Found. Comput. Math. 345–368 (2004)

DOI: 10.1007/s10208-004-0096-3

FOUNDATIONS OF COMPUTATIONAL MATHEMATICS
The Journal of the Society for the Foundations of Computational Mathematics

Whitney Estimates for Convex Domains with Applications to Multivariate Piecewise Polynomial Approximation

S. Dekel[1] and D. Leviatan[2]

[1]RealTimeImage,
6 Hamasger St.
Or-Yehuda 60408, Israel
shai.dekel@turboimage.com

[2]School of Mathematical Sciences,
Tel-Aviv University
Tel-Aviv 69978, Israel
leviatan@math.tau.ac.il

Abstract. We prove the following Whitney estimate. Given $0 < p \leq \infty$, $r \in \mathbb{N}$, and $d \geq 1$, there exists a constant $C(d, r, p)$, depending only on the three parameters, such that for every bounded convex domain $\Omega \subset \mathbb{R}^d$, and each function $f \in L_p(\Omega)$,

$$E_{r-1}(f, \Omega)_p \leq C(d, r, p)\omega_r(f, \operatorname{diam}(\Omega))_p,$$

where $E_{r-1}(f, \Omega)_p$ is the degree of approximation by polynomials of total degree, $r - 1$, and $\omega_r(f, \cdot)_p$ is the modulus of smoothness of order r. Estimates like this can be found in the literature but with constants that depend in an essential way on the geometry of the domain, in particular, the domain is assumed to be a Lipschitz domain and the above constant C depends on the minimal head-angle of the cones associated with the boundary.

The estimates we obtain allow us to extend to the multivariate case, the results on bivariate Skinny B-spaces of Karaivanov and Petrushev on characterizing nonlinear approximation from nested triangulations. In a sense, our results were anticipated by Karaivanov and Petrushev.

Date received: May 12, 2003. Final version received: January 18, 2004. Date accepted: March 3, 2004. Communicated by Arieh Iserles. Online publication: June 8, 2004.
AMS classification: 41A10, 41A25, 41A17, 41A63.
Key words and phrases: Piecewise polynomial approximation, Nonlinear approximation, Whitney estimates, John's theorem.

1. Introduction

We begin by recalling classical smoothness measures over multivariate domains. Here and throughout the paper we assume that domains $\Omega \subset \mathbb{R}^d$ are bounded with a nonempty interior. Let $W_p^r(\Omega)$, $1 \le p \le \infty$, $r \in \mathbb{N}$, denote the **Sobolev spaces**, namely, the spaces of functions $g \in L_p(\Omega)$ which have all their distributional derivatives of order up to r, $D^\alpha g := \partial^k g/\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}$, $\alpha := (\alpha_1, \dots, \alpha_d) \in \mathbb{Z}_+^d$, $|\alpha| := \sum_{i=1}^d \alpha_i = k$, $0 \le k \le r$, are in $L_p(\Omega)$. The kth seminorm, $0 \le k \le r$, is given by $|g|_{k,p} := \sum_{|\alpha|=k} \|D^\alpha g\|_{L_p(\Omega)} < \infty$ and, in particular, the rth seminorm may be regarded as a measure of the smoothness of order r of the function, provided it is in the appropriate Sobolev space. The ***K*-functional of order *r*** of $f \in L_p(\Omega)$, $1 \le p \le \infty$, (see, e.g., [**8**], [**9**]) is defined by

$$K_r(f,t)_p := K(f,t,L_p(\Omega), W_p^r(\Omega)) := \inf_{g \in W_p^r(\Omega)} \{\|f-g\|_p + t|g|_{r,p}\}. \quad (1.1)$$

We denote

$$K_r(f,\Omega)_p := K(f, \operatorname{diam}(\Omega)^r)_p. \quad (1.2)$$

For $f \in L_p(\Omega)$, $0 < p \le \infty$, $h \in \mathbb{R}^d$, and $r \in \mathbb{N}$ we recall the rth-order difference operator $\Delta_h^r(f) : \Omega \to \mathbb{R}$, defined by

$$\Delta_h^r(f,x) := \Delta_h^r(f,\Omega,x) := \begin{cases} \sum_{k=0}^{r} (-1)^{r+k} \binom{r}{k} f(x+kh), & [x, x+rh] \subset \Omega, \\ 0, & \text{otherwise,} \end{cases}$$

where $[x,y]$ denotes the line segment connecting any two points $x, y \in \mathbb{R}^d$. The **modulus of smoothness of order *r*** (see, e.g., [**8**], [**9**]) is defined by

$$\omega_r(f,t)_p := \sup_{|h| \le t} \|\Delta_h^r(f,\Omega,\cdot)\|_{L_p(\Omega)}, \qquad t > 0, \quad (1.3)$$

where for $h \in \mathbb{R}^d$, $|h|$ denotes the norm of h. As above we also denote

$$\omega_r(f,\Omega)_p := \omega_r(f, \operatorname{diam}(\Omega))_p. \quad (1.4)$$

It is known that, for $1 \le p \le \infty$, the above two notions of smoothness, (1.1) and (1.3), are sometimes equivalent. In particular, it is shown in [**14**] that if $\Omega \subset \mathbb{R}^d$ has the uniform cone property (see Definition 2.1 below), then for $1 \le p \le \infty$ and $r \in \mathbb{N}$, there exist $C_1 > 0$ and C_2 such that, for any $f \in L_p(\Omega)$ and $t > 0$,

$$C_1 K_r(f,t^r)_p \le \omega_r(f,t)_p \le C_2 K_r(f,t^r)_p. \quad (1.5)$$

However, while it is easy to show that C_2 in (1.5) depends only on r, the constant C_1 may further depend on the geometry of Ω (see [**14**]). Also, it is important to note that the K-functional is unsuitable as a measure of smoothness if $0 < p < 1$ (see [**11**]).

Let $\Pi_{r-1} := \Pi_{r-1}(\mathbb{R}^d)$ denote the multivariate polynomials of total degree $r-1$ (order r) in d variables. Given $\Omega \subset \mathbb{R}^d$, our initial goal is to estimate the degree of approximation of a function $f \in L_p(\Omega)$, $0 < p \le \infty$,

$$E_{r-1}(f, \Omega)_p := \inf_{P \in \Pi_{r-1}} \|f - P\|_{L_p(\Omega)},$$

using either one of the above notions of smoothness. One of the classical results in this direction is the **Bramble–Hilbert lemma** [**3**]. To introduce it we require the following definitions.

A domain Ω is **star-shaped with respect to a ball** if for each point $x \in \Omega$, the closed convex hull of $\{x\} \cup B$ is contained in Ω. Let $\rho_{\max} = \max\{\rho \mid \Omega$ is star-shaped with respect to a ball $B \subseteq \Omega$ of radius $\rho\}$. The **chunkiness parameter** of Ω is defined by

$$\gamma := \frac{\operatorname{diam}(\Omega)}{\rho_{\max}}. \tag{1.6}$$

This leads to the following formulation of the Bramble–Hilbert lemma (see [**4**, Chapter 4]):

Proposition 1.1. *Let Ω be star-shaped with respect to some ball B and let $g \in W_p^r(\Omega)$, $1 \le p \le \infty$, $r \in \mathbb{N}$. Then there exists a polynomial $P \in \Pi_{r-1}$ for which*

$$|g - P|_{k,p} \le C(d, r, \gamma) \operatorname{diam}(\Omega)^{r-k} |g|_{r,p}, \qquad k = 0, 1, \ldots, r. \tag{1.7}$$

The Bramble–Hilbert lemma implies that for Ω, a star-shaped domain with respect to some ball B, and $f \in L_p(\Omega)$, $1 \le p \le \infty$, we have

$$K_r(f, \Omega)_p \le E_{r-1}(f, \Omega)_p \le C(d, r, \gamma) K_r(f, \Omega)_p. \tag{1.8}$$

By (1.5) if we further assume that the domain has the uniform cone property, then the equivalence

$$E_{r-1}(f, \Omega)_p \approx K_r(f, \Omega)_p \approx \omega_r(f, \Omega)_p, \tag{1.9}$$

holds for $1 \le p \le \infty$ with constants that depend on the shape of the domain Ω.

When $d \ge 2$, the main drawback of (1.7) and (1.8) is that the constant depends on the chunkiness parameter (1.6) which "blows-up," for example, in the case of a sequence of triangles of equivalent diameter that become thinner and thinner. Also (1.9) may further depend on the geometry of the domain. This dependence is too restrictive to be applied in estimates in nonlinear approximation by piecewise polynomials (see [**8**] for a survey on nonlinear approximation). For instance, a problem that is motivated by image compression applications is trying to characterize the degree of nonlinear approximation (see [**15**], [**16**] and [**7**]),

$$\sigma_{n,r}(f)_p := \inf_{S \in \Sigma_n^r} \|f - S\|_{L_p([0,1]^d)},$$

where $f \in L_p([0,1]^d)$ and $\Sigma_n^r := \Sigma_n^r(\mathbb{R}^d)$ is the collection

$$\sum_{k=1}^{n} \mathbf{1}_{\Delta_k} P_k, \tag{1.10}$$

where $\{\Delta_k\}$ are d-simplices such that $\bigcup_{k=1}^n \Delta_k = [0,1]^d$, and $P_k \in \Pi_{r-1}(\mathbb{R}^d)$.

In order to apply (1.9), we would need to assume that all of the simplices $\{\Delta_k\}$ have some sort of uniform geometric properties. In the finite-elements community [**4**], one says that the mesh $\{\Delta_k\}$ is required to be "quasi-uniform." This limitation is in contradiction to the main idea behind piecewise polynomial approximation which is to adaptively place the simplices $\{\Delta_k\}$ over subdomains where the function is smooth. These subdomains could be long and narrow and thus, under the constraint of a "quasi-uniform" mesh, it may take many simplices to cover them.

