High-Dimensional Approximations for Parametric and Random PDEs

Markus Bachmayr
RWTH Aachen

Summer School on Applied Analysis 2025

Funding acknowledgements: DFB project 501389786, DFB SFB 1481, ERC CoG COCOA

Outline

- (Elliptic) PDEs with many parameters and the connection to PDEs with random coefficients
- An overview of methods applied to a simple model problem
- Notions of approximability: Kolmogorov widths and summability of coefficients
- Series expansions of random fields as inputs
- Adaptive methods and the interactions with spatial discretization refinements

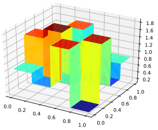
Parameter-dependent PDEs: Find $u = u(a) \in V$ such that $\mathcal{P}(a; u) = 0$, $a \in \mathcal{A}$

Elliptic model problem: $u \in V = H_0^1(D)$, $D \subset \mathbb{R}^d$, such that

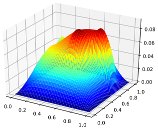
$$-\nabla \cdot (a \nabla u) = f \text{ in } D, \quad u = 0 \text{ on } \partial D$$

- Model order reduction:

  efficient approximation of $a \mapsto u(a)$



$\longrightarrow$

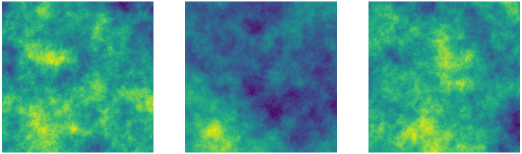


a $\qquad$ u

- Uncertainty quantification:

  probability measure on $\mathcal{A}$ modelling uncertainty in a,

  extract information on distribution of $u(a)$



Coefficient parametrizations: for $y \in Y$, find $u(y) \in V = H_0^1(D)$ such that

$$\int_D a(y)\nabla u(y) \cdot \nabla v \, dx = \int_D f v \, dx \quad \forall v \in V$$

- Piecewise constant model case: with partition $\{D_i\}$ of D, for $y \in Y = [-1,1]^P$,

$$a(y) = 1 + \theta \sum_{i=1}^{P} y_i \, \chi_{D_i}, \quad \theta \in (0,1)$$

- Affine parametrization with $y \in Y = [-1,1]^{\mathbb{N}}$,

$$a(y) = \bar{a} + \sum_{j=1}^{\infty} y_j \psi_j, \quad \bar{a}, \psi_j \in L^\infty(D)$$

  such that (uniform ellipticity): $0 < r \leq a(y) \leq R < \infty$ in D for all $y \in Y$.

- Lognormal coefficients: with $Y = \mathbb{R}^{\mathbb{N}}$,

$$a(y) = \exp\Big(\sum_{j\in\mathbb{N}} y_j \psi_j\Big), \quad y_j \sim \mathcal{N}(0,1),\ \psi_j \in L^\infty(D)$$

Aim: efficient approximations of $Y \ni y \mapsto u(y) \in V = H_0^1(D)$ or $y \mapsto Q(u(y))$

A. Cohen and R. DeVore, *Approximation of high-dimensional parametric PDEs*, Acta Numerica, 2015.

An overview of methods

- Model reduction:
  reduced bases (RB), proper orthogonal decomposition (POD), ...
- Polynomial approximations:
  stochastic Galerkin, stochastic collocation, discrete least squares, ...
- Sparse grids based on piecewise polynomials
- (Quasi-)Monte Carlo methods
- Kernel-based methods
- Low-rank tensor approximations
- Operator learning using neural networks

Suitability of methods can depend strongly on the type of PDE! For example, for transport problems such as

$$b \cdot \nabla u = f \quad + \text{inflow boundary conditions}$$

with parameter-dependent transport direction b, many of the above methods may become inefficient.

A simple example: one parameter

- $u \in V := H_0^1(D)$ with $\|v\|_V = \|\nabla v\|_{L^2(D)}$ such that

$$\int_D a \nabla u \cdot \nabla v \, dx = \int_D f \, v \, dx \quad \forall v \in V$$

- Take $\mathcal{A} = \{a(\cdot, y) \in L^\infty(D) \colon y \in [-1, 1]\}$ where

$$a(\cdot, y) = 1 + \theta y \chi_{D_1} \quad \text{with subdomain } D_1 \subset D,\ \theta \in (0, 1).$$

- To be solved:

$$\bar{B}\big(u(\cdot, y), v\big) + \theta y B_1\big(u(\cdot, y), v\big) = \langle f, v \rangle \quad \forall v \in V, y \in [-1, 1]$$

with $\bar{B}(w, v) = \int_D \nabla w \cdot \nabla v \, dx$, $B_1(w, v) = \int_{D_1} \nabla w \cdot \nabla v \, dx$, $\langle f, v \rangle = \int_D f \, v \, dx$.

- Uniform ellipticity: $0 < r \le a(\cdot, y) \le R$ for all $y \in [-1, 1]$ with $r := 1 - \theta$, $R := 1 + \theta$

- How to efficiently approximate $y \mapsto u(y) \in V$?
- Also: efficient evaluation of $Q(u(y))$ for some $Q \in V'$

Grid-based approximation

For grid points $-1 = y_0 < y_1 < \ldots < y_N = 1$,

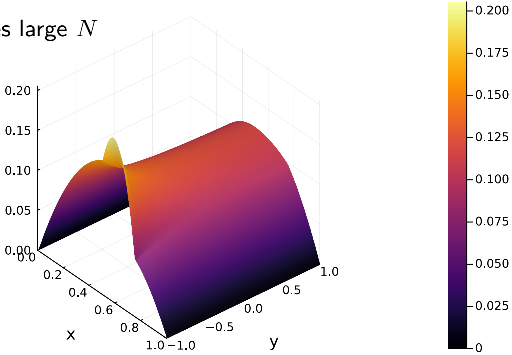
- Compute $u(y_i)$, $i = 0, \ldots, N$,
- Approximations for $y \in [y_i, y_{i+1}]$: with $Q \in V'$,

$$u(y) \approx \frac{y - y_i}{y_{i+1} - y_i} u(y_{i+1}) + \frac{y_{i+1} - y}{y_{i+1} - y_i} u(y_i)$$

$$Q(u(y)) \approx \frac{y - y_i}{y_{i+1} - y_i} Q(u(y_{i+1})) + \frac{y_{i+1} - y}{y_{i+1} - y_i} Q(u(y_i))$$

Limitation: small error requires large N

Example: $D = (0, 1)$,
$D_1 = (\frac{1}{2}, \frac{3}{4})$, $\theta = \frac{9}{10}$



Derivatives

Observation: $\partial_y^k u(\cdot, y) \in V$ exists for all $k \in \mathbb{N}$, $y \in [-1,1]$

Computing derivatives:

- $u(y) = u(\cdot, y)$ solves

$$\bar{B}\big(u(y), v\big) + \theta y B_1\big(u(y), v\big) = \langle f, v\rangle \quad \forall v \in V,$$

that is: $\displaystyle\int_D \nabla u(y) \cdot \nabla v \, dx + \theta y \int_D \chi_{D_1} \nabla u(y) \cdot \nabla v \, dx = \int_D f \, v \, dx \quad \forall v \in H_0^1(D).$

- Formal application of ∂_y (assuming differentiability) gives

$$\bar{B}\big(\partial_y u(y), v\big) + \theta y B_1\big(\partial_y u(y), v\big) = -\theta B_1(u(y), v) \quad \forall v \in V$$

- Differentiating once more:

$$\bar{B}\big(\partial_y^2 u(y), v\big) + \theta y B_1\big(\partial_y^2 u(y), v\big) = -2\theta B_1(\partial_y u(y), v) \quad \forall v \in V$$

- By induction: for $k \in \mathbb{N}$,

$$\bar{B}\big(\partial_y^k u(y), v\big) + \theta y B_1\big(\partial_y^k u(y), v\big) = -k\theta B_1(\partial_y^{k-1} u(y), v) \quad \forall v \in V$$

Taylor approximation

Note $\|v\|_V^2 = \bar{B}(v,v)$. Thus

$$\|\partial_y^k u(0)\|_V = \sup_{\|v\|_V=1} \bar{B}(\partial_y^k u(0), v) = \sup_{\|v\|_V=1} k\theta B_1(\partial_y^{k-1} u(0), v)$$
$$\leq k\theta \|\nabla \partial_y^{k-1} u(0)\|_{L^2(D_1)} \|\nabla v\|_{L^2(D_1)} \leq k\theta \|\partial_y^{k-1} u(0)\|_V$$

Applying this recursively and using the Lax-Milgram lemma,

$$\|\partial_y^k u(0)\|_V \leq k!\theta^k \|u(0)\|_V \leq k!\theta^k \|f\|_{V'}$$

Taylor expansion:

$$u(y) = \sum_{k=0}^{\infty} \frac{y^k}{k!} \partial_y^k u(0)$$

where

$$\left\| u(y) - \sum_{k=0}^{K} \frac{y^k}{k!} \partial_y^k u(0) \right\|_V \leq \frac{\sup_{y\in[-1,1]} \|\partial_y^{K+1} u(y)\|_V}{(K+1)!} |y|^{K+1},$$

and consequently

$$\sup_{y\in[-1,1]} \left\| u(y) - \sum_{k=0}^{K} \frac{y^k}{k!} \partial_y^k u(0) \right\|_V \leq \theta^{K+1}$$

Again: $N+1$ distinct $y_0, \dots, y_N \in [-1,1]$,

- Compute $u(y_i)$, $i = 0, \dots, N$,
- Interpolate by polynomial of degree N:

$$u \approx u_N, \quad u_N = \sum_{i=0}^{N} \left(\prod_{j \neq i} \frac{y - y_j}{y_i - y_j} \right) u(y_i)$$

(also known as stochastic collocation)

Lebesgue's lemma:

$$\sup_{y \in [-1,1]} \|u(y) - u_N(y)\|_V \leq \left(1 + \Lambda_N(y_0, \dots, y_N)\right) \min_{c_0, \dots, c_N \in V} \sup_{y \in [-1,1]} \left\| u(y) - \sum_{i=0}^{N} c_i y^i \right\|_V$$

with Lebesgue constant $\Lambda_N(y_0, \dots, y_N) = \sup_{y \in [-1,1]} \sum_{i=0}^{N} \left| \prod_{j \neq i} \frac{y - y_j}{y_i - y_j} \right|$.

