ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MATHICSE

Mathematics Institute of Computational Science and Engineering

School of Basic Sciences - Section of Mathematics

MATHICSE Technical Report
Nr. 28.2015
November 2015

Discrete least-squares approximations over optimized downward closed polynomial spaces in arbitrary dimension

Albert Cohen, Giovanni Migliorati, Fabio Nobile

http://mathicse.epfl.ch

Phone: +41 21 69 37648

Address:
EPFL - SB - MATHICSE (Bâtiment MA)
Station 8 - CH-1015 - Lausanne - Switzerland

Fax: +41 21 69 32545

Discrete least-squares approximations over optimized downward closed polynomial spaces in arbitrary dimension *

Albert Cohen[†] Giovanni Migliorati[‡] Fabio Nobile[§]

October 23, 2015

Abstract

We analyze the accuracy of the discrete least-squares approximation of a function u in multivariate polynomial spaces $\mathbb{P}_\Lambda := \mathrm{span}\{y \mapsto y^\nu \,:\, \nu \in \Lambda\}$ with $\Lambda \subset \mathbb{N}_0^d$ over the domain $\Gamma := [-1,1]^d$, based on the sampling of this function at points $y^1, \dots, y^m \in \Gamma$. The samples are independently drawn according to a given probability density ρ belonging to the class of multivariate beta densities, which includes the uniform and Chebyshev densities as particular cases. Motivated by recent results on high-dimensional parametric and stochastic PDEs, we restrict our attention to polynomial spaces associated with *downward closed* sets Λ of *prescribed* cardinality n and we optimize the choice of the space for the given sample. This implies in particular that the selected polynomial space depends on the sample. We are interested in comparing the error of this least-squares approximation measured in $L^2(\Gamma,\rho)$ with the best achievable polynomial approximation error when using downward closed sets of cardinality n. We establish conditions between the dimension n and the size m of the sample, under which these two errors are proven to be comparable. Our main finding is that the dimension d enters only moderately in the resulting trade-off between m and n, in terms of a logarithmic factor $\ln(d)$, and is even absent when the optimization is restricted to a relevant subclass of downward closed sets, named *anchored* sets. In principle, this allows one to use these methods in arbitrarily high or even infinite dimension. Our analysis builds upon [3] which considered fixed and non-optimized downward closed multi-index sets. Potential applications of the proposed results are found in the development and analysis of efficient numerical methods for computing the solution of high-dimensional parametric or stochastic PDEs, but is not limited to this area.

Keywords: high dimensional approximation, convergence rate, discrete least squares, best n-term approximation, downward closed set, anchored set, multivariate polynomial approximation.

MSC: 41A10, 41A25, 41A50, 41A63, 65M70.

1 Introduction

In recent years it has become clear that many interesting engineering applications feature an intrinsic dependence on a large number of parameters $y_1, \dots, y_d$, leading to a major concentration of efforts in

*This research was supported by the Institut Universitaire de France, the ERC AdG BREAD, and the center for advanced modeling science (CADMOS).

[†]Sorbonne Universités, UPMC Univ Paris 06, CNRS, UMR 7598, Laboratoire Jacques-Louis Lions, 4, place Jussieu 75005, Paris, France. email: cohen@ann.jussieu.fr

[‡]Sorbonne Universités, UPMC Univ Paris 06, CNRS, UMR 7598, Laboratoire Jacques-Louis Lions, 4, place Jussieu 75005, Paris, France. email: migliorati@ann.jussieu.fr

[§]MATHICSE-CSQI, École Polytechnique Fédérale de Lausanne, Lausanne CH-1015, Switzerland. email: fabio.nobile@epfl.ch

the development and analysis of high-dimensional approximation methods. In many relevant situations, the parameters y_j are independent random variables distributed on intervals I_j according to probability measures ρ_j. We are then typically interested in approximating a function

$$y = (y_1, \dots, y_d) \mapsto u(y), \tag{1}$$

depending on these parameters and measuring the error in $L^2(\Gamma, \rho)$, where $\Gamma = I_1 \times \cdots \times I_d$ and $\rho = \rho_1 \otimes \cdots \otimes \rho_d$. Up to a renormalization, we may assume that $I_j = [-1, 1]$ for all j, so that $\Gamma = [-1, 1]^d$. In certain cases, the number of parameters may even be countably infinite, in which case $\Gamma = [-1, 1]^{\mathbb{N}}$. Examples where such problems occur are recurrent in the numerical treatment of parametric and stochastic PDEs, where a fast and accurate approximation of the parameter-to-solution map over high-dimensional parameter sets is useful to tackle more complex optimization, control and inverse problems.

In this context, the potential of specific high-dimensional *polynomial* approximation methods has been demonstrated in [6, 5, 17, 11, 3]. In these methods, the approximation is picked from a multivariate polynomial space

$$\mathbb{P}_\Lambda := \operatorname{span}\{y \mapsto y^\nu \ : \ \nu \in \Lambda\}, \tag{2}$$

where Λ is a given finite subset of $\mathbb{N}_0^d$. In the case of countably many variables, $d = \infty$, we replace $\mathbb{N}_0^d$ by the set of finitely supported sequences of nonnegative integers.

Such a set Λ is said to be *downward closed* if and only if

$$\nu \in \Lambda \quad \text{and} \quad \mu \leq \nu \implies \mu \in \Lambda, \tag{3}$$

where $\mu \leq \nu$ is meant component-wise. Polynomial spaces $\mathbb{P}_\Lambda$ associated to downward closed index sets Λ have been studied in various contexts, see [1, 9, 8, 12, 13].

There exist two main approaches to polynomial approximation of a given function u based on pointwise evaluations. The first one relies on *interpolation* of the function u at a given set of points $\{y^1, \dots, y^n\}$ where $n := \#(\Lambda) = \dim(\mathbb{P}_\Lambda)$, that is, find $v \in \mathbb{P}_\Lambda$ such that $v(y^i) = u(y^i)$ for $i = 1, \dots, n$. The second one relies on *projection*, which aims at minimizing the $L^2(\Gamma, \rho)$ error between u and its approximation in $\mathbb{P}_\Lambda$. Since the exact projection is not available, one typical approach consists in using the discrete least-squares method, that is, solving the problem

$$\min_{v \in \mathbb{P}_\Lambda} \sum_{i=1}^{m} |v(y^i) - u(y^i)|^2, \tag{4}$$

where now $m > n$. Discrete least-squares methods are often preferred to interpolation methods when the observed point values are polluted by noise. Their convergence analysis has been studied in the the general context of learning theory, see for example [7, 19, 10, 20, 21].

