High dimensional sparse polynomial approximations
of parametric and stochastic PDE’s

Albert Cohen
Laboratoire Jacques-Louis Lions
Université Pierre et Marie Curie
Paris

with Ronald DeVore and Christoph Schwab
numerical results by Abdellah Chkifa

Banff, 2011
The curse of dimensionality

Consider a continuous function $y \mapsto u(y)$ with $y \in [0, 1]$. Sample at equispaced points. Reconstruct, for example by piecewise linear interpolation.

Error in terms of point spacing $h > 0$: if $u$ has $C^2$ smoothness

$$\|u - R(u)\|_{L^\infty} \leq C\|u''\|_{L^\infty} h^2.$$ 

Using piecewise polynomials of higher order, if $u$ has $C^m$ smoothness

$$\|u - R(u)\|_{L^\infty} \leq C\|u^{(m)}\|_{L^\infty} h^m.$$ 

In terms of the number of samples $N \sim h^{-1}$, the error is estimated by $N^{-m}$.

In $d$ dimensions: $u(y) = u(y_1, \ldots, y_d)$ with $y \in [0,1]^d$. With a uniform sampling, we still have

$$\|u - R(u)\|_{L^\infty} \leq C\|d^m u\|_{L^\infty} h^m,$$

but the number of samples is now $N \sim h^{-d}$, and the error estimate is in $N^{-m/d}$.
The curse of dimensionality

Consider a continuous function $y \mapsto u(y)$ with $y \in [0, 1]$. Sample at equispaced points. Reconstruct, for example by piecewise linear interpolation.

Error in terms of point spacing $h > 0$: if $u$ has $C^2$ smoothness

$$\|u - R(u)\|_{L^\infty} \leq C\|u''\|_{L^\infty} h^2.$$  

Using piecewise polynomials of higher order, if $u$ has $C^m$ smoothness

$$\|u - R(u)\|_{L^\infty} \leq C\|u^{(m)}\|_{L^\infty} h^m.$$  

In terms of the number of samples $N \sim h^{-1}$, the error is estimated by $N^{-m}$.

In $d$ dimensions: $u(y) = u(y_1, \cdots, y_d)$ with $y \in [0, 1]^d$. With a uniform sampling, we still have

$$\|u - R(u)\|_{L^\infty} \leq C\|d^m u\|_{L^\infty} h^m,$$

but the number of samples is now $N \sim h^{-d}$, and the error estimate is in $N^{-m/d}$. 
The curse of dimensionality

Consider a continuous function \( y \mapsto u(y) \) with \( y \in [0, 1] \). Sample at equispaced points. Reconstruct, for example by piecewise linear interpolation.

Error in terms of point spacing \( h > 0 \) : if \( u \) has \( C^2 \) smoothness
\[
\|u - R(u)\|_{L^\infty} \leq C\|u''\|_{L^\infty} h^2.
\]

Using piecewise polynomials of higher order, if \( u \) has \( C^m \) smoothness
\[
\|u - R(u)\|_{L^\infty} \leq C\|u^{(m)}\|_{L^\infty} h^m.
\]

In terms of the number of samples \( N \sim h^{-1} \), the error is estimated by \( N^{-m} \).

In \( d \) dimensions : \( u(y) = u(y_1, \ldots, y_d) \) with \( y \in [0, 1]^d \). With a uniform sampling, we still have
\[
\|u - R(u)\|_{L^\infty} \leq C\|d^m u\|_{L^\infty} h^m,
\]
but the number of samples is now \( N \sim h^{-d} \), and the error estimate is in \( N^{-m/d} \).
The curse of dimensionality

Consider a continuous function $y \mapsto u(y)$ with $y \in [0, 1]$. Sample at equispaced points. Reconstruct, for example by piecewise linear interpolation.

Error in terms of point spacing $h > 0$ : if $u$ has $C^2$ smoothness

$$
\|u - R(u)\|_\infty \leq C \|u''\|_\infty h^2.
$$

Using piecewise polynomials of higher order, if $u$ has $C^m$ smoothness

$$
\|u - R(u)\|_\infty \leq C \|u^{(m)}\|_\infty h^m.
$$

In terms of the number of samples $N \sim h^{-1}$, the error is estimated by $N^{-m}$.

In $d$ dimensions : $u(y) = u(y_1, \cdots, y_d)$ with $y \in [0, 1]^d$. With a uniform sampling, we still have

$$
\|u - R(u)\|_\infty \leq C \|d^m u\|_\infty h^m,
$$

but the number of samples is now $N \sim h^{-d}$, and the error estimate is in $N^{-m/d}$.
The curse of dimensionality

Consider a continuous function \( y \mapsto u(y) \) with \( y \in [0, 1] \). Sample at equispaced points. Reconstruct, for example by piecewise linear interpolation.

Error in terms of point spacing \( h > 0 \) : if \( u \) has \( C^2 \) smoothness

\[
\|u - R(u)\|_{L^\infty} \leq C\|u''\|_{L^\infty} h^2.
\]

Using piecewise polynomials of higher order, if \( u \) has \( C^m \) smoothness

\[
\|u - R(u)\|_{L^\infty} \leq C\|u^{(m)}\|_{L^\infty} h^m.
\]

In terms of the number of samples \( N \sim h^{-1} \), the error is estimated by \( N^{-m} \).

In \( d \) dimensions : \( u(y) = u(y_1, \cdots, y_d) \) with \( y \in [0, 1]^d \). With a uniform sampling, we still have

\[
\|u - R(u)\|_{L^\infty} \leq C\|d^m u\|_{L^\infty} h^m,
\]

but the number of samples is now \( N \sim h^{-d} \), and the error estimate is in \( N^{-m/d} \).
The curse of dimensionality

Consider a continuous function $y \mapsto u(y)$ with $y \in [0, 1]$. Sample at equispaced points. Reconstruct, for example by piecewise linear interpolation.

Error in terms of point spacing $h > 0$: if $u$ has $C^2$ smoothness

$$\|u - R(u)\|_{L^\infty} \leq C \|u''\|_{L^\infty} h^2.$$  

Using piecewise polynomials of higher order, if $u$ has $C^m$ smoothness

$$\|u - R(u)\|_{L^\infty} \leq C \|u^{(m)}\|_{L^\infty} h^m.$$  

In terms of the number of samples $N \sim h^{-1}$, the error is estimated by $N^{-m}$.

In $d$ dimensions: $u(y) = u(y_1, \cdots, y_d)$ with $y \in [0, 1]^d$. With a uniform sampling, we still have

$$\|u - R(u)\|_{L^\infty} \leq C \|d^m u\|_{L^\infty} h^m,$$

but the number of samples is now $N \sim h^{-d}$, and the error estimate is in $N^{-m/d}$. 


The curse of dimensionality

Consider a continuous function \( y \mapsto u(y) \) with \( y \in [0, 1] \). Sample at equispaced points. Reconstruct, for example by piecewise linear interpolation.

Error in terms of point spacing \( h > 0 \) : if \( u \) has \( C^2 \) smoothness

\[ \| u - R(u) \|_{L^\infty} \leq C \| u'' \|_{L^\infty} h^2. \]

Using piecewise polynomials of higher order, if \( u \) has \( C^m \) smoothness

\[ \| u - R(u) \|_{L^\infty} \leq C \| u^{(m)} \|_{L^\infty} h^m. \]

In terms of the number of samples \( N \sim h^{-1} \), the error is estimated by \( N^{-m} \).

In \( d \) dimensions : \( u(y) = u(y_1, \cdots, y_d) \) with \( y \in [0, 1]^d \). With a uniform sampling, we still have

\[ \| u - R(u) \|_{L^\infty} \leq C \| d^m u \|_{L^\infty} h^m, \]

but the number of samples is now \( N \sim h^{-d} \), and the error estimate is in \( N^{-m/d} \).
Other sampling/reconstruction methods cannot do better!

Can be explained by nonlinear manifold width (DeVore-Howard-Micchelli).

Let $X$ be a normed space and $K \subset X$ a compact set.

Consider maps $E : K \mapsto \mathbb{R}^N$ (encoding) and $R : \mathbb{R}^N \mapsto X$ (reconstruction).

