



# Decomposition of spaces of distributions induced by tensor product bases <sup>☆</sup>

Kamen Ivanov <sup>a</sup>, Pencho Petrushev <sup>b,a,\*</sup>, Yuan Xu <sup>c</sup>

<sup>a</sup> *Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria*

<sup>b</sup> *Department of Mathematics, University of South Carolina, Columbia, SC 29208, United States*

<sup>c</sup> *Department of Mathematics, University of Oregon, Eugene, OR 97403-1222, United States*

Received 23 December 2010; accepted 2 June 2012

Available online 14 June 2012

Communicated by I. Rodnianski

---

## Abstract

Rapidly decaying kernels and frames (needlets) in the context of tensor product Jacobi polynomials are developed based on several constructions of multivariate  $C^\infty$  cutoff functions. These tools are further employed to the development of the theory of weighted Triebel–Lizorkin and Besov spaces on  $[-1, 1]^d$ . It is also shown how kernels induced by cross product bases can be constructed and utilized for the development of weighted spaces of distributions on products of multidimensional ball, cube, sphere or other domains.

© 2012 Elsevier Inc. All rights reserved.

*Keywords:* Product spaces; Frames; Triebel–Lizorkin spaces; Besov spaces

---

## 1. Introduction

The purpose of this article is to introduce and study Triebel–Lizorkin and Besov spaces on the  $d$ -dimensional cube  $Q^d = [-1, 1]^d$  with Jacobi weights and discuss the respective spaces

---

<sup>☆</sup> The second author has been supported by NSF Grant DMS-1211528 and the third author by NSF Grant DMS-1106113.

\* Corresponding author at: Department of Mathematics, University of South Carolina, Columbia, SC 29208, United States.

*E-mail addresses:* [kamen@math.bas.bg](mailto:kamen@math.bas.bg) (K. Ivanov), [pencho@math.sc.edu](mailto:pencho@math.sc.edu) (P. Petrushev), [yuan@math.uoregon.edu](mailto:yuan@math.uoregon.edu) (Y. Xu).

on the product domains  $B^{d_2} \times Q^{d_1}$ ,  $B^{d_1} \times B^{d_2}$  with  $B^d$  being the unit ball in  $\mathbb{R}^d$  as well as sets of the form  $Q^{d_1} \times S^{d_2}$ ,  $S^{d_1} \times B^{d_2}$ ,  $Q^{d_1} \times T^{d_2}$ ,  $Q^{d_1} \times \mathbb{R}^{d_2}$ ,  $B^{d_1} \times \mathbb{R}_+^{d_2}$ ,  $Q^{d_1} \times B^{d_2} \times \mathbb{R}_+^{d_3}$ , etc. Here  $S^d$  is the unit sphere in  $\mathbb{R}^{d+1}$ ,  $T^d$  is the simplex in  $\mathbb{R}^d$ , and  $\mathbb{R}_+^d := [0, \infty)^d$ . In short, we are interested in developing the theory of distribution spaces on some products of  $Q^{d_1}$ ,  $B^{d_2}$ ,  $S^{d_3}$ ,  $T^{d_4}$ ,  $\mathbb{R}^{d_5}$ , or  $\mathbb{R}_+^{d_6}$  with weights. There are two important components of such undertaking: (i) the spaces need to be properly defined and (ii) building blocks need to be constructed and used for characterization of the spaces. We maintain that for both tasks tensor product orthogonal bases should be used.

*1.1. The principle distinction between the spaces on  $[-1, 1]^d$  and  $B^d$*

It seems to us natural to introduce weighted smoothness spaces on  $[-1, 1]^d$  or  $B^d$  with weights by means of orthogonal polynomials. However, there is a surprising difference between the orthogonal polynomial expansions on  $[-1, 1]^d$  and  $B^d$  which we would like to describe next.

Let us first briefly review the definition of Triebel–Lizorkin and Besov spaces on  $B^d$ , given in [14]. Denote by  $\mathcal{V}_n^d$  the space of all polynomials of total degree  $n$  which are orthogonal to lower degree polynomials in  $L^2(B^d, w_\mu)$  with weight  $w_\mu(x) := (1 - \|x\|_2^2)^{\mu-1/2}$ . The orthogonal projector  $\text{Proj}_n : L^2(B^d, w_\mu) \mapsto \mathcal{V}_n^d$  can be written in the form

$$(\text{Proj}_n f)(x) = \int_{B^d} f(y) P_n(w_\mu; x, y) w_\mu(y) dy. \tag{1.1}$$

To introduce weighted Triebel–Lizorkin ( $F$ -spaces) and Besov spaces ( $B$ -spaces) on  $B^d$  (see [17,24] for the general idea), let

$$\Phi_0(x, y) := 1 \quad \text{and} \quad \Phi_j(x, y) := \sum_{n=0}^{\infty} \hat{a} \left( \frac{n}{2^{j-1}} \right) P_n(w_\mu; x, y), \quad j \geq 1, \tag{1.2}$$

where  $\hat{a} \in C^\infty[0, \infty)$  is a cutoff function such that  $\text{supp } \hat{a} \subset [\frac{1}{2}, 2]$  and  $|\hat{a}| \geq c > 0$  on  $[3/5, 5/3]$ .

The weighted  $F$ -space  $F_p^{s,q}$  on  $B^d$  with  $s \in \mathbb{R}$ ,  $0 < p < \infty$ ,  $0 < q \leq \infty$ , is defined as the space of all distributions  $f$  on  $B^d$  such that

$$\|f\|_{F_p^{s,q}} := \left\| \left( \sum_{j=0}^{\infty} (2^{sj} |\Phi_j * f(\cdot)|^q)^{1/q} \right) \right\|_{L^p(w_\mu)} < \infty, \tag{1.3}$$

where  $\Phi_j * f(x) := \langle f, \overline{\Phi_j(x, \cdot)} \rangle$  (as in (5.15)). The corresponding scale of weighted Besov spaces  $B_p^{s,q}$  is defined via the (quasi-)norms

$$\|f\|_{B_p^{s,q}} := \left( \sum_{j=0}^{\infty} (2^{sj} \|\Phi_j * f(\cdot)\|_{L^p(w_\mu)})^q \right)^{1/q}. \tag{1.4}$$

We refer the reader to [14] for more detailed account of weighed  $F$ - and  $B$ -spaces on the ball.

A “natural” attempt to introduce Triebel–Lizorkin and Besov spaces on  $[-1, 1]^d$  with weight

$$w_{\alpha,\beta}(x) := \prod_{i=1}^d (1 - x_i)^{\alpha_i} (1 + x_i)^{\beta_i} \tag{1.5}$$

would be to use directly the same idea as above. Namely, for multi-indices  $\alpha, \beta, \nu$  the  $d$ -dimensional tensor product Jacobi polynomials are defined by

$$\tilde{P}_\nu^{(\alpha,\beta)}(x) := \prod_{j=1}^d \tilde{P}_{\nu_j}^{(\alpha_j,\beta_j)}(x_j). \tag{1.6}$$

Set  $\tilde{P}_n^{(\alpha,\beta)}(x, y) := \sum_{|\nu|=n} \tilde{P}_\nu^{(\alpha,\beta)}(x) \tilde{P}_\nu^{(\alpha,\beta)}(y)$  and with  $\hat{a}$  as in (1.2) define

$$\Phi_0(x, y) := 1 \quad \text{and} \quad \Phi_j(x, y) := \sum_{n=0}^\infty \hat{a} \left( \frac{n}{2^{j-1}} \right) \tilde{P}_n^{(\alpha,\beta)}(x, y), \quad j \geq 1, \tag{1.7}$$

which can be viewed as an analogue of the kernels from (1.2).

The next step would be to define weighted Triebel–Lizorkin and Besov spaces on  $[-1, 1]^d$  with weight  $w_{\alpha,\beta}(x)$  exactly as in (1.3) and (1.4) using the kernels  $\Phi_j(x, y)$  from (1.7). Such a definition, however, is completely unacceptable due to the *poor localization* of the kernels  $\Phi_j(x, y)$  from (1.7). As is shown in [11] in the particular case of Legendre or Chebyshev polynomials, kernels of the form (1.7) have no localization whatsoever for some points  $x, y \in [-1, 1]^d$ . In contrast, the kernels  $\Phi_j(x, y)$  from (1.2) decay rapidly away from the main diagonal in  $B^d \times B^d$ . Interestingly enough, the situation is quite the same on the interval [13], sphere [15], simplex [11], and more surprisingly in the context of tensor product Hermite [20] and Laguerre functions [12].

### 1.2. The remedy for the problem

It appears that the tensor product Jacobi polynomials are in a sense of a different nature compared to orthogonal polynomials on the interval, ball or simplex as well as spherical harmonics and tensor product Hermite and Laguerre functions. Truly multivariate cutoff functions need to be employed. Our primary goal in this paper is to identify a natural class of cutoff functions which will enable us to develop a meaningful theory of Triebel–Lizorkin and Besov spaces on  $[-1, 1]^d$  with weight  $w_{\alpha,\beta}(x)$  via tensor product Jacobi polynomials.

The key is to consider multivariate cutoff functions  $\hat{A}$  with dyadic dilations covering the whole spectrum and such that the kernels

$$\Phi_j(x, y) := \sum_{\nu \in \mathbb{N}_0^d} \hat{A} \left( \frac{\nu}{2^{j-1}} \right) \tilde{P}_\nu^{(\alpha,\beta)}(x) \tilde{P}_\nu^{(\alpha,\beta)}(y) \tag{1.8}$$

decay rapidly away from the main diagonal in  $[-1, 1]^d \times [-1, 1]^d$ . It turns out that it suffices to consider compactly supported  $C^\infty$  cutoff functions  $\hat{A} : [0, \infty)^d \mapsto \mathbb{C}$  which obey the following

**First Boundary Condition.** For any  $t \in [0, \infty)^d$  which belongs to the coordinate planes, i.e.  $t$  of the form  $t = (t_1, \dots, t_{k-1}, 0, t_{k+1}, \dots, t_d)$  for some  $1 \leq k \leq d$ ,

$$\frac{\partial^m}{\partial t_k^m} \hat{A}(t) = 0 \quad \text{for } m = 1, 2, \dots \tag{1.9}$$

Sometimes, instead of this condition it will be more convenient to use the following slightly more restrictive but for certain purposes better and easier to deal with

**Second Boundary Condition.** There exists a constant  $c_* > 0$  such that for any  $\tau \in [0, \infty)^d$  of the form  $\tau = (\tau_1, \dots, \tau_{k-1}, 0, \tau_{k+1}, \dots, \tau_d)$ ,  $1 \leq k \leq d$ ,  $\hat{A}(t) = \text{constant}$  for  $t \in [\tau, \tau + c_* e_k]$  with  $e_k$  being the  $k$ th coordinate vector.

The point is that either of these conditions combined with  $\hat{A}$  being  $C^\infty$  and compactly supported yields the rapid decay of the kernels  $\Phi_j(x, y)$  from (1.8) (see Theorem 4.1). Then these kernels can be deployed to the definition of weighted Triebel–Lizorkin and Besov spaces on  $[-1, 1]^d$  by means of norms similar to the norms in (1.3)–(1.4).

As will be seen the weights

$$W_{\alpha,\beta}(n; x) := \prod_{i=1}^d (1 - x_i + n^{-2})^{\alpha_i+1/2} (1 + x_i + n^{-2})^{\beta_i+1/2} \tag{1.10}$$

will appear naturally in most estimates and results related to spaces on  $[-1, 1]^d$  with weight  $w_{\alpha,\beta}(x)$ . Moreover, the inhomogeneity created by  $w_{\alpha,\beta}(x)$  and the boundary of  $[-1, 1]^d$  leads us to the introduction via  $W_{\alpha,\beta}(\cdot; \cdot)$  of a fourth parameter  $\rho$  in the definition of weighted Triebel–Lizorkin and Besov spaces on  $[-1, 1]^d$ . Thus we introduce  $F$ -spaces by the norms

$$\|f\|_{F_{p,q}^{s,\rho}} := \left\| \left( \sum_{j=0}^{\infty} [2^{sj} W_{\alpha,\beta}(2^j; \cdot)^{-\rho/d} |\Phi_j * f(\cdot)|]^q \right)^{1/q} \right\|_{L^p(w_{\alpha,\beta})}$$

and  $B$ -spaces by the norms

$$\|f\|_{B_{p,q}^{s,\rho}} := \left( \sum_{j=0}^{\infty} [2^{sj} \|W_{\alpha,\beta}(2^j; \cdot)^{-\rho/d} \Phi_j * f(\cdot)\|_{L^p(w_{\alpha,\beta})}]^q \right)^{1/q}$$

(see Sections 7–8). This allows to use for different purposes various scales of weighted  $F$ - and  $B$ -spaces on  $[-1, 1]^d$ . For instance, as will be seen the Besov spaces  $B_{\tau\tau}^{s,s}$  appear naturally in nonlinear approximation in  $L^p(w_{\alpha,\beta})$ .

As a next step we use kernels of the form (1.8) for the construction of building blocks (needlets)  $\{\varphi_\xi\}, \{\psi_\xi\}$ . These are multiscale dual frames which enable us to characterize the  $F$ - and  $B$ -space norms by the size of the needlet coefficients  $\{(f, \varphi_\xi)\}$  in appropriate sequence norms. They can be viewed as an analogue of the  $\varphi$ -transform of Frazier and Jawerth [6–8].

The theory of weighted Triebel–Lizorkin and Besov spaces on  $[-1, 1]^d$  and needlet decompositions in dimensions  $d > 1$  run parallel to their theory in dimension  $d = 1$ , developed in [13], and on the ball [14]. Therefore, to spare the reader the repetition of well-established arguments

we shall only exhibit the essential differences and refer for the rest to [13,14]. We shall place the emphasis on the development of multivariable cutoff functions and the associated tensor product Jacobi kernels which differ substantially from the ones in the univariate case and are the main reason for writing this paper. We shall also discuss the main points of the development of Triebel–Lizorkin and Besov spaces and needlets on products of  $[-1, 1]^{d_1}$ ,  $B^{d_2}$ ,  $\mathbb{S}^{d_3}$ ,  $\mathbb{R}^{d_4}$ , or  $\mathbb{R}_+^{d_5}$  with weights as mentioned above.

This paper is part of a broader undertaking for development of spaces of distributions in non-standard settings such as on the sphere [15], ball [14] as well as in the frameworks of Hermite [20] and Laguerre [12] expansions. It is also closely related to the development of sub-exponentially localized Jacobi and other kernels and needlets in [11].

### 1.3. Outline of the paper

A substantial part of the paper is devoted to the development of multivariate cutoff functions and related tensor product Jacobi kernels. In Section 2 we review some basic results from [11,13] and prove new results about admissible univariate cutoff functions and the localization of the respective kernels induced by univariate Jacobi polynomials. In Section 3 we present several constructions of multivariate admissible cutoff functions. We also construct cutoff functions of “small” derivatives which enables us to develop tensor product kernels with sub-exponential localization. In Section 4 the localization results of the corresponding tensor product Jacobi polynomial are established. In Section 5 we give some auxiliary results concerning a maximal operator and distributions on  $[-1, 1]^d$ . We also establish some  $L^p$ -multipliers for tensor product Jacobi polynomial expansions. In Section 6 we utilize kernels associated to cutoff functions of types (b) and (c) to the construction of frame elements (needlets). In Sections 7–8 we further use these kernels to define “correctly” the weighted Triebel–Lizorkin and Besov spaces on  $[-1, 1]^d$  with weight  $w_{\alpha,\beta}(x)$ . We also establish needlet decomposition of the  $F$ - and  $B$ -spaces. Section 9 is devoted to nonlinear approximation from Jacobi needlets. In Section 10 we briefly consider weighted spaces of distributions on  $B^{d_1} \times [-1, 1]^d$ . In Section 11 we discuss various aspects of distribution spaces on product domains and tensor product bases. In Appendix A we place the lengthy proof of a lemma from Section 10.

**Some useful notation.** Throughout the paper (if not specified otherwise) we shall denote

$$\|f\|_p := \left( \int_{[-1,1]^d} |f(x)|^p w_{\alpha,\beta}(x) dx \right)^{1/p}, \quad 0 < p < \infty,$$

and  $\|f\|_\infty := \text{ess sup}_{x \in [-1,1]^d} |f(x)|$ . For  $x \in \mathbb{R}^d$  we shall use the norms  $\|x\| = \|x\|_\infty := \max_i |x_i|$ ,  $\|x\|_2 := (\sum_i |x_i|^2)^{1/2}$ , and  $|x| = \|x\|_1 := \sum_i |x_i|$ .  $\Pi_n^d$  will denote the set of all algebraic polynomials of total degree  $\leq n$  in  $d$  variables. Positive constants will be denoted by  $c$ ,  $c_1$ ,  $c'$ ,  $\dots$  and they may vary at every occurrence,  $a \sim b$  will stand for  $c_1 a \leq b \leq c_2 a$ .

## 2. Localized Jacobi kernels induced by univariate cutoff functions

Here we introduce the notion of admissible univariate cutoff functions and review the localization properties of the associated kernels induced by Jacobi polynomials established in [11,13,18]. We also obtain some new localization results.

2.1. Admissible univariate cutoff functions

**Definition 2.1.** A function  $\hat{a} \in C^\infty[0, \infty)$  is said to be admissible if  $\text{supp } \hat{a} \subset [0, 2]$  and  $\hat{a}^{(m)}(0) = 0$  for  $m \geq 1$ . Furthermore,  $\hat{a}$  is said to be admissible of type (a), (b) or (c) if  $\hat{a}$  is admissible and in addition obeys the respective condition:

- (a)  $\hat{a}(t) = 1, t \in [0, 1]$ ,
- (b)  $\text{supp } \hat{a} \subset [1/2, 2]$ , or
- (c)  $\text{supp } \hat{a} \subset [1/2, 2]$  and  $\sum_{j=0}^\infty |\hat{a}(2^{-j}t)|^2 = 1$  for  $t \in [1, \infty)$ .

We next introduce sets of  $C^\infty$  functions with “small” derivatives. As a tool for measuring the derivatives’ growth we use functions  $\mathcal{L}$  satisfying the conditions:

$$\mathcal{L} : [0, \infty) \mapsto [1, \infty) \text{ is monotone increasing, } \mathcal{L}(0) = 1 \text{ and}$$

$$M = M(\mathcal{L}) := 1 + \int_0^\infty \frac{dt}{(t + 1)\mathcal{L}(t)} < \infty. \tag{2.1}$$

Typical examples of functions  $\mathcal{L}$  satisfying (2.1) are  $\mathcal{L}_{0,\varepsilon}(t) := (1 + t)^\varepsilon, \varepsilon > 0$ , and

$$\mathcal{L}_{\ell,\varepsilon}(t) := \ln(e + t) \cdots \underbrace{\ln \cdots \ln}_{\ell-1}(\underbrace{\exp \cdots \exp}_{\ell-1} 1 + t) \left[ \underbrace{\ln \cdots \ln}_{\ell}(\underbrace{\exp \cdots \exp}_{\ell} 1 + t) \right]^{1+\varepsilon}, \tag{2.2}$$

where  $\ell \in \mathbb{N}$  and  $0 < \varepsilon \leq 1$ . Evidently,  $M(\mathcal{L}_{\ell,\varepsilon}) \leq c(\ell)\varepsilon^{-1}$ .

We shall use the standard notation  $D_j^k := \frac{\partial^k}{\partial x_j^k}$ .

**Definition 2.2.** Let  $\mathcal{L}$  satisfy (2.1). Given constants  $\gamma, \tilde{\gamma} > 0$  and  $d \geq 1$ , we define  $\mathcal{S}(d, \mathcal{L}; \gamma, \tilde{\gamma})$  to be the set of all functions  $\hat{A} \in C^\infty[0, \infty)^d$ , such that  $\|\hat{A}\|_\infty \leq \gamma$  and

$$\frac{1}{k!} \|D_j^k \hat{A}\|_\infty \leq \gamma (\tilde{\gamma} \mathcal{L}(k - 1))^k, \quad \forall k \in \mathbb{N}, 1 \leq j \leq d. \tag{2.3}$$

The next statement asserts the existence of admissible univariate cutoff functions with “small” derivatives.

**Theorem 2.3.** Let  $\mathcal{L}$  and  $M$  be given by (2.1). Then the sets  $\mathcal{S}(1, \mathcal{L}; 1, 2M)$ ,  $\mathcal{S}(1, \mathcal{L}; 2, 4M)$  and  $\mathcal{S}(1, \mathcal{L}; 8, 8M)$  contain admissible cutoff functions  $\hat{a}$  of types (a), (b), and (c), respectively, (see Definition 2.1) with values in  $[0, 1]$ .

**Proof.** We shall proceed quite similarly as in the proof of Theorem 3.1 in [11]. We let  $\chi_\delta := \frac{1}{2\delta} \mathbb{1}_{[-\delta, \delta]}$  and select  $\delta_j := \frac{1}{(j+1)\mathcal{L}(j)}$  for  $j \geq 0$ . Apparently

$$\sum_{j=0}^\infty \delta_j \leq 1 + \int_0^\infty \frac{dt}{(t + 1)\mathcal{L}(t)} = M.$$

We define

$$\varphi_m := \chi_{\delta_0} * \cdots * \chi_{\delta_m} \quad \text{and} \quad \varphi(t) := \lim_{m \rightarrow \infty} \varphi_m(t).$$

Just as in [10, Theorem 1.3.5] we have  $\varphi \in C^\infty$ ,  $\varphi \geq 0$ ,  $\text{supp } \varphi \subset [-M, M]$  and

$$\|\varphi^{(k-1)}\|_\infty \leq \frac{1}{\prod_{j=0}^{k-1} \delta_j} \leq k! \mathcal{L}(k-1)^k \quad \text{for } k \geq 1.$$

Furthermore, since  $\int_{\mathbb{R}} \chi_\delta = 1$ , we have  $\int_{\mathbb{R}} \varphi = 1$  and  $0 \leq \varphi \leq 1/2$ .

We now set  $\psi(t) := 2M\varphi(2Mt)$  and define  $g(t) := \frac{\pi}{2} \int_{-\infty}^t \psi(s) ds$ . Evidently,  $g \in C^\infty(\mathbb{R})$ ,  $\text{supp } g' \subset [-\frac{1}{2}, \frac{1}{2}]$ ,  $g(t) + g(-t) = \frac{\pi}{2}$  for  $t \in \mathbb{R}$ ,  $0 \leq g \leq \pi/2$ ,  $\|g'\|_\infty \leq (\pi/2)\|\psi\|_\infty \leq (\pi/2)M$  and

$$\|g^{(k)}\|_\infty \leq \frac{\pi}{2} \|\psi^{(k-1)}\|_\infty \leq \frac{\pi}{2} (2M)^k k! \mathcal{L}(k-1)^k \quad \text{for } k \geq 2. \tag{2.4}$$

Apparently  $\hat{a}(t) := \frac{2}{\pi} g(\frac{3}{2} - t)$  is an admissible function of type (a) and  $\hat{a}$  belongs to  $\mathcal{S}(1, \mathcal{L}; 1, 2M)$ . Also  $\hat{a}(t) - \hat{a}(2t)$  is an admissible function of type (b) belonging to  $\mathcal{S}(1, \mathcal{L}; 2, 4M)$ .

