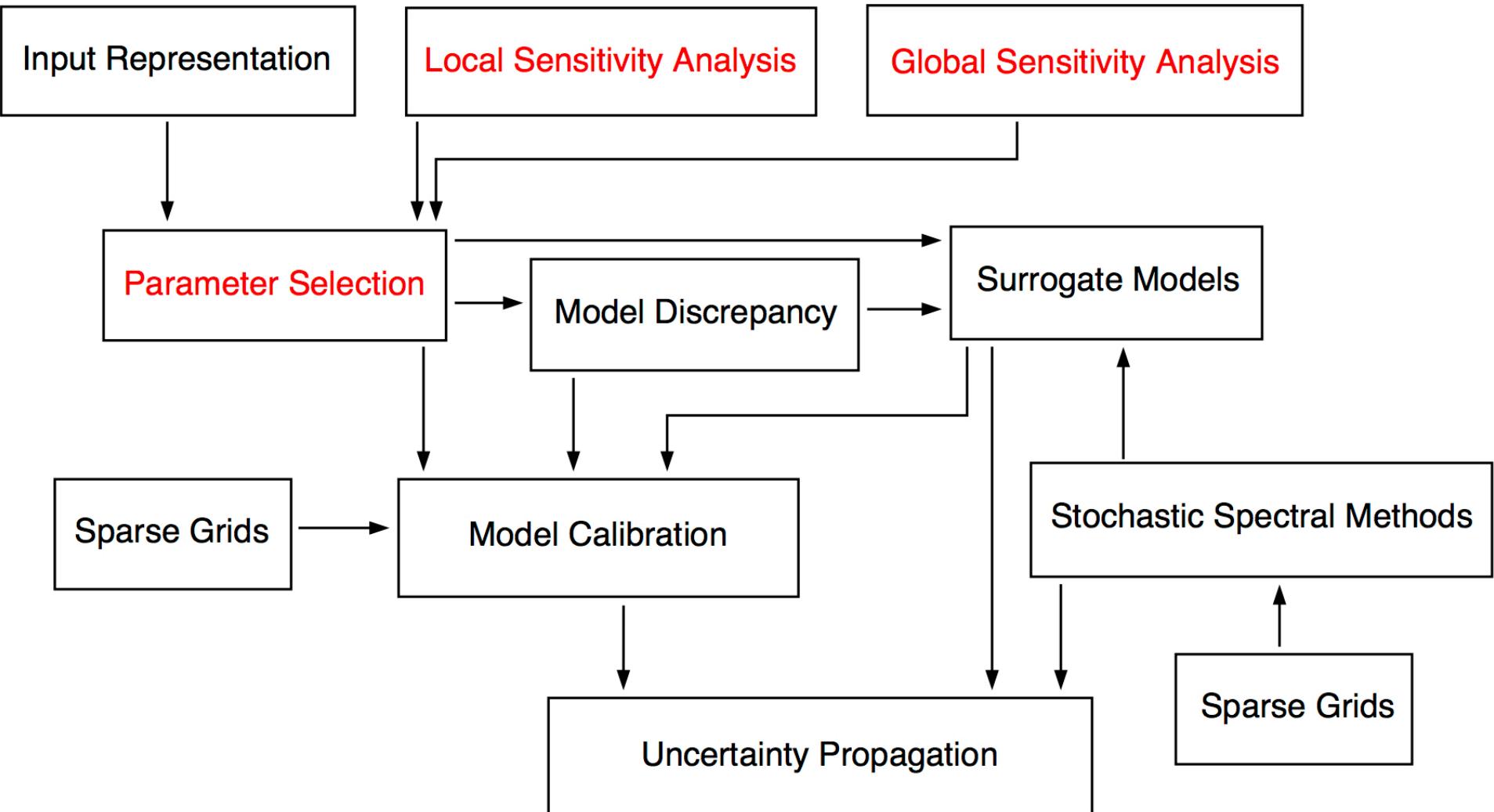


Steps in Uncertainty Quantification



Parameter Selection: Required for models with unidentifiable or noninfluential inputs

- e.g., SIR model

Parameter Selection Techniques and Surrogate Models

Parameter Space Reduction: SIR Model

$$\frac{dS}{dt} = \delta N - \delta S - \underline{\gamma kIS} , \quad S(0) = S_0 \quad \text{Susceptible}$$

$$\frac{dI}{dt} = \underline{\gamma kIS} - (r + \delta)I , \quad I(0) = I_0 \quad \text{Infectious}$$

$$\frac{dR}{dt} = rI - \delta R , \quad R(0) = R_0 \quad \text{Recovered}$$

Parameters:

- γ : Infection coefficient
- k : Interaction coefficient
- r : Recovery rate
- δ : Birth/death rate

Response:

$$y = \int_0^5 R(t, q) dt$$

Note: Parameters $q = [\gamma, k, r, \delta]$ not uniquely determined by data

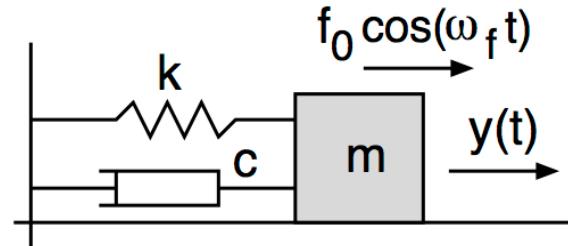
Parameter Selection Techniques

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

Example: Spring model

$$\underline{m} \frac{d^2z}{dt^2} + \underline{c} \frac{dz}{dt} + \underline{k} z = \underline{f_0} \cos(\omega_F t)$$

$$z(0) = z_0, \frac{dz}{dt}(0) = z_1$$



Problem: Parameters $q = [m, c, k, f_0]$ and $q = [1, \frac{c}{m}, \frac{k}{m}, \frac{f_0}{m}]$ yield same displacements

Parameter Selection Techniques

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

Example: Spring model

$$\underline{m} \frac{d^2z}{dt^2} + \underline{c} \frac{dz}{dt} + \underline{k} z = \underline{f_0} \cos(\omega_F t)$$

$$z(0) = z_0, \frac{dz}{dt}(0) = z_1$$

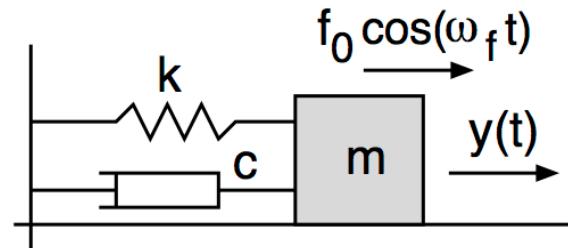
Problem: Parameters $q = [m, c, k, f_0]$ and $q = [1, \frac{c}{m}, \frac{k}{m}, \frac{f_0}{m}]$ yield same displacements

Solution: Reformulate problem as

$$\frac{d^2z}{dt^2} + \underline{C} \frac{dz}{dt} + \underline{K} z = \underline{F_0} \cos(\omega_F t)$$

$$z(0) = z_0, \frac{dz}{dt}(0) = z_1$$

where $C = \frac{c}{m}$, $K = \frac{k}{m}$ and $F_0 = \frac{f_0}{m}$



Techniques for General Models:

- Linear algebra analysis;
 - e.g., SVD or QR algorithms
- Sensitivity analysis
- Active Subspaces

Parameter Selection Techniques and Surrogate Models

Second Issue: Models can have thousands to millions of parameters

3-D Neutron Transport Equations:

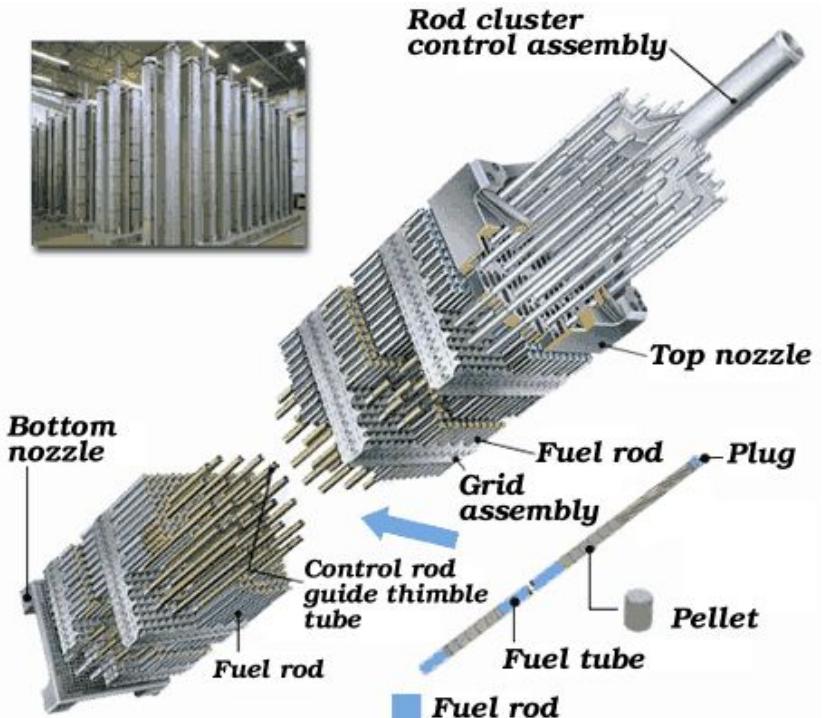
$$\begin{aligned} \frac{1}{|v|} \frac{\partial \varphi}{\partial t} + \Omega \cdot \nabla \varphi + \Sigma_t(r, E) \varphi(r, E, \Omega, t) \\ = \int_{4\pi} d\Omega' \int_0^\infty dE' \Sigma_s(E' \rightarrow E, \Omega' \rightarrow \Omega) \varphi(r, E', \Omega', t) \\ + \underline{\frac{\chi(E)}{4\pi}}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

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



Global Sensitivity Analysis

Example: Portfolio model

$$Y = c_1 Q_1 + c_2 Q_2$$

Note:

- Q_1 and Q_2 represent hedged portfolios
- c_1 and c_2 amounts invested in each portfolio

Take

$$c_1 = 2, c_2 = 1$$

$$Q_1 \sim N(0, 1)$$

$$Q_2 \sim N(0, 9)$$

Local Sensitivities:

$$\frac{\partial Y}{\partial Q_1} = 2, \quad \frac{\partial Y}{\partial Q_2} = 1$$

Conclusion: Investment is more sensitive to Portfolio 1 than to Portfolio 2

Limitations:

- Does not accommodate potential uncertainty in parameters.
- Sensitive to units and magnitudes of parameters.

