Analysis of SPDEs and numerical methods for UQ

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

John Burkardt† & Clayton Webster⋆

Thanks to Max Gunzburger & Guannan Zhang (FSU), Fabio Nobile (MOX), Raul Tempone (KAUST)

†Department of Scientific Computing
Florida State University

⋆CEES, CASL, Computer Science and Mathematics Division
Oak Ridge National Laboratory

April 2-3, 2012

Outline

Stochastic formulation of uncertainty

A simplified general setting

Consider an operator $\mathcal{L}$, linear or nonlinear, on a domain $D \subset \mathbb{R}^d$, which depends on some coefficients $a(\omega, x)$ with $x \in D$, $\omega \in \Omega$ and $(\Omega, \mathcal{F}, P)$ a complete probability space. The forcing $f = f(\omega, x)$ and the solution $u = u(\omega, x)$ are random fields s.t.

$$\mathcal{L}(a)(u) = f \quad \text{a.e. in } D \tag{1}$$

equipped with suitable boundary conditions.

A_1. the solution to (1) has realizations in the Banach space $W(D)$, i.e. $u(\cdot, \omega) \in W(D)$ almost surely

$$\|u(\cdot, \omega)\|_{W(D)} \leq C \|f(\cdot, \omega)\|_{W^*(D)}$$

A_2. the forcing term $f \in L^2_{\mathbb{P}}(\Omega) \otimes W^*(D) \equiv L^2_{\mathbb{P}}(\Omega; W^*(D))$ is such that the solution u is unique and bounded in $L^2_{\mathbb{P}}(\Omega) \otimes W(D) \equiv L^2_{\mathbb{P}}(\Omega; W(D))$

A_3. $\mathbb{P}\left[a(\omega, x) \in (a_{min}, a_{max})\, \forall x \in \overline{D}\right] = 1$, $a_{min} > 0$, $a_{max} < \infty$

Stochastic formulation of uncertainty

A simplified general setting

OAK RIDGE National Laboratory

Consider an operator $\mathcal{L}$, linear or nonlinear, on a domain $D \subset \mathbb{R}^d$, which depends on some coefficients $a(\omega, x)$ with $x \in D$, $\omega \in \Omega$ and $(\Omega, \mathcal{F}, P)$ a complete probability space. The forcing $f = f(\omega, x)$ and the solution $u = u(\omega, x)$ are random fields s.t.

$$\mathcal{L}(a)(u) = f \quad \text{a.e. in } D \tag{1}$$

equipped with suitable boundary conditions.

A_1. the solution to (1) has realizations in the Banach space $W(D)$, i.e. $u(\cdot, \omega) \in W(D)$ almost surely

$$\|u(\cdot, \omega)\|_{W(D)} \leq C \|f(\cdot, \omega)\|_{W^*(D)}$$

A_2. the forcing term $f \in L^2_{\mathbb{P}}(\Omega) \otimes W^*(D) \equiv L^2_{\mathbb{P}}(\Omega; W^*(D))$ is such that the solution u is unique and bounded in $L^2_{\mathbb{P}}(\Omega) \otimes W(D) \equiv L^2_{\mathbb{P}}(\Omega; W(D))$

A_3. $\mathbb{P}\left[a(\omega, x) \in (a_{min}, a_{max}) \, \forall x \in \overline{D}\right] = 1$, $a_{min} > 0$, $a_{max} < \infty$

Examples

Linear and Nonlinear Elliptic SPDEs

Example: The linear elliptic problem

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

with $f(\omega,\cdot)$ square integrable with respect to $\mathbb{P}$, satisfies assumptions A_1, A_2 and A_3 with $W(D) = H_0^1(D)$

Example: The nonlinear elliptic problem

Similarly, for $k \in \mathbb{N}^+$,

$$\begin{cases} -\nabla \cdot (a(\omega,\cdot)\nabla u(\omega,\cdot)) + u(\omega,\cdot)|u(\omega,\cdot)|^k &= f(\omega,\cdot) \quad \text{in } \Omega \times D, \\ u(\omega,\cdot) &= 0 \quad \text{on } \Omega \times \partial D, \end{cases}$$

satisfies assumptions A_1, A_2 and A_3 with $W(D) = H_0^1(D) \cap L^{k+2}(D)$

Goal of the computations

Stochastic QoI

OAK RIDGE National Laboratory

Forward Problem: to approximate u or some statistical QoI depending on u:

$$\mathbf{\Phi}_u = \langle \mathbf{\Phi}(u) \rangle := \mathbb{E}\left[\mathbf{\Phi}(u)\right] = \int_\Omega \int_D \Phi(u(\omega,x),\omega,x)\,dx\,d\mathbb{P}(\omega)$$

e.g. $\overline{u}(x_0) = \mathbb{E}[u](x_0)$, OR $\mathbb{V}ar[u](x_0) = \mathbb{E}[(\widetilde{u})^2](x_0)$, where $\widetilde{u} = u - \overline{u}$,

OR $\mathbb{P}\left[u \geq u_0\right] = \mathbb{P}\left[\{\omega \in \Omega \,:\, u(\omega,x_0) \geq u_0\}\right] = \mathbb{E}\left[\chi_{\{u \geq u_0\}}\right]$,

OR even statistics of functionals of u, i.e. $\phi(u) = \int_{\Sigma \subset D} u(\cdot,x)dx$

where Σ is a subdomain of interest.

Goal: to develop highly efficient, robust and scalable techniques that include uncertainty in the models, allow us to quantify uncertainty in the outputs and provide reliable and verifiable predictions. **Probability Theory** provides an effective tool to describe and propagate uncertainty.

Goal of the computations

Stochastic QoI

Forward Problem: to approximate u or some statistical QoI depending on u:

$$\mathbf{\Phi}_u = \langle \mathbf{\Phi}(u) \rangle := \mathbb{E}\left[\mathbf{\Phi}(u)\right] = \int_{\Omega} \int_{D} \Phi(u(\omega, x), \omega, x) dx \, d\mathbb{P}(\omega)$$

e.g. $\overline{u}(x_0) = \mathbb{E}[u](x_0)$, OR $\mathbb{V}ar[u](x_0) = \mathbb{E}[(\widetilde{u})^2](x_0)$, where $\widetilde{u} = u - \overline{u}$,

OR $\mathbb{P}\left[u \geq u_0\right] = \mathbb{P}\left[\{\omega \in \Omega \,:\, u(\omega, x_0) \geq u_0\}\right] = \mathbb{E}\left[\chi_{\{u \geq u_0\}}\right]$,

OR even statistics of functionals of u, i.e. $\phi(u) = \int_{\Sigma \subset D} u(\cdot, x) dx$

where Σ is a subdomain of interest.

Goal: to develop highly efficient, robust and scalable techniques that include uncertainty in the models, allow us to quantify uncertainty in the outputs and provide reliable and verifiable predictions. **Probability Theory** provides an effective tool to describe and propagate uncertainty.

Finite dimensional noise assumption

Nonlinear input data: $a \to a_N$, $f \to f_N$

WLOG assume the random fields $a(\omega, x)$ and $f(\omega, x)$ depends on a finite number of random variables $\mathbf{Y}(\omega) = [Y_1(\omega), \ldots, Y_N(\omega)] : \Omega \to \mathbb{R}^N$:

$$a_N(\omega, x) = a(\mathbf{Y}(\omega), x), \quad f_N(\omega, x) = f(\mathbf{Y}(\omega), x)$$

1. Piecewise constant material properties: Let $\{D_n\}_{n=1}^N$ be a partition of D then define $a_N(\omega, x) = \sum_{i=1}^N \sigma_i Y_i(\omega) \chi_{D_i}(x)$
2. ∞-dimensional random field suitably truncated, e.g. lognormal permeability model in groundwater flows
   - $\forall \omega \in \Omega$, $a(\omega, \cdot) \in L^\infty(D)$
   - $\forall x_0 \in D$, $a(\cdot, x_0)$ is a random variable, e.g. $a(\cdot, x_0) \sim N(\mu, \sigma)$
   - the interaction between points is described by a covariance function, e.g. $\mathbb{C}ov[a](x_1, x_2) = \mathbb{E}\left[\widetilde{a}(\cdot, x_1)\widetilde{a}(\cdot, x_2)\right] = \sigma^2 \exp\left(-\frac{\|x_1 - x_2\|^2}{L_c^2}\right)$

   Expand a in a Karhunen-Loève expansion and retain the first N terms, denoted a_N, to capture most of the variability

Finite dimensional noise assumption

Nonlinear input data: $a \to a_N$, $f \to f_N$

WLOG assume the random fields $a(\omega, x)$ and $f(\omega, x)$ depends on a finite number of random variables $\mathbf{Y}(\omega) = [Y_1(\omega), \ldots, Y_N(\omega)] : \Omega \to \mathbb{R}^N$:

$$a_N(\omega, x) = a(\mathbf{Y}(\omega), x), \quad f_N(\omega, x) = f(\mathbf{Y}(\omega), x)$$

1. Piecewise constant material properties: Let $\{D_n\}_{n=1}^N$ be a partition of D then define $a_N(\omega, x) = \sum_{i=1}^N \sigma_i Y_i(\omega) \chi_{D_i}(x)$
2. ∞-dimensional random field suitably truncated, e.g. lognormal permeability model in groundwater flows
   - $\forall \omega \in \Omega$, $a(\omega, \cdot) \in L^\infty(D)$
   - $\forall x_0 \in D$, $a(\cdot, x_0)$ is a random variable, e.g. $a(\cdot, x_0) \sim N(\mu, \sigma)$
   - the interaction between points is described by a covariance function, e.g. $\mathbb{C}ov[a](x_1, x_2) = \mathbb{E}\left[\widetilde{a}(\cdot, x_1)\widetilde{a}(\cdot, x_2)\right] = \sigma^2 \exp\left(-\frac{\|x_1 - x_2\|^2}{L_c^2}\right)$

   Expand a in a Karhunen-Loève expansion and retain the first N terms, denoted a_N, to capture most of the variability

Example random fields

b_n: eigenfunctions of $\int_D \mathbb{C}_a(x_1, x_2) b_n(x_2) dx_2 = \lambda_n b_n(x_1)$, $\|b_n\|_{L^2(D)} = 1$
Y_n: uncorrelated RVs with $\mathbb{V}ar[Y_n] = \lambda_n$

Example: Uniform random field

$$a(\omega, x) = a_0 + \sigma \sum_{n=1}^{\infty} b_n(x) Y_n(\omega)$$

- $Y_n \sim U(-\sqrt{3}, \sqrt{3})$, $\mathbb{E}[Y_n] = 0$, $\mathbb{V}ar[Y_n] = 1$
- $a_{min} = a_0 - \sigma \sum_{n=1}^{\infty} \sqrt{3} \|b_n\|_{L^\infty(D)} > 0$ if σ is not too large

Example: Lognormal random field

$$a(\omega, x) = a_0 + \exp\left(\sum_{n=1}^{\infty} b_n(x) Y_n(\omega) \right)$$

- $Y_n \sim N(0, 1)$, $\mathbb{E}[Y_n] = 0$, $\mathbb{V}ar[Y_n] = 1$
- $a_{min} = 0$ and $a_{max} = \infty$

Approximating a Stochastic PDE

Transform SBVP to Parameterized deterministic BVP

- Given $a_N(\mathbf{Y}(\omega), x)$, $f_N(\mathbf{Y}(\omega), x) \Rightarrow u_N(Y_1(\omega), \ldots, Y_N(\omega), x)$ s.t.

$$\mathcal{L}(a_N)(u_N) = f_N \quad \text{in } D \text{ a.s.}$$

- $\Gamma_n \equiv Y_n(\Omega) \subset \mathbb{R}$ and $\Gamma = \prod_{n=1}^{N} \Gamma_n \subset \mathbb{R}^N$ - image of the random vector $\mathbf{Y}(\Omega)$ (curse of dimensionality when N is large)
- $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_N)$ has a joint PDF $\rho : \Gamma \to \mathbb{R}_+$, with $\rho \in L^\infty(\Gamma)$, i.e. for $\mathbf{y} \in \Gamma$

$$\mathbb{P}\left[Z \in \gamma \subset \Gamma\right] = \int_\gamma \rho(\mathbf{y})\, d\mathbf{y},$$

i.e. transform the measure $\mathbb{P}$ to $\mathbb{R}^N$

Quantities of interest (QoI)

Our goal of predicting the statistical behavior of a physical system often requires the approximation of multi-dimensional statistical QoIs, e.g.:

$$\mathbb{E}[u](x) = \int_\Gamma u(\mathbf{y}, x)\rho(\mathbf{y})d\mathbf{y}, \quad \text{where } \mathbf{y} \in \Gamma^N \text{ and } x \in \overline{D}$$

