Constr. Approx. (2008) 27: 121–148
DOI: 10.1007/s00365-007-0678-9

CONSTRUCTIVE APPROXIMATION

Localized Polynomial Frames on the Ball

Pencho Petrushev and Yuan Xu

Abstract. Almost exponentially localized polynomial kernels are constructed on the unit ball B^d in $\mathbf{R}^d$ with weights $w_\mu(x) = (1-|x|^2)^{\mu-1/2}$, $\mu \geq 0$, by smoothing out the coefficients of the corresponding orthogonal projectors. These kernels are utilized to the design of cubature formulas on B^d with respect to $w_\mu(x)$ and to the construction of polynomial tight frames in $L^2(B^d, w_\mu)$ (called needlets) whose elements have nearly exponential localization.

1. Introduction

The construction of bases and frames on various domains, in particular on $\mathbf{R}^d$ and on the d-dimensional cube, sphere, and ball, is important from many prospectives and has numerous applications. The example of Meyers' wavelets [10] and the φ- transform of Frazier and Jawerth (see [6]) clearly shows the advantage of using localized bases or frames for decomposition of function and distribution spaces on $\mathbf{R}^d$ in contrast to other means such as atomic decompositions or Fourier series (in the periodic case). Three of their features: (i) infinite smoothness; (ii) almost exponential space localization; and (iii) infinitely vanishing moments, make them a universal tool for decomposing most of the classical spaces on $\mathbf{R}^d$, including Besov and Triebel–Lizorkin spaces. The key to this is that the coefficients in the wavelet or φ-transform expansions precisely capture the information in the norms defining the corresponding spaces.

Our primary goal in this paper is to develop a similar tool for decomposition of weighted spaces of functions or distributions on the unit ball B^d in $\mathbf{R}^d$ ($d > 1$) with weights

$$w_\mu(x) := (1-|x|^2)^{\mu-1/2}, \qquad \mu \geq 0, \tag{1.1}$$

where $|x|$ is the Euclidean norm of $x \in \mathbf{R}^d$. The situation here, however, is much more complicated than on $\mathbf{R}^d$ (the shift invariant case) or on the torus or even on the sphere due to several reasons: (i) there are no dilation or translation operators on B^d; (ii) the boundary of B^d in combination with the weight $w_\mu(x)$ creates a great deal of inhomogeneity; (iii) orthogonal systems such as orthogonal polynomials on B^d are much

Date received: February 8, 2006. Dates revised: October 10, 2006; January 9, 2007; March 9, 2007. Date accepted: March 13, 2007. Communicated by Edward B. Saff. Online publication: July 5, 2007.
AMS classification: 42C40, 42B35.
Key words and phrases: Localized kernels, Frames, Needlets, Unit ball.

less friendly than the trigonometric system; and (iv) there are no uniformly distributed points on B^d or on the d-dimensional unit sphere S^d.

Our approach to the problem at hand will heavily rely on orthogonal polynomials in the weighted spaces $L^2(B^d, w_\mu)$. The standard Hilbert space theory gives the orthogonal decomposition

$$L^2(B^d, w_\mu) = \bigoplus_{\nu=0}^{\infty} \mathcal{V}_n^d, \qquad \mathcal{V}_n^d \subset \Pi_n^d, \tag{1.2}$$

where $\mathcal{V}_n^d$ is the subspace of all polynomials of degree n which are orthogonal to lower degree polynomials in $L^2(B^d, w_\mu)$. Note that $\dim \mathcal{V}_n^d = \binom{n+d-1}{n} \sim n^{d-1}$, so $\mathcal{V}_n^d$ is a large subspace of L^2. The orthogonal projector $\mathrm{Proj}_n : L^2(B^d, w_\mu) \mapsto \mathcal{V}_n^d$ can be written as

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

where $P_n(w_\mu; x, y)$ is its kernel. It is crucial for our further development that the kernels $P_n(w_\mu; x, y)$ have an explicit representation [21] in terms of Gegenbauer polynomials (see (4.1)–(4.2) below). Now,

$$K_n(w_\mu; x, y) := \sum_{\nu=0}^{n} P_\nu(w_\mu; x, y) \tag{1.3}$$

is the kernel of the orthogonal projector of $L^2(B^d, w_\mu)$ onto $\bigoplus_{\nu=0}^{n} \mathcal{V}_\nu^d$.

Consider the kernel

$$L_n^\mu(x, y) = \sum_{j=0}^{\infty} \widehat{a}\Big(\frac{j}{n}\Big) P_j(w_\mu; x, y), \tag{1.4}$$

obtained by smoothing out the coefficients in the definition of the kernel $K_n(w_\mu; x, y)$ in (1.3) by sampling a smooth function $\widehat{a}$. One of our main results in this paper essentially asserts that if $\widehat{a} \in C^\infty[0, \infty)$ is compactly supported, then $L_n^\mu(x, y)$ has almost exponential (faster than any polynomial) rate of decay away from the main diagonal $y = x$ in $B^d \times B^d$. To state this result more precisely, let us introduce the distance (see (4.7))

$$d(x, y) := \arccos\{\langle x, y\rangle + \sqrt{1-|x|^2}\sqrt{1-|y|^2}\} \quad \text{on } B^d \tag{1.5}$$

and set

$$\mathcal{W}_\mu(n; x) := (\sqrt{1-|x|^2} + n^{-1})^{2\mu}, \qquad x \in B^d.$$

Then (see Section 4) for any $k > 0$ there exists a constant $c_k > 0$ depending only on k, d, μ, and $\widehat{a}$ such that

$$|L_n^\mu(x, y)| \le c_k \frac{n^d}{\sqrt{\mathcal{W}_\mu(n; x)}\sqrt{\mathcal{W}_\mu(n; y)}(1 + n\, d(x, y))^k}. \tag{1.6}$$

The localized kernels L_n^μ provide a powerful tool for constructing cubature formulas on B^d with weights $w_\mu(x)$, $\mu \ge 0$, that are exact for all polynomials of degree n and have positive coefficients of the right size. It is an important feature of our cubature formulas (see Section 5) that for all $\mu \geq 0$ the knots are obtained by projecting onto B^d sets of "almost equally" distributed points on the upper hemisphere S^d_+ in $\mathbf{R}^{d+1}$; the knots are in fact almost equally distributed on B^d with respect to the distance $d(\cdot, \cdot)$ defined in (1.5). Currently, very few families of cubature formulas with positive weights are known on B^d, among them is the family of the product-type formulas [18], [14]. However, the knots in these formulas are not almost equally distributed.

Most importantly, the kernels L_n^μ enable us to construct localized polynomial frames in $L^2(B^d, w_\mu)$ which is our primary goal in this paper. Our construction is based on a semidiscrete Calderón-type decomposition combined with our cubature formulas on the ball from Section 5. If we denote by $\Psi = \{\psi_\xi\}_{\xi \in \mathcal{X}}$ our frame on B^d, where $\mathcal{X} = \bigcup_{j=0}^\infty \mathcal{X}_j$ is an index set consisting of the localization points (poles) of the frame elements, then we have the following representation of each $f \in L^2(B^d, w_\mu)$:

$$f = \sum_{\xi \in \mathcal{X}} \langle f, \psi_\xi \rangle \psi_\xi \qquad \text{and} \qquad \|f\|_{L^2(B^d, w_\mu)} = \left(\sum_{\xi \in \mathcal{X}} |\langle f, \psi_\xi \rangle|^2 \right)^{1/2}.$$

The above clearly indicates that Ψ is a tight frame for $L^2(B^d, w_\mu)$. The most important feature of the frame elements ψ_ξ is their almost exponential localization: For $\xi \in \mathcal{X}_j$ (the jth level in $\mathcal{X}$),

$$(1.7) \qquad |\psi_\xi(x)| \leq c_k \frac{2^{jd/2}}{\sqrt{\mathcal{W}_\mu(2^j; x)}(1 + 2^j d(x, y))^k}, \quad \forall\, k > 0.$$

Here the presence of the factor $\sqrt{\mathcal{W}_\mu(2^j; x)}$ is critical; it reflects the expected influence of the boundary of B^d and the weight $w_\mu(x)$ on the localization of the frame elements. Notice that the distance $d(\cdot, \cdot)$ is also affected by the boundary of B^d. This localization of the ψ_ξ's is the reason for calling them **needlets**.

The superb localization of the needlets along with their semiorthogonal structure and increasing (with the levels) number of vanishing moments enables one to utilize them for decomposition of spaces of functions or distributions on B^d other than $L^2(B^d, w_\mu)$ such as $L^p(B^d, w_\mu)$ $(1 < p < \infty)$ and the more general weighted Triebel–Lizorkin and Besov spaces on B^d. This paper is Phase 1 of a bigger project. In [9] we use the results from this paper for characterization of the weighted Triebel–Lizorkin and Besov spaces on B^d. Consequently, some of the results here get beyond the immediate needs of this paper.

These ideas were first used in [15] for the construction of frames on the unit sphere S^d in $\mathbf{R}^{d+1}$. In [16] the spherical frames were utilized for decomposition of Besov and Triebel–Lizorkin spaces on the sphere. Further, this scheme has been applied in [17] for the development of frames on $[-1, 1]$ with Jacobi weights and then used in [8] for decomposition of weighted Besov and Triebel–Lizorkin spaces on the interval.

This paper is organized as follows. In Section 2 we outline the general principles which guide us in constructing localized kernels and frames on domains other than $\mathbf{R}^d$. In Section 3 we present some results on localized polynomial kernels on $[-1, 1]$ with Jacobi weights. In Section 4 we prove our main results on localized polynomial kernels on B^d with weights $w_\mu(x)$, $\mu \geq 0$. In Section 5 we construct cubature formulas on B^d

with weights $w_\mu(x)$. In Section 6 we construct our needlet system and give some of its properties.

Throughout this paper positive constants are denoted by $c, c_1, \ldots$ and they may vary at every occurrence. As usual the constants may depend on some parameters, which are indicated explicitly in some important cases. The notation $A \sim B$ means $c_1 A \le B \le c_2 A$. The set of all algebraic polynomials of total degree n in d variables is denoted by Π_n^d.

2. General Principles for Constructing Localized Kernels and Frames

Let (E, μ) be a measure space with E a metric space and suppose that there is an orthogonal decomposition of $L^2(E, \mu)$,

$$L^2(E, \mu) = \bigoplus_{n=0}^{\infty} \mathcal{V}_n, \tag{2.1}$$

where $\mathcal{V}_n$ is a subspace of dimension $\dim \mathcal{V}_n \sim n^\gamma$, $\gamma > 0$. Let P_n be the kernel of the orthogonal projector $\mathrm{Proj}_n : L^2(E, \mu) \to \mathcal{V}_n$, i.e.,

$$(\mathrm{Proj}_n f)(x) = \int_E P_n(x, y) f(y)\, d\mu, \qquad f \in L^2(E, \mu).$$

Notice that P_n can be written in the form $P_n(x, y) = \sum_{j=1}^{\dim \mathcal{V}_n} p_j(x)\overline{p_j(y)}$, where $\{p_j\}$ is an orthonormal basis for $\mathcal{V}_n$. Then $K_n := \sum_{j=0}^{n} P_\nu$ is the kernel of the orthogonal projector onto $\bigoplus_{\nu=0}^{n} \mathcal{V}_\nu$. In most cases of interest the kernel $K_n(x, y)$ has poor localization, examples include the trigonometric system, orthogonal polynomials in one or several variables on various domains.

Localization principle. Consider now the kernel

$$L_n(x, y) := \sum_{j=0}^{\infty} \widehat{a}\left(\frac{j}{n}\right) P_j(x, y), \tag{2.2}$$

where $\widehat{a} \in C^\infty(\mathbf{R})$, $\widehat{a}$ is even, and $\widehat{a}$ is compactly supported or $\widehat{a} \in \mathcal{S}$ (the Schwartz class of rapidly decreasing C^∞ functions on $\mathbf{R}$). It seems that there is a general localization principle, which says that for all "natural" orthogonal systems, the kernel $L_n(x, y)$ decays away from the main diagonal $y = x$ at nearly exponential (faster than any polynomial) rate with respect to the distance in E.