Introducing the distortion of the pair $(E, R)$ over $K$

$$\max_{u \in K} \| u - R(E(u)) \|_X,$$

we define the nonlinear $N$-width of $K$ as

$$d_N(K) := \inf_{E, R} \max_{u \in K} \| u - R(E(u)) \|_X,$$

where the infimum is taken over all continuous maps $(E, R)$.

If $X = L^\infty$ and $K$ is the unit ball of $C^m([0,1]^d)$ it is known that

$$cN^{-m/d} \leq d_N(K) \leq CN^{-m/d}.$$
Other sampling/reconstruction methods cannot do better!

Can be explained by nonlinear manifold width (DeVore-Howard-Micchelli).

Let $X$ be a normed space and $K \subset X$ a compact set.

Consider maps $E : K \mapsto \mathbb{R}^N$ (encoding) and $R : \mathbb{R}^N \mapsto X$ (reconstruction).

Introducing the distorsion of the pair $(E, R)$ over $K$

$$\max_{u \in K} \| u - R(E(u)) \|_X,$$

we define the nonlinear $N$-width of $K$ as

$$d_N(K) := \inf_{E, R} \max_{u \in K} \| u - R(E(u)) \|_X,$$

where the infimum is taken over all continuous maps $(E, R)$.

If $X = L^\infty$ and $K$ is the unit ball of $C^m([0,1]^d)$ it is known that

$$cN^{-m/d} \leq d_N(K) \leq CN^{-m/d}.$$
Other sampling/reconstruction methods cannot do better!

Can be explained by nonlinear manifold width (DeVore-Howard-Micchelli).

Let $X$ be a normed space and $K \subset X$ a compact set.

Consider maps $E : K \mapsto \mathbb{R}^N$ (encoding) and $R : \mathbb{R}^N \mapsto X$ (reconstruction).

Introducing the distortion of the pair $(E, R)$ over $K$

$$\max_{u \in K} \|u - R(E(u))\|_X,$$

we define the nonlinear $N$-width of $K$ as

$$d_N(K) := \inf_{E, R} \max_{u \in K} \|u - R(E(u))\|_X,$$

where the infimum is taken over all continuous maps $(E, R)$.

If $X = L^\infty$ and $K$ is the unit ball of $C^m([0,1]^d)$ it is known that

$$cN^{-m/d} \leq d_N(K) \leq CN^{-m/d}.$$
Other sampling/reconstruction methods cannot do better!

Can be explained by **nonlinear manifold width** (DeVore-Howard-Micchelli).

Let $X$ be a normed space and $\mathcal{K} \subset X$ a compact set.

Consider maps $E : \mathcal{K} \mapsto \mathbb{R}^N$ (encoding) and $R : \mathbb{R}^N \mapsto X$ (reconstruction).

Introducing the distortion of the pair $(E, R)$ over $\mathcal{K}$

$$\max_{u \in \mathcal{K}} \| u - R(E(u)) \|_X,$$

we define the nonlinear $N$-width of $\mathcal{K}$ as

$$d_N(\mathcal{K}) := \inf_{E,R} \max_{u \in \mathcal{K}} \| u - R(E(u)) \|_X,$$

where the infimum is taken over all continuous maps $(E, R)$.

If $X = L^\infty$ and $\mathcal{K}$ is the unit ball of $C^m([0,1]^d)$ it is known that

$$cN^{-m/d} \leq d_N(\mathcal{K}) \leq CN^{-m/d}.$$
Other sampling/reconstruction methods cannot do better!

Can be explained by **nonlinear manifold width** (DeVore-Howard-Micchelli).

Let $X$ be a normed space and $\mathcal{K} \subset X$ a compact set.

Consider maps $E : \mathcal{K} \mapsto \mathbb{R}^N$ (encoding) and $R : \mathbb{R}^N \mapsto X$ (reconstruction).

Introducing the distortion of the pair $(E, R)$ over $\mathcal{K}$

$$\max_{u \in \mathcal{K}} \| u - R(E(u)) \|_X,$$

we define the nonlinear $N$-width of $\mathcal{K}$ as

$$d_N(\mathcal{K}) := \inf_{E, R} \max_{u \in \mathcal{K}} \| u - R(E(u)) \|_X,$$

where the infimum is taken over all **continuous** maps $(E, R)$.

If $X = L^\infty$ and $\mathcal{K}$ is the unit ball of $C^m([0, 1]^d)$ it is known that

$$cN^{-m/d} \leq d_N(\mathcal{K}) \leq CN^{-m/d}.$$
High dimensional problems occur frequently

PDE’s with solutions \( u(x, v, t) \) defined in phase space: \( d = 7 \).

Post-processing of numerical codes: \( u \) solver with input parameters \( (y_1, \cdots, y_d) \).

Learning theory: \( u \) regression function of input parameters \( (y_1, \cdots, y_d) \).

In these applications \( d \) may be of the order up to \( 10^3 \).

Approximation of stochastic-parametric PDEs (this talk): \( d = +\infty \).

Smoothness properties of functions should be revisited by other means than \( C^m \) classes, and appropriate approximation tools should be used.

Key ingredients:

(i) Sparsity

(ii) Variable reduction

(iii) Anisotropy
High dimensional problems occur frequently

PDE’s with solutions $u(x, v, t)$ defined in phase space: $d = 7$.

Post-processing of numerical codes: $u$ solver with input parameters $(y_1, \cdots, y_d)$.

Learning theory: $u$ regression function of input parameters $(y_1, \cdots, y_d)$.

In these applications $d$ may be of the order up to $10^3$.

Approximation of stochastic-parametric PDEs (this talk): $d = +\infty$.

Smoothness properties of functions should be revisited by other means than $C^m$ classes, and appropriate approximation tools should be used.

Key ingredients:

(i) Sparsity

(ii) Variable reduction

(iii) Anisotropy
**High dimensional problems occur frequently**

PDE’s with solutions $u(x, v, t)$ defined in phase space: $d = 7$.

Post-processing of numerical codes: $u$ solver with input parameters $(y_1, \cdots, y_d)$.

Learning theory: $u$ regression function of input parameters $(y_1, \cdots, y_d)$.

In these applications $d$ may be of the order up to $10^3$.

Approximation of stochastic-parametric PDEs (this talk): $d = +\infty$.

Smoothness properties of functions should be revisited by other means than $C^m$ classes, and appropriate approximation tools should be used.

Key ingredients:

(i) Sparsity

(ii) Variable reduction

(iii) Anisotropy
High dimensional problems occur frequently

PDE’s with solutions $u(x,v,t)$ defined in phase space: $d = 7$.

Post-processing of numerical codes: $u$ solver with input parameters $(y_1, \ldots, y_d)$.

Learning theory: $u$ regression function of input parameters $(y_1, \ldots, y_d)$.

In these applications $d$ may be of the order up to $10^3$.

Approximation of stochastic-parametric PDEs (this talk): $d = +\infty$.

Smoothness properties of functions should be revisited by other means than $C^m$ classes, and appropriate approximation tools should be used.

Key ingredients:

(i) Sparsity

(ii) Variable reduction

(iii) Anisotropy
We consider the steady state diffusion equation

\[-\text{div}(a \nabla u) = f \text{ in } D \subset \mathbb{R}^m \text{ and } u = 0 \text{ on } \partial D,\]

where \( f = f(x) \in L^2(D) \) and \( a = a(x, y) \) are variable coefficients depending on \( x \in D \) and on a vector \( y \) of parameters in an affine manner:

\[ a = a(x, y) = \overline{a}(x) + \sum_{j>0} y_j \psi_j(x), \quad x \in D, y = (y_j)_{j>0} \in U := [-1, 1]^N, \]

where \((\psi_j)_{j>0}\) is a given family of functions.

The parameters may be deterministic (control, optimization) or random (uncertainty modeling and propagation, reliability assessment).

