Eigenbasis and Approximation

Gabriele Santin\textsuperscript{*1} and Robert Schaback\textsuperscript{†2}

\textsuperscript{1}Department of Mathematics, University of Padova
\textsuperscript{2}Institut für Numerische und Angewandte Mathematik, Universität Göttingen

Draft of Nov. 14, 2014

Contents

1 Introduction 1
2 $L_2$-minimal power function 2
3 Point-based sets 3
3.1 Construction of the basis ............... 6
3.2 Greedy approximation ................. 7
4 Approximation of linear functionals 8
5 Experiments 9
5.1 Decay of the Power Function ............ 10
5.2 Approximation of the eigencouples ........ 10

1 Introduction

Consider a $n$-dimensional subspace $V_n \subset \mathcal{N}_K(\Omega)$, $\Omega \subset \mathbb{R}^d$ being a compact set, and the corresponding projection operator $\Pi_{V_n} : \mathcal{N}_K(\Omega) \rightarrow V_n$. We

\textsuperscript{*}gsantin@math.unipd.it, http://orcid.org/0000-0001-6959-1070
\textsuperscript{†}schaback@math.uni-goettingen.de
have for any \( f \in \mathcal{N}_K(\Omega) \)

\[
|f(x) - \Pi_{V_n}(x)| \leq \|P_{V_n}(x)\| \quad \forall x \in \Omega
\]

where \( P_{V_n} \) is the Power Function corresponding to \( V_n \). If \( H_{V_n} \) is the reproducing kernel of \( V_n^\perp \) we have \( P_{V_n}(x)^2 = \|H_{V_n}(\cdot, x)\|^2 = H_{V_n}(x, x) \).

The problem we consider is to find an optimal subspace wrt the \( L_2 \) norm, i.e., a subspace which minimizes \( \|P_{V_n}\|_{L_2} \) among all \( n \)-dimensional subspaces \( V_n \subset \mathcal{N}_K(\Omega) \).

Let \( T : L_2 \to L_2 \) be the usual integral operator

\[
T[f](x) = \int_\Omega K(x, y)f(y)dy
\]

with corresponding eigencouples \( \{\varphi_j\}_{j \in \mathbb{N}}, \{\lambda_j\}_{j \in \mathbb{N}} \), and define the subspace \( E_n = \text{span}\{\sqrt{\lambda_j}\varphi_j, 1 \leq j \leq n\} \). The basis \( \{\sqrt{\lambda_j}\varphi_j\}_{j \in \mathbb{N}} \) is called the eigenbasis.

**Remark 1.** The operator can be defined using any positive and finite measure \( \mu \) with full support on \( \Omega \) (see \([9]\)) , but for simplicity we will use here only the Lebesgue measure.

## 2 \( L_2 \)-minimal power function

First, we prove that \( E_n \) is the \( n \)-dimensional subset whose power function is of minimal \( L_2 \) norm.

**Theorem 2.** For any \( n \)-dimensional subspace \( V_n \subset \mathcal{N}_K(\Omega) \) we have

\[
\|P_{V_n}\|_{L_2} \geq \sqrt{\sum_{j=n+1}^{\infty} \lambda_j}, \quad (2)
\]

and the subspace \( E_n \) is the unique optimal subset for any \( n > 0 \), i.e., equality holds.

**Proof.** The proof is the same of \([3, \text{Theorem 1}]\).

For a subset \( V_n \) we can consider a onb \( \{\psi_k\}_{k=1}^n \) and complete it to a onb \( \{\psi_k\}_{k \in \mathbb{N}} \) of \( \mathcal{N}_K(\Omega) \). We can move from this basis to the eigenbasis using a
matrix $A = (a_{ij})$ as follows

$$\psi_k = \sum_{j=1}^{\infty} a_{jk} \sqrt{\lambda_j} \varphi_j,$$

where $\sum_{j=1}^{\infty} a_{jk}^2 = \sum_{k=1}^{\infty} a_{jk}^2 = 1$.

