A piecewise constant algorithm for weighted L_1 approximation over bounded or unbounded regions in $\mathbb{R}^s$

Hickernell, Fred J.; Sloan, Ian H.; Wasilkowski, Grzegorz W.

Published in:
SIAM Journal on Numerical Analysis

DOI:
10.1137/S0036142903427664

Published: 01/01/2005

Document Version:
Publisher's PDF, also known as Version of record

SIAM J. NUMER. ANAL.
Vol. 43, No. 3, pp. 1003–1020

A PIECEWISE CONSTANT ALGORITHM FOR WEIGHTED L_1 APPROXIMATION OVER BOUNDED OR UNBOUNDED REGIONS IN $\mathbb{R}^s$*

FRED J. HICKERNELL†, IAN H. SLOAN‡, AND GRZEGORZ W. WASILKOWSKI§

Abstract. Using Smolyak's construction [S. A. Smolyak, *Dokl. Akad. Nauk SSSR,* 4 (1963), pp. 240–243], we derive a new algorithm for approximating multivariate functions over bounded or unbounded regions in $\mathbb{R}^s$ with the error measured in a weighted L_1-norm. We provide upper bounds for the algorithm's cost and error for a class of functions whose mixed first order partial derivatives are bounded in the L_1-norm. In particular, we prove that the error and the cost (measured in terms of the number of function evaluations) satisfy the relation

$$\text{error} \le \frac{s\,\exp\left(\frac{1}{12(s-1)}\right)}{(s-1)\pi}\left(\frac{e\,\ln(\text{cost})}{(s-1)\sqrt{2}\ln(2)}\right)^{2(s-1)}\frac{1}{\text{cost}}$$

whenever the cost is sufficiently large relative to the number s of variables. More specifically, the inequality holds when $q \ge 2(s-1)$, where q is a special parameter defining the refinement level in Smolyak's algorithm, and hence the number of function evaluations used by the algorithm. We also discuss extensions of the results to the spaces with the derivatives bounded in L_p-norms.

Key words. Banach spaces, mixed first order partial derivatives, multivariate functions, Smolyak's construction

AMS subject classifications. 65D05, 65D15

DOI. 10.1137/S0036142903427664

1. Introduction.

In this paper, we derive a simple-to-use, piecewise constant algorithm for approximating functions in a weighted L_1 sense. Function approximation has been studied quite extensively; see, e.g., [13, 17, 22, 24] and the papers cited therein. However, such problems were considered mainly for functions with a bounded domain D, say $D = [0,1]^s$.

The worst case complexity of weighted approximation over unbounded domains D has recently been studied in, e.g., [11, 27], assuming that the corresponding function classes $\mathcal{F}$ are isotropic. The analysis of the approximation problem for tensor product spaces $\mathcal{F}$ is quite straightforward if $\mathcal{F}$ is a Hilbert space, since then desirable properties of Smolyak's construction could be used; see, e.g., [25].

In this paper, we study a weighted approximation problem with an emphasis on unbounded domains D and tensor product function classes $\mathcal{F}$ in a non-Hilbert-space setting. More specifically, we study a ρ-weighted L_1 approximation problem with the

*Received by the editors May 5, 2003; accepted for publication (in revised form) September 24, 2004; published electronically September 20, 2005. This work was partially supported by the Hong Kong Research Grants Council under grant HKBU/2020/02P, the Hong Kong Baptist University Faculty Research under grant FRG/00-01/II-62, the Australian Research Council, and the National Science Foundation under grant CCR-0095709.

http://www.siam.org/journals/sinum/43-3/42766.html

†Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong. Current address: Department of Applied Mathematics, Illinois Institute of Technology, Chicago, IL 60616 (hickernell@iit.edu).

‡School of Mathematics, University of New South Wales, Sydney 2052, Australia (sloan@maths.unsw.edu.au).

§Department of Computer Science, University of Kentucky, 773 Anderson Hall, Lexington, KY 40506-0046 (greg@cs.uky.edu).

error between f and the approximation $\mathcal{A}(f)$ measured in the following seminorm:

$$\|(f - \mathcal{A}(f))\,\rho\|_{L_1(D)} = \int_D \rho(\mathbf{x})\,|f(\mathbf{x}) - \mathcal{A}(f)(\mathbf{x})|\,d\mathbf{x}. \tag{1}$$

We have chosen the L_1-norm here (as opposed to the L_r-norm for $r > 1$) since then the approximation problem is related to weighted integration. This relation is briefly discussed in section 6; here we only mention that any algorithm $\mathcal{A}$ for L_1 approximation yields an integration algorithm with the error at least as small as the error of $\mathcal{A}$. Another reason for choosing $r = 1$ is the simplicity of analysis; for $r \neq 1, 2$ the corresponding weighted L_r approximation problem is more difficult to analyze.

Throughout this article, it is assumed that $\rho(\mathbf{x}) = \prod_{k=1}^{s} \rho_k(x_k)$ is a given integrable weight function. The domain D of functions can be an arbitrary box, even $D = \mathbb{R}^s$. The class $\mathcal{F}$ is a space of functions whose dominating mixed first order derivatives are bounded in the L_1-norm. This seems to be about the minimum amount of smoothness needed to get a convergence rate of essentially $O(n^{-1})$. For example, if only the functions have bounded L_1-norm, then function values may not even be finite, and algorithms depending on function values are not be to guaranteed to converge. The case of derivatives bounded in L_p-norms seems to be harder than the case with the L_1-norm (especially for unbounded D), and is briefly addressed in section 5.

We stress that classes $\mathcal{F}$ as defined here are commonly assumed in the context of integration problems over $D = [0,1]^s$, especially when dealing with quasi-Monte Carlo methods and discrepancy; see, e.g., [5, 12, 19] and papers cited therein. This point is further discussed in section 2. These $\mathcal{F}$ have been used even for integration problems with unbounded domains [9]. Moreover, for bounded D, these classes contain the classes MW_p^1 of periodic functions considered for approximation in [22, Chap. 4].

Instead of a condition on the dominating mixed derivatives, one sometimes assumes that all derivatives of total order α are bounded in the L_1-norm. Error analysis using such conditions suffer from the "curse of dimensionality." For example, to have a convergence rate of $O(n^{-1})$ requires all derivatives of total order s to be bounded in the L_1-norm. The condition assumed here is weaker.

The main result of the paper is the derivation and analysis of a family $\{\mathcal{A}_{q,s}\}_{q=s}^{\infty}$ of algorithms that provide approximations that are special piecewise constant functions. They are obtained by applying Smolyak's construction (see [21]) to scalar piecewise constant interpolation methods. As we shall see, given a number s of variables and a parameter $q \geq s$, the algorithm has the worst case error bounded by

$$\mathrm{error}(\mathcal{A}_{q,s}) \leq \begin{cases} s\,2^{-2q+2s-3}\binom{2q-2s+3}{q-s+1} & \text{if } q < 2(s-1), \\ s\,2^{-q}\binom{q}{s-1} & \text{if } q \geq 2(s-1). \end{cases}$$

Here q is the refinement parameter in Smolyak's algorithm. Under an additional symmetry assumption (4) that is stated later, we get

$$\mathrm{error}(\mathcal{A}_{q,s}) \leq s\,2^{-q+s-1-a}\binom{q-s}{a} \quad \text{with} \quad a = \begin{cases} \lfloor \frac{q-s+1}{3} \rfloor & \text{if } q < 4(s-1), \\ s-1 & \text{otherwise.} \end{cases}$$

Let $n = \mathrm{card}(\mathcal{A}_{q,s})$ denote the number of function evaluations used by $\mathcal{A}_{q,s}$. From [25, Lem. 7] we have $n \leq 2^{q-s+1}\binom{q-1}{s-1}$, and we conclude that for every $s \geq 2$ and $q \geq 2(s-1)$,

$$\mathrm{error}(\mathcal{A}_{q,s}) \leq \frac{s}{(s-1)\pi}\left(\frac{e\,\ln(n)}{(s-1)\sqrt{2}\ln(2)}\right)^{2(s-1)}\frac{1}{n}.$$

This rate $n^{-1}\ln^{2(s-1)}(n)$ of convergence is as good as the best known rate for L_1 approximation of periodic functions from the class $\mathcal{F} = MW_1^1$; see [22, Thm. 5.1 of Chap. 4] and [23]. Recall that MW_1^1 is contained in $\mathcal{F}$ when D^s is bounded and the approximation is with $\rho \equiv 1$. Hence $\mathcal{A}_{q,s}$ works for more general classes of problems and spaces of functions. Moreover, the result above shows in an explicit way the dependence of the errors on the dimension s.

We do not know whether $n^{-1}\ln^{2(s-1)}(n)$ is the best possible rate. However, because the Kolmogorov width for the corresponding problem over MW_1^1 equals $\Theta(n^{-1}\ln^{3(s-1)/2}(n))$ (see [22, Thm. 4.5 of Chap. 3]), the difference could only be in the exponent of the $\ln(n)$-term.

