High dimensional approximation of parametric PDE's

Albert Cohen

Laboratoire Jacques-Louis Lions
Université Pierre et Marie Curie
Paris

EPFL, 7-04-2016

Overview

1. Introduction to the main themes : high dimensional parametric PDEs
2. Sparse polynomial approximation for a model example
3. Other models
4. Towards sparse polynomial algorithms

References

R. DeVore, "Nonlinear approximation", Acta Numerica, 1998.

A. Cohen, R. DeVore and C. Schwab, "Analytic regularity and polynomial approximation of parametric and stochastic PDEs", Analysis and Application, 2011.

A. Cohen and R. DeVore, "High dimensional approximation of parametric PDEs", Acta Numerica, 2015.

Parametric/Stochastic PDEs

We are interested in PDE's of the general form

$$\mathcal{D}(u, y) = 0,$$

where $\mathcal{D}$ is a partial differential operator, u is the unknown and $y = (y_j)_{j=1,\dots,d}$ is a parameter vector of dimension $d >> 1$ or $d = \infty$ ranging in some domain U.

We assume well-posedness of the solution in some Banach space V for every $y \in U$,

$$y \mapsto u(y)$$

is the solution map from U to V.

Solution manifold $\mathcal{M} := \{u(y) \; : \; y \in U\} \subset V$.

The parameters may be deterministic (control, optimization, inverse problems) or random (uncertainty modeling and quantification, risk assessment). In the second case the solution $u(y)$ is a V-valued random variable.

These applications often requires many queries of $u(y)$, and therefore in principle running many times a numerical solver.

Objective : economical numerical approximation of the map $y \mapsto u(y)$.

Related objectives : numerical approximation of scalar quantities of interest $y \mapsto Q(y) = Q(u(y))$, or of averaged quantities $\overline{u} = \mathbb{E}(u(y))$ or $\overline{Q} = \mathbb{E}(Q(y))$.

We are interested in PDE's of the general form

$$\mathcal{D}(u, y) = 0,$$

where $\mathcal{D}$ is a partial differential operator, u is the unknown and $y = (y_j)_{j=1,\dots,d}$ is a parameter vector of dimension $d >> 1$ or $d = \infty$ ranging in some domain U.

We assume well-posedness of the solution in some Banach space V for every $y \in U$,

$$y \mapsto u(y)$$

is the solution map from U to V.

Solution manifold $\mathcal{M} := \{u(y) \ : \ y \in U\} \subset V$.

The parameters may be deterministic (control, optimization, inverse problems) or random (uncertainty modeling and quantification, risk assessment). In the second case the solution $u(y)$ is a V-valued random variable.

These applications often requires many queries of $u(y)$, and therefore in principle running many times a numerical solver.

Objective : economical numerical approximation of the map $y \mapsto u(y)$.

Related objectives : numerical approximation of scalar quantities of interest $y \mapsto Q(y) = Q(u(y))$, or of averaged quantities $\overline{u} = \mathbb{E}(u(y))$ or $\overline{Q} = \mathbb{E}(Q(y))$.

We are interested in PDE's of the general form

$$\mathcal{D}(u, y) = 0,$$

where $\mathcal{D}$ is a partial differential operator, u is the unknown and $y = (y_j)_{j=1,\dots,d}$ is a parameter vector of dimension $d >> 1$ or $d = \infty$ ranging in some domain U.

We assume well-posedness of the solution in some Banach space V for every $y \in U$,

$$y \mapsto u(y)$$

is the solution map from U to V.

Solution manifold $\mathcal{M} := \{u(y) \ : \ y \in U\} \subset V$.

The parameters may be deterministic (control, optimization, inverse problems) or random (uncertainty modeling and quantification, risk assessment). In the second case the solution $u(y)$ is a V-valued random variable.

These applications often requires many queries of $u(y)$, and therefore in principle running many times a numerical solver.

Objective : economical numerical approximation of the map $y \mapsto u(y)$.

Related objectives : numerical approximation of scalar quantities of interest $y \mapsto Q(y) = Q(u(y))$, or of averaged quantities $\overline{u} = \mathbb{E}(u(y))$ or $\overline{Q} = \mathbb{E}(Q(y))$.

Parametric/Stochastic PDEs

We are interested in PDE's of the general form

$$\mathcal{D}(u, y) = 0,$$

where $\mathcal{D}$ is a partial differential operator, u is the unknown and $y = (y_j)_{j=1,\dots,d}$ is a parameter vector of dimension $d >> 1$ or $d = \infty$ ranging in some domain U.

We assume well-posedness of the solution in some Banach space V for every $y \in U$,

$$y \mapsto u(y)$$

is the solution map from U to V.

Solution manifold $\mathcal{M} := \{u(y) \ : \ y \in U\} \subset V$.

The parameters may be deterministic (control, optimization, inverse problems) or random (uncertainty modeling and quantification, risk assessment). In the second case the solution $u(y)$ is a V-valued random variable.

These applications often requires many queries of $u(y)$, and therefore in principle running many times a numerical solver.

Objective : economical numerical approximation of the map $y \mapsto u(y)$.

Related objectives : numerical approximation of scalar quantities of interest $y \mapsto Q(y) = Q(u(y))$, or of averaged quantities $\overline{u} = \mathbb{E}(u(y))$ or $\overline{Q} = \mathbb{E}(Q(y))$.

We are interested in PDE's of the general form

$$\mathcal{D}(u, y) = 0,$$

where $\mathcal{D}$ is a partial differential operator, u is the unknown and $y = (y_j)_{j=1,\dots,d}$ is a parameter vector of dimension $d >> 1$ or $d = \infty$ ranging in some domain U.

We assume well-posedness of the solution in some Banach space V for every $y \in U$,

$$y \mapsto u(y)$$

is the solution map from U to V.

Solution manifold $\mathcal{M} := \{u(y) \ : \ y \in U\} \subset V$.

The parameters may be deterministic (control, optimization, inverse problems) or random (uncertainty modeling and quantification, risk assessment). In the second case the solution $u(y)$ is a V-valued random variable.

These applications often requires many queries of $u(y)$, and therefore in principle running many times a numerical solver.

Objective : economical numerical approximation of the map $y \mapsto u(y)$.

Related objectives : numerical approximation of scalar quantities of interest $y \mapsto Q(y) = Q(u(y))$, or of averaged quantities $\overline{u} = \mathbb{E}(u(y))$ or $\overline{Q} = \mathbb{E}(Q(y))$.

Guiding example : elliptic PDEs

We consider the steady state diffusion equation

$$-\mathrm{div}(a\nabla u) = f \quad \text{on} \quad D \subset \mathbf{R}^m \quad \text{and} \quad u_{|\partial D} = 0,$$

set on a domain $D \subset \mathbb{R}^m$, where $f = f(x) \in L^2(D)$ and $a \in L^\infty(D)$

Lax-Milgram lemma : assuming $a_{\min} := \min_{x \in D} a(x) > 0$, unique solution $u \in V = H_0^1(D)$ with

$$\|u\|_V := \|\nabla u\|_{L^2(D)} \leq \frac{1}{a_{\min}} \|f\|_{V'}.$$

Proof of the estimate : multiply equation by u and integrate

$$a_{\min}\|u\|_V^2 \leq \int_D a\nabla u \cdot \nabla u = -\int_D u\, \mathrm{div}(a\nabla u) = \int_D uf \leq \|u\|_V \|f\|_{V'}.$$

We may extend this theory to the solution of the weak (or variational) formulation

$$\int_D a\nabla u \cdot \nabla v = \langle f, v \rangle, \quad v \in V = H_0^1(D),$$

if $f \in V' = H^{-1}(D)$

Guiding example : elliptic PDEs

We consider the steady state diffusion equation

$$-\mathrm{div}(a\nabla u) = f \quad \text{on} \quad D \subset \mathbf{R}^m \quad \text{and} \quad u_{|\partial D} = 0,$$

set on a domain $D \subset \mathbb{R}^m$, where $f = f(x) \in L^2(D)$ and $a \in L^\infty(D)$

Lax-Milgram lemma : assuming $a_{\min} := \min_{x\in D} a(x) > 0$, unique solution $u \in V = H_0^1(D)$ with

$$\|u\|_V := \|\nabla u\|_{L^2(D)} \le \frac{1}{a_{\min}} \|f\|_{V'}.$$

Proof of the estimate : multiply equation by u and integrate

$$a_{\min}\|u\|_V^2 \le \int_D a\nabla u \cdot \nabla u = -\int_D u\,\mathrm{div}(a\nabla u) = \int_D uf \le \|u\|_V \|f\|_{V'}.$$

We may extend this theory to the solution of the weak (or variational) formulation

$$\int_D a\nabla u \cdot \nabla v = \langle f, v\rangle, \quad v \in V = H_0^1(D),$$

if $f \in V' = H^{-1}(D)$

We consider the steady state diffusion equation

$$-\mathrm{div}(a\nabla u) = f \quad \text{on} \quad D \subset \mathbf{R}^m \quad \text{and} \quad u_{|\partial D} = 0,$$

set on a domain $D \subset \mathbb{R}^m$, where $f = f(x) \in L^2(D)$ and $a \in L^\infty(D)$

Lax-Milgram lemma : assuming $a_{\min} := \min_{x\in D} a(x) > 0$, unique solution $u \in V = H_0^1(D)$ with

$$\|u\|_V := \|\nabla u\|_{L^2(D)} \leq \frac{1}{a_{\min}} \|f\|_{V'}.$$

Proof of the estimate : multiply equation by u and integrate

$$a_{\min}\|u\|_V^2 \leq \int_D a\nabla u \cdot \nabla u = -\int_D u\,\mathrm{div}(a\nabla u) = \int_D uf \leq \|u\|_V \|f\|_{V'}.$$

We may extend this theory to the solution of the weak (or variational) formulation

$$\int_D a\nabla u \cdot \nabla v = \langle f, v\rangle, \quad v \in V = H_0^1(D),$$

if $f \in V' = H^{-1}(D)$

Assume diffusion coefficients in the form of an expansion

$$a = a(y) = \bar{a} + \sum_{j \geq 1} y_j \psi_j, \quad y = (y_j)_{j \geq 1} \in U,$$

with $d >> 1$ or $d = \infty$ terms, where $\bar{a}$ and $(\psi_j)_{j \geq 1}$ are functions from L^∞,

Note that $a(y)$ is a function for each given y. We may also write

$$a = a(x, y) = \bar{a}(x) + \sum_{j \geq 1} y_j \psi_j(x), \quad x \in D, y \in U,$$

where x and y are the spatial and parametric variable, respectively. Likewise, the corresponding solution $u(y)$ is a function $x \mapsto u(y, x)$ for each given y. We often ommit the reference to the spatial variable.

Up to a change of variable, we assume that all y_j range in $[-1, 1]$, therefore

$$y \in U = [-1, 1]^d \text{ or } [-1, 1]^{\mathbb{N}}.$$

Uniform ellipticity assumption :

$$(UEA) \qquad 0 < r \leq a(x, y) \leq R, \;\; x \in D, \; y \in U$$

Then the solution map is bounded from U to $V := H_0^1(D)$, that is, $u \in L^\infty(U, V)$:

$$\|u(y)\|_V \leq C_r := \frac{\|f\|_{V'}}{r}, \;\; y \in U,$$

Assume diffusion coefficients in the form of an expansion

$$a = a(y) = \bar{a} + \sum_{j \geq 1} y_j \psi_j, \quad y = (y_j)_{j \geq 1} \in U,$$

with $d >> 1$ or $d = \infty$ terms, where $\bar{a}$ and $(\psi_j)_{j \geq 1}$ are functions from L^∞,

Note that $a(y)$ is a function for each given y. We may also write

$$a = a(x, y) = \bar{a}(x) + \sum_{j \geq 1} y_j \psi_j(x), \quad x \in D, y \in U,$$

where x and y are the spatial and parametric variable, respectively. Likewise, the corresponding solution $u(y)$ is a function $x \mapsto u(y, x)$ for each given y. We often ommit the reference to the spatial variable.

Up to a change of variable, we assume that all y_j range in $[-1, 1]$, therefore

$$y \in U = [-1, 1]^d \text{ or } [-1, 1]^{\mathbb{N}}.$$

Uniform ellipticity assumption :

$$(UEA) \qquad 0 < r \leq a(x, y) \leq R, \quad x \in D, \ y \in U$$

Then the solution map is bounded from U to $V := H_0^1(D)$, that is, $u \in L^\infty(U, V)$:

$$\|u(y)\|_V \leq C_r := \frac{\|f\|_{V'}}{r}, \quad y \in U,$$

Assume diffusion coefficients in the form of an expansion

$$a = a(y) = \bar{a} + \sum_{j\geq 1} y_j \psi_j, \quad y = (y_j)_{j\geq 1} \in U,$$

with $d >> 1$ or $d = \infty$ terms, where $\bar{a}$ and $(\psi_j)_{j\geq 1}$ are functions from L^∞,

Note that $a(y)$ is a function for each given y. We may also write

$$a = a(x, y) = \bar{a}(x) + \sum_{j\geq 1} y_j \psi_j(x), \quad x \in D, y \in U,$$

where x and y are the spatial and parametric variable, respectively. Likewise, the corresponding solution $u(y)$ is a function $x \mapsto u(y, x)$ for each given y. We often ommit the reference to the spatial variable.

Up to a change of variable, we assume that all y_j range in $[-1, 1]$, therefore

$$y \in U = [-1, 1]^d \text{ or } [-1, 1]^{\mathbb{N}}.$$

Uniform ellipticity assumption :

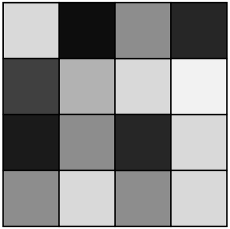
$$(UEA) \qquad 0 < r \leq a(x, y) \leq R, \;\; x \in D, \; y \in U$$

Then the solution map is bounded from U to $V := H_0^1(D)$, that is, $u \in L^\infty(U, V)$:

$$\|u(y)\|_V \leq C_r := \frac{\|f\|_{V'}}{r}, \;\; y \in U,$$

Example of parametrization : piecewise constant coefficients

Assume that a is piecewise constant over a partition $\{D_1, \dots, D_d\}$ of D, and such that on each D_j the value of a varies on $[c - c_j, c + c_j]$ for some $c > 0$ and $0 < c_j < c$.



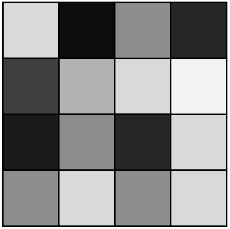
Then a natural parametrization is

$$a(y) = \overline{a} + \sum_{j=1}^{d} y_j \psi_j, \quad \overline{a} = c, \quad \psi_j = c_j \chi_{D_j},$$

with $y = (y_j)_{j=1,\dots,d} \in U = [-1, 1]^d$.

Example of parametrization : piecewise constant coefficients

Assume that a is piecewise constant over a partition $\{D_1, \dots, D_d\}$ of D, and such that on each D_j the value of a varies on $[c - c_j, c + c_j]$ for some $c > 0$ and $0 < c_j < c$.



Then a natural parametrization is

$$a(y) = \overline{a} + \sum_{j=1}^{d} y_j \psi_j, \quad \overline{a} = c, \quad \psi_j = c_j \chi_{D_j},$$

with $y = (y_j)_{j=1,\dots,d} \in U = [-1, 1]^d$.

Example of parametrization : Karhunen-Loeve representation

Assume $a = (a(x))_{x \in D}$ is a random process with average

$$\overline{a}(x) = \mathbb{E}(a(x)),$$

and covariance function

$$C_a(x,z) = \mathbb{E}\Big(\tilde{a}(x)\tilde{a}(z)\Big), \quad \tilde{a} := a - \overline{a}, \quad x, z \in D.$$

Define the integral operator by

$$Tv(x) = \int_D C_a(x,z)v(z)dz,$$

self-adjoint, positive and compact in $L^2(D)$. Therefore it admits an L^2 orthonormal basis $(\varphi_j)_{j \geq 1}$ of eigenfunctions, associated to eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$, such that $\lambda_n \to 0$ as $n \to +\infty$.

Karhunen-Loeve (KL) decomposition (a.k.a. principal component analysis) :

$$a = \overline{a} + \sum_{j \geq 1} \xi_j \varphi_j, \quad \xi_j := \int_D a(x)\varphi_j(x)dx.$$

Example of parametrization : Karhunen-Loeve representation

Assume $a = (a(x))_{x\in D}$ is a random process with average

$$\overline{a}(x) = \mathbb{E}(a(x)),$$

and covariance function

$$C_a(x,z) = \mathbb{E}\Big(\tilde{a}(x)\tilde{a}(z)\Big), \quad \tilde{a} := a - \overline{a}, \quad x,z \in D.$$

Define the integral operator by

$$Tv(x) = \int_D C_a(x,z)v(z)dz,$$

self-adjoint, positive and compact in $L^2(D)$. Therefore it admits an L^2 orthonormal basis $(\varphi_j)_{j\geq 1}$ of eigenfunctions, associated to eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$, such that $\lambda_n \to 0$ as $n \to +\infty$.

Karhunen-Loeve (KL) decomposition (a.k.a. principal component analysis) :

$$a = \overline{a} + \sum_{j\geq 1} \xi_j \varphi_j, \quad \xi_j := \int_D a(x)\varphi_j(x)dx.$$

Example of parametrization : Karhunen-Loeve representation

Assume $a = (a(x))_{x \in D}$ is a random process with average

$$\overline{a}(x) = \mathbb{E}(a(x)),$$

and covariance function

$$C_a(x,z) = \mathbb{E}\Big(\tilde{a}(x)\tilde{a}(z)\Big), \quad \tilde{a} := a - \overline{a}, \quad x, z \in D.$$

Define the integral operator by

$$Tv(x) = \int_D C_a(x,z)v(z)dz,$$

self-adjoint, positive and compact in $L^2(D)$. Therefore it admits an L^2 orthonormal basis $(\varphi_j)_{j\geq 1}$ of eigenfunctions, associated to eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$, such that $\lambda_n \longrightarrow 0$ as $n \longrightarrow +\infty$.

Karhunen-Loeve (KL) decomposition (a.k.a. principal component analysis) :

$$a = \overline{a} + \sum_{j\geq 1} \xi_j \varphi_j, \quad \xi_j := \int_D a(x)\varphi_j(x)dx.$$

The ξ_j are centered and decorelated scalar random variables, with

$$\mathbb{E}(\xi_j) = 0, \quad \mathbb{E}(\xi_i \xi_j) = 0 \quad \text{if} \quad j \neq i, \quad \mathbb{E}(|\xi_j|^2) = \lambda_j.$$

If the random process a is bounded, then the variables ξ_j have bounded range $|\xi_j| \leq c_j$, so that with $y_j := \xi_j / c_j$ and $\psi_j := c_j \varphi_j$ we may also write

$$a = \bar{a} + \sum_{j \geq 1} y_j \psi_j, \quad y = (y_j)_{j \geq 1} \in U = [-1, 1]^{\mathbb{N}}.$$

The KL representation is optimal for trunctation in mean-square $L^2(D)$-error :

$$\inf_{\dim(E)=J} \mathbb{E}(\|\tilde{a} - P_E \tilde{a}\|_{L^2}^2),$$

is attained by $E = E_J := \mathrm{span}\{\psi_1, \dots, \psi_J\}$ with

$$\mathbb{E}(\|\tilde{a} - P_{E_J} \tilde{a}\|_{L^2}^2) = \mathbb{E}\Big(\|\sum_{j > J} y_j \psi_j\|_{L^2}^2\Big) = \sum_{j > J} \lambda_j.$$

Case of a stationary process : $C_a(x, z) = \kappa(x - z)$, that is T is a convolution operator. If D is the m-dimensional 2π-periodic torus, the KL basis is of Fourier type

$$x \mapsto \varphi_k(x) := (2\pi)^{-m/2} e^{ik\dot{x}}, \quad k \in \mathbb{Z}^m.$$

The ξ_j are centered and decorelated scalar random variables, with

$$\mathbb{E}(\xi_j) = 0, \quad \mathbb{E}(\xi_i\xi_j) = 0 \quad \text{if} \quad j \neq i, \quad \mathbb{E}(|\xi_j|^2) = \lambda_j.$$

If the random process a is bounded, then the variables ξ_j have bounded range $|\xi_j| \leq c_j$, so that with $y_j := \xi_j/c_j$ and $\psi_j := c_j\varphi_j$ we may also write

$$a = \bar{a} + \sum_{j\geq 1} y_j\psi_j, \quad y = (y_j)_{j\geq 1} \in U = [-1,1]^{\mathbb{N}}.$$

The KL representation is optimal for trunctation in mean-square $L^2(D)$-error :

$$\inf_{\dim(E)=J} \mathbb{E}(\|\tilde{a} - P_E\tilde{a}\|_{L^2}^2),$$

is attained by $E = E_J := \mathrm{span}\{\psi_1, \dots, \psi_J\}$ with

$$\mathbb{E}(\|\tilde{a} - P_{E_J}\tilde{a}\|_{L^2}^2) = \mathbb{E}\Big(\|\sum_{j>J} y_j\psi_j\|_{L^2}^2\Big) = \sum_{j>J} \lambda_j.$$

Case of a stationary process : $C_a(x,z) = \kappa(x-z)$, that is T is a convolution operator. If D is the m-dimensional 2π-periodic torus, the KL basis is of Fourier type

$$x \mapsto \varphi_k(x) := (2\pi)^{-m/2}e^{ik\cdot x}, \quad k \in \mathbb{Z}^m.$$

The ξ_j are centered and decorelated scalar random variables, with

$$\mathbb{E}(\xi_j) = 0, \quad \mathbb{E}(\xi_i \xi_j) = 0 \quad \text{if} \quad j \neq i, \quad \mathbb{E}(|\xi_j|^2) = \lambda_j.$$

If the random process a is bounded, then the variables ξ_j have bounded range $|\xi_j| \leq c_j$, so that with $y_j := \xi_j / c_j$ and $\psi_j := c_j \varphi_j$ we may also write

$$a = \overline{a} + \sum_{j \geq 1} y_j \psi_j, \quad y = (y_j)_{j \geq 1} \in U = [-1, 1]^{\mathbb{N}}.$$

The KL representation is optimal for trunctation in mean-square $L^2(D)$-error :

$$\inf_{\dim(E)=J} \mathbb{E}(\|\tilde{a} - P_E \tilde{a}\|_{L^2}^2),$$

is attained by $E = E_J := \text{span}\{\psi_1, \dots, \psi_J\}$ with

$$\mathbb{E}(\|\tilde{a} - P_{E_J} \tilde{a}\|_{L^2}^2) = \mathbb{E}\Big(\|\sum_{j>J} y_j \psi_j\|_{L^2}^2\Big) = \sum_{j>J} \lambda_j.$$

Case of a stationary process : $C_a(x, z) = \kappa(x - z)$, that is T is a convolution operator. If D is the m-dimensional 2π-periodic torus, the KL basis is of Fourier type

$$x \mapsto \varphi_k(x) := (2\pi)^{-m/2} e^{ik\dot{x}}, \quad k \in \mathbb{Z}^m.$$

The ξ_j are centered and decorelated scalar random variables, with

$$\mathbb{E}(\xi_j) = 0, \quad \mathbb{E}(\xi_i \xi_j) = 0 \quad \text{if} \quad j \neq i, \quad \mathbb{E}(|\xi_j|^2) = \lambda_j.$$

If the random process a is bounded, then the variables ξ_j have bounded range $|\xi_j| \leq c_j$, so that with $y_j := \xi_j / c_j$ and $\psi_j := c_j \varphi_j$ we may also write

$$a = \bar{a} + \sum_{j \geq 1} y_j \psi_j, \quad y = (y_j)_{j \geq 1} \in U = [-1, 1]^{\mathbb{N}}.$$

The KL representation is optimal for trunctation in mean-square $L^2(D)$-error :

$$\inf_{\dim(E)=J} \mathbb{E}(\|\tilde{a} - P_E \tilde{a}\|_{L^2}^2),$$

is attained by $E = E_J := \mathrm{span}\{\psi_1, \dots, \psi_J\}$ with

$$\mathbb{E}(\|\tilde{a} - P_{E_J} \tilde{a}\|_{L^2}^2) = \mathbb{E}\Big(\| \sum_{j>J} y_j \psi_j \|_{L^2}^2\Big) = \sum_{j>J} \lambda_j.$$

Case of a stationary process : $C_a(x, z) = \kappa(x - z)$, that is T is a convolution operator. If D is the m-dimensional 2π-periodic torus, the KL basis is of Fourier type

$$x \mapsto \varphi_k(x) := (2\pi)^{-m/2} e^{ik\dot{x}}, \quad k \in \mathbb{Z}^m.$$

Objective : fast approximate computation of $y \mapsto u(y)$ for many queries of y.

