Polynomial Chaos Methods

Oliver Ernst

USC/QUEST
Uncertainty Quantification Summer School
University of Southern California
August 16, 2016
## Polynomial Chaos Expansion

1. **Introduction**
2. **Orthogonal Polynomials**
3. **Polynomial Chaos Expansion**

## Random PDEs

1. **A model problem**
2. **Representation of random fields**
3. **Stochastic Galerkin Discretization**
4. **Stochastic Collocation**
5. **Further PC-based methods**
Contents

1 Polynomial Chaos Expansion

2 Random PDEs
1 Polynomial Chaos Expansion
1.1 Introduction
1.2 Orthogonal Polynomials
1.3 Polynomial Chaos Expansion
Given **probability space** $(\Omega, \mathcal{A}, P)$, denote by

$$\xi : \Omega \to \mathbb{R}, \quad \omega \mapsto \xi(\omega)$$

a real-valued **random variable**, i.e., a measurable mapping from $(\Omega, \mathcal{A})$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. 

---

Oliver Ernst (TU Chemnitz)  
Polynomial Chaos Methods  
USC, August 2016  
5 / 111
Random Variables

Given probability space $(\Omega, \mathcal{A}, P)$, denote by

$$\xi : \Omega \rightarrow \mathbb{R}, \quad \omega \mapsto \xi(\omega)$$

a real-valued random variable, i.e., a measurable mapping from $(\Omega, \mathcal{A})$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

The $\sigma$-algebra $\sigma(\xi)$ generated by a random variable $\xi$ is the smallest $\sigma$-algebra containing the sets $\xi^{-1}(B), \; B \in \mathcal{B}(\mathbb{R})$. 
Given probability space \((Ω, Α, P)\), denote by
\[ ξ : Ω \to ℝ, \quad ω \mapsto ξ(ω) \]
a real-valued random variable, i.e., a measurable mapping from \((Ω, Α)\) to \((ℝ, Β(ℝ))\).

The \(σ\)-algebra \(σ(ξ)\) generated by a random variable \(ξ\) is the smallest \(σ\)-algebra containing the sets \(ξ^{-1}(B)\), \(B ∈ Β(ℝ)\).

Probability distribution (measure) \(μ = μ_ξ\) of \(ξ\) on \(Β(ℝ)\) defined by
\[ μ(B) = P(ξ^{-1}(B)), \quad B ∈ Β(ℝ). \]
Given probability space \((\Omega, \mathcal{A}, P)\), denote by

\[
\xi : \Omega \rightarrow \mathbb{R}, \quad \omega \mapsto \xi(\omega)
\]
a real-valued random variable, i.e., a measurable mapping from \((\Omega, \mathcal{A})\) to \((\mathbb{R}, \mathcal{B}(\mathbb{R}))\).

The \(\sigma\)-algebra \(\sigma(\xi)\) generated by a random variable \(\xi\) is the smallest \(\sigma\)-algebra containing the sets \(\xi^{-1}(B), B \in \mathcal{B}(\mathbb{R})\).

Probability distribution (measure) \(\mu = \mu_\xi\) of \(\xi\) on \(\mathcal{B}(\mathbb{R})\) defined by

\[
\mu(B) = P(\xi^{-1}(B)), \quad B \in \mathcal{B}(\mathbb{R}).
\]

Cumulative distribution function (CDF): \(F = F_\xi : \mathbb{R} \rightarrow [0, 1]\) such that

\[
F(x) = P(\xi \leq x), \quad x \in \mathbb{R}.
\]
If \( \mu_\xi \) is absolutely continuous, then \( \xi \) possesses probability density function (PDF) \( f = f_\xi : \mathbb{R} \to \mathbb{R}_0^+ \) such that

\[
F(x) = \int_{-\infty}^{x} f(t) \, dt, \quad x \in \mathbb{R}, \quad \text{i.e., } F' = f.
\]
Uniform distribution,

\[ \xi \sim U[a, b], \quad a < b, \quad a, b, \in \mathbb{R}. \quad f_\xi(x) = 1_{[a,b]}, \quad x \in \mathbb{R}. \]
Standard normal (Gaussian) distribution,

\[ \xi \sim N(0, 1), \quad f_{\xi}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}. \]
\( \chi^2 \) (Chi-squared) distribution with \( k \in \mathbb{N} \) degrees of freedom,

\[
\xi \sim \chi_k^2, \quad f_\xi(x) = \frac{x^{k/2-1}}{2^{k/2} \Gamma(k/2)} e^{-x/2}, \quad x \in \mathbb{R}.
\]
Introduction

Random Variables: moments

Expectation of RV $\xi$:

$$E[\xi] := \int_\Omega \xi(\omega) \, dP(\omega) = \int_\mathbb{R} \xi \, d\mu(\omega),$$

$k$-th moment

$$E[\xi^k], \quad k \in \mathbb{N},$$

Variance

$$\text{Var} \, \xi := E[(\xi - E[\xi])^2] \quad (2\text{nd central moment}).$$

$L^p$ norms ($1 \leq p \leq \infty$):

$$\|\xi\|_p := E[|\xi|^p]^{1/p}, \quad 1 \leq p < \infty, \quad \|\xi\|_\infty := \text{ess sup} |\xi|$$

$L^p$ spaces: $L^p(\Omega) = L^p(\Omega, \mathcal{A}, P) := \{ \xi : \Omega \to \mathbb{R} \text{ RV}, \|\xi\|_p < \infty \}$

$$\|\xi\|_p \leq \|\xi\|_q, \quad L^p \supset L^q, \quad 1 \leq p \leq q \leq \infty.$$
Given RV $\xi$, measurable function $g : \mathbb{R} \to \mathbb{R}$, construct new RV

$$\eta := g(\xi).$$

**Lemma 1.1 (Doob-Dynkin)**

For a random variable $\xi$, each $\sigma(\xi)$-measurable random variable $\eta$ can be written as $\eta = g(\xi)$ for some measurable function $g$.

**Examples:**

- $\xi \sim N(0, 1), \quad \eta = g(\xi) = a + b\xi \; (b > 0); \quad \Rightarrow \quad \eta \sim N(a, b^2)$
- $\xi \sim N(0, 1), \quad \eta = g(\xi) = \xi^2; \quad \Rightarrow \quad \eta \sim \chi_1^2.$
- $\xi$ RV with strictly increasing CDF $F. \quad \Rightarrow \eta := F(\xi) \sim U[0, 1]$. 
**Idea:** Represent arbitrary RV $\eta$ as a function of a given (e.g. Gaussian) RV $\xi$ by approximating the mapping $g$ connecting $\xi$ with $\eta$ by a polynomial $p$, i.e.,

$$\eta = g(\xi) \approx p(\xi).$$

**Note:**

- Polynomials dense in continuous functions on bounded intervals.  
  [Weierstrass, 1885]
- For $(-\infty, \infty)$ replace polynomial by entire function (power series).  
  [Carleman, 1927]
- Expansions in orthogonal polynomials behave as Fourier series.
- If $\xi \sim N(0, 1)$, $\xi(\Omega) = \mathbb{R}$, use Hermite polynomials.  
  These are dense in $L^2_w(\mathbb{R})$, wr.t. weight function

\[
w(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}. \quad [\text{Riesz, 1921}]\]
1 Polynomial Chaos Expansion
  1.1 Introduction
  1.2 Orthogonal Polynomials
  1.3 Polynomial Chaos Expansion
Orthogonal Polynomials

Definition

Denote by $\mathcal{P}_n$: polynomials of degree $\leq n$, $n \in \mathbb{N}_0$, $\mathcal{P}$: polynomials of arbitrary degree, and by $F: \mathbb{R} \to \mathbb{R}$ a non-decreasing, bounded, continuous real-valued distribution function.

Theorem 1.2

If $F$ takes infinitely many values and possesses finite moments of all orders, i.e.,

$$\int_{\mathbb{R}} x^n \, dF(x) < \infty \quad \text{for all } n \in \mathbb{N}_0,$$

then there exists a unique system $\{p_n\}_{n \in \mathbb{N}_0} \subset \mathcal{P}$ of (real) orthogonal polynomials such that

(a) $p_n$ has exact degree $n$ with positive leading coefficient.

(b) The system $\{p_n\}_{n \in \mathbb{N}_0}$ is orthogonal in $L^2(\mathbb{R}, \mathcal{B}(\mathbb{R}), dF)$, i.e.,

$$\int_{\mathbb{R}} p_n(x)p_m(x) \, dF(x) = d_n\delta_{m,n}, \quad m, n \in \mathbb{N}_0, \quad d_n > 0.$$
Orthogonal Polynomials

Remarks

(1) If $F$ has only $N \in \mathbb{N}$ points of increase then the associated system of orthogonal polynomials is finite and spans $\mathcal{P}_{N-1}$.

(2) The case of piecewise constant $F$ corresponds to discrete orthogonal polynomials, which we will not consider.

(3) We will consider only absolutely continuous distribution functions; these possess a weight function or density $\rho = F'$ such that

$$\int_{\mathbb{R}} g(x) \, dF(x) = \int_{\mathbb{R}} g(x) \rho(x) \, dx \quad \text{for all } g \in L^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), dF).$$

(4) We will denote the inner product in the Hilbert space $L^2_\rho(\mathbb{R}) := L^2(\mathbb{R}, \mathcal{B}(\mathbb{R}), dF)$ by

$$(f, g) := (f, g)_\rho := \int_{\mathbb{R}} f(x)g(x) \rho(x) \, dx, \quad f, g \in L^2_\rho(\mathbb{R}),$$

and omit $\rho$ when the context is clear.
Theorem 1.3

A system of (real) orthogonal polynomials obeys a three-term recurrence

\[ \gamma_n p_n(x) = (x - \alpha_n) p_{n-1}(x) - \beta_n p_{n-2}(x), \quad n = 1, 2, \ldots, \]

\[ p_0 \equiv \text{const}, \quad p_{-1} \equiv 0. \]

The recurrence coefficients are given by

\[ \alpha_n = \frac{(xp_{n-1}, p_{n-1})}{(p_{n-1}, p_{n-1})}, \quad \gamma_n = \frac{(xp_{n-1}, p_n)}{(p_n, p_n)}, \quad n = 1, 2, \ldots \]

\[ \beta_n = \frac{(xp_{n-2}, p_{n-1})}{(p_{n-2}, p_{n-2})} = \gamma_{n-1} \frac{(p_{n-1}, p_{n-1})}{(p_{n-2}, p_{n-2})}, \quad n = 2, 3, \ldots, \quad \beta_1 \text{ arbitrary.} \]

This property rests on the self-adjointness of the multiplication operator \( p \mapsto xp \) w.r.t. \((\cdot, \cdot)\), i.e., that

\[ (xp, q) = (p, xq) \quad p, q \in \mathcal{P}. \]
Orthogonal Polynomials

Symmetry, scaling

**Note:** Symmetric densities \( \rho(-x) = \rho(x) \) result in \( \alpha_n \equiv 0 \), giving even/odd symmetry for the associated orthogonal polynomials: \( p_n(x) = (-1)^n p_n(-x) \).

