Argonne NATIONAL LABORATORY

... for a brighter future

U.S. Department of Energy

THE UNIVERSITY OF CHICAGO

Office of Science U.S. DEPARTMENT OF ENERGY

A U.S. Department of Energy laboratory managed by The University of Chicago

Stochastic Finite Element Approaches in Nuclear Reactor Applications

Mihai Anitescu, Pino Palmiotti, Won-Sik Yang, (ANL)

Monika Neda (Pitt and UNLV)

MC & SNA Monterey

April 19, 2007

Uncertainty quantification (characterization?) of parametric equations

- Consider the parametric nonlinear equation

$$F(x,\omega) = 0, \quad x \in R^n,\ \omega \in \Omega \subset R^m,$$
$$F : R^n \times R^m \to R^n, \text{ a smooth function.}$$

- **Question:** How can I compute (or at least, efficiently approximate) the solution mapping $x(\omega)$ (and its distribution if ω has a PDF)?
- Sensitivity approach: Assume $F(x,\omega)$ is linear and ω Gaussian
- Monte Carlo answer:
  - Use parameter PDF and sample
    $\{\omega_i\},\ i = 1, 2, \ldots, M \Rightarrow \{\tilde{x}(\omega)_i = F^{-1}(0, \omega_i)\}$
  - Followed by “histogram” or other post processing.
- Difficulties: Monte Carlo is slow and sensitivity approach is restrictive to relatively small uncertainty.

Discussion of Objectives

- **Can we derive an uncertainty characterization method that is more general than sensitivity and quicker than Monte Carlo?**
- There are several promising avenues (including combinations of them)
  - Randomized Quasi Monte Carlo?
  - Higher order sensitivity
  - Stochastic Finite Element
- In this work we investigate the promise of Stochastic Finite Element
- We concentrate on the parametric aspect alone (that is, approximating the mapping $x(\omega)$ and not its distribution (i.e. we look at the “hard” part, since with a good model, there is no need for further interaction with model)
- We discuss issues appearing in evaluating uncertainty in NR with respect to crossections and other parameters.
- Contributions: **Theoretical fundamentals for SFEM Galerkin approach and uncovered potential of collocation approach.**

Stochastic Finite Element (SFEM) Notions and Notations

- The average operator

$$E_\omega[g(\omega)] = \int_\Omega W(\omega) g(\omega) d\omega. \text{ Here } W(\omega) \text{ may be PDF}$$

- The orthogonal polynomials

$$E_\omega \left[P_i(\omega) P_j(\omega) \right] = \delta_{ij}, \quad 0 \le i, j \le K$$

- Typical multidimensional polynomial families are created by tensor products of Legendre, Laguerre, Chebyshev …

$$\omega = (\omega_1, \omega_2, \ldots, \omega_m), \quad P_i(\omega) = L_{i_1}(\omega_1) L_{i_2}(\omega_2) \ldots L_{i_m}(\omega_m) \quad 0 \le i \le K$$

- The stochastic connection leads to the name of **polynomial chaos**
- Stochastic finite dimensional approximation space

$$\sum_{k=1}^{K} c_k P_k(\omega) = \tilde{x}(\omega) = \tilde{x}(\omega, \{c_k\}) \in \mathcal{V}_K$$

- **Note that PDF of outcome can be done now without using model!**

SFEM approximation procedures

- The model equation $F(x(\omega), \omega) = 0, \quad \omega \in \Omega$
- **The Galerkin (projection)** approach to approximate $x(\omega)$ enforces that the residual be orthogonal to the SFEM space $\mathcal{V}_K$

$$\left.\begin{array}{c} E_\omega \left[F(\tilde{x}(\omega, \{c_k\}) P_k(\omega)\right] = 0, \\ k = 1, 2, \ldots, K \end{array}\right\} \Rightarrow \{c_k\}_{k=1,2,\ldots,K}$$