In [**6**] we have proved the following variant of the Bramble–Hilbert lemma:

Proposition 1.2. *Let $\Omega \subset \mathbb{R}^d$ be a bounded convex domain and let $g \in W_p^r(\Omega)$, $r \in \mathbb{N}$, $1 \le p \le \infty$. Then there exists a polynomial $P \in \Pi_{r-1}$ for which*

$$|g - P|_{k,p} \le C(d,r)\,\mathrm{diam}(\Omega)^{r-k}|g|_{r,p}, \qquad k = 0, 1, \ldots, r. \tag{1.11}$$

A direct consequence of Proposition 1.2 is the following ([**6**]):

Corollary 1.3. *For all convex domains $\Omega \subset \mathbb{R}^d$ and functions $f \in L_p(\Omega)$, $1 \le p \le \infty$,*

$$E_{r-1}(f,\Omega)_p \approx K_r(f,\Omega)_p,$$

where $K_r(f,\Omega)_p$ is defined in (1.2) *and the constants of equivalency depend only on d and r.*

Our main result which we prove in Section 2 is the following Whitney-type theorem that generalizes a result by Karaivanov and Petrushev [**15**] for triangles in the plane and a result by Storozhenko and Oswald [**19**] for multivariate boxes (see also [**5**] for the case of continuous functions).

Theorem 1.4. *Let $\Omega \subset \mathbb{R}^d$ be a bounded convex domain and let $f \in L_p(\Omega)$, $0 < p \le \infty$. Then, for any $r \in \mathbb{N}$,*

$$E_{r-1}(f,\Omega)_p \le C(d,r,p)\omega_r(f,\Omega)_p, \tag{1.12}$$

where $\omega_r(f,\Omega)_p$ is defined in (1.4).

Theorem 1.4 implies that for all bounded convex domains $\Omega \subset \mathbb{R}^d$ and functions $f \in L_p(\Omega)$, if $1 \le p \le \infty$, then we have the equivalence

$$E_{r-1}(f,\Omega)_p \approx K_r(f,\Omega)_p \approx \omega_r(f,\Omega)_p, \tag{1.13}$$

and for $0 < p < 1$, we have the equivalence

$$E_{r-1}(f, \Omega)_p \approx \omega_r(f, \Omega)_p, \tag{1.14}$$

where the constants of equivalency depend only on d, r, and p.

In Section 3 we show how these equivalences can be used in the analysis of piecewise polynomial approximation where arbitrarily long and thin simplices are allowed in (1.10). We prove, in the multivariate case, the results anticipated by Karaivanov and Petrushev [**15**], who introduced bivariate smoothness spaces, the so-called bivariate skinny B-spaces, for the purpose of characterizing the approximation spaces corresponding to nonlinear n-term piecewise polynomials over nested triangulations (see also [**16**] for a survey on such spaces).

2. Local Polynomial Approximation over Convex Domains

In this section we prove the Whitney Theorem 1.4, separately for $1 \le p \le \infty$, and for $0 < p < 1$. In the former case we can use the equivalence (1.5) and the full machinery of K-functionals. In the latter case we have to work harder as the classical K-functional in L_p, $0 < p < 1$, is trivial (see [**11**]), and thus inappropriate for our purposes. We begin by recalling the definitions of domains with the uniform cone property, and of Lipschitz domains (see [**1**, p. 66]).

Definition 2.1. A bounded domain Ω has the **uniform cone property** if there exist numbers $\delta > 0$, $L > 0$, a finite cover of open sets $\{U_j\}_{j=1}^J$ of $\partial\Omega$, and a corresponding collection $\{V_j\}_{j=1}^J$ of finite cones, each congruent to some fixed cone V, such that:

(i) $\operatorname{diam}(U_j) \le L$, $1 \le j \le J$.
(ii) For any $x \in \Omega$ such that $\operatorname{dist}(x, \partial\Omega) < \delta$, we have $x \in \bigcup_{j=1}^J U_j$.
(iii) For every j, $\bigcup_{x \in \Omega \cap U_j} (x + V_j) \subseteq \Omega$.

Definition 2.2. A bounded domain Ω is called a **Lipschitz domain** if there exist numbers $\delta > 0$, $M > 0$, and a finite cover of open sets $\{U_j\}_{j=1}^J$ of $\partial\Omega$ such that:

(i) For every pair of points $x_1, x_2 \in \Omega$ such that $|x_1 - x_2| < \delta$ and $\operatorname{dist}(x_i, \partial\Omega) < \delta$, $i = 1, 2$, there exists an index j such that $x_i \in U_j$, $i = 1, 2$, and $\operatorname{dist}(x_i, \partial U_j) > \delta$, $i = 1, 2$.
(ii) For each j there exists some Cartesian coordinate system $(\xi_{j,1}, \dots, \xi_{j,d})$ in U_j, such that the set $\Omega \cap U_j$ can be represented by the inequality $\xi_{j,d} \le f_j(\xi_{j,1}, \dots, \xi_{j,d-1})$, where $f_j : \mathbb{R}^{d-1} \to \mathbb{R}$ satisfies a Lipschitz condition with constant M, namely,

$$|f_j(x) - f_j(y)| \le M|x - y|, \qquad x, y \in \mathbb{R}^{d-1}.$$

It is known that a Lipschitz domain has the uniform cone property (see [**1**, p. 66]). Later, in the proof of Lemma 2.9, which is one of the crucial lemmas of this paper, we find it expedient to apply the uniform cone property. However, in order to show that the absolute constants we obtain in the Whitney estimates are valid uniformly for all bounded convex domains, we find it more convenient, in Lemma 2.3, to obtain estimates on the Lipschitz constants of the domain which depend solely on certain given parameters.

For $x \in \mathbb{R}^d$ let $B(x, \rho) := B_d(x, \rho)$ denote the ball in $\mathbb{R}^d$, of radius ρ with center at x.

Lemma 2.3. *Let $\Omega \subset \mathbb{R}^d$ be a convex domain such that $B(0, 1) \subseteq \Omega \subseteq B(0, R)$ for some $R > 1$. Then Ω is a Lipschitz domain with $\delta = \delta(d, R)$, $M := M(d, R)$, and an open cover $\{U_j\}_{j=1}^{J}$, $J = J(d, R)$, that are fixed for all such domains.*

Proof. For each d and R, we construct a finite set of overlapping finite cylinders U_j that cover the ring $\overline{B(0, R)\backslash B(0, 1)}$ and therefore cover the boundary of any convex domain $B(0, 1) \subseteq \Omega \subseteq B(0, R)$ (see Figure 2.1 for illustration). Indeed, we cover the surface of the ball of radius R, $S^{d-1}(0, R)$, by overlapping $(d - 1)$-dimensional open balls of small enough radius ε, in such a way that the boundary of each such ball is completely covered by the adjacent balls. If $\mathbf{n}$ denotes the unit vector emanating from the origin in the direction of the center of such a ball, then $U := \{x = t\mathbf{n} + y \mid t \geq 0, y \in B_{d-1}(0, \varepsilon)\}$ is the cylinder of radius ε whose intersection with $S^{d-1}(0, R)$ is the boundary of this ball. The collection of these cylinders we denote by $\{U_j\}$. Then property (i) is evident, and for each j, the coordinate system in condition (ii) is simply the map that transforms the cylinder U_j to align with the direction of e_d, where e_d is the last vector of the standard basis of $\mathbb{R}^d$. Now, for any $x \in \partial\Omega \cap U_j$, the cone defined by the convex closure of $\{x\} \cup B(0, 1)$ is contained in the closure of Ω. Any head-angle α of this cone satisfies $\sin(\alpha/2) \geq 1/R$ and thus the boundary is a Lip 1 function with $M := M(d, R)$. □

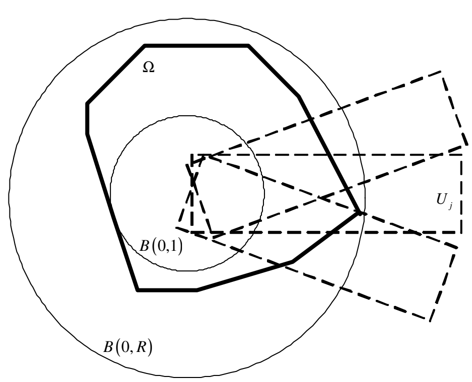

Fig. 2.1. The cover of $\partial\Omega$ by cylinders $\{U_j\}$.

The equivalence between the K-functional and the modulus of smoothness may further depend on the geometry of the domain. However, one can provide uniform equivalency constants for a class of domains that are of the same "Lipschitz-type."

Lemma 2.4. *Let $\Omega \subset \mathbb{R}^d$ fulfill the conditions of Lemma* 2.3 *for some $R > 1$. Then, for $f \in L_p(\Omega)$, $1 \le p \le \infty$, $r \in \mathbb{N}$, and $t \in (0,1)$,*

$$C_1(d,r,p,R)K_r(f,t^r)_{L_p(\Omega)} \le \omega_r(f,t)_{L_p(\Omega)} \le C_2(r)K_r(f,t^r)_{L_p(\Omega)}. \qquad (2.1)$$

Proof. As already mentioned in the Introduction, the right-hand side of (2.1) holds for arbitrary domains. The left-hand side inequality is almost immediate from [**14**, Theorem 1], where it is proved for domains with the uniform cone property, and where the constant C_1 depends on r and p, and on the uniform cone properties of Ω. Since Lemma 2.3 implies that the Lipschitz properties of Ω only depend on d and R, and the Lipschitz property implies the uniform cone property, we get that $C_1 = C_1(d,r,p,R)$ in (2.1). □

Recall that an **ellipsoid** E is the image of the closed unit ball in $\mathbb{R}^d$ under a nonsingular affine map $A(x) = Mx + b$, $M \in M_{d\times d}(\mathbb{R})$, $b \in \mathbb{R}^d$. The **center** of E is $b = A(0)$. The following result by Fritz John [**13**] (see also [**2**]) is an important tool in this work:

Proposition 2.5 (John's theorem). *Let $\Omega \subset \mathbb{R}^d$ be a bounded convex domain. Then there exists an ellipsoid $E \subseteq \Omega$ so that if x_0 is the center of E, then the inclusions*

$$E \subseteq \Omega \subseteq x_0 + d(E - x_0).$$

hold. Here $x_0 + d(E - x_0) := \{z \mid z = x_0 + d(x - x_0), x \in E\}$.

John's theorem implies that for each convex domain Ω one can find a nonsingular affine map A such that

$$B(0,1) \subseteq \widetilde{\Omega} := A^{-1}(\Omega) \subseteq B(0,d). \qquad (2.2)$$

It is interesting to note that John's ellipsoid is the ellipsoid $E \subseteq \Omega$ with maximal volume. In some sense this means that it "covers" Ω sufficiently well.