Global Sensitivity Analysis

Example: Portfolio model

$$Y = c_1 Q_1 + c_2 Q_2$$

Note:

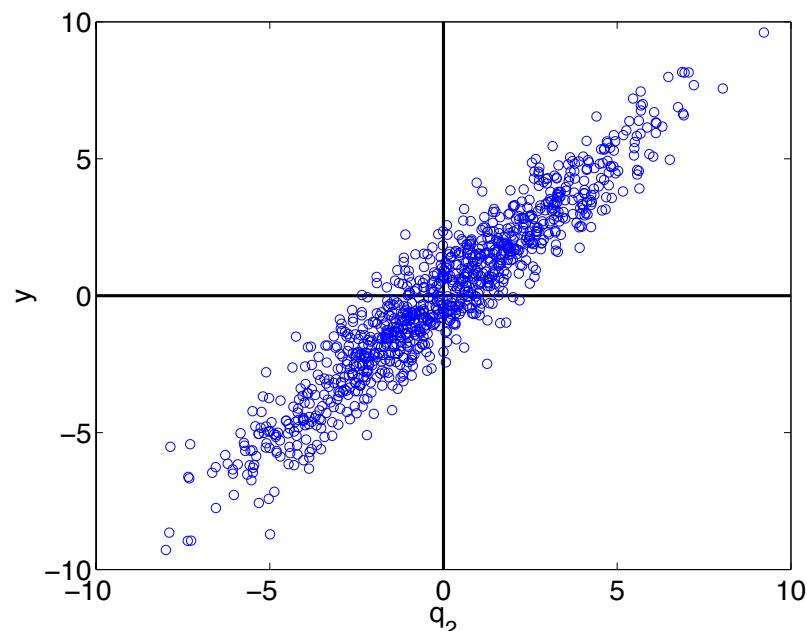
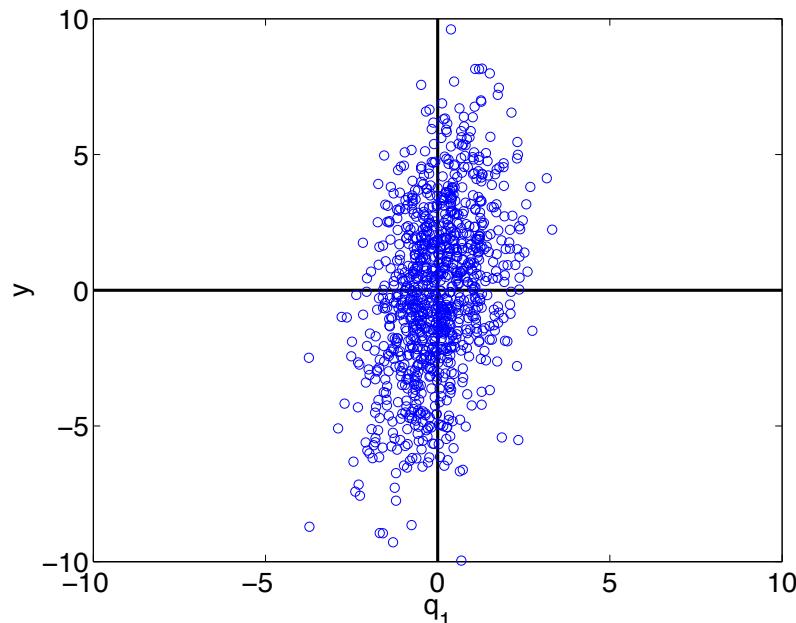
- Q_1 and Q_2 represent hedged portfolios
- c_1 and c_2 amounts invested in each portfolio

Take

$$c_1 = 2, c_2 = 1$$

$$Q_1 \sim N(0, 1)$$

$$Q_2 \sim N(0, 9)$$



Local Sensitivities:

$$\frac{\partial Y}{\partial Q_1} = 2, \quad \frac{\partial Y}{\partial Q_2} = 1$$

Solutions:

- Response correlation
- Variance-based methods
- Random sampling of local sensitivities

Global Sensitivity Analysis: Variance-Based Methods

Example: Portfolio model

$$Y = c_1 Q_1 + c_2 Q_2$$

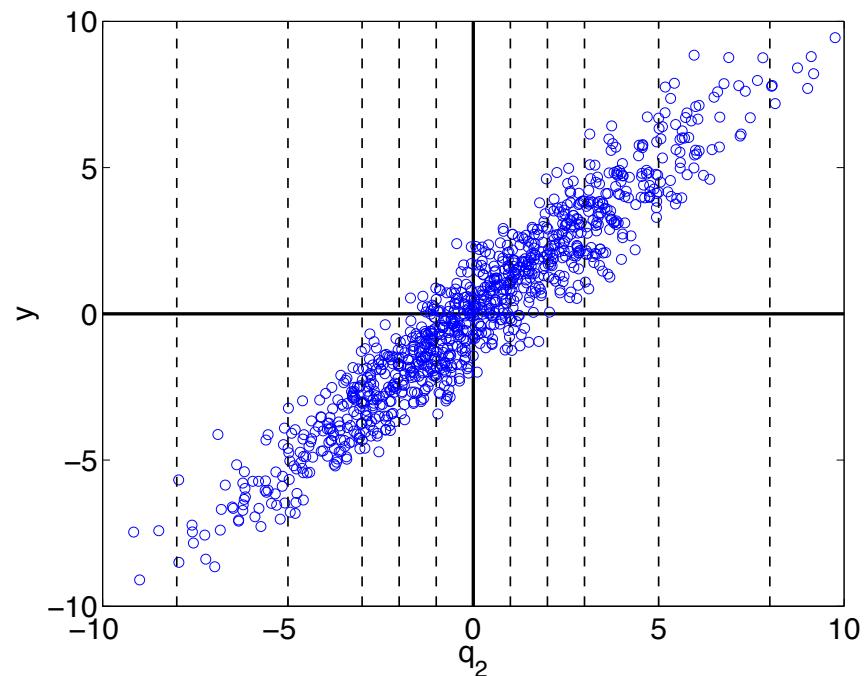
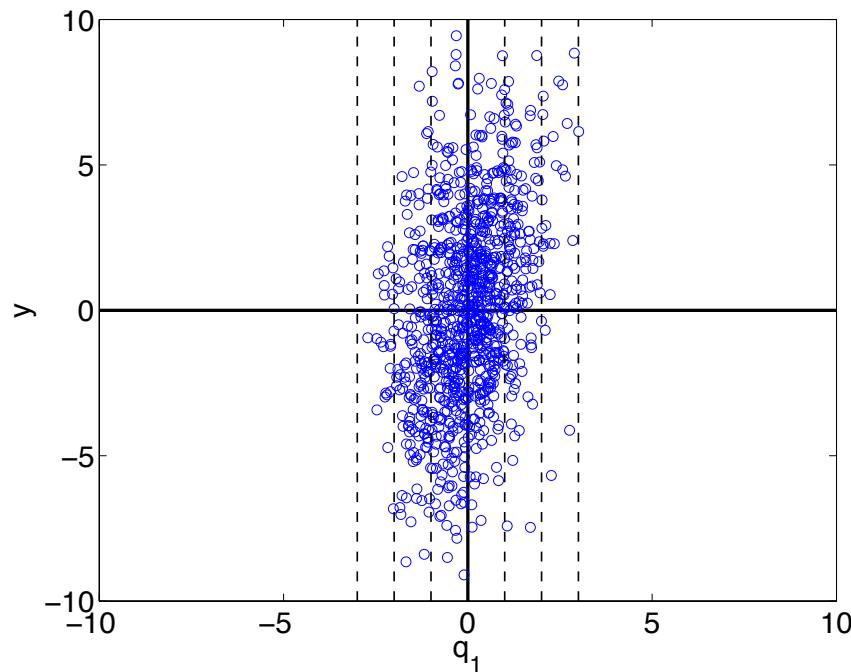
Take $c_1 = 2, c_2 = 1$

$$Q_1 \sim N(0, 1)$$

$$Q_2 \sim N(0, 9)$$

Statistical Motivation: Consider variability of expected values

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



Note: Here $D_2 > D_1$

Variance-Based Methods

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

Take

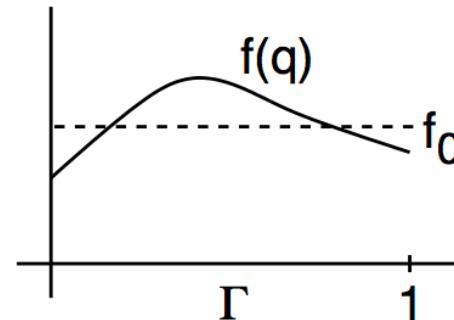
$$f(q) = f_0 + \sum_{i=1}^p f_i(q_i) + \sum_{1 \leq i < j \leq p} f_{ij}(q_i, q_j)$$

Analogy: Taylor or Fourier series

With appropriate assumptions,

$$f_0 = \int_{\Gamma} f(q) dq$$

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



Variance-Based Methods

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

Take

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

Analogy: Taylor or Fourier series

Assumption: Mutually independent parameters

With appropriate assumptions,

$$f_0 = \int_{\Gamma} f(q) dq$$

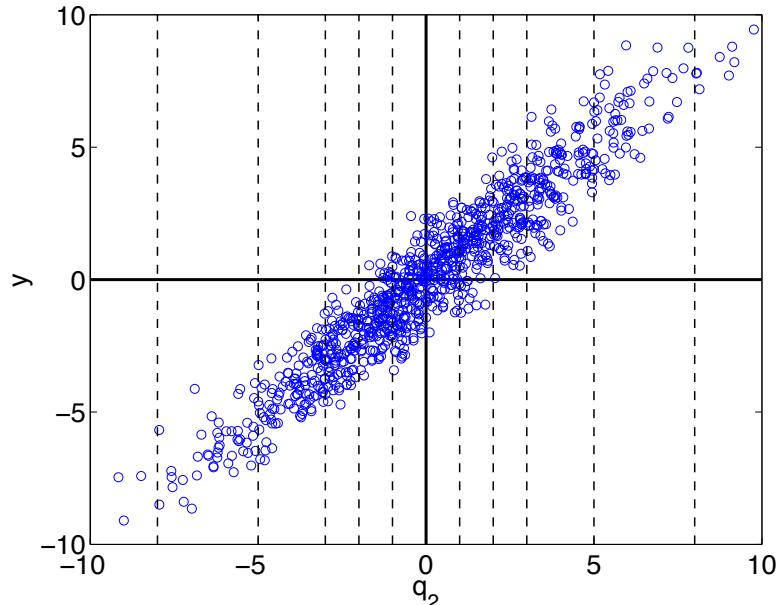
$$f_i(q_i) = \int_{\Gamma^{p-1}} f(q) dq_{\sim i} - f_0$$

Variances:

$$D_i = \int_0^1 f_i^2(q_i) dq_i$$

$$D = \text{var}(Y)$$

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

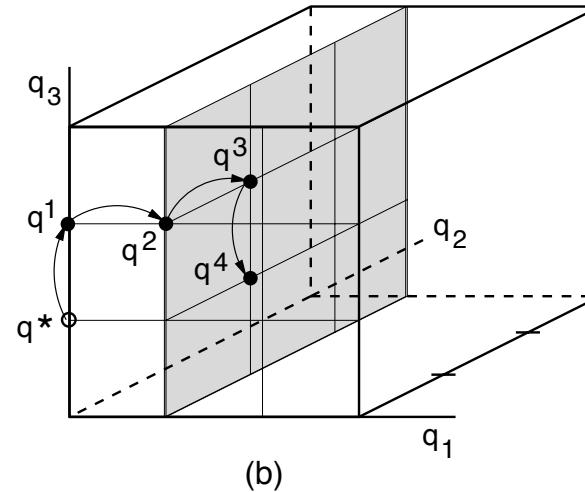
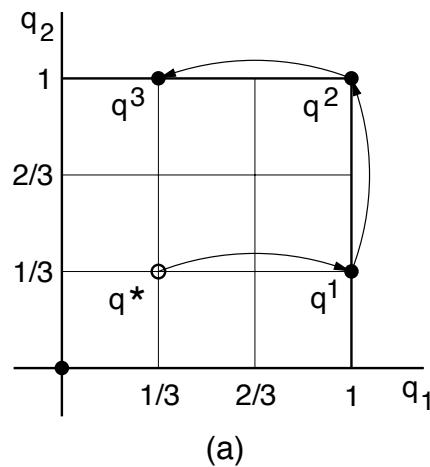


Statistical Interpretation:

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

Morris Screening: Random Sampling of Approximated Derivatives

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



Elementary Effect:

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

Global Sensitivity Measures: r samples

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

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

SIR Disease Example

SIR Model:

$$\frac{dS}{dt} = \delta N - \delta S - \underline{\gamma kIS} , \quad S(0) = S_0 \quad \text{Susceptible}$$

$$\frac{dI}{dt} = \underline{\gamma kIS} - (r + \delta)I , \quad I(0) = I_0 \quad \text{Infectious}$$

$$\frac{dR}{dt} = rI - \delta R , \quad R(0) = R_0 \quad \text{Recovered}$$

Note: Parameter set $q = [\gamma, k, r, \delta]$ is not identifiable

Assumed Parameter Distribution:

$$\gamma \sim \mathcal{U}(0, 1) , \quad k \sim \text{Beta}(\alpha, \beta) , \quad r \sim \mathcal{U}(0, 1) , \quad \delta \sim \mathcal{U}(0, 1)$$

Infection Coefficient	Interaction Coefficient	Recovery Rate	Birth/death Rate

Response:

$$y = \int_0^5 R(t, q) dt$$

SIR Disease Example

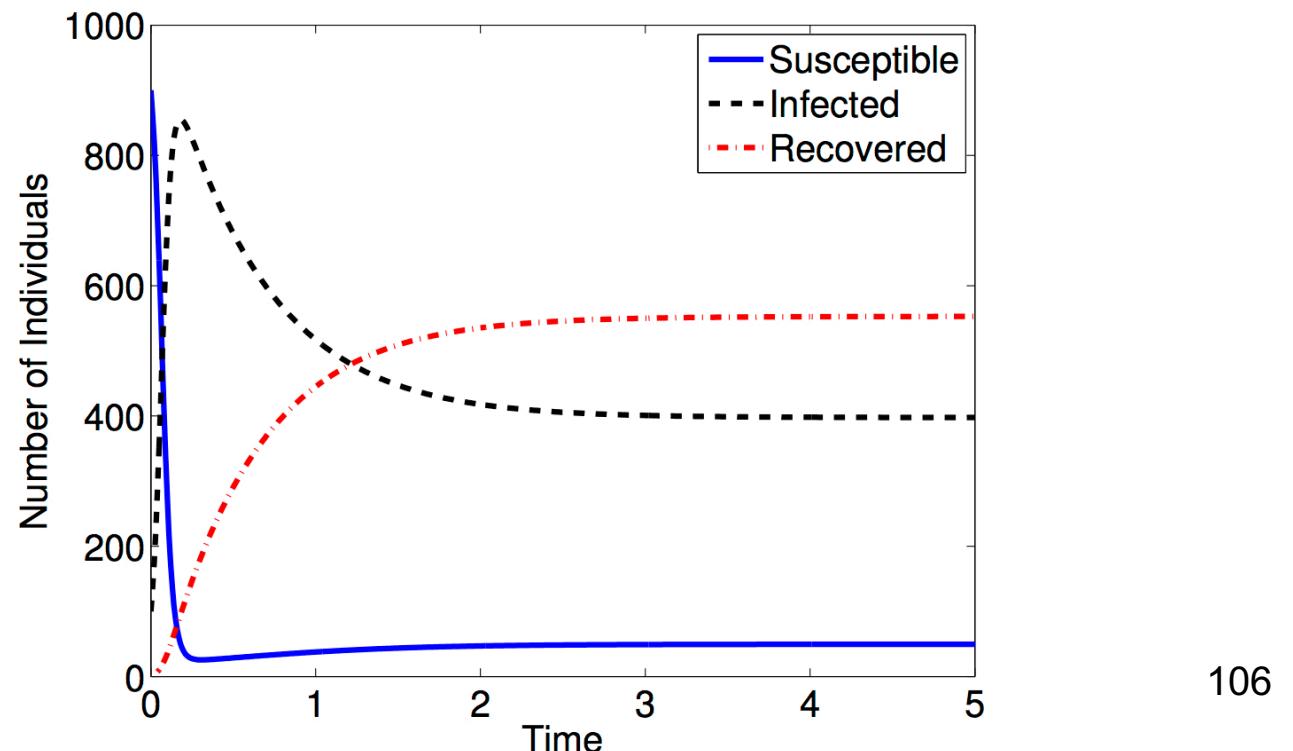
SIR Model:

$$\frac{dS}{dt} = \delta N - \delta S - \underline{\gamma kIS} , \quad S(0) = S_0 \quad \text{Susceptible}$$

$$\frac{dI}{dt} = \underline{\gamma kIS} - (r + \delta)I , \quad I(0) = I_0 \quad \text{Infectious}$$

$$\frac{dR}{dt} = rI - \delta R , \quad R(0) = R_0 \quad \text{Recovered}$$

Typical Realization:



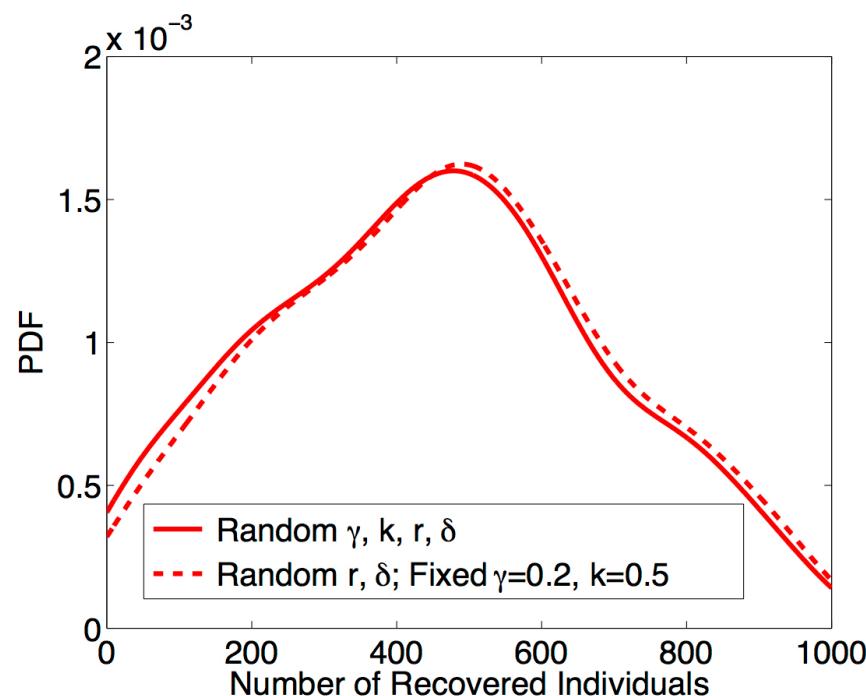
SIR Disease Example

Global Sensitivity Measures:

		γ	k	r	δ
Sobol	S_i	0.0997	0.0312	0.7901	0.1750
	S_{T_i}	-0.0637	-0.0541	0.5634	0.2029
Morris	$\mu_i^* (\times 10^3)$	0.2532	0.2812	2.0184	1.2328
	$\sigma_i (\times 10^3)$	0.9539	1.6245	6.6748	3.9886

Result: Densities for $R(t_f)$ at $t_f = 5$

Influential Parameters



Note: Can fix non-influential parameters γ, k

Parameter Selection: Nuclear Power Plant Design

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_tmags	-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_xdge	-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

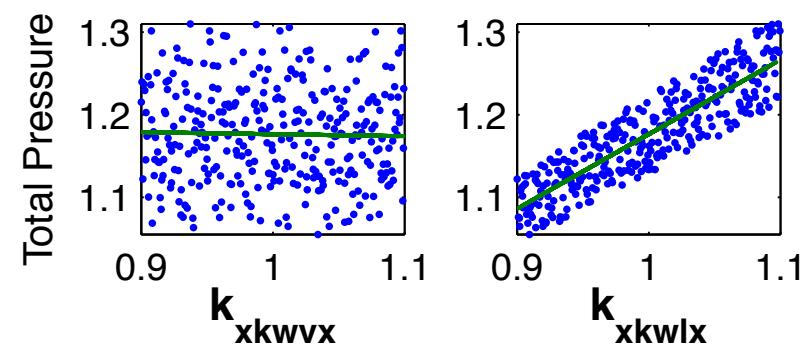
k_xkwlx: Vertical liquid wall drag coefficient

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 VUQ parameters reduced to 5 via sensitivity analysis

Global Sensitivity Analysis: Potential Pitfalls

Example: Quantum-informed continuum model

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

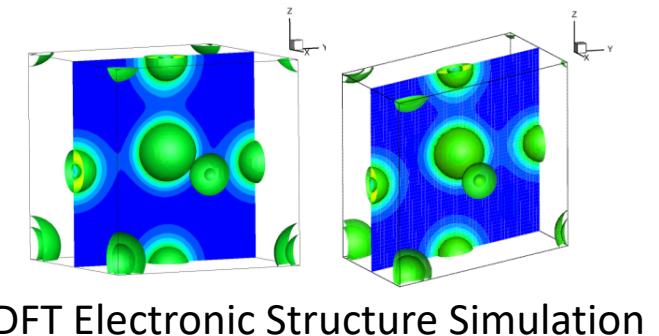
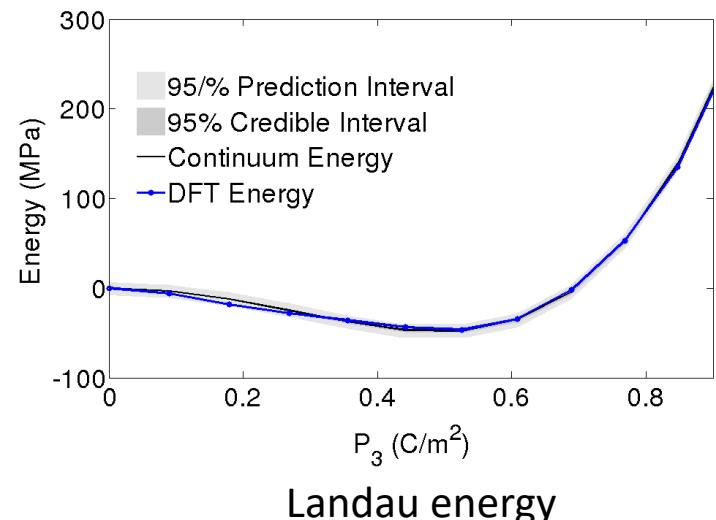
$$\psi(P, q) = \alpha_1 P^2 + \alpha_{11} P^4 + \alpha_{111} P^6$$

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:

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

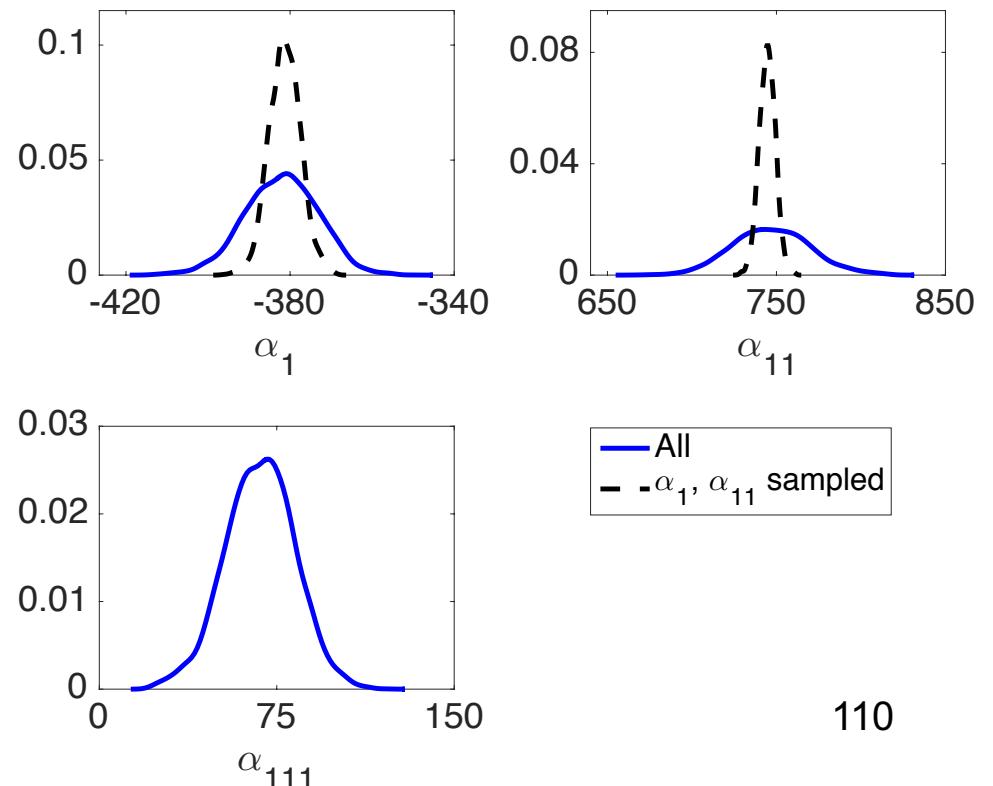
Global Sensitivity Analysis:

	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?