Equidistant points	$\Lambda_N \approx 2^N$
Gauß points	$\Lambda_N \approx \sqrt{N}$
Chebyshev, Clenshaw-Curtis points	$\Lambda_N \approx \log(N+1)$

I. Babuška, F. Nobile, R. Tempone, *A stochastic collocation method for elliptic partial differential equations with random input data*, SIAM Rev, 2010.

Offline phase: choose $y_1, \dots, y_N \in [-1,1]$, set

$$V_N = \operatorname{span}\big\{u(y_i) \colon i = 1, \dots, N\big\}$$

Online phase: for each given y, find $u_N(y) \in V_N$ by solving

$$\bar B(u_N(y), v_N) + \theta y B_1(u_N(y), v_N) = \langle f, v_N \rangle \quad \forall v_N \in V_N$$

Precompute $N \times N$-matrices and N-vectors, then solve (dense) $N \times N$ system for each given y.

- Greedy basis refinement strategy:

$$y_{N+1} = \operatorname*{arg\,max}_{y \in [-1,1]} \big\| \langle f, \cdot \rangle - \bar B(u_N(y), \cdot) - \theta y B_1(u_N(y), \cdot) \big\|_{V'}$$

  (Difficulty: maximization over higher-dimensional parameter domains!)

- Quasi-optimality for each y, by Ceá's lemma: for all $y \in [-1,1]$,

$$\|u_N(y) - u(y)\|_V \le \sqrt{\frac{1+\theta}{1-\theta}} \min\big\{\|w_N - u(y)\|_V \colon w_N \in V_N\big\}.$$

Random coefficients

Probability measure μ on $[-1,1]$ (e.g. uniform measure $d\mu = \frac{1}{2}dy$)

Expectations:

$$\mathbb{E}u = \int_{-1}^{1} u(y)\, d\mu(y), \qquad \mathbb{E}Q(u) = \int_{-1}^{1} Q\big(u(y)\big)\, d\mu(y),$$

Monte Carlo Approximation: with $y_1, \dots, y_N \sim \mu$ i.i.d.,

$$\bar{u}_N := \frac{1}{N}\sum_{i=1}^{N} u(y_i), \quad \bar{Q}_N := \frac{1}{N}\sum_{i=1}^{N} Q\big(u(y_i)\big)$$

Standard convergence bound:

$$\mathbb{E}_{\otimes_{i=1}^{N}\mu}\|\bar{u}_N - \mathbb{E}u\|_V^2 = \frac{\mathbb{E}_\mu\|u - \mathbb{E}u\|_V^2}{N}$$

- Improved complexity by Multilevel MC and/or QMC methods
- Approximations of u in $\mathcal{V} = L^2([-1,1], V; \mu)$? Given $\varepsilon > 0$, find $\tilde{u}$ such that

$$\|u - \tilde{u}\|_{\mathcal{V}} = \left(\int_{-1}^{1} \|u(y) - \tilde{u}(y)\|_V^2\, d\mu(y)\right)^{1/2} \le \varepsilon$$

$$\mathbb{P}_N = \{p \text{ polynomial}\colon\ \deg p \le N\}$$

- Sequence of orthonormal polynomials for μ: $P_0 = 1$ and for $n \in \mathbb{N}$, $P_n \in \mathbb{P}_n$ such that $\|P_n\|_{L^2([-1,1];\mu)} = 1$, $P_n \perp \mathbb{P}_{n-1}$
- For uniform measure with $d\mu = \frac{1}{2}dy$, orthonormal Legendre polynomials given by three-term recurrence: $L_0 = 1$, $L_{-1} = 0$,

$$\sqrt{\beta_{k+1}}L_{k+1}(y) = yL_k(y) - \sqrt{\beta_k}L_{k-1}(y), \quad \beta_k := (4 - k^{-2})^{-1}.$$

  Rodrigues formula:

$$L_k(y) = \partial_y^k\left(\frac{\sqrt{2k+1}}{k!\,2^k}(y^2-1)^k\right)$$

- For any $u \in L^2([-1,1], V; \mu)$,

$$u(y) = \sum_{k=0}^{\infty} u_k L_k(y), \qquad u_k = \int_{-1}^{1} u(y)\, L_k(y)\, d\mu(y) \in V$$

Discrete least squares

- Polynomial approximations from $V \otimes \mathbb{P}_N$?
- One option: variational construction from samples
- Take $y_1, \dots, y_m \sim \rho$ i.i.d. with sampling measure ρ,

$$\tilde{u} = \operatorname*{arg\,min}_{v \in \mathbb{P}_N} \frac{1}{m} \sum_{i=1}^{m} \omega(y_i) \| u(y_i) - v(y_i) \|_V^2$$

- With $\tilde{u} = \sum_{k=0}^{N} \tilde{u}_k L_k$ and $\tilde{\mathbf{u}} = (u_k)_{k=0,\dots,N}$,

$$\mathbf{G} = \left(\sum_{i=1}^{m} \omega(y_i) L_k(y_i) L_\ell(y_i) \right)_{k,\ell=0,\dots,N}, \quad \mathbf{c} = \left(\frac{1}{m} \sum_{i=1}^{m} \omega(y_i) u(y_i) L_k(y_i) \right)_{k=0,\dots,N}$$

  solve $\mathbf{G}\tilde{\mathbf{u}} = \mathbf{c}$ in V^{N+1}
- With the right choice[1] of ρ and ω, we need $m \gtrsim N \log N$ (can be further improved, e.g. Nagel, Schäfer, Ullrich '20; Dolbeault, Cohen '22, ...) for quasi-optimal convergence in $\mathcal{V}$ in expectation
- Also need a spatial discretization: finite-dimensional subspace $\tilde{V} \subset V$
- Side note: closely related to "operator learning"

[1] A. Cohen and G. Migliorati, *Optimal weighted least-squares methods*, SMAI-JCM, 2017.

- Define, for all $v, w \in \mathcal{V}$,

$$\mathcal{B}(v,w) = \int_{-1}^{1} \bar{B}\big(v(y), w(y)\big) + \theta y B_1\big(v(y), w(y)\big)\, d\mu(y), \quad \langle F, v\rangle = \int_{-1}^{1} \langle f, v(y)\rangle\, d\mu(y)$$

- Then $\mathcal{B}$ is bounded and elliptic on $\mathcal{V}$, here:

$$\mathcal{B}(v,v) \geq (1-\theta)\|v\|_{\mathcal{V}}^2, \quad |\mathcal{B}(v,w)| \leq (1+\theta)\|v\|_{\mathcal{V}}\|w\|_{\mathcal{V}} \qquad \forall v, w \in \mathcal{V}$$

- By Lax-Milgram: $\mathcal{B}(u,v) = \langle F, v\rangle$ for all $v \in \mathcal{V}$ has unique solution $u \in \mathcal{V}$, agrees with $y \mapsto u(y)$ for μ-almost all y
- Let $\mathcal{V}_N = V \otimes \mathbb{P}_N$, that is, $\mathcal{V}_N \subset \mathcal{V} = L^2([-1,1], V; \mu)$ with

$$\mathcal{V}_N = \Big\{ \sum_{k=0}^{N} v_k L_k : v_i \in V,\ i = 0, \dots, N \Big\}$$

Stochastic Galerkin approximation $\tilde{u} \in \mathcal{V}_N$ defined by

$$\mathcal{B}(\tilde{u}, \tilde{v}) = \langle F, \tilde{v}\rangle \quad \forall \tilde{v} \in \mathcal{V}_N$$

Insert expansion $\tilde{u} = \sum_{\ell=0}^{N} u_\ell L_\ell$, test with vL_k for $v \in V$, $k = 0, \dots, N$:

$$\sum_{\ell=0}^{N} \left\{ \bar{B}(u_\ell, v) \int_{-1}^{1} L_\ell L_k \, d\mu + \theta B_1(u_\ell, v) \int_{-1}^{1} y L_\ell L_k \, d\mu \right\} = \langle f, v \rangle \int_{-1}^{1} L_k \, d\mu$$

Using

$$y L_k(y) = \sqrt{\beta_{k+1}} L_{k+1}(y) + \sqrt{\beta_k} L_{k-1}(y), \quad \beta_k = (4 - k^{-2})^{-1}$$

results in

$$\bar{B}(u_k, v) + \theta B_1\big(\sqrt{\beta_{k+1}} u_{k+1} + \sqrt{\beta_k}\, u_{k-1}, v\big) = \langle f, v \rangle \delta_{0,k}, \quad \forall v \in V, \ k = 0, \dots, N,$$

with discretizations $\bar{\mathbf{A}}$, $\mathbf{A}_1$ for $\bar{B}$, B_1:

$$\begin{pmatrix} \bar{\mathbf{A}} & \theta\sqrt{\beta_1}\mathbf{A}_1 & & & \\ \theta\sqrt{\beta_1}\mathbf{A}_1 & \bar{\mathbf{A}} & \theta\sqrt{\beta_2}\mathbf{A}_1 & & \\ & \theta\sqrt{\beta_2}\mathbf{A}_1 & \bar{\mathbf{A}} & \ddots & \\ & & \ddots & & \theta\sqrt{\beta_K}\mathbf{A}_1 \\ & & & \theta\sqrt{\beta_K}\mathbf{A}_1 & \bar{\mathbf{A}} \end{pmatrix} \begin{pmatrix} \mathbf{u}_0 \\ \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_K \end{pmatrix} = \begin{pmatrix} \mathbf{f} \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Multi-parametric case

General affine parameter-dependence: $\mathcal{I} = \{1, \dots, P\}$ or $\mathcal{I} = \mathbb{N}$,

$$a(y) = \bar{a} + \sum_{j \in \mathcal{I}} y_j \psi_j, \quad y \in Y = [-1, 1]^{\mathcal{I}}$$

with uniform ellipticity: $0 < r \leq a(y) \leq R < \infty$ in D for all $y \in Y$.

