Uncertainty quantification using stochastic finite element method.

Oleg Roderick*
Mihai Anitescu
Paul Fischer
Mathematics and Computer Science Division
October 2008

*Portland State University
Department of Mathematics & Statistics
Topics of discussion

- Need for uncertainty analysis
- An example of a complex system with uncertain inputs
- Limited applicability of linear model, random sampling
- Applicability of stochastic finite element (SFEM) approach

- Improvement of SFEM performance through the use of goal-oriented polynomial basis, goal-oriented collocation procedure
Uncertainty calculations in high dimensional systems

- How to approximate the stochastic distribution of functions over very large uncertain spaces?

\[ 0 = F(x, p) : R^n \times R^q \rightarrow R^n \Rightarrow x = x(p) \]

\[ p = p(\omega) : D(\mu(\omega)) \Rightarrow E_\omega [G(x(p))] = ? \]

- \( F \) could be a coupled system of several differential equations, describing a complex engineering structure, \( p \) the physical parameters (\( p \sim 10, 100, 1000... \)).

- \( G \) could be an observation on performance of the model (maximal temperature), or a characteristic function which computes the probability for the observed quantity to be in the prescribed range (chance of overheating).
Uncertainty calculations in high dimensional systems

- Computing the correct average is strongly related to approximating the function $G$.

- Available solutions:
  - Random sampling (directly evaluate $G$ for a sample from parameter space)
    Global, but slow.
  - Sensitivity analysis (approximate $G$ by a linear function, using derivative information)
    Fast, but local, hard to adjust. What if the precision is insufficient?

- Can we create a method that efficiently uses the advantages of both approaches? We think so: using adapted stochastic finite element method fitted with derivative information.
Stochastic Finite Element Method, preview

- A stochastic finite element model is an explicit approximation
  \[ G \approx \hat{J}(\alpha) = \sum x_q \Psi_q \] in basis \( \{ \Psi(\alpha) \} \), dependent on direct (or derived) uncertain parameters \( \{\alpha\} \).

- Stochastic Finite Element Method (SFEM):
  - Choose a set of multi-variable orthogonal polynomials \( \Psi \). Use some subset \( \{\Psi_q\} \) to approximate the output function:
    \[ \hat{J} \approx \sum x_q \Psi_q \]
    
    \[ \Psi_q = P_0 + \sum P_i \alpha_i + \sum P_{ij} \alpha_i \alpha_j + \sum P_{ijk} \alpha_i \alpha_j \alpha_k + \ldots \]
  
  - The coefficients \( P \) in the definition of each polynomial are chosen to satisfy the orthogonality condition in some measure \( \pi \):
    \[ \int_{\Omega} \Psi_p \Psi_q d\pi = 0 \quad p \neq q \]
  
  - The coefficients \( x_q \) are found by solving the collocation equation
    \[ G(\alpha) = \sum x_q \Psi_q(\alpha) \]
There are two aspects of heat exchange in the reactor core: thermal hydraulics, and neutron interaction.

Basic element of thermo-hydraulic model is a cylindrical pin surrounded by flowing coolant. Reactor core contains a hexagonal assembly of pins.

Finite volumes description of temperature distribution includes:
- a partition of the core into horizontal layers of volume elements;
- a heat flux equilibrium equation producing temperature $T$ in each element;
- temperature dependencies of the material properties $R$ of each element.
**Distribution and transport of heat in the reactor core**

- The parameters/material properties $R$ of the model include heat capacity $c_p$, heat conductivity $K$ for the coolant and fuel; convective heat transfer coefficient $h$.

- A fixed-point iteration procedure $R:=R(T), T:=T(R)$ is used to couple the dependence of the temperature distribution on the material parameters, and the dependence of the material parameters on temperature.

- Uncertainty in the performance $J(T)$ of the nuclear reactor is attributed to the uncertainty in the values of parameters $R$.