In the case of the trigonometric system this principle is well known and widely used. It is a simple but fundamental fact in Harmonic Analysis that the Fourier transform of every function f in the Schwartz space $\mathcal{S}$ belongs to the same space. As a consequence, one standardly shows that any trigonometric polynomial of the form

$$L_n(t) := \sum_{\nu \in \mathbf{Z}} \widehat{a}\left(\frac{\nu}{n}\right) e^{i\nu t},$$

where $\widehat{a}$ is a compactly supported C^∞ function, has nearly exponential rate of decay away from zero. More precisely, for any $k > 0$ and $r \geq 0$, there exists a constant $c_k > 0$ depending only on k, r, and $\widehat{a}$ such that

$$|L_n^{(r)}(t)| \leq c_k \frac{n^{r+1}}{(1+n|t|)^k}, \qquad t \in [-\pi, \pi]. \tag{2.3}$$

This estimate will serve as a prototype for our further localization results.

For Gegenbauer polynomials and spherical harmonics the localization principle is established and used in [15] and also follows by the general result in [7] on the spectral properties of elliptic operators. For Jacobi polynomials it is proved in [17] and [3], and can be extracted from [11] (see Theorem 3.1 below). For Hermite and Laguerre polynomials the localization principle is established in [5]. We will establish it here for multivariate orthogonal polynomials in $L^2(B^d, w_\mu)$ (see Theorem 4.2). We believe that the localization principle is valid in other and in more general settings as well.

For our purposes we restrict our attention to "smoothing functions" $\widehat{a}$ satisfying:

Definition 2.1. A function $\widehat{a}$ is said to be admissible if $\widehat{a} \in C^\infty[0,\infty)$, $\widehat{a}(t) \geq 0$, and $\widehat{a}$ satisfies one of the following two conditions:

(a) $\operatorname{supp}\widehat{a} \subset [0,2]$, $\widehat{a}(t) = 1$ on $[0,1]$, and $0 \leq \widehat{a}(t) \leq 1$ on $[1,2]$; or
(b) $\operatorname{supp}\widehat{a} \subset [\frac{1}{2}, 2]$.

There are two important applications of the localized kernels $L_n(x,y)$ from (2.2):

(i) If $\widehat{a}$ is admissible of type (a), then the operator

$$(\mathcal{L}_n f)(x) := \int_E L_n(x,y) f(y)\, d\mu(y)$$

apparently satisfies $\mathcal{L}_n f = f$ for all $f \in \bigoplus_{\nu=0}^n \mathcal{V}_\nu$ and $\mathcal{L}_n f \in \bigoplus_{\nu=0}^{2n} \mathcal{V}_\nu$. These along with the superb localization of L_n (to be established) make $\mathcal{L}_n$ a useful tool. We will see this operator at work in the construction of cubature formulas on the ball in Section 5.

(ii) More importantly, kernels $L_n(x,y)$ with $\widehat{a}$ admissible of type (b) are a valuable tool for constructing localized frames. Let, in addition, $\widehat{a}$ satisfy the conditions $\widehat{a}(t) \geq 0$ and

$$\widehat{a}^2(t) + \widehat{a}^2(2t) = 1, \qquad t \in [\tfrac{1}{2}, 1]. \tag{2.4}$$

Then

$$\sum_{\nu=0}^\infty \widehat{a}^2(2^{-\nu} t) = 1, \qquad t \in [1,\infty). \tag{2.5}$$

Define

$$L_0(x,y) := P_0(x,y) \tag{2.6}$$

and

$$L_j(x,y) := \sum_{\nu=0}^\infty \widehat{a}\left(\frac{\nu}{2^{j-1}}\right) P_\nu(x,y), \qquad j = 1, 2, \ldots,$$

and denote briefly

$$(L_j * f)(x) := \int_E L_j(x, y) f(y)\, d\mu(y).$$

One easily obtains the following semidiscrete Calderón-type decomposition (see, e.g., [17])

$$f = \sum_{j=0}^{\infty} L_j * L_j * f \qquad \text{for} \quad f \in L_2(E, \mu). \tag{2.7}$$

To get a completely discretized decomposition of $L_2(E, \mu)$ one can use quadrature (cubature) formulas, if available. Assume that there is a quadrature formula

$$\int_E f\, d\mu \sim \sum_{\xi \in \mathcal{X}_j} \lambda_\xi f(\xi) \tag{2.8}$$

with $\mathcal{X}_j \subset E$ and $\lambda_\xi > 0$, which is exact for all functions f of the form $f = gh$ with $g, h \in \bigoplus_{\nu=0}^{2^j} \mathcal{V}_\nu$. Assume also that if $g \in \mathcal{V}_n$, then $\overline{g} \in \mathcal{V}_n$.

After these preparations we now define the frame elements by

$$\psi_\xi(x) := \sqrt{\lambda_\xi} \cdot L_j(x, \xi) \qquad \text{for} \quad \xi \in \mathcal{X}_j, \qquad j = 0, 1, \ldots. \tag{2.9}$$

The ψ's inherit the localization of the kernels L_j, which is almost exponential in all cases of interest. This is the reason for calling them needlets.

We write $\mathcal{X} := \bigcup_{j=0}^{\infty} \mathcal{X}_j$, where any two points $\xi, \omega \in \mathcal{X}$ (from levels $\mathcal{X}_j \neq \mathcal{X}_k$) are considered to be different elements of $\mathcal{X}$ even if they coincide. We use $\mathcal{X}$ as an index set in the definition of the needlet system

$$\Psi := \{\psi_\xi\}_{\xi \in \mathcal{X}}.$$

One easily shows that Ψ is a tight frame in $L^2(E, \mu)$ (see [15]): If $f \in L^2(E, \mu)$, then

$$f = \sum_{j=0}^{\infty} \sum_{\xi \in \mathcal{X}_j} \langle f, \psi_\xi \rangle \psi_\xi = \sum_{\xi \in \mathcal{X}} \langle f, \psi_\xi \rangle \psi_\xi \quad \text{in } L^2(E, \mu) \tag{2.10}$$

and

$$\|f\|_{L^2(E,\mu)} = \left(\sum_{\xi \in \mathcal{X}} |\langle f, \psi_\xi \rangle|^2 \right)^{1/2}. \tag{2.11}$$

This scheme for the construction of frames was first introduced in [15] and further utilized in [16], [17].

3. Localized Polynomial Kernels on $[-1, 1]$

The Jacobi polynomials $\{P_n^{(\alpha,\beta)}\}_{n=0}^{\infty}$ constitute an orthogonal basis for the weighted space $L^2([-1, 1], w_{\alpha,\beta})$ with $w_{\alpha,\beta}(t) := (1-t)^\alpha (1+t)^\beta$, $\alpha, \beta > -1$. It is well known that [19]

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

where

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

For $f \in L^2([-1,1], w_{\alpha,\beta})$ the Fourier expansion of f in Jacobi polynomials is

$$f(t) = \sum_{n=0}^{\infty} d_n(f)(h_n^{(\alpha,\beta)})^{-1} P_n^{(\alpha,\beta)}(t), \qquad d_n(f) = \int_{-1}^{1} f(t) P_n^{(\alpha,\beta)}(t) w_{\alpha,\beta}(t)\, dt.$$

The nth partial sum of this expansion can be written as

$$(S_n f)(x) = \sum_{j=0}^{n} d_j(f)(h_j^{(\alpha,\beta)})^{-1} P_j^{(\alpha,\beta)}(x) = \int_{-1}^{1} f(t) K_n^{(\alpha,\beta)}(x,t) w_{\alpha,\beta}(t)\, dt,$$

where the kernel is given by

$$K_n^{(\alpha,\beta)}(x,t) = \sum_{j=0}^{n} (h_j^{(\alpha,\beta)})^{-1} P_j^{(\alpha,\beta)}(x) P_j^{(\alpha,\beta)}(y). \tag{3.1}$$

The grand question here is: What is the localization around the main diagonal $y = x$ in $[-1,1]^2$ of a polynomial kernel of the form

$$L_n^{\alpha,\beta}(x,y) = \sum_{j=0}^{\infty} \widehat{a}\left(\frac{j}{n}\right) (h_j^{(\alpha,\beta)})^{-1} P_j^{(\alpha,\beta)}(x) P_j^{(\alpha,\beta)}(y), \tag{3.2}$$

where $\widehat{a} \in C^\infty$?

To address this question, denote

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

Theorem 3.1 [17]. *Let $\alpha, \beta > -\frac{1}{2}$ and let $\widehat{a}$ be admissible according to Definition* 2.1. *Then for every $k > 0$ there is a constant $c_k > 0$ depending only on k, α, β, and $\widehat{a}$ such that, for $0 \le \theta, \varphi \le \pi$,*

$$|L_n^{\alpha,\beta}(\cos\theta, \cos\varphi)| \le c_k \frac{n}{\sqrt{w_{\alpha,\beta}(n;\cos\theta)}\sqrt{w_{\alpha,\beta}(n;\cos\varphi)}(1 + n|\theta - \varphi|)^k}. \tag{3.3}$$

Here the dependence of c_k on $\widehat{a}$ is of the form $c_k = c(\alpha,\beta,k)\max_{0\le\nu\le k}\|\widehat{a}^{(\nu)}\|_{L^\infty}$.

For the proof of this theorem it is important to establish estimate (3.3) first in the particular case when $\varphi = 0$ (the localization of $L_n^{\alpha,\beta}(x,1)$). Set

$$Q_n^{\alpha,\beta}(x) := L_n^{\alpha,\beta}(x,1) = \sum_{j=0}^{\infty} \widehat{a}\left(\frac{j}{n}\right) (h_j^{(\alpha,\beta)})^{-1} P_j^{(\alpha,\beta)}(1) P_j^{(\alpha,\beta)}(x). \tag{3.4}$$

Since [19, (4.1.1), p. 58],

$$P_n^{(\alpha,\beta)}(1) = \binom{n+\alpha}{n} = \frac{\Gamma(n+\alpha+1)}{\Gamma(\alpha+1)\Gamma(n+1)},$$

it is easy to verify that

$$(3.5)\quad Q_n^{\alpha,\beta}(x) = c^\diamond \sum_{j=0}^{\infty} \widehat{a}\left(\frac{j}{n}\right) \frac{(2j+\alpha+\beta+1)\Gamma(j+\alpha+\beta+1)}{\Gamma(j+\beta+1)} P_j^{(\alpha,\beta)}(x),$$

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

Now the key role is played by the following theorem, which will also be critical for the proof of our main localization result (Theorem 4.2).

Theorem 3.2. *Let $\widehat{a}$ be admissible and assume that $\alpha \ge \beta > -\frac{1}{2}$. Then for every $k > 0$ and $r \ge 0$ there exists a constant $c_k > 0$ depending only on k, r, α, β, and $\widehat{a}$ such that*

$$(3.6)\qquad \left|\left(\frac{d}{dx}\right)^r Q_n^{\alpha,\beta}(\cos\theta)\right| \le c_k \frac{n^{2\alpha+2r+2}}{(1+n\theta)^k}, \qquad 0 \le \theta \le \pi.$$

The dependence of c_k on $\widehat{a}$ is of the form $c_k = c(\alpha,\beta,k,r)\max_{0\le\nu\le k}\|\widehat{a}^{(\nu)}\|_{L^\infty}$.

This theorem is proved in [3] with $\widehat{a}$ admissible of type (a) (but the proof in [3] is valid in general) and in [17] with $\widehat{a}$ admissible of type (b). Estimate (3.6), when $r = 0$, can also be extracted from [11, Lemma 4.10]. Estimate (3.6) was proved earlier in [15] in the case $\alpha = \beta = \lambda - \frac{1}{2}$ (with λ a half-integer) and utilized for the construction of frames on the n-dimensional sphere. Theorem 3.1 is established in [17]. Its proof rests on Theorem 3.2.