- Many sample-based methods (MC, RB, ...) keep their basic form, but selecting/drawing samples generally more difficult
- Methods based on polynomial approximations (stochastic collocation, stochastic Galerkin, ...) need to avoid the curse of (parametric) dimension

  $\rightsquigarrow$ sparse selection of polynomial degrees

Separation of variables

Rank-n expansions of parameter-dependent solution u,

$$u(y) \approx u_n(y) = \sum_{j=1}^{n} v_j \, \phi_j(y) \,, \quad v_j \in V,\ \phi_j : Y \to \mathbb{R}$$

- Reduced basis methods: solution snapshots $v_j := u(y^j)$, with $\phi_j(y)$ determined implicitly by Galerkin projection
- Approximation in $L^\infty(Y,V)$: Kolmogorov n-widths of $u(Y) \subset V$,

$$d_n\big(u(Y)\big)_V := \inf_{\substack{V_n \subset V \\ \dim(V_n)=n}} \sup_{y\in Y} \min_{v\in V_n} \|u(y) - v\|_V$$

- Controlling errors in $L^\infty(Y,V)$ problematic for high-dimensional Y

- Approximation in $L^2(Y, V; \mu)$, μ probability measure: Hilbert-Schmidt decomposition / SVD,

$$u = \sum_{j=1}^{n} \sigma_j \, \hat{v}_j \otimes \hat{\phi}_j, \quad \{\hat{v}_j\}, \{\hat{\phi}_j\} \text{ orthonormal},$$

best approximation by truncation, where

$$\sqrt{\sum_{j>n} \sigma_j^2} \leq d_n(u(Y))_V .$$

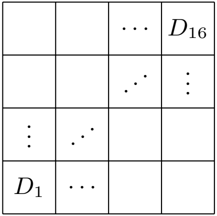
- Upper bounds for σ_j by prescribing $\hat{\phi}_j$, e.g. product orthonormal polynomial expansions in $L^2(Y, V; \mu)$: with $\mathcal{I} = \{1, \dots, P\}$ or $\mathcal{I} = \mathbb{N}$,

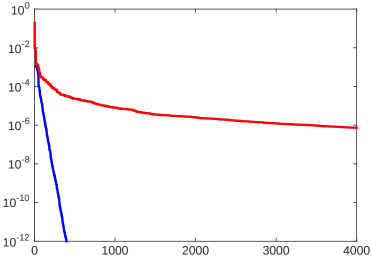
$$u(x, y) \approx \sum_{\nu \in \Lambda \subset \mathbb{N}_0^{\mathcal{I}}} u_\nu(x) \, L_\nu(y), \qquad L_\nu(y) := \prod_{i \in \mathcal{I}} L_{\nu_i}(y_i),$$

then $\sigma_j \leq \|u_{\nu_j^*}\|_V$ with decreasing rearrangement $\|u_{\nu_j^*}\|_V$

Piecewise constant a on partition $\{D_i\}$, with $\bar{a} := 1$:

$$a(y) = 1 + \sum_{i=1}^{P} y_i \, \psi_i \,, \qquad \psi_i := \theta \chi_{D_i} \,, \quad \theta < 1 \,.$$





red: ordered norms $\|u_\nu\|_V$ of Legendre coefficients in $u(y) = \sum_\nu u_\nu L_\nu(y)$,

blue: singular values σ_j in SVD $u(y) = \sum_j \sigma_j \hat{v}_j \, \hat{\phi}_j(y)$

Upper bounds for Kolmogorov widths (B., Cohen '17):

recombining linearly dependent terms in Taylor polynomial expansions in y

- Trivial: $d_n(u(Y)) \lesssim \exp\bigl(-|\ln\theta| n^{-1/P}\bigr)$
- For piecewise constant parameters: when $\sum_{j=1}^{P} \psi_j = \theta\bar{a}$,

$$d_n(u(Y)) \lesssim \exp\bigl(-|\ln\theta| n^{-1/(P-1)}\bigr)$$

- Using further spatial symmetries:

D_3	D_4
D_1	D_2

$P = 4$ with regular 2×2 checkerboard. Then for any $f \in V'$,

$$d_n\bigl(u(Y)\bigr)_V \leq C \exp\Bigl(-\frac{|\ln\theta|}{8} n\Bigr).$$

($\rightsquigarrow$ Autio, Hannukainen '25)

- Related: higher-order low-rank tensor approximations

M. Bachmayr and A. Cohen, *Kolmogorov widths and low-rank approximations of parametric elliptic PDEs*, Math Comp, 2017

Affinely parametrized linear elliptic PDEs

Parametric diffusion problem: for $y \in Y = [-1,1]^{\mathbb{N}}$, find $u(y) \in V = H_0^1(D)$ such that

$$\int_D a(y)\nabla u(y) \cdot \nabla v \, dx = \langle f, v \rangle, \quad \forall v \in V,$$

$$\text{where} \quad a(y) = \bar{a} + \sum_{j=1}^{\infty} y_j \psi_j, \quad \bar{a}, \psi_j \in L^\infty(D)$$

Uniform ellipticity assumption:

$$0 < r \le a(y) \le R < \infty\,, \quad \text{in } D, \text{ for all } y \in Y.$$

Here: for an $r > 0$,

$$\sum_{j \ge 1} |\psi_j| \le \bar{a} - r. \tag{UEA}$$

Objective: Approximate u in $L^\infty(Y, V)$ or $L^2(Y, V; \mu)$, with μ prob. measure on Y.

Product polynomial expansions

We set

$$\mathcal{F} := \{\nu \in \mathbb{N}_0^{\mathbb{N}} : \#\operatorname{supp}\nu < \infty\},$$

multi-index notation:

$$|\nu| := \sum_{j\geq 1} \nu_j, \quad \nu! := \prod_{j\geq 1} \nu_j!, \quad y^\nu = \prod_{j\geq 1} y_j^{\nu_j}.$$

Taylor expansion: $\quad u = \sum_{\nu\in\mathcal{F}} t_\nu\, y^\nu \quad$ with $t_\nu = \frac{1}{\nu!}\partial_y^\nu u(0) \in V$

Recursion for Taylor coefficients:

$$\int_D \bar{a}\nabla t_\nu \cdot \nabla v\, dx = -\sum_{j\in\operatorname{supp}\nu} \int_D \psi_j \nabla t_{\nu-e_j} \cdot \nabla v\, dx \quad \forall v \in V$$

Product orthonormal polynomial expansions

- In what follows: μ uniform measure on Y
- Univariate Legendre polynomials $\{L_k\}_{k\in\mathbb{N}_0}$, orthonormal in $L^2([-1,1];\mu)$ (analogously: general beta distributions and Jacobi polynomials)
- For $\nu \in \mathcal{F}$, define product polynomials

$$L_\nu(y) := \prod_{j\geq 1} L_{\nu_j}(y_j), \quad y \in Y = [-1,1]^{\mathbb{N}},$$

then $\{L_\nu\}_{\nu\in\mathcal{F}}$ is orthonormal basis of $L^2(Y;\mu)$

- Legendre expansion (for μ uniform measure):

$$u = \sum_{\nu\in\mathcal{F}} u_\nu \, L_\nu(y), \quad u_\nu = \int_Y u(y)\, L_\nu(y)\, d\mu(y)$$

- By orthonormality, for any $v \in \mathcal{V} = L^2(Y, V; \mu)$,

$$\|v\|_{\mathcal{V}}^2 = \int_Y \|v(y)\|_V^2 \, d\mu(y) = \sum_{\nu\in\mathcal{F}} \left\| \int_Y v(y)\, L_\nu(y)\, d\mu(y) \right\|_V^2 .$$

Summability (norm-based)

"Stechkin's lemma:" $0 < p < q < \infty$, $(c_\nu)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F})$ a sequence of positive numbers, $\Lambda_n \subset \mathcal{F}$ a set of indices with n largest c_ν. Then

$$\Big(\sum_{\nu\notin\Lambda_n} c_\nu^q\Big)^{1/q} \le C(n+1)^{-s}, \quad C := \|(c_\nu)_{\nu\in\mathcal{F}}\|_{\ell^p}, \quad s := \frac{1}{p} - \frac{1}{q}.$$

Theorem (Cohen, DeVore, Schwab '11).
Assume that (UEA) holds and $(\|\psi_j\|_{L^\infty})_{j\ge 1} \in \ell^p(\mathbb{N})$ for a $p \in (0,1)$, then $(\|t_\nu\|_V)_{\nu\in\mathcal{F}}$ and $(\|u_\nu\|_V)_{\nu\in\mathcal{F}}$ belong to $\ell^p(\mathcal{F})$.

