



Nonlinear Nonnested Spline Approximation

Martin Lind¹ · Pencho Petrushev²

Received: 16 June 2015 / Revised: 22 April 2016 / Accepted: 26 August 2016 /
Published online: 28 December 2016
© Springer Science+Business Media New York 2016

Abstract Nonlinear approximation from regular piecewise polynomials (splines) supported on rings in \mathbb{R}^2 is studied. By definition, a ring is a set in \mathbb{R}^2 obtained by subtracting a compact convex set with polygonal boundary from another such a set, but without creating uncontrollably narrow elongated subregions. Nested structure of the rings is not assumed; however, uniform boundedness of the eccentricities of the underlying convex sets is required. It is also assumed that the splines have maximum smoothness. Bernstein type inequalities for this sort of splines are proved that allow us to establish sharp inverse estimates in terms of Besov spaces.

Keywords Spline approximation · Multivariate approximation · Nonlinear approximation · Besov spaces · Jackson estimate · Bernstein estimate

Mathematics Subject Classification Primary 41A15; Secondary 41A17 · 41A25 · 41A27

1 Introduction

Nonlinear approximation from piecewise polynomials (splines) in dimensions $d > 1$ is important from theoretical and practical points of view. We are interested in charac-

Communicated by Ronald A. DeVore.

✉ Pencho Petrushev
pencho@math.sc.edu

Martin Lind
martin.lind@kau.se

¹ Department of Mathematics, Karlstad University, 651 88 Karlstad, Sweden

² Department of Mathematics, University of South Carolina, Columbia, SC 29208, USA

terizing the rates of nonlinear spline approximation in L^p . While this theory is simple and well understood in the univariate case, it is underdeveloped and challenging in dimensions $d > 1$.

In this article, we focus on nonlinear approximation in L^p , $0 < p < \infty$, from regular piecewise polynomials in \mathbb{R}^2 or on compact subsets of \mathbb{R}^2 with polygonal boundaries. Our goal is to obtain complete characterization of the rates of approximation (the associated approximation spaces). To describe our results, we begin by introducing in more detail our

Setting and approximation tool. We are interested in approximation in the space L^p , $0 < p < \infty$, from the class of regular piecewise polynomials $\mathcal{S}(n, k)$ of degree $k - 1$ with $k \geq 1$ of maximum smoothness over n rings. More specifically, with Ω being a compact polygonal domain in \mathbb{R}^2 or $\Omega = \mathbb{R}^2$, we denote by $\mathcal{S}(n, k)$ the set of all piecewise polynomials S of the form

$$S = \sum_{j=1}^n P_j \cdot \mathbb{1}_{R_j}, \quad S \in C^{k-2}(\Omega), \quad P_j \in \Pi_{k-1}, \tag{1.1}$$

where R_1, \dots, R_n are rings with disjoint interiors. Here Π_{k-1} denotes the set of all algebraic polynomials of (total) degree $\leq k - 1$ in two variables, and as usual $S \in C^{k-2}(\Omega)$ means that all partial derivatives $\partial^\alpha S \in C(\Omega)$, $|\alpha| \leq k - 2$. The elements of $\mathcal{S}(n, 1)$ are simply piecewise constants.

A set $R \subset \mathbb{R}^2$ is called a ring if R is a compact convex set with polygonal boundary or the difference of two such sets. All convex sets we consider are with uniformly bounded eccentricity, and we do not allow uncontrollably narrow elongated subregions. For the precise definitions, see Sects. 3.1 and 4.1.

It is important to point out that although regular, our tool for approximation is highly nonlinear. In particular, the rings in (1.1) may vary with S , and we do not assume any nested structure of the rings involved in the definition of different splines S in (1.1). Consequently, if $S_1, S_2 \in \mathcal{S}(n, k)$, then in general $S_1 \pm S_2 \notin \mathcal{S}(N, k)$ for any N . The case of approximation from splines over nested (anisotropic) rings induced by hierarchical nested triangulations is developed in [3, 6].

Given a function $f \in L^p(\Omega)$, we denote by $S_n^k(f)_p$ the best L^p -approximation of f from $\mathcal{S}(n, k)$. Our goal is to completely characterize the approximation spaces A_q^α , $\alpha > 0, 0 < q \leq \infty$, defined by the (quasi)norm

$$\|f\|_{A_q^\alpha} := \|f\|_{L^p} + \left(\sum_{n=1}^\infty \left(n^\alpha S_n^k(f)_p \right)^q \frac{1}{n} \right)^{1/q},$$

with the ℓ^q -norm replaced by the sup-norm if $q = \infty$. To this end, we utilize the standard machinery of Jackson and Bernstein estimates. The Besov spaces $B_\tau^{s,k} := B_{\tau\tau}^{s,k}$ with $1/\tau = s/2 + 1/p$ naturally appear in our regular setting, see (2.1). The Jackson estimate takes the form: For any $f \in B_\tau^{s,k}$,

$$S_n^k(f)_p \leq cn^{-s/2} |f|_{B_\tau^{s,k}}. \tag{1.2}$$

For $k = 1, 2$, this estimate follows readily from the results in [6]. It is an open problem to establish it for $k > 2$. Estimate (1.2) implies the direct estimate

$$S_n^k(f)_p \leq cK\left(f, n^{-s/2}\right), \tag{1.3}$$

where $K(f, t) = K(f, t; L^p, B_\tau^{s,k})$ is the K -functional induced by L^p and $B_\tau^{s,k}$, see (3.6). Note that estimate (1.2), for any $k > 2$, is well known and easy to prove when approximating from discontinuous piecewise polynomials over rings. For example, it follows from Theorems 2.25 and 3.10 in [6]. For smoother splines (but not splines of maximal smoothness), (1.2) follows by Theorems 2.15 and 3.1 in [3].

It is a major problem to establish a companion matching inverse estimate. The following Bernstein estimate would imply such an estimate:

$$|S_1 - S_2|_{B_\tau^{s,k}} \leq cn^{s/2} \|S_1 - S_2\|_{L^p}, \quad S_1, S_2 \in \mathcal{S}(n, k). \tag{1.4}$$

However, as is easy to show, this estimate is not valid. The problem is that $S_1 - S_2$ may have one or more uncontrollably elongated parts such as $\mathbb{1}_{[0,\varepsilon] \times [0,1]}$ with small ε , which create problems for the Besov norm, see Example 3.3 below.

The main idea of this article is to replace (1.4) by the Bernstein type estimate

$$|S_1|_{B_\tau^{s,k}}^\lambda \leq |S_2|_{B_\tau^{s,k}}^\lambda + cn^{\lambda s/2} \|S_1 - S_2\|_{L^p}^\lambda, \quad \lambda := \min\{\tau, 1\}, \tag{1.5}$$

where $0 < s/2 < k - 1 + 1/p$. This estimate leads to the needed inverse estimate

$$K(f, n^{-s/2}) \leq cn^{-s/2} \left(\sum_{v=1}^n \frac{1}{v} \left[v^{s/2} S_v(f)_p \right]^\lambda + \|f\|_{L^p}^\lambda \right)^{1/\lambda}. \tag{1.6}$$

In turn, this estimate and (1.3) yield a characterization of the associated approximation spaces A_q^α in terms of real interpolation spaces

$$A_q^\alpha = \left(L^p, B_\tau^{s,k} \right)_{\frac{\alpha}{s}, q}, \quad 0 < \alpha < s, \quad 0 < q \leq \infty. \tag{1.7}$$

See, e.g., [4, 8].

A natural restriction on the Bernstein estimate (1.5) is the requirement that the splines $S_1, S_2 \in \mathcal{S}(n, k)$ have maximum smoothness. For instance, if we consider approximation from piecewise linear functions S ($k = 2$), it is assumed that S is continuous. As will be shown in Example 4.4, estimate (1.5) is no longer valid for discontinuous piecewise linear functions.

Motivation. Our setting would simplify considerably if the rings R_j in (1.1) are replaced by regular convex sets with polygonal boundaries or simply triangles. For example, in many cases, people do adaptive approximation from piecewise linear functions by local refinements resulting in nested triangulations. However, this would restrict the approximation power of our approximation tool. The isotropic refinement

schemes can give rate $O(n^{-s/2})$ of L^p -approximation for functions in the Besov space $B_\tau^{s,k}$ with $1/\tau < s/2 + 1/p$, which is off the Sobolev embedding line. For more details, see [1] ([5] is also relevant). In contrast, the piecewise polynomials over rings as defined above allow one to obtain the Jackson estimate (1.2), where $1/\tau = s/2 + 1/p$; i.e., in this case the Besov space $B_\tau^{s,k}$ is just on the Sobolev embedding line. This leads to a complete characterization of the associated approximation spaces. The idea of using rings has already been utilized in [2]. In concluding, there are two main points that we would like to make:

(i) In order to achieve the full strength of nonlinear spline approximation in dimension $d = 2$, the underlying partition should be in rings or a partition compatible with rings.

(ii) In nonlinear approximation from regular splines in L^p , $p < \infty$, in dimension $d = 2$ the rates of approximation are not sensitive to the particular underlying partitions as long as these are in rings. For example, in regular piecewise linear approximation asymptotically one cannot do better than if approximating by using a particular hierarchy of Courant hat functions over regular nested triangulations.

The proof of the Bernstein estimate (1.5) is quite involved. To make it more understandable, we first prove it in Sect. 3 in the easier case of piecewise constants and then in Sect. 4 for smoother splines. Our method is not limited to splines in dimension $d = 2$. However, there is a great deal of geometric arguments involved in these proofs, and to avoid more complicated considerations, we focus only on spline approximation in dimension $d = 2$ here.

Useful notation. Throughout this article, we shall use $|G|$ to denote the Lebesgue measure of a set $G \subset \mathbb{R}^2$, G° , \overline{G} , and ∂G will denote the interior, closure, and boundary of G , $d(G) := \sup_{x,y \in G} |x - y|$ will stand for the diameter of G , and $\mathbb{1}_G$ will denote the characteristic function of G ; as usual, $|x|$ will stand for the Euclidean norm of $x \in \mathbb{R}^2$. If G is finite, then $\#G$ will stand for the number of elements of G . If γ is a polygonal curve in \mathbb{R}^2 , then $\ell(\gamma)$ will denote its length. Positive constants will be denoted by c_1, c_2, c', \dots , and they may vary at every occurrence. Some important constants will be denoted by c_0, N_0, β, \dots , and will remain unchanged throughout. The notation $a \sim b$ will stand for $c_1 \leq a/b \leq c_2$.

2 Background

2.1 Besov Spaces

Besov spaces appear naturally in nonlinear spline approximation. For spline approximation in $L^p(\Omega)$, $0 < p < \infty$, we will utilize the Besov spaces $B_\tau^{s,k} = B_{\tau\tau}^{s,k}$, where $s > 0, k \geq 1, 1/\tau := s/2 + 1/p$. The space $B_\tau^{s,k}$ is defined as the set of all functions $f \in L^\tau(\Omega)$ such that

$$|f|_{B_\tau^{s,k}} := \left(\int_0^\infty [t^{-s} \omega_k(f, t)_\tau]^\tau \frac{dt}{t} \right)^{1/\tau} < \infty. \tag{2.1}$$

Here $\omega_k(f, t)_\tau := \sup_{|h| \leq t} \|\Delta_h^k f(\cdot)\|_{L^\tau(\Omega)}$, with

$$\Delta_h^k f(x) := \sum_{v=0}^k (-1)^{k+v} \binom{k}{v} f(x + vh)$$

if the segment $[x, x + kh] \subset \Omega$ and $\Delta_h^k f(x) := 0$ otherwise.

Observe that for the standard Besov spaces B_{pq}^s with $s > 0$ and $1 \leq p, q \leq \infty$, the norm is independent of the index $k > s$. However, for the above Besov spaces in general $\tau < 1$, which changes the nature of the Besov space and k should no longer be directly connected to s . For more details, see the discussion in [6, pp. 202–203].

2.2 Nonlinear Spline Approximation in Dimension $d = 1$

For comparison, here we provide a brief account of nonlinear spline approximation in the univariate case. Denote by $\tilde{S}_n^k(f)_p$ the best L^p -approximation of $f \in L^p(\mathbb{R})$ from the set $\tilde{S}(n, k)$ of all piecewise polynomials S of degree $\leq k - 1$ with $n + 1$ free knots. Thus, $S \in \tilde{S}(n, k)$ if $S = \sum_{j=1}^n P_j \mathbb{1}_{I_j}$, where $P_j \in \Pi_{k-1}$ and $I_j, j = 1, \dots, n$, are arbitrary compact intervals with disjoint interiors and $\cup_j I_j$ is an interval. No smoothness of S is required.

Let $s > 0, 0 < p < \infty$, and $1/\tau = s + 1/p$. The following Jackson and Bernstein estimates hold (see [7]): If $f \in L^p(\mathbb{R})$ and $n \geq 1$, then

$$\tilde{S}_n^k(f)_p \leq cn^{-s} |f|_{B_\tau^{s,k}} \tag{2.2}$$

and

$$|S|_{B_\tau^{s,k}} \leq cn^s \|S\|_{L^p}, \quad S \in \tilde{S}(n, k), \tag{2.3}$$

where $c > 0$ is a constant depending only on s and p . These estimates imply direct and inverse estimates that allow the complete characterization of the respective approximation spaces. For more details, see [7] or [4, 8].

Several remarks are in order. (1) Above no smoothness is imposed on the piecewise polynomials from $\tilde{S}(n, k)$. The point is that the rates of approximation from smooth splines are the same as for nonsmooth splines. A key observation is that in dimension $d = 1$, the discontinuous piecewise polynomials are infinitely smooth with respect to the Besov spaces $B_\tau^{s,k}$. This is not the case in dimensions $d > 1$, where smoothness matters. (2) Unlike in the multivariate case, estimates (2.2–2.3) hold for every $s > 0$. (3) If $S_1, S_2 \in \tilde{S}(n, k)$, then $S_1 - S_2 \in \tilde{S}(2n, k)$, and hence (2.3) is sufficient for establishing the respective inverse estimate. This is not true in the multivariate case, and one needs estimates like (1.4) (if valid) or (1.5) (in our case). (4) There is a great deal of geometry involved in multivariate spline approximation, while in dimension $d = 1$ there is none.

2.3 Nonlinear Nested Spline Approximation in Dimension $d = 2$

The rates of approximation in $L^p, 0 < p < \infty$, from splines generated by multilevel anisotropic nested triangulations in \mathbb{R}^2 are studied in [3, 6]. The respective approximation spaces are completely characterized in terms of Besov type spaces (B-spaces)

defined via local piecewise polynomial approximation. The setting in [3, 6] allows one to deal with piecewise polynomials over triangulations with arbitrarily sharp angles. However, the nested structure of the underlying triangulations is quite restrictive. In this article, we consider nonlinear approximation from nonnested splines, but in a regular setting. This is a setting that frequently appears in applications.

3 Nonlinear Approximation from Piecewise Constants

3.1 Setting

Here we describe all components of our setting, including the region Ω where the approximation will take place and the tool for approximation we consider.

The region Ω . We shall consider two scenarios for Ω : (a) Ω is a compact polygonal domain in \mathbb{R}^2 , or (b) $\Omega = \mathbb{R}^2$. More explicitly, in the first case, we assume that Ω can be represented as the union of finitely many rings in the sense of Definition 3.1 with disjoint interiors. Therefore, the boundary $\partial\Omega$ of Ω is the union of finitely many polygonal curves consisting of finitely many segments (edges).

The approximation tool. To describe our tool for approximation, we first introduce *rings* in \mathbb{R}^2 .

Definition 3.1 We say that $R \subset \mathbb{R}^2$ is a ring if R can be represented in the form $R = Q_1 \setminus Q_2$, where Q_1, Q_2 satisfy the following conditions:

- (a) $Q_2 \subset Q_1$ or $Q_2 = \emptyset$;
- (b) Each of Q_1 and Q_2 is a compact regular convex set in \mathbb{R}^2 whose boundary is a polygonal curve consisting of no more than N_0 ($N_0 \geq 3$ is fixed) line segments. Here a compact convex set $Q \subset \mathbb{R}^2$ is deemed *regular* if Q has a *bounded eccentricity*; that is, there exist balls $B_1, B_2, B_j = B(x_j, r_j)$, such that $B_2 \subset Q \subset B_1$ and $r_1 \leq c_0 r_2$, where $c_0 > 0$ is a universal constant.
- (c) R contains no uncontrollably narrow and elongated subregions, which is specified as follows: Each edge (segment) E of the boundary of R can be subdivided into the union of at most two segments E_1, E_2 ($E = E_1 \cup E_2$) with disjoint (one dimensional) interiors such that there exist triangles Δ_1 with a side E_1 and adjacent to E_1 angles of magnitude β , and Δ_2 with a side E_2 and adjacent to E_2 angles of magnitude β such that $\Delta_j \subset R, j = 1, 2$, where $0 < \beta \leq \pi/3$ is a fixed constant.

Figure 1 illustrates the above definition of rings.

Remark Observe that from the above definition, it readily follows that for any ring R in \mathbb{R}^2 ,

$$|R| \sim d(R)^2, \quad (3.1)$$

with constants of equivalence depending only on the parameters N_0, c_0 , and β .

Condition 3.2 *In the case, when Ω is a compact polygonal domain in \mathbb{R}^2 , we assume that there exists a constant $n_0 \geq 1$ such that Ω can be represented as the union of n_0 rings R_j with disjoint interiors: $\Omega = \cup_{j=1}^{n_0} R_j$. If $\Omega = \mathbb{R}^2$, then we set $n_0 := 1$.*

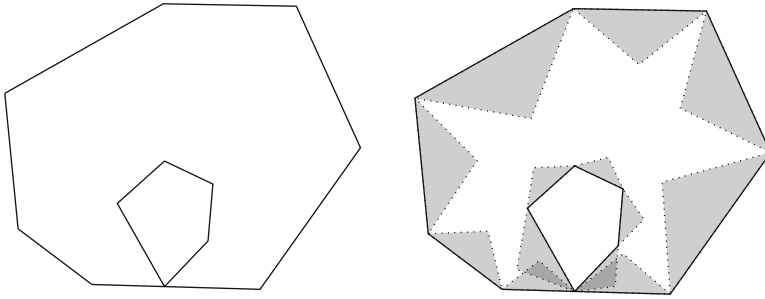


Fig. 1 Left a ring $R = Q_1 \setminus Q_2$. Right R with the triangles associated to the segments of ∂R

We now can introduce the class of regular piecewise constants.

Case 1: Ω is a compact polygonal domain in \mathbb{R}^2 . We denote by $\mathcal{S}(n, 1)$ ($n \geq n_0$) the set of all piecewise constants S of the form

$$S = \sum_{j=1}^n c_j \mathbb{1}_{R_j}, \quad c_j \in \mathbb{R}, \tag{3.2}$$

where R_1, \dots, R_n are rings with disjoint interiors such that $\Omega = \cup_{j=1}^n R_j$.

Case 2: $\Omega = \mathbb{R}^2$. In this case, we denote by $\mathcal{S}(n, 1)$ the set of all piecewise constant functions S of the form (3.2), where R_1, \dots, R_n are rings with disjoint interiors such that the support $R := \cup_{j=1}^n R_j$ of S is a ring in the sense of Definition 3.1.

A simple case of the above setting is when $\Omega = [0, 1]^2$ and the rings R are of the form $R = Q_1 \setminus Q_2$, where Q_1, Q_2 are dyadic squares in \mathbb{R}^2 . These kind of dyadic rings have been used in [2].

A bit more general is the setting when Ω is a regular rectangle in \mathbb{R}^2 with sides parallel to the coordinate axes or $\Omega = \mathbb{R}^2$ and the rings R are of the form $R = Q_1 \setminus Q_2$, where Q_1, Q_2 are regular rectangles with sides parallel to the coordinate axes, and no narrow and elongated subregions are allowed in the sense of Definition 3.1 (c).

Clearly, the set $\mathcal{S}(n, 1)$ is nonlinear since the rings $\{R_j\}$ and the constants $\{c_j\}$ in (3.2) may vary with S .

We denote by $S_n^1(f)_p$ the best approximation of $f \in L^p(\Omega)$ from $\mathcal{S}(n, 1)$ in $L^p(\Omega)$, $0 < p < \infty$; i.e.,

$$S_n^1(f)_p := \inf_{S \in \mathcal{S}(n,1)} \|f - S\|_{L^p}. \tag{3.3}$$

Besov spaces. When approximating in L^p , $0 < p < \infty$, from piecewise constants the Besov spaces $B_\tau^{s,1}$ with $1/\tau = s/2 + 1/p$ naturally appear. In this section, we shall use the abbreviated notation $B_\tau^s := B_\tau^{s,1}$.

3.2 Direct and Inverse Estimates

The following *Jackson estimate* is quite easy to establish (see [6]): If $f \in B_\tau^s$, $s > 0$, $1/\tau := s/2 + 1/p$, $0 < p < \infty$, then $f \in L^p(\Omega)$ and

$$S_n^1(f)_p \leq cn^{-s/2} |f|_{B_\tau^s} \text{ for } n \geq n_0, \tag{3.4}$$

where $c > 0$ is a constant depending only on s, p and the structural constants $N_0, c_0,$ and β of the setting.

This estimate leads immediately to the following direct estimate: If $f \in L^p(\Omega),$ then

$$S_n^1(f)_p \leq cK(f, n^{-s/2}), \quad n \geq n_0, \tag{3.5}$$

where $K(f, t)$ is the K -functional induced by L^p and $B_\tau^s;$ namely,

$$K(f, t) = K(f, t; L^p, B_\tau^s) := \inf_{g \in B_\tau^s} \{ \|f - g\|_p + t \|g\|_{B_\tau^s} \}, \quad t > 0. \tag{3.6}$$

The main problem here is to prove a matching inverse estimate. Observe that the following Bernstein estimate holds: If $S \in \mathcal{S}(n, 1), n \geq n_0,$ and $0 < p < \infty,$ $0 < s < 2/p, 1/\tau = s/2 + 1/p,$ then

$$|S|_{B_\tau^s} \leq cn^{s/2} \|S\|_{L^p}, \tag{3.7}$$

where the constant $c > 0$ depends only on $s, p,$ and the structural constants of the setting (see the proof of Theorem 4.5). The point is that this estimate does not imply a companion to (3.5) inverse estimate. The following estimate would imply such an estimate:

$$|S_1 - S_2|_{B_\tau^s} \leq cn^{s/2} \|S_1 - S_2\|_{L^p}, \quad S_1, S_2 \in \mathcal{S}(n, 1). \tag{3.8}$$

However, as the following example shows this estimate is in general not valid.

