OAK RIDGE National Laboratory
COMPUTATIONAL & APPLIED MATHEMATICS

Uncertainty quantification & approximation theory for parameterized (stochastic) PDEs

Part II: Well-posed SPDEs, regularity, and numerical approximations

Clayton G. Webster[†⋆]

Hoang Tran[⋆], Guannan Zhang[⋆]

[†]Department of Mathematics, University of Tennessee

[⋆]Department of Computational & Applied Mathematics (CAM)
Oak Ridge National Laboratory

Supporting agencies: **DOE** (ASCR, BES), **DOD** (AFOSR, DARPA), **NSF** (CM)

`equinox.ornl.gov` `formulate.ornl.gov` `tasmanian.ornl.gov`

Part II Outline

An overview of methods and algorithms

1. Motivation - parameterized / stochastic equations
2. Brief taxonomy of deterministic and stochastic numerical strategies
3. Monte Carlo FEM (MCFEM)
4. Stochastic Galerkin FEM (SGFEM)
5. Comparisons to SCFEM
6. Computational complexity of solving the SGFEM
7. Numerical illustrations of complexity results

Motivation: Parameterized PDE models

Deterministic and stochastic coefficients

parameters $\boldsymbol{y} \in \mathcal{U} \subset \mathbb{R}^d$ $\longrightarrow$ **PDE model:** $\mathcal{F}(a(y))[u(\boldsymbol{y})] = 0$ in $D \subset \mathbb{R}^n$, $n = 1, 2, 3$ $\longrightarrow$ quantity of interest $Q[u(\boldsymbol{y})]$

- The operator $\mathcal{F}$, linear or nonlinear, depends on a vector of d parameters $\boldsymbol{y} = (y_1, y_2, \ldots, y_d) \in \mathcal{U} = \prod_{i=1}^d \mathcal{U}_i$, which can be deterministic or stochastic.
- **Deterministic setting**: $\boldsymbol{y}$ are known or controlled by the user.
  - **Goal**: a query $\boldsymbol{y} \in \mathcal{U}$, quickly approximation the solution map $\boldsymbol{y} \mapsto u(\boldsymbol{y}) \in \mathcal{V}$.
- **Stochastic setting**: $\boldsymbol{y}$ may be affected by uncertainty and are modeled as a random vector $\boldsymbol{y} : \Omega \to \mathcal{U}$ with joint PDF $\varrho : \mathcal{U} \to \mathbb{R}_+$ s.t. $\varrho(\boldsymbol{y}) = \prod_{i=1}^d \varrho_i(y_i)$.

$$\mathbb{P}\left[Z \in I \subset \mathcal{U}\right] = \int_I \rho(\boldsymbol{y})\, d\boldsymbol{y}, \text{ i.e., transform the measure } \mathbb{P} \text{ to } \mathbb{R}^d$$

  - **Remark:** replace $(\Omega, \mathcal{F}, P)$ with $(\mathcal{U}, \mathcal{B}(\mathcal{U}), \varrho(\boldsymbol{y})d\boldsymbol{y})$, where $\mathcal{B}(\mathcal{U})$ denotes the Borel σ-algebra on $\mathcal{U}$ and $\rho(\boldsymbol{y})d\boldsymbol{y}$ is the distribution measure of $\boldsymbol{y}$.

UQ for parameterized PDE models

Goals of forward UQ

Goal: Approximate u or some statistical QoI depending on u, i.e.

$$\mathbb{E}[u],\ \mathbb{V}\mathrm{ar}[u],\ \mathbb{P}[u > u_0] = \mathbb{E}[\mathbb{1}_{\{u>u_0\}}]$$

with as minimal computational cost as possible.

Quantity of interest (QoI) $Q[u]$, e.g., multi-dimensional expectation

$$\mathbb{E}[u](x) = \int_{\mathcal{U}} u(\boldsymbol{y}, x)\rho(\boldsymbol{y})\, d\boldsymbol{y}, \quad \text{where } \boldsymbol{y} \in \mathcal{U} \text{ and } x \in \overline{D}$$

1. directly approximate $Q[u]$
2. find a surrogate of the solution u (approximating the map $\boldsymbol{y} \mapsto u(\,\cdot\,, \boldsymbol{y})$)
   $\rightarrow$ use the surrogate to cheaply compute any desired quantity of interest

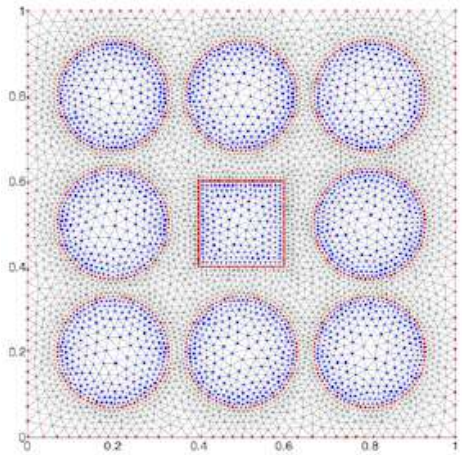
Parameterized PDEs - coefficients

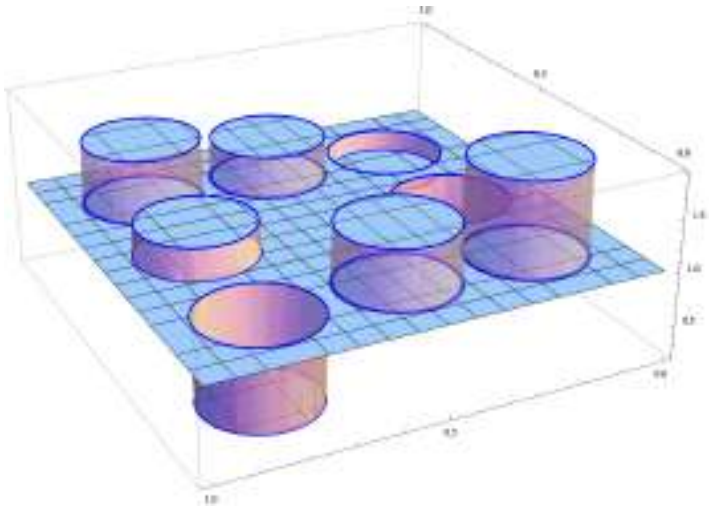
$a(x, \boldsymbol{y})$ satisifies **CC** and **AN**

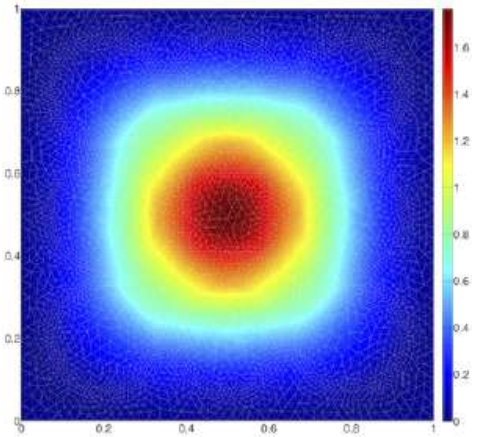
Piecewise constant random fields: Let $\{D_k\}_{k=1}^d$ be a non-overlapping partition of D. We consider

$$a(x, \boldsymbol{y}) = a_0(x) + \sum_{k=1}^{d} \sigma_k y_k \chi_{D_k}(x)$$

where $\sigma_k > 0$ for all k, $a_0(x)$ is large enough to satisfy **(CC)**, and χ_{D_k} is the indicator function of the set D_k.







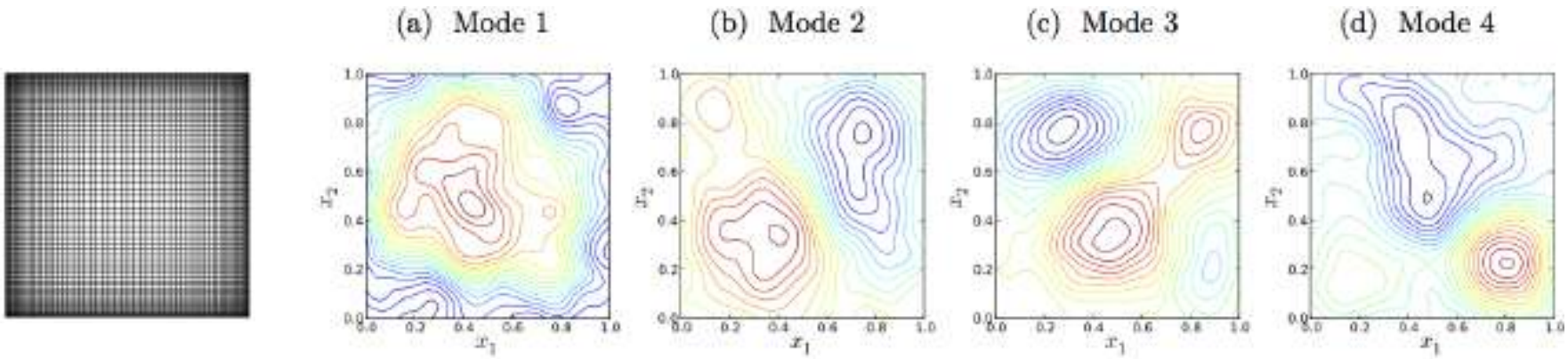
Parameterized PDEs - coefficients

$a(x, \boldsymbol{y})$ satisifies **CC** and **AN**

Karhunen-Loève expansion: a 2^{nd} order random field with continuous covariance function can be represented as an infinite sum of r.v.s via a KL expansion. When the expansion decays quickly, we may truncate

$$a(x, \boldsymbol{y}) \approx \varphi_0(x) + \sum_{k=1}^{d} \varphi_k(x) y_k.$$

Here φ_0 is the mean, $\{\lambda_k = \mathbb{V}ar[y_k], \varphi_k\}_{k=1}^{d}$ are the largest eigenpairs of $\mathbb{C}\text{ov}[a](x_1, x_2)$.



Sargsyan, Safta, Chowdhary, Castorena, de Bord, Debusschere, "UQTk v3.0.1 Manual," SAND2016-9215

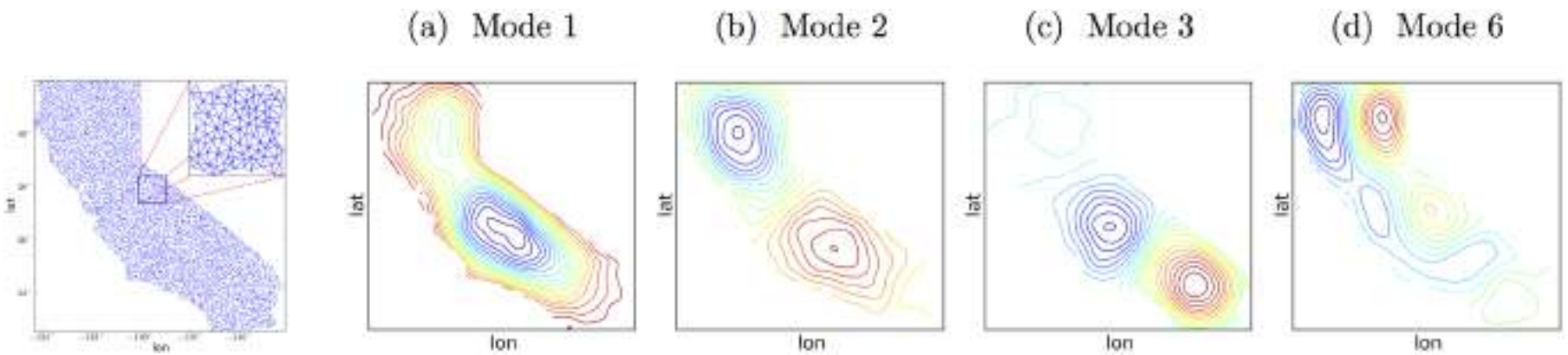
Parameterized PDEs - coefficients

$a(x, \boldsymbol{y})$ satisifies **CC** and **AN**

Karhunen-Loève expansion: a 2^{nd} order random field with continuous covariance function can be represented as an infinite sum of r.v.s via a KL expansion. When the expansion decays quickly, we may truncate

$$a(x, \boldsymbol{y}) \approx \varphi_0(x) + \sum_{k=1}^{d} \varphi_k(x) y_k.$$

Here φ_0 is the mean, $\{\lambda_k = \mathbb{V}ar[y_k], \varphi_k\}_{k=1}^{d}$ are the largest eigenpairs of $\mathbb{C}\text{ov}[a](x_1, x_2)$.

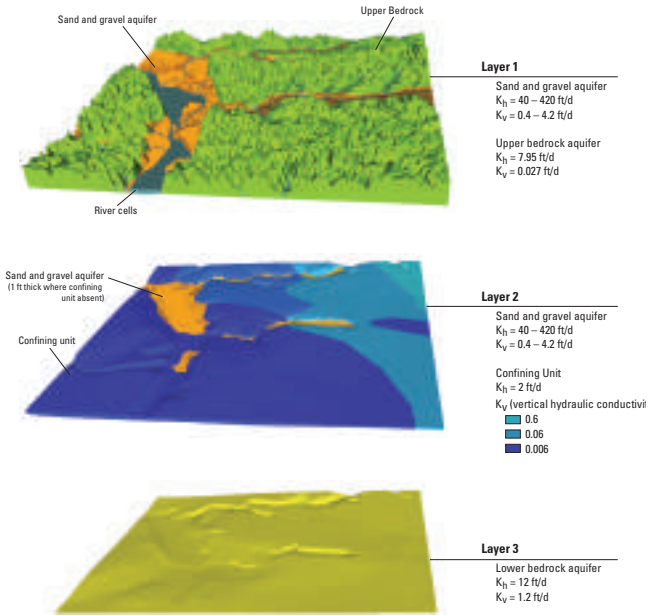


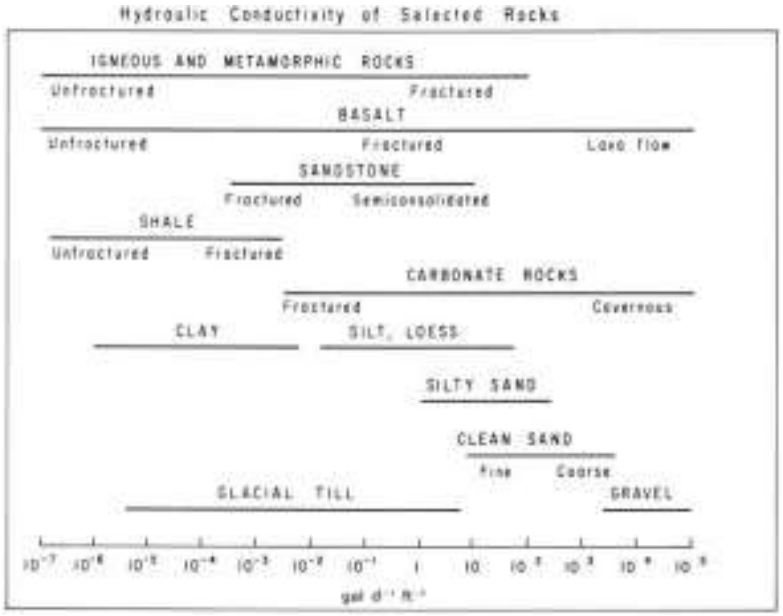
Sargsyan, Safta, Chowdhary, Castorena, de Bord, Debusschere, "UQTk v3.0.1 Manual," SAND2016-9215

Parameterized PDEs - coefficients

$a(x, \boldsymbol{y})$ satisifies **CC** and **AN**

In certain models, it is more appropriate to perform a Karhunen-Loève expansion on the logarithmic scale: for a constant $a_0 > 0$, $\log(a - a_0)(x, \boldsymbol{y}) \approx \varphi_0 + \sum_{k=1}^{d} \varphi_k y_k$.





Hunt, Saad, Chapel, "Numerical Simulation of Ground-Water Flow in La Crosse County, Wisconsin" Report 03-4154 (2003)
Heath, "Basic Ground-Water Hydrology," USGS: Water-Supply Paper 2220 (1983)

UQ for parameterized PDE models

Some assumptions

Continuity and coercivity **(CC)**

For all $x \in \overline{D}$ *and* $\boldsymbol{y} \in \mathcal{U}$, $0 < a_{\min} \leq a(x, \boldsymbol{y}) \leq a_{\max}$.

Analyticity **(AN)**

The complex continuation of a, *represented as the map* $a : \mathbb{C}^d \to L^\infty(D)$, *is an* $L^\infty(D)$*-valued analytic function on* $\mathbb{C}^d$.

Existence and uniqueness of solutions **(EU)**

For all $\boldsymbol{y} \in \mathcal{U}$ *the PDE problem admits an unique solution* $u \in \mathcal{V}$, *where* $\mathcal{V}$ *is a suitable finite or infinite dimensional Hilbert or Banach space. In addition*

$$\forall \boldsymbol{y} \in \mathcal{U}, \quad \exists\, C(\boldsymbol{y}) > 0 \text{ such that } \|u(\boldsymbol{y})\|_{\mathcal{V}} \leq C(\boldsymbol{y})$$

Some simple consequences:

- The PDE induces a map $u = u(\boldsymbol{y}) : \mathcal{U} \to \mathcal{V}$.
- If $\int_{\mathcal{U}} C(\boldsymbol{y})^p \varrho(\boldsymbol{y}) d\boldsymbol{y} < \infty$ then $u \in L^p_\varrho(\mathcal{U}, \mathcal{V})$.

A simple illustrative example

Parameterized elliptic problems: $\mathcal{U} = [-1,1]^d$, $\mathcal{V} = H_0^1(D)$

$$\begin{cases} -\nabla \cdot (a(x,\boldsymbol{y})\nabla u(x,\boldsymbol{y})) &= f(x) \quad x \in D,\ \boldsymbol{y} \in \mathcal{U} \\ u(x,\boldsymbol{y}) &= 0 \qquad x \in \partial D,\ \boldsymbol{y} \in \mathcal{U} \end{cases}$$

Assume $a(x,\boldsymbol{y})$ satisfies **(CC)** and **(AN)**, and that $f \in L^2(D)$, then:

$$\forall y \in \mathcal{U}, \quad u(\boldsymbol{y}) \in H_0^1(D) \equiv \mathcal{V} \quad \text{and} \quad \|u(\boldsymbol{y})\|_{\mathcal{V}} \leq \frac{C_P}{a_{\min}} \|f\|_{L^2(D)}$$

- Lax-Milgram ensures the existence and uniqueness of solution $u \in L^2_\varrho(\mathcal{U},\mathcal{V})$.

Affine and non-affine coefficients:

1. $a(x,\boldsymbol{y}) = a_0(x) + \sum_{i=1}^d y_i \psi_i(x)$.
2. $a(x,\boldsymbol{y}) = a_0(x) + \left(\sum_{i=1}^d y_i \psi_i(x)\right)^q$, $q \in$.
3. $a(x,\boldsymbol{y}) = a_0(x) + \exp\left(\sum_{i=1}^d y_i \psi_i(x)\right)$ (e.g., truncated KL expansion in the log scale).

Remark. In what follows - can be extended to nonlinear elliptic (u^k), parabolic, and some hyperbolic PDEs, all defined on unbounded high-dimensional domains.

