Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Approximation Theory 136 (2005) 60–83

Journal of Approximation Theory

www.elsevier.com/locate/jat

Bivariate n-term rational approximation

Kyungwon Park*

IMI-Department of Mathematics, University of South Carolina, Columbia, SC 29208, USA

Received 2 June 2004; accepted 24 May 2005

Communicated by Dany Leviatan
Available online 27 July 2005

Abstract

We study nonlinear approximation in $L_p(\mathbb{R}^2)$, $0 < p < \infty$, from n-term rational functions. Our main result relates n-term rational approximation in L_p to nonlinear approximation from a broad class of piecewise polynomials over multilevel triangulations allowing a lot of flexibility and, in particular, arbitrarily sharp angles. This relationship and the existing estimates for spline approximation give a Jackson estimate for n-term rational approximation in terms of a minimal smoothness norm over a large collection of anisotropic Besov-type spaces (B-spaces).

1. Introduction

While the theory of univariate rational approximation is considerably well developed area in Approximation theory (see, e.g., [9]), the theory of multivariate rational approximation is just emerging. The reason for this is that it is extremely hard to deal with multivariate rational functions. Apparently rational functions of the form $R = P/Q$, where P and Q are algebraic polynomials in d variables ($d > 1$), are powerful tool for approximation but very little is known about them. It seems natural to consider nonlinear n-term approximation

* Fax: +1 803 777 6527.
E-mail address: kpark001@math.sc.edu.

0021-9045/$ - see front matter
doi:10.1016/j.jat.2005.05.005

from the dictionary $\mathcal{R}$ consisting of all functions on $\mathbb{R}^d$ of the form

$$R = \sum_{j=1}^{n} r_j, \tag{1.1}$$

where r_j are partial fractions. In [8], it is considered the case when the r_j's are of the form $r(x) = \prod_{\mu=1}^{d} \frac{a_\mu x_\mu + b_\mu}{(x_\mu - \alpha)^2 + \beta_\mu^2}$. The main result from [8] relates this type of n-term rational approximation with nonlinear piecewise polynomial approximation over arbitrary dyadic partitions of $\mathbb{R}^d$.

In this article we obtain similar results for the more complicated case of n-term rational approximation in $\mathbb{R}^2$, when the r_j's are of the form

$$r(x) = \prod_{\mu=1}^{6} \frac{a_\mu x_1 + b_\mu x_2 + c_\mu}{1 + (\alpha_\mu x_1 + \beta_\mu x_2 + \gamma_\mu)^2} \quad \text{with } a_\mu, b_\mu, c_\mu, \alpha_\mu, \beta_\mu, \gamma_\mu \in \mathbb{R}. \tag{1.2}$$

Our main result relates the bivariate n-term rational approximation to nonlinear approximation from a broad class of piecewise polynomials over multilevel nested triangulations. To be more specific, let us consider a sequence of nested triangulations $(\mathcal{T}_m)_{m\in\mathbb{Z}}$ such that each level $\mathcal{T}_m$ is a partition of $\mathbb{R}^2$ into triangles and a refinement of the previous level $\mathcal{T}_{m-1}$. Denote $\mathcal{T} := \bigcup_{m\in\mathbb{Z}} \mathcal{T}_m$. Natural mild conditions are imposed on the triangulations in order to prevent them from possible deterioration. These conditions, however, allow the triangles in $\mathcal{T}$ to change in size, shape, and orientation quickly when moving around at a given level or through the levels. In particular, triangles with arbitrarily sharp angles are allowed in any location and at any level. Let $\Sigma_n^k(\mathcal{T})$ denote the nonlinear set of all n-term piecewise polynomial functions S of the form $\sum_{\triangle\in\Lambda_n} \mathbb{1}_\triangle \cdot P_\triangle$, where each $P_\triangle$ is a polynomial of degree $< k$ and Λ_n consists of n triangles from $\mathcal{T}$. Further, denote by $\sigma_n(f, \mathcal{T})_p$ the error of L_p-approximation to f from $\Sigma_n^k(\mathcal{T})$. Denote by $R_n(f)_p$ the error of L_p-approximation of f from n-term rational functions of form (1.1) with r_j of form (1.2).

Our main result says that $(R_n(f)_p)$ has the rate of $(\sigma_n(f, \mathcal{T})_p)$ or a better rate for any $0 < p < \infty$, $k \geqslant 1$, and multilevel triangulation $\mathcal{T}$. This relationship and the existing estimates for anisotropic piecewise polynomial approximation (see [6]) give a Jackson estimate for n-term rational approximation in terms of the minimal smoothness norm over a wide collection of anisotropic Besov-type smoothness spaces (B-spaces).

Results of the same character are obtained also by Dekel and Leviatan [3] under the restrictive condition that the piecewise polynomials are over triangulations satisfying the minimal angle condition (regular triangulations, see Section 2.1) when $1 < p < \infty$.

The main tools in proving our result are the famous result of Newman on the rational uniform approximation of $|x|$ and an anisotropic version of the Fefferman–Stein vector-valued maximal inequality.

The outline of the paper is the following. In Section 2 we gather all necessary auxiliary definition and results. Thus in Section 2.1 we give the definition and some basic properties of the multilevel triangulations considered. In Section 2.2 we give the needed simple facts about polynomials. In Section 2.3 we give some known facts about B-spaces and nonlinear piecewise polynomial approximation. In Section 2.4 we provide everything we need about maximal functions. Finally, in Section 3 we state and prove our main results on n-term rational approximation.

Throughout this article, for a set $E \subset \mathbb{R}^d$, $\mathbb{1}_E$ denotes the characteristic function of E, and $|E|$ denotes the Lebesgue measure of E, while E° means the interior of E. For a finite set E, $\#E$ denotes the cardinality of E. For a vector (point) $x \in \mathbb{R}^2$, $|x|$ denotes the Euclidean norm of x. Positive constants are denoted by $c, c_1, c', \ldots$ and if not specified they may vary at every occurrence. Further, $A \approx B$ means $c_1 \leqslant A/B \leqslant c_2$, and $A := B$ or $B =: A$ stands for "A is by definition equal to B". Whenever the L_p-norm of a function is on $\mathbb{R}^2$, we write briefly $\|\cdot\|_p$, whereas $\|\cdot\|_{L_p(E)}$ denotes the L_p-norm on a particular set $E \subset \mathbb{R}^2$. The set of all algebraic polynomials in two variables of total degree $< k$ is denoted by Π_k.

2. Preliminary results

2.1. Multilevel nested triangulations

Here we introduce several types of *multilevel nested triangulations* following the development in [6]. Let $\mathcal{T} = \bigcup_{m\in\mathbb{Z}} \mathcal{T}_m$ be a set of closed triangles in $\mathbb{R}^2$ with *levels* $(\mathcal{T}_m)_{m\in\mathbb{Z}}$. We say that $\mathcal{T}$ is a *hierarchical nested triangulation* or simply *triangulation* of $\mathbb{R}^2$ if the following conditions are satisfied:

(a) Every level $\mathcal{T}_m,\ m \in \mathbb{Z}$, is a set of triangles with disjoint interiors which cover $\mathbb{R}^2$, i.e.

$$\mathbb{R}^2 = \bigcup_{\triangle\in\mathcal{T}_m} \triangle.$$

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

(c) Each triangle $\triangle \in \mathcal{T}_m$ has at least two and at most M_0 subtriangles in $\mathcal{T}_{m+1}$, where $M_0 \geqslant 4$ is a constant independent of m.

(d) The *valence* N_v of each vertex $v \in \mathcal{V}_m$ (the number of triangles $\triangle \in \mathcal{T}_m$ which share v as a vertex) is less than N_0, where $N_0 \geqslant 3$ is a constant.

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

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

$$K = \bigcup_{\triangle\subset\Lambda_n\subset\mathcal{T}_m} \triangle \quad \text{where } \#\Lambda_n < \infty.$$

We denote by $\mathcal{V}_m$ and $\mathcal{E}_m$ the set of all vertices and edges of triangles from $\mathcal{T}_m$, respectively. We set $\mathcal{V} := \bigcup_{m\in\mathbb{Z}} \mathcal{V}_m$ and $\mathcal{E} := \bigcup_{m\in\mathbb{Z}} \mathcal{E}_m$.

Note that any two triangles in $\mathcal{T}$ either have disjoint interiors or one of them contains the other. If $\triangle$ and $\triangle'$ are two different triangles in $\mathcal{T}$ and $\triangle' \subset \triangle$, then we say that $\triangle$ is an *ancestor* of $\triangle'$, while $\triangle'$ is a *descendant* of $\triangle$. Also if $\triangle' \in \mathcal{T}_{m+1}$ and $\triangle' \subset \triangle,\ \triangle \in \mathcal{T}_m$, then $\triangle'$ is called a *child* of $\triangle$. Now we define two types of triangulations by imposing more conditions in addition to conditions (a)–(f) above.

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

(i) There exists constants $0 < r < \rho < 1$ ($r \leqslant \frac{1}{4}$) such that for each $\triangle \in \mathcal{T}$ and any child $\triangle' \in \mathcal{T}$ of $\triangle$,

$$r|\triangle| \leqslant |\triangle'| \leqslant \rho|\triangle|. \tag{2.1}$$

(ii) There exists a constant $0 < \delta \leqslant 1$ such that for any $\triangle', \triangle'' \in \mathcal{T}_m$ ($m \in \mathbb{Z}$) with an edge,

$$\delta \leqslant \frac{|\triangle'|}{|\triangle''|} \leqslant \delta^{-1}. \tag{2.2}$$

For $v \in \mathcal{V}_m$, $m \in \mathbb{Z}$ we denote by θ_v the *cell* associated to v, i.e. the union of all triangles from $\mathcal{T}_m$ which have v as a common vertex. We denote by Θ_m the set of all cells generated by $\mathcal{T}_m$ and $\Theta := \bigcup_{m\in\mathbb{Z}} \Theta_m$.

Strong locally regular triangulations. A triangulation $\mathcal{T} = \bigcup_{m\in\mathbb{Z}} \mathcal{T}_m$ is called a *strong locally regular triangulation* of $\mathbb{R}^2$ or, briefly, an *SLR-triangulation* if $\mathcal{T}$ satisfies the following two additional conditions:

(i) There exists a constant $0 < r < \rho < 1 (r \leqslant \frac{1}{4})$ such that for each $\triangle \in \mathcal{T}$ and any child $\triangle' \in \mathcal{T}$ of $\triangle$,

$$r|\triangle| \leqslant |\triangle'| \leqslant \rho|\triangle|. \tag{2.3}$$

(ii) *Affine transform angle condition:* There exists a constant $\beta = \beta(\mathcal{T})$, $0 < \beta \leqslant \pi/3$, such that if $\triangle_0 \in \mathcal{T}_m$, $m \in \mathbb{Z}$, and $A : \mathbb{R}^2 \to \mathbb{R}^2$ is an affine transform that maps $\triangle_0$ one-to-one onto an equilateral reference triangle, then for every $\triangle \in \mathcal{T}_m$ which has at least one common vertex with $\triangle_0$ and for every child $\triangle \in \mathcal{T}_{m+1}$ of $\triangle_0$, we have

$$\min \text{angle}\,(A(\triangle)) \geqslant \beta, \tag{2.4}$$

where $A(\triangle)$ is the image of $\triangle$ by the affine transform A.