Example 3.3 Consider the function $f := \mathbb{1}_{[0,\varepsilon] \times [0,1]},$ where $\varepsilon > 0$ is sufficiently small. It is easy to see that

$$\omega_1(f, t)_\tau \sim \begin{cases} t & \text{if } t \leq \varepsilon, \\ \varepsilon & \text{if } t > \varepsilon, \end{cases}$$

and hence for $0 < s < 2/p$ and $1/\tau = s/2 + 1/p,$ we have

$$|f|_{B_\tau^s} \sim \varepsilon^{1/\tau-s} \sim \varepsilon^{1/p-s/2} \sim \varepsilon^{-s/2} \|f\|_{L^p}, \text{ implying } |f|_{B_\tau^s} \not\leq c \|f\|_{L^p},$$

since ε can be arbitrarily small. It is easy to see that one comes to the same conclusion if f is the characteristic function of any convex elongated set in $\mathbb{R}^2.$ The point is that if $S_1, S_2 \in \mathcal{S}(n, 1),$ then $S_1 - S_2$ can be a constant multiple of the characteristic function of one or more elongated convex sets in $\mathbb{R}^2,$ and, therefore, estimate (3.8) is in general not possible.

We overcome the problem with estimate (3.8) by establishing the following main result:

Theorem 3.4 *Let $0 < p < \infty$, $0 < s < 2/p$, and $1/\tau = s/2 + 1/p$. Then for any $S_1, S_2 \in \mathcal{S}(n, 1)$, $n \geq n_0$, we have*

$$|S_1|_{B_\tau^s} \leq |S_2|_{B_\tau^s} + cn^{s/2} \|S_1 - S_2\|_{L^p} \text{ if } \tau \geq 1, \text{ and} \tag{3.9}$$

$$|S_1|_{B_\tau^s}^\tau \leq |S_2|_{B_\tau^s}^\tau + cn^{\tau s/2} \|S_1 - S_2\|_{L^p}^\tau \text{ if } \tau < 1, \tag{3.10}$$

where the constant $c > 0$ depends only on s, p , and the structural constants N_0, c_0 , and β ; n_0 is from Condition 3.2.

In the limiting case, we have this result:

Theorem 3.5 *If $S_1, S_2 \in \mathcal{S}(n, 1)$, $n \geq n_0$, then*

$$|S_1|_{BV} \leq |S_2|_{BV} + cn^{1/2} \|S_1 - S_2\|_{L^2}, \tag{3.11}$$

where the constant $c > 0$ depends only on the structural constants N_0, c_0 , and β .

The proof of this theorem is easier than the one of Theorem 3.4 and will be omitted.

We next show that estimates (3.9–3.10) and (3.11) imply the desired inverse estimate.

Theorem 3.6 *Let p, s , and τ be as in Theorem 3.4, and set $\lambda := \min\{\tau, 1\}$. Then for any $f \in L^p(\Omega)$, we have*

$$K(f, n^{-s/2}) \leq cn^{-s/2} \left(\sum_{\ell=n_0}^n \frac{1}{\ell} \left[\ell^{s/2} S_\ell^1(f)_p \right]^\lambda + \|f\|_p^\lambda \right)^{1/\lambda}, \quad n \geq n_0. \tag{3.12}$$

Here $K(f, t) = K(f, t; L^p, B_\tau^s)$ is the K -functional defined in (3.6), and $c > 0$ is a constant depending only on s, p , and the structural constants of the setting.

Furthermore, in the case when $p = 2$ and $s = 1$, estimate (3.12) holds with B_τ^s replaced by BV and $\lambda = 1$.

Proof Let $\tau < 1$ and $f \in L^p(\Omega)$. We may assume that for any $n \geq n_0$, there exists $S_n \in \mathcal{S}(n, 1)$ such that $\|f - S_n\|_p = S_n^1(f)_p$. Clearly, for any $m \geq m_0$ with $m_0 := \lceil \log_2 n_0 \rceil$, we have

$$K(f, 2^{-ms/2}) \leq \|f - S_{2^m}\|_p + 2^{-ms/2} |S_{2^m}|_{B_\tau^s}. \tag{3.13}$$

We now estimate $|S_{2^m}|_{B_\tau^s}^\tau$ using iteratively estimate (3.10). For $v \geq m_0 + 1$, we get

$$\begin{aligned} |S_{2^v}|_{B_\tau^s}^\tau &\leq |S_{2^{v-1}}|_{B_\tau^s}^\tau + c2^{\tau v s/2} \|S_{2^v} - S_{2^{v-1}}\|_p^\tau \\ &\leq |S_{2^{v-1}}|_{B_\tau^s}^\tau + c2^{\tau v s/2} \left(\|f - S_{2^v}\|_p^\tau + \|f - S_{2^{v-1}}\|_p^\tau \right) \\ &\leq |S_{2^{v-1}}|_{B_\tau^s}^\tau + c'2^{\tau v s/2} S_{2^{v-1}}^1(f)_p^\tau. \end{aligned}$$

From (3.7) we also have

$$|S_{2^{m_0}}|_{B_{\tau}^s} \leq c \|S_{2^{m_0}}\|_p \leq c \|f - S_{2^{m_0}}\|_p + c \|f\|_p = c S_{2^{m_0}}^1(f)_p + c \|f\|_p.$$

Summing up these estimates, we arrive at

$$|S_{2^m}|_{B_{\tau}^s} \leq c \sum_{v=m_0}^{m-1} 2^{\tau v s/2} S_{2^v}^1(f)_p^{\tau} + c \|f\|_p^{\tau}.$$

Clearly, this estimate and (3.13) imply (3.12). The proof in the cases $\lambda \geq 1$ or $p = 2$, $s = 1$, and B_{τ}^s replaced by BV is similar; we omit it. □

Observe that the direct and inverse estimates (3.5) and (3.9–3.11) imply immediately a characterization of the approximation spaces A_q^{α} associated with piecewise constant approximation from above just like in (1.7).

3.3 Proof of Theorem 3.4

We shall only consider the case when $\Omega \subset \mathbb{R}^2$ is a compact polygonal domain, obeying Condition 3.2. The proof in the case $\Omega = \mathbb{R}^2$ is similar.

Assume $S_1, S_2 \in \mathcal{S}(n, 1)$, $n \geq n_0$. Then S_1, S_2 can be represented in the form $S_j = \sum_{R \in \mathcal{R}_j} c_R \mathbb{1}_R$, where \mathcal{R}_j is a set of at most n rings in the sense of Definition 3.1 with disjoint interiors and such that $\Omega = \cup_{R \in \mathcal{R}_j} R$, $j = 1, 2$.

We denote by \mathcal{U} the set of all maximal compact connected subsets U of Ω obtained by intersecting all rings from \mathcal{R}_1 and \mathcal{R}_2 with the property $\overline{U^{\circ}} = U$ (the closure of the interior of U is U). Here U being maximal means that it is not contained in another such set.

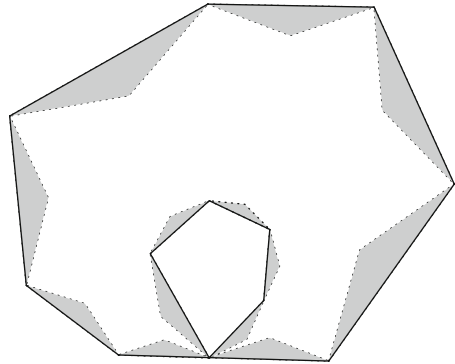
Observe first that each $U \in \mathcal{U}$ is obtained from the intersection of exactly two rings $R' \in \mathcal{R}_1$ and $R'' \in \mathcal{R}_2$, and is a subset of Ω with polygonal boundary ∂U consisting of $\leq 2N_0$ line segments (edges). Secondly, the sets in \mathcal{U} have disjoint interiors and $\Omega = \cup_{U \in \mathcal{U}} U$.

It is easy to see that there exists a constant $c > 0$ such that

$$\#\mathcal{U} \leq cn. \tag{3.14}$$

Indeed, each $U \in \mathcal{U}$ is obtained by intersecting two rings, say, $R' \in \mathcal{R}_1$ and $R'' \in \mathcal{R}_2$. If $|R'| < |R''|$, we associate R' with U , if $|R'| > |R''|$ we associate R'' with U , and if $|R'| = |R''|$, we associate either R' or R'' with U . However, because of condition (b) in Definition 3.1, every ring R from \mathcal{R}_1 or \mathcal{R}_2 can be intersected by only finitely many, say, N^* rings from \mathcal{R}_2 or \mathcal{R}_1 , respectively, of area $\geq |R|$. Here N^* depends only on the structural constants N_0 and c_0 . Also, the intersection of any two rings may have only finitely many, say N^{**} , connected components. Therefore, every ring $R \in \mathcal{R}_1 \cup \mathcal{R}_2$ can be associated with no more than $N^* N^{**}$ sets $U \in \mathcal{U}$, which implies (3.14) with $c = 2N^* N^{**}$.

Fig. 2 The ring from Fig. 1 with good triangles (angles = $\beta/2$)



Example 3.3 clearly indicates that our main problem will be in dealing with sets $U \in \mathcal{U}$ or parts of them with (diameter)² much larger than their area. To overcome the problem with these sets, we shall subdivide each of them using the following

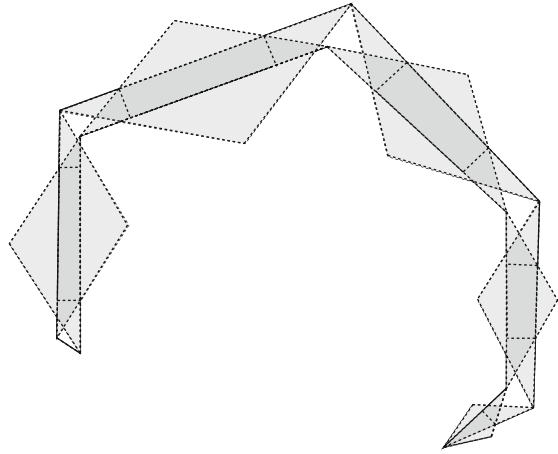
Construction of good triangles. According to Definition 3.1, each segment E from the boundary of every ring $R \in \mathcal{R}_j$ can be subdivided into the union of at most two segments E_1, E_2 ($E = E_1 \cup E_2$) with disjoint interiors such that there exist triangles Δ_1 with a side E_1 and adjacent to E_1 angles of size $\beta > 0$ and Δ_2 with a side E_2 and adjacent to E_2 angles β such that $\Delta_\ell \subset R, \ell = 1, 2$. We now associate with Δ_1 the triangle $\tilde{\Delta}_1 \subset \Delta_1$ with one side E_1 and adjacent to E_1 angles of size $\beta/2$; just in the same way we construct the triangle $\tilde{\Delta}_2 \subset \Delta_2$ with a side E_2 . We proceed in the same way for each edge E from $\partial R, R \in \mathcal{R}_j, j = 1, 2$. We denote by \mathcal{T}_R the set of all triangles $\tilde{\Delta}_1, \tilde{\Delta}_2$ associated in the above manner with all edges E from ∂R . We shall call the triangles from \mathcal{T}_R the good triangles associated with R . Observe that due to $\Delta_1, \Delta_2 \subset R$ for the triangles from above it readily follows that the good triangles associated with R ($R \in \mathcal{R}_j, j = 1, 2$) have disjoint interiors; this was the purpose of the above construction. To see this, one has simply to consider two arbitrary segments on ∂R and the associated triangles.

From now on, for every segment E from ∂R that has been subdivided into E_1 and E_2 as above, we shall consider E_1 and E_2 as segments from ∂R in place of E . We denote by \mathcal{E}_R the set of all (new) segments from ∂R . We now associate with each $E \in \mathcal{E}_R$ the good triangle that has E as a side and denote it by Δ_E .

To summarize, we have subdivided the boundary ∂R of each ring $R \in \mathcal{R}_j, j = 1, 2$, into a set \mathcal{E}_R of segments with disjoint interiors ($\partial R = \cup_{E \in \mathcal{E}_R} E$) and associated with each $E \in \mathcal{E}_R$ a good triangle $\Delta_E \subset R$ such that E is a side of Δ_E and the triangles $\{\Delta_E\}_{E \in \mathcal{E}_R}$ have disjoint interiors. In addition, if $E' \subset E$ is a subsegment of E , then we associate to E' the triangle $\Delta_{E'} \subset \Delta_E$ with one side E' and the other two sides parallel to the respective sides of Δ_E ; hence $\Delta_{E'}$ is similar to Δ_E . We shall call $\Delta_{E'}$ a good triangle as well. Fig. 2 illustrates the construction of good triangles (compare with Fig. 1).

Subdivision of the sets from \mathcal{U} . We next subdivide each set $U \in \mathcal{U}$ by using the good triangles constructed above. Suppose $U \in \mathcal{U}$ is obtained from the intersection of rings $R' \in \mathcal{R}_1$ and $R'' \in \mathcal{R}_2$. Then the boundary ∂U of U consists of two sets of

Fig. 3 A set U with its good triangles. Note also the trapezoids



segments \mathcal{E}'_U and \mathcal{E}''_U , where each $E \in \mathcal{E}'_U$ is a segment or subsegment of a segment from $\mathcal{E}_{R'}$ and each $E \in \mathcal{E}''_U$ is a segment or subsegment of a segment from $\mathcal{E}_{R''}$. Clearly, $\partial U = \cup_{E \in \mathcal{E}'_U \cup \mathcal{E}''_U} E$, and the segments from $\mathcal{E}'_U \cup \mathcal{E}''_U$ have disjoint interiors. For each $E \in \mathcal{E}'_U \cup \mathcal{E}''_U$, we denote by Δ_E the good triangle with a side E , defined above.

Definition of the set \mathcal{T}_U of trapezoids associated with $U \in \mathcal{U}$. We consider the collection of all nonempty sets of the form $\Delta_{E_1} \cap \Delta_{E_2}$ with the properties: (a) $E_1 \in \mathcal{E}'_U$, $E_2 \in \mathcal{E}''_U$, and (b) There exists an isosceles trapezoid or an isosceles triangle $T \subset \Delta_{E_1} \cap \Delta_{E_2}$ such that its two legs (of equal length) are contained in E_1 and E_2 , respectively, and its height is not smaller than its larger base. We assume that T is a maximal isosceles trapezoid (or triangle) with these properties. We denote by \mathcal{T}_U the set of all trapezoids as above.

Definition of the collection \mathcal{A}_U . We denote by \mathcal{A}_U the set of all maximal compact connected subsets A of $U \setminus \cup_{T \in \mathcal{T}_U} T^\circ$.

Clearly, $U = \cup_{T \in \mathcal{T}_U} T \cup_{A \in \mathcal{A}_U} A$, and the sets in $\mathcal{T}_U \cup \mathcal{A}_U$ have disjoint interiors. Figs. 3 and 8 illustrate of the above construction.

In the next lemma, we prove the “obvious” fact that as a result of the above subdivision of every set $U \in \mathcal{U}$ uncontrollably narrow and elongated subregions of U can only be realized as trapezoids from \mathcal{T}_U .

Lemma 3.7 *There exist constants $c^* > 1$ and $\beta^* > 0$ depending only on N_0, c_0 , and β , such that if $A \in \mathcal{A}_U$ for some $U \in \mathcal{U}$, then $d(A)^2 \leq c^*|A|$, and there exists a triangle $\Delta \subset A$ whose minimum angle is $\geq \beta^*$ such that $|A| \leq c^*|\Delta|$. Here $d(A)$ stands for the diameter of A .*

Proof Let $U \in \mathcal{U}$. There are only two possibilities for U : either U is of the form $U = (Q_2 \setminus \tilde{Q}_2) \setminus \tilde{Q}_1$ or of the form $U = (Q_1 \cap Q_2) \setminus (\tilde{Q}_1 \cup \tilde{Q}_2)$, where $R_1 = Q_1 \setminus \tilde{Q}_1$ and $R_2 = Q_2 \setminus \tilde{Q}_2$ are two rings (see Definition 3.1), one of which belongs to \mathcal{R}_1 and the other to \mathcal{R}_2 , see Fig. 4.

We shall only consider the case $U = (Q_2 \setminus \tilde{Q}_2) \setminus \tilde{Q}_1$; the other case is similar. Let $A \in \mathcal{A}_U$ (observe that if $\mathcal{T}_U = \emptyset$, then $\mathcal{A}_U = \{U\}$ and $A = U$). Define $\gamma_2 := \partial Q_2 \cap A$,

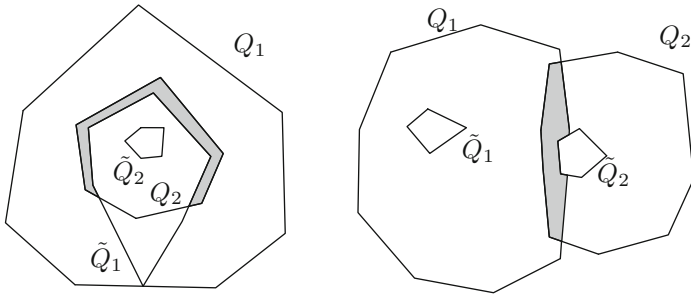
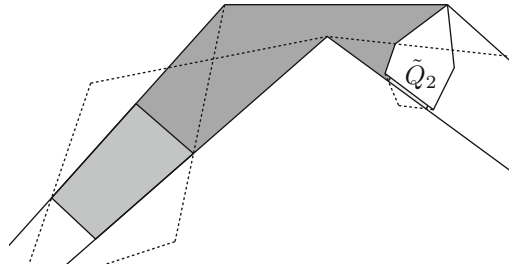


Fig. 4 Two possible configurations for U : $U = (Q_2 \setminus \tilde{Q}_2) \setminus \tilde{Q}_1$ (left) or $U = (Q_1 \cap Q_2) \setminus (\tilde{Q}_1 \cup \tilde{Q}_2)$ (right)

Fig. 5 One instance of $A \in \mathcal{A}_U$ (dark shade) and trapezoids (light shade). Observe that in this case, one of the trapezoids arises from a segment of \tilde{Q}_2



$\tilde{\gamma}_1 := \partial \tilde{Q}_1 \cap A$, and $\tilde{\gamma}_2 := \partial \tilde{Q}_2 \cap A$. Clearly, ∂A consists of $\gamma_2, \tilde{\gamma}_1, \tilde{\gamma}_2$ (if $\tilde{\gamma}_2 \neq \emptyset$) and at most two base segments of trapezoids from \mathcal{T}_U , see Fig. 5. Observe that from Definition 3.1, it follows that the number of edges of ∂A is $\leq 3N_0 + 2$.

Let E^* be the longest edge (line segment) of ∂A . There are four possibilities for E^* that we consider separately below.

Case 1: E^* is the base of a trapezoid in \mathcal{T}_U . Then from the construction of the trapezoids in \mathcal{T}_U , it readily follows that there exists a triangle $\Delta \subset A$ with a side E^* and minimal angle $\geq \beta/2$. Hence,

$$d(A)^2 \leq (3N_0 + 2)^2 \ell(E^*)^2 \leq c(3N_0 + 2)^2 |\Delta| \leq c'|A|, \tag{3.15}$$

and $|A| \leq d(A)^2 \leq c|\Delta|$ as claimed.

Case 2: E^* is an edge (line segment) of γ_2 . Let $\Delta_{E^*} \subset R_2$ be the good triangle with a side E^* . As such it follows that $\Delta_{E^*} \cap \tilde{Q}_2^\circ = \emptyset$. Denote by u_1, u_2, u_3 the vertices of Δ_{E^*} , where u_1, u_2 are the end points of E^* . Further, let u_4 be the point on the side $[u_1, u_3]$ of Δ_{E^*} such that $|u_1 - u_4| = |u_1 - u_3|/4$. Similarly, let u_5 be the point on the side $[u_2, u_3]$ such that $|u_2 - u_5| = |u_2 - u_3|/4$. Also, denote by u_6 and u_7 the points on E^* such that $|u_1 - u_6| = |u_1 - u_2|/4$ and $|u_2 - u_7| = |u_1 - u_2|/4$. Let $\Delta' := [u_1, u_4, u_6]$ be the triangle with vertices u_1, u_4, u_6 , and let $\Delta'' := [u_2, u_5, u_7]$. See Fig. 5.

If $\tilde{\gamma}_1 \cap \Delta' = \emptyset$, then $\Delta' \subset A$, and as in (3.15), we conclude that

$$d(A)^2 \leq 4^2(3N_0 + 2)^2 (\ell(E^*)/4)^2 \leq c|\Delta'| \leq c|A|$$

as claimed. The same argument applies whenever $\tilde{\gamma}_1 \cap \Delta'' = \emptyset$.

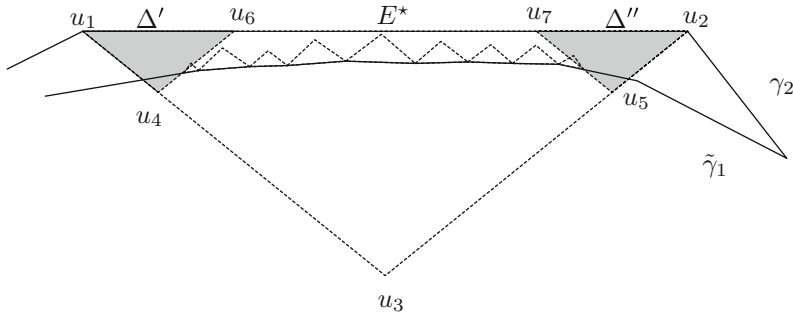


Fig. 6 Illustration of Case 2

Consider the case when $\tilde{\gamma}_1 \cap \Delta' \neq \emptyset$ and $\tilde{\gamma}_1 \cap \Delta'' \neq \emptyset$, see Fig. 6. Define $\tilde{\gamma}_1^\diamond := \tilde{\gamma}_1 \setminus (\Delta' \cup \Delta'')$. Let $\tilde{\mathcal{E}}_1^\diamond$ be the set of all edges of $\tilde{\gamma}_1^\diamond$. Clearly, $\ell(\tilde{\gamma}_1^\diamond) \geq \ell(E^*)/2$ and $\#\tilde{\mathcal{E}}_1^\diamond \leq N_0$. Since each good triangle $\Delta_E \subset R_1$ associated with an edge $E \in \tilde{\mathcal{E}}_1^\diamond$ does not form a trapezoid in \mathcal{T}_U , there exists a constant $\hat{\beta}$, depending only on β , such that $0 < \hat{\beta} < \beta/2$ and the triangle $\hat{\Delta}_E$ ($\hat{\Delta}_E \subset \Delta_E$) with angles adjacent to E of size $\hat{\beta}$ does not intersect E^* . Also, from the fact that E is contained in the trapezoid $[u_4, u_5, u_7, u_6]$, it follows that $\hat{\Delta}_E$ cannot intersect the other two sides of Δ_{E^*} . Hence, $\hat{\Delta}_E \subset \Delta_{E^*}$. Using this, we obtain

$$\begin{aligned}
 d(A)^2 &\leq (3N_0 + 2)^2 \ell(E^*)^2 \leq 4(3N_0 + 2)^2 \ell(\tilde{\gamma}_1^\diamond)^2 = 4(3N_0 + 2)^2 \left(\sum_{E \in \tilde{\mathcal{E}}_1^\diamond} \ell(E) \right)^2 \\
 &\leq 4(3N_0 + 2)^2 N_0 \sum_{E \in \tilde{\mathcal{E}}_1^\diamond} \ell(E)^2 \leq c \sum_{E \in \tilde{\mathcal{E}}_1^\diamond} |\hat{\Delta}_E| \leq c|A|,
 \end{aligned}
 \tag{3.16}$$

where we used that the triangles $\hat{\Delta}_E, E \in \tilde{\mathcal{E}}_1^\diamond$, are with disjoint interiors and $\hat{\Delta}_E \subset A$. Observe also that if $\Delta \in \{\hat{\Delta}_E : E \in \tilde{\mathcal{E}}_1^\diamond\}$ is a triangle of largest area from this set of triangles, then it follows from above that $|A| \leq d(A)^2 \leq c|\Delta|$. This completes the proof of the lemma in Case 2.