A simple illustrative example

Parameterized elliptic weak formulation

The parameterized (stochastic) weak form of problem is given by:

Find $u \in L^2_\varrho(\mathcal{U}; H^1_0(D))$ *such that* $\forall v \in L^2_\varrho(\mathcal{U}; H^1_0(D))$

$$\int_{\mathcal{U}} \mathcal{B}[u,v](\boldsymbol{y})\varrho(\boldsymbol{y})\, d\boldsymbol{y} = \int_{\mathcal{U}} F(v)\varrho(\boldsymbol{y})\, d\boldsymbol{y},$$

where

$$\mathcal{B}[u,v](\boldsymbol{y}) = \int_D a(x,\boldsymbol{y})\nabla u(x,\boldsymbol{y})\cdot\nabla v(x,\boldsymbol{y})dx,$$

and

$$F(v) = \int_D f(x)v(x,\boldsymbol{y})dx.$$

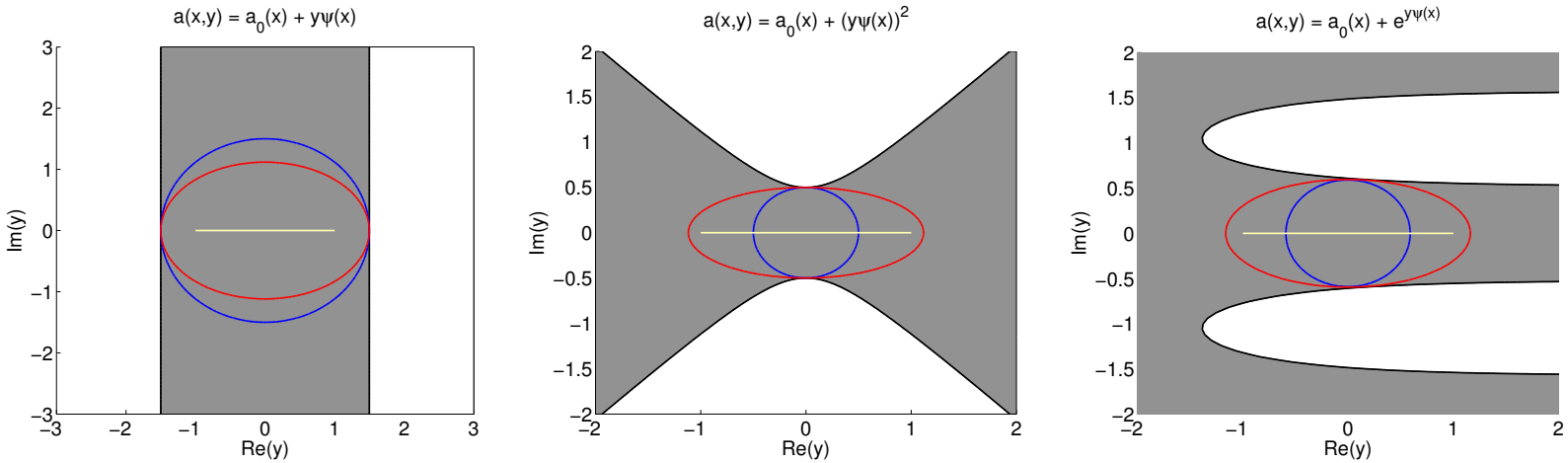
- It follows from **(CC)** that $\mathcal{B}(\boldsymbol{y})$ is a symmetric, uniformly coercive, and continuous bilinear operator on $H^1_0(D)$ for every $\boldsymbol{y} \in \mathcal{U}$.
- Lax-Milgram ensures the existence and uniqueness of solution $u \in L^2_\varrho(\mathcal{U}, H^1_0(D))$.

Analyticity of the solution

$\boldsymbol{\rho} = (\rho_i)_{1\le i\le d},\ \rho_i > 1\ \ \forall i$

- Polydisc: $\mathcal{O}_{\boldsymbol{\rho}} = \bigotimes_i \{z_i \in \mathbb{C}; |z_i| \le \rho_i\}$.
- Polyellipse: $\mathcal{E}_{\boldsymbol{\rho}} = \bigotimes_i \left\{ \frac{z_i + z_i^{-1}}{2}; z_i \in \mathbb{C}, |z_i| = \rho_i \right\}$.

Theorem. [Tran, W., Zhang '16]
Assume $a(x, \boldsymbol{y})$ satisfies **CC** and **AN**. Then the function $\boldsymbol{z} \mapsto u(\boldsymbol{z})$ is well-defined and analytic in an open neighborhood of some polyellipse $\mathcal{E}_{\boldsymbol{\rho}}$ (or polydisc $\mathcal{O}_{\boldsymbol{\rho}}$).



Domain of complex uniform ellipticity for some random fields.

Remark. The high-dimensional discontinuous case is analyzed in:
[Gunzburger, W., Zhang '14], [Burkardt, Gunzburger, W., Zhang '15 (SINUM), '16 (SIREV)]

Brief taxonomy of numerical strategies

General approaches I

There have been many formulations and approaches to solve parameterized deterministic and stochastic PDEs:

1 Statistical sampling methods:

- *Brute-force* Monte Carlo (MC): convergence rate independent of the number of random variables, robust, embarrassingly parallel - very slow convergence
- Quasi MC (QMC), Latin Hypercube Sampling (LHS), Lattice Rules, etc.
- Variance reduction techniques: important, conditional and correlated sampling - limitations when confronted with large number of RVs

2 Indirect methods (require closure approx.):

- Moment methods: derive equations for the moments of the quantities of interest - not applicable to nonlinear problems or non-Gaussian RVs
- PDEs for PDFs (e.g., Fokker-Planck equations): derive a system of PDEs whose solution approximates the probability distributions / densities - boundary conditions and higher dimensions are challenging

Brief taxonomy of numerical strategies

General approaches II

3. **Direct methods:** compute an approximate surrogate to $u(x, \boldsymbol{y})$ in a suitable subspace and use this solution to compute the desired statistics, e.g., stochastic Galerkin (projections), stochastic collocation, etc.

- Interval analysis : maximum bounds of output uncertainty - can dramatically overestimated to uncertainties
- Perturbation-based methods : Taylor expansion around a mean solution - can only be used for linear QoIs and when the variance in solution is small
- Operator-based methods : compute the inverse of a given operator, if it exists, by using a Neumann series expansion or the weighted integral method - restricted to small magnitude uncertainties and often limited to static problems
- Stochastic polynomial approximations: Taylor, Galerkin projections, interpolation and collocation, discrete least squares, and compressed sensing - challenges include: optimal polynomial subspaces, *curse of dimensionality*, adaptive and anisotropic refinement, low stochastic regularity and discontinuities, etc.

Brief taxonomy of numerical strategies

Stochastic FEMs [Gunzburger, W., Zhang, '14 (Acta Numerica)]

- **Monte Carlo methods:** Let $\{\boldsymbol{y}_k \in \mathcal{U}\}_{k=1}^m$ denote a set of random sample points

$$\mathbb{E}[u] = \frac{1}{m}\sum_{k=1}^{m} u(\boldsymbol{y}_k)$$

  - Simple to implement, parallelize, and convergence rate is independent of d.
  - Asymptotic rate is $\mathcal{O}(1/\sqrt{m})$.
  - Unable to simultaneously approximate $\boldsymbol{y} \mapsto u(\boldsymbol{y})$.

- **Polynomial approximations**: Let $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_d) \in \Lambda \subset {}^d$ a multi-index set, and $\boldsymbol{\Psi}_{\boldsymbol{\nu}}$ be multivariate polynomials in $\mathbb{P}_\Lambda(\mathcal{U}) = span\left\{\prod_{i=1}^d {y_i}^{\mu_i}, \mu_i \leq \nu_i \ \forall i\right\}$. Approximate the solution u by:

$$u_\Lambda(x, \boldsymbol{y}) = \sum_{\boldsymbol{\nu}\in\Lambda} c_{\boldsymbol{\nu}}(x)\boldsymbol{\Psi}_{\boldsymbol{\nu}}(\boldsymbol{y}) \in \mathcal{V} \otimes \mathbb{P}_\Lambda(\mathcal{U})$$

  - Takes advantage of the smoothness and/or the sparsity structure of u.
  - Can feature faster convergence than MC.
  - The evaluation of u_Λ requires the computation of $c_{\boldsymbol{\nu}}$ (in possibly) high-dimensions.

Brief taxonomy of numerical strategies

Stochastic FEMs [Gunzburger, W., Zhang, '14 (Acta Numerica)]

- **Monte Carlo methods:** Let $\{\boldsymbol{y}_k \in \mathcal{U}\}_{k=1}^m$ denote a set of random sample points

$$\mathbb{E}[u] = \frac{1}{m}\sum_{k=1}^{m} u(\boldsymbol{y}_k)$$

  - Simple to implement, parallelize, and convergence rate is independent of d.
  - Asymptotic rate is $\mathcal{O}(1/\sqrt{m})$.
  - Unable to simultaneously approximate $\boldsymbol{y} \mapsto u(\boldsymbol{y})$.

- **Polynomial approximations**: Let $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_d) \in \Lambda \subset {}^d$ a multi-index set, and $\boldsymbol{\Psi}_{\boldsymbol{\nu}}$ be multivariate polynomials in $\mathbb{P}_\Lambda(\mathcal{U}) = span\left\{\prod_{i=1}^d {y_i}^{\mu_i}, \mu_i \leq \nu_i \ \forall i\right\}$. Approximate the solution u by:

$$u_\Lambda(x, \boldsymbol{y}) = \sum_{\boldsymbol{\nu}\in\Lambda} c_{\boldsymbol{\nu}}(x)\boldsymbol{\Psi}_{\boldsymbol{\nu}}(\boldsymbol{y}) \in \mathcal{V} \otimes \mathbb{P}_\Lambda(\mathcal{U})$$

  - Takes advantage of the smoothness and/or the sparsity structure of u.
  - Can feature faster convergence than MC.
  - The evaluation of u_Λ requires the computation of $c_{\boldsymbol{\nu}}$ (in possibly) high-dimensions.

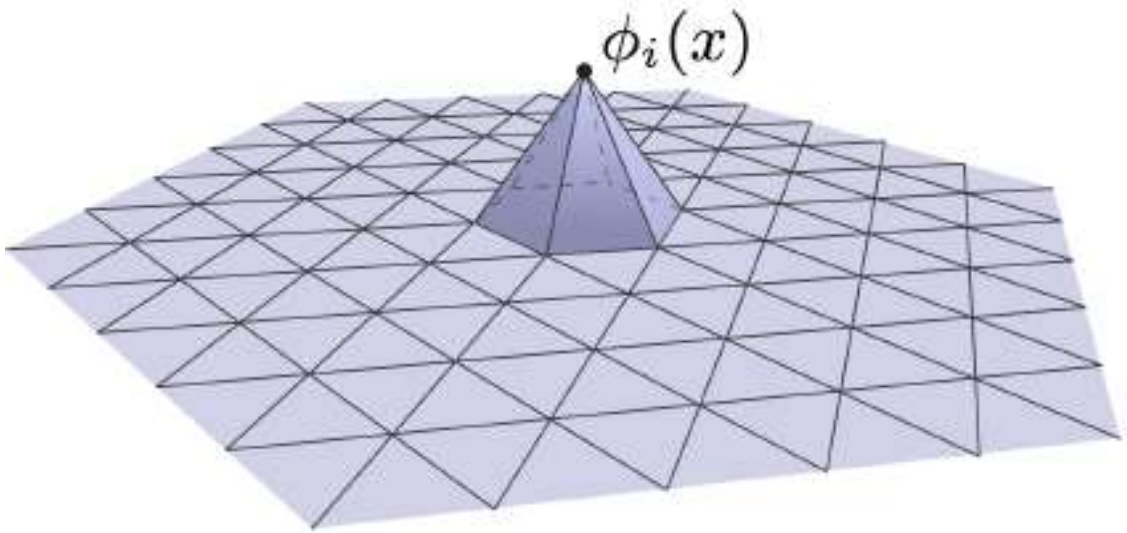
Discretizing the physical domain

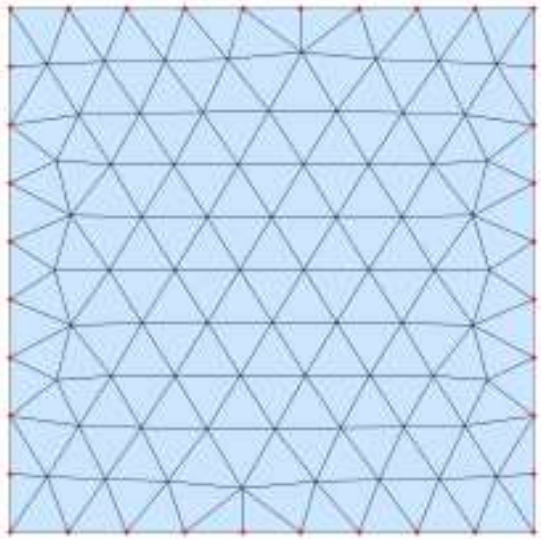
Parameterized finite element approximations

All methods require a further discretization over the physical domain D. Here we rely on the finite element method, though finite differences and finite volume may be used when appropriate. Let

- $\mathcal{T}_h$, triangulation of D
- $u_h(\boldsymbol{y}) \in V_h(D) \subset H_0^1(D)$, finite element space
- $\{\phi_i(x)\}_{i=1}^{J_h}$, piecewise linear polynomial basis for $V_h(D)$ having cardinality J_h

Discretization results in the linear system: $\mathbf{A}(\boldsymbol{y})\boldsymbol{c}(\boldsymbol{y}) = \mathbf{F} \quad \forall \boldsymbol{y} \in \mathcal{U}$.





Schröder, Crane, "Caltech:CS177 Discrete Differential Geometry Course Notes" http://brickisland.net/cs177/?p=309 (2011)

Monte Carlo FEM (MCFEM)

Approximation statistics of QoIs $Q[u]$

1. **Classical approach**: Choose a number of realizations, $m \in \mathbb{N}_+$, and let $\{\boldsymbol{y}_k\}_{k=1}^m$ be a given sample set of random abscissas

2. For each $k = 1, \ldots, m$ sample iid realizations of the diffusion $a(\boldsymbol{y}_k, x)$, the load $f(\boldsymbol{y}_k, x)$ and find a FEM approximation $u_h(\boldsymbol{y}_k, \cdot) \in W_h(D)$ s.t.

$$\begin{cases} -\nabla \cdot (a(\boldsymbol{y}_k, \cdot)\nabla u_h(\boldsymbol{y}_k, \cdot)) = f(\boldsymbol{y}_k, \cdot), & \text{in } D \\ u_h(\boldsymbol{y}_k, \cdot) = 0, & \text{on } \partial D \end{cases}$$

   If desired evaluate the QoI $Q\left(u_h(\boldsymbol{y}_k, \cdot)\right)$

3. Approximate statistics, e.g. expectations $\mathbb{E}[u_h](x)$, by sample averages:

$$\mathbb{E}\left[u_h(\boldsymbol{y})\right](x) \approx \frac{1}{m}\sum_{k=1}^{m} u_h(\boldsymbol{y}_k)\rho(\boldsymbol{y}_k) := \mathscr{E}\left(u_h; m\right), \quad \boldsymbol{y}_k \in \mathcal{U}$$

Goal: Compute, with high probability, sample statistics, e.g.

$$\|\mathbb{E}[u_h] - \mathscr{E}\left(u_h; m\right)\| \leq \text{TOL}$$

Monte Carlo FEM (MCFEM)

Approximation statistics of QoIs $Q[u]$

1. **Classical approach**: Choose a number of realizations, $m \in \mathbb{N}_+$, and let $\{\boldsymbol{y}_k\}_{k=1}^m$ be a given sample set of random abscissas

2. For each $k = 1, \ldots, m$ sample iid realizations of the diffusion $a(\boldsymbol{y}_k, x)$, the load $f(\boldsymbol{y}_k, x)$ and find a FEM approximation $u_h(\boldsymbol{y}_k, \cdot) \in W_h(D)$ s.t.

$$\begin{cases} -\nabla \cdot (a(\boldsymbol{y}_k, \cdot)\nabla u_h(\boldsymbol{y}_k, \cdot)) = & f(\boldsymbol{y}_k, \cdot), & \text{in } D \\ u_h(\boldsymbol{y}_k, \cdot) = & 0, & \text{on } \partial D \end{cases}$$

   If desired evaluate the QoI $Q\left(u_h(\boldsymbol{y}_k, \cdot)\right)$

3. Approximate statistics, e.g. expectations $\mathbb{E}[u_h](x)$, by sample averages:

$$\mathbb{E}\left[u_h(\boldsymbol{y})\right](x) \approx \frac{1}{m}\sum_{k=1}^{m} u_h(\boldsymbol{y}_k)\rho(\boldsymbol{y}_k) := \mathscr{E}\left(u_h; m\right), \quad \boldsymbol{y}_k \in \mathcal{U}$$

Goal: Compute, with high probability, sample statistics, e.g.

$$\|\mathbb{E}[u_h] - \mathscr{E}\left(u_h; m\right)\| \leq \text{TOL}$$

Monte Carlo FEM (MCFEM)

Approximation statistics of QoIs $Q[u]$

1. **Classical approach**: Choose a number of realizations, $m \in \mathbb{N}_+$, and let $\{\boldsymbol{y}_k\}_{k=1}^m$ be a given sample set of random abscissas

2. For each $k = 1, \ldots, m$ sample iid realizations of the diffusion $a(\boldsymbol{y}_k, x)$, the load $f(\boldsymbol{y}_k, x)$ and find a FEM approximation $u_h(\boldsymbol{y}_k, \cdot) \in W_h(D)$ s.t.

$$\begin{cases} -\nabla \cdot (a(\boldsymbol{y}_k, \cdot)\nabla u_h(\boldsymbol{y}_k, \cdot)) = & f(\boldsymbol{y}_k, \cdot), & \text{in } D \\ u_h(\boldsymbol{y}_k, \cdot) = & 0, & \text{on } \partial D \end{cases}$$

   If desired evaluate the QoI $Q\left(u_h(\boldsymbol{y}_k, \cdot)\right)$

3. Approximate statistics, e.g. expectations $\mathbb{E}[u_h](x)$, by sample averages:

$$\mathbb{E}\left[u_h(\boldsymbol{y})\right](x) \approx \frac{1}{m}\sum_{k=1}^{m} u_h(\boldsymbol{y}_k)\rho(\boldsymbol{y}_k) := \mathscr{E}\left(u_h; m\right), \quad \boldsymbol{y}_k \in \mathcal{U}$$

Goal: Compute, with high probability, sample statistics, e.g.

$$\|\mathbb{E}[u_h] - \mathscr{E}\left(u_h; m\right)\| \leq \text{TOL}$$

Monte Carlo FEM (MCFEM)

Approximation statistics of QoIs $Q[u]$

1. **Classical approach**: Choose a number of realizations, $m \in \mathbb{N}_+$, and let $\{\boldsymbol{y}_k\}_{k=1}^m$ be a given sample set of random abscissas

2. For each $k = 1, \ldots, m$ sample iid realizations of the diffusion $a(\boldsymbol{y}_k, x)$, the load $f(\boldsymbol{y}_k, x)$ and find a FEM approximation $u_h(\boldsymbol{y}_k, \cdot) \in W_h(D)$ s.t.

$$\begin{cases} -\nabla \cdot (a(\boldsymbol{y}_k, \cdot)\nabla u_h(\boldsymbol{y}_k, \cdot)) = & f(\boldsymbol{y}_k, \cdot), & \text{in } D \\ u_h(\boldsymbol{y}_k, \cdot) = & 0, & \text{on } \partial D \end{cases}$$

   If desired evaluate the QoI $Q\left(u_h(\boldsymbol{y}_k, \cdot)\right)$

3. Approximate statistics, e.g. expectations $\mathbb{E}[u_h](x)$, by sample averages:

$$\mathbb{E}\left[u_h(\boldsymbol{y})\right](x) \approx \frac{1}{m}\sum_{k=1}^{m} u_h(\boldsymbol{y}_k)\rho(\boldsymbol{y}_k) := \mathscr{E}\left(u_h; m\right), \quad \boldsymbol{y}_k \in \mathcal{U}$$

Goal: Compute, with high probability, sample statistics, e.g.

$$\|\mathbb{E}[u_h] - \mathscr{E}\left(u_h; m\right)\| \leq \text{TOL}$$

Convergence of the MCFEM

Error splitting

$$\mathbb{E}[u] - \mathscr{E}(u_h; m) = \underbrace{\left(\mathbb{E}[u - u_h]\right)}_{\text{Spatial Discret.}} + \underbrace{\left(\mathbb{E}[u_h] - \frac{1}{m}\sum_{k=1}^{m} u_h(\boldsymbol{y}_k)\rho(\boldsymbol{y}_k)\right)}_{\text{Statistical Error}}$$

- **Spatial discretization error**:

$$\|\mathbb{E}[u - u_h]\|_{L^2(D)} + h\,\|\mathbb{E}[u - u_h]\|_{H_0^1(D)} \leq Ch^2\sqrt{\mathbb{E}\left[\|f\|^2_{L^2(D)}\right]}$$