It can be proved (see [2]) that condition (ii) is equivalent to the following condition:

(ii′) There exists a constant $0 < \delta_1 \leqslant 1/2$ such that for any $\triangle', \triangle'' \in \mathcal{T}_m$ ($m \in \mathbb{Z}$) sharing an edge,

$$|\text{conv}\,(\triangle' \cup \triangle'')|/|\triangle'| \leqslant \delta_1^{-1}, \tag{2.5}$$

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

Note that condition (ii′) implies (2.2) with $\delta_1 = \delta$. Therefore, each SLR-triangulation is an LR-triangulation, however, the inverse statement is not true (see [6]).

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

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

Evidently, every regular triangulation is an SLR-triangulation but the converse statement is not true.

With the next remarks we clarify several important issues concerning different types of multilevel triangulations.

(a) For each of the triangulations there are constants which are assumed fixed. We refer to them as *parameters*. Thus the parameters of an LR-triangulation are M_0, N_0, ρ, δ, and r and the parameters of an SLR-triangulation are M_0, N_0, ρ, δ, r and β.

(b) The most important observation is that the collection of all SLR-triangulation with given (fixed) parameters is invariant under affine transforms. More precisely, if $\mathcal{T}$ is an SLR-triangulation in $\mathbb{R}^2$ and $\mathbf{A}$ is an affine transform of $\mathbb{R}^2$, then $\mathbf{A}(\mathcal{T}) := \{\mathbf{A}(\triangle) : \triangle \in \mathcal{T}\}$ is an SLR-triangulation with the same parameters. The LR-triangulations with fixed parameters are also invariant under affine transforms.

(c) If $\mathcal{T}$ is an LR-triangulation and $\triangle', \triangle'' \in \mathcal{T}_m$ have a common edge, then it may happen that $\triangle'$ is an equilateral triangle (or close to an equilateral triangle) but $\triangle''$ has an uncontrollably sharp angle. Such a configuration on an SLR-triangulation is impossible, however, at any level and location there can be triangles with uncontrollably sharp angles. For more details, see [6].

(d) In an SLR-triangulation $\mathcal{T}$ there can be an equilateral (or close to such) triangle $\triangle^\diamond$ at any level $\mathcal{T}_m$ with descendants $\triangle_1 \supset \triangle_2 \supset \cdots$ such that $\min \text{angle}(\triangle_j) \to 0$ as $j \to \infty$, and also a sequence $(\triangle'_j)_{j=0}^{\infty} \subset \mathcal{T}_m$ with $\triangle'_0 = \triangle^\diamond$ and $\triangle'_j \cap \triangle'_{j+1} \neq \emptyset$ $(j = 0, 1, \ldots)$ such that $\min \text{angle}\,(\triangle'_j) \to 0$ as $j \to \infty$.

(e) For an SLR-triangulation $\mathcal{T}$, conditions (2.3)–(2.5) suggest geometric rates of change of $|\triangle|$, $\min \text{angle}\,(\triangle)$, and $\max \text{edge}\,(\triangle)$ as $\triangle \in \mathcal{T}_m$ moves away from a fixed triangle $\triangle' \in \mathcal{T}_m$. However, as it will be shown later in this section, the rates of change are powers of the number of the connecting edges. A similar observation is true for LR-triangulations.

In the following we show how $|\triangle|$, $|\max \text{edge}\,(\triangle)|$, and $\min \text{angle}\,(\triangle)$ may change as $\triangle \in \mathcal{T}$ moves away and in depth from a fixed triangle. (See [6, Lemma 2.4] for the proof.)

Proposition 2.1. *Let $\mathcal{T}$ be an LR-triangulation of $\mathbb{R}^2$. Suppose that $\triangle', \triangle'' \in \mathcal{T}_m$, $m \in \mathbb{Z}$, and $\triangle'$ and $\triangle''$ can be connected by $< 2^\nu$ intermediate edges from $\mathcal{E}_m$ with (pairwise) common vertices. Then there exist $\triangle_1, \triangle_2 \in \mathcal{T}_{m-2N_0\nu}$ with a common vertex such that $\triangle' \subset \triangle_1$ and $\triangle'' \subset \triangle_2$.*

Lemma 2.2. *Let $\mathcal{T}$ be an SLR-triangulation of $\mathbb{R}^2$ with parameter $\beta = \beta(\mathcal{T})$, $0 < \beta \leqslant \pi/3$.*

(a) *If $\triangle', \triangle'' \in \mathcal{T}_m$, $m \in \mathbb{Z}$, and $\triangle' \cap \triangle'' \neq \emptyset$, then*

$$\eta_1^{-1} \leqslant |\max \text{edge}\,(\triangle')|/|\max \text{edge}\,(\triangle'')| \leqslant \eta_1, \tag{2.6}$$

where η_1 depends only on β and N_0.

(b) *If $\triangle \in \mathcal{T}_m$, $\triangle' \in \mathcal{T}_{m+1}$, and $\triangle' \subset \triangle$, then*

$$1 \leqslant |\max \text{edge}\,(\triangle)|/|\max \text{edge}\,(\triangle')| \leqslant \eta_2, \tag{2.7}$$

where η_2 depends only on the parameters of $\mathcal{T}$.

Proof. (a) It suffices to prove that if $\triangle', \triangle'' \in \mathcal{T}_m$ have a common edge, then

$$\eta_0^{-1} \leqslant |\max \text{edge}\,(\triangle')|/|\max \text{edge}\,(\triangle'')| \leqslant \eta_0, \quad \eta_0 > 1. \tag{2.8}$$

Then, since every vertex can have valence at most N_0, (2.6) follows with $\eta_1 = \eta_0^{\lceil N_0/2 \rceil}$ by applying the above estimate $\lceil N_0/2 \rceil$ times.

Suppose that $\triangle', \triangle'' \in \mathcal{T}_m$ have a common edge. Let $\triangle_1$ be an equilateral reference triangle of side length one and let $\mathbf{A}$ be an affine transform which maps $\triangle'$ one-to-one onto $\triangle_1$. Write $\triangle_2 := \mathbf{A}(\triangle'')$. Let S_1^- be the circle inscribed in $\triangle_1$ and let S_1^+ be the circle circumscribed around $\triangle_1$. Similarly, we let S_2^- and S_2^+ be the circles inscribed in $\triangle_2$ and circumscribed around $\triangle_2$, respectively. Denote by r_j^-, r_j^+ $(j = 1, 2)$ the radii of the circles S_j^-, S_j^+ $(j = 1, 2)$, respectively. Simple geometric argument shows that

$$r_1^+ = \frac{1}{\sqrt{3}} \quad \text{and} \quad r_2^- \geqslant 2 \sin \frac{\beta}{2}, \tag{2.9}$$

where β is from condition (2.4) on the SLR-triangulations.

Write $E_j^- := \mathbf{A}^{-1}(S_j^-)$, $E_j^+ := \mathbf{A}^{-1}(S_j^+)$, $j = 1, 2$. Since $\mathbf{A}$ is an affine transform, then $\mathbf{A}^{-1}$ is also an affine transform and, therefore, E_j^-, E_j^+ $(j = 1, 2)$ are ellipses. Denote by d_j^-, d_j^+, $j = 1, 2$, (the lengths of) the major diameters of the above ellipses. Since A^{-1} is an affine transform and $E_j^\pm$ $(j = 1, 2)$ are images of circles, then

$$\frac{d_1^\pm}{d_2^\pm} = \frac{r_1^\pm}{r_2^\pm}$$

Using this and (2.9), we obtain

$$\frac{d_1^+}{d_2^-} = \frac{r_1^+}{r_2^-} \leqslant \frac{1}{2\sqrt{3} \sin \frac{\beta}{2}} =: \eta_0.$$

We have $\triangle' \subset E_1^+$ and $E_2^- \subset \triangle''$, and hence

$$|\text{max edge } (\triangle')| \leqslant d_1^+ \leqslant \eta_0 d_2^- \leqslant \eta_0 |\text{max edge } (\triangle'')|,$$

which yields (2.8), using also the symmetry.

(b) The right-hand-side estimate in (2.7) follows immediately by (2.1) and the fact that any triangle $\triangle \in \mathcal{T}$ can have at most M_0 children. The left-hand-side estimate in (2.7) is obvious. $\square$

Theorem 2.3. (a) *Let $\mathcal{T}$ be an LR-triangulation of $\mathbb{R}^2$ with parameters $0 < r < \rho < 1$ and N_0. Suppose that $\triangle', \triangle'' \in \mathcal{T}_m$, $m \in \mathbb{Z}$, and $\triangle'$ and $\triangle''$ can be connected by n intermediate triangles* (*or edges*) *with common vertices from $\mathcal{T}_m$. Then*

$$c_1^{-1} n^{-s} \leqslant \frac{|\triangle'|}{|\triangle''|} \leqslant c_1 n^s \tag{2.10}$$

with $s := 2N_0 \log_2(\rho/r)$ and $c_1 := \delta^{-N_0}(\rho/r)^{2N_0}$.

(b) *Let $\mathcal{T}$ be an SLR-triangulation of $\mathbb{R}^2$ with parameter $\beta = \beta(\mathcal{T}), 0 < \beta \leqslant \pi/3$. Suppose that $\triangle', \triangle'' \in \mathcal{T}_m$, $m \in \mathbb{Z}$, and $\triangle'$ and $\triangle''$ can be connected by n intermediate triangles* *with common vertices from $\mathcal{T}_m$. Then*

$$c_2^{-1} n^{-u} \leqslant \frac{|\text{max edge}\,(\triangle')|}{|\text{max edge}\,(\triangle'')|} \leqslant c_2 n^u \tag{2.11}$$

with $u := 2N_0 \log_2(\eta_2)$ *and* $c_2 := \eta_1 \eta_2^{2N_0}$ *where* η_1, η_2 *are from Lemma* 2.2.

Proof. (a) See [6, Theorem 2.5].

(b) Choose $v \geqslant 1$ so that $2^{v-1} \leqslant n < 2^v$. By Proposition 2.1, there exist $\triangle_1, \triangle_2 \in \mathcal{T}_{m-2N_0 v}$ with a common vertex such that $\triangle' \subset \triangle_1$ and $\triangle'' \subset \triangle_2$. Using (2.6), we have

$$|\text{max edge}\,(\triangle_1)| \leqslant \eta_1 |\text{max edge}\,(\triangle_2)|.$$

On the other hand, applying (2.7) repeatedly, we infer

$$|\text{max edge}\,(\triangle_2)| \leqslant \eta_2^{2N_0 v} |\text{max edge}\,(\triangle'')|.$$

Combining these estimates, we obtain

$$|\text{max edge}\,(\triangle')| \leqslant |\text{max edge}\,(\triangle_1)| \leqslant \eta_1 \eta_2^{2N_0 v} |\text{max edge}\,(\triangle'')|$$

which implies (2.11) since $2^{v-1} \leqslant n$. □

Theorem 2.4. (a) *Let $\mathcal{T}$ be an SLR-triangulation of $\mathbb{R}^2$ with parameters $\beta = \beta(\mathcal{T})$, $0 < \beta \leqslant \pi/3$. There exists a constant $0 < \vartheta < 1$ depending only on β such that if $\triangle \in \mathcal{T}_m$ $(m \in \mathbb{Z})$, $\triangle' \in \mathcal{T}_{m+\ell}$, $\ell \geqslant 1$, and $\triangle' \subset \triangle$, then*

$$\vartheta^{\ell} \leqslant \frac{\text{min angle}\,(\triangle')}{\text{min angle}\,(\triangle)} \leqslant \vartheta^{-\ell}. \tag{2.12}$$

(b) *Let $\mathcal{T}$ be an LR-triangulation of $\mathbb{R}^2$. There exist constants $0 < r_1 < \rho_1 < 1$ depending only on the parameters of $\mathcal{T}$ (see the definition of LR-triangulations) such that if $\triangle \in \mathcal{T}_m$ $(m \in \mathbb{Z})$, $\triangle' \in \mathcal{T}_{m+3N_0\ell}$, $\ell \geqslant 1$, and $\triangle' \subset \triangle$, then*

$$r_1^{\ell} \leqslant \frac{|\text{max edge}\,(\triangle')|}{|\text{max edge}\,(\triangle)|} \leqslant \rho_1^{\ell}. \tag{2.13}$$

Proof. (a) See [6] (see Lemma 2.3).