Case 3: E^* is an edge of $\tilde{\gamma}_2$. In this case, the argument is just as the one in Case 2. We omit the details.

Case 4: E^* is an edge of $\tilde{\gamma}_1$ (recall that E^* is the longest edge of ∂A). Let $\Delta_{E^*} \subset R_1$ be the good triangle with a side E^* . Two subcases present themselves here depending on whether $\Delta_{E^*} \cap \tilde{Q}_2^\diamond = \emptyset$ or $\Delta_{E^*} \cap \tilde{Q}_2^\diamond \neq \emptyset$.

Case 4 (a): $\Delta_{E^*} \cap \tilde{Q}_2^\diamond = \emptyset$. Let u_1, u_2, u_3 be the vertices of Δ_{E^*} , where u_1, u_2 are the end points of E^* . We define the points u_4, u_5, u_6, u_7 on the sides of Δ_{E^*} just as in Case 2 above, see Fig. 7.

Assume $\gamma_2 \cap [u_4, u_5, u_3] \neq \emptyset$, where $[u_4, u_5, u_3]$ stands for the triangle with vertices u_4, u_5, u_3 . Pick a point $u \in \gamma_2 \cap [u_4, u_5, u_3]$. Because of the convexity of Q_2 , the triangle $\Delta := [u_1, u_2, u]$ is contained in A , and hence

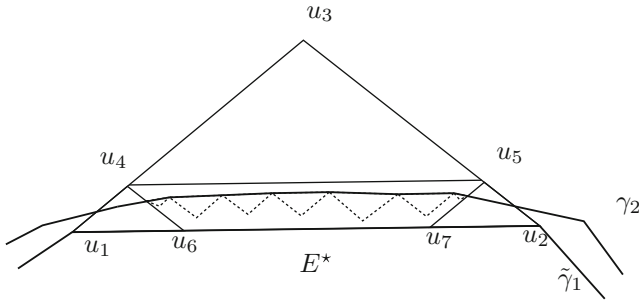


Fig. 7 Illustration of Case 4 (a)

$$d(A)^2 \leq (3N_0 + 2)^2 \ell(E^*)^2 \leq c(3N_0 + 2)^2 |\Delta| \leq c'|A|$$

as claimed.

Assume $\gamma_2 \cap [u_4, u_5, u_3] = \emptyset$. Then γ_2 intersects the segments $[u_4, u_6]$ and $[u_5, u_7]$. Set $\gamma_2^\diamond := \gamma_2 \cap [u_6, u_7, u_5, u_4]$, where $[u_6, u_7, u_5, u_4]$ is the trapezoid with vertices u_6, u_7, u_5, u_4 . Let \mathcal{E}_2^\diamond be the set of all edges of γ_2^\diamond . Clearly, $\ell(\gamma_2^\diamond) > \ell(E^*)/2$ and $\#\mathcal{E}_2^\diamond \leq N_0$. Just as in Case 2, we note that each good triangle $\Delta_E \subset R_2$ associated with an edge $E \in \mathcal{E}_2^\diamond$ does not form a trapezoid in \mathcal{T}_U , and hence the triangle $\hat{\Delta}_E$ ($\hat{\Delta}_E \subset \Delta_E$) with angles adjacent to E of size $\hat{\beta}$ with $0 < \hat{\beta} < \beta/2$ as in Case 2 is contained in $Q_2 \cap \Delta_{E^*}$. Therefore, as in (3.16),

$$\begin{aligned} d(A)^2 &\leq (3N_0 + 2)^2 \ell(E^*)^2 \leq 4(3N_0 + 2)^2 \ell(\gamma_2^\diamond)^2 = 4(3N_0 + 2)^2 \left(\sum_{E \in \mathcal{E}_2^\diamond} \ell(E) \right)^2 \\ &\leq 4(3N_0 + 2)^2 N_0 \sum_{E \in \mathcal{E}_2^\diamond} \ell(E)^2 \leq c \sum_{E \in \mathcal{E}_2^\diamond} |\hat{\Delta}_E| \leq c|A|. \end{aligned}$$

Furthermore, if $\Delta \in \{\hat{\Delta}_E : E \in \mathcal{E}_2^\diamond\}$ is a triangle of largest area from this set of triangles, then $|A| \leq d(A)^2 \leq c|\Delta|$, which completes the proof in this subcase.

Case 4 (b): $\Delta_{E^*} \cap \tilde{Q}_2^\diamond \neq \emptyset$. Observe that $A \cap \Delta_{E^*} = (Q_2 \setminus \tilde{Q}_2) \cap \Delta_{E^*}$. Just as in Case 4 (a), one shows that

$$\ell(E^*)^2 \leq c|Q_2 \cap \Delta_{E^*}|. \tag{3.17}$$

We next prove that there exists a constant $c' > 0$ such that

$$|Q_2 \cap \Delta_{E^*}| \leq c'|(Q_2 \cap \Delta_{E^*}) \setminus \tilde{Q}_2| = c'|A \cap \Delta_{E^*}|. \tag{3.18}$$

Define $\tilde{Q}_2^\diamond := \Delta_{E^*} \cap \tilde{Q}_2$, $\tilde{\gamma}_2^\diamond := \tilde{\gamma}_2 \cap \Delta_{E^*}$, and let $\tilde{\mathcal{E}}_2^\diamond$ be the set of all edges of $\tilde{\gamma}_2^\diamond$. Note that just as in Case 2 and Case 4 (a), each good triangle $\Delta_E \subset R_2$ associated with an edge $E \in \tilde{\mathcal{E}}_2^\diamond$ does not form with E^* a trapezoid in \mathcal{T}_U , and hence the triangle $\hat{\Delta}_E$ ($\hat{\Delta}_E \subset \Delta_E$) with angles adjacent to E of size $\hat{\beta}$ with $0 < \hat{\beta} < \beta/2$ as in Case 2

does not intersect E^* . We claim that there exists a constant $c'' > 0$ such that

$$|\tilde{Q}_2 \cap \Delta_{E^*}| \leq c'' |(\cup_{E \in \tilde{\mathcal{E}}_2^\circ} \hat{\Delta}_E) \cap \Delta_{E^*}|. \tag{3.19}$$

As in Case 4 (a), let $\Delta_{E^*} =: [u_1, u_2, u_3]$, where $E^* = [u_1, u_2]$. Denote by n_1 and n_2 the unit vectors that are orthogonal to the sides $[u_1, u_3]$ and $[u_2, u_3]$, respectively, and exterior to Δ_{E^*} . Observe that since Δ_{E^*} is a good triangle, the angle made by the sides $[u_1, u_3]$ and $[u_2, u_3]$ is of size $\geq \pi - \beta \geq 2\pi/3$. Further, denote by $\tilde{\mathcal{E}}_2^b$ the set of all edges $E \in \tilde{\mathcal{E}}_2^\circ$ whose exterior (to \tilde{Q}_2) normal vectors make angles $\geq \pi/2$ with n_1 and n_2 . Clearly, $\hat{\Delta}_E \cap ([u_1, u_3] \cup [u_2, u_3]) = \emptyset, \forall E \in \tilde{\mathcal{E}}_2^b$, and hence

$$\hat{\Delta}_E \subset (Q_2 \cap \Delta_{E^*}) \setminus \tilde{Q}_2, \quad \forall E \in \tilde{\mathcal{E}}_2^b. \tag{3.20}$$

On the other hand, since the convex set \tilde{Q}_2 is with bounded eccentricity (Definition 3.1), it readily follows that there exist constants $c_1, c_2 > 0$ such that

$$\sum_{E \in \tilde{\mathcal{E}}_2^b} \ell(E) \geq c_1 \sum_{E \in \tilde{\mathcal{E}}_2^\circ} \ell(E) \geq c_2 d(\tilde{Q}_2 \cap \Delta_{E^*}). \tag{3.21}$$

From (3.20–3.21) it follows that

$$\begin{aligned} |\tilde{Q}_2 \cap \Delta_{E^*}| &\leq d(\tilde{Q}_2 \cap \Delta_{E^*})^2 \leq c \left(\sum_{E \in \tilde{\mathcal{E}}_2^b} \ell(E) \right)^2 \leq cN_0 \sum_{E \in \tilde{\mathcal{E}}_2^b} \ell(E)^2 \\ &\leq cN_0 \sum_{E \in \tilde{\mathcal{E}}_2^b} |\hat{\Delta}_E| \leq c \sum_{E \in \tilde{\mathcal{E}}_2^\circ} |\hat{\Delta}_E \cap \Delta_{E^*}|, \end{aligned}$$

which confirms (3.19).

To prove (3.18), we consider two cases. If $|\tilde{Q}_2 \cap \Delta_{E^*}| \leq |Q_2 \cap \Delta_{E^*}|/2$, then (3.18) follows trivially. Assume $|\tilde{Q}_2 \cap \Delta_{E^*}| > |Q_2 \cap \Delta_{E^*}|/2$. Then using (3.19),

$$|Q_2 \cap \Delta_{E^*}| \leq 2|\tilde{Q}_2 \cap \Delta_{E^*}| \leq c |(\cup_{E \in \tilde{\mathcal{E}}_2^\circ} \hat{\Delta}_E) \cap \Delta_{E^*}| \leq c' |(Q_2 \cap \Delta_{E^*}) \setminus \tilde{Q}_2|,$$

which completes the proof of (3.18).

Finally, (3.17) and (3.18) imply

$$d(A)^2 \leq (3N_0 + 2)^2 \ell(E^*)^2 \leq c|A \cap \Delta_{E^*}| \leq c|A|.$$

Therefore, $d(A)^2 \leq c|A|$ as claimed.

In the case when $|\tilde{Q}_2 \cap \Delta_{E^*}| \leq |Q_2 \cap \Delta_{E^*}|/2$, just as in Case 4 (a), the triangle $\Delta \in \{\hat{\Delta}_E : E \in \tilde{\mathcal{E}}_2^\circ\}$ of largest area has the property $|A| \leq d(A)^2 \leq c|\Delta|$. In the other case, the triangle $\Delta \in \{\hat{\Delta}_E : E \in \tilde{\mathcal{E}}_2^\circ, \hat{\Delta}_E \subset \Delta_{E^*}\}$ of largest area has this property. The proof of Lemma 3.7 is complete. \square

In what follows, we shall need the following obvious property of the trapezoids from \mathcal{T} .

Property 3.8 There exists a constant $0 < \hat{c} < 1$ such that if $L = [v_1, v_2]$ is one of the legs of a trapezoid $T \in \mathcal{T}$ and $T \subset \Delta_{E_1} \cap \Delta_{E_2}$ (see the construction of trapezoids), then for any $x \in L$ with $|x - v_j| \geq \rho$, $j = 1, 2$, for some $\rho > 0$ we have $B(x, \hat{c}\rho) \subset \Delta_{E_1} \cup \Delta_{E_2}$. Moreover, if $D = [v_1, v_2]$ is one of the bases of the trapezoid T , then for any $x \in D$ with $|x - v_j| \geq \rho$, $j = 1, 2$, for some $\rho > 0$ we have $B(x, \hat{c}\rho) \subset \Delta_{E_1} \cap \Delta_{E_2}$.

Let $\mathcal{A} := \cup_{U \in \mathcal{U}} \mathcal{A}_U$ and $\mathcal{T} := \cup_{U \in \mathcal{U}} \mathcal{T}_U$. We have $\Omega = \cup_{A \in \mathcal{A}} A \cup_{T \in \mathcal{T}} T$, and, clearly, the sets in $\mathcal{A} \cup \mathcal{T}$ have disjoint interiors. From these we obtain the following representation of $S_1(x) - S_2(x)$ for $x \in \Omega$ which is not on any of the edges:

$$S_1(x) - S_2(x) = \sum_{A \in \mathcal{A}} c_A \mathbb{1}_A(x) + \sum_{T \in \mathcal{T}} c_T \mathbb{1}_T(x), \tag{3.22}$$

where c_A and c_T are constants.

For future reference, we note that

$$\#\mathcal{A} \leq cn \quad \text{and} \quad \#\mathcal{T} \leq cn. \tag{3.23}$$

These estimates follow readily by (3.14) and the fact that the number of edges of each $U \in \mathcal{U}$ is $\leq 2N_0$.

Let $0 < s/2 < 1/p$, and assume $\tau \leq 1$. Fix $t > 0$, and let $h \in \mathbb{R}^2$ with norm $|h| \leq t$. Write $v := |h|^{-1}h$, and assume $v =: (\cos \theta, \sin \theta)$, $-\pi < \theta \leq \pi$.

We shall frequently use the following obvious identities: If S is a constant on a measurable set $G \subset \mathbb{R}^2$ and $H \subset G$ (H measurable), then

$$\|S\|_{L^\tau(G)} = |G|^{1/\tau-1/p} \|S\|_{L^p(G)} = |G|^{s/2} \|S\|_{L^p(G)} \tag{3.24}$$

and

$$\|S\|_{L^\tau(H)} = (|H|/|G|)^{1/\tau} \|S\|_{L^\tau(G)}. \tag{3.25}$$

We next estimate $\|\Delta_h S_1\|_{L^\tau(G)}^\tau - \|\Delta_h S_2\|_{L^\tau(G)}^\tau$ for different subsets G of Ω .

Case 1

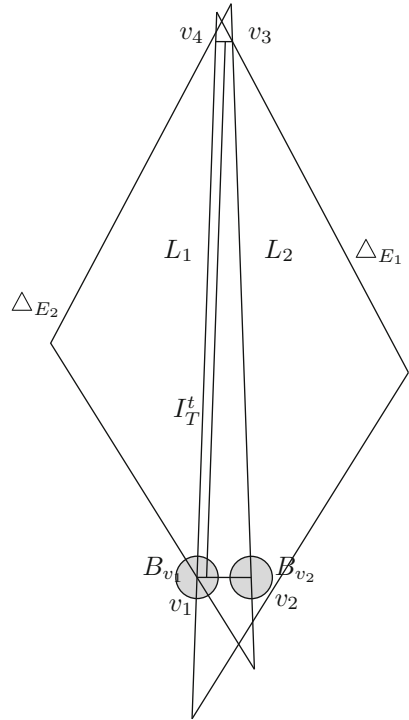
Let $T \in \mathcal{T}$ be such that $d(T) > 2t/\hat{c}$ with \hat{c} the constant from Property 3.8. Define

$$T_h := \{x \in \Omega : [x, x + h] \subset \Omega \text{ and } [x, x + h] \cap T \neq \emptyset\}.$$

We now estimate $\|\Delta_h S_1\|_{L^\tau(T_h)}^\tau - \|\Delta_h S_2\|_{L^\tau(T_h)}^\tau$.

We may assume that T is an isosceles trapezoid contained in $\Delta_{E_1} \cap \Delta_{E_2}$, where Δ_{E_j} ($j = 1, 2$) is a good triangle for a ring $R_j \in \mathcal{R}_j$ and T is positioned so that its vertices are the points

Fig. 8 A trapezoid T



$$v_1 := (-\delta_1/2, 0), \quad v_2 := (\delta_1/2, 0), \quad v_3 := (\delta_2/2, H), \quad v_4 := (-\delta_2/2, H),$$

where $0 \leq \delta_2 \leq \delta_1$ and $H > \delta_1$. Let $L_1 := [v_1, v_4]$ and $L_2 := [v_2, v_3]$ be the two equal (long) legs of T . We assume that $L_1 \subset E_1$ and $L_2 \subset E_2$. We denote by $D_1 := [v_1, v_2]$ and $D_2 := [v_3, v_4]$ the two bases of T . Set $\mathcal{V}_T := \{v_1, v_2, v_3, v_4\}$. See Fig. 8.

Furthermore, let $\gamma \leq \pi/2$ be the angle between D_1 and L_1 , and assume that $v = (\cos \theta, \sin \theta)$ with $\theta \in [\gamma, \pi]$. The case $\theta \in [-\gamma, 0]$ is just the same. The case when $\theta \in [0, \gamma] \cup [-\pi, -\gamma]$ is considered similarly.

Define $B_v := B(v, 2t/\hat{c})$, $v \in \mathcal{V}_T$,

$$\begin{aligned} \mathcal{A}_T^t &:= \{A \in \mathcal{A} : d(A) > t \text{ and } A \cap (T + B(0, t)) \neq \emptyset\}, \\ \mathfrak{A}_T^t &:= \{A \in \mathcal{A} : d(A) \leq t \text{ and } A \cap (T + B(0, t)) \neq \emptyset\}, \end{aligned}$$

and

$$\begin{aligned} \mathcal{T}_T^t &:= \{T' \in \mathcal{T} : d(T') > 2t/\hat{c} \text{ and } T' \cap (T + B(0, t)) \neq \emptyset\}, \\ \mathfrak{T}_T^t &:= \{T' \in \mathcal{T} : d(T') \leq 2t/\hat{c} \text{ and } T' \cap (T + B(0, t)) \neq \emptyset\}. \end{aligned}$$

Case 1 (a). If $[x, x + h] \in \Delta_{E_1}^\circ$, then $\Delta_h S_1(x) = 0$ because S_1 is a constant on Δ_{E_1} . Hence no estimate is needed.

Case 1 (b). If $[x, x + h] \subset \cup_{v \in \mathcal{V}_T} B_v$, we estimate $|\Delta_h S_1(x)|$ using the obvious inequality

$$|\Delta_h S_1(x)| \leq |\Delta_h S_2(x)| + |S_1(x) - S_2(x)| + |S_1(x + h) - S_2(x + h)|. \tag{3.26}$$

Clearly, the contribution of this case to estimating $\|\Delta_h S_1\|_{L^\tau(T_h)}^\tau - \|\Delta_h S_2\|_{L^\tau(T_h)}^\tau$ is

$$\begin{aligned} &\leq c \sum_{v \in \mathcal{V}_T} \sum_{A \in \mathcal{A}_T^t} \|S_1 - S_2\|_{L^\tau(B_v \cap A)}^\tau + c \sum_{v \in \mathcal{V}_T} \sum_{T' \in \mathcal{T}_T^t} \|S_1 - S_2\|_{L^\tau(B_v \cap T')}^\tau \\ &\quad + c \sum_{v \in \mathcal{V}_T} \sum_{A \in \mathcal{A}_T^t} \|S_1 - S_2\|_{L^\tau(B_v \cap A)}^\tau + c \sum_{v \in \mathcal{V}_T} \sum_{T' \in \mathcal{T}_T^t} \|S_1 - S_2\|_{L^\tau(B_v \cap T')}^\tau \\ &\leq \sum_{A \in \mathcal{A}_T^t} ct^2 d(A)^{\tau s - 2} \|S_1 - S_2\|_{L^p(A)}^\tau + \sum_{T' \in \mathcal{T}_T^t} ct^{1 + \tau s / 2} d(T')^{\tau s / 2 - 1} \|S_1 - S_2\|_{L^p(T')}^\tau \\ &\quad + \sum_{A \in \mathcal{A}_T^t} cd(A)^{\tau s} \|S_1 - S_2\|_{L^p(A)}^\tau + \sum_{T' \in \mathcal{T}_T^t} cd(T')^{\tau s} \|S_1 - S_2\|_{L^p(T')}^\tau. \end{aligned}$$

Here we used these estimates, obtained using Lemma 3.7 and (3.24) or/and (3.25):

(1) If $A \in \mathcal{A}_T^t$ and $v \in \mathcal{V}_T$, then

$$\begin{aligned} \|S_1 - S_2\|_{L^\tau(B_v \cap A)}^\tau &= (|B_v|/|A|) \|S_1 - S_2\|_{L^\tau(A)}^\tau \\ &\leq ct^2 d(A)^{-2} \|S_1 - S_2\|_{L^\tau(A)}^\tau \leq ct^2 d(A)^{\tau s - 2} \|S_1 - S_2\|_{L^p(A)}^\tau. \end{aligned}$$

(2) If $T' \in \mathcal{T}_T^t$ and $\delta_1(T') > 2t/\hat{c}$ with $\delta_1(T')$ being the maximal base of T' , then for any $v \in \mathcal{V}_T$, we have

$$\begin{aligned} \|S_1 - S_2\|_{L^\tau(B_v \cap T')}^\tau &= (|B_v|/|T'|) \|S_1 - S_2\|_{L^\tau(T')}^\tau \leq ct^2 |T'|^{\tau s / 2 - 1} \|S_1 - S_2\|_{L^p(T')}^\tau \\ &\leq ct^2 \delta_1(T')^{\tau s / 2 - 1} d(T')^{\tau s / 2 - 1} \|S_1 - S_2\|_{L^p(T')}^\tau \\ &\leq ct^{1 + \tau s / 2} d(T')^{\tau s / 2 - 1} \|S_1 - S_2\|_{L^p(T')}^\tau, \end{aligned}$$

where we used that $\tau s / 2 < 1$, which is equivalent to $s < s + 2/p$.

