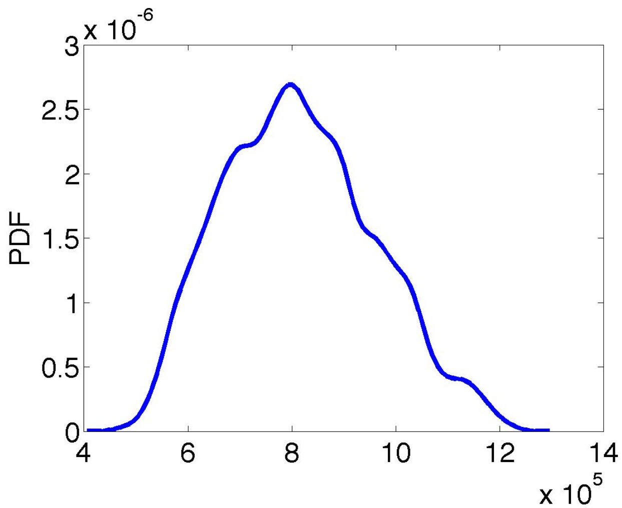
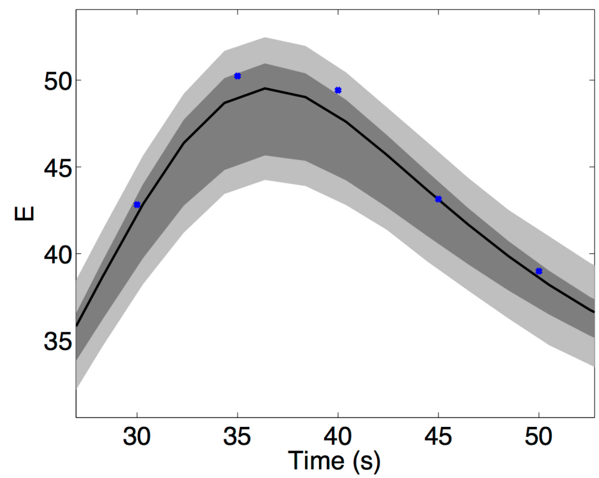
Uncertainty Quantification and Sensitivity Analysis

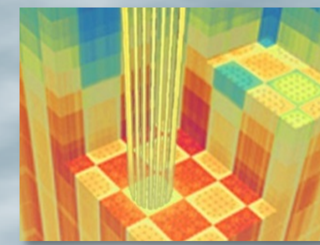
Ralph C. Smith, Department of Mathematics,
North Carolina State University

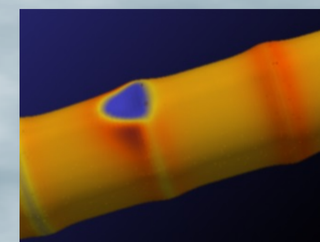


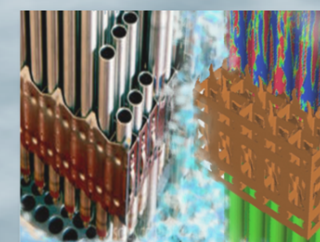


Essentially, all models are wrong, but some are useful, George E.P. Box, Industrial Statistician.

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





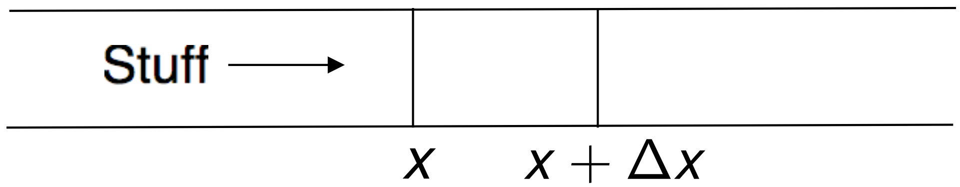


CASL The Consortium for Advanced Simulation of LWRs
A DOE Energy Innovation Hub

U.S. DEPARTMENT OF ENERGY

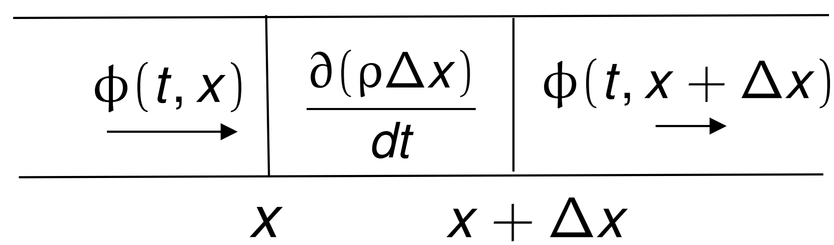
Modeling Strategy

General Strategy: Conservation of stuff



$$\boxed{\frac{dStuff}{dt} = \text{Stuff in - Stuff out + Stuff created - Stuff destroyed}}$$

Continuity Equation:



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

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

$$\boxed{\Rightarrow \frac{\partial \rho}{\partial t} + \frac{\partial \phi}{\partial x} = 0}$$

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

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

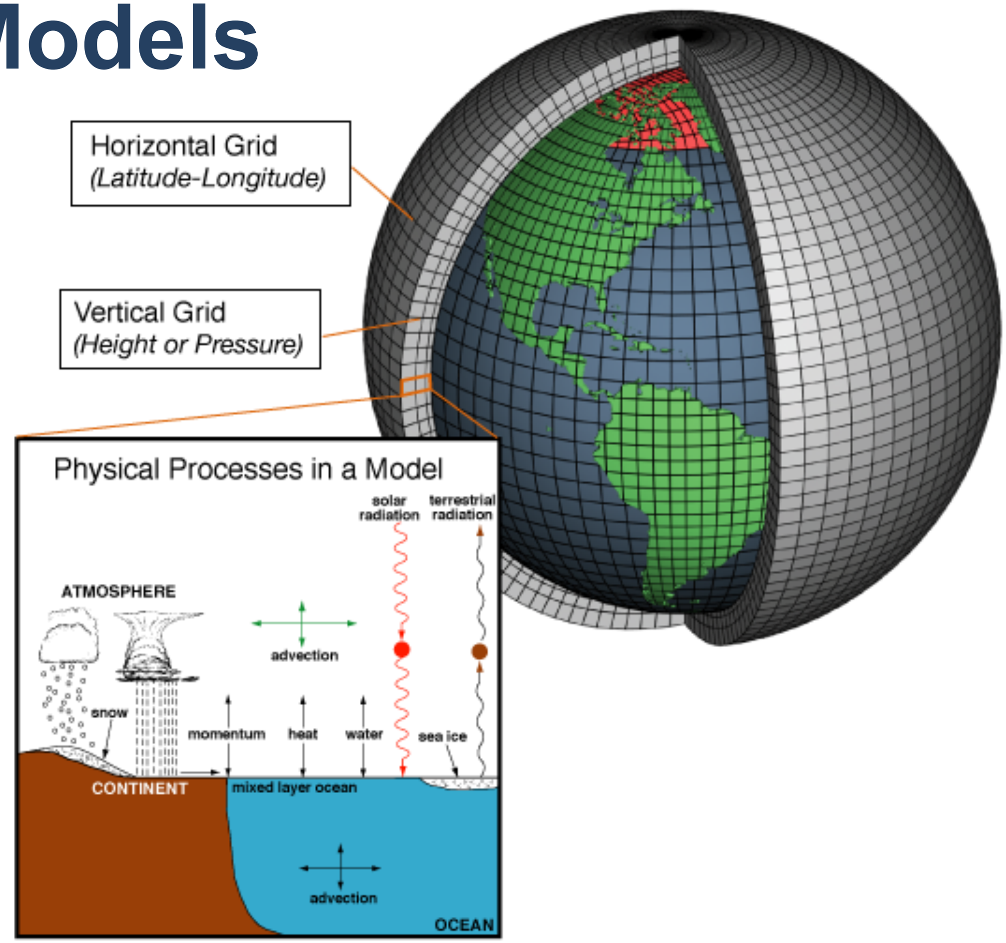
More Generally:

$$\boxed{\Rightarrow \frac{\partial \rho}{\partial t} + \frac{\partial \phi}{\partial x} = \text{Sources - Sinks}}$$

Example 1: Weather Models

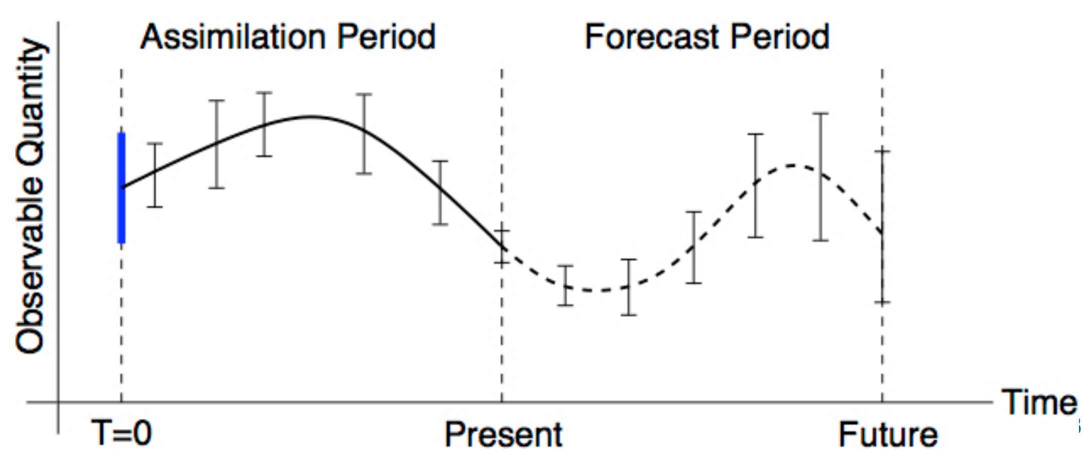
Challenges:

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



Goal:

- Assimilate data to quantify uncertain initial conditions and parameters
- Make predictions with quantified uncertainties.



Equations of Atmospheric Physics

Conservation Relations:

$$\boxed{\frac{\partial \rho}{\partial t} + \frac{\partial \phi}{\partial x} = \text{Sources - Sinks}}$$

Mass
$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho v) = 0$$

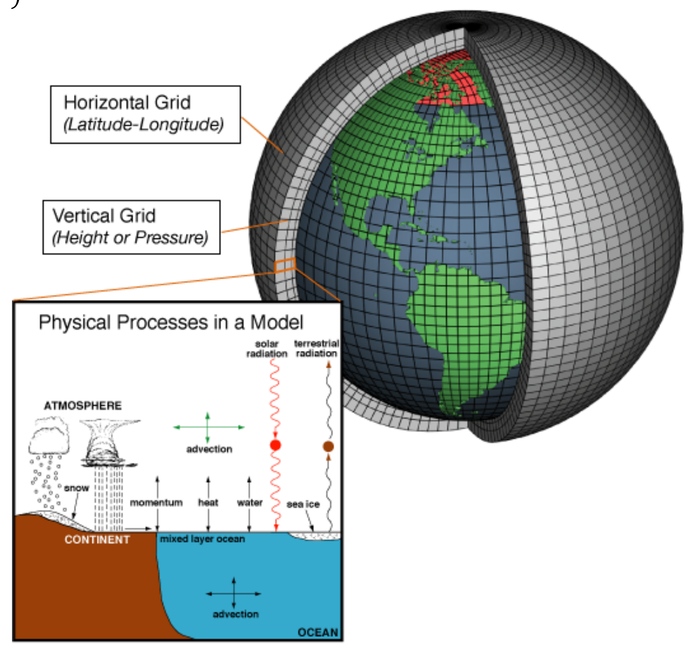
Momentum
$$\frac{\partial v}{\partial t} = -v \cdot \nabla v - \frac{1}{\rho}\nabla p - g\hat{k} - 2\Omega \times v$$

Energy
$$\rho c_V \frac{\partial T}{\partial t} + p\nabla \cdot v = -\nabla \cdot F + \nabla \cdot (k\nabla T) + \rho\dot{q}(T, p, \rho)$$

$$p = \rho R T$$

Water
$$\frac{\partial m_j}{\partial t} = -v \cdot \nabla m_j + S_{m_j}(T, m_j, \chi_j, \rho)\ ,\ j = 1, 2, 3,$$

Aerosol
$$\frac{\partial \chi_j}{\partial t} = -v \cdot \nabla \chi_j + S_{\chi_j}(T, \chi_j, \rho)\ ,\ j = 1, \cdots, J,$$



Constitutive Closure Relations: e.g.,

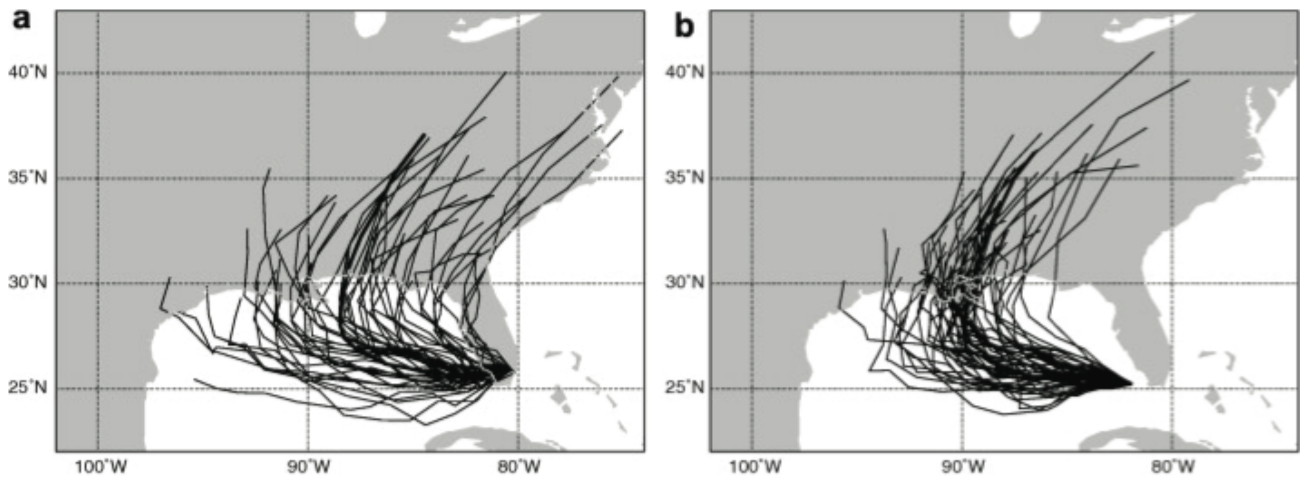
$$S_{m_2} = S_1 + S_2 + S_3 - S_4$$

where

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

Ensemble Predictions

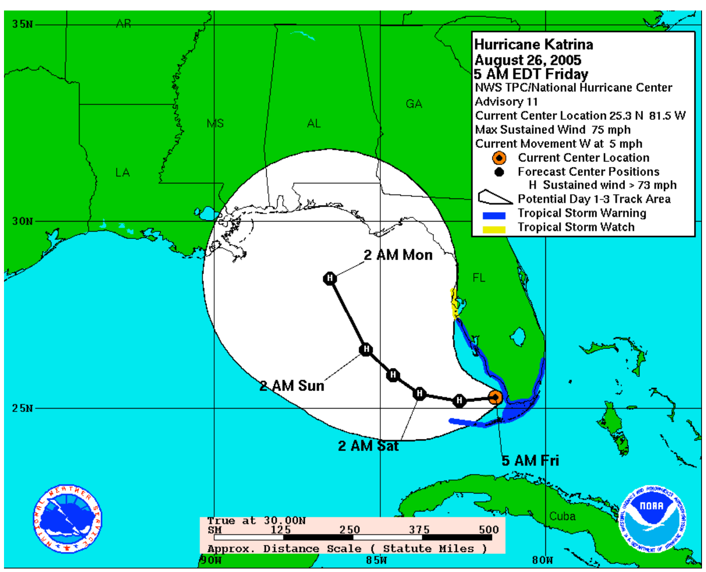
Ensemble Predictions:



00 UTC on August 26, 2005

12 UTC on August 26, 2005

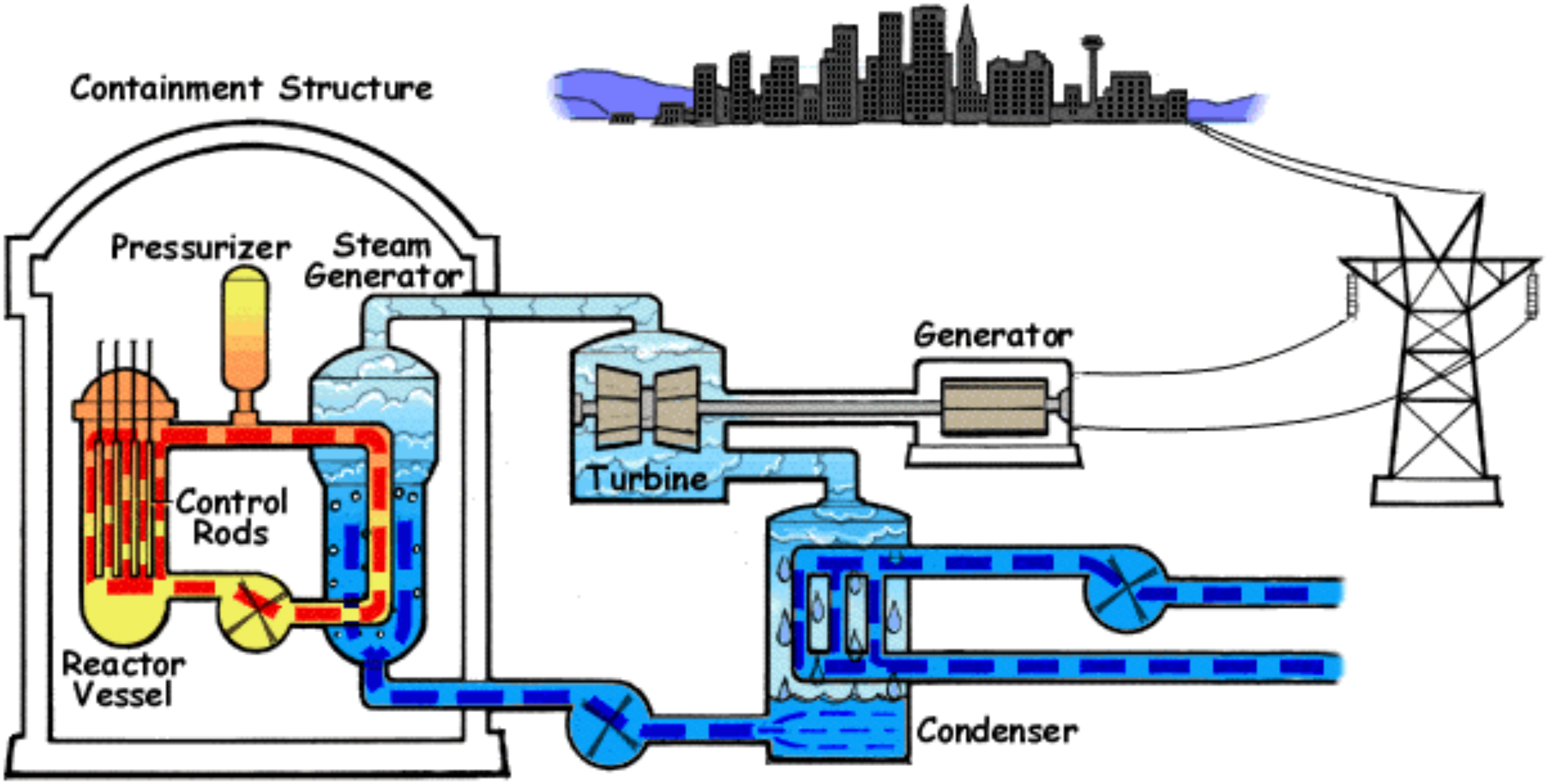
Cone of Uncertainty:



General Questions:

- What is expected rainfall on August 13?
- What is average high temperature?
- Note: Quantities are statistical in nature.

Example 2: Pressurized Water Reactors (PWR)



Models:

- Involve neutron transport, thermal-hydraulics, chemistry, fuels
- Inherently multi-scale, multi-physics.

Example: PWR

3-D Neutron Transport Equations:

$$\frac{1}{|v|}\frac{\partial \varphi}{\partial t} + \Omega \cdot \nabla \varphi + \Sigma_t(r,E)\varphi(r,E,\Omega,t)$$

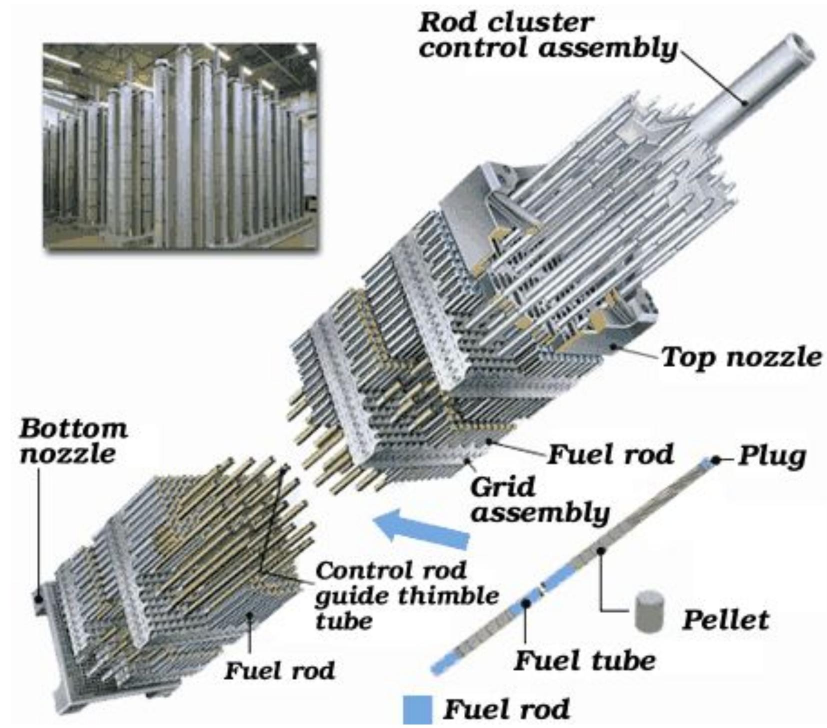
$$= \int_{4\pi} d\Omega' \int_0^\infty dE' \Sigma_s(E' \to E, \Omega' \to \Omega)\varphi(r,E',\Omega',t)$$

$$+ \underline{\frac{\chi(E)}{4\pi}} \int_{4\pi} d\Omega' \int_0^\infty dE' \underline{\nu(E')}\,\underline{\Sigma_f(E')}\varphi(r,E',\Omega',t)$$

$$\boxed{\frac{\partial \rho}{\partial t} + \frac{\partial \phi}{\partial x} = \text{Sources - Sinks}}$$

Challenges:

- Very large number of inputs; e.g., 100,000; Active subspace construction critical.
- ORNL Code SCALE: Can take hours to run.
- Numerical errors often difficult to quantify.
- Predicting future requires extrapolatory or out-of-data predictions.



Example: PWR

Thermo-Hydraulic Equations: Mass, momentum and energy balance for fluid

$$\frac{\partial}{\partial t}(\alpha_f \rho_f) + \nabla \cdot (\alpha_f \rho_f v_f) = -\Gamma$$

$$\alpha_f \rho_f \frac{\partial v_f}{\partial t} + \alpha_f \rho_f v_f \cdot \nabla v_f + \nabla \cdot \sigma_f^R + \alpha_f \nabla \cdot \sigma + \alpha_f \nabla p_f$$
$$= -F^R - F + \Gamma(v_f - v_g)/2 + \alpha_f \rho_f g$$

$$\frac{\partial}{\partial t}(\alpha_f \rho_f e_f) + \nabla \cdot (\alpha_f \rho_f e_f v_f + Th) = (T_g - T_f)H + T_f \Delta_f$$
$$-T_g(H - \alpha_g \nabla \cdot h) + h \cdot \nabla T - \Gamma[e_f + T_f(s^* - s_f)]$$
$$-p_f \left(\frac{\partial \alpha_f}{\partial t} + \nabla \cdot (\alpha_f v_f) + \frac{\Gamma}{\rho_f} \right)$$

$$\boxed{\frac{\partial \rho}{\partial t} + \frac{\partial \phi}{\partial x} = \text{Sources - Sinks}}$$

Notes:

- Similar relations for gas and bubbly phases

Challenges:

- Codes can have 15-30 closure relations and up to 75 parameters.
- Codes and closure relations often ”borrowed” from other physical phenomena; e.g., single phase fluids, airflow over a car (CFD code STAR-CCM+)
- Calibration necessary and closure relations can conflict.

Example: PWR

Thermo-Hydraulic Equations: Mass, momentum and energy balance for fluid

$$\frac{\partial}{\partial t}(\alpha_f \rho_f) + \nabla \cdot (\alpha_f \rho_f v_f) = -\Gamma$$

$$\alpha_f \rho_f \frac{\partial v_f}{\partial t} + \alpha_f \rho_f v_f \cdot \nabla v_f + \nabla \cdot \sigma_f^R + \alpha_f \nabla \cdot \sigma + \alpha_f \nabla p_f$$
$$= -F^R - F + \Gamma(v_f - v_g)/2 + \alpha_f \rho_f g$$

$$\frac{\partial}{\partial t}(\alpha_f \rho_f e_f) + \nabla \cdot (\alpha_f \rho_f e_f v_f + Th) = (T_g - T_f)H + T_f \Delta_f$$
$$-T_g(H - \alpha_g \nabla \cdot h) + h \cdot \nabla T - \Gamma[e_f + T_f(s^* - s_f)]$$
$$-p_f \left(\frac{\partial \alpha_f}{\partial t} + \nabla \cdot (\alpha_f v_f) + \frac{\Gamma}{\rho_f} \right)$$

$$\boxed{\frac{\partial \rho}{\partial t} + \frac{\partial \phi}{\partial x} = \text{Sources - Sinks}}$$

Notes:

- Similar relations for gas and bubbly phases

Challenges:

- Codes can have 15-30 closure relations and up to 75 parameters.

Example: Dittus—Boelter Relation

$$Nu = 0.023 Re^{0.8} Pr^{0.4}$$

Nu: Nusselt number
Re: Reynolds number
Pr: Prandtl number

Example: PWR

Example: Shearon Harris outside Raleigh

UQ Questions:

- What is peak operating temperature?
- What is expected level of CRUD buildup?
- What is risk associated with operating regime?
- What is expected profit for new design?

Example 3: Quantum-Informed Continuum Models

Objectives:

- Employ density function theory (DFT) to construct/calibrate continuum energy relations.
  - e.g., Helmholtz energy

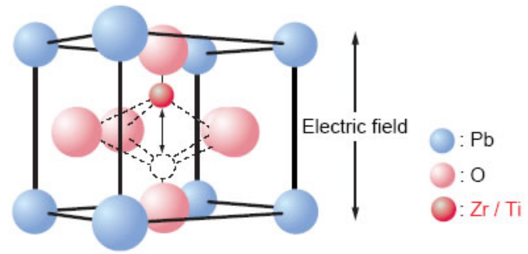
$$\psi(P) = \alpha_1 P^2 + \alpha_{11} P^4 + \alpha_{111} P^6$$

Note:

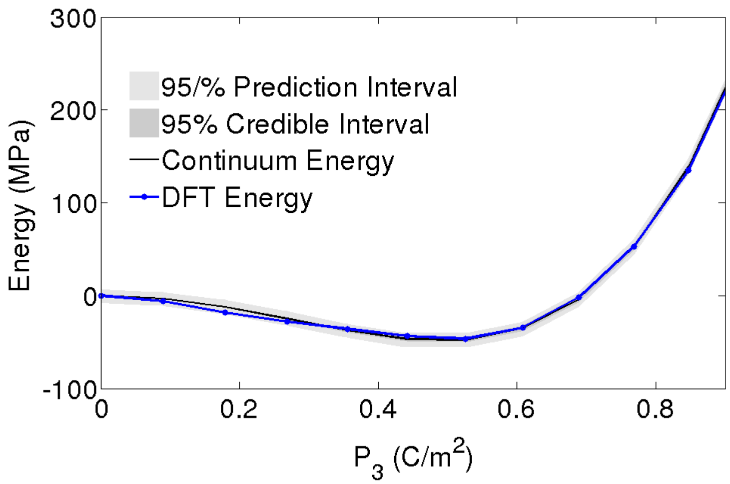
- Linearly parameterized

UQ and SA Issues:

- Is 6^{th} order term required to accurately characterize material behavior?
- **Note:** Determines molecular structure



Lead Titanate Zirconate (PZT)



Helmholtz Energy

Challenge: Terminology and Notation

Terminology:

- Inputs: Parameters, initial conditions, boundary conditions, exogenous forces; e.g., parameters in closure relations, initial conditions in transient models.
- Outputs or Responses: Quantities that we experimentally or numerically measure; e.g., 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.