In recent years, an analysis of discrete least-squares methods has been proposed [3, 16, 11, 18], specifically targeted to the above described case of multivariate polynomial spaces associated with downward closed sets, in the case where the ρ_j are identical Jacobi-type measures. This analysis, which builds upon the general results from [4], gives conditions ensuring that, in the absence of noise in the pointwise evaluation of u, the accuracy of the discrete least-squares approximation is comparable to the best approximation error achievable in $\mathbb{P}_\Lambda$, that is,

$$e_\Lambda(u) := \inf_{v \in \mathbb{P}_\Lambda} \|u - v\|_{L^2(\Gamma, \rho)}. \tag{5}$$

These conditions are stated in terms of a relation between the size m of the sample and the dimension n of $\mathbb{P}_\Lambda$. A similar analysis also covers the case of an additive noise in the evaluation of the samples which results in additional terms in the error estimate, see *e.g.* [18].

One remarkable result from the above analysis is that the conditions ensuring that the least-square method has accuracy comparable to $e_\Lambda(u)$ only involve the dimension of $\mathbb{P}_\Lambda$. These conditions are independent of the specific shape of the set Λ (as long as it is downward closed), and in particular independent of the dimension d.

The possibility of using arbitrary sets Λ is critical in the context of parametric PDEs in view of the recent results on high-dimensional polynomial approximation obtained in [6, 2, 5]. These results show that for relevant class of parametric PDEs, the functions $y \mapsto u(y)$ describing either the full solution or scalar quantities of interest can be approximated with convergence rates $\mathcal{O}(n^{-s})$ which are *independent* of the parametric dimension d, when using polynomial spaces $\mathbb{P}_{\Lambda_n}$ associated to *specific* sequences of downward closed multi-index sets $(\Lambda_n)_{n\geq 1}$ with $\#(\Lambda_n) = n$. In summary, we have

$$e_n(u) := \min_{\#(\Lambda)=n} e_\Lambda(u) \leq Cn^{-s}, \tag{6}$$

where the minimum is taken over all downward closed sets of given cardinality n.

For each value of n, the optimal set Λ_n is the one that achieves the minimum in (6) among all downward closed Λ of cardinality n. This set is unknown to us when observing only the samples $u(y^i)$. Therefore, a legitimate objective is to develop least-squares methods for which the accuracy is comparable to the quantity $e_n(u)$.

In this paper, we discuss least-square approximations on multivariate polynomial spaces for which the choice of Λ is optimized based on the sample. In particular we prove that the performance of such approximation is comparable to the quantity in (6), under a relation between m and n where the dimension d enters as a logarithmic factor. We show that this logarithmic dependence on d can be fully removed by considering a more restricted class of downward closed sets called *anchored sets*, for which similar approximation rates as in (6) can be achieved. The resulting least-square methods are thus immune to the curse of dimensionality.

The outline of the paper is the following: in Section 2 we introduce the notation and briefly review some of the previous results achieved in the analysis of discrete least squares on *fixed* multivariate polynomial spaces. In Section 3 we present the main results of the paper concerning discrete least-square approximations on *optimized* polynomial spaces. Our analysis is based on establishing upper bounds on the number of downward closed or anchored sets of a given cardinality, or on the cardinality of their union.

The selection of the optimal polynomial space is based on minimizing the least-squares error among all possible choices of downward closed or anchored sets of a given cardinality n. Let us stress that in the form of an exhaustive search, this task becomes computationally intensive when n and d are simultaneously large. Our results should therefore mainly be viewed as a benchmark in arbitrary dimension d for assessing the performance of fast selection algorithms, such as greedy algorithms, that still need to be developed and analyzed in this context.

2 Least-squares approximation by multivariate polynomials

In this section we introduce some useful notation, and recall from [3] the main results achieved for the analysis of the stability and accuracy of discrete least-squares approximations in multivariate polynomial spaces.

2.1 Notation

In any given dimension $d \in \mathbb{N}$, we consider the domain $\Gamma := [-1,1]^d$, and for some given real numbers $\theta_1, \theta_2 > -1$, the tensorized Jacobi measure

$$\rho(y) := \otimes_{j=1}^d \beta(y_j) \tag{7}$$

where

$$d\beta = c(1-t)^{\theta_1}(1+t)^{\theta_2}dt, \quad c := \left(\int_{-1}^1 (1-t)^{\theta_1}(1+t)^{\theta_2}dt\right)^{-1}. \tag{8}$$

We may also consider the case $\Gamma := [-1,1]^{\mathbb{N}}$ for which $d = +\infty$ and ρ is the Jacobi measure defined over Γ in the usual manner. We denote by $L^2(\Gamma, \rho)$ the Hilbert space of real-valued square-integrable functions with respect to ρ and denote by $\|\cdot\|$ the associated norm, i.e.

$$\|v\| := \left(\int_\Gamma |v(y)|^2 d\rho(y)\right)^{1/2}. \tag{9}$$

Moreover, let $\mathcal{F}$ be defined as the set $\mathbb{N}_0^d$, where $\mathbb{N}_0 := \{0,1,2,\dots\}$, in the case $d < +\infty$, or as the countable set of all finitely supported sequences from $\mathbb{N}_0^{\mathbb{N}}$ in the case $d = +\infty$. Sometimes we refer to $\mathcal{F}$ as the universe multi-index set. For any $\nu \in \mathcal{F}$ we define

$$\text{supp}(\nu) := \{j \geq 1 : \nu_j \neq 0\}, \tag{10}$$

and for any multi-index set $\Lambda \subseteq \mathcal{F}$ we define

$$\text{supp}(\Lambda) := \cup_{\nu\in\Lambda} \text{supp}\{\nu\}. \tag{11}$$

We say that a variable y_j is *active* in the space $\mathbb{P}_\Lambda$ when $j \in \text{supp}(\Lambda)$.

For the given real parameters $\theta_1, \theta_2 > -1$, we introduce the family $(J_n)_{n\geq 0}$ of univariate orthonormal Jacobi polynomials associated with the measure β, and their tensorized counterpart

$$J_\nu(y) = \prod_{j\geq 1} J_{\nu_j}(y_j), \quad y = (y_j)_{j\geq 1}, \tag{12}$$

for any $\nu \in \mathcal{F}$. The $(J_\nu)_{\nu\in\mathcal{F}}$ are an $L^2(\Gamma,\rho)$-orthonormal basis. Particular instances of these polynomials are tensorized Legendre polynomials when $\theta_1 = \theta_2 = 0$ and tensorized Chebyshev polynomials of the first kind when $\theta_1 = \theta_2 = -1/2$.

