

NONLINEAR APPROXIMATION FROM DIFFERENTIABLE PIECEWISE POLYNOMIALS*

OLEG DAVYDOV[†] AND PENCHO PETRUSHEV[‡]

Abstract. We study nonlinear n -term approximation in $L_p(\mathbb{R}^2)$ ($0 < p \leq \infty$) from hierarchical sequences of stable local bases consisting of differentiable (i.e., C^r with $r \geq 1$) piecewise polynomials (splines). We construct such sequences of bases over multilevel nested triangulations of \mathbb{R}^2 , which allow arbitrarily sharp angles. To quantize nonlinear n -term spline approximation, we introduce and explore a collection of smoothness spaces (B-spaces). We utilize the B-spaces to prove companion Jackson and Bernstein estimates and then characterize the rates of approximation by interpolation. Even when applied on uniform triangulations with well-known families of basis functions such as box splines, these results give a more complete characterization of the approximation rates than the existing ones involving Besov spaces. Our results can easily be extended to properly defined multilevel triangulations in \mathbb{R}^d , $d > 2$.

Key words. nonlinear approximation, Jackson and Bernstein estimates, multivariate splines, multilevel nested triangulations, multilevel bases, stable local spline bases

AMS subject classifications. 41A15, 41A25, 41A63, 65D07, 65D17

DOI. 10.1137/S0036141002409374

1. Introduction. Nonlinear approximation of functions in dimensions $d > 1$ is a challenging area, especially if one moves away from tensor product-type approaches in order to more adequately approximate functions with singularities along curves and with other anisotropies. One of the most natural tools for approximation is piecewise polynomials over triangulations, and a fundamental problem is to characterize the rate of nonlinear approximation in L_p ($0 < p \leq \infty$) in terms of properly defined global smoothness conditions. This problem is disheartening if one allows the nonlinear approximation to be from any piecewise polynomial over an arbitrary triangulation. The difficulty stems from the highly nonlinear nature of piecewise polynomials in dimensions $d > 1$. For instance, if s_1 and s_2 are two piecewise polynomials over n triangles in \mathbb{R}^2 each, then $s_1 + s_2$ is in general a piecewise polynomial over many more than n (even $> n^2$) pieces. Therefore, the well-known recipe of proving Jackson and Bernstein estimates and then applying interpolation is useless.

The problem becomes even harder when *differentiable* piecewise polynomials are needed, which, for instance, is the case for numerous practical applications of surface modeling and for the conforming finite element methods for higher order PDEs. Moreover, there is an intrinsic demand for differentiability of the approximating tools from the point of view of the nonlinear approximation theory itself. Indeed, this property, together with local reproduction of higher degree polynomials, is crucial for the ability to represent higher order smoothness spaces, such as classical Sobolev or Besov spaces in regular settings (see Theorem 2.15). The desirable differentiability of the approximating piecewise polynomials, however, leads to additional difficulties

*Received by the editors June 9, 2002; accepted for publication (in revised form) March 14, 2003; published electronically October 2, 2003.

<http://www.siam.org/journals/sima/35-3/40937.html>

[†]Mathematisches Institut, Justus-Liebig-Universität Giessen, D-35392 Giessen, Germany (oleg.davydov@math.uni-giessen.de, <http://www.uni-giessen.de/~gcn5/davydov/>).

[‡]Department of Mathematics, University of South Carolina, Columbia, SC 29208, (pencho@math.sc.edu, <http://www.math.sc.edu/~pencho>). This author was supported by National Science Foundation grant DMS-0200665.

because of the complicated structure of spaces of multivariate splines. For example, the dimension is not known and stable local bases are impossible in general already for the space of all piecewise polynomials of degree $< k$ and smoothness $r \geq 1$ with respect to a finite triangulation of a polygonal domain in \mathbb{R}^2 if $k \leq 3r + 2$ [10].

A reasonable alternative to “spline approximation with free triangulations” is to consider nonlinear n -term approximation from hierarchical sequences of spline bases over *multilevel nested triangulations* of \mathbb{R}^d . (For the sake of simplicity, we shall restrict ourselves in this article to the case $d = 2$.) To explain this concept more precisely, consider a sequence $(\mathcal{T}_m)_{m \in \mathbb{Z}}$ of partitions of \mathbb{R}^2 into triangles with disjoint interiors such that each *level* \mathcal{T}_m is a refinement of the previous one \mathcal{T}_{m-1} . Let $\mathcal{T} := \bigcup_{m \in \mathbb{Z}} \mathcal{T}_m$. We impose certain mild (and natural) conditions on the triangulations which prevent them from deterioration but still allow the triangles to change in size, shape, and orientation quickly when moving around at a given level or through the levels. In particular, triangles with arbitrarily sharp angles may occur at any location. We denote by $\mathcal{S}^{k,r}(\mathcal{T}_m)$ the set of all r -times differentiable piecewise polynomials with respect to \mathcal{T}_m of degree $< k$. Given a ladder of spaces

$$(1.1) \quad \cdots \subset \mathcal{S}_{-1} \subset \mathcal{S}_0 \subset \mathcal{S}_1 \subset \cdots, \quad \mathcal{S}_m \subset \mathcal{S}^{k,r}(\mathcal{T}_m),$$

and bases Φ_m of \mathcal{S}_m , $m \in \mathbb{Z}$, we set $\Phi := \Phi_{\mathcal{T}} := \bigcup_{m \in \mathbb{Z}} \Phi_m$. Using the standard wavelet terminology, we can describe such a nested sequence of spaces with bases as “*spline multiresolution*” (or “*multiresolution analysis*”).

Consider now the problem for nonlinear (n -term) approximation from the set Σ_n of all piecewise polynomials of the form $s = \sum_{j=1}^n c_j \varphi_j$, where $\varphi_j \in \Phi$ may come from different levels and locations. Once a particular multilevel triangulation has been selected, the variety of piecewise polynomial approximations significantly reduces. However, a great deal of flexibility is retained, and the problem remains highly nonlinear. For instance, thin and elongated basis functions are allowed. On the other hand, the advantages of multilevel approximation methods can be exploited in full.

Our program consists of the following basic steps:

1. We construct hierarchical sequences of bases $(\Phi_m)_{m \in \mathbb{Z}}$ on multilevel triangulations satisfying certain requirements of *local regularity* allowing anisotropically shaped triangles.

2. To quantify the approximation process, we introduce and develop a family (library) of smoothness spaces $B_{\tau}^{\alpha}(\Phi_{\mathcal{T}})$ depending on $\Phi_{\mathcal{T}}$ and as a consequence on the triangulation \mathcal{T} . We call them B-spaces since they have some resemblance to Besov spaces. So, the idea is to measure the smoothness of the functions using a family of space scales $B_{\tau}^{\alpha}(\Phi_{\mathcal{T}})$ (which vary with $\Phi_{\mathcal{T}}$) instead of a single scale of smoothness spaces like the scale of Besov spaces.

3. We develop a coherent theory of nonlinear n -term approximation from $\Phi_{\mathcal{T}}$ based on the idea of proving Jackson and Bernstein estimates and interpolation.

4. We utilize this theory in the development of algorithms for nonlinear piecewise polynomial (spline) approximation which capture the rate of the best approximation.

The logic of the resulting *approximation scheme* is the following: Suppose $\{\Phi_{\mathcal{T}}\}_{\mathcal{T}}$ is a collection of multilevel sequences of (spline) bases as above.

- (i) For a given function f , find the “right” triangulation $\mathcal{T} := \mathcal{T}_f$ such that f exhibits the most smoothness (sparsity of its representation) when measured via the scale $B^{\alpha}(\Phi_{\mathcal{T}})$.

(ii) Find an optimal or near optimal representation of f using $\Phi_{\mathcal{T}}$. (Note that $\Phi_{\mathcal{T}}$ is redundant, i.e., linearly dependent.)

(iii) Using this representation, run an algorithm for n -term L_p -approximation which achieves the rate of the best n -term approximation.

Naturally, the first step presents the most challenging problem in this scheme. We do not have a completely satisfactory algorithm for this step. (Note that this problem has a complete and efficient solution in the simpler case of nonlinear approximation from piecewise polynomials over dyadic partitions; see [54].) As it will be shown in this article, the other steps are now well understood and have complete solutions.

The above program has been suggested in [38] and implemented in [38, 39] in the cases of approximation from discontinuous piecewise polynomials and continuous piecewise linear functions ($r = -1$, $k \geq 1$, and $r = 0$, $k = 2$, where $r = -1$ corresponds to the discontinuous case). The simplest example of a hierarchical family of continuous basis functions is the set of all Courant elements generated by a multilevel nested triangulation \mathcal{T} , that is, the set of all piecewise linear and continuous functions $\Phi_{\mathcal{T}} = \{\varphi_{\theta}\}$ supported on the cells $\{\theta\}$ (each θ is the union of all triangles of a particular level \mathcal{T}_m attached to a vertex); see [38].

In the present article, we develop the theory of nonlinear n -term approximation for basis families consisting of differentiable piecewise polynomials ($r \geq 1$). The construction of such basis functions suitable for application is hampered by the fact that both the classical differentiable finite elements [14] and the earlier polynomial spline basis constructions on arbitrary triangulations [1, 8, 16, 17, 18, 35, 36, 44, 48, 57] are difficult to use for our purposes; see Remark 4.12 and the discussion in section 5.3. The stable local spline bases of [27] can in principle be used in two variables. However, all other arguments of our article are basically “dimension independent,” and we refrain here from treating the case $d > 2$ only for the sake of simplicity and clarity. Therefore, we build upon the nodal spline bases of [22], which is the only known approach that produces stable local bases for nested spline spaces on general triangulations in all dimensions.

However, these bases are stable only for triangulations satisfying (in \mathbb{R}^2) the minimal angle condition. We extend the construction of [22] to a wider class of *strong locally regular* triangulations; see section 2 for a definition. Note that the new basis functions are invariant under affine transforms (see Remark 4.9). In the case $r = 0$ our construction reduces to the classical continuous Lagrange finite elements and is valid for any *locally regular* triangulation; see Remark 4.13.

A focal point of our development is the characterization of nonlinear n -term approximation from families of differentiable spline basis functions, including the development of B-spaces, proof of Jackson and Bernstein estimates, and characterization of the approximation spaces by interpolation (see sections 2–3). In [39], there are three algorithms developed for nonlinear n -term approximation in L_p ($0 < p \leq \infty$) from Courant elements. These can be immediately implemented for n -term approximation from differentiable spline bases, and it can be shown similarly as in [39] that they achieve the rate of the best approximation. We do not pursue this goal here.

The B-spaces from the present article can be viewed as a generalization of the “approximation spaces” from section 3.4 of [51] (see also the references therein). More precisely, in the specific setting of “quasi-uniform partitions” and the basis functions used in [51], our B-spaces coincide with the approximation spaces of [51].

The theory of nonlinear n -term approximation from box splines (on uniform triangulations) has been developed in [29] ($p < \infty$) and [30] ($p = \infty$) (for nonlinear

spline approximation in dimension $d = 1$, see [53]). In these articles, direct, inverse, and characterization theorems have been proved utilizing certain Besov spaces. Even in this case, our results, which utilize B-spaces (in place of Besov spaces), are more complete since they characterize nonlinear n -term box spline approximation for all rates of approximation, while in the above-mentioned articles the rate is restricted by the Besov smoothness of the box splines.

There is an apparent connection between our developments here and multilevel finite element methods for PDEs; see, e.g., [51]. Therefore, it seems an interesting task to develop finite element algorithms for solving PDEs which achieve the rate of the best n -term approximation of the solution.

The outline of the article is the following. In section 2, we introduce and develop the B-spaces needed for the characterization of nonlinear approximation for any family of basis functions with certain properties. In section 3, we develop the general theory of nonlinear n -term approximation from piecewise polynomials, where the global smoothness of functions is measured by means of our B-spaces. In section 4, we construct hierarchical sequences of bases consisting of differentiable piecewise polynomials. In section 5, we review a number of alternative constructions fitting into our scheme, based on box splines and some other spline bases on special triangulations. The final section A is an appendix containing some of the proofs.

Throughout the article, we use the following notation: $L_\infty^{\text{loc}}(\mathbb{R}^2) := C(\mathbb{R}^2)$ and $L_\infty(\mathbb{R}^2) := C_0(\mathbb{R}^2) := \{f \in C(\mathbb{R}^2) : \lim_{x \rightarrow \infty} f(x) = 0\}$, $L_q^{\text{loc}} := L_q^{\text{loc}}(\mathbb{R}^2)$, $0 < q \leq \infty$, $C := C(\mathbb{R}^2)$, $\|\cdot\|_q := \|\cdot\|_{L_q(\mathbb{R}^2)}$, $0 < q \leq \infty$; Π_k denotes the set of all algebraic polynomials in two variables of total degree $< k$. For any $\Omega \subset \mathbb{R}^2$, $\mathbb{1}_\Omega$ denotes the characteristic function of Ω and $|\Omega|$ denotes the Lebesgue measure of Ω . Positive constants are denoted by c, c_1, \dots (they may vary at every occurrence), $\alpha \approx \beta$ means $c_1\alpha \leq \beta \leq c_2\alpha$, and $\alpha := \beta$ or $\beta =: \alpha$ stands for “ α is by definition equal to β .”

2. B-spaces generated by spline multiresolution. In the present section, we introduce and explore the smoothness spaces we need for the characterization of nonlinear n -term spline approximation generated by families of differentiable basis functions over multilevel nested triangulations.

2.1. Triangulations. In our development, we utilize three types of multilevel nested triangulations. We shall call each of them simply a *triangulation*, although such a triangulation does not form a single partition of \mathbb{R}^2 but rather an infinite nested family of partitions (each of them is a triangulation of \mathbb{R}^2 in the more commonly used sense).

Let $\mathcal{T} = \bigcup_{m \in \mathbb{Z}} \mathcal{T}_m$ be a set of closed triangles in \mathbb{R}^2 with *levels* \mathcal{T}_m , $m \in \mathbb{Z}$. Denote by \mathcal{V}_m the set of all vertices (nodal points) of triangles from \mathcal{T}_m and set $\mathcal{V} := \bigcup_{m \in \mathbb{Z}} \mathcal{V}_m$. We say that \mathcal{T} is a *triangulation* of \mathbb{R}^2 if the following conditions are fulfilled:

(a) Every level \mathcal{T}_m is a set of triangles with disjoint interiors which cover \mathbb{R}^2 : $\mathbb{R}^2 = \bigcup_{\Delta \in \mathcal{T}_m} \Delta$.

(b) The levels $(\mathcal{T}_m)_{m \in \mathbb{Z}}$ of \mathcal{T} are *nested*; i.e., \mathcal{T}_{m+1} is a refinement of \mathcal{T}_m obtained by splitting each $\Delta \in \mathcal{T}_m$ into subtriangles with disjoint interiors called *children* of Δ .

(c) Each triangle $\Delta \in \mathcal{T}_m$ has at least two and at most M_0 children in \mathcal{T}_{m+1} , where $M_0 \geq 2$ is a constant independent of m .

(d) *No hanging vertices condition*: No vertex of any triangle $\Delta \in \mathcal{T}_m$ lies in the interior of an edge of another triangle from \mathcal{T}_m .

(e) The *valence* N_v of each vertex $v \in \mathcal{V}_m$ (the number of triangles $\Delta \in \mathcal{T}_m$ which share v as a vertex) is $\leq N_0$, where N_0 is a constant.

(f) For any compact $K \subset \mathbb{R}^2$ and any fixed $m \in \mathbb{Z}$, there is a finite collection of triangles from \mathcal{T}_m which cover K .

Note that any two triangles in \mathcal{T} either have disjoint interiors or one of them contains the other. In particular, $\Delta' \in \mathcal{T}_{m+1}$ is a child of $\Delta \in \mathcal{T}_m$ ($m \in \mathbb{Z}$) if and only if $\Delta' \subset \Delta$. If Δ and Δ' are two different triangles in \mathcal{T} and $\Delta' \subset \Delta$, then we say that Δ is an *ancestor* of Δ' , while Δ' is a *descendant* of Δ .

Locally regular triangulations. We call a triangulation $\mathcal{T} = \bigcup_{m \in \mathbb{Z}} \mathcal{T}_m$ a *locally regular triangulation* of \mathbb{R}^2 , or briefly an *LR-triangulation*, if \mathcal{T} satisfies the following additional conditions:

(g) There exists a constant $1/2 \leq \rho < 1$ such that for each $\Delta \in \mathcal{T}$ and any child $\Delta' \in \mathcal{T}$ of Δ ,

$$(2.1) \quad (1 - \rho)|\Delta| \leq |\Delta'| \leq \rho|\Delta|.$$

(h1) There exists a constant $0 < \delta_1 \leq 1$ independent of m such that for any $\Delta', \Delta'' \in \mathcal{T}_m$ ($m \in \mathbb{Z}$) with a common edge,

$$(2.2) \quad \delta_1 \leq |\Delta'|/|\Delta''| \leq \delta_1^{-1}.$$

By (e), it follows that for any $\Delta', \Delta'' \in \mathcal{T}_m$ with at least one common vertex, (2.2) holds with δ_1 replaced by $\delta_1^{N_0/2}$.

Strong locally regular triangulations. We call a triangulation $\mathcal{T} = \bigcup_{m \in \mathbb{Z}} \mathcal{T}_m$ a *strong locally regular triangulation* of \mathbb{R}^2 , or briefly an *SLR-triangulation*, if \mathcal{T} satisfies (2.1) and the following condition that replaces (2.2):

(h2) There exists a constant $0 < \delta_2 \leq 1/2$ such that for any $\Delta', \Delta'' \in \mathcal{T}_m$ ($m \in \mathbb{Z}$) sharing an edge,

$$(2.3) \quad |\text{conv}(\Delta' \cup \Delta'')|/|\Delta'| \leq \delta_2^{-1},$$

where $\text{conv}(G)$ denotes the convex hull of $G \subset \mathbb{R}^2$.

Obviously, (2.3) implies (2.2) with $\delta_1 = \delta_2$. Therefore, each SLR-triangulation is an LR-triangulation.

Regular triangulations. By definition, a triangulation $\mathcal{T} = \bigcup_{m \in \mathbb{Z}} \mathcal{T}_m$ is called a *regular triangulation* if \mathcal{T} satisfies the following condition:

(h3) There exists a constant $\beta = \beta(\mathcal{T}) > 0$ such that the minimal angle of each triangle $\Delta \in \mathcal{T}$ is $\geq \beta$.

Next, we make a few remarks which will help understand better the nature of the triangulations that we utilize.

(i) For each of the three types of triangulations there is a number of constants that are assumed fixed. In what follows we refer to them as *parameters*. Thus the parameters of an SLR-triangulation are M_0 , N_0 , ρ , and δ_2 . Notice that because of (2.1), we can set $M_0 := 1/(1-\rho)$ and remove M_0 from the list of parameters. However, this would tend to obscure the actual role of ρ and M_0 .

(ii) It is a key observation that the collection of all SLR-triangulations with given (fixed) parameters is invariant under affine transforms. The same is true for LR-triangulations.

(iii) It is easy to see that (2.3) is equivalent to the following condition introduced in [38]. *Affine transform angle condition:* There exists a constant $\beta = \beta(\mathcal{T})$, $0 < \beta \leq \pi/3$, such that if $\Delta_0 \in \mathcal{T}_m$, $m \in \mathbb{Z}$, and $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is an affine transform that maps

Δ_0 one-to-one onto an equilateral reference triangle, then for every $\Delta \in \mathcal{T}_m$ which has at least one common vertex with Δ_0 , we have

$$(2.4) \quad \min \text{angle}(A(\Delta)) \geq \beta,$$

where $A(\Delta)$ is the image of Δ by the affine transform A . (The equivalence of the two conditions follows easily from the obvious but important fact that both conditions are invariant under affine transforms. Note that we prefer to use (2.3) rather than (2.4) in the definition of SLR-triangulations in this article since the constant δ_2 appears naturally when estimating norms of the basis functions constructed in section 4 (see (4.8)) and also (2.3) is easier to verify in practical situations.)

(iv) As we have already mentioned, every SLR-triangulation is an LR-triangulation, but the converse statement is not true. Also, every regular triangulation is an SLR-triangulation but not the other way around. Counterexamples are given in [38].

(v) The *maximal angle (MA) condition*

$$(2.5) \quad \pi - \max \text{angle}(\Delta) \geq \beta > 0, \quad \Delta \in \mathcal{T},$$

known from the finite element method [2] is totally different from our conditions of regularity. It is easy to see that there are SLR-triangulations that do not satisfy MA, and there are triangulations that satisfy the MA and fail to be locally regular. As we shall see below (Example 4.7), our construction of stable differentiable basis functions does not extend to triangulations satisfying the MA condition but failing to be SLR.

(vi) The rate of change of the size of the elements ($|\Delta|$, $\min \text{angle}(\Delta)$, and $\text{diam}(\Delta)$) of a triangle $\Delta \in \mathcal{T}$ as Δ moves away from a fixed triangle $\Delta^\diamond \in \mathcal{T}$ for different types of triangulations \mathcal{T} is explored in [38]. We shall briefly discuss this issue for SLR-triangulations, which are the most important type of triangulations for the present article. An SLR-triangulation \mathcal{T} may have an equilateral (or close to such) triangle Δ^\diamond at any level T_m with descendants $\Delta_1 \supset \Delta_2 \supset \dots$ such that $\min \text{angle}(\Delta_j) \rightarrow 0$ as $j \rightarrow \infty$, and also a sequence $(\Delta'_j)_{j=0}^\infty \subset \mathcal{T}_m$ with $\Delta'_0 = \Delta$ and $\Delta'_j \cap \Delta'_{j+1} \neq \emptyset$ ($j = 0, 1, \dots$) such that $\min \text{angle}(\Delta'_j) \rightarrow 0$. Conditions (2.1) and (2.3) suggest geometric rates of change of $|\Delta|$, $\min \text{angle}(\Delta)$, and $\text{diam}(\Delta)$ as $\Delta \in \mathcal{T}_m$ moves away from a fixed $\Delta^\diamond \in \mathcal{T}_m$. In fact, the rate of change is a power of the minimal number of edges connecting Δ and Δ^\diamond ; see [38].

(vii) We shall need to know what happens with the levels \mathcal{T}_m of a triangulation \mathcal{T} as $m \rightarrow -\infty$. By Lemma 2.1 from [38], for each LR-triangulation \mathcal{T} there exists a finite cover $\mathcal{T}_-\infty$ of \mathbb{R}^2 such that either $\mathcal{T}_-\infty = \{\mathbb{R}^2\}$ or $\mathcal{T}_-\infty = (\Delta_\infty^j)_{j=1}^{N_\infty}$, $N_\infty \leq N_0$, where each Δ_∞^j is an *infinite triangle*, i.e., the set of all points on and between two rays which are not collinear and have a common beginning. Moreover, in the second case, the infinite triangles $(\Delta_\infty^j)_{j=1}^{N_\infty}$ have a single common vertex and disjoint interiors, and also each triangle $\Delta \in \mathcal{T}$ and all its ancestors are contained in an infinite triangle $\Delta_\infty^j \in \mathcal{T}_-\infty$.

For more details about multilevel triangulations, see [38].

Some additional notation and preliminaries. We denote by $[v_1, v_2]$ the interval (straight line segment) with endpoints v_1, v_2 and by $|e|$ the length of $e = [v_1, v_2]$. Furthermore, we let $[v_1, v_2, v_3]$ denote the triangle with vertices v_1, v_2, v_3 , and let $|\Delta|$ denote the area of $\Delta = [v_1, v_2, v_3]$. Throughout the article, we assume that the vertices v_1, v_2, v_3 of any triangle $[v_1, v_2, v_3]$ are ordered counterclockwise.

For a triangle $\Delta \in \mathcal{T}_m$ ($m \in \mathbb{Z}$), we define $\text{level}(\Delta) := m$.

For any vertex $v \in \mathcal{V}_m$, we let $\text{star}(v) = \text{star}^1(v)$ denote the *star* of v , i.e., the union of all triangles $\Delta \in \mathcal{T}_m$ attached to v . Moreover, for each $\ell \geq 2$, we denote

by $\text{star}^\ell(v)$ the union of $\text{star}^{\ell-1}(v)$ and the stars of the vertices of $\text{star}^{\ell-1}(v)$. (Note that $\text{star}^\ell(v)$ depends also on the level m , but we do not indicate this in the notation since it is always clear from the context what level is meant.) We also set

$$(2.6) \quad \Omega_\Delta^\ell := \cup\{\text{star}^\ell(v) : v \in \mathcal{V}_m, \Delta \subset \text{star}^\ell(v)\}, \quad \Delta \in \mathcal{T}_m.$$

It is easy to check that $\Omega_\Delta^\ell := \cup\{\text{star}^{2\ell-1}(v) : v \text{ is a vertex of } \Delta\}, \Delta \in \mathcal{T}_m.$

It is readily seen that there exists a constant $c^* = c^*(N_0, \ell) \leq N_0^\ell$ such that

$$(2.7) \quad \#\{\Delta \in \mathcal{T}_m : \Delta \subset \text{star}^\ell(v)\} \leq c^*, \quad v \in \mathcal{V}_m,$$

and hence there exists a constant $c^{**} = c^{**}(N_0, \ell) \leq 3c^*(N_0, 2\ell - 1) - 5 \leq 3N_0^{2\ell-1}$ such that

$$(2.8) \quad \#\{\Delta \in \mathcal{T}_m : \Delta \subset \Omega_{\Delta'}^\ell\} \leq c^{**}, \quad \Delta' \in \mathcal{T}_m.$$

We denote by \mathcal{E}_m the set of all edges of triangles of \mathcal{T}_m and set $\mathcal{E} := \cup_{m \in \mathbb{Z}} \mathcal{E}_m.$ We let $\text{star}(e)$ denote the union of the two triangles attached to $e \in \mathcal{E}_m.$

For future use, we state the following inequality:

$$(2.9) \quad \sum_{\Delta \in \mathcal{T}, \Delta \supset \Delta'} (|\Delta'|/|\Delta|)^\gamma \leq \sum_{j=0}^\infty \rho^{j\gamma} = c(\rho, \gamma) < \infty, \quad \Delta' \in \mathcal{T}, \gamma > 0,$$

which is immediate from the properties of LR-triangulations ($|\Delta'| \leq \rho|\Delta|$ if Δ' is a child of Δ).

2.2. Basis functions: The general setting. Let $\mathcal{T} = \cup_{m \in \mathbb{Z}} \mathcal{T}_m$ be a locally regular (or better) triangulation. For $m \in \mathbb{Z}, r \geq 0,$ and $k \geq 1,$ we denote by $\mathcal{S}_m^{k,r} = \mathcal{S}_m^{k,r}(\mathcal{T}_m)$ the set of all r times differentiable piecewise polynomial functions of degree $< k$ over $\mathcal{T}_m;$ i.e., $s \in \mathcal{S}_m^{k,r}$ if and only if $s \in C^r(\mathbb{R}^2)$ and $s = \sum_{\Delta \in \mathcal{T}_m} \mathbb{1}_\Delta \cdot P_\Delta$ with $P_\Delta \in \Pi_k.$ Naturally, $\mathcal{S}_m^{k,-1}$ will denote the set of all piecewise polynomials of degree $< k$ over \mathcal{T}_m which are, in general, discontinuous across the edges from $\mathcal{E}_m.$

We assume that for each $m \in \mathbb{Z}$ there is a subspace \mathcal{S}_m of $\mathcal{S}_m^{k,r}$ ($r \geq 0, k \geq 2$) and a family $\Phi_m = \{\varphi_\theta : \theta \in \Theta_m\} \subset \mathcal{S}_m$ of *basis functions* satisfying the following conditions:

1. $\Pi_{\tilde{k}} \subset \mathcal{S}_m$ for some $1 \leq \tilde{k} \leq k$ (\tilde{k} independent of m).
2. $\mathcal{S}_m \subset \mathcal{S}_{m+1}$ ($m \in \mathbb{Z}$).
3. For any $s \in \mathcal{S}_m$ there exists a unique sequence of real coefficients $a(s) = (a_\theta(s))_{\theta \in \Theta_m}$ such that

$$s = \sum_{\theta \in \Theta_m} a_\theta(s) \varphi_\theta.$$

(Thus, Φ_m is a basis for \mathcal{S}_m and $(a_\theta(\cdot))_{\theta \in \Theta_m}$ are the dual functionals.)

4. For each $\theta \in \Theta_m$ there is a vertex $v = v_\theta \in \mathcal{V}_m$ such that

$$(2.10) \quad \text{supp } \varphi_\theta \subset \text{star}^\ell(v) =: E_\theta,$$

$$(2.11) \quad \|\varphi_\theta\|_{L_\infty(\mathbb{R}^2)} = \|\varphi_\theta\|_{L_\infty(E_\theta)} \leq M_1,$$

$$(2.12) \quad |a_\theta(s)| \leq M_2 \|s\|_{L_\infty(E_\theta)}, \quad s \in \mathcal{S}_m,$$

where $\ell \geq 1$ and M_1, M_2 are positive constants, all independent of θ and $m.$

Let

$$\Phi := \bigcup_{m \in \mathbb{Z}} \Phi_m \quad \text{and} \quad \Theta := \bigcup_{m \in \mathbb{Z}} \Theta_m.$$

We shall refer to $r, k, \tilde{k}, \ell, M_1,$ and M_2 as *parameters* of Φ .