Since $\mathcal{A}_{q,s}(f)$ is a piecewise constant function, the algorithm is easy to implement. Its only drawback is in exponential gaps between consecutive numbers $\mathrm{card}(\mathcal{A}_{q,s})$ of function evaluations. However, we believe that it is of a practical interest, especially for small to moderate values of s. Implementation of the algorithm and numerical tests will be reported later.

We stress that Smolyak's construction, also referred to as *sparse grid*, *hyperbolic cross*, and *Boolean blending*, has been applied to many problems, including differential and integral equations, integration and approximation of multivariate functions, and to wavelets construction; see, e.g., [1, 2, 3, 4, 6, 7, 8, 10, 16, 18, 14, 22, 23, 25, 26] and the papers cited therein. However, with only a few exceptions, the considered functions are defined over bounded domains (e.g., $D = [0,1]^s$), and sometimes are of a special form (e.g., $f = g * B$ for a fixed function B as in MW_p^α classes). By analyzing ρ-weighted L_1 approximation, we are able to propose a simple algorithm that works well for any probability density function ρ of a tensor product form and even for $D = \mathbb{R}^s$. The proposed algorithm uses a finite number n of function values and approximates the function with error $O(n^{-1}\ln^{2(s-1)}(n))$. It is linear, simple, and easy to implement.

We would like to contrast this to a possible wavelet approach with $D = \mathbb{R}^s$ as proposed in [4]. There, the approximating algorithm is defined as the function from a special linear subspace $\mathcal{H}_n$ that minimizes the L_p distance from f, and the space $\mathcal{H}_n$ has infinite dimension. Hence, it is very difficult to implement. This should explain our interest in ρ-weighted approximation. We have chosen the L_1-norm since for the L_r-norm ($r > 1$) the analysis would be much harder and getting sharp bounds (including explicit constants) rather impossible.

We now explain the choice of low regularity $\alpha = 1$. Recall that we assume first order mixed derivatives to be bounded. As in the papers cited in the previous paragraph that deal with classes of functions of arbitrary (but fixed) regularity and bounded domains, it is possible to extend the results even to the case of $D = \mathbb{R}^s$. Indeed, using general properties of Smolyak's construction (e.g., from [25]) and results on the complexity of scalar ($s = 1$) weighted L_r approximation over $\mathbb{R}$ (see [27]), it is possible to achieve a convergence rate $O(n^{-(\alpha+(1/r-1/p)_-)}\ln^{(\alpha+1)(s-1)}(n))$. This is when the mixed partial derivatives of order α are bounded in the (weighted) L_p-norm. However, this extension to general α would (i) hold only under special assumptions on ρ (see [27, Thm. 1]), (ii) make explicit dependence on the dimension s very difficult to obtain, and (iii) make the resulting algorithms more difficult to implement and less applicable. In particular, the gaps between consecutive values of $\mathrm{card}(\mathcal{A}_{q,s})$ would increase with increasing regularity α, making the algorithm applicable only for very small values of s (say $s \leq 3$). Moreover, by assuming $\alpha = 1$ we make our algorithm applicable to a larger class of functions also because the classes with $\alpha > 1$ are contained in the class with $\alpha = 1$. The explicitness of the results, as well as the simplicity and applicability of the algorithm, were our primary reasons for choosing $\alpha = 1$.

We summarize the content of this paper. Section 2 provides some basic definitions and assumptions, as well as an error bound for an arbitrary algorithm $\mathcal{A}$. Since Smolyak's construction depends on the specific choice of scalar algorithms, section 3 considers very special scalar algorithms based on a piecewise constant interpolation. The corresponding algorithm and its properties are presented in section 4. An extension to functions with derivatives bounded in L_p-norm is provided in section 5. Section 6 briefly explains why the error bounds obtained for the ρ-weighted L_1 approximation also hold for the corresponding ρ-weighted integration problem.

2. Basic definitions. In this section, we briefly present some definitions and basic facts concerning the worst case setting. A more detailed discussion can be found, e.g., in [13, 24].

We consider a weighted L_1 approximation of functions of s variables whose domain D is a box,

$$D = \overline{(a_1, b_1)} \times \cdots \times \overline{(a_s, b_s)}.$$

The values a_i and b_i might be infinite; this is why we write $\overline{(a_i, b_i)}$ instead of $[a_i, b_i]$.

Let $\mathcal{F}$ be a Banach space of functions $f : D \to \mathbb{R}$ that will be specified later. The approximation problem depends on a weight function ρ which is assumed to have the following properties:

$$\rho(\mathbf{x}) = \prod_{k=1}^{s} \rho_k(x_k) \quad \text{and} \quad \rho_k \geq 0. \tag{2}$$

For simplicity of presentation, we also assume that

$$\int_{a_k}^{b_k} \rho_k(t)\, dt = 1 \quad \forall\, k = 1, \ldots, s.$$

However, it is enough to assume that the integrals of ρ_k are finite; in such a case, all error bounds derived in this paper should be multiplied by the constant $c = \int_D \rho(\mathbf{x})\, d\mathbf{x}$.

Functions from $\mathcal{F}$ are approximated by an algorithm $\mathcal{A}$,

$$f \sim \mathcal{A}(f) = \sum_{i=1}^{n} f(\mathbf{x}^i)\, g_i$$

for some points $\mathbf{x}^i$ and functions g_i, with the error between f and $\mathcal{A}(f)$ measured in the ρ-weighted L_1-norm; see (1). The *worst case error* (with respect to $\mathcal{F}$) of $\mathcal{A}$ is defined by

$$\text{error}(\mathcal{A}) := \sup_{\|f\| \leq 1} \|(f - \mathcal{A}(f))\, \rho\|_{L_1(D)},$$

where $\|f\|$ denotes the norm of f in the space $\mathcal{F}$. The importance of this definition is that due to linearity of $\mathcal{A}$ we have

$$\|(f - \mathcal{A}(f))\, \rho\|_{L_1} \leq \|f\|\ \text{error}(\mathcal{A}) \quad \forall\, f \in \mathcal{F}.$$

Each algorithm uses a finite number n of function evaluations. That number is called the *cardinality* and is denoted by $\mathrm{card}(\mathcal{A})$.

With the exception of sections 5 and 6, the following space $\mathcal{F}=\mathcal{F}_{1,s}$ is considered. Let $\mathcal{H}_k$ be the space of absolutely continuous functions on $\overline{(a_k,b_k)}$ whose first derivative is in $L_1((a_k,b_k))$. Let $\mathcal{H}^s = \bigotimes_{k=1}^s \mathcal{H}_k$ be the space consisting of linear combinations of functions f of the tensor product form

$$f : D \to \mathbb{R} \qquad \text{and} \qquad f(\mathbf{x}) = \prod_{k=1}^{s} h_k(x_k) \quad \text{with} \quad h_k \in \mathcal{H}_k.$$

The space $\mathcal{F}_{1,s}$ is the completion of $\mathcal{H}^s$ with respect to the following norm:

$$\|f\|_{1,s} := |f(\mathbf{c})| + \sum_{U \neq \emptyset} \|f'_U\|_{L_1(D_U)}. \tag{3}$$

Here $\mathbf{c} = [c_1, \ldots, c_s] \in D$ is a fixed point, called an *anchor*. The summation is with respect to nonempty subsets U of $\{1, \ldots, s\}$, and

$$f'_U(\mathbf{x}_U) := \frac{\partial^{|U|}}{\prod_{k \in U} \partial x_k} f(\mathbf{x}_U, \mathbf{c}),$$

where $(\mathbf{x}_U, \mathbf{c})$ denotes the s-dimensional vector whose kth component is x_k if $k \in U$, and c_k otherwise. By $\mathbf{x}_U$ we mean the $|U|$-dimensional vector obtained from $\mathbf{x}$ by removing all components x_k with $k \notin U$. This means that f'_U is a function defined on $D_U := \prod_{k \in U} \overline{(a_k, b_k)}$ and $\mathbf{x}_U \in D_U$. To simplify the notation, we will also write $f'_\emptyset$ and $\|f'_\emptyset\|_{L_1}$ to denote $f(\mathbf{c})$ and $|f(\mathbf{c})|$, respectively; and we often drop D_U by writing $\|\cdot\|_{L_1}$ instead of $\|\cdot\|_{L_1(D_U)}$. This allows the more concise formula

$$\|f\|_{1,s} = \sum_{U} \|f'_U\|_{L_1}.$$

We illustrate this for $s = 2$:

$$\begin{aligned}\|f\|_{1,2} = |f(c_1, c_2)| &+ \int_{a_1}^{b_1} \left| \frac{\partial}{\partial x_1} f(x_1, c_2) \right| dx_1 + \int_{a_2}^{b_2} \left| \frac{\partial}{\partial x_2} f(c_1, x_2) \right| dx_2 \\ &+ \int_{a_1}^{b_1} \int_{a_2}^{b_2} \left| \frac{\partial^2}{\partial x_1 \partial x_2} f(x_1, x_2) \right| dx_2\, dx_1.\end{aligned}$$

Although in general the anchor $\mathbf{c}$ and the weight ρ are not related, we shall obtain stronger results under the following symmetry condition:

$$\int_{a_k}^{c_k} \rho_k(x)\, dx = \frac{1}{2} \quad \forall k = 1, \ldots, s. \tag{4}$$

Of course, this condition is satisfied if (a_k, b_k) and ρ_k are symmetric with respect to c_k ($k = 1, \ldots, s$), e.g., if $a_k = -b_k$, $c_k = 0$, and $\rho_k(x) = \rho_k(-x)$. In general, for a given ρ_k one may always choose c_k to satisfy (4).

