The Computational Complexity of Stochastic Galerkin and Collocation Methods for PDEs with Random Coefficients

Nick Dexter[†],
& Clayton Webster[‡]

[†]Distinguished Graduate Fellow, Dept. of Mathematics, University of Tennessee
[‡]Computational and Applied Mathematics, Oak Ridge National Laboratory

March 8, 2014

Stochastic Finite Element Method

Consider the stochastic elliptic problem defined on $(\Omega, \mathcal{F}, \mathbb{P})$ and $D = [0, b]^2 \subseteq \mathbb{R}^d$, $d = 2$:

$$\begin{cases} -\nabla \cdot (a(\mathbf{x}, \omega) \nabla u(\mathbf{x}, \omega)) = f(\mathbf{x}) & \text{in } \Omega \times D, \\ u(\omega, \mathbf{x}) = 0 & \text{on } \Omega \times \partial D, \end{cases} \tag{1}$$

with deterministic forcing $f(\mathbf{x})$ and $a(\mathbf{x}, \omega) = a(\mathbf{x}, \boldsymbol{y}(\omega))$, $\boldsymbol{y} = (y_1, \ldots, y_N) \in \Gamma$, with $\Gamma_n = y_n(\omega), \Gamma = \Gamma_1 \times \cdots \times \Gamma_N \subseteq \mathbb{R}^N$.
We impose the additional assumptions on $a(\mathbf{x}, \boldsymbol{y})$, that

(A1) $a(\mathbf{x}, \boldsymbol{y}(\omega)) = a_{\min} + h(\mathbf{x}, \boldsymbol{y}(\omega))$ *where the y_j's are independent random variables, and $h : \mathbb{R}^N \times \mathbb{R}^d \to \mathbb{R}$.*

(A2) $\exists\, 0 < a_{\min} \leq a_{\max} < \infty$ *such that*

$$P(a_{\min} \leq a(\mathbf{x}, \boldsymbol{y}(\omega)) \leq a_{\max}) = 1, \qquad \forall \mathbf{x} \in \overline{D}$$

Also, let $\rho(\boldsymbol{y}) = \prod_{n=1}^{N} \rho_n(\boldsymbol{y})$ be the joint density of the vector $\boldsymbol{y}$.

Stochastic Finite Element Method

The weak form of problem (1) is now given by: *find* $u \in H_0^1(D) \otimes L_\rho^2(\Gamma)$ *such that* $\forall v \in H_0^1(D) \otimes L_\rho^2(\Gamma)$

$$\int_\Gamma \int_D a(\mathbf{x},\boldsymbol{y}) \nabla u(\mathbf{x},\boldsymbol{y}) \cdot \nabla v(\mathbf{x},\boldsymbol{y}) d\mathbf{x} \rho(\boldsymbol{y}) d\boldsymbol{y} = \int_\Gamma \int_D f(\mathbf{x}) v(\mathbf{x},\boldsymbol{y}) d\mathbf{x} \rho(\boldsymbol{y}) d\boldsymbol{y}. \tag{2}$$

With some additional assumptions on the smoothness of the data $a(\mathbf{x},\boldsymbol{y})$ it is well known that the solution depends analytically on the parameters $y_n \in \Gamma_n$.

Let $\{\phi_j\}_{j=1}^{J_h}$ be a finite basis for $W_h(D) \subset H_0^1(D)$, and set $J_h = \dim(W_h(D))$. We are interested in the semi-discrete approximation

$$u_{J_h}(\mathbf{x}, \boldsymbol{y}) = \sum_{j=1}^{J_h} u_j(\boldsymbol{y}) \phi_j(\mathbf{x}). \tag{3}$$

given by: find $u_{J_h} \in W_h(D) \bigotimes L^2_\rho(\Gamma)$ such that

$$\int_D a(\mathbf{x}, \boldsymbol{y}) \nabla u_{J_h}(\mathbf{x}, \boldsymbol{y}) \cdot \nabla v_{J_h}(\mathbf{x}) \, \mathrm{d}\mathbf{x} = \int_D f(\mathbf{x}) v_{J_h}(\mathbf{x}) \, \mathrm{d}\mathbf{x} \quad \rho\text{-a.e. in } \Gamma, \tag{4}$$

for all $v_{J_h} \in W_h(D)$. For any $\boldsymbol{y} \in \Gamma$, define

$$\boldsymbol{u}(\boldsymbol{y}) = [u_1(\boldsymbol{y}), u_2(\boldsymbol{y}), \ldots, u_{J_h}(\boldsymbol{y})].$$

Then, the semi-discrete problem (4) can be written algebraically as

$$\boldsymbol{A}(\boldsymbol{y}) \boldsymbol{u}(\boldsymbol{y}) = \boldsymbol{f} \quad \rho\text{-a.e. in } \Gamma. \tag{5}$$

where $\boldsymbol{A}(\boldsymbol{y})$ is the stochastic finite element stiffness matrix.

