Numerical Methods for Partial Differential Equations with Random Data

Howard Elman
University of Maryland
Outline

I. Problem statement and discretization
   • Example: diffusion equation with random diffusion coefficient
   • Discretization by stochastic Galerkin method
   • Discretization by stochastic collocation method

II. Solution algorithms
   • Multigrid-style methods for various discretizations
   • Comparison of solution costs for different discretizations
I. Stochastic Differential Equations with Random Data

Example: diffusion equation

\[- \nabla \cdot (a \nabla u) = f \quad \text{in} \quad \mathcal{D} \subset \mathbb{R}^d \]
\[u = g_D \quad \text{on} \quad \partial \mathcal{D}_D, \quad (a \nabla u) \cdot n = 0 \quad \text{on} \quad \partial \mathcal{D}_N = \partial \mathcal{D} \setminus \partial \mathcal{D}_D \]

Uncertainty / randomness:
\[a = a(x, \omega) \quad \text{a random field} \]
For each fixed \(x\), \(a(x, \omega)\) a random variable

Other possibly uncertain quantities:
Forcing function \(f\)
Boundary data \(g_D\)
Viscosity \(\nu\) in Navier-Stokes equations
\[- \nu \nabla^2 u + (u \cdot \text{grad})u + \text{grad} \ p = f \]
\[-\text{div} \ u = 0 \]
Depictions: Random Data on Unit Square
Diffusion Equation with Random Diffusion Coefficient

\[ -\nabla \cdot (a \nabla u) = f \quad \text{in } D \]

Assumptions:

1. Spatial correlation of random field: For \( x, y \in D \):

   Random field \( a(x, \omega) \)

   Mean \( \mu(x) = E(a(x, \cdot)) \)

   Variance \( \sigma(x) = E(a(x, \cdot)^2) - \mu^2 \)

   Covariance function
   \[ c(x, y) = E( (a(x, \cdot) - \mu(x))(a(y, \cdot) - \mu(y))) \]
   is finite

   \text{vs. white noise, where } c \text{ is a } \delta \text{-function}

2. Coercivity \( 0 < \alpha_1 \leq a \leq \alpha_2 < \infty \)

   \[ \Rightarrow \text{Problem is well-posed} \]
Monte-Carlo Simulation

Sample $a(x, \omega)$ at all $x \in \mathcal{D}$, solve in usual way

Standard weak formulation: find $u \in H^1_E(\mathcal{D})$ such that

$$a(u, v) = \ell(v)$$

for all $v \in H^1_{E_0}(\mathcal{D})$,

$$a(u, v) = \int_{\mathcal{D}} a \nabla u \cdot \nabla v \, dx, \quad \ell(v) = \int_{\mathcal{D}} f \, v \, dx$$

Multiple realizations (samples) of $a(x, \cdot)$
Multiple realizations of $u$
Statistical properties of $u$

Problem: convergence is slow, requires many solves
Another Point of View

\[-\nabla \cdot (a \nabla u) = f \text{ in } \mathcal{D}\]

Covariance function is finite \(\implies\) random field (diffusion coefficient) has \textit{Karhunen-Loève} expansion:

\[a(x, \omega) = a_0(x) + \sigma \sum_{r=1}^{\infty} \sqrt{\lambda_r} \ a_r(x) \xi_r(\omega)\]

\[a_0(x) = \mu(x) = E(a(x, \cdot)) \text{ mean}\]

\[a_r(x), \lambda_r = \text{eigenfunctions/eigenvalues of covariance operator}\]

\[(Ca)(x) = \lambda a(x), \quad (Ca)(x) = \int_D c(x, y)a(y)dy\]

\[\xi_r(\omega) = \text{identically distributed uncorrelated random variables with mean 0 and variance 1}\]
Finite Noise Assumption

\[- \nabla \cdot (a \nabla u) = f \quad \text{in } D\]

Truncated *Karhunen-Loève* expansion:

\[a(x, \omega) = a_0(x) + \sigma \sum_{r=1}^{m} \sqrt{\lambda_r} \, a_r(x) \xi_r(\omega)\]

~ Principal components analysis

Requires: \( m \) large enough so that the fluctuation of \( a \)

is well-represented, i.e. \( \lambda_{m+1} / \lambda_1 \) is small

More precisely: error from truncation is

\[\frac{|D| \sigma^2 - \sum_{j=1}^{m} \lambda_j}{|D| \sigma^2}\]

Choose \( m \) to make this small
Various Ways to Use This

\[ a(x, \omega) = a_0(x) + \sigma \sum_{r=1}^{m} \sqrt{\lambda_r} \ a_r(x) \xi_r(\omega) \]

1. **Stochastic Finite Element (Galerkin) Method:**
   Introduce a weak formulation analogous to finite elements in space that handles the “stochastic” component of the problem

2. **Stochastic Collocation Method:**
   Devise a special strategy for sampling \( \xi \) that converges more quickly than Monte Carlo simulation; derived from interpolation