We are ready to prove Theorem 1.4 for $1 \le p \le \infty$.

Proof of Theorem 1.4 *for the Case* $1 \le p \le \infty$. For a bounded convex domain Ω, let $A(x) = Mx + b$ be the affine transform for which (2.2) holds. Corollary 1.3 implies that for $\widetilde{\Omega} := A^{-1}(\Omega)$ and $\widetilde{f} := f(A\,\cdot)$ there exists a polynomial $\widetilde{P} \in \Pi_{r-l}(\mathbb{R}^d)$ such that

$$\|\widetilde{f} - \widetilde{P}\|_{L_p(\widetilde{\Omega})} \le C(d,r)K_r(\widetilde{f},\widetilde{\Omega})_p.$$

Since the domain $\widetilde{\Omega}$ fulfills the conditions of Lemma 2.3, we may further apply Lemma 2.4 to obtain

$$\begin{aligned}\|\widetilde{f}-\widetilde{P}\|_{L_p(\widetilde{\Omega})} &\le C(d,r)K_r(\widetilde{f},\widetilde{\Omega})_p\\ &\le C(d,r,p)\omega_r(\widetilde{f},\widetilde{\Omega})_p.\end{aligned}$$

Denoting $P := \widetilde{P}(A^{-1}\cdot)$ yields

$$\begin{aligned}\|f-P\|_{L_p(\Omega)} &= |\det M|^{1/p}\|\widetilde{f}-\widetilde{P}\|_{L_p(\widetilde{\Omega})}\\ &\le C|\det M|^{1/p}\omega_r(\widetilde{f},\widetilde{\Omega})_p\\ &= C\omega_r(f,\Omega)_p.\end{aligned}$$

This shows that for $1 \le p \le \infty$, $E_{r-1}(f,\Omega)_p \le C(d,r,p)\omega_r(f,\Omega)_p$. □

We now turn to the proof of Theorem 1.4 for the case $0 < p < 1$. We follow the method to prove the univariate Whitney estimate, which was used in [**17**, Section 7.1] (see also [**9**, Section 12.5]), but we apply it in the multivariate setting. We first prove the case $r = 1$ (compare with [**17**, Lemma 7.6]).

Lemma 2.6. *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain and let $f \in L_p(\Omega)$, $0 < p < \infty$. Then there exists a constant c such that*

$$\int_\Omega |f(x)-c|^p\,dx \le \frac{1}{|\Omega|}\int_{|h|\le \operatorname{diam}(\Omega)}\int_\Omega |\Delta_h(f,\Omega,x)|^p\,dx\,dh, \tag{2.3}$$

where $|\Omega|$ denotes the volume of the domain Ω.

Proof. Consider the function $\varphi(y) := \int_\Omega |f(x)-f(y)|^p\,dx$, $y \in \Omega$. Clearly, there exists $y_0 \in \Omega$ such that

$$\varphi(y_0) \le \frac{1}{|\Omega|}\int_\Omega \varphi(y)\,dy.$$

Therefore with $c := f(y_0)$ we get

$$\begin{aligned}\int_\Omega |f(x)-c|^p\,dx &= \varphi(y_0)\\ &\le \frac{1}{|\Omega|}\int_\Omega \varphi(y)\,dy\\ &= \frac{1}{|\Omega|}\int_\Omega\int_\Omega |f(x)-f(y)|^p\,dx\,dy.\end{aligned}$$

By definition, for any domain Ω and every $x \in \Omega$, if $x+h \notin \Omega$, then $\Delta_h(f,\Omega,x) = 0$. Therefore, the substitution $h = y - x$ yields (2.3). □

Corollary 2.7. *Let $\Omega \subset \mathbb{R}^d$ be a bounded convex domain and let $f \in L_p(\Omega)$, $0 < p < \infty$. Then there exists a constant c such that*

$$\|f - c\|_{L_p(\Omega)} \leq (2d)^{d/p}\omega_1(f, \Omega)_p. \tag{2.4}$$

Proof. Let $\widetilde{\Omega} := A^{-1}(\Omega)$ be the convex domain for which (2.2) holds and let $\widetilde{f} := f(A\,\cdot)$. By Lemma 2.6 there exists a constant c such that

$$\int_{\widetilde{\Omega}} |\widetilde{f}(x) - c|^p\,dx \leq \frac{1}{|\widetilde{\Omega}|}\int_{|h|\leq 2d}\int_{\widetilde{\Omega}} |\Delta_h(\widetilde{f}, \widetilde{\Omega}, x)|^p\,dx\,dh.$$

Hence

$$\begin{aligned}\int_{\widetilde{\Omega}} |\widetilde{f}(x) - c|^p\,dx &\leq \frac{|B(0, 2d)|}{|B(0, 1)|}\omega_1(\widetilde{f}, \widetilde{\Omega})_p^p \\ &= (2d)^d\omega_1(\widetilde{f}, \widetilde{\Omega})_p^p. \end{aligned} \tag{2.5}$$

As we have seen in the proof of Theorem 1.4 for the case $1 \leq p \leq \infty$, the Whitney inequality is invariant under affine maps and therefore (2.5) implies (2.4). □

Next we require the following piecewise constant approximation estimate which is similar to the univariate estimate in [**17**, Lemma 7.7]):

Lemma 2.8. *Let $\Omega \subset \mathbb{R}^d$ be a convex domain such that $B(0, 1) \subseteq \Omega \subseteq B(0, d)$ and let $f \in L_p(\Omega)$, $0 < p < \infty$. Then, for each $n \in \mathbb{N}$, there exists a step function*

$$\varphi = \sum_{k=1}^{K} \mathbf{1}_{Q_k} c_k,$$

with the following properties:

(1) *Q_k, $1 \leq k \leq K \leq C_1(d)n^d$, are cubes taken from the uniform grid of side length n^{-1} and thus have disjoint interiors*;
(2) $\Omega \subseteq \bigcup_{k=1}^{K} Q_k$;
(3) $\|f - \varphi\|_{L_p(\Omega)} \leq C(d, p)\omega_1(f, 1/n)_{L_p(\Omega)}$;
(4) $\|\varphi\|_{L_p(\mathbb{R}^d)} \leq C(d, p)\|f\|_{L_p(\Omega)}$.

Proof. For $n \in \mathbb{N}$, we select from the uniform grid of length n^{-1} all the cubes Q_k, $1 \leq k \leq \widetilde{K} \leq (2d)^d n^d$, for which $\mathrm{int}(Q_k \cap \Omega) \neq \emptyset$. For each $1 \leq k \leq \widetilde{K}$, we construct from Q_k, by a symmetric extension, the cube $\widetilde{Q}_k$ with side-length $3n^{-1}$. We claim that there exists a constant $C_2(d)$ such that

$$|\widetilde{Q}_k \cap \Omega| \geq C_2(d)n^{-d}, \qquad 1 \leq k \leq \widetilde{K}. \tag{2.6}$$

Indeed, given $1 \leq k \leq \widetilde{K}$, take a point $x_0 \in Q_k \cap \Omega$. If $x_0 \in B(0, 1)$, then it is easy to see that there exists a constant $C_3(d)$ for which

$$|\Omega \cap \widetilde{Q}_k| \geq |B(0, 1) \cap \widetilde{Q}_k| \geq C_3(d)n^{-d}.$$

Otherwise, $x_0 \notin B(0,1)$, and we denote by $V(x_0)$ the cone defined by the convex closure of the set $\{x_0\} \cup B(0,1) \subseteq \Omega$. From the properties of the domain Ω, it follows that the head angle α of the cone satisfies $\sin(\alpha/2) \geq 1/d$. Therefore, as the cone $V(x_0)$ is not too "thin," there exists a constant $C_4(d)$ such that

$$|\Omega \cap \widetilde{Q}_k| \geq |V(x_0) \cap \widetilde{Q}_k| \geq C_4(d) n^{-d}.$$

We conclude that (2.6) holds with $C_2 := \min(C_3, C_4)$.

Next we augment cubes Q_k, $\widetilde{K} < k \leq K \leq C_1(d) n^d$, taken from the uniform grid, to ensure that $\bigcup_{k=1}^{K} Q_k = \bigcup_{k=1}^{\widetilde{K}} \widetilde{Q}_k$.

We first assume that $f \geq 0$, and we take $0 < p \leq 1$. Lemma 2.6 implies that for each $1 \leq j \leq \widetilde{K}$ there exists a constant $\widetilde{c}_j$ that satisfies

$$\int_{\widetilde{Q}_j \cap \Omega} |f(x) - \widetilde{c}_j|^p \, dx \leq \frac{1}{|\widetilde{Q}_j \cap \Omega|} \int_{|h| \leq 3\sqrt{d} n^{-1}} \int_{\Omega} |\Delta_h(f, \widetilde{Q}_j \cap \Omega, x)|^p \, dx \, dh.$$