Stochastic Global Polynomial Subspaces

Let $p \in \mathbb{N}$ denote the polynomial order of an approximation and consider a sequence of increasing, nested multi-index sets $\mathcal{J}(p)$ such that

$$\mathcal{J}(0) = \{(0, \ldots, 0)\} \qquad \text{and} \qquad \mathcal{J}(p) \subseteq \mathcal{J}(p+1).$$

Let $\mathcal{P}_{\mathcal{J}(p)}(\Gamma) \subset L^2_\rho(\Gamma)$ denote the multivariate polynomial space over Γ corresponding to the index set $\mathcal{J}(p)$, defined by

$$\mathcal{P}_{\mathcal{J}(p)}(\Gamma) = \text{span}\left\{ \prod_{n=1}^{N} y_n^{p_n} \,\middle|\, \boldsymbol{p} = (p_1, \ldots, p_N) \in \mathcal{J}(p), y_n \in \Gamma_n \right\}. \tag{6}$$

We set $M_p = \dim \left\{\mathcal{P}_{\mathcal{J}(p)}\right\}$. The fully-discrete global polynomial approximation is now denoted by $u_{J_h M_p} \in W_h(D) \bigotimes \mathcal{P}_{\mathcal{J}(p)}(\Gamma)$.

Examples of $\mathcal{J}(p)$:

- Tensor Products (TP):

$$\mathcal{J}_{TP}(p) = \left\{ \boldsymbol{p} \in \mathbb{N}^N \,\middle|\, \max_n p_n \leq p \right\}, \qquad M_p^{TP} = (p+1)^N$$

- Total Degree (TD):

$$\mathcal{J}_{TD}(p) = \left\{ \boldsymbol{p} \in \mathbb{N}^N \,\middle|\, \sum_{n=1}^{N} p_n \leq p \right\}, \qquad M_p^{TD} = (N+p)!/(N!\,p!),$$

- Hyperbolic Cross (HC):

$$\mathcal{J}_{HC}(p) = \left\{ \boldsymbol{p} \in \mathbb{N}^N \,\middle|\, \sum_{n=1}^{N} \log_2(p_n + 1) \leq \log_2(p+1) \right\}$$

- Sparse Smolyak (SS):

$$\mathcal{J}_{SS}(p) = \left\{ \boldsymbol{p} \in \mathbb{N}^N \,\middle|\, \sum_{n=1}^{N} \gamma(p_n) \leq \gamma(p) \right\}, \quad \gamma(p) = \begin{cases} 0 & \text{for } p = 0 \\ 1 & \text{for } p = 1 \\ \lceil \log_2(p) \rceil & \text{for } p \geq 2. \end{cases}$$

Global stochastic Galerkin methods

Let $\{\psi_{p_n}(y_n)\}_{p=0}^{\infty}$ denote a set of $L^2_{\rho_n}$-orthonormal polynomials in Γ_n. For $\boldsymbol{p} \in \mathcal{J}(p)$, we define

$$\psi_{\boldsymbol{p}}(\boldsymbol{y}) = \prod_{n=1}^{N} \psi_{p_n}(y_n).$$

Then we see that

$$\int_{\Gamma} \psi_{\boldsymbol{p}}(\boldsymbol{y})\psi_{\boldsymbol{p}'}(\boldsymbol{y})\rho(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} = \prod_{n=1}^{N}\int_{\Gamma_n} \psi_{p,n}(y_n)\psi_{p',n}(y_n)\rho_n(y_n)\,\mathrm{d}y_n = \prod_{n=1}^{N}\delta_{p_n p_n'}.$$

