Applied Mathematics and Computation 416 (2022) 126749

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

Best kernel approximation in Bergman spaces[☆]

Wei Qu[a], Tao Qian[b,*], Haichou Li[c], Kehe Zhu[d]

[a] Department of Applied Mathematics, School of Mathematics and Physics, University of Science and Technology Beijing, China
[b] Macau Center for Mathematical Sciences, Macau University of Science and Technology, Macau
[c] School of Mathematics and Informatics, South China Agricultural University, Guangzhou, Guangdong 510640, China
[d] Department of Mathematics and Statistics, SUNY, Albany, NY 12222, USA

ARTICLE INFO

Article history:
Received 23 June 2021
Revised 17 September 2021
Accepted 17 October 2021
Available online 10 November 2021

2010 MSC:
30H20
30C40
30J10
46E22
46E20

Keywords:
Bergman space
Reproducing kernel hilbert space
Blaschke product
Best kernel approximation

ABSTRACT

Let H be a reproducing kernel Hilbert space of analytic functions on the unit disk $\mathbb{D}$. The best kernel approximation problem for H is the following: given any positive integer n and any function $f\in H$ find the best norm approximation of f by a linear combination of no more than n kernel functions $K(z,z_k)$, $1\le k\le n$. The purpose of this paper is to prove the existence of best kernel approximation for weighted Bergman spaces with standard weights.

1. Introduction

Rational approximation over a long period of time has been an active research area [22,23]. It has fruitful results and ample applications in mathematics itself, as well as in practical problems, including, in particular, in system identification [12–15,24]. In the study of rational approximation great interest has been devoted to the classical Hardy spaces, namely the Hardy spaces of the unit disc and that of a half of the complex plane. Among others the question of best approximations to functions in the Hardy spaces by rational functions of degree not exceeding n was brought attention. Existence of a solution to this n-best problems has long been proved [22], yet, new proofs have been explored in order to obtain a mathematically ultimate algorithm [16–18,25]. The same type problems in weighted Bergman and weighted Hardy spaces naturally arise. One of the motivations is that in control theory some system functions belong to those spaces [2,3]. In those non-Hardy cases the original question in terms of rational function is naturally adapted to whether there exists a best approximation by linear combinations of at most n reproducing kernels. The present study establishes such existence result for all weighted

☆ Qian is funded by the Science and Technology Development Fund, Macau SAR (File no. 0123/2018/A3). Li is supported by NNSF of China (Grant no. 11901205). Zhu is partially supported by NNSF of China (Grant no. 11720101003), NSF of Guangdong Province (Grant no. 2018A030313512), Key projects of fundamental research in universities of Guangdong Province (Grant no. 2018KZDXM034) and STU Scientific Research Foundation (Grant no. NTF17009).

* Corresponding author.
E-mail addresses: quwei2math@qq.com (W. Qu), tqian@must.edu.mo (T. Qian), hcl2016@scau.edu.cn (H. Li), kzhu@albany.edu (K. Zhu).

https://doi.org/10.1016/j.amc.2021.126749
0096-3003/

Bergman spaces. The main results and the related kernel estimations are seen to be new, and significant to both the analysis of the analytic function spaces, and algorithms solutions of system identification questions.

Let $\mathbb{D}$ be the open unit disk in the complex plane $\mathbb{C}$ and let $H(\mathbb{D})$ be the space of all analytic functions on $\mathbb{D}$. A complex Hilbert space $H \subset H(\mathbb{D})$ is called a reproducing kernel Hilbert space if for every point $a \in \mathbb{D}$ the point-evaluation $f \mapsto f(a)$ is a bounded linear functional on H. Then by Riesz representation, there must exist a unique function $K_a \in H$ such that $f(a) = \langle f, K_a \rangle$ for all $f \in H$. The function $K(z,w) = K_w(z)$ on $\mathbb{D} \times \mathbb{D}$ is called the reproducing kernel of H.

It is easy to show that the kernel function K has the following properties: $K(z,w)$ is analytic in z, conjugate analytic in w, and $\overline{K(z,w)} = K(w,z)$. Furthermore,

$$K(z,z) = K_z(z) = \langle K_z, K_z \rangle = \|K_z\|^2, \qquad z \in \mathbb{D}.$$

In the case when $K(z,z) > 0$, we will consider the unit vector $k_z = K_z/\|K_z\|$, which is called the normalized reproducing kernel of H at z.

The most classical examples of reproducing kernel Hilbert spaces of analytic functions on the unit disk include the Hardy space H^2 and the weighted Bergman spaces A^2_α with $\alpha > -1$. Recall that H^2 consists of analytic functions $f(z) = \sum_{k=0}^{\infty} c_k z^k$ on $\mathbb{D}$ with the property that

$$\|f\|^2_{H^2} = \sum_{k=0}^{\infty} |c_k|^2 < \infty.$$

It is well known that H^2 is a reproducing kernel Hilbert space with

$$K(z,w) = \frac{1}{1 - z\overline{w}}.$$

For any $\alpha > -1$ we define

$$A^2_\alpha = L^2(\mathbb{D}, dA_\alpha) \cap H(\mathbb{D}),$$

where

$$dA_\alpha(z) = (\alpha+1)(1-|z|^2)^\alpha \, dA(z),$$

and dA is the normalized area measure on $\mathbb{D}$. Each A^2_α is a reproducing kernel Hilbert space with its inner product (and norm) inherited from $L^2(\mathbb{D}, dA_\alpha)$. The kernel of A^2_α is

$$K(z,w) = \frac{1}{(1 - z\overline{w})^{2+\alpha}}.$$

Given any positive integer n and any finite sequence $Z = \{z_1, \cdots, z_n\}$ of distinct points in $\mathbb{D}$, let H_Z denote the closed subspace of a reproducing kernel Hilbert space H spanned by the kernel functions K_{z_k}, $1 \le k \le n$, namely,

$$H_Z = \left\{ \sum_{k=1}^{n} c_k K(z, z_k) : c_k \in \mathbb{C}, 1 \le k \le n \right\}. \tag{1}$$

It is clear that $H_Z = H \ominus I_Z$, where I_Z is the space of all functions in H that vanish on Z.

If we allow repetition in Z, which is the case for most of our discussion below, we need to make some adjustments to the definitions of H_Z and I_Z. The necessary adjustment for the definition of I_Z is obvious: if a point λ appears m times in Z, then we require a function $f \in I_Z$ to vanish at λ to the order of m, that is,

$$f(\lambda) = f'(\lambda) = \cdots = f^{(m-1)}(\lambda) = 0.$$

The corresponding adjustment for H_Z is then as follows: if λ appears m times in Z, we then use the following m functions in the spanning set of H_Z:

$$\left. \frac{\partial^j}{\partial \overline{w}^j} K(z,w) \right|_{w=\lambda}, \qquad 0 \le j \le m-1. \tag{2}$$

These will be called generalized kernel functions for H. Consequently, if Z consists of n points from $\mathbb{D}$, counting multiplicity, then H_Z is still spanned by a total of n (generalized kernel) functions. With the understanding of this adjustment we will still denote H_Z by (1), although this is a slight abuse of notation.

