axioms

Article

# Some Upper Bounds for RKHS Approximation by Bessel Functions

**Mingdang Tian [1], Baohuai Sheng [1,*] and Shuhua Wang [2]**

[1] Department of Economic Statistics, School of International Business, Zhejiang Yuexiu University, Shaoxing 312000, China; 20112016@zyufl.edu.cn
[2] School of Information Engineering, Jingdezhen Ceramic University, Jingdezhen 333403, China; w614sh@126.com
* Correspondence: bhsheng@usx.edu.cn or shengbaohuai@163.com

**Abstract:** A reproducing kernel Hilbert space (RKHS) approximation problem arising from learning theory is investigated. Some $K$-functionals and moduli of smoothness with respect to RKHSs are defined with Fourier–Bessel series and Fourier–Bessel transforms, respectively. Their equivalent relation is shown, with which the upper bound estimate for the best RKHS approximation is provided. The convergence rate is bounded with the defined modulus of smoothness, which shows that the RKHS approximation can attain the same approximation ability as that of the Fourier–Bessel series and Fourier–Bessel transform. In particular, it is shown that for a RKHS produced by the Bessel operator, the convergence rate sums up to the bound of a corresponding convolution operator approximation. The investigations show some new applications of Bessel functions. The results obtained can be used to bound the approximation error in learning theory.

**Keywords:** bessel function; Fourier–Bessel series; Fourier–Bessel transform; $K$-functional; modulus of smoothness; semigroup of operators; reproducing kernel Hilbert space (RKHS); best approximation error; learning theory



**Citation:** Tian, M.; Sheng, B.; Wang, S. Some Upper Bounds for RKHS Approximation by Bessel Functions. *Axioms* **2022**, *11*, 233. https://doi.org/10.3390/axioms11050233

Academic Editor: Stefano De Marchi

Received: 19 April 2022
Accepted: 9 May 2022
Published: 17 May 2022



## 1. Introduction

The error analysis in learning theory shows that the learning rate of the kernel regularized regression depends upon the approximation ability of the kernel function spaces (see, for example, [1–3]).

Let $X$ be a complete metric space and $\mu$ be a Borel measure on $X$. Denoted by $L^2_\mu(X)$, the Hilbert space consisting of (real) square integrable functions with the inner product

$$\langle f,\ g\rangle_{L^2_\mu(X)} = \int_X f(x)g(x)\, d\mu(x), \qquad f,g \in L^2_\mu(X).$$

Suppose that $K : X \times X \to R = (-\infty, +\infty)$ is continuous, symmetric and strictly positive definite, i.e., for any given integers $m \geq 1, (K(x_i, x_j))_{i,j=1}^m$ are positive definite matrices for given finite sets $\{x_1, x_2, \cdots, x_m\} \subset X$. Assume that $K \in L^2_{\mu\times\mu}(X \times X)$, i.e.,

$$\int_X \int_X |K(x,t)|^2 d\mu(x)d\mu(t) < +\infty.$$

Then the linear operator $L_K :\ L^2_\mu(X) \to L^2_\mu(X)$ defined by

$$L_K(f,x) = \int_X K(x,\ t)f(t)d\mu(t), \qquad x \in X \tag{1}$$

is positive, and its range lies in $C(X)$. Take $L_K^{\frac{1}{2}}$ to be the linear operator on $L^2_\mu(X)$ satisfying $L_K^{\frac{1}{2}} \circ L_K^{\frac{1}{2}} = L_K$ and $L_K^{-\frac{1}{2}}$, the inverse of $L_K^{\frac{1}{2}}$. Additionally, define $\mathcal{H}_K = L_K^{\frac{1}{2}}(L^2_\mu(X))$. Then $(\mathcal{H}_K, \|\cdot\|_{\mathcal{H}_K})$ is a reproducing kernel Hilbert space associated with $K_x(y) = K(x, y)$, i.e., (see [1,4–7]),

$$f(x) = \langle f,\ K_x\rangle_{\mathcal{H}_K}, \qquad f \in \mathcal{H}_K,\ x \in X, \tag{2}$$

where the inner product $\langle\cdot,\cdot\rangle_{\mathcal{H}_K}$ is induced by a norm defined as

$$\|f\|_{\mathcal{H}_K} = \|L_K^{-\frac{1}{2}} f\|_{L^2_\mu(X)}, \qquad f \in \mathcal{H}_K, \tag{3}$$

i.e.,

$$\|L_K^{\frac{1}{2}} f\|_{\mathcal{H}_K} = \|f\|_{L^2_\mu(X)}, \qquad f \in L^2_\mu(X). \tag{4}$$

One of the targets of learning theory is to find an unknown function $f : X \to R$ from the random observations $\{(x_i,\ y_i)\}_{i=1}^m$ drawn i.i.d. (identically and independently distributed) according to a unknown probability $\rho(x, y) = \rho_X(x)\rho(y|x)$ defined on $X \times R$ (see [1,6]). A usual algorithm to realize this aim is to solve the following kernel regularized optimization problem:

$$f_{z,\lambda} = arg \min_{f\in\mathcal{H}_K} \frac{1}{m}\sum_{i=1}^m (f(x_i) - y_i)^2 + \lambda\|f\|^2_{\mathcal{H}_K}, \tag{5}$$

where $\mathcal{H}_K$ is taken as the hypothesis space, $\lambda > 0$ is a parameter which balances the relationship between the empirical error term $\sum_{i=1}^m (f(x_i) - y_i)^2$ and the penalty term $\|f\|^2_{\mathcal{H}_K}$. Let $f_\rho(x) = \int_R y\ d\rho(y|x)$ be the regression function. Then $f_\rho$ is the least-squares-best predictor (see Section 9.4 in Section 9 of [8]), i.e.,

$$E((f_\rho(\cdot) - y)^2) = \inf_{g\in L^2_{\rho_X}(X)} E\Big((f(\cdot) - y)^2\Big).$$

It is known that the convergence analysis of model (5) sums up to bound the convergence rate for error $\|f_{z,\lambda} - f_\rho\|_{L^2_{\rho_X}(X)}$, which depends upon the decay of the best approximation $I(f,\ \gamma)_{L^2_{\rho_X}(X)}$ defined as (see e.g., [1,2,6])

$$I(f,\gamma)_{L^2_{\rho_X}(X)} \quad = \inf_{g\in\mathcal{H}_K,\ \|g\|_{\mathcal{H}_K}\le\gamma} \|f - g\|_{L^2_{\rho_X}(X)}, \qquad \gamma > 0 \tag{6}$$

as $\gamma \to +\infty$.

Formula (6) deals with a decay rate which depends upon the approximation property of $\mathcal{H}_K$. Many mathematicians have performed investigations on it. For example, D. X. Zhou gives the decay of (6) with the RKHS interpolation theory (see [2,3]). P.X. Ye gives the decay using convolution operators in the Euclidean space $R^d$ (see [9]). H.W. Sun gives a decay for (6) with the help of operator theory in a Hilbert space (see [10]). It is known that the Fourier–Bessel series is a good approximation tool and has been studied by many mathematicians (see for example, [11–16]). Additionally, we found that approximation by RBF networks of Delsarte translates was studied by some mathematicians. The essence of RBF is summed up as the approximation of Fourier–Bessel transforms (see, for example, [17–20]). So it is of interest for us to conduct investigations on the decay of $I(f, R)_{L^2_{\rho_X}(X)}$ with both the Fourier–Bessel series and the Fourier–Bessel transforms.

Let $\alpha > -\frac{1}{2}$ and $1 \leq p \leq +\infty$ be given real numbers, and $L^p(R_+,\ d\mu_\alpha)$ denote the space of all measurable real functions on $R_+ = [0, +\infty)$ such that

$$\|f\|_{p,\alpha} = \begin{cases} \left( \int_{R_+} \left|f(x)\right|^p d\mu_\alpha \right)^{\frac{1}{p}} < +\infty, & 1 \leq p < +\infty, \\ \\ ess \sup\limits_{x \in R_+} \left|f(x)\right| < +\infty, & p = +\infty, \end{cases}$$

where $d\mu_\alpha(x) = \frac{x^{2\alpha+1}}{2^\alpha \Gamma(\alpha+1)}\ dx$. The normalized Bessel function $j_\alpha(z)$ of the first kind and order $\alpha$ is

$$\begin{aligned} j_\alpha(z) &= \Gamma(\alpha+1) \sum_{n=0}^{+\infty} \frac{(-1)^n (\frac{z}{2})^{2n}}{n!\ \Gamma(n+\alpha+1)} \\ &= 2^\alpha \Gamma(\alpha+1) \frac{J_\alpha(x)}{x^\alpha}, \qquad z \in R_+, \end{aligned} \tag{7}$$

where

$$J_\alpha(x) = \left(\frac{x}{2}\right)^\alpha \sum_{n=0}^{+\infty} \frac{(-1)^n (\frac{z}{2})^{2n}}{n!\ \Gamma(n+\alpha+1)}$$

is the Bessel function of first kind and order $\alpha$, and $\Gamma(\alpha+1)$ is the Gamma function.

For $f \in L^1(R_+,\ d\mu_\alpha)$, the usual Fourier–Bessel transform $\mathcal{F}_B^{(\alpha)}(f)$ is defined as

$$\mathcal{F}_B^{(\alpha)}(f)(\lambda) = \int_{R_+} f(x)\ j_\alpha(\lambda x)\ d\mu_\alpha, \qquad \lambda \in R_+.$$

In the present paper, some investigations on the decay of $I(f,\ \gamma)_{L^2_{\rho_X}(X)}$ in the case that $H_K$ are constructed with $j_\alpha(z)(z \in [0,1])$ and $\mathcal{F}_B^{(\alpha)}(f)$ are provided. Some $K$-functional and moduli of smoothness are defined with the help of the semigroup of operators, and their equivalences are shown, with which the error for the decay is bounded. The results obtained are two kinds of upper bound estimates associated with Fourier–Bessel series and Fourier–Bessel transforms, respectively.