Ghanem, Spanos, Babuška, Deb, Oden, Matthies, Keese, Karniadakis, Xiu, Hesthaven, Tempone, Nobile, Webster, Schwab, Todor, Ernst, Powell, Furnival, E., Ullmann, Rosseel, Vandewalle
Stochastic Finite Element (Stochastic Galerkin) Method

Probability space \((\Omega, \mathcal{F}, P)\)

\[ L^2_P(\Omega) \equiv \{ \text{square integrable functions wrt } dP(\omega) \} \]

Inner product on \(L^2_P(\Omega)\):

\[ \langle v, w \rangle = E(vw) = \int_{\Omega} v(\omega)w(\omega) dP(\omega) \]

Use to concoct weak formulation on product space \(H^1_E(D) \otimes L^2_P(\Omega)\)

Find \(u \in H^1_E(D) \otimes L^2_P(\Omega)\) such that

\[ \langle a(u,v) \rangle = \langle \ell(v) \rangle \quad \text{for all } \quad v \in H^1_{E_0}(D) \otimes L^2_P(\Omega) \]

\[ \int_{\Omega} \int_{D} a \nabla u \cdot \nabla v \, dx \, dP(\omega) \]

Solution \(u= u(x, \omega)\) is itself a random field
For Computation: Return to Finite Noise Assumption

Truncated Karhunen-Loève expansion

\[ a(x, \xi(\omega)) = a_0(x) + \sigma \sum_{r=1}^{m} \sqrt{\lambda_r} a_r(x) \xi_r(\omega) \]

Stochastic weak formulation uses

\[ \langle a(u, v) \rangle = \int_{\Omega} \int_{D} a \nabla u \cdot \nabla v \, dx \, dP(\omega) = \int_{\xi(\Omega)} \int_{D} a(x, \xi) \nabla u \cdot \nabla v \, dx \, \rho(\xi) \, d\xi \]

Bilinear form entails integral over image of random variables \( \xi \)

Require joint density function associated with \( \xi \)

\( \xi \) plays the role of a Cartesian coordinate
Statement of Problem Becomes

Find \( u \in H^1_E(\mathcal{D}) \otimes L^2_P(\Gamma) \) such that

\[
\int_\Gamma \int_{\mathcal{D}} a(x, \xi) \nabla u \cdot \nabla v \, dx \, \rho(\xi) \, d\xi = \int_\Gamma \int_{\mathcal{D}} f v \, dx \, \rho(\xi) \, d\xi
\]

for all \( v \in H^1_{E_0}(\mathcal{D}) \otimes L^2_P(\Gamma) \) \( (\Gamma = \xi(\Omega)) \)

Like an ordinary Galerkin (or Petrov-Galerkin) problem on a \((d+m)\)-dimensional “continuous” space

\[
d = \text{dimension of spatial domain} \\
\]

\[
m = \text{dimension of stochastic space}
\]
Discretization

\[ \int_{\Gamma} \int_{\mathcal{D}} a(x, \xi) \nabla u \cdot \nabla v \, dx \rho(\xi) d\xi = \int_{\Gamma} \int_{\mathcal{D}} f \, v \, dx \rho(\xi) d\xi \]

Finite dimensional spaces:

- **spatial discretization**: \( S_h \subset H^1_0(\mathcal{D}) \), spanned by \( \{ \phi_j \}_{j=1}^{N_x} \)
  for example: piecewise linear on triangles

- **stochastic discretization**: \( T_p \subset L^2(\Gamma) \), spanned by \( \{ \psi_l \}_{l=1}^{N_\xi} \)
  for example: polynomial chaos = \( m \)-variate Hermite polynomials (orthogonal wrt Gaussian measure)

Discrete weak formulation:

\[ a(u_{hp}, \nu_{hp}) = \ell(\nu_{hp}) \quad \text{for all } \nu_{hp} \in S_h \otimes T_p \]

\[ u_{hp} = \sum_{j=1}^{N_x} \sum_{l=1}^{N_\xi} u_{jl} \phi_j(x) \psi_l(\xi) \]
Basis Functions for Stochastic Space

Underlying space: \[ L^2(\Gamma) = \left\{ v(\xi) \right\} \bigg| \int_{\Gamma} v(\xi)^2 \rho(\xi) d\xi < \infty \]

\[ \rho(\xi) = \rho_1(\xi_1)\rho_2(\xi_2) \cdots \rho_M(\xi_M) \]

Let \( q_j^{(k)}(\xi_k) \) = polynomial of degree \( j \) orthogonal wrt \( \rho_k \)

Examples: if \( \rho_k \sim \text{Gaussian measure} \rightarrow \text{Hermite polynomials} \)

\( \rho_k \sim \text{uniform distribution} \rightarrow \text{Legendre polynomials} \)

Any \( \rho_k \) can be handled computationally (Gautschi)

\( \rightarrow \text{Rys polynomials} \)

\( T_p \subset L^2(\Gamma) \) spanned by \( \{ q_{j_1}^{(1)}(\xi_1)q_{j_2}^{(2)}(\xi_2) \cdots q_{j_m}^{(m)}(\xi_m) \} \)

