JOURNAL OF APPROXIMATION THEORY 36, 119–140 (1982)

Adaptive Approximation by Multivariate Smooth Splines

WOLFGANG A. DAHMEN

Institut für Angewandte Mathematik der Universität Bonn, 5300 Bonn, West Germany

Communicated by Charles A. Micchelli

Received March 18, 1981

In this paper we analyze the potential of adaptive approximation by globally smooth multivariate splines. To this end, certain variants of an adaptive scheme representing "optimal" piecewise polynomial approximation are shown to be, on the one hand, still "equivalent" to the original algorithm while, on the other hand, they turn out to realize adaptive approximation by certain smooth multivariate splines.

1. INTRODUCTION

Several investigations [2, 4, 11] affirm that functions with singularities are efficiently approximated by piecewise polynomials on adaptively refined "grid configurations." Specifically, de Boor and Rice [4] proposed a simple adaptive scheme providing an optimal convergence rate. More precisely, even for multivariate functions with certain "natural" singularities the approximation rate which is achieved by piecewise polynomials of total degree k, say, on N cubes partitioning the respective domain in $\mathbb{R}^s$, is still

$$\mathcal{O}(N^{-(k+1)/s}), \qquad N \to \infty. \tag{1.1}$$

This is known to be optimal for uniform grids and functions in C^{k+1} (the space of functions possessing continuous partial derivatives of order $k+1$) (cf. [8]).

However, all these approaches do not seem to apply directly to smooth multivariate spline approximation and in view of the necessarily "less local" structure of globally smooth splines de Boor and Rice [4] suggested the use of blending methods (cf. [11]) when posing the problem to analyse the potential of smooth adaptive approximation.

Yet, the objective of this paper is to show that in spite of the loss of local structure the potential of spline approximation with even highest possible

0021-9045/82/100119-22$02.00/0

global smoothness is, for any spatial dimension s and arbitrary degree k, still essentially the same as that of the corresponding piecewise polynomial schemes.

To this end we analyse in a rather general setting in Section 2 certain modifications of the simple adaptive algorithm for piecewise polynomial approximation proposed in [4, 12]. While on the one hand these modifications will be shown to be still "equivalent" to the original algorithm we shall point out in Section 3 using the results in [7] that, in particular, these modifications can be realized by adaptive approximation with splines of highest possible global (nontrivial) smoothness.

2. Modifications of an Adaptive Algorithm

The following simple adaptive scheme which was proposed by de Boor and Rice [4, 12] will play an essential role in this section. Given (cf. [4])

(i) a collection $\mathscr{C}$ of "allowable" cells in $\mathbb{R}^s$, $s \geqslant 1$;

(ii) a non-negative function $E: \mathscr{C} \to \mathbb{R}$ with $E(C)$ giving the error (bound) for the approximation on $C \in \mathscr{C}$;

(iii) an initial subdivision of the domain Ω into "allowable" cells;

(iv) a division algorithm for subdividing a cell C into a fixed number of "allowable" cells;

the adaptive algorithm consists in subdividing some cell C with $E(C) > \varepsilon$ in a current partition according to (iv) until $E(C) \leqslant \varepsilon$ for all cells in a current partition. The prescribed "tolerance" ε reflects the desired final accuracy of the approximation.

We will from now on assume that $\mathscr{C}$ denotes the collection of all s-cubes $[a_1, b_1] \times \cdots \times [a_s, b_s]$, $b_i - a_i = h$, $i = 1,..., s$, which is a typical representative of a collection of "allowable" cells. As for the precise general meaning of "allowable" we refer to [4]. For the sake of simplicity Ω will always be the unit s-cube $[0, 1]^s$ and $\{\Omega\}$ the initial partition. Concerning (iv) we will consider only "m-type partitions" of Ω arising from successive "elementary m-type subdivisions" of a cube C into m^s congruent subcubes forming the set $d_m(C)$ of "children" of C. As in [4], C is called the parent of $C' \in d_m(C)$ we write this as $C = \hat{C}$. More generally, C is an "ancestor" of C' if C' is obtained by subsequent subdivisions (including the trivial one) of C. For a given (m-type) partition $\Theta \subset \mathscr{C}$, $\hat{\Theta}$ denotes the set of all ancestors of elements of Θ. In particular we have $\Theta \subset \hat{\Theta}$. For any collection $X \subset \mathscr{C}$ we will always denote by $\tilde{X}$ the union of the cubes in X.

In view of the above conventions the scheme (i)–(iv) merely depending on m and E may be briefly denoted by $A(E, m)$. Clearly, $A(E, m)$ typically represents adaptive approximation by piecewise polynomials on cubes, i.e., for a given function f, $E = E_f$ is the local error (bound)

$$\operatorname{dist}_p(f, \Pi_k)_C \equiv \inf_{g \in \Pi_k} \|f - g\|_p(C) \overset{(\leqslant)}{=} E(C)$$

where $\|\cdot\|_p(D)$ denotes the usual L_p-norm with respect to some domain $D \subset \mathbb{R}^s$ and

$$\Pi_k = \left\{ \sum_{|\alpha| \leqslant k} c_\alpha x^\alpha : c_\alpha \in \mathbb{R} \right\}$$

is the space of polynomials of (total) degree less than or equal to k. Here we have used for $x = (x_1, ..., x_s) \in \mathbb{R}^s$, $\alpha = (\alpha_1, ..., \alpha_s) \in \mathbb{Z}_+^s$ the standard multiindex notation, i.e., $|\alpha| = \alpha_1 + \cdots + \alpha_s$, $x^\alpha = x_1^{\alpha_1} \cdot \cdots \cdot x_s^{\alpha_s}$.

Let $\mathscr{C}_0$ be the collection of all closed convex sets C in $\mathbb{R}^s$ with

$$\operatorname{diam}(C)/\delta(C) \leqslant 2,$$

where $\delta(C)$ is the diameter of the largest cube contained in $C \in \mathscr{C}_0$. Note that the following assumptions on E are typically satisfied for polynomial approximation (cf. [2, 4]):

There exists a constant $b > 0$ and a function $F: \mathscr{C}_0 \to \mathbb{R}_+$ such that for all $C \in \mathscr{C}_0$

(i) $E(C) = (\operatorname{diam}(C))^b F(C)$;

(ii) $F(C_1) \leqslant F(C_2)$ if $C_1 \subseteq C_2$; (2.1)

(iii) for any fixed $m \in \mathbb{N}$ there is a constant $d > 0$ such that $C = \bigcup \{C_i : i = 1, ..., m\} \in \mathscr{C}_0$, with $C_i \in \mathscr{C}_0$, $i = 1, ..., m$, implies

$$dF(C) \leqslant \sum_{i=1}^{m} F(C_i).$$

In particular, (ii) ensures that E is monotone, i.e., for $C_1 \subseteq C_2$

$$E(C_1) \leqslant E(C_2). \tag{2.2}$$

The analysis of $A(E, m)$ relies of course crucially on the completely local nature of such schemes, namely, on the fact that the approximations on distinct cubes do not affect each other. This is no longer true when dealing with globally smooth spline approximation. However the following result from [7] (which will be presented in more detail in Section 3) provides a useful link between the "non-smooth" and the "smooth" case. To a given partition $\Theta \subset \mathscr{C}$ of Ω we may assign a space $\mathscr{S}_k(\Theta)$ of splines of degree k with the following properties

— $\dim \mathscr{S}_k(\Theta) = \mathscr{O}(|\Theta|)$, $|\Theta| \to \infty$ ($|\Theta|$ denoting the cardinality of Θ);

— $\mathscr{S}_k(\Theta) \subset C^{k-1}(\Omega)$;

— for $f \in L_p(\Omega)$ there is a spline $S(f) \in \mathscr{S}_k(\Theta)$ and a constant γ independent of Θ such that for any $C \in \Theta$

$$\|f - S(f)\|_p(C) \leqslant \gamma \operatorname{dist}_p(f, \Pi_k)_{\tilde{X}(C \mid \Theta)}, \tag{2.3}$$

where $X(C \mid \Theta)$ is a collection of cubes depending on C and Θ which reflects, roughly speaking, to what an extent the local nature of the spline approximation is spoiled by the global smoothness conditions.