To construct an admissible function of type (c) we write  $\phi(t) := \sin g(t)$ ,  $t \in \mathbb{R}$ . From above, we have  $\phi(t)^2 + \phi(-t)^2 = 1$  for  $t \in \mathbb{R}$ . We define

$$\hat{a}(t) := \begin{cases} \phi(2t - \frac{3}{2}) & \text{if } t \in [\frac{1}{2}, 1], \\ \phi(\frac{3}{2} - t) & \text{if } t \in (1, 2], \\ 0 & \text{if } \mathbb{R} \setminus [\frac{1}{2}, 2]. \end{cases}$$

We claim that  $\hat{a}$  is an admissible cutoff function of type (c) and  $\hat{a} \in \mathcal{S}(1, \mathcal{L}; 8, 8M)$ . All required conditions on  $\hat{a}$  are trivial to verify but the estimate

$$\frac{1}{k!} \|\hat{a}^{(k)}\|_\infty \leq 8(8M\mathcal{L}(k-1))^k, \quad k \geq 1. \tag{2.5}$$

Let  $t_0 \in (-\frac{1}{2}, \frac{1}{2})$  and set  $g_k(t) := \sum_{j=0}^k \frac{(t-t_0)^j}{j!} g^{(j)}(t_0)$ . It is easy to see that  $\phi^{(k)}(t_0) = [\sin g_k]^{(k)}(t_0)$  and since  $\sin g_k(z)$  is an entire function, by the Cauchy formula,

$$\phi^{(k)}(t_0) = \frac{k!}{2\pi i} \int_{\mathcal{C}} \frac{\sin g_k(z)}{(z-t_0)^{k+1}} dz, \tag{2.6}$$

where  $\mathcal{C} := \{z \in \mathbb{C}: |z-t_0|=r\}$  with  $r = \frac{1}{4M\mathcal{L}(k-1)}$ . By (2.4) we have for  $z \in \mathcal{C}$  and  $k \geq 1$

$$|g_k(z)| \leq \frac{\pi}{2} \left( 1 + \frac{M}{4M\mathcal{L}(k-1)} + \sum_{j=2}^k \frac{(2M)^j j! \mathcal{L}(j-1)^j}{j! [4M\mathcal{L}(k-1)]^j} \right) \leq \frac{\pi}{2} \left( 1 + \frac{1}{4} + \sum_{j=2}^k \frac{1}{2^j} \right) = \frac{7\pi}{8}$$

and hence  $|\sin g_k(z)| \leq (e^{7\pi/8} + e^{-7\pi/8})/2 < 8$  for  $z \in \mathcal{C}$ . From this and (2.6) we get

$$|\phi^{(k)}(t_0)| \leq 8k! [4M\mathcal{L}(k-1)]^k,$$

which implies (2.5).  $\square$

**Remark 2.4.** Theorem 2.3 is sharp in the sense that if  $\int_0^\infty \frac{dt}{(t+1)\mathcal{L}(t)} = \infty$ , then there is no admissible cutoff function  $\hat{a}$  belonging to  $\mathcal{S}(1, \mathcal{L}; \gamma, \tilde{\gamma})$  for any  $\gamma, \tilde{\gamma} > 0$ . The argument is precisely the same as in [11, Remark 3.2].

2.2. Localized kernels induced by Jacobi polynomials

The Jacobi polynomials  $P_n^{(\alpha, \beta)}$ ,  $n = 0, 1, \dots$ , form an orthogonal basis for the weighted space  $L^2([-1, 1], w_{\alpha, \beta})$  with weight  $w_{\alpha, \beta}(t) := (1-t)^\alpha(1+t)^\beta$ . For various technical reasons we shall assume that  $\alpha, \beta \geq -1/2$ . The Jacobi polynomials are traditionally normalized by  $P_n^{(\alpha, \beta)}(1) = \binom{n+\alpha}{n}$ . It is well known that [23, (4.3.3)]

$$\int_{-1}^1 P_n^{(\alpha, \beta)}(t) P_m^{(\alpha, \beta)}(t) w_{\alpha, \beta}(t) dt = \delta_{n,m} h_n^{(\alpha, \beta)},$$

where

$$h_n^{(\alpha, \beta)} = \frac{2^{\alpha+\beta+1}}{(2n + \alpha + \beta + 1)} \frac{\Gamma(n + \alpha + 1)\Gamma(n + \beta + 1)}{\Gamma(n + 1)\Gamma(n + \alpha + \beta + 1)}. \tag{2.7}$$

Hence

$$\tilde{P}_n^{(\alpha, \beta)} = (h_n^{(\alpha, \beta)})^{-1/2} P_n^{(\alpha, \beta)} \tag{2.8}$$

is the  $n$ th degree Jacobi polynomial normalized in  $L^2([-1, 1], w_{\alpha, \beta})$ .

We are interested in kernels of the form

$$L_n^{\alpha, \beta}(x, y) = \sum_{j=0}^\infty \hat{a}\left(\frac{j}{n}\right) \tilde{P}_j^{(\alpha, \beta)}(x) \tilde{P}_j^{(\alpha, \beta)}(y), \tag{2.9}$$

for smooth cutoff functions  $\hat{a} : [0, \infty) \mapsto \mathbb{C}$ .

In [18] (see also [1]) it was proved that the kernels  $L_n^{\alpha, \beta}(x, y)$  decay rapidly away from the main diagonal in  $[-1, 1]^2$  for compactly supported  $C^\infty$  cutoff functions  $\hat{a}$  which are constants around  $t = 0$ . It was also proved in [11] that for such cutoff functions with “small” derivatives the localization of these kernels is sub-exponential. Furthermore, it was shown that the behavior of  $\hat{a}$  at  $t = 0$  plays a critical role for the localization of  $L_n^{\alpha, \beta}(x, y)$ , in particular, the fact that  $\hat{a}$  is  $C^\infty$  and compactly supported does not guarantee rapid decay of the kernels  $L_n^{\alpha, \beta}(x, y)$ .

Here we extend that localization result from [18] to smooth cutoff functions  $\hat{a}$  with multiple zeros of their first derivatives at  $t = 0$ . To give this result we need the quantities:  $w_{\alpha,\beta}(0; x) := 1$  and

$$w_{\alpha,\beta}(n; x) := (1 - x + n^{-2})^{\alpha+1/2} (1 + x + n^{-2})^{\beta+1/2}, \quad n \geq 1. \tag{2.10}$$

We shall also use the distance  $\rho(x, y) := |\arccos x - \arccos y|$  on  $[-1, 1]$ .

**Theorem 2.5.** *Let  $\hat{a} \in C^{3k-1}[0, \infty)$  for some integer  $k \geq 1$ ,  $\text{supp } \hat{a} \subset [0, 2]$ , and  $\hat{a}^{(m)}(0) = 0$  for  $m = 1, 2, \dots, 3k - 1$ . Then there exists a constant  $c > 0$  of the form  $c = c(k, \alpha, \beta) \|\hat{a}^{(3k-1)}\|_\infty$  such that the kernels from (2.9) satisfy*

$$|L_n^{\alpha,\beta}(x, y)| \leq c \frac{n}{\sqrt{w_{\alpha,\beta}(n; x)}\sqrt{w_{\alpha,\beta}(n; y)}} (1 + n\rho(x, y))^{-k}, \quad x, y \in [-1, 1]. \tag{2.11}$$

Consequently, if  $\hat{a}$  is an admissible cutoff function, then the above estimate holds for any  $k \geq 1$ .

As in [18] estimate (2.11) follows by the localization of  $L_n^{\alpha,\beta}(x, 1)$  given in the next theorem. Denote  $\mathcal{Q}_n^{\alpha,\beta}(x) := L_n^{\alpha,\beta}(x, 1)$ . It is readily seen that (see e.g. [18])

$$\mathcal{Q}_n^{\alpha,\beta}(x) = c^* \sum_{j=0}^\infty \hat{a}\left(\frac{j}{n}\right) \frac{(2j + \alpha + \beta + 1)\Gamma(j + \alpha + \beta + 1)}{\Gamma(j + \beta + 1)} P_j^{(\alpha,\beta)}(x), \tag{2.12}$$

where  $c^* := 2^{-\alpha-\beta-1} \Gamma(\alpha + 1)^{-1}$ .

**Theorem 2.6.** *Let  $\hat{a}$  be as in Theorem 2.5 and  $\alpha \geq \beta \geq -1/2$ . Then for any  $r \geq 0$*

$$\left| \frac{d^r}{dx^r} \mathcal{Q}_n^{\alpha,\beta}(\cos \theta) \right| \leq c \frac{n^{2\alpha+2r+2}}{(1 + n\theta)^k}, \quad 0 \leq \theta \leq \pi. \tag{2.13}$$

Here  $c$  is of the form  $c = c(k, r, \alpha) \|\hat{a}^{(3k-1)}\|_\infty$ .

**Proof.** We shall proceed quite similarly as in the proof of Theorem 4.2 in [11] and, therefore, we shall use some notation and facts from that proof.

We shall only prove (2.13) for  $r = 0$ ; then in general (2.13) follows by using Markov’s inequality as in [11].

We trivially have (see (4.8) in [11])  $|\mathcal{Q}_n^{\alpha,\beta}(\cos \theta)| \leq cn^{2\alpha+2}$ , which gives (2.13) ( $r = 0$ ) for  $0 \leq \theta \leq 1/n$ .

The following identity is crucial in estimating  $|\mathcal{Q}_n^{\alpha,\beta}(\cos \theta)|$  [23, (4.5.3)]:

$$\begin{aligned} & \sum_{\nu=0}^n \frac{(2\nu + \alpha + k + \beta + 1)\Gamma(\nu + \alpha + k + \beta + 1)}{\Gamma(\nu + \beta + 1)} P_\nu^{(\alpha+k,\beta)}(x) \\ &= \frac{\Gamma(n + \alpha + k + 1 + \beta + 1)}{\Gamma(n + \beta + 1)} P_n^{(\alpha+k+1,\beta)}(x). \end{aligned} \tag{2.14}$$

We define  $A_0(t) := (2t + \alpha + \beta + 1)\hat{a}(\frac{t}{n})$  and inductively

$$A_{k+1}(t) := \frac{A_k(t)}{2t + \alpha + k + \beta + 1} - \frac{A_k(t+1)}{2t + \alpha + k + \beta + 3}, \quad k \geq 0. \tag{2.15}$$

We apply summation by parts  $k$  times starting from (2.12) and using every time (2.14) and (2.15) to obtain

$$\mathcal{Q}_n^{\alpha, \beta}(x) = c^* \sum_{j=0}^{\infty} A_k(j) \frac{\Gamma(j + \alpha + k + \beta + 1)}{\Gamma(j + \beta + 1)} P_j^{(\alpha+k, \beta)}(x). \tag{2.16}$$

Observe first that  $A_1(t) = \hat{a}(\frac{t}{n}) - \hat{a}(\frac{t+1}{n}) = \frac{1}{n} \int_0^1 \hat{a}'(\frac{t+s}{n}) ds$  and hence  $A_1^{(m)}(t) = \frac{1}{n^{m+1}} \times \int_0^1 \hat{a}^{(m+1)}(\frac{t+s}{n}) ds$ , which leads to

$$|A_1^{(m)}(t)| \leq \frac{1}{n^{m+1}} \|\hat{a}^{(m+1)}\|_{L^\infty[\frac{t}{n}, \frac{t+1}{n}]}$$

On the other hand, since  $\hat{a}^{(\ell)}(0) = 0$  for  $\ell = 1, 2, \dots, 3k - 1$ , then by Taylor’s theorem

$$|\hat{a}^{(m+1)}(z)| \leq \frac{z^{2k-1}}{(2k-1)!} \|\hat{a}^{(2k+m)}\|_{L^\infty[0, z]} \quad \text{whenever } m+1 \leq k, z > 0. \tag{2.17}$$

Therefore,

$$|A_1^{(m)}(t)| \leq \frac{1}{n^{m+1}} \left(\frac{t+1}{n}\right)^{2k-1} \|\hat{a}^{(2k+m)}\|_{L^\infty[0, \frac{t+1}{n}]}, \quad m+1 \leq k, t > 0. \tag{2.18}$$

We next estimate  $|A_l^{(m)}(t)|$  by induction on  $l$ . We claim that

$$|A_l^{(m)}(t)| \leq \frac{c}{(t+1)^{m+2l-1}} \left(\frac{t+l}{n}\right)^{2k-1} \max_{2k \leq \ell \leq 2k+m+l-1} \|\hat{a}^{(\ell)}\|_{L^\infty[0, \frac{t+l}{n}]} \tag{2.19}$$

if  $m+l \leq k, m \geq 0, l \geq 1$ , and  $0 \leq t \leq 2n$ , where  $c = c(l, m)$ , and hence, using (2.17),

$$\begin{aligned} |A_k(t)| &\leq \frac{c(k)}{(t+1)^{2k-1}} \left(\frac{t+k}{n}\right)^{2k-1} \max_{2k \leq \ell \leq 3k-1} \|\hat{a}^{(\ell)}\|_{L^\infty[0, \frac{t+k}{n}]} \\ &\leq \frac{c(k)}{n^{2k-1}} \|\hat{a}^{(3k-1)}\|_{\infty}. \end{aligned} \tag{2.20}$$

Indeed, estimate (2.18) gives (2.19) for  $l = 1$ . Suppose (2.19) holds for some  $l \geq 1$  and all  $m \geq 0$  such that  $m+l \leq k$ . Then by (2.15)  $A_{l+1}(t) = -\int_0^1 G_l'(t+s) ds$  with  $G_l(t) := \frac{A_l(t)}{2t+\alpha+l+\beta+1}$  and hence  $A_{l+1}^{(m)}(t) = -\int_0^1 G_l^{(m+1)}(t+s) ds$ . We have

$$G_l^{(m+1)}(t) = \sum_{v=0}^{m+1} \binom{m+1}{v} A_l^{(v)}(t) \frac{(-2)^{m+1-v} (m+1-v)!}{(2t + \alpha + l + \beta + 1)^{m+2-v}}$$

and using the inductive assumption

$$\begin{aligned} |A_{l+1}^{(m)}(t)| &\leq c \max_{2k \leq \ell \leq 2k+m+l} \|\hat{a}^{(\ell)}\|_{L^\infty[0, \frac{t+l+1}{n}]} \\ &\quad \times \sum_{v=0}^{m+1} \frac{1}{(t+1)^{v+2l-1}} \left(\frac{t+l+1}{n}\right)^{2k-1} \frac{1}{(t+1)^{m+2-v}} \\ &\leq \frac{c}{(t+1)^{m+2l+1}} \left(\frac{t+l+1}{n}\right)^{2k-1} \max_{2k \leq \ell \leq 2k+m+l} \|\hat{a}^{(\ell)}\|_{L^\infty[0, \frac{t+l+1}{n}]}, \end{aligned}$$

which confirms (2.19).

We next prove (2.13) ( $r = 0$ ) for  $1/n \leq \theta \leq \pi/2$ . By (2.7) it readily follows that  $h_n^{(\alpha+k, \beta)} \leq c2^k/n$  and it is well known that (see e.g. (4.18) in [11])

$$|P_n^{(\alpha+k, \beta)}(\cos \theta)| \leq \frac{c}{n^{1/2} \theta^{k+\alpha+1/2}}, \quad 0 < \theta \leq \pi/2.$$

We use the above and (2.20) in (2.16) to obtain for  $1/n \leq \theta \leq \pi/2$

$$|Q_n^{\alpha, \beta}(\cos \theta)| \leq c \|\hat{a}^{(3k-1)}\|_\infty \sum_{j=1}^{2n} \frac{j^{\alpha+k}}{n^{2k-1} j^{1/2} \theta^{k+\alpha+1/2}} \leq c \|\hat{a}^{(3k-1)}\|_\infty \frac{n^{2\alpha+2}}{(n\theta)^{k+\alpha+1/2}}.$$

Hence, estimate (2.13) (with  $r = 0$ ) holds for  $1/n \leq \theta \leq \pi/2$ .

Let  $\pi/2 < \theta \leq \pi - 1/n$ . Similarly as in [11]

$$|P_n^{(\alpha+k, \beta)}(\cos \theta)| \leq c2^k n^\beta, \quad \pi/2 \leq \theta \leq \pi - 1/n.$$

Combining this with (2.16) and (2.20) we get for  $\pi/2 \leq \theta \leq \pi - 1/n$

$$|Q_n^{\alpha, \beta}(\cos \theta)| \leq c \|\hat{a}^{(3k-1)}\|_\infty n^{-2k+1} \sum_{j=1}^{2n} j^{\beta+\alpha+k} \leq c \|\hat{a}^{(3k-1)}\|_\infty \frac{n^{\alpha+\beta+2}}{n^k},$$

which implies (2.13).

In the case  $\pi - 1/n \leq \theta \leq \pi$  estimate (2.13) follows from the above estimate exactly as in [11]. This completes the proof of estimate (2.13) in the case  $r = 0$ .  $\square$

Estimate (2.11) can be improved for admissible cutoff functions which are constant around  $t = 0$  and have “small” derivatives as in Theorem 2.3:

**Theorem 2.7.** *Let  $\mathcal{L}$  and  $M$  be as in (2.1). Suppose  $\hat{a}$  is an admissible cutoff function of type (a), (b) or (c) which belongs to  $\mathcal{S}(1, \mathcal{L}; \gamma, \tilde{\gamma}M)$  for some  $\gamma, \tilde{\gamma} > 0$  (see Theorem 2.3). Then the kernels from (2.9) satisfy*

$$|L_n^{\alpha, \beta}(x, y)| \leq \frac{cn}{\sqrt{w_{\alpha, \beta}(n; x)} \sqrt{w_{\alpha, \beta}(n; y)}} \exp \left\{ -\frac{\tilde{c}n\rho(x, y)}{\mathcal{L}(n\rho(x, y))} \right\} \tag{2.21}$$

for  $x, y \in [-1, 1]$ , where  $\tilde{c} = c'/\tilde{\gamma}M$  with  $c' > 0$  being an absolute constant and  $c$  depends continuously only on  $\alpha, \beta, \gamma, \tilde{\gamma}$  and  $M$ .

In particular, the above result holds for  $\mathcal{L}_{\ell,\varepsilon}(t)$  from (2.2) with  $M = c(\ell)\varepsilon^{-1}$ .

For the proof of this theorem one first uses Theorem 2.3 to prove the following estimate for the kernels from (2.12) with  $\hat{a}$  from above

$$\left| \frac{d^r}{dx^r} \mathcal{Q}_n^{\alpha,\beta}(\cos \theta) \right| \leq cn^{2\alpha+2r+2} \exp\left\{ -\frac{\tilde{c}n\theta}{\mathcal{L}(n\theta)} \right\}, \quad 0 \leq \theta \leq \pi,$$

and then proceeds exactly as in the proof of Theorem 4.1 in [11]. The proofs are nearly identical to the ones in [11] and will be omitted.

**Remark 2.8.** Theorem 2.7 remains true if we require that  $\hat{a} \in \mathcal{S}(1, \mathcal{L}; \gamma, \tilde{\gamma}M)$ ,  $\text{supp } \hat{a} \subset [0, 2]$ , and  $\hat{a}$  be a constant on  $[0, \delta]$  for a fixed  $\delta \in (0, 1)$ . Then  $c$  and  $c'$  will depend on  $\delta$  as well. However, we are not aware if Theorem 2.7 holds for admissible cutoff functions which are not constants around  $t = 0$ . The method of proof of Theorem 2.6 does not give such a result for all admissible cutoff functions.

In [11] results similar to Theorem 2.7 are proved on the sphere, ball, simplex and in the context of Hermite and Laguerre functions with  $\mathcal{L}$  replaced by  $\mathcal{L}_{\ell,\varepsilon}$ . We would like to point out here that with the same proofs these results hold for a general function  $\mathcal{L}$  as above.

We shall need

**Lemma 2.9.** *There exists a constant  $c$  depending only on  $\alpha, \beta$  such that*

$$|\tilde{P}_n^{(\alpha,\beta)}(x)| \leq \frac{c}{\sqrt{w_{\alpha,\beta}(n; x)}}, \quad x \in [-1, 1], \quad n \geq 1. \tag{2.22}$$

**Proof.** For  $x \in [-1 + n^{-2}, 1 - n^{-2}]$ , using

$$w_{\alpha,\beta}(n; x) \leq 2^{\alpha+\beta+1}(1-x)^{\alpha+1/2}(1+x)^{\beta+1/2},$$

we get (2.22) from the inequality

$$\sup_{x \in [-1, 1]} (1-x)^{\alpha+1/2}(1+x)^{\beta+1/2} |\tilde{P}_n^{(\alpha,\beta)}(x)|^2 \leq \frac{2e}{\pi} (2 + \sqrt{\alpha^2 + \beta^2})$$

established in [5, Theorem 1]. For the remaining  $x$  estimate (2.22) follows from above invoking Theorem 8.4.8 in [4, p. 108].  $\square$

The next theorem shows that the kernels  $L_n^{\alpha,\beta}$  from (2.9) are Lip 1 with respect to the distance  $\rho(\cdot, \cdot)$ .

**Theorem 2.10.** *Let  $\hat{a} \in C^{3k-1}[0, \infty)$  for some  $k > 2\alpha + 2\beta + 5$ ,  $\text{supp } \hat{a} \subset [0, 2]$ , and  $\hat{a}^{(r)}(0) = 0$  for  $r = 1, 2, \dots, 3k - 1$ . Then for any  $x, y, \xi \in [-1, 1]$  such that  $\rho(x, \xi) \leq c_*n^{-1}$ ,  $n \geq 1$ ,  $c_* > 0$ , the kernel  $L_n^{\alpha,\beta}$  from (2.9) satisfies*

$$|L_n^{\alpha,\beta}(x, y) - L_n^{\alpha,\beta}(\xi, y)| \leq \frac{cn^2\rho(x, \xi)}{\sqrt{w_{\alpha,\beta}(n; x)}\sqrt{w_{\alpha,\beta}(n; y)}}(1 + n\rho(x, y))^{-\sigma}, \tag{2.23}$$

where  $\sigma = k - 2\alpha - 2\beta - 5$  and  $c$  depends only on  $k, \alpha, \beta, c_*$ , and  $\|\hat{a}^{(3k-1)}\|_\infty$ . Consequently, if  $\hat{a}$  is an admissible cutoff function, then the above estimate holds for any  $\sigma > 0$ .

The proof of this theorem for  $\alpha, \beta > -1/2$  utilizes estimate (2.13) and is identical with the proof of Theorem 2.2 in [13]. The limit cases  $\alpha = -1/2$  or  $\beta = -1/2$  are treated as in the proof of [11, Theorem 4.1]. We omit the details.

### 3. Multivariate cutoff functions

As was explained in the introduction, cutoff functions in  $d$ -variables will play a prominent role in the development of weighted  $F$ - and  $B$ -spaces on  $[-1, 1]^d$ . In this section we introduce two kinds of admissible  $d$ -dimensional cutoff functions and give several constructions of such functions.

#### 3.1. Admissible $d$ -dimensional cutoff functions

To define multivariate cutoff functions we need to introduce some convenient notation. Given  $1 \leq k \leq d$  we define  $\text{proj}_k : \mathbb{R}^d \rightarrow \mathbb{R}^d$  by

$$\text{proj}_k(t_1, \dots, t_d) := (t_1, \dots, t_{k-1}, 0, t_{k+1}, \dots, t_d). \tag{3.1}$$

We also denote by  $\mathcal{B}_p$  the part of the unit ball of the standard  $\ell^p(\mathbb{R}^d)$  norm contained in the first octant, i.e.

$$\mathcal{B}_p := \{t \in [0, \infty)^d : \|t\|_p \leq 1\}, \quad 1 \leq p \leq \infty.$$

**Definition 3.1.** A cutoff function  $\hat{A} \in C^\infty[0, \infty)^d$  is said to be *admissible of first kind* or simply *admissible* if  $\text{supp } \hat{A} \subset [0, 2]^d$  and  $\hat{A}$  obeys the First Boundary Condition, introduced in Section 1.2, i.e. for any  $t \in [0, \infty)^d$  of the form  $t = \text{proj}_k t$  for some  $1 \leq k \leq d$  we have  $D_k^m \hat{A}(t) = 0$  for  $m = 1, 2, \dots$

Furthermore,  $\hat{A}$  is said to be of type (a), (b), or (c) if in addition

- (a)  $\hat{A}(t) = 1$  for  $t \in \mathcal{B}_1$ ,
- (b)  $\hat{A}(t) = 0$  for  $t \in \frac{1}{2}\mathcal{B}_1$ , or
- (c)  $\hat{A}$  is of type (b) and  $\sum_{j=0}^\infty |\hat{A}(2^{-j}t)|^2 = 1$  for  $t \in [0, \infty)^d \setminus [0, 1)^d$ .

**Definition 3.2.** A cutoff function  $\hat{A} \in C^\infty[0, \infty)^d$  is said to be *admissible of second kind* and type (a), (b), or (c) if  $\text{supp } \hat{A} \subset [0, 2]^d$ ,  $\hat{A}(t) = \hat{A}(\text{proj}_k t)$  for every  $t = (t_1, \dots, t_d) \in [0, \infty)^d$  such that  $t_k \leq \frac{1}{2}\|t\|_\infty$  and

- (a)  $\hat{A}(t) = 1$  if  $t \in \mathcal{B}_1$ ,
- (b)  $\hat{A}(t) = 0$  if  $t \in \frac{1}{2}\mathcal{B}_1$ , or
- (c)  $\hat{A}$  is of type (b) and  $\sum_{j=0}^\infty |\hat{A}(2^{-j}t)|^2 = 1$  for  $t \in [0, \infty)^d \setminus [0, 1)^d$ .