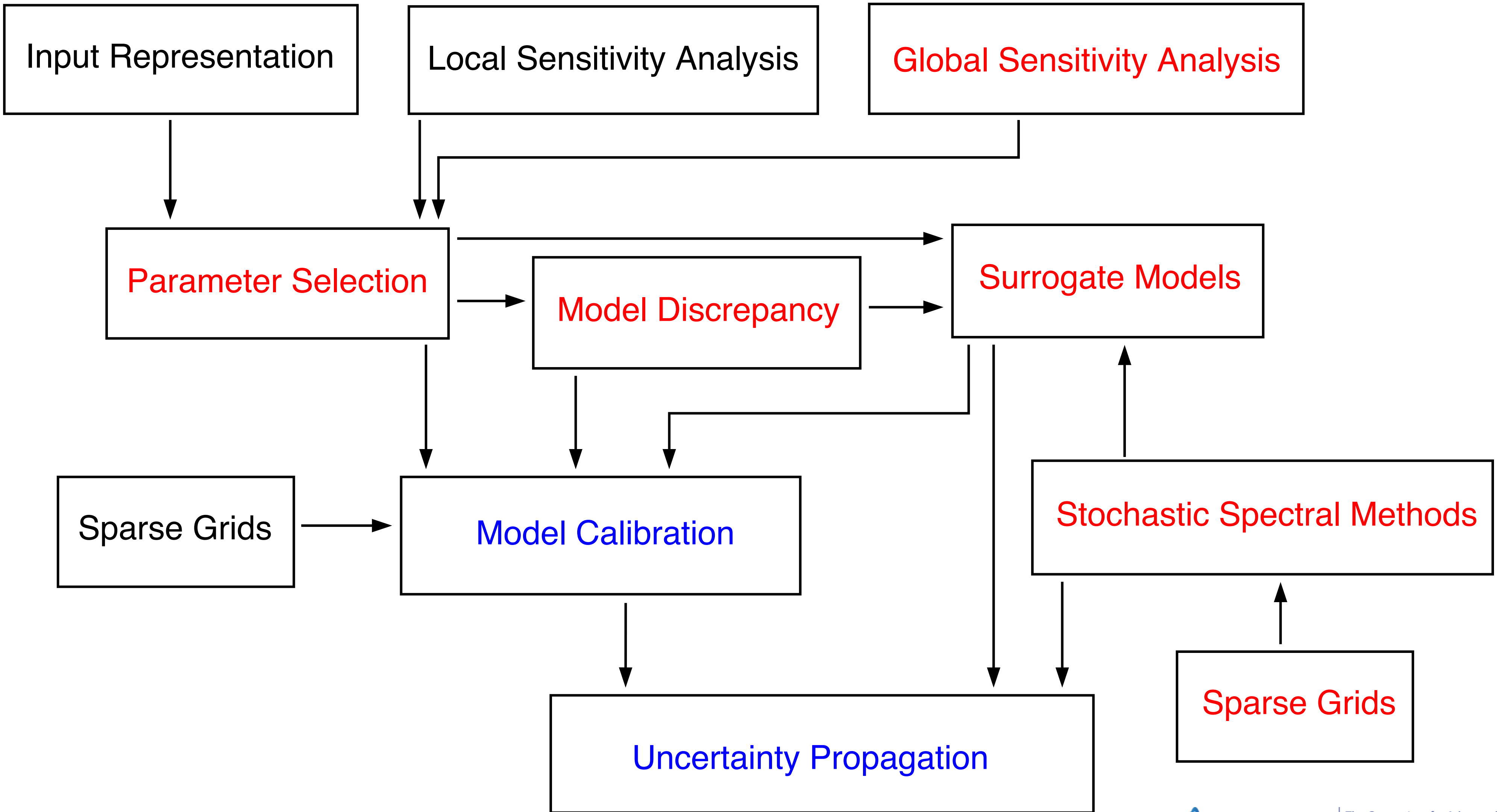
Input Notation: Can vary even within disciplines!

- 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

Steps in Uncertainty Quantification

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



Deterministic Model Calibration

Example: Helmholtz energy $\psi(P,q) = \underline{\alpha_1} P^2 + \underline{\alpha_{11}} P^4 + \underline{\alpha_{111}} P^6$

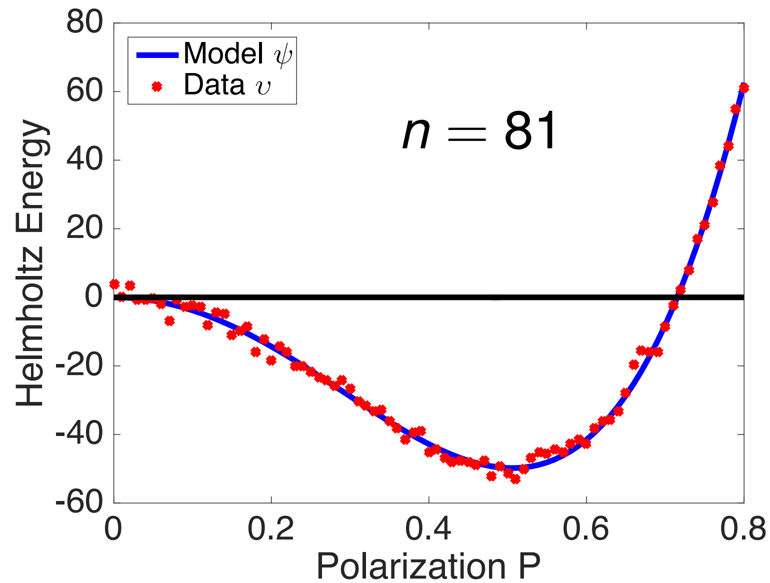
$$q = [\alpha_1, \alpha_{11}, \alpha_{111}]$$

Statistical Model: Describes observation process

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

Point Estimates: Ordinary least squares

$$q^0 = \arg\min_q \frac{1}{2} \sum_{j=1}^{n} [\upsilon_i - \psi(P_i, q)]^2$$



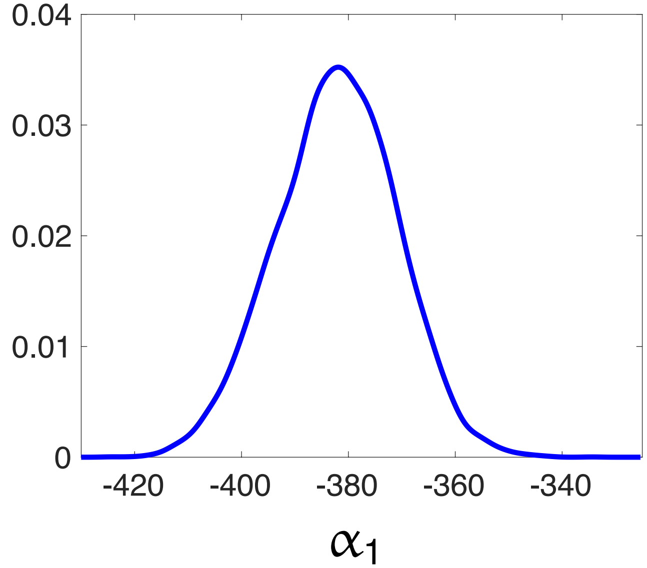
Note: Provides point estimates but no quantification of uncertainty in:

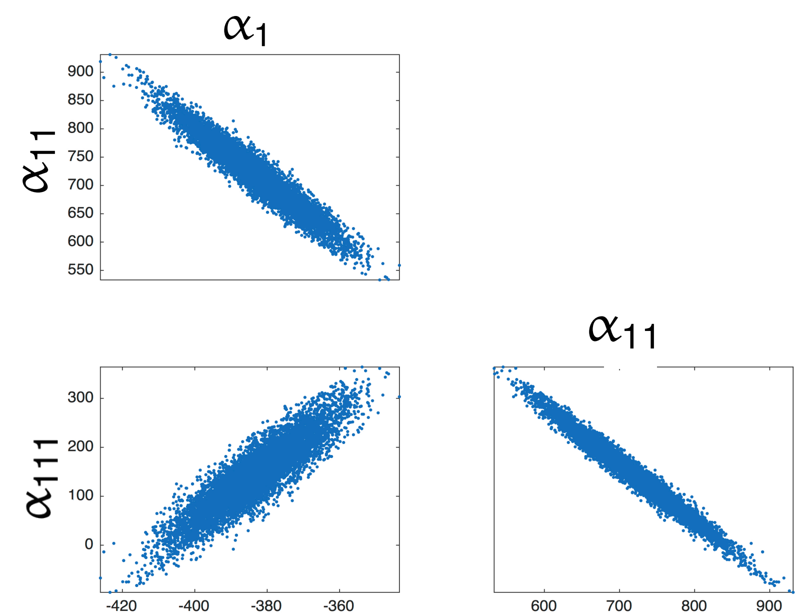
- Model
- Parameters
- Data

Objectives for Uncertainty Quantification

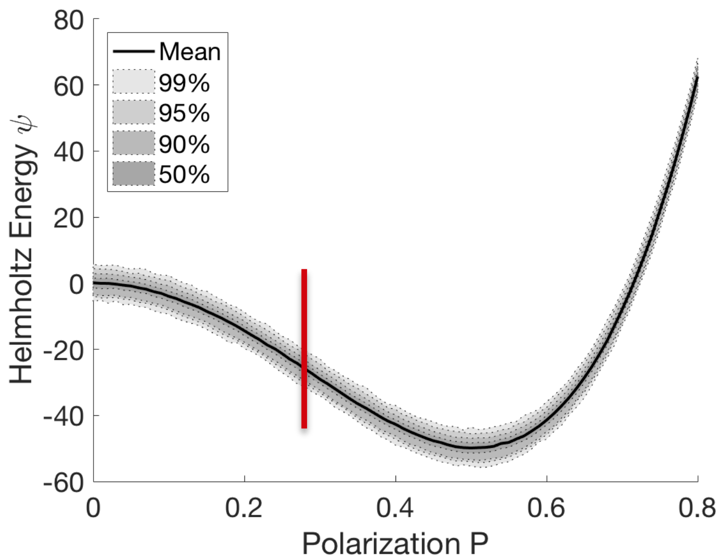
Goal: Replace point estimates with distributions or credible intervals

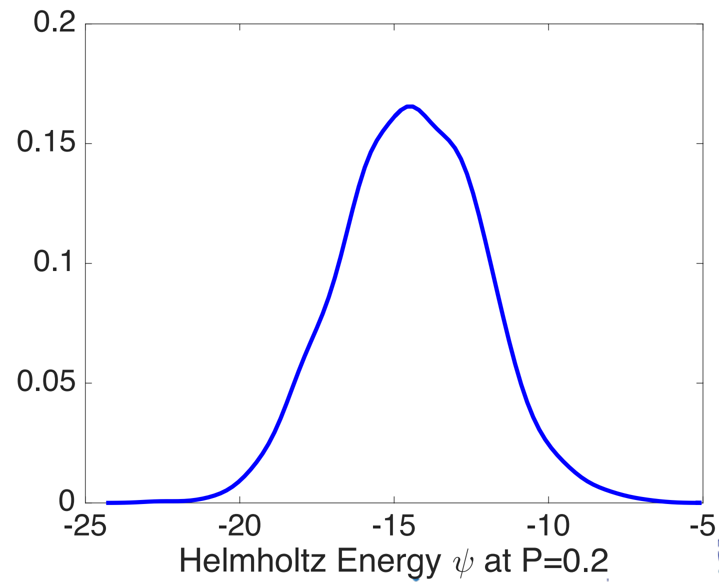
E.g., Parameter Densities





E.g., Response Intervals





Objectives for Uncertainty Quantification

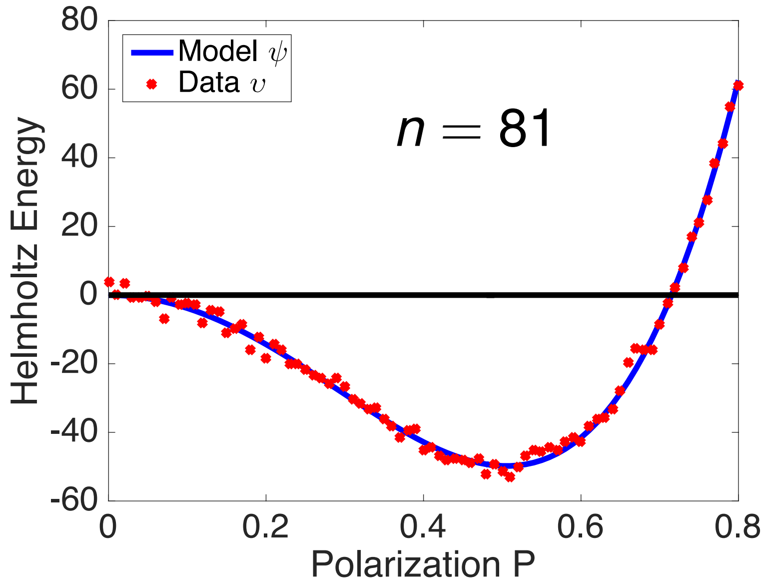
Example: Helmholtz energy $\psi(P, q) = \alpha_1 P^2 + \alpha_{11} P^4$, $q = [\alpha_1, \alpha_{11}]$

Statistical Model: Describes observation process

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

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

UQ Goals: Quantify parameter and response uncertainties

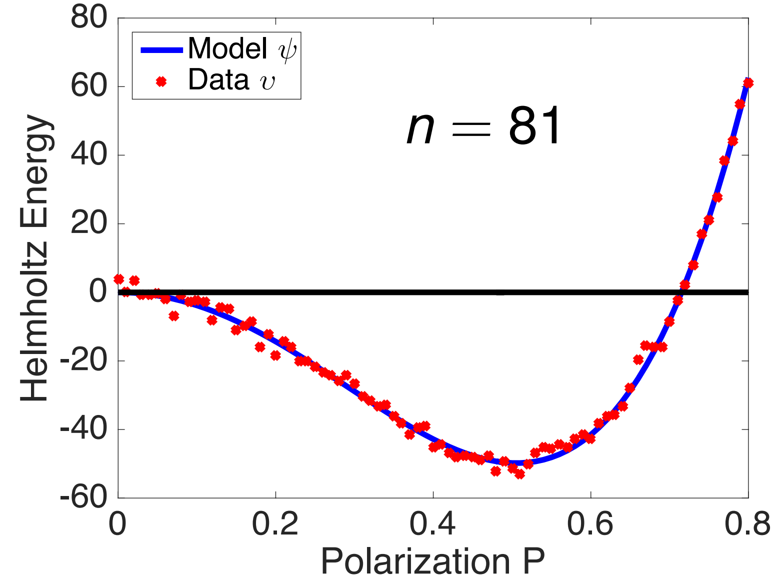


Strategy 1: Perform Experiments

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

Statistical Model: Describes observation process

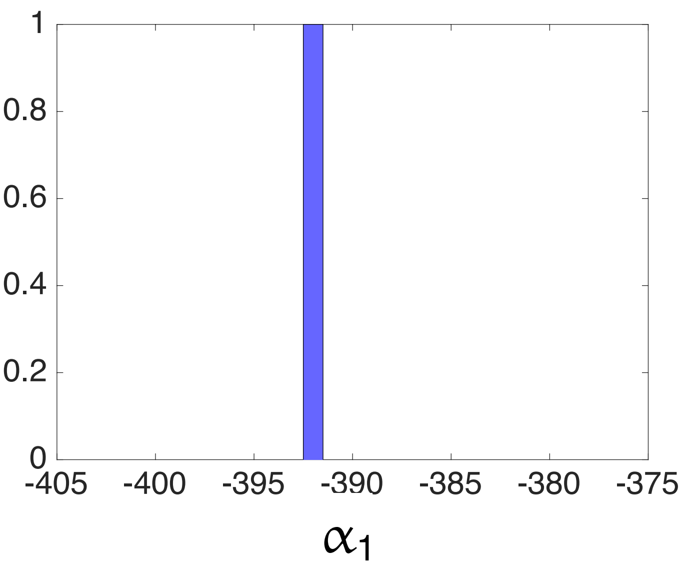
$$\upsilon_i = \psi(P_i, q) + \varepsilon_i \quad , \; i = 1, \ldots, n$$

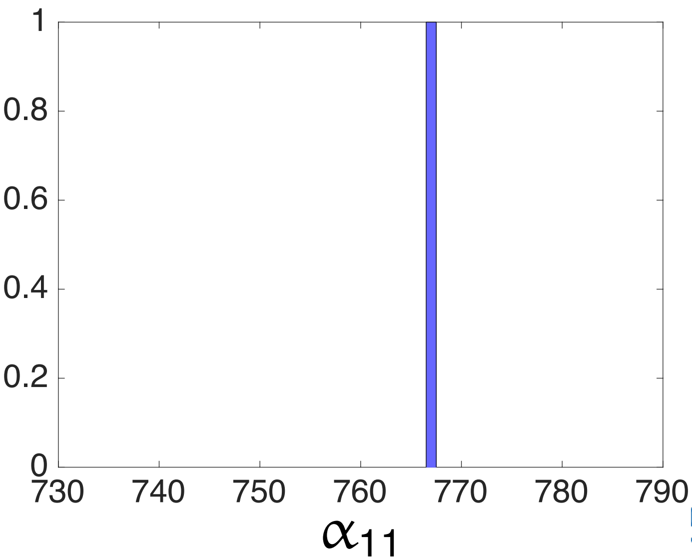


Common Assumption: $\varepsilon_i \sim N(0, \sigma^2)$

UQ Goals: Quantify parameter and response uncertainties

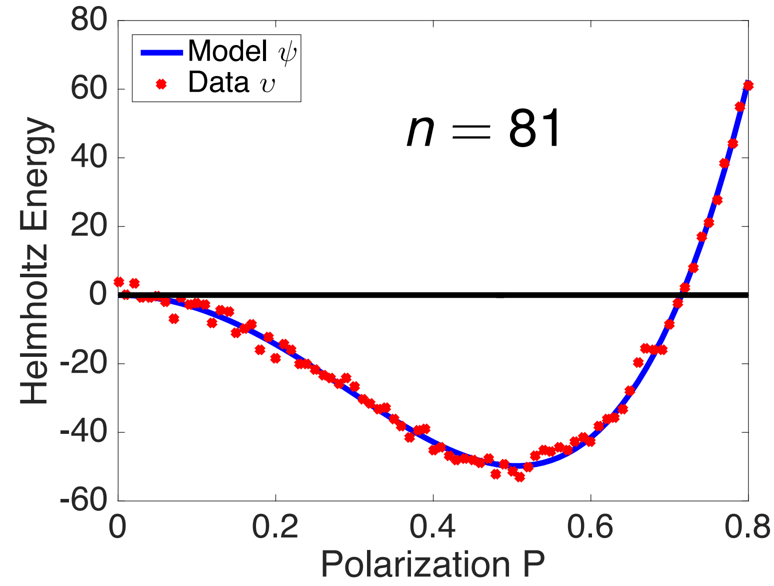
Strategy 1: Perform experiments; e.g., 1





Strategy 1: Perform Experiments

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



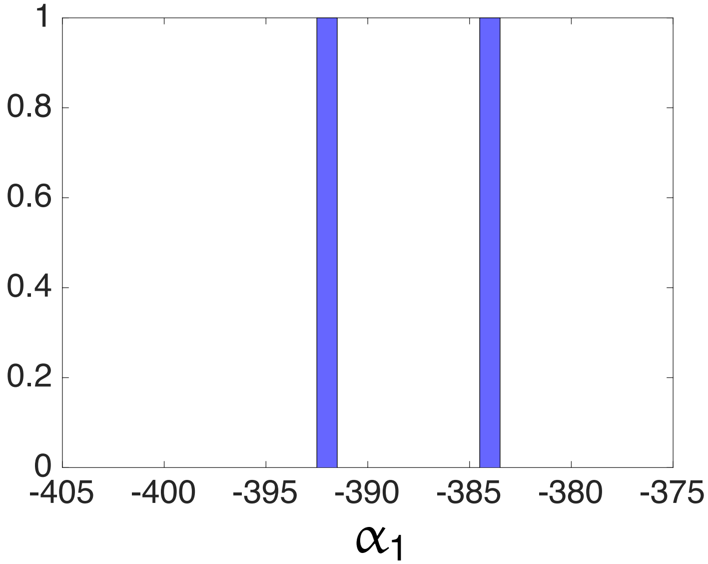
Statistical Model: Describes observation process

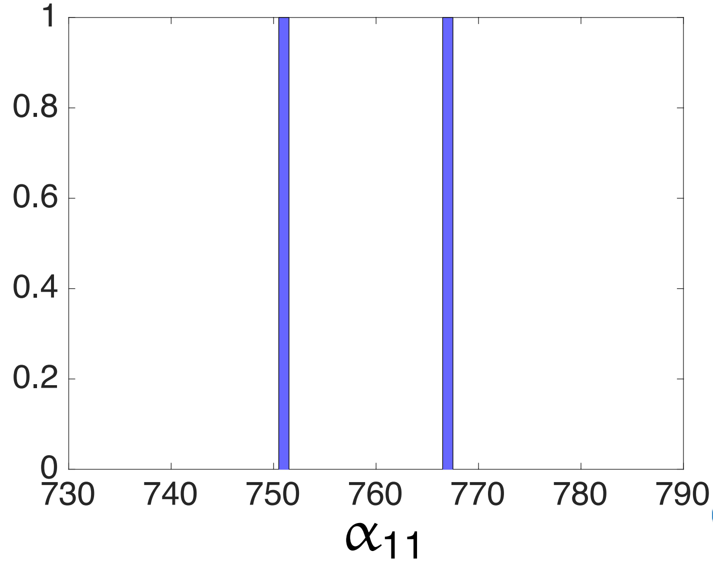
$$\upsilon_i = \psi(P_i, q) + \varepsilon_i \quad , \ i = 1, \dots, n$$

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

UQ Goals: Quantify parameter and response uncertainties

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



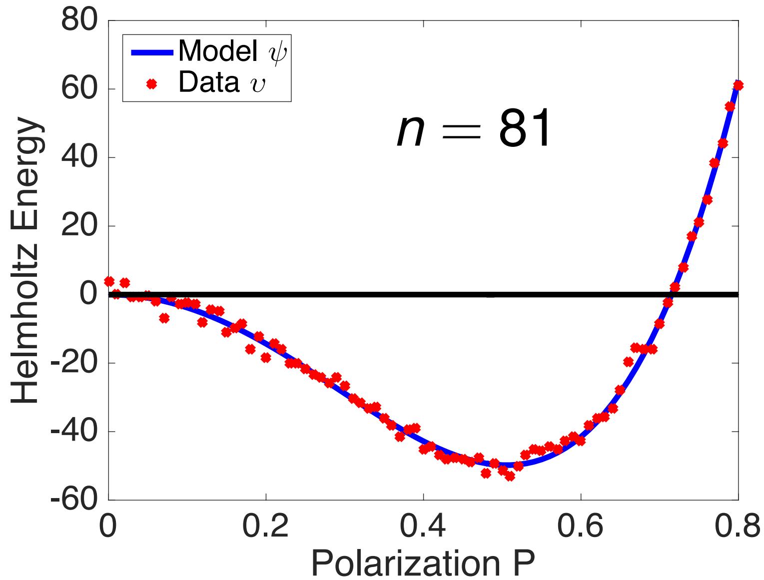


Strategy 1: Perform Experiments

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

Statistical Model: Describes observation process

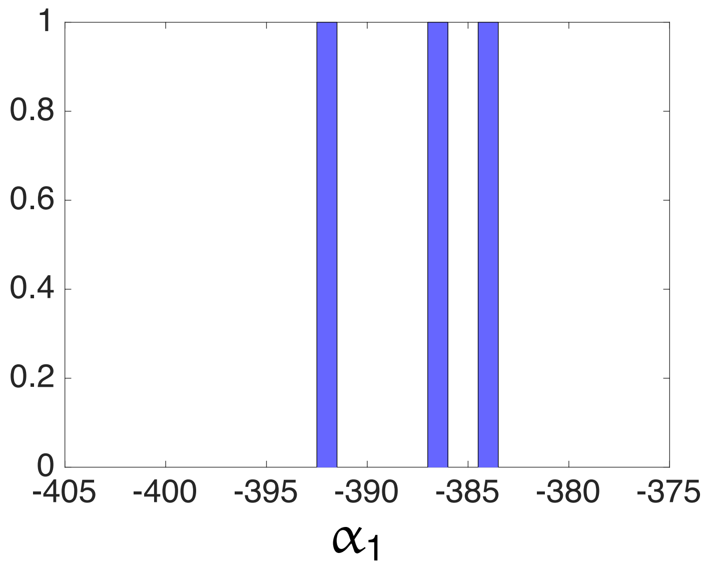
$$\upsilon_i = \psi(P_i, q) + \varepsilon_i \quad , \; i = 1, \ldots, n$$

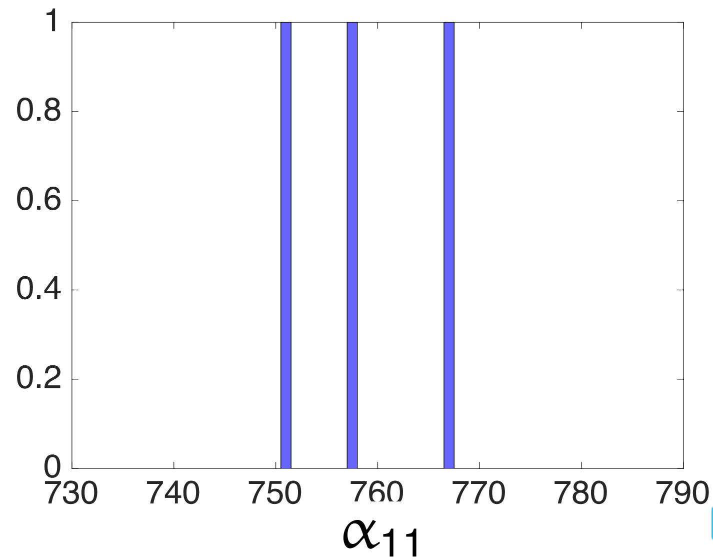


Common Assumption: $\varepsilon_i \sim N(0, \sigma^2)$

UQ Goals: Quantify parameter and response uncertainties

Strategy 1: Perform experiments; e.g., 3





Strategy 1: Perform Experiments

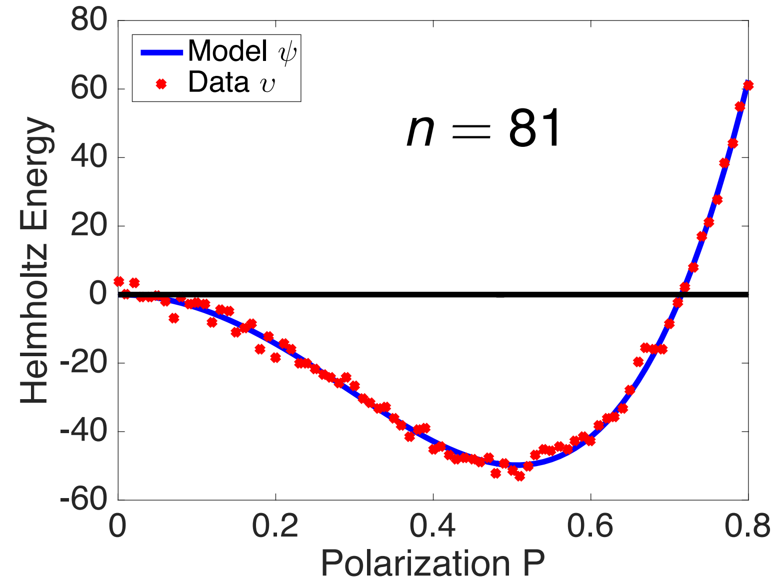
Example: Helmholtz energy $\psi(P,q) = \alpha_1 P^2 + \alpha_{11} P^4 \;,\; q = [\alpha_1, \alpha_{11}]$

Statistical Model: Describes observation process

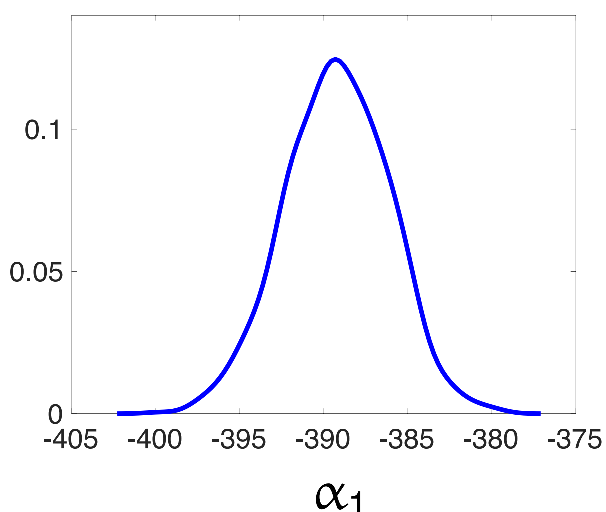
$$\upsilon_i = \psi(P_i, q) + \varepsilon_i \quad , \; i = 1, \ldots, n$$

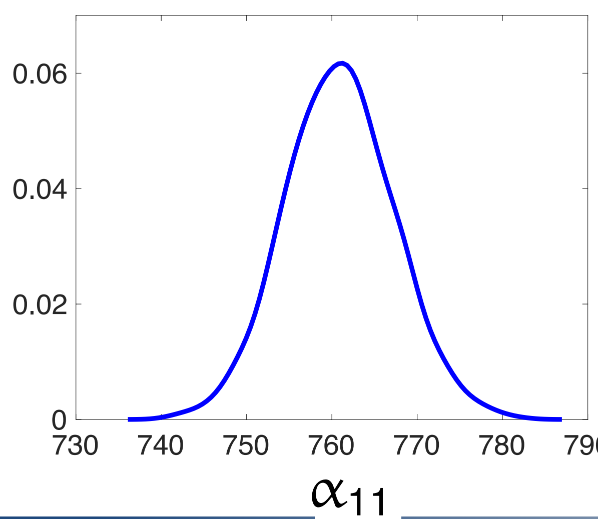
Common Assumption: $\varepsilon_i \sim N(0, \sigma^2)$

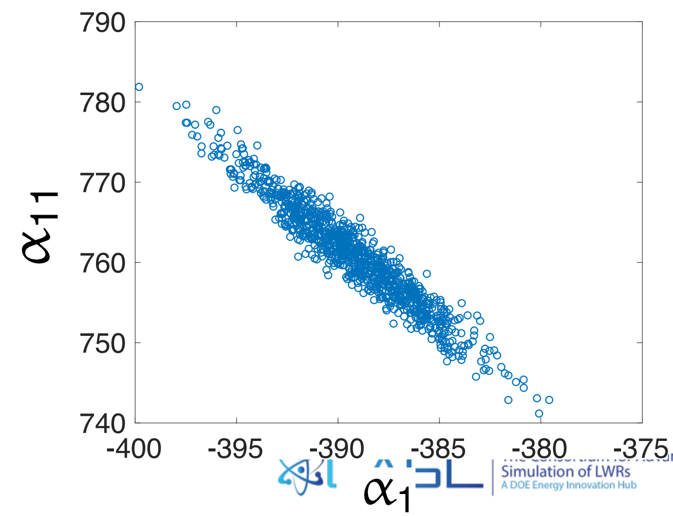
UQ Goals: Quantify parameter and response uncertainties