(3) If $T' \in \mathcal{T}_T^t$ and $\delta_1(T') \leq 2t/\hat{c}$, then for any $v \in \mathcal{V}_T$,

$$\begin{aligned} \|S_1 - S_2\|_{L^\tau(B_v \cap T')}^\tau &= (|B_v \cap T'|/|T'|) \|S_1 - S_2\|_{L^\tau(T')}^\tau \\ &= |B_v \cap T'| |T'|^{\tau s / 2 - 1} \|S_1 - S_2\|_{L^p(T')}^\tau \\ &\leq ct \delta_1(T') [\delta_1(T') d(T')]^{\tau s / 2 - 1} \|S_1 - S_2\|_{L^p(T')}^\tau \\ &= ct \delta_1(T')^{\tau s / 2} d(T')^{\tau s / 2 - 1} \|S_1 - S_2\|_{L^p(T')}^\tau \\ &\leq ct^{1 + \tau s / 2} d(T')^{\tau s / 2 - 1} \|S_1 - S_2\|_{L^p(T')}^\tau. \end{aligned}$$

(4) If $A \in \mathfrak{A}_T^t$, then

$$\begin{aligned} \|S_1 - S_2\|_{L^\tau(B_v \cap A)}^\tau &\leq \|S_1 - S_2\|_{L^\tau(A)}^\tau \leq c|A|^{\tau s/2} \|S_1 - S_2\|_{L^p(A)}^\tau \\ &\leq cd(A)^{\tau s} \|S_1 - S_2\|_{L^p(A)}^\tau. \end{aligned}$$

(5) If $T' \in \mathfrak{T}_T^t$, then

$$\begin{aligned} \|S_1 - S_2\|_{L^\tau(B_v \cap T')}^\tau &\leq \|S_1 - S_2\|_{L^\tau(T')}^\tau \leq c|T'|^{\tau s/2} \|S_1 - S_2\|_{L^p(T')}^\tau \quad (3.27) \\ &\leq cd(T')^{\tau s} \|S_1 - S_2\|_{L^p(T')}^\tau. \end{aligned}$$

Case 1 (c). If $[x, x + h] \not\subset \cup_{v \in \mathcal{V}_T} B_v$ and $[x, x + h]$ intersects D_1 or D_2 , then $\delta_1 > 2t/\hat{c} > 2t$ or $\delta_2 > 2t$ and hence $[x, x + h] \subset \Delta_{E_1} \cap \Delta_{E_2}$, which implies $\Delta_h S_1(x) = 0$. No estimate is needed.

Case 1 (d). Let I_T^t be the set defined by

$$I_T^t := \{x \in T : x \text{ is between } L_1 \text{ and } L_1 + \varepsilon e_1\} \setminus (B(v_1, t/\hat{c}) \cup B(v_4, t/\hat{c})), \quad (3.28)$$

where $\varepsilon := (\delta_1 - \delta_2)M^{-1}t$, $e_1 := \langle 1, 0 \rangle$, and $M := |L_1| = |L_2|$. Set $J_T^h := I_T^t + [0, h]$. See Fig. 8.

In this case, we again use (3.26) to estimate $|\Delta_h S_1(x)|$. We obtain

$$\begin{aligned} \|\Delta_h S_1\|_{L^\tau(I_T^t)}^\tau &\leq \|\Delta_h S_2\|_{L^\tau(I_T^t)}^\tau + \|S_1 - S_2\|_{L^\tau(I_T^t)}^\tau \\ &\quad + \sum_{A \in \mathfrak{A}_T^t} \|S_1 - S_2\|_{L^\tau(J_T^h \cap A)}^\tau + \sum_{A \in \mathfrak{A}_T^t} \|S_1 - S_2\|_{L^\tau(J_T^h \cap A)}^\tau. \end{aligned}$$

Clearly, $|I_T^t| \leq ct\delta_1(T)$ and $|T| \sim \delta_1(T)d(T)$. Then using (3.24–3.25), we infer

$$\begin{aligned} \|S_1 - S_2\|_{L^\tau(I_T^t)}^\tau &= (|I_T^t|/|T|) \|S_1 - S_2\|_{L^\tau(T)}^\tau \leq ctd(T)^{-1} \|S_1 - S_2\|_{L^\tau(T)}^\tau \\ &= ctd(T)^{-1} |T|^{\tau s/2} \|S_1 - S_2\|_{L^p(T)}^\tau \leq ctd(T)^{\tau s - 1} \|S_1 - S_2\|_{L^p(T)}^\tau. \end{aligned}$$

Similarly, for $A \in \mathfrak{A}_T^t$, we use that $|J_T^h \cap A| \leq ctd(A)$ and $|A| \sim d(A)^2$ to obtain

$$\begin{aligned} \|S_1 - S_2\|_{L^\tau(J_T^h \cap A)}^\tau &\leq ctd(A) \|S_1 - S_2\|_{L^\infty(A)}^\tau = ctd(A) |A|^{-\tau/p} \|S_1 - S_2\|_{L^p(A)}^\tau \\ &\leq ctd(A)^{1 - 2\tau/p} \|S_1 - S_2\|_{L^p(A)}^\tau \leq ctd(A)^{\tau s - 1} \|S_1 - S_2\|_{L^p(A)}^\tau. \end{aligned}$$

For $A \in \mathfrak{A}_T^t$, we have

$$\begin{aligned} \|S_1 - S_2\|_{L^\tau(J_T^h \cap A)}^\tau &\leq \|S_1 - S_2\|_{L^\tau(A)}^\tau = |A|^{\tau s/2} \|S_1 - S_2\|_{L^p(A)}^\tau \\ &\leq cd(A)^{\tau s} \|S_1 - S_2\|_{L^p(A)}^\tau. \end{aligned}$$

Putting the above estimates together, we get

$$\begin{aligned} \|\Delta_h S_1\|_{L^\tau(I_T^t)}^\tau &\leq \|\Delta_h S_2\|_{L^\tau(I_T^t)}^\tau + ct d(T)^{\tau s-1} \|S_1 - S_2\|_{L^p(T)}^\tau \\ &\quad + \sum_{A \in \mathcal{A}_T^t} ct d(A)^{\tau s-1} \|S_1 - S_2\|_{L^\infty(A)}^\tau + \sum_{A \in \mathfrak{A}_T^t} cd(A)^\tau \|S_1 - S_2\|_{L^p(A)}^\tau. \end{aligned}$$

Case 1 (e) (Main). Let $T_h^* \subset T_h$ be defined by

$$T_h^* := \{x \in T_h : [x, x + h] \cap L_1 \neq \emptyset, x \notin I_T^t, [x, x + h] \not\subset \bigcup_{v \in \mathcal{V}_T} B_v\}. \tag{3.29}$$

We next estimate $\|\Delta_h^k S_1\|_{L^\tau(T_h^*)}^\tau$.

Recall that by assumption $h = |h|v$ with $v =: (\cos \theta, \sin \theta)$ and $\theta \in [\gamma, \pi]$, where $\gamma \leq \pi/2$ is the angle between D_1 and L_1 .

Let $x \in T_h^*$. With the notation $x = (x_1, x_2)$, we let $(-a, x_2) \in L_1$ and $(a, x_2) \in L_2$, $a > 0$, be the points of intersection of the horizontal line through x with L_1 and L_2 . Set $b := 2a - \varepsilon$ with $\varepsilon := (\delta_1 - \delta_2)M^{-1}t$, see (3.28).

We associate the points $x + be_1$ and $x + be_1 + h$ with x and $x + h$. A simple geometric argument shows that $x + be_1 \in \Delta_{E_1} \setminus T$, while $x + be_1 + h \in T^\circ$.

Now, using that $S_1 = \text{constant}$ on $\Delta_{E_1}^\circ$, we have $S_1(x) = S_1(x + be_1)$, and since $S_2 = \text{constant}$ on $\Delta_{E_2}^\circ$, we have $S_2(x + h) = S_2(x + be_1 + h)$. We use these two identities to obtain

$$\begin{aligned} S_1(x + h) - S_1(x) &= S_2(x + be_1 + h) - S_2(x + be_1) \\ &\quad + [S_1(x + h) - S_2(x + h)] - [S_1(x + be_1) - S_2(x + be_1)], \end{aligned}$$

and, therefore,

$$\begin{aligned} |\Delta_h S_1(x)| &\leq |\Delta_h S_2(x + be_1)| \\ &\quad + |S_1(x + h) - S_2(x + h)| + |S_1(x + be_1) - S_2(x + be_1)|. \end{aligned} \tag{3.30}$$

Some words of explanation are in order here. The purpose of the set I_T^t is that there is one-to-one correspondence between pairs of points $x \in T^\circ \setminus I_T^t, x + h \in \Delta_{E_2} \setminus T$ and $x + be_1 \in \Delta_{E_1} \setminus T, x + be_1 + h \in T^\circ$. Due to $\delta_2 < \delta_1$, this would not be true if I_T^t was not removed from T° . Thus there is one-to-one correspondence between the differences $|\Delta_h S_1(x)|$ and $|\Delta_h S_2(x + be_1)|$ in the case under consideration. Also, it is important that $\Delta_h S_1(x + be_1) = 0$, and hence $|\Delta_h S_2(x + be_1)|$ need not be used to estimate $|\Delta_h S_1(x + be_1)|$.

Another important point here is that $x + h \notin T^\circ$ and $x + be_1 \notin T^\circ$. Therefore, no quantities $|S_1(x) - S_2(x)|$ with $x \in T^\circ \setminus I_T^t$ are involved in (3.30), which is critical.

Observe that for $x \in T_h^*$, we have

$$[x, x + h] \not\subset \bigcup_{v \in \mathcal{V}_T} B_v, \quad \text{and hence} \quad [x + be_1, x + be_1 + h] \not\subset \bigcup_{v \in \mathcal{V}_T} B_v.$$

Therefore, by Property 3.8, it follows that $[x, x + h]$ and $[x + be_1, x + be_1 + h]$ do not intersect any trapezoid $T' \in \mathcal{T}, T' \neq T$.

Let $T_h^{**} := \{x + be_1 : x \in T_h^*\}$. For any $A \in \mathcal{A}$ and $t > 0$, define

$$A_t := \{x \in A : \text{dist}(x, \partial A) \leq t\}. \tag{3.31}$$

From all of the above, we get

$$\|\Delta_h S_1\|_{L^\tau(T_h^*)}^\tau \leq \|\Delta_h S_2\|_{L^\tau(T_h^{**})}^\tau + \sum_{A \in \mathcal{A}_T^t} \|S_1 - S_2\|_{L^\tau(A_t)}^\tau + \sum_{A \in \mathfrak{A}_T^t} \|S_1 - S_2\|_{L^\tau(A)}^\tau.$$

Now, using that $|A_t| \leq ctd(A)$ and $|A| \sim d(A)^2$ for $A \in \mathcal{A}_T^t$, we obtain

$$\begin{aligned} \|S_1 - S_2\|_{L^\tau(A_t)}^\tau &= (|A_t|/|A|)|A|^{\tau s/2} \|S_1 - S_2\|_{L^p(A)}^\tau \\ &\leq ctd(A)^{\tau s-1} \|S_1 - S_2\|_{L^p(A)}^\tau. \end{aligned} \tag{3.32}$$

For $A \in \mathfrak{A}_T^t$, we use that $|A| \sim d(A)^2$ and obtain

$$\|S_1 - S_2\|_{L^\tau(A)}^\tau = |A|^{\tau s/2} \|S_1 - S_2\|_{L^p(A)}^\tau \leq cd(A)^{\tau s} \|S_1 - S_2\|_{L^p(A)}^\tau. \tag{3.33}$$

Inserting these estimates above, we get

$$\begin{aligned} \|\Delta_h S_1\|_{L^\tau(T_h^*)}^\tau &\leq \|\Delta_h S_2\|_{L^\tau(T_h^{**})}^\tau + \sum_{A \in \mathcal{A}_T^t} ctd(A)^{\tau s-1} \|S_1 - S_2\|_{L^p(A)}^\tau \\ &\quad + \sum_{A \in \mathfrak{A}_T^t} cd(A)^{\tau s} \|S_1 - S_2\|_{L^p(A)}^\tau. \end{aligned} \tag{3.34}$$

Case 2

Let Ω_h^* be the set of all $x \in \Omega$ such that $[x, x + h] \subset \Omega$ and $[x, x + h] \cap T = \emptyset$ for all $T \in \mathcal{T}$ with $d(T) \geq 2t/\hat{c}$. To estimate $|\Delta_h S_1(x)|$, we again use (3.26). With the notation from (3.31), we get

$$\begin{aligned} \|\Delta_h S_1\|_{L^\tau(\Omega_h^*)}^\tau &\leq \|\Delta_h S_2\|_{L^\tau(\Omega_h^*)}^\tau + \sum_{T \in \mathcal{T}: d(T) \leq 2t/\hat{c}} \|S_1 - S_2\|_{L^\tau(T)}^\tau \\ &\quad + \sum_{A \in \mathcal{A}: d(A) > t} \|S_1 - S_2\|_{L^\tau(A_t)}^\tau + \sum_{A \in \mathcal{A}: d(A) \leq t} \|S_1 - S_2\|_{L^\tau(A)}^\tau. \end{aligned}$$

For the first sum above, we have just as in (3.27),

$$\sum_{T \in \mathcal{T}: d(T) \leq 2t/\hat{c}} \|S_1 - S_2\|_{L^\tau(T)}^\tau \leq \sum_{T \in \mathcal{T}: d(T) \leq 2t/\hat{c}} cd(T)^{\tau s} \|S_1 - S_2\|_{L^p(T)}^\tau.$$

We estimate the other two sums as in (3.32) and (3.33). We obtain

$$\begin{aligned} & \|\Delta_h S_1\|_{L^\tau(\Omega_h^*)}^\tau \leq \|\Delta_h S_2\|_{L^\tau(\Omega_h^*)}^\tau + \sum_{T \in \mathcal{T}: d(T) \leq 2t/\hat{c}} cd(T)^{\tau s} \|S_1 - S_2\|_{L^p(T)}^\tau \\ & + \sum_{A \in \mathcal{A}: d(A) > t} ctd(A)^{\tau s-1} \|S_1 - S_2\|_{L^p(A)}^\tau + \sum_{A \in \mathcal{A}: d(A) \leq t} cd(A)^{\tau s} \|S_1 - S_2\|_{L^p(A)}^\tau. \end{aligned}$$

It is an important observation that each trapezoid $T \in \mathcal{T}$ with $d(T) > 2t/\hat{c}$ may share trapezoids $T' \in \mathfrak{T}_T^t$ and sets $A \in \mathcal{A}_T^t$ with only finitely many trapezoids with the same properties. Also, for every such trapezoid T , we have $\#\mathcal{T}_T^t \leq c$ and $\#\mathcal{A}_T^t \leq c$ with $c > 0$ a constant depending only on the structural constants of the setting. Therefore, in the above estimates, only finitely many norms may overlap at a time. Putting all of them together, we obtain

$$\omega_1(S_1, t)_\tau^\tau \leq \omega_1(S_2, t)_\tau^\tau + Y_1 + Y_2,$$

where

$$\begin{aligned} Y_1 &= \sum_{A \in \mathcal{A}: d(A) > t} ctd(A)^{\tau s-1} \|S_1 - S_2\|_{L^p(A)}^\tau \\ &+ \sum_{A \in \mathcal{A}: d(A) > t} ct^2d(A)^{\tau s-2} \|S_1 - S_2\|_{L^p(A)}^\tau \\ &+ \sum_{A \in \mathcal{A}: d(A) \leq t} cd(A)^{\tau s} \|S_1 - S_2\|_{L^p(A)}^\tau \end{aligned}$$

and

$$\begin{aligned} Y_2 &= \sum_{T \in \mathcal{T}: d(T) > 2t/\hat{c}} ctd(T)^{\tau s-1} \|S_1 - S_2\|_{L^p(T)}^\tau \\ &+ \sum_{T \in \mathcal{T}: d(T) > 2t/\hat{c}} ct^{1+\tau s/2} d(T)^{\tau s/2-1} \|S_1 - S_2\|_{L^p(T)}^\tau \\ &+ \sum_{T \in \mathcal{T}: d(T) \leq 2t/\hat{c}} cd(T)^{\tau s} \|S_1 - S_2\|_{L^p(T)}^\tau. \end{aligned}$$

We now turn to the estimation of $|S_1|_{B_\tau^s}$. Using the above and interchanging the order of integration and summation, we get

$$|S_1|_{B_\tau^s}^\tau = \int_0^\infty t^{-s\tau-1} \omega_1(S_1, t)_\tau^\tau dt \leq |S_2|_{B_\tau^s}^\tau + Z_1 + Z_2,$$

where

$$\begin{aligned} Z_1 &= \sum_{A \in \mathcal{A}} cd(A)^{\tau s-1} \|S_1 - S_2\|_{L^p(A)}^\tau \int_0^{d(A)} t^{-\tau s} dt \\ &+ \sum_{A \in \mathcal{A}} cd(A)^{\tau s-2} \|S_1 - S_2\|_{L^p(A)}^\tau \int_0^{d(A)} t^{-\tau s+1} dt \\ &+ \sum_{A \in \mathcal{A}} cd(A)^{\tau s} \|S_1 - S_2\|_{L^p(A)}^\tau \int_{d(A)}^\infty t^{-\tau s-1} dt \end{aligned}$$

and

$$\begin{aligned} Z_2 &= \sum_{T \in \mathcal{T}} cd(T)^{\tau s-1} \|S_1 - S_2\|_{L^p(T)}^\tau \int_0^{\hat{cd}(T)/2} t^{-\tau s} dt \\ &+ \sum_{T \in \mathcal{T}} cd(T)^{\tau s/2-1} \|S_1 - S_2\|_{L^p(T)}^\tau \int_0^{\hat{cd}(T)/2} t^{-\tau s/2} dt \\ &+ \sum_{T \in \mathcal{T}} cd(T)^{\tau s} \|S_1 - S_2\|_{L^p(T)}^\tau \int_{\hat{cd}(T)/2}^\infty t^{-\tau s-1} dt. \end{aligned}$$

Observe that $-\tau s > -1$ is equivalent to $s/2 < 1/p$, which is one of the assumptions, and $-\tau s/2 > -1$ is equivalent to $s < s + 2/p$, which is obvious. Therefore, all of the above integrals are convergent, and we obtain

$$|S_1|_{B_s^\tau} \leq |S_2|_{B_s^\tau} + \sum_{A \in \mathcal{A}} c \|S_1 - S_2\|_{L^p(A)}^\tau + \sum_{T \in \mathcal{T}} c \|S_1 - S_2\|_{L^p(T)}^\tau.$$

Finally, applying Hölder’s inequality and using (3.23), we arrive at

$$\begin{aligned} |S_1|_{B_s^\tau} &\leq |S_2|_{B_s^\tau} + c (\#\mathcal{A})^{\tau(1/\tau-1/p)} \left(\sum_{A \in \mathcal{A}} \|S_1 - S_2\|_{L^p(A)}^p \right)^{\tau/p} \\ &+ c (\#\mathcal{T})^{\tau(1/\tau-1/p)} \left(\sum_{T \in \mathcal{T}} \|S_1 - S_2\|_{L^p(T)}^p \right)^{\tau/p} \\ &\leq cn^{\tau(1/\tau-1/p)} \|S_1 - S_2\|_{L^p(\Omega)}^\tau = cn^{\tau s/2} \|S_1 - S_2\|_{L^p(\Omega)}^\tau. \end{aligned}$$

This confirms estimate (3.10). The proof in the case when $\tau > 1$ is the same. □

4 Nonlinear Approximation from Smooth Splines

In this section, we focus on Bernstein estimates in nonlinear approximation in L^p , $0 < p < \infty$, from regular nonnested smooth piecewise polynomial functions in \mathbb{R}^2 .

4.1 Setting and Approximation Tool

We first elaborate on our setting and consider examples. As in Sect. 3, we consider two versions of the class of regular piecewise polynomials $\mathcal{S}(n, k)$ of degree $k - 1$ with $k \geq 2$ over n rings of maximum smoothness, depending on whether Ω is compact or $\Omega = \mathbb{R}^2$.

Case 1: Assume Ω is a compact polygonal domain in \mathbb{R}^2 that can be represented as the union of n_0 rings with disjoint interiors, see Condition 3.2. We denote by $\mathcal{S}(n, k)$ ($n \geq n_0$) the set of all piecewise polynomials S of the form

$$S = \sum_{j=1}^n P_j \mathbb{1}_{R_j}, \quad S \in C^{k-2}(\Omega), \quad P_j \in \Pi_{k-1}, \tag{4.1}$$

where R_1, \dots, R_n are rings in the sense of Definition 3.1 with disjoint interiors such that $\Omega = \cup_{j=1}^n R_j$. Recall that Π_{k-1} stands for the set of all polynomials of degree $\leq k - 1$ in two variables and $S \in C^{k-2}(\Omega)$ means that all partial derivatives $\partial^\alpha S \in C(\Omega)$, $|\alpha| \leq k - 2$.

Case 2: $\Omega = \mathbb{R}^2$. In this case, we denote by $\mathcal{S}(n, k)$ the set of all piecewise polynomials S of degree $k - 1$ on \mathbb{R}^2 of the form (4.1), where R_1, \dots, R_n are rings with disjoint interiors such that the support $\Lambda = \cup_{j=1}^n R_j$ of S is a ring in the sense of Definition 3.1.

We denote by $S_n^k(f)_p$ the best approximation of $f \in L^p(\Omega)$ from $\mathcal{S}(n, k)$ in $L^p(\Omega)$, $0 < p < \infty$; i.e.,

$$S_n^k(f)_p := \inf_{S \in \mathcal{S}(n, k)} \|f - S\|_{L^p}. \tag{4.2}$$

Remark Observe that in our setting, the splines are of maximum smoothness, and this is critical for our development. As will be shown in Example 4.4 below in the nonnested case our Bernstein type inequality is not valid in the case when the smoothness of the splines is not maximal.

We next consider several scenarios for constructing regular piecewise polynomials of maximum smoothness:

1. *Piecewise linear functions induced by nested triangulations.* Suppose that \mathcal{T}_0 is an initial subdivision of Ω into triangles that obey the minimum angle condition and is with no hanging vertices in the interior of Ω . In the case of $\Omega = \mathbb{R}^2$, we assume for simplicity that the triangles $\Delta \in \mathcal{T}_0$ are of similar areas; i.e., $c_1 \leq |\Delta_1|/|\Delta_2| \leq c_2$ for all $\Delta_1, \Delta_2 \in \mathcal{T}_0$. Next we subdivide each triangle $\Delta \in \mathcal{T}_0$ into 4 triangles by introducing the midpoints on the sides of Δ . The result is a triangulation \mathcal{T}_1 of Ω . In the same way, we define the triangulations $\mathcal{T}_2, \mathcal{T}_3$, etc. Each triangulation \mathcal{T}_j supports Courant hat functions (linear finite elements) φ_θ , each of them supported on the union θ of all triangles from \mathcal{T}_j that have a common vertex, say, v . Thus $\varphi_\theta(v) = 1$, φ_θ takes values zero at all other vertices of triangles from \mathcal{T}_j , and φ_θ is continuous and piecewise linear over the triangles from \mathcal{T}_j . Clearly, each piecewise linear function over the triangles from \mathcal{T}_j can be represented as a linear combination of Courant hat functions like these.

Denote by Θ_j the set of all supports θ of Courant elements supported by \mathcal{T}_j and set $\Theta := \cup_{j \geq 0} \Theta_j$. Consider the nonlinear set \mathbb{S}_n of all piecewise linear functions S of the form

$$S = \sum_{\theta \in \mathcal{M}_n} c_\theta \varphi_\theta,$$

where $\mathcal{M}_n \subset \Theta$ and $\#\mathcal{M}_n \leq n$; the elements $\theta \in \mathcal{M}_n$ may come from different levels and locations. It is not hard to see that $\mathbb{S}_n \subset S(cn, 2)$, see [6].

