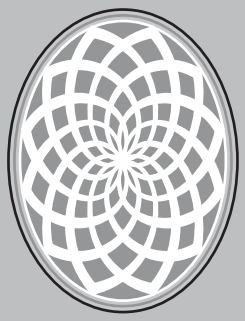

Industrial Mathematics Institute

2009:02

Decomposition of spaces of distributions induced by tensor product bases

K. Ivanov, P. Petrushev and Y. Xu

IMI

Preprint Series

Department of Mathematics
University of South Carolina

DECOMPOSITION OF SPACES OF DISTRIBUTIONS INDUCED BY TENSOR PRODUCT BASES

KAMEN IVANOV, PENCHO PETRUSHEV AND YUAN XU

Abstract. Rapidly decaying kernels and frames (needlets) in the context of tensor product Jacobi polynomials are developed based on several constructions of multivariate C^∞ 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.

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 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 $Q^{d_1} \times \mathbb{S}^{d_2}$, $\mathbb{S}^{d_1} \times B^{d_2}$, $Q^{d_1} \times T^{d_2}$, $Q^{d_1} \times \mathbb{R}^{d_2}$, $B^{d_1} \times \mathbb{R}^{d_2}_+$, $Q^{d_1} \times B^{d_2} \times \mathbb{R}^{d_3}_+$, etc. Here $\mathbb{S}^d$ 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}, $\mathbb{S}^{d_3}$, 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 [13]. 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 $\mathrm{Proj}_n : L^2(B^d, w_\mu) \mapsto \mathcal{V}_n^d$ can be written in the form

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

Date: February 9, 2009.

The second author has been supported by NSF Grant DMS-0709046 and the thid author by NSF Grant DMS-0604056.

To introduce weighted Triebel-Lizorkin (F-spaces) and Besov spaces (B-spaces) on B^d (see [16], [23] for the general idea), let

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

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

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

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

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

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

We refer the reader to [13] 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 α, β, ν 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

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

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 [10] 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 [12], sphere [14], simplex [10], and more surprisingly in the context of tensor product Hermite [19] and Laguerre functions [11].

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

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

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^∞ 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. $t=(t_1,\dots,t_{k-1},0,t_{k+1},\dots,t_d)$ for some $1\le k\le d$,*

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

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 $\tau=(\tau_1,\dots,\tau_{k-1},0,\tau_{k+1},\dots,\tau_d)$, $1\le k\le d$, $\hat{A}(t)=$ constant for $t\in[\tau,\tau+c_*e_k]$ with e_k being the kth coordinate vector.*

The point is that either of these conditions combined with $\hat{A}$ being C^∞ 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 ρ in the definition of weighted Triebel-Lizorkin and Besov spaces on $[-1,1]^d$. Thus we introduce F-spaces by the norms

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

and B-spaces by the norms

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

(see §§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 $B^{ss}_{\tau\tau}$ 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 $\{\langle f, \varphi_\xi\rangle\}$ in appropriate sequence norms. They can be viewed as an analogue of the φ-transform of Frazier and Jawerth [5, 6, 7].

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 [12], and on the ball [13]. Therefore, to spare the reader the repetition of well established arguments we shall only exhibit the essential differences and refer for the rest to [12, 13]. We shall place the emphasis on the development of multivariable cutoff functions and the associated tensor product Jacobi kernels which defer 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 nonstandard settings such as on the sphere [14], ball [13] as well as in the frameworks of Hermite [19] and Laguerre [11] expansions. It is also closely related to the development of sub-exponentially localized Jacobi and other kernels and needlets in [10].

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 §2 we review some basic results from [10, 12] and prove new results about admissible univariate cutoff functions and the localization of the respective kernels induced by univariate Jacobi polynomials. In §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 §4 the localization results of the corresponding tensor product Jacobi polynomial are established. In §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 §6 we utilize kernels associated to cutoff functions of type (b) and (c) to the construction of frame elements (needlets). In §§7-8 we further use these kernels to define "correctly" the weighted Triebel-Lizorlin 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 §10 we briefly consider weighted spaces of distributions on $B^{d_1} \times [-1,1]^d$. In §11 we discuss various aspects of distribution spaces on product domains and tensor product bases. Section 12 is an appendix, where we place the lengthy proof of a lemma from §10.

Some useful notation. Throughout we shall denote

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

and $\|f\|_\infty := \sup_{x\in[-1,1]^d} |f(x)|$. 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 $|x| = \|x\|_1 := \sum_i |x_i|$. Π^d_n will denote the set

of all algebraic polynomials of total degree $\le n$ in d variables. Positive constants will be denoted by c, c_1, c', ... and they may vary at every occurrence, $a \sim b$ will stand for $c_1 a \le b \le c_2 a$.

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 [10, 12, 17]. 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* $\operatorname{supp}\hat{a} \subset [0,2]$ *and $\hat{a}^{(m)}(0) = 0$ for $m \ge 1$. 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) $\operatorname{supp}\hat{a} \subset [1/2, 2]$ *or*

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

We next introduce sets of C^∞ 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

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

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

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 $d \ge 1$, 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 $\|\hat{A}\|_\infty \le \gamma$ and*

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

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 [10]. We let $\chi_\delta := \frac{1}{2\delta}\mathbb{1}_{[-\delta,\delta]}$ and select $\delta_j := \frac{1}{(j+1)\mathcal{L}(j)}$ for $j \ge 0$. Apparently

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

We define

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

Just as in [9, Theorem 1.3.5] we have $\varphi \in C^\infty$, $\varphi \ge 0$, $\operatorname{supp}\varphi \subset [-M, M]$ and

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

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

We now set $\psi(t) := 2M\varphi(2Mt)$ and define $g(t) := \frac{\pi}{2}\int_{-\infty}^t \psi(s)ds$. Evidently, $g \in C^\infty(\mathbb{R})$, $\operatorname{supp} g' \subset [-\frac12, \frac12]$, $g(t) + g(-t) = \frac{\pi}{2}$ for $t \in \mathbb{R}$, $0 \le g \le \pi/2$, $\|g'\|_\infty \le (\pi/2)\|\psi\|_\infty \le (\pi/2)M$ and

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

Apparently $\hat a(t) := \frac{2}{\pi} g(\frac32 - t)$ 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, $\phi(t)^2 + \phi(-t)^2 = 1$ for $t \in \mathbb{R}$. We define

$$\hat a(t) := \begin{cases} \phi(2t - \frac32) & \text{if} \quad t \in [\frac12, 1], \\ \phi(\frac32 - t) & \text{if} \quad t \in (1, 2], \\ 0 & \text{if} \quad \mathbb{R} \setminus [\frac12, 2]. \end{cases}$$

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

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

Let $t_0 \in (-\frac12, \frac12)$ 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 $k \ge 1$

$$\begin{aligned} |g_k(z)| &\le \frac{\pi}{2}\Big(1 + \frac{M}{4M\mathcal{L}(k-1)} + \sum_{j=2}^k \frac{(2M)^j j!\mathcal{L}(j-1)^j}{j![4M\mathcal{L}(k-1)]^j}\Big) \\ &\le \frac{\pi}{2}\Big(1 + \frac14 + \sum_{j=2}^k \frac{1}{2^j}\Big) = \frac{7\pi}{8} \end{aligned}$$

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

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

which implies (2.5). □

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 [10, 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 $\alpha, \beta \ge -1/2$. The Jacobi polynomials are traditionally normalized by $P_n^{(\alpha,\beta)}(1) = \binom{n+\alpha}{n}$. It is well known that [22, (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 nth degree Jacobi polynomial normalized in $L^2([-1,1], w_{\alpha,\beta})$.