(b) For the proof of the upper bound in (2.13) the argument is quite similar to the argument of the proof of Lemma 2.7 in [2] and will be omitted.

The argument for the proof of the lower bound in (2.13) is simpler. Suppose $\triangle \in \mathcal{T}_m$, $\triangle' \in \mathcal{T}_{m+1}$, and $\triangle' \subset \triangle$. Let $e_{\max}$ and $e'_{\max}$ be the largest edges of $\triangle$ and $\triangle'$, respectively. Denote by h the length of the height to $e_{\max}$ in $\triangle$ and by h' the length of the height to $e'_{\max}$ in $\triangle'$. Further, let R and R' be the radii of the circles inscribed in $\triangle$ and $\triangle'$, respectively. A simple geometric argument shows that $R < h < 3R$ as well as $R' < h' < 3R'$. Since $\triangle' \subset \triangle$, then $R' \leqslant R$ and hence $h' < 3h$. We use this and (2.1) to obtain

$$(1/2) r |e_{\max}| h = r|\triangle| \leqslant |\triangle'| \leqslant (1/2)|e'_{\max}| h' \leqslant (3/2)|e'_{\max}| h$$

which implies $|e'_{\max}| \geqslant (r/3)|e_{\max}|$ where r is the original parameter of $\mathcal{T}$. This obviously implies the lower bound in (2.13). □

Stars. In order to deal with graph distances and neighborhood relations it is convenient to employ the notion of mth-level star of a set.

Definition 2.5. For any set $E \subset \mathbb{R}^2$ and $m \in \mathbb{Z}$, we define $\text{star}_m^1(E)$ *by*

$$\text{star}_m^1(E) := \cup\{\theta \in \Theta_m : \theta^\circ \cap E \neq \emptyset\} \tag{2.14}$$

and inductively

$$\text{star}_m^k(E) := \text{star}_m^1(\text{star}_m^{k-1}(E)), \quad k > 1. \tag{2.15}$$

When E consists of a single point x, in slight abuse of notation, we shall write $\text{star}_m^k(x)$ instead of $\text{star}_m^k(\{x\})$.

2.2. Local polynomial approximation

We shall frequently use the equivalence of norms of polynomials over various sets in $\mathbb{R}^2$, which we give in the following proposition. See [6] for the proof.

Proposition 2.6. *Let* $P \in \Pi_k$, $k \geqslant 1$, *and* $0 < p, q \leqslant \infty$.

(a) *For any triangle* $\triangle$,

$$\|P\|_{L_p(\triangle)} \approx |\triangle|^{1/p-1/q}\|P\|_{L_q(\triangle)}. \tag{2.16}$$

where $c = c(p, q, k)$.

(b) *If* $\triangle$ *and* $\triangle'$ *are two triangles such that* $\triangle' \subset \triangle$ *and* $|\triangle| \leqslant c_0|\triangle'|$, *then*

$$\|P\|_{L_p(\triangle)} \leqslant c\|P\|_{L_p(\triangle')}, \tag{2.17}$$

where $c = c(p, k, c_0)$.

(c) *If* $\triangle$ *and* $\triangle'$ *are two triangles such that* $\triangle' \subset \triangle$ *and* $|\triangle| \leqslant c_1|\triangle'|$, *then*

$$\|P\|_{L_p(\triangle)} \leqslant c\|P\|_{L_p(\triangle\setminus\triangle')} \approx c|\triangle|^{1/p-1/q}\|P\|_{L_q(\triangle\setminus\triangle')}, \tag{2.18}$$

where $c = c(p, k, c_1)$.

In the following, $\triangle^\diamond$ will denote an equilateral (reference) triangle of side one, centered at the origin. We shall need an estimate on the growth of a polynomial $P(x)$ as x moves away from the origin.

Lemma 2.7. *Let* $P \in \Pi_k$ *and* $0 < p \leqslant \infty$. *Then*

$$|P(x)| \leqslant c\|P\|_{L_p(\triangle^\diamond)}(1 + |x|)^{k-1} \quad \text{for } x \in \mathbb{R}^2, \tag{2.19}$$

where $c = c(p, k)$.

Proof. Let $P(x) = \sum_{|\alpha|<k} a_\alpha x^\alpha$. Then for $x \in \mathbb{R}^2$,

$$|P(x)| \leqslant \sum_{|\alpha|<k} |a_\alpha||x|^\alpha \leqslant k^2 \max_\alpha\{|a_\alpha|\}(1 + |x|)^{k-1} \leqslant c\|P\|_{L_p(\triangle^\diamond)}(1 + |x|)^{k-1}$$

since all norms in a finite-dimensional space are equivalent. $\square$

For $f \in L_p(E)$, $E \subset \mathbb{R}^2$, $0 < p \leqslant \infty$, and $k \geqslant 1$, we denote by $E_k(f, E)_p$ the error of L_p-approximation to f from Π_k, i.e.

$$E_k(f, E)_p := \inf_{P \in \Pi_k} \|f - P\|_{L_p(E)}. \tag{2.20}$$

We also denote by $w_k(f, E)_p$ the kth modulus of smoothness of $f \in L_p(E)$, defined by

$$w_k(f, E)_p := \sup_{h \in \mathbb{R}^2} \|\triangle_h^k(f, \cdot)\|_{L_p(E)}, \tag{2.21}$$

where

$$\triangle_h^k(f, x) := \begin{cases} \sum_{j=0}^{k} (-1)^{j+k} \binom{k}{j} f(x + jh) & \text{if } [x, x + kh] \subset E, \\ 0 & \text{otherwise} \end{cases}$$

and $[x, x + kh]$ denotes the line segment between x and $x + kh$.

Proposition 2.8 (*Whitney*). *Let $f \in L_p(\triangle)$ for some triangle $\triangle$, $0 < p \leqslant \infty$, and $k \geqslant 1$. Then*

$$E_k(f, \triangle)_p \leqslant c w_k(f, \triangle)_p, \tag{2.22}$$

where $c = c(p, k)$.

We refer the reader to [6] for the proof of Proposition 2.8.

2.3. Nonlinear piecewise polynomial approximation and B-spaces

In this section we provide the basic results of the theory of nonlinear n-term approximation from piecewise polynomials generated by multilevel nested triangulations, developed in [6]. This theory provides important ingredients for our theory of n-term rational approximation.

B-spaces. We begin with the definition of a collection of spaces (B-spaces) needed for the theory of nonlinear piecewise polynomial approximation in $L_p(\mathbb{R}^2)$ $(0 < p < \infty)$. In [6] they are termed "skinny" B-spaces.

Taking into consideration our further needs, we shall be assuming in the following that $\mathcal{T}$ is an LR-triangulation or an SLR-triangulation (see Section 2.1). Throughout this section we assume that $0 < p < \infty$, $\alpha > 0$, $k \geqslant 1$, and τ is determined from $1/\tau := \alpha + 1/p$.

Definition 2.9. The B-spaces $\mathcal{B}_\tau^{\alpha k}(\mathcal{T})$ is defined as the set of all functions $f \in L_p(\mathbb{R}^2)$ such that

$$\|f\|_{\mathcal{B}_\tau^{\alpha k}(\mathcal{T})} := \left(\sum_{\triangle \in \mathcal{T}} (|\triangle|^{-\alpha} w_k(f, \triangle)_\tau)^\tau \right)^{1/\tau} < \infty, \tag{2.23}$$

where $w_k(f, \triangle)$ is a kth modulus of smoothness of f on $\triangle$ (see (2.21)).

Whitney's estimate (Proposition 2.8) implies

$$\|f\|_{\mathcal{B}_\tau^{\alpha k}(\mathcal{T})} \approx \left(\sum_{\triangle \in \mathcal{T}} (|\triangle|^{-\alpha} E_k(f, \triangle)_\tau)^\tau \right)^{1/\tau}. \tag{2.24}$$

Nonlinear piecewise polynomial approximation. Let $\Sigma_n^k(\mathcal{T})$, $k \geqslant 1$, denote the nonlinear set of all n-term piecewise polynomial function of the form

$$S = \sum_{\Delta \in \Lambda_n} \mathbb{1}_\Delta \cdot P_\Delta,$$

where $P_\Delta \in \Pi_k$, $\Lambda_n \subset \mathcal{T}$, $\#\Lambda_n \leqslant n$, and Λ_n may vary with S. We denote by $\sigma_n(f, \mathcal{T})_p$ the error of L_p-approximation to $f \in L_p(\mathbb{R}^2)$ from $\Sigma_n^k(\mathcal{T})$:

$$\sigma_n(f, \mathcal{T})_p := \inf_{S \in \Sigma_n^k(\mathcal{T})} \| f - S \|_p. \tag{2.25}$$

In [6] for the characterization of the approximation spaces generated by $(\sigma_n(f, \mathcal{T})_p)$ the machinery of Jackson–Bernstein estimates and interpolation are used.

Proposition 2.10 (*Jackson estimate*). *If $f \in \mathcal{B}_\tau^{\alpha k}(\mathcal{T})$, then*

$$\sigma_n(f, \mathcal{T})_p \leqslant c n^{-\alpha} \| f \|_{\mathcal{B}_\tau^{\alpha k}(\mathcal{T})}$$

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

Proposition 2.11 (*Bernstein estimate*). *If $S \in \Sigma_n^k(\mathcal{T})$, then*

$$\| S \|_{\mathcal{B}_\tau^{\alpha k}(\mathcal{T})} \leqslant c n^\alpha \| S \|_p \tag{2.26}$$

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

Denote by $A_q^\gamma := A_q^\gamma(L_p, \mathcal{T})$ the approximation space generated by $(\sigma_n(f, \mathcal{T})_p)$, consisting of all functions $f \in L_p$ such that

$$\| f \|_{A_q^\gamma} := \| f \|_p + \left(\sum_{n=1}^{\infty} (n^\gamma \sigma_n(f))^q \frac{1}{n} \right)^{1/q} < \infty \tag{2.27}$$

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

The following characterization of the approximation spaces A_q^γ follows in a standard way by Propositions 2.10–2.11.

Proposition 2.12. *If $0 < \gamma < \alpha$ and $0 < q \leqslant \infty$, then*

$$A_q^\gamma(L_p, \mathcal{T}) = (L_p, \mathcal{B}_\tau^{\alpha k}(\mathcal{T}))_{\frac{\gamma}{\alpha}, q}$$

with equivalent norms. Here $(X, Y)_{\lambda, q}$ denotes the real interpolation space between the spaces X and Y (see e.g. [1]).

Denote

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

where the infimum is taken over all LR-triangulations $\mathcal{T}$ with fixed parameters. The following result is immediate from Proposition 2.10.

Proposition 2.13. *Suppose* $\inf_{\mathcal{T}} \|f\|_{\mathcal{B}^{\alpha k}_{\tau}(\mathcal{T})} < \infty$, *where the infimum is taken over all LR-triangulations with fixed parameters, and let* $f \in L_p(\mathbb{R}^2)$. *Then*

$$\sigma_n(f)_p \leqslant cn^{-\alpha} \inf_{\mathcal{T}} \|f\|_{\mathcal{B}^{\alpha k}_{\tau}(\mathcal{T})}.$$

For more details, see [6].