The paper is organized as follows. In Section 2, some notions and results of the Fourier–Bessel series and Fourier–Bessel transforms are provided, with which two kinds of RKHSs are constructed; the corresponding best RKHS approximation problem in these setting is restated. Some $K$-functionals and moduli of smoothness associated with Fourier–Bessel series and Fourier–Bessel transforms are provided, and their equivalence is shown, with which some upper bounds for the best approximation are shown in Sections 3 and 4, respectively. All the proofs for the propositions, the theorems and lemmas are given in Section 5. Some further analysis for the results of the present paper are given in Section 6, from which one can see the value of writing this manuscript. A general proposition for the strong equivalence of $K$-functionals and moduli of smoothness is listed in the Appendix A.

## 2. Preliminaries

Let $\lambda_1, \lambda_2, \cdots$, be the positive zeros of $J_\alpha(u)$ arranged in increasing order. It is well known that $j_\alpha(\lambda_n x),\ n = 1, 2, \cdots$, form a complete orthogonal system in $L^2_\alpha = \{f : \|f\|_{L^2_\alpha} = (\int_0^1 x^{2\alpha+1} |f(x)|^2\ dx)^{\frac{1}{2}} < +\infty\}$ (see, for example, [12,16,21]), i.e.,

$$\int_0^1 x^{2\alpha+1} j_\alpha(\lambda_n u)\ j_\alpha(\lambda_m u) du = \|j_\alpha(\lambda_i \cdot)\|^2_{L^2_\alpha}\ \delta_{m,n}.$$

Take $j_\alpha^*(\lambda_i x) = \frac{j_\alpha(\lambda_i x)}{\|j_\alpha(\lambda_i \cdot)\|_{L^2_\alpha}}$. Then

$$\int_0^1 x^{2\alpha+1} j_\alpha^*(\lambda_n u) j_\alpha^*(\lambda_m u)\, du = \delta_{m,n}, \tag{8}$$

$\{j_\alpha^*(\lambda_i x)\}_{i=1}^{\infty}$ forms an orthonormal basis of $L^2_\alpha$ and for any $f \in L^2_\alpha$, there holds Fourier–Bessel series

$$f(x) = \sum_{i=1}^{+\infty} a_i(f)\, j_\alpha^*(\lambda_i x), \qquad x \in [0,1], \tag{9}$$

where $a_i(f) = \int_0^1 x^{2\alpha+1} f(x) j_\alpha^*(\lambda_i x) dx$ and

$$\|f\|_{L^2_\alpha} = \left( \sum_{i=1}^{+\infty} |a_i(f)|^2 \right)^{\frac{1}{2}}. \tag{10}$$

**Lemma 1.** *We have the following results:*

*(i)* *Let $\Lambda \subset \mathcal{N}$. Then*

$$\| \sum_{i \in \Lambda} c_i\, j_\alpha^*(\lambda_i x) \|_{L^2_\alpha} = \left( \sum_{i \in \Lambda} c_i^2 \right)^{\frac{1}{2}}. \tag{11}$$

*(ii)* *The generalized translation operator $T_x$ on $L^2_\alpha$ defined as*

$$T_x(f)(y) = \frac{\Gamma(\alpha+1)}{\sqrt{\pi}\Gamma(\alpha+\frac{1}{2})} \int_0^\pi f\left( \sqrt{x^2+y^2-2xy\cos\theta} \right) (\sin\theta)^{2\alpha} d\theta, \quad x, y \in [0,1]$$

*has the expansion of*

$$T_x(f)(y) = \sum_{i=1}^{+\infty} a_i^*(f)\, j_\alpha^*(\lambda_i x)\, j_\alpha^*(\lambda_i y), \qquad x,\ y \in [0,1], \tag{12}$$

*where $a_i^*(f) = \int_0^1 x^{2\alpha+1} f(x)\, j_\alpha(\lambda_i x)\, dx$, and*

$$\|T_h(f)(\cdot)\|_{L^2_\alpha} \le \|f\|_{L^2_\alpha}, \qquad \forall h \in [0,1]. \tag{13}$$

*(iii)* *The zeros $\{\lambda_1, \lambda_2, \cdots, \}$ satisfy*

$$\lambda_n = n\pi + \frac{\alpha\pi}{2} - \frac{\pi}{4} + O(\frac{1}{n}). \tag{14}$$

**Proof.** See it from Section 5. □

Inequality (13) is a theoretical basis for defining the moduli of smoothness with translation operators $T_x(f)(y)$.

Let $\{h_i\}_{i=1}^{+\infty}$ be the set of given positive real sequences such that the right side of the series

$$K_x^{(\alpha)}(y) = K^{(\alpha)}(x,y) = \sum_{i=1}^{+\infty} h_i\, j_\alpha^*(\lambda_i x)\, j_\alpha^*(\lambda_i y), \qquad x,\ y \in [0,1], \tag{15}$$

has uniform convergence for all $x \in R_+$. It therefore is a Mercer kernel. Then

$$L_{K^{(\alpha)}}(f,x) = \sum_{i=1}^{+\infty} h_i \, a_i(f) \, j^*_\alpha(\lambda_i x), \qquad x \in [0,1]. \tag{16}$$

Take

$$L^{\frac{1}{2}}_{K^{(\alpha)}}(f,x) = \sum_{i=1}^{+\infty} \sqrt{h_i} \, a_i(f) \, j^*_\alpha(\lambda_i x), \qquad x \in [0,1]. \tag{17}$$

Then it is easy to verify that $L_{K^{(\alpha)}} = L^{\frac{1}{2}}_{K^{(\alpha)}} \circ L^{\frac{1}{2}}_{K^{(\alpha)}}$, and

$$\mathcal{H}_{K^{(\alpha)}} = L^{\frac{1}{2}}_{K^{(\alpha)}}(L^2_\alpha) = \{g \in L^2_\alpha : \ \|g\|_{K^{(\alpha)}} = \|L^{-\frac{1}{2}}_{K^{(\alpha)}}(g)\|_{L^2_\alpha} = \left(\sum_{i=1}^{+\infty} \frac{|a_i(g)|^2}{h_i}\right)^{\frac{1}{2}} < +\infty\}$$

is a RKHS in $L^2_\alpha$ associating with reproducing kernel $K^{(\alpha)}(x,y)$ and an inner product $\langle \cdot, \cdot \rangle_{K^{(\alpha)}}$ defined as

$$\langle f, \ g\rangle_{K^{(\alpha)}} = \sum_{i=1}^{+\infty} \frac{a_i(f) \, a_i(g)}{h_i}, \qquad f, \ g \in \mathcal{H}_{K^{(\alpha)}}.$$

Since

$$\begin{aligned} a_i(K^{(\alpha)}_x(\cdot)) &= \int_0^1 y^{2\alpha+1} K^{(\alpha)}(x,y) j^*_\alpha(\lambda_i y) dy \\ &= \int_0^1 y^{2\alpha+1} \left(\sum_{k=1}^{+\infty} h_k \, j^*_\alpha(\lambda_k x) \, j^*_\alpha(\lambda_k y)\right) j^*_\alpha(\lambda_i y) dy \\ &= h_i \, j^*_\alpha(\lambda_i x), \end{aligned}$$

we have

$$\begin{aligned} \langle f, \ K^{(\alpha)}_x(\cdot)\rangle_{K^{(\alpha)}} &= \sum_{i=1}^{+\infty} \frac{a_i(f) \, a_i(K^{(\alpha)}_x(\cdot))}{h_i} \\ &= \sum_{i=1}^{+\infty} \frac{a_i(f) \, h_i \, j^*_\alpha(\lambda_i x)}{h_i} \\ &= \sum_{i=1}^{+\infty} a_i(f) \, j^*_\alpha(\lambda_i x) = f(x). \end{aligned}$$

Equality (6) becomes

$$I(f,\gamma)_{L^2_\alpha} = \inf_{g \in \mathcal{H}_{K^{(\alpha)}}, \, \|g\|_{K^{(\alpha)}} \le \gamma} \|f-g\|_{L^2_\alpha}, \qquad \gamma > 0 \tag{18}$$

as $\gamma \to +\infty$.

Let $C_*(R)$ be the class of even $C^\infty$-functions on $R = \{-\infty, \ +\infty\}$. Denoted by $A_*(R)$, the space of even $C^\infty$-functions on $R$ which are rapidly decreasing together with all their derivatives, i.e.,

$$\forall p,k \in \mathcal{N}, \quad \sup_{x \ge 0}\Big(|x^p \, f^{(k)}(x)|\Big) < +\infty,$$

where $\mathcal{N}$ is the set of natural numbers.

Let $D_{*,a}$ denote the space of even $C^\infty$-functions on $R$ with support in $[-a, \ a]$, $a \ge 0$ and

$$D_*(R) = \bigcup_{a \ge 0} D_{*,a}.$$

Additionally, define the generalized translation operator $T_x$ on $L^1(R_+, d\mu_\alpha)$ as

$$T_x(f)(y) = \frac{\Gamma(\alpha+1)}{\sqrt{\pi}\Gamma(\alpha+\frac{1}{2})} \int_0^\pi f\left(\sqrt{x^2+y^2-2xy\cos\theta}\right)(\sin\theta)^{2\alpha} d\theta, \quad x,\ y \in R_+.$$

and define a convolution on $L^1(R_+, d\mu_\alpha)$ by

$$(f *_{\mathcal{B}} g)(x) = \int_{R_+} T_x(f)(y) g(y) d\mu_\alpha(y), \ \ f,\ g \in L^1(R_+, d\mu_\alpha),\ x \in R_+.$$

For the Bessel operators

$$l_\alpha = \frac{d^2}{dx^2} + \frac{2\alpha+1}{x}\frac{d}{dx}$$

we have (see p. 12 or p. 177 of [22])

$$(-l_\alpha)(j_\alpha(\lambda\cdot))(x) = \lambda^2\, j_\alpha(\lambda x),\ (-l_\alpha)^{-1}(j_\alpha(\lambda\cdot))(x) = \frac{1}{\lambda^2}\, j_\alpha(\lambda x), \quad \lambda, \quad x \in R_+. \tag{19}$$

and therefore

$$(-l_\alpha)^{\mp\frac{1}{2}}(j_\alpha(\lambda\cdot))(x) = \lambda^{\mp}\, j_\alpha(\lambda x), \qquad x \in R_+.$$

Moreover, we have the following lemma.