Approximating a Stochastic PDE

Transform SBVP to Parameterized deterministic BVP

- Given $a_N(\mathbf{Y}(\omega), x)$, $f_N(\mathbf{Y}(\omega), x) \Rightarrow u_N(Y_1(\omega), \ldots, Y_N(\omega), x)$ s.t.

$$\mathcal{L}(a_N)(u_N) = f_N \quad \text{in } D \text{ a.s.}$$

- $\Gamma_n \equiv Y_n(\Omega) \subset \mathbb{R}$ and $\Gamma = \prod_{n=1}^N \Gamma_n \subset \mathbb{R}^N$ - image of the random vector $\mathbf{Y}(\Omega)$ (curse of dimensionality when N is large)
- $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_N)$ has a joint PDF $\rho : \Gamma \to \mathbb{R}_+$, with $\rho \in L^\infty(\Gamma)$, i.e. for $\mathbf{y} \in \Gamma$

$$\mathbb{P}\left[Z \in \gamma \subset \Gamma\right] = \int_\gamma \rho(\mathbf{y})\, d\mathbf{y},$$

i.e. transform the measure $\mathbb{P}$ to $\mathbb{R}^N$

Quantities of interest (QoI)

Our goal of predicting the statistical behavior of a physical system often requires the approximation of multi-dimensional statistical QoIs, e.g.:

$$\mathbb{E}[u](x) = \int_\Gamma u(\mathbf{y}, x)\rho(\mathbf{y})d\mathbf{y}, \quad \text{where } \mathbf{y} \in \Gamma^N \text{ and } x \in \overline{D}$$

Applications to linear elliptic SPDEs

Parametrized equivalent (deterministic) formulation

- By **Lax-Milgram** $\exists!\ u \in H_P = L^2_P(\Omega; H^1_0(D))$ to the linear SPDE s.t.

$$\|u\|_{H_P} \leq \frac{C_P}{a_{min}} \left(\int_D \mathbb{E}[f^2]\, dx \right)^{1/2}$$

Strong formulation: find $u(\mathbf{y}, x) \in H_\rho = L^2_\rho(\Gamma; H^1_0(D))$ s.t.

$$\begin{cases} -\nabla \cdot (a(\mathbf{y}, x)\nabla u(\mathbf{y}, x)) = f(\mathbf{y}, x) & \text{for a.e. } x \in D, \\ u(\mathbf{y}, x) = 0 & \text{for a.e. } x \in \partial D, \end{cases}$$

where $\mathbf{y} \in \Gamma \subset \mathbb{R}^N$ and $x \in \overline{D}$

Weak formulation: find $u \in H_\rho$ s.t., $\forall v \in H_\rho$

$$\mathbb{E}\left[\int_D a(\mathbf{y}, x)\nabla u(\mathbf{y}, x) \cdot \nabla v(\mathbf{y}, x)\, dx \right] = \mathbb{E}\left[\int_D f(\mathbf{y}, x) \cdot v(\mathbf{y}, x)\, dx \right]$$

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

Applications to linear elliptic SPDEs

Parametrized equivalent (deterministic) formulation

- By **Lax-Milgram** $\exists!$ $u \in H_P = L^2_P(\Omega; H^1_0(D))$ to the linear SPDE s.t.

$$\|u\|_{H_P} \le \frac{C_P}{a_{min}} \left(\int_D \mathbb{E}[f^2]\, dx \right)^{1/2}$$

Strong formulation: find $u(\mathbf{y}, x) \in H_\rho = L^2_\rho(\Gamma; H^1_0(D))$ s.t.

$$\begin{cases} -\nabla \cdot (a(\mathbf{y}, x) \nabla u(\mathbf{y}, x)) = f(\mathbf{y}, x) & \text{for a.e. } x \in D, \\ u(\mathbf{y}, x) = 0 & \text{for a.e. } x \in \partial D, \end{cases}$$

where $\mathbf{y} \in \Gamma \subset \mathbb{R}^N$ and $x \in \overline{D}$

Weak formulation: find $u \in H_\rho$ s.t., $\forall v \in H_\rho$

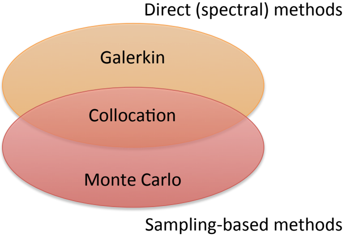
$$\mathbb{E}\left[\int_D a(\mathbf{y}, x) \nabla u(\mathbf{y}, x) \cdot \nabla v(\mathbf{y}, x)\, dx \right] = \mathbb{E}\left[\int_D f(\mathbf{y}, x) \cdot v(\mathbf{y}, x)\, dx \right]$$

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

A brief taxonomy of methods

For numerical soln. of PDEs with random input data

Stochastic finite element methods (SFEMs)



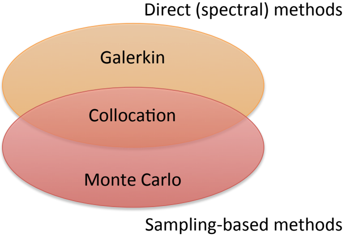
- methods for which spatial discretization is effected using finite element methods (FEMs)†
- Stochastic sampling methods (SSMs): random samples in Γ of PDE inputs are used to compute ensemble averages of statistical QoIs, e.g. MCFEM - *non-intrusive*

- 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) - *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*

A brief taxonomy of methods

For numerical soln. of PDEs with random input data

Stochastic finite element methods (SFEMs)

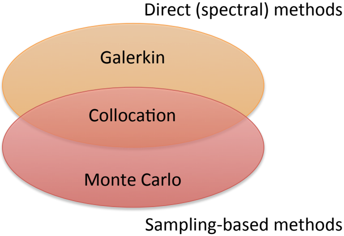


- methods for which spatial discretization is effected using finite element methods (FEMs)[†]
- Stochastic sampling methods (SSMs): random samples in Γ of PDE inputs are used to compute ensemble averages of statistical QoIs, e.g. MCFEM - *non-intrusive*
- Stochastic polynomial approximation
  1. Stochastic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

A brief taxonomy of methods

For numerical soln. of PDEs with random input data

Stochastic finite element methods (SFEMs)



- methods for which spatial discretization is effected using finite element methods (FEMs)†
- Stochastic sampling methods (SSMs): random samples in Γ 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) - *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*

NOTICE

1. **Spatial discretization**:
   †Throughout, we assume that spatial discretization is effected using finite element methods; most of what we say also holds for other spatial discretization approaches, e.g., finite differences, finite volumes, spectral, etc.
2. **Probabilistic discretization**:
   Throughout, we assume that the probabilistic discretization is effected using globally supported spectral or interpolatory basis functions. In this sense adaptive refinement refers to the anisotropic polynomial order used by the global basis.
   - *Adaptive wavelet stochastic collocation method for non-smooth solutions of SPDEs*, Max Gunzburger, Clayton Webster, Guannan Zhang
     Wednesday 2:30pm: MS53 - Recent Advances in Numerical SPDEs

Monte Carlo FEM (MCFEM)

Approximation statistics of QoIs $Q(u(\mathbf{y},x))$

1. **Classical approach**: Choose a number of realizations, $M \in \mathbb{N}_+$, and let $\{\mathbf{y}_k\}_{k=1}^{M}$ be a given sample set of random abscissas
2. For each $k = 1,\ldots,M$ sample iid realizations of the diffusion $a(\mathbf{y}_k,x)$, the load $f(\mathbf{y}_k,x)$ and find a FEM approximation $u^h(\mathbf{y}_k,\cdot) \in W_h(D)$ s.t.

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

   If desired evaluate the QoI $Q\left(u^h(\mathbf{y}_k,\cdot)\right)$
3. Approximate statistics, e.g. expectations $\mathbb{E}[u^h](x)$, by sample averages:

$$\mathbb{E}\left[u^h(\mathbf{y})\right](x) \approx \frac{1}{M}\sum_{k=1}^{M} u^h(\mathbf{y}_k)\rho(\mathbf{y}_k) := \mathscr{E}\left(u^h;M\right), \quad \mathbf{y}_k \in \Gamma$$

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

$$\left\|\mathbb{E}[u^h] - \mathscr{E}\left(u^h;M\right)\right\| \leq \text{TOL}$$

Monte Carlo FEM (MCFEM)

Approximation statistics of QoIs $Q(u(\mathbf{y},x))$

1. **Classical approach**: Choose a number of realizations, $M \in \mathbb{N}_+$, and let $\{\mathbf{y}_k\}_{k=1}^{M}$ be a given sample set of random abscissas
2. For each $k = 1,\dots,M$ sample iid realizations of the diffusion $a(\mathbf{y}_k,x)$, the load $f(\mathbf{y}_k,x)$ and find a FEM approximation $u^h(\mathbf{y}_k,\cdot) \in W_h(D)$ s.t.

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

   If desired evaluate the QoI $Q\left(u^h(\mathbf{y}_k,\cdot)\right)$
3. Approximate statistics, e.g. expectations $\mathbb{E}[u^h](x)$, by sample averages:

$$\mathbb{E}\left[u^h(\mathbf{y})\right](x) \approx \frac{1}{M}\sum_{k=1}^{M} u^h(\mathbf{y}_k)\rho(\mathbf{y}_k) := \mathscr{E}\left(u^h;M\right), \quad \mathbf{y}_k \in \Gamma$$

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

$$\left\|\mathbb{E}[u^h] - \mathscr{E}\left(u^h;M\right)\right\| \leq \text{TOL}$$

Monte Carlo FEM (MCFEM)

Approximation statistics of QoIs $Q(u(\mathbf{y},x))$

1. **Classical approach**: Choose a number of realizations, $M \in \mathbb{N}_+$, and let $\{\mathbf{y}_k\}_{k=1}^{M}$ be a given sample set of random abscissas
2. For each $k = 1,\ldots,M$ sample iid realizations of the diffusion $a(\mathbf{y}_k,x)$, the load $f(\mathbf{y}_k,x)$ and find a FEM approximation $u^h(\mathbf{y}_k,\cdot) \in W_h(D)$ s.t.

$$\begin{cases} -\nabla\cdot(a(\mathbf{y}_k,\cdot)\nabla u^h(\mathbf{y}_k,\cdot)) = & f(\mathbf{y}_k,\cdot), \quad \text{in } D \\ u^h(\mathbf{y}_k,\cdot) = & 0, \quad \text{on } \partial D \end{cases}$$

   If desired evaluate the QoI $Q\left(u^h(\mathbf{y}_k,\cdot)\right)$
3. Approximate statistics, e.g. expectations $\mathbb{E}[u^h](x)$, by sample averages:

$$\mathbb{E}\left[u^h(\mathbf{y})\right](x) \approx \frac{1}{M}\sum_{k=1}^{M} u^h(\mathbf{y}_k)\rho(\mathbf{y}_k) := \mathscr{E}\left(u^h;M\right), \quad \mathbf{y}_k \in \Gamma$$

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

$$\left\|\mathbb{E}[u^h] - \mathscr{E}\left(u^h;M\right)\right\| \leq \text{TOL}$$

Monte Carlo FEM (MCFEM)

Approximation statistics of QoIs $Q(u(\mathbf{y},x))$

1. **Classical approach**: Choose a number of realizations, $M \in \mathbb{N}_+$, and let $\{\mathbf{y}_k\}_{k=1}^{M}$ be a given sample set of random abscissas
2. For each $k = 1, \ldots, M$ sample iid realizations of the diffusion $a(\mathbf{y}_k, x)$, the load $f(\mathbf{y}_k, x)$ and find a FEM approximation $u^h(\mathbf{y}_k, \cdot) \in W_h(D)$ s.t.

$$\begin{cases} -\nabla \cdot (a(\mathbf{y}_k,\cdot)\nabla u^h(\mathbf{y}_k,\cdot)) = & f(\mathbf{y}_k,\cdot), \quad \text{in } D \\ u^h(\mathbf{y}_k,\cdot) = & 0, \quad \text{on } \partial D \end{cases}$$

   If desired evaluate the QoI $Q\left(u^h(\mathbf{y}_k,\cdot)\right)$
3. Approximate statistics, e.g. expectations $\mathbb{E}[u^h](x)$, by sample averages:

$$\mathbb{E}\left[u^h(\mathbf{y})\right](x) \approx \frac{1}{M}\sum_{k=1}^{M} u^h(\mathbf{y}_k)\rho(\mathbf{y}_k) := \mathscr{E}\left(u^h; M\right), \quad \mathbf{y}_k \in \Gamma$$

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

$$\left\|\mathbb{E}[u^h] - \mathscr{E}\left(u^h; M\right)\right\| \leq \text{TOL}$$

Convergence of the MCFEM

Error splitting

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

- **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.