Proof based on holomorphic extension in the parameter domain.

Best n-term approximation: Take $\Lambda_{\mathsf{T},n}, \Lambda_{\mathsf{L},n} \subset \mathcal{F}$ corresponding to n largest coefficients,

$$\sup_{y\in Y} \Big\| u(y) - \sum_{\nu\in\Lambda_{\mathsf{T},n}} t_\nu\, y^\nu \Big\|_V \le C n^{-\frac{1}{p}+1}, \quad \Big\| u - \sum_{\nu\in\Lambda_{\mathsf{L},n}} u_\nu\, L_\nu \Big\|_{L^2(Y,V,\mu)} \le C n^{-\frac{1}{p}+\frac{1}{2}}.$$

A. Cohen, R. DeVore, and Ch. Schwab, *Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDE's*, Analysis and Applications, 2011.

Summability (using weights)

Basic idea: improved results for ψ_j with spatial localization, still with basic assumption

$$\sum_{j\geq 1} |\psi_j| \leq \bar{a} - r. \tag{UEA}$$

Theorem (B., Cohen, Migliorati '17).
Let (UEA) hold and with $\rho_j > 1$, $j \in \mathbb{N}$, let

$$\sum_{j\geq 1} \rho_j |\psi_j| \leq \bar{a} - s \quad \text{for some } s > 0. \tag{UEA*}$$

Then

$$\sum_{\nu\in\mathcal{F}} \rho^{2\nu} \|t_\nu\|_V^2 < \infty, \qquad \sum_{\nu\in\mathcal{F}} \Big(\prod_{j\geq 1} (2\nu_j + 1) \Big)^{-1} \rho^{2\nu} \|u_\nu\|_V^2.$$

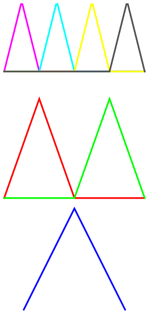
M. Bachmayr, A. Cohen, and G. Migliorati, *Sparse polynomial approximation of parametric elliptic PDEs. Part I: affine coefficients*, ESAIM M2AN, 2017

$y = (y_{\ell,m})_{\ell,m}$ with $y_{\ell,m} \sim \mathcal{U}(-1,1)$ i.i.d., and a with affine parameterization,

$$a(y) = a_0 + \sum_{\ell,m} y_{\ell,m}\psi_{\ell,m},$$

where $\sup_{x\in D} \sum_m \bigl|\psi_{\ell,m}(x)\bigr| \lesssim 2^{-\alpha\ell}$ for all $\ell \geq 0$

$\rightsquigarrow$ choose weights with $\rho_{\ell,m} \eqsim 2^{\beta\ell}$ with $\beta < \alpha$



Generally: $\sum_{j\geq 1} \rho_j|\psi_j| \leq \bar{a} - s$ with $\rho_j \to \infty$ is possible even when the decay of $\|\psi_j\|_{L^\infty}$ is slow, due to localized supports of the ψ_j.

M. Bachmayr, A. Cohen, and G. Migliorati, *Sparse polynomial approximation of parametric elliptic PDEs. Part I: affine coefficients*, ESAIM M2AN, 2017

Estimates for Taylor coefficients

1. ℓ^2-estimates, claim: (UEA) implies $\sum_{\nu\in\mathcal{F}} \|t_\nu\|_V^2 < \infty$.

(UEA) $\sum_{j\geq 1}|\psi_j| \leq \bar{a} - r$ with $r > 0$ is equivalent to $\theta := \left\| \frac{1}{\bar{a}} \sum_{j\geq 1} |\psi_j| \right\|_{L^\infty} < 1$, from which we obtain

$$\sum_{j\geq 1} \int_D |\psi_j||\nabla t_\nu|^2\, dx \leq \theta \int_D \bar{a}|\nabla t_\nu|^2\, dx, \quad \nu \in \mathcal{F}. \tag{$*$}$$

Recursion for Taylor coefficients: $\int_D \bar{a}\nabla t_0 \cdot \nabla v\, dx = \langle f, v\rangle_{V',V}$ and, for $\nu \neq 0$,

$$\int_D \bar{a}\nabla t_\nu \cdot \nabla v\, dx = - \sum_{j\in \operatorname{supp}\nu} \int_D \psi_j \nabla t_{\nu-e_j} \cdot \nabla v\, dx, \quad v \in V,$$

which gives, with Young's inequality,

$$\begin{aligned}\int_D \bar{a}|\nabla t_\nu|^2\, dx &\leq \sum_{j\in \operatorname{supp}\nu} \int_D |\psi_j||\nabla t_{\nu-e_j}||\nabla t_\nu|\, dx \\ &\leq \frac{1}{2} \sum_{j\in \operatorname{supp}\nu} \left(\int_D |\psi_j||\nabla t_{\nu-e_j}|^2\, dx + \int_D |\psi_j||\nabla t_\nu|^2\, dx \right).\end{aligned}$$

By ($*$),

$$\left(1 - \frac{\theta}{2}\right) \int_D \bar{a}|\nabla t_\nu|^2\, dx \leq \frac{1}{2} \sum_{j\in \operatorname{supp}\nu} \int_D |\psi_j||\nabla t_{\nu-e_j}|^2\, dx.$$

Summing over $|\nu| = k$,

$$\left(1-\frac{\theta}{2}\right)\sum_{|\nu|=k}\int_D \bar{a}|\nabla t_\nu|^2\,dx \le \frac{1}{2}\sum_{|\nu|=k}\sum_{j\in\operatorname{supp}\nu}\int_D |\psi_j||\nabla t_{\nu-e_j}|^2\,dx$$
$$= \frac{1}{2}\sum_{|\nu|=k-1}\sum_{j\ge 1}\int_D |\psi_j||\nabla t_\nu|^2\,dx \le \frac{\theta}{2}\sum_{|\nu|=k-1}\int_D \bar{a}|\nabla t_\nu|^2\,dx.$$

Therefore, since $\|v\|_V^2 := \int_D \bar{a}|\nabla v|^2\,dx$,

$$\sum_{|\nu|=k}\|t_\nu\|_V^2 \le \kappa \sum_{|\nu|=k-1}\|t_\nu\|_V^2, \quad \kappa := \frac{\theta}{2-\theta} < 1, \qquad \text{in particular} \quad \sum_{\nu\in\mathcal{F}}\|t_\nu\|_V^2 < \infty.$$

2. Weighted ℓ^2-estimates from strengthened UEA

The assumption $\sum_{j\ge 1}\rho_j|\psi_j| \le \bar{a} - s$ is equivalent to $\left\|\frac{1}{\bar{a}}\sum_{j\ge 1}\rho_j|\psi_j|\right\|_{L^\infty} < 1.$

This is (UEA) for the modified coefficient $a_\rho(y) := a(D_\rho y)$, where $D_\rho y := (\rho_j y_j)_{j\ge 1}$, with Taylor coefficients

$$t_{\rho,\nu} = \frac{1}{\nu!}\partial^\nu u_\rho(0) = \rho^\nu t_\nu, \quad u_\rho(y) = u(D_\rho y).$$

Applying the above gives

$$\sum_{\nu\in\mathcal{F}}\left(\rho^\nu\|t_\nu\|_V\right)^2 = \sum_{\nu\in\mathcal{F}}\|t_{\rho,\nu}\|_V^2 < \infty.$$

Sketch: Estimates for Legendre coefficients

$$u_\nu = \int_Y u(y)\, L_\nu(y)\, d\mu(y), \quad L_\nu(y) = \prod_{j\ge 1} L_{\nu_j}(y_j) = \prod_{j\ge 1} \partial_{y_j}^{\nu_j}\left(\frac{\sqrt{2\nu_j+1}}{\nu_j!\, 2^{\nu_j}}(y_j^2-1)^{\nu_j}\right)$$

By the latter Rodrigues' formula,

$$u_\nu = \Big(\prod_{j\ge 1}\sqrt{2\nu_j+1}\Big)\int_Y \frac{1}{\nu!}\partial^\nu u(y) \prod_{j\ge 1}\frac{(1-|y_j|)^{\nu_j}(1+|y_j|)^{\nu_j}}{2^{\nu_j}}\, d\mu(y) \qquad (*)$$

For fixed $y \in Y$: set $w_y(z) := u(T_y z)$ for $z \in Y$, where $T_y z := \big(y_j + (1-|y_j|)\rho_j z_j\big)_{j\ge 1}$, then

$$\partial^\nu w_y(0) = \Big(\prod_{j\ge 1}(1-|y_j|)^{\nu_j}\Big)\rho^\nu \partial^\nu u(y)\,.$$

Using modified equation with affine structure for w_y, previous arguments yield

$$\sum_{\nu\in\mathcal{F}} \Big\| \frac{1}{\nu!}\partial^\nu w_y(0)\Big\|_V^2 \le C < \infty$$

with $C > 0$ independent of y.

