# Uncertainty Quantification and Approximation Theory for Parameterized PDEs

## Part III: Orthogonal polynomials and best $s$-term approximations

**Hoang A. Tran**

Clayton G. Webster, Guannan Zhang

Computer Science and Mathematics Division
Oak Ridge National Laboratory, USA

Supported by: **DOE** (ASCR)

Vietnam Institute for Advanced Study in Mathematics
November 15, 2016

# Outline

1. High dimensional parameterized PDEs
2. Global polynomial approximations
3. Best $s$-term approximations
4. Convergence analysis of best $s$-term approximations
5. Strategies for finding near optimal polynomial space

# Parameterized PDEs

parameters $\boldsymbol{y} \in \mathcal{U} \subset \mathbb{R}^d$ $\longrightarrow$ PDE model: $\mathcal{L}(u, \boldsymbol{y}) = f$ in $D \subset \mathbb{R}^n$, $n = 1, 2, 3$ $\longrightarrow$ quantity of interest $Q[u(\cdot, \boldsymbol{y})]$

- The operator $\mathcal{L}$, linear or nonlinear, depends on a vector of $d$ parameters $\boldsymbol{y} = (y_1, y_2, \ldots, y_d) \in \mathcal{U} = \prod_{i=1}^d \mathcal{U}_i$, which can be deterministic or stochastic.
- **Deterministic setting**: The parameters $\boldsymbol{y}$ are known or controlled by the user.
  - **Goal**: Given a query $\boldsymbol{y} \in \mathcal{U}$, quickly approximate the solution map $\boldsymbol{y} \mapsto u(\cdot, \boldsymbol{y})$.
- **Stochastic setting**: The parameters $\boldsymbol{y}$ may be affected by uncertainty (measurement error, incomplete description of parameters, unresolved scales), and are modeled as a random vector $\boldsymbol{y} : \Omega \to \mathcal{U}$ with joint PDF $\varrho(\boldsymbol{y}) = \prod_{i=1}^d \varrho_i(y_i)$
  - **Goal**: Uncertainty quantification of $u$ or some statistical QoI depending on $u$, i.e.,

$$\mathbb{E}[u],\ \mathbb{V}\mathrm{ar}[u],\ \mathbb{P}[u > u_0].$$

# An example problem: elliptic PDE

We consider the steady state diffusion equation

$$\begin{cases} -\nabla \cdot (a(x)\nabla u(x)) &= f(x) \quad \forall x \in D \subset \mathbb{R}^n, \\ u(x) &= 0 \quad \forall x \in \partial D. \end{cases} \tag{1}$$

where $f(x)$ is the given forcing term, $a(x)$ describes the diffusion coefficient and $u(x)$ is the solution.

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

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

Proof of the estimate: multiplying equation (1) by $u$ and integrate

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

# Parameterization

Let $\mathcal{U} := [-1,1]^d$, $\boldsymbol{y} = (y_1, \ldots, y_d) \in \mathcal{U}$.

Assume the diffusion coefficient $a = a(x, \boldsymbol{y})$ depends on the parametric variable $\boldsymbol{y}$.

We consider the following parameterized boundary value problem:

for all $\boldsymbol{y} \in \mathcal{U}$, find $u(\cdot, \boldsymbol{y}) : \overline{D} \to \mathbb{R}$ such that

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

Continuity and coercivity (CC) assumption:

$$0 < a_{\min} \leq a(x, \boldsymbol{y}) \leq a_{\max}, \quad \forall x \in \overline{D} \text{ and } \boldsymbol{y} \in \mathcal{U}.$$

Lax-Milgram ensures the existence and uniqueness of solution $u \in L^2(\mathcal{U}) \otimes \mathcal{V}$.

Analyticity (AN) assumption:

The complex continuation of $a$, represented as the map $a : \mathbb{C}^d \to L^\infty(D)$, is an $L^\infty(D)$-valued **analytic** function on $\mathbb{C}^d$.

# Parameterization

Let $\mathcal{U} := [-1,1]^d$, $\boldsymbol{y} = (y_1, \ldots, y_d) \in \mathcal{U}$.

Assume the diffusion coefficient $a = a(x, \boldsymbol{y})$ depends on the parametric variable $\boldsymbol{y}$.

We consider the following parameterized boundary value problem:

for all $\boldsymbol{y} \in \mathcal{U}$, find $u(\cdot, \boldsymbol{y}) : \overline{D} \to \mathbb{R}$ such that

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

## Continuity and coercivity (CC) assumption:

$$0 < a_{\min} \le a(x, \boldsymbol{y}) \le a_{\max}, \quad \forall x \in \overline{D} \text{ and } \boldsymbol{y} \in \mathcal{U}.$$

Lax-Milgram ensures the existence and uniqueness of solution $u \in L^2(\mathcal{U}) \otimes \mathcal{V}$.

## Analyticity (AN) assumption:

The complex continuation of $a$, represented as the map $a : \mathbb{C}^d \to L^\infty(D)$, is an $L^\infty(D)$-valued **analytic** function on $\mathbb{C}^d$.

# Parameterization

## Smoothness of the solutions

**Examples:** Affine and non-affine coefficients

1. $a(x, \boldsymbol{y}) = a_0(x) + \sum_{i=1}^{d} y_i \psi_i(x)$.
2. $a(x, \boldsymbol{y}) = a_0(x) + \left(\sum_{i=1}^{d} y_i \psi_i(x)\right)^p$, $p \in \mathbb{N}$.
3. $a(x, \boldsymbol{y}) = a_0(x) + \exp\left(\sum_{i=1}^{d} y_i \psi_i(x)\right)$

### Theorem

*Assume $a(x, \boldsymbol{y})$ satisfies* **CC** *and* **AN**. *Then*

$\boldsymbol{z} \mapsto u(\boldsymbol{z})$ *is well-defined and analytic in an open neighborhood of $\mathcal{U}$ in $\mathbb{C}^d$.*

**Key observation:**

- If $a(x, \boldsymbol{y})$ satisfies CC and AN, then for some open neighborhood $\widehat{\mathcal{U}}$ of $\mathcal{U}$ in $\mathbb{C}^d$,

$$0 < \delta \leq Re(a(x, \boldsymbol{z})), \quad \forall x \in \overline{D} \text{ and } \boldsymbol{z} \in \widehat{\mathcal{U}}.$$

- $u(\boldsymbol{z})$ is well-defined and analytic in $\widehat{\mathcal{U}}$.
- We call $\widehat{\mathcal{U}}$ the domain of uniform ellipticity.

# Domain of uniform ellipticity

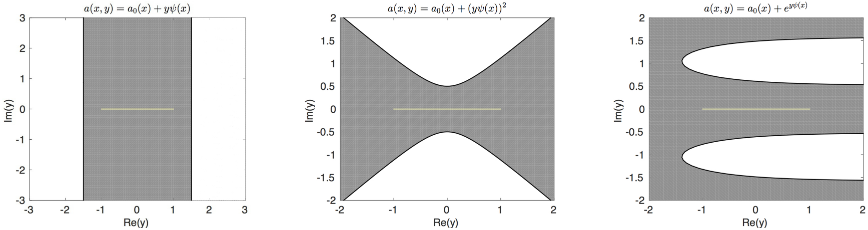
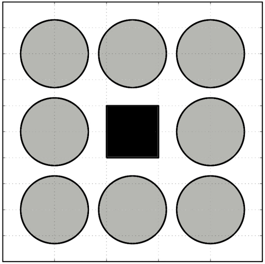



Figure : Domain of complex uniform ellipticity for some random fields.

# Example of parameterization

## Inclusion problem

Let $D = [0,1]^2$. Partitioning $D$ into 8 circular subdomains $D_1, D_2, \ldots, D_8$ arrayed around 1 square subdomain $F$.



Assume $a$ is a piecewise constant function over $D$ and on each $D_i$, the value of $a$ varies on $[c - c_i, c + c_i]$ for some $c$ and $c_i$ satisfying $0 < c_i < c$.

Then $a$ can be represented as

$$a(x, \boldsymbol{y}) = a_0(x) + \sum_{i=1}^{8} y_i \psi_i(x),$$

where $a_0(x) = c, \quad \psi_i = c_i \chi_{D_i}, \quad \boldsymbol{y} = (y_1, \ldots, y_8) \in \mathcal{U} = [-1,1]^8$.

# Global polynomial approximations

Observation: Parameterized solutions are often smooth with respect to the parametric variables.

Basic idea: approximate the response $u(\boldsymbol{y}, \cdot)$ by multi-variate global polynomials.

- Let $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_d) \in \mathbb{N}_0^d$ and $\Lambda \subset \mathbb{N}_0^d$ denote a multi-index set.
- Multivariate polynomial space:

$$\mathcal{P}_\Lambda(\mathcal{U}) = \text{span}\left\{ \prod_{i=1}^{d} {y_i}^{\nu_i}, \text{ with } \boldsymbol{\nu} \in \Lambda \right\}.$$

- Assume $N = \#\Lambda$ and $\{\boldsymbol{\Psi}_{\boldsymbol{\nu}}\}_{\boldsymbol{\nu}\in\Lambda}$ form a basis for $\mathcal{P}_\Lambda(\mathcal{U})$, e.g., multivariate Taylor, Legendre, Jacobi, Hermite, Lagrange, etc.
- Parametric discretization: $u^\# \in \mathcal{P}_\Lambda(\mathcal{U}) \otimes \mathcal{V}$ s.t.

$$u^\#(x, \boldsymbol{y}) = \sum_{\boldsymbol{\nu}\in\Lambda} c_{\boldsymbol{\nu}}(x) \boldsymbol{\Psi}_{\boldsymbol{\nu}}(\boldsymbol{y}).$$

Remarks:

- Take advantage of the smooth dependence of $u$ on parametric variables.
- Feature faster convergence than MC.
- The evaluation of $u^\#$ requires the computation of coefficients $c_{\boldsymbol{\nu}}$.
- We omit the spatial discretization for now.

