An Overview of Solution Strategies for Stochastic Galerkin Discretizations

Elisabeth Ullmann

Institut für Numerische Mathematik und Optimierung
Technische Universität Bergakademie Freiberg

Computational Linear Algebra for PDEs
LMS Durham Symposium 2008

Outline

- Stochastic Galerkin (sG) method
  - Variational formulation of stochastic diffusion equation
  - Discretization and structure of Galerkin matrix
- Solution strategies for Galerkin system
  - Decoupling the Galerkin system ?
  - Preconditioning

Model problem

Stochastic diffusion equation

Given: bounded spatial domain $D \subset \mathbb{R}^2$ with boundary $\Gamma = \Gamma_D \cup \Gamma_N$ and a complete probability space $(\Omega, \mathcal{A}, P)$.

$$\begin{aligned} -\nabla \cdot (T(\mathbf{x},\omega)\nabla p(\mathbf{x},\omega)) &= F(\mathbf{x}), && \mathbf{x} \in D,\ P - a.s. \\ p(\mathbf{x}) &= p_D(\mathbf{x}), && \mathbf{x} \in \Gamma_D, \\ \mathbf{n} \cdot (T\nabla p)(\mathbf{x}) &= 0, && \mathbf{x} \in \Gamma_N. \end{aligned}$$

Note: T and therefore p are **random fields**.

Assume: $T(\mathbf{x},\omega) \geq T_\ell > 0 \quad \mathbf{x} \in D,\ P - a.s.$

Stochastic variational formulation

Stochastic diffusion equation

Find a function $p \in H^1_{\Gamma_D}(D) \otimes L^2_P(\Omega)$, such that for all test functions $v \in H^1_0(D) \otimes L^2_P(\Omega)$ there holds

$$\left\langle \int_D T(\mathbf{x},\omega)\nabla p(\mathbf{x},\omega) \cdot \nabla v(\mathbf{x},\omega)\, d\mathbf{x} \right\rangle = \left\langle \int_D F(\mathbf{x})\, v(\mathbf{x},\omega) d\mathbf{x} \right\rangle .$$

$\langle \cdot \rangle$ denotes the expectation operator w.r.t. the measure P,

$$\langle \xi \rangle := \int_\Omega \xi(\omega)\, dP(\omega).$$

Discretization steps of stochastic variational formulation

- Input random fields depend on M mutually *independent* random variables $\{\xi_m\}_{m=1}^M$ with given probability density functions $\rho_m : \mathbb{R} \supseteq \Gamma_m \to [0,\infty)$.

$$\rho(\boldsymbol{\xi}) := \rho_1(\xi_1)\cdots\rho_M(\xi_M), \qquad \boldsymbol{\xi} \in \Gamma := \Gamma_1 \times \cdots \times \Gamma_M.$$

- Identify $L^2_P(\Omega)$ with $L^2_\rho(\Gamma)$ and $\langle\cdot\rangle$ with $\int_\Gamma \rho(\boldsymbol{\xi}) \cdot d\boldsymbol{\xi}$.
- Choose $X_h = \text{ span } \{\phi_1, \phi_2, \ldots, \phi_{N_\mathbf{x}}\} \subset H^1_{\Gamma_D}(D)$
- Choose $S_d = \text{ span } \left\{\psi_1, \psi_2, \ldots, \psi_{N_{\boldsymbol{\xi}}}\right\} \subset L^2_\rho(\Gamma)$
- Tensor product variational space: $X_h \otimes S_d \subset H^1_{\Gamma_D}(D) \otimes L^2_\rho(\Gamma)$
- $N_\mathbf{x} \cdot N_{\boldsymbol{\xi}}$ degrees of freedom !

Stochastic shape functions

Multivariate polynomials on Γ

$$\psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \prod_{m=1}^{M} \psi_{\alpha_m}^{(m)}(\xi_m), \quad \boldsymbol{\alpha} \in \mathcal{I} := \mathbb{N}_0^M$$

$\psi_0^{(m)}, \psi_1^{(m)}, \ldots, \psi_n^{(m)}, \ldots,$ are the (univariate) *orthonormal* polynomials of exact degree n associated with weight function ρ_m defined on Γ_m, $m = 1, \ldots, M$.