We mention that multiple zeros and the corresponding generalized kernel functions in (2) arise naturally in function theory and optimization problems. See [6,11,19] for example. For those who are not familiar enough with the relationship between multiple zeros and the generalized kernel functions in (2), we illustrate it with A^2_α as an example. Given any $f \in A^2_\alpha$, we have

$$f(w) = \int_{\mathbb{D}} K(w,z) f(z)\, dA_\alpha(z) = \int_{\mathbb{D}} f(z) \overline{K(z,w)}\, dA_\alpha(z).$$

Differentiating under the integral sign, we obtain

$$f^{(j)}(w) = \int_{\mathbb{D}} f(z) \overline{\left[\frac{\partial^j K(z,w)}{\partial \overline{w}^j} \right]} \, dA_\alpha(z).$$

Therefore, $f^{(j)}(\lambda) = 0$ if and only if f is orthogonal to the following function in A^2_α:

$$z \mapsto \left. \frac{\partial^j K(z,w)}{\partial \overline{w}^j} \right|_{w=\lambda}.$$

The best kernel approximation problem for H is the following: given any positive integer n and any function $f \in H$ find the best possible approximation of f in norm by a linear combination of no more than n kernel functions. Thus, we consider the extremal problem

$$\inf \left\{ \left\| f - \sum_{k=1}^{n} c_k K_{z_k} \right\| : c_k \in \mathbb{C}, \ z_k \in \mathbb{D}, \ 1 \le k \le n \right\}.$$

If $Z = \{z_1, \cdots, z_n\}$ is fixed and we first take the infimum over all coefficients $\{c_k\}$, then by elementary functional analysis,

$$\inf \left\{ \left\| f - \sum_{k=1}^{n} c_k K_{z_k} \right\| : c_k \in \mathbb{C}, \ 1 \le k \le n \right\} = \|f - P_Z f\|,$$

where $P_Z : H \to H_Z$ is the orthogonal projection. Therefore, the best kernel approximation problem can be stated as

$$\inf\{\|f - P_Z f\| : Z = \{z_1, \cdots, z_n\} \subset \mathbb{D}\}. \tag{3}$$

Since

$$\|f\|^2 = \|P_Z f\|^2 + \|(I - P_Z) f\|^2,$$

the extremal problem in (3) is equivalent to the following: Given any positive integer n and any $f \in H$ find

$$\sup\{\|P_Z f\| : Z = \{z_1, \cdots, z_n\} \subset \mathbb{D}\}. \tag{4}$$

Our main concern is whether or not the extremal values above are actually attained by a certain sequence Z^* of n points in $\mathbb{D}$. The difficulty of the problem stems from the fact that $\mathbb{D}^n$ is non-compact.

The best kernel approximation problem for general reproducing kernel Hilbert spaces arises from system identification of stable systems. In particular, the weighted Bergman spaces A^2_α appear as special cases [2,3,20]. From the respective reproducing kernel formulas earlier, it is clear that the Hardy space H^2 can be thought of as the limit case of the weighted Bergman spaces A^2_α as $\alpha \to -1^+$. The best kernel approximation for H^2 corresponds to stable linear time-invariant system identification, whose identification has been studied by numerous authors (see, for instance, [12,23]). The existence of extremal sets in (3) and (4) for H^2, which is also equivalent to the problem of best n-degree rational function approximations, was established in [4,16,17,22].

The purpose of this paper is to prove the existence of extremal sets in (3) or (4) for the Bergman spaces A^2_α. Our main result is the following:

Main Theorem. Suppose $\alpha > -1$, n is a positive integer, and $f \in A^2_\alpha$. Then there exists a sequence Z^* of n points in $\mathbb{D}$ (with possible multiplicities) such that the infimum in (3) and the supremum in (4) are attained at Z^*.

We will see that the extremal set Z^* is not unique in general. Practical algorithms to find the extremal sets Z^* will not be considered in this paper. To the best of our knowledge, no mathematical algorithm has been found to locate the extremal sets Z^* even in the case of the Hardy space H^2. There are some related algorithms in engineering applications, but they can only find local extremes [4,16,18].

The best kernel approximation problem for the weighted Bergman spaces is much more complicated than that for the Hardy space. This is because the solution of the problem depends on a very careful analysis of the reproducing kernels of the spaces H_Z and I_Z. The structure of those kernel functions in the case of H^2 is relatively simple, while it is well known that the situation is completely different in the case of (weighted) Bergman spaces. In fact, our proof of the main result relies on several new estimates for the kernel functions of zero-based invariant subspaces of A^2_α, which are of some independent interest.

Unlike the Hardy spaces and in spite of several major breakthroughs in the past few decades (including [1,8,21] for example), the theory of Bergman spaces is far from mature. We refer the reader to [7,11] for an introduction to and more information about the spaces A^2_α and their zero-based invariant subspaces.

2. Preliminaries on reproducing kernel Hilbert spaces

In this section, we collect some preliminary material on reproducing kernel Hilbert spaces. In particular, we will need a result from [28], which is not readily available to people outside of Hong Kong. Thus, we include it here for convenience of the reader.

Throughout the paper we assume that any reproducing kernel Hilbert space H satisfies the following additional assumptions: $K(a,a) > 0$ for every $a \in \mathbb{D}$, and for every finite sequence $\{z_1, \cdots, z_n\}$ of distinct points in $\mathbb{D}$ the functions $K(z,z_1), \cdots, K(z,z_n)$ are linearly independent in H. These assumptions are obviously satisfied by the Hardy space H^2 and the weighted Bergman spaces A^2_α.

Recall that for a sequence Z of points in $\mathbb{D}$ we use I_Z to denote the space of all functions $f \in H$ such that $f|_Z = 0$ and we write $H_Z = H \ominus I_Z$. Since I_Z and H_Z are closed subspaces of H, they are reproducing kernel Hilbert spaces themselves. We will use $K_Z(z,w)$ to denote the reproducing kernel of I_Z.

In this section, all sequences will consist of distinct points. So we will be working with the original reproducing kernel $K(z,w)$ of H instead of the generalized kernel functions in (2).

For each $k \geq 1$ let $Z_k = \{z_1, \cdots, z_k\}$. Then the reproducing kernels $K_{Z_k}(z,w)$ can be constructed inductively as follows:

$$K_{Z_{k+1}}(z,w) = K_{Z_k}(z,w) - \frac{K_{Z_k}(z,z_{k+1})K_{Z_k}(z_{k+1},w)}{K_{Z_k}(z_{k+1},z_{k+1})}, \quad k \geq 0, \tag{5}$$

where $Z_0 = \emptyset$ and $K_{Z_0}(z,w) = K(z,w)$ is the reproducing kernel of the full Hilbert space H. In fact, if we fix w and let h denote the function on the right-hand side of (5) (with z being the variable), then it is clear that $h \in I_{Z_{k+1}}$. Moreover, h has the reproducing property at w for functions $f \in I_{Z_{k+1}}$:

$$\begin{aligned} \langle f, h\rangle &= \langle f, K_{Z_k}(w)\rangle - \frac{K_{Z_k}(w,z_{n+1})}{K_{Z_k}(z_{n+1},z_{n+1})}\langle f, K_{Z_k}(z_{n+1})\rangle \\ &= f(w) - \frac{K_{Z_k}(w,z_{n+1})}{K_{Z_k}(z_{n+1},z_{n+1})} f(z_{k+1}) \\ &= f(w). \end{aligned}$$

By uniqueness, the function h must be the reproducing kernel of $I_{Z_{k+1}}$ at the point w.

Proposition 1. *With the above notation, if we apply the Gram-Schmidt orthogonalization process to the kernel functions*

$$K(z,z_1), K(z,z_2), \cdots, K(z,z_k), \cdots,$$

the result is the following orthonormal system:

$$\frac{K(z,z_1)}{\sqrt{K(z_1,z_1)}}, \frac{K_{Z_1}(z,z_2)}{\sqrt{K_{Z_1}(z_2,z_2)}}, \cdots, \frac{K_{Z_k}(z,z_{k+1})}{\sqrt{K_{Z_k}(z_{k+1},z_{k+1})}}, \cdots. \tag{6}$$

Proof. This was proved in [28]. Again, for convenience of the reader, we include a brief proof here (different from the one given in [28]).