We now comment on the norm (3) and the role of the anchor $\mathbf{c}$.

Remark 1. As already mentioned in the introduction, the norm (3) and the space $\mathcal{F}_{1,s}$ are frequently assumed/studied in the context of quasi-Monte Carlo integration and discrepancies when the domain D is bounded, $D = [0, 1]^s$. Then, classically, the anchor $\mathbf{c} = [1, \ldots, 1]$. Of course, when $D = \mathbb{R}^s$ such a choice of the anchor seems unjustified and this is why we prefer to deal with arbitrary $\mathbf{c}$; see [9] for additional discussion.

There are a number of important results when, instead of the norm (3), a seminorm

$$|\!|\!| f |\!|\!|_{\alpha,p,s} := \|f^{(\alpha,\ldots,\alpha)}\|_{L_p} \quad \text{with} \quad f^{(\alpha,\ldots,\alpha)} = \left(\prod_{k=1}^{s} \frac{\partial^\alpha}{\partial x_k^\alpha}\right) f$$

is assumed. Of course, for $\alpha = 1$ and $p = 1$ this seminorm is very much related to (3) since it is equivalent to $\|f'_U\|_{L_1}$ with $U = \{1, \ldots, s\}$. In particular, for $s = 1$, the complexities of problems with bounded $\|f\|_{1,s}$ or $|\!|\!| f |\!|\!|_{1,1,s}$ are equivalent. The situation changes drastically in the the multivariate case $s \geq 2$. This is because the subspace of functions with vanishing $|\!|\!| \cdot |\!|\!|_{\alpha,p,s}$-seminorm has infinite dimension. Therefore, to guarantee finite errors, one has to add additional restrictions on the considered class of functions. Examples of such restrictions include periodicity as in classes MW_p^α, or boundary conditions such as the vanishing of f and its partial derivatives at the points $\mathbf{x}$ with at least one component equal to zero. The presence of $\|f'_U\|_{L_1}$-terms in the definition of (3) guarantees that it is a well-defined norm without additional restrictions on the class of functions.

The following fact will play an important role. Let

$$M_k(x,t) := \begin{cases} 1 & \text{if } c_k < t < x, \\ -1 & \text{if } x < t < c_k, \\ 0 & \text{otherwise,} \end{cases}$$

$$M(\mathbf{x},\mathbf{t}) := \prod_{k=1}^{s} M_k(x_k, t_k) \quad \text{and} \quad M_U(\mathbf{x}_U, \mathbf{t}_U) := \prod_{k \in U} M_k(x_k, t_k),$$

with the convention that $M_\emptyset \equiv 1$. Then for every $f \in \mathcal{F}_{1,s}$ and every $\mathbf{x} \in D$,

$$f(\mathbf{x}) = \sum_U \int_{D_U} f'_U(\mathbf{t}_U)\, M_U(\mathbf{x}_U, \mathbf{t}_U)\, d\mathbf{t}_U. \tag{5}$$

The representation (5) has been used, at least implicitly, in a number of papers, and its short proof can be found in [9]. From it we have

$$\|f\,\rho\|_{L_1} \leq \sum_U \int_{D_U} |f'_U(\mathbf{t}_U)| \int_D |\rho(\mathbf{x})\, M_U(\mathbf{x}_U, \mathbf{t}_U)|\, d\mathbf{x}\, d\mathbf{t}_U \leq \|f\|_{1,s}.$$

This means that the approximation problem is well defined since the corresponding embedding operator is bounded. Actually, the following theorem, when applied to the zero algorithm $\mathcal{A} \equiv 0$, implies that the norm of the embedding is equal to one, i.e.,

$$\sup_{f \in \mathcal{F}} \frac{\|f\,\rho\|_{L_1}}{\|f\|_{1,s}} = 1.$$

THEOREM 1. *The error of any $\mathcal{A}$ is bounded by*

$$\text{error}(\mathcal{A}) \leq \sup_{\mathbf{t} \in D} \max_U \int_{D_U} h_U(\mathbf{x}_U, \mathbf{t}_U)\, d\mathbf{x}_U, \tag{6}$$

where

$$h_U(\mathbf{x}_U, \mathbf{t}_U) = \rho_U(\mathbf{x}_U)\,|M_U(\mathbf{x}_U, \mathbf{t}_U) - \mathcal{A}(M_U(\cdot, \mathbf{t}_U))(\mathbf{x}_U)|.$$

If $\mathcal{A}$ is based on piecewise constant interpolation then (6) *holds with equality.*

Proof. We defer the proof to section 5, where a more general result is proven. □

3. Scalar functions. Since Smolyak's construction depends on specific algorithms for the scalar cases, we consider now approximating univariate functions whose domain is $\overline{(a_k, b_k)}$ and whose weight function is ρ_k. To simplify the notation in this section, we write a, b, and ω instead of a_k, b_k, and ρ_k.

For $i = 1, 2, \dots$, consider the set of points $x_{i,j}^k$ ($j = 0, \dots, 2^i$) such that

$$a = x_{i,0}^k < x_{i,1}^k < \cdots < x_{i,2^i}^k = b$$

and

$$\int_{x_{i,j-1}^k}^{x_{i,j}^k} \omega(t)\,dt = 2^{-i}. \tag{7}$$

For simplicity, we will write in this section $x_{i,j}$ instead $x_{i,j}^k$.

Of course, $x_{i-1,j} = x_{i,2j}$. We take the following family of algorithms A_i ($i = 1, 2, \dots$) based on piecewise constant interpolation:

$$A_i(f)(x) = f(x_{i,j}) \qquad \text{if } c_k \le x_{i,j} \le x < x_{i,j+1} \quad \text{or} \quad x_{i,j-1} < x \le x_{i,j} \le c_k.$$

Moreover, when $x, c_k \in (x_{i,j}, x_{i,j+1})$ then $A_i(f)(x)$ equals $f(x_{i,j})$ or $f(x_{i,j+1})$ depending on whether or not $x \le c_k$. Note that under the symmetry condition (4), $x_{1,1} = x_{i,2^{i-1}} = c_k$ and $A_i(f)(x) = f(c_k)$ for $x \in (x_{i,2^{i-1}-1}, x_{i,2^{i-1}+1})$.

For given k, define

$$\delta_{k,1}(x,t) := A_1(M_k(\cdot, t))(x)$$

and

$$\delta_{k,i}(x,t) := A_i(M_k(\cdot, t))(x) - A_{i-1}(M_k(\cdot, t))(x), \quad i \ge 2. \tag{8}$$

The following result is needed in section 4.

PROPOSITION 1. *For every t, we have*

$$\|\omega\,\delta_{k,i}(\cdot, t)\|_{L_1} \le 2^{-i} \quad \forall\, i \ge 2.$$

Moreover, $A_1(M_k(\cdot, t))$ is a constant function equal to ± 1 or zero, and if (4) *holds, then*

$$A_1(M_k(\cdot, t)) = 0.$$

The second part of the proposition can be directly checked. The first part of the proposition is an immediate consequence of the following lemma that will be used in section 5. For simplicity of presentation, we state the lemma only for arguments x and t greater than c_k. The analogous result is true for arguments smaller than c_k with the only difference being that $\delta_{i,k}(x,t) \in \{-1, 0\}$. Moreover, $\delta_{k,i}(x,t) = 0$ when c_k is between x and t.

LEMMA 1. *The following statements hold for any* $i \geq 2$ *and* $x, t > c_k$.

(i) $\delta_{k,i}(x,t) \in \{0,1\}$.

(ii) *If* $\delta_{k,i}(x,t) = 1$, *then there exists* $j \leq 2^{i-1} - 1$ *such that*

$$t \in (x_{i,2j}, x_{i,2j+1}] \qquad \text{and} \qquad x \in (x_{i,2j+1}, x_{i,2j+2}].$$

(iii) *There is at most one* $i \geq 2$ *such that* $\delta_{k,i}(x,t) = 1$.

Proof. (i) follows immediately from the facts that $M_k(x,t) \in \{0,1\}$, $M_k(\cdot,t)$ is nondecreasing, and A_i uses the points used by A_{i-1}.