A simple example of a family of basis functions satisfying the above conditions is the set of well-known Courant elements (continuous piecewise linear basis functions, $r = 0, k = 2$) associated with \mathcal{T} (see [38]). Concrete constructions of differentiable basis functions ($r \geq 1$) will be discussed below in sections 4–5.

Although Θ and Θ_m ($m \in \mathbb{Z}$) are simply index sets, in the case of Courant elements, Θ can be identified as the set of all cells (supports of basis functions). As we shall see in sections 4–5, in general, several basis functions of Φ_m may have the same support. However, the supports of only \leq constant of them may overlap.

LEMMA 2.1. *There is a constant L depending only on $k, \ell,$ and N_0 such that for any $\Delta \in \mathcal{T}_m$ ($m \in \mathbb{Z}$),*

$$(2.13) \quad \#\{\theta \in \Theta_m : E_\theta \supset \Delta\} \leq L,$$

where E_θ is defined in (2.10).

Proof. We have by (2.10) and (2.8)

$$\begin{aligned} \#\{\theta \in \Theta_m : \Delta \subset E_\theta\} &\leq \dim \mathcal{S}_m^{k,r}|_{\Omega_\Delta^\ell} \leq \dim \mathcal{S}^{k,-1}(\mathcal{T}_m)|_{\Omega_\Delta^\ell} \\ &= \binom{k+1}{2} \#\{\Delta' \in \mathcal{T}_m : \Delta' \subset \Omega_\Delta^\ell\} \\ &\leq \binom{k+1}{2} c^{**}. \quad \square \end{aligned}$$

We shall frequently use the equivalence of different norms of polynomials as stated in the following lemma (see also [38]).

LEMMA 2.2. *Let $P \in \Pi_k, k \geq 1,$ and $0 < p, q \leq \infty.$*

(a) *For any triangle $\Delta \subset \mathbb{R}^2, \|P\|_{L_p(\Delta)} \approx |\Delta|^{1/p-1/q} \|P\|_{L_q(\Delta)}$ with constants of equivalence depending only on $p, q,$ and $k.$*

(b) *If Δ and Δ' are two triangles such that $\Delta' \subset \Delta$ and $|\Delta| \leq c_1 |\Delta'|,$ then $\|P\|_{L_p(\Delta)} \leq c \|P\|_{L_p(\Delta')}$ with $c = c(p, k, c_1).$*

(c) *If Δ' and Δ are two triangles such that $\Delta' \subset \Delta$ and $|\Delta'| \leq c_2 |\Delta|$ with $0 < c_2 < 1,$ then $\|P\|_{L_p(\Delta)} \leq c \|P\|_{L_p(\Delta \setminus \Delta')} \approx |\Delta|^{1/p-1/q} \|P\|_{L_q(\Delta \setminus \Delta')}$ with constants depending only on $p, q, k,$ and $c_2.$*

By (2.2) and (2.7), $|E_\theta| \approx |\Delta|$ if $\Delta \subset E_\theta, \Delta \in \mathcal{T}_m,$ and $\theta \in \Theta_m.$ Using this and Lemma 2.2, we obtain that, for $0 < p, q \leq \infty,$

$$(2.14) \quad \|s\|_{L_p(E_\theta)} \approx |E_\theta|^{1/p-1/q} \|s\|_{L_q(E_\theta)}, \quad s \in \mathcal{S}_m, \quad \theta \in \Theta_m,$$

where the constants of equivalence depend on $p, q, k,$ and $\delta_1.$ In particular, we shall need (2.14) with $s = \varphi_\theta,$ when it takes the form $\|\varphi_\theta\|_p \approx |E_\theta|^{1/p-1/q} \|\varphi_\theta\|_q,$ in view of (2.10).

LEMMA 2.3. *The bases Φ_m are L_q -stable for all $0 < q \leq \infty.$ That is, if $g := \sum_{\theta \in \Theta_m} b_\theta \varphi_\theta,$ where $(b_\theta)_{\theta \in \Theta_m}$ is an arbitrary sequence of real numbers, then*

$$\|g\|_q \approx \left(\sum_{\theta \in \Theta_m} \|b_\theta \varphi_\theta\|_q^q \right)^{1/q}.$$

Moreover, for any $\gamma \in \mathbb{R}$ and $0 < \tau \leq \infty$,

$$(2.15) \quad \left(\sum_{\Delta \in \mathcal{T}_m} (|\Delta|^\gamma \|g\|_{L_q(\Delta)})^\tau \right)^{1/\tau} \approx \left(\sum_{\theta \in \Theta_m} (|E_\theta|^\gamma \|b_\theta \varphi_\theta\|_q)^\tau \right)^{1/\tau},$$

where the constants of equivalence are independent of m and g . In the case $q = \infty$ (or $\tau = \infty$) the ℓ_q -norm (ℓ_τ -norm) above is replaced by the sup-norm as usual.

Proof. We have to prove only (2.15), since the first statement of the lemma then follows with $\gamma = 0$ and $\tau = q$. For each $\Delta \in \mathcal{T}_m$, we have by (2.10)

$$\|g\|_{L_q(\Delta)} = \left\| \sum_{\theta \in \Theta_m, E_\theta \supset \Delta} b_\theta \varphi_\theta \right\|_q \leq c \sum_{\theta \in \Theta_m, E_\theta \supset \Delta} \|b_\theta \varphi_\theta\|_q.$$

Therefore, by Lemma 2.1 and (2.7),

$$\begin{aligned} \sum_{\Delta \in \mathcal{T}_m} (|\Delta|^\gamma \|g\|_{L_q(\Delta)})^\tau &\leq c \sum_{\Delta \in \mathcal{T}_m} \sum_{\theta \in \Theta_m, E_\theta \supset \Delta} (|E_\theta|^\gamma \|b_\theta \varphi_\theta\|_q)^\tau \\ &\leq c \sum_{\theta \in \Theta_m} (|E_\theta|^\gamma \|b_\theta \varphi_\theta\|_q)^\tau. \end{aligned}$$

In the other direction, since Φ is a basis of \mathcal{S}_m and $g \in \mathcal{S}_m$, we have $b_\theta = a_\theta(g)$, $\theta \in \Theta_m$, and hence, by (2.12), (2.14), and (2.11),

$$\begin{aligned} \|b_\theta \varphi_\theta\|_q &= \|a_\theta(g) \varphi_\theta\|_q \leq c \|g\|_{L_\infty(E_\theta)} \|\varphi_\theta\|_q \leq c \|g\|_{L_\infty(E_\theta)} |E_\theta|^{1/q} \\ &\leq c \|g\|_{L_q(E_\theta)} \leq c \sum_{\Delta \in \mathcal{T}_m, \Delta \subset E_\theta} \|g\|_{L_q(\Delta)}. \end{aligned}$$

Since $|E_\theta| \approx |\Delta|$ if $\Delta \in \mathcal{T}_m$ and $\Delta \subset E_\theta$, we have, by (2.7) and Lemma 2.1,

$$\begin{aligned} \sum_{\theta \in \Theta_m} (|E_\theta|^\gamma \|b_\theta \varphi_\theta\|_q)^\tau &\leq c \sum_{\theta \in \Theta_m} \sum_{\Delta \in \mathcal{T}_m, \Delta \subset E_\theta} (|\Delta|^\gamma \|g\|_{L_q(\Delta)})^\tau \\ &\leq c \sum_{\Delta \in \mathcal{T}_m} (|\Delta|^\gamma \|g\|_{L_q(\Delta)})^\tau. \quad \square \end{aligned}$$

Local polynomial approximation is an important tool in spline approximation. For a function $f \in L_q(G)$, $G \subset \mathbb{R}^2$, we denote by $E_k(f, G)_q$ the error of the best L_q -approximation to f on G from Π_k and by $\omega_k(f, G)_q$ the k th local modulus of smoothness of f on G :

$$E_k(f, G)_q := \inf_{P \in \Pi_k} \|f - P\|_{L_q(G)}, \quad \omega_k(f, G)_q := \sup_{h \in \mathbb{R}^2} \|\Delta_h^k(f, \cdot)\|_{L_q(G)}.$$

Whitney's theorem gives an important relation between these two quantities: If $f \in L_q(G)$, $0 < q \leq \infty$, where $G = \Delta$ is an arbitrary triangle or $G = \Omega_\Delta$ with $\Delta \in \mathcal{T}$, \mathcal{T} is an SLR-triangulation, then

$$(2.16) \quad E_k(f, G)_q \leq c \omega_k(f, G)_q,$$

where $c = c(q, k)$ if $G = \Delta$ and $c = c(q, k, \delta_2)$ if $G = \Omega_\Delta$ (δ_2 is from (2.3)). For a proof of this estimate, see, e.g., the appendix of [38]. Note that this estimate holds

for much more general regions G , but then the constant $c = c(G)$ may become hard to control.

For $0 < q \leq \infty$ and a triangle Δ , we let $P_{\Delta,q} : L_q(\Delta) \rightarrow \Pi_k$ be a projector such that

$$(2.17) \quad \|f - P_{\Delta,q}(f)\|_{L_q(\Delta)} \leq cE_k(f, \Delta)_q \quad \text{for } f \in L_q(\Delta).$$

Note that $P_{\Delta,q}$ can be realized as a linear projector if $q \geq 1$. For instance, one can utilize the averaged Taylor polynomial. Namely, suppose Δ_0 is an equilateral reference triangle and A is an affine transform mapping Δ onto Δ_0 . Let now $P(g) \in \Pi_k$ be the averaged Taylor polynomial of the function $g := f \circ A^{-1}$ (the composition of f with A^{-1}) over the disc B inscribed in Δ_0 (see, e.g., section 4.1 of [12]). Clearly, $P : L_q(B) \rightarrow \Pi_k$ is a linear operator, $\|P(g)\|_{L_q(B)} \leq c\|g\|_{L_q(B)}$ ($q \geq 1$), and P is a projector, i.e., $P(Q) = Q$ for $Q \in \Pi_k$. From these properties of P , it follows that for an arbitrary $Q \in \Pi_k$,

$$\begin{aligned} \|g - P(g)\|_{L_q(\Delta_0)} &\leq \|g - Q\|_{L_q(\Delta_0)} + \|Q - P(g)\|_{L_q(\Delta_0)} \\ &\leq \|g - Q\|_{L_q(\Delta_0)} + c\|P(g - Q)\|_{L_q(B)} \leq c\|g - Q\|_{L_q(\Delta_0)}, \end{aligned}$$

which implies $\|g - P(g)\|_{L_q(\Delta_0)} \leq cE_k(g, \Delta_0)_q$. Substituting back, one easily obtains $\|f - (P \circ A)(f)\|_{L_q(\Delta)} \leq cE_k(f, \Delta)_q$. Finally, we set $P_{\Delta,q} := P \circ A$, which is the desired linear projector of $L_q(\Delta)$ into Π_k .

Note that $P_{\Delta,q}$ cannot be realized as a linear operator if $0 < q < 1$ (otherwise, we would be able to construct a nonzero bounded linear functional on L_q).

We define a linear operator $Q_m : \mathcal{S}^{k,-1}(\mathcal{T}_m) \rightarrow \mathcal{S}_m$ as follows. For each $\theta \in \Theta_m$, let $\lambda_\theta : \mathcal{S}^{k,-1}(\mathcal{T}_m)|_{E_\theta} \rightarrow \mathbb{R}$ be a linear functional such that

$$\lambda_\theta(s|_{E_\theta}) = a_\theta(s), \quad s \in \mathcal{S}_m, \quad \text{and}$$

$$|\lambda_\theta(f)| \leq M_2\|f\|_{L_\infty(E_\theta)}, \quad f \in \mathcal{S}^{k,-1}(\mathcal{T}_m)|_{E_\theta}.$$

Such linear functional always exists by the Hahn–Banach theorem. We set

$$(2.18) \quad Q_m(s) := \sum_{\theta \in \Theta_m} \lambda_\theta(s|_{E_\theta})\varphi_\theta, \quad s \in \mathcal{S}^{k,-1}(\mathcal{T}_m).$$

Clearly, $Q_m(s) = s$ if $s \in \mathcal{S}_m$, and thus Q_m is a linear projector of $\mathcal{S}^{k,-1}(\mathcal{T}_m)$ into \mathcal{S}_m .

LEMMA 2.4. *For any $s \in \mathcal{S}^{k,-1}(\mathcal{T}_m)$, $0 < q \leq \infty$, and $\Delta \in \mathcal{T}_m$,*

$$(2.19) \quad \|Q_m(s)\|_{L_q(\Delta)} \leq c\|s\|_{L_q(\Omega_\Delta^e)},$$

with a constant c independent of m , Δ , and s .

Proof. By Lemma 2.2 and (2.14), we have

$$\begin{aligned} \|\varphi_\theta\|_{L_q(\Delta)} &\leq c_1|\Delta|^{1/q}\|\varphi_\theta\|_{L_\infty(\Delta)} \leq c_1M_1|\Delta|^{1/q}, \\ \|s\|_{L_\infty(E_\theta)} &\leq c_2|\Delta|^{-1/q}\|s\|_{L_q(E_\theta)}, \end{aligned}$$

where c_1 and c_2 depend only on q and k . Therefore,

$$\begin{aligned} \|Q_m(s)\|_{L_q(\Delta)} &= \left\| \sum_{\substack{\theta \in \Theta_m \\ \Delta \subset E_\theta}} \lambda_\theta(s|_{E_\theta})\varphi_\theta \right\|_{L_q(\Delta)} \leq c \sum_{\substack{\theta \in \Theta_m \\ \Delta \subset E_\theta}} |\lambda_\theta(s|_{E_\theta})| \|\varphi_\theta\|_{L_q(\Delta)} \\ &\leq c \sum_{\substack{\theta \in \Theta_m \\ \Delta \subset E_\theta}} \|s\|_{L_\infty(E_\theta)} |\Delta|^{1/q} \leq c \sum_{\substack{\theta \in \Theta_m \\ \Delta \subset E_\theta}} \|s\|_{L_q(E_\theta)} \leq c\|s\|_{L_q(\Omega_\Delta^e)}. \quad \square \end{aligned}$$

We now extend Q_m to $L_q^{\text{loc}}(\mathbb{R}^2)$, $0 < q \leq \infty$. Let $P_{\Delta,q} : L_q(\Delta) \rightarrow \Pi_k$ be a projector satisfying (2.17) (linear if $q \geq 1$). We define

$$(2.20) \quad p_{m,q}(f) := \sum_{\Delta \in \mathcal{T}_m} \mathbb{1}_{\Delta} \cdot P_{\Delta,q}(f) \quad \text{for } f \in L_q^{\text{loc}},$$

which is a projector of L_q^{loc} into $\mathcal{S}_m^{k,-1}$.

We put

$$(2.21) \quad Q_{m,q}(f) := Q_m(p_{m,q}(f)) \quad \text{for } f \in L_q^{\text{loc}},$$

which is evidently a projector of L_q^{loc} into \mathcal{S}_m (linear if $q \geq 1$ and all $P_{\Delta,q}$ are linear).

We next show that $Q_{m,q}$ provides a good local L_q -approximation from \mathcal{S}_m . We let $\mathbb{S}_{\Delta}(f)_q$ denote the error of $L_q(\Omega_{\Delta}^{\ell})$ -approximation from \mathcal{S}_m , i.e.,

$$(2.22) \quad \mathbb{S}_{\Delta}(f)_q := \inf_{s \in \mathcal{S}_m} \|f - s\|_{L_q(\Omega_{\Delta}^{\ell})}, \quad \Delta \in \mathcal{T}_m.$$

Thus, $\mathbb{S}_{\Delta}(f)_q$ is the error of approximation to f from *restrictions* to Ω_{Δ}^{ℓ} of functions from \mathcal{S}_m , which is not necessarily the same as the approximation by all r times differentiable piecewise polynomials of degree $< k$ defined only on Ω_{Δ}^{ℓ} , even if \mathcal{S}_m coincides with $\mathcal{S}_m^{k,r}$. However, since $\Pi_{\tilde{k}} \subset \mathcal{S}_m$, $\mathbb{S}_{\Delta}(f)_q$ does not exceed the error of $L_q(\Omega_{\Delta}^{\ell})$ -approximation to f from polynomials of degree $< \tilde{k}$.

LEMMA 2.5. *If $f \in L_q^{\text{loc}}(\mathbb{R}^2)$, $0 < q \leq \infty$ ($f \in C$ if $q = \infty$), then*

$$\|f - Q_{m,q}(f)\|_{L_q(\Delta)} \leq c \mathbb{S}_{\Delta}(f)_q, \quad \Delta \in \mathcal{T}_m \quad (m \in \mathbb{Z}),$$

with c independent of f , m , and Δ .

Proof. Let $s_{\Delta} \in \mathcal{S}_m$ be such that $\|f - s_{\Delta}\|_{L_q(\Omega_{\Delta}^{\ell})} \leq c \mathbb{S}_{\Delta}(f)_q$. Using the properties of Q_m (see Lemma 2.4), we find

$$\begin{aligned} \|f - Q_{m,q}(f)\|_{L_q(\Delta)} &= \|f - Q_m(p_{m,q}(f))\|_{L_q(\Delta)} \\ &\leq c \|f - s_{\Delta}\|_{L_q(\Delta)} + c \|s_{\Delta} - Q_m(p_{m,q}(f))\|_{L_q(\Delta)} \\ &\leq c \mathbb{S}_{\Delta}(f)_q + c \|Q_m(s_{\Delta} - p_{m,q}(f))\|_{L_q(\Delta)} \\ &\leq c \mathbb{S}_{\Delta}(f)_q + c \|s_{\Delta} - p_{m,q}(f)\|_{L_q(\Omega_{\Delta}^{\ell})} \\ &\leq c \mathbb{S}_{\Delta}(f)_q + c \|f - s_{\Delta}\|_{L_q(\Omega_{\Delta}^{\ell})} + c \|f - p_{m,q}(f)\|_{L_q(\Omega_{\Delta}^{\ell})} \\ &\leq c \mathbb{S}_{\Delta}(f)_q. \quad \square \end{aligned}$$

LEMMA 2.6. (a) *If $f \in L_q^{\text{loc}}(\mathbb{R}^2)$, $0 < q \leq \infty$, then for every compact $K \subset \mathbb{R}^2$,*

$$(2.23) \quad \|f - Q_{m,q}(f)\|_{L_q(K)} \rightarrow 0 \quad \text{as } m \rightarrow \infty.$$

(b) *If $f \in L_q(\mathbb{R}^2)$, $0 < q \leq \infty$, then*

$$(2.24) \quad \|f - Q_{m,q}(f)\|_{L_q(\mathbb{R}^2)} \rightarrow 0 \quad \text{as } m \rightarrow \infty.$$

For the proof of this lemma, we need the following result.

LEMMA 2.7. *If \mathcal{T} is an LR-triangulation, then for each triangle $\Delta^{\diamond} \in \mathcal{T}$*

$$(2.25) \quad \max\{\text{diam}(\Delta) : \Delta \in \mathcal{T}_m, \Delta \subset \Delta^{\diamond}\} \rightarrow 0 \quad \text{as } m \rightarrow \infty.$$

Proof. Let $m_0 := \text{level}(\Delta^\diamond)$. We set $d_m := \max\{\text{diam}(\Delta) : \Delta \in \mathcal{T}_m, \Delta \subset \Delta^\diamond\}$. Since $(d_m)_{m=m_0}^\infty$ is nonincreasing, it suffices to show the existence of a subsequence tending to zero. Let e be an edge of a triangle $\Delta \in \mathcal{T}_m, \Delta \subset \Delta^\diamond$. If it is also an edge of a child of Δ , then the valence of at least one of the two endpoints of e will increase by one at level $m + 1$. (Recall that there are always at least two children, so that a child and a parent cannot be the same triangle.) Therefore, e will be subdivided at least once after at most $S := 2(N_0 - 3) + 1$ steps of refinement. By (2.1), it readily follows that any edge e' obtained by subdividing e satisfies $|e'| \leq \rho|e| \leq \rho d_m$.

We call an edge of a descendant of Δ^\diamond a *cutting edge* for Δ^\diamond if one of its endpoints is a vertex of Δ^\diamond and the other lies in the interior of the opposite edge of Δ^\diamond . Since all cutting edges must emanate from the same vertex of Δ^\diamond , there are totally no more than $M := N_0 - 3$ such edges for Δ^\diamond . Therefore, no *new* cutting edges for Δ^\diamond will be created at levels $m > m_0 + M$. (It is easy to see that as soon as no new cutting edges are created at a level m , they cannot be created on any further level.) Using this and the above observation, we conclude that there will be no cutting edges at levels $m > m_0 + M + S$ since they all will be subdivided. Therefore, each edge e inside Δ^\diamond at these levels is either a proper part of an edge of Δ^\diamond or has both of its endpoints in the interiors of two different edges of Δ^\diamond , or else it has at least one endpoint in the interior of Δ^\diamond . In all cases, condition (2.1) ensures that $|e| \leq \rho d_{m_0}$, which implies $d_{m_1} \leq \rho d_{m_0}$, where $m_1 = m_0 + M + S + 1$. It is clear now that there is an increasing sequence $\{m_k\}_{k=1}^\infty$ such that

$$d_{m_k} \leq \rho^k d_{m_0} \rightarrow 0 \quad \text{as } k \rightarrow \infty,$$

which completes the proof. \square

Proof of Lemma 2.6. (a) By condition (f) on triangulations, it suffices to prove the lemma for $K = \Delta^\diamond$, an arbitrary triangle from \mathcal{T} . By Lemma 2.7,

$$(2.26) \quad \max\{\text{diam}(\Omega_\Delta^\ell) : \Delta \in \mathcal{T}_m, \Delta \subset \Omega_{\Delta^\diamond}^\ell\} \rightarrow 0 \quad \text{as } m \rightarrow \infty.$$

Case 1. $q < \infty$. Fix $\varepsilon > 0$. In view of (2.26), there exists a piecewise constant function S_ε of the form

$$S_\varepsilon = \sum_{\Delta \in \mathcal{T}_{m_\varepsilon}, \Delta \subset \Omega_{\Delta^\diamond}^\ell} c_\Delta \mathbf{1}_\Delta, \quad m_\varepsilon \geq \text{level}(\Delta^\diamond),$$

such that

$$(2.27) \quad \|f - S_\varepsilon\|_{L_q(\Omega_{\Delta^\diamond}^\ell)} < \varepsilon$$

(choose first $g \in C(\Omega_{\Delta^\diamond}^\ell)$ so that $\|f - g\|_{L_q(\Omega_{\Delta^\diamond}^\ell)} < \varepsilon/2$ and then choose S_ε so that $\|g - S_\varepsilon\|_{L_\infty(\Omega_{\Delta^\diamond}^\ell)} < \frac{\varepsilon}{2} |\Omega_{\Delta^\diamond}^\ell|^{-1/q}$). Then $Q_{m,q}(S_\varepsilon) = Q_m(S_\varepsilon)$.

We have, for $m \geq m_\varepsilon$,

$$(2.28) \quad \begin{aligned} \|f - Q_{m,q}(f)\|_{L_q(\Delta^\diamond)} &\leq c\|f - S_\varepsilon\|_{L_q(\Delta^\diamond)} + c\|S_\varepsilon - Q_{m,q}(S_\varepsilon)\|_{L_q(\Delta^\diamond)} \\ &\quad + c\|Q_m(S_\varepsilon - p_{m,q}(f))\|_{L_q(\Delta^\diamond)}. \end{aligned}$$

For the third term above, we have

$$(2.29) \quad \begin{aligned} \|Q_m(S_\varepsilon - p_{m,q}(f))\|_{L_q(\Delta^\diamond)} &\leq c\|S_\varepsilon - p_{m,q}(f)\|_{L_q(\Omega_{\Delta^\diamond}^\ell)} \\ &\leq c\|f - S_\varepsilon\|_{L_q(\Omega_{\Delta^\diamond}^\ell)} + c\|f - p_{m,q}(f)\|_{L_q(\Omega_{\Delta^\diamond}^\ell)} \\ &\leq c\|f - S_\varepsilon\|_{L_q(\Omega_{\Delta^\diamond}^\ell)} \leq c\varepsilon, \end{aligned}$$

where we used Lemma 2.4 and that $\|f - p_{m,q}(f)\|_{L_q(\Omega_{\Delta^\circ}^\ell)} \leq c\|f - S_\varepsilon\|_{L_q(\Omega_{\Delta^\circ}^\ell)}$ ($m \geq m_\varepsilon$), by (2.17).

It remains to show that $\|S_\varepsilon - Q_{m,q}(S_\varepsilon)\|_{L_q(\Delta^\circ)} \leq c\varepsilon$ for sufficiently large m . Denote by G the union of the edges of all triangles $\Delta \in \mathcal{T}_{m_\varepsilon}$ such that $\Delta \subset \Delta^\circ$ and by $G_\delta := \{x \in \mathbb{R}^2 : \text{dist}(x, G) \leq \delta\}$ the δ -neighborhood of G . Clearly, there exists $\delta > 0$ such that $\|S_\varepsilon\|_{L_q(G_\delta)} < \varepsilon$.

By (2.26), there exists $m_1 \geq m_\varepsilon$ such that $\text{diam}(\Omega_\Delta^\ell) < \delta$ for all triangles $\Delta \in T_m$ ($m \geq m_1$) such that $\Delta \subset \Delta^\circ$ and $\Omega_\Delta^\ell \cap G \neq \emptyset$. Since $\Pi_1 \subset \mathcal{S}_m$, $Q_m(S_\varepsilon)|_\Delta = S_\varepsilon|_\Delta$ if $S_\varepsilon|_{\Omega_\Delta^\ell} = \text{constant}$. Using this, we obtain by Lemma 2.5

$$\|S_\varepsilon - Q_{m,q}(S_\varepsilon)\|_{L_q(\Delta^\circ)} \leq c \left(\sum_{\Delta \in T_m, \Omega_\Delta^\ell \cap G \neq \emptyset} \mathbb{S}_\Delta(S_\varepsilon)_q^q \right)^{1/q} \leq c\|S_\varepsilon\|_{L_q(G_\delta)} \leq c\varepsilon.$$

We substitute this estimate together with (2.27) and (2.29) in (2.28) to obtain

$$\|f - Q_{m,q}(f)\|_{L_q(\Delta^\circ)} \leq c\varepsilon \quad \text{for } m \geq m_1.$$

This implies (2.23) if $q < \infty$.

Case 2. $q = \infty$. We have, by Lemma 2.5 and the fact that $\Pi_1 \subset \mathcal{S}_m$,

$$\|f - Q_{m,q}(f)\|_{L_\infty(\Delta^\circ)} \leq c \max_{\Delta \in T_m, \Delta \subset \Delta^\circ} \inf_{C \in \Pi_1} \|f - C\|_{L_\infty(\Omega_\Delta^\ell)}.$$

Now the result follows, using (2.26) and the fact that f is uniformly continuous on $\Omega_{\Delta^\circ}^\ell$.

Part (b) of the lemma is immediate from part (a). \square

We denote $\mathcal{S}_{-\infty} := \bigcap_{m \in \mathbb{Z}} \mathcal{S}_m$. As we already mentioned, there are only two possibilities for $\mathcal{T}_{-\infty}$: $\mathcal{T}_{-\infty} = \{\mathbb{R}^2\}$ or $\mathcal{T}_{-\infty} = (\Delta_\infty^j)_{j=1}^{N_\infty}$, $N_\infty \leq N_0$, where $\{\Delta_\infty^j\}$ are infinite triangles with disjoint interiors and a common vertex which cover \mathbb{R}^2 . If $\mathcal{T}_{-\infty} = \{\mathbb{R}^2\}$, then obviously \mathbb{R}^2 is the union of a sequence of nested triangles, and hence each $s \in \mathcal{S}_{-\infty}$ is a polynomial of degree $< k$ on \mathbb{R}^2 . Therefore, if $\mathcal{T}_{-\infty} = \{\mathbb{R}^2\}$, then $\mathcal{S}_{-\infty}$ a subspace of Π_k .

Suppose $\mathcal{T}_{-\infty} = (\Delta_\infty^j)_{j=1}^{N_\infty}$ and $s \in \mathcal{S}_{-\infty}$. Then each triangle Δ_∞^j can be represented as the union of a sequence of nested triangles, and hence s is a polynomial of degree $< k$ on Δ_∞^j . Therefore, in this case, $s \in \mathcal{S}_{-\infty}$ implies $s \in C^r(\mathbb{R}^2)$ and $s|_{\Delta_\infty^j} = P_j|_{\Delta_\infty^j}$ for some $P_j \in \Pi_k$, $j = 1, \dots, N_\infty$.

Furthermore, if $s \in \mathcal{S}_{-\infty}$ and $|\{x \in \mathbb{R}^2 : |s(x)| > t\}| < \infty$ for some $t > 0$, then $s = \text{const}$. In particular, if $s \in \mathcal{S}_{-\infty} \cap L_p$ ($p < \infty$), then $s \equiv 0$.