Uniform ellipticity assumption:

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

Then \( u : y \mapsto u(y) = u(\cdot, y) \) is a bounded map from \( U \) to \( V := H^1_0(\Omega) : \)

\[ \|u(y)\|_V \leq C_0 := \frac{\|f\|_{V^*}}{r}, \quad y \in U, \text{ where } \|v\|_V := \|\nabla v\|_{L^2}. \]

Proof: multiply equation by \( u \) and integrate

\[ r\|u\|^2_V \leq \int_D a \nabla u \cdot \nabla u = -\int_D u \text{div}(a \nabla u) = \int_D uf \leq \|u\|_V \|f\|_{V^*}. \]

Objective: build a computable approximation to this map at reasonable cost, i.e. simultaneaously approximate \( u(y) \) for all \( y \in U. \)
A model elliptic PDE

We consider the steady state diffusion equation

\[-\text{div}(a \nabla u) = f \quad \text{in} \quad D \subset \mathbb{R}^m \quad \text{and} \quad u = 0 \quad \text{on} \quad \partial D,\]

where \( f = f(x) \in L^2(D) \) and \( a = a(x, y) \) are variable coefficients depending on \( x \in D \) and on a vector \( y \) of parameters in an affine manner:

\[a = a(x, y) = \bar{a}(x) + \sum_{j>0} y_j \psi_j(x), \quad x \in D, \quad y = (y_j)_{j>0} \in U := [-1, 1]^\mathbb{N},\]

where \((\psi_j)_{j>0}\) is a given family of functions.

The parameters may be deterministic (control, optimization) or random (uncertainty modeling and propagation, reliability assessment).

Uniform ellipticity assumption:

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

Then \( u : y \mapsto u(y) = u(\cdot, y) \) is a bounded map from \( U \) to \( V := H^1_0(\Omega) :\)

\[\|u(y)\|_V \leq C_0 := \frac{\|f\|_{V^*}}{r}, \quad y \in U, \quad \text{where} \quad \|v\|_V := \|\nabla v\|_{L^2}.\]

Proof: multiply equation by \( u \) and integrate

\[r\|u\|^2_V \leq \int_D a \nabla u \cdot \nabla u = -\int_D u \text{div}(a \nabla u) = \int_D uf \leq \|u\|_V \|f\|_{V^*}.\]

Objective: build a computable approximation to this map at reasonable cost, i.e. simultaneously approximate \( u(y) \) for all \( y \in U \).
A model elliptic PDE

We consider the steady state diffusion equation
\[-\text{div}(a\nabla u) = f \quad \text{in} \quad D \subset \mathbb{R}^m \quad \text{and} \quad u = 0 \quad \text{on} \quad \partial D,\]
where \( f = f(x) \in L^2(D) \) and \( a = a(x, y) \) are variable coefficients depending on \( x \in D \) and on a vector \( y \) of parameters in an affine manner:
\[a = a(x, y) = \bar{a}(x) + \sum_{j>0} y_j \psi_j(x), \quad x \in D, y = (y_j)_{j>0} \in U := [-1,1]^N,\]
where \((\psi_j)_{j>0}\) is a given family of functions.

The parameters may be deterministic (control, optimization) or random (uncertainty modeling and propagation, reliability assessment).

Uniform ellipticity assumption :

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

Then \( u : y \mapsto u(y) = u(\cdot, y) \) is a bounded map from \( U \) to \( V := H^1_0(\Omega) : \|
\|u(y)\|_V \leq C_0 := \frac{\|f\|_{V^*}}{r}, \quad y \in U, \quad \text{where} \quad \|v\|_V := \|\nabla v\|_{L^2}.\]

Proof: multiply equation by \( u \) and integrate
\[r\|u\|^2_V \leq \int_D a\nabla u \cdot \nabla u = -\int_D u \text{div}(a\nabla u) = \int_D uf \leq \|u\|_V \|f\|_{V^*}.\]

Objective: build a computable approximation to this map at reasonable cost, i.e. simultaneously approximate \( u(y) \) for all \( y \in U \).
A model elliptic PDE

We consider the steady state diffusion equation

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

where $f = f(x) \in L^2(D)$ and $a = a(x, y)$ are variable coefficients depending on $x \in D$ and on a vector $y$ of parameters in an affine manner:

$$a = a(x, y) = \bar{a}(x) + \sum_{j>0} y_j \psi_j(x), \quad x \in D, y = (y_j)_{j>0} \in U := [-1, 1]^\mathbb{N},$$

where $(\psi_j)_{j>0}$ is a given family of functions.

The parameters may be deterministic (control, optimization) or random (uncertainty modeling and propagation, reliability assessment).

Uniform ellipticity assumption:

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

Then $u : y \mapsto u(y) = u(\cdot, y)$ is a bounded map from $U$ to $V := H^1_0(\Omega)$:

$$\|u(y)\|_V \leq C_0 := \frac{\|f\|_{V^*}}{r}, \quad y \in U, \quad \text{where } \|v\|_V := \|
abla v\|_{L^2}.$$

Proof: multiply equation by $u$ and integrate

$$r\|u\|_V^2 \leq \int_D a\nabla u \cdot \nabla u = -\int_D u \text{div}(a\nabla u) = \int_D uf \leq \|u\|_V \|f\|_{V^*}.$$

Objective: build a computable approximation to this map at reasonable cost, i.e. simultaneously approximate $u(y)$ for all $y \in U$. 
We consider the steady state diffusion equation

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

where $f = f(x) \in L^2(D)$ and $a = a(x, y)$ are variable coefficients depending on $x \in D$ and on a vector $y$ of parameters in an affine manner:

$$a = a(x, y) = \overline{a}(x) + \sum_{j>0} y_j \psi_j(x), \ x \in D, y = (y_j)_{j>0} \in U := [-1,1]^N,$$

where $(\psi_j)_{j>0}$ is a given family of functions.

The parameters may be deterministic (control, optimization) or random (uncertainty modeling and propagation, reliability assessment).

Uniform ellipticity assumption:

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

Then $u : y \mapsto u(y) = u(\cdot, y)$ is a bounded map from $U$ to $V := H^1_0(\Omega)$:

$$\|u(y)\|_V \leq C_0 := \frac{\|f\|_{V^*}}{r}, \ y \in U, \text{ where } \|v\|_V := \|\nabla v\|_{L^2}.$$

Proof: multiply equation by $u$ and integrate

$$r\|u\|_V^2 \leq \int_D a\nabla u \cdot \nabla u = -\int_D u \text{div}(a\nabla u) = \int_D uf \leq \|u\|_V \|f\|_{V^*}.$$

Objective: build a computable approximation to this map at reasonable cost, i.e. simultaneously approximate $u(y)$ for all $y \in U$. 
Polynomial expansions

Use of multivariate polynomials in the $y$ variable.

Sometimes referred to as “polynomial chaos” in the random setting (Ghanem-Spanos, Babushka-Tempone-Nobile-Zouharis, Karniadakis, Schwab...).

We study the convergence of the Taylor development

$$u(y) = \sum_{\nu \in \mathcal{F}} t_{\nu} y^{\nu},$$

where

$$y^{\nu} := \prod_{j>0} y_j^{\nu_j}.$$

Here $\mathcal{F}$ is the set of all finitely supported sequences $\nu = (\nu_j)_{j>0}$ of integers (only finitely many $\nu_j$ are non-zero). The Taylor coefficients $t_{\nu} \in \mathcal{V}$ are

$$t_{\nu} := \frac{1}{\nu!} \partial^{\nu} u|_{y=0} \text{ with } \nu! := \prod_{j>0} \nu_j! \text{ and } 0! := 1.$$

We also studied Legendre series $u(y) = \sum_{\nu \in \mathcal{F}} u_{\nu} L_{\nu}$ where

$$L_{\nu} (y) := \prod_{j>0} L_{\nu_j} (y_j).$$
Polynomial expansions

Use of multivariate polynomials in the $y$ variable.

Sometimes referred to as “polynomial chaos” in the random setting (Ghanem-Spanos, Babushka-Tempone-Nobile-Zouharis, Karniadakis, Schwab...).

We study the convergence of the Taylor development

$$u(y) = \sum_{\nu \in \mathcal{F}} t_\nu y^\nu,$$

where

$$y^\nu := \prod_{j > 0} y_j^{\nu_j}.$$

Here $\mathcal{F}$ is the set of all finitely supported sequences $\nu = (\nu_j)_{j > 0}$ of integers (only finitely many $\nu_j$ are non-zero). The Taylor coefficients $t_\nu \in \mathcal{V}$ are

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

We also studied Legendre series $u(y) = \sum_{\nu \in \mathcal{F}} u_\nu L_\nu$ where $L_\nu(y) := \prod_{j > 0} L_{\nu_j}(y_j)$. 
Sparse $N$-term polynomial approximation

The sequence $(t_\nu)_{\nu \in \mathcal{F}}$ is indexed by countably many integers.