We are interested in kernels of the form

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

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

In [17] (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^∞ cutoff functions $\hat{a}$ which are constants around $t = 0$. It was also proved in [10] 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^∞ and compactly supported does not guarantee rapid decay of the kernels $L_n^{\alpha,\beta}(x,y)$.

Here we extend that localization result from [17] 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

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

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 $k \ge 1$, $\operatorname{supp} \hat{a} \subset [0,2]$, 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*

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

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

As in [17] 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. [17])

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

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

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

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

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 [10] 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 [10].

We trivially have (see (4.8) in [10]) $|\mathcal{Q}_n^{\alpha,\beta}(\cos\theta)| \le cn^{2\alpha+2}$, which gives (2.13) ($r = 0$) for $0 \le \theta \le 1/n$.

The following identity is crucial in estimating $|\mathcal{Q}_n^{\alpha,\beta}(\cos\theta)|$ [22, (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 $A_0(t) := (2t+\alpha+\beta+1)\hat{a}\left(\frac{t}{n}\right)$ and inductively

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

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

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

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 $A_1^{(m)}(t) = \frac{1}{n^{m+1}}\int_0^1 \hat{a}^{(m+1)}(\frac{t+s}{n})ds$, which leads to

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

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

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

Therefore,

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

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

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

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

$$|A_k(t)| \le \frac{c(k)}{(t+1)^{2k-1}}\Big(\frac{t+k}{n}\Big)^{2k-1} \max_{2k\le\ell\le 3k-1}\big\|\hat a^{(\ell)}\big\|_{L^\infty[0,\frac{t+k}{n}]} \le \frac{c(k)}{n^{2k-1}}\big\|\hat a^{(3k-1)}\big\|_\infty. \tag{2.20}$$

Indeed, estimate (2.18) gives (2.19) for $l=1$. Suppose (2.19) holds for some $l\ge 1$ and all $m\ge 0$ such that $m+l\le k$. 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

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

and using the inductive assumption

$$\begin{aligned}|A_{l+1}^{(m)}(t)| &\le c\max_{2k\le\ell\le 2k+m+l}\big\|\hat a^{(\ell)}\big\|_{L^\infty[0,\frac{t+l+1}{n}]}\\ &\qquad\times\sum_{\nu=0}^{m+1}\frac{1}{(t+1)^{\nu+2l-1}}\Big(\frac{t+l+1}{n}\Big)^{2k-1}\frac{1}{(t+1)^{m+2-\nu}}\\ &\le \frac{c}{(t+1)^{m+2l+1}}\Big(\frac{t+l+1}{n}\Big)^{2k-1}\max_{2k\le\ell\le 2k+m+l}\big\|\hat a^{(\ell)}\big\|_{L^\infty[0,\frac{t+l+1}{n}]},\end{aligned}$$

which confirms (2.19).

We next prove (2.13) ($r=0$) for $1/n\le\theta\le\pi/2$. By (2.7) it readily follows that $h_n^{(\alpha+k,\beta)}\le c2^k/n$ and it is well known that (see e.g. (4.18) in [10])

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

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

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

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

Let $\pi/2<\theta\le\pi-1/n$. Similarly as in [10]

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

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

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

which implies (2.13).

In the case $\pi-1/n\le\theta\le\pi$ estimate (2.13) follows from the above estimate exactly as in [10]. This completes the proof of estimate (2.13) in the case $r=0$. □

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*

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

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 α, β, $\gamma,\tilde{\gamma}$ 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

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

and then proceeds exactly as in the proof of Theorem 4.1 in [10]. The proofs are nearly identical to the ones in [10] and will be omitted.

Remark 2.8. Theorem 2.7 remains true if we require that $\hat{a}\in\mathcal{S}(1,\mathcal{L};\gamma,\tilde{\gamma}M)$, $\operatorname{supp}\hat{a}\subset[0,2]$, and $\hat{a}$ be a constant on $[0,\delta]$ for a fixed $\delta\in(0,1)$. Then c and c' will depend on δ 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 [10] results similar to Theorem 2.7 are proved on the sphere, ball, simplex and in the context of Hermite and Jaguerre 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 α,β such that*

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

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

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

we get (2.22) from the inequality

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

established in [4, Theorem 1]. For the remaining x estimate (2.22) follows from above invoking Theorem 8.4.8 in [3, p. 108]. □

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 $k>2\alpha+2\beta+5$, $\operatorname{supp}\hat{a}\subset[0,2]$, and $\hat{a}^{(r)}(0)=0$ for $r=1,2,\dots,3k-1$. Then for any $x,y,\xi\in[-1,1]$ such that $\rho(x,\xi)\le c_*n^{-1}$, $n\ge1$, $c_*>0$, the kernel $L_n^{\alpha,\beta}$ from* (2.9) *satisfies*

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

where $\sigma = k - 2\alpha - 2\beta - 5$ and c depends only on k, α, β, 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 [12]. The limit cases $\alpha = -1/2$ or $\beta = -1/2$ are treated as in the proof of [10, 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 $1 \le k \le d$ we define $\mathrm{proj}_k : \mathbb{R}^d \to \mathbb{R}^d$ by

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

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.

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

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 $\operatorname{supp} \hat{A} \subset [0,2]^d$ and $\hat{A}$ obeys the First Boundary Condition, introduced in §1.2, i.e. for any $t \in [0,\infty)^d$ of the form $t = \mathrm{proj}_k t$ for some $1 \le k \le d$ 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 $\operatorname{supp} \hat{A} \subset [0,2]^d$, $\hat{A}(t) = \hat{A}(\mathrm{proj}_k t)$ for every $t = (t_1, \dots, t_d) \in [0,\infty)^d$ such that $t_k \le \frac{1}{2}\|t\|_\infty$ 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 $\tau = \mathrm{proj}_k \tau$ for some $1 \le k \le d$, 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 §1.2 with constant $c_* = \frac{1}{4d-2}$. Consequently, any admissible cutoff function of second kind is admissible of first kind as well. To see the above one simply has to observe that $\frac{1}{4d-2}$ is the ℓ^∞-distance of the set

$$\Big(\bigcap_{k=1}^d \Big\{ t_k \ge \frac{1}{2}\|t\|_\infty \Big\} \Big) \bigcap \Big\{ \|t\|_1 \ge \frac{1}{2} \Big\}$$

from the coordinate hyperplanes. Note also that for $d = 1$ 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 for 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 §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 (§3.4, §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 univatiate 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 $\hat{A} \in C^\infty[0,\infty)^d$, $\operatorname{supp} \hat{A} \subset 2\mathcal{B}_\infty$ and $\hat{A}(t) = 1$ if $t \in \mathcal{B}_\infty \supset \mathcal{B}_1$. Furthermore, $\hat{A}(t) = \hat{A}(\operatorname{proj}_k t)$ for all $t \in [0,\infty)^d$ such that $\|t\|_\infty < 2$ and $t_k \le 1$. From this and $\operatorname{supp} \hat{A} \subset 2\mathcal{B}_\infty$ it follows that $\hat{A}(t) = \hat{A}(\operatorname{proj}_k t)$ for all $t = (t_1, \dots, t_d) \in [0,\infty)^d$ such that $t_k \le \frac{1}{2}\|t\|_\infty$. $\square$

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 $\operatorname{supp} \hat{A}$ essentially cover the whole spectrum. More precisely, we shall need admissible functions $\hat{A}$, which obey the following **dyadic covering condition**:

$$\begin{gathered}\text{For any } t \in [0,\infty)^d \text{ with } \|t\|_\infty = 1 \text{ there is } 0 < \gamma \le 1 \text{ such that} \\ \inf_{\lambda \in [\gamma, 2\gamma]} |\hat{A}(\lambda t)| \ge c > 0.\end{gathered} \tag{3.4}$$