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

Problem:

- Parameters correlated
- Cannot fix α_{111}

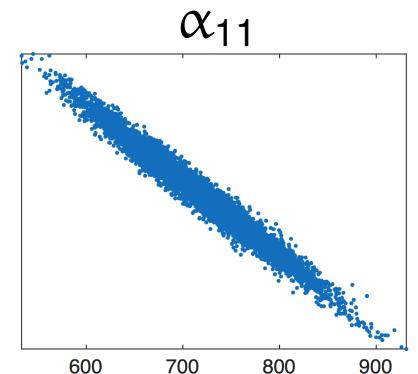
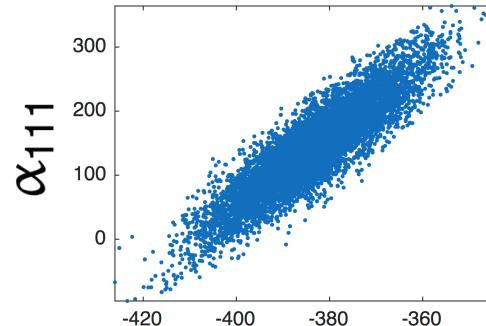
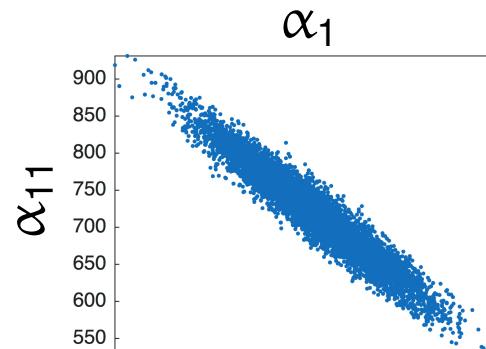
Parameters:

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

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



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

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

Sobol' Indices:

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

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.

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.

Active Subspaces

Note:

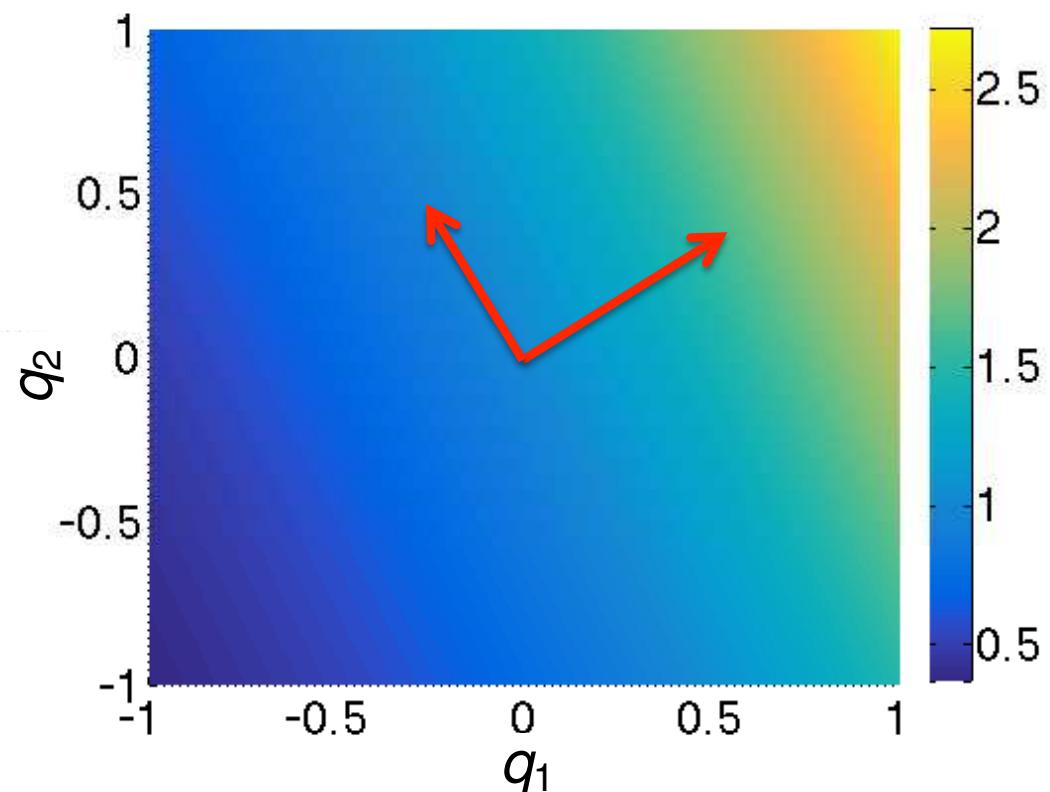
- Functions may vary significantly in only a few directions
- “Active” directions may be linear combination of inputs

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:

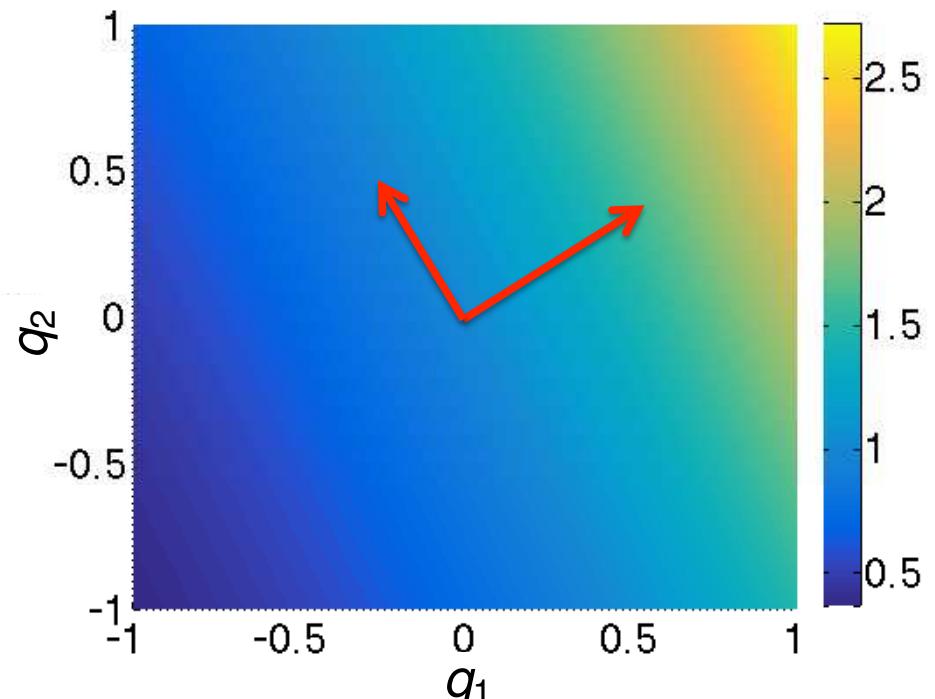
- Functions may vary significantly in only a few directions
- “Active” directions may be linear combination of inputs

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.



Parameter Space Reduction Techniques: 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

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

Singular Value Decomposition (SVD):

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

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

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.

Random Range Finding Algorithms: Linear Problems

Algorithm: Halko, Martinsson and Tropp, SIAM Review, 2011

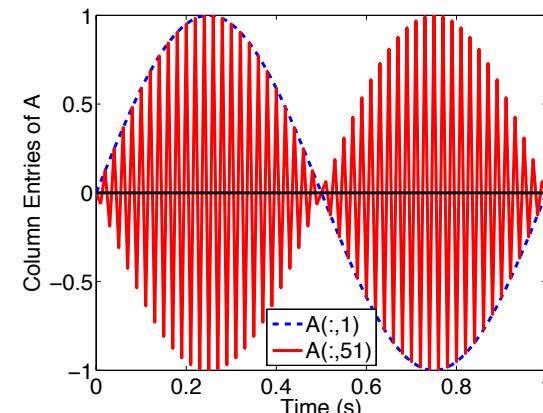
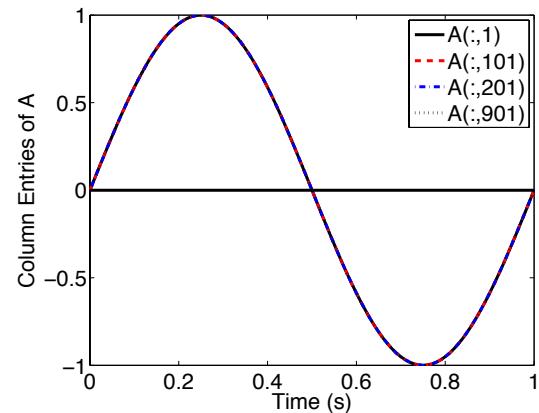
1. Choose ℓ random inputs q^i and compute outputs $y^i = Aq^i$ which are compiled in the $m \times \ell$ matrix Y .
2. Take a pivoted QR factorization $Y = QR$ to construct a matrix Q whose columns form an orthonormal basis for the range of Y .

Example: $y_i = \sum_{k=1}^p q_k \sin(2\pi k t_i)$, $i = 1, \dots, m$

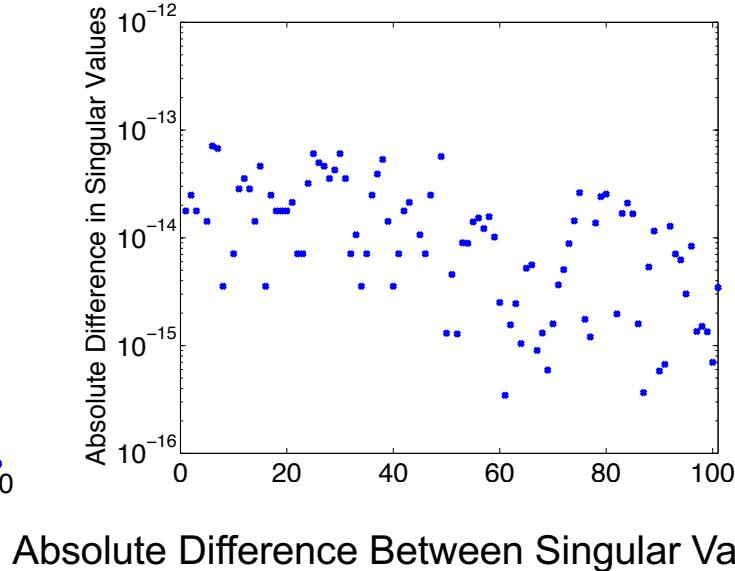
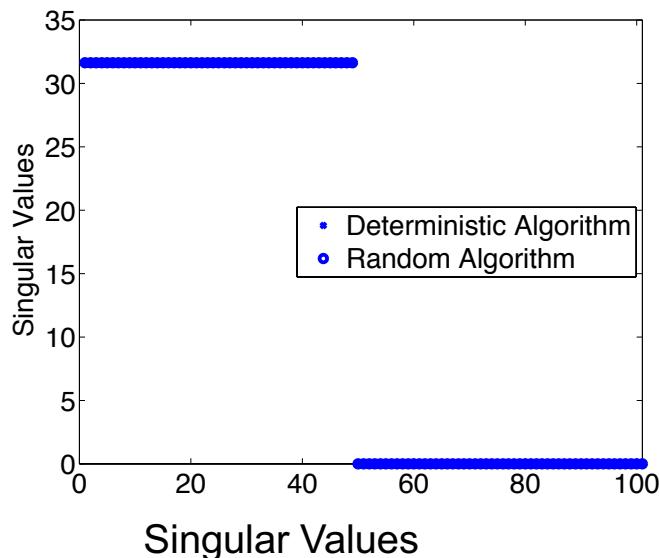
$$\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} \sin(2\pi t_1) & \cdots & \sin(2\pi p t_1) \\ \vdots & & \vdots \\ \sin(2\pi t_m) & \cdots & \sin(2\pi p t_m) \end{bmatrix} \begin{bmatrix} q_1 \\ \vdots \\ q_p \end{bmatrix}$$

