

# DFG-Schwerpunktprogramm 1324

„Extraktion quantifizierbarer Information aus komplexen Systemen“

## Convergence Rates for Greedy Algorithms in Reduced Basis Methods

P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova,  
P. Wojtaszczyk

Preprint 47



Edited by

AG Numerik/Optimierung  
Fachbereich 12 - Mathematik und Informatik  
Philipps-Universität Marburg  
Hans-Meerwein-Str.  
35032 Marburg

# DFG-Schwerpunktprogramm 1324

„Extraktion quantifizierbarer Information aus komplexen Systemen“

## Convergence Rates for Greedy Algorithms in Reduced Basis Methods

P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova,  
P. Wojtaszczyk

Preprint 47



The consecutive numbering of the publications is determined by their chronological order.

The aim of this preprint series is to make new research rapidly available for scientific discussion. Therefore, the responsibility for the contents is solely due to the authors. The publications will be distributed by the authors.

# Convergence Rates for Greedy Algorithms in Reduced Basis Methods \*

Peter Binev, Albert Cohen, Wolfgang Dahmen,  
Ronald DeVore, Guergana Petrova, and Przemyslaw Wojtaszczyk

May 18, 2010

## Abstract

The reduced basis method was introduced for the accurate online evaluation of solutions to a parameter dependent family of elliptic partial differential equations. Abstractly, it can be viewed as determining a “good”  $n$  dimensional space  $\mathcal{H}_n$  to be used in approximating the elements of a compact set  $\mathcal{F}$  in a Hilbert space  $\mathcal{H}$ . One, by now popular, computational approach is to find  $\mathcal{H}_n$  through a greedy strategy. It is natural to compare the approximation performance of the  $\mathcal{H}_n$  generated by this strategy with that of the Kolmogorov widths  $d_n(\mathcal{F})$  since the latter gives the smallest error that can be achieved by subspaces of fixed dimension  $n$ . The first such comparisons, given in [1], show that the approximation error,  $\sigma_n(\mathcal{F}) := \text{dist}(\mathcal{F}, \mathcal{H}_n)$ , obtained by the greedy strategy satisfies  $\sigma_n(\mathcal{F}) \leq Cn2^n d_n(\mathcal{F})$ . In this paper, various improvements of this result will be given. Among these, it is shown that whenever  $d_n(\mathcal{F}) \leq Mn^{-\alpha}$ , for all  $n > 0$ , and some  $M, \alpha > 0$ , we also have  $\sigma_n(\mathcal{F}) \leq C_\alpha Mn^{-\alpha}$  for all  $n > 0$ , where  $C_\alpha$  depends only on  $\alpha$ . Similar results are derived for generalized exponential rates of the form  $Me^{-an^\alpha}$ . The exact greedy algorithm is not always computationally feasible and a commonly used computationally friendly variant can be formulated as a “weak greedy algorithm”. The results of this paper are established for this version as well.

**Key Words:** Parameter dependent elliptic PFEs, reduced basis, greedy algorithm, weak greedy algorithm, Kolmogorov widths, convergence rates.

**Mathematical Subject Classification:** 41A25, 41A45, 41A65, 65N15.

---

\*This research was supported by the Fondation Sciences Mathématiques de Paris, the Office of Naval Research Contracts ONR-N00014-07-1-0978 ONR-N00014-08-1-1113, and ONR N00014-09-1-0107; the AFOSR Contract FA95500910500; the ARO/DoD Contracts W911NF-05-1-0227 and W911NF-07-1-0185; the NSF Grants DMS-0721621, DMS-0810869, DMS-0915104, and DMS 0915231; the Special Priority Program SPP 1324 funded by DFG; and the Polish MNiSW grant N201 269335.

# 1 Introduction

## 1.1 The reduced basis method

Many design problems in engineering can be formulated as optimal control problems with parametric PDE constraints. Typically, solving these control problems requires the frequent numerical solution of a PDE depending on dynamically updated parameters. A model situation that should serve as a guiding example, may be sketched as follows. Let  $\mathcal{H}$  be a Hilbert space with norm  $\|\cdot\|$  and let  $(a_\mu)_{\mu \in D}$  be a family of symmetric, continuous and elliptic bilinear forms

$$a_\mu : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{R},$$

depending on parameters  $\mu \in D$ , where  $D$  is a compact set in  $\mathbb{R}^d$ , such that all the norms  $\|\cdot\|_\mu := a_\mu(\cdot, \cdot)^{1/2}$  are uniformly equivalent to  $\|\cdot\|$ , i.e.

$$c^* \|v\| \leq \|v\|_\mu \leq C^* \|v\|, \quad v \in \mathcal{H},$$

holds for some constants  $0 < c^* \leq C^*$  independent of  $\mu$ . For a given linear functional  $L \in \mathcal{H}'$ , we denote by  $u_\mu$  the solution of the variational problem

$$a_\mu(u_\mu, v) = L(v), \quad v \in \mathcal{H}. \tag{1.1}$$

A typical problem is to control a functional quantity  $J(u_\mu)$  through the parameter  $\mu$ . For example, one may wish to drive the values  $\ell(u_\mu)$  towards some target value  $\ell^*$ , where  $\ell \in \mathcal{H}'$  is a given linear functional. If such an optimization is to be performed online, it would be infeasible to employ a standard highly accurate numerical solution of the PDE for each parameter value  $\mu \in D$ , arising e.g. in the course of a descent procedure.

The reduced basis method (see [1, 3, 4, 5, 6, 8]) has been introduced as a strategy to deal with such tasks by exploiting the fact that the set

$$\mathcal{F} := \{u_\mu : \mu \in D\}$$

of all parameter dependent solutions is compact in the energy space  $\mathcal{H}$ . This allows a model reduction along the following lines. One determines first offline (possibly at high computational cost) a fixed subspace  $\mathcal{H}_n$  of the energy space (of possibly small dimension  $n$ ) from which any element in the compact set  $\mathcal{F}$  can be well approximated. For instance, a low dimensional Galerkin scheme based on  $\mathcal{H}_n$  could be used for rapidly computing an approximate solution for any parameter query. The subspace itself is often generated as the span of solutions  $f_j := u_{\mu_j}$ ,  $j = 0, \dots, n-1$ , to specific parameters  $\mu_j$ , forming the *reduced basis*.

A key question therefore is to find “good” parameters  $\mu_j$ , or equivalently, good basis functions  $f_j \in \mathcal{F}$  in the preprocessing step. Here “good” means that relatively few basis functions suffice in order to warrant a desired accuracy for any parameter query.

## 1.2 A Greedy Algorithm

One by now prominent approach to the problem of finding a good parameter selection in the above setting is to employ a *greedy* strategy, see e.g [1, 3, 4, 5, 8]. For conceptual clarity we shall describe first a somewhat idealized version in the following abstract setting. We are given a compact set  $\mathcal{F}$  in a Hilbert space  $\mathcal{H}$ . The problem is then to find functions  $\{f_0, f_1, \dots, f_{n-1}\}$  so that each  $f \in \mathcal{F}$  is well approximated by the elements of the subspace  $F_n := \text{span}\{f_0, \dots, f_{n-1}\}$ .

The greedy algorithm for generating  $f_0, f_1, \dots$  proceeds as follows. First, one chooses a function  $f_0$  such that  $\|f_0\| = \max_{f \in \mathcal{F}} \|f\|$ , where again here and later  $\|\cdot\|$  is the norm on  $\mathcal{H}$  induced by its inner product  $\langle \cdot, \cdot \rangle$ . We define

$$\sigma_0 := \sigma_0(\mathcal{F}) := \|f_0\|.$$

The space  $F_1 := \text{span}\{f_0\}$  is the first subspace used to approximate the elements of  $\mathcal{F}$ . At the general step, if  $f_0, f_1, \dots, f_{n-1}$  have been chosen,  $F_n := \text{span}\{f_0, f_1, \dots, f_{n-1}\}$  and  $P_n$  is the projector onto  $F_n$ , then for any  $f \in \mathcal{F}$ , we denote by

$$\sigma_n(f) := \|f - P_n f\|$$

its error of best approximation in  $F_n$ . We define

$$\sigma_n := \sigma_n(\mathcal{F}) := \max_{f \in \mathcal{F}} \sigma_n(f),$$

which is the error in approximating the set  $\mathcal{F}$  using the elements of  $F_n$ . We now choose  $f_n$  as

$$f_n := \operatorname{argmax}_{f \in \mathcal{F}} \sigma_n(f), \tag{1.2}$$

as the next element in the greedy selection. Notice that the sequence  $f_0, f_1, \dots$  is generally not unique and neither is the sequence  $(\sigma_n(\mathcal{F}))_{n \geq 0}$  since it depends on this selection.

We emphasize that this greedy algorithm differs from those which aim at approximating *one given* function  $f$  by a finite linear combination of elements from a given dictionary (see [7] for a complete survey on this type of greedy algorithms). In the algorithm described above the objective is to *simultaneously* approximate all the functions in the class  $\mathcal{F}$  using finite dimensional subspaces generated from elements of  $\mathcal{F}$ .

The space  $F_n$  is an  $n$ -dimensional subspace of  $\mathcal{H}$  and to assess its quality it is natural to compare this subspace with the *best*  $n$ -dimensional subspace that would minimize the projection error for the whole class  $\mathcal{F}$  over all  $n$ -dimensional subspaces. This minimal error is given by the *Kolmogorov width*

$$d_n(\mathcal{F}) := \inf_{\dim(Y)=n} \sup_{f \in \mathcal{F}} \operatorname{dist}(f, Y), \tag{1.3}$$

where  $\operatorname{dist}(f, Y) := \min_{g \in Y} \|f - g\| = \|f - P_Y f\|$  with  $P_Y$  denoting the orthogonal projector onto  $Y$ , and where the infimum is taken over all  $n$  dimensional subspaces  $Y$  of  $\mathcal{H}$ . We refer to [2] for a general discussion on Kolmogorov width. We also define

$$d_0(\mathcal{F}) := \max_{f \in \mathcal{F}} \|f\| = \sigma_0(\mathcal{F}),$$

which corresponds to approximating by zero dimensional spaces.

Of course, if  $\sigma_n(\mathcal{F})$  decayed at a rate comparable to  $d_n(\mathcal{F})$ , this would mean that the greedy selection provides essentially the best possible accuracy attainable by  $n$ -dimensional subspaces. The central objective of this paper is to establish estimates for the sequence  $\sigma_n(\mathcal{F})$  in terms of the widths  $d_n(\mathcal{F})$ .

A startling result in this regard was recently given in [1] where it was proved that

$$\sigma_n(\mathcal{F}) \leq Cn2^n d_n(\mathcal{F}), \quad (1.4)$$

with  $C$  an absolute constant. While this is an interesting comparison, it is only useful if  $d_n(\mathcal{F})$  decays to zero faster than  $n^{-1}2^{-n}$ . On the other hand, since by assumption  $\mathcal{F}$  is compact, we always have  $\sigma_n(\mathcal{F}) \rightarrow 0$  as  $n \rightarrow \infty$ . A first central theme in this paper is to improve on (1.4) in the general Hilbert space setting regarding the following central questions: first, is (1.4) sharp? and second, what can be said for the whole range of possible decay rates of the widths  $d_n(\mathcal{F})$  between polynomial and exponential where (1.4) gives no information?

### 1.3 The Weak Greedy Algorithm

Before explaining our results in a bit more detail, let us pause to remark that, for the original motivation of the reduced basis method for solving parametric PDEs, the execution of the algorithm in the above form (1.2) is computationally not feasible, even in offline mode. In fact, to solve (1.2), either one would have to solve the PDE for essentially all parameters in (a highly resolved discretized) parameter domain or possibly employ a descent strategy to find this maximum, which would also involve a large number of such solves.

However, a variant of the greedy strategy (1.2) has been shown to become practically feasible in the context of (1.1) (see e.g. [5]) when replacing the error  $\|f - P_n f\|$  by a surrogate  $r_n(f)$  satisfying

$$c_r r_n(f) \leq \|f - P_n f\| \leq C_r r_n(f), \quad f \in \mathcal{H}, \quad (1.5)$$

for some fixed constants  $c_r, C_r$ . In fact, in the context of (1.1), one can find surrogates satisfying (1.5) such that

$$f_n := \operatorname{argmax}_{f \in \mathcal{F}} r_n(f), \quad (1.6)$$

can be computed in a much more efficient way (at least for a wide class of bilinear forms  $a_\mu$ ) than the exact greedy step (1.2).

What is important for us, is that the numerically feasible variant (1.6) produces an element  $f_n \in \mathcal{F}$  that actually satisfies

$$\|f_n - P_n f_n\| \geq \frac{c_r}{C_r} \max_{f \in \mathcal{F}} \|f - P_n f\|,$$

as one easily checks, where  $c_r, C_r$  are the constants from (1.5).

This allows us to embed such practical variants into a general abstract setting by considering the following *weak greedy algorithm*. This algorithm depends on a constant  $0 < \gamma \leq 1$  (in the above example  $\gamma = c_r/C_r$ ) which represents the relative accuracy to which we can compute norms. At the first step of the algorithm, one chooses a function  $f_0$  such that

$$\|f_0\| \geq \gamma \sigma_0, \quad \text{where} \quad \sigma_0 := \sigma_0(\mathcal{F}) := \max_{f \in \mathcal{F}} \|f\|.$$

At the general step, if  $f_0, \dots, f_{n-1}$  have been chosen,  $F_n := \text{span}\{f_0, \dots, f_{n-1}\}$ , and  $P_n$  is the projector onto  $F_n$ , then for any  $f \in \mathcal{F}$  we again denote by

$$\sigma_n(f) := \|f - P_n f\|$$

its best approximation error onto  $F_n$  and define

$$\sigma_n := \sigma_n(\mathcal{F}) := \max_{f \in \mathcal{F}} \sigma_n(f),$$

which is the error in approximating the class  $\mathcal{F}$  using the elements of  $F_n$ . We now choose  $f_n$  as an element in  $\mathcal{F}$  such that

$$\sigma_n(f_n) \geq \gamma \max_{f \in \mathcal{F}} \sigma_n(f), \tag{1.7}$$

as the next element in the greedy selection. Note that if  $\gamma = 1$ , then the weak greedy algorithm reduces to the greedy algorithm that we have introduced above.

Notice that similar to the greedy algorithm,  $(\sigma_n(\mathcal{F}))_{n \geq 0}$  is also monotone decreasing. Of course, the weak greedy algorithm does not give a unique sequence  $f_0, f_1, \dots$ , nor is the sequence  $(\sigma_n(\mathcal{F}))_{n \geq 0}$  unique. In all that follows, the notation reflects any sequences which can arise in the implementation of the weak greedy selection for the fixed value of  $\gamma$ .

Although not explicitly phrased in the framework of weak greedy schemes, it was also shown in [1] that, employing a surrogate as in (1.6), one obtains an estimate of the form

$$\sigma_n(\mathcal{F}) \leq C n \beta^n d_n(\mathcal{F}),$$

where  $\beta > 2$  is a constant depending on  $c_r, C_r$  in (1.5). Since  $\beta > 2$ , this further reduces the range of decay rates for which one obtains any information. Therefore, the same questions asked for the greedy algorithm can be posed for the weak greedy algorithm as well.

## 1.4 Objective, results and layout of the paper

Our central objective is to quantify the performance of the greedy and weak greedy algorithm by relating the distances  $\sigma_n(\mathcal{F})$  to the widths  $d_n(\mathcal{F})$ . We address these issues in the general setting where  $\mathcal{H}$  is an arbitrary separable Hilbert space and  $\mathcal{F} \subset \mathcal{H}$  is any compact subset. Since the greedy algorithm can be viewed as the particular instance  $\gamma = 1$  of the weak greedy algorithm, we formulate most of our results in the setting of the weak greedy algorithm.

In §2, we reformulate the problem in sequence space and present a matrix formulation of the weak greedy algorithm that will be used throughout the remainder of the paper. Using this formulation, we establish a fundamental lemma that expresses the following property: if  $\sigma_n(\mathcal{F})$  is flat in the sense that it does not decay more than a prescribed amount  $\theta$  between the index  $n$  and  $n + qm$  where  $q$  is a given constant, then  $\sigma_n(\mathcal{F})$  is controlled by  $d_m(\mathcal{F})$ .

In §3, we use this lemma to compare the rate of decay of  $\sigma_n(\mathcal{F})$  and  $d_n(\mathcal{F})$ . Our main result shows that for any  $\alpha > 0$ , whenever  $d_n(\mathcal{F}) \leq M n^{-\alpha}$  for all  $n > 0$  then  $\sigma_n(\mathcal{F}) \leq C M n^{-\alpha}$  for all  $n > 0$ , where  $C$  only depends on  $\alpha$  and  $\gamma$ . In the same spirit, we show that the exponential decay  $d_n(\mathcal{F}) \leq M e^{-an}$  with  $M, a, \alpha > 0$  implies a certain (different) exponential decay of  $\sigma_n(\mathcal{F})$ , giving a somewhat better resolution of the range of decay rates slower than  $2^{-n}$ . We show that these estimates for  $\sigma_n(\mathcal{F})$  are robust in the following sense:

if the  $f_j$  are computed up to an error of the order  $\varepsilon$ , we obtain similar estimates up to the addition of  $C\varepsilon$ . In the context of (1.1), such error typically results from the space discretization process when computing the  $f_j$ , e.g. by the finite element method.

As to directly comparing  $\sigma_n(\mathcal{F})$  with  $d_n(\mathcal{F})$ , note first that the optimal subspaces in (1.3) need not be spanned by elements of  $\mathcal{F}$ . In fact, we show in §4 that subspaces spanned by elements of  $\mathcal{F}$  itself will in general *not* give rise to distances that decay as fast as the  $n$ -widths, see Theorem 4.1. In particular, we can therefore not expect any result of the form  $\sigma_n(\mathcal{F}) \leq Cd_n(\mathcal{F})$ . On the other hand, we give for the exact greedy algorithm ( $\gamma = 1$ ) an improvement of (1.4) in that the factor  $n$  can be removed, and we show that the resulting bound cannot be improved.

## 2 A matrix formulation of the weak greedy algorithm

In general, the weak greedy algorithm does not terminate and we obtain an infinite sequence  $f_0, f_1, f_2, \dots$ . In order to have a consistent notation in what follows, we shall define  $f_m := 0$ ,  $m > N$ , if the algorithm terminates at  $N$ , i.e. if  $\sigma_N = 0$ . By  $(f_n^*)_{n \geq 0}$  we denote the orthonormal system obtained from the  $f_n$  by Gram-Schmidt orthogonalization, so that

$$P_n f = \sum_{i=0}^{n-1} \langle f, f_i^* \rangle f_i^*,$$

and in particular

$$f_i = P_{i+1} f_i = \sum_{j=0}^i a_{i,j} f_j^*, \quad a_{i,j} = \langle f_i, f_j^* \rangle, \quad j \leq i.$$

There is no loss of generality in assuming that the infinite dimensional Hilbert space  $\mathcal{H}$  is  $\ell_2(\mathbb{N} \cup \{0\})$  and that  $f_j^* = e_j$ , where  $e_j$  is the vector with a one in the coordinate indexed by  $j$  and all other coordinates zero, i.e.  $(e_j)_i = \delta_{j,i}$ . We adhere to this assumption throughout the paper.

We consider the lower triangular matrix

$$A := (a_{i,j})_{i,j=0}^{\infty}, \quad a_{i,j} := 0, \quad j > i.$$

This matrix incorporates all the information about the greedy algorithm on  $\mathcal{F}$ . The following two properties characterize a lower triangular matrix  $A$  generated by such a greedy algorithm. With the notation  $\sigma_n := \sigma_n(\mathcal{F})$ , we have

**P1:** The diagonal elements of  $A$  satisfy  $\gamma\sigma_n \leq |a_{n,n}| \leq \sigma_n$ .

**P2:** For every  $m \geq n$  one has  $\sum_{j=n}^m a_{m,j}^2 \leq \sigma_n^2$ .

Indeed, **P1** follows from

$$a_{n,n}^2 = \|f_n\|^2 - \|P_n f_n\|^2 = \|f_n - P_n f_n\|^2,$$

combined with the weak greedy selection property (1.7). To see **P2** we note that for  $m \geq n$

$$\sum_{j=n}^m a_{m,j}^2 = \|f_m - P_n f_m\|^2 \leq \max_{f \in \mathcal{F}} \|f - P_n f\|^2 = \sigma_n^2.$$

**Remark 2.1** *If  $A$  is any matrix satisfying **P1** and **P2** with  $\sigma_n$  a strictly decreasing sequence, then the rows of  $A$  form a compact subset of  $\ell_2(\mathbb{N} \cup \{0\})$ . If  $\mathcal{F}$  is the set consisting of these rows, then one of the possible realizations of the weak greedy algorithm with constant  $\gamma$  will choose the rows in order and  $A$  will be the resulting matrix.*

The following result plays a key role in our subsequent analysis of the convergence rate of the weak greedy algorithm.

**Lemma 2.2** *Fix any value  $0 < \theta < 1$  and define  $q := \lceil 2\gamma^{-1}\theta^{-1} \rceil^2$ . Suppose that for some  $m, n \in \mathbb{N}$ , we have*

$$\sigma_{n+qm} \geq \theta \sigma_n. \quad (2.1)$$

Then,

$$\sigma_n(\mathcal{F}) \leq q^{\frac{1}{2}} d_m(\mathcal{F}).$$

**Proof:** We consider the  $(qm + 1) \times (qm + 1)$  submatrix  $A'$  of  $A$  given by the entries of  $A$  in the rows and columns with indices in  $\{n, \dots, n + qm\}$ . We denote by  $g_j$ ,  $j = 0, 1, \dots, qm$ , the rows of  $A'$ . These are each vectors in  $\mathbb{R}^{qm+1}$  whose coordinates we index by  $0, 1, \dots, qm$ . Let  $Y$  be an  $m$ -dimensional space which realizes  $d_m(\mathcal{F})$  and let  $Y_m$  be the restriction of  $Y$  to the coordinates in  $\{n, n + 1, \dots, n + qm\}$ . Then, for each  $j$  the projection  $y_j$  of  $g_j$  onto  $Y_m$  satisfies

$$\|g_j - y_j\| \leq d_m := d_m(\mathcal{F}), \quad j = 0, 1, \dots, \dots, qm.$$

We can without loss of generality assume that  $\dim(Y_m) = m$ . Let  $\phi_1, \dots, \phi_m$  be an orthonormal basis for  $Y_m$ . Since each  $\phi_i = (\phi_i(j))_{j=0}^{qm}$ ,  $i = 1, \dots, m$ , has norm one, for at least one value of  $j \in \{0, \dots, qm\}$ , we have

$$\sum_{i=1}^m |\phi_i(j)|^2 \leq q^{-1}. \quad (2.2)$$

We fix this value of  $j$  and write  $y_j = \sum_{i=1}^m \langle g_j, \phi_i \rangle \phi_i$ . This means that the  $j$ -th coordinate  $y_j(j)$  of  $y_j$  satisfies

$$|y_j(j)| = \left| \sum_{i=1}^m \langle g_j, \phi_i \rangle \phi_i(j) \right| \leq \left( \sum_{i=1}^m |\langle g_j, \phi_i \rangle|^2 \right)^{1/2} \left( \sum_{i=1}^m |\phi_i(j)|^2 \right)^{1/2} \leq q^{-\frac{1}{2}} \|g_j\| \leq q^{-\frac{1}{2}} \sigma_n,$$

where we have used (2.2) and Property **P2**. From **P1**, (2.1) and the definition of  $q$ , we have

$$|g_j(j)| = |a_{n+j, n+j}| \geq \gamma \sigma_{n+j} \geq \gamma \theta \sigma_n \geq 2q^{-\frac{1}{2}} \sigma_n.$$

It follows that

$$q^{-\frac{1}{2}} \sigma_n \leq |g_j(j) - y_j(j)| \leq \|g_j - y_j\| \leq d_m,$$

which completes the proof of the lemma.  $\square$

### 3 Convergence estimates

In this section we use Lemma 2.2 to estimate the rate of decay of  $\sigma_n(\mathcal{F})$  for the weak greedy algorithm from the rate of decay of the Kolmogorov width  $d_n(\mathcal{F})$ .

#### 3.1 Polynomial Decay

We first consider the case where  $d_n(\mathcal{F})$  has polynomial decay and show that in such a case  $\sigma_n(\mathcal{F})$  inherits the same rate of decay.

**Theorem 3.1** *Let  $0 < \gamma \leq 1$  be the parameter in the weak greedy algorithm. Suppose that  $d_0(\mathcal{F}) \leq M$  and*

$$d_n(\mathcal{F}) \leq Mn^{-\alpha}, \quad n > 0,$$

for some  $M > 0$  and  $\alpha > 0$ . Then,

$$\sigma_n = \sigma_n(\mathcal{F}) \leq CMn^{-\alpha}, \quad n > 0, \tag{3.1}$$

with  $C := q^{\frac{1}{2}}(4q)^\alpha$  and  $q := \lceil 2^{\alpha+1}\gamma^{-1} \rceil^2$ .

**Proof:** First, we choose  $\theta = 2^{-\alpha}$ . It follows that  $q = \lceil 2\gamma^{-1}\theta^{-1} \rceil^2$  which is the relation needed in Lemma 2.2. Next, we define  $N_0 := 4q$ . Note that  $q \geq 9$  and  $N_0 \geq 36$ . Then, the definition of  $C$  shows that (3.1) holds for  $1 \leq n \leq N_0$  since

$$\sigma_n \leq \sigma_0 \leq M \leq MN_0^\alpha n^{-\alpha} \leq CMn^{-\alpha}, \quad n \leq N_0.$$

We suppose now that (3.1) does not hold for some  $n > N_0$  and draw a contradiction. To start with we let  $N$  be the smallest integer  $N > N_0$  for which (3.1) does not hold. Then

$$CMN^{-\alpha} < \sigma_N. \tag{3.2}$$

We define  $n$  as the smallest integer for which  $n^\alpha N^{-\alpha} \geq \theta$ , i.e.  $n := \lceil \theta^{1/\alpha} N \rceil = \lceil N/2 \rceil$ . We then have

$$\sigma_n \leq CMn^{-\alpha} = CMN^\alpha n^{-\alpha} N^{-\alpha} \leq \theta^{-1} \sigma_N.$$

We now take this value of  $n$  and take  $m$  as the largest integer such that  $n + qm \leq N$ , i.e.  $m := \lfloor (N - n)/q \rfloor$ . We can apply Lemma 2.2 and conclude that

$$\sigma_N \leq \sigma_n \leq q^{\frac{1}{2}} d_m(\mathcal{F}) \leq q^{\frac{1}{2}} M m^{-\alpha} = (q^{\frac{1}{2}} N^\alpha m^{-\alpha}) MN^{-\alpha}. \tag{3.3}$$

Since  $N > N_0 > 4q - 2$ , it follows that

$$\begin{aligned} 4qm &= 4q \lfloor (N - n)/q \rfloor \\ &\geq 4(N - n - q + 1) \\ &= 4N - 2(2 \lceil N/2 \rceil) - 4q + 4 \\ &\geq 2N - 2 - 4q + 4 \geq N, \end{aligned}$$

where we have used the fact that if  $a$  and  $b$  are positive integers, then  $a \lfloor \frac{b}{a} \rfloor \geq b - a + 1$  and  $a \lceil \frac{b}{a} \rceil \leq b + a - 1$ . Therefore, we have  $\frac{N}{m} \leq 4q$ . Combining (3.2) with (3.3), we obtain

$$C < q^{\frac{1}{2}} (N/m)^\alpha \leq q^{\frac{1}{2}} (4q)^\alpha.$$

This contradicts the definition of  $C$ . □

## 3.2 Exponential Rates

We next use the same ideas to derive bounds when  $d_n(\mathcal{F})$  decays to zero like  $Me^{-an^\alpha}$  for some  $M, a, \alpha > 0$ .

**Theorem 3.2** *Let  $0 < \gamma \leq 1$  be the parameter in the weak greedy algorithm. Suppose that*

$$d_n(\mathcal{F}) \leq Me^{-an^\alpha}, \quad n \geq 0,$$

for some  $M, a, \alpha > 0$ . Then setting  $\beta := \frac{\alpha}{\alpha+1}$ , one has

$$\sigma_n(\mathcal{F}) \leq CMe^{-cn^\beta}, \quad n \geq 0, \quad (3.4)$$

whenever for any fixed  $0 < \theta < 1$ , one takes  $c := \min\{|\ln \theta|, (4q)^{-\alpha}a\}$ ,  $C := \max\{e^{cN_0^\beta}, q^{\frac{1}{2}}\}$ ,  $q := \lceil 2\gamma^{-1}\theta^{-1} \rceil^2$  and  $N_0 := \lceil (8q)^{\frac{1}{1-\beta}} \rceil = \lceil (8q)^{\alpha+1} \rceil$ .

**Proof:** The proof is similar to that of Theorem 3.1. By our definition of  $C$  we find that (3.4) holds for  $n \leq N_0$  since we have

$$\sigma_n \leq \sigma_0 \leq \sigma_0 e^{cN_0^\beta} e^{-cn^\beta} \leq CMe^{-cn^\beta}, \quad n \leq N_0.$$

We suppose now that (3.4) does not hold for some  $n > N_0$  and draw a contradiction. Let  $N > N_0$  be the smallest integer for which

$$CMe^{-cN^\beta} < \sigma_N.$$

We now let  $m$  be any positive integer for which

$$e^{c(N-qm)^\beta} e^{-cN^\beta} \geq \theta, \quad \text{or equivalently} \quad c(N^\beta - (N - qm)^\beta) \leq |\ln \theta|. \quad (3.5)$$

Then, using that (3.4) holds for  $N - qm$ , (3.5) yields

$$\sigma_{N-qm} \leq \theta^{-1} \sigma_N.$$

Now we can apply Lemma 2.2 and conclude that

$$\sigma_N \leq \sigma_{N-qm} \leq q^{\frac{1}{2}} d_m(\mathcal{F}) \leq q^{\frac{1}{2}} Me^{-am^\alpha}.$$

We are left to show that there is a choice of  $m$  satisfying (3.5) and such that

$$q^{\frac{1}{2}} Me^{-am^\alpha} \leq CMe^{-cN^\beta}, \quad (3.6)$$

since this will be the desired contradiction. Taking logarithms, (3.6) is equivalent to

$$\frac{1}{2} \ln q - \ln C \leq am^\alpha - cN^\beta. \quad (3.7)$$

Now to show that there is an  $m$  which satisfies both (3.7) and (3.5), we consider

$$m := \lfloor \frac{N^{1-\beta}}{2q} \rfloor.$$

Clearly  $m \leq \frac{N}{2q}$  and therefore  $N - qm \geq N/2$ . From the mean value theorem, we have for some  $\xi \in (N - qm, N)$

$$N^\beta - (N - qm)^\beta = \beta \xi^{\beta-1} qm \leq qm \beta (N/2)^{\beta-1} \leq \frac{N^{1-\beta}}{2} \beta (N/2)^{\beta-1} \leq 2^{-\beta} \beta \leq 1.$$

Thus, (3.5) will be satisfied for the definition of  $c$ . Now let us check (3.7). We first remark that since  $N > N_0$ , we have  $m \geq 4$  and therefore  $m \geq \frac{N^{1-\beta}}{4q}$ . From the definition of  $c$ , we thus have

$$am^\alpha - cN^\beta \geq a \left( \frac{N^{1-\beta}}{4q} \right)^\alpha - cN^\beta = (a(4q)^{-\alpha} - c)N^\beta \geq 0.$$

Since by definition  $C \geq q^{\frac{1}{2}}$ , we have verified (3.7) and completed the proof.  $\square$

**Remark 3.3** *Let us also note that under the assumptions of Theorem 3.2 for the exact greedy algorithm, (1.4) gives*

$$\sigma_n(\mathcal{F}) \leq CMn2^n e^{-an^\alpha}. \quad (3.8)$$

*This latter estimate is asymptotically sharper than (3.4) in Theorem 3.2 as long as  $\alpha > 1$ , whereas (3.4) is asymptotically sharper for  $\alpha < 1$ . For  $\alpha = 1$ , (3.8) is asymptotically sharper provided  $a > \ln 2$ , and not if  $a \leq \ln 2$ .*

### 3.3 Robustness

In a PDE context, for example as described by (1.1), the reduced basis functions  $f_j$  cannot be computed exactly. Instead, one relies on a space discretization, such as by finite elements, in order to approximate them by functions  $\hat{f}_j$  with some prescribed accuracy. This leads to a modification of the spaces  $F_n$  generated by the weak greedy algorithm that we shall now describe and analyze.

We denote by  $\varepsilon > 0$  the accuracy to which a given element  $f \in \mathcal{F}$  is computed when it is chosen as a member of the reduced basis. At the first step of the weak greedy algorithm we determine  $f_0 \in \mathcal{F}$  such that  $\|f_0\| \geq \gamma \sup_{f \in \mathcal{F}} \|f\|$ . However, rather than receiving  $f_0$  we receive the noisy version  $\hat{f}_0$  which is not even necessarily in  $\mathcal{F}$ . All we know about  $\hat{f}_0$  is that

$$\|f_0 - \hat{f}_0\| \leq \varepsilon.$$

This means that our first space is now  $\hat{F}_1 := \text{span}\{\hat{f}_0\}$  and it is used for the selection of  $f_1$ . In general, if  $\hat{f}_0, \dots, \hat{f}_{n-1}$  have been found and  $\hat{F}_n := \text{span}\{\hat{f}_0, \dots, \hat{f}_{n-1}\}$  then we determine  $f_n$  such that  $\text{dist}(f_n, \hat{F}_n) \geq \gamma \sup_{f \in \mathcal{F}} \text{dist}(f, \hat{F}_n)$ . Rather than  $f_n$  we receive the noisy version  $\hat{f}_n$  for which we only know  $\|f_n - \hat{f}_n\| \leq \varepsilon$ . Let us define

$$\hat{\sigma}_n(\mathcal{F}) := \sup_{f \in \mathcal{F}} \text{dist}(f, \hat{F}_n),$$

which is the performance of this noisy-weak greedy algorithm on  $\mathcal{F}$ .

The following shows that the result of Theorem 3.1 is robust with respect to this perturbation.

**Theorem 3.4** *Suppose that  $d_0(\mathcal{F}) \leq M$  and*

$$d_n(\mathcal{F}) \leq Mn^{-\alpha}, \quad n > 0, \quad (3.9)$$

for some  $M, \alpha > 0$ . Then,

$$\hat{\sigma}_n(\mathcal{F}) \leq C \max\{Mn^{-\alpha}, \varepsilon\}, \quad n > 0, \quad (3.10)$$

with  $C := (1 + 2^\alpha)C(\alpha, \frac{\gamma}{2})$ , where  $C(\alpha, \frac{\gamma}{2})$  is the constant of Theorem 3.1 for  $\alpha$  and  $\frac{\gamma}{2}$ .

**Proof:** Let  $N$  be the largest integer such that  $MN^{-\alpha} \geq \varepsilon$ . We shall first prove

$$\hat{\sigma}_n(\mathcal{F}) \leq C \max\{Mn^{-\alpha}, \varepsilon\}, \quad 1 \leq n \leq N + 1. \quad (3.11)$$

Fix a value of  $n \in \{1, \dots, N\}$ . If  $\hat{\sigma}_n(\mathcal{F}) \leq C\varepsilon$ , then we have verified (3.11) for this value of  $n$ , as well as for the next ones. So we can assume  $\hat{\sigma}_n(\mathcal{F}) > C\varepsilon$ . This implies that  $\hat{\sigma}_j(\mathcal{F}) > C\varepsilon$  for all  $j = 1, \dots, n$ .

We define  $\mathcal{F}_n := \mathcal{F} \cup \{\hat{f}_0, \dots, \hat{f}_{n-1}\}$  and first examine the Kolmogorov widths of  $\mathcal{F}_n$ . Since each of the  $\hat{f}_j$  are within  $\varepsilon$  of an element from  $\mathcal{F}$ , using (3.9) we have for any  $1 \leq m \leq 2N$ ,

$$d_m(\mathcal{F}_n) \leq Mm^{-\alpha} + \varepsilon \leq (1 + 2^\alpha)Mm^{-\alpha}.$$

For  $m > 2N$ , using the fact that  $\mathcal{F}_n$  has been obtained by adding  $n$  functions to  $\mathcal{F}$ , we have

$$d_m(\mathcal{F}_n) \leq d_{m-n}(\mathcal{F}) \leq M(m-n)^{-\alpha} \leq M(m/2)^{-\alpha} \leq 2^\alpha Mm^{-\alpha}.$$

In summary,

$$d_m(\mathcal{F}_n) \leq (1 + 2^\alpha)Mm^{-\alpha}, \quad m > 0.$$

Next, we examine the weak greedy algorithm applied to  $\mathcal{F}_n$  with the weakness parameter  $\frac{\gamma}{2}$ . We claim that the functions  $\hat{f}_0, \dots, \hat{f}_n$  could be chosen in this order by this algorithm. First, we define  $\hat{F}_0 := \{0\}$ , and therefore for any  $f$ ,  $\text{dist}(f, \hat{F}_0) = \|f\|$ . Observe that for  $m = 0, \dots, n$ ,

$$\text{dist}(\hat{f}_m, \hat{F}_m) \geq \text{dist}(f_m, \hat{F}_m) - \varepsilon \geq \gamma \sup_{f \in \mathcal{F}} \text{dist}(f, \hat{F}_m) - \varepsilon \geq \gamma \sup_{f \in \mathcal{F}_n} \text{dist}(f, \hat{F}_m) - 2\varepsilon,$$

and since  $\sup_{f \in \mathcal{F}_n} \text{dist}(f, \hat{F}_m) \geq \hat{\sigma}_m(\mathcal{F}) > C\varepsilon$  and  $C = (1 + 2^\alpha)C(\alpha, \frac{\gamma}{2}) \geq 4\gamma^{-1}$ , we also have

$$\varepsilon < \frac{\gamma}{4} \sup_{f \in \mathcal{F}_n} \text{dist}(f, \hat{F}_m).$$

Therefore,

$$\text{dist}(\hat{f}_m, \hat{F}_m) \geq \frac{\gamma}{2} \sup_{f \in \mathcal{F}_n} \text{dist}(f, \hat{F}_m).$$

We have thus verified that  $\hat{f}_0, \dots, \hat{f}_n$  is an allowable sequence to choose under the hypothesis of Theorem 3.1 for the set  $\mathcal{F}_n$  and the parameters  $\alpha$  and  $\frac{\gamma}{2}$ . If we apply the theorem, we thus obtain that for  $m > 0$ ,

$$\hat{\sigma}_m(\mathcal{F}_n) \leq C(\alpha, \frac{\gamma}{2})(1 + 2^\alpha)Mm^{-\alpha} \leq CMm^{-\alpha}.$$

Since  $\hat{f}_0, \dots, \hat{f}_n$  is an allowable sequence,

$$\hat{\sigma}_m(\mathcal{F}) = \sup_{f \in \mathcal{F}} \text{dist}(f, \hat{F}_m) \leq \hat{\sigma}_m(\mathcal{F}_n), \quad 1 \leq m \leq n+1.$$

Thus, we have proved (3.11) for  $1 \leq n \leq N+1$ . Finally, if  $n > N+1$ , then

$$\hat{\sigma}_n(\mathcal{F}) \leq \hat{\sigma}_{N+1}(\mathcal{F}) \leq CM(N+1)^{-\alpha} \leq C\varepsilon.$$

This shows that (3.10) holds for  $n \geq N$  as well. □

By similar arguments, one obtains the following robust version of Theorem 3.2.

**Theorem 3.5** *Suppose that*

$$d_n(\mathcal{F}) \leq Me^{-an^\alpha}, \quad n \geq 0,$$

for some  $M, a, \alpha > 0$ . Then, setting  $\beta := \frac{\alpha}{1+\alpha}$ , one has

$$\hat{\sigma}_n(\mathcal{F}) \leq C \max\{Me^{-cn^\beta}, \varepsilon\}, \quad n \geq 0,$$

where  $c, C$  are constants that depend on  $a, \alpha$  and on the weakness parameter  $\gamma$ .

We do not provide the proof which is similar to the proof of Theorem 3.4.

## 4 Direct comparison between $\sigma_n(\mathcal{F})$ and $d_n(\mathcal{F})$

The results of the previous section show that estimates on the rate of decay of  $\sigma_n(\mathcal{F})$  may be derived from estimates on the rate of decay of  $d_n(\mathcal{F})$ , but they do not provide a direct comparison between the individual quantities  $\sigma_n(\mathcal{F})$  and  $d_n(\mathcal{F})$ . In this section, we discuss such comparison and show, in particular, that estimates of the form  $\sigma_n(\mathcal{F}) \leq Cd_n(\mathcal{F})$  cannot be met.

### 4.1 Good approximating subspaces

The optimal subspace for the Kolmogorov width  $d_n(\mathcal{F})$  is not necessarily generated by elements from  $\mathcal{F}$ . However, we are considering only subspaces of the latter type in our approximation algorithms. Thus, it is of interest to understand how well we can approximate  $\mathcal{F}$  when using spaces  $Y$  of dimension  $n$ , where  $Y$  is spanned by  $n$  elements from  $\mathcal{F}$ . That is,  $Y = \text{span}\{g_1, \dots, g_n\}$ , where  $g_i \in \mathcal{F}$ ,  $i = 1, \dots, n$ . We define  $\mathcal{M}_n(\mathcal{F})$  to be the collection of all such subspaces  $Y$  and set

$$\bar{d}_n(\mathcal{F}) := \inf_{Y \in \mathcal{M}_n(\mathcal{F})} \sup_{f \in \mathcal{F}} \text{dist}(f, Y).$$

Clearly  $\bar{d}_n(\mathcal{F}) \geq d_n(\mathcal{F})$ . The following result draws a comparison between  $d_n$  and  $\bar{d}_n$ .

**Theorem 4.1** *The following holds:*

- (i) For any compact set  $\mathcal{F}$  and any  $n \geq 0$ , we have  $\bar{d}_n(\mathcal{F}) \leq (n+1)d_n(\mathcal{F})$ .
- (ii) Given any  $n > 0$  and  $\epsilon > 0$ , there is a set  $\mathcal{F}$  such that  $\bar{d}_n(\mathcal{F}) \geq (n-1-\epsilon)d_n(\mathcal{F})$ .

**Proof:** To see (i), let  $Y$  be the optimal Kolmogorov space for  $\mathcal{F}$  of dimension  $n$ . We let  $\phi_1, \dots, \phi_n$  be an orthonormal basis for  $Y$  and let  $P$  be the projection onto  $Y$ . Therefore, for any  $f \in \mathcal{F}$ ,

$$Pf = \sum_{j=1}^n \langle f, \phi_j \rangle \phi_j,$$

and  $\|f - Pf\| \leq d_n(\mathcal{F})$ ,  $f \in \mathcal{F}$ . For any  $\{f_1, \dots, f_n\} \subset \mathcal{F}$ , we consider the determinant  $D(f_1, \dots, f_n) := \det(\langle f_i, \phi_j \rangle)$  and choose  $f_1, \dots, f_n$  to maximize the absolute value of this determinant. For any  $f \in \mathcal{F}$ , we can write  $Pf = \sum_{i=1}^n \alpha_i Pf_i$ , where

$$\alpha_i = \frac{D(Pf_1, \dots, Pf_{i-1}, Pf, Pf_{i+1}, \dots, Pf_n)}{D(Pf_1, \dots, Pf_n)} = \frac{D(f_1, \dots, f_{i-1}, f, f_{i+1}, \dots, f_n)}{D(f_1, \dots, f_n)},$$

because  $P$  is an orthogonal projector. From the definition of  $f_1, \dots, f_n$ , we have that  $|\alpha_i| \leq 1$ . Now, we can write

$$f - \sum_{i=1}^n \alpha_i f_i = f - Pf + \sum_{i=1}^n \alpha_i [Pf_i - f_i].$$

It follows that

$$\|f - \sum_{i=1}^n \alpha_i f_i\| \leq (n+1)d_n(\mathcal{F}),$$

as claimed.

As for (ii) let  $\mu > 0$  and  $\mathcal{F} = \{f_1, f_2, \dots, f_n, f_{n+1}, 0, \dots, 0, \dots\}$ , where  $f_j := e_j + \mu e_{n+1}$ ,  $j = 1, \dots, n$ , and  $f_{n+1} := \sum_{j=1}^n e_j + \mu e_{n+1}$ . By choosing  $Y^* := \text{span}\{e_1, \dots, e_n\}$ , we see that  $d_n(\mathcal{F}) \leq \mu$ . Now consider any  $n$  dimensional space  $Y = \text{span}\{g_1, \dots, g_n\}$  with the  $g_j \in \mathcal{F}$ . If  $f_{n+1}$  is not one of the  $g_j$ , then the only choice of  $Y$  is  $Y = \text{span}\{f_1, \dots, f_n\}$ . For this choice of  $Y$ , the vector  $y := -e_{n+1} + \mu \sum_{j=1}^n e_j$  is orthogonal to  $Y$  and has norm  $\|y\| = \sqrt{1 + n\mu^2}$ . So the error in approximating  $f_{n+1}$  by the elements of  $Y$  is  $(1 + n\mu^2)^{-1/2} \langle y, f_{n+1} \rangle = (1 + n\mu^2)^{-1/2} (n-1)\mu$ . On the other hand, if  $f_{n+1}$  is included as one of the  $g_1, \dots, g_n$ , then one of the  $f_j$ ,  $j = 1, \dots, n$ , is not included. Without loss of generality, we can assume that this is  $f_1$ . Now, we can take  $y := (n-2)\mu e_1 - \mu \sum_{j=2}^n e_j + e_{n+1}$  which has norm  $\sqrt{1 + \mu^2(n^2 - 3n + 3)}$ . This gives that the error in approximating  $f_1$  is  $(n-1)\mu(1 + \mu^2(n^2 - 3n + 3))^{-1/2}$ . So for either of these two cases, if we take  $\mu$  sufficiently small we obtain (ii).  $\square$

**Remark 4.2** Since  $\sigma_n(\mathcal{F}) \geq \bar{d}_n(\mathcal{F})$ , an immediate consequence of the second item in the above theorem is that given any  $n > 0$  and  $\epsilon > 0$ , there is a set  $\mathcal{F}$  such that  $\sigma_n(\mathcal{F}) \geq (n-1-\epsilon)d_n(\mathcal{F})$ . In particular, this shows that one cannot expect a general estimate of the form  $\sigma_n(\mathcal{F}) \leq Cd_n(\mathcal{F})$ . Note also that there is no contradiction between this second item and the result of Theorem 3.1 concerning polynomial rate of decay.

## 4.2 An improvement of (1.4)

We shall now show that (1.4) can be improved by replacing the factor  $n2^n$  by  $2^n$ . Here, we work in the specific setting of the exact greedy algorithm corresponding to  $\gamma = 1$ , and use the matrix formulation of the greedy algorithm introduced in §2. The main vehicle for proving this upper estimate is to understand how well the coordinate vectors  $e_i$  can be represented as a linear combination of the  $f_j$ ,  $j = 0, 1, \dots, i$ . This information is provided by the inverse of the matrix  $A$ . We shall need the following lemma reflecting this fact.

**Lemma 4.3** *For each  $n \geq 1$ , we can write*

$$\sigma_n e_n = \sum_{j=0}^n b_n(j) f_j, \quad (4.1)$$

where

$$\sum_{j=0}^n |b_n(j)| \leq 2^n. \quad (4.2)$$

**Proof:** Bounds of this type were given in [1]. For completeness, we repeat the simple proof. Let  $b_n := (b_n(j))_{j=0}^n$  and  $A_n$  be the submatrix of  $A$  formed by all entries  $a_{i,j}$ ,  $0 \leq i, j \leq n$ . The  $j$ -th row  $(a_{j,i})_{i=0}^n$  of  $A$  can be identified with  $f_j$  and (4.1) can be then rephrased as

$$A_n^t b_n = \sigma_n e_n.$$

Since this is an upper triangular system, it can be solved by back substitution which yields  $b_n(n) = 1$  and for  $j = n - 1, \dots, 0$ ,

$$\sigma_j b_n(j) = - \sum_{i=j+1}^n b_n(i) a_{i,j}. \quad (4.3)$$

From **P2**, we have  $|a_{i,j}| \leq \sigma_j$ ,  $i \geq j + 1$ . Using this in (4.3) gives

$$|b_n(j)| \leq \sum_{i=j+1}^n |b_n(i)|.$$

Since  $b_n(n) = 1$ , we obtain  $|b_n(n - 1)| \leq 1$  and in general  $|b_n(n - j)| \leq 2^{j-1}$ ,  $1 \leq j \leq n$ . We obtain (4.2) by adding these inequalities.  $\square$

**Theorem 4.4** *Let  $\mathcal{F}$  be an arbitrary compact set in a Hilbert space  $\mathcal{H}$ . For each  $n = 1, 2, \dots$  we have*

$$\sigma_n(\mathcal{F}) \leq \frac{2^{n+1}}{\sqrt{3}} d_n(\mathcal{F}). \quad (4.4)$$

**Proof.** Let  $Y$  be a space of dimension  $n$  which attains  $d_n(\mathcal{F})$ . Then, in particular, we have

$$\text{dist}(f_j, Y) \leq d_n(\mathcal{F}), \quad j = 0, \dots, n. \quad (4.5)$$

Since the  $k$ -th coordinate of each of the vectors  $f_j$ ,  $j = 0, \dots, n$ , is zero for  $k \geq n + 1$ , we can achieve (4.5) for a space  $Y$  of dimension  $n$  contained in the  $\text{span}\{e_0, \dots, e_n\}$ . By (4.5), we can find for each  $j = 0, \dots, n$ , a  $y_j \in Y$  such that  $\|f_j - y_j\| \leq d_n(\mathcal{F})$ . Setting  $v_i := \sum_{k=0}^i b_i(k)y_k \in Y$ , it follows from Lemma 4.3 that

$$\text{dist}(\sigma_i e_i, Y) = \|\sigma_i e_i - v_i\| \leq 2^i d_n(\mathcal{F}), \quad i = 0, 1, \dots, n.$$

Now  $Y$  can be characterized by a unit vector  $y = (y(0), \dots, y(n)) \in \text{span}\{e_0, \dots, e_n\}$  which is orthogonal to  $Y$ . For this  $y$ , we have from (4.2) that

$$\sigma_n |y(i)| \leq |\sigma_i y(i)| = |\langle \sigma_i e_i, y \rangle| = |\langle \sigma_i e_i - v_i, y \rangle| \leq \|\sigma_i e_i - v_i\| \leq 2^i d_n(\mathcal{F}), \quad i = 0, \dots, n.$$

Thus,

$$\sigma_n^2 = \sigma_n^2 \sum_{i=0}^n |y(i)|^2 \leq d_n(\mathcal{F})^2 \sum_{i=0}^n 2^{2i} \leq \frac{2^{2n+2} - 1}{3} d_n(\mathcal{F})^2 \leq \frac{2^{2n+2}}{3} d_n(\mathcal{F})^2,$$

which completes the proof of the theorem.  $\square$

For any  $n > 0$ , we finally present an example of a set  $\mathcal{F} = \mathcal{F}_n$  such that for the exact greedy algorithm

$$\sigma_n(\mathcal{F}) \geq (1 - \varepsilon) 2^n d_n(\mathcal{F}) \quad \text{for all } \varepsilon > 0,$$

which reveals that Theorem 4.4 is sharp safe for the constant  $2/\sqrt{3}$ . An example with a similar type of implication was also given in [1]. In our example, the lower triangular matrix  $A$ , defined in §2, has the following form: the  $j$ -th column has  $\sigma_j$  in position  $(j, j)$  and all entries in positions  $(i, j)$  with  $0 \leq j < i \leq n$  will equal a fixed constant  $a_j$ . All rows of  $A$  with indices  $i > n$  are identically zero.

Let  $t_0 = 1, t_1, t_2, \dots, t_n$  be positive real numbers. We define

$$\sigma_0 := 1, \quad \sigma_{k+1} := \prod_{j=0}^k \frac{2\sqrt{t_j t_{j+1}}}{t_j + t_{j+1}} = \sigma_k \frac{2\sqrt{t_k t_{k+1}}}{t_k + t_{k+1}}, \quad k = 0, \dots, n-1.$$

Since each factor in the product is less or equal to 1, the sequence  $\sigma_k$  is non-increasing. Furthermore, we define  $a_k := \sigma_k \frac{t_k - t_{k+1}}{t_k + t_{k+1}}$ . These values for  $a_k$  serve to define the matrix  $A$  as described above. We note that  $\sigma_k^2 = a_k^2 + \sigma_{k+1}^2$ ,  $k = 0, \dots, n-1$ , from which we derive

$$\sigma_k^2 + \sum_{j=0}^{k-1} a_j^2 = 1, \quad k = 0, \dots, n. \quad (4.6)$$

We define  $f_k$  as the  $k$ -th row of  $A$  and  $\mathcal{F} = \{f_0, f_1, f_2, \dots\}$ . Thus,  $A$  satisfies properties **P1** and **P2** and, according to Remark 2.1, it follows that  $A$  is a matrix that represents a possible scenario of the weak greedy algorithm applied to  $\mathcal{F}$ . In fact, since  $a_{j,j} = \sigma_j$ , this scenario corresponds to the particular case of the exact greedy algorithm. Note also that by (4.6) each  $f_k$  has norm one.

We want now to show that with a suitable choice of the  $t_k$ , the Kolmogorov width  $d_n(\mathcal{F})$  will be quite small. For this, it will be enough to construct a unit length vector

$y = \sum_{j=0}^n y(j)e_j$ , whose inner product with each of the  $f_k$ ,  $k = 0, \dots, n$ , is small. We define  $\varepsilon := \frac{1}{\sqrt{\sum_{j=0}^n t_j}}$  and  $y(j) := \varepsilon\sqrt{t_j}$  which gives that  $\|y\| = 1$ . We first note that  $\langle f_0, y \rangle = \sigma_0\varepsilon\sqrt{t_0} = \varepsilon$ . We can compute the other inner products of  $y$  with the  $f_k$  by using the following equality

$$\begin{aligned}
\langle f_{k+1}, y \rangle &= \sigma_{k+1}\varepsilon\sqrt{t_{k+1}} + \sum_{j=0}^k a_j\varepsilon\sqrt{t_j} \\
&= \sigma_{k+1}\varepsilon\sqrt{t_{k+1}} + a_k\varepsilon\sqrt{t_k} + \sum_{j=0}^{k-1} a_j\varepsilon\sqrt{t_j} \\
&= \sigma_k\varepsilon \left\{ \sqrt{t_{k+1}} \frac{2\sqrt{t_k t_{k+1}}}{t_k + t_{k+1}} + \sqrt{t_k} \frac{t_k - t_{k+1}}{t_k + t_{k+1}} \right\} + \sum_{j=0}^{k-1} a_j\varepsilon\sqrt{t_j} \\
&= \sigma_k\varepsilon \frac{2t_{k+1}\sqrt{t_k} + t_k\sqrt{t_k} - t_{k+1}\sqrt{t_k}}{t_k + t_{k+1}} + \sum_{j=0}^{k-1} a_j\varepsilon\sqrt{t_j} \\
&= \sigma_k\varepsilon\sqrt{t_k} + \sum_{j=0}^{k-1} a_j\varepsilon\sqrt{t_j} \\
&= \langle f_k, y \rangle, \quad k = 0, \dots, n-1.
\end{aligned}$$

Thus  $\langle f_k, y \rangle = \varepsilon$  for  $k = 0, \dots, n$ . If we take  $Y$  as the linear space contained in  $\text{span}\{e_0, \dots, e_n\}$  which is orthogonal to  $y$ , then  $Y$  has dimension  $n$  and therefore

$$d_n(\mathcal{F}) \leq \varepsilon.$$

As a particular choice, we now choose  $t_k := r^k$ ,  $k = 0, \dots, n$ , where  $r$  is large. Then, one checks that  $\sigma_n = 2^n r^{-n/2} (1 + 1/r)^{-n}$ . In addition,  $\frac{1}{\varepsilon} = \sqrt{\sum_{j=0}^n r^j} = \sqrt{\frac{r^{n+1}-1}{r-1}} \geq r^{n/2}$ . Therefore,

$$\frac{\sigma_n(\mathcal{F})}{d_n(\mathcal{F})} \geq 2^n (1 + 1/r)^{-n}.$$

If we let  $r \rightarrow \infty$  with  $n$  fixed, we see that in general we cannot have an estimate better than that given in (4.4) save for the factor  $\sqrt{4/3}$ .

## References

- [1] A. Buffa, Y. Maday, A.T. Patera, C. Prud'homme, and G. Turinici, *A Priori convergence of the greedy algorithm for the parameterized reduced basis*, preprint.
- [2] G.G. Lorentz, M. von Golitschek, and Y. Makovoz, *Constructive Approximation: Advanced Problems*, Springer Verlag, New York, 1996.

- [3] Y. Maday, A.T. Patera, and G. Turinici, *A priori convergence theory for reduced-basis approximations of single-parametric elliptic partial differential equations*, J. Sci. Comput., **17**(2002), 437–446.
- [4] Y. Maday, A. T. Patera, and G. Turinici, *Global a priori convergence theory for reduced-basis approximations of single-parameter symmetric coercive elliptic partial differential equations*, C. R. Acad. Sci., Paris, Ser. I, Math., **335**(2002), 289–294.
- [5] G. Rozza, D.B.P. Huynh, and A.T. Patera, *Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations application to transport and continuum mechanics*, Arch. Comput Method E, **15**(2008), 229–275.
- [6] S. Sen, *Reduced-basis approximation and a posteriori error estimation for many-parameter heat conduction problems*, Numer. Heat Tr. B-Fund, **54**(2008), 369–389.
- [7] V. Temlyakov, *Greedy Approximation*, Acta Numerica **17**(2008), 235–409.
- [8] K. Veroy, C. Prudhomme, D. V. Rovas, and A. T. Patera, *A Posteriori error bounds for reduced-basis approximation of parametrized noncoercive and nonlinear elliptic partial differential equations*, in: Proceedings of the 16th AIAA Computational Fluid Dynamics Conference, 2003, Paper 2003-3847.

Peter Binev, Department of Mathematics, University of South Carolina, Columbia, SC, 29208, binev@math.sc.edu

Albert Cohen, UPMC Univ Paris 06, UMR 7598, Laboratoire Jacques-Louis Lions, F-75005, Paris, France, cohen@ann.jussieu.fr

Wolfgang Dahmen, Institut für Geometrie und Praktische Mathematik, RWTH Aachen, Templergraben 55, D-52056 Aachen Germany, dahmen@igpm.rwth-aachen.de

Ronald DeVore, Department of Mathematics, Texas A&M University, College Station, TX, rdevore@math.tamu.edu

Guergana Petrova, Department of Mathematics, Texas A&M University, College Station, TX, gpetrova@math.tamu.edu

Przemyslaw Wojtaszczyk, Institute of Applied Mathematics, University of Warsaw, Warsaw, Poland, wojtaszczyk@mimuw.edu.pl

# Preprint Series DFG-SPP 1324

<http://www.dfg-spp1324.de>

## Reports

- [1] R. Ramlau, G. Teschke, and M. Zhariy. A Compressive Landweber Iteration for Solving Ill-Posed Inverse Problems. Preprint 1, DFG-SPP 1324, September 2008.
- [2] G. Plonka. The Easy Path Wavelet Transform: A New Adaptive Wavelet Transform for Sparse Representation of Two-dimensional Data. Preprint 2, DFG-SPP 1324, September 2008.
- [3] E. Novak and H. Woźniakowski. Optimal Order of Convergence and (In-) Tractability of Multivariate Approximation of Smooth Functions. Preprint 3, DFG-SPP 1324, October 2008.
- [4] M. Espig, L. Grasedyck, and W. Hackbusch. Black Box Low Tensor Rank Approximation Using Fibre-Crosses. Preprint 4, DFG-SPP 1324, October 2008.
- [5] T. Bonesky, S. Dahlke, P. Maass, and T. Raasch. Adaptive Wavelet Methods and Sparsity Reconstruction for Inverse Heat Conduction Problems. Preprint 5, DFG-SPP 1324, January 2009.
- [6] E. Novak and H. Woźniakowski. Approximation of Infinitely Differentiable Multivariate Functions Is Intractable. Preprint 6, DFG-SPP 1324, January 2009.
- [7] J. Ma and G. Plonka. A Review of Curvelets and Recent Applications. Preprint 7, DFG-SPP 1324, February 2009.
- [8] L. Denis, D. A. Lorenz, and D. Trede. Greedy Solution of Ill-Posed Problems: Error Bounds and Exact Inversion. Preprint 8, DFG-SPP 1324, April 2009.
- [9] U. Friedrich. A Two Parameter Generalization of Lions' Nonoverlapping Domain Decomposition Method for Linear Elliptic PDEs. Preprint 9, DFG-SPP 1324, April 2009.
- [10] K. Bredies and D. A. Lorenz. Minimization of Non-smooth, Non-convex Functionals by Iterative Thresholding. Preprint 10, DFG-SPP 1324, April 2009.
- [11] K. Bredies and D. A. Lorenz. Regularization with Non-convex Separable Constraints. Preprint 11, DFG-SPP 1324, April 2009.

- [12] M. Döhler, S. Kunis, and D. Potts. Nonequispaced Hyperbolic Cross Fast Fourier Transform. Preprint 12, DFG-SPP 1324, April 2009.
- [13] C. Bender. Dual Pricing of Multi-Exercise Options under Volume Constraints. Preprint 13, DFG-SPP 1324, April 2009.
- [14] T. Müller-Gronbach and K. Ritter. Variable Subspace Sampling and Multi-level Algorithms. Preprint 14, DFG-SPP 1324, May 2009.
- [15] G. Plonka, S. Tenorth, and A. Iske. Optimally Sparse Image Representation by the Easy Path Wavelet Transform. Preprint 15, DFG-SPP 1324, May 2009.
- [16] S. Dahlke, E. Novak, and W. Sickel. Optimal Approximation of Elliptic Problems by Linear and Nonlinear Mappings IV: Errors in  $L_2$  and Other Norms. Preprint 16, DFG-SPP 1324, June 2009.
- [17] B. Jin, T. Khan, P. Maass, and M. Pidcock. Function Spaces and Optimal Currents in Impedance Tomography. Preprint 17, DFG-SPP 1324, June 2009.
- [18] G. Plonka and J. Ma. Curvelet-Wavelet Regularized Split Bregman Iteration for Compressed Sensing. Preprint 18, DFG-SPP 1324, June 2009.
- [19] G. Teschke and C. Borries. Accelerated Projected Steepest Descent Method for Nonlinear Inverse Problems with Sparsity Constraints. Preprint 19, DFG-SPP 1324, July 2009.
- [20] L. Grasedyck. Hierarchical Singular Value Decomposition of Tensors. Preprint 20, DFG-SPP 1324, July 2009.
- [21] D. Rudolf. Error Bounds for Computing the Expectation by Markov Chain Monte Carlo. Preprint 21, DFG-SPP 1324, July 2009.
- [22] M. Hansen and W. Sickel. Best m-term Approximation and Lizorkin-Triebel Spaces. Preprint 22, DFG-SPP 1324, August 2009.
- [23] F.J. Hickernell, T. Müller-Gronbach, B. Niu, and K. Ritter. Multi-level Monte Carlo Algorithms for Infinite-dimensional Integration on  $\mathbb{R}^N$ . Preprint 23, DFG-SPP 1324, August 2009.
- [24] S. Dereich and F. Heidenreich. A Multilevel Monte Carlo Algorithm for Lévy Driven Stochastic Differential Equations. Preprint 24, DFG-SPP 1324, August 2009.
- [25] S. Dahlke, M. Fornasier, and T. Raasch. Multilevel Preconditioning for Adaptive Sparse Optimization. Preprint 25, DFG-SPP 1324, August 2009.

- [26] S. Dereich. Multilevel Monte Carlo Algorithms for Lévy-driven SDEs with Gaussian Correction. Preprint 26, DFG-SPP 1324, August 2009.
- [27] G. Plonka, S. Tenorth, and D. Roşca. A New Hybrid Method for Image Approximation using the Easy Path Wavelet Transform. Preprint 27, DFG-SPP 1324, October 2009.
- [28] O. Koch and C. Lubich. Dynamical Low-rank Approximation of Tensors. Preprint 28, DFG-SPP 1324, November 2009.
- [29] E. Faou, V. Gradinaru, and C. Lubich. Computing Semi-classical Quantum Dynamics with Hagedorn Wavepackets. Preprint 29, DFG-SPP 1324, November 2009.
- [30] D. Conte and C. Lubich. An Error Analysis of the Multi-configuration Time-dependent Hartree Method of Quantum Dynamics. Preprint 30, DFG-SPP 1324, November 2009.
- [31] C. E. Powell and E. Ullmann. Preconditioning Stochastic Galerkin Saddle Point Problems. Preprint 31, DFG-SPP 1324, November 2009.
- [32] O. G. Ernst and E. Ullmann. Stochastic Galerkin Matrices. Preprint 32, DFG-SPP 1324, November 2009.
- [33] F. Lindner and R. L. Schilling. Weak Order for the Discretization of the Stochastic Heat Equation Driven by Impulsive Noise. Preprint 33, DFG-SPP 1324, November 2009.
- [34] L. Kämmerer and S. Kunis. On the Stability of the Hyperbolic Cross Discrete Fourier Transform. Preprint 34, DFG-SPP 1324, December 2009.
- [35] P. Cerejeiras, M. Ferreira, U. Kähler, and G. Teschke. Inversion of the noisy Radon transform on  $SO(3)$  by Gabor frames and sparse recovery principles. Preprint 35, DFG-SPP 1324, January 2010.
- [36] T. Jahnke and T. Udrescu. Solving Chemical Master Equations by Adaptive Wavelet Compression. Preprint 36, DFG-SPP 1324, January 2010.
- [37] P. Kittipoom, G. Kutyniok, and W.-Q Lim. Irregular Shearlet Frames: Geometry and Approximation Properties. Preprint 37, DFG-SPP 1324, February 2010.
- [38] G. Kutyniok and W.-Q Lim. Compactly Supported Shearlets are Optimally Sparse. Preprint 38, DFG-SPP 1324, February 2010.
- [39] M. Hansen and W. Sickel. Best  $m$ -Term Approximation and Tensor Products of Sobolev and Besov Spaces – the Case of Non-compact Embeddings. Preprint 39, DFG-SPP 1324, March 2010.

- [40] B. Niu, F.J. Hickernell, T. Müller-Gronbach, and K. Ritter. Deterministic Multi-level Algorithms for Infinite-dimensional Integration on  $\mathbb{R}^N$ . Preprint 40, DFG-SPP 1324, March 2010.
- [41] P. Kittipoom, G. Kutyniok, and W.-Q Lim. Construction of Compactly Supported Shearlet Frames. Preprint 41, DFG-SPP 1324, March 2010.
- [42] C. Bender and J. Steiner. Error Criteria for Numerical Solutions of Backward SDEs. Preprint 42, DFG-SPP 1324, April 2010.
- [43] L. Grasedyck. Polynomial Approximation in Hierarchical Tucker Format by Vector-Tensorization. Preprint 43, DFG-SPP 1324, April 2010.
- [44] M. Hansen und W. Sickel. Best  $m$ -Term Approximation and Sobolev-Besov Spaces of Dominating Mixed Smoothness - the Case of Compact Embeddings. Preprint 44, DFG-SPP 1324, April 2010.
- [45] P. Binev, W. Dahmen, and P. Lamby. Fast High-Dimensional Approximation with Sparse Occupancy Trees. Preprint 45, DFG-SPP 1324, May 2010.
- [46] J. Ballani and L. Grasedyck. A Projection Method to Solve Linear Systems in Tensor Format. Preprint 46, DFG-SPP 1324, May 2010.
- [47] P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova, and P. Wojtaszczyk. Convergence Rates for Greedy Algorithms in Reduced Basis Methods. Preprint 47, DFG-SPP 1324, May 2010.