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

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

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

Convergence of the MCFEM

Error splitting

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

- **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.

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

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

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

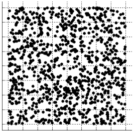
Other sampling-based methods

Attempting to cope with the *curse of dimensionality*

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

$$\mathbb{E}[Q(u(\mathbf{y}))] \approx \frac{1}{M}\sum_{k=1}^{M} Q(u(\mathbf{y}_k))\rho(\mathbf{y}_k), \quad \mathbf{y}_k \in \Gamma$$

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



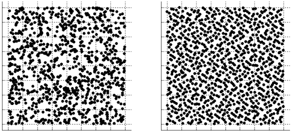
Other sampling-based methods

Attempting to cope with the *curse of dimensionality*

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

$$\mathbb{E}[Q(u(\mathbf{y}))] \approx \frac{1}{M}\sum_{k=1}^{M} Q(u(\mathbf{y}_k))\rho(\mathbf{y}_k), \quad \mathbf{y}_k \in \Gamma$$

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



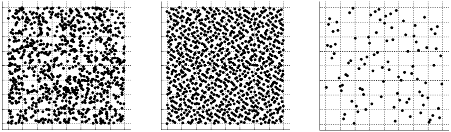
Other sampling-based methods

Attempting to cope with the *curse of dimensionality*

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

$$\mathbb{E}[Q(u(\mathbf{y}))] \approx \frac{1}{M}\sum_{k=1}^{M} Q(u(\mathbf{y}_k))\rho(\mathbf{y}_k), \quad \mathbf{y}_k \in \Gamma$$

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



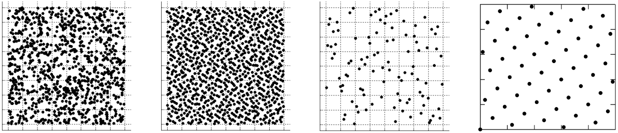
Other sampling-based methods

Attempting to cope with the *curse of dimensionality*

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

$$\mathbb{E}[Q(u(\mathbf{y}))] \approx \frac{1}{M}\sum_{k=1}^{M} Q(u(\mathbf{y}_k))\rho(\mathbf{y}_k), \quad \mathbf{y}_k \in \Gamma$$

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



Other sampling-based methods

Attempting to cope with the *curse of dimensionality*

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

$$\mathbb{E}[Q(u(\mathbf{y}))] \approx \frac{1}{M}\sum_{k=1}^{M} Q(u(\mathbf{y}_k))\rho(\mathbf{y}_k), \quad \mathbf{y}_k \in \Gamma$$

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

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

Other sampling-based methods

Attempting to cope with the *curse of dimensionality*

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

$$\mathbb{E}[Q(u(\mathbf{y}))] \approx \frac{1}{M}\sum_{k=1}^{M} Q(u(\mathbf{y}_k))\rho(\mathbf{y}_k), \quad \mathbf{y}_k \in \Gamma$$

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

Regularity

With respect to *the noise*

- Let $\Gamma_n^* = \prod_{\substack{j=1 \\ j\neq n}}^N \Gamma_j$, and let $\mathbf{y}_n^*$ denote an arbitrary element of Γ_n^*

Theorem: Regularity [Linear: BNT07], [Nonlinear: W07, GW11]

For each $y_n \in \Gamma_n$, there exists $\tau_n > 0$ such that the function $u(y_n, \mathbf{y}_n^*, x)$ as a function of y_n, $u : \Gamma_n \to C^0(\Gamma_n^*; W(D))$ admits an analytic extension $u(z, \mathbf{y}_n^*, x)$, $z \in \mathbb{C}$, in the region of the complex plane

$$\Sigma(\Gamma_n; \tau_n) \equiv \{z \in \mathbb{C}, \ \text{dist}(z, \Gamma_n) \leq \tau_n\}.$$

Moreover, $\forall z \in \Sigma(\Gamma_n; \tau_n)$,

$$\|u(z)\|_{C^0(\Gamma_n^*; W(D))} \leq \lambda$$

with λ a constant independent of n.

Remark: The analyticity of the solution $u(\mathbf{y}, x)$ w.r.t. each random direction y_n suggests the use of (multivariate) polynomial approximation

Region of analyticity

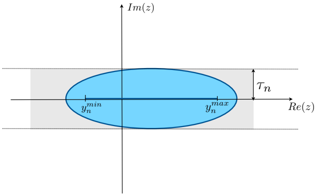
An example: bounded RVs [Babuska et. al 2007]

- Assume a_N is an exponential Karhunen-Loève expansion and f_N deterministic: $a_N(\omega, x) = a_{min} + e^{b_0(x) + \sum_{n=1}^{N} \sqrt{\lambda_n}\, b_n(x) Y_n(\omega)}$
- Γ_n bounded: $\Gamma_n = \left[y_n^{min}, y_n^{max}\right]$

The analyticity region is given by:
$\Sigma(\Gamma_n; \tau_n) = \{z \in \mathbb{C} : \ |Im(z)| \le \tau_n\}$,

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

$\delta = 4$ (linear), $\delta = 12$ (nonlinear, $k = 1$)



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

$$0 < g_n = \log \left[\frac{2\tau_n}{|\Gamma_n|} + \sqrt{1 + \frac{4{\tau_n}^2}{|\Gamma_n|^2}} \right]$$

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

Multivariate polynomial approximation

With respect to the *noise*

- The analyticity of the solution $u(\mathbf{y}, x)$ 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 $Y_1(\omega), \ldots, Y_N(\omega) \Rightarrow$ Possible 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?

E.g. Curse of dimensionality: (Isotropic) TP's of degree p in N dimensions

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

$\Updownarrow$

error $\leq Ce^{-gM^{\frac{1}{N}}}$ Impractical in higher dimensions

Multivariate polynomial approximation

With respect to the *noise*

- The analyticity of the solution $u(\mathbf{y}, x)$ 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 $Y_1(\omega), \ldots, Y_N(\omega) \Rightarrow$ Possible 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?

E.g. Curse of dimensionality: (Isotropic) TP's of degree p in N dimensions

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

$\Updownarrow$

error $\leq Ce^{-gM^{\frac{1}{N}}}$ Impractical in higher dimensions

Multivariate polynomial approximation

Fully discrete *stochastic* solutions

Basic idea: approximate the response $u(\mathbf{y}, \cdot)$ by multi-variate global polynomials. The numerical solution should convergence quickly since the solution is analytic in $\mathbf{y}$.

Approximating spaces:

1. Let $\mathcal{T}_h$ be a triangulation of D and $W^h(D) \subset W(D)$ contains cont. piecewise polynomials defined in $\mathcal{T}_h$
   - Assume $J = \dim[W^h(D)]$ and $\{\phi_j(x)\}_{j=1}^J \subset W_h(D)$ is a FE basis for the deterministic domain
2. Let $\mathbf{p} = (p_1, \ldots, p_N)$ be a multi-index, $\mathcal{J}(p) \subset \mathbb{N}^N$ a multi-index set, with $p \in \mathbb{N}_+$, and define:

Multivariate polynomial space

$$\mathcal{P}_{\mathcal{J}(p)}(\Gamma) = \text{span}\left\{ \prod_{n=1}^N y_n^{p_n}, \quad \text{with } \mathbf{p} \in \mathcal{J}(p) \right\} \subset L^2_\rho(\Gamma)$$

- Assume $M = \dim\left[\mathcal{P}_{\mathcal{J}(p)}(\Gamma)\right]$ and $\{\psi_k\}_{k=1}^M$ form a basis for $\mathcal{P}_{\mathcal{J}(p)}(\Gamma)$, e.g. multivariate Legendre, Hermite, Lagrange, etc.

Multivariate polynomial approximation

Fully discrete *stochastic* solutions

Basic idea: approximate the response $u(\mathbf{y}, \cdot)$ by multi-variate global polynomials. The numerical solution should convergence quickly since the solution is analytic in $\mathbf{y}$.

Approximating spaces:

1. Let $\mathcal{T}_h$ be a triangulation of D and $W^h(D) \subset W(D)$ contains cont. piecewise polynomials defined in $\mathcal{T}_h$
   - Assume $J = \dim[W^h(D)]$ and $\{\phi_j(x)\}_{j=1}^J \subset W_h(D)$ is a FE basis for the deterministic domain
2. Let $\mathbf{p} = (p_1, \ldots, p_N)$ be a multi-index, $\mathcal{J}(p) \subset \mathbb{N}^N$ a multi-index set, with $p \in \mathbb{N}_+$, and define:

Multivariate polynomial space

$$\mathcal{P}_{\mathcal{J}(p)}(\Gamma) = \operatorname{span}\left\{\prod_{n=1}^N y_n^{p_n}, \quad \text{with } \mathbf{p} \in \mathcal{J}(p)\right\} \subset L^2_\rho(\Gamma)$$

- Assume $M = \dim\left[\mathcal{P}_{\mathcal{J}(p)}(\Gamma)\right]$ and $\{\psi_k\}_{k=1}^M$ form a basis for $\mathcal{P}_{\mathcal{J}(p)}(\Gamma)$, e.g. multivariate Legendre, Hermite, Lagrange, etc.

Multivariate polynomial approximation

Fully discrete *stochastic* solutions II

Multivariate polynomial space

$$\mathcal{P}_{\mathcal{J}(p)}(\Gamma) = \text{span}\left\{\prod_{n=1}^{N} y_n^{p_n}, \quad \text{with } \mathbf{p} \in \mathcal{J}(p)\right\} \subset L^2_\rho(\Gamma)$$

Fully discrete approximations: $u_p \in \mathcal{P}_{\mathcal{J}(p)}(\Gamma) \otimes W^h(D)$ s.t.

$$u_p(\mathbf{y}, x) = \sum_{j=1}^{J}\sum_{k=1}^{M} c_{jk}\phi_j(x)\psi_k(\mathbf{y}) = \sum_{k=1}^{M} u_k(x)\psi_k(\mathbf{y}), \quad \text{with } u_k(x) \in W^h(D)$$

- To compute the fully discrete approximation using SFEMs requires the resolution of the coefficients u_k which can be accomplished via:
  - intrusive methods by solving the fully coupled $JM \times JM$ system (i.e. JM equations and JM degrees of freedom
  - non-intrusive methods by de-coupling the above expression and solving a M system's of size J
- *Our goal* is to estimate the error $\|u - u_p\|_{L^2_\rho(\Gamma)}$

Multivariate polynomial approximation

Fully discrete *stochastic* solutions II

Multivariate polynomial space

$$\mathcal{P}_{\mathcal{J}(p)}(\Gamma) = \text{span}\left\{\prod_{n=1}^{N} y_n^{p_n}, \quad \text{with } \mathbf{p} \in \mathcal{J}(p)\right\} \subset L^2_\rho(\Gamma)$$

Fully discrete approximations: $u_p \in \mathcal{P}_{\mathcal{J}(p)}(\Gamma) \otimes W^h(D)$ s.t.

$$u_p(\mathbf{y}, x) = \sum_{j=1}^{J} \sum_{k=1}^{M} c_{jk}\phi_j(x)\psi_k(\mathbf{y}) = \sum_{k=1}^{M} u_k(x)\psi_k(\mathbf{y}), \quad \text{with } u_k(x) \in W^h(D)$$

- To compute the fully discrete approximation using SFEMs requires the resolution of the coefficients u_k which can be accomplished via:
  - intrusive methods by solving the fully coupled $JM \times JM$ system (i.e. JM equations and JM degrees of freedom
  - non-intrusive methods by de-coupling the above expression and solving a M system's of size J
- *Our goal* is to estimate the error $\|u - u_p\|_{L^2_\rho(\Gamma)}$

Polynomial spaces

Anisotropic representations

Rewrite the fully discrete approximation using the multi-index notation:

$$u_p = \sum_{\mathbf{p} \in \mathcal{J}(p)} u_{\mathbf{p}}(x) \psi_{\mathbf{p}}(\mathbf{y}), \quad \mathbf{p} = (p_1, \ldots, p_N)$$

where, if $\mathbf{y} = (y_1, \ldots, y_N)$ independent then

$$\psi_{\mathbf{p}}(\mathbf{y}) = \prod_{n=1}^{N} \psi_{p_n}^{(n)}(y_n), \quad \psi_{p_n}^{(n)} \in L^2_{\rho_n}(\Gamma_n)$$

Several choices for polynomial multi-index $\mathbf{p} \in \mathcal{J}(p)$ [BNTT11]:

- Tensor products (TP): $\max_n \alpha_n p_n \leq p$ (Intractable for large N),
- Total degree (TD): $\sum_{n=1}^{N} \alpha_n p_n \leq p$,
- Hyperbolic cross (HC): $\prod_{n=1}^{N} (p_n + 1)^{\alpha_n} \leq p + 1$,
- Smolyak method (SM): $\sum_{n=1}^{N} \alpha_n f(p_n) \leq f(p)$ 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 $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N) \in \mathbb{R}^N_+$, with $\alpha_{min} = 1$

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

Polynomial spaces

Anisotropic representations

Rewrite the fully discrete approximation using the multi-index notation:

$$u_p = \sum_{\mathbf{p} \in \mathcal{J}(p)} u_{\mathbf{p}}(x)\psi_{\mathbf{p}}(\mathbf{y}), \quad \mathbf{p} = (p_1, \ldots, p_N)$$

where, if $\mathbf{y} = (y_1, \ldots, y_N)$ independent then

$$\psi_{\mathbf{p}}(\mathbf{y}) = \prod_{n=1}^{N} \psi_{p_n}^{(n)}(y_n), \quad \psi_{p_n}^{(n)} \in L^2_{\rho_n}(\Gamma_n)$$

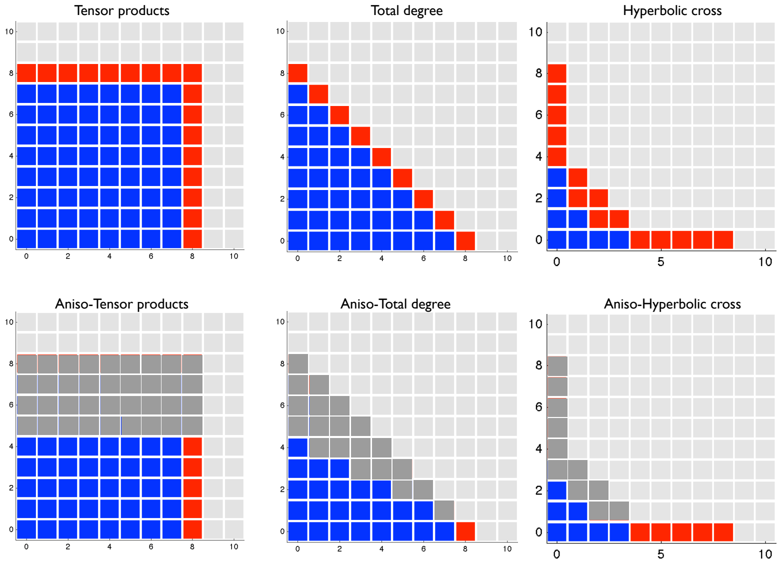
Several choices for polynomial multi-index $\mathbf{p} \in \mathcal{J}(p)$ [BNTT11]:

- Tensor products (TP): $\max_n \alpha_n p_n \leq p$ (Intractable for large N),
- Total degree (TD): $\sum_{n=1}^{N} \alpha_n p_n \leq p$,
- Hyperbolic cross (HC): $\prod_{n=1}^{N} (p_n + 1)^{\alpha_n} \leq p + 1$,
- Smolyak method (SM): $\sum_{n=1}^{N} \alpha_n f(p_n) \leq f(p)$ 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 $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N) \in \mathbb{R}^N_+$, with $\alpha_{min} = 1$
TD, HC & SM all reduce the *curse of dimensionality* w.r.t. TP methods

Example (anisotropic) polynomial spaces

$N = 2$, $p = 8$ with $\boldsymbol{\alpha} = (1, 2)$



Example: $N = 2$ with monomial basis

TD space vs. TP space

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

$p_1 + p_2 \le 0$					1				
$p_1 + p_2 \le 1$				y_1		y_2			
$p_1 + p_2 \le 2$			y_1^2		$y_1 y_2$		y_2^2		
$p_1 + p_2 \le 3$		y_1^3		$y_1^2 y_2$		$y_1 y_2^2$		y_2^3	
$p_1 + p_2 \le 4$	y_1^4		$y_1^3 y_2$		$y_1^2 y_2^2$		$y_1 y_2^3$		y_2^4
		$y_1^4 y_2$		$y_1^3 y_2^2$		$y_1^2 y_2^3$		$y_1 y_2^4$	
			$y_1^4 y_2^2$		$y_1^3 y_2^3$		$y_1^2 y_2^4$		
				$y_1^4 y_2^3$		$y_1^3 y_2^4$			
$\max(p_1, p_2) \le 4$					$y_1^4 y_2^4$				

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

Recall $M = \dim\left[\mathcal{P}_{\mathcal{J}(p)}(\Gamma)\right] \Longrightarrow M_{TD} = \frac{(N+p)!}{N!p!} << M_{TP} = (p+1)^N$

Example: $N = 2$ with monomial basis

TD space vs. TP space

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

$p_1 + p_2 \leq 0$					1				
$p_1 + p_2 \leq 1$				y_1		y_2			
$p_1 + p_2 \leq 2$			y_1^2		$y_1 y_2$		y_2^2		
$p_1 + p_2 \leq 3$		y_1^3		$y_1^2 y_2$		$y_1 y_2^2$		y_2^3	
$p_1 + p_2 \leq 4$	y_1^4		$y_1^3 y_2$		$y_1^2 y_2^2$		$y_1 y_2^3$		y_2^4
		$y_1^4 y_2$		$y_1^3 y_2^2$		$y_1^2 y_2^3$		$y_1 y_2^4$	
			$y_1^4 y_2^2$		$y_1^3 y_2^3$		$y_1^2 y_2^4$		
				$y_1^4 y_2^3$		$y_1^3 y_2^4$			
$\max(p_1, p_2) \leq 4$					$y_1^4 y_2^4$				

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

Recall $M = \dim \left[\mathcal{P}_{\mathcal{J}(p)}(\Gamma)\right] \Longrightarrow M_{TD} = \frac{(N+p)!}{N!p!} << M_{TP} = (p+1)^N$

General basis in N dimensions

Total degree vs. Tensor products

N = # RVs, $\dim(\Gamma)$	p = maximal degree of polynomials	K = 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

Stochsatic Galerkin (SG) FEM

Galerkin projection onto the subspace $\mathcal{P}_{\mathcal{J}(p)}(\Gamma) \otimes W^h(D)$

Stochastic Galerkin approximation

Find $u_p^{SG} \in \mathcal{P}_{\mathcal{J}(p)}(\Gamma) \otimes W^h(D)$ s.t.

$$\mathbb{E}\left[\int_D a(\mathbf{y}, x) \nabla u_p^{SG}(\mathbf{y}, x) \cdot \nabla v(\mathbf{y}, x)\, dx\right] = \mathbb{E}\left[\int_D f(\mathbf{y}, x) \cdot v(\mathbf{y}, x)\, dx\right]$$

for all $v \in \mathcal{P}_{\mathcal{J}(p)}(\Gamma) \otimes W^h(D)$

- Typically u_p^{SG} is defined using an L^2_ρ-orthogonal basis $\{\psi_k\}_{k=1}^{M_{SG}}$ constructed from univariate $L^2_{\rho_n}$-orthogonal polynomials - i.e. assuming independent RVs, e.g. $\rho(\mathbf{y}) = \prod_{n=1}^{N} \rho_n(y_n)$ -Wiener (polynomial) chaos
- Optimal index set of cardinality M_{TD} corresponds to the TD subspace

---

[Ghanem-Spanos], [Karniadakis-Xiu], [Matthies-Keese], [Schwab-Todor et. al], [Knio-Le Maître et. al], [Babuška et. al], etc.

Stochsatic Galerkin (SG) FEM

Galerkin projection onto the subspace $\mathcal{P}_{\mathcal{J}(p)}(\Gamma) \otimes W^h(D)$

Stochastic Galerkin approximation

Find $u_p^{SG} \in \mathcal{P}_{\mathcal{J}(p)}(\Gamma) \otimes W^h(D)$ s.t.

$$\mathbb{E}\left[\int_D a(\mathbf{y},x)\nabla u_p^{SG}(\mathbf{y},x) \cdot \nabla v(\mathbf{y},x)\,dx\right] = \mathbb{E}\left[\int_D f(\mathbf{y},x) \cdot v(\mathbf{y},x)\,dx\right]$$

for all $v \in \mathcal{P}_{\mathcal{J}(p)}(\Gamma) \otimes W^h(D)$

- Typically u_p^{SG} is defined using an L^2_ρ-orthogonal basis $\{\psi_k\}_{k=1}^{M_{SG}}$ constructed from univariate $L^2_{\rho_n}$-orthogonal polynomials - i.e. assuming independent RVs, e.g. $\rho(\mathbf{y}) = \prod_{n=1}^N \rho_n(y_n)$ -Wiener (polynomial) chaos
- Optimal index set of cardinality M_{TD} corresponds to the TD subspace

Pros: Guaranteed L^2 optimality of the projection and spectral convergence for smooth stochastic solutions

[Ghanem-Spanos], [Karniadakis-Xiu], [Matthies-Keese], [Schwab-Todor et. al], [Knio-Le Maître et. al], [Babuška et. al], etc.

Stochsatic Galerkin (SG) FEM

Galerkin projection onto the subspace $\mathcal{P}_{\mathcal{J}(p)}(\Gamma) \otimes W^h(D)$

Stochastic Galerkin approximation

Find $u_p^{SG} \in \mathcal{P}_{\mathcal{J}(p)}(\Gamma) \otimes W^h(D)$ s.t.

$$\mathbb{E}\left[\int_D a(\mathbf{y}, x)\nabla u_p^{SG}(\mathbf{y}, x) \cdot \nabla v(\mathbf{y}, x)\, dx\right] = \mathbb{E}\left[\int_D f(\mathbf{y}, x) \cdot v(\mathbf{y}, x)\, dx\right]$$

for all $v \in \mathcal{P}_{\mathcal{J}(p)}(\Gamma) \otimes W^h(D)$

- Typically u_p^{SG} is defined using an L^2_ρ-orthogonal basis $\{\psi_k\}_{k=1}^{M_{SG}}$ constructed from univariate $L^2_{\rho_n}$-orthogonal polynomials - i.e. assuming independent RVs, e.g. $\rho(\mathbf{y}) = \prod_{n=1}^N \rho_n(y_n)$ -Wiener (polynomial) chaos
- Optimal index set of cardinality M_{TD} corresponds to the TD subspace

Cons: Inherently a fully coupled *intrusive* approximation that requires the resolution of **large** systems (integrals and residuals) as well as novel (problem specific) pre-conditioning techniques

---

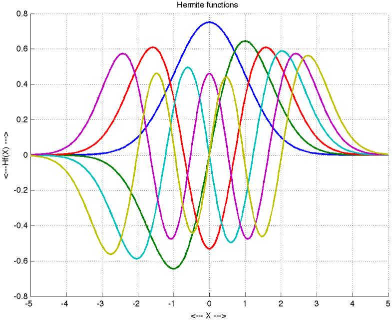
[Ghanem-Spanos], [Karniadakis-Xiu], [Matthies-Keese], [Schwab-Todor et. al], [Knio-Le Maître et. al], [Babuška et. al], etc.

SGFEM

Spectral orthogonal basis

The L^2_ρ-orthogonal basis was originally developed 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 $\rho(y) = \frac{1}{\sqrt{2\pi}} e^{\frac{-y^2}{2}}$

Hermite polynomials

Generalized Multi-dimensional orthonormal polynomials

Let $\left\{H_{p_n}^{(n)}\right\}_{p_n=0}^{p}$ denote the set of univariate Hermite polynomials (of degree $\le p$) defined in $L^2_{\rho_n}(\Gamma_n)$, that are orthonormal w.r.t. the Gaussian PDF $\rho_n(y_n)$, for each $n = 1, \ldots, N$:

$$\int_{\Gamma_n} H_{p_n}^{(n)}(y_n) H_{r_n}^{(n)}(y_n) \rho_n(y_n)\, dy_n = \delta_{p_n r_n}, \quad p_n, r_n \in \{0, \ldots, p\}$$

The multi-variate $L^2_\rho(\Gamma)$-orthogonal Hermite basis is defined as a tensor-product of the univariate polynomials with $\mathbf{p} \in \mathcal{J}(p)$:

$$H_{\mathbf{p}}(\mathbf{y}) = \prod_{n=1}^{N} H_{p_n}^{(n)}(y_n), \quad \text{s.t. } \rho(\mathbf{y}) = \prod_{n=1}^{N} \rho_n(y_n),$$

where $\rho(\mathbf{y})$ is the Gaussian joint-PDF

- identical construction for other orthonormal bases (generalized PC) e.g. $\mathbf{Y}$ uniform RVs $\to$ Legendre polynomial basis, etc.