4. Localized Polynomial Kernels on the Unit Ball

It is known (see [21]) that the orthogonal projector $\mathrm{Proj}_n : L^2(B^d, w_\mu) \mapsto \mathcal{V}_n^d$ can be written as

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

where if $\mu > 0$ the kernel $P_n(w_\mu; x, y)$ has the following explicit representation:

$$(4.1)\quad P_n(w_\mu; x, y) = b_d^\mu b_1^{\mu-1/2} \frac{\lambda+n}{\lambda} \times \int_{-1}^{1} C_n^\lambda(\langle x, y\rangle + u\sqrt{1-|x|^2}\sqrt{1-|y|^2})(1-u^2)^{\mu-1}\, du.$$

Here $\langle x, y\rangle$ is the usual Euclidean inner product, the constants $b_d^\mu, b_1^{\mu-1/2}$ are defined by $(b_d^\gamma)^{-1} := \int_{B^d} w_\gamma(x)\, dx$, where $w_\gamma(x)$ is as in (1.1), C_n^λ is the nth-degree Gegenbauer polynomial, and

$$\lambda = \mu + \frac{d-1}{2}.$$

The case $\mu = 0$ is a limit case and we have

$$(4.2)\qquad P_n(w_0; x, y) = b_d^0 \frac{\lambda+n}{2\lambda}[C_n^\lambda(\langle x, y\rangle + \sqrt{1-|x|^2}\sqrt{1-|y|^2}) + C_n^\lambda(\langle x, y\rangle - \sqrt{1-|x|^2}\sqrt{1-|y|^2})].$$

For an admissible $\widehat{a}$ (according to Definition 2.1) we define

$$L_n^{\mu}(x, y) = \sum_{j=0}^{\infty} \widehat{a}\left(\frac{j}{n}\right) P_j(w_\mu; x, y), \qquad x, y \in B^d. \tag{4.3}$$

The explicit representation (4.1) gives

$$L_n^{\mu}(x, y) = b_d^{\mu} b_1^{\mu-1/2} \int_{-1}^{1} Q_n^{\lambda}(\langle x, y\rangle + u\sqrt{1-|x|^2}\sqrt{1-|y|^2})(1-u^2)^{\mu-1}\, du, \tag{4.4}$$

where Q_n^{λ} is defined by

$$Q_n^{\lambda}(x) := \sum_{j=0}^{\infty} \widehat{a}\left(\frac{j}{n}\right) \frac{j+\lambda}{\lambda} C_j^{\lambda}(x).$$

Since

$$C_n^{\lambda}(x) = \frac{\Gamma(\lambda + \frac{1}{2})}{\Gamma(2\lambda)} \frac{\Gamma(n+2\lambda)}{\Gamma(n+\lambda+\frac{1}{2})} P_n^{(\lambda-1/2, \lambda-1/2)}(x),$$

it readily follows from (3.5) that

$$Q_n^{\lambda-1/2,\lambda-1/2}(x) = c\, Q_n^{\lambda}(x), \qquad \text{where} \quad c = 2^{-2\lambda}\Gamma(2\lambda+1)\Gamma(\lambda+\tfrac{1}{2})^{-2}.$$

Then by Theorem 3.2 we get the following estimate: For all $k, \lambda > 0$ and $r \geq 0$ there exists a constant $c_k > 0$ depending only on k, r, λ, and $\widehat{a}$, such that

$$\left| \left(\frac{d}{dx}\right)^r Q_n^{\lambda}(\cos\theta) \right| \leq c_k \frac{n^{2\lambda+2r+1}}{(1+n\theta)^k}, \qquad 0 \leq \theta \leq \pi. \tag{4.5}$$

Distance on B^d. In order to show that L_n^{μ} is a well-localized kernel and for our further development, we need to introduce an appropriate distance in B^d that takes into account the fact that B^d has a boundary. In [20] it is shown that the orthogonal polynomials on the unit ball and those on the unit sphere are closely related by the simple map

$$x \in B^d \mapsto x' := (x, \sqrt{1-|x|^2}) \in S^d, \tag{4.6}$$

which "lifts" the points from B^d to the upper hemisphere S_+^d in $\mathbf{R}^{d+1}$, that is, $S_+^d := \{x \in S^d : x_{d+1} \geq 0\}$. This relation leads us to the following distance on B^d, which will play a vital role in the following:

$$d(x, y) := \arccos\{\langle x, y\rangle + \sqrt{1-|x|^2}\sqrt{1-|y|^2}\}. \tag{4.7}$$

In fact this is the geodesic distance between $x' := (x, \sqrt{1-|x|^2})$ and $y' := (y, \sqrt{1-|y|^2})$ on $S_+^d \subset \mathbf{R}^{d+1}$ and, consequently, it is a true distance on B^d. This distance has been used to prove various polynomial inequalities, see the discussions in [2] and the references therein.

The map (4.6) also leads to a close relation between the spaces $L^p(B^d, w_0)$ and $L^p(S^d, d\omega)$, where $d\omega$ is the surface measure on S^d. This allows us to derive results on

$L^p(B^d, w_0)$ from those on $L^p(S^d, d\omega)$, which are also easier to prove. For these reasons we will prove our results only in the case $\mu > 0$.

The following lemma provides an important relation between $d(\cdot, \cdot)$ and the Euclidean norm $|\cdot|$ in B^d.

Lemma 4.1. *For $x, y \in B^d$, we have*

$$||x| - |y|| \le \frac{1}{\sqrt{2}} d(x, y)(\sqrt{1 - |x|^2} + \sqrt{1 - |y|^2}) \tag{4.8}$$

and hence

$$|\sqrt{1 - |x|^2} - \sqrt{1 - |y|^2}| \le \sqrt{2}\, d(x, y). \tag{4.9}$$

Proof. Let $0 \le \alpha, \beta \le \pi/2$ be defined from $|x| = \cos\alpha$ and $|y| = \cos\beta$. Using spherical-polar coordinates $x = |x|\xi$ and $y = |y|\zeta$, where $\xi, \zeta \in S^{d-1}$, we see that

$$d(x, y) = \arccos(\cos\alpha\cos\beta\langle\xi, \zeta\rangle + \sin\alpha\sin\beta) \ge \arccos(\cos(\alpha - \beta))$$

which yields $d(x, y) \ge |\alpha - \beta|$. On the other hand, since $0 \le \alpha, \beta \le \pi/2$, we have $\cos(\alpha - \beta)/2 \ge \cos(\pi/4) = \sqrt{2}/2$ and, consequently,

$$\sin\alpha + \sin\beta = 2\sin\frac{\alpha + \beta}{2}\cos\frac{\alpha - \beta}{2} \ge \sqrt{2}\sin\frac{\alpha + \beta}{2}.$$

Using the above we obtain

$$\begin{aligned} ||x| - |y|| \ &= \ |\cos\alpha - \cos\beta| = 2\sin\frac{|\alpha - \beta|}{2}\sin\frac{\alpha + \beta}{2} \\ &\le \ \frac{1}{\sqrt{2}}|\alpha - \beta|(\sin\alpha + \sin\beta) \le \frac{1}{\sqrt{2}} d(x, y)(\sqrt{1 - |x|^2} + \sqrt{1 - |y|^2}). \end{aligned}$$

Thus (4.8) is established. Estimate (4.9) follows immediately from (4.8). ■

Let us define

$$\mathcal{W}_\mu(n; x) := (\sqrt{1 - |x|^2} + n^{-1})^{2\mu}, \qquad x \in B^d. \tag{4.10}$$

Our next theorem shows that the kernels L_n^μ are almost exponentially localized around the main diagonal $y = x$ in $B^d \times B^d$.

Theorem 4.2. *Let $\widehat{a}$ be admissible. Then for any $k > 0$ there exists a constant $c_k > 0$ depending only on k, d, μ, and $\widehat{a}$ such that*

$$|L_n^\mu(x, y)| \le c_k \frac{n^d}{\sqrt{\mathcal{W}_\mu(n; x)}\sqrt{\mathcal{W}_\mu(n; y)}(1 + n\, d(x, y))^k}, \qquad x, y \in B^d. \tag{4.11}$$

Remark 4.3. Theorem 4.2 as well as Theorems 3.1–3.2 can obviously be modified for the case when $\widehat{a} \in C^k$ (k sufficiently large) in place of $\widehat{a} \in C^\infty$.

We will derive Theorem 4.2 when $\mu > 0$ from estimate (4.5) and the following lemma, using representation (4.4) of L_n^μ. The proof in the case $\mu = 0$ is easier and will be omitted; it utilizes (4.2).

Let us denote briefly

$$t(x, y; u) := \langle x, y\rangle + u\sqrt{1-|x|^2}\sqrt{1-|y|^2}. \tag{4.12}$$

Lemma 4.4. *Let $\gamma > -1$, $k > 3\gamma + 4$, and $n \ge 1$. Then, for $x, y \in B^d$,*

$$\int_{-1}^{1} \frac{(1-u^2)^\gamma\, du}{(1+n\sqrt{1-t(x,y;u)})^k} \le \frac{cn^{-2\gamma-2}}{(\sqrt{1-|x|^2}+n^{-1})^{\gamma+1}(\sqrt{1-|y|^2}+n^{-1})^{\gamma+1}(1+n\,d(x,y))^{k-3\gamma-4}}, \tag{4.13}$$

where $c > 0$ depends only on γ, k, and d.

Proof. Denote briefly $t := t(x, y; u)$. Then we can write

$$1 - t = 1 - \langle x, y\rangle - \sqrt{1-|x|^2}\sqrt{1-|y|^2} + (1-u)\sqrt{1-|x|^2}\sqrt{1-|y|^2},$$

which implies

$$\begin{aligned} 1 - t \;&\ge\; 1 - \langle x, y\rangle - \sqrt{1-|x|^2}\sqrt{1-|y|^2} \\ &=\; 1 - \cos d(x, y) = 2\sin^2\frac{d(x,y)}{2} \ge \frac{2}{\pi^2} d(x,y)^2 \end{aligned} \tag{4.14}$$

and

$$\begin{aligned} 1 - t \;&\ge\; \frac{2}{\pi^2} d(x,y)^2 + (1-u)\sqrt{1-|x|^2}\sqrt{1-|y|^2} \\ &\ge\; (1-u)\sqrt{1-|x|^2}\sqrt{1-|y|^2}. \end{aligned} \tag{4.15}$$

By (4.14), we have

$$\int_{-1}^{1} \frac{(1-u^2)^\gamma\, du}{(1+n\sqrt{1-t})^k} \le \frac{c}{(1+n\,d(x,y))^k}. \tag{4.16}$$

Inequality (4.13) will follow from this and the estimate:

$$\int_{-1}^{1} \frac{(1-u^2)^\gamma\, du}{(1+n\sqrt{1-t})^k} \le \frac{cn^{-2\gamma-2}}{(\sqrt{1-|x|^2})^{\gamma+1}(\sqrt{1-|y|^2})^{\gamma+1}(1+n\,d(x,y))^{k-2\gamma-3}}. \tag{4.17}$$

To establish this last estimate, we split the integral over $[-1, 1]$ into two integrals: one over $[-1, 0]$ and the other over $[0, 1]$. For the integral over $[-1, 0]$ we write the factor

$(1+n\sqrt{1-t})^k$ as the product of $(1+n\sqrt{1-t})^{k-2\gamma-2}$ and $(1+n\sqrt{1-t})^{2\gamma+2}$. Then we apply inequalities (4.14) and (4.15) to the first and second terms, respectively. This gives

$$\begin{aligned}\int_{-1}^{0} &\le \frac{c}{(1+n\,d(x,y))^{k-2\gamma-2}}\int_{-1}^{0}\frac{(1-u^2)^{\gamma}}{[n^2\sqrt{1-|x|^2}\sqrt{1-|y|^2}(1-u)]^{\gamma+1}}\,du\\ &\le \frac{cn^{-2\gamma-2}}{(\sqrt{1-|x|^2})^{\gamma+1}(\sqrt{1-|y|^2})^{\gamma+1}(1+n\,d(x,y))^{k-2\gamma-2}}.\end{aligned}$$

We now estimate the integral over $[0,1]$. Denote briefly $A:=\sqrt{1-|x|^2}\sqrt{1-|y|^2}$. Using (4.15) and applying the substitution $s=An^2(1-u)$, we get

$$\begin{aligned}\int_{0}^{1} &\le c\int_0^1\frac{(1-u^2)^{\gamma}}{(1+n\sqrt{d(x,y)^2+A(1-u)})^k}\,du\\ &\le \frac{c}{(An^2)^{\gamma+1}}\int_0^{An^2}\frac{s^{\gamma}}{(1+\sqrt{n^2\,d(x,y)^2+s})^k}\,ds\\ &\le \frac{cn^{-2\gamma-2}}{A^{\gamma+1}(1+n\,d(x,y))^{k-2\gamma-3}}\int_0^{\infty}\frac{s^{\gamma}\,ds}{(1+\sqrt{n^2\,d(x,y)^2+s})^{2\gamma+3}}\\ &\le \frac{cn^{-2\gamma-2}}{A^{\gamma+1}(1+n\,d(x,y))^{k-2\gamma-3}}.\end{aligned}$$

Putting these estimates together gives (4.17).