# Taylor approximations

Consider the parameterized elliptic PDEs with $a(x,\boldsymbol{y}) = a_0(x) + \sum_{i=1}^{d} y_i \psi_i(x)$.

The Taylor expansion of $u(x,\boldsymbol{y})$:

$$u(x,\boldsymbol{y}) = \sum_{\boldsymbol{\nu}\in\mathbb{N}_0^d} t_{\boldsymbol{\nu}}(x)\boldsymbol{y}^{\boldsymbol{\nu}},$$

$$\text{where} \quad \boldsymbol{y}^{\boldsymbol{\nu}} := \prod_{i=1}^{d} y_i^{\nu_i} \quad \text{and} \quad t_{\boldsymbol{\nu}}(\,\cdot\,) := \frac{1}{\boldsymbol{\nu}!}\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u(\,\cdot\,,\mathbf{0}) \in \mathcal{V}.$$

Recursive formula for the Taylor coefficients:

with $\{\boldsymbol{e}_i\}_{1\le i\le d}$ the canonical basis of $\mathbb{R}^d$, the coefficient $t_{\boldsymbol{\nu}}$ is solution to

$$\int_D a_0 \nabla t_{\boldsymbol{\nu}} \nabla v = -\sum_{\{i:\,\nu_i\neq 0\}} \int_D \psi_i \nabla t_{\boldsymbol{\nu}-\boldsymbol{e}_i} \nabla v, \quad \forall v \in \mathcal{V}$$

Objective: approximate $u$ by the partial expansion

$$u_\Lambda(x,\boldsymbol{y}) = \sum_{\boldsymbol{\nu}\in\Lambda} t_{\boldsymbol{\nu}}(x)\boldsymbol{y}^{\boldsymbol{\nu}},$$

# Orthogonal polynomials

For stochastic Galerkin, discrete least square and compressed sensing approaches: we choose $\{\boldsymbol{\Psi}_{\boldsymbol{\nu}}\}_{\boldsymbol{\nu}\in\Lambda}$ as an **orthonormal basis**.

- Let $\varrho = \varrho(\boldsymbol{y})$ be a weight function on $\mathcal{U}$, i.e., nonnegative integrable in $\mathcal{U}$.
- $L^2(\mathcal{U}, d\varrho) := \left\{\boldsymbol{\Psi} : \mathcal{U} \to \mathbb{R},\ \int_{\mathcal{U}} \boldsymbol{\Psi}^2(\boldsymbol{y})\varrho(\boldsymbol{y})d\boldsymbol{y} < \infty\right\}$.
- Define the inner product and norm on $L^2(\mathcal{U}, d\varrho)$: for $h_1$, $h_2 \in L^2(\mathcal{U}, d\varrho)$,

$$\langle h_1, h_2\rangle := \int_{\mathcal{U}} h_1(\boldsymbol{y})h_2(\boldsymbol{y})\varrho(\boldsymbol{y})d\boldsymbol{y}, \qquad \|h_1\|_{\varrho} = \left(\int_{\mathcal{U}} |h_1(\boldsymbol{y})|^2\varrho(\boldsymbol{y})d\boldsymbol{y}\right)^{1/2}.$$

- The system $\{\boldsymbol{\Psi}_{\boldsymbol{\nu}}\}_{\boldsymbol{\nu}\in\Lambda}$ is called orthogonal with respect to $\varrho$ if

$$\langle \boldsymbol{\Psi}_{\boldsymbol{\nu}}, \boldsymbol{\Psi}_{\boldsymbol{\nu}'}\rangle = 0, \text{ if } \boldsymbol{\nu} \neq \boldsymbol{\nu}'.$$

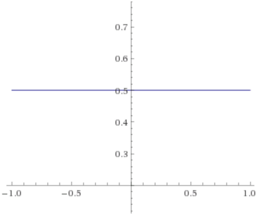
  If, in additional,

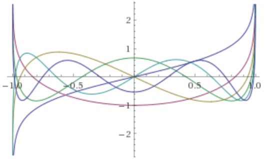
$$\langle \boldsymbol{\Psi}_{\boldsymbol{\nu}}, \boldsymbol{\Psi}_{\boldsymbol{\nu}}\rangle = 1 \text{ for all } \boldsymbol{\nu} \in \Lambda,$$

  then $\{\boldsymbol{\Psi}_{\boldsymbol{\nu}}\}_{\boldsymbol{\nu}\in\Lambda}$ is orthonormal.

# Orthogonal polynomials

## Example: Legendre polynomials




Univariate: The Legendre polynomials $\{L_j\}_{j=1}^{\infty}$ are orthogonal polynomials over the interval $\mathcal{U} = [-1, 1]$ with respect to the weight function $\varrho(y) = 1/2$

$$\int_{-1}^{1} L_j(y) L_{j'}(y) \frac{dy}{2} = 0, \quad \forall j \neq j'.$$

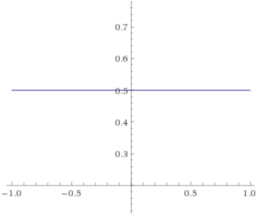
Multivariate: The multivariate Legendre polynomials $\{L_{\boldsymbol{\nu}}\}_{\boldsymbol{\nu} \in \mathbb{N}_0^d}$, given by

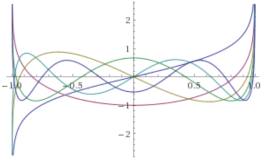
$$L_{\boldsymbol{\nu}}(\boldsymbol{y}) = \prod_{i=1}^{d} L_{\nu_i}(y_i), \qquad \forall \boldsymbol{\nu} \in \mathbb{N}_0^d,$$

are orthogonal polynomials over the domain $\mathcal{U} = [-1, 1]^d$ with respect to the weight function $\varrho(\boldsymbol{y}) = (1/2)^d$.

# Orthogonal polynomials

## Example: Legendre polynomials




Univariate: The Legendre polynomials $\{L_j\}_{j=1}^{\infty}$ are orthogonal polynomials over the interval $\mathcal{U} = [-1, 1]$ with respect to the weight function $\varrho(y) = 1/2$

$$\int_{-1}^{1} L_j(y) L_{j'}(y) \frac{dy}{2} = 0, \quad \forall j \neq j'.$$

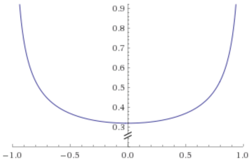
Multivariate: The multivariate Legendre polynomials $\{L_{\boldsymbol{\nu}}\}_{\boldsymbol{\nu} \in \mathbb{N}_0^d}$, given by

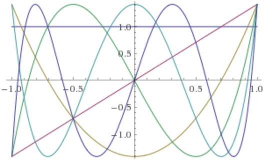
$$L_{\boldsymbol{\nu}}(\boldsymbol{y}) = \prod_{i=1}^{d} L_{\nu_i}(y_i), \quad \forall \boldsymbol{\nu} \in \mathbb{N}_0^d,$$

are orthogonal polynomials over the domain $\mathcal{U} = [-1, 1]^d$ with respect to the weight function $\varrho(\boldsymbol{y}) = (1/2)^d$.

# Orthogonal polynomials

## Example: Chebyshev polynomials




Univariate: The Chebyshev polynomials $\{T_j\}_{j=1}^{\infty}$ are orthogonal polynomials over the interval $\mathcal{U} = [-1,1]$ with respect to the weight function $\varrho(y) = \pi^{-1}(1-y^2)^{-1/2}$

$$\int_{-1}^{1} T_j(y) T_{j'}(y) \frac{dy}{\pi\sqrt{1-y^2}} = 0, \quad \forall j \neq j'.$$

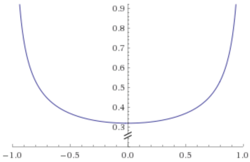
Multivariate: The multivariate Chebyshev polynomials $\{T_{\boldsymbol{\nu}}\}_{\boldsymbol{\nu}\in\mathbb{N}_0^d}$, given by

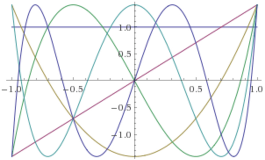
$$T_{\boldsymbol{\nu}}(\boldsymbol{y}) = \prod_{i=1}^{d} T_{\nu_i}(y_i), \qquad \forall \boldsymbol{\nu} \in \mathbb{N}_0^d,$$

are orthogonal polynomials over the domain $\mathcal{U} = [-1,1]^d$ with respect to the weight function $\varrho(\boldsymbol{y}) = \prod_{i=1}^{d} \varrho_i(y_i)$, where $\varrho_i(y_i) = \pi^{-1}(1-y_i^2)^{-1/2}$, $\forall i = 1, \dots, d$.

# Orthogonal polynomials

## Example: Chebyshev polynomials




Univariate: The Chebyshev polynomials $\{T_j\}_{j=1}^{\infty}$ are orthogonal polynomials over the interval $\mathcal{U} = [-1,1]$ with respect to the weight function $\varrho(y) = \pi^{-1}(1-y^2)^{-1/2}$

$$\int_{-1}^{1} T_j(y) T_{j'}(y) \frac{dy}{\pi\sqrt{1-y^2}} = 0, \quad \forall j \neq j'.$$

Multivariate: The multivariate Chebyshev polynomials $\{T_{\boldsymbol{\nu}}\}_{\boldsymbol{\nu}\in\mathbb{N}_0^d}$, given by

$$T_{\boldsymbol{\nu}}(\boldsymbol{y}) = \prod_{i=1}^{d} T_{\nu_i}(y_i), \qquad \forall \boldsymbol{\nu} \in \mathbb{N}_0^d,$$

are orthogonal polynomials over the domain $\mathcal{U} = [-1,1]^d$ with respect to the weight function $\varrho(\boldsymbol{y}) = \prod_{i=1}^{d} \varrho_i(y_i)$, where $\varrho_i(y_i) = \pi^{-1}(1-y_i^2)^{-1/2}$, $\forall i = 1,\dots,d$.

