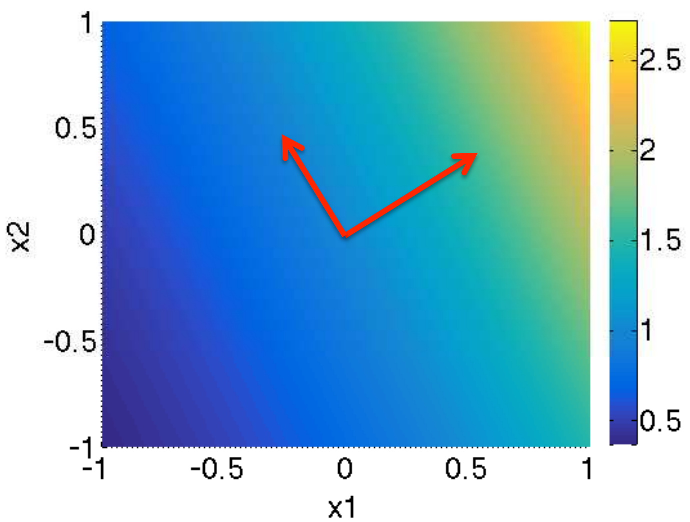
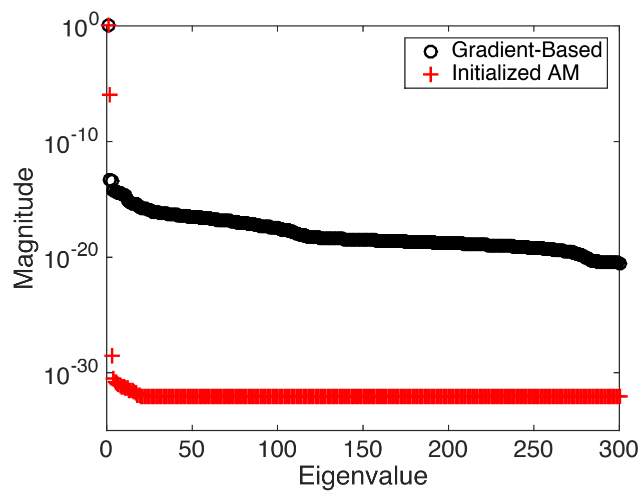
Sensitivity Analysis and Active Subspace Construction for Surrogate Models Employed for Bayesian Inference

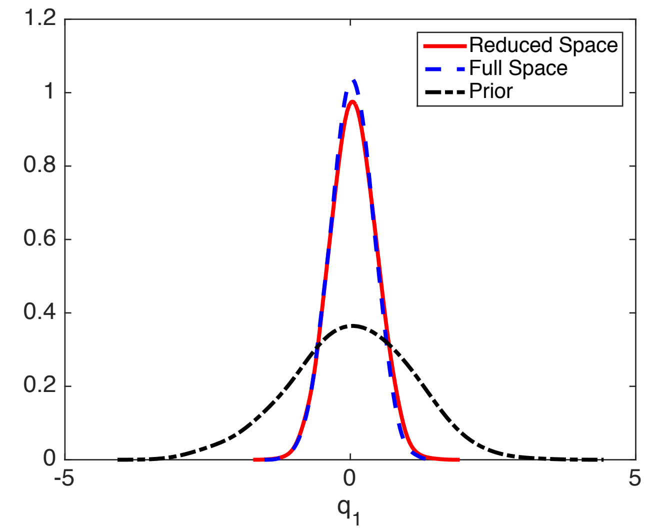
Ralph C. Smith

Department of Mathematics

North Carolina State University







Support: DOE Consortium for Advanced Simulation of LWR (CASL)

NNSA Consortium for Nonproliferation Enabling Capabilities (CNEC)

NSF Grant CMMI-1306290, Collaborative Research CDS&E

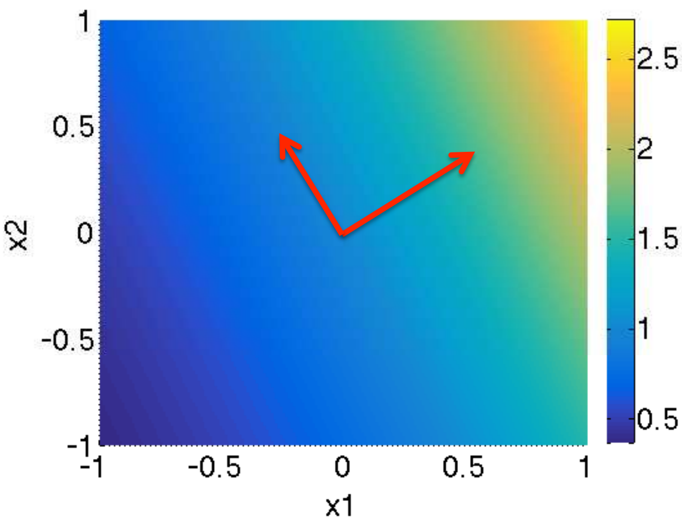
AFOSR Grant FA9550-15-1-0299

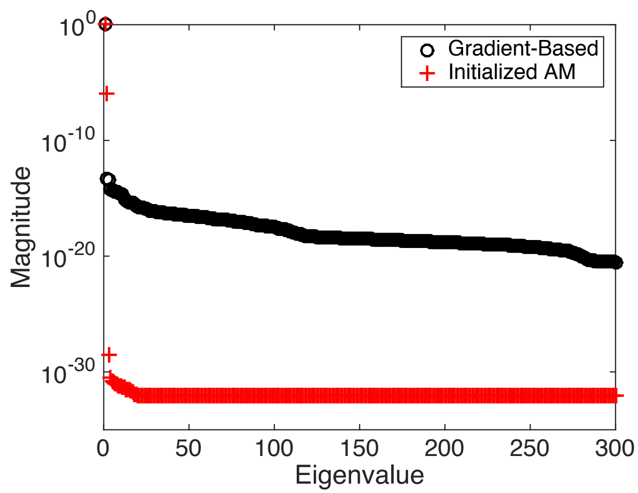
Sensitivity Analysis and Active Subspace Construction for Surrogate Models Employed for Bayesian Inference

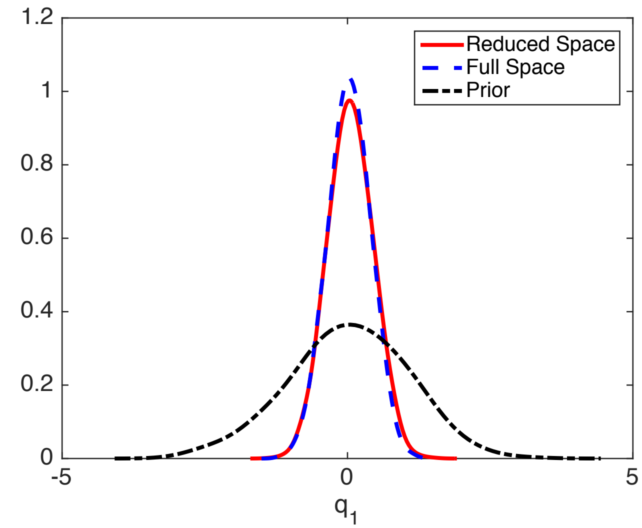
Ralph C. Smith

Department of Mathematics

North Carolina State University







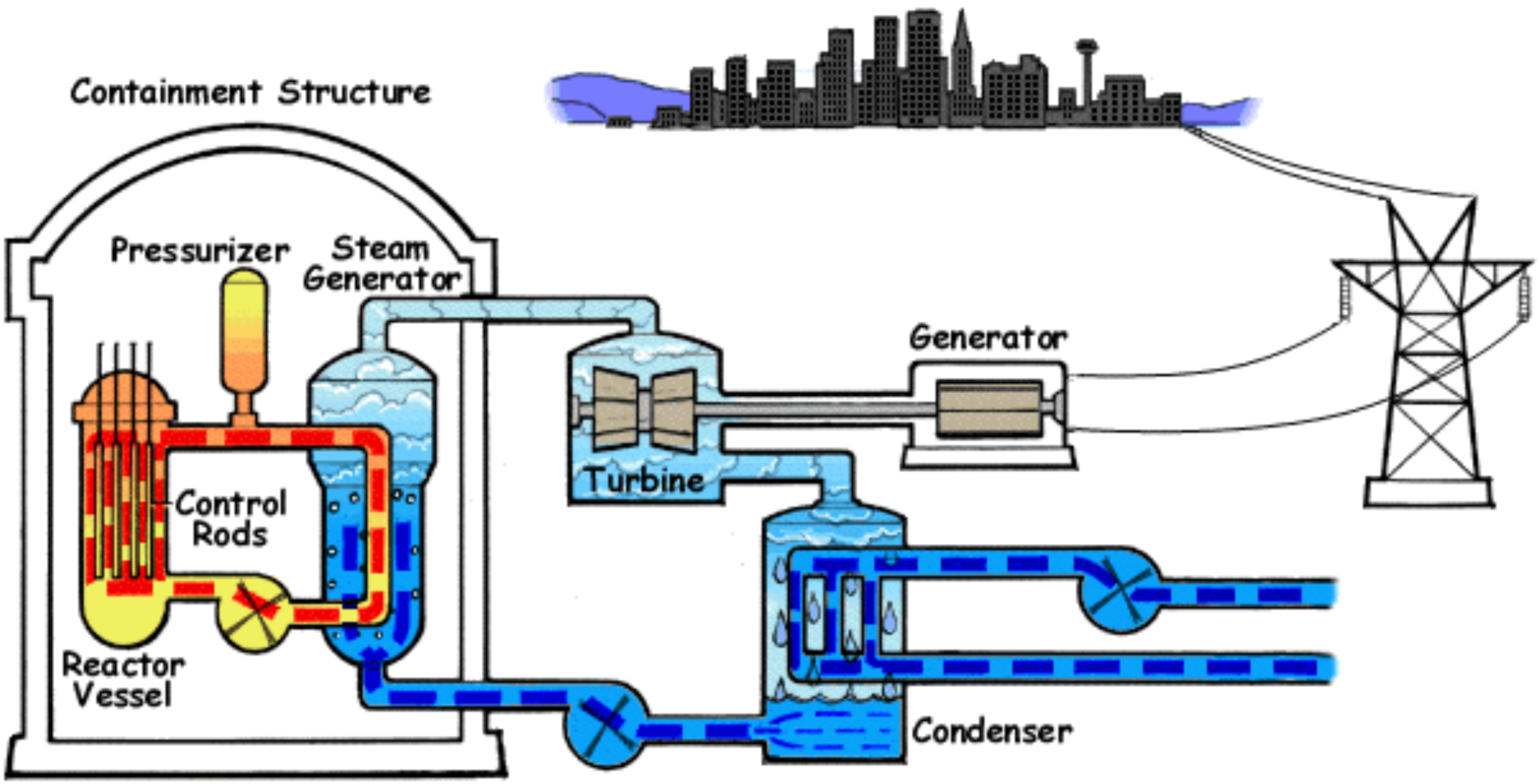
"We":

Kayla Coleman, Lider Leon, Allison Lewis, Mohammad Abdo (NCSU)

Brian Williams (LANL), Max Morris (Iowa State University)

Billy Oates, Paul Miles (Florida State University)

Example 1: Pressurized Water Reactors (PWR)



Models:

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

Objective: Develop Virtual Environment for Reactor Applications (VERA)

Motivation for Active Subspace Construction

3-D Neutron Transport Equations:

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

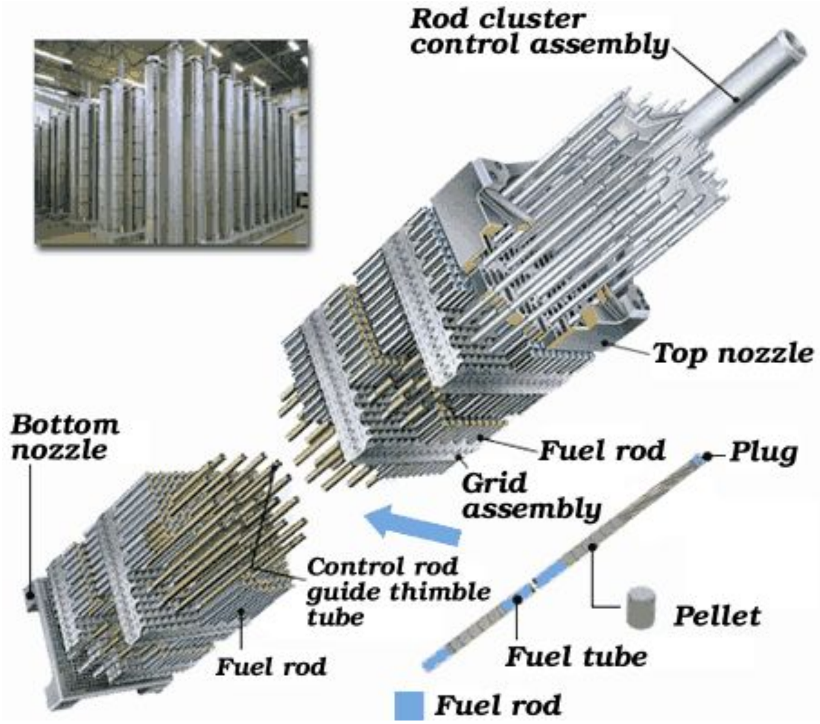
Challenges:

• Linear in the state but function of 7 independent variables:

$$r = x, y, z; E; \Omega = \theta, \phi; t$$

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

• ORNL Code SCALE: can take minutes to hours to run.