- **The collocation approach** (which is a deterministic sampling, or response surface approach).
  - Choose sample points $\{\omega_i\} \;\; i = 1, 2, \ldots, K$ generally tensor product of “good” one-dimensional points (Chebyshev zeros, Gauss-Legendre nodes, etc..)
  - Compute the exact solution at these points $F(x(\omega_i), \omega_i) = 0$
  - Collocate the approximate solution at these point

$$\left.\begin{array}{c} \tilde{x}(\omega_i, \{c_k\}) = x(\omega_i), \\ k = 1, 2, \ldots, K \end{array}\right\} \Rightarrow \{c_k\}_{k=1,2,\ldots,K}$$

Comparison of the two approaches

- Galerkin:
  - Pros:
    - *Works even if the problem is discontinuous or nonsmooth.*
    - *Exponential convergence with increasing degree of polynomials if the problem is smooth.*
  - Cons:
    - *Needs to solve coupled nonlinear system of equations.*
    - *For complicated functions F, equations are hard to set up and need quadrature (since the averages cannot be exactly computed)*
- Collocation:
  - Pros
    - *The problems to be solved are decoupled*
    - *Easy to implement and setup.*
  - Cons
    - *Its convergence properties are not understood for nonsmooth cases or incomplete tensor subset.*

Galerkin approach for constrained optimization

$$(O) \qquad \min_{x \in R^n} f(x, \omega) \text{ subject to } g(x, \omega) = 0$$

- SFEM formulation

$$(SO(K) \qquad \min_{c_1,\ldots,c_K} E_\omega \left[f(\tilde{x}(\omega, \{c_k\}, \omega) \right]$$
$$\text{subject to} \quad E_\omega [g(\tilde{x}(\omega, \{c_k\}), \omega) P_k(\omega)] = 0, \qquad k = 1, 2, \ldots, K$$

- **Result: For sufficiently large K, under suitable assumptions, (SO(K)) has well posed constraint and the objective function is bounded below.**
- We can then solve the SFEM equation by optimization techniques.
- **Result: Under suitable assumptions, the solution converges to the solution of (O) as K increases.**

Optimization-Galerkin formulation of symmetric eigenvalue problems

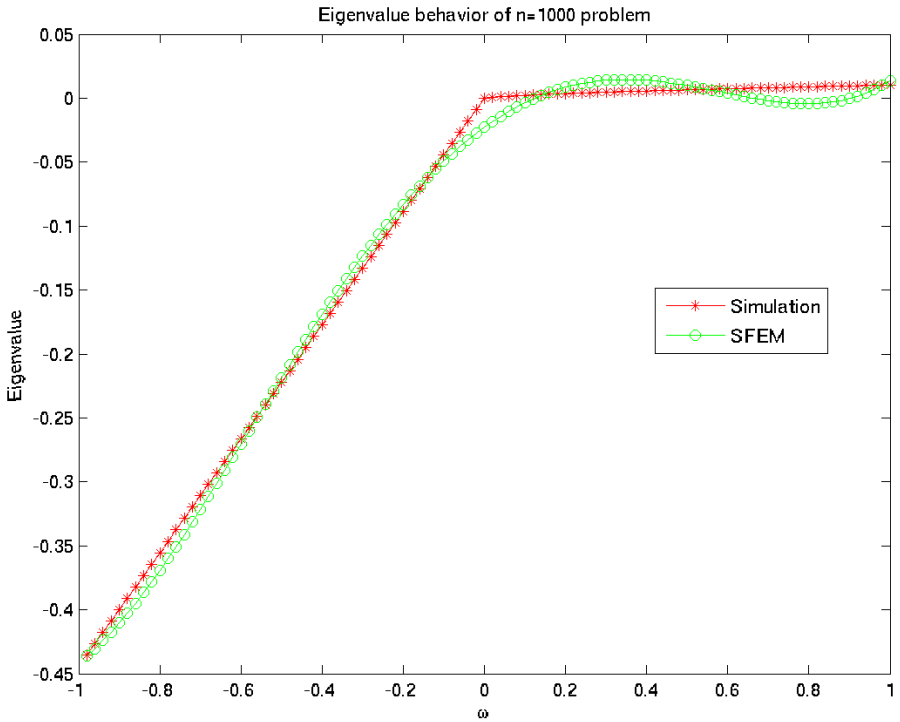
$$\left.\begin{array}{ll}\min & x^T Q(\omega) x \\ \text{s.t.} & x^T x = 1\end{array}\right\} \Rightarrow \left\{\begin{array}{ll}\min_{x_i, i=1,2,\ldots,K} & E_\omega\left[\left(\sum_{i=0}^{K} x_i P_i(\omega)\right)^T (Q+\omega D_Q)\left(\sum_{i=0}^{K} x_i P_i(\omega)\right)\right] \\ \text{s.t.} & E_\omega\left[\left(\sum_{i=0}^{K} x_i P_i(\omega)\right)^T \left(\sum_{i=0}^{K} x_i P_i(\omega)\right) P_k(\omega)\right] = E_\omega\left[P_k(\omega)\right] \\ & k = 1, 2, \ldots, K\end{array}\right.$$