Combine with $(*)$ to show the weighted ℓ^2-estimate

$$\sum_{\nu\in\mathcal{F}}\Big(\prod_{j\ge 1}(2\nu_j+1)\Big)^{-1}\rho^{2\nu}\|u_\nu\|_V^2 \le \int_Y \sum_{\nu\in\mathcal{F}}\Big\|\frac{1}{\nu!}\partial^\nu w_y(0)\Big\|_V^2\, d\mu(y) \le C.$$

We thus have

$$\sum_{j\geq 1} \rho_j |\psi_j| \leq \bar{a}-s \quad \text{implies} \quad \sum_{\nu\in\mathcal{F}} \rho^{2\nu} \|t_\nu\|_V^2 < \infty, \ \sum_{\nu\in\mathcal{F}} \Big(\prod_{j\geq 1} (2\nu_j+1) \Big)^{-1} \rho^{2\nu} \|u_\nu\|_V^2 < \infty.$$

Corollary. Let $0 < p < 2$ and assume that for $q = q(p) := \frac{2p}{2-p}$, there exists a sequence $\rho = (\rho_j)_{j\geq 1}$ with $\rho_j > 1$ satisfying (UEA*) and $\big(\rho_j^{-1}\big)_{j\geq 1} \in \ell^q(\mathbb{N})$. Then $\big(\|t_\nu\|_V\big)_{\nu\in\mathcal{F}}$ and $\big(\|u_\nu\|_V\big)_{\nu\in\mathcal{F}}$ belong to $\ell^p(\mathcal{F})$.

Proof: By Hölder's inequality,

$$\sum_{\nu\in\mathcal{F}} \|t_\nu\|_V^p \leq \Big(\sum_{\nu\in\mathcal{F}} \rho^{2\nu} \|t_\nu\|_V^2 \Big)^{p/2} \Big(\sum_{\nu\in\mathcal{F}} \rho^{-q\nu} \Big)^{(2-p)/2}.$$

Now use that $\rho_j > 1$ and $(\rho_j^{-1})_{j\geq 1} \in \ell^q(\mathbb{N})$ imply $\sum_{\nu\in\mathcal{F}} \rho^{-q\nu} < \infty$ (very similar for Legendre coefficients). $\square$

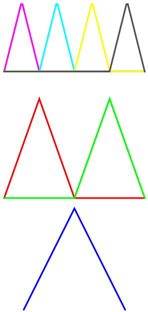
Wavelet-type parametrization

$y = (y_{\ell,m})_{\ell,m}$ with $y_{\ell,m} \sim \mathcal{U}(-1,1)$ i.i.d., and a with affine parameterization,

$$a(y) = a_0 + \sum_{\ell,m} y_{\ell,m}\psi_{\ell,m},$$

where $\sup_{x\in D} \sum_m \left|\psi_{\ell,m}(x)\right| \lesssim 2^{-\alpha\ell}$ for all $\ell \geq 0$

$\rightsquigarrow$ weights with $\rho_{\ell,m} \eqsim 2^{\beta\ell}$ with $\beta < \alpha$



Convergence of product Legendre expansions

Take $\Lambda_n \subset \mathcal{F}$ as indices of n largest $\|u_\nu\|_V$ in the expansion $u = \sum_{\nu\in\mathcal{F}} u_\nu L_\nu$.

Then

$$\left\| u - \sum_{\nu\in\Lambda_n} u_\nu L_\nu \right\|_{L^2(Y,V,\mu)} \lesssim n^{-s} \quad \text{for any } s < \frac{\alpha}{d}$$

Note: $a \in C^{0,\beta}(\overline{D})$ a.s. for any $\beta < \alpha$, implies $u \in H^{1+s}(D)$ a.s. for any $s < \alpha$

$\rightsquigarrow$ Work (at best) $\mathcal{O}(\varepsilon^{-\frac{d}{s}})$ for finite element approximation of one realisation of u

Lognormal coefficients: $a = \exp(b)$ with b Gaussian random field

Starting point: expansion

$$b = \sum_{j \in \mathbb{N}} y_j \psi_j \quad \text{i.i.d. } y_j \sim \mathcal{N}(0,1),\ \psi_j \in L^\infty(D)$$

- b with Hölder continuous realizations $\rightsquigarrow$ $u \in L^2(\mathbb{R}^{\mathbb{N}}, V; \gamma)$ with $\gamma = \bigotimes_{j=1}^{\infty} \mathcal{N}(0,1)$
- Product Hermite polynomials

$$H_\nu(y) = \prod_{j \geq 1} H_{\nu_j}(y_j) \quad \text{with univariate Hermite polynomials } H_{\nu_j}$$

  are orthonormal basis of $L^2(\mathbb{R}^{\mathbb{N}}; \gamma)$
- Product Hermite expansion of u,

$$u(y) = \sum_{\nu \in \mathcal{F}} u_\nu H_\nu(y) \approx \sum_{\nu \in \Lambda \subset \mathcal{F}} u_\nu H_\nu(y), \quad u_\nu = \int_{\mathbb{R}^{\mathbb{N}}} u(y)\, H_\nu(y)\, d\gamma(y)$$

Theorem (B., Cohen, DeVore, Migliorati '17). Let $0 < q < \infty$ and $0 < p < 2$ such that $\frac{1}{q} = \frac{1}{p} - \frac{1}{2}$. Assume there exists a positive sequence $\rho = (\rho_j)_{j\geq 1}$ such that

$$(\rho_j^{-1})_{j\geq 1} \in \ell^q(\mathbb{N}) \quad \text{und} \quad \sup_{x\in D} \sum_{j\geq 1} \rho_j |\psi_j(x)| < \infty.$$

Then $\big(\|u_\nu\|_V\big)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F})$.

- For $\{\psi_j\}$ with multilevel structure such that $\|\psi_j\|_{L^\infty} \lesssim 2^{-\alpha\ell(j)}$,

$$\Big\| u - \sum_{\nu\in\Lambda_n} u_\nu H_\nu \Big\|_{L^2(\mathbb{R}^{\mathbb{N}},V,\gamma)} \lesssim n^{-s} \quad \text{for any } s < \frac{\alpha}{d}$$

M. Bachmayr, A. Cohen, R. DeVore, and G. Migliorati, *Sparse polynomial approximation of parametric elliptic PDEs. Part II: lognormal coefficients*, ESAIM M2AN, 2017

Gaussian random fields

$D \subset \mathbb{R}^d$, centered Gaussian random field $\big(b(x)\big)_{x\in D}$ with covariance function

$$\mathbb{E}\big(b(x)b(x')\big) = K(x,x'), \quad x, x' \in D.$$

Given K, find $\{\psi_j\}$ such that $\qquad b(x) = \sum_{j=1}^{\infty} y_j \psi_j(x), \quad y_j \sim \mathcal{N}(0,1)$ i.i.d.

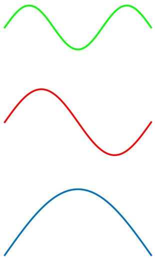
- Classical choice: Karhunen-Loève decomposition,

$$b(x) = \sum_{j=1}^{\infty} \sqrt{\lambda_j}\varphi_j(x)\, y_j \qquad \text{with } y_j \sim \mathcal{N}(0,1) \text{ i.i.d.}$$

with (λ_j, φ_j) eigenpairs of covariance operator, where φ_j is L^2-orthonormal

- Not the only option! Precise criterion (Luschgy, Pagès '09):

ψ_j provide an expansion with y_j i.i.d. precisely when ψ_j Parseval frame in reproducing kernel Hilbert space of K



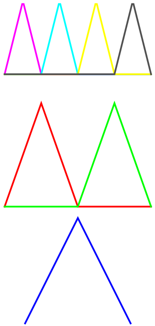
$K(s,t) = \min\{s,t\} - st$, with RKHS $H_0^1(0,1)$,
series $b = \sum_{j\geq 1} y_j \psi_j$ on $D = (0,1)$:

- **KL expansion:** $\psi_j(x) = \dfrac{\sqrt{2}}{\pi j} \sin(\pi j x)$,
  $\|\psi_j\|_{L^\infty} \sim j^{-1}$ with $|\operatorname{supp} \psi_j| = 1$.

- **Lévy-Ciesielski representation:**
  using Schauder basis (primitives of Haar system)

$$\psi_{\ell,m}(x) := 2^{-\ell/2} \psi(2^\ell x - m), \quad m = 0, \dots, 2^\ell - 1,\ \ell \geq$$

where $\psi(x) := \frac{1}{2}\big(1 - |2x-1|\big)_+$.
Ordering from coarse to fine, $\psi_j := \psi_{\ell,m}$ for $j = 2^\ell + m$,
$\|\psi_j\|_{L^\infty} \sim j^{-\frac{1}{2}}$ and $|\operatorname{supp} \psi_j| \sim j^{-1}$.