**Remark 1.4**

1. **Rescaling** a system of orthogonal polynomials \( \{p_n\} \) to \( \{\tilde{p}_n = \delta_n p_n\} \), \( \delta_n > 0 \), results in another system of orthogonal polynomials with recurrence coefficients

\[
\tilde{\alpha}_n = \alpha_n, \quad \tilde{\gamma}_n = \frac{\delta_{n-1}}{\delta_n} \gamma_n \quad (n \geq 1), \quad \tilde{\beta}_n = \frac{\delta_{n-1}}{\delta_{n-2}} \beta_n \quad (n \geq 2).
\]

2. Scaling for leading coefficient one gives **monic** orthogonal polynomials with

\[
\gamma_n = 1 \quad (n \geq 1) \quad \text{and} \quad \beta_n = \frac{(p_{n-1}, p_{n-1})}{(p_{n-2}, p_{n-2})} > 0, \quad (n \geq 2).
\]

3. Scaling for \( (p_n, p_n) = 1 \) for all \( n \) gives **orthonormal** polynomials for which

\[
\beta_n = \gamma_{n-1}, \quad (n \geq 2).
\]
The most well-studied orthogonal polynomials, defined by their densities:

1. **Jacobi polynomials** $P_n^{\alpha,\beta}(x)$ with
   \[
   \rho(x) = (1 - x)^{\alpha}(1 - x)^{\beta}, \quad x \in [-1, 1], \quad \alpha, \beta > -1.
   \]
   Special cases:
   - $\alpha = \beta = 0$, $\rho(x) \equiv 1$ Legendre polynomials.
   - $\alpha = \beta = -1/2$, $\rho(x) = (1 - x^2)^{-1/2}$ Chebyshev polynomials.

2. **Laguerre polynomials** $L_n(x)$ with
   \[
   \rho(x) = e^{-x}, \quad x \in [0, \infty).
   \]

3. **Hermite polynomials** $H_n(x)$ with
   \[
   \rho(x) = e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}.
   \]
   These are known as the probabilist’s Hermite polynomials due to their occurrence in statistics and probability theory. The physicist’s Hermite polynomials, with density $\rho(x) = e^{-x^2}$, are a rescaling of these.
Orthogonal Polynomials

Jacobi polynomials and the standard Beta distribution

\[ f(x) = \frac{x^{\alpha-1}(1 - x)^{\beta-1}}{B(\alpha, \beta)}, \quad B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}, \quad x \in [0, 1]. \]
Orthogonal Polynomials
Legendre polynomials

The Legendre polynomials $P_n$ on $[-1, 1]$ can be expressed by the Rodrigues’ formula

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n}[(x^2 - 1)^n], \quad n \in \mathbb{N}_0,$$

and the three-term recurrence

$$nP_n(x) = (2n - 1)xP_{n-1}(x) - (n - 1)P_{n-2}(x), \quad n \in \mathbb{N},$$

giving

$$P_0(x) = 1, \quad P_1(x) = x, \quad P_2(x) = \frac{1}{2}(3x^2 - 1), \ldots$$

They satisfy the orthogonality relation

$$\int_{-1}^{1} P_n(x)P_m(x) \, dx = \frac{2}{2n + 1} \delta_{n,m}, \quad m, n \in \mathbb{N}_0,$$

Hence $\{\sqrt{2n + 1}P_n\}_{n \in \mathbb{N}_0}$ are the orthonormal Legendre polynomials with respect to the density $\rho(x) \equiv \frac{1}{2}$ on $[-1, 1]$. 
A simple rescaling yields the orthonormal polynomial associated with the uniform distribution on a general interval.

**Proposition 1.5**

The orthonormal polynomials associated with a uniform distribution on a finite interval $[a, b]$ with density $\rho(x) = \frac{1}{b-a}$ are given by

$$p_n(x) = \sqrt{2n+1} P_n \left( \frac{2x - b - a}{b - a} \right), \quad n \in \mathbb{N}_0.$$
Orthogonal Polynomials
Generalized Laguerre polynomials and the Gamma distribution

\[ f(x) = \frac{x^{\alpha-1}}{\Gamma(\alpha)\beta^\alpha} e^{-x/\beta}, \quad \alpha, \beta > 0, \quad x \in [0, \infty). \]

Standard Laguerre obtained for \( \alpha = \beta = 1 \) (black curves).
Orthogonal Polynomials
Hermite polynomials

The Hermite polynomials $H_n$ on $\mathbb{R}$ can be expressed by the Rodrigues’ formula

$$H_n(x) = (-1)^n e^{\frac{x^2}{2}} \frac{d^n}{dx^n} e^{-\frac{x^2}{2}}, \quad n \in \mathbb{N}_0,$$

and the three-term recurrence

$$H_n(x) = xH_{n-1}(x) - (n - 1)H_{n-2}(x), \quad n \in \mathbb{N},$$

giving

$$H_0(x) = 1, \quad H_1(x) = x, \quad H_2(x) = x^2 - 1, \ldots$$

They satisfy the orthogonality relation

$$\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} H_n(x)H_m(x) e^{-\frac{x^2}{2}} dx = n!\delta_{n,m}, \quad m, n \in \mathbb{N}_0,$$

hence $\{H_n/\sqrt{n!}\}_{n \in \mathbb{N}_0}$ are the orthonormal Hermite polynomials with respect to the density $\rho(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. 
A simple rescaling yields the orthonormal polynomials associated with a general normal distribution.

**Proposition 1.6**

The orthonormal polynomials associated with a normal distribution with mean \( \mu \in \mathbb{R} \) and standard deviation \( \sigma > 0 \), i.e., with density

\[
\rho(x) = \frac{1}{\sqrt{2\pi \sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}},
\]

are given by

\[
p_n(x) = \frac{1}{\sqrt{n!}} H_n \left( \frac{x - \mu}{\sigma} \right), \quad n \in \mathbb{N}_0.
\]
Orthogonal Polynomials
The Askey scheme

[Askey & Wilson, 1985]: Classification of all (?) known orthogonal polynomials as hypergeometric orthogonal polynomials, i.e., as polynomial solutions of a differential equation of hypergeometric type

\[ p(x)y'' + q(x)y' + \lambda y = 0, \quad p \in \mathcal{P}_2, \ q \in \mathcal{P}_1, \ \lambda \equiv \text{const.} \quad (1.1) \]

(1.1) has particular solutions \( y_n \in \mathcal{P}_n \) if

\[ \lambda = \lambda_n = -nq' - \frac{1}{2}n(n-1)p''. \]

These polynomial solutions satisfy

\[ \int y_m(x)y_n(x)\rho(x) \, dx = d_n^2 \delta_{n,m}, \quad d_n \neq 0, \]

density \( \rho \) satisfies. \((p\rho)' = q\rho. \) (Pearson differential equation).
Orthogonal Polynomials
The Askey scheme: hypergeometric functions

These polynomials have (finite) hypergeometric series representation

\[ pF_q(a_1, \ldots, a_p; b_1, \ldots, b_q; x) = \sum_{j=0}^{\infty} \frac{(a_1)_j \cdots (a_p)_j x^j}{(b_1)_j \cdots (b_q)_j j!} \]

in terms of shifted factorial (Pochhammer’s symbol)

\[ (a)_0 = 1, \quad a_k = a(a+1) \cdots (a+k-1), \quad k \in \mathbb{N}. \]

Example: Jacobi polynomials

\[ P_n^{(\alpha,\beta)}(x) = \frac{(\alpha + 1)_n}{n!} 2F_1 \left( -n, n + \alpha + \beta + 1; \alpha + 1; \frac{1-x}{2} \right). \]
Orthogonal Polynomials

The Askey scheme: overview

[Koekoek, Lesky & Swarttouw, 2010]
Arrows in graph: limiting relations.

**Examples:**

**Jacobi→Laguerre**

\[
\lim_{\beta \to \infty} P_n^{(\alpha,\beta)}(1 - 2\beta^{-1}x) = L_n^{(\alpha)}(x)
\]

**Laguerre→Hermite**

\[
\lim_{\alpha \to \infty} \left(\frac{2}{\alpha}\right)^\frac{n}{2} L_n^{(\alpha)}(\sqrt{2\alpha}x + \alpha) = \frac{(-1)^n}{n!} H_n(x).
\]
The Askey scheme also contains a number of discrete probability distributions, for example

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Polynomials</th>
<th>Density</th>
<th>Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poisson((\mu))</td>
<td>Charlier</td>
<td>(\rho(k) = e^{-\mu} \frac{\mu^k}{k!})</td>
<td>(\mathbb{N}_0)</td>
</tr>
<tr>
<td>Binomial((N, p))</td>
<td>Krawtchouk</td>
<td>(\rho(k) = \binom{N}{k} p^k q^{N-k})</td>
<td>(\mathbb{N}_0)</td>
</tr>
<tr>
<td>Pascal((\gamma, \mu))</td>
<td>Meixner</td>
<td>(\rho(k) = (1 - \mu)^\gamma \mu^k \frac{\gamma}{k!})</td>
<td>({0, 1, \ldots, N})</td>
</tr>
<tr>
<td>Hyp((\alpha, \beta, N))</td>
<td>Hahn</td>
<td>(\rho(k) = \binom{N}{k} \frac{(\alpha+1)k(\beta+1)N-k}{(\alpha+\beta+2)N})</td>
<td>({0, 1, \ldots, N})</td>
</tr>
</tbody>
</table>
Given a function $f$ in the Hilbert space $L^2_\rho(\Gamma)$, where $\Gamma = \mathbb{R}$, $\Gamma = \mathbb{R}^+$ or $\Gamma = [-1, 1]$ and $\rho$ is the density/weight function for a sequence of orthonormal polynomials $\{p_n\}_{n \in \mathbb{N}}$, we may expand $f \in L^2_\rho(\Gamma)$ as

$$f \sim \sum_{n=0}^{\infty} a_n p_n, \quad a_n = (f, p_n)_\rho.$$ 

- If the system $\{p_n\}$ is dense in $L^2_\rho(\Gamma)$ the expansion converges in $\| \cdot \|_\rho$, the $\rho$-weighted $L^2$-norm on $\Gamma$.
- In this case the approximation error of the truncated expansion $f_N = \sum_{n=0}^{N} a_n p_n$ is given by

$$\| f - f_N \|^2_\rho = \sum_{n=N+1}^{\infty} |a_n|^2,$$

hence results on the decay of $a_n$ immediately translate to approximation error rates.
Orthogonal Polynomials
Orthogonal polynomial expansions

For many families of classical orthogonal polynomials one can prove also uniform convergence on domains of the complex plane which are larger than the convergence region for Taylor series:

Analytic functions on $\Gamma = (-1, 1)$

Analytic functions on $\Gamma = \mathbb{R}$
For a RV $\xi \sim N(0,1)$ and a function $f \in L^2_\rho(\mathbb{R})$ with $\rho$ the standard normal pdf, we have

$$a_n = (f, H_n)_\rho = \int_{\mathbb{R}} f(x) H_n(x) \rho(x) \, dx$$

$$= \int_\Omega f(\xi(\omega)) H_n(\xi(\omega)) \, dP(\omega) = \mathbb{E} [f(\xi)p_n(\xi)],$$

where $\{H_n\}_{n \in \mathbb{N}_0}$ now denote the orthonormal Hermite polynomials.

