \begin{bmatrix}
\text{var}(q_1) & \text{cov}(q_1, q_2) & \cdots & \text{cov}(q_1, q_n) \\
\text{cov}(q_2, q_1) & \text{var}(q_2) & \text{cov}(q_2, q_3) \\
\vdots & \vdots & \ddots & \vdots \\
\text{cov}(q_n, q_1) & \cdots & \cdots & \text{var}(q_n)
\end{bmatrix} \]
One Solution: Parameter Subset Selection

Note:

\[
J(q^* + \Delta q) \approx \frac{1}{n} \Delta q^T \chi^T \chi \Delta q
\]

Strategy: Take \( \Delta q \) to be eigenvector of \( \chi^T \chi \) Fisher Information

\[
\Rightarrow \chi^T \chi \Delta q = \lambda \Delta q
\]

\[
\Rightarrow J(q^* + \Delta q) \approx \frac{\lambda}{n} \| \Delta q \|_2^2
\]

\( \lambda \approx 0 \) \Rightarrow Perturbations \( J(q^* + \Delta q) \approx 0 \)

\Rightarrow Nonidentifiable

Example:

\[
\psi(P, q) = \alpha_1 P^2 + \alpha_{11} P^4 + \alpha_{111} P^6
\]

Parameters:

\[
q = [\alpha_1, \alpha_{11}, \alpha_{111}]
\]

Result: \( \text{rank}(\chi^T \chi) = 3 \) so all parameters identifiable
Case Study: Spring Model

**Example:** Spring model

\[ m \frac{d^2 z}{dt^2} + kz = 0 \]

\[ z(0) = 1, \quad \frac{dz}{dt}(0) = 0 \]

Responses: For \( q = [k,m] \), consider

\[ y = f(q) = \cos \left( \sqrt{\frac{k}{m}} \cdot \frac{\pi}{2} \right) \]

\[ y = \int_0^{\pi/2} \cos \left( \sqrt{\frac{k}{m}} t \right) dt = \sqrt{\frac{m}{k}} \]

**Exercise:** Download MATLAB software from the website

https://rsmith.math.ncsu.edu/SPIE_SHORT_COURSE/

and run the codes spring_morris.m and spring_Saltelli.m. What do you conclude about the parameter sensitivity?
Steps in Uncertainty Quantification

**Challenge 2:**
- How do we do uncertainty quantification for computationally expensive models?
- Example:
  - We have a computational budget of 5000 model evaluations.
  - Bayesian inference and uncertainty propagation require 120,000 evaluations.
Uncertainty Quantification Challenges

**Example:** MFC model – Fourth-order PDE

\[ \rho \frac{\partial^2 w}{\partial t^2} + \gamma \frac{\partial w}{\partial t} - \frac{\partial^2 M}{\partial x^2} = 0 \]

\[ M = -c^E l \frac{\partial^2 w}{\partial x^2} - c_D l \frac{\partial^3 w}{\partial x^2 \partial t} - [k_1 e(E, \sigma_0) E + k_2 \varepsilon_{irr}(E, \sigma_0)] \chi_{MFC}(x) \]

**Bayesian Inference:** 20 parameters -- Took 6 days!

**Problem:**

\[ 1.2 \times 10^5 \text{ PDE solutions} \]

**Solution:** Highly efficient surrogate models
Surrogate Models: Motivation

**Example:** Consider the heat equation

\[
\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} + f(q)
\]

Boundary Conditions
Initial Conditions
with the response

\[
y(q) = \int_0^1 \int_0^1 \int_0^1 \int_0^1 u(t, x, y, z) \, dx \, dy \, dz \, dt
\]

**Notes:**
- Requires approximation of PDE in 3-D
- What would be a simple surrogate?
Surrogate Models

Example: Consider the heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} + f(q)$$

Boundary Conditions

Initial Conditions

with the response

$$y(q) = \int_0^1 \int_0^1 \int_0^1 \int_0^1 u(t, x, y, z) dx dy dz dt$$

Surrogate: Quadratic

$$y_s(q) = (q - 0.25)^2 + 0.5$$

Question: How do you construct a polynomial surrogate?