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







Strategy 1: Perform Experiments

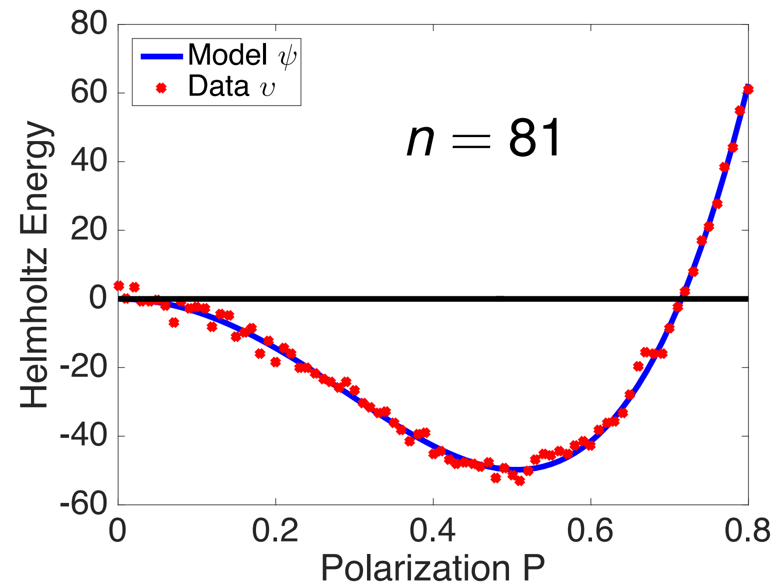
Example: Helmholtz energy $\psi(P,q) = \alpha_1 P^2 + \alpha_{11} P^4$, $q = [\alpha_1, \alpha_{11}]$

Statistical Model: Describes observation process

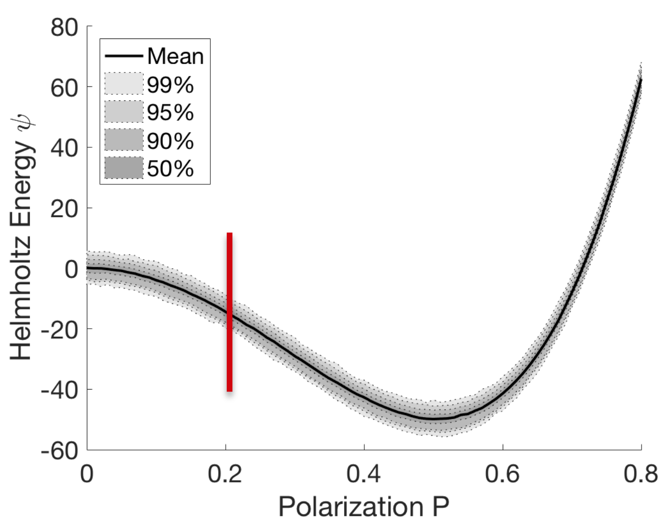
$$\upsilon_i = \psi(P_i, q) + \varepsilon_i \quad , \; i = 1, \ldots, n$$

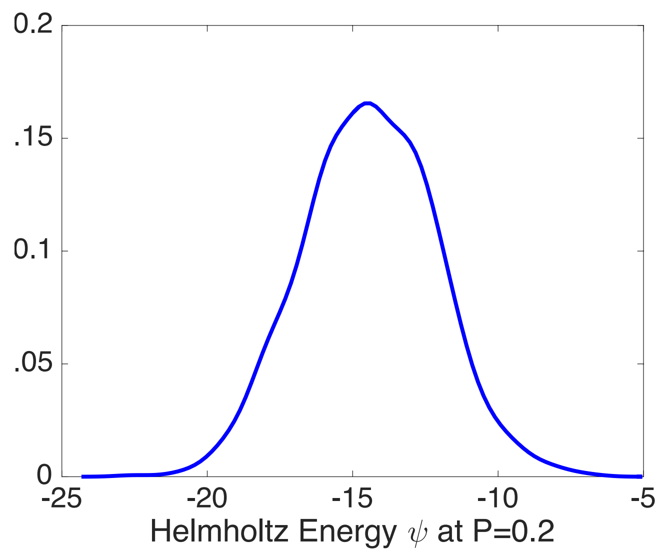
Common Assumption: $\varepsilon_i \sim N(0, \sigma^2)$

UQ Goals: Quantify parameter and response uncertainties



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





Strategy 1: Perform Experiments

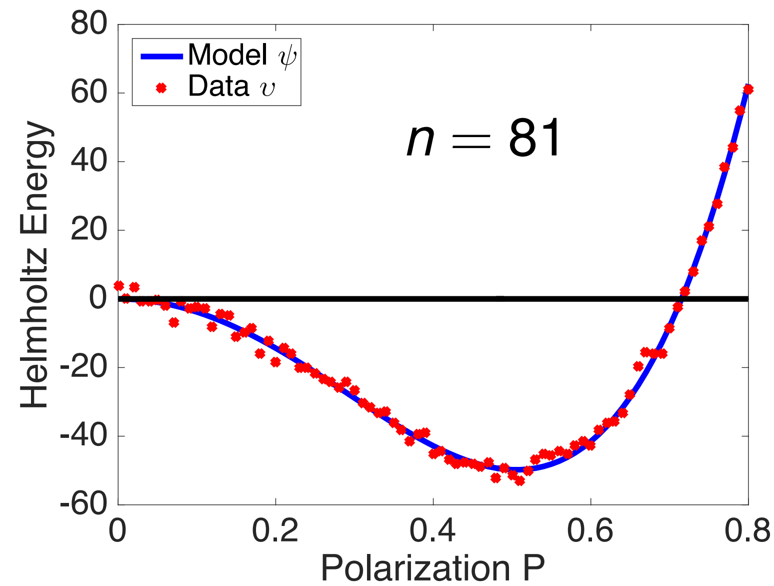
Example: Helmholtz energy $\psi(P,q) = \alpha_1 P^2 + \alpha_{11} P^4$, $q = [\alpha_1, \alpha_{11}]$

Statistical Model: Describes observation process

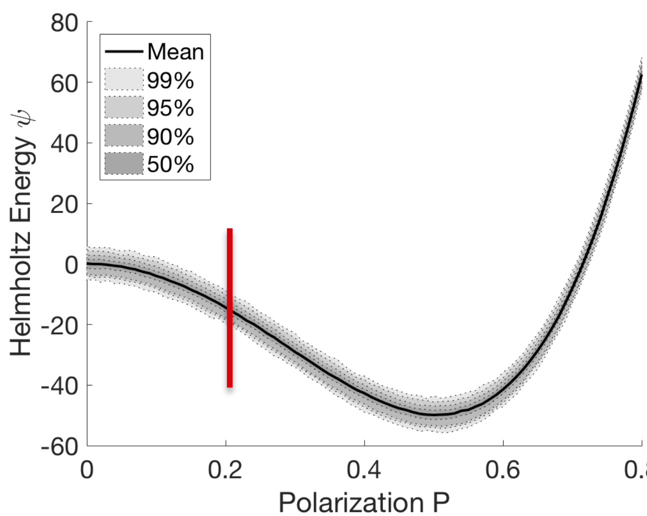
$$\upsilon_i = \psi(P_i, q) + \varepsilon_i \quad , \; i = 1, \ldots, n$$

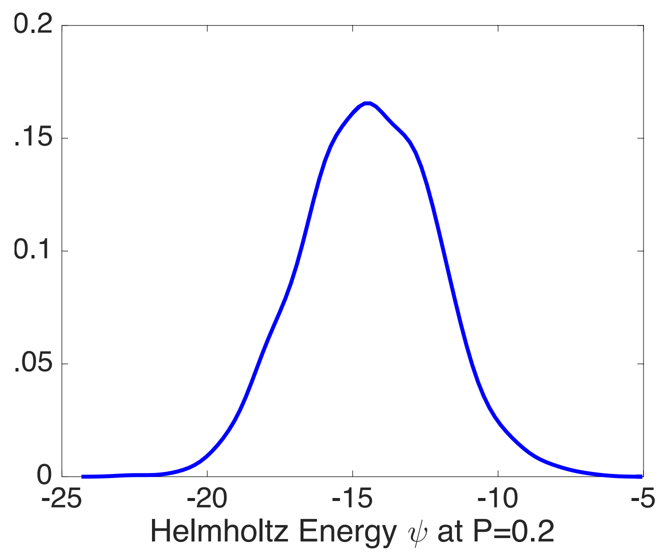
Common Assumption: $\varepsilon_i \sim N(0, \sigma^2)$

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

Polarization Example

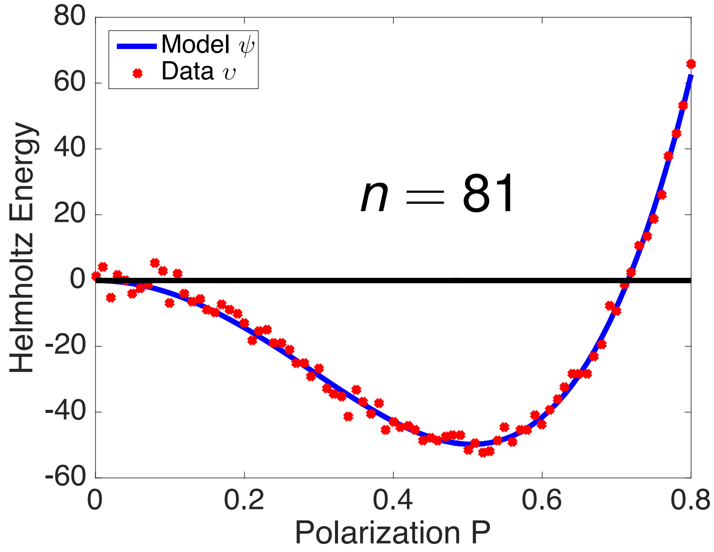
Statistical Model: For i = 1, …, n

$$\upsilon_i = \psi(P_i, q) + \varepsilon_i \longleftarrow \varepsilon_i \sim N(0, \sigma^2)$$

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

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

$$\Rightarrow \upsilon = Xq + \varepsilon$$



Statistical Quantities:

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

Polarization Example

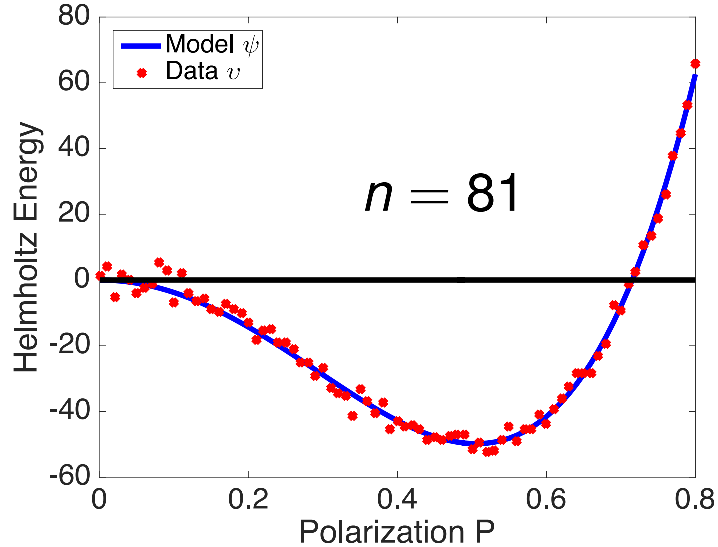
Statistical Model: For i = 1, …, n

$$\upsilon_i = \psi(P_i, q) + \varepsilon_i \qquad \varepsilon_i \sim N(0, \sigma^2)$$

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

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

$$\Rightarrow \upsilon = Xq + \varepsilon$$



Statistical Quantities:

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

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

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

$$= q_0$$

$$\upsilon = Xq_0 + \varepsilon$$

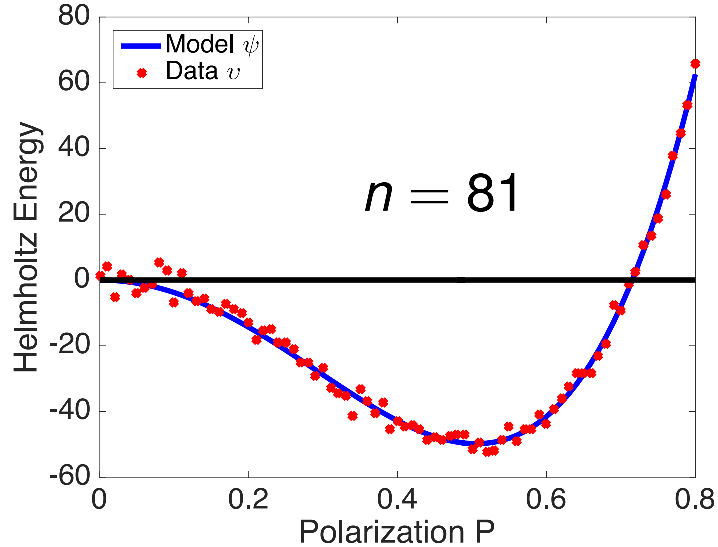
Polarization Example

Statistical Model: For i = 1, …, n

$$\upsilon_i = \psi(P_i, q) + \varepsilon_i \longleftarrow \varepsilon_i \sim N(0, \sigma^2)$$
$$= \alpha_1 P_i^2 + \alpha_{11} P_i^4 + \varepsilon_i$$

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

$$\Rightarrow \upsilon = Xq + \varepsilon$$



Statistical Quantities:

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

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

$$= (X^T X)^{-1} X^T \mathbb{E}(\upsilon)$$
$$= q_0$$

$$\upsilon = Xq_0 + \varepsilon$$

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

$$\begin{aligned} V(q) &= \mathbb{E}[(q - q_0)(q - q_0)^T] \\ &= \mathbb{E}[(q_0 + A\varepsilon - q_0)(q_0 + A\varepsilon - q_0)^T] \ \text{since } q = A\Upsilon = A(Xq_0 + \varepsilon) \\ &= A\mathbb{E}(\varepsilon\varepsilon^T)A^T \\ &= \sigma^2 (X^T X)^{-1} \end{aligned}$$

Polarization Example

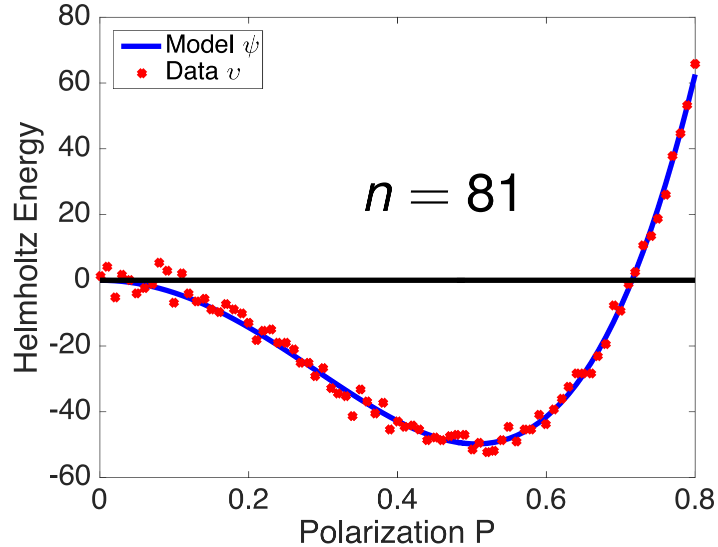
Statistical Model: For i = 1, …, n

$$\upsilon_i = \psi(P_i, q) + \varepsilon_i \longleftarrow \varepsilon_i \sim N(0, \sigma^2)$$

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

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

$$\Rightarrow \upsilon = Xq + \varepsilon$$



Statistical Quantities:

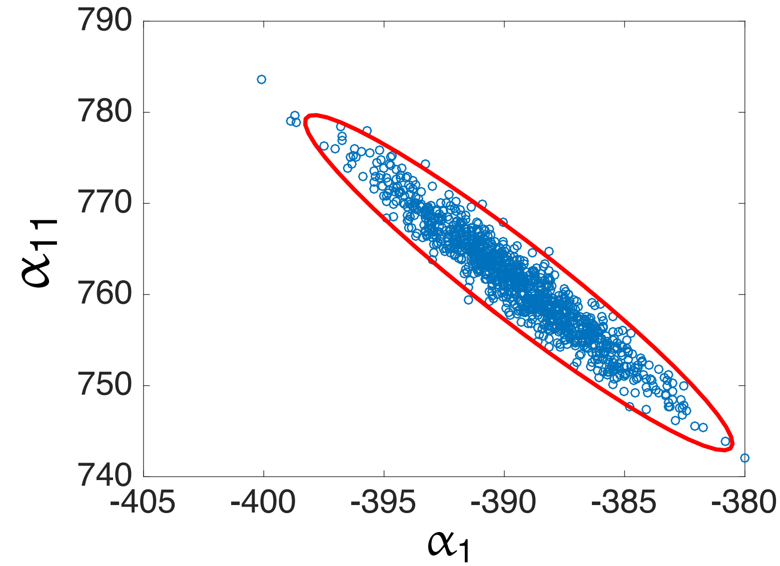
$$q = (X^T X)^{-1} X^T \upsilon$$

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

var(α_1)

cov(α_1, α_{11})

var(α_{11})



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:

o Relies on estimators derived from different data sets and a specific sampling distribution.

o Parameters may be unknown but are fixed and deterministic.

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

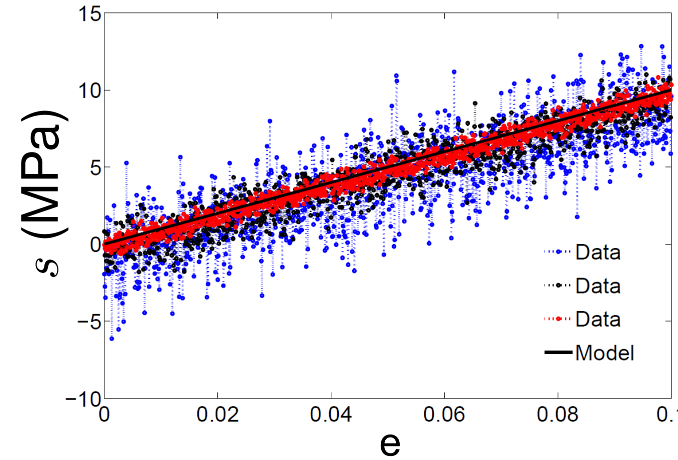
• Parameter Estimation: Parameters are considered to be random variables having associated densities.

Bayesian Inference: Simpler Example

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

$$s_i = Ee_i + \varepsilon_i \ , \ i = 1, \ldots, N$$

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



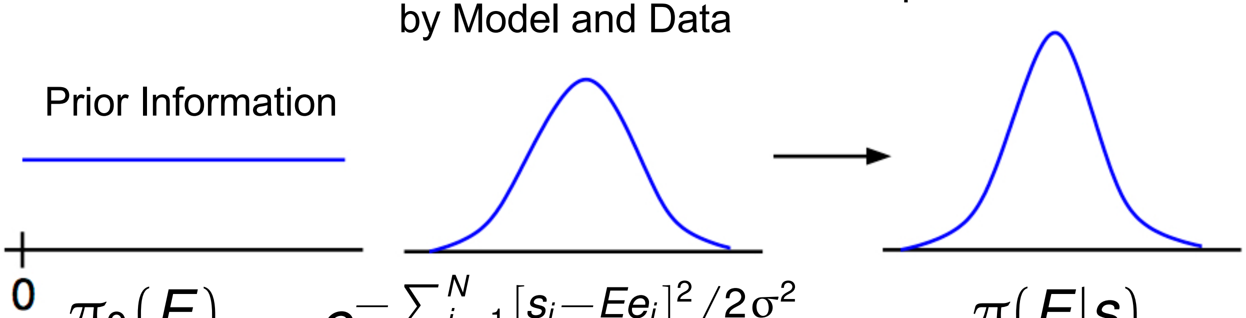
Parameter: Stiffness E

Strategy: Use model fit to data to update prior information

Prior Information

Information Provided by Model and Data

Updated Information



$\pi_0(E)$ $\qquad e^{-\sum_{i=1}^N [s_i - Ee_i]^2/2\sigma^2}$ $\qquad \pi(E|s)$

Data $\qquad$ Model

Non-normalized Bayes' Relation:

$$\pi(E|s) = e^{-\sum_{i=1}^N [s_i - Ee_i]^2/2\sigma^2} \pi_0(E)$$

Bayesian Inference

Bayes' Relation: Specifies posterior in terms of likelihood and prior

Likelihood: $e^{-\sum_{i=1}^{N}[s_i - Ee_i]^2/2\sigma^2}$, $q = E$

$\upsilon = [s_1, \ldots, s_N]$

Posterior Distribution

$$\boxed{\pi(q|\upsilon) = \frac{\pi(\upsilon|q)\pi_0(q)}{\int_{\mathbb{R}^p} \pi(\upsilon|q)\pi_0(q)dq}}$$

Prior Distribution

Normalization Constant

- Prior Distribution: Quantifies prior knowledge of parameter values
- Likelihood: Probability of observing a data given set of parameter values.
- Posterior Distribution: Conditional distribution of parameters given observed data.

Problem: Can require high-dimensional integration

- e.g., Thermal-hydraulics and chemistry codes: p = 5-20!
- 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 Model Calibration

Bayes' Relation:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Bayesian Model Calibration:

- Parameters assumed to be random variables

$$\pi(q|\upsilon) = \frac{\pi(\upsilon|q)\pi_0(q)}{\int_{\mathbb{R}^p} \pi(\upsilon|q)\pi_0(q)dq}$$

Example: Coin Flip

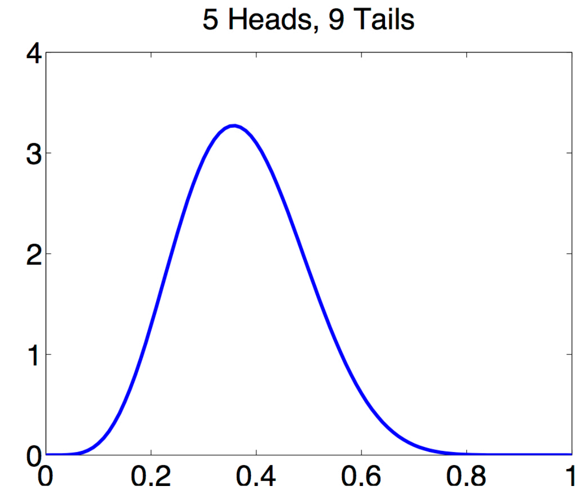
$$\Upsilon_i(\omega) = \begin{cases} 0 & , \quad \omega = T \\ 1 & , \quad \omega = H \end{cases}$$

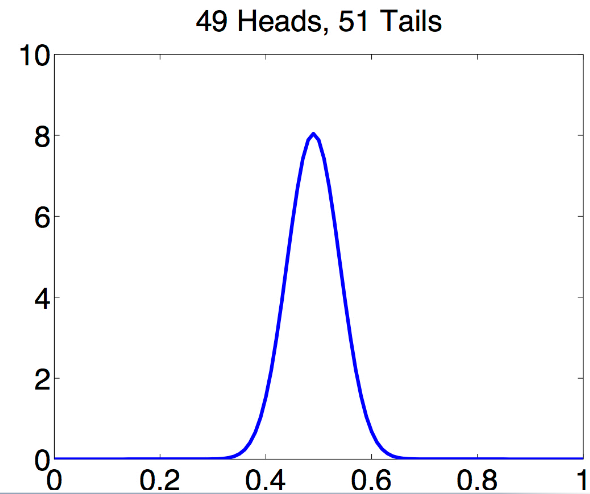
Likelihood:

$$\begin{aligned} \pi(\upsilon|q) &= \prod_{i=1}^{N} q^{\upsilon_i}(1-q)^{1-\upsilon_i} \\ &= q^{N_1}(1-q)^{N_0} \end{aligned}$$

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

$$\pi(q|\upsilon) = \frac{q^{N_1}(1-q)^{N_0}}{\int_0^1 q^{N_1}(1-q)^{N_0}dq} = \frac{(N+1)!}{N_0!N_1!} q^{N_1}(1-q)^{N_0}$$



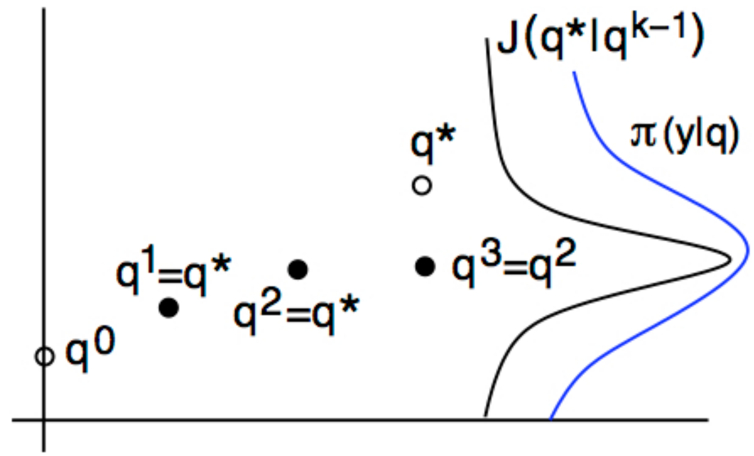


Bayesian Model Calibration

Bayesian Model Calibration:

• Parameters considered to be random variables with associated densities.

$$\pi(q|\upsilon) = \frac{\pi(\upsilon|q)\pi_0(q)}{\int_{\mathbb{R}^p} \pi(\upsilon|q)\pi_0(q)dq}$$



Problem:

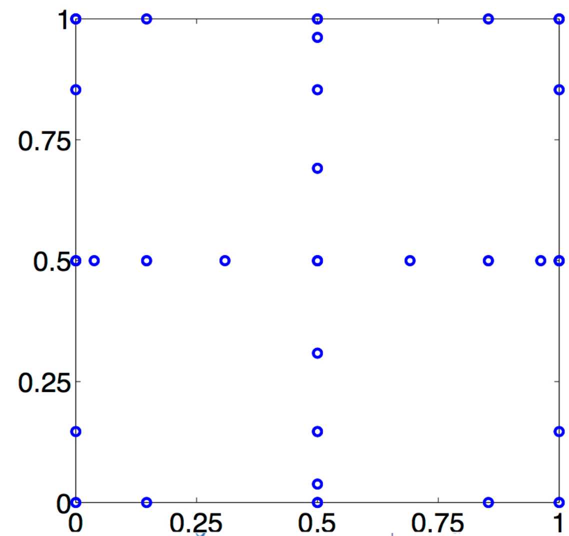
• Often requires high dimensional integration;

  ○ p = hundreds to thousands for some models

Strategies:

• Sampling methods

• Sparse grid quadrature techniques



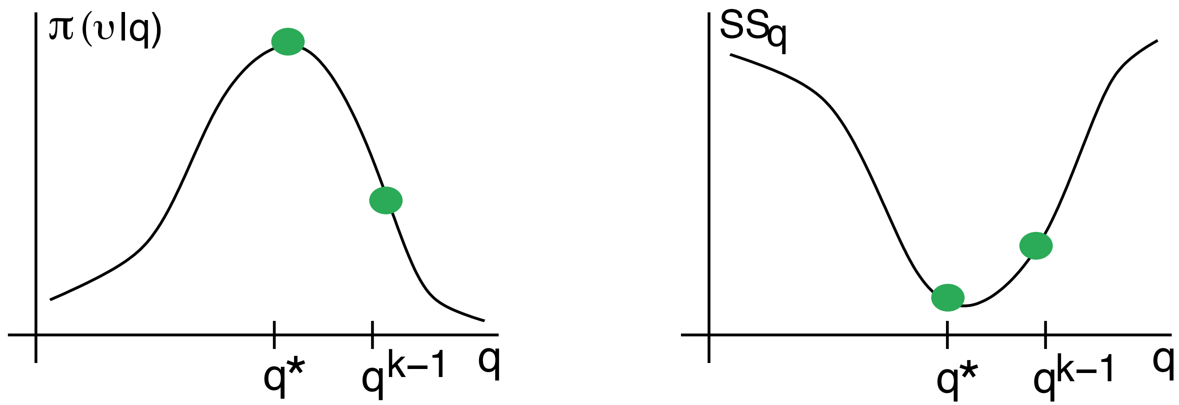
Markov Chain Monte Carlo Methods

Strategy:

- Sample values from proposal distribution $J(q^*|q^{k-1})$ that reflects geometry of posterior distribution
- Compute $r(q^*|q^{k-1}) = \frac{\pi(\upsilon|q^*)\pi_0(q^*)}{\pi(\upsilon|q^{k-1})\pi_0(q^{k-1})}$
  - $*$ If $r \geqslant 1$, 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

$$\pi(\upsilon|q) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-SS_q/2\sigma^2} \qquad SS_q = \sum_{i=1}^{N} [\upsilon_i - f(t_i, q)]^2$$



Delayed Rejection Adaptive Metropolis (DRAM)

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

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

Example: Helmholtz energy

$$\upsilon_i = \psi(P_i, q) + \varepsilon_i \longleftarrow \varepsilon_i \sim N(0, \sigma^2)$$