Orthogonality of basis functions \( \rightarrow \) sparsity of coefficient matrix
Matrix Equation \( Au = f \)

\[
a(x, \xi(\omega)) = a_0(x) + \sigma \sum_{r=1}^{m} \sqrt{\lambda_r} a_r(x) \xi_r(\omega)
\]

\[
A = G_0 \otimes A_0 + \sum_{r=1}^{m} G_r \otimes A_r
\]

\[
[A_0]_{jk} = \int_D a_0(x) \nabla \varphi_j(x) \cdot \nabla \varphi_k(x) \, dx
\]

\[
[A_r]_{jk} = \sqrt{\lambda_r} \int_D \sigma(x) a_r(x) \nabla \varphi_j(x) \cdot \nabla \varphi_k(x) \, dx
\]

\[
[G_0]_{lq} = \langle \psi_l, \psi_q \rangle, \quad [G_r]_{lq} = \langle \xi_r \psi_l, \psi_q \rangle
\]

\[
[f]_{kq} = \iint_{\Gamma D} f(x, \xi) \varphi_k(x) \psi_q(\xi) \, dx \rho(\xi) d\xi
\]

Properties of \( A \):

- order = \( N_x \times N_\xi = \) (size of spatial basis) \( \times \) (size of stochastic basis)
- sparsity: inherited from that of \( \{G_r\} \) and \( \{A_r\} \)
Dimensions of Discrete Stochastic Space

\[ T_p \subset L^2(\Gamma) \text{ spanned by } \{ q^{(1)}_{j_1}(\xi_1)q^{(2)}_{j_2}(\xi_2)\cdots q^{(m)}_{j_m}(\xi_m) \} \]

Full tensor product basis: \( 0 \leq j_i \leq p, \quad i = 1, \ldots, m \)
Dimension: \( (p+1)^m \) Too large

“Complete” polynomial basis: \( j_1 + j_2 + \cdots + j_m \leq p \)
Dimension: \( \binom{m+p}{p} = \frac{(m+p)!}{m! \ p!} \) More manageable

Order these in a systematic way →

\( \psi_1(\xi), \psi_2(\xi), \ldots, \psi_{N_\xi}(\xi) \)
Example

\( T_p \subseteq L^2(\Gamma) \) spanned by \( \{ q_{j_1}^{(1)}(\xi_1)q_{j_2}^{(2)}(\xi_2)\cdots q_{j_m}^{(m)}(\xi_m) \} \)

“Complete” polynomial basis: \( j_1 + j_2 + \cdots + j_m \leq p \)

\[ m=2, \ p=3 \quad \rightarrow \quad \binom{m+p}{p} = \binom{5}{2} = 10 \]

Orthogonal (Hermite) polynomials in 1D:

\( H_0(\xi) = 1, \ H_1(\xi) = \xi, \ H_2(\xi) = \xi^2 - 1, \ H_3(\xi) = \xi^3 - 3\xi \)

Gives basis set:

\[
\begin{align*}
\psi_1(\xi) &= 1 \\
\psi_2(\xi) &= \xi_1 \\
\psi_3(\xi) &= \xi_1^2 - 1 \\
\psi_4(\xi) &= \xi_1^3 - 3\xi_1 \\
\psi_5(\xi) &= \xi_2 \\
\psi_6(\xi) &= \xi_1\xi_2 \\
\psi_7(\xi) &= (\xi_1^2 - 1)\xi_2 \\
\psi_8(\xi) &= (\xi_2^2 - 1) \\
\psi_9(\xi) &= (\xi_2^2 - 1)\xi_1 \\
\psi_{10}(\xi) &= \xi_2^3 - 3\xi_2 
\end{align*}
\]
Example of Sparsity Pattern

For $m$-variate polynomials of total degree $p$:

$$N = \frac{(m+p)!}{m!p!} = \frac{10!}{6!4!} = 210$$
Uses of the Computed Solution:

\[ u_{hp} = \sum_{l=1}^{N_\xi} \sum_{j=1}^{N_x} u_{jl} \phi_j(x) \psi_l(\xi) = \sum_{l=1}^{N_\xi} u_l(x) \psi_l(\xi) \]

1. **Moments:** First moment of \( u \) (expected value):

\[
E(u_{hp}) = \sum_{l=1}^{M} u_l(x) \int_{\Gamma} \psi_l(\xi) \rho(\xi) \, d\xi
\]

\[
= u_1(x) = \sum_{j=1}^{N} u_{j1} \phi_j(x)
\]

using orthogonality of stochastic basis functions

Similarly for second moment / covariance
Uses of the Computed Solution:

\[ u_{hp} = \sum_{l=1}^{N_\xi} \sum_{j=1}^{N_x} u_{jl} \phi_j(x) \psi_l(\xi) = \sum_{l=1}^{N_\xi} u_l(x) \psi_l(\xi) \]

2. Cumulative distribution functions

E.g.: \( P(u_{hp}(x, \xi) > \alpha) \) at some point \( x \)