2. *General piecewise linear functions.* More generally, one can consider piecewise linear functions S of the form

$$S = \sum_{\theta \in \mathcal{M}_n} c_\theta \varphi_\theta,$$

where $\{\varphi_\theta\}$ are Courant hat functions as above, $\#\mathcal{M}_n \leq n$, and \mathcal{M}_n consists of cells θ as above that are not necessarily induced by a hierarchical collection of triangulations of Ω ; however, there exists an underlying subdivision of Ω into rings obeying the conditions from Sect. 3.1.

3. *Piecewise quadratic or cubic splines.* The C^1 quadratic box-splines on the four-directional mesh (the so-called ‘‘Powell–Zwart finite elements’’) and the piecewise cubics in \mathbb{R}^2 or on a rectangular domain, endowed with the Powell–Sabin triangulation generated by a uniform 6-direction mesh, provide examples of quadratic and cubic splines of maximum smoothness.

Other examples are to be identified or developed.

Splines with defect. To make the difference between approximation from nonnested and nested splines more transparent and for future references, we now introduce the splines with arbitrary smoothness. Given a set $\Omega \subset \mathbb{R}^2$ with polygonal boundary or $\Omega := \mathbb{R}^2$, $k \geq 2$, and $0 \leq r \leq k - 1$, we denote by $\mathcal{S}(n, k, r)$ ($n \geq n_0$) the set of all piecewise polynomials S of the form

$$S = \sum_{j=1}^n P_j \mathbb{1}_{R_j}, \quad S \in C^{r-1}(\Omega), \quad P_j \in \Pi_{k-1}, \tag{4.3}$$

where R_1, \dots, R_n are rings with disjoint interiors such that $\Omega = \cup_{j=1}^n R_j$. We set

$$S_n^{k,r}(f)_p := \inf_{S \in \mathcal{S}(n,k,r)} \|f - S\|_{L^p}. \tag{4.4}$$

4.2 Jackson Estimate

Jackson estimates in spline approximation are relatively easy to prove. Such estimates (also in anisotropic settings) are established in [3, 6]. For example, the Jackson estimate we need in the case of approximation from piecewise linear functions ($k = 2$) follows from [6, Theorem 3.6] and takes the form:

Theorem 4.1 *Let $0 < p < \infty$, $s > 0$, and $1/\tau = s/2 + 1/p$. Assume $\Omega = \mathbb{R}^2$ or $\Omega \subset \mathbb{R}^2$ is a compact set with polygonal boundary and an initial triangulation consisting of $\leq n_0$ triangles with no hanging interior vertices and obeying the minimum angle condition. Then for any $f \in B_{\tau}^{s,2}$, we have $f \in L^p(\Omega)$ and*

$$S_n^2(f)_p \leq cn^{-s/2} |f|_{B_{\tau}^{s,2}}, \quad n \geq n_0. \tag{4.5}$$

Consequently, for any $f \in L^p(\Omega)$,

$$S_n^2(f)_p \leq cK(f, n^{-s/2}), \quad n \geq n_0. \tag{4.6}$$

Here $K(f, t) = K(f, t; L^p, B_{\tau}^s)$ is the K -functional defined in (3.6) and $c > 0$ is a constant depending only on s, p , and the structural constants of the setting.

Similar Jackson and direct estimates for nonlinear approximation from splines of degrees ≥ 2 and of maximum smoothness do not follow automatically from the results in [3], the reason being the fact that the basis functions for splines of degree 2 and 3 that we are familiar with are not stable. The stability is required in [3]. The problem for establishing Jackson estimates for approximation from splines of degree ≥ 2 of maximum smoothness remains open.

4.3 Bernstein Estimate in the Nonnested Case

We come now to one of the main results of this article. Here we operate in the setting described above in Sect. 4.1.

Theorem 4.2 *Let $0 < p < \infty, k \geq 1, 0 < s/2 < k - 1 + 1/p$, and $1/\tau = s/2 + 1/p$. Then for any $S_1, S_2 \in \mathcal{S}(n, k), n \geq n_0$, we have*

$$|S_1|_{B_{\tau}^{s,k}} \leq |S_2|_{B_{\tau}^{s,k}} + cn^{s/2} \|S_1 - S_2\|_{L^p} \quad \text{if } \tau \geq 1, \quad \text{and} \tag{4.7}$$

$$|S_1|_{B_{\tau}^{s,k}}^{\tau} \leq |S_2|_{B_{\tau}^{s,k}}^{\tau} + cn^{\tau s/2} \|S_1 - S_2\|_{L^p}^{\tau} \quad \text{if } \tau < 1, \tag{4.8}$$

where the constant $c > 0$ depends only on s, p, k , and the structural constants of the setting; n_0 is from Condition 3.2.

An immediate consequence of this theorem is the inverse estimate given in

Corollary 4.3 *Let $0 < p < \infty, k \geq 1, 0 < s/2 < k - 1 + 1/p$, and $1/\tau = s/2 + 1/p$. Set $\lambda := \min\{\tau, 1\}$. Then for any $f \in L^p(\Omega)$, we have*

$$K(f, n^{-s/2}) \leq cn^{-s/2} \left(\sum_{\ell=n_0}^n \frac{1}{\ell} \left[\ell^{s/2} S_{\ell}^k(f)_p \right]^{\lambda} + \|f\|_p^{\lambda} \right)^{1/\lambda}, \quad n \geq n_0. \tag{4.9}$$

Here $K(f, t) = K(f, t; L^p, B_{\tau}^s)$ is the K -functional defined just as in (3.6), and $c > 0$ is a constant depending only on s, p, k , and the structural constants of the setting.

The proof of this corollary is just a repetition of the proof of Theorem 3.6. We omit it.

In turn, estimates (4.6) and (4.9) imply a characterization of the approximation spaces associated with nonlinear nonnested piecewise linear approximation, see (1.7).

The proof of Theorem 4.2 relies on the idea we used in the proof of Theorem 3.4. However, there is an important complication to overcome. The fact that many rings with relatively small supports can be located next to a large ring is a major obstacle in implementing this idea in the case of smooth splines. An additional construction is needed. To make the proof more accessible, we shall proceed in two steps. We first develop the needed additional construction and implement it in Sect. 4.4 to prove the respective Bernstein estimate in the nested case, and then we present the proof of Theorem 4.2 in Sect. 4.5.

Before we proceed with the proofs of the Bernstein estimates, we show in the next example that the assumption that in our setting the splines are of maximum smoothness is essential.

Example 4.4 We now show that estimates (4.7–4.8) fail without the assumption that $S_1, S_2 \in C^{k-2}(\Omega)$ (i.e., both splines have maximum smoothness). We shall only consider the case when $k = 2$ and $\tau \leq 1$. Let $\Omega = [-1, 1] \times [0, 1]$ and $0 < \varepsilon < 1/4$. Set

$$S_1(x) := x_1 \mathbb{1}_{[0,1]^2}(x), \quad S_2(x) := x_1 \mathbb{1}_{[\varepsilon,1] \times [0,1]}(x), \quad x = (x_1, x_2).$$

Clearly, S_1 is continuous on Ω , while S_2 is discontinuous along $x_1 = \varepsilon$. A straightforward calculation shows that

$$\omega_2(S_1, t)_\tau^\tau = \frac{2t^{\tau+1}}{\tau + 1} \quad \text{and} \quad \omega_2(S_2, t)_\tau^\tau = \int_{-t}^t |w + \varepsilon|^\tau dw \quad \text{for } 0 \leq t \leq 1/4. \tag{4.10}$$

Further,

$$\int_{-t}^t |w + \varepsilon|^\tau dw = \frac{1}{\tau + 1} \left[(t + \varepsilon)^{\tau+1} + \text{sign}(t - \varepsilon) |t - \varepsilon|^{\tau+1} \right]. \tag{4.11}$$

On the other hand, obviously $\omega_2(S_1 - S_2, t)_\tau^\tau \leq 4\|S_1 - S_2\|_{L^\tau}^\tau \leq 4\varepsilon^{\tau+1}$, yielding

$$\omega_2(S_2, t)_\tau^\tau \geq \omega_2(S_1, t)_\tau^\tau - 4\varepsilon^{\tau+1}. \tag{4.12}$$

We shall use this estimate for $t > 1/4$. From (2.1) and (4.10–4.12), we obtain

$$\begin{aligned} |S_2|_{B_\tau^{s,2}}^\tau - |S_1|_{B_\tau^{s,2}}^\tau &\geq \frac{1}{\tau + 1} \left[\int_0^\varepsilon t^{-s\tau-1} [(t + \varepsilon)^{\tau+1} - (\varepsilon - t)^{\tau+1} - 2t^{\tau+1}] dt \right. \\ &\quad \left. + \int_\varepsilon^{1/4} t^{-s\tau-1} [(\varepsilon + t)^{\tau+1} + (t - \varepsilon)^{\tau+1} - 2t^{\tau+1}] dt \right] - 4\varepsilon^{\tau+1} \int_{1/4}^\infty t^{-s\tau-1} dt \\ &=: I_1 + I_2 - (4^{s\tau+1}/s\tau)\varepsilon^{\tau+1}. \end{aligned}$$

Substituting $t = \varepsilon u$ in I_1 and I_2 , we get

$$I_1 + I_2 = \frac{\varepsilon^{\tau-s\tau+1}}{\tau + 1} \left[\int_0^1 u^{-s\tau-1} \phi_1(u) du + \int_1^{1/4\varepsilon} u^{-s\tau-1} \phi_2(u) du \right],$$

where

$$\phi_1(u) = (1 + u)^{\tau+1} - (1 - u)^{\tau+1} - 2u^{\tau+1}$$

and

$$\phi_2(u) = (1 + u)^{\tau+1} + (u - 1)^{\tau+1} - 2u^{\tau+1}.$$

We clearly have $\phi_1 \geq 0$ on $[0, 1]$ and $\phi_2 \geq 0$ on $[1, \infty)$. Therefore,

$$|S_2|_{B_\tau^{s,2}}^\tau - |S_1|_{B_\tau^{s,2}}^\tau \geq c_1 \varepsilon^{\tau-s\tau+1} - c_0 \varepsilon^{\tau+1} = \varepsilon^{\tau-s\tau+1} (c_1 - c_0 \varepsilon^{s\tau}),$$

where

$$c_1 := \frac{1}{\tau + 1} \int_0^1 t^{-s\tau-1} \phi_1(u) du > 0 \quad \text{and} \quad c_0 := 4^{s\tau+1} / s\tau.$$

By taking ε sufficiently small, we get

$$|S_2|_{B_\tau^{s,2}}^\tau - |S_1|_{B_\tau^{s,2}}^\tau \geq (c_1/2) \varepsilon^{\tau-s\tau+1}. \tag{4.13}$$

Evidently, $\|S_2 - S_1\|_{L^p} \leq \varepsilon^{1+1/p}$. This estimate coupled with (4.13) implies

$$\frac{|S_2|_{B_\tau^{s,2}}^\tau - |S_1|_{B_\tau^{s,2}}^\tau}{\|S_2 - S_1\|_{L^p}^\tau} \geq (c_1/2) \varepsilon^{1-s\tau-\tau/p} = (c_1/2) \varepsilon^{-s\tau/2}.$$

Therefore, since $\varepsilon^{-s\tau/2} \rightarrow \infty$ as $\varepsilon \rightarrow 0$, estimate (4.8) cannot hold.

4.4 Additional Subdivision and Bernstein Estimate in the Nested Case

As already indicated above, the idea of the proof of the Bernstein estimate from Theorem 3.4 is insufficient for the proof of the Bernstein estimate for approximation from smooth splines (Theorem 4.2). In the case of smooth splines, we hit a snag when “small” rings are located next to “large” rings. To overcome this obstacle, we next introduce an additional subdivision of the underlying rings. As an application of this construction and for comparison, we prove the following Bernstein estimate, which yields an inverse estimate, in the case of nested spline approximation.

Theorem 4.5 *Let $0 < p < \infty$, $k \geq 2$, $0 \leq r \leq k - 1$, $0 < s/2 < r + 1/p$, and $1/\tau = s/2 + 1/p$. Then for any $S \in \mathcal{S}(n, k, r)$, $n \geq n_0$, we have*

$$|S|_{B_\tau^{s,k}} \leq cn^{s/2} \|S\|_{L^p}, \tag{4.14}$$

where the constant $c > 0$ depends only on s, p, k, r , and the structural constant of our setting.

Additional subdivision of Ω . We subdivide Ω in two steps.

Subdivision of all rings $R \in \mathcal{R}_n$ into nested hierarchies of rings.

Lemma 4.6 *There exists a subdivision of every ring $R \in \mathcal{R}_n$ into a nested multilevel collection of rings*

$$\mathcal{K}^R = \bigcup_{m=m_R}^\infty \mathcal{K}_m^R$$

with the following properties, where we use the abbreviated notation $\mathcal{K}_m := \mathcal{K}_m^R$:

- (a) Every level \mathcal{K}_m defines a partition of R into rings with disjoint interiors such that $R = \bigcup_{K \in \mathcal{K}_m} K$.
- (b) The levels $\{\mathcal{K}_m\}_{m \geq m_R}$ are nested; i.e., \mathcal{K}_{m+1} is a refinement of \mathcal{K}_m , and each $K \in \mathcal{K}_m$ has at least 4 and at most M children in \mathcal{K}_{m+1} , where $M \geq 4$ is a constant.
- (c) $|R| \leq c_1 |K|$ for all $K \in \mathcal{K}_{m_R}$.
- (d) We have

$$c_2^{-1} 4^{-m} \leq |K| \leq c_2 4^{-m}, \quad \forall K \in \mathcal{K}_m, \quad \forall m \geq m_R.$$

As a consequence, we have $c_3^{-1} 4^{-m_R} \leq |R| \leq c_3 4^{-m_R}$ and

$$c_4^{-1} 2^{-m} \leq d(K) \leq c_4 2^{-m}, \quad \forall K \in \mathcal{K}_m, \quad \forall m \geq m_R.$$

- (e) All rings $K \in \mathcal{K}^R$ are rings without a hole, except for finitely many of them in the case when $R = Q_1 \setminus Q_2$ and Q_2 is small relative to Q_1 . Then the rings with a hole form a chain $R \supset K_1 \supset K_2 \supset \dots \supset K_\ell \supset Q_2$. All sets $K \in \mathcal{K}^R$ are rings in the sense of Definition 3.1 with structural constants (parameters) N_0^* , c_0^* , and β^* . These and the constants M and $c_1, c_2, c_3, c_4 > 0$ from above depend only on the initial structural constants N_0, c_0 , and β .

Proof Observe first that if we are in a setting as the one described in Scenario 1 from Sect. 4.1, then the needed subdivision is given by the hierarchy of triangulations described there.

In the general case, let $R = Q_1 \setminus Q_2$ be a ring in the sense of Definition 3.1, and assume that $Q_2 \neq \emptyset$. We subdivide the polygonal convex set Q_1 into subrings by connecting the center of eccentricity of Q_1 with, say, 6 points from the boundary ∂R of R , preferably end points of segments on the boundary, so that the minimum angle condition is obeyed. After that we subdivide the resulting rings using midpoints and connecting them with segments. Necessary adjustments are made around Q_2 , depending on the size and location of Q_2 . □

Subdivision of all rings from \mathcal{R}_n into subrings with disjoint interiors. We first pick up all rings from each \mathcal{K}^R , $R \in \mathcal{R}_n$, see Lemma 4.6, that are needed to handle situations where many small rings are located next to a large ring.

We shall only need the rings in \mathcal{K}^R that intersect the boundary ∂R of R . Denote the set of all such rings by Γ^R , and set $\Gamma_m^R := \Gamma^R \cap \mathcal{K}_m^R$. We shall make use of the tree structure in Γ^R . More precisely, we shall use the parent-child relation in Γ^R induced by the inclusion relation: Each ring $K \in \Gamma_m^R$ has (contains) at least 1 and at most M children in Γ_{m+1}^R and has a single parent in Γ_{m-1}^R or no parent.

We now construct a set Λ^R of rings from Γ^R which will help prevent situations where a ring may have many small neighbors.

Given $R \in \mathcal{R}_n$, we denote by \mathcal{R}_n^R the set of all rings $\tilde{R} \in \mathcal{R}_n$, $\tilde{R} \neq R$, such that $\tilde{R} \cap R \neq \emptyset$ and $d(\tilde{R}) \leq d(R)$. These are all rings from \mathcal{R}_n that are small relative to R and intersect R (are neighbors of R).

It will be convenient to introduce the following somewhat geometric terminology: We say that a ring $K \in \Gamma^R$ can see $\tilde{R} \in \mathcal{R}_n^R$ or that \tilde{R} is in the range of K if $d(K) \geq d(\tilde{R})$ and $K \cap \tilde{R} \neq \emptyset$.

We now construct Λ^R by applying the following

Rule: We place $K \in \Gamma^R$ in Λ^R if K can see some (at least one) rings from \mathcal{R}_n^R but neither of the children of K in Γ^R can see all of them.

We now extend Λ^R to $\tilde{\Lambda}^R$ by adding to Λ^R all same level neighbors of all $K \in \Lambda^R$; i.e., if $K \in \Lambda^R$ and $K \in \Gamma_m^R$, then we add to Λ^R each $K' \in \Gamma_m^R$ such that $K' \cap K \neq \emptyset$.

The next step is to construct a subdivision of each $R \in \mathcal{R}_n$ into rings by using $\tilde{\Lambda}^R$. We fix $R \in \mathcal{R}_n$ and shall suppress the superscript R for the new sets that will be introduced next and depend on R .

Let $\tilde{\Gamma} \subset \Gamma^R$ be the minimal subtree of Γ^R that contains $\tilde{\Lambda}^R$; i.e., $\tilde{\Gamma}$ is the set of all $K \in \Gamma^R$ such that $K \supset K'$ for some $K' \in \tilde{\Lambda}^R$. We denote by $\tilde{\Gamma}_b$ the set of all *branching rings* in $\tilde{\Gamma}$ (rings with more than one child in $\tilde{\Gamma}$) and by $\tilde{\Gamma}'_b$ the set of all *children in $\tilde{\Gamma}$ of branching rings* (each of them may or may not belong to $\tilde{\Gamma}$). Furthermore, we let $\tilde{\Gamma}_\ell$ denote the set of all *leaves* in $\tilde{\Gamma}$ (rings in $\tilde{\Gamma}$ containing no other rings from $\tilde{\Gamma}$).

Evidently, $\tilde{\Gamma}_\ell \subset \tilde{\Lambda}^R$. However, rings from $\tilde{\Gamma}_b$ and $\tilde{\Gamma}'_b$ may or may not belong to $\tilde{\Lambda}^R$. We extend $\tilde{\Lambda}^R$ to $\tilde{\tilde{\Lambda}}^R := \tilde{\Lambda}^R \cup \tilde{\Gamma}_b \cup \tilde{\Gamma}'_b$. In addition, we add to $\tilde{\tilde{\Lambda}}^R$ all rings from $\mathcal{K}_{m_R}^R$, if they are not there yet.

It is readily seen that each ring $\tilde{R} \in \mathcal{R}_n^R$ can be in the range of only finitely many $K \in \tilde{\Gamma}_\ell$ and each ring $\tilde{R} \in \mathcal{R}_n$ may have only finitely many neighbors $R \in \mathcal{R}_n$ such that $d(R) \geq d(\tilde{R})$. Therefore,

$$\sum_{R \in \mathcal{R}_n} \#\tilde{\Gamma}_\ell^R \leq cn.$$

Obviously, $\#\tilde{\Gamma}_b \leq \#\tilde{\Gamma}_\ell, \#\tilde{\Gamma}'_b \leq M\#\tilde{\Gamma}_b \leq M\#\tilde{\Gamma}_\ell$, implying $\#\tilde{\tilde{\Lambda}}^R \leq \#\tilde{\Gamma}_\ell + \#\tilde{\Gamma}_b \leq c\#\tilde{\Gamma}_\ell$, and hence $\#\tilde{\tilde{\Lambda}}^R \leq c'\#\tilde{\Gamma}_\ell$. Putting these estimates together implies

$$\sum_{R \in \mathcal{R}_n} \#\tilde{\tilde{\Lambda}}^R \leq cn.$$

Observe that, with the exception of all branching rings in $\tilde{\Lambda}^R$, by construction every other ring $K \in \tilde{\Lambda}^R$ is either a leaf, and hence contains no other rings from $\tilde{\Lambda}^R$, or contains only one ring $K' \in \tilde{\Lambda}^R$ of minimum level; i.e., K has one descendent K' in $\tilde{\Lambda}^R$.

We now make the *final step* in our construction: We denote by \mathcal{F}^R the set of all rings from $\tilde{\Gamma}_\ell^R$ along with all new rings of the form $\overline{K \setminus K'}$, where $K \in \tilde{\Gamma}_b^R$, $K' \in \tilde{\Lambda}^R$, $K' \subset K$ and K' is of minimum level with these properties. Set $\mathcal{F} := \cup_{R \in \mathcal{R}_n} \mathcal{F}^R$.

The purpose of the above construction becomes clear from the following

Lemma 4.7 *The set \mathcal{F} consists of rings in the sense of Definition 3.1 with parameters depending only on the structural constants N_0 , c_0 , and β . Also, for any $R \in \mathcal{R}_n$, the rings in \mathcal{F}^R have disjoint interiors, $R = \cup_{K \in \mathcal{F}^R} K$, and $\#\mathcal{F}^R \leq c\#\tilde{\Lambda}^R$. Hence,*

$$\Omega = \bigcup_{R \in \mathcal{R}_n} \bigcup_{K \in \mathcal{F}^R} K \quad \text{and} \quad \sum_{R \in \mathcal{R}_n} \#\mathcal{F}^R \leq cn.$$

Most importantly, each ring $K \in \mathcal{F}$ has only finitely many neighbors in \mathcal{F} ; that is, there exists a constant N_1 such that for any $K \in \mathcal{F}$ there are at most N_1 rings in \mathcal{F} intersecting K .

To prove the most important property of the set of rings \mathcal{F} , namely, that each ring $K \in \mathcal{F}$ has only finitely many neighbors in \mathcal{F} , we shall need the following technical

Lemma 4.8 *Suppose $K \supset K_1 \supset K_2$, $K \in \Gamma^R$, $K_1, K_2 \in \tilde{\Lambda}^R$, and both K_1 and K_2 share parts of an edge E of K located in the interior of R . Then there exists $K^* \in \tilde{\Lambda}^R$ such that $K^* \cap K^\circ = \emptyset$, $K^* \cap E \neq \emptyset$, and K^* is either a neighbor of K_1 or K_2 , or K^* is a neighbor of the parent of K_1 in Γ^R .*

Proof If $K_1 \in \Lambda^R$, then by construction all same level neighbors of K_1 belong to $\tilde{\Lambda}^R$ and hence the one that shares the edge of K_1 contained in E will be in $\tilde{\Lambda}^R$. We denote this ring by K^* , and apparently it has the claimed properties. By the same token, if $K_2 \in \Lambda^R$, then one of its neighbors will do the job.