We have obtained an expansion of the random variable

$$\eta = f(\xi) \in L^2(\Omega, \mathcal{A}, P)$$

in Hermite polynomials of a standard normal RV $\xi$. 
Classical approximation theory (e.g. [Boyd, 2001]) relates smoothness of functions to the decay of the Hermite coefficients $a_n$, i.e., the convergence rate of the Hermite series.

**Definition 1.7**

The *order of real axis decay* $k$ of a function $f : \mathbb{R} \to \mathbb{R}$ is the least upper bound for $r$ for which there exists a constant $C$ such that

$$f(x) = O(e^{-C|x|^r}), \quad |x| \to \infty.$$  

Depending on this decay order, $f$ is called

$$\begin{cases} 
  \text{sub-Gaussian} & \text{if } k < 2 \text{ and } \\
  \text{super-Gaussian} & \text{if } k > 2.
\end{cases}$$
Theorem 1.8 (Hermite series convergence domain)

1. The domain of convergence of a Hermite series is an infinite strip in the complex plane about the real axis bounded by $|\text{Im } z| = w$.

2. For an **entire** function (no finite singularities) the convergence width for its Hermite expansion is

\[
w = \begin{cases} 
\infty, & k > 1, \\
0, & k < 1. 
\end{cases}
\]

3. For a function with finite singularities the convergence width for its Hermite expansion is

\[
w = \begin{cases} 
\tau, & k > 1, \\
\min\{\tau, C\}, & k = 1, \\
0, & k < 1, 
\end{cases}
\]

where $C$ denotes the constant in the bound for its real axis decay $k$ and $\tau$ denotes the absolute value of the singularity closest to the real axis.
Theorem 1.9 (Hermite coefficient decay)

1. If the strip of convergence has finite width $w$, then

$$a_n = O(e^{-w\sqrt{2n+1}}), \quad n \to \infty$$

[Hille, 1939, 1940].

(subgeometric convergence with exponential index $r = 1/2$).

2. For entire functions the exponential convergence index of $a_n$ is

$$r = \begin{cases} \frac{k}{2(k-1)} & f \text{ super-Gaussian}, \\ \frac{k}{2} & f \text{ sub-Gaussian}, \end{cases}$$

[Boyd, 1984]

3. If $f$ decays algebraically with $x$, i.e., $f(x) = O(|x|^{-s})$ as $|x| \to \infty$ and if $f$ has $p$ continuous derivatives on $\mathbb{R}$ (with the $(p+1)$-st derivative continuous except at a finite number of points), then

$$a_n = O(n^{-q/2}), \quad q = \min\{p + 3/2, s - 5/6\},$$

[Bain, 1978].
One can extend this idea to multivariate functions by tensorization:

- Assume RV $\eta$ is a function $g$ of $M$ independent standard normal RV $\xi_1, \ldots, \xi_M$.
- Note that the tensorized Hermite polynomials

$$H_\alpha(\xi) := \prod_{m=1}^{M} H_{\alpha_m}(\xi_m), \quad \alpha = (\alpha_1, \ldots, \alpha_M) \in \mathbb{N}_0^M, \quad \xi = (\xi_1, \ldots, \xi_M) \in \mathbb{R}^M$$

form a (complete) orthonormal system in $L^2_\rho(\mathbb{R}^M)$, where

$$\rho(\xi) = (2\pi)^{-M/2} \exp(-|\xi|^2/2) = \prod_{m=1}^{M} \rho_m(\xi_m),$$

is the joint density of the random vector $\xi(\omega) = (\xi_1(\omega), \ldots, \xi_M(\omega))$:

$$(H_\alpha, H_\beta)_\rho = \int_{\mathbb{R}^M} H_\alpha(\xi) H_\beta(\xi) \rho(\xi) \, d\xi = \prod_{m=1}^{M} (H_{\alpha_m}, H_{\beta_m})_{\rho_m} = \delta_{\alpha, \beta}.$$
1 Polynomial Chaos Expansion

1.1 Introduction

1.2 Orthogonal Polynomials

1.3 Polynomial Chaos Expansion
Polynomial Chaos Expansion

Early history

- [Kolmogorov, 1933] Formalization of probability as measure theory
  Attempt at modeling nonlinear phenomena in statistical mechanics, turbulence.
  - “homogeneous chaos”: stationary random measure
  - polynomial chaoses through repeated (Wiener) integration
  - general stochastic processes approximated by (nonlinear) functionals of multidimensional Wiener process.
- [Cameron & Martin, 1947] Wiener-Hermite orthogonal expansion of 2nd order random processes
- [Itô, 1953] Connection between Itô Integral, polynomial chaos expansion and expansions with multiple Wiener integrals. See also [Kallianpur, 1980].
Polynomial Chaos Expansion

More recent history

- **1980s**: Uncertainty Quantification via *Stochastic Finite Element Methods*, i.e., PDEs with random data, spatial part discretized via FE; randomness treated by Monte Carlo method, perturbation expansions, response surface methods

- **[Ghanem & Spanos, 1991]** *Spectral Stochastic Finite Element Method*
  - Seek random field solution to PDE with random input in tensor product space $X \otimes \Xi$
  
    \[ X : \text{function space appropriate for deterministic version of PDE} \]
  
    \[ \Xi := L^2(\Omega, \mathcal{A}, P), \quad \text{for probability space } (\Omega, \mathcal{A}, P) \]

Discretization

- finite dimensional noise assumption
- $L^2$-RV approximated by multivariate Hermite polynomials in finite number of Gaussian RVs, inspired by PC expansions.
Polynomial Chaos Expansion

More recent developments

[Xiu & Karniadakis, 2002-03] *Generalized Polynomial Chaos (gPC)*

**Observation:** Multivariate polynomials in non-Gaussian basic random variables sometimes have better approximation properties than PC expansions.

[E., Mugler, Starkloff & Ullmann, 2012]

**Question:** gPC same approximation properties as PC? A: not always. Extension of Cameron Martin theorem to gPC.
(Ω, ℂ, P) : probability space, sufficiently rich that it is possible to define on it non-degenerate Gaussian random variables ξ ∼ N(0, σ²) with mean value zero and variance σ² > 0.

ξ : Ω → ℜ : random variable

E[ξ] : expectation

σ(S) : σ-algebra generated by set of RV S

L²(Ω, ℂ, P) : Hilbert space of real-valued RV w. finite second moments

∥ξ∥²_{L²} = E[ξ²] : associated norm (mean-square convergence)

ℋ : Gaussian linear (Hilbert) space: (complete) subspace of L²(Ω, ℂ, P) consisting of centered Gaussian RV

Note: ℋ cannot contain all Gaussian RV in underlying space.
Polynomial Chaos Expansion

Polynomial spaces

- For $M \in \mathbb{N}$ and $k \in \mathbb{N}_0$ denote by $\mathcal{P}^M_k$ the set of all $M$-variate polynomials of (complete) degree $n$, i.e.,

$$\mathcal{P}^M_k = \left\{ \sum_{|\alpha| \leq n} a_\alpha \xi^\alpha : \xi^\alpha = \xi_1^{\alpha_1} \cdots \xi_M^{\alpha_M}, a_\alpha \in \mathbb{R} \right\},$$

$$\alpha \in \mathbb{N}_0^M, \ |\alpha| = \alpha_1 + \cdots + \alpha_M.$$

- For Gaussian linear space $\mathcal{H}$ and $n \in \mathbb{N}_0$, set

$$\mathcal{P}_n(\mathcal{H}) := \left\{ p(\xi_1, \ldots, \xi_M) : p \in \mathcal{P}^M_n, M \in \mathbb{N}, \xi_j \in \mathcal{H}, j = 1, \ldots, M, \right\}.$$

- Then

$$\mathcal{P}_n(\mathcal{H}), \overline{\mathcal{P}}_n(\mathcal{H}) \subset L^2(\Omega, \mathcal{A}, P),$$

$$\mathcal{P}_0(\mathcal{H}) = \overline{\mathcal{P}}_0(\mathcal{H}) \quad \text{a.s. constant RV},$$

$$\mathcal{P}_1(\mathcal{H}), \overline{\mathcal{P}}_1(\mathcal{H}) \quad \text{Gaussian RV},$$

$$\left\{ \overline{\mathcal{P}}_n(\mathcal{H}) \right\}_{n \in \mathbb{N}_0} \quad \text{strictly increasing subspaces of } L^2(\Omega, \mathcal{A}, P).$$
Setting

\[ \mathcal{H}_0 := \mathcal{P}_0(\mathcal{H}) = \overline{\mathcal{P}}_0(\mathcal{H}), \]
\[ \mathcal{H}_n := \overline{\mathcal{P}}_n(\mathcal{H}) \cap \mathcal{P}_{n-1}(\mathcal{H})^\perp, \quad n \in \mathbb{N}, \]

we have

\[ \overline{\mathcal{P}}_n(\mathcal{H}) = \bigoplus_{k=0}^{n} \mathcal{H}_k. \]

We also set

\[ \bigoplus_{n=0}^{\infty} \mathcal{H}_n := \bigcup_{n=0}^{\infty} \mathcal{P}_n(\mathcal{H}). \]
Theorem 1.10 (Cameron & Martin, 1947)

\[ \bigoplus_{n=0}^{\infty} \mathcal{H}_n = L^2(\Omega, \sigma(\mathcal{H}), \mathcal{P}). \]

In particular, if \( \sigma(\mathcal{H}) = \mathcal{A} \), then

\[ L^2(\Omega, \mathcal{A}, \mathcal{P}) = \bigoplus_{n=0}^{\infty} \mathcal{H}_n. \]

**Note:** Condition \( \sigma(\mathcal{H}) = \mathcal{A} \) crucial.

Consider \( \xi \sim \mathcal{N}(0, 1) \), \( \mathcal{H} = \text{span} \{ \xi \} \), and \( \eta \in L^2(\Omega, \mathcal{A}, \mathcal{P}) \), \( \mathbf{E}[\eta] = 0 \), \( \xi, \eta \) independent. Then all orthogonal projections of \( \eta \) on \( \mathcal{H}_n \) vanish a.s., with approximation error \( \mathbf{E}[\eta^2] \).