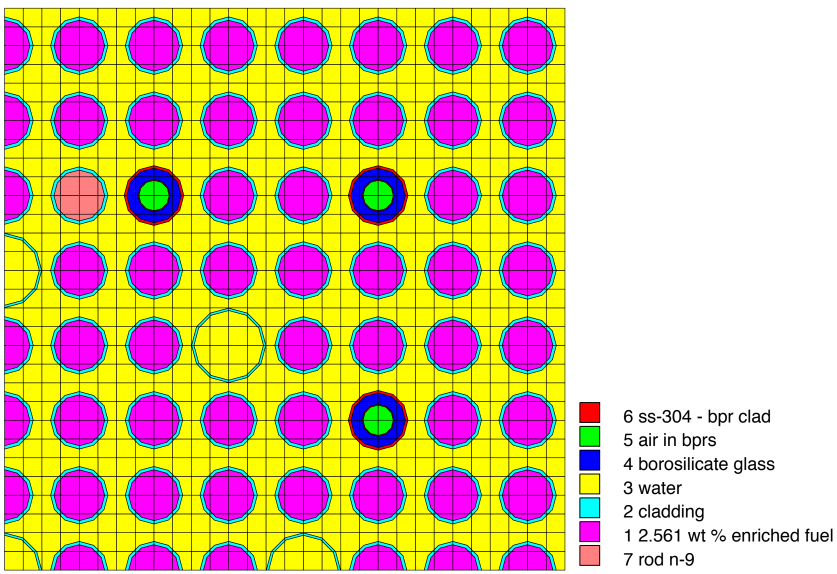
• SCALE TRITON has adjoint capabilities via TSUNAMI-2D and NEWT.

SCALE6.1: High-Dimensional Example

Setup: Cross-section computations SCALE6.1

- Input Dimension: 7700
- Output k_{eff} : Magnitude governs reactions

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



PWR Quarter Fuel Lattice

Note:

- Requires determination of active subspace to reduce input dimensions.
- Finite-difference approximations of gradient ineffective due to dimension

Motivation for Inference on Active Subspaces

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

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

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

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

Notes:

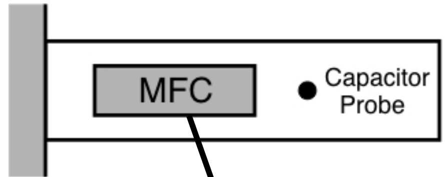
- Similar relations for gas and bubbly phases
- Reduced models must conserve mass, momentum and energy

Note:

- CFD and sub-channel 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 is necessary and closure relations can conflict.

Example 2. Multiscale Model Development

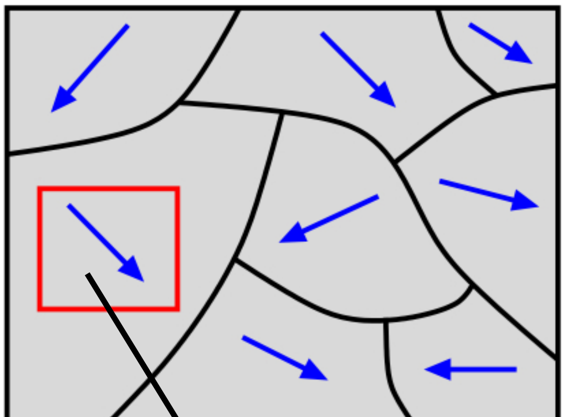
Example: PZT-Based Macro-Fiber Composites

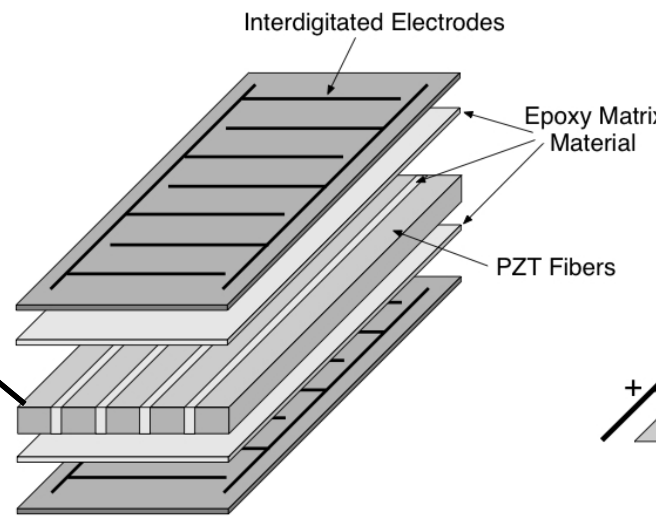


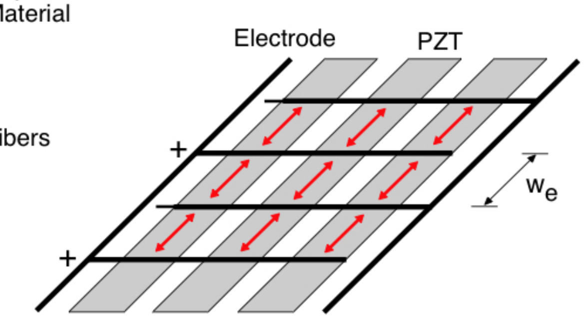
$$\rho \ddot{u} = \nabla \cdot \sigma + F$$

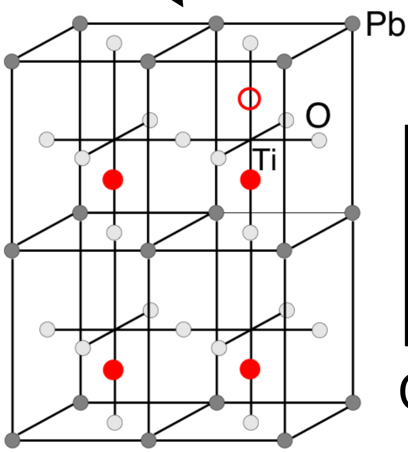
$$\nabla \cdot D = 0 \; , \; D = \varepsilon_0 E + P$$

$$\nabla \times E = 0 \; , \; E = -\nabla \varphi$$









$$P^\alpha = d_\alpha \sigma + \chi_\alpha^\sigma E + P_R^\alpha$$

$$\varepsilon^\alpha = s_\alpha^E \sigma + d_\alpha E + \varepsilon_R^\alpha$$

Continuum Energy Relations

$$P = d(E,\sigma)\sigma + \chi^\sigma E + P_{irr}(E,\sigma)$$

$$\varepsilon = s^E \sigma + d(E,\sigma) E + \varepsilon_{irr}(E,\sigma)$$

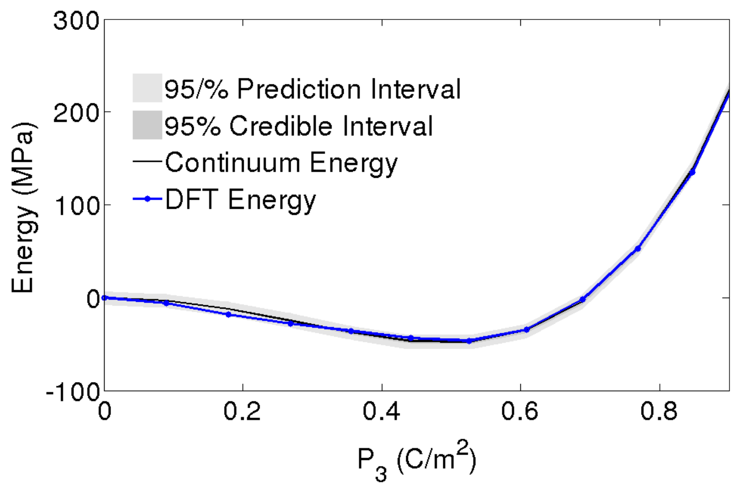
Homogenized Energy Model (HEM)

Quantum-Informed Continuum Models

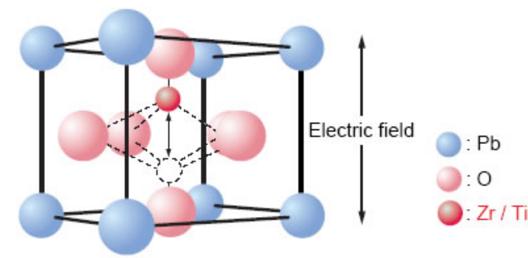
Objectives:

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

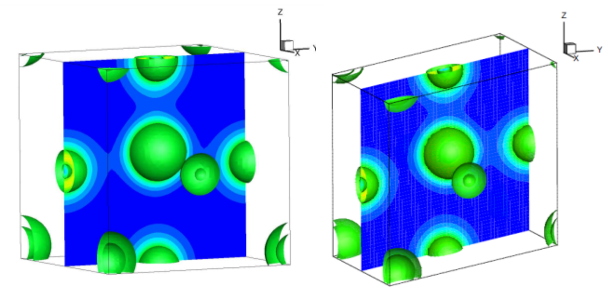
$$\psi(P) = \alpha_1 P^2 + \alpha_{11} P^4 + \alpha_{111} P^6$$



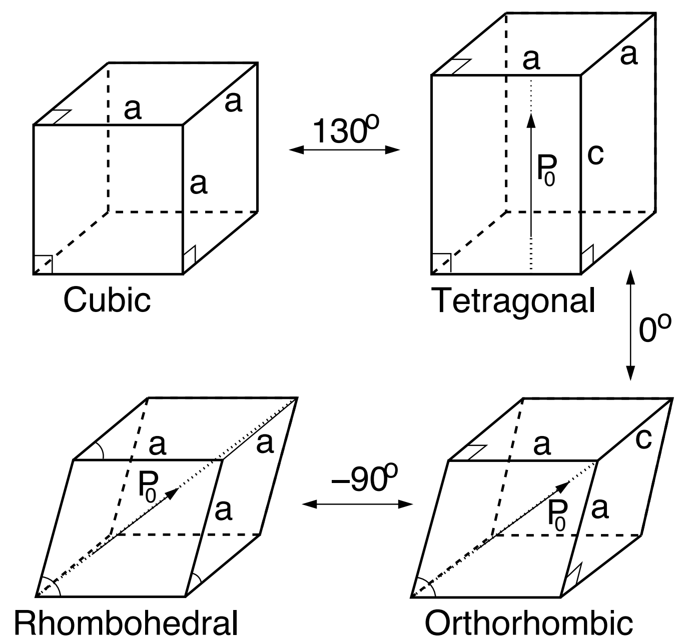
Landau energy



Lead Titanate Zirconate (PZT)



DFT Electronic Structure Simulation



UQ and SA Issues:

- Is 6^{th} order term required to accurately characterize material behavior?