- **Statistical Error**: Within confidence level $\alpha \in (0,1)$, $\exists\, \delta(\alpha) > 0$ s.t.

$$\mathbb{P}\left[\left\|\mathbb{E}[u_h] - \frac{1}{m}\sum_{k=1}^{m} u_h(\boldsymbol{y}_k)\rho(\boldsymbol{y}_k)\right\|_{H_0^1(D)} \leq \delta\frac{C_u}{\sqrt{m}}\right] \geq \alpha$$

$$(M_n)^{\beta}\ \|\mathbb{E}[u_h] - \mathscr{E}(u_h; m)\|_{H_0^1(D)} \to 0\ , n \to \infty \text{ a.s.}$$

for all $\beta \in (0, 1/2)$ with $M_n = 2^d$

Convergence of the MCFEM

Error splitting

$$\mathbb{E}[u] - \mathscr{E}\left(u_h; m\right) = \underbrace{\left(\mathbb{E}\big[u - u_h\big]\right)}_{\text{Spatial Discret.}} + \underbrace{\left(\mathbb{E}[u_h] - \frac{1}{m}\sum_{k=1}^{m} u_h(\boldsymbol{y}_k)\rho(\boldsymbol{y}_k)\right)}_{\text{Statistical Error}}$$

- **Spatial discretization error**:

$$\left\|\mathbb{E}[u - u_h]\right\|_{L^2(D)} + h\left\|\mathbb{E}[u - u_h]\right\|_{H_0^1(D)} \leq Ch^2\sqrt{\mathbb{E}\left[\|f\|^2_{L^2(D)}\right]}$$

- **Statistical Error**: Within confidence level $\alpha \in (0,1)$, $\exists\ \delta(\alpha) > 0$ s.t.

$$\mathbb{P}\left[\left\|\mathbb{E}[u_h] - \frac{1}{m}\sum_{k=1}^{m} u_h(\boldsymbol{y}_k)\rho(\boldsymbol{y}_k)\right\|_{H_0^1(D)} \leq \delta\frac{C_u}{\sqrt{m}}\right] \geq \alpha$$

$$(M_n)^{\beta}\ \left\|\mathbb{E}[u_h] - \mathscr{E}\left(u_h; m\right)\right\|_{H_0^1(D)} \to 0\ , n \to \infty \text{ a.s.}$$

for all $\beta \in (0, 1/2)$ with $M_n = 2^d$

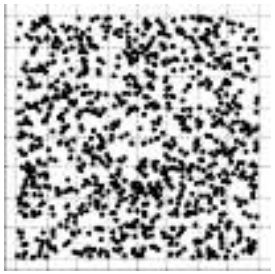
Other sampling-based methods

Attempting to cope with the *curse of dimensionality*

Let $\{\boldsymbol{y}_k\}_{k=1}^m$ be iid samples. Approximate expectations of QoIs by sample averages:

$$\mathbb{E}[Q(u(\boldsymbol{y}))] \approx \frac{1}{m}\sum_{k=1}^{m} Q(u(\boldsymbol{y}_k))\varrho(\boldsymbol{y}_k), \quad \boldsymbol{y}_k \in \mathcal{U}$$

1. Monte Carlo methods: $\varepsilon(m) \approx \mathcal{O}\left(m^{-1/2}\right)$
   abscissas are (pseudo) random numbers
2. Quasi-Monte Carlo methods: $\varepsilon(m) \approx \mathcal{O}\left(m^{-1}(\log(m))^d\right)$
   abscissas are low discrepancy sequences
3. Latin Hypercube Sampling: $\varepsilon(m) \approx \mathcal{O}\left(m^{-1}(\log(m))^d\right)$
   abscissas are chosen to ensure "good" spacing in each 1-D component
4. Lattice rules: $\varepsilon(m) \approx \mathcal{O}\left(m^{-1}(\log(m))^{(d+1)/2}\right)$
   abscissas are "good" lattice points



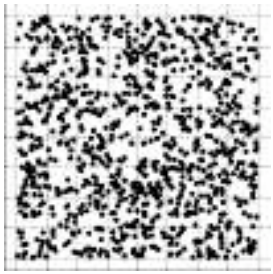
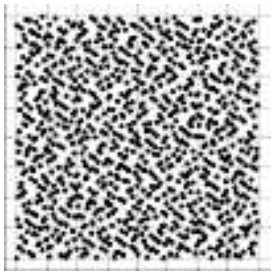
Other sampling-based methods

Attempting to cope with the *curse of dimensionality*

Let $\{\boldsymbol{y}_k\}_{k=1}^m$ be iid samples. Approximate expectations of QoIs by sample averages:

$$\mathbb{E}[Q(u(\boldsymbol{y}))] \approx \frac{1}{m}\sum_{k=1}^{m} Q(u(\boldsymbol{y}_k))\varrho(\boldsymbol{y}_k), \quad \boldsymbol{y}_k \in \mathcal{U}$$

1. Monte Carlo methods: $\varepsilon(m) \approx \mathcal{O}\left(m^{-1/2}\right)$
   abscissas are (pseudo) random numbers
2. Quasi-Monte Carlo methods: $\varepsilon(m) \approx \mathcal{O}\left(m^{-1}(\log(m))^d\right)$
   abscissas are low discrepancy sequences
3. Latin Hypercube Sampling: $\varepsilon(m) \approx \mathcal{O}\left(m^{-1}(\log(m))^d\right)$
   abscissas are chosen to ensure "good" spacing in each 1-D component
4. Lattice rules: $\varepsilon(m) \approx \mathcal{O}\left(m^{-1}(\log(m))^{(d+1)/2}\right)$
   abscissas are "good" lattice points

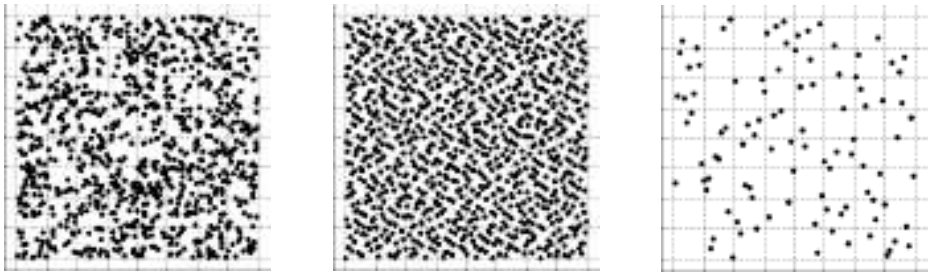
Other sampling-based methods

Attempting to cope with the *curse of dimensionality*

Let $\{\boldsymbol{y}_k\}_{k=1}^m$ be iid samples. Approximate expectations of QoIs by sample averages:

$$\mathbb{E}[Q(u(\boldsymbol{y}))] \approx \frac{1}{m}\sum_{k=1}^{m} Q(u(\boldsymbol{y}_k))\varrho(\boldsymbol{y}_k), \quad \boldsymbol{y}_k \in \mathcal{U}$$

1. Monte Carlo methods: $\varepsilon(m) \approx \mathcal{O}\left(m^{-1/2}\right)$
   abscissas are (pseudo) random numbers
2. Quasi-Monte Carlo methods: $\varepsilon(m) \approx \mathcal{O}\left(m^{-1}(\log(m))^d\right)$
   abscissas are low discrepancy sequences
3. Latin Hypercube Sampling: $\varepsilon(m) \approx \mathcal{O}\left(m^{-1}(\log(m))^d\right)$
   abscissas are chosen to ensure "good" spacing in each 1-D component
4. Lattice rules: $\varepsilon(m) \approx \mathcal{O}\left(m^{-1}(\log(m))^{(d+1)/2}\right)$
   abscissas are "good" lattice points



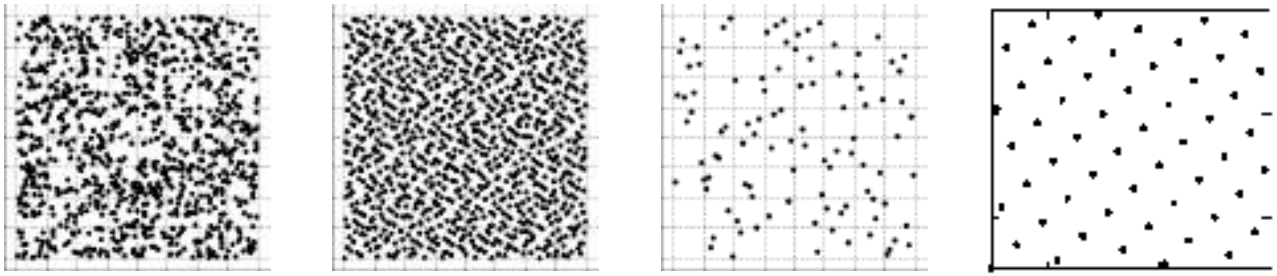
Other sampling-based methods

Attempting to cope with the *curse of dimensionality*

Let $\{\boldsymbol{y}_k\}_{k=1}^m$ be iid samples. Approximate expectations of QoIs by sample averages:

$$\mathbb{E}[Q(u(\boldsymbol{y}))] \approx \frac{1}{m}\sum_{k=1}^{m} Q(u(\boldsymbol{y}_k))\varrho(\boldsymbol{y}_k), \quad \boldsymbol{y}_k \in \mathcal{U}$$

1. Monte Carlo methods: $\varepsilon(m) \approx \mathcal{O}\left(m^{-1/2}\right)$
   abscissas are (pseudo) random numbers
2. Quasi-Monte Carlo methods: $\varepsilon(m) \approx \mathcal{O}\left(m^{-1}(\log(m))^d\right)$
   abscissas are low discrepancy sequences
3. Latin Hypercube Sampling: $\varepsilon(m) \approx \mathcal{O}\left(m^{-1}(\log(m))^d\right)$
   abscissas are chosen to ensure "good" spacing in each 1-D component
4. Lattice rules: $\varepsilon(m) \approx \mathcal{O}\left(m^{-1}(\log(m))^{(d+1)/2}\right)$
   abscissas are "good" lattice points



Other sampling-based methods

Attempting to cope with the *curse of dimensionality*

Let $\{\boldsymbol{y}_k\}_{k=1}^m$ be iid samples. Approximate expectations of QoIs by sample averages:

$$\mathbb{E}[Q(u(\boldsymbol{y}))] \approx \frac{1}{m}\sum_{k=1}^{m} Q(u(\boldsymbol{y}_k))\varrho(\boldsymbol{y}_k), \quad \boldsymbol{y}_k \in \mathcal{U}$$

1. Monte Carlo methods: $\varepsilon(m) \approx \mathcal{O}\left(m^{-1/2}\right)$
   abscissas are (pseudo) random numbers
2. Quasi-Monte Carlo methods: $\varepsilon(m) \approx \mathcal{O}\left(m^{-1}(\log(m))^d\right)$
   abscissas are low discrepancy sequences
3. Latin Hypercube Sampling: $\varepsilon(m) \approx \mathcal{O}\left(m^{-1}(\log(m))^d\right)$
   abscissas are chosen to ensure "good" spacing in each 1-D component
4. Lattice rules: $\varepsilon(m) \approx \mathcal{O}\left(m^{-1}(\log(m))^{(d+1)/2}\right)$
   abscissas are "good" lattice points

Pros: Allow for reusability of deterministic codes and the convergence rate is independent of the regularity of $u(\boldsymbol{y})$ (and dimension with MC methods)

Other sampling-based methods

Attempting to cope with the *curse of dimensionality*

Let $\{\boldsymbol{y}_k\}_{k=1}^m$ be iid samples. Approximate expectations of QoIs by sample averages:

$$\mathbb{E}[Q(u(\boldsymbol{y}))] \approx \frac{1}{m}\sum_{k=1}^{m} Q(u(\boldsymbol{y}_k))\varrho(\boldsymbol{y}_k), \quad \boldsymbol{y}_k \in \mathcal{U}$$

1. Monte Carlo methods: $\varepsilon(m) \approx \mathcal{O}\left(m^{-1/2}\right)$
   abscissas are (pseudo) random numbers
2. Quasi-Monte Carlo methods: $\varepsilon(m) \approx \mathcal{O}\left(m^{-1}(\log(m))^d\right)$
   abscissas are low discrepancy sequences
3. Latin Hypercube Sampling: $\varepsilon(m) \approx \mathcal{O}\left(m^{-1}(\log(m))^d\right)$
   abscissas are chosen to ensure "good" spacing in each 1-D component
4. Lattice rules: $\varepsilon(m) \approx \mathcal{O}\left(m^{-1}(\log(m))^{(d+1)/2}\right)$
   abscissas are "good" lattice points

Cons: The sampling methods do not yield fully discrete approximations and slow convergence rates do not exploit the possible regularity of the functional

Multivariate polynomial approximations

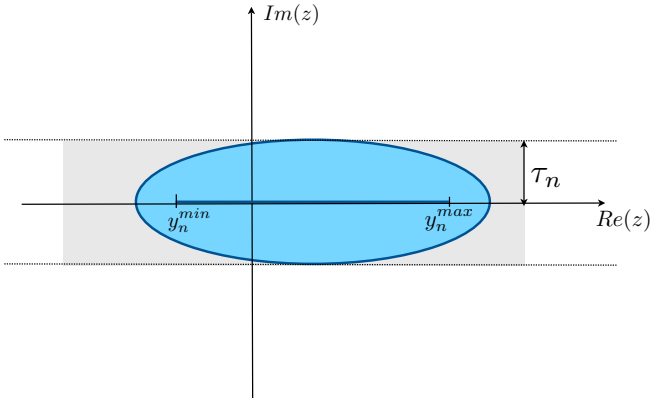
First: $d = 1$-dimensional example with bounded RVs

- Assume a is an exponential Karhunen-Loève expansion and f deterministic: $a(x, \boldsymbol{y}) = a_{min} + \exp[b_0(x) + \sum_{n=1}^{d} b_n(x) y_n]$
- $\mathcal{U}_n$ bounded: $\mathcal{U}_n = \left[y_n^{min}, y_n^{max}\right]$

The analyticity region is given by:

$\mathcal{E}(\mathcal{U}_n; \rho_n) = \{z \in \mathbb{C} : \ |Im(z)| \leq \rho_n\}$,

$$\rho_n = \frac{1}{4\sqrt{\lambda_n}\|b_n\|_{L^\infty(D)}}$$



- Approximate by Chebyshev/Legendre polynomials in y_n yields exponential convergence: *error* $\leq Ce^{-g_n p}$

$$0 < g_n = \log\left[\frac{2\rho_n}{|\mathcal{U}_n|} + \sqrt{1 + \frac{4{\rho_n}^2}{|\mathcal{U}_n|^2}}\right]$$

- **Anisotropic** behavior with respect to the "direction" n
- Similar results for unbounded RVs and various random expansions

Be careful of the *curse*!

Tensor product polynomial approximation in higher dimensions

- The analyticity of the solution $u(\boldsymbol{y})$ w.r.t. each random direction y_n suggests the use of multivariate polynomial approximation.

  what is the correct polynomial approximation subspace?

- The solution must be approximated w.r.t. **all** RV's $y_1, \ldots, y_d \Rightarrow$ possibly high-dimensional problem!

  how do we compute numerical approximations within those subspaces?

- The numerical method must convergence using as few d.o.f.'s as possible

  what is the resulting complexity of my polynomial approximation?

Curse of dimensionality: (Isotropic) TP's of degree p in d dimensions

$$\text{error} \leq Ce^{-gp}, \qquad \#\text{d.o.f. } M = (p+1)^d$$

$$\Updownarrow$$

$$\text{error} \leq Ce^{-gm^{\frac{1}{d}}} \qquad \text{Impractical in higher dimensions}$$

Be careful of the *curse*!

Tensor product polynomial approximation in higher dimensions

- The analyticity of the solution $u(\boldsymbol{y})$ w.r.t. each random direction y_n suggests the use of multivariate polynomial approximation.

  what is the correct polynomial approximation subspace?

- The solution must be approximated w.r.t. **all** RV's $y_1, \ldots, y_d \Rightarrow$ possibly high-dimensional problem!

  how do we compute numerical approximations within those subspaces?

- The numerical method must convergence using as few d.o.f.'s as possible

  what is the resulting complexity of my polynomial approximation?

Curse of dimensionality: (Isotropic) TP's of degree p in d dimensions

$$\text{error} \leq Ce^{-gp}, \qquad \#\text{d.o.f. } M = (p+1)^d$$

$$\Updownarrow$$

$$\text{error} \leq Ce^{-gm^{\frac{1}{d}}} \qquad \text{Impractical in higher dimensions}$$

Multivariate polynomial approximations

Major challenge: *curse of dimensionality*

1. **Taylor approximations:** [Cohen et. al. '10, '11; Tran, W., Zhang '14, '15]
   - $\boldsymbol{\Psi}_{\boldsymbol{\nu}}(\boldsymbol{y}) = \boldsymbol{y}^{\boldsymbol{\nu}}$ and $c_{\boldsymbol{\nu}} = \frac{1}{\boldsymbol{\nu}!}\partial^{\boldsymbol{\nu}} u(\mathbf{0})$ can be computed recursively.
   - useful when ψ_i have non-overlapping supports (affine "inclusion problems")

2. **Galerkin projection methods:** [Wiener '38, Ghanem, Spanos '99; Xiu, Karniadakis '02; Babuška et. al. '02; Todor, Schwab '03; Tran, W., Zhang '14; Dexter, W. '15]
   - $\{\boldsymbol{\Psi}_{\boldsymbol{\nu}}\}$ is a multivariate orthonormal polynomial basis in $\boldsymbol{y}$, e.g., Legendre polynomials, Hermite polynomials, etc.
   - u_Λ is the L^2_ϱ projection of u on $\mathbb{P}_\Lambda(\mathcal{U})$, with $\dim(\mathbb{P}_\Lambda) = \#(\Lambda) \equiv N$.
   - Couples the parametric and physical degrees of freedom.

3. **Interpolation methods:** [Smolyak, '63; Griebel et. al '99,'04; Nobile, Tempone, W. '08a, b; Jantsch, W., Zhang '13, '15; Gunzburger, Jantsch, Teckentrup, W., '15]
   - Given $m \geq \#(\Lambda)$ evaluations $\{u(\boldsymbol{y}_k)\}_{k=1}^m$, and $\{\boldsymbol{\Psi}_{\boldsymbol{\nu}}\}$ a Lagrange basis.
   - u_Λ is the interpolant of u over an associated grid (structured vs. unstructured).
   - Non-intrusive, sample-based approaches. Allow the use of legacy code.
   - May be unstable if the interpolation nodes are poorly chosen (i.e., $m >> \#(\Lambda)$).

Multivariate polynomial approximations

continued...

4. **Discrete least squares:** [Cohen et. al. '13; Migliorati et. al. '13, Narayan et. al. '13; Zhou et. al. '14; Chkifa et. al. '15]

   - Given m evaluations $\{u(\boldsymbol{y}_k)\}_{k=1}^m$, find $(c_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\Lambda}$ by minimizing

$$\sum_{k=1}^{m} \|u(\boldsymbol{y}_k) - u_\Lambda(\boldsymbol{y}_k)\|^2_{\mathcal{V},\ell^2}.$$

   - Mitigate Runge's phenomenon.
   - Reconstruct statistics of u, and stability of the design matrix requires $m >> \#(\Lambda)$.

5. **Compressed sensing:** [Doostan, Owhadi '11; Mathelin, Gallivan '12; Yang, Karniadakis '13; Rauhut, Schwab '14; Adcock '15, '16; Chkifa, Dexter, Tran, W. '15]

   - Given an enriched set Λ_0, and $m << \#(\Lambda_0)$ evaluations $\{u(\boldsymbol{y}_k)\}_{k=1}^m$, find $(c_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\Lambda_0}$ by solving the following minimization problem:

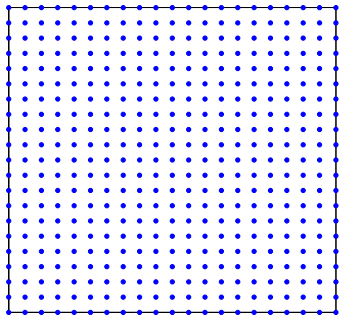
$$\text{argmin } \|\hat{c}_{\boldsymbol{\nu}}\|_{\mathcal{V},\ell^1(\Lambda_0)}, \text{ subject to } u(\boldsymbol{y}_k) = \sum_{\boldsymbol{\nu}\in\Lambda_0} \hat{c}_{\boldsymbol{\nu}}(x)\boldsymbol{\Psi}_{\boldsymbol{\nu}}(\boldsymbol{y}_k).$$

   - Number of samples to recover the best s-term scales linearly in s (up to log factors).
   - ℓ^1 minimization may be impractical in high dimensional problems.