- Regression
- Interpolation
Surrogate Models

Recall: Consider the model

\[
\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} + f(q)
\]

Boundary Conditions
Initial Conditions

with the response

\[
y(q) = \int_0^1 \int_0^1 \int_0^1 \int_0^1 u(t, x, y, z) \, dx \, dy \, dz \, dt
\]

Question: How do you construct a polynomial surrogate?

- Interpolation
- Regression

Surrogate: Quadratic

\[
y_s(q) = (q - 0.25)^2 + 0.5
\]
Data-Fit Models

Notes:
- Often termed response surface models, emulators, meta-models.
- Constructed via interpolation or regression.
- Data can consist of high-fidelity simulations or experiments.

Example: Steady-state Euler-Bernouilli beam model with PZT patch

\[ Yl \frac{d^4 w}{dx^4} (x) = k_p V \chi_{pzt}(x) \]

Data: Displacement observations
Parameter: \( Yl \)

Simulations: Nikolas Bravo
Data-Fit Models

**Example:** Steady-state Euler-Bernouilli beam model with PZT patch

\[ Y_1 \frac{d^4 w}{dx^4}(x) = k_p V_{\chi_{\text{pzt}}}(x) \]

Data: Displacement observations
Parameter: Y1
Training points: 5000
Polynomial surrogate: 6\textsuperscript{th} order

Bayesian Inference
Data-Fit Models

Notes:
- Often termed response surface models, surrogates, emulators, meta-models.
- Rely on interpolation or regression.
- Data can consist of high-fidelity simulations or experiments.
- Common techniques: polynomial models, kriging (Gaussian process regression), orthogonal polynomials.

Strategy: Consider high fidelity model

\[ y = f(q) \]

with M model evaluations

\[ y_m = f(q^m) , \ m = 1, \ldots, M \]

Statistical Model: \( f_s(q) \): Surrogate for \( f(q) \)

\[ y_m = f_s(q^m) + \epsilon_m , \ m = 1, \ldots, M \]

Surrogate:

\[ y^K(Q) = f_s(Q) = \sum_{k=0}^{K} \alpha_k \psi_k(Q) \]

Note: \( \psi_k(Q) \) orthogonal with respect to inner product associated with pdf

e.g., \( Q \sim N(0, 1) \): Hermite polynomials
\( Q \sim U(-1, 1) \): Legendre polynomials
Orthogonal Polynomial Representations

**Representation:**

\[ y^K(Q) = \sum_{k=0}^{K} \alpha_k \psi_k(Q) \]

**Note:** \( \psi_0(Q) = 1 \) implies that

\[ \mathbb{E}[\psi_0(Q)] = 1 \]

\[ \mathbb{E}[\psi_i(Q)\psi_j(Q)] = \int_{\Gamma} \psi_i(q)\psi_j(q)\rho(q) dq \]

\[ = \delta_{ij} \gamma_i \]

where \( \gamma_i = \mathbb{E}[\psi_i^2(Q)] \)

**Properties:**

(i) \( \mathbb{E}[y^K(Q)] = \alpha_0 \)

(ii) \( \text{var}[y^K(Q)] = \sum_{k=1}^{K} \alpha_k^2 \gamma_k \)

**Note:** Can be used for:

- Uncertainty propagation
- Sobol-based global sensitivity analysis

**Issue:** How does one compute \( \alpha_k, k = 0, \ldots, K \)?

- Stochastic Galerkin techniques (Polynomial Chaos Expansion – PCE)
- Nonintrusive PCE (Discrete projection)
- Stochastic collocation
- Regression-based methods with sparsity control (Lasso)

**Note:** Methods nonintrusive and treat code as blackbox.
Orthogonal Polynomial Representations

**Nonintrusive PCE:** Take weighted inner product of \( y(q) = \sum_{k=0}^{\infty} \alpha_k \psi_k(q) \) to obtain

\[
\alpha_k = \frac{1}{\gamma_k} \int_{\Gamma} y(q) \psi_k(q) \rho(q) dq
\]

**Quadrature:**

\[
\alpha_k \approx \frac{1}{\gamma_k} \sum_{r=1}^{R} y(q^r) \psi_k(q^r) w^r
\]