We first note that if  $\hat{A}$  is admissible of second kind, then for all  $\tau \in [0, \infty)^d$  such that  $\tau = \text{proj}_k \tau$  for some  $1 \leq k \leq d$ , the function  $\hat{A}(t)$  is a constant on the segment  $t \in [\tau, \tau + \frac{1}{4d-2}e_k]$ , i.e.  $\hat{A}$  obeys the Second Boundary Condition from Section 1.2 with constant  $c_* = \frac{1}{4d-2}$ . Indeed, if  $\|\tau\|_\infty \leq \frac{1}{2d-1}$ , then

$$\|t\|_1 = \|\tau\|_1 + t_k \leq \frac{d-1}{2d-1} + \frac{1}{4d-2} = \frac{1}{2}$$

and  $\hat{A}(t) = \hat{A}(\tau)$  because  $\hat{A}(t) = 1$  or  $\hat{A}(t) = 0$  for  $t \in \frac{1}{2}\mathcal{B}_1$ . If  $\|\tau\|_\infty \geq \frac{1}{2d-1}$ , then  $t_k \leq \frac{1}{4d-2} \leq \frac{1}{2}\|\tau\|_\infty = \frac{1}{2}\|t\|_\infty$  and we have  $\hat{A}(t) = \hat{A}(\tau)$  by the definition. Consequently, any admissible cutoff function of second kind is admissible of first kind as well.

Note also that in dimension  $d = 1$  the set of admissible cutoff functions of first kind and type (a) coincides with the set of admissible cutoff functions of second kind and type (a); the same is true for cutoff functions of types (b) and (c).

**Remark 3.3.** As was explained in the introduction the fact that the admissible cutoff functions satisfy the First Boundary Condition (see Section 1.2) is crucial for the rapid decay of the associated tensor product Jacobi polynomial kernel from (2.9); this will be established in the next section.

An important reason for introducing admissible cutoff functions of second kind is that such cutoff functions with “small” derivatives (Section 3.4, Section 3.6) allow to construct tensor product Jacobi polynomial kernel of sub-exponential localization (see Theorems 4.2 and 10.5), while as for now we are unable to achieve such localization with admissible cutoff functions of first kind.

It is easy to construct admissible cutoff functions of type (a) as products of univariate cutoff functions of type (a).

**Lemma 3.4.** *Let  $\hat{a}_j, j = 1, \dots, d$ , be admissible univariate functions of type (a). Then*

$$\hat{A}(t) = \prod_{j=1}^d \hat{a}_j(t_j) \tag{3.2}$$

*is an admissible  $d$ -dimensional cutoff function of second kind and type (a).*

**Proof.** By the definition evidently  $\hat{A} \in C^\infty[0, \infty)^d$ ,  $\text{supp } \hat{A} \subset 2\mathcal{B}_\infty$  and  $\hat{A}(t) = 1$  if  $t \in \mathcal{B}_\infty \supset \mathcal{B}_1$ . Furthermore,  $\hat{A}(t) = \hat{A}(\text{proj}_k t)$  for all  $t \in [0, \infty)^d$  such that  $\|t\|_\infty < 2$  and  $t_k \leq 1$ . This and  $\text{supp } \hat{A} \subset 2\mathcal{B}_\infty$  imply  $\hat{A}(t) = \hat{A}(\text{proj}_k t)$  for all  $t = (t_1, \dots, t_d) \in [0, \infty)^d$  such that  $t_k \leq \frac{1}{2}\|t\|_\infty$ .  $\square$

The construction of admissible cutoff functions of type (b) is straightforward using admissible cutoff functions of type (a):

**Lemma 3.5.** *If  $\hat{A}_1, \hat{A}_2$  are admissible of type (a) (any kind), then*

$$\hat{A}(t) = \hat{A}_1(t) - \hat{A}_2(2t) \tag{3.3}$$

*is admissible of type (b).*

For the definition of  $F$ - and  $B$ -spaces on  $[-1, 1]^d$  we shall utilize admissible cutoff functions  $\hat{A}$  of type (b) with the property that the dyadic dilations of  $\text{supp } \hat{A}$  essentially cover the whole spectrum. More precisely, we shall need admissible functions  $\hat{A}$ , which obey the following *dyadic covering condition*:

for any  $t \in [0, \infty)^d$  with  $\|t\|_\infty = 1$  there is  $0 < \gamma \leq 1$  such that

$$\inf_{\lambda \in [\gamma, 2\gamma]} |\hat{A}(\lambda t)| \geq c > 0. \tag{3.4}$$

Note that this condition yields  $\sum_{j=0}^\infty |\hat{A}(2^{-j}t)| \geq c > 0$  for  $t \in [0, \infty)^d \setminus \mathcal{B}_\infty$  which justifies our terminology.

From the constructions of admissible cutoff functions below it will be clear that it is easy to construct admissible functions  $\hat{A}$  of type (b) which satisfy condition (3.4).

The following lemma will be instrumental in the development of  $F$ - and  $B$ -spaces.

**Lemma 3.6.** *For any admissible function  $\hat{A}$  of first or second kind and type (b) satisfying the dyadic covering condition (3.4) there exists an admissible function  $\hat{B}$  of type (b) (and the same kind) such that*

$$\sum_{j=0}^\infty \overline{\hat{A}(2^{-j}t)} \hat{B}(2^{-j}t) = 1 \quad \text{for } t \in [0, \infty)^d \setminus \mathcal{B}_\infty. \tag{3.5}$$

**Proof.** We shall only prove this lemma for an admissible function of second kind, since the case of first kind cutoff functions is easier.

We define  $\hat{B}(t) := 0$  for  $t \in \frac{1}{2}\mathcal{B}_1$  and  $t \in [0, \infty)^d \setminus 2\mathcal{B}_\infty$ . For the remaining  $t \in [0, \infty)^d$  we set

$$\hat{B}(t) := \frac{\hat{A}(t)}{\sum_{j=-\infty}^\infty |\hat{A}(2^{-j}t)|^2}. \tag{3.6}$$

For every  $t \in [0, \infty)^d$  the sum in the denominator of (3.6) is non-zero on account of property (3.4) and contains no more than  $2 + \log_2 d$  non-zero terms. Hence  $\hat{B} \in C^\infty[0, \infty)^d$ . On the other hand, for  $t \in [0, \infty)^d \setminus \mathcal{B}_\infty$  we have  $2^{-j}t \notin 2\mathcal{B}_\infty$  for  $j < 0$  and the sum in the denominator of (3.6) reduces to  $j \geq 0$ . Hence (3.5) is trivially satisfied. Finally, if  $t \in [0, \infty)^d$  and  $t_k \leq \frac{1}{2}\|t\|_\infty$  for some  $1 \leq k \leq d$ , then  $2^{-j}t_k \leq \frac{1}{2}\|2^{-j}t\|_\infty$  for  $j \in \mathbb{Z}$  and  $\hat{A}(2^{-j}t) = \hat{A}(\text{proj}_k(2^{-j}t))$ , which implies  $\hat{B}(t) = \hat{B}(\text{proj}_k t)$ .  $\square$

The construction of admissible cutoff functions of type (c) will require some care. We shall give several constructions of cutoff functions below.

3.2. Construction of admissible cutoff functions via quasi-norms

One approach for constructing admissible  $d$ -dimensional cutoff functions is based on the following lemma.

**Lemma 3.7.** *Suppose the function  $\mathcal{N} : \mathbb{R}^d \rightarrow \mathbb{R}$  is in  $C^\infty(\mathbb{R}^d \setminus \{0\})$  and for  $t \in \mathbb{R}^d$  obeys*

$$\mathcal{N}(\alpha t) = \alpha \mathcal{N}(t), \quad \alpha > 0, \tag{3.7}$$

$$\|t\|_\infty \leq \mathcal{N}(t) \leq \|t\|_1, \tag{3.8}$$

$$\mathcal{N}(t) = \mathcal{N}(\text{proj}_m t) \quad \text{provided } |t_m| \leq \frac{1}{2} \|t\|_\infty, \quad m = 1, \dots, d. \tag{3.9}$$

If  $\hat{a}$  is an admissible univariate function of type (a), (b) or (c), then  $\hat{A}(t) = \hat{a}(\mathcal{N}(t))$  is an admissible  $d$ -dimensional function of second kind and type (a), (b) or (c), respectively.

The proof of this lemma is straightforward.

A simple way to construct a function  $\mathcal{N}$  satisfying the conditions of Lemma 3.7 is the following. Let  $\hat{c}$  be an even real-valued function, whose restriction on  $[0, \infty)$  is an admissible univariate function of type (a), satisfying  $0 \leq \hat{c} \leq 1$ . For  $d \in \mathbb{N}$  and  $t \in \mathbb{R}^d \setminus \{0\}$  set

$$\mathcal{N}(t) := \sum_{m=1}^d |t_m| \prod_{j=1}^d \hat{c}\left(\frac{t_j}{t_m}\right), \tag{3.10}$$

where  $\hat{c}(\frac{\tau}{0}) := 0$  for every real  $\tau$ , including  $\tau = 0$ . For  $t = 0$  by continuity we set  $\mathcal{N}(0) = 0$ .

Given  $t \neq 0$  let  $k$  be such that  $\|t\|_\infty = |t_k|$ . If  $|t_m| \leq \frac{1}{2} \|t\|_\infty$ , then  $|t_k/t_m| \geq 2$  and hence  $\hat{c}(\frac{t_k}{t_m}) = 0$ . Observing also that  $\hat{c}(\frac{t_j}{t_m}) = 1$  if  $|t_j| \leq |t_m|$  we see that (3.10) can be rewritten as

$$\mathcal{N}(t) = \sum_{1 \leq m \leq d, |t_m| > \frac{1}{2} \|t\|_\infty} |t_m| \prod_{1 \leq j \leq d, |t_j| > |t_m|} \hat{c}\left(\frac{t_j}{t_m}\right). \tag{3.11}$$

As indicated the  $m$ th term in the above sum vanishes if  $|t_m| \leq \frac{1}{2} \|t\|_\infty$  and hence it belongs to  $C^\infty(\mathbb{R}^d \setminus \{0\})$ . Therefore,  $\mathcal{N} \in C^\infty(\mathbb{R}^d \setminus \{0\})$ . From (3.10) it readily follows that  $\mathcal{N}$  satisfies condition (3.7). If  $m$  is such that  $|t_m| \leq \frac{1}{2} \|t\|_\infty$ , then  $|t_m|$  does not participate in the right-hand side of (3.11) and hence  $\mathcal{N}$  satisfies (3.9). The inequality  $\mathcal{N}(t) \leq \|t\|_1$  follows from  $0 \leq \hat{c} \leq 1$ . Finally, if  $\|t\|_\infty = |t_k|$ , then from (3.11) and  $0 \leq \hat{c} \leq 1$  we get  $\mathcal{N}(t) \geq |t_k| = \|t\|_\infty$  and thus (3.8) is also satisfied. Thus, we have proved

**Corollary 3.8.** *Let  $\mathcal{N}$  be given by (3.10), where  $\hat{c}$  is an even real-valued function, whose restriction on  $[0, \infty)$  is an admissible univariate function of type (a), satisfying  $0 \leq \hat{c} \leq 1$ . If  $\hat{a}$  is an admissible univariate function of type (a), (b) or (c), then  $\hat{A}(t) = \hat{a}(\mathcal{N}(t))$  is an admissible  $d$ -dimensional function of second kind and type (a), (b) or (c), respectively.*

### 3.3. Construction of admissible $d$ -dimensional cutoff functions via norms

From (3.10)–(3.11) it follows that  $\mathcal{N}$  is a quasi-norm. A necessary and sufficient condition for  $\mathcal{N}$  to be a norm is the convexity of the unit ball  $\mathcal{B} = \{t: \mathcal{N}(t) \leq 1\}$ .

The construction of the boundary  $\partial\mathcal{B} = \{t: \mathcal{N}(t) = 1\}$  of the unit ball of a norm  $\mathcal{N}$  satisfying the conditions of Lemma 3.7 can be carried out by induction on the dimension. First, one gets the boundaries of the  $d - 1$ -dimensional unit balls on every coordinate hyperplane. Second, one extends them into the first octant by line segments of length  $\frac{1}{2}$ . Third, one completes the surface of the unit ball boundary in the first octant by convex  $C^\infty$  blending. Finally, one extends it by symmetry to the remaining octants and defines the norm from the ball in a standard way.

If instead of convex  $C^\infty$  blending in the above scheme it is used a  $C^\infty$  blending satisfying (3.8), then one obtains a quasi-norm  $\mathcal{N}$  satisfying all conditions of Lemma 3.7. We shall not further elaborate on this construction.

### 3.4. Construction of admissible cutoff functions via quasi-norms with “small” derivatives

In analogy to Theorem 2.3 we construct here admissible  $d$ -dimensional cutoff functions with “small” derivatives. In this construction we shall utilize classes of  $C^\infty$  functions of this type:

$$\mathcal{R}(a, b, F; \gamma, \tilde{\gamma}) := \left\{ f \in C^\infty[a, b]: \frac{1}{k!} \|f^{(k)}\|_{L^\infty[a, b]} \leq \gamma (\tilde{\gamma} F(k))^k, \forall k \in \mathbb{N} \right\},$$

where  $F$  is a given positive non-decreasing function defined at least on  $\mathbb{N}$ , and  $\gamma, \tilde{\gamma} > 0$  are parameters independent of  $k$ . Obviously, the sum and the product of two functions from such classes also belong to a class like that (as the parameters  $\gamma, \tilde{\gamma}$  may vary). More importantly, the composition of two functions also belongs to such a class as the following lemma shows.

**Lemma 3.9.** *Let  $F(v) > 0$  for  $v \in [1, \infty)$  and let  $\ln F(v)^v$  be convex on  $[1, \infty)$ . If  $f \in \mathcal{R}(a_1, b_1, F; \gamma_1, \tilde{\gamma}_1)$ ,  $g \in \mathcal{R}(a_2, b_2, F; \gamma_2, \tilde{\gamma}_2)$  and the range of  $f$  is in  $[a_2, b_2]$ , then  $g \circ f \in \mathcal{R}(a_1, b_1, F; \gamma_2, \tilde{\gamma}_1(\gamma_1 \tilde{\gamma}_2 F(1) + 1))$ . Here  $g \circ f$  is the composition of  $g$  and  $f$ .*

**Proof.** In order to find an estimate for  $D^k(g \circ f)$  we apply Faà di Bruno’s formula in the form

$$\frac{1}{k!} D^k(g \circ f) = \sum_{m \in \mathcal{M}_k} \frac{|m|!}{m_1! \dots m_k!} \frac{(D^{|m|}g) \circ f}{|m|!} \prod_{j=1}^k \left( \frac{D^j f}{j!} \right)^{m_j}, \tag{3.12}$$

where  $\mathcal{M}_k = \{m \in \mathbb{N}_0^k: \sum_{j=1}^k jm_j = k\}$ . Note that  $\sum_{j=1}^k jm_j = k$  implies that at most  $O(\sqrt{k})$  of  $m_j$  can be non-zero. We assume that  $(\cdot)^0 = 1$  in the product in (3.12) even if the argument is zero.

Applying the estimates on the derivatives of  $f$  and  $g$  we get from (3.12)

$$\begin{aligned} \frac{1}{k!} \|D^k(g \circ f)\|_\infty &\leq \sum_{m \in \mathcal{M}_k} \frac{|m|!}{m_1! \dots m_k!} \gamma_2 (\tilde{\gamma}_2 F(|m|))^{|m|} \prod_{j=1}^k \gamma_1^{m_j} (\tilde{\gamma}_1 F(j))^{jm_j} \\ &= \gamma_2 \tilde{\gamma}_1^k \sum_{n=1}^k (\gamma_1 \tilde{\gamma}_2)^n F(n)^n \sum_{m \in \mathcal{M}_k, |m|=n} \frac{n!}{m_1! \dots m_k!} \prod_{j=1}^k F(j)^{jm_j}. \end{aligned} \tag{3.13}$$

Now, from the convexity of  $v \ln F(v)$  we get

$$\ell \ln F(\ell) + j \ln F(j) \leq \ln F(1) + (\ell + j - 1) \ln F(\ell + j - 1), \quad \forall \ell, j > 1. \tag{3.14}$$

If in a multi-index  $m$  we increase  $m_1$  and  $m_{\ell+j-1}$  by 1 and decrease  $m_\ell$  and  $m_j$  by 1, then the quantities  $\sum_{j=1}^k j m_j$  and  $\sum_{j=1}^k m_j$  remain unchanged. Observing that this operation decreases  $\sum_{j=2}^k m_j$  by 1 and applying inductively (3.14) we obtain that among all  $m \in \mathcal{M}_k$  with  $|m| = n$  the largest value of the product  $\prod_{j=1}^k F(j)^{j m_j}$  is attained for  $m_1 = n - 1$ ,  $m_{k-n+1} = 1$ , and  $m_j = 0$  if  $j \neq 1$  and  $j \neq k - n + 1$ , i.e.

$$\prod_{j=1}^k F(j)^{j m_j} \leq F(1)^{n-1} F(k - n + 1)^{k-n+1}. \tag{3.15}$$

For the rigorous proof of estimate (3.15) one should take into account the following properties of the multi-indices  $m \in \mathcal{M}_k$  with  $|m| = n$ : (i) if  $m_j \geq 1$ , then  $j \leq k - n + 1$ ; (ii) if  $\sum_{j=2}^k m_j = 1$ , then  $m_{k-n+1} = 1$ ; (iii) if  $m_j \geq 1$  and  $m_\ell \geq 1$  for  $1 < j < \ell$ , then  $j + \ell - 1 \leq k - n + 1$ ; (iv) if  $m_j \geq 2$  for  $1 < j$ , then  $2j - 1 \leq k - n + 1$ . Note also that (3.15) is trivial for  $n = k$ .

Using (3.15) and

$$\sum_{m \in \mathcal{M}_k, |m|=n} \frac{n!}{m_1! \dots m_k!} = \binom{k-1}{n-1}$$

(see e.g. [21, Section 5.5]) in (3.13) and further applying (3.14) with  $\ell = n$ ,  $j = k - n + 1$  we finally get

$$\begin{aligned} \frac{1}{k!} \|D^k(g \circ f)\|_\infty &\leq \gamma_2 \tilde{\gamma}_1^k \sum_{n=1}^k (\gamma_1 \tilde{\gamma}_2)^n F(n)^n \binom{k-1}{n-1} F(1)^{n-1} F(k-n+1)^{k-n+1} \\ &\leq \gamma_2 \tilde{\gamma}_1^k F(k)^k \sum_{n=1}^k \binom{k-1}{n-1} F(1)^n (\gamma_1 \tilde{\gamma}_2)^n \\ &= \gamma_2 \tilde{\gamma}_1^k F(k)^k \gamma_1 \tilde{\gamma}_2 F(1) (\gamma_1 \tilde{\gamma}_2 F(1) + 1)^{k-1} \\ &\leq \gamma_2 [\tilde{\gamma}_1 (\gamma_1 \tilde{\gamma}_2 F(1) + 1) F(k)]^k. \quad \square \end{aligned}$$

We shall utilize Lemma 3.9 to the composition of admissible cutoff functions with “small” derivatives in the sense of Definition 2.2 (see Theorem 2.3), where  $\mathcal{L}$  obeys an additional *convexity condition*. Namely, we shall assume that

$$\mathcal{L} \text{ satisfies (2.1) and } (t + 1) \ln \mathcal{L}(t) \text{ is convex on } [0, \infty). \tag{3.16}$$

The functions  $\mathcal{L}_{0,\varepsilon}$  and  $\mathcal{L}_{\ell,\varepsilon}$  from (2.2) are examples of functions  $\mathcal{L}$  satisfying this condition.

**Theorem 3.10.** *Let  $\mathcal{L}$  satisfy (3.16) and let  $M$  be given by (2.1). Then the set*

$$S(d, \mathcal{L}; 8, 10d(2d - 1)M(8(d + 2)M + 1))$$

*contains an admissible cutoff function  $\hat{A}$ ,  $0 \leq \hat{A} \leq 1$ , of second kind and any type: (a), (b) or (c) (see Definition 3.2).*

**Proof.** Set  $\hat{A}(t) = \hat{a}(\mathcal{N}(t))$ , where  $\hat{a} \in S(1, \mathcal{L}; 8, 8M)$ ,  $0 \leq \hat{a} \leq 1$ , is an admissible function of any type (a), (b) or (c) from Theorem 2.3 and  $\mathcal{N} = \mathcal{N}_d$  is given by (3.10) with  $\hat{c}|_{[0, \infty)} \in S(1, \mathcal{L}; 1, 2M)$ ,  $0 \leq \hat{c} \leq 1$ , being an admissible function of type (a) from Theorem 2.3. Then  $\hat{A}$  is an admissible multivariate cutoff function of the same type as  $\hat{a}$  according to Corollary 3.8. Moreover,  $0 \leq \hat{A} \leq 1$ .

In estimating  $D_j^k \hat{A}(t_1, \dots, t_d)$  we may assume without loss of generality that  $j = d$ . Further, we consider only

$$\frac{1}{4d - 2} \leq t_d \leq 2, \tag{3.17}$$

because  $D_d^k \hat{A}(t_1, \dots, t_d) = 0$  if  $0 \leq t_d < \frac{1}{4d-2}$  or  $t_d > 2$ .

In order to apply Lemma 3.9 with  $g = \hat{a}$  and  $f = \mathcal{N}$  (as a function of  $t_d$ ) we need upper bounds for  $D_d^k \mathcal{N}(t_1, \dots, t_d)$ . From (3.10) we write

$$\mathcal{N}(t) =: \sum_{m=1}^d F_m(t), \quad F_m(t) := t_m \prod_{j=1}^d \hat{c}\left(\frac{t_j}{t_m}\right).$$

Then for  $m = 1, \dots, d - 1$  we have

$$D_d^k F_m(t) = (D^k \hat{c})\left(\frac{t_d}{t_m}\right) t_m^{1-k} \prod_{j=1}^{d-1} \hat{c}\left(\frac{t_j}{t_m}\right),$$

which on account of (3.17) and since  $D^k \hat{c}(\tau) = 0$  for  $\tau \notin [1, 2]$  and  $0 \leq \hat{c} \leq 1$  implies

$$\frac{1}{k!} |D_d^k F_m(t)| \leq (8(2d - 1)M\mathcal{L}(k - 1))^k \tag{3.18}$$

for all  $(t_1, \dots, t_{d-1}) \in [0, \infty)^{d-1}$ . Using the formulas for derivatives of a product we get

$$\begin{aligned} D_d^k F_d(t) &= t_d D_d^k \left( \prod_{j=1}^{d-1} \hat{c}\left(\frac{t_j}{t_d}\right) \right) + k D_d^{k-1} \left( \prod_{j=1}^{d-1} \hat{c}\left(\frac{t_j}{t_d}\right) \right) \\ &= t_d \sum_{|m|=k} \frac{k!}{m_1! \dots m_{d-1}!} \prod_{j=1}^{d-1} D_d^{m_j} \left( \hat{c}\left(\frac{t_j}{t_d}\right) \right) \\ &\quad + \sum_{|m|=k-1} \frac{k!}{m_1! \dots m_{d-1}!} \prod_{j=1}^{d-1} D_d^{m_j} \left( \hat{c}\left(\frac{t_j}{t_d}\right) \right). \end{aligned} \tag{3.19}$$

For  $1 \leq t_j/t_d \leq 2$  the function  $t_j/t_d$  of  $t_d$  belongs to  $\mathcal{S}(1, \mathcal{L}; 2, 2(2d - 1))$ . Hence, by Lemma 3.9 with  $F(v) = \mathcal{L}(v - 1)$  it follows that  $\hat{c}(t_j/t_d) \in \mathcal{S}(1, \mathcal{L}; 1, 10(2d - 1)M)$ . Using this and (3.17) in (3.19) we get

$$\begin{aligned} \frac{1}{k!} |D_d^k F_d(t)| &\leq 2 \sum_{|m|=k} \prod_{j=1}^{d-1} (10(2d - 1)M\mathcal{L}(m_j - 1))^{m_j} \\ &\quad + \sum_{|m|=k-1} \prod_{j=1}^{d-1} (10(2d - 1)M\mathcal{L}(m_j - 1))^{m_j} \\ &\leq 3 \left( \sum_{|m|=k} 1 \right) (10(2d - 1)M\mathcal{L}(k - 1))^k \\ &\leq 3(10d(2d - 1)M\mathcal{L}(k - 1))^k. \end{aligned} \tag{3.20}$$

We recall that the terms in (3.20) with  $m_j = 0$  are considered equal 1. Now, combining (3.18) and (3.20) we get for all  $(t_1, \dots, t_{d-1}) \in [0, \infty)^{d-1}$  and  $t_d$  as in (3.17)

$$\frac{1}{k!} |D_d^k \mathcal{N}(t)| \leq (d + 2)(10d(2d - 1)M\mathcal{L}(k - 1))^k, \tag{3.21}$$

i.e.  $\mathcal{N} \in \mathcal{S}(d, \mathcal{L}; d + 2, 10d(2d - 1)M)$ . Now, Lemma 3.9 with  $F(v) = \mathcal{L}(v - 1)$ ,  $g = \hat{a}$ ,  $f = \mathcal{N}$  and (3.21) prove the theorem.  $\square$

**Remark 3.11.** The arguments from the above proof also imply that (2.3) holds for the mixed derivatives of order  $k$ . However, Theorem 3.10 is sufficient for our purposes in this paper.