Hermite polynomials

Generalized Multi-dimensional orthonormal polynomials

Let $\left\{H_{p_n}^{(n)}\right\}_{p_n=0}^{p}$ denote the set of univariate Hermite polynomials (of degree $\leq p$) defined in $L^2_{\rho_n}(\Gamma_n)$, that are orthonormal w.r.t. the Gaussian PDF $\rho_n(y_n)$, for each $n = 1, \ldots, N$:

$$\int_{\Gamma_n} H_{p_n}^{(n)}(y_n) H_{r_n}^{(n)}(y_n) \rho_n(y_n)\, dy_n = \delta_{p_n r_n}, \quad p_n, r_n \in \{0, \ldots, p\}$$

The multi-variate $L^2_\rho(\Gamma)$-orthogonal Hermite basis is defined as a tensor-product of the univariate polynomials with $\mathbf{p} \in \mathcal{J}(p)$:

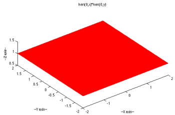
$$H_{\mathbf{p}}(\mathbf{y}) = \prod_{n=1}^{N} H_{p_n}^{(n)}(y_n), \quad \text{s.t. } \rho(\mathbf{y}) = \prod_{n=1}^{N} \rho_n(y_n),$$

where $\rho(\mathbf{y})$ is the Gaussian joint-PDF

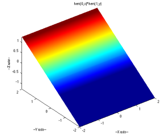
- identical construction for other orthonormal bases (generalized PC) e.g. $\mathbf{Y}$ uniform RVs $\rightarrow$ Legendre polynomial basis, etc.

$(N = 2,\, p = 5)$ Hermite polynomials

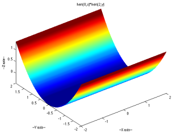
TD subspace: $p_1 + p_2 \leq 5$



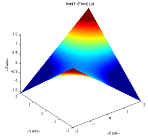
$H_{(0,0)}$



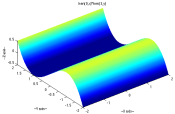
$H_{(0,1)}, H_{(1,0)}$



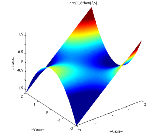
$H_{(0,2)}, H_{(2,0)}$



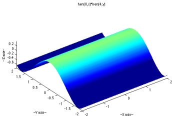
$H_{(1,1)}$



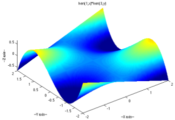
$H_{(0,3)}, H_{(3,0)}$



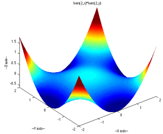
$H_{(1,2)}, H_{(2,1)}$



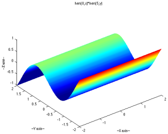
$H_{(0,4)}, H_{(4,0)}$



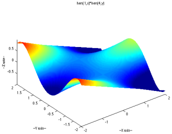
$H_{(1,3)}, H_{(3,1)}$



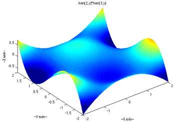
$H_{(2,2)}$



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



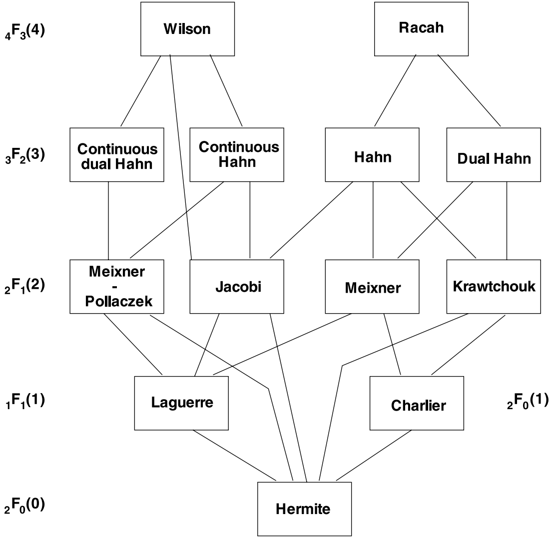
$H_{(1,4)}, H_{(4,1)}$



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

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 $L_n^{(\alpha)}(y)$	$[0, \infty]$

SGFEM

Applications to linear elliptic SPDEs

- Leads to a single large coupled system (of size $JM \times JM$ - typically $M = M_{TD}$) which requires preconditioning methods, e.g. CG [Ghanem-Pellisetti], [Helman-Powell et.al], [Ullman et.al] etc.
  - most preconditioning methods are very problem specific
- terms in the coupled matrix will require computations of the form:

$$A_{kl}(x) = \mathbb{E}\left[a(\cdot, x)\psi_k\psi_l\right] = \int_\Gamma a(\mathbf{y}, x)\psi_k(\mathbf{y})\psi_l(\mathbf{y})\rho(\mathbf{y})\, d\mathbf{y}$$

  - could be High-Dimensional integration problem!
    If $a(\mathbf{y}, x)$ is nonlinear in $\mathbf{y}$ or if $\mathbf{Y}(\omega)$ is not independent then this is certainly possible
- For linear problems and independent RV's, i.e. $\rho(\mathbf{y}) = \prod_{n=1}^{N} \rho_n(y_n)$, using the multivariate orthonormal basis - computations simplify significantly
  - sparse matrix approaches can be exploited

SGFEM

Applications to linear elliptic SPDEs

- Leads to a single large coupled system (of size $JM \times JM$ - typically $M = M_{TD}$) which requires preconditioning methods, e.g. CG [Ghanem-Pellisetti], [Helman-Powell et.al], [Ullman et.al] etc.
  - most preconditioning methods are very problem specific
- terms in the coupled matrix will require computations of the form:

$$A_{kl}(x) = \mathbb{E}\left[a(\cdot, x)\psi_k\psi_l\right] = \int_\Gamma a(\mathbf{y}, x)\psi_k(\mathbf{y})\psi_l(\mathbf{y})\rho(\mathbf{y})\,d\mathbf{y}$$

  - could be High-Dimensional integration problem!
    If $a(\mathbf{y}, x)$ is nonlinear in $\mathbf{y}$ or if $\mathbf{Y}(\omega)$ is not independent then this is certainly possible
- For linear problems and independent RV's, i.e. $\rho(\mathbf{y}) = \prod_{n=1}^{N} \rho_n(y_n)$, using the multivariate orthonormal basis - computations simplify significantly
  - sparse matrix approaches can be exploited

SGFEM

Applications to linear elliptic SPDEs

- Leads to a single large coupled system (of size $JM \times JM$ - typically $M = M_{TD}$) which requires preconditioning methods, e.g. CG [Ghanem-Pellisetti], [Helman-Powell et.al], [Ullman et.al] etc.
  - most preconditioning methods are very problem specific
- terms in the coupled matrix will require computations of the form:

$$A_{kl}(x) = \mathbb{E}\left[a(\cdot, x)\psi_k\psi_l\right] = \int_\Gamma a(\mathbf{y}, x)\psi_k(\mathbf{y})\psi_l(\mathbf{y})\rho(\mathbf{y})\, d\mathbf{y}$$

  - could be High-Dimensional integration problem!
    If $a(\mathbf{y}, x)$ is nonlinear in $\mathbf{y}$ or if $\mathbf{Y}(\omega)$ is not independent then this is certainly possible
- For linear problems and independent RV's, i.e. $\rho(\mathbf{y}) = \prod_{n=1}^{N} \rho_n(y_n)$, using the multivariate orthonormal basis - computations simplify significantly
  - sparse matrix approaches can be exploited

SGFEM

Linear problems with independent RVs

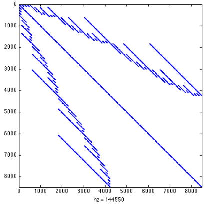
Recall from above:

$$A_{kl}(x) = \int_\Gamma a(\mathbf{y},x)\psi_k(\mathbf{y})\psi_l(\mathbf{y})\rho(\mathbf{y})\,d\mathbf{y}$$

$$= \mathbb{E}[a](x)\int_\Gamma \psi_k\psi_l\rho(\mathbf{y})d\mathbf{y} + \sum_{n=1}^N b_n(x)\int_\Gamma y_n\psi_k\psi_l\rho(\mathbf{y})d\mathbf{y}$$

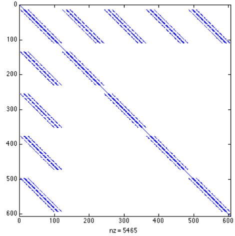
$$= \mathbb{E}[a](x)\delta_{kl} + \sum_{n=1}^N b_n(x)\prod_{m=1}^N \int_{\Gamma_m} y_n^{\delta_{nm}}\psi_{k_m}^{(m)}\psi_{l_m}^{(m)}\rho_m(y_m)dy_m$$

- matrix becomes dense when the coefficient and/or PDE is nonlinear
- condition number deteriorates when approximating using non-uniform RVs

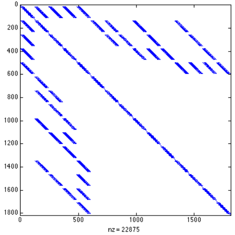


Applications to linear elliptic SPDEs

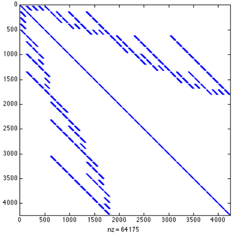
$N = 4$ Uniform RF with TP Legendre polynomials



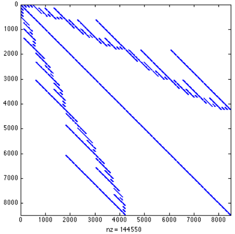
$p = 1$



$p = 2$



$p = 3$



$p = 4$

Non-intrusive approaches

We can use the orthogonality of the basis to yield an expression for the unknown coefficients - that is non-intrusive

$$\int_\Gamma u(\mathbf{y},\cdot)\psi_{k'}(\mathbf{y})\rho(\mathbf{y})\,d\mathbf{y} = \sum_{k}^{M_{SG}} u_k(x)\int_\Gamma \psi_k(\mathbf{y})\psi_{k'}(\mathbf{y})\rho(\mathbf{y})\,d\mathbf{y} = u_{k'},$$

$k' = 1, \ldots M_{SG}$.

- the integrals could be high dimensional - for each of the M_{SG} coefficients
- the convergence of $u_p \to u$ is dictated by the accurate calculation of the coefficients which becomes dominated by the "error" in the integration scheme

Instead we will consider a powerful *non-intrusive* alternative approach whose polynomial coefficients are the nodal values of the deterministic FEM, i.e. a sampling-based approach with the added benefit of remaining a polynomial approximation

Non-intrusive approaches

We can use the orthogonality of the basis to yield an expression for the unknown coefficients - that is non-intrusive

$$\int_\Gamma u(\mathbf{y},\cdot)\psi_{k'}(\mathbf{y})\rho(\mathbf{y})\,d\mathbf{y} = \sum_k^{M_{SG}} u_k(x)\int_\Gamma \psi_k(\mathbf{y})\psi_{k'}(\mathbf{y})\rho(\mathbf{y})\,d\mathbf{y} = u_{k'},$$

$k' = 1, \ldots M_{SG}$.

- the integrals could be high dimensional - for each of the M_{SG} coefficients
- the convergence of $u_p \to u$ is dictated by the accurate calculation of the coefficients which becomes dominated by the "error" in the integration scheme

Instead we will consider a powerful *non-intrusive* alternative approach whose polynomial coefficients are the nodal values of the deterministic FEM, i.e. a sampling-based approach with the added benefit of remaining a polynomial approximation

Stochastic Collocation (SC) FEM

General description

1. Choose a set of points $H_M = \{\mathbf{y}_k \in \Gamma\}_{k=1}^M$
2. For each k solve the FE solution $u_k(x) = u(\mathbf{y}_k, x)$, given $a_k(x) = a(\mathbf{y}_k, x)$ and $f_k(x) = f(\mathbf{y}_k, x)$
3. Interpolate the sampled values: $u_p(\mathbf{y}, x) = \sum_{k=1}^M u_k(x) L_k(\mathbf{y})$, yielding the fully discrete SC approximation $u_p \in \mathcal{P}_{\mathcal{J}(p)}(\Gamma) \otimes W^h(D)$, where $L_k \in \mathcal{P}_{\mathcal{J}(p)}(\Gamma)$ are suitable combinations of Lagrange interpolants

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

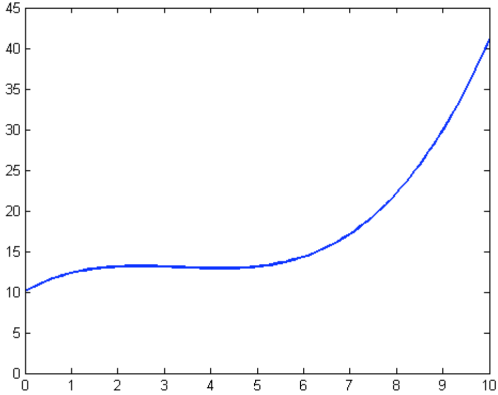
$$\mathbb{E}[u](x) \approx \int_\Gamma u_p(\mathbf{y}, x)\rho(\mathbf{y})d\mathbf{y} = \sum_{k=1}^M u_k(x) \underbrace{\int_\Gamma L_k(y)\rho(\mathbf{y})dy}_{\text{precomputed weights}} = \sum_{k=1}^M u_k(x) w_k$$

[Tatang], [Mathelin-Hussani], [Hesthaven-Xiu], [Babuška et. al], [Zabaras et.al], [Nobile-Tempone-CW], etc.