The first function in (6) is obvious. The second function in (6) follows from the recursion formula in (5). Now assume that applying the Gram-Schmidt process to the functions $\{K(z,z_k) : 1 \leq k \leq n\}$ results in the orthonormal set

$$\left\{ e_k(z) = \frac{K_{Z_{k-1}}(z,z_k)}{\sqrt{K_{Z_{k-1}}(z_k,z_k)}} : 1 \leq k \leq n \right\}.$$

Then the last function e_{n+1} from the Gram-Schmidt process when applied to $\{K(z,z_k) : 1 \leq k \leq n+1\}$ is the normalization of the function

$$\begin{aligned} h(z) &= K(z,z_{n+1}) - \sum_{k=1}^{n} \langle K(z_{n+1}), e_k\rangle e_k(z) \\ &= K(z,z_{n+1}) - \sum_{k=1}^{n} e_k(z)\overline{e_k(z_{n+1})}. \end{aligned}$$

By assumption, $\{e_k : 1 \leq k \leq n\}$ is an orthonormal basis for H_{Z_n}, so the summation above represents the reproducing kernel of H_{Z_n} at z_{n+1}. Since $H \ominus H_{Z_n} = I_{Z_n}$, the function h is the reproducing kernel of I_{Z_n} at z_{n+1}, that is, $h(z) = K_{Z_n}(z,z_{n+1})$. Thus, a normalization of h yields

$$e_{n+1}(z) = \frac{K_{Z_n}(z,z_{n+1})}{\sqrt{K_{Z_n}(z_{n+1},z_{n+1})}}.$$

The desired result then follows from induction. $\square$

See [26] for additional information about these kernel functions $K_{Z_k}(z,w)$ and their application to interpolation problems for Bergman spaces.

Corollary 2. *With the above notation and $a \in \mathbb{D} \setminus Z$, if we apply the Gram-Schmidt process to the functions*

$$K(z,z_1), K(z,z_2), \cdots, K(z,z_k), K(z,a),$$

then the last unit vector

$$G_{Z_k}(z) = \frac{K_{Z_k}(z,a)}{\sqrt{K_{Z_k}(a,a)}}$$

is the (unique) solution to the following extremal problem:

$$\sup\{\mathrm{Re}\, f(a) : f \in I_{Z_k}, \|f\| \le 1\}. \tag{7}$$

Proof. It follows from the reproducing property

$$f(a) = \langle f, K_{Z_k}(a) \rangle, \qquad f \in I_{Z_k},$$

and the Cauchy-Schwarz inequality that the extremal problem in (7) has a unique solution and the solution is simply the normalized reproducing kernel of I_{Z_k} at the point a. The desired result then follows from Proposition 1. □

The extremal problem in (7) plays a critical role in the modern theory of analytic function spaces. In the case of the Hardy space, the solution of (7) is exactly the Blaschke product with Z being its zero set. In the case of the ordinary Bergman space A^2 ($\alpha = 0$), this problem was first studied by Hedenmalm in [8] and then by many authors since then. See [1,7,11] and references therein.

3. Weighted Bergman spaces

Although our main result is about weighted Bergman space A^2_α with standard weights, its proof relies on certain results about more general weighted Bergman spaces. In this section, we prove a kernel estimate for weighted Bergman spaces induced by finite Blaschke products. This estimate will then be used in the next section to obtain kernel estimates for zero-based invariant subspaces in A^2_α.

For any finite positive measure ω on $\mathbb{D}$ we write

$$A^2_\omega = L^2(\mathbb{D}, d\omega) \cap H(\mathbb{D})$$

for the weighted Bergman space induced by ω. Under very mild conditions on ω it is easy to show that A^2_ω is a reproducing kernel Hilbert space. The weighted area measures

$$d\omega(z) = dA_\alpha(z) = (\alpha+1)(1-|z|^2)^\alpha \, dA(z), \qquad \alpha > -1,$$

are called standard weights, and the resulting spaces are of course A^2_α. See [11,30] for an extensive study of the spaces A^2_α.

In this paper, in addition to the spaces A^2_α, we also need another family of weighted Bergman spaces. More specifically, for any Blaschke product B we consider the measure $d\omega(z) = |B(z)|^2 \, dA_\alpha(z)$ and use $A^2_{\alpha,B}$ to denote the corresponding weighted Bergman space. For $f \in A^2_{\alpha,B}$ we will write

$$\|f\|^2_{A^2_{\alpha,B}} = \int_{\mathbb{D}} |f(z)|^2 |B(z)|^2 \, dA_\alpha(z).$$

We begin with the following identification of $A^2_{\alpha,B}$ as a set.

Lemma 3. *For any finite Blaschke product B of n zeros we have $A^2_{\alpha,B} = A^2_\alpha$. Furthermore,*

$$\frac{1}{(\alpha+2)^{n/2}} \|f\|_{A^2_\alpha} \le \|f\|_{A^2_{\alpha,B}} \le \|f\|_{A^2_\alpha}$$

for all $f \in H(\mathbb{D})$. The constant above is best possible.

Proof. The inclusion $A^2_\alpha \subset A^2_{\alpha,B}$ and the inequality $\|f\|_{A^2_{\alpha,B}} \le \|f\|_{A^2_\alpha}$ is trivial, because $|B(z)| \le 1$ on $\mathbb{D}$.

To prove the other direction, choose a positive constant c such that

$$\int_{\mathbb{D}} |f(z)|^2 |z|^2 \, dA_\alpha(z) \ge c \int_{\mathbb{D}} |f(z)|^2 \, dA_\alpha(z)$$

for all $f \in H(\mathbb{D})$. See Lemma 4.26 of [30] for the existence of such a constant. Computing with the Taylor expansion $f(z) = \sum_{k=0}^\infty a_k z^k$, we can actually find the best such constant. More specifically, we have

$$\|f\|^2_{A^2_\alpha} = \sum_{k=0}^\infty |a_k|^2 \int_{\mathbb{D}} |z|^{2k} \, dA_\alpha(z) = (\alpha+1) \sum_{k=0}^\infty |a_k|^2 B(k+1, \alpha+1),$$

where $B(x, y)$ is the beta function. Similarly,

$$\|zf\|^2_{A^2_\alpha} = (\alpha+1)\sum_{k=0}^{\infty} |a_k|^2 B(k+2, \alpha+1).$$

Thus, we just need to find a constant $c > 0$ such that

$$B(k+2, \alpha+1)/B(k+1, \alpha+1) \geq c$$

for all $k \geq 0$. Since

$$\frac{B(k+2, \alpha+1)}{B(k+1, \alpha+1)} = \frac{k+1}{k+\alpha+2}$$

is an increasing function of k, we just need to set $k = 0$ to get the constant $c = 1/(\alpha+2)$.

Next, for any $a \in \mathbb{D}$ we consider the Möbius map $\varphi_a : \mathbb{D} \to \mathbb{D}$ defined by

$$\varphi_a(z) = \frac{a-z}{1-\bar{a}z}, \qquad z \in \mathbb{D}.$$

By the change of variables $w = \varphi_a(z)$, which is the same as $z = \varphi_a(w)$, we have

$$\begin{aligned}
\|f\varphi_a\|^2_{A^2_\alpha} &= \int_{\mathbb{D}} |f(z)|^2 |\varphi_a(z)|^2 \, dA_\alpha(z) \\
&= \int_{\mathbb{D}} \left| f(\varphi_a(z)) \frac{(1-|a|^2)^{(2+\alpha)/2}}{(1-\bar{a}z)^{2+\alpha}} \right|^2 |z|^2 \, dA_\alpha(z) \\
&\geq \frac{1}{\alpha+2} \int_{\mathbb{D}} \left| f(\varphi_a(z)) \frac{(1-|a|^2)^{(2+\alpha)/2}}{(1-\bar{a}z)^{2+\alpha}} \right|^2 dA_\alpha(z) \\
&= \frac{1}{\alpha+2} \int_{\mathbb{D}} |f(z)|^2 \, dA_\alpha(z) = \frac{1}{\alpha+2} \|f\|^2_{A^2_\alpha}
\end{aligned}$$

for all $f \in H(\mathbb{D})$. The desired result then follows from applying the estimate above n times. □

Lemma 4. *Suppose H is any reproducing kernel Hilbert space of analytic functions on $\mathbb{D}$ and $K(z, w)$ is its reproducing kernel. Then*

$$K(a, a) = \sup\left\{ |f(a)|^2 : f \in H, \|f\| \leq 1 \right\}$$

for every $a \in \mathbb{D}$.