As in the univariate case (cf. e.g., [3]) (2.3) says that the local error for smooth approximations can be estimated by an error function for polynomial approximation with respect to a "slightly" larger domain $\tilde{X}(C \mid \Theta)$. So, using the above terms, the task of analyzing smooth adaptive schemes and, in particular, of comparing them to piecewise polynomial approximation means in view of (2.3) to analyze the effect of changing E to $\gamma \cdot E \circ \tilde{X}$ in $A(E, m)$, i.e., to compare formally the schemes

$$A(E, m) \qquad \text{and} \qquad A(\gamma E \circ \tilde{X}, m).$$

As a suitable criterion for such a comparison we introduce the following notion of "equivalence."

Definition 2.1. An algorithm A_1 is called equivalent to an algorithm A_2, in symbols $A_1 \sim A_2$, if there are constants $0 < c_1 \leqslant c_2 < \infty$ such that for all $\varepsilon > 0$

$$c_1 |\Phi_\varepsilon(A_1)| \leqslant |\Phi_\varepsilon(A_2)| \leqslant c_2 |\Phi_\varepsilon(A_1)|,$$

where, for any algorithm $A = A(E, m)$, $\Phi_\varepsilon(A)$ denotes the "final" partition produced by A; therefore $E(C) \leqslant \varepsilon$ for all $C \in \Phi_\varepsilon(A)$.

We state the following simple observation.

Lemma 2.1. *Suppose that $E(C) \leqslant \varepsilon$ for all $C \in \Theta$, where Θ is some m-type partition of Ω. Then*

$$|\Phi_\varepsilon(A(E, m))| \leqslant |\Theta|.$$

Heading for the abovementioned comparison we start now to modify $A(E, m)$ step by step.

LEMMA 2.2. *For any two fixed integers $m_1, m_2 \geqslant 2$ and any E satisfying* (2.2) *one has*

$$A(E, m_1) \sim A(E, m_2).$$

Proof. It is certainly sufficient to prove the assertion for $m_2 = l \cdot m_1$, for some $l \in \mathbb{N}$, since this would imply that $A(E, m_i) \sim A(E, m_1 \cdot m_2)$, $i = 1, 2$. But setting $A_1 = A(E, m_1)$, $A_2 = A(E, m_2)$, for $m_2 = lm_1$, the estimate

$$|\Phi_\varepsilon(A_1)| \leqslant |\Phi_\varepsilon(A_2)|$$

is then obvious in view of (2.2) and since any m_2-type partition is also of m_1-type. This fact, (2.2) and the definition of $\Phi_\varepsilon(A)$ also lead to the converse estimate

$$|\Phi_\varepsilon(A_2)| \leqslant l^s \, |\Phi_\varepsilon(A_1)|$$

which completes the proof.

LEMMA 2.3. *For any error function E satisfying* (2.1) *and any fixed constant $\gamma > 0$ one has*

$$A(E, m) \sim A(\gamma E, m).$$

Proof. Without loss of generality we may assume $\gamma > 1$. Setting $A_1 = A(E, m)$, $A_2 = A(\gamma E, m)$ we trivially have for all $\varepsilon > 0$

$$|\Phi_\varepsilon(A_1)| \leqslant |\Phi_\varepsilon(A_2)|.$$

On the other hand, we may assume in view of Lemma 2.2 that $(\gamma/m^b) < 1$ where b is the constant occurring in (2.1)(i). Now let Θ be the partition which is obtained by subdividing each $C \in \Phi_\varepsilon(A_1)$ once again. Then $|\Theta| = m^s \, |\Phi_\varepsilon(A_1)|$. Hence if $C \in \Theta$ and $\hat{C} \in \Phi_\varepsilon(A_1)$ is its parent, (2.1)(i,ii) yield

$$\begin{aligned} \gamma E(C) &= \gamma(\operatorname{diam}(C))^b \, F(C) \leqslant (\gamma/m^b)(\operatorname{diam}(\hat{C}))^b \, F(\hat{C}) \\ &= (\gamma/m^b) \, E(\hat{C}) < \varepsilon. \end{aligned}$$

Thus, Lemma 2.1 provides

$$|\Phi_\varepsilon(A_2)| \leqslant m^s \, |\Phi_\varepsilon(A_1)|$$

which finishes the proof.

The discussion of further modifications of $A(E, m)$ as indicated by (2.1) requires some more notation and preliminary remarks.

Two cubes $C_1, C_2 \in \Theta$ are called neighbors (in Θ) iff $C_1 \cap C_2$ is not empty.

$U(C \mid \Theta)$ denotes the set of all neighbors of C in Θ and

$$\tilde{U}(C \mid \Theta) = \bigcup \{C' : C' \in U(C \mid \Theta)\}.$$

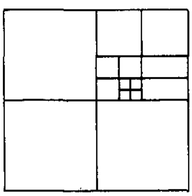
As before, for any subset X of a partition Θ the union over its elements will be denoted by $\tilde{X}$.

A cube $C \in \Theta$ is called a "large" cube in Θ iff all its neighbors in Θ belong to the same or a later generation, i.e., their size is equal to or smaller than the size of C. Accordingly, $C \in \Theta$ is called "small" in Θ if C is not large in the above sense.

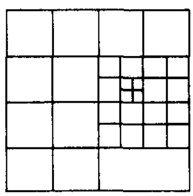
A partition is called graded (cf. [4]) iff the difference in generations between any two neighboring cubes is at most one. We shall have to deal with the following slightly stronger restriction.

DEFINITION 2.2. An m-type partition $\Theta \subset \mathscr{C}$ is said to be "properly nested" if Θ is graded and contains with any small cube C also all its "brothers" in $d_m(\hat{C})$.

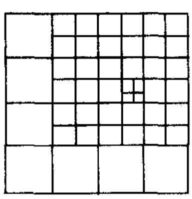
Definition 2.2 is illustrated by the following examples for $s = m = 2$.



not graded



graded but not properly nested



properly nested

It turns out that any m-type partition is "almost" properly nested in the following sense.

LEMMA 2.4. *Any m-type partition $\Theta \subseteq \mathscr{C}$ can be extended to a properly nested partition $\mathring{\Theta}$ so that*

$$|\mathring{\Theta}| \leqslant \delta |\Theta|$$

where the constant δ depends only on m and s but not on Θ.

Proof. Let Θ be an arbitrary m-type partition of Ω and let $\{\Theta_i\}_{i=0}^N$ be a "generating" sequence for Θ, i.e., $\Theta_0 = \{\Omega\}$, $\Theta_N = \Theta$, and Θ_{i+1} arises from Θ_i by one elementary m-type subdivision of some $C_i \in \Theta_i$. Hence one has for $i = 0,..., N-1$

$$|\Theta_{i+1}| - |\Theta_i| = m^s - 1. \tag{2.4}$$

We construct now a sequence of graded partitions $\{\Theta_i'\}_{i=0}^N$ in the following way. Setting $\Theta_0' = \Theta_0 = \{\Omega\}$, Θ_{i+1}' is obtained from Θ_i' as follows: if $C_i \in \Theta_i \cap \Theta_i'$ subdivide C_i and all its neighbors in Θ_i' which have exactly the same size as C_i. If $C_i \notin \Theta_i'$, i.e., C_i was already subdivided before, subdivide only those neighbors of C_i in Θ_i' which have the same size as C_i. Hence we have

$$|\Theta_{i+1}'| - |\Theta_i'| \leqslant (m^s - 1)\, 3^s. \tag{2.5}$$

One may verify by induction that for every $i = 0,..., N$ every neighbor C' of any cube C in $\Theta_i' \cap \Theta_i$ has exactly the same size as C. Moreover, Θ_i' contains together with such a C' all cubes in $d_m(\hat{C}')$. Hence $C \in \Theta_i' \cap \Theta_i$ as well as all its neighbors in Θ_i' are large cubes in Θ_i'. Thus the subdivisions leading to Θ_{i+1}' may cause at most a difference in generations of one. Hence all the Θ_i' and in particular Θ_N' are graded. Furthermore, it follows from (2.4) and (2.5) that

$$|\Theta_N'| \leqslant 3^s\, |\Theta_N| = 3^s\, |\Theta|. \tag{2.6}$$

Subdividing every cube in Θ_N' once again provides a partition Θ'' which is properly nested and satisfies

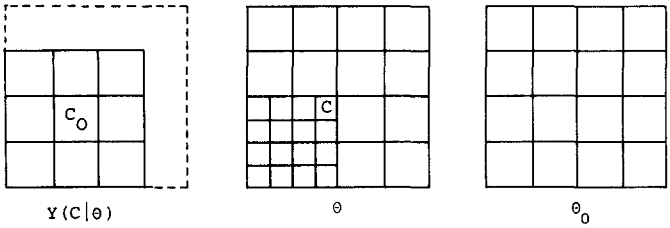
$$|\Theta''| \leqslant m^s\, |\Theta_N'| \leqslant (3m)^s\, |\Theta|.$$

This completes the proof.