**Lemma 2.** *There hold the following:*

*(i)* *$D_*(R)$ is dense in $A_*(R)$;*
*(ii)* *Both $D_*(R)$ and $A_*(R)$ are dense in $L^p(R_+, d\mu_\alpha)$, $1 \le p < +\infty$, and*

$$D_*(R) \subset A_*(R) \subset L^p(R_+, d\mu_\alpha), \qquad 1 \le p < +\infty; \tag{20}$$

*(iii)* *If $f \in A_*(R)$, then $\mathcal{F}_B^{(\alpha)}(f) \in A_*(R)$ and $T_x(f) \in A_*(R)$;*
*(iv)* *$\mathcal{F}_B^{(\alpha)}$ is a topological isomorphism from $A_*(R)$ to itself and ${\mathcal{F}^{(\alpha)}}_B^{-1} = \mathcal{F}_B^{(\alpha)}$.*
*(v)* *There hold*

$$\mathcal{F}_B^{(\alpha)}(f *_{\mathcal{B}} g) = \mathcal{F}_B^{(\alpha)}(f)\mathcal{F}_B^{(\alpha)}(g), \qquad f, g \in L^1(R_+, d\mu_\alpha), \tag{21}$$

$$(f *_B g)(x) = \int_{R_+} \mathcal{F}_B^{(\alpha)}(f)(\lambda)\mathcal{F}_B^{(\alpha)}(g)(\lambda) j_\alpha(\lambda x) d\mu_\alpha(\lambda) \tag{22}$$

*and*

$$\mathcal{F}_B^{(\alpha)}(T_x(f))(\lambda) = j_\alpha(\lambda x)\mathcal{F}_B^{(\alpha)}(f)(\lambda), \quad f \in L^1(R_+, d\mu_\alpha). \tag{23}$$

*It follows*

$$T_x(f, y) = \int_{R_+} \mathcal{F}_B^{(\alpha)}(f)(\lambda)\ j_\alpha(\lambda x)\ j_\alpha(\lambda y)\ d\mu_\alpha(\lambda), \quad f \in L^1(R_+, d\mu_\alpha). \tag{24}$$

*(vi)* *If $f, \mathcal{F}_B^{(\alpha)}(f) \in L^1(R_+,\ d\mu_\alpha)$, then*

$$f(x) = \int_{R_+} \mathcal{F}_B^{(\alpha)}(f)(\lambda)\ j_\alpha(\lambda x)\ d\mu_\alpha(\lambda), \quad a.e. x \in R_+; \tag{25}$$

*(vii)* *Let $f \in A_*(R)$ or $f \in L^2(R_+, d\mu_\alpha)$. Then*

$$\int_{R_+} |f(x)|^2 d\mu_\alpha = \int_{R_+} \left|\mathcal{F}_B^{(\alpha)}(f)(\lambda)\right|^2 d\mu_\alpha(\lambda); \tag{26}$$

*(viii) There hold the following relations*

$$\mathcal{F}_B^{(\alpha)}(l_\alpha^p(f))(\lambda) = (-1)^p \lambda^{2p} \mathcal{F}_B^{(\alpha)}(f)(\lambda), \quad f \in L^1(R_+, d\mu_\alpha), \tag{27}$$

$$\|T_x(f)\|_{p,\alpha} \leq \|f\|_{p,\alpha}, \quad f \in L^p(R_+, d\mu_\alpha),\ 1 \leq p < +\infty, \tag{28}$$

$$\mathcal{F}_B^{(\alpha)}(j_\alpha(\lambda\cdot))(y) = j_\alpha(\lambda x)\, j_\alpha(\lambda y), \quad \forall x, y, \lambda \in R_+. \tag{29}$$

Proposition 2.1 of [23] shows that if $\phi \in L^1(R_+,\ d\mu_\alpha)$ satisfies $\mathcal{F}_B^{(\alpha)}(\phi) \geq 0$ and $\mathcal{F}_B^{(\alpha)}(\phi) \in L^1(R_+,\ d\mu_\alpha)$, then

$$K(\phi, x, y) = K_x(\phi, y) = T_x(\phi, y) = \int_{R_+} \mathcal{F}_B^{(\alpha)}(\phi)(\lambda) j_\alpha(\lambda x)\, j_\alpha(\lambda y)\, d\mu_\alpha, \qquad y \in R_+.$$

defines a Mercer kernel on $R_+$. We give an assumption

**Assumption 1.** *Let $\phi \in L^1(R_+,\ d\mu_\alpha)$ satisfy $\mathcal{F}_B^{(\alpha)}(\phi) > 0, \mathcal{F}_B^{(\alpha)}(\phi) \in L^1(R_+,\ d\mu_\alpha)$ and for any $\mu > 0$ there is a real number $a \in R_+$ such that*

$$\{\lambda \in R_+ :\ \mathcal{F}_B^{(\alpha)}(\phi)(\lambda) \leq \frac{1}{\mu}\} \subset [0,\ a]. \tag{30}$$

*We point here that the functions $\phi$ satisfying Assumption 1 are existent, and give two examples.*

*Example 1.* For $t \in (0,\ +\infty)$ the function $p_t : [0, +\infty) \to R_+$ defined by

$$p_t(x) = \frac{2^{\alpha+1}\Gamma(\alpha + \frac{3}{2})}{\sqrt{\pi}} \frac{t}{(t^2 + x^2)^{\alpha+\frac{3}{2}}}$$

satisfies $\|p_t\|_{L^1(R_+,\, d\mu_\alpha)} = 1, p_t *_{\mathcal{B}} p_s = p_{t+s}$ and $\mathcal{F}_B^{(\alpha)}(p_t)(\lambda) = e^{-t\lambda}$ for $\lambda \in R_+$ (see Problem 5. VIII 2 in Section 5.VIII Problems of [22]).

*Example 2.* For $t, s \in (0,\ +\infty)$ the function $k_t : R_+ \to R_+$ defined by

$$k_t(x) = \frac{e^{-\frac{x^2}{4t}}}{(2t)^{\alpha+1}}$$

satisfies $\|k_t\|_{L^1(R_+,\, d\mu_\alpha)} = 1,\ k_t *_{\mathcal{B}} k_s = k_{t+s}$ and $\mathcal{F}_B^{(\alpha)}(k_t)(\lambda) = e^{-t\lambda^2}$ for $\lambda \in R_+$ (see Problem 5. VIII 1 in Section 5.VIII Problems of [22]).

Define

$$\mathcal{H}_{K(\phi)} = \{g \in L^2(R_+,\ d\mu_\alpha) \cap C_*(R) : \quad \frac{\mathcal{F}_B^{(\alpha)}(g)}{\mathcal{F}_B^{(\alpha)}(\phi)^{\frac{1}{2}}} \in L^2(R_+,\ d\mu_\alpha), \quad g(u) = \int_{R_+} \mathcal{F}_B^{(\alpha)}(g)(\lambda)\, j_\alpha(\lambda u)\, d\mu_\alpha(\lambda)\}$$

with norm $\|g\|_{\mathcal{H}_{K(\phi)}} = \left( \int_{R_+} \frac{|\mathcal{F}_B^{(\alpha)}(g)(\lambda)|^2}{\mathcal{F}_B^{(\alpha)}(\phi)(\lambda)} d\mu_\alpha \right)^{\frac{1}{2}}$

Define an inner product on $\mathcal{H}_{K(\phi)}$ as

$$\langle g,\ f \rangle_{K(\phi)} = \int_{R_+} \frac{\mathcal{F}_B^{(\alpha)}(f)(\lambda)\mathcal{F}_B^{(\alpha)}(g)(\lambda)}{\mathcal{F}_B^{(\alpha)}(\phi)(\lambda)} d\mu_\alpha, \qquad f, g \in \mathcal{H}_{K(\phi)}.$$

It is known that $K(\phi, x, y)$ is a reproducing kernel of $\mathcal{H}_{K(\phi)}$ (see [24]), i.e.,

$$\langle g, K(\phi, x, \cdot)\rangle_{K(\phi)} = g(x), \qquad g \in \mathcal{H}_{K(\phi)}, \qquad x \in R_+. \tag{31}$$

We have

$$\begin{aligned} L_{K(\phi)}(f, x) &= \int_{R_+} K_x(\phi, u) f(u)\, d\mu_\alpha(u) \\ &= \int_{R_+} \mathcal{F}_B^{(\alpha)}(\phi)(\lambda)\mathcal{F}_B^{(\alpha)}(f)(\lambda)\, j_\alpha(\lambda x)\, d\mu_\alpha(\lambda), \quad f \in L^1(R_+, d\mu_\alpha). \end{aligned} \tag{32}$$

Define for a given real number $r \in R$ an operator as

$$L^r_{K(\phi)}(f, x) = \int_{R_+} \left(\mathcal{F}_B^{(\alpha)}(\phi)(\lambda)\right)^r \mathcal{F}_B^{(\alpha)}(f)(\lambda)\, j_\alpha(\lambda x)\, d\mu_\alpha(\lambda), f \in L^1(R_+, d\mu_\alpha). \tag{33}$$

Then it is easy to show that $L_{K(\phi)} = L^{\frac{1}{2}}_{K(\phi)}(L^{\frac{1}{2}}_{K(\phi)}) = L^{\frac{1}{2}}_{K(\phi)} \circ L^{\frac{1}{2}}_{K(\phi)}$,

$$L^{\frac{1}{2}}_{K(\phi)}(L^2(R_+, d\mu_\alpha)) = \{g \in L^2(R_+, d\mu_\alpha) : \left(\int_{R_+} \frac{|\mathcal{F}_B^{(\alpha)}(g)(\lambda)|^2}{\mathcal{F}_B^{(\alpha)}(\phi)(\lambda)} d\mu_\alpha\right)^{\frac{1}{2}} < +\infty\} = \mathcal{H}_{K(\phi)},$$

and

$$\|f\|_{K(\phi)} = \|L^{-\frac{1}{2}}_{K(\phi)}(f)\|_{L^2(R_+,\, d\mu_\alpha)}, \quad f \in \mathcal{H}_{K(\phi)}.$$

In this case, the decay (6) becomes

$$I(f, \gamma)_{L^2(R_+,\, d\mu_\alpha)} = \inf_{g \in H_{K(\phi)},\, \|g\|_{K(\phi)} \le \gamma} \|f - g\|_{L^2(R_+,\, d\mu_\alpha)}, \quad f \in L^2(R_+,\, d\mu_\alpha) \tag{34}$$

for $\gamma \to +\infty$.