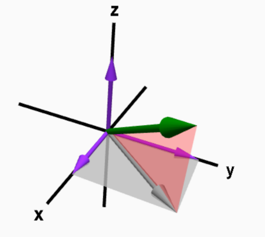
# Orthogonal polynomials

## The Askey scheme

| Distribution | Density function | Polynomial | Support |
|---|---|---|---|
| Normal | $\frac{1}{\sqrt{2\pi}}e^{\frac{-y^2}{2}}$ | Hermite $H_n(y)$ | $[-\infty, \infty]$ |
| Uniform | $\frac{1}{2}$ | Legendre $P_n(y)$ | $[-1, 1]$ |
| Beta | $\frac{(1-y)^\alpha(1+y)^\beta}{2^{\alpha+\beta+1}B(\alpha+1,\beta+1)}$ | Jacobi $P_n^{(\alpha,\beta)}(y)$ | $[-1, 1]$ |
| Exponential | $e^{-y}$ | Laguerre $L_n(y)$ | $[0, \infty]$ |
| Gamma | $\frac{y^\alpha e^{-y}}{\Gamma(\alpha+1)}$ | Generalized Laguerre $L_n^{(\alpha)}(y)$ | $[0, \infty]$ |

# Orthogonal polynomials

## Properties




1. For any function $h \in L^2(\mathcal{U}, d\varrho)$,

$$h(\boldsymbol{y}) = \sum_{\boldsymbol{\nu} \in \mathbb{N}_0^d} \widehat{h}_{\boldsymbol{\nu}} \boldsymbol{\Psi}_{\boldsymbol{\nu}}(\boldsymbol{y}), \quad \text{with } \widehat{h}_{\boldsymbol{\nu}} = \langle h, \boldsymbol{\Psi}_{\boldsymbol{\nu}} \rangle.$$

2. Parseval's equality:

$$\|h\|_{\varrho}^2 = \sum_{\boldsymbol{\nu} \in \mathbb{N}_0^d} |\widehat{h}_{\boldsymbol{\nu}}|^2.$$

3. For $\Lambda \subset \mathbb{N}_0^d$ and $h_\Lambda(\boldsymbol{y}) := \sum_{\boldsymbol{\nu} \in \Lambda} \widehat{h}_{\boldsymbol{\nu}} \boldsymbol{\Psi}_{\boldsymbol{\nu}}(\boldsymbol{y})$,

$$\|h - h_\Lambda\|_{\varrho} = \min_{q \in \mathcal{P}_\Lambda(\mathcal{U})} \|h - q\|_{\varrho}.$$

# Orthogonal polynomials

1. For a solution $u \in L^2(\mathcal{U}, \varrho) \otimes \mathcal{V}$,

$$u(x, \boldsymbol{y}) = \sum_{\boldsymbol{\nu} \in \mathbb{N}_0^d} \widehat{u}_{\boldsymbol{\nu}}(x) \boldsymbol{\Psi}_{\boldsymbol{\nu}}(\boldsymbol{y}), \quad \text{with } \widehat{u}_{\boldsymbol{\nu}}(x) = \langle u(x, \cdot), \boldsymbol{\Psi}_{\boldsymbol{\nu}} \rangle.$$

2. Parseval's equality:

$$\begin{aligned} \|u\|^2_{L^2(\mathcal{U},\varrho)\otimes\mathcal{V}} &= \int_{\mathcal{U}} \int_D |\nabla_x u(x, \boldsymbol{y})|^2 dx \varrho(\boldsymbol{y}) d\boldsymbol{y} \\ &= \int_D \int_{\mathcal{U}} \left( \sum_{\boldsymbol{\nu} \in \mathbb{N}_0^d} \nabla \widehat{u}_{\boldsymbol{\nu}}(x) \boldsymbol{\Psi}_{\boldsymbol{\nu}}(\boldsymbol{y}) \right)^2 \varrho(\boldsymbol{y}) d\boldsymbol{y} dx = \sum_{\boldsymbol{\nu} \in \mathbb{N}_0^d} \int_D |\nabla \widehat{u}_{\boldsymbol{\nu}}(x)|^2 dx \\ &= \sum_{\boldsymbol{\nu} \in \mathbb{N}_0^d} \|\widehat{u}\|^2_{\mathcal{V}}. \end{aligned}$$

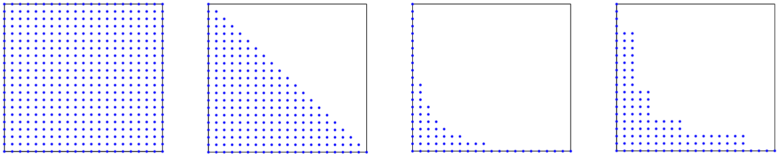
3. The best approximation of $u$ on $\mathcal{P}_\Lambda(\mathcal{U}) \otimes \mathcal{V}$ is $u_\Lambda(x, \boldsymbol{y}) := \sum_{\boldsymbol{\nu} \in \Lambda} \widehat{u}_{\boldsymbol{\nu}}(x) \boldsymbol{\Psi}_{\boldsymbol{\nu}}(\boldsymbol{y})$

$$\left\|u - u_\Lambda\right\|_{L^2(\mathcal{U},\varrho)\otimes\mathcal{V}} = \min_{q \in \mathcal{P}_\Lambda(\mathcal{U})\otimes\mathcal{V}} \left\|u - q\right\|_{L^2(\mathcal{U},\varrho)\otimes\mathcal{V}}.$$

What is the good polynomial approximation subspace?

# Polynomial space

## Selection of the index set



**Tensor Product** $\Lambda(w) = \{\nu \in \mathbf{N}^N : \max_{1\le i\le N} \nu_i \le w\}$

**Total Degree** $\Lambda(w) = \{\nu \in \mathbf{N}^N : \sum \nu_i \le w\}$

**Hyperbolic Cross** $\Lambda(w) = \{\nu \in \mathbf{N}^N : \prod(\nu_i + 1) \le w + 1\}$

**Smolyak** $\Lambda(w) = \{\nu \in \mathbf{N}^N : \sum f(\nu_i) \le f(w)\}$, with $f(\nu) = \lceil \log_2(\nu) \rceil$, $\nu \ge 2$.

Natural choice in 1d: $\Lambda = \{1, 2, 3, \dots, w\}$ and $u_\Lambda(x, \boldsymbol{y}) := \sum_{j=1}^{w} \widehat{u}_j(x) \boldsymbol{\Psi}_j(\boldsymbol{y})$.

Several choices for multi-index $\boldsymbol{\nu} \in \Lambda(w)$:

- Tensor product (TP): $\max_{1\le i\le d} \nu_i \le w$.
- Total degree (TD): $\sum_{i=1}^{d} \nu_i \le w$.
- Hyperbolic cross (HC): $\prod_{i=1}^{d} (\nu_i + 1) \le w + 1$.
- Smolyak (SM): $\sum_{i=1}^{d} f(\nu_i) \le f(w)\}$, with $f(\nu) = \lceil \log_2(\nu) \rceil$, $\nu \ge 2$.

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

$$\#\Lambda^{TP}(w) = (w+1)^d, \quad \#\Lambda^{TD}(w) = \frac{(d+w)!}{d!w!}, \quad \#\Lambda^{HC}(w) \lesssim \min\{w^3 4^d, w^{2+\log(d)}\}.$$

# Best $s$-term approximations

Observe the following estimates

- Taylor expansions:

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

- Legendre expansions:

$$\|u - u_\Lambda\|^2_{L^2(\mathcal{U},\varrho)\otimes\mathcal{V}} = \sum_{\boldsymbol{\nu}\notin\Lambda} \|\widehat{u}_{\boldsymbol{\nu}}\|^2_{\mathcal{V}}$$

**Best $s$-term approximations** [Chkifa, Cohen, DeVore, Schwab '10,'11,'13,'14]

choose $\Lambda = \Lambda_s^{\mathrm{opt}}$, the set of $s$ largest coefficients $\|t_{\boldsymbol{\nu}}\|_{\mathcal{V}}$ (or $\|\widehat{u}_{\boldsymbol{\nu}}\|_{\mathcal{V}}$).

- Theoretical approximation, generally inaccessible in practice.

# Stechkin estimate

**Lemma**

Assume $(\|t_{\boldsymbol{\nu}}\|_{\mathcal{V}})_{\boldsymbol{\nu}\in\mathbb{N}_0^d} \in \ell^p$ for some $0 < p < 1$, then

$$\sum_{\boldsymbol{\nu}\notin\Lambda_s^{\rm opt}} \|t_{\boldsymbol{\nu}}\|_{\mathcal{V}} \le C(p)s^{1-\frac{1}{p}}, \quad \text{where } C(p) = \|(\|t_{\boldsymbol{\nu}}\|_{\mathcal{V}})_{\boldsymbol{\nu}}\|_{\ell^p}.$$

$$\begin{aligned}\sum_{\boldsymbol{\nu}\notin\Lambda_s^{\rm opt}} \|t_{\boldsymbol{\nu}}\|_{\mathcal{V}} &= \sum_{\boldsymbol{\nu}\notin\Lambda_s^{\rm opt}} \|t_{\boldsymbol{\nu}}\|_{\mathcal{V}}^{1-p}\|t_{\boldsymbol{\nu}}\|_{\mathcal{V}}^{p} \\ &\le \sum_{\boldsymbol{\nu}\notin\Lambda_s^{\rm opt}} \max_{\boldsymbol{\nu}^*\notin\Lambda_s^{\rm opt}} \left(\|t_{\boldsymbol{\nu}^*}\|_{\mathcal{V}}^{1-p}\right)\|t_{\boldsymbol{\nu}}\|_{\mathcal{V}}^{p} = \max_{\boldsymbol{\nu}^*\notin\Lambda_s^{\rm opt}} \left(\|t_{\boldsymbol{\nu}^*}\|_{\mathcal{V}}^{1-p}\right)[C(p)]^p.\end{aligned}$$

We have

$$\max_{\boldsymbol{\nu}^*\notin\Lambda_s^{\rm opt}} \left(\|t_{\boldsymbol{\nu}^*}\|_{\mathcal{V}}^{p}\right) \le \left(\sum_{\boldsymbol{\nu}\in\Lambda_s^{\rm opt}} \|t_{\boldsymbol{\nu}}\|^p\right)/s \le \frac{[C(p)]^p}{s}$$

$$\Rightarrow \quad \max_{\boldsymbol{\nu}^*\notin\Lambda_s^{\rm opt}} \left(\|t_{\boldsymbol{\nu}^*}\|_{\mathcal{V}}^{1-p}\right) \le [C(p)]^{1-p}s^{1-1/p}. \quad QED.$$

- Similarly, $\left(\sum_{\boldsymbol{\nu}\notin\Lambda_s^{\rm opt}} \|\widehat{u}_{\boldsymbol{\nu}}\|_{\mathcal{V}}^2\right)^{1/2} \le C(p)s^{\frac{1}{2}-\frac{1}{p}}$.