Note that this condition yields $\sum_{j=0}^{\infty} |\hat{A}(2^{-j}t)| \ge c > 0$ 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*

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

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 $t \in [0,\infty)^d \backslash 2\mathcal{B}_\infty$. 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 that $2 + \log_2 d$ non-zero terms. Hence $\hat{B} \in C^\infty[0,\infty)^d$. On the other hand, for $t \in [0,\infty)^d \backslash \mathcal{B}_\infty$ 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 $\hat{A}(2^{-j}t) = \hat{A}(\mathrm{proj}_k(2^{-j}t))$, which implies $\hat{B}(t) = \hat{B}(\mathrm{proj}_k\, t)$. □

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 $\mathcal{N} : \mathbb{R}^d \to \mathbb{R}$ is in $C^\infty(\mathbb{R}^d \backslash \{0\})$ and for $t \in \mathbb{R}^d$ obeys*

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

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

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

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 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 $t \in \mathbb{R}^d \backslash \{0\}$ set

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

where $\hat{c}\big(\frac{\tau}{0}\big) := 0$ for every real τ, 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 $\hat{c}\big(\frac{t_k}{t_m}\big) = 0$. Observing also that $\hat{c}\big(\frac{t_j}{t_m}\big) = 1$ if $|t_j| \leq |t_m|$ we see that (3.10) can be rewritten as

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

It follows immediately from (3.11) that $\mathcal{N}$ belongs to $C^\infty(\mathbb{R}^d\backslash\{0\})$ and 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, 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^∞ 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^∞ blending in the above scheme it is used a C^∞ 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^∞ functions of this type:

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

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 the composition $g\circ f \in \mathcal{R}(a_1,b_1,F;\gamma_2,\tilde\gamma_1(\gamma_1\tilde\gamma_2 F(1)+1))$.*

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

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

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 may be 0.

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

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

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

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

If in a multi-index m we increase m_1 and $m_{\ell+j-1}$ by 1 and decrease m_ℓ and m_j by 1, then the quantities $\sum_{j=1}^k jm_j$ 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 $\prod_{j=1}^k F(j)^{jm_j}$ is attained for $m_1 = n-1, m_{k-n+1} = 1$, and $m_j = 0$ if $j\ne 1$ and $j \ne k-n+1$, i.e.

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

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

$$\begin{aligned}\frac{1}{k!}\|D^k(g\circ f)\|_\infty &\le \gamma_2\tilde\gamma_1^k\sum_{n=1}^k(\gamma_1\tilde\gamma_2)^n F(n)^n\binom{k-1}{n-1}F(1)^{n-1}F(k-n+1)^{k-n+1}\\ &\le \gamma_2\tilde\gamma_1^k F(k)^k\sum_{n=1}^k\binom{k-1}{n-1}F(1)^n(\gamma_1\tilde\gamma_2)^n\\ &= \gamma_2\tilde\gamma_1^k F(k)^k\gamma_1\tilde\gamma_2F(1)(\gamma_1\tilde\gamma_2F(1)+1)^{k-1}\\ &\le \gamma_2\big[\tilde\gamma_1(\gamma_1\tilde\gamma_2F(1)+1)F(k)\big]^k.\quad\square\end{aligned}$$

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 $\mathcal{S}(d,\mathcal{L};8,10d(2d-1)M(8(d+2)M+1))$ contains an admissible cutoff function $\hat A$, $0\le\hat A\le 1$, 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 $\hat a\in\mathcal{S}(1,\mathcal{L};8,8M)$, $0\le\hat a\le 1$, 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 $\hat c|_{[0,\infty)}\in\mathcal{S}(1,\mathcal{L};1,2M)$, $0\le\hat c\le 1$, 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, $0\le\hat A\le 1$.

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

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

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

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

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

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

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

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

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

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

For $1\le t_j/t_d\le 2$ 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

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

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 in (3.17)

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

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. □

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 ε^{-2} and not of order ε^{-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 Subsections 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,*

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

or

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

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 §3.5 allow better estimates on the derivatives than those in §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), $0 \le \hat{a}(t) \le 1$ 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 $1/2 \le \|t\|_\infty \le 2$ we fix $1 \le j \le d$. Consider $\hat{C}(t) = (g \circ f)(t_j)$ as function of $t_j \in [1/2, 1]$, where for $\|t\|_\infty \le 1$ we set $f(t_j) = \hat{a}(2t_j)$ and $g(x) = \cos(\lambda x)$ with $\lambda = \frac{\pi}{2}\prod_{m=1,m\ne j}^d \hat{a}(2t_m)$ and for $1 < \|t\|_\infty \le 2$ we set $f(t_j) = \hat{a}(t_j)$ and $g(x) = \sin(\lambda x)$ with $\lambda = \frac{\pi}{2}\prod_{m=1,m\ne j}^d \hat{a}(t_m)$. 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 $g \circ f \in \mathcal{S}(1, \mathcal{L}; 1, (\pi\gamma + 2)\tilde{\gamma}M)$. For $t_j \in [1, 2]$ we use $f(t_j) = \hat{a}(t_j)$ and $g(x) = \sin(\lambda x)$ with $\lambda = \frac{\pi}{2}\prod_{m=1,m\ne j}^d \hat{a}(t_m)$. We apply Lemma 3.9 as $f \in \mathcal{S}(1, \mathcal{L}; \gamma, \tilde{\gamma}M)$, $g \in \mathcal{S}(1, \mathcal{L}; 1, \pi/2)$ and get $g \circ f \in \mathcal{S}(1, \mathcal{L}; 1, (\pi\gamma/2 + 1)\tilde{\gamma}M)$. Consequently, in all cases $\hat{C} \in \mathcal{S}(d, \mathcal{L}; 1, (\pi\gamma + 2)\tilde{\gamma}M)$. □

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

4. Localized tensor product Jacobi polynomial kernels

Denote by $\tilde{P}_n^{(\alpha_j,\beta_j)}$ $(1 \le j \le d)$ the nth degree Jacobi polynomial normalized in $L^2([-1, 1], w_{\alpha_j,\beta_j})$, see §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

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

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

We are interested in kernels of the form

$$\Lambda_n(x, y) := \sum_{\nu\in\mathbb{N}_0^d} \hat{A}\Big(\frac{\nu}{n}\Big)\tilde{P}_\nu^{(\alpha,\beta)}(x)\tilde{P}_\nu^{(\alpha,\beta)}(y), \quad x, y \in [-1, 1]^d. \tag{4.2}$$

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

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

Theorem 4.1. *Suppose $\hat{A} \in C^{3k-1}[0, \infty)^d$ for some $k \ge 1$, $\operatorname{supp}\hat{A} \subset [0, 2]^d$, and for any $t \in [0, 2]^d$ of the form $t = (t_1, \dots, t_{\ell-1}, 0, t_{\ell+1}, \dots, t_d)$, $1 \le \ell \le d$, i.e.*

$t = \mathrm{proj}_\ell\, t$, $\hat{A}$ *satisfies* $D_\ell^m \hat{A}(t) = 0$ *for* $m = 1, 2, \dots, 3k-1$. *Then the kernels from* (4.2) *satisfy*

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

Here the constant c *depends on* k, d, α, β *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 Λ_n from (4.2) as

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

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

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

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

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

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. □

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 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*

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

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, α, β, γ *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 Lip1 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 $\rho(x,\xi) \le c_* n^{-1}$, $n \ge 1$, $c_* > 0$, the kernel Λ_n from (4.2) satisfies*

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

where $\sigma = k - 2\max_i\{\alpha_i+\beta_i\} - 5$ and $c > 0$ depends only on k,d,α,β, 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 $\rho(x, x+\delta e_i) \le c_* n^{-1}$ and $1 \le i \le d$ with e_i being the ith coordinate vector.