We now introduce a typical candidate for a map $X: C \mapsto X(C \mid \Theta)$ for $C \in \Theta \subset \mathscr{C}$ which may appear in (2.3). For any cube $C \in \hat{\Theta}$ let $C_0 \in \hat{\Theta}$ be the smallest ancestor of C which is a large cube in some partition $\Theta_0 \subset \hat{\Theta}$. We also assume that Θ_0 is minimal in the sense that the elements of $U(C_0 \mid \Theta_0)$ belong all to the same generation as C_0, i.e., all the neighbors of C_0 in Θ_0 have the same size. Then we set

$$Y(C \mid \Theta) := U(C_0 \mid \Theta_0). \tag{2.7}$$

The following drawings visualize a typical situation in the case $s = m = 2$.



Note that $\tilde{Y}(C \mid \Theta) = \tilde{U}(C \mid \Theta)$ if C is a large cube in Θ. Moreover, if Θ is properly nested, then C_0 is either C or $\hat{C}$. Let us list some further properties of Y which follow immediately from its definition. For any partition $\Theta \subset \mathscr{C}$ one has

(i) $\tilde{Y}(C_1 \mid \Theta) \subseteq \tilde{Y}(C_2 \mid \Theta)$ if $C_1 \subseteq C_2$.

(ii) $|Y(C \mid \Theta)| \leqslant 3^s$ and, if $|Y(C \mid \Theta)| = 3^s$, the "central" cube contains C.

(iii) For any $C \in \Theta$ there is at least one cube in $Y(C \mid \Theta) \cap \Theta$ which is not an ancestor of C.

(iv) For every small cube $C \in \Theta$ one has $Y(C \mid \Theta) = Y(\hat{C} \mid \Theta)$. (2.8)

The map X occurring in (2.3) will be actually slightly different from Y although it will still share the properties (2.8). So we will call a map $X\colon \hat{\Theta} \to 2^{\mathscr{C}}$ a "cover" if X satisfies (2.8)(i–iv) as well as the following two relations: For any partition Θ one has

$$\tilde{Y}(C \mid \Theta) \subseteq \tilde{X}(C \mid \Theta) \qquad \text{for} \quad C \in \Theta. \tag{2.9}$$

If in addition Θ is properly nested, one has also

$$\tilde{X}(C \mid \Theta) \subseteq \tilde{Y}(\hat{C} \mid \Theta), \qquad C \in \Theta. \tag{2.10}$$

We are now ready to discuss the effect of replacing the error function E in $A(E, m)$ by $E \circ \tilde{X}$ where X is a cover. However, note that, according to the properties of a cover (cf. (2.8)(iv)), one may have $X(C \mid \Theta) = X(C' \mid \Theta)$ for some child $C' \in d_m(C) \subset \Theta$ and consequently $E(\tilde{X}(C \mid \Theta)) = E(\tilde{X}(C' \mid \Theta))$. So subdividing only the cube C if $E(\tilde{X}(C \mid \Theta)) > \varepsilon$ may not decrease the error and the adaptive algorithm in the above simple form may therefore produce redundant subdivisions, as a little thought will confirm. In order to avoid such redundant subdivisions one will have to subdivide sometimes only certain neighbors of C instead of C itself. This leads to the following extended version of the original simple adaptive scheme.

For any cover X and any error E satisfying (2.1), $A(E \circ \tilde{X}, m)$ will always denote an adaptive procedure according to the following steps. Let ε be the tolerance.

(a) Set $\Theta_0 = \{\Omega\}$.

(b) Compute the approximation with respect to Θ_j and set $\varepsilon_j = \max\{E(\tilde{X}(C \mid \Theta_j)): C \in \Theta_j\}$.

(c) If $\varepsilon_j \leqslant \varepsilon$, stop.

(d) Set $\varepsilon^*_{j+1} = \max\{\varepsilon, m^{-b}\varepsilon_j\}$.

(e) Let $\Theta^*_j = \{C' \in X(C \mid \Theta_j) \cap \Theta_j: C \in \Theta_j$ and $E(\tilde{X}(C \mid \Theta_j)) > \varepsilon^*_{j+1}\}$.

(f) Extend Θ_j to Θ_{j+1} by an elementary m-type subdivision of each cube in Θ^*_j.

(g) Set $j \leftarrow j+1$ and go to (b). (2.11)

As mentioned before, note that when $E(\tilde{X}(C \mid \Theta)) > \varepsilon$, then sometimes, namely when $C \notin X(C \mid \Theta)$, only certain neighbors of C instead of C itself will be subdivided (cf. (2.11)(e,f)). Some properties of this algorithm are stated in the following:

LEMMA 2.5. *Using the above notation we have*

(i) $\max\{E(\tilde{X}(C \mid \Theta_{j+1})): C \in \Theta_{j+1}\} = \varepsilon_{j+1} \leqslant \varepsilon^*_{j+1}$. *Hence there is an* $N \in \mathbb{N}$ *such that* $\varepsilon^*_N \leqslant \varepsilon$.

(ii) Θ^*_j *contains the minimal number of cubes which have to be subdivided in order to achieve* (i), *i.e., it makes sense to write*

$$\Theta_j = \Phi_{\varepsilon^*_j}(A(E \circ \tilde{X}, m)), \qquad j = 0,..., N,$$

where Lemma 2.1 *still holds for* $\Phi_\varepsilon(A(E \circ \tilde{X}, m))$.

(iii) *Each partition* Θ_j, $j = 0,..., N$, *is properly nested.*

Proof. (i) From (2.8)(i,iii) we conclude that when subdividing all the cubes in $\Theta \cap X(C \mid \Theta)$ and calling the resulting extended partition Θ' we get

$$\operatorname{diam}(\tilde{X}(C' \mid \Theta')) \leqslant \frac{1}{m} \operatorname{diam}(\tilde{X}(C' \mid \Theta)),$$

where $C' = C$ if $C \notin X(C \mid \Theta)$ and $\hat{C}' = C$, otherwise. Combining this with (2.1)(i) provides (i).

As for (ii), let us denote in the following by $\dot{C}$ the "central" cube in $X(C \mid \Theta)$ (cf. (2.8)(ii)). Suppose now that some $C^* \in \Theta^*_j$ is not subdivided. We will show that then $\varepsilon_{j+1} > \varepsilon^*_{j+1}$. By definition of Θ^*_j there must be a $C \in \Theta_j$ so that

$$C^* \in \Theta_j \cap X(C \mid \Theta_j) \qquad \text{and} \qquad E(\tilde{X}(C \mid \Theta_j)) > \varepsilon^*_{j+1}. \qquad (2.12)$$

Let us first consider the case $\mathring{C} = C$, i.e., $C \in \Theta^*_{j+1}$. Note that $\mathring{C} \neq C^*$ since otherwise (2.7), (2.9) and (2.8)(ii,iii) would imply that $\tilde{X}(C \mid \Theta_{j+1}) \supseteq \tilde{X}(C \mid \Theta_j)$ which because of (2.12) and the monotonicity (2.2) of E yields the contradiction $\varepsilon_{j+1} > \varepsilon^*_{j+1}$. So, assume that $C = \mathring{C} \neq C^*$ and let $C' \in d_m(C)$ be a neighbor of C^*. Again (2.7) and (2.9) ensure that $\tilde{X}(C' \mid \Theta_{j+1}) \supseteq C^*$. This implies, because of (2.8)(ii) and the fact that $C^* \in \Theta_{j+1} \cap \Theta_j$, the inclusion

$$\tilde{X}(C' \mid \Theta_{j+1}) \supseteq \tilde{X}(C \mid \Theta_j).$$

This yields by (2.2) again that $\varepsilon_{j+1} \geqslant E(\tilde{X}(C' \mid \Theta_{j+1})) \geqslant E(\tilde{X}(C \mid \Theta_j)) > \varepsilon^*_{j+1}$, contradicting (i). The case $C \neq \mathring{C}$, i.e., $C \notin X(C \mid \Theta_j)$, follows now in a similar way. In fact, $\mathring{C}$ must contain a cube C'' which is not smaller than C and is a neighbor of C^*. From (2.8)(i,iv) we conclude that $\tilde{X}(C \mid \Theta_j) \subseteq \tilde{X}(C'' \mid \Theta_j)$. The same reasoning which was applied before to C' works now for C''.