Proof. If $f(a) = 0$ for every $f \in H$, then $K(a, a) = \|K_a\|^2 = 0$, as $K_a(z) = \overline{K_z(a)} = 0$ for every z. In this case, the desired result is obvious.

Suppose $a \in \mathbb{D}$ is a point such that $K(a, a) > 0$. For any vector f in H with $\|f\| \leq 1$ we have

$$|f(a)|^2 = |\langle f, K_a \rangle|^2 \leq \|f\|^2 \|K_a\|^2 \leq K(a, a)$$

by the Cauchy-Schwarz inequality. Furthermore, equality is achieved by the normalized reproducing kernel

$$f(z) = K_a(z)/\|K_a\| = K(z, a)/\sqrt{K(a, a)}.$$

This proves the desired result. □

The simple result above is certainly well known to experts. We included it here for convenience of reference. Lemma 4 plays a crucial role for our next estimate.

Theorem 5. *Let B be a Blaschke product of n zeros. Then $A^2_{\alpha,B}$ is a reproducing kernel Hilbert space and its reproducing kernel $K_B(z, w)$ satisfies*

$$\frac{1}{(1-|a|^2)^{2+\alpha}} \leq K_B(a, a) \leq \frac{(\alpha+2)^n}{(1-|a|^2)^{2+\alpha}}$$

for all $a \in \mathbb{D}$.

Proof. Denote by $K(z, w)$ the reproducing kernel of A^2_α. By the last two lemmas, we have

$$\begin{aligned}
K(a, a) &= \sup\left\{ |f(a)|^2 : f \in A^2_\alpha, \|f\|_{A^2_\alpha} \leq 1 \right\} \\
&\leq \sup\left\{ |f(a)|^2 : f \in A^2_{\alpha,B}, \|f\|_{A^2_{\alpha,B}} \leq 1 \right\} \\
&= K_B(a, a).
\end{aligned}$$

Similarly, we have

$$\begin{aligned} K_B(a,a) &= \sup\left\{|f(a)|^2 : f \in A^2_{\alpha,B},\ \|f\|_{A^2_{\alpha,B}} \le 1\right\} \\ &\le \sup\left\{|f(a)|^2 : f \in A^2_{\alpha},\ \|f\|_{A^2_{\alpha}} \le (\alpha+2)^{n/2}\right\} \\ &= (\alpha+2)^n \sup\left\{|f(a)|^2 : f \in A^2_{\alpha},\ \|f\|_{A^2_{\alpha}} \le 1\right\} \\ &= (\alpha+2)^n K(a,a). \end{aligned}$$

Since

$$K(a,a) = \frac{1}{(1-|a|^2)^{2+\alpha}},$$

the desired result follows. □

The above proofs work for weight functions more general than finite Blaschke products. For example, if B is an infinite Blaschke product whose zero set is the union of finitely many interpolating sequences, then there exists a positive constant c such that

$$\int_{\mathbb{D}} |f(z)|^2 |B(z)|^2\, dA_\alpha(z) \ge c \int_{\mathbb{D}} |f(z)|^2\, dA_\alpha(z), \qquad f \in H(\mathbb{D}).$$

See [5,27]. Thus, Lemma 3 and Theorem 5 are still valid, except that the involved constant c may depend on the location of the zeros of B.

4. Kernel estimates for zero-based invariant subspaces

In this section, we use the kernel estimates for weighted Bergman spaces from the previous section to obtain sharp kernel estimates for zero-based invariant subspaces of A^2_α. Recall that for any A^2_α-zero set Z we use I_Z to denote the space of all functions in A^2_α that vanish on Z, counting multiplicity. We also write $H_Z = A^2_\alpha \ominus I_Z$. The closed subspace I_Z is often called a zero-based invariant subspace, because it is invariant under the action of the multiplication operator M_z (a weighted shift operator). See [11] for an introduction to the theory of invariant subspaces of A^2_α.

Proposition 6. *Suppose $Z = \{z_1, \cdots, z_n, \cdots\}$ is a sequence of distinct points in $\mathbb{D}$. If we perform the Gram-Schmidt orthogonalization process to the kernel functions $\{K(z,z_1), \cdots, K(z,z_n), \cdots\}$ of A^2_α, then we obtain an orthonormal system $\{e_1, \cdots, e_n, \cdots\}$ in A^2_α with the following properties:*

(i) *For each $k \ge 1$, the function e_k depends only on $z_1, \cdots z_k$.*
(ii) *For each $k \ge 1$, the vectors in $\{e_1, \cdots, e_k\}$ form an orthonormal basis for H_{Z_k}, where $Z_k = \{z_1, \cdots, z_k\}$.*
(iii) *For each $k \ge 1$ and any fixed $a \in \mathbb{D}$, the kernel function $K_{Z_k}(z,a)$ is a meromorphic function with poles outside the closed unit disk. In particular, each $K_{Z_k}(z,a)$ is continuous up to the boundary.*

Proof. The assertions clearly follow from Proposition 1 and the recursion formula in (5). □

Proposition 7. *Suppose Z is a Blaschke sequence in $\mathbb{D}$ and B is the corresponding Blaschke product. Then the reproducing kernel K_Z for $I_Z \subset A^2_\alpha$ and the reproducing kernel K_B for $A^2_{\alpha,B}$ are related by*

$$K_Z(z,w) = B(z)K_B(z,w)\overline{B(w)}.$$

Proof. Consider the functions

$$K^\dagger_w(z) =: K^\dagger(z,w) =: B(z)K_B(z,w)\overline{B(w)}.$$

For any $a \notin Z$ and $f \in I_Z$ we have

$$\begin{aligned} \langle f, K^\dagger_a\rangle_{A^2_\alpha} &= \int_{\mathbb{D}} f(z)\overline{K^\dagger(z,a)}\, dA_\alpha(z) \\ &= B(a)\int_{\mathbb{D}} f(z)K_B(a,z)\overline{B(z)}\, dA_\alpha(z) \\ &= B(a)\int_{\mathbb{D}} \frac{f(z)}{B(z)} K_B(a,z)|B(z)|^2\, dA_\alpha(z) \\ &= B(a)[f(a)/B(a)] = f(a). \end{aligned}$$

By the uniqueness of the reproducing kernel of I_Z at a, we must have $K^\dagger_a(z) = K_Z(z,a)$. A limit argument then shows that $K^\dagger(z,a) = K_Z(z,a)$ for all z and a in $\mathbb{D}$. □

Corollary 8. *Let $Z = \{z_1, \cdots, z_n\}$ be a Blaschke sequence in $\mathbb{D}$ and B be the correpsonding Blaschke product. Then for any $a \in \mathbb{D}\setminus Z$ the unique solution G to the extremal problem*

$$\sup\{\mathrm{Re}\, f(a) : f \in I_Z,\ \|f\|_{A^2_\alpha} \le 1\}$$

is given by

$$G(z) = K_B(a,a)^{-\frac{1}{2}} B(z) K_B(z,a) \frac{\overline{B(a)}}{|B(a)|},$$

where $K_B(z,w)$ is the reproducing kernel of the weighted Bergman space $A^2_{\alpha,B}$.