- Note that this problem can be *exactly* represented, even when multiple parameters are involved, modulo a series of quadratures, (if not, Galerkin is hard to formulate and may need quadrature

$$\begin{array}{ll}\min & \sum_{i=0}^{K} x_i^T Q x_i + \sum_{i=0}^{K}\sum_{j=0}^{K} x_i^T D_Q x_i E_\omega\left[P_i(\omega)(\omega) P_j(\omega)\right] \\ \text{s.t.} & \sum_{i=0}^{K}\sum_{j=0}^{K} x_i^T x_i E_\omega\left[P_i(\omega) P_k(\omega) P_j(\omega)\right] = E_\omega\left[P_k(\omega)\right] \quad k = 0,1,2,\ldots,K\end{array}$$

Galerkin Numerical Example for n=1000

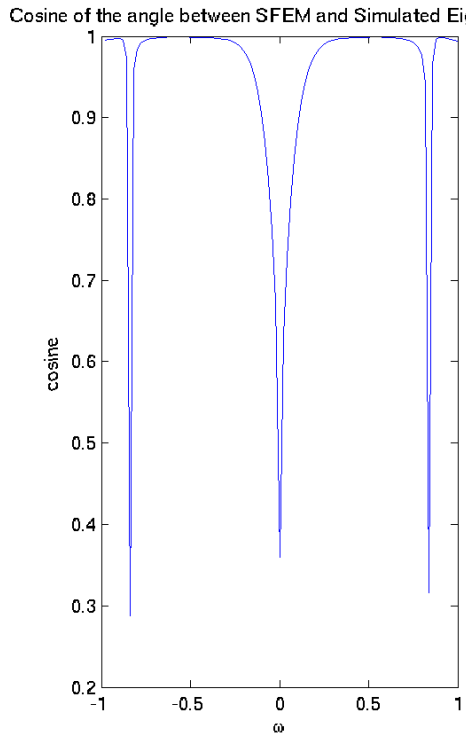
- Mimics the minimum eigenvalue problem for a one-dimensional diffusion equation, with uncertainty in absorption, akin to neutronics problems (where uncertainty in parameters also enters the matrix linearly)
- For K=4 The SCO formulation was solved in 7 seconds by KNITRO (one eig operation on same computer takes 13 seconds!)