So far we have seen that subdividing at times only certain neighbors of a cube, for which the error was found to be still too large, was just the right strategy to avoid redundant subdivisions. On the other hand one may intuitively expect that this strategy of subdividing, e.g., the large neighbors of a small cube, instead of the small cube itself, automatically keeps the partitions very gradual which is essentially the claim of (iii). In view of (ii) it suffices now to show that $\Theta_\varepsilon := \Phi_\varepsilon(A(E \circ \tilde{X}, m))$ is properly nested for any $\varepsilon > 0$. Suppose C is a small cube in Θ_ε and $C^* \in U(C \mid \Theta_\varepsilon)$ has maximal size in $U(C \mid \Theta_\varepsilon)$. By virtue of (2.7), (2.9) and (2.8)(ii), $\tilde{X}(C \mid \Theta_\varepsilon)$ contains 3^s cubes which have at least the same size as C^*. Let $\bar{C}$ be the ancestor of C which has the same size as C^*. Furthermore, let $C' \in d_m(\bar{C})$ and, in particular, $C \subseteq C'_0 \in d_m(\bar{C})$. Then (2.8)(i,iv) ensures that

$$\tilde{X}(C' \mid \Theta_\varepsilon) \subseteq \tilde{X}(C'_0 \mid \Theta_\varepsilon) = \tilde{X}(C \mid \Theta_\varepsilon).$$

On the other hand, if some $C' \in d_m(\bar{C})$ belongs to $X(C'' \mid \Theta_\varepsilon)$ for some $C'' \in \Theta_\varepsilon$ we still have by (2.8)(ii) that $\tilde{X}(C'' \mid \Theta_\varepsilon) \subseteq \tilde{X}(C \mid \Theta_\varepsilon)$. So, since $E(\tilde{X}(C \mid \Theta_\varepsilon)) \leqslant \varepsilon$ holds by assumption, any further subdivision of the cubes in $d_m(\bar{C})$ would be redundant. Hence, by (ii), $C'_0 = C \in d_m(\bar{C})$, i.e., C differs from C^* at most in one generation and all cubes in $d_m(\hat{C}) = d_m(\bar{C})$ belong to Θ_ε. This completes the proof.

Lemma 2.5 justifies calling an algorithm A properly nested if $\Phi_\varepsilon(A)$ is properly nested for all $\varepsilon > 0$.

THEOREM 2.1. *Let E satisfy* (2.1) *and let m, $l \geqslant 2$ be any two fixed integers. Then*

$$A(E, l) \sim A(\gamma E \circ \tilde{X}, m)$$

for any cover X and any fixed constant $\gamma > 0$. Furthermore, $A(\gamma E \circ \tilde{X}, m)$ is properly nested.

Proof. In view of Lemmas 2.2 and 2.3 we may assume without loss of generality that $m = l$, $\gamma = 1$. Setting $A_1 = A(E, m)$ and $A_2 = A(E \circ \tilde{X}, m)$, the estimate $|\Phi_\varepsilon(A_1)| \leqslant |\Phi_\varepsilon(A_2)|$ is trivial, since, by (2.2), $E(C) \leqslant E(\tilde{X}(C \mid \Theta))$ for all $C \in \Theta$ and any partition Θ.

As for the converse estimate let Θ be obtained by extending $\Phi_\varepsilon(A_1)$ to a properly nested partition (cf. Lemma 2.4). Subdividing each cube in Θ into $(m^s)^3$ congruent subcubes provides a further extension Θ_ε of $\Phi_\varepsilon(A_1)$ which is still properly nested and whose cardinality may be estimated according to Lemma 2.4 as

$$|\Theta_\varepsilon| \leqslant \delta\, |\Phi_\varepsilon(A_1)|. \tag{2.13}$$

Here the constant δ depends only on m and s but not on $\Phi_\varepsilon(A_1)$. Conditions (2.7) and (2.10) ensure now that for all $C \in \Theta_\varepsilon$ the cubes in $X(C \mid \Theta_\varepsilon)$ are children or even grandchildren of cubes in Θ. (2.1), (2.8)(ii) and (2.10) yield now the following estimates:

$$\begin{aligned}
E(\tilde{X}(C \mid \Theta_\varepsilon)) &\leqslant E(\tilde{Y}(\hat{C} \mid \Theta_\varepsilon)) \leqslant \frac{1}{d} \sum_{C' \in Y(\hat{C} \mid \Theta_\varepsilon)} E(C') \\
&\leqslant (dm^b)^{-1} \sum_{C' \in Y(\hat{C} \mid \Theta_\varepsilon)} (\operatorname{diam}(\hat{C}'))^b\, F(\hat{C}') \\
&= (dm^b)^{-1} \sum_{C' \in Y(\hat{C} \mid \Theta_\varepsilon)} E(\hat{C}') \leqslant \frac{3^s}{dm^b}\, \varepsilon.
\end{aligned}$$

Again we may assume by Lemma 2.2 that m is sufficiently large, i.e., $(3^s/dm^b) < 1$ which implies that

$$E(\tilde{X}(C \mid \Theta_\varepsilon)) \leqslant \varepsilon \qquad \text{for all} \quad C \in \Theta_\varepsilon.$$

But Lemma 2.5(ii) says then that

$$|\Phi_\varepsilon(A_2)| \leqslant |\Theta_\varepsilon|.$$

The first part of the assertion follows now from (2.13). The rest is an immediate consequence of Lemma 2.5(iii).

3. An Application to Certain Multivariate Splines of Highest Possible Global Smoothness

In this section we concentrate on one possible realization of Theorem 2.1 which is based on multivariate splines of arbitrary degree k which even belong to $C^{k-1}(\Omega)$. The construction of the spaces $\mathscr{S}_k(\Theta)$ satisfying (2.3) with an appropriate cover X is essentially based on the results in [7].

Reviewing briefly the main facts from [7] requires the following notation. The convex hull of a given set V is denoted by $[V]$, whereas when $V \subseteq \mathbb{R}^n$, $n \geqslant s$, $(V)_s$ means the orthogonal projection of V to $\mathbb{R}^s$. The three integers n, k, s will be consistently interrelated by $k = n - s \geqslant 0$ where s and k will always refer to the spatial dimension and to the (total) degree of splines or polynomials, respectively.

A crucial role will be played by the multivariate B-spline which may be defined as follows (cf. [3, 5, 10]). For any (non-degenerate) simplex $\sigma = [\{v^0,..., v^n\}] \subset \mathbb{R}^n$ let

$$M_\sigma(x) = \mathrm{vol}_k(\{u \in \sigma\colon (u)_s = x\})$$

and

$$P(\sigma) = \{(v^0)_s,..., (v^n)_s\}$$

be the set of "knots" associated with σ. One can show [5, 10] that the B-spline

$$M_\sigma(x)/\mathrm{vol}_n(\sigma) = M(x \mid P(\sigma))$$

only depends on the knots in $P(\sigma)$. Clearly $M(x \mid P)$ is non-negative and supported on $[P]$. As for the various properties and representations of $M(x \mid P)$, in particular practical recurrence relations, we refer to [5, 6, 10, 11]. Here we emphasize only that $M(x \mid P)$ is a piecewise polynomial of degree $k = n - s$ which even belongs to $C^{k-1}(\mathbb{R}^s)$ whenever the knots $x^i \in P$ are in "general position," i.e., every $s + 1$ points in P are affinely independent (cf. [5, 10]).

