# 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

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

Uncertainty / randomness:

$a = a(\mathrm{x},\omega)$ a *random field*

For each fixed $x$, $a(x,\omega)$ a random variable

Other possibly uncertain quantities :

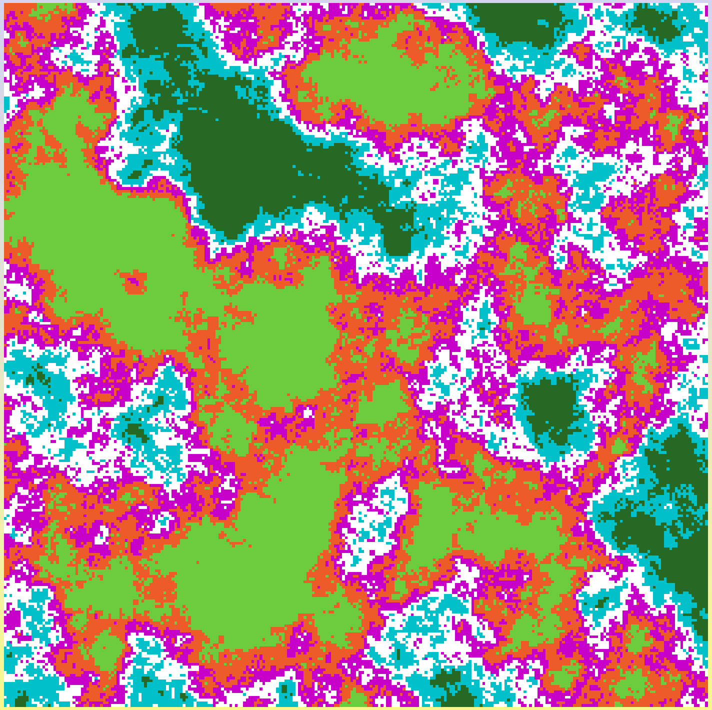
Forcing function $f$

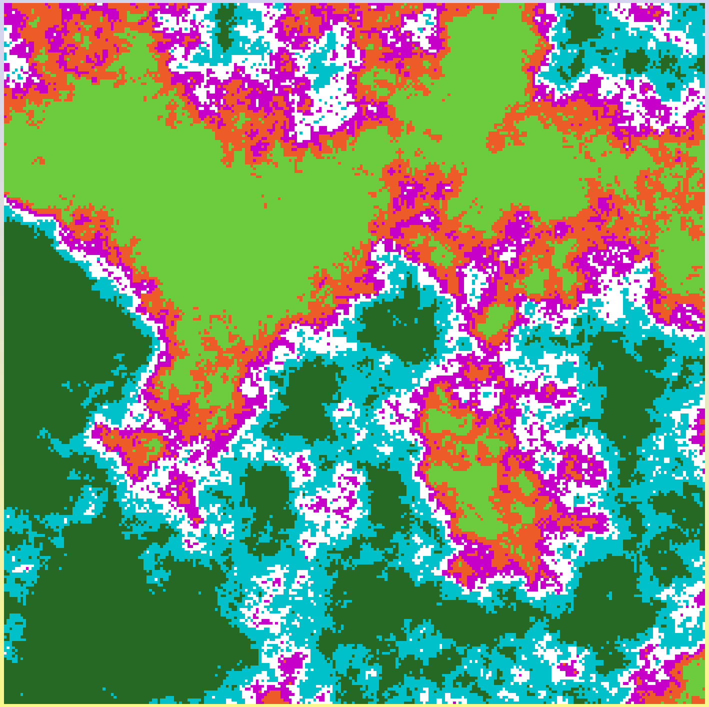
Boundary data $g_D$

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

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

**Assumptions:**

1. Spatial correlation of random field: For $x, y \in \mathcal{D}$:

   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

   vs. *white noise*, where $c$ is a $\delta$-function

2. Coercivity $0 < \alpha_1 \le a \le \alpha_2 < \infty$

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

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

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

**Problem: convergence is slow, requires many solves**

# Another Point of View

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

Covariance function is finite $\Longrightarrow$
random field (diffusion coefficient) has *Karhunen-Loève* expansion:

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

$$a_0(x) = \mu(x) = E(a(x,\cdot)) \quad \text{mean}$$

$a_r(x), \lambda_r =$ eigenfunctions/eigenvalues of covariance operator

$$(Ca)(x) = \lambda a(x), \qquad (Ca)(x) = \int_{\mathcal{D}} c(x,y)a(y)dy$$

$\xi_r(\omega)$ = identically distributed uncorrelated random variables with mean 0 and variance 1

# Finite Noise Assumption

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

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 $\dfrac{|\mathcal{D}|\sigma^2 - \sum_{j=1}^{m} \lambda_j}{|\mathcal{D}|\sigma^2}$

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 $\underline{\xi}$ 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 $H^1_E(\mathcal{D}) \otimes L^2_P(\Omega)$

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

$$\langle a(u,v) \rangle = \langle \ell(v) \rangle \qquad \langle a(u,v) \rangle = \int_\Omega \int_D a \nabla u \cdot \nabla v \, dx \, dP(\omega)$$

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

Solution ***u*=*u*(*x*,*ω*)** 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

$$\langle a(u,v)\rangle = \int_\Omega \int_D a \nabla u \cdot \nabla v \, dx \, dP(\omega) = \int_{\xi(\Omega)} \int_D a(x,\underline{\xi}) \nabla u \cdot \nabla v \, dx \, \rho(\underline{\xi}) d\underline{\xi}$$

Bilinear form entails integral over *image* of random variables $\underline{\xi}$

Require joint density function associated with $\underline{\xi}$

$\underline{\xi}$ plays the role of a Cartesian coordinate

# Statement of Problem Becomes

Find $u \in H_E^1(\mathcal{D}) \otimes L_P^2(\Gamma)$ such that

$$\int_\Gamma \int_\mathcal{D} a(x,\underline{\xi})\nabla u \cdot \nabla v \, dx \, \rho(\underline{\xi}) d\underline{\xi} = \int_\Gamma \int_\mathcal{D} fv \, dx \, \rho(\underline{\xi}) d\underline{\xi}$$

for all $v \in H_{E_0}^1(\mathcal{D}) \otimes L_P^2(\Gamma) \quad (\Gamma = \underline{\xi}(\Omega))$

Like an ordinary Galerkin (or Petrov-Galerkin) problem on a ($d$+$m$)-dimensional "continuous" space

$d$ = dimension of spatial domain
$m$ = dimension of stochastic space

# Discretization

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

**Finite dimensional spaces:**

- spatial discretization: $S_h \subset H_0^1(\mathcal{D})$, spanned by $\{\varphi_j\}_{j=1}^{N_x}$

  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:**

$$a(u_{hp}, v_{hp}) = \ell(v_{hp}) \quad \text{for all } v_{hp} \in S_h \otimes T_p$$

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

# Basis Functions for Stochastic Space

Underlying space: $L^2(\Gamma) = \left\{ \mathrm{v}(\underline{\xi}) \middle| \int_\Gamma \mathrm{v}(\underline{\xi})^2 \rho(\underline{\xi})\, d\underline{\xi} < \infty \right\}$

$$\rho(\underline{\xi}) = \rho_1(\xi_1)\rho_2(\xi_2)\cdots\rho_M(\xi_M)$$

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

Examples: if $\rho_k$ ~ *Gaussian measure* ⟶ Hermite polynomials

$\rho_k$ ~ *uniform distribution* ⟶ Legendre polynomials

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

⟶ 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 ⟶ sparsity of coefficient matrix

# **Matrix Equation *Au=f***

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

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

$$[A_0]_{jk} = \int_{\mathcal{D}} a_0(x) \nabla \varphi_j(x) \cdot \nabla \varphi_k(x)\, dx$$

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

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

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

Properties of *A*:

- order = $N_x$ x $N_\xi$ = (size of spatial basis) x (size of stochastic basis)
- sparsity: inherited from that of $\{G_r\}$ and $\{A_r\}$

# Dimensions of Discrete Stochastic Space

$T_p \subset L^2(\Gamma)$ 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: $0 \le j_i \le p, \quad i = 1, \ldots, m$

Dimension: $(p+1)^m$ Too large

"Complete" polynomial basis: $j_1 + j_2 + \cdots + j_m \le p$

Dimension: $\binom{m+p}{p} = \frac{(m+p)!}{m!\, p!}$ More manageable

Order these in a systematic way $\longrightarrow$

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

# Example

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

"Complete" polynomial basis: $j_1 + j_2 + \cdots + j_m \leq p$

$$m=2,\ p=3 \longrightarrow \binom{m+p}{p} = \binom{5}{2} = 10$$

Orthogonal (Hermite) polynomials in 1D:

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

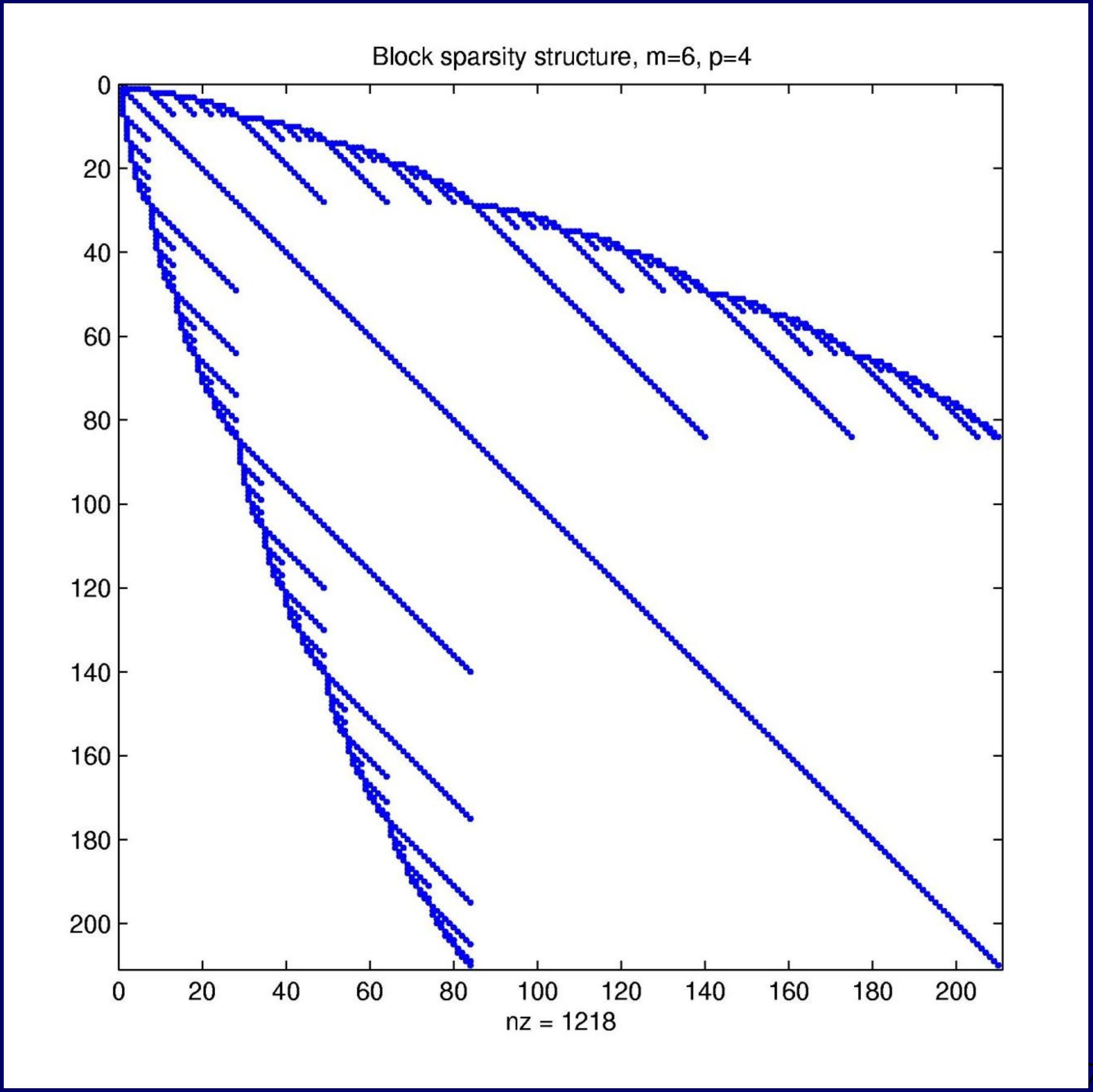
Gives basis set:

$$\psi_1(\underline{\xi}) = 1$$
$$\psi_2(\underline{\xi}) = \xi_1$$
$$\psi_3(\underline{\xi}) = \xi_1^2 - 1$$
$$\psi_4(\underline{\xi}) = \xi_1^3 - 3\xi_1$$
$$\psi_5(\underline{\xi}) = \xi_2$$
$$\psi_6(\underline{\xi}) = \xi_1 \xi_2$$
$$\psi_7(\underline{\xi}) = (\xi_1^2 - 1)\xi_2$$
$$\psi_8(\underline{\xi}) = (\xi_2^2 - 1)$$
$$\psi_9(\underline{\xi}) = (\xi_2^2 - 1)\xi_1$$
$$\psi_{10}(\underline{\xi}) = \xi_2^3 - 3\xi_2$$