2.3. Definition of B-spaces. Equivalent norms. Interpolation. Suppose \mathcal{T} is an LR(or better)-triangulation and $\Phi = \Phi_{\mathcal{T}}$ is a family of differentiable piecewise polynomial basis functions over \mathcal{T} as described in sections 2.1–2.2. For the characterization of nonlinear n -term L_p -approximation from Φ , we need the *B-spaces* $B_\tau^\alpha(\Phi)$ which we shall introduce and explore in this subsection. In fact, the spaces $B_\tau^\alpha(\Phi)$ depend only on the underlying ladder of spaces $\dots \subset \mathcal{S}_{-1} \subset \mathcal{S}_0 \subset \mathcal{S}_1 \subset \dots$ associated with the bases $(\Phi_m)_{m \in \mathbb{Z}}$, but as it will be shown below these spaces have atomic representations using Φ , which justifies our notation.

We shall need the B-spaces $B_\tau^\alpha(\Phi)$ in two cases: (a) $0 < p < \infty$ and $\alpha > 0$, or (b) $p = \infty$ and $\alpha \geq 1$ (see Remark 2.14). In both cases, we define τ from the identity $1/\tau = \alpha + 1/p$ ($1/\infty := 0$).

Definition of $B_\tau^\alpha(\Phi)$ via local approximation. We define the B-space $B_\tau^\alpha(\Phi)$ as the set of all functions $f \in L_\tau(\mathbb{R}^2)$ such that

$$(2.30) \quad \|f\|_{B_\tau^\alpha(\Phi)} := \left(\sum_{\Delta \in \mathcal{T}} (|\Delta|^{-\alpha} \mathbb{S}_\Delta(f)_\tau)^\tau \right)^{1/\tau} < \infty,$$

where $\mathbb{S}_\Delta(f)_\tau$ is the error of L_τ -approximation of f on Ω_Δ^ℓ from \mathcal{S}_m if $\Delta \in \mathcal{T}_m$ (see (2.22)).

It is readily seen that $B_\tau^\alpha(\Phi)$ is a linear space, $\|cf\|_{B_\tau^\alpha} = |c|\|f\|_{B_\tau^\alpha}$, and $\|f + g\|_{B_\tau^\alpha}^\lambda \leq \|f\|_{B_\tau^\alpha}^\lambda + \|g\|_{B_\tau^\alpha}^\lambda$, with $\lambda := \min\{\tau, 1\}$. Clearly, see Theorem 2.8, if $\|f\|_{B_\tau^\alpha} = 0$, then $f = 0$ a.e. Therefore, $\|\cdot\|_{B_\tau^\alpha}$ is a norm if $\tau \geq 1$ and a quasi norm if $\tau < 1$.

We next define other equivalent norms in $B_\tau^\alpha(\Phi)$. We define

$$(2.31) \quad N_{\Phi, \mathbb{S}, \eta}(f) := \left(\sum_{\Delta \in \mathcal{T}} (|\Delta|^{1/p-1/\eta} \mathbb{S}_\Delta(f)_\eta)^\tau \right)^{1/\tau},$$

where we have taken into account that $1/\tau := \alpha + 1/p$. Thus, $N_{\Phi, \mathbb{S}, \tau}(f) = \|f\|_{B_\tau^\alpha(\Phi)}$. Moreover, we shall show that $N_{\Phi, \mathbb{S}, \eta}(f) \approx \|f\|_{B_\tau^\alpha(\Phi)}$ if $0 < \eta < p$ (see Theorem 2.10).

Definition of norms in $B_\tau^\alpha(\Phi)$ via basis functions (atomic decomposition). For $f \in L_\tau(\mathbb{R}^2)$, we define

$$(2.32) \quad N_\Phi(f) := \inf_{f = \sum_{\theta \in \Theta} c_\theta \varphi_\theta} \left(\sum_{\theta \in \Theta} (|E_\theta|^{-\alpha} \|c_\theta \varphi_\theta\|_\tau)^\tau \right)^{1/\tau},$$

where the infimum is over all representations of f in the form $f = \sum_{\theta \in \Theta} c_\theta \varphi_\theta$ in L_τ . (Note that the existence of such representations for each $f \in L_\tau$ follows by Lemma 2.6.) By Theorem 2.9,

$$\sum_{\theta \in \Theta} (|E_\theta|^{-\alpha} \|c_\theta \varphi_\theta\|_\tau)^\tau < \infty \quad \text{implies} \quad \left\| \sum_{\theta \in \Theta} |c_\theta \varphi_\theta(\cdot)| \right\|_p < \infty,$$

and hence $\sum_{\theta \in \Theta} c_\theta \varphi_\theta(x)$ converges absolutely a.e. Therefore, the specific type of convergence that we use in the definition of $N_\Phi(f)$ above is not essential. Using (2.14), we have

$$(2.33) \quad \begin{aligned} N_\Phi(f) &\approx \inf_{f = \sum_{\theta \in \Theta} c_\theta \varphi_\theta} \left(\sum_{\theta \in \Theta} (|E_\theta|^{1/p-1/\eta} \|c_\theta \varphi_\theta\|_\eta)^\tau \right)^{1/\tau} \\ &\approx \inf_{f = \sum_{\theta \in \Theta} c_\theta \varphi_\theta} \left(\sum_{\theta \in \Theta} \|c_\theta \varphi_\theta\|_p^\tau \right)^{1/\tau}. \end{aligned}$$

Definition of norms in $B_\tau^\alpha(\Phi)$ via projections. For $f \in L_\eta^{\text{loc}}$, we set

$$(2.34) \quad q_{m, \eta}(f) := Q_{m, \eta}(f) - Q_{m-1, \eta}(f) \in \mathcal{S}_m,$$

where $Q_{m, \eta}$ is from (2.21), and let $(b_{\theta, \eta}(f))_{\theta \in \Theta_m}$ be defined by the identity

$$(2.35) \quad q_{m, \eta}(f) = \sum_{\theta \in \Theta_m} b_{\theta, \eta}(f) \varphi_\theta, \quad \text{i.e.,} \quad b_{\theta, \eta}(f) := a_\theta(q_{m, \eta}(f)), \quad \theta \in \Theta_m.$$

We define

$$(2.36) \quad N_{\Phi, Q, \tau}(f) := \left(\sum_{\theta \in \Theta} (|E_\theta|^{-\alpha} \|b_{\theta, \tau}(f)\varphi_\theta\|_\tau)^\tau \right)^{1/\tau}$$

and, more generally (see (2.31)),

$$(2.37) \quad N_{\Phi, Q, \eta}(f) := \left(\sum_{\theta \in \Theta} (|E_\theta|^{1/p-1/\eta} \|b_{\theta, \eta}(f)\varphi_\theta\|_\eta)^\tau \right)^{1/\tau}, \quad 0 < \eta < p.$$

By Lemmas 2.2–2.3, it follows that

$$(2.38) \quad N_{\Phi, Q, \eta}(f) \approx \left(\sum_{m \in \mathbb{Z}} \sum_{\Delta \in \mathcal{T}_m} (|\Delta|^{1/p-1/\eta} \|q_{m, \eta}(f)\|_{L_\eta(\Delta)})^\tau \right)^{1/\tau}$$

and, for $0 < \mu \leq \infty$,

$$(2.39) \quad N_{\Phi, Q, \eta}(f) \approx \left(\sum_{\theta \in \Theta} (|E_\theta|^{1/p-1/\mu} \|b_{\theta, \eta}(f)\varphi_\theta\|_\mu)^\tau \right)^{1/\tau} \approx \left(\sum_{\theta \in \Theta} \|b_{\theta, \eta}(f)\varphi_\theta\|_p^\tau \right)^{1/\tau}.$$

We shall show (see Theorem 2.10) that all of the above norms are equivalent. To this end, we need the following embedding theorem.

THEOREM 2.8. *If $f \in L_\tau(\mathbb{R}^2)$ and $N_{\Phi, Q, \eta}(f) < \infty$, $0 < \eta < p$, then*

$$(2.40) \quad f = \sum_{m \in \mathbb{Z}} q_{m, \eta}(f) = \sum_{\theta \in \Theta} b_{\theta, \eta}(f)\varphi_\theta,$$

with the series converging absolutely a.e. and in L_p , and

$$(2.41) \quad \|f\|_p \leq c \left\| \sum_{m \in \mathbb{Z}} |q_{m, \eta}(f)(\cdot)| \right\|_p \leq c \left\| \sum_{\theta \in \Theta} |b_{\theta, \eta}(f)\varphi_\theta(\cdot)| \right\|_p \leq c N_{\Phi, Q, \eta}(f),$$

with c independent of f .

The proof of Theorem 2.8 hinges on the following more general embedding theorem, which is a special case of Theorem 2.5 from [54].

THEOREM 2.9. *If $0 < \tau < p < \infty$, or $p = \infty$, and $0 < \tau \leq 1$, then for any sequence of real numbers $(c_\theta)_{\theta \in \Theta}$ we have*

$$(2.42) \quad \left\| \sum_{\theta \in \Theta} |c_\theta \varphi_\theta(\cdot)| \right\|_p \leq c \left(\sum_{\theta \in \Theta} \|c_\theta \varphi_\theta\|_p^\tau \right)^{1/\tau},$$

with c independent of $(c_\theta)_{\theta \in \Theta}$.

For completeness, we give the simple proof of this theorem in the appendix.

Proof of Theorem 2.8. We introduce the following abbreviated notation: $Q_m := Q_{m, \eta}(f)$, $q_m := q_{m, \eta}(f)$, $b_\theta := b_{\theta, \eta}(f)$, and $N(f) := N_{\Phi, Q, \eta}(f)$. By (2.35), (2.39), and Theorem 2.9, we have

$$(2.43) \quad \left\| \sum_{m \in \mathbb{Z}} |q_m(\cdot)| \right\|_p \leq c \left\| \sum_{\theta \in \Theta} |b_\theta \varphi_\theta(\cdot)| \right\|_p \leq c N(f) < \infty,$$

and hence $\sum_{m \in \mathbb{Z}} |q_m(x)| < \infty$ a.e. On the other hand, by Lemma 2.6, we have $\|f - Q_m\|_{L_\eta(\Delta)} \rightarrow 0$ as $m \rightarrow \infty$ for each $\Delta \in \mathcal{T}$. The above two facts imply

$$(2.44) \quad f - Q_0 = \sum_{m=1}^{\infty} q_m \quad \text{absolutely a.e. on } \mathbb{R}^2.$$

We use Lemmas 2.1 and 2.2 to obtain, for $\Delta \in \mathcal{T}_m$ ($m \in \mathbb{Z}$),

$$\|q_m\|_{L_\infty(\Delta)} \leq c|\Delta|^{-\frac{1}{p}} \|q_m\|_{L_p(\Delta)} \leq c|\Delta|^{-\frac{1}{p}} \sum_{\theta \in \Theta_m, E_\theta \supset \Delta} \|b_\theta \varphi_\theta\|_p \leq c|\Delta|^{-\frac{1}{p}} N(f).$$

Therefore, for a fixed $\Delta' \in \mathcal{T}_\nu$ ($\nu \in \mathbb{Z}$),

$$(2.45) \quad \begin{aligned} \sum_{m=-\infty}^{\nu} \|q_m\|_{L_\infty(\Delta')} &\leq cN(f) \sum_{\Delta \in \mathcal{T}, \Delta \supset \Delta'} |\Delta|^{-1/p} \\ &= cN(f) |\Delta'|^{-1/p} \sum_{\Delta \in \mathcal{T}, \Delta \supset \Delta'} (|\Delta'|/|\Delta|)^{1/p} \\ &\leq c|\Delta'|^{-1/p} N(f) < \infty, \end{aligned}$$

where we used (2.9). We set

$$(2.46) \quad s_\infty := Q_0 - \sum_{m=-\infty}^0 q_m \quad \text{pointwise in } \mathbb{R}^2.$$

From (2.45), it follows that s_∞ is well defined and the series in (2.46) converges uniformly on every compact in \mathbb{R}^2 . Evidently, (2.46) yields $s_\infty = Q_\nu - \sum_{m=-\infty}^{\nu} q_m$ for each $\nu \in \mathbb{Z}$.

Fix $n \in \mathbb{Z}$. Using Theorem 2.9, we obtain, for $\nu \leq n$,

$$\begin{aligned} \inf_{s \in \mathcal{S}_n} \|s_\infty - s\|_p &\leq \|s_\infty - Q_\nu\|_p = \left\| \sum_{m=-\infty}^{\nu} q_m \right\|_p \\ &\leq c \left(\sum_{\theta \in \bigcup_{m=-\infty}^{\nu} \Theta_m} \|b_\theta \varphi_\theta\|_p^\tau \right)^{1/\tau} \rightarrow 0 \quad \text{as } \nu \rightarrow -\infty, \end{aligned}$$

where we used that $(\sum_{\theta \in \Theta} \|b_\theta \varphi_\theta\|_p^\tau)^{1/\tau} \approx N(f) < \infty$. Therefore, $s_\infty \in \mathcal{S}_n$ for every $n \in \mathbb{Z}$, and hence $s_\infty \in \bigcap_{n \in \mathbb{Z}} \mathcal{S}_n = \mathcal{S}_{-\infty}$.

Identities (2.44) and (2.46) yield

$$(2.47) \quad f - s_\infty = \sum_{m \in \mathbb{Z}} q_{m,\eta}(f) = \sum_{\theta \in \Theta} b_{\theta,\eta}(f) \varphi_\theta \quad \text{absolutely a.e.,}$$

and hence, using (2.43),

$$(2.48) \quad \begin{aligned} \|f - s_\infty\|_p &\leq c \left\| \sum_{m \in \mathbb{Z}} |q_{m,\eta}(f)(\cdot)| \right\|_p \\ &\leq c \left\| \sum_{\theta \in \Theta} |b_{\theta,\eta}(f) \varphi_\theta(\cdot)| \right\|_p \leq cN_{\Phi,Q,\eta}(f) < \infty. \end{aligned}$$

Since $f \in L_\tau$ and $f - s_\infty \in L_p$, it readily follows that, for $t > 0$,

$$\begin{aligned} |\{x : |s_\infty(x)| > t\}| &\leq |\{x : |f(x)| > t/2\}| + |\{x : |f(x) - s_\infty(x)| > t/2\}| \\ &\leq (t/2)^{-\tau} \|f\|_\tau^\tau + (t/2)^{-p} \|f - s_\infty\|_p^p < \infty, \end{aligned}$$

which implies $s_\infty \equiv 0$ (see the end of section 2.2). From this, (2.47), and (2.48), we infer (2.40) and (2.41). The proof is complete. \square

THEOREM 2.10. *The norms $\|\cdot\|_{B_\tau(\Phi)}$, $N_{\Phi, \mathbb{S}, \eta}(\cdot)$ ($0 < \eta < p$), $N_\Phi(\cdot)$, and $N_{\Phi, Q, \eta}(\cdot)$ ($0 < \eta < p$), defined in (2.30)–(2.32) and (2.37), are equivalent with constants of equivalence depending only on p , α , η , and the parameters of \mathcal{T} and Φ .*

Proof. Theorem 2.8 readily implies

$$(2.49) \quad N_\Phi(f) \leq N_{\Phi, Q, \eta}(f), \quad 0 < \eta < p,$$

if $N_{\Phi, Q, \eta}(f) < \infty$.

Suppose $N_{\Phi, \mathbb{S}, \eta}(f) < \infty$. For each $\Delta \in \mathcal{T}_m$ ($m \in \mathbb{Z}$), we have, by (2.34) and Lemma 2.5,

$$\|q_{m, \eta}(f)\|_{L_\eta(\Delta)} \leq c \|f - Q_{m, \eta}\|_{L_\eta(\Delta)} + c \|f - Q_{m-1, \eta}\|_{L_\eta(\Delta)} \leq c \mathbb{S}_\Delta(f)_\eta + c \mathbb{S}_{\Delta^\circ}(f)_\eta,$$

where $\Delta^\circ \supset \Delta$, $\Delta^\circ \in \mathcal{T}_{m-1}$, is the only parent of Δ . These estimates readily imply

$$(2.50) \quad N_{\Phi, Q, \eta}(f) \leq N_{\Phi, \mathbb{S}, \eta}(f), \quad 0 < \eta < p.$$

It remains to prove that

$$(2.51) \quad N_{\Phi, \mathbb{S}, \eta}(f) \leq N_\Phi(f), \quad 0 < \eta < p,$$

provided $N_\Phi(f) < \infty$. Evidently, (2.49)–(2.51) imply the desired equivalence of norms.

Notice first that, by Hölder's inequality, $N_{\Phi, \mathbb{S}, \mu}(f) \leq N_{\Phi, \mathbb{S}, \eta}(f)$ if $0 < \mu \leq \eta$, and hence it suffices to prove (2.51) only for $\tau < \eta < p$.

Suppose $f \in L_\tau$ and $0 < N_\Phi(f) < \infty$. Then it follows by the definition of $N_\Phi(f)$ that there exists a sequence $(c_\theta)_{\theta \in \Theta}$ such that

$$(2.52) \quad f = \sum_{\theta \in \Theta} c_\theta \varphi_\theta \quad \text{in } L_\tau$$

and $(\sum_{\theta \in \Theta} (|E_\theta|^{-\alpha} \|c_\theta \varphi_\theta\|_\tau)^\tau)^{1/\tau} \leq 2N_\Phi(f)$. Theorem 2.9 implies that in (2.52) we have absolute convergence a.e. We next estimate

$$(2.53) \quad N_{\Phi, \mathbb{S}, \eta}(f) := \left(\sum_{\Delta \in \mathcal{T}} [|\Delta|^{1/p-1/\eta} \mathbb{S}_\Delta(f)_\eta]^\tau \right)^{1/\tau},$$

using that $\mathbb{S}_\Delta(g)_\eta = 0$ if $g \in \mathcal{S}_m$ and $\Delta \in \mathcal{T}_m$, and $\mathbb{S}_\Delta(g)_\eta \leq \|g\|_{L_\eta(\Omega_\Delta^\ell)}$, in general.

We denote $f_j := \sum_{\theta \in \Theta_j} c_\theta \varphi_\theta$. Fix $\Delta' \in \mathcal{T}$ and assume that $\Delta' \in \mathcal{T}_m$ ($m \in \mathbb{Z}$). We have, using Theorem 2.9 ($\tau < \eta < \infty$) and (2.14),

$$\begin{aligned} \mathbb{S}_{\Delta'}(f)_\eta^\tau &= \mathbb{S}_{\Delta'} \left(\sum_{j=m+1}^{\infty} f_j \right)_\eta^\tau \leq \left\| \sum_{j=m+1}^{\infty} f_j \right\|_{L_\eta(\Omega_{\Delta'}^\ell)}^\tau \\ &\leq \left\| \sum_{j=m+1}^{\infty} \sum_{\theta \in \Theta_j, E_\theta \subset \Omega_{\Delta'}^{2\ell}} c_\theta \varphi_\theta \right\|_{L_\eta(\Omega_{\Delta'}^\ell)}^\tau \leq c \sum_{\theta \in \Theta, E_\theta \subset \Omega_{\Delta'}^{2\ell}} \|c_\theta \varphi_\theta\|_\eta^\tau \\ &\leq c \sum_{\theta \in \Theta, E_\theta \subset \Omega_{\Delta'}^{2\ell}} |E_\theta|^{\tau(1/\eta-1/\tau)} \|c_\theta \varphi_\theta\|_\tau^\tau. \end{aligned}$$

Substituting this in (2.53), we obtain

$$\begin{aligned} N_{\Phi, \mathcal{S}, \eta}(f)^\tau &\leq c \sum_{\Delta' \in \mathcal{T}} |\Delta'|^{\tau(1/p-1/\eta)} \sum_{\theta \in \Theta, E_\theta \subset \Omega_{\Delta'}^{2\ell}} |E_\theta|^{\tau(1/\eta-1/\tau)} \|c_\theta \varphi_\theta\|_\tau^\tau \\ &= c \sum_{\Delta' \in \mathcal{T}} \sum_{\theta \in \Theta, E_\theta \subset \Omega_{\Delta'}^{2\ell}} (|E_\theta|/|\Delta'|)^{\tau(1/\eta-1/p)} (|E_\theta|^{-\alpha} \|c_\theta \varphi_\theta\|_\tau)^\tau \\ &\leq c \sum_{\theta \in \Theta} (|E_\theta|^{-\alpha} \|c_\theta \varphi_\theta\|_\tau)^\tau \sum_{\Delta' \in \mathcal{T}, \Omega_{\Delta'}^{2\ell} \supset E_\theta} (|E_\theta|/|\Delta'|)^{\tau(1/\eta-1/p)}, \end{aligned}$$

where we once switched the order of summation. By (2.1)–(2.2),

$$\#\{\Delta' \in \mathcal{T}_\nu : \Omega_{\Delta'}^{2\ell} \supset E_\theta\} \leq c(N_0, \ell), \quad \nu \in \mathbb{Z}, \quad \theta \in \Theta,$$

and $|E_\theta| \leq c\rho^j|\Delta'|$ if $E_\theta \subset \Omega_{\Delta'}^{2\ell}$, with $\Delta' \in \mathcal{T}_m$ and $\theta \in \Theta_{m+j}$ ($m \in \mathbb{Z}, j \geq 0$). Using these, we obtain

$$\sum_{\Delta' \in \mathcal{T}, \Omega_{\Delta'}^{2\ell} \supset E_\theta} (|E_\theta|/|\Delta'|)^{\tau(1/\eta-1/p)} \leq c \sum_{j=0}^\infty \rho^{j\tau(1/\eta-1/p)} \leq c < \infty.$$

Therefore, $N_{\Phi, \mathcal{S}, \eta}(f)^\tau \leq c \sum_{\theta \in \Theta} (|E_\theta|^{-\alpha} \|c_\theta \varphi_\theta\|_\tau)^\tau \leq cN_\Phi(f)^\tau$ which yields (2.51). \square

The following embedding result is quite obvious.

THEOREM 2.11. *For $0 < \alpha_0 < \alpha_1$ and $\tau_j := (\alpha_j + 1/p)^{-1}, j = 0, 1$, we have the continuous embedding*

$$(2.54) \quad B_{\tau_1}^{\alpha_1}(\Phi) \subset B_{\tau_0}^{\alpha_0}(\Phi);$$

i.e., if $f \in B_{\tau_1}^{\alpha_1}(\Phi)$, then $f \in B_{\tau_0}^{\alpha_0}(\Phi)$ and $\|f\|_{B_{\tau_0}^{\alpha_0}(\Phi)} \leq c\|f\|_{B_{\tau_1}^{\alpha_1}(\Phi)}$.

Proof. By Theorem 2.8, if $f \in B_{\tau_1}^{\alpha_1}(\Phi)$, then $f \in L_{\tau_1} \cap L_p \subset L_{\tau_0}$. Fix $0 < \eta < p$. Then by (2.39), we have

$$\|f\|_{B_{\tau_j}^{\alpha_j}(\Phi)} \approx \left(\sum_{\theta \in \Theta} \|b_{\theta, \eta}(f)\varphi_\theta\|_p^{\tau_j} \right)^{1/\tau_j}, \quad j = 0, 1,$$

and the theorem follows since $\tau_1 < \tau_0$. \square

Interpolation of B-spaces. We first recall some basic definitions from the real interpolation method. We refer the reader to [3] and [4] as general references for interpolation theory. For a pair of quasi-normed spaces X_0, X_1 , embedded in a Hausdorff space, the space $X_0 + X_1$ is defined as the collection of all functions f that can be represented as $f_0 + f_1$ with $f_0 \in X_0$ and $f_1 \in X_1$. The quasi norm in $X_0 + X_1$ is defined by

$$\|f\|_{X_0+X_1} := \inf_{f=f_0+f_1} \|f_0\|_{X_0} + \|f_1\|_{X_1}.$$

Peetre’s K -functional is defined for each $f \in X_0 + X_1$ and $t > 0$ by

$$(2.55) \quad K(f, t) := K(f, t; X_0, X_1) := \inf_{f=f_0+f_1} \|f_0\|_{X_0} + t\|f_1\|_{X_1}.$$

The real interpolation space $(X_0, X_1)_{\lambda, q}$ with $0 < \lambda < 1$ and $0 < q \leq \infty$ is defined as the set of all $f \in X_0 + X_1$ such that

$$\|f\|_{(X_0, X_1)_{\lambda, q}} := \|f\|_{X_0 + X_1} + \left(\int_0^\infty (t^{-\lambda} K(f, t))^q \frac{dt}{t} \right)^{1/q} < \infty$$

with the L_q -norm replaced by the sup-norm if $q = \infty$.

It is easily seen that if $X_1 \subset X_0$ (X_1 continuously embedded in X_0), then $K(f, t) \approx \|f\|_{X_0}$ for $f \in X_0$ and $t \geq 1$ and, consequently,

$$(2.56) \quad \|f\|_{(X_0, X_1)_{\lambda, q}} \approx \|f\|_{X_0} + \left(\sum_{\nu=0}^{\infty} [2^{\nu\lambda} K(f, 2^{-\nu})]^q \right)^{1/q}.$$

THEOREM 2.12. *Suppose $0 < p < \infty$ and $\alpha_0, \alpha_1 > 0$, or $p = \infty$ and $\alpha_0, \alpha_1 \geq 1$. Let $\tau_j := (\alpha_j + 1/p)^{-1}$, $j = 0, 1$. Then*

$$(2.57) \quad (B_{\tau_0}^{\alpha_0}(\Phi), B_{\tau_1}^{\alpha_1}(\Phi))_{\lambda, \tau} = B_\tau^\alpha(\Phi)$$

with equivalent norms, provided $\alpha = (1 - \lambda)\alpha_0 + \lambda\alpha_1$ with $0 < \lambda < 1$ and $\tau := (\alpha + 1/p)^{-1}$.

Proof. We shall use some ideas from [32]. We may assume that $\alpha_0 < \alpha_1$. We denote briefly $B^\alpha := B_\tau^\alpha(\Phi)$ and $B^{\alpha_j} := B_{\tau_j}^{\alpha_j}(\Phi)$, $j = 0, 1$. Furthermore, we denote by ℓ_q the space of all sequences $\mathbf{a} = (a_\theta)_{\theta \in \Theta}$ of real numbers such that

$$\|\mathbf{a}\|_{\ell_q} := \left(\sum_{\theta \in \Theta} |a_\theta|^q \right)^{1/q} < \infty.$$

We shall utilize the following well-known interpolation result (see, e.g., [3]):

$$(2.58) \quad (\ell_{\tau_0}, \ell_{\tau_1})_{\lambda, \tau} = \ell_\tau, \quad \text{where } \frac{1}{\tau} = \frac{1-\lambda}{\tau_0} + \frac{\lambda}{\tau_1} \text{ with } 0 < \lambda < 1.$$

We fix $0 < \eta < p$. Then we normalize the basis functions from Φ in L_p , that is, $\|\varphi_\theta\|_p = 1$ (we use the same notation for the normalized basis functions). We also renormalize the dual functionals λ_θ in the definition of Q_m in (2.18) accordingly.

We denote by $\mathbf{b}(f) = (b_\theta(f))_{\theta \in \Theta}$ the sequence of numbers defined by (see (2.34)–(2.35))

$$q_{m, \eta}(f) =: \sum_{\theta \in \Theta_m} b_\theta(f) \varphi_\theta, \quad m \in \mathbb{Z} \quad (\|\varphi_\theta\|_p = 1).$$

By Theorem 2.8, Theorem 2.10, and (2.39), if $f \in B^{\alpha_j}$ ($j = 0, 1$), then

$$(2.59) \quad f \stackrel{L_p}{=} \sum_{\theta \in \Theta} b_\theta(f) \varphi_\theta \quad \text{and} \quad \|f\|_{B^{\alpha_j}} \approx \|\mathbf{b}(f)\|_{\ell_{\tau_j}},$$

and similarly for $f \in B^\alpha$.

The theorem will follow by (2.58) and the following lemma.

LEMMA 2.13. *For $f \in B^{\alpha_0} + B^{\alpha_1} = B^\alpha$ ($\alpha_0 < \alpha_1$), we have*

$$(2.60) \quad K(f, t; B^{\alpha_0}, B^{\alpha_1}) \approx K(\mathbf{b}(f), t; \ell_{\tau_0}, \ell_{\tau_1}), \quad t > 0.$$

Proof. We first prove that

$$(2.61) \quad K(f, t; B^{\alpha_0}, B^{\alpha_1}) \leq cK(\mathbf{b}(f), t; \ell_{\tau_0}, \ell_{\tau_1}), \quad t > 0.$$

Indeed, let $\mathbf{a} = (a_\theta)_{\theta \in \Theta} \in \ell_{\tau_1}$. Then $\mathbf{a} \in \ell_{\tau_0}$ ($\tau_0 > \tau_1$) and since $\mathbf{b}(f) \in \ell_{\tau_0}$ ($f \in B^{\alpha_0}$), we have $\mathbf{b}(f) - \mathbf{a} \in \ell_{\tau_0}$. We define $g \stackrel{L_p}{=} \sum_{\theta \in \Theta} a_\theta \varphi_\theta$. Then by Theorem 2.9, g is well defined, and hence

$$f - g \stackrel{L_p}{=} \sum_{\theta \in \Theta} (b_\theta(f) - a_\theta) \varphi_\theta.$$

By (2.33) and Theorem 2.10, we infer

$$\|g\|_{B^{\alpha_1}} \leq c\|\mathbf{a}\|_{\ell_{\tau_1}} \quad \text{and} \quad \|f - g\|_{B^{\alpha_0}} \leq c\|\mathbf{b}(f) - \mathbf{a}\|_{\ell_{\tau_0}}.$$

Since $\mathbf{a} \in \ell_{\tau_1}$ is arbitrary, the last two estimates give (2.61).