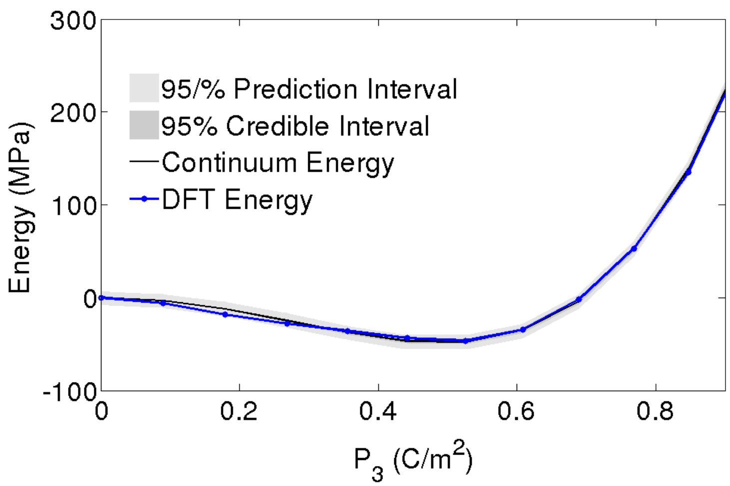
- **Note:** Determines molecular structure

Quantum-Informed Continuum Models

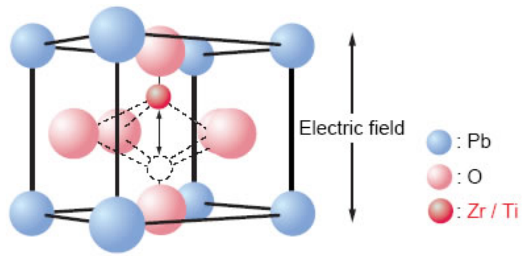
Objectives:

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

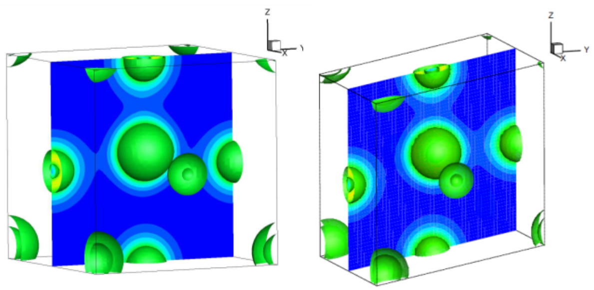
$$\psi(P) = \alpha_1 P^2 + \alpha_{11} P^4 + \alpha_{111} P^6$$



Landau energy



Lead Titanate Zirconate (PZT)



DFT Electronic Structure Simulation

UQ and SA Issues:

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

Broad Objective:

- Use UQ/SA to help bridge scales from quantum to system

Global Sensitivity Analysis: Analysis of Variance

Sobol' Representation: $Y = f(q)$

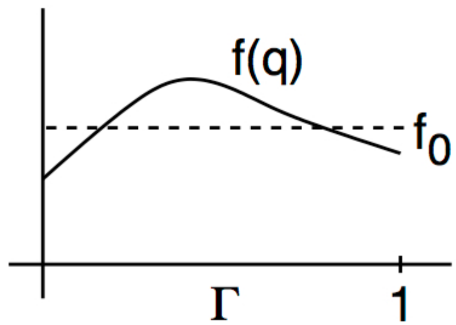
$$
\begin{aligned}
f(q) &= f_0 + \sum_{i=1}^{p} f_i(q_i) + \sum_{i \leqslant i < j \leqslant p} f_{ij}(q_i, q_j) + \cdots + f_{12\cdots p}(q_1, \ldots, q_p) \\
&= f_0 + \sum_{i=1}^{p} \sum_{|u|=i} f_u(q_u)
\end{aligned}
$$

where

$$f_0 = \int_\Gamma f(q)\rho(a)dq = \mathbb{E}[f(q)]$$

$$f_i(q_i) = \mathbb{E}[f(q)|q_i] - f_0$$

$$f_{ij}(q_i, q_j) = \mathbb{E}[f(q)|q_i, q_j] - f_i(q_i) - f_j(q_j) - f_0$$



Typical Assumption: $q_1, q_2, \ldots, q_p$ independent. Then

$$\int_\Gamma f_u(q_u) f_v(q_v) \rho(q) dq = 0 \quad \text{for } u \neq v$$

$$\Rightarrow \mathrm{var}[f(q)] = \sum_{i=1}^{p} \sum_{|u|=i} \mathrm{var}[f_u(q_u)]$$

Sobol' Indices:

$$S_u = \frac{\mathrm{var}[f_u(q_u)]}{\mathrm{var}[f(q)]} \quad , \quad T_u = \sum_{v \subseteq u} S_v$$

Note: Magnitude of S_i, T_i quantify contributions of q_i to $\mathrm{var}[f(q)]$

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

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?

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

Parameters:

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

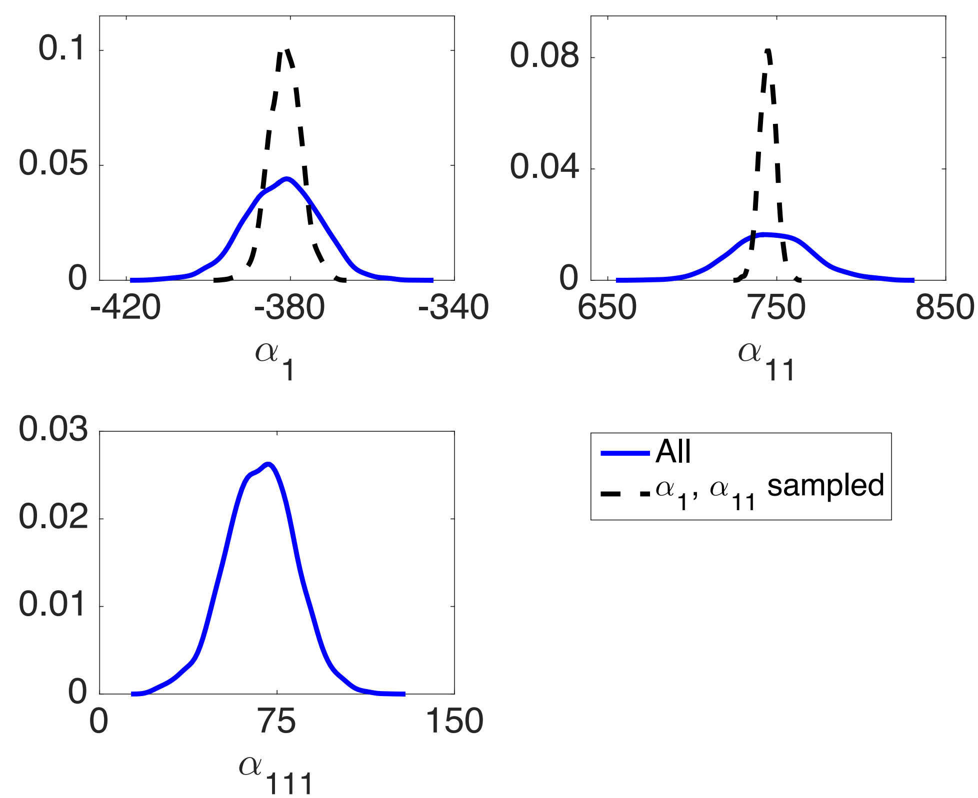
Global Sensitivity Analysis:

	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

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:

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

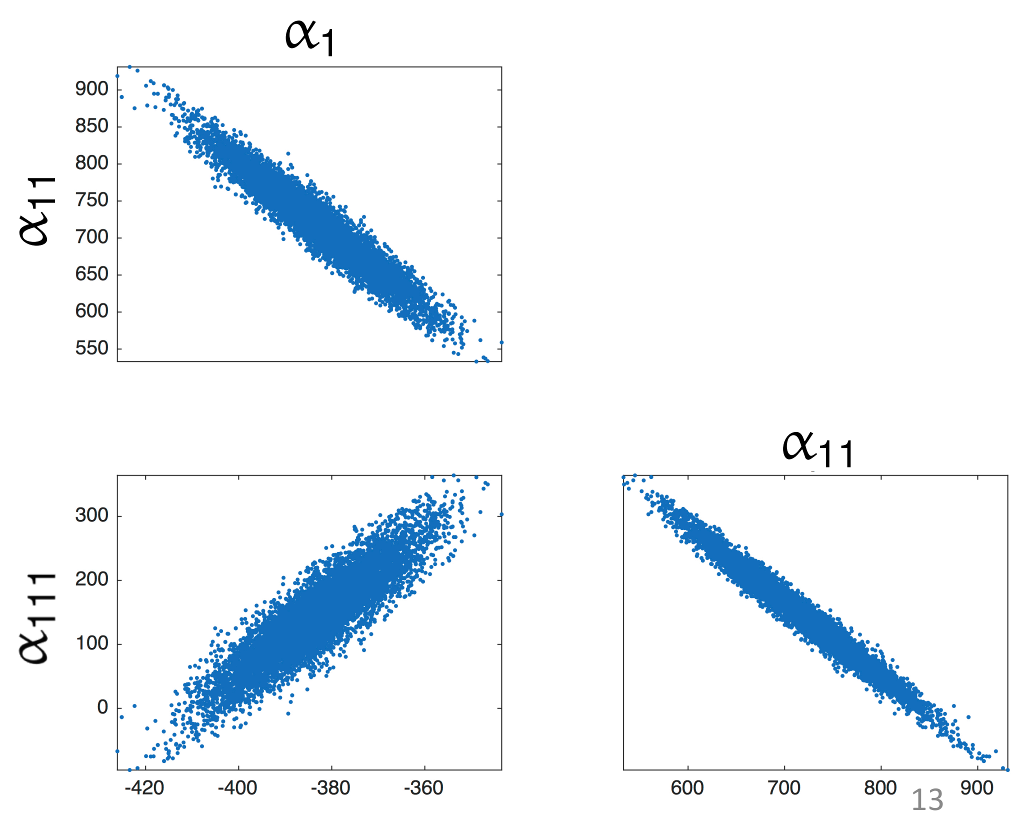
Global Sensitivity Analysis:

	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}



Global Sensitivity Analysis: Analysis of Variance

Sobol' Representation:

$$f(q) = f_0 + \sum_{i=1}^{p} \sum_{|u|=i} f_u(q_u)$$

One Solution: Take variance to obtain

$$\text{var}[f(q)] = \sum_{i=1}^{p} \sum_{|u|=i} \text{cov}[f_u(q_u), f(q)]$$

Sobol' Indices:

$$S_u = \frac{\text{cov}[f_u(q_u), f(q)]}{\text{var}[f(q)]}$$

Pros:

- Provides variance decomposition that is analogous to independent case

Cons:

- Indices can be negative and difficult to interpret
- Often difficult to determine underlying distribution
- Monte Carlo approximation often prohibitively expensive.

Global Sensitivity Analysis: Analysis of Variance

Sobol' Representation:

$$f(q) = f_0 + \sum_{i=1}^{p} \sum_{|u|=i} f_u(q_u)$$

One Solution: Take variance to obtain

$$\mathrm{var}[f(q)] = \sum_{i=1}^{p} \sum_{|u|=i} \mathrm{cov}[f_u(q_u), f(q)]$$

Sobol' Indices:

$$S_u = \frac{\mathrm{cov}[f_u(q_u), f(q)]}{\mathrm{var}[f(q)]}$$

Pros:

- Provides variance decomposition that is analogous to independent case

Cons:

- Indices can be negative and difficult to interpret
- Often difficult to determine underlying distribution
- Monte Carlo approximation often prohibitively expensive.

Alternative: Construct active subspaces

- Can accommodate parameter correlation
- Often effective in high-dimensional space; e.g., p = 7700 for neutronics example

Additional Goal: Use Bayesian analysis on active subspace to construct posterior densities for physical parameters.

