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

\[
\frac{dS}{dt} = \delta N - \delta S - \gamma k l S \quad , \quad S(0) = S_0 \quad \text{Susceptible}
\]

\[
\frac{dl}{dt} = \gamma k l S - (r + \delta) l \quad , \quad l(0) = l_0 \quad \text{Infectious}
\]

\[
\frac{dR}{dt} = r l - \delta R \quad , \quad R(0) = R_0 \quad \text{Recovered}
\]

**Parameters:**
- \(\gamma\): Infection coefficient
- \(k\): Interaction coefficient
- \(r\): Recovery rate
- \(\delta\): 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

\[
\begin{align*}
\frac{m}{dt^2} \frac{d^2 z}{dt^2} + \frac{c}{dt} \frac{dz}{dt} + k z &= f_0 \cos(\omega_f t) \\
z(0) &= z_0, \quad \frac{dz}{dt}(0) = z_1
\end{align*}
\]

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

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

\[ z(0) = z_0, \quad \frac{dz}{dt}(0) = z_1 \]

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

Solution: Reformulate problem as

\[ \frac{d^2 z}{dt^2} + \frac{C}{m} \frac{dz}{dt} + Kz = F_0 \cos(\omega_F t) \]

\[ z(0) = z_0, \quad \frac{dz}{dt}(0) = z_1 \]

where \( C = \frac{c}{m}, \quad K = \frac{k}{m} \) 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:**

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

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

\[
+ \frac{\chi(E)}{4\pi} \int_{4\pi} d\Omega' \int_0^\infty dE' \mathbf{v}(E') \Sigma_f(E') \varphi(r, E', \Omega', t)
\]

**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

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

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:**

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

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

**Sobol Indices:**

$$S_i = \frac{D_i}{D}$$

**Analogy:** Taylor or Fourier series

**Assumption:** Mutually independent parameters

**Statistical Interpretation:**

$$D_i = \text{var}[\mathbb{E}(Y|q_i)] \Rightarrow S_i = \frac{\text{var}[\mathbb{E}(Y|q_i)]}{\text{var}(Y)}$$
Morris Screening: Random Sampling of Approximated Derivatives

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

![Diagram](image)

**Elementary Effect:**

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

**Global Sensitivity Measures:** $r$ samples

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

$$\sigma^2_i = \frac{1}{r-1} \sum_{j=1}^{r} \left( d_i^j(q) - \mu_i \right)^2, \quad \mu_i = \frac{1}{r} \sum_{j=1}^{r} d_i^j(q)$$
SIR Disease Example

SIR Model:

\[
\frac{dS}{dt} = \delta N - \delta S - \gamma kI S, \quad S(0) = S_0 \quad \text{Susceptible}
\]

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

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

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

Assumed Parameter Distribution:

\[\begin{align*}
\gamma & \sim \mathcal{U}(0, 1), \\
k & \sim \text{Beta}(\alpha, \beta), \\
r & \sim \mathcal{U}(0, 1), \\
\delta & \sim \mathcal{U}(0, 1)
\end{align*}\]

Infection Coefficient \quad Interaction Coefficient \quad Recovery Rate \quad Birth/death Rate

Response:

\[y = \int_0^5 R(t, q) dt\]
SIR Disease Example

SIR Model:

\[
\begin{align*}
\frac{dS}{dt} &= \delta N - \delta S - \gamma kIS, \quad S(0) = S_0 \quad \text{Susceptible} \\
\frac{dI}{dt} &= \gamma kIS - (r + \delta)I, \quad I(0) = I_0 \quad \text{Infectious} \\
\frac{dR}{dt} &= rI - \delta R, \quad R(0) = R_0 \quad \text{Recovered}
\end{align*}
\]

Typical Realization:
SIR Disease Example

Global Sensitivity Measures:

<table>
<thead>
<tr>
<th></th>
<th>$\gamma$</th>
<th>$k$</th>
<th>$r$</th>
<th>$\delta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_i$</td>
<td>0.0997</td>
<td>0.0312</td>
<td>0.7901</td>
<td>0.1750</td>
</tr>
<tr>
<td>$S_{T_i}$</td>
<td>$-0.0637$</td>
<td>$-0.0541$</td>
<td>0.5634</td>
<td>0.2029</td>
</tr>
<tr>
<td>$\mu_i^* \left( \times 10^3 \right)$</td>
<td>0.2532</td>
<td>0.2812</td>
<td>2.0184</td>
<td>1.2328</td>
</tr>
<tr>
<td>$\sigma_i \left( \times 10^3 \right)$</td>
<td>0.9539</td>
<td>1.6245</td>
<td>6.6748</td>
<td>3.9886</td>
</tr>
</tbody>
</table>