We denote by $\{\widetilde{Q}_{j,k} : 1 \leq j \leq J(k) \leq 3^d\}$, the collection of larger cubes that contain the cube Q_k, $1 \leq k \leq K$, and set

$$c_k := \frac{1}{J(k)} \sum_{j=1}^{J(k)} \tilde{c}_{j,k}.$$

If $\varphi := \sum_{k=1}^{K} \mathbf{1}_{Q_k} c_k$, then

$$\begin{aligned}
\|f - \varphi\|_{L_p(\Omega)}^p &= \sum_{k=1}^{\widetilde{K}} \int_{Q_k \cap \Omega} |f(x) - c_k|^p \, dx \\
&= \sum_{k=1}^{\widetilde{K}} \int_{Q_k \cap \Omega} \left| \frac{1}{J(k)} \sum_{j=1}^{J(k)} (f(x) - \tilde{c}_{j,k}) \right|^p dx \\
&\leq \sum_{k=1}^{\widetilde{K}} \sum_{j=1}^{J(k)} \int_{Q_k \cap \Omega} |f(x) - \tilde{c}_{j,k}|^p \, dx \\
&= \sum_{j=1}^{\widetilde{K}} \int_{\widetilde{Q}_j \cap \Omega} |f(x) - \tilde{c}_j|^p \, dx \\
&\leq C \sum_{j=1}^{\widetilde{K}} \frac{1}{|\widetilde{Q}_j \cap \Omega|} \int_{|h| \leq 3\sqrt{d} n^{-1}} \int_{\widetilde{Q}_j \cap \Omega} |\Delta_h(f, \widetilde{Q}_j \cap \Omega, x)|^p \, dx \\
&\leq C n^d \sum_{k=1}^{\widetilde{K}} \int_{|h| \leq 3\sqrt{d} n^{-1}} \int_{Q_k \cap \Omega} |\Delta_h(f, \Omega, x)|^p \, dx \\
&= C n^d \int_{|h| \leq 3\sqrt{d} n^{-1}} \int_{\Omega} |\Delta_h(f, \Omega, x)|^p \, dx
\end{aligned}$$

$$\le C\omega_1(f, 3\sqrt{d}/n)^p_{L_p(\Omega)}$$
$$\le C(d, r, p)\omega_1(f, 1/n)^p_{L_p(\Omega)},$$

where we have used

$$\omega_r(f, \lambda t)_p \le C(r, p)(\lambda + 1)^{r-1+1/p}\omega_r(f, t)_p, \qquad \lambda > 0. \tag{2.7}$$

This proves (3). In order to prove property (4), we note that since we assumed that $f \ge 0$, it follows from the proof of Lemma 2.6 that we may take $\tilde{c}_j \ge 0$, $1 \le j \le \widetilde{K}$, and hence that $c_k \ge 0$, $1 \le k \le K$. Applying (2.6) yields

$$\begin{aligned}
\|\varphi\|^p_{L_p(\mathbb{R}^d)} &= n^{-d}\sum_{j=1}^{K} c_k^p \\
&= n^{-d}\sum_{j=1}^{K}\left(\frac{1}{J(k)}\sum_{j=1}^{j(k)}\tilde{c}_{j,k}\right)^p \\
&\le C\sum_{j=1}^{K}\sum_{j=1}^{J(k)}\tilde{c}^p_{j,k}|\widetilde{Q_{j,k}}\cap\Omega| \\
&\le C\sum_{j=1}^{\widetilde{K}}\tilde{c}^p_j|\widetilde{Q}_j\cap\Omega|.
\end{aligned}$$

Using the norm equivalence of finite-dimensional spaces we may proceed with

$$\begin{aligned}
\sum_{j=1}^{\widetilde{K}}\tilde{c}^p_j|\widetilde{Q}_j\cap\Omega| &= \sum_{k=1}^{\widetilde{K}}\left(\sum_{j=1}^{J(k)}\tilde{c}^p_{k,j}\right)|Q_k\cap\Omega| \\
&\le C(d, p)\sum_{k=1}^{\widetilde{K}}\int_{Q_k\cap\Omega} c_k^p\,dx \\
&= C(d, p)\|\varphi\|^p_{L_p(\Omega)} \\
&\le C(d, p)\left(\|f\|^p_{L_p(\Omega)} + \|f-\varphi\|^p_{L_p(\Omega)}\right) \\
&\le C(d, p)\left(\|f\|^p_{L_p(\Omega)} + \omega_1(f, 1/n)^p_{L_p(\Omega)}\right) \\
&\le C(d, p)\|f\|^p_{L_p(\Omega)}.
\end{aligned}$$

The proof of the case $1 \le p < \infty$ is similar, and this completes the proof of (4) for nonnegative functions.

For an arbitrary function $f \in L_p(\Omega)$, $0 < p < \infty$, we use the representation $f = f_+ - f_-$ where $f_+, f_- \ge 0$. Using the above method, we construct approximating step functions φ_1, φ_2 such that

$$\|f_+ - \varphi_1\|_{L_p(\Omega)} \le C\omega_1(f_+, 1/n)_p, \qquad \|f_- - \varphi_2\|_{L_p(\Omega)} \le C\omega_1(f_-, 1/n)_p,$$

and

$$\|\varphi_1\|_{L_p(\mathbb{R}^d)} \le C\|f_+\|_{L_p(\Omega)}, \qquad \|\varphi_2\|_{L_p(\mathbb{R}^d)} \le C\|f_-\|_{L_p(\Omega)}.$$

Recalling that $\omega_1(f_\pm\cdot)_p \le \omega_1(f,\cdot)_p$ and $\|f_\pm\|_{L_p(\Omega)} \le \|f\|_{L_p(\Omega)}$, we conclude that the step function $\varphi := \varphi_1 - \varphi_2$ fulfills the properties (1)–(4). □

We have not been able to find a reference to a multidimensional Marchaud inequality for $0 < p < 1$, on Lipschitz domains. The only known results related to this that we are aware of are Marchaud inequalities for the unit cube in $\mathbb{R}^d$ (see [**10**], [**12**]). Thus we give here a proof of a particular case of the Marchaud inequality that suffices for our purposes. Namely, we have

Lemma 2.9. *Let Ω be a bounded Lipschitz domain and let $f \in L_p(\Omega)$, $0 < p < \infty$. Then for any $r \ge 2$, there exists $0 < t_0 < 1$ such that, for $0 < t \le t_0$,*

$$\omega_1(f,t)_p^p \;\le\; C_1 t^p \left(\int_t^1 u^{-p}\omega_r(f,u)_p^p \frac{du}{u} + \|f\|_p^p\right), \qquad 0 < p < 1,$$

$$\omega_1(f,t)_p \;\le\; C_2 t \left(\int_t^1 u^{-1}\omega_r(f,u)_p \frac{du}{u} + \|f\|_p\right), \qquad 1 \le p < \infty, \quad (2.8)$$

where the constants t_0, C_1, and C_2 depend on d, p, r, and the Lipschitz properties of Ω.

Proof. The case $1 \le p < \infty$ is well-known (see [**14**]), so we shall prove the case $0 < p < 1$. From the Lipschitz properties of Ω it is easy to see that we may augment the cover of $\partial\Omega$ (see Definition 2.1) to a finite open cover $\{U_j\}$ of Ω, with corresponding finite cones $\{V_j\}$, each congruent to a fixed finite cone V, such that for each $x \in U_j$, we have $x + V_j \subset \Omega$. Furthermore, there exists $t_0 > 0$, such that if $x, y \in \Omega$ with $|x - y| \le t_0$, then $x, y \in U_j$ for some j, the cones $x + V_j$ and $y + V_j$ intersect, and there is a $z \in (x+V_j)\cap(y+V_j)$ such that $|x-z|, |y-z| \le C|x-y|$, where C depends only on the head-angle of V. Moreover, we can choose that z so that the length of the intersection of the ray from x through z with Ω is proportional to diam(Ω), and the same for y, with the proportion depending on the Lipschitz properties. Indeed, if either x or y is in the cone whose vertex is the other, say, $y \in x + V_j$, then we take $z := y$. Otherwise, we take for z the vertex of the cone $(x + V_j) \cap (y + V_j)$. Clearly in both cases z satisfies the above requirements (for illustration, see Fig. 2.2).

Let $0 < t \le t_0$, fix $h \in \mathbb{R}^d$ with $|h| \le t$. Since the direction of h is arbitrary, if $x + h \in \Omega$, we cannot simply connect x and $x + h$ by a segment and proceed, as in the proof below to $x + mh$, to sufficiently large m, because we clearly may get out of Ω too soon. Thus, we proceed differently and denote

$$U_{h,j} := \{x \in \Omega : x + h \notin \Omega \text{ or } x, x + h \in U_j\}.$$

It follows from the discussion above that there exist two unit vectors $h_{i,1}$ and $h_{j,2}$ such that if $x, x + h \in U_j$, then:

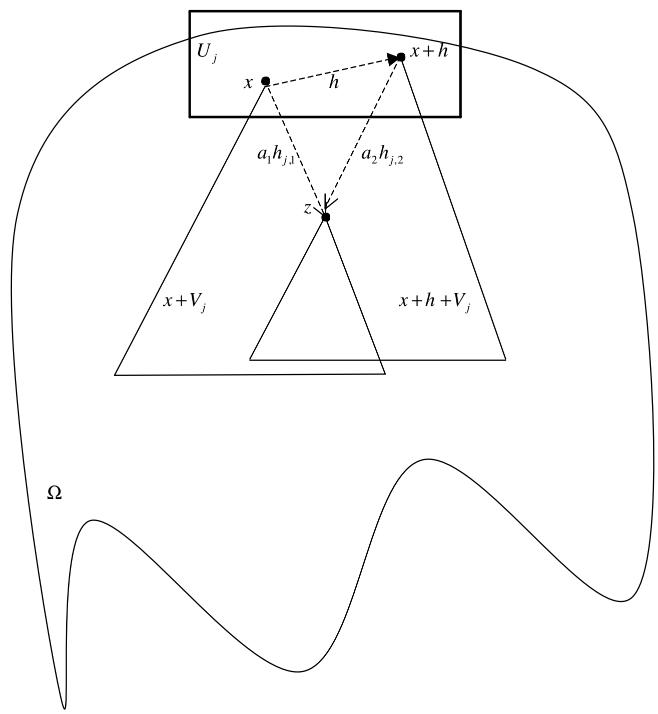

Fig. 2.2.

(i) $h = a_1 h_{j,1} - a_2 h_{j,2}$ with $0 \le a_1, a_2 \le C|h|$, where C depends only on the head-angle of V;

(ii) $[x, x + a_1 h_{j,1}] \subset x + V_j$ and $[x + h, x + h + a_2 h_{j,2}] \subset x + h + V_j$.

For $k \ge 1$ and a unit vector $H \in \mathbb{R}^d$ we denote

$$\omega_k^H(f,t)_p^p := \sup_{0<s\le t} \int_{U_{h,j}} |\Delta_{sH}^k(f,\Omega,x)|^p \, dx.$$

Also we let

$$\omega_k^{j,l}(f,t)_p := \omega_k^{h_{j,l}}(f,t)_p, \qquad l = 1, 2.$$

Then we clearly have

$$\int_{U_{h,j}} |\Delta_h(f,\Omega,x)|^p \, dx \le C(\omega_1^{j,1}(f,t)_p^p + \omega_1^{j,2}(f,t)_p^p). \tag{2.9}$$

Note that an inequality like (2.9) is invalid for higher order moduli, and is the reason why we concentrate on proving only the particular case of the Marchaud inequality.