In the present paper we focus on finite multi-index sets Λ which are downward closed in the sense of (3). We also say that a polynomial space $\mathbb{P}_\Lambda$ is downward closed when it is associated with a downward closed multi-index set $\Lambda \subset \mathcal{F}$. Recall that $\mathbb{P}_\Lambda$ has been defined as the span of the monomials $y \mapsto y^\nu$ for $\nu \in \Lambda$. Therefore it admits $(J_\nu)_{\nu\in\Lambda}$ as an $L^2(\Gamma,\rho)$-orthonormal basis in the case of Λ downward closed. Sometimes we enumerate the indices ν using the lexicographical ordering, and denote this basis by $(\psi_k)_{k=1,\dots,n}$, where

$$n := \#(\Lambda) = \dim(\mathbb{P}_\Lambda). \tag{13}$$

Given a finite downward closed multi-index set $\Lambda \subset \mathcal{F}$, we would like to approximate the target function $u : \Gamma \to \mathbb{R}$ in the L^2 sense, using the noiseless evaluations $(u(y^i))_{i=1,\dots,m}$ of u at the points $(y^i)_{i=1,\dots,m}$, where the y^i are i.i.d. random variables distributed according to ρ. We introduce the continuous L^2 projection of u on the polynomial space $\mathbb{P}_\Lambda$ as

$$\Pi_\Lambda u := \underset{v\in\mathbb{P}_\Lambda}{\text{argmin}} \|u - v\|, \tag{14}$$

and the discrete least-squares approximation $\Pi_\Lambda^m u$ as

$$\Pi_\Lambda^m u := \underset{v\in\mathbb{P}_\Lambda}{\operatorname{argmin}} \|u - v\|_m, \tag{15}$$

where we have used the notation

$$\|v\|_m := \left(\frac{1}{m}\sum_{i=1}^m v(y^i)^2\right)^{\frac{1}{2}}. \tag{16}$$

We introduce the $m \times \#(\Lambda)$ design matrix $\mathbf{D}$ and the vector $\boldsymbol{b} \in \mathbb{R}^m$ whose entries are given by $\mathbf{D}_{i,k} = \psi_k(y^i)$ and $\boldsymbol{b}_i = u(y^i)$. We define the Gramian matrix $\mathbf{G} = m^{-1}\mathbf{D}^{\mathrm{T}}\mathbf{D}$. The discrete least-squares projection in (15) is then given by

$$\Pi_\Lambda^m u = \sum_{k=1}^{\#(\Lambda)} \mathbf{a}_k \psi_k, \tag{17}$$

where the vector $\mathbf{a} = (\mathbf{a}_k) \in \mathbb{R}^{\#(\Lambda)}$ is the solution to the normal equations

$$\mathbf{G}\mathbf{a} = m^{-1}\mathbf{D}^{\mathrm{T}}\boldsymbol{b}. \tag{18}$$

2.2 Previous results on the stability and accuracy of discrete least squares

We introduce the quantity

$$K(\mathbb{P}_\Lambda) := \sup_{y\in\Gamma} \sum_{\nu\in\Lambda} |J_\nu(y)|^2. \tag{19}$$

It is proven in [3] that discrete least squares in multivariate polynomial spaces are stable and accurate provided a precise proportionality between m and $K(\mathbb{P}_\Lambda)$ is satisfied. Similar results have been proven in [18] for the case of noisy observations of the target function, with several noise models.

For any $\delta \in]0,1[$ we introduce the positive quantity

$$\zeta(\delta) := \delta + (1-\delta)\ln(1-\delta). \tag{20}$$

Given a threshold $\tau \in \mathbb{R}_0^+$, we introduce the truncation operator

$$T_\tau(t) := \operatorname{sign}(t)\min\{\tau, |t|\}, \qquad \text{for any } t \in \mathbb{R},$$

and use it to define the truncated discrete least-squares projection operator $u \mapsto T_\tau(\Pi_\Lambda^m u)$. The main results from [3] concerning stability and accuracy of the discrete least-squares approximation with noiseless evaluations can be summarized as follows.

Theorem 1. *In any dimension d, for any $r > 0$, any $\delta \in]0,1[$ and any downward closed multi-index set $\Lambda \subset \mathbb{N}_0^d$, one has*

$$\Pr\left(\left\{(1-\delta)\|v\|^2 \le \|v\|_m^2 \le (1+\delta)\|v\|^2,\ \forall v \in \mathbb{P}_\Lambda\right\}\right) \ge 1 - 2n\exp(-\zeta(\delta)m/K(\mathbb{P}_\Lambda)). \tag{21}$$

If the following condition between m and $K(\mathbb{P}_\Lambda)$ is satisfied

$$\frac{m}{\ln m} \ge \frac{(1+r)}{\zeta(\delta)} K(\mathbb{P}_\Lambda), \tag{22}$$

then

$$\Pr\left(\left\{(1-\delta)\|v\|^2 \le \|v\|_m^2 \le (1+\delta)\|v\|^2,\ \forall v \in \mathbb{P}_\Lambda\right\}\right) \ge 1 - 2m^{-r}. \tag{23}$$

Moreover, for any $u \in L^\infty(\Gamma)$ with $\|u\|_{L^\infty(\Gamma)} \le \tau$, the following holds:

$$\Pr\left(\|u - \Pi_\Lambda^m u\| \le \left(1 + \sqrt{\frac{1}{1-\delta}}\right) \inf_{v \in \mathbb{P}_\Lambda} \|u - v\|_{L^\infty(\Gamma)}\right) \ge 1 - 2m^{-r}, \tag{24}$$

$$\mathbb{E}\left(\|u - T_\tau(\Pi_\Lambda^m u)\|^2\right) \le \left(1 + \frac{4\zeta(\delta)}{(1+r)\ln m}\right) \|u - \Pi_\Lambda u\|^2 + 8\tau^2 m^{-r}. \tag{25}$$

The above theorem states that under condition (22) the discrete least-squares approximation is stable, since one has

$$(1-\delta)\|v\|^2 \le \|v\|_m^2 \le (1+\delta)\|v\|^2, \ \forall v \in \mathbb{P}_\Lambda \Leftrightarrow (1-\delta)\mathbf{I} \le \mathbf{G} \le (1+\delta)\mathbf{I}. \tag{26}$$

Under the same condition, the discrete least-squares approximation is also accurate in probability, from (24), and in expectation, from (25), since the approximation error behaves like the best approximation error in L^∞ or in L^2.

The quantity $K(\mathbb{P}_\Lambda)$ depends both on Λ and on the chosen Jacobi measure, and therefore on the parameters θ_1, θ_2. The following result from [3] and [14] shows that, once these two parameters are fixed, the quantity $K(\mathbb{P}_\Lambda)$ only depends on $\#(\Lambda)$, independently of the particular shape of Λ and of the dimension d.

Lemma 1. *In any dimension d and for any finite downward closed set $\Lambda \subset \mathcal{F}$, one has*

$$\#(\Lambda) \le K(\mathbb{P}_\Lambda) \le \begin{cases} (\#(\Lambda))^{\ln 3/\ln 2}, & \text{if } \theta_1 = \theta_2 = -1/2, \\ (\#(\Lambda))^{2\max\{\theta_1,\theta_2\}+2} & \text{if } \theta_1, \theta_2 \in \mathbb{N}_0. \end{cases} \tag{27}$$

Combining the two results, one therefore obtains sufficient conditions for stability and optimal accuracy expressed only in terms of a relation between $\#(\Lambda)$ and m. For example, in the case of the uniform measure that corresponds to $\theta_1 = \theta_2 = 0$, this relation is of the form

$$\frac{m}{\ln m} \ge c\,(\#(\Lambda))^2, \quad c := c(\delta, r). \tag{28}$$

3 Optimal selection of downward closed polynomial spaces

The results recalled in the previous section hold for a given downward closed set $\Lambda \subset \mathcal{F}$. We now consider the problem of optimizing the choice of Λ, or equivalently that of the space $\mathbb{P}_\Lambda$.