Suppose $K_1, K_2 \in \tilde{\Lambda}^R \setminus \Lambda^R$. Then K_1 has a neighbor, say, \hat{K}_1 that belongs to Λ^R and \hat{K}_1 is at the level of K_1 . If \hat{K}_1 has an edge contained in E , then $K^* := \hat{K}_1$ has the claimed property. Similarly, K_2 has a neighbor $\hat{K}_2 \in \Lambda^R$ at the level of K_2 . If \hat{K}_2 has an edge contained in E , then $K^* := \hat{K}_2$ will do the job.

Assume that neither of the above is true. Then since $K_1, \hat{K}_1 \in \Gamma^R$, they must have the same parent in Γ^R that has an edge contained in E . Denote this common parent by K^\sharp . For the same reason, $K_2, \hat{K}_2 \in \Gamma^R$ have a common parent, say, $K^{\sharp\sharp}$ in Γ^R . Clearly, K^\sharp and $K^{\sharp\sharp}$ have some edges contained in E . Also, $\hat{K}_1 \subset K^\sharp$, $\hat{K}_2 \subset K^{\sharp\sharp}$, and $\hat{K}_1^\circ \cap \hat{K}_2^\circ = \emptyset$.

We claim that K^\sharp belongs to Λ^R . Indeed, the rings from \mathcal{R}_n that are in the range of \hat{K}_1 are also in the range of K^\sharp . Also, the rings from \mathcal{R}_n that are in the range of \hat{K}_2 are also in the range of $K^{\sharp\sharp}$. However, obviously neither of the children of K^\sharp can have the range of K^\sharp . Therefore, K^\sharp belongs to Λ^R . Now, just as above, we conclude that one of the neighbors of K^\sharp has the claimed property. □

Proof of Lemma 4.7 All properties of the newly constructed set of rings \mathcal{F} , given in Lemma 4.7, but the last one follow readily from their construction.

We now show that each ring $K \in \mathcal{F}$ has only finitely many neighbors in \mathcal{F} . Indeed, by the construction any $K \in \mathcal{F}^R$, $R \in \mathcal{R}_n$, has only finitely many neighbors that do not belong to \mathcal{F}^R . Thus, it remains to show that it cannot happen that there exist rings $K_1 \subset K_2 \subset \dots \subset K_J$, $K_j \in \tilde{\Lambda}^R$, with J uncontrollably large that have edges contained in an edge of a single ring $K \in \tilde{\Lambda}^R$ whose interior does not intersect K_j , $j = 1, \dots, J$. But this assertion readily follows by Lemma 4.8. \square

The following lemma will be instrumental in the proof of Theorem 4.5.

Lemma 4.9 *Assume $0 < p, q \leq \infty$, $k \geq 1$, $r \geq 0$, and $v \in \mathbb{R}^2$ with $|v| = 1$. Let the sets $G, H \subset \mathbb{R}^2$ be measurable, $G \subset H$, and such that there exist balls B_1, B_2, B_3, B_4 , $B_j = B(x_j, r_j)$, with the properties: $B_2 \subset G \subset B_1$, $r_1 \leq c^b r_2$, and $B_4 \subset H \subset B_3$, $r_3 \leq c^b r_4$, where $c^b \geq 1$ is a constant. Then for any $P \in \Pi_{k-1}$,*

$$\|P\|_{L^p(G)} \leq c|G|^{1/p-1/q}\|P\|_{L^q(G)}, \tag{4.15}$$

$$\|D_v^r P\|_{L^p(G)} \leq cd(G)^{-r}\|P\|_{L^p(G)}, \tag{4.16}$$

and

$$\|P\|_{L^p(G)} \leq c(|G|/|H|)^{1/p}\|P\|_{L^p(H)}, \tag{4.17}$$

where $c > 0$ is a constant depending on p, q, k, r, c^b , and the parameters N_0, c_0 , and β from Definition 3.1. Here $D_v^r S$ is the r th directional derivative of S in the direction of v .

Furthermore, inequality (4.17) holds with Q and H replaced by their images $L(G)$ and $L(H)$, where L is a nonsingular linear transformation of \mathbb{R}^2 .

Proof Inequality (4.15) holds whenever $B_2 = B(0, 1)$ and $B_1 = B(0, c_\diamond)$ with $c_\diamond =$ constant by the fact that any two (quasi)norms on Π_{k-1} are equivalent. This implies that (4.15) is valid in the case when $B_2 = B(0, 1)$ and $B_2 \subset B_1$, where $B_1 = B(x_2, c_\diamond/2)$. Then (4.15), in general, follows by rescaling. Inequality (4.17) is obvious when $p = \infty$. In general, it follows from the case $p = \infty$ and application of (4.15) to G with p and $q = \infty$ and to H with $p = \infty, q = p$. Inequality (4.16) is an easy consequence of the Markov inequality for univariate polynomials whenever G is a square. Then in general it follows by inscribing B_1 in a smallest possible cube and then applying it for the cube and using (4.17). The last claim in the lemma is obvious. \square

Proof of Theorem 4.5 With the preparations from above, we are ready to carry out this proof. We shall only consider the case when $\Omega \subset \mathbb{R}^2$ is a compact polygonal domain. Let $S \in \mathcal{S}(n, k, r)$, and suppose S is represented as in (4.1); that is,

$$S = \sum_{R \in \mathcal{R}_n} P_R \mathbb{1}_R, \quad S \in C^{r-1}(\Omega), \quad P_R \in \Pi_{k-1}, \tag{4.18}$$

where \mathcal{R}_n is a collection of $\leq n$ rings with disjoint interiors such that $\Omega = \cup_{R \in \mathcal{R}_n} R$. Let \mathcal{F} be the set of rings constructed above starting with the rings from \mathcal{R}_n . Then from

(4.18) and because \mathcal{F} is a refinement of \mathcal{R}_n , it follows that S can be represented in the form

$$S = \sum_{K \in \mathcal{F}} P_K \mathbb{1}_K, \quad S \in C^{r-1}(\Omega), \quad P_K \in \Pi_{k-1}. \tag{4.19}$$

Recall that \mathcal{F} is the collection of at most cn rings with disjoint interiors such that $\Omega = \cup_{K \in \mathcal{F}} K$ (see Lemma 4.7).

We next introduce some convenient notation. For any ring $K \in \mathcal{F}$, we denote by \mathcal{N}_K the set of all rings $K' \in \mathcal{F}$ such that $K \cap K' \neq \emptyset$; \mathcal{E}_K will denote the set of all segments (edges) from the boundary ∂K of K ; and \mathcal{V}_K will be the set of all vertices of the polygonal curve ∂K (end points of edges from \mathcal{E}_K).

The fact that \mathcal{F} consists of rings in the sense of Definition 3.1 implies the following

Property 4.10 There exists a constant $0 < \check{c} < 1$ such that if $E = [v_1, v_2]$ is an edge shared by two rings $K, K' \in \mathcal{F}$, then for any $x \in E$ with $|x - v_j| \geq \rho, j = 1, 2$ for some $\rho > 0$, we have $B(x, \check{c}\rho) \subset K \cup K'$.

Fix $t > 0$. For each ring $K \in \mathcal{F}$, we define

$$K_t := \{x \in K : \text{dist}(x, \partial K) \leq kt\}.$$

Write $\Omega_t := \cup_{K \in \mathcal{F}} K_t$.

Let $h \in \mathbb{R}^2$ with norm $|h| \leq t$, and set $v := |h|^{-1}h$. For S is a polynomial of degree $\leq k - 1$ on each $K \in \mathcal{F}$, we have $\Delta_h^k S(x) = 0$ for $x \in \cup_{K \in \mathcal{F}} K \setminus K_t$. Therefore,

$$\|\Delta_h^k S\|_{L^r(\Omega)} = \|\Delta_h^k S\|_{L^r(\Omega_t)}.$$

Let $K \in \mathcal{F}$, and assume $d(K) > 2kt/\check{c}$ with $0 < \check{c} < 1$ the constant from Property 4.10. Define $\mathcal{N}_K^t := \{K' \in \mathcal{N}_K : d(K) > 2kt/\check{c}\}$, $B_v := B(v, 2kt/\check{c})$, $v \in \mathcal{V}_K$, and

$$\mathfrak{N}_K^t := \{K' \in \mathcal{F} : d(K') > 2kt/\check{c} \text{ and } K' \cap (K + B(0, 2kt/\check{c})) \neq \emptyset\}.$$

Observe that because $d(K) > 2kt/\check{c}$, the number of rings in \mathfrak{N}_K^t is uniformly bounded.

Let $x \in \Omega_t$ be such that $[x, x + kh] \cap K \neq \emptyset$. Two cases are to be considered here.

(a) Let $[x, x + kh] \not\subset \cup_{v \in \mathcal{V}_K} B_v$. Then $[x, x + kh]$ intersects some edge $E \in \mathcal{E}_K$ such that $\ell(E) \geq 2kt/\check{c}$, and $[x, x + kh]$ cannot intersect another edge $E' \in \mathcal{E}_K$ with this property or an edge $E' \in \mathcal{E}_K$ with $\ell(E') < 2kt/\check{c}$.

Suppose that the edge $E =: [v_1, v_2]$ is shared with $K' \in \mathcal{F}$ and $y := E \cap [x, x + kh]$. Evidently, $|y - v_j| > kt/\check{c}, j = 1, 2$, and in light of Property 4.10, we have $[x, x + kh] \subset B(y, kt) \subset K \cup K'$. Clearly,

$$|\Delta_h^k S(x)| \leq ct^r \|D_v^r S\|_{L^\infty([x, x+kh])} \leq ct^r \|D_v^r S\|_{L^\infty(K)} + ct^r \|D_v^r S\|_{L^\infty(K')}. \tag{4.20}$$

(b) Let $[x, x + kh] \subset \cup_{v \in \mathcal{V}_K} B_v$. Then we estimate $|\Delta_h^k S(x)|$ trivially:

$$|\Delta_h^k S(x)| \leq 2^k \sum_{\ell=0}^k |S(x + \ell h)|. \tag{4.21}$$

Using (4.20–4.21), we obtain

$$\begin{aligned} \|\Delta_h^k S\|_{L^\tau(K_t)}^\tau &\leq c \sum_{K' \in \mathcal{N}_K^t} td(K')t^{r\tau} \|D_v^r S\|_{L^\infty(K')}^\tau \\ &+ c \sum_{K' \in \mathfrak{N}_K^t} \sum_{v \in \mathcal{V}_K} \|S\|_{L^\tau(B_v \cap K')}^\tau + c \sum_{K'' \in \mathcal{F}: d(K'') \leq 2kt/\check{c}} \|S\|_{L^\tau(K'' \cap (K + [0, kh]))}^\tau. \end{aligned} \tag{4.22}$$

Note that the number of rings $K' \in \mathfrak{N}_K^t$ such that $K' \cap B_v \neq \emptyset$ for some $v \in \mathcal{V}_K$ is uniformly bounded.

By Lemma 4.9, it follows that $\|D_v^r S\|_{L^\infty(K')} \leq cd(K')^{-r-2/p} \|S\|_{L^p(K')}$, and if the ring $K' \in \mathfrak{N}_K^t$ and $v \in \mathcal{V}_K$, then

$$\begin{aligned} \|S\|_{L^\tau(B_v \cap K')}^\tau &\leq c(|B_v|/|K'|) \|S\|_{L^\tau(K')}^\tau \leq ct^2|K'|^{-1} \|S\|_{L^\tau(K')}^\tau \\ &\leq ct^2|K'|^{-1+\tau(1/\tau-1/p)} \|S\|_{L^p(K')}^\tau \leq ct^2d(K')^{\tau s-2} \|S\|_{L^p(K')}^\tau. \end{aligned}$$

We use the above estimates in (4.22) to obtain

$$\begin{aligned} \|\Delta_h^k S\|_{L^\tau(K_t)}^\tau &\leq c \sum_{K' \in \mathcal{N}_K^t} t^{1+r\tau} d(K')^{1-r\tau-2\tau/p} \|S\|_{L^p(K')}^\tau \\ &+ c \sum_{K' \in \mathfrak{N}_K^t} t^2 d(K')^{\tau s-2} \|S\|_{L^p(K')}^\tau \\ &+ c \sum_{K'' \in \mathcal{F}: d(K'') \leq 2kt/\check{c}} \|S\|_{L^\tau(K'' \cap (K + [0, kh]))}^\tau. \end{aligned} \tag{4.23}$$

Denote by Ω_t^* the set of all $x \in \Omega_t$ such that $[x, x + kh] \subset \Omega$ and

$$[x, x + kh] \subset \cup\{K \in \mathcal{F} : d(K) \leq 2kt/\check{c}\}.$$

In this case, we shall use the obvious estimate

$$\|\Delta_h^k S\|_{L^\tau(\Omega_t^*)}^\tau \leq c \sum_{K \in \mathcal{F}: d(K) \leq 2kt/\check{c}} \|S\|_{L^\tau(K)}^\tau.$$

This estimate along with (4.23) yields

$$\begin{aligned} \omega_k(S, t)^\tau &\leq c \sum_{K \in \mathcal{F}: d(K) \geq 2kt/\check{c}} t^{1+r\tau} d(K)^{1-r\tau-2\tau/p} \|S\|_{L^p(K)}^\tau \\ &+ c \sum_{K \in \mathcal{F}: d(K) \geq 2kt/\check{c}} t^2 d(K)^{s\tau-2} \|S\|_{L^p(K)}^\tau + c \sum_{K \in \mathcal{F}: d(K) \leq 2kt/\check{c}} \|S\|_{L^\tau(K)}^\tau. \end{aligned}$$

Here we used the fact that only finitely many (uniformly bounded number) of the rings involved in the above estimates may overlap at a time due to Lemma 4.7. For the

norms involved in the last sum, we use the estimate $\|S\|_{L^\tau(K)}^\tau \leq cd(K)^{s\tau} \|S\|_{L^p(K)}^\tau$, which follows by Lemma 4.9, to obtain

$$\begin{aligned} \omega_k(S, t)^\tau &\leq c \sum_{K \in \mathcal{F}: d(K) \geq 2kt/\check{c}} t^{1+r\tau} d(K')^{1-r\tau-2\tau/p} \|S\|_{L^p(K')}^\tau \\ &+ c \sum_{K \in \mathcal{F}: d(K) \geq 2kt/\check{c}} t^2 d(K)^{s\tau-2} \|S\|_{L^p(K)}^\tau + c \sum_{K \in \mathcal{F}: d(K) \leq 2kt/\check{c}} d(K)^{s\tau} \|S\|_{L^p(K)}^\tau. \end{aligned}$$

We insert this estimate in (2.1) and interchange the order of integration and summation to obtain

$$\begin{aligned} |S|_{B_\tau^{s,k}}^\tau &= \int_0^\infty t^{-s\tau-1} \omega_k(S, t)^\tau dt \\ &\leq c \sum_{K \in \mathcal{F}} d(K)^{1-r\tau-2\tau/p} \|S\|_{L^p(K)}^\tau \int_0^{\check{c}d(K)/2k} t^{-s\tau+r\tau} dt \\ &+ c \sum_{K \in \mathcal{F}} d(K)^{s\tau-2} \|S\|_{L^p(K)}^\tau \int_0^{\check{c}d(K)/2k} t^{-s\tau+1} dt \\ &+ c \sum_{K \in \mathcal{F}} d(K)^{s\tau} \|S\|_{L^p(K)}^\tau \int_{\check{c}d(K)/2k}^\infty t^{-s\tau-1} dt. \end{aligned}$$

Observe that $-s\tau + r\tau > -1$ is equivalent to $s/2 < r + 1/p$ and $-s\tau + 1 > -1$ is equivalent to $s < 2/\tau = s + 2/p$. Therefore, the above integrals are convergent, and taking into account that $2 - 2\tau/p - s\tau = 2\tau(1/\tau - 1/p - s/2) = 0$, we obtain

$$|S|_{B_\tau^{s,k}}^\tau \leq c \sum_{K \in \mathcal{F}} \|S\|_{L^p(K)}^\tau \leq cn^{\tau(1/\tau-1/p)} \left(\sum_{K \in \mathcal{F}} \|S\|_{L^p(K)}^\tau \right)^{\tau/p} = cn^{\tau s/2} \|S\|_{L^p(\Omega)}^\tau,$$

where we used Hölder’s inequality. This completes the proof of Theorem 4.5. □

4.5 Proof of the Bernstein Estimate (Theorem 4.2) in the Nonnested Case

For the proof of Theorem 4.2, we combine ideas from the proofs of Theorem 3.4 and Theorem 4.5. We shall adhere to a large extent to the notation introduced in the proof of Theorem 3.4 in Sect. 3.3. An important distinction between this proof and the proof of Theorem 3.4 is that the directional derivatives $D_v^{k-1}S$ of any $S \in \mathcal{S}(n, k)$ are piecewise constants along the respective straight lines rather than S being a piecewise constant.

We consider the case when $\Omega \subset \mathbb{R}^2$ is a compact polygonal domain. Assume $S_1, S_2 \in \mathcal{S}(n, k)$, $n \geq n_0$. Then each S_j ($j = 1, 2$) can be represented in the form $S_j = \sum_{R \in \mathcal{R}_j} P_R \mathbb{1}_R$, where \mathcal{R}_j is a set of at most n rings in the sense of Definition 3.1 with disjoint interiors and such that $\Omega = \cup_{R \in \mathcal{R}_j} R$, $P_R \in \Pi_k$, and $S_j \in W^{k-2}(\Omega)$.

Just as in the proof of Theorem 4.5, there exist subdivisions $\mathcal{F}_1, \mathcal{F}_2$ of the rings from $\mathcal{R}_1, \mathcal{R}_2$ with the following properties, for $j = 1, 2$:

- (a) \mathcal{F}_j consists of rings in the sense of Definition 3.1 with parameters N_0^* , c_0^* , and β^* depending only on the structural constants N_0 , c_0 , and β .
- (b) $\cup_{R \in \mathcal{F}_j} R = \Omega$ and $\#F_j \leq cn$.
- (c) There exists a constant N_1 such that for any $R \in \mathcal{F}_j$, there are at most N_1 rings in \mathcal{F}_j intersecting R (R has $\leq N_1$ neighbors in \mathcal{F}_j).
- (d) S_j can be represented in the form $S_j = \sum_{R \in \mathcal{F}_j} P_R \mathbb{1}_R$ with $P_R \in \Pi_k$.

Now, just as in the proof of Theorem 3.4, we denote by \mathcal{U} the collection of all maximal connected sets obtained by intersecting rings from \mathcal{F}_1 and \mathcal{F}_2 . By (3.14), there exists a constant $c > 0$ such that $\#\mathcal{U} \leq cn$.

We claim that there exists a constant N_2 such that for any $U \in \mathcal{U}$ there are no more than N_2 sets $U' \in \mathcal{U}$ that intersect U ; i.e., U has at most N_2 neighbors in \mathcal{U} . Indeed, let $U \in \mathcal{U}$ be a maximal connected component of $R_1 \cap R_2$ with $R_1 \in \mathcal{F}_1$ and $R_2 \in \mathcal{F}_2$. Then using the fact that the ring R_1 has finitely many neighbors in \mathcal{F}_1 and R_2 has finitely many neighbors in \mathcal{F}_2 , we conclude that U has finitely many neighbors in \mathcal{U} .

Further, we introduce the sets \mathcal{A} and \mathcal{T} just as in the proof of Theorem 3.4.

Trapezoids. Our main concern will be in dealing with the trapezoids $T \in \mathcal{T}$. We next use the fact that any ring from \mathcal{F}_j , $j = 1, 2$, has at most N_1 neighbors in \mathcal{F}_j to additionally subdivide the trapezoids from \mathcal{T} into trapezoids whose long sides are sides of good triangles for rings in \mathcal{F}_1 or \mathcal{F}_2 .

Consider an arbitrary trapezoid $T \in \mathcal{T}$. Just as in Sect. 3.3, we may assume that T is a maximal isosceles trapezoid contained in $\Delta_{E_1} \cap \Delta_{E_2}$, where Δ_{E_j} ($j = 1, 2$) is a good triangle for a ring $R_j \in \mathcal{F}_j$, and T is positioned so that its vertices are the points

$$v_1 := (-\delta_1/2, 0), \quad v_2 := (\delta_1/2, 0), \quad v_3 := (\delta_2/2, H), \quad v_4 := (-\delta_2/2, H),$$

where $0 \leq \delta_2 \leq \delta_1$ and $H > \delta_1$. Let $L_1 := [v_1, v_4]$ and $L_2 := [v_2, v_3]$ be the two equal (long) legs of T . We assume that $L_1 \subset E_1$ and $L_2 \subset E_2$. See Fig. 8.

By Lemma 4.7, it follows that there exist less than N_1 rings $K'_\ell \in \mathcal{F}_1$, $\ell = 1, \dots, m'$, each of them with an edge or part of an edge contained in L_1 . By Definition 3.1, each of them can be subdivided into at most two segments so that each of these is a side of a good triangle. Denote by I'_ℓ , $\ell = 1, \dots, m'$, these segments, and by $\Delta_{I'_\ell}$, $\ell = 1, \dots, m'$, the respective good triangles attached to them. More precisely, I'_ℓ is a side of $\Delta_{I'_\ell} \subset K'_\ell$, and $\Delta_{I'_\ell}$ is a good triangle for K'_ℓ . Thus we have $L_1 = \cup_{\ell=1}^{m'} I'_\ell$, where the segments I'_ℓ , $\ell = 1, \dots, m'$, are with disjoint interiors.

Similarly, there exist segments I''_ℓ , $\ell = 1, \dots, m''$, and attached to them good triangles $\Delta_{I''_\ell}$, $\ell = 1, \dots, m''$, for rings from \mathcal{F}_2 , so that $L_2 = \cup_{\ell=1}^{m''} I''_\ell$.

Denote by v'_ℓ , $\ell = 1, \dots, m' + 1$, the vertices of the triangles $\Delta_{I'_\ell}$, $\ell = 1, \dots, m'$, on L_1 so that $I'_\ell = [v'_\ell, v'_{\ell+1}]$, and assume that their orthogonal projections onto the x_2 -axis p'_ℓ , $\ell = 1, \dots, m' + 1$, are ordered so that $0 = p'_1 < p'_2 < \dots < p'_{m'+1} = H$. Exactly in the same way, we define the vertices v''_ℓ , $\ell = 1, \dots, m'' + 1$, of the triangles $\Delta_{I''_\ell}$ and their projections onto the x_2 -axis $0 = p''_1 < p''_2 < \dots < p''_{m''+1} = H$.