As in the proof of Theorem 4.1, without loss of generality 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

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

Applying Theorem 2.10 to the inner sum we get

$$\begin{aligned}
&\Big|\sum_{\nu_d=0}^{\infty}\hat{A}\Big(\frac{\nu_1}{n},\dots,\frac{\nu_d}{n}\Big)\Big(\tilde{P}_{\nu_d}^{(\alpha_d,\beta_d)}(x_d) - \tilde{P}_{\nu_d}^{(\alpha_d,\beta_d)}(x_d+\delta)\Big)\tilde{P}_{\nu_d}^{(\alpha_d,\beta_d)}(y_d)\Big|\\
&\qquad\qquad \le \frac{cn^2\rho(x,x+\delta e_i)}{\sqrt{w_{\alpha_d,\beta_d}(n;x_d)}\sqrt{w_{\alpha_d,\beta_d}(n;y_d)}}\big(1+n\rho_d(x,y)\big)^{-\sigma}.
\end{aligned}$$

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

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

$$\begin{aligned}
&\Lambda_n(x,y) - \Lambda_n(x+\delta e_i, y)\\
&= \sum_{\nu_1=0}^{2n-1}\cdots\sum_{\nu_{d-1}=0}^{2n-1}\Big[\sum_{\nu_d=0}^{\infty}\hat{A}\Big(\frac{\nu_1}{n},\dots,\frac{\nu_d}{n}\Big)\tilde{P}_{\nu_d}^{(\alpha_d,\beta_d)}(x_d)\tilde{P}_{\nu_d}^{(\alpha_d,\beta_d)}(y_d)\Big]\\
&\quad\times \prod_{j=1,\,j\ne i}^{d-1}\tilde{P}_{\nu_j}^{(\alpha_j,\beta_j)}(x_j)\Big(\tilde{P}_{\nu_i}^{(\alpha_i,\beta_i)}(x_i) - \tilde{P}_{\nu_i}^{(\alpha_i,\beta_i)}(x_i+\delta)\Big)\prod_{j=1}^{d-1}\tilde{P}_{\nu_j}^{(\alpha_j,\beta_j)}(y_j).
\end{aligned} \tag{4.11}$$

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 [22, (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

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

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

$$|\tilde P_m^{(\alpha,\beta)}(\cos\theta) - \tilde P_m^{(\alpha,\beta)}(\cos\theta')| \le \frac{cm|\cos\theta - \cos\theta'|}{\sqrt{w_{\alpha,\beta}(m,\cos\theta)}\big(\sin\theta + m^{-1}\big)} \tag{4.12}$$

$$\le \frac{cm \sin|\frac{\theta-\theta'}{2}| \sin|\frac{\theta+\theta'}{2}|}{\sqrt{w_{\alpha,\beta}(m,\cos\theta)}\big(\sin\theta + m^{-1}\big)} \le \frac{cm|\theta-\theta'|}{\sqrt{w_{\alpha,\beta}(m,\cos\theta)}}.$$

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

$$|\tilde P_{\nu_i}^{(\alpha_i,\beta_i)}(x_i) - \tilde P_{\nu_i}^{(\alpha_i,\beta_i)}(x_i+\delta)| \le \frac{c\nu_i \rho(x, x+\delta e_i)}{\sqrt{w_{\alpha_i,\beta_i}(\nu_i, x_i)}} \le \frac{cn\rho(x,\xi)}{\sqrt{w_{\alpha_i,\beta_i}(n,x_i)}}. \tag{4.13}$$

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

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 $|\hat A(t)| \ge c > 0$ for $t \in [1, 1+\delta]^d$, $\delta > 0$. Then for $n \ge 1/\delta$*

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

where $c > 0$ depends only on δ, α, β, and d.

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

$$\int_{[-1,1]^d} |\Lambda_n(x,y)|^2 w_{\alpha,\beta}(y)dy = \sum_{\nu\in\mathbb{N}_0^d} |\hat A(\nu/n)|^2 [\tilde P_\nu^{(\alpha,\beta)}(x)]^2$$

$$\ge \sum_{\nu\in[n,n+\delta n]^d} |\hat A(\nu/n)|^2 [\tilde P_\nu^{(\alpha,\beta)}(x)]^2 \ge c \prod_{i=1}^d \sum_{\nu_i=n}^{n+\lfloor \delta n\rfloor} [\tilde P_{\nu_i}^{(\alpha_i,\beta_i)}(x_i)]^2$$

and the stated lower bound follows from the respective result in the univariate case, given in [12, Proposition 2.4]. □

The rapidly decaying polynomial kernels $\Lambda_n(x,y)$ from (4.2) can be utilized as in the univariate case [12, Proposition 2.6] for establishing Nikolski type inequalities:

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

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

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

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

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 axces 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 $0 < \delta \le \pi$

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

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

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

which confirms (5.3).

By (5.3) it follows that $\mu(B(y,2\delta)) \le c\mu(B(y,\delta))$, i.e. μ is a doubling measure on $[-1,1]^d$ and, therefore, the Fefferman-Stein vector-valued maximal inequality is valid (see [21]): Assuming that $0 < p < \infty, 0 < q \le \infty$ and $0 < t < \min\{p,q\}$, then for any sequence of functions $\{f_k\}_{k=1}^\infty$ on $[-1,1]^d$,

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

We need to estimate $(\mathcal{M}_t \mathbb{1}_{B(y,\delta)})(x)$. Such estimates readily follow by (5.3) and the respective univaruate result in [12, Lemma 2.7].

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

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

and hence

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

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 [12, (2.22)]:

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

for $x, y \in [-1,1]^d$ and $n \ge 1$, 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 $\mathcal{N}_k(\cdot)$, $k \ge 0$.

The proofs of this lemma is easy and similar to the proof of Lemma 2.8 in [12].

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 $\langle f, \phi\rangle := f(\overline{\phi})$. 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 $k \ge 0$ such that*

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

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

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

$$f(\phi) = \lim_{n\to\infty} \langle S_n, \overline{\phi}\rangle = \sum_{\nu\in\mathbb{N}_0^d} \langle f, \tilde P_\nu^{(\alpha,\beta)}\rangle \langle \phi, \tilde P_\nu^{(\alpha,\beta)}\rangle \quad \text{for} \quad \phi \in \mathcal{D}, \tag{5.12}$$

where $S_n := \sum_{|\nu|\le n} \langle f, \tilde P_\nu^{(\alpha,\beta)}\rangle \tilde P_\nu^{(\alpha,\beta)}$ and the series converges absolutely.