To complete the proof of (4.13) we need the following simple inequality (see inequality (2.21) in [17]):

$$(4.18)\quad (a+n^{-1})(b+n^{-1})\le 3(ab+n^{-2})(1+n|a-b|),\qquad a,b\ge 0,\quad n\ge 1.$$

Inequalities (4.9) and (4.18) yield

$$\begin{aligned}&(\sqrt{1-|x|^2}+n^{-1})(\sqrt{1-|y|^2}+n^{-1})\\ &\qquad\le 3(\sqrt{1-|x|^2}\sqrt{1-|y|^2}+n^{-2})(1+n\,d(x,y)).\end{aligned}$$

This along with (4.16) and (4.17) implies (4.13). ∎

Proof of Theorem 4.2. For $t=\cos\theta$, $0\le\theta\le\pi$, we have $\theta/2\sim\sin\theta/2\sim\sqrt{1-t}$. Therefore, estimate (4.5) with $r=0$ is equivalent to

$$\left|Q_n^{\lambda}(t)\right|\le c_k\frac{n^{2\lambda+1}}{(1+n\sqrt{1-t})^k},\qquad -1\le t\le 1.$$

Now, (4.11) follows readily by Lemma 4.4. ∎

The estimate of $|L_n^{\mu}(x,y)|$ from Theorem 4.2 enables us to control its L^p-norm.

Proposition 4.5. *For* $0 < p \le \infty$, *we have*

$$\left(\int_{B^d} |L_n^\mu(x,y)|^p w_\mu(y)\,dy\right)^{1/p} \le c\left(\frac{n^d}{\mathcal{W}_\mu(n;x)}\right)^{1-1/p}, \qquad x \in B^d. \tag{4.19}$$

Proof. If $0 < p < \infty$ this proposition is an immediate consequence of Theorem 4.2 and Lemma 4.6 below. In the case $p = \infty$, estimate (4.19) follows by (4.11) and (4.9) (see also estimate (4.23) below). Note also that for $1 < p < \infty$, estimate (4.19) follows readily from the cases $p = 1$ and $p = \infty$. ∎

Lemma 4.6. *If* $\sigma > d/p + 2\mu|1/p - 1/2|$, $\mu \ge 0$, $0 < p < \infty$, *then*

$$J_p := \int_{B^d} \frac{w_\mu(y)\,dy}{\mathcal{W}_\mu(n;y)^{p/2}(1+n\,d(x,y))^{\sigma p}} \le cn^{-d}\mathcal{W}_\mu(n;x)^{1-p/2}, \tag{4.20}$$

where $c > 0$ *depends only on* p, μ, *and* d.

Proof. Let $\mu > 0$ (the case $\mu = 0$ is easier). Three cases present themselves here.

Case 1. $p = 2$. Using spherical-polar coordinates and the fact that

$$\int_{S^{d-1}} g(\langle x,y\rangle)\,d\omega(y) = \sigma_{d-2}\int_{-1}^{1} g(|x|t)(1-t^2)^{(d-3)/2}\,dt,$$

where σ_{d-2} is the surface area of S^{d-2}, it follows that

$$J_2 = c\int_0^1 \frac{r^{d-1}(1-r^2)^{\mu-1/2}}{(n^{-1}+\sqrt{1-r^2})^{2\mu}}\int_{-1}^{1}\frac{(1-s^2)^{(d-3)/2}\,ds}{(1+n\arccos(rs|x|+\sqrt{1-|x|^2}\sqrt{1-r^2}\,))^{2\sigma}}\,dr.$$

Write briefly $F(r,t) := 1/[1+n\arccos(t|x|+\sqrt{1-|x|^2}\sqrt{1-r^2})]^{2\sigma}$. Then

$$J_2 = c\int_0^1 \frac{r^{d-1}(1-r^2)^{\mu-1/2}}{(n^{-1}+\sqrt{1-r^2})^{2\mu}}\int_{-1}^{1} F(r,rs)(1-s^2)^{(d-3)/2}\,ds\,dr. \tag{4.21}$$

Next, we apply the substitution $u = rs$, then switch the order of integration and, finally, substitute $t = \sqrt{1-r^2}$. This gives

$$\begin{aligned}
J_2 &= c\int_0^1 \frac{r(1-r^2)^{\mu-1/2}}{(n^{-1}+\sqrt{1-r^2})^{2\mu}}\int_{-r}^{r} F(r,u)(r^2-u^2)^{(d-3)/2}\,du\,dr\\
&= c\int_{-1}^{1}\int_{|u|}^{1} F(r,u)\frac{r(1-r^2)^{\mu-1/2}}{(n^{-1}+\sqrt{1-r^2})^{2\mu}}(r^2-u^2)^{(d-3)/2}\,dr\,du\\
&= c\int_{-1}^{1}\int_0^{\sqrt{1-u^2}} F(\sqrt{1-t^2},u)\frac{t^{2\mu}(1-t^2-u^2)^{(d-3)/2}}{(n^{-1}+t)^{2\mu}}\,dt\,du.
\end{aligned}$$

Using the trivial inequality $t/(t+n^{-1}) \le 1$ we conclude that

$$J_2 \le c\int_{-1}^{1}\int_0^{\sqrt{1-u^2}} F(\sqrt{1-t^2},u)(1-t^2-u^2)^{(d-3)/2}\,du\,dt.$$

Since $\theta \sim \sin\theta/2 \sim \sqrt{1-\cos\theta}$ for $0 \le \theta \le \pi$, we have

$$F(\sqrt{1-t^2}, u) \sim (1 + n\sqrt{1 - u|x| - t\sqrt{1-|x|^2}})^{-2\sigma}, \qquad 0 \le t \le \sqrt{1-u^2}.$$

But $1 - u|x| - t\sqrt{1-|x|^2} \ge 0$ if $-\sqrt{1-u^2} \le t \le 0$. Therefore, we can enlarge the domain of integration to obtain

$$J_2 \le c \int_{B^2} \frac{(1-t^2-u^2)^{(d-3)/2}\,du\,dt}{(1+n\sqrt{1-u|x|-t\sqrt{1-|x|^2}})^{2\sigma}}.$$

Here B^2 is the unit disk in $\mathbf{R}^2$. We now change the variables $(u,t) \mapsto (a,b)$, where

$$a = \sqrt{1-|x|^2}\,t + |x|u, \qquad b = -|x|t + \sqrt{1-|x|^2}\,u.$$

It is easy to see that this is an orthogonal transformation so that $da\,db = du\,dt$. Hence

$$\begin{aligned}
J_2 \;&\le\; c\int_{B^2} \frac{(1-a^2-b^2)^{(d-3)/2}}{(1+n\sqrt{1-a})^{2\sigma}}\,da\,db \\
&=\; c\int_{-1}^{1} \frac{1}{(1+n\sqrt{1-a})^{2\sigma}} \int_{-\sqrt{1-a^2}}^{\sqrt{1-a^2}} (1-a^2-b^2)^{(d-3)/2}\,db\,da \\
&\le\; c\int_{-1}^{1} \frac{(1-a^2)^{(d-2)/2}}{(1+n\sqrt{1-a})^{2\sigma}}\,da \\
&\le\; \frac{c}{n^{2\sigma}} + c\int_0^1 \frac{(1-a)^{(d-2)/2}}{(1+n\sqrt{1-a})^{2\sigma}}\,da \\
&\le\; \frac{c}{n^{2\sigma}} + \frac{c}{n^d}\int_0^n \frac{s^{d-1}}{(1+s)^{2\sigma}}\,ds \le \frac{c}{n^d},
\end{aligned}$$

since $2\sigma > d$. Thus (4.20) is established when $p = 2$.

To prove (4.20) when $p \ne 2$ we will need the inequalities

$$\begin{aligned}
(4.22)\qquad \frac{\sqrt{1-|x|^2}+n^{-1}}{\sqrt{2}(1+n\,d(x,y))} \;&\le\; \sqrt{1-|y|^2}+n^{-1} \\
&\le\; \sqrt{2}(\sqrt{1-|x|^2}+n^{-1})(1+n\,d(x,y)), \qquad x, y \in B^d,
\end{aligned}$$

which follow readily from (4.9). From this and the definition of $\mathcal{W}_\mu(x;n)$ in (4.10) we get

$$(4.23)\quad c\mathcal{W}_\mu(n;x)(1+n\,d(x,y))^{-2\mu} \le \mathcal{W}_\mu(n;y) \le c\mathcal{W}_\mu(n;x)(1+n\,d(x,y))^{2\mu}.$$

Case 2. $0 < p < 2$. Using (4.23) we obtain

$$\mathcal{W}_\mu(n;y)^{p/2} = \mathcal{W}_\mu(n;y)\mathcal{W}_\mu(n;y)^{p/2-1} \ge \frac{c\mathcal{W}_\mu(n;y)}{\mathcal{W}_\mu(n;x)^{1-p/2}(1+nd(x,y))^{2\mu(1-p/2)}}$$

and hence

$$\int_{B^d} \frac{w_\mu(y)\,dy}{\mathcal{W}_\mu(n;y)^{p/2}(1+n\,d(x,y))^{\sigma p}} \le c\mathcal{W}_\mu(n;x)^{1-p/2} \int_{B^d} \frac{w_\mu(y)\,dy}{\mathcal{W}_\mu(n;y)(1+n\,d(x,y))^{\tau}},$$

where $\tau := (\sigma - 2\mu(1/p - 1/2))p$. By the hypothesis of the lemma, $\tau > d$. Then the above inequality and (4.20) with $p = 2$ imply (4.20) in this case.