Given the bases $\{\phi_j\}_{j=1}^{J_h} \subset W_h(D)$ and $\{\psi_{\boldsymbol{p}}\}_{\boldsymbol{p}\in\mathcal{J}(p)} \subset \mathcal{P}_{\mathcal{J}(p)}(\Gamma)$, the gSGM approximation is defined by

$$u_{J_h M_p}^{gSG}(\mathbf{x},\boldsymbol{y}) = \sum_{\boldsymbol{p}\in\mathcal{J}(p)} u_{\boldsymbol{p}}(\mathbf{x})\psi_{\boldsymbol{p}}(\boldsymbol{y}) = \sum_{\boldsymbol{p}\in\mathcal{J}(p)}\sum_{j=1}^{J_h} u_{\boldsymbol{p},j}\,\phi_j(\mathbf{x})\psi_{\boldsymbol{p}}(\boldsymbol{y}). \tag{7}$$

Our goal is then to solve for the coefficients $\{u_{\boldsymbol{p},j}\}$, $\boldsymbol{p} \in \mathcal{J}(p), j = 1,\ldots,J_h$ which requires the substitution of (7) into the weak formulation (2), resulting in a (possibly nonlinear) coupled system of size $J_h M_p \times J_h M_p$.

Global stochastic Galerkin methods - An Example

- $\boldsymbol{u_p} := [u_{\boldsymbol{p},1}, \ldots, u_{\boldsymbol{p},J_h}]$, the vector of nodal values of the FEM solution corresponding to the $\boldsymbol{p}$-th stochastic mode.
- A Galerkin projection onto the span of $\{\psi_{\boldsymbol{p}}\}_{\boldsymbol{p}\in\mathcal{J}(p)}$ yields the following linear algebraic system: for all $\boldsymbol{p} \in \mathcal{J}(p)$

$$\sum_{\boldsymbol{p}'\in\mathcal{J}(p)} \underbrace{\left(\int_\Gamma \boldsymbol{A}(\boldsymbol{y})\psi_{\boldsymbol{p}}(\boldsymbol{y})\psi_{\boldsymbol{p}'}(\boldsymbol{y})\rho(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} \right)}_{\boldsymbol{K}_{\boldsymbol{p},\boldsymbol{p}'}} \boldsymbol{u}_{\boldsymbol{p}'} = \underbrace{\int_\Gamma \boldsymbol{f}\psi_{\boldsymbol{p}}(\boldsymbol{y})\rho(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}}_{\boldsymbol{F}_{\boldsymbol{p}}}. \qquad (8)$$

1. The coefficient matrix $\boldsymbol{K}$ of the system (8) consists of $(M_p)^2$ block matrices, each of size $J_h \times J_h$, i.e., the size of $\boldsymbol{A}(\boldsymbol{y})$.
2. Even if $\boldsymbol{K}$ is sparse, it is impractical to form and store the matrix explicitly.
3. The structure and sparsity of $\boldsymbol{K}$ depends entirely on $a(\mathbf{x},\boldsymbol{y})$.
4. This approach requires rewriting the Galerkin solver for each new choice of $a(\mathbf{x},\boldsymbol{y})$.

Global stochastic Galerkin methods - NISP

A more convenient and robust choice is to perform an "offline" projection of $a(\mathbf{x},\boldsymbol{y})$ onto span$\{\psi_{\boldsymbol{q}}(\boldsymbol{y})\}_{\boldsymbol{q}\in\mathcal{J}(w)}$, i.e. write $a(\mathbf{x},\boldsymbol{y})$ as

$$a(\mathbf{x},\boldsymbol{y}) = \sum_{n=1}^{\infty} a_n(\boldsymbol{x})\psi_n(\boldsymbol{y}) \approx \sum_{\boldsymbol{q}\in\mathcal{J}(w)} a_{\boldsymbol{q}}(\boldsymbol{x})\psi_{\boldsymbol{q}}(\boldsymbol{y})$$

truncating the expansion on some finite basis. Then for all $\boldsymbol{q} \in \mathcal{J}(w)$,

$$\int_\Gamma \sum_{n=1}^{\infty} a_n(\mathbf{x})\psi_n(\boldsymbol{y})\psi_{\boldsymbol{q}}(\boldsymbol{y})\rho(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} = \sum_{n=1}^{\infty} a_n(\mathbf{x})\delta_{n,\boldsymbol{q}} = \int_\Gamma a(\mathbf{x},\boldsymbol{y})\psi_{\boldsymbol{q}}(\boldsymbol{y})\rho(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}.$$