(ii) Let $x \in (x_{i,\ell}, x_{i,\ell+1}]$ for some ℓ. Then $A_i(M_k(\cdot,t))(x) = 1$ only if $t < x_{i,\ell}$. However, for $A_{i-1}(M_k(\cdot,t))(x) = 0$, $x_{i,\ell}$ has to be different than any point $x_{i-1,j}$ used by A_{i-1}. This means that ℓ has to be odd.

(iii) follows from (ii). Indeed, let (x,t) be fixed. If $\delta_{k,i}(x,t) = 1$, then t and x are in two neighboring subintervals with the evaluation point $x_{i,\ell}$ between them and $\ell = 2j+1$. Consider now $\delta_{k,i+n}(x,t)$ for positive n. Then the points $x_{i,\ell} = x_{i+n,m}$ with $m = 2^n \ell$ is between t and x, yet m is even. Hence, due to part (ii), $\delta_{k,i+n}(x,t)$ cannot be equal to one. This completes the proof. □

4. The algorithm. Let $\{A_{k,i}\}$ be the families of algorithms from the previous section, each for $\omega = \rho_k$ and $(a,b) = (a_k, b_k)$, respectively. Recall that $A_{k,i}$ uses function values at points $x^k_{i,1}, \ldots, x^k_{i,2^i-1}$. Define

$$\Delta_{k,1} := A_{k,1}, \qquad \Delta_{k,i} := A_{k,i} - A_{k,i-1} \quad \text{for} \quad i \geq 2$$

and

$$\mathcal{A}_{q,s} := \sum_{|\mathbf{i}| \leq q} \bigotimes_{k=1}^{s} \Delta_{k,i_k} \tag{9}$$

for $q \geq s$. Here and elsewhere, $\mathbf{i} = [i_1, \ldots, i_s] \in \mathbb{N}^s_+$ is a multi-index with $i_k \geq 1$ and $|\mathbf{i}| = \sum_{k=1}^s i_k$.

THEOREM 2. *Let* $s \geq 2$ *and* $q \geq s$. *Then*

$$\operatorname{error}(\mathcal{A}_{q,s}) \leq \begin{cases} s\, 2^{-q} \binom{q}{\lfloor \frac{q+1}{2} \rfloor} & \text{if } q < 2(s-1), \\ s\, 2^{-q} \binom{q}{s-1} & \text{if } q \geq 2(s-1). \end{cases}$$

If, additionally, (4) *holds, then*

$$\operatorname{error}(\mathcal{A}_{q,s}) \leq s\, 2^{-q+s-1-a} \binom{q-s}{a} \quad \text{with} \quad a = \begin{cases} \lfloor \frac{q-s+1}{3} \rfloor & \text{if } q < 4(s-1), \\ s-1 & \text{if } q \geq 4(s-1). \end{cases}$$

To prove this theorem, we need the following lemma. We think it is known; however, we have not found it in the literature. For $q \geq s-1$, define

$$\mathrm{B}(q,s) := \sum_{|\mathbf{i}| \geq q+1} 2^{-|\mathbf{i}|}. \tag{10}$$

LEMMA 2. *For every* $q \geq s-1$,

$$\mathrm{B}(q,s) = 2^{-q} \sum_{j=0}^{s-1} \binom{q}{j} \leq \overline{\mathrm{B}}(q,s), \tag{11}$$

where

$$\overline{\mathrm{B}}(q,s) := s\,2^{-q} \begin{cases} \binom{q}{\lfloor \frac{q+1}{2} \rfloor} & \textit{if } 2s \ge q+3, \\ \binom{q}{s-1} & \textit{otherwise.} \end{cases} \tag{12}$$

Proof of Lemma 2. Indeed,

$$\mathrm{B}(q,s) = \sum_{j=q+1}^{\infty} 2^{-j} \binom{j-1}{s-1}$$

and

$$\begin{aligned} 2^{-j}\binom{j-1}{s-1} &= 2^{-s} \cdot 2^{-(j-s)} \cdot \frac{(j-1)\cdots(j-s+1)}{(s-1)!} \\ &= \frac{x^s}{(s-1)!} \cdot \left(x^{j-1}\right)^{(s-1)} \Big|_{x=1/2}. \end{aligned}$$

Taking the summation with respect to j inside the differentiation, we get

$$\begin{aligned} \mathrm{B}(q,s) &= \frac{x^s}{(s-1)!} \left(\sum_{j=q+1}^{\infty} x^{j-1} \right)^{(s-1)} \Bigg|_{x=1/2} = \frac{x^s}{(s-1)!} \left(\frac{x^q}{1-x} \right)^{(s-1)} \Bigg|_{x=1/2} \\ &= \frac{x^s}{(s-1)!} \sum_{j=0}^{s-1} \binom{s-1}{j} (x^q)^{(j)} \left((1-x)^{-1}\right)^{(s-1-j)} \Bigg|_{x=1/2} \\ &= x^s \sum_{j=0}^{s-1} \frac{1}{j!(s-1-j)!} \cdot x^{q-j} \cdot q \cdots (q-j+1) \cdot \frac{(s-1-j)!}{(1-x)^{s-j}} \Bigg|_{x=1/2}, \end{aligned}$$

which, after some elementary manipulation, can be shown to be equal to $2^{-q} \sum_{j=0}^{s-1} \binom{q}{j}$. This completes the proof of (11). The upper bound $\overline{\mathrm{B}}(q,s)$ follows from this and the well-known fact that $\binom{q}{j-1} \le \binom{q}{j}$ iff $2j \le q+1$. This completes the proof of Lemma 2. □

Proof of Theorem 2. Note that for any scalar function $f \in \mathcal{H}_k$, $A_{k,n}(f)$ converges pointwise to f with $n \to \infty$ and that $\sum_{i=1}^{n} \Delta_{k,i} = A_{k,n}$. Therefore,

$$f(\mathbf{x}) = \sum_{\mathbf{i} \in \mathbb{N}_+^s} \bigotimes_{k=1}^{s} \Delta_{k,i_k}(f)(\mathbf{x})$$

for any $\mathbf{x}$ and any $f \in \mathcal{F}_{1,s}$. Here and elsewhere, by $\sum_{\mathbf{i} \in \mathbb{N}^s}$ we mean a double sum $\sum_{\ell=s}^{\infty} \sum_{|\mathbf{i}|=\ell}$. Hence

$$f - \mathcal{A}_{q,s}(f) = \mathcal{E}_{q,s}(f) \quad \text{with} \quad \mathcal{E}_{q,s}(f) := \sum_{|\mathbf{i}| \ge q+1} \bigotimes_{k=1}^{s} \Delta_{k,i_k}(f).$$

Due to (5) and the fact that $\mathcal{E}_{q,s}$ vanishes on constant functions,

$$\mathcal{E}_{q,s}(f) = \sum_{U \ne \emptyset} \int_{D_U} f_U'(\mathbf{t}_U)\, \mathcal{E}_{q,s}(M_U(\cdot, \mathbf{t}_U))\, d\mathbf{t}_U,$$

which yields

$$\|(f-\mathcal{A}_{q,s}(f))\,\rho\|_{L_1} \le \|f\|_{1,s}\cdot \sup_{\mathbf{t}\in D}\max_{U\neq\emptyset}\|\rho\,\mathcal{E}_{q,s}(M_U(\cdot,\mathbf{t}_U))\|_{L_1}.$$

Hence, to complete the proof, we only need to estimate the above supremum. For that purpose, note that

$$\bigotimes_{k=1}^{s}\Delta_{k,i_k}(M_U(\cdot,\mathbf{t}_U))\equiv 0 \quad \text{if } i_k\ge 2 \text{ for some } k\notin U.$$

This follows from the fact that $\Delta_{k,i}(1)\equiv 0$ for any $i\ge 2$, which, in turn, is a consequence of the fact that $A_{k,i}(1)\equiv 1$ for any $i\ge 1$. Otherwise, i.e., when $i_k=1$ for all $k\notin U$,

$$\left\|\rho\bigotimes_{k=1}^{s}\Delta_{k,i_k}(M_U(\cdot,\mathbf{t}_U))\right\|_{L_1}=\prod_{k\in U}\|\rho_k\,\Delta_{k,i_k}(M_k(\cdot,t_k))\|_{L_1}\le\prod_{k\in U}2^{-i_k},$$

independently of $\mathbf{t}$, due to Proposition 1. Therefore

$$\|\rho\,\mathcal{E}_{q,s}(M_U(\cdot,\mathbf{t}_U))\|_{L_1}\le\sum_{\mathbf{i}}2^{-|\mathbf{i}|}\le \mathrm{B}(q-s+|U|,|U|)\le\overline{\mathrm{B}}(q-s+|U|,|U|) \tag{13}$$

with the summation in (13) over $\mathbf{i}\in\mathbb{N}^{|U|}$ such that $|\mathbf{i}|\ge q+1-(s-|U|)$. The first result in the theorem is proved by showing that $\overline{\mathrm{B}}(q-s+|U|,|U|)$ is increasing with $|U|$, i.e., $\max_U\overline{\mathrm{B}}(q-s+|U|,|U|)=\overline{\mathrm{B}}(q,s)$.