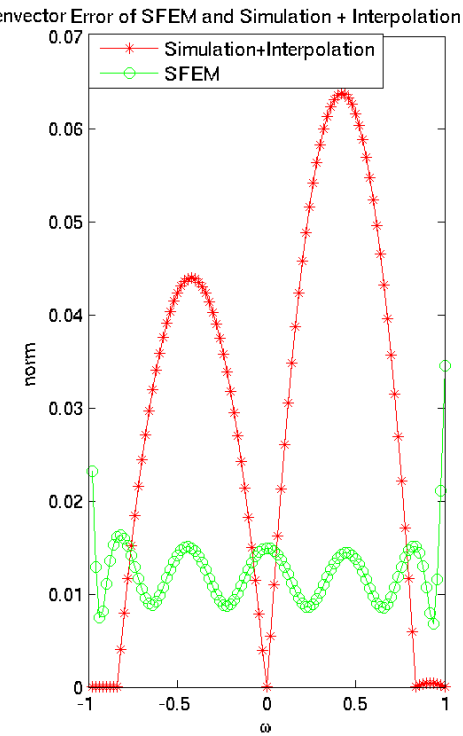


$$Q_{i,j} = \begin{cases} 1 & i = j = 1 \\ 1 & i = j = n \\ -1 & |i-j| = 1 \\ 2 & 1 < i = j < n \\ 0 & \text{otherwise} \end{cases}, \quad D_{Q\,i,j} = \begin{cases} 2\frac{i}{n}\frac{n-i}{n}\cos\left(\frac{i}{n}\right) & i = j \\ 0 & \text{otherwise} \end{cases}$$

Galerkin Numerical Example, eigenvector behavior

- At first sight, approximation is awful (left panel) but $\left\|\left(Q+\omega D_Q\right)\tilde{x}(\omega)-\tilde{\lambda}(\omega)\tilde{x}(\omega)\right\|$ has typical value of is 1%!
- The problem is the occasional degeneracy of the eigenvalue problem (though we have not encountered this case for keff calculations yet).
- If we sample eigenvector at 5 (well-chosen ☺) points, and interpolate linearly, we get an error that is typically larger by a factor of 3-4.
- **This shows that black-box approaches (collocation, Monte Carlo,..), even enhanced with AD, may be unadvisable.**
- **Sensitivity approaches would not be applicable for such eigenvector uncertainty problems!**





Galerkin formulation, generalized nonsymmetric eigenvalue problems—embedded optimization form

$$Q^1(\omega)x^1(\omega)+Q^2(\omega)x^2(\omega)=Q^3(\omega)x^3(\omega)$$

$$Q^5(\omega)x^2(\omega)=Q^4(\omega)x^1(\omega)$$

$$x^3(\omega)=\lambda(\omega)x^1(\omega)$$

$$x^1(\omega)^T\,x^1(\omega)=1$$

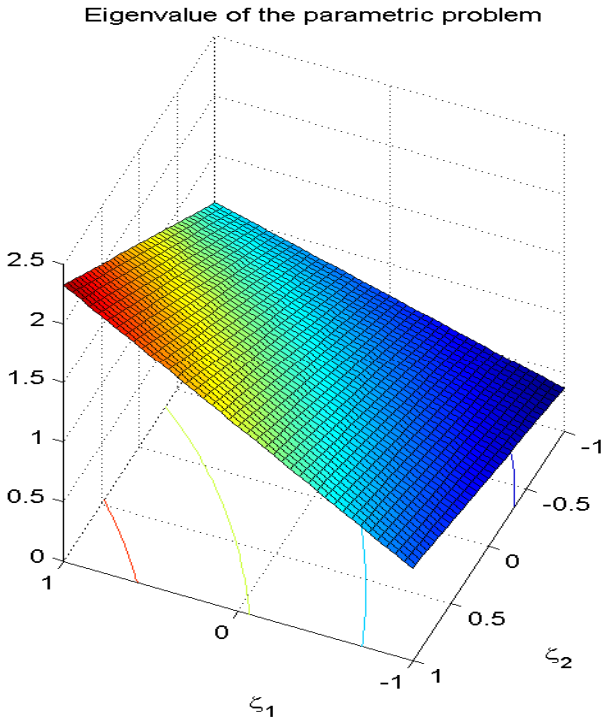
- Embed an optimization problem – allows us to use previous results and define a power-method-like iteration (**note that rescaling is far more complicated here!!—it is a projection**, details in paper)