Letting ϵ_{SG} be the error in SG approximation, we chose w such that $\|a(\mathbf{x},\boldsymbol{y}) - \sum_{\boldsymbol{q}\in\mathcal{J}(w)} a_{\boldsymbol{q}}(\mathbf{x})\psi_{\boldsymbol{q}}(\boldsymbol{y})\|_{L^2} < \epsilon_{SG}$, Substituting the finite expansion of $a(\mathbf{x},\boldsymbol{y})$ yields, for all $j, j' = 1\ldots, J_h$,

$$\boldsymbol{A}_{j,j'}(\boldsymbol{y}) \approx \sum_{\boldsymbol{q}\in\mathcal{J}(w)} \psi_{\boldsymbol{q}}(\boldsymbol{y}) \int_D a_{\boldsymbol{q}}(\mathbf{x})\nabla\phi_j(\mathbf{x})\cdot\nabla\phi_{j'}(\mathbf{x})\,\mathrm{d}\mathbf{x} = \sum_{\boldsymbol{q}\in\mathcal{J}(w)} \psi_{\boldsymbol{q}}(\boldsymbol{y})[\boldsymbol{A}_{\boldsymbol{q}}]_{j,j'},$$

where $[\boldsymbol{A}_{\boldsymbol{q}}]_{j,j'} = \int_D a_{\boldsymbol{q}}(\mathbf{x})\ \nabla\phi_j(\mathbf{x})\cdot\nabla\phi_{j'}(\mathbf{x})\,\mathrm{d}\mathbf{x}$ can be computed component-wise.

Global stochastic Galerkin methods - NISP

Given a sufficiently resolved stochastic finite element stiffness matrix $\boldsymbol{A}(\boldsymbol{y}) \approx \sum_{\boldsymbol{q}\in\mathcal{J}(w)} [\boldsymbol{A}_{\boldsymbol{q}}]\psi_{\boldsymbol{q}}(\boldsymbol{y})$, we substitute $\boldsymbol{A}(\boldsymbol{y})$ into (8) and obtain, for all $\boldsymbol{p}' \in \mathcal{J}(p)$,

$$\sum_{\boldsymbol{p}\in\mathcal{J}(p)} \sum_{\boldsymbol{q}\in\mathcal{J}(w)} \left[\int_\Gamma [\boldsymbol{A}_{\boldsymbol{q}}]\psi_{\boldsymbol{q}}(\boldsymbol{y})\psi_{\boldsymbol{p}'}(\boldsymbol{y})\psi_{\boldsymbol{p}}(\boldsymbol{y})\rho(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\right] \boldsymbol{u}_{\boldsymbol{p}} = \boldsymbol{F}_{\boldsymbol{p}'}. \tag{9}$$

By defining

$$[\boldsymbol{G}_{\boldsymbol{q}}]_{\boldsymbol{p}',\boldsymbol{p}} = \int_\Gamma \psi_{\boldsymbol{q}}\psi_{\boldsymbol{p}'}\psi_{\boldsymbol{p}}\rho\,\mathrm{d}\boldsymbol{y} \quad \text{and} \quad \boldsymbol{K} = \sum_{\boldsymbol{r}\in\mathcal{J}(w)} [\boldsymbol{G}_{\boldsymbol{q}}] \bigotimes [\boldsymbol{A}_{\boldsymbol{q}}], \tag{10}$$

where $[\boldsymbol{G}_{\boldsymbol{q}}] \bigotimes [\boldsymbol{A}_{\boldsymbol{q}}]$ denotes the Kronecker product of $[\boldsymbol{G}_{\boldsymbol{q}}]$ and $[\boldsymbol{A}_{\boldsymbol{q}}]$, we obtain the gSGM coupled system of equations, namely,

$$\boldsymbol{K}\boldsymbol{u} = \boldsymbol{F}, \tag{11}$$

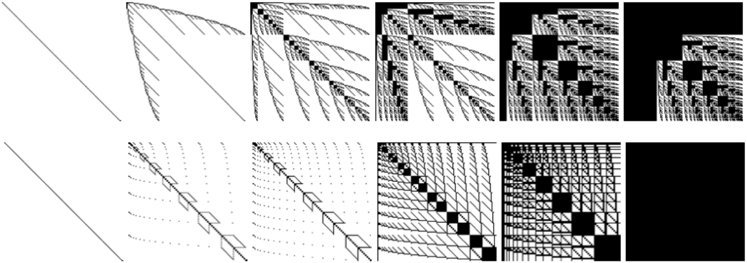
with $\boldsymbol{K}$ symmetric and positive definite.