Sample \( \xi \)
Evaluate \( u_{hp}(x, \xi) = \sum_{l=1}^{N_\xi} u_l(x) \psi_l(\xi) \)
Repeat

Not free, but **no solves required**
Stochastic Collocation Method

Given \( a(x, \xi) = a_0(x) + \sigma \sum_{r=1}^{m} \sqrt{\lambda_r} a_r(x) \xi_r \) as above

Let \( \xi \) be a specified realization (\( \sim \) Monte Carlo)

Weak formulation:

\[
\int_{\mathcal{D}} (a_0(x) + \sigma \sum_{r=1}^{m} \sqrt{\lambda_r} a_r(x) \xi_r) \nabla u \cdot \nabla v \, dx = \int_{\mathcal{D}} f \, v \, dx
\]

Discretize in space in usual way.

Stochastic collocation: choose special set \( \xi^{(1)}, \xi^{(2)}, \ldots, \xi^{(N_\xi)} \) from considerations of interpolation

Advantage: Spatial systems are decoupled
Multi-Dimensional Interpolation

Given \( \xi^{(1)}, \xi^{(2)}, \ldots, \xi^{(N_\xi)} \), and \( v(\xi) \), consider an interpolant

\[
(Iv)(\xi) \equiv \sum_{k=1}^{N_\xi} v(\xi^{(k)}) L_k(\xi) \approx v(\xi),
\]

where \( L_k(\xi^{(j)}) = \delta_{jk} \), Lagrange interpolating polynomial

If \( u_h^{(k)} \) solves the discrete (in space) version of

\[
\int_{\mathcal{D}} \left( a_0(x) + \sigma \sum_{r=1}^{m} \sqrt{\lambda_r} a_r(x) \xi_r \right) \nabla u \cdot \nabla v \, dx = \int_{\mathcal{D}} f \, v \, dx
\]

with \( \xi = \xi^{(k)} \), then the \textit{collocated} solution is

\[
u_{hp}(x, \xi) = \sum_{k=1}^{N_\xi} u_h^{(k)}(x) L_k(\xi)
\]
To Compute Statistical Quantities

Solution

\[ u_{hp}(x, \xi) = \sum_{k=1}^{N_{\xi}} u_h^{(k)}(x)L_k(\xi) \]

1. Moments

\[ E(u_{hp})(x) = \sum_{k=1}^{N_{\xi}} u_h^{(k)}(x) \int_{\Gamma} L_k(\xi) \rho(\xi) d\xi \]

Not free but can be precomputed

2. Distribution functions

Obtained by sampling, cheap
Strategy for Interpolation

\[(Iv)(\xi) \equiv \sum_{k=1}^{N_\xi} v(\xi^{(k)}) L_k(\xi) \approx v(\xi),\]

One choice of \(\{L_k\} : \quad L_k(\xi) = \ell_{k_1}(\xi_1) \ell_{k_2}(\xi_2) \cdots \ell_{k_m}(\xi_m)\)

\[\ell_{k_j} = \text{1D interpolating polynomial}\]

0 \leq k_j \leq p

Advantage: easy to construct

Disadvantage: “curse of dimensionality,”

dimension = \((p+1)^m\)
Detour: Sparse Grids

Given: 1D interpolation rule \( (U^{(k)}v)(y^{(k)}) = \sum_{j=1}^{m_k} v(y_j^{(k)}) \ell_j(y^{(k)}) \)

Derived from (1D) grid \( Y^{(k)} = \{ y_1^{(k)}, \ldots, y_{m_k}^{(k)} \} \)

Multidimensional rule above is induced by \textit{fully populated} multidimensional grid \( Y^{(1)} \times Y^{(2)} \times \cdots \times Y^{(m)} \).

\[ |Y^{(k)}| = m_k = p + 1 \]

Alternative: multidimensional \textit{sparse grid} (Smolyak)

\[ \mathcal{H}(m + p, m) \equiv \bigcup_{p-m+1 \leq i_1 + \cdots + i_m \leq p} (Y^{(i_1)} \times Y^{(i_2)} \times \cdots \times Y^{(i_m)}) \]
Sparse Grid Interpolation

Example of sparse grid for $m=3$, $p=16$

For $v$ of the form $v(\xi) = v_1(\xi_1)v_2(\xi_2)\cdots v_m(\xi_m)$, interpolating function takes the form

$$(Iv)(\xi) = \sum_{i_1+\cdots+i_m\leq p} (U^{(i_1)} - U^{(i_1-1)})v_1(\xi_1) \otimes (U^{(i_2)} - U^{(i_2-1)})v_2(\xi_2) \otimes \cdots \otimes (U^{(i_m)} - U^{(i_m-1)})v_m(\xi_m)$$
Sparse Grid Interpolation

**Theorem** (Novak, Ritter, Wasilkowski, Wozniakowski)

For $\xi \in$ sparse grid and $\nu(\xi)$ a tensor product polynomial of total degree at most $p$,

$$\nu(\xi) = q_{j_1}^{(1)}(\xi_1)q_{j_2}^{(2)}(\xi_2)\cdots q_{j_m}^{(m)}(\xi_m), \quad j_1 + j_2 + \cdots + j_m \leq p$$

$$(I\nu)(\xi) = \nu(\xi).$$

That is: sparse grid interpolation evaluates the set of complete $m$-variate polynomials exactly.