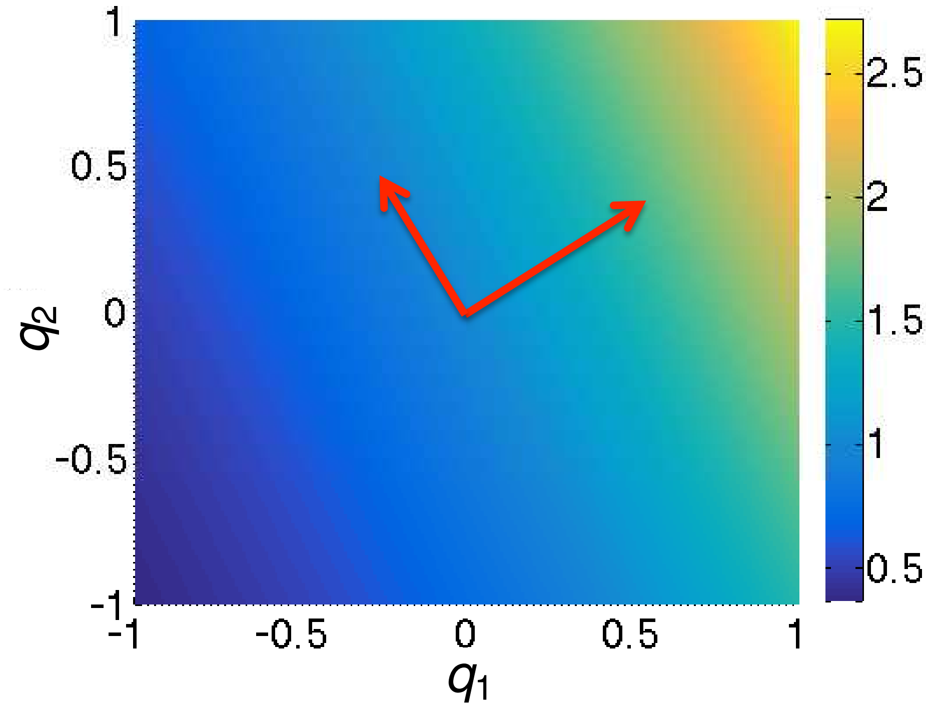
Active Subspaces

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



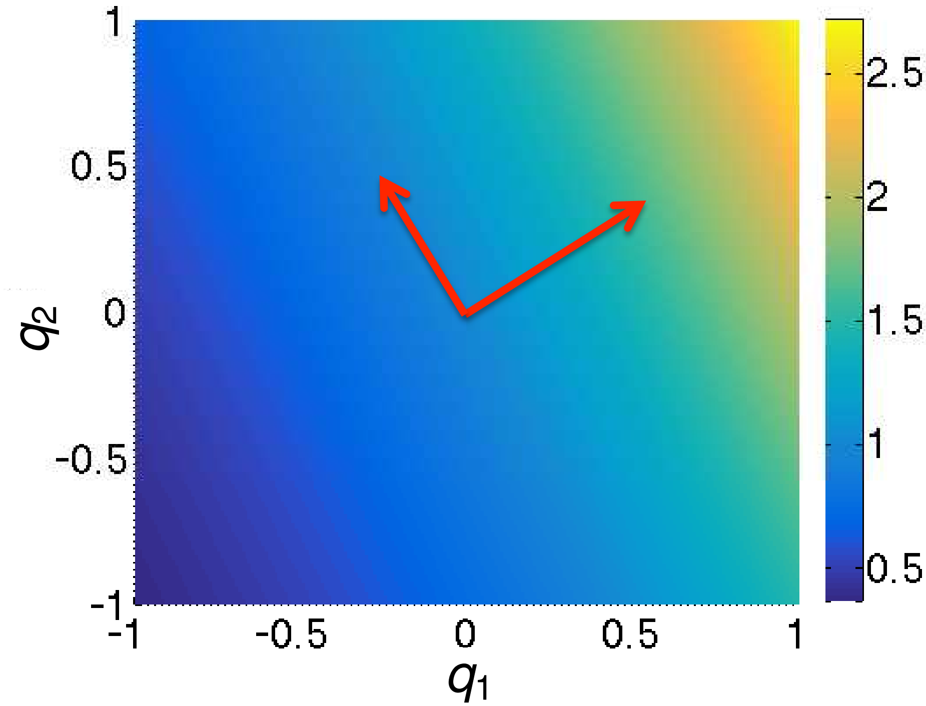
Active Subspaces

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).

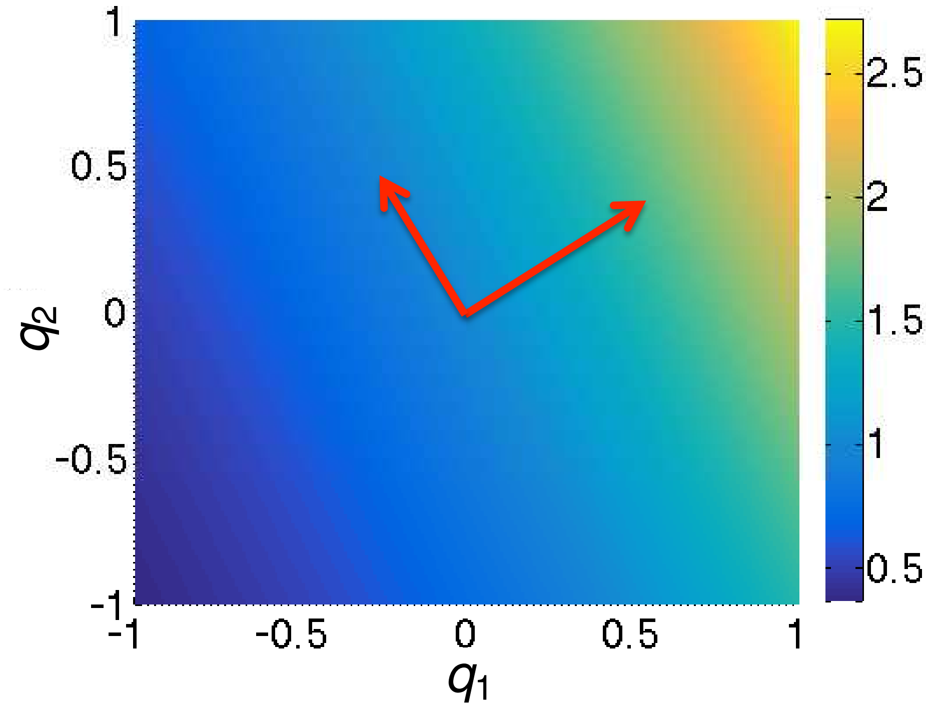
Active Subspaces

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).

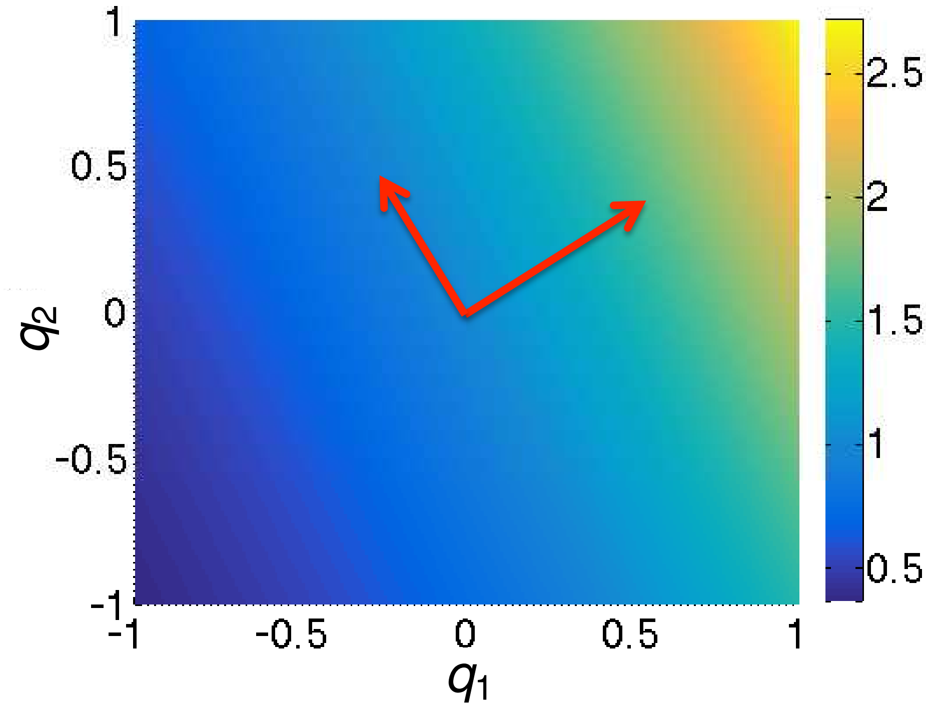
Active Subspaces

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.

Gradient-Based Active Subspace Construction

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

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

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

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

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 Construction

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

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

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

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

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

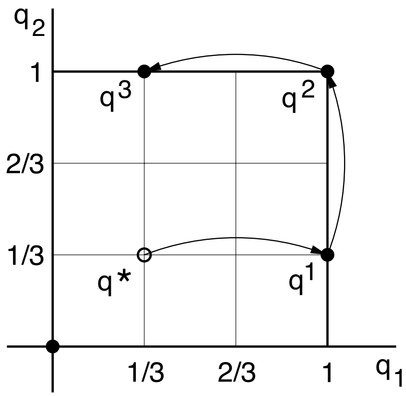
One Goal: Develop efficient algorithm for codes that do not have adjoint capabilities

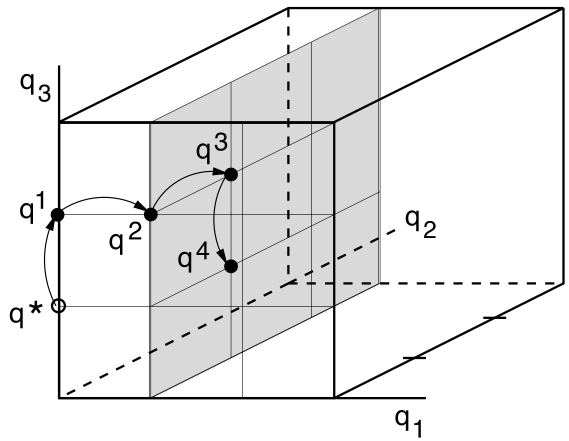
Note: Finite-difference approximations tempting but not effective for high-D

Strategy: Algorithm based on initialized adaptive Morris indices

Morris Screening: Random Sampling of Approximated Derivatives

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





Adaptive Algorithm:

- Use SVD to adapt stepsizes and directions to reflect active subspace.
- Reduce dimension of differencing as active subspace is discovered.