Hence, the power function of $V_n$ is

$$P_{V_n}(x)^2 = \sum_{k=n+1}^{\infty} \psi(x)^2 = \sum_{k=n+1}^{\infty} \left( \sum_{j=1}^{\infty} a_{jk} \sqrt{\lambda_j} \varphi_j(x) \right)^2$$

$$= \sum_{i,j=1}^{\infty} \sqrt{\lambda_i} \varphi_i(x) \sqrt{\lambda_j} \varphi_j(x) \sum_{k=n+1}^{\infty} a_{ik} a_{jk},$$

and in particular,

$$\|P_{V_n}\|_{L_2}^2 = \int_\Omega P_{V_n}(x)^2 dx = \int_\Omega \sum_{k=n+1}^{\infty} \left( \sum_{j=1}^{\infty} a_{jk} \sqrt{\lambda_j} \varphi_j(x) \right)^2 dx$$

$$= \sum_{k=n+1}^{\infty} \int_\Omega \left( \sum_{j=1}^{\infty} a_{jk} \sqrt{\lambda_j} \varphi_j(x) \right)^2 d\mu(x) = \sum_{k=n+1}^{\infty} \sum_{j=1}^{\infty} a_{jk}^2 \lambda_j,$$

where we used the orthonormality of the eigenbasis in $L_2$.

Now let $q_j = \sum_{k=1}^{n} a_{jk}^2$. We can bound the last term as

$$\sum_{k=n+1}^{\infty} \sum_{j=1}^{\infty} a_{jk}^2 \lambda_j = \sum_{j=1}^{\infty} \lambda_j - \sum_{j=1}^{n} q_j \lambda_j \geq \sum_{j=n+1}^{\infty} \lambda_j,$$

(4)

since $\sum_{j=1}^{\infty} q_j \lambda_j \leq \sum_{j=1}^{n} \lambda_j$.

In the case of the eigenbasis, it suffices to notice that if we take $V_n = E_n$ and $\{ \sqrt{\lambda_j} \varphi_j \}_{j=1}^{n}$ as its basis, the matrix $A$ is the (infinite) identity matrix, thus in (4) equality holds.

3 Point-based sets

The eigenbasis cannot be computed in general and, moreover, it can not be used for truly scattered data approximation, since there exist at least a set...
$X_n \subset \Omega$ such that the collocation matrix of $E_n$ on $X_n$ is singular (by the Mairhuber-Curtis Theorem, see e.g. [10, Theorem 2.3]).

The usual setting is instead to consider subsets of $\mathcal{N}_K(\Omega)$ of the form $V_m = \text{span}\{K(\cdot, y) : y \in X_m\}$, where $X_m = \{x_1, \ldots, x_m\}$ is a possibly large set of points in $\Omega$.

We can hence define a constrained version of the problem of the previous section, i.e., the problem of finding, for $n < m$, a $n$-dimensional subset $E_{n,m} \subset V_m$ which minimizes the $L_2$ norm of the power function among all the subspace of $V_m$ of dimension $n$. In this view, the subset $X_m$ (or the subspace $V_m$) is used as a sampling set, and $E_{n,m}$ should provide a point-dependent approximation of $E_n$. In fact, as we will see, the optimal subspace $E_{n,m}$ is not given by the best approximation of $E_n$, but it will approximate it in a near-optimal sense.

The problem can be solved in the same way of the general case. Consider the operator $T_m : L_2 \to L_2$ defined by

$$T_m[f](x) = \int_{\Omega} K_m(x,y)f(x)dy,$$

where $K_m$ is the reproducing kernel of $V_m$.

Since $V_m$ is the native space on $\Omega$ of the kernel $K_m$, the same result of the previous section applies here. In particular the operator has a set of eigenvalues $\{\lambda_{j,m}\}_{j=1}^{m}$ and eigenvectors $\{\varphi_{j,m}\}_{j=1}^{m}$ such that $\langle \varphi_{i,m}, \varphi_{j,m}\rangle_{L_2} = \delta_{ij}, \langle \varphi_{i,m}, \varphi_{j,m}\rangle = \lambda_{j,m}\delta_{ij}$ and $\{\sqrt{\lambda_{j,m}}\varphi_{j,m}\}_{j=1}^{m}$ is a o.n. basis of $V_m$.