- Stochastic trial and test functions

  $S_d =$ span $\{\psi_{\boldsymbol{\alpha}}, \boldsymbol{\alpha} \in \mathcal{I}_S \subset \mathcal{I}\}$, $N_{\boldsymbol{\xi}} = |\mathcal{I}_S| < \infty$ d.o.f.s

- Representation of random diffusion coefficient

  $T(\mathbf{x}, \boldsymbol{\xi}) = \sum_{\boldsymbol{\alpha} \in \mathcal{I}_T} t_{\boldsymbol{\alpha}}(\mathbf{x}) \psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi})$, $|\mathcal{I}_T| < \infty$ expansion terms

Choice of stochastic shape functions: implications

Complete polynomials	Tensor product polynomials
$\mathcal{I}_S = \{\boldsymbol{\alpha} \in \mathcal{I},\ \|\boldsymbol{\alpha}\| \leq d\}$	$\mathcal{I}_S = \{\boldsymbol{\alpha} \in \mathcal{I},\ 0 \leq \alpha_m \leq d_m\}$
$N_{\boldsymbol{\xi}} = \binom{M+d}{M}$ d.o.f.s	$N_{\boldsymbol{\xi}} = \prod_{m=1}^{M}(1+d_m)$ d.o.f.s

T linear in $\boldsymbol{\xi}$ (e.g. Karhunen-Loève (KL) expansion)
$\mathcal{I}_T = \{\boldsymbol{\alpha} \in \mathcal{I},\ \|\boldsymbol{\alpha}\| \leq 1\} \Rightarrow \|\mathcal{I}_T\| = M+1$

T nonlinear in $\boldsymbol{\xi}$ (e.g. lognormal)	
$\|\mathcal{I}_T\| \leq \binom{M+2d}{M}$	$\|\mathcal{I}_T\| \leq \prod_{m=1}^{M}(1+2d_m)$

Structure of Galerkin matrix

$$A = \sum_{\boldsymbol{\alpha} \in \mathcal{I}_T} G_{\boldsymbol{\alpha}} \otimes K_{\boldsymbol{\alpha}}$$

where $G_{\boldsymbol{\alpha}} \in \mathbb{R}^{N_{\boldsymbol{\xi}} \times N_{\boldsymbol{\xi}}}$ and $K_{\boldsymbol{\alpha}} \in \mathbb{R}^{N_{\mathbf{x}} \times N_{\mathbf{x}}}$, $\boldsymbol{\alpha} \in \mathcal{I}_T$.

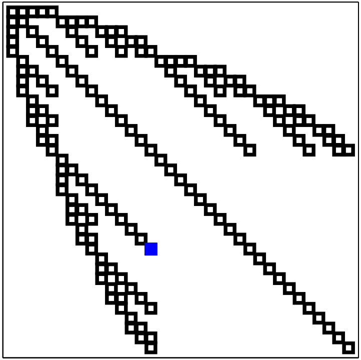
- Stochastic Galerkin matrices

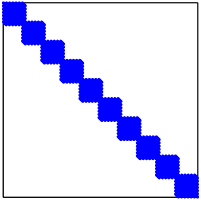
$$[G_{\boldsymbol{\alpha}}]_{\boldsymbol{\beta},\boldsymbol{\gamma}} = \langle \psi_{\boldsymbol{\alpha}} \psi_{\boldsymbol{\beta}} \psi_{\boldsymbol{\gamma}} \rangle, \quad \boldsymbol{\alpha} \in \mathcal{I}_T, \quad \boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathcal{I}_S.$$

- Deterministic stiffness matrices

$$[K_{\boldsymbol{\alpha}}]_{i,k} = \int_D t_{\boldsymbol{\alpha}}(\mathbf{x}) \nabla \phi_i(\mathbf{x}) \cdot \nabla \phi_k(\mathbf{x}) \, d\mathbf{x}, \; \boldsymbol{\alpha} \in \mathcal{I}_T, \; i, k = 1, \ldots, N_{\mathbf{x}}.$$

Structure of Galerkin matrix - Example





Solution strategies

T linear in ξ

Solve system in $N_{\mathbf{x}} \cdot N_{\xi}$ unknowns.

[Ghanem & Kruger], [Ghanem & Pellissetti], [Le Maître et al.], [Matthies & Keese], [Seynaeve et al.], [Elman & Furnival], [Elman & Powell], [Rosseel et al.]

T linear in ξ

Construct **doubly orthogonal** stochastic shape functions [Babuška et al.]. Solve N_{ξ} systems in $N_{\mathbf{x}}$ unknowns.

[Eiermann, Ernst & U.], [Cai et al.], [Ernst, Powell, Silvester, U.]

T nonlinear in ξ

Solve system in $N_{\mathbf{x}} \cdot N_{\xi}$ unknowns.

[Matthies & Keese], [Rosseel et al.]

T nonlinear in ξ

–

complete polynomials

tensor product polynomials

Thesis work

- Tensor product polynomials, T linear in $\boldsymbol{\xi}$:
  - After decoupling study of **Krylov subspace recycling** techniques for solving the resulting sequence of linear systems.
- Complete polynomials, T linear in $\boldsymbol{\xi}$:
  - **Decoupling** of global Galerkin matrix w.r.t. the stochastic d.o.f.s **impossible**.
- Complete polynomials, T nonlinear in $\boldsymbol{\xi}$:
  - Introduction of **Kronecker product preconditioner**.

