STUDIA MATHEMATICA 102 (3) (1992)

ε-Entropy and moduli of smoothness in L^p-spaces

by

A. KAMONT (Sopot)

Abstract. The asymptotic behaviour of ε-entropy of classes of Lipschitz functions in $L^p(\mathbb{I}^d)$ is obtained. Moreover, the asymptotics of ε-entropy of classes of Lipschitz functions in $L^p(\mathbb{R}^d)$ whose tail function decreases as $O(\lambda^{-\gamma})$ is obtained. In case $p = 1$ the relation between the ε-entropy of a given class of probability densities on $\mathbb{R}^d$ and the minimax risk for that class is discussed.

1. Introduction. First we recall the notions of ε-entropy and ε-capacity (cf. [8]).

Let (X, ϱ) be a metric space, and let A be a subset of X. A family $\{U_t\}_{t\in T}$ of subsets of X is called an *ε-covering* of A if $A \subset \bigcup_{t\in T} U_t$ and the diameter of each U_t does not exceed 2ε. A subset $U \subset X$ is called an *ε-net* of A if for each $a \in A$ there exists $u \in U$ such that $\varrho(a, u) \le \varepsilon$. A subset $U \subset X$ is called *ε-distinguishable* if $\varrho(u_1, u_2) > \varepsilon$ for any distinct $u_1, u_2 \in U$.

A subset $A \subset X$ is called *totally bounded* if for each $\varepsilon > 0$ there exists a finite ε-net of A. For a totally bounded $A \subset X$ define

$$\mathcal{N}_\varepsilon(A) = \min\{\#T : \{U_t\}_{t\in T} \text{ is an } \varepsilon\text{-covering of } A\},$$
$$\mathcal{N}_\varepsilon^X(A) = \min\{\#U : U \text{ is an } \varepsilon\text{-net of } A,\ U \subset X\},$$
$$\mathcal{M}_\varepsilon(A) = \sup\{\#M : M \text{ is an } \varepsilon\text{-distinguishable subset of } A\},$$

where $\#T$ denotes the cardinality of T, and

$$\mathcal{H}_\varepsilon(A) = \ln \mathcal{N}_\varepsilon(A), \quad \mathcal{H}_\varepsilon^X(A) = \ln \mathcal{N}_\varepsilon^X(A), \quad \mathcal{C}_\varepsilon(A) = \ln \mathcal{M}_\varepsilon(A),$$

where ln is the natural logarithm.

$\mathcal{H}_\varepsilon(A)$ and $\mathcal{C}_\varepsilon(A)$ are called the *ε-entropy* and *ε-capacity* of A respectively. Note that if $A \subset B$ then $\mathcal{N}_\varepsilon(A) \le \mathcal{N}_\varepsilon(B)$, $\mathcal{M}_\varepsilon(A) \le \mathcal{M}_\varepsilon(B)$, $\mathcal{H}_\varepsilon(A) \le \mathcal{H}_\varepsilon(B)$ and $\mathcal{C}_\varepsilon(A) \le \mathcal{C}_\varepsilon(B)$.

The following inequalities are satisfied ([8]):

$$\mathcal{M}_{2\varepsilon}(A) \le \mathcal{N}_\varepsilon(A) \le \mathcal{N}_\varepsilon^X(A) \le \mathcal{N}_\varepsilon^A(A) \le \mathcal{M}_\varepsilon(A), \tag{1.1}$$
$$\mathcal{C}_{2\varepsilon}(A) \le \mathcal{H}_\varepsilon(A) \le \mathcal{H}_\varepsilon^X(A) \le \mathcal{H}_\varepsilon^A(A) \le \mathcal{C}_\varepsilon(A). \tag{1.2}$$

1991 *Mathematics Subject Classification*: 41A25, 41A29, 41A46, 41A63, 62G05.

To formulate the main result we recall the definitions of moduli of smoothness of order m in $L^p(\mathbb{I}^d)$ and $L^p(\mathbb{R}^d)$ for $d \in \mathbb{N}$, $1 \le p < \infty$, where $\mathbb{I} = [0,1]$. For $f \in L^p(\mathbb{R}^d)$, $m \in \mathbb{N}$ and $u \in \mathbb{R}^d$ let

$$\Delta_u^m f(x) = \sum_{i=0}^{m} (-1)^{m+i} \binom{m}{i} f(x + iu)$$

and for $\delta > 0$ define

$$\omega_{m,p}(f,\delta) = \sup_{\|u\| \le \delta} \|\Delta_u^m f\|_p \,,$$

where $\|u\| = (u_1^2 + \ldots + u_d^2)^{1/2}$ for $u = (u_1, \ldots, u_d) \in \mathbb{R}^d$.

For $f \in L^p(\mathbb{I}^d)$ and $u \in \mathbb{R}^d$ let

$$\mathbb{I}^d(u) = \{x \in \mathbb{I}^d : x + u \in \mathbb{I}^d\}$$

and for $m \in \mathbb{N}$, $0 < \delta \le 1/m$ define

$$\omega_{m,p}(f,\delta) = \sup_{\|u\| \le \delta} \Big(\int_{I^d(mu)} |\Delta_u^m f(x)|^p \, dx \Big)^{1/p} .$$

The functional $\omega_{m,p}(f,\delta)$ is the modulus of smoothness of the function f of order m in the L^p-norm. In addition we put $\omega_{m,p}(f,\delta) = \|f\|_p$ for $m = 0$.

For $d, m \in \mathbb{N}$, $0 < \alpha < m$, $1 \le p < \infty$ and $C > 0$ define

$$A(p,d,m,\alpha,C) = \{f \in L^p(\mathbb{I}^d) : \|f\|_p \le C\,, \quad \omega_{m,p}(f,\delta) \le C\delta^\alpha \text{ for } 0 < \delta \le 1/m\}.$$

For $f \in L^p(\mathbb{R}^d)$ the *tail function* $\Phi_p(f,\lambda)$ is defined as follows:

$$Z_\lambda = [-\lambda, \lambda]^d \quad \text{for } \lambda > 0\,,$$

$$\Phi_p(f,\lambda) = \Big(\int_{\mathbb{R}^d \setminus Z_\lambda} |f(x)|^p \, dx \Big)^{1/p} \quad \text{for } f \in L^p(\mathbb{R}^d).$$

Now for $d, m \in \mathbb{N}$, $0 < \alpha < m$, $\gamma > 0$ $1 \le p < \infty$ and $C > 0$ define

$$U(p,d,m,\alpha,\gamma,C) = \{f \in L^p(\mathbb{R}^d) : \|f\|_p \le C,\ \omega_{m,p}(f,\delta) \le C\delta^\alpha \text{ for } \delta > 0,\ \Phi_p(f,\lambda) \le C\lambda^{-\gamma} \text{ for } \lambda > 0\}.$$

We will discuss the asymptotic behaviour of the ε-entropy of $A(p,d,m,\alpha,C)$ and $U(p,d,m,\alpha,\gamma,C)$ in the metric induced from the spaces $L^p(\mathbb{I}^d)$ and $L^p(\mathbb{R}^d)$ respectively. It will be shown that there exist constants $a_1, a_2 > 0$ and $\varepsilon_0 > 0$, independent of C and ε, such that for all $0 < \varepsilon \le C\varepsilon_0$

$$a_1(C/\varepsilon)^{d/\alpha} \le \mathcal{H}_\varepsilon(A(p,d,m,\alpha,C)) \le a_2(C/\varepsilon)^{d/\alpha}.$$

Results of this type are already known ([9], Theorem 10); however, we present the proof, since the dependence of $\mathcal{H}_\varepsilon(A(p,d,m,\alpha,C))$ on C is needed for estimating $\mathcal{H}_\varepsilon(U(p,d,m,\alpha,\gamma,C))$. Moreover, it will be shown that there exist constants $k_1, k_2 > 0$ and $\varepsilon_0 > 0$, independent of C and ε, such that for all $0 < \varepsilon \le C\varepsilon_0$

$$k_1(C/\varepsilon)^{d(1/\alpha+1/\gamma)} \le \mathcal{H}_\varepsilon(U(p,d,m,\alpha,\gamma,C)) \le k_2(C/\varepsilon)^{d(1/\alpha+1/\gamma)}.$$

The proofs rely very much on the following theorem of G. G. Lorentz ([9], Theorem 2):

THEOREM 1.3. *Let $(X, \|\cdot\|)$ be a real separable Banach space, and let $\Phi = \{\varphi_1, \varphi_2, \ldots\}$ be a sequence of linearly independent elements of X such that $X = \operatorname{span}\Phi$. For a given $f \in X$ set*

$$E_0(f) = \|f\|\,, \qquad E_n(f) = \inf\Big\{\Big\|f - \sum_{k=1}^{n} a_k\varphi_k\Big\| : a_1, \ldots, a_n \in \mathbb{R}\Big\}.$$

Let $\Delta = \{\delta_0, \delta_1, \ldots\}$ be a nonincreasing sequence of positive numbers such that $\lim_{n\to\infty} \delta_n = 0$, and define

$$\mathcal{A}(\Delta, \Phi) = \{f \in X : E_n(f) \le \delta_n\,,\ n = 0, 1, \ldots\}.$$

For $c \in \mathbb{R}$ with $1 < c < 4$ put

$$N_0 = 0\,, \qquad N_i = \min\{k : \delta_k \le c^{-i}\} \quad \textit{for } i = 1, 2, \ldots$$

Given $0 < \varepsilon < 1$, let $j \in \mathbb{N}$ be such that $c^{-(j-1)} < \varepsilon \le c^{-(j-2)}$. Then

$$N_{j-3} \ln c \le \mathcal{C}_\varepsilon(\mathcal{A}(\Delta, \Phi))\,, \tag{1.4}$$

$$\mathcal{H}_\varepsilon(\mathcal{A}(\Delta,\Phi)) \le \mathcal{H}_\varepsilon^X(\mathcal{A}(\Delta,\Phi)) \le (N_1 + \ldots + N_j)(1 + \ln c) + N_j \ln\frac{12}{c-1} + N_1 \ln \delta_0\,. \tag{1.5}$$

In the sequel the following notation will be used. For a multi-index $\mathbf{a} = (a_1, \ldots, a_d) \in (\mathbb{N} \cup \{0\})^d$ define $|\mathbf{a}| = a_1 + \ldots + a_d$, $D^{\mathbf{a}} = \partial^{|\mathbf{a}|}/\partial x_1^{a_1} \ldots \partial x_d^{a_d}$; if in addition $\mathbf{b} = (b_1, \ldots, b_d) \in (\mathbb{N} \cup \{0\})^d$, then we write $\mathbf{a} \le \mathbf{b}$ iff $a_j \le b_j$ for $j = 1, \ldots, d$.

2. The asymptotics of $\mathcal{H}_\varepsilon(A(p,d,m,\alpha,C))$. In this section the asymptotics for $\mathcal{H}_\varepsilon(A(p,d,m,\alpha,C))$ will be obtained by means of Theorem 1.3. We will find a sequence Φ in $L^p(\mathbb{I}^d)$ and two sequences of positive numbers Δ' and Δ'' such that

$$\mathcal{A}(\Delta', \Phi) \subset A(p,d,m,\alpha,C)) \subset \mathcal{A}(\Delta'', \Phi).$$

It occurs that the sequence Φ can be chosen in such a way that its elements are spline functions with dyadic knots. Some necessary definitions are recalled below.

For each $n \in \mathbb{N}$, $n \ge 2$, there exists exactly one pair of integers μ, k such that $n = 2^\mu + k$, $\mu \ge 0$, $1 \le k \le 2^\mu$; put

$$s_{n,j} = \begin{cases} j/2^{\mu+1} & \text{for } j \le 2k\,, \\ (j-k)/2^\mu & \text{for } j \ge 2k+1\,. \end{cases}$$

For $r \in \mathbb{N}$ the B-splines $N_{n,j}^{(r)}$ are defined by the formula

$$N_{n,j}^{(r)}(t) = (s_{n,j+r} - s_{n,j})[s_{n,j}, \ldots, s_{n,j+r}; (s-t)_+^{r-1}]$$

(where the square brackets denote the divided difference of $(s-t)_+^{r-1}$, taken in the variable s at the points $s_{n,j}, \ldots, s_{n,j+r}$). For the properties of the B-splines we refer e.g. to [1]; some of those properties are listed below:

(2.1) $N_{n,j}^{(r)}$ is a function of class C^{r-2} and is a polynomial of degree at most $r-1$ on each interval $[s_{n,i}, s_{n,i+1}]$.