**Remark 3.12.** In Definition 3.2  $\mathcal{B}_1$  can be replaced by  $\mathcal{B}_\infty$ , but this will lead to some complications in the construction of admissible functions by semi-norms, as well as bigger constants in Theorem 3.10.

For  $\mathcal{L} = \mathcal{L}_{\ell, \varepsilon}$  the admissible multivariate cutoff function in Theorem 3.10 is from the class  $\mathcal{S}(d, \mathcal{L}_{\ell, \varepsilon}; \gamma_0, \tilde{\gamma}_0/\varepsilon^2)$ , where the second parameter is of order  $\varepsilon^{-2}$  and not of order  $\varepsilon^{-1}$  as in the univariate case. This is due to the method of construction via composition of two functions from  $\mathcal{S}(1, \mathcal{L}_{\ell, \varepsilon}; \gamma, \tilde{\gamma}/\varepsilon)$ . The composition necessarily belongs to  $\mathcal{S}(1, \mathcal{L}_{\ell, \varepsilon}; \gamma_0, \tilde{\gamma}_0/\varepsilon^2)$  unless better estimates for the derivatives are known. A different construction that leads to a smaller value of the second parameter is given in Sections 3.5–3.6.

### 3.5. Construction of admissible cutoff functions by univariate products

Another natural approach for constructing admissible  $d$ -dimensional cutoff functions resembles the construction of  $d$ -dimensional wavelets from univariate father wavelets.

From Lemmas 3.4–3.5 we immediately get

**Lemma 3.13.** *Let  $\hat{a}_{1,j}, \hat{a}_{2,j}, j = 1, \dots, d$ , be admissible univariate functions of type (a). Then*

$$\hat{B}(t) = \prod_{j=1}^d \hat{a}_{1,j}(t_j) - \prod_{j=1}^d \hat{a}_{2,j}(2t_j) \tag{3.22}$$

*is an admissible  $d$ -dimensional cutoff function of second kind and type (b).*

In the univariate case all admissible functions of type (c) are among the admissible functions of type (b) constructed via (3.22). Unfortunately, in dimensions  $d \geq 2$  representation (3.22) does not provide any admissible  $d$ -dimensional function of type (c). In order to get such cutoff functions we employ two other one-dimensional techniques.

**Lemma 3.14.** *Let  $\hat{A}$  be given by (3.2) with  $\hat{a}_j$  satisfying  $0 \leq \hat{a}_j(t) \leq 1$ . We define a cutoff function  $\hat{C}$  in two ways, namely,*

$$\hat{C}(t) := \begin{cases} 1 - \hat{A}^2(2t), & t \in \mathcal{B}_\infty, \\ \hat{A}(t)\sqrt{2 - \hat{A}^2(t)}, & t \in 2\mathcal{B}_\infty \setminus \mathcal{B}_\infty, \\ 0, & t \notin 2\mathcal{B}_\infty, \end{cases} \tag{3.23}$$

or

$$\hat{C}(t) := \begin{cases} 0, & t \in \frac{1}{2}\mathcal{B}_\infty, \\ \cos(\frac{\pi}{2}\hat{A}(2t)), & t \in \mathcal{B}_\infty \setminus \frac{1}{2}\mathcal{B}_\infty, \\ \sin(\frac{\pi}{2}\hat{A}(t)), & t \in 2\mathcal{B}_\infty \setminus \mathcal{B}_\infty, \\ 0, & t \notin 2\mathcal{B}_\infty. \end{cases} \tag{3.24}$$

*Then the function  $\hat{C} \geq 0$  from (3.23) or (3.24) is admissible of second kind and type (c).*

The proof of this lemma is straightforward.

**Remark 3.15.** The cutoff functions constructed in this subsection satisfy a stronger form of Definition 3.2 with  $\mathcal{B}_1$  replaced by  $\mathcal{B}_\infty$ .

### 3.6. Construction of admissible cutoff functions from univariate products with “small” derivatives

The admissible cutoff functions from univariate products from Section 3.5 allow better estimates on the derivatives than those in Section 3.4.

**Theorem 3.16.** *Let  $\mathcal{L}$  satisfy (3.16) and let  $M$  be given by (2.1). Let  $\gamma, \tilde{\gamma} > 0$  be such that the set  $\mathcal{S}(1, \mathcal{L}; \gamma, \tilde{\gamma}M)$  contains an admissible univariate cutoff function  $\hat{a}, 0 \leq \hat{a}(t) \leq 1$ , of type (a) according to Theorem 2.3.*

- (a) If  $\hat{A}$  is defined by (3.2) with  $\hat{a}_j = \hat{a}$ , then  $\hat{A}$  is an admissible cutoff function of second kind and type (a) and  $\hat{A} \in \mathcal{S}(d, \mathcal{L}; \gamma, \tilde{\gamma}M)$ .
- (b) If  $\hat{B}$  is defined by (3.22) with  $\hat{a}_{1,j} = \hat{a}_{1,j} = \hat{a}$ , then  $\hat{B}$  is an admissible cutoff function of second kind and type (b) and  $\hat{B} \in \mathcal{S}(d, \mathcal{L}; 2\gamma, 2\tilde{\gamma}M)$ .
- (c) If  $\hat{C}$  is defined by (3.24) with  $\hat{A}$  from (3.2), where  $\hat{a}_j = \hat{a}$ , then  $\hat{C}$  is admissible of second kind and type (c) and  $\hat{C} \in \mathcal{S}(d, \mathcal{L}; 1, (\pi\gamma + 2)\tilde{\gamma}M)$ .

**Proof.** It is established in Lemmas 3.4, 3.13 and 3.14 that  $\hat{A}, \hat{B}, \hat{C}$  are admissible cutoff functions of the respective type. The fact that  $\hat{A} \in \mathcal{S}(d, \mathcal{L}; \gamma, \tilde{\gamma}M)$  follows immediately by (3.2),  $0 \leq \hat{a}(t) \leq 1$  and Definition 2.2. Also  $\hat{B} \in \mathcal{S}(d, \mathcal{L}; 2\gamma, 2\tilde{\gamma}M)$  follows by (3.22) and Definition 2.2, as the constant  $2\tilde{\gamma}$  replaces  $\tilde{\gamma}$  because of the multiplier 2 in the arguments of the functions in the second product in (3.22).

To find bounds on the derivatives of  $\hat{C}(t)$  for  $1/2 \leq \|t\|_\infty \leq 2$  we fix  $1 \leq j \leq d$ . Consider  $\hat{C}(t) = (g \circ f)(t_j)$  as a function of  $t_j \in [1/2, 1]$ , where for  $\|t\|_\infty \leq 1$  we set  $f(t_j) = \hat{a}(2t_j)$  and  $g(x) = \cos(\lambda x)$  with  $\lambda = \frac{\pi}{2} \prod_{m=1, m \neq j}^d \hat{a}(2t_m)$  and for  $1 < \|t\|_\infty \leq 2$  we set  $f(t_j) = \hat{a}(t_j)$  and  $g(x) = \sin(\lambda x)$  with  $\lambda = \frac{\pi}{2} \prod_{m=1, m \neq j}^d \hat{a}(t_m)$ . We apply Lemma 3.9 as  $f \in \mathcal{S}(1, \mathcal{L}; \gamma, 2\tilde{\gamma}M)$ ,  $g \in \mathcal{S}(1, \mathcal{L}; 1, \pi/2)$  and get  $g \circ f \in \mathcal{S}(1, \mathcal{L}; 1, (\pi\gamma + 2)\tilde{\gamma}M)$ .

If  $t_j \in [1, 2]$  we use  $f(t_j) = \hat{a}(t_j)$  and  $g(x) = \sin(\lambda x)$  with  $\lambda = \frac{\pi}{2} \prod_{m=1, m \neq j}^d \hat{a}(t_m)$ . We apply Lemma 3.9 as  $f \in \mathcal{S}(1, \mathcal{L}; \gamma, \tilde{\gamma}M)$  and  $g \in \mathcal{S}(1, \mathcal{L}; 1, \pi/2)$  to obtain  $g \circ f \in \mathcal{S}(1, \mathcal{L}; 1, (\pi\gamma/2 + 1)\tilde{\gamma}M)$ . Consequently, in all cases  $\hat{C} \in \mathcal{S}(d, \mathcal{L}; 1, (\pi\gamma + 2)\tilde{\gamma}M)$ .  $\square$

**Remark 3.17.** In cases (a) and (b) of Theorem 3.16 it suffices to require  $\mathcal{L}$  to satisfy (2.1) instead of (3.16).

#### 4. Localized tensor product Jacobi polynomial kernels

Denote by  $\tilde{P}_n^{(\alpha_j, \beta_j)}$  ( $1 \leq j \leq d$ ) the  $n$ th degree Jacobi polynomial normalized in  $L^2([-1, 1], w_{\alpha_j, \beta_j})$ , see Section 2.2. Then for multi-indexes  $\alpha = (\alpha_1, \dots, \alpha_d)$  and  $\beta = (\beta_1, \dots, \beta_d)$  the  $d$ -dimensional tensor product Jacobi polynomials are defined by

$$\tilde{P}_v^{(\alpha, \beta)}(x) := \prod_{j=1}^d \tilde{P}_{v_j}^{(\alpha_j, \beta_j)}(x_j). \tag{4.1}$$

Recall our standing assumption:  $\alpha_j, \beta_j \geq -1/2$ . Evidently,  $\{\tilde{P}_v^{(\alpha, \beta)}\}_{v \in \mathbb{N}_0^d}$  is an orthonormal basis for the weighted space  $L^2([-1, 1]^d, w_{\alpha, \beta})$  with  $w_{\alpha, \beta}$  being the product Jacobi weight defined in (1.5).

We are interested in kernels of the form

$$A_n(x, y) := \sum_{v \in \mathbb{N}_0^d} \hat{A}\left(\frac{v}{n}\right) \tilde{P}_v^{(\alpha, \beta)}(x) \tilde{P}_v^{(\alpha, \beta)}(y), \quad x, y \in [-1, 1]^d. \tag{4.2}$$

Define

$$W_{\alpha,\beta}(n; x) := \prod_{j=1}^d w_{\alpha_j,\beta_j}(n; x_j), \tag{4.3}$$

where  $w_{\alpha_j,\beta_j}(n; x_j)$  is given in (2.10). We shall also use the distance on  $[-1, 1]^d$  defined by

$$\rho(x, y) = \max_{1 \leq j \leq d} |\arccos x_j - \arccos y_j|. \tag{4.4}$$

**Theorem 4.1.** *Suppose  $\hat{A} \in C^{3k-1}[0, \infty)^d$  for some  $k \geq 1$ ,  $\text{supp } \hat{A} \subset [0, 2]^d$ , and for any  $t \in [0, 2]^d$  of the form  $t = (t_1, \dots, t_{\ell-1}, 0, t_{\ell+1}, \dots, t_d)$ ,  $1 \leq \ell \leq d$ , i.e.  $t = \text{proj}_\ell t$ ,  $\hat{A}$  satisfies  $D_\ell^m \hat{A}(t) = 0$  for  $m = 1, 2, \dots, 3k - 1$ . Then the kernels from (4.2) satisfy*

$$|\Lambda_n(x, y)| \leq \frac{cn^d}{\sqrt{W_{\alpha,\beta}(n; x)}\sqrt{W_{\alpha,\beta}(n; y)}} (1 + n\rho(x, y))^{-k}, \quad x, y \in [-1, 1]^d. \tag{4.5}$$

Here the constant  $c$  depends on  $k, d, \alpha, \beta$  and  $\|D_\ell^{3k-1} \hat{A}\|_\infty$ ,  $\ell = 1, \dots, d$ , but not on  $x, y$  and  $n$ . Consequently, for an admissible cutoff function  $\hat{A}$  the above estimate holds for any  $k > 0$ .

**Proof.** Without loss of generality we may assume that  $\rho(x, y) = |\arccos x_d - \arccos y_d|$ . We write  $\Lambda_n$  from (4.2) as

$$\begin{aligned} \Lambda_n(x, y) &= \sum_{\nu_1=0}^{2n-1} \cdots \sum_{\nu_{d-1}=0}^{2n-1} \left[ \sum_{\nu_d=0}^\infty \hat{A}\left(\frac{\nu_1}{n}, \dots, \frac{\nu_d}{n}\right) \tilde{P}_{\nu_d}^{(\alpha_d, \beta_d)}(x_d) \tilde{P}_{\nu_d}^{(\alpha_d, \beta_d)}(y_d) \right] \\ &\quad \times \prod_{j=1}^{d-1} \tilde{P}_{\nu_j}^{(\alpha_j, \beta_j)}(x_j) \prod_{j=1}^{d-1} \tilde{P}_{\nu_j}^{(\alpha_j, \beta_j)}(y_j). \end{aligned} \tag{4.6}$$

For any  $\nu_1, \dots, \nu_{d-1}$  we estimate the inner sum in (4.6) by using Theorem 2.5. We get

$$\begin{aligned} &\left| \sum_{\nu_d=0}^\infty \hat{A}\left(\frac{\nu_1}{n}, \dots, \frac{\nu_d}{n}\right) \tilde{P}_{\nu_d}^{(\alpha_d, \beta_d)}(x_d) \tilde{P}_{\nu_d}^{(\alpha_d, \beta_d)}(y_d) \right| \\ &\leq \frac{cn}{\sqrt{w_{\alpha_d, \beta_d}(n; x_d)}\sqrt{w_{\alpha_d, \beta_d}(n; y_d)}} (1 + n\rho(x, y))^{-k}. \end{aligned} \tag{4.7}$$

For the Jacobi polynomials from the outer products we apply (2.22),  $\alpha_j, \beta_j \geq -1/2$ , and use that  $\nu_j < 2n$  to obtain

$$|\tilde{P}_{\nu_j}^{(\alpha_j, \beta_j)}(t)| \leq \frac{c}{\sqrt{w_{\alpha_j, \beta_j}(\nu_j; t)}} \leq \frac{c}{\sqrt{w_{\alpha_j, \beta_j}(n; t)}}, \quad t = x_j, y_j. \tag{4.8}$$

Combining the above two estimates and the fact that the total number of terms in the outer sums in (4.6) is  $(2n)^{d-1}$  proves the theorem.  $\square$

We next show that estimate (4.5) can be improved for cutoff functions of “small” derivatives given by Theorem 3.10 or Theorem 3.16.

**Theorem 4.2.** *Let  $\hat{A}$  be an admissible cutoff function of second kind which belongs to  $\mathcal{S}(d, \mathcal{L}; \gamma, \tilde{\gamma}M)$  with  $\mathcal{L}$  and  $M$  as in (2.1) and  $\gamma, \tilde{\gamma} > 0$  (see Definition 2.2). Then the kernels from (4.2) satisfy*

$$|\Lambda_n(x, y)| \leq \frac{cn^d}{\sqrt{W_{\alpha,\beta}(n; x)}\sqrt{W_{\alpha,\beta}(n; y)}} \exp\left\{-\frac{\tilde{c}n\rho(x, y)}{\mathcal{L}(n\rho(x, y))}\right\} \tag{4.9}$$

for  $x, y \in [-1, 1]^d$ . Here  $\tilde{c} = c'/\tilde{\gamma}M$  with  $c' > 0$  being an absolute constant and the constant  $c > 0$  depends on  $d, M, \alpha, \beta, \gamma$  and  $\tilde{\gamma}$ , but not on  $x, y$  and  $n$ .

The proof of Theorem 4.2 is the same as the proof of Theorem 4.1 with the role of Theorem 2.5 played by Theorem 2.7 and Remark 2.8.

The next theorem shows that the kernels  $\Lambda_n(x, y)$  from (4.2) are Lip 1 in  $x$  and  $y$  with respect to the distance  $\rho(\cdot, \cdot)$ ; it is needed for our further development.

**Theorem 4.3.** *Under the hypotheses of Theorem 4.1 with  $k > 2 \max_i \{\alpha_i + \beta_i\} + 5$  for all  $x, y, \xi \in [-1, 1]^d$  such that  $\rho(x, \xi) \leq c_*n^{-1}, n \geq 1, c_* > 0$ , the kernel  $\Lambda_n$  from (4.2) satisfies*

$$|\Lambda_n(x, y) - \Lambda_n(\xi, y)| \leq \frac{cn^{d+1}\rho(x, \xi)}{\sqrt{W_{\alpha,\beta}(n; x)}\sqrt{W_{\alpha,\beta}(n; y)}} (1 + n\rho(x, y))^{-\sigma}, \tag{4.10}$$

where  $\sigma = k - 2 \max_i \{\alpha_i + \beta_i\} - 5$  and  $c > 0$  depends only on  $k, d, \alpha, \beta, c_*$ , and  $\|D_\ell^{3k-1} \hat{A}\|_\infty, \ell = 1, \dots, d$ . Therefore, for an admissible cutoff function  $\hat{A}$  the above estimate holds for any  $\sigma > 0$ .

**Proof.** Apparently it suffices to prove estimate (4.10) for all  $\xi \in [-1, 1]^d$  of the form  $\xi = x + \delta e_i$  such that  $\rho(x, x + \delta e_i) \leq c_*n^{-1}$  and  $1 \leq i \leq d$  with  $e_i$  being the  $i$ th coordinate vector.

As in the proof of Theorem 4.1, we may assume that  $\rho(x, y) = |\arccos x_d - \arccos y_d| =: \rho_d(x, y)$ . Assuming that  $\xi = x + \delta e_i$  is as above, we consider two cases for  $i$ .

Case 1:  $i = d$ . Then we have

$$\begin{aligned} &\Lambda_n(x, y) - \Lambda_n(x + \delta e_d, y) \\ &= \sum_{v_1=0}^{2n-1} \dots \sum_{v_{d-1}=0}^{2n-1} \left[ \sum_{v_d=0}^{\infty} \hat{A}\left(\frac{v_1}{n}, \dots, \frac{v_d}{n}\right) (\tilde{P}_{v_d}^{(\alpha_d, \beta_d)}(x_d) - \tilde{P}_{v_d}^{(\alpha_d, \beta_d)}(x_d + \delta)) \tilde{P}_{v_d}^{(\alpha_d, \beta_d)}(y_d) \right] \\ &\quad \times \prod_{j=1}^{d-1} \tilde{P}_{v_j}^{(\alpha_j, \beta_j)}(x_j) \prod_{j=1}^{d-1} \tilde{P}_{v_j}^{(\alpha_j, \beta_j)}(y_j). \end{aligned}$$

Applying Theorem 2.10 to the inner sum we get

$$\left| \sum_{v_d=0}^{\infty} \hat{A}\left(\frac{v_1}{n}, \dots, \frac{v_d}{n}\right) (\tilde{P}_{v_d}^{(\alpha_d, \beta_d)}(x_d) - \tilde{P}_{v_d}^{(\alpha_d, \beta_d)}(x_d + \delta)) \tilde{P}_{v_d}^{(\alpha_d, \beta_d)}(y_d) \right| \leq \frac{cn^2 \rho(x, x + \delta e_i)}{\sqrt{w_{\alpha_d, \beta_d}(n; x_d)} \sqrt{w_{\alpha_d, \beta_d}(n; y_d)}} (1 + n\rho_d(x, y))^{-\sigma}.$$

For the Jacobi polynomials  $\tilde{P}_{v_j}^{(\alpha_j, \beta_j)}(x_j)$  and  $\tilde{P}_{v_j}^{(\alpha_j, \beta_j)}(y_j)$  from the outer products we apply estimates (4.8) and combining these with the above we arrive at (4.10).

Case 2:  $i \neq d$ . Let  $x, x + \delta e_i \in [-1, 1]^d$  and  $\rho(x, x + \delta e_i) \leq c_* n^{-1}$ . We have

$$\begin{aligned} & A_n(x, y) - A_n(x + \delta e_i, y) \\ &= \sum_{v_1=0}^{2n-1} \cdots \sum_{v_{d-1}=0}^{2n-1} \left[ \sum_{v_d=0}^{\infty} \hat{A}\left(\frac{v_1}{n}, \dots, \frac{v_d}{n}\right) \tilde{P}_{v_d}^{(\alpha_d, \beta_d)}(x_d) \tilde{P}_{v_d}^{(\alpha_d, \beta_d)}(y_d) \right] \\ & \times \prod_{j=1, j \neq i}^{d-1} \tilde{P}_{v_j}^{(\alpha_j, \beta_j)}(x_j) (\tilde{P}_{v_i}^{(\alpha_i, \beta_i)}(x_i) - \tilde{P}_{v_i}^{(\alpha_i, \beta_i)}(x_i + \delta)) \prod_{j=1}^{d-1} \tilde{P}_{v_j}^{(\alpha_j, \beta_j)}(y_j). \end{aligned} \tag{4.11}$$

As is well known that  $\frac{d}{dt}[P_m^{(\alpha, \beta)}(t)] = \frac{m + \alpha + \beta + 1}{2} P_{m-1}^{(\alpha+1, \beta+1)}(t)$  (see [23, (4.21.7)]). Combining this with estimate (2.22) from Lemma 2.9 and  $h_m^{(\alpha, \beta)} \sim h_{m-1}^{(\alpha, \beta)} \sim m^{-1}$  (see (2.7)) give

$$\left| \frac{d}{dt} \tilde{P}_m^{(\alpha, \beta)}(t) \right| \leq \frac{cm}{\sqrt{w_{\alpha+1, \beta+1}(m-1, t)}} \leq \frac{cm}{\sqrt{w_{\alpha, \beta}(m, t)} (\sqrt{1-t^2} + m^{-1})}.$$

We use this to obtain for  $\theta, \theta' \in [0, \pi]$  with  $|\theta - \theta'| \leq c_* m^{-1}, m \geq 2$ ,

$$\begin{aligned} \left| \tilde{P}_m^{(\alpha, \beta)}(\cos \theta) - \tilde{P}_m^{(\alpha, \beta)}(\cos \theta') \right| &\leq \frac{cm |\cos \theta - \cos \theta'|}{\sqrt{w_{\alpha, \beta}(m, \cos \theta)} (\sin \theta + m^{-1})} \\ &\leq \frac{cm \sin \left| \frac{\theta - \theta'}{2} \right| \left| \sin \left| \frac{\theta + \theta'}{2} \right| \right|}{\sqrt{w_{\alpha, \beta}(m, \cos \theta)} (\sin \theta + m^{-1})} \\ &\leq \frac{cm |\theta - \theta'|}{\sqrt{w_{\alpha, \beta}(m, \cos \theta)}}. \end{aligned} \tag{4.12}$$

Note that (4.12) is trivial for  $m = 0, 1$ . Therefore,

$$\left| \tilde{P}_{v_i}^{(\alpha_i, \beta_i)}(x_i) - \tilde{P}_{v_i}^{(\alpha_i, \beta_i)}(x_i + \delta) \right| \leq \frac{cv_i \rho(x, x + \delta e_i)}{\sqrt{w_{\alpha_i, \beta_i}(v_i, x_i)}} \leq \frac{cn\rho(x, \xi)}{\sqrt{w_{\alpha_i, \beta_i}(n, x_i)}}. \tag{4.13}$$

Now, we use (4.7) to estimate the inner sum in (4.11), (4.8) to estimate the Jacobi polynomials  $\tilde{P}_{v_j}^{(\alpha_j, \beta_j)}(x_j)$  ( $j \neq i$ ) and  $\tilde{P}_{v_j}^{(\alpha_j, \beta_j)}(y_j)$  from the outer products in (4.11), and we also use (4.13) to obtain again (4.10). Here as in Case 1 we took into account that the number of terms in the outer sums is  $(2n)^{d-1}$ .  $\square$

Lower bound estimates for the  $L^p$ -norms of the kernels  $\Lambda_n(x, y)$  in  $x$  or  $y$  can also be easily derived from the corresponding results in dimension one.