[Janson, 1997]
Polynomial Chaos Expansion

Chaos expansion as sum of projections

\[ \mathcal{H} : \text{Gaussian linear space,} \]

\[ P_k : L^2(\Omega, \mathcal{A}, \mathbf{P}) \to \mathcal{H}_k : \text{orthogonal projection onto} \ \mathcal{H}_k \]

Polynomial chaos expansion of \( \eta \in L^2(\Omega, \sigma(\mathcal{H}), \mathbf{P}) \) given by

\[ \eta = \sum_{k=0}^{\infty} P_k \eta. \]

When \( \sigma(\mathcal{H}) \subsetneq \mathcal{A} \) expansion still (mean-square) convergent, but in this case to orthogonal projection of \( \eta \) onto \( L^2(\Omega, \sigma(\mathcal{H}), \mathbf{P}) \).
Polynomial Chaos Expansion

Computational realization

In applications typically have

\[ \mathcal{H} = \text{span}\{\xi_m : m \in \mathbb{N}\}, \quad \xi_m \sim \mathcal{N}(0,1) \text{ independent "basic" RV}. \]

Orthonormal basis of \( \mathcal{H} \) given by \( \{H_\alpha : |\alpha|_0 < \infty\} \), where

\[ \alpha \in \{(\alpha_1, \alpha_2, \ldots) : \alpha_m \in \mathbb{N}_0\}, \quad |\alpha|_0 := |\{m : \alpha_m > 0\}|, \]

\[ H_\alpha(\xi) = \prod_{\alpha_m \neq 0} H_{\alpha_m}(\xi_m), \quad \xi = (\xi_m)_{m \in \mathbb{N}}, \]

where \( \{H_k\}_{k \in \mathbb{N}_0} \) denotes the system of univariate normalized Hermite polynomials.

For finitely many basic RV \( \xi_1, \ldots, \xi_M \) and \( P_n^M(\xi_1, \ldots, \xi_M) \) the \( M \)-variate polynomials in \( \{\xi_m\}_{m=1}^M \) of degree at most \( n \), there holds

\[ \eta_n^M := P_n^M \eta \xrightarrow{n,M \to \infty} \eta \quad \forall \eta \in L^2(\Omega, \sigma(\{\xi_m\}_{m \in \mathbb{N}})), \mathcal{P}). \]
Consider a smooth transformation

\[
a = a(x, \omega) = f(G(x, \omega)), \quad x \in D \subset \mathbb{R}^d,
\]

of a Gaussian random field \( G = G(x, \omega) \) given by its Karhunen-Loève expansion

\[
G(x, \omega) = \mathbb{E}[G(x)] + \sum_{m=1}^{\infty} \sqrt{\lambda_m} g_m(x) \xi_m(\omega), \quad \xi_m \sim \mathcal{N}(0, 1) \text{ i.i.d.}
\]

The coefficients \( a_\alpha(x) \) of the polynomial chaos expansion

\[
a(x, \omega) = \sum_\alpha a_\alpha(x) H_\alpha(\xi(\omega))
\]

satisfy (cf. [Malliavin, 1997])

\[
a_\alpha(x) = \mathbb{E}[a(x, \omega) H_\alpha(\xi(\omega))] = \frac{1}{\sqrt{\alpha!}} \mathbb{E}[D^\alpha f(G(x, \xi(\omega)))].
\]
Polynomial Chaos Expansion

Example: PC expansion of lognormal random field

Special case: lognormal random field \( a(x, \omega) = e^{G(x, \omega)} \).

Here we obtain

\[
a_\alpha(x) = \frac{E[a(x)]}{\sqrt{\alpha!}} \prod_{m=1}^{\infty} \left( \sqrt{\lambda_m g_m(x)} \right)^{\alpha_m}.
\]
Polynomial Chaos Expansion
Generalized polynomial chaos

**Idea:** If RV \( \eta \) “far from Gaussian”, expand it in polynomials of RV with non-Gaussian distributions.

Many common probability distributions correspond to classical real orthogonal polynomials, e.g.,

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Polynomials</th>
<th>Density</th>
<th>Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian</td>
<td>Hermite</td>
<td>( \rho(\xi) = \frac{1}{\sqrt{2\pi}} e^{-\xi^2/2} )</td>
<td>( \mathbb{R} )</td>
</tr>
<tr>
<td>Gamma(( \alpha, \lambda ))</td>
<td>Laguerre</td>
<td>( \rho(\xi) = \frac{\lambda}{\Gamma(\alpha)} (\lambda\xi)^{\alpha-1} e^{-\lambda\xi} )</td>
<td>(0, ( \infty ))</td>
</tr>
<tr>
<td>Beta(( \alpha, \beta ))</td>
<td>Jacobi</td>
<td>( \rho(\xi) = \frac{(1-\xi)^{\alpha}(1+\xi)^{\beta}}{2^{\alpha+\beta+1}B(\alpha+1,\beta+1)} )</td>
<td>((-1, 1))</td>
</tr>
<tr>
<td>Uniform(( \alpha, \beta ))</td>
<td>Legendre</td>
<td>( \rho(\xi) = \frac{1}{\beta-\alpha} )</td>
<td>(( \alpha, \beta ))</td>
</tr>
<tr>
<td>Arcsin</td>
<td>Chebyshev</td>
<td>( \rho(\xi) = \frac{1}{\sqrt{1-\xi^2}} )</td>
<td>((-1, 1))</td>
</tr>
</tbody>
</table>

[Ogura, 1972] Poisson chaos (Charlier polynomials)
Assumption 1.11

Basic RV $\xi$ with

- finite moments $E[|\xi|^k]$ of all orders and
- continuous distribution function $F_\xi$.

Then there exists sequence $\{p_k\}_{k \in \mathbb{N}_0}$ of polynomials (deg $p_k = k$) orthonormal with respect to the distribution of $\xi$, i.e., in $L^2(\mathbb{R}, \mathcal{B}(\mathbb{R}), F_\xi(dx))$.

For any $\eta \in L^2(\Omega, \mathcal{A}, P)$ the coefficients $a_k$ of the expansion

$$\eta \sim \sum_{k=0}^{\infty} a_k p_k(\xi), \quad a_k = E[\eta p_k(\xi)]$$

are defined.

**Question:** does the expansion converge to $\eta$ in quadratic mean?
Equality in

\[ \eta = \sum_{k=0}^{\infty} a_k p_k(\xi) \quad \text{for all} \quad \eta \in L^2(\Omega, \sigma(\xi), P) \]

equivalent with density of polynomials

\[ p(\xi) \text{ in } L^2(\Omega, \sigma(\xi), P) \text{ or} \]
\[ p(x) \text{ in } L^2(\mathbb{R}, \mathcal{B}, F_\xi(dx)) \text{, respectively.} \]

**Theorem 1.12 (M. Riesz, 1923)**

The polynomials span \( \{\xi^k\}_{k \in \mathbb{N}_0} \) are dense in \( L^2(\Omega, \sigma(\xi), P) \) if and only if the Hamburger moment problem is uniquely solvable for the distribution of \( \xi \).
Definition 1.13

The moment problem is uniquely solvable for a probability distribution on \((\mathbb{R}, \mathcal{B}(\mathbb{R}))\) or the distribution is **determinate in the Hamburger sense**, if the distribution function is uniquely defined by the sequence of its moments

\[
\mu_k := \mathbb{E}[\xi^k] = \int_{\mathbb{R}} x^k F_{\xi}(dx), \quad k \in \mathbb{N}_0.
\]

**Thus:** generalized polynomial chaos expansions in one basic RV \(\xi\) converge if and only if the distribution of \(\xi\) is determinate.
Some determinate/indeterminate distributions

- determinate:
  - normal
  - uniform
  - beta
  - gamma
  - ...

- indeterminate:
  - lognormal
  - certain powers of Gaussian RV, e.g.
    \[ \xi^{2k+1} \text{ for any } k = 1, 2, \ldots \text{ or} \]
    \[ \xi^{2k} \text{ for any } k = 3, 4, \ldots \quad (\xi \sim N(0, 1)) \]
  - certain powers of exponentially distributed RV
  - ...

\(\xi\)
Theorem 1.14 (Generalized Cameron-Martin Theorem)

Let \( \{\xi_m\}_{m \in \mathbb{N}} \) be independent RV with continuous distributions and possessing moments all orders. Furthermore let \( \{\mathcal{H}_n\}_{n \geq 0} \) be the polynomial subspaces as in the Cameron-Martin theorem. Then the spaces \( \{\mathcal{H}_n\}_{n \geq 0} \) are mutually orthogonal closed subspaces of \( L^2(\Omega, \mathcal{A}, \mathcal{P}) \) and there holds

\[
\bigoplus_{n=0}^{\infty} \mathcal{H}_n = L^2(\Omega, \sigma(\{\xi_m\}_{m \in \mathbb{N}}), \mathcal{P})
\]

if and only if for each basic random variable \( \{\xi_m\}_{m \in \mathbb{N}} \) the moment problem for its distribution is uniquely solvable.

[E., Mugler, Starkloff & Ullmann, 2012]
Idea of proof:

- For one basic random variable $\xi_m$ orthonormal polynomials yield an orthonormal basis in $L^2(\mathbb{R}, \mathcal{B}(\mathbb{R}), F_{\xi_k}(dx_k))$.

- For finitely many independent basic random variables tensor products of univariate orthonormal polynomials yield an orthonormal basis in $L^2(\mathbb{R}^M, \mathcal{B}(\mathbb{R}^M), F_{\xi_1}(dx_1) \times \ldots \times F_{\xi_M}(dx_M))$.

- General case: approximation of random variables depending on $(\xi_1, \xi_2, \ldots)$ by random variables depending on a finite number of basic random variables.

**Note:** For nonindependent basic RV the condition is sufficient.
Consider standard lognormal RV $\xi = \exp(\gamma)$, $\gamma \sim N(0, 1)$.