(2.2) $N_{n,j}^{(r)} \geq 0$, $\operatorname{supp} N_{n,j}^{(r)} = [s_{n,j}, s_{n,j+r}]$.

(2.3) For a given $n \in \mathbb{N}$ and an interval (a,b) those functions $N_{n,j}^{(r)}$ which are nontrivial when restricted to (a,b), are linearly independent over that interval.

In the following the restrictions of $N_{n,j}^{(r)}$ to $\mathbb{I}$ will be considered. Let

$$S_n^{(r)} = \operatorname{span}\{N_{n,j}^{(r)} : j = -r+1, \ldots, n-1\} \quad \text{for } n \geq 2,$$

and for $n = -r+2, \ldots, 1$ let $S_n^{(r)}$ be the space of polynomials of degree at most $n+r-1$, restricted to $\mathbb{I}$. Then $\dim S_n^{(r)} = n+r-1$ and $S_n^{(r)} \subset S_{n+1}^{(r)}$. Let the sequence of functions $(f_j^{(r)}, j \geq 2-r)$ be defined as follows: $f_{2-r}^{(r)} = 1$, $f_{j+1}^{(r)} \in S_{j+1}^{(r)}$, $f_{j+1}^{(r)}$ is orthogonal to $S_j^{(r)}$ (with respect to the inner product in $L^2(\mathbb{I})$: $(f,g) = \int_0^1 f(x)g(x)\,dx$), $\|f_{j+1}^{(r)}\|_2 = 1$. Using the notation $Hf(t) = \int_t^1 f(x)\,dx$, $Df(t) = \frac{d}{dt}f(t)$, define for $k \in \mathbb{Z}$, $|k| < r$,

$$f_j^{(r,k)} = \begin{cases} f_j^{(r)} & \text{for } k = 0, \\ D^k f_j^{(r)} & \text{for } 0 < k < r, \\ H^{-k} f_j^{(r)} & \text{for } -r < k < 0. \end{cases}$$

Now we introduce the tensor product spline functions on $\mathbb{I}^d$.

For $\mathbf{r} = (r_1, \ldots, r_d) \in \mathbb{N}^d$, $\mathbf{k} = (k_1, \ldots, k_d) \in \mathbb{Z}^d$, $|k_i| < r_i$, $\mathbf{j} = (j_1, \ldots, j_d)$, $j_i \geq 2 - r_i + k_i$, $\mathbf{t} = (t_1, \ldots, t_d) \in \mathbb{I}^d$ define

$$f_{\mathbf{j}}^{(\mathbf{r},\mathbf{k})}(\mathbf{t}) = f_{j_1}^{(r_1,k_1)}(t_1) \ldots f_{j_d}^{(r_d,k_d)}(t_d).$$

For $m \in \mathbb{N}$ let $\mathbf{k}_m = (m+2, \ldots, m+2) \in \mathbb{N}^d$, $\mathbf{r}_m = 2\,\mathbf{k}_m$; let

$$F_{\mathbf{j}}^{(m)} = f_{\mathbf{j}}^{(\mathbf{r}_m,\mathbf{k}_m)}, \qquad G_{\mathbf{j}}^{(m)} = f_{\mathbf{j}}^{(\mathbf{r}_m,-\mathbf{k}_m)}$$

for $\mathbf{j} = (j_1, \ldots, j_d)$ such that $j_i \geq -m$ for $i = 1, \ldots, d$; observe that $F_{\mathbf{j}}^{(m)}$ are functions of class C^m and $(F_{\mathbf{j}}^{(m)}, G_{\mathbf{i}}^{(m)}) = \delta_{\mathbf{j}\mathbf{i}}$. Now, for $f \in L^p(\mathbb{I}^d)$, $n \geq -m$, $\mathbf{n} = (n, \ldots, n)$, $\mathbf{m} = (m, \ldots, m)$ set

$$Q_n^{(m)}(f) = \sum_{-\mathbf{m} \leq \mathbf{j} \leq \mathbf{n}} (f, G_{\mathbf{j}}^{(m)}) F_{\mathbf{j}}^{(m)}.$$

Note that $Q_n^{(m)}$ is a projection onto the space

$$V_n^{(m)} = \operatorname{span}\{F_{\mathbf{j}}^{(m)} : -\mathbf{m} \leq \mathbf{j} \leq \mathbf{n}\}.$$

For $f \in L^p(\mathbb{I}^d)$ let

$$E_{n,p}^{(m)}(f) = \inf\{\|f-g\|_p : g \in V_n^{(m)}\}.$$

The following lemmas establish some relations between $E_{n,p}^{(m)}(f)$ and $\omega_{m,p}(f,\delta)$. Lemma 2.4 was proved in [5] as Theorem 9.18. The proof of Lemma 2.5 is omitted, since it can be proved similarly to Theorem 10 of [2], with the help of (5.14) of [4] and the inequality

$$\Big(\int_{\mathbb{I}^d((m+k)u)} |\Delta_u^{m+k} f(x)|^p\,dx\Big)^{1/p} \leq A_{k,d}\|u\|^k \sum_{|\mathbf{a}|=k} \Big(\int_{\mathbb{I}^d(mu)} |\Delta_u^m D^{\mathbf{a}} f(x)|^p\,dx\Big)^{1/p},$$

which holds for $k \in \mathbb{N}$, $f \in C^k(\mathbb{I}^d)$, $m \in \mathbb{N} \cup \{0\}$, $u \in \mathbb{R}^d$, $\|u\| \leq 1/(m+k)$, $1 \leq p < \infty$, where $A_{k,d}$ is a constant independent of f, u, m and p.

LEMMA 2.4. *Let $1 \leq p < \infty$ and $m, d \in \mathbb{N}$ be given. There exists a constant $M_{m,p,d}$ such that*

$$E_{n,p}^{(m)}(f) \leq M_{m,p,d}\,\omega_{m,p}(f, 1/n)$$

for every $n \geq m$ and $f \in L^p(\mathbb{I}^d)$.

LEMMA 2.5. *Let $1 \leq p < \infty$ and $m, d \in \mathbb{N}$ be given. There exists a constant $M'_{m,p,d}$ such that*

$$\omega_{m,p}(f, 1/n) \leq M'_{m,p,d}\, n^{-m}\Big(\|f\|_p + \sum_{i=m}^n i^{m-1} E_{i,p}^{(m)}(f)\Big)$$

for every $n \geq m$ and $f \in L^p(\mathbb{I}^d)$.

Let

$$B(p,d,m,\alpha,C) = \{f \in L^p(\mathbb{I}^d) : \|f\|_p \leq C,\ E_{n,p}^{(m)}(f) \leq C/n^\alpha \text{ for } n \geq m\}.$$

COROLLARY 2.6. *Let $1 \le p < \infty$, $m, d \in \mathbb{N}$ and $0 < \alpha < m$ be given. There exist constants $c_1, c_2 > 0$, depending only on p, m, α and d, such that for every $C > 0$*

$$B(p,d,m,\alpha,c_1C) \subset A(p,d,m,\alpha,C) \subset B(p,d,m,\alpha,c_2C).$$

Proof. The existence of c_2 follows from Lemma 2.4.

As $0 < \alpha < m$, there exists $a_{m,\alpha}$ such that $1 + \sum_{i=m}^{n} i^{m-1-\alpha} \le a_{m,\alpha} n^{m-\alpha}$ for every $n \ge m$. The existence of c_1 now follows from Lemma 2.5 and the inequality $\omega_{m,p}(f,\delta_1) \le \omega_{m,p}(f,\delta_2)$ for $0 < \delta_1 \le \delta_2 \le 1/m$. ∎

Now we will find the asymptotic behaviour of the ε-entropy of $B(p,d,m,\alpha,M)$.

LEMMA 2.7. *Let $1 \le p < \infty$, $m, d \in \mathbb{N}$ and $0 < \alpha < m$ be given. There exist constants $a_1, a_2 > 0$ such that the following inequalities hold for each $M > 0$:*

$$a_1(M/\varepsilon)^{d/\alpha} \le \mathcal{H}_\varepsilon(B(p,d,m,\alpha,M)) \quad \textit{for } 0 < \varepsilon \le M/2,$$
$$\mathcal{H}_\varepsilon(B(p,d,m,\alpha,M)) \le a_2(M/\varepsilon)^{d/\alpha} \quad \textit{for } 0 < \varepsilon \le M.$$

Proof. As $B(p,d,m,\alpha,M) = \{Mf : f \in B(p,d,m,\alpha,1)\}$ and

$$\mathcal{H}_\varepsilon(B(p,d,m,\alpha,M)) = \mathcal{H}_{\varepsilon/M}(B(p,d,m,\alpha,1)),$$

it is enough to estimate $\mathcal{H}_\varepsilon(B(p,d,m,\alpha,1))$. We will use Theorem 1.3.

Consider $V_n^{(m)}$ for $n \ge m$ as subspaces of $X = L^p(\mathbb{I}^d)$ and set

$$d_n^{(m)} = \dim V_n^{(m)} = (n+m+1)^d.$$

The sequence $\Phi = \{\varphi_1, \varphi_2, \ldots\}$ is obtained by ordering $\{F_{\mathbf{j}}^{(m)} : \mathbf{j} = (j_1,\ldots,j_d) \ge (-m,\ldots,-m)\}$ so that $\operatorname{span}\{\varphi_1,\ldots,\varphi_{d_n^{(m)}}\} = V_n^{(m)}$ for every $n \ge m$. The sequence Δ is defined as follows:

$$\delta_0 = \ldots = \delta_{d_m^{(m)}-1} = 1, \quad \delta_k = 1/n^\alpha \quad \text{for } d_n^{(m)} \le k < d_{n+1}^{(m)}.$$

Then $B(p,d,m,\alpha,1) = \mathcal{A}(\Delta,\Phi)$. Let $c = e$ in Theorem 1.3 and write $n_i = \max(\lceil e^{i/\alpha}\rceil, m)$ (where $\lceil a\rceil = \min\{k \in \mathbb{Z} : a \le k\}$); then $N_i = d_{n_i}^{(m)}$.

First the constant a_1 will be found.