Consider first $q\ge 2(s-1)$. To see that $\overline{\mathrm{B}}(q-s+|U|,|U|)$ is increasing with $|U|$, note that

$$\begin{aligned}\frac{\overline{\mathrm{B}}(q-s+|U|+1,|U|+1)}{\overline{\mathrm{B}}(q-s+|U|,|U|)}&=\frac{(|U|+1)(q-s+|U|+1)}{2|U|^2}\\ &\ge\frac{q-s+|U|+1}{2|U|}\ge 1\end{aligned}$$

with the last inequality due to the fact that $|U|+1\le s$ and hence $|U|\le s-1\le q-s+1$.

Suppose now $q<2(s-1)$. To show that $\overline{\mathrm{B}}(q-s+|U|,|U|)$ increases with $|U|$ also in this case, we need to consider three different cases. Consider first $|U|>q-s+3$ (i.e., $2|U|>q-s+|U|+3$). Then

$$\begin{aligned}\frac{\overline{\mathrm{B}}(q-s+|U|+1,|U|+1)}{\overline{\mathrm{B}}(q-s+|U|,|U|)}&=\frac{|U|+1}{2|U|}\binom{q-s+|U|+1}{\lfloor\frac{q-s+|U|+2}{2}\rfloor}\Big/\binom{q-s+|U|}{\lfloor\frac{q-s+|U|+1}{2}\rfloor}\\ &=\frac{|U|+1}{|U|}\,\frac{2\ell+1}{2\ell+2}>1\end{aligned}$$

with the last equality due to an extra assumption that $q-s+|U|=2\ell$ (the proof for odd $q-s+|U|$ is very similar).

Consider next the case of $|U|=q-s+2$. Then

$$\frac{\overline{\mathrm{B}}(q-s+|U|+1,|U|+1)}{\overline{\mathrm{B}}(q-s+|U|,|U|)}=\frac{(q-s+3)(2q-2s+3)}{(q-s+2)(2q-2s+4)}\ge 1.$$

Consider now $|U|+1 < q-s+3$. Then

$$\frac{\overline{\mathrm{B}}(q-s+|U|+1,|U|+1)}{\overline{\mathrm{B}}(q-s+|U|,|U|)} = \frac{|U|+1}{2|U|}\binom{q-s+|U|+1}{|U|}\bigg/\binom{q-s+|U|}{|U|-1}$$
$$= \frac{(|U|+1)(q-s+|U|+1)}{2|U|^2} \geq 1.$$

This completes the proof of the first part of the theorem.

Assume now that (4) holds. Then

$$\bigotimes_{k=1}^{s} \Delta_{k,i_k}(M_U(\cdot,\mathbf{t}_U)) \equiv 0 \quad \text{if } i_k = 1 \text{ for some } k \in U$$

as follows directly from the second part of Proposition 1. This means that the sum in (13) is now over $\mathbf{i} \in \mathbb{N}^{|U|}$ such that $|\mathbf{i}| \geq \max\{q+1-s+|U|, 2|U|\}$ and $\mathbf{i} \geq \mathbf{2}$. Replacing $\mathbf{i}$ by $\mathbf{j} = \mathbf{i} - \mathbf{1}$, the sum becomes

$$(14) \qquad 2^{-|U|} \sum_{|\mathbf{j}| \geq \max\{q+1-s,|U|\}} 2^{-|\mathbf{j}|} = 2^{-|U|} \cdot \mathrm{B}(\max\{q-s,|U|-1\},|U|).$$

Therefore, for $|U| < q-s+1$ we have

$$(15) \qquad \|\rho\,\mathcal{E}_{q,s}(M_U(\cdot,\mathbf{t}_U))\|_{L_1} \leq 2^{-q-|U|+s} \sum_{j=0}^{|U|-1} \binom{q-s}{j}$$

and for $|U| \geq q-s+1$ we have

$$(16) \qquad \|\rho\,\mathcal{E}_{q,s}(M_U(\cdot,\mathbf{t}_U))\|_{L_1} \leq 2^{-|U|} \leq 2^{-q+s-1}.$$

Since $|U| \leq s$, we have to have $q \leq 2s-1$ for the latter case to happen. To estimate the maximum of the upper bound with respect to $|U|$, we consider the following cases.

Case $q \geq 3(s-1)$. Then, because $q-s \geq 2|U|-3$, the right-hand side of (15) is bounded from above by $s\,2^{-q-|U|+s}\binom{q-s}{|U|-1}$, which can be shown to increase with $|U|$ as long as $|U| \leq 1 + \lfloor (q-s+1)/3 \rfloor$. Since $|U| \leq s$, this yields the upper bounds

$$\|\rho\,\mathcal{E}_{q,s}(M_U(\cdot,\mathbf{t}_U))\|_{L_1} \leq s\,2^{-q}\binom{q-s}{s-1},$$

if $q \geq 4(s-1)$, and

$$\|\rho\,\mathcal{E}_{q,s}(M_U(\cdot,\mathbf{t}_U))\|_{L_1} \leq s\,2^{-q+s-1-\lfloor\frac{q-s+1}{3}\rfloor}\binom{q-s}{\lfloor (q-s+1)/3 \rfloor},$$

if $q/(s-1) \in [3,4)$.

Case $2(s-1) < q < 3(s-1)$. First note that the right-hand side of (15) decreases with $|U|$ when $|U| > (q-s)/2$. This follows from the fact that the value of the right-hand side for $|U|$ minus the value for $|U|+1$ equals

$$2^{-q-|U|-1+s}\left(\sum_{j=0}^{|U|-1}\binom{q-s}{j} - \binom{q-s}{|U|}\right) \geq 0$$

as claimed. Moreover, for $|U| \le (q-s)/2$, the right-hand side of (15) is bounded by $s\,2^{-q-|U|+s}\binom{q-s}{|U|-1}$, which attains its maximum for $|U| = 1 + \lfloor (q-s+1)/3 \rfloor$. Hence again,

$$\|\rho\,\mathcal{E}_{q,s}(M_U(\cdot,\mathbf{t}_U))\|_{L_1} \le s\,2^{-q+s-1-\lfloor\frac{q-s+1}{3}\rfloor}\binom{q-s}{\lfloor (q-s+1)/3\rfloor}.$$

Case $q \le 2(s-1)$. Due to (16), we need only to estimate the right-hand side of (15) for $|U| < q-s+1$. However, as in the previous case, it is bounded by $s\,2^{-q-|U^*|+s}\binom{q-s}{|U^*|-1}$ with $|U^*| = 1 + \lfloor (q-s+1)/3 \rfloor$. Since

$$s\,2^{-q+s-1-\lfloor\frac{q-s+1}{3}\rfloor}\binom{q-s}{\lfloor (q-s+1)/3\rfloor} \ge 2^{-q+s-1},$$

the left-hand side of the above inequality is an upper bound on $\|\rho\,\mathcal{E}_{q,s}(M_U(\cdot_U,\mathbf{t}_U))\|_{L_1}$ also in this case. This completes the proof. □

We end this section by relating the error of $\mathcal{A}_{q,s}$ to its cardinality $\mathrm{card}(\mathcal{A}_{q,s})$.

THEOREM 3. *For every* $s \ge 2$ *and* $q \ge 2(s-1)$,

$$\text{(17)} \qquad \mathrm{error}(\mathcal{A}_{q,s}) \le \frac{s\,\exp\left(\frac{1}{12(s-1)}\right)}{(s-1)\pi}\left(\frac{e\,\ln(\mathrm{card}(\mathcal{A}_{q,s}))}{(s-1)\sqrt{2}\ln(2)}\right)^{2(s-1)}\frac{1}{\mathrm{card}(\mathcal{A}_{q,s})}.$$

Proof. Since the information used by $\mathcal{A}_{q,s}$ is nested, we know from [25, Lem. 7] that

$$\text{(18)} \qquad 2^{q-s}\binom{q-1}{s-1} \le \mathrm{card}(\mathcal{A}_{q,s}) \le 2^{q-s+1}\binom{q-1}{s-1}.$$

Let's represent q as $q = (t+1)(s-1)$ with $t \ge 1$. We apply Stirling's formula ($n! = (n/e)^n\sqrt{2\pi n}\exp(\theta/(12n))$ for $\theta \in (0,\pi)$) to the error bound from Theorem 2; one can show that

$$\begin{aligned}\mathrm{error}(\mathcal{A}_{q,s}) &\le s\,2^{-(t+1)(s-1)}\binom{(t+1)(s-1)}{s-1}\\ &\le s\,2^{-(t+1)(s-1)}\,((t+1)e)^{s-1}\sqrt{\frac{t+1}{t2\pi(s-1)}}\,\exp\left(\frac{1}{12q}\right)\\ &\le s\,2^{-(t+1)(s-1)}\,((t+1)e)^{s-1}\sqrt{\frac{\exp\left(\frac{1}{12(s-1)}\right)}{\pi(s-1)}},\end{aligned}$$

with the last inequality due to the fact that $t \ge 1$. Similarly,

$$\mathrm{card}(\mathcal{A}_{q,s}) \le 2^{t(s-1)}\,(e(t+1))^{s-1}\sqrt{\frac{\exp\left(\frac{1}{12(s-1)}\right)}{\pi(s-1)}}.$$