PDF given by

$$f_\xi(x) = \begin{cases} \frac{1}{x\sqrt{2\pi}} e^{-\log^2 x} \frac{1}{2}, & x \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

This distribution is indeterminate in the Hamburger sense.
Polynomial Chaos Expansion

Generalized polynomial chaos: “lognormal chaos”

Associated orthonormal polynomials

\[ p_0(x) \equiv 1, \]
\[ p_k(x) = \frac{(-1)^k e^{k(k-1)/4}}{\sqrt{\prod_{i=1}^k (e^i - 1)}} \sum_{j=0}^k (-1)^j \begin{bmatrix} k \\ j \end{bmatrix} e^{-j^2 + j/2} x^j, \quad k \geq 1, \]

with

\[ \begin{bmatrix} k \\ j \end{bmatrix} = \frac{(1 - e^{-k})(1 - e^{-(k-1)}) \cdots (1 - e^{-(k-j+1)})}{(1 - e^{-j})(1 - e^{-(j-1)}) \cdots (1 - e^{-1})}. \]

Can be derived from Stieltjes-Wigert polynomials (cf. [Szegö, 1939])
Consider lognormal RV $\xi$ and $g : \mathbb{R} \to \mathbb{R}$ measurable, odd and 1-periodic for which
\[
\mathbb{E} \left[ [g(\log \xi)]^2 \right] < \infty,
\text{ e.g. } g(x) = \sin(2\pi x).
\]

Then for all $k \in \mathbb{N}$ we have
\[
a_k = \mathbb{E} [p_k(\xi) g(\log(\xi))] = \int_0^\infty p_k(x) g(\log(x)) f\xi(x) \, dx = 0,
\]
and therefore, for $\eta = g(\log(\xi)) \in L^2(\Omega, \sigma(\xi), \mathbb{P})$, $\eta \neq \sum_{k=0}^\infty a_k p_k(\xi)$.

**By contrast:** Hermite chaos expansion
\[
\sin(2\pi \log \xi) = \sum_{k=0}^\infty a_k H_k(\log \xi), \quad a_k = \begin{cases} 
\frac{(-1)^{(k-1)/2}(2\pi)^k}{\sqrt{k!}} e^{-2\pi^2}, & k \text{ odd}, \\
0, & k \text{ even},
\end{cases}
\]
with normalized Hermite polynomials converges in mean square.
Consider RV $\eta = \frac{1}{\xi}$, $\xi$ lognormal.

Lognormal chaos coefficients of $\eta$ given by

$$a_0 = \sqrt{e}, \quad a_k = (-1)^k e^{-(k^2 + 3k - 2)/4} \left( \prod_{j=1}^{k} (e^j - 1) \right)^{1/2}, \quad k \geq 1.$$ 

Partial sums of chaos expansion $\eta_n := \sum_{k=0}^{n} a_k p_k(\xi)$ can be bounded by

$$\|\eta_n\|_{L^2}^2 \leq \frac{e^2}{e - 1}.$$ 

Since $\|\eta\|_{L^2}^2 = e^2$ remainders of partial sums satisfy

$$\|\eta - \eta_n\|_{L^2}^2 = \|\eta\|_{L^2}^2 - \|\eta_n\|_{L^2}^2 \geq e^2 - \frac{e^2}{e - 1} > 0.$$
Therefore, for $\eta = \frac{1}{\xi} \in L^2(\Omega, \sigma(\xi), P)$, $\xi$ lognormal, we again have

$$\eta \neq \sum_{k=0}^{\infty} a_k p_k(\xi).$$
1 Polynomial Chaos Expansion

2 Random PDEs
Random PDEs
UQ and the scientific computing paradigm

Physical Phenomenon
Uncertain Data
Lack of Knowledge
Variability

Mathematical Model
SDEs
Random Fields

Prediction
Insight
Optimization
Control
Decision

Computer Implementation

Numerical Approximation

Oliver Ernst (TU Chemnitz)
Random PDEs
UQ and the scientific computing paradigm

Physical Phenomenon

Uncertain Data
Lack of Knowledge
Variability

Mathematical Model

SDEs
Random Fields

Prediction
Insight
Optimization
Control
Decision

Quantified

Computer Implementation

Numerical Approximation

Oliver Ernst (TU Chemnitz)
2 Random PDEs

2.1 A model problem

2.2 Representation of random fields

2.3 Stochastic Galerkin Discretization

2.4 Stochastic Collocation

2.5 Further PC-based methods
Random PDEs
Model Problem

The most popular model problem in the UQ community has become the second-order elliptic PDE with an uncertain coefficient function:

\[-\nabla \cdot (a \nabla u) = f \quad + \quad \text{domain } D \subset \mathbb{R}^d \quad + \quad \text{BC}.\]

Rather than the solution \(u\) (whatever that may be), typical UQ problems center on some functional \(Q\) of the solution, e.g. its value at a point in the computational domain, average over a subdomain, flux across a boundary etc. Such a functional is known as a quantity of interest (QoI).

**Examples:**

\[Q(u) = u(x_0), \quad Q(u) = \frac{1}{|D_0|} \int_{D_0} u(x) \, dx.\]

Introduce input set \(V\) containing all possible inputs \(a\) and associated output set \(W = \{Q(u(a)) : a \in V\}\) as well as mapping \(G : V \rightarrow W\) of inputs to outputs.

In what way might uncertainty in the coefficient \(a\) be addressed?
Idea: Some values (functions) $a \in V$ are more likely than others.

Purely probabilistic approach:

- Introduce probability measure on $V$.
- (Measurable) mapping $G : V \rightarrow W$ induces probability measure on $W$. ("uncertainty propagation").
- Big issue: choice of distribution, too much subjective information?
- Statistical tradition of “eliciting probability distributions from expert opinion”.
- Some classical guidelines: Laplace’s principle of insufficient reason, maximum entropy (information theory), etc.
- Choosing distribution based on data is point of departure for Bayesian inverse problem.

Alternatives: replace all uncertain quantities by a nominal value, worst-case analysis, fuzzy logic, evidence theory, ...  
[Halpern, 2003]
Random PDEs
Stochastic vs. random PDEs

• What we shall not consider:

  Stochastic PDEs, i.e., generalisations to function spaces of stochastic (ordinary) evolution equations driven by the Wiener process.

  [Da Prato & Zabczyk, 1992], [Prévot & Röckner, 2007].

  Characteristics: White noise, uncorrelated randomness.

• Random PDEs:

  PDEs with random data in which randomness displays significant correlations ("colored noise").

• Randomness enters PDE problems through

  • source terms,
  • boundary data,
  • domain.
  • We focus on random coefficients.
Consider model problem (isotropic stationary diffusion)

\[- \nabla \cdot (a \nabla u) = f \text{ on } D, \quad u|_{\partial D} = 0,\]

for bounded, polygonal domain $D \subset \mathbb{R}^d$, $f \in L^2(D)$, $a \in L^\infty(D)$ with uniform bounds

\[0 < a_{\min} \leq a(x) \leq a_{\max} < \infty, \quad x \in D.\]

Variational problem of finding $u \in X := H^1_0(D)$ such that

\[\int_D a \nabla u \cdot \nabla v \, dx = \int_D fv \, dx \quad \forall v \in X\]

well-posed,

\[\|u\|_X \leq C\|f\|_{X'} .\]
Now let $a = a(x, \omega)$ be a random field w.r.t. probability space $(\Omega, \mathcal{A}, P)$, i.e.,

$$a : D \times \Omega \to \mathbb{R}, \quad \text{jointly measurable w.r.t. } \mathcal{B}(D) \times \mathcal{A}.$$  

Interpretation:

- for each $x \in D$, the quantity $a(x, \cdot)$ is a random variable,
- for each $\omega \in \Omega$, the realization $a(\cdot, \omega)$ is an element of $L^\infty(D)$.

Stochastic problem (strong formulation):

seek RF $u : D \times \Omega \to \mathbb{R}$ such that, $P$-a.s.,

$$- \nabla \cdot \left( a(x, \omega) \nabla u \right) = f(x), \quad x \in \Omega, \quad u(x, \omega) = 0, \quad x \in \partial D.$$
Restrict all RV to space $L_p^2(\Omega)$, i.e., RV’s with **finite second moments**.

Solution is RV taking values in $H_0^1(D)$; require finite second moments in the sense

$$u \in \{ v : \Omega \to H_0^1(D) : \mathbb{E} \left[ \| v \|^2_{H_0^1(D)} \right] < \infty \} = L_p^2(\Omega; H_0^1(D)) \quad \simeq \quad H_0^1(D) \otimes L_p^2(\Omega) =: \mathcal{V}$$

with inner product

$$(u, v)_{\mathcal{V}} := \mathbb{E} \left[ (\nabla u, \nabla v)_{L^2(D)} \right].$$
Simplest assumptions on $a$: uniform bounds with probability 1:

$$0 < a_{\text{min}} \leq a(x, \omega) \leq a_{\text{max}} < \infty, \quad x \in D, \quad \mathcal{P}\text{-a.s.}$$

Then bilinear form

$$B(u, v) := \mathbf{E} \left[ (a \nabla u, \nabla v)_{L^2(D)} \right], \quad u, v \in \mathcal{V},$$

is coercive and bounded

$$|B(u, v)| \leq a_{\text{max}} \|u\|_{\mathcal{V}} \|v\|_{\mathcal{V}}, \quad B(v, v) \geq a_{\text{min}} \|v\|^2_{\mathcal{V}}, \quad u, v \in \mathcal{V},$$

Linear form

$$\ell(v) := \mathbf{E} \left[ (f, v)_{L^2(D)} \right], \quad v \in \mathcal{V}$$

is continuous on $\mathcal{V}$. 
Standard application of Lax-Milgram lemma now yields:

**Stochastic variational problem** of finding $u \in \mathcal{V}$ such that

$$B(u, v) = \ell(v) \quad \forall v \in \mathcal{V},$$

i.e.,

$$\mathbb{E} \left[ \int_D a \nabla u \cdot \nabla v \, dx \right] = \mathbb{E} \left[ \int_D f v \, dx \right] \quad \forall v \in \mathcal{V}$$

possesses a unique solution which satisfies

$$\|u\|_{\mathcal{V}} \leq \frac{1}{a_{\text{min}}} \left( \mathbb{E} \left[ \|f\|_{L^2(D)}^2 \right] \right)^{1/2}.$$

[Babuška et al., 2004], [Babuška et al., 2005], [Frauenfelder et al., 2005], [Todor & Schwab, 2007], [Bieri et al., 2009], [Bieri & Schwab, 2009]
Random PDEs
Weaker assumptions

In many applications a not bounded uniformly away from 0, ∞.