Quantities of interest

Interpolatory quadrature

A simple function to integrate

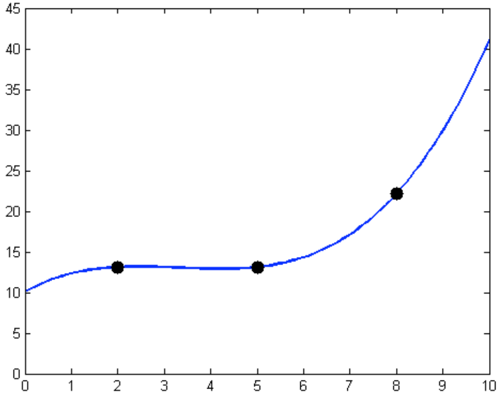


$u(y)$ is a given function

Quantities of interest

Interpolatory quadrature

Selected function values

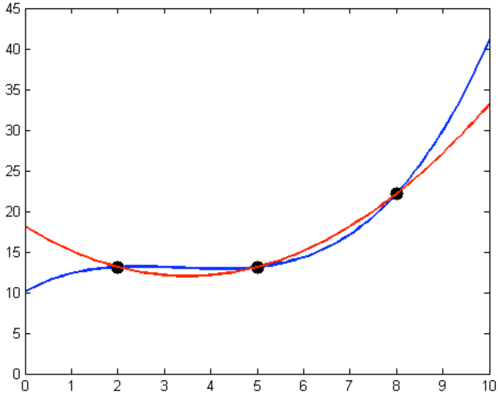


Evaluate $u(y)$ at M values $\{u(y_1), u(y_2), \ldots, u(y_M)\}$

Quantities of interest

Interpolatory quadrature

The interpolant

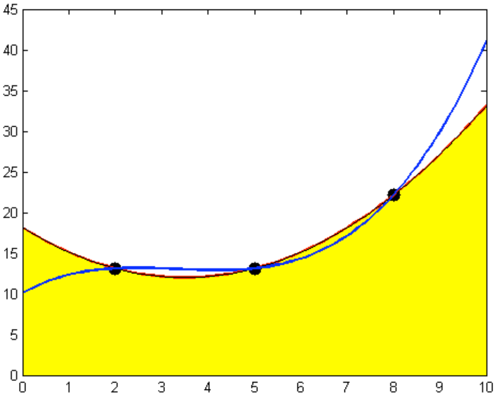


Determine the approximate polynomial u_p

Quantities of interest

Interpolatory quadrature

The area under the curve



The QoI = Integrating the approximating polynomial EXACTLY

Stochastic Collocation (SC) FEM

Advantages compared with SGFEM

- preserves the convergence rate to the stochastic Galerkin FEM (SGFEM)

Pros: there are several advantages wrt the SGFEM approach:

- completely decouples computations as Monte Carlo does,
- efficiently treats the case of non-independent RVs by introducing an auxiliary density
- No difficulty in treating nonlinear problems, exponential expansions of RFs, unbounded RVs (e.g. Gaussian), etc.,
- effectively handle problems that depend on random input data described by a moderately large number of RVs with the use of sparse gird collocation ([Smolyak '63], [Griebel et al '98-'03-'04], [Hesthaven-Xiu '05], [Barthelmann-Novak-Ritter '00], [Zabaras et al '07], [Nobile-Tempone-W. '08])

Cons:

- can use more DoFs then the SGFEM to represent the same polynomial subspace
- the (interpolation) Lebesgue constant can impact the rate of convergence

Stochastic Collocation (SC) FEM

Advantages compared with SGFEM

- preserves the convergence rate to the stochastic Galerkin FEM (SGFEM)

Pros: there are several advantages wrt the SGFEM approach:

- completely decouples computations as Monte Carlo does,
- efficiently treats the case of non-independent RVs by introducing an auxiliary density
- No difficulty in treating nonlinear problems, exponential expansions of RFs, unbounded RVs (e.g. Gaussian), etc.,
- effectively handle problems that depend on random input data described by a moderately large number of RVs with the use of sparse gird collocation ([Smolyak '63], [Griebel et al '98-'03-'04], [Hesthaven-Xiu '05], [Barthelmann-Novak-Ritter '00], [Zabaras et al '07], [Nobile-Tempone-W. '08])

Cons:

- can use more DoFs then the SGFEM to represent the same polynomial subspace
- the (interpolation) Lebesgue constant can impact the rate of convergence

Stochastic Collocation (SC) FEM

Advantages compared with SGFEM

- preserves the convergence rate to the stochastic Galerkin FEM (SGFEM)

Pros: there are several advantages wrt the SGFEM approach:

- completely decouples computations as Monte Carlo does,
- efficiently treats the case of non-independent RVs by introducing an auxiliary density
- No difficulty in treating nonlinear problems, exponential expansions of RFs, unbounded RVs (e.g. Gaussian), etc.,
- effectively handle problems that depend on random input data described by a moderately large number of RVs with the use of sparse gird collocation ([Smolyak '63], [Griebel et al '98-'03-'04], [Hesthaven-Xiu '05], [Barthelmann-Novak-Ritter '00], [Zabaras et al '07], [Nobile-Tempone-W. '08])

Cons:

- can use more DoFs then the SGFEM to represent the same polynomial subspace
- the (interpolation) Lebesgue constant can impact the rate of convergence

Tensor product SCFEM

The simplest multi-dimensional interpolant

Basic idea is to construct the point set H^M for each variable y_n: $H_n^{m(i_n)}$

- choose the level i_n of interpolation in the nth direction
- set the number of points used by the i_nth interpolant, denoted $m(i_n)$
- define the set $H_n^{m(i_n)} = \left\{ y_n^1, y_n^2, \ldots, y_n^{m(i_n)} \right\}$ of 1d interpolating points:
  - according to the measure $\rho(y_n)dy_n$, e.g. Gauss-Hermite (Normal), Gauss-Legendre, Clenshaw-Curtis (Uniform), etc.
- $H^M = H_1^{m(i_1)} \times \cdots \times H_N^{m(i_N)}$ where $M_{TP} = m(i_1)m(i_2)\ldots m(i_N)$
- $\mathbf{y_k} = \left(y_1^{k_1}, y_2^{k_2}, \ldots, y_N^{k_N} \right)$, where $\mathbf{k} \in \text{TP} \equiv \left\{ \mathbf{k} \in \mathbb{N}_+^N \,:\, k_n < m(i_n) \right\}$

The tensor product (TP) Lagrange-interpolant is defined by:

$$u_p^{TP}(\mathbf{y}, x) = \sum_{\mathbf{k} \in \text{TP}}^{M_{TP}} u_{\mathbf{k}}(x) L_{\mathbf{k}}(\mathbf{y}), \quad \text{with } L_{\mathbf{k}}(\mathbf{y}) = \prod_{n=1}^{N} \prod_{s=1, s \neq k_n}^{m(i_n)} \frac{y_n - y_n^s}{y_n^{k_n} - y_n^s}$$

Tensor product SCFEM

The simplest multi-dimensional interpolant

Basic idea is to construct the point set H^M for each variable y_n: $H_n^{m(i_n)}$

- choose the level i_n of interpolation in the nth direction
- set the number of points used by the i_nth interpolant, denoted $m(i_n)$
- define the set $H_n^{m(i_n)} = \left\{ y_n^1, y_n^2, \dots, y_n^{m(i_n)} \right\}$ of 1d interpolating points:
  - according to the measure $\rho(y_n)dy_n$, e.g. Gauss-Hermite (Normal), Gauss-Legendre, Clenshaw-Curtis (Uniform), etc.
- $H^M = H_1^{m(i_1)} \times \cdots \times H_N^{m(i_N)}$ where $M_{TP} = m(i_1)m(i_2)\dots m(i_N)$
- $\mathbf{y_k} = \left(y_1^{k_1}, y_2^{k_2}, \dots, y_N^{k_N} \right)$, where $\mathbf{k} \in \mathsf{TP} \equiv \left\{ \mathbf{k} \in \mathbb{N}_+^N \ : \ k_n < m(i_n) \right\}$

The tensor product (TP) Lagrange-interpolant is defined by:

$$u_p^{TP}(\mathbf{y}, x) = \sum_{\mathbf{k} \in \mathsf{TP}}^{M_{TP}} u_{\mathbf{k}}(x) L_{\mathbf{k}}(\mathbf{y}), \quad \text{with } L_{\mathbf{k}}(\mathbf{y}) = \prod_{n=1}^{N} \prod_{s=1, s\neq k_n}^{m(i_n)} \frac{y_n - y_n^s}{y_n^{k_n} - y_n^s}$$

Tensor product SCFEM

Let $\mathscr{U}_n^{m(i_n)}$ be the ith level interpolant in the direction y_n using $m(i_n)$ points:

$$\mathscr{U}_n^{m(i_n)}[u](y_n) = \sum_{k=1}^{m(i_n)} u(y_n^k) l_n^k(y_n), \quad \{y_n^1, \dots, y_n^{m(i_n)}\} \in \Gamma_n$$

- $\mathscr{U}_n^{m(i_n)} : C^0(\Gamma_n) \to \mathcal{P}_{m(i_n)-1}(\Gamma_n), \quad \mathscr{U}_n^0[u] = 0 \; \forall u \in C^0(\Gamma_n)$
- The degree in the y_n direction is $p_n = m(i_n) - 1$

The TP-SC approximation is given by

$$u_p^{TP}(\mathbf{y}) = \bigotimes_{n=1}^{N} \mathscr{U}_n^{m(i_n)}[u](\mathbf{y}), \quad \max_n \alpha_n p_n \leq p$$

- the interpolation requires $M_{TP} = \prod_{n=1}^{N} m(i_n)$ function evaluations (In this case, solutions of the PDE)

Tensor product SCFEM

Let $\mathscr{U}_n^{m(i_n)}$ be the ith level interpolant in the direction y_n using $m(i_n)$ points:

$$\mathscr{U}_n^{m(i_n)}[u](y_n) = \sum_{k=1}^{m(i_n)} u(y_n^k) l_n^k(y_n), \quad \{y_n^1, \ldots, y_n^{m(i_n)}\} \in \Gamma_n$$

- $\mathscr{U}_n^{m(i_n)} : C^0(\Gamma_n) \to \mathcal{P}_{m(i_n)-1}(\Gamma_n), \quad \mathscr{U}_n^0[u] = 0 \; \forall u \in C^0(\Gamma_n)$
- The degree in the y_n direction is $p_n = m(i_n) - 1$

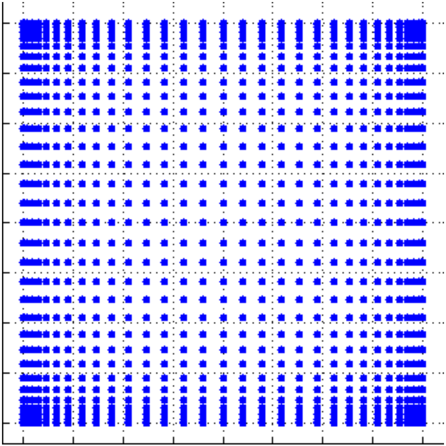
The TP-SC approximation is given by

$$u_p^{TP}(\mathbf{y}) = \bigotimes_{n=1}^{N} \mathscr{U}_n^{m(i_n)}[u](\mathbf{y}), \quad \max_n \alpha_n p_n \leq p$$

- the interpolation requires $M_{TP} = \prod_{n=1}^{N} m(i_n)$ function evaluations (In this case, solutions of the PDE)