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

$$s_n := \sum_{|\nu|\le n} c_\nu P_\nu^{(\alpha,\beta)}$$

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

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 $\phi \in \mathcal{D}$

$$\lim_{n\to\infty} \langle S_n, \overline{\phi}\rangle = \lim_{n\to\infty} f\Big(\sum_{|\nu|\le n} \langle \phi, \tilde P_\nu^{(\alpha,\beta)}\rangle P_\nu^{(\alpha,\beta)}\Big) = f(\phi),$$

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

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

$$F(\phi) := \lim_{n\to\infty} \langle s_n, \overline{\phi}\rangle = \sum_{\nu\in\mathbb{N}_0^d} c_\nu \langle \phi, \tilde{P}_\nu^{(\alpha,\beta)}\rangle, \quad \phi \in \mathcal{D}, \tag{5.13}$$

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

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

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

Finally, $F(\tilde{P}_\nu^{(\alpha,\beta)}) = \lim_{n\to\infty}\langle s_n, \overline{\phi}\rangle = c_\nu$ is immediate by (5.13). □

To simplify our notation, we introduce the following "convolution": For functions $\Phi : [-1,1]^d \times [-1,1]^d \to \mathbb{C}$ and $f : [-1,1]^d \to \mathbb{C}$, 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 $\Phi : [-1,1]^d \times [-1,1]^d \to \mathbb{C}$ 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 (§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 $t = (t_1,\dots,t_{\ell-1},0,t_{\ell+1},\dots,t_d)$, $1 \le \ell \le d$, we have $D_\ell^s m(t) = 0$ for $s = 1,2,\dots,r$. Also, assume*

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

with $c > 0$ independent of t. Then the operator $T_m f := \sum_{\nu\in\mathbb{N}_0^d} m(\nu)\langle f, \tilde{P}_\nu^{(\alpha,\beta)}\rangle \tilde{P}_\nu^{(\alpha,\beta)}$ 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 $\hat{B} = |\hat{C}|^2 \ge 0$ 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

$$\Phi_j(x,y) := \sum_{\nu\in\mathbb{N}_0^d} \hat{B}\Big(\frac{\nu}{2^{j-1}}\Big) m(\nu) \tilde{P}_\nu^{(\alpha,\beta)}(x)\tilde{P}_\nu^{(\alpha,\beta)}(y), \quad j \ge 1.$$

Consider the kernels $K_N := \sum_{j=0}^N \Phi_j$. We shall prove that

$$\|K_N * f\|_p \le c\|f\|_p \quad \text{for } f \in L^p(w_{\alpha,\beta}) \tag{5.17}$$

with $c > 0$ a constant independent of f and N. As a consequence of this, it is easy to show that for any $f \in L^p(w_{\alpha,\beta})$ one has $T_m f = \lim_{N\to\infty} K_N f$ in $L^p(w_{\alpha,\beta})$ and $\|T_m f\|_p \le c\|f\|_p$ as claimed.

To prove (5.17) we shall employ the theory of generalized Caldeón-Zygmund operators. Note first that by Parseval's identity

$$\|K_N * f\|_2 \le c\|f\|_2 \quad \text{for } \ f \in L^2(w_{\alpha,\beta}). \tag{5.18}$$

Following Stein [21], p. 29, denote

$$V(x,y) := \inf\{\mu(B(y,\delta)) : x \in B(y,\delta)\} = \mu(B(y,\rho(x,y))), \tag{5.19}$$

where the last equality follows from the definition of $\rho(\cdot,\cdot)$. We shall show that

$$|K_N(x,y) - K_N(x,\bar y)| \le c\frac{\rho(y,\bar y)}{\rho(x,\bar y)}[V(x,\bar y)]^{-1} \tag{5.20}$$

whenever $\rho(x,\bar y) \ge 2\rho(y,\bar y)$. Then (5.17) will follow for $1 < p \le 2$ by (5.18) and (5.20) using the proposition on pp. 29-30 and Theorem 3 on p. 19 in [21]. After that a standard duality argument leads to estimate (5.17) in the case $2 < p < \infty$.

We now turn to the proof of (5.20). Fix $x, y, \bar y \in [-1,1]^d$, $x \ne \bar y$, and define

$$\gamma_i := \begin{cases} \alpha_i & \text{if} \quad 0 \le \bar y_i \le 1, \\ \beta_i & \text{if} \quad -1 \le \bar y_i < 0. \end{cases}$$

By (5.19) and (5.3) it follows that

$$V(x,\bar y) \sim \rho(x,\bar y)^d \prod_{i=1}^d \left(\sqrt{1-\bar y_i^2} + \rho(x,\bar y)\right)^{2\gamma_i+1}. \tag{5.21}$$

Let $\hat A_j(t) := \hat B(t) m(2^{j-1}t)$. We have $\operatorname{supp} \hat B \subset [0,2]^d \setminus \frac12 \mathcal{B}_1$ and by (5.16) it readily follows that $|D^\tau[m(2^{j-1}t)]| \le c$ for $t \in [0,2]^d \setminus \frac12\mathcal{B}_1$ and $|\tau| \le r$, where the constant $c > 0$ is independent of j. Therefore, $\|D^\tau \hat A_j\|_\infty = \|D^\tau[\hat B(\cdot) m(2^{j-1}\cdot)]\|_\infty \le c$ for $|\tau| \le r$ 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

$$|\Phi_j(x,y) - \Phi_j(x,\bar y)| \le \frac{c2^{j(d+1)}\rho(y,\bar y)}{W_{\alpha,\beta}(2^j,\bar y)(1+2^j\rho(x,\bar y))^\sigma} \tag{5.22}$$

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

If $\rho(y,\bar y) > 2^{-j}$ and $\rho(x,\bar y) \ge 2\rho(y,\bar y)$ (hence $\rho(x,y) \ge \rho(y,\bar y)$ and $\rho(x,\bar y) \le 2\rho(x,y) \le 3\rho(x,\bar y)$), then estimate (5.22) follows by Theorem 4.1 applied separately to $\Phi_j(x,y)$ and $\Phi_j(x,\bar y)$ and using (5.7). Therefore, (5.22) holds whenever $\rho(x,\bar y) \ge 2\rho(y,\bar y)$.

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

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

For F_1 we have using (5.22) and (5.21)

$$\begin{aligned} F_1 &\le \frac{c\rho(y,\bar y)}{\prod_{i=1}^d \left(\sqrt{1-\bar y_i^2}+2^{-j_1}\right)^{2\gamma_i+1}} \sum_{j=1}^{j_1} 2^{j(d+1)} \\ &\le \frac{c\rho(y,\bar y)2^{j_1(d+1)}}{\prod_{i=1}^d \left(\sqrt{1-\bar y_i^2}+2^{-j_1}\right)^{2\gamma_i+1}} \\ &\le \frac{c\rho(y,\bar y)}{\rho(x,\bar y)^{d+1}\prod_{i=1}^d \left(\sqrt{1-\bar y_i^2}+\rho(x,\bar y)\right)^{2\gamma_i+1}} \le c\frac{\rho(y,\bar y)}{\rho(x,\bar y)}[V(x,\bar y)]^{-1}. \end{aligned}$$

To estimate F_2 we first observe that $\left(\sqrt{1-\bar y_i^2}+2^{-j}\right)(1+2^{j-j_1}) \ge \sqrt{1-\bar y_i^2}+2^{-j_1}$. Then, using again (5.22) and (5.21), we get

$$\begin{aligned} F_2 &\le c\rho(y,\bar y) \sum_{j=j_1+1}^{N} \frac{2^{j(d+1)}}{\prod_{i=1}^d \left(\sqrt{1-\bar y_i^2}+2^{-j}\right)^{2\gamma_i+1}\left(1+2^{j-j_1}\right)^{\sigma}} \\ &\le c\rho(y,\bar y) \sum_{j=j_1+1}^{\infty} \frac{2^{j(d+1)}}{\prod_{i=1}^d \left(\sqrt{1-\bar y_i^2}+2^{-j_1}\right)^{2\gamma_i+1}\left(1+2^{j-j_1}\right)^{\sigma-2\sum_i\gamma_i-d}} \\ &\le \frac{c\rho(y,\bar y)2^{j_1(d+1)}}{\prod_{i=1}^d \left(\sqrt{1-\bar y_i^2}+2^{-j_1}\right)^{2\gamma_i+1}} \sum_{j=j_1+1}^{\infty} 2^{-(j-j_1)(\sigma-2\sum_i\gamma_i-2d-1)} \\ &\le \frac{c\rho(y,\bar y)}{\rho(x,\bar y)^{d+1}\prod_{i=1}^d \left(\sqrt{1-\bar y_i^2}+\rho(x,\bar y)\right)^{2\gamma_i+1}} \le c\frac{\rho(y,\bar y)}{\rho(x,\bar y)}[V(x,\bar y)]^{-1}, \end{aligned}$$

where we used that $\sigma > 2\sum_i \gamma_i + 2d + 1$. The above estimates of F_1 and F_2 yield (5.20). This completes the proof of the proposition. □