**Example:** lognormal RF, $a(x, \omega) = \exp(g(x, \omega))$, $g$ Gaussian RF.

If $0 < a_{\text{min}}$ but $a_{\text{max}} = \infty$ can obtain well-posed problem in the stochastic energy (Hilbert) space

$$\mathcal{V}_a := \left\{ v \in \mathcal{V} : \mathbf{E} \left[ (a \nabla v, \nabla v)_{L^2(D)} \right] < \infty \right\}.$$

Lower bound yields continuous imbedding of $\mathcal{V}_a$ in $\mathcal{V}$

$$\|v\|_{\mathcal{V}} \leq \frac{1}{\sqrt{a_{\text{min}}}} \|v\|_{\mathcal{V}_a},$$

and thus, with Poincaré-Friedrichs constant $C_{\text{PF}},$

$$\|u\|_{\mathcal{V}} \leq \frac{C_{\text{PF}}}{a_{\text{min}}} \left( \mathbf{E} \left[ \|f\|_{L^2(D)} \right] \right)^{1/2}.$$

[Babuška et al., 2007]
Random PDEs

Weaker assumptions

For the case $a_{\text{min}} = 0$ and $a_{\text{max}} = \infty$ formulations with weighted $L^2$-spaces have been recently introduced.

Consider the realization-wise bounded case

$$0 < a_{\text{min}}(\omega) \leq a(x, \omega) \leq a_{\text{max}}(\omega) < \infty \quad \text{a.e. and a.s}$$

for two real-valued RV $a_{\text{min}}, a_{\text{max}}$, and allow for a random source term $f \in L^2_P(\Omega) \otimes H^{-1}(D)$.

Introducing the bilinear form $B_\omega : H^1_0(D) \times H^1_0(D) \to \mathbb{R}$

$$B_\omega(u, v) := \int_D a(x, \omega) \nabla u \cdot \nabla v \, dx, \quad u, v \in H^1_0(D),$$

we can formulate the realization-wise problem of finding an $H^1_0(D)$-valued RV $u$ such that

$$B_\omega(u, v) = \langle f, v \rangle_{H^{-1}(D) \times H^1_0(D)} \quad \forall v \in H^1_0(D), \text{ a.s.}$$
Realization-wise application of Lax-Milgram yields unique solution \( u = u(\omega) \in H^1_0(D) \) \( P \)-a.e. satisfying

\[
\| u(\omega) \|_{H^1(D)} \leq \frac{C}{a_{\text{min}}(\omega)} \| f(\omega) \|_{H^{-1}(D)} \quad \text{a.s.} \quad (*)
\]

with \( C \) independent of \( \omega \).

[Mugler & Starkloff, 2010]: Realization-wise solution \( u : \Omega \rightarrow H^1_0(D) \) is a random variable measurable with respect to the \( \sigma \)-algebra \( \sigma(f, a) \).

Squaring \((*)\) and taking expectations yields

\[
E \left[ \| u \|_{H^1_0(D)}^2 \right] \leq C^2 E \left[ \frac{\| f \|_{H^{-1}(D)}^2}{a_{\text{min}}^2} \right]
\]

Thus: \( u \) has finite variance if second moment of \( \| f \|_{H^{-1}(D)} \) weighted by \( a_{\text{min}} \) finite.
Define for RV \( \nu > 0 \) a.s. the weighted \( L^2 \)-spaces

\[
L^2(\Omega, \mathcal{A}, \nu \, d\mathbf{P}) := \{ \xi : \Omega \rightarrow \mathbb{R} \text{ measurable} : \mathbb{E} [\nu \xi^2] < \infty \}
\]
as well as the tensor-product space

\[
\mathcal{V}_\nu := H^1_0(D) \otimes L^2(\Omega, \mathcal{A}, \nu \, d\mathbf{P})
\]

with dual space

\[
\mathcal{V}'_\nu := H^{-1}(D) \otimes L^2(\Omega, \mathcal{A}, \frac{1}{\nu} \, d\mathbf{P})
\]

**Stochastic variational formulation:** Given \( f \in \mathcal{V}'_{a_{\min}^2} \), find \( u \in \mathcal{V} \) such that

\[
B(u, v) = \langle f, v \rangle \quad \forall v \in \mathcal{V}_{a_{\min}^2}
\]

with duality pairing between \( \mathcal{V}'_{a_{\min}^2} \) and \( \mathcal{V}_{a_{\min}^2} \) given by

\[
\langle f, v \rangle = \mathbb{E} \left[ \langle f, v \rangle_{H^{-1}(D) \times H^1_0(D)} \right].
\]
Theorem 2.1 (Mugler & Starkloff, 2010)

For any \( f \in \mathcal{V}_{a_{\min}^2} \) there exists a unique solution \( u \in \mathcal{V} \) to the weighted stochastic variational problem satisfying

\[
\| u \|_{\mathcal{V}} \leq C \| f \|_{\mathcal{V}_{a_{\min}^2}}.
\]

Proof: inf-sup argument with continuous and dense subspaces.

Random PDEs

2.1 A model problem

2.2 Representation of random fields

2.3 Stochastic Galerkin Discretization

2.4 Stochastic Collocation

2.5 Further PC-based methods
Most computational methods employ a parametrization of RF based on an expansion separating deterministic and random quantities.

\[ a(x, \omega) = \bar{a}(x) + \sum_{m=1}^{\infty} a_m(x) \xi_m(\omega), \quad a_m \in L^\infty(D), \]

typically \( \bar{a}(x) = \mathbb{E}[a(x, \cdot)] \); then \( a = \bar{a} + \tilde{a} \) with \( \mathbb{E}[\tilde{a}] = 0 \).

Various possibilities for \( a_m \):

- Fourier modes
- eigenfunctions of (spatial) differential operator
- physically relevant quantities (e.g. buckling modes of a column)
- multiresolution analysis [Dahlke et al., 2010]
- eigenfunctions of covariance operator (Karhunen-Loève expansion)
- generalized spectral decomposition [Nouy, 2007]
**Random PDEs**

Representation of random fields: Karhunen-Loève expansion

**Covariance function** of RF $a \in L^2(D) \otimes L^2_P(\Omega)$

$$c(x, y) = c_a(x, y) := E \left[ \left( a(x, \cdot) - \bar{a}(x) \right) \left( a(y, \cdot) - \bar{a}(y) \right) \right], \quad x, y \in D,$$

is symmetric in $x, y$, positive semidefinite, and continuous on $D \times D$ if continuous along ‘diagonal’ {$(x, x) : x \in D$}.

The **covariance operator**

$$C = C_a : L^2(\Omega) \to L^2(\Omega), \quad (Cu)(x) = \int_D u(y) c(x, y) \, dy$$

is therefore selfadjoint, compact, nonnegative. Its eigenvalues $\{ \lambda_m \}_{m \in \mathbb{N}}$ form a nonincreasing sequence accumulating at most at 0.
Denoting eigenfunctions by \( \{a_m\}_{m \in \mathbb{N}} \) there exists sequence of RV

\[
\{\xi_m\}_{m \in \mathbb{N}} \subset L^2_\mathbb{P}(\Omega), \quad E[\xi_m] = 0, \quad E[\xi_k \xi_m] = \delta_{k,m},
\]

such that the expansion

\[
a(x, \omega) = \bar{a}(x) + \sum_{m=1}^{\infty} \sqrt{\lambda_m} a_m(x) \xi_m(\omega)
\]

converges uniformly w.r.t. \( x \) and in mean-square w.r.t. \( \omega \).

[Karhunen, 1947], [Loève, 1948]
For normalized eigenfunctions $a_m(x)$,

$$\text{Var}_a(x) := c(x, x) = \sum_{m=1}^{\infty} \lambda_m a_m(x)^2,$$

$$\int_D \text{Var}_a(x) \, dx = \sum_{m=1}^{\infty} \lambda_m (a_m, a_m)_D = \text{trace } C.$$

For constant variance (e.g., stationary RF),

$$\text{Var}_a \equiv \sigma^2 > 0, \quad \sum_m \lambda_m = |D| \sigma^2.$$

**Interpretation:** $M$ first covariance eigenmodes form best rank-$M$ approximation to $C$ in sense of retaining maximal amount of variance.
Random PDEs
Representation of random fields: truncated Karhunen-Loève expansion

For computations KL expansion often truncated after $M$ terms:

$$a^{(M)}(x, \omega) = \bar{a}(x) + \sum_{m=1}^{M} \sqrt{\lambda_m} a_m(x) \xi_m(\omega).$$

Truncation error

$$E \left[ \|a - a^{(M)}\|_{L^2(D)}^2 \right] = \sum_{m=M+1}^{\infty} \lambda_m.$$

Choose $M$ to retain sufficient fraction $\delta \in (0, 1)$ of total variance, i.e.,

$$\frac{E \left[ \|a - a^{(M)}\|_{L^2(D)}^2 \right]}{E \left[ \|a\|_{L^2(D)}^2 \right]} = \frac{\sum_{m=M+1}^{\infty} \lambda_m}{\sum_{m=1}^{\infty} \lambda_m} = 1 - \frac{\sum_{m=1}^{M} \lambda_m}{|D|\sigma^2} < \delta.$$
Random PDEs
Representation of random fields: isotropic covariance functions

\[ c(x, y) = c(r), \quad r = \|x - y\|_2 \]

Convenient parametrization: Matérn class of covariance kernels:

\[ c(r) = \frac{\sigma^2}{2^{\nu-1} \Gamma(\nu)} \left( \frac{2\sqrt{\nu} r}{\rho} \right)^\nu K_\nu \left( \frac{2\sqrt{\nu} r}{\rho} \right) \]

- \( K_\nu \): modified Bessel function of order \( \nu \)
- \( \nu \): smoothness parameter
- \( \rho \): correlation length parameter

Special cases:

1. \( \nu = \frac{1}{2} \):
   \[ c(r) = \sigma^2 \exp(-\sqrt{2}r/\rho) \]
   exponential covariance

2. \( \nu = 1 \):
   \[ c(r) = \sigma^2 \left( \frac{2r}{\rho} \right) K_1 \left( \frac{2r}{\rho} \right) \]
   Bessel covariance

3. \( \nu \to \infty \):
   \[ c(r) = \sigma^2 \exp(-r^2/\rho^2) \]
   Gaussian covariance
Random PDEs
Representation of random fields: isotropic covariance functions

\[ \rho = 1 \]

Smoothness of realizations: RF \( a \) is \( s \) times mean-square differentiable if and only if \( \nu > s \).
**Eigenvalue Decay:** the smoother the kernel, the faster $\{\lambda_m\}_{m \in \mathbb{N}} \to 0$.

More precisely: if $D \subset \mathbb{R}^d$, then if the kernel function $c$ is

- piecewise $H^s$: $\lambda_m \leq c_1 m^{-s/d}$
- piecewise smooth: $\lambda_m \leq c_2 m^{-s}$ for any $s > 0$
- piecewise analytic: $\lambda_m \leq c_3 \exp(-c_4 m^{1/d})$

for suitable constants $c_1, c_2, c_3, c_4$.