Vehicle : separable (low rank) approximations of the form

$$u(x,y) \approx u_n(x,y) := \sum_{k=1}^{n} v_k(x)\phi_k(y),$$

where $v_k : D \to \mathbb{R}$ with $v_k \in V$ and $\phi_k : U \to \mathbb{R}$. Equivalently

$$u_n(y) := \sum_{k=1}^{n} v_k\phi_k(y) = \sum_{k=1}^{n} \phi_k(y)v_k \in V_n := \text{span}\{v_1, \dots, v_n\} \subset V, \quad y \in U.$$

Thus we approximate simultaneously all solutions $u(y)$ in the same n-dimensional space $V_n \subset V$.

By the way, this is what we do when we use a finite element solver :

$$y \mapsto u_h(y) \in V_h \subset V.$$

So what's new here ?

Accurate solutions may require V_h of very large dimension $N_h = \dim(V_h) >> 1$ and each query $y \mapsto u_h(y)$ is expensive.

We hope to achieve same order of accuracy $n << N_h$ by a choice of V_n adapted to the parametric problem. In practice the functions $v_1, \dots, v_n$ are typically picked from such a finite element space V_h, so that $u_n(y) \in V_h$ for all y but actually belongs to the much smaller space $V_n \subset V_h$.

Objective : fast approximate computation of $y \mapsto u(y)$ for many queries of y.

Vehicle : separable (low rank) approximations of the form

$$u(x,y) \approx u_n(x,y) := \sum_{k=1}^{n} v_k(x)\phi_k(y),$$

where $v_k : D \to \mathbb{R}$ with $v_k \in V$ and $\phi_k : U \to \mathbb{R}$. Equivalently

$$u_n(y) := \sum_{k=1}^{n} v_k \phi_k(y) = \sum_{k=1}^{n} \phi_k(y) v_k \in V_n := \mathrm{span}\{v_1, \dots, v_n\} \subset V, \quad y \in U.$$

Thus we approximate simultaneously all solutions $u(y)$ in the same n-dimensional space $V_n \subset V$.

By the way, this is what we do when we use a finite element solver :

$$y \mapsto u_h(y) \in V_h \subset V.$$

So what's new here ?

Accurate solutions may require V_h of very large dimension $N_h = \dim(V_h) >> 1$ and each query $y \mapsto u_h(y)$ is expensive.

We hope to achieve same order of accuracy $n << N_h$ by a choice of V_n adapted to the parametric problem. In practice the functions $v_1, \dots, v_n$ are typically picked from such a finite element space V_h, so that $u_n(y) \in V_h$ for all y but actually belongs to the much smaller space $V_n \subset V_h$.

Objective : fast approximate computation of $y \mapsto u(y)$ for many queries of y.

Vehicle : separable (low rank) approximations of the form

$$u(x,y) \approx u_n(x,y) := \sum_{k=1}^{n} v_k(x)\phi_k(y),$$

where $v_k : D \to \mathbb{R}$ with $v_k \in V$ and $\phi_k : U \to \mathbb{R}$. Equivalently

$$u_n(y) := \sum_{k=1}^{n} v_k\phi_k(y) = \sum_{k=1}^{n} \phi_k(y)v_k \in V_n := \mathrm{span}\{v_1,\dots,v_n\} \subset V, \quad y \in U.$$

Thus we approximate simultaneously all solutions $u(y)$ in the same n-dimensional space $V_n \subset V$.

By the way, this is what we do when we use a finite element solver :

$$y \mapsto u_h(y) \in V_h \subset V.$$

So what's new here ?

Accurate solutions may require V_h of very large dimension $N_h = \dim(V_h) >> 1$ and each query $y \mapsto u_h(y)$ is expensive.

We hope to achieve same order of accuracy $n << N_h$ by a choice of V_n adapted to the parametric problem. In practice the functions $v_1, \dots, v_n$ are typically picked from such a finite element space V_h, so that $u_n(y) \in V_h$ for all y but actually belongs to the much smaller space $V_n \subset V_h$.

Objective : fast approximate computation of $y \mapsto u(y)$ for many queries of y.

Vehicle : separable (low rank) approximations of the form

$$u(x,y) \approx u_n(x,y) := \sum_{k=1}^{n} v_k(x)\phi_k(y),$$

where $v_k : D \to \mathbb{R}$ with $v_k \in V$ and $\phi_k : U \to \mathbb{R}$. Equivalently

$$u_n(y) := \sum_{k=1}^{n} v_k \phi_k(y) = \sum_{k=1}^{n} \phi_k(y) v_k \in V_n := \mathrm{span}\{v_1, \dots, v_n\} \subset V, \quad y \in U.$$

Thus we approximate simultaneously all solutions $u(y)$ in the same n-dimensional space $V_n \subset V$.

By the way, this is what we do when we use a finite element solver :

$$y \mapsto u_h(y) \in V_h \subset V.$$

So what's new here ?

Accurate solutions may require V_h of very large dimension $N_h = \dim(V_h) >> 1$ and each query $y \mapsto u_h(y)$ is expensive.

We hope to achieve same order of accuracy $n << N_h$ by a choice of V_n adapted to the parametric problem. In practice the functions $v_1, \dots, v_n$ are typically picked from such a finite element space V_h, so that $u_n(y) \in V_h$ for all y but actually belongs to the much smaller space $V_n \subset V_h$.

Measure of performance

1. Uniform sense

$$\|u-u_n\|_{L^\infty(U,V)} := \sup_{y\in U} \|u(y)-u_n(y)\|_V,$$

2. Mean-square sense, for some measure μ on U,

$$\|u-u_n\|^2_{L^2(U,V,\mu)} := \int_U \|u(y)-u_n(y)\|_V^2 d\mu(y).$$

If μ is a probability measure, and y randomly distributed according to this measure, we have

$$\|u-u_n\|^2_{L^2(U,V,\mu)} = \mathbb{E}(\|u(y)-u_n(y)\|_V^2).$$

Note that we always have

$$\mathbb{E}(\|u(y)-u_n(y)\|_V^2) \le \|u-u_n\|^2_{L^\infty(U,V)}.$$

A "worst case" estimate is always above an "average" estimate.

Measure of performance

1. Uniform sense

$$\|u - u_n\|_{L^\infty(U,V)} := \sup_{y \in U} \|u(y) - u_n(y)\|_V,$$

2. Mean-square sense, for some measure μ on U,

$$\|u - u_n\|^2_{L^2(U,V,\mu)} := \int_U \|u(y) - u_n(y)\|^2_V d\mu(y).$$

If μ is a probability measure, and y randomly distributed according to this measure, we have

$$\|u - u_n\|^2_{L^2(U,V,\mu)} = \mathbb{E}(\|u(y) - u_n(y)\|^2_V).$$

Note that we always have

$$\mathbb{E}(\|u(y) - u_n(y)\|^2_V) \leq \|u - u_n\|^2_{L^\infty(U,V)}.$$

A "worst case" estimate is always above an "average" estimate.

Measure of performance

1. Uniform sense

$$\|u-u_n\|_{L^\infty(U,V)} := \sup_{y\in U} \|u(y)-u_n(y)\|_V,$$

2. Mean-square sense, for some measure μ on U,

$$\|u-u_n\|^2_{L^2(U,V,\mu)} := \int_U \|u(y)-u_n(y)\|_V^2 d\mu(y).$$

If μ is a probability measure, and y randomly distributed according to this measure, we have

$$\|u-u_n\|^2_{L^2(U,V,\mu)} = \mathbb{E}(\|u(y)-u_n(y)\|_V^2).$$

Note that we always have

$$\mathbb{E}(\|u(y)-u_n(y)\|_V^2) \le \|u-u_n\|^2_{L^\infty(U,V)}.$$

A "worst case" estimate is always above an "average" estimate.

Optimal spaces ?

Best n-dimensional space for approximation in the uniform sense : the space F_n one that reaches the Kolmogorov n-width of the solution manifold in the V norm

$$d_n = d_n(\mathcal{M}) := \inf_{\dim(E)\leq n} \sup_{v\in\mathcal{M}} \min_{w\in E} \|v-w\|_V = \inf_{\dim(E)\leq n} \sup_{y\in U} \min_{w\in E} \|u(y)-w\|_V.$$

Best n-dimensional space for approximation in the mean-square sense : principal component analysis in V (instead of L^2 with KL basis). Consider an orthonormal basis $(e_k)_{k\geq 1}$ of V and decompose

$$u(y) := \sum_{k\geq 1} u_k(y) e_k, \quad u_k(y) := \langle u(y), e_k\rangle_V.$$

Introduce the infinite correlation matrix $M = (\mathbb{E}(u_k u_l))_{k,l\geq 1}$. It has eigenvalues $(\lambda_k)_{k\geq 1}$ and associated eigenvectors $g_k = (g_{k,l})_{l\in\mathbb{N}}$ which form an orthonormal basis of $\ell^2(\mathbb{N})$. The best space is

$$G_n := \mathrm{span}\{v_1,\dots,v_n\}, \quad v_k := \sum_{l\geq 1} g_{k,l} e_l,$$

and has performance

$$\varepsilon_n^2 := \inf_{\dim(E)\leq n} \mathbb{E}\Big(\min_{w\in E} \|u(y)-w\|_V^2\Big) = \sum_{k>n} \lambda_k \leq d_n^2.$$

Optimal spaces ?

Best n-dimensional space for approximation in the uniform sense : the space F_n one that reaches the Kolmogorov n-width of the solution manifold in the V norm

$$d_n = d_n(\mathcal{M}) := \inf_{\dim(E)\leq n} \sup_{v\in\mathcal{M}} \min_{w\in E} \|v-w\|_V = \inf_{\dim(E)\leq n} \sup_{y\in U} \min_{w\in E} \|u(y)-w\|_V.$$

Best n-dimensional space for approximation in the mean-square sense : principal component analysis in V (instead of L^2 with KL basis). Consider an orthonormal basis $(e_k)_{k\geq 1}$ of V and decompose

$$u(y) := \sum_{k\geq 1} u_k(y) e_k, \quad u_k(y) := \langle u(y), e_k\rangle_V.$$

Introduce the infinite correlation matrix $M = (\mathbb{E}(u_k u_l))_{k,l\geq 1}$. It has eigenvalues $(\lambda_k)_{k\geq 1}$ and associated eigenvectors $g_k = (g_{k,l})_{l\in\mathbb{N}}$ which form an orthonormal basis of $\ell^2(\mathbb{N})$. The best space is

$$G_n := \mathrm{span}\{v_1,\dots,v_n\}, \quad v_k := \sum_{l\geq 1} g_{k,l} e_l,$$

and has performance

$$\varepsilon_n^2 := \inf_{\dim(E)\leq n} \mathbb{E}\left(\min_{w\in E} \|u(y)-w\|_V^2\right) = \sum_{k>n} \lambda_k \leq d_n^2.$$

Realistic strategies

The optimal spaces F_n and G_n are usually out of reach. There are two main computational approaches to realistically design the approximation $u_n = \sum_{k=1}^n v_k \phi_k$.

1. Expand formally the solution map $y \mapsto u(y)$ in a given "basis" $(\phi_k)_{k\geq 1}$ of high dimensional functions

$$u(y) = \sum_{k\geq 1} v_k \phi_k(y),$$

where $v_k \in V$ are viewed as the coefficients in this expansion.

Compute these coefficients for $k = 1, \dots, n$ approximately by some numerical procedure.

Main representative (this lecture) : Polynomial methods (the ϕ_k are multivariate polynomials).

2. Compute first a "good" basis $\{v_1, \dots, v_n\}$ and define V_n as their span. Then, for any given instance y, compute $u_n(y) \in V_n$ by a numerical method.

Main representative : Reduced Bases (RB) methods emulate the n-width spaces F_n for uniform, or $L^\infty(U, V)$, approximation. Proper Orthogonal Decompositions (POD) methods emulate the principal component spaces G_n for mean-square, or $L^2(U, V, \mu)$, approximation.

Realistic strategies

The optimal spaces F_n and G_n are usually out of reach. There are two main computational approaches to realistically design the approximation $u_n = \sum_{k=1}^n v_k \phi_k$.

1. Expand formally the solution map $y \mapsto u(y)$ in a given "basis" $(\phi_k)_{k\geq 1}$ of high dimensional functions

$$u(y) = \sum_{k\geq 1} v_k \phi_k(y),$$

where $v_k \in V$ are viewed as the coefficients in this expansion.

Compute these coefficients for $k = 1, \dots, n$ approximately by some numerical procedure.

Main representative (this lecture) : Polynomial methods (the ϕ_k are multivariate polynomials).

2. Compute first a "good" basis $\{v_1, \dots, v_n\}$ and define V_n as their span. Then, for any given instance y, compute $u_n(y) \in V_n$ by a numerical method.

Main representative : Reduced Bases (RB) methods emulate the n-width spaces F_n for uniform, or $L^\infty(U, V)$, approximation. Proper Orthogonal Decompositions (POD) methods emulate the principal component spaces G_n for mean-square, or $L^2(U, V, \mu)$, approximation.

Remarks

In the second approach, the functions v_k are typically computed in an heavy offline stage, then for any given y, the computation of $u_n(y)$ is done in a cheap online stage.

The first approach gives immediate access to the approximation u_n for all values of y since the functions v_k and ϕ_k are both precomputed offline, the online stage is then a trivial recombination.

Other important distinction : intrusive versus non-intrusive methods. The latter are based on post-processing individual solution instances

$$u(y^i), \quad y^i \in U, \quad i = 1, \dots, m.$$

They may benefit of a pre-existing numerical solver viewed as a blackbox and do not necessarily require full knowledge of PDE model.

In practice, the v_k are typically chosen in a discrete (finite element) space $V_h \subset V$, with $N_h = \dim(V_h) >> n$. Equivalently, we apply the above technique to the discrete solution map $y \mapsto u_h(y) \in V_h$. The error may thus be decomposed into the finite element discretization error and the model reduction error.

Remarks

In the second approach, the functions v_k are typically computed in an heavy offline stage, then for any given y, the computation of $u_n(y)$ is done in a cheap online stage.

The first approach gives immediate access to the approximation u_n for all values of y since the functions v_k and ϕ_k are both precomputed offline, the online stage is then a trivial recombination.

Other important distinction : intrusive versus non-intrusive methods. The latter are based on post-processing individual solution instances

$$u(y^i), \quad y^i \in U, \quad i = 1, \dots, m.$$

They may benefit of a pre-existing numerical solver viewed as a blackbox and do not necessarily require full knowledge of PDE model.

In practice, the v_k are typically chosen in a discrete (finite element) space $V_h \subset V$, with $N_h = \dim(V_h) >> n$. Equivalently, we apply the above technique to the discrete solution map $y \mapsto u_h(y) \in V_h$. The error may thus be decomposed into the finite element discretization error and the model reduction error.

Remarks

In the second approach, the functions v_k are typically computed in an heavy offline stage, then for any given y, the computation of $u_n(y)$ is done in a cheap online stage.

The first approach gives immediate access to the approximation u_n for all values of y since the functions v_k and ϕ_k are both precomputed offline, the online stage is then a trivial recombination.

Other important distinction : intrusive versus non-intrusive methods. The latter are based on post-processing individual solution instances

$$u(y^i), \quad y^i \in U, \quad i = 1, \dots, m.$$

They may benefit of a pre-existing numerical solver viewed as a blackbox and do not necessarily require full knowledge of PDE model.

In practice, the v_k are typically chosen in a discrete (finite element) space $V_h \subset V$, with $N_h = \dim(V_h) >> n$. Equivalently, we apply the above technique to the discrete solution map $y \mapsto u_h(y) \in V_h$. The error may thus be decomposed into the finite element discretization error and the model reduction error.

How to defeat the curse of dimensionality ?

The map $y \mapsto u(y)$ is high dimensional, or even infinite dimensional $y = (y_j)_{j \geq 1}$.

We are thus facing the curse of dimensionality when trying to approximate it with conventional discretization tools in the y variable (Fourier series, finite elements).

A general function of d variable with m bounded derivatives cannot be approximated in L^∞ with rate better than $n^{-m/d}$ where n is the number of degrees of freedom.

A possible way out : exploit anisotropic features in the function $y \mapsto u(y)$.

The PDE is parametrized by a function a (diffusion coefficient, velocity, domain boundary) and y_j are the coordinates of a in a certain basis representation $a = \overline{a} + \sum_{j \geq 1} y_j \psi_j$.

If the ψ_j decays as $j \to +\infty$ (for instance if a has some smoothness) then the variable y_j are less active for large j.

We shall see that in certain relevant instances, this mechanism allows to break the curse of dimensionality by using suitable expansions : we obtain approximation rates $\mathcal{O}(N^{-s})$ that are independent of d in the sense that they hold when $d = \infty$.

One key tool for obtaining such result is the concept of sparse approximation.

How to defeat the curse of dimensionality ?

The map $y \mapsto u(y)$ is high dimensional, or even infinite dimensional $y = (y_j)_{j \geq 1}$.

We are thus facing the curse of dimensionality when trying to approximate it with conventional discretization tools in the y variable (Fourier series, finite elements).

A general function of d variable with m bounded derivatives cannot be approximated in L^∞ with rate better than $n^{-m/d}$ where n is the number of degrees of freedom.

A possible way out : exploit anisotropic features in the function $y \mapsto u(y)$.

The PDE is parametrized by a function a (diffusion coefficient, velocity, domain boundary) and y_j are the coordinates of a in a certain basis representation $a = \overline{a} + \sum_{j \geq 1} y_j \psi_j$.

If the ψ_j decays as $j \to +\infty$ (for instance if a has some smoothness) then the variable y_j are less active for large j.

We shall see that in certain relevant instances, this mechanism allows to break the curse of dimensionality by using suitable expansions : we obtain approximation rates $\mathcal{O}(N^{-s})$ that are independent of d in the sense that they hold when $d = \infty$.

One key tool for obtaining such result is the concept of sparse approximation.

How to defeat the curse of dimensionality ?

The map $y \mapsto u(y)$ is high dimensional, or even infinite dimensional $y = (y_j)_{j \geq 1}$.

We are thus facing the curse of dimensionality when trying to approximate it with conventional discretization tools in the y variable (Fourier series, finite elements).

A general function of d variable with m bounded derivatives cannot be approximated in L^∞ with rate better than $n^{-m/d}$ where n is the number of degrees of freedom.

A possible way out : exploit anisotropic features in the function $y \mapsto u(y)$.

The PDE is parametrized by a function a (diffusion coefficient, velocity, domain boundary) and y_j are the coordinates of a in a certain basis representation $a = \bar{a} + \sum_{j \geq 1} y_j \psi_j$.

If the ψ_j decays as $j \to +\infty$ (for instance if a has some smoothness) then the variable y_j are less active for large j.

We shall see that in certain relevant instances, this mechanism allows to break the curse of dimensionality by using suitable expansions : we obtain approximation rates $\mathcal{O}(N^{-s})$ that are independent of d in the sense that they hold when $d = \infty$.

One key tool for obtaining such result is the concept of sparse approximation.

Consider sequences $\mathbf{d} = (d_\nu)_{\nu\in\mathcal{F}}$ in $\ell^q(\mathcal{F})$ where $\mathcal{F}$ is a countable index set.

Best n-term approximation : we seek to approximate $\mathbf{d}$ by a sequence supported on a set of size n.

Best choice : $\mathbf{d}_n$ defined by leaving d_ν unchanged for the n largest $|d_\nu|$ and setting the others to 0.

Lemma : for $0 < p < q \leq \infty$, one has

$$\mathbf{d} \in \ell^p(\mathcal{F}) \Rightarrow \|\mathbf{d} - \mathbf{d}_n\|_{\ell^q} \leq C(n+1)^{-s}, \quad s = \frac{1}{p} - \frac{1}{q}, \quad C := \|\mathbf{d}\|_{\ell^p}.$$

Proof : introduce $(d_k^*)_{k\geq 1}$ the decreasing rearrangement of $(|d_\nu|)_{\nu\in\mathcal{F}}$, and combine

$$\|\mathbf{d} - \mathbf{d}_n\|_{\ell^q}^q = \sum_{k>n} |d_k^*|^q = \sum_{k>n} |d_k^*|^{q-p} |d_k^*|^p \leq C^p |d_{n+1}^*|^{q-p}$$

with

$$(n+1)|d_{n+1}^*|^p \leq \sum_{k=1}^{n+1} |d_k^*|^p \leq C^p.$$

Note that a large value of s corresponds to a value $p < 1$ (non-convex spaces).

Consider sequences $\mathbf{d} = (d_\nu)_{\nu\in\mathcal{F}}$ in $\ell^q(\mathcal{F})$ where $\mathcal{F}$ is a countable index set.

Best n-term approximation : we seek to approximate $\mathbf{d}$ by a sequence supported on a set of size n.

Best choice : $\mathbf{d}_n$ defined by leaving d_ν unchanged for the n largest $|d_\nu|$ and setting the others to 0.

Lemma : for $0 < p < q \leq \infty$, one has

$$\mathbf{d} \in \ell^p(\mathcal{F}) \Rightarrow \|\mathbf{d} - \mathbf{d}_n\|_{\ell^q} \leq C(n+1)^{-s}, \quad s = \frac{1}{p} - \frac{1}{q}, \quad C := \|\mathbf{d}\|_{\ell^p}.$$

Proof : introduce $(d_k^*)_{k\geq 1}$ the decreasing rearrangement of $(|d_\nu|)_{\nu\in\mathcal{F}}$, and combine

$$\|\mathbf{d} - \mathbf{d}_n\|_{\ell^q}^q = \sum_{k>n} |d_k^*|^q = \sum_{k>n} |d_k^*|^{q-p} |d_k^*|^p \leq C^p |d_{n+1}^*|^{q-p}$$

with

$$(n+1)|d_{n+1}^*|^p \leq \sum_{k=1}^{n+1} |d_k^*|^p \leq C^p.$$

Note that a large value of s corresponds to a value $p < 1$ (non-convex spaces).

Consider sequences $\mathbf{d} = (d_\nu)_{\nu\in\mathcal{F}}$ in $\ell^q(\mathcal{F})$ where $\mathcal{F}$ is a countable index set.

Best n-term approximation : we seek to approximate $\mathbf{d}$ by a sequence supported on a set of size n.

Best choice : $\mathbf{d}_n$ defined by leaving d_ν unchanged for the n largest $|d_\nu|$ and setting the others to 0.

Lemma : for $0 < p < q \leq \infty$, one has

$$\mathbf{d} \in \ell^p(\mathcal{F}) \Rightarrow \|\mathbf{d} - \mathbf{d}_n\|_{\ell^q} \leq C(n+1)^{-s}, \quad s = \frac{1}{p} - \frac{1}{q}, \quad C := \|\mathbf{d}\|_{\ell^p}.$$

Proof : introduce $(d_k^*)_{k\geq 1}$ the decreasing rearrangement of $(|d_\nu|)_{\nu\in\mathcal{F}}$, and combine

$$\|\mathbf{d} - \mathbf{d}_n\|_{\ell^q}^q = \sum_{k>n} |d_k^*|^q = \sum_{k>n} |d_k^*|^{q-p}|d_k^*|^p \leq C^p |d_{n+1}^*|^{q-p}$$

with

$$(n+1)|d_{n+1}^*|^p \leq \sum_{k=1}^{n+1} |d_k^*|^p \leq C^p.$$

Note that a large value of s corresponds to a value $p < 1$ (non-convex spaces).

Consider sequences $\mathbf{d} = (d_\nu)_{\nu\in\mathcal{F}}$ in $\ell^q(\mathcal{F})$ where $\mathcal{F}$ is a countable index set.

Best n-term approximation : we seek to approximate $\mathbf{d}$ by a sequence supported on a set of size n.