# Analyticity of the solution revisited

- Polydisc: $\mathcal{O}_{\boldsymbol{\rho}} = \bigotimes_i \{z_i \in \mathbb{C}; |z_i| \leq \rho_i\}$.
- Polyellipse: $\mathcal{E}_{\boldsymbol{\rho}} = \bigotimes_i \left\{\frac{z_i + z_i^{-1}}{2}; z_i \in \mathbb{C}, |z_i| = \rho_i\right\}$.

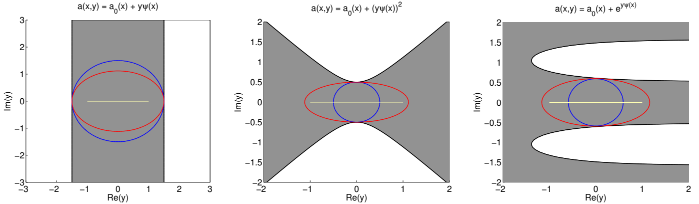
## Theorem

*Assume $a(x, \boldsymbol{y})$ is* **affine** *in $\boldsymbol{y}$ and satisfied* **CC**. *Then $\widehat{\mathcal{U}}$ contains some polydisc $\mathcal{O}_{\boldsymbol{\rho}}$, with $\boldsymbol{\rho} = (\rho_i)_{1 \leq i \leq N}$, $\rho_i > 1 \ \forall i$.*

- *$(\|t_{\boldsymbol{\nu}}\|_{\mathcal{V}})_{\boldsymbol{\nu}}$ $\ell^p$-summable for all $p < 1$.*

*Assume $a(x, \boldsymbol{y})$ satisfies* **CC** *and* **AN**. *Then the domain of uniform ellipticity $\widehat{\mathcal{U}}$ contains some polyellipse $\mathcal{E}_{\boldsymbol{\rho}}$, with $\boldsymbol{\rho} = (\rho_i)_{1 \leq i \leq N}$, $\rho_i > 1 \ \forall i$.*

- *$(\|\widehat{u}_{\boldsymbol{\nu}}\|_{\mathcal{V}})_{\boldsymbol{\nu}}$ $\ell^p$-summable for all $p < 1$.*



# Taylor coefficient estimate

## Proposition

*Assume* $\boldsymbol{z} \mapsto u(\boldsymbol{z})$ *is analytic in an open neighborhood of* $\mathcal{O}_{\boldsymbol{\rho}}$, *with* $\rho_i > 1 \ \forall i$, *and* $\Re(a(x, \boldsymbol{z})) \geq \delta > 0$, $\forall x \in \overline{D}$, $\boldsymbol{z} \in \mathcal{O}_{\boldsymbol{\rho}}$. *Then,* $\sum_{\boldsymbol{\nu} \in \mathbb{N}_0^d} t_{\boldsymbol{\nu}} \boldsymbol{y}^{\boldsymbol{\nu}}$ *converges uniformly towards* $u(\boldsymbol{y})$ *in* $\mathcal{U}$ *and*

$$\|t_{\boldsymbol{\nu}}\|_{\mathcal{V}} \leq C_\delta \boldsymbol{\rho}^{-\boldsymbol{\nu}}.$$

*where* $C_\delta = \frac{\|f\|_{\mathcal{V}'}}{\delta}$.

Simple case $d = 1$: applying Cauchy formula

$$u(\boldsymbol{z}) = \frac{1}{2i\pi} \int_{|\boldsymbol{z}'|=b} \frac{u(\boldsymbol{z}')}{\boldsymbol{z}' - \boldsymbol{z}} d\boldsymbol{z}' \quad \Rightarrow \quad u^{(n)}(\boldsymbol{z}) = \frac{n!}{2i\pi} \int_{|\boldsymbol{z}'|=b} \frac{u(\boldsymbol{z}')}{(\boldsymbol{z}' - \boldsymbol{z})^{n+1}} d\boldsymbol{z}'$$

$$\Rightarrow \quad |u^{(n)}(0)| \leq n!\, b^{-n} \max_{|\boldsymbol{z}| \leq b} |u(\boldsymbol{z})|$$

Multi-dimensional case $d > 1$:

$$\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u(\,\cdot\,, \boldsymbol{z}) = \frac{\boldsymbol{\nu}!}{(2i\pi)^d} \int_{|z_1'|=\rho_1} \dots \int_{|z_d'|=\rho_d} \frac{u(\,\cdot\,, z_1', \dots, z_d')}{(z_1' - z_1)^{\nu_1+1} \dots (z_d' - z_d)^{\nu_d+1}} dz_1' \dots dz_d'$$

# Legendre coefficient estimate

## Proposition

*Assume* $\boldsymbol{z} \mapsto u(\boldsymbol{z})$ *is analytic in an open neighborhood of* $\mathcal{E}_{\boldsymbol{\rho}}$, *with* $\rho_i > 1 \ \forall i$, *and* $\Re(a(x,\boldsymbol{z})) \geq \delta > 0$, $\forall x \in \overline{D}$, $\boldsymbol{z} \in \mathcal{E}_{\boldsymbol{\rho}}$. *Then,* $\sum_{\boldsymbol{\nu} \in \mathbb{N}_0^d} \widehat{u}_{\boldsymbol{\nu}} L_{\boldsymbol{\nu}}(\boldsymbol{y})$ *converges towards* $u(\boldsymbol{y})$ *in* $\mathcal{V} \otimes L^2(\mathcal{U})$ *and*

$$\|\widehat{u}_{\boldsymbol{\nu}}\|_{\mathcal{V}} \leq C_{\boldsymbol{\rho},\delta} \boldsymbol{\rho}^{-\boldsymbol{\nu}} \prod_{i=1}^{d} \sqrt{2\nu_i + 1},$$

*where* $C_{\boldsymbol{\rho},\delta} = \frac{\|f\|_{\mathcal{V}'}}{\delta} \prod_{i=1}^{N} \frac{\ell(\mathcal{E}_{\rho_i})}{4(\rho_i - 1)}$ *with* $\ell(\mathcal{E}_{\rho_i})$ *denoting the perimeter of* $\mathcal{E}_{\rho_i}$.

# Optimal coefficient upper bounds

Let $\mathcal{A}$ denote the set of all $(\boldsymbol{\rho}, \delta)$, $\rho_i > 1$, $\delta > 0$, such that the polydisc $\mathcal{O}_{\boldsymbol{\rho}}$ (polyellipse $\mathcal{E}_{\boldsymbol{\rho}}$) is contained in $\widehat{\mathcal{U}}$ and $\Re(a(x, \boldsymbol{z})) \geq \delta > 0$, $\forall x \in \overline{D}$, $\boldsymbol{z} \in \mathcal{O}_{\boldsymbol{\rho}}$ (correspondingly $\mathcal{E}_{\boldsymbol{\rho}}$). Then

$$\|t_{\boldsymbol{\nu}}\|_{\mathcal{V}} \leq C_\delta \boldsymbol{\rho}^{-\boldsymbol{\nu}} \quad \left( \text{or} \quad \|\widehat{u}_{\boldsymbol{\nu}}\|_{\mathcal{V}} \leq C_{\boldsymbol{\rho},\delta} \boldsymbol{\rho}^{-\boldsymbol{\nu}} \prod_{i=1}^{N} \sqrt{2\nu_i + 1} \right), \quad \forall (\boldsymbol{\rho}, \delta) \in \mathcal{A}.$$

Optimal Taylor/Legendre coefficient bounds

$$\inf_{(\boldsymbol{\rho},\delta)\in\mathcal{A}} C_\delta \boldsymbol{\rho}^{-\boldsymbol{\nu}} \quad \text{and} \quad \inf_{(\boldsymbol{\rho},\delta)\in\mathcal{A}} C_{\boldsymbol{\rho},\delta} \boldsymbol{\rho}^{-\boldsymbol{\nu}} \prod_{i=1}^{N} \sqrt{2\nu_i + 1}.$$

- Generally inaccessible.
- In case basis functions $\psi_i$ have **non-overlapping supports** ("inclusion problems"), $\boldsymbol{\rho}$ solving the minimization problems is independent of $\boldsymbol{\nu}$ and can be found easily.

Similar coefficient bounds were derived for nonlinear elliptic PDEs, initial value problems, parabolic PDEs [Hansen, Schwab '13; Hoang, Schwab '13 '14].