The proof of the Marchaud inequality for each of the components $\omega_1^{j,l}(f,t)_p^p$, $1 \le l \le 2$, respectively, along the lines containing the segment $[x, x + a_1 h_{j,1}]$ and

$[x+h, x+h+a_2h_{j,2}]$, namely,

$$\omega_1^{j,l}(f,t)_p^p \le Ct^p\left(\int_t^1 u^{-p}\omega_r^{j,l}(f,u)_p^p\,\frac{du}{u} + \|f\|^p_{L_p(U_{h,j})}\right), \qquad l = 1,2, \tag{2.10}$$

now follows the proof of the univariate Marchaud inequality (see, e.g., [**9**, (2.8.3), (2.8.5), and Proof of Theorem 8.1]). We remark that we do not need to worry about (a fixed number of) the translates of any x in $U_{j,h}$, by multiples of H getting out of Ω. Combining (2.9) and (2.10) yields

$$\|\Delta_h(f,\Omega,x)\|^p_{L_p(U_{h,j})} \le Ct^p\left(\int_t^1 u^{-p}\omega_r(f,\Omega,u)_p^p\,\frac{du}{u} + \|f\|^p_{L_p(\Omega)}\right). \tag{2.11}$$

Finally, (2.8) is obtained by summing (2.11) over all $U_{h,j}$, and then taking supremum on h. We note that we used again the relation (2.7). □

Combining Lemmas 2.3 and 2.9 we obtain:

Corollary 2.10. *Let $\Omega \subset \mathbb{R}^d$ be a convex domain such that $B(0,1) \subseteq \Omega \subseteq B(0,d)$, and let $f \in L_p(\Omega)$, $0 < p < \infty$. Then* (2.8) *holds with constants that depend only on d, p, and r.*

Definition 2.11. Let $\Omega \subset \mathbb{R}^d$ be a bounded convex domain containing the origin. Then we denote by $\varphi_\Omega \in C(\mathbb{T}^{d-1})$, where $\mathbb{T}^{d-1}$ is the $(d-1)$-dimensional torus, the unique continuous function that describes $\partial\Omega$. Namely, $\varphi_\Omega(\theta) = r$ if and only if (r,θ) is the unique point in $\mathbb{R}^d$, in polar representation, for which $(r,\theta) \in \partial\Omega$. Observe that the norm $C(\mathbb{T}^{d-1})$ induces a metric on the collection of such domains.

Lemma 2.12. *Let $\{\Omega_m\}_{m\ge1}$ be convex domains in $\mathbb{R}^d$ such that $B(0,1) \subseteq \Omega_m \subseteq B(0,R)$, for some $R > 0$. Then there exists a subsequence $\{\Omega_{m_i}\}_{i\ge1}$ that converges in the sense of Definition* 2.11 *to a convex domain Ω, such that $B(0,1) \subseteq \Omega \subseteq B(0,R)$.*

Proof. Let $\varphi_{\Omega_m}(\theta)$, $m \ge 1$, $\theta \in \mathbb{T}^{d-1}$, be the corresponding continuous function that describes the boundary of Ω_m. A similar argument to the one used in Lemma 2.3 shows that all the functions $\{\varphi_{\Omega_m}\}$ are uniformly bounded in the Lip-1 norm, with a constant $M := M(d,R)$. By the Arzela–Ascoli theorem there exists a convergent subsequence, say to a function φ. It is easy to verify that the function φ describes the boundary of a convex domain Ω, with $B(0,1) \subseteq \Omega \subseteq B(0,R)$. □

The following is a generalization of [**17**, Lemma 7.8] (see also [**9**, Theorem 12.5.3]):

Lemma 2.13. *Let $\Omega \subset \mathbb{R}^d$ be a Lipschitz domain, and let $f \in L_p(\Omega)$, $0 < p \le \infty$, be such that $\omega_r(f, \Omega)_p = 0$ for some $r \in \mathbb{N}$. Then there exists a polynomial $P \in \Pi_{r-1}(\mathbb{R}^d)$ such that $f = P$ a.e. on Ω.*

Proof. Since we already proved (1.12) for the case $1 \le p \le \infty$, we may apply it to conclude that $E_{r-1}(f, \Omega)_p = 0$ and hence that f is a.e. a polynomial. To obtain the result for $0 < p < 1$ one partitions Ω to a countable number of overlapping boxes $\{Q_k\}_{k\ge1}$, $Q_k \subseteq \Omega$, such that $\Omega = \bigcup_{k=1}^{\infty} Q_k$. This can be done, for example, by taking a decomposition of Ω into interior disjoint dyadic cubes (see, e.g., [**18**, p. 167]), and then extending the cubes to boxes so that they overlap with at least one of their neighbors while remain contained in Ω. Since for $0 < p < 1$, (1.12) is known for multivariate boxes [**19**], the inequality $\omega_r(f, Q_k)_p \le \omega_r(f, \Omega)_p = 0$, $k \ge 1$, implies that f is a.e. a polynomial P_k on each Q_k. For any two overlapping boxes Q_i, Q_j we have that $P_i = P_j$ a.e. on $Q_i \cap Q_j$, which implies $P_i = P_j$. Therefore there exists a polynomial $P \in \Pi_{r-1}(\mathbb{R}^d)$ such that $f = P$ a.e. on Ω. □

Proof of Theorem 1.4 *for the case* $0 < p < 1$. We generally follow the proof of [**17**, Theorem 7.1] (see also [**9**, Theorem 12.5.5]). The estimate (2.4) is (1.12) for $r = 1$, so assume to the contrary that, for some $r > 1$ and fixed parameters d and p, there does not exist a constant $C(d, r, p)$ for which (1.12) holds for all bounded convex domains $\Omega \subset \mathbb{R}^d$ and functions $f \in L_p(\Omega)$. In view of the invariance of the Whitney estimate under affine maps, and by John's theorem, this implies the existence of a sequence of convex domains $\{\Omega_m\}_{m\ge1}$, $B(0, 1) \subseteq \Omega_m \subseteq B(0, d)$, and functions $f_m \in L_p(\Omega_m)$ for which

$$E_{r-1}(f_m, \Omega_m)_p^p > m\omega_r(f_m, \Omega_m)_p^p, \qquad m \ge 1.$$

By Lemma 2.12 we may assume that $\{\Omega_m\}_{m\ge1}$ converges to a convex domain Ω, $B(0, 1) \subseteq \Omega \subseteq B(0, d)$, in the sense of Definition 2.11. Furthermore, we may assume that $\Omega_m \subseteq \Omega$, $m \ge 1$. Indeed, let $\varepsilon_k \downarrow 0$, then there exist $m_k \uparrow \infty$ such that

$$B\left(0, \tfrac{1}{2}\right) \subseteq \widetilde{\Omega}_{m_k} := (1 - \varepsilon_k)\Omega_{m_k} \subseteq \Omega \subseteq B(0, d).$$

Hence, for the functions $\widetilde{f}_{m_k} := (1 - \varepsilon_k)^{-d/p} f_{m_k}\left((1 - \varepsilon_k)^{-1} \cdot\right)$, we have

$$\begin{aligned} E_{r-1}(\widetilde{f}_{m_k}, \widetilde{\Omega}_{m_k})_p^p &= E_{r-1}(f_{m_k}, \Omega_{m_k})^p \\ &> m_k\omega_r(f_{m_k}, \Omega_{m_k})_p^p \\ &= m_k\omega_r(\widetilde{f}_{m_k}, \widetilde{\Omega}_{m_k})_p^p. \end{aligned}$$

Clearly $\{\widetilde{\Omega}_{m_k}\}_{k\ge1}$ converges to Ω in the sense of Definition 2.11. Thus, we are justified in our assumption. Therefore, we let $P_m \in \Pi_{r-1}(\mathbb{R}^d)$ be the best approximation to f_m on Ω_m, i.e.,

$$\|f_m - P_m\|_{L_p(\Omega_m)}^p = E_{r-1}(f_m, \Omega_m)_p^p > m\omega_r(f_m, \Omega_m)_p^p.$$

Setting $g_m := \lambda_m(f_m - P_m)$, with λ_m defined by $\|g_m\|_{L_p(\Omega_m)} = 1$, we have a sequence of domains $\{\Omega_m\}_{m\geq 1}$ and functions $\{g_m\}_{m\geq 1}$ with the properties:

(i) $\|g_m\|_{L_p(\Omega_m)} = E_{r-1}(g_m, \Omega_m)_p = 1$;
(ii) $\omega_r(g_m, \Omega_m)_p^p \leq 1/m$;
(iii) $B(0, 1) \subseteq \Omega_m \subseteq \Omega$, and $\{\Omega_m\}$ converges to Ω in the sense of Definition 2.11.