Solving the decoupled Galerkin system

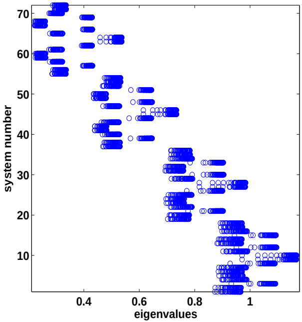
Tensor product polynomials, T linear in $\boldsymbol{\xi}$

Use **Krylov subspace recycling** methods for solving sequence of linear systems $A^{(\ell)}\mathbf{x}^{(\ell)} = \mathbf{b}^{(\ell)}$, where

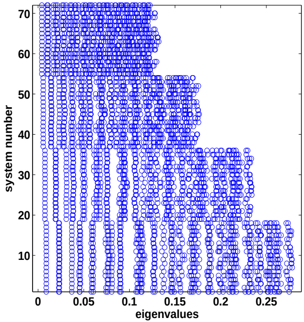
$$A^{(\ell)} = K_0 + \sum_{m=1}^{M} c_{m,\ell} K_m, \quad \ell = 1, \ldots, N_{\boldsymbol{\xi}}.$$

- GCROT-rec., GCRO-DR, R-MINRES [De Sturler et al.]
- success of recycling depends on ordering/grouping of systems and preconditioner
- recycling not useful (= save iteration counts) if preconditioner removes spectrum of system matrix away from zero and results in a strong clustering of eigenvalues

Smallest eigenvalues of preconditioned system matrices



$$P = I_{N_{\boldsymbol{\xi}}} \otimes \texttt{amg}(K_0)$$



$$P = I_{N_{\boldsymbol{\xi}}} \otimes \texttt{cholinc}(K_0,\texttt{'0'})$$

Decoupling the global Galerkin matrix

Task: Find $X \in \mathbb{R}^{N_{\boldsymbol{\xi}} \times N_{\boldsymbol{\xi}}}$ orthogonal, s.t. $\widehat{G}_{\boldsymbol{\alpha}} = X^T G_{\boldsymbol{\alpha}} X$, $|\boldsymbol{\alpha}| > 0$, are diagonal $\Leftrightarrow X^T X = I_{N_{\boldsymbol{\xi}}}$ and $G_{\boldsymbol{\alpha}} X = X \widehat{G}_{\boldsymbol{\alpha}}$, $|\boldsymbol{\alpha}| > 0$.

Theorem (Complete polynomials, T linear in $\boldsymbol{\xi}$)

For $d > 0$ and $M \geq 2$ the stochastic Galerkin matrices $G_{\boldsymbol{\alpha}}$ with $|\boldsymbol{\alpha}| = 1$ are not simultaneously diagonalizable by orthogonal congruence.

Sketch of proof: Let $|\boldsymbol{\alpha}| = 1$.

- For $n \neq m$, $\psi_d^{(n)} \in \mathcal{N}(G_{\boldsymbol{\alpha}})$, $\alpha_m = 1$, but $\psi_d^{(n)}$ is not an eigenfunction of the operator $G_{\boldsymbol{\alpha}}$ where $\alpha_n = 1$.

Mean-based preconditioner P_0

[Ghanem & Kruger, 1996] [Pellissetti & Ghanem, 2000]

Mean stiffness matrix: $[K_0]_{i,k} = \int_D \langle T(\mathbf{x})\rangle \nabla\, \phi_i(\mathbf{x}) \cdot \nabla\, \phi_k(\mathbf{x})\, d\mathbf{x}$.

$$\begin{aligned} A &= I_{N_{\boldsymbol{\xi}}} \otimes K_0 + \sum_{\substack{\boldsymbol{\alpha}\in\mathcal{I}_T,\\ |\boldsymbol{\alpha}|>0}} G_{\boldsymbol{\alpha}} \otimes K_{\boldsymbol{\alpha}} \\ P_0 &= I_{N_{\boldsymbol{\xi}}} \otimes K_0 \end{aligned}$$

- P_0 is symmetric positive definite (s.p.d.)
- $P_0^{-1} = I_{N_{\boldsymbol{\xi}}} \otimes K_0^{-1} \Rightarrow$ solve $N_{\boldsymbol{\xi}}$ systems in $N_{\mathbf{x}}$ unknowns

New approach: Kronecker product preconditioner P_1

Note: P_0 performs no preconditioning w.r.t. the matrices $G_{\boldsymbol{\alpha}}$.