Random Range Finding Algorithms: Linear Problems

Example: $m = 101$, $p = 1000$: Analytic value for rank is 49



Aliasing



Example: $m = 101$, $p = 1,000,000$: Random algorithm still viable

Active Subspaces

Note:

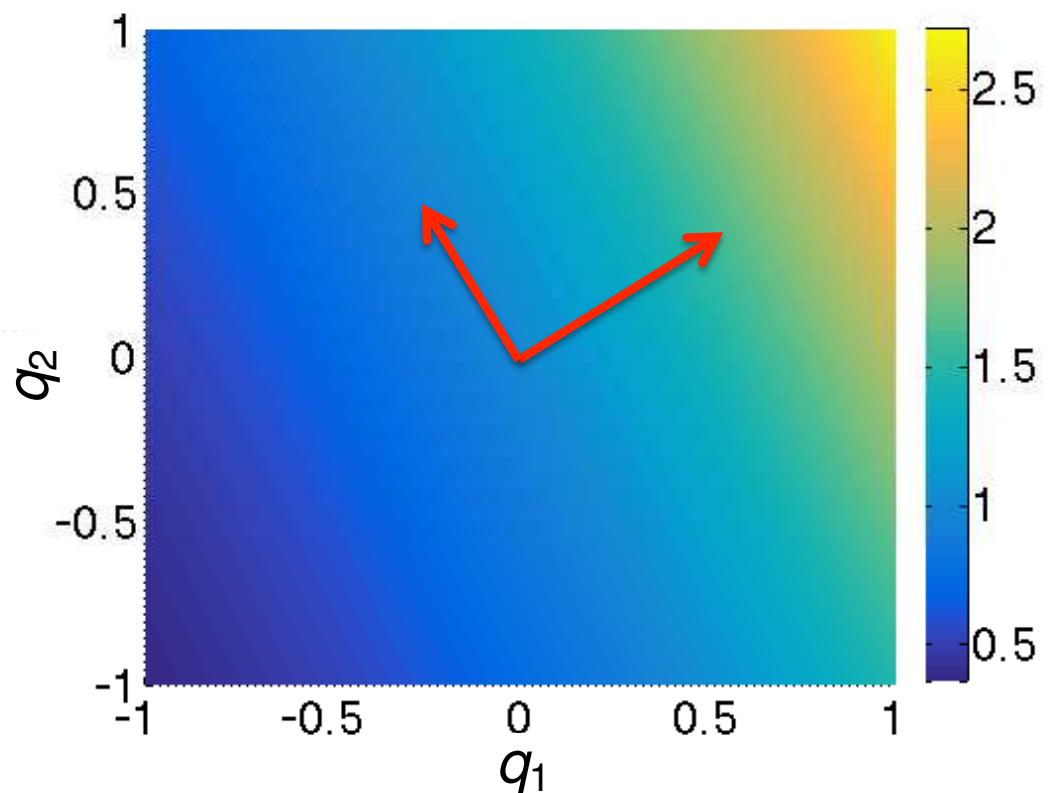
- Functions may vary significantly in only a few directions
- “Active” directions may be linear combination of inputs

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 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

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

Construct outer product

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

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

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 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

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

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

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

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

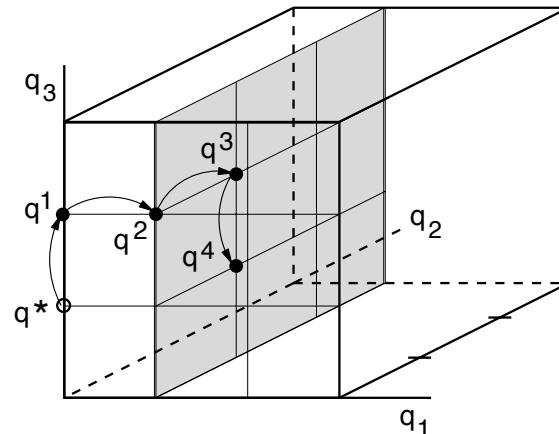
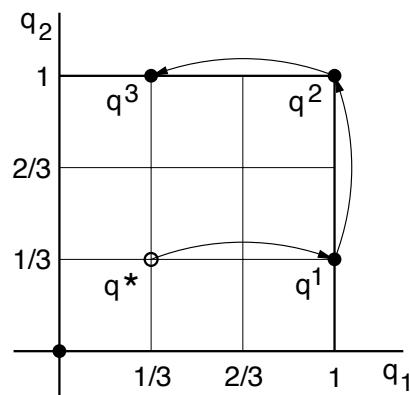
Goal: Develop efficient algorithm for codes that do not have adjoint capabilities

Note: Finite difference approximations tempting but not very effective

Strategy: Algorithm based on initialized adaptive Morris indices

Morris Screening: Random Sampling of Approximated Derivatives

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



Elementary Effect:

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

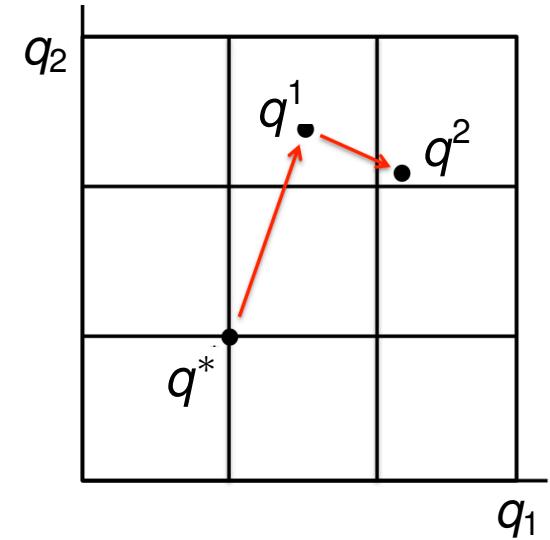
Global Sensitivity Measures: r samples

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

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

Adaptive Algorithm:

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



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

Initialization Algorithm

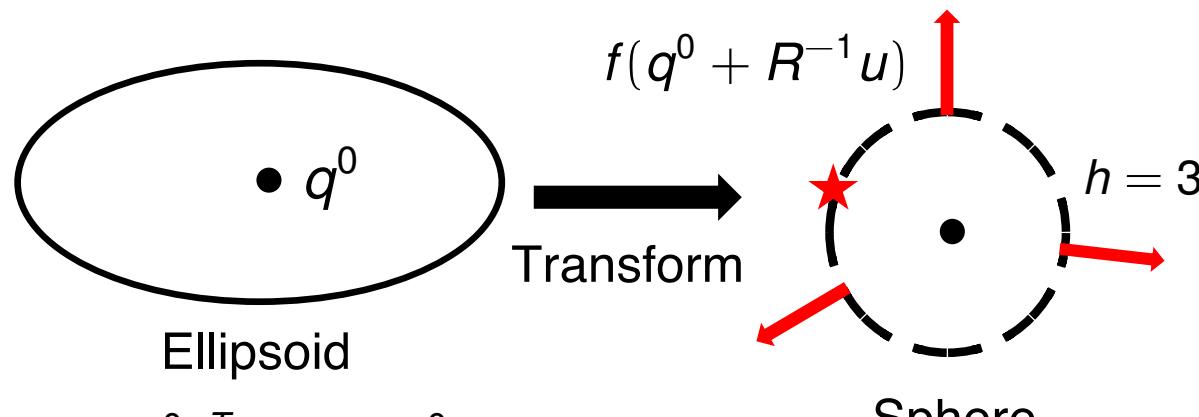
1. Inputs: ℓ iterations, h function evaluations per iteration
2. Sample w^1 from surface of unit sphere where approximately linear

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

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

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

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

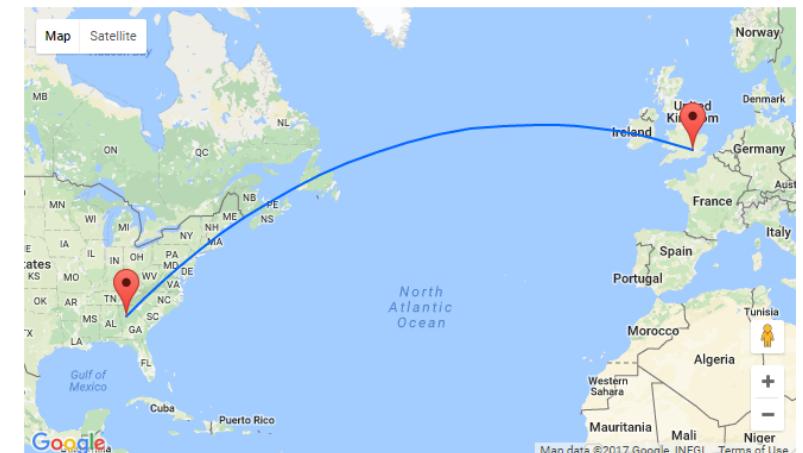


$$(z - q^0)^T S(z - q^0) = 1$$

$$S = R^T R$$

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

Example: Let $w^1 = \text{Atlanta}$, $v^1 = \text{London}$, and $g(u) = \text{'QUIETness'}$ of seatmate on flight