Proof. It is easy to see that the extremal function G is exactly the normalized reproducing kernel of I_Z at a. The desired result then follows from Proposition 7. □

Corollary 9. *Let n be any positive integer. Then*

$$\frac{|B(a)|^2}{(1-|a|^2)^{2+\alpha}} \le K_Z(a,a) \le (\alpha+2)^n \frac{|B(a)|^2}{(1-|a|^2)^{2+\alpha}}$$

for every set $Z \subset \mathbb{D}$ of n points, where B is the Blaschke product corresponding to Z.

Proof. This follows from Theorem 5 and Proposition 7. □

We will also prove an off-diagonal estimate for the reproducing kernel functions $K_Z(z,w)$ of I_Z, where Z is a finite sequence in $\mathbb{D}$. We need the following elementary estimate first.

Lemma 10. *We have the inequality*

$$\frac{|1-z\bar{a}|(1-|b|^2)}{|1-z\bar{b}||1-a\bar{b}|} \le 2$$

for all $a,b,z \in \mathbb{D}$. The constant 2 is best possible.

Proof. This follows easily from Lemma 1.3 of [29]. □

Theorem 11. *For any positive integer n there exists a positive constant $C = C(n,\alpha)$ such that*

$$|K_Z(z,w)| \le C \frac{|B(z)B(w)|}{|1-z\overline{w}|^{2+\alpha}} \tag{8}$$

for all $Z = \{z_1, \cdots, z_n\} \subset \mathbb{D}$ and $z,w \in \mathbb{D}$, where B is the Blaschke product corresponding to Z.

Proof. By a standard limit argument we may assume that the points in Z are distinct. Under this extra assumption, we prove the estimate in (8) by induction on n.

When $n = 1$, we write $Z = \{z_1\}$ and let $K(z,w) = (1-z\overline{w})^{-(2+\alpha)}$ be the reproducing kernel for A^2_α. It follows from the recursive formula in (5) that

$$\begin{aligned} K_Z(z,w) &= K(z,w) - \frac{K(z,z_1)K(z_1,w)}{K(z_1,z_1)} \\ &= \frac{1}{(1-z\overline{w})^{2+\alpha}} - \frac{(1-|z_1|^2)^{2+\alpha}}{(1-z\bar{z}_1)^{2+\alpha}(1-z_1\overline{w})^{2+\alpha}} \\ &= \frac{1}{(1-z\overline{w})^{2+\alpha}} \left[1 - \frac{(1-z\overline{w})^{2+\alpha}(1-|z_1|^2)^{2+\alpha}}{(1-z\bar{z}_1)^{2+\alpha}(1-z_1\overline{w})^{2+\alpha}} \right]. \end{aligned}$$

It follows from Lemma 10 that for the constant $C_1 = 1 + 2^{2+\alpha}$ we have

$$|K_Z(z,w)| \le \frac{C_1}{|1-z\overline{w}|^{2+\alpha}}, \qquad z,w \in \mathbb{D}.$$

Since $K_Z(z_1,w) = 0$ and $K_Z(z,z_1) = 0$ for any fixed z or w in $\mathbb{D}$, an application of the maximum modulus principle twice then gives

$$|K_Z(z,w)| \le \frac{C_1|B(z)B(w)|}{|1-z\overline{w}|^{2+\alpha}}, \qquad z,w \in \mathbb{D}.$$

This proves the desired result for $n = 1$.

Now suppose that for some positive integer n there exists a positive constant $C_n = C(n,\alpha)$ such that (8) holds. We then consider

$$\widetilde{Z} = \{z_1, \cdots, z_n, z_{n+1}\} = \{z_1, \cdots, z_n\} \cup \{z_{n+1}\} =: Z \cup \{z_{n+1}\}.$$

Let B and $\widetilde{B}$ be the Blaschke products corresponding to Z and $\widetilde{Z}$, respectively. By the recursion formula in (5), we have

$$K_{\widetilde{Z}}(z,w) = K_Z(z,w) - \frac{K_Z(z,z_{n+1})K_Z(z_{n+1},w)}{K_Z(z_{n+1},z_{n+1})}.$$

This together with the induction hypothesis and Corollary 9 gives

$$\begin{aligned}
&|K_{\widetilde{Z}}(z,w)| \\
&\le C_n \frac{|B(z)B(w)|}{|1-z\overline{w}|^{2+\alpha}} + \frac{C_n^2 |B(z)B(z_{n+1})B(z_{n+1})B(w)|}{|1-z\overline{z}_{n+1}|^{2+\alpha}|1-z_{n+1}\overline{w}|^{2+\alpha}K_Z(z_{n+1},z_{n+1})} \\
&= C_n \frac{|B(z)B(w)|}{|1-z\overline{w}|^{2+\alpha}}\left[1+\frac{C_n|B(z_{n+1})|^2|1-z\overline{w}|^{2+\alpha}}{|1-z\overline{z}_{n+1}|^{2+\alpha}|1-z_{n+1}\overline{w}|^{2+\alpha}K_Z(z_{n+1},z_{n+1})}\right] \\
&\le C_n \frac{|B(z)B(w)|}{|1-z\overline{w}|^{2+\alpha}}\left[1+\frac{C_n|1-z\overline{w}|^{2+\alpha}(1-|z_{n+1}|^2)^{2+\alpha}}{|1-z\overline{z}_{n+1}|^{2+\alpha}|1-z_{n+1}\overline{w}|^{2+\alpha}}\right]
\end{aligned}$$

Another application of Lemma 10 yields the positive constant

$$C_{n+1} = C_n(1+2^{2+\alpha}C_n)$$

such that

$$|K_{\widetilde{Z}}(z,w)| \le C_{n+1}\frac{|B(z)B(w)|}{|1-z\overline{w}|^{2+\alpha}}.$$

Since $K_{\widetilde{Z}}(z_{n+1},w) = K_{\widetilde{Z}}(z,z_{n+1}) = 0$, the maximum modulus principle again implies that

$$|K_{\widetilde{Z}}(z,w)| \le C_{n+1}\frac{|B(z)B(w)||\varphi_{z_{n+1}}(z)\varphi_{z_{n+1}}(w)|}{|1-z\overline{w}|^{2+\alpha}},$$

where $\varphi_{z_{n+1}}(z)$ is the Blaschke factor corresponding to the single zero z_{n+1}. It is clear that $|B(z)\varphi_{z_{n+1}}(z)| = |\widetilde{B}(z)$. Therefore, we have

$$|K_{\widetilde{Z}}(z,w)| \le C_{n+1}\frac{|\widetilde{B}(z)\widetilde{B}(w)|}{|1-z\overline{w}|^{2+\alpha}}.$$

This shows that the desired result holds for $n+1$ as well. The proof is then complete by induction. □

Corollary 12. *For any positive integer n there exists a positive constant $C = C(n,\alpha)$ such that*

$$|K_Z(a,w)| \le \frac{C\,|B(a)B(w)|}{(1-|a|^2)^{2+\alpha}}$$

for all $Z = \{z_1,\cdots,z_n\} \subset \mathbb{D}$ and all $w,a \in \mathbb{D}$.

Proof. This follows from Theorem 11 and the elementary inequalities

$$|1-a\overline{w}| \ge 1-|a| \ge \frac{1}{2}(1-|a|^2).$$

□

Corollary 13. *For any positive integer n there exists a positive constant $C = C(n,\alpha)$ such that for every Blaschke product B of n zeros we have*

$$|K_B(z,w)| \le \frac{C}{|1-z\overline{w}|^{2+\alpha}}$$

for all $z,w \in \mathbb{D}$. Consequently, there exisists another positive constant $C = C(n,\alpha)$ such that

$$|K_B(z,a)| \le \frac{C}{(1-|a|^2)^{2+\alpha}}$$

for all $z,a \in \mathbb{D}$.