$D \subset \mathbb{R}^d$, centered and stationary Gaussian random field $\big(b(x)\big)_{x \in D}$ with covariance function

$$\mathbb{E}\big(b(x)b(x')\big) = K(x, x') = k(x - x'), \quad x, x' \in D.$$

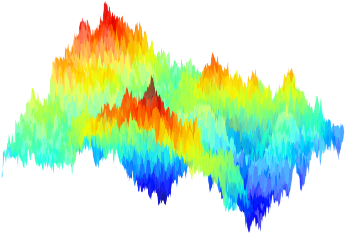
- **Matérn covariances**

$$k(x) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu}|x|}{\lambda}\right)^{\nu} K_\nu\left(\frac{\sqrt{2\nu}|x|}{\lambda}\right), \quad \nu, \lambda > 0,$$

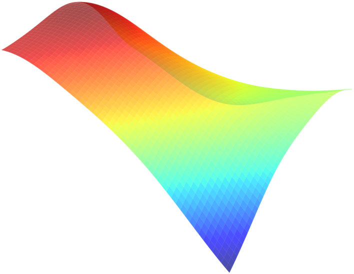
where K_ν is the modified Bessel function of the second kind, Fourier transform:

$$\hat{k}(\omega) = c_{\nu,\lambda} \left(\frac{2\nu}{\lambda^2} + |\omega|^2\right)^{-(\nu + d/2)}, \quad c_{\nu,\lambda} := \frac{2^d \pi^{d/2} \Gamma(\nu + d/2)(2\nu)^{\nu}}{\Gamma(\nu)\lambda^{2\nu}}.$$

(Exponential covariance $\nu = \frac{1}{2}$, Gaussian covariance $\nu \to \infty$)



$\nu = \frac{1}{2}$



$\nu = 4$

Construction of wavelet expansions

For class of stationary random fields including Matérn:

- Embed D into a torus $\mathbb{T}$, periodized random field with covariance k_{p},
- Apply square root of covariance operator to periodic L^2-orthonormal Meyer wavelets (in terms of Fourier exponentials), yields orthonormal basis of RKHS of k
- Difficult part: verify localization properties
- Restrict back to D to obtain Parseval frame of RKHS of k

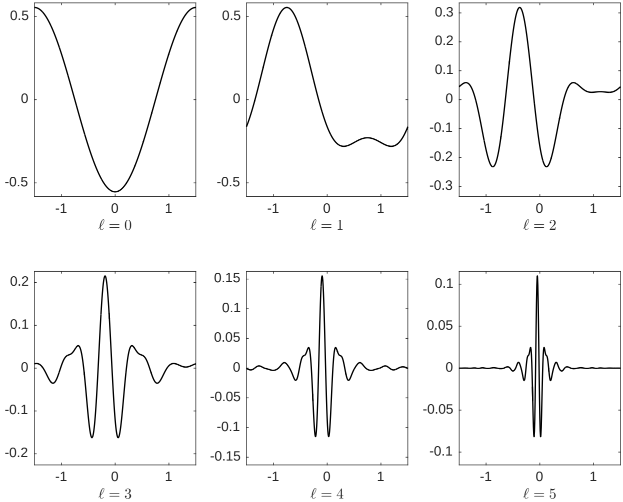
Conclusion: For $a = \exp(b)$, Matérn-type b with realizations in $C^{0,\beta}(\overline{D})$ for $\beta < \alpha$ in wavelet representation, where $\|\psi_j\|_{L^\infty} \lesssim 2^{-\alpha\ell(j)}$,

$$\Big\| u - \sum_{\nu\in\Lambda_n} u_\nu H_\nu \Big\|_{L^2(\mathbb{R}^{\mathbb{N}},V,\gamma)} \lesssim n^{-s} \quad \text{for any } s < \frac{\alpha}{d}$$

M. Bachmayr, A. Cohen, and G. Migliorati, *Representations of Gaussian random fields and approximation of elliptic PDEs with lognormal coefficients*, JFAA, 2018

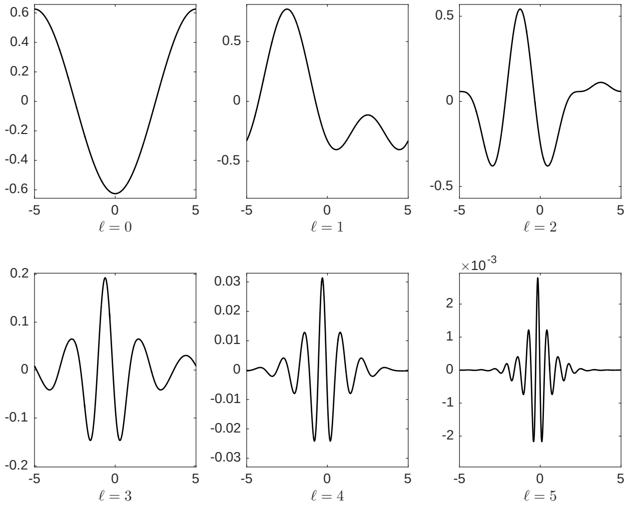
Matérn wavelets, 1D case on $D = [-\frac{1}{2}, \frac{1}{2}]$

Matérn covariance with $\lambda = 1$, $\nu = \frac{1}{2}$: plots of ψ_ℓ, where $\psi_{\ell,m}(x) = \psi_\ell(2^\ell x - m)$



Matérn wavelets, 1D case on $D = [-\frac{1}{2}, \frac{1}{2}]$

Matérn covariance with $\lambda = 1$, $\nu = 4$: plots of ψ_ℓ, where $\psi_{\ell,m}(x) = \psi_\ell(2^\ell x - m)$



Fully discrete approximability

Back to affine case: $a(y) = \bar{a} + \sum_{\ell,m} y_{\ell,m}\psi_{\ell,m}$ uniformly elliptic, $Y \simeq [-1,1]^{\mathbb{N}}$

Legendre expansion of $u \in \mathcal{V}$:

$$u(y) = \sum_{\nu\in\mathcal{F}} u_\nu L_\nu(y)$$

For each $\nu \in \mathcal{F}$, choose $V_\nu \subset V$ with $N_\nu := \dim V_\nu < \infty$,

$$\mathcal{V}_N = \Big\{ \sum_{\nu\in\mathcal{F}} v_\nu L_\nu : v_\nu \in V_\nu \Big\}, \qquad N = \sum_{\nu\in\mathcal{F}} N_\nu .$$

Approximations $u \approx u_N \in \mathcal{V}_N$: $F \subset \mathcal{F}$ and $u_\nu \approx \tilde{u}_\nu \in V_\nu$ for $\nu \in F$,

$$u(y) \approx u_N(y) = \sum_{\nu\in F} \tilde{u}_\nu L_\nu(y)$$

Adaptive approximations ($d \geq 2$) (B., Cohen, Dũng, Schwab '17)

Let $d \geq 2$ and $\alpha \in (0,1]$, let a be given in multilevel expansion with

$$\sup_D \sum_m |\psi_{\ell,m}| \lesssim 2^{-\alpha\ell}, \quad \sup_D \sum_m |\nabla\psi_{\ell,m}| \lesssim 2^{-(\alpha-1)\ell} \quad \text{for all } \ell \geq 0,$$

let D be convex or smooth and let $f \in L^2(D)$. Then for each N there exist $(V_\nu)_{\nu\in\mathcal{F}}$ such that for the corresponding $\mathcal{V}_N$,

$$\inf_{u_N \in \mathcal{V}_N} \|u - u_N\|_{L^2(Y,V,\mu)} \lesssim N^{-s} \quad \text{for any } s < \frac{\alpha}{d}.$$

Proof via summability

$$\sum_{\nu\in\mathcal{F}} \bigl(\rho^\nu \|\Delta t_\nu\|_{L^\tau(D)}\bigr)^\tau < \infty$$

with $\tau \in [1,2)$ and interpolation arguments.

M. Bachmayr, A. Cohen, D. Dũng, and Ch. Schwab, *Fully discrete approximation of parametric and stochastic elliptic PDEs*, SINUM, 2017

Space-parameter adaptivity

- How to choose $(V_\nu)_{\nu\in\mathcal{F}}$, total number of degrees of freedom $N = \sum_{\nu\in\mathcal{F}} N_\nu$?

- Adaptive wavelet approximation for each ν:

  $\{\Psi_\lambda\}_{\lambda\in\mathcal{S}}$ wavelet Riesz basis of $V = H_0^1(D)$,

$$\Big\|\sum_{\lambda,\nu} \mathbf{v}_{\lambda,\nu}\Psi_\lambda \otimes L_\nu\Big\|^2_{L^2(Y,V,\mu)} \eqsim \sum_{\lambda,\nu} |\mathbf{v}_{\lambda,\nu}|^2, \quad \mathbf{v} \in \ell^2(\mathcal{S}\times\mathcal{F})$$

$$\rightsquigarrow \text{ expansion } \quad u = \sum_{\lambda,\nu} \mathbf{u}_{\lambda,\nu}\Psi_\lambda \otimes L_\nu$$

- Best N-term approximation by keeping (λ,ν) with N largest $|\mathbf{u}_{\lambda,\nu}|$:

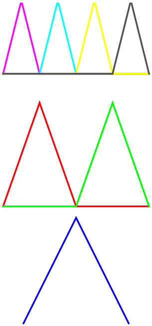
$$\|u - u_{[N]}\|_{L^2(Y,V,\mu)} \eqsim \|\mathbf{u} - \mathbf{u}_{[N]}\|_{\ell_2} \le N^{-s}\|\mathbf{u}\|_{\mathcal{A}^s} \qquad \rightsquigarrow \qquad N(\varepsilon) = \|\mathbf{u}\|_{\mathcal{A}^s}^{\frac{1}{s}} \varepsilon^{-\frac{1}{s}}$$

Example

Multiscale representation in $d = 1$, with $\alpha = 1$,

$$\psi_{\ell,m}(x) := c2^{-\ell}\psi(2^{\ell}x - m)$$

$$a(y) = 1 + \sum_{\ell,m} y_{\ell,m}\psi_{\ell,m} \quad \rightsquigarrow \quad u(y) = \sum_{\nu \in \mathcal{F}} u_\nu L_\nu(y)$$