Objective: identify a set $\Lambda \subset \mathcal{F}$ with $\#(\Lambda) \leq N$ such that $u$ is well approximated in the space

$$V_\Lambda := \left\{ \sum_{\nu \in \Lambda} c_\nu y^\nu ; \ u_\nu \in V \right\},$$

for example by the partial Taylor expansion

$$u_\Lambda (y) := \sum_{\nu \in \Lambda} t_\nu y^\nu.$$
Sparse \( N \)-term polynomial approximation

The sequence \((t_{\nu})_{\nu \in \mathcal{F}}\) is indexed by countably many integers.

Objective: identify a set \(\Lambda \subset \mathcal{F}\) with \(#(\Lambda) \leq N\) such that \(u\) is well approximated in the space

\[ V_{\Lambda} := \left\{ \sum_{\nu \in \Lambda} c_{\nu} y^{\nu} \mid u_{\nu} \in V \right\}, \]

for example by the partial Taylor expansion

\[ u_{\Lambda}(y) := \sum_{\nu \in \Lambda} t_{\nu} y^{\nu}. \]
Best $N$-term approximation

A-priori choices for $\Lambda$ have been proposed: (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 study a choice of $\Lambda$ optimally adapted to $u$.

For all $y \in U = [-1, 1]^N$ we have

$$\|u(y) - u_{\Lambda}(y)\|_V \leq \|\sum_{\nu \not\in \Lambda} t_{\nu} y^\nu\|_V \leq \sum_{\nu \not\in \Lambda} \|t_{\nu}\|_V$$

Best $N$-term approximation in the $l^1(F)$ norm: use for $\Lambda$ the $N$ largest $\|t_{\nu}\|_V$.

Observation (Stechkin): if $(\|t_{\nu}\|_V)_{\nu \in F} \in l^p(F)$ for some $p < 1$, then for this $\Lambda$,

$$\sum_{\nu \not\in \Lambda} \|t_{\nu}\|_V \leq CN^{-s}, \quad s := \frac{1}{p} - 1, \quad C := \|(\|t_{\nu}\|_V)\|_p.$$  

Proof: with $(t_n)_{n \geq 0}$ the decreasing rearrangement, we combine

$$\sum_{\nu \not\in \Lambda} \|t_{\nu}\|_V = \sum_{n \geq N} t_n = \sum_{n \geq N} t_n^{1-p} t_n^p \leq t_N^{1-p} C^p \quad \text{and} \quad N t_N^p \leq \sum_{n=1}^N t_n^p \leq C^p.$$  

Question: do we have $(\|t_{\nu}\|_V)_{\nu \in F} \in l^p(F)$ for some $p < 1$?
Best $N$-term approximation

A-priori choices for $\Lambda$ have been proposed: (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 study a choice of $\Lambda$ optimally adapted to $u$.

For all $y \in U = [-1, 1]^N$ we have

$$\|u(y) - u_\Lambda(y)\|_V \leq \left\| \sum_{\nu \notin \Lambda} t_{\nu} y^\nu \right\|_V \leq \sum_{\nu \notin \Lambda} \|t_{\nu}\|_V$$

Best $N$-term approximation in the $\ell^1(F)$ norm: use for $\Lambda$ the $N$ largest $\|t_{\nu}\|_V$.

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

$$\sum_{\nu \notin \Lambda} \|t_{\nu}\|_V \leq CN^{-s}, \quad s := \frac{1}{p} - 1, \quad C := \|\|t_{\nu}\|_V\|_p.$$

Proof: with $(t_n)_{n>0}$ the decreasing rearrangement, we combine

$$\sum_{\nu \notin \Lambda} \|t_{\nu}\|_V = \sum_{n>N} t_n = \sum_{n>N} t_{n}^{1-p} t_{n}^p \leq t_N^{1-p} C^p \text{ and } Nt_N^p \leq \sum_{n=1}^{N} t_n^p \leq C^p.$$

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

A-priori choices for $\Lambda$ have been proposed: (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 study a choice of $\Lambda$ optimally adapted to $u$.

For all $y \in U = [-1,1]^N$ we have

$$\|u(y) - u_{\Lambda}(y)\|_V \leq \|\sum_{\nu \notin \Lambda} t_{\nu} y^\nu\|_V \leq \sum_{\nu \notin \Lambda} \|t_{\nu}\|_V$$

Best $N$-term approximation in the $\ell^1(F)$ norm: use for $\Lambda$ the $N$ largest $\|t_{\nu}\|_V$.

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

$$\sum_{\nu \notin \Lambda} \|t_{\nu}\|_V \leq CN^{-s}, \quad s := \frac{1}{p} - 1, \quad C := \|(\|t_{\nu}\|_V)\|_p.$$  

Proof: with $(t_n)_{n>0}$ the decreasing rearrangement, we combine

$$\sum_{\nu \notin \Lambda} \|t_{\nu}\|_V = \sum_{n>N} t_n = \sum_{n>N} t_n^{1-p} t_n^p \leq t_N^{1-p} C^p \quad \text{and} \quad N t_n^p \leq \sum_{n=1}^N t_n^p \leq C^p.$$  

Question: do we have $(\|t_{\nu}\|_V)_{\nu \in F} \in \ell^p(F)$ for some $p < 1$?
**Best $N$-term approximation**

A-priori choices for $\Lambda$ have been proposed: (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 study a choice of $\Lambda$ optimally adapted to $u$.

For all $y \in U = [-1, 1]^N$ we have

$$\| u(y) - u_{\Lambda}(y) \|_V \leq \| \sum_{\nu \notin \Lambda} t_\nu y^\nu \|_V \leq \sum_{\nu \notin \Lambda} \| t_\nu \|_V$$

Best $N$-term approximation in the $\ell^1(F)$ norm: use for $\Lambda$ the $N$ largest $\| t_\nu \|_V$.

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

$$\sum_{\nu \notin \Lambda} \| t_\nu \|_V \leq CN^{-s}, \quad s := \frac{1}{p} - 1, \quad C := \| (\| t_\nu \|_V) \|_p.$$  

Proof: with $(t_n)_{n>0}$ the decreasing rearrangement, we combine

$$\sum_{\nu \notin \Lambda} \| t_\nu \|_V = \sum_{n>N} t_n = \sum_{n>N} t_n^{1-p} t_n^p \leq t_N^{1-p} C^p \quad \text{and} \quad N t_n^p \leq \sum_{n=1}^N t_n^p \leq C^p.$$  

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

Theorem (Cohen-DeVore-Schwab, 2009) : under the uniform ellipticity assumption (UAE), then for any $p < 1$,

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

Interpretations :

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

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

(iii) The set $\mathcal{K} := \{u(y) ; y \in U\}$ is compact in $V$ and has small $N$-width $d_N(\mathcal{K}) := \inf_{\dim(E) \leq N} \max_{\nu \in \mathcal{K}} \text{dist}(\nu, E)_V$ : for all $y$

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

With $\Lambda$ corresponding to the $N$ largest $\|t_\nu\|_V$, we find that

$$d_N(\mathcal{K}) \leq \max_{y \in U} \text{dist}(u(y), E_\Lambda)_V \leq \max_{y \in U} \|u(y) - u_\Lambda(y)\|_V \leq CN^{-s}.$$  

Such approximation rates cannot be proved for the usual a-priori choices of $\Lambda$. 