**Overhead:** number of sparse grid points to achieve this ($= \#$ stochastic dof) is larger than for Galerkin

$$\approx 2^p \binom{m+p}{p} \text{ vs. } \binom{m+p}{p}$$
Monte-Carlo: \[ E(u) - E_s(u_h) = (E(u) - E(u_h)) + (E(u_h) - E_s(u_h)) \leq c_1 h E(\|u\|_2) \sim 1/\sqrt{s} \]

Convergence is slow wrt number of samples but independent of number of random variables \( m \)

Stochastic Galerkin and Collocation:

\[ E(u) - E(u_{hp}) = (E(u) - E(u_h)) + (E(u_h) - E(u_{hp})) \leq c_1 h E(\|u\|_2) \leq c_2 r^p, \ r < 1 \]

Exponential in polynomial degree \( p \)

Constants \( (c_2, r) \) depend on \( m \)

Rule of thumb: the same \( p \) gives the same error

(for all versions of SG and collocation)

More dof for collocation than SG
Recapitulating

Monte-Carlo methods:
Many samples needed for statistical quantities
Many systems to solve
Systems are independent
Statistical quantities are free (once data is accumulated)

With $s$ realizations: \[ E_s(u_h) = \frac{1}{s} \sum_{r=1}^{s} u_h^{(r)}(x) \]

Convergence is slow but independent of $m$

Stochastic Galerkin methods:
One large system to solve
Statistical quantities are free or (relatively) cheap

Stochastic collocation methods:
Systems are independent
Fewer systems than Monte Carlo
More degrees of freedom than Galerkin
Statistical quantities are (relatively) cheap

Similar convergence behavior
Faster than MC
Depends on $m$
II. Computing with the Stochastic Galerkin and Collocation Methods

For both: compute a discrete solution, a random field $u_{hp}(x, \xi)$

Stochastic Galerkin:

$$u_{hp}(x, \xi) = \sum_{l=1}^{N\xi} \sum_{j=1}^{N_x} u_{jl} \varphi_j(x) \psi_l(\xi) = \sum_{l=1}^{N\xi} u_l(x) \psi_l(\xi)$$

Stochastic Collocation:

$$u_{hp}(x, \xi) = \sum_{l=1}^{N\xi} \sum_{j=1}^{N_x} u_{jl} \varphi_j(x) L_l(\xi) = \sum_{l=1}^{N\xi} u_l(x) L_l(\xi)$$

Postprocess to get statistics
Computational Issues

**Stochastic Galerkin:** Solve one large system of order $N_x \times N_\xi$

$$N_\xi = \binom{m+p}{p}$$

Frequently cited as a problem for this methodology

**Stochastic Collocation:** Solve $N_\xi$ “ordinary” algebraic systems (of order $N_x$), one for each sparse grid point

Here: $N_\xi^{(collocation)} \sim 2^p N_\xi^{(Galerkin)}$

Some savings possible
Multigrid Solution of Matrix Equation I  

(E. & Furnival)

Solving $Au=f$

\[
A = G_0 \otimes A_0 + \sum_{r=1}^{m} G_r \otimes A_r
\]

\[
[A_r]_{jk} = \sqrt{\lambda_r} \sigma \int_{D} a_r(x) \nabla \varphi_j(x) \cdot \nabla \varphi_k(x) \, dx,
\]

\[
[G_r]_{lq} = \int_{\Omega} \psi_l(\xi) \psi_q(\xi) \xi_r \rho(\xi) \, d\xi
\]

$A_r = A_r^{(h)}$,  \hspace{1cm} $A = A^{(h)}$,  \hspace{1cm} spatial discretization parameter $h$

$A_r = A_r^{(2h)}$,  \hspace{1cm} $A = A^{(2h)}$,  \hspace{1cm} spatial discretization parameter $2h$

Develop MG algorithm for spatial component of the problem
Multigrid Algorithm (Two-grid)

Let $A^{(h)} = Q - N$, $Q =$ smoothing operator

for $i=0,1,...$
  for $j=1:k$  \hspace{1cm} k smoothing steps
    $u^{(h)} \leftarrow (I - Q^{-1}A^{(h)})u^{(h)} + Q^{-1}f^{(h)}$
  end

$r^{(2h)} = R(f^{(h)} - A^{(h)}u^{(h)})$  \hspace{1cm} Restriction

Solve $A^{(2h)}c^{(2h)} = r^{(2h)}$  \hspace{1cm} Coarse grid correction

$u^{(h)} \leftarrow u^{(h)} + Pc^{(2h)}$  \hspace{1cm} Prolongation

end

Prolongation and restriction:

$P = I \otimes P$, \hspace{1cm} induced by natural inclusion in spatial domain

$R = P^T = I \otimes R$, $R = P^T$
Convergence Analysis: Use “Standard” Approach

Error propagation matrix:

\[ e^{(i+1)} = [(A^{(h)})^{-1} - P(A^{(2h)})^{-1} R)] [A^{(h)}(I - Q^{-1} A^{(h)})^k] e^{(i)} \]

Establish \textit{approximation property}

\[ \left\| [ (A^{(h)})^{-1} - P(A^{(2h)})^{-1} R] y \right\|_{A^{(h)}} \leq C \left\| y \right\|_2 \quad \forall y \]

and \textit{smoothing property}

\[ \left\| [A^{(h)}(I - Q^{-1} A^{(h)})^k] y \right\|_2 \leq \eta(k) \left\| y \right\|_{A^{(h)}} \quad \forall y, \quad \eta(k) \to 0 \quad k \text{ increases} \]

Analysis is:

\[ \left\| e^{(i+1)} \right\|_{A^{(h)}} \leq \left\| (A^{(h)})^{-1} - P(A^{(2h)})^{-1} R)] [A^{(h)}(I - Q^{-1} A^{(h)})^k] e^{(i)} \right\|_{A^{(h)}} \]

\[ \leq C \left\| [A^{(h)}(I - Q^{-1} A^{(h)})^k] e^{(i)} \right\|_2 \]

\[ \leq C \eta(k) \left\| e^{(i)} \right\|_{A^{(h)}} \]
Approximation Property

“Standard” MG analysis for deterministic problem:

\[ \left\| \left( A^{(h)} \right)^{-1} - \mathcal{P} \left( A^{(2h)} \right)^{-1} \mathcal{R} \right\|_{A^{(h)}} \left( A^{(h)} \right) = \left\| u^{(h)} - u^{(2h)} \right\|_{A^{(h)}} \]

\[ = \left\| u_h - u_{2h} \right\|_a \quad (= a(u_h - u_{2h}, u_h - u_{2h})^{1/2}) \]

\[ \leq \left\| u_h - u \right\|_a + \left\| u - u_{2h} \right\|_a \]

Approximability

\[ \leq \sqrt{\alpha_2} \left( C h \left\| D^2 u \right\|_{L^2(\mathcal{D})} + C 2h \left\| D^2 u \right\|_{L^2(\mathcal{D})} \right) \]

Regularity

\[ \leq C h \left\| f \right\|_{L^2(\mathcal{D})} \]

Property of mass matrix

\[ \leq C \left\| y \right\|_2 \]
For Approximation Property in Stochastic Case

Introduce *semi-discrete* space \( H_0^1(\mathcal{D}) \otimes T_p \)

Weak formulation

\[
a(u_p, v_p) = \ell(v_p) \quad \text{for all } v_p \in H_0^1(\mathcal{D}) \otimes T_p
\]

Solution \( u_p \)

\[
\left\| \left( (A^{(h)})^{-1} - P(A^{(2h)})^{-1} \mathcal{R} \right) y \right\|_{A^{(h)}} = \left\| u_{hp} - u_{2h,p} \right\|_a \leq \left\| u_h - u_p \right\|_a + \left\| u_p - u_{2h} \right\|_a
\]

Approximation (in 2D):

\[
\left\| u_p - u_{hp} \right\|_a \leq C h \left\| D^2 u_p \right\|_{L^2(\mathcal{D}) \otimes L^2(\Gamma)}
\]

Established using best approximation property of \( u_{hp} \) and interpolant \( \tilde{u}_p(x_j, \xi) = u_p(x_j, \xi) \quad \forall \xi \)

Similarly for other steps used for deterministic analysis
Comments

• Establishes convergence of multigrid with rate independent of spatial discretization size $h$

• No dependence on stochastic parameters $m, p$

• Applies to any basis of stochastic space

• Coarse grid operator: $G = a_0 G_0 + \sigma \sum_{r=1}^{m} a_r \sqrt{\lambda_r} G_r$, size $O(N_\xi)$

  $G_r$ derives from basis of multivariate polynomials of total degree $p$, orthogonal wrt probability measure $\rho(\xi)d\xi$

  Maximum eigenvalue $\eta = \max$ root of orthogonal polynomial, bounded for bounded measure

  $0 < a_0^{l \times l} - \sigma \eta (\sum_{r=1}^{m} a_r^{l \times l} \sqrt{\lambda_r}) \leq \lambda(G) \leq a_0^{l \times l} + \sigma \eta (\sum_{r=1}^{m} a_r^{l \times l} \sqrt{\lambda_r})$,

  CG iteration is an option
## Iteration Counts / Normal Distribution

<table>
<thead>
<tr>
<th>Polynomial degree</th>
<th># terms (m) in KL-expansion</th>
<th>m=1</th>
<th>m=2</th>
<th>m=3</th>
<th>m=4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>h=1/16</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polynomial degree</td>
<td>p=1</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>p=2</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>p=3</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>p=4</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>h=1/32</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polynomial degree</td>
<td>p=1</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>p=2</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>p=3</td>
<td>8</td>
<td>8</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>p=4</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
</tbody>
</table>
Multigrid Solution of Matrix Equation II