Idea: Construct $P_1 = G \otimes K_0$ s.t. $||A - G \otimes K_0||_F \to \min!$

Solution: [Van Loan & Pitsianis, 1993]

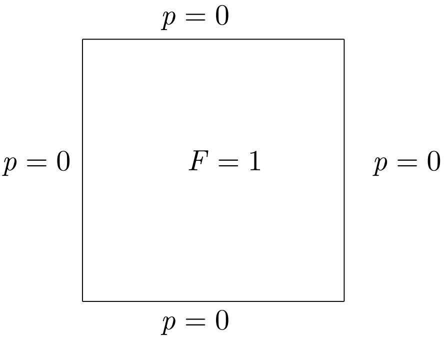
$$G = \sum_{\boldsymbol{\alpha} \in \mathcal{I}_T} \frac{\operatorname{tr}\left(K_0^T K_{\boldsymbol{\alpha}}\right)}{\operatorname{tr}\left(K_0^T K_0\right)} G_{\boldsymbol{\alpha}}$$

Matrix trace: $\operatorname{tr}(A) = \sum_{i=1}^n [A]_{i,i}$

- If A is s.p.d. so is P_1 !
- $P_1^{-1} = G^{-1} \otimes K_0^{-1} \Rightarrow$ solve $N_{\boldsymbol{\xi}}$ systems in $N_{\mathbf{x}}$ unknowns plus $N_{\mathbf{x}}$ systems in $N_{\boldsymbol{\xi}}$ unknowns

Numerical example (1/4)

Stochastic diffusion equation in unit square $D = (0,1) \times (0,1)$.



Deterministic discretization:

- spectral element method
- 6th degree tensor polynomials on 10×10 axis parallel squares
- d.o.f.s: Gauss-Legendre-Lobatto (GLL) nodes on each element
- $N_{\mathbf{x}} = 3,481$

Numerical example (2/4)

Random field model:

- lognormal diffusion $T = \exp(G)$
- underlying Gaussian field G with mean $\langle G \rangle = 1$ and covariance function

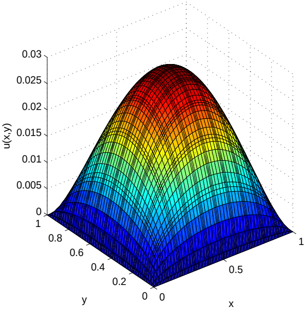
$$\text{Cov}_G(\mathbf{x}, \mathbf{y}) = \sigma^2 \exp\left(-\frac{r^2}{c^2}\right), \quad r = ||\mathbf{x} - \mathbf{y}||_2, \ c = 1.$$

- $M = 4$, captures 98 % of the Gaussian field's total variance
- complete polynomials

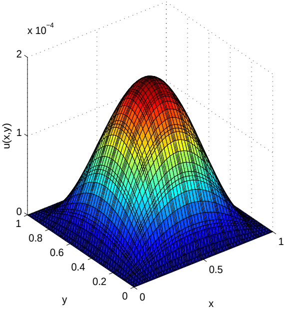
Solver:

- preconditioned conjugate gradient method (PCG)
- action of K_0^{-1} replaced by one AMG V-cycle with K_0
- stopping criterion: $\frac{||\mathbf{r}_k||_2}{||\mathbf{r}_0||_2} < 10^{-8}$

Numerical example (3/4)



mean $\langle p \rangle$



variance $\langle (p - \langle p \rangle)^2 \rangle$

Numerical example (4/4)

$$t_{\boldsymbol{\alpha}}(\mathbf{x}) \quad = \quad \langle T(\mathbf{x})\rangle \frac{\sigma^{|\boldsymbol{\alpha}|}}{\sqrt{\boldsymbol{\alpha}!}} \prod_{m=1}^{M} (\sqrt{\lambda_m} g_m(\mathbf{x}))^{\alpha_m}$$

σ	d=1	2	3	4	5	6
0.1	10	12	13	14	14	15
0.3	14	19	22	26	30	34
0.5	18	27	36	47	58	71
0.7	21	36	56	79	109	144

PCG iteration count P_0

σ	d=1	2	3	4	5	6
0.1	10	10	11	11	11	12
0.3	12	14	16	18	20	22
0.5	14	19	24	29	33	38
0.7	17	25	34	43	54	65

PCG iteration count P_1

Summary

- Tensor product polynomials, T linear in $\boldsymbol{\xi}$:
  - After decoupling solve sequence of $N_{\boldsymbol{\xi}}$ linear systems in $N_{\mathbf{x}}$ unknowns using **Krylov subspace recycling** methods. Success of recycling depends on ordering/grouping of linear systems and preconditioner.
- Complete polynomials, T linear in $\boldsymbol{\xi}$:
  - **Decoupling** of global Galerkin matrix w.r.t. the stochastic d.o.f.s **impossible**.
- Complete polynomials, T nonlinear in $\boldsymbol{\xi}$:
  - Introduction of **Kronecker product preconditioner:** general approach for (linear) SPDEs. Performs favourable compared to popular mean-based preconditioner, few additional costs.