We next prove that

$$(2.62) \quad K(\mathbf{b}(f), t; \ell_{\tau_0}, \ell_{\tau_1}) \leq cK(f, t; B^{\alpha_0}, B^{\alpha_1}), \quad t > 0.$$

Suppose $g \in B^{\alpha_1}$; then by Theorem 2.11, $g \in B^{\alpha_0}$ ($\alpha_0 < \alpha_1$), and hence $f - g \in B^{\alpha_0}$. We shall show that there exists a sequence $\mathbf{b}(g) = (b_\theta(g))_{\theta \in \Theta} \in \ell_{\tau_1}$ such that

$$(2.63) \quad g \stackrel{L_p}{=} \sum_{\theta \in \Theta} b_\theta(g) \varphi_\theta \quad \text{with} \quad \|g\|_{B^{\alpha_1}} \approx \|\mathbf{b}(g)\|_{\ell_{\tau_1}}$$

and

$$(2.64) \quad f - g \stackrel{L_p}{=} \sum_{\theta \in \Theta} (b_\theta(f) - b_\theta(g)) \varphi_\theta \quad \text{with} \quad \|f - g\|_{B^{\alpha_0}} \approx \|\mathbf{b}(f) - \mathbf{b}(g)\|_{\ell_{\tau_0}}.$$

Clearly, estimate (2.62) follows by (2.63)–(2.64).

Notice that if $\eta \geq 1$, then $\mathbf{b}(\cdot)$ can be realized as a linear operator, and hence $\mathbf{b}(f - g) = \mathbf{b}(f) - \mathbf{b}(g)$. Therefore, (2.63)–(2.64) are immediate from $g \in B^{\alpha_1}$ and $f - g \in B^{\alpha_0}$.

Suppose $\eta < 1$. For $\Delta \in \mathcal{T}$, we let $P_\Delta(f) := P_{\Delta, \eta}(f) \in \Pi_k$ be the polynomial from the definition of $p_{m, \eta}(f)$ in (2.20) ($P_\Delta(f)$ is not unique). Thus $P_\Delta(f) \in \Pi_k$ is such that

$$(2.65) \quad \|f - P_\Delta(f)\|_{L_\eta(\Delta)} \leq cE_k(f, \Delta)_\eta.$$

We shall next show that for each $\Delta \in \mathcal{T}$ there exists a polynomial $P_\Delta(g) \in \Pi_k$ such that

$$(2.66) \quad \|g - P_\Delta(g)\|_{L_\eta(\Delta)} \leq cE_k(g, \Delta)_\eta$$

and

$$(2.67) \quad \|f - g - (P_\Delta(f) - P_\Delta(g))\|_{L_\eta(\Delta)} \leq cE_k(f - g, \Delta)_\eta.$$

We consider two cases.

Case 1. $E(f - g) \leq E(g)$, where $E(\cdot) := E_k(\cdot, \Delta)_\eta$. Let $R \in \Pi_k$ be such that

$$(2.68) \quad \|f - g - R\| = E(f - g), \quad \text{where} \quad \|\cdot\| := \|\cdot\|_{L_\eta(\Delta)}.$$

We define $P_\Delta(g) := P_\Delta(f) - R \in \Pi_k$. Then (2.67) holds, by (2.68). We use (2.65) and (2.68) to obtain

$$\begin{aligned} \|g - P_\Delta(g)\| &\leq c\|f - P_\Delta(f)\| + c\|f - g - R\| \leq cE(f) + cE(f - g) \\ &\leq cE(f - g) + cE(g) + cE(f - g) \leq cE(g), \end{aligned}$$

which gives (2.66).

Case 2. $E(g) < E(f - g)$. This time we choose $P_\Delta(g) \in \Pi_k$ so that $\|g - P_\Delta(g)\| = E(g)$. Similarly as above, one can show that

$$\|f - g - (P_\Delta(f) - P_\Delta(g))\| \leq cE(f - g).$$

Thus the existence of $P_\Delta(g) \in \Pi_k$ satisfying (2.66) and (2.67) is established.

Using the polynomials $P_\Delta(g)$ from above, we define, for $m \in \mathbb{Z}$,

$$p_{m,\eta}(g) := \sum_{\Delta \in \mathcal{T}_m} \mathbb{1}_\Delta \cdot P_\Delta(g) \quad \text{and} \quad p_{m,\eta}(f - g) := \sum_{\Delta \in \mathcal{T}_m} \mathbb{1}_\Delta \cdot (P_\Delta(f) - P_\Delta(g)).$$

Furthermore, as in (2.21) and (2.34), we define

$$Q_{m,\eta}(g) := Q_m(p_{m,\eta}(g)) \quad \text{and} \quad q_{m,\eta}(g) := Q_{m,\eta}(g) - Q_{m-1,\eta}(g).$$

We define $Q_{m,\eta}(f - g)$ and $q_{m,\eta}(f - g)$ in the same way. Finally, we define $\mathbf{b}(g) = (b_\theta(g))_{\theta \in \Theta}$ and $\mathbf{b}(f - g) = (b_\theta(f - g))_{\theta \in \Theta}$ from

$$q_{m,\eta}(g) =: \sum_{\theta \in \Theta_m} b_\theta(g) \varphi_\theta \quad \text{and} \quad q_{m,\eta}(f - g) =: \sum_{\theta \in \Theta_m} b_\theta(f - g) \varphi_\theta, \quad m \in \mathbb{Z}.$$

Evidently, $p_{m,\eta}(f - g) = p_{m,\eta}(f) - p_{m,\eta}(g)$ and since Q_m is a linear operator, it follows that $\mathbf{b}(f - g) = \mathbf{b}(f) - \mathbf{b}(g)$. From this and the fact that $P_\Delta(g)$ satisfies (2.66) and (2.67), using Theorem 2.8, Theorem 2.10, and (2.39), we obtain that $\mathbf{b}(g)$ satisfies (2.63) and (2.64), and hence (2.62) holds. This completes the proof of the lemma. \square

By Lemma 2.13, (2.58), and (2.59) (with α_j replaced by α), we obtain

$$\|f\|_{(B^{\alpha_0, B^{\alpha_1}})_{\lambda, \tau}} \approx \|\mathbf{b}(f)\|_{(\ell_{\tau_0, \ell_{\tau_1}})_{\lambda, \tau}} \approx \|\mathbf{b}(f)\|_{\ell_\tau} \approx \|f\|_{B^\alpha}.$$

Thus the proof of Theorem 2.12 is complete. \square

Several remarks are in order.

Remark 2.14. (a) If $p = \infty$, then the B-space $B_\tau^\alpha(\Phi)$ ($\tau := 1/\alpha$) is useful for our goals only if $\alpha \geq 1$. The reason for this is that $B_\tau^\alpha(\Phi)$ is not embedded in C if $\alpha < 1$. Indeed, consider the function $f := \sum_{j=1}^\infty j^{-1} \varphi_{\theta_j}$, where $\theta_j \in \Theta_{m_j}$, $m_1 < m_2 < \dots$, and $\{\varphi_{\theta_j}\}$ are Courant (or other) elements which overlap so that $\|f\|_\infty \approx \sum_{j=1}^\infty j^{-1} = \infty$. On the other hand (see (2.33)), $|f|_{B_\tau^\alpha(\Phi)} \leq c(\sum_{j=1}^\infty j^{-\tau})^{1/\tau} < \infty$, since $\tau := 1/\alpha > 1$.

(b) We introduced the B-norms $N_{\Phi, \mathcal{S}, \eta}(\cdot)$ and $N_{\Phi, \mathcal{Q}, \eta}(\cdot)$ with $0 < \eta < p$ (see (2.31) and (2.37)) for the following reason. As we shall see in section 3, normally $\alpha > 1$, and hence $\tau < 1$, which compels us to work in L_τ with $\tau < 1$, which is not a very friendly space. At the same time, if $p > 1$ we can choose $1 \leq \eta < p$ and work in L_η instead.

(c) We also want to explain why we introduce the B-spaces over locally regular (or better) triangulations but not over more general ones. The reason is that if we relax the main conditions (2.1)–(2.2) in the definition of LR-triangulations, then we can hardly work with the B-spaces. In particular, the equivalence of the norms (see Theorem 2.10) fails to exist, which makes it impossible to prove all the results from section 3.

General B-spaces. Given an LR(or better)-triangulation \mathcal{T} and a family of basis functions $\Phi = \Phi_\mathcal{T}$ over \mathcal{T} as in section 2.2, we define the more general B-space $B_{pq}^\alpha(\Phi) = B_{pq}^\alpha(\mathcal{S})$, $\alpha > 0$, $0 < p, q \leq \infty$, as the set of all $f \in L_p(\mathbb{R}^2)$ such that

$$\|f\|_{B_{pq}^\alpha(\Phi)} := \|f\|_p + \left(\sum_{m \in \mathbb{Z}} \left[2^{m\alpha} \left(\sum_{\Delta \in \mathcal{T}, 2^{-m} \leq |\Delta| < 2^{-m+1}} \mathbb{S}_\Delta(f)_p^p \right)^{1/p} \right]^q \right)^{1/q} < \infty,$$

with the ℓ_q -norm replaced by the sup-norm if $q = \infty$, where $\mathbb{S}_\Delta(f)_p$ is as above (see (2.22)). Evidently, $B_p^\alpha(\Phi) = B_{pp}^\alpha(\Phi)$. In going further, the norms in $B_\tau^\alpha(\Phi)$ from (2.31), (2.32), and (2.37) can be generalized accordingly. In the present article, we do not explore the B-spaces in such generality because the space scale $B_\tau^\alpha(\Phi)$ is sufficient for our goal of characterizing the approximation rates of nonlinear n -term approximation from differentiable piecewise polynomials.

Fat B-spaces: The link to Besov spaces. Suppose \mathcal{T} is an arbitrary SLR-triangulation of \mathbb{R}^2 . The *fat B-space* $\mathbb{B}_\tau^{\alpha k}(\mathcal{T})$ with $k \geq 1$ and α, τ as in the definition of $B_\tau^\alpha(\mathcal{T})$ (section 2.3) is defined (see [38]) as the set of all functions $f \in L_\tau(\mathbb{R}^2)$ such that

$$\|f\|_{\mathbb{B}_\tau^{\alpha k}(\mathcal{T})} := \left(\sum_{\Delta \in \mathcal{T}} [|\Delta|^{-\alpha} E_k(f, \Omega_\Delta)_\tau]^\tau \right)^{1/\tau} \approx \left(\sum_{\Delta \in \mathcal{T}} [|\Delta|^{-\alpha} \omega_k(f, \Omega_\Delta)_\tau]^\tau \right)^{1/\tau} < \infty,$$

where $E_k(f, \Omega_\Delta)_\tau$ is the error of L_τ -approximation to f on $\Omega_\Delta := \Omega_\Delta^1$ from Π_k and $\omega_k(f, \Omega_\Delta)_\tau$ is the local L_τ -modulus of smoothness of f on Ω_Δ . (Recall that $E_k(f, \Omega_\Delta)_\tau \approx \omega_k(f, \Omega_\Delta)_\tau$ by Whitney’s theorem (2.16), since \mathcal{T} is an SLR-triangulation.) Furthermore, other equivalent norms in $\mathbb{B}_\tau^{\alpha k}(\mathcal{T})$ as well as more general fat B-spaces $\mathbb{B}_{pq}^{\alpha k}(\mathcal{T})$ can be defined as in [38].

Suppose that $\Phi = \Phi_\mathcal{T}$ is a hierarchical family of basis functions over \mathcal{T} as described in section 2.2. Assuming $\Pi_k \subset \mathcal{S}_m \subset \mathcal{S}_m^{k,r}(\mathcal{T})$ for all $m \in \mathbb{Z}$ (that is, $\tilde{k} = k$ in the notation of section 2.2), we have for $f \in L_\tau$ and $\Delta \in \mathcal{T}_m$,

$$E_k(f, \Omega_\Delta^\ell)_\tau \leq c \sum_{\Delta' \in \mathcal{T}_m, \Delta' \subset \Omega_\Delta^\ell} E_k(f, \Omega_{\Delta'})_\tau,$$

which implies $\|f\|_{B_\tau^\alpha(\Phi_{\mathcal{T}})} \leq c\|f\|_{\mathbb{B}_\tau^{\alpha k}(\mathcal{T})}$. Therefore, the space $\mathbb{B}_\tau^{\alpha k}(\mathcal{T})$ is a good candidate to replace $B_\tau^\alpha(\Phi)$ in nonlinear spline approximation, but this is only possible if $0 < \alpha < \alpha_0$ for some $\alpha_0 < \infty$, which we do not compute here. The problem with the space $\mathbb{B}_\tau^{\alpha k}(\mathcal{T})$ is that $\|\varphi_\theta\|_{\mathbb{B}_\tau^{\alpha k}(\mathcal{T})} < \infty$ only for $0 < \alpha < \alpha_0$. (See Theorem 2.15 in the case of regular triangulations.) Therefore, the basic norm equivalence results (Theorem 2.10) hold only for a restricted range of α . Thus, $\mathbb{B}_\tau^{\alpha k}(\mathcal{T})$ is simply not the “right” space for the specific problem at hand if $\alpha \geq \alpha_0$. It is too “fat.” However, the spaces $\mathbb{B}_\tau^{\alpha k}(\mathcal{T})$ are still noteworthy since they are less sensitive to small perturbations of the triangulation \mathcal{T} and are technically easier. We believe that a situation will present itself when they will be the “right” spaces.

Comparison between regular B-spaces and Besov spaces. We begin by recalling the definition of the classical Besov space by moduli of smoothness. So, the space $B_q^s(L_p) := B_q^s(L_p(\mathbb{R}^2))$, $s > 0$, $1 \leq p, q \leq \infty$, is defined as the set of all functions $f \in L_p(\mathbb{R}^2)$ such that

$$(2.69) \quad \|f\|_{B_q^s(L_p)} := \left(\int_0^\infty (t^{-s} \omega_k(f, t)_p)^q \frac{dt}{t} \right)^{1/q} < \infty$$

($\|f\|_p$ is usually added to the right-hand side above), where $k := [s] + 1$ and $\omega_k(f, t)_p$ is the k th modulus of smoothness of f in $L_p(\mathbb{R}^2)$, i.e., $\omega_k(f, t)_p := \sup_{|h| \leq t} \|\Delta_h^k(f, \cdot)\|_p$. It is well known that whenever $1 \leq p \leq \infty$, if in (2.69) k is replaced by any other $k > s$, then the resulting space would be the same with an equivalent norm. However, the situation is totally different when $p < 1$, and this is a reason for introducing k as a parameter of the Besov spaces in the following.

As elsewhere, let us assume that $0 < p < \infty$ and $\alpha > 0$, or $p = \infty$ and $\alpha \geq 1$, and in both cases $1/\tau := \alpha + 1/p$. Let $k \geq 1$. We define the space $B_\tau^{2\alpha, k}(L_\tau)$ as the Besov space $B_\tau^{2\alpha}(L_\tau)$ (see (2.69)), where k and α are independent of each other. These are the spaces that naturally occur in nonlinear spline approximation (see [53]).

Suppose that \mathcal{T}^* is a regular triangulation of \mathbb{R}^2 (see section 2.1). Then as shown in [38], $\mathbb{B}_\tau^{\alpha k}(\mathcal{T}^*) = B_\tau^{2\alpha, k}(L_\tau)$ with equivalent norms. (Notice that the smoothness parameters of B-spaces and Besov spaces are normalized differently and α corresponds to 2α .)

Let us now assume that $\Phi_{\mathcal{T}^*} = \{\varphi_\theta\}$ is a family of basis functions over \mathcal{T}^* as in section 2.2 such that $\Pi_k \subset \mathcal{S}_m \subset \mathcal{S}_m^{k, r}$ ($m \in \mathbb{Z}$), where $r \geq 0$ and $k > r$. As we mentioned above, the fat B-space $\mathbb{B}_\tau^{\alpha k}(\mathcal{T}^*)$, and hence the Besov space $B_\tau^{2\alpha, k}(L_\tau)$, is a good candidate to replace the B-space $B_\tau^\alpha(\Phi_{\mathcal{T}^*})$ in nonlinear n -term approximation from $\Phi_{\mathcal{T}^*}$. We next spell out the exact conditions for equivalence of the corresponding norms.

THEOREM 2.15. *Under the above assumptions, if $0 < \alpha < r + 1 + 1/p$, then*

$$(2.70) \quad B_\tau^{2\alpha, k}(L_\tau) = B_\tau^\alpha(\Phi_{\mathcal{T}^*})$$

with equivalent norms. Furthermore, if a single basis function $\varphi_\theta \in \Phi_{\mathcal{T}^}$ does not belong to C^{r+1} , then the equivalence is no longer true when $\alpha \geq r + 1 + 1/p$. More precisely, for such φ_θ and α , $\|\varphi_\theta\|_{B_\tau^{2\alpha, k}(L_\tau)} = \infty$, while $\|\varphi_\theta\|_{B_\tau^\alpha(\Phi_{\mathcal{T}^*})} \approx \|\varphi_\theta\|_p$.*

Proof. As we mentioned before, $\|f\|_{B_\tau^\alpha(\Phi_{\mathcal{T}^*})} \leq c\|f\|_{\mathbb{B}_\tau^{\alpha k}(\mathcal{T}^*)}$ for $f \in \mathbb{B}_\tau^{\alpha k}(\mathcal{T}^*)$, and also we have $\|f\|_{\mathbb{B}_\tau^{\alpha k}(\mathcal{T}^*)} \approx \|f\|_{B_\tau^{2\alpha, k}(L_\tau)}$, exactly as in Theorem 2.25 from [38]. Therefore,

$$\|f\|_{B_\tau^\alpha(\Phi_{\mathcal{T}^*})} \leq c\|f\|_{B_\tau^{2\alpha, k}(L_\tau)} \text{ for } f \in B_\tau^{2\alpha, k}(L_\tau).$$

The proof of the reverse estimate follows in the footsteps of the proof of Theorem 2.28 from [38], and we shall indicate only the differences. Using the conditions on $\Phi_{\mathcal{T}^*}$ and the fact that \mathcal{T}^* is regular, one can show by straightforward calculations that, for each $\theta \in \Theta(\mathcal{T}^*)$,

$$(2.71) \quad \omega_k(\varphi_\theta, t)_\tau^r \leq \begin{cases} c|E_\theta|^{\frac{1}{2}(1-(r+1)\tau)} \cdot t^{1+(r+1)\tau} & \text{if } 0 < t < |E_\theta|^{1/2}, \\ c|E_\theta| & \text{if } t \geq |E_\theta|^{1/2}. \end{cases}$$

Moreover, both sides of (2.71) are equivalent if φ_θ does not belong to C^{r+1} . In going further, one uses (2.71) exactly as in [38] to complete the proof of the theorem. \square

Remark 2.16. An interesting situation occurs when $p = \infty$ and $r = 0$. Then there is no α for which (2.70) holds. This is the case when $\Phi_{\mathcal{T}^*}$ is the set of all Courant elements generated by \mathcal{T}^* (a regular triangulation).

Comparison between different B-spaces and Besov spaces. Suppose $\Phi_{\mathcal{T}}$ is a family of basis functions associated with an SLR-triangulation \mathcal{T} which allows arbitrarily sharp angles. Then some extremely “skinny” basis functions $\varphi_\theta \in \Phi_{\mathcal{T}}$ (with elongated level curves) will occur. It is easily seen that such functions have huge Besov norms (see [38]) compared to their L_p -norms as well as their $B(\Phi_{\mathcal{T}})$ -norms (see Theorem 3.2 below) for any smoothness $\alpha > 0$. Therefore, the B-spaces for such a triangulation are essentially different from Besov spaces. The situation is quite similar when comparing two B-spaces over different triangulations. Therefore, the B-spaces change substantially with the triangulations, thus making the search for the “right” triangulation mentioned in the introduction a meaningful task. In contrast to this, the standard Besov spaces can be used only to characterize the approximation power of piecewise polynomials over *regular* triangulations.

B-spaces over compact domains. B-spaces can be introduced on an arbitrary compact polygonal domain $E \subset \mathbb{R}^2$. A substantial difference would be in assuming that each triangulation \mathcal{T} of E is of the form $\mathcal{T} = \bigcup_{m=0}^\infty \mathcal{T}_m$, where \mathcal{T}_0 is an initial level (triangulation of E) and $\mathcal{T}_1, \mathcal{T}_2, \dots$ are consecutive refinements of \mathcal{T}_0 . This approach is important for the applications (see [39]).

B-spaces in dimensions $d > 2$. Multilevel triangulations and B-spaces can be introduced in much the same way in dimensions $d > 2$. Of course, then the triangles should be replaced by simplices, thus making some geometric argumentation of this section essentially more involved. In particular, the property (e) of a multilevel triangulation should be extended to all faces of the simplices in \mathcal{T}_m , thus saying that there are at most N_0 simplices in \mathcal{T}_m attached to a particular face. The “no hanging vertices” condition (d) should be replaced by the condition that each facet of a simplex in \mathcal{T}_m is a common facet of exactly two simplices in \mathcal{T}_m . The minimal angle condition appearing in the definition of regular triangulations and in (2.4) should be replaced by the *shape regularity* condition that postulates the existence of an upper bound on the ratio of the diameter of a simplex and the diameter of the inscribed sphere. In conditions (2.1)–(2.3) the area should be replaced by the d -dimensional volume.

B-spaces in dimension $d = 1$. B-spaces can be introduced in the univariate case, but none will give anything new, and hence they are not needed. The key fact is that, in the univariate case, the Bernstein inequality involving Besov spaces holds with no restrictions on the smoothness parameter $\alpha < \infty$ (see [53]).

In a nutshell, the essence of the spaces we considered in this section is the following. The Besov spaces are based on local polynomial approximation over regular multilevel triangulations, which is explicitly shown in [38]. When the regular triangulations are replaced by SLR-triangulations, then the Besov spaces become fat

B-spaces, which further evolve to B-spaces when the local polynomial approximation is replaced by local spline approximation.

The B-spaces are closely related to certain anisotropic maximal functions, non-classical differentiability, and other problems, which are beyond the scope of this article.

3. Nonlinear n -term spline approximation. In this section, we assume that \mathcal{T} is a locally regular (or better) triangulation of \mathbb{R}^2 . Also, we assume that $\Phi = \Phi_{\mathcal{T}}$ is a hierarchical family of basis functions over \mathcal{T} (see section 2.2). Notice that Φ is not a basis; Φ is redundant. We consider nonlinear n -term approximation from Φ in $L_p(\mathbb{R}^2)$ ($0 < p \leq \infty$), where we identify $L_\infty(\mathbb{R}^2)$ as $C_0(\mathbb{R}^2)$. We let $\Sigma_n(\Phi)$ denote the nonlinear set consisting of all splines s of the form

$$s = \sum_{\theta \in \mathcal{M}} a_\theta \varphi_\theta,$$

where $\mathcal{M} \subset \Theta(\mathcal{T})$, $\#\mathcal{M} \leq n$, and \mathcal{M} may vary with s . We denote by $\sigma_n(f, \Phi)_p$ the error of L_p -approximation to $f \in L_p(\mathbb{R}^2)$ from $\Sigma_n(\Phi)$:

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

Our goal is to characterize the approximation spaces generated by nonlinear n -term approximation from Φ . To this end we next prove a pair of companion Jackson and Bernstein estimates. We shall utilize the B-spaces $B_\tau^\alpha(\Phi)$ introduced in section 2. We assume that $0 < p < \infty$ and $\alpha > 0$, or $p = \infty$ and $\alpha \geq 1$. In both cases, $1/\tau := \alpha + 1/p$ ($1/\infty := 0$).

THEOREM 3.1 (Jackson estimate). *If $f \in B_\tau^\alpha(\Phi)$, then*

$$(3.1) \quad \sigma_n(f, \Phi)_p \leq cn^{-\alpha} \|f\|_{B_\tau^\alpha(\Phi)},$$

with c independent of f and n .

In the case $0 < p < \infty$, this theorem follows by the general Theorem 3.4 from [38], in view of the results of section 2. For completeness, we shall give its short proof in the appendix. The proof when $p = \infty$ can be carried out as the proof of Theorem 4.1 from [39] but is a little longer, and so we shall skip it.

THEOREM 3.2 (Bernstein estimate). *If $s \in \Sigma_n(\Phi)$, then*

$$(3.2) \quad \|s\|_{B_\tau^\alpha(\Phi)} \leq cn^\alpha \|s\|_p,$$

with c independent of s and n .

The proof of this (vital for our development) theorem utilizes the ideas of the proofs of Theorem 3.6 from [38] ($0 < p < \infty$) and Theorem 4.2 from [39] ($p = \infty$) but is not identical to them. We shall give the proof in the appendix.

For a fixed \mathcal{T} and $\Phi := \Phi_{\mathcal{T}}$, we set $K(f, t) := K(f, t; L_p, B_\tau^\alpha(\Phi))$ ($L_p := C_0$ if $p = \infty$); see (2.55). The Jackson and Bernstein estimates from Theorems 3.1 and 3.2 imply in a standard way (see, e.g., [55]) the following direct and inverse estimates: For any $\alpha > 0$, if $f \in L_p$, then

$$(3.3) \quad \sigma_n(f, \Phi)_p \leq cK(f, n^{-\alpha})$$

and

$$(3.4) \quad K(f, n^{-\alpha}) \leq cn^{-\alpha} \left(\|f\|_p + \left[\sum_{\nu=1}^n \frac{1}{\nu} (\nu^\alpha \sigma_\nu(f, \Phi)_p)^\mu \right]^{1/\mu} \right),$$

where $\mu := \min\{p, 1\}$ and c is independent of f and n .

An immediate consequence of (3.3) and (3.4) is that $\sigma_n(f, \Phi)_p = O(n^{-\gamma})$, $0 < \gamma < \alpha$, if and only if $K(f, n^{-\alpha}) = O(n^{-\gamma})$. More generally, these estimates enable us to characterize the approximation spaces generated by nonlinear n -term approximation from Φ . We define the approximation space $A_q^\gamma := A_q^\gamma(\Phi, L_p)$, $\alpha > 0$, $0 < q \leq \infty$, as the set of all functions $f \in L_p$ such that

$$\|f\|_{A_q^\gamma} := \|f\|_p + \left(\sum_{n=1}^{\infty} (n^\gamma \sigma_n(f, \Phi)_p)^q \frac{1}{n} \right)^{1/q} < \infty$$

with the ℓ_q -norm replaced by the sup-norm if $q = \infty$ as usual.

The direct and inverse estimates (3.3)–(3.4) readily imply (see, e.g., [55]) the following characterization of the approximation spaces.

THEOREM 3.3. *If $0 < \gamma < \alpha$ and $0 < q \leq \infty$, then*

$$A_q^\gamma(\Phi, L_p) = (L_p, B_\tau^\alpha(\Phi))_{\frac{\gamma}{\alpha}, q}$$

with equivalent norms.

In one specific case the interpolation spaces can be identified as B-spaces.

THEOREM 3.4. *Suppose $0 < p < \infty$ and $\alpha > 0$, or $p = \infty$ and $\alpha > 1$, and let $\tau := (\alpha + 1/p)^{-1}$. Then*

$$(3.5) \quad A_\tau^\alpha(\Phi, L_p) = B_\tau^\alpha(\Phi)$$

with equivalent norms.

The following interpolation result is immediate from Theorems 3.3 and 3.4.

COROLLARY 3.5. *Suppose p , α , and $\tau := \tau(\alpha)$ are as in the hypothesis of Theorem 3.4, and let $\beta > \alpha$ and $\tau(\beta) := (\beta + 1/p)^{-1}$. Then*

$$(3.6) \quad (L_p, B_{\tau(\beta)}^\beta(\Phi))_{\frac{\alpha}{\beta}, \tau(\alpha)} = B_{\tau(\alpha)}^\alpha(\Phi)$$

with equivalent norms.

Proof of Theorem 3.4. We shall employ the idea of the proof of Theorem 3.3 in [30]. We shall use abbreviated notation: $A_q^\alpha := A_q^\alpha(\Phi, L_p)$, $B_\tau^\alpha := B_\tau^\alpha(\Phi)$, and the like. For any $\beta > 0$, we denote $\tau(\beta) := (\beta + 1/p)^{-1}$.