We can define for $V_n \subset V_m$ the squared power function restricted to $V_m$ as (see [5]),

$$P_{V_n}^{(m)}(x)^2 = K_m(x,x) - \sum_{j=1}^{n} v_j(x)^2, \quad x \in \Omega,$$

where $\{v_j\}_{j=1}^{n}$ is any o.n. basis of $V_n$.

By Theorem 2, $E_{n,m} = \text{span}\{\sqrt{\lambda_{1,m}}\varphi_{1,m}, \ldots, \sqrt{\lambda_{n,m}}\varphi_{n,m}\}$ is the $n$-dimensional subset of $V_m$ which minimizes the $L_2$ norm of $P^{(m)}$. Moreover, for any $V_n \subset V_m$ and for any $x \in \Omega$, the squared power functions of $\mathcal{N}_K(\Omega)$ and of $V_m$ differ by an additive constant, i.e.,

$$P_{V_n}(x)^2 - P_{V_n}^{(m)}(x)^2 = P_{V_m}(x)^2.$$

This means that the minimality of $E_{n,m}$ does not change if we consider the standard power function on $\mathcal{N}_K(\Omega)$. We proved the following theorem:
**Theorem 3.** Let $V_m \subset \mathcal{N}_K(\Omega)$ be a subspace of dimension $m$, and let $n \leq m$. For any $n$-dimensional subspace $V_n \subseteq V_m$ we have

$$\|P_{V_n}\|_{L_2} \geq \sqrt{\sum_{j=1}^{\infty} \lambda_j - \sum_{j=1}^{n} \lambda_{j,m}},$$

and the subspace $E_{n,m}$ is the unique optimal subset for any $n > 0$, i.e., equality holds.

Regarding the relation between $E_{n,m}$ and $E_n$, observe that $T_m = \Pi_{V_m} T$, since, if $\{v_j\}_{j=1}^{n}$ is any o.n. basis of $V_m$, so that $K_m(x,y) = \sum_{j=1}^{m} v_j(x)v_j(y)$, we have

$$T_m[f](x) = \int_{\Omega} K_m(x,y)f(y)dy = \sum_{j=1}^{m} v_j(x) \langle v_j, f \rangle_{L_2}$$

$$= \sum_{j=1}^{m} v_j(x) \langle v_j, T[f] \rangle = \Pi_{V_m} T[f](x).$$

This means that the couples $(\lambda_{j,m}, \varphi_{j,m})$, $1 \leq j \leq m$, are precisely the Bubnov - Galerkin approximations (see e.g. [4, Sec. 18.4]) of the solutions $(\lambda_j, \varphi_j)$, $j \in \mathbb{N}$, of the eigenvalue problem for the restricted operator $T : \mathcal{N}_K(\Omega) \to \mathcal{N}_K(\Omega)$ (which is still a compact, positive and self-adjoint operator). We can then use the well known estimates on the convergence of the Bubnov - Galerkin method to express the distance between $E_n$ and $E_{n,m}$.

The following Proposition collects convergence results which follows from [4, Th. 18.5, Th. 18.6].

**Proposition 4.** Let $m \geq 1$ and let $V_1 \subset V_2 \subset \ldots \subset V_m \subset \ldots$ be a sequence of subsets which becomes dense in $\mathcal{N}_K(\Omega)$. For $1 \leq j \leq m$ we have

(i) $\lambda_{j,m} \leq \lambda_{j,m+1} \leq \lambda_j$,

(ii) $\lambda_j - \lambda_{j,m} \leq \|id - \Pi_{V_m}\|_1$[

(iii) Let $r \in \mathbb{N}$ be the multiplicity of $\lambda_j$ and $F_{j,m} = \{f \in V_m : T_m f = \lambda_{i,m} f \text{ and } \lim_{m \to \infty} \lambda_{i,m} = \lambda_j\}$. For $m$ sufficiently large $\dim F_{j,m} = r$ and there exists $c_{j,m} > 1/\lambda_j$, $c_{j,m} \to m 1/\lambda_j$ s.t.