Initialization Algorithm

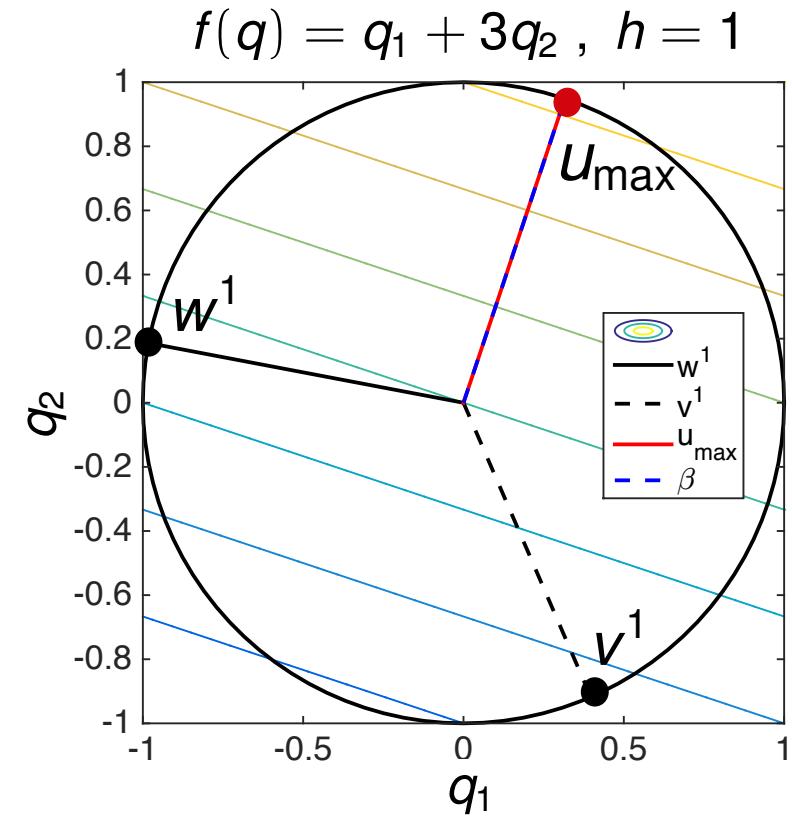
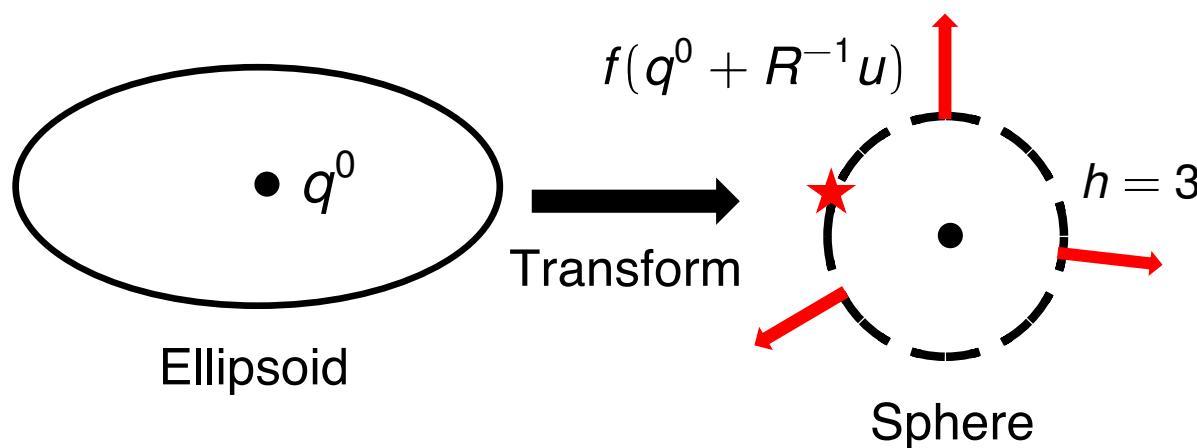
1. Inputs: ℓ iterations, h function evaluations per iteration
2. Sample w^1 from surface of unit sphere where approximately linear

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

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

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

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 $j = 1, \dots, \ell$

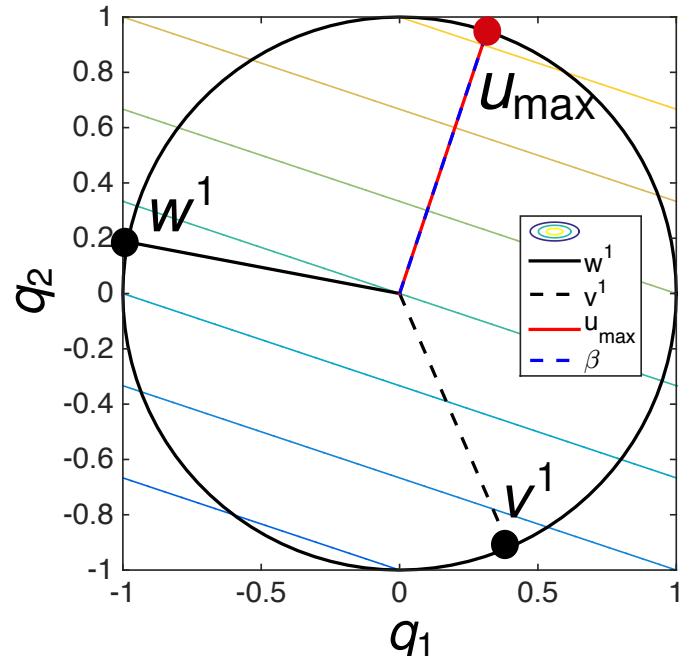
3. Sample $\{\tilde{v}_1^j, \dots, \tilde{v}_h^j\}$ from surface of sphere
4. Use Lagrange multiplier to determine

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

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 $C_{j\perp} = [\text{span}(C_{(j-1)\perp}, (I_m - P_{u_{\max}^j}) C)]$ and set $P_{C_{j\perp}} = C_{j\perp} (C_{j\perp}^T C_{j\perp})^{-1} C_{j\perp}^T$
7. Take $v_i^j = \frac{(I_m - P_{C_{j\perp}}) \tilde{v}_i^j}{\|(I_m - P_{C_{j\perp}}) \tilde{v}_i^j\|}, \quad i = 1, \dots, h$ and repeat



Ortho-complement
of u_{\max}^j

Example: Initialization Algorithm to Approximate Gradient

Example: Family of elliptic PDE's

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

with the random field representations

$$a(q, s, \ell) = a_{\min} + e^{\bar{a}(s, \ell) + \sum_{i=1}^p q_k^\ell \gamma_i \phi_i(s)}$$

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

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

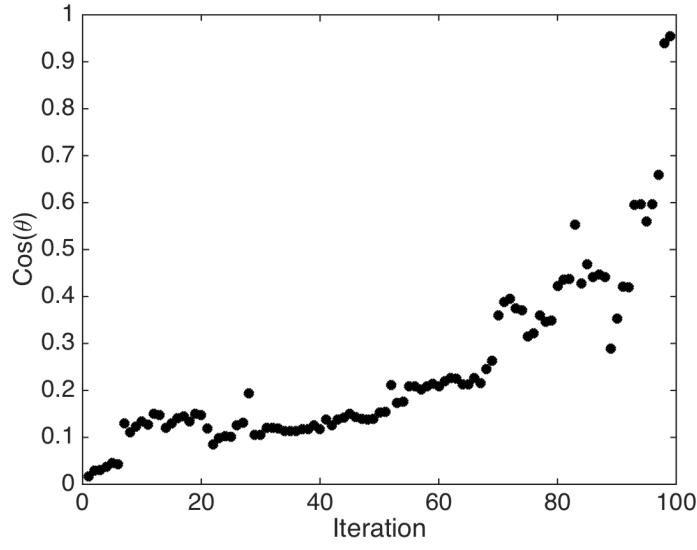


Problem Dimensions:

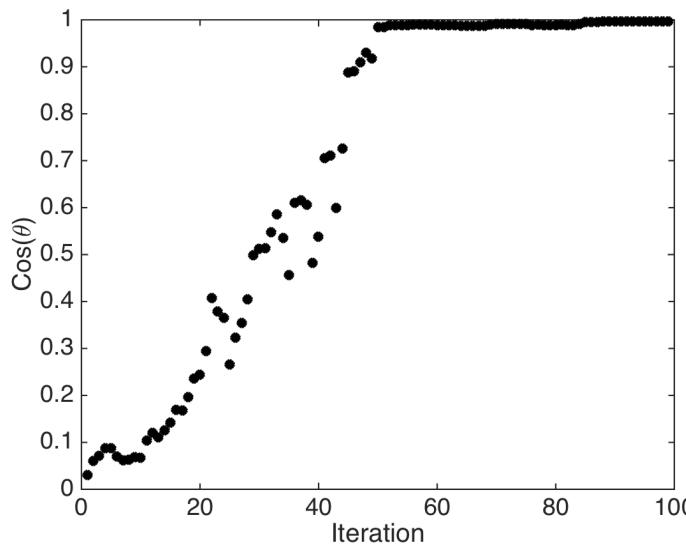
- Parameter dimension: $p = 100$
- Active subspace dimension: $n = 1$
- Finite element approximation

Example: Initialization Algorithm to Approximate Gradient

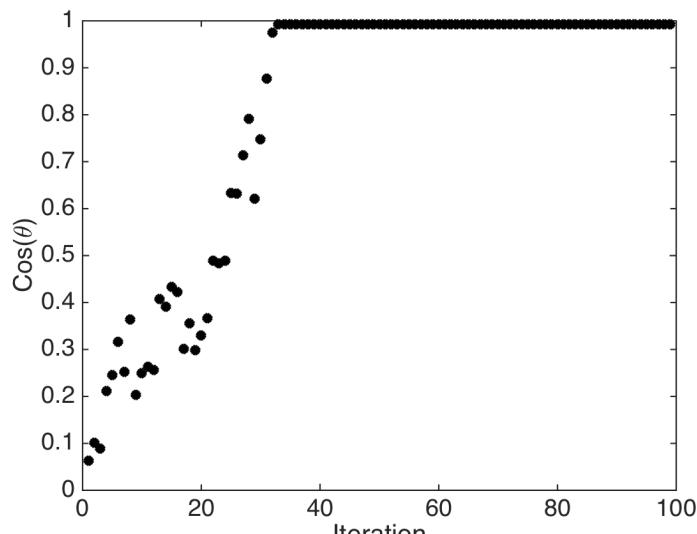
Results: Cosine of angle between 'analytic' and computed gradient



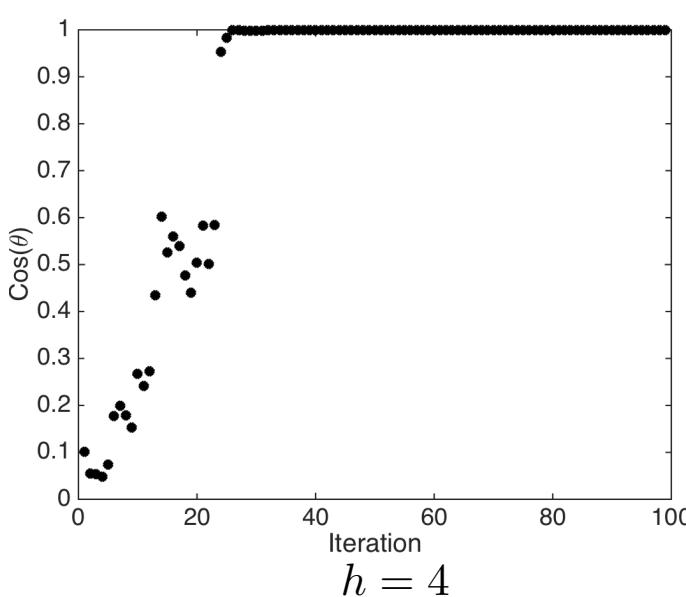
$h = 1$



$h = 2$



$h = 3$



$h = 4$

Recall: $p=100$

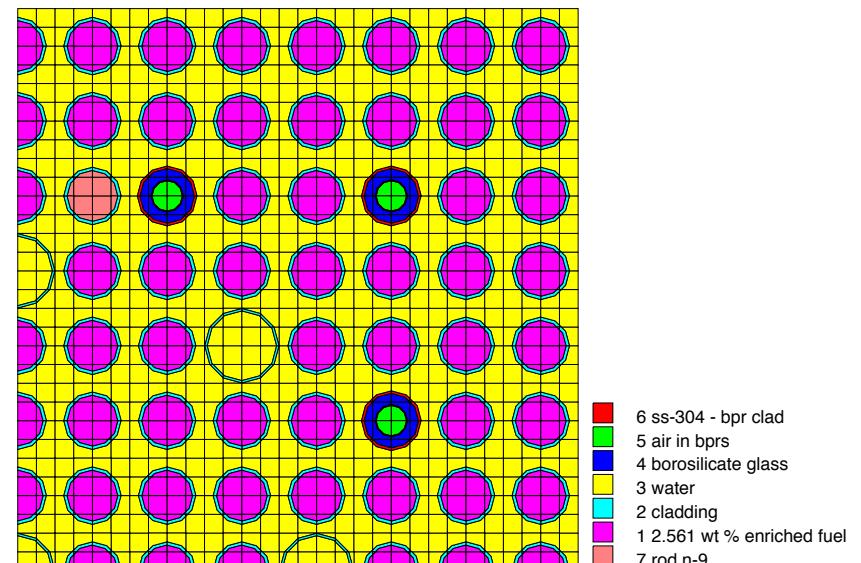
Note: Convergence within $h \cdot \ell$ iterations