Case 3. $2 < p < \infty$. Similarly, as above by (4.23),

$$\mathcal{W}_\mu(n;y)^{p/2} = \mathcal{W}_\mu(n;y)\mathcal{W}_\mu(n;y)^{p/2-1} \ge \frac{c\mathcal{W}_\mu(n;y)\mathcal{W}_\mu(n;x)^{p/2-1}}{(1+n\,d(x,y))^{2\mu(p/2-1)}}.$$

Consequently,

$$\int_{B^d} \frac{w_\mu(y)\,dy}{\mathcal{W}_\mu(n;y)^{p/2}(1+n\,d(x,y))^{\sigma p}} \le c\mathcal{W}_\mu(n;x)^{1-p/2} \int_{B^d} \frac{w_\mu(y)\,dy}{\mathcal{W}_\mu(n;y)(1+n\,d(x,y))^{\tau}},$$

where this time $\tau := (\sigma - 2\mu(1/2 - 1/p))p$. Since $\tau > d$, the above inequality and (4.20) with $p = 2$ imply (4.20) in the case $2 < p < \infty$. ■

It will be vital for our further development that $L_n^\mu(x,y)$ is a *Lip* 1 function in x (or y) with respect to the distance $d(\cdot,\cdot)$. Throughout the rest of the paper, we denote by $B_\xi(r)$ the closed ball centered at ξ of radius $r > 0$ with respect to the distance $d(\cdot,\cdot)$ on B^d, i.e.,

$$B_\xi(r) := \{x \in B^d : d(x,\xi) \le r\}, \qquad \xi \in B^d, \quad r > 0.$$

Proposition 4.7. *Let $\xi, y \in B^d$, $n \ge 1$, and $c^* > 0$. Then, for all $x, z \in B_\xi(c^* n^{-1})$ and an arbitrary k, we have*

$$|L_n^\mu(x,y) - L_n^\mu(\xi,y)| \le c_k \frac{n^{d+1}\,d(x,\xi)}{\sqrt{\mathcal{W}_\mu(n;y)}\sqrt{\mathcal{W}_\mu(n;z)}(1+n\,d(y,z))^k}, \tag{4.24}$$

where c_k depends only on k, μ, d, $\widehat{a}$, and c^*.

Proof. Let $\mu > 0$. We will use the notation $t(x,y;u) := \langle x,y\rangle + u\sqrt{1-|x|^2}\sqrt{1-|y|^2}$, introduced in (4.12). By (4.4) it follows that

$$\begin{aligned} &|L_n^\mu(x,y) - L_n^\mu(\xi,y)| \\ &\qquad \le c\int_{-1}^{1} |Q_n^\lambda(t(x,y;u)) - Q_n^\lambda(t(\xi,y;u))|(1-u^2)^{\mu-1}\,du \\ &\qquad \le c\int_{-1}^{1} \|\partial Q_n^\lambda(\cdot)\|_{L^\infty(I_u)}|t(x,y;u) - t(\xi,y;u)|(1-u^2)^{\mu-1}\,du, \end{aligned} \tag{4.25}$$

where $\partial f = f'$ and I_u is the interval with endpoints $t(x,y;u)$ and $t(\xi,y;u)$.

As in the proof of Theorem 4.2, by estimate (4.5) with $r = 1$, it follows that

$$\begin{aligned}
\|\partial Q_n^\lambda(\cdot)\|_{L^\infty(I_u)} &\le c_k n^{2\lambda+3} \max_{\tau\in I_u}(1+n\sqrt{1-\tau})^{-k} \\
&\le c_k n^{2\lambda+3}((1+n\sqrt{1-t(x,y;u)})^{-k} \\
&\quad + (1+n\sqrt{1-t(\xi,y;u)})^{-k}),
\end{aligned} \tag{4.26}$$

using the fact that $(1+n\sqrt{1-\tau})^{-k}$ is an increasing function of τ.

By the definition of $t(x, y; u)$ it follows that (recall $x' := (x, \sqrt{1-|x|^2})$),

$$\begin{aligned}
&|t(x,y;u) - t(\xi,y;u)| \\
&\qquad \le |\langle x', y'\rangle - \langle \xi', y'\rangle| + |1-u|\sqrt{1-|y|^2}|\sqrt{1-|x|^2} - \sqrt{1-|\xi|^2}| \\
&\qquad \le |\cos d(x,y) - \cos d(\xi,y)| + \sqrt{2}\,|1-u|\sqrt{1-|y|^2}\, d(x,\xi),
\end{aligned}$$

where we used inequality (4.9) from Lemma 4.1. Denote briefly $\alpha := d(x, y)$ and $\beta := d(\xi, y)$. Then

$$\begin{aligned}
|\cos d(x,y) - \cos d(\xi,y)| &= 2\sin\frac{|\alpha-\beta|}{2}\sin\frac{\alpha+\beta}{2} \le \tfrac{1}{2}|\alpha-\beta|(\alpha+\beta) \\
&\le \tfrac{1}{2}|d(x,y) - d(\xi,y)|(d(x,y) + d(\xi,y)) \\
&\le \tfrac{1}{2} d(x,\xi)(d(x,y) + d(\xi,y)) \\
&\le d(x,\xi)(d(y,z) + c^*n^{-1})
\end{aligned}$$

for $z \in B_\xi(c^*n^{-1})$. Hence,

$$|t(x,y;u) - t(\xi,y;u)| \le d(x,\xi)(d(y,z) + c^*n^{-1}) + \sqrt{2}\,|1-u|\sqrt{1-|y|^2}\, d(x,\xi).$$

We use this and (4.26) in (4.25) to obtain

$$\begin{aligned}
&|L_n^\mu(x,y) - L_n^\mu(\xi,y)| \\
&\qquad \le cn^{2\lambda+3}\, d(x,\xi)(d(y,z) + c^*n^{-1}) \\
&\qquad\quad \times \left(\int_{-1}^1 \frac{(1-u^2)^{\mu-1}\,du}{(1+n\sqrt{1-t(x,y;u)})^k} + \int_{-1}^1 \frac{(1-u^2)^{\mu-1}\,du}{(1+n\sqrt{1-t(\xi,y;u)})^k}\right) \\
&\qquad\quad + cn^{2\lambda+3}\sqrt{1-|y|^2}\, d(x,\xi) \\
&\qquad\quad \times \left(\int_{-1}^1 \frac{(1-u)(1-u^2)^{\mu-1}\,du}{(1+n\sqrt{1-t(x,y;u)})^k} + \int_{-1}^1 \frac{(1-u)(1-u^2)^{\mu-1}\,du}{(1+n\sqrt{1-t(\xi,y;u)})^k}\right) \\
&\qquad =: A_1 + A_2 + A_3 + A_4.
\end{aligned}$$

By Lemma 4.4 with $\gamma = \mu - 1$, we have

$$A_1 \le cn^{2\lambda+3}\, d(x,\xi)(d(y,z) + c^*n^{-1}) \frac{n^{-2\mu}}{\sqrt{\mathcal{W}_\mu(n;x)}\sqrt{\mathcal{W}_\mu(n;y)}(1+n\,d(x,y))^\sigma}$$

with $\sigma := k - 3\mu - 1$. Note that, for $y \in B^d$ and all $z \in B_\xi(c^* n^{-1})$, we have $1 + n\,d(z, y) \sim 1 + n\,d(\xi, y)$ and $\sqrt{1-|z|^2} + c^* n^{-1} \sim \sqrt{1-|\xi|^2} + n^{-1}$, using (4.9). Consequently,

$$A_1 \le \frac{cn^{d+1}\,d(x,\xi)}{\sqrt{\mathcal{W}_\mu(n;x)}\sqrt{\mathcal{W}_\mu(n;z)}(1+n\,d(y,z))^{\sigma-1}}. \tag{4.27}$$

We similarly obtain the same bound for A_2.

To estimate A_3 we employ Lemma 4.4 with $\gamma = \mu$ and obtain

$$\begin{aligned} A_3 \;\le\; & cn^{2\lambda+3}\sqrt{1-|y|^2}\,d(x,\xi) \\ & \times \frac{n^{-2\mu-2}}{(\sqrt{1-|x|^2}+n^{-1})^{\mu+1}(\sqrt{1-|y|^2}+n^{-1})^{\mu+1}(1+n\,d(x,y))^{\sigma}} \end{aligned} \tag{4.28}$$

with $\sigma := k - 3\mu - 4$. By cancelling appropriate terms we conclude that (4.27) holds for A_3 as well. In exactly the same way one can see that A_4 also satisfies (4.28) and hence (4.27). The proof of the proposition is complete. ■

Remark 4.8. For the sake of completeness, we record next some simple properties of the operators with kernels $L_n^\mu(x, y)$ from (4.3), where $\widehat{a}$ is admissible of type (a). Denote

$$(\mathcal{L}_n^\mu f)(x) := \int_{B^d} f(y) L_n^\mu(x, y) w_\mu(y)\,dy, \qquad \mu \ge 0, \tag{4.29}$$

Evidently, $\mathcal{L}_n^\mu f \in \Pi_{2n}^d$ and $\mathcal{L}_n^\mu g = g$ for $g \in \Pi_n^d$. A classical argument using Proposition 4.5 shows that, for $1 \le p \le \infty$, $\|\mathcal{L}_n^\mu\|_{L_\mu^p \to L_\mu^p} \le c$ and hence

$$\|f - \mathcal{L}_n^\mu f\|_{L_\mu^p} \le cE_n(f)_p, \qquad f \in L_\mu^p,$$

where $E_n(f)_p$ denotes the best approximation of f from Π_n^d in $L_\mu^p := L^p(B^d, w_\mu)$. In the next section we will put the operators $\mathcal{L}_n^\mu$ to work for their primary purpose, namely, for the construction of cubature formulas on B^d.

5. Cubature Formula on B^d

Cubature formulas on B^d with weights $w_\mu(x)$, $\mu \ge 0$, which are exact for all polynomials of degree n, are valuable from many prospectives. Those with positive coefficients are preferred for numerical computation and are called positive cubature formulas. In the literature, only a handful of positive cubature formulas are known. For our purpose of constructing polynomial frames on B^d (see Section 6)) we will need positive cubature whose knots are almost equally distributed with respect to the distance $d(\cdot,\cdot)$ introduced in (4.7). To the best of our knowledge there are no such cubature formulas available up to now. There is a close relation between cubature formulas on the unit ball and those on the unit sphere S^d [20]. In the following we will apply the method used in [12], [13] (see also [15]) for constructing cubature formulas on the unit sphere.

One of the difficulties in constructing cubature formulas on B^d is the lack of uniformly distributed points on B^d. We shall use, as a substitute, sets of "almost equally distributed points" with respect to the distance $d(\cdot,\cdot)$ in B^d which we describe next.

Definition 5.1. We say that a set $\mathcal{X}_\varepsilon \subset B^d$, along with an associated partition $\mathcal{R}_\varepsilon$ of B^d consisting of measurable subsets of B^d, is a *set of almost uniformly ε-distributed points* on B^d if:

(i) $B^d = \bigcup_{R\in\mathcal{R}_\varepsilon} R$ and the sets in $\mathcal{R}_\varepsilon$ do not overlap ($R_1^\circ \cap R_2^\circ = \emptyset$ if $R_1 \neq R_2$).
(ii) For each $R \in \mathcal{R}_\varepsilon$ there is a unique $\xi \in \mathcal{X}_\varepsilon$ such that $B_\xi(c^*\varepsilon) \subset R \subset B_\xi(\varepsilon)$.

Hence $\#\mathcal{X}_\varepsilon = \#\mathcal{R}_\varepsilon \le c^{**}\varepsilon^{-d}$. Here the constant $c^* > 0$, depending only on d, is fixed but sufficiently small, so that the existence of sets of almost uniformly ε-distributed points on B^d is guaranteed (see the next lemma).