We first prove the following continuous embedding:

$$(3.7) \quad A_\mu^\beta \subset B_{\tau(\beta)}^\beta \quad \text{with} \quad \mu := \min\{\tau(\beta), 1\}.$$

Indeed, suppose $f \in A_\mu^\beta$, and let $s_m \in \Sigma_m$ be such that

$$(3.8) \quad \|f - s_m\|_p \leq 2\sigma_m(f)_p.$$

Since $\sigma_m(f)_p \rightarrow 0$, we have $f = s_1 + \sum_{\nu=1}^{\infty} (s_{2^\nu} - s_{2^{\nu-1}})$ in L_p (uniformly if $p = \infty$), and hence ($\mu \leq 1$)

$$(3.9) \quad \|f\|_{B_{\tau(\beta)}^\beta}^\mu \leq \|s_1\|_{B_{\tau(\beta)}^\beta}^\mu + \sum_{\nu=1}^{\infty} \|s_{2^\nu} - s_{2^{\nu-1}}\|_{B_{\tau(\beta)}^\beta}^\mu.$$

We apply the Bernstein estimate from Theorem 3.2 to $s_{2^\nu} - s_{2^{\nu-1}} \in \Sigma_{2^{\nu+1}}$ and use (3.8) to obtain

$$\|s_{2^\nu} - s_{2^{\nu-1}}\|_{B_{\tau(\beta)}^\beta} \leq c2^{\nu\beta} \|s_{2^\nu} - s_{2^{\nu-1}}\|_p \leq c2^{\nu\beta} (\sigma_{2^\nu}(f)_p + \sigma_{2^{\nu-1}}(f)_p),$$

and similarly $\|s_1\|_{B_{\tau(\beta)}^\beta} \leq c(\|f\|_p + \sigma_1(f)_p)$. Using these in (3.9) implies

$$\|f\|_{B_{\tau(\beta)}^\beta}^\mu \leq c\|f\|_p^\mu + c \sum_{\nu=1}^{\infty} (2^{\nu\beta} \sigma_{2^\nu}(f)_p)^\mu \leq c\|f\|_{A_\mu^\beta}^\mu,$$

which is (3.7).

Second, the Jackson estimate from Theorem 3.1 gives the continuous embedding

$$(3.10) \quad B_{\tau(\beta)}^\beta \subset A_\infty^\beta.$$

A third important ingredient in this proof is the fact that the approximation spaces A_q^α are invariant under interpolation (see [31, 52]): If $\alpha_0, \alpha_1 > 0$ and $0 < q_1, q_2, q \leq \infty$, then

$$(3.11) \quad (A_{q_0}^{\alpha_0}, A_{q_1}^{\alpha_1})_{\lambda, q} = A_q^\alpha, \quad \text{where } \alpha = (1 - \lambda)\alpha_0 + \lambda\alpha_1 \text{ with } 0 < \lambda < 1.$$

Now we choose α_0 and α_1 so that $0 < \alpha_0 < \alpha < \alpha_1$ ($\alpha_0 := 1$ if $p = \infty$). Also, we select $0 < \lambda < 1$ so that $\alpha = (1 - \lambda)\alpha_0 + \lambda\alpha_1$. Furthermore, we set $\tau_j := (\alpha_j + 1/p)^{-1}$ and $\mu_j := \min\{\tau_j, 1\}$, $j = 0, 1$. By Theorem 2.12, we have

$$(B_{\tau_0}^{\alpha_0}, B_{\tau_1}^{\alpha_1})_{\lambda, \tau} = B_\tau^\alpha.$$

We use this, (3.7), (3.10), and (3.11) to obtain the following continuous embeddings:

$$A_\tau^\alpha = (A_{\mu_0}^{\alpha_0}, A_{\mu_1}^{\alpha_1})_{\lambda, \tau} \subset B_\tau^\alpha = (B_{\tau_0}^{\alpha_0}, B_{\tau_1}^{\alpha_1})_{\lambda, \tau} \subset (A_\infty^{\alpha_0}, A_\infty^{\alpha_1})_{\lambda, \tau} = A_\tau^\alpha,$$

which give (3.5). \square

Algorithms. In [39], there are three algorithms developed for n -term Courant element approximation in L_p ($0 < p \leq \infty$). These algorithms can be immediately adapted to nonlinear n -term approximation from any family of differentiable spline basis functions $\Phi_{\mathcal{T}}$ on a compact polygonal domain $E \subset \mathbb{R}^2$. It is an integral part of our program that using the machinery of the B-spaces, Jackson and Bernstein estimates, interpolation, etc. developed in this article, we can prove that these algorithms achieve the rate of the best n -term approximation. This aspect of our theory will not be elaborated on here (see [39]).

Approximation from the libraries $\{\Phi_{\mathcal{T}}\}_{\mathcal{T}}$. An important element of our concept for nonlinear spline approximation is the introduction of another level of nonlinearity by allowing the triangulation \mathcal{T} to vary. For a given SRL(or LR)-triangulation \mathcal{T} , let $\Phi_{\mathcal{T}}$ be a family of spline basis functions like the ones considered in section 2.2. Now, without changing the nature of the basis elements from $\Phi_{\mathcal{T}}$, we let \mathcal{T} vary and obtain a collection (library) of basis families $\{\Phi_{\mathcal{T}}\}_{\mathcal{T}}$. We denote

$$\sigma_n(f)_p := \inf_{\mathcal{T}} \sigma_n(f, \Phi_{\mathcal{T}})_p,$$

where the infimum is taken over all SLR-triangulations \mathcal{T} with fixed parameters, and we also assume that the parameters of $\Phi_{\mathcal{T}}$ are fixed. The following theorem is immediate from the Jackson estimate in Theorem 3.1. We shall assume again that $0 < p < \infty$ and $\alpha > 0$, or $p = \infty$ and $\alpha \geq 1$, and in both cases, $1/\tau := \alpha + 1/p$.

THEOREM 3.6. *If $\inf_{\mathcal{T}} \|f\|_{B_{\tau}^{\alpha}(\Phi_{\mathcal{T}})} < \infty$, then*

$$\sigma_n(f)_p \leq cn^{-\alpha} \inf_{\mathcal{T}} \|f\|_{B_{\tau}^{\alpha}(\Phi_{\mathcal{T}})}$$

with c depending only on p, α , and the parameters of \mathcal{T} and $\Phi_{\mathcal{T}}$.

The ultimate *open problem* here is to characterize the approximation spaces generated by $\{\sigma_n(f)_p\}$ for a given library of basis functions $\{\Phi_{\mathcal{T}}\}_{\mathcal{T}}$.

Global smoothness of functions: How to measure it? Here we come to one of the fundamental questions in approximation theory (and not only there) of how the global smoothness of the functions should be measured.

In the case of nonlinear n -term L_p -approximation from a *single* basis family $\Phi_{\mathcal{T}}$, a function f should be considered of smoothness $\alpha > 0$ if $\|f\|_{B_{\tau}^{\alpha}(\Phi_{\mathcal{T}})} < \infty$. Then the rate of n -term L_p -approximation of f from $\Phi_{\mathcal{T}}$ is $O(n^{-\alpha})$ (roughly). If we consider nonlinear n -term approximation from a given *library* of basis families $\{\Phi_{\mathcal{T}}\}_{\mathcal{T}}$ (\mathcal{T} is allowed to vary), then a function f should naturally be considered of smoothness $\alpha > 0$ if $\inf_{\mathcal{T}} \|f\|_{B_{\tau}^{\alpha}(\Phi_{\mathcal{T}})} < \infty$, which means that there exists a triangulation $\mathcal{T} := \mathcal{T}_f$ such that $\|f\|_{B_{\tau}^{\alpha}(\Phi_{\mathcal{T}})} < \infty$. Then the rate of n -term L_p -approximation of f from the library $\{\Phi_{\mathcal{T}}\}_{\mathcal{T}}$ is $O(n^{-\alpha})$. It is crystal clear to us that no single (super) space can do the job in this case. It is an *open problem* to develop an algorithm for finding, for a given function f , an optimal (or near optimal) triangulation, i.e., a triangulation \mathcal{T}_f for which f exhibits maximal (near maximal) smoothness, using the space scale $B_{\tau}^{\alpha}(\Phi_{\mathcal{T}_f})$. It is also an *open problem* whether, for a given function $f \in L_p$, there exists a single triangulation \mathcal{T}_f such that, for all $n \geq 1$, the n -term L_p -approximation of f from the library $\{\Phi_{\mathcal{T}}\}_{\mathcal{T}}$ can be realized by n -term approximation from $\Phi_{\mathcal{T}_f}$ and, consequently, characterized by the B-spaces $B_{\tau}^{\alpha}(\Phi_{\mathcal{T}_f})$ via interpolation.

Another important related issue for discussion is the smoothness of the approximating tool $\Phi_{\mathcal{T}} := \{\varphi_{\theta}\}$ (\mathcal{T} fixed). Clearly, in nonlinear approximation, there is no saturation, which means that the corresponding approximation spaces A_q^{γ} are non-trivial for all $0 < \gamma < \infty$. Therefore, the smoothness spaces to be used should naturally be designed so that the functions $\{\varphi_{\theta}\}$ are infinitely smooth with respect to these spaces. This has been one of the guiding principles to us in constructing the B-spaces. Thus each basis function $\varphi_{\theta} \in \Phi$ is infinitely smooth with respect to the scale of B-spaces $B_{\tau}^{\alpha}(\Phi)$, which is reflected in the fact that $\|\varphi_{\theta}\|_{B_{\tau}^{\alpha}(\Phi)} \leq c\|\varphi_{\theta}\|_p$ for $0 < \alpha < \infty$ (see Theorem 3.2). This makes it possible that in our direct, inverse, and characterization theorems we impose no restrictions on the rate of approximation $\alpha < \infty$ (see Theorems 3.1–3.4). Also, this explains the complete success of Besov spaces in univariate nonlinear spline approximation (see [53]) and why Besov spaces are not quite suitable in dimensions $d > 1$. The latter remark needs a few words of explanation: First, by allowing triangulations with arbitrarily sharp angles, we allow very “skinny” basis functions with huge Besov norms compared to their L_p -norms (see [38]), which precludes the use of Besov spaces in such situations. Second, even when working on regular triangulations, the use of Besov spaces is restricted by the Besov smoothness (regularity) of the basis functions (see Theorem 2.15), while B-spaces impose no restrictions on the rates of approximation.

Spline wavelets (prewavelets) and frames. In the case of uniform triangulations, spline wavelets play an essential role in practical algorithms. It would be desirable to have compactly supported wavelet (prewavelet) bases or frames generated by (differentiable) spline basis families $\Phi_{\mathcal{T}}$ over LR- or SLR-triangulations \mathcal{T} . To our knowledge there are no constructions of this type available, as for now. Moreover, there is some evidence that such constructions would be too complicated and impractical for general triangulations. However, continuous spline prewavelets on regular triangulations with uniform dyadic refinements are available from [21, 34, 58]. (See also [47].) Evidently, nonlinear n -term approximation from compactly supported spline wavelets or frames, generated by Courant elements or a smoother spline basis

family $\Phi_{\mathcal{T}}$, cannot give a better rate of convergence than nonlinear n -term approximation from $\Phi_{\mathcal{T}}$. We hope that efficient algorithms for n -term approximation from such families may provide a substitute for wavelet methods in situations where the latter are difficult to apply and, in particular, for approximation in L_{∞} .

Adaptive tree approximation. This is a method for nonlinear approximation from piecewise polynomials on (single level) triangular partitions, which has been developed recently in [5, 7]. In [5], algorithms are developed which achieve the rate of the best adaptive tree approximation, while in [7] the rates of approximation are related to the smoothness of the functions in terms of Besov spaces. There are substantial distinctions between this approach and the one in the present article. Namely, the approximation schemes from [5, 7] use “single level” piecewise polynomials on triangulations which satisfy the minimal angle condition, while here we use multilevel (multiscale) piecewise polynomial bases over triangulations which allow arbitrarily sharp angles. Therefore, the notion of “best approximation” in [5, 7] is quite different from the one used here. Substantial progress has been made in [6] in applying the adaptive tree approximation method for numerical (finite element) solution of PDEs.

4. Construction of differentiable basis functions. In this section, we give, for any SLR-triangulation, a construction of differentiable spline basis in $\mathcal{S}_m^{k,r}$, $r \geq 1$, $k > 4r + 1$, satisfying the conditions from section 2.2. In general, we follow the scheme of [22]; however, appropriate modifications in the construction and in the proofs have to be made since we do not assume that the triangulation is regular. In particular, we replace the standard normal derivatives to the edges by derivatives in affine invariant directions; see the definition of $D_{\mu(e,\Delta)}$ below. Since our construction is also applicable to nonnested triangulations (see Remark 4.8), we formulate the results here for a *fixed* level \mathcal{T}_m assuming only conditions (a), (d)–(f), and (2.3) of section 2.1 and making sure that the constants in (2.11) and (2.12) depend only on k, r, N_0 , and δ_2 .

4.1. Nodal functionals. As before, let \mathcal{V}_m and \mathcal{E}_m be the sets of all vertices and all edges of \mathcal{T}_m , respectively. We shall describe the basis functions for $\mathcal{S}_m = \mathcal{S}^{k,r}(\mathcal{T}_m)$, $k > 4r + 1$, with the aid of the so-called *nodal functionals* defined on $\mathcal{S}^{k,r}(\mathcal{T}_m)$. These are certain linear functionals involving the values of the splines and their derivatives at specific points in \mathbb{R}^2 . The functional corresponding to the simple evaluation of the splines at $\xi \in \mathbb{R}^2$ will be denoted by δ_{ξ} .

Of particular interest as evaluation points are the vertices $v \in \mathcal{V}_m$, where we also need the derivative evaluation functionals of type $\delta_v D_e^{\alpha}$ with e being any edge in \mathcal{E}_m emanating from v , and $\delta_v D_{e_1}^{\alpha} D_{e_2}^{\beta}$, where e_1, e_2 are adjacent edges emanating from v . Here $D_{[v,\tilde{v}]}^{\alpha} s$ denotes the derivative of s of order α in the direction of the interval $[v, \tilde{v}]$, weighted with the length of $[v, \tilde{v}]$, namely,

$$D_{[v,\tilde{v}]}^{\alpha} s := \left((\tilde{v}_x - v_x) D_x + (\tilde{v}_y - v_y) D_y \right)^{\alpha} s, \\ v = (v_x, v_y), \quad \tilde{v} = (\tilde{v}_x, \tilde{v}_y).$$

Note that, due to this weighting, the corresponding *Markov inequality* reads as follows:

$$(4.1) \quad \|D_{[v,\tilde{v}]}^{\alpha} p\|_{L_{\infty}[v,\tilde{v}]} \leq c \|p\|_{L_{\infty}[v,\tilde{v}]}, \quad p \in \Pi_k,$$

where c depends only on k and α .

Let $\Delta_1, \Delta_2 \in \mathcal{T}_m$ share an edge e . Since every $s \in \mathcal{S}^{k,r}(\mathcal{T}_m)$ is continuous, the two polynomial patches $s|_{\Delta_1}$ and $s|_{\Delta_2}$ coincide along e . Therefore, $\delta_v D_e^{\alpha} s$ may be computed for any $\alpha = 0, 1, \dots$ as either $\delta_v D_e^{\alpha}(s|_{\Delta_1})$ or $\delta_v D_e^{\alpha}(s|_{\Delta_2})$ with the same

result. Similarly, let $e_1, e_2 \in \mathcal{E}_m$ be two edges of a triangle $\Delta \in \mathcal{T}_m$ with a common vertex v . Then $\delta_v D_{e_1}^\alpha D_{e_2}^\beta s$ denotes the mixed derivative of s at v in the directions of e_1 and e_2 away from v . If $\alpha + \beta \leq r$, this derivative is uniquely defined. If $\alpha + \beta > r$, the result may depend on the choice of the polynomial patch of s attached to v . We follow the convention to always take $\delta_v D_{e_1}^\alpha D_{e_2}^\beta s := \delta_v D_{e_1}^\alpha D_{e_2}^\beta (s|_\Delta)$, where Δ is the above triangle formed by e_1, e_2 .

We shall also need functionals evaluating at some points on any edge e the derivatives of the spline in an affine invariant direction not parallel to e . Let $e = [v_1, v_2] \in \mathcal{E}_m$, and let $\Delta_e = [v_1, v_2, v_3] \in \mathcal{T}_m$ be a triangle attached to e . Denote by $\mu(e, \Delta)$ the median of Δ connecting the middle point $(v_1 + v_2)/2$ of e with the third vertex v_3 of Δ . For any point $\xi \in e$, $\delta_\xi D_{\mu(e, \Delta)}$ will denote the derivative at ξ in the direction pointing into the half-plane containing Δ parallel to $\mu(e, \Delta)$, weighted with the length of $\mu(e, \Delta)$. For each edge $e \in \mathcal{E}_m$, we choose one of the two triangles attached to e and denote it by Δ_e . (Note that this selection of Δ_e is not unique, but as will be seen it will cause no problems for the basis construction.)

Remark 4.1. For later references, we note here that any nodal functional $\eta : \mathcal{S}^{k,r}(\mathcal{T}_m) \rightarrow \mathbb{R}$ of the above type can be *extended* to a linear functional $\tilde{\eta} : \mathcal{S}^{k,-1}(\mathcal{T}_m) \rightarrow \mathbb{R}$ such that $\tilde{\eta}(s) = \eta(s)$ as long as $s \in \mathcal{S}^{k,r}(\mathcal{T}_m)$. Indeed, if the definition of η involves δ_ξ for some point $\xi \in \cup_{e \in \mathcal{E}_m} e$, then we choose one of the triangles $\Delta \in \mathcal{T}_m$ containing ξ and use the corresponding value of $s|_\Delta$ or its derivatives at ξ to define $\tilde{\eta}(s)$ for any $s \in \mathcal{S}^{k,-1}(\mathcal{T}_m)$. The only restriction on the choice of Δ is that it must be consistent with the above rules for $\delta_v D_e^\alpha$, $\delta_v D_{e_1}^\alpha D_{e_2}^\beta$, and $\delta_\xi D_{\mu(e, \Delta)}$. Clearly, the extension of this type is *not unique*. Moreover, convex combinations of evaluations of the restrictions of s to different triangles can also be used.

4.2. Characterization of differentiability. Let L be a straight line dividing \mathbb{R}^2 into two half-planes H, \tilde{H} . Given $p, \tilde{p} \in \Pi_k$, we define a piecewise polynomial function s by setting $s|_H = p$, $s|_{\tilde{H}} = \tilde{p}$. To check whether s is differentiable across L , we choose two points u, v on L , as well as two points w, \tilde{w} in the interiors of H and \tilde{H} , respectively. We set $\Delta := [u, v, w]$, $\tilde{\Delta} := [u, v, \tilde{w}]$, $e := [u, v]$, $\mu := [u, w]$, $\tilde{\mu} := [u, \tilde{w}]$, $\theta := \angle e\mu$, $\tilde{\theta} := \angle \tilde{e}\tilde{\mu}$. The proof of the following lemma can be found in [17, 25].

LEMMA 4.2. *Let $0 \leq r < k$. Then $s \in C^r(\mathbb{R}^2)$ if and only if*

$$(4.2) \quad \delta_u D_{\tilde{\mu}}^\alpha D_e^{q-\alpha} \tilde{p} = \sum_{\beta=0}^\alpha (-1)^\beta \binom{\alpha}{\beta} \left(\frac{\sin(\theta+\tilde{\theta})}{|e|} \right)^{\alpha-\beta} \left(\frac{\sin \tilde{\theta}}{|\mu|} \right)^\beta \left(\frac{\sin \theta}{|\tilde{\mu}|} \right)^{-\alpha} \delta_u D_\mu^\beta D_e^{q-\beta} p$$

for all $\alpha = 0, \dots, r$ and $q = \alpha, \dots, k-1$.

It is readily seen that (4.2) can be reformulated as follows:

$$(4.3) \quad \delta_u D_{\tilde{\mu}}^\alpha D_e^{q-\alpha} \tilde{p} = \sum_{\beta=0}^\alpha (-1)^\beta \binom{\alpha}{\beta} \left(\sigma |\Delta^*| \right)^{\alpha-\beta} \frac{|\tilde{\Delta}|^\beta}{|\Delta|^\alpha} \delta_u D_\mu^\beta D_e^{q-\beta} p,$$

where $\sigma := \text{sgn} \sin(\theta + \tilde{\theta})$ and $\Delta^* := [u, w, \tilde{w}]$. (This identity simplifies in an obvious way when $|\Delta^*| = 0$.)

See [22] for a discussion of the relationship between these *nodal conditions of differentiability* and the well-known Bernstein–Bézier conditions.

4.3. Construction of basis splines. Consider the following set \mathcal{N}_m of nodal functionals on $\mathcal{S}^{k,r}(\mathcal{T}_m)$,

$$(4.4) \quad \mathcal{N}_m := \left(\bigcup_{v \in \mathcal{V}_m} \mathcal{N}_m^v \right) \cup \left(\bigcup_{e \in \mathcal{E}_m} \mathcal{N}_m^e \right) \cup \left(\bigcup_{\Delta \in \mathcal{T}_m} \mathcal{N}_m^\Delta \right),$$

where for each $\Delta = [v_1, v_2, v_3] \in \mathcal{T}_m$,

$$\mathcal{N}_m^\Delta := \{\eta_\xi^\Delta := \delta_\xi : \xi \in \Xi_\Delta\},$$

$$\Xi_\Delta := \left\{ \frac{i_1 v_1 + i_2 v_2 + i_3 v_3}{k-1} : i_1 + i_2 + i_3 = k-1, \quad i_1, i_2, i_3 > r \right\} \subset \Delta,$$

for each edge $e = [v_1, v_2] \in \mathcal{E}_m$,

$$\mathcal{N}_m^e := \{\eta_{q,\xi}^e := \delta_\xi D_{\mu(e,\Delta_e)}^q : q = 0, \dots, r, \quad \xi \in \Xi_{e,q}\},$$

$$\Xi_{e,q} := \left\{ \frac{i_1 v_1 + i_2 v_2}{k-q-1} : i_1 + i_2 = k-q-1, \quad i_1, i_2 > 2r-q \right\} \subset e,$$

and for each vertex $v \in \mathcal{V}_m$,

$$\mathcal{N}_m^v := \bigcup_{q=0}^{2r} \mathcal{N}_m^{v,q},$$

with $\mathcal{N}_m^{v,q}$, $q = 0, \dots, 2r$, being defined as follows. Let $\Delta^{[i]} = [v, v_i, v_{i+1}]$, $i = 1, \dots, N_v$, be the triangles in \mathcal{T}_m attached to v in counterclockwise order, $v_{N_v+l} = v_l$, and let $e_i = [v, v_i]$. We set

$$\mathcal{N}_m^{v,0} := \{\eta^{v,0} := \delta_v\},$$

$$\mathcal{N}_m^{v,q} := \{\eta_{i,\alpha}^{v,q} := \delta_v D_{e_i}^{q-\alpha} D_{e_{i+1}}^\alpha : i = 1, \dots, N_v, \alpha = 0, \dots, q-1\}, \quad q \geq 1.$$

Note that \mathcal{N}_m^Δ or \mathcal{N}_m^e might be empty for some combinations of r, k , e.g., $\mathcal{N}_m^\Delta = \mathcal{N}_m^e = \emptyset$ if $r = 0, k = 2$, or $\mathcal{N}_m^\Delta = \emptyset$ if $r = 1, k = 6$. This, however, does not cause any problem for the construction below.

In view of (4.2), the functionals in $\mathcal{N}_m^{v,q}$ are not linearly independent on $\mathcal{S}^{k,r}(\mathcal{T}_m)$ if $q \geq 1$. Namely, the following conditions hold for all $s \in \mathcal{S}^{k,r}(\mathcal{T}_m)$, $v \in \mathcal{V}_m$, $q = 1, \dots, 2r$:

$$(4.5) \quad \eta_{i,\alpha}^{v,q}(s) = \sum_{\beta=0}^{\alpha} (-1)^\beta \binom{\alpha}{\beta} \left(\frac{\sin(\theta_{i-1} + \theta_i)}{|e_i|} \right)^{\alpha-\beta} \left(\frac{\sin \theta_i}{|e_{i-1}|} \right)^\beta \left(\frac{\sin \theta_{i-1}}{|e_{i+1}|} \right)^{-\alpha} \eta_{i-1,q-\beta}^{v,q}(s),$$

$$\alpha = 1, \dots, \min\{r, q\}, \quad i = 1, \dots, N_v,$$

where $\theta_i := \angle e_i e_{i+1}$, $\eta_{i,q}^{v,q} := \eta_{i+1,0}^{v,q}$.

The following key lemma is instrumental in constructing the basis functions.

LEMMA 4.3. *There is a unique spline $s \in \mathcal{S}^{k,r}(\mathcal{T}_m)$ such that*

$$(4.6) \quad \begin{cases} \eta_\xi^\Delta(s) = a_\xi^\Delta, & \xi \in \Xi_\Delta, \Delta \in \mathcal{T}_m, \\ \eta_{q,\xi}^e(s) = a_{q,\xi}^e, & \xi \in \Xi_{e,q}, q = 0, \dots, r, e \in \mathcal{E}_m, \\ \eta^{v,0}(s) = a^{v,0}, & v \in \mathcal{V}_m, \\ \eta_{i,\alpha}^{v,q}(s) = a_{i,\alpha}^{v,q}, & i = 1, \dots, N_v, \alpha = 0, \dots, q-1, q = 1, \dots, 2r, v \in \mathcal{V}_m, \end{cases}$$

for any given $a_\xi^\Delta, a_{q,\xi}^e, a^{v,0} \in \mathbb{R}$ and any $a_{i,\alpha}^{v,q} \in \mathbb{R}$ satisfying

$$(4.7) \quad a_{i,\alpha}^{v,q} = \sum_{\beta=0}^{\alpha} (-1)^\beta \binom{\alpha}{\beta} \left(\frac{\sin(\theta_{i-1} + \theta_i)}{|e_i|} \right)^{\alpha-\beta} \left(\frac{\sin \theta_i}{|e_{i-1}|} \right)^\beta \left(\frac{\sin \theta_{i-1}}{|e_{i+1}|} \right)^{-\alpha} a_{i-1,q-\beta}^{v,q},$$

$$\alpha = 1, \dots, \min\{r, q\}, \quad i = 1, \dots, N_v.$$

Moreover, for each $\Delta \in \mathcal{T}_m$,

$$(4.8) \quad \|s|_\Delta\|_{L^\infty(\Delta)} \leq c \delta_2^{-2r} \max_{\eta \in \mathcal{N}_m(\Delta)} |\eta(s)|,$$

where c is a constant depending only on k , and

$$\mathcal{N}_m(\Delta) := \left(\bigcup_{v \in \mathcal{V}_m \cap \Delta} \mathcal{N}_m^v \right) \cup \left(\bigcup_{\substack{e \in \mathcal{E}_m \\ e \subset \Delta}} \mathcal{N}_m^e \right) \cup \mathcal{N}_m^\Delta.$$

Proof. We first determine $s|_e$ for each $e = [v_1, v_2] \in \mathcal{E}_m$ using the fact that $s|_e$, as a univariate function on the interval e , is a polynomial $s_{e,0}$ of degree at most $k-1$. Therefore, $s_{e,0}$ is uniquely determined by the following k Hermite interpolation conditions:

$$(4.9) \quad \begin{cases} \delta_{v_1} s_{e,0} = a^{v_1,0}, & \delta_{v_2} s_{e,0} = a^{v_2,0}, \\ \delta_{v_1} D_e^\gamma s_{e,0} = a_{i,0}^{v_1,\gamma}, & \delta_{v_2} D_e^\gamma s_{e,0} = a_{j,0}^{v_2,\gamma}, & \gamma = 1, \dots, 2r, \\ \delta_\xi s_{e,0} = a_{0,\xi}^e, & \xi \in \Xi_{e,0}, \end{cases}$$

where we assume that e is the i th edge emanating from v_1 and the j th edge emanating from v_2 .