The main result

Theorem (Cohen-DeVore-Schwab, 2009): under the uniform ellipticity assumption (UAE), then for any $p < 1$,

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

Interpretations:

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

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

(iii) The set $\mathcal{K} := \{u(y); y \in U\}$ is compact in $V$ and has small $N$-width $d_N(\mathcal{K}) := \inf_{\dim(E) \leq N} \max_{\nu \in \mathcal{K}} \text{dist}(\nu, E)_V$:

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

With $\Lambda$ corresponding to the $N$ largest $\|t_\nu\|_V$, we find that

$$d_N(\mathcal{K}) \leq \max_{y \in U} \text{dist}(u(y), E_\Lambda)_V \leq \max_{y \in U} \|u(y) - u_\Lambda(y)\|_V \leq CN^{-s}.$$  

Such approximation rates cannot be proved for the usual a-priori choices of $\Lambda$. 
The main result

Theorem (Cohen-DeVore-Schwab, 2009): under the uniform ellipticity assumption (UAE), then for any \( p < 1 \),

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

Interpretations:

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

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

(iii) The set \( \mathcal{K} := \{ u(y) ; y \in U \} \) is compact in \( V \) and has small \( N \)-width

\[
d_N(\mathcal{K}) := \inf_{\dim(E) \leq N} \max_{\nu \in \mathcal{K}} \text{dist}(\nu, E)_V : \text{ for all } y
\]

\[
u \in \Lambda \]

\[
\sum_{\nu \in \Lambda} t_\nu y^\nu = \sum_{\nu \in \Lambda} y^\nu t_\nu \in E_\Lambda := \text{Span}\{t_\nu ; \nu \in \Lambda\}.
\]

With \( \Lambda \) corresponding to the \( N \) largest \( \|t_\nu\|_V \), we find that

\[
d_N(\mathcal{K}) \leq \max_{y \in U} \text{dist}(u(y), E_\Lambda)_V \leq \max_{y \in U} \|u(y) - u_\Lambda(y)\|_V \leq CN^{-s}.
\]

Such approximation rates cannot be proved for the usual a-priori choices of \( \Lambda \).
The main result

Theorem (Cohen-DeVore-Schwab, 2009): under the uniform ellipticity assumption (UAE), then for any $p < 1$,

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

Interpretations:

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

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

(iii) The set $\mathcal{K} := \{u(y) ; y \in U\}$ is compact in $V$ and has small $N$-width $d_N(\mathcal{K}) := \inf_{\dim(E) \leq N} \max_{\nu \in \mathcal{K}} \text{dist}(\nu, E)_V$ : for all $y$

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

With $\Lambda$ corresponding to the $N$ largest $\|t_\nu\|_V$, we find that

$$d_N(\mathcal{K}) \leq \max_{y \in U} \text{dist}(u(y), E_\Lambda)_V \leq \max_{y \in U} \|u(y) - u_\Lambda(y)\|_V \leq CN^{-s}.$$  

Such approximation rates cannot be proved for the usual a-priori choices of $\Lambda$. 
Idea of proof: extension to complex variable

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

Uniform ellipticity \( 0 < r \leq \bar{a}(x) + \sum_{j>0} y_j \psi_j(x) \) for all \( x \in D, y \in U \) is equivalent to

\[
\sum_{j>0} |\psi_j(x)| \leq \bar{a}(x) - r, \quad x \in D.
\]

This allows to say that with \( a(x, z) = \bar{a}(x) + \sum_{j>0} z_j \psi_j(x) \),

\[
0 < r \leq \Re(a(x, z)) \leq |a(x, z)| \leq 2R,
\]

for all \( z \in U := \{ |z| \leq 1 \}^N = \otimes \{ |z_j| \leq 1 \} \).

Lax-Milgram theory applies: \( \| u(z) \| \leq C_0 = \frac{\| f \|_{V^*}}{r} \) for all \( z \in U \). The function \( u \mapsto u(z) \) is holomorphic in each variable \( z_j \) at any \( z \in U \).

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

\[
\sum_{j>0} \rho_j |\psi_j(x)| \leq \bar{a}(x) - \delta, \quad x \in D,
\]

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

\[
U_\rho := \otimes \{ |z_j| \leq \rho_j \},
\]

We call such sequences \( \rho \) “\( \delta \)-admissible”. 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 $z = (z_j) \in \mathbb{C}^N$.

Uniform ellipticity $0 < r \leq \Re(a(x) + \sum_{j > 0} y_j \psi_j(x))$ for all $x \in D, y \in U$ is equivalent to

$$\sum_{j > 0} |\psi_j(x)| \leq \Re(a(x)) - r, \quad x \in D.$$ 

This allows to say that with $a(x, z) = \Re(a(x) + \sum_{j > 0} z_j \psi_j(x))$,

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

for all $z \in U := \{|z| \leq 1\}^N = \otimes \{|z_j| \leq 1\}$.

Lax-Milgram theory applies: $\|u(z)\| \leq C_0 = \frac{\|f\|_{V^*}}{r}$ for all $z \in U$. The function $u \mapsto u(z)$ is holomorphic in each variable $z_j$ at any $z \in U$.

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

$$\sum_{j > 0} \rho_j |\psi_j(x)| \leq \Re(a(x)) - \delta, \quad x \in D,$$

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

We call such sequences $\rho$ “$\delta$-admissible”. 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 $z = (z_j) \in \mathbb{C}^N$.

Uniform ellipticity $0 < r \leq \overline{a}(x) + \sum_{j > 0} y_j \psi_j(x)$ for all $x \in D, y \in U$ is equivalent to

$$\sum_{j > 0} |\psi_j(x)| \leq \overline{a}(x) - r, \ x \in D.$$

This allows to say that with $a(x, z) = \overline{a}(x) + \sum_{j > 0} z_j \psi_j(x)$,

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

for all $z \in U := \{|z| \leq 1\}^N \otimes \{|z_j| \leq 1\}$.

Lax-Milgram theory applies: $\|u(z)\| \leq C_0 = \frac{\|f\|_V^*}{r}$ for all $z \in U$. The function $u \mapsto u(z)$ is holomorphic in each variable $z_j$ at any $z \in U$.

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

$$\sum_{j > 0} \rho_j |\psi_j(x)| \leq \overline{a}(x) - \delta, \ x \in D,$$

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

$$\mathcal{U}_\rho := \otimes \{|z_j| \leq \rho_j\},$$

We call such sequences $\rho$ “$\delta$-admissible”. If $\delta < r$, we can take $\rho_j > 1$. 
Idea of proof: extension to complex variable

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

Uniform ellipticity \( 0 < r \leq \bar{a}(x) + \sum_{j>0} y_j \psi_j(x) \) for all \( x \in D, y \in U \) is equivalent to

\[
\sum_{j>0} |\psi_j(x)| \leq \bar{a}(x) - r, \quad x \in D.
\]

This allows to say that with \( a(x,z) = \bar{a}(x) + \sum_{j>0} z_j \psi_j(x) \),

\[
0 < r \leq \Re(a(x,z)) \leq |a(x,z)| \leq 2R,
\]

for all \( z \in \mathcal{U} := \{ |z| \leq 1 \}^N = \otimes \{ |z_j| \leq 1 \} \).

Lax-Milgram theory applies: \( \| u(z) \| \leq C_0 = \frac{\| f \| \nu^*}{r} \) for all \( z \in \mathcal{U} \). The function \( u \mapsto u(z) \) is holomorphic in each variable \( z_j \) at any \( z \in \mathcal{U} \).

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

\[
\sum_{j>0} \rho_j |\psi_j(x)| \leq \bar{a}(x) - \delta, \quad x \in D,
\]

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

\[
\mathcal{U}_\rho := \otimes \{ |z_j| \leq \rho_j \},
\]

We call such sequences \( \rho \) “\( \delta \)-admissible”. If \( \delta < r \), we can take \( \rho_j > 1 \).
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 \( \{ |z| \leq a \} \), then for all \( z \) in this disc

\[
  u(z) = \frac{1}{2i\pi} \int_{|z'|=a} \frac{u(z')}{z-z'} dz',
\]

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

Recursive application of this to all variables \( z_j \) such that \( \nu_j \neq 0 \), with \( a = \rho_j \), for a \( \delta \)-admissible sequence \( \rho \) gives

\[
  \| \partial^{\nu} u|_{z=0} \|_{\nu} \leq C_\delta \nu! \prod_{j>0} \rho_j^{-\nu_j}.
\]

and therefore

\[
  \| t_\nu \|_{\nu} \leq C_\delta \prod_{j>0} \rho_j^{-\nu_j} = C_0 \rho^{-\nu}.
\]

Since \( \rho \) is not fixed we have

\[
  \| t_\nu \|_{\nu} \leq C_\delta \inf\{ \rho^{-\nu} ; \ \rho \text{ is } \delta-\text{admissible} \}.
\]

We do not know the general solution to this problem, except when the \( \psi_j \) have disjoint supports. Instead design a particular choice \( \rho = \rho(\nu) \) of \( \delta \)-admissible sequences with \( \delta = r/2 \), for which we prove that

\[
  (\| \psi_j \|_{L^\infty})_{j \geq 0} \in \ell^p(N) \Rightarrow (\rho(\nu)^{-\nu})_{\nu \in F} \in \ell^p(F).
\]
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 \( \{|z| \leq a\} \), then for all \( z \) in this disc

\[
    u(z) = \frac{1}{2i\pi} \int_{|z'|=a} \frac{u(z')}{z-z'} dz',
\]