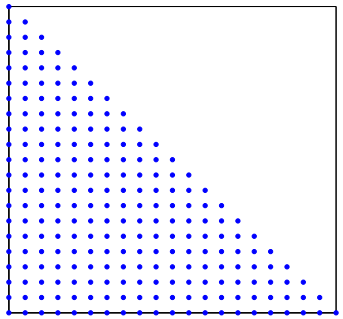
Multivariate polynomial approximations

Selection of (**lower**) index sets in high-dimensions

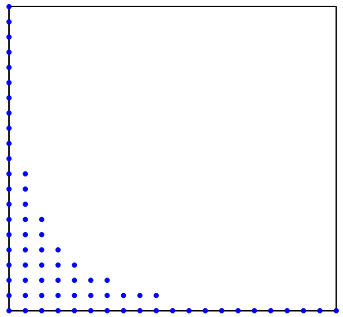
- The efficiency of polynomial approximations depends on the selection of Λ.
- **Standard approaches:** impose index sets Λ *a priori*. The cardinality of the polynomial space $\mathbb{P}_\Lambda(\mathcal{U})$ can grow quickly with respect to the dimension d.
- Some most common choices of index sets Λ



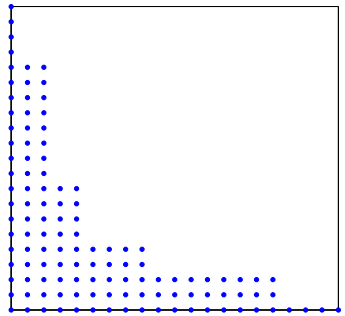
Tensor Product
$\Lambda(w) = \{\nu \in \mathbf{N}^N : \max_{1 \le i \le N} \nu_i \le w\}$



Total Degree
$\Lambda(w) = \{\nu \in \mathbf{N}^N : \sum \nu_i \le w\}$



Hyperbolic Cross
$\Lambda(w) = \{\nu \in \mathbf{N}^N : \prod(\nu_i + 1) \le w + 1\}$



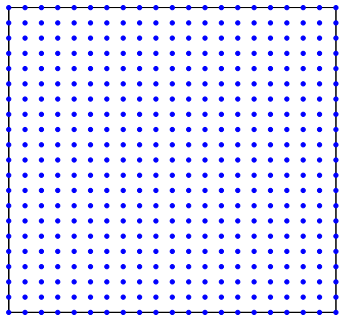
Smolyak
$\Lambda(w) = \{\nu \in \mathbf{N}^N : \sum f(\nu_i) \le f(w)\}$,
with $f(\nu) = \lceil \log_2(\nu) \rceil$, $\nu \ge 2$.

- **Ideally**, the "optimal" $\Lambda \subset \mathbb{N}^d$ has minimal cardinality and enables the approximation of $\boldsymbol{y} \mapsto u(\boldsymbol{y})$ (in high dimensions) with maximum accuracy for a given given computational cost.

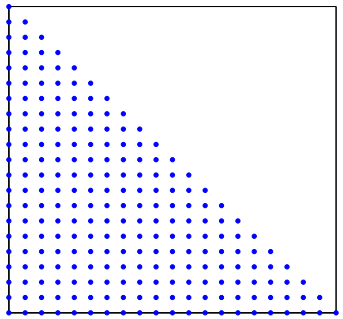
Multivariate polynomial approximations

Selection of (**lower**) index sets in high-dimensions

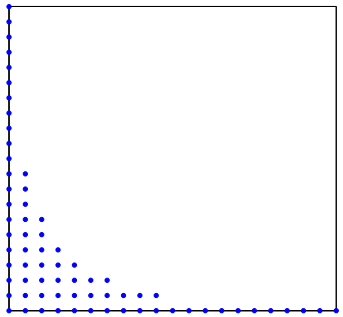
- The efficiency of polynomial approximations depends on the selection of Λ.
- **Standard approaches:** impose index sets Λ *a priori*. The cardinality of the polynomial space $\mathbb{P}_\Lambda(\mathcal{U})$ can grow quickly with respect to the dimension d.
- Some most common choices of index sets Λ



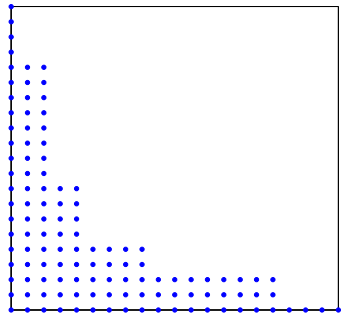
Tensor Product
$\Lambda(w) = \{\nu \in \mathbf{N}^N : \max_{1\le i\le N} \nu_i \le w\}$



Total Degree
$\Lambda(w) = \{\nu \in \mathbf{N}^N : \sum \nu_i \le w\}$



Hyperbolic Cross
$\Lambda(w) = \{\nu \in \mathbf{N}^N : \prod(\nu_i + 1) \le w + 1\}$



Smolyak
$\Lambda(w) = \{\nu \in \mathbf{N}^N : \sum f(\nu_i) \le f(w)\}$,
with $f(\nu) = \lceil \log_2(\nu) \rceil$, $\nu \ge 2$.

- **Ideally**, the "optimal" $\Lambda \subset \mathbb{N}^d$ has minimal cardinality and enables the approximation of $\boldsymbol{y} \mapsto u(\boldsymbol{y})$ (in high dimensions) with maximum accuracy for a given given computational cost.

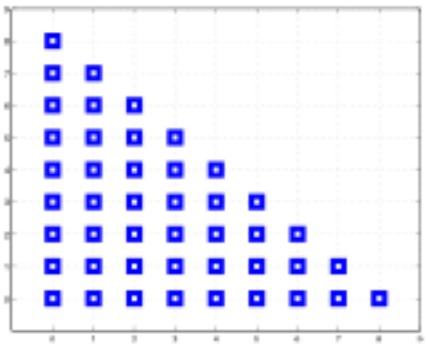
A priori selection of polynomial spaces

Several choices for polynomial multi-index $\boldsymbol{\nu} \in \Lambda_p$:

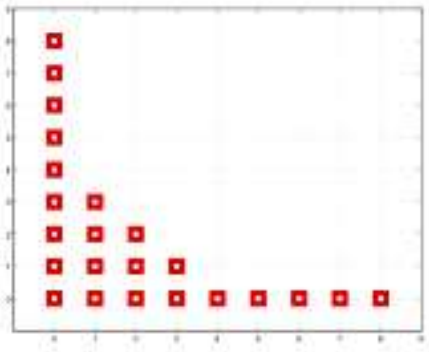
- Tensor products (TP): $\max_n \alpha_n p_n \leq p$ (Intractable for large d),
- Total degree (TD): $\sum_{n=1}^{d} \alpha_n p_n \leq p$,
- Hyperbolic cross (HC): $\prod_{n=1}^{d} (p_n + 1)^{\alpha_n} \leq p + 1$,
- Smolyak method (SM): $\sum_{n=1}^{d} \alpha_n f(p_n) \leq f(p)$ with $f(p) = \begin{cases} 0, & p = 0 \\ 1, & p = 1 \\ \lceil \log_2(p) \rceil, & p \geq 2 \end{cases}$

Anisotropic: introduce weight vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}_+^d$, with $\alpha_{min} = 1$

TD: $\sum_n p_n \leq p$



HC: $\prod_n (p_n + 1) \leq (p + 1)$



TD, HC & SM all reduce the *curse of dimensionality* w.r.t. TP methods.

A priori selection of polynomial spaces

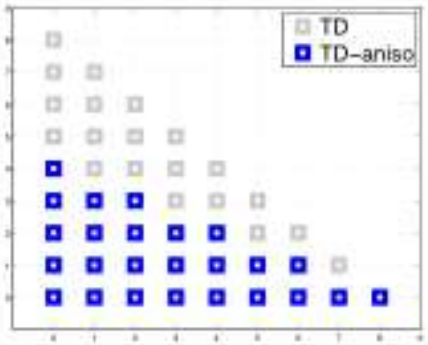
Anisotropic representations

Several choices for polynomial multi-index $\boldsymbol{\nu} \in \Lambda_p$:

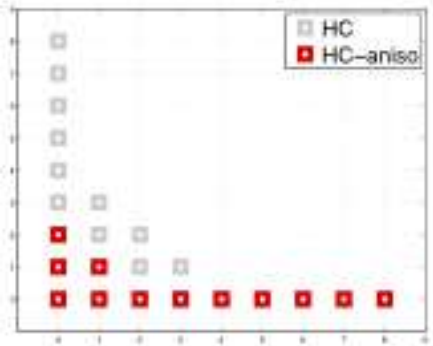
- Tensor products (TP): $\max_n \alpha_n p_n \leq p$ (Intractable for large d),
- Total degree (TD): $\sum_{n=1}^{d} \alpha_n p_n \leq p$,
- Hyperbolic cross (HC): $\prod_{n=1}^{d} (p_n + 1)^{\alpha_n} \leq p + 1$,
- Smolyak method (SM): $\sum_{n=1}^{d} \alpha_n f(p_n) \leq f(p)$ with $f(p) = \begin{cases} 0, & p = 0 \\ 1, & p = 1 \\ \lceil \log_2(p) \rceil, & p \geq 2 \end{cases}$

Anisotropic: introduce weight vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}_+^d$, with $\alpha_{min} = 1$

ATD: $\sum_n \alpha_n p_n \leq p$



AHC: $\prod_n (p_n + 1)^{\alpha_n} \leq p + 1$



TD, HC & SM all reduce the *curse of dimensionality* w.r.t. TP methods.

Example: $d = 2$ with monomial basis

TD space vs. TP space

4th order accurate TD space compared with the TP space:

$\nu_1+\nu_2 \le 0$				1				
$\nu_1+\nu_2 \le 1$			y_1		y_2			
$\nu_1+\nu_2 \le 2$		y_1^2		y_1y_2		y_2^2		
$\nu_1+\nu_2 \le 3$	y_1^3		$y_1^2y_2$		$y_1y_2^2$		y_2^3	
$\nu_1+\nu_2 \le 4$	y_1^4	$y_1^3y_2$		$y_1^2y_2^2$		$y_1y_2^3$		y_2^4
	$y_1^4y_2$		$y_1^3y_2^2$		$y_1^2y_2^3$		$y_1y_2^4$	
		$y_1^4y_2^2$		$y_1^3y_2^3$		$y_1^2y_2^4$		
			$y_1^4y_2^3$		$y_1^3y_2^4$			
$\max(\nu_1,\nu_2) \le 4$				$y_1^4y_2^4$				

Monomials up to 4th degree. Those below the line are the useless monomials we capture (using tensor products) and **are not needed** (and not possible) in higher dimensions - they don't add the asymptotic accuracy and the cost increases exponential as the dimensions increase.

$$m = \dim\left[\mathbb{P}_{\Lambda_p}(\mathcal{U})\right] \Longrightarrow m_{\text{TD}} = \frac{(d+p)!}{d!p!} << m_{\text{TP}} = (p+1)^d$$

Example: $d = 2$ with monomial basis

TD space vs. TP space

OAK RIDGE National Laboratory
COMPUTATIONAL & APPLIED MATHEMATICS

4th order accurate TD space compared with the TP space:

$\nu_1+\nu_2 \le 0$					1				
$\nu_1+\nu_2 \le 1$				y_1		y_2			
$\nu_1+\nu_2 \le 2$			y_1^2		y_1y_2		y_2^2		
$\nu_1+\nu_2 \le 3$		y_1^3		$y_1^2y_2$		$y_1y_2^2$		y_2^3	
$\nu_1+\nu_2 \le 4$	y_1^4		$y_1^3y_2$		$y_1^2y_2^2$		$y_1y_2^3$		y_2^4
		$y_1^4y_2$		$y_1^3y_2^2$		$y_1^2y_2^3$		$y_1y_2^4$	
			$y_1^4y_2^2$		$y_1^3y_2^3$		$y_1^2y_2^4$		
				$y_1^4y_2^3$		$y_1^3y_2^4$			
$\max(\nu_1,\nu_2) \le 4$					$y_1^4y_2^4$				

Monomials up to 4th degree. Those below the line are the useless monomials we capture (using tensor products) and **are not needed** (and not possible) in higher dimensions - they don't add the asymptotic accuracy and the cost increases exponential as the dimensions increase.

$$m = \dim\left[\mathbb{P}_{\Lambda_p}(\mathcal{U})\right] \Longrightarrow m_{\text{TD}} = \frac{(d+p)!}{d!p!} << m_{\text{TP}} = (p+1)^d$$

General basis in d dimensions

Total degree versus Tensor products

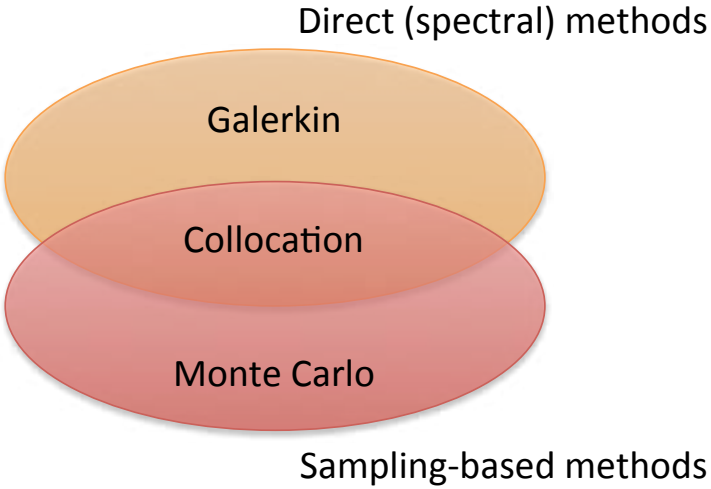
$d =$ # RVs, $\dim(\mathcal{U})$	$p =$ maximal degree of polynomials	$m =$ total # of probabilistic degrees of freedom	
		using total degree basis	using tensor product basis
3	3	20	64
	5	56	216
5	3	56	$1,024$
	5	252	$7,776$
10	3	286	$1,048,576$
	5	$3,003$	$60,046,176$
20	3	$1,771$	$> 1 \times 10^{12}$
	5	$53{,}130$	$> 3 \times 10^{15}$
100	3	$176,851$	$> 1 \times 10^{60}$
	5	$96,560,646$	$> 6 \times 10^{77}$

- tensor products become computational infeasible in higher dimensions

A brief taxonomy of methods

For numerical solution of parameterized stochastic PDEs input

Stochastic finite element methods (SFEMs)



- methods for which spatial discretization is effected using finite element methods (FEMs)[†]
- **Stochastic sampling methods (SSMs)**: random samples in $\mathcal{U}$ of PDE inputs are used to compute ensemble averages of statistical QoIs, e.g. MCFEM - *non-intrusive*

Stochastic polynomial approximation

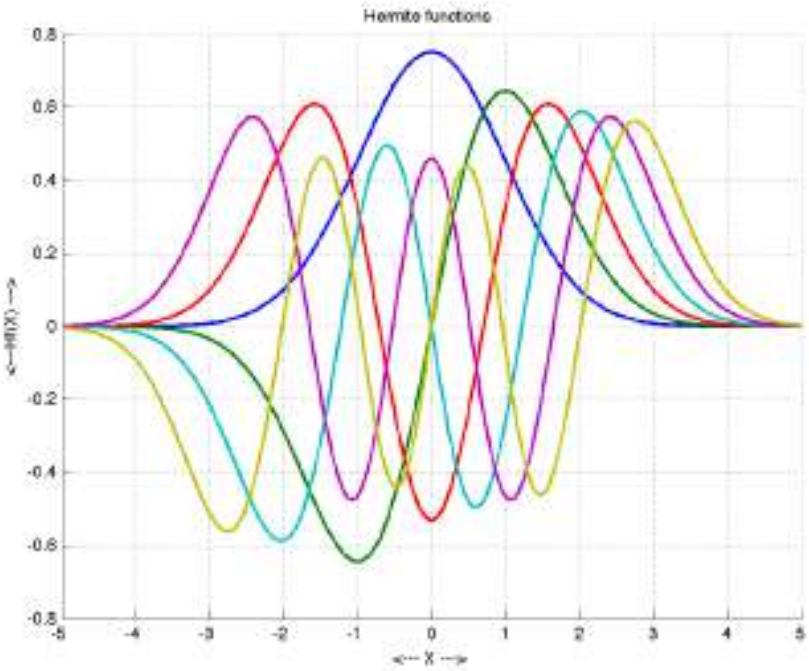
1. Stochastic Galerkin methods (SGMs): probabilistic discretization is also effected by a spectral Galerkin projection onto, e.g., an L^2_ρ-orthogonal basis (Wiener or polynomial chaos) - *fully intrusive*
2. Stochastic Collocation methods (SCMs): probabilistic discretization is effected by collocating the FE solution on a particular set of of points and then connect the realizations with suitable interpolatory basis (Lagrangian) - *non-intrusive*

Stochastic Galerkin FEM (SGFEM)

Motivation: univariate Hermite polynomials

Motivation: The L^2_ϱ-orthogonal basis was originally proposed to approximate white noise processes with Gaussian measure [Wiener, 1938].