The construction of certain spline spaces based on the B-splines involves the so called "Kuhn's triangulation" $\mathscr{K}_n(\varphi)$ which decomposes any parallelepiped $\varphi \subset \mathbb{R}^n$ into $n!$ simplices of equal voume (cf. [9]). This makes it possible to associate with any set $\mathscr{V}$ of parallelepipeds in $\mathbb{R}^n$ the space of splines of degree k

$$\mathscr{S}_k(\mathscr{V}, D) = \mathrm{span}\{M(x \mid P(\sigma))\colon \sigma \in \mathscr{K}_n(\varphi), \varphi \in \mathscr{V}, (\varphi)_s \cap D \neq \varnothing\}, \quad (3.1)$$

where D is a given domain in $\mathbb{R}^s$.

In particular, we shall deal with special collections $\mathscr{V}$ of parallelepipeds whose construction involves an $(n \times n)$ matrix of the following type:

$$H = \begin{pmatrix} (\delta_{ij})_{i,j=1}^{s} & (c_{il})_{i=1,l=1}^{s,k} \\ 0 & (\delta_{ij})_{i,j=1}^{k} \end{pmatrix}, \tag{3.2}$$

where

$$\max_{i=1,\ldots,s} \sum_{j=k}^{k} |c_{ij}| < \frac{1}{2}. \tag{3.3}$$

Whenever we write $u + x$ for $u \in \mathbb{R}^n$, $x \in \mathbb{R}^s$, x will be understood to be extended to an n-vector by setting $x_{s+1} = \cdots = x_n = 0$. Parallelepipeds of the form $\mathscr{q} = aH(Q) + u$, where $a \in \mathbb{R}_+$, $u \in \mathbb{R}^n$, $Q = [0, 1]^n$ will be briefly called "H-cubes."

Setting

$$\mathscr{H} = \{H(Q) + v : v \in \mathbb{Z}^s, (H(Q) + v)_s \cap \Omega \neq \varnothing\} \tag{3.4}$$

we note that one can choose H such that

$$\mathscr{S}_k(\mathscr{H}, \Omega) \subset C^{k-1}(\Omega). \tag{3.5}$$

We shall need the following consequences of (3.3). There is a cube $C' \subset \mathbb{R}^s$ such that for $\Omega = [0, 1]^s$

$$\begin{gathered} \Omega \subset (H(Q))_s \subseteq C', \\ \operatorname{diam}(C') < 2 \operatorname{diam}(\Omega). \end{gathered} \tag{3.6}$$

Moreover, setting $\mu(\mathscr{q}) := \max\{\operatorname{vol}_k(\{u \in \mathscr{q} : (u)_s = x\}) : x \in (\mathscr{q})_s\}$ and

$$I(\mathscr{q}) := \{x \in (\mathscr{q})_s : \operatorname{vol}_k(\{u \in \mathscr{q} : (u)_s = x\}) = \mu(\mathscr{q})\}, \tag{3.7}$$

condition (3.3) ensures that

$$\operatorname{vol}_s(I(\mathscr{q})) > 0. \tag{3.8}$$

Note that $I(\mathscr{q})$ is exactly the domain for which the B-splines $M(x \mid P(\sigma))$, $\sigma \in \mathscr{K}_n(\mathscr{q})$, form a partition of unity.

It was pointed out in [6, 7] how the spaces $\mathscr{S}_k(\mathscr{H}, D)$ may be "locally refined" without interfering with the global smoothness properties, namely, by simply refining the H-cubes in $\mathscr{H}$. In fact, assigning to every $\mathscr{q} \in \mathscr{H}$ a partition $\varphi(\mathscr{q})$ of a fixed m-type we obtain a "refined" partition

$$\mathscr{R} = \mathscr{R}(H) = \bigcup \{\varphi(\mathscr{q}) : \mathscr{q} \in \mathscr{H}\} \tag{3.9}$$

for $\Omega \times [0, 1]^k$. $\mathscr{R}(H)$ is called an "m-type refinement" of $\mathscr{H}$. We shall use in the following the terminology of Section 2 also for partitions consisting of H-cubes.

It was shown in [7] that when $\mathscr{S}_k(\mathscr{H}, \Omega) \subset C^{k-1}(\Omega)$, then

$$\mathscr{S}_k(\mathscr{R}, \Omega) \subset C^{k-1}(\Omega) \tag{3.10}$$

holds for any refinement $\mathscr{R} = \mathscr{R}(H)$ of $\mathscr{H}$.

Since we may state in view of the definition of $\mathscr{H}_n(q)$, (3.8) and (3.9) that one has in general

$$|\mathscr{R}| \leqslant \dim \mathscr{S}_k(\mathscr{R}, \Omega) \leqslant n! \, |\mathscr{R}|, \tag{3.11}$$

it is very desirable to refine as few H-cubes as are necessary in order to improve the local approximation behavior. This suggests consideration of the following restricted type of refinements. To this end we call any H-cube in $\mathscr{R}$ sharing an s-face, viz., the "bottom face," with $\mathscr{R}^s$ a "bottom cube." A refinement $\mathscr{R}$ is then called "economic" if $\hat{\mathscr{R}} \backslash \mathscr{R}$ contains only bottom cubes, i.e., if $\mathscr{R}$ is obtained by subsequently subdividing only bottom cubes.

In order to state next local error bounds for the approximation by elements of $\mathscr{S}_k(\mathscr{R}, \Omega)$ we have to introduce the notion of a "protected" H-cube which is slightly stronger than that of a large cube (cf. Section 2). For any bottom cube q in $\hat{\mathscr{R}} = \hat{\mathscr{R}}(H)$, let $\mathscr{L}_q$ denote the collection of all bottom cubes in $\hat{\mathscr{R}}$ belonging to the same generation as q. The bottom cube q is called "protected" in $\mathscr{R}$ iff (see (3.7))

$$(q)_s \subseteq I\left(\bigcup \{\hat{q}' : q' \in \mathscr{L}_q\}\right). \tag{3.12}$$

In analogy to (2.7), we denote, for any bottom cube q in $\hat{\mathscr{R}}$, by q_0 the minimal ancestor of q in $\hat{\mathscr{R}}$ which is protected with respect to some "minimal" refinement $\mathscr{R}_0 \subset \hat{\mathscr{R}}$. As in (2.7) we define

$$G(q \mid \mathscr{R}) = U(q_0 \mid \mathscr{L}_{q_0}). \tag{3.13}$$

Let $\mathscr{V}_i = \mathscr{V}_i(\mathscr{R})$, $i = 0,\ldots, l = l(\mathscr{R})$ be the collection of H-cubes in $\hat{\mathscr{R}}$ which differ from those in $\mathscr{H}$ by exactly i generations and let $S(f, \mathscr{V}_i)$ be a best approximation (with respect to the L_p-norm) to $f \in L_p$ in $\mathscr{S}_k(\mathscr{V}_i, \Omega)$. We consider the following approximation scheme which is obtained by updating "coarser grid approximations" locally on "finer grids."

$$\begin{gathered} S_0(f) = S(f, \mathscr{H}), \\ S_{j+1}(f) = S(f - S_j(f), \mathscr{V}_{j+1}) + S_j(f), \qquad j = 0,\ldots, l-1, \\ T(f, \mathscr{R}) = S_l(f). \end{gathered} \tag{3.14}$$

The following result from [7] provides local error estimates of the desired type even for economic refinements $\mathscr{R}(H)$.

LEMMA 3.1. *For any H given by* (3.2), (3.3) (*with rational entries c_{ij}*) *there is an integer $m = m(H)$ such that the following estimate holds for any $f \in L_p(\Omega)$ and each bottom cube φ in any economic and properly nested* (*cf. Definition* 2.2) *m-type refinement $\mathscr{R} = \mathscr{R}(H)$* (*cf.* (3.2)):

$$\|f - T(f, \mathscr{R})\|_p((\varphi)_s \cap \Omega) \leqslant \gamma \operatorname{dist}_p(f, \Pi_k)_{(\tilde{G}(\varphi \mid \mathscr{R}))_s \cap \Omega}.$$

Here the constant γ does not depend on f, φ and $\mathscr{R}$.