Note that our kernel estimates in this section are in the same spirit as the results obtained by Hedenmalm in [9]. However, the main theorem in [9] does not cover our results, because our measure $|B(z)|^2\,dA_\alpha(z)$ does not satisfy the assumptions for weight functions in [9]. In the unweighted case $\alpha = 0$, if Z is a finite sequence in $\mathbb{D}$ (not containing the origin) and G is the unique solution to the extremal problem

$$\sup\{\operatorname{Re} f(0) : f \in I_Z, \|f\| \le 1\},$$

then Hedenmalm's results in [9] are valid for the weighted Bergman space corresponding to the measure $|G(z)|^2\,dA(z)$. When Z is fixed, the function $|G(z)/B(z)|$ is bounded from both below and above on $\mathbb{D}$. But the structure of the function G is very complicated, and it is not clear how to control the lower and upper bounds of $|G(z)/B(z)|$. Potentially, those bounds could depend on the location of the zeros in Z.

The main point of this section is that we are able to prove kernel estimates with constants that only depend on the number of zeros but not on the location of the zeros. These estimates are of independent interest and they will be critical for our analysis in the next section.

Our results here hold for all weighted Bergman spaces A^2_α, $\alpha > -1$. It is well known that many techniques and ideas based on the extremal function for invariant subspaces and the bi-harmonic Green function (such as those used in [7,9,11]) do not work for A^2_α in general. One such example can be found in [10].

Finally in this section we mention that it is impossible to obtain off-diagonal lower estimates like

$$|K_B(z,w)| \geq \frac{c}{|1-z\overline{w}|^{2+\alpha}},$$

or equivalently,

$$|K_Z(z,w)| \geq \frac{c|B(z)B(w)|}{|1-z\overline{w}|^{2+\alpha}},$$

where c is a positive constant. It was shown in [10] that, for any fixed w ($w = 0$ for example), the kernel function $z \mapsto K_Z(z,w)$ can have more zeros than those in Z when α is large. This clearly shows that the off-diagonal lower estimates above are not possible for general $\alpha > -1$, although they do hold on the diagonal!

5. Best kernel approximation in A^2_α

In this section, we consider the best kernel approximation problem for weighted Bergman spaces with standard weights. We will focus on the equivalent extremal problem in (4) and prove that there is always an extremal set Z^* inside $\mathbb{D}$. All norms and inner products appearing in this section are those in A^2_α.

To see that the problem is non-trivial and also to motivate our discussions below, we first consider the case when $n = 1$, so $Z = \{a\}$ is a singleton. In this case, the space H_Z is one-dimensional, consisting of all constant multiples of the normalized reproducing kernel

$$k_a(z) = \frac{(1-|a|^2)^{(2+\alpha)/2}}{(1-z\overline{a})^{2+\alpha}}.$$

Thus,

$$P_Z f = \langle f, k_a\rangle k_a = (1-|a|^2)^{(2+\alpha)/2} f(a) k_a,$$

so the extremal problem in (4) becomes

$$\sup_{a\in\mathbb{D}} (1-|a|^2)^{(2+\alpha)/2}|f(a)|, \tag{9}$$

where $f \in A^2_\alpha$ is fixed. The existence of an extremal set $Z^* = \{a^*\}$ inside the non-compact set $\mathbb{D}$ follows from the fact that the non-negative function in (9) is continuous on $\mathbb{D}$ and tends to 0 as $|a| \to 1^-$, which is well-known and follows, for example, from the fact that $k_a \to 0$ weakly in A^2_α as $|a| \to 1^-$. See [11,30].

The arguments above also show that, even in the case when $n = 1$, the best kernel approximation is not unique in general, because the function in (9) can attain its maximum value at more than one interior points. For example, if $f(z) = z^n$ for some positive integer n, then the function in (9) attains it maximum value at infinitely many points (a whole circle centered at the origin) in $\mathbb{D}$.

We can now prove the main result of the paper.

Theorem 14. *Suppose n is a positive integer and $f \in A^2_\alpha$. Then there exists a set $Z^* = \{z_1^*, \cdots, z_n^*\} \subset \mathbb{D}$ (with possible multiplicities) such that*

$$\|P_{Z^*} f\| = \sup\{\|P_Z f\| : Z = \{z_1, \cdots, z_n\} \subset \mathbb{D}\}.$$

Proof. We prove the desired result by induction on n, the number of points in Z (counting multiplicity). Recall that the case $n = 1$ follows from the fact that the normalized reproducing kernels k_a in A^2_α converge weakly to 0 as $|a| \to 1^-$.

For any $m \geq 1$ we will consider the function

$$F_m(Z) = \|P_Z f\|^2, \qquad Z = (z_1, \cdots, z_m) \in \mathbb{D}^m.$$

So here and below we will think of Z as a point in $\mathbb{D}^m$ instead of a set in $\mathbb{D}$. It is easy to see that each F_m is continuous on $\mathbb{D}^m$ (see Proposition 6) and $|F_m(Z)| \leq \|f\|^2$ for all $m \geq 1$ and all $Z \in \mathbb{D}^m$. We will assume that for some positive integer n the function $F_n(Z)$ attains its maximal value inside $\mathbb{D}^n$ and proceed to show the same for $F_{n+1}(Z)$ on $\mathbb{D}^{n+1}$.

Choose a sequence $\{Z^{(k)}\}$ in $\mathbb{D}^{n+1}$,

$$Z^{(k)} = \{z_1^{(k)}, \cdots, z_n^{(k)}, z_{n+1}^{(k)}\}, \qquad k \geq 1,$$

such that

$$\lim_{k\to\infty} F_{n+1}(Z^{(k)}) = \sup\{F_{n+1}(Z) : Z \in \mathbb{D}^{n+1}\} =: M_{n+1}. \tag{10}$$

By the continuity of F_{n+1} on $\mathbb{D}^{n+1}$ we may assume that the $n+1$ coordinates in each $Z^{(k)}$ are distinct. Going down to a subsequence if necessary, we may also assume that

$$Z^{(k)} \to Z^* = (z_1^*, \cdots, z_n^*, z_{n+1}^*) \in \overline{\mathbb{D}}^{n+1}$$

as $k \to \infty$.

If $Z^* \in \mathbb{D}^{n+1}$, we are done. If Z^* is not contained in $\mathbb{D}^{n+1}$, then there exists at least one z_k^* that is on the unit circle. The function $F_{n+1}(Z)$ is clearly symmetric in the sense that the values $F_{n+1}(Z)$ remain the same when we permute the coordinates of Z. Thus, we may as well assume that $z_{n+1}^* \in \partial\mathbb{D}$.

For each $k \geq 1$ we let

$$\{e_1^{(k)}, \cdots, e_n^{(k)}, e_{n+1}^{(k)}\}$$

be the orthonormal basis for $H_{Z^{(k)}}$ generated by applying the Gram-Schmidt process to the following A_α^2-kernel functions:

$$K(z, z_1^{(k)}), \quad \cdots, \quad K(z, z_n^{(k)}), \quad K(z, z_{n+1}^{(k)}).$$

Recall from Proposition 6 that for each $1 \leq j \leq n$ the set $\{e_1^{(k)}, \cdots, e_j^{(k)}\}$ is also an orthonormal basis for $H_{\{z_1^{(k)}, \cdots, z_j^{(k)}\}}$. We can write

$$P_{Z^{(k)}} f = \sum_{j=1}^{n+1} \langle f, e_j^{(k)} \rangle e_j^{(k)} = P_{W^{(k)}} f + \langle f, e_{n+1}^{(k)} \rangle e_{n+1}^{(k)},$$

where $W^{(k)} = (z_1^{(k)}, \cdots, z_n^{(k)}) \in \mathbb{D}^n$, and

$$F_{n+1}(Z^{(k)}) = \sum_{j=1}^{n+1} |\langle f, e_j^{(k)} \rangle|^2 = F_n(W^{(k)}) + |\langle f, e_{n+1}^{(k)} \rangle|^2. \tag{11}$$

We will show that the assumption $z_{n+1}^* \in \partial\mathbb{D}$ implies $\langle f, e_{n+1}^{(k)} \rangle \to 0$ as $k \to \infty$. Since the kernel functions $\{K_w : w \in \mathbb{D}\}$ span A_α^2, a simple approximation argument tells us that we can assume f is a finite linear combination of kernel functions in A_α^2, which allows us to further simply to the case where $f = K_a$ is the reproducing kernel of A_α^2 at some fixed $a \in \mathbb{D}$.