If $\mathcal{F}_B^{(\alpha)}(\phi)(\lambda) = \frac{1}{\lambda^2}$, then we define the corresponding RKHS

$$\mathcal{H}^*_{K(\phi)} = L^{\frac{1}{2}}_{K(\phi)}(A_*(R)) = \{g \in A_*(R) : \left(\int_{R_+} \lambda^2 \left|\mathcal{F}_B^{(\alpha)}(g)(\lambda)\right|^2 d\mu_\alpha\right)^{\frac{1}{2}} < +\infty\}$$

and for $g \in \mathcal{H}^*_{K(\phi)}$, there holds

$$\begin{aligned} \|g\|_{K(\phi)} &= \|L^{-\frac{1}{2}}_{K(\phi)}(g)\|_{L^2(R_+, d\mu_\alpha)} \\ &= \left(\int_{R_+} \lambda^2 \left|\mathcal{F}_B^{(\alpha)}(g)(\lambda)\right|^2 d\mu_\alpha(\lambda)\right)^{\frac{1}{2}} \\ &= \|(-l_\alpha)^{\frac{1}{2}} g\|_{L^2(R_+,\, d\mu_\alpha)}. \end{aligned}$$

We have by (34) that

$$I(f, \gamma)_{L^2(R_+,\, d\mu_\alpha)} = \inf_{\|(-l_\alpha)^{\frac{1}{2}} g\|_{L^2(R_+,\, d\mu_\alpha)} \le \gamma} \|f - g\|_{L^2(R_+,\, d\mu_\alpha)} \tag{35}$$

for $\gamma \to +\infty$.

## 3. An Upper Bound Estimate with Fourier–Bessel Series

To bound the decay of (18), we define a $K$-functional

$$D_{\mathcal{H}_{K(\alpha)}}(f,\, t)_{L^2_\alpha} = \inf_{g \in \mathcal{H}_{K(\alpha)}} \left(\|f - g\|_{L^2_\alpha} + t\|g\|_{K(\alpha)}\right), \qquad f \in L^2_\alpha, t > 0 \tag{36}$$

and a modulus of smoothness

$$\omega_{\mathcal{H}_{K(\alpha)}}(f,t)_{L^2_\alpha} = \|(T_{K(\alpha)}(t) - I)f\|_{L^2_\alpha}, \qquad f \in L^2_\alpha,\ t > 0, \tag{37}$$

where

$$T_{K(\alpha)}(t)f(x) = \sum_{i=1}^{\infty} e^{-\frac{t}{\sqrt{h_i}}}\, a_i(f)\, j^*_\alpha(\lambda_i x), \qquad x \in [0,1].$$

Then we have the following Proposition 1 whose proofs can be found from Section 5.

**Proposition 1.** *There holds an equivalent relation*

$$D_{\mathcal{H}_{K(\alpha)}}(f,t)_{L^2_\alpha} \sim \omega_{\mathcal{H}_{K(\alpha)}}(f,t)_{L^2_\alpha}, \qquad f \in L^2_\alpha,\ t > 0. \tag{38}$$

**Proof.** See it from Section 5. □

**Theorem 1.** *There is a constant $C > 0$ such that*

$$I(f,\gamma)_{L^2_\alpha} \le C\omega_{\mathcal{H}_{K(\alpha)}}\left(f, \frac{\|f\|_{L^2_\alpha}}{\gamma}\right)_{L^2_\alpha}, \qquad f \in L^2_\alpha \tag{39}$$

*if $\gamma \to +\infty$.*

**Proof.** See it from Section 5. □

Taking $h_i = \frac{1}{\lambda_i^2}$ into (15), we have a kernel

$$K^*_x(y) = K^*(x,y) = \sum_{i=1}^{+\infty} \frac{1}{\lambda_i^2}\, j^*_\alpha(\lambda_i x)\, j^*_\alpha(\lambda_i y), \qquad x,y \in [0,1],$$

It follows that

$$\begin{aligned} \mathcal{H}_{K^*} &= L^{\frac{1}{2}}_{K^*}(L^2_\alpha) \\ &= \{g \in L^2_\alpha : \quad \|g\|_{K^*} = \left(\sum_{i=1}^{+\infty} \lambda_i^2 |a_i(g)|^2\right)^{\frac{1}{2}} < +\infty\}, \end{aligned}$$

which shows that $\|g\|_{K^*} = \|(-l_\alpha)^{\frac{1}{2}}(g)\|_{L^2_\alpha}$ and

$$D_{\mathcal{H}_{K^*}}(f,t)_{L^2_\alpha} = \inf_{g \in \mathcal{H}_{K^*}} \left(\|f-g\|_{L^2_\alpha} + t\|(-l_\alpha)^{\frac{1}{2}}(g)\|_{L^2_\alpha}\right), \qquad f \in L^2_\alpha, t > 0$$

and

$$\omega_{\mathcal{H}_{K^*}}(f,t)_{L^2_\alpha} = \|(T_{K^*}(t) - I)f\|_{L^2_\alpha}, \qquad f \in L^2_\alpha, t > 0,$$

where

$$T_{K^*}(t)f(x) = \sum_{i=1}^{\infty} e^{-t\lambda_i}\, a_i(f)\, j^*_\alpha(\lambda_i x), \qquad x \in [0,1].$$

We have two corollaries.

**Corollary 1.** *For any $f \in L^2_\alpha$, there holds*

$$D_{\mathcal{H}_{K^*}}(f,t)_{L^2_\alpha} \sim \omega_{\mathcal{H}_{K^*}}(f,t)_{L^2_\alpha}, \qquad f \in L^2_\alpha, t > 0$$

**Corollary 2.** *For any $f \in L^2_\alpha$, there holds*

$$I(f,\gamma)_{L^2_\alpha} \le C\omega_{\mathcal{H}_{K^*}}\left(f, \frac{\|f\|_{L^2_\alpha}}{\gamma}\right)_{L^2_\alpha}, \qquad \gamma \to +\infty.$$

## 4. An Upper Bound Estimate with the Fourier–Bessel Transform

To bound $I(f,\ \gamma)_{L^2(R_+,\, d\mu_\alpha)}$, we define a $K$-functional $D_{K(\phi)}(f,t)_{L^2(R_+,d\mu_\alpha)}$ and a modulus $\omega_{K(\phi)}(f,t)_{L^2(R_+,\, d\mu_\alpha)}$ respectively corresponding to $\mathcal{H}_{K(\phi)}$ as

$$\begin{aligned} &D_{K(\phi)}(f,t)_{L^2(R_+,\, d\mu_\alpha)} \\ = &\inf_{g\in\mathcal{H}_{K(\phi)}}\left(\|f-g\|_{L^2(R_+,\, d\mu_\alpha)} + t\|g\|_{K(\phi)}\right) \\ = &\inf_{g\in L^{\frac{1}{2}}_{K(\phi)}(L^2(R_+,d\mu_\alpha))}\left(\|f-g\|_{L^2(R_+,d\mu_\alpha)} + t\ \|L^{-\frac{1}{2}}_{K(\phi)}(f)\|_{L^2(R_+,d\mu_\alpha)}\right), f\in L^2(R_+,d\mu_\alpha), \end{aligned}$$

and

$$\omega_{K(\phi)}(f,t)_{L^2(R_+,\, d\mu_\alpha)} = \|(T_{K(\phi)}(t)-I)f\|_{L^2(R_+,\, d\mu_\alpha)}, \quad f\in L^2(R_+,\ d\mu_\alpha), t>0,$$

where

$$T_{K(\phi)}(t)f(x) = \int_{R_+} e^{-\frac{t}{\sqrt{\mathcal{F}^{(\alpha)}_B(\phi)(\lambda)}}}\mathcal{F}^{(\alpha)}_B(f)(\lambda)\ j_\alpha(\lambda x)\ d\mu_\alpha(\lambda).$$

The K-functional and the modulus are equivalent, i.e., we have the following proposition.

**Proposition 2.** *Let $\phi \in L^1(R_+,\ d\mu_\alpha)$ satisfy Assumption 1. Then there holds the equivalence*

$$D_{K(\phi)}(f,t)_{L^2(R_+,\, d\mu_\alpha)} \sim \omega_{K(\phi)}(f,t)_{L^2(R_+,\, d\mu_\alpha)},\ f\in L^2(R_+,\ d\mu_\alpha), t>0. \tag{40}$$

We now give an upper bound estimate for (34).

**Theorem 2.** *Under the conditions of Proposition 2, there is a constant $C > 0$ such that*

$$I(f,\gamma)_{L^2(R_+,\, d\mu_\alpha)} \quad \le \quad C\omega_{K(\phi)}\left(f, \frac{\|f\|_{L^2(R_+,\, d\mu_\alpha)}}{\gamma}\right)_{L^2(R_+,\, d\mu_\alpha)},\ f\in L^2(R_+,\ d\mu_\alpha) \tag{41}$$

*if $\gamma \to +\infty$.*

For $\mathcal{F}^{(\alpha)}_B(\phi)(\lambda) = \frac{1}{\lambda^2}$ we define a $K$-functional on $L^2(R_+,\ d\mu_\alpha)$ as

$$D_{l^{\frac{1}{2}}_\alpha}(f,t)_{L^2(R_+,\, d\mu_\alpha)} = \inf_{g\in\mathcal{H}^*_{K(\phi)}}\left(\|f-g\| + t\|(-l_\alpha)^{\frac{1}{2}}g\|_{L^2(R_+,\, d\mu_\alpha)}\right), t>0.$$

Define a modulus of smoothness as

$$\omega_{l^{\frac{1}{2}}_\alpha}(f,t)_{L^2(R_+,\, d\mu_\alpha)} = \|(T_{l^{\frac{1}{2}}_\alpha}(t)-I)f\|_{L^2(R_+,\, d\mu_\alpha)}, t>0,$$

where

$$T_{l^{\frac{1}{2}}_\alpha}(t)f(x) = \int_{R_+} e^{-\lambda t}\mathcal{F}^{(\alpha)}_B(f)(\lambda)\ j_\alpha(\lambda x)d\mu_\alpha(\lambda).$$

Then we have the following two corollaries.