Lemma 5.2. *For sufficiently small constant $c^* > 0$, depending only on d, and an arbitrary $0 < \varepsilon \le \pi$, there exists a set $\mathcal{X}_\varepsilon \subset B^d$ of almost uniformly ε-distributed points on B^d, where the associated partition $\mathcal{R}_\varepsilon$ of B^d consists of projections of spherical simplices.*

Proof. As we already mentioned, the distance $d(x, y)$ ($x, y \in B^d$) is the geodesic distance between $x', y' \in S^d_+$. So, we need to subdivide properly S^d_+. We first divide S^d_+ into 2^d spherical simplices analogous to the intersections of S^2 with the octants in $\mathbf{R}^3$. Let $\mathcal{O}_1$ be the spherical simplex on which all coordinates of $\xi \in \mathcal{O}_1$ are nonnegative and let

$$\mathcal{T}_1 := \left\{ \sum_{j=1}^{d+1} t_j e_j : t_j \ge 0,\ \sum_{j=1}^{d+1} t_j = 1 \right\},$$

where $\{e_j\}$ are the standard unit vectors in $\mathbf{R}^{d+1}$. If $v := (1, \dots, 1)$, then the map $x(\xi) := \xi/\langle \xi, v\rangle$ gives an one-to-one correspondence between $\mathcal{O}_1$ and $\mathcal{T}_1$. It is readily seen that, for $\xi, \zeta \in \mathcal{O}_1$,

$$\frac{1}{2\sqrt{d}} d(\xi, \zeta) \le |x(\xi) - x(\zeta)| \le 2\sqrt{d}\, d(\xi, \zeta). \tag{5.1}$$

Here $|\cdot|$ denotes the Euclidean norm in $\mathbf{R}^{d+1}$ and $d(\cdot,\cdot)$ is the geodesic distance on $S^d \subset \mathbf{R}^{d+1}$.

We set $M := \lceil 2\sqrt{d}\varepsilon^{-1} \rceil$ and divide the equilateral simplex $\mathcal{T}_1$ into M^d equal equilateral subsimplices of side length $L = \sqrt{2}/M$. We denote by $\widetilde{\mathcal{R}}^1_\varepsilon$ the set of all spherical simplices obtained by applying the inverse map x^{-1} to the simplices defined above. We similarly define the set $\widetilde{\mathcal{X}}^1_\varepsilon$ of the "centers" of all spherical simplices in $\widetilde{\mathcal{R}}^1_\varepsilon$ by applying the inverse map x^{-1} to the midpoints of the corresponding Euclidean simplices. After these preparations, we define $\mathcal{R}^1_\varepsilon$ as the set of projections onto B^d of all spherical simplices from $\widetilde{\mathcal{R}}^1_\varepsilon$ and we similarly define $\mathcal{X}^1_\varepsilon$.

It is straightforward to show that an equilateral Euclidean simplex with each side of length L contains the ball of radius $L/\sqrt{2d(d+1)}$ centered at its midpoint and is contained in a ball of radius $< L/\sqrt{2}$ with the same center. Then (5.1) yields that the corresponding spherical simplex contains the spherical cap centered at its center and of radius $L/(2d\sqrt{2(d+1)})$ and is contained in a spherical cap with the same center and radius $< \sqrt{2d}L \le 2\sqrt{d}/M \le \varepsilon$. This establishes property (ii) of Lemma 5.2 for the spherical simplices in $\mathcal{R}^1_\varepsilon$. Also, we have $\#\mathcal{X}^1_\varepsilon = \#\mathcal{R}^1_\varepsilon = M^d \le (4\sqrt{d}\varepsilon^{-1})^d$.

Repeating this procedure with all other initial simplices, we establish the existence of the desired partition $\mathcal{R}_\varepsilon$. ■

The following lemma contains additional information about the partition $\mathcal{R}_\varepsilon$ from above.

Lemma 5.3. *Let $\mathcal{R}_\varepsilon$ be as in Lemma 5.2. Then, for any $\xi \in \mathcal{X}_\varepsilon$,*

$$(5.2) \qquad |R_\xi| := \int_{R_\xi} 1\,dx \sim \varepsilon^d \sqrt{1-|\xi|^2} \sim \varepsilon^d(\sqrt{1-|\xi|^2}+\varepsilon)$$

and

$$(5.3) \quad m_\mu(R_\xi) := \int_{R_\xi} w_\mu(x)\,dx \sim \varepsilon^d(1-|\xi|^2)^\mu = \varepsilon^d \frac{w_\mu(\xi)}{w_0(\xi)} \sim \varepsilon^d(\sqrt{1-|\xi|^2}+\varepsilon)^{2\mu}.$$

Here the constants of equivalence depend only on d and μ.

Proof. To prove (5.2) we use Lemma 5.2. Hence property (ii) in Definition 5.1 yields

$$(5.4) \qquad |R_\xi| \sim |B_\xi(\varepsilon)| \qquad \text{and} \qquad d(\xi, \partial B^d) \geq c^*\varepsilon.$$

We can assume without loss of generality that ξ lies on the positive x_1-axis, i.e., $\xi = (\xi_1, 0, \dots, 0)$ and $0 < \xi_1 < 1$. The boundary $\partial B_\xi(\varepsilon)$ of $B_\xi(\varepsilon)$ is given by the equation $x_1\xi_1 + \sqrt{1-|x|^2}\sqrt{1-\xi_1^2} = \cos\varepsilon$. A simple manipulation of this identity shows that $\partial B_\xi(\varepsilon)$ is the ellipsoid

$$\frac{(x_1 - \xi_1\cos\varepsilon)^2}{1-|\xi_1|^2} + x_2^2 + \dots + x_d^2 = \sin^2\varepsilon.$$

From this it follows that $|B_\xi(\varepsilon)| \sim \varepsilon^d\sqrt{1-|\xi|^2}$ (using that $\sin\varepsilon \sim \varepsilon$) and then (5.2) follows.

We now turn to the proof of (5.3). There are two cases to be considered.

Case 1. $\mu \geq \frac{1}{2}$. Denote $R_\xi^- := R_\xi \cap \{x \in B^d : |x| \leq |\xi|\}$. It is easily seen that $|R_\xi^-| \sim |R_\xi| \sim \varepsilon^d\sqrt{1-|\xi|^2}$. Then

$$\int_{R_\xi} w_\mu(x)\,dx \geq \int_{R_\xi^-} w_\mu(x)\,dx \geq w_\mu(\xi)|R_\xi^-| \sim w_\mu(\xi)\varepsilon^d\sqrt{1-|\xi|^2} = \varepsilon^d(1-|\xi|^2)^\mu.$$

Since ξ is in the center of $\mathcal{R}_\xi$ by construction, we have $\sqrt{1-|\xi|^2} \geq c\varepsilon$. Hence, for $x \in R_\xi \subset B_\xi(\varepsilon)$, inequality (4.9) shows that

$$w_\mu(x) \leq (\sqrt{1-|\xi|^2}+\varepsilon)^{2\mu-1} \leq c w_\mu(\xi),$$

which yields

$$\int_{R_\xi} w_\mu(x)\,dx \leq c w_\mu(\xi)|R_\xi| \sim w_\mu(\xi)\varepsilon^d\sqrt{1-|\xi|^2} = \varepsilon^d(1-|\xi|^2)^\mu.$$

Case 2. $0 \le \mu < \frac{1}{2}$. Denote $R_\xi^+ := R_\xi \cap \{x \in B^d : |x| \ge |\xi|\}$. Proceeding as above we again get (5.3).

Finally, using (5.4) we obtain $\sqrt{1-|\xi|^2} \ge \sin c^*\varepsilon \ge c\varepsilon$ which implies the last equivalence in (5.3). The proof of the lemma is complete. ■

Theorem 5.4. *There exists a constant* $c^\diamond > 0$ (*depending only on* d) *such that for any* $n \ge 1$ *and a set* $\mathcal{X}_\varepsilon$ *of almost uniformly* ε*-distributed points on* B^d *with* $\varepsilon := c^\diamond/n$, *there exist positive coefficients* $\{\lambda_\xi\}_{\xi\in\mathcal{X}_\varepsilon}$ *such that the cubature formula*

$$\int_{B^d} f(x)\,dx \sim \sum_{\xi\in\mathcal{X}_\varepsilon} \lambda_\xi f(\xi)$$

is exact for all polynomials of degree $\le n$. *In addition*,

$$\lambda_\xi \sim n^{-d}\mathcal{W}_\mu(n;\xi) \sim \varepsilon^d(1-|\xi|^2)^\mu \sim m_\mu(B_\xi(\varepsilon))$$

with constants of equivalence depending only on μ *and* d. *Here* $m_\mu(E) := \int_E w_\mu(x)\,dx$.

Note that when $\mu = 0$ the cubature formula of Theorem 5.4 can be derived from the cubature formula on S^{d+1} from [12], [13], [15] by applying [20, Theorem 4.2].

Assume that $\mathcal{X}_\varepsilon$ (with associated partition $\mathcal{R}_\varepsilon$) is a set of almost uniformly ε-distributed points on B^d (see Definition 5.1), where $\varepsilon = \delta/n$ with $n \ge 1$ and δ will be selected later on. We introduce the following weighted ℓ_1-norm for functions defined on B^d:

$$\|f\|_{\ell^1_\mu(\mathcal{X}_\varepsilon)} := \sum_{\xi\in\mathcal{X}_\varepsilon} |f(\xi)| m_\mu(R_\xi). \tag{5.5}$$

Also, recall the notation $\|f\|_{L^1_\mu} = \|f\|_{L^1(w_\mu,B^d)} := \int_{B^d} |f(x)|w_\mu(x)\,dx$.

We need a couple of additional results.

Lemma 5.5. *If* $g \in \Pi_n^d$, *then*

$$\big|\|g\|_{L^1_\mu} - \|g\|_{\ell^1_\mu(\mathcal{X}_\varepsilon)}\big| \le \sum_{\xi\in\mathcal{X}_\varepsilon}\int_{R_\xi} |g(x)-g(\xi)|w_\mu(x)\,dx \le c^\star\delta\|g\|_{L^1_\mu} \tag{5.6}$$

and hence

$$(1-c^\star\delta)\|g\|_{L^1_\mu} \le \|g\|_{\ell^1_\mu(\mathcal{X}_\varepsilon)} \le (1+c^\star\delta)\|g\|_{L^1_\mu}, \tag{5.7}$$

where $c^\star$ *depends only on* d *and* μ.

Proof. Let $\mathcal{L}_n^\mu$ be the operator from (4.29). Using that $g = \mathcal{L}_n^\mu g$ and the fact that $\mathcal{R}_\varepsilon$ is a partition of B^d (see Lemma 5.2), we obtain

$$\begin{aligned}
\big|\|g\|_{L^1_\mu} - \|g\|_{\ell^1_\mu(\mathcal{X}_\varepsilon)}\big| &\le \sum_{\xi\in\mathcal{X}_\varepsilon}\int_{R_\xi} |g(x)-g(\xi)|w_\mu(x)\,dx \\
&\le \sum_{\xi\in\mathcal{X}_\varepsilon}\int_{R_\xi}\int_{B^d} |L_n^\mu(x,y)-L_n^\mu(\xi,y)||g(y)|w_\mu(y)\,dy w_\mu(x)\,dx \\
&\le \|g\|_{L^1_\mu} \sup_{y\in B^d}\sum_{\xi\in\mathcal{X}_\varepsilon}\int_{R_\xi} |L_n^\mu(x,y)-L_n^\mu(\xi,y)|w_\mu(x)\,dx.
\end{aligned}$$

By Proposition 4.7 with $z = x$, it follows that

$$\int_{R_\xi} |L_n^\mu(x, y) - L_n^\mu(\xi, y)| w_\mu(x)\, dx$$

$$\leq \int_{R_\xi} \frac{c_k n^{d+1}\, d(x, \xi) w_\mu(x)\, dx}{\sqrt{\mathcal{W}_\mu(n; x)}\sqrt{\mathcal{W}_\mu(n; y)}(1 + n\, d(y, x))^k}.$$

Choosing k sufficiently large ($k > d + \mu$ will do) we apply Lemma 4.6 with $p = 1$ and use that $d(x, \xi) \leq \delta n^{-1}$ for $x \in R_\xi$ to obtain

$$\sup_{y \in B^d} \sum_{\xi \in \mathcal{X}_\varepsilon} \int_{R_\xi} |L_n^\mu(x, y) - L_n^\mu(\xi, y)| w_\mu(x)\, dx$$

$$\leq c\delta n^d \int_{B^d} \frac{w_\mu(x)\, dx}{\sqrt{\mathcal{W}_\mu(n; x)}\sqrt{\mathcal{W}_\mu(n; y)}(1 + nd(y, x))^k} \leq c\delta.$$

The lemma follows. ∎

The Farkas Lemma. A variant of the well known in Optimization Farkas lemma will play an important role in the proof of Theorem 5.4.