- **Note:** the available temperature-dependencies are built as a best fit to experimental data. Statistical information about the uncertainty may be unavailable.

\[
c_p \approx 1.6582 - 8.470 \cdot 10^{-4} T + 4.4541 \cdot 10^{-7} T^2 - 2992.6 T^{-2}
\]
with uncertainty $\frac{\Delta c_p}{c_p}$ estimated at 0.1% at 300 K, 3% at 1000 K, 8% at 2000 K.
Choose a single output $J(T)$ to characterize the performance of the model. For example: maximal temperature of the core.

Evaluation of the model:
- For current values of thermodynamical parameters, compute thermal fluxes $F$ over all types of interfaces (pin-pin, pin-coolant, coolant-coolant, coolant-outflow). Temperature gradient is estimated by a finite difference, all fluxes are linear in $T$.
- For a given nuclear reaction source term $q''$, assemble the conservation law $0 = \sum_{\partial \Omega} F - \int_{\Omega} q'' dV$ into the form $\Lambda T = \sum_{\partial \Omega} q'' dV$.
- Repeat the iterations $R := R(T), T := T(R)$ until convergence of the output.
Distribution and transport of heat in the reactor core

\[ 0 = -\nabla \cdot K \nabla T - \rho c_p \vec{u} \cdot \nabla T + q'' \]

When discretized we obtain

\[ \Phi = h(T_J - T_{\text{surface}}) = c_p \frac{T_{\text{surface}} - T_I}{D/2} \]

\[ \int q'' dV = \sum_j 2 \frac{h c_p}{h D + 2 c_p} T_J - \sum_j 2 \frac{h c_p}{h D + 2 c_p} T_I + \sum K \frac{1}{H_{l\rightarrow l^+}} \int ndST_{l^-,l^+} + \ldots \]

\[ ... - \sum K \frac{1}{H_{l^+ \rightarrow l^-}} \int ndST_l \]

\[ 0 = \sum_j K \frac{1}{H_{l \rightarrow J}} \int ndST_J + \sum_{J^+J^-} K \frac{1}{H_{l \rightarrow J}} \int ndST_{J^+,J^-} - \sum_j K \frac{1}{H_{l \rightarrow J}} \int ndST_I + \ldots \]

\[ - \sum_{J^+J^-} K \frac{1}{H_{l \rightarrow J}} \int ndST_I + \frac{1}{2} \sum_j \rho c_p \int \bar{u} ndST_J + \frac{1}{2} \sum_{J^+J^-} \rho c_p \int \bar{u} ndST_{J^+,J^-} + \ldots \]

\[ + \frac{1}{2} \sum_j \rho c_p \int \bar{u} ndST_I + \frac{1}{2} \sum_{J^+J^-} \rho c_p \int \bar{u} ndST_{I^+,I^-} + \sum J_{l^+J^-} 2 \frac{h c_p}{h D + 2 c_p} T_{J^*} - \sum J_{l^+J^-} 2 \frac{h c_p}{h D + 2 c_p} T_I \]
Distribution and transport of heat in the reactor core

Fig. 1.1-11 Deviations of Values from Other Assessments from the Recommended Values for the Heat Capacity at Constant Pressure of Liquid Sodium
Approximating the output of the model

- Assume a temperature-dependent structure for the uncertainty:

  \[ R = \left( \sum_i r^{(i)} T^i \right) \cdot \left( 1 + \alpha^{(0)} C^{(0)}(T) + \alpha^{(1)} C^{(1)}(T) + \alpha^{(2)} C^{(2)}(T) \right) + \ldots \]

  \[ h = \left( \sum_i r^{(i)} T^i \right) \cdot \left( 1 + \alpha^{(0)} C^{(0)}(Pe(T)) + \alpha^{(1)} C^{(1)}(Pe(T)) + \alpha^{(2)} C^{(2)}(Pe(T)) \right) + \ldots \]

  in the Chebyshev polynomial basis

  \[ C^{(0)}(T) = 1 \quad C^{(1)}(T) = T + 1 \quad C^{(2)}(T) = 2T^2 - 1 \]

  \[ C^{(3)}(T) = 4T^3 - 3T^2 \]

- With no oscillations in uncertainty, use 2\textsuperscript{nd} order expansion, resulting in 3 uncertainty parameters per thermo-dynamical property.