Inequality (1.2) implies

$$\mathcal{C}_{2\varepsilon}(B(p,d,m,\alpha,1)) \le \mathcal{H}_\varepsilon(B(p,d,m,\alpha,1)).$$

For $\varepsilon \le 1/2$ choose j such that $e^{-(j-1)} < 2\varepsilon \le e^{-(j-2)}$; then (1.4) yields

$$\begin{aligned}\mathcal{C}_{2\varepsilon}(B(p,d,m,\alpha,1)) &\ge N_{j-3} = d_{n_{j-3}}^{(m)} = (n_{j-3}+m+1)^d \\ &\ge (e^{(j-3)/\alpha}+m+1)^d \ge e^{(j-3)d/\alpha} \\ &= e^{-2d/\alpha}e^{(j-1)d/\alpha} \ge e^{-2d/\alpha}\left(\frac{1}{2\varepsilon}\right)^{d/\alpha}.\end{aligned}$$

Setting $a_1 = e^{-2d/\alpha}(1/2)^{d/\alpha}$ we obtain for any $0 < \varepsilon \le 1/2$

$$a_1(1/\varepsilon)^{d/\alpha} \le \mathcal{C}_\varepsilon(B(p,d,m,\alpha,1)). \tag{2.8}$$

Now the constant a_2 will be found. For $\varepsilon \le 1$ choose j such that $e^{-(j-1)} < \varepsilon \le e^{-(j-2)}$. As $N_i \le (e^{i/\alpha}+m+2)^d + (2m+1)^d$, it follows from (1.5) that

$$\begin{aligned}\mathcal{H}_\varepsilon(B(p,d,m,\alpha,1)) &\le 2\sum_{i=1}^{j} N_i + N_j \ln\frac{12}{e-1} \le b\sum_{i=1}^{j} N_i \\ &\le b(2m+1)^d j + b\sum_{i=1}^{j}(e^{i/\alpha}+m+2)^d \\ &\le b\sum_{k=1}^{d}\binom{d}{k}(m+2)^{d-k}\frac{e^{k/\alpha}}{e^{k/\alpha}-1}e^{jk/\alpha} \\ &\quad + b((m+2)^d+(2m+1)^d)j,\end{aligned}$$

where $b = 2 + \ln(12/(e-1))$. Setting

$$b_1 = b\sum_{k=1}^{d}\binom{d}{k}(m+2)^{d-k}\frac{e^{k/\alpha}}{e^{k/\alpha}-1}, \quad b_2 = b((m+2)^d+(2m+1)^d),$$

we get

$$\mathcal{H}_\varepsilon(B(p,d,m,\alpha,1)) \le b_1 e^{jd/\alpha} + b_2 j < b_1 e^{2d/\alpha}\left(\frac{1}{\varepsilon}\right)^{d/\alpha} + b_2\left(\ln\frac{1}{\varepsilon}+2\right).$$

As

$$\ln\frac{1}{\varepsilon} \le \frac{\alpha}{de}\left(\frac{1}{\varepsilon}\right)^{d/\alpha},$$

by setting $a_2 = b_1 e^{2d/\alpha} + \frac{\alpha}{de}b_2 + 2b_2$ we obtain

$$\mathcal{H}_\varepsilon(B(p,d,m,\alpha,1)) \le a_2(1/\varepsilon)^{d/\alpha} \tag{2.9}$$

for any $0 < \varepsilon \le 1$. Our lemma now follows from (2.8) and (2.9). ∎

COROLLARY 2.10. *Let $1 \le p < \infty$, $m, d \in \mathbb{N}$ and $0 < \alpha < m$ be given. There exist constants $a_1, a_2 > 0$ and $\varepsilon_0 > 0$ such that*

$$a_1(M/\varepsilon)^{d/\alpha} \le \mathcal{H}_\varepsilon(A(p,d,m,\alpha,M)) \le a_2(M/\varepsilon)^{d/\alpha}$$

for each $M > 0$ and $0 < \varepsilon \le M\varepsilon_0$.

This follows from Lemma 2.7 and Corollary 2.6.

The case $m = 1$ will be considered more carefully. Define

$$A^+(p,d,1,\alpha,M) = \{f \in A(p,d,1,\alpha,M) : f \ge 0\}.$$

For $f \in L^p(\mathbb{I}^d)$ set $f_+ = \max(f,0)$, $f_- = \max(-f,0)$; then $f_+, f_- \geq 0$, $f = f_+ - f_-$. As $g(a) = \max(a,0)$ is a Lipschitz function with constant 1, for any $x, y \in \mathbb{I}^d$

$$|f_+(x) - f_+(y)| \leq |f(x) - f(y)|, \quad |f_-(x) - f_-(y)| \leq |f(x) - f(y)|,$$

which implies

$$\omega_{1,p}(f_+, \delta) \leq \omega_{1,p}(f, \delta) \quad \text{and} \quad \omega_{1,p}(f_-, \delta) \leq \omega_{1,p}(f, \delta)$$

for each $0 < \delta \leq 1$; in addition $\|f_+\|_p \leq \|f\|_p$, $\|f_-\|_p \leq \|f\|_p$. Therefore, if $f \in A(p,d,1,\alpha,M)$ then $f_+, f_- \in A^+(p,d,1,\alpha,M)$. Let $U = \{U_1, \ldots, U_k\}$ be an $\varepsilon/2$-covering of $A^+(p,d,1,\alpha,M)$, and $V_{ij} = \{f \in A(p,d,1,\alpha,M) : f_+ \in U_i,\ f_- \in U_j\}$ for $i,j = 1,\ldots,k$. Then $V = \{V_{ij} : i,j = 1,\ldots,k\}$ is an ε-covering of $A(p,d,1,\alpha,M)$ with $\#V \leq k^2 = (\#U)^2$; this implies

$$\mathcal{H}_\varepsilon(A(p,d,1,\alpha,M)) \leq 2\mathcal{H}_{\varepsilon/2}(A^+(p,d,1,\alpha,M)).$$

As $A^+(p,d,1,\alpha,M) \subset A(p,d,1,\alpha,M)$, the last inequality and Corollary 2.10 yield

COROLLARY 2.11. *Let $1 \leq p < \infty$, $d \in \mathbb{N}$ and $0 < \alpha < 1$ be given. There exist constants $a_1, a_2 > 0$ and $\varepsilon_0 > 0$ such that*

$$a_1(M/\varepsilon)^{d/\alpha} \leq \mathcal{H}_\varepsilon(A^+(p,d,1,\alpha,M)) \leq a_2(M/\varepsilon)^{d/\alpha}$$

for every $M > 0$ and $0 < \varepsilon \leq M\varepsilon_0$.

LEMMA 2.12. *Let $1 \leq p < \infty$, $d \in \mathbb{N}$, $0 < \alpha < 1$, $M_1, M_2 > 0$ be given; let*

$$M(p,d,\alpha,M_1,M_2) = \{f \in L^p(\mathbb{I}^d) : f \geq 0,\ \|f\|_p = M_1,\ \omega_{1,p}(f,\delta) \leq M_1M_2\delta^\alpha \text{ for } 0 < \delta \leq 1\}.$$

Then there are constants $m_1, m_2 > 0$ and $\varepsilon_0 > 0$ such that

$$m_1(1/\varepsilon)^{d/\alpha} \leq \mathcal{H}_\varepsilon(M(p,d,\alpha,M_1,M_2)) \leq m_2(1/\varepsilon)^{d/\alpha}$$

for $0 < \varepsilon \leq \varepsilon_0$.

Proof. As for $M_0 = \max(M_1, M_1M_2)$ we have $M(p,d,\alpha,M_1,M_2) \subset A^+(p,d,1,\alpha,M_0)$, the existence of $m_2 > 0$ and $\varepsilon_2 > 0$ such that

$$\mathcal{H}_\varepsilon(M(p,d,\alpha,M_1,M_2)) \leq m_2(1/\varepsilon)^{d/\alpha}$$

for every $0 < \varepsilon \leq \varepsilon_2$ follows from Corollary 2.11.

Now we prove the existence of m_1. Define

$$\widetilde{M}(p,d,\alpha,M_1,M_2) = \{f \in L^p(\mathbb{I}^d) : f \geq 0,\ \|f\|_p \leq M_1,\ \omega_{1,p}(f,\delta) \leq M_2\|f\|_p\delta^\alpha \text{ for } 0 < \delta \leq 1\}.$$

Set $C_0 = \min(M_1/2, M_1M_2/2)$; for every constant c and $f \in L^p(\mathbb{I}^d)$ we have $\omega_{1,p}(f,\delta) = \omega_{1,p}(f+c,\delta)$, and so $\{f + C_0 : f \in A^+(p,d,1,\alpha,C_0)\} \subset \widetilde{M}(p,d,\alpha,M_1,M_2)$. Thus $\widetilde{M}(p,d,\alpha,M_1,M_2)$ contains a subset isometric to $A^+(p,d,1,\alpha,C_0)$ and it follows from Corollary 2.11 that there are $m^* > 0$ and $\eta > 0$ such that

$$m^*(1/\varepsilon)^{d/\alpha} \leq \mathcal{H}_\varepsilon(\widetilde{M}(p,d,\alpha,M_1,M_2)) \tag{2.13}$$

for each $0 < \varepsilon \leq \eta$.

Now, given an $\varepsilon/2$-net of $M(p,d,\alpha,M_1,M_2)$, we construct an ε-net of $\widetilde{M}(p,d,\alpha,M_1,M_2)$.

First notice that, for any $a > 0$, $\|f-g\|_p \leq \varepsilon \leftrightarrow \|af - ag\| \leq a\varepsilon$ and

$$M(p,d,\alpha,aM_1,M_2) = \{af : f \in M(p,d,\alpha,M_1,M_2)\}. \tag{2.14}$$

Let $\mathcal{N}^*_\varepsilon(a)$ denote the minimal number of elements of an ε-net of $M(p,d,\alpha,aM_1,M_2)$, consisting only of elements of this set. It follows from (2.14) that $\mathcal{N}^*_\varepsilon(a) = \mathcal{N}^*_{\varepsilon/a}(1)$. Let $\varepsilon > 0$ be given, and $r = [M_1/\varepsilon] + 1$. For given $f \in \widetilde{M}(p,d,\alpha,M_1,M_2)$ choose an integer k, $0 \leq k \leq r$, such that

$$M_1\frac{2k-1}{2r} \leq \|f\|_p \leq M_1\frac{2k+1}{2r}.$$

For $\|f\|_p \neq 0$ set

$$f^* = \frac{M_1k}{r}\frac{f}{\|f\|_p};$$

then $f^* \geq 0$, $\|f^*\|_p = (k/r)M_1$,

$$\omega_{1,p}(f^*,\delta) = \frac{M_1k}{r}\frac{1}{\|f\|_p}\omega_{1,p}(f,\delta) \leq \frac{k}{r}M_1M_2\delta^\alpha$$

and $f^* \in M(p,d,\alpha,(k/r)M_1,M_2)$. In addition,

$$\|f - f^*\|_p = \left|\|f\|_p - \frac{M_1k}{r}\right| \leq \frac{M_1}{2r} \leq \frac{M_1}{2}\frac{\varepsilon}{M_1} = \frac{\varepsilon}{2}.$$

Set $U = \{0\} \cup \bigcup_{k=1}^r \{f_{1,k}, \ldots, f_{n_k,k}\}$, where $n_k = \mathcal{N}^*_{\varepsilon/2}(k/r)$, $k = 1,\ldots,r$, and $\{f_{1,k},\ldots,f_{n_k,k}\} \subset M(p,d,\alpha,(k/r)M_1,M_2)$ is an $\varepsilon/2$-net of $M(p,d,\alpha,(k/r)M_1,M_2)$. Then U is an ε-net of $\widetilde{M}(p,d,\alpha,M_1,M_2)$. This and (1.1) yield

$$\begin{aligned}\mathcal{N}_\varepsilon(\widetilde{M}(p,d,\alpha,M_1,M_2)) &\leq \#U = 1 + \sum_{k=1}^r \mathcal{N}^*_{\varepsilon/2}(k/r) = 1 + \sum_{k=1}^r \mathcal{N}^*_{\varepsilon/2\cdot r/k}(1)\\ &\leq (1+r)\mathcal{N}_{\varepsilon/4}(M(p,d,\alpha,M_1,M_2)),\end{aligned}$$

which implies

$$\mathcal{H}_\varepsilon(\widetilde{M}(p,d,\alpha,M_1,M_2)) \leq \ln(M_1/\varepsilon + 2) + \mathcal{H}_{\varepsilon/4}(M(p,d,\alpha,M_1,M_2)).$$

As $\lim_{\varepsilon\to 0}\varepsilon^s \ln(M_1/\varepsilon+2)=0$ for any $s>0$, it follows from the last inequality and (2.13) that there exist $\varepsilon_1>0$ and $m_1>0$ such that

$$m_1(1/\varepsilon)^{d/\alpha}\le \mathcal{H}_\varepsilon(M(p,d,\alpha,M_1,M_2))$$

for every $0<\varepsilon\le\varepsilon_1$. ■

3. The asymptotics of $\mathcal{H}_\varepsilon(U(p,d,m,\alpha,\gamma,C))$. Recall that for $m,d\in\mathbb{N}$, $1\le p<\infty$, $0<\alpha<m$, $\gamma>0$ and $C>0$

$$U(p,d,m,\alpha,\gamma,C)=\{f\in L^p(\mathbb{R}^d):\|f\|_p\le C,\ \omega_{m,p}(f,\delta)\le C\delta^\alpha \text{ for } \delta>0,\ \Phi_p(f,\lambda)\le C\lambda^{-\gamma} \text{ for } \lambda>0\},$$

where $\omega_{m,p}(f,\delta)=\sup_{\|u\|\le\delta}\|\Delta_u^m f\|_p$ is the modulus of smoothness of f of order m in $L^p(\mathbb{R}^d)$, and $\Phi_p(f,\lambda)=(\int_{\mathbb{R}^d\setminus Z_\lambda}|f(x)|^p\,dx)^{1/p}$ is the tail function (recall that $Z_\lambda=[-\lambda,\lambda]^d$).