3.1 Optimized index sets

We define the family

$$\mathcal{M}_n^d := \{\Lambda \subset \mathcal{F} \,:\, \Lambda \text{ is downward closed and } \#(\Lambda) = n\}, \tag{29}$$

of all downward closed sets of cardinality n in d variables. Note that, in contrast to the family of *all* sets of cardinality n, the family $\mathcal{M}_n^d$ is finite.

The error of best n-term polynomial approximation by downward closed sets is then defined by

$$\sigma_n(u) := \min_{\Lambda \in \mathcal{M}_n^d} \min_{v \in \mathbb{P}_\Lambda} \|u - v\|. \tag{30}$$

A best n-term downward closed set is a $\Lambda \in \mathcal{M}_n^d$ that achieves this minimum, that is, such that

$$\Lambda^{opt} := \operatorname*{argmin}_{\Lambda \in \mathcal{M}_n^d} \min_{v \in \mathbb{P}_\Lambda} \|u - v\|. \tag{31}$$

In view of the Parseval identity, we find that Λ^{opt} is also given by

$$\Lambda^{opt} := \operatorname*{argmin}_{\Lambda \in \mathcal{M}_n^d} \sum_{\nu \notin \Lambda} |u_\nu|^2, \quad u_\nu = \int_\Gamma u(y) J_\nu(y) d\rho(y). \tag{32}$$

Note that such a set may not be unique due to possible ties in the values of the coefficients, in which case we consider a unique choice by breaking the ties in some arbitrary but fixed way. We set

$$u_n = \Pi_{\Lambda^{opt}} u = \operatorname*{argmin}_{v \in \mathbb{P}_{\Lambda^{opt}}} \|u - v\|. \tag{33}$$

Of course, in the least-squares method, the discrete data do not allow us to identify Λ^{opt}. Instead, we rely on

$$\Lambda_m^{opt} := \operatorname*{argmin}_{\Lambda \in \mathcal{M}_n^d} \min_{v \in \mathbb{P}_\Lambda} \|u - v\|_m. \tag{34}$$

and compute

$$w_n = \Pi^m_{\Lambda_m^{opt}} u = \operatorname*{argmin}_{v \in \mathbb{P}_{\Lambda_m^{opt}}} \|u - v\|_m. \tag{35}$$

Our objective is now to compare the accuracy of the polynomial least-squares approximation based on Λ_m^{opt} with the above optimal error $\sigma_n(u)$. For this purpose, we shall use the random variable

$$C_n^d := \max_{\Lambda \in \mathcal{M}_n^d} \max_{v \in \mathbb{P}_\Lambda} \frac{\|v\|^2}{\|v\|_m^2}. \tag{36}$$

Note that the search of Λ_m^{opt} remains a difficult task from the computational point of view, due to the fact that $\#(\mathcal{M}_n^d)$ becomes very large even for moderate values of n and d. As we discuss further, this cardinality also affects the conditions between m and n which guarantee the optimality of the least-squares approximation based on Λ_m^{opt}.

For this reason, it is useful to introduce an additional restriction on the potential index sets. We say that Λ is *anchored* if and only if it is downward closed and satisfies in addition

$$e_j \in \Lambda \text{ and } j' \le j \implies e_{j'} \in \Lambda, \tag{37}$$

where e_j and $e_{j'}$ are the Kronecker sequences with 1 at position j and j', respectively. We also say that a polynomial space $\mathbb{P}_\Lambda$ is anchored when Λ is anchored. Likewise, we define the family

$$\mathcal{A}_n := \{\Lambda \subset \mathcal{F} \ : \ \Lambda \text{ is anchored and } \#(\Lambda) = n\}. \tag{38}$$

The property of anchored set introduces an order of priority between the variables: given any $j \ge 1$, the variable y_j is active in Λ only if all the variables y_k for $k < j$ are also active. In particular, for any set $\Lambda \in \mathcal{A}_n$ we have

$$\operatorname{supp}(\Lambda) = \{1, \dots, k\}, \tag{39}$$

for some $k \le n - 1$.

It is proven in [5] that, for relevant classes of parametric PDEs, the same algebraic convergence rates $\mathcal{O}(n^{-s})$ can be obtained when imposing the anchored structure on the optimally selected sets $(\Lambda_n)_{n \ge 0}$ with $\#(\Lambda_n) = n$. As we shall see further, one specific advantage of anchored sets is to completely remove the dependence in the dimension d in the convergence analysis of the least-squares method.

Using the same notation as before with obvious modifications, we introduce the following entities:

$$\tilde{\Lambda}^{opt} := \operatorname*{argmin}_{\Lambda \in \mathcal{A}_n} \min_{v \in \mathbb{P}_\Lambda} \|u - v\|, \qquad \tilde{u}_n := \Pi_{\tilde{\Lambda}^{opt}} u = \operatorname*{argmin}_{v \in \mathbb{P}_{\tilde{\Lambda}^{opt}}} \|u - v\|, \tag{40}$$

$$\tilde{\Lambda}_m^{opt} := \operatorname*{argmin}_{\Lambda \in \mathcal{A}_n} \min_{v \in \mathbb{P}_\Lambda} \|u - v\|_m, \qquad \tilde{w}_n := \Pi^m_{\tilde{\Lambda}_m^{opt}} u = \operatorname*{argmin}_{v \in \mathbb{P}_{\tilde{\Lambda}_m^{opt}}} \|u - v\|_m. \tag{41}$$

and

$$\widetilde{C}_n := \max_{\Lambda \in \mathcal{A}_n} \max_{v \in \mathbb{P}_\Lambda} \frac{\|v\|^2}{\|v\|_m^2}. \tag{42}$$

We now would like to study the best n-term approximation in the aforementioned classes of multi-index sets $\mathcal{M}_n^d$ and $\mathcal{A}_n$. The following lemma shows the role played by C_n^d and $\widetilde{C}_n$ in quantifying the relation between the error achieved by the optimal discrete least-square projection and the error achieved by the optimal L^2 projection.

Lemma 2. *It holds that, for any* $\Lambda \in \mathcal{M}_n^d$,

$$\|u - w_n\| \leq \|u - v\| + 2\sqrt{C_{2n-1}^d}\|u - v\|_m, \quad v \in \mathbb{P}_\Lambda, \tag{43}$$

and for any $\Lambda \in \mathcal{A}_n$,

$$\|u - \tilde{w}_n\| \leq \|u - \tilde{v}\| + 2\sqrt{\widetilde{C}_{2n-1}}\|u - \tilde{v}\|_m, \quad \tilde{v} \in \mathbb{P}_\Lambda. \tag{44}$$

Proof. Let $\Lambda \in \mathcal{M}_n^d$ and define $\widehat{\Lambda} := \Lambda \cup \Lambda_m^{opt}$. We first observe that $\widehat{\Lambda}$ is also downward closed and $\#(\widehat{\Lambda}) \leq 2n - 1$ because any downward closed set contains the null multi-index. Since $w_n \in \mathbb{P}_{\Lambda_m^{opt}}$, we have $v - w_n \in \mathbb{P}_{\widehat{\Lambda}}$ for any $v \in \mathbb{P}_\Lambda$. It follows that

$$\|v - w_n\| \leq \sqrt{C_{2n-1}^d}\|v - w_n\|_m \leq \sqrt{C_{2n-1}^d}(\|u - v\|_m + \|u - w_n\|_m) \leq 2\sqrt{C_{2n-1}^d}\|u - v\|_m,$$

and therefore

$$\|u - w_n\| \leq \|u - v\| + \|v - w_n\| \leq \|u - v\| + 2\sqrt{C_{2n-1}^d}\|u - v\|_m, \tag{45}$$

which is (43). The proof of (44) is analogous. □

Note that the estimates in the above lemma imply in particular that

$$\|u - w_n\| \leq \|u - u_n\| + 2\sqrt{C_{2n-1}^d}\|u - u_n\|_m, \tag{46}$$

and

$$\|u - \tilde{w}_n\| \leq \|u - \tilde{u}_n\| + 2\sqrt{\widetilde{C}_{2n-1}}\|u - \tilde{u}_n\|_m, \tag{47}$$

with u_n and $\tilde{u}_n$ defined by (33) and (40). Note that they also imply

$$\|u - w_n\| \leq \left(1 + 2\sqrt{C_{2n-1}^d}\right)\|u - v\|_{L^\infty}, \quad v \in \mathbb{P}_\Lambda, \tag{48}$$

for any $\Lambda \in \mathcal{M}_n^d$, and

$$\|u - \tilde{w}_n\| \leq \left(1 + 2\sqrt{\widetilde{C}_{2n-1}}\right)\|u - \tilde{v}\|_{L^\infty}, \quad \tilde{v} \in \mathbb{P}_\Lambda, \tag{49}$$

for any $\Lambda \in \mathcal{A}_n$.

3.2 Probabilistic bounds

In view of Lemma 2, we are interested in bounding the random variables C_n^d and $\widetilde{C}_n$. In this section we give probabilistic bounds, which ensure that under certain conditions between m and n, these variable do not exceed a fixed value, here set to 2, with high probability. In the whole section we choose $\delta = 1/2$, so that, with the notation (20), one has

$$\zeta := \zeta(\delta) = \zeta(1/2) = (1 - \ln 2)/2 \approx 0.153. \tag{50}$$

We define, for any $\nu \in \mathcal{F}$, the "rectangular" set $\mathcal{R}_\nu := \{\mu \in \mathcal{F},\ \mu \leq \nu\}$; and for any $n \geq 1$, the hyperbolic cross set

$$\mathcal{H}_n^d := \left\{ \mu \in \mathcal{F} : \ \prod_{j=1}^{d} (\mu_j + 1) \leq n \right\}. \tag{51}$$

Note that

$$\mathcal{H}_n^d = \bigcup_{\#(\mathcal{R}_\nu) \leq n} \mathcal{R}_\nu. \tag{52}$$

The cardinality of $\mathcal{H}_n^d$ is bounded by

$$\#(\mathcal{H}_n^d) \leq n(1 + \ln(n))^{d-1}, \tag{53}$$

see [11, Appendix A.2] for a proof and some remarks on the accuracy of this upper bound.

Let us observe that the union of all downward closed sets of cardinality at most n coincides with $\mathcal{H}_n^d$, that is,

$$\bigcup_{\Lambda \in \mathcal{M}_n^d} \Lambda = \mathcal{H}_n^d. \tag{54}$$

Indeed, on the one hand, all rectangles $\mathcal{R}_\nu$ such that $\#(\mathcal{R}_\nu) \leq n$ belong to $\mathcal{M}_n^d$, so that inclusion holds from right to left. On the other hand, inclusion from left to right follows by observing that for any $\Lambda \in \mathcal{M}_n^d$, one has $\Lambda = \cup_{\mu \in \Lambda} \mathcal{R}_\mu$ and $\mathcal{R}_\mu \subset \mathcal{H}_n^d$ for all $\mu \in \Lambda$.

This leads us to a first probabilistic bound for the random variable C_n^d. Indeed, using (54) we obtain that

$$\Pr\left(C_n^d > 2\right) = \Pr\left(\max_{\Lambda \in \mathcal{M}_n^d} \max_{v \in \mathbb{P}_\Lambda} \frac{\|v\|^2}{\|v\|_m^2} > 2\right) \leq \Pr\left(\max_{v \in \mathbb{P}_{\mathcal{H}_n^d}} \frac{\|v\|^2}{\|v\|_m^2} > 2\right).$$

Thus, using Theorem 1 with $\delta = 1/2$ combined with the estimates in Lemma 1, we obtain that, in any dimension d and for any $r > 0$, if m and n satisfy

$$\frac{m}{\ln m} \geq \begin{cases} \frac{(1+r)}{\zeta} (\#(\mathcal{H}_n^d))^{\ln 3/\ln 2}, & \text{with Chebyshev polynomials,} \\ \frac{(1+r)}{\zeta} (\#(\mathcal{H}_n^d))^{2\max\{\theta_1,\theta_2\}+2}, & \text{with Jacobi polynomials and } \theta_1, \theta_2 \in \mathbb{N}_0, \end{cases} \tag{55}$$

then

$$\Pr(C_n^d > 2) \leq 2m^{-r}. \tag{56}$$

From (54) and (39) we also find that the union of all anchored sets of cardinality at most n satisfies the following inclusion

$$\bigcup_{\Lambda \in \mathcal{A}_n} \Lambda \subset \mathcal{H}_n^{n-1}. \tag{57}$$

By similar arguments, we obtain the following probabilistic bound for the random variable $\widetilde{C}_n$: in any dimension d, for any $r > 0$, if m and n satisfy

$$\frac{m}{\ln m} \geq \begin{cases} \frac{(1+r)}{\zeta}(\#(\mathcal{H}_n^{n-1}))^{\ln 3/\ln 2}, & \text{with Chebyshev polynomials,} \\ \frac{(1+r)}{\zeta}(\#(\mathcal{H}_n^{n-1}))^{2\max\{\theta_1,\theta_2\}+2}, & \text{with Jacobi polynomials and } \theta_1, \theta_2 \in \mathbb{N}_0, \end{cases} \tag{58}$$

then

$$\Pr(\widetilde{C}_n > 2) \leq 2m^{-r}. \tag{59}$$

The above results describe the regimes of m and n such that C_n^d and $\widetilde{C}_n$ do not exceed 2 with high probability. In the case of downward closed sets, this regime is quite restrictive due to the presence of $\ln(n)^{d-1}$ in the cardinality of $\mathcal{H}_n^d$, which enforces the sample size m to be extremely large as d grows. Likewise, m has to be extremely large compared to n in the case of anchored sets.