Since $x/(\ln(x))^{2(s-1)}$ increases for $x \ge e^{2(s-1)}$, and since $\mathrm{card}(\mathcal{A}_{q,s}) \ge e^{2(s-1)}$ due to (18), we can replace $\mathrm{card}(\mathcal{A}_{q,s})$ by the right-hand side of the above inequality in the following estimation:

$$\begin{aligned}L &:= \frac{\mathrm{card}(\mathcal{A}_{q,s})\ \mathrm{error}(\mathcal{A}_{q,s})}{(\ln(\mathrm{card}(\mathcal{A}_{q,s})))^{2(s-1)}}\\ &\le \frac{s\,\exp\left(\frac{1}{12(s-1)}\right)}{\pi(s-1)}\,2^{-(s-1)}\left(\frac{e}{s-1}\,g(t)\right)^{2(s-1)},\end{aligned}$$

with

$$g(t) := \frac{t+1}{t\ln(2) + \ln(t+1)}$$

since $s-1-\ln(\pi(s-1))/2$ is positive. It is easy to verify that $\max_{t\geq 1} g(t) = 1/\ln(2)$. This yields

$$\begin{aligned} L &\leq \frac{s\,\exp\left(\frac{1}{12(s-1)}\right)}{(s-1)\pi}\, 2^{-(s-1)} \left(\frac{e}{(s-1)\ln(2)}\right)^{2(s-1)} \\ &= \frac{s\,\exp\left(\frac{1}{12(s-1)}\right)}{(s-1)\pi} \left(\frac{e}{(s-1)\sqrt{2}\ln(2)}\right)^{2(s-1)}, \end{aligned}$$

which completes the proof. □

5. Extensions. In this section, we extend some of the previous results assuming now that the functions f are from the space $\mathcal{F} = \mathcal{F}_{p,s,\gamma}$, which is the completion of $\mathcal{H}^s$ with respect to the norm

$$\|f\|_{p,s,\gamma} := \left(|f(\mathbf{c})|^p + \sum_{U\neq\emptyset} \gamma_{s,U}^{-p}\, \|f_U'\|_{L_p}^p\right)^{1/p} = \left(\sum_U \gamma_{s,u}^{-p} \|f_U'\|_{L_p}^p\right)^{1/p},$$

where $p \in [1,\infty]$ and $\gamma = \{\gamma_{s,u}\}_{s,U}$ is a family of nonnegative numbers, called weights. By a convention $0/0 = 0$. Of course, for $p = \infty$ we have $\|f\|_{\infty,s} = \max_U \gamma_{s,U}^{-1} \|f_U'\|_{L_\infty}$. This norm differs from $\|\cdot\|_{1,s}$ by using L_p instead of L_1-norms and by adding the weights $\gamma_{s,U}$.

The role of $\gamma_{s,U}$ is that they model how important certain variables and their groups are. For instance, the condition of small enough $\gamma_{s,U}$ means that $\|f_U'\|_{L_p}$ cannot be too large, and $\gamma_{s,U} = 0$ implies that $\|f_U'\|_{L_p} = 0$. Since the introduction of weighted norms in [20], high dimensional problems with such norms have been investigated in a number of papers; see, e.g., [15] and papers cited there. It is often the case that if the weights are small enough, then the error in solving a high dimensional problem is essentially no worse than that for solving a one-dimensional problem.

To stress that now a different space from $\mathcal{F}_{1,s}$ is being considered, we will write error$(\mathcal{A};\mathcal{F}_{p,s,\gamma})$ instead error$(\mathcal{A})$. When all weights equal 1, we will simply write $\mathcal{F}_{p,s}$. For simplicity, we assume throughout this section that (4) is satisfied and that $\mathbf{c} = \mathbf{0}$. Hence, in particular $a_k < 0 < b_k$.

Note that for $p > 1$ and unbounded D it could happen that the approximation problem is not well defined since the corresponding embedding operator could be unbounded. As follows from [27, Thm. 1], the problem is well defined iff

$$\left(\int_{a_k}^{b_k} \psi_k^{p^*}(x)\,dx\right)^{1/p^*} < \infty \quad \forall k, \tag{19}$$

with

$$\psi_k(t) = \begin{cases} \int_t^{b_k} \rho_k(x)\,dx & \text{for } t \geq 0, \\ \int_{a_k}^t \rho_k(x)\,dx & \text{otherwise}, \end{cases} \tag{20}$$

where here and elsewhere p^* denotes the conjugate to p, i.e.,

$$\frac{1}{p} + \frac{1}{p^*} = 1.$$

This is why we assume (19) throughout the rest of this paper. Of course, (19) trivially holds when D is bounded. It also holds for $p = 1$ since then $p^* = \infty$ and the left-hand side of (19) should formally be replaced by $\operatorname{ess\,sup}_{t \geq c_k} \int_t^{b_k} \rho_k(x)\,dx$, which obviously is equal to 1/2.

THEOREM 4. *The error of any algorithm $\mathcal{A}$ is bounded by*

$$\text{error}(\mathcal{A}; \mathcal{F}_{p,s,\gamma}) \leq \left(\sum_U \gamma_{s,U}^{p^*} \int_{D_U} \left(\int_{D_U} h_U(\mathbf{x}_U, \mathbf{t}_U)\, d\mathbf{x}_U \right)^{p^*} d\mathbf{t}_U \right)^{1/p^*}, \tag{21}$$

where

$$h_U(\mathbf{x}_U, \mathbf{t}_U) = \rho_U(\mathbf{x}_U)\, |M_U(\mathbf{x}_U, \mathbf{t}_U) - \mathcal{A}(M_U(\cdot, \mathbf{t}_U))(\mathbf{x}_U)|.$$

If $\mathcal{A}$ is based on piecewise constant interpolation, then we have equality in (21).

Proof. To simplify the notation, we write $m(\mathbf{x}_U, \mathbf{t}_U)$ to denote

$$m(\mathbf{x}_U, \mathbf{t}_U) := M(\mathbf{x}_U, \mathbf{t}_U) - \mathcal{A}(M(\cdot, \mathbf{t}_U))(\mathbf{x}_U).$$

Of course, $h_U(\mathbf{x}_U, \mathbf{t}_U) = \rho_U(\mathbf{x}_U)\, |m_U(\mathbf{x}_U, \mathbf{t}_U)|$.

Suppose that $p > 1$. We begin with the proof of (21). Using (5), we have by Hölder's inequality

$$\begin{aligned}
&\int_D \rho(\mathbf{x})\, |f(\mathbf{x}) - \mathcal{A}(f)(\mathbf{x})|\, d\mathbf{x} \\
&= \int_D \rho(\mathbf{x}) \left| \sum_U \int_{D_U} f'_U(\mathbf{t}_U)\, m_U(\mathbf{x}_U, \mathbf{t}_U)\, d\mathbf{t}_U \right| d\mathbf{x}_U \\
&\leq \sum_U \int_{D_U} \rho_U(\mathbf{x}_U) \int_{D_U} |f'_U(\mathbf{t}_U)\, m_U(\mathbf{x}_U, \mathbf{t}_U)|\, d\mathbf{t}_U\, d\mathbf{x}_U \\
&= \sum_U \int_{D_U} |f'_U(\mathbf{t}_U)| \int_{D_U} h_U(\mathbf{x}_U, \mathbf{t}_U)\, d\mathbf{x}_U\, d\mathbf{t}_U \\
&\leq \sum_U \|f'_U\|_{L_p} \gamma_{s,U}^{-1} \left(\gamma_{s,U}^{p^*} \int_{D_U} \left(\int_{D_U} h_U(\mathbf{x}_U, \mathbf{t}_U)\, d\mathbf{x}_U \right)^{p^*} d\mathbf{t}_U \right)^{1/p^*} \\
&\leq \|f\|_{p,s,\gamma} \left(\sum_U \gamma_{s,U}^{p^*} \int_{D_U} \left(\int_{D_U} h_U(\mathbf{x}_U, \mathbf{t}_U)\, d\mathbf{x}_U \right)^{p^*} d\mathbf{t}_U \right)^{1/p^*}.
\end{aligned}$$

This proves (21) for $p > 1$. We now show equality when $\mathcal{A}(m_U(\cdot, \mathbf{t}_U))$ is a piecewise constant function interpolating $m_U(\cdot, t_U)$. To that end, define

$$g_U(\mathbf{t}_U) := \text{sign}(\mathbf{t}_U) \left(\int_{D_U} h(\mathbf{x}_U, \mathbf{t}_U)\, d\mathbf{x}_U \right)^{p^*-1} \quad \text{with} \quad \text{sign}(\mathbf{t}_U) := \prod_{k \in U} \text{sign}(t_k),$$

and

$$\widetilde{f}(\mathbf{y}) := \sum_U \gamma_{s,U}^{p^*} \int_{D_U} g_U(\mathbf{t}_U)\, M_U(\mathbf{y}_U, \mathbf{t}_U)\, d\mathbf{t}_U.$$