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 $j \ge 0$, 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 $0 < \theta_1 < \cdots < \theta_{2^{j+1}} < \pi$ and set

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

It is well known that uniformly (see [8])

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

As is well known [22] the zeros of the Jacobi polynomial $P_{2^{j+1}}^{(\alpha,\beta)}$ serve as knots of the Gaussian quadrature

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

which is exact for all algebraic polynomials that are of degree $2^{j+2}-1$. Furthermore, the coefficients c_ξ are all positive and satisfy (see e.g. [15])

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

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

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

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

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

For multi-indices $\alpha = (\alpha_1,\dots,\alpha_d)$, $\beta = (\beta_1,\dots,\beta_d)$ and $j \ge 0$, $1 \le i \le d$, 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

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

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

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

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

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

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

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

Evidently, $[-1,1]^d = \cup_{\xi\in\mathcal{X}_j} I_\xi$ 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

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

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

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

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

where $c_\sigma > 0$ depends only on σ, α, β, d, and $c_\star$.

The proof of this lemma is merely a repetition of the proof of the univariate result in [12, 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 $\Phi_0(x,y) = \Psi_0(x,y) := \tilde{P}_0(x)\tilde{P}_0(y)$,

$$\Phi_j(x,y) := \sum_{\nu\in\mathbb{N}_0^d} \hat{A}\Big(\frac{\nu}{2^{j-1}}\Big)\tilde{P}_\nu(x)\tilde{P}_\nu(y), \quad j \ge 1, \text{ and} \tag{6.11}$$

$$\Psi_j(x,y) := \sum_{\nu\in\mathbb{N}_0^d} \hat{B}\Big(\frac{\nu}{2^{j-1}}\Big)\tilde{P}_\nu(x)\tilde{P}_\nu(y), \quad j \ge 1. \tag{6.12}$$

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 jth level *needlets* by

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

We write $\mathcal{X} := \cup_{j=0}^\infty \mathcal{X}_j$, 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 Φ and Ψ 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,

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

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 $0 < p \le \infty$,

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

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

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

The needlet decomposition of $\mathcal{D}'$ and L^p follows as in the univariate case (see [12, Proposition 3.1]) by the definition of needlets and their superb localization.

Proposition 6.2. *(i) For $f \in \mathcal{D}'$, we have*

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

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

(ii) If $f \in L^p(w_{\alpha,\beta})$, $1 \le p \le \infty$, 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) Pick $\hat{A} \geq 0$ a cutoff function of type (c) (see Definition 3.1). Then we can choose $\hat{B} = \hat{A}$ in the constuction of needlets in (6.10)-(6.13) and 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. [12]) $\|f\|_2 = \Big(\sum_{\xi\in\mathcal{X}} |\langle f, \psi_\xi\rangle|^2\Big)^{1/2}$ for $f \in L^2(w_{\alpha,\beta})$, which shows that Ψ 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 super-exponential localization of the needlets, namely,

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

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

We next utilize the general idea of using spectral or orthogonal decompositions (see e.g. [16, 23]) 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 [12]. Therefore, we shall only state the main results, provide the important ingredients and refer the reader to [12] 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

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

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*

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

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 $\mathcal{X} := \cup_{j=0}^\infty \mathcal{X}_j$ 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*

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

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 Φ, Ψ are defined by

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

As in [12] one shows that the operator T_ψ is well defined on $f^{s\rho}_{pq}$, namely, for any $h \in f^{s\rho}_{pq}$, $T_\psi h := \sum_{\xi\in\mathcal{X}} h_\xi\psi_\xi$ converges in $\mathcal{D}'$. Moreover, the operator $T_\psi : f^{s\rho}_{pq} \to \mathcal{D}'$ is continuous, i.e. there exist constants $k > 0$ and $c > 0$ such that

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

Our main result in this section asserts that the weighet 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 $0 < q \le \infty$. The operators $S_\varphi : F^{s\rho}_{pq} \to f^{s\rho}_{pq}$ and $T_\psi : f^{s\rho}_{pq} \to F^{s\rho}_{pq}$ are bounded and $T_\psi \circ S_\varphi = Id$ on $F^{s\rho}_{pq}$. Consequently, $f \in F^{s\rho}_{pq}$ if and only if $\{\langle f, \varphi_\xi\rangle\}_{\xi\in\mathcal{X}} \in f^{s\rho}_{pq}$. Furthermore,*

(7.6)

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

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

To us the spaces F^{ss}_{pq} are more natural than the spaces $F^{s\rho}_{pq}$ with $\rho \ne s$ since they embed "correctly" with respect to the smoothness index s.

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

$$F^{ss}_{pq} \subset F^{s_1 s_1}_{p_1 q_1} \quad \text{if} \quad s/d - 1/p = s_1/d - 1/p_1. \tag{7.7}$$

The proof of this proposition is quite similar to the proof of the respective embedding result on B^d in [13, Proposition 4.11] and will be omitted.

We have the following identification of spaces F^{00}_{p2}.

Proposition 7.5. *We have*

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

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

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

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 [14, 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 [12].

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

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

where the ℓ^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 $B^{s\rho}_{pq}$ is a (quasi-)Banach space which is continuously embedded in $\mathcal{D}'$.

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

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

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

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

The analysis and synthesis operators S_φ and T_ψ defined in (7.4) play an important role here. As for weighted Triebel-Lizorkin spaces the operator T_ψ is well defined on $b^{s\rho}_{pq}$, i.e. for any $h\in b^{s\rho}_{pq}$, $T_\psi h:=\sum_{\xi\in\mathcal{X}}h_\xi\psi_\xi$ converges in $\mathcal{D}'$. Also, the operator $T_\psi: b^{s\rho}_{pq}\to\mathcal{D}'$ 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 $0<p,q\le\infty$. Then the operators $S_\varphi: B^{s\rho}_{pq}\to b^{s\rho}_{pq}$ and $T_\psi: b^{s\rho}_{pq}\to B^{s\rho}_{pq}$ are bounded and $T_\psi\circ S_\varphi=Id$ on $B^{s\rho}_{pq}$. Consequently, for $f\in\mathcal{D}'$ we have that $f\in B^{s\rho}_{pq}$ if and only if $\{\langle f,\varphi_\xi\rangle\}_{\xi\in\mathcal{X}}\in b^{s\rho}_{pq}$. Moreover,*

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

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

The parameter ρ in the definition of $B^{s\rho}_{pq}$ allows to consider various scales of weighted Besov spaces. The spaces B^{s0}_{pq} can be regarded as "classical" Besov spaces. However, to us more natural are the spaces B^{ss}_{pq} ($\rho=s$) which in contrast to B^{s0}_{pq}, 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^{ss}_{pq} (see §9 below).