# Best $s$-term approximations

## Existing convergence estimate applied to finite dimensional parametric domains

Let $C(p) = \|(\|t_{\boldsymbol{\nu}}\|_{\mathcal{V}})_{\boldsymbol{\nu}}\|_{\ell^p}$ (or $\|(\|\widehat{u}_{\boldsymbol{\nu}}\|_{\mathcal{V}})_{\boldsymbol{\nu}}\|_{\ell^p}$).

Taylor: $\|u - u_{\Lambda_s^{\rm opt}}\|_{L^\infty(\mathcal{U},\mathcal{V})} \leq C(p) s^{1-\frac{1}{p}}$ for all $0 < p < 1$.

Legendre: $\left\|u - u_{\Lambda_s^{\rm opt}}\right\|_{L^2(\mathcal{U},\varrho)\otimes\mathcal{V}} \leq C(p) s^{\frac{1}{2}-\frac{1}{p}}$ for all $0 < p < 1$.

- Legendre approximation features faster convergence rate than that of Taylor expansion.
- $p$ small: stronger convergence rate, but bigger $C(p)$.
- $C(p)$ is implicit. Realizable estimate:
  bound $C(p)$ by $\|(B(\boldsymbol{\nu}))_{\boldsymbol{\nu}}\|_{\ell^p}$, where $B(\boldsymbol{\nu})$ is the upper bound of $\|t_{\boldsymbol{\nu}}\|_{\mathcal{V}}$ (or $\|\widehat{u}_{\boldsymbol{\nu}}\|_{\mathcal{V}}$).

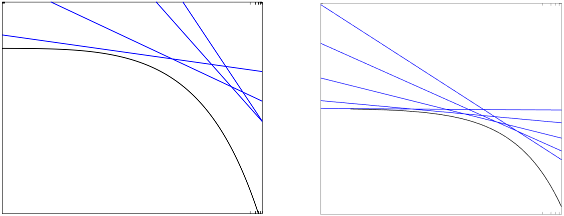


Figure : Estimate the truncation error $\sum_{\boldsymbol{\nu}\in\Lambda_s^{\rm opt}} \boldsymbol{\rho}^{-\boldsymbol{\nu}}$.

# A new convergence analysis

- Assume $\|c_{\boldsymbol{\nu}}\|_{\mathcal{V}} \le B(\boldsymbol{\nu}) =: e^{-b(\boldsymbol{\nu})}$, we have

$$\sup_{\boldsymbol{y}\in\mathcal{U}} \left\| u - \sum_{\boldsymbol{\nu}\in\Lambda_s^{\text{opt}}} c_{\boldsymbol{\nu}} \boldsymbol{\Psi}_{\boldsymbol{\nu}} \right\|_{\mathcal{V}} \le \sum_{\boldsymbol{\nu}\notin\Lambda_s^{\text{opt}}} \|c_{\boldsymbol{\nu}}\|_{\mathcal{V}} \le \sum_{\boldsymbol{\nu}\notin\Lambda_s^{\text{Qopt}}} e^{-b(\boldsymbol{\nu})},$$

where $\Lambda_s^{\text{Qopt}}$ is the quasi-optimal index set corresponding to $s$ largest coefficient bounds.

- Split $\mathbb{N}_0^d \setminus \Lambda_s^{\text{Qopt}}$ into a family $(\mathcal{Q}_j)_{j\in\mathbb{N},\, j\ge J}$ of disjoint subsets of $\mathbb{N}^d$ based on values of $e^{-b(\boldsymbol{\nu})}$, where $\mathcal{Q}_j$ contains $\boldsymbol{\nu}$ satisfying $e^{-j-1} \le e^{-b(\boldsymbol{\nu})} < e^{-j}$, so that the truncation error can be bounded as

$$\sum_{\boldsymbol{\nu}\notin\Lambda_s^{\text{Qopt}}} e^{-b(\boldsymbol{\nu})} = \sum_{j\ge J} \sum_{\boldsymbol{\nu}\in\mathcal{Q}_j} e^{-b(\boldsymbol{\nu})} \le \sum_{j\ge J} \#(\mathcal{Q}_j)\cdot e^{-j}.$$

**Central task**: Find a sharp estimate of $\#(\mathcal{Q}_j)$.

- We can instead estimate $\#(\mathcal{P}_j \cap \mathbb{Z}^d)$, the number of integer points in $\mathcal{P}_j$, where

$$\mathcal{P}_j := \{\boldsymbol{\nu} \in [0,\infty)^d : e^{-b(\boldsymbol{\nu})} \ge e^{-j}\}$$
$$= \{\boldsymbol{\nu} \in [0,\infty)^d : b(\boldsymbol{\nu}) \le j\},$$

as $\#(\mathcal{Q}_j) = \#(\mathcal{P}_{j+1} \cap \mathbb{Z}^d) - \#(\mathcal{P}_j \cap \mathbb{Z}^d)$.

# A new convergence analysis

- Assume $\|c_{\boldsymbol{\nu}}\|_{\mathcal{V}} \le B(\boldsymbol{\nu}) =: e^{-b(\boldsymbol{\nu})}$, we have

$$\sup_{\boldsymbol{y}\in\mathcal{U}} \left\| u - \sum_{\boldsymbol{\nu}\in\Lambda_s^{\rm opt}} c_{\boldsymbol{\nu}} \boldsymbol{\Psi}_{\boldsymbol{\nu}} \right\|_{\mathcal{V}} \le \sum_{\boldsymbol{\nu}\notin\Lambda_s^{\rm opt}} \|c_{\boldsymbol{\nu}}\|_{\mathcal{V}} \le \sum_{\boldsymbol{\nu}\notin\Lambda_s^{\rm Qopt}} e^{-b(\boldsymbol{\nu})},$$

  where $\Lambda_s^{\rm Qopt}$ is the quasi-optimal index set corresponding to $s$ largest coefficient bounds.

- Split $\mathbb{N}_0^d \setminus \Lambda_s^{\rm Qopt}$ into a family $(\mathcal{Q}_j)_{j\in\mathbb{N},\, j\ge J}$ of disjoint subsets of $\mathbb{N}^d$ based on values of $e^{-b(\boldsymbol{\nu})}$, where $\mathcal{Q}_j$ contains $\boldsymbol{\nu}$ satisfying $e^{-j-1} \le e^{-b(\boldsymbol{\nu})} < e^{-j}$, so that the truncation error can be bounded as

$$\sum_{\boldsymbol{\nu}\notin\Lambda_s^{\rm Qopt}} e^{-b(\boldsymbol{\nu})} = \sum_{j\ge J}\sum_{\boldsymbol{\nu}\in\mathcal{Q}_j} e^{-b(\boldsymbol{\nu})} \le \sum_{j\ge J} \#(\mathcal{Q}_j)\cdot e^{-j}.$$

  **Central task**: Find a sharp estimate of $\#(\mathcal{Q}_j)$.

- We can instead estimate $\#(\mathcal{P}_j \cap \mathbb{Z}^d)$, the number of integer points in $\mathcal{P}_j$, where

$$\mathcal{P}_j := \{\boldsymbol{\nu}\in[0,\infty)^d : e^{-b(\boldsymbol{\nu})} \ge e^{-j}\} = \{\boldsymbol{\nu}\in[0,\infty)^d : b(\boldsymbol{\nu}) \le j\},$$

  as $\#(\mathcal{Q}_j) = \#(\mathcal{P}_{j+1}\cap\mathbb{Z}^d) - \#(\mathcal{P}_j\cap\mathbb{Z}^d)$.

# A new convergence analysis

- Assume $\|c_{\boldsymbol{\nu}}\|_{\mathcal{V}} \le B(\boldsymbol{\nu}) =: e^{-b(\boldsymbol{\nu})}$, we have

$$\sup_{\boldsymbol{y}\in\mathcal{U}} \left\| u - \sum_{\boldsymbol{\nu}\in\Lambda_s^{\mathrm{opt}}} c_{\boldsymbol{\nu}} \boldsymbol{\Psi}_{\boldsymbol{\nu}} \right\|_{\mathcal{V}} \le \sum_{\boldsymbol{\nu}\notin\Lambda_s^{\mathrm{opt}}} \|c_{\boldsymbol{\nu}}\|_{\mathcal{V}} \le \sum_{\boldsymbol{\nu}\notin\Lambda_s^{\mathrm{Qopt}}} e^{-b(\boldsymbol{\nu})},$$

  where $\Lambda_s^{\mathrm{Qopt}}$ is the quasi-optimal index set corresponding to $s$ largest coefficient bounds.

- Split $\mathbb{N}_0^d \setminus \Lambda_s^{\mathrm{Qopt}}$ into a family $(\mathcal{Q}_j)_{j\in\mathbb{N},\, j\ge J}$ of disjoint subsets of $\mathbb{N}^d$ based on values of $e^{-b(\boldsymbol{\nu})}$, where $\mathcal{Q}_j$ contains $\boldsymbol{\nu}$ satisfying $e^{-j-1} \le e^{-b(\boldsymbol{\nu})} < e^{-j}$, so that the truncation error can be bounded as

$$\sum_{\boldsymbol{\nu}\notin\Lambda_s^{\mathrm{Qopt}}} e^{-b(\boldsymbol{\nu})} = \sum_{j\ge J} \sum_{\boldsymbol{\nu}\in\mathcal{Q}_j} e^{-b(\boldsymbol{\nu})} \le \sum_{j\ge J} \#(\mathcal{Q}_j)\cdot e^{-j}.$$

  **Central task**: Find a sharp estimate of $\#(\mathcal{Q}_j)$.

- We can instead estimate $\#(\mathcal{P}_j \cap \mathbb{Z}^d)$, the number of integer points in $\mathcal{P}_j$, where

$$\mathcal{P}_j := \{\boldsymbol{\nu} \in [0,\infty)^d : e^{-b(\boldsymbol{\nu})} \ge e^{-j}\}$$
$$= \{\boldsymbol{\nu} \in [0,\infty)^d : b(\boldsymbol{\nu}) \le j\},$$

  as $\#(\mathcal{Q}_j) = \#(\mathcal{P}_{j+1}\cap\mathbb{Z}^d) - \#(\mathcal{P}_j\cap\mathbb{Z}^d)$.