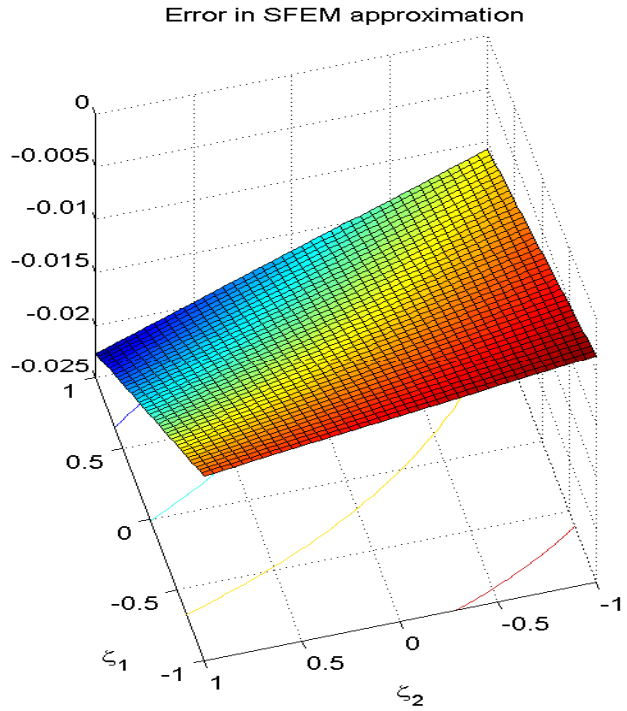
$$Q^1(\omega)x^1(\omega)+Q^2(\omega)x^2(\omega)=Q^3(\omega)x^3(\omega)$$

$$Q^5x^2(\omega)=Q^4(\omega)x^1(\omega)$$

$$x^1(\omega)=\arg\min_{z,z(\omega)^T z(\omega)=1}\left(x^3(\omega)-z(\omega)\right)^T\left(x^3(\omega)-z(\omega)\right)$$

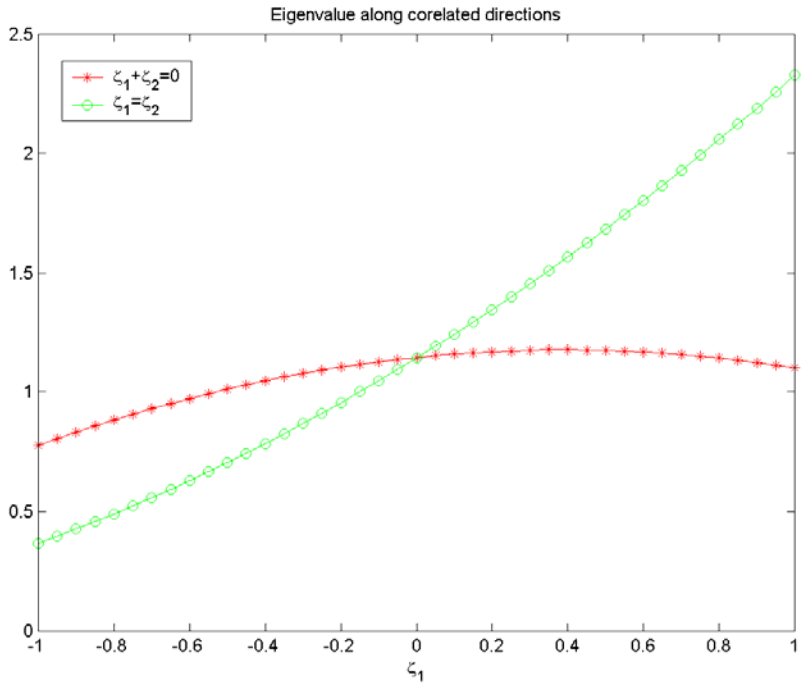
Results for keff for two-group 1D diffusion model