By Corollary 2.10, the Marchaud inequality (2.8) holds with a uniform constant for all the above domains $\{\Omega_m\}$. Thus, for sufficiently small $0 < \delta < 1$ and $m \geq 1$ we get, from property (ii) above,

$$\begin{aligned}\omega_1(g_m, \delta)_{L_p(\Omega_m)}^p &\leq C(d, r, p)\delta^p \left(\int_\delta^1 u^{-p}\frac{1}{m}\frac{du}{u} + 1\right)\\ &\leq C(d, r, p)\left(\frac{1}{m} + \delta^p\right).\end{aligned}$$

It follows that for each $\varepsilon > 0$ there exist δ_0 and m_0 such that

$$\omega_1(g_m, \delta)_{L_p(\Omega_m)}^p \leq \varepsilon \qquad \text{for} \quad \delta < \delta_0 \quad \text{and} \quad m \geq m_0.$$

Applying Lemma 2.8 we get that for any $\varepsilon > 0$ there exist functions $\varphi_{m,n}$, $m \geq m_0$, $n := n(\varepsilon)$, that are piecewise constant over the grid of length n^{-1} and for which

$$\|g_m - \varphi_{m,n}\|_{L_p(\Omega_m)}^p \leq C\omega_1(g_m, n^{-1})_{L_p(\Omega_m)}^p \leq \varepsilon, \qquad m \geq m_0(\varepsilon). \tag{2.12}$$

Lemma 2.8(4) and property (i) above yield

$$\|\varphi_{m,n}\|_{L_p(\mathbb{R}^d)}^p \leq C(d, p). \tag{2.13}$$

Since $\varphi_{m,n}$ is constant over the cubes of side length n^{-1} we have, by (2.13),

$$\begin{aligned}\|\varphi_{m,n}\|_{L_\infty(\Omega)} &\leq C\left(n^d \int_\Omega |\varphi_{m,n}(x)|^p\,dx\right)^{1/p}\\ &\leq Cn^{d/p} =: M.\end{aligned}$$

Consider the set $\Phi := \Phi(\varepsilon)$ of all step functions over the uniform grid of side length n^{-1} that take the values

$$k\varepsilon^{1/p}|B(0, d)|^{-1/p}, \qquad k = 0, \pm 1, \ldots, \pm\lceil \varepsilon^{-1/p}|B(0, d)|^{1/p}M\rceil,$$

Clearly,

$$\inf_{\varphi\in\Phi} \|\varphi_{m,n} - \varphi\|_{L_p(\Omega)}^p \leq \int_\Omega (\varepsilon^{1/p}|B(0, d)|^{-1/p})^p dx \leq \varepsilon,$$

hence, the set Φ is a finite ε-net for $\{\varphi_{m,n}\}_{m=m_0(\varepsilon)}^\infty$ in $L_p(\Omega)$. Thus, there exists $\varphi_\varepsilon \in \Phi$ and infinite subsequences $\{\varphi_{m,n}^\varepsilon\}_{m\geq 1}$ and $\{g_m^\varepsilon\}_{m\geq 1}$, such that $\left\|\varphi_{m,n}^\varepsilon - \varphi_\varepsilon\right\|_{L_p(\Omega)}^p \leq \varepsilon$ and, in turn, $\left\|g_m^\varepsilon - \varphi_\varepsilon\right\|_{L_p(\Omega_m)}^p \leq 2\varepsilon$. Applying the above process for $\varepsilon_k := 1/(2k)$, $k \geq 2$, we can construct a sequence $\{\varphi_k\}_{k\geq 2}$ with the following properties:

(i) $0 < C_1 \le \|\varphi_k\|_{L_p(\Omega)} \le C_2 < \infty$.
(ii) For each $k \ge 2$, $\|\varphi_k - g_{k,j}\|_{L_p(\Omega_{k,j})} \le 1/k$, $\forall j \ge 1$, where $\{g_{k,j}\}_{j\ge 1}$ is an infinite subsequence of $\{g_m\}$.
(iii) $E_{r-1}(\varphi_k, \Omega)_p^p \ge \frac{1}{2}$.
(For, since $\Omega_{k,j} \subseteq \Omega$, $j \ge 1$, it follows that

$$\begin{aligned} E_{r-1}(\varphi_k, \Omega)_p^p &\ge \inf_{Q\in\Pi_{r-1}} \|\varphi_k - Q\|_{L_p(\Omega_{k,j})}^p \\ &\ge \inf_{Q\in\Pi_{r-1}} \|g_{k,j} - Q\|_{L_p(\Omega_{k,j})}^p - \|\varphi_k - g_{k,j}\|_{L_p(\Omega_{k,j})}^p \\ &\ge 1 - 1/k \ge \tfrac{1}{2}.) \end{aligned}$$

Finally,
(iv) $\omega_r(\varphi_k, \Omega)_p^p \le C/k$, where $C = C(r)$.

Indeed, for a fixed $k \ge 2$, let $h \in \mathbb{R}^d$, $|h| \le \operatorname{diam}(\Omega)$, be such that

$$\omega_r(\varphi_k, \Omega)_p^p \le 2 \int_\Omega |\Delta_h^r(\varphi_k, \Omega, x)|^p \, dx.$$

Now let

$$\Omega_{k,j,h} := \{x \in \Omega : [x, x+rh] \subset \Omega,\ [x, x+rh] \not\subset \Omega_{k,j}\},$$

and

$$\widetilde{\Omega}_{k,j,h} := \bigcup_{i=0}^{r} (\Omega_{k,j,h} + ih),$$

where $\Omega_{k,j,h} + ih := \{x + ih : x \in \Omega_{k,j,h}\}$.

As the domains $\Omega_{k,j}$ converge to Ω as $j \to \infty$, in the sense of Definition 2.11, it follows that the measure of the sets $\widetilde{\Omega}_{k,j,h}$ tends to zero as $j \to \infty$. Consequently,

$$\int_{\widetilde{\Omega}_{k,j,h}} |\varphi_k(x)|^p \, dx \to 0, \qquad j \to \infty, \tag{2.14}$$

This gives

$$\begin{aligned} \omega_r(\varphi_k, \Omega)_p^p &\le 2 \int_\Omega |\Delta_h^r(\varphi_k, \Omega, x)|^p \, dx \\ &\le 2 \left(\int_{\Omega\setminus\Omega_{k,j,h}} |\Delta_h^r(\varphi_k, \Omega, x)|^p \, dx + \int_{\Omega_{k,j,h}} |\Delta_h^r(\varphi_k, \Omega, x)|^p \, dx \right) \\ &\le C \left(\int_{\Omega_{k,j}} |\Delta_h^r(\varphi_k, \Omega_{k,j}, x)|^p \, dx + \int_{\widetilde{\Omega}_{k,j,h}} |\varphi_k(x)|^p \, dx \right) \end{aligned}$$

$$\le C\left(\int_{\Omega_{k,j}} |\Delta_h^r(g_{k,j},\Omega_{k,j},x)|^p\,dx + \|\varphi_k - g_{k,j}\|^p_{L_p(\Omega_{k,j})}\right.$$
$$\left.+\int_{\widetilde{\Omega}_{k,j,h}} |\varphi_k(x)|^p\,dx\right)$$
$$\le C\left(\omega_r(g_{k,j},\Omega_{k,j})_p^p + \|\varphi_k - g_{k,j}\|^p_{L_p(\Omega_{k,j})} + \int_{\widetilde{\Omega}_{k,j,h}} |\varphi_k(x)|^p\,dx\right)$$
$$=: C(I_1 + I_2 + I_3).$$

Finally, $I_1 = \omega_r(g_{k,j},\Omega_{k,j})_p^p \to 0$, as $j \to \infty$, and by (2.14), $I_3 \to 0$, as $j \to \infty$, while, by (ii), $I_2 \le 1/k$ for all $j \ge 1$. This completes the proof of (iv).

We now repeat the proof with the sequence $\{\varphi_k\}_{k\ge 2}$ on the fixed domain Ω, in place of sequences $\{g_m\}_{m\ge 1}$, $\{\Omega_m\}_{m\ge 1}$. This can be done because the properties (i), (iii), and (iv) of $\{\varphi_k\}$ are almost the same as the properties (i) and (ii) of $\{g_m\}$ and, in addition, we have the major advantage of a fixed domain Ω. Thus we obtain sequences $\{\Psi_{k,n}\}$, of piecewise constants on the grid of length n^{-1}, for which

$$\|\varphi_k - \Psi_{k,n}\|^p_{L_p(\Omega)} \le \varepsilon,$$

and that these sequences possess the finite ε-net property. That is, for each $\varepsilon > 0$, we have a Ψ^ε such that $\|\varphi_k^\varepsilon - \Psi^\varepsilon\|^p_{L_p(\Omega)} \le 2\varepsilon$, for an infinite subsequence of the φ_k's. Taking $\varepsilon_l = 1/(2l)$ and repeating the argument for $l = 2, 3, \ldots$, each time taking a subsequence of the previous one. In summary, we obtain a sequence $\{\Psi_l\}_{l\ge 2}$ and a sequence $\{\varphi_{k_j}\}_{j\ge 2}$ such that

$$\|\Psi_l - \varphi_{k_j}\|^p_{L_p(\Omega)} \le \frac{1}{l}, \qquad \forall j \ge l.$$

Hence $\{\Psi_l\}_{l\ge 2}$ is a Cauchy sequence in $L_p(\Omega)$, and therefore converges to, say, $\Psi \in L_p(\Omega)$. This implies that $\varphi_{k_j} \to \Psi$ in $L_p(\Omega)$ and, in turn, that on the one hand $\omega_r(\Psi,\Omega)_p = 0$, while on the other hand,

$$E_{r-1}(\Psi,\Omega)_p^p \ge \inf_{Q\in\Pi_{r-1}} \|\varphi_{k_j} - Q\|^p_{L_p(\Omega)} - \|\Psi - \varphi_{k_j}\|^p_{L_p(\Omega)}$$
$$\ge \tfrac{1}{2} - \|\Psi - \varphi_{k_j}\|^p_{L_p(\Omega)} \to \tfrac{1}{2} \quad \text{as } j \to \infty,$$

contradicting Lemma 2.13.

We conclude that there exists a constant $C(d,r,p)$ such that for all bounded convex domains Ω and all functions $f \in L_p(\Omega)$, $0 < p < 1$,

$$E_{r-1}(f)_p \le C(d,r,p)\,\omega_r(f,\Omega)_p. \qquad \square$$

3. Multivariate Skinny B-Spaces

Karaivanov and Petrushev [**15**] introduce the bivariate skinny B-spaces and remark that they can be extended to higher dimensions but they do not consider them since they want to avoid some complications (see [**15**, end of Section 2]). The machinery that we have developed here enables us to extend the skinny B-spaces to arbitrary dimension $d \geq 2$, and to obtain their approximation properties without further complications.

A set $\mathcal{T}$ of d-simplices is called a **weak locally regular (WLR-)triangulation** of $\mathbb{R}^d$ with levels $\{\mathcal{T}_m\}_{m\in\mathbb{Z}}$ if $\mathcal{T} = \bigcup_{m\in\mathbb{Z}} \mathcal{T}_m$ satisfies the following conditions:

(i) Every level $\mathcal{T}_m$ is a set of simplices with disjoint interiors such that

$$\mathbb{R}^d = \bigcup_{\Delta\in\mathcal{T}_m} \Delta.$$

Note that, by definition, simplices are compact and convex.