Best choice : $\mathbf{d}_n$ defined by leaving d_ν unchanged for the n largest $|d_\nu|$ and setting the others to 0.

Lemma : for $0 < p < q \leq \infty$, one has

$$\mathbf{d} \in \ell^p(\mathcal{F}) \Rightarrow \|\mathbf{d} - \mathbf{d}_n\|_{\ell^q} \leq C(n+1)^{-s}, \quad s = \frac{1}{p} - \frac{1}{q}, \quad C := \|\mathbf{d}\|_{\ell^p}.$$

Proof : introduce $(d_k^*)_{k\geq 1}$ the decreasing rearrangement of $(|d_\nu|)_{\nu\in\mathcal{F}}$, and combine

$$\|\mathbf{d} - \mathbf{d}_n\|_{\ell^q}^q = \sum_{k>n} |d_k^*|^q = \sum_{k>n} |d_k^*|^{q-p}|d_k^*|^p \leq C^p|d_{n+1}^*|^{q-p}$$

with

$$(n+1)|d_{n+1}^*|^p \leq \sum_{k=1}^{n+1} |d_k^*|^p \leq C^p.$$

Note that a large value of s corresponds to a value $p < 1$ (non-convex spaces).

If a V-valued u has an expansion of the form $u(y) = \sum_{\nu\in\mathcal{F}} u_\nu \phi_\nu(y)$, in a given basis $(\phi_\nu)_{\nu\in\mathcal{F}}$, we use Stechkin's lemma to study the approximation of u by

$$u_n := \sum_{\nu\in\Lambda_n} u_\nu \phi_\nu,$$

where $\Lambda_n \subset \mathcal{F}$ corresponds to the n-largest $\|u_\nu\|_V$.

If $\sup_{y\in U} |\phi_\nu(y)| = 1$, then by triangle inequality

$$\|u - u_n\|_{L^\infty(U,V)} \leq \sum_{\nu\notin\Lambda_n} \|u_\nu\phi_\nu\|_{L^\infty(U,V)} = \sum_{\nu\notin\Lambda_n} \|u_\nu\|_V,$$

If $(\phi_\nu)_{\nu\in\mathcal{F}}$ is an orthonormal basis of $L^2(U,\mu)$, then by Parseval equality

$$\|u - u_n\|^2_{L^2(U,V,\mu)} = \sum_{\nu\notin\Lambda_n} \|u_\nu\|^2_V,$$

For concrete choices of bases a relevant question is thus : what smoothness properties of a function ensure that its coefficient sequence belongs to ℓ^p for small values of p ?

In the case of wavelet bases, such properties are characterized by Besov spaces.

In our present setting of high-dimensional functions $y \mapsto u(y)$ we shall rather use tensor-product polynomial bases instead of wavelet bases. Sparsity properties will follow to the anisotropic features of these functions.

From sequence approximation to Banach space valued function approximation

If a V-valued u has an expansion of the form $u(y) = \sum_{\nu\in\mathcal{F}} u_\nu \phi_\nu(y)$, in a given basis $(\phi_\nu)_{\nu\in\mathcal{F}}$, we use Stechkin's lemma to study the approximation of u by

$$u_n := \sum_{\nu\in\Lambda_n} u_\nu \phi_\nu,$$

where $\Lambda_n \subset \mathcal{F}$ corresponds to the n-largest $\|u_\nu\|_V$.

If $\sup_{y\in U} |\phi_\nu(y)| = 1$, then by triangle inequality

$$\|u - u_n\|_{L^\infty(U,V)} \leq \sum_{\nu\notin\Lambda_n} \|u_\nu\phi_\nu\|_{L^\infty(U,V)} = \sum_{\nu\notin\Lambda_n} \|u_\nu\|_V,$$

If $(\phi_\nu)_{\nu\in\mathcal{F}}$ is an orthonormal basis of $L^2(U,\mu)$, then by Parseval equality

$$\|u - u_n\|^2_{L^2(U,V,\mu)} = \sum_{\nu\notin\Lambda_n} \|u_\nu\|^2_V,$$

For concrete choices of bases a relevant question is thus : what smoothness properties of a function ensure that its coefficient sequence belongs to ℓ^p for small values of p ?

In the case of wavelet bases, such properties are characterized by Besov spaces.

In our present setting of high-dimensional functions $y \mapsto u(y)$ we shall rather use tensor-product polynomial bases instead of wavelet bases. Sparsity properties will follow to the anisotropic features of these functions.

If a V-valued u has an expansion of the form $u(y) = \sum_{\nu\in\mathcal{F}} u_\nu \phi_\nu(y)$, in a given basis $(\phi_\nu)_{\nu\in\mathcal{F}}$, we use Stechkin's lemma to study the approximation of u by

$$u_n := \sum_{\nu\in\Lambda_n} u_\nu \phi_\nu,$$

where $\Lambda_n \subset \mathcal{F}$ corresponds to the n-largest $\|u_\nu\|_V$.

If $\sup_{y\in U} |\phi_\nu(y)| = 1$, then by triangle inequality

$$\|u - u_n\|_{L^\infty(U,V)} \leq \sum_{\nu\notin\Lambda_n} \|u_\nu\phi_\nu\|_{L^\infty(U,V)} = \sum_{\nu\notin\Lambda_n} \|u_\nu\|_V,$$

If $(\phi_\nu)_{\nu\in\mathcal{F}}$ is an orthonormal basis of $L^2(U,\mu)$, then by Parseval equality

$$\|u - u_n\|^2_{L^2(U,V,\mu)} = \sum_{\nu\notin\Lambda_n} \|u_\nu\|_V^2,$$

For concrete choices of bases a relevant question is thus : what smoothness properties of a function ensure that its coefficient sequence belongs to ℓ^p for small values of p ?

In the case of wavelet bases, such properties are characterized by Besov spaces.

In our present setting of high-dimensional functions $y \mapsto u(y)$ we shall rather use tensor-product polynomial bases instead of wavelet bases. Sparsity properties will follow to the anisotropic features of these functions.

Steady state diffusion equation

$$-\mathrm{div}(a\nabla u) = f \quad \text{on} \quad D \subset \mathbb{R}^m \quad \text{and} \quad u_{|\partial D} = 0,$$

where $f = f(x) \in L^2(D)$ and the diffusion coefficients are given by

$$a = a(x, y) = \bar{a}(x) + \sum_{j\geq 1} y_j \psi_j(x),$$

where $\bar{a}$ and the $(\psi_j)_{j\geq 1}$ are given functions and $y \in U := [-1, 1]^{\mathbb{N}}$. Uniform ellipticity assumption :

$$(UEA) \qquad 0 < r \leq a(x, y) \leq R, \quad x \in D, \; y \in U.$$

Equivalent expression of (UEA) : $\bar{a} \in L^\infty(D)$ and

$$\sum_{j\geq 1} |\psi_j(x)| \leq \bar{a}(x) - r, \quad x \in D,$$

or

$$\left\| \frac{\sum_{j\geq 1} |\psi_j|}{\bar{a}} \right\|_{L^\infty(D)} \leq \theta < 1.$$

Lax-Milgram : solution map is well-defined from U to $V := H_0^1(D)$ with uniform bound

$$\|u(y)\|_V \leq C_r := \frac{\|f\|_{V'}}{r}, \quad y \in U, \quad \text{where} \quad \|v\|_V := \|\nabla v\|_{L^2}.$$

Return to the main guiding example

Steady state diffusion equation

$$-\mathrm{div}(a\nabla u) = f \quad \text{on} \quad D \subset \mathbb{R}^m \quad \text{and} \quad u_{|\partial D} = 0,$$

where $f = f(x) \in L^2(D)$ and the diffusion coefficients are given by

$$a = a(x,y) = \bar{a}(x) + \sum_{j\geq 1} y_j \psi_j(x),$$

where $\bar{a}$ and the $(\psi_j)_{j\geq 1}$ are given functions and $y \in U := [-1,1]^{\mathbb{N}}$. Uniform ellipticity assumption :

$$(UEA) \qquad 0 < r \leq a(x,y) \leq R, \quad x \in D,\ y \in U.$$

Equivalent expression of (UEA) : $\bar{a} \in L^\infty(D)$ and

$$\sum_{j\geq 1} |\psi_j(x)| \leq \bar{a}(x) - r, \quad x \in D,$$

or

$$\left\| \frac{\sum_{j\geq 1} |\psi_j|}{\bar{a}} \right\|_{L^\infty(D)} \leq \theta < 1.$$

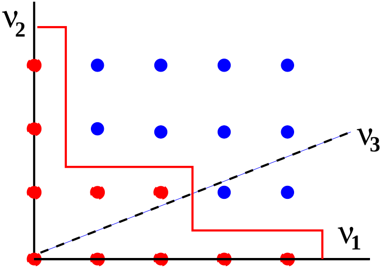
Lax-Milgram : solution map is well-defined from U to $V := H_0^1(D)$ with uniform bound

$$\|u(y)\|_V \leq C_r := \frac{\|f\|_{V'}}{r}, \quad y \in U, \quad \text{where} \quad \|v\|_V := \|\nabla v\|_{L^2}.$$

We consider the expansion of $u(y) = \sum_{\nu\in\mathcal{F}} t_\nu y^\nu$, where

$$y^\nu := \prod_{j\geq 1} y_j^{\nu_j} \text{ and } t_\nu := \frac{1}{\nu!}\partial^\nu u_{|y=0} \in V \text{ with } \nu! := \prod_{j\geq 1} \nu_j! \text{ and } 0! := 1.$$

where $\mathcal{F}$ is the set of all finitely supported sequences of integers (finitely many $\nu_j \neq 0$). The sequence $(t_\nu)_{\nu\in\mathcal{F}}$ is indexed by countably many integers.



Objective : identify a set $\Lambda \subset \mathcal{F}$ with $\#(\Lambda) = n$ such that u is well approximated by the partial expansion

$$u_\Lambda(y) := \sum_{\nu\in\Lambda} t_\nu y^\nu.$$

Best n-term approximation

A-priori choices for Λ have been proposed, e.g. (anisotropic) sparse grid defined by restrictions of the type $\sum_j \alpha_j \nu_j \leq A(n)$ or $\prod_j (1+\beta_j \nu_j) \leq B(n)$.

Instead we want to choose Λ optimally adapted to u. By triangle inequality we have

$$\|u-u_\Lambda\|_{L^\infty(U,V)} = \sup_{y\in U} \|u(y)-u_\Lambda(y)\|_V \leq \sup_{y\in U} \sum_{\nu\notin\Lambda} \|t_\nu y^\nu\|_V = \sum_{\nu\notin\Lambda} \|t_\nu\|_V$$

Best n-term approximation in $\ell^1(\mathcal{F})$ norm : use $\Lambda = \Lambda_n$ index set of n largest $\|t_\nu\|_V$.

Stechkin lemma : if $(\|t_\nu\|_V)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F})$ for some $p<1$, then for this Λ_n,

$$\sum_{\nu\notin\Lambda_n} \|t_\nu\|_V \leq Cn^{-s}, \quad s := \frac{1}{p}-1, \quad C := \|(\|t_\nu\|_V)\|_{\ell^p}.$$

Question : do we have $(\|t_\nu\|_V)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F})$ for some $p<1$?

Best n-term approximation

A-priori choices for Λ have been proposed, e.g. (anisotropic) sparse grid defined by restrictions of the type $\sum_j \alpha_j \nu_j \leq A(n)$ or $\prod_j (1+\beta_j \nu_j) \leq B(n)$.

Instead we want to choose Λ optimally adapted to u. By triangle inequality we have

$$\|u - u_\Lambda\|_{L^\infty(U,V)} = \sup_{y \in U} \|u(y) - u_\Lambda(y)\|_V \leq \sup_{y \in U} \sum_{\nu \notin \Lambda} \|t_\nu y^\nu\|_V = \sum_{\nu \notin \Lambda} \|t_\nu\|_V$$

Best n-term approximation in $\ell^1(\mathcal{F})$ norm : use $\Lambda = \Lambda_n$ index set of n largest $\|t_\nu\|_V$.

Stechkin lemma : if $(\|t_\nu\|_V)_{\nu \in \mathcal{F}} \in \ell^p(\mathcal{F})$ for some $p < 1$, then for this Λ_n,

$$\sum_{\nu \notin \Lambda_n} \|t_\nu\|_V \leq C n^{-s}, \quad s := \frac{1}{p} - 1, \quad C := \|(\|t_\nu\|_V)\|_{\ell^p}.$$

Question : do we have $(\|t_\nu\|_V)_{\nu \in \mathcal{F}} \in \ell^p(\mathcal{F})$ for some $p < 1$?

Best n-term approximation

A-priori choices for Λ have been proposed, e.g. (anisotropic) sparse grid defined by restrictions of the type $\sum_j \alpha_j \nu_j \leq A(n)$ or $\prod_j (1+\beta_j \nu_j) \leq B(n)$.

Instead we want to choose Λ optimally adapted to u. By triangle inequality we have

$$\|u - u_\Lambda\|_{L^\infty(U,V)} = \sup_{y \in U} \|u(y) - u_\Lambda(y)\|_V \leq \sup_{y \in U} \sum_{\nu \notin \Lambda} \|t_\nu y^\nu\|_V = \sum_{\nu \notin \Lambda} \|t_\nu\|_V$$

Best n-term approximation in $\ell^1(\mathcal{F})$ norm : use $\Lambda = \Lambda_n$ index set of n largest $\|t_\nu\|_V$.

Stechkin lemma : if $(\|t_\nu\|_V)_{\nu \in \mathcal{F}} \in \ell^p(\mathcal{F})$ for some $p < 1$, then for this Λ_n,

$$\sum_{\nu \notin \Lambda_n} \|t_\nu\|_V \leq C n^{-s}, \quad s := \frac{1}{p} - 1, \quad C := \|(\|t_\nu\|_V)\|_{\ell^p}.$$

Question : do we have $(\|t_\nu\|_V)_{\nu \in \mathcal{F}} \in \ell^p(\mathcal{F})$ for some $p < 1$?

Best n-term approximation

A-priori choices for Λ have been proposed, e.g. (anisotropic) sparse grid defined by restrictions of the type $\sum_j \alpha_j \nu_j \leq A(n)$ or $\prod_j (1+\beta_j \nu_j) \leq B(n)$.

Instead we want to choose Λ optimally adapted to u. By triangle inequality we have

$$\|u - u_\Lambda\|_{L^\infty(U,V)} = \sup_{y\in U} \|u(y) - u_\Lambda(y)\|_V \leq \sup_{y\in U} \sum_{\nu\notin\Lambda} \|t_\nu y^\nu\|_V = \sum_{\nu\notin\Lambda} \|t_\nu\|_V$$

Best n-term approximation in $\ell^1(\mathcal{F})$ norm : use $\Lambda = \Lambda_n$ index set of n largest $\|t_\nu\|_V$.

Stechkin lemma : if $(\|t_\nu\|_V)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F})$ for some $p < 1$, then for this Λ_n,

$$\sum_{\nu\notin\Lambda_n} \|t_\nu\|_V \leq Cn^{-s}, \quad s := \frac{1}{p} - 1, \quad C := \|(\|t_\nu\|_V)\|_{\ell^p}.$$

Question : do we have $(\|t_\nu\|_V)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F})$ for some $p < 1$?

Theorem (Cohen-DeVore-Schwab, 2011) : under the uniform ellipticity assumption (UAE), then for any $p < 1$,

$$(\|\psi_j\|_{L^\infty})_{j>0} \in \ell^p(\mathbb{N}) \Rightarrow (\|t_\nu\|_V)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F}).$$

Interpretations :

(i) The Taylor expansion of $u(y)$ inherits the sparsity properties of the expansion of $a(y)$ into the ψ_j.

(ii) We approximate $u(y)$ in $L^\infty(U, V)$ with algebraic rate $\mathcal{O}(n^{-s})$ despite the curse of (infinite) dimensionality, due to the fact that y_j is less influencial as j gets large.

(iii) The solution manifold $\mathcal{M} := \{u(y)\ ;\ y \in U\}$ is uniformly well approximated by the n-dimensional space $V_n := \mathrm{span}\{t_\nu\ :\ \nu \in \Lambda_n\}$. Its n-width satisfies the bound

$$d_n(\mathcal{M})_V \le \max_{y\in U} \mathrm{dist}(u(y), V_n)_V \le \max_{y\in U} \|u(y) - u_{\Lambda_n}(y)\|_V \le Cn^{-s}.$$

Such approximation rates cannot be proved for the usual a-priori choices of Λ.

Same result for more general linear equations $Au = f$ with affine operator dependance : $A = \overline{A} + \sum_{j\ge1} y_j A_j$ uniformly invertible over $y \in U$, and $(\|A_j\|_{V\to W})_{j\ge1} \in \ell^p(\mathbb{N})$, as well as other models.

Theorem (Cohen-DeVore-Schwab, 2011) : under the uniform ellipticity assumption (UAE), then for any $p < 1$,

$$(\|\psi_j\|_{L^\infty})_{j>0} \in \ell^p(\mathbb{N}) \Rightarrow (\|t_\nu\|_V)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F}).$$

Interpretations :

(i) The Taylor expansion of $u(y)$ inherits the sparsity properties of the expansion of $a(y)$ into the ψ_j.

(ii) We approximate $u(y)$ in $L^\infty(U, V)$ with algebraic rate $\mathcal{O}(n^{-s})$ despite the curse of (infinite) dimensionality, due to the fact that y_j is less influencial as j gets large.

(iii) The solution manifold $\mathcal{M} := \{u(y)\ ;\ y \in U\}$ is uniformly well approximated by the n-dimensional space $V_n := \mathrm{span}\{t_\nu\ :\ \nu \in \Lambda_n\}$. Its n-width satisfies the bound

$$d_n(\mathcal{M})_V \leq \max_{y\in U} \mathrm{dist}(u(y), V_n)_V \leq \max_{y\in U} \|u(y) - u_{\Lambda_n}(y)\|_V \leq Cn^{-s}.$$

Such approximation rates cannot be proved for the usual a-priori choices of Λ.

Same result for more general linear equations $Au = f$ with affine operator dependance : $A = \overline{A} + \sum_{j\geq 1} y_j A_j$ uniformly invertible over $y \in U$, and $(\|A_j\|_{V\to W})_{j\geq 1} \in \ell^p(\mathbb{N})$, as well as other models.

Theorem (Cohen-DeVore-Schwab, 2011) : under the uniform ellipticity assumption (UAE), then for any $p < 1$,

$$(\|\psi_j\|_{L^\infty})_{j>0} \in \ell^p(\mathbb{N}) \Rightarrow (\|t_\nu\|_V)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F}).$$

Interpretations :

(i) The Taylor expansion of $u(y)$ inherits the sparsity properties of the expansion of $a(y)$ into the ψ_j.

(ii) We approximate $u(y)$ in $L^\infty(U,V)$ with algebraic rate $\mathcal{O}(n^{-s})$ despite the curse of (infinite) dimensionality, due to the fact that y_j is less influencial as j gets large.

(iii) The solution manifold $\mathcal{M} := \{u(y) \; ; \; y \in U\}$ is uniformly well approximated by the n-dimensional space $V_n := \mathrm{span}\{t_\nu \; : \; \nu \in \Lambda_n\}$. Its n-width satisfies the bound

$$d_n(\mathcal{M})_V \le \max_{y\in U} \mathrm{dist}(u(y), V_n)_V \le \max_{y\in U} \|u(y) - u_{\Lambda_n}(y)\|_V \le Cn^{-s}.$$

Such approximation rates cannot be proved for the usual a-priori choices of Λ.

Same result for more general linear equations $Au = f$ with affine operator dependance : $A = \overline{A} + \sum_{j\ge1} y_j A_j$ uniformly invertible over $y \in U$, and $(\|A_j\|_{V\to W})_{j\ge1} \in \ell^p(\mathbb{N})$, as well as other models.

Estimates on $\|t_\nu\|_V$ by complex analysis : extend $u(y)$ to $u(z)$ with $z = (z_j) \in \mathbb{C}^{\mathbb{N}}$.

Uniform ellipticity $\sum_{j\geq 1} |\psi_j| \leq \overline{a} - r$ implies that with $a(z) = \overline{a} + \sum_{j\geq 1} z_j\psi_j$,

$$0 < r \leq \Re(a(x,z)) \leq |a(x,z)| \leq 2R, \quad x \in D,$$

for all $z \in \mathcal{U} := \{|z| \leq 1\}^{\mathbb{N}} = \otimes_{j\geq 1}\{|z_j| \leq 1\}$.

Lax-Milgram theory applies : $\|u(z)\|_V \leq C_0 = \frac{\|f\|_{V^*}}{r}$ for all $z \in \mathcal{U}$.

The function $u \mapsto u(z)$ is holomorphic in each variable z_j at any $z \in \mathcal{U}$: its first derivative $\partial_{z_j} u(z)$ is the unique solution to

$$\int_D a(z)\nabla\partial_{z_j} u(z) \cdot \nabla v = -\int_D \psi_j \nabla u(z) \cdot \nabla v, \quad v \in V.$$

Note that ∇ is with respect to spatial variable $x \in D$.

Extended domains of holomorphy : if $\rho = (\rho_j)_{j\geq 0}$ is any positive sequence such that for some $\delta > 0$

$$\sum_{j\geq 1} \rho_j|\psi_j(x)| \leq \overline{a}(x) - \delta, \quad x \in D,$$

then u is holomorphic with uniform bound $\|u(z)\| \leq C_\delta = \frac{\|f\|_{V^*}}{\delta}$ in the polydisc

$$\mathcal{U}_\rho := \otimes_{j\geq 1}\{|z_j| \leq \rho_j\},$$

If $\delta < r$, we can take $\rho_j > 1$.

Idea of proof : extension to complex variable

Estimates on $\|t_\nu\|_V$ by complex analysis : extend $u(y)$ to $u(z)$ with $z = (z_j) \in \mathbb{C}^{\mathbb{N}}$.

Uniform ellipticity $\sum_{j\geq 1} |\psi_j| \leq \bar{a} - r$ implies that with $a(z) = \bar{a} + \sum_{j\geq 1} z_j \psi_j$,

$$0 < r \leq \Re(a(x,z)) \leq |a(x,z)| \leq 2R, \quad x \in D,$$

for all $z \in \mathcal{U} := \{|z| \leq 1\}^{\mathbb{N}} = \otimes_{j\geq 1}\{|z_j| \leq 1\}$.

Lax-Milgram theory applies : $\|u(z)\|_V \leq C_0 = \frac{\|f\|_{V^*}}{r}$ for all $z \in \mathcal{U}$.

The function $u \mapsto u(z)$ is holomorphic in each variable z_j at any $z \in \mathcal{U}$: its first derivative $\partial_{z_j} u(z)$ is the unique solution to

$$\int_D a(z) \nabla \partial_{z_j} u(z) \cdot \nabla v = -\int_D \psi_j \nabla u(z) \cdot \nabla v, \quad v \in V.$$

Note that ∇ is with respect to spatial variable $x \in D$.

Extended domains of holomorphy : if $\rho = (\rho_j)_{j\geq 0}$ is any positive sequence such that for some $\delta > 0$

$$\sum_{j\geq 1} \rho_j |\psi_j(x)| \leq \bar{a}(x) - \delta, \;\; x \in D,$$

then u is holomorphic with uniform bound $\|u(z)\| \leq C_\delta = \frac{\|f\|_{V^*}}{\delta}$ in the polydisc

$$\mathcal{U}_\rho := \otimes_{j\geq 1}\{|z_j| \leq \rho_j\},$$

If $\delta < r$, we can take $\rho_j > 1$.

Estimates on $\|t_\nu\|_V$ by complex analysis : extend $u(y)$ to $u(z)$ with $z = (z_j) \in \mathbb{C}^{\mathbb{N}}$.

Uniform ellipticity $\sum_{j\geq 1} |\psi_j| \leq \overline{a} - r$ implies that with $a(z) = \overline{a} + \sum_{j\geq 1} z_j \psi_j$,

$$0 < r \leq \Re(a(x,z)) \leq |a(x,z)| \leq 2R, \quad x \in D,$$

for all $z \in \mathcal{U} := \{|z| \leq 1\}^{\mathbb{N}} = \otimes_{j\geq 1}\{|z_j| \leq 1\}$.