TP grid for SCFEM

Isotropic grid $\max(p_1, p_2) \leq 32$



Isotropic TP SC grid constructed from C-C points for $(y_1, y_2) \in U(-1, 1)$

Choices for interpolation

Based on 1-d interpolation formulas

Clenshaw-Curtis abscissas (Γ_n bounded):

- $\{y_n^k\}_{k=1}^{m(i_n)}$: extrema of Chebyshev polynomials
- optimal for uniform convergence in Γ_n
- if $m(i_n) = 2^{i_n-1} + 1$ lead to nested sets, i.e. $H_n^{m(i_n)} \subset H_n^{m(i_n+1)}$

Gaussian abscissas (Γ_n bounded or unbounded): Assume, either

- Y_n independent, i.e. $\rho(\mathbf{y}) = \prod_{n=1}^N \rho_n(y_n)$, or
- construct an auxiliary joint PDF $\hat{\rho}(\mathbf{y}) = \prod_{n=1}^N \hat{\rho}_n(y_n)$ such that $\|\rho/\hat{\rho}\|_{L^\infty(\Gamma)} < \infty$ and small enough.
- $\{y_n^k\}_{k=1}^{m(i_n)}$: zeros of orthogonal polynomials with respect to $\hat{\rho}$ e.g. abscissas become roots of Gauss-Legendre, -Hermite, -Jacobi, -Laguerre polynomials corresponding to uniform, normal, beta, exponential distributions, respectively
- optimal for L^2_ρ convergence

Convergence of the TP-SCFEM

w.r.t. the polynomial order and the # of PDE solves

Recall the $\mathbf{p} = (p_1, \ldots, p_N)$ is the polynomial degree used in each direction $\mathbf{y}_n$

Theorem [Babuška-Nobile-Tempone, 2007]:

Let $L^2_\rho \equiv L^2_\rho(\Gamma; H^1_0(D))$ then since u is analytic in $\mathbf{y}$ you get:

- Γ_n bounded:

$$\left\|u - u_p^{TP}\right\|_{L^2_\rho} \leq C \sum_{n=1}^{N} e^{-g_n p_n}, \quad \text{with } g_n = \log\left[\frac{2\tau_n}{|\Gamma_n|} + \sqrt{1 + \frac{4{\tau_n}^2}{|\Gamma_n|^2}}\right]$$

- Γ_n unbounded, $\hat{\rho}_n \approx e^{-(\delta_n y_n)^2}$ at infinity:

$$\left\|u - u_p^{TP}\right\|_{L^2_\rho} \leq C \sum_{n=1}^{N} \sqrt{p_n} e^{-g_n \sqrt{p_n}}, \quad g_n = \frac{\sqrt{2}\tau_n}{\delta_n}$$

Error in # of samples M: $\varepsilon_{TP}(M) = \left\|u - u_p^{TP}\right\|_{L^2_\rho} \leq C(N) M^{-g_{min}/N}$

How to construct anisotropic weights?

Example: Γ_n bounded

1-dimensional analysis: polynomial approximation (L^2 projection or interpolation using Gauss points) in y_n (only) yields exponential convergence

$$\varepsilon_n = \|u - u_p\|_{L^2_\rho} \leq C e^{-g_n p_n}$$

Optimal choice for anisotropic weights: $\alpha_n = g_n$

- The decay rates g_n can be estimated theoretically (*a priori*),

$$g_n = \log\left(\frac{2\tau_n}{|\Gamma_n|} + \sqrt{1 + \frac{4{\tau_n}^2}{|\Gamma_n|^2}}\right)$$

  and numerically (*a posteriori*), $\log_{10}(\varepsilon_n) \approx \log_{10}(d_n) - p_n \log_{10}(e)\alpha_n$
- Theoretical estimates for linear and several nonlinear PDEs available [BNT07, W07, NTW08a, NTW08b, GW11]
- Dimension-adaptivity without paying the cost of searching and evaluating the multi-indices $\{\mathbf{p} + e_j, 1 \leq j \leq N\}$ using an heuristic error estimator [Gerstner-Griebel '03]

How to construct anisotropic weights?

Example: Γ_n bounded

1-dimensional analysis: polynomial approximation (L^2 projection or interpolation using Gauss points) in y_n (only) yields exponential convergence

$$\varepsilon_n = \|u - u_p\|_{L^2_\rho} \leq C e^{-g_n p_n}$$

Optimal choice for anisotropic weights: $\alpha_n = g_n$

- The decay rates g_n can be estimated theoretically (*a priori*),

$$g_n = \log\left(\frac{2\tau_n}{|\Gamma_n|} + \sqrt{1 + \frac{4\tau_n{}^2}{|\Gamma_n|^2}}\right)$$

  and numerically (*a posteriori*), $\log_{10}(\varepsilon_n) \approx \log_{10}(d_n) - p_n \log_{10}(e)\alpha_n$
- Theoretical estimates for linear and several nonlinear PDEs available [BNT07, W07, NTW08a, NTW08b, GW11]
- Dimension-adaptivity without paying the cost of searching and evaluating the multi-indices $\{\mathbf{p} + e_j, 1 \leq j \leq N\}$ using an heuristic error estimator [Gerstner-Griebel '03]

Numerical example
TP-SCFEM

We let $\mathbf{x} = (x_1, x_2)$ and consider the following nonlinear elliptic SPDE:

$$\begin{cases} -\nabla \cdot (a(\omega, x_1)\nabla u(\omega, \mathbf{x})) = \cos(x_1)\sin(x_2) & \mathbf{x} \in [0,1]^2 \\ u(\omega, \mathbf{x}) = 0 & \text{on } \partial D \end{cases}$$

The diffusion coefficient is a 1d random field (varies only in x_1) and is $a(\omega, x_1) = 0.5 + \exp\{\gamma(\omega, x_1)\}$, where γ is a truncated 1d random field with correlation length L and covariance

$$Cov[\gamma](x_1, \tilde{x}_1) = \exp\left(-\frac{(x_1 - \tilde{x}_1)^2}{L^2}\right), \quad \forall (x_1, \tilde{x}_1) \in [0,1]$$

$$\gamma(\omega, x_1) = 1 + Y_1(\omega)\left(\frac{\sqrt{\pi}L}{2}\right)^{1/2} + \sum_{n=2}^{N} \beta_n \, \varphi_n(x_1)\, Y_n(\omega)$$

$$\beta_n := (\sqrt{\pi}L)^{1/2}\, e^{\frac{-(\lfloor \frac{n}{2} \rfloor \pi L)^2}{8}}, \quad \varphi_n(x_1) := \begin{cases} \sin\left(\lfloor \frac{n}{2} \rfloor \pi x_1\right), & \text{if } n \text{ even,} \\ \cos\left(\lfloor \frac{n}{2} \rfloor \pi x_1\right), & \text{if } n \text{ odd} \end{cases}$$

- $\mathbb{E}[Y_n] = 0$ and $\mathbb{E}[Y_n Y_m] = \delta_{nm}$ for $n, m \in \mathbb{N}_+$ and iid in $U(-\sqrt{3}, \sqrt{3})$

Numerical example

TP-SCFEM

We let $\mathbf{x} = (x_1, x_2)$ and consider the following nonlinear elliptic SPDE:

$$\begin{cases} -\nabla \cdot (a(\omega, x_1)\nabla u(\omega, \mathbf{x})) = \cos(x_1)\sin(x_2) & \mathbf{x} \in [0,1]^2 \\ u(\omega, \mathbf{x}) = 0 & \text{on } \partial D \end{cases}$$

The diffusion coefficient is a 1d random field (varies only in x_1) and is $a(\omega, x_1) = 0.5 + \exp\{\gamma(\omega, x_1)\}$, where γ is a truncated 1d random field with correlation length L and covariance

$$Cov[\gamma](x_1, \tilde{x}_1) = \exp\left(-\frac{(x_1 - \tilde{x}_1)^2}{L^2}\right), \quad \forall (x_1, \tilde{x}_1) \in [0,1]$$

$$\gamma(\omega, x_1) = 1 + Y_1(\omega)\left(\frac{\sqrt{\pi}L}{2}\right)^{1/2} + \sum_{n=2}^{N} \beta_n \, \varphi_n(x_1)\, Y_n(\omega)$$

$$\beta_n := \left(\sqrt{\pi}L\right)^{1/2} e^{\frac{-\left(\lfloor \frac{n}{2} \rfloor \pi L\right)^2}{8}}, \quad \varphi_n(x_1) := \begin{cases} \sin\left(\lfloor \frac{n}{2} \rfloor \pi x_1\right), & \text{if } n \text{ even,} \\ \cos\left(\lfloor \frac{n}{2} \rfloor \pi x_1\right), & \text{if } n \text{ odd} \end{cases}$$

- $\mathbb{E}[Y_n] = 0$ and $\mathbb{E}[Y_n Y_m] = \delta_{nm}$ for $n, m \in \mathbb{N}_+$ and iid in $U(-\sqrt{3}, \sqrt{3})$

Calculating the weighting parameters

A priori selection: $N = 11$

A priori of the dimension weights $\alpha_n = g_n$:

$$g_n = \log\left(\frac{2\tau_n}{|\Gamma_n|} + \sqrt{1 + \frac{4{\tau_n}^2}{|\Gamma_n|^2}}\right) \quad \text{and } \tau_n = \frac{1}{12\sqrt{\lambda_n}\|b_n\|_{L^\infty(D)}}$$

For this problem we have

$$g_n = \begin{cases} \log\left(1 + c/\sqrt{L}\right), & \text{for } n << L^{-2} \\ n^2 L^2, & \text{for } n > L^{-2} \end{cases}$$

	α_1	α_2, α_3	α_4, α_5	α_6, α_7	α_8, α_9	α_{10}, α_{11}
$L = 1/2$	0.20	0.19	0.42	1.24	3.1	5.8
$L = 1/64$	0.79	0.62	0.62	0.62	0.62	0.62

Goal: $\|\mathbb{E}[\epsilon]\|_{L^2(D)} \approx \|\mathbb{E}\left[u_p^{TP}(\mathbf{y}, x) - u_{p_{max}+1}^{TP}(\mathbf{y}, x)\right]\|_{L^2(D)}$

- $p = 0, 1, \ldots, p_{max}$ and $u_{p_{max}+1}$ is an overkilled solution

Calculating the weighting parameters

A priori selection: $N = 11$

A priori of the dimension weights $\alpha_n = g_n$:

$$g_n = \log\left(\frac{2\tau_n}{|\Gamma_n|} + \sqrt{1 + \frac{4{\tau_n}^2}{|\Gamma_n|^2}}\right) \quad \text{and } \tau_n = \frac{1}{12\sqrt{\lambda_n}\|b_n\|_{L^\infty(D)}}$$

For this problem we have

$$g_n = \begin{cases} \log\left(1 + c/\sqrt{L}\right), & \text{for } n << L^{-2} \\ n^2 L^2, & \text{for } n > L^{-2} \end{cases}$$

	α_1	$\alpha_2,\ \alpha_3$	$\alpha_4,\ \alpha_5$	$\alpha_6,\ \alpha_7$	$\alpha_8,\ \alpha_9$	$\alpha_{10},\ \alpha_{11}$
$L = 1/2$	0.20	0.19	0.42	1.24	3.1	5.8
$L = 1/64$	0.79	0.62	0.62	0.62	0.62	0.62

Goal: $\|\mathbb{E}[\epsilon]\|_{L^2(D)} \approx \|\mathbb{E}\left[u_p^{TP}(\mathbf{y}, x) - u_{p_{max}+1}^{TP}(\mathbf{y}, x)\right]\|_{L^2(D)}$

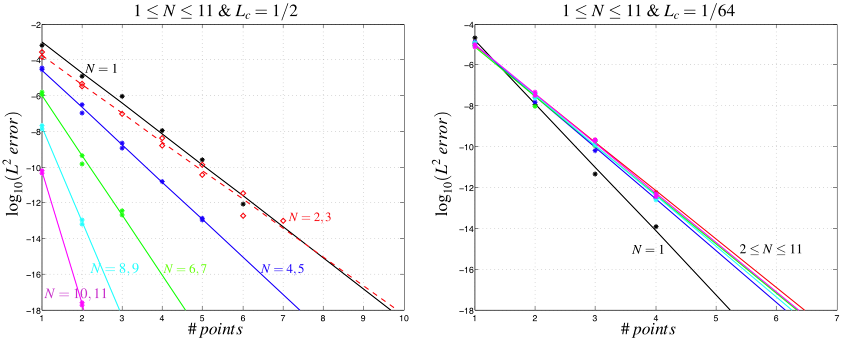
- $p = 0, 1, \ldots, p_{max}$ and $u_{p_{max}+1}$ is an overkilled solution