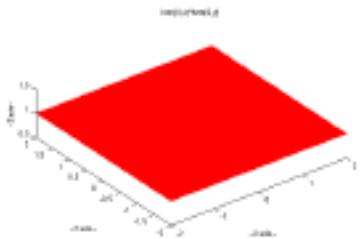
- the univariate Hermite polynomials $H(y)$ serve as the foundation for the construction of the multi-dimensional Hermite polynomials - orthogonal with respect to the Gaussian measure



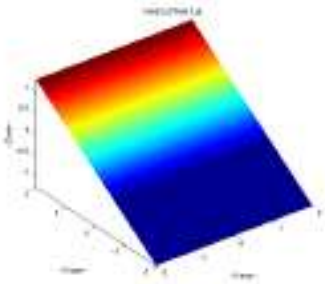
The PDF of a Gaussian RV is $\varrho(y) = \frac{1}{\sqrt{2\pi}} e^{\frac{-y^2}{2}}$

$(d = 2, p = 5)$ Hermite polynomials

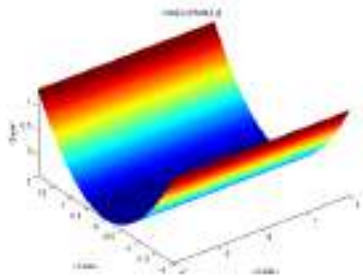
TD subspace: $\nu_1 + \nu_2 \leq 5$



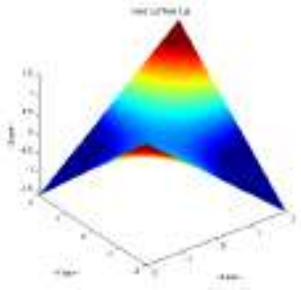
$H_{(0,0)}$



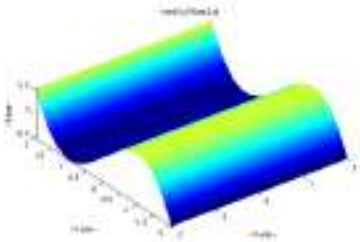
$H_{(0,1)}, H_{(1,0)}$



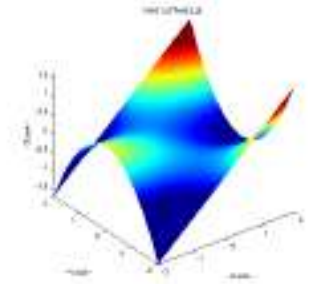
$H_{(0,2)}, H_{(2,0)}$



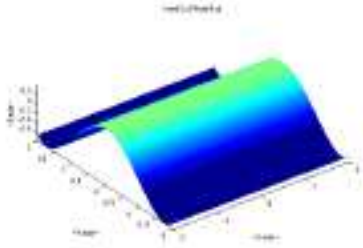
$H_{(1,1)}$



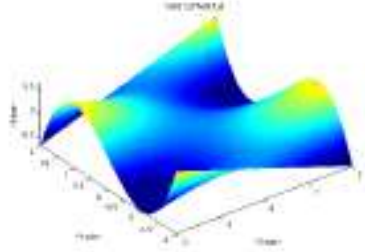
$H_{(0,3)}, H_{(3,0)}$



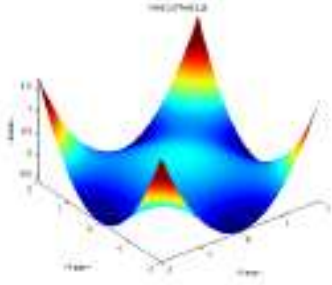
$H_{(1,2)}, H_{(2,1)}$



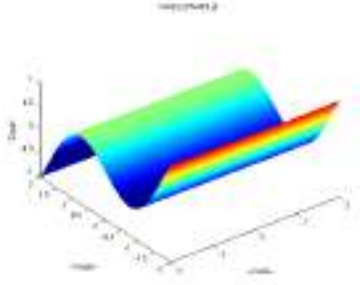
$H_{(0,4)}, H_{(4,0)}$



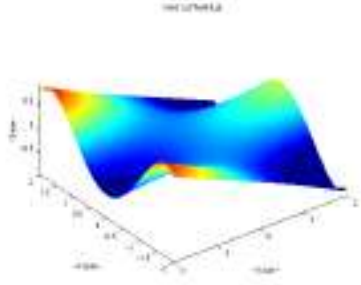
$H_{(1,3)}, H_{(3,1)}$



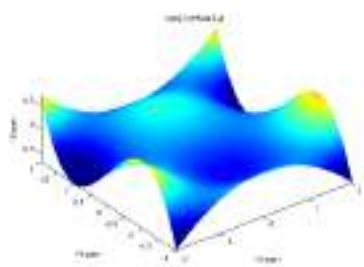
$H_{(2,2)}$



$H_{(0,5)}, H_{(5,0)}$



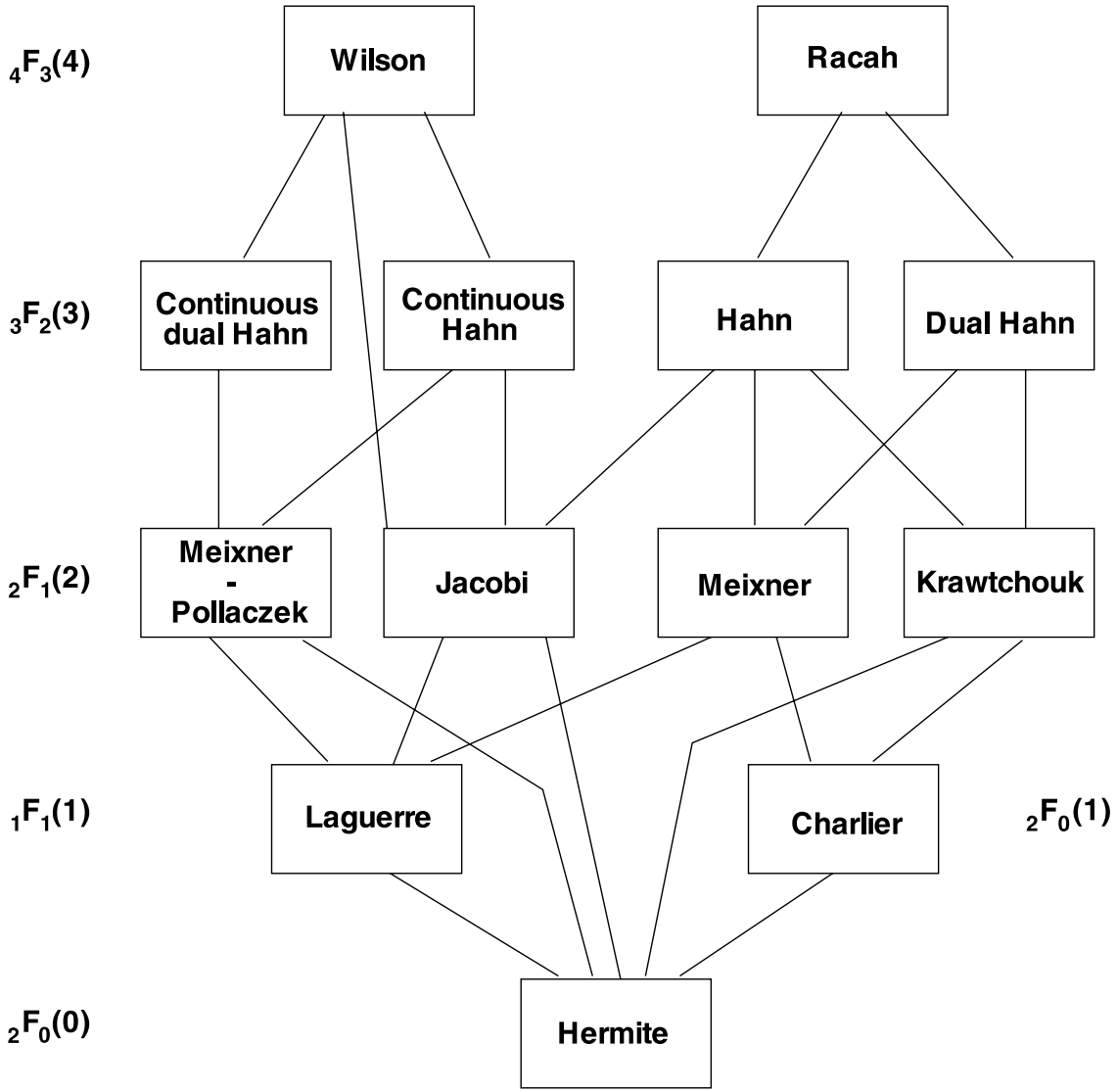
$H_{(1,4)}, H_{(4,1)}$



$H_{(2,3)}, H_{(3,2)}$

The Askey scheme

Classification of hypergeometric orthogonal polynomials



The Askey scheme

Connections between PDF and the orthogonal polynomials

Distribution	Density function	Polynomial	Support
Normal	$\frac{1}{\sqrt{2\pi}} e^{\frac{-y^2}{2}}$	Hermite $H_n(y)$	$[-\infty, \infty]$
Uniform	$\frac{1}{2}$	Legendre $P_n(y)$	$[-1, 1]$
Beta	$\frac{(1-y)^\alpha (1+y)^\beta}{2^{\alpha+\beta+1} B(\alpha+1, \beta+1)}$	Jacobi $P_n^{(\alpha,\beta)}(y)$	$[-1, 1]$
Exponential	e^{-y}	Laguerre $L_n(y)$	$[0, \infty]$
Gamma	$\frac{y^\alpha e^{-y}}{\Gamma(\alpha+1)}$	Generalized Laguerre $L_n^{(\alpha)}(y)$	$[0, \infty]$

Stochastic Galerkin FEM (SGFEM)

Orthogonal projection, implementation and solving challenges, and convergence rate

Let $\{\Psi_{\boldsymbol{p}}\}_{\boldsymbol{p}\in\Lambda_p}$ be a global orthonormal basis w.r.t. ϱ, then the Galerkin projection onto $\text{span}\{\Psi_{\boldsymbol{p}}\}_{\boldsymbol{p}\in\Lambda_p}$ yields the **coupled** system of equations:

$$\sum_{\boldsymbol{q}\in\Lambda_p} \underbrace{\langle \Psi_{\boldsymbol{p}}(\boldsymbol{y})\mathbf{A}(\boldsymbol{y})\Psi_{\boldsymbol{q}}(\boldsymbol{y})\rangle_{\varrho}}_{=:\,\mathbf{K}_{\boldsymbol{p},\boldsymbol{q}}\in\mathbb{R}^{J_h\times J_h}} \boldsymbol{c}_{\boldsymbol{q}} = \underbrace{\langle\, \mathbf{F}, \Psi_{\boldsymbol{p}}(\boldsymbol{y})\rangle_{\varrho}}_{=:\,\mathbf{F}\delta_{\mathbf{0},\boldsymbol{p}}\in\mathbb{R}^{J_h}} \qquad \forall \boldsymbol{p}\in\Lambda_p$$

- Too large to store and solve directly, $\mathbf{K}\in\mathbb{R}^{N_p J_h\times N_p J_h}$ for $N_p=\#(\Lambda_p)$
- Number of nonzero blocks $\mathbf{K}_{\boldsymbol{p},\boldsymbol{q}}$ depends on $a(x,\boldsymbol{y})$, can be fully **block-dense**
- Computing the entries of $\mathbf{K}$ may require computing a $d\times n$ dimensional integral

Convergence is sub-exponential w.r.t. to the stochastic discretization.

Proposition (spectral convergence). [Todor, Schwab '07]
When the map $\boldsymbol{z}\mapsto u(\boldsymbol{z})$ is analytic, and Λ_p total degree then:

$$\|u-u_{\Lambda_p}\|_{L^\infty_\varrho(\mathcal{U};H^1_0(D))}\leq C_1\exp(-C_2 p)\quad \forall p\in\mathbb{N},$$

for some constants $C_1, C_2>0$ depending only on a, f, d.

Stochastic Galerkin method

Orthogonal expansion of the coefficient $a(x, \boldsymbol{y})$

When $a(x, \boldsymbol{y})$ is not affine it is advantageous to approximate, i.e.,

$$a^r(x, \boldsymbol{y}) := \sum_{\boldsymbol{k} \in \Lambda_r} a_{\boldsymbol{k}}(x) \Psi_{\boldsymbol{k}}(\boldsymbol{y}) \to a(x, \boldsymbol{y}) \quad \text{sub-exponentially as } r \to \infty.$$

Substituting $a^r(x, \boldsymbol{y})$ into the Galerkin equations for $a(x, \boldsymbol{y})$ we obtain

$$\sum_{\boldsymbol{q} \in \Lambda_p} [\mathbf{K}_r]_{\boldsymbol{p},\boldsymbol{q}} \boldsymbol{c}_{\boldsymbol{q}}^r = \mathbf{F} \delta_{\mathbf{0},\boldsymbol{p}} \qquad \forall \boldsymbol{p} \in \Lambda_p$$

where $\mathbf{K}_r = \sum_{\boldsymbol{k} \in \Lambda_r} \mathbf{G}_{\boldsymbol{k}} \otimes \mathbf{A}_{\boldsymbol{k}}$ has Kronecker product structure, and

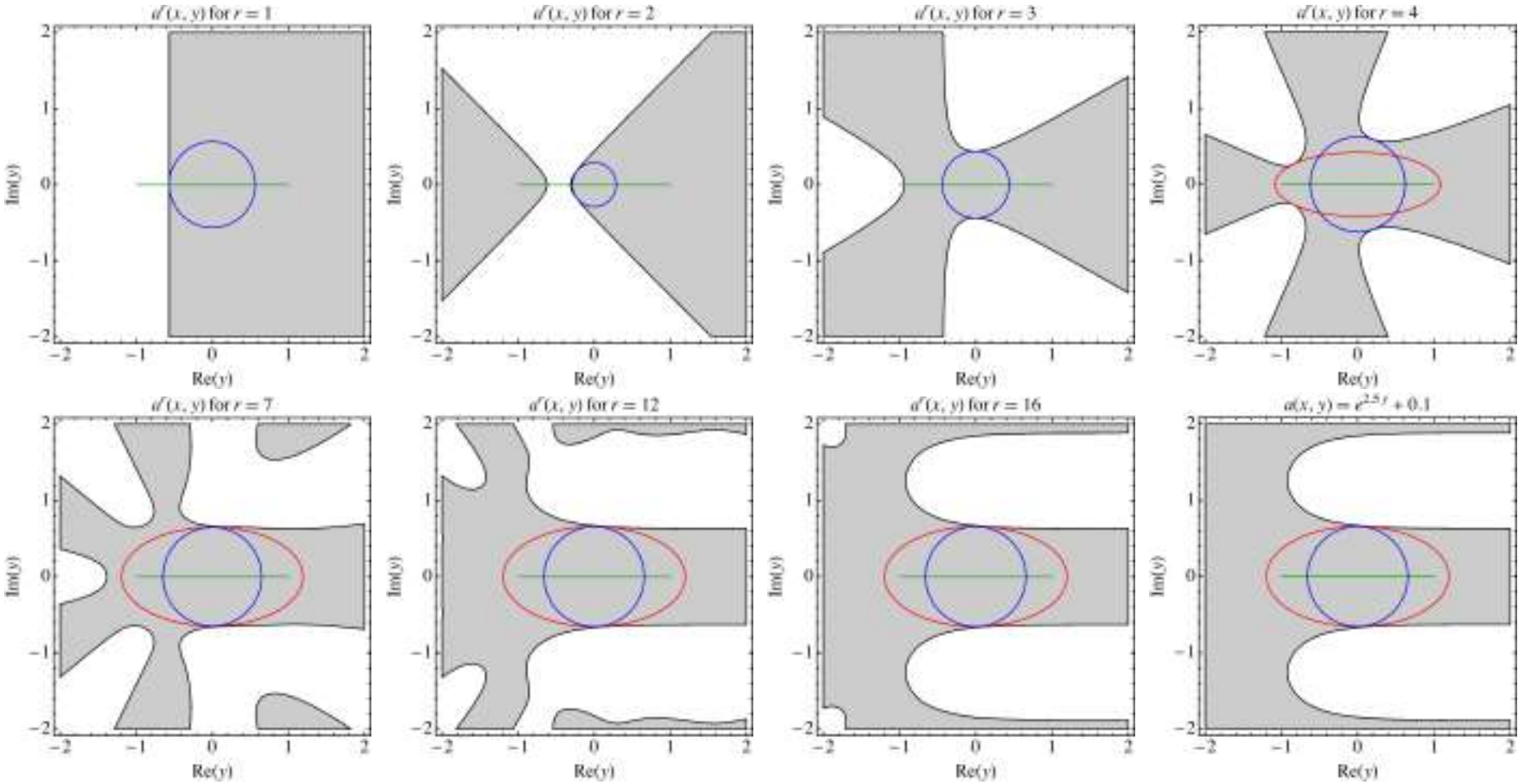
$$[\mathbf{G}_{\boldsymbol{k}}]_{\boldsymbol{p},\boldsymbol{q}} = \langle \Psi_{\boldsymbol{k}} \Psi_{\boldsymbol{p}} \Psi_{\boldsymbol{q}} \rangle_\varrho \qquad \text{and} \qquad [\mathbf{A}_{\boldsymbol{k}}]_{i,j} = \int_D a_{\boldsymbol{k}}(x) \nabla \phi_j(x) \cdot \nabla \phi_i(x) dx.$$

- $\mathbf{K}_r$ approximates the full Galerkin system $\mathbf{K}$ when $a(x, \boldsymbol{y}) \notin \mathcal{P}_{\Lambda_r}(\mathcal{U})$
- Letting $r = 2p$ yields $\mathbf{K}_r = \mathbf{K}$ due to orthogonality of $\{\Psi_{\boldsymbol{p}}\}$ [Matthies, Keese '05]
- Only need to store $\{\mathbf{G}_{\boldsymbol{k}}\} \in \mathbb{R}^{N_p \times N_p}$ and $\{\mathbf{A}_{\boldsymbol{k}}\} \in \mathbb{R}^{J_h \times J_h}$
- Allows control over the cost of solving by varying r, since

$$\|u - u^r\|_{L^2_\varrho(\mathcal{U}; H^1_0(D))} \leq \frac{\|f\|_{H^{-1}(D)}}{a^2_{\min}} \|a - a^r\|_{L^2_\varrho(\mathcal{U}; L^\infty(D))}$$

Stochastic Galerkin method

Orthogonal expansion of $a(x, \boldsymbol{y})$ and well-posedness



Must be careful when choosing projection order to assure that the problem is well-posed. One way to guarantee this is to choose $\tilde{r} \leq r \leq 2p$, where

$$\tilde{r} := \min\{r \in \mathbb{N}_0 : \|a - a^{\nu}\|_{L^{\infty}_{\varrho}(\mathcal{U};L^{\infty}(D))} \leq a_{\min},\ \forall \nu \in \mathbb{N}_0,\ \nu \geq r\}.$$

Comparison to interpolation methods

Stochastic collocation FEM: general setting

1. Choose a set of points $\{\boldsymbol{y}_k \in \Gamma\}_{k=1}^{m_p}$ according to the measure $\varrho(\boldsymbol{y})d(\boldsymbol{y}) = \prod_{n=1}^{d} \varrho_n(y_n)d(y_n)$.
2. For each k solve the FE solution $u_k(x) = u(\boldsymbol{y}_k, x)$, given $a_k(x) = a(\boldsymbol{y}_k, x)$ and $f_k(x) = f(\boldsymbol{y}_k, x)$.
3. Interpolate the sampled values:

$$\mathcal{I}_{\Lambda_p}[u] = \sum_{k=1}^{m_p} u_k(x)\ell_k(\boldsymbol{y}) \in \mathbb{P}_{\Lambda_p}(\mathcal{U}) \otimes \mathcal{V}_h,$$

yielding the fully discrete SC approximation in, where $\ell_k \in \mathbb{P}_{\Lambda_p}(\mathcal{U})$ are suitable combinations of **global** (Lagrange) interpolants.

Compute a quantity of interest, e.g., $\mathbb{E}[u](x)$

$$\mathbb{E}[u](x) \approx \int_\Gamma \mathcal{I}_{\Lambda_p}[u](\cdot, \boldsymbol{y})\varrho(\boldsymbol{y})d\boldsymbol{y} = \sum_{k=1}^{m_p} u_k(x) \underbrace{\int_\Gamma \ell_k(\boldsymbol{y})\varrho(\boldsymbol{y})dy}_{\text{precomputed weights}} = \sum_{k=1}^{m_p} u_k(x)w_k$$

Comparison to (interpolatory) SCFEM

Optimality of the SGFEM - in terms of DOFs

Given a fixed multi-index set Λ_p, let u_{Λ_p} be the SGFEM on $\mathbb{P}_{\Lambda_p}(\mathcal{U})$, then:

$$\|u - u_{\Lambda_p}\|_{L^2_\varrho(\mathcal{U};H^1_0(D))} \lesssim C_a \min_{v\in H^1_0(D)\otimes\mathbb{P}_{\Lambda_p}(\mathcal{U})} \|u - v\|_{L^2_\varrho(\mathcal{U};H^1_0(D))}.$$

We can construct an interpolation operator $\mathcal{I}_{\Lambda_p} : C^0(\mathcal{U}) \to \mathbb{P}_{\Lambda_p}(\mathcal{U})$ for which

$$\begin{aligned}\|u - \mathcal{I}_{\Lambda_p}[u]\|_{L^\infty_\varrho(\mathcal{U};H^1_0(D))} &\leq (\mathbb{L}_{\Lambda_p} + 1) \min_{v\in H^1_0(D)\otimes\mathbb{P}_{\Lambda_p}(\mathcal{U})} \|u - v\|_{L^\infty_\varrho(\mathcal{U};H^1_0(D))} \\ &\lesssim (\mathbb{L}_{\Lambda_p} + 1)\|u - u_{\Lambda_p}\|_{L^\infty_\varrho(\mathcal{U};H^1_0(D))},\end{aligned}$$

where $\mathbb{L}_{\Lambda_p} = \|\mathcal{I}_{\Lambda_p}\|_{L^\infty_\varrho(\mathcal{U})\to L^\infty_\varrho(\mathcal{U})}$ is the Lebesgue constant of $\mathcal{I}_{\Lambda_p}$.

Recall:

- $m \geq \#(\Lambda_p)$ in the construction of $\mathcal{I}_{\Lambda_p}$
- Implies the optimality of the Galerkin projection in terms of the stochastic degrees of freedom $\#(\Lambda_p)$
- How do they relate in terms of computational complexity or stability?