Lax-Milgram theory applies : $\|u(z)\|_V \leq C_0 = \frac{\|f\|_{V^*}}{r}$ for all $z \in \mathcal{U}$.

The function $u \mapsto u(z)$ is holomorphic in each variable z_j at any $z \in \mathcal{U}$: its first derivative $\partial_{z_j} u(z)$ is the unique solution to

$$\int_D a(z) \nabla \partial_{z_j} u(z) \cdot \nabla v = -\int_D \psi_j \nabla u(z) \cdot \nabla v, \quad v \in V.$$

Note that ∇ is with respect to spatial variable $x \in D$.

Extended domains of holomorphy : if $\rho = (\rho_j)_{j\geq 0}$ is any positive sequence such that for some $\delta > 0$

$$\sum_{j\geq 1} \rho_j |\psi_j(x)| \leq \overline{a}(x) - \delta, \quad x \in D,$$

then u is holomorphic with uniform bound $\|u(z)\| \leq C_\delta = \frac{\|f\|_{V^*}}{\delta}$ in the polydisc

$$\mathcal{U}_\rho := \otimes_{j\geq 1}\{|z_j| \leq \rho_j\},$$

If $\delta < r$, we can take $\rho_j > 1$.

Estimates on $\|t_\nu\|_V$ by complex analysis : extend $u(y)$ to $u(z)$ with $z = (z_j) \in \mathbb{C}^{\mathbb{N}}$.

Uniform ellipticity $\sum_{j\geq 1} |\psi_j| \leq \bar{a} - r$ implies that with $a(z) = \bar{a} + \sum_{j\geq 1} z_j \psi_j$,

$$0 < r \leq \Re(a(x,z)) \leq |a(x,z)| \leq 2R, \quad x \in D,$$

for all $z \in \mathcal{U} := \{|z| \leq 1\}^{\mathbb{N}} = \otimes_{j\geq 1}\{|z_j| \leq 1\}$.

Lax-Milgram theory applies : $\|u(z)\|_V \leq C_0 = \frac{\|f\|_{V^*}}{r}$ for all $z \in \mathcal{U}$.

The function $u \mapsto u(z)$ is holomorphic in each variable z_j at any $z \in \mathcal{U}$: its first derivative $\partial_{z_j} u(z)$ is the unique solution to

$$\int_D a(z)\nabla \partial_{z_j} u(z) \cdot \nabla v = -\int_D \psi_j \nabla u(z) \cdot \nabla v, \quad v \in V.$$

Note that ∇ is with respect to spatial variable $x \in D$.

Extended domains of holomorphy : if $\rho = (\rho_j)_{j\geq 0}$ is any positive sequence such that for some $\delta > 0$

$$\sum_{j\geq 1} \rho_j |\psi_j(x)| \leq \bar{a}(x) - \delta, \quad x \in D,$$

then u is holomorphic with uniform bound $\|u(z)\| \leq C_\delta = \frac{\|f\|_{V^*}}{\delta}$ in the polydisc

$$\mathcal{U}_\rho := \otimes_{j\geq 1}\{|z_j| \leq \rho_j\},$$

If $\delta < r$, we can take $\rho_j > 1$.

Use Cauchy formula. In 1 complex variable if $z \mapsto u(z)$ is holomorphic and bounded in a neighbourhood of disc $\{|z| \leq b\}$, then for all z in this disc

$$u(z) = \frac{1}{2i\pi} \int_{|z'|=b} \frac{u(z')}{z - z'} dz',$$

which leads by n differentiation at $z = 0$ to $|u^{(n)}(0)| \leq n! b^{-n} \max_{|z| \leq b} |u(z)|$.

This yields exponential convergence rate $b^{-n} = \exp(-cn)$ of Taylor series for 1-d holomorphic functions. Curse of dimensionality : in d dimension, this yields sub-exponential rate $\exp(-cn^{1/d})$ where n is the number of retained terms.

Recursive application of this to all variables z_j such that $\nu_j \neq 0$, with $b = \rho_j$ gives

$$\|\partial^\nu u_{|z=0}\|_V \leq C_\delta \nu! \prod_{j \geq 1} \rho_j^{-\nu_j},$$

and thus

$$\|t_\nu\|_V \leq C_\delta \prod_{j \geq 1} \rho_j^{-\nu_j} = C_\delta \rho^{-\nu},$$

for any sequence $\rho = (\rho_j)_{j \geq 1}$ such that

$$\sum_{j \geq 1} \rho_j |\psi_j(x)| \leq \overline{a}(x) - \delta.$$

Estimate on the Taylor coefficients

Use Cauchy formula. In 1 complex variable if $z \mapsto u(z)$ is holomorphic and bounded in a neighbourhood of disc $\{|z| \leq b\}$, then for all z in this disc

$$u(z) = \frac{1}{2i\pi} \int_{|z'|=b} \frac{u(z')}{z - z'} dz',$$

which leads by n differentiation at $z = 0$ to $|u^{(n)}(0)| \leq n! b^{-n} \max_{|z| \leq b} |u(z)|$.

This yields exponential convergence rate $b^{-n} = \exp(-cn)$ of Taylor series for 1-d holomorphic functions. Curse of dimensionality : in d dimension, this yields sub-exponential rate $\exp(-cn^{1/d})$ where n is the number of retained terms.

Recursive application of this to all variables z_j such that $\nu_j \neq 0$, with $b = \rho_j$ gives

$$\|\partial^\nu u_{|z=0}\|_V \leq C_\delta \nu! \prod_{j \geq 1} \rho_j^{-\nu_j},$$

and thus

$$\|t_\nu\|_V \leq C_\delta \prod_{j \geq 1} \rho_j^{-\nu_j} = C_\delta \rho^{-\nu},$$

for any sequence $\rho = (\rho_j)_{j \geq 1}$ such that

$$\sum_{j \geq 1} \rho_j |\psi_j(x)| \leq \overline{a}(x) - \delta.$$

Use Cauchy formula. In 1 complex variable if $z \mapsto u(z)$ is holomorphic and bounded in a neighbourhood of disc $\{|z| \leq b\}$, then for all z in this disc

$$u(z) = \frac{1}{2i\pi} \int_{|z'|=b} \frac{u(z')}{z - z'} dz',$$

which leads by n differentiation at $z = 0$ to $|u^{(n)}(0)| \leq n! b^{-n} \max_{|z| \leq b} |u(z)|$.

This yields exponential convergence rate $b^{-n} = \exp(-cn)$ of Taylor series for 1-d holomorphic functions. Curse of dimensionality : in d dimension, this yields sub-exponential rate $\exp(-cn^{1/d})$ where n is the number of retained terms.

Recursive application of this to all variables z_j such that $\nu_j \neq 0$, with $b = \rho_j$ gives

$$\|\partial^\nu u_{|z=0}\|_V \leq C_\delta \nu! \prod_{j \geq 1} \rho_j^{-\nu_j},$$

and thus

$$\|t_\nu\|_V \leq C_\delta \prod_{j \geq 1} \rho_j^{-\nu_j} = C_\delta \rho^{-\nu},$$

for any sequence $\rho = (\rho_j)_{j \geq 1}$ such that

$$\sum_{j \geq 1} \rho_j |\psi_j(x)| \leq \bar{a}(x) - \delta.$$

Optimization

Since ρ is not fixed we have

$$\|t_\nu\|_V \leq C_\delta \inf\{\rho^{-\nu} \ : \ \rho \ \text{s.t.} \ \sum_{j\geq 1} \rho_j |\psi_j(x)| \leq \overline{a}(x) - \delta, \ \ x \in D\}.$$

We do not know the general solution to this problem, except in particular case, for example when the ψ_j have disjoint supports.

Instead design a particular choice $\rho = \rho(\nu)$ satisfying the constraint with $\delta = r/2$, for which we prove that

$$(\|\psi_j\|_{L^\infty})_{j\geq 1} \in \ell^p(\mathbb{N}) \Rightarrow (\rho(\nu)^{-\nu})_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F}),$$

therefore proving the main theorem.

Optimization

Since ρ is not fixed we have

$$\|t_\nu\|_V \leq C_\delta \inf\{\rho^{-\nu} \;:\; \rho \text{ s.t. } \sum_{j\geq 1} \rho_j |\psi_j(x)| \leq \overline{a}(x) - \delta, \;\; x \in D\}.$$

We do not know the general solution to this problem, except in particular case, for example when the ψ_j have disjoint supports.

Instead design a particular choice $\rho = \rho(\nu)$ satisfying the constraint with $\delta = r/2$, for which we prove that

$$(\|\psi_j\|_{L^\infty})_{j\geq 1} \in \ell^p(\mathbb{N}) \Rightarrow (\rho(\nu)^{-\nu})_{\nu \in \mathcal{F}} \in \ell^p(\mathcal{F}),$$

therefore proving the main theorem.

Optimization

Since ρ is not fixed we have

$$\|t_\nu\|_V \leq C_\delta \inf\{\rho^{-\nu} \ : \ \rho \ \text{s.t.} \ \sum_{j\geq 1} \rho_j |\psi_j(x)| \leq \overline{a}(x) - \delta, \ \ x \in D\}.$$

We do not know the general solution to this problem, except in particular case, for example when the ψ_j have disjoint supports.

Instead design a particular choice $\rho = \rho(\nu)$ satisfying the constraint with $\delta = r/2$, for which we prove that

$$(\|\psi_j\|_{L^\infty})_{j\geq 1} \in \ell^p(\mathbb{N}) \Rightarrow (\rho(\nu)^{-\nu})_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F}),$$

therefore proving the main theorem.

A simple case

Assume that the ψ_j have disjoint supports. Then we maximize separately the ρ_j so that

$$\sum_{j\geq 1} \rho_j |\psi_j(x)| \leq \overline{a}(x) - \frac{r}{2}, \quad x \in D,$$

which leads to

$$\rho_j := \min_{x\in D} \frac{\overline{a}(x) - \frac{r}{2}}{|\psi_j(x)|}.$$

We have, with $\delta = \frac{r}{2}$,

$$\|t_\nu\|_V \leq C_\delta \rho^{-\nu} = C_\delta b^\nu,$$

where $b = (b_j)$ and

$$b_j := \rho_j^{-1} = \max_{x\in D} \frac{|\psi_j(x)|}{\overline{a}(x) - \frac{r}{2}} \leq \frac{\|\psi_j\|_{L^\infty}}{R - \frac{r}{2}}.$$

Therefore $b \in \ell^p(\mathbb{N})$. From (UEA), we have $|\psi_j(x)| \leq \overline{a}(x) - r$ and thus $\|b\|_{\ell^\infty} < 1$. We finally observe that

$$b \in \ell^p(\mathbb{N}) \text{ and } \|b\|_{\ell^\infty} < 1 \Leftrightarrow (b^\nu)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F}).$$

Proof : factorize

$$\sum_{\nu\in\mathcal{F}} b^{p\nu} = \prod_{j\geq 1} \sum_{n\geq 0} b_j^{pn} = \prod_{j\geq 1} \frac{1}{1 - b_j^p}.$$

Assume that the ψ_j have disjoint supports. Then we maximize separately the ρ_j so that

$$\sum_{j\geq 1} \rho_j |\psi_j(x)| \leq \overline{a}(x) - \frac{r}{2}, \quad x \in D,$$

which leads to

$$\rho_j := \min_{x\in D} \frac{\overline{a}(x) - \frac{r}{2}}{|\psi_j(x)|}.$$

We have, with $\delta = \frac{r}{2}$,

$$\|t_\nu\|_V \leq C_\delta \rho^{-\nu} = C_\delta b^\nu,$$

where $b = (b_j)$ and

$$b_j := \rho_j^{-1} = \max_{x\in D} \frac{|\psi_j(x)|}{\overline{a}(x) - \frac{r}{2}} \leq \frac{\|\psi_j\|_{L^\infty}}{R - \frac{r}{2}}.$$

Therefore $b \in \ell^p(\mathbb{N})$. From (UEA), we have $|\psi_j(x)| \leq \overline{a}(x) - r$ and thus $\|b\|_{\ell^\infty} < 1$.
We finally observe that

$$b \in \ell^p(\mathbb{N}) \ \text{ and } \ \|b\|_{\ell^\infty} < 1 \Leftrightarrow (b^\nu)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F}).$$

Proof : factorize

$$\sum_{\nu\in\mathcal{F}} b^{p\nu} = \prod_{j\geq 1} \sum_{n\geq 0} b_j^{pn} = \prod_{j\geq 1} \frac{1}{1 - b_j^p}.$$

Assume that the ψ_j have disjoint supports. Then we maximize separately the ρ_j so that

$$\sum_{j\geq 1} \rho_j |\psi_j(x)| \leq \overline{a}(x) - \frac{r}{2}, \quad x \in D,$$

which leads to

$$\rho_j := \min_{x\in D} \frac{\overline{a}(x) - \frac{r}{2}}{|\psi_j(x)|}.$$

We have, with $\delta = \frac{r}{2}$,

$$\|t_\nu\|_V \leq C_\delta \rho^{-\nu} = C_\delta b^\nu,$$

where $b = (b_j)$ and

$$b_j := \rho_j^{-1} = \max_{x\in D} \frac{|\psi_j(x)|}{\overline{a}(x) - \frac{r}{2}} \leq \frac{\|\psi_j\|_{L^\infty}}{R - \frac{r}{2}}.$$

Therefore $b \in \ell^p(\mathbb{N})$. From (UEA), we have $|\psi_j(x)| \leq \overline{a}(x) - r$ and thus $\|b\|_{\ell^\infty} < 1$. We finally observe that

$$b \in \ell^p(\mathbb{N}) \text{ and } \|b\|_{\ell^\infty} < 1 \Leftrightarrow (b^\nu)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F}).$$

Proof : factorize

$$\sum_{\nu\in\mathcal{F}} b^{p\nu} = \prod_{j\geq 1} \sum_{n\geq 0} b_j^{pn} = \prod_{j\geq 1} \frac{1}{1 - b_j^p}.$$

One defect of the previous result is that it depends on the ψ_j only through $\|\psi_j\|_{L^\infty}$, without taking their support into account.Improved results can be obtained, without relying on complex variable, by better exploiting the specific structure of PDE.

Recursive formula for the Taylor coefficients : with $e_j = (0,\dots,0,1,0,\dots)$ the Kroeneker sequence of index j, the coefficient t_ν is solution to

$$\int_D \bar{a}\nabla t_\nu \nabla v = -\sum_{j:\ \nu_j\neq 0}\int_D \psi_j \nabla t_{\nu-e_j}\nabla v, \quad v\in V.$$

We introduce the quantities

$$d_\nu := \int_D \bar{a}|\nabla t_\nu|^2 \quad \text{and} \quad d_{\nu,j} := \int_D |\psi_j|\,|\nabla t_\nu|^2.$$

Recall that (UEA) implies that $\left\|\frac{\sum_{j\geq 1}|\psi_j|}{\bar{a}}\right\|_{L^\infty(D)} \leq \theta < 1$. In particular

$$\sum_{j\geq 1} d_{\nu,j} \leq \theta d_\nu.$$

We use here the equivalent norm $\|v\|_V^2 := \int_D \bar{a}|\nabla v|^2$.

Lemma : under (UEA), one has $\sum_{\nu\in\mathcal{F}} d_\nu = \sum_{\nu\in\mathcal{F}}\|t_\nu\|_V^2 < \infty$.

One defect of the previous result is that it depends on the ψ_j only through $\|\psi_j\|_{L^\infty}$, without taking their support into account.Improved results can be obtained, without relying on complex variable, by better exploiting the specific structure of PDE.

Recursive formula for the Taylor coefficients : with $e_j = (0, \dots, 0, 1, 0, \dots)$ the Kroeneker sequence of index j, the coefficient t_ν is solution to

$$\int_D \bar{a} \nabla t_\nu \nabla v = - \sum_{j:\, \nu_j \neq 0} \int_D \psi_j \nabla t_{\nu - e_j} \nabla v, \quad v \in V.$$

We introduce the quantities

$$d_\nu := \int_D \bar{a} |\nabla t_\nu|^2 \quad \text{and} \quad d_{\nu,j} := \int_D |\psi_j| \, |\nabla t_\nu|^2.$$

Recall that (UEA) implies that $\left\| \frac{\sum_{j \geq 1} |\psi_j|}{\bar{a}} \right\|_{L^\infty(D)} \leq \theta < 1$. In particular

$$\sum_{j \geq 1} d_{\nu,j} \leq \theta d_\nu.$$

We use here the equivalent norm $\|v\|_V^2 := \int_D \bar{a} |\nabla v|^2$.

Lemma : under (UEA), one has $\sum_{\nu \in \mathcal{F}} d_\nu = \sum_{\nu \in \mathcal{F}} \|t_\nu\|_V^2 < \infty$.

One defect of the previous result is that it depends on the ψ_j only through $\|\psi_j\|_{L^\infty}$, without taking their support into account.Improved results can be obtained, without relying on complex variable, by better exploiting the specific structure of PDE.

Recursive formula for the Taylor coefficients : with $e_j = (0,\dots,0,1,0,\dots)$ the Kroeneker sequence of index j, the coefficient t_ν is solution to

$$\int_D \bar{a}\nabla t_\nu \nabla v = -\sum_{j:\ \nu_j\neq 0}\int_D \psi_j \nabla t_{\nu-e_j}\nabla v, \quad v\in V.$$

We introduce the quantities

$$d_\nu := \int_D \bar{a}|\nabla t_\nu|^2 \quad \text{and} \quad d_{\nu,j} := \int_D |\psi_j|\,|\nabla t_\nu|^2.$$

Recall that (UEA) implies that $\left\| \frac{\sum_{j\geq 1}|\psi_j|}{\bar{a}} \right\|_{L^\infty(D)} \leq \theta < 1$. In particular

$$\sum_{j\geq 1} d_{\nu,j} \leq \theta d_\nu.$$

We use here the equivalent norm $\|v\|_V^2 := \int_D \bar{a}|\nabla v|^2$.

Lemma : under (UEA), one has $\sum_{\nu\in\mathcal{F}} d_\nu = \sum_{\nu\in\mathcal{F}} \|t_\nu\|_V^2 < \infty$.

Proof

Taking $v = t_\nu$ in the recursion gives

$$d_\nu = \int_D \overline{a} |\nabla t_\nu|^2 = - \sum_{j:\ \nu_j \neq 0} \int_D \psi_j \nabla t_{\nu - e_j} \nabla t_\nu .$$

Apply Young's inequality on the right side gives

$$d_\nu \leq \sum_{j:\ \nu_j \neq 0} \left(\frac{1}{2} \int_D |\psi_j| |\nabla t_\nu|^2 + \frac{1}{2} \int_D |\psi_j| |\nabla t_{\nu - e_j}|^2 \right) = \frac{1}{2} \sum_{j:\ \nu_j \neq 0} d_{\nu, j} + \frac{1}{2} \sum_{j:\ \nu_j \neq 0} d_{\nu - e_j, j} .$$

The first sum is bounded by θd_ν, therefore

$$\left(1 - \frac{\theta}{2}\right) d_\nu \leq \frac{1}{2} \sum_{j:\ \nu_j \neq 0} d_{\nu - e_j, j} .$$

Now summing over all $|\nu| = k$ gives

$$\left(1 - \frac{\theta}{2}\right) \sum_{|\nu| = k} d_\nu \leq \frac{1}{2} \sum_{|\nu| = k} \sum_{j:\ \nu_j \neq 0} d_{\nu - e_j, j} = \frac{1}{2} \sum_{|\nu| = k-1} \sum_{j \geq 1} d_{\nu, j} \leq \frac{\theta}{2} \sum_{|\nu| = k-1} d_\nu .$$

Therefore $\sum_{|\nu|=k} d_\nu \leq \kappa \sum_{|\nu|=k-1} d_\nu$ with $\kappa := \frac{\theta}{2-\theta} < 1$, and thus $\sum_{\nu \in \mathcal{F}} d_\nu < \infty$.

Proof

Taking $v = t_\nu$ in the recursion gives

$$d_\nu = \int_D \overline{a}|\nabla t_\nu|^2 = -\sum_{j:\ \nu_j \neq 0} \int_D \psi_j \nabla t_{\nu - e_j} \nabla t_\nu.$$

Apply Young's inequality on the right side gives

$$d_\nu \leq \sum_{j:\ \nu_j \neq 0} \left(\frac{1}{2} \int_D |\psi_j|\, |\nabla t_\nu|^2 + \frac{1}{2} \int_D |\psi_j|\, |\nabla t_{\nu - e_j}|^2 \right) = \frac{1}{2} \sum_{j:\ \nu_j \neq 0} d_{\nu,j} + \frac{1}{2} \sum_{j:\ \nu_j \neq 0} d_{\nu - e_j, j}.$$

The first sum is bounded by θd_ν, therefore

$$\left(1 - \frac{\theta}{2}\right) d_\nu \leq \frac{1}{2} \sum_{j:\ \nu_j \neq 0} d_{\nu - e_j, j}.$$

Now summing over all $|\nu| = k$ gives

$$\left(1 - \frac{\theta}{2}\right) \sum_{|\nu| = k} d_\nu \leq \frac{1}{2} \sum_{|\nu| = k} \sum_{j:\ \nu_j \neq 0} d_{\nu - e_j, j} = \frac{1}{2} \sum_{|\nu| = k-1} \sum_{j \geq 1} d_{\nu, j} \leq \frac{\theta}{2} \sum_{|\nu| = k-1} d_\nu.$$

Therefore $\sum_{|\nu| = k} d_\nu \leq \kappa \sum_{|\nu| = k-1} d_\nu$ with $\kappa := \frac{\theta}{2 - \theta} < 1$, and thus $\sum_{\nu \in \mathcal{F}} d_\nu < \infty$.

Rescaling

Now let $\rho = (\rho_j)_{j\geq 1}$ be any sequence with $\rho_j > 1$ such that $\sum_{j\geq 1} \rho_j |\psi_j| \leq \bar{a} - \delta$ for some $\delta > 0$, or equivalently such that $\left\| \frac{\sum_{j\geq 1} \rho_j |\psi_j|}{\bar{a}} \right\|_{L^\infty(D)} \leq \theta < 1$.

Considered the rescaled solution map $\tilde{u}(y) = u(\rho y)$ where $\rho y := (\rho_j y_j)_{j\geq 1}$ which is the solution of the same problem as u with ψ_j replaced by $\rho_j \psi_j$.

Since (UEA) holds for for these rescaled functions, the previous lemma shows that

$$\sum_{\nu\in\mathcal{F}} \|\tilde{t}_\nu\|_V^2 < \infty,$$

where

$$\tilde{t}_\nu := \frac{1}{\nu!}\partial^\nu \tilde{u}(0) = \frac{1}{\nu!}\rho^\nu \partial^\nu u(0) = \rho^\nu t_\nu.$$

This therefore gives the weighted ℓ^2 estimate

$$\sum_{\nu\in\mathcal{F}} (\rho^\nu \|t_\nu\|_V)^2 \leq C < \infty.$$

In particular, we retrieve the estimate $\|t_\nu\|_V \leq C\rho^{-\nu}$ that was obtained by the complex variable approach, however the above estimate is stronger.

Rescaling

Now let $\rho = (\rho_j)_{j\geq 1}$ be any sequence with $\rho_j > 1$ such that $\sum_{j\geq 1} \rho_j |\psi_j| \leq \overline{a} - \delta$ for some $\delta > 0$, or equivalently such that $\left\| \frac{\sum_{j\geq 1} \rho_j |\psi_j|}{\overline{a}} \right\|_{L^\infty(D)} \leq \theta < 1$.

Considered the rescaled solution map $\tilde{u}(y) = u(\rho y)$ where $\rho y := (\rho_j y_j)_{j\geq 1}$ which is the solution of the same problem as u with ψ_j replaced by $\rho_j \psi_j$.

Since (UEA) holds for for these rescaled functions, the previous lemma shows that

$$\sum_{\nu\in\mathcal{F}} \|\tilde{t}_\nu\|_V^2 < \infty,$$

where

$$\tilde{t}_\nu := \frac{1}{\nu!}\partial^\nu \tilde{u}(0) = \frac{1}{\nu!}\rho^\nu \partial^\nu u(0) = \rho^\nu t_\nu.$$

This therefore gives the weighted ℓ^2 estimate

$$\sum_{\nu\in\mathcal{F}} (\rho^\nu \|t_\nu\|_V)^2 \leq C < \infty.$$

In particular, we retrieve the estimate $\|t_\nu\|_V \leq C\rho^{-\nu}$ that was obtained by the complex variable approach, however the above estimate is stronger.