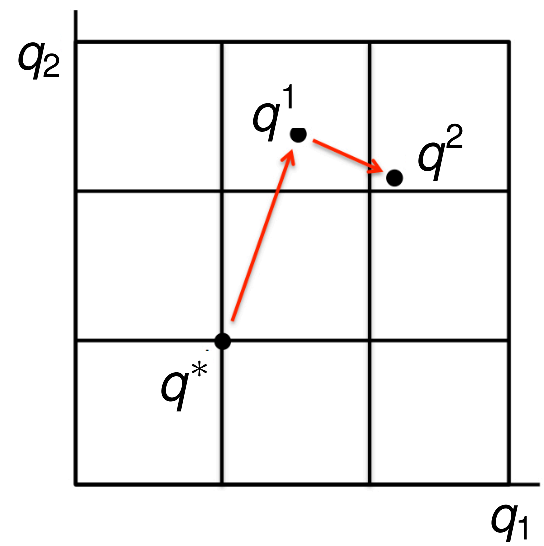
Elementary Effect:

$$d_i = \frac{f(q^j + \Delta e_i) - f(q^j)}{\Delta}$$

Global Sensitivity Measures: r samples

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

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



Note: Gets us to moderate-D but initialization required for high-D

Initialization Algorithm

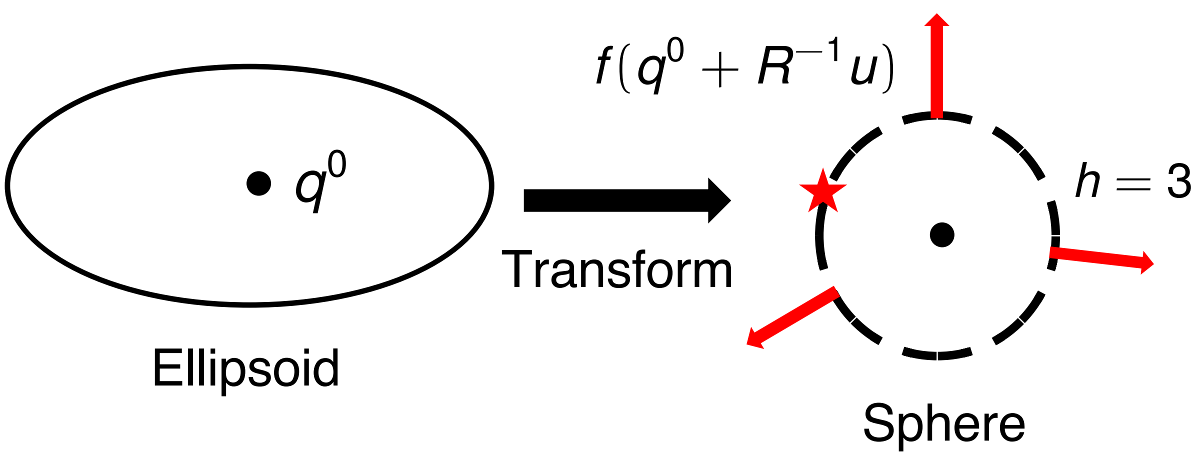
1. Inputs: ℓ iterations, h function evaluations per iteration
2. Sample w^1 from surface of unit sphere where approximately linear

For $j = 1, \ldots, \ell$

3. Sample $\{\tilde{v}_1^j, \ldots, \tilde{v}_h^j\}$ from surface of sphere
4. Use Lagrange multiplier to determine

$$u_{\max}^j = a_0^+ w^j + \sum_{i=1}^{h} a_i^+ v_i^j \;,\; v_i^1 = \tilde{v}_i^1$$

that maximizes $g(u) = f(q^0 + R^{-1} u)$.



Initialization Algorithm

1. Inputs: ℓ iterations, h function evaluations per iteration
2. Sample w^1 from surface of unit sphere where approximately linear

For $j = 1, \ldots, \ell$

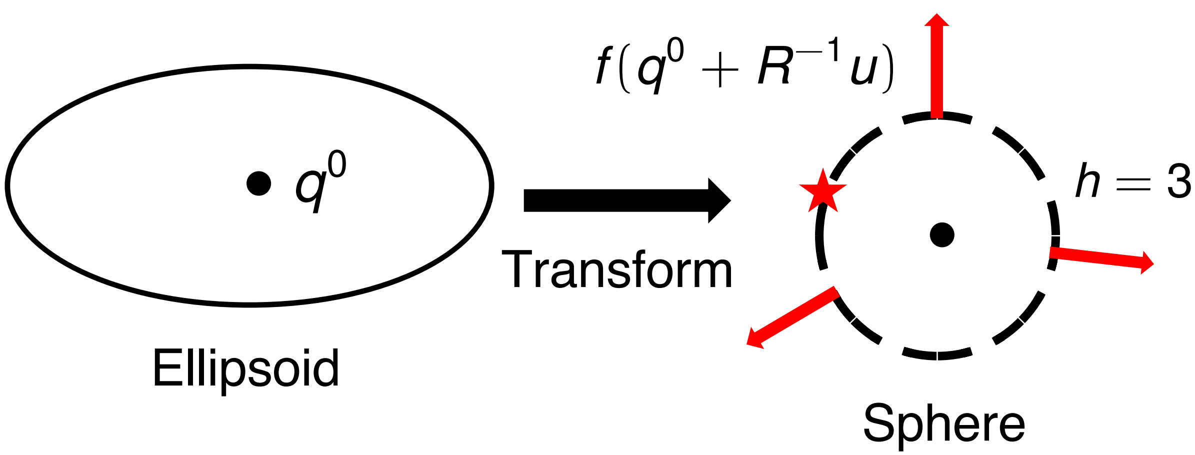
3. Sample $\{\tilde{v}_1^j, \ldots, \tilde{v}_h^j\}$ from surface of sphere
4. Use Lagrange multiplier to determine

$$u^j_{\max} = a_0^+ w^j + \sum_{i=1}^{h} a_i^+ v_i^j \ , \ v_i^1 = \tilde{v}_i^1$$

that maximizes $g(u) = f(q^0 + R^{-1} u)$.

Note: For h=1, maximizing great circle through w^1, v^1

Example: Let w^1 = Atlanta, v^1 = Venice, and g(u) = 'QUIETness' of seatmate on flight



Initialization Algorithm

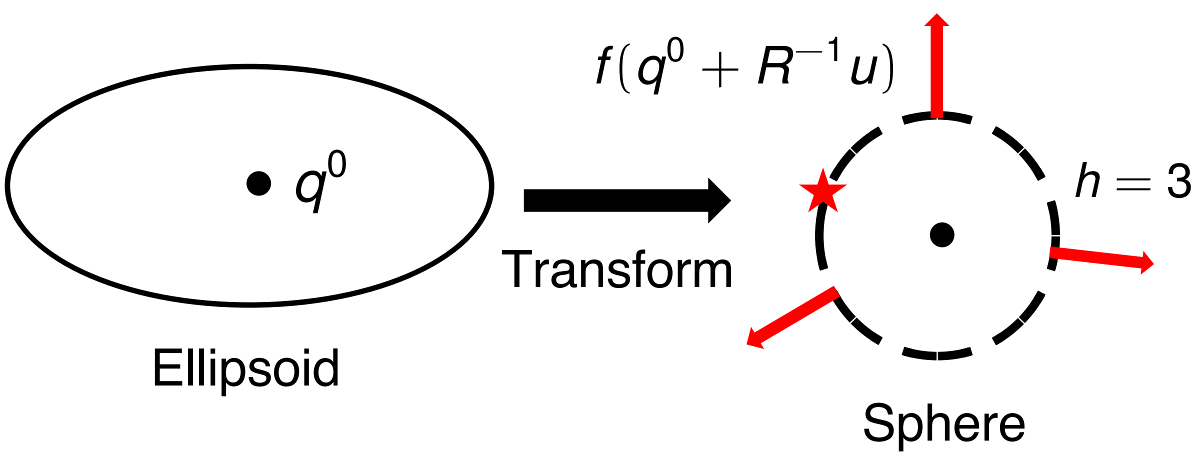
1. Inputs: ℓ iterations, h function evaluations per iteration
2. Sample w^1 from surface of unit sphere where approximately linear

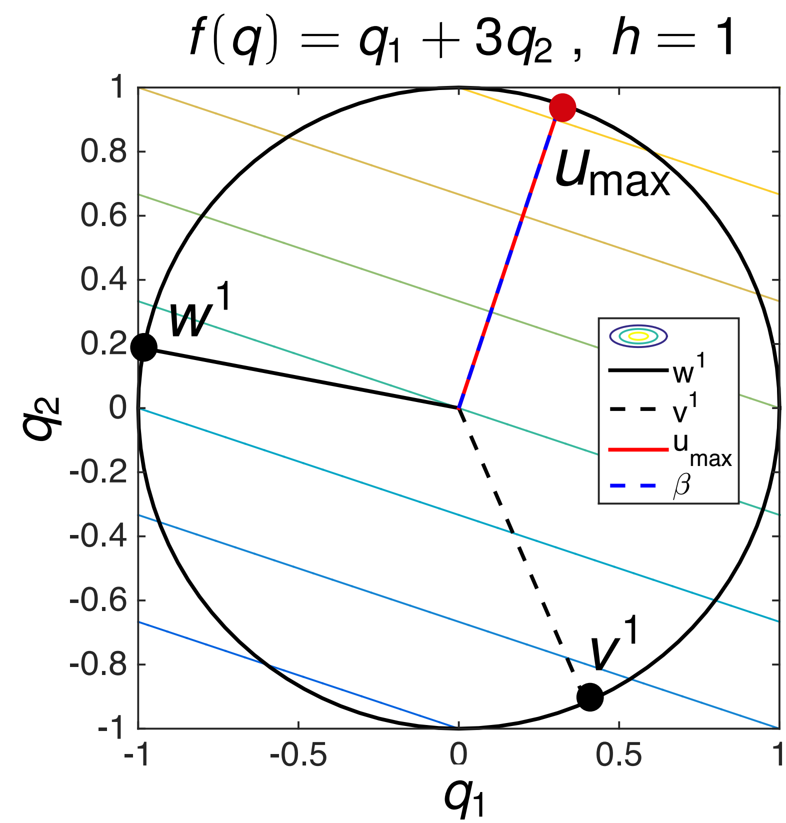
For $j = 1, \ldots, \ell$

3. Sample $\{\tilde{v}_1^j, \ldots, \tilde{v}_h^j\}$ from surface of sphere
4. Use Lagrange multiplier to determine

$$u_{\max}^j = a_0^+ w^j + \sum_{i=1}^{h} a_i^+ v_i^j \ , \ v_i^1 = \tilde{v}_i^1$$

that maximizes $g(u) = f(q^0 + R^{-1}u)$.





Initialization Algorithm

1. Inputs: ℓ iterations, h function evaluations per iteration
2. Sample w^1 from surface of unit sphere where approximately linear

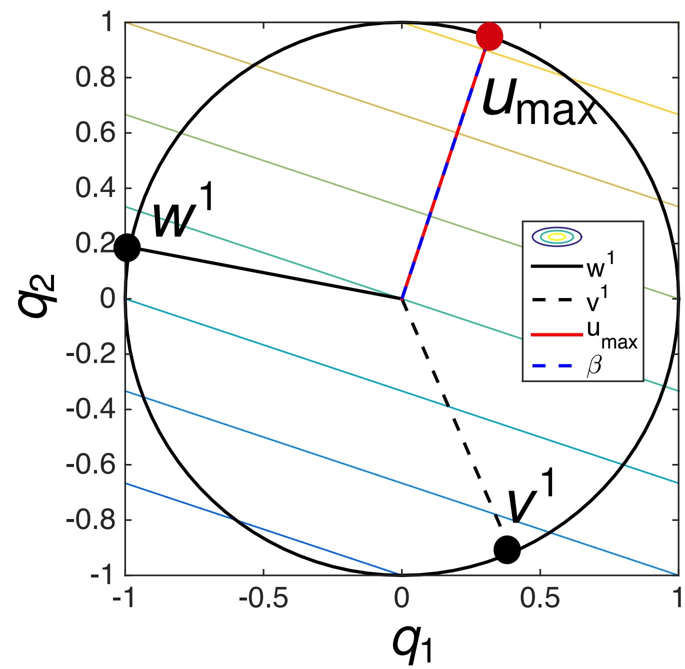
For $j = 1, \dots, \ell$

3. Sample $\{\tilde{v}_1^j, \dots, \tilde{v}_h^j\}$ from surface of sphere
4. Use Lagrange multiplier to determine

$$u_{\max}^j = a_0^+ w^j + \sum_{i=1}^{h} a_i^+ v_i^j \ , \ v_i^1 = \tilde{v}_i^1$$

that maximizes $g(u) = f(q^0 + R^{-1}u)$.

Set $w^{j+1} = u_{\max}^j$.