# Asymptotic upper estimate

**Idea:** Connect $\#(\mathcal{P}_j \cap \mathbb{Z}^d)$ and the continuous volume (Lebesgue measure) of $\mathcal{P}_j$

**Lemma [Gruber '07]**

Suppose $\mathcal{P} \subset \mathbb{R}^d$ is a bounded Jordan measurable set. For $j \in \mathbb{N}$, $j > 0$, there holds

$$|\mathcal{P}| = \lim_{j\to\infty} \frac{1}{j^d} \cdot \#(\mathcal{P} \cap \frac{1}{j}\mathbb{Z}^d) = \lim_{j\to\infty} \frac{1}{j^d} \cdot \#(j\mathcal{P} \cap \mathbb{Z}^d).$$

Observe: If $b(\boldsymbol{\nu})$ is defined such that $\mathcal{P}_j = j\mathcal{P}$, $\forall j \in \mathbb{N}$, with some $\mathcal{P} \subset \mathbb{R}^d$, one obtains a simple asymptotic formula

$$\#(\mathcal{P}_j \cap \mathbb{Z}^d) \simeq j^d |\mathcal{P}|.$$

# Asymptotic upper estimate

**Lemma [T, Webster, Zhang '14]**

Suppose $(\mathcal{P}_\tau)_{\tau\in\mathbb{R}^+}$ is a family of bounded Lebesgue measurable sets in $\mathbb{R}^d$ such that

$$\frac{1}{\tau_1}\mathcal{P}_{\tau_1} \subset \frac{1}{\tau_2}\mathcal{P}_{\tau_2},\ \forall\tau_1 \geq \tau_2 > 0, \quad (2)$$

or

$$\frac{1}{\tau_1}\mathcal{P}_{\tau_1} \supset \frac{1}{\tau_2}\mathcal{P}_{\tau_2},\ \forall\tau_1 \geq \tau_2 > 0. \quad (3)$$

Denote $\mathcal{P} = \bigcap_{\tau\in\mathbb{R}^+} \frac{1}{\tau}\mathcal{P}_\tau$ if (2) holds and $\mathcal{P} = \bigcup_{\tau\in\mathbb{R}^+} \frac{1}{\tau}\mathcal{P}_\tau$ otherwise. If $\mathcal{P}$ is bounded Jordan measurable, $|\mathcal{P}| > 0$, and $\mathcal{P}_\tau$ is Jordan measurable for countably infinite $\tau \in (0,\infty)$, there follows

$$|\mathcal{P}| = \lim_{\tau\to\infty} \frac{1}{\tau^d} \cdot \#(\mathcal{P}_\tau \cap \mathbb{Z}^d).$$

- $\#(\mathcal{P}_\tau \cap \mathbb{Z}^d) \simeq \tau^d|\mathcal{P}|$ .
- **Which $b(\boldsymbol{\nu})$ makes the sublevel sets $(\mathcal{P}_\tau)_{\tau\in\mathbb{R}^+}$ satisfy the above lemma?**

# Asymptotic upper estimate

Jordan measurability $\Leftarrow$ Continuity of $b$.

Ascending/descending property of $\left(\frac{1}{\tau}\mathcal{P}_\tau\right)_{\tau\in\mathbb{R}^+}$ $\Leftarrow$ Monotonicity of $\tau \mapsto \frac{1}{\tau}b(\tau\boldsymbol{\nu})$.

$\mathcal{P}$ is bounded and not null $\Leftarrow B(\mathbf{0}, 1/C) \subset \mathcal{P} \subset B(\mathbf{0}, 1/c)$.

## Assumption on coefficient bounds (AoB)

The map $b : [0,\infty)^d \to \mathbb{R}$ satisfies

1. $b(\mathbf{0}) = 0$ and $b$ is **continuous** in $[0,\infty)^d$,
2. $H_{\boldsymbol{\nu}} : \tau \mapsto \frac{1}{\tau}b(\tau\boldsymbol{\nu})$ is either **increasing** in $(0,\infty)$ for all $\boldsymbol{\nu} \in [0,\infty)^d$ or **decreasing** in $(0,\infty)$ for all $\boldsymbol{\nu} \in [0,\infty)^d$,
3. $b(\boldsymbol{\nu}) \in \Theta(|\boldsymbol{\nu}|)$. In other words, there exists $0 < c < C$ such that $c|\boldsymbol{\nu}| < b(\boldsymbol{\nu}) < C|\boldsymbol{\nu}|$ as $\boldsymbol{\nu} \to \infty$.

Given $\boldsymbol{\rho} = (\rho_i)_{1\le i\le d}$ with $\rho_i > 1\ \forall i$, we define $\boldsymbol{\lambda} = (\lambda_i)_{1\le i\le d}$ such that $\lambda_i = \log \rho_i\ \forall i$.

**AoB** is satisfied by

- $B(\boldsymbol{\nu}) = \boldsymbol{\rho}^{-\boldsymbol{\nu}}$; $\qquad b(\boldsymbol{\nu}) = \sum_{i=1}^{d} \lambda_i \nu_i$.
- $B(\boldsymbol{\nu}) = \boldsymbol{\rho}^{-\boldsymbol{\nu}} \prod_{i=1}^{d} \sqrt{2\nu_i + 1}$; $\qquad b(\boldsymbol{\nu}) = \sum_{i=1}^{d} \left(\lambda_i \nu_i - \frac{1}{2}\log(2\nu_i + 1)\right)$.

  [Beck, Nobile, Tamellini, Tempone '14]
- $B(\boldsymbol{\nu}) = \inf_{\boldsymbol{\rho},\delta} C_\delta \boldsymbol{\rho}^{-\boldsymbol{\nu}}$; $\qquad b(\boldsymbol{\nu}) = \sup_{\boldsymbol{\rho},\delta} \left( \sum_{i=1}^{d} \lambda_i \nu_i - \log C_\delta \right)$.
- $B(\boldsymbol{\nu}) = \inf_{\boldsymbol{\rho},\delta} C_{\rho,\delta} \boldsymbol{\rho}^{-\boldsymbol{\nu}} \prod_{i=1}^{d} \sqrt{2\nu_i + 1}$; $\qquad b(\boldsymbol{\nu}) = \sup_{\boldsymbol{\rho},\delta} \left( \sum_{i=1}^{d} \left(\lambda_i \nu_i - \frac{1}{2}\log(2\nu_i + 1)\right) - \log C_{\rho,\delta} \right)$.

  [Cohen, DeVore, Schwab '11]
- $B(\boldsymbol{\nu}) = \boldsymbol{\rho}^{-\boldsymbol{\nu}} \frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!}$; $\qquad b(\boldsymbol{\nu}) = \sum_{i=1}^{d} \lambda_i \nu_i - \log \frac{\Gamma(|\boldsymbol{\nu}|+1)}{\prod_{i=1}^{d} \Gamma(\nu_i+1)}$.

  [Beck, Tempone, Nobile, Tamellini '12]

> Current analysis has been limited to $b(\boldsymbol{\nu}) = \sum_{i=1}^{d} \lambda_i \nu_i$ [BNTT '14].

# Asymptotic upper estimate

Under **AoB**,

$$\sum_{\boldsymbol{\nu} \notin \Lambda_M^{\mathrm{Qopt}}} e^{-b(\boldsymbol{\nu})} \leq \sum_{j \geq J} (\#(\mathcal{P}_{j+1} \cap \mathbb{Z}^d) - \#(\mathcal{P}_j \cap \mathbb{Z}^d)) \cdot e^{-j}$$

$$\leq \sum_{j \geq J} \left[ (1+\varepsilon)(j+1)^d |\mathcal{P}| - \frac{1}{2} j^d |\mathcal{P}| \right] e^{-j}$$

$$\lesssim |\mathcal{P}| \sum_{j \geq J} j^d e^{-j}.$$

## Lemma

For any $d, J \in \mathbb{N}$, if $J \geq \frac{2}{e^{1/d}-1}$, it gives

$$\sum_{j \geq J} j^d e^{-j} \leq 2 J^d e^{-J} \frac{e}{e-1}.$$

Lower bound: $\sum_{j \geq J} j^d e^{-j} \geq J^d e^{-J} \frac{e}{e-1}, \forall J \in \mathbb{N}.$

# Main result

## Main theorem [T, Webster, Zhang '14]

Consider the multi-indexed series $\sum_{\boldsymbol{\nu}\in\mathbb{N}_0^d} e^{-b(\boldsymbol{\nu})}$ with $b : [0,\infty)^d \to \mathbb{R}$ satisfying **AoB**. Let $\Lambda_s^{\mathrm{Qopt}}$ be the set of indices corresponding to $s$ largest $e^{-b(\boldsymbol{\nu})}$. Then, there exists $s_\varepsilon > 0$ depending on $\varepsilon$ such that

$$\sum_{\boldsymbol{\nu}\notin\Lambda_s^{\mathrm{Qopt}}} e^{-b(\boldsymbol{\nu})} \le C_u(\varepsilon)\, s \exp\left(-\left(\frac{s}{|\mathcal{P}|(1+\varepsilon)}\right)^{1/d}\right)$$

for all $s > s_\varepsilon$. Here, $C_u(\varepsilon) = (4e + 4\varepsilon e - 2)\,\frac{e}{e-1}$.