**Proposition 4.4.** *Let  $\hat{A}$  be admissible and  $|\hat{A}(t)| \geq c > 0$  for  $t \in [1, 1 + \delta]^d$ ,  $\delta > 0$ . Then for  $n \geq 1/\delta$*

$$\int_{[-1,1]^d} |\Lambda_n(x, y)|^2 w_{\alpha,\beta}(y) dy \geq cn^d W_{\alpha,\beta}(n; x)^{-d}, \quad x \in [-1, 1]^d, \tag{4.14}$$

where  $c > 0$  depends only on  $\delta, \alpha, \beta$ , and  $d$ .

**Proof.** By the definition of  $\Lambda_n(x, y)$  in (4.2) and the orthogonality of the Jacobi polynomials, it follows that

$$\begin{aligned} \int_{[-1,1]^d} |\Lambda_n(x, y)|^2 w_{\alpha,\beta}(y) dy &= \sum_{v \in \mathbb{N}_0^d} |\hat{A}(v/n)|^2 [\tilde{P}_v^{(\alpha,\beta)}(x)]^2 \\ &\geq \sum_{v \in [n, n+\delta n]^d} |\hat{A}(v/n)|^2 [\tilde{P}_v^{(\alpha,\beta)}(x)]^2 \\ &\geq c \prod_{i=1}^d \sum_{v_i=n}^{n+[\delta n]} [\tilde{P}_{v_i}^{(\alpha_i,\beta_i)}(x_i)]^2 \end{aligned}$$

and the stated lower bound follows from the respective result in the univariate case, given in [13, Proposition 2.4].  $\square$

The rapidly decaying polynomial kernels  $\Lambda_n(x, y)$  from (4.2) can be utilized as in the univariate case [13, Proposition 2.6] for establishing Nikolski type inequalities:

**Proposition 4.5.** *For  $0 < q \leq p \leq \infty$  and  $g \in \Pi_n^d$ ,*

$$\|g\|_p \leq cn^{(2d+2 \sum_{i=1}^d \min\{0, \max\{\alpha_i, \beta_i\}\})(1/q-1/p)} \|g\|_q, \tag{4.15}$$

furthermore, for any  $s \in \mathbb{R}$ ,

$$\|W_{\alpha,\beta}(n; \cdot)^s g(\cdot)\|_p \leq cn^{d(1/q-1/p)} \|W_{\alpha,\beta}(n; \cdot)^{s+1/p-1/q} g(\cdot)\|_q. \tag{4.16}$$

**5. Additional auxiliary results**

*5.1. The maximal inequality*

We let  $\mathcal{M}_t$  ( $0 < t < \infty$ ) be the maximal operator defined by

$$\mathcal{M}_t f(x) := \sup_{I \ni x} \left( \frac{1}{\mu(I)} \int_I |f(y)|^t w_{\alpha,\beta}(y) dy \right)^{1/t}, \quad x \in [-1, 1]^d, \tag{5.1}$$

where the sup is over all boxes (rectangles)  $I \subset [-1, 1]^d$  with sides parallel to the coordinate axes containing  $x$ . Here  $\mu(E) := \int_E w_{\alpha, \beta}(y) dy$ .

We denote by  $B(\xi, r)$  the “ball” (box) centered at  $\xi \in [-1, 1]^d$  of radius  $r > 0$  with respect to the distance  $\rho(\cdot, \cdot)$  on  $[-1, 1]^d$ , i.e.

$$B(\xi, r) := \{x \in [-1, 1]^d : \rho(x, \xi) < r\}. \tag{5.2}$$

We next show that for  $0 < \delta \leq \pi$

$$\mu(B(y, \delta)) \sim \delta^d \prod_{i=1}^d \left(\sqrt{1 - y_i^2} + \delta\right)^{2\gamma_i+1}, \quad \gamma_i := \begin{cases} \alpha_i & \text{if } 0 \leq y_i \leq 1, \\ \beta_i & \text{if } -1 \leq y_i < 0. \end{cases} \tag{5.3}$$

Let  $y_i =: \cos \phi_i$ ,  $0 \leq \phi_i \leq \pi$ , and  $\phi'_i := \max\{\phi_i - \delta, 0\}$ ,  $\phi''_i := \min\{\phi_i + \delta, \pi\}$ . Evidently

$$\begin{aligned} \mu(B(y, \delta)) &= \prod_{i=1}^d \int_{\cos \phi'_i}^{\cos \phi_i} (1 - x_i)^{\alpha_i} (1 + x_i)^{\beta_i} dx_i \\ &= \prod_{i=1}^d \int_{\phi'_i}^{\phi_i} (1 - \cos \theta_i)^{\alpha_i} (1 + \cos \theta_i)^{\beta_i} \sin \theta_i d\theta_i \\ &\sim \delta^d \prod_{i=1}^d (\sin \phi_i + \delta)^{2\gamma_i+1} = \delta^d \prod_{i=1}^d \left(\sqrt{1 - y_i^2} + \delta\right)^{2\gamma_i+1}, \end{aligned}$$

which confirms (5.3).

By (5.3) it follows that  $\mu(B(y, 2\delta)) \leq c\mu(B(y, \delta))$ , i.e.  $\mu$  is a doubling measure on  $[-1, 1]^d$  and, therefore, the Fefferman–Stein vector-valued maximal inequality is valid (see [22]): Assuming that  $0 < p < \infty$ ,  $0 < q \leq \infty$  and  $0 < t < \min\{p, q\}$ , then for any sequence of functions  $\{f_k\}_{k=1}^\infty$  on  $[-1, 1]^d$ ,

$$\left\| \left( \sum_{k=1}^\infty |\mathcal{M}_t f_k(\cdot)|^q \right)^{1/q} \right\|_p \leq c \left\| \left( \sum_{k=1}^\infty |f_k(\cdot)|^q \right)^{1/q} \right\|_p. \tag{5.4}$$

We need to estimate  $(\mathcal{M}_t \mathbb{1}_{B(y, \delta)})(x)$ . Such estimates readily follow by (5.3) and the respective univariate result in [13, Lemma 2.7].

**Lemma 5.1.** *Let  $y \in [-1, 1]^d$  and  $0 < r \leq \pi$ , and suppose  $\gamma_i$ ,  $i = 1, \dots, d$ , are defined as in (5.3). Then for any  $x \in [-1, 1]^d$*

$$(\mathcal{M}_t \mathbb{1}_{B(y, r)})(x) \sim \prod_{j=1}^d \left(1 + \frac{\rho(y_j, x_j)}{r}\right)^{-1/t} \left(1 + \frac{\rho(y_j, x_j)}{r + \rho(y_j, 1)}\right)^{-(2\gamma_j+1)/t} \tag{5.5}$$

and hence

$$\begin{aligned}
 (\mathcal{M}_t \mathbb{1}_{B(y,r)})(x) &\geq c \prod_{j=1}^d \left( 1 + \frac{\rho(y_j, x_j)}{r} \right)^{-(2\gamma_j+2)/t} \\
 &\geq c \left( 1 + \frac{\rho(y, x)}{r} \right)^{-(2|\gamma|+2d)/t}.
 \end{aligned}
 \tag{5.6}$$

Here  $\rho(y_j, x_j) := |\arccos y_j - \arccos x_j|$  and  $\rho(y, x)$  is defined in (4.4).

We also want to record the following useful inequality which follows easily from the case  $d = 1$ , proved in [13, (2.22)]:

$$W_{\alpha,\beta}(n; x) \leq c W_{\alpha,\beta}(n; y) (1 + n\rho(x, y))^{d+2\sum_{i=1}^d \max\{\alpha_i, \beta_i\}},
 \tag{5.7}$$

for  $x, y \in [-1, 1]^d$  and  $n \geq 1$ , where  $W_{\alpha,\beta}(n; x)$  is from (4.3).

### 5.2. Distributions on $[-1, 1]^d$

Here we introduce and give some basic facts about distributions on  $[-1, 1]^d$ . We shall use as test functions the set  $\mathcal{D} := C^\infty[-1, 1]^d$ , where the topology is induced by the semi-norms

$$\|\phi\|_\mu := \|D^\mu \phi(t)\|_\infty \quad \text{for all multi-indices } \mu.
 \tag{5.8}$$

Observe that the tensor product Jacobi polynomials  $\{\tilde{P}_\nu^{(\alpha,\beta)}\}$  belong to  $\mathcal{D}$  and more importantly the test functions  $\phi \in \mathcal{D}$  can be completely characterized by the coefficients of their Jacobi expansions. Denote

$$\mathcal{N}_k(\phi) := \sup_{\nu \in \mathbb{N}_0^d} (|\nu| + 1)^k |\langle \phi, \tilde{P}_\nu^{(\alpha,\beta)} \rangle|,
 \tag{5.9}$$

where  $\langle f, g \rangle := \int_{[-1,1]^d} f(x) \overline{g(x)} w_{\alpha,\beta}(x) dx$ .

#### Lemma 5.2.

- (i)  $\phi \in \mathcal{D}$  if and only if  $|\langle \phi, \tilde{P}_\nu^{(\alpha,\beta)} \rangle| = \mathcal{O}((|\nu| + 1)^{-k})$  for all  $k$ .
- (ii) For every  $\phi \in \mathcal{D}$  we have  $\phi = \sum_{\nu \in \mathbb{N}_0^d} \langle \phi, \tilde{P}_\nu^{(\alpha,\beta)} \rangle \tilde{P}_\nu^{(\alpha,\beta)}$ , where the convergence is in the topology of  $\mathcal{D}$ .
- (iii) The topology in  $\mathcal{D}$  can be equivalently defined by the norms  $\mathcal{N}_k(\cdot)$ ,  $k \geq 0$ .

The proof of this lemma is easy and similar to the proof of Lemma 2.8 in [13].

The space  $\mathcal{D}'$  of distributions on  $[-1, 1]^d$  is defined as the set of all continuous linear functionals on  $\mathcal{D}$ . The pairing of  $f \in \mathcal{D}'$  and  $\phi \in \mathcal{D}$  will usually be denoted by  $\langle f, \phi \rangle := f(\bar{\phi})$ . As will

be shown it is in a sense consistent with the inner product  $\langle f, g \rangle$  in  $L^2(w_{\alpha,\beta})$ . We shall need the representation of distributions from  $\mathcal{D}'$  in terms of Jacobi polynomials.

**Lemma 5.3.**

(i) A linear functional  $f$  on  $\mathcal{D}$  belongs to  $\mathcal{D}'$  if and only if there exists  $k \geq 0$  such that

$$|f(\phi)| = |\langle f, \bar{\phi} \rangle| \leq c_k \mathcal{N}_k(\phi) \quad \text{for all } \phi \in \mathcal{D}. \tag{5.10}$$

(ii) For any  $f \in \mathcal{D}'$  there exist constants  $c > 0$  and  $k \geq 0$  such that

$$|f(\tilde{P}_v^{(\alpha,\beta)})| = |\langle f, \tilde{P}_v^{(\alpha,\beta)} \rangle| \leq c_k (|v| + 1)^k \quad \text{for all } v \in \mathbb{N}_0^d, \quad \text{and} \tag{5.11}$$

$$f(\phi) = \lim_{n \rightarrow \infty} \langle S_n, \bar{\phi} \rangle = \sum_{v \in \mathbb{N}_0^d} \langle f, \tilde{P}_v^{(\alpha,\beta)} \rangle \langle \phi, \tilde{P}_v^{(\alpha,\beta)} \rangle \quad \text{for } \phi \in \mathcal{D}, \tag{5.12}$$

where  $S_n := \sum_{|v| \leq n} \langle f, \tilde{P}_v^{(\alpha,\beta)} \rangle \tilde{P}_v^{(\alpha,\beta)}$  and the series converges absolutely.

(iii) For any sequence  $\{c_v\}_{v \in \mathbb{N}_0^d}$  satisfying  $|c_v| \leq A(|v| + 1)^\ell$  for  $v \in \mathbb{N}_0^d$  and some constants  $A$  and  $\ell$ , the sequence

$$s_n := \sum_{|v| \leq n} c_v P_v^{(\alpha,\beta)}$$

converges in  $\mathcal{D}'$  as  $n \rightarrow \infty$  to some distribution  $F \in \mathcal{D}'$  such that  $\langle F, \tilde{P}_v^{(\alpha,\beta)} \rangle = c_v$  for  $v \in \mathbb{N}_0^d$ .

**Proof.** Part (i) of the lemma follows by the definition of  $\mathcal{D}'$  and Lemma 5.2 as in the classical case.

Estimate (5.11) is immediate from (5.10) and (5.9). Further, we have for  $\phi \in \mathcal{D}$

$$\lim_{n \rightarrow \infty} \langle S_n, \bar{\phi} \rangle = \lim_{n \rightarrow \infty} f \left( \sum_{|v| \leq n} \langle \phi, \tilde{P}_v^{(\alpha,\beta)} \rangle P_v^{(\alpha,\beta)} \right) = f(\phi),$$

which confirms (5.12). Here we used Lemma 5.2, (ii).

To prove part (iii), we observe that  $\langle s_n, \bar{\phi} \rangle = \sum_{|v| \leq n} c_v \langle \phi, \tilde{P}_v^{(\alpha,\beta)} \rangle$  for  $\phi \in \mathcal{D}$  and using the assumption and Lemma 5.2 we get  $|c_v| |\langle \phi, \tilde{P}_v^{(\alpha,\beta)} \rangle| \leq c(|v| + 1)^{\ell-k}$  for an arbitrary  $k \geq 0$ . Therefore, the series  $\sum_{v \in \mathbb{N}_0^d} c_v \langle \phi, \tilde{P}_v^{(\alpha,\beta)} \rangle$  converges absolutely and hence

$$F(\phi) := \lim_{n \rightarrow \infty} \langle s_n, \bar{\phi} \rangle = \sum_{v \in \mathbb{N}_0^d} c_v \langle \phi, \tilde{P}_v^{(\alpha,\beta)} \rangle, \quad \phi \in \mathcal{D}, \tag{5.13}$$

is a well-defined linear functional. We claim that  $F$  is bounded. Indeed, for  $\phi \in \mathcal{D}$

$$\begin{aligned}
 |F(\phi)| &\leq \sum_{v \in \mathbb{N}_0^d} |c_v| |\langle \phi, \tilde{P}_v^{(\alpha, \beta)} \rangle| \leq A \sum_{v \in \mathbb{N}_0^d} (|v| + 1)^\ell |\langle \phi, \tilde{P}_v^{(\alpha, \beta)} \rangle| \\
 &\leq A \mathcal{N}_{\ell+d+1}(\phi) \sum_{v \in \mathbb{N}_0^d} (|v| + 1)^{-d-1} \leq c \mathcal{N}_{\ell+d+1}(\phi),
 \end{aligned}$$

which shows that  $F \in \mathcal{D}'$ .

Finally,  $F(\tilde{P}_v^{(\alpha, \beta)}) = \lim_{n \rightarrow \infty} \langle s_n, \bar{\phi} \rangle = c_v$  is immediate by (5.13).  $\square$

To simplify our notation, we introduce the following “convolution”: For functions  $\Phi : [-1, 1]^d \times [-1, 1]^d \rightarrow \mathbb{C}$  and  $f : [-1, 1]^d \rightarrow \mathbb{C}$ , we define

$$\Phi * f(x) := \int_{[-1, 1]^d} \Phi(x, y) f(y) w_{\alpha, \beta}(y) dy \tag{5.14}$$

and extend it to  $\mathcal{D}'$  by duality, i.e. assuming that  $f \in \mathcal{D}'$  and  $\Phi : [-1, 1]^d \times [-1, 1]^d \rightarrow \mathbb{C}$  is such that  $\Phi(x, y)$  belongs to  $\mathcal{D}$  as a function of  $y$ , we define  $\Phi * f$  by

$$\Phi * f(x) := \langle f, \overline{\Phi(x, \cdot)} \rangle. \tag{5.15}$$

Here on the right  $f$  acts on  $\overline{\Phi(x, y)}$  as a function of  $y$ .

### 5.3. $L^p$ -multipliers

We shall need  $L^p$ -multipliers for tensor product Jacobi polynomial expansions. Since we cannot find any such multipliers in the literature we next derive simple but non-optimal multipliers satisfying the First Boundary Condition (Section 1.2) of a certain order.

**Theorem 5.4.** *Let  $m \in C^r[0, \infty)^d$  for  $r$  sufficiently large ( $r > 6 \max_i \{\alpha_i + \beta_i\} + 6 \sum_i \max\{\alpha_i, \beta_i\} + 6d + 20$  will do) and suppose  $m$  satisfies the following condition: For any  $t \in [0, \infty)^d$  of the form  $t = (t_1, \dots, t_{\ell-1}, 0, t_{\ell+1}, \dots, t_d)$ ,  $1 \leq \ell \leq d$ , we have  $D_\ell^s m(t) = 0$  for  $s = 1, 2, \dots, r$ . Also, assume*

$$|D^\tau m(t)| \leq c(1 + \|t\|_\infty)^{-|\tau|} \quad \text{for } t \in [0, \infty)^d \text{ and } |\tau| \leq r, \tag{5.16}$$

with  $c > 0$  independent of  $t$ . Then the operator  $T_m f := \sum_{v \in \mathbb{N}_0^d} m(v) \langle f, \tilde{P}_v^{(\alpha, \beta)} \rangle \tilde{P}_v^{(\alpha, \beta)}$  is bounded on  $L^p(w_{\alpha, \beta})$  for  $1 < p < \infty$ .

**Proof.** We shall utilize a standard decomposition of unity argument. Let  $\hat{C}$  be an admissible cutoff function of type (c). Then  $\hat{B} = |\hat{C}|^2 \geq 0$  is admissible of type (b) and  $\sum_{j=0}^\infty \hat{B}(2^{-j}t) = 1$  for  $t \in [0, \infty)^d \setminus [0, 1)^d$ . We define  $\Phi_0(x, y) := m(0) \tilde{P}_0^{(\alpha, \beta)}(x) \tilde{P}_0^{(\alpha, \beta)}(y)$  and

$$\Phi_j(x, y) := \sum_{v \in \mathbb{N}_0^d} \hat{B}\left(\frac{v}{2^{j-1}}\right) m(v) \tilde{P}_v^{(\alpha, \beta)}(x) \tilde{P}_v^{(\alpha, \beta)}(y), \quad j \geq 1.$$

Set  $K(x, y) := \sum_{j \geq 0} \Phi_j(x, y)$ .

We shall show that  $K(x, y)$  is well defined for  $x \neq y$  and  $|K(x, y)| \leq cV(x, y)^{-1}$ , where  $V(x, y) := \mu(B(y, \rho(x, y)))$ , and furthermore  $K(x, y)$  obeys the following Hörmander condition

$$\int_{M \setminus B(y, 2\delta)} |K(x, y) - K(x, \bar{y})| d\mu(x) \leq c, \quad \text{whenever } \bar{y} \in B(y, \delta), \tag{5.17}$$

for all  $y \in M$  and  $\delta > 0$ . To this end it suffices to show that

$$|K(x, y) - K(x, \bar{y})| \leq c \frac{\rho(y, \bar{y})}{\rho(x, \bar{y})} V(x, \bar{y})^{-1}, \tag{5.18}$$

whenever  $\rho(x, \bar{y}) \geq 2\rho(y, \bar{y})$ , see [2,22].

For the proof of (5.18), fix  $x, y, \bar{y} \in [-1, 1]^d$ ,  $x \neq \bar{y}$ , and define

$$\gamma_i := \begin{cases} \alpha_i & \text{if } 0 \leq \bar{y}_i \leq 1, \\ \beta_i & \text{if } -1 \leq \bar{y}_i < 0. \end{cases}$$

By (5.3) it follows that

$$V(x, \bar{y}) \sim \rho(x, \bar{y})^d \prod_{i=1}^d \left( \sqrt{1 - \bar{y}_i^2} + \rho(x, \bar{y}) \right)^{2\gamma_i+1}. \tag{5.19}$$

Let  $\hat{A}_j(t) := \hat{B}(t)m(2^{j-1}t)$ . We have  $\text{supp } \hat{B} \subset [0, 2]^d \setminus \frac{1}{2}\mathcal{B}_1$  and by (5.16) one gets  $|D^\tau[m(2^{j-1}t)]| \leq c$  for  $t \in [0, 2]^d \setminus \frac{1}{2}\mathcal{B}_1$  and  $|\tau| \leq r$ , where the constant  $c > 0$  is independent of  $j$ . Therefore,  $\|D^\tau \hat{A}_j\|_\infty = \|D^\tau[\hat{B}(\cdot)m(2^{j-1}\cdot)]\|_\infty \leq c$  for  $|\tau| \leq r$  with  $c > 0$  independent of  $j$ . Now, it is evident that  $\hat{A}_j$  satisfies the assumptions of Theorem 4.3 for some  $k > 2 \max_i \{\alpha_i + \beta_i\} + 2 \sum_i \max\{\alpha_i, \beta_i\} + 2d + 6$  and hence, using also (5.7), we get

$$|\Phi_j(x, y) - \Phi_j(x, \bar{y})| \leq \frac{c2^{j(d+1)}\rho(y, \bar{y})}{W_{\alpha,\beta}(2^j, \bar{y})(1 + 2^j\rho(x, \bar{y}))^\sigma} \tag{5.20}$$

if  $\rho(y, \bar{y}) \leq 2^{-j}$ , where  $\sigma = k - 2 \max_i \{\alpha_i + \beta_i\} - 5$ .

If  $\rho(y, \bar{y}) > 2^{-j}$  and  $\rho(x, \bar{y}) \geq 2\rho(y, \bar{y})$  (hence  $\rho(x, y) \geq \rho(y, \bar{y})$  and  $\rho(x, \bar{y}) \leq 2\rho(x, y) \leq 3\rho(x, \bar{y})$ ), then estimate (5.20) follows by Theorem 4.1 applied separately to  $\Phi_j(x, y)$  and  $\Phi_j(x, \bar{y})$  and using (5.7). Therefore, (5.20) holds whenever  $\rho(x, \bar{y}) \geq 2\rho(y, \bar{y})$ .