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

Recursive application of this to all variables \( z_j \) such that \( \nu_j \neq 0 \), with \( a = \rho_j \), for a \( \delta \)-admissible sequence \( \rho \) gives

\[
    \|\partial^\nu u|_{z=0}\|_V \leq C_\delta \nu! \prod_{j>0} \rho_j^{-\nu_j}.
\]

and therefore

\[
    \|t_\nu\|_V \leq C_\delta \prod_{j>0} \rho_j^{-\nu_j} = C_0 \rho^{-\nu}.
\]

Since \( \rho \) is not fixed we have

\[
    \|t_\nu\|_V \leq C_\delta \inf\{\rho^{-\nu} \mid \rho \text{ is } \delta-\text{admissible}\}.
\]

We do not know the general solution to this problem, except when the \( \psi_j \) have disjoint supports. Instead design a particular choice \( \rho = \rho(\nu) \) of \( \delta \)-admissible sequences with \( \delta = r/2 \), for which we prove that

\[
    (\|\psi_j\|_{\ell^\infty})_{j \geq 0} \in \ell^p(\mathbb{N}) \Rightarrow (\rho(\nu)^{-\nu})_{\nu \in \mathcal{F}} \in \ell^p(\mathcal{F}).
\]
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 \( \{ |z| \leq a \} \), then for all \( z \) in this disc

\[
u(z) = \frac{1}{2i\pi} \int_{|z'|=a} \frac{u(z')}{z - z'} dz',
\]

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

Recursive application of this to all variables \( z_j \) such that \( \nu_j \neq 0 \), with \( a = \rho_j \), for a \( \delta \)-admissible sequence \( \rho \) gives

\[
\| \partial^\nu u |_{z=0} \|_V \leq C_\delta \nu! \prod_{j>0} \rho_j^{-\nu_j}.
\]

and therefore

\[
\| t_\nu \|_V \leq C_\delta \prod_{j>0} \rho_j^{-\nu_j} = C_0 \rho^{-\nu}.
\]

Since \( \rho \) is not fixed we have

\[
\| t_\nu \|_V \leq C_\delta \inf \{ \rho^{-\nu} ; \ \rho \text{ is } \delta - \text{admissible} \}.
\]

We do not know the general solution to this problem, except when the \( \psi_j \) have disjoint supports. Instead design a particular choice \( \rho = \rho(\nu) \) of \( \delta \)-admissible sequences with \( \delta = r/2 \), for which we prove that

\[
(\| \psi_j \|_{\ell^\infty})_{j \geq 0} \in \ell^p(\mathbb{N}) \Rightarrow (\rho(\nu)^{-\nu})_{\nu \in \mathcal{F}} \in \ell^p(\mathcal{F}).
\]
A simple case

Assume that the $\psi_j$ have disjoint supports. Then we maximize separately the $\rho_j$ so that

$$\sum_{j>0} \rho_j |\psi_j(x)| \leq \overline{\alpha}(x) - \frac{r}{2}, \ x \in D,$$

which leads to

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

We have

$$\|t_\nu\|_V \leq 2C_0 \rho^{-\nu} = 2C_0 b^{\nu},$$

where $b = (b_j)$ and

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

Therefore $b \in \ell^p(\mathbb{N})$. From (UEA), we have $|\psi_j(x)| \leq \overline{\alpha}(x) - r$ 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

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

Assume that the $\psi_j$ have disjoint supports. Then we maximize separately the $\rho_j$ so that

$$\sum_{j>0} \rho_j |\psi_j(x)| \leq \bar{a}(x) - \frac{r}{2}, \quad x \in D,$$

which leads to

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

We have

$$\|t_{\nu}\|_V \leq 2C_0 \rho^{-\nu} = 2C_0 b^\nu,$$

where $b = (b_j)$ and

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

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

We finally observe that

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

Proof: factorize

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

Assume that the $\psi_j$ have disjoint supports. Then we maximize separately the $\rho_j$ so that

$$\sum_{j > 0} \rho_j |\psi_j(x)| \leq \overline{a}(x) - \frac{r}{2}, \quad x \in D,$$

which leads to

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

We have

$$\|t_\nu\|_V \leq 2C_0 \rho^{-\nu} = 2C_0 b^\nu,$$

where $b = (b_j)$ and

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

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

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

Proof : factorize

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

Strategies to build the set $\Lambda$:

(i) **Non-adaptive**, based on the available a-priori estimates for the $\|t_\nu\|_V$.

(ii) **Adaptive**, based on a-posteriori information gained in the computation $\Lambda_1 \subset \Lambda_2 \subset \ldots \subset \Lambda_N$.

Objective: develop adaptive strategies that converge with optimal rate (similar to adaptive wavelet methods for elliptic PDE's: Cohen-Dahmen-DeVore, Stevenson).

Recursive computation of the Taylor coefficients: with $e_j$ the Kroenecker sequence

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

We compute the $t_\nu$ on sets $\Lambda$ with monotone structure: $\nu \in \Lambda$ and $\mu \leq \nu \Rightarrow \mu \in \Lambda$.

Given such a $\Lambda_k$ and the $(t_\nu)_{\nu \in \Lambda_k}$ we compute the $t_\nu$ for $\nu$ in the margin

$$\mathcal{M}_k := \{\nu \notin \Lambda_k ; \nu - e_j \in \Lambda_k \text{ for some } j\},$$

and build the new set by **bulk search**: $\Lambda_{k+1} = \Lambda_k \cup S_k$, with $S_k \subset \mathcal{M}_k$ smallest such that $\sum_{\nu \in S_k} \|t_\nu\|_V^2 \geq \theta \sum_{\nu \in \mathcal{M}_k} \|t_\nu\|_V^2$, with $\theta \in (0,1)$.

Such a strategy can be proved to converge with optimal convergence rate $\#(\Lambda_k)^{−s}$. 

An adaptive algorithm

Strategies to build the set $\Lambda$:

(i) **Non-adaptive**, based on the available a-priori estimates for the $\|t_\nu\|_V$.

(ii) **Adaptive**, based on a-posteriori information gained in the computation $\Lambda_1 \subset \Lambda_2 \subset \cdots \subset \Lambda_N$.

Objective: develop adaptive strategies that converge with optimal rate (similar to adaptive wavelet methods for elliptic PDE’s: Cohen-Dahmen-DeVore, Stevenson).

Recursive computation of the Taylor coefficients: with $e_j$ the Kroenecker sequence

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

We compute the $t_\nu$ on sets $\Lambda$ with monotone structure: $\nu \in \Lambda$ and $\mu \leq \nu \Rightarrow \mu \in \Lambda$.

Given such a $\Lambda_k$ and the $(t_\nu)_{\nu \in \Lambda_k}$ we compute the $t_\nu$ for $\nu$ in the margin

$$\mathcal{M}_k := \{\nu \notin \Lambda_k \; ; \; \nu - e_j \in \Lambda_k \; \text{ for some } j\},$$

and build the new set by bulk search: $\Lambda_{k+1} = \Lambda_k \cup S_k$, with $S_k \subset \mathcal{M}_k$ smallest such that $\sum_{\nu \in S_k} \|t_\nu\|^2_V \geq \theta \sum_{\nu \in \mathcal{M}_k} \|t_\nu\|^2_V$, with $\theta \in (0,1)$.

Such a strategy can be proved to converge with optimal convergence rate $\#(\Lambda_k)^{-s}$. 
An adaptive algorithm

Strategies to build the set $\Lambda$ :

(i) **Non-adaptive**, based on the available a-priori estimates for the $\|t_\nu\|_V$.

(ii) **Adaptive**, based on a-posteriori information gained in the computation $\Lambda_1 \subset \Lambda_2 \subset \cdots \subset \Lambda_N$.

Objective : develop adaptive strategies that converge with optimal rate (similar to adaptive wavelet methods for elliptic PDE’s : Cohen-Dahmen-DeVore, Stevenson).

Recursive computation of the Taylor coefficients : with $e_j$ the Kroenecker sequence

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

We compute the $t_\nu$ on sets $\Lambda$ with **monotone** structure : $\nu \in \Lambda$ and $\mu \leq \nu \Rightarrow \mu \in \Lambda$.