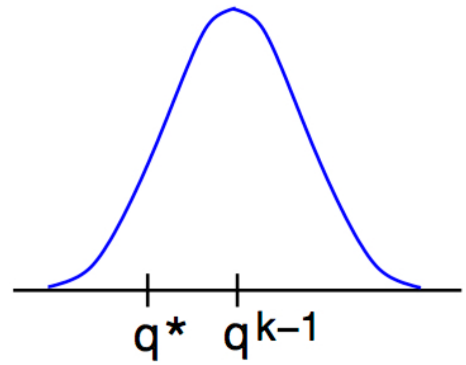
$$= \alpha_1 P_i^2 + \alpha_{11} P_i^4 + \varepsilon_i$$

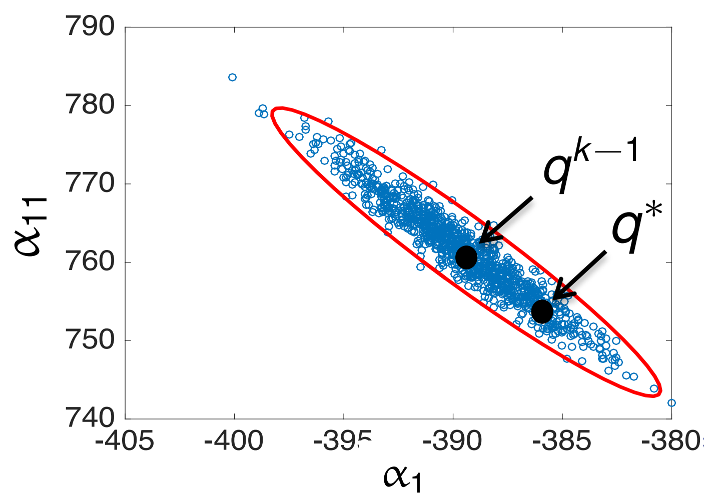
Delayed Rejection Adaptive Metropolis

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

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

   (a) Construct candidate $q^* \sim N(q^{k-1}, \underline{V})$





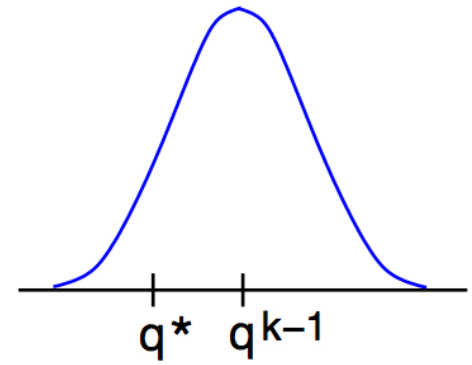
Example: Helmholtz energy

$$\upsilon_i = \psi(P_i, q) + \varepsilon_i \leftarrow \varepsilon_i \sim N(0, \sigma^2)$$
$$= \alpha_1 P_i^2 + \alpha_{11} P_i^4 + \varepsilon_i$$

Recall: Covariance V incorporates geometry

Delayed Rejection Adaptive Metropolis

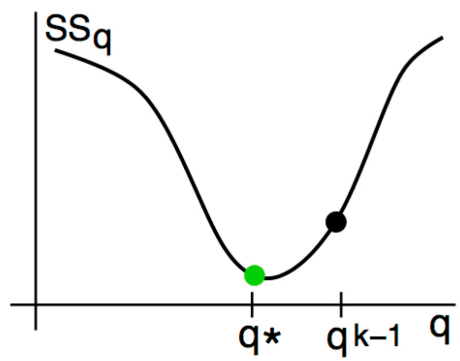
Algorithm: [Haario et al., 2006] – MATLAB, Python, R

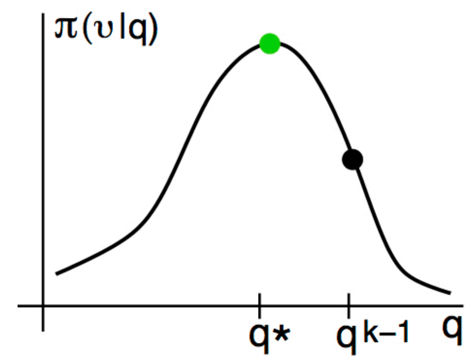


1. Determine $q^0 = \arg\min_q \sum_{i=1}^{N} [\upsilon_i - \psi(P_i, q)]^2]$
2. For $k = 1, \ldots, M$
   (a) Construct candidate $q^* \sim N(q^{k-1}, V)$
   (b) Compute likelihood

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

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



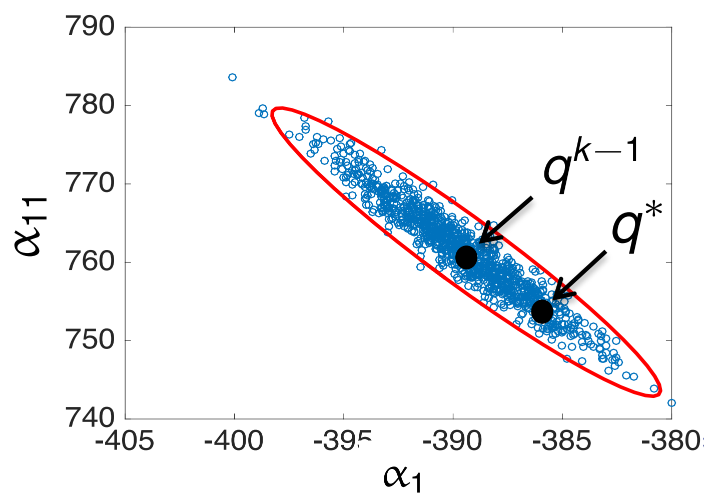


   (c) Accept q^* with probability dictated by likelihood

Example: Helmholtz energy

$$\upsilon_i = \psi(P_i, q) + \varepsilon_i \longleftarrow \varepsilon_i \sim N(0, \sigma^2)$$

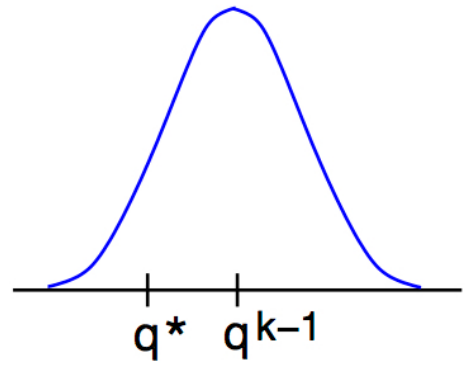
$$= \alpha_1 P_i^2 + \alpha_{11} P_i^4 + \varepsilon_i$$



Recall: Covariance V incorporates geometry

Delayed Rejection Adaptive Metropolis

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



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

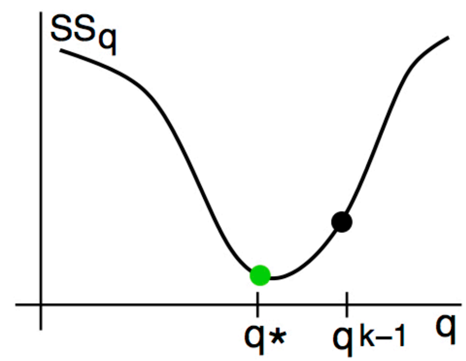
2. For $k = 1, \ldots, M$

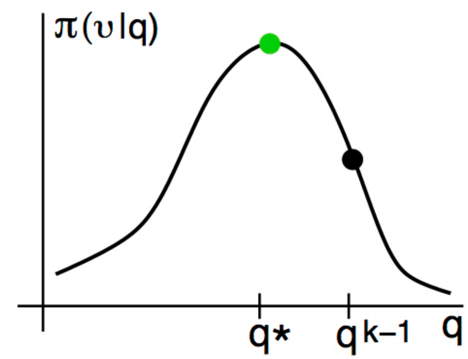
   (a) Construct candidate $q^* \sim N(q^{k-1}, V)$

   (b) Compute likelihood

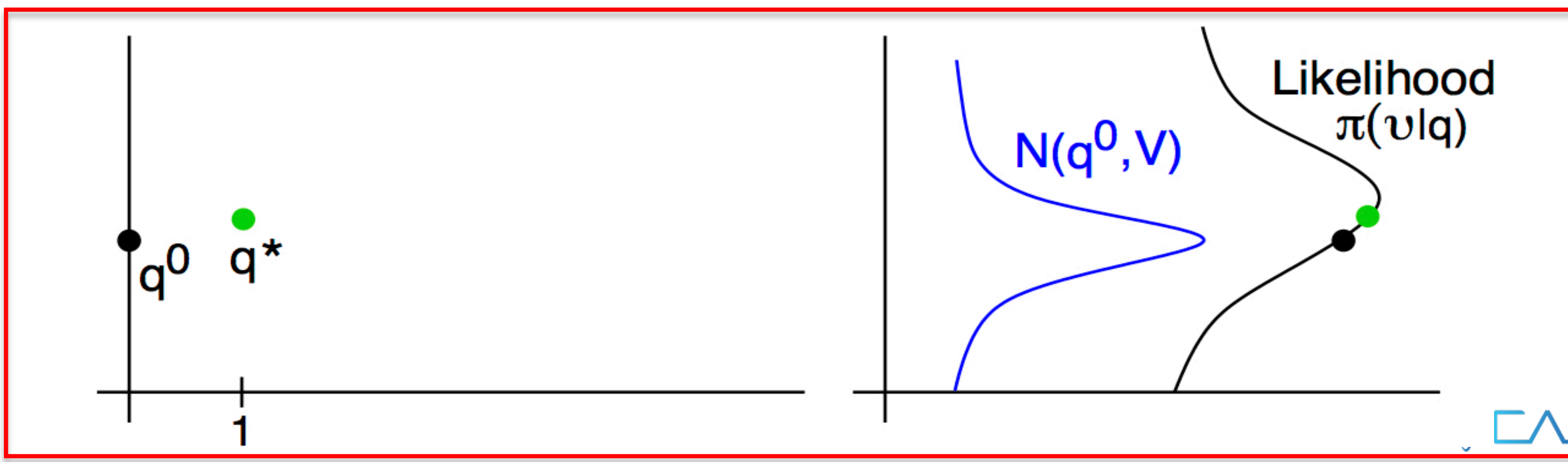
$$SS_{q^*} = \sum_{i=1}^{N} \upsilon_i - \psi(P_i, q^*)]^2$$

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



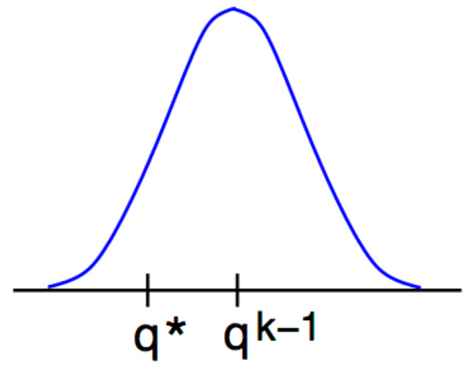


   (c) Accept q^* with probability dictated by likelihood



Delayed Rejection Adaptive Metropolis

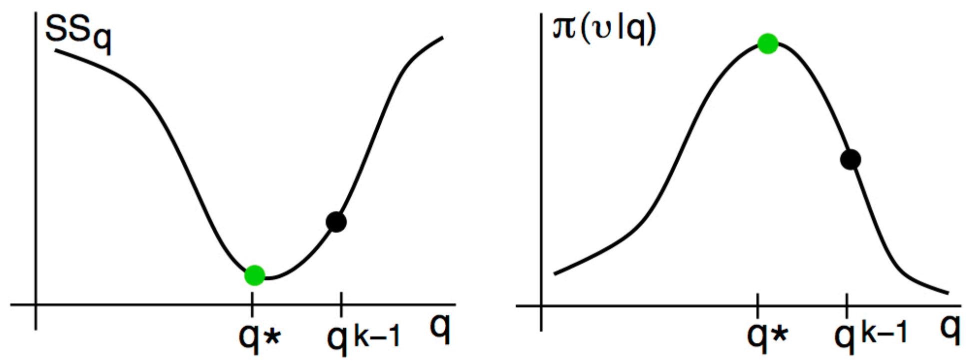
Algorithm: [Haario et al., 2006] – MATLAB, Python, R



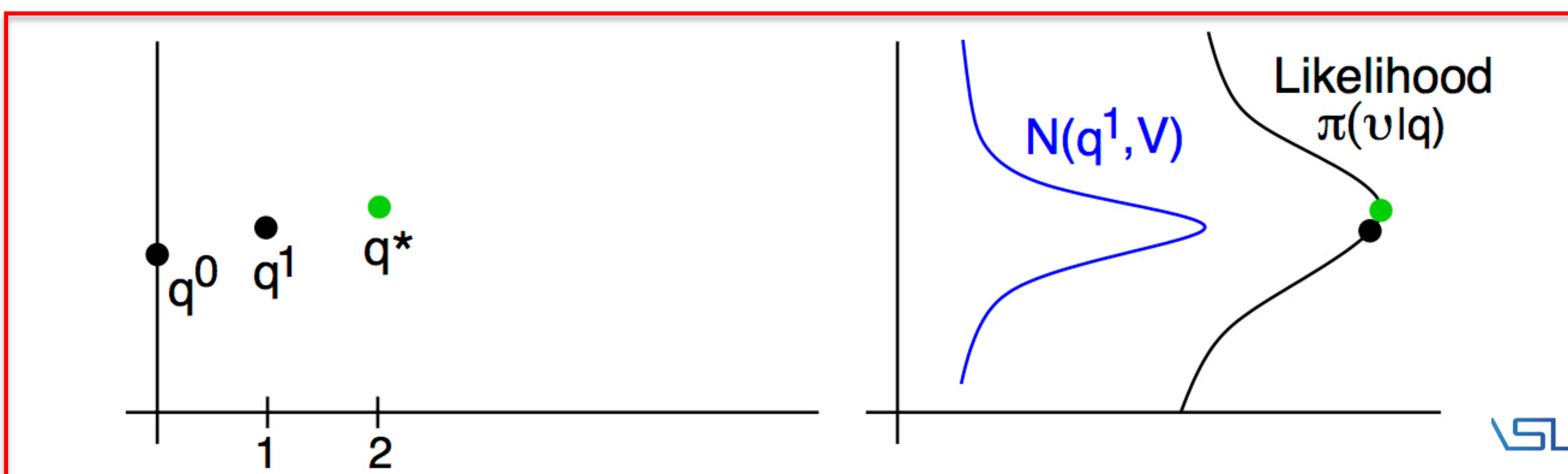
1. Determine $q^0 = \arg\min_q \sum_{i=1}^{N} [\upsilon_i - \psi(P_i, q)]^2]$
2. For $k = 1, \ldots, M$
   (a) Construct candidate $q^* \sim N(q^{k-1}, V)$
   (b) Compute likelihood

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

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



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



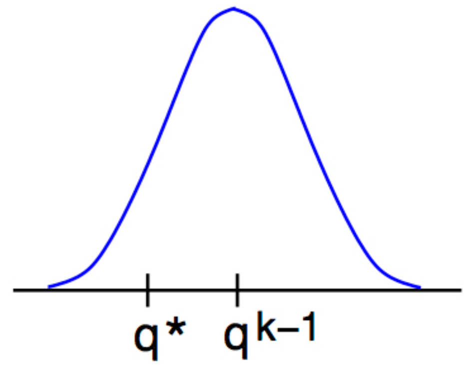
Delayed Rejection Adaptive Metropolis

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

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

2. For $k = 1, \dots, M$

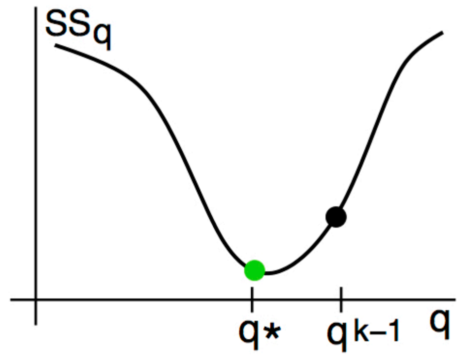
   (a) Construct candidate $q^* \sim N(q^{k-1}, V)$

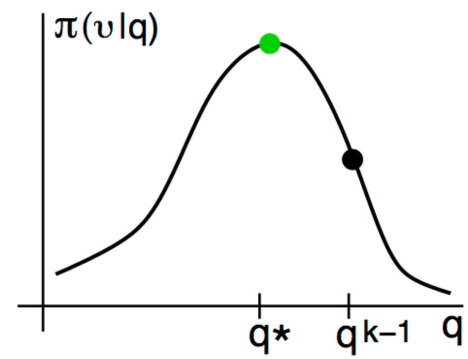


   (b) Compute likelihood

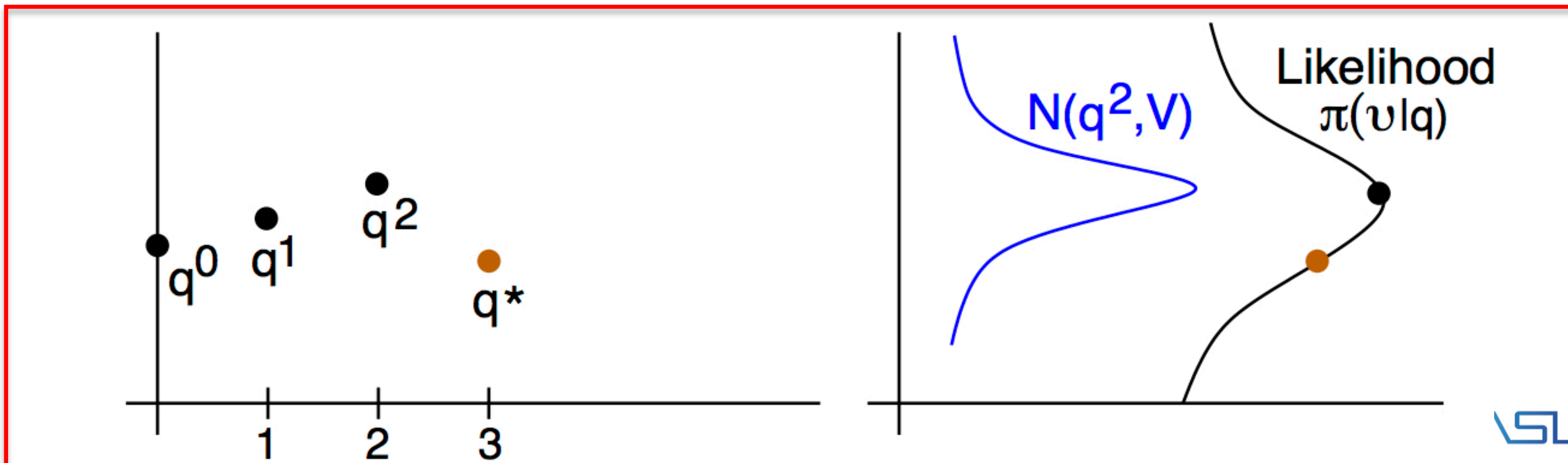
$$SS_{q^*} = \sum_{i=1}^{N} \upsilon_i - \psi(P_i, q^*)]^2$$

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



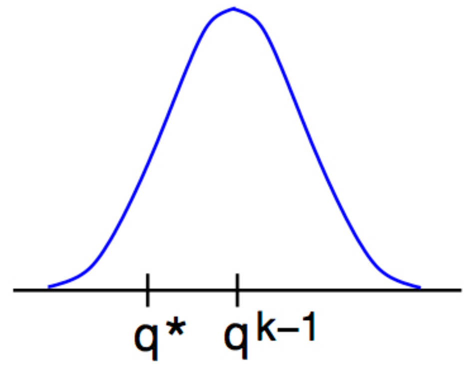


   (c) Accept q^* with probability dictated by likelihood



Delayed Rejection Adaptive Metropolis

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



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

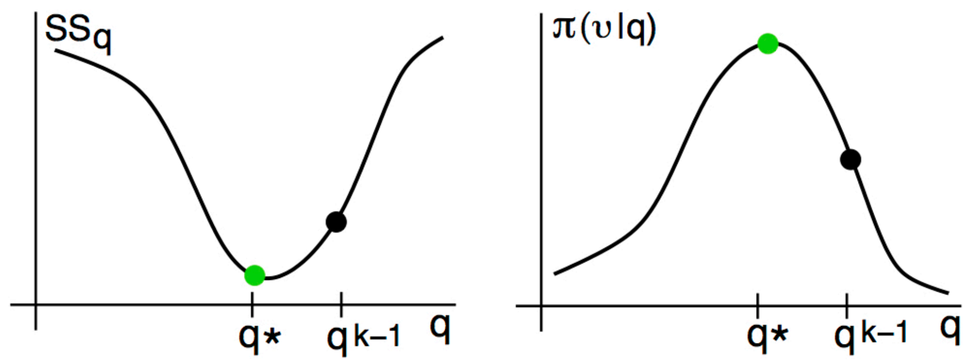
2. For $k = 1, \ldots, M$

   (a) Construct candidate $q^* \sim N(q^{k-1}, V)$

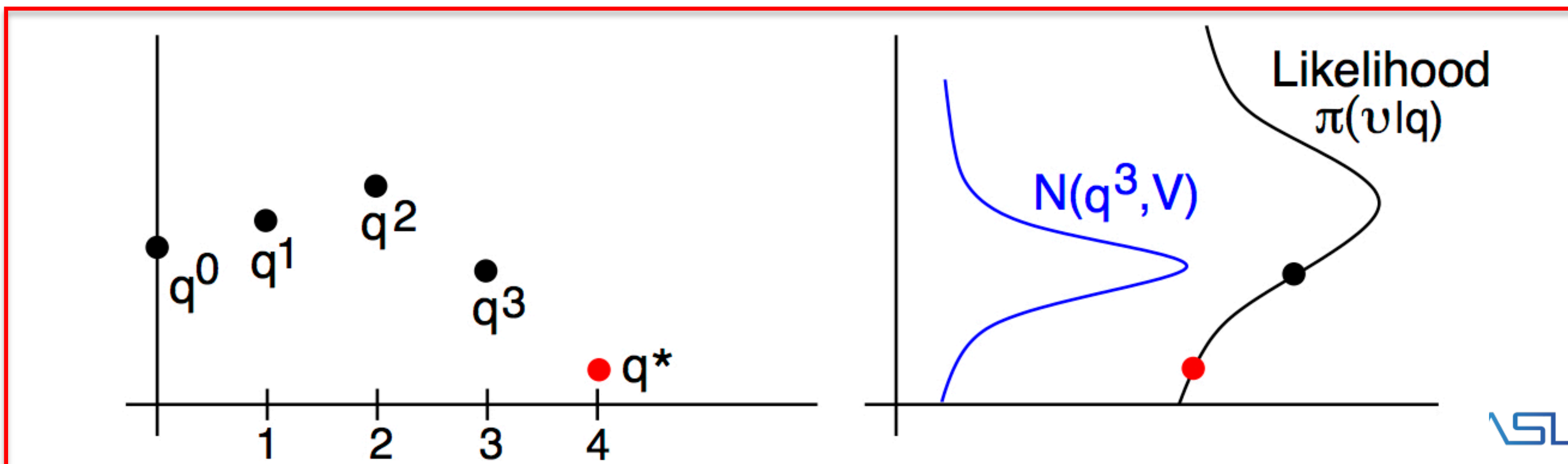
   (b) Compute likelihood

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

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

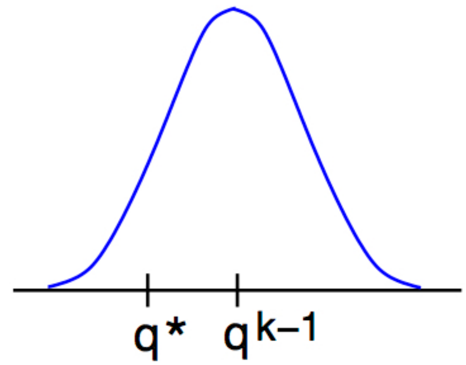


   (c) Accept q^* with probability dictated by likelihood



Delayed Rejection Adaptive Metropolis

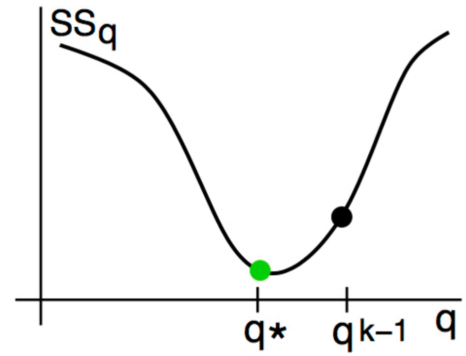
Algorithm: [Haario et al., 2006] – MATLAB, Python, R

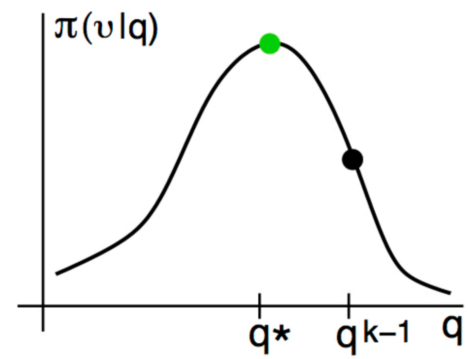


1. Determine $q^0 = \arg\min_q \sum_{i=1}^{N} [\upsilon_i - \psi(P_i, q)]^2]$
2. For $k = 1, \ldots, M$
   (a) Construct candidate $q^* \sim N(q^{k-1}, V)$
   (b) Compute likelihood

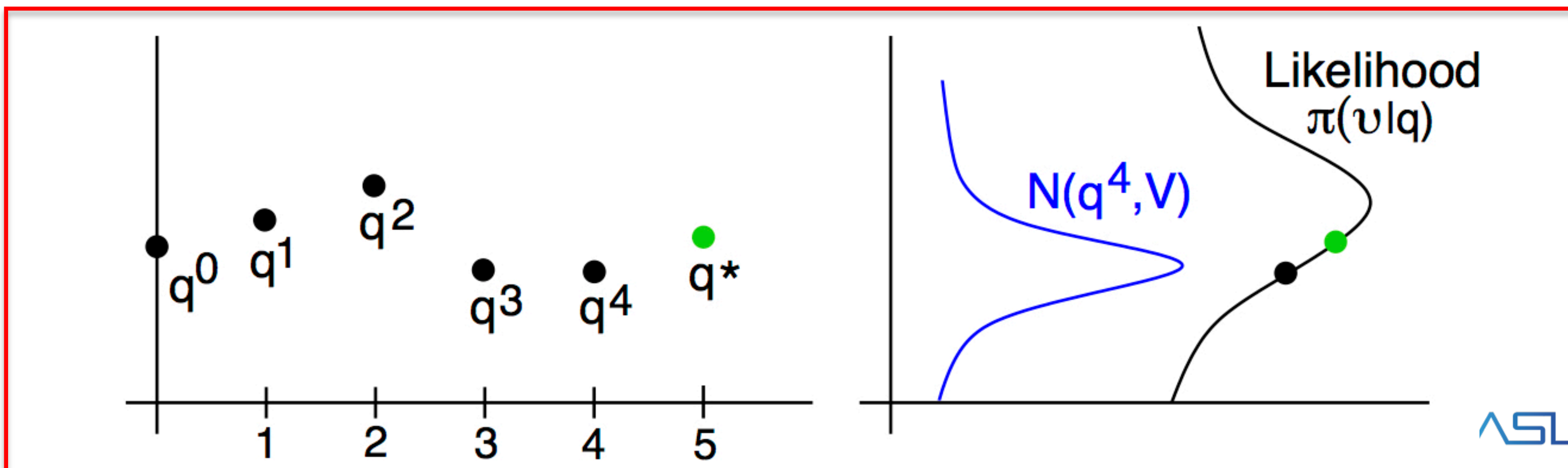
$$SS_{q^*} = \sum_{i=1}^{N} \upsilon_i - \psi(P_i, q^*)]^2$$

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





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



Delayed Rejection Adaptive Metropolis

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

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

2. For $k = 1, \ldots, M$

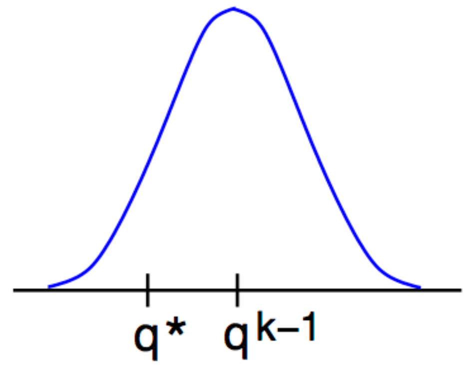
   (a) Construct candidate $q^* \sim N(q^{k-1}, V)$

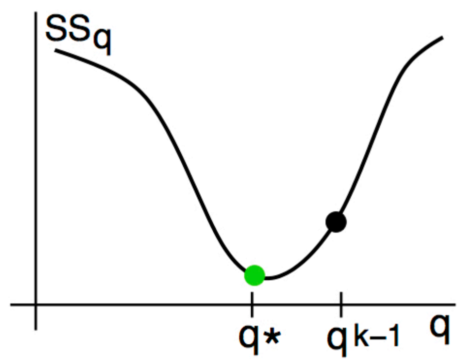
   (b) Compute likelihood

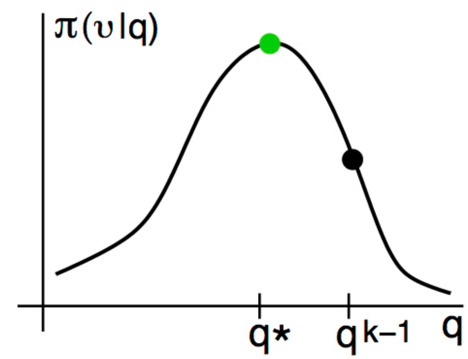
$$SS_{q^*} = \sum_{i=1}^{N} \upsilon_i - \psi(P_i, q^*)]^2$$