- We achieve **sub-exponential** convergence rates of the form $s\exp(-(\kappa s)^{1/d})$, with **optimized** $\kappa$.
- $|\mathcal{P}|$ can be determined computationally
  1. $\mathcal{P} = \left\{\boldsymbol{\nu}\in[0,\infty)^d : \sum_{i=1}^d \lambda_i\nu_i \le 1\right\}$, for $B(\boldsymbol{\nu}) = \boldsymbol{\rho}^{-\boldsymbol{\nu}}\prod_{i=1}^d\sqrt{2\nu_i+1}$.
  2. $\mathcal{P} = \left\{\boldsymbol{\nu}\in[0,\infty)^d : \sum_{i=1}^d \lambda_i\nu_i \le 1 \;\; \forall(\boldsymbol{\rho},\delta)\in\mathcal{A}\right\}$, for $B(\boldsymbol{\nu}) = \inf_{\boldsymbol{\rho},\delta} C_\delta\boldsymbol{\rho}^{-\boldsymbol{\nu}}$.
  3. $\mathcal{P} = \{\boldsymbol{\nu}\in(0,\infty)^d : \sum_{i=1}^d \lambda_i\nu_i - \log\frac{|\boldsymbol{\nu}|^{|\boldsymbol{\nu}|}}{\prod_{i=1}^d {\nu_i}^{\nu_i}} < 1\}$, for $B(\boldsymbol{\nu}) = \boldsymbol{\rho}^{-\boldsymbol{\nu}}\frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!}$.
- **Stronger** rates are realized at **larger** cardinality.

# Comparison with existing estimates

**Proposition: Best $s$-term Taylor expansion [T, Webster, Zhang '14]**

Consider the Taylor series $\sum_{\boldsymbol{\nu}\in\mathbb{N}_0^d} t_{\boldsymbol{\nu}}\boldsymbol{y}^{\boldsymbol{\nu}}$ of $u$. Recall that

$$\|t_{\boldsymbol{\nu}}\|_{\mathcal{V}} \leq C\boldsymbol{\rho}^{-\boldsymbol{\nu}}, \quad \forall \boldsymbol{\nu} \in \mathbb{N}_0^d. \tag{4}$$

Denote by $\Lambda_s^{\mathrm{opt}}$ the set of indices corresponding to $s$ largest coefficients. For any $\varepsilon > 0$, there exists $s_\varepsilon > 0$ depending on $\varepsilon$ such that

$$\sup_{\boldsymbol{y}\in\mathcal{U}} \left\| u(\boldsymbol{y}) - \sum_{\boldsymbol{\nu}\in\Lambda_s^{\mathrm{opt}}} t_{\boldsymbol{\nu}}\boldsymbol{y}^{\boldsymbol{\nu}} \right\|_{\mathcal{V}} \leq C_u(\varepsilon)s\exp\left( -\left( \frac{sd!\prod_{i=1}^d \lambda_i}{(1+\varepsilon)} \right)^{1/d} \right)$$

for all $s > s_\varepsilon$.

**Previous rates:**

1. Applying on Stechkin estimation [CDS '11] to our setting: $\left(\prod_{i=1}^{d} \frac{1}{1-e^{-p\lambda_i}}\right)^{1/p} s^{1-\frac{1}{p}}$.

   This rate is **non-asymptotic** and applicable for **infinite** dimensional parameter space.
2. Optimization of Stechkin rate [BNTT '14]: $s\exp\left(-\frac{1}{e}\left(s\prod_{i=1}^d \lambda_i\right)^{1/d} d\xi\right)$.

   $\xi \nearrow \frac{e-1}{e} \simeq 0.63$ as $s \nearrow \infty$.

# Numerical illustrations

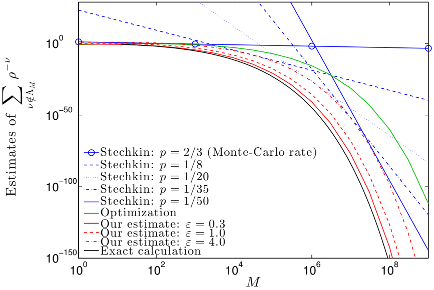
## Example 1

Estimate the truncation error of $\sum_{\boldsymbol{\nu}\in\mathbb{N}_0^4} \boldsymbol{\rho}^{-\boldsymbol{\nu}}$, where

$$\rho_1 = \rho_2 = 2,\ \rho_3 = 4,\ \rho_4 = 16.$$

This problem arises in error analysis of best $s$-term Taylor approximations for parameterized elliptic PDEs with **non-overlapping** basis functions.

- The best theoretical coefficient bound has the form: $\|t_{\boldsymbol{\nu}}\|_{\mathcal{V}} \leq \boldsymbol{\rho}^{-\boldsymbol{\nu}}$.
- **Anisotropic** case, 4-dimensional parametric domain.

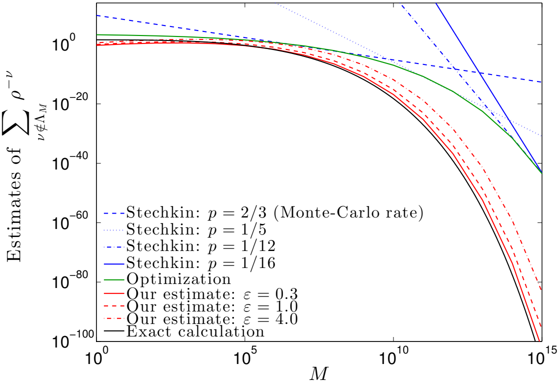


# Numerical illustrations

## Example 2

Estimate the truncation error of $\sum_{\boldsymbol{\nu}\in\mathbb{N}_0^8} \boldsymbol{\rho}^{-\boldsymbol{\nu}}$, where $\rho_i = 2, \quad \forall 1 \le i \le 8.$

- **Isotropic** case, 8-dimensional parametric domain.



# Lower bound

- **Goal:** derive a lower bound of $\sum_{\boldsymbol{\nu} \notin \Lambda_s^{\mathrm{Qopt}}} e^{-b(\boldsymbol{\nu})}$.
- If exploiting $\#(\mathcal{P}_\tau \cap \mathbb{Z}^d) \simeq \tau^d |\mathcal{P}|$ :

$$\sum_{\boldsymbol{\nu} \notin \Lambda_s^{\mathrm{Qopt}}} e^{-b(\boldsymbol{\nu})} \geq \sum_{j \geq J} (\#(\mathcal{P}_{j+1} \cap \mathbb{Z}^d) - \#(\mathcal{P}_j \cap \mathbb{Z}^d)) \cdot e^{-j-1}$$

$$\geq \sum_{j \geq J} \left[ (1-\varepsilon)(j+1)^d |\mathcal{P}| - (1+\varepsilon) j^d |\mathcal{P}| \right] e^{-j-1}$$

- However, $(1-\varepsilon)(j+1)^d < (1+\varepsilon) j^d$. **AoB is not enough!**

# Lower bound

## A simplified case

1. $\mathcal{P}$ is a **rational** convex polytope,
2. $\mathcal{P}_\tau = \tau\mathcal{P}$ for all $\tau \in (0, \infty)$.

- This setting arises from the multi-indexed sequence $(e^{-b(\boldsymbol{\nu})})_{\boldsymbol{\nu}\in\mathbb{N}_0^d}$ with

$$b(\boldsymbol{\nu}) = \sup_{\boldsymbol{\lambda}\in\mathcal{A}} \Big(\sum_{i=1}^{d} \lambda_i \nu_i\Big), \text{ where } \mathcal{A} \text{ is a } \textbf{finite} \text{ subset of } (\mathbb{Q}^+)^d,$$

- Accommodate **Taylor** coefficient estimates.

## Lemma: Ehrhart polynomial (LEp) [Gruber '07]

- Let $\mathcal{P}$ is an $d$-dimensional convex polytope with integer vertices. There exist rational numbers $c_0, \dots, c_{d-1}$ such that

$$\#(j\mathcal{P} \cap \mathbb{Z}^d) = |\mathcal{P}| j^d + c_{d-1} j^{d-1} + \dots + c_0, \ \ \forall j \in \mathbb{N}. \ \textbf{(Ehrhart polynomial)}$$

- If, instead, $\mathcal{P}$ is a rational convex polytope, the above assertion holds with $c_i$'s are periodic functions with integral periods.

# Lower bound

Applying **LEp**:

$$\sum_{\boldsymbol{\nu}\notin\Lambda_s^{\mathrm{Qopt}}} e^{-b(\boldsymbol{\nu})} \geq \sum_{j\geq J}(\#(\mathcal{P}_{j+1}\cap\mathbb{Z}^d)-\#(\mathcal{P}_j\cap\mathbb{Z}^d))\cdot e^{-j-1}$$

$$\gtrsim d|\mathcal{P}|\sum_{j\geq J} j^{d-1}e^{-j-1}.$$

**Theorem: Asymptotic lower bound [T, Webster, Zhang '14]**

There exists a constant $s^*>0$ such that

$$\sum_{\boldsymbol{\nu}\notin\Lambda_s^{\mathrm{Qopt}}} e^{-b(\boldsymbol{\nu})} \geq C_\ell\, s^{1-\frac{1}{d}}\exp\left(-\left(\frac{s}{|\mathcal{P}|}\right)^{1/d}\right)$$

for all $s>s^*$. Here, $C_\ell=\frac{1}{2}\left(\frac{2}{3}\right)^{1-\frac{1}{d}}\frac{d|\mathcal{P}|^{\frac{1}{d}}q}{e^q-1}$ where $q$ is the period of $c_i$'s.