**Corollary 3.** *There holds the equivalent relation*

$$D_{l_\alpha^{\frac{1}{2}}}(f,t)_{L^2(R_+,\,d\mu_\alpha)} \sim \omega_{l_\alpha^{\frac{1}{2}}}(f,t)_{L^2(R_+,\,d\mu_\alpha)}, f \in L^2(R_+,\, d\mu_\alpha), t > 0.$$

**Corollary 4.** *There is a constant $C > 0$ such that*

$$I(f,R)_{L^2(R_+,\,d\mu_\alpha)} \leq C\omega_{l_\alpha^{\frac{1}{2}}}\left(f, \frac{\|f\|_{L^2(R_+,\,d\mu_\alpha)}}{R}\right)_{L^2(R_+,\,d\mu_\alpha)}, \; f \in L^2(R_+,\, d\mu_\alpha). \tag{42}$$

We give further computations for $T_{l_\alpha^{\frac{1}{2}}}(t)f(x)$. By Example 1, we know $\mathcal{F}_B^{(\alpha)}(p_t)(\lambda) = e^{-\lambda t}$, which, together with (21), gives

$$\begin{aligned} T_{l_\alpha^{\frac{1}{2}}}(t)f(x) &= \int_{R_+} \mathcal{F}_B^{(\alpha)}(p_t)(\lambda)\mathcal{F}_B^{(\alpha)}(f)(\lambda)\, j_\alpha(\lambda x)d\mu_\alpha(\lambda) \\ &= \int_{R_+} \mathcal{F}_B^{(\alpha)}(f *_B p_t)(\lambda)\, j_\alpha(\lambda x)d\mu_\alpha(\lambda) \\ &= (f *_B p_t)(x), \qquad x \in R_+, \end{aligned}$$

which with (42) shows that

$$\omega_{l_\alpha^{\frac{1}{2}}}(f,t)_{L^2(R_+,\,d\mu_\alpha)} = \|(f *_B p_t) - f\|_{L^2(R_+,\,d\mu_\alpha)}, t > 0. \tag{43}$$

Take (43) into (42). Then

$$I(f,\, \gamma)_{L^2(R_+,\,d\mu_\alpha)} \leq C\|(f *_B p_t) - f\|_{L^2(R_+,\,d\mu_\alpha)}\Big|_{t=\frac{\|f\|_{L^2(R_+,\,d\mu_\alpha)}}{\gamma}}, \; f \in L^2(R_+,\, d\mu_\alpha). \tag{44}$$

(44) shows that the decay of $I(f,\gamma)_{L^2(R_+,\,d\mu_\alpha)}$ is controlled by the approximation order of convolution operator $f *_B p_t$ for $t = \frac{\|f\|_{L^2(R_+,\,d\mu_\alpha)}}{\gamma}$.

For $\mathcal{F}_B^{(\alpha)}(\phi)(\lambda) = \frac{1}{\lambda^4}$ we define

$$\begin{aligned} \mathcal{H}_{K^\sharp(\phi)} &= L^{\frac{1}{2}}_{K^\sharp(\phi)}(A_*(R)) \\ &= \{g \in A_*(R) : \left(\int_{R_+} \lambda^4 \left|\mathcal{F}_B^{(\alpha)}(f)(\lambda)d\mu_\alpha\right|^2\right)^{\frac{1}{2}} < +\infty\}. \end{aligned} \tag{45}$$

Then

$$\begin{aligned} \|g\|_{K^\sharp(\phi)} &= \left(\int_{R_+} \lambda^4 \left|\mathcal{F}_B^{(\alpha)}(f)(\lambda)d\mu_\alpha\right|^2\right)^{\frac{1}{2}} \\ &= \|(-l_\alpha)g\|_{L^2(R_+,\,d\mu_\alpha)}. \end{aligned} \tag{46}$$

Define a $K$-functional on $L^2(R_+,\, d\mu_\alpha)$ as

$$D_{l_\alpha}(f,t)_{L^2(R_+,\,d\mu_\alpha)} = \inf_{g \in \mathcal{H}_{K^\sharp(\phi)}} \left(\|f - g\| + t\|(-l_\alpha)g\|_{L^2(R_+,\,d\mu_\alpha)}\right), \qquad t > 0.$$

Define a modulus of smoothness as

$$\omega_{l_\alpha}(f,t)_{L^2(R_+,\,d\mu_\alpha)} = \|(T_{l_\alpha}(t) - I)f\|_{L^2(R_+,\,d\mu_\alpha)}, \qquad t > 0,$$

where

$$T_{l_\alpha}(t)f(x) = \int_{R_+} e^{-\lambda^2 t}\mathcal{F}_B^{(\alpha)}(f)(\lambda)\, j_\alpha(\lambda x)\, d\mu_\alpha(\lambda).$$

Then we have the following two corollaries.

**Corollary 5.** *There holds*

$$D_{l_\alpha}(f,t)_{L^2(R_+,\, d\mu_\alpha)} \sim \omega_{l_\alpha}(f,t)_{L^2(R_+,\, d\mu_\alpha)}, \quad f \in L^2(R_+,\, d\mu_\alpha), t > 0.$$

**Corollary 6.** *There is a constant $C > 0$ such that*

$$I(f,\, \gamma)_{L^2(R_+,\, d\mu_\alpha)} \le C\omega_{l_\alpha}\left(f,\, \frac{\|f\|_{L^2(R_+,\, d\mu_\alpha)}}{\gamma}\right)_{L^2(R_+,\, d\mu_\alpha)},\ f \in L^2(R_+,\, d\mu_\alpha). \tag{47}$$

Additionally, by Example 2, we know $\mathcal{F}_B^{(\alpha)}(k_t)(\lambda) = e^{-\lambda^2 t}$, which, together with (21), gives

$$\begin{aligned} T_{l_\alpha}(t)f(x) &= \int_{R_+} \mathcal{F}_B^{(\alpha)}(k_t)(\lambda)\mathcal{F}_B^{(\alpha)}(f)(\lambda)\, j_\alpha(\lambda x)\, d\mu_\alpha(\lambda) \\ &= \int_{R_+} \mathcal{F}_B^{(\alpha)}(f *_B k_t)(\lambda)\, j_\alpha(\lambda x)\, d\mu_\alpha(\lambda) \\ &= (f *_B k_t)(x), \qquad x \in R_+, \end{aligned}$$

which, with (47), shows that

$$\omega_{l_\alpha}(f,t)_{L^2(R_+,\, d\mu_\alpha)} = \|(f *_B k_t) - f\|_{L^2(R_+,\, d\mu_\alpha)}, \qquad t > 0. \tag{48}$$

Take (48) into (47), we have

$$I(f,\gamma)_{L^2(R_+,\, d\mu_\alpha)} \le C\|(f *_B k_t) - f\|_{L^2(R_+,\, d\mu_\alpha)}\Big|_{t=\frac{\|f\|_{L^2(R_+,\, d\mu_\alpha)}}{\gamma}},\ f \in L^2(R_+,\, d\mu_\alpha). \tag{49}$$

We know by (49) that the decay of $I(f,\gamma)_{L^2(R_+,\, d\mu_\alpha)}$ is controlled by the approximation order of the convolution operator $f *_B k_t$ for $t = \frac{\|f\|_{L^2(R_+,\, d\mu_\alpha)}}{\gamma}$.

## 5. Proofs

**Proof of Lemma 1.** Formula (11) can be obtained by the orthonormal of $\{j_\alpha^*(\lambda_i x)\}_{i=1}^{+\infty}$. Formula (13) can be seen from [11] or Lemma 1 in [12]. Formula (14) can be seen from [16]. □

**Proof of Lemma 2.** *Proof of (i).* See Proposition 2.III.1 in P51 of [22].

*Proof of (ii).* See Corollary 4.III.2 in P104 and Corollary 4.III.3 in P105 of [22].

*Proof of (iii).* See Theorem 5.III.1 in P127 and Proposition 5.II.4 in P129 of [22].

*Proof of (iv).* See Theorem 5.III.1 in P127 and (5.III.3) in P128 of [22].

*Proof of (v).* See Proposition 5.II.2 in P120 of [22] and (4.III.10) in Proposition 4.III.4 of [22].

*Proof of (vi).* See Theorem 5.II.2 in P126 of [22].

*Proof of (vii).* See (5.III.5) and (5.III.6) in Proposition 5.III.2 in P128,(5.V.2) in P139 of [22], and Proposition 2.2 in [25].