We have to reinterpret the above result in terms of Section 2, in particular with regard to (2.3), i.e., we have to relate first the spaces $\mathscr{S}_k(\mathscr{R}, \Omega)$ and hence the $(s+k)$-dimensional refinements $\mathscr{R}$ to an s-dimensional partition $\Theta \subset \mathscr{C}$ of Ω.

To this end, let, for any cube $C \in \mathscr{C}$, $Q(C) \subset \mathbb{R}^n$ be the n-dimensional cube having C as an s-face and let

$$\varphi(C, H) = H(Q(C))$$

be the corresponding H-cube. Conversely, we denote, for any bottom cube φ in some refinement $\mathscr{R}$, its bottom face by

$$C(\varphi) := \varphi \cap \mathbb{R}^s.$$

LEMMA 3.2. *Any m-type refinement $\mathscr{R} = \mathscr{R}(H)$ induces an m-type partition $\Theta(\mathscr{R})$ of Ω. Conversely, to any m-type partition $\Theta \subset \mathscr{C}$ of Ω one may assign an m-type refinement $\mathscr{R}(H, \Theta)$ with the following properties*

(i) *$\mathscr{R}(H, \Theta)$ is economic.*

(ii) *$\mathscr{R}(H, \Theta)$ is properly nested iff Θ is so.*

(iii) *The inequalities*

$$|\Theta| \leqslant |\mathscr{R}(H, \Theta)| \leqslant \gamma\, |\Theta|$$

hold uniformly in Θ, i.e., the constant γ depends only on m, k and s.

Proof. As for the first part of the assertion we let $\Theta(\mathscr{R})$ be the collection of all bottom faces (in Ω) of the bottom cubes in $\mathscr{R}$. This defines an m-type partition of Ω. Conversely choose for a given $\Theta \subset \mathscr{C}$ a sequence $\{\Theta_i\}_{i=0}^N$ such that $\Theta_0 = \{\Omega\}$, $\Theta_N = \Theta$ and Θ_{i+1} is obtained by one elementary m-type subdivision of some $C_i \in \Theta_i$, say. Setting $\mathscr{R}_0 = \mathscr{H}$ we inductively define $\mathscr{R}_{i+1}$ for $i = 1,..., N-1$ as follows. If some lower dimensional face of C_i is part of the boundary $\partial\Omega$ we subdivide $\varphi_i = \varphi(C_i, H) \in \mathscr{R}_i$ as well as all the neighboring bottom cubes φ with at least the same size as φ_i but with

$C(\varphi) \not\subset \Omega$. If C_i does not intersect the boundary, only φ_i is subdivided. It is not hard to see that $\mathscr{R}_N$ does not depend on the particular sequence $\{\Theta_i\}_{i=0}^N$. Hence it makes sense to define

$$\mathscr{R}(H, \Theta) := \mathscr{R}_N.$$

(i) and (ii) follow now immediately from the construction of $\mathscr{R}(H, \Theta)$.

As for (iii) we set $\mathscr{R}_\Omega = \{\varphi \in \mathscr{R}(H, \Theta): C(\varphi) \subset \Omega\}$ and denote by $\underline{\hat{\mathscr{R}}}_\Omega$ the collection of all bottom cubes in $\hat{\mathscr{R}}_\Omega$. It is then easily verified that

$$\begin{aligned} |\mathscr{R}(H, \Theta)| &\leqslant (3^s - 1)\,|\mathscr{R}_\Omega| \leqslant (3^s - 1)\,|\hat{\mathscr{R}}_\Omega| \leqslant (3^s - 1)\,m^k\,|\underline{\hat{\mathscr{R}}}_\Omega| \\ &= (3^s - 1)\,m^k\,|\hat{\Theta}| \leqslant 2(3^s - 1)\,m^k\,|\Theta|. \end{aligned}$$

Since the estimate $|\Theta| \leqslant |\mathscr{R}(H, \Theta)|$ is obvious, the proof is complete.

Combining (3.11) and Lemma 3.2(iii) we have that for any economic refinement $\mathscr{R}$ and $\Theta = \Theta(\mathscr{R})$ (cf. (2.3))

$$\dim \mathscr{S}_k(\mathscr{R}, \Omega) \leqslant \gamma\,|\Theta| \tag{3.15}$$

where γ depends only on s, k and the type m of Φ and $\mathscr{R}$. Hence the complexity of

$$\mathscr{S}_H(\Theta) := \mathscr{S}_k(\mathscr{R}(H, \Theta), \Omega) \tag{3.16}$$

is comparable to the space of piecewise polynomials

$$\Pi_{k,\Theta} := \{g: g\,|_C \in \Pi_k, C \in \Theta\}.$$

Defining the set

$$B_H(C \mid \Theta) := (\tilde{G}(\varphi(C, H) \mid \mathscr{R}(H, \Theta)))_s$$

we may summarize the previous results in

THEOREM 3.1. *Let the matrix H (see* (3.2)) *satisfy* (3.3), (3.5) *such that $m = m(H) < \infty$. Let Θ be an arbitrary properly nested m-type partition of Ω. The spline space $\mathscr{S}_H(\Theta)$* (3.16) *of degree k has the following properties*:

(i) $\dim \mathscr{S}_H(\Theta) \leqslant c_1\,|\Theta|$, *where c_1 does not depend on Θ*;

(ii) $\mathscr{S}_H(\Theta) \subseteq C^{k-1}(\Omega)$;

(iii) *there is an operator $T(\circ, \Theta): L_p(\Omega) \mapsto \mathscr{S}_H(\Theta)$ such that the estimate*

$$\|f - T(f, \Theta)\|_p(C) \leqslant c_2 \operatorname{dist}_p(f, \Pi_k)_{B_H(C \mid \Theta)}$$

holds for any $f \in L_p(\Omega)$ and every $C \in \Theta$, where the constant c_2 does not depend on f, C and Θ.

In fact, (i) follows from Definition (3.16) and (3.15) while (ii) is ensured by (3.10). Since $C \subseteq (\mathcal{q}(C, H))_s$, (iii) is an immediate consequence of Lemma 3.1.

Let us consider now, for any error function E satisfying (2.1) and any fixed $\gamma > 0$, the function

$$E_H(C) = \gamma E((\mathcal{q}(C, H))_s), \qquad C \in \mathscr{C}. \tag{3.17}$$

In fact, (3.6) implies that $(\mathcal{q}(C, H))_s \in \mathscr{C}_0$ for H satisfying (3.2), (3.3) so that E_H is well defined as a function on $\mathscr{C}$.

Lemma 3.3. *Suppose E satisfies* (2.1) *and E_H is defined by* (3.17). *Then*

(a) *E_H satisfies the conditions* (2.1) *as a function of cubes*;

(b) *for every fixed integer $m \geqslant 2$ one has*

$$A(E, m) \sim A(E_H, m).$$

Proof. (a) Since for any fixed H the ratio $r = \operatorname{diam}((\mathcal{q}(C, H))_s)/\operatorname{diam}(C)$ does not depend on $C \in \mathscr{C}$, we have $E_H(C) = (\operatorname{diam}(C))^b F_H(C)$ where $F_H(C) = \gamma r^b F((\mathcal{q}(C, H))_s)$. (2.1)(ii,iii) for E_H follow immediately from the assumptions on E.

(b) Let $A = A(E, m)$ and $A_H = A(E_H, m)$. Since $\gamma \geqslant 1$ and $C \subseteq (\mathcal{q}(C, H))_s$ we trivially have for all $\varepsilon > 0$

$$|\Phi_\varepsilon(A_1)| \leqslant |\Phi_\varepsilon(A_H)|.$$

On the other hand recalling (3.3), (3.6) the property (2.8)(ii) of the cover Y (cf. (2.7)) implies that

$$(\mathcal{q}(C, H))_s \subseteq \tilde{Y}(C \mid \Theta)$$

holds for any C in any partition Θ of Ω. This yields, because of (2.2) and (3.17),

$$E_H(C) \leqslant \gamma E(\tilde{Y}(C \mid \Theta))$$

for all $C \in \Theta$. Therefore

$$|\Phi_\varepsilon(A_H)| \leqslant |\Phi_\varepsilon(A(\gamma E \circ \tilde{Y}, m))|.$$

Since Y is a cover, Theorem 2.1 and finally Lemma 2.3 yield

$$|\Phi_\varepsilon(A(\gamma E \circ \tilde{Y}, m))| \leqslant \beta\, |\Phi_\varepsilon(A)|$$

where β is independent of ε. This finishes the proof.