Recall the **upper bound** of $\sum_{\boldsymbol{\nu}\notin\Lambda_s^{\mathrm{Qopt}}} e^{-b(\boldsymbol{\nu})}$:

$$\sum_{\boldsymbol{\nu}\notin\Lambda_s^{\mathrm{Qopt}}} e^{-b(\boldsymbol{\nu})} \leq C_u(\varepsilon)\, s\exp\left(-\left(\frac{s}{|\mathcal{P}|(1+\varepsilon)}\right)^{1/d}\right).$$

# Pre-asymptotic analysis

Two **asymptotic** estimates were employed (and need to be replaced in pre-asymptotic analysis):

$$(1-\varepsilon)j^d|\mathcal{P}| \leq \#(\mathcal{P}_j \cap \mathbb{Z}^d) \leq (1+\varepsilon)j^d|\mathcal{P}| \quad \text{if} \quad s > s_\varepsilon{}^{\ddagger}, \tag{5}$$

and

$$\sum_{j \geq J} j^d e^{-j} \leq 2J^d e^{-J} \frac{e}{e-1} \quad \text{if} \quad J \geq \frac{2}{e^{1/d}-1} \simeq d. \tag{6}$$

**Lemma: Non-asymptotic estimate of $\#(\mathcal{P}_j \cap \mathbb{Z}^d)$**

Assume that $b : [0,\infty)^d \to \mathbb{R}$ is continuous and

- $b(\boldsymbol{\nu}) \leq b(\boldsymbol{\mu})$ for all $\boldsymbol{\nu}, \boldsymbol{\mu} \in [0,\infty)^d$ such that $\boldsymbol{\nu} \leq \boldsymbol{\mu}$,
- $b(\tau\boldsymbol{\nu}) = \tau b(\boldsymbol{\nu})$ for all $\tau \in (0,\infty)$, $\boldsymbol{\nu} \in [0,\infty)^d$,

then

$$j^d|\mathcal{P}| \leq \#(\mathcal{P}_j \cap \mathbb{Z}^d) \leq j^d \cdot \#(\mathcal{P} \cap \mathbb{Z}^d), \quad \forall j \in \mathbb{N}.$$

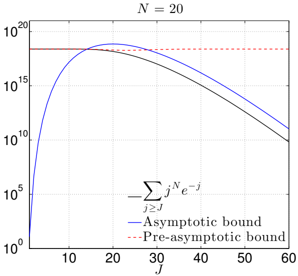
‡ *can be specified computationally by means of Ehrhart polynomials.*

# Pre-asymptotic analysis

- **Task:** Establish an upper bound of $\sum_{j\geq J} j^d e^{-j}$ for $J \leq d$.
- Observe that $\sum_{j<J} j^d e^{-j} \ll \sum_{j\geq J} j^d e^{-j}$ for small $J$, we estimate $\sum_{j<J} j^d e^{-j}$ instead:

$$\sum_{j\geq J} j^d e^{-j} \leq \sum_{j=1}^{\infty} j^d e^{-j} - \frac{(J-1)^{d+1}}{d+1} \exp\left(-\frac{(J-1)(d+1)}{d+2}\right).$$

- Comparison of the asymptotic bound and the pre-asymptotic bound in estimating $\sum_{j\geq J} j^d e^{-j}$ for $d = 20$.




- **Sub-exponential convergence rate is not effective for $J \leq d$!**

# Pre-asymptotic analysis

## Theorem [T, Webster, Zhang '14]

Consider the multi-indexed series $\sum_{\boldsymbol{\nu}\in\mathbb{N}_0^s} e^{-b(\boldsymbol{\nu})}$ with $b(\boldsymbol{\nu})$ being continuous and satisfying

- $b(\boldsymbol{\nu}) \le b(\boldsymbol{\mu})$ for all $\boldsymbol{\nu}, \boldsymbol{\mu} \in [0,\infty)^d$ such that $\boldsymbol{\nu} \le \boldsymbol{\mu}$,
- $b(\tau\boldsymbol{\nu}) = \tau b(\boldsymbol{\nu})$ for all $\tau \in (0,\infty)$, $\boldsymbol{\nu} \in [0,\infty)^d$.

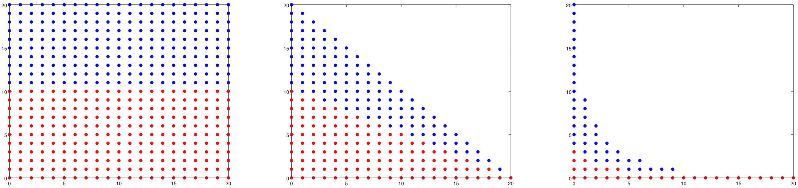
Let $\Lambda_s^{\rm Qopt}$ be the set of indices corresponding to $s$ largest $e^{-b(\boldsymbol{\nu})}$. If $s \le \#(\mathcal{P}_d \cap \mathbb{Z}^d)$, there holds

$$\sum_{\boldsymbol{\nu}\notin\Lambda_s^{\rm Qopt}} e^{-b(\boldsymbol{\nu})} \le e\sigma\left[\sum_{j=1}^{\infty} j^d e^{-j} - \frac{1}{d+1}\left(\frac{s}{\sigma}\right)^{\frac{d+1}{d}} \exp\left(-\frac{s^{\frac{1}{d}}(d+1)}{\sigma^{\frac{1}{d}}(d+2)}\right)\right].$$

Here, $\sigma = \#(\mathcal{P}_1 \cap \mathbb{Z}^d)$.

# How to choose polynomial space efficiently?

Anisotropic representation [Nobile, Tempone, Webster '08]




Anisotropic: introduce weight vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d_+$, with $\alpha_{\min} = 1$:

- Tensor product (TP): $\max_{1 \le i \le d} \alpha_i \nu_i \le w$.
- Total degree (TD): $\sum_{i=1}^d \alpha_i \nu_i \le w$.
- Hyperbolic cross (HC): $\prod_{i=1}^d (\nu_i + 1)^{\alpha_i} \le w + 1$.
- Smolyak (SM): $\sum_{i=1}^d \alpha_i f(\nu_i) \le f(w)\}$, with $f(\nu) = \lceil \log_2(\nu) \rceil$, $\nu \ge 2$.

# How to choose a good polynomial space efficiently?

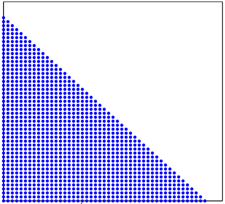
Quasi-optimal strategy [Beck, Nobile, Tamellini, Tempone '12 '14; T., Webster, Zhang '14]

Choose the polynomial space corresponding to the quasi-optimal index set $\Lambda_s^{\mathrm{Qopt}}$ with respect to the sharp estimates $B(\boldsymbol{\nu})$ of $\|c_{\boldsymbol{\nu}}\|_{\mathcal{V}}$, rather than $\|c_{\boldsymbol{\nu}}\|_{\mathcal{V}}$ themselves.

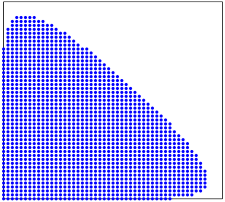
**Advantage:**

- Practical substitution for best $s$-term approximations.
- Index sets are determined near optimally and adaptively w.r.t. diffusion coefficients.

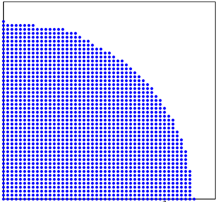
**Challenge:** Sharp bounds are still very difficult and somewhat problem dependent.



$\Lambda_M^{opt}: B(\nu) = \rho^{-\nu}$



$\Lambda_M^{opt}: B(\nu) = \rho^{-\nu}\prod_{i=1}^{N}\sqrt{2\nu_i+1}$



$\Lambda_M^{opt}: B(\nu) = \inf_{(\rho,\delta)\in\mathbf{A}} \frac{1}{\delta}\rho^{-\nu}$

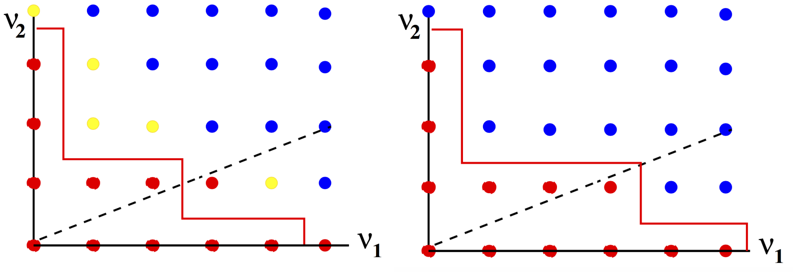
# How to choose a good polynomial space efficiently?

## Greedy algorithm [Chkifa, Cohen, DeVore, Schwab '13 '14]

### Recursive formula for the Taylor coefficients:

with $\{\boldsymbol{e}_i\}_{1\le i\le d}$ the canonical basis of $\mathbb{R}^d$, the coefficient $t_{\boldsymbol{\nu}}$ is solution to

$$\int_D a_0 \nabla t_{\boldsymbol{\nu}} \nabla v = - \sum_{\{i:\, \nu_i \neq 0\}} \int_D \psi_i \nabla t_{\boldsymbol{\nu}-\boldsymbol{e}_i} \nabla v, \quad \forall v \in \mathcal{V}$$




Source: [Chkifa, Cohen, DeVore, Schwab '13]

# References


- J. Beck, F. Nobile, L. Tamellini, R. Tempone, *Convergence of quasi-optimal stochastic Galerkin methods for a class of PDEs with random coefficients*, Computers and Mathematics with Applications, Vol. 67 (4), 732–751, 2014.
- A. Chkifa, A. Cohen, R. DeVore, C. Schwab, *Sparse adaptive Taylor approximation algorithms for parametric and stochastic elliptic PDEs*, Modél. Math. Anal. Numér. 47(1), 253–280 (2013).
- A. Cohen, R. DeVore, C. Schwab, *Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDEs*, Analysis and Applications, Vol. 9, No. 1 (2011), 11–47.
- P. Gruber, *Convex and Discrete Geometry*, Springer, 2007.
- M. Gunzburger, C. G. Webster, G. Zhang, *Stochastic finite element methods for PDEs with random input data*, Acta Numerica, 23:521-650, 2014.
- H. Tran, C. G. Webster, G. Zhang, *Analysis of quasi-optimal polynomial approximations for parameterized PDEs with deterministic and stochastic coefficients*, Numerische Mathematik, to appear, 2016.