5. Take $C = [w^j, v_1^j, \dots, v_h^j]$ and set $P_{u_{\max}^j} = u_{\max}^j (u_{\max}^j)^T$

6. Let $C_{j\perp} = \left[\text{span}\left(C_{(j-1)\perp}, (I_m - P_{u_{\max}^j} C\right)\right]$ and set $P_{C_{j\perp}} = C_{j\perp}(C_{j\perp}^T C_{j\perp})^{-1} C_{j\perp}^T$

7. Take $v_i^j = \dfrac{(I_m - P_{C_{j\perp}})\tilde{v}_i^j}{\|(I_m - P_{C_{j\perp}})\tilde{v}_i^j\|}$, $i = 1, \dots, h$ and repeat

Example: Initialization Algorithm to Approximate Gradient

Example: Family of elliptic PDE's

$$-\nabla_s \cdot (a(s,\ell)\nabla_s u(s,a(s,\ell)) = 1 \quad , \; s \in [0,1]^2 \quad , \ell = 1, \ldots, n$$

with the random field representations

$$\log(a(s,\ell)) = \sum_{i=1}^{p} q_i^{\ell} \gamma_i \phi_i(s)$$

Quantity of interest: e.g., strain along edge at n levels

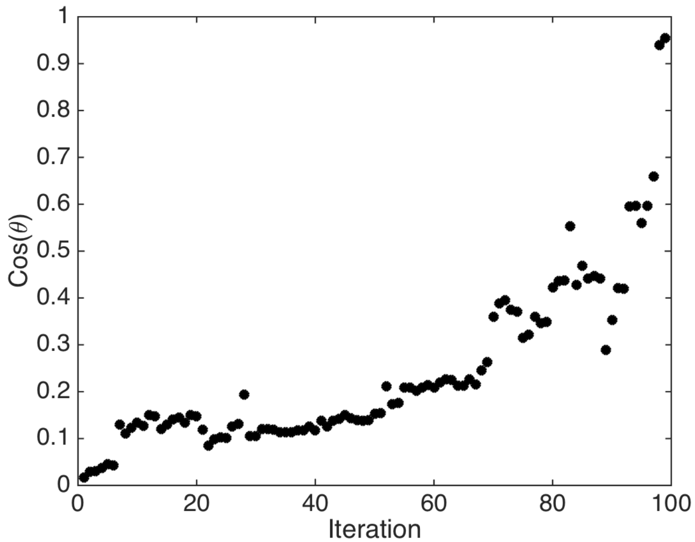
$$f(\mathbf{q}^1, \ldots, \mathbf{q}^n) \approx \sum_{\ell=1}^{n} \frac{1}{|\Gamma_2|} \int_{\Gamma_2} u(q(s,\ell)ds$$

Problem Dimensions:

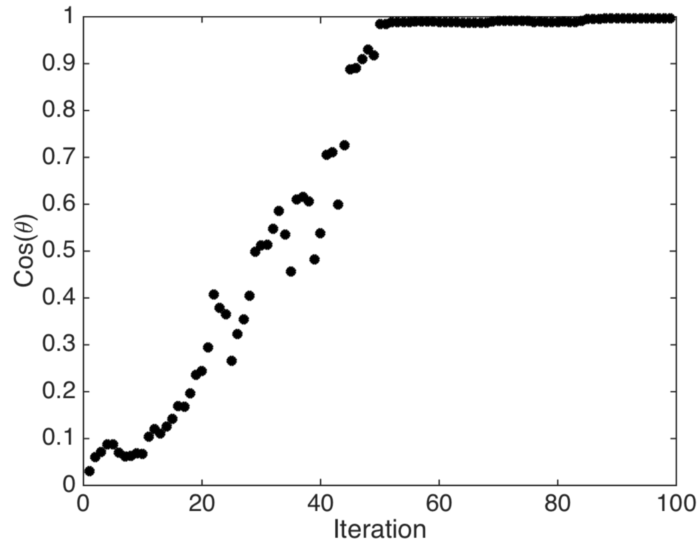
- Parameter dimension: p = 100
- Active subspace dimension: n = 1
- Finite element approximation

Example: Initialization Algorithm to Approximate Gradient

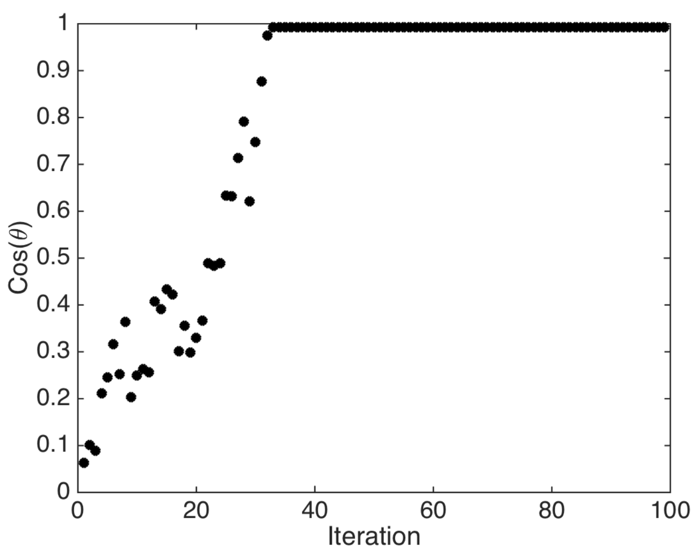
Results: Cosine of angle between ’analytic’ and computed gradient



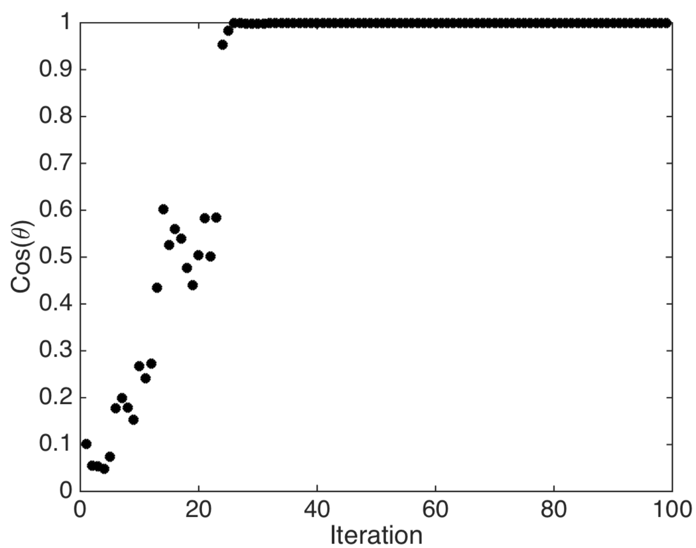
$h = 1$



$h = 2$



$h = 3$



$h = 4$

Recall: p=100

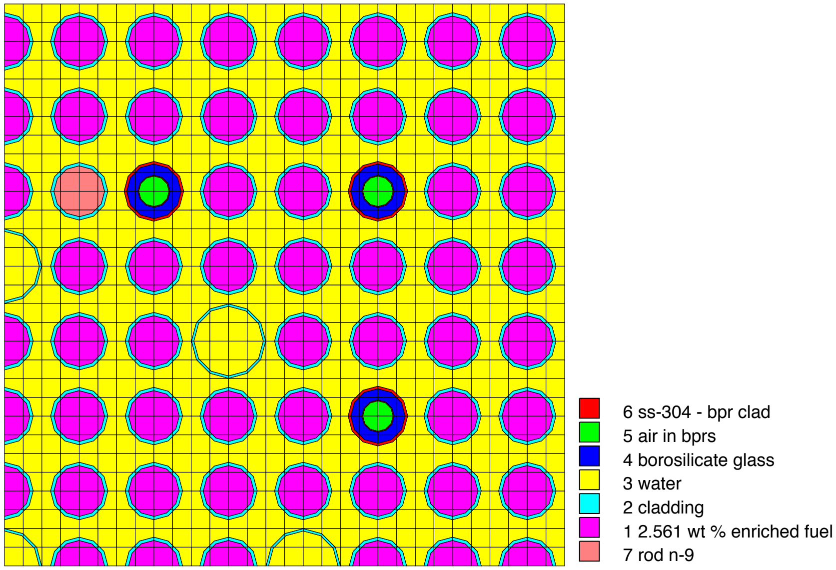
Note: Convergence within $h \cdot \ell$ iterations

SCALE6.1: High-Dimensional Example

Setup: Cross-section computations SCALE6.1

- Input Dimension: 7700
- Output k_{eff} : Governs reactions

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

Really Annoying Reality for Allie and Kayla: Cross-section libraries are binary and require conversion to floating point for perturbations.

SCALE6.1: High-Dimensional Example

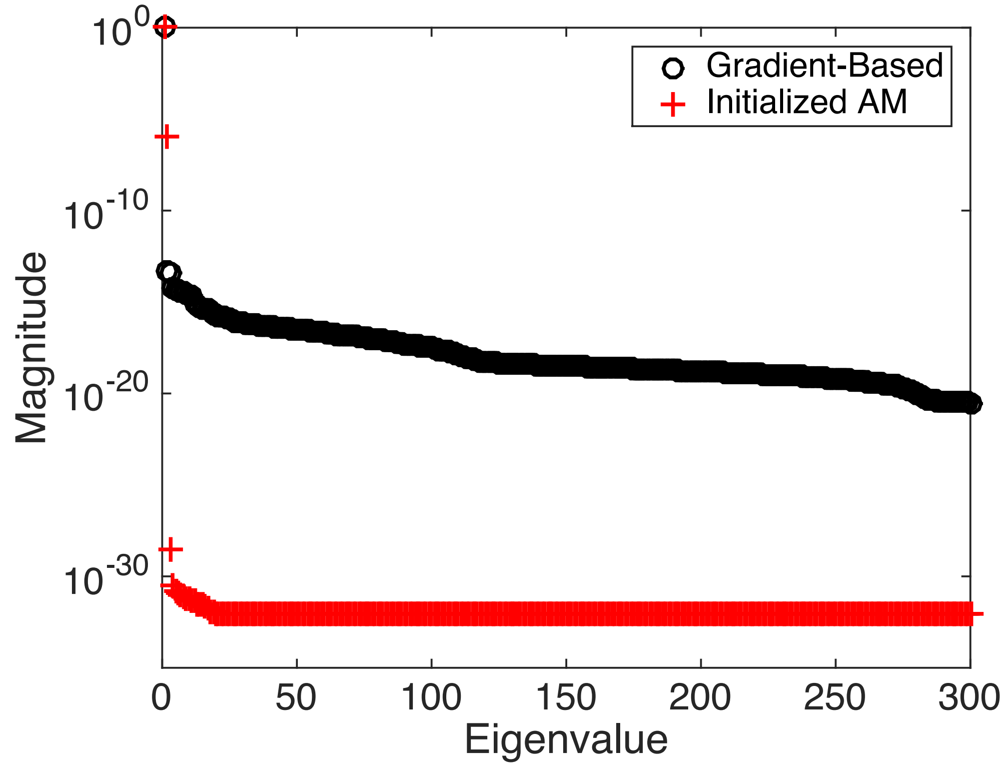
Setup:

- Input Dimension: 7700

SCALE Evaluations:

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

Note: Analytic eigenvalues: 0, 1



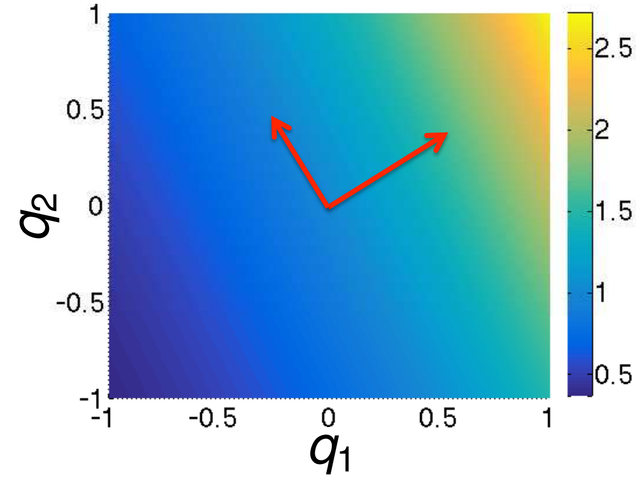
Active Subspace Dimensions:

	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

Note: Computing *converged* adjoint solution is expensive and often not achieved

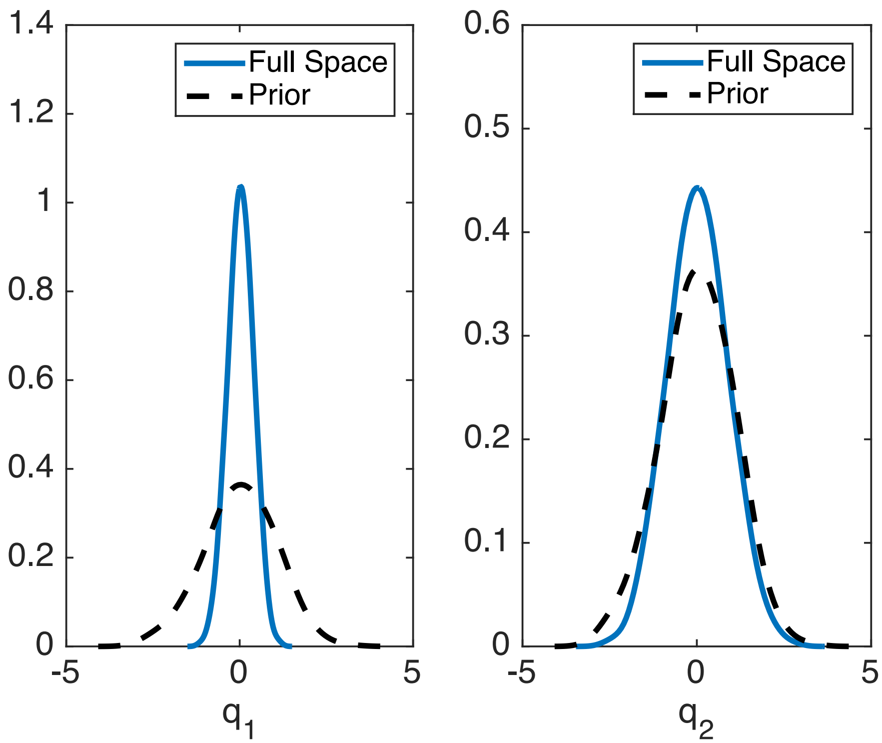
Bayesian Inference on Active Subspaces

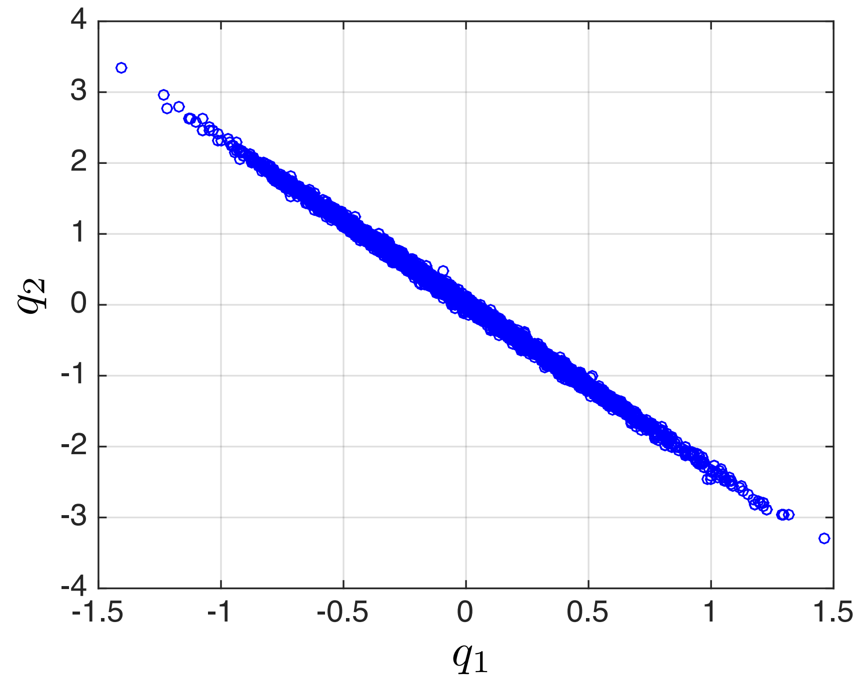
Example: $y = \exp(0.7q_1 + 0.3q_2)$



Full Space Inference:

- Parameters not jointly identifiable
- Result: Prior for 2nd parameter is minimally informed.
- **Goal:** Use active subspace to quantify parameter sensitivity and guide inference.





Bayesian Inference on Active Subspaces

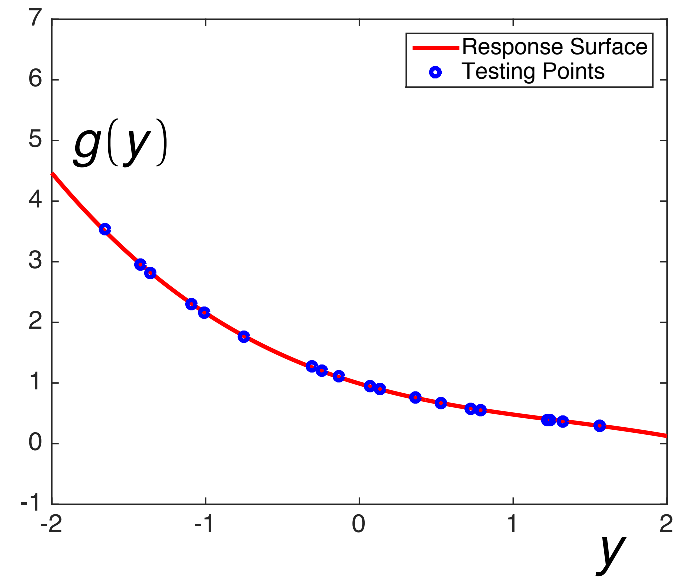
Example: $y = \exp(0.7q_1 + 0.3q_2)$

Active Subspace: For gradient matrix G, form SVD

$$G = U \Lambda V^T$$

Eigenvalue spectrum indicates 1-D active subspace with basis

$$U(:,1) = [0.91 \;,\; 0.39]$$

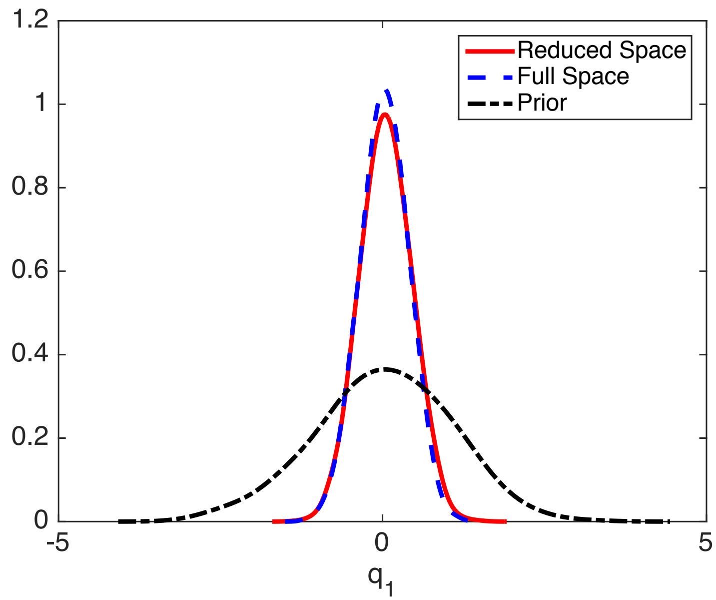


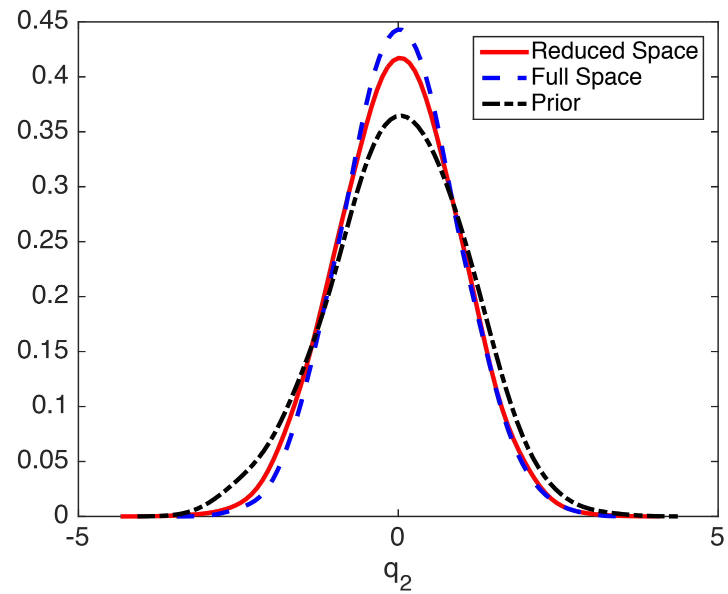
Strategy: Inference based on active subspace

- For values $\{q^j\}_{j=1}^M$, compute $y^j = U(:,1)^T q^j$ and fit response surface $g(y)$
- Perform Bayesian inference for y
- Because model is "invariant" to $z = U(:,2)^T q$, draw $\{z^j\} \sim N(0,1)$
- Transform to $q^j = U(:,1)y^j + U(:,2)z^j$ to obtain posterior densities for physical parameters

Bayesian Inference on Active Subspaces

Results: Inference based on active subspace





Global Sensitivity: For active subspace of dimension n, consider vector of activity scores

$$\alpha_i(n) = \sum_{j=1}^{n} \lambda_j w_{i,j}^2 \ , \ \ i = 1, \ldots , p$$

Present Example: Here n = 1 and $w_1^2 = U(:,1) .* U(:,1) = [0.91^2 \ , \ 0.39^2]$

Conclusion: First parameter is more influential and better informed during Bayesian inference.

Bayesian Inference on Active Subspaces

Example: Family of elliptic PDE's – Same as initialization example

$$-\nabla_s \cdot (a(s,\ell)\nabla_s u(s,a(s,\ell)) = 1 \quad , \; s \in [0,1]^2 \quad , \ell = 1, \ldots, n$$

with the random field representations

$$\log(a(s,\ell)) = \sum_{i=1}^{p} q_i^\ell \gamma_i \phi_i(s)$$

Quantity of interest: e.g., strain along edge at n levels

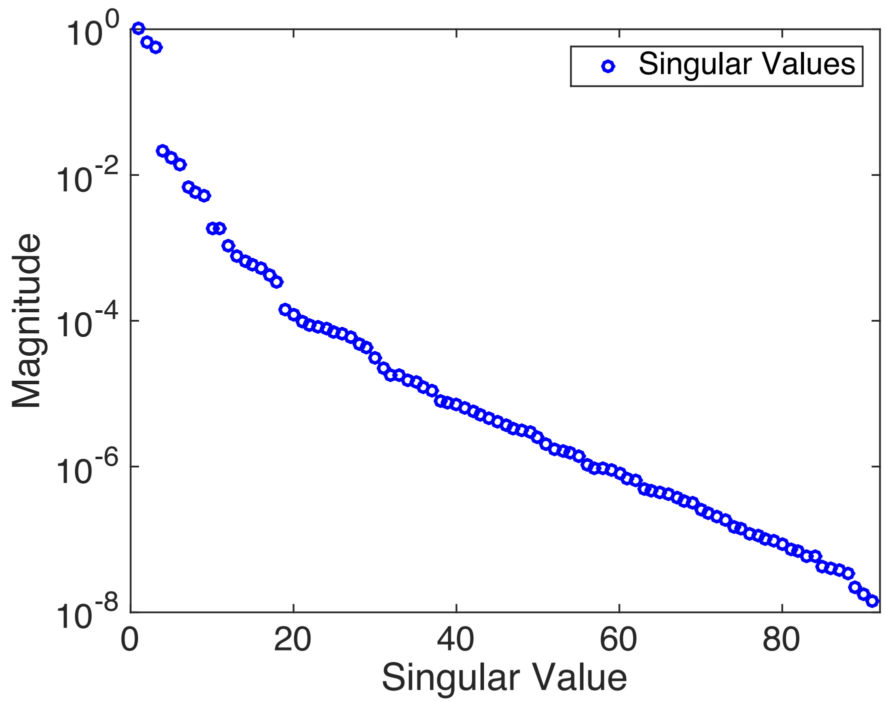
$$f(\mathbf{q}^1, \ldots, \mathbf{q}^n) \approx \sum_{\ell=1}^{n} \frac{1}{|\Gamma_2|}\int_{\Gamma_2} u(q(s,\ell)ds$$