We next describe another strategy which yields similar probabilistic bounds under less restrictive regimes. It is based on estimating the cardinality of $\mathcal{M}_n^d$ and $\mathcal{A}_n$ and using union bounds. Recalling the definition of $K(\mathbb{P}_\Lambda)$ from (19), we introduce the following notation:

$$K_n = \max_{\Lambda \in \mathcal{M}_n^d} K(\mathbb{P}_\Lambda), \tag{60}$$

$$\widetilde{K}_n = \max_{\Lambda \in \mathcal{A}_n} K(\mathbb{P}_\Lambda). \tag{61}$$

According to Lemma 1, one has the estimate

$$\widetilde{K}_n \leq K_n \leq \begin{cases} n^{\ln 3/\ln 2}, & \text{if } \theta_1 = \theta_2 = -1/2, \\ n^{2\max\{\theta_1,\theta_2\}+2} & \text{if } \theta_1, \theta_2 \in \mathbb{N}_0. \end{cases} \tag{62}$$

Our way of estimating $\#(\mathcal{M}_n^d)$ and $\#(\mathcal{A}_n)$ is based on bitstream models for encoding any lower or anchored set.

One first model to encode any given lower set $\Lambda \subset \mathcal{F}$ in d dimensions, consists in associating d bits to each multi-index $\nu \in \Lambda$, where the value of the jth bit is equal to one if $\nu + e_j \in \Lambda$, and equal to zero if $\nu + e_j \notin \Lambda$. By ordering these blocks of bits according to the lexicographic order of appearance of ν in Λ, one obtains a bitstream which uniquely encodes Λ. Hence we can represent any lower set containing n multi-indices by means of nd bits. The representation provided by this model is encoded with redundancy, *i.e.* there exist many different ways to encode the same set. Using this model we have the upper bound

$$\#(\mathcal{M}_n^d) \leq 2^{nd}. \tag{63}$$

Using a union bound and (21) we obtain

$$\begin{aligned} \Pr\left(C_n^d > 2\right) &= \Pr\left(\max_{\Lambda \in \mathcal{M}_n^d} \max_{v \in \mathbb{P}_\Lambda} \frac{\|v\|^2}{\|v\|_m^2} > 2\right) \\ &\leq \sum_{\Lambda \in \mathcal{M}_n^d} \Pr\left(\max_{v \in \mathbb{P}_\Lambda} \frac{\|v\|^2}{\|v\|_m^2} > 2\right) \\ &\leq 2^{nd} 2n \exp\{-\zeta m/K_n\} \\ &= 2n \exp\{-\zeta m/K_n + nd\ln(2)\}. \end{aligned} \tag{64}$$

Combining with (62), we obtain the following probabilistic bound for the random variable C_n^d: in any dimension d, for any $r > 0$, if m and n satisfy

$$\frac{m}{\ln m} \geq \begin{cases} \left(1 + r + \dfrac{nd \ln 2}{\ln m}\right) \dfrac{n^{\ln 3/\ln 2}}{\zeta}, & \text{with Chebyshev polynomials,} \\ \left(1 + r + \dfrac{nd \ln 2}{\ln m}\right) \dfrac{n^{2\max\{\theta_1,\theta_2\}+2}}{\zeta}, & \text{with Jacobi polynomials and } \theta_1, \theta_2 \in \mathbb{N}_0, \end{cases} \tag{65}$$

then

$$\Pr\left(C_n^d > 2\right) \leq 2nm^{-(r+1)} \leq 2m^{-r}. \tag{66}$$

The same encoding of course works for anchored sets, and thus by similar arguments we obtain

$$\#(\mathcal{A}_n) \leq \#(\mathcal{M}_n^{n-1}) \leq 2^{n^2}, \tag{67}$$

and

$$\begin{aligned} \Pr\left(\widetilde{C}_n > 2\right) &\leq 2^{n^2} 2n \exp\left\{-\zeta m/\widetilde{K}_n\right\} \\ &= 2n \exp\left\{-\zeta m/\widetilde{K}_n + n^2 \ln(2)\right\} \end{aligned} \tag{68}$$

in the case of anchored sets. Likewise, we obtain the following probabilistic bound for the random variable $\widetilde{C}_n$: in any dimension d, for any $r > 0$, if m and n satisfy

$$\frac{m}{\ln m} \geq \begin{cases} \left(1 + r + \dfrac{n^2 \ln 2}{\ln m}\right) \dfrac{n^{\ln 3/\ln 2}}{\zeta}, & \text{with Chebyshev polynomials,} \\ \left(1 + r + \dfrac{n^2 \ln 2}{\ln m}\right) \dfrac{n^{2\max\{\theta_1,\theta_2\}+2}}{\zeta}, & \text{with Jacobi polynomials and } \theta_1, \theta_2 \in \mathbb{N}_0, \end{cases} \tag{69}$$

then

$$\Pr\left(\widetilde{C}_n > 2\right) \leq 2nm^{-(r+1)} \leq 2m^{-r}. \tag{70}$$

The regimes of m and n described by the above results, are in principle less restrictive than those previously obtained using the cardinality of $\mathcal{H}_n^d$ or $\mathcal{H}_n^{n-1}$ since factors of the form $\ln(n)^d$ or $\ln(n)^n$ have been replaced by nd and n^2 respectively.

We next discuss another bitstream model, which allow us to get a further improvement. Given any downward closed multi-index set Λ with $\#(\Lambda) = n$, we order the elements of Λ such that the set

$$\Lambda^k := \{\nu^1, \ldots, \nu^k\}, \tag{71}$$

obtained by retaining only the first k elements of Λ, is downward closed for any $k = 1, \ldots, n$. Of course such an ordering always exists, since for any $k = 1, \ldots, n$ we can take

$$\nu^k = \nu^l + e_j \tag{72}$$

where $l \in \{1, \ldots, k-1\}$ and $j \in \{1, \ldots, d\}$. Hence

$$\#(\mathcal{M}_n^d) \leq d(2d) \cdots (n-1)d = d^{n-1}(n-1)!. \tag{73}$$

and using the inequality $n! \leq e\sqrt{n}(n/e)^n$ which holds for any $n \geq 1$, we obtain by a union bound, for any $n \geq 2$,

$$\begin{aligned}
\Pr\left(C_n^d > 2\right) &\leq d^{n-1}(n-1)!\ (2n)\ \exp\left\{-\zeta m/K_n\right\} \\
&\leq d^{n-1} e\sqrt{n-1}\left(\frac{n-1}{e}\right)^{n-1}\ (2n)\ \exp\left\{-\zeta m/K_n\right\} \\
&= 2n\exp\left\{-\zeta m/K_n + (n-1/2)\ln\left(\frac{d(n-1)}{e}\right) - \frac{1}{2}\ln\left(\frac{d}{e}\right) + 1\right\}. \\
&\leq 2n\exp\left\{-\zeta m/K_n + n\ln(dn)\right\}.
\end{aligned}$$