*Proof of (viii).* Formula (27) may be found from (5.II.12) of Proposition 5.II.3 in P122 of [22]; (28) may be found from (4.II.9) of Proposition 4.II.2 in P94 of [22]; (29) may be found from (4.II.8) in P93 of [22]. □

**Proof of Proposition 1.** We show it with the help of Proposition A1 in the Appendix A.

It is easy to see that $T_{K^{(\alpha)}}(t)$ satisfies (A1) and (A2). Simple computations show

$$\begin{aligned} Ef(x) &= \lim_{t\to 0} \frac{T_{K^{(\alpha)}}(t)f(x)-f(x)}{t} \\ &= \sum_{i=1}^{+\infty} a_i(f) \lim_{t\to 0} \frac{(e^{-\frac{t}{\sqrt{h_i}}}-1)}{t} j^*_\alpha(\lambda_i x) \\ &= \sum_{i=1}^{+\infty} \left(-\frac{1}{\sqrt{h_i}}\right) a_i(f)\, j^*_\alpha(\lambda_i x) \end{aligned}$$

and

$$tET_{K^{(\alpha)}}(t)f(x) = \sum_{i=1}^{+\infty} \left(-\frac{t}{\sqrt{h_i}}\right) e^{-\frac{t}{\sqrt{h_i}}}\, a_i(f)\, j^*_\alpha(\lambda_i x).$$

It follows

$$\begin{aligned} \|tET_{K^{(\alpha)}}(t)f\|_{L^2_\alpha} &= \left(\sum_{i=1}^{+\infty} \left|\left(-\frac{t}{\sqrt{h_i}}\right)e^{-\frac{t}{\sqrt{h_i}}}\right|^2 a_i^2(f)\right)^{\frac{1}{2}} \\ &\le \left(\sum_{i=1}^{+\infty} a_i^2(f)\right)^{\frac{1}{2}} = \|f\|_{L^2_\alpha}. \end{aligned} \tag{50}$$

Collecting (50), and (A5), we have (38). □

**Proof of Theorem 1.** Because $h_i \to 0^+ (i \to +\infty)$, defining

$$f^{(\alpha)}_\mu(x) = \sum_{\frac{1}{h_i}<\mu} a_i(f)\, j^*_\alpha(\lambda_i x), \tag{51}$$

we have for any $g \in \mathcal{H}_{K^{(\alpha)}}$ that

$$f(x) - f^{(\alpha)}_\mu(x) = \sum_{\frac{1}{h_i}\ge\mu} a_i(f)\, j^*_\alpha(\lambda_i x) = \sum_{\frac{1}{h_i}\ge\mu} a_i(f-g)\, j^*_\alpha(\lambda_i x) + \sum_{\frac{1}{h_i}\ge\mu} a_i(g)\, j^*_\alpha(\lambda_i x)$$

and

$$\begin{aligned} \|f - f^{(\alpha)}_\mu\|_{L^2_\alpha} &\le \left(\sum_{\frac{1}{h_i}\ge\mu} |a_i(f-g)|^2\right)^{\frac{1}{2}} + \left(\sum_{\frac{1}{h_i}\ge\mu} |a_i(g)|^2\right)^{\frac{1}{2}} \\ &\le \|f-g\|_{L^2_\alpha} + \left(\sum_{\frac{1}{h_i}\ge\mu} \frac{h_i}{h_i}|a_i(g)|^2\right)^{\frac{1}{2}} \\ &\le \|f-g\|_{L^2_\alpha} + \frac{1}{\sqrt{\mu}}\left(\sum_{\frac{1}{h_i}\ge\mu} \frac{1}{h_i}|a_i(g)|^2\right)^{\frac{1}{2}} \\ &\le \|f-g\|_{L^2_\alpha} + \frac{1}{\sqrt{\mu}}\|g\|_{K^{(\alpha)}}. \end{aligned} \tag{52}$$

Since the arbitrariness of $g \in \mathcal{H}_{K^{(\alpha)}}$, we have

$$\|f - f_\mu^{(\alpha)}\|_{L^2_\alpha} \le \inf_{g \in \mathcal{H}_{K^{(\alpha)}}} \left( \|f - g\|_{L^2_\alpha} + \frac{1}{\sqrt{\mu}} \|g\|_{K^{(\alpha)}} \right). \tag{53}$$

Take $h_\mu^{(\alpha)}(x) = \sum_{\frac{1}{h_i} < \mu} \frac{a_i(f)}{\sqrt{h_i}} j_\alpha^*(\lambda_i x)$. Then $f_\mu^{(\alpha)}(x) = L^{\frac{1}{2}}_{K^{(\alpha)}}(h_\mu^{(\alpha)}, x) \in \mathcal{H}_{K^{(\alpha)}}$ and

$$\begin{aligned} \|f_\mu\|_{K^{(\alpha)}} &= \|h_\mu\|_{L^2_\alpha} \\ &= \left( \sum_{\frac{1}{h_i} < \mu} \frac{|a_i(f)|^2}{h_i} \right)^{\frac{1}{2}} \\ &\le \sqrt{\mu} \left( \sum_{\frac{1}{h_i} < \mu} |a_i(f)|^2 \right)^{\frac{1}{2}} \le \sqrt{\mu} \|f\|_{L^2_\alpha}. \end{aligned}$$

Take $\sqrt{\mu}\|f\|_{L^2_\alpha} = \gamma$. Then $\frac{1}{\sqrt{\mu}} = \frac{\|f\|_{L^2_\alpha}}{\gamma}$. By the definition of $I(f, \gamma)_{L^2_\alpha}$, we have (39). □

**Proof of Proposition 2.** It is easy to see that $T_{K(\phi)}(t)$ satisfies (A1) and (A2). Simple computations show

$$\begin{aligned} Ef(x) &= \lim_{t \to 0} \frac{T_{K(\phi)}(t) f(x) - f(x)}{t} \\ &= \lim_{t \to 0} \frac{\int_{R_+} (e^{-\frac{t}{\sqrt{\mathcal{F}_B^{(\alpha)}(\phi)(\lambda)}}} - 1) \mathcal{F}_B^{(\alpha)}(f)(\lambda) j_\alpha(\lambda x) d\mu_\alpha(\lambda)}{t} \\ &= \int_{R_+} \left( -\frac{1}{\sqrt{\mathcal{F}_B^{(\alpha)}(\phi)(\lambda)}} \right) \mathcal{F}_B^{(\alpha)}(f)(\lambda) j_\alpha(\lambda x) \, d\mu_\alpha(\lambda) \end{aligned}$$

and

$$(tET_{K(\phi)}(t)f)(x) = \int_{R_+} \left( -\frac{t}{\sqrt{\mathcal{F}_B^{(\alpha)}(\phi)(\lambda)}} \right) e^{-\frac{t}{\sqrt{\mathcal{F}_B^{(\alpha)}(\phi)(\lambda)}}} \mathcal{F}_B^{(\alpha)}(f)(\lambda) \, j_\alpha(\lambda x) d\mu_\alpha(\lambda).$$

Since $f \in L^2(R_+, \, d\mu_\alpha)$, we know by (26) that $\mathcal{F}_B^{(\alpha)}(\phi)(\cdot) \in L^2(R_+, \, d\mu_\alpha)$. Additionally, since

$$\left| \left( -\frac{t}{\sqrt{\mathcal{F}_B^{(\alpha)}(\phi)(\lambda)}} \right) e^{-\frac{t}{\sqrt{\mathcal{F}_B^{(\alpha)}(\phi)(\lambda)}}} \right| \le 1, \qquad \forall t \ge 0,$$

we know

$$h_t(\cdot) = \left( -\frac{t}{\sqrt{\mathcal{F}_B^{(\alpha)}(\phi)(\cdot)}} \right) e^{-\frac{t}{\sqrt{\mathcal{F}_B^{(\alpha)}(\phi)(\cdot)}}} \mathcal{F}_B^{(\alpha)}(f)(\cdot) \in L^2(R_+, \, d\mu_\alpha).$$

It is easy to see that

$$(tET_{K(\phi)}(t)f)(x) = \mathcal{F}_B^{(\alpha)}(h_t)(x).$$

It follows by (26) again that

$$\begin{aligned}\|tET_{K(\phi)}(t)f\|^2_{L^2(R_+,\, d\mu_\alpha)} &= \|\mathcal{F}_B^{(\alpha)}(h_t)\|_{L^2(R_+,\, d\mu_\alpha)} \\ &= \|h_t\|_{L^2(R_+,\, d\mu_\alpha)} \\ &\le \|\mathcal{F}_B^{(\alpha)}(f)\|_{L^2(R_+,\, d\mu_\alpha)} = \|f\|^2_{L^2(R_+,\, d\mu_\alpha)}. \end{aligned} \tag{54}$$

By the same method, we have

$$\begin{aligned}\|T_{K(\phi)}(t)f\|^2_{L^2(R_+,\, d\mu_\alpha)} &= \int_{R_+}\left(e^{-\frac{t}{\sqrt{\mathcal{F}_B^{(\alpha)}(\phi)(\lambda)}}}\right)^2 \left|\mathcal{F}_B^{(\alpha)}(f)(\lambda)\right|^2 d\mu_\alpha(\lambda) \\ &\le \int_{R_+}\left|\mathcal{F}_B^{(\alpha)}(f)(\lambda)\right|^2 d\mu_\alpha(\lambda) = \|f\|^2_{L^2(R_+,\, d\mu_\alpha)}. \end{aligned} \tag{55}$$

Collect (54), (55) and (A6) we have (40). □

**Proof of Theorem 2.** Define $\Re_{\mu,\lambda} = \{\lambda \in R_+ : \frac{1}{\mathcal{F}_B^{(\alpha)}(\phi)(\lambda)} < \mu\}$ and

$$f_*(x) = \int_{\Re_{\mu,\lambda}} \mathcal{F}_B^{(\alpha)}(f)(\lambda) j_\alpha(\lambda x)\, d\mu_\alpha(\lambda).$$

Then

$$f(x) - f_*(x) = \int_{R_+\backslash\Re_{\mu,\lambda}} \mathcal{F}_B^{(\alpha)}(f)(\lambda) j_\alpha(\lambda x)\, d\mu_\alpha(\lambda).$$

It follows that for any $g \in \mathcal{H}_{K(\phi)}$, there holds

$$\begin{aligned}& f(x) - f_*(x) \\ =& \int_{R_+\backslash\Re_{\mu,\lambda}} \mathcal{F}_B^{(\alpha)}(f-g)(\lambda)\, j_\alpha(\lambda x)\, d\mu_\alpha(\lambda) + \int_{R_+\backslash\Re_{\mu,\lambda}} \mathcal{F}_B^{(\alpha)}(g)(\lambda) j_\alpha(\lambda x)\, d\mu_\alpha(\lambda).\end{aligned}$$

Define the characteristic of $R_+\backslash\Re_{\mu,\lambda}$ as $\chi_{R_+\backslash\Re_{\mu,\lambda}}(\lambda)$. Then

$$\begin{aligned}& f(x) - f_*(x) \\ =& \int_{R_+} \chi_{R_+\backslash\Re_{\mu,\lambda}}(\lambda)\mathcal{F}_B^{(\alpha)}(f-g)(\lambda)\, j_\alpha(\lambda x)\, d\mu_\alpha(\lambda) \\ & + \int_{R_+} \chi_{R_+\backslash\Re_{\mu,\lambda}}(\lambda)\mathcal{F}_B^{(\alpha)}(g)(\lambda)\, j_\alpha(\lambda x)\, d\mu_\alpha(\lambda) \\ =& \mathcal{F}_B^{(\alpha)}(g_\mu)(x) + \mathcal{F}_B^{(\alpha)}(b_\mu)(x), \end{aligned} \tag{56}$$

where

$$g_\mu(\lambda) = \chi_{R_+\backslash\Re_{\mu,\lambda}}(\lambda)\mathcal{F}_B^{(\alpha)}(f-g)(\lambda), \quad b_\mu(\lambda) = \chi_{R_+\backslash\Re_{\mu,\lambda}}(\lambda)\mathcal{F}_B^{(\alpha)}(g)(\lambda).$$