**Note:**

(i) Low-dimensional: Tensored 1-D quadrature rules – e.g., Gaussian

(ii) Moderate-dimensional: Sparse grid (Smolyak) techniques

(iii) High-dimensional: Monte Carlo or quasi-Monte Carlo (QMC) techniques

**Regression-Based Methods with Sparsity Control (Lasso):** Solve

\[
\min_{\alpha \in \mathbb{R}^{K+1}} \|\Lambda \alpha - d\|^2 \quad \text{subject to} \quad \sum_{k=0}^{K} |\alpha_k| \leq \tau
\]

**Note:** Sample points \( \{q^m\}_{m=1}^{M} \)

\[
\Lambda \in \mathbb{R}^{M \times (K+1)} \quad \text{where} \quad \Lambda_{jk} = \psi_k(q^j)
\]

\[
d = [y(q^1), \ldots, y(q^m)]
\]

**e.g., SPGL1**

- MATLAB Solver for large-scale sparse reconstruction
Surrogate Models – Grid Choice

**Example:** Consider the Runge function \( f(q) = \frac{1}{1 + 25q^2} \) with points

\[
q^j = -1 + (j - 1) \frac{2}{M}, \quad j = 1, \ldots, M
\]
Surrogate Models – Grid Choice

**Example:** Consider the Runge function \( f(q) = \frac{1}{1 + 25q^2} \) with points

\[
q^j = -1 + (j - 1) \frac{2}{M}, \quad j = 1, \ldots, M
\]
Surrogate Models – Grid Choice

**Example:** Consider the Runge function \( f(q) = \frac{1}{1 + 25q^2} \) with points

\[
q^j = -1 + (j - 1) \frac{2}{M}, \quad j = 1, \ldots, M
\]

\[
q^j = -\cos \frac{\pi(j - 1)}{M - 1}, \quad j = 1, \ldots, M
\]
Surrogate Models – Grid Choice

**Example:** Consider the Runge function $f(q) = \frac{1}{1 + 25q^2}$ with points

$$q^j = -1 + (j - 1) \frac{2}{M}, \ j = 1, \ldots, M$$

$$q^j = -\cos \frac{\pi (j - 1)}{M - 1}, \ j = 1, \ldots, M$$
Sparse Grid Techniques

**Tensored Grids:** Exponential growth

Sparse Grids: Same accuracy

<table>
<thead>
<tr>
<th>$p$</th>
<th>$R_\ell$</th>
<th>$\mathcal{R}$ (Sparse Grid)</th>
<th>$R = (R_\ell)^p$ (Tensored Grid)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>9</td>
<td>29</td>
<td>81</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>241</td>
<td>59,049</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>1581</td>
<td>$&gt; 3 \times 10^9$</td>
</tr>
<tr>
<td>50</td>
<td>9</td>
<td>171,901</td>
<td>$&gt; 5 \times 10^{47}$</td>
</tr>
<tr>
<td>100</td>
<td>9</td>
<td>1,353,801</td>
<td>$&gt; 2 \times 10^{95}$</td>
</tr>
</tbody>
</table>
Surrogate Models

**Question:** How do we keep from fitting noise?

- Akaike Information Criterion (AIC)
  \[
  AIC = 2k - 2 \log[\pi(y|q)]
  \]
- Bayesian Information Criterion (BIC)
  \[
  BIC = k \log(M) - 2 \log[\pi(y|q)]
  \]

Likelihood:

\[
\pi(y|q) = \frac{1}{(2\pi\sigma^2)^{M/2}} e^{-SS_q/2\sigma^2}
\]

\[
SS_q = \sum_{m=1}^{M} [y_m - y_s(q^m)]^2
\]

**Example:** \( y = \exp(0.7q_1 + 0.3q_2) \)

**Exercise:**

- Construct a polynomial surrogate using the code `response_surface.m`.
- What order seems appropriate?
Concluding Remarks

Notes:

• UQ requires a synergy between engineering, statistics, and applied mathematics.

• Model calibration, model selection, uncertainty propagation and experimental design are natural in a Bayesian framework.

• Goal is to predict model responses with quantified and reduced uncertainties.

• Parameter selection is critical to isolate identifiable and influential parameters.

• Surrogate models critical for computationally intensive simulation codes.

• Codes and packages: Sandia Dakota, R, MATLAB, Python, nanoHUB.

• Algorithms are new and evolving.

• Prediction is very difficult, especially if it’s about the future, Niels Bohr.
Books:


Papers:


