A Greedy Method for Solving Classes of PDE Problems

Robert Schaback[1]

Abstract: Motivated by the successful use of greedy algorithms for Reduced Basis Methods, a greedy method is proposed that selects N input data in an asymptotically optimal way to solve well-posed operator equations using these N data. The operator equations are defined as infinitely many equations given via a compact set of functionals in the dual of an underlying Hilbert space, and then the greedy algorithm, defined directly in the dual Hilbert space, selects N functionals step by step. When N functionals are selected, the operator equation is numerically solved by projection onto the span of the Riesz representers of the functionals. Orthonormalizing these yields useful Reduced Basis functions. By recent results on greedy methods in Hilbert spaces, the convergence rate is asymptotically given by Kolmogoroff N-widths and therefore optimal in that sense. However, these N-widths seem to be unknown in PDE applications. Numerical experiments show that for solving elliptic second-order Dirichlet problems, the greedy method of this paper behaves like the known P-greedy method for interpolation, applied to second derivatives. Since the latter technique is known to realize Kolmogoroff N-widths for interpolation, it is hypothesized that the Kolmogoroff N-widths for solving second-order PDEs behave like the Kolmogoroff N-widths for second derivatives, but this is an open theoretical problem.

Keywords: Operator equations, Greedy methods, Reduced Basis Methods, Kolmogoroff N-widths, Partial Differential Equations, meshless methods, collocation, discretization, error bounds, well-posedness, stability, convergence

Mathematics Subject Classification (2000): 65M12, 65M70, 65N12, 65N35, 65M15, 65J10, 41A25, 41A63

[1]Institut für Numerische und Angewandte Mathematik
Georg-August-Universität Göttingen, Lotzestraße 16–18, D-37083 Göttingen
schaback@math.uni-goettingen.de, http://num.math.uni-goettingen.de/schaback/

1 Introduction

For illustration of the application scope of this paper, consider the class of all second-order elliptic boundary value problems

$$
\begin{aligned}
Lu &= f \text{ on } \overline{\Omega} \subset \mathbb{R}^d \\
u &= g \text{ on } \Gamma := \partial\Omega
\end{aligned}
\tag{1}
$$

with arbitrary Dirichlet data and a fixed second-order strongly elliptic operator L on a fixed bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$. We keep this problem class in strong form and pose it in Sobolev space $W_2^m(\mathbb{R}^d)$ with $m > 2 + d/2$ and spaces of data functions f and g of corresponding smoothness. Similar to Reduced Basis (e.g. [14, 1, 3, 6]) or Proper Orthogonal Decomposition methods (e.g. [15, 13, 10]), we focus on a *class* of PDE problems, not on single problems. The output of this paper will be connected to both areas, since a "reduced" orthonormal basis is produced that is adapted to the given class of PDE problems.

The next section will generalize such problems to operator equations defined by sets Λ of infinitely many functionals on an underlying Hilbert space of functions, e.g. a Sobolev space. In this context, well-posedness can be formulated, and for N selected functionals in a set $\Lambda_N \subset \Lambda$, numerical solutions can be obtained by Hilbert space projection on the Riesz representers of these functionals. This is the well-known Rayleigh-Ritz idea. For kernel-based spaces, it coincides with Symmetric Collocation and yields the optimal recovery technique in Hilbert space for the given data functionals.

Section 3 analyzes the error in terms of the *Generalized Power Function*

$$
P_{\Lambda_N}(\lambda) := \text{dist}(\lambda, \text{ span } (\Lambda_N))
$$

on the dual of the Hilbert space and introduces

$$
\sigma_{\Lambda_N} := \max_{\lambda \in \Lambda} P_{\Lambda_N}(\lambda)
$$

that controls the error of the numerical solution in Hilbert space.

The greedy method of Section 4 now selects

$$
\lambda_{N+1} := \arg \max P_{\Lambda_N}(\lambda)
$$

and follows the results of the literature on Reduced Basis Methods, linking the decay of σ_{Λ_N} to Kolmogoroff N-widths.

For operator equations defined on Hilbert spaces of functions on a bounded domain Ω, the additional quantity

$$\rho_{\Lambda_N} := \max_{x \in \overline{\Omega}} P_{\Lambda_N}(\delta_x)$$

directly controls the pointwise and uniform error in the domain, but is not useful for greedy methods. For well-posed problems, Section 3.2 will show that ρ_{Λ_N} is bounded form above by σ_{Λ_N} up to a factor, and numerical results in the final Section 6 suggest that this bound is asymptotically sharp.

Before that, Section 5 gives a partial analysis of expectable Kolmogoroff N-widths for second-order Dirichlet problems in Sobolev spaces W_2^m. It is hypothesized that both σ_{Λ_N} and the N-width behave like $\mathscr{O}(N^{-\frac{m-2-d/2}{d}})$ for $N \to \infty$, which is the Kolmogoroff N-width for interpolation problems in W_2^{m-2} with respect to the supremum norm. Section 6 shows supporting examples and concludes the paper.

2 Hilbert Space Theory

Following [11, 21] the problem class is written in terms of infinitely many constraints, each defined by a linear functional. The functional sets then are

$$\begin{aligned}
\Lambda_1 &:= \{\delta_x \circ L \,:\, x \in \overline{\Omega}\} \\
\Lambda_2 &:= \{\delta_x \,:\, x \in \Gamma := \partial\Omega\}
\end{aligned} \tag{2}$$

combined into $\Lambda := \Lambda_1 \cup \Lambda_2$. Since all single functionals are continuous on $W_2^m(\mathbb{R}^d)$, the above sets are images of compact sets by continuous maps, thus compact. This brings us into line with the literature on reduced basis methods.

The above problems are *well-posed* in the sense that there is a standard *well-posedness inequality* of the form [2, 1.5, p. 30]

$$\|u\|_{\infty,\overline{\Omega}} \le \|u\|_{\infty,\partial\Omega} + C\|Lu\|_{\infty,\overline{\Omega}} \le (C+1)\sup_{\lambda \in \Lambda} |\lambda(u)| \text{ for all } u \in C^2(\overline{\Omega}) \cup C(\Gamma). \tag{3}$$

2.1   Abstract Problem

Generalizing this case, and following [11, 21], we assume a Hilbert space $\mathscr{H}$ and a subset Λ of its dual $\mathscr{H}^*$ that is *total* in the sense that

$$\lambda(u) = 0 \text{ for all } \lambda \in \Lambda \text{ implies } u = 0.$$

If we formally introduce the linear *data map* $D_\Lambda : \mathscr{H} \to \mathbb{R}^\Lambda$ with

$$D_\Lambda(u) := \{\lambda(u)\}_{\lambda \in \Lambda} \text{ for all } u \in \mathscr{H},$$

this means that elements $u \in \mathscr{H}$ are uniquely identifiable from their *data* $D_\Lambda(u)$. The central background problem in this paper is to recover elements u from their data $D_\Lambda(u)$ in practice, i.e. the approximate numerical inversion of the data map. In view of the preceding example we assume that the set Λ is compact.

The invertibility of the data map is quantified by assuming a *well-posedness inequality*

$$\|u\|_{WP} \le C_{WP} \|D_\Lambda(u)\|_\infty = C_{WP} \sup_{\lambda \in \Lambda} |\lambda(u)| \text{ for all } u \in \mathscr{H} \qquad (4)$$

in some *well-posedness* norm $\|.\|_{WP}$ on $\mathscr{H}$ that usually is weaker than the norm on $\mathscr{H}$. Recall that [21] allows also to handle weakly formulated problems as well this way. Furthermore, the framework applies to general operator equations, including the case of interpolation if the operator is the identity.

3   Error Analysis of Projection Methods

For a finite subset $\Lambda_N := \{\lambda_1, \ldots, \lambda_N\}$ of Λ we can define the subspace

$$L_N := \text{ span } \{\Lambda_N\} \subseteq \mathscr{H}^*$$

and use the Riesz representers $v_{\lambda_1}, \ldots, v_{\lambda_N}$ of $\lambda_1, \ldots, \lambda_N$ as trial functions. They span a space V_{Λ_N} that will lead later to reduced bases.

The standard optimal recovery of an element $u \in \mathscr{H}$ from finite data $\lambda_1(u), \ldots, \lambda_N(u)$ then proceeds by Hilbert space projection, i.e. by solving the linear system

$$\lambda_k(u) = \sum_{j=1}^N \alpha_j \lambda_k(v_{\lambda_j}) = \sum_{j=1}^N \alpha_j (v_{\lambda_k}, v_{\lambda_j})_{\mathscr{H}} = \sum_{j=1}^N \alpha_j (\lambda_k, \lambda_j)_{\mathscr{H}^*}, \ 1 \le k \le N$$

to get the numerical approximation

$$\tilde{u}_N := \sum_{j=1}^{N} \alpha_j v_{\lambda_j}.$$

This satisfies the orthogonality relation

$$\|u\|_{\mathcal{H}}^2 = \|u - \tilde{u}_N\|_{\mathcal{H}}^2 + \|\tilde{u}_N\|_{\mathcal{H}}^2 \tag{5}$$

that implies uniform stability in Hilbert space. In case of the example in the beginning, the method is known as *Symmetric Collocation*. This is a numerical technique [7] based on [23] with certain optimality properties [20] and a convergence theory [9, 8]. Within the Hilbert space framework, it produces pointwise optimal approximations to the true solution under all possible methods that use the same data [20].

3.1 Power Function

The standard error analysis in Hilbert Spaces uses the *generalized Power Function* defined as

$$P_{\Lambda_N}(\mu) := \min_{\lambda \in L_N} \|\mu - \lambda\|_{\mathcal{H}^*} =: \operatorname{dist}(\mu, L_N)_{\mathcal{H}^*} \text{ for all } \mu \in \mathcal{H}^*. \tag{6}$$

This is continuous and attains its maximum on the compact set Λ. In particular, we are interested in

$$\sigma_{\Lambda_N} := \sup_{\lambda \in \Lambda} P_{\Lambda_N}(\lambda) = \sup_{\lambda \in \Lambda} \operatorname{dist}(\lambda, L_N)_{\mathcal{H}^*}. \tag{7}$$

3.2 Error Analysis

This quantity leads to error bounds for the discretized recovery problem that we described in the beginning of this section.

Lemma 1. *Let $u \in \mathcal{H}$ supply the finite data $\lambda_1(u), \ldots, \lambda_N(u)$ that is used to construct $\tilde{u}_N$ by projection, and assume well-posedness in the sense of (4). Then*

$$\|u - \tilde{u}_N\|_{WP} \leq C_{WP} \sigma_{\Lambda_N} \|u\|_{\mathcal{H}}.$$

Proof. Since u and $\tilde{u}_N$ share the same data, the assertion follows from

$$
\begin{aligned}
\|u - \tilde{u}_N\|_{WP} &\leq C_{WP}\sup_{\lambda \in \Lambda}|\lambda(u) - \lambda(\tilde{u}_N)| \\
&= C_{WP}\sup_{\lambda \in \Lambda}|(\lambda - \mu)(u - \tilde{u}_N)| \text{ for all } \mu \in L_N
\end{aligned}
\tag{8}
$$

using orthogonality (5) and well-posedness (4). $\qquad\qquad\qquad\square$

Therefore we are interested to find sets Λ_N that minimize σ_{Λ_N} under all sets of N functionals from Λ.

By the same argument, for all continuous test functionals μ the inequality

$$
|\mu(u) - \mu(\tilde{u}_N)| \leq P_{\Lambda_N}(\mu)\|u\|_{\mathscr{H}},
\tag{9}
$$

holds and we can check derivative errors once we evaluate the Generalized Power Function on derivative functionals. Taking all functionals of Λ here, we get

$$
\|D_\Lambda(u - \tilde{u}_N)\|_\infty = \sup_{\lambda \in \Lambda}|\lambda(u - \tilde{u}_N)| \leq \sigma_{\Lambda_N}\|u\|_{\mathscr{H}},
$$

i.e. σ_{Λ_N} bounds the error in the data norm.

In case of a Hilbert space $\mathscr{H}$ containing continuous functions on some compact set $\overline{\Omega}$, there is a special case of (9) in the pointwise form

$$
|u(x) - \tilde{u}_N(x)| \leq P_{\Lambda_N}(\delta_x)\|u\|_{\mathscr{H}}
\tag{10}
$$

that may be checked by evaluation of the Generalized Power Function over all delta functionals for all points of the domain. If we define

$$
\rho_{\Lambda_N} := \max_{x \in \overline{\Omega}} P_{\Lambda_N}(\delta_x)
\tag{11}
$$

we get a uniform bound

$$
\|u - \tilde{u}_N\|_{\infty,\overline{\Omega}} \leq \rho_{\Lambda_N}\|u\|_{\mathscr{H}}
\tag{12}
$$

that is numerically available, and similar to the well-posedness inequality (3).

By definition, the Generalized Power Function decreases at all functionals when we extend the set Λ_N. All of these error bounds will then improve. For fixed Λ_N and $\mathscr{H}$, they are optimal, but here we are interested in finding good finite subsets Λ_N of Λ.

Note that both Lemma 1 and (12) furnish similar error bounds with associated convergence rates, and they might have the same behavior if the well-posedness norm $\|.\|_{WP}$ coincides with $\|.\|_{L_\infty(\overline{\Omega})}$. We shall have a closer look at this now, delaying experiments to Section 6.

To find a connection between ρ_{Λ_N} and σ_{Λ_N} in well-posed situations, we assume that the well-posedness norm $\|.\|_{WP}$ of (4) can be written as

$$\|u\|_{WP} = \sup_{\mu \in M} |\mu(u)|, \ u \in \mathscr{H} \tag{13}$$

with a compact set $M \subset \mathscr{H}^*$, i.e. M is *norming* for $\|.\|_{WP}$. This yields (3) when taking M as the set of delta functionals on $\overline{\Omega}$. The generalization of (11) then is

$$\rho_{M,\Lambda_N} := \max_{\mu \in M} P_{\Lambda_N}(\mu). \tag{14}$$

Lemma 2. *Assuming (4) and (13), we have*

$$\rho_{M,\Lambda_N} \leq C_{WP}\sigma_{\Lambda_N}.$$

Proof. The maximum in (14) is attained at some $\tilde{\mu}_N \in M$ and then

$$\begin{aligned}
\rho_{M,\Lambda_N} &= P_{\Lambda_N}(\tilde{\mu}_N) \\
&= \text{dist}(\tilde{\mu}_N, \text{ span } \Lambda_N) \\
&= \left\| \tilde{\mu}_N - \sum_{j=1}^{N} (\tilde{\mu}_N, \mu_j)\mu_j \right\|_{\mathscr{H}^*}.
\end{aligned}$$

Applied to an arbitrary $u \in \mathscr{H}$ this yields

$$\begin{aligned}
\left| \tilde{\mu}_N(u) - \sum_{j=1}^{N}(\tilde{\mu}_N, \mu_j)\mu_j(u) \right| &= |\tilde{\mu}_N(u) - \tilde{\mu}_N(\tilde{u}_N)| \\
&\leq \sup_{\mu \in M} |\mu(u) - \mu(\tilde{u}_N)| \\
&= \|u - \tilde{u}_N\|_{WP} \\
&\leq C_{WP} \sup_{\lambda \in \Lambda} |\lambda(u) - \lambda(\tilde{u}_N)| \\
&\leq C_{WP}\sigma_{\Lambda_N}\|u\|_{\mathscr{H}}
\end{aligned}$$

where the last line follows like (9) using a generalization of the argument within (10). $\square$

4 Greedy Method

Given a set $\Lambda_N := \{\lambda_1, \ldots, \lambda_N\}$, the *P-greedy algorithm* [5, 18] in its abstract form defines λ_{N+1} recursively via

$$P_{\Lambda_N}(\lambda_{N+1}) = \sup_{\lambda \in \Lambda} P_{\Lambda_N}(\lambda). \tag{15}$$

In the context of Reduced Basis Methods in Hilbert spaces, this algorithm coincides with the one in [14, 1, 3, 6]. The main difference is that we work in the dual here, focusing on a single class of PDEs in applications instead of a parametrized family. A first application of Reduced Basis Methods within a Reproducing Kernel Hilbert Space setting is in [4], implementing a greedy method based on discrete least-squares, not on the dual of the basic Hilbert space of functions.

The cited papers provide a useful error analysis of the method that we state here for completeness. The main ingredient is the *Kolmogorov N-width*

$$d_N(\Lambda) := \inf_{\text{all } H_N} \sup_{\mu \in \Lambda} \inf_{\lambda \in H_N} \|\lambda - \mu\|_{\mathscr{H}^*} = \inf_{\text{all } H_N} \sup_{\mu \in \Lambda} \text{dist}(\mu, H_N)_{\mathscr{H}^*}$$

where the first infimum is taken over all N-dimensional subspaces $H_N \subseteq \mathscr{H}^*$, not only those spanned by N functionals from Λ. Then [6] proves

$$\sigma_{2N}^2(\Lambda) \leq 2d_N(\Lambda),$$

while [1] has

$$\sigma_{\Lambda_N} \leq C(\alpha)N^{-\alpha} \text{ for } n \in \mathbb{N}, \text{ if } d_N(\Lambda) \leq C'(\alpha)N^{-\alpha} \text{ for } n \in \mathbb{N}$$

with suitable constants. This links the behaviour of the *P*-greedy algorithm to Kolmogoroff *N*-widths. The paper [18] exploits this connection for function recovery by interpolation, while this paper extends [18] to classes of operator equations, including PDE solving.

Consequently, the greedy method converges roughly like the Kolmogoroff *N*-widths. Asymptotically, there are no better choices for selecting *N* functionals out of Λ. Note that this works for all well-posed problems stated in abstract form in Hilbert space via infinitely many constraints. The differential and boundary operators can be easily generalized.

On the downside, the literature does not provide much information about the Kolmogoroff *N*-widths in such situations. This is why on has to look at special cases. We postpone this to Section 5.

If we run the greedy algorithm numerically, we should get very good candidates for reduced bases. Choosing them with additional orthogonality properties will then bring us close to Proper Orthogonal Decomposition methods. The next section will explain how to do that.

4.1 Implementation

Throughout, we assume that the Hilbert space $\mathscr{H}$ has a kernel K such that inner products in $\mathscr{H}^*$ can be numerically calculated via

$$(\lambda,\mu)_{\mathscr{H}^*} = \lambda^x \mu^y K(x,y) \text{ for all } \lambda,\mu \in \mathscr{H}^*$$

where the upper index stands for the variable the functional acts on. In case of Sobolev spaces $W_2^m(\mathbb{R}^d)$ with $m > d/2$, we use the standard Whittle-Matérn kernel

$$K_{m,d}(x,y) = \|x-y\|_2^{m-d/2} K_{m-d/2}(\|x-y\|_2), \; x,y \in \mathbb{R}^d$$

with the modified Bessel function K_ν of second kind. In what follows, we treat Λ as being very large and finite, but extensions to infinite compact Λ will be possible if functions on Λ are discretized somehow.

A direct way to assess the generalized Power Function at λ for given $\Lambda_N := \{\lambda_1,\dots,\lambda_N\}$ is to use its definition via the approximation problem

$$\inf_{\alpha \in \mathbb{R}^N} \left\| \lambda - \sum_{j=1}^N \alpha_j \lambda_j \right\|_{\mathscr{H}^*}^2 = P_{\Lambda_N}^2(\lambda).$$

We assume that we have turned $\lambda_1,\dots,\lambda_N$ into an $\mathscr{H}^*$-orthonormal basis $\mu_1,\dots,\mu_N$ already, and then the solution is

$$P_{\Lambda_N}^2(\lambda) = \left\| \lambda - \sum_{j=1}^N (\lambda,\mu_j)_{\mathscr{H}^*} \mu_j \right\|_{\mathscr{H}^*}^2 = (\lambda,\lambda)_{\mathscr{H}^*} - \sum_{j=1}^N (\lambda,\mu_j)_{\mathscr{H}^*}^2. \qquad (16)$$

Furthermore, we store the orthonormalization system for getting the μ_j from the λ_j as

$$\mu_k = \sum_{j=1}^{k-1} c_{kj}\mu_j + c_{k,k}\lambda_k, \; 1 \le k \le N. \qquad (17)$$

Besides the $N \times N$ triangular matrix C, we store the values $\{(\lambda, \lambda)_{\mathscr{H}^*}\}_{\lambda \in \Lambda}$ and $\{(\lambda, \mu_k)_{\mathscr{H}^*}\}_{\lambda \in \Lambda}$ for $k = 1, \ldots, N$. These make up the method's bulk storage of order $(N+1)|\Lambda|$. Note that $\{(\lambda, \lambda)_{\mathscr{H}^*}\}_{\lambda \in \Lambda}$ simplifies considerably for translation-invariant kernels.

After maximizing (16) over all $\lambda \in \Lambda$, we assume to have some λ_{N+1} at which a nonzero maximum is attained. This can then not be one of the old λ_j, and we retrieve the values $(\lambda_{N+1}, \mu_k)_{\mathscr{H}^*}$ from what we have. We now orthonormalize

$$\mu_{N+1} = \sum_{j=1}^{N} c_{N+1,j} \mu_j + c_{N+1,N+1} \lambda_{N+1}$$

via $c_{N+1,j} = -(\lambda_{N+1}, \mu_k)_{\mathscr{H}^*} c_{N+1,N+1}$, $1 \le k \le N$ and

$$
\begin{aligned}
1 = (\mu_{N+1}, \mu_{N+1})_{\mathscr{H}^*} &= c_{N+1,N+1}^2 \left((\lambda_{N+1}, \lambda_{N+1})_{\mathscr{H}^*} - \sum_{k=1}^{N} (\lambda_{N+1}, \mu_k)_{\mathscr{H}^*}^2 \right) \\
&= c_{N+1,N+1}^2 P_{\Lambda_N}^2(\lambda_{N+1})
\end{aligned}
$$

to update the C matrix. Finally

$$(\lambda, \mu_{N+1})_{\mathscr{H}^*} = \sum_{j=1}^{N} c_{N+1,j} (\lambda, \mu_j)_{\mathscr{H}^*} + c_{N+1,N+1} (\lambda, \lambda_{N+1})_{\mathscr{H}^*}$$

can be calculated from what we have, if we first calculate all $\{(\lambda, \lambda_{N+1})_{\mathscr{H}^*}\}_{\lambda \in \Lambda}$ and overwrite them with $\{(\lambda, \mu_{N+1})_{\mathscr{H}^*}\}_{\lambda \in \Lambda}$ after use. This extends the *Newton basis* technique in [16] to general functionals.

4.2   Bases and Postprocessing

Useful reduced bases for PDE solving are the Riesz representers v_{μ_k} of the μ_k, being orthonormal in Hilbert space. From Riesz representers $v_{\lambda_j}(\cdot) = \lambda_j^x K(x, \cdot)$ of the λ_j, we can calculate them recursively via

$$v_{\mu_k}(x) = \sum_{j=1}^{k-1} c_{kj} v_{\mu_j}(x) + c_{k,k} v_{\lambda_k}(x), \ 1 \le k \le n \tag{18}$$

on whatever point sets we like, using our triangular matrix C. Via (16) in the form

$$P_{\Lambda_N}^2(\delta_x) = K(x,x) - \sum_{j=1}^{N} v_{\mu_j}^2(x), \tag{19}$$

this allows to calculate the pointwise error bounds (10) described in Section 3.2 explicitly, up to the term $\|u\|_{\mathscr{H}}$.

In view of Proper Orthogonal Decomposition, one can also apply a Singular Value Decomposition to the partial Gramian matrix with entries $(\lambda_j, \lambda_k)_{\mathscr{H}^*}$, $1 \leq j, k \leq N$ and construct a different $\mathscr{H}$-orthonormal basis. This possibility is not pursued here, because we want to keep the recursive structure of the algorithm. In view of [17], this basis may be closer to what happens for Kolmogoroff N-widths, because the spaces for the latter will not necessarily have a recursive structure.

Like in reduced basis techniques, the orthonormal bases are a simple tool to solve a variety of similar problems, namely all problems of the form (1). If data $\lambda_j(u)$ are known for an unknown function u, we go over to $\mu_j(u)$ via the triangular system (17) and then form the optimal projection

$$\tilde{u}_N := \sum_{j=1}^{N} \mu_j(u) v_{\mu_j} \tag{20}$$

using the $\mathscr{H}$-orthonormal basis we have constructed.

Note that this algorithm neither stores nor solves a large $|\Lambda| \times |\Lambda|$ system. It works "on–the–fly". For N steps, storage is of order $\mathscr{O}(N \cdot |\Lambda| + N^2)$ and calculations are of order $\mathscr{O}(N^2 \cdot |\Lambda|)$. For a given accuracy requirement, the number N of steps will depend on the Kolmogoroff N-width for the set Λ and the Hilbert space $\mathscr{H}^*$.

There are no square systems to be solved. Instead, there are N orthonormalization steps on vectors of length $|\Lambda|$ that may require standard stabilization precautions for the basic Gram-Schmidt technique. The resulting triangular $N \times N$ matrix C is not explicitly inverted, but used via (17) to transform input data in terms of the λ_j functionals into data in terms of the orthonormalized functionals μ_j. The error behavior of this is comparable to backsubstitution after an *LR* or *QR* factorization, but it will pay the price when some of the λ_j are strongly correlated. This will be unavoidable for large N, but (6) and (15) show that λ_{N+1} will be kept away from the zeros $\lambda_1, \ldots, \lambda_N$ of P_{Λ_N} and the space they span, by construction.

4.3 Extended Greedy Method

One can get somewhat closer to the error analysis in Section 3.2 and in particular to (9) by a modification of the selection strategy of functionals. Given a set

$\Lambda_N := \{\lambda_1, \ldots, \lambda_N\}$, and a set $M = \{\mu_1, \ldots, \mu_M\}$ with nonempty intersection, we calculate two maxima

$$
\begin{aligned}
\tilde{\lambda} &:= \operatorname{arg\,sup}_{\lambda \in \Lambda} P_{\Lambda_N}(\lambda) &=& \sigma_{\Lambda_N} \\
\tilde{\mu} &:= \operatorname{arg\,sup}_{\mu \in M} P_{\Lambda_N}(\mu) &=:& \rho_{\Lambda_N}
\end{aligned}
$$

and set

$$
\lambda_{N+1} := \begin{cases} \tilde{\mu} & \text{if } \mu \in \Lambda \cap M \\ \tilde{\lambda} & \text{else.} \end{cases}
$$

This can be called an *extended P-greedy algorithm*. It tries to keep some additional control of ρ_{M,Λ_N} from (14) by selecting functionals from $\Lambda \cap M$ whenever ρ_{M,Λ_N} is attained on them.

For solving Dirichlet problems, the set M will consist of delta functionals in $\overline{\Omega}$, the intersection of M and Λ being Λ_2, the set of functionals for Dirichlet boundary values. The new technique will make sure that if $P_{\Lambda_N}(M)$ attains its maximum on the boundary, the corresponding functional is preferred over the functional where $P_{\lambda_N}(\Lambda)$ attains its maximum.

Now for some inplementation details for the Dirichlet case. Define the set Z of boundary points via

$$
\{\delta_z \, : \, z \in Z\} = \Lambda_2
$$

and add some other point set $Y \subset \overline{\Omega}$ to get

$$
M := \{\delta_y \, : \, y \in Y\} \cup \Lambda_2 = \{\delta_x \, : \, x \in Y \cup Z\}.
$$

We need the additional value $\|P_{\Lambda_N}\|_{\infty,Y}$, while the usual Greedy method provides $\|P_{\Lambda_N}\|_{\infty,Z}$ as part of the calculation of ρ_{Λ_N}. If we use (19) on Y, we have the necessary data, but for that we have to evaluate (18) on Y as well, and have to store the values of the orthonormal basis on Y. This requires an additional storage of size $|Y| \cdot N$ for N steps, and additional calculations of order $|Y| \cdot N^2$. On can choose $|Y|$ smaller than $|\Lambda_1|$ to keep the complexity at bay.

5   Sobolev case

We now go back to the example at the beginning. We need information on the Kolmogoroff N-width

$$
d_N(\Lambda) = \inf_{\text{all } H_N} \sup_{\mu \in \Lambda} \operatorname{dist}(\mu, H_N)_{W_2^m(\Omega)^*}
$$

where the infimum is taken over all N-dimensional subspaces H_N of $\mathscr{H}^* = W_2^m(\Omega)^*$, and the set $\Lambda = \Lambda_1 \cup \Lambda_2 \subset \mathscr{H}^*$ is formed by (2). There should be an m- and d-dependent decay rate $\kappa(m,d)$ in the sense

$$d_N(\Lambda) \leq CN^{-\kappa(m,d)} \text{ for } N \to \infty$$

but no explicit results on this were found yet.

If we restrict attention to spaces H_N being generated by N functionals from $\Lambda = \Lambda_1 \cup \Lambda_2 \subset \mathscr{H}^*$, we only get an upper bound for $d_N(\Lambda)$, and we do not know the splitting $N = N_1 + N_2$ if an optimal choice of N functionals from Λ takes N_1 functionals out of Λ_1 and N_2 functionals out of Λ_2. The rest of the chapter will give some arguments supporting the hypothesis

$$\kappa(m,d) \geq \frac{m-2-d/2}{d} \tag{21}$$

that will be observed in the numerical behavior of the P-greedy method in Section 6.

If Λ is the union of two disjoint compact sets Λ_1 and Λ_2, then

$$
\begin{aligned}
d_N(\Lambda) &= \inf_{\text{all } H_N} \sup_{\mu \in \Lambda} \text{dist}(\mu, H_N)_{\mathscr{H}^*} \\
&= \inf_{\text{all } H_N} \max \left(\sup_{\mu \in \Lambda_1} \text{dist}(\mu, H_N)_{\mathscr{H}^*}, \sup_{\mu \in \Lambda_2} \text{dist}(\mu, H_N)_{\mathscr{H}^*} \right) \\
&\leq \inf_{H_{N_1}, H_{N_2}, N_1+N_2 \leq N, H_N = H_{N_1}+H_{N_2}} \max(d_{N_1}(\Lambda_1), d_{N_2}(\Lambda_2))
\end{aligned}
$$

where H_{N_i} is used for approximation of Λ_i, and the sum of spaces is direct. This *splitting argument* is similar to the technique in [8]. However, we do not know a priori how the greedy algorithm selects functionals from either set, and how the dimension splits into $N_1 + N_2 \leq N$.

We first aim at $d_N(\Lambda_1)$ in the space $W_2^m(\Omega)^*$. We have

$$d_N(\Lambda_1) = \inf_{\text{all } H_N} \sup_{x \in \Omega} \inf_{\lambda \in H_N} \|\lambda - \delta_x \circ L\|_{W_2^m(\Omega)^*}$$

and can majorise it by choosing N asymptotically uniformly placed points $x_1, \ldots x_N$ in $\overline{\Omega}$ at fill distance h_Ω and taking H_N to be the span of the corresponding functionals $\delta_{x_j} \circ L$. Then

$$d_N(\Lambda_1) \leq \sup_{x \in \Omega} \inf_{\alpha \in \mathbb{R}^N} \|\delta_x \circ L - \sum_{j=1}^N \alpha_j \delta_{x_j} \circ L\|_{W_2^m(\Omega)^*}$$

holds and we can expect

$$
\begin{aligned}
d_N(\Lambda_1) &\leq C \sup_{x \in \Omega} \inf_{\alpha \in \mathbb{R}^N} \| \delta_x - \sum_{j=1}^{N} \alpha_j \delta_{x_j} \|_{W_2^{m-2}(\Omega)^*} \\
&\leq C h_\Omega^{m-2-d/2} \leq C N^{-(m-2-d/2)/d}
\end{aligned}
$$

due to standard results on error bounds for interpolation [22], and with generic constants depending on m, d and the domain.

On the boundary Γ, we can argue similarly to expect

$$
d_N(\Lambda_2) \leq C h_\Gamma^{m-d/2} \leq C N^{-\frac{m-d/2}{d-1}}
$$

for a fill distance h_Γ on the boundary, either by working in $W_2^m(\Omega)^*$ directly, or via trace theorems, which would give the same rate due to $m - 1/2 - (d-1)/2 = m - d/2$.

If we now consider splittings $N = N_1 + N_2$, we roughly have the upper bound

$$
d_N(\Lambda) \leq C \max \left(N_1^{-(m-2-d/2)/d}, N_2^{-\frac{m-d/2}{d-1}} \right)
$$

and the crude split $N_1 \approx N/2 \approx N_2$ will already support (21).

For purposes of asymptotics, we can minimize the sum of the above quantities instead of the maximum. Using standard optimization arguments under the constraint $N_1 + N_2 \leq N$, the result after some calculations is that one should expect

$$
c_1 \frac{m-2-d/2}{d} N_1^{-\frac{m-2+d/2}{d}} = c_2 \frac{m-d/2}{d-1} N_2^{-\frac{m-1+d/2}{d-1}}
$$

with constants depending on the domain and the space dimension, but not on N and m. For the typical case $m = 4$, $d = 2$ this implies $N_2 \approx N_1^{3/8}$. For large m and d we get $N_1 \approx N_2$, i.e. the domain and the boundary require roughly the same degrees of freedom. This is no miracle, because the volumes of balls of high dimension are increasingly accounted for by the boundary layer. In general, the above argument implies $N_2 \leq c N_1$ for N large enough, and then (21) will hold. Section 6 will reveal that the greedy algorithm selects unexpectedly small values of N_2, and that the hypothesis (21) is supported.

6 Numerical Results

No matter which PDE examples are selected, the most interesting question is the behavior of the greedy method as a function of the N steps it takes. If Λ is chosen large enough but still finite, the crucial quantities are σ_{Λ_N} and ρ_{Λ_N} as defined in (7) and (11). By Section 4 we can expect that σ_{Λ_N} decays at an optimal rate comparable to the Kolmogoroff N-width with respect to Λ, but since the latter is still unknown, we can not yet assess how close we come to it. Since ρ_{Λ_N} is controlling the error in the sup norm, without being usable for greedy refinement, and in view of Lemma 2, we would like to confirm that ρ_{Λ_N} decays as fast as σ_{Λ_N}. And, the decay rates should improve with smoothness of the functions in the basic Hilbert space, i.e. with m if we work in $W_2^m(\mathbb{R}^d)$, our hypothesis being (21). Another interesting question is how the greedy method chooses between boundary functionals from Λ_2 and domain functionals from Λ_1, and whether the corresponding points are roughly uniformly distributed in both cases. Finally, the shape and the behaviour of the basis functions v_{μ_j} should be demonstrated. MATLAB code for all examples can be requested from the author.

6.1 Observations for varying N

We start on the 2D unit disk, with L being the Laplace operator, carrying the greedy method out for up to 500 steps, offering 17570 functionals for Λ_1 and 150 functionals for Λ_2. The Hilbert space will be $W_2^m(\mathbb{R}^2)$, but we fix $m > 2 + d/2$ first, to study the behaviour of the greedy method for varying the number N of steps. Classical results on kernel-based interpolation lets us expect rates for σ_{Λ_N} and ρ_{Λ_N} that are determined by fill distances. If h_Γ and h_Ω are fill distances for points on the boundary Γ and in the domain Ω, we can compare σ_{Λ_N} with plain interpolation of $\Delta u = f$ with smoothness $m - 2$ and behavior

$$h_\Omega^{m-2-d/2} = h_\Omega^{m-3} \approx N_\Omega^{-\frac{m-2-d/2}{d}} = N_\Omega^{-\frac{m-3}{2}} \tag{22}$$

on the domain. The error on the boundary in L_∞ should, if it were a plain interpolation, behave like

$$h_\Gamma^{m-1/2-(d-1)/2} = h_\Gamma^{m-d/2} = h_\Gamma^{m-1} \approx N_\Gamma^{-\frac{m-d/2}{d-1}} = N_\Gamma^{-(m-1)}, \tag{23}$$

if N_Γ points are asymptotically equally spaced on the boundary. But we do not know a priori how the greedy algorithm chooses between boundary and domain functionals.

The first experiments are for $m = 4 > 2 + d/2 = 3$, and we ignore the split $N = N_\Omega + N_\Gamma$ in the beginning. The scale of the Whittle-Matérn kernel is chosen to be 1, and then both types of functionals happen to have the same norm. Scaling changes the relation between function value evaluation and Laplace operator evaluation and must be used with care. We wanted to eliminate scaling effects for what follows, in order to let the choice between boundary and domain functionals be unbiased by possibly different norms.

Figure 1 shows σ_{Λ_N} (left) and ρ_{Λ_N} (right) as functions of N, with the observed rates -0.45 and -0.54, respectively. We see that ρ_{Λ_N} decays as fast as σ_{Λ_N} as functions of N.

Figure 1: σ_{Λ_N} and ρ_{Λ_N} and their rates as functions of N

The strange drops of ρ_{Λ_N} at certain N occur exactly when the fill density h_Γ on the boundary drops, namely after the greedy method has chosen another boundary functional. This can be read off the two plots of Figure 2 that show both σ_{Λ_N} and ρ_{Λ_N} (left) and both $h_\Gamma(N)$ and $h_\Omega(N)$ as functions of the N functionals that were greedily selected. Figure 4 shows the strange fact that the greedy method selects only rather few boundary functionals compared to domain functionals (6 versus 494). It allows two large peaks on the boundary, because it still fights for getting small on the domain functionals. These effects were observed in many other cases that we suppress here for brevity. Choosing different weights for domain and boundary functionals makes some sense in view of well-posedness inequalities like (3), and then the effect will be less apparent.

Figure 3 shows decays like

$$P_{\Lambda_N} \approx N_\Omega^{-0.53} \text{ and } P_{\Lambda_N} \approx N_\Gamma^{-1.42}$$
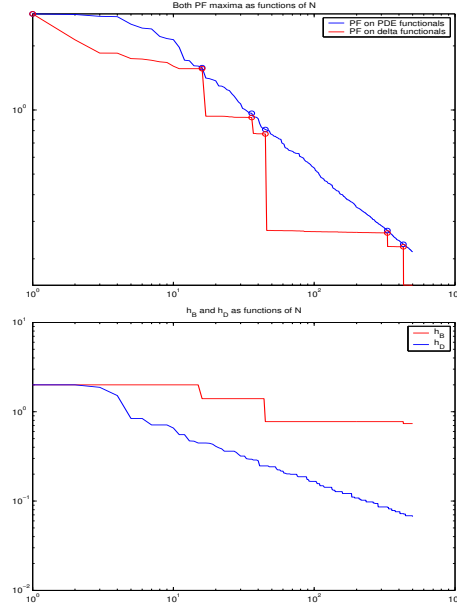
Figure 2: σ_{Λ_N} and ρ_{Λ_N} (top) compared with $h_\Gamma(N)$ and $h_\Omega(N)$ (bottom) as functions of N. The drops of ρ_{Λ_N} and $h_\Gamma(N)$ occur at the same N.

in comparison to the expectations in (23) and (22) that suggest rates -0.5 and -3, respectively, for $m = 4$. This shows again that the greedy method focuses on domain points, and is able to maintain a small error on the boundary by adding "domain" functionals there and close to the boundary.

If one looks at the Power Function on δ functionals all over the domain, like for the ρ_{Λ_N} calculation, the results are in Figure 5. It shows that for a better overall L_∞ bound it would be useful to pick boundary functionals at the boundary peaks of the left plot. This calls for a variation of the greedy method that monitors P_{Λ_N} on all δ functionals as well, and picks a boundary functional as soon as the maximum is on the boundary. This will work for Dirichlet problems, but not in general circumstances in Hilbert spaces. See Sections 4.3 for theory and 6.4 for numerical results, respectively.

The basis functions v_{μ_N} are orthonormal in Hilbert space but not in L_2, decay with N, and show a sharp bell-shape for large N. Figure 6 shows a case like Figure 5, but for $N = 150$. The new point, marked with an asterisk, is preferred over any

Figure 3: P_{Λ_N} with $N = N_\Omega + N_\Gamma$ as a function of N_Ω and N_Γ, respectively.

boundary point by the greedy method, though it is close to the boundary. The greedy method, as is, does not need many δ functionals on the boundary, if it has plenty of $\delta \circ L$ functionals on or near the boundary.

The norm and the condition estimate of the transformation matrix that takes the $\lambda_j(u)$ values into the $\mu_j(u)$ values, based on (17), are given in Figure 7, behaving roughly like $N^{0.69}$ and $N^{1.7}$, respectively. This is like $h_\Omega^{-1/3}$ and $h_\Omega^{-2/3}$, respectively, in terms of the fill distance h_Ω in the domain.

To get an idea of the decay of the basis functions v_{μ_j} in L_2 context, the matrix of the values in a fine point set on the domain is calculated and a singular value decomposition is done on that matrix. The result is shown in Figure 8, the decay behavior of singular values being roughly like $N^{-2.4}$.

Note that each basis function is a worst case for the preceding steps, because it has zero data for them and is approximated by the zero function. Thus L_∞ or L_2 norms of the basis functions are closely related to the worst-case L_∞ or L_2 norms of solutions with Hilbert space norm one. Figure 9 shows the RMSQ and L_∞ norms of the basis functions v_{μ_j} as functions of j, the estimated decay being like $N^{-1.65}$ and $N^{-1.44}$, respectively, but with serious roundoff pollution for large j. The peaks are exactly where boundary points are chosen by the greedy method,
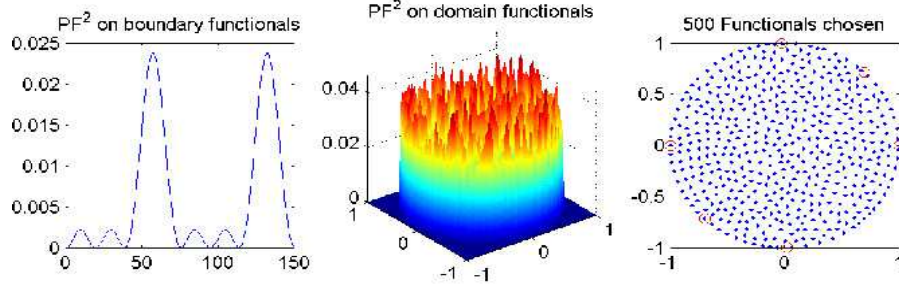
Figure 4: $P_{\Lambda_{500}}^2$ on the boundary functionals, the domain functionals, and the selected 500 functionals
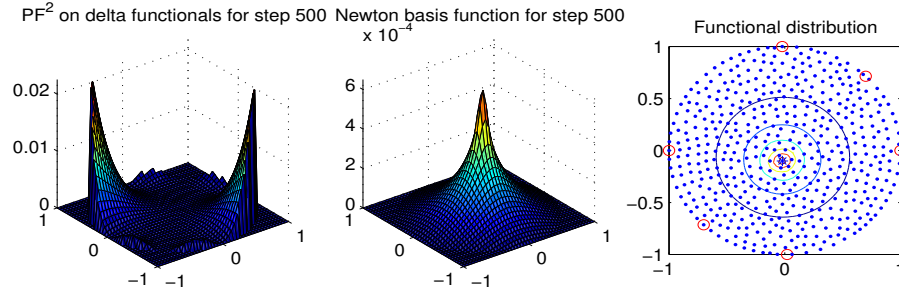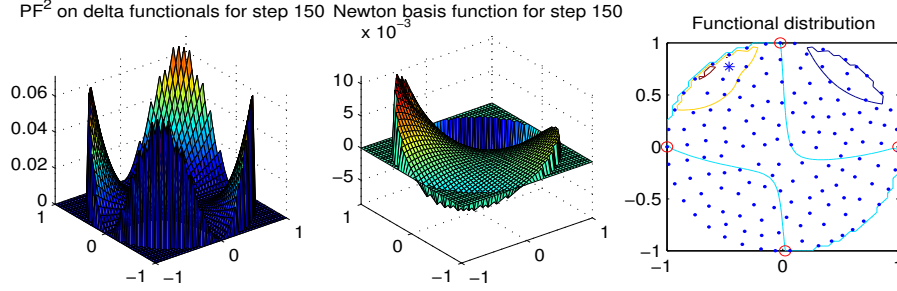

Figure 5: $P_{\Lambda_{500}}^2(\delta_x)$ on the domain, the basis function $v_{\mu_{500}}$, and the selected 500 functionals, with the final selected domain functional marked with a blue asterisk in the center, the contours being those of $v_{\mu_{500}}$.

and this is explained via (19) by the identity

$$v_{\mu_j}^2(x) = P_{\Lambda_{j-1}}^2(\delta_x) - P_{\Lambda_j}^2(\delta_x), \; P_{\emptyset}^2(\delta_x) = K(x,x)$$

that calls for a large v_{μ_j} when there is a sharp drop from $\rho_{\Lambda_{j-1}}$ to ρ_{Λ_j}.

6.2   Observations for varying smoothness

We now check the behavior of the greedy method when the smoothness parameter m of the Hilbert spaces $W_2^m(\mathbb{R}^d)$ changes. In view of (22) and (23), and since we saw before that the greedy method focuses on the differential operator and the domain, not on boundary values, Figure 10 shows rates -0.67 and -0.71 as

Figure 6: $P^2_{\Lambda_{500}}(\delta_x)$ on the domain, the basis function $v_{\mu_{150}}$, and the selected 150 functionals, with the final selected domain functional marked with a blue asterisk in the northwest, the contours being those of $v_{\mu_{150}}$.

functions of m that confirm the $-(m-3)/2$ rate of (22) for both $\sigma_{\Lambda_{500}}$ and $\rho_{\Lambda_{500}}$ as functions of m after 500 steps of the greedy method.

6.3 Use of Basis Functions

To get examples of solving specific Poisson problems via (20) using the basis obtained by the Greedy Method, we selected a run for smoothness $m = 6$ to generate the basis first, the other parameters being as in the previous examples. Then the functions f and g of (1) were defined to let the true solution be of the form $K(\cdot, z)$ for a different kernel K and a point $z = (-\pi/10, 0)$. This allows to check cases with solutions of different smoothness.

The first case is the infinitely smooth situation where K is a Gaussian. Figure 11 shows a very fast decay of the error, a fast increase of the cumulative sum of the coefficients $\mu_j^2(u)$ from (20), and the final maximal error $8 \cdot 10^{-6}$, roughly. A less smooth solution is of the form $r^{2.5}$ with $r = \|\cdot -z\|_2$ with a derivative singularity at z, and the corresponding results are in Figure 12. The μ_j^2 decay much more slowly, and the solver has to fight with the derivative singularity at z. The two cases were aligned by factors to start with an error of roughly one using the same precalculated basis. Recall that classical Harmonic Analysis shows that decay rates of coefficients of orthonormal expansions depend on smoothness and determine convergence rates.
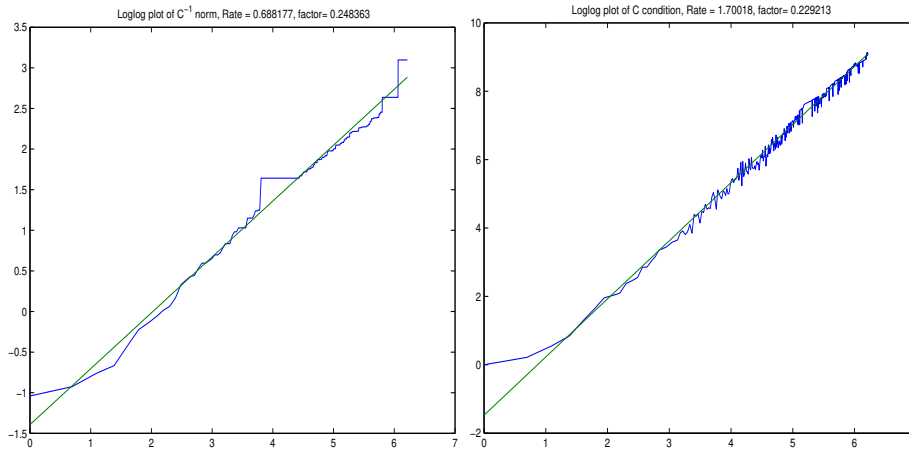
Figure 7: Norm and condition estimate of matrix $C(N)$ as functions of N.

6.4 The Extended Greedy Method

The Extended Greedy Method from Section 4.3 was run in the same situations as in Sections 6.1 and 6.2. The results are given from Figure 13 on.

The rates in the left parts of Figures 1 and 13 are similar, but not in the right parts. Since the extended method picks more boundary points than the original method, the convergence rate on the delta functionals is now much better. Similarly, Figures 2 and 14 differ considerably. The error in the interior decays much better, and the fill distances in the domain and on the boundary show a better alignment.

Figure 15 adds a plot of the Power Function on the domain to the three plots of Figure 4, and it should be compared to Figure 5. One can see that the improved selection of boundary functionals now avoids large values of the Power Function on the boundary. The other results are very similar, and plots are omitted, except for the condition. The extended method is less stable, if Figures 7 and 16 are compared.

From Lemma 2 one might hope that the extended Greedy Method performs better as a function of smoothness, but this cannot be supported by experiments. The rates for the situation corresponding to Figure 10 come out to be roughly -0.69 and -0.71, respectively, but the right-hand plot is much smoother. The actual plot is suppressed.
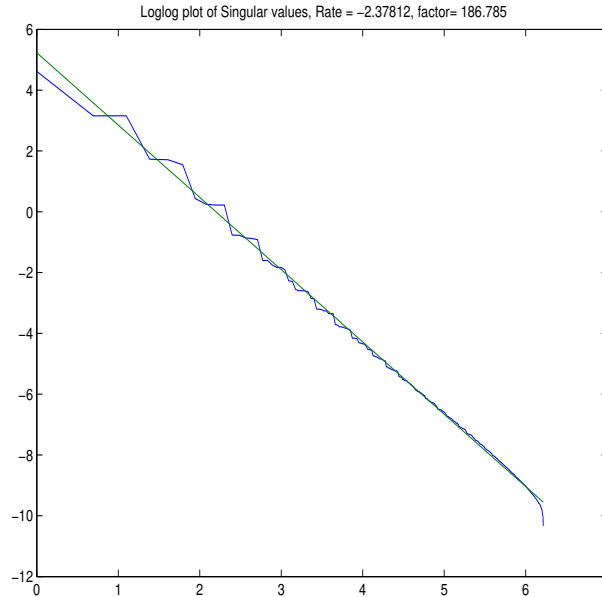
Loglog plot of Singular values, Rate = –2.37812, factor= 186.785

Figure 8: Singular value decay of the matrix of suggested basis functions.

7   Summary and Open Problems

Roughly, the *P*-greedy method for solving Dirichlet problems for second-order elliptic operators in $W_2^m(\mathbb{R}^d)$ seems to behave like the comparable *P*-greedy method for interpolation of functions in $W_2^{m-2}(\mathbb{R}^d)$, with all its pros and cons. It focuses on the domain, not on the boundary, and it tends to produce an asymptotically uniform distribution of evaluation points there, with an unexpectedly small number of points for sampling the boundary values.

On the theoretical side, this opens the quest for a thorough analysis of Kolmogoroff *N*-widths for such PDE problems. A reasonable hypothesis is that these behave like those without differential operators, but for spaces of functions with lower-order smoothness.

Another observation is that the maximum of the generalized Power Function taken on all delta functionals, being a central quantity for pointwise error bounds, shows the same asymptotics as the maximum of the generalized Power Function taken on all chosen PDE data functionals. This is no surprise for well-posed problems,
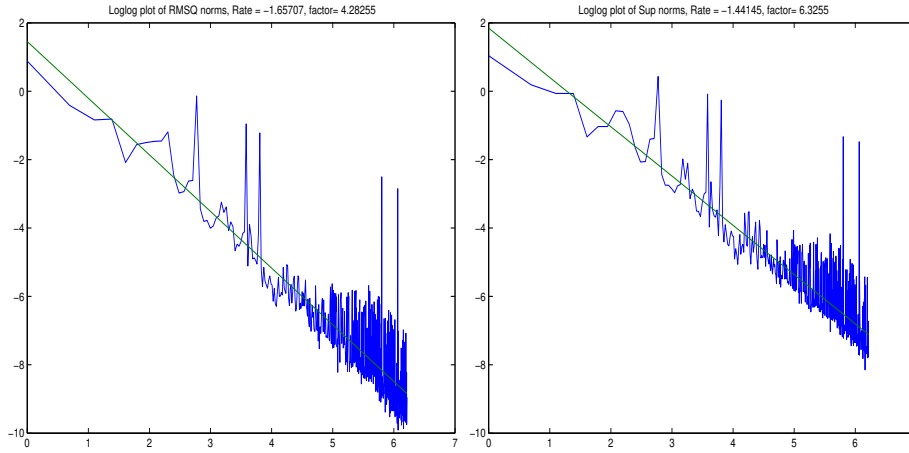
Figure 9: RMSQ and Sup

but it opens a way for explicitly computable factors for error bounds in terms of the Hilbert space norm of the true solution.

We supplied an Extended Greedy Method that cares for the delta functionals in a better way, but it falls out of the Hilbert space foundation, so far. Its analysis is another open problem.

Using the standard Hilbert space background [19, 12], we also can formulate a *P*-greedy method for solving Dirichlet problems for harmonic functions in 2D or 3D. It will follow the Kolmogoroff *N*-width theory for such cases, but the latter seems to be open.

There are other greedy techniques on the market ("*f*–greedy" and "*f*/*P*-greedy") that apply to specific problems of the form (1), not uniformly to the whole class. But their convergence analysis is less far developed, see [18].

Acknowledgment

Plot of rate of PFmax on Lambdas, Slope = –0.673522, Constant= 2.1414

Plot of rate of PFmax on deltas, Slope = –0.711287, Constant= 2.23952

Figure 10: $\sigma_{\Lambda_{500}}$ and $\rho_{\Lambda_{500}}$ as functions of smoothness m, with rates

References

[1] Peter Binev, Albert Cohen, Wolfgang Dahmen, Ronald DeVore, Guergana Petrova, and Przemyslaw Wojtaszczyk. Convergence rates for greedy algorithms in reduced basis methods. *SIAM J. Math. Anal.*, 43(3):1457–1472, 2011.

[2] D. Braess. *Finite Elements. Theory, Fast Solvers and Applications in Solid Mechanics*. Cambridge University Press, 2001. Second edition.

[3] Annalisa Buffa, Yvon Maday, Anthony T. Patera, Christophe Prud'homme, and Gabriel Turinici. *A priori* convergence of the greedy algorithm for the parametrized reduced basis method. *ESAIM Math. Model. Numer. Anal.*, 46(3):595–603, 2012.

[4] Y.L. Chen, S. Gottlieb, A. Heryudono, and A. Narayan. A reduced radial basis function method for partial differential equations on irregular domains. *J. Sci. Comput.*, 66:67–90, 2016.

[5] Stefano De Marchi, R. Schaback, and H. Wendland. Near-optimal data-independent point locations for radial basis function interpolation. *Adv. Comput. Math.*, 23(3):317–330, 2005.
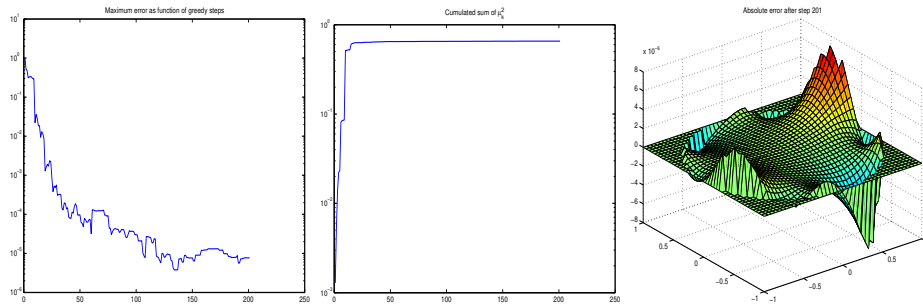
Figure 11: Maximum error and cumulative sum of the $\mu_j^2(u)$ as functions of N, and final error about $8 \cdot 10^{-6}$, Gaussian case
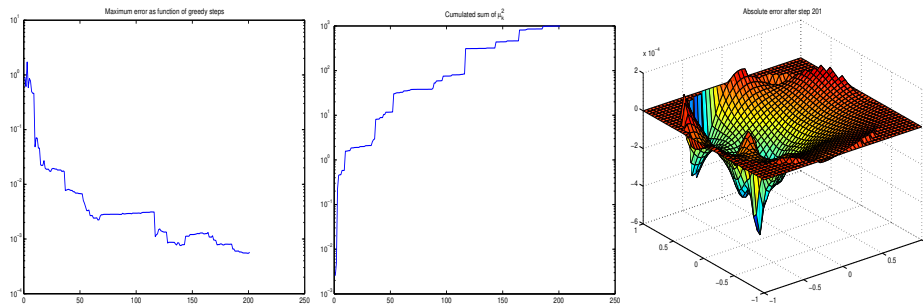

Figure 12: Maximum error and cumulative sum of the $\mu_j^2(u)$ as functions of N, and final error about $4 \cdot 10^{-5}$, Power case $r^{2.5}$

[6] Ronald DeVore, Guergana Petrova, and Przemyslaw Wojtaszczyk. Greedy algorithms for reduced bases in Banach spaces. *Constr. Approx.*, 37(3):455–466, 2013.

[7] G. Fasshauer. Solving partial differential equations by collocation with radial basis functions. In A. LeMéhauté, C. Rabut, and L.L. Schumaker, editors, *Surface Fitting and Multiresolution Methods*, pages 131–138. Vanderbilt University Press, Nashville, TN, 1997.

[8] C. Franke and R. Schaback. Convergence order estimates of meshless collocation methods using radial basis functions. *Advances in Computational Mathematics*, 8:381–399, 1998.
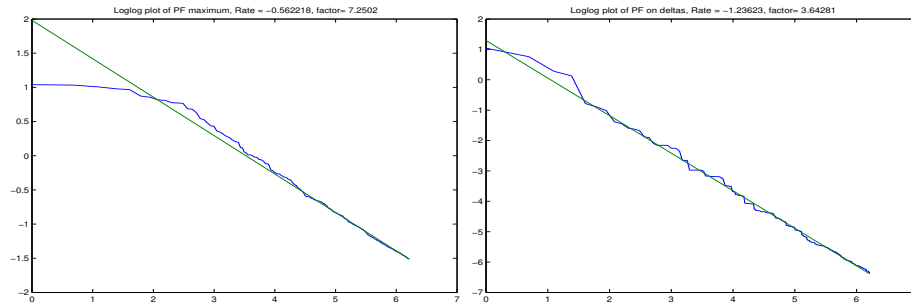
Figure 13: σ_{Λ_N} and ρ_{Λ_N} and their rates as functions of N, for the extended Greedy Method

[9] C. Franke and R. Schaback. Solving partial differential equations by collocation using radial basis functions. *Appl. Math. Comp.*, 93:73–82, 1998.

[10] M. Gubisch and S. Volkwein. Chapter 1: Proper orthogonal decomposition for linear– quadratic optimal control. In *Model Reduction and Approximation: Theory and Algorithms*, pages 3–63. SIAM, Philadelphia, 2016.

[11] Y. C. Hon, R. Schaback, and X. Zhou. An adaptive greedy algorithm for solving large RBF collocation problems. *Numer. Algorithms*, 32(1):13–25, 2003.

[12] Y.C. Hon and R. Schaback. Solving the 3D Laplace equation by meshless collocation via harmonic kernels. *Adv. in Comp. Math.*, pages 1–19, 2013.

[13] K. Kunisch and S. Volkwein. Control of Burgers' equation by reduced order approach using proper orthogonal decomposition. *J. Optim. Theory Appl.*, 102:345–371, 1999.

[14] Yvon Maday, Anthony T. Patera, and Gabriel Turinici. A priori convergence theory for reduced-basis approximations of single-parameter elliptic partial differential equations. In *Proceedings of the Fifth International Conference on Spectral and High Order Methods (ICOSAHOM-01) (Uppsala)*, volume 17, pages 437–446, 2002.

[15] B. Moore. Principal component analysis in nonlinear systems: Preliminary results. In *18th IEEE Conference on Decision and Control Including the Symposium on Adaptive Processes*, volume 2, pages 1057–1060, 1979.
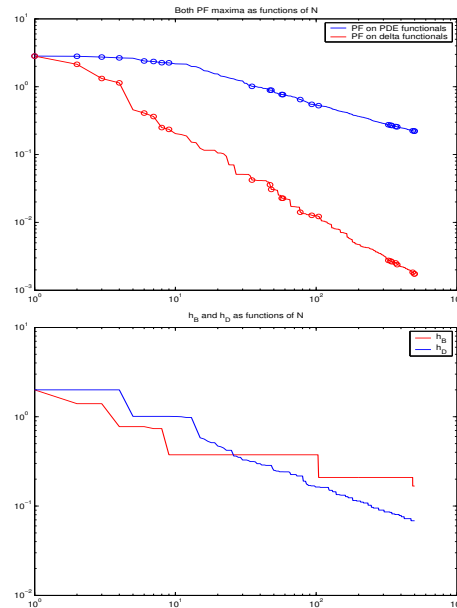
Figure 14: σ_{Λ_N} and ρ_{Λ_N} (top) compared with $h_\Gamma(N)$ and $h_\Omega(N)$ (bottom) as functions of N, for the extended Greedy Method.

[16] M. Pazouki and R. Schaback. Bases for kernel-based spaces. *Computational and Applied Mathematics*, 236:575–588, 2011.

[17] G. Santin and R. Schaback. Approximation of eigenfunctions in kernel-based spaces. *Adv. Comput. Math.*, 42(4):973–993, 2016.

[18] Gabriele Santin and Bernard Haasdonk. Convergence rate of the data-independent *P*-greedy algorithm in kernel-based approximation. *Dolomites Res. Notes Approx.*, 10(Special Issue):68–78, 2017.

[19] R. Schaback. Solving the Laplace equation by meshless collocation using harmonic kernels. *Adv. in Comp. Math.*, 31:457–470, 2009. DOI 10.1007/s10444-008-9078-3.

[20] R. Schaback. A computational tool for comparing all linear PDE solvers. *Advances of Computational Mathematics*, 41:333–355, 2015.

[21] R. Schaback. All well–posed problems have uniformly stable and convergent discretizations. *Numerische Mathematik*, 132:597–630, 2016.
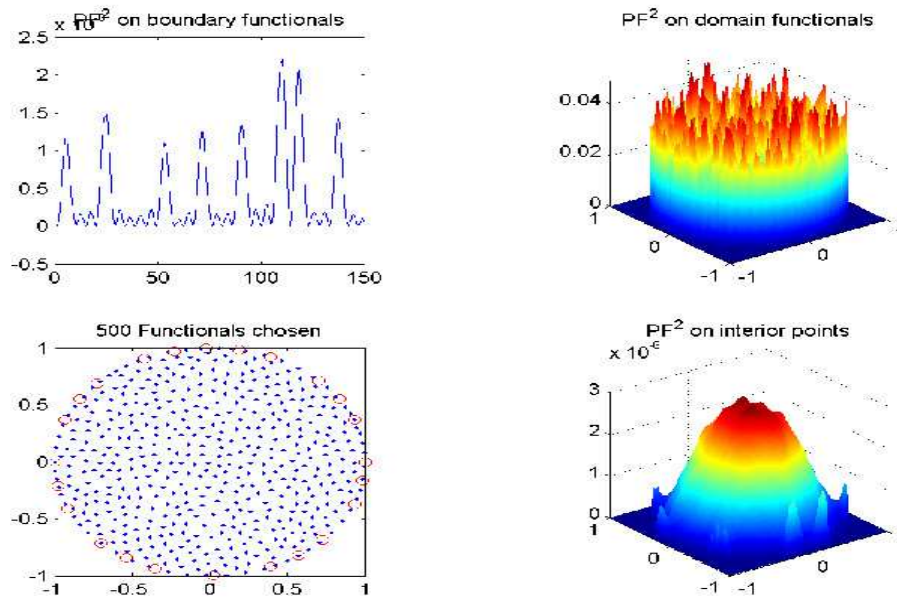
Figure 15: $P^2_{\Lambda_{500}}$ on the boundary functionals, the domain functionals, and the selected 500 functionals, for the extended Greedy Method.

[22] H. Wendland. *Scattered Data Approximation*. Cambridge University Press, 2005.

[23] Z. Wu. Hermite–Birkhoff interpolation of scattered data by radial basis functions. *Approximation Theory and its Applications*, 8/2:1–10, 1992.
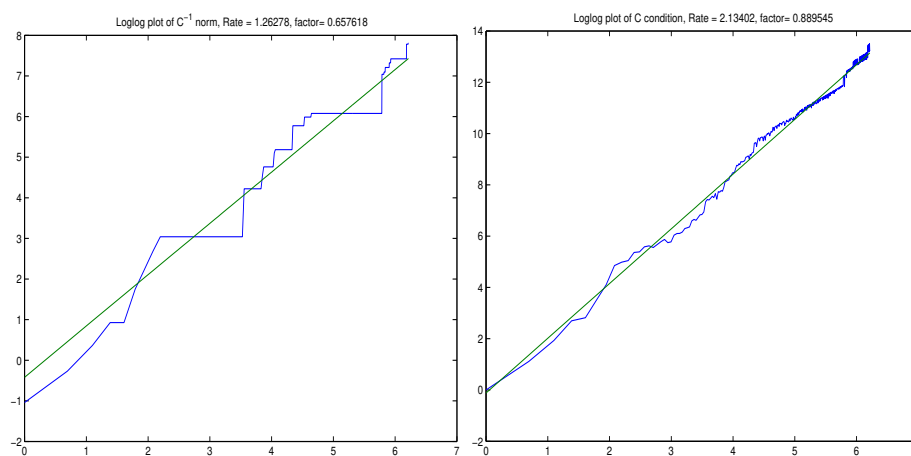
Figure 16: Norm and condition estimate of matrix $C(N)$ as functions of N, for the extended Greedy Method.