- Note: other structures are less suitable for expressing the uncertainty conditions.

\[ R = \left( \sum_i r^{(i)} T^i \right) + \left( \alpha^{(0)} C^{(0)}(T) + \alpha^{(1)} C^{(1)}(T) + \alpha^{(2)} C^{(2)}(T) \right) \]

\[ R = \sum_i (r^{(i)} + \alpha^{(i)} T^i) \]
Approximating the output of the model

- Find the validity region for the uncertainty coefficients \( \{\alpha\} \) by random sampling. Start with a large uniform sample of values, reject the points that violate the uncertainty condition

\[
\frac{\Delta R}{R} \leq \xi \%
\]

- In the multiplicative uncertainty model,

\[
\frac{\Delta R}{R} = 1 + \alpha^{(0)}C^{(0)} + \alpha^{(1)}C^{(1)} + \alpha^{(2)}C^{(2)}
\]

\[
C = C(T), C = C(\text{Pe}(T))
\]

- Check the uncertainty condition at selected values of \( T \), or of \( \text{Pe} \) to obtain the uncertainty validity region.
Stochastic Finite Element Method

- Stochastic Finite Element Method (SFEM):
  - Choose a set of multi-variable orthogonal polynomials $\Psi$. Use some subset $\{\Psi_q\}$ to approximate the output function: $J \approx \hat{J} = \sum_q x_q \Psi_q$

- The coefficients $P$ in the definition of each polynomial are chosen to satisfy the orthogonality condition in some measure $\pi$:
  $$\Psi_q = P_0 + \sum P_1 \alpha_i + \sum P_{ij} \alpha_i \alpha_j + \sum P_{ijk} \alpha_i \alpha_j \alpha_k + ...$$

- For Gaussian probability measure, $\int \Psi_p \Psi_q d\pi = 0$ $p \neq q$ the basis is a set of Hermite polynomials:
  $$\Psi = (H_1, H_2, ...) = \prod_i H_i^{(k_i)}$$

- The coefficients $x_q$ are found by collocation.

\[ H^{(0)}(\alpha) = 1 \quad H^{(1)}(\alpha) = 2\alpha \quad H^{(2)}(\alpha) = 4\alpha^2 - 1 \]

\[ H^{(3)}(\alpha) = 8\alpha^3 - 12\alpha \quad H^{(4)}(\alpha) = 16\alpha^4 - 48\alpha^2 + 12 \]
SFEM: Derivative-based Regression

- Collocation procedure: evaluate the basis polynomials at the sample points in the parameter space, run full model to compute the outputs \( S \) at the sample points, assemble the collocation system \( \Psi x = S \):

\[
\begin{pmatrix}
\Psi(S_1)
\Psi(S_2)
\vdots
\Psi(S_m)
\end{pmatrix}
\begin{pmatrix}
\chi
\end{pmatrix}
=
\begin{pmatrix}
J(S_1)
J(S_2)
\vdots
J(S_m)
\end{pmatrix}
\]

- Issue: we would like to use high-order polynomials. The number of sample points required to assemble \( \Psi \) grows rapidly.

- Suggestions:
  - For each sample point, include derivative information.
  - Use an incomplete basis.
SFEM: Derivative-based Regression

- Collocation with derivative information:

\[
\begin{pmatrix}
\Psi(S_1) \\
\frac{\partial}{\partial \alpha} \Psi(S_1) \\
\Psi(S_2) \\
\frac{\partial}{\partial \alpha} \Psi(S_1) \\
\Psi(S_n) \\
\frac{\partial}{\partial \alpha} \Psi(S_n)
\end{pmatrix}
\begin{pmatrix}
J(S_1) \\
\frac{\partial}{\partial \alpha} J(S_1) \\
J(S_2) \\
\frac{\partial}{\partial \alpha} J(S_1) \\
J(S_n) \\
\frac{\partial}{\partial \alpha} J(S_n)
\end{pmatrix}
= x
\]

- Only right-hand side requires direct evaluations of the model.
- It is possible to evaluate the derivative efficiently, and get blocks of additional information for not even an extra model evaluation cost!
Approximating the output of the model

- SFEM setup choices:
  - “Full” basis vs. “truncated” basis.
  - “Tall” $\Psi$ with over-sampling vs. “square” matrix $\Psi$ with a minimal number of sample points

- Goal-oriented basis: polynomials of high degree are only included for “important” variables. Importance is defined as sensitivity of the output function to a particular parameter.