Given such a $\Lambda_k$ and the $(t_\nu)_{\nu \in \Lambda_k}$ we compute the $t_\nu$ for $\nu$ in the margin

$$
\mathcal{M}_k := \{\nu \notin \Lambda_k ; \nu - e_j \in \Lambda_k \text{ for some } j\},
$$

and build the new set by **bulk search** : $\Lambda_{k+1} = \Lambda_k \cup S_k$, with $S_k \subset \mathcal{M}_k$ smallest such that

$$
\sum_{\nu \in S_k} \|t_\nu\|_V^2 \geq \theta \sum_{\nu \in \mathcal{M}_k} \|t_\nu\|_V^2, \quad \text{with } \theta \in (0,1).
$$

Such a strategy can be proved to converge with optimal convergence rate $\#(\Lambda_k)^{-s}$. 
An adaptive algorithm

Strategies to build the set $\Lambda$:

(i) **Non-adaptive**, based on the available a-priori estimates for the $\|t_\nu\|_V$.

(ii) **Adaptive**, based on a-posteriori information gained in the computation $\Lambda_1 \subset \Lambda_2 \subset \cdots \subset \Lambda_N$.

Objective: develop adaptive strategies that converge with optimal rate (similar to adaptive wavelet methods for elliptic PDE’s: Cohen-Dahmen-DeVore, Stevenson).

Recursive computation of the Taylor coefficients: with $e_j$ the Kroenecker sequence

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

We compute the $t_\nu$ on sets $\Lambda$ with **monotone** structure: $\nu \in \Lambda$ and $\mu \leq \nu \Rightarrow \mu \in \Lambda$.

Given such a $\Lambda_k$ and the $(t_\nu)_{\nu \in \Lambda_k}$ we compute the $t_\nu$ for $\nu$ in the **margin**

\[
\mathcal{M}_k := \{\nu \notin \Lambda_k ; \nu - e_j \in \Lambda_k \text{ for some } j\},
\]

and build the new set by **bulk search**: $\Lambda_{k+1} = \Lambda_k \cup S_k$, with $S_k \subset \mathcal{M}_k$ smallest such that

\[
\sum_{\nu \in S_k} \|t_\nu\|_V^2 \geq \theta \sum_{\nu \in \mathcal{M}_k} \|t_\nu\|_V^2,
\]

with $\theta \in (0,1)$.

Such a strategy can be proved to converge with optimal convergence rate $\#(\Lambda_k)^{-5}$. 
An adaptive algorithm

Strategies to build the set \( \Lambda \):

(i) **Non-adaptive**, based on the available a-priori estimates for the \( \| t_\nu \|_V \).

(ii) **Adaptive**, based on a-posteriori information gained in the computation \( \Lambda_1 \subset \Lambda_2 \subset \cdots \subset \Lambda_N \).

Objective: develop adaptive strategies that converge with optimal rate (similar to adaptive wavelet methods for elliptic PDE’s: Cohen-Dahmen-DeVore, Stevenson).

Recursive computation of the Taylor coefficients: with \( e_j \) the Kroenecker sequence

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

We compute the \( t_\nu \) on sets \( \Lambda \) with monotone structure: \( \nu \in \Lambda \) and \( \mu \leq \nu \Rightarrow \mu \in \Lambda \).

Given such a \( \Lambda_k \) and the \( (t_\nu)_{\nu \in \Lambda_k} \) we compute the \( t_\nu \) for \( \nu \) in the margin

\[
\mathcal{M}_k := \{ \nu \notin \Lambda_k; \nu - e_j \in \Lambda_k \text{ for some } j \},
\]

and build the new set by bulk search: \( \Lambda_{k+1} = \Lambda_k \cup S_k \), with \( S_k \subset \mathcal{M}_k \) smallest such that \( \sum_{\nu \in S_k} \| t_\nu \|_V^2 \geq \theta \sum_{\nu \in \mathcal{M}_k} \| t_\nu \|_V^2 \), with \( \theta \in (0,1) \).

Such a strategy can be proved to converge with optimal convergence rate \( \#(\Lambda_k)^{-s} \).
Test case in moderate dimension $d = 16$

Physical domain $D = [0, 1]^2 = \bigcup_{j=1}^{d} D_j$.

Diffusion coefficients $a(x, y) = 1 + \sum_{j=1}^{d} y_j \left( \frac{0.9}{j^2} \right) \chi_{D_j}$.

Adaptive search of $\Lambda$ implemented in C++, spatial discretization by FreeFem++.

Comparison between the $\Lambda_k$ generated by the adaptive algorithm (red) and non-adaptive choices $\{\sup \nu_j \leq k\}$ (blue) or $\{\sum \nu_j \leq k\}$ (green) or $k$ largest a-priori bounds on the $\|t_\nu\|_V$ (pink)

Highest polynomial degree with $\#(\Lambda) = 1000$ coefficients: 1, 4, 115 and 81.
What I did not speak about

Use of Legendre polynomials instead of Taylor series: leads to approximation error estimates in $L^2(U, d\mu)$ with $d\mu$ the tensor product probability measure on $U$.

Computation of the approximate Legendre coefficients: either use a Galerkin (projection) method or a Collocation (interpolation) method. For the second one, designing optimal collocation points is an open problem.

Strategies to build the set $\Lambda$:

(i) **A-priori**, based on the available estimates for the $\|t_\nu\|_V$.

(ii) **A-posteriori**, based on error indicators in the Galerkin framework: $\Lambda_1 \subset \Lambda_2 \subset \cdots \subset \Lambda_N$. Optimal convergence of this strategy may be proved by similar techniques as for adaptive wavelet methods for elliptic PDE’s (Cohen-Dahmen-DeVore, Stevenson).

(iii) Reconstruction a sparse orthogonal series from random sampling: techniques from **Compressed Sensing** (Sparse Fourier series: Gilbert-Strauss-Tropp, Candes-Romberg-Tao, Rudelson-Vershynin. Sparse Legendre series: Rauhut-Ward 2010).

**Space discretization**: should be properly tuned (use different resolution for each $t_\nu$) and injected in the final error analysis.

Our results can be used in the analysis of **reduced basis methods**.
What I did not speak about

Use of Legendre polynomials instead of Taylor series: leads to approximation error estimates in $L^2(U, d\mu)$ with $d\mu$ the tensor product probability measure on $U$.

Computation of the approximate Legendre coefficients: either use a Galerkin (projection) method or a Collocation (interpolation) method. For the second one, designing optimal collocation points is an open problem.

Strategies to build the set $\Lambda$:

(i) A-priori, based on the available estimates for the $\|t_\nu\|_V$.

(ii) A-posteriori, based on error indicators in the Galerkin framework: $\Lambda_1 \subset \Lambda_2 \subset \cdots \subset \Lambda_N$. Optimal convergence of this strategy may be proved by similar techniques as for adaptive wavelet methods for elliptic PDE's (Cohen-Dahmen-DeVore, Stevenson).


Space discretization: should be properly tuned (use different resolution for each $t_\nu$) and injected in the final error analysis.

Our results can be used in the analysis of reduced basis methods.
What I did not speak about

Use of Legendre polynomials instead of Taylor series leads to approximation error estimates in $L^2(U, d\mu)$ with $d\mu$ the tensor product probability measure on $U$.

Computation of the approximate Legendre coefficients: either use a Galerkin (projection) method or a Collocation (interpolation) method. For the second one, designing optimal collocation points is an open problem.

Strategies to build the set $\Lambda$:

(i) **A-priori**, based on the available estimates for the $\|t_\nu\|_V$.

(ii) **A-posteriori**, based on error indicators in the Galerkin framework: $\Lambda_1 \subset \Lambda_2 \subset \cdots \subset \Lambda_N$. Optimal convergence of this strategy may be proved by similar techniques as for adaptive wavelet methods for elliptic PDE’s (Cohen-Dahmen-DeVore, Stevenson).


Space discretization: should be properly tuned (use different resolution for each $t_\nu$) and injected in the final error analysis.

Our results can be used in the analysis of reduced basis methods.
What I did not speak about

Use of Legendre polynomials instead of Taylor series: leads to approximation error estimates in $L^2(U, d\mu)$ with $d\mu$ the tensor product probability measure on $U$.

Computation of the approximate Legendre coefficients: either use a Galerkin (projection) method or a Collocation (interpolation) method. For the second one, designing optimal collocation points is an open problem.