Let  $2^{-j_1-1} \leq \rho(x, \bar{y}) < 2^{-j_1}$ . Then using  $\Phi_0(x, y) = \Phi_0(x, \bar{y})$  we write

$$\begin{aligned} |K(x, y) - K(x, \bar{y})| &\leq \sum_{j=1}^{j_1} |\Phi_j(x, y) - \Phi_j(x, \bar{y})| + \sum_{j=j_1+1}^{\infty} |\Phi_j(x, y) - \Phi_j(x, \bar{y})| \\ &=: F_1 + F_2. \end{aligned}$$

For  $F_1$  we have using (5.20) and (5.19)

$$\begin{aligned}
 F_1 &\leq \frac{c\rho(y, \bar{y})}{\prod_{i=1}^d (\sqrt{1 - \bar{y}_i^2} + 2^{-j_i})^{2\gamma_i+1}} \sum_{j=1}^{j_1} 2^{j(d+1)} \leq \frac{c\rho(y, \bar{y})2^{j_1(d+1)}}{\prod_{i=1}^d (\sqrt{1 - \bar{y}_i^2} + 2^{-j_i})^{2\gamma_i+1}} \\
 &\leq \frac{c\rho(y, \bar{y})}{\rho(x, \bar{y})^{d+1} \prod_{i=1}^d (\sqrt{1 - \bar{y}_i^2} + \rho(x, \bar{y}))^{2\gamma_i+1}} \leq c \frac{\rho(y, \bar{y})}{\rho(x, \bar{y})} V(x, \bar{y})^{-1}.
 \end{aligned}$$

To estimate  $F_2$  we first observe that  $(\sqrt{1 - \bar{y}_i^2} + 2^{-j})(1 + 2^{j-j_i}) \geq \sqrt{1 - \bar{y}_i^2} + 2^{-j_i}$ . Then, using again (5.20) and (5.19), we get

$$\begin{aligned}
 F_2 &\leq c\rho(y, \bar{y}) \sum_{j=j_1+1}^{\infty} \frac{2^{j(d+1)}}{\prod_{i=1}^d (\sqrt{1 - \bar{y}_i^2} + 2^{-j})^{2\gamma_i+1} (1 + 2^{j-j_i})^\sigma} \\
 &\leq c\rho(y, \bar{y}) \sum_{j=j_1+1}^{\infty} \frac{2^{j(d+1)}}{\prod_{i=1}^d (\sqrt{1 - \bar{y}_i^2} + 2^{-j_i})^{2\gamma_i+1} (1 + 2^{j-j_i})^{\sigma-2\sum_i \gamma_i-d}} \\
 &\leq \frac{c\rho(y, \bar{y})2^{j_1(d+1)}}{\prod_{i=1}^d (\sqrt{1 - \bar{y}_i^2} + 2^{-j_i})^{2\gamma_i+1}} \sum_{j=j_1+1}^{\infty} 2^{-(j-j_i)(\sigma-2\sum_i \gamma_i-2d-1)} \\
 &\leq \frac{c\rho(y, \bar{y})}{\rho(x, \bar{y})^{d+1} \prod_{i=1}^d (\sqrt{1 - \bar{y}_i^2} + \rho(x, \bar{y}))^{2\gamma_i+1}} \leq c \frac{\rho(y, \bar{y})}{\rho(x, \bar{y})} V(x, \bar{y})^{-1},
 \end{aligned}$$

where we used that  $\sigma > 2\sum_i \gamma_i + 2d + 1$ . The above estimates of  $F_1$  and  $F_2$  yield (5.18) and hence the kernel  $K(x, y)$  satisfies the Hörmander condition (5.17).

The estimate

$$|K(x, y)| \leq \sum_{j \geq 0} |\Phi_j(x, y)| \leq cV(x, y)^{-1}, \quad x \neq y, \tag{5.21}$$

follows similarly as above from

$$|\Phi_j(x, y)| \leq \frac{c2^{jd}}{W_{\alpha,\beta}(2^j, y)(1 + 2^j \rho(x, y))^\sigma},$$

which is a consequence of Theorem 4.1, using (5.7). We omit the details.

By Parseval’s identity  $\|T_m f\|_2 \leq c\|f\|_2$  for  $f \in L^2(\omega_{\alpha,\beta})$ . Using this and (5.21), a standard argument yields that  $K(x, y)$  is the kernel of the operator  $T_m$ . Also, as indicated in Section 5.1, the measure has the doubling property (see (5.3)). Therefore,  $T_m$  is a generalized Calderón–Zygmund operator and hence  $T_m$  is bounded on  $L^p(\omega_{\alpha,\beta})$ ,  $1 < p < \infty$  (see [2,22]). The proof is complete.  $\square$

### 6. Construction of building blocks (needlets)

The construction of frames (needlets) on  $[-1, 1]^d$  has two basic components: (i) a Calderón type decomposition formula and (ii) a cubature formula.

6.1. Cubature formula and subdivision of  $[-1, 1]^d$

For the construction of needlets we shall employ the Gaussian quadrature formula on  $[-1, 1]$  with weight  $w_{\alpha,\beta}(t) := (1 - t)^\alpha(1 + t)^\beta$ . Given  $j \geq 0$ , denote by  $\xi^m =: \cos \theta_m$ ,  $m = 1, 2, \dots, 2^{j+1}$ , the zeros of the Jacobi polynomial  $P_{2^{j+1}}^{(\alpha,\beta)}$  ordered so that  $0 < \theta_1 < \dots < \theta_{2^{j+1}} < \pi$  and set

$$\mathcal{X}_j^{\alpha,\beta} := \{\xi^m : 1 \leq m \leq 2^{j+1}\}.$$

It is well known that uniformly (see [9])

$$\theta_1 \sim 2^{-j}, \quad \pi - \theta_{2^{j+1}} \sim 2^{-j}, \quad \theta_{m+1} - \theta_m \sim 2^{-j}, \quad \text{and hence} \quad \theta_m \sim m2^{-j}. \tag{6.1}$$

As is well known [23] the zeros of the Jacobi polynomial  $P_{2^{j+1}}^{(\alpha,\beta)}$  serve as knots of the Gaussian quadrature

$$\int_{[-1,1]} f(t)w_{\alpha,\beta}(t) dt \sim \sum_{\xi \in \mathcal{X}_j^{\alpha,\beta}} c_\xi f(\xi), \tag{6.2}$$

which is exact for all algebraic polynomials that are of degree  $2^{j+2} - 1$ . Furthermore, the coefficients  $c_\xi$  are all positive and satisfy (see e.g. [16])

$$c_\xi \sim 2^{-j} w_{\alpha,\beta}(\xi)(1 - \xi^2)^{1/2}. \tag{6.3}$$

**Tiling of  $[-1, 1]$ .** With  $\{\xi^m\}$  as above we write

$$I_{\xi^m} := [(\xi^{m+1} + \xi^m)/2, (\xi^{m-1} + \xi^m)/2], \quad m = 2, 3, \dots, 2^{j+1} - 1, \quad \text{and} \\ I_{\xi^1} := [(\xi^2 + \xi^1)/2, 1], \quad I_{\xi^{2^{j+1}}} := [-1, (\xi^{2^{j+1}} + \xi^{2^{j+1}-1})/2].$$

We define

$$\mathcal{I}_j^{\alpha,\beta} := \{I_{\xi^m} : 1 \leq m \leq 2^{j+1}\}.$$

For multi-indices  $\alpha = (\alpha_1, \dots, \alpha_d)$ ,  $\beta = (\beta_1, \dots, \beta_d)$  and  $j \geq 0$ ,  $1 \leq i \leq d$ , we denote by  $\mathcal{X}_j^{\alpha_i,\beta_i}$  the zeroes of the Jacobi polynomial  $P_{2^{j+1}}^{(\alpha_i,\beta_i)}$  and write

$$\mathcal{X}_j := \mathcal{X}_j^{\alpha_1,\beta_1} \times \dots \times \mathcal{X}_j^{\alpha_d,\beta_d}. \tag{6.4}$$

Now, for  $\xi = (\xi_1, \dots, \xi_d) \in \mathcal{X}_j$  we set  $c_\xi := c_{\xi_1} \dots c_{\xi_d}$ , where  $c_{\xi_i}$  is the corresponding coefficient of the Gaussian quadrature (6.2) with  $\alpha = \alpha_i$  and  $\beta = \beta_i$ . Evidently, the cubature formula

$$\int_{[-1,1]^d} f(x)w_{\alpha,\beta}(x) dx \sim \sum_{\xi \in \mathcal{X}_j} c_\xi f(\xi) \tag{6.5}$$

is exact for all polynomials in  $d$ -variables of degree  $2^{j+2} - 1$  in each variable and by (6.3) the coefficients  $\{c_\xi\}$  are positive and satisfy

$$c_\xi \sim 2^{-dj} W_{\alpha,\beta}(2^j; \xi), \tag{6.6}$$

where  $W_{\alpha,\beta}(2^j; \xi)$  is defined in (4.3).

**Tiling of  $[-1, 1]^d$ .** For  $\xi = (\xi_1, \dots, \xi_d) \in \mathcal{X}_j$ , we write

$$I_\xi := I_{\xi_1} \times \dots \times I_{\xi_d}, \quad I_{\xi_i} \in \mathcal{I}_j^{\alpha_i, \beta_i}. \tag{6.7}$$

Evidently,  $[-1, 1]^d = \bigcup_{\xi \in \mathcal{X}_j} I_\xi$  and the interiors of the tiles  $\{I_\xi\}_{\xi \in \mathcal{X}_j}$  do not overlap.

With  $B(y, r)$  defined in (5.2) it easily follows from the univariate case that there exist constants  $c_1, c_2 > 0$  such that

$$B(\xi, c_1 2^{-j}) \subset I_\xi \subset B(\xi, c_2 2^{-j}), \quad \xi \in \mathcal{X}_j. \tag{6.8}$$

By (5.3) it follows that

$$\mu(I_\xi) := \int_{I_\xi} w_{\alpha,\beta}(x) dx \sim 2^{-j} W_{\alpha,\beta}(2^j; \xi) \sim c_\xi, \quad \xi \in \mathcal{X}_j, \quad j \geq 0. \tag{6.9}$$

The next lemma is of an independent interest and is instrumental in the subsequent development.

**Lemma 6.1.** *Suppose  $P \in \Pi_{2^j}^d$ ,  $\xi \in \mathcal{X}_j$ ,  $j \geq 0$ , and let  $x', x'' \in [-1, 1]^d$  be such that  $\rho(x', \xi) \leq c_\star 2^{-j}$ ,  $\rho(x'', \xi) \leq c_\star 2^{-j}$ . Then for any  $\sigma > 0$*

$$|P(x') - P(x'')| \leq c 2^j \rho(x', x'') \sum_{\eta \in \mathcal{X}_j} \frac{|P(\eta)|}{(1 + 2^j \rho(\xi, \eta))^\sigma},$$

where  $c > 0$  depends only on  $\sigma, \alpha, \beta, d$ , and  $c_\star$ .

The proof of this lemma is merely a repetition of the proof of the univariate result in [13, Lemma 9.2] and will be omitted.

6.2. *Needlets on  $[-1, 1]^d$*

The construction of needlet systems is now standard and follows a well-established scheme. We begin with two cutoff functions  $\hat{A}, \hat{B}$  of type (b) which satisfy (see Lemma 3.6):

$$\sum_{j=0}^{\infty} \overline{\hat{A}(2^{-j}t)} \hat{B}(2^{-j}t) = 1, \quad t \in [0, \infty)^d \setminus \mathcal{B}_\infty. \tag{6.10}$$

We define  $\Phi_0(x, y) = \Psi_0(x, y) := \tilde{P}_0(x) \tilde{P}_0(y)$ ,

$$\Phi_j(x, y) := \sum_{v \in \mathbb{N}_0^d} \hat{A}\left(\frac{v}{2^{j-1}}\right) \tilde{P}_v(x) \tilde{P}_v(y), \quad j \geq 1, \quad \text{and} \tag{6.11}$$

$$\Psi_j(x, y) := \sum_{v \in \mathbb{N}_0^d} \hat{B}\left(\frac{v}{2^{j-1}}\right) \tilde{P}_v(x) \tilde{P}_v(y), \quad j \geq 1. \tag{6.12}$$

Let  $\mathcal{X}_j$  be the set of knots of cubature formula (6.5), defined in (6.4), and let  $\{c_\xi\}$  be its coefficients. We define the  $j$ th level *needlets* by

$$\varphi_\xi(x) := c_\xi^{1/2} \Phi_j(x, \xi) \quad \text{and} \quad \psi_\xi(x) := c_\xi^{1/2} \Psi_j(x, \xi), \quad \xi \in \mathcal{X}_j. \tag{6.13}$$

We write  $\mathcal{X} := \bigcup_{j=0}^\infty \mathcal{X}_j$ , where equal points from different levels  $\mathcal{X}_j$  are considered as distinct elements of  $\mathcal{X}$ , so that  $\mathcal{X}$  can be used as an index set. We define the *analysis* and *synthesis* needlet systems  $\Phi$  and  $\Psi$  by

$$\Phi := \{\varphi_\xi\}_{\xi \in \mathcal{X}}, \quad \Psi := \{\psi_\xi\}_{\xi \in \mathcal{X}}. \tag{6.14}$$

Theorem 4.1 and (5.7) imply that the needlets decay rapidly, namely,

$$|\varphi_\xi(x)|, |\psi_\xi(x)| \leq \frac{c_\sigma 2^{jd/2}}{\sqrt{W_{\alpha,\beta}(2^j; \xi)}} (1 + 2^j \rho(\xi, x))^{-\sigma}, \quad x \in [-1, 1]^d, \quad \forall \sigma. \tag{6.15}$$

We next give estimates on the norms of the needlets, which can be proved exactly as in the case  $d = 1$ , upon using (6.15) and the lower bound estimate from Proposition 4.4: For  $0 < p \leq \infty$ ,

$$\|\varphi_\xi\|_p \sim \|\psi_\xi\|_p \sim \|\tilde{\mathbb{1}}_{I_\xi}\|_p \sim \left(\frac{2^{dj}}{W_{\alpha,\beta}(2^j; \xi)}\right)^{1/2-1/p}, \quad \xi \in \mathcal{X}_j. \tag{6.16}$$

Here  $\tilde{\mathbb{1}}_{I_\xi} := \mu(I_\xi)^{-1/2} \mathbb{1}_{I_\xi}$  with  $\mathbb{1}_E$  being the characteristic function of the set  $E$ . Moreover, there exist constants  $c^*, c^\diamond > 0$  such that

$$\|\varphi_\xi\|_{L^\infty(B(\xi, c^* 2^{-j}))}, \|\psi_\xi\|_{L^\infty(B(\xi, c^* 2^{-j}))} \geq c^\diamond \left(\frac{2^{dj}}{W_{\alpha,\beta}(2^j; \xi)}\right)^{1/2}. \tag{6.17}$$

The needlet decomposition of  $\mathcal{D}'$  and  $L^p$  follows as in the univariate case (see [13, Proposition 3.1]) by the definition of needlets and their superb localization.

**Proposition 6.2.**

(i) For  $f \in \mathcal{D}'$ , we have

$$f = \sum_{j=0}^\infty \Psi_j * \bar{\Phi}_j * f \quad \text{in } \mathcal{D}', \quad \text{and} \tag{6.18}$$

$$f = \sum_{\xi \in \mathcal{X}} \langle f, \varphi_\xi \rangle \psi_\xi \quad \text{in } \mathcal{D}'. \tag{6.19}$$

(ii) If  $f \in L^p(w_{\alpha,\beta})$ ,  $1 \leq p \leq \infty$ , then (6.18)–(6.19) hold in  $L^p(w_{\alpha,\beta})$ . Moreover, if  $1 < p < \infty$ , then the convergence in (6.18)–(6.19) is unconditional.

**Remark 6.3.**

- (i) Select  $\hat{A} \geq 0$  a cutoff function of type (c) (see Definition 3.1). Then choosing  $\hat{B} = \hat{A}$  in the construction of needlets in (6.10)–(6.13) we obtain  $\varphi_\xi = \psi_\xi$ . Consequently, (6.19) becomes  $f = \sum_{\xi \in \mathcal{X}} \langle f, \psi_\xi \rangle \psi_\xi$  and it is easy to prove that (see e.g. [13])  $\|f\|_2 = (\sum_{\xi \in \mathcal{X}} |\langle f, \psi_\xi \rangle|^2)^{1/2}$  for  $f \in L^2(w_{\alpha,\beta})$ , which shows that  $\Psi$  is a tight frame for  $L^2(w_{\alpha,\beta})$ .
- (ii) If  $\hat{A} \geq 0$  is an admissible cutoff function of second kind and type (c) (see Definition 3.2) which belongs to  $\mathcal{S}(d, \mathcal{L}; \gamma, \tilde{\gamma}M)$ , then Theorem 4.2 implies sup-exponential localization of the needlets, namely,

$$|\psi_\xi(x)| \leq \frac{c2^{jd/2}}{\sqrt{W_{\alpha,\beta}(2^j; \xi)}} \exp\left\{-\frac{\tilde{c}2^j \rho(\xi, x)}{\mathcal{L}(2^j \rho(\xi, x))}\right\}, \quad x \in [-1, 1]^d. \tag{6.20}$$

**7. Weighted Triebel–Lizorkin spaces on  $[-1, 1]^d$**

We next utilize the general idea of using spectral or orthogonal decompositions (see e.g. [17,24]) to introduce weighted Triebel–Lizorkin spaces on  $[-1, 1]^d$ . The theory of these spaces is entirely parallel to their theory in the univariate case, developed in [13]. Therefore, we shall only state the main results, provide the important ingredients and refer the reader to [13] for the proofs.

Given an admissible cutoff function  $\hat{A}$  of type (b) (see Definition 3.1) satisfying the dyadic covering condition (3.4) we define a sequence of kernels  $\{\Phi_j\}$  by  $\Phi_0(x, y) := \tilde{P}_0(x)\tilde{P}_0(y)$  and

$$\Phi_j(x, y) := \sum_{v \in \mathbb{N}_0^d} \hat{A}\left(\frac{v}{2^{j-1}}\right) \tilde{P}_v(x)\tilde{P}_v(y), \quad j \geq 1. \tag{7.1}$$

**Definition 7.1.** For  $s, \rho \in \mathbb{R}$ ,  $0 < p < \infty$ , and  $0 < q \leq \infty$  the weighted Triebel–Lizorkin space  $F_{pq}^{s\rho} := F_{pq}^{s\rho}(w_{\alpha,\beta})$  is defined as the set of all  $f \in \mathcal{D}'$  such that

$$\|f\|_{F_{pq}^{s\rho}} := \left\| \left( \sum_{j=0}^{\infty} [2^{sj} W_{\alpha,\beta}(2^j; \cdot)^{-\rho/d} |\Phi_j * f(\cdot)|^q] \right)^{1/q} \right\|_p < \infty \tag{7.2}$$

with the usual modification when  $q = \infty$ .

Note that the above definition is independent of the choice of  $\hat{A}$  as long as  $\hat{A}$  is an admissible function of type (b), satisfying (3.4) (see Theorem 7.3 below).

Also,  $F_{pq}^{s\rho}$  is a (quasi-)Banach space which is continuously embedded in  $\mathcal{D}'$ , i.e. there exist  $k$  and  $c > 0$  such that

$$|\langle f, \phi \rangle| \leq c \|f\|_{F_{pq}^{s\rho}} \mathcal{N}_k(\phi) \quad \text{for all } f \in F_{pq}^{s\rho}, \phi \in \mathcal{D}.$$

We next introduce the sequence spaces  $f_{pq}^{s\rho}$  associated to  $F_{pq}^{s\rho}$ . Here we assume that  $\{\mathcal{X}_j\}_{j=0}^\infty$  and  $\mathcal{X} := \bigcup_{j=0}^\infty \mathcal{X}_j$  are the sets of points from the definition of needles with associated neighborhoods  $\{I_\xi\}$ , given in (6.7).

**Definition 7.2.** Suppose  $s, \rho \in \mathbb{R}$ ,  $0 < p < \infty$ , and  $0 < q \leq \infty$ . Then  $f_{pq}^{s\rho}$  is defined as the space of all complex-valued sequences  $h := \{h_\xi\}_{\xi \in \mathcal{X}}$  such that

$$\|h\|_{f_{pq}^{s\rho}} := \left\| \left( \sum_{j=0}^\infty 2^{sjq} \sum_{\xi \in \mathcal{X}_j} [|h_\xi| W_{\alpha,\beta}(2^j; \xi)^{-\rho/d} \tilde{\mathbb{1}}_{I_\xi}(\cdot)]^q \right)^{1/q} \right\|_p < \infty \tag{7.3}$$

with the usual modification for  $q = \infty$ . Here as before  $\tilde{\mathbb{1}}_{I_\xi} := \mu(I_\xi)^{-1/2} \mathbb{1}_{I_\xi}$ .

The “analysis” and “synthesis” operators associated to the needlet systems  $\Phi, \Psi$  are defined by

$$S_\varphi : f \rightarrow \{\langle f, \varphi_\xi \rangle\}_{\xi \in \mathcal{X}} \quad \text{and} \quad T_\psi : \{h_\xi\}_{\xi \in \mathcal{X}} \rightarrow \sum_{\xi \in \mathcal{X}} h_\xi \psi_\xi. \tag{7.4}$$

As in [13] one shows that the operator  $T_\psi$  is well defined on  $f_{pq}^{s\rho}$ , namely, for any  $h \in f_{pq}^{s\rho}$ ,  $T_\psi h := \sum_{\xi \in \mathcal{X}} h_\xi \psi_\xi$  converges in  $\mathcal{D}'$ . Moreover, the operator  $T_\psi : f_{pq}^{s\rho} \rightarrow \mathcal{D}'$  is continuous, i.e. there exist constants  $k > 0$  and  $c > 0$  such that

$$|\langle T_\psi h, \phi \rangle| \leq c \mathcal{N}_k(\phi) \|h\|_{f_{pq}^{s\rho}} \quad \text{for all } h \in f_{pq}^{s\rho}, \phi \in \mathcal{D}. \tag{7.5}$$

Our main result in this section asserts that the weighted  $F$ -spaces can be characterized by the needlet coefficients of the distributions.

**Theorem 7.3.** Let  $s, \rho \in \mathbb{R}$ ,  $0 < p < \infty$  and  $0 < q \leq \infty$ . The operators  $S_\varphi : F_{pq}^{s\rho} \rightarrow f_{pq}^{s\rho}$  and  $T_\psi : f_{pq}^{s\rho} \rightarrow F_{pq}^{s\rho}$  are bounded and  $T_\psi \circ S_\varphi = Id$  on  $F_{pq}^{s\rho}$ . Consequently,  $f \in F_{pq}^{s\rho}$  if and only if  $\{\langle f, \varphi_\xi \rangle\}_{\xi \in \mathcal{X}} \in f_{pq}^{s\rho}$ . Furthermore,

$$\|f\|_{F_{pq}^{s\rho}} \sim \left\| \{\langle f, \varphi_\xi \rangle\}_{\xi \in \mathcal{X}} \right\|_{f_{pq}^{s\rho}} \sim \left\| \left( \sum_{j=0}^\infty 2^{sjq} \sum_{\xi \in \mathcal{X}_j} [|\langle f, \varphi_\xi \rangle| W_{\alpha,\beta}(2^j; \xi)^{-\rho/d} |\psi_\xi(\cdot)|]^q \right)^{1/q} \right\|_p. \tag{7.6}$$

In addition, the definition of  $F_{pq}^{s\rho}$  is independent of the particular selection of the type (b) cutoff function  $\hat{A}$  satisfying (3.4).

To us the spaces  $F_{pq}^{ss}$  are more natural than the spaces  $F_{pq}^{s\rho}$  with  $\rho \neq s$  since they embed “correctly” with respect to the smoothness index  $s$ .

**Proposition 7.4.** *Let  $0 < p < p_1 < \infty$ ,  $0 < q, q_1 \leq \infty$ , and  $-\infty < s_1 < s < \infty$ . Then we have the continuous embedding*

$$F_{pq}^{ss} \subset F_{p_1q_1}^{s_1s_1} \quad \text{if } s/d - 1/p = s_1/d - 1/p_1. \tag{7.7}$$

The proof of this proposition is quite similar to the proof of the respective embedding result on  $B^d$  in [14, Proposition 4.11] and will be omitted.

We have the following identification of spaces  $F_{p2}^{00}$ .

**Proposition 7.5.** *We have*

$$F_{p2}^{00} \sim L^p(w_{\alpha,\beta}), \quad 1 < p < \infty,$$

with equivalent norms. Consequently, for any  $f \in L^p(w_{\alpha,\beta})$ ,  $1 < p < \infty$ ,

$$\|f\|_p \sim \left\| \left( \sum_{j=0}^{\infty} \sum_{\xi \in \mathcal{X}_j} (|\langle f, \varphi_{\xi} \rangle| |\psi_{\xi}(\cdot)|)^2 \right)^{1/2} \right\|_p.$$

The proof of this proposition uses the multipliers from Theorem 5.4 and can be carried out exactly as in the case of spherical harmonic expansions in [15, Proposition 4.3]. We omit it.

### 8. Weighted Besov spaces on $[-1, 1]^d$

To define weighted Besov spaces on  $[-1, 1]^d$  we use again the sequence of kernels  $\{\Phi_j\}$  introduced in (7.1) with  $\hat{A}$  a cutoff function of type (b) obeying (3.4). We shall keep the development of these spaces short since the proofs of the results are the same as in the univariate case, given in [13].