THEOREM 3.1. *Let $1\le p<\infty$, $m,d\in\mathbb{N}$, $0<\alpha<m$ and $\gamma>0$ be given. There exist constants $k_1,k_2>0$ and $\varepsilon_0>0$, depending only on p, d, m, α and γ, such that*

$$k_1(C/\varepsilon)^{d(1/\alpha+1/\gamma)}\le\mathcal{H}_\varepsilon(U(p,d,m,\alpha,\gamma,C))\le k_2(C/\varepsilon)^{d(1/\alpha+1/\gamma)}$$

for every $C>0$ and $0<\varepsilon\le C\varepsilon_0$.

The proof of Theorem 3.1 is split into several lemmas. First the upper estimate for $\mathcal{H}_\varepsilon(U(p,d,m,\alpha,\gamma,C))$ is obtained.

LEMMA 3.2. *Let $1\le p<\infty$, $m,d\in\mathbb{N}$, $0<\alpha<m$ and $\gamma>0$ be given. There exist constants $k>0$ and $\varepsilon_0>0$, depending only on p, d, m, α and γ, such that*

$$\mathcal{H}_\varepsilon(U(p,d,m,\alpha,\gamma,C))\le k(C/\varepsilon)^{d(1/\alpha+1/\gamma)}$$

for every $C>0$ and $0<\varepsilon\le C\varepsilon_0$.

Proof. For $f\in L^p(Z_\lambda)$ and $0<\delta\le 2\lambda/m$ define

$$\omega_{m,p}^{(\lambda)}(f,\delta)=\sup_{\|u\|\le\delta}\Big(\int_{Z_\lambda(mu)}|\Delta_u^m f(x)|^p\,dx\Big)^{1/p},$$

where $Z_\lambda(mu)=\{x\in Z_\lambda: x+mu\in Z_\lambda\}$. Set

$$U_\lambda(p,d,m,\alpha,C)=\{f\in L^p(Z_\lambda):\|f\|_{L^p(Z_\lambda)}\le C,\ \omega_{m,p}^{(\lambda)}(f,\delta)\le C\delta^\alpha \text{ for } 0<\delta\le 2\lambda/m\}.$$

Then for $f\in U(p,d,m,\alpha,\gamma,C)$ and $\lambda>0$ we have $f_\lambda=f|_{Z_\lambda}\in U_\lambda(p,d,m,\alpha,C)$. For $\varepsilon>0$ write $\lambda_\varepsilon=(2C/\varepsilon)^{1/\gamma}$; then for $f,g\in U(p,d,m,\alpha,\gamma,C)$

$$\|f-g\|_p\le\Big(\int_{Z_{\lambda_\varepsilon}}|f(x)-g(x)|^p\,dx\Big)^{1/p}+\Phi_p(f-g,\lambda_\varepsilon)\le\|f_{\lambda_\varepsilon}-g_{\lambda_\varepsilon}\|_{L^p(Z_{\lambda_\varepsilon})}+\varepsilon,$$

so if $\|f_{\lambda_\varepsilon}-g_{\lambda_\varepsilon}\|_{L^p(Z_{\lambda_\varepsilon})}\le\varepsilon$ then $\|f-g\|_p\le 2\varepsilon$; therefore

$$\mathcal{H}_\varepsilon(U(p,d,m,\alpha,\gamma,C))\le\mathcal{H}_{\varepsilon/2}(U_{\lambda_\varepsilon}(p,d,m,\alpha,C)). \tag{3.3}$$

Now $\mathcal{H}_{\varepsilon/2}(U_{\lambda_\varepsilon}(p,d,m,\alpha,C))$ will be estimated from above; we will find $\eta_\varepsilon>0$ and a subset $A_\varepsilon\subset L^p(\mathbb{I}^d)$ such that $\mathcal{H}_{\varepsilon/2}(U_{\lambda_\varepsilon}(p,d,m,\alpha,C))\le\mathcal{H}_{\eta_\varepsilon}(A_\varepsilon)$.

Define

$$\psi_\varepsilon:[0,1]\to[-\lambda_\varepsilon,\lambda_\varepsilon],\quad \psi_\varepsilon(t)=\lambda_\varepsilon(2t-1),$$
$$\Psi_\varepsilon:[0,1]^d\to Z_{\lambda_\varepsilon},\quad \Psi_\varepsilon(t_1,\ldots,t_d)=(\psi_\varepsilon(t_1),\ldots,\psi_\varepsilon(t_d)).$$

Then for any $f\in L^p(Z_{\lambda_\varepsilon})$

$$\|f\|_{L^p(Z_{\lambda_\varepsilon})}=(2\lambda_\varepsilon)^{d/p}\|f\circ\Psi_\varepsilon\|_{L^p(I^d)},$$
$$\Delta_u^m f(x)=\Delta_{u/(2\lambda_\varepsilon)}^m(f\circ\Psi_\varepsilon)(\Psi_\varepsilon^{-1}(x)),$$
$$Z_{\lambda_\varepsilon}(mu)=\Psi_\varepsilon\Big(\mathbb{I}^d\Big(m\frac{u}{2\lambda_\varepsilon}\Big)\Big),$$

which implies

$$\omega_{m,p}^{(\lambda_\varepsilon)}(f,\delta)=(2\lambda_\varepsilon)^{d/p}\omega_{m,p}\Big(f\circ\Psi_\varepsilon,\frac{\delta}{2\lambda_\varepsilon}\Big)_{L^p(\mathbb{I}^d)},$$

and for $f\in U_{\lambda_\varepsilon}(p,d,m,\alpha,C)$

$$\|f\circ\Psi_\varepsilon\|_{L^p(\mathbb{I}^d)}\le(2\lambda_\varepsilon)^{-d/p}C,$$
$$\omega_{m,p}(f\circ\Psi_\varepsilon,\delta)_{L^p(\mathbb{I}^d)}\le(2\lambda_\varepsilon)^{\alpha-d/p}C\delta^\alpha.$$

Set $\eta_\varepsilon=\varepsilon(2\lambda_\varepsilon)^{-d/p}$, $A_\varepsilon=A(p,d,m,\alpha,(2\lambda_\varepsilon)^{\alpha-d/p}C)$. Notice that if $f,g\in U_{\lambda_\varepsilon}(p,d,m,\alpha,C)$ are such that $\|f\circ\Psi_\varepsilon-g\circ\Psi_\varepsilon\|_{L^p(\mathbb{I}^d)}\le\eta_\varepsilon$, then $\|f-g\|_{L^p(Z_{\lambda_\varepsilon})}\le\varepsilon$. In addition, $\lambda_\varepsilon\ge 1$ for $\varepsilon\le 2C$, and if $f\in U_{\lambda_\varepsilon}(p,d,m,\alpha,C)$ then $f\circ\Psi_\varepsilon\in A_\varepsilon$. Therefore

$$\mathcal{H}_{\varepsilon/2}(U_{\lambda_\varepsilon}(p,d,m,\alpha,C))\le\mathcal{H}_{\eta_\varepsilon}(A_\varepsilon). \tag{3.4}$$

Corollary 2.10 implies the existence of constants $\xi>0$ and $a>0$, independent of C and ε, such that if $\eta_\varepsilon\le(2\lambda_\varepsilon)^{\alpha-d/p}C\xi$ then

$$\mathcal{H}_{\eta_\varepsilon}(A_\varepsilon)\le a((2\lambda_\varepsilon)^{\alpha-d/p}C/\eta_\varepsilon)^{d/\alpha}=k(C/\varepsilon)^{d(1/\alpha+1/\gamma)}, \tag{3.5}$$

where $k=2^{d(1+1/\gamma)}a$.

ε_0 must be chosen so that the inequalities $\varepsilon \le 2C$ and $\eta_\varepsilon \le (2\lambda_\varepsilon)^{\alpha-d/p}C\xi$ hold for any $0<\varepsilon\le C\varepsilon_0$. It is enough to take $\varepsilon_0=\min(2,2\xi^{\gamma/(\alpha+\gamma)})$. Now our lemma follows from (3.3)–(3.5). ∎

To estimate $\mathcal{H}_\varepsilon(U(p,d,m,\alpha,\gamma,C))$ from below it is convenient to use Theorem 1.3. Some spaces of spline functions are also needed, and the classical relation between the distances from those subspaces and the moduli of smoothness will be recalled. The following notation will be used:

$$N^{(s)}(x)=s[0,\ldots,s;(\cdot-x)_+^{s-1}]\quad \text{for } s\in\mathbb{N},\ x\in\mathbb{R},$$

$$N^{(s)}_{i,h}(x)=N^{(s)}\left(\frac{x-ih}{h}\right)\quad \text{for } h>0,\ i\in\mathbb{Z},\ x\in\mathbb{R},$$

$$N^{(\mathbf{z})}_{\mathbf{j},\mathbf{h}}(t_1,\ldots,t_d)=N^{(z_1)}_{j_1,h_1}(t_1)\ldots N^{(z_d)}_{j_d,h_d}(t_d)$$

for $\mathbf{z}=(z_1,\ldots,z_d)\in\mathbb{N}^d$, $\mathbf{j}=(j_1,\ldots,j_d)\in\mathbb{Z}^d$, $\mathbf{h}=(h_1,\ldots,h_d)\in\mathbb{R}^d$, $h_i>0$, $\mathbf{t}=(t_1,\ldots,t_d)\in\mathbb{R}^d$.

Some useful properties of spline functions $N^{(\mathbf{z})}_{\mathbf{j},\mathbf{h}}$ are mentioned below. They are multivariable analogues of Schoenberg's result ([10]).

(3.6) There exists a constant $c_{\mathbf{z}}>0$ such that

$$c_{\mathbf{z}}^{-1}(h_1\ldots h_d)^{1/p}\Big(\sum_{\mathbf{j}\in\mathbb{Z}^d}|a_{\mathbf{j}}|^p\Big)^{1/p} \le \Big\|\sum_{\mathbf{j}\in\mathbb{Z}^d}a_{\mathbf{j}}N^{(\mathbf{z})}_{\mathbf{j},\mathbf{h}}\Big\|_p\le c_{\mathbf{z}}(h_1\ldots h_d)^{1/p}\Big(\sum_{\mathbf{j}\in\mathbb{Z}^d}|a_{\mathbf{j}}|^p\Big)^{1/p}$$

for every sequence $(a_{\mathbf{j}})_{\mathbf{j}\in\mathbb{Z}^d}$.

(3.7) For $\mathbf{a}\in(\mathbb{N}\cup\{0\})^d$ and $\mathbf{z}\in\mathbb{N}^d$ with $(0,\ldots,0)\le\mathbf{a}<\mathbf{z}$ and $g=\sum_{\mathbf{j}\in\mathbb{Z}^d}g_{\mathbf{j}}N^{(\mathbf{z})}_{\mathbf{j},\mathbf{h}}$ we have

$$h_1^{a_1}\ldots h_d^{a_d}D^{\mathbf{a}}g=\sum_{\mathbf{j}\in\mathbb{Z}^d}(\Delta^{\mathbf{a}}g_{\mathbf{j}})N^{(\mathbf{z}-\mathbf{a})}_{\mathbf{j}+\mathbf{a},\mathbf{h}},$$

where

$$\Delta^{\mathbf{a}}g_{\mathbf{j}}=\sum_{0\le\mathbf{b}\le\mathbf{a}}\binom{a_1}{b_1}\ldots\binom{a_d}{b_d}(-1)^{|\mathbf{a}|+|\mathbf{b}|}g_{\mathbf{b}+\mathbf{j}}.$$

(3.8) For each $m,d\in\mathbb{N}$ there exists a constant $M_{m,d}>0$ such that

$$\|\Delta^m_u f\|_p\le M_{m,d}\|u\|^m\sum_{|\mathbf{a}|=m}\|D^{\mathbf{a}}f\|_p$$

for any $1\le p<\infty$, $f\in C^m(\mathbb{R}^d)\cap L^p(\mathbb{R}^d)$ and $u\in\mathbb{R}^d$.