Global stochastic Galerkin methods - Cost (solve)

Given $\boldsymbol{K}\boldsymbol{u} = \boldsymbol{F}$, where $\boldsymbol{K} = \sum_{\boldsymbol{q}\in\mathcal{J}(w)} [\boldsymbol{G}_{\boldsymbol{q}}] \bigotimes [\boldsymbol{A}_{\boldsymbol{q}}]$, we define

$$N_G = \sum_{\boldsymbol{q}\in\mathcal{J}(w)} \text{number of nonzeros in } [\boldsymbol{G}_{\boldsymbol{q}}], \tag{12}$$

pictorially $N_G = \#$ of black pixels in the matrices



where each pixel represents a block matrix of the size of the original finite element system.

Global stochastic Galerkin methods - Cost (solve)

Given $\boldsymbol{K}\boldsymbol{u} = \boldsymbol{F}$, where $\boldsymbol{K} = \sum_{\boldsymbol{q}\in\mathcal{J}(w)} [\boldsymbol{G_q}] \bigotimes [\boldsymbol{A_q}]$, we define

$$N_G = \sum_{\boldsymbol{q}\in\mathcal{J}(w)} \text{number of nonzeros in } [\boldsymbol{G_q}], \tag{12}$$

then N_G is the total number of nonzeros in the $\{[\boldsymbol{G_q}]\}_{\boldsymbol{q}\in\mathcal{J}(w)}$.

The cost of solving the gSGM method with CG without preconditioning is then given by

$$W_{solve}^{gSGM} \approx N_G * N_{iter}, \tag{13}$$

where N_{iter} is the number of iterations of the system (11) required to converge to a given tolerance in CG. With preconditioning for a block diagonal Jacobi preconditioner, this becomes

$$W_{solve}^{gSGM} \approx (N_G + M_p) * N_{iter}. \tag{14}$$

Global stochastic Collocation methods

1. Choose a set of points $\mathcal{H}_{M_L} = \{\boldsymbol{y}_k \in \Gamma\}_{k=1}^{M_L}$ according to the measure $\rho(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} = \prod_{n=1}^{N} \rho_n(y_n)\,\mathrm{d}y_n$

2. For each k solve the FE solution $u_{J_h}(\mathbf{x}, \boldsymbol{y}_k)$ given $a(\mathbf{x}, \boldsymbol{y}_k)$

3. Interpolate the sampled values: $u_{J_h M_L}^{gSGM}(\mathbf{x}, \boldsymbol{y}) = \sum_{k=1}^{M_L} u_{J_h}(\mathbf{x}, \boldsymbol{y}_k)\mathcal{L}_k(\boldsymbol{y})$, yielding the fully discrete gSCM approximation $u_{J_h M_L}^{gSGM} \in W_h(D) \otimes \mathcal{P}_{\mathcal{J}(L)}(\Gamma)$, where $\mathcal{L}_k \in \mathcal{P}_{\mathcal{J}(L)}(\Gamma)$ are suitable combinations of global (Lagrange) interpolants

$$\mathbb{E}[u](\mathbf{x}) \approx \int_\Gamma u_{M_L}(\mathbf{x}, \boldsymbol{y})\rho(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} = \sum_{k=1}^{M_L} u^h(\mathbf{x}, \boldsymbol{y}_k) \underbrace{\int_\Gamma \mathcal{L}_k(\boldsymbol{y})\rho(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}}_{\text{precomputed weights } w_k}$$

gSCM - Cost (solve)

Therefore, the cost of solving for gSCM is given by

$$W^{gSCM}_{solve} \approx \sum_{k=1}^{M_L} N^{(k)}_{iter},$$

where $N^{(k)}_{iter}$ is the number of iterations required by CG to solve the kth FEM solution $u_{J_h}(\mathbf{x}, \boldsymbol{y}_k)$. Again, the cost of solving the gSGM is

$$W^{gSGM}_{solve} \approx N_G * N_{iter}.$$

Both costs are in terms of total number of matrix vector products required to find the corresponding approximation.

gSCM - Cost (solve)

Therefore, the cost of solving for gSCM is given by

$$W_{solve}^{gSCM} \approx 2 \sum_{k=1}^{M_L} N_{iter}^{(k)},$$

where $N_{iter}^{(k)}$ is the number of iterations required by CG to solve the kth FEM solution $u_{J_h}(\mathbf{x}, \boldsymbol{y}_k)$. Again, the cost of solving the gSGM is

$$W_{solve}^{gSGM} \approx (N_G + M_p) * N_{iter}.$$

Both costs are in terms of total number of matrix vector products required to find the corresponding approximation.