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

$$B_{pq}^{ss} \subset B_{p_1 q_1}^{s_1 s_1} \quad \text{if} \quad s/d - 1/p = s_1/d - 1/p_1. \tag{8.3}$$

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 $\hat{B} = \hat{A}$, $\hat{A} \ge 0$, i.e. $\hat{A} \ge 0$ 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 Σ_n be the nonlinear set of all functions g of the form

$$g = \sum_{\xi\in\Lambda} a_\xi \psi_\xi,$$

where $\Lambda \subset \mathcal{X}$, $\#\Lambda \le n$, and Λ 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 Σ_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

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

Exactly as in [14, Proposition 6.1] this leads to the embedding of B_τ^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*

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

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 [14, 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

$$\|g\|_{B_\tau^s} \le c n^{s/d} \|g\|_p. \tag{9.3}$$

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 [18] and utilized in [13] to the development of Triebel-Lizorkin and Besov spaces on B^d with weight

$$w_\mu(x) := (1-\|x\|_2^2)^{\mu-1/2}, \quad \mu \ge 0.$$

Here, we compile all needed results from [13, 18] 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 $\mathrm{Proj}_n : L^2(B^d, w_\mu) \to V_n$, i.e.

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

An explicit representation of the reproducing kernel $P_n(w_\mu, x, y)$ is given in [24]:

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

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

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

Analogues of Theorems 4.1 and 4.2 on B^d are established in [18]. Denote

$$W_\mu(n; x) := \Big(\sqrt{1-\|x\|_2^2} + n^{-1}\Big)^{2\mu} \tag{10.4}$$

and

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

which is a distance on B^d.

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

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

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

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

This theorem was established in [18] (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 $\mathcal{Q}_n^{\alpha,\beta}$ 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 [18].

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

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

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

Proof. The proof of this lemma relies on the following estimate that follows from Theorem 3.1 in [2]: If a and b are constants such that $|a| + |b| \le 1$, then

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

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 $\|y\|_2 =: \cos\phi$, $0 \le \theta, \phi \le \pi$, we have

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

and hence

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

Here we used that $\left(1 + n\sqrt{1-|a|-|b|}\right)^{\lambda-\mu} \ge 1$. Now, from $A(x), A(y) \ge 0$ 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). □

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 $\rho(x, \xi) \le c^* n^{-1}$,*

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

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

The proof of this lemma is somewhat lengthy and will be given in the appendix.

10.2. **Localized cross product basis kernels.** We consider orthogonal polynomials on $B^{d_1} \times [-1,1]^{d_2}$ with weight

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

where $w_\mu(x') := (1 - \|x'\|_2^2)^{\mu-1/2}$, $\mu \ge 0$, and $w_{\alpha,\beta}(x'') := \prod_{j=1}^{d_2} w_{\alpha_j,\beta_j}(x_j'')$ with $\alpha_j, \beta_j \ge -1/2$ 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 $\mathrm{Proj}_n : L^2(B^{d_1}, w_\mu) \to V_n$, see (10.1)-(10.2).

We are interested in kernels of the form

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

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

$$\rho_*(x,y) := \max\Big\{\rho(x',y'), \max_{1\le j\le d_2} |\arccos x_j'' - \arccos y_j''|\Big\},$$

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*

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

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*

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

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 [10, 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 $\rho(x,\xi)\le c_*n^{-1}$, $n\ge 1$, $c_*>0$, the kernel Λ_n from* (10.12) *satisfies*

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

where $c>0$ depends only on σ,d,α,β, 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 [18, §5] and the cubature on $[-1,1]^{d_2}$ from §6.1. Once the components are in place, the construction is carried out exactly as in §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 μ. Just as in the case of tensor product Jacobi polynomials (§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

$$\Phi_j(x,y):=\sum_{(m,\nu)\in\mathbb{N}_0\times\mathbb{N}_0^{d_2}}\hat{A}\Big(\frac{m}{2^{j-1}},\frac{\nu}{2^{j-1}}\Big)P_m(w_\mu;x',y')\tilde{P}_\nu^{(\alpha,\beta)}(x'')\tilde{P}_\nu^{(\alpha,\beta)}(y''),\ j\ge1.$$

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 $0<q\le\infty$, is defined as the set of all $f\in\mathcal{D}'$ such that

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

with the usual modification when $q=\infty$. Here Φ_j*f is defined as in (5.15).

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

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

where the ℓ^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 §§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 §§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 belive 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 [14], ball [13], simplex [10], and in the context of tensor product Hermite [19] and Laguerre functions [11]. 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^∞ univariate cutoff function $\hat{a}$ satisfies

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

These are: (1) Theorem 4.2 in [18] on the ball, (2) Theorem 2.2 in [14] on the sphere, (3) Theorem 7.1 in [10] on the simplex, (4) Corollary 1 in [19] for tensor product Hermite functions, (5) Theorems 3.2, 3.7, 3.8 in [11] 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 §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.

12. Appendix: Proof of Lemma 10.3.

For $\mu = 0$ the expression of $P_n(w_\mu; x, y)$ in (10.2) simplifies considerably as $\mu \to 0$; the integral becomes a sum of two terms, as shown in [24]. 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 [2].

Lemma 12.1. *Suppose $\mu > 0$, $0 < |a| < 1$, $\eta \in C^\infty[-1,1]$ with $\operatorname{supp}\eta \subset [-\frac{1}{2}, 1]$. If $|b| \le 1 - |a|$, then*

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

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

Case 1: $A(x)A(y) \le 16c^*(n^{-1}\rho(x,y) + n^{-2})$, 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):

$$|C_n^\lambda(t)| \le cn^{2\lambda-1}\left(1+n\sqrt{1-t^2}\right)^{-\lambda}, \quad t\in[-1,1]. \tag{12.1}$$

Denote by I_u the interval with end points $t(x,y;u)$ and $t(\xi,y;u)$. Then using the identity $\frac{d}{dt}C_n^\lambda(t) = 2\lambda C_{n-1}^{\lambda+1}(t)$ [22, (4.7.27)], we obtain

$$\begin{aligned} E :&= |P_n(w_\mu;x,y) - P_n(w_\mu;\xi,y)| \\ &\le cn\int_{-1}^1 \left|C_n^\lambda(t(x,y;u)) - C_n^\lambda(t(\xi,y;u))\right|(1-u^2)^{\mu-1}du \qquad (12.2)\\ &\le cn\int_{-1}^1 \|C_{n-1}^{\lambda+1}\|_{L^\infty(I_u)}|t(x,y;u)-t(\xi,y;u)|(1-u^2)^{\mu-1}du. \end{aligned}$$

By (12.1) it follows that

$$\begin{aligned} \|C_{n-1}^{\lambda+1}\|_{L^\infty(I_u)} \le c\,n^{2\lambda+1}\Bigg[&\left(1+n\sqrt{1-t(x,y;u)^2}\right)^{-\lambda-1} \\ &+\left(1+n\sqrt{1-t(\xi,y;u)^2}\right)^{-\lambda-1}\Bigg]. \end{aligned} \tag{12.3}$$

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

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

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

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

The above estimates along with (12.3) and $\rho(x,\xi)\le c^*n^{-1}$ yield

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

We use these inequality in (12.2) to obtain

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

To estimate E_1 and E_2 we shall need the inequality (see [18, Lemma 4.1])

$$|A(x)-A(y)| \le \sqrt{2}\rho(x,y), \quad x,y\in B^d, \tag{12.4}$$

which implies

$$\begin{aligned} A(y)+n^{-1} &\le A(x)+n^{-1}+\sqrt{2}\rho(x,y) \\ &\le \sqrt{2}(A(x)+n^{-1})(1+n\rho(x,y)). \end{aligned} \tag{12.5}$$

On the other hand, by (10.10), (12.4), and our assumption it follows that

$$\begin{aligned} (A(x)+n^{-1})(A(y)+n^{-1}) &\le c\big(A(x)A(y)+n^{-1}|A(x)-A(y)|+n^{-2}\big) \\ &\le cn^{-2}(1+n\rho(x,y)). \end{aligned} \tag{12.6}$$