Define

$$\mathbf{r}_m=(m+2,\ldots,m+2)\in\mathbb{N}^d\quad\text{for } m\in\mathbb{N},$$

$$\mathbf{h}_n=(h_n,\ldots,h_n)\in\mathbb{R}^d,\quad h_n=1/2^n\quad\text{for } n\in\mathbb{N}\cup\{0\},$$
$$V^{(m)}_{n,p}=\operatorname{span}\{N^{(\mathbf{r}_m)}_{\mathbf{j},\mathbf{h}_n}:\mathbf{j}\in\mathbb{Z}^d\}\cap L^p(\mathbb{R}^d).$$

Notice that $V^{(m)}_{n,p}\subset V^{(m)}_{n+1,p}$.

For $f\in L^p(\mathbb{R}^d)$ define

$$E^{(m)}_{n,p}(f)=\inf\{\|f-g\|_p: g\in V^{(m)}_{n,p}\}.$$

The following lemma recalls the well-known relation between $E^{(m)}_{n,p}(f)$ and $\omega_{m,p}(f,\delta)$ (cf. [3]).

LEMMA 3.9. *Let $m,d\in\mathbb{N}$ be given. There exists a constant $C_{m,d}$ such that*

$$\omega_{m,p}\left(f,\frac{1}{2^k}\right)\le\frac{C_{m,d}}{2^{km}}\Big(\|f\|_p+\sum_{i=0}^k 2^{im}E^{(m)}_{i,p}(f)\Big)$$

for every $1\le p<\infty$, $f\in L^p(\mathbb{R}^d)$ and $k\in\mathbb{N}\cup\{0\}$.

For $m,d\in\mathbb{N}$, $0<\alpha<m$ and $\gamma>0$ put

$$k_n=k_n(m,\alpha,\gamma)=(m+3)2^{n\alpha/\gamma}\quad\text{for } n\in\mathbb{N}\cup\{0\},$$
$$U^{(m)}_n=\operatorname{span}\{N^{(\mathbf{r}_m)}_{\mathbf{j},\mathbf{h}_n}:\mathbf{j}\in\mathbb{Z}^d,\ \operatorname{supp}N^{(\mathbf{r}_m)}_{\mathbf{j},\mathbf{h}_n}\subset[-k_n,k_n]^d\},$$

and for $1\le p<\infty$, $f\in L^p(\mathbb{R}^d)$ set

$$D^{(m)}_{n,p}(f)=\inf\{\|f-g\|_p: g\in U^{(m)}_n\}.$$

It follows from the definitions of $U^{(m)}_n$ and $V^{(m)}_{n,p}$ that $U^{(m)}_n\subset V^{(m)}_{n,p}$, which implies $E^{(m)}_{n,p}\le D^{(m)}_{n,p}(f)$. The properties of the functions $N^{(\mathbf{r}_m)}_{\mathbf{j},\mathbf{h}_i}$ imply that $U^{(m)}_n\subset U^{(m)}_{n+1}$.

Define

$$W(p,d,m,\alpha,\gamma,C)=\{f\in L^p(\mathbb{R}^d):\|f\|_p\le C,\ D^{(m)}_{n,p}(f)\le C/2^{n\alpha}\text{ for } n\in\mathbb{N}\cup\{0\}\}.$$

LEMMA 3.10. *Let $d,m\in\mathbb{N}$, $0<\alpha<m$, $\gamma>0$ and $1\le p<\infty$ be given. There exists a constant $a>0$ such that for any $C>0$*

$$W(p,d,m,\alpha,\gamma,aC)\subset U(p,d,m,\alpha,\gamma,C).$$

Proof. First notice that for $f\in W(p,d,m,\alpha,\gamma,aC)$

$$\|f\|_p\le aC.\tag{3.11}$$

Now we estimate $\omega_{m,p}(f,\delta)$. For $\delta>1$ we have

$$\omega_{m,p}(f,\delta)\le 2^m\|f\|_p\le 2^m aC\delta^\alpha.\tag{3.12}$$

For $0 < \delta \le 1$ choose $k \in \mathbb{N} \cup \{0\}$ such that $1/2^{k+1} < \delta \le 1/2^k$; then it follows from Lemma 3.9 that

$$\omega_{m,p}(f,\delta) \le \omega_{m,p}\left(f, \frac{1}{2^k}\right) \le \frac{C_{m,d}}{2^{km}}\Big(\|f\|_p + \sum_{i=0}^{k} 2^{im} E_{i,p}^{(m)}(f)\Big) \tag{3.13}$$

$$\le \frac{C_{m,d}}{2^{km}}\Big(\|f\|_p + \sum_{i=0}^{k} 2^{im} D_{i,p}^{(m)}(f)\Big) \le 2^{m+1} C_{m,d} aC\delta^\alpha .$$

We also have to estimate $\Phi_p(f,\lambda)$. As $f \in W(p,d,m,\alpha,\gamma,aC)$, the definition of $U_n^{(m)}$ implies

$$\Phi_p(f,k_n) \le D_{n,p}^{(m)}(f) \le \frac{(m+3)^\gamma}{k_n^\gamma} aC .$$

For $0 < \lambda < k_0$ we have

$$\Phi_p(f,\lambda) \le \|f\|_p \le aC \le \frac{(m+3)^\gamma}{\lambda^\gamma} aC . \tag{3.14}$$

For $\lambda \ge k_0$ choose $n \in \mathbb{N} \cup \{0\}$ such that $k_n \le \lambda < k_{n+1}$; then

$$\Phi_p(f,\lambda) \le \Phi_p(f,k_n) \le \frac{(m+3)^\gamma}{k_n^\gamma} aC \tag{3.15}$$

$$= \frac{2^\alpha (m+3)^\gamma}{k_{n+1}^\gamma} aC \le \frac{2^\alpha (m+3)^\gamma}{\lambda^\gamma} aC .$$

It follows from (3.11)–(3.15) that it is enough to take

$$a = \min(2^{-m}, 2^{-(m+1)} C_{m,d}, 2^{-\alpha}(m+3)^\gamma) . \ ■$$

Now the ε-entropy of $W(p,d,m,\alpha,\gamma,C)$ will be estimated from below.

LEMMA 3.16. *Let $d, m \in \mathbb{N}$, $0 < \alpha < m$, $\gamma > 0$ and $1 \le p < \infty$ be given. There exist constants $k > 0$ and $\varepsilon_0 > 0$ such that*

$$k(C/\varepsilon)^{d(1/\alpha+1/\gamma)} \le \mathcal{H}_\varepsilon(W(p,d,m,\alpha,\gamma,C))$$

for every $C > 0$ and $0 < \varepsilon \le C\varepsilon_0$.

Proof. We will use Theorem 1.3.

First let $C = 1$. Write $d_n^{(m)} = \dim U_n^{(m)}$; then $(2^n k_n)^d \le d_n^{(m)} < \infty$ for $n \in \mathbb{N} \cup \{0\}$. Choose a sequence $\Phi = \{\varphi_1, \varphi_2, \ldots\}$ in $L^p(\mathbb{R}^d)$ so that $U_n^{(m)} = \operatorname{span}\{\varphi_1, \ldots, \varphi_{d_n^{(m)}}\}$ for each $n \in \mathbb{N} \cup \{0\}$. The sequence $\Delta = \{\delta_0, \delta_1, \ldots\}$ is defined as follows:

$$\delta_0 = \ldots = \delta_{d_0^{(m)}-1} = 1, \quad \delta_k = \frac{1}{2^{n\alpha}} \quad \text{for } d_n^{(m)} \le k < d_{n+1}^{(m)},\ n \ge 0 .$$

Then $W = W(p,d,m,\alpha,\gamma,1) = \mathcal{A}(\Delta,\Phi)$.

Let $c = 2$ in Theorem 1.3. For $0 < \varepsilon \le 1/2$ choose $j \in \mathbb{N}$ such that $2^{-(j-1)} < 2\varepsilon \le 2^{-(j-2)}$; then $N_{j-3} = d_{n_{j-3}}^{(m)}$, where $n_i = \max(m, \lceil i/\alpha \rceil)$.

(1.2) and (1.4) now imply

$$\mathcal{H}_\varepsilon(W) \ge \mathcal{C}_{2\varepsilon}(W) \ge N_{j-3} \ln 2 = d_{n_{j-3}}^{(m)} \ln 2$$

$$\ge (2^{n_{j-3}} k_{n_{j-3}})^d \ln 2 \ge 2^{-3d(1/\alpha+1/\gamma)} (m+3)^d (1/\varepsilon)^{d(1/\alpha+1/\gamma)} \ln 2 .$$

Take $k = 2^{-3d(1/\alpha+1/\gamma)}(m+3)^d \ln 2$, $\varepsilon_0 = 1/2$. As $W(p,d,m,\alpha,\gamma,C) = \{Cf : f \in W(p,d,m,\alpha,\gamma,1)\}$, it follows that $\mathcal{H}_\varepsilon(W(p,d,m,\alpha,\gamma,C)) = \mathcal{H}_{\varepsilon/C}(W(p,d,m,\alpha,\gamma,1))$, and so for $0 < \varepsilon \le C\varepsilon_0$

$$k(C/\varepsilon)^{d(1/\alpha+1/\gamma)} \le \mathcal{H}_\varepsilon(W(p,d,m,\alpha,\gamma,C)) . \ ■$$

Theorem 3.1 now follows from Lemmas 3.2, 3.10 and 3.16.

Now let $U^+(p,d,m,\alpha,\gamma,C) = \{f \in U(p,d,m,\alpha,\gamma,C) : f \ge 0\}$ for $m = 1, 2$. The following result can be proved similarly to Corollary 2.11:

COROLLARY 3.17. *Let $d \in \mathbb{N}$, $0 < \alpha < 1$, $\gamma > 0$ and $1 \le p < \infty$ be given. There exist constants $k_1, k_2 > 0$ and $\varepsilon_0 > 0$ such that*

$$k_1(C/\varepsilon)^{d(1/\alpha+1/\gamma)} \le \mathcal{H}_\varepsilon(U^+(p,d,m,\alpha,\gamma,C)) \le k_2(C/\varepsilon)^{d(1/\alpha+1/\gamma)}$$

for every $C > 0$ and $0 < \varepsilon \le C\varepsilon_0$.

Similar inequalities will be proved for $m = 2$ and $0 < \alpha < 2$. First some more notation will be introduced. Define

$$M_{\mathbf{j},\mathbf{l}}^{(\mathbf{z})} = (l_1 \ldots l_d)^{-1} N_{\mathbf{j},\mathbf{l}}^{(\mathbf{z})} \quad \text{for } \mathbf{j} \in \mathbb{Z}^d,\ \mathbf{z} \in \mathbb{N}^d,\ \mathbf{l} = (l_1, \ldots, l_d) \in \mathbb{R}^d,\ l_i > 0,$$

$$Q_{\mathbf{l}}^{(\mathbf{z})} f(x) = \sum_{\mathbf{j} \in \mathbb{Z}^d} \int_{\mathbb{R}^d} M_{\mathbf{j},\mathbf{l}}^{(\mathbf{z})}(t) f(t)\, dt \cdot N_{\mathbf{j},\mathbf{l}}^{(\mathbf{z})}(x) \quad \text{for } f \in L^p(\mathbb{R}^d),$$

$$\omega_{k,e_i,p}(f,\delta) = \sup_{|t| \le \delta} \|\Delta_{te_i}^k f\|_p \quad \text{for } f \in L^p(\mathbb{R}^d) .$$

where $e_i = (\delta_{1i}, \ldots, \delta_{di})$. It was proved in [3] that for any $f \in L^p(\mathbb{R}^d)$

$$\|f - Q_{\mathbf{l}}^{(\mathbf{z})} f\|_p \le 16 \sum_{i=1}^{d} \omega_{2,e_i,p}(f, z_i l_i) . \tag{3.18}$$

For $m, k \in \mathbb{N} \cup \{0\}$, $f \in C^k(\mathbb{R}^d) \cap L^p(\mathbb{R}^d)$ and $\delta > 0$ the following inequality holds:

$$\omega_{m+k,e_i,p}(f,\delta) \le \delta^k \omega_{m,e_i,p}(D^{ke_i} f, \delta) . \tag{3.19}$$