Combining with (62), we obtain the following probabilistic bound for the random variable C_n^d: in any dimension d, for any $r > 0$, if m and n satisfy, for Chebyshev and Jacobi polynomials, respectively,

$$\frac{m}{\ln m} \geq \begin{cases} \left(1 + r + \dfrac{n\ln(dn)}{\ln m}\right)\dfrac{n^{\ln 3/\ln 2}}{\zeta}, \\ \left(1 + r + \dfrac{n\ln(dn)}{\ln m}\right)\dfrac{n^{2\max\{\theta_1,\theta_2\}+2}}{\zeta}, \end{cases} \tag{74}$$

then

$$\Pr\left(C_n^d > 2\right) \leq 2nm^{-(r+1)} \leq 2m^{-r}. \tag{75}$$

The alternate bitstream model can also be used to encode any anchored set. In this case, the range of j in (72) is $\{1,\ldots,k\}$ which gives

$$\#(\mathcal{A}_n) \leq ((n-1)!)^2, \tag{76}$$

leading by a similar computation, for any $n \geq 2$, to

$$\Pr\left(\widetilde{C}_n > 2\right) \leq 2n\exp\left\{-\zeta m/\tilde{K}_n + 2n\ln n\right\}, \tag{77}$$

which is a sharper bound that the one obtained by using the first bitstream.

Combining with (62), we obtain the following probabilistic bound for the random variable $\tilde{C}_n$: in any dimension d, for any $r > 0$, if m and n satisfy, for Chebyshev and Jacobi polynomials, respectively,

$$\frac{m}{\ln m} \geq \begin{cases} \left(1 + r + \dfrac{2n\ln n}{\ln m}\right)\dfrac{n^{\ln 3/\ln 2}}{\zeta}, \\ \left(1 + r + \dfrac{2n\ln n}{\ln m}\right)\dfrac{n^{2\max\{\theta_1,\theta_2\}+2}}{\zeta}, \end{cases} \tag{78}$$

then

$$\Pr\left(\tilde{C}_n > 2\right) \leq 2nm^{-(r+1)} \leq 2m^{-r}. \tag{79}$$

We may summarize the probabilistic bounds established in this section as follows: for any $r > 0$ and any $n \geq 2$, one has $C_n^d \leq 2$ with probability larger than $1 - 2m^{-r}$ provided that (55) or (65) or (74) holds. Likewise, one has $\tilde{C}_n \leq 2$ with probability larger than $1 - 2m^{-r}$ provided that (58) or (69) or (78) holds.

3.3 Accuracy of the optimized discrete least-squares approximation

We are now in position to state our main results concerning the accuracy of the discrete least-squares approximation w_n and $\tilde{w}_n$ over the optimized index set Λ_m^{opt} and $\tilde{\Lambda}_m^{opt}$. These results show that the accuracy compares favorably with the best approximation error of the function u, measured either in L^∞ or L^2, using optimal choices of downward closed or anchored sets (which might differ from the sets Λ_m^{opt} and $\tilde{\Lambda}_m^{opt}$). We begin with a result expressed in probability.

Theorem 2. *Consider a function u defined on Γ in arbitrary dimension d and let $r > 0$. Under condition* (55)*, or* (65)*, or* (74)*, with n replaced by $2n-1$, it holds that*

$$\Pr\left(\|u - w_n\| \le (1 + 2\sqrt{2}) \min_{\Lambda \in \mathcal{M}_n^d} \min_{v \in \mathbb{P}_\Lambda} \|u - v\|_{L^\infty(\Gamma)}\right) \ge 1 - 2m^{-r}. \tag{80}$$

Under condition (58)*, or* (69)*, or* (78)*, it holds that*

$$\Pr\left(\|u - \tilde{w}_n\| \le (1 + 2\sqrt{2}) \min_{\Lambda \in \mathcal{A}_n} \min_{v \in \mathbb{P}_\Lambda} \|u - u_n\|_{L^\infty(\Gamma)}\right) \ge 1 - 2m^{-r}. \tag{81}$$

Proof. These estimates immediately follow from (48) and (49) combined with the probabilistic bounds from the previous section. □

We next give a result expressed in expectation for the truncated discrete least-square projection $T_\tau(w_n)$ and $T_\tau(\tilde{w}_n)$.

Theorem 3. *Consider a function u defined on Γ in arbitrary dimension d, such that $|u(y)| \le \tau$ for any $y \in \Gamma$, and let $r > 0$. Under condition* (55)*, or* (65)*, or* (74)*, it holds that*

$$\mathbb{E}(\|u - T_\tau(w_n)\|^2) \le 8\sqrt{2}\|u - u_n\|^2 + 8\tau^2 m^{-r}, \tag{82}$$

Under condition (58)*, or* (69)*, or* (78)*, it holds that*

$$\mathbb{E}(\|u - T_\tau(\tilde{w}_n)\|^2) \le 8\sqrt{2}\|u - u_n\|^2 + 8\tau^2 m^{-r}, \tag{83}$$

Proof. For (82), we distinguish between the two complementary events $\Omega_1 := \{C_n^d \le 2\}$ and $\Omega_2 := \{C_n^d > 2\}$ and write

$$\mathbb{E}(\|u - T_\tau(w_n)\|^2) = \mathbb{E}(\|u - T_\tau(w_n)\|^2 \,|\, \Omega_1) \Pr(\Omega_1) + \mathbb{E}(\|u - T_\tau(w_n)\|^2 \,|\, \Omega_2) \Pr(\Omega_2) =: E_1 + E_2. \tag{84}$$

Since $|u - T_\tau(w_n)| \le 2\tau$ and $\Pr(\Omega_2) \le 2m^{-r}$, the second term E_2 is bounded by $8\tau^2 m^{-r}$. For the first term E_1, we combine (46) and the fact that $|u - T_\tau(w_n)| \le |u - w_n|$ to obtain the bound

$$E_1 \le \mathbb{E}\Big((\|u - u_n\| + 2\sqrt{2}\|u - u_n\|_m)^2\Big) \le 4\sqrt{2}\|u - u_n\|^2 + 4\sqrt{2}\mathbb{E}(\|u - u_n\|_m^2) = 8\sqrt{2}\|u - u_n\|^2. \tag{85}$$

The proof of (83) is analogous. □

Remark 1. *The constants $1 + 2\sqrt{2}$ and $8\sqrt{2}$ in the above theorems can be reduced if one further restricts the regime between m and n so that C_n^d and $\tilde{C}_n$ are close to 1 with high probability.*

References

[1] A. Chkifa, A. Cohen, and C. Schwab: *High-dimensional adaptive sparse polynomila interpolation and applications to parametric PDEs*, Found. Comput. Math., 14: 601-633, 2014.