Solving  \( Au = f \)

\[
A = G_0 \otimes A_0 + \sum_{r=1}^{m} G_r \otimes A_r
\]

\[
[A_r]_{jk} = \sqrt{\lambda_r} \sigma \int_{\mathcal{D}} a_r(x) \nabla \varphi_j(x) \cdot \nabla \varphi_k(x) \, dx,
\]

\[
[G_r]_{lq} = \int_{\Omega} \psi_l(\xi)\psi_q(\xi)\xi_r\rho(\xi) \, d\xi
\]

Preconditioner for use with CG:  \( Q = G_0 \otimes A_0 \)  (Kruger, Pellissetti, Ghanem)

\[
A_0 \sim \int_{\mathcal{D}} a_0(x) \nabla \varphi_j(x) \cdot \nabla \varphi_k(x) \, dx \quad \text{Deterministic diffusion, from mean}
\]

\[
G_0 = I
\]
Analysis (Powell & E.)

Recall \( a(x, \omega) = a_0(x) + \sigma \sum_{r=1}^{m} \sqrt{\lambda_r} \ a_r(x) \xi_r(\omega) \)

\[ A = G_0 \otimes A_0 + \sum_{r=1}^{m} G_r \otimes A_r \]

\[ Q = G_0 \otimes A_0 \]

**Theorem**: For \( \mu \) constant, the Rayleigh quotient satisfies

\[
1 - \tau \leq \frac{(w, Aw)}{(w, Qw)} \leq 1 + \tau
\]

\[
\tau = (\sigma / \mu) c(p) \sum_{r=1}^{m} \sqrt{\lambda_r} \| a_r \|_\infty
\]

Consequence: \( \kappa \leq \frac{1 + \tau}{1 - \tau} \) dictates convergence of PCG
Sketch of Proof

\[ \tau = \left( \frac{\sigma}{\mu} \right) c(p) \sum_{r=1}^{m} \sqrt{\lambda_r} \| a_r \|_\infty \]

\[ A = G_0 \otimes A_0 + \sum_{r=1}^{m} G_r \otimes A_r \]

In spatial domain:

\[ (\varphi, A_r \varphi) \sim \sigma \sqrt{\lambda_r} \int_{D} a_r(x) \nabla \varphi(x) \cdot \nabla \varphi(x) \, dx \]

\[ \leq \sigma \sqrt{\lambda_r} \| a_r \|_\infty \int_{D} \nabla \varphi(x) \cdot \nabla \varphi(x) \, dx \]

\[ = \left( \frac{\sigma}{\mu} \right) \sqrt{\lambda_r} \| a_r \|_\infty (\varphi, A_0 \varphi) \]

From stochastic component: as above

\( c(p) \) bounded by largest root of scalar orthogonal polynomial
Multigrid Variant of this Idea

Replace action of $A_0^{-1}$ with multigrid preconditioner

$$Q_{MG} = G_0 \otimes A_{0, MG}$$

(Le Maitre, et al.)

Analysis:

$$\frac{(w, Aw)}{(w, Q_{MG} w)} = \frac{(w, Aw)}{(w, Qw)} \frac{(w, Qw)}{(w, Q_{MG} w)}$$

Spectral equivalence of MG approximation to diffusion operator

$$\in [\beta_1, \beta_2]$$

$$\implies \kappa \leq \frac{(1 + \tau)}{(1 - \tau)} \frac{\beta_2}{\beta_1}$$
Experiment

Starting with \( a \) with specified covariance and small \( \sigma (=.01) \):

Compare Monte-Carlo simulation with SFEM, for

\[-\nabla \cdot (a \nabla u) = f\]

N.B.: No negative samples of diffusion obtained in MC

<table>
<thead>
<tr>
<th></th>
<th># Samples ( s )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Max SFEM 100 1000 10,000 40,000</td>
</tr>
<tr>
<td>Max</td>
<td>SFEM 100 1000 10,000 40,000</td>
</tr>
<tr>
<td>Mean</td>
<td>.06311 .06361 .06330 .06313 .06313</td>
</tr>
<tr>
<td>Variance</td>
<td>2.360(-5) 2.161(-5) 2.407(-5) 2.258(-5) 2.316(-5)</td>
</tr>
</tbody>
</table>

Solve one system of order 210x225

Solve \( s \) systems of size 225
Comparison of Galerkin and Collocation

Recall, for stochastic collocation

Discrete solution

\[ u_{hp}(x, \xi) = \sum_{k=1}^{N_\xi} u_h^{(k)}(x) L_k(\xi) \]

Obtained by solving

\[ \int_{\Omega} (a_0(x) + \sigma \sum_{r=1}^{m} \sqrt{\lambda_r} a_r(x) \xi_r) \nabla u \cdot \nabla v \, dx = \int_{\Omega} f \, v \, dx \]

For set of samples \( \{\xi^{(k)}\} \) situated in a sparse grid

Advantage of this approach: simpler (decoupled) systems
Disadvantage: larger stochastic space for comparable accuracy
larger by factor approximately \( 2^p \)
## Dimensions of Stochastic Space