- Goal-oriented sample set: mostly an open question. Sample points may be chosen: in the directions of highest sensitivity of the output function; for the best condition of $\Psi$; for optimal approximation error; for the best condition of $\Lambda$. 
Approximating the output of the model

- For a moderate number of parameters (3-15), a good choice is “tall” matrix, “truncated” basis.

- Possible definitions of “importance” of a parameter $r^{(i)}$:
  - Derivative (at some “typical” temperature distribution): $\left| \frac{\partial J}{\partial r^{(i)}} \right|
  - Derivative adjusted by parameter variance: $\left| \frac{\partial J}{\partial r^{(i)}} \right| \sigma_i$

- We start with a full basis of order 3, separate the variables, by “importance”, into groups I, II and III of sizes $n_I > n_{II} >> n_{III}$. We allow polynomials that include variables from group III to have degree 3; allow the polynomials that include variables from group II have degree 2; only keep polynomials of degree 1 in the variables from group I.
Performance of SFEM model

- Size of the finite volume model: 7 pins, 20 horizontal layers.

- The output function is an estimation of maximal temperature in the centerline of the reactor core (evaluation may require additional calculations).

- The computational budget for the reduced SFEM basis is kept approximately constant. We use 100 exact outputs for validation.

- Linear approximation is defined as a tangent model obtained at a “typical” temperature distribution.

\[ \tilde{J} = J_0 + \sum_i \frac{\partial J}{\partial a_i} a_i \]

- We output: Range (lowest and highest observed outputs)
  Variance (variance of the observed outputs)
  Error variance (variance of the difference between surrogate and exact outputs)
### Performance of SFEM model

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Range</th>
<th>Variance</th>
<th>Error Variance</th>
<th># Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cp-coolant</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random sampling</td>
<td>1123.11</td>
<td>36.86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear model</td>
<td>1211.31</td>
<td>1.58</td>
<td>23.19</td>
<td></td>
</tr>
<tr>
<td>SFEM, full</td>
<td>1124.10</td>
<td>36.54</td>
<td>0.0013</td>
<td>12</td>
</tr>
<tr>
<td>SFEM, truncated</td>
<td>1124.45</td>
<td>39.03</td>
<td>0.0015</td>
<td>4</td>
</tr>
</tbody>
</table>
## Performance of SFEM model

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Range</th>
<th>Variance</th>
<th>Error variance</th>
<th># points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cp-coolant, K-fuel, K-coolant, Cp-fuel</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random sampling</td>
<td>1132.06</td>
<td>39.49</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1153.95</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear model</td>
<td>1130.92</td>
<td>1.57</td>
<td>25.28</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1135.28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SFEM, full</td>
<td>1133.71</td>
<td>40.12</td>
<td>0.0032</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>1155.75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SFEM, truncated</td>
<td>1135.81</td>
<td>45.21</td>
<td>0.0028</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>1155.79</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Performance of SFEM model

12 parameters: \( c_{p,\text{coolant}}, K_{\text{fuel}}, K_{\text{coolant}}, h \)

- Full model
- SFEM, full basis
- SFEM, truncated
- Linear model
Conclusion

- We have defined a SFEM method for high-order approximation of the response of a multiphysics system.
- The method uses derivatives to fit the SFEM polynomial, a first, to our knowledge. This results in an advantage in computational efficiency over a generic SFEM approach.
- For a simplified reactor core model this results in significant improvement in variance over the linear model.
- The method has several challenges: optimal sampling, improvement of the collocation condition number, basis truncation.
- We will extend the developed approach to larger models, incorporating additional physical effects.
Thank you for your attention.