Proposition 5.6. *Let V be a finite-dimensional real vector space and denote by V^* its dual. Let $u_1, u_2, \dots, u_n \in V^*$ and suppose $u \in V^*$ has the property that $u(x) \geq 0$ for all $x \in V$ such that $u_j(x) \geq 0$ for $j = 1, 2, \dots, n$. Then there exist $a_j \geq 0$, $j = 1, 2, \dots, n$, such that*

$$u = \sum_{j=1}^{n} a_j u_j. \tag{5.8}$$

For the proof of this proposition, see, e.g., [1].

Proof of Theorem 5.4. First, we choose $\delta := 1/3c^\star$, where $c^\star$ is the constant from Lemma 5.5. In applying Proposition 5.6, we take $V := \Pi_n^d$ and $\{u_j\}$ to be the set of all point evaluation functionals $\{\delta_\xi\}_{\xi \in \mathcal{X}_\varepsilon}$.

Let the linear functionals u and u_γ be defined by

$$u(g) := \int_{B^d} g(x) w_\mu(x)\, dx \qquad \text{and} \qquad u_\gamma(g) := u(g) - \gamma \sum_{\xi \in \mathcal{X}_\varepsilon} g(\xi) m_\mu(R_\xi).$$

Since $c^\star\delta = \frac{1}{3}$, the left-hand side estimate in (5.7) yields

$$\|g\|_{L_\mu^1} \leq \tfrac{3}{2} \|g\|_{\ell_\mu^1(\mathcal{X}_\varepsilon)}, \qquad g \in \Pi_n^d. \tag{5.9}$$

Suppose $g \in \Pi_n^d$ and $g(\xi) \geq 0$ for all $\xi \in \mathcal{X}_\varepsilon$. Then using (5.6) with $c^*\delta = \frac{1}{3}$ and (5.9), we obtain

$$|u(g) - \|g\|_{\ell_\mu^1(\mathcal{X}_\varepsilon)}| = \sum_{\xi \in \mathcal{X}_\varepsilon} \int_{R_\xi} |g(x) - g(\xi)| w_\mu(x)\, dx \leq c^\star\delta \|g\|_{L_\mu^1} \leq \tfrac{1}{2} \|g\|_{\ell_\mu^1(\mathcal{X}_\varepsilon)}$$

and hence $u(g) \geq \frac{1}{2}\|g\|_{\ell^1_\mu(\mathcal{X}_\varepsilon)}$. Choose $\gamma := \frac{1}{3}$. Then since $g(\xi) \geq 0$, $\xi \in \mathcal{X}_\varepsilon$, we obtain

$$u_\gamma(g) = u(g) - \tfrac{1}{3}\|g\|_{\ell^1_\mu(\mathcal{X}_\varepsilon)} \geq \tfrac{1}{6}\|g\|_{\ell^1_\mu(\mathcal{X}_\varepsilon)} \geq 0.$$

Applying Proposition 5.6 to u_γ, there exist numbers $a_\xi \geq 0$, $\xi \in \mathcal{X}_\varepsilon$, such that

$$u_\gamma(g) = \sum_{\xi \in \mathcal{X}_\varepsilon} a_\xi g(\xi), \qquad g \in \Pi_n^d,$$

and hence

$$u(g) = \sum_{\xi \in \mathcal{X}_\varepsilon} (a_\xi + \tfrac{1}{3} m_\mu(R_\xi)) g(\xi) =: \sum_{\xi \in \mathcal{X}_\varepsilon} \lambda_\xi g(\xi), \qquad g \in \Pi_n^d.$$

Therefore, the linear functional $\sum_{\xi \in \mathcal{X}_\varepsilon} \lambda_\xi g(\xi)$ provides a cubature formula exact for all polynomials of degree n.

Clearly, $\lambda_\xi \geq m_\mu(R_\xi)/3$ and the estimate $\lambda_\xi \leq c m_\mu(R_\xi)$ follows from Lemma 5.3 and Proposition 5.7 below. ■

The last ingredient in bounding λ_ξ from above is the following general result that is of independent interest.

Proposition 5.7. *If a positive cubature formula*

$$\int_{B^d} f(x) w_\mu(x)\, dx \sim \sum_{\xi \in \mathcal{X}_\varepsilon} \lambda_\xi f(\xi), \qquad \lambda_\xi > 0, \quad |\xi| < 1, \tag{5.10}$$

is exact for all polynomials of degree $\leq n$, *then*

$$\lambda_\xi \leq c n^{-d} \mathcal{W}_\mu(n;\xi) = c n^{-d} (\sqrt{1 - |\xi|^2} + n^{-1})^{2\mu}, \qquad \xi \in \mathcal{X}_\varepsilon, \tag{5.11}$$

where $c > 0$ *depends only on* μ *and* d.

Proof. Recall the kernel $K_m(w_\mu; x, y)$ defined in (1.3). Evidently, $K_m(w_\mu; \xi, \xi) > 0$ and

$$\int_{B^d} [K_m(w_\mu; x, y)]^2 w_\mu(y)\, dy = K_m(w_\mu; x, x).$$

Let $m = \lfloor n/2 \rfloor$. Then it follows that

$$\lambda_\xi \leq \sum_{\eta \in \mathcal{X}_\varepsilon} \lambda_\eta \left[\frac{K_m(w_\mu; \xi, \eta)}{K_m(w_\mu; \xi, \xi)}\right]^2 = \int_{B^d} \left[\frac{K_m(w_\mu; \xi, x)}{K_m(w_\mu; \xi, \xi)}\right]^2 w_\mu(x)\, dx = \frac{1}{K_m(w_\mu; \xi, \xi)}.$$

Hence, the stated result is a consequence of an upper bound for $[K_m(w_\mu; x, x)]^{-1}$, to be established in Proposition 5.9 below. ■

In order to establish the needed upper bound for $[K_n(w_\mu; x, x)]^{-1}$ we now construct a family of well-localized polynomials.

Lemma 5.8. *For any $k, m \ge 1$ and $\xi \in B^d$ there exists a polynomial $P_\xi \in \Pi^d_{2km}$ and a constant $c^* > 0$ depending only on k and d such that $P_\xi(\xi) = 1$ and for $0 \le \gamma \le k$, $x \in B^d$,*

$$0 \le P_\xi(x) \le \frac{c^*}{(1 + m\, d(\xi, x))^{2k}} \le \frac{c(\sqrt{1-|\xi|^2} + m^{-1})^\gamma}{(\sqrt{1-|x|^2} + m^{-1})^\gamma (1 + m\, d(\xi, x))^k}. \tag{5.12}$$

Proof. Let $q(\theta) := (\sin(m\theta/2)/m \sin(\theta/2))^{2k}$. Evidently, q is a trigonometric polynomial of degree less than km, $q(0) = 1$, and

$$0 \le q(\theta) \le \frac{c}{(1 + m|\theta|)^{2k}}, \qquad |\theta| \le \pi. \tag{5.13}$$

For $0 \le \alpha \le \pi$, we define the algebraic polynomial $Q_\alpha(t)$ by

$$Q_\alpha(\cos\theta) := \frac{q(\theta - \alpha) + q(\theta + \alpha)}{1 + q(2\alpha)}.$$

It is readily seen that deg $Q_\alpha < km$, $Q_\alpha(\cos\alpha) = 1$, and

$$0 \le Q_\alpha(\cos\theta) \le \frac{c}{(1 + m|\theta - \alpha|)^{2k}}, \qquad 0 \le \theta \le \pi. \tag{5.14}$$

Also, $Q_{\pi/2}$ is even and

$$0 \le Q_{\pi/2}(t) \le \frac{c}{(1 + m|\arccos t - \pi/2|)^{2k}} \le \frac{c}{(1 + m|t|)^{2k}}, \qquad |t| \le 1. \tag{5.15}$$

Without loss of generality we may assume that $\xi = (\xi_1, 0, \dots, 0)$ with $0 < \xi_1 < 1$. We choose $\alpha \in (0, \pi/2)$ so that $\xi_1 = \cos\alpha$. Then (5.14) gives

$$0 \le Q_\alpha(t) \le \frac{c}{(1 + m\, d_1(\xi_1, t))^{2k}}, \qquad |t| \le 1, \tag{5.16}$$

where $d_1(\xi_1, t) := \arccos(\xi_1 t + \sqrt{1 - \xi_1^2}\sqrt{1 - t^2})$ is the univariate version of the distance $d(\cdot, \cdot)$ (see (4.7)). We define

$$P_\xi(x) := Q_\alpha(x_1) Q_{\pi/2}(\sqrt{x_2 + \cdots + x_d^2}).$$

Clearly, $P_\xi \in \Pi^d_{2km}$, $P_\xi(\xi) = 1$ and, by (5.15)–(5.16),

$$0 \le P_\xi(x) \le \frac{c}{[(1 + m|x_*|)(1 + m\, d_1(\xi_1, x_1))]^{2k}}, \qquad x \in B^d, \tag{5.17}$$

where $x_* := (x_2, \dots, x_d)$ and $|x_*| := (x_2^2 + \cdots + x_d^2)^{1/2}$.

It remains to show that P_ξ obeys (5.12). To this end, we first show that

$$d(\xi, x) \le 2(|x_*| + d_1(\xi_1, x_1)). \tag{5.18}$$

Denote briefly $x_\diamond := (x_1, 0, \dots, 0)$. We have

$$d(\xi, x) \le d(\xi, x_\diamond) + d(x_\diamond, x) = d_1(\xi_1, x_1) + d(x_\diamond, x).$$

Our next step is to prove the inequality

$$d(x_\diamond, x) \le 2|x_*|. \tag{5.19}$$

Evidently,

$$d(x_\diamond, x) = \arccos(\langle x'_\diamond, x'\rangle) = \arccos(x_1^2 + \sqrt{1-x_1^2}\sqrt{1-x_1^2-x_2^2-\cdots-x_d^2}).$$

One easily verifies the inequality $\arccos t \le 2\sqrt{1-t}$, $0 \le t \le 1$, and hence (5.19) will be established if we show that

$$(1 - x_1^2 - \sqrt{1-x_1^2}\sqrt{1-x_1^2-|x_*|^2})^{1/2} \le |x_*|.$$

Denote briefly $a := \sqrt{1-x_1^2}$ and $b := |x_*|$. Then the above inequality is equivalent to $a^2 - a\sqrt{a^2-b^2} \le b^2$ or $a\sqrt{a^2-b^2} - (a^2-b^2) \ge 0$. But the latter inequality is apparently valid since

$$a\sqrt{a^2-b^2} - (a^2-b^2) = \frac{b^2\sqrt{a^2-b^2}}{a+\sqrt{a^2-b^2}} \ge 0.$$

Thus (5.19) is established and hence (5.18) holds. Combining (5.17) with (5.18) gives

$$0 \le P_\xi(x) \le \frac{c}{(1+m\, d(\xi,x))^{2k}}, \qquad x \in B^d, \tag{5.20}$$

which is the first estimate of $P_\xi(x)$ in (5.12).

To prove the second estimate in (5.12) we need the estimate

$$\frac{1}{1+m\, d(\xi,x)} \le c\frac{\sqrt{1-|\xi|^2}+m^{-1}}{\sqrt{1-|x|^2}+m^{-1}}, \qquad x \in B^d, \tag{5.21}$$

which apparently follows by inequality (4.9) in Lemma 4.1 (see also (4.22)).