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

Influential Parameters

Note: Can fix non-influential parameters $\gamma$, $k$
Parameter Selection: Nuclear Power Plant Design

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

<table>
<thead>
<tr>
<th>parameter</th>
<th>partial correlation</th>
<th>simple correlation</th>
<th>morris main</th>
<th>morris interaction</th>
<th>CPS variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>k_eta</td>
<td>0.07</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_gama</td>
<td>-0.03</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_sent</td>
<td>-0.03</td>
<td>-0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_sdent</td>
<td>-0.07</td>
<td>-0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_tmasv</td>
<td>-0.03</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_tmasl</td>
<td>0.11</td>
<td>0.00</td>
<td>6.48E-05</td>
<td>2.28E-05</td>
<td>medium</td>
</tr>
<tr>
<td>k_tmasg</td>
<td>-0.19</td>
<td>-0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_tmomv</td>
<td>-0.12</td>
<td>-0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_tmome</td>
<td>0.02</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_tmoml</td>
<td>0.02</td>
<td>-0.02</td>
<td>2.23E-04</td>
<td>1.30E-04</td>
<td>medium</td>
</tr>
<tr>
<td>k_xk</td>
<td>0.08</td>
<td>-0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_xkes</td>
<td>-0.05</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_xkge</td>
<td>-0.07</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_xkl</td>
<td>0.04</td>
<td>-0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_xkge</td>
<td>-0.03</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_xkvs</td>
<td>0.11</td>
<td>-0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_xkvw</td>
<td>-0.10</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_xkwlm</td>
<td>0.14</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_xkwe</td>
<td>-0.01</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_qvapl</td>
<td>-0.09</td>
<td>-0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_tngv</td>
<td>-0.03</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_tngl</td>
<td>-0.01</td>
<td>0.03</td>
<td>9.00E-06</td>
<td>9.49E-06</td>
<td>low</td>
</tr>
<tr>
<td>k_rodqq</td>
<td>0.02</td>
<td>-0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_qradd</td>
<td>-0.02</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_qradv</td>
<td>-0.01</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_qliht</td>
<td>-0.01</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_sphts</td>
<td>-0.05</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_cond</td>
<td>-0.04</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_xkvx</td>
<td>0.03</td>
<td>-0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_xkwl</td>
<td>1.00</td>
<td>0.88</td>
<td>1.80E-01</td>
<td>7.07E-03</td>
<td>high</td>
</tr>
<tr>
<td>k_cd</td>
<td>1.00</td>
<td>0.46</td>
<td>9.59E-02</td>
<td>7.88E-03</td>
<td>high</td>
</tr>
<tr>
<td>k_cdfb</td>
<td>-0.02</td>
<td>-0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_wkr</td>
<td>0.02</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5 Identified Active Inputs:

k_cd: Pressure loss coefficient of space in sub-channel

k_xkwvlx: 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_tngl: 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 4\textsuperscript{th} or 6\textsuperscript{th}-order Landau energy?

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

**Parameters:**

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

**Global Sensitivity Analysis:**

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

**Conclusion:**

\(\alpha_{111}\) 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:

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

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

Conclusion:

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

Example: Quantum-informed continuum model

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

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

Parameters:

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

Global Sensitivity Analysis:

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

Note: Must accommodate correlation
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:**

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

**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

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 depend 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):**

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

\[ S = \begin{bmatrix}
\sigma_1 \\
\vdots \\
\sigma_r \\
0
\end{bmatrix}, \quad \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \geq \varepsilon \]

**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.
Random Range Finding Algorithms: Linear Problems


1. Choose $\ell$ 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: $y_i = \sum_{k=1}^p q_k \sin(2\pi kt_i)$, $i = 1, \ldots, m$

\[
\begin{bmatrix}
y_1 \\
\vdots \\
y_m
\end{bmatrix} = \begin{bmatrix}
sin(2\pi t_1) & \cdots & \sin(2\pi pt_1) \\
\vdots & \ddots & \vdots \\
\sin(2\pi t_m) & \cdots & \sin(2\pi pt_m)
\end{bmatrix} \begin{bmatrix}
q_1 \\
\vdots \\
q_p
\end{bmatrix}
\]
Random Range Finding Algorithms: Linear Problems