# Example of Sparsity Pattern

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

$$N_\xi = \frac{(m+p)!}{m!p!} = \frac{10!}{6!4!} = 210$$



# Uses of the Computed Solution:

$$u_{hp} = \sum_{l=1}^{N_\xi} \underbrace{\sum_{j=1}^{N_x} u_{jl}\varphi_j(x)}_{u_l(x)} \psi_l(\underline{\xi}) = \sum_{l=1}^{N_\xi} u_l(x)\psi_l(\underline{\xi})$$

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

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

**Free!**

$$= u_1(x) = \sum_{j=1}^{N} u_{j1}\varphi_j(x)$$

using orthogonality of stochastic basis functions

Similarly for second moment / covariance

# Uses of the Computed Solution:

$$u_{hp} = \underbrace{\sum_{l=1}^{N_\xi} \sum_{j=1}^{N_x} u_{jl} \varphi_j(x)}_{u_l(x)} \psi_l(\underline{\xi}) = \sum_{l=1}^{N_\xi} u_l(x) \psi_l(\underline{\xi})$$

2. **Cumulative distribution functions**

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

Sample $\underline{\xi}$

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

$u_l(x)$ → Precomputed

Repeat

Not free, but **<u>no solves required</u>**

# 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 $\underline{\xi}$ be a specified realization (~ Monte Carlo) $\longrightarrow$

Weak formulation:

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

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

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

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

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

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

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

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

# To Compute Statistical Quantities

Solution $$u_{hp}(x,\underline{\xi}) = \sum_{k=1}^{N_\xi} u_h^{(k)}(x) L_k(\underline{\xi})$$

1. **Moments**

$$E(u_{hp})(x) = \sum_{k=1}^{N_\xi} u_h^{(k)}(x) \underbrace{\int_\Gamma L_k(\underline{\xi})\rho(\underline{\xi})d\underline{\xi}}_{\text{Not free but can be precomputed}}$$

2. **Distribution functions**

Obtained by sampling, cheap

# Strategy for Interpolation

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

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

$\ell_{k_j}$ = 1D interpolating polynomial

$0 \le k_j \le p$

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 ***fully populated*** multidimensional grid $Y^{(1)} \times Y^{(2)} \times \cdots \times Y^{(m)}$.

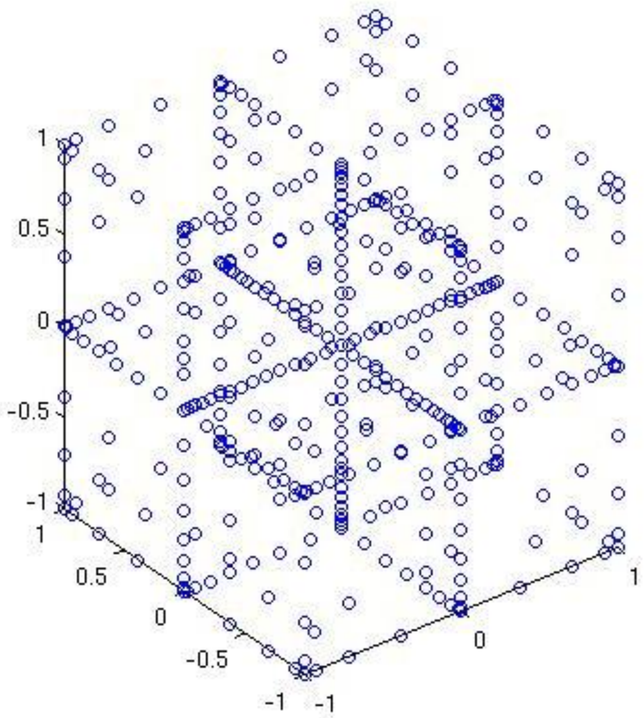
$$|Y^{(k)}| = m_k = p + 1$$

Alternative: multidimensional ***sparse grid*** (Smolyak)

$$\mathcal{H}(m+p, m) \equiv \bigcup_{p-m+1 \le i_1 + \cdots + i_m \le p} (Y^{(i_1)} \times Y^{(i_2)} \times \cdots \times Y^{(i_m)})$$