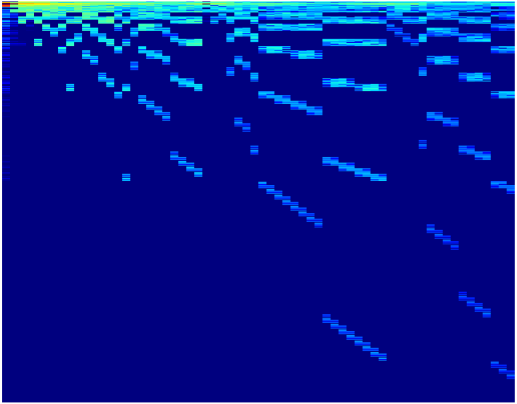


$d = 1$: $\quad a(y) = \bar{a} + \sum_{j \geq 1} y_j \, \psi_j$, ψ_j hierarchical hat functions, $\|\psi_j\|_{L^\infty} \lesssim 2^{-\alpha \ell(j)}$

Values $|\mathbf{u}_{\mu,\nu}|$ (for $\alpha = 1$):

$\xrightarrow{\nu}$ (decreasing $\|u_\nu\|_V$)

$\mu \downarrow$



Stochastic Galerkin discretization: $u_N \in \mathcal{V}_N$ such that

$$\int_Y \int_D a \nabla u_N \cdot \nabla v \, dx \, d\mu(y) = \int_Y \langle f, v \rangle \, d\mu(y), \quad \text{for all } v \in \mathcal{V}_N$$

Operator representation w.r.t. spatial-parametric Riesz basis $\{\Psi_\lambda \otimes L_\nu\}_{\lambda \in \mathcal{S}, \nu \in \mathcal{F}}$,

$$\mathbf{A} = \sum_{j \geq 0} \mathbf{A}_j \otimes \mathbf{M}_j : \ell^2(\mathcal{S} \times \mathcal{F}) \to \ell^2(\mathcal{S} \times \mathcal{F})$$

where

$$\mathbf{A}_0 = \left(\int_D \bar{a} \nabla \Psi_{\lambda'} \cdot \nabla \Psi_\lambda \right)_{\lambda, \lambda' \in \mathcal{S}}, \quad \mathbf{M}_0 = \left(\delta_{\nu, \nu'} \right)_{\nu, \nu' \in \mathcal{F}}$$

$$\mathbf{A}_j = \left(\int_D \psi_j \nabla \Psi_{\lambda'} \cdot \nabla \Psi_\lambda \right)_{\lambda, \lambda' \in \mathcal{S}}, \quad \mathbf{M}_j = \left(\int_Y y_j L_\nu(y) L_{\nu'}(y) \, d\mu(y) \right)_{\nu, \nu' \in \mathcal{F}}, \quad j \geq 1.$$

$\rightsquigarrow$ well-conditioned sequence-space formulation $\mathbf{A}\mathbf{u} = \mathbf{f}$.

Standard adaptive Galerkin scheme
(Cohen, Dahmen, DeVore '01; Gantumur, Harbrecht, Stevenson '07)

Given $\Lambda^k \subset \mathcal{S} \times \mathcal{F}$, compute Galerkin solution $\mathbf{u}_k$ on Λ^k, approximate $\mathbf{r}_k = \mathbf{A}\mathbf{u}_k - \mathbf{f}$, and with fixed $\mu \in (0, 1)$ set

$$\Lambda^{k+1} = \Lambda^k \cup \hat{\Lambda} \quad \text{with } \hat{\Lambda} \text{ of minimal size such that } \|\mathbf{r}|_{\hat{\Lambda}}\|_{\ell^2} \geq \mu \|\mathbf{r}\|_{\ell^2}$$

Direct residual approximation

- Residual approximation for stochastic Galerkin systems can be done based on standard compression techniques for $\mathbf{A}$ (using s^*-compressibility)
- For ψ_j with global supports, rates generally not optimal (Gittelson '13, '14)
- Observation[2] for $\{\psi_j\}$ with multilevel structure such that $\|\psi_j\|_{L^\infty} \lesssim 2^{-\alpha\ell(j)}$ (ordered by level): $\mathbf{A} = \sum_{j\geq 0} \mathbf{A}_j \otimes \mathbf{M}_j$ satisfies

$$\Big\| \sum_{j>M} \mathbf{A}_j \otimes \mathbf{M}_j \Big\| \lesssim M^{-\frac{\alpha}{d}}.$$

- Compression based on approximations $\sum_{j\leq M} \mathbf{A}_j \otimes \mathbf{M}_j$ combined with spatial s^*-compressibility of the $\mathbf{A}_j$: sub-optimal rates

$$s^* = \frac{t}{t+d}\frac{\alpha}{d}$$

when $\psi_j \nabla \Psi_\lambda \in H^t$.

[2] M. Bachmayr, A. Cohen, and W. Dahmen, *Parametric PDEs: Sparse or low-rank approximations?*, IMA JNA, 2018

- Iteratively refined stochastic Galerkin discretizations with spatial approximation by H^2-regular spline wavelets, piecewise polynomial (approximations of) ψ_j

New residual approximation strategy:

- Adaptive semidiscrete operator compression in parametric variables, based on $\sum_{j\le M} \mathbf{A}_j \otimes \mathbf{M}_j$,
- Spatial error estimation using tree index sets and piecewise polynomial structure without adaptive operator compression (Stevenson '14; Binev '18)

Optimality (B., Voulis '22)

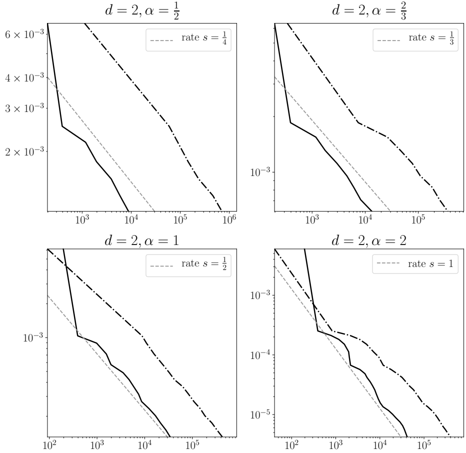
If the best approximation to u converges at rate $s < \frac{\alpha}{d}$ then for each $\varepsilon > 0$, the adaptive scheme with appropriately chosen parameters finds an approximation u_ε with $\|u - u_\varepsilon\|_{\mathcal{V}} \le \varepsilon$ using $\mathcal{O}\big(1 + \varepsilon^{-\frac{1}{s}}(1 + |\log \varepsilon|)\big)$ operations.

(see also Bespalov, Praetorius, Ruggeri '21: optimal cardinality under saturation assumption)

M. Bachmayr and I. Voulis, *An adaptive stochastic Galerkin method based on multilevel expansions of random fields: Convergence and optimality*, ESAIM M2AN, 2022

Numerical experiments: wavelets, $d = 2$ (B., Voulis '22)

$D = (0,1)^2$, $\psi_{\ell,m}$ hierarchical piecewise linear hat functions with $\|\psi_{\ell,m}\|_{L_\infty} \lesssim 2^{-\alpha\ell}$, spatial discretization by C^1 piecewise polynomial DGH multiwavelets of order 6; expected fully discrete rate $\frac{\alpha}{2}$.



Residual estimates as a function of #dof (—) and of computation time (- -)

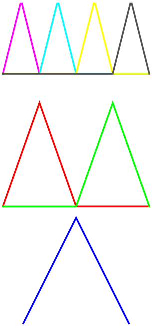
Finite element approximations in space?

Aim: $u(y) = \sum_{\nu \in \Lambda} u_\nu L_\nu(y)$ with $u_\nu \in \mathbb{P}_1(\mathcal{T}_\nu) \cap V$, separate mesh $\mathcal{T}_\nu$ for each ν

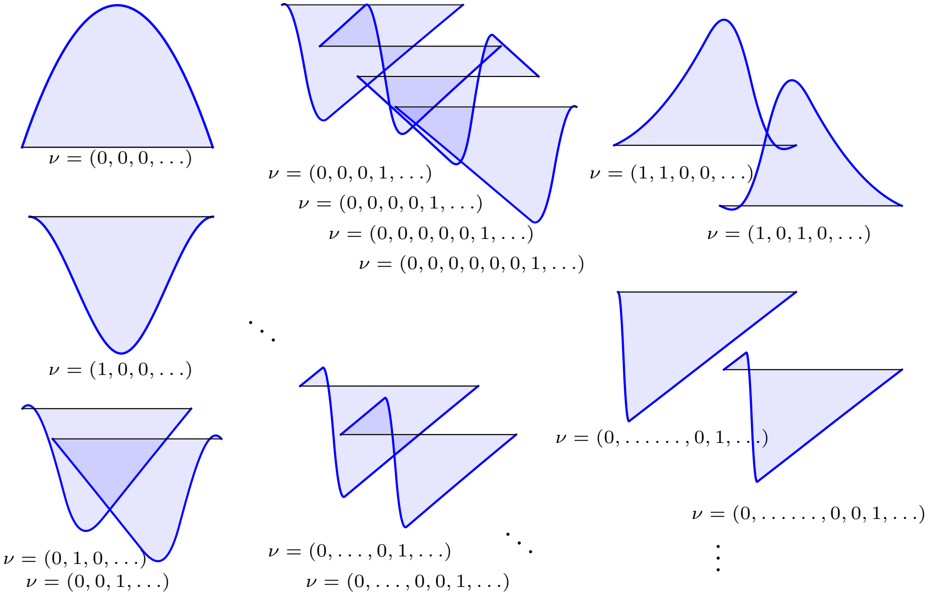
Same example with $d = 1$, $\alpha = 1$,