2.4. Maximal functions

In this section we introduce and explore two types of maximal functions. They will be our main vehicle in proving out results for nonlinear n-term rational approximation.

Definition 2.14. Let $\mathcal{T}$ be a multilevel triangulation in $\mathbb{R}^2$ (for the definition, see Section 2.1). For a Lebesgue measurable function *f*, defined on $\mathbb{R}^2$, and $s > 0$, we define the maximal function $\mathcal{M}^s_{\mathcal{T}} f$ by

$$(\mathcal{M}^s_{\mathcal{T}} f)(x) := \sup_{\theta \in \Theta : x \in \theta} \left(\frac{1}{|\theta|} \int_{\theta} |f(y)|^s dy \right)^{1/s} \tag{2.28}$$

where the sup is taken over all cells $\theta \in \Theta$ containing *x*.

We next associate with any triangle $\triangle \subset \mathbb{R}^2$ a *collection of ellipses* $\mathcal{E}_{\triangle}$, which will be used in the definition of another type of maximal function. Let $\triangle^{\diamond}$ be a fixed equilateral reference triangle of side length one. Denote by B^- the circle inscribed in $\triangle^{\diamond}$ and by B^+ the circle circumscribed around $\triangle^{\diamond}$.

For a given triangle $\triangle$, let **A** be an affine transform which maps $\triangle^{\diamond}$ one-to-one onto $\triangle$. Denote $E^- = \mathbf{A}(B^-)$ and $E^+ = \mathbf{A}(B^+)$, which are apparently ellipses. It is also readily seen that E^- can be obtained by dilating and shifting E^+. Now, we let $\mathcal{E}_{\triangle}$ denote the set of all ellipses in $\mathbb{R}^2$ which can be obtained by dilating and shifting E^- or E^+.

Definition 2.15. Suppose $\triangle$ is a fixed triangle in $\mathbb{R}^2$ and $s > 0$. For any Lebesgue measurable function *f*, we define the maximal function $\mathcal{M}^s_{\mathcal{E}_{\triangle}} f$ by

$$(\mathcal{M}^s_{\mathcal{E}_{\triangle}} f)(x) := \sup_{E \in \mathcal{E}_{\triangle} : x \in E} \left(\frac{1}{|E|} \int_{E} |f(y)|^s dy \right)^{1/s} \tag{2.29}$$

where the sup is taken over all ellipses $E \in \mathcal{E}_{\triangle}$ containing *x*.

We first note that if $\triangle$ is an equilateral triangle and $s = 1$, then $\mathcal{M}^s_{\mathcal{E}_{\triangle}} f$ is the standard maximal function.

If $s = 1$, we denote $\mathcal{M}_{\mathcal{T}} f := \mathcal{M}^1_{\mathcal{T}} f$ and $\mathcal{M}_{\mathcal{E}_{\triangle}} f := \mathcal{M}^1_{\mathcal{E}_{\triangle}} f$. Note that $\mathcal{M}^s_{\mathcal{T}} f = (\mathcal{M}_{\mathcal{T}} |f|^s)^{1/s}$.

Remark 2.16. One of the most important properties of the maximal functions $\mathcal{M}^s_{\mathcal{T}} f$ and $\mathcal{M}^s_{\mathcal{E}_{\triangle}} f$ is that they are invariant under affine transforms. Thus if **A** is an arbitrary affine transform on $\mathbb{R}^2$, then

$$(\mathcal{M}^s_{\mathcal{T}} f)(x) = (\mathcal{M}^s_{\mathbf{A}(\mathcal{T})} f(A^{-1}))(\mathbf{A}(x)), \quad x \in \mathbb{R}^2,$$

where $\mathbf{A}(\mathcal{T}) := \{\mathbf{A}(\triangle) : \triangle \in \mathcal{T}\}$. The maximal functions $\mathcal{M}^s_{\mathcal{E}_\triangle} f$ are invariant in a similar sense.

Recall that if $\mathcal{T}$ is an SLR-triangulation (LR-triangulation), then $\mathbf{A}(\mathcal{T})$ is also an SLR-triangulation (LR-triangulation) with the same parameters. Consequently, the set of all maximal functions $\{\mathcal{M}^s_{\mathcal{T}}\}$, where the $\mathcal{T}$'s are SLR-triangulations with the same fixed parameters is invariant under affine transforms.

The next theorem provides a very important relation between the above defined maximal functions.

Theorem 2.17. *Let $\mathcal{T}$ be an SLR-triangulation and let $s > 0$. Then there exists $s' > 0$, depending only on s and the parameters of $\mathcal{T}$ such that if $\triangle \in \mathcal{T}$, then*

$$(\mathcal{M}^{s'}_{\mathcal{E}_\triangle} \mathbb{1}_\triangle)(x) \leqslant c(\mathcal{M}^s_{\mathcal{T}} \mathbb{1}_\triangle)(x), \quad x \in \mathbb{R}^2, \tag{2.30}$$

where c depends only on s and the parameters of $\mathcal{T}$. Here s' $(s' < s)$ can be defined e.g. by $s' := s\ \ln(1/\rho_1)/[\ln(1/\rho_1) + 3N_0\ \ln(1/\vartheta)]$, *where ϑ and ρ_1 are from Theorem* 2.4.

Proof. An important ingredient in the proof of this theorem will be the fact that (2.30) is invariant under affine transforms (see Remark 2.16).

Suppose $\triangle \in \mathcal{T}_m$ $(m \in \mathbb{Z})$ and let $x \in \mathbb{R}^2$. Two cases are to be considered here.

Case 1: $x \in \operatorname{star}^1_m(\triangle)$. Evidently, $(\mathcal{M}^{s'}_{\mathcal{E}_\triangle} \mathbb{1}_\triangle)(x) \leqslant \|\mathbb{1}_\triangle\|_{L_\infty} = 1$. On the other hand, by the definition of $\operatorname{star}^1_m(\triangle)$ in (2.14) there exists a cell $\theta \in \mathcal{T}_m$ such that $x \in \theta$ and $\triangle \subset \theta$. Here $\triangle$ is one of the triangles in $\mathcal{T}_m$ which make up θ. Then using Definition 2.14 we obtain

$$(\mathcal{M}^s_{\mathcal{T}} \mathbb{1}_\triangle)(x) \geqslant \left(\frac{1}{|\theta|}\int_\theta |\mathbb{1}_\triangle(y)|^s\,dy\right)^{1/s} \geqslant (|\triangle|/|\theta|)^{1/s} \geqslant c > 0,$$

where we used that $|\theta| \leqslant c|\triangle|$ which follows by conditions (i)–(ii) on SLR-triangulations (see also (2.1)–(2.2)). The above estimates imply (2.30).

Case 2: $x \in \mathbb{R}^2 \setminus \operatorname{star}^1_m(\triangle)$. Let l $(l \leqslant m)$ be the minimum level such that $x \in \mathbb{R}^2 \setminus \operatorname{star}^1_l(\triangle)$. The existence of such level $l \leqslant m$ follows by property (f) of SLR-triangulations and Proposition 2.1. Then $x \in \operatorname{star}^1_{l-1}(\triangle)$.

Denote by $\triangle_0$ the unique triangle in $\mathcal{T}_l$ such that $\triangle \subset \triangle_0$. Since (2.30) is invariant under affine transforms, we may assume that $\triangle_0$ is an equilateral triangle of side length one. Let $e_{\max}$ be the maximal edge of $\triangle$ and write $a := |e_{\max}|$. Also, let h be (the length of) the height in $\triangle$ to $e_{\max}$.

Since $x \in \operatorname{star}^1_{l-1}(\triangle)$, then there exists $\theta \in \mathbf{\Theta}_{l-1}$ such that $x \in \theta$ and $\triangle \subset \theta$. By conditions (i)–(ii) on SLR-triangulations and since $\triangle_0$ is an equilateral triangle of side length one, it follows that $|\theta| \approx 1$. Consequently,

$$(\mathcal{M}^s_{\mathcal{T}} \mathbb{1}_\triangle)(x) \geqslant \left(\frac{1}{|\theta|}\int_\theta |\mathbb{1}_\triangle(y)|^s\,dy\right)^{1/s} \geqslant \left(\frac{|\triangle|}{|\theta|}\right)^{1/s} \geqslant c|\triangle|^{1/s} \geqslant c'(ah)^{1/s}. \tag{2.31}$$

To estimate $(\mathcal{M}^{s'}_{\mathcal{E}_\triangle}\mathbb{1}_\triangle)(x)$ from above, we first show that $d := \mathrm{dist}(\triangle, x) \geqslant c_1$ for some constant $c_1 > 0$, where $\mathrm{dist}(\triangle, x)$ is the Euclidean distance from x to $\triangle$ in $\mathbb{R}^2$. Since $\triangle \subset \triangle_0$ and $\triangle_0 \in \mathcal{T}_l$, then $\mathrm{star}^1_l(\triangle) = \mathrm{star}^1_l(\triangle_0)$. Therefore, it suffices to show that $\mathrm{dist}(\triangle_0, y) \geqslant c_1$ for all y from the boundary of $\mathbb{R}^2 \setminus \mathrm{star}^1_l(\triangle_0)$. But the boundary of $\mathbb{R}^2 \setminus \mathrm{star}^1_l(\triangle_0)$ is obviously the same as the boundary $\partial(\mathrm{star}^1_l(\triangle_0))$ of $\mathrm{star}^1_l(\triangle_0)$. A simple geometric argument shows that the Euclidean distance between $\triangle_0$ and $\partial(\mathrm{star}^1_l(\triangle_0))$ is bounded below by the minimum height to an edge in a triangle from $\mathcal{T}_l$ which has a common vertex with $\triangle_0$. But by condition (ii) on SLR-triangulations, $\min\mathrm{angle}\,(\triangle') \geqslant \beta > 0$ and hence $|\min\mathrm{edge}\,(\triangle')| \geqslant c(\beta) > 0$ for all triangles $\triangle' \in \mathcal{T}_l$ which have a common vertex with $\triangle_0$. Here we use that $\triangle_0 \in \mathcal{T}_l$ is an equilateral triangle of side length one. These inequalities yield that the minimum height to an edge in a triangle from $\mathcal{T}_l$ which has a common vertex with $\triangle_0$ is bounded below by a constant $c_1 > 0$ depending only on β, which in turn implies $\mathrm{dist}(\triangle, x) \geqslant \mathrm{dist}(\triangle_0, \partial(\mathrm{star}^1_l(\triangle_0))) \geqslant c_1$.

Let $\mathbf{A}$ be an affine transform which maps an equilateral reference triangle $\triangle^\diamond$ one-to-one onto $\triangle$. Let $E^\pm$ be the images of the inscribed $(-)$ and subscribed $(+)$ circles of $\triangle^\diamond$ (see the construction above Definition 2.15). Evidently the major diameters of $E^\pm$ are $\approx a$ and the minor diameters of $E^\pm$ are $\approx h$. Let E^* be the smallest ellipse in $\mathcal{E}_\triangle$ such that $x \in E^*$ and $\triangle \cap E^* \neq \emptyset$. Denote by D and H the major and minor diameters of E^*. Evidently, $D \geqslant d \geqslant c_1$, where $d := \mathrm{dist}(\triangle, x)$. Since E^* can be obtained from E^+ (or E^-) by a dilation and a shift, then $H/D \approx h/a$ and hence $|E^*| \geqslant cDH \geqslant cD^2h/a \geqslant cd^2h/a \geqslant c''h/a$.