LEMMA 3.20. *Let $d \in \mathbb{N}$, $0 < \alpha < 2$, $\gamma > 0$ and $1 \le p < \infty$ be given. There exist constants $k > 0$ and $\varepsilon_0 > 0$ such that for every $C > 0$ and $0 < \varepsilon \le C\varepsilon_0$*

$$k(C/\varepsilon)^{d(1/\alpha+1/\gamma)} \le \mathcal{C}_\varepsilon(U^+(p,d,2,\alpha,\gamma,C)) .$$

Proof. Let $\mathbf{z} = (4,\dots,4) \in \mathbb{N}^d$ and recall that

$$U_n^{(2)} = \operatorname{span}\left\{ N_{\mathbf{j},\mathbf{h}_n}^{(\mathbf{z})} : \mathbf{h}_n = \left(\frac{1}{2^n},\dots,\frac{1}{2^n}\right) \in \mathbb{R}^d, \operatorname{supp} N_{\mathbf{j},\mathbf{h}_n}^{(\mathbf{z})} \subset [-k_n,k_n]^d \right\}.$$

For $f \in L^p(\mathbb{R}^d)$ set

$$D_{n,p}^{(+)}(f) = \inf\{\|f-g\|_p : g \in U_n^{(2)},\ g \ge 0\}$$

and define

$$W^+(p,d,\alpha,\gamma,C) = \{f \in L^p(\mathbb{R}^d) : \|f\|_p \le C,\ D_{n,p}^{(+)}(f) \le C/2^{n\alpha} \text{ for } n \in \mathbb{N}\cup\{0\}\}.$$

Notice that if $f \in W^+(p,d,\alpha,\gamma,C)$ then $f \ge 0$. As in the proof of Lemma 3.10, it is possible to find a constant $a > 0$ such that for every $C > 0$

$$W^+(p,d,\alpha,\gamma,aC) \subset U^+(p,d,2,\alpha,\gamma,C). \tag{3.21}$$

Now the ε-capacity of $W^+(p,d,\alpha,\gamma,C)$ will be estimated from below.

It follows from (3.6), (3.7), (3.18) and (3.19) that there exists a constant $A > 0$ such that for any $s \in \mathbb{N}\cup\{0\}$, $g \in U_s^{(2)}$ and $\mathbf{l} = (l_1,\dots,l_d)$ with $l_i > 0$

$$\|g - Q_{\mathbf{l}}^{(\mathbf{z})} g\|_p \le A\|g\|_p \frac{\sum_{i=1}^d l_i^2}{h_s^2}$$

(recall that $h_s = 1/2^s$). Let $s \in \mathbb{N}$ be chosen so that $2\cdot 2^{s(1+\alpha/\gamma)}(2^{\alpha/\gamma}-1) \ge 1$; then for $g \in U_s^{(2)}$ and $n > s$ we have $\operatorname{supp} Q_{\mathbf{h}_n}^{(\mathbf{z})} g \subset [-k_n,k_n]^d$. As $Q_{\mathbf{l}}^{(\mathbf{z})} g \ge 0$ for $g \ge 0$, the above inequality gives for $n > s$ and $g \in U_s^{(2)}$, $g \ge 0$,

$$D_{n,p}^{(+)}(g) \le \|g - Q_{\mathbf{h}_n}^{(\mathbf{z})} g\|_p \le Ad\frac{h_n^2}{h_s^2}\|g\|_p \le Ad\frac{h_n^\alpha}{h_s^\alpha}\|g\|_p.$$

As $D_{n,p}^{(+)}(g) \le \|g\|_p$ for $n \le s$, putting $C_0 = \min(1, 1/(Ad))$ we obtain for $g \in U_s^{(2)}$, $g \ge 0$, $\|g\|_p \le C_0 C h_s^\alpha$,

$$D_{n,p}^{(+)}(g) \le C h_n^\alpha \quad \text{for } n \in \mathbb{N}\cup\{0\}.$$

Set $E_s = \{\mathbf{j} \in \mathbb{Z}^d : \operatorname{supp} N_{\mathbf{j},\mathbf{h}_s}^{(\mathbf{z})} \subset [-k_s,k_s]^d\}$; the last inequality implies

$$W_{s,C} = \Big\{ g = \sum_{\mathbf{j}\in E_s} a_{\mathbf{j}} N_{\mathbf{j},\mathbf{h}_s}^{(\mathbf{z})} : a_{\mathbf{j}} \ge 0,\ \|g\|_p \le C_0 C h_s^\alpha \Big\} \subset W^+(p,d,\alpha,\gamma,C),$$

so for any $\varepsilon > 0$

$$\mathcal{C}_\varepsilon(W_{s,C}) \le \mathcal{C}_\varepsilon(W^+(p,d,\alpha,\gamma,C)). \tag{3.22}$$

Now the ε-capacity of $W_{s,C}$ will be estimated from below. It follows from (3.6) that for some $c_{\mathbf{z}} > 0$

$$c_{\mathbf{z}}^{-1} h_s^{d/p}\Big(\sum_{\mathbf{j}\in E_s} |a_{\mathbf{j}}|_p\Big)^{1/p} \le \Big\|\sum_{\mathbf{j}\in E_s} a_{\mathbf{j}} N_{\mathbf{j},\mathbf{h}_s}^{(\mathbf{z})}\Big\|_p \le c_{\mathbf{z}} h_s^{d/p}\Big(\sum_{\mathbf{j}\in E_s} |a_{\mathbf{j}}|_p\Big)^{1/p}.$$

Let

$$G = \Big\{ a = (a_{\mathbf{j}})_{\mathbf{j}\in E_s} : \|a\|_p = \Big(\sum_{\mathbf{j}\in E_s} |a_{\mathbf{j}}|_p\Big)^{1/p} \le \frac{C_0 C}{c_{\mathbf{z}}} h_s^{\alpha-d/p},\ a_{\mathbf{j}} \ge 0 \Big\}.$$

Then for any $a = (a_{\mathbf{j}})_{\mathbf{j}\in E_s} \in G$ we have $g_a = \sum_{\mathbf{j}\in E_s} a_{\mathbf{j}} N_{\mathbf{j},\mathbf{h}_s}^{(\mathbf{z})} \in W_{s,C}$, and for any $a,b \in G$, if $\|a-b\|_p \ge \eta_\varepsilon = \varepsilon c_{\mathbf{z}} h_s^{-d/p}$ then $\|g_a - g_b\|_p \ge \varepsilon$, so $\mathcal{C}_\varepsilon(W_{s,C}) \ge \mathcal{C}_{\eta_\varepsilon}(G)$. G is a subset of $\mathbb{R}^{d_s^{(2)}}$, where $d_s^{(2)} = \dim U_s^{(2)} = \# E_s$. It can be checked (using the method of the proof of Lemma 1 of [9]) that

$$\mathcal{M}_\eta(G) \ge \left(\frac{1}{2\eta}\right)^{d_s^{(2)}} \left(\frac{C_0 C}{c_{\mathbf{z}}} h_s^{\alpha-d/p}\right)^{d_s^{(2)}},$$

which implies

$$\mathcal{C}_\eta(G) \ge d_s^{(2)} \ln\left(\frac{C_0 C}{2c_{\mathbf{z}}}\frac{1}{\eta} h_s^{\alpha-d/p}\right).$$

Therefore

$$\mathcal{C}_\varepsilon(W_{s,C}) \ge \mathcal{C}_{\eta_\varepsilon}(G) \ge d_s^{(2)} \ln\left(\frac{C_0 C}{2c_{\mathbf{z}}^2}\frac{1}{\varepsilon} h_s^{\alpha}\right).$$

Let

$$s_{\varepsilon,C} = \left[\frac{1}{\alpha}\log_2\left(\frac{C_0}{2c_{\mathbf{z}}^2 e}\frac{C}{\varepsilon}\right)\right];$$

then

$$\frac{C_0 C}{2c_{\mathbf{z}}^2}\frac{1}{\varepsilon} h_{s_{\varepsilon,C}}^\alpha \ge e,$$

which implies

$$\mathcal{C}_\varepsilon(W_{s_{\varepsilon,C},C}) \ge d_{s_{\varepsilon,C}}^{(2)} \ge (2^{s_{\varepsilon,C}} k_{s_{\varepsilon,C}})^d \ge k(C/\varepsilon)^{d(1/\alpha+1/\gamma)}. \tag{3.23}$$

Choose $\varepsilon_0 > 0$ in such a way that

$$2\cdot 2^{s_{\varepsilon,C}(1+\alpha/\gamma)}(2^{\alpha/\gamma}-1) \ge 1$$

for every $0 < \varepsilon \le C\varepsilon_0$. The lemma now follows from (3.21)–(3.23). ∎

Lemmas 3.2, 3.20 and inequalities (1.2) imply

LEMMA 3.24. *Let $d \in \mathbb{N}$, $0 < \alpha < 2$, $\gamma > 0$ and $1 \le p < \infty$ be given. There exist constants $k_1, k_2 > 0$ and $\varepsilon_0 > 0$ such that*

$$k_1(C/\varepsilon)^{d(1/\alpha+1/\gamma)} \le \mathcal{H}_\varepsilon(U^+(p,d,2,\alpha,\gamma,C)) \le k_2(C/\varepsilon)^{d(1/\alpha+1/\gamma)}$$

for every $C > 0$ and $0 < \varepsilon \le C\varepsilon_0$.

LEMMA 3.25. *Let $d \in \mathbb{N}$, $m \in \{1,2\}$, $0 < \alpha < m$, $\gamma > 0$, $1 \le p < \infty$ and $M_1, M_2, M_3 > 0$ be given. Set*

$$M(p,d,m,\alpha,\gamma,M_1,M_2,M_3) = \{f \in L^p(\mathbb{R}^d) : f \ge 0,\ \|f\|_p = M_1,\ \omega_{m,p}(f,\delta) \le M_1 M_2 \delta^\alpha \text{ for } \delta > 0,\ \Phi_p(f,\lambda) \le M_1 M_3 \lambda^{-\gamma} \text{ for } \lambda > 0\}.$$

Assume that there exists $\varphi \in L^p(\mathbb{R}^d)$ such that $\varphi \ge 0$, $\|\varphi\|_p > 0$,

$$\omega_{m,p}(\varphi,\delta) \le a_1 \|\varphi\|_p \delta^\alpha \quad \text{for } \delta > 0,$$
$$\Phi_p(\varphi,\lambda) \le a_2 \|\varphi\|_p \lambda^{-\gamma} \quad \text{for } \lambda > 0$$

for some $0 < a_1 < M_2$, $0 < a_2 < M_3$. Then there exist $k_1, k_2, \varepsilon_0 > 0$ such that for any $0 < \varepsilon < \varepsilon_0$

$$k_1(1/\varepsilon)^{d(1/\alpha+1/\gamma)} \le \mathcal{H}_\varepsilon(M(p,d,m,\alpha,\gamma,M_1,M_2,M_3)) \le k_2(1/\varepsilon)^{d(1/\alpha+1/\gamma)}.$$

Proof. Define

$$\widetilde{M}(p,d,m,\alpha,\gamma,M_1,M_2,M_3) = \{f \in L^p(\mathbb{R}^d) : f \ge 0,\ \|f\|_p \le M_1,\ \omega_{m,p}(f,\delta) \le M_2 \|f\|_p \delta^\alpha \text{ for } \delta > 0,\ \Phi_p(f,\lambda) \le M_3 \|f\|_p \lambda^{-\gamma} \text{ for } \lambda > 0\},$$

and set $\varphi_{M_1} = M_1 \varphi/(2\|\varphi\|_p)$ and

$$b_1 = \frac{M_1}{2}, \quad b_2 = \frac{M_1(M_2 - a_1)}{2}, \quad b_3 = \frac{M_1(M_3 - a_2)}{2}, \quad b = \min(b_1, b_2, b_3).$$

Then $\{\varphi_{M_1} + f : f \in U^+(p,d,m,\alpha,\gamma,b)\} \subset \widetilde{M}(p,d,m,\alpha,\gamma,M_1,M_2,M_3)$ and the rest of the proof is similar to the proof of Corollary 2.12. ∎

4. ε-Entropy and nonparametric density estimation. A measurable function $F_n : \mathbb{R}^{nd} \times \mathbb{R}^d \to \mathbb{R}$ is called *a density estimator* if for any $x_1, \ldots, x_n \in \mathbb{R}^d$, $F_n(x_1, \ldots, x_n; \cdot) : \mathbb{R}^d \to \mathbb{R}$ is a probability density on $\mathbb{R}^d$.