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

![Graphs showing column entries of A and absolute difference in singular values](image)

**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

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

and

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

Construct outer product

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

Partition eigenvalues: \( C = W \Lambda W^T \)

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

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 \quad Active Subspace: Range of eigenvectors in \( W_1 \)

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

**Question**: How sensitive are results to distribution, which is typically not known?
Gradient-Based Active Subspace Construction

**Active Subspace**: Construction based on random sampling

1. Draw $M$ independent samples $\{q^i\}$ from $\rho$

2. Evaluate $\nabla q f_j = \nabla q f(q^i)$

3. Approximate outer product

   $$C \approx \tilde{C} = \frac{1}{M} \sum_{j=1}^{M} (\nabla q f_j)(\nabla q f_j)^T$$  
   
   **Monte Carlo Quadrature**

   Note: $\tilde{C} = GG^T$ 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 $W$

**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:**

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

**Global Sensitivity Measures:** $r$ samples

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

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

**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: \( \ell \) 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 \( \{ \tilde{v}^j_1, \ldots, \tilde{v}^j_h \} \) from surface of sphere
4. Use Lagrange multiplier to determine
   \[
   u^j_{\text{max}} = a_0^+ w^j + \sum_{i=1}^{h} a_i^+ v^j_i , \quad v^1_i = \tilde{v}^j_i
   \]
   that maximizes \( g(u) = f(q^0 + R^{-1}u) \).

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

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

\[
(z - q^0)^T S(z - q^0) = 1
\]
\[
S = R^T R
\]
Initialization Algorithm

1. Inputs: \( \ell \) 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 \( \{\tilde{v}_j^1, \ldots, \tilde{v}_h^j\} \) from surface of sphere
4. Use Lagrange multiplier to determine
   \[
   u_{\max}^j = a_0^+ w_j^i + \sum_{i=1}^h a_i^+ v_i^j, \quad v_i^1 = \tilde{v}_i^1
   \]
   that maximizes \( g(u) = f(q^0 + R^{-1}u) \).

---

\[ f(q) = q_1 + 3q_2, \ h = 1 \]
Initialization Algorithm

1. Inputs: $\ell$ 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

$$u_{\text{max}}^j = a_0^+ w^j + \sum_{i=1}^h a_i^+ v_i^j, \quad v_1^j = \tilde{v}_1^j$$

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

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

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

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

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

Ortho-complement of $u_{\text{max}}^j$
Example: Initialization Algorithm to Approximate Gradient

**Example:** Family of elliptic PDE’s

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

with the random field representations

\[a(q, s, \ell) = a_{\min} + e^{\overline{a}(s, \ell) + \sum_{i=1}^{p} q_i^\ell \gamma_i \Phi_i(s)}\]

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

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

**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

Note: Convergence within $h \cdot \ell$ iterations
**SCALE6.1: High-Dimensional Example**

**Setup:** Cross-section computations SCALE6.1

- Input Dimension: 7700
- Output $k_{\text{eff}}$

<table>
<thead>
<tr>
<th>Materials</th>
<th>Reactions</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^{234}\text{U}$</td>
<td>$\Sigma_t$</td>
</tr>
<tr>
<td>$^{235}\text{U}$</td>
<td>$\Sigma_e$</td>
</tr>
<tr>
<td>$^{236}\text{U}$</td>
<td>$\Sigma_f$</td>
</tr>
<tr>
<td>$^{238}\text{U}$</td>
<td>$\Sigma_c$</td>
</tr>
<tr>
<td>$^{1}\text{H}$</td>
<td>$\bar{\nu}$</td>
</tr>
<tr>
<td>$^{16}\text{O}$</td>
<td>$\chi$</td>
</tr>
<tr>
<td>$^6\text{C}$</td>
<td>$n \rightarrow n'$</td>
</tr>
</tbody>
</table>

**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:

<table>
<thead>
<tr>
<th>Method</th>
<th>Gap</th>
<th>PCA</th>
<th>Error Tolerance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$10^{-3}$</td>
</tr>
<tr>
<td>Gradient-Based</td>
<td>1</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Initialized AM</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

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 2\textsuperscript{nd} 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 \( \{z^j\} \sim \mathcal{N}(0, 1) \)
- 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

![Graphs showing q1 and q2 distributions](image)

**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

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

with the random field representations

\[\log(a(s, n)) = \sum_{i=1}^{p} q^n_i \gamma_i \phi_i(s)\]

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

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

**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