Rescaling

Now let $\rho = (\rho_j)_{j\geq 1}$ be any sequence with $\rho_j > 1$ such that $\sum_{j\geq 1} \rho_j |\psi_j| \leq \bar{a} - \delta$ for some $\delta > 0$, or equivalently such that $\left\| \frac{\sum_{j\geq 1} \rho_j |\psi_j|}{\bar{a}} \right\|_{L^\infty(D)} \leq \theta < 1$.

Considered the rescaled solution map $\tilde{u}(y) = u(\rho y)$ where $\rho y := (\rho_j y_j)_{j\geq 1}$ which is the solution of the same problem as u with ψ_j replaced by $\rho_j \psi_j$.

Since (UEA) holds for for these rescaled functions, the previous lemma shows that

$$\sum_{\nu\in\mathcal{F}} \|\tilde{t}_\nu\|_V^2 < \infty,$$

where

$$\tilde{t}_\nu := \frac{1}{\nu!} \partial^\nu \tilde{u}(0) = \frac{1}{\nu!} \rho^\nu \partial^\nu u(0) = \rho^\nu t_\nu.$$

This therefore gives the weighted ℓ^2 estimate

$$\sum_{\nu\in\mathcal{F}} (\rho^\nu \|t_\nu\|_V)^2 \leq C < \infty.$$

In particular, we retrieve the estimate $\|t_\nu\|_V \leq C\rho^{-\nu}$ that was obtained by the complex variable approach, however the above estimate is stronger.

Applying Hölder's inequality gives

$$\sum_{\nu\in\mathcal{F}} \|t_\nu\|_V^p \leq \Big(\sum_{\nu\in\mathcal{F}} (\rho^\nu \|t_\nu\|_V)^2\Big)^{p/2} \Big(\sum_{\nu\in\mathcal{F}} \rho^{-q\nu}\Big)^{1-p/2},$$

with $q = \frac{2p}{2-p} > p$, or equivalently $\frac{1}{q} = \frac{1}{p} - \frac{1}{2}$.

The sum in second factor is finite provided that $(\rho_j^{-1})_{j\geq 1} \in \ell^q$. Therefore, the following result holds.

Theorem (Bachmayr-Cohen-Migliorati, 2015) : Let p and q be such that $\frac{1}{q} = \frac{1}{p} - \frac{1}{2}$. Assume that there exists a sequence $\rho = (\rho_j)_{j\geq 1}$ of numbers larger than 1 such that

$$\sum_{j\geq 1} \rho_j |\psi_j| \leq \overline{a} - \delta,$$

for some $\delta > 0$ and

$$(\rho_j^{-1})_{j\geq 1} \in \ell^q.$$

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

The above conditions ensuring ℓ^p summability of $(\|t_\nu\|_V)_{\nu\in\mathcal{F}}$ are significantly weaker than those in the first summability theorem especially for locally supported ψ_j.

Applying Hölder's inequality gives

$$\sum_{\nu\in\mathcal{F}} \|t_\nu\|_V^p \leq \Big(\sum_{\nu\in\mathcal{F}} (\rho^\nu \|t_\nu\|_V)^2\Big)^{p/2} \Big(\sum_{\nu\in\mathcal{F}} \rho^{-q\nu}\Big)^{1-p/2},$$

with $q = \frac{2p}{2-p} > p$, or equivalently $\frac{1}{q} = \frac{1}{p} - \frac{1}{2}$.
The sum in second factor is finite provided that $(\rho_j^{-1})_{j\geq 1} \in \ell^q$. Therefore, the following result holds.

Theorem (Bachmayr-Cohen-Migliorati, 2015) : Let p and q be such that $\frac{1}{q} = \frac{1}{p} - \frac{1}{2}$. Assume that there exists a sequence $\rho = (\rho_j)_{j\geq 1}$ of numbers larger than 1 such that

$$\sum_{j\geq 1} \rho_j |\psi_j| \leq \bar{a} - \delta,$$

for some $\delta > 0$ and

$$(\rho_j^{-1})_{j\geq 1} \in \ell^q.$$

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

The above conditions ensuring ℓ^p summability of $(\|t_\nu\|_V)_{\nu\in\mathcal{F}}$ are significantly weaker than those in the first summability theorem especially for locally supported ψ_j.

Applying Hölder's inequality gives

$$\sum_{\nu\in\mathcal{F}} \|t_\nu\|_V^p \le \Big(\sum_{\nu\in\mathcal{F}} (\rho^\nu \|t_\nu\|_V)^2\Big)^{p/2} \Big(\sum_{\nu\in\mathcal{F}} \rho^{-q\nu}\Big)^{1-p/2},$$

with $q = \frac{2p}{2-p} > p$, or equivalently $\frac{1}{q} = \frac{1}{p} - \frac{1}{2}$.
The sum in second factor is finite provided that $(\rho_j^{-1})_{j\ge 1} \in \ell^q$. Therefore, the following result holds.

Theorem (Bachmayr-Cohen-Migliorati, 2015) : Let p and q be such that $\frac{1}{q} = \frac{1}{p} - \frac{1}{2}$. Assume that there exists a sequence $\rho = (\rho_j)_{j\ge 1}$ of numbers larger than 1 such that

$$\sum_{j\ge 1} \rho_j |\psi_j| \le \bar{a} - \delta,$$

for some $\delta > 0$ and

$$(\rho_j^{-1})_{j\ge 1} \in \ell^q.$$

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

The above conditions ensuring ℓ^p summability of $(\|t_\nu\|_V)_{\nu\in\mathcal{F}}$ are significantly weaker than those in the first summability theorem especially for locally supported ψ_j.

Disjoint supports

Assume that the ψ_j have disjoint supports.

Then with $\delta = \frac{r}{2}$, we choose

$$\rho_j := \min_{x \in D} \frac{\overline{a}(x) - \frac{r}{2}}{|\psi_j(x)|} > 1.$$

so that $\sum_{j \geq 1} \rho_j |\psi_j| \leq \overline{a} - \delta$ holds.

We have

$$b_j := \rho_j^{-1} = \frac{|\psi_j(x)|}{\overline{a}(x) - \frac{r}{2}} \leq \frac{\|\psi_j\|_{L^\infty}}{R - \frac{r}{2}}.$$

Thus in this case, our result gives for any $0 < q < \infty$,

$$(\|\psi_j\|_{L^\infty})_{j \geq 1} \in \ell^q(\mathbb{N}) \Rightarrow (\|t_\nu\|_V)_{\nu \in \mathcal{F}} \in \ell^p(\mathcal{F}),$$

with $\frac{1}{q} = \frac{1}{p} - \frac{1}{2}$.

Similar improved results if the ψ_j have supports with limited overlap, such as wavelets.

No improvement in the case of globally supported functions, such as typical KL bases.

Disjoint supports

Assume that the ψ_j have disjoint supports.

Then with $\delta = \frac{r}{2}$, we choose

$$\rho_j := \min_{x \in D} \frac{\overline{a}(x) - \frac{r}{2}}{|\psi_j(x)|} > 1.$$

so that $\sum_{j \geq 1} \rho_j |\psi_j| \leq \overline{a} - \delta$ holds.

We have

$$b_j := \rho_j^{-1} = \frac{|\psi_j(x)|}{\overline{a}(x) - \frac{r}{2}} \leq \frac{\|\psi_j\|_{L^\infty}}{R - \frac{r}{2}}.$$

Thus in this case, our result gives for any $0 < q < \infty$,

$$(\|\psi_j\|_{L^\infty})_{j \geq 1} \in \ell^q(\mathbb{N}) \Rightarrow (\|t_\nu\|_V)_{\nu \in \mathcal{F}} \in \ell^p(\mathcal{F}),$$

with $\frac{1}{q} = \frac{1}{p} - \frac{1}{2}$.

Similar improved results if the ψ_j have supports with limited overlap, such as wavelets.

No improvement in the case of globally supported functions, such as typical KL bases.

Disjoint supports

Assume that the ψ_j have disjoint supports.

Then with $\delta = \frac{r}{2}$, we choose

$$\rho_j := \min_{x \in D} \frac{\overline{a}(x) - \frac{r}{2}}{|\psi_j(x)|} > 1.$$

so that $\sum_{j \geq 1} \rho_j |\psi_j| \leq \overline{a} - \delta$ holds.

We have

$$b_j := \rho_j^{-1} = \frac{|\psi_j(x)|}{\overline{a}(x) - \frac{r}{2}} \leq \frac{\|\psi_j\|_{L^\infty}}{R - \frac{r}{2}}.$$

Thus in this case, our result gives for any $0 < q < \infty$,

$$(\|\psi_j\|_{L^\infty})_{j \geq 1} \in \ell^q(\mathbb{N}) \Rightarrow (\|t_\nu\|_V)_{\nu \in \mathcal{F}} \in \ell^p(\mathcal{F}),$$

with $\frac{1}{q} = \frac{1}{p} - \frac{1}{2}$.

Similar improved results if the ψ_j have supports with limited overlap, such as wavelets.

No improvement in the case of globally supported functions, such as typical KL bases.

Other models

Model 1 : same PDE but no affine dependence, e.g. $a(x,y)=\overline{a}(x)+(\sum_{j\geq 0} y_j\psi_j(x))^2$. Assuming that $\overline{a}(x)\geq r>0$ guarantees ellipticity uniformly over $y\in U$.

Model 2 : similar problems + non-linearities, e.g.

$$g(u)-\text{div}(a\nabla u)=f \quad \text{on} \quad D=D(y) \quad u_{|\partial D}=0,$$

with same assumptions on a and f. Well-posedness in $V=H^1_0(D)$ for all $f\in V'$ is ensured for certain nonlinearities, e.g. $g(u)=u^3$ of u^5 in dimension $m=3$ ($V\subset L^6$).

Model 3 : PDE's on domains with parametrized boundaries, e.g.

$$-\Delta v=f \quad \text{on} \quad D=D_y \quad u_{|\partial D}=0.$$

where the boundary of D_y is parametrized by y, e.g.

$$D_y:=\{(x_1,x_2)\in\mathbb{R}^2 \ : \ 0<x_1<1 \ \text{ and } \ 0<x_2<b(x_1,y)\},$$

where $b=b(x,y)=\overline{b}(x)+\sum_j y_j\psi_j(x)$ satisfies $0<r<b(x,y)<R$. We transport this problem on the reference domain $[0,1]^2$ and study

$$u(y):=v(y)\circ\phi_y, \quad \phi_y:[0,1]^2\to D_y, \quad \phi_y(x_1,x_2):=(x_1,x_2b(x_1,y)).$$

which satisfies a diffusion equation with coefficient $a=a(x,y)$ non-affine in y.

Other models

Model 1 : same PDE but no affine dependence, e.g. $a(x,y) = \overline{a}(x) + (\sum_{j\geq 0} y_j \psi_j(x))^2$. Assuming that $\overline{a}(x) \geq r > 0$ guarantees ellipticity uniformly over $y \in U$.

Model 2 : similar problems + non-linearities, e.g.

$$g(u) - \mathrm{div}(a\nabla u) = f \;\; \text{on} \;\; D = D(y) \;\; u_{|\partial D} = 0,$$

with same assumptions on a and f. Well-posedness in $V = H_0^1(D)$ for all $f \in V'$ is ensured for certain nonlinearities, e.g. $g(u) = u^3$ of u^5 in dimension $m = 3$ ($V \subset L^6$).

Model 3 : PDE's on domains with parametrized boundaries, e.g.

$$-\Delta v = f \;\; \text{on} \;\; D = D_y \;\; u_{|\partial D} = 0.$$

where the boundary of D_y is parametrized by y, e.g.

$$D_y := \{(x_1, x_2) \in \mathbb{R}^2 \;\; : \;\; 0 < x_1 < 1 \;\; \text{and} \;\; 0 < x_2 < b(x_1, y)\},$$

where $b = b(x,y) = \overline{b}(x) + \sum_j y_j \psi_j(x)$ satisfies $0 < r < b(x,y) < R$. We transport this problem on the reference domain $[0,1]^2$ and study

$$u(y) := v(y) \circ \phi_y, \quad \phi_y : [0,1]^2 \to D_y, \quad \phi_y(x_1, x_2) := (x_1, x_2 b(x_1, y)).$$

which satisfies a diffusion equation with coefficient $a = a(x,y)$ non-affine in y.

Model 1 : same PDE but no affine dependence, e.g. $a(x,y) = \overline{a}(x) + (\sum_{j\geq 0} y_j\psi_j(x))^2$. Assuming that $\overline{a}(x) \geq r > 0$ guarantees ellipticity uniformly over $y \in U$.

Model 2 : similar problems + non-linearities, e.g.

$$g(u) - \mathrm{div}(a\nabla u) = f \ \text{ on } \ D = D(y) \ \ u_{|\partial D} = 0,$$

with same assumptions on a and f. Well-posedness in $V = H_0^1(D)$ for all $f \in V'$ is ensured for certain nonlinearities, e.g. $g(u) = u^3$ of u^5 in dimension $m = 3$ ($V \subset L^6$).

Model 3 : PDE's on domains with parametrized boundaries, e.g.

$$-\Delta v = f \ \text{ on } \ D = D_y \ \ u_{|\partial D} = 0.$$

where the boundary of D_y is parametrized by y, e.g.

$$D_y := \{(x_1, x_2) \in \mathbb{R}^2 \ : \ 0 < x_1 < 1 \ \text{ and } \ 0 < x_2 < b(x_1, y)\},$$

where $b = b(x,y) = \overline{b}(x) + \sum_j y_j\psi_j(x)$ satisfies $0 < r < b(x,y) < R$. We transport this problem on the reference domain $[0,1]^2$ and study

$$u(y) := v(y) \circ \phi_y, \quad \phi_y : [0,1]^2 \longrightarrow D_y, \quad \phi_y(x_1, x_2) := (x_1, x_2 b(x_1, y)).$$

which satisfies a diffusion equation with coefficient $a = a(x,y)$ non-affine in y.

In contrast to our guiding example (which we refer to as model 0), bounded holomorphic extension is generally not feasible in a complex domain containing the polydisc $\mathcal{U} = \otimes_{j\geq 1}\{|z_j| \leq 1\}$. For this reason, Taylor series are not expected to converge.

Instead we consider the tensorized Legendre expansion

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

where $L_\nu(y) := \prod_{j\geq 1} L_{\nu_j}(y_j)$ and $(L_k)_{k\geq 0}$ are the Legendre polynomials normalized in $L^2\left([-1,1], \frac{dt}{2}\right)$.

Thus $(L_\nu)_{\nu\in\mathcal{F}}$ is an orthonormal basis for $L^2(U, V, \mu)$ where $\mu := \otimes_{j\geq 1}\frac{dy_j}{2}$ is the uniform probability measure and we have

$$v_\nu = \int_U u(y) L_\nu(y) d\mu(y).$$

We also consider the L^∞-normalized Legendre polynomials $P_k = (1+2k)^{-1/2} L_k$ and their tensorized version P_ν, so

$$u(y) = \sum_{\nu\in\mathcal{F}} w_\nu P_\nu(y),$$

where $w_\nu := \left(\prod_{j\geq 1}(1+\nu_j)^{1/2}\right) v_\nu$.

In contrast to our guiding example (which we refer to as model 0), bounded holomorphic extension is generally not feasible in a complex domain containing the polydisc $\mathcal{U} = \otimes_{j\geq 1}\{|z_j| \leq 1\}$. For this reason, Taylor series are not expected to converge.

Instead we consider the tensorized Legendre expansion

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

where $L_\nu(y) := \prod_{j\geq 1} L_{\nu_j}(y_j)$ and $(L_k)_{k\geq 0}$ are the Legendre polynomials normalized in $L^2\left([-1,1], \frac{dt}{2}\right)$.

Thus $(L_\nu)_{\nu\in\mathcal{F}}$ is an orthonormal basis for $L^2(U, V, \mu)$ where $\mu := \otimes_{j\geq 1}\frac{dy_j}{2}$ is the uniform probability measure and we have

$$v_\nu = \int_U u(y)L_\nu(y)d\mu(y).$$

We also consider the L^∞-normalized Legendre polynomials $P_k = (1+2k)^{-1/2}L_k$ and their tensorized version P_ν, so

$$u(y) = \sum_{\nu\in\mathcal{F}} w_\nu P_\nu(y),$$

where $w_\nu := \left(\prod_{j\geq 1}(1+\nu_j)^{1/2}\right)v_\nu$.

In contrast to our guiding example (which we refer to as model 0), bounded holomorphic extension is generally not feasible in a complex domain containing the polydisc $\mathcal{U} = \otimes_{j\geq 1}\{|z_j| \leq 1\}$. For this reason,Taylor series are not expected to converge.

Instead we consider the tensorized Legendre expansion

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

where $L_\nu(y) := \prod_{j\geq 1} L_{\nu_j}(y_j)$ and $(L_k)_{k\geq 0}$ are the Legendre polynomials normalized in $L^2\left([-1,1], \frac{dt}{2}\right)$.

Thus $(L_\nu)_{\nu\in\mathcal{F}}$ is an orthonormal basis for $L^2(U,V,\mu)$ where $\mu := \otimes_{j\geq 1}\frac{dy_j}{2}$ is the uniform probability measure and we have

$$v_\nu = \int_U u(y)L_\nu(y)d\mu(y).$$

We also consider the L^∞-normalized Legendre polynomials $P_k = (1+2k)^{-1/2}L_k$ and their tensorized version P_ν, so

$$u(y) = \sum_{\nu\in\mathcal{F}} w_\nu P_\nu(y),$$

where $w_\nu := \left(\prod_{j\geq 1}(1+\nu_j)^{1/2}\right)v_\nu$.

Main result

Theorem (Chkifa-Cohen-Schwab, 2013) : For models 0, 1, 2 and 3, and for any $p < 1$,

$$(\|\psi_j\|_X)_{j>0} \in \ell^p(\mathbb{N}) \Rightarrow (\|v_\nu\|_V)_{\nu\in\mathcal{F}} \text{ and } (\|w_\nu\|_V)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F}).$$

with $X = L^\infty$ for models 0, 1, 2, and $X = W^{1,\infty}$ for model 3.

By the same application of Stechkin's argument as for Taylor expansions, best n-term truncations for the L^∞ normalized expansion converge rate $\mathcal{O}(n^{-s})$ in $L^\infty(U, V)$ where $s = \frac{1}{p} - 1$.

Best n-term truncations for the L^2 normalized expansion converge rwith ate $\mathcal{O}(n^{-r})$ in $L^2 U, V, \mu)$ where $r = \frac{1}{p} - \frac{1}{2}$.

In the particular case of our guiding example, model 0, we can obtain improved summability results for Legendre expansions, similar to Taylor expansions.

Key ingredient in the proof of the above theorem : estimates of Legendre coefficients for holomorphic functions in a "small" complex neighbourhood of U.

Main result

Theorem (Chkifa-Cohen-Schwab, 2013) : For models 0, 1, 2 and 3, and for any $p < 1$,

$$(\|\psi_j\|_X)_{j>0} \in \ell^p(\mathbb{N}) \Rightarrow (\|v_\nu\|_V)_{\nu\in\mathcal{F}} \text{ and } (\|w_\nu\|_V)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F}).$$

with $X = L^\infty$ for models 0, 1, 2, and $X = W^{1,\infty}$ for model 3.

By the same application of Stechkin's argument as for Taylor expansions, best n-term truncations for the L^∞ normalized expansion converge rate $\mathcal{O}(n^{-s})$ in $L^\infty(U, V)$ where $s = \frac{1}{p} - 1$.

Best n-term truncations for the L^2 normalized expansion converge rwith ate $\mathcal{O}(n^{-r})$ in $L^2 U, V, \mu)$ where $r = \frac{1}{p} - \frac{1}{2}$.

In the particular case of our guiding example, model 0, we can obtain improved summability results for Legendre expansions, similar to Taylor expansions.

Key ingredient in the proof of the above theorem : estimates of Legendre coefficients for holomorphic functions in a "small" complex neighbourhood of U.

Main result

Theorem (Chkifa-Cohen-Schwab, 2013) : For models 0, 1, 2 and 3, and for any $p < 1$,

$$(\|\psi_j\|_X)_{j>0} \in \ell^p(\mathbb{N}) \Rightarrow (\|v_\nu\|_V)_{\nu\in\mathcal{F}} \text{ and } (\|w_\nu\|_V)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F}).$$

with $X = L^\infty$ for models 0, 1, 2, and $X = W^{1,\infty}$ for model 3.

By the same application of Stechkin's argument as for Taylor expansions, best n-term truncations for the L^∞ normalized expansion converge rate $\mathcal{O}(n^{-s})$ in $L^\infty(U, V)$ where $s = \frac{1}{p} - 1$.

Best n-term truncations for the L^2 normalized expansion converge rwith ate $\mathcal{O}(n^{-r})$ in $L^2 U, V, \mu)$ where $r = \frac{1}{p} - \frac{1}{2}$.

In the particular case of our guiding example, model 0, we can obtain improved summability results for Legendre expansions, similar to Taylor expansions.

Key ingredient in the proof of the above theorem : estimates of Legendre coefficients for holomorphic functions in a "small" complex neighbourhood of U.

Main result

Theorem (Chkifa-Cohen-Schwab, 2013) : For models 0, 1, 2 and 3, and for any $p < 1$,

$$(\|\psi_j\|_X)_{j>0} \in \ell^p(\mathbb{N}) \Rightarrow (\|v_\nu\|_V)_{\nu\in\mathcal{F}} \text{ and } (\|w_\nu\|_V)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F}).$$

with $X = L^\infty$ for models 0, 1, 2, and $X = W^{1,\infty}$ for model 3.

By the same application of Stechkin's argument as for Taylor expansions, best n-term truncations for the L^∞ normalized expansion converge rate $\mathcal{O}(n^{-s})$ in $L^\infty(U, V)$ where $s = \frac{1}{p} - 1$.

Best n-term truncations for the L^2 normalized expansion converge rwith ate $\mathcal{O}(n^{-r})$ in $L^2 U, V, \mu)$ where $r = \frac{1}{p} - \frac{1}{2}$.

In the particular case of our guiding example, model 0, we can obtain improved summability results for Legendre expansions, similar to Taylor expansions.

Key ingredient in the proof of the above theorem : estimates of Legendre coefficients for holomorphic functions in a “small” complex neighbourhood of U.

Taylor vs Legendre expansions

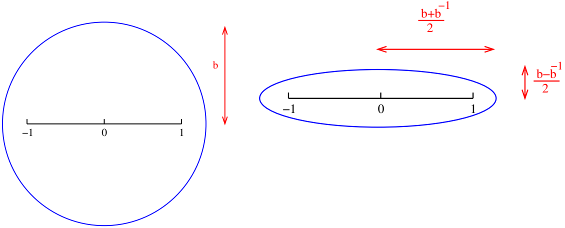
In one variable :

- If u is holomorphic in an open neighbourhood of the disc $\{|z| \leq b\}$ and bounded by M on this disc, then the n-th Taylor coefficient of u is bounded by

$$|t_n| := \left| \frac{u^{(n)}(0)}{n!} \right| \leq M b^{-n}$$

- If u is holomorphic in an open neighbourhood of the domain $\mathcal{E}_b$ limited by the ellipse of semi axes of length $(b + b^{-1})/2$ and $(b - b^{-1})/2$, for some $b > 1$, and bounded by M on this domain, then the n-th Legendre coefficent w_n of u is bounded by

$$|w_n| \leq M b^{-n}(1 + 2n)\phi(b), \qquad \phi(b) := \frac{\pi b}{b - 1}$$



A general assumption for sparsity of Legendre expansions

We say that the solution to a parametric PDE $\mathcal{D}(u,y)=0$ satisfies the (p,ε)-holomorphy property if and only if there exist a sequence $(c_j)_{j\geq 1}\in \ell^p(\mathbb{N})$, a constant $\varepsilon>0$ and $C_0>0$, such that : for any sequence $\rho=(\rho_j)_{j\geq 1}$ such that $\rho_j>1$ and

$$\sum_{j\geq 1}(\rho_j-1)c_j\leq \varepsilon,$$

the solution map has a complex extension

$$z\mapsto u(z),$$

of the solution map that is holomorphic with respect to each variable on a domain of the form $\mathcal{O}_\rho=\otimes_{j\geq 1}\mathcal{O}_{\rho_j}$ where $\mathcal{O}_{\rho_j}$ is an open neigbourhood of the elliptical domain $\mathcal{E}_{\rho_j}$, with bound

$$\sup_{z\in\mathcal{E}_\rho}\|u(z)\|_V\leq C_0,$$

where $\mathcal{E}_\rho=\otimes_{j\geq 1}\mathcal{E}_{\rho_j}$.