Acknowledgement of information provided to us:

- Nuclear reactor analysis (Duderstadt, Hamilton)
- Thermo-hydraulics of a reactor core (lecture by Paul Fischer)
- Stochastic finite element methods (Ghanem, Spanos)
- Material properties (Fink)
Choose a single output $J(T)$ to characterize the performance of the model. For example: (maximal, average) temperature of coolant.

Evaluation of the model:
- For current values of thermodynamical parameters, compute thermal fluxes $F$ over all types of interfaces (pin-pin, pin-coolant, coolant-coolant, coolant-outflow). Temperature gradient is estimated by a finite difference, all fluxes are linear in $T$.
- For a given nuclear reaction source term $q''$, assemble the conservation law $0 = \sum_{\partial \Omega} F - \int_{\Omega} q'' dV$ into the form $\Lambda T = \sum_{\partial \Omega} q'' dV$.
- Repeat the iterations $R:=R(T)$, $T:=T(R)$ until convergence of the output.
Approximating the output of the model

- Assume a temperature-dependent structure for the uncertainty:

\[
R = \left( \sum_i r^{(i)} T^i \right) \cdot (1 + \alpha^{(0)} C^{(0)}(T) + \alpha^{(1)} C^{(1)}(T) + \alpha^{(2)} C^{(2)}(T)) + \ldots
\]

\[
h = \left( \sum_i r^{(i)} T^i \right) \cdot (1 + \alpha^{(0)} C^{(0)}(Pe(T)) + \alpha^{(1)} C^{(1)}(Pe(T)) + \alpha^{(2)} C^{(2)}(Pe(T))) + \ldots
\]

in the Chebyshev polynomial basis

\[
C^{(0)}(T) = 1 \quad C^{(1)}(T) = T + 1 \quad C^{(2)}(T) = 2T^2 - 1
\]

\[
C^{(3)}(T) = 4T^3 - 3T^2
\]

- With no oscillations in uncertainty, use 2\textsuperscript{nd} order expansion, resulting in 3 uncertainty parameters per thermo-dynamical property.

- Note: other structures are less suitable for expressing the uncertainty conditions.

\[
R = \left( \sum_i r^{(i)} T^i \right) + (\alpha_0^{(1)} C^{(0)}(T) + \alpha_1^{(1)} C^{(1)}(T) + \alpha_2^{(1)} C^{(2)}(T))
\]

\[
R = \sum_i (r^{(i)} + \alpha^{(i)} T)^i
\]
Approximating the output of the model

- Find the validity region for the uncertainty coefficients \( \{ \alpha \} \) by random sampling. Start with a large uniform sample of values, reject the points that violate the uncertainty condition

\[
\frac{\Delta R}{R} \leq \xi \%
\]

- In the multiplicative uncertainty model,

\[
\frac{\Delta R}{R} = 1 + \alpha^{(0)} C^{(0)} + \alpha^{(1)} C^{(1)} + \alpha^{(2)} C^{(2)}
\]

where

\[
C = C(T), \ C = C(Pe(T))
\]

- Check the uncertainty condition at selected values of \( T \), or of \( Pe \).
Approximating the output of the model

A surrogate model is an explicit approximation \( J \approx \tilde{J}(\alpha) = \sum x_q \psi_q \) in some basis \( \psi(\alpha) \).