SCALE6.1: High-Dimensional Example

Setup: Cross-section computations SCALE6.1

- Input Dimension: 7700
- Output k_{eff}

Materials			Reactions	
$^{234}_{92}\text{U}$	$^{10}_{5}\text{B}$	$^{31}_{15}\text{P}$	Σ_t	$n \rightarrow \gamma$
$^{235}_{92}\text{U}$	$^{11}_{5}\text{B}$	$^{55}_{25}\text{Mn}$	Σ_e	$n \rightarrow p$
$^{236}_{92}\text{U}$	$^{14}_{7}\text{N}$	$^{26}_{26}\text{Fe}$	Σ_f	$n \rightarrow d$
$^{238}_{92}\text{U}$	$^{15}_{7}\text{N}$	$^{116}_{50}\text{Sn}$	Σ_c	$n \rightarrow t$
^1_1H	$^{23}_{11}\text{Na}$	$^{120}_{50}\text{Sn}$	$\bar{\nu}$	$n \rightarrow {}^3\text{He}$
$^{16}_8\text{O}$	$^{27}_{13}\text{Al}$	$^{40}_{20}\text{Zr}$	χ	$n \rightarrow \alpha$
^6_6C	$^{14}_{14}\text{Si}$	$^{19}_{19}\text{K}$	$n \rightarrow n'$	$n \rightarrow 2n$



Note: We cannot efficiently approximate all directional derivatives required to approximate the gradient matrix. Requires efficient initialization algorithm.

SCALE6.1: High-Dimensional Example

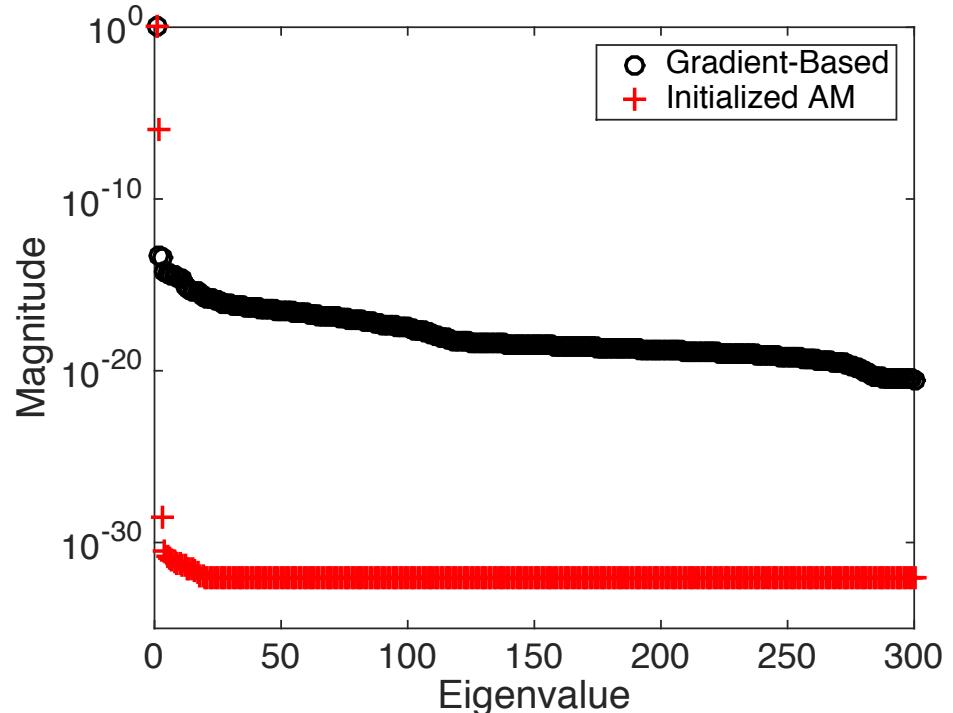
Setup:

- Input Dimension: 7700

SCALE Evaluations:

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

Note: Analytic eigenvalues: 0, 1



Active Subspace Dimensions:

For surrogate sampled off space

Method	Gap	PCA				Error Tolerance			
		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

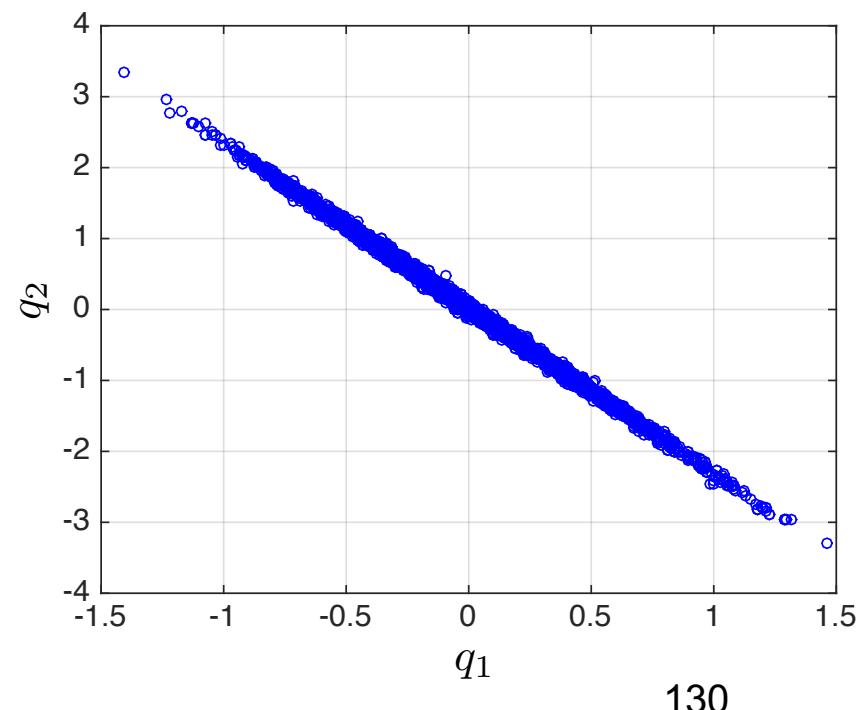
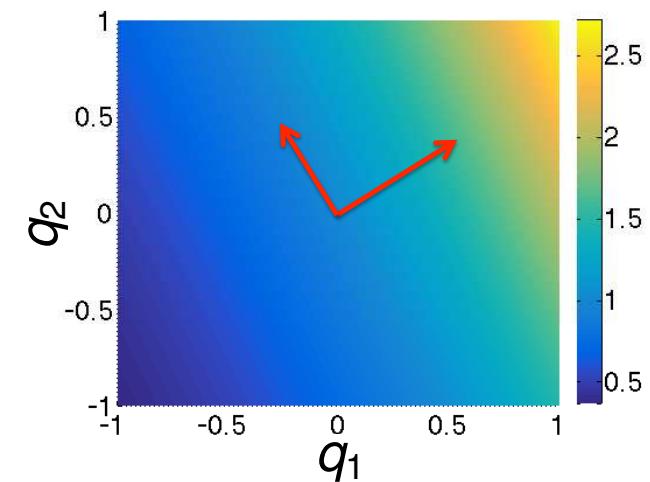
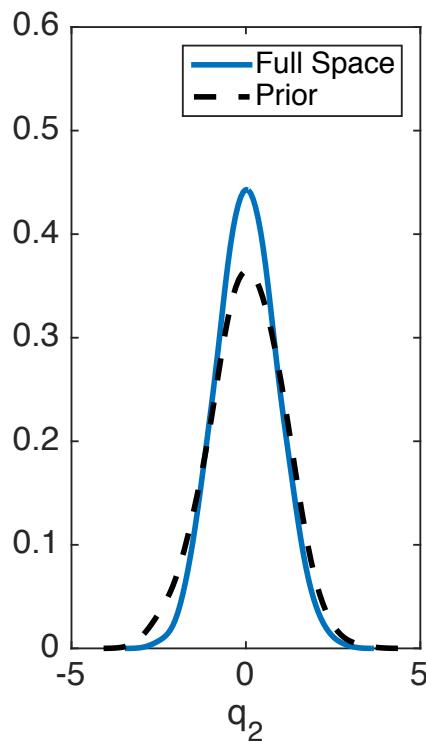
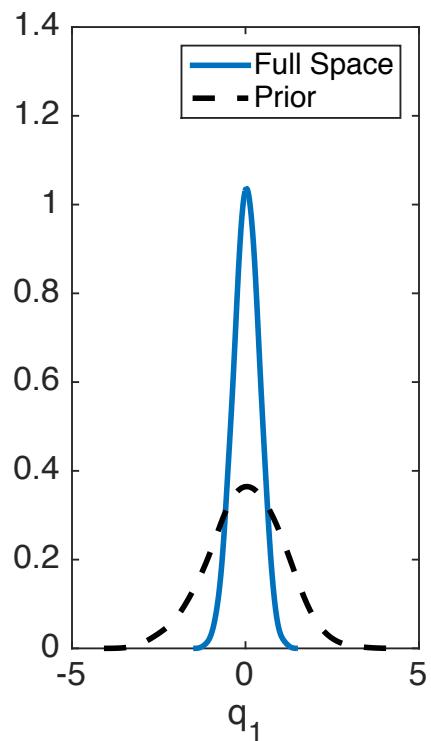
Notes: Computing *converged* adjoint solution is expensive and *often not achieved*