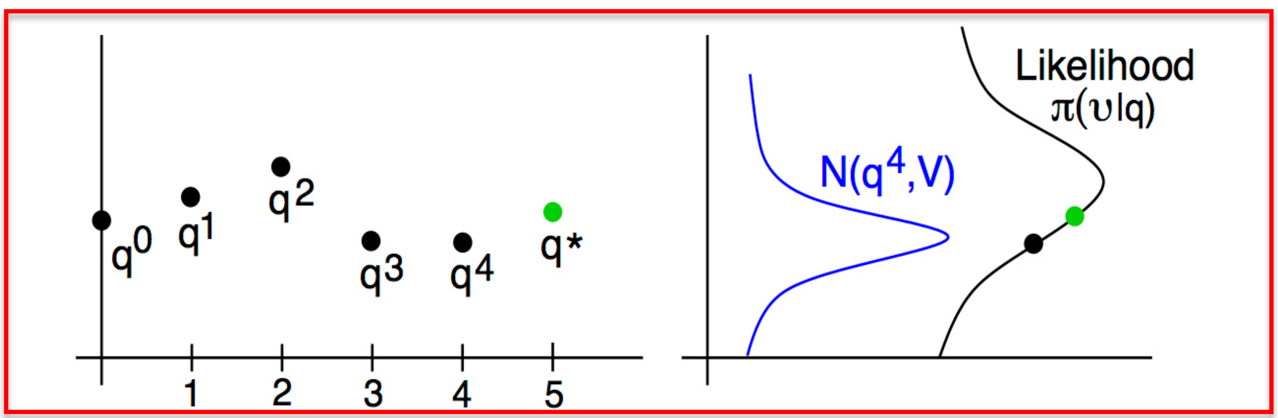
$$\pi(\upsilon|q) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-SS_q/2\sigma^2}$$

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









Note:

- Delayed Rejection: Shrink proposal: γV
- Adaptive Metropolis: Update proposal as samples are accepted

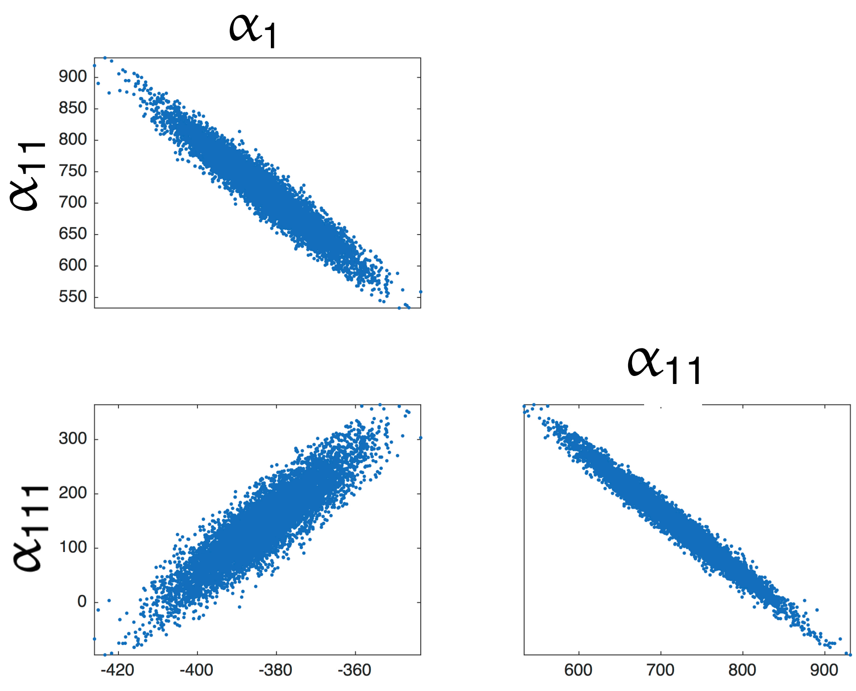
Delayed Rejection Adaptive Metropolis

Example: Helmholtz energy with 3 parameters

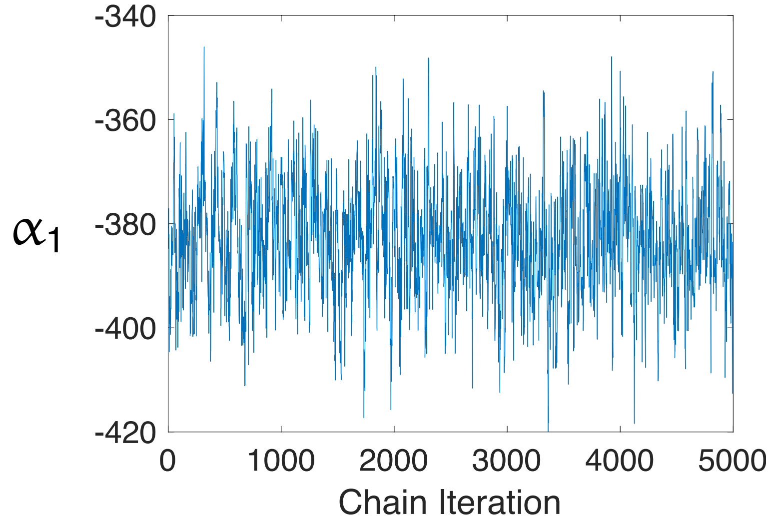
$$\psi(P,q) = \underline{\alpha_1} P^2 + \underline{\alpha_{11}} P^4 + \underline{\alpha_{111}} P^6$$

Note: Similar results for α_{11} and α_{111}

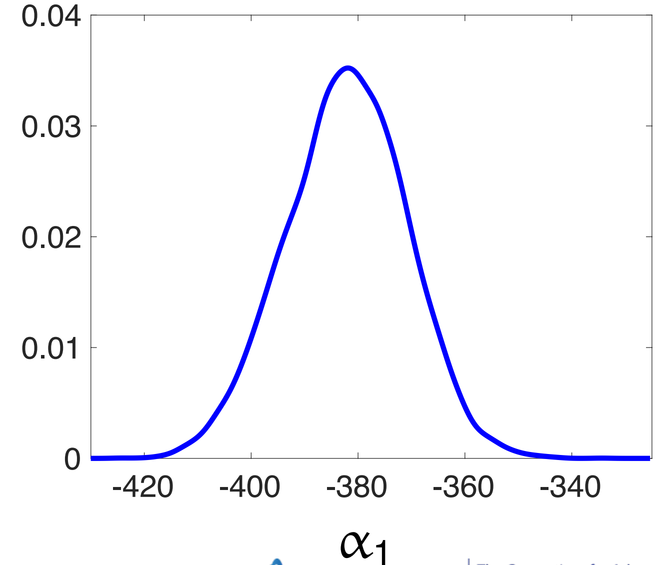
Pairwise Plots: Quantify correlation



Chain for α_1 with 5000 samples



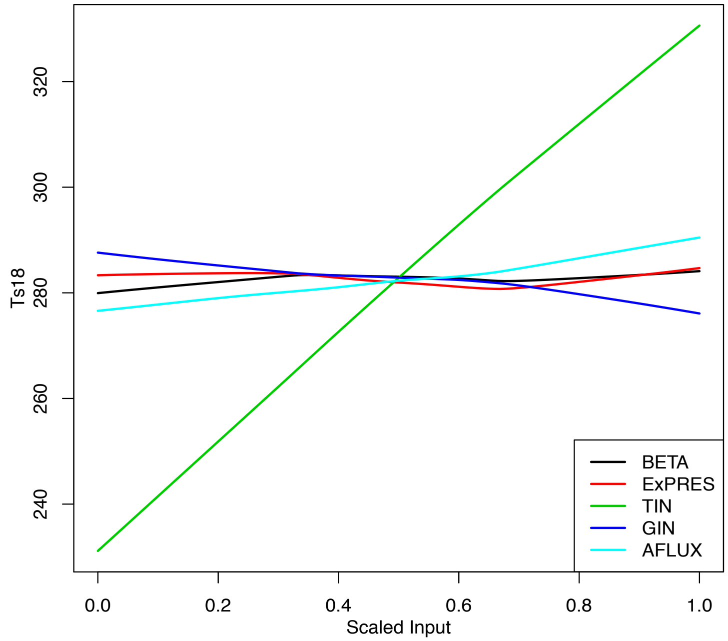
Marginal density for α_1



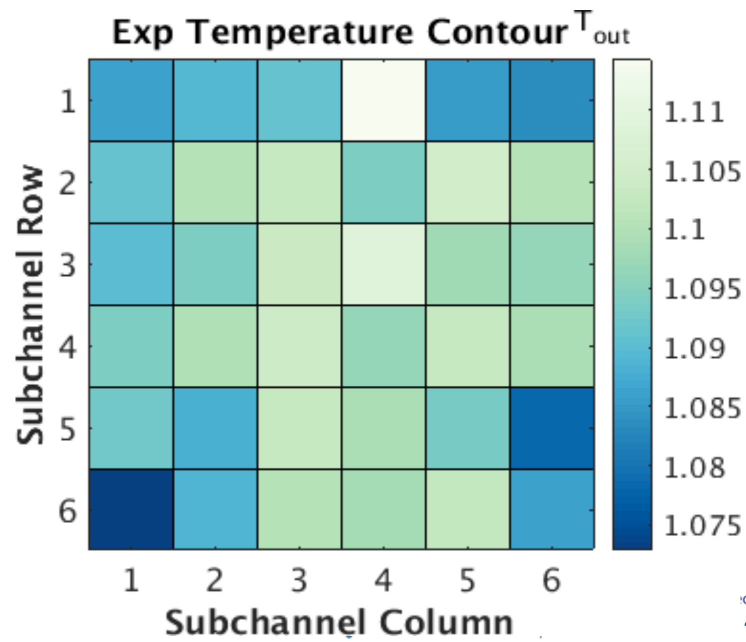
Bayesian Calibration: Beta in CTF

Problem Setup:

- Configuration (Design) Variables in STAR
  - ExPRES: Initial pressure of fluid domain
  - TIN: Initial temperature in fluid domain
  - GIN: Inlet mass flow rate
  - AFLUX: Average linear heat rate per rod



- Calibration Variable in CTF
  - BETA: Turbulent mixing factor
- Experimental Data from WEC
  - 21 tests each of which produce 36 outlet temperatures



Surrogate Construction for CTF

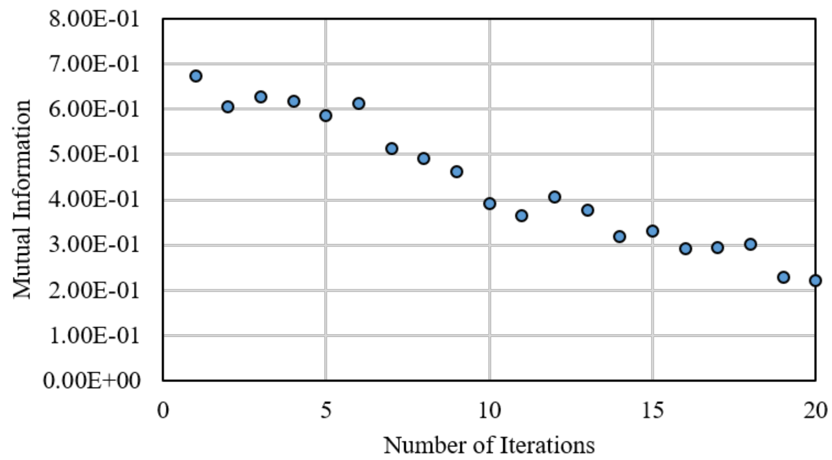
Bayesian Inference:

- MCMC with 20% burn-in removed and subsampling rate of 3 requires minimum of 18,750 iterations.
- Mutual information computation requires 5000 independent samples.
- Each CTF takes approximately 5 minutes.
- This necessitates construction and verification of fast surrogate for CTF – will discuss later
- Gaussian process (GP) surrogate trained and verified for all 36 subchannels.
  - 1000 LHS samples used to compute surrogate, 300 LHS used for verification.
  - Difference between surrogate and CTF-computed outlet temperatures within 1.8%.
  - Surrogate runs in seconds.

Bayesian Calibration of Beta

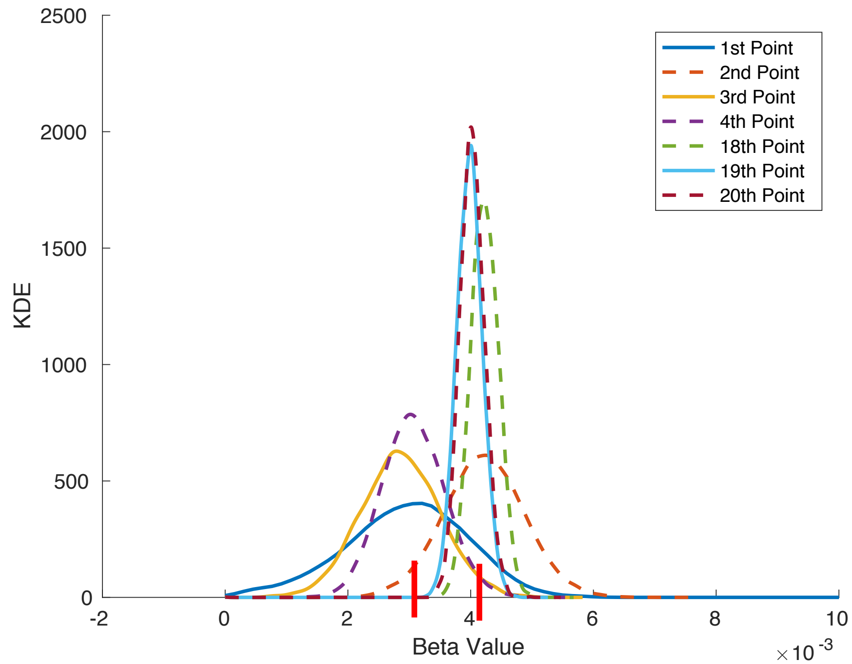
Hi2Lo Workflow and Results:

- Calibrate Beta to initial simulation and/or experiment.
- Performed Hi2Lo calibration using both experimental data and STAR simulations.
- Estimate MI between Beta samples and HiFi predictions at each candidate. Select candidate with largest MI.
- Repeat until MI is sufficiently small or design budget is exhausted.



Results:

- Mean Beta value increased from 0.0028 to 0.004 with *reduced uncertainty*.



Bayesian Inference

Advantages:

• Advantageous over frequentist inference when data is limited.

• Directly provides parameter densities, which can subsequently be propagated to construct response uncertainties.

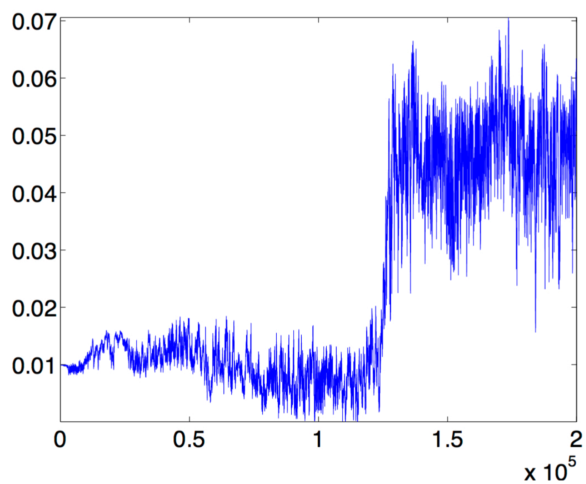
• Can be used to infer non-identifiable parameters if priors are tight.

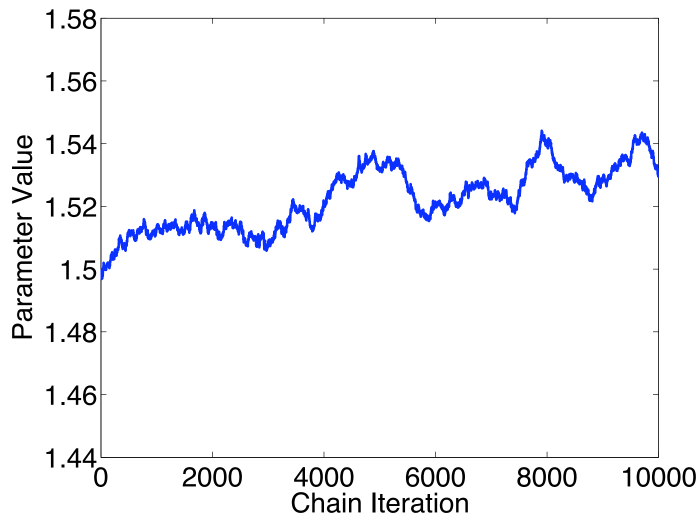
• Provides natural framework for experimental design.

Disadvantages:

• More computationally intense than frequentist inference.

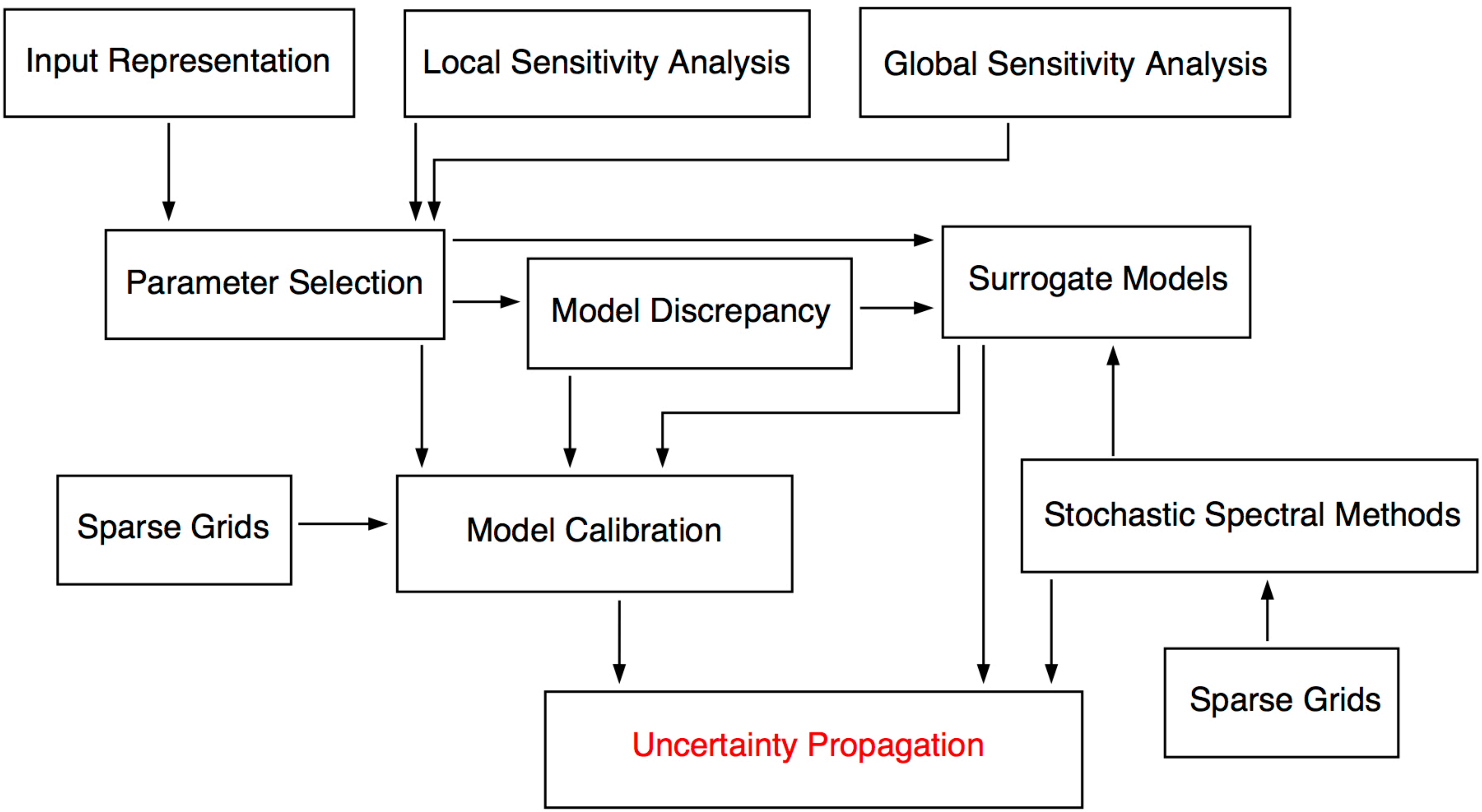
• Can be difficult to confirm that chains have burned-in or converged.





Steps in Uncertainty Quantification

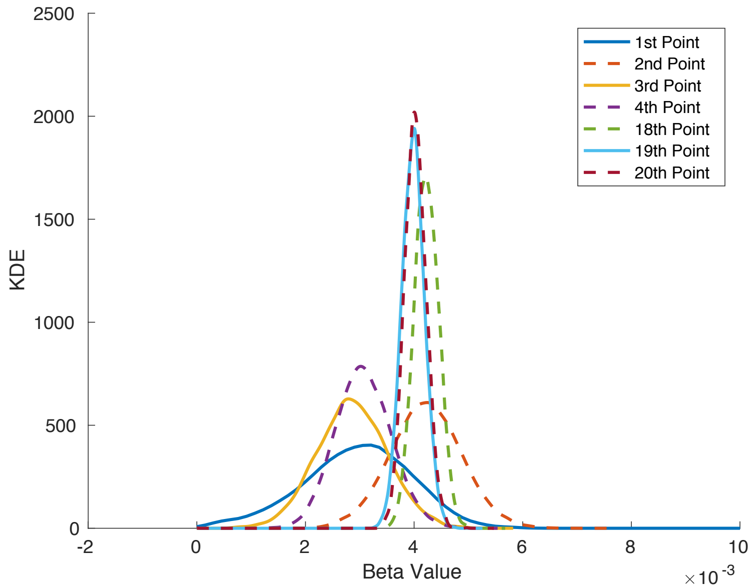
Note: Uncertainty quantification requires synergy between statistics, mathematics and application area.



Uncertainty Propagation

Setting:

- We assume that we have determined distributions for parameters
    - e.g., Bayesian inference, prior experiments, expert opinion



Goal: Construct statistics for quantities of interest (QoI)

- e.g., Void fraction, peak clad temperature, total pressure drop
- Note: Often involves moderate to high-dimensional integration

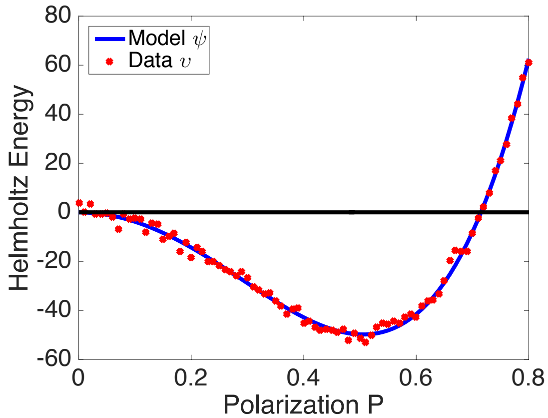
$$\mathbb{E}[u(t,x)] = \int_{\mathbb{R}^p} u(t,x,q)\rho(q)dq$$

Uncertainty Propagation: Linear Models

Note: Analytic mean and variance relations

Example: Helmholtz energy

$$\Upsilon_i = \alpha_1 P_i^2 + \alpha_{11} P_i^4 + \varepsilon_i \; , \; \text{var}[\varepsilon_i] = \sigma^2$$



Model Statistics:

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

$$\mathbb{E}[\alpha_1 P_i^2 + \alpha_{11} P_i^4] = \overline{\alpha}_1 P_i^2 + \overline{\alpha}_{11} P_i^4$$

$$\text{var}[\alpha_1 P_i^2 + \alpha_{11} P_i^4] = P_i^4 \text{var}[\alpha_1] + P_i^8 \text{var}[\alpha_{11}] + 2P_i^6 \text{cov}[\alpha_1, \alpha_{11}]$$

Response Statistics: Assume measurement errors uncorrelated from model response.

$$\mathbb{E}[\Upsilon] = \overline{\alpha}_1 P_i^2 + \overline{\alpha}_{11} P_i^4$$

$$\text{var}[\Upsilon] = P_i^4 \text{var}[\alpha_1] + P_i^8 \text{var}[\alpha_{11}] + 2P_i^6 \text{cov}[\alpha_1, \alpha_{11}] + \sigma^2$$

Problem: Models almost always nonlinearly parameterized

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.

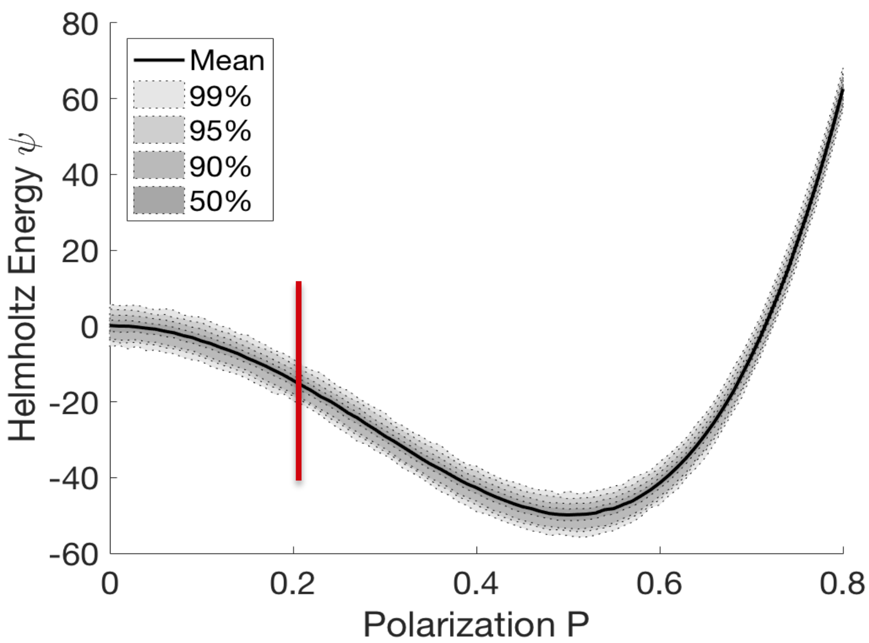
Strategy 2: Employ numerical surrogate representations to analytically propagate uncertainties.

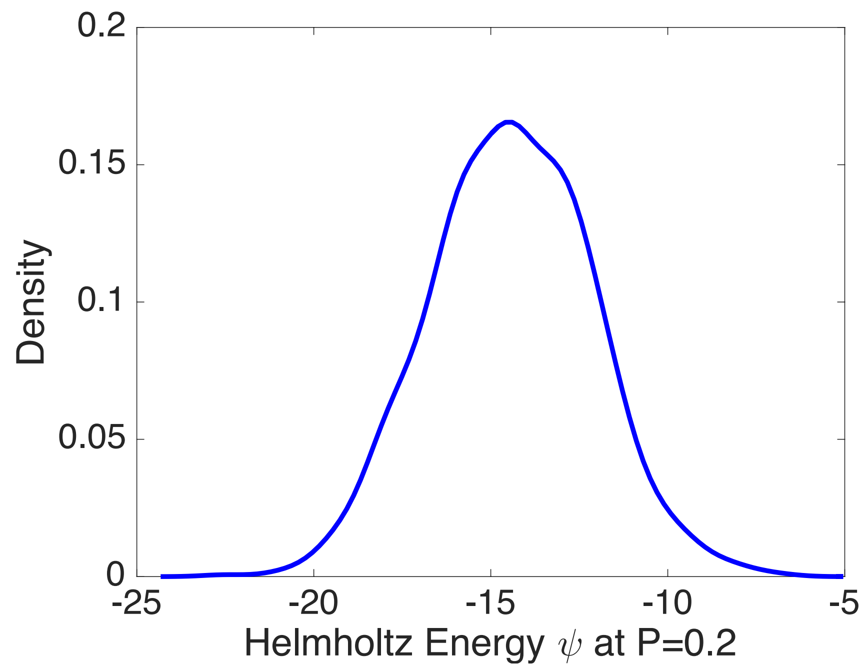
Prediction Intervals

Note:

- We now know how to compute the mean response for the QoI.
- Sample to compute prediction intervals.