**Definition 8.1.** Let  $s, \rho \in \mathbb{R}$  and  $0 < p, q \leq \infty$ . The weighted Besov space  $B_{pq}^{s\rho} := B_{pq}^{s\rho}(w_{\alpha,\beta})$  is defined as the set of all  $f \in \mathcal{D}'$  such that

$$\|f\|_{B_{pq}^{s\rho}} := \left( \sum_{j=0}^{\infty} (2^{sj} \|W_{\alpha,\beta}(2^j; \cdot)^{-\rho/d} \Phi_j * f(\cdot)\|_p)^q \right)^{1/q} < \infty, \tag{8.1}$$

where the  $\ell^q$ -norm is replaced by the sup-norm if  $q = \infty$ .

Note that as in the case of weighted Triebel–Lizorkin spaces the above definition is independent of the particular choice of  $\hat{A}$  and  $B_{pq}^{s\rho}$  is a (quasi-)Banach space which is continuously embedded in  $\mathcal{D}'$ .

We next introduce the sequence spaces  $b_{pq}^{s\rho}$  associated to  $B_{pq}^{s\rho}$ . To this end we use some of the notation established in the previous section.

**Definition 8.2.** Let  $s, \rho \in \mathbb{R}$  and  $0 < p, q \leq \infty$ . Then  $b_{pq}^{s\rho}$  is defined to be the space of all complex-valued sequences  $h := \{h_{\xi}\}_{\xi \in \mathcal{X}}$  such that

$$\|h\|_{b_{pq}^{s\rho}} := \left( \sum_{j=0}^{\infty} 2^{j(s-d/p+d/2)q} \left[ \sum_{\xi \in \mathcal{X}_j} (W_{\alpha,\beta}(2^j; \xi))^{-\rho/d+1/p-1/2} |h_{\xi}| \right]^p \right)^{1/q} \tag{8.2}$$

is finite, with the usual modification for  $p = \infty$  or  $q = \infty$ .

The analysis and synthesis operators  $S_{\varphi}$  and  $T_{\psi}$  defined in (7.4) play an important role here. As for weighted Triebel–Lizorkin spaces the operator  $T_{\psi}$  is well defined on  $b_{pq}^{s\rho}$ , i.e. for any  $h \in b_{pq}^{s\rho}$ ,  $T_{\psi}h := \sum_{\xi \in \mathcal{X}} h_{\xi} \psi_{\xi}$  converges in  $\mathcal{D}'$ . Also, the operator  $T_{\psi} : b_{pq}^{s\rho} \rightarrow \mathcal{D}'$  is continuous.

The following characterization of weighted Besov spaces is the main result of this section.

**Theorem 8.3.** *Let  $s, \rho \in \mathbb{R}$  and  $0 < p, q \leq \infty$ . Then the operators  $S_{\varphi} : B_{pq}^{s\rho} \rightarrow b_{pq}^{s\rho}$  and  $T_{\psi} : b_{pq}^{s\rho} \rightarrow B_{pq}^{s\rho}$  are bounded and  $T_{\psi} \circ S_{\varphi} = Id$  on  $B_{pq}^{s\rho}$ . Consequently, for  $f \in \mathcal{D}'$  we have that  $f \in B_{pq}^{s\rho}$  if and only if  $\{\langle f, \varphi_{\xi} \rangle\}_{\xi \in \mathcal{X}} \in b_{pq}^{s\rho}$ . Moreover,*

$$\|f\|_{B_{pq}^{s\rho}} \sim \|\{\langle f, \varphi_{\xi} \rangle\}\|_{b_{pq}^{s\rho}} \sim \left( \sum_{j=0}^{\infty} 2^{sjq} \left[ \sum_{\xi \in \mathcal{X}_j} (W_{\alpha,\beta}(2^j; \xi))^{-\rho/d} \|\langle f, \varphi_{\xi} \rangle \psi_{\xi}\|_p \right]^p \right)^{1/q}.$$

In addition, the definition of  $B_{pq}^{s\rho}$  is independent of the particular selection of the type (b) cutoff function  $\hat{A}$  satisfying (3.4).

The parameter  $\rho$  in the definition of  $B_{pq}^{s\rho}$  allows to consider various scales of weighted Besov spaces. The spaces  $B_{pq}^{s0}$  can be regarded as “classical” Besov spaces. However, to us more natural are the spaces  $B_{pq}^{ss}$  ( $\rho = s$ ) which in contrast to  $B_{pq}^{s0}$ , first, embed “correctly” with respect to the smoothness index  $s$ , and secondly, the right smoothness spaces in nonlinear  $n$ -term weighted approximation from needles are defined in terms of spaces  $B_{pq}^{ss}$  (see Section 9 below).

**Proposition 8.4.** *Let  $0 < p \leq p_1 \leq \infty$ ,  $0 < q \leq q_1 \leq \infty$ , and  $-\infty < s_1 \leq s < \infty$ . Then we have the continuous embedding*

$$B_{pq}^{ss} \subset B_{p_1q_1}^{s_1s_1} \quad \text{if } s/d - 1/p = s_1/d - 1/p_1. \tag{8.3}$$

This proposition is an immediate consequence of estimate (4.16).

### 9. Application of weighted Besov spaces to nonlinear approximation

We now consider nonlinear  $n$ -term approximation for a needlet system  $\{\psi_{\eta}\}_{\eta \in \mathcal{X}}$  with  $\varphi_{\eta} = \psi_{\eta}$ , defined as in (6.11)–(6.14) with  $\hat{B} = \hat{A}$ ,  $\hat{A} \geq 0$ , i.e.  $\hat{A} \geq 0$  is a first or second kind admissible cutoff function of type (c) (see Definitions 3.1–3.2). Then  $\{\psi_{\eta}\}$  are real-valued.

Let  $\Sigma_n$  be the nonlinear set of all functions  $g$  of the form  $g = \sum_{\xi \in \Lambda} a_{\xi} \psi_{\xi}$ , where  $\Lambda \subset \mathcal{X}$ ,  $\#\Lambda \leq n$ , and  $\Lambda$  is allowed to vary with  $g$ . Denote by  $\sigma_n(f)_p$  the error of best  $L^p(w_{\alpha,\beta})$ -approximation to  $f \in L^p(w_{\alpha,\beta})$  from  $\Sigma_n$ , i.e.

$$\sigma_n(f)_p := \inf_{g \in \Sigma_n} \|f - g\|_p.$$

We consider approximation in  $L^p(w_{\alpha,\beta})$ ,  $0 < p < \infty$ .

Assume  $0 < p < \infty$ ,  $s > 0$ , and  $1/\tau := s/d + 1/p$  and denote briefly  $B_\tau^s := B_{\tau\tau}^{ss}$ . By Theorem 8.3 and (6.16) it follows that

$$\|f\|_{B_\tau^s} \sim \left( \sum_{\xi \in \mathcal{X}} \|\langle f, \psi_\xi \rangle \psi_\xi\|_p^\tau \right)^{1/\tau}. \tag{9.1}$$

Exactly as in [15, Proposition 6.1] this leads to the embedding of  $B_\tau^s$  into  $L^p(w_{\alpha,\beta})$ , which plays an important role in the proof of the main result of this section:

**Theorem 9.1** (*Jackson estimate*). *If  $f \in B_\tau^s$ , then*

$$\sigma_n(f)_p \leq cn^{-s/d} \|f\|_{B_\tau^s}, \quad n \geq 1, \tag{9.2}$$

where  $c > 0$  depends only on  $s$ ,  $p$ , and  $\hat{A}$ .

The proofs of this theorem can be carried out exactly as the proofs of the Jackson estimate in [15, Theorem 6.2]. We omit it.

It is an important open problem to prove the companion to (9.2) Bernstein estimate: If  $g \in \Sigma_n$  and  $1 < p < \infty$ , then

$$\|g\|_{B_\tau^s} \leq cn^{s/d} \|g\|_p. \tag{9.3}$$

If true this estimate would enable one to characterize the rates (approximation spaces) of nonlinear  $n$ -term approximation in  $L^p(w_{\alpha,\beta})$  ( $1 < p < \infty$ ) from needlet systems.

### 10. Weighted Triebel–Lizorkin and Besov spaces on $B^{d_1} \times [-1, 1]^{d_2}$

Our aim is to briefly describe how the theory of weighted spaces of distributions on the product set  $B^{d_1} \times [-1, 1]^{d_2}$  can be developed via tensor product orthogonal polynomials.

#### 10.1. Localized kernels for orthogonal polynomials on the ball

Localized polynomial kernels on the unit ball  $B^d$  in  $\mathbb{R}^d$  have been developed in [19] and utilized in [14] to the development of Triebel–Lizorkin and Besov spaces on  $B^d$  with weight

$$w_\mu(x) := (1 - \|x\|_2^2)^\mu, \quad \mu \geq 0.$$

Here, we compile all needed results from [14,19] and give some new facts.

Denote by  $V_n$  the set of all polynomials of degree  $n$  in  $d$  variables which are orthogonal to the lower degree polynomials in  $L^2(B^d, w_\mu)$  and let  $P_n(w_\mu, x, y)$  be the kernel of the orthogonal projector  $\text{Proj}_n : L^2(B^d, w_\mu) \rightarrow V_n$ , i.e.

$$(\text{Proj}_n f)(x) = \int_{B^d} f(y) P_n(w_\mu; x, y) w_\mu(y) dy. \tag{10.1}$$

An explicit representation of the reproducing kernel  $P_n(w_\mu, x, y)$  is given in [25]:

$$P_n(w_\mu; x, y) = c_\mu \frac{n + \lambda}{\lambda} \int_{-1}^1 C_n^\lambda(t(x, y; u))(1 - u^2)^{\mu-1} du, \quad \mu > 0, \tag{10.2}$$

where

$$t(x, y; u) := \langle x, y \rangle + u\sqrt{1 - \|x\|_2^2}\sqrt{1 - \|y\|_2^2} \quad \text{and} \quad \lambda := \mu + \frac{d-1}{2}.$$

For an admissible univariate cutoff function  $\hat{a}$  (see Definition 2.1), denote

$$L_n^\mu(x, y) := \sum_{j=0}^\infty \hat{a}\left(\frac{j}{n}\right) P_j(w_\mu; x, y). \tag{10.3}$$

Analogue of Theorems 4.1 and 4.2 on  $B^d$  are established in [19]. Denote

$$W_\mu(n; x) := \left(\sqrt{1 - \|x\|_2^2} + n^{-1}\right)^{2\mu}, \quad \text{and} \tag{10.4}$$

$$\rho(x, y) := \arccos\left(\langle x, y \rangle + \sqrt{1 - \|x\|_2^2}\sqrt{1 - \|y\|_2^2}\right), \tag{10.5}$$

which is a distance on  $B^d$ .

**Theorem 10.1.** *Given an admissible univariate cutoff function  $\hat{a}$ , for any  $\sigma > 0$  there exists a constant  $c > 0$  such that*

$$|L_n^\mu(x, y)| \leq \frac{cn^d}{\sqrt{W_\mu(n; x)}\sqrt{W_\mu(n; y)}}(1 + n\rho(x, y))^{-\sigma}, \quad x, y \in B^d. \tag{10.6}$$

Furthermore, for any  $x, y, \xi \in B^d$  such that  $\rho(x, \xi) \leq c_*n^{-1}$

$$|L_n^\mu(x, y) - L_n^\mu(\xi, y)| \leq \frac{cn^{d+1}\rho(x, \xi)}{\sqrt{W_\mu(n; x)}\sqrt{W_\mu(n; y)}}(1 + n\rho(x, y))^{-\sigma}. \tag{10.7}$$

This theorem was established in [19, Theorem 4.2 and Proposition 4.7] in the case of admissible cutoff functions  $\hat{a}$  which are constant around  $t = 0$ . Its proof hinges on the localization of the kernels  $Q_n^{\alpha, \beta}$  from (2.12). Due to Theorem 2.6 now Theorem 10.1 holds for admissible cutoff functions  $\hat{a}$  in the sense of Definition 2.1 with the proof from [19].

We shall need two additional estimates with the first being the analogue of Lemma 2.9 on  $B^d$ .

**Lemma 10.2.** *For  $x, y \in B^d$ ,*

$$|P_n(w_\mu; x, y)| \leq \frac{cn^{d-1}}{\sqrt{W_\mu(n; x)}\sqrt{W_\mu(n; y)}}. \tag{10.8}$$

**Proof.** The proof of this lemma relies on the following estimate that follows from Theorem 3.1 in [3]: If  $a$  and  $b$  are constants such that  $|a| + |b| \leq 1$ , then

$$\left| \int_{-1}^1 C_n^\lambda(au + b)(1 - u^2)^{\mu-1} du \right| \leq cn^{2\lambda-2\mu-1} \frac{(|a| + n^{-1}\sqrt{1 - |a| - |b|} + n^{-2})^{-\mu}}{(1 + n\sqrt{1 - |a| - |b|})^{\lambda-\mu}}.$$

Denote briefly  $A(x) := \sqrt{1 - \|x\|_2^2}$ . We apply the above inequality with  $a = A(x)A(y)$  and  $b = \langle x, y \rangle$ . Setting  $\|x\|_2 =: \cos \theta$  and  $\|y\|_2 =: \cos \phi$ ,  $0 \leq \theta, \phi \leq \pi$ , we have

$$\begin{aligned} 1 - |a| - |b| &\geq 1 - \|x\|_2 \cdot \|y\|_2 - \sqrt{1 - \|x\|_2^2} \sqrt{1 - \|y\|_2^2} = 1 - \cos(\theta - \phi) \\ &= 2 \sin^2 \frac{\theta - \phi}{2} \geq c(\theta - \phi)^2 \geq c(\sin \theta - \sin \phi)^2 = c(A(x) - A(y))^2, \end{aligned}$$

and hence

$$|P_n(w_\mu; x, y)| \leq cn^{2\lambda-2\mu} (A(x)A(y) + n^{-1}|A(x) - A(y)| + n^{-2})^{-\mu}. \tag{10.9}$$

Here we used that  $(1 + n\sqrt{1 - |a| - |b|})^{\lambda-\mu} \geq 1$ . Now, from  $A(x), A(y) \geq 0$  it easily follows that

$$A(x)A(y) + n^{-1}|A(x) - A(y)| + n^{-2} \sim (A(x) + n^{-1})(A(y) + n^{-1}). \tag{10.10}$$

This coupled with (10.9) yields (10.8).  $\square$

The next lemma gives an analogue of estimate (4.12) on the ball.

**Lemma 10.3.** *For any  $x, y, \xi \in B^d$  such that  $\rho(x, \xi) \leq c^*n^{-1}$ ,*

$$|P_n(w_\mu; x, y) - P_n(w_\mu; \xi, y)| \leq \frac{cn^d \rho(x, \xi)}{\sqrt{W_\mu(n; x)}\sqrt{W_\mu(n; y)}}, \tag{10.11}$$

where the constant  $c > 0$  depends only on  $\mu, d$ , and  $c^*$ .

The proof of this lemma is somewhat lengthy and will be given in Appendix A.

### 10.2. Localized cross product basis kernels

We consider orthogonal polynomials on  $B^{d_1} \times [-1, 1]^{d_2}$  with weight

$$w_{\mu, \alpha, \beta}(x) := w_\mu(x')w_{\alpha, \beta}(x''), \quad x = (x', x''), \quad x' \in B^{d_1}, \quad x'' \in [-1, 1]^{d_2},$$

where  $w_\mu(x') := (1 - \|x'\|_2^2)^{\mu-1/2}$ ,  $\mu \geq 0$ , and  $w_{\alpha, \beta}(x'') := \prod_{j=1}^{d_2} w_{\alpha_j, \beta_j}(x''_j)$  with  $\alpha_j, \beta_j \geq -1/2$  as in (1.5).

Denote by  $V_n$  the set of all algebraic polynomials of degree  $n$  in  $d_1$  variables which are orthogonal to the lower degree polynomials in  $L^2(B^{d_1}, w_\mu)$  and let  $P_n(w_\mu, x', y')$  be the kernel of the orthogonal projector  $\text{Proj}_n : L^2(B^{d_1}, w_\mu) \rightarrow V_n$ , see (10.1)–(10.2).

We are interested in kernels of the form

$$\Lambda_n(x, y) := \sum_{(j, \nu) \in \mathbb{N}_0 \times \mathbb{N}_0^{d_2}} \hat{A}\left(\frac{j}{n}, \frac{\nu}{n}\right) P_j(w_\mu; x', y') \tilde{P}_\nu^{(\alpha, \beta)}(x'') \tilde{P}_\nu^{(\alpha, \beta)}(y''). \tag{10.12}$$

Here  $\hat{A} \in C^\infty[0, \infty)^{1+d_2}$  is an admissible cutoff function in the sense of Definition 3.1 and  $\tilde{P}_\nu^{(\alpha, \beta)}$  are the tensor product Jacobi polynomials defined as in (4.1). To estimate the localization of  $\Lambda_n(x, y)$  we need the weight

$$W_{\mu, \alpha, \beta}(n; x) := W_\mu(n; x') W_{\alpha, \beta}(n; x''), \tag{10.13}$$

where  $W_\mu(n; x')$  is defined as in (10.4) and  $W_{\alpha, \beta}(n; x'')$  as in (4.3). We also need the distance  $\rho_*(x, y)$  on  $B^{d_1} \times [-1, 1]^{d_2}$  defined by

$$\rho_*(x, y) := \max\left\{\rho(x', y'), \max_{1 \leq j \leq d_2} |\arccos x''_j - \arccos y''_j|\right\},$$

where  $\rho(x', y')$  is the distance on  $B^{d_1}$  defined as in (10.5).

We now give the localization of the kernels  $\Lambda_n(x, y)$  from (10.12):

**Theorem 10.4.** *If  $\hat{A} \in C^\infty[0, \infty)^{1+d_2}$  is an admissible cutoff function in the sense of Definition 3.1, then for any  $\sigma > 0$  there exists a constant  $c > 0$  such that*

$$|\Lambda_n(x, y)| \leq \frac{cn^{d_1+d_2}}{\sqrt{W_{\mu, \alpha, \beta}(n; x)}\sqrt{W_{\mu, \alpha, \beta}(n; y)}} (1 + n\rho_*(x, y))^{-\sigma} \tag{10.14}$$

for  $x, y \in B^{d_1} \times [-1, 1]^{d_2}$ .

This theorem is an immediate consequence of Theorems 4.1, 10.1, and Lemma 10.2. (see the proof of Theorem 4.1).

The analogue of Theorem 4.2 reads as follows:

**Theorem 10.5.** *Let  $\hat{A}$  be an admissible cutoff function which belongs to the class  $\mathcal{S}(d_2 + 1, \mathcal{L}; \gamma, \tilde{\gamma}M)$  for some  $\mathcal{L}$  and  $M$  as in (2.1) and  $\gamma, \tilde{\gamma} > 0$  (see Definition 2.2). Then the kernels from (10.12) satisfy*

$$|\Lambda_n(x, y)| \leq \frac{cn^{d_1+d_2}}{\sqrt{W_{\mu, \alpha, \beta}(n; x)}\sqrt{W_{\mu, \alpha, \beta}(n; y)}} \exp\left\{-\frac{\tilde{c}n\rho_*(x, y)}{\mathcal{L}(n\rho_*(x, y))}\right\} \tag{10.15}$$

for  $x, y \in B^{d_1} \times [-1, 1]^{d_2}$ . Here  $\tilde{c} = c'/\tilde{\gamma}M$ , where  $c' > 0$  is an absolute constant.

Here the argument is the same as for the proof of Theorem 4.2 (see the proof of Theorem 4.1) using Theorem 4.2 and [11, Theorem 6.1].

An analogue of Theorem 4.3 is also valid:

**Theorem 10.6.** *If  $\hat{A} \in C^\infty[0, \infty)^{1+d_2}$  is an admissible cutoff function, then for any  $\sigma > 0$  and for all  $x, y, \xi \in B^{d_1} \times [-1, 1]^{d_2}$  such that  $\rho(x, \xi) \leq c_* n^{-1}$ ,  $n \geq 1$ ,  $c_* > 0$ , the kernel  $\Lambda_n$  from (10.12) satisfies*

$$|\Lambda_n(x, y) - \Lambda_n(\xi, y)| \leq \frac{cn^{d_1+d_2+1}\rho(x, \xi)}{\sqrt{W_{\mu,\alpha,\beta}(n; x)}\sqrt{W_{\mu,\alpha,\beta}(n; y)}}(1 + n\rho(x, y))^{-\sigma}, \quad (10.16)$$

where  $c > 0$  depends only on  $\sigma, d, \alpha, \beta, c_*$ , and  $\hat{A}$ .

The proof of this theorem is quite similar to the proof of Theorem 4.3 and relies on Theorems 4.3, 10.1, and Lemma 10.3.

### 10.3. Construction of needlets on $B^{d_1} \times [-1, 1]^{d_2}$

An important component of our theory is the construction of frames on  $B^{d_1} \times [-1, 1]^{d_2}$ . To this end one uses a Caldeón type formula based on localized kernels as the kernels in (10.12) and a cubature formula. A cubature formula on  $B^{d_1} \times [-1, 1]^{d_2}$  exact for sufficiently large degree polynomials can be constructed as product of the cubature formula on  $B^{d_1}$  from [19, §5] and the cubature on  $[-1, 1]^{d_2}$  from Section 6.1. Once the components are in place, the construction is carried out exactly as in Section 6.2. We skip the details.

### 10.4. Spaces of distributions on $B^{d_1} \times [-1, 1]^{d_2}$

It is natural to use as test functions the set  $\mathcal{D} := C^\infty(B^{d_1} \times [-1, 1]^{d_2})$ , where the topology is defined by the semi-norms  $|\phi|_\mu := \|D^\mu \phi\|_\infty$  for all multi-indices  $\mu$ . Just as in the case of tensor product Jacobi polynomials (Section 5.2) the test functions  $\phi \in \mathcal{D}$  can be characterized by their cross polynomial expansions on  $B^{d_1} \times [-1, 1]^{d_2}$ . The space  $\mathcal{D}'$  of distributions on  $B^{d_1} \times [-1, 1]^{d_2}$  is defined as the set of all continuous linear functionals on  $\mathcal{D}$ .

For an admissible cutoff function  $\hat{A} : [0, \infty)^{1+d_2} \mapsto \mathbb{C}$  of type (b) obeying condition (3.4) we define  $\Phi_0(x, y) := P_0(w_\mu; x', y')\tilde{P}_0^{(\alpha,\beta)}(x'')\tilde{P}_0^{(\alpha,\beta)}(y'')$  and

$$\Phi_j(x, y) := \sum_{(m,v) \in \mathbb{N}_0 \times \mathbb{N}_0^{d_2}} \hat{A}\left(\frac{m}{2^{j-1}}, \frac{v}{2^{j-1}}\right) P_m(w_\mu; x', y') \tilde{P}_v^{(\alpha,\beta)}(x'') \tilde{P}_v^{(\alpha,\beta)}(y''), \quad j \geq 1.$$

Then the weighted Triebel–Lizorkin space  $F_{pq}^{s\rho} := F_{pq}^{s\rho}(w_{\mu,\alpha,\beta})$  with  $s, \rho \in \mathbb{R}$ ,  $0 < p < \infty$ , and  $0 < q \leq \infty$ , is defined as the set of all  $f \in \mathcal{D}'$  such that

$$\|f\|_{F_{pq}^{s\rho}} := \left\| \left( \sum_{j=0}^{\infty} [2^{sj} W_{\mu,\alpha,\beta}(2^j; \cdot)^{-\rho/(d_1+d_2)} |\Phi_j * f(\cdot)|]^q \right)^{1/q} \right\|_p < \infty \quad (10.17)$$

with the usual modification when  $q = \infty$ . Here  $\Phi_j * f$  is defined as in (5.15) and  $\|\cdot\|_p$  is the weighted  $L^p(w_{\mu,\alpha,\beta})$  norm on  $B^{d_1} \times [-1, 1]^{d_2}$ .