$$\|\varphi_j - \Pi F_{j,m} \varphi_j\| \leq c_{j,m} \lambda_j \|\varphi_j - \Pi_{V_m} \varphi_j\|. \quad (7)$$

1There are sharper estimates for this, done by taking into account (iii).
Moreover, recall that the eigenbasis can be approximated with double rate of convergence (see \[8\]), i.e., for \(j \in \mathbb{N}\),

\[
|\varphi_j(x) - \Pi_{V_m}\varphi_j(x)| \leq \lambda_j^{-1}P_{V_m}(x)\|P_{V_m}\|_{L_2}.
\]

Using this fact and (7), for \(m\) sufficiently large and \(1 \leq j \leq m\), we have also

\[
|\varphi_j(x) - \Pi_{F_j,m}\varphi_j(x)| \leq c_{j,m}P_{F_j,m}(x)\|P_{V_m}\|_{L_2},
\]

where

\[
P_{F_j,m}(x)^2 = K(x,x) - \sum_{i: \lambda_{i,m} \rightarrow \lambda_j} \lambda_{i,m}\varphi_{i,m}(x)^2.
\]

### 3.1 Construction of the basis

Notice that \(\{\varphi_{j,m}\}_{j=1}^m\) is the unique set of \(m\) functions in \(V_m\) which is orthonormal in \(L_2\) and orthogonal in \(N_K(\Omega)\), indeed, any such set needs to be a set of eigenvectors of \(T_m\) whose \(N_K(\Omega)\) norms equal the eigenvalues of \(T_m\).

Moreover, given any couple of inner products \(\langle \cdot, \cdot \rangle_a\) and \(\langle \cdot, \cdot \rangle_b\) on \(V_m\), it is always possible to build a basis \(\{u_j\}_{j=1}^m\) of \(V_m\) which is \(b\)-o.n. and \(a\)-orthogonal with norms \(\{\gamma_j\}_{j=1}^m\). Indeed, let \(A = ((K(\cdot, x_i), K(\cdot, x_j))_a)_{i,j=1}^m\) and \(B = ((K(\cdot, x_i), K(\cdot, x_j))_b)_{i,j=1}^m\) be the gramians of the standard basis wrt the two inner products. Following the notation of [7], we need to construct an invertible matrix \(C_U\) s.t. \(C_U^TAC_U = \Gamma\) and \(C_U^TBC_U = I\), where \(\Gamma\) is the diagonal matrix having on the diagonal the \(a\)-norms of the new basis. This means to simultaneously diagonalize the two gramians, and since they are symmetric and positive definite this is always possible, e.g. in the following way:

- let \(B = LL^T\) be a Cholesky decomposition,
- define \(C = L^{-1}AL^{-T}\) (which is symmetric and positive definite),
- let \(C = UTU^T\) be a SVD decomposition,
- define \(C_U = L^{-T}U\).

Observe that, for practical use, is more convenient to swap the role of \(A\) and \(B\). In this way we construct the basis \(\{\sqrt{\lambda_{j,m}}\varphi_{j,m}\}_{j=1}^m\), which is \(N_K(\Omega)\)-o.n. hence more suitable for approximation purposes, and, moreover, we obtain directly the eigenvalues of order \(m\) as \(\Gamma = \text{diag}(\lambda_{1,m}, \ldots, \lambda_{m,m})\).

In both ways we are just computing a generalized diagonalization of the pencil \((A, B)\), up to a proper scaling of the diagonals.
Remark 5. If the two inner products are the one of $N_K(\Omega)$ and of $\ell_2(X_m)$, the two gramians are the kernel matrix $A$ and $A^2$, so we are in the situation of [7]. In particular, a single SVD (e.g. of $A$) is enough to diagonalize both the matrices. This is also the situation when the $L_2$ inner product is approximated by $\ell_2(X_m)$.

In our case $A$ is the usual kernel matrix, while $B_{ij} = \langle K(\cdot, x_i), K(\cdot, x_j) \rangle_{L_2}$, i.e.,

$$A = \Phi(X_m) \Lambda \Phi(X_m)^T, \quad B = \Phi(X_m) \Lambda^2 \Phi(X_m)^T,$$

where $\Phi(X_m)$ is the $m \times \infty$ collocation matrix of the eigenbasis on $X_m$ and $\Lambda$ is the infinite and square diagonal matrix with the eigenvalues of $T$ on the diagonal.