By the definition of E^*, for any ellipse $E \in \mathcal{E}_\triangle$ such that $x \in E$ and $\triangle \cap E \neq \emptyset$ we have $|E| \geqslant |E^*|$. Then by Definition 2.15, it follows that

$$(\mathcal{M}^{s'}_{\mathcal{E}_\triangle}\mathbb{1}_\triangle)(x) = \sup_{E\in\mathcal{E}_\triangle :\, x\in E,\ \triangle\cap E\neq\emptyset} \left(\frac{1}{|E|}\int_E \mathbb{1}_\triangle(x)\,dx\right)^{1/s'} \leqslant \left(\frac{|\triangle|}{|E^*|}\right)^{1/s'} \leqslant c\left(\frac{a|\triangle|}{h}\right)^{1/s'} \leqslant c_2 a^{2/s'}. \tag{2.32}$$

Taking into account (2.31)–(2.32), it remains to show that $a^{2/s'} \leqslant c(ah)^{1/s}$ or equivalently

$$a^{2/s'-2/s} \leqslant c(h/a)^{1/s}. \tag{2.33}$$

Denote $v := m - l$. Using Theorem 2.4(b), it follows that

$$a = |\max\mathrm{edge}\,(\triangle)| \leqslant \rho_1^{[v/3N_0]}|\max\mathrm{edge}\,(\triangle_0)| = \rho_1^{[v/3N_0]}. \tag{2.34}$$

Let $\alpha := \min\mathrm{angle}\,(\triangle)$. By Theorem 2.4 (a), $\alpha \geqslant \vartheta^v \min\mathrm{angle}\,(\triangle_0) \geqslant c\vartheta^v$, which yields

$$h/a \geqslant (1/2)\sin\alpha \geqslant (1/\pi)\alpha \geqslant c\vartheta^v. \tag{2.35}$$

We are now prepared to show that (2.33) holds true. If $0 \leqslant v < 6N_0$, then by (2.34)–(2.35) we have $a \leqslant 1$ and $h/a \geqslant c$. Hence (2.33) holds with some constant $c > 0$ depending only on s and the parameters of $\mathcal{T}$. Suppose $v \geqslant 6N_0$. Then $[v/3N_0] \geqslant v/6N_0$ and using (2.34)–(2.35) we obtain

$$a^{2/s'-2/s} \leqslant \rho_1^{[v/3N_0](2/s'-2/s)} \leqslant \rho_1^{(v/3N_0)(1/s'-1/s)} = \vartheta^{v/s} \leqslant c(h/a)^{1/s},$$

where the constant $c > 0$ is again depending only on s and the parameters of $\mathcal{T}$. Thus in both cases (2.33) holds true and this completes the proof. $\square$

The maximal inequality. Here we extend the usual L_p maximal inequalities (boundedness of maximal operators) to maximal functions generated by multilevel nested triangulations. In essence these are well-known results. We present in the form that we need.

Suppose that $d : \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty)$ is a *quasi-distance in* $\mathbb{R}^n$, i.e. d satisfies

$$\begin{aligned} &\text{(a)} \quad d(x, y) = 0 \Longleftrightarrow x = y, \\ &\text{(b)} \quad d(x, y) = d(y, x), \\ &\text{(c)} \quad d(x, z) \leqslant c(d(x, y) + d(y, z)) \quad \text{with } c \geqslant 1. \end{aligned} \tag{2.36}$$

We denote by $B(y, a)$ $(a > 0)$ the "ball" with respect to this quasi-distance of radius a centered at y, that is, $B(y, a) := \{x : d(x, y) < a\}$.

In this setting the maximal function (operator) $\mathcal{M}^s$ is defined by

$$(\mathcal{M}^s f)(x) := \sup_{B : x \in B} \left(\frac{1}{|B|} \int_B |f(y)|^s \, dy \right)^{1/s}, \tag{2.37}$$

where the infimum is over all balls B containing x.

The Fefferman–Stein vector-valued maximal inequality (see [5,10]) reads as follows:

Proposition 2.18. *If* $0 < p < \infty$, $0 < q \leqslant \infty$, *and* $0 < s < \min\{p, q\}$, *then for any sequence of functions* $(f_j)_{j=1}^{\infty}$ *on* $\mathbb{R}^2$

$$\left\| \left(\sum_{j=1}^{\infty} |\mathcal{M}^s f_j|^q \right)^{1/q} \right\|_p \leqslant c \left\| \left(\sum_{j=1}^{\infty} |f_j|^q \right)^{1/q} \right\|_p, \tag{2.38}$$

where c depends only on p, q, and s.

As a matter of fact, in [5,10] the maximal inequality (2.38) is stated and proved in the case $s = 1$ but since $\mathcal{M}^s f = (\mathcal{M}^1 |f|^s)^{1/s}$ the proposition follows.

Definition 2.19. For a given LR-triangulation $\mathcal{T}$, we define a quasi-distance $d_{\mathcal{T}} : \mathbb{R}^2 \times \mathbb{R}^2 \to [0, \infty)$ by

$$d_{\mathcal{T}}(x, y) := \inf\{|\theta| : \theta \in \Theta \text{ and } x, y \in \theta\}. \tag{2.39}$$

Lemma 2.20. *The mapping* $d_{\mathcal{T}} : \mathbb{R}^2 \times \mathbb{R}^2 \to [0, \infty)$ *defined in* (2.39) *is a quasi-distance in* $\mathbb{R}^2$.

Proof. Condition (a) on quasi-distances (see (2.36)) follows by the properties of the LR-triangulations (see Section 2.1). Condition (b) is obvious.

To prove that condition (c) holds let x, y, z be three distinct points in $\mathbb{R}^2$. Assume that $d(x, z) = |\theta'|$, where $\theta' \in \Theta_m$ is a cell containing x, z. Similarly let $d(y, z) = |\theta''|$ for some cell $\theta'' \in \Theta_n$ which contains both y and z. Without loss of generality we assume that $m \leqslant n$. Obviously x and z lie in triangles in $\mathcal{T}_m$ with a common vertex (or in the same triangle). Since $m \leqslant n$, the same is true for y and z. In other words there exist triangles $\triangle_1, \triangle_2 \in \mathcal{T}_m$ which can be connected with < 2 intermediate triangles from $\mathcal{T}_m$ (with common vertices), so that $x \in \triangle_1$, $y \in \triangle_2$. By Proposition 2.1 that there exists $\theta \in$

Θ_{m-2N_0} such that $\triangle_1, \triangle_2 \subset \theta$ and hence $d(x, y) \leqslant |\theta|$. By properties (2.1)–(2.2) of the LR-triangulations there exists a constant $c := c(\delta, r, N_0)$ such that $|\theta| \leqslant c|\theta'|$. Consequently, $d(x, y) \leqslant c(d(x, z) + d(z, y))$. $\square$

Denote by $\mathcal{M}^s_{d_\mathcal{T}}$ the maximal function generated by the quasi-distance defined in (2.39).

Lemma 2.21. *If $\mathcal{T}$ is an LR-triangulation, then for any measurable function f*

$$\mathcal{M}^s_\mathcal{T} f(x) \approx \mathcal{M}^s_{d_\mathcal{T}} f(x), \quad x \in \mathbb{R}^2, \tag{2.40}$$

where the constants of equivalence depend only on s and the parameters of $\mathcal{T}$.

Proof. Fix a ball $B = B(x, \delta)$, $x \in \mathbb{R}^2$, $\delta > 0$. Let m be the minimum level such that for some $\theta' \in \Theta_m$, we have $x \in \theta' \subset B$. Since every $\theta \in \Theta_l$ with $l > m$ is contained in a cell from Θ_m,

$$B(x, \delta) = \bigcup_{|\theta|<\delta,\, x\in\theta} \theta \subset \bigcup_{\theta\in\Theta_m,\, x\in\theta} \theta \subset \mathrm{star}^2_m(x).$$

But any two triangles from $\mathcal{T}_m$ which are contained in $\mathrm{star}^2_m(x)$ can be connected by $< 2^2$ intermediate edges from $\mathcal{E}_m$. Then by Proposition 2.1 it follows that $\mathrm{star}^2_m(x) \subset \theta''$ for some $\theta'' \in \Theta_{m-4N_0}$. Thus $\theta' \subset B \subset \theta''$ with $\theta' \in \Theta_m$ and $\theta'' \in \Theta_{m-4N_0}$. By properties (i), (ii) (see (2.1), (2.2)), we conclude that $|\theta''| \leqslant c|\theta'|$ with c depending on r, δ, and N_0.

In the other direction, for any cell $\theta \in \Theta_n$ ($n \in \mathbb{Z}$) with "central" vertex v, we have $\theta \subset \mathrm{star}^2_n(v)$. Let $\delta' = \max\{|\theta| : \theta \subset \mathrm{star}^2_n(v)\}$. Then

$$\theta \subset B(v, \delta') = \bigcup_{|\theta^\diamond|<\delta',\, v\in\theta^\diamond} \theta^\diamond \subset \mathrm{star}^2_n(v),$$

which yields $|B(v, \delta')| \leqslant c|\theta|$. This completes the proof. $\square$

We now couple Proposition 2.18 with the above lemma to obtain the following modification of the Fefferman–Stein maximal inequality:

Proposition 2.22. *Let $\mathcal{T}$ be an LR-triangulation of $\mathbb{R}^2$. If $0 < p < \infty$, $0 < q \leqslant \infty$, and $0 < s < \min\{p, q\}$, then for any sequence of functions $(f_j)_{j=1}^\infty$ on $\mathbb{R}^2$*

$$\left\| \left(\sum_{j=1}^\infty |\mathcal{M}^s_\mathcal{T} f_j|^q \right)^{1/q} \right\|_p \leqslant c \left\| \left(\sum_{j=1}^\infty |f_j|^q \right)^{1/q} \right\|_p, \tag{2.41}$$

where c depends only on p, q, s, and the parameters of $\mathcal{T}$.

3. Main results

We denote by $\mathcal{R}_n$ the set of all n-term rational functions on $\mathbb{R}^2$ of the form

$$R = \sum_{i=1}^n r_i,$$

where each r_i is of the form

$$r_i = \prod_{\mu=1}^{6} \frac{a_\mu x_1 + b_\mu x_2 + c_\mu}{1 + (\alpha_\mu x_1 + \beta_\mu x_2 + \gamma_\mu)^2}$$

with $a_\mu, b_\mu, c_\mu, \alpha_\mu, \beta_\mu, \gamma_\mu \in \mathbb{R}$.

Denote by $R_n(f)_p$ the error of L_p-approximation to *f* from $\mathcal{R}_n$:

$$R_n(f)_p := \inf_{R\in\mathcal{R}_n} \|f - R\|_p.$$

Clearly, each $R \in \mathcal{R}_n$ depends on $\leqslant 36n$ parameters and $\mathcal{R}_n$ is a nonlinear set, however, $c\mathcal{R}_n = \mathcal{R}_n$ $(c \neq 0)$ and $\mathcal{R}_n + \mathcal{R}_m = \mathcal{R}_{n+m}$. A fundamental property of $\mathcal{R}_n$ is that it is invariant under affine transforms, i.e. if $R \in \mathcal{R}_n$, then $R \circ \mathbf{A} \in \mathcal{R}_n$ for every affine transform $\mathbf{A}$.

Our primary goal in this chapter is to relate n-term rational approximation and n-term piecewise polynomial approximation. We shall use all the notation from Section 2.3. Throughout this section, we assume that $\mathcal{T}$ is an SLR-triangulation on $\mathbb{R}^2$ (see Section 2.1).

The following theorem contains our main result.