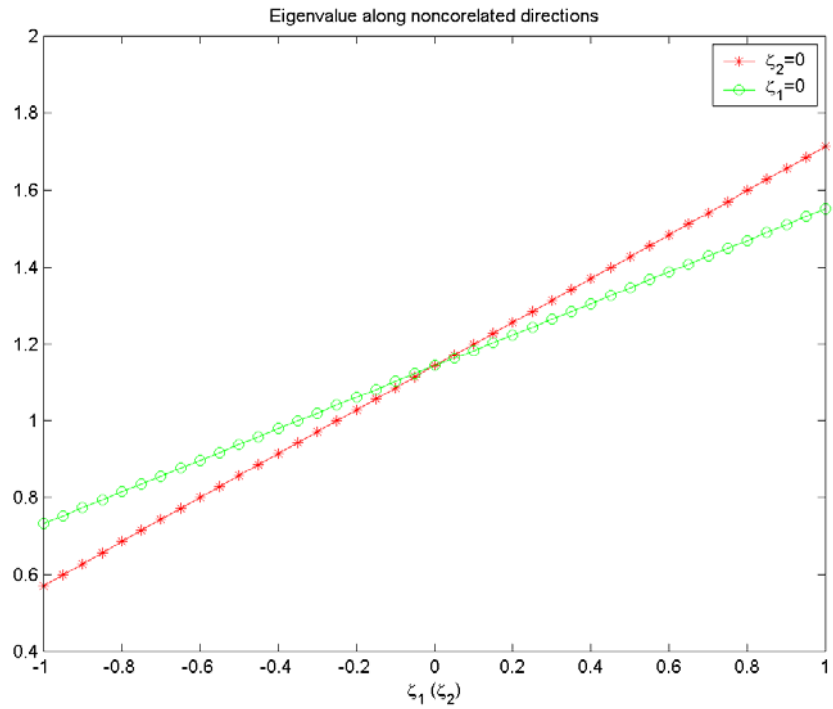




- Direct coupling only, correlated fission uncertainty and uncertainty in $\Sigma_{s,1\to 2}$
- Model takes 40s to compute; about the time to take 100 samples.
- If the goal is computation of distribution (for which 100 samples would be insufficient to compute distribution with such accuracy), method more competitive than Monte Carlo, but less spectacular than preceding example.

Further (interesting) observation:





- Nonlinearity is invisible along the coordinate directions, but it is very significant along the correlated directions $\zeta_1 = \zeta_2,\ \zeta_2 = -\zeta_1$
- This shows that first-order derivative approximation may miss substantial information.

Keff for 4 group diffusion model--Collocation Approach

Method	Number of Parameters= Nr. Calibration Samples	RMSE Error	Extra Calibration Samples/Equivalent Monte Carlo Sample Reduction
Sensitivity	12	0.3352	1/1
Collocation degree 1	12	0.1320	1/6.5
Collocation degree 2	78	0.0301	6.5/121

- 11 crossections, direct transfer only.
- The error is estimated by averaging it over 100 random samples.
- The upshot is **not** that sensitivity is not useful, but that, **given a certain budget of information, we may be better off from including info from more than one point** (including sensitivity for higher order approaches)

How did we choose the collocation points and polynomials?

- Tensor Legendre polynomials of homogeneous degree up to d

$$\tilde{x}(\omega) = \tilde{x}(\omega_1, \omega_2, \ldots, \omega_m) = \sum_{i=0}^{d} \sum_{\substack{i_1 + i_2 \ldots + i_m = i \\ i_j \geq 0}} x_{i_1 i_2 \ldots i_m} \prod_{j=1}^{m} L_{i_j}(\omega_j)$$

- The collocation points are the simplex Chebyshev tensor roots

$$\omega^{i_1 i_2 \ldots i_m} = \left(\omega^{i_1}, \omega^{i_2}, \ldots, \omega^{i_m}\right); C_d\left(\omega^j\right) = 0,\ j = 1, 2, \ldots, m$$

- For large dimension m, one must prune the polynomials (cannot choose either simplex or full tensor), since it is subject to curse of dimensionality … though people have been doing 50 variables in UQ in aerodynamics.

Conclusions and future work

- SFEM Galerkin models **are very robust** (though not inexpensive to compute), and work even in cases of nonsmooth problems where sampling-based or sensitivity methods would fail.
- **For smooth models, SFEM collocation models can be very competitive,** faster and more accurate than sensitivity approaches even for the same computational budget.
- To apply the work to 1000s of parameters, the set of parameters must be reduced by using correlation information.
- SFEM problems still have open questions
  - Well-posedness of constraints for less than small variation assumption for either Galerkin and collocation.
  - Choice of collocation points that results in well conditioning of the linear of the collocation for incomplete basis sets.