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

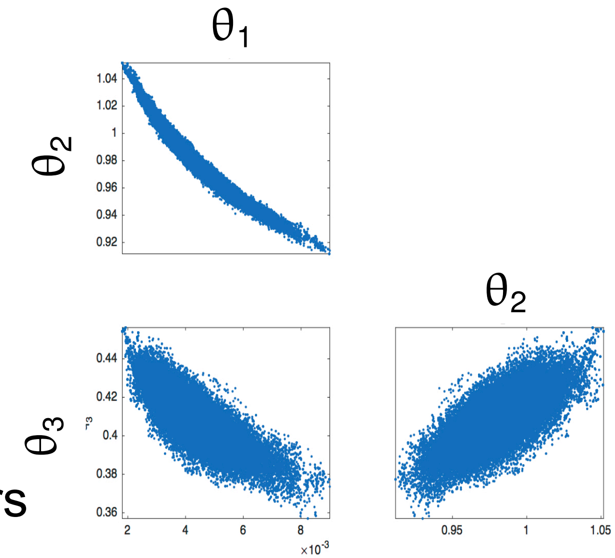




Bayesian Calibration: CASL Application

Example: Dittus—Boelter Relation

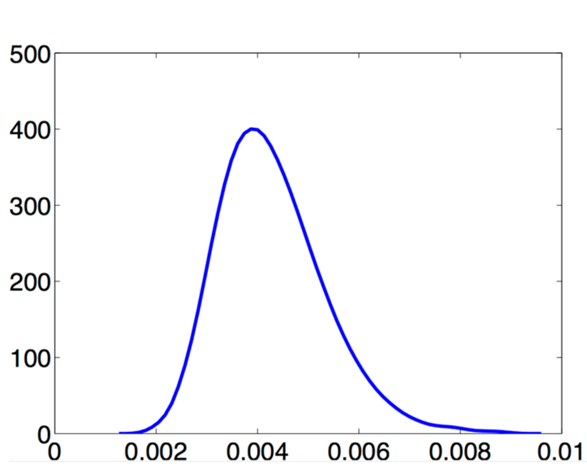
$$Nu = 0.023 Re^{0.8} Pr^{0.4}$$



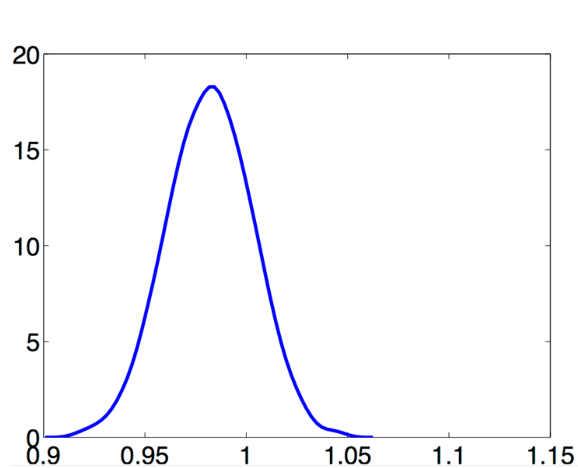
Industry Standard: Conservative, uniform, bounds

i.e., [0, 0.046], [0, 1.6], [0,0.8]

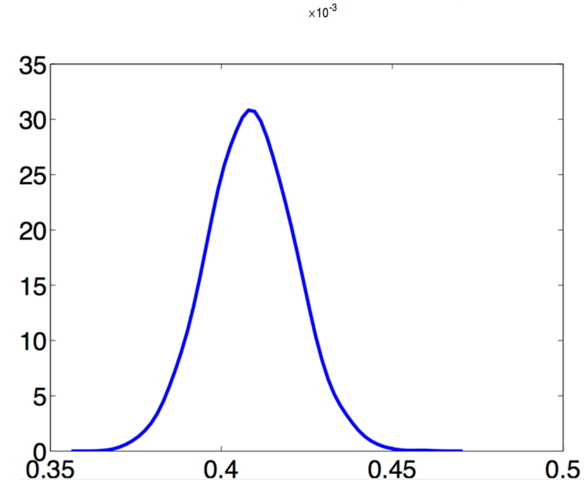
Bayesian Analysis: Employ conservative bounds as priors



$2\sigma \approx 0.0035$



$2\sigma \approx 0.06$



$2\sigma \approx 0.03$

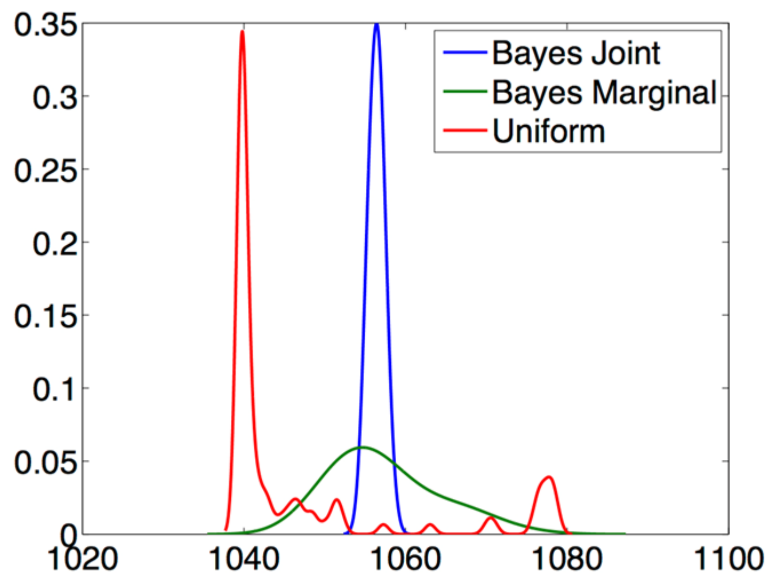
Note:

- Substantial reduction in parameter uncertainty
- Quantifies correlation between parameters

The Consortium for Advanced Simulation of LWRs
A DOE Energy Innovation Hub

Use of Prediction Intervals: CASL

Strategy: Propagate parameter uncertainties through COBRA-TF to determine uncertainty in maximum fuel temperature



Ramifications:

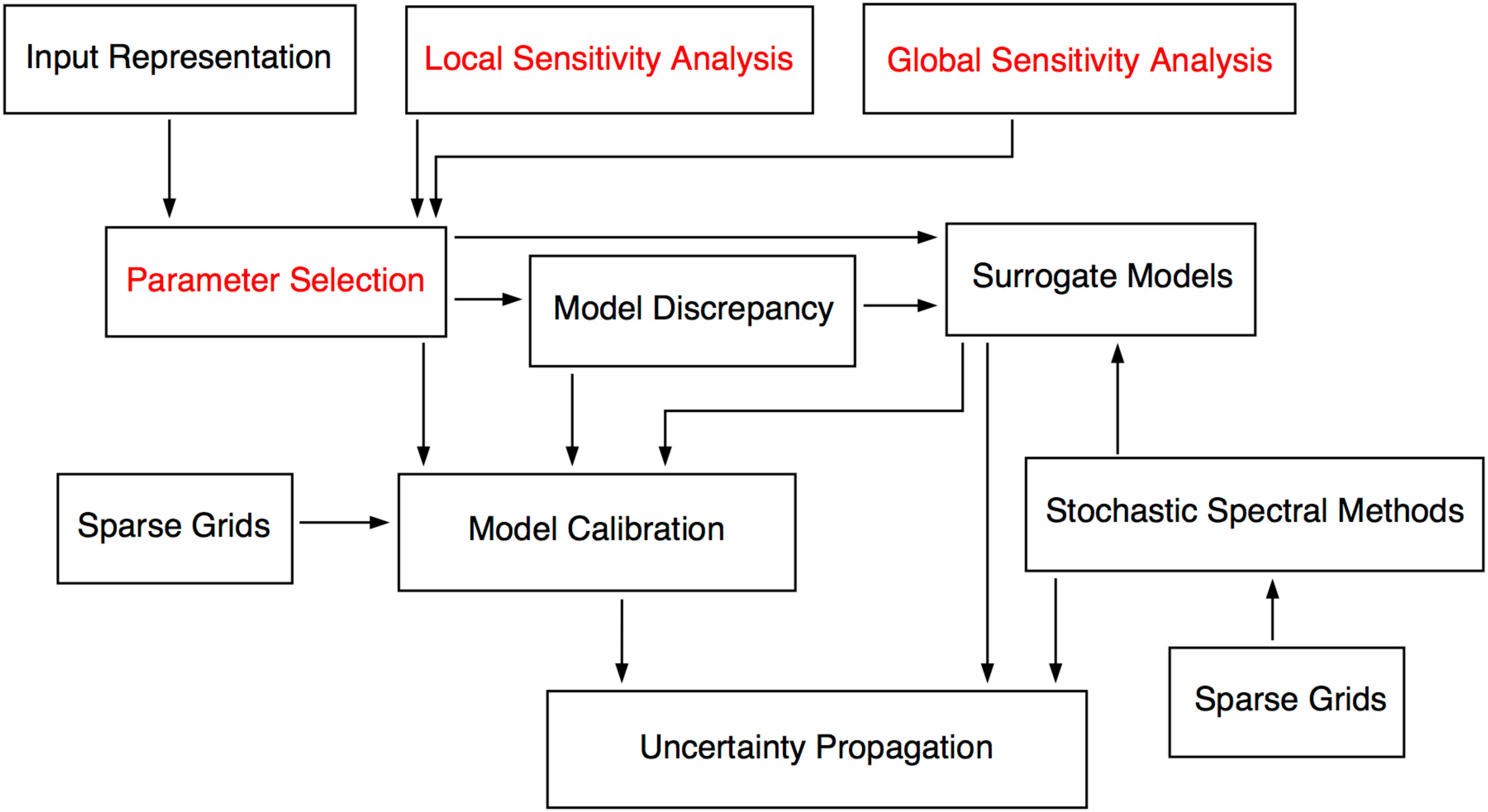
- Temperature uncertainty reduced from 40 degrees to 5 degrees.
- Can run plant 20 degrees hotter, which significantly improves efficiency.
- Warranted continued calibration of closure relations.
- Accommodates disparate data sets.

Potential Ramification: Savings of 10 billion dollars per year for US power plants

Issues:

- We considered only one of many closure relations
- Nuclear regulatory commission takes years to change requirements and codes

Steps in Uncertainty Quantification



Parameter Selection: Required for models with unidentifiable or noninfluential inputs

- e.g., Many closure relations, thermal-hydraulics

Parameter Subset/Subspace Selection

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

Example 1: Spring model

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

$$y(0) = y_0 \, , \; \frac{dy}{dt}(0) = 0$$

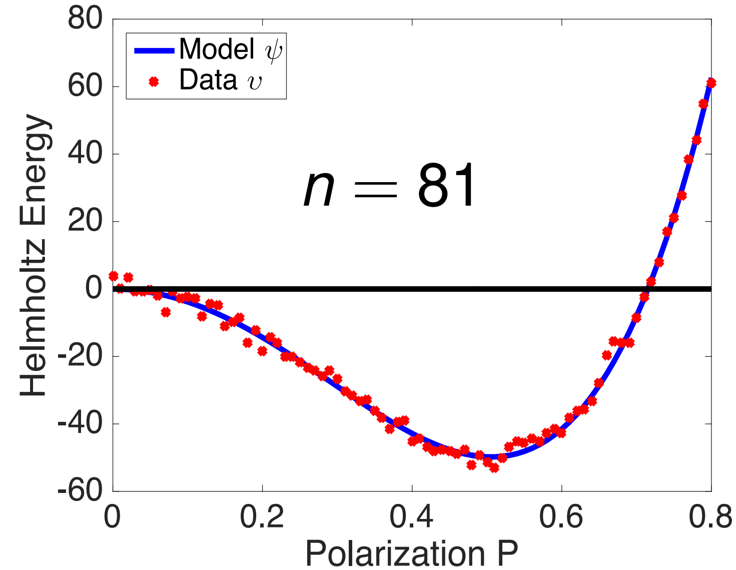
Solution: $y(t,q) = y_0 \cos\left(\sqrt{k/m} \cdot t\right)$

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

Example 2: Helmholtz energy

$$\psi(P) = \alpha_1 P^2 + \alpha_{11} P^4 + \alpha_{111} P^6$$

Question: Are $q = [\alpha_1, \alpha_{11}, \alpha_{111}]$ identifiable for $P \in [0, 0.8]$?



Techniques:

- Global Sensitivity analysis
- Parameter subset selection
- Active subspace techniques (SVD,QR)

Parameter Subset/Subspace Selection

Second Issue: Models can have thousands to millions of parameters

3-D Neutron Transport Equations:

$$
\begin{aligned}
&\frac{1}{|v|}\frac{\partial \varphi}{\partial t} + \Omega \cdot \nabla \varphi + \Sigma_t(r,E)\varphi(r,E,\Omega,t) \\
&\quad = \int_{4\pi} d\Omega' \int_0^\infty dE' \Sigma_s(E' \to E, \Omega' \to \Omega)\varphi(r,E',\Omega',t) \\
&\quad + \underline{\frac{\chi(E)}{4\pi}} \int_{4\pi} d\Omega' \int_0^\infty dE' \underline{\nu(E')}\,\underline{\Sigma_f(E')}\varphi(r,E',\Omega',t)
\end{aligned}
$$

Challenges:

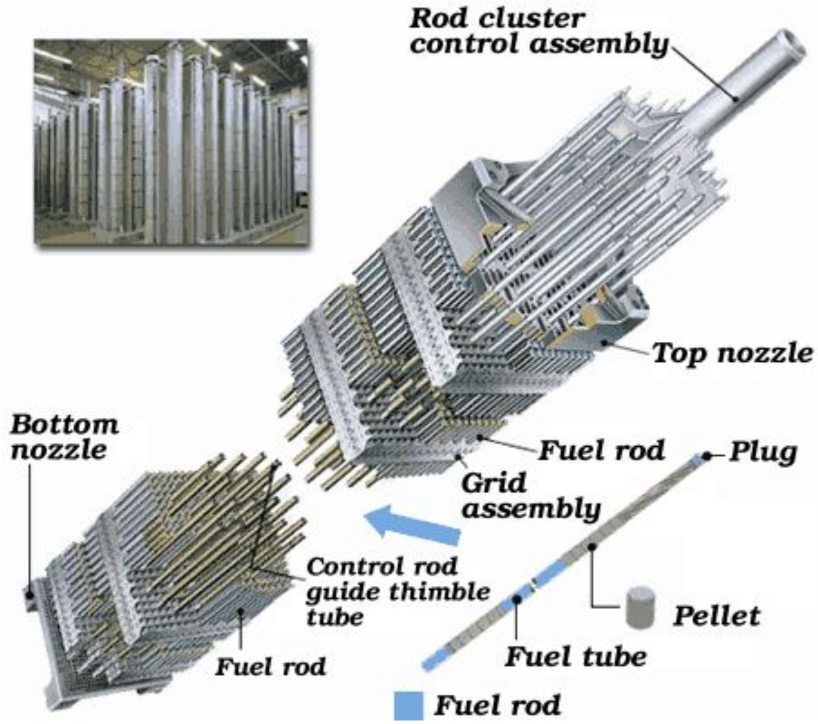
- Very large number of inputs; e.g., 100,000;
Active subspace construction critical.

- ORNL Code SCALE: Can take hours to run

Techniques for General Models:

- Identifiability and sensitivity analysis

- Active Subspaces



Sensitivity Analysis: Motivation

Example: Linear elastic constitute relation

$$\sigma = Ee + c\frac{de}{dt}$$

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:

$$\frac{\partial \sigma}{\partial E} = e = 0.001$$

$$\boxed{\frac{\partial \sigma}{\partial c} = \frac{de}{dt} = 0.1}$$

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: $E = 100,\ c = 0.1$

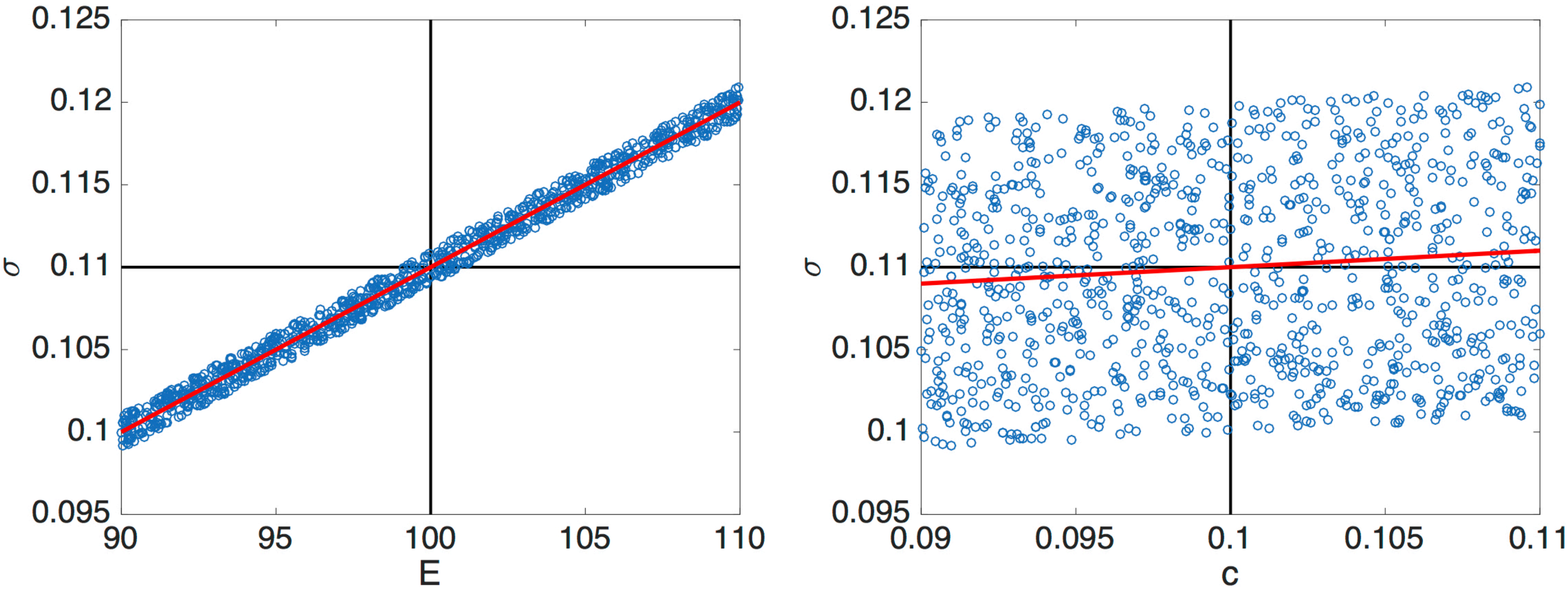
Uncertainty: 10% of nominal values

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

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 constitute relation

$$\sigma = Ee + c\frac{de}{dt}$$

Nominal Values: $E = 100, \; c = 0.1$

Uncertainty: 10% of nominal values

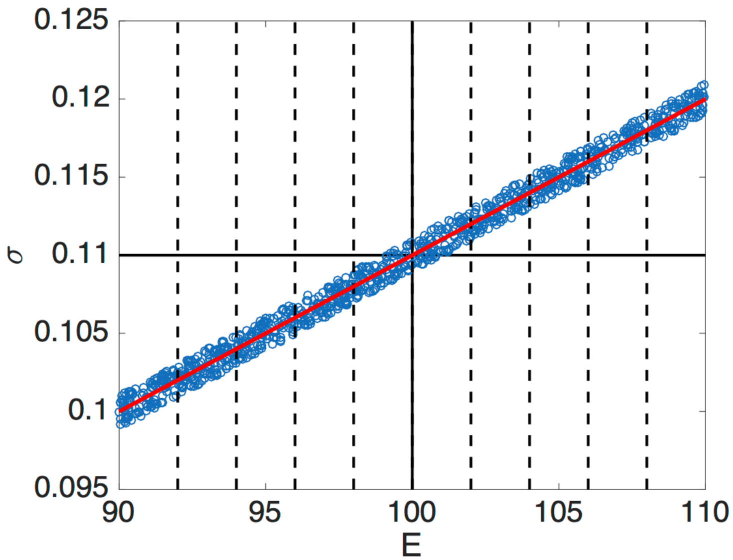
$$E \sim \mathcal{U}(90, 110)\;, \; c \sim \mathcal{U}(0.09, 0.11)$$

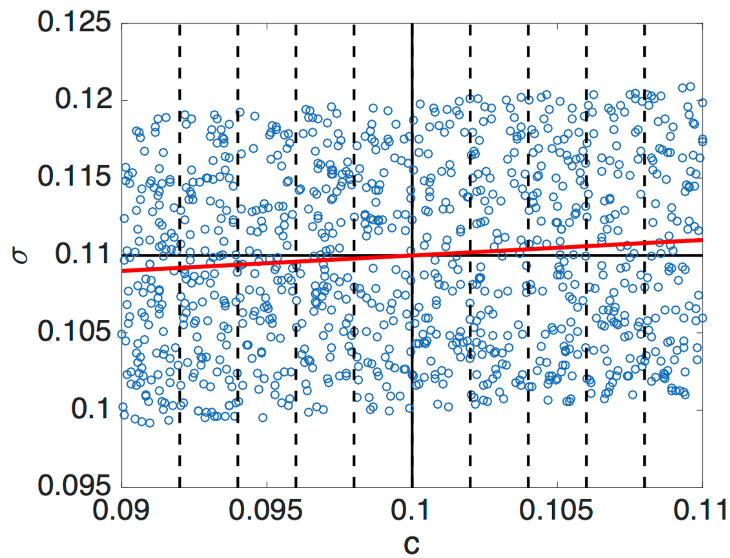
Assumption: Mutually independent parameters

Statistical Interpretation:

$$D_i = \mathrm{var}[\mathbb{E}(Y|q_i)]$$

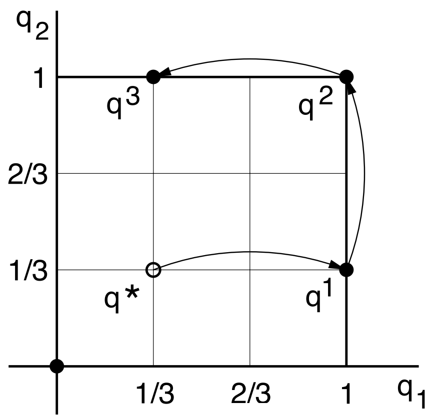
$$S_i = \frac{\mathrm{var}[\mathbb{E}(Y|q_i)]}{\mathrm{var}(Y)}$$

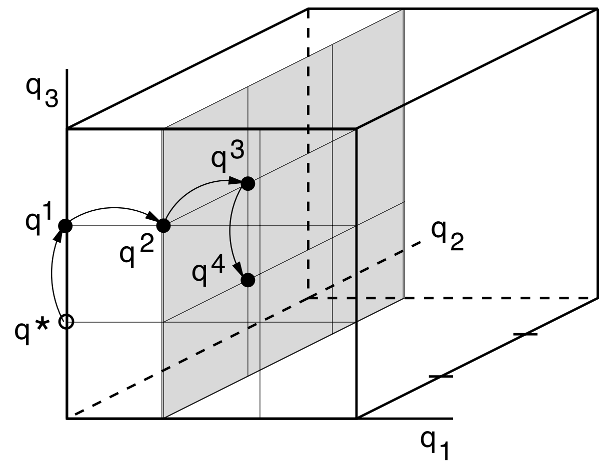




Global Sensitivity Analysis: Morris

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





Elementary Effect:

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

Global Sensitivity Measures: r samples

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

$$\sigma_i^2 = \frac{1}{r-1}\sum_{j=1}^{r}\left(d_i^j(q) - \mu_i\right)^2 \quad , \quad \mu_i = \frac{1}{r}\sum_{j=1}^{r} d_i^j(q)$$

Global Sensitivity Analysis: CASL

Subchannel Code (COBRA-TF): numerous closure relation and parameters

parameter	partial correlation	simple correlation	morris main	morris interaction	CPS variation
k_eta	0.07	0.03			
k_gama	-0.03	0.04			
k_sent	-0.03	-0.02			
k_sdent	-0.07	-0.01			
k_tmasv	-0.03	0.00			
k_tmasl	0.11	0.00	6.48E-05	2.28E-05	medium
k_tmasg	-0.19	-0.01			
k_tmomv	-0.12	-0.01			
k_tmome	0.02	0.00			
k_tmoml	0.02	-0.02	2.23E-04	1.30E-04	medium
k_xk	0.08	-0.02			
k_xkes	-0.05	0.00			
k_xkge	-0.07	0.01			
k_xkl	0.04	-0.01			
k_xkle	-0.03	0.00			
k_xkvls	0.11	-0.01			
k_xkwvw	-0.10	0.01			
k_xkwlw	0.14	0.01			
k_xkwew	-0.01	0.03			
k_qvapl	-0.09	-0.01			
k_tnrgv	-0.03	0.00			
k_tnrgl	-0.01	0.03	9.00E-06	9.49E-06	low
k_rodqq	0.02	-0.01			
k_qradd	-0.02	0.00			
k_qradv	-0.01	0.00			
k_qliht	-0.01	0.00			
k_sphts	-0.05	0.03			
k_cond	-0.04	0.00			
k_xkwvx	0.03	-0.02			
k_xkwlx	1.00	0.88	1.80E-01	7.07E-03	high
k_cd	1.00	0.46	9.59E-02	7.88E-03	high
k_cdfb	-0.02	-0.01			
k_wkr	0.02	0.02			

5 Identified Active Inputs:

k_cd: Pressure loss coefficient of space in sub-channel

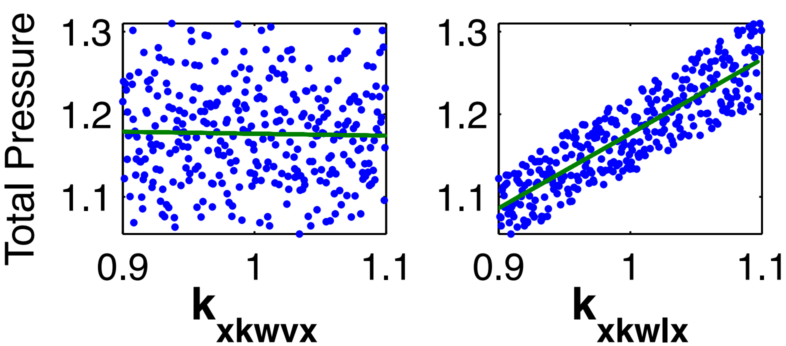
k_xkwlx: Vertical liquid wall drag coefficient

k_tmasl: Loss of liquid mass due to mixing and void drift

k_tmoml: Loss of liquid momentum due to mixing and void drift

k_tnrgl: Loss of liquid enthalpy due to mixing and void drift

Partial Correlation:



Note: 33 initial parameters reduced to 5 via sensitivity analysis

Global Sensitivity Analysis

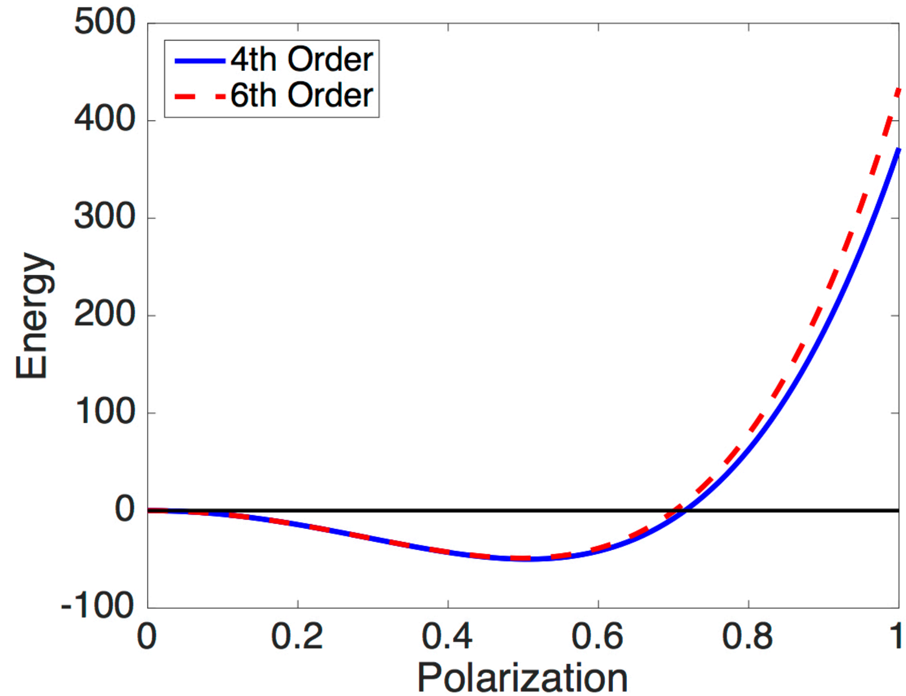
Example: Quantum-informed continuum model

Question: Do we use 4th or 6th-order Landau energy?

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

Parameters:

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



Global Sensitivity Analysis:

	α_1	α_{11}	α_{111}
S_i	0.62	0.39	0.01
S_{T_i}	0.66	0.38	0.06
μ_i^*	0.17	0.07	0.03

Conclusion: α_{111} insignificant and can be fixed

Global Sensitivity Analysis

Example: Quantum-informed continuum model

Question: Do we use 4th or 6th-order Landau energy?

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

Parameters:

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

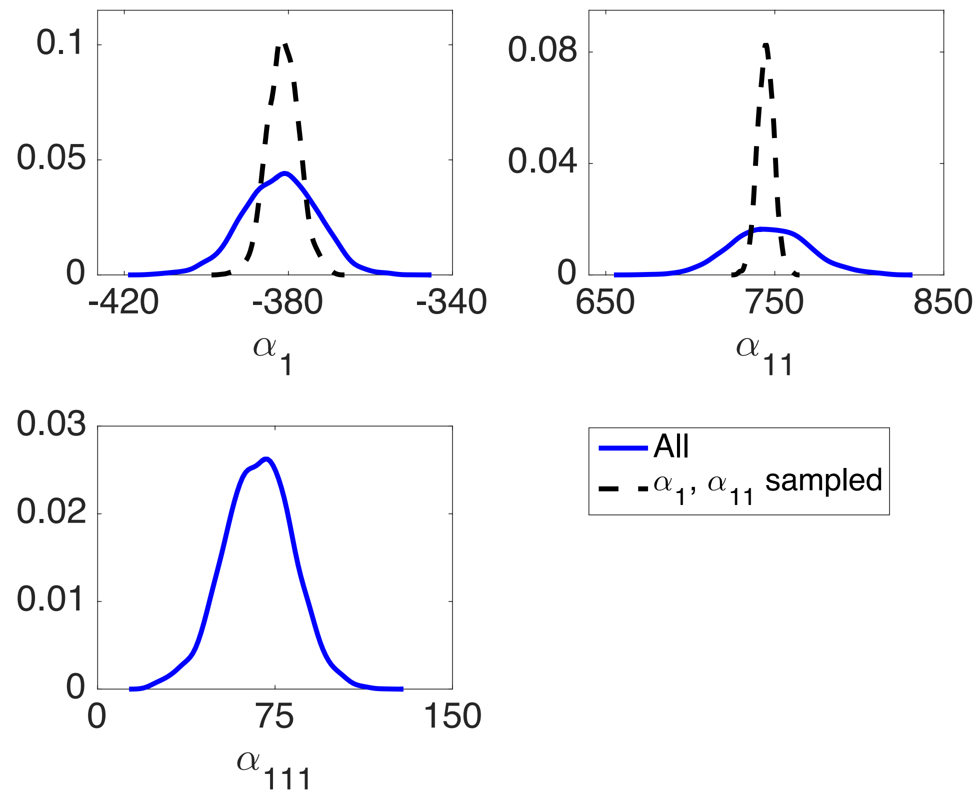
Global Sensitivity Analysis:

	α_1	α_{11}	α_{111}
S_i	0.62	0.39	0.01
S_{T_i}	0.66	0.38	0.06
μ_i^*	0.17	0.07	0.03

Conclusion:

α_{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 4th or 6th-order Landau energy?