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

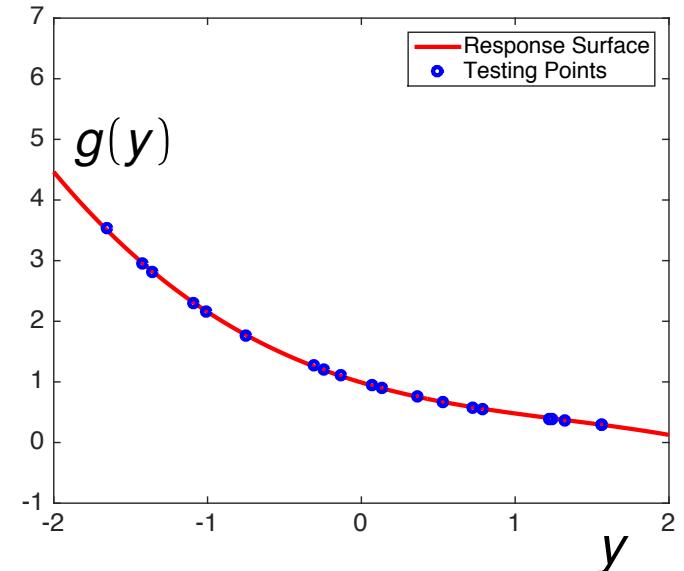
Example: $y = \exp(0.7q_1 + 0.3q_2)$

Active Subspace: For gradient matrix G , form SVD

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

Eigenvalue spectrum indicates 1-D active subspace with basis

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

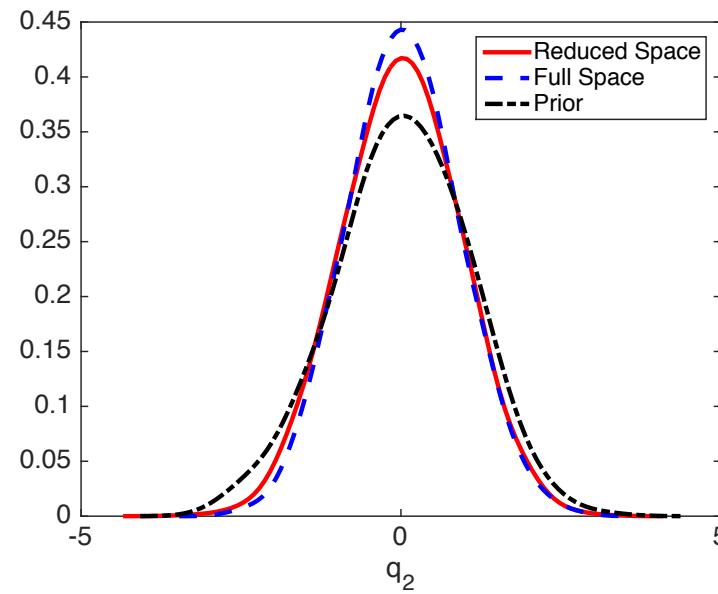
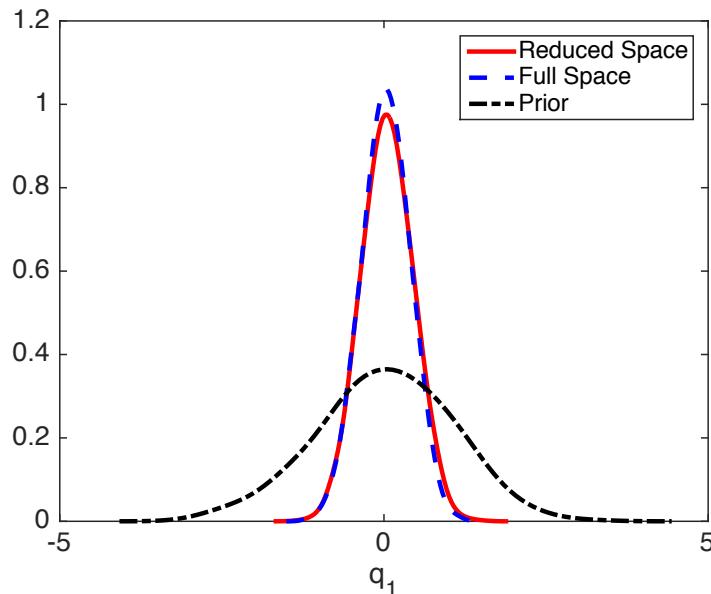


Strategy: Inference based on active subspace

- For values $\{q^j\}_{j=1}^M$, compute $y^j = U(:, 1)^T q^j$ and fit response surface $g(y)$
- Use DRAM to calibrate 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

$$\alpha(N) = \sum_{j=1}^N \lambda_j w_j^2$$

Note: Here $N = 1$ and $w_j^2 = U(:, 1) \cdot * U(:, 1) = [0.91^2, 0.39^2]$

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

Bayesian Inference on Active Subspaces

Example: Family of elliptic PDE's

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

with the random field representations

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

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

$$f(q^1, \dots, q^N) \approx \sum_{n=1}^N \frac{1}{|\Gamma_2|} \int_{\Gamma_2} u(a(s, n)) ds$$

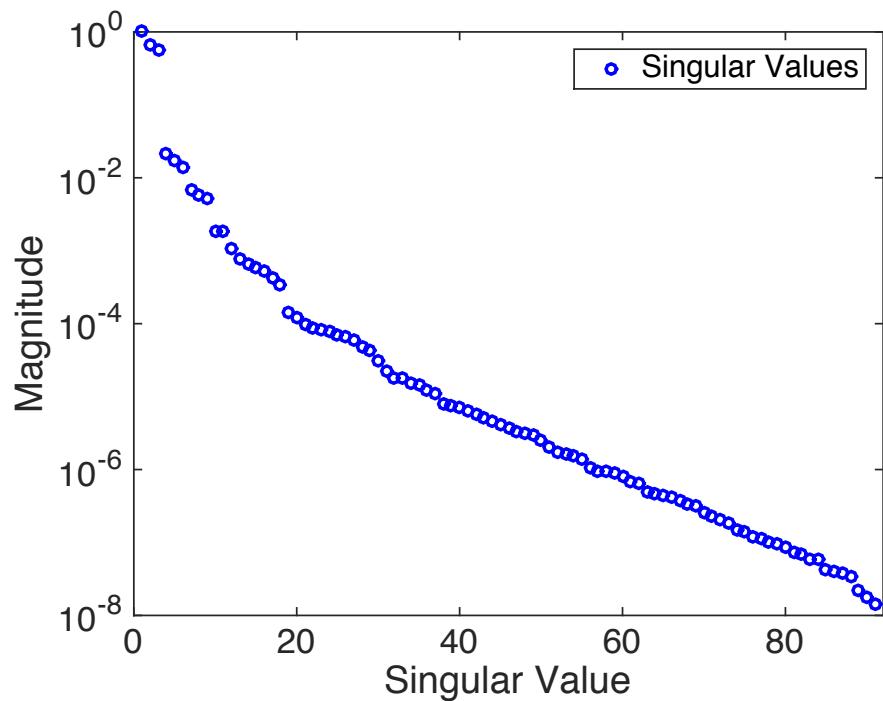


Problem Dimensions:

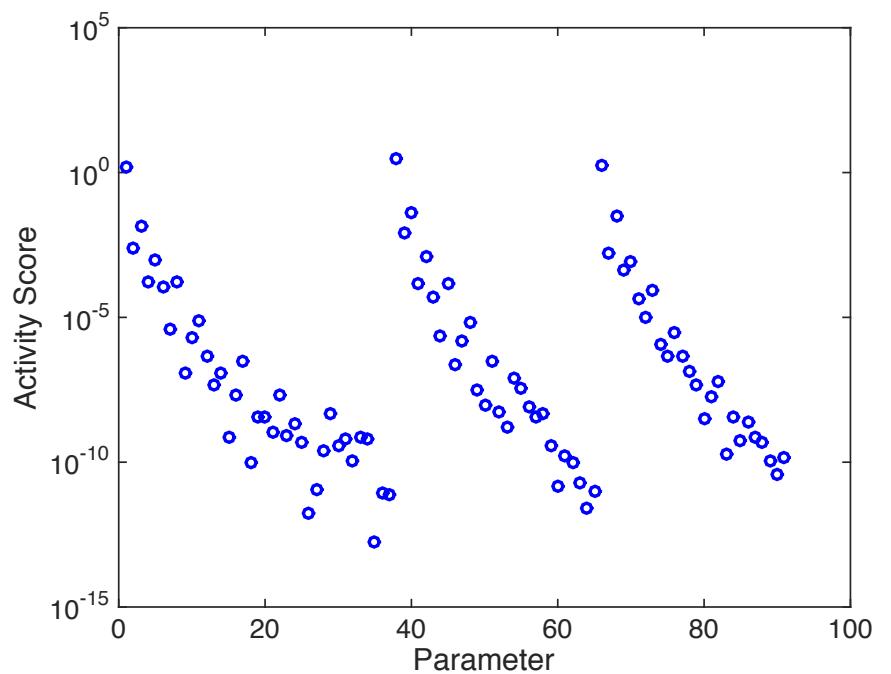
- Parameter dimension: $p = 91$
- Active subspace dimension: $N = 3$
- Finite element space: 1372 triangular elements, 727 nodes

Bayesian Inference on Active Subspaces

Singular Values: Recall $N = 3$



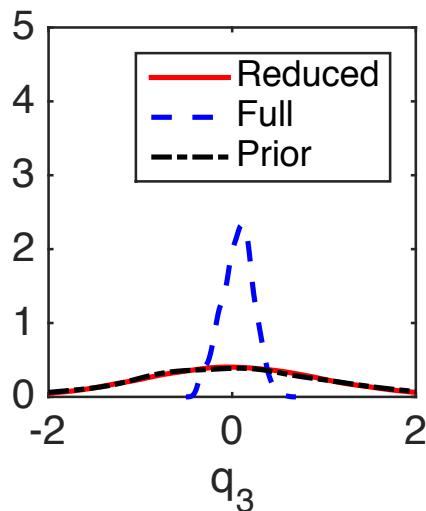
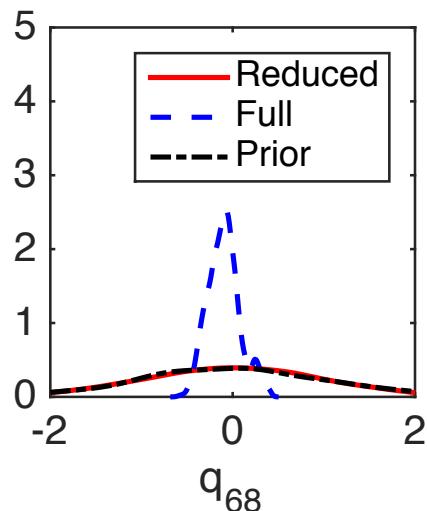
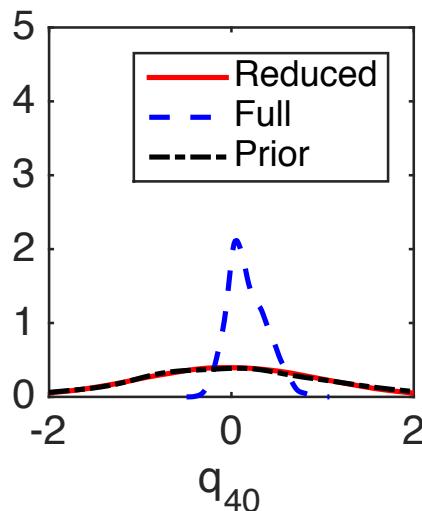
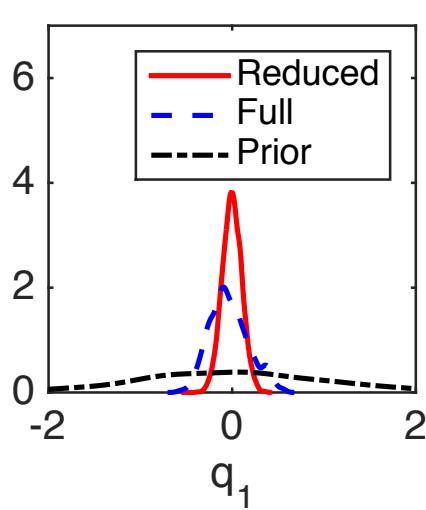
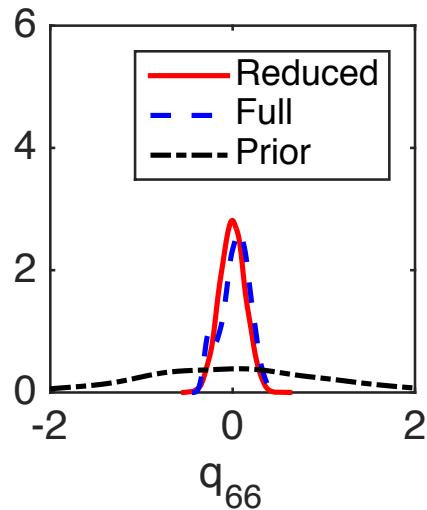
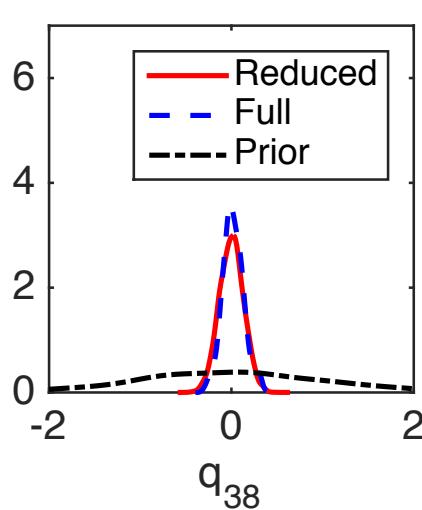
Activity Scores: Quantify global sensitivity



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