Comparison to (interpolatory) SCFEM

Optimality of the SGFEM - in terms of DOFs

Given a fixed multi-index set Λ_p, let u_{Λ_p} be the SGFEM on $\mathbb{P}_{\Lambda_p}(\mathcal{U})$, then:

$$\|u - u_{\Lambda_p}\|_{L^2_\varrho(\mathcal{U};H^1_0(D))} \lesssim C_a \min_{v\in H^1_0(D)\otimes\mathbb{P}_{\Lambda_p}(\mathcal{U})} \|u - v\|_{L^2_\varrho(\mathcal{U};H^1_0(D))}.$$

We can construct an interpolation operator $\mathcal{I}_{\Lambda_p} : C^0(\mathcal{U}) \to \mathbb{P}_{\Lambda_p}(\mathcal{U})$ for which

$$\begin{aligned}\|u - \mathcal{I}_{\Lambda_p}[u]\|_{L^\infty_\varrho(\mathcal{U};H^1_0(D))} &\le (\mathbb{L}_{\Lambda_p} + 1) \min_{v\in H^1_0(D)\otimes\mathbb{P}_{\Lambda_p}(\mathcal{U})} \|u - v\|_{L^\infty_\varrho(\mathcal{U};H^1_0(D))} \\ &\lesssim (\mathbb{L}_{\Lambda_p} + 1)\|u - u_{\Lambda_p}\|_{L^\infty_\varrho(\mathcal{U};H^1_0(D))},\end{aligned}$$

where $\mathbb{L}_{\Lambda_p} = \|\mathcal{I}_{\Lambda_p}\|_{L^\infty_\varrho(\mathcal{U})\to L^\infty_\varrho(\mathcal{U})}$ is the Lebesgue constant of $\mathcal{I}_{\Lambda_p}$.

Recall:

- $m \ge \#(\Lambda_p)$ in the construction of $\mathcal{I}_{\Lambda_p}$
- Implies the optimality of the Galerkin projection in terms of the stochastic degrees of freedom $\#(\Lambda_p)$
- How do they relate in terms of computational complexity or stability?

Computational complexity of solving the SGFEM

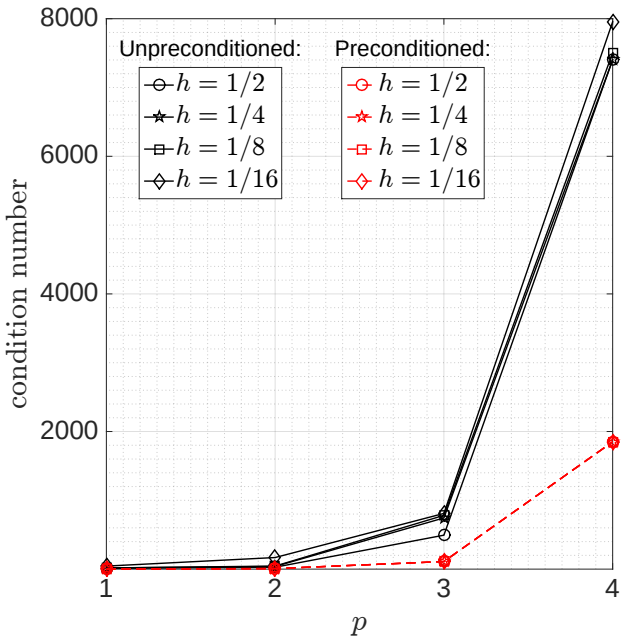
Iterative solvers and conditioning challenges

$\mathbf{K}_r$ is symmetric positive-definite, hence preconditioned conjugate gradients can be used to solve and we have the error at the k-th step:

$$\|\boldsymbol{c}^r - \boldsymbol{c}^{r,(k)}\|_{\mathbf{K}_r} \leq 2\left(\frac{\sqrt{\boldsymbol{\kappa}_r}-1}{\sqrt{\boldsymbol{\kappa}_r}+1}\right)^k \|\boldsymbol{c}^r - \boldsymbol{c}^{r,(0)}\|_{\mathbf{K}_r}. \tag{1}$$

- $\boldsymbol{c}^{r,(0)}$ and $\boldsymbol{c}^{r,(k)}$ are the initial guess and the output at the k-th iteration, resp.
- $\boldsymbol{\kappa}_r$ is the condition number of $\mathbf{K}_r$, depends on h, p, and r

We use a simple preconditioner: $\mathbf{P} := \mathbf{G_0} \otimes \mathbf{A_0}$, (mean-based block-diagonal)



- $\mathbf{G_0} = \mathbf{I} \in \mathbb{R}^{N_p \times N_p}$ for orthonormal $\{\Psi_{\boldsymbol{p}}\}$
- Easy to invert since $\mathbf{P}^{-1} = \mathbf{G_0}^{-1} \otimes \mathbf{A_0}^{-1}$
- Complexity of applying is $\mathcal{O}(J_h) * N_p$ when incomplete Cholesky is used
- Removes dependence of h in $\boldsymbol{\kappa}_r$, condition number of preconditioned system still depends on r and p
- Better preconditioners are available, but are more challenging to implement and analyze cost

Computational complexity of solving the SGFEM

Given $\mathbf{K}_r = \sum_{\boldsymbol{k}\in\Lambda_r} \mathbf{G}_{\boldsymbol{k}} \otimes \mathbf{A}_{\boldsymbol{k}}$, we define

$$\mathcal{M}(p,r) = \sum_{\boldsymbol{k}\in\Lambda_r} \text{nnz}(\mathbf{G}_{\boldsymbol{k}}) = \#\left\{(\boldsymbol{k},\boldsymbol{p},\boldsymbol{q}) : \langle \Psi_{\boldsymbol{k}}\Psi_{\boldsymbol{p}}\Psi_{\boldsymbol{q}}\rangle_\varrho \neq 0, \boldsymbol{k}\in\Lambda_r, \boldsymbol{p},\boldsymbol{q}\in\Lambda_p\right\}$$

Pictorially, $\mathcal{M}(p,r) = \#$ of black pixels in the matrices:

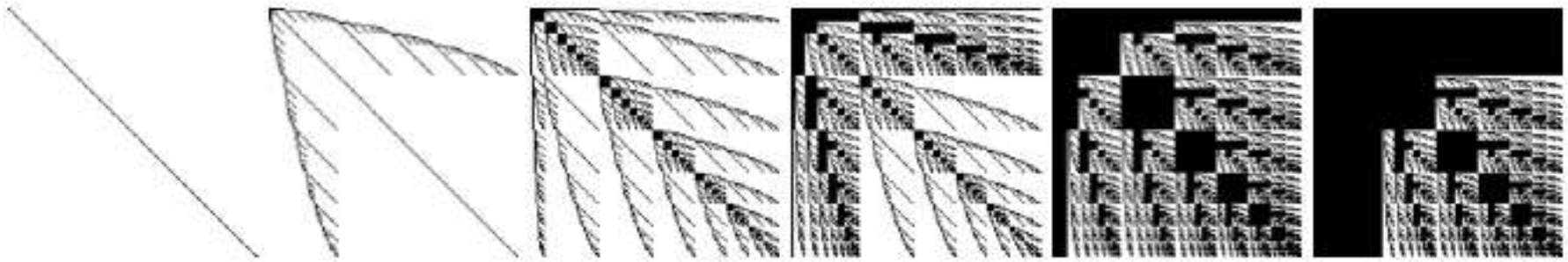


Figure Block sparsity of $\mathbf{K}_r$ for fixed $p = 3$, increasing $r = 0, 1, 2, 3, 4, 5$.

Each CG iteration requires $\mathcal{M}(p,r)$ matrix-vector products of complexity $\mathcal{O}(J_h)$, where $\mathcal{O}(J_h)$ depends on the connectivity of $\mathcal{T}_h$, hence

- cost without preconditioning: $W^{\text{SG}} \approx \mathcal{O}(J_h) * \mathcal{M}(p,r) * N^{\text{SG}}_{\text{iter}}$
- cost with preconditioner $\mathbf{P}$: $W^{\text{pSG}} \approx \mathcal{O}(J_h) * (N_p + \mathcal{M}(p,r)) * N^{\text{pSG}}_{\text{iter}}$.

$N^{\text{pSG}}_{\text{iter}}$, $N^{\text{SG}}_{\text{iter}}$ are number of CG iterations required to converge to a given tolerance with and without preconditioning, respectively.

Basic unit of cost is in terms FLOPS.

SGFEM - How to count $\mathcal{M}(p,r)$?

Theorem ((sparsity of $\mathbf{G}_{\boldsymbol{k}}$) [Dexter, W. Zhang '16].

Let $d, p, r \in \mathbb{N}$, $d \geq 1$, $0 \leq r \leq 2p$, $\boldsymbol{k} \in \Lambda_r$, and ϱ_i be even $\forall i$. Then

$$\text{nnz}(\mathbf{G}_{\boldsymbol{k}}) = \sum_{\ell=\lceil |\boldsymbol{k}|/2 \rceil}^{|\boldsymbol{k}|} c(\boldsymbol{k},\ell) \binom{d+p-\ell}{p-\ell},$$

where

$$c(\boldsymbol{k},\ell) = \begin{cases} \#(\boldsymbol{S}_{\boldsymbol{k},\ell}) & \text{for } \ell = \lceil |\boldsymbol{k}|/2 \rceil \text{ and } |\boldsymbol{k}| \text{ even}, \\ 2\#(\boldsymbol{S}_{\boldsymbol{k},\ell}) & \text{otherwise}, \end{cases}$$

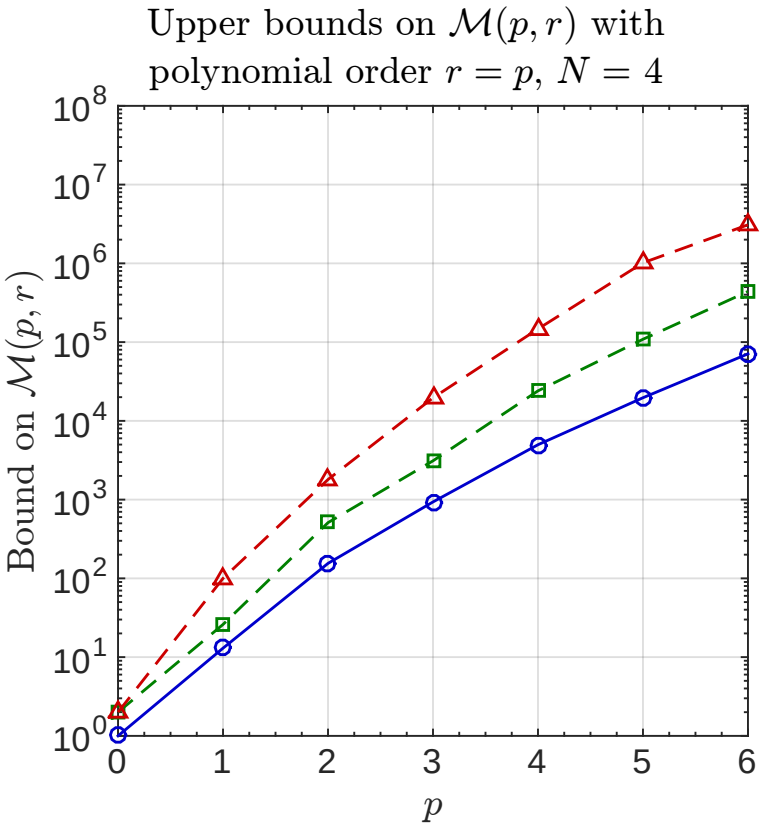
and $\boldsymbol{S}_{\boldsymbol{k},\ell} = \left\{ \boldsymbol{s} \in \mathbb{N}_0^d : |\boldsymbol{s}| = \ell,\ s_i \leq k_i \ \ \forall i \right\}$, with $\#(\boldsymbol{S}_{\boldsymbol{k},\ell})$ equal to the coefficient of t^ℓ in the generating function $P_{\boldsymbol{k}}(t) = \prod_{i=1}^d \sum_{j=1}^{k_i} t^j$.

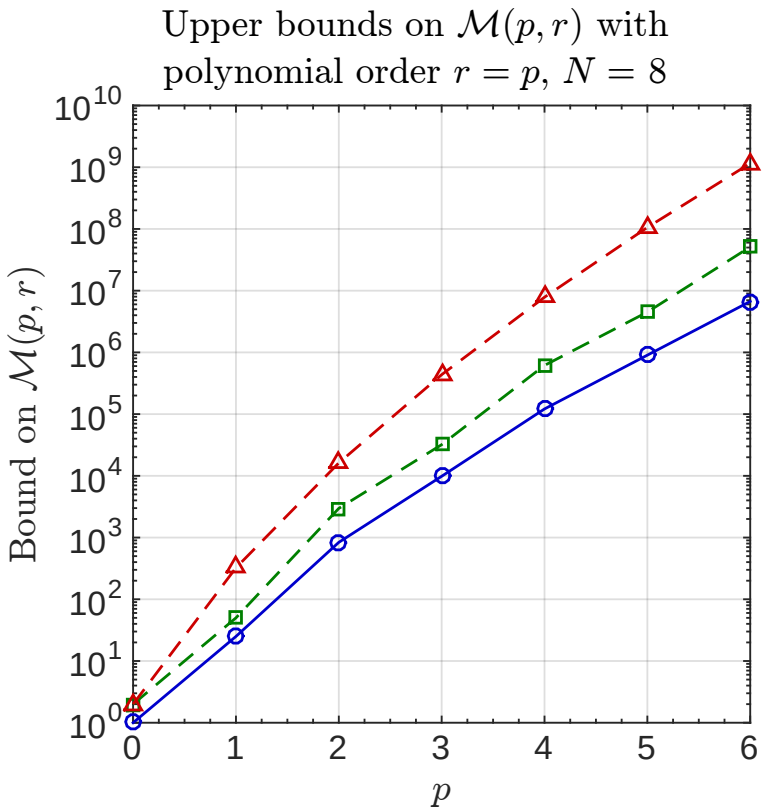
Moreover,

$$\begin{aligned} \mathcal{M}(p,r) = \sum_{\boldsymbol{k}\in\Lambda_r} \text{nnz}(\mathbf{G}_{\boldsymbol{k}}) &= \sum_{\boldsymbol{k}\in\Lambda_r} \sum_{\ell=\lceil |\boldsymbol{k}|/2 \rceil}^{|\boldsymbol{k}|} c(\boldsymbol{k},\ell) \binom{d+p-\ell}{p-\ell} \\ &\leq \sum_{j=0}^{r} 2\min\left\{ 2^j, \binom{d+\lceil j/2 \rceil}{d} \right\} \binom{d-1+j}{d-1} \binom{d+p-\lceil j/2 \rceil}{d} \\ &\leq 2\min\left\{ 2^r, \binom{d+\lceil r/2 \rceil}{d} \right\} \binom{d+p}{d} \binom{d+r}{d}. \end{aligned}$$

SGFEM - bounds for $\mathcal{M}(p,r)$

Q: How good are the bounds? A: Not sharp. In these figures $N = d$, dimension.





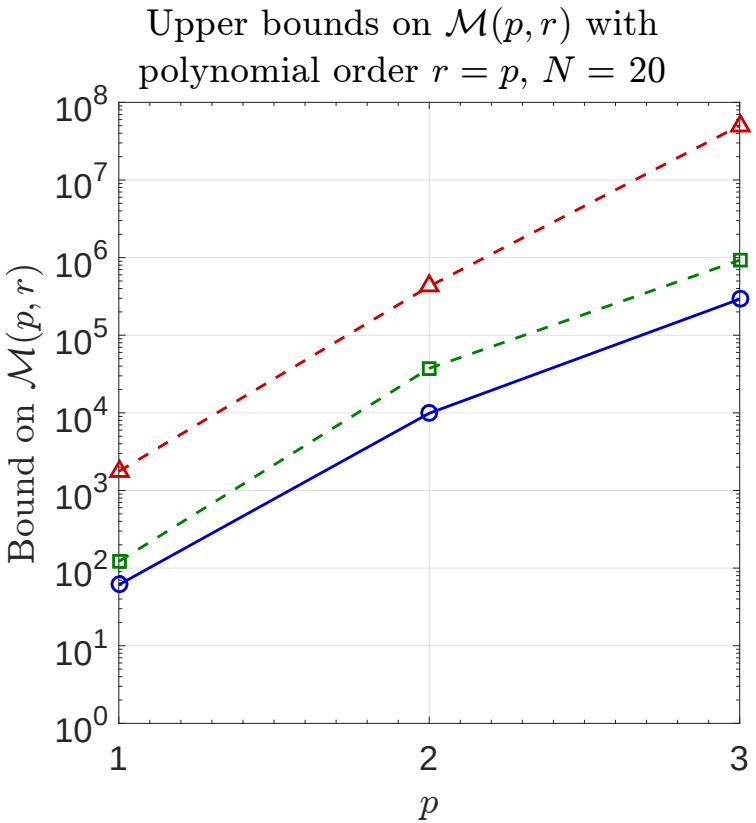
Actual count of $\mathcal{M}(p,r)$

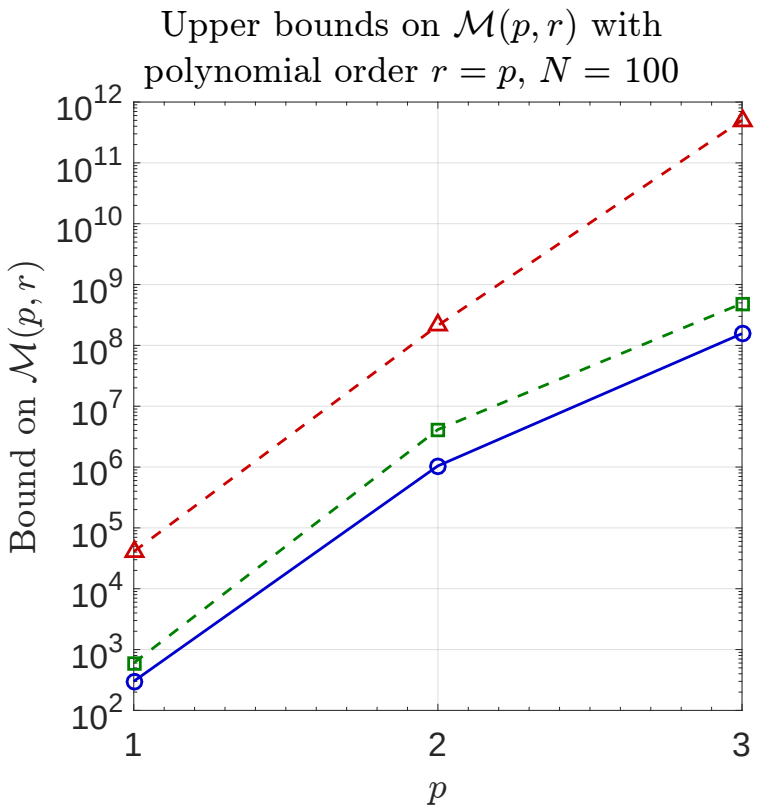
Sum bound (used when r is fixed): $\sum_{j=0}^{r} 2\min\left\{2^j, \binom{d+\lceil j/2\rceil}{d}\right\}\binom{d-1+j}{d-1}\binom{d+p-\lceil j/2\rceil}{d}$

Bound by largest (used when error depends on r): $2\min\left\{2^r, \binom{d+\lceil r/2\rceil}{d}\right\}\binom{d+p}{d}\binom{d+r}{d}$

SGFEM - bounds for $\mathcal{M}(p,r)$

Q: How good are the bounds? A: Not sharp. In these figures $N = d$, dimension.