Thus, provided the knowledge of $B$, we can explicitly construct the basis. This can be done, for a general kernel, using a large set of points (not $X_m$ itself, because, in view of the previous remark, this would just lead to the (W)SVD basis) to approximate the $L_2$ inner product.

Numerical experiments (see Section 5) suggest that this construction of $E_{n,m}$ is highly unstable also in simple cases. Another way to face the problem is the following greedy procedure.

3.2 Greedy approximation

Instead of directly constructing the subspace $U_n$ as described in the previous section, we can build another $n$-dimensional subset $G_n \subset V_m$ through a greedy minimization of the $L_2$ norm of the power function. Also in this case, except for some particular kernel, we need to approximate the $L_2$ inner product.

The algorithm works as follows. Since minimizing the $L_2$ norm of the power function of a subspace is equivalent to maximize the sum of the $L_2$ norms of a basis, the first point can be chosen as

$$x_1 = \arg \max_{x \in X_m} \left\| \frac{K(\cdot, x)}{\sqrt{K(x, x)}} \right\|_{L_2}.$$

Denoting by $X_i$ the already chosen points, and by $\{v_1, v_2, \ldots, v_i\}$ the Newton basis of $G_i$ (see [6]), the $(i+1)$-th point is selected as

$$x_{i+1} = \arg \max_{x \in X_m \setminus X_i} \sum_{j=1}^{i+1} \|v_j\|^2_{L_2} = \arg \max_{x \in X_m \setminus X_i} \|v_{i+1}\|^2_{L_2}.$$
Of course we obtain \( \|P_{G_n}\|_{L^2} > \|P_{E_{n,m}}\|_{L^2} \). The problem now is to find an upper bound.

This is equivalent to study the relation between \( \sum_{j=1}^{n} \lambda_{j,m} \) and \( \sum_{j=1}^{n} \lambda_{j,n} \), where \( \{\lambda_{j,n}\}_{j=1}^{n} \) are the eigenvalues of the operator \( T_n \) corresponding to the subspace \( G_n \), and \( \{\lambda_{j,m}\}_{j=1}^{n} \) are the first \( n \) eigenvalues of \( T_m \).

### 4 Approximation of linear functionals

The previous results can be easily extended to the approximation of linear functionals (see e.g. [10, Ch. 16]). Namely, let \( L : \mathcal{N}_K(\Omega) \to \mathcal{C}(\Omega) \) be a linear operator such that \( \delta_x \circ L \in \mathcal{N}_K(\Omega)^* \) for any \( x \in \Omega \) and \( \{\delta x \circ L : x \in \Omega\} \) is linearly independent over \( \mathcal{N}_K(\Omega) \). In this situation it is known that \( L(\mathcal{N}_K(\Omega)) = \mathcal{N}_{KL} \), where \( \mathcal{N}_{KL} \) is the native space of the symmetric and positive definite kernel

\[
K_L(x,y) = L^x L^y K(x,y), \quad x, y \in \Omega.
\]

Given \( g \in \mathcal{N}_{KL} \) and \( V_m = \text{span}\{v_j, 1 \leq j \leq m\} \subset \mathcal{N}_{KL} \), we look for a solution \( s \in V_m \) of the problem

\[
Lf = g,
\]

in the sense that

\[
\langle Lf, v_j \rangle_{\mathcal{N}_{KL}} = \langle g, v_j \rangle_{\mathcal{N}_{KL}}, \quad 1 \leq j \leq m.
\]

The standard choice for \( V_m \), as in the interpolation case, is \( V_m = \text{span}\{LK(\cdot, x_i), 1 \leq i \leq m\} \), but we want to consider more general subspaces of \( \mathcal{N}_{KL} \).