Since $\phi$ satisfies Assumption 1, by (30) we know $g_\mu \in \mathcal{D}_*(R) \subset A_*(R) \subset L^2(R_+, d\mu_\alpha)$. By (26), we have

$$\|\mathcal{F}_B^{(\alpha)}(g_\mu)\|_{L^2(R_+,\, d\mu_\alpha)} = \|g_\mu\|_{L^2(R_+,\, d\mu_\alpha)}. \tag{57}$$

By the same method, we have

$$\|\mathcal{F}_B^{(\alpha)}(b_\mu)\|_{L^2(R_+,\, d\mu_\alpha)} = \|b_\mu\|_{L^2(R_+,\, d\mu_\alpha)}. \tag{58}$$

It follows from (56), (57) and (58) that

$$\begin{aligned}
& \|f-f_*\|_{L^2(R_+,\,d\mu_\alpha)} \\
\leq\ & \|\chi_{R_+\backslash\Re_{\mu,\lambda}}(\cdot)\mathcal{F}_B^{(\alpha)}(f-g)(\cdot)\|_{L^2(R_+,\,d\mu_\alpha)} + \|\chi_{R_+\backslash\Re_{\mu,\lambda}}(\cdot)\mathcal{F}_B^{(\alpha)}(g)(\cdot)\|_{L^2(R_+,\,d\mu_\alpha)} \\
=\ & \left(\int_{R_+\backslash\Re_{\mu,\lambda}}\left|\mathcal{F}_B^{(\alpha)}(f-g)(\lambda)\right|^2 d\mu_\alpha\right)^{\frac{1}{2}} + \left(\int_{R_+\backslash\Re_{\mu,\lambda}}\left|\mathcal{F}_B^{(\alpha)}(g)(\lambda)\right|^2 d\mu_\alpha\right)^{\frac{1}{2}} \\
\leq\ & \left(\int_{R_+}\left|\mathcal{F}_B^{(\alpha)}(f-g)(\lambda)\right|^2 d\mu_\alpha\right)^{\frac{1}{2}} + \left(\int_{R_+\backslash\Re_{\mu,\lambda}}\left|\mathcal{F}_B^{(\alpha)}(g)(\lambda)\right|^2 d\mu_\alpha\right)^{\frac{1}{2}}.
\end{aligned}$$

Since (26), we have by the definition of $\Re_{\mu,\lambda}$ that

$$\begin{aligned}
& \|f-f_*\|_{L^2(R_+,\,d\mu_\alpha)} \\
\leq\ & \|f-g\|_{L^2(R_+,\,d\mu_\alpha)} + \left(\int_{R_+\backslash\Re_{\mu,\lambda}}\frac{\mathcal{F}_B^{(\alpha)}(\phi)(\lambda)}{\mathcal{F}_B^{(\alpha)}(\phi)(\lambda)}\left|\mathcal{F}_B^{(\alpha)}(g)(\lambda)\right|^2 d\mu_\alpha\right)^{\frac{1}{2}} \\
\leq\ & \|f-g\|_{L^2(R_+,\,d\mu_\alpha)} + \left(\max_{\lambda\in R_+\backslash\Re_{\mu,\lambda}}\mathcal{F}_B^{(\alpha)}(\phi)(\lambda)\right)^{\frac{1}{2}}\left(\int_{R_+\backslash\Re_{\mu,\lambda}}\frac{\left|\mathcal{F}_B^{(\alpha)}(g)(\lambda)\right|^2}{\mathcal{F}_B^{(\alpha)}(\phi)(\lambda)}d\mu_\alpha\right)^{\frac{1}{2}} \\
\leq\ & \|f-g\|_{L^2(R_+,\,d\mu_\alpha)} + \left(\max_{\lambda\in R_+\backslash\Re_{\mu,\lambda}}\mathcal{F}_B^{(\alpha)}(\phi)(\lambda)\right)^{\frac{1}{2}}\left(\int_{R_+}\frac{\left|\mathcal{F}_B^{(\alpha)}(g)(\lambda)\right|^2}{\mathcal{F}_B^{(\alpha)}(\phi)(\lambda)}d\mu_\alpha\right)^{\frac{1}{2}} \\
=\ & \|f-g\|_{L^2(R_+,\,d\mu_\alpha)} + \frac{1}{\sqrt{\mu}}\|g\|_{\mathcal{H}_{K(\phi)}}.
\end{aligned}$$

Because of the arbitrariness of $g\in\mathcal{H}_{K(\phi)}$, we have

$$\|f-f_*\|_{L^2(R_+,\,d\mu_\alpha)} \leq \inf_{g\in\mathcal{H}_{K(\phi)}}\left(\|f-g\|_{L^2(R_+,\,d\mu_\alpha)} + \frac{1}{\sqrt{\mu}}\|g\|_{\mathcal{H}_{K(\phi)}}\right). \tag{59}$$

Let $h_*(x) = \int_{\Re_{\mu,\lambda}}\frac{\mathcal{F}_B^{(\alpha)}(f)(\lambda)\, j_\alpha(\lambda x)}{\sqrt{\mathcal{F}_B^{(\alpha)}(\phi)(\lambda)}}d\mu_\alpha$. Then by (20) we have $h_*\in L^2(R_+,\mu_\alpha)$ and

$$f_*(x) = L_{K(\phi)}^{\frac{1}{2}}(h_*,x) = \int_{\Re_{\mu,\lambda}}\mathcal{F}_B^{(\alpha)}(f)(\lambda)\, j_\alpha(\lambda x)\, d\mu_\alpha(\lambda).$$

Therefore, $f_*\in\mathcal{H}_{K(\phi)}$. It follows that

$$\begin{aligned}
\|f_*\|_{K(\phi)} &= \|h_*\|_{L^2(R_+,\,d\mu_\alpha)} \\
&= \left(\int_{\Re_{\mu,\lambda}}\frac{|\mathcal{F}_B^{(\alpha)}(f)(\lambda)|^2}{\mathcal{F}_B^{(\alpha)}(\phi)(\lambda)}d\mu_\alpha\right)^{\frac{1}{2}} \\
&\leq \sqrt{\mu}\|\mathcal{F}_B^{(\alpha)}(f)\|_{L^2(R_+,\,d\mu_\alpha)} = \sqrt{\mu}\|f\|_{L^2(R_+,\,d\mu_\alpha)}.
\end{aligned} \tag{60}$$

Take $\sqrt{\mu}\|f\|_{L^2(R_+,\,d\mu_\alpha)} = \gamma$. Then $\sqrt{\mu} = \frac{\gamma}{\|f\|_{L^2(R_+,\,d\mu_\alpha)}}$. Collecting (60) and (59), together with the definition of $I(f;\gamma)_{L^2(R_+,\,d\mu_\alpha)}$ we arrive at

$$\begin{aligned} I(f;\gamma)_{L^2(R_+,\,d\mu_\alpha)} &\leq \inf_{g\in\mathcal{H}_{K(\phi)}}\left(\|f-g\|_{L^2(R_+,\,d\mu_\alpha)} + \frac{\|f\|_{L^2(R_+,\,d\mu_\alpha)}}{\gamma}\|g\|_{\mathcal{H}_{K(\phi)}}\right) \\ &= D\left(f, \frac{\|f\|_{L^2(R_+,\,d\mu_\alpha)}}{\gamma}\right)_{L^2(R_+,\,d\mu_\alpha)} \\ &\sim \omega\left(f, \frac{\|f\|_{L^2(R_+,\,d\mu_\alpha)}}{\gamma}\right)_{L^2(R_+,\,d\mu_\alpha)}. \end{aligned}$$

□

## 6. Further Discussions

We now give some comments on the results obtained in the present paper.

A more general problem arising from learning theory is to bound the decay rate of the function (see [2])

$$I(a,R) = \inf_{\|g\|_H\leq R}(\|a-b\|), \qquad a\in B, R\to +\infty, \tag{61}$$

where $(B, \|\cdot\|)$ is a Banach space and $(H, \|\cdot\|_H)$ is a dense subspace with $\|b\| \leq \|b\|_H$ for $b \in H$.

It is known that the approximation ability of a function class is determined by the smoothness of its functions. So the decay of $I(a,R)$ is influenced by the smoothness of the functions in $H$.

Smale and Zhou (see [2]) give the first estimate for the decay of (61) in the case that $a \in (B,H)_{\theta,\infty}$, which is a particular Besov space (in fact, it is the interpolation space of $B$ and $H$). This work is improved in [9]. For $B = H^s(R^d)(s>0)$ (the Sobolev space, see [2] for the definition) and the reproducing kernel Hilbert space $H = \mathcal{H}_{K_\sigma}$, Zhou gives an estimate as (see [3])

$$\inf_{\|g\|_{K_\sigma}\leq R}\|f-g\| \leq B_{d,s}(logR)^{-s} \tag{62}$$

if $R \geq A\|f\|_{L^2(R^d)}$, where $K_\sigma$ is the Gaussian kernels

$$K_\sigma(x,y) = \exp\{-\frac{\|x-y\|^2}{\sigma^2}\}, \qquad x,y\in[0,1]^d, \sigma>0.$$

The tools used is the RKHS function interpolation.