By Corollary 2, we have

$$e_{n+1}^{(k)}(z) = \frac{K_{W^{(k)}}(z, z_{n+1}^{(k)})}{\sqrt{K_{W^{(k)}}(z_{n+1}^{(k)}, z_{n+1}^{(k)})}},$$

where $K_{W^{(k)}}(z, w)$ is the reproducing kernel of $I_{W^{(k)}}$, the zero-based invariant subspace corresponding to the zeros in $W^{(k)}$. It follows from the reproducing property of K_a in A_α^2 that, for $f = K_a$, we have

$$|\langle f, e_{n+1}^{(k)} \rangle| = \frac{|K_{W^{(k)}}(a, z_{n+1}^{(k)})|}{K_{W^{(k)}}(z_{n+1}^{(k)}, z_{n+1}^{(k)})^{1/2}}.$$

Since a is fixed and $|z_{n+1}^{(k)}| \to 1$ as $k \to \infty$, it follows from Corollary 12 and the first inequality in Corollary 9 that

$$\lim_{k \to \infty} |\langle f, e_{n+1}^{(k)} \rangle|^2 = 0, \tag{12}$$

and the convergence here is uniform with respect to the location of the coordinates of $Z^{(k)}$.

By the induction hypothesis, there exists some $W^* = (w_1^*, \cdots, w_n^*) \in \mathbb{D}^n$ such that

$$M_n =: \sup\{\|P_W f\|^2 : W \in \mathbb{D}^n\} = \|P_{W^*} f\|^2.$$

It follows from (10), (11), and (12) that $M_{n+1} \leq \|P_{W^*} f\|^2$. On the other hand, it is clear that adding an extra coordinate to W^* will expand the space H_{W^*} and hence increase the norm of $\|P_{W^*} f\|$, so

$$\|P_{W^*} f\|^2 \leq \|P_{(w_1^*, \cdots, w_n^*, w_{n+1})} f\|^2 \leq M_{n+1}$$

for all $w_{n+1} \in \mathbb{D}$. This shows that

$$M_{n+1} = \|P_{W^*} f\| = \|P_{(w_1^*, \cdots, w_n^*, w_{n+1})} f\|^2$$

for any $w_{n+1} \in \mathbb{D}$, so the function $F_{n+1}(Z)$ attains its maximum value at $(w_1^*, \cdots, w_n^*, w_{n+1}) \in \mathbb{D}^{n+1}$ for any choice of $w_{n+1} \in \mathbb{D}$ (under the earlier assumption that $z_{n+1}^* \in \partial\mathbb{D}$). This completes the proof of the theorem by induction. □

Recall that when we write H_Z with $Z = \{z_1, \cdots, z_n\} \subset \mathbb{D}$, it is implicit that the points in Z can repeat. For example, for $Z = \left\{\frac{1}{2}, \frac{1}{3}, \frac{1}{2}\right\}$, we have

$$H_Z = \operatorname{span}\left\{K(z, \frac{1}{2}), \quad \left.\frac{\partial K(z, w)}{\partial \overline{w}}\right|_{w=\frac{1}{2}}, \quad K(z, \frac{1}{3})\right\}.$$

With this understanding, we can say that the infimum in (3) and the supremum in (4) are actually *attained*.

On the other hand, if we are only concerned the value of the infimum or supremum, we can use sequences Z of distinct points in $\mathbb{D}$. For example, for any $f \in A^2_\alpha$ and any positive integer n, it follows from the continuity of the function F_n in the above proof that

$$\delta(f, n) = \widetilde{\delta}(f, n),$$

where

$$\delta(f, n) = \inf\{\|f - P_Z f\| : Z = \{z_1, \cdots, z_n\} \subset \mathbb{D}\}$$

with possible repetition in Z, and

$$\widetilde{\delta}(f, n) = \inf\{\|f - P_Z f\| : Z = \{z_1, \cdots, z_n\} \subset \mathbb{D},\ z_i \neq z_j \text{ for } i \neq j\}.$$

The same remark applies to the supremum in (4). If, in (3) and (4), we use sequences Z of distinct points in $\mathbb{D}$, then technically we cannot say that the infimum and supremum are *attained*, because repetition can definitely occur in the extremal set Z^*.

We illustrate this with a simple example. Consider the function $f(z) = z$ and $n = 2$. In this case, we have $\delta(z, 2) = \widetilde{\delta}(z, 2) = 0$. The infimum in the definition of $\delta(z, 2)$ is attained by $Z^* = \{0, 0\}$, because $H_{Z^*} = \text{span}\,\{1, z\}$. On the other hand, it is easy to see that with $Z^{(k)} = \{0, 1/k\}$, we have

$$z = \lim_{k\to\infty} \frac{k}{2+\alpha}[K(z, 1/k) - K(z, 0)]$$

and the convergence is in norm. Thus, $\|f - P_{Z^{(k)}} f\| \to 0$ as $k \to \infty$, so $\widetilde{\delta}(z, 2)$ is indeed equal to 0. However, the infimum in the definition of $\widetilde{\delta}(z, 2)$ cannot be attained, because there is clearly no way we can write

$$z = \frac{c_1}{(1 - z\bar{z}_1)^{2+\alpha}} + \frac{c_2}{(1 - z\bar{z}_2)^{2+\alpha}}$$

for two distinct points z_1 and z_2 in $\mathbb{D}$.

6. Further remarks

The proof of Theorem 14 works for the Hardy space H^2 as well. In fact, the proof works for many other reproducing kernel Hilbert spaces up until (12). It is well known that the reproducing kernel for $I_Z \subset H^2$ is given by

$$K_Z(z, w) = \frac{B(z)\overline{B(w)}}{1 - z\overline{w}},$$

where B is the Blaschke product with zeros in Z. This immediately leads to (12) in the Hardy space case.

Recall that in the special case $n = 1$ the best kernel approximation problem for A^2_α, $\alpha \geq -1$, amounts to finding the maximum value of the function

$$F_1(a) = (1 - |a|^2)^{2+\alpha}|f(a)|^2$$

on $\mathbb{D}$. Here we think of the Hardy space H^2 as A^2_{-1}. The maximum value of F_1 is attained inside $\mathbb{D}$ because $F_1(a) \to 0$ as $|a| \to 1^-$. It is natural to wonder if for any $n > 1$ we might have a certain closed-form formula for the function $F_n(Z) = \|P_Z f\|^2_{A^2_\alpha}$ on $\mathbb{D}^n$ that we can use to solve the best kernel approximation problem.