# Sparse Grid Interpolation

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




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

$$(\mathcal{I}v)(\underline{\xi}) = \sum_{i_1+\cdots+i_m \le p} (U^{(i_1)} - U^{(i_1-1)})v_1(\xi_1) \otimes (U^{(i_2)} - U^{(i_2-1)})v_2(\xi_2) \otimes \cdots \otimes (U^{(i_m)} - U^{(i_m-1)})v_m(\xi_m)$$

# Sparse Grid Interpolation

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

For $\underline{\xi} \in$ sparse grid and $v(\underline{\xi})$ a tensor product polynomial of total degree at most *p*,

$$v(\underline{\xi}) = q_{j_1}^{(1)}(\xi_1) q_{j_2}^{(2)}(\xi_2) \cdots q_{j_m}^{(m)}(\xi_m), \qquad j_1 + j_2 + \cdots + j_m \le p$$

$$(\mathcal{I}v)(\underline{\xi}) = v(\underline{\xi}).$$

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

$$\approx 2^{p} \binom{m+p}{p} \qquad \text{vs.} \qquad \binom{m+p}{p}$$

# **Analysis** (Babuška, Tempone, Zouraris, Nobile, Webster)

**Monte-Carlo:**

$$E(u)-E_s(u_h)=\underbrace{(E(u)-E(u_h))}_{\le c_1 h\, E(|u|_2)}+\underbrace{(E(u_h)-E_s(u_h))}_{\sim 1/\sqrt{s}}$$

Convergence is slow wrt number of samples but independent of number of random variables $m$

**Stochastic Galerkin and Collocation:**

$$E(u)-E(u_{hp})=\underbrace{(E(u)-E(u_h))}_{\le c_1 h\, E(|u|_2)}+\underbrace{(E(u_h)-E(u_{hp}))}_{\le c_2 r^p,\ r<1}$$

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 $u_{hp}(x,\underline{\xi})$

Stochastic Galerkin:

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

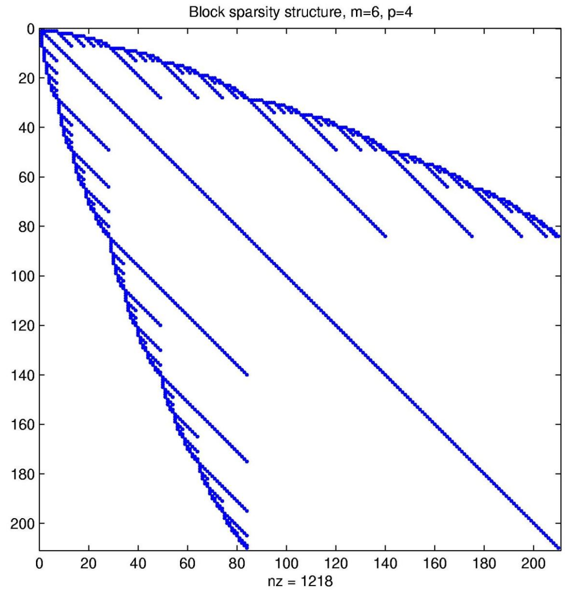
Stochastic Collocation:

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

Postprocess to get statistics

# Computational Issues

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




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

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 **$Au=f$**

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

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

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

$A_r = A_r^{(h)}, \quad A = A^{(h)},$ spatial discretization parameter $h$

$A_r = A_r^{(2h)}, \quad A = A^{(2h)},$ 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* k smoothing steps

$$u^{(h)} \leftarrow (I - Q^{-1}A^{(h)})u^{(h)} + Q^{-1}f^{(h)}$$

end

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

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

$u^{(h)} \leftarrow u^{(h)} + \mathcal{P}c^{(2h)}$ Prolongation

end

Prolongation and restriction:

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

$\mathcal{R} = \mathcal{P}^T = I \otimes R, \quad R = P^T$

# Convergence Analysis: Use "Standard" Approach

Error propagation matrix:

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

Establish ***approximation property***

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

and ***smoothing property***

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

Analysis is:

$$\begin{aligned} \| e^{(i+1)} \|_{A^{(h)}} &\le \| (A^{(h)})^{-1} - \mathcal{P}(A^{(2h)})^{-1}\mathcal{R})][A^{(h)}(I - Q^{-1}A^{(h)})^{k}]\,e^{(i)} \|_{A^{(h)}} \\ &\le C \| [A^{(h)}(I - Q^{-1}A^{(h)})^{k}]\,e^{(i)} \|_2 \\ &\le C\eta(k) \| e^{(i)} \|_{A^{(h)}} \end{aligned}$$

# Approximation Property

"Standard" MG analysis for deterministic problem:

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

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

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

Approximability

$$\leq \sqrt{\alpha_2}\left( Ch \left\| D^2 u \right\|_{L^2(\mathcal{D})} + C2h \left\| D^2 u \right\|_{L^2(\mathcal{D})} \right)$$

Regularity

$$\leq Ch \left\| f \right\|_{L^2(\mathcal{D})}$$

Property of mass matrix

$$\leq C \left\| y \right\|_2$$

# For Approximation Property in Stochastic Case

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

Discrete stochastic space

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 $u_p$

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

$$\leq \left\| u_h - u_p \right\|_a + \left\| u_p - u_{2h} \right\|_a$$

Approximation (in 2D):

$$\left\| u_p - u_{hp} \right\|_a \leq Ch \left\| D^2 u_p \right\|_{L^2(\mathcal{D}) \otimes L^2(\Gamma)}$$

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

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

  $$\Rightarrow 0 < a_0^{1x1} - \sigma\eta\left(\sum_{r=1}^{m} a_r^{1x1}\sqrt{\lambda_r}\right) \le \lambda(G) \le a_0^{1x1} + \sigma\eta\left(\sum_{r=1}^{m} a_r^{1x1}\sqrt{\lambda_r}\right),$$

  CG iteration is an option

# Iteration Counts / Normal Distribution

# terms (m) in KL-expansion

h=1/16

Polynomial degree

| | m=1 | m=2 | m=3 | m=4 |
|---|---|---|---|---|
| p=1 | 8 | 8 | 8 | 8 |
| p=2 | 8 | 8 | 8 | 8 |
| p=3 | 9 | 9 | 9 | 9 |
| p=4 | 9 | 10 | 10 | 10 |

h=1/32

Polynomial degree

| | m=1 | m=2 | m=3 | m=4 |
|---|---|---|---|---|
| p=1 | 7 | 7 | 8 | 8 |
| p=2 | 8 | 8 | 8 | 8 |
| p=3 | 8 | 8 | 9 | 9 |
| p=4 | 9 | 9 | 9 | 9 |

# Multigrid Solution of Matrix Equation II

Solving ***Au=f***

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

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

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

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

$$A_0 \sim \int_{\mathcal{D}} a_0(x)\nabla\varphi_j(x)\cdot\nabla\varphi_k(x)\,dx$$

Deterministic diffusion, from mean

$$G_0 = I$$

# Analysis (Powell & E.)

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

$$\longrightarrow \quad A = G_0 \otimes A_0 + \sum_{r=1}^{\mathrm{m}} G_r \otimes A_r$$

$$Q = G_0 \otimes A_0$$

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

$$1 - \tau \le \frac{(w, Aw)}{(w, Qw)} \le 1 + \tau$$

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

Consequence: $\kappa \le \dfrac{1+\tau}{1-\tau}$ dictates convergence of PCG

**Sketch of Proof** $\tau = (\sigma/\mu)\, c(p) \sum_{r=1}^{m} \sqrt{\lambda_r}\, \| a_r \|_\infty$

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

In spatial domain:

$$(\varphi, A_r \varphi) \sim \sigma \sqrt{\lambda_r} \int_D a_r(x) \nabla \varphi(x) \cdot \nabla \varphi(x)\, dx$$

$$\leq \sigma \sqrt{\lambda_r}\, \| a_r \|_\infty \int_D \nabla \varphi(x) \cdot \nabla \varphi(x)\, dx$$

$$= (\sigma / \mu) \sqrt{\lambda_r}\, \| a_r \|_\infty\, (\varphi, A_0 \varphi)$$