We next determine $s_{e,q} := (D_{\mu(e,\Delta_e)}^q s)|_e$, $q = 1, \dots, r$. Let $\Delta_e = [v_1, v_2, v_3]$. Then $D_{\mu(e,\Delta_e)}^q = (D_{[v_1,v_3]} - \frac{1}{2}D_{[v_1,v_2]})^q$. Therefore, for $\gamma = 0, \dots, 2r - q$,

$$\delta_{v_1} D_e^\gamma s_{e,q} = \sum_{\ell=0}^q (-1)^\ell 2^{-\ell} \delta_{v_1} D_{e_i}^{\gamma+\ell} D_{e_{i+1}}^{q-\ell} s = \sum_{\ell=0}^q (-1)^\ell 2^{-\ell} \eta_{i,q-\ell}^{v_1,q+\gamma}(s).$$

Similarly, since $D_{\mu(e,\Delta_e)}^q = (D_{[v_2,v_3]} - \frac{1}{2}D_{[v_2,v_1]})^q$, we have for $\gamma = 0, \dots, 2r - q$,

$$\begin{aligned} \delta_{v_2} D_e^\gamma s_{e,q} &= \sum_{\ell=0}^q (-1)^\ell 2^{-\ell} \delta_{v_2} D_{e_j}^{\gamma+\ell} D_{e_{j-1}}^{q-\ell} s \\ &= \sum_{\ell=0}^{q-1} (-1)^\ell 2^{-\ell} \eta_{j-1,\gamma+\ell}^{v_2,q+\gamma}(s) + (-1)^q 2^{-q} \eta_{j,0}^{v_2,q+\gamma}(s). \end{aligned}$$

In addition, we have for each $\xi \in \Xi_{e,q}$,

$$\delta_\xi s_{e,q} = \delta_\xi D_{\mu(e,\Delta_e)}^q s = \eta_{q,\xi}^e(s).$$

Thus, for each $q = 1, \dots, r$, the univariate polynomial $s_{e,q}$ of degree $k-1-q$ is uniquely determined by the $k-q$ Hermite interpolation conditions

$$(4.10) \quad \begin{cases} \delta_{v_1} D_e^\gamma s_{e,q} = \sum_{\ell=0}^q (-1)^\ell 2^{-\ell} a_{i,q-\ell}^{v_1,q+\gamma}, & \gamma = 0, \dots, 2r - q, \\ \delta_{v_2} D_e^\gamma s_{e,q} = \sum_{\ell=0}^{q-1} (-1)^\ell 2^{-\ell} a_{j-1,\gamma+\ell}^{v_2,q+\gamma} + (-1)^q 2^{-q} a_{j,0}^{v_2,q+\gamma}, & \gamma = 0, \dots, 2r - q, \\ \delta_\xi s_{e,q} = a_{q,\xi}^e, & \xi \in \Xi_{e,q}. \end{cases}$$

Let $\tilde{\Delta} = [v_1, v_2, \tilde{v}_3] \in \mathcal{T}_m$ be the second triangle attached to e . We set

$$\begin{aligned} \tilde{a}_{0,\xi}^e &:= a_{0,\xi}^e, & \xi \in \Xi_{e,0}, \\ \tilde{a}_{q,\xi}^e &:= \sum_{\ell=0}^q (-1)^\ell \binom{q}{\ell} (2\sigma|\Delta^*|)^{q-\ell} |\tilde{\Delta}|^\ell |\Delta_e|^{-q} \delta_\xi D_e^{q-\ell} s_{e,\ell}, \\ & \xi \in \Xi_{e,q}, \quad q = 1, \dots, r, \end{aligned}$$

where $\Delta^* := [\bar{v}, v_3, \tilde{v}_3]$, $\bar{v} = (v_1 + v_2)/2$, and $\sigma := \text{sgn} \sin(\widehat{v_3 \bar{v} v_2} + \widehat{\bar{v}_3 \bar{v} v_2})$, where \widehat{uvw} denotes the angle determined by u, v, w with vertex at v . (It may happen that $|\Delta^*| = 0$.) Since $\Delta^* \subset \text{conv}(\Delta_e \cup \tilde{\Delta})$, we have

$$(4.11) \quad |\Delta^*|^{q-\ell} |\tilde{\Delta}|^\ell |\Delta_e|^{-q} \leq \delta_2^{-q+\ell} |\tilde{\Delta}|^q |\Delta_e|^{-q} \leq \delta_2^{-2q+\ell}.$$

We now construct each polynomial patch $s|_\Delta$, $\Delta \in \mathcal{T}_m$, of the spline s as the unique solution of the following interpolation problem:

$$(4.12) \quad \begin{cases} \delta_\xi(s|_\Delta) = a_\xi^\Delta, & \xi \in \Xi_\Delta, \\ \delta_\xi D_{\mu(e,\Delta)}^q(s|_\Delta) = a_{q,\xi}^e, & \xi \in \Xi_{e,q}, \quad q = 0, \dots, r, \quad e \subset \Delta \text{ if } \Delta_e = \Delta, \\ \delta_\xi D_{\mu(e,\Delta)}^q(s|_\Delta) = \tilde{a}_{q,\xi}^e, & \xi \in \Xi_{e,q}, \quad q = 0, \dots, r, \quad e \subset \Delta \text{ if } \Delta_e \neq \Delta, \\ \delta_v(s|_\Delta) = a^{v,0}, & v \in \Delta, \\ \delta_v D_{e_i}^{q-\alpha} D_{e_{i+1}}^\alpha(s|_\Delta) = a_{i,\alpha}^{v,q}, & \alpha = 0, \dots, q, \quad q = 1, \dots, 2r, \quad v \in \Delta, \\ & (i \text{ is such that } e_i, e_{i+1} \subset \Delta). \end{cases}$$

Since (4.12) is a standard finite element interpolation scheme for bivariate polynomials of degree $k - 1$ (see, e.g., [57] or Lemma 3.7 in [25]), the polynomial $s|_\Delta$ is uniquely determined.

We now show that the piecewise polynomial s constructed in this way lies in the space $\mathcal{S}^{k,r}(\mathcal{T}_m)$; i.e., it is r times differentiable. To this end we consider any edge $e = [v_1, v_2] \in \mathcal{E}_m$. As before, let $\Delta_e = [v_1, v_2, v_3]$, and let $\tilde{\Delta} = [v_2, v_1, \tilde{v}_3]$ be the second triangle attached to e , and we again assume that e is the i th edge $e_{1,i}$ emanating from v_1 and at the same time the j th edge $e_{2,j}$ emanating from v_2 . Then we have

$$\begin{aligned} e_{1,i-1} &= [v_1, \tilde{v}_3], & e_{1,i} &= [v_1, v_2], & e_{1,i+1} &= [v_1, v_3], \\ e_{2,j-1} &= [v_2, v_3], & e_{2,j} &= [v_2, v_1], & e_{2,j+1} &= [v_2, \tilde{v}_3]. \end{aligned}$$

Obviously, for each $q = 0, \dots, r$, $D_{\mu(e,\Delta_e)}^q(s|_{\Delta_e})|_e = s_{e,q}$ satisfies the interpolation conditions (4.9) if $q = 0$ or (4.10) if $q > 0$. We set

$$\hat{s}_{e,q} := D_{\mu(e,\tilde{\Delta})}^q(s|_{\Delta_e})|_e.$$

The desired differentiability of s will follow if we show that

$$(4.13) \quad \hat{s}_{e,q} = \tilde{s}_{e,q} := D_{\mu(e,\tilde{\Delta})}^q(s|_{\tilde{\Delta}})|_e, \quad q = 0, \dots, r.$$

By (4.12) we have

$$\begin{aligned} \delta_{v_1}(s|_{\Delta_e}) &= \delta_{v_1}(s|_{\tilde{\Delta}}) = a^{v_1,0}, \\ \delta_{v_1} D_{e_{1,i}}^{q-\alpha} D_{e_{1,i+1}}^\alpha(s|_{\Delta_e}) &= a_{i,\alpha}^{v_1,q}, \quad \alpha = 0, \dots, q-1, \quad q = 1, \dots, 2r, \\ \delta_{v_1} D_{e_{1,i-1}}^{q-\alpha} D_{e_{1,i}}^\alpha(s|_{\tilde{\Delta}}) &= a_{i-1,\alpha}^{v_1,q}, \quad \alpha = 0, \dots, q-1, \quad q = 1, \dots, 2r, \end{aligned}$$

which in view of (4.7) imply

$$\begin{aligned} \delta_{v_1} D_{e_{1,i}}^{q-\alpha} D_{e_{1,i+1}}^\alpha (s|_{\Delta_e}) &= \delta_{v_1} D_{e_{1,i}}^{q-\alpha} D_{e_{1,i+1}}^\alpha (s|_{\bar{\Delta}}), \\ \alpha &= 0, \dots, \min\{r, q\}, \quad q = 0, \dots, 2r, \end{aligned}$$

and hence

$$\delta_{v_1} D_e^\gamma (\hat{s}_{e,q} - \tilde{s}_{e,q}) = 0, \quad \gamma = 0, \dots, 2r - q, \quad q = 0, \dots, r.$$

Similarly, we get

$$\delta_{v_2} D_e^\gamma (\hat{s}_{e,q} - \tilde{s}_{e,q}) = 0, \quad \gamma = 0, \dots, 2r - q, \quad q = 0, \dots, r.$$

In addition, a simple calculation relying on (4.3) shows that

$$\delta_\xi \hat{s}_{e,q} = \tilde{a}_{q,\xi}^e, \quad \xi \in \Xi_{e,q}, \quad q = 0, \dots, r,$$

so that by (4.12),

$$\delta_\xi (\hat{s}_{e,q} - \tilde{s}_{e,q}) = 0, \quad \xi \in \Xi_{e,q}, \quad q = 0, \dots, r.$$

Since $\hat{s}_{e,q} - \tilde{s}_{e,q}$ satisfies homogeneous interpolation conditions of a well-posed Hermite scheme, (4.13) follows.

The uniqueness of s is clear from the above proof, since $s = 0$ if the numbers in the right-hand side of (4.6) are all zeros.

It remains to prove (4.8). Since $s_{e,q}$ satisfies the interpolation conditions (4.9) if $q = 0$ or (4.10) if $q > 0$,

$$\|s_{e,q}\|_{L_\infty(e)} \leq c \max\{\eta(s) : \eta \in \mathcal{N}_m^{v_1} \cup \mathcal{N}_m^{v_2} \cup \mathcal{N}_m^e\}, \quad q = 0, \dots, r,$$

where c depends only on k . In view of (4.11) and Markov inequality (4.1), we have

$$|\tilde{a}_{q,\xi}^e| \leq c \delta_2^{-2q} \|s_{e,q}\|_{L_\infty(e)}, \quad q = 0, \dots, r,$$

and (4.8) follows by the properties of the interpolation problem (4.12); see Lemma 3.9 in [25]. \square

For each $v \in \mathcal{V}_m$ and $q = 1, \dots, 2r$, we denote by $R_m^{v,q}$ the $(\min\{r, q\}N_v \times qN_v)$ -matrix of differentiability conditions (4.5). Let the vectors

$$a^{v,q,j}, \quad j = 1, \dots, \rho_{v,q} := qN_v - \text{rank}(R_m^{v,q}),$$

form an orthonormal basis for the null space of $R_m^{v,q}$,

$$\text{null}(R_m^{v,q}) := \{a \in \mathbb{R}^{qN_v} : R_m^{v,q}a = 0\}.$$

For convenience, we shall use the double indices introduced in the definition of $\mathcal{N}_m^{v,q}$ also for the components of $a^{v,q,j}$:

$$(4.14) \quad a_{i,\alpha}^{v,q,j}, \quad i = 1, \dots, N_v, \quad \alpha = 0, \dots, q - 1.$$

We set

$$(4.15) \quad \eta^{v,q,j} := \sum_{i=1}^{N_v} \sum_{\alpha=0}^{q-1} a_{i,\alpha}^{v,q,j} \eta_{i,\alpha}^{v,q}, \quad j = 1, \dots, \rho_{v,q},$$

$$\begin{aligned} \tilde{\mathcal{N}}_m^{v,q} &:= \{\eta^{v,q,j} : j = 1, \dots, \rho_{v,q}\}, \quad q = 1, \dots, 2r, \\ \tilde{\mathcal{N}}_m^v &:= \mathcal{N}_m^{v,0} \cup \bigcup_{q=1}^{2r} \tilde{\mathcal{N}}_m^{v,q}, \quad v \in \mathcal{V}_m, \\ \tilde{\mathcal{N}}_m &:= \left(\bigcup_{v \in \mathcal{V}_m} \tilde{\mathcal{N}}_m^v \right) \cup \left(\bigcup_{e \in \mathcal{E}_m} \mathcal{N}_m^e \right) \cup \left(\bigcup_{\Delta \in \mathcal{T}_m} \mathcal{N}_m^\Delta \right), \end{aligned}$$

and define the set

$$\Phi_m = \{\varphi_\eta : \eta \in \tilde{\mathcal{N}}_m\}$$

of the *basis functions* for $\mathcal{S}^{k,r}(\mathcal{T}_m)$ by the duality condition,

$$(4.16) \quad \mu(\varphi_\eta) = \begin{cases} 1 & \text{if } \mu = \eta, \\ 0 & \text{if } \mu \in \tilde{\mathcal{N}}_m \setminus \{\eta\}. \end{cases}$$

To see that the above definition is correct we have to check that for each $\eta \in \tilde{\mathcal{N}}_m$ there exists a unique φ_η satisfying (4.16). This follows by Lemma 4.3. Indeed, since the vectors $a^{v,q,j}$ are orthonormal, we have

$$\eta_{i,\alpha}^{v,q} = \sum_{j=1}^{\rho_{v,q}} a_{i,\alpha}^{v,q,j} \eta^{v,q,j}, \quad i = 1, \dots, N_v, \quad \alpha = 0, \dots, q-1.$$

Therefore, for a fixed η , the numbers

$$a_{i,\alpha}^{v,q} := \eta_{i,\alpha}^{v,q}(\varphi_\eta), \quad i = 1, \dots, N_v, \quad \alpha = 0, \dots, q-1,$$

satisfy (4.7), which ensures the applicability of Lemma 4.3.

4.4. Properties of basis splines. It follows by Lemma 4.3 that every spline $s \in \mathcal{S}^{k,r}(\mathcal{T}_m)$ is uniquely determined by the sequence $(\eta(s))_{\eta \in \tilde{\mathcal{N}}_m}$; i.e., s has a unique representation

$$s = \sum_{\eta \in \tilde{\mathcal{N}}_m} a_\eta \varphi_\eta, \quad a_\eta = \eta(s) \in \mathbb{R}.$$

Furthermore, (4.8) immediately implies

$$(4.17) \quad \text{supp } \varphi_\eta \subseteq \begin{cases} \text{star}(v) & \text{if } \eta \in \tilde{\mathcal{N}}_m^v \text{ for a vertex } v \in \mathcal{V}_m, \\ \text{star}(e) & \text{if } \eta \in \mathcal{N}_m^e \text{ for an edge } e \in \mathcal{E}_m, \\ \Delta & \text{if } \eta \in \mathcal{N}_m^\Delta \text{ for a triangle } \Delta \in \mathcal{T}_m, \end{cases}$$

$$(4.18) \quad \|\varphi_\eta\|_{L_\infty(\mathbb{R}^2)} \leq c \delta_2^{-2r}.$$

By using Markov inequality it is easy to show that

$$(4.19) \quad |\eta(s)| \leq \tilde{c} \begin{cases} \|s\|_{L_\infty(\text{star}(v))} & \text{if } \eta \in \tilde{\mathcal{N}}_m^v \text{ for a vertex } v \in \mathcal{V}_m, \\ \|s\|_{L_\infty(\text{star}(e))} & \text{if } \eta \in \mathcal{N}_m^e \text{ for an edge } e \in \mathcal{E}_m, \\ \|s\|_{L_\infty(\Delta)} & \text{if } \eta \in \mathcal{N}_m^\Delta \text{ for a triangle } \Delta \in \mathcal{T}_m, \end{cases}$$

with \tilde{c} a constant depending only on k, r and N_0 .

Thus, we showed that the basis $\Phi_m = \{\varphi_\eta : \eta \in \tilde{\mathcal{N}}_m\}$ satisfies all requirements of section 2.2 with $\mathcal{S}_m = \mathcal{S}^{k,r}(\mathcal{T}_m)$ and $\tilde{k} = k$. (Obviously, $\Pi_k \subset \mathcal{S}_m$ and $\mathcal{S}^{k,r}(\mathcal{T}_m) \subset \mathcal{S}^{k,r}(\mathcal{T}_{m+1})$ if \mathcal{T}_{m+1} is a refinement of \mathcal{T}_m .) More precisely, we have the following result.

THEOREM 4.4. *Let $r \geq 0, k > 4r + 1$. Suppose that \mathcal{T}_m satisfies (a), (d)–(f), and (2.3) of section 2.1. Then the basis functions $\varphi_\eta \in \mathcal{S}^{k,r}(\mathcal{T}_m)$ ($\eta \in \tilde{\mathcal{N}}_m$) constructed above have the following properties:*

(a) *For any $s \in \mathcal{S}^{k,r}(\mathcal{T}_m)$ there exists a unique sequence of real coefficients $(a_\eta)_{\eta \in \tilde{\mathcal{N}}_m}$ such that*

$$s = \sum_{\eta \in \tilde{\mathcal{N}}_m} a_\eta \varphi_\eta,$$

with $a_\eta = \eta(s), \eta \in \tilde{\mathcal{N}}_m$.

(b) *For each $\eta \in \tilde{\mathcal{N}}_m$ there is a vertex $v = v_\eta \in \mathcal{V}_m$ such that*

$$\begin{aligned} \text{supp } \varphi_\eta \subset \text{star}(v) &=: E_\eta, \\ \|\varphi_\eta\|_{L_\infty(\mathbb{R}^2)} &= \|\varphi_\eta\|_{L_\infty(E_\eta)} \leq M_1, \\ |\eta(s)| &\leq M_2 \|s\|_{L_\infty(E_\eta)}, \quad s \in \mathcal{S}^{k,r}(\mathcal{T}_m), \end{aligned}$$

where M_1, M_2 are positive constants depending only on k, r, δ_2 , and N_0 .

In particular, by the proof of Lemma 2.3, we have the following stability property of Φ_m .

THEOREM 4.5. *The basis Φ_m is L_p -stable for all $0 < p \leq \infty$; i.e., for any sequence $(a_\eta)_{\eta \in \tilde{\mathcal{N}}_m}$,*

$$\left\| \sum_{\eta \in \tilde{\mathcal{N}}_m} a_\eta \varphi_\eta \right\|_{L_p(\mathbb{R}^2)} \approx \left(\sum_{\eta \in \tilde{\mathcal{N}}_m} \|a_\eta \varphi_\eta\|_{L_p(\mathbb{R}^2)}^p \right)^{1/p},$$

where the constants of equivalence depend only on p, k, r, δ_2 , and N_0 . In the case $p = \infty$ the ℓ_p -norm in the right-hand side is replaced by the sup-norm.

The linear functionals $\lambda_\eta : \mathcal{S}^{k,-1}(\mathcal{T}_m) \cap L_\infty(E_\eta) \rightarrow \mathbb{R}, \eta \in \tilde{\mathcal{N}}_m$, with properties

$$\lambda_\eta(s|_{E_\eta}) = \eta(s), \quad s \in \mathcal{S}^{k,r}(\mathcal{T}_m),$$

$$|\lambda_\eta(f)| \leq M_2 \|f\|_{L_\infty(E_\eta)}, \quad f \in \mathcal{S}^{k,-1}(\mathcal{T}_m)|_{E_\eta} \cap L_\infty(E_\eta),$$

needed in the definition of the projector Q_m (see (2.18)) can now be defined in a constructive manner. Indeed, we first extend each functional $\eta \in \tilde{\mathcal{N}}_m$ to a functional $\tilde{\eta}$ defined on $\mathcal{S}^{k,-1}(\mathcal{T}_m)$, according to Remark 4.1, and then set

$$\lambda_\eta := \tilde{\eta} \quad \text{if} \quad \eta \in \left(\bigcup_{e \in \mathcal{E}_m} \mathcal{N}_m^e \right) \cup \left(\bigcup_{\Delta \in \mathcal{T}_m} \mathcal{N}_m^\Delta \right)$$

and

$$\lambda_\eta := \sum_{i=1}^{N_v} \sum_{\alpha=0}^{q-1} a_{i,\alpha}^{v,q,j} \tilde{\eta}_{i,\alpha}^{v,q} \quad \text{if} \quad \eta^{v,q,j} = \sum_{i=1}^{N_v} \sum_{\alpha=0}^{q-1} a_{i,\alpha}^{v,q,j} \eta_{i,\alpha}^{v,q} \in \bigcup_{v \in \mathcal{V}_m} \tilde{\mathcal{N}}_m^v.$$

By (2.22), Q_m can be extended to the operator $Q_{m,p} : L_p^{\text{loc}} \rightarrow \mathcal{S}^{k,r}(\mathcal{T}_m)$ whose local approximation power is described in the following theorem (see Lemma 2.5).

THEOREM 4.6. *Suppose $f \in L_p^{\text{loc}}$, $0 < p \leq \infty$ ($f \in C$ if $p = \infty$). Then*

$$\|f - Q_{m,p}(f)\|_{L_p(\Delta)} \leq c\mathcal{S}_\Delta(f)_p \leq cE_k(f, \Omega_\Delta)_p, \quad \Delta \in \mathcal{T}_m,$$

where $\Omega_\Delta := \Omega_\Delta^1$ is the union of all triangles in \mathcal{T}_m that have a common vertex with Δ , and the constant c depends only on p, k, r, δ_2 , and N_0 .

To show that the assumption that the triangulations \mathcal{T}_m satisfy (2.3) cannot be omitted, we consider the following example.

Example 4.7. Suppose \mathcal{T}_m has an edge $e = [v, u]$ with two triangles $\Delta = \Delta_e = [v, u, w]$ and $\tilde{\Delta} = [v, u, \tilde{w}]$ attached to e such that $u = v + (2^{-M}\alpha, 0)$, $w = v + (-\alpha, \alpha)$, $\tilde{w} = v + (-\alpha, -\alpha)$, where the positive numbers M, α depend on m . Evidently, $|\text{conv}(\Delta_e \cup \tilde{\Delta})|/|\Delta_e| = 2(2^M + 1)$, and (2.3) will be violated if M grows unboundedly with m , while the maximal angle of the two triangles is $3\pi/4$, thus allowing the maximal angle condition (2.5) to hold. Note that such configurations of triangles are possible for a sequence of levels of an LR-triangulation \mathcal{T} with the corresponding M 's tending to infinity; see section 2.1 of [38]. Choosing $k = 6$ and $r = 1$, we consider the basis functions $\varphi_\eta \in \mathcal{S}^{6,1}(\mathcal{T}_m)$, $\eta \in \tilde{\mathcal{N}}_m$, constructed according to the above algorithm. We next show that the basis $\Phi_m = \{\varphi_\eta : \eta \in \tilde{\mathcal{N}}_m\}$ is *unstable*; i.e., Theorem 4.5 does not hold for it. (Therefore, neither Φ_m nor a renorming of it satisfies the requirements of section 2.2.) More precisely, we show that the constant function $\mathbb{1}_{\mathbb{R}^2}(x) \equiv 1$, $x \in \mathbb{R}^2$, does not have an L_∞ -stable expansion with respect to Φ_m . We have

$$\|\mathbb{1}_{\mathbb{R}^2}\|_{L_\infty(\mathbb{R}^2)} = 1, \quad \mathbb{1}_{\mathbb{R}^2} = \sum_{\eta \in \tilde{\mathcal{N}}_m} \eta(\mathbb{1}_{\mathbb{R}^2})\varphi_\eta.$$

Now choose $\eta = \eta^{v,0} = \delta_v \in \mathcal{N}_m^{v,0}$. Since $\eta(\mathbb{1}_{\mathbb{R}^2}) = 1$, the instability of Φ_m will follow if we show that $\|\varphi_\eta\|_{L_\infty(\mathbb{R}^2)}$ is unbounded as $M \rightarrow \infty$. By (4.12),

$$\begin{aligned} \delta_\xi D_{\mu(e,\tilde{\Delta})}^1(\varphi_\eta|_{\tilde{\Delta}}) &= \tilde{a}_{1,\xi}^e \\ &= -2 \frac{|\Delta^*|}{|\Delta_e|} \delta_\xi D_e^1 s_{e,0} - \frac{|\tilde{\Delta}|}{|\Delta_e|} \delta_\xi s_{e,1}, \end{aligned}$$

where $\xi = (v + u)/2$, $\Delta^* = [\xi, w, \tilde{w}]$, $s_{e,0} = \varphi_\eta|_e$, $s_{e,1} = (D_{\mu(e,\Delta_e)}^1 \varphi_\eta)|_e$. Obviously, $|\tilde{\Delta}|/|\Delta_e| = 1$, and

$$|\Delta^*|/|\Delta_e| = \left(|\text{conv}(\Delta_e \cup \tilde{\Delta})| - \frac{|\Delta_e| + |\tilde{\Delta}|}{2} \right) / |\Delta_e| = 2^{M+1} + 1.$$

The univariate polynomial $s_{e,0}$ of degree 5 is determined by the Hermite interpolation conditions (4.9) that take in our case the form

$$\delta_v s_{e,0} = 1, \quad \delta_u s_{e,0} = \delta_u D_e^1 s_{e,0} = \delta_u D_e^2 s_{e,0} = \delta_v D_e^1 s_{e,0} = \delta_v D_e^2 s_{e,0} = 0.$$

An elementary computation shows that $\delta_\xi D_e^1 s_{e,0} = -15/8$. By (4.10), we immediately get $\delta_\xi s_{e,1} = a_{1,\xi}^e = \eta_{1,\xi}^e(\varphi_\eta) = 0$. Thus,

$$\delta_\xi D_{\mu(e,\tilde{\Delta})}^1(\varphi_\eta|_{\tilde{\Delta}}) = \frac{15}{4}(2^{M+1} + 1) \rightarrow \infty \quad \text{as} \quad M \rightarrow \infty.$$

In view of Markov inequality, $\|\varphi_\eta\|_{L_\infty(\mathbb{R}^2)} \geq c|\delta_\xi D_{\mu(e,\bar{\Delta})}^1(\varphi_\eta|_{\bar{\Delta}})|$, and we get the desired unboundedness of $\|\varphi_\eta\|_{L_\infty(\mathbb{R}^2)}$ for sufficiently large M .

Remark 4.8. It is clear that Theorems 4.4–4.6 are valid for any sequence of levels \mathcal{T}_m satisfying the hypotheses of Theorem 4.4; i.e., nestedness and other additional assumptions on $\{\mathcal{T}_m\}$ stated in section 2.1 are not needed for these results.

Remark 4.9. It is an important property of the basis functions φ_η constructed above that they are invariant under affine transforms. More precisely, let \mathcal{T}_m satisfy the hypotheses of Theorem 4.4, and let $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be an affine transform. We set $A(\mathcal{T}_m) = \{A(\Delta) : \Delta \in \mathcal{T}_m\}$, and $A\eta(s) := \delta_{A(v)}D_{A(e_1)}^\alpha D_{A(e_2)}^\beta s$, for each nodal functional η of the form $\eta(s) = \delta_v D_{e_1}^\alpha D_{e_2}^\beta s$ and extend the operator A linearly to the linear combinations of the nodal functionals such as those occurring in (4.15). Then, clearly, the sets of nodal functionals \mathcal{N}_m and \mathcal{N}_m^A defined by (4.4) for \mathcal{T}_m and $A(\mathcal{T}_m)$, respectively, satisfy $\mathcal{N}_m^A = \{A\eta : \eta \in \mathcal{N}_m\}$. (We used here, in particular, the fact that $\mu(A(e), A(\Delta_e)) = A(\mu(e, \Delta_e))$.) Moreover, since the matrices $R_m^{v,q}$ of the differentiability conditions (4.5) are affine invariant (see (4.3)), we also have $\tilde{\mathcal{N}}_m^A = \{A\eta : \eta \in \tilde{\mathcal{N}}_m\}$ for the appropriate sets $\tilde{\mathcal{N}}_m, \tilde{\mathcal{N}}_m^A$ defined as in the construction above, provided we choose the same orthonormal vectors (4.14) in both cases. Let now $\Phi_m = \{\varphi_\eta : \eta \in \tilde{\mathcal{N}}_m\} \subset \mathcal{S}^{k,r}(\mathcal{T}_m)$ and $\Phi_m^A = \{\varphi_{A\eta} : \eta \in \tilde{\mathcal{N}}_m\} \subset \mathcal{S}^{k,r}(A(\mathcal{T}_m))$ be the spline bases dual to $\tilde{\mathcal{N}}_m$ and $\tilde{\mathcal{N}}_m^A$, respectively. Since $\varphi_\eta(A \cdot), \eta \in \tilde{\mathcal{N}}_m$, obviously satisfy the same duality relations, we conclude that $\varphi_{A\eta} = \varphi_\eta(A \cdot), \eta \in \tilde{\mathcal{N}}_m$, which is the desired affine invariance.

Remark 4.10. Our construction is extendable to the spaces $\mathcal{S}^{k,r}(\mathcal{T}_m), k > r2^d + 1$, in dimensions $d > 2$. To this end the algorithm given in [22] should be extended to SLR-triangulations in \mathbb{R}^d . In particular, the orthogonal directions of derivatives used in [22] should be replaced by appropriate affine invariant directions.

Remark 4.11. If the triangulation covers only a compact domain E , then usual modifications of basis functions corresponding to boundary edges or vertices (see [22, 23]) lead to the desired stable local bases.

Remark 4.12. In this section, we extended to the setting of SLR-triangulations the bivariate version of nodal stable local basis construction of [22, 23], which was originally designed for regular triangulations. The scheme from [27] can be used as an alternative means of constructing stable local bases for $\mathcal{S}^{k,r}(\mathcal{T}_m), k > 3r + 2$, in dimension $d = 2$. Such a development would take advantage of the affine invariance of the Bernstein–Bézier representation of piecewise polynomials. We elected to utilize the scheme from [22] instead, since it is available for any number of variables and allows an effective numerical implementation as shown (for $r = 1, 2, d = 2$) in [23]. Also, we want to pay heed to two more spline basis constructions (for regular triangulations in dimension $d = 2$) that allow the same kind of extension to SLR-triangulations: (a) stable local bases for $\mathcal{S}^{k,1}(\mathcal{T}_m), k > 5$, constructed in [26]; (b) locally stable bases on nested triangulations ($k > 4r + 1$) [24]. Note that the stable local bases for superspline subspaces of $\mathcal{S}^{k,r}(\mathcal{T}_m)$ [16, 17, 44, 57] cannot be used since these spaces are not nested for nested triangulations, while the earlier local spline bases for $\mathcal{S}^{k,r}(\mathcal{T}_m)$ [1, 8, 18, 35, 36, 48] are known to be unstable for certain triangulations.

Remark 4.13. It is easy to see that, in the case $r = 0$, the above basis reduces to the classical Lagrange finite element basis for $\mathcal{S}^{k,0}(\mathcal{T}_m), k > 1$. Since δ_2 disappears from (4.8) when $r = 0$, Theorems 4.4–4.6 hold for locally regular triangulations; i.e., the SLR assumption (2.3) is not needed in this case. (Note that δ_2 and N_0 completely disappear from Theorem 4.4, and δ_2 is replaced by δ_1 in Theorems 4.5–4.6.) For $r = 0, k = 2$, we get the Courant elements, and the only essential difference

to the construction from [38] is that we rely here on the extensions of linear functionals described in Remark 4.1 rather than on the explicit quasi interpolant for continuous piecewise linear functions adopted in [38]. Both approaches obviously lead to the same B-spaces.