(ii) The levels $\mathcal{T}_m$ are nested, that is, for every $\Delta \in \mathcal{T}_m$,

$$\Delta = \bigcup_{\substack{\Delta'\in\mathcal{T}_{m+1}\\ \Delta'\subset\Delta}} \Delta'.$$

Therefore, any two simplices in $\mathcal{T}$ either have disjoint interiors or one of them contains the other. We shall call Δ' a **child** of $\Delta \in \mathcal{T}_m$ if $\Delta' \in \mathcal{T}_{m+1}$ and $\Delta' \subset \Delta$.

(iii) There exist constants $0 < \rho_1 < \rho_2 < 1$ ($\rho_1 \leq \frac{1}{4}$) such that, for each $\Delta \in \mathcal{T}$ and any child Δ' of Δ,

$$\rho_1|\Delta| \leq |\Delta'| \leq \rho_2|\Delta|.$$

In particular, the number of children of any $\Delta \in \mathcal{T}$ satisfies

$$1 < \lfloor\rho_2^{-1}\rfloor \leq \#\, child(\Delta) \leq \lceil\rho_1^{-1}\rceil.$$

The following are generalizations of Lemma 2.7(a) and (b) in [**15**]:

Lemma 3.1. *Let $P \in \Pi_{r-1}(\mathbb{R}^d)$ and let $\Omega_1, \Omega_2 \subset \mathbb{R}^d$ be bounded convex domains such that $\Omega_1 \subseteq \Omega_2$ and $|\Omega_2| \leq \rho|\Omega_1|$ for some $\rho > 1$. Then, for $0 < p \leq \infty$,*

$$\|P\|_{L_p(\Omega_2)} \leq C(d, r, p, \rho)\|P\|_{L_p(\Omega_1)}.$$

Proof. Let $Ax = Mx + b$ be the affine transform for which (2.2) holds for Ω_1. Since $A^{-1}(\Omega_1) \subseteq B(0, d)$ we have

$$\begin{aligned} |A^{-1}(\Omega_2)| &= |A^{-1}(\Omega_1)|\frac{|A^{-1}(\Omega_2)|}{|A^{-1}(\Omega_1)|} \\ &\leq |B(0, d)|\rho := C(d, \rho). \end{aligned} \tag{3.1}$$

Observe that $A^{-1}(\Omega_2)$ is a convex domain that contains $A^{-1}(\Omega_1)$ and therefore also contains $B(0, 1)$. Together with (3.1) this implies that the diameter of such a domain must also be bounded by a constant that depends on d and ρ, i.e., $A^{-1}(\Omega_2) \subseteq B(0, R)$, $R := R(d, \rho)$. Hence

$$\begin{aligned}
\|P\|_{L_p(\Omega_2)} &= |\det M|^{1/p} \|P\|_{L_p(A^{-1}(\Omega_2))} \\
&\le |\det M|^{1/p} \|P\|_{L_p(B(0,R))} \\
&\le C|\det M|^{1/p} \|P\|_{L_p(B(0,1))} \\
&\le C|\det M|^{1/p} \|P\|_{L_p(A^{-1}(\Omega_1))} \\
&= C\|P\|_{L_p(\Omega_1)}.
\end{aligned}$$

□

Lemma 3.2. *For any bounded convex domain* $\Omega \subset \mathbb{R}^d$, $P \in \Pi_{r-1}(\mathbb{R}^d)$, *and* $0 < p, q \le \infty$, *we have*

$$\|P\|_{L_q(\Omega)} \approx |\Omega|^{1/q-1/p} \|P\|_{L_p(\Omega)}, \tag{3.2}$$

with constants of equivalency depending only on d, r, p, *and* q. *Also, if* Ω', Ω *are* d*-simplices such that* $\Omega' \subset \Omega$, *and* $|\Omega'| \le \rho|\Omega|$ *with* $0 < \rho < 1$, *then*

$$\|P\|_{L_q(\Omega)} \approx \|P\|_{L_q(\Omega\setminus\Omega')} \approx |\Omega|^{1/q-1/p} \|P\|_{L_p(\Omega\setminus\Omega')}, \tag{3.3}$$

with constants of equivalency depending only on d, r, p, q, *and* ρ.

Proof. We begin with a proof of (3.2). Let $Ax = Mx + b$ be the affine transform for which (2.2) holds. Since $A(B(0, 1)) = E$, we get from the properties of John's ellipsoid, $|\det M| \approx |\Omega|$, with constants of equivalency depending only on d. Also, for any polynomial $\widetilde{P} \in \Pi_{r-1}(\mathbb{R}^d)$ we have that $\|\widetilde{P}\|_{L_p(B(0,1))} \approx \|\widetilde{P}\|_{L_q(B(0,d))}$ with constants of equivalency that depend only on d, r, p, and q. Let $P \in \Pi_{r-1}(\mathbb{R}^d)$, and denote $\widetilde{P} := P(A\cdot)$. Then

$$\begin{aligned}
\|P\|_{L_q(\Omega)} &= |\det M|^{1/q} \|\widetilde{P}\|_{L_q(A^{-1}(\Omega))} \\
&\le |\det M|^{1/q} \|\widetilde{P}\|_{L_q(B(0,d))} \\
&\le C|\det M|^{1/q} \|\widetilde{P}\|_{L_p(B(0,1))} \\
&\le C|\det M|^{1/q} \|\widetilde{P}\|_{L_p(A^{-1}(\Omega))} \\
&\le C|\det M|^{1/q-1/p} \|P\|_{L_p(\Omega)} \\
&\le C|\Omega|^{1/q-1/p} \|P\|_{L_p(\Omega)}.
\end{aligned}$$

To prove (3.3) assume that Ω' and Ω are d-simplices such that $|\Omega'| \le \rho|\Omega|$ with $0 < \rho < 1$. We claim that there exists a constant $\rho'' := \rho''(d, \rho) > 1$ and a d-simplex $\Omega'' \subseteq \Omega\setminus\Omega'$ such that $|\Omega| \le \rho''|\Omega''|$. Indeed, we can triangulate $\Omega\setminus\Omega'$ into at most $C(d) = \binom{2d+2}{d+1}$, d-simplices, with a total volume $\ge (1 - \rho)|\Omega|$. Consequently, with $\rho'' := (1 - \rho)/C(d)$, there exists at least one simplex $\Omega'' \subseteq \Omega\setminus\Omega'$ such that $|\Omega| \le \rho''|\Omega''|$. By Lemma 3.1, we have

$$\begin{aligned}
\|P\|_{L_q(\Omega)} &\le C(d, r, q, \rho'')\|P\|_{L_q(\Omega'')} \\
&\le C(d, r, q, \rho)\|P\|_{L_q(\Omega\setminus\Omega')},
\end{aligned}$$

which implies the left-hand side equivalency. By the first part of the proof we also have that

$$\|P\|_{L_q(\Omega'')} \approx |\Omega''|^{1/q-1/p}\|P\|_{L_p(\Omega)},$$

which together with the equivalency $|\Omega''| \approx |\Omega|$ yields the right-hand side of (3.3). □

Definition 3.3. For a given WLR-triangulation $\mathcal{T}$ of $\mathbb{R}^d$ we define the **skinny B-space** $\mathcal{B}^{\alpha,r}_\tau(\mathcal{T})$, $\alpha > 0$, $0 < \tau < \infty$, $r \in \mathbb{N}$, as the set of functions $f \in L_\tau(\mathbb{R}^d)$ for which

$$\|f\|_{\mathcal{B}^{\alpha,r}_\tau(\mathcal{T})} := \left(\sum_{\Delta\in\mathcal{T}} (|\Delta|^{-\alpha}\omega_r(f,\Delta)_\tau)^\tau\right)^{1/\tau} < \infty.$$

Since the triangulation is composed of convex elements, we may apply (1.13) and (1.14) to obtain

$$\|f\|_{\mathcal{B}^{\alpha,r}_\tau(\mathcal{T})} \approx \left(\sum_{\Delta\in\mathcal{T}} (|\Delta|^{-\alpha}E_{r-1}(f,\Delta)_\tau)^\tau\right)^{1/\tau}. \tag{3.4}$$

To characterize nonlinear n-term piecewise polynomial approximation in the p-norm, over simplices taken from $\mathcal{T}$, where the error "decays" at the rate $n^{-\alpha}$, $\alpha > 0$, we will be interested in the spaces $\mathcal{B}^{\alpha,r}_\tau(\mathcal{T})$, $1/\tau = \alpha + 1/p$. Evidently, $\|\cdot\|_{\mathcal{B}^{\alpha,r}_\tau(\mathcal{T})}$ is a norm if $\tau \ge 1$ and a quasi-norm if $\tau < 1$.

Following [**15**] we introduce another equivalent norm in $\mathcal{B}^{\alpha_r}_\tau(\mathcal{T})$. For $f \in L^{\text{loc}}_\eta(\mathbb{R}^d)$, $\eta > 0$, and each $\Delta \in \mathcal{T}$, we let $P_{\Delta,\eta}(f)$ be a **near best** $L_\eta(\Delta)$-approximation to f from Π_{r-1}, namely,

$$\|f - P_{\Delta,\eta}\|_{L_\eta(\Delta)} \le AE_{r-1}(f,\Delta)_\eta, \qquad \text{for all } \Delta \in \mathcal{T}.$$

We denote $P_{m,\eta}(f) := \sum_{\Delta\in\mathcal{T}_m} \mathbf{1}_\Delta P_{\Delta,\eta} \in S^r_m(\mathcal{T})$, where $S^r_m(\mathcal{T})$ is the set of all piecewise polynomials of degree $r-1$ over the simplices in $\mathcal{T}_m$. Let

$$p_{m,\eta}(f) := p_{m,\eta}(f,\mathcal{T}) := P_{m,\eta}(f) - P_{m-1,\eta}(f) \in S^r_m(\mathcal{T}),$$

and set $p_{\Delta,\eta}(f) := \mathbf{1}_\Delta p_{m,\eta}$ for every $\Delta \in \mathcal{T}_m$. The elements $p_{\Delta,\eta}(f)$ play the role of local wavelet components of f. In fact, the strategy used below to obtain a near best n-term piecewise polynomial approximation is to pick the n elements $p_{\Delta,\eta}(f)$ with the biggest norm.