**Note:** Piecewise smoothness of kernel also leads to bounds on derivatives of eigenfunctions $a_m$ in $L^\infty(D)$.

Proven e.g. in [Schwab & Todor, 2006], [Todor, 2006]
Preasymptotic plateau (determined by correlation length $\rho$) before asymptotic decay sets in.

Rate:

$$\lambda_m \sim m^{-(1+2\nu/d)} \quad (m \to \infty)$$

[Lord, Powell & Shardlow, 2014], [Widom, 1963]

Eigenvalue decay, Matérn kernel,

$$D = [-1, 1].$$
Realizations of Gaussian field, Matérn covariance, $\nu = 1.5$ $\rho = 0.5$
Realizations of Gaussian field, Matérn covariance, $\nu = 1.5$ $\rho = 0.5$
Realizations of Gaussian field, Matérn covariance, $\nu = 1.5 \ \rho = 0.5$
Realizations of Gaussian field, Matérn covariance, $\nu = 1.5 \ \rho = 0.5$
Realizations of Gaussian field, Matérn covariance, $\nu = 1.5 \ \rho = 0.5$
Realizations of Gaussian field, Matérn covariance, $\nu = 1.5 \ \rho = 0.5$
Realizations of Gaussian field, Matérn covariance, \( \nu = 1.5 \ \rho = 0.02 \)
Realizations of Gaussian field, Matérn covariance, $\nu = 1.5 \quad \rho = 0.02$
Realizations of Gaussian field, Matérn covariance, $\nu = 1.5 \  \rho = 0.02$
Realizations of Gaussian field, Matérn covariance, $\nu = 1.5$ $\rho = 0.02$
Realizations of Gaussian field, Matérn covariance, $\nu = 1.5 \quad \rho = 0.02$
Realizations of Gaussian field, Matérn covariance, $\nu = 1.5 \quad \rho = 0.02$
Random PDEs

Deterministic parametric representation

- Parametrize input RF by sequence of independent basic RV
  \[ \{\xi_m\}_{m \in \mathbb{N}} =: \xi. \]

- If \( \xi_m \) has density \( \rho_m \) and image \( \Gamma_m := \xi_m(\Omega) \), then (Doob-Dynkin lemma)
  \[ L^2_p(\Omega) \simeq L^2_p(\Gamma), \quad \text{where} \quad \Gamma := \times_{m=1}^{\infty} \Gamma_m, \quad \rho = \prod_{m} \rho_m. \]

- Replace \( a(x, \omega), u(x, \omega) \) with \( a(x, \xi), u(x, \xi) \).

BVP becomes purely deterministic with (possibly) high-dimensional parameter.
2 Random PDEs

2.1 A model problem
2.2 Representation of random fields
2.3 Stochastic Galerkin Discretization
2.4 Stochastic Collocation
2.5 Further PC-based methods
Random PDEs
Deterministic parametric representation

Pointwise strong form: for every \( \xi \in \Gamma \) find \( u = u(x, \xi) \) such that

\[
- \nabla \cdot \left( a(x, \xi) \nabla u(x, \xi) \right) = f(x) \text{ in } D, \quad u|_{\partial D} = 0.
\]

Pointwise variational form: for every \( \xi \in \Gamma \) find \( u = u(\xi) \in H^1_0(\Omega) \) such that

\[
\int_{D} a(\cdot, \xi) \nabla u \cdot \nabla v \, dx = \int_{D} fv \, dx \quad \forall v \in H^1_0(D).
\]

Parametric variational form: find \( u \in H^1_0(\Omega) \otimes L^2_{\rho}(\Gamma) \) such that

\[
\int_{\Gamma} \rho(\xi) \int_{D} a(x, \xi) \nabla u(x, \xi) \cdot \nabla v(x, \xi) \, dx \, d\xi = \int_{\Gamma} \rho(\xi) \int_{D} f(x) v(x, \xi) \, dx \, d\xi \quad \forall v \in H^1_0(D) \otimes L^2_{\rho}(\Gamma).
\]
Random PDEs
Stochastic Galerkin Discretization


- **Deterministic variables:** use any discretization technique appropriate for problem without randomness, e.g., Galerkin FE with finite-dimensional subspace $X^h \subset H^1_0(D)$.

- **Parameter (stochastic) variables:** choose finite-dimensional subspace $\Xi^J \subset L^2(\Gamma)$, where

\[ J \subset \mathcal{F} := \{ \alpha \in \mathbb{N}_0^N : |\text{supp } \alpha| < \infty \}. \]

- Joint Galerkin subspace $\mathcal{V}^{h,J} = X^h \otimes \Xi^J \subset H^1_0(D) \otimes L^2(\Gamma) = \mathcal{V}$.

**Stochastic Galerkin problem:** find $u = u^{h,J} \in \mathcal{V}^{h,J}$ such that

\[ B(u, \nu) = \ell(\nu) \quad \forall \nu \in \mathcal{V}^{h,J}. \]

$N = N_D \cdot N_\Gamma$ degrees of freedom, $N_D := |X^h|$, $N_\Gamma := |\Xi^J|$. 

Oliver Ernst (TU Chemnitz)
Random PDEs
Stochastic Galerkin Discretization: Choices for parametric subspace

(1) $M$ fixed, “finite-dimensional noise”, $\xi = (\xi_1, \xi_2, \ldots, \xi_M)$
   
   (a) $\Gamma$ bounded: can use piecewise polynomials, cf. [Deb et al., 2001], [Babuška et al., 2004].
   
   (b) $\Gamma$ unbounded (e.g. Gaussian basic RV): can use wavelets, cf. [Knio et al., 2004]
       [Le Maître et al, 2004], [Le Maître et al, 2007]
       global polynomials
       - tensor product polynomials of (coordinate) degree $k$: $N_\Gamma = (k + 1)^M$
       - complete polynomials of (total) degree $k$: $N_\Gamma = \binom{M+k}{k}$
       - Smolyak grids (see collocation)
       - sparse polynomial chaos, cf. [Bieri & Schwab, 2009], [Bieri et al., 2009]

(2) $M$ unbounded, $\xi = (\xi_1, \xi_2, \ldots)$
   Best $N$-term approximation [Bieri et al., 2009], [Cohen et al., 2010],
Random PDEs
Stochastic Galerkin Discretization: Convergence rates

Strongest results so far [Cohen, De Vore & Schwab, 2010],
[Gittelson & Schwab, 2011]

• $0 < a_{\min} \leq a(x, \omega) \leq a_{\max} < \infty$, $f \in L^2(D)$,
• $\Gamma = [-1, 1]^\infty$, $\xi_m \sim U[-1, 1]$,
• Scaled KL eigenfunctions $\tilde{a}_m(x) := \sqrt{\lambda_m} a_m(x)$ satisfy
  $\{\|\tilde{a}_m\|_{L^\infty(D)}\}_{m \in \mathbb{N}} \in \ell^p(\mathbb{N})$
• Spatial FE spaces have approximation property
  \[
  \inf_{v_h \in X^h} \| w - v_h \|_X \leq CN_h^{-t} |w|_{H^2(D)}, \quad N_h := \dim X^h, \quad 0 < t < \frac{1}{d}
  \]
• then
  \[
  \| u - u_h,\mathcal{F} \|_{\mathcal{Y}} \leq CN_{\text{dof}}^{-\min\{s,t\}}, \quad s := 1 - \frac{1}{p}.
  \]
Recall Galerkin equations

\[ B(u, v) = \ell(v) \quad \forall v \in X^h \otimes \Xi \]

\[ B(u, v) = E \left[ \int_D a(x, \xi) \nabla u(x, \xi) \cdot \nabla v(x, \xi) \, dx \right], \]

\[ \ell(v) = E \left[ \int_D f(x)v(x, \xi) \, dx \right]. \]

Given bases

\[ X^h = \text{span}\{\phi_1(x), \ldots, \phi_{N_D}(x)\}, \quad \Xi = \text{span}\{\psi_1(\xi), \ldots, \psi_{N_\xi}(\xi)\}, \]

trial and test functions

\[ u(x, \xi) = \sum_{\ell,j} u_{\ell,j} \phi_\ell(x) \psi_j(\xi), \quad v(x, \xi) = \phi_k(x) \psi_i(\xi) \]
... Galerkin system has block form

\[
A = \begin{bmatrix} A_{1,1} & \cdots & A_{1,N_{\Gamma}} \\ \vdots & \ddots & \vdots \\ A_{N_{\xi},1} & \cdots & A_{N_{\xi},N_{\Gamma}} \end{bmatrix}, \quad u = \begin{bmatrix} u_1 \\ \vdots \\ u_{N_{\Gamma}} \end{bmatrix}, \quad f = \begin{bmatrix} f_1 \\ \vdots \\ f_{N_{\Gamma}} \end{bmatrix},
\]

where

\[
A_{i,j} = E[\psi_i \psi_j K(\xi)] \in \mathbb{R}^{N_{D} \times N_{D}}, \quad i,j = 1, \ldots, N_{\Gamma},
\]

\[
f_i = E[\psi_i] f \in \mathbb{R}^{N_{D}}, \quad i = 1, \ldots, N_{\Gamma},
\]

\[
u_j = \begin{bmatrix} u_{1,j} \\ \vdots \\ u_{N_{D},j} \end{bmatrix} \in \mathbb{R}^{N_{D}}, \quad j = 1, \ldots, N_{\Gamma}.
\]
Random PDEs
Stochastic Galerkin Discretization: stochastic Galerkin systems

Representation of $K(\xi)$

(a) Truncated KL Expansion of RF (linear in $\xi$).

$$a(x, \xi) = \bar{a}(x) + \sum_{m=1}^{M} a_m(x)\xi_m$$

$$[K(\xi)]_{k,\ell} = [K_0]_{k,\ell} + \sum_{m=1}^{M} [K_m]_{k,\ell}\xi_m, \quad [K_m]_{k,\ell} = (a_m\nabla\phi_k, \nabla\phi_\ell),$$

$$A = G_0 \otimes K_0 + \sum_{m=1}^{M} G_m \otimes K_m,$$

$$[G_0]_{i,j} = E[\psi_i\psi_j] = I, \quad [G_m]_{i,j} = E[\xi_m\psi_i\psi_j].$$
Random PDEs
Stochastic Galerkin systems, Example: Complete polynomials in $\xi$

$M = 4$ Gaussian RVs,

$nnz = 55$

$nnz = 155$

degree 2

color coding: $G_0, G_1, G_2, G_3, G_4$. 

degree 3
(b) Expansion of RF in PC basis (nonlinear in $\xi$).

$$a(x, \xi) = \bar{a}(x) + \sum_{\alpha} a_{\alpha}(x) \psi_{\alpha}(\xi), \quad \xi = \xi(\omega) \in \mathbb{R}^{M},$$

Results in Galerkin matrix

$$A = G_0 \otimes K_0 + \sum_{\alpha} G_{\alpha} \otimes K_{\alpha},$$

$$[G_{\alpha}]_{\beta,\gamma} = E[\psi_{\alpha} \psi_{\beta} \psi_{\gamma}], \quad [K_{\alpha}]_{k,\ell} = (a_{\alpha} \nabla \phi_k, \nabla \phi_\ell).$$

Lots of structure in stochastic Galerkin matrices $G_m, G_{\alpha}$

[E. & Ullmann, 2010].
lognormal RV $a$ expanded in Hermite PC of $M = 4$ Gaussian RV

degree 2

degree 3

$nz = 1070$

$nz = 1990$
Random PDEs
Stochastic Galerkin discretization: Algebraic equations

- **Preconditioned Krylov solvers:**

- **Multigrid Methods:**

- **Low-Rank Approximation:**
  [Blatman & Sudret, 2009]
  [Khoromskij & Schwab, 2011]
  [Grasedyck, Kressner & Tobler, 2013]
  [Khoromskij, 2015]
  [Dahmen, DeVore, Grasedyck & Süli, 2015]
Random PDEs

2.1 A model problem
2.2 Representation of random fields
2.3 Stochastic Galerkin Discretization
2.4 Stochastic Collocation
2.5 Further PC-based methods
Collocation methods are a long-established technique for solving integral or differential equations and are based on requiring the equation under consideration to hold at a finite number of collocation points sufficient to determine an approximate solution in an appropriate finite-dimensional function space.