The weighted Besov space  $B_{pq}^{s\rho} := B_{pq}^{s\rho}(w_{\mu,\alpha,\beta})$  with  $s, \rho \in \mathbb{R}$  and  $0 < p, q \leq \infty$ , is defined as the set of all  $f \in \mathcal{D}'$  such that

$$\|f\|_{B_{pq}^{s\rho}} := \left( \sum_{j=0}^{\infty} [2^{sj} \|W_{\mu,\alpha,\beta}(2^j; \cdot)^{-\rho/(d_1+d_2)} \Phi_j * f(\cdot)\|_p]^q \right)^{1/q} < \infty, \tag{10.18}$$

where the  $\ell^q$ -norm is replaced by the sup-norm if  $q = \infty$ .

Without going into further details, we note that the theory of Triebel–Lizorkin and Besov space on  $B^{d_1} \times [-1, 1]^{d_2}$  with weight  $w_{\mu,a,b}(x)$  can be further developed in analogy to the spaces on  $[-1, 1]^d$  from Sections 7–8. Also, needlets on  $B^{d_1} \times [-1, 1]^{d_2}$  can be deployed for the decomposition of the  $F$ - and  $B$ -spaces on  $B^{d_1} \times [-1, 1]^{d_2}$  as in Sections 7–8. The point is that all ingredients needed for this theory are either in place or can easily be developed.

### 11. Discussion

Although this paper is mainly concerned with weighted Triebel–Lizorkin and Besov space on  $[-1, 1]^d$  it is one of our goals to show how the theory of  $F$ - and  $B$ -spaces can be developed on products of  $[-1, 1]^{d_1}$ ,  $B^{d_2}$ ,  $\mathbb{S}^{d_3}$ ,  $T^{d_4}$ ,  $\mathbb{R}^{d_5}$ , or  $\mathbb{R}_+^{d_6}$  with weights. For  $B^{d_1} \times [-1, 1]^{d_2}$  a sketch of the main ingredients of the theory was given in the previous section. We believe that the most natural way to define and develop this sort of spaces is via orthogonal decompositions, where kernels like the ones from (2.9), (4.2) or (10.12) play a prominent role.

We would like to turn again our attention to the fundamental question of what kind of cutoff functions  $\hat{A}$  can be used in the case of cross product bases. As was already mentioned in the introduction, as for univariate Jacobi polynomials (see (2.9)) univariate cutoff functions  $\hat{a}$  induce rapidly decaying kernels on the sphere [15], ball [14], simplex [11], and in the context of tensor product Hermite [20] and Laguerre functions [12]. Note that cutoff functions  $\hat{a}$  which are constants around  $t = 0$  are sufficient for the development of the theory in these cases. However, as was already seen truly multivariate cutoff functions  $\hat{A}$  need to be used in the case of product Jacobi polynomials or cross product bases. Moreover, the localization of the respective kernels depends on the behavior of  $\hat{A}$  at the boundary of  $[0, \infty)^d$ , i.e. at the coordinate planes. This is intimately related to the impact of the behavior of the univariate cutoff functions  $\hat{a}$  at  $t = 0$  on the localization of the kernels on the interval, ball, sphere, etc. This behavior appears as a boundary condition on  $\hat{A}$  and becomes an important issue.

The key observation is that (as in Theorem 2.5) the localization results given in the theorems described below hold under the condition that the compactly supported  $C^\infty$  univariate cutoff function  $\hat{a}$  satisfies

$$\hat{a}^{(m)}(0) = 0 \quad \text{for } m = 1, 2, \dots$$

These are: (1) Theorem 4.2 in [19] on the ball, (2) Theorem 2.2 in [15] on the sphere, (3) Theorem 7.1 in [11] on the simplex, (4) Corollary 1 in [20] for tensor product Hermite functions, (5) Theorems 3.2, 3.7, 3.8 in [12] for tensor product Laguerre functions. The proofs of these results utilize the scheme of the proof of Theorem 2.5 with very little variations and will be omitted. Consequently, the cross product basis kernels induced by an admissible cutoff function  $\hat{A}$

(see Definition 3.1) obtained from any combination of the above mentioned bases on  $[-1, 1]^{d_1}$ ,  $B^{d_2}$ ,  $\mathbb{S}^{d_3}$ ,  $T^{d_4}$ ,  $\mathbb{R}^{d_5}$ , or  $\mathbb{R}_+^{d_6}$  will decay rapidly as in Theorems 4.1, 10.4. Further modifications and extensions as in Theorems 4.2, 4.3, 10.5, 10.6 are also almost automatic.

The construction of needlets on products of two or more of the sets  $[-1, 1]^{d_1}$ ,  $B^{d_2}$ ,  $\mathbb{S}^{d_3}$ ,  $T^{d_4}$ ,  $\mathbb{R}^{d_5}$ , or  $\mathbb{R}_+^{d_6}$  follows easily the pattern of the construction on  $[-1, 1]^d$  from Section 6, based on tensor product basis kernels and product cubature formulas.

The ensuing program for developing weighted Triebel–Lizorkin and Besov spaces on products of sets as above can be carried out as for the spaces on  $[-1, 1]^d$  developed in this article.

**Appendix A. Proof of Lemma 10.3.**

For  $\mu = 0$  the expression of  $P_n(w_\mu; x, y)$  in (10.2) simplifies considerably as  $\mu \rightarrow 0$ ; the integral becomes a sum of two terms, as shown in [25]. This case is easier than the case  $\mu > 0$ . We omit its proof.

Assume  $\mu > 0$ . The proof hinges on the following lemma which is an immediate consequence of Lemma 3.5 in [3].

**Lemma A.1.** *Suppose  $\mu > 0$ ,  $0 < |a| < 1$ ,  $\eta \in C^\infty[-1, 1]$  with  $\text{supp } \eta \subset [-\frac{1}{2}, 1]$ . If  $|b| \leq 1 - |a|$ , then*

$$\left| \int_{-1}^1 C_n^\lambda(at + b)\eta(t)(1 - t)^{\mu-1} dt \right| \leq \frac{cn^{2\lambda-2\mu-1}}{|a|^\mu(1 + n\sqrt{1 - |a + b|})^{\lambda-\mu}}.$$

The proof of Lemma 10.3 will be divided into two parts.

Case 1:  $A(x)A(y) \leq 16c^*(n^{-1}\rho(x, y) + n^{-2})$ , where  $c^*$  is the constant from the hypothesis of Lemma 10.3. We shall need the following estimate for Gegenbauer polynomials, which follows from (2.22):

$$|C_n^\lambda(t)| \leq cn^{2\lambda-1}(1 + n\sqrt{1 - t^2})^{-\lambda}, \quad t \in [-1, 1]. \tag{A.1}$$

Denote by  $I_u$  the interval with end points  $t(x, y; u)$  and  $t(\xi, y; u)$ . Then using the identity  $\frac{d}{dt}C_n^\lambda(t) = 2\lambda C_{n-1}^{\lambda+1}(t)$  [23, (4.7.27)], we obtain

$$\begin{aligned} E &:= |P_n(w_\mu; x, y) - P_n(w_\mu; \xi, y)| \\ &\leq cn \int_{-1}^1 |C_n^\lambda(t(x, y; u)) - C_n^\lambda(t(\xi, y; u))|(1 - u^2)^{\mu-1} du \\ &\leq cn \int_{-1}^1 \|C_{n-1}^{\lambda+1}\|_{L^\infty(I_u)} |t(x, y; u) - t(\xi, y; u)|(1 - u^2)^{\mu-1} du. \end{aligned} \tag{A.2}$$

By (A.1) it follows that

$$\begin{aligned} \|C_{n-1}^{\lambda+1}\|_{L^\infty(I_u)} &\leq cn^{2\lambda+1}[(1+n\sqrt{1-t(x,y;u)^2})^{-\lambda-1} \\ &\quad + (1+n\sqrt{1-t(\xi,y;u)^2})^{-\lambda-1}]. \end{aligned} \tag{A.3}$$

If  $t(x, y; u) \geq 0$ , then

$$\begin{aligned} 1 - t(x, y; u)^2 &\geq 1 - t(x, y; u) \geq 1 - \langle x, y \rangle - A(x)A(y) \\ &= 1 - \cos \rho(x, y) = 2 \sin^2(\rho(x, y)/2) \geq (2/\pi^2)\rho(x, y)^2, \end{aligned}$$

and similarly if  $t(x, y; u) < 0$ , then

$$\begin{aligned} 1 - t(x, y; u)^2 &\geq 1 + t(x, y; u) \geq 1 + \langle x, y \rangle - A(x)A(y) \\ &= 1 - \langle x, -y \rangle - A(x)A(-y) \geq (2/\pi^2)\rho(x, -y)^2. \end{aligned}$$

The above estimates along with (A.3) and  $\rho(x, \xi) \leq c^*n^{-1}$  yield

$$\begin{aligned} \|C_{n-1}^{\lambda+1}\|_{L^\infty(I_u)} &\leq cn^{2\lambda+1}(1+n\rho(x,y))^{-\lambda-1} \quad \text{if } t(x,y;u) \geq 0, \quad \text{and} \\ \|C_{n-1}^{\lambda+1}\|_{L^\infty(I_u)} &\leq cn^{2\lambda+1}(1+n\rho(x,-y))^{-\lambda-1} \quad \text{if } t(x,y;u) < 0. \end{aligned}$$

We use these inequalities in (A.2) to obtain

$$\begin{aligned} E &\leq cn^{2\lambda+2} \int_{-1}^1 \frac{|t(x,y;u) - t(\xi,y;u)|}{(1+n\rho(x,y))^{\lambda+1}} (1-u^2)^{\mu-1} du \\ &\quad + cn^{2\lambda+2} \int_{-1}^1 \frac{|t(x,y;u) - t(\xi,y;u)|}{(1+n\rho(x,-y))^{\lambda+1}} (1-u^2)^{\mu-1} du \\ &=: E_1 + E_2. \end{aligned}$$

To estimate  $E_1$  and  $E_2$  we shall need the inequality

$$|A(x) - A(y)| \leq \rho(x, y), \quad x, y \in B^d, \tag{A.4}$$

given in [19, Lemma 4.1] with a factor  $\sqrt{2}$  on the right. To establish this inequality, we may assume (by rotation) that  $x = (x_1, 0, \dots, 0)$ ,  $x_1 \geq 0$ . Observe that the left-hand side of (A.4) depends only on  $\|x\|_2$  and  $\|y\|_2$ , while the minimum of  $\rho(x, y)$  over all  $y \in B^d$  such that  $\|y\|_2 = \text{constant}$  occurs when  $y = (y_1, 0, \dots, 0)$ ,  $y_1 \geq 0$ . This leads us to the conclusion that it suffices to prove (A.4) in dimension  $d = 1$  only. But in this case (A.4) is trivial.

Inequality (A.4) implies

$$A(y) + n^{-1} \leq A(x) + n^{-1} + \rho(x, y) \leq (A(x) + n^{-1})(1 + n\rho(x, y)). \tag{A.5}$$

On the other hand, by (10.10), (A.4), and our assumption it follows that

$$\begin{aligned} (A(x) + n^{-1})(A(y) + n^{-1}) &\leq c(A(x)A(y) + n^{-1}|A(x) - A(y)| + n^{-2}) \\ &\leq cn^{-2}(1 + n\rho(x, y)). \end{aligned} \tag{A.6}$$

This along with (A.5) gives

$$A(y)^2 \leq (A(x) + n^{-1})(A(y) + n^{-1})(1 + n\rho(x, y)) \leq cn^{-2}(1 + n\rho(x, y))^2. \tag{A.7}$$

As in [19, p. 136] we have using (A.4)

$$\begin{aligned} |t(x, y; u) - t(\xi, y; u)| &\leq |\cos \rho(x, y) - \cos \rho(\xi, y)| + |1 - u|A(y)\rho(x, \xi) \\ &\leq \rho(x, \xi)(\rho(x, y) + \rho(\xi, y)) + A(y)\rho(x, \xi). \end{aligned} \tag{A.8}$$

Combining this with (A.7) and  $\rho(x, \xi) \leq cn^{-1}$  we get

$$|t(x, y; u) - t(\xi, y; u)| \leq cn^{-1}\rho(x, \xi)(1 + n\rho(x, y)).$$

This estimate coupled with (A.6) leads to

$$E_1 \leq cn^d \rho(x, \xi) \frac{n^{2\mu}}{(1 + n\rho(x, y))^\mu} \leq \frac{cn^d \rho(x, \xi)}{(A(x) + n^{-1})^\mu (A(y) + n^{-1})^\mu}. \tag{A.9}$$

To estimate  $E_2$  we observe that  $t(x, y; u) = -t(x, -y; -u)$  and hence

$$|t(x, y; u) - t(\xi, y; u)| = |t(x, -y; -u) - t(\xi, -y; -u)|.$$

Consequently,  $E_2$  can be estimated exactly as  $E_1$  with the same bound as in (A.9). These two estimates yield (10.11).

*Case 2:*  $A(x)A(y) > 8c^*(n^{-1}\rho(x, y) + n^{-2})$ . In this case by (10.10) and (A.4) it readily follows that

$$A(x)A(y) \sim (A(x) + n^{-1})(A(x) + n^{-1}). \tag{A.10}$$

Let  $\eta_+$  be a  $C^\infty$  function such that  $\eta_+(u) = 1$  for  $\frac{1}{2} \leq u \leq 1$ , and  $\eta_+(u) = 0$  for  $-1 \leq u \leq -\frac{1}{2}$ . Define  $\eta_-(u) := 1 - \eta_+(u)$ . Then on account of (10.2), we can write

$$P_n(w_\mu; x, y) = P_n^+(w_\mu; x, y) + P_n^-(w_\mu; x, y),$$

where

$$P_n^\pm(w_\mu; x, y) := c_\mu \frac{n + \lambda}{\lambda} \int_{-1}^1 C_n^\lambda(t(x, y; u)) \eta_\pm(u) (1 - u^2)^{\mu-1} du.$$

Since  $t(x, y; -u) = -t(x, -y; u)$  and  $C_n^\lambda(-t) = (-1)^n C_n^\lambda(t)$ , we only need to prove (10.11) for  $P_n^+(w_\mu; \cdot, \cdot)$ . We write  $t(x, y; u)$  as

$$t(x, y; u) = B(x, y) + A(x)A(y)(u - 1) \quad \text{with } B(x, y) := \cos \rho(x, y).$$

In going further, we have

$$P_n^+(w_\mu; x, y) - P_n^+(w_\mu; \xi, y) = J_1 + J_2, \tag{A.11}$$

where

$$\begin{aligned} J_1 &:= c_\mu \frac{n + \lambda}{\lambda} \int_{-1}^1 [C_n^\lambda(B(x, y) + A(x)A(y)(u - 1)) \\ &\quad - C_n^\lambda(B(\xi, y) + A(x)A(y)(u - 1))] \eta_+(u) (1 - u^2)^{\mu-1} du, \\ J_2 &:= c_\mu \frac{n + \lambda}{\lambda} \int_{-1}^1 [C_n^\lambda(B(\xi, y) + A(x)A(y)(u - 1)) \\ &\quad - C_n^\lambda(B(\xi, y) + A(\xi)A(y)(u - 1))] \eta_+(u) (1 - u^2)^{\mu-1} du. \end{aligned}$$

To estimate  $|J_1|$ , we again use  $\frac{d}{ds} C_n^\lambda(s) = 2\lambda C_{n-1}^{\lambda+1}(s)$  to write

$$J_1 = 2c_\mu(n + \lambda) \int_{B(\xi, y)}^{B(x, y)} \int_{-1}^1 C_{n-1}^{\lambda+1}(s + A(x)A(y)(u - 1)) \eta_+(u) (1 - u^2)^{\mu-1} du ds.$$

We estimate the inner integral above using Lemma A.1 with  $\eta(t) = \eta_+(t)(1 + t)^{\mu-1}$ ,  $b = s - A(x)A(y)$  and  $a = A(x)A(y)$ . We get

$$|J_1| \leq c \frac{n^{2(\lambda+1)-2\mu}}{[A(x)A(y)]^\mu} \left| \int_{B(\xi, y)}^{B(x, y)} \frac{1}{(1 + n\sqrt{1 - |s|})^{\lambda+1-\mu}} ds \right|.$$

As in (A.8)

$$|B(x, y) - B(\xi, y)| \leq \rho(x, \xi)(\rho(x, y) + \rho(\xi, y)) \leq c\rho(x, \xi)(\rho(x, y) + n^{-1}).$$

On the other hand  $1 - B(x, y) = 1 - \cos \rho(x, y) \geq c\rho(x, y)^2$  and similarly  $1 - B(\xi, y) \geq c\rho(\xi, y)^2$ . Therefore,

$$|J_1| \leq c \frac{n^{d+1}}{[A(x)A(y)]^\mu} \frac{(\rho(x, y) + n^{-1})\rho(x, \xi)}{(1 + n \min\{\rho(x, y), \rho(\xi, y)\})^{\lambda+1-\mu}} \leq \frac{cn^d \rho(x, \xi)}{[A(x)A(y)]^\mu}, \tag{A.12}$$

where we used that  $\rho(x, y) \leq \rho(\xi, y) + \rho(x, \xi) \leq \rho(\xi, y) + c^*n^{-1}$ .

To estimate  $|J_2|$ , we again use  $\frac{d}{ds}C_n^\lambda(s) = 2\lambda C_{n-1}^{\lambda+1}(s)$  to express  $J_2$  as

$$-2c_\mu(n + \lambda)A(y) \int_{A(\xi)-1}^{A(x)-1} \int_{A(\xi)-1}^{A(x)-1} C_{n-1}^{\lambda+1}(B(\xi, y) + sA(y)(u - 1))\eta_+(u)(1 + u)^{\mu-1}(1 - u)^\mu du ds.$$

We estimate the inner integral by using Lemma A.1 with  $\eta(t) = \eta_+(t)(1 + t)^{\mu-1}$ ,  $b = B(\xi, y) - sA(y)$ ,  $a = sA(y)$ , and  $\lambda, \mu$  replaced by  $\lambda + 1, \mu + 1$  to obtain

$$|J_2| \leq \frac{cn^{2(\lambda+1)-2(\mu+1)}}{A(y)^\mu(1 + n\sqrt{1 - B(\xi, y)})^{\lambda-\mu}} \left| \int_{A(\xi)}^{A(x)} s^{-\mu-1} ds \right|$$

and using that  $|A(x) - A(\xi)| \leq \sqrt{2}\rho(x, \xi)$  (see (A.4))

$$|J_2| \leq \frac{cn^{d-1}\rho(x, \xi)}{A(y)^\mu \min\{A(x)^{\mu+1}, A(\xi)^{\mu+1}\}}. \tag{A.13}$$

By the same token and since by assumption  $\rho(x, \xi) \leq c^*n^{-1}$  we have

$$A(\xi) \geq A(x) - |A(x) - A(\xi)| \geq A(x) - \sqrt{2}\rho(x, \xi) \geq A(x) - \sqrt{2}c^*n^{-1}.$$

If  $A(x) \geq 2\sqrt{2}c^*n^{-1}$ , then from above  $A(\xi) \geq A(x)/2$ . These two estimates and (A.13) imply that  $|J_2|$  has the bound of  $|J_1|$  from (A.12), and using (A.10) estimate (10.11) holds for  $|P_n^+(w_\mu; x, y) - P_n^+(w_\mu; \xi, y)|$ .

Let  $A(x) < 2\sqrt{2}c^*n^{-1}$ . We claim that  $A(y) < 4\sqrt{2}c^*n^{-1}$ . Indeed, suppose  $A(y) \geq 4\sqrt{2}c^*n^{-1}$ . Then  $A(y) \geq A(x)/2$  and using (A.4), we get

$$\sqrt{2}\rho(x, y) \geq |A(x) - A(y)| \geq A(y) - A(x) \geq A(y)/2$$

and hence  $A(x)A(y) > 8c^*n^{-1}\rho(x, y) \geq 2\sqrt{2}c^*n^{-1}A(y)$  yielding  $A(x) > 2\sqrt{2}c^*n^{-1}$ , that is a contradiction. Therefore,  $A(x)A(y) < 16c^*n^{-2}$ . Thus  $A(x), A(y)$  obey the conditions of Case 1 and hence estimate (10.11) holds true. This completes the proof of Lemma 10.3.

**References**

[1] G. Brown, F. Dai, Approximation of smooth functions on compact two-point homogeneous spaces, *J. Funct. Anal.* 220 (2005) 401–423.  
 [2] R. Coifman, G. Weiss, *Analyse harmonique non-commutative sur certains espaces homogènes*, Lecture Notes in Math., vol. 242, Springer, Berlin, 1971.  
 [3] F. Dai, Yuan Xu, Cesàro means of orthogonal expansions in several variables, *Constr. Approx.* 29 (2009) 129–155.  
 [4] Z. Ditzian, V. Totik, *Moduli of Smoothness*, Springer-Verlag, New York, 1987.  
 [5] T. Erdelyi, A. Magnus, P. Nevai, Generalized Jacobi weights, Christoffel functions, and Jacobi polynomials, *SIAM J. Math. Anal.* 25 (1994) 602–614.  
 [6] M. Frazier, B. Jawerth, Decomposition of Besov spaces, *Indiana Univ. Math. J.* 34 (1985) 777–799.  
 [7] M. Frazier, B. Jawerth, A discrete transform and decompositions of distribution, *J. Funct. Anal.* 93 (1990) 34–170.

- [8] M. Frazier, B. Jawerth, G. Weiss, Littlewood–Paley Theory and the Study of Function Spaces, CBMS Reg. Conf. Ser. Math., vol. 79, Amer. Math. Soc., 1991.
- [9] C.L. Frenzen, R. Wong, A uniform asymptotic expansion of the Jacobi polynomials with error bounds, *Canad. J. Math.* 37 (1985) 979–1007.
- [10] L. Hörmander, *The Analysis of Linear Partial Differential Operators I*, Springer-Verlag, 1983.
- [11] K. Ivanov, P. Petrushev, Yuan Xu, Sub-exponentially localized kernels and frames induced by orthogonal expansions, *Math. Z.* 264 (2010) 361–397.
- [12] G. Kerkyacharian, P. Petrushev, D. Picard, Yuan Xu, Decomposition of Triebel–Lizorkin and Besov spaces in the context of Laguerre expansions, *J. Funct. Anal.* 256 (2009) 1137–1188.
- [13] G. Kyriazis, P. Petrushev, Yuan Xu, Jacobi decomposition of weighted Triebel–Lizorkin and Besov spaces, *Studia Math.* 186 (2008) 161–202.
- [14] G. Kyriazis, P. Petrushev, Yuan Xu, Decomposition of weighted Triebel–Lizorkin and Besov spaces on the ball, *Proc. London Math. Soc.* 97 (2008) 477–513.
- [15] F.J. Narcowich, P. Petrushev, J.D. Ward, Decomposition of Besov and Triebel–Lizorkin spaces on the sphere, *J. Funct. Anal.* 238 (2006) 530–564.
- [16] P. Nevai, Orthogonal polynomials, *Mem. Amer. Math. Soc.* 18 (1979).
- [17] J. Peetre, *New Thoughts on Besov Spaces*, Duke Univ. Math. Ser., Mathematics Department, Duke University, Durham, NC, 1976.
- [18] P. Petrushev, Yuan Xu, Localized polynomial frames on the interval with Jacobi weights, *J. Fourier Anal. Appl.* 11 (2005) 557–575.
- [19] P. Petrushev, Yuan Xu, Localized polynomial frames on the ball, *Constr. Approx.* 27 (2008) 121–148.
- [20] P. Petrushev, Yuan Xu, Decomposition of spaces of distributions induced by Hermite expansion, *J. Fourier Anal. Appl.* 14 (2008) 372–414.
- [21] J. Riordan, *Combinatorial Identities*, John Wiley, New York, 1968.
- [22] E. Stein, *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*, Princeton University Press, Princeton, NJ, 1993.
- [23] G. Szegő, *Orthogonal Polynomials*, fourth ed., Amer. Math. Soc. Colloq. Publ., vol. 23, Amer. Math. Soc., Providence, RI, 1975.
- [24] H. Triebel, *Theory of Function Spaces*, Monogr. Math., vol. 78, Birkhäuser Verlag, Basel, 1983.
- [25] Yuan Xu, Summability of Fourier orthogonal series for Jacobi weight on a ball in  $\mathbb{R}^d$ , *Trans. Amer. Math. Soc.* 351 (1999) 2439–2458.