Denoting

$$\mathcal{N}_{p,\eta}(f,\mathcal{T}) := \left(\sum_{\Delta\in\mathcal{T}} (|\Delta|^{1/p-1/\eta}\|p_{\Delta,\eta}(f)\|_\eta)^\tau\right)^{1/\tau}, \qquad 1/\tau = \alpha + 1/p, \tag{3.5}$$

we may apply Lemma 3.2 to obtain, for the particular case $\eta = \tau$,

$$\mathcal{N}_{p,\tau}(f,\mathcal{T}) = \left(\sum_{\Delta\in\mathcal{T}} (|\Delta|^{-\alpha}\|p_{\Delta,\tau}(f)\|_\tau)^\tau\right)^{1/\tau} \approx \left(\sum_{\Delta\in\mathcal{T}} \|p_{\Delta,\tau}(f)\|_p^\tau\right)^{1/\tau}. \quad (3.6)$$

Theorem 3.4. *If* $0 < \eta < p$, *then the norms* $\|\cdot\|_{\mathcal{B}^{\alpha,r}_\tau(\mathcal{T})}$ *and* $\mathcal{N}_{p,\eta}(f,\mathcal{T})$ *are equivalent with constants of equivalence depending only on* α, d, r, p, ρ_1, *and* ρ_2.

Sketch of proof. The proof is almost identical to the proof of [**15**, Theorem 2.18], except for the fact that the generalized Whitney inequality (1.12) and Lemma 3.2, replace the Whitney inequality for triangles and [**15**, Lemma 2.7(b)]. □

Let $\Sigma^r_n(\mathcal{T})$ be the collection

$$\sum_{k=1}^n \mathbf{1}_{\Delta_k} P_k,$$

where $\Delta_k \in \mathcal{T}$ and $P_k \in \Pi_{r-1}(\mathbb{R}^d)$, $1 \le k \le n$, and let

$$\sigma_{n,r}(f,\mathcal{T})_p := \inf_{S\in\Sigma^r_n(\mathcal{T})} \|f - S\|_p,$$

denote the degree of nonlinear approximation from $\Sigma^r_n(\mathcal{T})$. We have

Theorem 3.5 (Jackson Estimate). *Let* $\mathcal{T}$ *be a WLR-triangulation*, $0 < p < \infty$, $\alpha > 0$, *and* $r \in \mathbb{N}$. *If* $f \in \mathcal{B}^{\alpha,r}_\tau(\mathcal{T})$, *and* $1/\tau = \alpha + 1/p$, *then*

$$\sigma_{n,r}(f,\mathcal{T})_p \le Cn^{-\alpha}\|f\|_{\mathcal{B}^{\alpha,r}_\tau(\mathcal{T})}, \quad (3.7)$$

with $C := C(\alpha, d, r, p, \rho_1, \rho_2)$.

Sketch of proof. The proof is very similar to the proof of [**15**, Theorem 3.10]. The main tools are [**15**, Theorem 3.4], which is a general Jackson inequality for nonlinear approximation and the norm (3.5), that by Theorem 3.4 is equivalent to the skinny B-space norm. □

Remark. From (3.7) we get the stronger Jackson estimate

$$\widetilde{\sigma}_{n,r}(f)_p := \inf_{\mathcal{T}} \sigma_{n,r}(f,\mathcal{T})_p \le Cn^{-\alpha} \inf_{\mathcal{T}} \|f\|_{\mathcal{B}^{\alpha,r}_\tau(\mathcal{T})} \quad (3.8)$$

where the infimum in (3.8) is taken over all WLR-triangulations with fixed parameters ρ_1, ρ_2 and with $C := C(\alpha, d, r, p, \rho_1, \rho_2)$.

Theorem 3.6 (Bernstein Estimate). *Let* $\mathcal{T}$ *be a WLR-triangulation and let* $S \in \Sigma^r_n(\mathcal{T})$. *Then, for* $0 < p < \infty$, $\alpha > 0$, *and* $1/\tau = \alpha + 1/p$,

$$\|S\|_{\mathcal{B}^{\alpha,r}_\tau(\mathcal{T})} \le Cn^\alpha \|S\|_p, \quad (3.9)$$

with $C := C(\alpha, d, r, p, \rho_1, \rho_2)$.

Sketch of proof. The proof is identical to the proof of [**15**, Theorem 3.11] with the exception that Lemmas 3.1 and 3.2 replace [**15**, Lemma 2.7(a),(b)]. □

The Jackson and Bernstein theorems above enable us to characterize the approximation spaces associated with nonlinear n-term approximation over a fixed triangulation. For a WLR-triangulation $\mathcal{T}$, we denote the **approximation space** $A_q^{\gamma,r}(L_p,\mathcal{T})$, $\gamma>0$, $0<q\le\infty$, to be the set of functions $f\in L_p(\mathbb{R}^d)$ for which

$$|f|_{A_q^{\gamma,r}(L_p,\mathcal{T})} := \begin{cases} \left(\sum_{m=0}^{\infty}\left(2^{m\gamma}\sigma_{2^m,r}(f,\mathcal{T})_p\right)^q\right)^{1/q}, & 0<q<\infty,\\ \sup_{m\ge 0}\left\{2^{m\gamma}\sigma_{2^m,r}(f,\mathcal{T})_p\right\}, & q=\infty,\end{cases}$$

is finite. For a skinny B-space, $\mathcal{B}$, we introduce the K-functional corresponding to the pair $L_p(\mathbb{R}^d)$ and $\mathcal{B}$,

$$K(f,t):=K(f,t,L_p,\mathcal{B}):=\inf_{g\in\mathcal{B}}\{\|f-g\|_p+t|g|_{\mathcal{B}}\},\qquad t>0.$$

The **interpolation space** $(L_p,\mathcal{B})_{\lambda,q}$, $\lambda>0$, $0<q\le\infty$, is defined as the set of all $f\in L_p(\mathbb{R}^d)$ such that

$$|f|_{(L_p,\mathcal{B})_{\lambda,q}} := \begin{cases} \left(\sum_{m=0}^{\infty}\left(2^{m\lambda}K(f,2^{-m})\right)^q\right)^{1/q}, & 0<q<\infty,\\ \sup_{m\ge 0}\left\{2^{m\lambda}K(f,2^{-m})\right\}, & q=\infty,\end{cases}$$

is finite. The norm in $(L_p,\mathcal{B})_{\lambda,q}$ is defined by $\|f\|_{(L_p,\mathcal{B})_{\lambda,q}}:=\|f\|_p+|f|_{(L_p,\mathcal{B})_{\lambda,q}}$. As in [**15**] (see also [**8**] for a survey of this technique), the pair of Jackson and Bernstein estimates above gives the following characterization:

Theorem 3.7. *If $\mathcal{T}$ is a WLR-triangulation*, $0<\gamma<\alpha$, $0<q\le\infty$, $0<p<\infty$, *and* $r\in\mathbb{N}$, *then*

$$A_q^{\gamma,r}(L_p,\mathcal{T})\approx(L_p,\mathcal{B}_\tau^{\alpha,r}(\mathcal{T}))_{\gamma/\alpha,q},$$

where $\tau=(\alpha+1/p)^{-1}$.

References

[1] R. Adams, *Sobolev Spaces*, Academic Press, San Diego, 1975.

[2] K. Ball, Ellipsoids of maximal volume in convex bodies, *Geom. Dedicata* **41** (1992), 241–250.

[3] J. H. Bramble and S. R. Hilbert, Estimation of linear functionals on Sobolev spaces with applications to Fourier transforms and spline interpolations, *SIAM J. Numer. Anal.* **7** (1970), 113–24.

[4] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Elements Methods*, Springer-Verlag, Berlin, 1994.

[5] Y. A. Brudnyi and N. J. Kalton, Polynomial approximation on convex subsets of $\mathbb{R}^n$, *Constr. Approx.* **16** (2000), 161–199.

[6] S. Dekel and D. Leviatan, The Bramble–Hilbert Lemma for convex domains, *SIAM J. Math. Anal.* **35** (2004), 1203–1212.
[7] S. Dekel, D. Leviatan, and M. Sharir, On bivariate smoothness spaces associated with nonlinear approximation, *Constr. Approx*. (to appear).
[8] R. A. DeVore, Nonlinear approximation, *Acta Numerica* **7** (1998), 51–150.
[9] R. A. DeVore and G. G. Lorentz, *Constructive Approximation*, Springer-Verlag, Berlin, 1991.
[10] R. A. DeVore and V. Popov, Interpolation of Besov spaces, *Trans. Amer. Math. Soc.* **305** (1988), 397–414.
[11] Z. Ditzian, V. H. Hristov, and K. Ivanov, Moduli of smoothness and K-functionals in L_p, $0 < p < 1$, *Constr. Approx.* **11** (1995), 67–83.
[12] I. P. Irodova, Generalization of the Marchaud inequality, in *Studies in the Theory of Functions of Several Variables*, Yaroslav. Gos. Univ. Yaroslavl/Cprime, Vol. 152, 1984, pp. 63–69 (in Russian).
[13] F. John, Extremum problems with inequalities as subsidiary conditions, in *Studies and Essays Presented to R. Courant on his* 60*th Birthday*, Interscience, New York, 1948, pp. 187–204.
[14] H. Johnen and K. Scherer, *On the Equivalence of the K-Functional and the Moduli of Continuity and Some Applications*, Lecture Notes in Mathematics, Vol. 571, pp. 119–140, Springer-Verlag, Berlin, 1976.
[15] B. Karaivanov and P. Petrushev, Nonlinear piecewise polynomial approximation beyond Besov spaces, *Appl. Comput. Harmonic Anal.* **15** (2003), 177–223.
[16] P. Petrushev, Nonlinear n-term approximation from hierarchical spline bases, in *Constructive Theory of Functions*, Varna 2002 (B. Bojanov, Ed.), DARBA, Sofia 2003, pp. 33–85.
[17] P. Petrushev and V. Popov, *Rational Approximation of Real Functions*, Cambridge University Press, Cambridge, UK, 1987.
[18] E. M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, NJ, 1970.
[19] E. A. Storozhenko and P. Oswald, Jackson's theorem in the spaces $L^p(\mathbb{R}^k)$, $0 < p < 1$, *Siberian Math. J.* **19** (1978), 630–639.