Since by Definition (3.13), $G(\mathfrak{q} \mid \mathscr{R})$ contains, for any bottom cube $\mathfrak{q}$, only bottom cubes we may define because of Lemma 3.2 for any partition $\Theta \subset \mathscr{C}$

$$X_H(C \mid \Theta) = \{C(\mathfrak{q}'): \mathfrak{q}' \in G(\mathfrak{q}(C, H) \mid \mathscr{R}(H, \Theta))\} \tag{3.18}$$

so that we have in view of (3.17)

$$E_H(\tilde{X}_H(C \mid \Theta)) = \gamma E((\tilde{G}(\mathfrak{q}(C, H) \mid \mathscr{R}(H, \Theta))_s) = E(B_H(\Theta \mid H)). \tag{3.19}$$

LEMMA 3.4. *The map X_H defined by* (3.18) *is a cover* (*cf. Section* 2).

Proof. In view of (3.12), (3.13) and (3.18) conditions (2.8)(i)–(iv) and (2.9) are easily checked. As for (2.10) we note first that $\mathscr{R}(H, \Theta)$ is by Lemma 3.2(ii) properly nested iff Θ is so, too. Observing (3.3), (3.12) it is not hard to verify that in a properly nested refinement $\mathscr{R}$ either $\mathfrak{q}$ or $\hat{\mathfrak{q}}$ is protected, where $\mathfrak{q}$ is any bottom cube in $\hat{\mathscr{R}}$. Hence $G(\mathfrak{q} \mid \mathscr{R}(H, \Theta))$ involves only bottom cubes which differ from $\mathfrak{q}$ in at most one generation. This confirms that (2.10) holds, too.

Note that, since a large H-cube in $\mathscr{R}(H, \Theta)$ is not necessarily protected, we have in general $X_H(C \mid \Theta) \neq Y(C \mid \Theta)$ (cf. (2.7)).

We are now in a position to state

PROPOSITION 3.1. *Suppose that the error function E satisfies* (2.1) *and let E_H and X_H be defined by* (3.17) *and* (3.18), *respectively. Then*

$$A(E, m_1) \sim A(E_H \circ \tilde{X}_H, m_2)$$

for any two fixed integers $m_1, m_2 \geqslant 2$. Moreover, $A(E_H \circ \tilde{X}_H, m_2)$ is properly nested.

Proof. Note that in the proof of Theorem 2.1 conditions (2.1) are required to hold for E only as a function acting on $\mathscr{C}$. Hence combining Lemma 3.3(a), Lemma 3.4 and Theorem 2.1 ensures that $A(E_H \circ \tilde{X}_H, m_2)$ is properly nested and that

$$A(E_H, m_2) \sim A(E_H \circ \tilde{X}_H, m_2)$$

whence the assertion follows from Lemmas 2.2 and 3.3(b).

Now choose a matrix H such that (3.5) and hence (3.10) holds. Let $m = m(H)$ be given by Lemma 3.1. Defining X_H by (3.18) with respect to this H it is readily seen that the scheme $A(E_H \circ \tilde{X}_H, m)$ represents an adaptive approximation procedure which is based on the smooth spline spaces $\mathscr{S}_H(\Theta) \subset C^{k-1}(\Omega)$ (cf. (3.16), Theorem 3.1(ii).

Indeed, $A(E_H \circ \tilde{X}_H, m)$ produces by Proposition 3.1 at each stage only properly nested partitions Θ. So, when $\operatorname{dist}_p(f, \Pi_k)_C \leqslant E(C)$, Theorem 3.1(iii) says that $E_H \circ \tilde{X}_H$ represents for every current partition Θ valid local error bounds (cf. (2.3)) for $\|f - T(f, \Theta)\|_p(C)$, $C \in \Theta$.

Setting

$$\Pi_{k,m,N} = \bigcup \{\Pi_{k,\Theta} : |\Theta| \leqslant N,\ \Theta \text{ is of } m\text{-type}\}$$

we may summarize the above results in

THEOREM 3.2. *Suppose that for a given $f \in L_p(\Omega)$ the error function $E(C) = \operatorname{dist}_p(f, \Pi_k)_C$ satisfies* (2.1) *as well as $\sum_{C' \in d_r(C)} F(C')^p \leqslant \beta F(C)^p$, $C \in \mathscr{C}$, $r \in \mathbb{N}$. Then there exists a sequence of spline spaces $\mathscr{S}_{k,N}$ having degree k such that the following properties hold for any fixed integer $q \geqslant 2$:*

(i) $\dim \mathscr{S}_{k,N} \leqslant \gamma_1 N$;

(ii) $\mathscr{S}_{k,N} \subset C^{k-1}(\mathbb{R}^s)$;

(iii) $\operatorname{dist}_p(f, \mathscr{S}_{k,N})_\Omega \leqslant \gamma_2 \operatorname{dist}_p(f, \Pi_{k,q,N})_\Omega$, *where the constants γ_1, γ_2 depend only on s, k, q, β and b* (2.1).

Proof. For a suitable matrix H (cf. (3.2), (3.3)) satisfying (3.5) and $m = m(H) < \infty$ we define (cf. (3.17), (3.18))

$$A_H = A(E_H \circ \tilde{X}_H, m) \tag{3.20}$$

so that one has for any $\varepsilon > 0$ by definition and Theorem 3.1(iii)

$$\|f - T(f, \Phi_\varepsilon(A_H))\|_p(C) \leqslant \varepsilon, \qquad C \in \Phi_\varepsilon(A_H). \tag{3.21}$$

Proposition 3.1 tells us that

$$|\Phi_\varepsilon(A_H)| = \mathscr{O}(|\Phi_\varepsilon(A(E, q))|), \qquad \varepsilon \to 0. \tag{3.22}$$

Let $\varepsilon(N) = \operatorname{dist}_p(f, \Pi_{k,q,N})_\Omega / N^{1/p}$. Setting $\mathscr{S}_{k,N} = \mathscr{S}_H(\Phi_{\varepsilon(N)}(A_H))$ the appropriate choice of H makes certain that (ii) holds (cf. Theorem 3.1(ii)). Because of our assumptions on E we have on the one hand $E(C) \leqslant (1/d) \sum_{C' \in d_q(C)} E(C')$, $d = d(q)$ (cf. (2.1)), while on the other hand $\sum_{C' \in d_q(C)} E(C')^p \leqslant (\beta/m^{bp}) E(C)$ holds for every $C \in \mathbb{C}$. It is not hard to conclude from this that when $\operatorname{dist}_p(f, \Pi_{k,q,N})_\Omega = \operatorname{dist}_p(f, \Pi_{k,\Theta})_\Omega$ the errors $E(C)$, $C \in \Theta$, are balanced (up to constants depending on q,. b, β). This, in turn, implies ultimately that $|\Phi_{\varepsilon(N)}(A(E, q))| \leqslant c(b, q, \beta) N$. Thus part (i) follows from Theorem 3.1(i) and (3.22). Finally (3.21) yields $\operatorname{dist}_p(f, \mathscr{S}_{k,N})_\Omega \leqslant c(N\varepsilon(N)^p/N)^{1/p} = c \operatorname{dist}_p(f, \Pi_{k,q,N})_\Omega$ confirming (iii).