Finally, applying (5.21) in (5.20), we get the second estimate in (5.12), which completes the proof. ∎

The function $\Lambda_n(x) := [K_n(w_\mu; x, x)]^{-1}$ is the so-called Christoffel function, which has the following characteristic property [4, p. 109]:

$$\Lambda_n(x) = \min_{P(x)=1,\, P\in\Pi_n^d} \int_{B^d} [P(y)]^2 w_\mu(y)\, dy, \qquad x \in B^d. \tag{5.22}$$

The localized polynomials in Lemma 5.8 give an upper bound for the Christoffel function, used in the proof of Proposition 5.7.

Proposition 5.9. *For any $\mu \ge 0$ and $d > 1$ there exists a constant $c > 0$ such that*

$$\Lambda_n(x) \le cn^{-d}\mathcal{W}_\mu(n; x), \qquad x \in B^d, \quad n \ge 1. \tag{5.23}$$

Proof. Write $k := [\max\{d/2, \mu\}] + 1$ and let $n \ge 4k$ (the case $1 \le n < 4k$ is trivial). Set $m := [n/2k]$. By Lemma 5.8 there exists a polynomial $P_x(y) \in \Pi_n^d$ such that $P_x(x) = 1$ and (5.12) holds with $\gamma = \mu$ and ξ, x replaced by x, y. Then, by (5.22), (5.12), and Lemma 4.6 with $p = 2$, we infer

$$\begin{aligned}\Lambda_n(x) \;&\le\; \int_{B^d} [P_x(y)]^2 w_\mu(y)\,dy \le c\int_{B^d} \frac{\mathcal{W}_\mu(m; x) w_\mu(y)\,dy}{\mathcal{W}_\mu(m; y)(1 + m\,d(x, y))^{2k}}\\ &\le\; cm^{-d}\mathcal{W}_\mu(m; x) \le cn^{-d}\mathcal{W}_\mu(n; x). \qquad \blacksquare\end{aligned}$$

For the construction of our frames, we will need the following result which is an immediate consequence of Lemma 5.2 and Theorem 5.4.

Corollary 5.10. *There exists a sequence $\{\mathcal{X}_j\}_{j=0}^\infty$ of sets of almost uniformly ε_j-distributed points on B^d ($\mathcal{X}_j := \mathcal{X}_{\varepsilon_j}$) with $\varepsilon_j := c^\diamond 2^{-j-2}$ and there exist positive coefficients $\{\lambda_\xi\}_{\xi\in\mathcal{X}_j}$ such that the cubature*

$$\int_{B^d} f(x) w_\mu(x)\,dx \sim \sum_{\xi\in\mathcal{X}_j} \lambda_\xi f(\xi) \tag{5.24}$$

is exact for all polynomials of degree $\le 2^{j+2}$. Moreover, $\lambda_\xi \sim m_\mu(B_\xi(2^{-j}))$ and $\#\mathcal{X}_j \sim 2^{jd}$ with constants of equivalence depending only on d and μ.

6. Localized Polynomial Frames (Needlets) in $L^2(B^d, w_\mu)$

We will apply the general scheme, described in Section 2, for construction of polynomial frames in $L^2_\mu := L^2(B^d, w_\mu)$. To this end, we will utilize the localized polynomials from Theorem 4.2 and the cubature formula from Theorem 5.4 (see Corollary 5.10).

Let $\widehat{a}$ satisfy the conditions

$$\widehat{a} \in C^\infty[0, \infty), \quad \widehat{a} \ge 0, \qquad \operatorname{supp}\widehat{a} \subset [\tfrac{1}{2}, 2], \tag{6.1}$$

$$\widehat{a}(t) > c > 0, \qquad \text{if} \quad t \in [\tfrac{3}{5}, \tfrac{5}{3}], \tag{6.2}$$

$$\widehat{a}^2(t) + \widehat{a}^2(2t) = 1, \qquad \text{if} \quad t \in [\tfrac{1}{2}, 1]. \tag{6.3}$$

For the construction of such functions, see, e.g., [15].

We introduce a sequence of polynomial "kernels": $L_0(x, y) := P_0(w_\mu; x, y)$ and (see Section 4)

$$L_j(x, y) := \sum_{\nu=0}^\infty \widehat{a}\left(\frac{\nu}{2^{j-1}}\right) P_\nu(w_\mu; x, y), \qquad j = 1, 2, \ldots.$$

We now define the needlets (frame elements) by

$$\psi_\xi(x) := \sqrt{\lambda_\xi}\cdot L_j(x, \xi) \qquad \text{for} \quad \xi \in \mathcal{X}_j, \qquad j = 0, 1, \ldots,$$

where $\mathcal{X}_j$ is the set of the knots and the λ_ξ's are the coefficients of the cubature formula (5.24) from Corollary 5.10. We write $\mathcal{X} := \bigcup_{j=0}^\infty \mathcal{X}_j$ (see Section 2) and define the needlet system Ψ by

$$\Psi := \{\psi_\xi\}_{\xi\in\mathcal{X}}.$$

Denoting

$$(L_j * f)(x) := \int_{B^d} L_j(x, y) f(y) w_\mu(y)\, dy,$$

we get as in (2.7) the semidiscrete needlet decomposition of L^2_μ:

$$f = \sum_{j=0}^\infty L_j * L_j * f \qquad \text{for} \quad f \in L^2_\mu.$$

It readily follows by (2.10)–(2.11) and Corollary 5.10 that the needlet system Ψ is a tight frame in L^2_μ:

Theorem 6.1. *If $f \in L^2_\mu$, then*

$$f = \sum_{j=0}^\infty \sum_{\xi\in\mathcal{X}_j} \langle f, \psi_\xi\rangle \psi_\xi = \sum_{\xi\in\mathcal{X}} \langle f, \psi_\xi\rangle \psi_\xi \quad \textit{in } L^2_\mu$$

and

$$\|f\|_{L^2_\mu} = \left(\sum_{\xi\in\mathcal{X}} |\langle f, \psi_\xi\rangle|^2 \right)^{1/2}.$$

Remark 6.2. We restricted our attention here to the needlet decomposition of L^2_μ only, however, much more is true. In [9] we show that the needlets from this paper can be used for characterization of $L^p_\mu(B^d)$ $(1 < p < \infty)$ and the more general weighted Triebel–Lizorkin and Besov spaces on the ball with weight $w_\mu(x)$.

We finally show that each needlet ψ_ξ has faster than any polynomial rate of decay away from its "center" ξ with respect to the distance $d(\cdot, \cdot)$ on B^d. This property of the needlets is critical for using them for decomposition of spaces other than L^2_μ.

Theorem 6.3. *For any $k > 0$ there exists a constant $c_k > 0$ depending only on k, μ, d, and $\widehat{a}$ such that, for $\xi \in \mathcal{X}_j$, $j = 0, 1, \ldots,$*

$$|\psi_\xi(x)| \le c_k \frac{2^{jd/2}}{\sqrt{\mathcal{W}_\mu(2^j; x)}(1 + 2^j\, d(x, \xi))^k}, \qquad x \in B^d. \tag{6.4}$$

Proof. Estimate (6.4) follows readily from (4.11) (see Theorem 4.2), taking into account that $\lambda_\xi \le c2^{-jd}\mathcal{W}_\mu(2^j; \xi)$ for $\xi \in \mathcal{X}_j$. ∎

Corollary 6.4. *For* $0 < p \le \infty$, *we have*

$$\|\psi_\xi\|_{L^p_\mu} \le c \left(\frac{2^{jd}}{\mathcal{W}_\mu(2^j;\xi)} \right)^{1/2-1/p}, \qquad \xi \in \mathcal{X}_j. \tag{6.5}$$

In particular, $\|\psi_\xi\|_{L^2_\mu} \le c$, *which shows that estimate* (6.4) *is sharp* (*in a sense*).

Estimate (6.5) follows readily by (6.4) and Lemma 4.6.

Acknowledgment. The first author has been supported by a grant from the National Geospatial-Intelligence Agency. The second author has been supported in part by NSF grant DMS-0604056.

References

[1] J. BORWEIN (1983): *A note on Farkas' lemma*. Utilitas Math., **24**:235–241.

[2] L. BOS, N. LEVENBERG, S. WALDRON (2004): *Metrics associated to multivariate polynomial inequalities*. In: Advances in Constructive Approximation (Vanderbilt 2003), pp. 133–147. Brentwood, TN: Nashboro Press.

[3] G. BROWN, F. DAI (2005): *Approximation of smooth functions on compact two-point homogeneous spaces*. J. Funct. Anal., **220**(2):401–423.

[4] C. F. DUNKL, YUAN XU (2001): Orthogonal Polynomials of Several Variables. Cambridge: Cambridge University Press.

[5] J. DZIUBAŃSKI (1997): *Triebel–Lizorkin spaces associated with Laguerre and Hernite expansions*. Proc. Amer. Math. Soc., **125**:3547–3554.

[6] M. FRAZIER, B. JAWERTH, G. WEISS (1991): Littlewood–Paley Theory and the Study of Function Spaces. CBMS, No. 79. Providence, RI: Amer. Math. Soc.

[7] L. HÖRMANDER (1968): *The spectral function of an elliptic operator*. Acta Math., **121**:193–218.

[8] G. KYRIAZIS, P. PETRUSHEV, YUAN XU (to appear): *Jacobi decomposition of weighted Triebel–Lizorkin and Besov spaces*. Preprint. (http://www.math.sc.edu/∼pencho/)

[9] G. KYRIAZIS, P. PETRUSHEV, YUAN XU (to appear): *Decomposition of weighted Triebel–Lizorkin and Besov spaces on the ball*. Preprint. (http://www.math.sc.edu/∼pencho/)

[10] Y. MEYERS (1990): Ondelletes et Opérateurs I: Ondelletes. Paris: Hermann.

[11] H. MHASKAR (2004): *Polynomial operators and local smoothness classes on the unit interval*. J. Approx. Theory, **131**:243–267.

[12] H. MHASKAR, F. NARCOWICH, J. WARD (2001): *Spherical Marcinkiewicz–Zygmund inequalities and positive quadrature*. Math. Comp., **70**:1113–1130.

[13] H. MHASKAR, F. NARCOWICH, J. WARD (2001): *Corrigendum to "Spherical Marcinkiewicz–Zygmund inequalities and positive quadrature"*. Math. Comp., **71**(#237):453–454.

[14] I. P. MYSOVSKIKH (1981): Interpolatory Cubature Formulas. In Russian. Moscow: Nauka.

[15] F. NARCOWICH, P. PETRUSHEV, J. WARD (2006): *Localized tight frames on spheres*. SIAM J. Math. Anal., **38**:347–692.

[16] F. NARCOWICH, J. WARD, P. PETRUSHEV (2006): *Decomposition of Besov and Triebel–Lizorkin spaces on the sphere*. J. Funct. Anal., **238**:530–564.

[17] P. PETRUSHEV, YUAN XU (2005): *Localized polynomial frames on the interval with Jacobi weights*. J. Fourier Anal. Appl., **5**:557–575.

[18] A. STROUD (1971): Approximation Calculation of Multiple Integrals. Englewood Cliffs, NJ: Prentice-Hall.

[19] G. SZEGŐ (1975): Orthogonal Polynomials, 4th ed. Amer. Math. Soc. Colloq. Publ., Vol. 23. Providence, RI: Amer. Math. Soc.

[20] YUAN XU (1998): *Orthogonal polynomials and cubature formulae on spheres and on balls*. SIAM J. Math. Anal., **29**:779–793.

[21] YUAN XU (1999): *Summability of Fourier orthogonal series for Jacobi weight on a ball in* $\mathbf{R}^d$. Trans. Amer. Math. Soc., **351**:2439–2458.

[22] YUAN XU (2005): *Weighted approximation of functions on the unit sphere*. Const. Approx., **21**:1–28.

P. Petrushev
Department of Mathematics
University of South Carolina
Columbia, SC 29208
USA
pencho@math.sc.edu

Yuan Xu
Department of Mathematics
University of Oregon
Eugene, OR 97403
USA
yuan@math.uoregon.edu