Theorem 3.1. *Let* $f \in L_p(\mathbb{R}^2)$, $0 < p < \infty$, $\alpha > 0$, *and* $k \geqslant 1$. *Then*

$$R_n(f)_p \leqslant cn^{-\alpha} \left(\sum_{m=1}^{n} \frac{1}{m} (m^\alpha \sigma_m(f, \mathcal{T})_p)^{p^*} + \|f\|_p^{p^*} \right)^{1/p^*}, \quad n = 1, 2, \ldots, \tag{3.1}$$

where $p^* = \min\{1, p\}$ *and c depends only on* α, *p*, *k*, *and the parameters of* $\mathcal{T}$.

It is an important observation that in Theorem 3.1 there is no restriction on $\alpha > 0$. The next corollary follows immediately from the above theorem.

Corollary 3.2. *If* $\sigma_n(f, \mathcal{T})_p = O(n^{-\gamma})$ *for an arbitrary SLR-triangulation* $\mathcal{T}$, $0 < p < \infty$, *and* $\gamma > 0$, *then* $R_n(f)_p = O(n^{-\gamma})$.

Combining the Jackson estimate for n-term piecewise polynomial approximation from Proposition 2.10 with Theorem 3.1, we infer the following result.

Corollary 3.3. *If* $f \in \bigcap_{\mathcal{T}} B_\tau^{\alpha k}(\mathcal{T})$, *where* $\alpha > 0$, $1/\tau := \alpha + 1/p$, $0 < p < \infty$, *then*

$$R_n(f)_p \leqslant cn^{-\alpha} \inf_{\mathcal{T}} \|f\|_{B_\tau^{\alpha k}(\mathcal{T})}, \tag{3.2}$$

where the infimum is taken over all SLR-triangulation with some fixed parameters.

We shall deduce Theorem 3.1 from the following result.

Theorem 3.4. *For each* $S \in \Sigma_m^k(\mathcal{T})$, $m \geqslant 1$, *and* $n \geqslant 1$, *there exists* $R \in \mathcal{R}_n$ *such that*

$$\|S - R\|_p \leqslant c \exp(-(n/m)^{1/12})\|S\|_p \tag{3.3}$$

where c depends only on p, k, and the parameters of $\mathcal{T}$.

The main vehicles in the proof of Theorem 3.4 will be the anisotropic version of the Fefferman–Stein vector-valued maximal inequality which was given in Proposition 2.22 and the following lemma which rests on the result of Newman [7] on the rational approximation of $|x|$.

Proposition 3.5. *For each* $\gamma > 0$, $0 < \delta < 1$, *and* μ *a positive integer, there exists a univariate rational function* σ *such that*

$$\deg\, \sigma \leqslant c\, \ln(e + 1/\gamma) \ln(e + 1/\delta) + 4\mu, \tag{3.4}$$

$$0 \leqslant 1 - \sigma(t) < \gamma \quad \textit{for } |t| \leqslant 1 - \delta, \tag{3.5}$$

$$0 \leqslant \sigma(t) < \gamma \left(\frac{1}{1 + |t|}\right)^{4\mu} \quad \textit{for } |t| \geqslant 1 \quad \textit{and} \tag{3.6}$$

$$0 \leqslant \sigma(t) \leqslant 1 \quad \textit{for } t \in \mathbb{R}, \tag{3.7}$$

where c is an absolute constant. Moreover, σ *has only simple poles. It follows by* (3.6) *that if* $\sigma = P/Q$ (*P*, *Q polynomials*) *then* $\deg\, Q \geqslant \deg\, P + 4\mu$.

For later use, we include the following lemma.

Lemma 3.6. *Suppose* $\sigma = P/Q$ *is a univariate rational function degree* $\leqslant l$ *such that* $\deg Q \geqslant \deg P + k + 1$ $(k \geqslant 1)$ *and* σ *has only simple poles. Let* $\tilde{P} \in \Pi_k(\mathbb{R}^2)$. *Suppose that* $\triangle := [v_1, v_2, v_3]$ *is a triangle in* $\mathbb{R}^2$ *and* $a_i x_1 + b_i x_2 + c_i = 0$ $(i = 1, 2, 3)$ *is an equation of the straight line containing the edge of* $\triangle$ *opposite to the vertex* v_i. *Denote* $T_i(x) = a_i x_1 + b_i x_2 + c_i$. *Then*

$$\prod_{i=1}^{3} \sigma(T_i)\tilde{P} \in \mathcal{R}_{cl^3}. \tag{3.8}$$

Proof. Each $x \in R^2$ has a representation of the form

$$x = b_1(x)v_1 + b_2(x)v_2 + b_3(x)v_3, \quad b_1(x) + b_2(x) + b_3(x) = 1,$$

where $b_1(x)$, $b_2(x)$, and $b_3(x)$ are the *barycentric coordinates* [4] of *x* with respect to $\triangle$. It is readily seen that $b_i(x) = A_i T_i(x)$. Then the Bernstein–Bezier representation of $\tilde{P}(x)$ is of the form

$$\tilde{P}(x) = \sum_{0 \leqslant \alpha+\beta+\gamma<k} c_{\alpha,\beta,\gamma} b_1(x)^\alpha b_2(x)^\beta b_3(x)^\gamma = \sum_{0 \leqslant \alpha+\beta+\gamma<k} d_{\alpha,\beta,\gamma} T_1(x)^\alpha T_2(x)^\beta T_3(x)^\gamma.$$

Therefore,

$$\prod_{i=1}^{3} \sigma(T_i(x))\tilde{P}(x) = \sum_{0 \leqslant \alpha+\beta+\gamma<k} d_{\alpha,\beta,\gamma}[T_1(x)^\alpha \sigma(T_1(x))][T_2(x)^\beta \sigma(T_2(x))][T_3(x)^\gamma \sigma(T_3(x))].$$

Since $\deg Q \geqslant \deg P + k + 1$ and σ has only simple poles then $T_1(x)^\alpha \sigma(T_1(x))$ has a representation of the form

$$T_1(x)^\alpha \sigma(T_1(x)) = \sum_{v=1}^{\mu_1} \frac{u_{1,v}T_1(x) + v_{1,v}}{t_{1,v} + (T_1(x) + s_{1,v})^2} \quad \text{with} \quad \mu_1 \leqslant l/2.$$

Evidently, $T_2(x)^\beta \sigma(T_2(x))$ and $T_3(x)^\gamma \sigma(T_3(x))$ have similar representations. Consequently,

$$\prod_{i=1}^{3} \sigma(T_i(x))\tilde{P}(x) = \sum_{0 \leqslant \alpha+\beta+\gamma<k} d_{\alpha,\beta,\gamma} \prod_{j=1}^{3} \sum_{v=1}^{\mu_j} \frac{u_{j,v}T_j(x) + v_{j,v}}{t_{j,v} + (T_j(x) + s_{j,v})^2} = \sum_{\mu=1}^{cl^3} \prod_{i=1}^{3} \frac{a_{i,\mu}x_1 + b_{i,\mu}x_2 + c_{i,\mu}}{1 + (\alpha_{i,\mu}x_1 + \beta_{i,\mu}x_2 + \gamma_{i,\mu})^2},$$

where $a_{i,\mu}, b_{i,\mu}, c_{i,\mu}, \alpha_{i,\mu}, \beta_{i,\mu}, \gamma_{i,\mu} \in \mathbb{R}$. The proof is complete. □

With the next lemma we show that every piecewise polynomial function $S \in \Sigma_n^k(\mathcal{T})$ can be represented as a piecewise polynomial function on $\leqslant cn$ non-overlapping "rings".

Lemma 3.7. *Suppose* $S := \sum_{\triangle\in\Lambda} \mathbb{1}_\triangle P_\triangle$, *where* $P_\triangle \in \Pi_k$, $\Lambda \subset \mathcal{T}$, *and* $\#\Lambda \leqslant n$. *Then S can be represented in the form*

$$S := \sum_{\triangle\in\tilde{\Lambda}} \mathbb{1}_{K_\triangle} P_{K_\triangle}, \tag{3.9}$$

where $\tilde{\Lambda} \subset \mathcal{T}$, $\#\tilde{\Lambda} \leqslant cn$ *with c depending only on the parameters of* $\mathcal{T}$, *each "ring"* $K_\triangle$ *is of the form* $K_\triangle = \triangle$ *or* $K_\triangle = \triangle \setminus \triangle'$, $\triangle' \in \mathcal{T}$, *and* $K_{\triangle_1}^\circ \cap K_{\triangle_2}^\circ = \emptyset$ *if* $\triangle_1 \neq \triangle_2$.

Proof. Since the levels of $\mathcal{T}$ are nested, there is a natural tree structure in $\mathcal{T}$ induced by the inclusion relation. Namely, if $\triangle_1, \triangle_2 \in \mathcal{T}$ then $\triangle_1 \subset \triangle_2$ or $\triangle_2 \subset \triangle_1$ or $\triangle_1^\circ \cap \triangle_2^\circ = \emptyset$. The set Λ generates a subtree in $\mathcal{T}$. Let $\mathcal{T}_\Lambda$ be the set of all triangles $\triangle \in \mathcal{T}$ for which there exist two triangles $\triangle_1, \triangle_2 \in \Lambda$ such that $\triangle_1 \subset \triangle \subset \triangle_2$. Clearly, $\Lambda \subset \mathcal{T}_\Lambda$.

We shall make the distinction between several types of triangles in $\mathcal{T}_\Lambda$. We say that $\triangle \in \mathcal{T}_\Lambda$ is a *leaf* in $\mathcal{T}_\Lambda$ if $\triangle$ does not contain any other triangle in $\mathcal{T}_\Lambda$. We denote by Λ_ℓ the set of all leaves in Λ. Evidently, $\Lambda_\ell \subset \Lambda$.

We say that $\triangle \in \mathcal{T}_\Lambda$ is a *branching triangle* for $\mathcal{T}_\Lambda$ if $\triangle$ has at least two children in $\mathcal{T}_\Lambda$, i.e., if at least two children of $\triangle$ in $\mathcal{T}$ lie in Λ or have descendants in Λ. We denote by Λ_b the set of all branching triangles in $\mathcal{T}$. We also denote by Λ_b' the set of all children in $\mathcal{T}$ of branching triangles. We extend Λ to $\tilde{\Lambda} := \Lambda \cup \Lambda_b \cup \Lambda_b'$. It is easy to see that in every tree the number of the branching elements does not exceed the number of the leaves. Therefore, $\#\Lambda_b \leqslant \#\Lambda_\ell \leqslant n$ and $\#\Lambda_b' \leqslant cn$ since the number of children of a triangle is bounded by M_0. Thus $\#\tilde{\Lambda} \leqslant cn$.

We denote by $\tilde{\Lambda}_\ell$ the set of all leaves in the tree $\mathcal{T}_\Lambda \cup \Lambda_{b'}$.