For any $q \in [0, H]$, we let $\delta(q)$ be the distance between the points where the line with equation $x_2 = q$ intersects L_1 and L_2 . Thus $\delta(0) = \delta_1$ and $\delta(H) = \delta_2$, and $\delta(q)$ is linear.

Inductively, starting from $q_1 = 0$, one can easily subdivide the interval $[0, H]$ by means of points

$$0 = q_1 < q_2 < \dots < q_{m+1} = H, \quad m \leq m' + m'' \leq 2N_1,$$

with the following properties, for $k = 1, \dots, m$, either

(a) $\delta(q_k) \leq q_{k+1} - q_k < 2\delta(q_k)$

or

(b) $q_{k+1} - q_k > \delta(q_k)$ and (q_k, q_{k+1}) contains no points p'_ℓ or p''_ℓ .

We use the above points to subdivide the trapezoid T . Let $T_k, k = 1, \dots, m$, be the trapezoid bounded by L_1, L_2 , and the lines with equations $x_2 = q_k$ and $x_2 = q_{k+1}$.

We now separate the “bad” from the “good” trapezoids T_k . Namely, if property (a) from above is valid, then T_k is a ring and we place T_k in \mathcal{A} ; if property (b) is valid, then T_k is a “bad” trapezoid and we place T_k in \mathcal{T} . We apply the above procedure to all trapezoids.

Properties of the new trapezoids. We now consider an arbitrary trapezoid T from the above defined \mathcal{T} (the set of bad trapezoids). We next summarise the properties of T . It will be convenient to us to use the same notation as above as well as in the proof of Theorem 3.4. We assume that T is an isosceles trapezoid contained in $\Delta_{E_1} \cap \Delta_{E_2}$, where $\Delta_{E_j}, j = 1, 2$, is a good triangle for a ring $R_j \in \mathcal{F}_j$, and T is positioned so that its vertices are the points

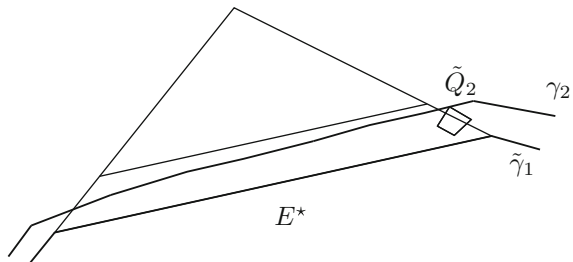
$$v_1 := (-\delta_1/2, 0), \quad v_2 := (\delta_1/2, 0), \quad v_3 := (\delta_2/2, H), \quad v_4 := (-\delta_2/2, H),$$

where $0 \leq \delta_2 \leq \delta_1$ and $H > \delta_1$. Let $L_1 := [v_1, v_4]$ and $L_2 := [v_2, v_3]$ be the two equal (long) sides of T . We assume that $L_1 \subset E_1$ and $L_2 \subset E_2$. See Fig. 9.

As a result of the above subdivision procedure, there exists a triangle Δ_{L_1} with a side L_1 such that Δ_{L_1} is a good triangle for some ring $\tilde{R}_1 \in \mathcal{F}_1$ and $\Delta_{L_1}^\circ \cap \Delta_{E_1}^\circ = \emptyset$. For the same reason, there exists a triangle Δ_{L_2} with a side L_2 such that Δ_{L_2} is a good triangle for some ring $\tilde{R}_2 \in \mathcal{F}_2$ and $\Delta_{L_2}^\circ \cap \Delta_{E_2}^\circ = \emptyset$.

Observe that Δ_{E_1} and Δ_{E_2} are good triangles, and hence the angles of Δ_{E_j} adjacent to E_j are of size $\beta^*/2, j = 1, 2$. Likewise, Δ_{L_1} and Δ_{L_2} are good triangles, and hence the angles of Δ_{L_j} adjacent to L_j are of size $\beta^*/2, j = 1, 2$. Therefore, we may assume that $\Delta_{L_1} \subset \Delta_{E_2}$ and $\Delta_{L_2} \subset \Delta_{E_1}$. Consequently, S_1 is a polynomial of degree $< k$

Fig. 9 Illustration of Case 4 (b)



on Δ_{L_1} and another polynomial of degree $< k$ on Δ_{L_2} . By the same token, S_2 is a polynomial of degree $< k$ on Δ_{L_1} and another polynomial of degree $< k$ on Δ_{L_2} . We shall assume that $\Delta_{L_1} \subset A_1$ and $\Delta_{L_2} \subset A_2$, where $A_1, A_2 \in \mathcal{A}$.

Further, denote by D_1 and D_2 the bottom and top sides of T . We shall denote by $\mathcal{V}_T = \{v_1, v_2, v_3, v_4\}$ the vertices of T , where v_1 is the point of intersection of L_1 and D_1 and the other vertices are indexed counterclockwise.

We shall use the notation $\delta_1(T) := \delta_1$ and $\delta_2(T) := \delta_2$. We always assume that $\delta_1(T) \geq \delta_2(T)$. Clearly, $d(T) \sim H$; more precisely, $H < d(T) < H + \delta_1 + \delta_2$.

Observe that by the construction of the sets \mathcal{T} , \mathcal{A} , and (3.14), it follows that $\mathcal{A} \cup \mathcal{T}$ consists of polygonal sets with disjoint interiors, $\cup_{A \in \mathcal{A}} A \cup_{T \in \mathcal{T}} T = \Omega$, there exists a constant $c > 0$ such that

$$\#\mathcal{A} \leq cn, \quad \#\mathcal{T} \leq cn,$$

and there exists a constant N_3 such that each set from $\mathcal{A} \cup \mathcal{T}$ has at most N_3 neighbors in $\mathcal{A} \cup \mathcal{T}$.

We summarize the most important properties of the sets from \mathcal{T} and \mathcal{A} in the following

Lemma 4.11 *The following properties hold for some constant $0 < \tilde{c} < 1$ depending only on the structural constants N_0, c_0 , and β of the setting:*

- (a) *Let $T \in \mathcal{T}$, and assume the notation related to T from above. If $x \in L_1$ with $|x - v_j| \geq \rho$, $j = 1, 4$, then $B(x, \tilde{c}\rho) \subset \Delta_{L_1} \cup \Delta_{L_2}$. Also, if $x \in L_2$ with $|x - v_j| \geq \rho$, $j = 2, 3$, then $B(x, \tilde{c}\rho) \subset \Delta_{L_1} \cup \Delta_{L_2}$. Furthermore, if $x \in D_1$ with $|x - v_j| \geq \rho$, $j = 1, 2$, then $B(x, \tilde{c}\rho) \subset \Delta_{E_1} \cap \Delta_{E_2}$, and similarly for $x \in D_2$.*
- (b) *Assume that $E = [w_1, w_2]$ is an edge shared by two sets $A, A' \in \mathcal{A}$. Let \mathcal{V}_A be the set of all vertices on ∂A (end points of edges) and let $\mathcal{V}_{A'}$ be the set of all vertices on $\partial A'$. If $x \in E$ with $|x - w_j| \geq \rho$, $j = 1, 2$, for some $\rho > 0$, then*

$$B(x, \tilde{c}\rho) \subset A \cup A' \cup_{v \in \mathcal{V}_A \cup \mathcal{V}_{A'}} B(v, \rho). \tag{4.24}$$

Proof Part (a) of this lemma follows readily from the properties of the trapezoids. Part (b) needs clarification. Suppose that for some $x \in E$ with $|x - w_j| \geq \rho$, $j = 1, 2$, $\rho > 0$, the inclusion (4.24) is not valid. Then there exists a point y from an edge $\tilde{E} = [u_1, u_2]$ of, say, ∂A such that $|y - x| < \rho$ and $|y - u_j| \geq \rho$, $j = 1, 2$. A simple geometric argument shows that if the constant \tilde{c} is sufficiently small (depending only on the parameter β of the setting), then there exists an isosceles trapezoid $\tilde{T} \subset \Delta_E \cap \Delta_{\tilde{E}}$ with two legs contained in E and \tilde{E} such that each leg is longer than its larger base. But then the subdivision of the sets from \mathcal{U} (see the proof of Theorem 3.4) would have created a trapezoid in \mathcal{T} that contains part of A . This is a contradiction, which shows that Part (b) holds true. □

We have the representation

$$S_1(x) - S_2(x) = \sum_{A \in \mathcal{A}} P_A(x) \mathbb{1}_A(x) + \sum_{T \in \mathcal{T}} P_T(x) \mathbb{1}_T(x), \tag{4.25}$$

where $P_A, P_T \in \Pi_k$. Note that $S_1 - S_2 \in C^{k-2}(\Omega)$.

Let $0 < s/2 < k - 1 + 1/p$ and $\tau \leq 1$. Fix $t > 0$, and let $h \in \mathbb{R}^2$ with norm $|h| \leq t$. Write $v := |h|^{-1}h$, and assume $v =: (\cos \theta, \sin \theta)$, $-\pi < \theta \leq \pi$.

Since $S_1, S_2 \in C^{k-2}(\Omega)$, we have the following representation of $\Delta_h^{k-1} S_j(x)$:

$$\Delta_h^{k-1} S_j(x) = |h|^{k-1} \int_{\mathbb{R}} D_v^{k-1} S_j(x + uv) M_{k-1}(u) du,$$

where $M_{k-1}(u)$ is the B-spline with knots $u_0, u_1, \dots, u_{k-1}, u_\ell := \ell|h|$. In fact, $M_{k-1}(u) = (k - 1)[u_0, \dots, u_{k-1}](\cdot - u)_+^{k-2}$ is the divided difference. As is well known, $0 \leq M_{k-1} \leq c|h|^{-1}$, $\text{supp } M_{k-1} \subset [0, (k - 1)|h|]$, and $\int_{\mathbb{R}} M_{k-1}(u) du = 1$. Therefore, by $\Delta_h^k S_j(x) = \Delta_h^{k-1} S_j(x + h) - \Delta_h^{k-1} S_j(x)$, whenever $[x, x + kh] \subset \Omega$, we arrive at the representation

$$\Delta_h^k S_j(x) = |h|^{k-1} \int_0^{k|h|} D_v^{k-1} S_j(x + uv) M_k^*(u) du, \tag{4.26}$$

where $M_k^*(u) := M_{k-1}(u - |h|) - M_{k-1}(u)$.

In what follows, we estimate $\|\Delta_h^k S_1\|_{L^\tau(G)}^\tau - \|\Delta_h^k S_2\|_{L^\tau(G)}^\tau$ for different subsets G of Ω .

Case 1

Let $T \in \mathcal{T}$ be such that $d(T) > 2kt/\tilde{c}$ with \tilde{c} the constant from Lemma 4.11. Denote

$$T_h := \{x \in \Omega : [x, x + kh] \subset \Omega \text{ and } [x, x + kh] \cap T \neq \emptyset\}.$$

We next estimate $\|\Delta_h^k S_1\|_{L^\tau(T_h)}^\tau - \|\Delta_h^k S_2\|_{L^\tau(T_h)}^\tau$.

Assume that $T \in \mathcal{T}$ is a trapezoid positioned as described above in Properties of New Trapezoids. We adhere to the notation introduced there.

In addition, let $v_4 - v_1 =: |v_4 - v_1|(\cos \gamma, \sin \gamma)$ with $\gamma \leq \pi/2$; i.e., γ is the angle between D_1 and L_1 . Assume that $v =: (\cos \theta, \sin \theta)$ with $\theta \in [\gamma, \pi]$. The case $\theta \in [-\gamma, 0]$ is just the same. The case when $\theta \in [0, \gamma] \cup [-\pi, -\gamma]$ is considered similarly.

We set $B_v := B(v, 2kt/\tilde{c})$, $v \in \mathcal{V}_T$. Also, denote

$$\begin{aligned} \mathcal{A}_T^t &:= \{A \in \mathcal{A} : d(A) > 2kt/\tilde{c} \text{ and } A \cap (T + B(0, kt)) \neq \emptyset\}, \\ \mathfrak{A}_T^t &:= \{A \in \mathcal{A} : d(A) \leq 2kt/\tilde{c} \text{ and } A \cap (T + B(0, kt)) \neq \emptyset\} \end{aligned}$$

and

$$\begin{aligned} \mathcal{T}_T^t &:= \{T' \in \mathcal{T} : d(T') > 2kt/\tilde{c} \text{ and } T' \cap (T + B(0, kt)) \neq \emptyset\}, \\ \mathfrak{T}_T^t &:= \{T' \in \mathcal{T} : d(T') \leq 2kt/\tilde{c} \text{ and } T' \cap (T + B(0, kt)) \neq \emptyset\}. \end{aligned}$$

Clearly, $\#\mathcal{A}_T^t \leq c$ and $\#\mathfrak{A}_T^t \leq c$ for some constant $c > 0$.

Case 1 (a). If $[x, x + kh] \subset \Delta_{E_1}$, then $\Delta_h^k S_1(x) = 0$ because S_1 is a polynomial of degree $< k$ on Δ_{E_1} . Hence no estimate is needed.

Case 1 (b). If $[x, x + kh] \subset \cup_{v \in \mathcal{V}_T} B_v$, we estimate $|\Delta_h^k S_1(x)|$ trivially:

$$|\Delta_h^k S_1(x)| \leq |\Delta_h^k S_2(x)| + 2^k \sum_{\ell=0}^k |S_1(x + \ell h) - S_2(x + \ell h)|. \tag{4.27}$$

Clearly, the contribution of this case to estimating $\|\Delta_h^k S_1\|_{L^\tau(T_h)}^\tau - \|\Delta_h^k S_2\|_{L^\tau(T_h)}^\tau$ is

$$\begin{aligned} &\leq c \sum_{v \in \mathcal{V}_T} \sum_{A \in \mathcal{A}_T^v} \|S_1 - S_2\|_{L^\tau(B_v \cap A)}^\tau + c \sum_{v \in \mathcal{V}_T} \sum_{T' \in \mathcal{T}_T^v} \|S_1 - S_2\|_{L^\tau(B_v \cap T')}^\tau \\ &\quad + c \sum_{v \in \mathcal{V}_T} \sum_{A \in \mathcal{A}_T^v} \|S_1 - S_2\|_{L^\tau(B_v \cap A)}^\tau + c \sum_{v \in \mathcal{V}_T} \sum_{T' \in \mathcal{T}_T^v} \|S_1 - S_2\|_{L^\tau(B_v \cap T')}^\tau \\ &\leq \sum_{A \in \mathcal{A}_T^v} ct^2 d(A)^{\tau s - 2} \|S_1 - S_2\|_{L^p(A)}^\tau + \sum_{T' \in \mathcal{T}_T^v} ct^{1 + \tau s / 2} d(T')^{\tau s / 2 - 1} \|S_1 - S_2\|_{L^p(T')}^\tau \\ &\quad + \sum_{A \in \mathcal{A}_T^v} cd(A)^{\tau s} \|S_1 - S_2\|_{L^p(A)}^\tau + \sum_{T' \in \mathcal{T}_T^v} cd(T')^{\tau s} \|S_1 - S_2\|_{L^p(T')}^\tau. \end{aligned}$$

Here we used the following estimates, which are a consequence of Lemma 4.9:

(1) If $A \in \mathcal{A}_T^v$, then

$$\begin{aligned} \|S_1 - S_2\|_{L^\tau(B_v \cap A)}^\tau &\leq c(|B_v|/|A|) \|S_1 - S_2\|_{L^\tau(A)}^\tau \\ &\leq ct^2 d(A)^{-2} \|S_1 - S_2\|_{L^\tau(A)}^\tau \leq ct^2 d(A)^{\tau s - 2} \|S_1 - S_2\|_{L^p(A)}^\tau. \end{aligned}$$

(2) If $T' \in \mathcal{T}_T^v$ and $\delta_1(T') > 2kt/\tilde{c}$, then for any $v \in \mathcal{V}_T$, we have

$$\begin{aligned} \|S_1 - S_2\|_{L^\tau(B_v \cap T')}^\tau &\leq c(|B_v|/|T'|) \|S_1 - S_2\|_{L^\tau(T')}^\tau \leq ct^2 |T'|^{\tau s / 2 - 1} \|S_1 - S_2\|_{L^p(T')}^\tau \\ &\leq ct^2 \delta_1(T')^{\tau s / 2 - 1} d(T')^{\tau s / 2 - 1} \|S_1 - S_2\|_{L^p(T')}^\tau \\ &\leq ct^{1 + \tau s / 2} d(T')^{\tau s / 2 - 1} \|S_1 - S_2\|_{L^p(T')}^\tau, \end{aligned}$$

where we used that $\tau s / 2 < 1$, which is equivalent to $s < s + 2/p$.

(3) If $T' \in \mathcal{T}_T^v$ and $\delta_1(T') \leq 2kt/\tilde{c}$, then for any $v \in \mathcal{V}_T$,

$$\begin{aligned} \|S_1 - S_2\|_{L^\tau(B_v \cap T')}^\tau &\leq c(|B_v \cap T'|/|T'|) \|S_1 - S_2\|_{L^\tau(T')}^\tau \\ &\leq ct \delta_1(T') [\delta_1(T') d(T')]^{-1} \|S_1 - S_2\|_{L^\tau(T')}^\tau \\ &\leq ct d(T')^{-1} [\delta_1(T') d(T')]^{\tau s / 2} \|S_1 - S_2\|_{L^p(T')}^\tau \\ &\leq ct^{1 + \tau s / 2} d(T')^{\tau s / 2 - 1} \|S_1 - S_2\|_{L^p(T')}^\tau. \end{aligned}$$

(4) If $A \in \mathfrak{A}_T^t$, then

$$\begin{aligned} \|S_1 - S_2\|_{L^\tau(B_v \cap A)}^\tau &\leq \|S_1 - S_2\|_{L^\tau(A)}^\tau \leq c|A|^{\tau s/2} \|S_1 - S_2\|_{L^p(A)}^\tau \\ &\leq cd(A)^{\tau s} \|S_1 - S_2\|_{L^p(A)}^\tau. \end{aligned}$$

(5) If $T' \in \mathfrak{T}_T^t$, then

$$\begin{aligned} \|S_1 - S_2\|_{L^\tau(B_v \cap T')}^\tau &\leq \|S_1 - S_2\|_{L^\tau(T')}^\tau \leq c|T'|^{\tau s/2} \|S_1 - S_2\|_{L^p(T')}^\tau \\ &\leq cd(T')^{\tau s} \|S_1 - S_2\|_{L^p(T')}^\tau. \end{aligned}$$

Case 1 (c). If $[x, x + kh] \not\subset \cup_{v \in \mathcal{V}_T} B_v$ and $[x, x + kh]$ intersects D_1 or D_2 , then $\delta_1 > 2kt/\tilde{c} > 2kt$ or $\delta_2 > 2kt$ and hence $[x, x + kh] \subset \Delta_{E_1} \cap \Delta_{E_2}$, which implies $\Delta_h^k S_1(x) = 0$. No estimate is needed.

Case 1 (d). Let $I_T^h \subset T$ be the quadrilateral bounded by the segments $L_1, L_1 - kh, D_1$ and the line with equation $x = v_2 + uh, u \in \mathbb{R}$, where v_2 is the point of intersection of L_2 with D_1 , whenever this straight line intersects L_1 . If the line $x = v_2 + uh, u \in \mathbb{R}$, does not intersect L_1 , then we replace it with the line $x = v_4 + uh, u \in \mathbb{R}$. Furthermore, we subtract B_{v_1} and B_{v_2} from I_T^h .

Set $J_T^h := I_T^h + [0, kh]$.

A simple geometric argument shows that $|J_T^h| \leq 2\delta_1 kt$.

In estimating $\|\Delta_h^k S_1\|_{L^\tau(I_T^h)}^\tau$ there are two subcases to be considered.

If $\delta_1(T) \leq 2kt/\tilde{c}$, we use (4.27) to obtain

$$\|\Delta_h^k S_1\|_{L^\tau(I_T^h)}^\tau \leq \|\Delta_h^k S_2\|_{L^\tau(I_T^h)}^\tau + \|S_1 - S_2\|_{L^\tau(I_T^h)}^\tau + \|S_1 - S_2\|_{L^\tau(J_T^h \cap A_1)}^\tau.$$

We estimate the above norms quite like in Case 1 (b), using Lemma 4.9. We have

$$\begin{aligned} \|S_1 - S_2\|_{L^\tau(I_T^h)}^\tau &\leq c(|I_T^h|/|T|) \|S_1 - S_2\|_{L^\tau(T)}^\tau \\ &\leq ct\delta_1(T)[\delta_1(T)d(T)]^{-1} \|S_1 - S_2\|_{L^\tau(T)}^\tau \leq ctd(T)^{-1}|T|^{\tau s/2} \|S_1 - S_2\|_{L^p(T)}^\tau \\ &\leq ctd(T)^{-1}(\delta_1(T)d(T))^{\tau s/2} \|S_1 - S_2\|_{L^p(T)}^\tau \leq ct^{1+\tau s/2}d(T)^{\tau s/2-1} \|S_1 - S_2\|_{L^p(T)}^\tau. \end{aligned}$$

For the second norm, we get

$$\begin{aligned} \|S_1 - S_2\|_{L^\tau(J_T^h \cap A_1)}^\tau &\leq c|J_T^h| \|S_1 - S_2\|_{L^\infty(A_1)}^\tau \leq ct^2|A_1|^{-\tau/p} \|S_1 - S_2\|_{L^p(A_1)}^\tau \\ &\leq ct^2d(A_1)^{-2\tau/p} \|S_1 - S_2\|_{L^p(A_1)}^\tau = ct^2d(A_1)^{\tau s-2} \|S_1 - S_2\|_{L^p(A_1)}^\tau, \end{aligned}$$

where as before we used the fact that $2\tau/p = 2 - \tau s$.