<table>
<thead>
<tr>
<th>m (#KL)</th>
<th>p</th>
<th>Galerkin</th>
<th>Collocation Sparse</th>
<th>Collocation Tensor</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1</td>
<td>5</td>
<td>9</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>15</td>
<td>41</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>35</td>
<td>137</td>
<td>256</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>70</td>
<td>401</td>
<td>625</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>11</td>
<td>21</td>
<td>1024</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>66</td>
<td>221</td>
<td>59,049</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>286</td>
<td>1582</td>
<td>1,048,576</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1001</td>
<td>8,801</td>
<td>9,765,625</td>
</tr>
<tr>
<td>30</td>
<td>1</td>
<td>31</td>
<td>61</td>
<td>1.07(9)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>496</td>
<td>1861</td>
<td>2.06(14)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>5456</td>
<td>37,941</td>
<td>1.15(18)</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>46,376</td>
<td>582,801</td>
<td>9.31(20)</td>
</tr>
</tbody>
</table>

~ size of coarse grid space for MG / Version 1

# systems for collocation MG / Version II
Experiment

(E., Miller, Phipps, Tuminaro)

• Solve the stochastic diffusion equation by both methods
• Compare the accuracy achieved for different parameter sets¹
• For parameter choices giving comparable accuracy, compare solution costs
• Spatial discretization fixed (32x32 finite difference grid)

Solution algorithm for both discretizations:
Preconditioned conjugate gradient with mean-based preconditioning, using AMG for the approximate diffusion solve

¹Estimated using a high-degree (p=10) Galerkin solution.
Experimental Results

Accuracy:
for fixed $m=4$: similar $p=$ polynomial degree for SG “level” for collocation produces comparable errors

Performance:
Experimental Results: Performance

Performed on a serial machine with C code and CG/AMG code from Trilinos
Observation: Galerkin faster, more so as number of stochastic variables (KL terms) grows

CPU times for larger $m = \#$KL terms:

<table>
<thead>
<tr>
<th>p</th>
<th>m=5</th>
<th>m=10</th>
<th>m=15</th>
<th>m=5</th>
<th>m=10</th>
<th>m=15</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.058</td>
<td>.147</td>
<td>.32-</td>
<td>.069</td>
<td>.163</td>
<td>.286</td>
</tr>
<tr>
<td>2</td>
<td>.269</td>
<td>1.20</td>
<td>3.80</td>
<td>.532</td>
<td>2.13</td>
<td>5.08</td>
</tr>
<tr>
<td>3</td>
<td>1.20</td>
<td>13.14</td>
<td>51.45</td>
<td>2.41</td>
<td>16.99</td>
<td>57.98</td>
</tr>
<tr>
<td>4</td>
<td>3.50</td>
<td>53.79</td>
<td>168.11</td>
<td>8.31</td>
<td>102.60</td>
<td>493.04</td>
</tr>
<tr>
<td>5</td>
<td>6.51</td>
<td>117.73</td>
<td>24.56</td>
<td>515.75</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
More General Problems

For the problem discussed, based on a KL expansion, has a *linear* dependence on the stochastic variable $\xi$.

Other models have *nonlinear* dependence. For example

$$a(x, \xi) = a_{\text{min}} + e^{c(x, \xi)}$$

$c(x, \xi) = a_0(x) + \sigma \sum_{r=1}^{m} \sqrt{\lambda_r} a_r(x) \xi_r$

Nonlinear

For Gaussian $c$, called a *log-normal* distribution

In particular: coercivity is guaranteed with this choice.
More General Problems

For stochastic Galerkin, need a finite term expansion for $a$

$$a(x, \xi) = a_0(x) + \sigma \sum_{r=1}^{M} \sqrt{\lambda_r} a_r(x) \psi_r(\xi)$$

Note: not $\xi_r$

$\longrightarrow$ matrix

$$A = G_0 \otimes A_0 + \sum_{r=1}^{M} G_r \otimes A_r$$

$$[G_r]_{ij} = \left\langle \psi_r \psi_i \psi_j \right\rangle \quad \text{Less sparse}$$

More importantly: # terms $M$ will be larger

perhaps as large as $2N\xi$

$\longrightarrow$ mvp will be more expensive

49
In Contrast

Collocation is less dependent on this expansion

\[ A^{(k)} \text{ comes from } \int_{\mathcal{D}} a(x, \xi^{(k)}) \nabla u \cdot \nabla v \, dx \text{ for each \xi^{(k)}} \]

sparse grid point \( \xi^{(k)} \)

Many matrices to assemble, but mvp is not a difficulty
Concluding Remarks

• Exciting new developments models of PDEs with uncertain coefficients
• Replace pure simulation (Monte Carlo) with finite-dimensional models that simulate sampling at potentially lower cost
• Two techniques, the *stochastic Galerkin* method and the *stochastic collocation* method, were presented, each with some advantages
• Solution algorithms are available for both methods, and work continues in this direction