Stochastic Finite Element Method (SFEM):
- Choose a set of multi-variable orthogonal polynomials \( \psi \). Use some subset \( \{\psi_q\} \) to approximate the output function: \( J \approx \tilde{J} = \sum x_q \psi_q \)

\[
\psi_q = P_0 + \sum P_i \alpha_i + \sum P_{ij} \alpha_i \alpha_j + \sum P_{ijk} \alpha_i \alpha_j \alpha_k + \cdots
\]

- The coefficients \( P \) in the definition of each polynomial are chosen to satisfy the orthogonality condition in some measure \( \pi \):

\[
\int_\Omega \psi_p \psi_q d\pi = 0 \quad p \neq q
\]

\[
\psi(\alpha_1, \alpha_2, \ldots) = \prod_i H^{(k_i)}(\alpha_i)
\]

- For Gaussian probability measure, the basis is a set of Hermite polynomials:

\[
H^{(0)}(\alpha) = 1 \quad H^{(1)}(\alpha) = 2\alpha \quad H^{(2)}(\alpha) = 4\alpha^2 - 1
\]

\[
H^{(3)}(\alpha) = 8\alpha^3 - 12\alpha \quad H^{(4)}(\alpha) = 16\alpha^4 - 48\alpha^2 + 12
\]

- The coefficients \( x_q \) are found by collocation.
Approximating the output of the model

- The dependencies \( J(T), T(\Lambda), (R,T), R(T,\alpha) \) can be studied directly, by random sampling.

- The derivative \( \nabla \alpha J \) can be used for sensitivity analysis.

- Derivative using the adjoint method:
  - Start with an algebraic form of the flux equilibrium equation:
    \[
    F(T,\alpha) = 0 \text{ with } F = \Lambda (\alpha) T - \sum q'' dV
    \]
  - Assemble a system for the adjoint variable \( \lambda \):
    \[
    D = \nabla_T F, \quad D^T \lambda = \nabla_T J
    \]
  - Evaluate the expression:
    \[
    \nabla \alpha J = -\lambda^T \nabla \alpha F
    \]
Approximating the output of the model

- Consider the finite volumes equation in the form \( F(T_n(\alpha), R(T_{n-1}, \alpha)) = 0 \)

\[
F(T_n, R(T_{n-1})) = \Lambda(R) \cdot T_n - B = 0
\]

\( T_n \approx T_{n-1} \)

- Differentiate to obtain

\[
\left( \frac{\partial F}{\partial T_n} + \frac{\partial F}{\partial R} \cdot \frac{\partial R}{\partial T_n} \right) \frac{dT_{n-1}}{d\alpha} + \frac{\partial F}{\partial R} \cdot \frac{\partial R}{\partial \alpha} = 0
\]

- We need two partial derivatives:

\[
\frac{\partial F}{\partial T_n} = \Lambda
\]

\[
\frac{\partial F}{\partial R} = \frac{\partial \Lambda}{\partial R} \cdot T_n
\]

- We have assembled the adjoint variable:

\[
\frac{dT_n}{d\alpha} = -\left( \frac{\partial F}{\partial T_n} + \frac{\partial F}{\partial R} \cdot \frac{\partial R}{\partial T_n} \right)^{-1} \frac{\partial F}{\partial R} \cdot \frac{\partial R}{\partial \alpha} = -\lambda^T \cdot \frac{\partial F}{\partial R} \cdot \frac{\partial R}{\partial \alpha}
\]
Approximating the output of the model

- The required components of the derivatives arrays

\[
\frac{\partial \Lambda}{\partial R^{(i,j)}}, \frac{\partial R^{(i,j)}}{\partial \alpha_k}, \frac{\partial R^{(i,j)}}{\partial T^{(l)}}
\]

for the volume cells \(I,J\) and parameter components \(R_j, \alpha_k\) are defined and stored during the last step of the iteration \(R_n := R(T_{n-1}), T_n := T(R_n)\).

- Finally, the derivative is expressed as:

\[
\frac{dJ}{d\alpha} = \frac{\partial J}{\partial T} \cdot \left( \Lambda + \frac{\partial \Lambda}{\partial R} \cdot T_n \cdot \frac{\partial R}{\partial T_n} \right)^{-1} \cdot \frac{\partial \Lambda}{\partial R} \cdot T_n \cdot \frac{\partial R}{\partial \alpha} \bigg|_{T_n = T, T_{n-1} = T_n}
\]

- Note: in Matlab, computing all derivatives for a single output typically produces an overhead of 10-40\%. 