[2] A. Chkifa, A. Cohen, and C. Schwab: *Breaking the curse of dimensionality in parametric and stochastic PDEs*, J. Math. Pures Appl., 103(2):400–428, 2015.

[3] A. Chkifa, A. Cohen, G. Migliorati, F. Nobile, R. Tempone: *Discrete least squares polynomial approximation with random evaluations - application to parametric and stochastic elliptic PDEs*, ESAIM Math. Model. Numer. Anal., 49(3):815–837, 2015.

[4] A. Cohen, M.A. Davenport, D. Leviatan: *On the stability and accuracy of least squares approximations*, Found. Comput. Math., 13:819–834, 2013.

[5] A. Cohen, R. DeVore: *Approximation of high-dimensional parametric PDEs*, Acta Numer., 24:1–159, 2015.

[6] A. Cohen, R. DeVore and C. Schwab: *Analytic reguality and polynomial approximation of parametric and stochastic PDEs*, Analysis and Applications, 9, 11–47, 2011.

[7] F.Cucker, D.Zhou: *Learning Theory: An Approximation Theory Viewpoint*, Cambridge University Press, 2007.

[8] C. de Boor and A. Ron: *Computational aspects of polynomial interpolation in several variables*, Math. of Comp., 58: 705–727, 1992.

[9] N. Dyn and M. S. Floater: *Multivariate polynomial interpolation on lower sets*, J. Approx. Theory, 177: 3–42, 2014.

[10] L.Györfi, M.Kohler, A.Krzyzak, H.Walk: *A distribution-free theory of nonparametric regression*, Springer-Verlag, 2002.

[11] G.Migliorati: *Polynomial approximation by means of the random discrete L2 projection and application to inverse problems for PDEs with stochastic data*, Ph.D. thesis, Dipartimento di Matematica "Francesco Brioschi", Politecnico di Milano, Milano, Italy, and Centre de Mathématiques Appliquées, École Polytechnique, Palaiseau, France, 2013.

[12] J. Kuntzman: *Méthodes numériques - Interpolation, dérivées*, Dunod, Paris, 1959.

[13] G. Lorentz and R. Lorentz: *Solvability problems of bivariate interpolation I*, Constr. Approx., 2: 153–169, 1986

[14] G.Migliorati: *Multivariate Markov-type and Nikolskii-type inequalities for polynomials associated with downward closed multi-index sets*, J. Approx. Theory, 189:137–159, 2015.

[15] G.Migliorati, F.Nobile: *Analysis of discrete least squares on multivariate polynomial spaces with evaluations at low-discrepancy point sets*, J. Complexity, 31(4):517–542, 2015.

[16] G.Migliorati, F.Nobile, E.von Schwerin, R.Tempone: *Analysis of discrete L^2 projection on polynomial spaces with random evaluations*, Found. Comput. Math., 14:419–456, 2014.

[17] G.Migliorati, F.Nobile, E.von Schwerin, R.Tempone: *Approximation of Quantities of Interest in stochastic PDEs by the random discrete L^2 projection on polynomial spaces*, SIAM J. Sci. Comput., 35(3):A1440–A1460, 2013.

[18] G.Migliorati, F.Nobile, R.Tempone: *Convergence estimates in probability and in expectation for discrete least squares with noisy evaluations at random points*, J. Multivar. Anal., 142:167–182, 2015.

[19] T.Poggio, S.Smale: *The mathematics of learning: Dealing with data.* Notices Amer.Math.Soc., 50:537–544, 2003.

[20] V.N.Temlyakov: *Approximation in Learning Theory*, Constr.Approx., 27:33–74, 2008.

[21] V.Vapnik: *Statistical learning theory*, John Wiley & Sons, 1998.

Recent publications:

MATHEMATICS INSTITUTE OF COMPUTATIONAL SCIENCE AND ENGINEERING
Section of Mathematics
Ecole Polytechnique Fédérale
CH-1015 Lausanne

15.2015 ANDREA MANZONI, STEFANO PAGANI:
A certified reduced basis method for PDE-constrained parametric optimization problems by an adjoint-based approach

16.2015 SIMONE DEPARIS, DAVIDE FORTI, ALFIO QUARTERONI:
A fluid-structure interaction algorithm using radial basis function interpolation between non-conforming interfaces

17.2015 ASSYR ABDULLE, ONDREJ BUDAC:
A reduced basis finite element heterogeneous multiscale method for Stokes flow in porous media

18.2015 DANIEL KRESSNER, MICHAEL STEINLECHNER, BART VANDEREYCKEN:
Preconditioned low-rank Riemannian optimization for linear systems with tensor product structure

19.2015 ALESSANDRO S. PATELLI, LUCA DEDÈ, TONI LASSILA, ANDREA BARTEZZAGHI, ALFIO QUARTERONI:
Isogeometric approximation of cardiac electrophysiology models on surfaces: an accuracy study with application to the human left atrium

20.2015 MATTHIEU WILHELM, LUCA DEDÈ, LAURA M. SANGALLI, PIERRE WILHELM:
IGS: an IsoGeometric approach for Smoothing on surfaces

21.2015 SIMONE DEPARIS, DAVIDE FORTI, PAOLA GERVASIO, ALFIO QUARTERONI:
INTERNODES: an accurate interpolation-based method for coupling the Galerkin solutions of PDEs on subdomains featuring non-conforming interfaces

22.2015 ABDUL-LATEEF HAJI-ALI, FABIO NOBILE, LORENZO TAMELLINI, RAÙL TEMPONE:
Multi-index stochastic collocation for random PDEs

23.2015 SIMONE BRUGIAPAGLIA, FABIO NOBILE, STEFANO MICHELETTI, SIMONA PEROTTO:
A theoretical study of COmpRessed SolvING for advection-diffusion-reaction problems

24.2015 ANA ŠUŠNJARA, NATHANAËL PERRAUDIN, DANIEL KRESSNER, PIERRE VANDERGHEYNST:
Accelerate filtering on graphs using Lanczos mehtod

25.2015 FRANCESCO BALLARIN, ELENA FAGGIANO, SONIA IPPOLITO, ANDREA MANZONI, ALFIO QUARTERONI, GIANLUIGI ROZZA, ROBERTO SCROFANI:
Fast simulation of patient-specific haemodynamics of coronary artery bypass grafts based on a Pod-Galerkin method and a vascular shape parametrization

26.2015 FRANCISCO MACEDO:
Benchmark problems on stochastic automata networks in tensor train format

27.2015 JONAS BALLANI, DANIEL KRESSNER:
Reduced basis methods: from low-rank matrices to low-rank tensors

28.2015 ALBERT COHEN, GIOVANNI MIGLIORATI, FABIO NOBILE:
Discrete least-squares approximations over optimized downward closed polynomial spaces in arbitrary dimension