Under such an assumption, one has (up to additional harmless factors) an estimate of the form

$$\|w_\nu\|_V\leq C_0\inf\{\rho^{-\nu}\ ;\ \rho \text{ s.t. } \sum_{j\geq 1}(\rho_j-1)c_j\leq\varepsilon\},$$

allowing us to prove that $(\|w_\nu\|_V)_{\nu\in\mathcal{F}}\in\ell^p(\mathcal{F})$.

A general assumption for sparsity of Legendre expansions

We say that the solution to a parametric PDE $\mathcal{D}(u, y) = 0$ satisfies the (p, ε)-holomorphy property if and only if there exist a sequence $(c_j)_{j\geq 1} \in \ell^p(\mathbb{N})$, a constant $\varepsilon > 0$ and $C_0 > 0$, such that : for any sequence $\rho = (\rho_j)_{j\geq 1}$ such that $\rho_j > 1$ and

$$\sum_{j\geq 1} (\rho_j - 1) c_j \leq \varepsilon,$$

the solution map has a complex extension

$$z \mapsto u(z),$$

of the solution map that is holomorphic with respect to each variable on a domain of the form $\mathcal{O}_\rho = \otimes_{j\geq 1} \mathcal{O}_{\rho_j}$ where $\mathcal{O}_{\rho_j}$ is an open neigbourhood of the elliptical domain $\mathcal{E}_{\rho_j}$, with bound

$$\sup_{z \in \mathcal{E}_\rho} \|u(z)\|_V \leq C_0,$$

where $\mathcal{E}_\rho = \otimes_{j\geq 1} \mathcal{E}_{\rho_j}$.

Under such an assumption, one has (up to additional harmless factors) an estimate of the form

$$\|w_\nu\|_V \leq C_0 \inf\{\rho^{-\nu} \;\; ; \;\; \rho \;\; \text{s.t.} \;\; \sum_{j\geq 1} (\rho_j - 1) c_j \leq \varepsilon\},$$

allowing us to prove that $(\|w_\nu\|_V)_{\nu \in \mathcal{F}} \in \ell^p(\mathcal{F})$.

A general framework for establishing the (p, ε)-holomorphy assumption

Assume a general problem of the form

$$\mathcal{P}(u, a) = 0,$$

with $a = a(y) = \overline{a} + \sum_{j \geq 1} y_j \psi_j$, where

$$\mathcal{P} : V \times X \to W,$$

with V, X, W a triplet of complex Banach spaces, and $\overline{a}$ and ψ_j are functions in X.

Theorem (Chkifa-Cohen-Schwab, 2013) : assume that

(i) The problem is well posed for all $a \in Q = a(U)$ with solution $u(y) = u(a(y)) \in V$.

(ii) The map $\mathcal{P}$ is differentiable (holomorphic) from $X \times V$ to W.

(iii) For any $a \in Q$, the differential $\partial_u \mathcal{P}(u(a), a)$ is an isomorphism from V to W

(iv) One has $(\|\psi_j\|_X)_{j \geq 1}$ in $\ell^p(\mathbb{N})$ for some $0 < p < 1$,

Then, for $\varepsilon > 0$ small enough, the (p, ε)-holomorphy property holds.

A general framework for establishing the (p,ε)-holomorphy assumption

Assume a general problem of the form

$$\mathcal{P}(u,a)=0,$$

with $a=a(y)=\overline{a}+\sum_{j\geq 1} y_j\psi_j$, where

$$\mathcal{P}: V\times X\to W,$$

with V,X,W a triplet of complex Banach spaces, and $\overline{a}$ and ψ_j are functions in X.

Theorem (Chkifa-Cohen-Schwab, 2013) : assume that

(i) The problem is well posed for all $a\in Q=a(U)$ with solution $u(y)=u(a(y))\in V$.

(ii) The map $\mathcal{P}$ is differentiable (holomorphic) from $X\times V$ to W.

(iii) For any $a\in Q$, the differential $\partial_u\mathcal{P}(u(a),a)$ is an isomorphism from V to W

(iv) One has $(\|\psi_j\|_X)_{j\geq 1}$ in $\ell^p(\mathbb{N})$ for some $0<p<1$,

Then, for $\varepsilon>0$ small enough, the (p,ε)-holomorphy property holds.

Idea of proof

Based on the holomorphic Banach valued version of the implicit function theorem (see e.g. Dieudonné).

1. For any $a \in Q = \{a(y) \ : \ y \in U\}$ we can find a $\varepsilon_a > 0$ such that the map $a \longrightarrow u(a)$ has an holomorphic extension on the ball $B(a, \varepsilon_a) := \{\tilde{a} \in X \ : \ \|\tilde{a} - a\|_X < \varepsilon_a\}$.

2. Using the decay properties of the $\|\psi_j\|_X$, we find that Q is compact in X. It can be covered by a finite union of balls $B(a_i, \varepsilon_{a_i})$, for $i = 1, \dots, M$.

3. Thus $a \longrightarrow u(a)$ has an holomorphic extension on a complex neighbourhood $\mathcal{N}$ of Q of the form

$$\mathcal{N} = \cup_{i=1}^{M} B(a_i, \varepsilon_{a_i}).$$

4. For ε small enough, one proves that if $\sum_{j\geq 1}(\rho_j - 1)c_j \leq \varepsilon$ with $c_j := \|\psi_j\|_L$ then with $\mathcal{O}_\rho = \otimes_{j\geq 1}\mathcal{O}_{\rho_j}$ where $\mathcal{O}_b := \{z \in \mathbb{C} \ : \ \mathrm{dist}(z, [-1,1])_{\mathbb{C}} \leq b - 1\}$ is a neighborhood of $\mathcal{E}_b$, one has

$$z \in \mathcal{O}_\rho \Rightarrow a(z) \in \mathcal{N}.$$

This gives holomorphy of $z \mapsto a(z) \mapsto u(z) = u(a(z))$ in each variable for $z \in \mathcal{O}_\rho$.

Lognormal coefficients

We assume diffusion coefficients are given by

$$a = \exp(b),$$

with b a random function defined by an affine expansion of the form

$$b = b(y) = \sum_{j \geq 1} y_j \psi_j,$$

where (ψ_j) is a given family of functions from $L^\infty(D)$ and $y = (y_j)_{j \geq 1}$ a sequence of i.i.d. standard Gaussians $\mathcal{N}(0,1)$ variables.

Thus y ranges in $U = \mathbb{R}^{\mathbb{N}}$ equipped with the probabilistic structure $(U, \mathcal{B}(U), \gamma)$ where $\mathcal{B}(U)$ is the cylindrical Borel Σ-algebra and γ the tensorized Gaussian measure.

Commonly used stochastic model for diffusion in porous media.

The solution $u(y)$ is well defined in V for those $y \in U$ such that $b(y) \in L^\infty(D)$, with

$$\|u(y)\|_V \leq \frac{1}{a_{\min}(y)} \|f\|_{V'} \leq \exp(\|b(y)\|_{L^\infty}) \|f\|_{V'}.$$

Lognormal coefficients

We assume diffusion coefficients are given by

$$a = \exp(b),$$

with b a random function defined by an affine expansion of the form

$$b = b(y) = \sum_{j\geq 1} y_j \psi_j,$$

where (ψ_j) is a given family of functions from $L^\infty(D)$ and $y = (y_j)_{j\geq 1}$ a sequence of i.i.d. standard Gaussians $\mathcal{N}(0,1)$ variables.

Thus y ranges in $U = \mathbb{R}^{\mathbb{N}}$ equipped with the probabilistic structure $(U, \mathcal{B}(U), \gamma)$ where $\mathcal{B}(U)$ is the cylindrical Borel Σ-algebra and γ the tensorized Gaussian measure.

Commonly used stochastic model for diffusion in porous media.

The solution $u(y)$ is well defined in V for those $y \in U$ such that $b(y) \in L^\infty(D)$, with

$$\|u(y)\|_V \leq \frac{1}{a_{\min}(y)} \|f\|_{V'} \leq \exp(\|b(y)\|_{L^\infty}) \|f\|_{V'}.$$

Lognormal coefficients

We assume diffusion coefficients are given by

$$a = \exp(b),$$

with b a random function defined by an affine expansion of the form

$$b = b(y) = \sum_{j\geq 1} y_j \psi_j,$$

where (ψ_j) is a given family of functions from $L^\infty(D)$ and $y = (y_j)_{j\geq 1}$ a sequence of i.i.d. standard Gaussians $\mathcal{N}(0,1)$ variables.

Thus y ranges in $U = \mathbb{R}^{\mathbb{N}}$ equipped with the probabilistic structure $(U, \mathcal{B}(U), \gamma)$ where $\mathcal{B}(U)$ is the cylindrical Borel Σ-algebra and γ the tensorized Gaussian measure.

Commonly used stochastic model for diffusion in porous media.

The solution $u(y)$ is well defined in V for those $y \in U$ such that $b(y) \in L^\infty(D)$, with

$$\|u(y)\|_V \leq \frac{1}{a_{\min}(y)} \|f\|_{V'} \leq \exp(\|b(y)\|_{L^\infty}) \|f\|_{V'}.$$

Lognormal coefficients

We assume diffusion coefficients are given by

$$a = \exp(b),$$

with b a random function defined by an affine expansion of the form

$$b = b(y) = \sum_{j\geq 1} y_j \psi_j,$$

where (ψ_j) is a given family of functions from $L^\infty(D)$ and $y = (y_j)_{j\geq 1}$ a sequence of i.i.d. standard Gaussians $\mathcal{N}(0,1)$ variables.

Thus y ranges in $U = \mathbb{R}^{\mathbb{N}}$ equipped with the probabilistic structure $(U, \mathcal{B}(U), \gamma)$ where $\mathcal{B}(U)$ is the cylindrical Borel Σ-algebra and γ the tensorized Gaussian measure.

Commonly used stochastic model for diffusion in porous media.

The solution $u(y)$ is well defined in V for those $y \in U$ such that $b(y) \in L^\infty(D)$, with

$$\|u(y)\|_V \leq \frac{1}{a_{\min}(y)} \|f\|_{V'} \leq \exp(\|b(y)\|_{L^\infty}) \|f\|_{V'}.$$

Given a centered Gaussian process $(b(x))_{x\in D}$ with covariance function $C_b(x,z) = \mathbb{E}(b(x)b(z))$, one frequently consider the Karhunen-Loeve expansion,

$$b = \sum_{j\geq 1} \xi_j \varphi_j,$$

where ξ_j are i.i.d. $\mathcal{N}(0,\sigma_j^2)$ and $(\varphi_j)_{j\geq 1}$ are $L^2(D)$-orthonormal, and normalize

$$\psi_j = \sigma_j \varphi_j \quad \text{and} \quad y_j = \sigma_j^{-1}\xi_j,$$

so that $b = \sum_{j\geq 1} y_j \psi_j$. However, other representations may be relevant.

Example : b the Brownian bridge on $D = [0,1]$ defined by $C_b(x,z) := \min\{x,z\} - xz$.

1. Normalized KL : $\psi_j(x) = \frac{\sqrt{2}}{\pi j}\sin(\pi j x)$.

2. Levy-Ciesielski representation : uses Schauder basis (primitives of Haar system)

$$\psi_{l,k}(x) := 2^{-l/2}\psi(2^l x - k), \quad k = 0,\dots,2^l - 1, \quad l \geq 0, \quad \psi(x) := \frac{1}{2}(1 - |2x-1|)_+.$$

Then with coarse to fine ordering $\psi_j = \psi_{l,k}$ for $j = 2^l + k$, one has $b = \sum_{j\geq 1} y_j \psi_j$.

Affine Gaussian representations

Given a centered Gaussian process $(b(x))_{x\in D}$ with covariance function $C_b(x,z) = \mathbb{E}(b(x)b(z))$, one frequently consider the Karhunen-Loeve expansion,

$$b = \sum_{j\geq 1} \xi_j \varphi_j,$$

where ξ_j are i.i.d. $\mathcal{N}(0,\sigma_j^2)$ and $(\varphi_j)_{j\geq 1}$ are $L^2(D)$-orthonormal, and normalize

$$\psi_j = \sigma_j \varphi_j \quad \text{and} \quad y_j = \sigma_j^{-1}\xi_j,$$

so that $b = \sum_{j\geq 1} y_j \psi_j$. However, other representations may be relevant.

Example : b the Brownian bridge on $D = [0,1]$ defined by $C_b(x,z) := \min\{x,z\} - xz$.

1. Normalized KL : $\psi_j(x) = \frac{\sqrt{2}}{\pi j}\sin(\pi j x)$.

2. Levy-Ciesielski representation : uses Schauder basis (primitives of Haar system)

$$\psi_{l,k}(x) := 2^{-l/2}\psi(2^l x - k), \quad k = 0,\dots,2^l - 1, \quad l \geq 0, \quad \psi(x) := \frac{1}{2}(1 - |2x-1|)_+.$$

Then with coarse to fine ordering $\psi_j = \psi_{l,k}$ for $j = 2^l + k$, one has $b = \sum_{j\geq 1} y_j \psi_j$.

Given a centered Gaussian process $(b(x))_{x\in D}$ with covariance function $C_b(x,z) = \mathbb{E}(b(x)b(z))$, one frequently consider the Karhunen-Loeve expansion,

$$b = \sum_{j\geq 1} \xi_j \varphi_j,$$

where ξ_j are i.i.d. $\mathcal{N}(0, \sigma_j^2)$ and $(\varphi_j)_{j\geq 1}$ are $L^2(D)$-orthonormal, and normalize

$$\psi_j = \sigma_j \varphi_j \quad \text{and} \quad y_j = \sigma_j^{-1} \xi_j,$$

so that $b = \sum_{j\geq 1} y_j \psi_j$. However, other representations may be relevant.

Example : b the Brownian bridge on $D = [0,1]$ defined by $C_b(x,z) := \min\{x,z\} - xz$.

1. Normalized KL : $\psi_j(x) = \frac{\sqrt{2}}{\pi j} \sin(\pi j x)$.

2. Levy-Ciesielski representation : uses Schauder basis (primitives of Haar system)

$$\psi_{l,k}(x) := 2^{-l/2}\psi(2^l x - k), \quad k = 0,\dots,2^l - 1, \quad l \geq 0, \quad \psi(x) := \frac{1}{2}(1 - |2x-1|)_+.$$

Then with coarse to fine ordering $\psi_j = \psi_{l,k}$ for $j = 2^l + k$, one has $b = \sum_{j\geq 1} y_j \psi_j$.

Main theoretical questions

1. Integrability : under which conditions is $y \mapsto u(y)$ Bochner measurable with values in V and satifies for $0 \leq k < \infty$.

$$\|u\|^k_{L^k(U,V,\gamma)} = \mathbb{E}(\|u(y)\|_V^k) < \infty,$$

In view of $\|u(y)\|_V \leq \exp(\|b(y)\|_{L^\infty})\|f\|_{V'}$, this holds if $\mathbb{E}(\exp(k\|b(y)\|_{L^\infty}) < \infty$.

2. Approximability : if $u \in L^2(U, V, \gamma)$, consider the multivariate Hermite expansion

$$u = \sum_{\nu \in \mathcal{F}} u_\nu H_\nu, \quad H_\nu(y) := \prod_{j \geq 1} H_{\nu_j}(y_j) \quad \text{and} \quad u_\nu := \int_U u(y) H_\nu(y) d\gamma(y)$$

where $\mathcal{F}$ is the set of finitely supported integer sequences $\nu = (\nu_j)_{j \geq 1}$.

Best n-term approximation : $u_n = \sum_{\nu \in \Lambda_n} u_\nu H_\nu$, with Λ_n indices of n largest $\|u_\nu\|_V$.

Stechkin lemma : if $(\|u_\nu\|_V)_{\nu \in \mathcal{F}} \in \ell^p(\mathcal{F})$ for some $0 < p < 2$ then

$$\|u - u_n\|_{L^2(U,V,\gamma)} \leq C n^{-s}, \quad s := \frac{1}{p} - \frac{1}{2}, \quad C := \|(\|u_\nu\|_V)_{\nu \in \mathcal{F}}\|_{\ell^p}$$

Main theoretical questions

1. Integrability : under which conditions is $y \mapsto u(y)$ Bochner measurable with values in V and satifies for $0 \leq k < \infty$.

$$\|u\|^k_{L^k(U,V,\gamma)} = \mathbb{E}(\|u(y)\|^k_V) < \infty,$$

In view of $\|u(y)\|_V \leq \exp(\|b(y)\|_{L^\infty})\|f\|_{V'}$, this holds if $\mathbb{E}(\exp(k\|b(y)\|_{L^\infty}) < \infty$.

2. Approximability : if $u \in L^2(U, V, \gamma)$, consider the multivariate Hermite expansion

$$u = \sum_{\nu \in \mathcal{F}} u_\nu H_\nu, \quad H_\nu(y) := \prod_{j \geq 1} H_{\nu_j}(y_j) \quad \text{and} \quad u_\nu := \int_U u(y) H_\nu(y) d\gamma(y)$$

where $\mathcal{F}$ is the set of finitely supported integer sequences $\nu = (\nu_j)_{j \geq 1}$.

Best n-term approximation : $u_n = \sum_{\nu \in \Lambda_n} u_\nu H_\nu$, with Λ_n indices of n largest $\|u_\nu\|_V$.

Stechkin lemma : if $(\|u_\nu\|_V)_{\nu \in \mathcal{F}} \in \ell^p(\mathcal{F})$ for some $0 < p < 2$ then

$$\|u - u_n\|_{L^2(U,V,\gamma)} \leq Cn^{-s}, \quad s := \frac{1}{p} - \frac{1}{2}, \quad C := \|(\|u_\nu\|_V)_{\nu \in \mathcal{F}}\|_{\ell^p}$$

Main theoretical questions

1. Integrability : under which conditions is $y \mapsto u(y)$ Bochner measurable with values in V and satifies for $0 \le k < \infty$.

$$\|u\|^k_{L^k(U,V,\gamma)} = \mathbb{E}(\|u(y)\|^k_V) < \infty,$$

In view of $\|u(y)\|_V \le \exp(\|b(y)\|_{L^\infty})\|f\|_{V'}$, this holds if $\mathbb{E}(\exp(k\|b(y)\|_{L^\infty}) < \infty$.

2. Approximability : if $u \in L^2(U, V, \gamma)$, consider the multivariate Hermite expansion

$$u = \sum_{\nu\in\mathcal{F}} u_\nu H_\nu, \quad H_\nu(y) := \prod_{j\ge1} H_{\nu_j}(y_j) \quad \text{and} \quad u_\nu := \int_U u(y)H_\nu(y)d\gamma(y)$$

where $\mathcal{F}$ is the set of finitely supported integer sequences $\nu = (\nu_j)_{j\ge1}$.

Best n-term approximation : $u_n = \sum_{\nu\in\Lambda_n} u_\nu H_\nu$, with Λ_n indices of n largest $\|u_\nu\|_V$.

Stechkin lemma : if $(\|u_\nu\|_V)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F})$ for some $0 < p < 2$ then

$$\|u - u_n\|_{L^2(U,V,\gamma)} \le Cn^{-s}, \quad s := \frac{1}{p} - \frac{1}{2}, \quad C := \|(\|u_\nu\|_V)_{\nu\in\mathcal{F}}\|_{\ell^p}$$

1. Integrability : under which conditions is $y \mapsto u(y)$ Bochner measurable with values in V and satifies for $0 \leq k < \infty$.

$$\|u\|^k_{L^k(U,V,\gamma)} = \mathbb{E}(\|u(y)\|_V^k) < \infty,$$

In view of $\|u(y)\|_V \leq \exp(\|b(y)\|_{L^\infty})\|f\|_{V'}$, this holds if $\mathbb{E}(\exp(k\|b(y)\|_{L^\infty}) < \infty$.

2. Approximability : if $u \in L^2(U,V,\gamma)$, consider the multivariate Hermite expansion

$$u = \sum_{\nu\in\mathcal{F}} u_\nu H_\nu, \quad H_\nu(y) := \prod_{j\geq 1} H_{\nu_j}(y_j) \quad \text{and} \quad u_\nu := \int_U u(y)H_\nu(y)d\gamma(y)$$

where $\mathcal{F}$ is the set of finitely supported integer sequences $\nu = (\nu_j)_{j\geq 1}$.

Best n-term approximation : $u_n = \sum_{\nu\in\Lambda_n} u_\nu H_\nu$, with Λ_n indices of n largest $\|u_\nu\|_V$.

Stechkin lemma : if $(\|u_\nu\|_V)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F})$ for some $0 < p < 2$ then

$$\|u - u_n\|_{L^2(U,V,\gamma)} \leq Cn^{-s}, \quad s := \frac{1}{p} - \frac{1}{2}, \quad C := \|(\|u_\nu\|_V)_{\nu\in\mathcal{F}}\|_{\ell^p}$$

Main theoretical questions

1. Integrability : under which conditions is $y \mapsto u(y)$ Bochner measurable with values in V and satifies for $0 \leq k < \infty$.

$$\|u\|^k_{L^k(U,V,\gamma)} = \mathbb{E}(\|u(y)\|_V^k) < \infty,$$

In view of $\|u(y)\|_V \leq \exp(\|b(y)\|_{L^\infty})\|f\|_{V'}$, this holds if $\mathbb{E}(\exp(k\|b(y)\|_{L^\infty}) < \infty$.

2. Approximability : if $u \in L^2(U,V,\gamma)$, consider the multivariate Hermite expansion

$$u = \sum_{\nu\in\mathcal{F}} u_\nu H_\nu, \quad H_\nu(y) := \prod_{j\geq 1} H_{\nu_j}(y_j) \quad \text{and} \quad u_\nu := \int_U u(y)H_\nu(y)d\gamma(y)$$

where $\mathcal{F}$ is the set of finitely supported integer sequences $\nu = (\nu_j)_{j\geq 1}$.

Best n-term approximation : $u_n = \sum_{\nu\in\Lambda_n} u_\nu H_\nu$, with Λ_n indices of n largest $\|u_\nu\|_V$.

Stechkin lemma : if $(\|u_\nu\|_V)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F})$ for some $0 < p < 2$ then

$$\|u - u_n\|_{L^2(U,V,\gamma)} \leq Cn^{-s}, \quad s := \frac{1}{p} - \frac{1}{2}, \quad C := \|(\|u_\nu\|_V)_{\nu\in\mathcal{F}}\|_{\ell^p}$$

Existing results

Integrability : sufficient conditions for $u \in L^k(U, V, \gamma)$ for all $0 \le k < \infty$ are known.

1. Smoothness : $C_b \in C^\alpha(D \times D)$ for some $\alpha > 0$ (Charrier).
2. Summability : $\sum_{j \ge 1} \|\psi_j\|_{L^\infty} < \infty$ (Schwab-Gittelson-Hoang)
3. $\sum_{j \ge 1} \|\psi_j\|_{L^\infty}^{2-\delta} \|\psi_j\|_{C^\alpha}^{\delta} < \infty$ for some $0 < \delta < 1$ (Dashti-Stuart)

Approximability : first available result is as follows.

Theorem (Hoang-Schwab, 2014) : for any $0 < p \le 1$, if $(j\|\psi_j\|_{L^\infty}) \in \ell^p(\mathbb{N})$ then $(\|u_\nu\|_V) \in \ell^p(\mathcal{F})$.

Remarks :

The condition $(j\|\psi_j\|_{L^\infty}) \in \ell^p(\mathbb{N})$ is strong, compared to L^2-integrability conditions.

It typically imposes high order of smoothness of the covariance function.