For each triangle $\triangle \in \tilde{\Lambda} \setminus (\Lambda_b \cup \tilde{\Lambda}_\ell)$ we denote by $\tilde{\triangle}$ the unique largest triangle $\tilde{\triangle} \subset \triangle$ such that $\tilde{\triangle} \in \tilde{\Lambda}$ and $\tilde{\triangle} \neq \triangle$. Finally, we introduce *rings* generated by Λ as follows. For $\triangle \in \tilde{\Lambda}$, we define

$$K_\triangle := \begin{cases} \emptyset & \text{if} \quad \triangle \in \Lambda_b, \\ \triangle \setminus \tilde{\triangle} & \text{if} \quad \triangle \in \tilde{\Lambda} \setminus (\Lambda_b \cup \tilde{\Lambda}_\ell), \\ \triangle & \text{if} \quad \triangle \in \tilde{\Lambda}_\ell. \end{cases}$$

It is readily seen that $K^\circ_{\triangle_1} \cap K^\circ_{\triangle_2} = \emptyset$ if $\triangle_1, \triangle_2 \in \tilde{\Lambda}$ and $\triangle_1 \neq \triangle_2$. Also, since all children of branching triangles belong to $\tilde{\Lambda}$, we have

$$\triangle = \bigcup_{\triangle' \in \tilde{\Lambda}, \triangle' \subset \triangle} K_{\triangle'}, \quad \triangle \in \tilde{\Lambda} \tag{3.10}$$

and, hence,

$$\bigcup_{\triangle \in \tilde{\Lambda}} \triangle = \bigcup_{\triangle' \in \tilde{\Lambda}} K_{\triangle'}. \tag{3.11}$$

Evidently, S is a polynomial of degree $< k$ on each ring $K_\triangle$ and therefore (3.9) holds. $\square$

The next lemma provides the main step in the proof of Theorem 3.4.

Lemma 3.8. *Suppose* $\varphi := \mathbb{1}_K \cdot P_K$, *where* $K = \triangle \setminus \triangle'$, $\triangle' \subset \triangle$, $\triangle, \triangle' \in \mathcal{T}$, *and* $P_K \in \Pi_k$, $k \geqslant 1$. *Then for* $\lambda > 0$ *and* $s > 0$ *there exists a rational function* $R \in \mathcal{R}_l$ *with* $l \leqslant c \ln^{12}(e + 1/\lambda)$ *such that*

$$\|\varphi - R\|_p \leqslant c\lambda \|\varphi\|_p \tag{3.12}$$

and

$$|R(x)| \leqslant c\lambda |K|^{-\frac{1}{p}} \|\varphi\|_p (\mathcal{M}^s_{\mathcal{T}} \mathbb{1}_K)(x) \quad \textit{for} \quad x \in \mathbb{R}^2 \setminus K, \tag{3.13}$$

where c depends on p, k, s, and the parameters of $\mathcal{T}$.

Proof. Let $\triangle^\diamond$ be an equilateral reference triangle with side length one, centered at the origin. Denote by v_1, v_2, and v_3 the vertices of $\triangle^\diamond$. Let l_3^- be the straight line in $\mathbb{R}^2$ through v_1 and v_2. Also, let l_3^+ be the line through v_3 which is parallel to l_3^- and let S_3 denote the strip bounded by l_3^- and l_3^+. We similarly define the lines l_j^-, l_j^+ $(j = 1, 2)$ and the strips S_1, S_2. Further, we denote by T_j ($j = 1, 2, 3$), the linear function of the form $T_j(x) = a_1^j x_1 + a_2^j x_2 + a_3^j$, so that $T_j(l_j^-) = -1$ and $T_j(l_j^+) = 1$.

For the given $s > 0$, we select s' so that $1/s' := 1/s + 3N_0 \ln(1/\vartheta)/[2s \ln(1/\rho_1)]$, where ϑ and ρ_1 are the constants from Theorem 2.4 (see Theorem 2.17).

Let σ be the univariate rational function from Proposition 3.5 applied with $\gamma := \lambda$, $\delta := \lambda^p$ and $\mu := \lceil (k + 2/s')/4 \rceil + 1$. We define $\kappa_{\triangle^\diamond}(x) := \prod_{i=1}^3 \sigma(T_i(x))$. By (3.4), we have

$$\deg \sigma \leqslant c \ln\left(e + \frac{1}{\lambda}\right) \ln\left(e + \frac{1}{\lambda^p}\right) + 4\mu \leqslant c \ln^2\left(e + \frac{1}{\lambda}\right), \quad c := c(k, s, p). \tag{3.14}$$

By (3.7), it follows that

$$0 \leqslant \kappa_{\triangle^\diamond}(x) < 1 \quad \text{for } x \in \mathbb{R}^2. \tag{3.15}$$

Denote $\triangle_\delta^\diamond := (1-\delta)\triangle^\diamond$, i.e. $\triangle_\delta^\diamond := \{x \in \mathbb{R}^2 : x = (1-\delta)y,\ y \in \triangle^\diamond\}$. Then (3.5) implies

$$0 \leqslant 1 - \kappa_{\triangle^\diamond}(x) \leqslant \sum_{i=1}^{3} (1 - \sigma(T_i(x))) \leqslant 3\lambda, \quad x \in \triangle_\delta^\diamond. \tag{3.16}$$

Let $x \in \mathbb{R}^2 \setminus \triangle^\diamond$. If $|x| \leqslant 2$, then by (3.6) we have $\kappa_{\triangle^\diamond}(x) < c\lambda$. Let $|x| > 2$. By the symmetry we may assume that $T_i(x) > 1$ for $i = 1, 2$, or 3. Then since $|x| \leqslant c \operatorname{dist}(x, S_i)$, we have

$$\kappa_{\triangle^\diamond}(x) \leqslant \sigma(T_i(x)) \leqslant c\lambda \left(\frac{1}{1 + \operatorname{dist}(x, S_i)}\right)^{4\mu} \leqslant c\lambda \left(\frac{1}{1 + |x|}\right)^{4\mu}.$$

These estimates imply that

$$\kappa_{\triangle^\diamond}(x) \leqslant c\lambda \left(\frac{1}{1 + |x|}\right)^{4\mu}, \quad x \in \mathbb{R}^2 \setminus \triangle^\diamond. \tag{3.17}$$

Clearly the statement of the lemma is invariant under affine transforms (see Remark 2.16). So, without loss of generality we shall assume that $\triangle$ is an equilateral triangle of side length one, namely, $\triangle = \triangle^\diamond$. Suppose $\triangle' \subset \triangle$ is any triangle. Let $\triangle_\delta := \triangle_\delta^\diamond := (1 - \delta)\triangle^\diamond$. Set $\kappa_\triangle := \kappa_{\triangle^\diamond}$.

Let $\mathbf{A}$ be an affine transform mapping one-to-one $\triangle'$ onto $\triangle_\delta^\diamond$. Then $\mathbf{A}^{-1}(\triangle_\delta^\diamond) = \triangle'$. Denote $\triangle'_\delta := \mathbf{A}^{-1}(\triangle^\diamond)$. Then $\triangle' \subset \triangle'_\delta$ and it is readily seen that $|\triangle'_\delta \setminus \triangle'| \leqslant \delta$.

Now, we define $\kappa_{\triangle'} := \kappa_{\triangle^\diamond} \circ \mathbf{A}$, the composition of $\kappa_{\triangle^\diamond}$ and $\mathbf{A}$. By the properties of $\kappa_{\triangle^\diamond}$ and $\mathbf{A}$ it follows that

$$0 \leqslant \kappa_{\triangle'}(x) < 1 \quad \text{for} \quad x \in \mathbb{R}^2, \quad 0 \leqslant 1 - \kappa_{\triangle'}(x) \leqslant 3\lambda \quad \text{for} \quad x \in \triangle' \tag{3.18}$$

and

$$\kappa_{\triangle'}(x) \leqslant c\lambda \quad \text{for} \quad x \in \mathbb{R}^2 \setminus \triangle'_\delta. \tag{3.19}$$

Let $\varphi := \mathbb{1}_K \cdot P_K$, $P_K \in \Pi_k$ with $K := \triangle \setminus \triangle'$. We set

$$R := \kappa_\triangle (1 - \kappa_{\triangle'}) P_K.$$

Note that $R = \kappa_\triangle P_K - \kappa_\triangle \kappa_{\triangle'} P_K =: R_1 + R_2$. By Lemma 3.6 and (3.14) we have $R_1 \in \mathcal{R}_n$ with $n := c \ln^6(1+1/\lambda)$. It follows from the fact that the univariate rational function σ from Lemma 3.5 has only simple poles and by (3.14) together with Lemma 3.6 that $R_2 \in \mathcal{R}_m$ with $m := c \ln^{12}(1 + 1/\lambda)$ and hence $R \in \mathcal{R}_l$ with $l \leqslant c \ln^{12}(1 + 1/\lambda)$.

We use Lemma 2.7, (3.16), and (3.18) to conclude that

$$\begin{aligned}\|\varphi - R\|_{L_p(\triangle_\delta \setminus \triangle'_\delta)} &= \|1 - \kappa_\triangle(1 - \kappa_{\triangle'})\|_{L_\infty(\triangle_\delta \setminus \triangle'_\delta)} \|\varphi\|_p \\ &= \left(\|1 - \kappa_\triangle\|_{L_\infty(\triangle_\delta)} + \|\kappa_{\triangle'}\|_{L_\infty(\mathbb{R}^2 \setminus \triangle'_\delta)}\right) \|\varphi\|_p \leqslant c\lambda \|\varphi\|_p.\end{aligned} \tag{3.20}$$

Write $K_\delta := (\triangle \setminus \triangle_\delta) \cup [\triangle \cap (\triangle'_\delta \setminus \triangle')]$. Then we have

$$\begin{aligned}\|\varphi - R\|_{L_p(K_\delta)} &\leqslant c\|\varphi\|_{L_\infty(\triangle)} (|\triangle \setminus \triangle_\delta| + |\triangle'_\delta \setminus \triangle'|)^{1/p} \\ &\leqslant c\delta^{1/p} \|\varphi\|_p \leqslant c\lambda \|\varphi\|_{L_p},\end{aligned}$$

where we used Proposition 2.6. This estimate and (3.20) imply

$$\|\varphi - R\|_{L_p(K)} \leqslant c\lambda \|\varphi\|_{L_p}. \tag{3.21}$$

It remains to prove estimate (3.13). Let first $x \in \triangle'$. Then

$$\begin{aligned}|R(x)| &\leqslant |1 - \kappa'_\triangle(x)||P_K(x)| \leqslant 3\lambda \|P_K\|_{L_\infty(\triangle')} \\ &\leqslant 3\lambda \|P_K\|_{L_\infty(\triangle)} \leqslant c\lambda \|P_K\|_{L_p(K)} = c\lambda \|\varphi\|_p,\end{aligned} \tag{3.22}$$

where we used again Proposition 2.6. For the estimate of $(\mathcal{M}^s_\mathcal{T} \mathbb{1}_K)(x)$ $(x \in \triangle')$ from below, assume that $\triangle \in \mathcal{T}_m$ for some $m \in \mathbb{Z}$. Let $\theta \in \mathcal{T}_m$ be such that $\triangle \subset \theta$. Then by (2.2) it follows that $|\theta| \leqslant c|\triangle|$ and hence, for $x \in \triangle'$,

$$\begin{aligned}(\mathcal{M}^s_\mathcal{T} \mathbb{1}_K)(x) &\geqslant \left(\frac{1}{|\theta|} \int_\theta |\mathbb{1}_K(y)|^s \, dy\right)^{1/s} \geqslant \left(\frac{|K|}{|\theta|}\right)^{1/s} \\ &\geqslant \left(\frac{(1-\rho)|\triangle|}{|\triangle|}\right)^{1/s} \geqslant c > 0,\end{aligned}$$

where we used (2.1). From this and (3.22), we infer

$$|R(x)| \leqslant c\lambda \|\varphi\|_p (\mathcal{M}^s_\mathcal{T} \mathbb{1}_K)(x), \quad x \in \triangle'. \tag{3.23}$$

Let now $x \in \mathbb{R}^2 \setminus \triangle$. Then using (3.17) and Proposition 2.6, we obtain

$$\begin{aligned}|R(x)| &\leqslant \kappa_\triangle(x)|P_K(x)| \leqslant c\lambda \|P_K\|_{L_p(\triangle)} \frac{(1 + |x|)^{k-1}}{(1 + |x|)^{4\mu}} \\ &\leqslant c\lambda \|\varphi\|_p \frac{1}{(1 + |x|)^{4\mu - k}}.\end{aligned} \tag{3.24}$$

Let $B_\triangle$ be the disc inscribed in $\triangle$ (of radius $1/\sqrt{3}$). Then using the definition of μ above, it is readily follows that

$$(\mathcal{M}^{s'}_{\mathcal{E}_\triangle} \mathbb{1}_\triangle)(x) \geqslant (\mathcal{M}^{s'}_{\mathcal{E}_\triangle} \mathbb{1}_{B_\triangle})(x) \geqslant \frac{c}{(1 + |x|)^{2/s'}} \geqslant \frac{c}{(1 + |x|)^{4\mu - k}}. \tag{3.25}$$

On the other hand, by Theorem 2.17, we have

$$(\mathcal{M}^{s'}_{\mathcal{E}_\triangle}\mathbb{1}_\triangle)(x) \leqslant c(\mathcal{M}^s_{\mathcal{T}}\mathbb{1}_\triangle)(x) \leqslant c(\mathcal{M}^s_{\mathcal{T}}\mathbb{1}_K)(x), \quad x \in \mathbb{R}^2, \tag{3.26}$$

where for the latter estimate we used that $|\triangle| \approx |K| \approx 1$. Finally, combining (3.24)–(3.26), we obtain

$$|R(x)| \leqslant c\lambda\|\varphi\|_p(\mathcal{M}^s_{\mathcal{T}}\mathbb{1}_K)(x), \quad x \in \mathbb{R}^2 \setminus \triangle.$$

This estimate coupled with (3.23) yields (3.13).