They were introduced for solving PDEs with random inputs in [Xiu & Hesthaven, 2005] and [Babuška, Nobile & Tempone, 2007] and offer a number of attractive features:

- Like MC, they reduce to a series of uncoupled deterministic subproblems for which legacy code can be used essentially unmodified.
- Unlike MC, collocation can take advantage of smooth dependence of the solution on the random parameters to yield spectral convergence.
- Nonlinear problems pose no additional difficulty (unlike SG).
To approximate a parameter-dependent object $u = u(\xi)$ with values in an abstract
space $V$, fix a finite-dimensional subspace $V_N = \text{span}\{u_1, \ldots, u_N\} \subset V$ and set

$$u(\xi) \approx u_N(\xi) = \sum_{j=1}^{N} u_j \psi_j(\xi)$$

with coefficient functions $\psi_j : \Gamma \to \mathbb{R}$ determined by a fixed set of

collocation points $\{\xi_j\}_{j=1}^{N} \subset \Gamma$.

**Simplest choice for $\psi_j$:** Lagrange basis of multivariate (global) polynomials with
respect to a system

$$\Xi := \{\xi_j\}_{j=1}^{N} \subset \Gamma$$

of unisolvent nodes.
Full tensor product collocation: [Xiu & Hesthaven, 2005]  
[Babuska, Nobile & Tempone, 2007]

\[
\chi_k = \{\xi^k_1, \ldots, \xi^k_{n_k}\} \quad \text{univariate node sequence, } k \in \mathbb{N},
\]

\[
\Xi_k = \chi_k \times \cdots \times \chi_k = \{\xi^k_\alpha = (\xi^k_{\alpha_1}, \ldots, \xi^k_{\alpha_M}) : 1 \leq \alpha_j \leq n_k\}
\]

\[
l_k f = \sum_{j=0}^{n_k} f(\xi^k_j) \ell^k_j(\xi), \quad \ell^k_j \in \mathcal{P}_{n_k-1} \text{ univariate Lagrange basis},
\]

\[
u(\xi) \approx \sum_{|\alpha|_\infty \leq n_k} \nu(\xi_\alpha)(\ell_{\alpha_1}^k \cdots \ell_{\alpha_M}^k)(\xi) = [(l_k \otimes \cdots \otimes l_k)u](\xi)
\]

\[
|\Xi_k| = n^M_k.
\]

Smolyak sparse grid collocation: [Xiu & Hesthaven, 2005]  
[Nobile, Tempone & Webster, 2008]

Replace equal-order interpolation by linear combination of mixed-order tensor-product formulas.
Random PDEs
Stochastic Collocation, Smolyak operator

Difference (detail) operators

\[ \Delta_k := I_k - I_{k-1} \quad (k \in \mathbb{N}), \quad I_0 := 0. \]

Smolyak sparse grid approximation operator

\[ A_{q,M} := \sum_{|i|_1 \leq q+M} \Delta_{i_1} \otimes \cdots \otimes \Delta_{i_M}, \quad i \in \mathbb{N}^M, \quad \text{level } q \in \mathbb{N}_0. \]

[Griebel et al., 1992], [Wasilkowski & Woźniakowski, 1995]

\[ A_{q,M} = \sum_{q+1 \leq |k|_1 \leq q+M} c_{q,M}(k) \ 1_{k_1} \otimes \cdots \otimes 1_{k_M}, \quad k \in \mathbb{N}^M, \]

\[ c_{q,M}(k) = (-1)^{q+M-|k|_1} \binom{M-1}{q+M-|k|_1} \]

\( A_{q,M} \) is interpolatory whenever \( \chi_k \subset \chi_{k+1} \ \forall k \) [Barthelmann, Novak & Ritter, 2000].
Sparse grid nodes

\[ \Xi_{q,M} = \bigcup_{q+1 \leq |k|_1 \leq q+M} \chi_{k_1} \times \cdots \times \chi_{k_M}. \]

For Gaussian field choose \( \chi_k \) to be \textbf{Gauss-Hermite} nodes of order

\[ n_k = \begin{cases} 
1, & k = 1, \\
1 + 2^{k-1}, & k > 1. 
\end{cases} \]

Number collocation points contained in sparse grid \( \Xi_{q,M} \) (crudely) bounded by

\[ |\Xi_{q,M}| \leq (3eM)^{q+1}. \]
Random PDEs
Stochastic Collocation, sparse grid nodes

M=2, n₁=1

M=2, q=0
Random PDEs
Stochastic Collocation, sparse grid nodes

$M=2, n_2=3$  

$M=2, q=1$
Random PDEs
Stochastic Collocation, sparse grid nodes

M=2, $n_3=5$

M=2, $q=2$
Random PDEs
Stochastic Collocation, sparse grid nodes

\[ M = 2, \ n_4 = 9 \]

\[ M = 2, \ q = 3 \]
Random PDEs

Stochastic Collocation, sparse grid nodes

\[ M=2, n_5=17 \]

\[ M=2, q=4 \]
Random PDEs
Stochastic Collocation, sparse grid nodes

\[ M = 2, n_6 = 33 \]

\[ M = 2, q = 5 \]
Random PDEs
Stochastic Collocation, sparse grid nodes

\[ M=2, \quad q=5, \quad (k_1,k_2) = (1,5) \]
Random PDEs
Stochastic Collocation, sparse grid nodes

\[ M=2, \quad q=5, \quad (k_1, k_2) = (2, 4) \]
Random PDEs
Stochastic Collocation, sparse grid nodes

\begin{align*}
M = 2, \quad q = 5, \quad (k_1, k_2) = (3, 3)
\end{align*}
$M=2, \quad q=5, \quad (k_1,k_2) = (4,2)$
Random PDEs
Stochastic Collocation, sparse grid nodes

$M=2, \quad q=5, \quad (k_1,k_2) = (5,1)$
Random PDEs
Stochastic Collocation, sparse grid nodes

\[ M = 2, \quad q = 5, \quad (k_1, k_2) = (1, 6) \]
Random PDEs
Stochastic Collocation, sparse grid nodes

\[ M=2, \quad q=5, \quad (k_1,k_2) = (2,5) \]
Random PDEs
Stochastic Collocation, sparse grid nodes

\[ M=2, \quad q=5, \quad (k_1, k_2) = (3, 4) \]
Random PDEs
Stochastic Collocation, sparse grid nodes

$M=2, \quad q=5, \quad (k_1, k_2) = (4,3)$
Random PDEs
Stochastic Collocation, sparse grid nodes

\[ M=2, \quad q=5, \quad (k_1, k_2) = (5, 2) \]
Random PDEs
Stochastic Collocation, sparse grid nodes

\[ M=2, \quad q=5, \quad (k_1,k_2) = (6,1) \]
Show that parameter-to-solution map

\[ \Gamma \ni \xi \mapsto u(\xi) \in H^1_0(D) \]

may be analytically extended to a subdomain of \( \mathbb{C}^M \).
(Weighted spaces, Hartogs’ theorem)

Use analyticity to show algebraic convergence with respect to number of collocation points. [Babuska et al., 2007], [E. & Sprungk, 2014], [Nobile, Tamellini, Tesei & Tempone, 2016]

Analyticity domain smallest for dominant modes of parameter vector \( \xi \).

Currently feasible for \( M \approx 50 \).
2 Random PDEs
  2.1 A model problem
  2.2 Representation of random fields
  2.3 Stochastic Galerkin Discretization
  2.4 Stochastic Collocation
  2.5 Further PC-based methods
Pseudospectral methods: (Non-Intrusive Spectral Projection, NISP)

\[ u(\xi) = \sum_{\alpha} a_{\alpha} H_{\alpha}(\xi), \quad a_{\alpha} = \mathbb{E}[uH_{\alpha}] \approx \sum_{j} u(\xi_j)H_{\alpha}(\xi_j)w_j \]

[Reagan, Najm, Ghanem & Knio, 2003]
[Winokur, Kim, Bisetti, Le Maitre & Knio, 2016],
[Constantine, Eldred & Phipps, 2012], [Conrad & Marzouk, 2013]

Least-squares methods: (Discrete $L^2$-Projection) Choose \( \{a_{\alpha}\}_{\alpha \in \mathcal{I}} \) to minimize

\[ \sum_{j=1}^{N} \|u(\xi_j) - \sum_{\alpha} a_{\alpha} H_{\alpha}(\xi_j)\|^2, \quad N > |\mathcal{I}| \]

[Cohen, Davenport & Leviatan, 2013]
[Migliorati, Nobile & Tempone, 2015]

Surrogate modeling: Use PC approximation as a surrogate for the random PDE model in optimization, inference, inverse problems etc.
Dongbin Xiu and George Em Karniadakis.
The Wiener-Askey polynomial chaos for stochastic differential equations.

On the convergence of generalized polynomial chaos expansions.

Oliver G. Ernst and Elisabeth Ullmann.
On stochastic Galerkin matrices.

Claude J. Gittelson and Christoph Schwab.
Sparse tensor discretization of high-dimensional parametric and stochastic PDEs.

Albert Cohen and Ronald DeVore.
Approximation of high-dimensional parametric PDEs.

**Books:** [LeMaitre & Knio, 2011],
[Lord, Powell & Schardlow, 2014],
[Sullivan, 2015]
[Janson, 1997]