Of course, from (5), $\widetilde{f} \in \mathcal{F}_{p,s,\gamma}$ and $\widetilde{f}'_U = \gamma_{s,U}^{p^*} g_U$. Since $pp^* - p = p^*$, it is easy to check that

$$\|\widetilde{f}\|_{p,s,\gamma} = \left(\sum_U \gamma_{s,U}^{p^*} \int_{D_U} \left(\int_{D_U} h_U(\mathbf{x}_U, \mathbf{t}_U)\, d\mathbf{x}_U \right)^{p^*} d\mathbf{t}_U \right)^{1/p}.$$

Moreover,

$$\begin{aligned}
\|(\widetilde{f} - \mathcal{A}(\widetilde{f}))\,\rho\|_{L_1} &= \int_D \rho(\mathbf{y})\, |\widetilde{f}(\mathbf{y}) - \mathcal{A}(\widetilde{f})(\mathbf{y})|\, d\mathbf{y} \\
&= \int_D \rho(\mathbf{y}) \left| \sum_U \gamma_{s,U}^{p^*} \int_{D_U} g_U(\mathbf{t}_U)\, m(\mathbf{y}_U, \mathbf{t}_U)\, d\mathbf{t}_U \right| d\mathbf{y} \\
&= \int_D \rho(\mathbf{y}) \left| \sum_U \gamma_{s,U}^{p^*} \int_{D_U} \left(\int_{D_U} h(\mathbf{x}_U, \mathbf{t}_U)\, d\mathbf{x}_U \right)^{p^*-1} \operatorname{sign}(\mathbf{t}_U)\, m(\mathbf{y}_U, \mathbf{t}_U)\, d\mathbf{t}_U \right| d\mathbf{y} \\
&= \sum_U \gamma_{s,U}^{p^*} \int_{D_U} \left(\int_{D_U} h(\mathbf{x}_U, \mathbf{t}_U)\, d\mathbf{x}_U \right)^{p^*-1} \int_{D_U} h(\mathbf{y}_U, \mathbf{t}_U)\, d\mathbf{y}_U\, d\mathbf{t}_U \\
&= \sum_U \gamma_{s,U}^{p^*} \int_{D_U} \left(\int_{D_U} h_U(\mathbf{x}_U, \mathbf{t}_U)\, d\mathbf{x}_U \right)^{p^*} d\mathbf{t}_U = \|\widetilde{f}\|_{p,s,\gamma}^p,
\end{aligned}$$

with the third-to-last equality due to the fact that

$$\operatorname{sign}(\mathbf{t}_U)\, m(\mathbf{y}_U, \mathbf{t}_U) = |m(\mathbf{y}_U, \mathbf{t}_U)| \ge 0 \quad \forall \mathbf{y}, \mathbf{t}.$$

Dividing $\|(\widetilde{f} - \mathcal{A}(\widetilde{f}))\,\rho\|_{L_1}$ by $\|\widetilde{f}\|_{p,s,\gamma}$ to obtain $\|\widetilde{f}\|_{p,s,\gamma}^{p-1} = \|\widetilde{f}\|_{p,s,\gamma}^{p/p^*}$ on the right completes the proof of the equality for $p > 1$.

For $p = 1$ the proof technique is the same with obvious modifications. □

Let $\mathcal{A}_{q,s}$ be the algorithm from section 4 and denote

$$\delta_{U,\mathbf{i}_U}(\mathbf{x}_U, \mathbf{t}_U) := \Delta_{U,\mathbf{i}_U}(M_U(\cdot, \mathbf{t}_U))(\mathbf{x}_U) = \prod_{k \in U} \delta_{k,i_k}(x_{i_k}, t_{i_k}).$$

From Lemma 1, we have the following proposition.

PROPOSITION 2. *Let $U \neq \emptyset$ and $\mathbf{t}_U \in D_U$. Then the following hold.*

(i) *For every $\mathbf{x}_U \in D_U$, $\delta_{U,\mathbf{i}_U}(\mathbf{x}_U, \mathbf{t}_U) \in \{-1, 0, 1\}$.*

(ii) *For every $\mathbf{x} \in D_U$ there exists at most one $\mathbf{i}_U$ such that $|\delta_{U,\mathbf{i}_U}(\mathbf{x}_U, \mathbf{t}_U)| = 1$.*

(iii) *For every $\mathbf{i}$, the ρ_U-probability of the set of $\mathbf{x}_U$'s with $|\delta_{U,\mathbf{i}_U}(\mathbf{x}_U, \mathbf{t}_U)| = 1$ is at most $2^{-|\mathbf{i}_U|}$.*

(iv) *If $\delta_{U,\mathbf{i}_U}(\mathbf{x}_U, \mathbf{t}_U) \neq 0$ for some $\mathbf{x}$ and $\mathbf{i}$, then $t_k \in [x_{i_k,1}^k, x_{i_k,2^{i_k}-1}^k]$ for every $k \in U$.*

As in the proof of Theorem 2, let $\mathcal{E}_{q,s}(f) = \sum_{|\mathbf{i}| \ge q+1} \bigotimes_{k=1}^s \Delta_{k,i_k}(f)$. Recall that $\Delta_{\mathbf{i}}(M_U(\cdot, \mathbf{t}_U)) \equiv 0$ if either $i_k = 1$ for some $k \in U$, or $i_k \ge 2$ for some $k \notin U$. Hence

$$\mathcal{E}_{q,s}(M_U(\cdot, \mathbf{t}_U))(\mathbf{x}_U) = \sum_{\mathbf{i}_U \in P(q,s,U)} \delta_{U,\mathbf{i}_U}(\mathbf{x}_U, \mathbf{t}_U)$$

with

$$P(q, s, U) = \left\{ \mathbf{j} \in \mathbb{N}_+^{|U|} \,:\, \mathbf{j} \ge \mathbf{2}, |\mathbf{j}| \ge q + 1 - s + |U| \right\}.$$

From Theorem 4, this yields

$$\mathrm{error}(\mathcal{A}_{q,s},\mathcal{F}) = \left(\sum_{U\neq\emptyset} \gamma_{s,U}^{p^*} \int_{D_U} b_U^{p^*}(\mathbf{t}_U)\, d\mathbf{t}_U\right)^{1/p^*},$$

where

$$(22)\qquad b_U(\mathbf{t}_U) := \left|\int_{D_U} \rho_U(\mathbf{x}_U) \sum_{\mathbf{i}_U\in P(q,s,U)} \delta_{U,\mathbf{i}_U}(\mathbf{x}_U,\mathbf{t}_U)\, d\mathbf{x}_U\right|.$$

Due to its definition (20), $\psi_k(t)$ converges to zero with $t \to b_k$ and/or $t \to a_k$. Moreover, from (iv) of Proposition 2, we know that

$$\delta_{U,\mathbf{i}_U}(\mathbf{x}_U,\mathbf{t}_U) \neq 0 \quad \text{implies } \psi_k(t_k) \geq 2^{-i_k} \text{ for every } k \in U.$$

This means that we can replace the set $P(q,s,U)$ by the even smaller set

$$P(q,s,U,\mathbf{t}_U) = \left\{\mathbf{j} \in \mathbb{N}_+^{|U|} \,:\, \mathbf{j} \geq \mathbf{2},\ |\mathbf{j}| \geq q+1-s+|U|, \text{ and } 2^{-j_k} \leq \psi_k(t_k),\ \forall\, k \in U\right\}.$$

This leads to

$$\begin{aligned} b_U(\mathbf{t}_U) &\leq \sum_{\mathbf{j}\in P(q,s,U,\mathbf{t}_U)} 2^{-\mathbf{j}} \\ &\leq \min\left\{2^{|U|}\,\psi_U(\mathbf{t}_U)\,,\ 2^{-|U|}\ \mathrm{B}(\max\{q-s,|U|-1\},|U|)\right\}, \end{aligned}$$

where, as always, $\psi_U(\mathbf{t}_U) = \prod_{k\in U} \psi_k(t_k)$. We summarize this in the following proposition.