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: $$\frac{(w, Aw)}{(w, Q_{MG}w)} = \frac{(w, Aw)}{(w, Qw)} \underbrace{\frac{(w, Qw)}{(w, Q_{MG}w)}}_{\in [\beta_1, \beta_2]}$$

Spectral equivalence of MG approximation to diffusion operator

$$\Longrightarrow \quad \kappa \le \frac{(1+\tau)}{(1-\tau)} \frac{\beta_2}{\beta_1}$$

# 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

| | | # Samples *s* | | | |
|---|---|---|---|---|---|
| Max | SFEM | 100 | 1000 | 10,000 | 40,000 |
| Mean | .06311 | .06361 | .06330 | .06313 | .06313 |
| Variance | 2.360(-5) | 2.161(-5) | 2.407(-5) | 2.258(-5) | 2.316(-5) |

Solve one system of order 210x225

Solve ***s*** systems of size 225

# Comparison of Galerkin and Collocation

Recall, for **stochastic collocation**

Discrete solution $u_{hp}(x,\underline{\xi}) = \sum_{k=1}^{N_\xi} u_h^{(k)}(x) L_k(\underline{\xi})$

Obtained by solving

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

For set of samples $\{\underline{\xi}^{(k)}\}$ 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

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

~ 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 sets[1]
- For parameter choices giving comparable accuracy, compare solution costs
- Spatial discretization fixed (32x32 finite difference grid)

Solution algorithm for both discretizations:

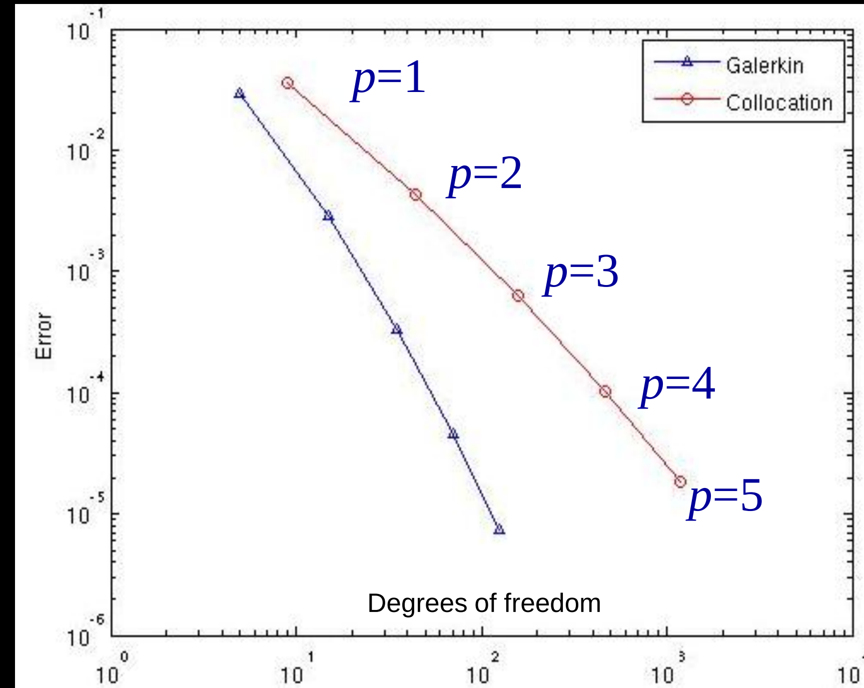
Preconditioned conjugate gradient with mean-based preconditioning, using AMG for the approximate diffusion solve

[1]Estimated using a high-degree (p=10) Galerkin solution.