For a given density $f : \mathbb{R}^d \to \mathbb{R}$ and density estimator $F_n : \mathbb{R}^{nd} \times \mathbb{R}^d \to \mathbb{R}$ define

$$E_f D(f, F_n) = \int_{\mathbb{R}^{nd}} D(f, F_n(x_1, \ldots, x_n; \cdot)) f(x_1) \ldots f(x_n)\, dx_1 \ldots dx_n,$$

where

$$D(f,g) = \int_{\mathbb{R}^d} |f(x) - g(x)|\, dx \quad \text{for } f, g \in L^1(\mathbb{R}^d).$$

Let $\mathcal{F}$ be a family of densities on $\mathbb{R}^d$. $\mathcal{F}$ is considered as a metric space with metric induced from $L^1(\mathbb{R}^d)$. Define

$$\mathcal{R}_{\mathcal{F}}(n) = \inf_{F_n} \sup_{f \in \mathcal{F}} E_f D(f, F_n),$$

where the infimum is taken over all density estimators F_n (based on a sample of size n) such that $F_n(x_1, \ldots, x_n; \cdot) \in \mathcal{F}$ for any $x_1, \ldots, x_n \in \mathbb{R}^d$. $\mathcal{R}_{\mathcal{F}}(n)$ is called the *minimax risk* for $\mathcal{F}$, corresponding to samples of size n and loss function D. Our aim is to establish some relations between $\mathcal{R}_{\mathcal{F}}(n)$ and $\mathcal{H}_\varepsilon(\mathcal{F})$.

THEOREM 4.1. *Let $\mathcal{F}$ be a family of densities on $\mathbb{R}^d$ which is totally bounded in the $L^1(\mathbb{R}^d)$ metric.*

(4.2) *If there exist $\varepsilon_0 > 0$, $C > 0$ and $\eta > 0$ such that*

$$\mathcal{H}_\varepsilon(\mathcal{F}) \le C(1/\varepsilon)^\eta \quad \text{for } 0 < \varepsilon \le \varepsilon_0,$$

then there exist $M > 0$ and $n_0 \in \mathbb{N}$ such that

$$\mathcal{R}_{\mathcal{F}}(n) \le M n^{-1/(2+\eta)} \quad \text{for } n \ge n_0.$$

(4.3) *If there exist $M > 0$, $\eta > 0$ and $n_0 \in \mathbb{N}$ such that*

$$M n^{-1/(2+\eta)} \le \mathcal{R}_{\mathcal{F}}(n) \quad \text{for } n \ge n_0,$$

then there exist $C > 0$ and $\varepsilon_0 > 0$ such that

$$C(1/\varepsilon)^\eta \le \mathcal{H}_\varepsilon(\mathcal{F}) \quad \text{for } 0 < \varepsilon \le \varepsilon_0.$$

The result analogous to (4.2), but under the additional assumption that the supports of all elements of $\mathcal{F}$ are contained in a compact set $S \subset \mathbb{R}^d$, was proved earlier by P. Groeneboom ([7]). The present proof is based on the main idea of the proof in [7]. The following lemma will be used ([7]):

LEMMA 4.4. *Let S be a compact subset of $\mathbb{R}^d$, and f_1, f_2 be two densities on $\mathbb{R}^d$ with supports contained in S. For $\xi > 0$ define*

$$B(f_i, \xi) = \Big\{h : \mathbb{R}^d \to \mathbb{R} : h \ge 0,\ \int_S h(x)dx = \int_{\mathbb{R}^d} h(x)\,dx = 1,\ D(f_i, h) \le \xi\Big\}.$$

For any $n \in \mathbb{N}$ there exists a measurable function $\varphi_{f_1,f_2} : S^n \to \{0,1\}$ such that

$$\text{(4.5)} \quad \sup_{h \in B(f_1,\xi)} \int_{S^n} \varphi_{f_1,f_2}(x_1, \ldots, x_n) \prod_{i=1}^n h(x_i)\, dx_1 \ldots dx_n + \sup_{h \in B(f_2,\xi)} \int_{S^n} (1 - \varphi_{f_1,f_2}(x_1, \ldots, x_n)) \prod_{i=1}^n h(x_i) dx_1 \ldots dx_n \le \exp\{-\tfrac{1}{8} n (D(f_1,f_2) - 2\xi)_+^2\}.$$

Proof of Theorem 4.1. The following notation will be used:

$$\mathbb{R}^d_\lambda = \{x \in \mathbb{R}^d : \|x\| \le \lambda\} \quad \text{for } \lambda > 0,$$

$$\Phi(f,\lambda) = \int_{\|x\|>\lambda} |f(x)|\, dx \quad \text{for } \lambda > 0,\ f \in L^1(\mathbb{R}^d),$$

and $g(\lambda)=\sup_{f\in\mathcal F}\Phi(f,\lambda)$. Notice that g is nonincreasing and nonnegative; we now show that $\lim_{\lambda\to\infty}g(\lambda)=0$. It is enough to check that for every $\xi>0$ there is a $\lambda_\xi>0$ such that $g(\lambda_\xi)\le\xi$. Let $\{f_1,\dots,f_k\}$ be a finite $\xi/2$-net of $\mathcal F$ (which exists since $\mathcal F$ is totally bounded); then for each i, $1\le i\le k$, there is a λ_i such that $\Phi(f_i,\lambda_i)\le\xi/2$. It is enough to put $\lambda_\xi=\max\{\lambda_1,\dots,\lambda_k\}$.

Now let $\lambda\ge\lambda_{1/2}$; then $\Phi(f,\lambda)\le g(\lambda)\le 1/2$ for $f\in\mathcal F$. Moreover, define

$$f_\lambda(x)=\begin{cases}0 & \text{for } \|x\|>\lambda,\\ \dfrac{1}{1-\Phi(f,\lambda)}f(x) & \text{for } \|x\|\le\lambda.\end{cases}$$

Clearly f_λ is a probability density on $\mathbb R^d$ with support contained in $\mathbb R^d_\lambda$. Note that for $f,h\in\mathcal F$ we have

$$D(f,h)\le 4g(\lambda)+D(f_\lambda,h_\lambda)\,, \tag{4.6}$$

$$D(f_\lambda,h_\lambda)\le 2g(\lambda)+D(f,h)\,. \tag{4.7}$$

Set $u(\varepsilon)=\mathcal H^{\mathcal F}_{\varepsilon-0}(\mathcal F)$; $u(\varepsilon)$ is a nonincreasing, left-continuous function and $u(\varepsilon)\ge\mathcal H^{\mathcal F}_\varepsilon(\mathcal F)$ for $\varepsilon>0$.

Now let $0<\varepsilon<1$ and $\lambda>0$ such that $g(\lambda)\le\varepsilon/4$ be given. We are going to construct a suitable density estimator.

Let $\{f_1,\dots,f_k\}$ be an ε-net of $\mathcal F$, consisting of elements of $\mathcal F$, such that $\ln k=\mathcal H^{\mathcal F}_\varepsilon(\mathcal F)\le u(\varepsilon)$. Write $h_i=(f_i)_\lambda$ and put $\xi=\varepsilon+2g(\lambda)$. For $i<j$ choose $\varphi_{h_i,h_j}:(\mathbb R^d_\lambda)^n\to\{0,1\}$ as in Lemma 4.4. Moreover, let $\varphi_{h_j,h_i}=1-\varphi_{h_i,h_j}$ for $i<j$. Clearly, in the latter case inequality (4.5) holds as well.

For $(x_1,\dots,x_n)\in(\mathbb R^d_\lambda)^n$ put

$$J_{h_i}(x_1,\dots,x_n)=\{h_j:\varphi_{h_i,h_j}(x_1,\dots,x_n)=1\}$$

and

$$L_{h_i}(x_1,\dots,x_n)=\begin{cases}\max\limits_{h_j\in J_{h_i}(x_1,\dots,x_n)}D(h_i,h_j) & \text{if } J_{h_i}(x_1,\dots,x_n)\ne\emptyset,\\ 0 & \text{if } J_{h_i}(x_1,\dots,x_n)=\emptyset.\end{cases}$$

Now define

$$\Theta(x_1,\dots,x_n;\cdot)=\begin{cases}f_1 & \text{if } \|x_l\|>\lambda \text{ for some } l,\ 1\le l\le n,\\ f_i & \text{where } L_{h_i}(x_1,\dots,x_n)=\min\limits_{h_1,\dots,h_k}L_{h_j}(x_1,\dots,x_n)\\ & \text{if } \|x_l\|\le\lambda \text{ for all } l,\ 1\le l\le n.\end{cases}$$

For $f\in\mathcal F$ and $i\in\mathbb N$ we have

$$\begin{aligned}P^n_f\{(x_1,\dots,x_n)\in(\mathbb R^d_\lambda)^n &: D(f,\Theta(x_1,\dots,x_n;\cdot))\ge(5+i)\varepsilon\}\\ &\le P^n_{f_\lambda}\{(x_1,\dots,x_n)\in(\mathbb R^d_\lambda)^n : D(f,\Theta(x_1,\dots,x_n;\cdot))\ge(5+i)\varepsilon\}\,,\end{aligned}\tag{4.8}$$

where P^n_f denotes the probability measure on $\mathbb R^{nd}$ with density $\prod_{i=1}^n f(x_i)$.

Choose r, $1\le r\le k$, such that $D(f,f_r)\le\varepsilon$. Then

$$D(f,\Theta(x_1,\dots,x_n;\cdot))\le D(f_r,\Theta(x_1,\dots,x_n;\cdot))+\varepsilon$$

and (by (4.6) and the choice of λ)

$$D(f_r,\Theta(x_1,\dots,x_n;\cdot))\le\varepsilon+D(h_r,\Theta(x_1,\dots,x_n;\cdot)_\lambda)\,,$$

so (4.8) implies

$$\begin{aligned}P^n_f\{(x_1,\dots,x_n)\in(\mathbb R^d_\lambda)^n &: D(f,\Theta(x_1,\dots,x_n;\cdot))\ge(5+i)\varepsilon\}\\ &\le P^n_{f_\lambda}\{(x_1,\dots,x_n)\in(\mathbb R^d_\lambda)^n : D(h_r,\Theta(x_1,\dots,x_n;\cdot)_\lambda)\ge(3+i)\varepsilon\}\,.\end{aligned}\tag{4.9}$$

It follows from the definition of Θ that if $D(h_r,\Theta(x_1,\dots,x_n;\cdot)_\lambda)\ge a$ then there is j, $1\le j\le k$, such that $\varphi_{h_r,h_j}(x_1,\dots,x_n)=1$ and $D(h_r,h_j)\ge a$. Therefore we obtain from (4.9)

$$\begin{aligned}P^n_f\{(x_1,\dots,x_n)\in(\mathbb R^d_\lambda)^n &: D(f,\Theta(x_1,\dots,x_n;\cdot))\ge(5+i)\varepsilon\}\\ &\le P^n_{f_\lambda}\{(x_1,\dots,x_n)\in(\mathbb R^d_\lambda)^n : \exists_{1\le s\le k}\,D(h_r,h_s)\ge(3+i)\varepsilon\\ &\qquad\qquad \text{and } \varphi_{h_r,h_s}(x_1,\dots,x_n)=1\}\,.\end{aligned}\tag{4.10}$$

Let

$$k_j=\#\{l:(3+j)\varepsilon\le D(h_l,h_r)<(4+j)\varepsilon\}\,;$$

then $\sum_{j\ge1}k_j\le k$.