It is known that the most commonly used tool in approximation theory is the $K$-functional. The most helpful relation is the strong equivalent relation between a $K$-functional and a corresponding modulus of smoothness (see, for example, [26]). The most commonly used quantity for describing the approximation ability of a function class is the Jackson inequality expressed with a $K$-functional or a modulus of smoothness (see also [26]). As far as we know from the literature, no Jackson inequality has been established for the decay of (6). There is little description for the smoothness of a RKHS. Recent research shows that any RKHS has some smoothness; it can be considered from the view of fractional derivative and orthogonal series and show that the well-known $K$-functional ([27])

$$D_{\mathcal{H}_K}(f,\,\lambda)_{L^2_{\rho_X}(X)} = \inf_{g\in\mathcal{H}_K}(\|f-g\|+\lambda\|g\|_{\mathcal{H}_K}), \qquad \lambda>0, \tag{63}$$

is equivalent to a modulus of smoothness, where $X$ is chosen as some compact sets, for example, $X = S^{d-1} = \{x\in R^d : \|x\|=1\}$ and $X = B^d = \{x\in R^d : \|x\|\leq 1\}$. It is valuable for us to extend these results to the RKHS defined on a noncompact set. The set $X$ used in the present paper is $X = R^1$, which is a noncompact set and has essential properties different from those of a compact set (see, for example, [5]). Moreover, it is the first time that a Jackson inequality is established to describe the decay (6). A advantage of this manuscript is the use of the Bessel series and Bessel transforms, which transforms the RKHS approximation problem into the classical Bessel–Fourier approximation problem and gives the decay rate with Bessel–Fourier approximation skills.

The Jackson inequalities in Theorem 1 and Theorem 2 show that the RKHSs constructed with Bessel series and Bessel transforms have the same approximation as that of the Bessel series and Bessel transforms.

The moduli of smoothness defined in this manuscript are one-order moduli. It is a valuable problem for us to define higher-order moduli of smoothness and show the Jackson inequality to describe the decay of (6).

**Author Contributions:** Formal analysis, B.S.; Investigation, S.W.; Writing—original draft, M.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work is supported partially by the NSF (Project No. 61877039), the NSFC/RGC Joint Research Scheme (Project No. 12061160462 and $N_$CityU102/20) of China, the NSF (Project No. LY19F020013) of Zhejiang Province, the Science and Technology Project in Jiangxi Province Department of Education (Project No. GJJ211334).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

It is known that the moduli of smoothness defined by a semi-group of operators have the same properties as those of the usual moduli of smoothness defined by the difference of the function (see Chapter Two of [28]) and have been used to describe the degree of approximation in approximation theory (see, for example, [27,29–32]). We restate here a proposition for a general strong equivalent relation.

Let $(\mathcal{B}, \|\cdot\|_{\mathcal{B}})$ be a normed linear space, $\left\{T(t) : (\mathcal{B}, and\|\cdot\|_{\mathcal{B}}) \to (\mathcal{B}, \|\cdot\|_{\mathcal{B}})\right\}_{t>0}$ be a strongly continuous semi-group of operators satisfying

$$T(s+t) = T(s)T(t), \qquad \lim_{t\to 0^+} T(t) = I, \tag{A1}$$

and

$$\|T(t)f\|_{\mathcal{B}} \le \|f\|_{\mathcal{B}}, \qquad f \in \mathcal{B}, t > 0. \tag{A2}$$

The infinitesimal generator $E$ is given by

$$Ef = \lim_{t\to 0^+} \frac{T(t)f - f}{t}, \qquad (\text{in } \mathcal{B}), \tag{A3}$$

whenever the limit exists. $\mathcal{D}(E)$ is the domain of $E$. Then we have the following proposition.

**Proposition A1.** (Theorem 5.1 of [33]) *Let $T(t)$ satisfy (A1), (A2) and (A3),*

$$T(t)f \in \mathcal{D}(E) \text{ for all } f \in \mathcal{B}, \tag{A4}$$

*and there exists a positive constant N independent of t and $T(t)$ such that*

$$t\|E\ T(t)\|_{\mathcal{B}} \leq N(N \ \textit{ is a constant independent of } t),\ ET(t): \ \ \mathcal{B} \to \mathcal{B} \textit{ for } t \geq 0, \tag{A5}$$

*Then for $r \in \mathcal{N}$ and $t > 0$, there holds*

$$\omega_r(f,\ t)_{\mathcal{B}} = \|(T(t)-I)^r f\|_{\mathcal{B}} \sim \inf_{g\in\mathcal{D}(E^r)} \left(\|f-g\|_{\mathcal{B}} + t^r\|E^r g\|_{\mathcal{B}}\right) = \mathcal{K}_{E^r}(f,\ t^r)_{\mathcal{B}}, \tag{A6}$$

*where*

$$(T(s)-I)^r f = \sum_{k=1}^{r} \left( \begin{array}{c} r \\ k \end{array} \right)(-1)^{r-k}T(ks)f + (-1)^r f.$$

## References


1. Cucker, F.; Smale, S. On the mathematical foundations of learning. *Bull. Amer. Math. Soc.* **2001**, *39*, 1–49. [CrossRef]
2. Smale, S.; Zhou, D.X. Estimating the approximation error in learning theory. *Anal. Appl.* **2003**, *1*, 17–41. [CrossRef]
3. Zhou, D.X. Density problem and approximation error in learning theory. *Abstr. Appl. Anal.* **2013**, 715683. [CrossRef]
4. Aronszajn, N. Theory of reproducing kernels. *Trans. Amer. Math. Soc.* **1950**, *68*, 337–404. [CrossRef]
5. Sun, H.W. Mercer theorem for RKHS on noncompact sets. *J. Complex.* **2005**, *21*, 337–349. [CrossRef]
6. Cucker, F.; Zhou, D.X. *Learning Theory: An Approximation Theory Viewpoint*; Cambridge University Press: Cambridge, UK, 2007.
7. Ferreiar, J.C.; Menegatto, V.A. Reproducing kernel Hilbert spaces associated with kernels on topological spaces. *Funct. Anal. Appl.* **2012**, *46*, 89–91. [CrossRef]
8. Williams, D. *Probability with Martingales*; Cambridge University Press: Cambridge, UK, 1990.
9. Ye, P.X. *Some Approximation Problems in Learning Theory, Post-Doctoral Research Work Report*; Chinese Academy of Sciences: Bejing, China, 2003.
10. Sun, H.W. Behavior of a functional in learning theory. *Front. Math. China* **2007**, *2*, 455–465. [CrossRef]
11. Abilov, V.A.; Abilova, F.V. Approximation of functions by Fourier-Bessel sums. *Izv. Vyssh. Uchebn. Zaved. Math.* **2001**, *8*, 3–9.
12. Abilov, V.A.; Abilova, F.V.; Kerimov, M.K. Some issues concerning approximation of functions by Fourier-Bessel sums. *Comput. Math. Math. Phy.* **2013**, *53*, 867–873. [CrossRef]
13. Abilov, V.A.; Abilova, F.V.; Kerimov, M.K. Sharp estimates for the convergence rate of Fourier-Bessel series. *Comput. Math. Math. Phy.* **2015**, *55*, 907–916. [CrossRef]
14. Abilov, V.A.; Abilova, F.V.; Kerimov, M.K. On sharp estimates of the convergence of double Fourier-Bessel series. *Comput. Math. Math. Phy.* **2017**, *57*, 1735–1740. [CrossRef]
15. Abilov, V.A.; Kerimov, M.K. Some estimates for the error in mixed Fourier-Bessel expansions of functions of two variables. *Comput. Math. Math. Phy.* **2006**, *46*, 1465–1486. [CrossRef]
16. Hochstadt, H. The mean convergence of Fourier-Bessel series. *SIAM Rev.* **1967**, *9*, 211–218. [CrossRef]
17. Arteaga, C.; Marrero, I. Universal approximation by radial basis function networks of Delsarte translates. *Neural Netw.* **2013**, *46*, 299–305. [CrossRef] [PubMed]
18. Arteaga, C.; Marrero, I. Approximation in weighted $p$-mean by RBF networks of Delsarte translates. *J. Math. Anal. Appl.* **2014**, *414*, 450–460. [CrossRef]
19. Dai, F.; Wang, H.P. Interpolation by weighted Paley-Wiener spaces associated with the Dunkl transform. *J. Math. Anal. Appl.* **2012**, *390*, 556–572. [CrossRef]
20. Marrero, I. Radial basisi function neural networks of Hankel translates as universal approximation. *Anal. Appl.* **2019**, *17*, 897–930. [CrossRef]
21. Vladimirov, V.S. *Equations of Matheamtical Physics*; Marcel Dekker: New York, NY, USA, 1971.
22. Triméche, K. *Generalized Harmonic Analysis and Wavelet Packets*; Gordon and Breach Science Publishers: Singapore, 2001.
23. Sheng, B.H. The weighted norm for some Mercer kernel matrices. *Acta Math. Sci.* **2013**, *33A*, 6–15. (In Chinese)
24. Sheng, B.H.; Zuo, L. Error analysis of the kernel regularized regression based on refined convex losses and RKBSs. *Int. J. Wavelets Multiresolut. Inform. Process* **2021**, *19*, 2150012. [CrossRef]
25. Quadih, S.E.; Daher, R. Estimates for the generalized Fourier-Bessel transform in the space $L^2_{\alpha,n}$. *Internat. J. Math. Model Comput.* **2016**, *6*, 269–275.
26. Ditzian, Z.; Totik, V. *Moduli of Smoothness*; Springer: New York, NY, USA, 1987.
27. Sheng, B.H.; Wang, J.L. On the K-functional in learning theory. *Anal. Appl.* **2020**, *18*, 423–446 [CrossRef]
28. Butzer, P.L.; Berens, H. *Semi-Group of Operators and Approximation*; Springer: New York, NY, USA, 1967.
29. Dai, F.; Ditzian, Z. Strong converse inequality for Poisson sums. *Proc. Amer. Math. Soc.* **2005**, *133*, 2609–2611. [CrossRef]
30. Dai, F.; Ditzian, Z. Cesaro summability and Marchaud inequality. *Constr. Approx.* **2007**, *25*, 73–88. [CrossRef]
31. Ditzian, Z. New moduli of smoothness on the unit ball and other domains, introduction and main properties. *Constr. Approx.* **2014**, *40*, 1–36. [CrossRef]

32. Ditzian, Z. New moduli of smoothness on the unit ball, applications and computability. *J. Approx. Theory* **2014**, *180*, 49–76. [CrossRef]
33. Ditzian, Z.; Ivanov, K.G. Strong converse inequalities. *J. Anal. Math.* **1993**, *61*, 61–111. [CrossRef]