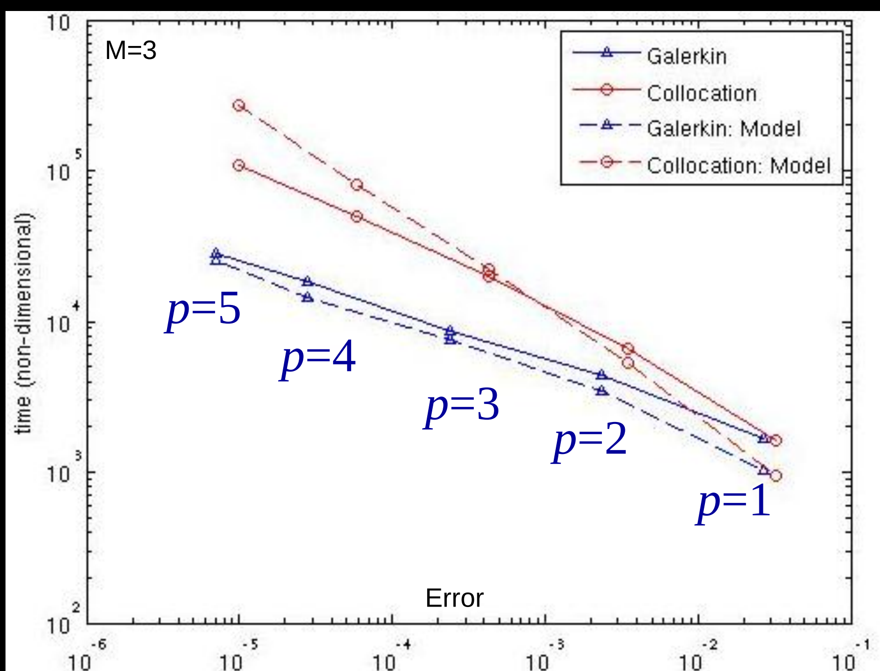
# Experimental Results

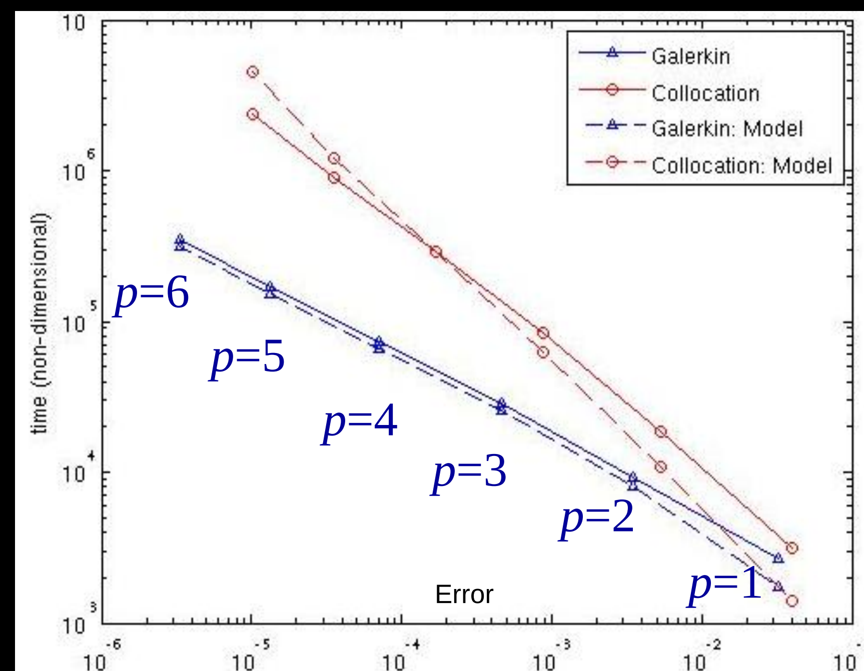
Accuracy:

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:

| | Galerkin | | | Collocation | | |
|---|---|---|---|---|---|---|
| p | m=5 | m=10 | m=15 | m=5 | m=10 | m=15 |
| 1 | .058 | .147 | .32- | .069 | .163 | .286 |
| 2 | .269 | 1.20 | 3.80 | .532 | 2.13 | 5.08 |
| 3 | 1.20 | 13.14 | 51.45 | 2.41 | 16.99 | 57.98 |
| 4 | 3.50 | 53.79 | 168.11 | 8.31 | 102.60 | 493.04 |
| 5 | 6.51 | 117.73 | | 24.56 | 515.75 | |

# More General Problems

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

Other models have *nonlinear* dependence. For example

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

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

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*

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

Note: not $\xi_r$

⟶ matrix

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

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

Less sparse

More importantly: # terms *M* will be larger
perhaps as large as $2N_\xi$

⟹ mvp will be more expensive

# In Contrast

Collocation is less dependent on this expansion

$A^{(k)}$ comes from $\int_{\mathcal{D}} a(x, \underline{\xi}^{(k)}) \nabla u \cdot \nabla v \, dx$ for each sparse grid point $\underline{\xi}^{(k)}$

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