For example it is not satified by KL or Schauder representation of Brownian bridge.

Condition based on $\|\psi_j\|_{L^\infty}$. Can we better exploit the support properties ?

Existing results

Integrability : sufficient conditions for $u \in L^k(U, V, \gamma)$ for all $0 \le k < \infty$ are known.

1. Smoothness : $C_b \in C^\alpha(D \times D)$ for some $\alpha > 0$ (Charrier).

2. Summability : $\sum_{j\ge1} \|\psi_j\|_{L^\infty} < \infty$ (Schwab-Gittelson-Hoang)

3. $\sum_{j\ge1} \|\psi_j\|_{L^\infty}^{2-\delta} \|\psi_j\|_{C^\alpha}^{\delta} < \infty$ for some $0 < \delta < 1$ (Dashti-Stuart)

Approximability : first available result is as follows.

Theorem (Hoang-Schwab, 2014) : for any $0 < p \le 1$, if $(j\|\psi_j\|_{L^\infty}) \in \ell^p(\mathbb{N})$ then $(\|u_\nu\|_V) \in \ell^p(\mathcal{F})$.

Remarks :

The condition $(j\|\psi_j\|_{L^\infty}) \in \ell^p(\mathbb{N})$ is strong, compared to L^2-integrability conditions.

It typically imposes high order of smoothness of the covariance function.

For example it is not satified by KL or Schauder representation of Brownian bridge.

Condition based on $\|\psi_j\|_{L^\infty}$. Can we better exploit the support properties ?

Existing results

Integrability : sufficient conditions for $u \in L^k(U, V, \gamma)$ for all $0 \le k < \infty$ are known.

1. Smoothness : $C_b \in C^\alpha(D \times D)$ for some $\alpha > 0$ (Charrier).
2. Summability : $\sum_{j \ge 1} \|\psi_j\|_{L^\infty} < \infty$ (Schwab-Gittelson-Hoang)
3. $\sum_{j \ge 1} \|\psi_j\|_{L^\infty}^{2-\delta} \|\psi_j\|_{C^\alpha}^{\delta} < \infty$ for some $0 < \delta < 1$ (Dashti-Stuart)

Approximability : first available result is as follows.

Theorem (Hoang-Schwab, 2014) : for any $0 < p \le 1$, if $(j\|\psi_j\|_{L^\infty}) \in \ell^p(\mathbb{N})$ then $(\|u_\nu\|_V) \in \ell^p(\mathcal{F})$.

Remarks :

The condition $(j\|\psi_j\|_{L^\infty}) \in \ell^p(\mathbb{N})$ is strong, compared to L^2-integrability conditions.

It typically imposes high order of smoothness of the covariance function.

For example it is not satified by KL or Schauder representation of Brownian bridge.

Condition based on $\|\psi_j\|_{L^\infty}$. Can we better exploit the support properties ?

Improved summability result

Theorem (Bachmayr-Cohen-DeVore-Migliorati, 2015) :

Let $0 < p < 2$ and define $q := q(p) = \frac{2p}{2-p} > p$ (or equivalently $\frac{1}{q} = \frac{1}{p} - \frac{1}{2}$).

Assume that there exists a positive sequence $\rho = (\rho_j)_{j\geq 1}$ such that

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

Then $y \mapsto u(y)$ is measurable and belongs $L^k(U, V, \gamma)$ for all $0 \leq k < \infty$ and

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

Remarks :

Similar result for the Taylor and Legendre coefficients for the affine parametric model $a(y) = \bar{a} + \sum_{j\geq 1} y_j \psi_j$ however by different arguments.

Proof is rather specific to the linear diffusion equation (yet extensions possible).

The above conditions for ℓ^p summability of $(\|u_\nu\|_V)_{\nu\in\mathcal{F}}$ are weaker than ℓ^p summability of $(j\|\psi_j\|_{L^\infty})_{j\geq 1}$ especially for locally supported ψ_j.

Theorem (Bachmayr-Cohen-DeVore-Migliorati, 2015) :

Let $0 < p < 2$ and define $q := q(p) = \frac{2p}{2-p} > p$ (or equivalently $\frac{1}{q} = \frac{1}{p} - \frac{1}{2}$).

Assume that there exists a positive sequence $\rho = (\rho_j)_{j\geq 1}$ such that

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

Then $y \mapsto u(y)$ is measurable and belongs $L^k(U, V, \gamma)$ for all $0 \leq k < \infty$ and

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

Remarks :

Similar result for the Taylor and Legendre coefficients for the affine parametric model $a(y) = \bar{a} + \sum_{j\geq 1} y_j\psi_j$ however by different arguments.

Proof is rather specific to the linear diffusion equation (yet extensions possible).

The above conditions for ℓ^p summability of $(\|u_\nu\|_V)_{\nu\in\mathcal{F}}$ are weaker than ℓ^p summability of $(j\|\psi_j\|_{L^\infty})_{j\geq 1}$ especially for locally supported ψ_j.

The case of the Brownian bridge

KL representation :

Globally supported functions $\psi_j(x) = \frac{\sqrt{2}}{\pi j}\sin(\pi j x)$.

The decay of $(\|\psi_j\|_{L^\infty})_{j\geq 1}$ is not sufficient to apply our results.

No provable approximability by best n-term Hermite series.

Schauder representation :

Wavelet type functions with decay in scale $\|\psi_\lambda\|_{L^\infty} \sim 2^{-l/2}$.

This allows to apply our result $\rho_\lambda = 2^{\beta l}$, for any $\beta < \frac{1}{2}$.

Our result imply that $(\|u_\nu\|_V)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F})$ for any p such that $\frac{1}{2} > \frac{1}{p} - \frac{1}{2}$.

In particular, best n-term Hermite approximations satisfy

$$\|u - u_{\Lambda_n}\|_{L^2(U,V,\gamma)} \leq Cn^{-s}, \quad s = \frac{1}{p} - \frac{1}{2} < \frac{1}{2}.$$

The case of the Brownian bridge

KL representation :

Globally supported functions $\psi_j(x) = \frac{\sqrt{2}}{\pi j}\sin(\pi j x)$.

The decay of $(\|\psi_j\|_{L^\infty})_{j\geq 1}$ is not sufficient to apply our results.

No provable approximability by best n-term Hermite series.

Schauder representation :

Wavelet type functions with decay in scale $\|\psi_\lambda\|_{L^\infty} \sim 2^{-l/2}$.

This allows to apply our result $\rho_\lambda = 2^{\beta l}$, for any $\beta < \frac{1}{2}$.

Our result imply that $(\|u_\nu\|_V)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F})$ for any p such that $\frac{1}{2} > \frac{1}{p} - \frac{1}{2}$.

In particular, best n-term Hermite approximations satisfy

$$\|u - u_{\Lambda_n}\|_{L^2(U,V,\gamma)} \leq Cn^{-s}, \quad s = \frac{1}{p} - \frac{1}{2} < \frac{1}{2}.$$

Main ingredient in the proof of the main result

1. Relate Hermite coefficients u_ν and partial derivatives $\partial^\mu u$. Base on 1-d Rodrigues formula : $H_n(t) = \frac{(-1)^n}{\sqrt{n!}} \frac{g^{(n)}(t)}{g(t)}$, where $g(t) := (2\pi)^{-1/2} \exp(-t^2/2)$. After some computation this leads to weighted ℓ^2 identity for any sequence $\rho := (\rho_j)_{j\geq 1}$.

$$\sum_{\|\mu\|_{\ell^\infty}\leq r} \frac{\rho^{2\mu}}{\mu!} \int_U \|\partial^\mu u(y)\|_V^2 \, d\gamma(y) = \sum_{\nu\in\mathcal{F}} b_\nu \|u_\nu\|_V^2,$$

where $b_\nu := \sum_{\|\mu\|_{\ell^\infty}\leq r} \binom{\nu}{\mu} \rho^{2\mu}$.

2. Prove finiteness of left hand side $\sum_{\|\mu\|_{\ell^\infty}\leq r} \frac{\rho^{2\mu}}{\mu!} \int_U \|\partial^\mu u(y)\|_V^2 \, d\gamma(y)$ when

$$\sup_{x\in D} \sum_{j\geq 1} \rho_j |\psi_j(x)| =: K < C_r := r^{-1/2} \ln 2.$$

Use PDE : $\int_D a(y)\nabla\partial^\mu u(y)\cdot\nabla v = -\sum_{\nu\leq\mu,\, \nu\neq\mu} \binom{\mu}{\nu} \int_D \psi^{\mu-\nu} a(y)\nabla\partial^\nu u(y)\cdot\nabla v$.

3. Derive ℓ^p estimate by mean of Hölder's inequality :

$$\Big(\sum_{\nu\in\mathcal{F}} \|u_\nu\|_V^p\Big)^{1/p} \leq \Big(\sum_{\nu\in\mathcal{F}} b_\nu \|u_\nu\|_V^2\Big)^{1/2} \Big(\sum_{\nu\in\mathcal{F}} b_\nu^{-q/2}\Big)^{1/q}.$$

We prove that the second factor is finite if $(\rho_j^{-1})_{j\geq 1} \in \ell^q(\mathbb{N})$ and r such that $\frac{2}{r+1} < p$.

Main ingredient in the proof of the main result

1. Relate Hermite coefficients u_ν and partial derivatives $\partial^\mu u$. Base on 1-d Rodrigues formula : $H_n(t) = \frac{(-1)^n}{\sqrt{n!}} \frac{g^{(n)}(t)}{g(t)}$, where $g(t) := (2\pi)^{-1/2} \exp(-t^2/2)$. After some computation this leads to weighted ℓ^2 identity for any sequence $\rho := (\rho_j)_{j \geq 1}$.

$$\sum_{\|\mu\|_{\ell^\infty} \leq r} \frac{\rho^{2\mu}}{\mu!} \int_U \|\partial^\mu u(y)\|_V^2 \, d\gamma(y) = \sum_{\nu \in \mathcal{F}} b_\nu \|u_\nu\|_V^2,$$

where $b_\nu := \sum_{\|\mu\|_{\ell^\infty} \leq r} \binom{\nu}{\mu} \rho^{2\mu}$.

2. Prove finiteness of left hand side $\sum_{\|\mu\|_{\ell^\infty} \leq r} \frac{\rho^{2\mu}}{\mu!} \int_U \|\partial^\mu u(y)\|_V^2 \, d\gamma(y)$ when

$$\sup_{x \in D} \sum_{j \geq 1} \rho_j |\psi_j(x)| =: K < C_r := r^{-1/2} \ln 2.$$

Use PDE : $\int_D a(y) \nabla \partial^\mu u(y) \cdot \nabla v = -\sum_{\nu \leq \mu, \, \nu \neq \mu} \binom{\mu}{\nu} \int_D \psi^{\mu-\nu} a(y) \nabla \partial^\nu u(y) \cdot \nabla v$.

3. Derive ℓ^p estimate by mean of Hölder's inequality :

$$\Big(\sum_{\nu \in \mathcal{F}} \|u_\nu\|_V^p\Big)^{1/p} \leq \Big(\sum_{\nu \in \mathcal{F}} b_\nu \|u_\nu\|_V^2\Big)^{1/2} \Big(\sum_{\nu \in \mathcal{F}} b_\nu^{-q/2}\Big)^{1/q}.$$

We prove that the second factor is finite if $(\rho_j^{-1})_{j \geq 1} \in \ell^q(\mathbb{N})$ and r such that $\frac{2}{r+1} < p$.

Main ingredient in the proof of the main result

1. Relate Hermite coefficients u_ν and partial derivatives $\partial^\mu u$. Base on 1-d Rodrigues formula : $H_n(t) = \frac{(-1)^n}{\sqrt{n!}} \frac{g^{(n)}(t)}{g(t)}$, where $g(t) := (2\pi)^{-1/2} \exp(-t^2/2)$. After some computation this leads to weighted ℓ^2 identity for any sequence $\rho := (\rho_j)_{j\geq 1}$.

$$\sum_{\|\mu\|_{\ell^\infty} \leq r} \frac{\rho^{2\mu}}{\mu!} \int_U \|\partial^\mu u(y)\|_V^2 \, d\gamma(y) = \sum_{\nu \in \mathcal{F}} b_\nu \|u_\nu\|_V^2,$$

where $b_\nu := \sum_{\|\mu\|_{\ell^\infty} \leq r} \binom{\nu}{\mu} \rho^{2\mu}$.

2. Prove finiteness of left hand side $\sum_{\|\mu\|_{\ell^\infty} \leq r} \frac{\rho^{2\mu}}{\mu!} \int_U \|\partial^\mu u(y)\|_V^2 \, d\gamma(y)$ when

$$\sup_{x \in D} \sum_{j \geq 1} \rho_j |\psi_j(x)| =: K < C_r := r^{-1/2} \ln 2.$$

Use PDE : $\int_D a(y) \nabla \partial^\mu u(y) \cdot \nabla v = -\sum_{\nu \leq \mu,\, \nu \neq \mu} \binom{\mu}{\nu} \int_D \psi^{\mu-\nu} a(y) \nabla \partial^\nu u(y) \cdot \nabla v$.

3. Derive ℓ^p estimate by mean of Hölder's inequality :

$$\Big(\sum_{\nu \in \mathcal{F}} \|u_\nu\|_V^p\Big)^{1/p} \leq \Big(\sum_{\nu \in \mathcal{F}} b_\nu \|u_\nu\|_V^2\Big)^{1/2} \Big(\sum_{\nu \in \mathcal{F}} b_\nu^{-q/2}\Big)^{1/q}.$$

We prove that the second factor is finite if $(\rho_j^{-1})_{j\geq 1} \in \ell^q(\mathbb{N})$ and r such that $\frac{2}{r+1} < p$.

From approximation results to numerical methods

The results so far are approximation results. They say that for several models of parametric PDEs, the solution map $y \mapsto u(y)$ can be accurately approximate (with rate n^{-s} for some $s > 0$) by multivariate polynomials having n terms.

These polynomials are computed by best n-term truncation of Taylor or Legendre or Hermite series, but this is not feasible in practical numercial methods.

Problem 1 : the best n-term index sets Λ_n are computationally out of reach. Their identification would require the knowledge of all coefficients in the expansion.

Objective : identify non-optimal yet good sets Λ_n.

Problem 2 : the exact polynomial coefficients t_ν (or v_ν, w_ν, u_ν) of u for the indices $\nu \in \Lambda_n$ cannot be computed exactly.

Objective : numerical strategy for approximately computing polynomial coefficients.

From approximation results to numerical methods

The results so far are approximation results. They say that for several models of parametric PDEs, the solution map $y \mapsto u(y)$ can be accurately approximate (with rate n^{-s} for some $s > 0$) by multivariate polynomials having n terms.

These polynomials are computed by best n-term truncation of Taylor or Legendre or Hermite series, but this is not feasible in practical numercial methods.

Problem 1 : the best n-term index sets Λ_n are computationally out of reach. Their identification would require the knowledge of all coefficients in the expansion.

Objective : identify non-optimal yet good sets Λ_n.

Problem 2 : the exact polynomial coefficients t_ν (or v_ν, w_ν, u_ν) of u for the indices $\nu \in \Lambda_n$ cannot be computed exactly.

Objective : numerical strategy for approximately computing polynomial coefficients.

From approximation results to numerical methods

The results so far are approximation results. They say that for several models of parametric PDEs, the solution map $y \mapsto u(y)$ can be accurately approximate (with rate n^{-s} for some $s > 0$) by multivariate polynomials having n terms.

These polynomials are computed by best n-term truncation of Taylor or Legendre or Hermite series, but this is not feasible in practical numercial methods.

Problem 1 : the best n-term index sets Λ_n are computationally out of reach. Their identification would require the knowledge of all coefficients in the expansion.

Objective : identify non-optimal yet good sets Λ_n.

Problem 2 : the exact polynomial coefficients t_ν (or v_ν, w_ν, u_ν) of u for the indices $\nu \in \Lambda_n$ cannot be computed exactly.

Objective : numerical strategy for approximately computing polynomial coefficients.

From approximation results to numerical methods

The results so far are approximation results. They say that for several models of parametric PDEs, the solution map $y \mapsto u(y)$ can be accurately approximate (with rate n^{-s} for some $s > 0$) by multivariate polynomials having n terms.

These polynomials are computed by best n-term truncation of Taylor or Legendre or Hermite series, but this is not feasible in practical numercial methods.

Problem 1 : the best n-term index sets Λ_n are computationally out of reach. Their identification would require the knowledge of all coefficients in the expansion.

Objective : identify non-optimal yet good sets Λ_n.

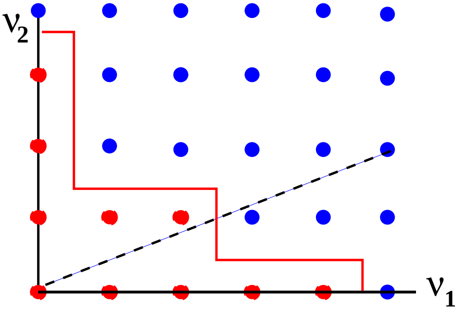
Problem 2 : the exact polynomial coefficients t_ν (or v_ν, w_ν, u_ν) of u for the indices $\nu \in \Lambda_n$ cannot be computed exactly.

Objective : numerical strategy for approximately computing polynomial coefficients.

Numerical methods : strategies to build the sets Λ_n

(i) Non-adaptive, based on the available a-priori estimates for the $\|t_\nu\|_V$ (or $\|v_\nu\|_V$, $\|w_\nu\|_V$, $\|u_\nu\|_V$). Take Λ_n to be the set corresponding to the n largest such estimates.

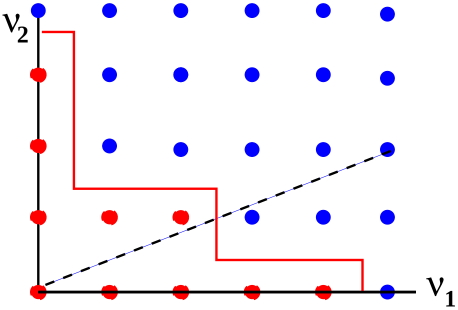
(ii) Adaptive, based on a-posteriori information gained in the computation $\Lambda_1 \subset \Lambda_2 \subset \cdots \subset \Lambda_n \cdots$.



Numerical methods : strategies to build the sets Λ_n

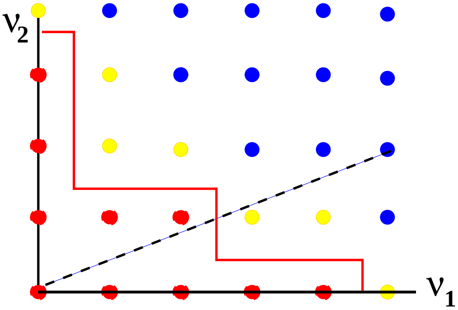
(i) Non-adaptive, based on the available a-priori estimates for the $\|t_\nu\|_V$ (or $\|v_\nu\|_V$, $\|w_\nu\|_V$, $\|u_\nu\|_V$). Take Λ_n to be the set corresponding to the n largest such estimates.

(ii) Adaptive, based on a-posteriori information gained in the computation $\Lambda_1 \subset \Lambda_2 \subset \cdots \subset \Lambda_n \cdots$.



(i) Non-adaptive, based on the available a-priori estimates for the $\|t_\nu\|_V$ (or $\|v_\nu\|_V$, $\|w_\nu\|_V$, $\|u_\nu\|_V$). Take Λ_n to be the set corresponding to the n largest such estimates.

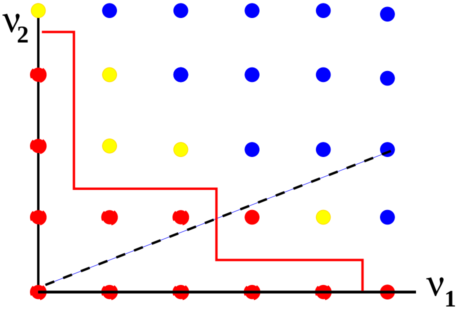
(ii) Adaptive, based on a-posteriori information gained in the computation $\Lambda_1 \subset \Lambda_2 \subset \cdots \subset \Lambda_n \cdots$.



Numerical methods : strategies to build the sets Λ_n

(i) Non-adaptive, based on the available a-priori estimates for the $\|t_\nu\|_V$ (or $\|v_\nu\|_V$, $\|w_\nu\|_V$, $\|u_\nu\|_V$). Take Λ_n to be the set corresponding to the n largest such estimates.

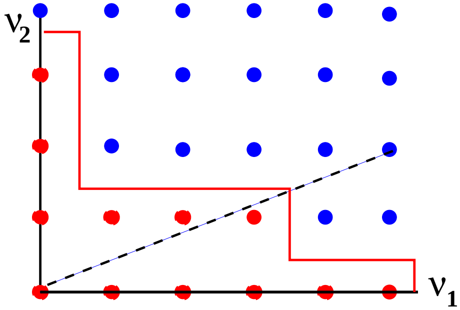
(ii) Adaptive, based on a-posteriori information gained in the computation $\Lambda_1 \subset \Lambda_2 \subset \cdots \subset \Lambda_n \cdots$.



Numerical methods : strategies to build the sets Λ_n

(i) Non-adaptive, based on the available a-priori estimates for the $\|t_\nu\|_V$ (or $\|v_\nu\|_V$, $\|w_\nu\|_V$, $\|u_\nu\|_V$). Take Λ_n to be the set corresponding to the n largest such estimates.

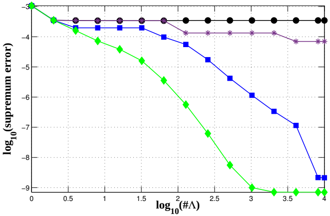
(ii) Adaptive, based on a-posteriori information gained in the computation $\Lambda_1 \subset \Lambda_2 \subset \cdots \subset \Lambda_n \cdots$.



Adaptive vs non-adaptive

Adaptive methods are known to converge better than non-adaptive ones, but their analysis is more difficult.

A test case for linear-affine model in dimension $d = 64$: comparison between the approximation performance with Λ_n given by standard choices $\{\sup \nu_j \leq k\}$ (black) or $\{\sum \nu_j \leq k\}$ (purple) and by anisotropic choices based on a-priori bounds (blue) or adaptively generated (green).



Highest polynomial degree for Λ_{1000} with different choices : 1, 2, 162 and 114.

Downward closed index sets

For adaptive algorithms it is critical that the index chosen sets are downward closed

$$\nu \in \Lambda \text{ and } \mu \leq \nu \Rightarrow \mu \in \Lambda,$$

where $\mu \leq \nu$ means that $\mu_j \leq \nu_j$ for all $j \geq 1$.

Such sets are also called lower sets. This property does not generally holds for the sets corresponding to the n largest estimates, however the same convergence rates as proved in the approximation theorems, can be proved when imposing such a structure.

If Λ is downward closed, we consider the polynomial space

$$\mathbb{P}_\Lambda = \text{span}\{y \to y^\nu \,:\, \nu \in \Lambda\} = \text{span}\{L_\nu \,:\, \nu \in \Lambda\} = \text{span}\{H_\nu \,:\, \nu \in \Lambda\}$$

and its V-valued version

$$V_\Lambda := \{\sum_{\nu \in \Lambda} v_\nu y^\nu \,:\, v_\nu \in V\} = V \otimes \mathbb{P}_\Lambda.$$

After having selected Λ_n we search for a computable approximation of u in V_{Λ_n}.

Note that $\dim(V_{\Lambda_n}) = \infty$. In practice we use $V_{\Lambda_n,h} = V_h \otimes \mathbb{P}_{\Lambda_n}$ which has dimension

$$\dim(V_{\Lambda_n,h}) = \dim(V_h)\dim(\mathbb{P}_{\Lambda_n}) = N_h n < \infty.$$

For adaptive algorithms it is critical that the index chosen sets are downward closed

$$\nu \in \Lambda \text{ and } \mu \leq \nu \Rightarrow \mu \in \Lambda,$$

where $\mu \leq \nu$ means that $\mu_j \leq \nu_j$ for all $j \geq 1$.

Such sets are also called lower sets. This property does not generally holds for the sets corresponding to the n largest estimates, however the same convergence rates as proved in the approximation theorems, can be proved when imposing such a structure.