This along with (12.5) gives

$$A(y)^2 \le \sqrt{2}(A(x)+n^{-1})(A(y)+n^{-1})(1+n\rho(x,y)) \le cn^{-2}(1+n\rho(x,y))^2. \tag{12.7}$$

As in [18, p. 136] we have using (12.4)

$$\begin{aligned} |t(x,y;u)-t(\xi,y;u)| &\le |\cos\rho(x,y)-\cos\rho(\xi,y)| + \sqrt{2}|1-u|A(y)\rho(x,\xi) \\ &\le \rho(x,\xi)\big(\rho(x,y)+\rho(\xi,y)\big) + \sqrt{2}A(y)\rho(x,\xi). \end{aligned} \tag{12.8}$$

Combining this with (12.7) and $\rho(x,\xi)\le cn^{-1}$ we get

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

This estimate coupled with (12.6) leads to

$$E_1 \le cn^d\rho(x,\xi)\frac{n^{2\mu}}{(1+n\rho(x,y))^\mu} \le \frac{cn^d\rho(x,\xi)}{(A(x)+n^{-1})^\mu(A(y)+n^{-1})^\mu}. \tag{12.9}$$

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 (12.9). 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 (12.4) it readily follows that

$$A(x)A(y) \sim (A(x)+n^{-1})(A(x)+n^{-1}). \tag{12.10}$$

Let η_+ be a C^∞ function such that $\eta_+(u)=1$ for $\frac12\le u\le 1$, and $\eta_+(u)=0$ for $-1\le u\le -\frac12$. 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 $C_n^\lambda(-t) = (-1)^nC_n^\lambda(t)$, we only need to prove (10.11) for $P_n^+(w_\mu;\cdot,\cdot)$. We write $t(x,y;u)$ as

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

In going further, we have

$$P_n^+(w_\mu;x,y) - P_n^+(w_\mu;\xi,y) = J_1 + J_2, \tag{12.11}$$

where

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

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

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

We estimate the inner integral above using Lemma 12.1 with $\eta(t) = \eta_+(t)(1+t)^{\mu-1}$, $b = s - A(x)A(y)$ and $a = A(x)A(y)$. We get

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

As in (12.8)

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

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

$$|J_1| \le c\frac{n^{d+1}}{[A(x)A(y)]^\mu}\frac{(\rho(x,y)+n^{-1})\rho(x,\xi)}{(1+n\min\{\rho(x,y),\rho(\xi,y)\})^{\lambda+1-\mu}} \le \frac{cn^d\rho(x,\xi)}{[A(x)A(y)]^\mu}, \tag{12.12}$$

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

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

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

We estimate the inner integral by using Lemma 12.1 with $\eta(t) = \eta_+(t)(1+t)^{\mu-1}$, $b = B(\xi,y) - sA(y)$, $a = sA(y)$, and λ, μ replaced by $\lambda+1$, $\mu+1$ to obtain

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

and using that $|A(x)-A(\xi)| \le \sqrt{2}\rho(x,\xi)$ (see (12.4))

$$|J_2| \le \frac{cn^{d-1}\rho(x,\xi)}{A(y)^\mu\min\{A(x)^{\mu+1}, A(\xi)^{\mu+1}\}}. \tag{12.13}$$

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

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

If $A(x) \ge 2\sqrt{2}c^*n^{-1}$, then from above $A(\xi) \ge A(x)/2$. These two estimates and (12.13) imply that $|J_2|$ has the bound of $|J_1|$ from (12.12), and using (12.10) 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 $A(y) \ge 4\sqrt{2}c^*n^{-1}$. Then $A(y) \ge A(x)/2$ and using (12.4), we get

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

and hence $A(x)A(y) > 8c^*n^{-1}\rho(x,y) \ge 2\sqrt{2}c^*n^{-1}A(y)$ 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 complete the proof of Lemma 10.3. $\square$

References

[1] G. Brown, F. Dai, Approximation of smooth functions on compact two-point homogeneous spaces, J. Funct. Anal. 220 (2005), 401–423.

[2] F. Dai and Yuan Xu, Cesàro means of orthogonal expansions in several variables, Const. Approx. 29 (2009), 129-155.

[3] Z. Ditzian, V. Totik, Moduli of Smoothness, New York, Springer Verlag, 1987.

[4] T. Erdelyi, A. Magnus, and P. Nevai, Generalized Jacobi weights, Christoffel functions, and Jacobi polynomials, SIAM J. Math. Anal. 25 (1994), 602–614.

[5] M. Frazier, B. Jawerth, Decomposition of Besov Spaces, Indiana Univ. Math. J. 34 (1985), 777–799.

[6] M. Frazier, B. Jawerth, A discrete transform and decompositions of distribution, J. of Funct. Anal. 93 (1990), 34–170.

[7] M. Frazier, B. Jawerth, and G. Weiss, Littlewood-Paley Theory and the Study of Function Spaces, CBMS 79 (1991), AMS.

[8] C. L. Frenzen and R. Wong, A uniform asymptotic expansion of the Jacobi polynomials with error bounds, Canad. J. Math. 37 (1985), 979-1007.

[9] L. Hörmander, The analysis of linear partial differential operators I, Springer-Verlag, 1983.

[10] K. Ivanov, P. Petrushev, Yuan Xu, Sub-exponentially localized kernels and frames induced by orthogonal expansions, Math. Z. (to appear).

[11] G. Kerkyacharian, P. Petrushev, D. Picard and Yuan Xu, Decomposition of Triebel-Lizorkin and Besov spaces in the context of Laguerre expansions, J. Funct. Anal. 256 (2009), 1137–1188.

[12] G. Kyriazis, P. Petrushev and Yuan Xu, Jacobi decomposition of weighted Triebel-Lizorkin and Besov spaces, Studia Math. 186 (2008), 161–202.

[13] G. Kyriazis, P. Petrushev and Yuan Xu, Decomposition of weighted Triebel-Lizorkin and Besov spaces on the ball, Proc. London Math. Soc. 97 (2008), 477–513.

[14] F. J. Narcowich, P. Petrushev and J. D. Ward, Decomposition of Besov and Triebel-Lizorkin spaces on the sphere, J. Funct. Anal. 238 (2006), 530–564.

[15] P. Nevai, Orthogonal Polynomials, Memoirs of AMS, Vol. 18, 1979.

[16] J. Peetre, New thoughts on Besov spaces, Duke Univ. Math. Series, Durham, N.C. 1976.

[17] P. Petrushev, Yuan Xu, Localized polynomial frames on the interval with Jacobi weights, J. Fourier Anal. and Appl. 11 (2005), 557–575.

[18] P. Petrushev, Yuan Xu, Localized polynomial frames on the ball, Constr. Approx. 27 (2008), 121–148.

[19] P. Petrushev, Yuan Xu, Decomposition of spaces of distributions induced by Hermite expansion, J. Fourier Anal. and Appl. 14 (2008), 372-414.

[20] J. Riordan, Combinatorial identities, John Wiley, New York, 1968.

[21] E. Stein, Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals, Princeton University Press, Princeton, NJ, 1993.

[22] G. Szegő, Orthogonal Polynomials, Amer. Math. Soc. Colloq. Publ. Vol.23, Providence, 4th edition, 1975.

[23] H. Triebel, Theory of function spaces, Monographs in Math. Vol. 78, Birkhäuser, Verlag, Basel, 1983.

[24] 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.

Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria

E-mail address: ivanov@math.sc.edu

Department of Mathematics, University of South Carolina, Columbia, SC 29208, and Institute of Mathematics and Informatics, Bulgarian Academy of Sciences

E-mail address: pencho@math.sc.edu

Department of Mathematics, University of Oregon, Eugene, Oregon 97403-1222.

E-mail address: yuan@math.uoregon.edu