Problem Dimensions:

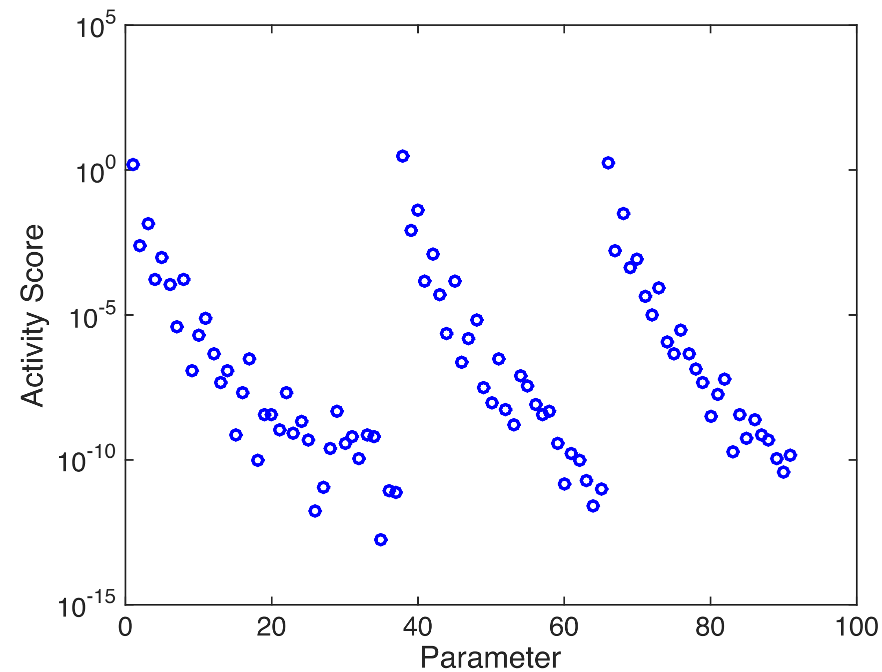
- Parameter dimension: p = 91
- Active subspace dimension: n = 3
- Finite element approximation

Bayesian Inference on Active Subspaces

Singular Values: Recall n = 3



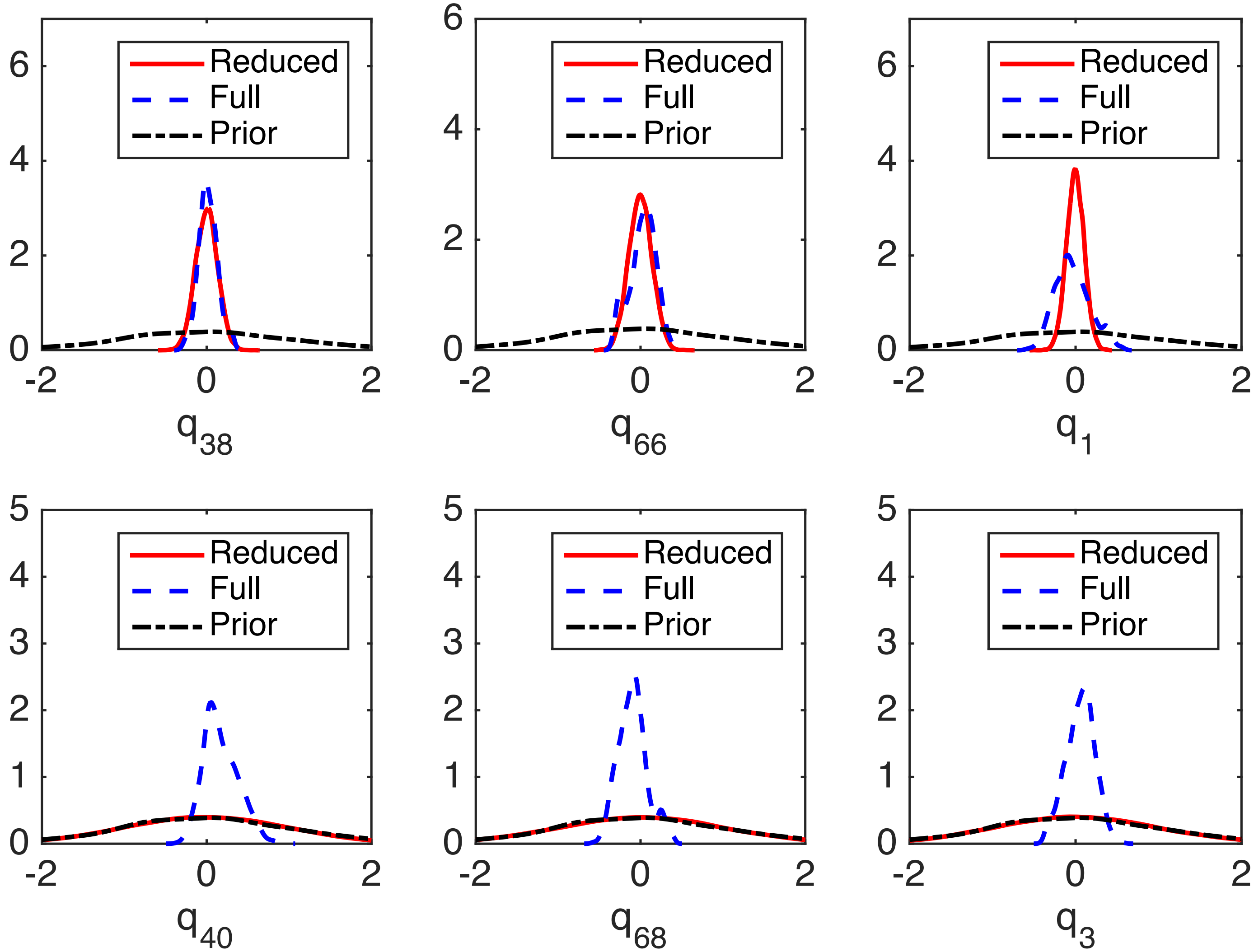
Activity Scores: Quantify global sensitivity



Conclusion: Parameters 1, 38, 66 are most influential and will be primarily informed during Bayesian inference

Bayesian Inference on Active Subspaces

Recall: Parameters 1, 38, 66 are most influential and will be primarily informed during Bayesian inference



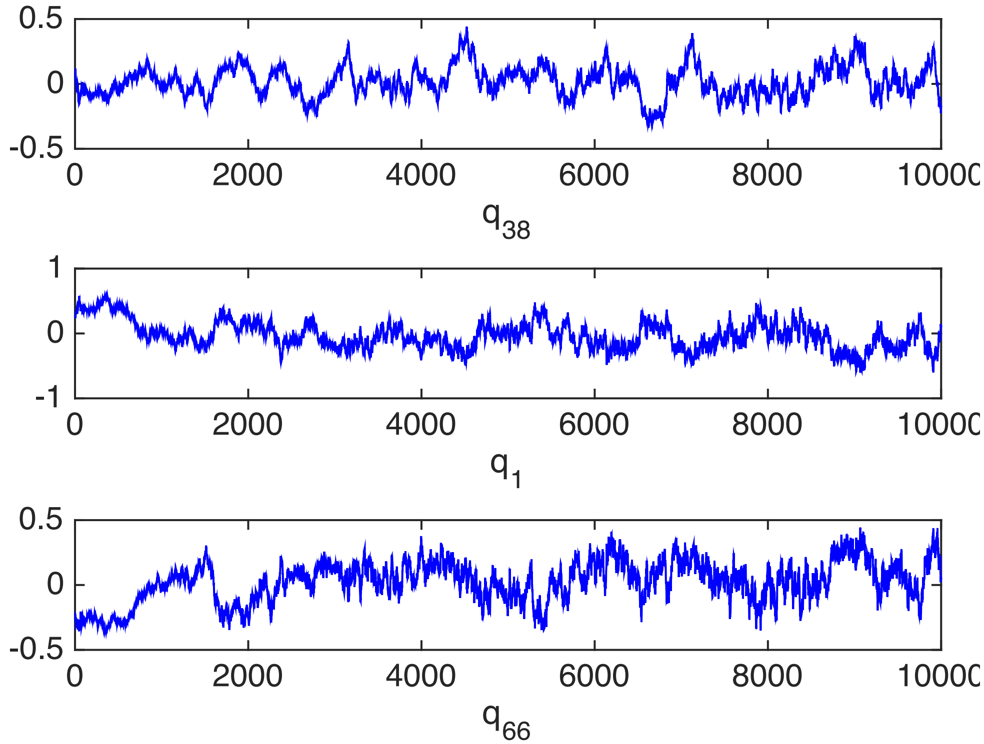
Note:

- Full space: 18 hours
- Reduced: 20 seconds

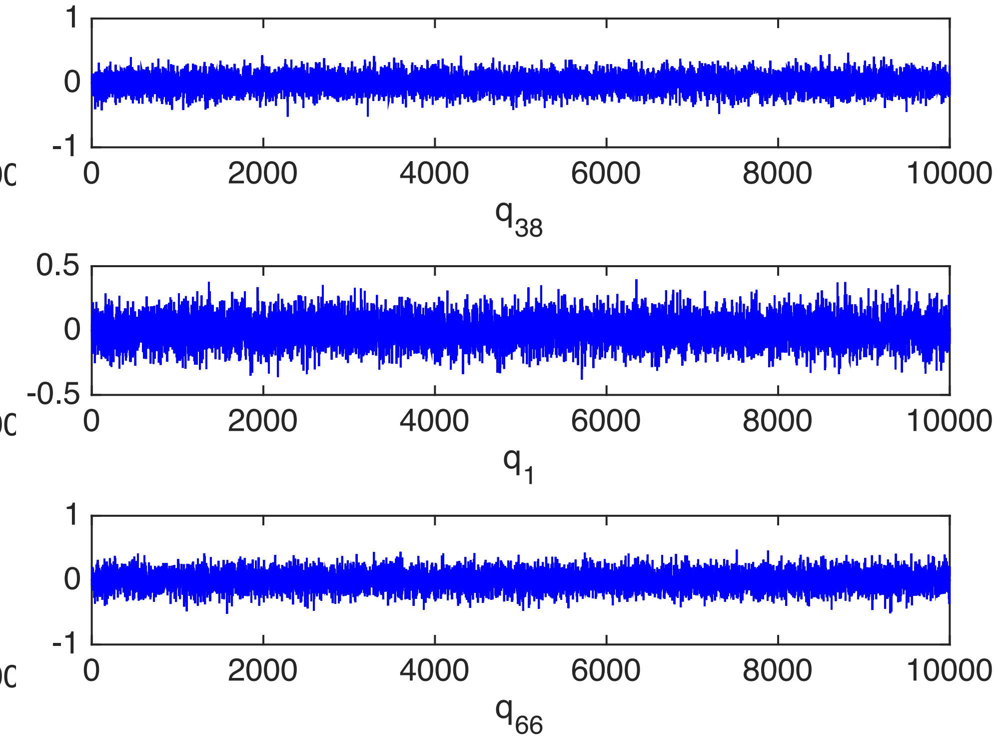
Bayesian Inference on Active Subspaces

Note:

- Chains for full space not converging well due to parameter nonidentifiability
- Hence full space inference is less reliable



Full Space



Active Subspace

Concluding Remarks

Notes:

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

• Active subspace construction necessary for models with high-dimensional parameter spaces; e.g., 7700.

• Due to complexity of models, surrogate or low-fidelity models typically required. Algorithms utilizing mutual information can maximize information gain when calibrating.

• Present and future work:

- Relax Gaussian constraints on priors when performing inference on active subspaces.
- Further analysis of activity scores.
- Construction of reduced-order models that conserve mass, momentum and energy.
- Reduced-order models for multi-physics problems.

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