Finally, by the maximal inequality and (3.13), it follows that

$$\|\varphi - R\|_{L_p(\mathbb{R}^2\setminus K)} \leqslant c\lambda\|\varphi\|_p,$$

which along with (3.21) yields (3.12). The proof is complete. □

Proof of Theorem 3.4. Suppose $S \in \Sigma^k_m(\mathcal{T})$. Then by Lemma 3.7, S can be represented in the form

$$S := \sum_{\triangle\in\tilde{\Lambda}} \mathbb{1}_{K_\triangle} P_{K_\triangle},$$

where $\#\tilde{\Lambda} \leqslant cm$ and $K_\triangle{}^\circ \cap K_{\triangle'}{}^\circ = \emptyset$ if $\triangle \neq \triangle'$.

Let $\varphi_K := \mathbb{1}_K P_K$ with $K := K_\triangle$. We apply the Lemma 3.8 with $\varphi := \varphi_K$, $\lambda := \exp(-(\frac{n}{m})^{1/12})$, and $s := \frac{1}{2}\min\{p, 1\}$ to infer that then there exists a rational function $R_K \in \mathcal{R}_l$ with $l \leqslant c \ \ln^{12}(e + 1/\lambda)$ such that

$$\|\varphi_K - R_K\|_p \leqslant c\lambda\|\varphi_K\|_p$$

and

$$|R_K(x)| \leqslant c\lambda|K|^{-1/p}\|\varphi_K\|_p\Big(\mathcal{M}^s_{\mathcal{T}}\mathbb{1}_K\Big)(x) \quad \text{for } x \in \mathbb{R}^2 \setminus K.$$

We set $R := \sum_{K\in\tilde{\Lambda}} R_K$. Clearly, $R \in \mathcal{R}_N$ with

$$N \leqslant \#\tilde{\Lambda}l \leqslant cml \leqslant cm \ \ln^{12}(e + e^{(\frac{n}{m})^{1/12}}) \leqslant cn.$$

Thus $R \in \mathcal{R}_{cn}$.

Now using Lemma 3.8, we have

$$\begin{aligned}
\|S - R\|_p &= \left\|\sum_K \varphi_K - \sum_K R_K\right\|_p \\
&\leqslant \left\|\sum_K(\varphi_K - R_K)\cdot\mathbb{1}_K + \sum_K R_K\cdot\mathbb{1}_{\mathbb{R}^2\setminus K}\right\|_p \\
&\leqslant c\left\|\sum_K(\varphi_K - R_K)\cdot\mathbb{1}_K\right\|_p + c\left\|\sum_K R_K\cdot\mathbb{1}_{\mathbb{R}^2\setminus K}\right\|_p \\
&\leqslant c\left(\sum_K\|\varphi_K - R_K\right)\|_p^p\Big)^{1/p} + c\lambda\left\|\sum_K\|\varphi_K\|_p|K|^{-\frac{1}{p}}(\mathcal{M}^s_{\mathcal{T}}\mathbb{1}_K)(\cdot)\right\|_p.
\end{aligned}$$

Applying the Fefferman–Stein maximal inequality (Proposition 2.22) with $q := 1$ and $s := \frac{1}{2}\min\{p, 1\} < \min\{p, 1\}$, we obtain

$$\begin{aligned}\|S-R\|_p &\leqslant c\lambda\left(\sum_K \|\varphi_K\|_p^p\right)^{1/p} + c\lambda\left\|\sum_K \|\varphi_K\|_p |K|^{-\frac{1}{p}} \mathbb{1}_K(\cdot)\right\|_p \\ &\leqslant c'\lambda\left(\sum_K \|\varphi_K\|_p^p\right)^{1/p} = c' \exp(-(n/m)^{1/12})\|S\|_p.\end{aligned}$$

The theorem follows. □

Proof of Theorem 3.1. Assume that $p \geqslant 1$. The case $0 < p < 1$ is similar. Choose $S_j \in \Sigma_j^k(\mathcal{T})$ so that $\|f - S_j\|_p \leqslant 2\sigma_j(f, \mathcal{T})_p$, $j = 1, 2, \ldots$ (see (2.25)) and set $\varphi_\nu := S_{2^\nu} - S_{2^{\nu-1}}$, $\nu \geqslant 1$, and $\varphi_0 := S_1$. Evidently, $\varphi_\nu \in \Sigma_{2^{\nu+1}}^k(\mathcal{T})$ and

$$\begin{aligned}\|\varphi_\nu\|_p &= \|S_{2^\nu} - S_{2^{\nu-1}}\|_p \leqslant \|f - S_{2^\nu}\|_p + \|f - S_{2^{\nu-1}}\|_p \\ &\leqslant 2\sigma_{2^\nu}(f, \mathcal{T})_p + 2\sigma_{2^{\nu-1}}(f, \mathcal{T})_p, \quad \nu \geqslant 1, \\ \|\varphi_0\|_p &= \|S_1\|_p \leqslant 2\sigma_1(f, \mathcal{T})_p + \|f\|_p.\end{aligned}$$

Fix $\mu \geqslant 0$. For $\nu = 0, 1, \ldots, \mu$, we apply Theorem 3.4 with $S := \varphi_\nu$, $m := m_\nu := 2^{\nu+1}$, and

$$n := n_\nu := \left\lceil 2^{\nu+1}\Big(\alpha(\mu-\nu)\ln 2\Big)^{12}\right\rceil + 1.$$

As a result, there exist rational functions $R_\nu \in \mathcal{R}_{n_\nu}$ such that for $\nu \geqslant 1$,

$$\begin{aligned}\|\varphi_\nu - R_\nu\|_p &\leqslant c\exp\Big(-(n_\nu/2^{\nu+1})^{1/12}\Big)\|\varphi_\nu\|_p \leqslant c2^{-\alpha(\mu-\nu)}\|\varphi_\nu\|_p \\ &\leqslant c2^{-\alpha(\mu-\nu)}\Big(\sigma_{2^\nu}(f,\mathcal{T})_p + \sigma_{2^{\nu-1}}(f,\mathcal{T})_p\Big)\end{aligned} \tag{3.27}$$

and

$$\|\varphi_0 - R_0\|_p \leqslant c2^{-\alpha\mu}\|\varphi_0\|_p \leqslant c2^{-\alpha\mu}(\sigma_1(f,\mathcal{T})_p + \|f\|_p). \tag{3.28}$$

Now we set $R := \sum_{\nu=0}^{\mu} R_\nu$. Then $R \in \mathcal{R}_N$ with

$$\begin{aligned}N &\leqslant \sum_{\nu=0}^{\mu} n_\nu \leqslant \sum_{\nu=0}^{\mu}\left[2^{\nu+1}\Big(\alpha(\mu-\nu)\ln 2/c^*\Big)^{12} + 1\right] \\ &\leqslant c\sum_{\nu=0}^{\mu} 2^\nu[(\mu-\nu)^{12} + 1] \leqslant c'2^\mu, \quad c' = \text{constant}.\end{aligned}$$

By (3.27) and (3.28), we obtain

$$\begin{aligned}\|f - R\|_p &\leqslant \|f - S_{2^\mu}\|_p + \sum_{\nu=1}^{\mu} \|\varphi_\nu - R_\nu\|_p + \|\varphi_0 - R_0\|_p \\ &\leqslant 2\sigma_{2^\mu}(f, \mathcal{T})_p + \sum_{\nu=1}^{\mu} c2^{-\alpha(\mu-\nu)}\sigma_{2^{\nu-1}}(f, \mathcal{T})_p \\ &\quad + c2^{-\alpha\mu}(\sigma_1(f, \mathcal{T})_p + \|f\|_p) \\ &\leqslant c2^{-\alpha\mu}\left(\sum_{\nu=0}^{\mu} 2^{\alpha\nu}\sigma_{2^\nu}(f, \mathcal{T})_p + \|f\|_p\right).\end{aligned}$$

Therefore, for any $\mu \geqslant 0$, we have

$$R_{N_\mu}(f)_p \leqslant c2^{-\alpha\mu}\left(\sum_{\nu=1}^{\mu} 2^{\alpha\nu}\sigma_{2^\nu}(f, \mathcal{T})_p + \|f\|_p\right) \quad \text{with} \quad N_\mu := c'2^\mu.$$

This estimate readily implies (3.1). □

Proof of Corollary 3.3. This corollary follows readily from Theorem 3.1 together with Proposition 2.10. □

References

[1] C. Bennett, R. Sharpley, Interpolation of operators, in: Pure and Applied Mathematics, vol. 129, Academic Press Inc., Boston, MA, 1988.
[2] O. Davydov, P. Petrushev, Nonlinear approximation from differentiable piecewise polynomials, SIAM J. Math. Anal. 35 (2004) 708–758.
[3] S. Dekel, D. Leviatan, On the relation between piecewise polynomial and rational approximation, Constr. Approx. 20 (2004) 73–91.
[4] G.E. Farin, Curves and Surfaces for Computer Aided Geometric Design: A Practical Guide, Academic Press Inc., Boston, MA, 1990.
[5] C. Fefferman, E. Stein, Some maximal inequalities, Amer. J. Math. 93 (1971) 107–115.
[6] B. Karaivanov, P. Petrushev, Nonlinear piecewise polynomial approximation beyond Besov spaces, Appl. Comput. Harmon. Anal. 15 (2003) 177–223.
[7] D. Newman, Rational approximation to $|x|$, Michigan Math. J. 11 (1964) 11–14.
[8] P. Petrushev, Multivariate n-term rational and piecewise polynomial approximation, J. Approx. Theory 121 (2003) 158–197.
[9] P. Petrushev, V. Popov, Rational Approximation of Real Functions, Cambridge University Press, Cambridge, 1987.
[10] E. Stein, Harmonic Analysis: Real-variable Methods, Orthogonality, and Oscillatory Integrals, Princeton Mathematical Series, No. 43, Princeton University Press, Princeton, NJ, 1993.