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

Parameters:

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

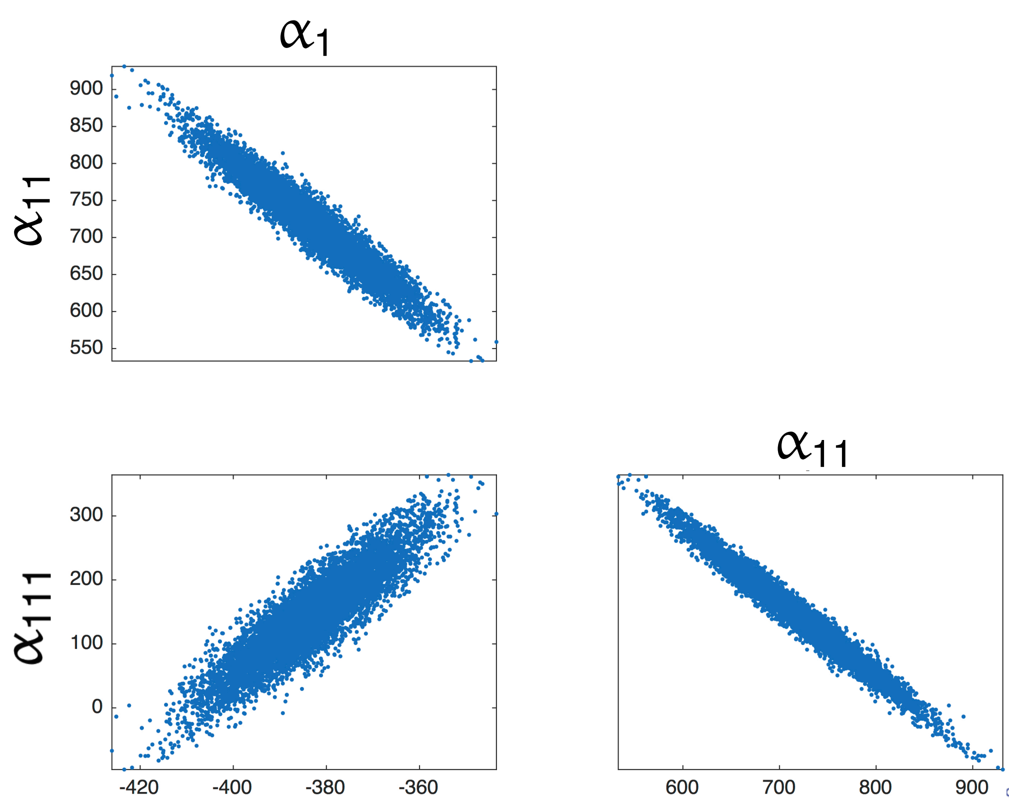
Global Sensitivity Analysis:

	α_1	α_{11}	α_{111}
S_k	0.62	0.39	0.01
T_k	0.66	0.38	0.06
μ_k^*	0.17	0.07	0.03

Note: Must accommodate correlation

Problem:

- Parameters correlated
- Cannot fix α_{111}



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 $\upsilon_i \approx \psi(P_i, q^*)$

$$\begin{aligned} J(q) &= \frac{1}{n}\sum_{i=1}^{n} [\upsilon_i - \psi(P_i, q)]]^2 \\ &\approx \frac{1}{n}\sum_{i=1}^{n} [\nabla_q \psi(P_i, q^*) \cdot \Delta q]^2 \\ &= \frac{1}{n}(\chi \Delta q)^T (\chi \Delta q) \end{aligned}$$

Sensitivity Matrix:

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

Note:

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

Parameter Subset Selection

Note:

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

Strategy: Take Δq to be eigenvector of $\boxed{\chi^T \chi}$ Fisher Information

$$\Rightarrow \chi^T \chi \Delta q = \lambda \Delta q$$

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

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

$\Rightarrow$ Nonidentifiable

Note: Estimator for covariance matrix

$$V = s^2 \left[\chi^T \chi\right]^{-1} = \begin{bmatrix} \mathrm{var}(q_1) & \mathrm{cov}(q_1, q_2) & \cdots & \mathrm{cov}(q_1, q_n) \\ \mathrm{cov}(q_2, q_1) & \mathrm{var}(q_2) & \mathrm{cov}(q_2, q_3) & \\ \vdots & & & \vdots \\ \mathrm{cov}(q_n, q_1) & \cdots & \cdots & \mathrm{var}(q_n) \end{bmatrix}$$

Parameter Subset Selection

Note:

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

Strategy: Take Δq to be eigenvector of $\boxed{\chi^T \chi}$ 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 $J(q^* + \Delta q) \approx 0$

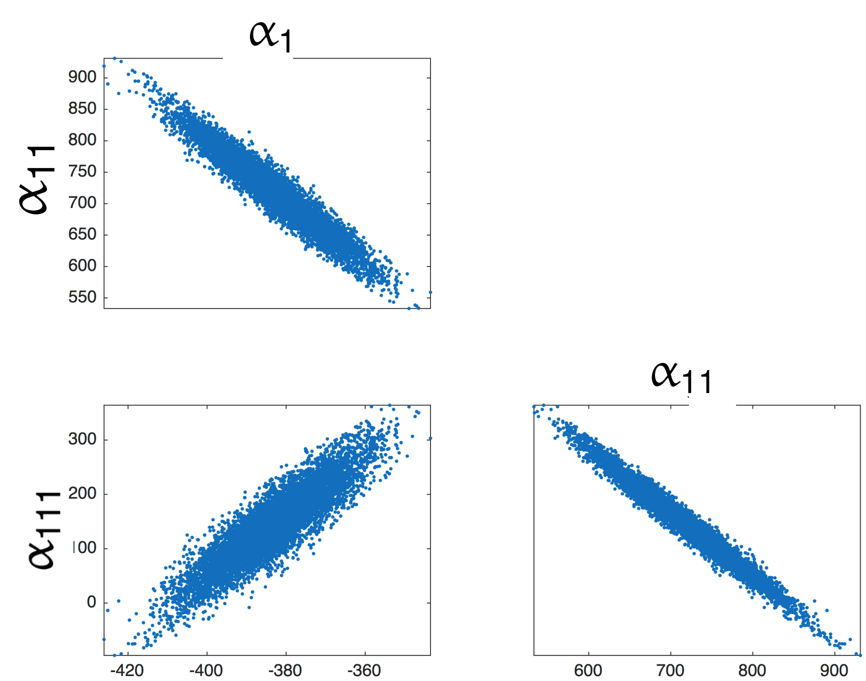
$\Rightarrow$ Nonidentifiable

Example:

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

Parameters:

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



Result: rank($\chi^T \chi$) = 3 so all parameters identifiable

Active Subspaces

Note:

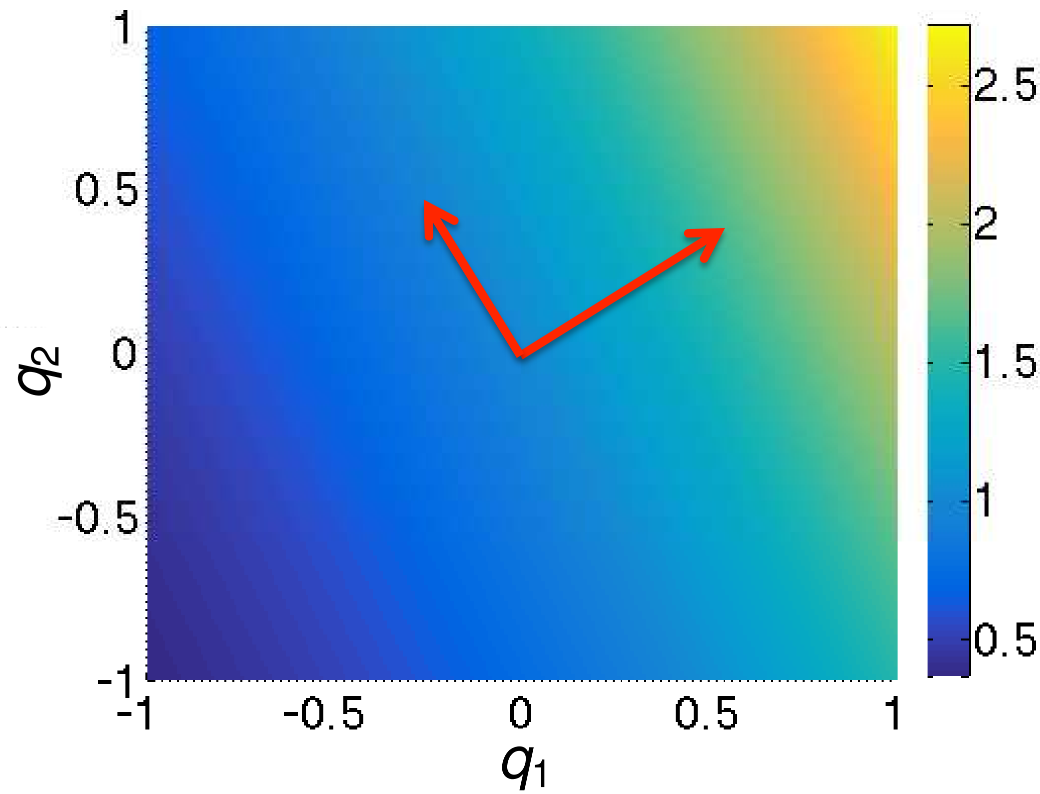
- Functions may vary significantly in only a few directions
- “Active” directions may be linear combination of inputs

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

- Varies most in [0.7, 0.3] direction
- No variation in orthogonal direction

Strategy:

- *Linearly parameterized problems:* Employ SVD or QR decomposition.
- *Nonlinear problems:* Construct approximate gradient matrix and employ SVD or QR.



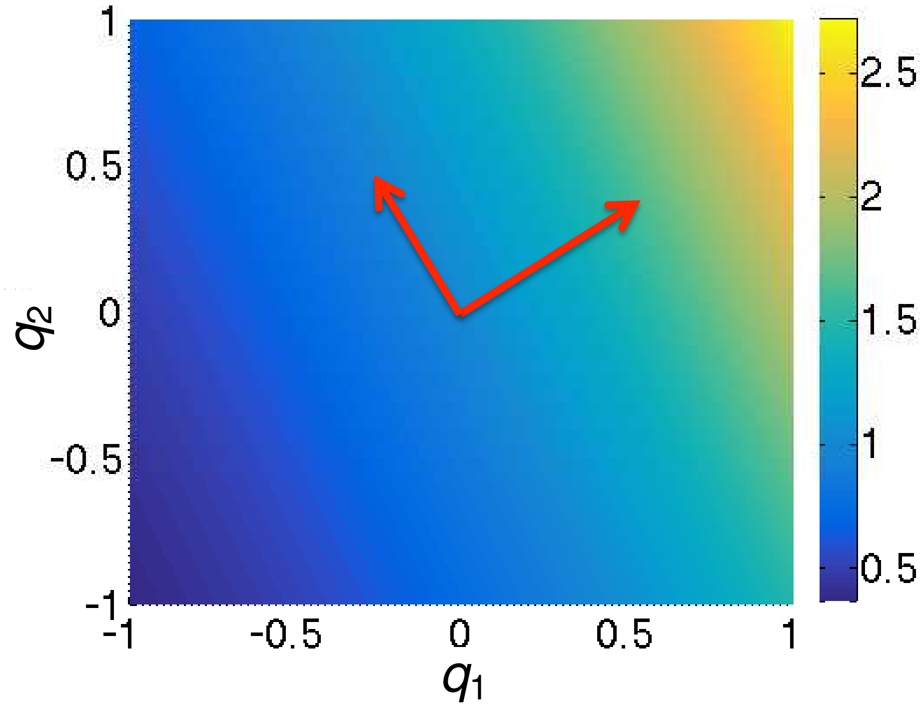
Active Subspaces

Note:

- Functions may vary significantly in only a few directions
- "Active" directions may be linear combination of inputs

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

- Varies most in [0.7, 0.3] direction
- No variation in orthogonal direction



A Bit of History:

- Often attributed to Russi (2010).
- Concept same as *identifiable subspaces* from systems and control; e.g., Reid (1977).
- For linearly parameterized problems, active subspace given by SVD or QR; Beltrami (1873), Jordan (1874), Sylvester (1889), Schmidt (1907), Weyl (1912). See 1993 *SIAM Review* paper by Stewart.

Linear Problems

Second Issue: Models depends on very large number of parameters – e.g., millions – but only a few are "significant".

Linear Algebra Techniques: Linearly parameterized problems

$$y = Aq \ , \ q \in \mathbb{R}^p \ , \ y \in \mathbb{R}^m$$

Singular Value Decomposition (SVD):

$$A = U\Sigma V^T \ , \ \Sigma = [S \quad 0]$$

$$S = \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & 0 \end{bmatrix} \ , \ \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \geq \varepsilon$$

Rank Revealing QR Decomposition: $A^T P = QR$

Problem: Neither is directly applicable when m or p are very large; e.g., millions.

Solution: Random range finding algorithms.

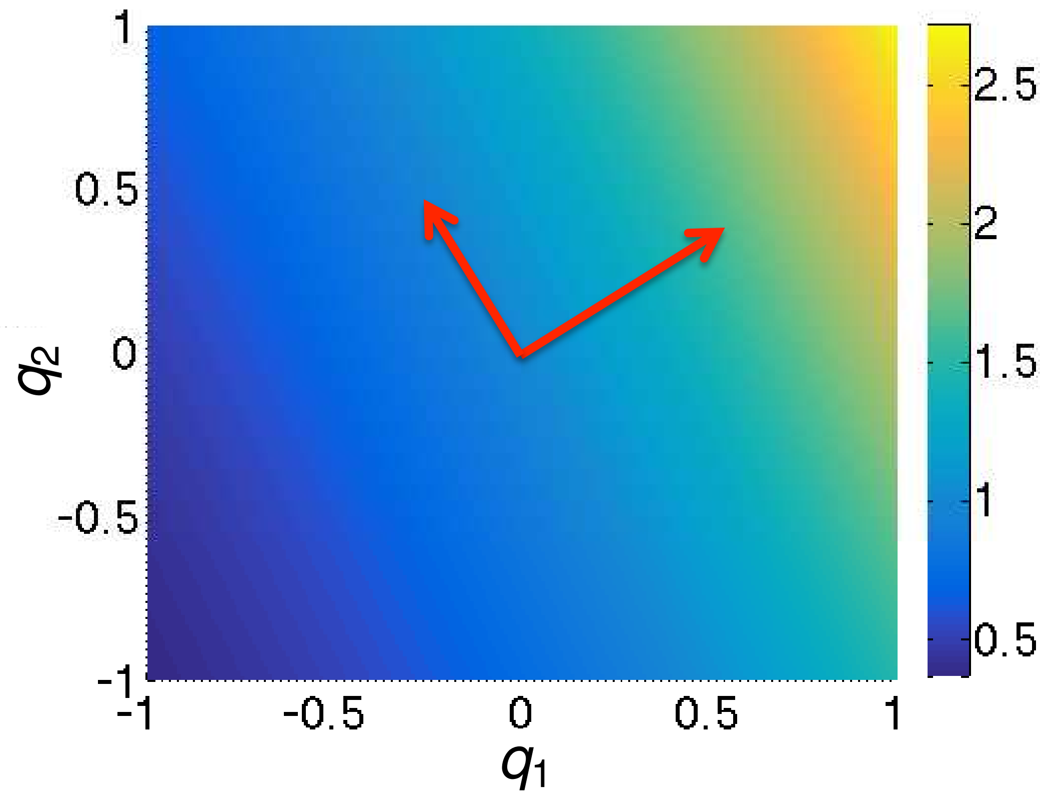
Active Subspaces

Note:

- Functions may vary significantly in only a few directions
- "Active" directions may be linear combination of inputs

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

- Varies most in [0.7, 0.3] direction
- No variation in orthogonal direction



Strategy:

- *Linearly parameterized problems:* Employ SVD or QR decomposition.
- *Nonlinear problems:* Construct approximate gradient matrix and employ SVD or QR.

Gradient-Based Active Subspace

Active Subspace: Consider

$$f = f(q)\ ,\, q \in \mathcal{Q} \subseteq \mathbb{R}^p$$

- E.g., see [Constantine, SIAM, 2015; Stoyanov & Webster, *IJUQ*, 2015]

and

$$\nabla_q f(q) = \left[\frac{\partial f}{\partial q_1}, \cdots, \frac{\partial f}{\partial q_p}\right]^T$$

Construct outer product

$\rho(q)$: Distribution of input parameters q

$$C = \int (\nabla_q f)(\nabla_q f)^T \rho dq$$

Question: How sensitive are results to distribution, which is typically not known?

Partition eigenvalues: $C = W \Lambda W^T$

$$\Lambda = \begin{bmatrix} \Lambda_1 & \\ & \Lambda_2 \end{bmatrix}\ ,\quad W = [W_1 \quad W_2]$$

Rotated Coordinates:

$$y = W_1^T q \in \mathbb{R}^n \quad \text{and} \quad z = W_2^T q \in \mathbb{R}^{p-n}$$

Active Variables

Active Subspace: Range of eigenvectors in W_1

Gradient-Based Active Subspace

Active Subspace: Construction based on random sampling

1. Draw M independent samples $\{q^j\}$ from ρ

2. Evaluate $\nabla_q f_j = \nabla_q f(q^j)$

3. Approximate outer product

$$C \approx \widetilde{C} = \frac{1}{M} \sum_{j=1}^{M} (\nabla_q f_j)(\nabla_q f_j)^T$$ Monte Carlo Quadrature

Note: $\widetilde{C} = GG^T$ where $G = \frac{1}{\sqrt{M}}[\nabla_q f_1, \ldots, \nabla_q f_M]$

4. Take SVD of $G = W\sqrt{\Lambda}V^T$

- Active subspace of dimension n is first n columns of W

Current Research: Develop efficient algorithms for codes that do not have adjoint capabilities

Note: Finite difference approximations tempting but not very effective

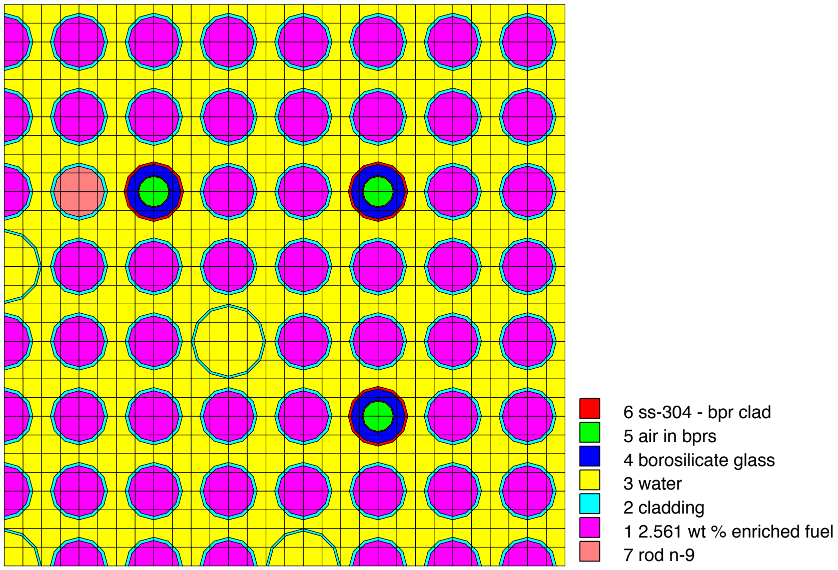
Strategy: Algorithm based on initialized adaptive Morris indices

SCALE6.1: High-Dimensional Example

Setup: Cross-section computations SCALE6.1

- Input Dimension: 7700
- Output k_{eff}

Materials			Reactions	
$^{234}_{92}\mathrm{U}$	$^{10}_{5}\mathrm{B}$	$^{31}_{15}\mathrm{P}$	Σ_t	$n \rightarrow \gamma$
$^{235}_{92}\mathrm{U}$	$^{11}_{5}\mathrm{B}$	$^{55}_{25}\mathrm{Mn}$	Σ_e	$n \rightarrow p$
$^{236}_{92}\mathrm{U}$	$^{14}_{7}\mathrm{N}$	$_{26}\mathrm{Fe}$	Σ_f	$n \rightarrow d$
$^{238}_{92}\mathrm{U}$	$^{15}_{7}\mathrm{N}$	$^{116}_{50}\mathrm{Sn}$	Σ_c	$n \rightarrow t$
$^{1}_{1}\mathrm{H}$	$^{23}_{11}\mathrm{Na}$	$^{120}_{50}\mathrm{Sn}$	$\bar{\nu}$	$n \rightarrow {}^{3}\mathrm{He}$
$^{16}_{8}\mathrm{O}$	$^{27}_{13}\mathrm{Al}$	$_{40}\mathrm{Zr}$	χ	$n \rightarrow \alpha$
$_{6}\mathrm{C}$	$_{14}\mathrm{Si}$	$_{19}\mathrm{K}$	$n \rightarrow n'$	$n \rightarrow 2n$



PWR Quarter Fuel Lattice

Note: Requires efficient initialization algorithm.

SCALE6.1: High-Dimensional Example

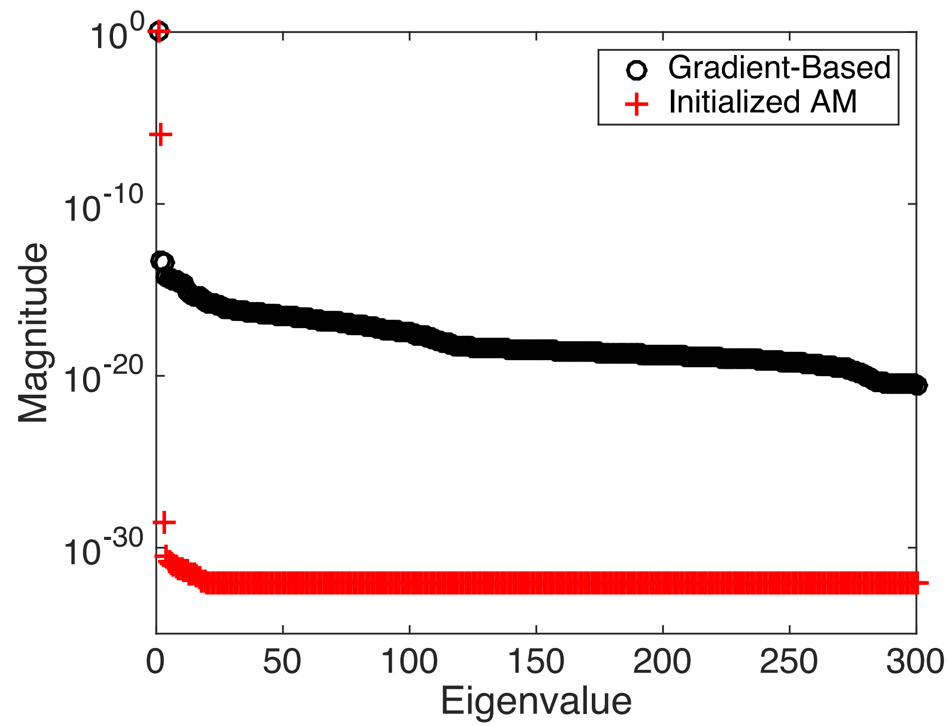
Setup:

- Input Dimension: 7700

SCALE Evaluations:

- Gradient-Based: 1000
- Initialized Adaptive Morris: 18,392
- Projected Finite-Difference: 7,701,000

Note: Analytic eigenvalues: 0, 1

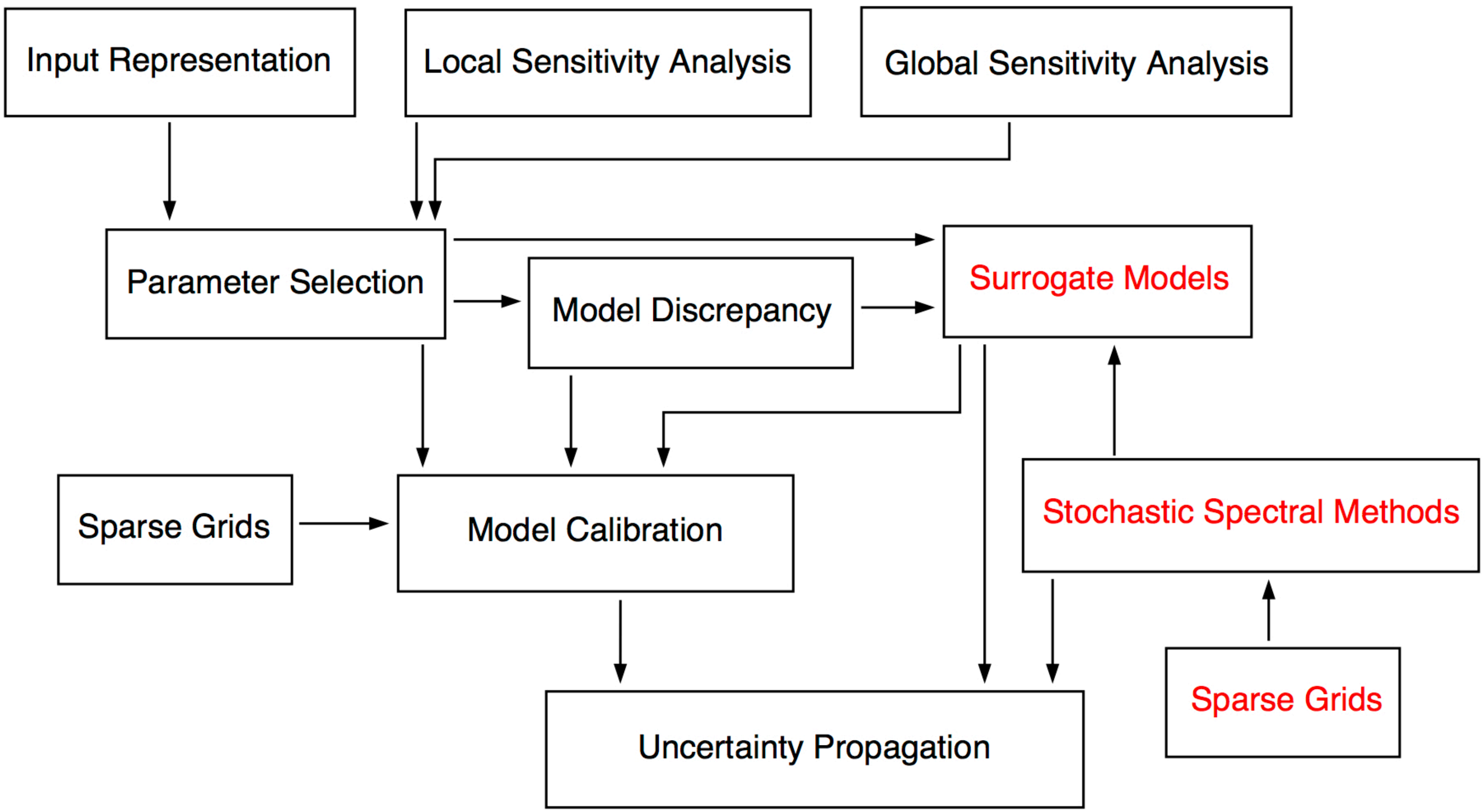


For surrogate sampled off space

Active Subspace Dimensions:

	Gap	PCA				Error Tolerance			
Method		0.75	0.90	0.95	0.99	10^{-3}	10^{-4}	10^{-5}	10^{-6}
Gradient-Based	1	2	6	9	24	1	13	90	233
Initialized AM	1	1	1	1	2	1	2	2	2

Steps in Uncertainty Quantification



Challenge:

- How do we do uncertainty quantification for computationally expensive models?

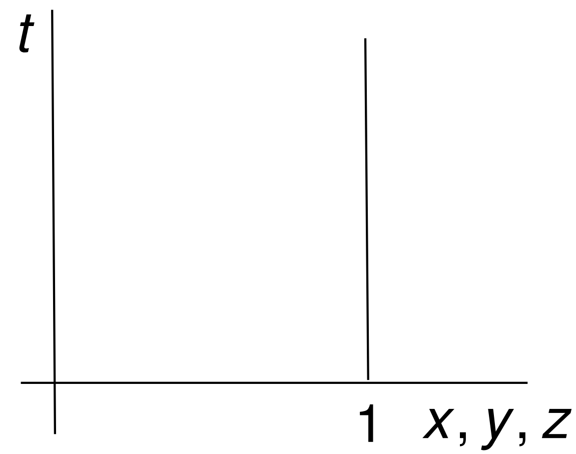
Surrogate Models: Motivation

Example: Consider the heat equation

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

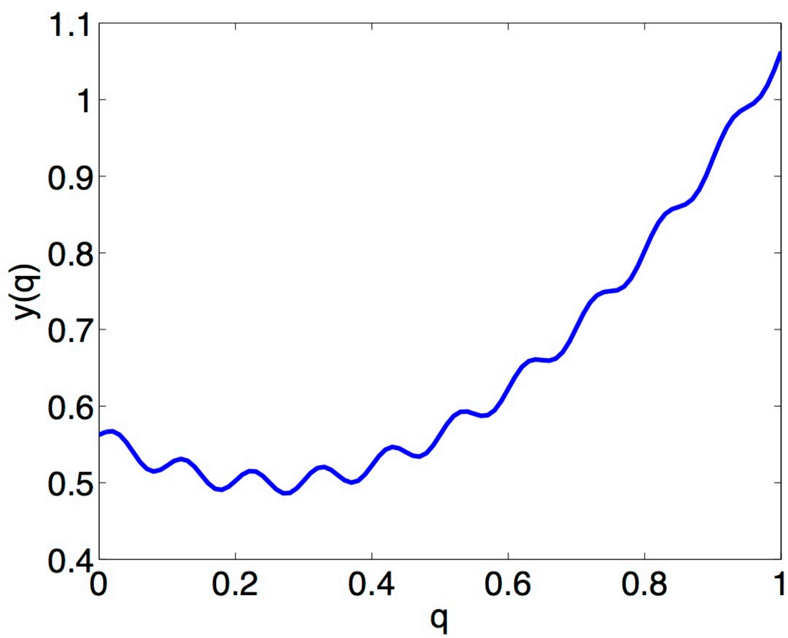
Boundary Conditions

Initial Conditions



with the response

$$y(q) = \int_0^1 \int_0^1 \int_0^1 \int_0^1 u(t,x,y,z)\,dxdydzdt$$



Notes:

- Requires approximation of PDE in 3-D
- What would be a simple surrogate?