PROPOSITION 3. *Let* (4) *hold. Then for any $s \geq 2$ and any $q \geq s$,*

$$\mathrm{error}(\mathcal{A}_{q,s,\gamma};\mathcal{F}_{p,s,\gamma}) = \left(\sum_{U\neq\emptyset} \gamma_{s,U}^{p^*} \int_{D_U} b_U^{p^*}(\mathbf{t}_U)\, d\mathbf{t}_U\right)^{1/p^*}.$$

Moreover,

$$b_U(\mathbf{t}_U) \leq \min\left\{2^{|U|}\,\psi_U(\mathbf{t}_U)\,,\ \frac{\mathrm{B}(\max\{q-s,|U|-1\},|U|)}{2^{|U|}}\right\}.$$

For $p > 1$ and unbounded D, the above error bound is quite complicated, due to the presence of integrals of $b_U^{p^*}$. Suppose now that D is bounded, say

$$D = [0,1]^s.$$

Then the integrals of b_U^* can be replaced by $\mathrm{B}(\max\{q-s,|U|-1\},|U|)\,2^{-|U|}$ leading to the following upper bound:

$$(23)\quad \mathrm{error}(\mathcal{A}_{q,s};\mathcal{F}_{p,s,\gamma}) \leq \left(\sum_{U\neq\emptyset}\left(2^{-|U|}\,\gamma_{s,U}\ \mathrm{B}(\max\{q-s,|U|-1\},|U|)\right)^{p^*}\right)^{1/p^*}.$$

Of course, if $p = 1$, then $p^* = \infty$ and the bound (23) takes the form

$$\mathrm{error}(\mathcal{A}_{q,s};\mathcal{F}_{1,s}) \le \max_{U\neq\emptyset} 2^{-|U|}\,\gamma_{s,U}\ \mathrm{B}(\max\{q-s,|U|-1\},|U|),$$

which coincides with the bound from section 4 for the case when (3) holds. For $p > 1$, (23) can be further estimated from above, leading to

$$\mathrm{error}(\mathcal{A}_{q,s};\mathcal{F}_{p,s,\gamma}) \le 2^{s/p^*}\ \max_{U\neq\emptyset} 2^{-|U|}\,\gamma_{s,U}\ \mathrm{B}(\max\{q-s,|U|-1\},|U|).$$

This means that error upper bounds from section 4 also hold for $p > 1$ modulo the multiplicative factor $2^{s/p^*}$. In particular, we have the following consequence of Theorem 3.

THEOREM 5. *Let $p > 1$, $\gamma_{s,U} \equiv 1$, and $D = [0,1]^s$. Then for every $s \ge 2$ and $q \ge 2(s-1)$,*

$$\mathrm{error}(\mathcal{A}_{q,s};\mathcal{F}_{p,s}) \le \frac{s\,\exp\left(\frac{1}{12(s-1)}\right), 2^{1/p^*}}{(s-1)\pi}\left(\frac{e\,\ln(\mathrm{card}(\mathcal{A}_{q,s}))}{(s-1)\,2^{1/(2p^*)}\,\ln(2)}\right)^{2(s-1)}\frac{1}{\mathrm{card}(\mathcal{A}_{q,s})}.$$

6. Integration problem. In this section, we briefly discuss the problem of approximating weighted integrals

$$I_\rho(f) = \int_D f(\mathbf{x})\,\rho(\mathbf{x})\,d\mathbf{x}$$

by algorithms (often called quadratures) $\mathcal{Q}$ of the form $\mathcal{Q}(f) = \sum_{i=1}^n f(\mathbf{x}^i)\,g_i$. The *worst case error* of $\mathcal{Q}$ (with respect to $\mathcal{F}_{p,s}$) is defined by

$$\mathrm{error}(\mathcal{Q};\mathcal{F}_{p,s},\mathrm{Int}) := \sup_{\|f\|_{p,s}\le 1} |I_\rho(f) - \mathcal{Q}(f)|.$$

Consider now the following quadrature

$$\mathcal{Q}_{q,s}(f) := I_\rho(\mathcal{A}_{q,s}(f)). \tag{24}$$

It is easy to see that

$$\mathcal{Q}_{q,s} = \sum_{|\mathbf{i}|\le q}\bigotimes_{k=1}^{s}(Q_{k,i_k} - Q_{k,i_k-1}), \quad \text{where} \quad Q_{k,i}(f) = 2^{-i}\left(f(c_k) + \sum_{j=1}^{2^i-1} f(x_{i,j}^k)\right)$$

$x_{i,j}^k$ defined by (7). Clearly,

$$\mathrm{error}(\mathcal{Q}_{q,s};\mathcal{F}_{p,s},\mathrm{Int}) \le \mathrm{error}(\mathcal{A}_{q,s};\mathcal{F}_{p,s}).$$

Hence all error bounds for $\mathcal{A}_{q,s}$ obtained in the previous sections also hold for $\mathcal{Q}_{q,s}$.

REFERENCES

[1] K.I. BABENKO, *Approximation by trigonometric polynomials in a certain class of periodic functions of several variables*, Soviet. Math. Dokl., 1 (1960), pp. 672–675.

[2] G. BASZENSKI, F.J. DELVOS, AND S. JESTER, *Blending approximations with sine functions*, in Numerical Methods in Approximation Theory, Vol. 9, D. Braess and L.L. Schumaker, eds., Internat. Ser. Numer. Math. 105, Birkhäuser, Basel, 1992, pp. 1–19.

[3] F.J. Delvos, *d-variate Boolean approximation*, J. Approx. Theory, 34 (1982), pp. 99–114.
[4] R.A. DeVore, S.V. Konyagin, and V.N. Temlyakov, *Hyperbolic wavelet approximation*, Constr. Approx., 14 (1998), pp. 1–26.
[5] M. Drmota and R.F. Tichy, *Sequences, Discrepancies and Applications*, Lecture Notes in Math. 1651, Springer-Verlag, Berlin, 1997.
[6] T. Gerstner and M. Griebel, *Dimension-adaptive tensor-product quadrature*, Computing, 71 (2003), pp. 65–87.
[7] M. Griebel and W. Knapek, *Optimized tensor-product approximation spaces*, Constr. Approx., 16 (2000), pp. 252–540.
[8] M. Griebel, P. Oswald, and T. Schiekofer, *Sparse grids for boundary integral equations*, Numer. Math., 83 (1999), pp. 279–312.
[9] F.J. Hickernell, I.H. Sloan, and G.W. Wasilkowski, *On tractability of weighted integration over bounded and unbounded regions in* $\mathbb{R}^s$, Math. Comp., 73 (2004), pp. 1885–1901.
[10] Y. Li, *Applicability of Smolyak's algorithm to certain Banach spaces of functions*, J. Complexity, 18 (2002), pp. 792–814.
[11] Y. Li and G.W. Wasilkowski, *Worst case complexity of weighted approximation and integration over* $\mathbb{R}^d$, J. Complexity, 18 (2002), pp. 330–345.
[12] H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods*, CBMS-NSF Regional Conf. Ser. in Appl. Math. 63, SIAM, Philadelphia, 1992.
[13] E. Novak, *Deterministic and Stochastic Error Bounds in Numerical Analysis*, Lecture Notes in Math. 1349, Springer-Verlag, Berlin, 1988.
[14] E. Novak and K. Ritter, *High dimensional integration of smooth functions over cubes*, Numer. Math., 75 (1996), pp. 79–97.
[15] E. Novak and H. Woźniakowski, *When are integration and discrepancy tractable?*, in Foundation of Computational Mathematics, Oxford, 1999, R.A. DeVore, A. Iserles, and E. Süli, eds., Cambridge University Press, Cambridge, UK, 2001, pp. 211–266.
[16] K. Petras, *On the Smolyak cubature error for analytical functions*, Adv. Comput. Math., 12 (2000), 71–93.
[17] A. Pinkus, *n-Widths in Approximation Theory*, Springer-Verlag, Berlin, 1985.
[18] L. Plaskota and G.W. Wasilkowski, *The exact exponent of sparse grid quadratures in the weighted case*, J. Complexity, 17 (2001), pp. 840–849.
[19] I.H. Sloan and S. Joe, *Lattice Methods for Multiple Integration*, Oxford University Press, Oxford, 1994.
[20] I.H. Sloan and H. Woźniakowski, *When are quasi-Monte Carlo algorithms efficient for high dimensional integrals*, J. Complexity, 14 (1998), pp. 1–33.
[21] S.A. Smolyak, *Quadrature and interpolation formulas for tensor products of certain classes of functions*, Dokl. Akad. Nauk SSSR, 4 (1963), pp. 240–243.
[22] V.N. Temlyakov, *Approximation of Periodic Functions*, Nova Science, New York, 1993.
[23] V.N. Temlyakov, *On approximate recovery of functions with bounded mixed derivative*, J. Complexity, 9 (1993), pp. 41–59.
[24] J.F. Traub, G.W. Wasilkowski, and H. Woźniakowski, *Information-Base Complexity*, Academic Press, New York, 1988.
[25] G.W. Wasilkowski and H. Woźniakowski, *Explicit cost bounds of algorithms for multivariate tensor product problems*, J. Complexity, 11 (1995), pp. 1–56.
[26] G.W. Wasilkowski and H. Woźniakowski, *Weighted tensor-product algorithms for linear multivariate problems*, J. Complexity, 15 (1999), pp. 402–447.
[27] G.W. Wasilkowski and H. Woźniakowski, *Complexity of weighted approximation over* $\mathbb{R}^1$, J. Approx. Theory, 103 (2000), pp. 223–251.