Actual count of $\mathcal{M}(p,r)$

Sum bound (used when r is fixed): $\sum_{j=0}^{r} 2\min\left\{2^j, \binom{d+\lceil j/2\rceil}{d}\right\}\binom{d-1+j}{d-1}\binom{d+p-\lceil j/2\rceil}{d}$

Bound by largest (used when error depends on r): $2\min\left\{2^r, \binom{d+\lceil r/2\rceil}{d}\right\}\binom{d+p}{d}\binom{d+r}{d}$

Complexity of the SGFEM

Bounding the error of the fully-discrete approximation

The error in SGFEM found using CG can be bounded by:

$$\left\|u-\tilde{u}_{h,p}^{r}\right\|_{L_{\varrho}^{2}(\mathcal{U};H_{0}^{1}(D))} \leq \underbrace{\left\|u-u^{r}\right\|_{L_{\varrho}^{2}(\mathcal{U};H_{0}^{1}(D))}}_{\text{SG(I)}} + \underbrace{\left\|u^{r}-u_{h}^{r}\right\|_{L_{\varrho}^{2}(\mathcal{U};H_{0}^{1}(D))}}_{\text{SG(II)}}$$
$$+ \underbrace{\left\|u_{h}^{r}-u_{h,p}^{r}\right\|_{L_{\varrho}^{2}(\mathcal{U};H_{0}^{1}(D))}}_{\text{SG(III)}} + \underbrace{\left\|u_{h,p}^{r}-\tilde{u}_{h,p}^{r}\right\|_{L_{\varrho}^{2}(\mathcal{U};H_{0}^{1}(D))}}_{\text{SG(IV)}}.$$

We can estimate the complexity by finding the minimum and maximum values of the following parameters to ensure the total error of each component is smaller than $\frac{\varepsilon}{4}$.

- SG(I) error from approximating $a(x,\boldsymbol{y})$: estimate the min. projection order $r_{\min}$
- SG(II) FE error: estimate the maximum mesh size $h_{\max}$
- SG(III) SG error: estimate the minimum polynomial order $p_{\min}$
- SG(IV) CG error: estimate the minimum number of iterations needed by PCG

Substitute into the cost metric: $W^{\text{pSG}} \approx \mathcal{O}(J_h) * (N_p + \mathcal{M}(p,r)) * N_{\text{iter}}^{\text{pSG}}$

Complexity of the SGFEM

Depends on the coefficient $a(x, \boldsymbol{y})$

Asymptotically, as $\varepsilon \to 0$, the complexity can be estimated as

$$\mathcal{O}\underbrace{\left(\frac{1}{\varepsilon}\right)^{\frac{\tilde{d}}{q}}}_{\text{(SG.1)}}\underbrace{\left[\log\left(\frac{1}{\varepsilon}\right)\right]^{g(d)}}_{\text{(SG.2)}}\underbrace{\left(\frac{\log\left(\frac{1}{\varepsilon}\right)}{\log\left(\frac{\sqrt{\kappa_r}+1}{\sqrt{\kappa_r}-1}\right)}\right)}_{\text{(SG.3)}}, \tag{2}$$

where $g(d)$ depends on the coefficient $a(x, \boldsymbol{y})$ in the elliptic operator:

- $g(d) = d$ if $a(x, \boldsymbol{y})$ is a polynomial function of $\boldsymbol{y}$ of fixed order $\overline{r} < \infty$
- $g(d) = 3d$ if $a(x, \boldsymbol{y})$ is a transcendental function of $\boldsymbol{y}$, requiring an orthogonal expansion of order $r \geq r_{\min}$ depending on ε

Here, we assume $u^r(\boldsymbol{y}) \in H_0^1(D) \cap H^{q+1}(D)\ \forall \boldsymbol{y} \in \mathcal{U}$, and

- (SG.1) corresponds to the work required by the finite element method
- (SG.2) corresponds to the work required by the SG method, coming from the estimates on the number of **coupled** finite element systems to solve
- (SG.3) corresponds to the work required by the PCG method

Complexity comparisons between SGFEM and SCFEM

SGFEM Complexity (Polynomial):

$$\mathcal{O}\underbrace{\left(\frac{1}{\varepsilon}\right)^{\frac{\tilde{d}}{q}}}_{\text{(SG.1)}}\underbrace{\left[\log\left(\frac{1}{\varepsilon}\right)\right]^{d}}_{\text{(SG.2)}}\underbrace{\left(\frac{\log\left(\frac{1}{\varepsilon}\right)}{\log\left(\frac{\sqrt{\kappa_r}+1}{\sqrt{\kappa_r}-1}\right)}\right)}_{\text{(SG.3)}},$$

SCFEM Complexity [Galindo, Jantsch, W., Zhang '16] :

$$\mathcal{O}\underbrace{\left(\frac{1}{\varepsilon}\right)^{\frac{\tilde{d}}{q}}}_{\text{(SC.1)}}\underbrace{\left[\log\left(\frac{1}{\varepsilon}\right)\right]^{d}\left[\log\log\left(\frac{1}{\varepsilon}\right)\right]^{d-1}}_{\text{(SC.2)}}\underbrace{\left(\frac{\log\left(\frac{1}{\varepsilon}\right)}{\log\left(\frac{\sqrt{\bar{\kappa}}+1}{\sqrt{\bar{\kappa}}-1}\right)}\right)}_{\text{(SC.3)}}$$

Here, $\bar{\kappa}$ is the supremum of the condition numbers of the SC systems, and

- (SC.1) corresponds to the work required by the finite element method
- (SC.2) corresponds to the work required by the SC method, coming from the estimates on the number of **decoupled** finite element systems to solve
- (SC.3) corresponds to the work required by the PCG method

Complexity comparisons between SGFEM and SCFEM

SGFEM Complexity (Transcendental):

$$\underbrace{\mathcal{O}\left(\frac{1}{\varepsilon}\right)^{\frac{\tilde{d}}{q}}}_{\text{(SG.1)}}\underbrace{\left[\log\left(\frac{1}{\varepsilon}\right)\right]^{3d}}_{\text{(SG.2)}}\underbrace{\left(\frac{\log\left(\frac{1}{\varepsilon}\right)}{\log\left(\frac{\sqrt{\kappa_r}+1}{\sqrt{\kappa_r}-1}\right)}\right)}_{\text{(SG.3)}},$$

SCFEM Complexity [Galindo, Jantsch, W., Zhang '16] :

$$\underbrace{\mathcal{O}\left(\frac{1}{\varepsilon}\right)^{\frac{\tilde{d}}{q}}}_{\text{(SC.1)}}\underbrace{\left[\log\left(\frac{1}{\varepsilon}\right)\right]^{d}\left[\log\log\left(\frac{1}{\varepsilon}\right)\right]^{d-1}}_{\text{(SC.2)}}\underbrace{\left(\frac{\log\left(\frac{1}{\varepsilon}\right)}{\log\left(\frac{\sqrt{\bar{\kappa}}+1}{\sqrt{\bar{\kappa}}-1}\right)}\right)}_{\text{(SC.3)}}$$

Here, $\bar{\kappa}$ is the supremum of the condition numbers of the SC systems, and

- (SC.1) corresponds to the work required by the finite element method
- (SC.2) corresponds to the work required by the SC method, coming from the estimates on the number of **decoupled** finite element systems to solve
- (SC.3) corresponds to the work required by the PCG method

Numerical illustrations of the complexity results

We solve both systems with the iterative Preconditioned Conjugate Gradient (PCG):

- For the SGFEM, we use block-diagonal mean-based preconditioner [Powell, Elman '08]
- For the SCFEM, we use the incomplete Cholesky factorization of $\mathbf{A}(\boldsymbol{y}_0)$ as preconditioner for each $\mathbf{A}(\boldsymbol{y}_k)$ for $\{\boldsymbol{y}_k\}_{k=1}^m$

We test convergence against a "highly enriched" approximation, obtained with stochastic collocation based on Clenshaw-Curtis abscissas, and then approximate

- $\|\mathbb{E}[\varepsilon_{\mathsf{SG}}]\|_{\ell^\infty} \approx \|\mathbb{E}[u_{\mathsf{ex}} - u_{\mathsf{SG}}]\|_{\ell^\infty}$ the error of the stochastic Galerkin approximation
- $\|\mathbb{E}[\varepsilon_{\mathsf{SC}}]\|_{\ell^\infty} \approx \|\mathbb{E}[u_{\mathsf{ex}} - u_{\mathsf{SC}}]\|_{\ell^\infty}$ the error of the stochastic collocation approximation

To ensure that we do not "over-resolve" either approximation with PCG, we use a tolerance of $\|\mathbb{E}[\varepsilon_{\mathsf{SG}}]\|_{\ell^\infty}/10$ and $\|\mathbb{E}[\varepsilon_{\mathsf{SC}}]\|_{\ell^\infty}/10$ for the stochastic Galerkin and stochastic collocation methods, respectively, after solving with an initial tolerance of $1\mathrm{E}-12$.

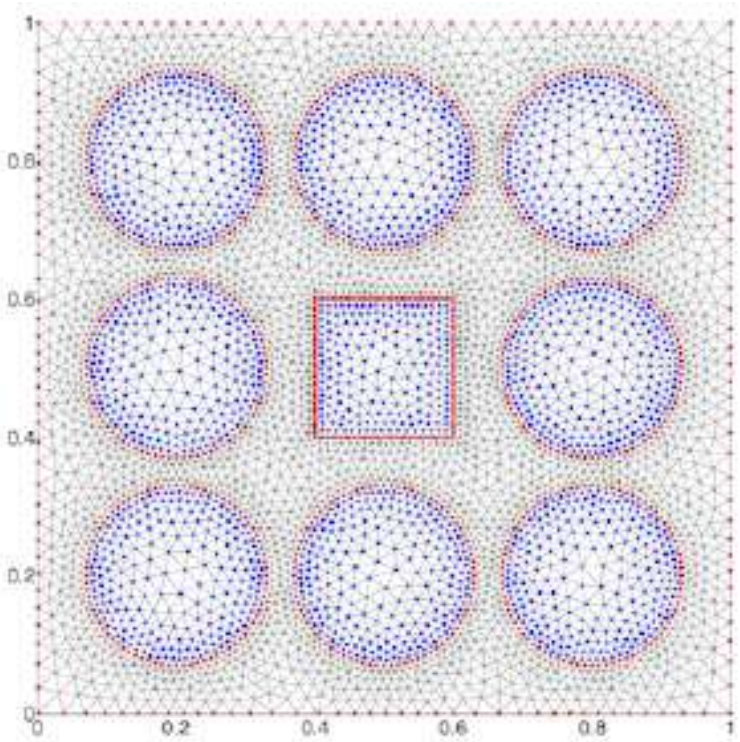
Numerical illustrations of the complexity results

Affine piecewise constant coefficient

Stochastic elliptic problem:

$$\begin{cases} -\nabla \cdot (a(x,\boldsymbol{y})\nabla u(x,\boldsymbol{y})) &= 100\chi_F(x) & \text{in } \mathcal{U}\times D, \\ u(x,\boldsymbol{y}) &= 0 & \text{on } \mathcal{U}\times \partial D, \end{cases}$$

- $a(x,\boldsymbol{y}) = 1 + \sum_{n=1}^{8} y_n \chi_n(x)$ and $y_n \sim \mathcal{U}(-0.99, -0.2)$
- χ_n, χ_F are indicators of the circles and the square



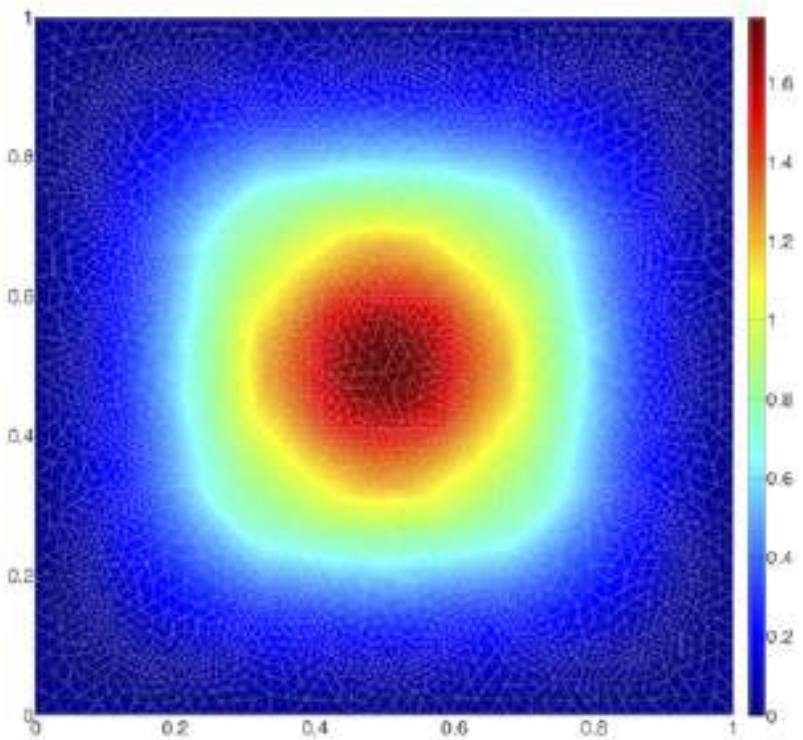


Figure Triangulation of $D = [0,1]^2$ and $\mathbb{E}[u]$.

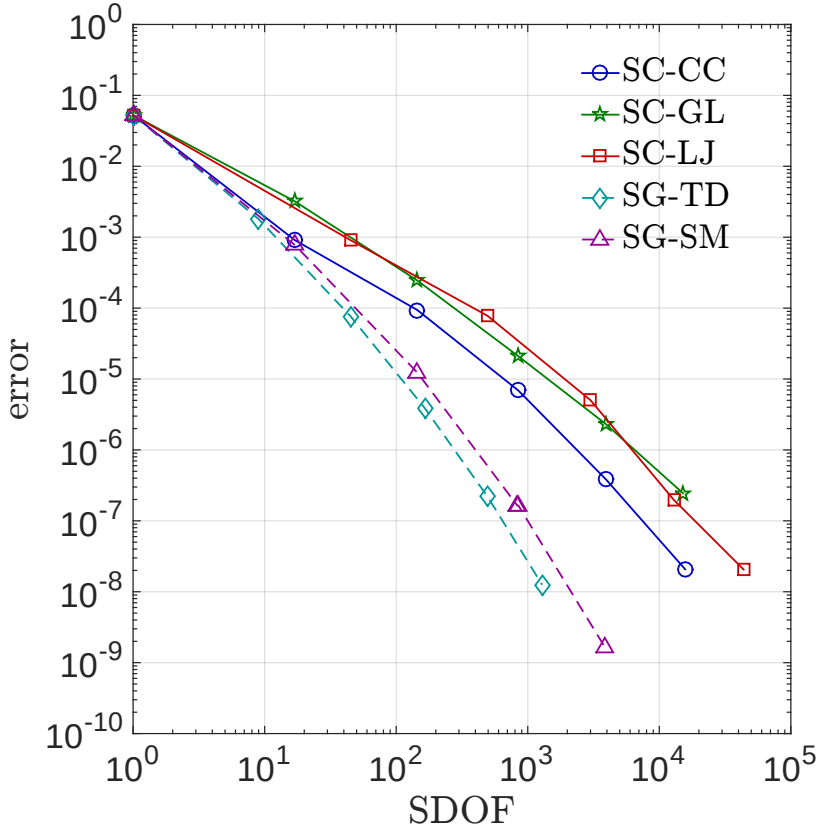
Numerical illustrations of the complexity results

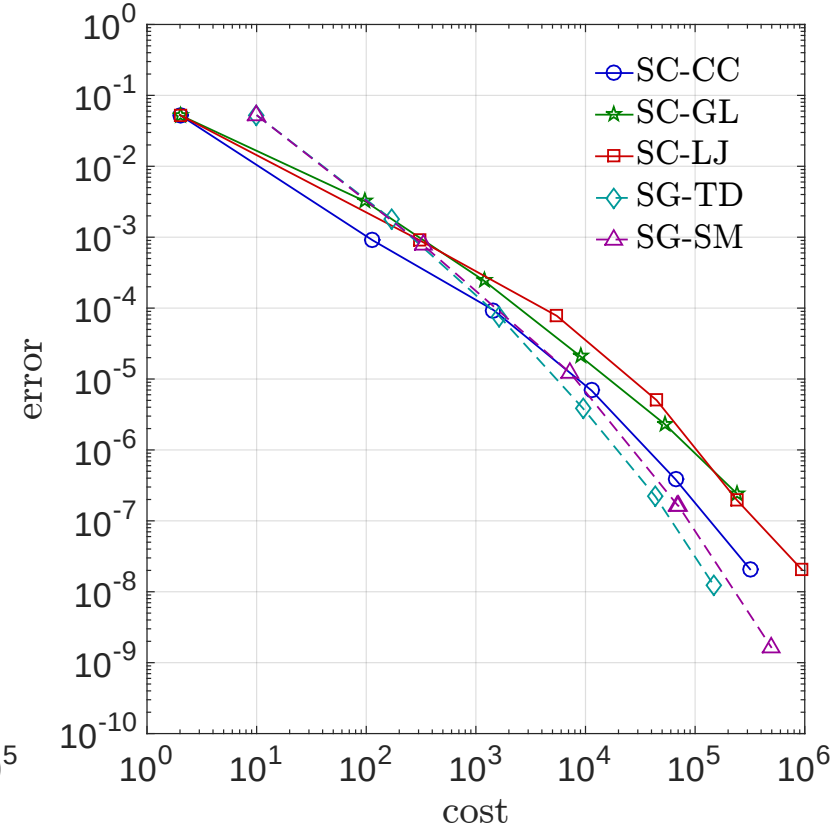
Affine piecewise constant coefficient

Stochastic elliptic problem:

$$\begin{cases} -\nabla \cdot (a(x,\boldsymbol{y})\nabla u(x,\boldsymbol{y})) &= 100\chi_F(x) & \text{in } \mathcal{U}\times D, \\ u(x,\boldsymbol{y}) &= 0 & \text{on } \mathcal{U}\times \partial D, \end{cases}$$

- $a(x,\boldsymbol{y}) = 1 + \sum_{n=1}^{8} y_n \chi_n(x)$ and $y_n \sim \mathcal{U}(-0.99, -0.2)$
- χ_n, χ_F are indicators of the circles and the square





Numerical illustrations of the complexity results

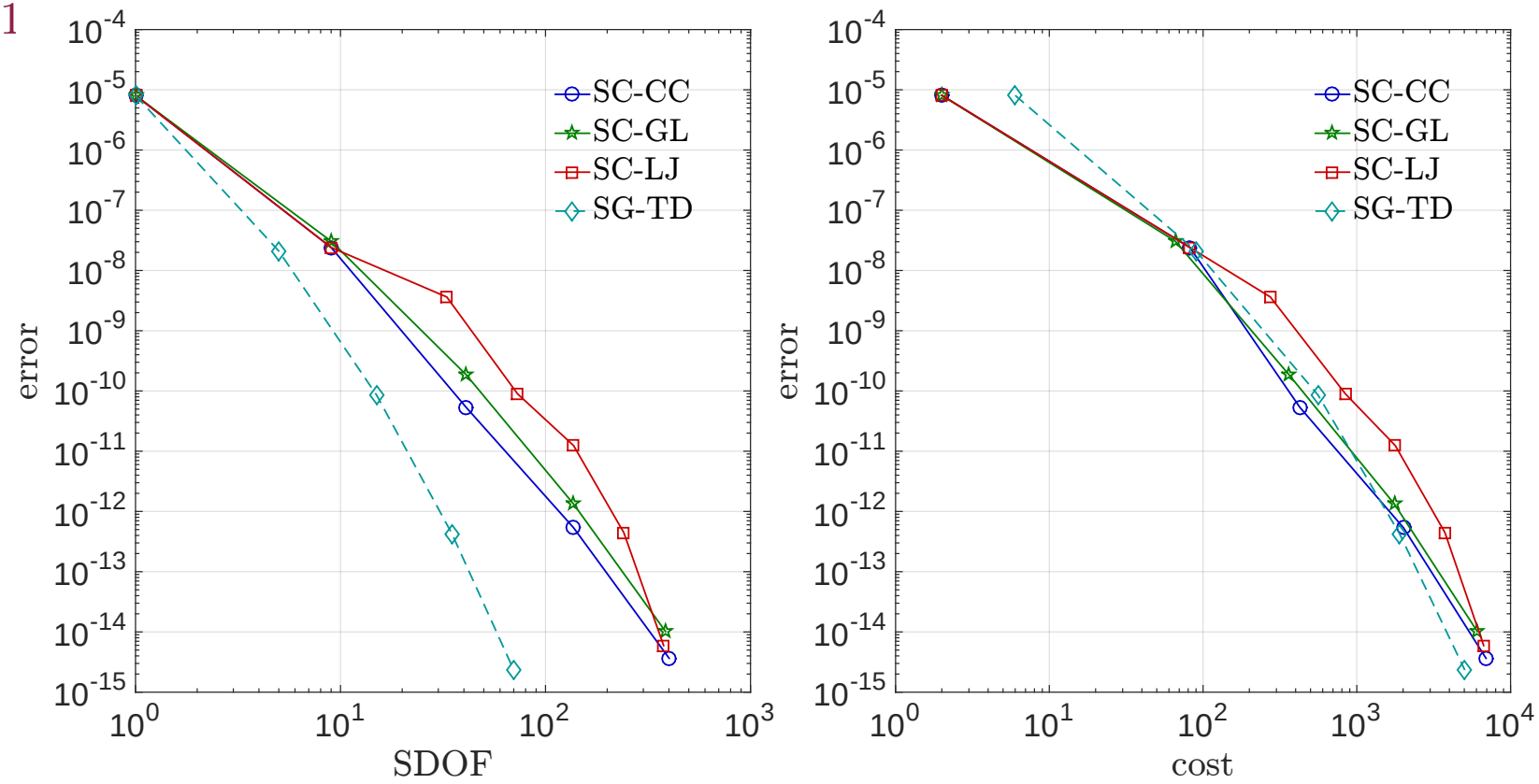
Polynomial coefficient

Stochastic elliptic problem:

$$\begin{cases} -\nabla \cdot (a(x,\boldsymbol{y})\nabla u(x,\boldsymbol{y})) &= 1 \quad \text{in } \mathcal{U}\times D, \\ u(x,\boldsymbol{y}) &= 0 \quad \text{on } \mathcal{U}\times \partial D, \end{cases}$$

$$a(x,\boldsymbol{y}) = 5+\sum_{|\boldsymbol{r}|\le \overline{r}} e^{-1.5|\boldsymbol{r}|}\varsigma_{\boldsymbol{r}}(x)\boldsymbol{y}^{\boldsymbol{r}} \qquad \varsigma_{\boldsymbol{r}}(x) = \begin{cases} \sin(|\boldsymbol{r}|\pi x_1)\cos(|\boldsymbol{r}|\pi x_2) & |\boldsymbol{r}| \text{ even}, \\ \cos(|\boldsymbol{r}|\pi x_1)\sin(|\boldsymbol{r}|\pi x_2) & |\boldsymbol{r}| \text{ odd}, \end{cases}$$