As an application we consider the following classes of functions (cf. [4]). Let M be a smooth manifold of dimension $r < s$ in Ω. Suppose $f \in C^{k+1}(\Omega \backslash M)$ satisfies, for $C \in \mathscr{C}_0$ and some fixed $a > 0$ such that

$$k + 1 > a > r(k+1)/s - (s-r)/p,$$

the estimate

$$\sup_{x \in C} \max_{|\alpha| = k+1} |(D^\alpha f)(x)| \leqslant \text{const}_f \operatorname{dist}(M, C)^{a-k-1} \tag{3.23}$$

where $\operatorname{dist}(M, C) = \inf\{|u - z|: u \in M, z \in C\}$ and $|\cdot|$ is the Euclidean norm. Moreover assume that we have in the case $C \cap M = \varnothing$

$$\operatorname{dist}_p(f, \Pi_k)_C \leqslant \text{const}_f \operatorname{vol}_s(C)^{1/p} \operatorname{diam}(C)^a. \tag{3.24}$$

COROLLARY 3.1. *Let f satisfy* (3.23), (3.24) *and* A_H *be defined by* (3.20). *Then one has*

$$|\Phi_\varepsilon(A_H)| \leqslant \text{const}_f \varepsilon^{-1/((k+1)/s + 1/p)}$$

as well as

$$\inf_{|\Theta| \leqslant N} \operatorname{dist}_p(f, \mathscr{S}_H(\Theta))_\Omega = \mathscr{O}(N^{-(k+1)/s}), \qquad N \to \infty.$$

Proof. It is not hard to check that the error bound for $\operatorname{dist}_p(f, \Pi_k)_C$ given in [4] satisfies (2.1). The assertion follows then by Theorem 3.2 and [4, Theorem 4.1, Corollary 4.1].

Final remarks. Since the smooth approximations (3.14) have to be recomputed only locally after refining a current partition the equivalence between smooth and non-smooth adaptive approximation is not only valid in the sense of Definition 2.1 but also (at least asymptotically) with respect to the total amount of computational work.

For the sake of simplicity we have restricted our considerations to refinements of uniform configurations for simple rectangular domains (and unions of such). Due to the structure of the B-splines these restrictions are clearly not essential and, e.g., appropriately distorted configurations near the boundary of the respective domain Ω would match with much more general domains than considered above, a fact which is for instance in contrast to refinements of rectangular grids say, for tensor product type splines.

However, when dispensing with highest possible global smoothness and when dealing with simple (rectangular) domains the results of Section 2 are easily applied to adaptive approximation with respect to smooth splines on local refinements of rectangular grids as well.

To keep things simple let Ω be again the unit s-cube and Θ be some m-type partition of Ω where $m/r = k + 1$, $r \in \mathbb{N}$, for a given $k \in \mathbb{N}$. As in the proof of Lemma 3.2 let Θ be enlarged to a collection Θ' of s-cubes by adding cubes outside Ω touching $\partial\Omega$ and whose size coincides with that of the corresponding neighbors in Ω. Clearly we have

$$|\Theta'| \leqslant c\,|\Theta|, \tag{3.25}$$

where c depends only on s. With each cube $C=[a, a+\underline{h}]=[a_1, a_1+h] \times \cdots \times [a_s, a_s+h]$ in Θ' we associate a uniform rectangular mesh with the lattice points $a+y\nu$, $0 \leqslant \nu_i \leqslant k+1$, $y=h/(k+1)$.

Denoting by Ω_h the union of all cubes in Θ' containing lattice points with step size $\leqslant h/(k+1)$ we set

$$\underline{\mathscr{S}}_{k,h}=\operatorname{span}\{M_{a,h,k}|_{\Omega}: a+h\nu/(k+1) \in \Omega_h, 0 \leqslant \nu_i \leqslant k+1\},$$

where $M_{a,h,k}(x)=\prod_{j=1}^{s} M(x_j \mid a_j, a_j+h/(k+1),..., a_j+h)$ is the usual tensor product B-spline belonging to $\underline{C}^{k-1}(\mathbb{R}^s)=\{f: D^\alpha f \in C(\mathbb{R}^s),\ \alpha_i \leqslant k-1,\ i=1,..., s\}$.

Denoting by $\{h_i\}_{i=0}^{d}$ the sequence of side lengths corresponding to the different generations of the elements of Θ' we may now assign to a given m-type partition Θ, $m/r=k+1$, the spline space

$$\underline{\mathscr{S}}_k(\Theta)=\bigoplus_{i=0}^{d} \underline{\mathscr{S}}_{k,h_i},$$

where $\oplus$ means the direct sum. Again we have by construction

$$\underline{\mathscr{S}}_k(\Theta) \subseteq \underline{C}^{k-1}(\Omega)$$

and in view of (3.25)

$$\dim \underline{\mathscr{S}}_k(\Theta) \leqslant c\,|\Theta|,$$

where c depends only on s and k. Moreover, suppose $C \in \hat{\Theta}'$ belongs to the jth generation and $d_m(C) \subseteq \Theta'$. When $C_0 \in d_m(C)$ is an inner cube in Θ' we have certainly

$$\underline{\mathscr{S}}_k(\Theta)|_{C_0}=\underline{\mathscr{S}}_{k,h_j}|_{C_0}.$$

Thus the same arguments as in the proof of Theorem 3.1 in [7] combined with known local estimates for tensor product spline approximation (cf., e.g., [3]) yield an analog to Theorem 3.1 above:

PROPOSITION 3.2. *Let Θ be a properly nested m-type partition of Ω, $m/r=k+1$. Then there is an operator $T: L_p(\Omega) \to \underline{\mathscr{S}}_k(\Theta)$ such that the following estimate holds for any $f \in L_p(\Omega)$, $C \in \Theta$*

$$\| Tf-f \|_p(C) \leqslant c \operatorname{dist}_p(f, \underline{\Pi}_k)_{\tilde{Y}(C|\Theta)}$$

(cf. (2.7)) where c depends only on s, k, m and $\underline{\Pi}_k$ is the space of polynomials of coordinate degree $\leqslant k$.

Since Y is a cover (cf. Section 2) the previous line of arguments provides analogous results for adaptive approximation by splines of the type $\underline{\mathscr{S}}_k(\Theta)$, in particular, using the refining strategy (2.11) or the respective interpretation when an error function for the smooth approximations is given directly.

ACKNOWLEDGMENTS

We would like to acknowledge here a private communication by R. DeVore who has established local error estimates similar to those in Proposition 3.2 for special approximation schemes based on univariate Hermite interpolation where the grid structure is similar to the one considered here.

REFERENCES

1. R. E. BARNHILL, Representation and approximation of surfaces, *in* "Mathematical Software III" (J. R. Rice, Ed.), pp. 68–119, Academic Press, New York, 1977.
2. M. S. BIRMAN AND M. Z. SOLOMIAK, Piecewise-polynomial approximations of functions of the classes W_p^α, *Mat. Sb.* **73** (1967), 295–317.
3. C. DE BOOR, Splines as linear combinations of B-splines, *in* "Approximation Theory II" (G. G. Lorentz, C. K. Chui, and L. L. Schumaker, Ed.), pp. 1–47, Academic Press, New York, 1976.
4. C. DE BOOR AND J. R. RICE, An adaptive algorithm for multivariate approximation giving optimal convergence rates, *J. Approx. Theory* **25** (1979), 337–359.
5. W. DAHMEN, On multivariate B-splines, *SIAM J. Numer. Anal.* **17** (1980), 179–191.
6. W. DAHMEN, Approximation by linear combinations of multivariate B-splines, *J. Approx. Theory* **31** (1981), 299–324.
7. W. DAHMEN, Approximation by smooth multivariate splines on nonuniform grids, *in* "Quantitative Approximation" (R. A. DeVore and K. Scherer, Eds.), pp. 99–114, Academic Press, New York, 1980.
8. W. DAHMEN, R. DEVORE, AND K. SCHERER, Multi-dimensional spline-approximation, *SIAM J. Numer. Anal.* **17** (1980), 380–402.
9. H. W. KUHN, Some combinatorial lemmas in topology, *IBM J. Res. Develop.* **45** (1960), 518–524.
10. C. A. MICCHELLI, A constructive approach to Kergin interpolation in $\mathbb{R}^k$: multivariate B-splines and Lagrange interpolation, *Rocky Mountain J. Math.* **10** (1980), 485–497.
11. C. A. MICCHELLI, On a numerically efficient method for computing multivariate B-splines, *in* "Proc. Conf. Multivariate Approximation Theory," pp. 211–248, Birkhäuser, Basel, 1979.
12. J. R. RICE, Multivariate piecewise polynomial approximation, *in* "Multivariate Approximation" (D. C. Handscomb, Ed.), pp. 261–277, Academic Press, New York, 1978.