Under the above assumptions the problem can be uniquely solved and we have the error bound

\[
|Lf(x) - Ls(x)| \leq P_{K,V_m \circ L}(\delta_x \circ L)\|f\|_{\mathcal{N}_K(\Omega)}
\]

where \( P_{K,V_m \circ L}(\delta_x \circ L) \) is the generalized power function, defined, for any \( \Lambda \subset \mathcal{N}_{KL}^* \) and \( \lambda \in \mathcal{N}_{KL}^* \), as

\[
P_{K,\Lambda}(\lambda) = \inf_{\mu \in \Lambda} \|\lambda - \mu\|_{\mathcal{N}_K(\Omega)^*}.
\]

So our problem is to find \( V_m \in \mathcal{N}_{KL} \) which minimizes \( \|P_{K,V_m \circ L}(\delta_x \circ L)\|_{L^2} \).
Since, for any $\Lambda$ and $\lambda$ as above, we can move from $N_K(\Omega)$ to $N_{KL}(\Omega)$ as

$$P_{KL,\lambda}(\Lambda \circ L) = P_{KL}(\lambda),$$

we have $P_{KL,V_m}(\delta_x \circ L) = P_{KL,V_m}(\delta_x)$, and $P_{KL,V_m}(\delta_x)$ is the usual power function of $V_m$ on $N_{KL}$.

Thus, the problem can be solved as in the previous sections by replacing $K$ with $KL$, i.e., by considering the operator $T_L : L^2 \to L^2$,

$$T_L[f](x) = \int_\Omega K_L(x,y)f(x)dy.$$ 

**Remark 6.** If we consider $L = \Pi V_m$, $V_m \subset N_K(\Omega)$, we obtain the results of Section 3 since $KL = \Pi_x \Pi_y V_m K(x,y) = K_m(x,y)$.

5 Experiments

In this section we make some experiment using the iterated Brownian bridge kernels (see [1]). This family of kernels is useful for our purposes because the exact eigenbasis is explicitly known and the smoothness of the kernel can be controlled using a parameter.

The kernels are defined, for $\beta \in \mathbb{N} \setminus \{0\}$, $\varepsilon \geq 0$ and $x, y \in [0,1]$, as

$$K_{\beta,\varepsilon}(x,y) = \sum_{j=1}^{\infty} \lambda_j(\varepsilon, \beta) \varphi_j(x) \varphi_j(y), \quad (9)$$

where

$$\lambda_j(\varepsilon, \beta) = \left(j^2 \pi^2 + \varepsilon^2\right)^{-\beta}, \quad \varphi_j(x) = \sin(j\pi x).$$

The kernel is of class $2\beta - 2$. For $\beta = 1$ and $\varepsilon = 0$, the kernel has the form $K_{1,0}(x,y) = \min(x,y) - xy$, but a general closed form is not known. In the following tests we will compute it using a truncation of the series \([9]\) at $j = N$, with a sufficiently large $N$.

For simplicity, we will consider for now only $\varepsilon = 0, 1$.

Thanks to the knowledge of the explicit expansion, we can also compute the $L^2$-gramian of $K_{\beta,\varepsilon}$ by squaring the eigenvalues in \([9]\), i.e.,

$$K^{(2)}_{\beta,\varepsilon}(x,y) = \langle K_{\beta,\varepsilon}(x, \cdot)K_{\beta,\varepsilon}(\cdot,y) \rangle_{L^2} = \sum_{j=1}^{\infty} \lambda_j(\varepsilon, 2\beta) \varphi_j(x) \varphi_j(y), \quad (10)$$

9
5.1 Decay of the Power Function

Now we want to compare the optimal decay of the power function with the one obtained by starting from a set of points \( X_m \), both in the direct and in the greedy way.

We take \( m = 500 \) randomly distributed points in \((0, 1)\) and we select, for \( n = 50 \), the subspaces \( E_{n,m} \) and \( G_n \) defined in Section 3. Moreover, we consider also the subspace \( N_n \) produced by the algorithm described in \([2, 6]\), which chooses points which maximizes the point-wise value of the power function.