Strategies to build the set $\Lambda$:

(i) **A-priori**, based on the available estimates for the $\| t_\nu \|_V$.

(ii) **A-posteriori**, based on error indicators in the Galerkin framework: $\Lambda_1 \subset \Lambda_2 \subset \cdots \subset \Lambda_N$. Optimal convergence of this strategy may be proved by similar techniques as for adaptive wavelet methods for elliptic PDE's (Cohen-Dahmen-DeVore, Stevenson).


Space discretization: should be properly tuned (use different resolution for each $t_\nu$) and injected in the final error analysis.

Our results can be used in the analysis of reduced basis methods.
What I did not speak about

Use of Legendre polynomials instead of Taylor series: leads to approximation error estimates in $L^2(U, d\mu)$ with $d\mu$ the tensor product probability measure on $U$.

Computation of the approximate Legendre coefficients: either use a Galerkin (projection) method or a Collocation (interpolation) method. For the second one, designing optimal collocation points is an open problem.

Strategies to build the set $\Lambda$:

(i) **A-priori**, based on the available estimates for the $\|t_\nu\|_V$.

(ii) **A-posteriori**, based on error indicators in the Galerkin framework: $\Lambda_1 \subset \Lambda_2 \subset \cdots \subset \Lambda_N$. Optimal convergence of this strategy may be proved by similar techniques as for adaptive wavelet methods for elliptic PDE’s (Cohen-Dahmen-DeVore, Stevenson).


Space discretization: should be properly tuned (use different resolution for each $t_\nu$) and injected in the final error analysis.

Our results can be used in the analysis of reduced basis methods.
What I did not speak about

Use of Legendre polynomials instead of Taylor series: leads to approximation error estimates in $L^2(U, d\mu)$ with $d\mu$ the tensor product probability measure on $U$.

Computation of the approximate Legendre coefficients: either use a Galerkin (projection) method or a Collocation (interpolation) method. For the second one, designing optimal collocation points is an open problem.

Strategies to build the set $\Lambda$:

(i) **A-priori**, based on the available estimates for the $\|t_\nu\|_V$.

(ii) **A-posteriori**, based on error indicators in the Galerkin framework: $\Lambda_1 \subset \Lambda_2 \subset \cdots \subset \Lambda_N$. Optimal convergence of this strategy may be proved by similar techniques as for adaptive wavelet methods for elliptic PDE’s (Cohen-Dahmen-DeVore, Stevenson).

(iii) Reconstruction a sparse orthogonal series from random sampling: techniques from **Compressed Sensing** (Sparse Fourier series: Gilbert-Strauss-Tropp, Candes-Romberg-Tao, Rudelson-Vershynin. Sparse Legendre series: Rauhut-Ward 2010).

Space discretization: should be properly tuned (use different resolution for each $t_\nu$) and injected in the final error analysis.

Our results can be used in the analysis of **reduced basis methods**.
What I did not speak about

Use of Legendre polynomials instead of Taylor series: leads to approximation error estimates in $L^2(U, d\mu)$ with $d\mu$ the tensor product probability measure on $U$.

Computation of the approximate Legendre coefficients: either use a Galerkin (projection) method or a Collocation (interpolation) method. For the second one, designing optimal collocation points is an open problem.

Strategies to build the set $\Lambda$:

(i) **A-priori**, based on the available estimates for the $\|t_\nu\|_V$.

(ii) **A-posteriori**, based on error indicators in the Galerkin framework: $\Lambda_1 \subset \Lambda_2 \subset \cdots \subset \Lambda_N$. Optimal convergence of this strategy may be proved by similar techniques as for adaptive wavelet methods for elliptic PDE’s (Cohen-Dahmen-DeVore, Stevenson).


**Space discretization**: should be properly tuned (use different resolution for each $t_\nu$) and injected in the final error analysis.

Our results can be used in the analysis of **reduced basis methods**.
Use of Legendre polynomials instead of Taylor series: leads to approximation error estimates in $L^2(U, d\mu)$ with $d\mu$ the tensor product probability measure on $U$.

Computation of the approximate Legendre coefficients: either use a Galerkin (projection) method or a Collocation (interpolation) method. For the second one, designing optimal collocation points is an open problem.

Strategies to build the set $\Lambda$:

(i) **A-priori**, based on the available estimates for the $\|t_\nu\|_V$.

(ii) **A-posteriori**, based on error indicators in the Galerkin framework: $\Lambda_1 \subset \Lambda_2 \subset \cdots \subset \Lambda_N$. Optimal convergence of this strategy may be proved by similar techniques as for adaptive wavelet methods for elliptic PDE’s (Cohen-Dahmen-DeVore, Stevenson).

(iii) Reconstruction a sparse orthogonal series from random sampling: techniques from **Compressed Sensing** (Sparse Fourier series: Gilbert-Strauss-Tropp, Candes-Romberg-Tao, Rudelson-Vershynin. Sparse Legendre series: Rauhut-Ward 2010).

**Space discretization**: should be properly tuned (use different resolution for each $t_\nu$) and injected in the final error analysis.

Our results can be used in the analysis of **reduced basis methods**.
**Conclusion and perspective**

**Rich topic:** involves a variety of tools such as stochastic processes, high dimensional approximation, complex analysis, sparsity and non-linear approximation, adaptivity and a-posteriori analysis.

**First numerical results in moderate dimensionality:** reveal the advantages of an adaptive approach. **Goal:** implementation for very high or infinite dimensionality.

Many applications in engineering.

Many other models to be studied:

(i) Non-affine dependence of \( a \) in the variable \( y \).

(ii) Other linear or non-linear PDE's.

Papers: www.ann.jussieu.fr/cohen

THANKS!
Conclusion and perspective

Rich topic: involves a variety of tools such as stochastic processes, high dimensional approximation, complex analysis, sparsity and non-linear approximation, adaptivity and a-posteriori analysis.

First numerical results in moderate dimensionality: reveal the advantages of an adaptive approach. Goal: implementation for very high or infinite dimensionality.

Many applications in engineering.

Many other models to be studied:

(i) Non-affine dependence of $a$ in the variable $y$.

(ii) Other linear or non-linear PDE's.

Papers: www.ann.jussieu.fr/cohen

THANKS!
Conclusion and perspective

Rich topic: involves a variety of tools such as stochastic processes, high dimensional approximation, complex analysis, sparsity and non-linear approximation, adaptivity and a-posteriori analysis.

First numerical results in moderate dimensionality: reveal the advantages of an adaptive approach. Goal: implementation for very high or infinite dimensionality.

Many applications in engineering.

Many other models to be studied:

(i) Non-affine dependence of $a$ in the variable $y$.

(ii) Other linear or non-linear PDE's.

Papers: www.ann.jussieu.fr/cohen

THANKS!
Conclusion and perspective

Rich topic: involves a variety of tools such as stochastic processes, high dimensional approximation, complex analysis, sparsity and non-linear approximation, adaptivity and a-posteriori analysis.

First numerical results in moderate dimensionality: reveal the advantages of an adaptive approach. Goal: implementation for very high or infinite dimensionality.

Many applications in engineering.

Many other models to be studied:

(i) Non-affine dependence of $a$ in the variable $y$.

(ii) Other linear or non-linear PDE’s.

Papers: www.ann.jussieu.fr/cohen

THANKS!
Conclusion and perspective

Rich topic: involves a variety of tools such as stochastic processes, high dimensional approximation, complex analysis, sparsity and non-linear approximation, adaptivity and a-posteriori analysis.

First numerical results in moderate dimensionality: reveal the advantages of an adaptive approach. Goal: implementation for very high or infinite dimensionality.

Many applications in engineering.

Many other models to be studied:

(i) Non-affine dependence of $a$ in the variable $y$.

(ii) Other linear or non-linear PDE’s.

Papers: www.ann.jussieu.fr/cohen

THANKS!
Conclusion and perspective

Rich topic: involves a variety of tools such as stochastic processes, high dimensional approximation, complex analysis, sparsity and non-linear approximation, adaptivity and a-posteriori analysis.

First numerical results in moderate dimensionality: reveal the advantages of an adaptive approach. Goal: implementation for very high or infinite dimensionality.

Many applications in engineering.

Many other models to be studied:
(i) Non-affine dependence of $a$ in the variable $y$.
(ii) Other linear or non-linear PDE's.

Papers: www.ann.jussieu.fr/cohen

THANKS!