5. Spline bases on special triangulations. There are several constructions of differentiable spline bases fitting into our scheme that are only available for specific multilevel triangulations. Since these triangulations have a special structure or even are uniform, the corresponding libraries $\{\Phi_{\mathcal{T}}\}$ of bases are not as rich as the one of the previous section associated with arbitrary SLR-triangulations. Moreover, the necessity to maintain the structure of the triangulation highly reduces the variety of refinement methods that can be used (whereas, e.g., *local* refinement by bisection can be used with bases on arbitrary triangulations). On the other hand, bases on special triangulations usually allow a smaller degree of piecewise polynomials for a given order of differentiability as well as a simpler and more efficient practical implementation.

In this section, we review some known constructions of this type. (Note that only box splines are available for more than two variables.)

5.1. Box splines. As usual, we consider only splines of two variables. Let $\Xi = [\xi_1 \cdots \xi_n]$ be a full rank $2 \times n$ matrix with columns ξ_i in $\mathbb{Z}^2 \setminus 0$. The *box spline* $M_{\Xi} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ associated with Ξ is defined by its Fourier transform

$$\hat{M}_{\Xi}(u) = \prod_{\nu=1}^n \frac{1 - e^{-i\xi_{\nu}u}}{i\xi_{\nu}u}, \quad u \in \mathbb{R}^2,$$

where $\xi_{\nu}u$ denotes the inner product of the two vectors.

We now review the basic properties of box splines (see [9]), in order to verify the requirements of section 2.2. It is well known that M_{Ξ} has a compact support,

$$(5.1) \quad \text{supp } M_{\Xi} = \left\{ \sum_{\nu=1}^n t_{\nu} \xi_{\nu} : 0 \leq t_{\nu} \leq 1 \right\}.$$

The box spline basis functions at the m th level are defined by

$$\varphi_{m,j} = M_{\Xi}(2^m \cdot -j), \quad j \in \mathbb{Z}^2.$$

We set

$$\Phi_m = \{\varphi_{m,j} : j \in \mathbb{Z}^2\}, \quad m \in \mathbb{Z},$$

and

$$\mathcal{S}_m = \left\{ \sum_{j \in \mathbb{Z}^2} a_{m,j} \varphi_{m,j} : a_{m,j} \in \mathbb{R} \right\}, \quad m \in \mathbb{Z},$$

where the series converges everywhere since for every $x \in \mathbb{R}^2$ and $m \in \mathbb{Z}$ only a finite number of $\varphi_{m,j}(x)$ ($j \in \mathbb{Z}^2$) are nonzero. Clearly, any affine change of variables $Q : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ gives rise to basis functions $\varphi_{m,j}(Qx)$ that satisfy the conditions of section 2.2 if and only if the $\varphi_{m,j}$ do. Therefore, we do not distinguish between constructions that can be transformed into each other by such a method.

Since

$$\frac{\hat{M}_\Xi(2u)}{\hat{M}_\Xi(u)} = \prod_{\nu=1}^n \frac{1 + e^{-i\xi_\nu u}}{2},$$

M_Ξ is a finite linear combination of $M_\Xi(2 \cdot -j)$, $j \in \mathbb{Z}^2$, which implies that

$$\mathcal{S}_m \subset \mathcal{S}_{m+1}, \quad m \in \mathbb{Z}.$$

Let

$$r(\Xi) := \max\{r : \text{any } 2 \times (n - r) \text{ submatrix of } \Xi \text{ has rank } 2\} - 1$$

and

$$k(\Xi) := n - 1.$$

The elements of \mathcal{S}_m are $r(\Xi)$ times differentiable piecewise polynomials of degree $k(\Xi) - 1$ with respect to the rectilinear partition \mathcal{T}_m^Ξ of \mathbb{R}^2 determined by the straight lines

$$H_\nu + 2^{-m}j, \quad j \in \mathbb{Z}^2, \quad \nu = 1, \dots, n,$$

where

$$H_\nu := \{t\xi_\nu : t \in \mathbb{R}\}.$$

Thus,

$$\mathcal{S}_m \subset \mathcal{S}^{k(\Xi), r(\Xi)}(\mathcal{T}_m^\Xi).$$

Moreover,

$$\Pi_{\tilde{k}(\Xi)} \subset \mathcal{S}_m, \quad m \in \mathbb{Z},$$

and $\Pi_{\tilde{k}(\Xi)+1} \not\subset \mathcal{S}_m$, where

$$\tilde{k}(\Xi) = r(\Xi) + 2.$$

It is well known that the translates of a box spline are not always linearly independent. In fact, Φ_m is a basis for \mathcal{S}_m ($m \in \mathbb{Z}$) if and only if the matrix Ξ is *unimodular*; i.e., each nonsingular 2×2 submatrix of Ξ has determinant ± 1 . This condition implies substantial restrictions on Ξ . Namely, up to an affine change of variables, Ξ must have the form

$$\Xi = \underbrace{[e_1 \cdots e_1]}_{n_1} \underbrace{[e_2 \cdots e_2]}_{n_2} \underbrace{[e_3 \cdots e_3]}_{n_3},$$

where $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $e_3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $n_1, n_2 \geq 1$, $n_3 \geq 0$, and $n_1 + n_2 + n_3 = n$. It is easy to see that

$$r(\Xi) = n - \max\{n_1, n_2, n_3\} - 2$$

and that \mathcal{T}_m^Ξ is either a tensor product mesh if $n_3 = 0$ or a *three-directional mesh* $\mathcal{T}_m^{(1)}$ defined by the straight lines $x_1 = 2^{-m}j$, $x_2 = 2^{-m}j$, $x_1 - x_2 = 2^{-m}j$ ($j \in \mathbb{Z}^2$) in \mathbb{R}^2 if $n_3 \geq 1$. Since only the latter case leads to a multilevel triangulation, we assume that $n_3 \geq 1$.

It remains to verify (2.10)–(2.12). By (5.1), the support of M_Ξ is the hexagon with vertices $(0, 0)$, $(n_1, 0)$, $(0, n_2)$, $(n_1 + n_3, n_3)$, $(n_3, n_2 + n_3)$, $(n_1 + n_3, n_2 + n_3)$, which implies (2.10) with $\ell \leq \lfloor n/2 \rfloor$. Obviously, (2.11) is valid with $M_1 = \|M_\Xi\|_{L^\infty}$. Finally, it is easy to show (2.12) by using the constructions of dual functionals $\lambda_j : \mathcal{S}_0 \rightarrow \mathbb{R}$ ($j \in \mathbb{Z}^2$), with $\lambda_j(\varphi_{0,k}) = \delta_{j,k}$, given, e.g., in [19, 37, 41].

Let us mention the following two cases that are perhaps most relevant in applications:

- (a) $n_1 = n_2 = 2$, $n_3 = 1$, $\mathcal{S}_m \subset \mathcal{S}^{4,1}(\mathcal{T}_m^{(1)})$, $\tilde{k} = 3$,
- (b) $n_1 = n_2 = n_3 = 2$, $\mathcal{S}_m \subset \mathcal{S}^{5,2}(\mathcal{T}_m^{(1)})$, $\tilde{k} = 4$.

5.2. Other spline bases on uniform triangulations. There are some other spline basis constructions for the three-directional mesh $\mathcal{T}_m^{(1)}$; see, e.g., [15, 56]. However, to our knowledge, none of them simultaneously satisfies the requirements of nestedness of the spaces, stability, and locality of the basis functions. The situation is better for the *four-directional mesh* $\mathcal{T}_m^{(2)}$ obtained from $\mathcal{T}_m^{(1)}$ by adding the straight lines $x_1 + x_2 = 2^{-m}j$ ($j \in \mathbb{Z}^2$). Since $\mathcal{T}_m^{(2)}$ is a special case of a so-called FVS-triangulation (see section 5.3), finite element bases for $\mathcal{S}^{4,1}(\mathcal{T}_m^{(2)})$ are available and satisfy the conditions of section 2.2. Some recent alternative constructions of stable local bases for $\mathcal{S}^{4,1}(\mathcal{T}_m^{(2)})$ can be found in [13, 28, 42, 49]. Moreover, a stable local basis for $\mathcal{S}^{7,2}(\mathcal{T}_m^{(2)})$ is also constructed in [28]. Finally, we want to mention the stable local basis from [33] for C^1 quadratic splines with respect to a sequence of triangulation levels that can be called the *six-directional meshes*.

5.3. Refinable composite finite elements. Multilevel and hierarchical bases play an important role in the modern theory and practice of numerical methods for PDEs; see, e.g., [51]. Classical smooth finite elements [14] give rise to stable local spline bases on triangulations satisfying the minimal angle condition. (Note that it should be possible to replace this condition of regularity with SLR.) However, there are difficulties in using them to build nested spline spaces on multilevel triangulations; see [11, 20]. Although the “polynomial” finite elements (e.g., the Argyris element) are available for arbitrary triangulations, they lead to *superspline* spaces [57] that lack nestedness for nested triangulations (*levels* in the terminology of our section 2). In contrast to them, “composite” finite elements require a special structure of the levels \mathcal{T}_m , e.g., a Clough–Tocher or Powell–Sabin split, which is not always compatible with nested refinements with other desirable properties like boundedness of the valence of the vertices. In fact, we are aware of only two cases when composite finite elements are *refinable*, i.e., provide stable local bases for certain multilevel triangulations. First, this is true for the triangulations obtained by the *Powell–Sabin 12-split*; see [50] for the relevant construction of stable local bases for C^1 quadratics and cubics. The other case is that of *FVS-triangulations* obtained from arbitrary strictly convex *quadrangulations* by adding two diagonals of each quadrilateral; see, e.g., [20, 43]. Here, a well-known composite finite element due to Fraeijns de Veubeke and Sander gives rise to a stable local basis for C^1 cubics, while for higher orders of differentiability only nonnested superspline-type constructions are known [40, 45, 46].

Appendix A.

Proof of Theorem 2.9. Denote briefly $N := (\sum_{\theta \in \Theta} \|c_\theta \varphi_\theta\|_p^\tau)^{1/\tau}$.

Case 1. $0 < p \leq 1$. Since $\tau < p \leq 1$, we have

$$\left\| \sum_{\theta \in \Theta} |c_\theta \varphi_\theta(\cdot)| \right\|_p \leq \left(\sum_{\theta \in \Theta} \|c_\theta \varphi_\theta\|_p^p \right)^{1/p} \leq \left(\sum_{\theta \in \Theta} \|c_\theta \varphi_\theta\|_p^\tau \right)^{1/\tau}.$$

Case 2. $p = \infty$. Since $\tau = 1/\alpha \leq 1$, then (2.42) is obvious.

Case 3. $1 < p < \infty$. We need the following lemma.

LEMMA A.1. *Let $g := \sum_{\theta \in \mathcal{M}} |c_\theta \varphi_\theta|$, where $\#\mathcal{M} < \infty$ and $\|c_\theta \varphi_\theta\|_p \leq L$ for $\theta \in \mathcal{M}$. Then*

$$\|g\|_p \leq cL(\#\mathcal{M})^{1/p},$$

with c independent of \mathcal{M} and $(c_\theta)_{\theta \in \mathcal{M}}$.

Proof. Using the properties of Φ , we have (recall that $\text{supp } \varphi_\theta \subset E_\theta := \text{star}^\ell(v_\theta)$ and $\|\varphi_\theta\|_\infty \approx |E_\theta|^{-1/p} \|\varphi_\theta\|_p$ by (2.14))

$$\|g\|_p \leq \left\| \sum_{\theta \in \mathcal{M}} \|c_\theta \varphi_\theta\|_\infty \cdot \mathbf{1}_{E_\theta}(\cdot) \right\|_p \leq cL \left\| \sum_{\theta \in \mathcal{M}} |E_\theta|^{-1/p} \cdot \mathbf{1}_{E_\theta}(\cdot) \right\|_p.$$

We define $E := \bigcup_{\theta \in \mathcal{M}} E_\theta$ and $E(x) := \min\{|E_\theta| : \theta \in \mathcal{M} \text{ and } E_\theta \ni x\}$ for $x \in E$. By the properties of the LR-triangulations, it follows that

$$\sum_{\theta \in \mathcal{M}} |E_\theta|^{-1/p} \cdot \mathbf{1}_{E_\theta}(x) \leq cE(x)^{-1/p} \mathbf{1}_E(x), \quad x \in \mathbb{R}^2.$$

On the other hand,

$$E(x)^{-1} = \max_{\theta \in \mathcal{M}, E_\theta \ni x} |E_\theta|^{-1} \leq \sum_{\theta \in \mathcal{M}} |E_\theta|^{-1} \mathbf{1}_{E_\theta}(x).$$

Therefore,

$$\begin{aligned} \|g\|_p &\leq cL \|E(\cdot)^{-1/p}\|_{L^p} = cL \left(\int_E E(x)^{-1} dx \right)^{1/p} \\ &\leq cL \left(\sum_{\theta \in \mathcal{M}} |E_\theta|^{-1} \int_{\mathbb{R}^2} \mathbf{1}_{E_\theta}(x) dx \right)^{1/p} = cL(\#\mathcal{M})^{1/p}. \quad \square \end{aligned}$$

We define

$$\mathcal{F}_\mu := \{\theta : 2^{-\mu}N \leq \|c_\theta \varphi_\theta\|_p < 2^{-\mu+1}N\},$$

where $N := (\sum_{\theta \in \Theta} \|c_\theta \varphi_\theta\|_p^\tau)^{1/\tau}$. Then

$$\bigcup_{\nu \leq \mu} \mathcal{F}_\nu = \{\theta : \|c_\theta \varphi_\theta\|_p \geq 2^{-\mu}N\},$$

and hence

$$(A.1) \quad \#\mathcal{F}_\mu \leq \sum_{\nu \leq \mu} \#\mathcal{F}_\nu = \#\left(\bigcup_{\nu \leq \mu} \mathcal{F}_\nu\right) \leq 2^{\mu\tau}.$$

We set $F_\mu := \sum_{\theta \in \mathcal{F}_\mu} |c_\theta \varphi_\theta|$. Using Lemma A.1 and (A.1), we obtain

$$\begin{aligned} \left\| \sum_{\theta \in \Theta} |c_\theta \varphi_\theta(\cdot)| \right\|_p &\leq \left\| \sum_{\mu=0}^{\infty} F_\mu(\cdot) \right\|_p \leq \sum_{\mu=0}^{\infty} \|F_\mu\|_p \leq c \sum_{\mu=0}^{\infty} 2^{-\mu} N(\#\mathcal{F}_\mu)^{1/p} \\ &\leq cN \sum_{\mu=0}^{\infty} 2^{-\mu(1-\tau/p)} \leq cN \sum_{\mu=0}^{\infty} 2^{-\mu\tau\alpha} \leq cN. \end{aligned}$$

This completes the proof of Theorem 2.9. \square

Proof of Theorem 3.1 (the case $0 < p < \infty$). Suppose $f \in B_\tau^\alpha(\Phi)$, where $\alpha > 0$, $1/\tau = \alpha + 1/p$, $0 < p < \infty$. By (2.40), f can be represented in the form $f = \sum_{\theta \in \Theta} b_\theta \varphi_\theta$ with the series converging absolutely a.e. in \mathbb{R}^2 and in L_p . We denote briefly $N(f) := N_{\Phi, Q, \tau}(f) := (\sum_{\theta \in \Theta} \|b_\theta \varphi_\theta\|_p^\tau)^{1/\tau} \approx \|f\|_{B_\tau^\alpha(\Phi)}$.

Suppose that $(b_{\theta_j} \varphi_{\theta_j})_{j=1}^\infty$ is a rearrangement of the sequence $(b_\theta \varphi_\theta)_{\theta \in \Theta}$ such that $\|b_{\theta_1} \varphi_{\theta_1}\|_p \geq \|b_{\theta_2} \varphi_{\theta_2}\|_p \geq \dots$. Set $s_n := \sum_{j=1}^n b_{\theta_j} \varphi_{\theta_j}$, $s_n \in \Sigma_n(\Phi)$.

Case 1. $0 < p \leq 1$. To estimate $\|f - s_n\|_p$ we shall use the following simple inequality [38]: If $x_1 \geq x_2 \geq \dots \geq 0$ and $0 < \tau < p$, then

$$(A.2) \quad \left(\sum_{j=n+1}^{\infty} x_j^p \right)^{1/p} \leq n^{1/p-1/\tau} \left(\sum_{j=1}^{\infty} x_j^\tau \right)^{1/\tau}.$$

We use Theorem 2.9 and apply (A.2) with $x_j := \|b_{\theta_j} \varphi_{\theta_j}\|_p$ to obtain

$$\begin{aligned} \|f - s_n\|_p &\leq \left\| \sum_{j=n+1}^{\infty} |b_{\theta_j} \varphi_{\theta_j}| \right\|_p \leq \left(\sum_{j=n+1}^{\infty} \|b_{\theta_j} \varphi_{\theta_j}\|_p^p \right)^{1/p} \\ &\leq n^{1/p-1/\tau} \left(\sum_{j=1}^{\infty} \|b_{\theta_j} \varphi_{\theta_j}\|_p^\tau \right)^{1/\tau} = n^{-\alpha} N(f), \end{aligned}$$

which proves Theorem 3.1 in Case 1.

Case 2. $1 < p < \infty$. We proceed quite similarly as in the proof of Theorem 2.9. We set $\mathcal{F}_\mu := \{\theta : 2^{-\mu} N(f) \leq \|b_\theta \varphi_\theta\|_p < 2^{-\mu+1} N(f)\}$ and $F_\mu := \sum_{\theta \in \mathcal{F}_\mu} |b_\theta \varphi_\theta|$.

Fix $m \geq 1$ and set $M := \lceil 2^{m\tau} \rceil$. As in the proof of Theorem 2.9 (see (A.1)), $\#\mathcal{F}_m \leq \sum_{\nu \leq m} \#\mathcal{F}_\nu \leq 2^{m\tau} \leq M$. Using Lemma A.1, we obtain

$$\begin{aligned} \|f - s_M\|_p &\leq \left\| \sum_{\mu=m+1}^{\infty} F_\mu \right\|_p \leq \sum_{\mu=m+1}^{\infty} \|F_\mu\|_p \\ &\leq c \sum_{\mu=m+1}^{\infty} 2^{-\mu} N(f) (\#\mathcal{F}_\mu)^{1/p} \leq cN(f) \sum_{\mu=m+1}^{\infty} 2^{-\mu(1-\tau/p)} \\ &\leq cN(f) 2^{-m(1-\tau/p)} \leq cM^{-1/\tau+1/p} N(f) = cM^{-\alpha} N(f). \end{aligned}$$

This estimate readily implies (3.1). \square

Proof of Theorem 3.2. Step 1. With this step we lay some groundwork that is needed for the proof of the Bernstein inequality. Let \mathcal{T} be an arbitrary LR-triangulation and suppose Λ is a finite subset of \mathcal{T} . The set Λ generates a certain tree structure that we want to bring up in what follows.

We say that $\Delta \in \mathcal{T}$ is a *branching triangle* if at least two children of Δ have descendants in Λ . Let $\tilde{\Lambda}$ denote the extension of Λ obtained by adding all branching triangles and all children of branching triangles if they are not already in Λ . By considering the *tree* of the ancestors of all triangles in Λ , it is not difficult to see that the total number of branching triangles does not exceed $\#\Lambda - 1$. Since the number of children of a triangle is bounded by M_0 , we conclude that $\#\tilde{\Lambda} \leq c\#\Lambda$.

Furthermore, for a later use in Step 3, we call $\Delta \in \mathcal{T} \setminus \tilde{\Lambda}$ a *chain triangle* if at least one of its descendants belongs to Λ . The set of all chain triangles will be denoted by Γ . By construction, for each $\Delta \in \Gamma$ there is a unique largest triangle $\tilde{\Delta} \in \tilde{\Lambda}$ contained in Δ . We set $K_\Delta := \Delta \setminus \tilde{\Delta}$ and $\mu_\Delta := m - \tilde{m}$, where $\Delta \in \mathcal{T}_m$ and $\tilde{\Delta} \in \mathcal{T}_{\tilde{m}}$. We denote by $\tilde{\Gamma}$ the set of all $\Delta \in \Gamma$ for which there is a $\Delta' \in \tilde{\Lambda}$ containing Δ . It is easy to see that $\tilde{\Gamma}$ is the disjoint union of *finite chains*, i.e., sets λ of the form $\lambda = \{\Delta_1, \dots, \Delta_\nu\} \subset \tilde{\Gamma}$ ($\nu \geq 1$), where $\Delta''_\lambda \supset \Delta_1 \supset \dots \supset \Delta_\nu \supset \Delta'_\lambda$ for some $\Delta'_\lambda, \Delta''_\lambda \in \tilde{\Lambda}$, and Δ_1 is a child of Δ''_λ , Δ_j is a child of Δ_{j-1} , $\nu = 2, \dots, \nu$, and Δ'_λ is a child of Δ_ν . Similarly, $\Gamma \setminus \tilde{\Gamma}$ is the disjoint union of *infinite chains* $\lambda = \{\dots, \Delta_{-2}, \Delta_{-1}\} \subset \Gamma$, where $\dots \supset \Delta_{-2} \supset \Delta_{-1} \supset \Delta'_\lambda$ for some $\Delta'_\lambda \in \tilde{\Lambda}$, and Δ_j is a child of Δ_{j-1} , $\nu = -1, -2, \dots$, and Δ'_λ is a child of Δ_{-1} . We let \mathcal{L} and \mathcal{L}^∞ denote the sets of all finite, respectively, infinite chains in Γ . Clearly, $\#\mathcal{L} \leq \#\tilde{\Lambda}$ and $\#\mathcal{L}^\infty \leq \#\tilde{\Lambda}$.

Step 2. For the proof of the theorem in the case $0 < p < \infty$, we need the following lemma.

LEMMA A.2. *Suppose $s = \sum_{\Delta \in \Lambda} \mathbb{1}_\Delta \cdot P_\Delta$, where $P_\Delta \in \Pi_k$ ($k \geq 1$), $\Lambda \subset \mathcal{T}$ with \mathcal{T} an LR-triangulation, and $\#\Lambda < \infty$. Then*

$$\left(\sum_{\Delta \in \Lambda} |\Delta|^{-\alpha\tau} \|s\|_{L^\tau(\Delta)}^\tau \right)^{1/\tau} \leq c(\#\Lambda)^\alpha \|s\|_p,$$

with c independent of s and Λ .

Proof. We adopt all necessary notation from Step 1 above with Λ from the hypotheses of the lemma. Since $\#\tilde{\Lambda} \leq c\#\Lambda$ and $s = \sum_{\Delta \in \tilde{\Lambda}} \mathbb{1}_\Delta \cdot P_\Delta$, where $P_\Delta = 0$ for $\Delta \in \tilde{\Lambda} \setminus \Lambda$, we may assume without loss of generality that $\tilde{\Lambda} = \Lambda$; i.e., the branching triangles and their children are contained in Λ .

Let $\Delta_1, \dots, \Delta_m$ be all nonbranching triangles in Λ . It is not difficult to see that for each of them there are only two possibilities: either Δ_i does not contain any other $\Delta \in \Lambda$ (in which case we call Δ_i a *final triangle*) or there is a unique largest triangle $\tilde{\Delta}_i \in \Lambda$ strictly contained in Δ_i . We define the *rings* $K_i := \Delta_i \setminus \tilde{\Delta}_i$, $i = 1, \dots, m$, where $\tilde{\Delta}_i := \emptyset$ for a final triangle Δ_i . Obviously, K_i have pairwise disjoint interiors, and $s|_{K_i} = P_i|_{K_i}$, for some $P_i \in \Pi_k$, $i = 1, \dots, m$. Since all children of branching triangles are in Λ , we have for each $\Delta \in \Lambda$,

$$\Delta = \bigcup_{\substack{i=1 \\ \Delta_i \subset \Delta}}^m K_i \quad \text{and} \quad s|_\Delta = \sum_{\substack{i=1 \\ \Delta_i \subset \Delta}}^m \mathbb{1}_{K_i} \cdot P_i.$$

Therefore,

$$\begin{aligned} \sum_{\Delta \in \Lambda} |\Delta|^{-\alpha\tau} \|s\|_{L_\tau(\Delta)}^\tau &= \sum_{\Delta \in \Lambda} |\Delta|^{-\alpha\tau} \sum_{\substack{i=1 \\ \Delta_i \subset \Delta}}^m \|s\|_{L_\tau(K_i)}^\tau \\ &= \sum_{i=1}^m \|s\|_{L_\tau(K_i)}^\tau \sum_{\Delta \in \Lambda, \Delta \supset \Delta_i} |\Delta|^{-\alpha\tau} \\ &= \sum_{i=1}^m \|s\|_{L_\tau(K_i)}^\tau |\Delta_i|^{-\alpha\tau} \sum_{\Delta \in \Lambda, \Delta \supset \Delta_i} (|\Delta_i|/|\Delta|)^{\alpha\tau} \\ &\leq c \sum_{i=1}^m \|s\|_{L_\tau(K_i)}^\tau |\Delta_i|^{-\alpha\tau}, \end{aligned}$$

where we once switched the order of summation and used (2.9). Since $|\tilde{\Delta}_i| \leq \rho|\Delta_i|$, we have by Lemma 2.2,

$$\|P_i\|_{L_\tau(K_i)} \approx |K_i|^{1/\tau-1/p} \|P_i\|_{L_p(K_i)} \approx |\Delta_i|^\alpha \|P_i\|_{L_p(K_i)},$$

which implies $\|s\|_{L_\tau(K_i)}^\tau |\Delta_i|^{-\alpha\tau} \approx \|s\|_{L_p(K_i)}^\tau$, $i = 1, \dots, m$. Now by Hölder's inequality,

$$\sum_{i=1}^m \|s\|_{L_p(K_i)}^\tau \leq \left(\sum_{i=1}^m \|s\|_{L_p(K_i)}^p \right)^{\tau/p} m^{1-\tau/p} \leq (\#\Lambda)^{\alpha\tau} \|s\|_p^\tau,$$

and the proof is complete. \square

Step 3. Let $s \in \Sigma_n(\Phi)$ and suppose that $s =: \sum_{\theta \in \mathcal{M}} c_\theta \varphi_\theta$, where $\mathcal{M} \subset \Theta(\mathcal{T})$ and $\#\mathcal{M} \leq n$. Let Λ be the set of all triangles $\Delta \in \mathcal{T}$ which are involved in all $E_\theta := \text{supp } \varphi_\theta$, $\theta \in \mathcal{M}$. Then $s = \sum_{\Delta \in \Lambda} s_\Delta$, where $s_\Delta =: \mathbf{1}_\Delta \cdot P_\Delta$, $P_\Delta \in \Pi_k$. Evidently, by (2.7), $\#\Lambda \leq c^*(N_0, \ell) \#\mathcal{M} \leq cn$.

We first extend Λ to $\tilde{\Lambda}$ as in Step 1 above and introduce some auxiliary sets of triangles needed for the forthcoming arguments. We set

$$\begin{aligned} \tilde{\Lambda}_m^* &:= \{\Delta \in \mathcal{T}_m : \Omega_\Delta^\ell \supset \Delta' \text{ for some } \Delta' \in \tilde{\Lambda} \cap \mathcal{T}_m\}, \\ \tilde{\Lambda}_m^{**} &:= \{\Delta \in \mathcal{T}_m : \Omega_\Delta^{2\ell} \supset \Delta' \text{ for some } \Delta' \in \tilde{\Lambda} \cap \mathcal{T}_m\}, \quad m \in \mathbb{Z}, \end{aligned}$$

and also

$$\tilde{\Lambda}^* := \bigcup_{m \in \mathbb{Z}} \tilde{\Lambda}_m^*, \quad \tilde{\Lambda}^{**} := \bigcup_{m \in \mathbb{Z}} \tilde{\Lambda}_m^{**}.$$

Note that $\Delta, \Delta' \in \mathcal{T}_m$ and $\Delta' \subset \Omega_\Delta^\ell$ imply $\Delta \subset \Omega_{\Delta'}^\ell$, and hence

$$\tilde{\Lambda}_m^* = \{\Delta \in \mathcal{T}_m : \Delta \subset \Omega_{\Delta'}^\ell \text{ for some } \Delta' \in \tilde{\Lambda} \cap \mathcal{T}_m\}.$$

Therefore, by (2.8), $\#\tilde{\Lambda}_m^* \leq c^{**}(N_0, \ell) \#(\tilde{\Lambda} \cap \mathcal{T}_m)$, and it follows that $\#\tilde{\Lambda}^* \leq cn$. Similarly, $\#\tilde{\Lambda}^{**} \leq c^{**}(N_0, 2\ell) \#(\tilde{\Lambda}) \leq cn$. It is clear that $\tilde{\Lambda} \subset \tilde{\Lambda}^* \subset \tilde{\Lambda}^{**}$.