$\overline{r} = 1$



Numerical illustrations of the complexity results

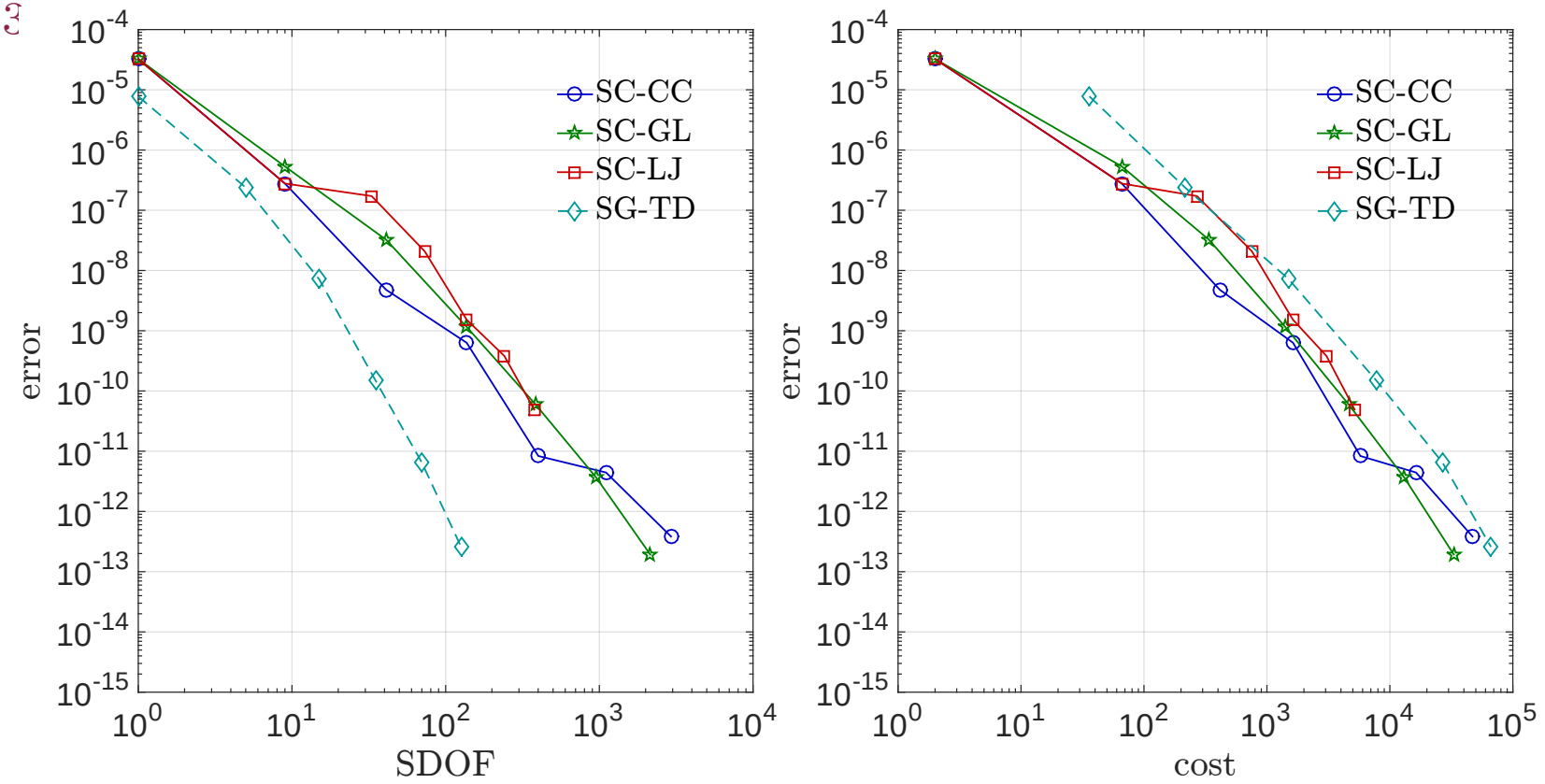
Polynomial coefficient

Stochastic elliptic problem:

$$\begin{cases} -\nabla \cdot (a(x,\boldsymbol{y})\nabla u(x,\boldsymbol{y})) &= 1 \quad \text{in } \mathcal{U}\times D, \\ u(x,\boldsymbol{y}) &= 0 \quad \text{on } \mathcal{U}\times \partial D, \end{cases}$$

$$a(x,\boldsymbol{y}) = 5+\sum_{|\boldsymbol{r}|\le \overline{r}} e^{-1.5|\boldsymbol{r}|}\varsigma_{\boldsymbol{r}}(x)\boldsymbol{y}^{\boldsymbol{r}} \qquad \varsigma_{\boldsymbol{r}}(x) = \begin{cases} \sin(|\boldsymbol{r}|\pi x_1)\cos(|\boldsymbol{r}|\pi x_2) & |\boldsymbol{r}| \text{ even}, \\ \cos(|\boldsymbol{r}|\pi x_1)\sin(|\boldsymbol{r}|\pi x_2) & |\boldsymbol{r}| \text{ odd}, \end{cases}$$

$\overline{r} = 3$



Numerical illustrations of the complexity results

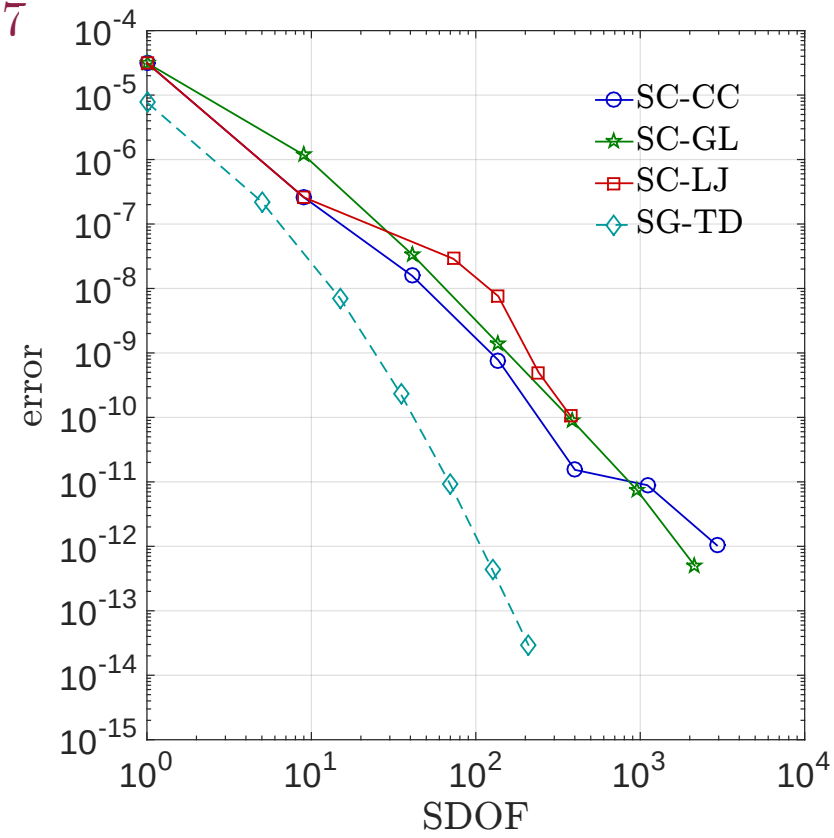
Polynomial coefficient

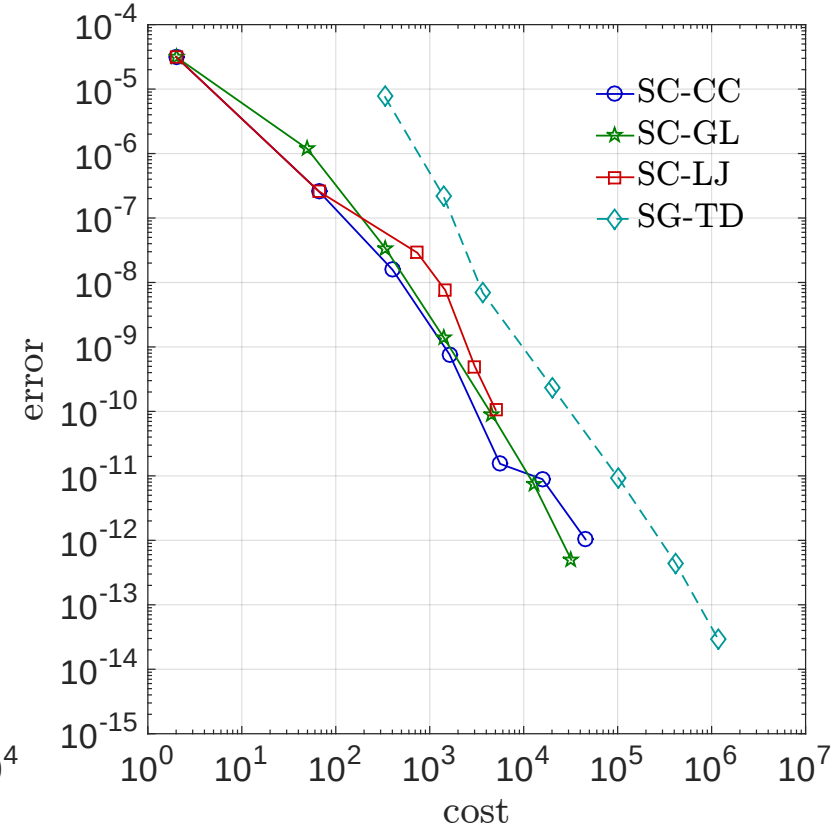
Stochastic elliptic problem:

$$\begin{cases} -\nabla \cdot (a(x,\boldsymbol{y})\nabla u(x,\boldsymbol{y})) &= 1 \quad \text{in } \mathcal{U}\times D, \\ u(x,\boldsymbol{y}) &= 0 \quad \text{on } \mathcal{U}\times \partial D, \end{cases}$$

$$a(x,\boldsymbol{y}) = 5 + \sum_{|\boldsymbol{r}|\le \overline{r}} e^{-1.5|\boldsymbol{r}|}\varsigma_{\boldsymbol{r}}(x)\boldsymbol{y}^{\boldsymbol{r}} \qquad \varsigma_{\boldsymbol{r}}(x) = \begin{cases} \sin(|\boldsymbol{r}|\pi x_1)\cos(|\boldsymbol{r}|\pi x_2) & |\boldsymbol{r}| \text{ even}, \\ \cos(|\boldsymbol{r}|\pi x_1)\sin(|\boldsymbol{r}|\pi x_2) & |\boldsymbol{r}| \text{ odd}, \end{cases}$$

$\overline{r} = 7$





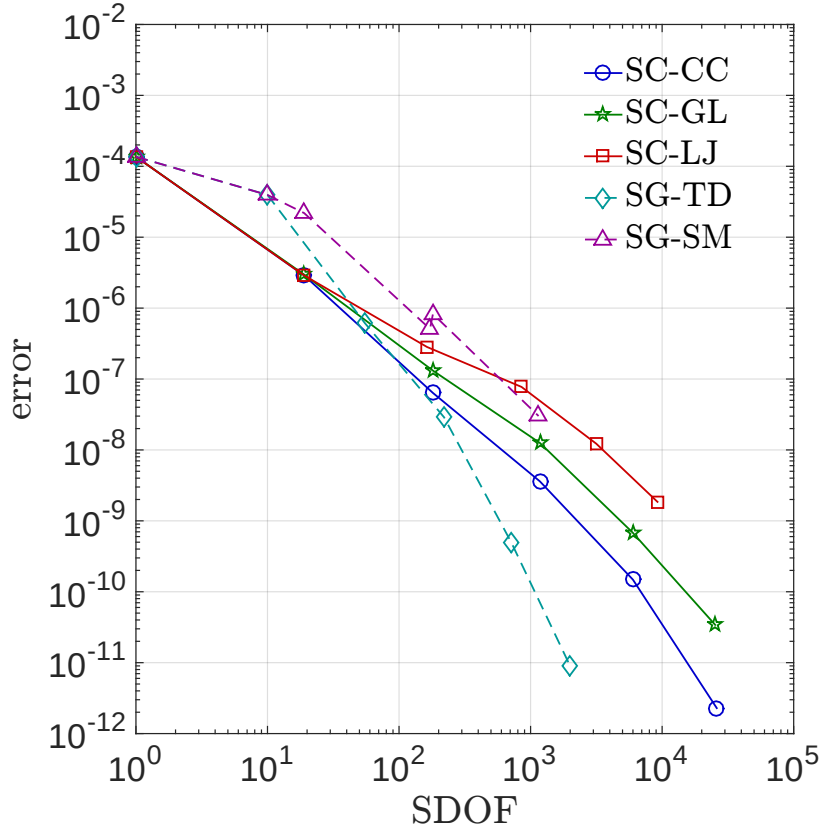
Numerical illustrations of the complexity results

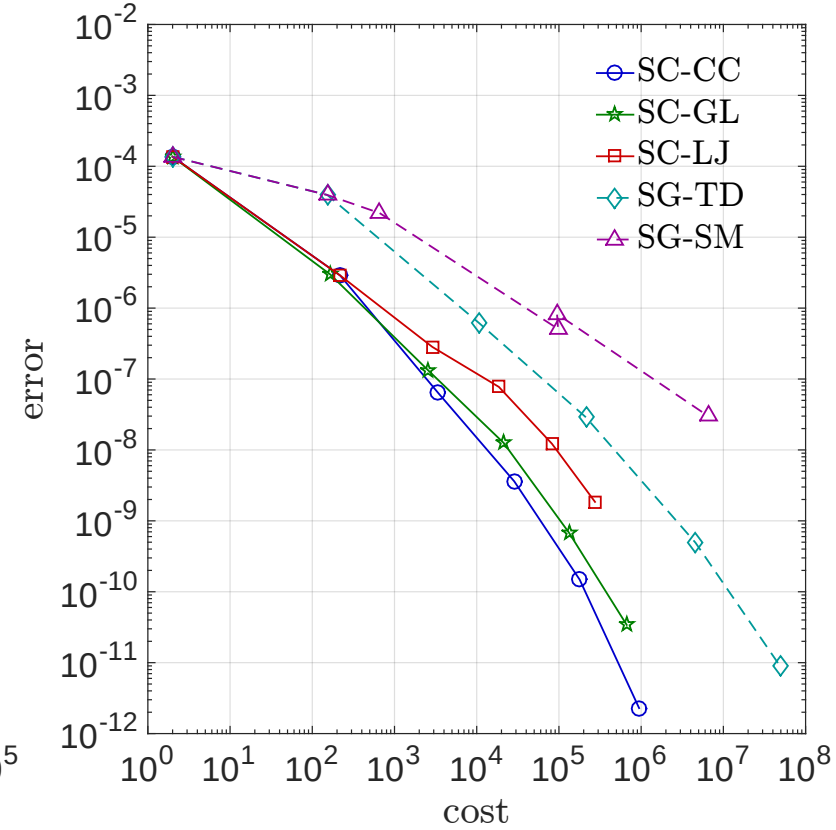
Transcendental "log-transformed" coefficient

Stochastic elliptic problem:

$$\begin{cases} -\nabla\cdot(a(x,\boldsymbol{y})\nabla u(x,\boldsymbol{y})) &= f(x) \quad \text{in } \mathcal{U}\times D, \\ u(x,\boldsymbol{y}) &= 0 \quad \text{on } \mathcal{U}\times\partial D, \end{cases}$$

$a(x,\boldsymbol{y}) \approx 0.5+\exp(\varphi_0+\sum_{k=1}^{d}\sqrt{\lambda_k}\varphi_k y_k)$, $\{\lambda_k,\varphi_k\}_{k=1}^{d}$ the largest eigenpairs of the squared exponential covariance kernel $\mathbb{C}\text{ov}[a]$ with correlation length $L_c = 1/64$, $f(x_1,x_2) = 2\cos(x_1)\sin(x_2)$, and $d = 9$.





Numerical illustrations of the complexity results

Transcendental "log-transformed" coefficient

Stochastic elliptic problem:

$$\begin{cases} -\nabla \cdot (a(x,\boldsymbol{y})\nabla u(x,\boldsymbol{y})) &= f(x) \quad \text{in } \mathcal{U}\times D, \\ u(x,\boldsymbol{y}) &= 0 \quad \text{on } \mathcal{U}\times \partial D, \end{cases}$$

$a(x,\boldsymbol{y}) \approx 0.5 + \exp(\varphi_0 + \sum_{k=1}^{d}\sqrt{\lambda_k}\varphi_k y_k)$, $\{\lambda_k,\varphi_k\}_{k=1}^{d}$ the largest eigenpairs of the squared exponential covariance kernel $\mathbb{C}\text{ov}[a]$ with correlation length $L_c = 1/64$, $f(x_1,x_2) = 2\cos(x_1)\sin(x_2)$, and $d = 9$.

SC-CC Level	SC-CC Error	Mat-vec cost of SC-CC	SG-TD Order	SG-TD Error	Mat-vec cost of SG-TD	Ratio SG/SC
0	1.3626×10^{-4}	2	0	1.3626×10^{-4}	4	2
1	2.8884×10^{-6}	218	1	3.9444×10^{-5}	152	0.69
2	6.3652×10^{-8}	3,398	2	6.1427×10^{-7}	10,710	3.15
3	3.6021×10^{-9}	28,638	3	2.8851×10^{-8}	213,010	7.43
4	1.4794×10^{-10}	178,894	4	4.9210×10^{-10}	4,579,575	25.59
5	2.2869×10^{-12}	944,220	5	8.9123×10^{-12}	49,089,051	51.98

- Stochastic Galerkin method features optimal error w.r.t. degrees of freedom
- Cost of the method is not optimal in every case
- Ignored issues associated with forming/solving the stochastic Galerkin system
  - High-dimensional integration problem in computing $\mathbf{K}_r$
  - Poorly conditioned with respect to p and r