Numerical Example

We now present some results using these methods to compare gSGM and gSCM. Recall the stochastic elliptic problem

$$\begin{cases} -\nabla \cdot (a(\mathbf{x}, \boldsymbol{y}) \nabla u(\mathbf{x}, \omega)) = \cos(x_1) \sin(x_2) & \text{in } \Omega \times D, \\ u(\mathbf{x}, \omega) = 0 & \text{on } \Omega \times \partial D, \end{cases}$$

with $D = [0, b]^2$, and random coefficient $a(\mathbf{x}, \omega)$ with one-dimensional spatial dependence given by

$$\log(a_N(\mathbf{x}, \boldsymbol{y}) - 0.5) = 1 + Y_1(\omega) \left(\frac{\sqrt{\pi} L}{2} \right)^{1/2} + \sum_{n=2}^{N} \zeta_n \varphi_n(\mathbf{x}) Y_n(\omega) \quad (15)$$

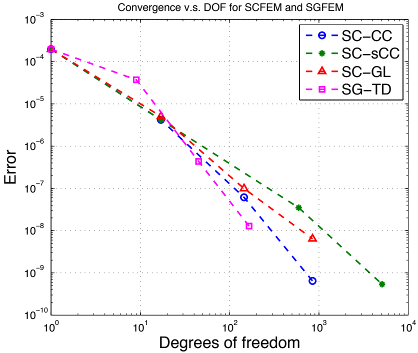
where

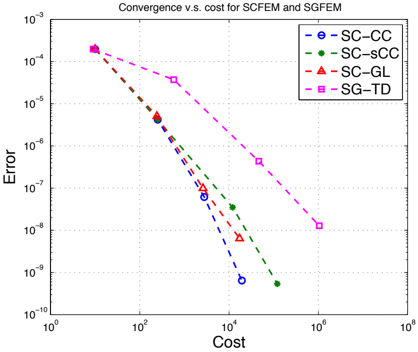
$$\zeta_n := (\sqrt{\pi} L)^{1/2} \exp\left(\frac{-(\lfloor \frac{n}{2} \rfloor \pi L)^2}{8} \right), \text{ if } n > 1 \quad (16)$$

and

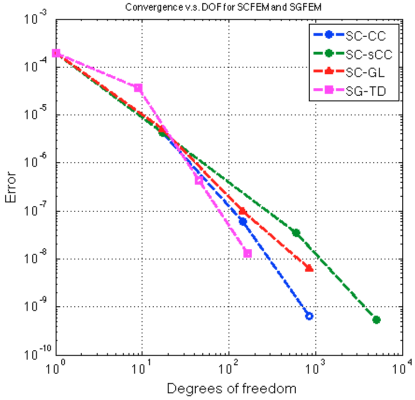
$$\varphi_n(\mathbf{x}) := \begin{cases} \sin\left(\frac{\lfloor \frac{n}{2} \rfloor \pi x_1}{L_p} \right), & \text{if } n \text{ even}, \\ \cos\left(\frac{\lfloor \frac{n}{2} \rfloor \pi x_1}{L_p} \right), & \text{if } n \text{ odd}. \end{cases} \quad (17)$$

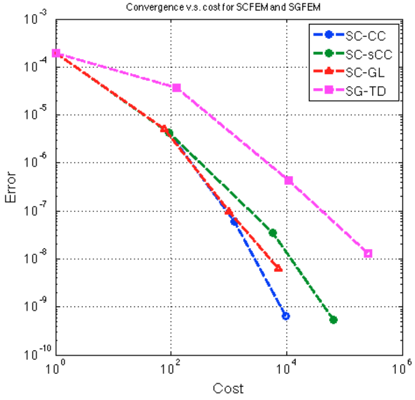
Numerical Results





Numerical Results





Conclusions and Future Work

- Need to apply this cost metric to preconditioning strategies. Cost then depends on the preconditioner used.
- Need to obtain error estimates for the spectral projection. We need complexity to reach desired error estimates.
- Discussion about strategies for preconditioning the Stochastic Collocation Method.

Gautschi, W. (2004).
Orthogonal polynomials: computation and approximation.
Numerical Mathematics and Scientific Computation. Oxford University Press, New York.
Oxford Science Publications.

Ghanem, R. G. and Spanos, P. D. (1991).
Stochastic finite elements: a spectral approach.
Springer-Verlag, New York.

Pellissetti, M. F. and Ghanem, R. G. (2000).
Iterative solution of systems of linear equations arising in the context of stochastic finite elements.
Advances in Engineering Software, 31(8):607–616.