To speed up the greedy algorithm for \( G_n \), we approximate the \( L_2 \) inner product with the discrete one on \( X_m \). This step, of course, introduces an error in the selection of the points. Nevertheless, in order to evaluate properly the performance of the algorithm, after the construction of \( G_n \) we compute the \( L_2 \) norms of the corresponding power function exactly, i.e., using (10).

The results for \( \beta = 1, \ldots, 4 \) and \( \epsilon = 0, 1 \) are shown in Figure 1. Lines not present in the plots mean that the corresponding power function is negative, because of numerical instability.

First, notice that the direct method is sufficiently stable only for \( \beta = 1 \), but in this case it is able to nearly reproduce the optimal rate of decay. The greedy algorithms, on the other hand, are feasible for all the choices of the parameters, and they have a convergence between \( \|P_{E_{n,m}}\|_{L_2} \) and \( \|P_{E_{2n,m}}\|_{L_2} \). As the kernel becomes smoother their performance is better, but they become unstable. Observe also that \( N_n \) and \( G_n \) behave essentially in the same way, even if the first one is not designed to minimize any \( L_2 \) norm.

Unlike for the smoothness parameter \( \beta \), there is not clear dependence on \( \epsilon \) of the performance of the algorithm, except in the case \( \beta = 4 \), where \( \epsilon = 1 \) gives a better stability of \( N_n \).

5.2 Approximation of the eigencouples

Here we test the convergence of the approximate eigencouples to the exact ones through the computation of \( E_{n,m} \).

Since the method becomes unstable, we use here a smaller set of \( m = 100 \) randomly distributed points in \((0, 1)\) and we check the approximation for the first \( n = 50 \) elements.

Figure 2 displays the results for \( \beta = 1, 2, 3, 4 \) and \( \epsilon = 0 \). We underline that
the eigenbasis is defined up to a change of sign, that is, we can expect to approximate $|\sqrt{\lambda_j} \varphi_j|$, $1 \leq j \leq n$.

As expected from Proposition 4, both for the eigenvalues and the eigenfunctions the convergence is faster for smoother kernels and for the smaller indexes $j$. Also in this example, for $\beta = 4$, the approximation becomes unstable.

References

[1] R. Cavoretto, G. Fasshauer, and M. McCourt. An introduction to the hilbert-schmidt svd using iterated brownian bridge kernels. *Numerical Algorithms*, pages 1–30, 2014.

[2] S. 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.

[3] R. Ismagilov. On n-dimensional diameters of compacts in a hilbert space. *Functional Analysis and Its Applications*, 2(2):125–132, 1968.

[4] M. A. Krasnosel'skiĭ, G. M. Vaĭnikko, P. P. Zabreĭko, Y. B. Rutitskiĭ, and V. Y. Stetsenko. *Approximate solution of operator equations*. Wolters-Noordhoff Publishing, Groningen, 1972. Translated from the Russian by D. Louvish.

[5] M. Mouattamid and R. Schaback. Recursive kernels. *Anal. Theory Appl.*, 25(4):301–316, 2009.

[6] S. M"uller and R. Schaback. A Newton basis for kernel spaces. *J. Approx. Theory*, 161(2):645–655, 2009.

[7] M. Pazouki and R. Schaback. Bases for kernel-based spaces. *Journal of Computational and Applied Mathematics*, 236(4):575–588, 2011.

[8] R. Schaback. A unified theory of radial basis functions. Native Hilbert spaces for radial basis functions. II. *J. Comput. Appl. Math.*, 121(1-2):165–177, 2000. Numerical analysis in the 20th century, Vol. I, Approximation theory.

[9] H. Sun and Q. Wu. Application of integral operator for regularized least-square regression. *Math. Comput. Modelling*, 49(1-2):276–285, 2009.
[10] H. Wendland. *Scattered Data Approximation*, volume 17 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge, 2005.
Figure 1: Decay of the power functions described in Section 5.1 for different parameters (from left to right: $\varepsilon = 0, 1$; from top to bottom: $\beta = 1, 2, 3, 4$).
Figure 2: Approximation of the eigenvalues (left) and the eigenbasis (right), for $\beta = 1, 2, 3$ (from top to bottom), $\varepsilon = 0$ and for $1 \leq j \leq n$, $n = 50$. 