For the Hardy space H^2 with $n = 2$, say $Z = \{z_1, z_2\}$ (assume $z_1 \neq z_2$ for the moment), we apply the Gram-Schmidt process to the two kernel functions $K(z, z_1)$ and $K(z, z_2)$ to obtain an orthonormal basis $\{e_1, e_2\}$ for H_Z, where

$$e_1(z) = \frac{K(z, z_1)}{\sqrt{K(z_1, z_1)}} = \frac{\sqrt{1 - |z_1|^2}}{1 - z\bar{z}_1},$$

and

$$e_2(z) = \frac{K_{Z_1}(z, z_2)}{\sqrt{K_{Z_1}(z_2, z_2)}}.$$

Recall that

$$K_{Z_1}(z, z_2) = K(z, z_2) - \frac{K(z, z_1)K(z_1, z_2)}{K(z_1, z_1)},$$

and consequently,

$$K_{Z_1}(z_2, z_2) = \frac{1}{1 - |z_2|^2} - \frac{1 - |z_1|^2}{|1 - z_1\bar{z}_2|^2} = \frac{|\varphi_{z_1}(z_2)|^2}{1 - |z_2|^2}. \tag{13}$$

It follows that

$$|\langle f, e_1\rangle|^2 = (1-|z_1|^2)|f(z_1)|^2,$$

and

$$|\langle f, e_2\rangle|^2 = \frac{1-|z_2|^2}{|\varphi_{z_1}(z_2)|^2}\left|f(z_2) - \frac{1-|z_1|^2}{1-z_1\bar{z}_2} f(z_1)\right|^2.$$

Therefore, after some simplification and with the new notation $\rho(z_1, z_2) = |\varphi_{z_1}(z_2)|$ in terms of the pseudo-hypobolic metric, we have

$$\begin{aligned} F_2(z_1, z_2) &= \|P_Z f\|^2 = |\langle f, e_1\rangle|^2 + |\langle f, e_2\rangle|^2 \\ &= \frac{1-|z_1|^2}{\rho(z_1, z_2)^2}|f(z_1)|^2 + \frac{1-|z_2|^2}{\rho(z_1, z_2)^2}|f(z_2)|^2 \\ &\qquad - \frac{2(1-|z_1|^2)(1-|z_2|^2)}{\rho(z_1, z_2)^2}\mathrm{Re}\,\frac{f(z_1)\overline{f(z_2)}}{1-z_1\bar{z}_2}, \end{aligned}$$

which is clearly a symmetric function of z_1 and z_2. Note that a different approach was used in [25] to tackle the best kernel approximation problem for the Hardy space H^2 using the pseudo-hyperbolic metric and TM systems. The present paper was motivated by [25].

If we fix $z_1 \in \mathbb{D}$ and let $z_2 \to z_1$, it can be shown directly that the function $F_2(z_1, z_2)$ above has a finite limit. Therefore, F_2 is a non-negative continuous function on $\mathbb{D}^2$. Also, for any fixed $z_1 \in \mathbb{D}$, it is clear that

$$\lim_{|z_2|\to 1^-} F_2(z_1, z_2) = (1-|z_1|^2)|f(z_1)|^2.$$

Therefore, the function $F_2(z_1, z_2)$ has limit values when one of the variables tends to the boundary while the other variable is held steady. But the behaviour of $F_2(z_1, z_2)$ is not that clear if both variables tend to the boundary at the same time without any control on them.

The example above shows the complexity of the best kernel approximation problem for $n > 1$. A similar calculation is possible for A^2_α, $\alpha > -1$, when $n = 2$. But the resulting function $F_2(z_1, z_2)$ is even more complicated. We were able to simplify $K_{Z_1}(z_2, z_2)$ in (13) for H^2, which is not possible for the weighted Bergman spaces A^2_α with $\alpha > -1$.

References

[1] A. Aleman, S. Richter, C. Sundberg, Beurling's theorem for the bergman space, Acta Math. 177 (1996) 275–310.
[2] J.A. Ball, V. Bolotnikov, Weighted bergman spaces: shift-invariant subspaces and input/state/output linear systems, Integ. Equ. Oper. Theory 76 (2013) 301–356.
[3] J.A. Ball, V. Bolotnikov, Weighted hardy spaces: shift invariant and coinvariant subspaces, linear systems and operator model theory, 2014. arXiv: 1405.2974.
[4] L. Baratchart, M. Cardelli, M. Olivi, Identification and rational l^2 approximation-a gradient algorithm, Automatica 27 (1991) 413–417.
[5] P. Bourdon, Similarity of parts to the whole for certain multiplication operators, Proc. Amer. Math. Soc. 99 (1987) 563–567.
[6] Q.-H. Chen, T. Qian, L.H. Tan, A theory on non-constant frequency decompositions and applications, in: Advancements in Complex Analysis, Springer, Cham, 2020, pp. 1–37.
[7] P. Duren, A. Schuster, Bergman spaces, Mathematical Surveys and Monographs, vol. 100, American Mathematical Society, 2004.
[8] H. Hedenmalm, A factorization theorem for square area-integrable analytic functions, J. Reine Angew. Math. 422 (1991) 45–68.
[9] H. Hedenmalm, An off-diagonal estimate of bergman kernels, J. Math. Pures Appl. 79 (2000) 163–172.
[10] H. Hedenmalm, K. Zhu, On the failure of optimal factorization for certain weighted bergman spaces, Complex Variables Theory Appl. 19 (1992) 165–176.
[11] B. Korenblum, H. Hedenmalm, K. Zhu, Theory of Bergman spaces, Graduate Texts in Mathematics, vol. 199, Springer-Verlag, New York, 2000.
[12] W. Mi, T. Qian, Frequency domain identification: an algorithm based on adaptive rational orthogonal system, Automatica 48 (2012) 1154–1162.
[13] B. Ninness, Frequency domain estimation using orthonormal bases, in: Proceedings of the 13th IFAC World Congress, 1996, pp. 381–386.
[14] B. Ninness, F. Gustafsson, A unifying construction of orthonormal bases for system identification, IEEE Trans. Automat. Control 42 (1997). 515–512
[15] B. Ninness, H. Hjalmarsson, F. Gustafsson, The fundamental role of general orthonormal bases in system identification, IEEE Trans. Automat. Control 44 (1997) 1384–1406.
[16] T. Qian, Cyclic AFD algorithm for best rational approximation, Math. Methods Appl. Sci. 37 (2014) 846–859.
[17] T. Qian, E. Wegert, Optimal approximation by Blaschke forms, Complex Var. Elliptic Equ. 58 (2013) 123–133.
[18] T. Qian, J.Z. Wang, W.X. Mai, An enhancement algorithm for cyclic adaptive fourier decomposition, Appl. Comput. Harmon. Anal. 47 (2019) 516–525.
[19] W. Qu, P. Dang, Rational approximation in a class of weighted hardy spaces, Complex Anal. Oper. Theory 13 (2019) 1827–1852.
[20] W. Qu, P. Dang, Reproducing kernel approximation in weighted bergman spaces: algorithm and applications, Math. Methods Appl. Sci. 42 (2019) 4292–4304.
[21] K. Seip, Beurling type density theorems in the unit disk, Invent. Math. 113 (1993) 21–39.
[22] J.L. Walsh, The existence of rational functions of best approximation, Ibid 33 (1931) 668–689.
[23] J.L. Walsh, Interpolation and Approximation by Rational Functions in the Complex Plane, American Mathematical Society, 1965.
[24] X.Y. Wang, T. Qian, L.M. Zhang, 2-D frequency-domain system identification, IEEE Trans. Autom. Control [0018-9286] 65 (2) (2020) 577–590.
[25] Y.-B. Wang, T. Qian, Pseudo-hyperbolic distance and n-best rational approximation in h^2 space, arXiv: 3510581 [Math. FA] (2020).
[26] K. Zhu, Interpolating sequences for the bergman space, Michigan Math. J. 41 (1994) 73–86.
[27] K. Zhu, Restriction of the bergman shift to an invariant subspace, Quart. J. Math. Oxford 48 (1997) 519–532.
[28] K. Zhu, Interpolating and recapturing in reproducing hilbert spaces, Bull. Hong Kong Math. Soc. 1 (1997) 21–33.
[29] K. Zhu, Spaces of Holomorphic Functions in the Unit Ball, Springer GTM, vol. 226, Springer-Verlag, New York, 2005.
[30] K. Zhu, Operator theory in function spaces, Mathematical Surveys and Monographs, vol. 138, American Mathematical Society, 2007.