$$\psi_{\ell,m}(x) := c2^{-\ell}\psi(2^\ell x - m)$$

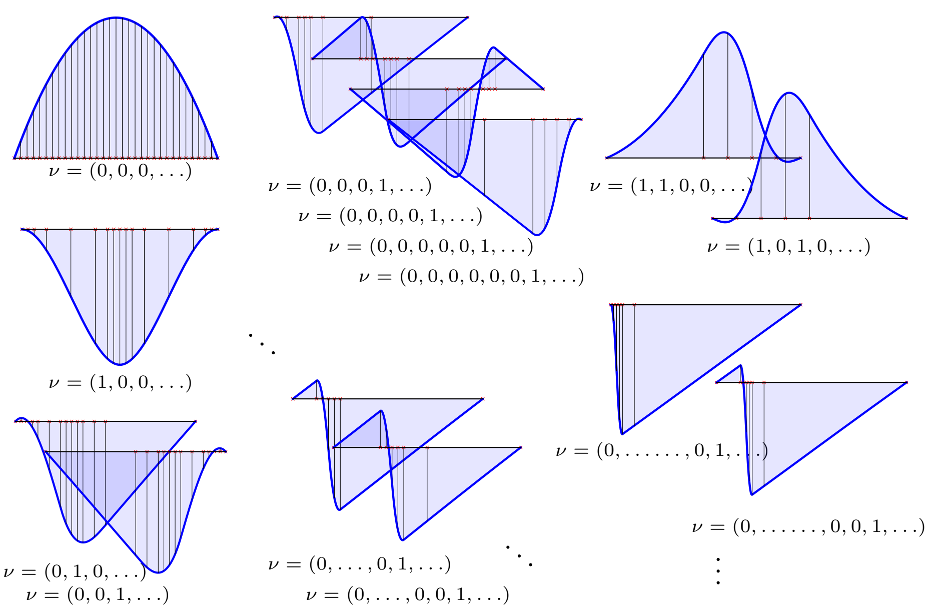
$$a(y) = 1 + \sum_{\ell,m} y_{\ell,m}\psi_{\ell,m} \quad \rightsquigarrow \quad u(y) = \sum_{\nu \in \mathcal{F}} u_\nu L_\nu(y)$$



Legendre coefficient functions $u_\nu : [0,1] \to \mathbb{R}$ in $u(y) = \sum_{\nu \in \mathcal{F}} u_\nu L_\nu(y)$ with diffusion coefficient a expanded in terms of hierarchical hat functions:



Best (dyadic) grids for piecewise linear approximations of u_ν:



Towards an optimal adaptive finite element solver

- Piecewise affine linear finite element approximation on independent adaptive mesh for each u_ν, refinement by standard newest vertex bisection
- Again using adaptive operator compression in the stochastic variables.
- Standard finite element error estimation strategies (e.g., residual estimators) not applicable due to interactions between meshes, lack of Galerkin orthogonality (see also Cohen, DeVore, Nochetto '12)
- Instead use BPX frame coefficients (cf. Harbrecht, Schneider '16): for $r \in V' = H^{-1}(D)$,

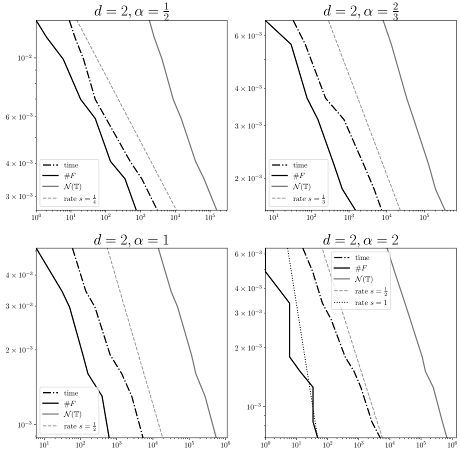
$$\|r\|_{V'}^2 \eqsim \sum_{j=0}^{\infty} \sum_{k \in \mathcal{N}_j} |\langle r, \varphi_{j,k} \rangle|^2$$

  with $\varphi_{j,k}$ piecewise linear hat function on level j (with $\|\varphi_{j,k}\|_{H_0^1(D)} \eqsim 1$)
- Choose refinements by tree-based selection of frame-based indicators (Binev '18)

First result[3]: reduction of stochastic Galerkin energy norm error by uniform factor in each step of the adaptive scheme, linear convergence to exact solution.

[3] M. Bachmayr, M. Eigel, H. Eisenmann and I. Voulis, *A convergent adaptive finite element stochastic Galerkin method based on multilevel expansions of random fields*, SINUM, 2025.

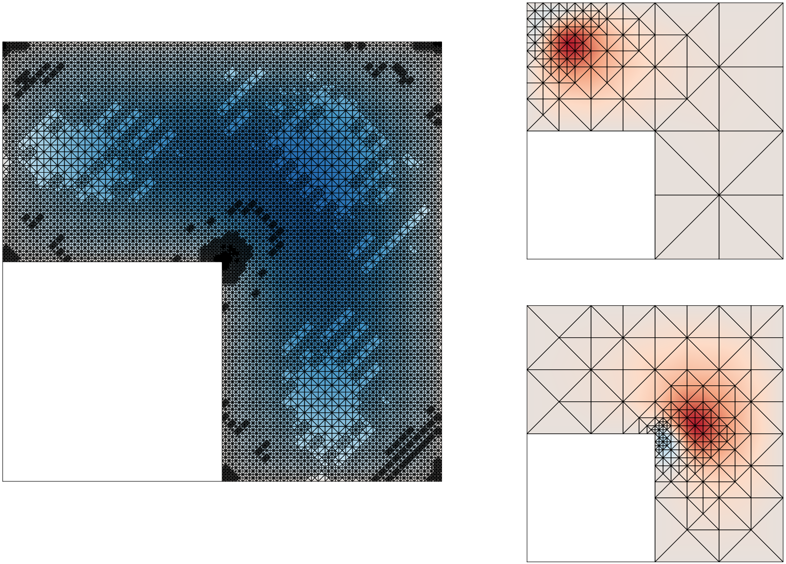
L-shaped domain, multilevel hat functions $\psi_{\ell,m}$ with $\|\psi_{\ell,m}\|_{L_\infty} \lesssim 2^{-\alpha\ell}$, spatial discretization by $\mathbb{P}_1$ elements on newest vertex bisection meshes.



Residual estimates as a function of #dof (parametric —, all —) and of computation time (- ·)

Finite elements, $d = 2$

(B., Eigel, Eisenmann, Voulis '25)



Optimal complexity

First consider optimality of generated discretizations assuming

- affine coefficients $a(y) = \bar{a} + \sum_{j \geq 1} y_j \psi_j$,
- best approximations of u in $\mathcal{V}_N$ converging as $\mathcal{O}(N^{-s})$ with $s < \alpha/d$.

Theorem (B., Eisenmann, Voulis '25; abridged).

- The meshes generated by the method have optimal cardinality:

$$\|u - u_N\|_{L^2(Y,V,\mu)} \leq \varepsilon \quad \text{with} \quad N \lesssim \varepsilon^{-1/s}.$$

- If $\{\psi_j\}$ have multilevel structure, near-optimal total number of operations

$$\mathcal{O}\big(\varepsilon^{-1/s}(1 + |\log \varepsilon|^3)\big) \qquad \text{for all } s < \alpha/d.$$

- Main new ingredient: stability property of finite element frames on adaptively refined (newest vertex bisection) meshes

M. Bachmayr, H. Eisenmann and I. Voulis, *Adaptive stochastic Galerkin finite element methods: Optimality and non-affine coefficients*, arXiv:2503.18704

Extension to non-affine coefficients

- Uniformly elliptic coefficients of the form (e.g., log-uniform case $g = \exp$)

$$a(y) = g\Big(\sum_{j\geq 1} y_j\theta_j\Big) \quad \text{with i.i.d. } y_j \sim \mathcal{U}(-1,1),$$

- Requires new semi-discrete operator compression
- Basic strategy: for g analytic in sufficiently large rectangle in $\mathbb{C}$, use polynomial approximations of g.

Theorem (B., Eisenmann, Voulis '25; abridged).

Assuming $\{\psi_j\}$ with multilevel structure as before and best approximation rate $s < \alpha/d$, then

$$\|u - u_N\|_{L^2(Y,V,\mu)} \leq \varepsilon \quad \text{with} \quad N \lesssim \varepsilon^{-1/s}$$

using a number of operations of order

$$\mathcal{O}\big(\varepsilon^{-1/s'}(1 + |\log \varepsilon|)^r\big) \qquad \text{for all } s' < s < \alpha/d$$

with $r > 0$ independent of s', s, k.

M. Bachmayr, H. Eisenmann and I. Voulis, *Adaptive stochastic Galerkin finite element methods: Optimality and non-affine coefficients*, arXiv:2503.18704

Summary and outlook

- Representations of parameterized (or random) coefficients in terms of localized functions lead to improved approximability of PDE solutions
- Taking full advantage of this requires a separately adapted mesh for each term in the Legendre expansion of u
- Stochastic Galerkin methods with optimal convergence (rates up to α/d) and computational costs optimal up to log-factors

---

- M. Bachmayr and I. Voulis, *An adaptive stochastic Galerkin method based on multilevel expansions of random fields: Convergence and optimality*, ESAIM M2AN, 2022.
- M. Bachmayr, M. Eigel, H. Eisenmann and I. Voulis, *A convergent adaptive finite element stochastic Galerkin method based on multilevel expansions of random fields*, SINUM, 2025.
- M. Bachmayr, H. Eisenmann and I. Voulis, *Adaptive stochastic Galerkin finite element methods: optimality and non-affine coefficients*, arXiv:2503.18704.
- M. Bachmayr and H. Yang, *Sparse and low-rank approximations of parametric elliptic PDEs: the best of both worlds*, arXiv:2506.19584.