We now proceed to estimate $|s|_{B_\tau^\alpha(\mathcal{T})}^\tau := \sum_{\Delta \in \mathcal{T}} |\Delta|^{-\alpha\tau} \mathbb{S}_\Delta(s)^\tau$. Let

$$s_m := \sum_{\mu \leq m} \sum_{\theta \in \mathcal{M} \cap \Theta_\mu} c_\theta \varphi_\theta, \quad m \in \mathbb{Z}.$$

Then $s_m \in \mathcal{S}_m$, and hence $\mathbb{S}_\Delta(s)_\tau = \mathbb{S}_\Delta(s - s_m)_\tau$ if $\Delta \in \mathcal{T}_m$. For each $\Delta \in \mathcal{T}$, we shall use one of the following two obvious bounds for $\mathbb{S}_\Delta(s)_\tau$:

$$(A.3) \quad \mathbb{S}_\Delta(s)_\tau \leq \|s\|_{L_\tau(\Omega_\Delta^\ell)},$$

$$(A.4) \quad \mathbb{S}_\Delta(s)_\tau \leq \|s - s_m\|_{L_\tau(\Omega_\Delta^\ell)}, \quad \Delta \in \mathcal{T}_m.$$

Namely, (A.3) will be applied to the triangles Δ in the set $\tilde{\Lambda}^* \subset \mathcal{T}$ defined above, while (A.4) will be used for all remaining triangles in \mathcal{T} .

For the next estimates, we shall consider separately the cases $0 < p < \infty$ and $p = \infty$.

Case 1. $0 < p < \infty$. We consider two possibilities for each $\Delta \in \mathcal{T}$: $\Delta \in \tilde{\Lambda}^*$ or $\Delta \in \mathcal{T} \setminus \tilde{\Lambda}^*$.

(a) If $\Delta \in \tilde{\Lambda}_m^*$, then for each $\Delta' \in \mathcal{T}_m$ such that $\Delta' \subset \Omega_\Delta^\ell$, we have $\Delta' \in \tilde{\Lambda}_m^{**}$ and, in view of (2.2), $|\Delta'| \leq c|\Delta|$. Hence, by (A.3),

$$\begin{aligned} \sum_{\Delta \in \tilde{\Lambda}_m^*} |\Delta|^{-\alpha\tau} \mathbb{S}_\Delta(s)_\tau^\tau &\leq \sum_{\Delta \in \tilde{\Lambda}_m^*} |\Delta|^{-\alpha\tau} \sum_{\Delta' \in \tilde{\Lambda}_m^{**}, \Delta' \subset \Omega_\Delta^\ell} \|s\|_{L_\tau(\Delta')}^\tau \\ &\leq c \sum_{\Delta \in \tilde{\Lambda}_m^*} \sum_{\Delta' \in \tilde{\Lambda}_m^{**}, \Delta' \subset \Omega_\Delta^\ell} |\Delta'|^{-\alpha\tau} \|s\|_{L_\tau(\Delta')}^\tau \\ &= c \sum_{\Delta' \in \tilde{\Lambda}_m^{**}} \sum_{\Delta \in \tilde{\Lambda}_m^*, \Omega_\Delta^\ell \supset \Delta'} |\Delta'|^{-\alpha\tau} \|s\|_{L_\tau(\Delta')}^\tau \\ &\leq c \sum_{\Delta' \in \tilde{\Lambda}_m^{**}} |\Delta'|^{-\alpha\tau} \|s\|_{L_\tau(\Delta')}^\tau, \end{aligned}$$

where we have switched the order of summation and taken into account the fact that $\#\{\Delta \in \tilde{\Lambda}_m^* : \Omega_\Delta^\ell \supset \Delta'\} = \#\{\Delta \in \tilde{\Lambda}_m^* : \Delta \subset \Omega_{\Delta'}^\ell\} \leq c^{**}(N_0, \ell)$, by (2.8). It follows that

$$(A.5) \quad \begin{aligned} \sum_{\Delta \in \tilde{\Lambda}^*} |\Delta|^{-\alpha\tau} \mathbb{S}_\Delta(s)_\tau^\tau &\leq c \sum_{\Delta \in \tilde{\Lambda}^{**}} |\Delta|^{-\alpha\tau} \|s\|_{L_\tau(\Delta)}^\tau \\ &\leq c(\#\tilde{\Lambda}^{**})^{\alpha\tau} \|s\|_p^\tau \leq cn^{\alpha\tau} \|s\|_p^\tau, \end{aligned}$$

where we applied Lemma A.2 to s with Λ replaced by $\tilde{\Lambda}^{**}$, which is obviously legitimate since $\tilde{\Lambda}^{**} \supset \Lambda$.

(b) Now suppose $\Delta \in \mathcal{T}_m \setminus \tilde{\Lambda}_m^*$. Then $\Omega_\Delta^\ell = \bigcup_{j=1}^{n_\Delta} \Delta_j$ for some $\Delta_j \in \mathcal{T}_m \setminus \tilde{\Lambda}$, $j = 1, \dots, n_\Delta$, with $n_\Delta \leq c^{**} \leq 3N_0^{2\ell-1}$ (see (2.8)). We have, using (A.4),

$$(A.6) \quad \mathbb{S}_\Delta(s)_\tau^\tau = \mathbb{S}_\Delta(s - s_m)_\tau^\tau \leq \sum_{j=1}^{n_\Delta} \|s - s_m\|_{L_\tau(\Delta_j)}^\tau.$$

If $\Delta_j \notin \Gamma$, then it has no descendants in Λ , and hence $s|_{\Delta_j} = s_m|_{\Delta_j}$, and

$$(A.7) \quad \|s - s_m\|_{L_\tau(\Delta_j)} = 0, \quad \Delta_j \notin \Gamma.$$

Suppose $\Delta_j \in \Gamma$; i.e., it is a chain triangle. Let $\tilde{\Delta}_j$ be the unique largest triangle of $\tilde{\Lambda}$ contained in Δ_j , and let $K_{\Delta_j} = \Delta_j \setminus \tilde{\Delta}_j$ and $\mu_{\Delta_j} = m - \tilde{m}$ be defined as in Step 1.

It is clear that in this case $s|_{K_{\Delta_j}} = s_m|_{K_{\Delta_j}} = \mathbb{1}_{K_{\Delta_j}} \cdot P_{\Delta_j}$ and $s_m|_{\Delta_j} = \mathbb{1}_{\Delta_j} \cdot P_{\Delta_j}$ for some $P_{\Delta_j} \in \Pi_k$. Therefore,

$$\|s - s_m\|_{L_\tau(\Delta_j)}^\tau = \|s - s_m\|_{L_\tau(\tilde{\Delta}_j)}^\tau \leq c\|s\|_{L_\tau(\tilde{\Delta}_j)}^\tau + c\|P_{\Delta_j}\|_{L_\tau(\tilde{\Delta}_j)}^\tau.$$

If $\Delta_j \in \Gamma \setminus \tilde{\Gamma}$, then clearly $s_m|_{\Delta_j} = 0$, and we have

$$(A.8) \quad \|s - s_m\|_{L_\tau(\Delta_j)} = \|s\|_{L_\tau(\tilde{\Delta}_j)}, \quad \Delta_j \in \Gamma \setminus \tilde{\Gamma}.$$

Assume that $\Delta_j \in \tilde{\Gamma}$. By Lemma 2.2,

$$\begin{aligned} \|P_{\Delta_j}\|_{L_\tau(\tilde{\Delta}_j)}^\tau &\leq |\tilde{\Delta}_j| \|P_{\Delta_j}\|_{L_\infty(\Delta_j)}^\tau \leq c|\tilde{\Delta}_j| \|P_{\Delta_j}\|_{L_\infty(K_{\Delta_j})}^\tau \\ &\leq c|\tilde{\Delta}_j| |K_{\Delta_j}|^{-\tau/p} \|P_{\Delta_j}\|_{L_p(K_{\Delta_j})}^\tau \leq c|\tilde{\Delta}_j| |\Delta_j|^{\alpha\tau-1} \|s\|_{L_p(K_{\Delta_j})}^\tau. \end{aligned}$$

By (2.1), $|\tilde{\Delta}_j| \leq \rho^{\mu_{\Delta_j}} |\Delta_j|$, and we arrive at the inequality

$$(A.9) \quad \|s - s_m\|_{L_\tau(\Delta_j)}^\tau \leq c\|s\|_{L_\tau(\tilde{\Delta}_j)}^\tau + c\rho^{\mu_{\Delta_j}} |\Delta_j|^{\alpha\tau} \|s\|_{L_p(K_{\Delta_j})}^\tau, \quad \Delta_j \in \tilde{\Gamma}.$$

From (A.6)–(A.9) and (2.2), we obtain

$$\begin{aligned} \sum_{\Delta \in \mathcal{T} \setminus \tilde{\Lambda}^*} |\Delta|^{-\alpha\tau} \mathbb{S}_\Delta(s)_\tau^\tau &= \sum_{m \in \mathbb{Z}} \sum_{\Delta \in \mathcal{T}_m \setminus \tilde{\Lambda}_m^*} |\Delta|^{-\alpha\tau} \mathbb{S}_\Delta(s)_\tau^\tau \\ &\leq c \sum_{\Delta \in \Gamma} |\Delta|^{-\alpha\tau} \|s\|_{L_\tau(\tilde{\Delta})}^\tau + c \sum_{\Delta \in \tilde{\Gamma}} \rho^{\mu_\Delta} \|s\|_{L_p(K_\Delta)}^\tau \\ &=: \Sigma_1 + \Sigma_2. \end{aligned}$$

Trivially,

$$\|s\|_{L_\tau(\tilde{\Delta})} \leq \sum_{\Delta' \in \tilde{\Lambda}, \Delta' \subset \Delta} \|s\|_{L_\tau(\Delta')}, \quad \Delta \in \Gamma.$$

Switching the order of summation, we find

$$\begin{aligned} \Sigma_1 &\leq c \sum_{\Delta' \in \tilde{\Lambda}} \|s\|_{L_\tau(\Delta')}^\tau \sum_{\Delta \in \Gamma, \Delta \supset \Delta'} |\Delta|^{-\alpha\tau} \\ (A.10) \quad &\leq c \sum_{\Delta' \in \tilde{\Lambda}} \|s\|_{L_\tau(\Delta')}^\tau |\Delta'|^{-\alpha\tau} \sum_{\Delta \in \Gamma, \Delta \supset \Delta'} (|\Delta'|/|\Delta|)^{\alpha\tau} \\ &\leq c \sum_{\Delta' \in \tilde{\Lambda}} |\Delta'|^{-\alpha\tau} \|s\|_{L_\tau(\Delta')}^\tau \leq c(\#\tilde{\Lambda})^{\alpha\tau} \|s\|_p^\tau, \end{aligned}$$

where we also used (2.9) and applied Lemma A.2 to s with Λ replaced by $\tilde{\Lambda}$.

To estimate Σ_2 we shall use the representation of $\tilde{\Gamma}$ as a disjoint union of chains: $\tilde{\Gamma} = \bigcup_{\lambda \in \mathcal{L}} \lambda$. Let $\lambda \in \mathcal{L}$ and suppose $\lambda = \{\Delta_1, \dots, \Delta_\nu\}$, where $\Delta''_\lambda \supset \Delta_1 \supset \dots \supset \Delta_\nu \supset \Delta'_\lambda$ with $\Delta'_\lambda, \Delta''_\lambda \in \tilde{\Lambda}$. Then $\mu_{\Delta_i} \geq \nu - i + 1$. Therefore,

$$\sum_{\Delta \in \lambda} \rho^{\mu_\Delta} \|s\|_{L_p(K_\Delta)}^\tau \leq \|s\|_{L_p(\Delta''_\lambda \setminus \Delta'_\lambda)}^\tau \sum_{j=1}^\nu \rho^{\nu-j+1} \leq c\|s\|_{L_p(K_\lambda)}^\tau,$$

where $K_\lambda := \Delta''_\lambda \setminus \Delta'_\lambda$. It is easy to see that the sets K_λ , $\lambda \in \mathcal{L}$, have pairwise disjoint interiors. Summing over all $\lambda \in \mathcal{L}$, we obtain by Hölder's inequality

$$\Sigma_2 \leq c \sum_{\lambda \in \mathcal{L}} \|s\|_{L_p(K_\lambda)}^\tau \leq c \left(\sum_{\lambda \in \mathcal{L}} \|s\|_{L_p(K_\lambda)}^p \right)^{\tau/p} (\#\mathcal{L})^{1-\tau/p} \leq c(\#\tilde{\Lambda})^{\alpha\tau} \|s\|_p^\tau.$$

From this estimate and (A.10), we find

$$\sum_{\Delta \in \mathcal{T} \setminus \tilde{\Lambda}^*} |\Delta|^{-\alpha\tau} \mathbb{S}_\Delta(s)_\tau^\tau \leq c(\#\tilde{\Lambda})^{\alpha\tau} \|s\|_p^\tau \leq cn^{\alpha\tau} \|s\|_p^\tau.$$

Combining this with (A.5) gives $\|s\|_{B_\tau^\alpha(\Phi)}^\tau \leq cn^{\alpha\tau} \|s\|_p^\tau$; i.e., (3.2) holds.

Case 2. $p = \infty$. The proof in this case is easier. We consider as before two possibilities for each $\Delta \in \mathcal{T}$: $\Delta \in \tilde{\Lambda}^*$ or $\Delta \in \mathcal{T} \setminus \tilde{\Lambda}^*$.

(a) For $\Delta \in \tilde{\Lambda}^*$, we obtain by (2.2)

$$|\Delta|^{-1} \mathbb{S}_\Delta(s)_\tau^\tau \leq |\Delta|^{-1} \|s\|_{L_\tau(\Omega_\Delta^\ell)}^\tau \leq |\Delta|^{-1} |\Omega_\Delta^\ell| \|s\|_\infty^\tau \leq c \|s\|_\infty^\tau.$$

Therefore,

$$(A.11) \quad \sum_{\Delta \in \tilde{\Lambda}^*} |\Delta|^{-1} \mathbb{S}_\Delta(s)_\tau^\tau \leq c \|s\|_\infty^\tau (\#\tilde{\Lambda}^*) \leq cn \|s\|_\infty^\tau.$$

(b) Let $\Delta \in \mathcal{T}_m \setminus \tilde{\Lambda}_m^*$. Then $\Omega_\Delta^\ell =: \bigcup_{j=1}^{n_\Delta} \Delta_j$ for some $\Delta_j \in \mathcal{T}_m \setminus \tilde{\Lambda}$, $j = 1, \dots, n_\Delta$, with $n_\Delta \leq c^{**} < 3N_0^{2\ell-1}$ (see (2.8)). We have (see (A.4))

$$\mathbb{S}_\Delta(s)_\tau^\tau = \mathbb{S}_\Delta(s - s_m)_\tau^\tau \leq \sum_{j=1}^{n_\Delta} \|s - s_m\|_{L_\tau(\Delta_j)}^\tau.$$

As in Case 1, if $\Delta_j \notin \Gamma$, then $\|s - s_m\|_{L_\tau(\Delta_j)} = 0$, and if $\Delta_j \in \Gamma$, then $s|_{K_{\Delta_j}} = s_m|_{K_{\Delta_j}} = \mathbb{1}_{K_{\Delta_j}} \cdot P_{\Delta_j}$ and $s_m|_{\Delta_j} = \mathbb{1}_{\Delta_j} \cdot P_{\Delta_j}$ for some $P_{\Delta_j} \in \Pi_k$. Therefore,

$$\begin{aligned} \|s - s_m\|_{L_\tau(\Delta_j)}^\tau &= \|s - s_m\|_{L_\tau(\tilde{\Delta}_j)}^\tau \\ &\leq c|\tilde{\Delta}_j|(\|s\|_\infty^\tau + \|P_{\Delta_j}\|_{L_\infty(\tilde{\Delta}_j)}^\tau) \leq c|\tilde{\Delta}_j| \|s\|_\infty^\tau, \end{aligned}$$

where we used the inequalities $\|P_{\Delta_j}\|_{L_\infty(\tilde{\Delta}_j)} \leq \|P_{\Delta_j}\|_{L_\infty(\Delta_j)} \leq c\|P_{\Delta_j}\|_{L_\infty(K_{\Delta_j})} \leq c\|s\|_\infty$ (see Lemma 2.2). From the above, we infer by (2.2)

$$|\Delta|^{-1} \mathbb{S}_\Delta(s)_\tau^\tau \leq c \|s\|_\infty^\tau \sum_{1 \leq j \leq n_\Delta, \Delta_j \in \Gamma \cap \mathcal{T}_m} |\tilde{\Delta}_j|/|\Delta_j|,$$

and hence, using (2.2) and the fact that each $\Delta' \in \Gamma \cap \mathcal{T}_m$ can belong to $\leq c^{**}$ sets Ω_Δ^ℓ , we obtain

$$\begin{aligned} \sum_{\Delta \in \mathcal{T}_m \setminus \tilde{\Lambda}_m^*} |\Delta|^{-1} \mathbb{S}_\Delta(s)_\tau^\tau &\leq c \|s\|_\infty^\tau \sum_{\Delta \in \Gamma \cap \mathcal{T}_m} |\tilde{\Delta}|/|\Delta| \\ &\leq c \|s\|_\infty^\tau \sum_{\Delta \in \Gamma \cap \mathcal{T}_m} \rho^{\mu\Delta}. \end{aligned}$$

Summing over $m \in \mathbb{Z}$, we find

$$\sum_{\Delta \in \mathcal{T} \setminus \tilde{\Lambda}^*} |\Delta|^{-1} \mathbb{S}_\Delta(s)_\tau^\tau \leq c \|s\|_\infty^\tau \sum_{\Delta \in \Gamma} \rho^{\mu_\Delta} \leq c \|s\|_\infty^\tau (\#\mathcal{L} + \#\mathcal{L}^\infty) \leq cn \|s\|_\infty^\tau.$$

We couple this with (A.11) to obtain $\|s\|_{B_\tau^\alpha(\mathcal{T})}^\tau \leq cn \|s\|_\infty^\tau$, which is (3.2). \square

REFERENCES

- [1] P. ALFELD, B. PIPER, AND L. L. SCHUMAKER, *Minimally supported bases for spaces of bivariate piecewise polynomials of smoothness r and degree $d \geq 4r+1$* , *Comput. Aided Geom. Design*, 4 (1987), pp. 105–123.
- [2] I. BABUŠKA AND A. K. AZIZ, *On the angle condition in the finite element method*, *SIAM J. Numer. Anal.*, 13 (1976), pp. 214–226.
- [3] J. BERGH AND J. LÖFSTRÖM, *Interpolation Spaces. An Introduction*, Grundlehren Math. Wiss. 223, Springer-Verlag, Berlin, New York, 1976.
- [4] C. BENNETT AND R. SHARPLEY, *Interpolation of Operators*, Pure Appl. Math. 129, Academic Press, Boston, 1988.
- [5] P. BINEV AND R. DEVORE, *Fast Computation in Adaptive Tree Approximation*, <http://www.math.sc.edu/~imip/02.html>.
- [6] P. BINEV, W. DAHMEN, AND R. DEVORE, *Adaptive Finite Element Methods with Convergence Rates*, IGPM Report 219, RWTH Aachen, Aachen, Germany, 2002, <http://elc2.igpm.rwth-aachen.de/~dahmen/>.
- [7] P. BINEV, W. DAHMEN, R. DEVORE, AND P. PETRUSHEV, *Approximation classes for adaptive methods*, *Serdica Math. J.*, 28 (2002), pp. 391–416.
- [8] C. DE BOOR AND K. HÖLLIG, *Approximation power of smooth bivariate pp functions*, *Math. Z.*, 197 (1988), pp. 343–363.
- [9] C. DE BOOR, K. HÖLLIG, AND S. D. RIEMENSCHNEIDER, *Box Splines*, Springer-Verlag, New York, 1993.
- [10] C. DE BOOR AND R. Q. JIA, *A sharp upper bound on the approximation order of smooth bivariate pp functions*, *J. Approx. Theory*, 72 (1993), pp. 24–33.
- [11] J. BRAMBLE AND X. ZHANG, *Multigrid methods for the biharmonic problem discretized by conforming C^1 finite elements on nonnested meshes*, *Numer. Funct. Anal. Optim.*, 1 (1995), pp. 835–846.
- [12] S. BRENNER AND L. R. SCOTT, *The Mathematical Theory of Finite Element Methods*, Texts Appl. Math. 15, Springer-Verlag, New York, 1994.
- [13] M. D. BUHMANN, O. DAVYDOV, AND T. N. T. GOODMAN, *Cubic spline prewavelets on the four-directional mesh*, *Found. Comp. Math.*, 3 (2003), pp. 113–133.
- [14] P. G. CIARLET, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.
- [15] C. K. CHUI, *Multivariate Splines*, CBMS-NSF Regional Conf. Ser. in Appl. Math. 54, SIAM, Philadelphia, 1988.
- [16] C. K. CHUI, D. HONG, AND R.-Q. JIA, *Stability of optimal order approximation by bivariate splines over arbitrary triangulations*, *Trans. Amer. Math. Soc.*, 347 (1995), pp. 3301–3318.
- [17] C. K. CHUI AND M.-J. LAI, *Multivariate vertex splines and finite elements*, *J. Approx. Theory*, 60 (1990), pp. 245–343.
- [18] C. K. CHUI AND M.-J. LAI, *On bivariate super vertex splines*, *Constr. Approx.*, 6 (1990), pp. 399–419.
- [19] W. DAHMEN AND C. A. MICHELLI, *On the local linear independence of translates of a box spline*, *Studia Math.*, 82 (1985), pp. 243–262.
- [20] W. DAHMEN, P. OSWALD, AND X.-Q. SHI, *C^1 -hierarchical bases*, *J. Comput. Appl. Math.*, 51 (1994), pp. 37–56.
- [21] W. DAHMEN AND R. STEVENSON, *Element-by-element construction of wavelets satisfying stability and moment conditions*, *SIAM J. Numer. Anal.*, 37 (1999), pp. 319–352.
- [22] O. DAVYDOV, *Stable local bases for multivariate spline spaces*, *J. Approx. Theory*, 111 (2001), pp. 267–297.
- [23] O. DAVYDOV, *On the computation of stable local bases for bivariate polynomial splines*, in *Trends in Approximation Theory*, K. Kopotun, T. Lyche, and M. Neamtu, eds., Vanderbilt University Press, Nashville, TN, 2001, pp. 85–94.

- [24] O. DAVYDOV, *Locally stable spline bases on nested triangulations*, in Approximation Theory X: Wavelets, Splines, and Applications, C. K. Chui, L. L. Schumaker, and J. Stöckler, eds., Vanderbilt University Press, Nashville, TN, 2002, pp. 231–240.
- [25] O. DAVYDOV, G. NÜRNBERGER, AND F. ZEILFELDER, *Bivariate spline interpolation with optimal approximation order*, Constr. Approx., 17 (2001), pp. 181–208.
- [26] O. DAVYDOV AND L. L. SCHUMAKER, *Stable local nodal bases for C^1 bivariate polynomial splines*, in Curve and Surface Fitting: Saint-Malo 1999, A. Cohen, C. Rabut, and L. L. Schumaker, eds., Vanderbilt University Press, Nashville, TN, 2000, pp. 171–180.
- [27] O. DAVYDOV AND L. L. SCHUMAKER, *On stable local bases for bivariate polynomial spline spaces*, Constr. Approx., 18 (2002), pp. 87–116.
- [28] O. DAVYDOV AND F. ZEILFELDER, *Scattered data fitting by direct extension of local polynomials with bivariate splines*, Adv. Comput. Math., to appear.
- [29] R. DEVORE, B. JAWERTH, AND V. POPOV, *Compression of wavelet decompositions*, Amer. J. Math., 114 (1992), pp. 737–785.
- [30] R. DEVORE, P. PETRUSHEV, AND X. YU, *Nonlinear wavelet approximation in the space $C(\mathbb{R}^d)$* , in Progress in Approximation Theory, A. A. Gonchar and E. B. Saff, eds., Springer-Verlag, New York, 1992, pp. 261–283.
- [31] R. DEVORE AND V. POPOV, *Interpolation spaces and nonlinear approximation*, in Function Spaces and Applications (Lund, 1986), Lecture Notes in Math. 1302, Springer-Verlag, Berlin, 1988, pp. 191–205.
- [32] R. DEVORE AND V. POPOV, *Interpolation of Besov spaces*, Trans. Amer. Math. Soc., 305 (1988), pp. 297–314.
- [33] M. DÆHLEN, T. LYCHE, K. MØRKEN, R. SCHNEIDER, AND H.-P. SEIDEL, *Multiresolution analysis over triangles, based on quadratic Hermite interpolation*, J. Comput. Appl. Math., 119 (2000), pp. 97–114.
- [34] M. S. FLOATER AND E. G. QUAK, *Linear independence and stability of piecewise linear pre-wavelets on arbitrary triangulations*, SIAM J. Numer. Anal., 38 (2000), pp. 58–79.
- [35] D. HONG, *Spaces of bivariate spline functions over triangulation*, Approx. Theory Appl., 7 (1991), pp. 56–75.
- [36] A. KH. IBRAHIM AND L. L. SCHUMAKER, *Super spline spaces of smoothness r and degree $d \geq 3r + 2$* , Constr. Approx., 7 (1991), pp. 401–423.
- [37] W. JIANZHONG, *On dual basis of bivariate box spline*, Approx. Theory Appl., 3 (1987), pp. 153–163.
- [38] B. KARAIVANOV AND P. PETRUSHEV, *Nonlinear piecewise polynomial approximation beyond Besov spaces*, Appl. Comput. Harmon. Anal., to appear.
- [39] B. KARAIVANOV, P. PETRUSHEV, AND R. SHARPLEY, *Algorithms for nonlinear piecewise polynomial approximation: Theoretical aspects*, Trans. Amer. Math. Soc., 355 (2003), pp. 2585–2631.
- [40] M. LAGHCHIM-LAHLLOU AND P. SABLONNIÈRE, *Quadrilateral finite elements of FVS type and class C^p* , Numer. Math., 70 (1995), pp. 229–243.
- [41] M.-J. LAI, *A remark on translates of a box spline*, Approx. Theory Appl., 5 (1989), pp. 97–104.
- [42] M.-J. LAI, *Approximation order from bivariate C^1 -cubics on a four-directional mesh is full*, Comput. Aided Geom. Design, 11 (1994), pp. 215–223.
- [43] M.-J. LAI, *Bivariate spline spaces on FVS-triangulations*, in Approximation Theory VIII, Vol. 1: Approximation and Interpolation, C. K. Chui and L. L. Schumaker, eds., World Scientific, Singapore, 1995, pp. 309–316.
- [44] M.-J. LAI AND L. L. SCHUMAKER, *On the approximation power of bivariate splines*, Adv. Comput. Math., 9 (1998), pp. 251–279.
- [45] M.-J. LAI AND L. L. SCHUMAKER, *On the approximation power of splines on triangulated quadrangulations*, SIAM J. Numer. Anal., 36 (1999), pp. 143–159.
- [46] M.-J. LAI AND L. L. SCHUMAKER, *Quadrilateral macroelements*, SIAM J. Math. Anal., 33 (2002), pp. 1107–1116.
- [47] T. LYCHE, K. MØRKEN, AND E. QUAK, *Theory and algorithms for non-uniform spline wavelets*, in Multivariate Approximation and Applications, N. Dyn, D. Leviatan, D. Levin, and A. Pinkus, eds., Cambridge University Press, Cambridge, UK, 2001, pp. 152–187.
- [48] J. MORGAN AND R. SCOTT, *A nodal basis for C^1 piecewise polynomials of degree $n \geq 5$* , Math. Comp., 29 (1975), pp. 736–740.
- [49] G. NÜRNBERGER, L. L. SCHUMAKER, AND F. ZEILFELDER, *Local Lagrange interpolation by bivariate C^1 cubic splines*, in Mathematical Methods in CAGD: Oslo 2000, T. Lyche and L. L. Schumaker, eds., Vanderbilt University Press, Nashville, TN, 2001, pp. 393–404.
- [50] P. OSWALD, *Hierarchical conforming finite element methods for the biharmonic equation*, SIAM J. Numer. Anal., 29 (1992), pp. 1610–1625.

- [51] P. OSWALD, *Multilevel Finite Element Approximation*, Teubner, Stuttgart, 1994.
- [52] J. PEETRE AND G. SPARR, *Interpolation of normed Abelian groups*, Ann. Mat. Pura Appl. (4), 92 (1972), pp. 217–262.
- [53] P. PETRUSHEV, *Direct and converse theorems for rational and spline approximation and Besov spaces*, in Function Spaces and Applications (Lund, 1986), Lecture Notes in Math. 1302, Springer-Verlag, Berlin, 1988, pp. 363–377.
- [54] P. PETRUSHEV, *Multivariate n -term rational and piecewise polynomial approximation*, J. Approx. Theory, 121 (2003), pp. 158–197.
- [55] P. PETRUSHEV AND V. POPOV, *Rational Approximation of Real Functions*, Cambridge University Press, Cambridge, UK, 1987.
- [56] P. SABLONNIÈRE, *Quasi-interpolants associated with H -splines on a three direction mesh*, J. Comput. Appl. Math., 66 (1996), pp. 433–442.
- [57] L. L. SCHUMAKER, *On super splines and finite elements*, SIAM J. Numer. Anal., 26 (1989), pp. 997–1005.
- [58] R. STEVENSON, *Piecewise linear (pre-)wavelets on non-uniform meshes*, in Multigrid Methods, V. W. Hackbusch and G. Wittum, eds., Springer-Verlag, Berlin, 1998, pp. 306–319.