Surrogate Models: Motivation

Example: Consider the heat equation

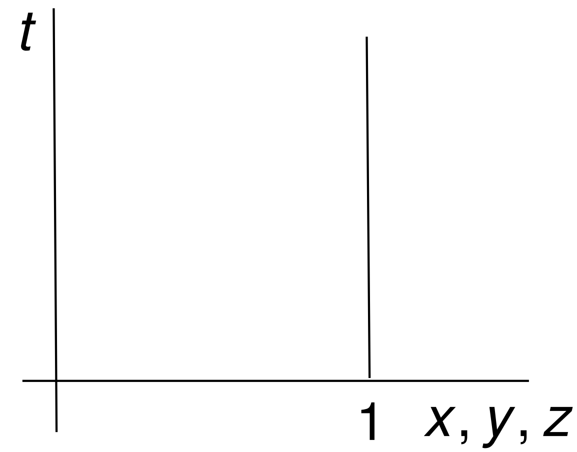
$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} + f(q)$$

Boundary Conditions

Initial Conditions

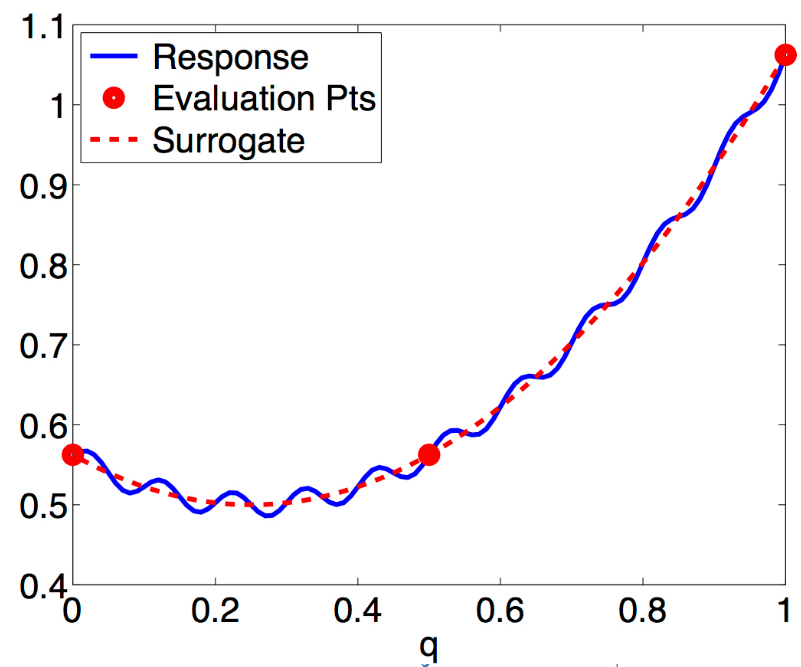
with the response

$$y(q) = \int_0^1 \int_0^1 \int_0^1 \int_0^1 u(t,x,y,z)\,dxdydzdt$$



Surrogate: Quadratic

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



Question: How do you construct a polynomial surrogate?

- Regression
- Interpolation

Surrogate Models: Motivation

Example: Consider the heat equation

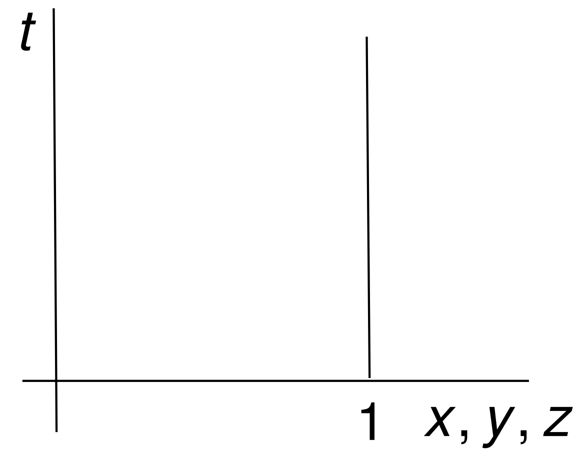
$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} + f(q)$$

Boundary Conditions

Initial Conditions

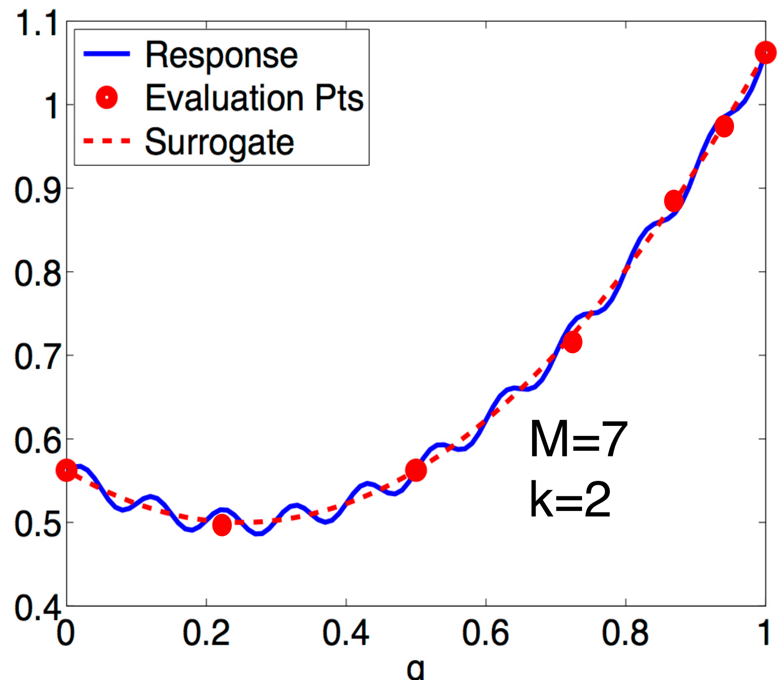
with the response

$$y(q) = \int_0^1 \int_0^1 \int_0^1 \int_0^1 u(t,x,y,z)\,dxdydzdt$$



Surrogate: Quadratic

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



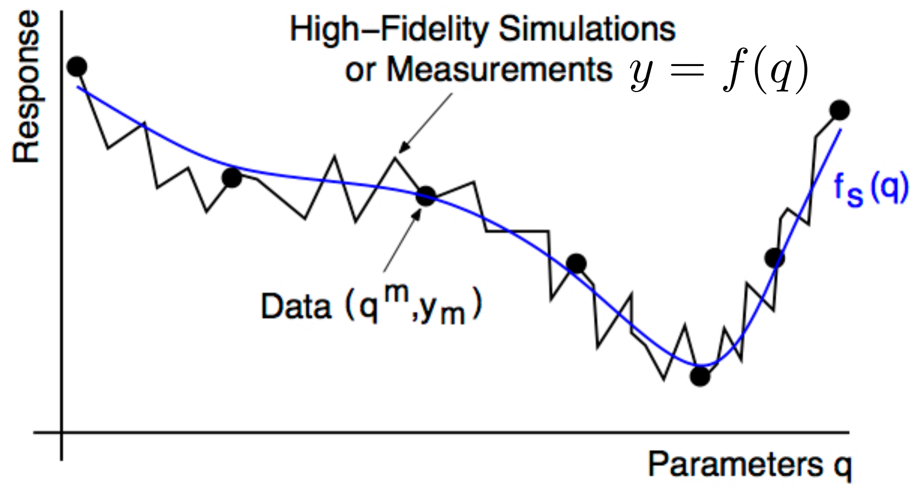
Question: How do you construct a polynomial surrogate?

- Interpolation
- Regression

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.



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 $f(q)$

$$y_m = f_s(q^m) + \varepsilon_m\ ,\ \ m = 1, \ldots, M$$

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

$$\mathbb{E}[\Psi_0(Q)] = 1$$

$$\begin{aligned}\mathbb{E}[\Psi_i(Q)\Psi_j(Q)] &= \int_\Gamma \Psi_i(q)\Psi_j(q)\rho(q)dq \\ &= \delta_{ij}\gamma_i\end{aligned}$$

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 α_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

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

Note: Sample points $\{q^m\}_{m=1}^{M}$

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

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

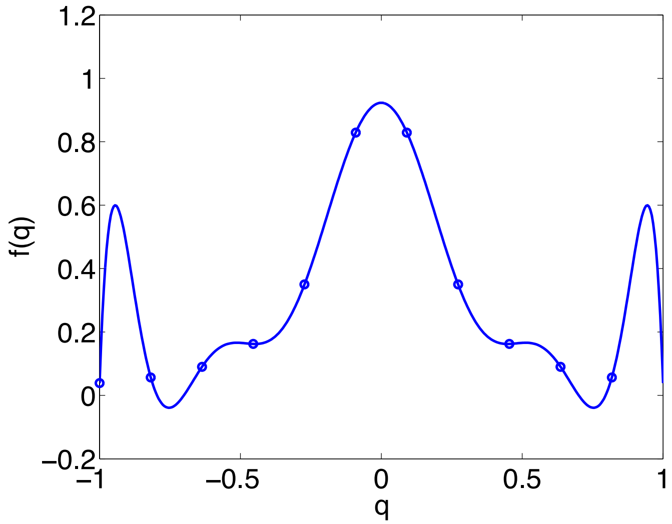
e.g., SPGL1

- MATLAB Solver for large-scale sparse reconstruction

Surrogate Models: Grid Choices

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

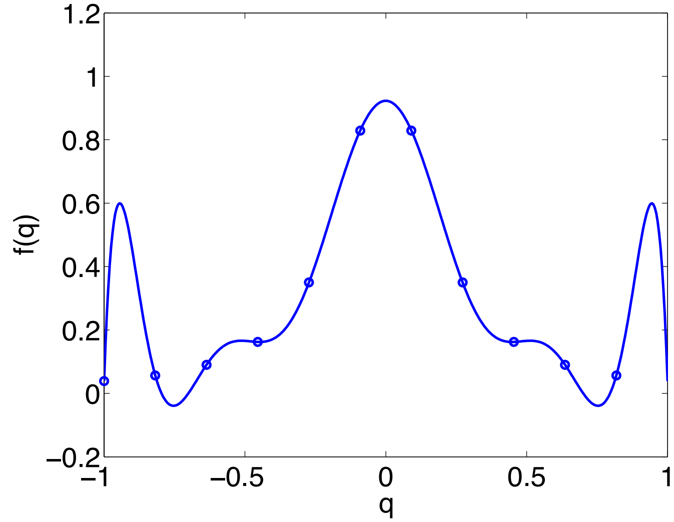
$$q^j = -1 + (j-1)\frac{2}{M} \; , \; j = 1, \ldots, M$$

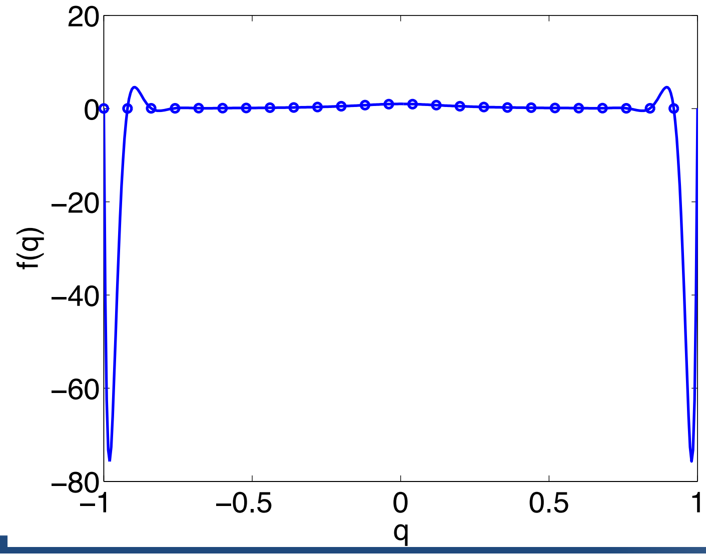


Surrogate Models: Grid Choices

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

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

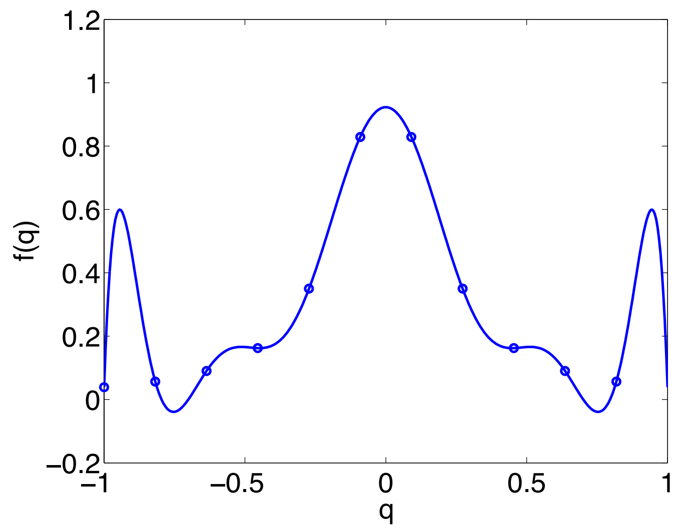


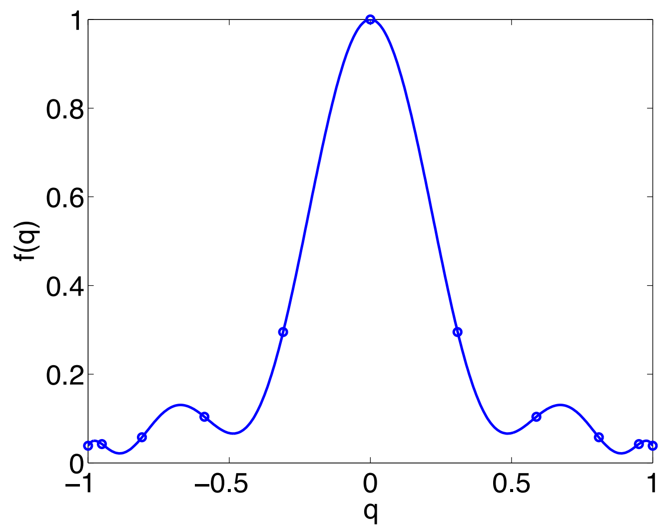


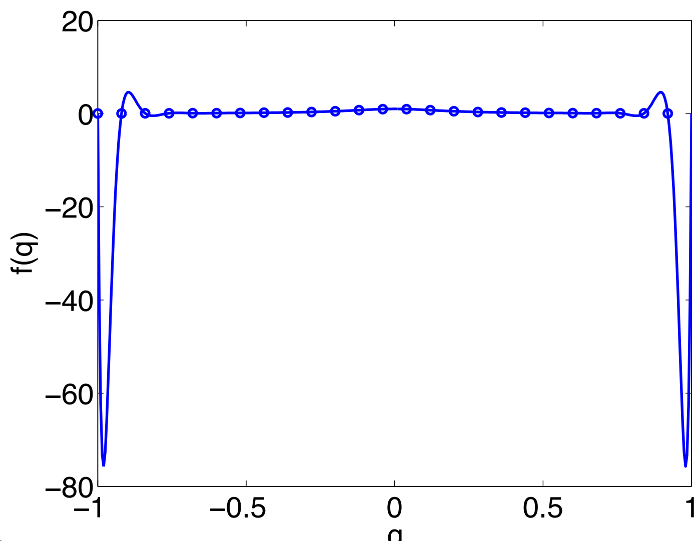
Surrogate Models: Grid Choices

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

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



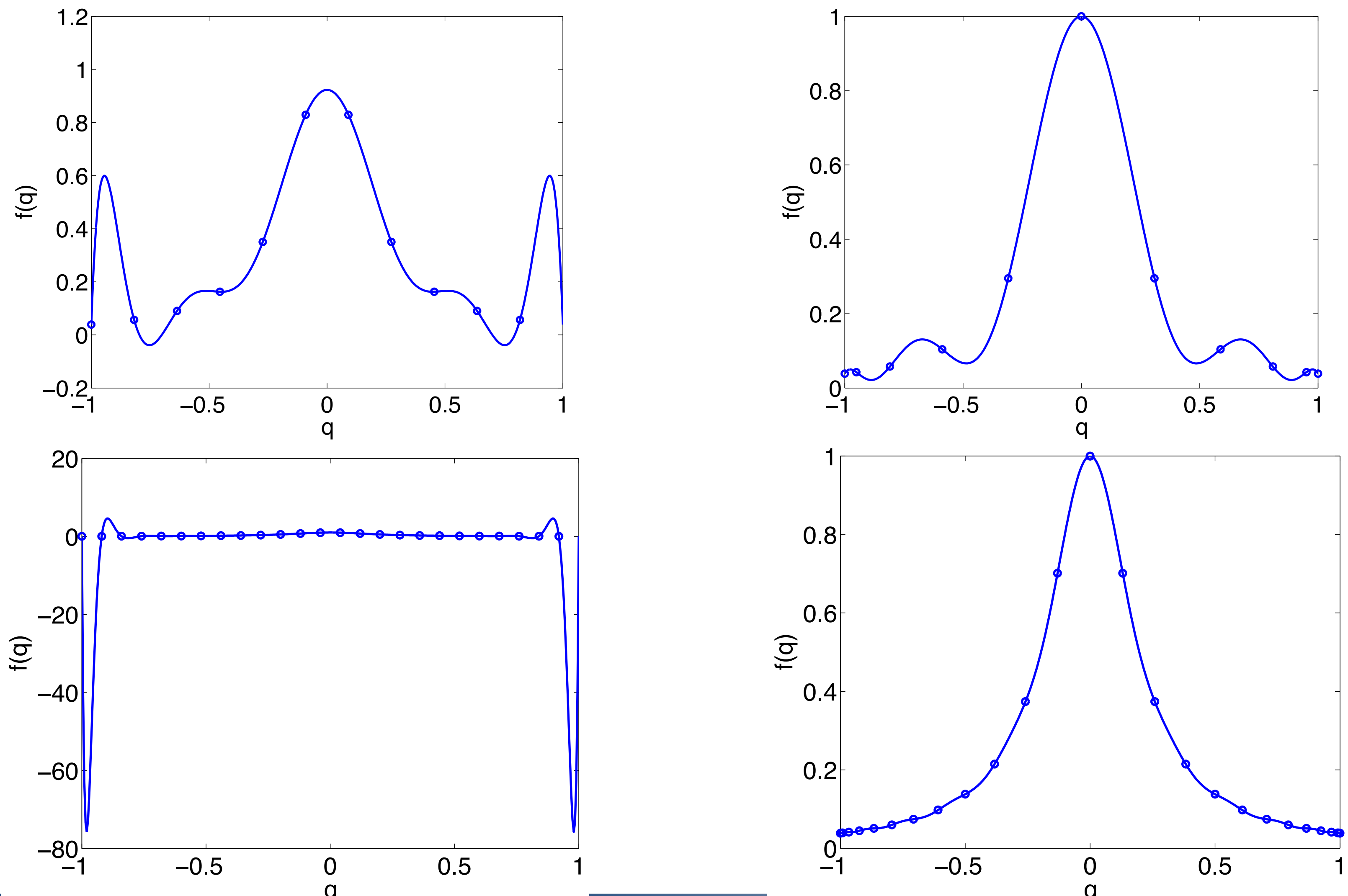




Surrogate Models: Grid Choices

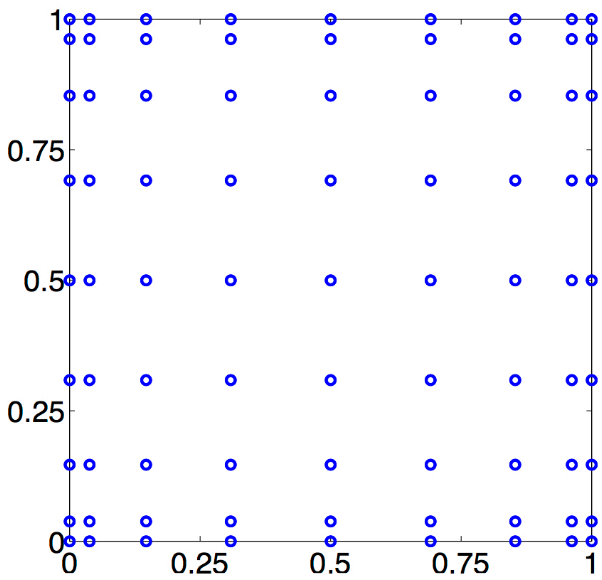
Example: Consider the Runge function $f(q) = \frac{1}{1+25q^2}$ with points

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

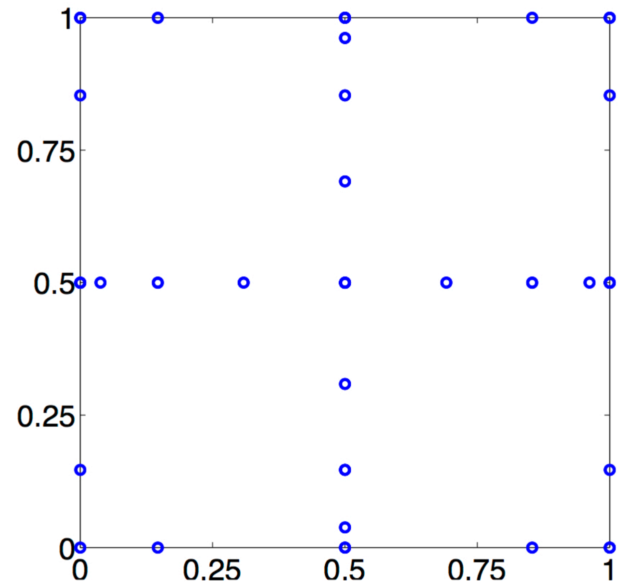


Sparse Grid Techniques

Tensored Grids: Exponential growth



Sparse Grids: Same accuracy



p	R_ℓ	Sparse Grid $\mathcal{R}$	Tensored Grid $R = (R_\ell)^p$
2	9	29	81
5	9	241	59,049
10	9	1581	$> 3 \times 10^9$
50	9	171,901	$> 5 \times 10^{47}$
100	9	1,353,801	$> 2 \times 10^{95}$

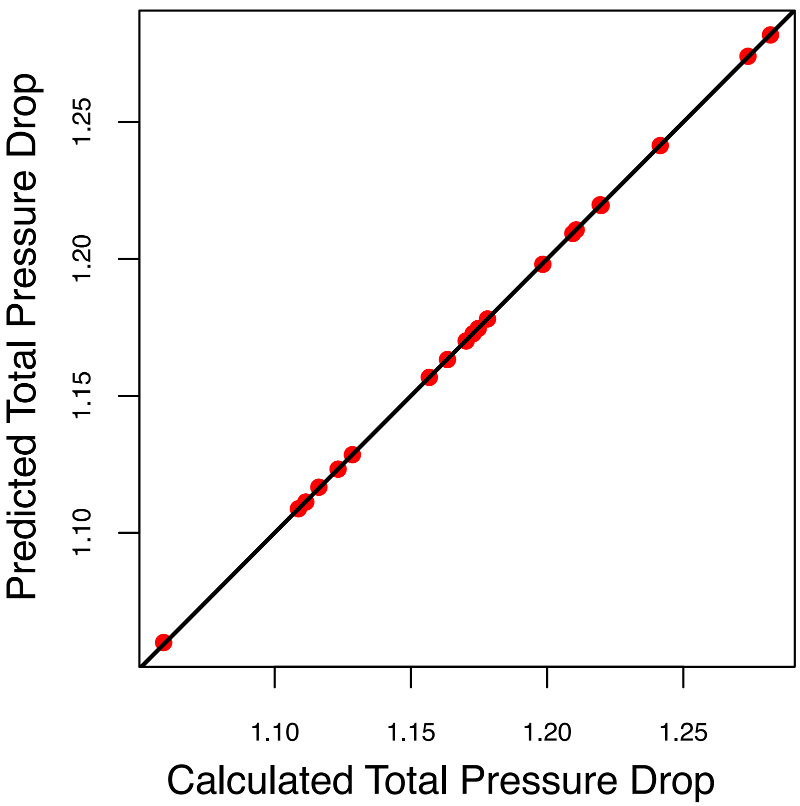
Surrogate Construction: CASL

Subchannel Code (COBRA-TF): 33 VUQ parameters reduced to 5 using SA

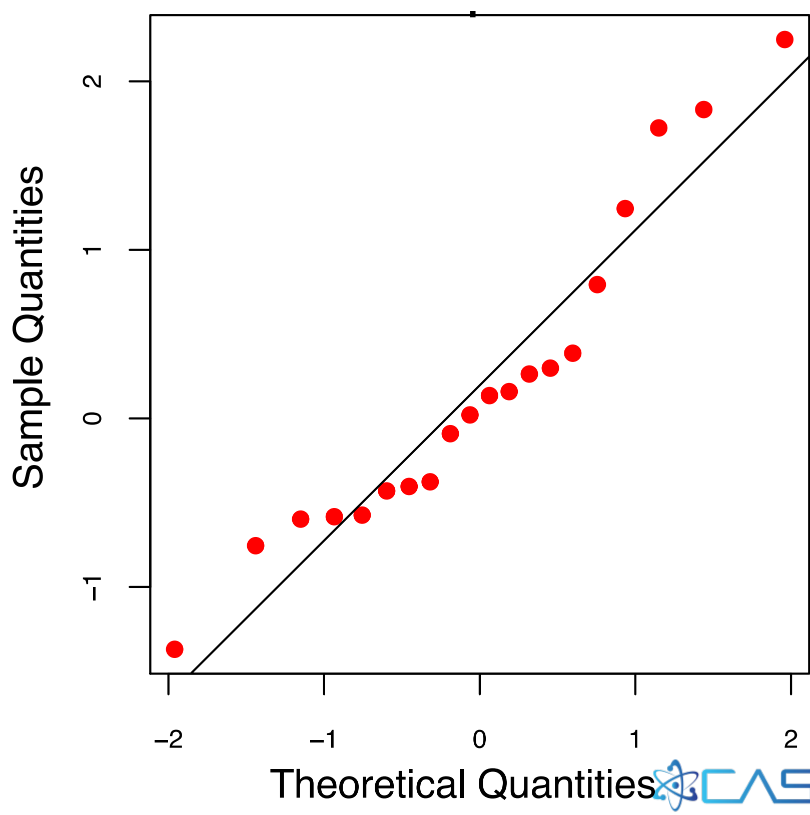
Surrogate: Total pressure drop

• Kriging (GP) emulator constructed using 50 COBRA-TF runs perturbing 5 active inputs.

• Use remaining computational budget to evaluate quality of surrogate using post-processed Dakota outputs.



Out-of-Sample Validation



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.

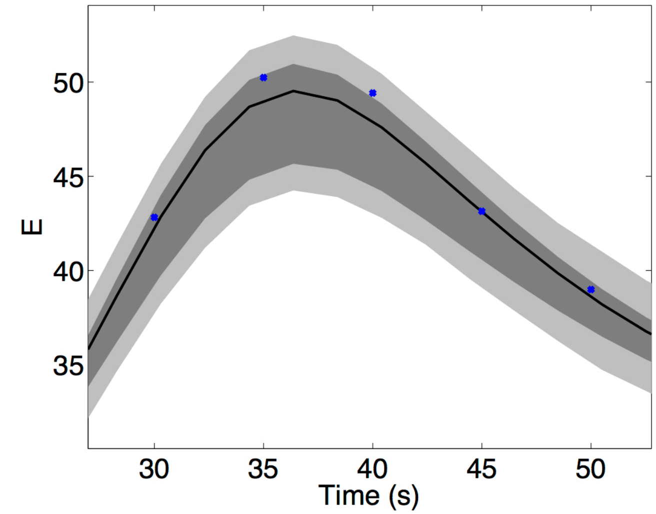
• Goal is to predict model responses with quantified and reduced uncertainties.

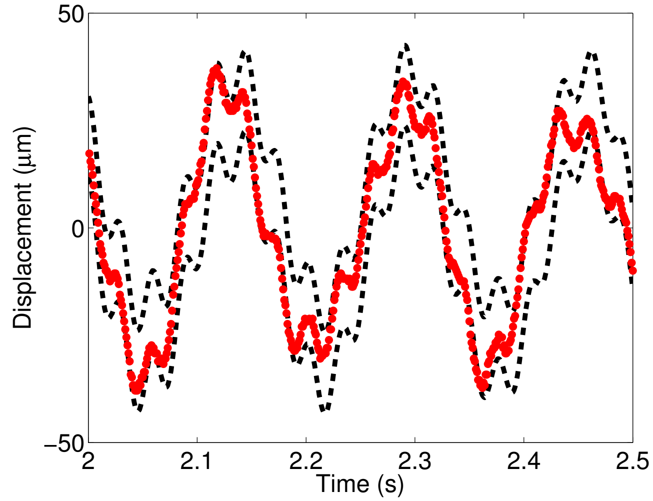
• Parameter selection is critical to isolate identifiable and influential parameters.

• Surrogate models critical for computationally intensive simulation codes.

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

• *Prediction is very difficult, especially if it's about the future*, Niels Bohr.





A DOE Energy Innovation Hub

www.casl.gov