If Λ is downward closed, we consider the polynomial space

$$\mathbb{P}_\Lambda = \text{span}\{y \to y^\nu : \nu \in \Lambda\} = \text{span}\{L_\nu : \nu \in \Lambda\} = \text{span}\{H_\nu : \nu \in \Lambda\}$$

and its V-valued version

$$V_\Lambda := \{ \sum_{\nu \in \Lambda} v_\nu y^\nu : v_\nu \in V\} = V \otimes \mathbb{P}_\Lambda.$$

After having selected Λ_n we search for a computable approximation of u in V_{Λ_n}.

Note that $\dim(V_{\Lambda_n}) = \infty$. In practice we use $V_{\Lambda_n,h} = V_h \otimes \mathbb{P}_{\Lambda_n}$ which has dimension

$$\dim(V_{\Lambda_n,h}) = \dim(V_h)\dim(\mathbb{P}_{\Lambda_n}) = N_h n < \infty.$$

Downward closed index sets

For adaptive algorithms it is critical that the index chosen sets are downward closed

$$\nu \in \Lambda \ \text{ and } \ \mu \leq \nu \Rightarrow \mu \in \Lambda,$$

where $\mu \leq \nu$ means that $\mu_j \leq \nu_j$ for all $j \geq 1$.

Such sets are also called lower sets. This property does not generally holds for the sets corresponding to the n largest estimates, however the same convergence rates as proved in the approximation theorems, can be proved when imposing such a structure.

If Λ is downward closed, we consider the polynomial space

$$\mathbb{P}_\Lambda = \mathrm{span}\{y \to y^\nu \ : \ \nu \in \Lambda\} = \mathrm{span}\{L_\nu \ : \ \nu \in \Lambda\} = \mathrm{span}\{H_\nu \ : \ \nu \in \Lambda\}$$

and its V-valued version

$$V_\Lambda := \{\sum_{\nu \in \Lambda} v_\nu y^\nu \ : \ v_\nu \in V\} = V \otimes \mathbb{P}_\Lambda.$$

After having selected Λ_n we search for a computable approximation of u in V_{Λ_n}.

Note that $\dim(V_{\Lambda_n}) = \infty$. In practice we use $V_{\Lambda_n,h} = V_h \otimes \mathbb{P}_{\Lambda_n}$ which has dimension

$$\dim(V_{\Lambda_n,h}) = \dim(V_h)\dim(\mathbb{P}_{\Lambda_n}) = N_h n < \infty.$$

For adaptive algorithms it is critical that the index chosen sets are downward closed

$$\nu \in \Lambda \text{ and } \mu \leq \nu \Rightarrow \mu \in \Lambda,$$

where $\mu \leq \nu$ means that $\mu_j \leq \nu_j$ for all $j \geq 1$.

Such sets are also called lower sets. This property does not generally holds for the sets corresponding to the n largest estimates, however the same convergence rates as proved in the approximation theorems, can be proved when imposing such a structure.

If Λ is downward closed, we consider the polynomial space

$$\mathbb{P}_\Lambda = \text{span}\{y \to y^\nu : \nu \in \Lambda\} = \text{span}\{L_\nu : \nu \in \Lambda\} = \text{span}\{H_\nu : \nu \in \Lambda\}$$

and its V-valued version

$$V_\Lambda := \{\sum_{\nu \in \Lambda} v_\nu y^\nu : v_\nu \in V\} = V \otimes \mathbb{P}_\Lambda.$$

After having selected Λ_n we search for a computable approximation of u in V_{Λ_n}.

Note that $\dim(V_{\Lambda_n}) = \infty$. In practice we use $V_{\Lambda_n,h} = V_h \otimes \mathbb{P}_{\Lambda_n}$ which has dimension

$$\dim(V_{\Lambda_n,h}) = \dim(V_h)\dim(\mathbb{P}_{\Lambda_n}) = N_h n < \infty.$$

For adaptive algorithms it is critical that the index chosen sets are downward closed

$$\nu \in \Lambda \ \text{and} \ \mu \leq \nu \Rightarrow \mu \in \Lambda,$$

where $\mu \leq \nu$ means that $\mu_j \leq \nu_j$ for all $j \geq 1$.

Such sets are also called lower sets. This property does not generally holds for the sets corresponding to the n largest estimates, however the same convergence rates as proved in the approximation theorems, can be proved when imposing such a structure.

If Λ is downward closed, we consider the polynomial space

$$\mathbb{P}_\Lambda = \text{span}\{y \to y^\nu \ : \ \nu \in \Lambda\} = \text{span}\{L_\nu \ : \ \nu \in \Lambda\} = \text{span}\{H_\nu \ : \ \nu \in \Lambda\}$$

and its V-valued version

$$V_\Lambda := \{ \sum_{\nu \in \Lambda} v_\nu y^\nu \ : \ v_\nu \in V\} = V \otimes \mathbb{P}_\Lambda.$$

After having selected Λ_n we search for a computable approximation of u in V_{Λ_n}.

Note that $\dim(V_{\Lambda_n}) = \infty$. In practice we use $V_{\Lambda_n,h} = V_h \otimes \mathbb{P}_{\Lambda_n}$ which has dimension

$$\dim(V_{\Lambda_n,h}) = \dim(V_h)\dim(\mathbb{P}_{\Lambda_n}) = N_h n < \infty.$$

1. Galerkin method : based on a space-parameter variational form (test the parametric PDE on arbitrary $y \mapsto v(y)$ and integrate both in x and y). Example for model 0 : find $u \in L^2(U, V, \mu)$ such that for all $v \in L^2(U, V, \mu)$,

$$A(u, v) := \int_U \int_D a(x, y) \nabla u(x, y) \nabla v(x, y) dx d\mu(y) = \int_U \langle f, v(y) \rangle d\mu(y) =: L(v),$$

The problem is coercive in $L^2(U, V, \mu)$. Galerkin formulation : find $u_n \in V_{\Lambda_n}$ such that

$$A(u_n, v_n) = L(v_n), \quad v_n \in V_{\Lambda_n}.$$

Cea's lemma gives error estimate

$$\|u - u_n\|_{L^2(U,V,\mu)} \leq (R/r)^{1/2} \min_{v \in V_{\Lambda_n}} \|u - v\|_{L^2(U,V,\mu)}.$$

After space discretization, Galerkin problem in $V_{\Lambda_n,h}$ gives a $(nN_h) \times (nN_h)$ system.

Lognormal case : lack of coervivity, Galerkin method needs some massaging.

2. Exact computation of the Taylor coefficients $\|t_\nu\|_V$, based on the recursive formula.

After space discretization, sequence of n systems of size $N_h \times N_h$.

Adaptive algorithms with optimal theoretical guarantees exist for both method 1 (Gittelson-Schwab) and 2 (Chkifa-Cohen-DeVore-Schwab).

These methods apply to other models, however mainly confined to linear PDEs, with affine parameter dependence.

1. Galerkin method : based on a space-parameter variational form (test the parametric PDE on arbitrary $y \mapsto v(y)$ and integrate both in x and y). Example for model 0 : find $u \in L^2(U, V, \mu)$ such that for all $v \in L^2(U, V, \mu)$,

$$A(u, v) := \int_U \int_D a(x, y) \nabla u(x, y) \nabla v(x, y) dx d\mu(y) = \int_U \langle f, v(y) \rangle d\mu(y) =: L(v),$$

The problem is coercive in $L^2(U, V, \mu)$. Galerkin formulation : find $u_n \in V_{\Lambda_n}$ such that

$$A(u_n, v_n) = L(v_n), \quad v_n \in V_{\Lambda_n}.$$

Cea's lemma gives error estimate

$$\|u - u_n\|_{L^2(U,V,\mu)} \leq (R/r)^{1/2} \min_{v \in V_{\Lambda_n}} \|u - v\|_{L^2(U,V,\mu)}.$$

After space discretization, Galerkin problem in $V_{\Lambda_n,h}$ gives a $(nN_h) \times (nN_h)$ system.

Lognormal case : lack of coervivity, Galerkin method needs some massaging.

2. Exact computation of the Taylor coefficients $\|t_\nu\|_V$, based on the recursive formula.

After space discretization, sequence of n systems of size $N_h \times N_h$.

Adaptive algorithms with optimal theoretical guarantees exist for both method 1 (Gittelson-Schwab) and 2 (Chkifa-Cohen-DeVore-Schwab).

These methods apply to other models, however mainly confined to linear PDEs, with affine parameter dependence.

1. Galerkin method : based on a space-parameter variational form (test the parametric PDE on arbitrary $y \mapsto v(y)$ and integrate both in x and y). Example for model 0 : find $u \in L^2(U, V, \mu)$ such that for all $v \in L^2(U, V, \mu)$,

$$A(u, v) := \int_U \int_D a(x, y) \nabla u(x, y) \nabla v(x, y) dx d\mu(y) = \int_U \langle f, v(y) \rangle d\mu(y) =: L(v),$$

The problem is coercive in $L^2(U, V, \mu)$. Galerkin formulation : find $u_n \in V_{\Lambda_n}$ such that

$$A(u_n, v_n) = L(v_n), \quad v_n \in V_{\Lambda_n}.$$

Cea's lemma gives error estimate

$$\|u - u_n\|_{L^2(U,V,\mu)} \le (R/r)^{1/2} \min_{v \in V_{\Lambda_n}} \|u - v\|_{L^2(U,V,\mu)}.$$

After space discretization, Galerkin problem in $V_{\Lambda_n,h}$ gives a $(nN_h) \times (nN_h)$ system.

Lognormal case : lack of coervivity, Galerkin method needs some massaging.

2. Exact computation of the Taylor coefficients $\|t_\nu\|_V$, based on the recursive formula.

After space discretization, sequence of n systems of size $N_h \times N_h$.

Adaptive algorithms with optimal theoretical guarantees exist for both method 1 (Gittelson-Schwab) and 2 (Chkifa-Cohen-DeVore-Schwab).

These methods apply to other models, however mainly confined to linear PDEs, with affine parameter dependence.

1. Galerkin method : based on a space-parameter variational form (test the parametric PDE on arbitrary $y \mapsto v(y)$ and integrate both in x and y). Example for model 0 : find $u \in L^2(U, V, \mu)$ such that for all $v \in L^2(U, V, \mu)$,

$$A(u, v) := \int_U \int_D a(x, y) \nabla u(x, y) \nabla v(x, y) dx d\mu(y) = \int_U \langle f, v(y) \rangle d\mu(y) =: L(v),$$

The problem is coercive in $L^2(U, V, \mu)$. Galerkin formulation : find $u_n \in V_{\Lambda_n}$ such that

$$A(u_n, v_n) = L(v_n), \quad v_n \in V_{\Lambda_n}.$$

Cea's lemma gives error estimate

$$\|u - u_n\|_{L^2(U,V,\mu)} \leq (R/r)^{1/2} \min_{v \in V_{\Lambda_n}} \|u - v\|_{L^2(U,V,\mu)}.$$

After space discretization, Galerkin problem in $V_{\Lambda_n,h}$ gives a $(nN_h) \times (nN_h)$ system.

Lognormal case : lack of coervivity, Galerkin method needs some massaging.

2. Exact computation of the Taylor coefficients $\|t_\nu\|_V$, based on the recursive formula.

After space discretization, sequence of n systems of size $N_h \times N_h$.

Adaptive algorithms with optimal theoretical guarantees exist for both method 1 (Gittelson-Schwab) and 2 (Chkifa-Cohen-DeVore-Schwab).

These methods apply to other models, however mainly confined to linear PDEs, with affine parameter dependence.

1. Galerkin method : based on a space-parameter variational form (test the parametric PDE on arbitrary $y \mapsto v(y)$ and integrate both in x and y). Example for model 0 : find $u \in L^2(U, V, \mu)$ such that for all $v \in L^2(U, V, \mu)$,

$$A(u, v) := \int_U \int_D a(x, y) \nabla u(x, y) \nabla v(x, y) dx d\mu(y) = \int_U \langle f, v(y) \rangle d\mu(y) =: L(v),$$

The problem is coercive in $L^2(U, V, \mu)$. Galerkin formulation : find $u_n \in V_{\Lambda_n}$ such that

$$A(u_n, v_n) = L(v_n), \quad v_n \in V_{\Lambda_n}.$$

Cea's lemma gives error estimate

$$\|u - u_n\|_{L^2(U,V,\mu)} \leq (R/r)^{1/2} \min_{v \in V_{\Lambda_n}} \|u - v\|_{L^2(U,V,\mu)}.$$

After space discretization, Galerkin problem in $V_{\Lambda_n, h}$ gives a $(nN_h) \times (nN_h)$ system.

Lognormal case : lack of coervivity, Galerkin method needs some massaging.

2. Exact computation of the Taylor coefficients $\|t_\nu\|_V$, based on the recursive formula.

After space discretization, sequence of n systems of size $N_h \times N_h$.

Adaptive algorithms with optimal theoretical guarantees exist for both method 1 (Gittelson-Schwab) and 2 (Chkifa-Cohen-DeVore-Schwab).

These methods apply to other models, however mainly confined to linear PDEs, with affine parameter dependence.

1. Galerkin method : based on a space-parameter variational form (test the parametric PDE on arbitrary $y \mapsto v(y)$ and integrate both in x and y). Example for model 0 : find $u \in L^2(U, V, \mu)$ such that for all $v \in L^2(U, V, \mu)$,

$$A(u, v) := \int_U \int_D a(x, y) \nabla u(x, y) \nabla v(x, y) dx d\mu(y) = \int_U \langle f, v(y) \rangle d\mu(y) =: L(v),$$

The problem is coercive in $L^2(U, V, \mu)$. Galerkin formulation : find $u_n \in V_{\Lambda_n}$ such that

$$A(u_n, v_n) = L(v_n), \quad v_n \in V_{\Lambda_n}.$$

Cea's lemma gives error estimate

$$\|u - u_n\|_{L^2(U,V,\mu)} \leq (R/r)^{1/2} \min_{v \in V_{\Lambda_n}} \|u - v\|_{L^2(U,V,\mu)}.$$

After space discretization, Galerkin problem in $V_{\Lambda_n,h}$ gives a $(nN_h) \times (nN_h)$ system.

Lognormal case : lack of coervivity, Galerkin method needs some massaging.

2. Exact computation of the Taylor coefficients $\|t_\nu\|_V$, based on the recursive formula.

After space discretization, sequence of n systems of size $N_h \times N_h$.

Adaptive algorithms with optimal theoretical guarantees exist for both method 1 (Gittelson-Schwab) and 2 (Chkifa-Cohen-DeVore-Schwab).

These methods apply to other models, however mainly confined to linear PDEs, with affine parameter dependence.

1. Galerkin method : based on a space-parameter variational form (test the parametric PDE on arbitrary $y \mapsto v(y)$ and integrate both in x and y). Example for model 0 : find $u \in L^2(U, V, \mu)$ such that for all $v \in L^2(U, V, \mu)$,

$$A(u, v) := \int_U \int_D a(x, y) \nabla u(x, y) \nabla v(x, y) dx d\mu(y) = \int_U \langle f, v(y) \rangle d\mu(y) =: L(v),$$

The problem is coercive in $L^2(U, V, \mu)$. Galerkin formulation : find $u_n \in V_{\Lambda_n}$ such that

$$A(u_n, v_n) = L(v_n), \quad v_n \in V_{\Lambda_n}.$$

Cea's lemma gives error estimate

$$\|u - u_n\|_{L^2(U,V,\mu)} \leq (R/r)^{1/2} \min_{v \in V_{\Lambda_n}} \|u - v\|_{L^2(U,V,\mu)}.$$

After space discretization, Galerkin problem in $V_{\Lambda_n,h}$ gives a $(nN_h) \times (nN_h)$ system.

Lognormal case : lack of coervivity, Galerkin method needs some massaging.

2. Exact computation of the Taylor coefficients $\|t_\nu\|_V$, based on the recursive formula.

After space discretization, sequence of n systems of size $N_h \times N_h$.

Adaptive algorithms with optimal theoretical guarantees exist for both method 1 (Gittelson-Schwab) and 2 (Chkifa-Cohen-DeVore-Schwab).

These methods apply to other models, however mainly confined to linear PDEs, with affine parameter dependence.

Based on snapshots $u(y^i)$ where $y^i \in U$ for $i = 1, \dots, m$..

1. Pseudo spectral methods : computation of $\sum_{\nu \in \Lambda_n} v_\nu L_\nu$ by quadrature

$$v_\nu = \int_U u(y) L_\nu(y) d\mu(y) \approx \sum_{i=1}^m w_i u(y^i) L_\nu(y^i).$$

2. Interpolation : with $m = n = \#(\Lambda_n) = \dim(\mathbb{P}_{\Lambda_n})$ search for a unique polynomial $u_n = I_{\Lambda_n} u \in V_{\Lambda_n}$ such that

$$u_n(y^i) = u(y^i), \quad i = 1, \dots, n.$$

3. Least-squares : with $m \geq n$, search for polynomial $u_n \in V_{\Lambda_n}$ minimizing

$$\sum_{i=1}^m \|u(y^i) - u_n(y^i)\|_V^2.$$

4. Underdetermined least-squares : with $m < n$ search for a polynomial $u_n \in V_{\Lambda_n}$ minimizing

$$\sum_{i=1}^m \|u(y^i) - u_n(y^i)\|_V^2 + \pi(u_n),$$

where π is a penalization functional. Compressed sensing : take for π the (weighted) ℓ^1 sum of V-norms of Legendre coefficients of u_n (promote sparse solutions).

Based on snapshots $u(y^i)$ where $y^i \in U$ for $i = 1, \dots, m$..

1. Pseudo spectral methods : computation of $\sum_{\nu \in \Lambda_n} v_\nu L_\nu$ by quadrature

$$v_\nu = \int_U u(y) L_\nu(y) d\mu(y) \approx \sum_{i=1}^{m} w_i u(y^i) L_\nu(y^i).$$

2. Interpolation : with $m = n = \#(\Lambda_n) = \dim(\mathbb{P}_{\Lambda_n})$ search for a unique polynomial $u_n = I_{\Lambda_n} u \in V_{\Lambda_n}$ such that

$$u_n(y^i) = u(y^i), \quad i = 1, \dots, n.$$

3. Least-squares : with $m \geq n$, search for polynomial $u_n \in V_{\Lambda_n}$ minimizing

$$\sum_{i=1}^{m} \|u(y^i) - u_n(y^i)\|_V^2.$$

4. Underdetermined least-squares : with $m < n$ search for a polynomial $u_n \in V_{\Lambda_n}$ minimizing

$$\sum_{i=1}^{m} \|u(y^i) - u_n(y^i)\|_V^2 + \pi(u_n),$$

where π is a penalization functional. Compressed sensing : take for π the (weighted) ℓ^1 sum of V-norms of Legendre coefficients of u_n (promote sparse solutions).

Based on snapshots $u(y^i)$ where $y^i \in U$ for $i = 1, \dots, m$..

1. Pseudo spectral methods : computation of $\sum_{\nu \in \Lambda_n} v_\nu L_\nu$ by quadrature

$$v_\nu = \int_U u(y) L_\nu(y) d\mu(y) \approx \sum_{i=1}^{m} w_i u(y^i) L_\nu(y^i).$$

2. Interpolation : with $m = n = \#(\Lambda_n) = \dim(\mathbb{P}_{\Lambda_n})$ search for a unique polynomial $u_n = I_{\Lambda_n} u \in V_{\Lambda_n}$ such that

$$u_n(y^i) = u(y^i), \quad i = 1, \dots, n.$$

3. Least-squares : with $m \geq n$, search for polynomial $u_n \in V_{\Lambda_n}$ minimizing

$$\sum_{i=1}^{m} \|u(y^i) - u_n(y^i)\|_V^2.$$

4. Underdetermined least-squares : with $m < n$ search for a polynomial $u_n \in V_{\Lambda_n}$ minimizing

$$\sum_{i=1}^{m} \|u(y^i) - u_n(y^i)\|_V^2 + \pi(u_n),$$

where π is a penalization functional. Compressed sensing : take for π the (weighted) ℓ^1 sum of V-norms of Legendre coefficients of u_n (promote sparse solutions).

Strategies to build the polynomial approximation : non-intrusive methods

Based on snapshots $u(y^i)$ where $y^i \in U$ for $i = 1, \dots, m$..

1. Pseudo spectral methods : computation of $\sum_{\nu \in \Lambda_n} v_\nu L_\nu$ by quadrature

$$v_\nu = \int_U u(y) L_\nu(y) d\mu(y) \approx \sum_{i=1}^m w_i u(y^i) L_\nu(y^i).$$

2. Interpolation : with $m = n = \#(\Lambda_n) = \dim(\mathbb{P}_{\Lambda_n})$ search for a unique polynomial $u_n = I_{\Lambda_n} u \in V_{\Lambda_n}$ such that

$$u_n(y^i) = u(y^i), \quad i = 1, \dots, n.$$

3. Least-squares : with $m \geq n$, search for polynomial $u_n \in V_{\Lambda_n}$ minimizing

$$\sum_{i=1}^m \|u(y^i) - u_n(y^i)\|_V^2.$$

4. Underdetermined least-squares : with $m < n$ search for a polynomial $u_n \in V_{\Lambda_n}$ minimizing

$$\sum_{i=1}^m \|u(y^i) - u_n(y^i)\|_V^2 + \pi(u_n),$$

where π is a penalization functional. Compressed sensing : take for π the (weighted) ℓ^1 sum of V-norms of Legendre coefficients of u_n (promote sparse solutions).

Advantages of non-intrusive methods

Applicable to a broad range of models, in particular non-linear PDEs.

Adaptive algorithms seem to work well for the interpolation and least squares approach, however with no theoretical guarantees.

Additional prescriptions for non-intrusive methods :

(i) Progressive : enrichment $\Lambda_n \to \Lambda_{n+1}$ requires only one or a few new snapshots.

(ii) Stable : moderate growth with n of the Lebesgue constant relative to the interpolation operator.

Main issue : how to best choose the point y^i ?

Advantages of non-intrusive methods

Applicable to a broad range of models, in particular non-linear PDEs.

Adaptive algorithms seem to work well for the interpolation and least squares approach, however with no theoretical guarantees.

Additional prescriptions for non-intrusive methods :

(i) Progressive : enrichment $\Lambda_n \to \Lambda_{n+1}$ requires only one or a few new snapshots.

(ii) Stable : moderate growth with n of the Lebesgue constant relative to the interpolation operator.

Main issue : how to best choose the point y^i ?

Advantages of non-intrusive methods

Applicable to a broad range of models, in particular non-linear PDEs.

Adaptive algorithms seem to work well for the interpolation and least squares approach, however with no theoretical guarantees.

Additional prescriptions for non-intrusive methods :

(i) Progressive : enrichment $\Lambda_n \longrightarrow \Lambda_{n+1}$ requires only one or a few new snapshots.

(ii) Stable : moderate growth with n of the Lebesgue constant relative to the interpolation operator.

Main issue : how to best choose the point y^i ?

Conclusions

The curse of dimensionality can be "defeated" by exploiting both smoothness and anisotropy in the different variables.

For certain models, this can be achieved by sparse polynomial approximations.

The way we parametrize the problem, or represent its solution, is crucial.

Adaptive algorithms with optimal theoretical guarantees are still to be developed, in particular for non-intrusive approaches (interpolation, collocation, least-squares).

Postdoc opportunities at Laboratoire Jacques-Louis Lions 2016/2017

ERC project BREAD (Breaking the curse of dimensionality in analysis and simulation)

Topics : High dimensional approximation, parametric and stochastic PDEs, reduced modelling, data assimilation, inverse problems.

Contact : cohen@ann.jussieu.fr

Conclusions

The curse of dimensionality can be "defeated" by exploiting both smoothness and anisotropy in the different variables.

For certain models, this can be achieved by sparse polynomial approximations.

The way we parametrize the problem, or represent its solution, is crucial.

Adaptive algorithms with optimal theoretical guarantees are still to be developed, in particular for non-intrusive approaches (interpolation, collocation, least-squares).

Postdoc opportunities at Laboratoire Jacques-Louis Lions 2016/2017

ERC project BREAD (Breaking the curse of dimensionality in analysis and simulation)

Topics : High dimensional approximation, parametric and stochastic PDEs, reduced modelling, data assimilation, inverse problems.

Contact : cohen@ann.jussieu.fr