Calculating the weighting parameters

A posteriori selection: $N = 11$

$$\|\mathbb{E}[\epsilon_n]\|_{L^2(D)} \approx \|\mathbb{E}\left[u_p(y_n,x) - u_{p_{max}+1}(y_n,x)\right]\|_{L^2(D)}$$

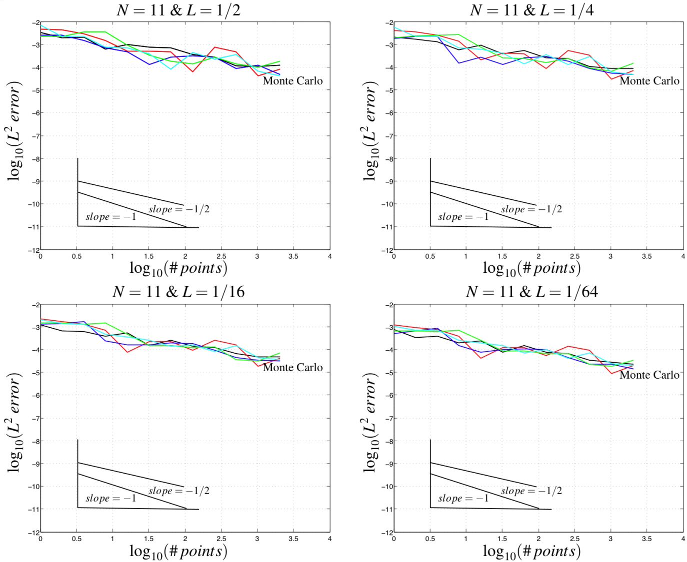
- $p = 0, 1, \ldots, p_{max}$ and $u_{p_{max}+1}$ is an overkilled solution



A linear least square approximation to fit $\log_{10}(\|E[\varepsilon_n]\|_{L^2(D)})$ versus p_n. For $n = 1, 2, \ldots, N = 11$ we plot: on the left, the highly anisotropic case $L_c = 1/2$ and on the right, the isotropic case $L_c = 1/64$

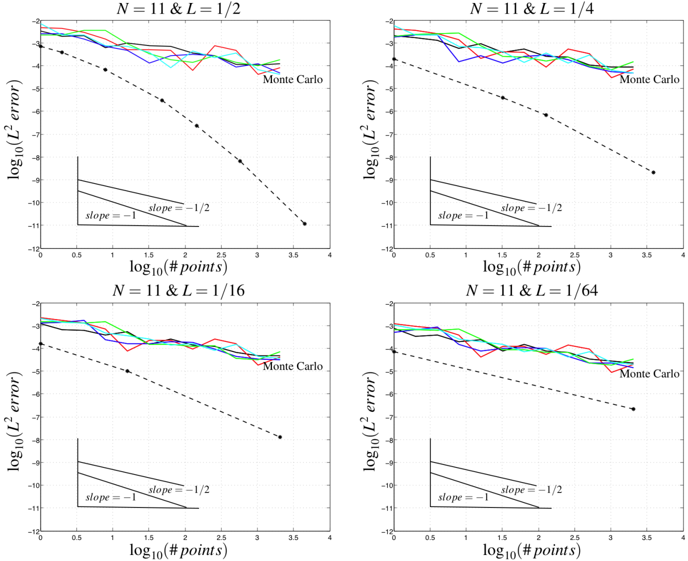
Convergence Comparisons

$N = 11$ random variables



Convergence Comparisons

$N = 11$ random variables



Summary

What's coming next?

- Proper input data representation/truncation is important to reduce the computational work
- Discussed the various intrusive and non-intrusive stochastic techniques for the for forward propagation of uncertainty, in particular Monte Carlo and Stochastic Galerkin/Collocation
- Global stochastic polynomial approximation is extremely effective for problems that smooth (analytic) dependence on the random variables
- Properly chosen Anisotropic polynomial spaces can improve considerably the convergence, when the input random variables have different influence on the output
- Can we construct a polynomial approximation that maintains the fast convergence even when N becomes large?
  - sparse grid SCFEM with anisotropic refinement
  - what about stochastic inverse problems and calibration?

What's coming next?

Sparse grid SCFEM

- Recall that $\mathscr{U}_n^{m(i_n)}$ be the ith level interpolant in the direction y_n using $m(i_n)$ points

$$\mathscr{U}_n^{m(i_n)} : C^0(\Gamma_n) \to \mathcal{P}_{m(i_n)-1}(\Gamma_n), \quad \mathscr{U}_n^0[u] = 0 \; \forall u \in C^0(\Gamma_n)$$

- The TP-SCFEM: $u_p^{TP}(\mathbf{y}) = \bigotimes_{n=1}^{N} \mathscr{U}_n^{m(i_n)}[u](\mathbf{y}), \quad \max_n \alpha_n p_n \le p$
- The nth difference operator: $\Delta_n^{m(i_n)}[u] = \mathscr{U}_n^{m(i_n)}[u] - \mathscr{U}_n^{m(i_n-1)}[u]$
- The hierarchical surplus: $\Delta_n^{m(\mathbf{i})}[u](\mathbf{y}) = \bigotimes_{n=1}^{N} \Delta_n^{m(i_n)}[u](\mathbf{y})$ where $\mathbf{i} = (i_1, \ldots, i_N) \in \mathbb{N}_+^N$ is a multi-index

What's coming next?

Sparse grid SCFEM

- Recall that $\mathscr{U}_n^{m(i_n)}$ be the ith level interpolant in the direction y_n using $m(i_n)$ points

$$\mathscr{U}_n^{m(i_n)} : C^0(\Gamma_n) \to \mathcal{P}_{m(i_n)-1}(\Gamma_n), \quad \mathscr{U}_n^0[u] = 0 \; \forall u \in C^0(\Gamma_n)$$

- The TP-SCFEM: $u_p^{TP}(\mathbf{y}) = \bigotimes_{n=1}^{N} \mathscr{U}_n^{m(i_n)}[u](\mathbf{y}), \quad \max_n \alpha_n p_n \leq p$
- The nth difference operator: $\Delta_n^{m(i_n)}[u] = \mathscr{U}_n^{m(i_n)}[u] - \mathscr{U}_n^{m(i_n-1)}[u]$
- The hierarchical surplus: $\Delta_n^{m(\mathbf{i})}[u](\mathbf{y}) = \bigotimes_{n=1}^{N} \Delta_n^{m(i_n)}[u](\mathbf{y})$ where $\mathbf{i} = (i_1, \ldots, i_N) \in \mathbb{N}_+^N$ is a multi-index

What's coming next?

Sparse grid SCFEM

- Recall that $\mathscr{U}_n^{m(i_n)}$ be the ith level interpolant in the direction y_n using $m(i_n)$ points

$$\mathscr{U}_n^{m(i_n)} : C^0(\Gamma_n) \to \mathcal{P}_{m(i_n)-1}(\Gamma_n), \quad \mathscr{U}_n^0[u] = 0 \; \forall u \in C^0(\Gamma_n)$$

- The TP-SCFEM: $u_p^{TP}(\mathbf{y}) = \bigotimes_{n=1}^{N} \mathscr{U}_n^{m(i_n)}[u](\mathbf{y}), \quad \max_n \alpha_n p_n \leq p$
- The nth difference operator: $\Delta_n^{m(i_n)}[u] = \mathscr{U}_n^{m(i_n)}[u] - \mathscr{U}_n^{m(i_n-1)}[u]$
- The hierarchical surplus: $\Delta_n^{m(\mathbf{i})}[u](\mathbf{y}) = \bigotimes_{n=1}^{N} \Delta_n^{m(i_n)}[u](\mathbf{y})$ where $\mathbf{i} = (i_1, \ldots, i_N) \in \mathbb{N}_+^N$ is a multi-index

What's coming next?

Sparse grid SCFEM

- Recall that $\mathscr{U}_n^{m(i_n)}$ be the ith level interpolant in the direction y_n using $m(i_n)$ points

$$\mathscr{U}_n^{m(i_n)} : C^0(\Gamma_n) \to \mathcal{P}_{m(i_n)-1}(\Gamma_n), \quad \mathscr{U}_n^0[u] = 0 \; \forall u \in C^0(\Gamma_n)$$

- The TP-SCFEM: $u_p^{TP}(\mathbf{y}) = \bigotimes_{n=1}^{N} \mathscr{U}_n^{m(i_n)}[u](\mathbf{y}), \quad \max_n \alpha_n p_n \le p$
- The nth difference operator: $\Delta_n^{m(i_n)}[u] = \mathscr{U}_n^{m(i_n)}[u] - \mathscr{U}_n^{m(i_n-1)}[u]$
- The hierarchical surplus: $\Delta_n^{m(\mathbf{i})}[u](\mathbf{y}) = \bigotimes_{n=1}^{N} \Delta_n^{m(i_n)}[u](\mathbf{y})$ where

  $\mathbf{i} = (i_1, \ldots, i_N) \in \mathbb{N}_+^N$ is a multi-index

Sparse grid SCFEM

High-dimensional interpolation approach

Basic idea: linear combination of tensor product grids, with a relatively low number of points (but maintain the asymptotic accuracy)

The tensor product SCFEM interpolant is defined as:

$$u_p^{TP}(\mathbf{y}) = \bigotimes_{n=1}^{N} \mathscr{U}_n^{m(i_n)}[u](\mathbf{y}), \quad \max_n \alpha_n p_n \leq p$$

The sparse grid SCFEM is defined as

$$u_p^{SG}(\mathbf{y}) = \sum_{g(\mathbf{i}) \leq p} \bigotimes_{n=1}^{N} \Delta_n^{m(i_n)}[u](\mathbf{y}) = \sum_{g(\mathbf{i}) \leq p} c(\mathbf{i}) \bigotimes_{n=1}^{N} \mathscr{U}_n^{m(i_n)}[u](\mathbf{y})$$

with $c(\mathbf{i}) = \sum_{\substack{\mathbf{j} \in \{0,1\}^N \\ g(\mathbf{i}+\mathbf{j}) \leq p}} (-1)^{|\mathbf{j}|_1}$ and $g : \mathbb{N}^N \to \mathbb{N}$ a strictly increasing function

Sparse grid SCFEM

High-dimensional interpolation approach

Basic idea: linear combination of tensor product grids, with a relatively low number of points (but maintain the asymptotic accuracy)

The tensor product SCFEM interpolant is defined as:

$$u_p^{TP}(\mathbf{y}) = \bigotimes_{n=1}^{N} \mathscr{U}_n^{m(i_n)}[u](\mathbf{y}), \quad \max_n \alpha_n p_n \leq p$$

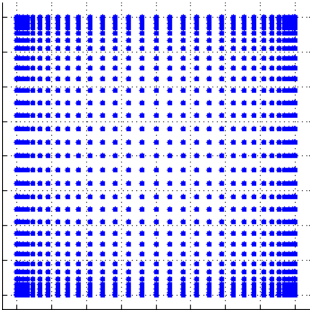
The sparse grid SCFEM is defined as

$$u_p^{SG}(\mathbf{y}) = \sum_{g(\mathbf{i})\leq p} \bigotimes_{n=1}^{N} \Delta_n^{m(i_n)}[u](\mathbf{y}) = \sum_{g(\mathbf{i})\leq p} c(\mathbf{i}) \bigotimes_{n=1}^{N} \mathscr{U}_n^{m(i_n)}[u](\mathbf{y})$$

with $c(\mathbf{i}) = \sum_{\substack{\mathbf{j}\in\{0,1\}^N \\ g(\mathbf{i}+\mathbf{j})\leq p}} (-1)^{|\mathbf{j}|_1}$ and $g : \mathbb{N}^N \to \mathbb{N}$ a strictly increasing function

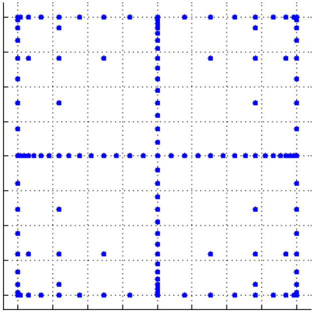
Asymptotic accuracy

What about the *curse of dimensionality*?



Tensor product grid

$$\varepsilon_{TP}(M) \le C(N) M^{-g_{min}/N}$$



Sparse grid

$$\varepsilon_{SG}(M) \le \widetilde{C}(N) M^{-?}$$

- The TP-SCFEM is a non-intrusive method with faster convergence than MCFEM (for smooth solutions)
- The number of samples grows exponentially fast with the number of RVs. Clearly unfeasible, even for moderate N