From the above estimates, we infer

$$\begin{aligned} \|\Delta_h^k S_1\|_{L^\tau(I_T^h)}^\tau &\leq \|\Delta_h^k S_2\|_{L^\tau(I_T^h)}^\tau + ct^{1+\tau s/2}d(T)^{\tau s/2-1} \|S_1 - S_2\|_{L^p(T)}^\tau \\ &\quad + ct^2d(A_1)^{\tau s-2} \|S_1 - S_2\|_{L^p(A_1)}^\tau. \end{aligned}$$

Let $\delta_1(T) > 2kt/\tilde{c}$. We use (4.26) to obtain

$$\begin{aligned} |\Delta_h^k S_1(x)| &\leq |\Delta_h^k S_2(x)| + |\Delta_h^k (S_1 - S_2)(x)| \\ &\leq |\Delta_h^k S_2(x)| + ct^{k-1} \|D_v^{k-1} (S_1 - S_2)\|_{L^\infty([x, x+kh])}, \end{aligned}$$

implying

$$\begin{aligned} \|\Delta_h^k S_1\|_{L^\tau(I_T^h)}^\tau &\leq \|\Delta_h^k S_2\|_{L^\tau(I_T^h)}^\tau + c|I_T^h|t^{\tau(k-1)} \|D_v^{k-1} (S_1 - S_2)\|_{L^\infty(I_T^h \cap T)}^\tau \\ &\quad + c|I_T^h|t^{\tau(k-1)} \|D_v^{k-1} (S_1 - S_2)\|_{L^\infty(A_1)}^\tau. \end{aligned}$$

Clearly,

$$\begin{aligned} \|D_v^{k-1} (S_1 - S_2)\|_{L^\infty(I_T^h \cap T)} &\leq c\delta_1(T)^{-(k-1)} \|S_1 - S_2\|_{L^\infty(T)} \\ &\leq c\delta_1(T)^{-(k-1)} |T|^{-1/p} \|S_1 - S_2\|_{L^p(T)} \leq c\delta_1(T)^{-(k-1)-2/p} \|S_1 - S_2\|_{L^p(T)}, \end{aligned}$$

and

$$\begin{aligned} \|D_v^{k-1} (S_1 - S_2)\|_{L^\infty(A_1)} &\leq cd(A_1)^{-(k-1)} \|S_1 - S_2\|_{L^\infty(A_1)} \\ &\leq cd(A_1)^{-(k-1)-2/p} \|S_1 - S_2\|_{L^p(A_1)}. \end{aligned} \tag{4.28}$$

Therefore,

$$\begin{aligned} \|\Delta_h^k S_1\|_{L^\tau(I_T^h)}^\tau &\leq \|\Delta_h^k S_2\|_{L^\tau(I_T^h)}^\tau + ct^{1+\tau(k-1)} \delta_1(T)^{1-\tau(k-1)-2\tau/p} \|S_1 - S_2\|_{L^p(T)}^\tau \\ &\quad + ct^{1+\tau(k-1)} d(A_1)^{1-\tau(k-1)-2\tau/p} \|S_1 - S_2\|_{L^p(A_1)}^\tau. \end{aligned}$$

Case 1 (e) (Main). Let $T_h^* \subset T_h$ be the set defined by

$$T_h^* := \left\{ x \in T_h : [x, x + kh] \cap L_1 \neq \emptyset \text{ and } [x, x + kh] \not\subset \bigcup_{v \in \mathcal{V}_T} I_T^h B_v \right\}. \tag{4.29}$$

We next estimate $\|\Delta_h^k S_1\|_{L^\tau(T_h^*)}^\tau$.

Let $x \in T_h^*$. Denote by b_1 and b_2 the points where the line through x and $x + kh$ intersects L_1 and L_2 . Set $b = b(x) := b_2 - b_1$. We associate the segment $[x + b, x + b + kh]$ with $[x, x + kh]$ and $\Delta_h^k S_2(x + b)$ with $\Delta_h^k S_1(x)$.

Since $S_1 \in \Pi_k$ on Δ_{E_1} , we have $D_v^{k-1} S_1(y) = \text{constant}$ on $[b_1, x + b]$, and hence

$$D_v^{k-1} S_1(b_1 - uv) = D_v^{k-1} S_1(b_2 - uv) \text{ for } 0 \leq u \leq |x - b_1|. \tag{4.30}$$

Similarly, since $S_2 \in \Pi_k$ on Δ_{E_2} , we have $D_v^{k-1} S_2(y) = \text{constant}$ on $[x + kh, b_2]$, and hence

$$D_v^{k-1} S_2(b_1 + uv) = D_v^{k-1} S_2(b_2 + uv) \text{ for } 0 \leq u \leq |x + kh - b_1|. \tag{4.31}$$

We use (4.26) and (4.30–4.31) to obtain

$$\begin{aligned} \Delta_h^k S_1(x) &= |h|^{k-1} \int_{|b_1-x|}^{k|h|} D_v^{k-1} S_1(x + uv) M_k^*(u) du \\ &\quad + |h|^{k-1} \int_0^{|b_1-x|} D_v^{k-1} S_1(x + uv) M_k^*(u) du \\ &= |h|^{k-1} \int_{|b_1-x|}^{k|h|} D_v^{k-1} S_1(x + uv) M_k^*(u) du \\ &\quad + |h|^{k-1} \int_0^{|b_1-x|} D_v^{k-1} S_1(x + b + uv) M_k^*(u) du \end{aligned}$$

and

$$\begin{aligned} \Delta_h^k S_2(x + b) &= |h|^{k-1} \int_{|b_1-x|}^{k|h|} D_v^{k-1} S_2(x + b + uv) M_k^*(u) du \\ &\quad + |h|^{k-1} \int_0^{|b_1-x|} D_v^{k-1} S_2(x + b + uv) M_k^*(u) du \\ &= |h|^{k-1} \int_{|b_1-x|}^{k|h|} D_v^{k-1} S_2(x + uv) M_k^*(u) du \\ &\quad + |h|^{k-1} \int_0^{|b_1-x|} D_v^{k-1} S_2(x + b + uv) M_k^*(u) du. \end{aligned}$$

Therefore,

$$\begin{aligned} \Delta_h^k S_1(x) &= \Delta_h^k S_2(x + b) + \Delta_h^k (S_1 - S_2)(x) \\ &= \Delta_h^k S_2(x + b) + |h|^{k-1} \int_{|b_1-x|}^{k|h|} D_v^{k-1} [S_1 - S_2](x + uv) M_k^*(u) du \\ &\quad + |h|^{k-1} \int_0^{|b_1-x|} D_v^{k-1} [S_1 - S_2](x + b + uv) M_k^*(u) du, \end{aligned}$$

and hence

$$\begin{aligned} |\Delta_h^k S_1(x)| &\leq |\Delta_h^k S_2(x + b)| + ct^{k-1} \|D_v^{k-1} (S_1 - S_2)\|_{L^\infty([b_1, x+kh])} \\ &\quad + ct^{k-1} \|D_v^{k-1} (S_1 - S_2)\|_{L^\infty([x+b, b_2])}. \end{aligned} \tag{4.32}$$

The key here is that $([b_1, x + kh] \cup [x + b, b_2]) \cap T^\circ = \emptyset$.

Let $T_h^{**} := \{x + b(x) : x \in T_h^*\}$, where T_h^* is from (4.29) and $b(x)$ is defined thereafter. By (4.32), we get

$$\begin{aligned} \|\Delta_h^k S_1\|_{L^\tau(T_h^{**})}^\tau &\leq \|\Delta_h^k S_2\|_{L^\tau(T_h^{**})}^\tau + ct d(A_1) t^{\tau(k-1)} \|D_v^{k-1} (S_1 - S_2)\|_{L^\infty(A_1)}^\tau \\ &\quad + ct d(A_2) t^{\tau(k-1)} \|D_v^{k-1} (S_1 - S_2)\|_{L^\infty(A_2)}^\tau. \end{aligned}$$

Just as (4.28) we have

$$\begin{aligned} \|D_v^{k-1}(S_1 - S_2)\|_{L^\infty(A_1)} &\leq cd(A_1)^{-(k-1)} \|S_1 - S_2\|_{L^\infty(A_1)} \\ &\leq cd(A_1)^{-(k-1)-2/p} \|S_1 - S_2\|_{L^p(A_1)}, \end{aligned}$$

and similar estimates hold with A_1 replaced by A_2 . We use all of the above to obtain

$$\begin{aligned} \|\Delta_h^k S_1\|_{L^\tau(T_h^*)}^\tau &\leq \|\Delta_h^k S_2\|_{L^\tau(T_h^{**})}^\tau + ct^{1+\tau(k-1)} d(A_1)^{1-\tau(k-1)-2\tau/p} \|S_1 - S_2\|_{L^p(A_1)}^\tau \\ &\quad + ct^{1+\tau(k-1)} d(A_2)^{1-\tau(k-1)-2\tau/p} \|S_1 - S_2\|_{L^p(A_2)}^\tau. \end{aligned}$$

It is an important observation that no part of $\|\Delta_h^k S_2\|_{L^\tau(T_h^{**})}^\tau$ has been used for estimation of quantities $\|\Delta_h^k S_1\|_{L^\tau(\cdot)}^\tau$ from previous cases.

Putting all of the above estimates together, we arrive at

$$\|\Delta_h^k S_1\|_{L^\tau(T_h)}^\tau \leq \|\Delta_h^k S_2\|_{L^\tau(T_h)}^\tau + Y_1 + Y_2 + Y_3 + Y_4, \tag{4.33}$$

where

$$\begin{aligned} Y_1 &:= \sum_{A \in \mathcal{A}_T^i} ct^2 d(A)^{\tau s - 2} \|S_1 - S_2\|_{L^p(A)}^\tau + \sum_{A \in \mathcal{A}_T^o} cd(A)^{\tau s} \|S_1 - S_2\|_{L^p(A)}^\tau, \\ Y_2 &:= ct^{1+\tau(k-1)} d(A_1)^{1-\tau(k-1)-2\tau/p} \|S_1 - S_2\|_{L^p(A_1)}^\tau \\ &\quad + ct^{1+\tau(k-1)} d(A_2)^{1-\tau(k-1)-2\tau/p} \|S_1 - S_2\|_{L^p(A_2)}^\tau, \\ Y_3 &:= \sum_{T' \in \mathcal{T}_T^i} ct^{1+\tau s/2} d(T')^{\tau s/2-1} \|S_1 - S_2\|_{L^p(T')}^\tau \\ &\quad + \sum_{T' \in \mathcal{T}_T^o} cd(T')^{\tau s} \|S_1 - S_2\|_{L^p(T')}^\tau + ct^{1+\tau s/2} d(T)^{\tau s/2-1} \|S_1 - S_2\|_{L^p(T)}^\tau, \end{aligned}$$

and

$$Y_4 := ct^{1+\tau(k-1)} \delta_1(T)^{1-\tau(k-1)-2\tau/p} \|S_1 - S_2\|_{L^p(T)}^\tau, \quad \text{if } \delta_1(T) > 2kt/\tilde{c},$$

otherwise $Y_4 := 0$.

Remark In all cases we considered above but *Case 1 (e)*, we used the simple inequality $|\Delta_h^k S_1(x)| \leq |\Delta_h^k S_2(x)| + |\Delta_h^k (S_1 - S_2)(x)|$ to estimate $\|\Delta_h^k S_1\|_{L^\tau(G)}^\tau$ for various sets G , and this works because these sets are of relatively small measure. As Example 3.3 shows, this approach in principle cannot be used in *Case 1 (e)*, and this is the main difficulty in this proof. The gist of our approach in going around is to estimate $|\Delta_h^k S_1(x)|$ by using $|\Delta_h^k S_2(x + b)|$ with some shift b , where $|\Delta_h^k S_2(x + b)|$ is not used to estimate other terms $|\Delta_h^k S_1(x')|$ (there is a one-to-one correspondence between these quantities).

Case 2

Let Ω_h^* be the set of all $x \in \Omega$ such that $[x, x + kh] \subset \Omega$, $[x, x + kh] \cap A \neq \emptyset$ for some $A \in \mathcal{A}$ with $d(A) > 2kt/\tilde{c}$, and $[x, x + kh] \cap T = \emptyset$ for all $T \in \mathcal{T}$ with $d(T) \geq 2kt/\tilde{c}$.

Denote by \mathcal{V}_A the set of all vertices on ∂A , and set $B_v := B(v, 4kt/\tilde{c})$, $v \in \mathcal{V}_A$.

We next indicate how we estimate $|\Delta_h^k S_1(x)|$ in different cases.

Case 2 (a). If $[x, x + kh] \subset A$, then $\Delta_h^k S_1(x) = \Delta_h^k S_2(x) = 0$ and no estimate is needed.

Case 2 (b). If $[x, x + kh] \subset \cup_{v \in \mathcal{V}_A} B(v, 2kt/\tilde{c})$, we estimate $|\Delta_h^k S_1(x)|$ trivially:

$$|\Delta_h^k S_1(x)| \leq |\Delta_h^k S_2(x)| + 2^k \sum_{\ell=0}^k |S_1(x + \ell h) - S_2(x + \ell h)|.$$

Case 2 (c). Let $[x, x + kh]$ intersect the edge $E =: [w_1, w_2]$ from ∂A , that is shared with $A' \in \mathcal{A}$ and $[x, x + kh] \not\subset \cup_{v \in \mathcal{V}_A} B_v$. Let $y := E \cap [x, x + kh]$. Evidently, $|y - w_j| > kt/\tilde{c}$, $j = 1, 2$, and in light of Lemma 4.11, we have $[x, x + kh] \subset B(y, kt) \subset A \cup A'$. In this case, we use the inequality

$$\begin{aligned} |\Delta_h^k S_1(x)| &\leq |\Delta_h^k S_2(x)| + |\Delta_h^k (S_1 - S_2)(x)| \\ &\leq |\Delta_h^k S_2(x)| + ct^{k-1} \|D_v^{k-1} (S_1 - S_2)\|_{L^\infty([x, x+kh])}, \end{aligned}$$

which follows by (4.26).

The case when $[x, x + kh]$ intersects an edge from ∂A that is shared with some $T \in \mathcal{T}$ is covered in Case 1 above.

We proceed further similarly as in Case 1 and in the proof of Theorem 4.5 to obtain

$$\|\Delta_h^k S_1\|_{L^\tau(\Omega_h^*)}^\tau \leq \|\Delta_h^k S_2\|_{L^\tau(\Omega_h^*)}^\tau + Y_1 + Y_2, \tag{4.34}$$

where

$$\begin{aligned} Y_1 &:= \sum_{A \in \mathcal{A}: d(A) \geq 2kt/\tilde{c}} t^{1+\tau(k-1)} cd(A)^{1-\tau(k-1)-2\tau/p} \|S_1 - S_2\|_{L^p(A)}^\tau \\ &+ \sum_{A \in \mathcal{A}: d(A) \geq 2kt/\tilde{c}} ct^2 d(A)^{\tau s-2} \|S_1 - S_2\|_{L^p(A)}^\tau \end{aligned}$$

and

$$\begin{aligned} Y_2 &:= \sum_{A \in \mathcal{A}: d(A) \leq 2kt/\tilde{c}} cd(A)^{\tau s} \|S_1 - S_2\|_{L^p(A)}^\tau \\ &+ \sum_{T \in \mathcal{T}: d(T) \leq 2kt/\tilde{c}} cd(T)^{\tau s} \|S_1 - S_2\|_{L^p(T)}^\tau. \end{aligned}$$

Case 3

Let Ω_h^{**} be the set of all $x \in \Omega$ such that

$$[x, x + kh] \subset \cup\{A \in \mathcal{A} : d(A) \leq 2kt/\tilde{c}\} \cup \{T \in \mathcal{T} : d(T) \leq 2kt/\tilde{c}\}.$$

In this case, we estimate $|\Delta_h^k S_1(x)|$ trivially just as in (4.27). We obtain

$$\begin{aligned} \|\Delta_h^k S_1\|_{L^\tau(\Omega_h^{**})}^\tau &\leq \|\Delta_h^k S_2\|_{L^\tau(\Omega_h^{**})}^\tau + \sum_{A \in \mathcal{A}:d(A) \leq 2kt/\tilde{c}} c \|S_1 - S_2\|_{L^\tau(A)}^\tau \\ &\quad + \sum_{T \in \mathcal{T}:d(T) \leq 2kt/\tilde{c}} c \|S_1 - S_2\|_{L^\tau(T)}^\tau \\ &\leq \|\Delta_h^k S_2\|_{L^\tau(\Omega_h^{**})}^\tau + \sum_{A \in \mathcal{A}:d(A) \leq 2kt/\tilde{c}} cd(A)^{\tau s} \|S_1 - S_2\|_{L^p(A)}^\tau \\ &\quad + \sum_{T \in \mathcal{T}:d(T) \leq 2kt/\tilde{c}} cd(T)^{\tau s} \|S_1 - S_2\|_{L^p(T)}^\tau. \end{aligned}$$

Just as in the proof of Theorem 3.4, it is important to note that in the above estimates only finitely many norms may overlap at a time. From above, (4.33), and (4.34), we obtain

$$\omega_k(S_1, t)^\tau \leq \omega_k(S_2, t)^\tau + \mathbb{A}_t + \mathbb{T}_t,$$

where

$$\begin{aligned} \mathbb{A}_t &:= \sum_{A \in \mathcal{A}:d(A) > 2kt/\tilde{c}} t^{1+\tau(k-1)} cd(A)^{1-\tau(k-1)-2\tau/p} \|S_1 - S_2\|_{L^p(A)}^\tau \\ &\quad + \sum_{A \in \mathcal{A}:d(A) > 2kt/\tilde{c}} ct^2 d(A)^{\tau s-2} \|S_1 - S_2\|_{L^p(A)}^\tau \\ &\quad + \sum_{A \in \mathcal{A}:d(A) \leq 2kt/\tilde{c}} cd(A)^{\tau s} \|S_1 - S_2\|_{L^p(A)}^\tau, \end{aligned}$$

and

$$\begin{aligned} \mathbb{T}_t &:= \sum_{T \in \mathcal{T}:\delta_1(T) > 2kt/\tilde{c}} ct^{1+\tau(k-1)} \delta_1(T)^{1-\tau(k-1)-2\tau/p} \|S_1 - S_2\|_{L^p(T)}^\tau \\ &\quad + \sum_{T \in \mathcal{T}:\delta_2(T) > 2kt/\tilde{c}} ct^{1+\tau(k-1)} \delta_2(T)^{1-\tau(k-1)-2\tau/p} \|S_1 - S_2\|_{L^p(T)}^\tau \\ &\quad + \sum_{T \in \mathcal{T}:d(T) > 2kt/\tilde{c}} ct^{1+\tau s/2} d(T)^{\tau s/2-1} \|S_1 - S_2\|_{L^p(T)}^\tau \\ &\quad + \sum_{T \in \mathcal{T}:d(T) \leq 2kt/\tilde{c}} cd(T)^{\tau s} \|S_1 - S_2\|_{L^p(T)}^\tau. \end{aligned}$$

We insert this estimate in (2.1) and interchange the order of integration and summation to obtain

$$|S_1|_{B_\tau^{s,k}}^\tau \leq |S_2|_{B_\tau^{s,k}}^\tau + Z_1 + Z_2,$$

where

$$\begin{aligned} Z_1 := & c \sum_{A \in \mathcal{A}} d(A)^{1-\tau(k-1)-2\tau/p} \|S_1 - S_2\|_{L^p(A)}^\tau \int_0^{\tilde{c}d(A)/2k} t^{-\tau s + \tau(k-1)} dt \\ & + c \sum_{A \in \mathcal{A}} d(A)^{\tau s - 2} \|S_1 - S_2\|_{L^p(A)}^\tau \int_0^{\tilde{c}d(A)/2k} t^{-\tau s + 1} dt \\ & + c \sum_{A \in \mathcal{A}} d(A)^{\tau s} \|S_1 - S_2\|_{L^p(A)}^\tau \int_{\tilde{c}d(A)/2k}^\infty t^{-\tau s - 1} dt \end{aligned}$$

and

$$\begin{aligned} Z_2 := & c \sum_{T \in \mathcal{T}} \delta_1(T)^{1-\tau(k-1)-2\tau/p} \|S_1 - S_2\|_{L^p(T)}^\tau \int_0^{\tilde{c}\delta_1(T)/2k} t^{-\tau s + \tau(k-1)} dt \\ & + c \sum_{T \in \mathcal{T}} \delta_2(T)^{1-\tau(k-1)-2\tau/p} \|S_1 - S_2\|_{L^p(T)}^\tau \int_0^{\tilde{c}\delta_2(T)/2k} t^{-\tau s + \tau(k-1)} dt \\ & + c \sum_{T \in \mathcal{T}} d(T)^{\tau s/2 - 1} \|S_1 - S_2\|_{L^p(T)}^\tau \int_0^{\tilde{c}d(T)/2k} t^{-\tau s/2} dt \\ & + c \sum_{T \in \mathcal{T}} d(T)^{s\tau} \|S_1 - S_2\|_{L^p(T)}^\tau \int_{\tilde{c}d(T)/2k}^\infty t^{-\tau s - 1} dt. \end{aligned}$$

Observe that $-\tau s + \tau(k-1) > -1$ is equivalent to $s/2 < k-1 + 1/p$, which holds true by the hypothesis, and $-\tau s/2 > -1$ is equivalent to $s < 2/\tau = s + 2/p$, which is obvious. Therefore, all integrals above are convergent, and taking into account that $2 - 2\tau/p - \tau s = 2\tau(1/\tau - 1/p - s/2) = 0$, we obtain

$$\begin{aligned} |S_1|_{B_\tau^{s,k}}^\tau & \leq |S_2|_{B_\tau^{s,k}}^\tau + c \sum_{A \in \mathcal{A} \cup \mathcal{T}} \|S_1 - S_2\|_{L^p(A)}^\tau \\ & \leq |S_2|_{B_\tau^{s,k}}^\tau + cn^{\tau(1/\tau - 1/p)} \left(\sum_{A \in \mathcal{A} \cup \mathcal{T}} \|S_1 - S_2\|_{L^p(A)}^\tau \right)^{\tau/p} \\ & = |S_2|_{B_\tau^{s,k}}^\tau + cn^{\tau s/2} \|S\|_{L^p(\Omega)}^\tau, \end{aligned}$$

where we used Hölder’s inequality. This completes the proof of Theorem 4.2. □

Acknowledgements We would like to give credit to Peter Petrov (Sofia University) with whom the second author discussed the theme of this article some years ago.

References

1. Binev, P., Dahmen, W., DeVore, R., Petrushev, P.: Approximation classes for adaptive methods. *Serdica Math. J.* **28**(4), 391–416 (2002)
2. Cohen, A., DeVore, R., Petrushev, P., Xu, H.: Nonlinear approximation and the space $BV(\mathbb{R}^2)$. *Am. J. Math.* **121**(3), 587–628 (1999)
3. Davydov, O., Petrushev, P.: Nonlinear approximation from differentiable piecewise polynomials. *SIAM J. Math. Anal.* **35**(3), 708–758 (2003)
4. DeVore, R., Lorentz, G.G.: *Constructive Approximation*, vol. 303. Springer Grundlehren, Heidelberg (1993)
5. DeVore, R., Yu, X.Y.: Degree of adaptive approximation. *Math. Comp.* **55**(192), 625–635 (1990)
6. Karaivanov, B., Petrushev, P.: Nonlinear piecewise polynomial approximation beyond Besov spaces. *Appl. Comput. Harmon. Anal.* **15**(3), 177–223 (2003)
7. Petrushev, P.: Direct and converse theorems for spline and rational approximation and Besov spaces. In: *Function Spaces and Applications. Lecture Notes in Mathematics*, Lund, vol. 1302, pp. 363–377. Springer, Berlin (1988)
8. Petrushev, P., Popov, V.: *Rational Approximation of Real Functions*. Cambridge University Press, Cambridge (1987)