As $D(f,f_r)\le\varepsilon$, (4.7) implies $D(f_\lambda,h_r)\le\varepsilon+2g(\lambda)=\xi$. Thus, (4.5) and (4.10) give

$$\begin{aligned}P^n_f\{(x_1,\dots,x_n)\in(\mathbb R^d_\lambda)^n &: D(f,\Theta(x_1,\dots,x_n;\cdot))\ge(5+i)\varepsilon\}\\ &\le\sum_{D(h_r,h_s)\ge(3+i)\varepsilon}\exp\{-\tfrac18 n(D(h_r,h_s)-2\xi)^2_+\}\le\sum_{j\ge i}k_j\exp\{-\tfrac18 nj^2\varepsilon^2\}\,.\end{aligned}$$

Now

$$\begin{aligned}E_fD(f,\Theta)&=\int\limits_{\mathbb R^{nd}}D(f,\Theta(x_1,\dots,x_n;\cdot))\prod_{i=1}^n f(x_i)\,dx_1\dots dx_n\\ &=\Big(\int\limits_{\mathbb R^{nd}\setminus(\mathbb R^d_\lambda)^n}+\int\limits_{(\mathbb R^d_\lambda)^n}\Big)D(f,\Theta(x_1,\dots,x_n;\cdot))\prod_{i=1}^n f(x_i)\,dx_1\dots dx_n\\ &\le 2ng(\lambda)+\varepsilon\Big(6+\sum_{j\ge1}jk_j\exp\{-\tfrac18 nj^2\varepsilon^2\}\Big)\,.\end{aligned}$$

For $\varepsilon\ge 2/\sqrt n$ we have $j\exp\{-\frac18 nj^2\varepsilon^2\}\le\exp\{-\frac18 n\varepsilon^2\}$, which together with the last inequality gives

$$\begin{aligned}E_fD(f,\Theta)&\le 2ng(\lambda)+\varepsilon\Big(6+\sum_{j\ge1}k_j\exp\{-\tfrac18 n\varepsilon^2\}\Big)\\ &\le 2ng(\lambda)+\varepsilon(6+\exp\{u(\varepsilon)-\tfrac18 n\varepsilon^2\})\,,\end{aligned}\tag{4.11}$$

and this implies

$$\mathcal{R}_{\mathcal{F}}(n) \le 2ng(\lambda) + \varepsilon(6 + \exp\{u(\varepsilon) - \tfrac{1}{8}n\varepsilon^2\}). \tag{4.12}$$

Since (4.12) holds for any $\lambda > 0$ such that $g(\lambda) \le \varepsilon/4$, and $\lim_{\lambda\to\infty} g(\lambda) = 0$, we get

$$\mathcal{R}_{\mathcal{F}}(n) \le \varepsilon(6 + \exp\{u(\varepsilon) - \tfrac{1}{8}n\varepsilon^2\}). \tag{4.13}$$

We recall that the inequalities (4.11)–(4.13) only hold for $2/\sqrt{n} \le \varepsilon \le 1$.

Now, let the assumption of (4.2) hold. It follows from (1.2) and the definition of u that

$$u(\varepsilon) \le \widetilde{C}(1/\varepsilon)^{\eta} \quad \text{for } 0 < \varepsilon \le \widetilde{\varepsilon}_0 = 2\varepsilon_0\,,\ \widetilde{C} = 2^{\eta}C\,. \tag{4.14}$$

If we put $\varepsilon_n = (8\widetilde{C}/n)^{1/(2+\eta)}$, then we find that there is an $n_0 \in \mathbb{N}$ such that $2/\sqrt{n} \le \varepsilon_n \le \min(1, 2\varepsilon_0)$ for $n > n_0$. Thus, from (4.13) and (4.14) we obtain

$$\mathcal{R}_{\mathcal{F}}(n) \le 7(8\widetilde{C})^{1/(2+\eta)}(1/n)^{1/(2+\eta)} \quad \text{for } n > n_0\,.$$

Now, let the assumption of (4.3) hold. This assumption and (4.13) imply that $\mathcal{F}$ contains infinitely many distinct elements, so that $\lim_{\varepsilon\to 0} u(\varepsilon) = \infty$.

Let $\psi(\varepsilon) = 8\varepsilon^{-2}u(\varepsilon)$. Then ψ is a left-continuous, strictly decreasing function on some interval $(0,\zeta)$ and $\lim_{\varepsilon\to 0}\psi(\varepsilon) = \infty$. For $m \in \mathbb{N}$, $m \ge \psi(\zeta)$, let

$$\varepsilon_m = \sup\{\varepsilon : \psi(\varepsilon) \ge m\}\,.$$

Then $\psi(\varepsilon_m + 0) \le m \le \psi(\varepsilon_m) = \psi(\varepsilon_m - 0)$. For convenience assume that ε_m is defined for all $m \in \mathbb{N}$. Note that $\varepsilon_{m+1} \le \varepsilon_m$ and $\lim_{m\to\infty}\varepsilon_m = 0$. Now, define

$$\begin{aligned}
\xi_1 &= \varepsilon_1 = \ldots = \varepsilon_{k_1 - 1} > \varepsilon_{k_1}\,,\\
\xi_2 &= \varepsilon_{k_1} = \ldots = \varepsilon_{k_2 - 1} > \varepsilon_{k_2}\,, \cdots\\
\xi_{n+1} &= \varepsilon_{k_n} = \ldots = \varepsilon_{k_{n+1}-1} > \varepsilon_{k_{n+1}}\,.
\end{aligned}$$

Note that $\lim_{n\to\infty} k_n = \infty$, $\lim_{n\to\infty}\xi_n = 0$ and $k_n - 1 \le \psi(\varepsilon_{k_n - 1}) = \psi(\xi_n) < k_n$ (if $k_n \le \psi(\xi_n)$, then $\varepsilon_{k_n} \ge \xi_n$, but the choice of k_n implies $\xi_n > \varepsilon_{k_n}$).

For n large enough we have $u(2/\sqrt{k_n}) \ge 1/2$ and $0 < 2/\sqrt{k_n} < \zeta$, hence

$$\psi\Big(\frac{2}{\sqrt{k_n}}\Big) = 8\frac{k_n}{4}u\Big(\frac{2}{\sqrt{k_n}}\Big) \ge k_n > \psi(\xi_n)\,,$$

which implies $\xi_n > 2/\sqrt{k_n}$. Note that $u(\varepsilon) - \frac{1}{8}k_n\varepsilon^2 < 0$ for $\varepsilon \ge \xi_n$. Therefore (4.13) implies for n large enough

$$\mathcal{R}_{\mathcal{F}}(k_n) \le 7\xi_n\,. \tag{4.15}$$

Our assumption implies that $M(1/n)^{1/(2+\eta)} \le \mathcal{R}_{\mathcal{F}}(n)$ for $n > n_0$. Choose $s \in \mathbb{N}$ such that for any $n \ge s$ we have $k_n > n_0$, $\psi(\xi_n) \ge 1$ and (4.15) holds.

Now $k_n \le 2\psi(\xi_n)$ for $n \ge s$ and

$$M\Big(\frac{1}{2\psi(\xi_n)}\Big)^{1/(2+\eta)} \le M\Big(\frac{1}{k_n}\Big)^{1/(2+\eta)} \le \mathcal{R}_{\mathcal{F}}(k_n) \le 7\xi_n\,.$$

The definition of ψ now implies

$$\widetilde{M}\xi_n^{-\eta} \le u(\xi_n)\,, \tag{4.16}$$

where $\widetilde{M} = \frac{1}{16}(M/7)^{2+\eta}$.

Now for $x > \psi(\zeta)$ define

$$\widetilde{\psi}^{-1}(x) = \begin{cases} y & \text{if } \psi(y) = x\,,\\ y & \text{if } \psi(y+0) \le x < \psi(y)\,. \end{cases}$$

Then $\widetilde{\psi}^{-1}$ is nonincreasing and $\widetilde{\psi}^{-1}(k_n) = \xi_{n+1}$. As u is nonincreasing, we have

$$2\psi(\varepsilon) = 16\varepsilon^{-2}u(\varepsilon) \le 8\Big(\frac{\varepsilon}{\sqrt{2}}\Big)^{-2}u\Big(\frac{\varepsilon}{\sqrt{2}}\Big) = \psi\Big(\frac{\varepsilon}{\sqrt{2}}\Big)\,,$$

and the inequality $k_n - 1 \le \psi(\xi_n) < k_n \le \psi(\xi_{n+1})$ implies

$$\begin{aligned}
\xi_{n+1} &= \widetilde{\psi}^{-1}(k_n) \ge \widetilde{\psi}^{-1}(2(k_n - 1))\\
&\ge \widetilde{\psi}^{-1}(2\psi(\xi_n)) \ge \widetilde{\psi}^{-1}(\psi(\xi_n/\sqrt{2})) = \xi_n/\sqrt{2}\,.
\end{aligned}$$

Now, let $0 < \varepsilon \le \xi_s$; choose $n \ge s$ such that $\xi_{n+1} < \varepsilon \le \xi_n$; then the last inequality and (4.16) imply

$$u(\varepsilon) \ge \widetilde{M}2^{-\eta/2}(1/\varepsilon)^{\eta}\,. \tag{4.17}$$

As $u(\varepsilon) = \mathcal{H}^{\mathcal{F}}_{\varepsilon-0}(\mathcal{F})$ and $\mathcal{H}^{\mathcal{F}}_{\varepsilon-0}(\mathcal{F})$ is nonincreasing (so that the set of its discontinuity points is countable) it follows from (4.17) that

$$\mathcal{H}^{\mathcal{F}}_{\varepsilon}(\mathcal{F}) \ge \widetilde{M}2^{-\eta/2}(1/\varepsilon)^{\eta} \quad \text{for } 0 < \varepsilon \le \xi_s\,.$$

The last inequality and (1.2) give (4.3). ∎

EXAMPLES.

(4.18) Let $d, m \in \mathbb{N}$, $0 < \alpha < m$, $\gamma > 0$, $C_1, C_2 > 0$ be given parameters, and

$$\begin{aligned}
\mathcal{F} = \{f \in L^1(\mathbb{R}^d) : f \ge 0,\ \|f\|_1 = 1,\ \omega_{m,1}(f,\delta) \le C_1\delta^{\alpha} \text{ for } \delta > 0,\\
\Phi_1(f,\lambda) \le C_2\lambda^{-\gamma} \text{ for } \lambda > 0\}\,.
\end{aligned}$$

It follows from Lemma 3.2 that there exist $C > 0$ and $\varepsilon_0 > 0$ such that

$$\mathcal{H}_{\varepsilon}(\mathcal{F}) \le C(1/\varepsilon)^{d(1/\alpha + 1/\gamma)} \quad \text{for } 0 < \varepsilon \le \varepsilon_0\,.$$

(4.2) now implies the existence of $M > 0$ and $n_0 \in \mathbb{N}$ such that

$$\mathcal{R}_{\mathcal{F}}(n) \le M(1/n)^{\alpha\gamma/(2\alpha\gamma + d(\alpha+\gamma))} \quad \text{for } n \ge n_0\,.$$

(4.19) Let $d, m \in \mathbb{N}$, $0 < \alpha < m$, $C > 0$ be given parameters, and

$$\mathcal{F} = \{f \in L^1(\mathbb{I}^d) : f \geq 0,\ \|f\|_1 = 1,\ \omega_{m,1}(f,\delta) \leq C\delta^\alpha \text{ for } 0 < \delta \leq 1/m\}$$

where the L^1-norm and the modulus of smoothness are taken in $L^1(\mathbb{I}^d)$. $\mathcal{F}$ can also be considered as a family of densities on $\mathbb{R}^d$ whose supports are contained in $\mathbb{I}^d$ (while the modulus of smoothness is still taken in $L^1(\mathbb{I}^d)$). It will be proved that there exist $A_1, A_2, M_1, M_2 > 0$, $\varepsilon_0 > 0$ and $n_0 \in \mathbb{N}$ such that

$$A_1(1/\varepsilon)^{d/\alpha} \leq \mathcal{H}_\varepsilon(\mathcal{F}) \leq A_2(1/\varepsilon)^{d/\alpha} \quad \text{for } 0 < \varepsilon \leq \varepsilon_0\,, \tag{4.20}$$

$$M_1(1/n)^{\alpha/(2\alpha+d)} \leq \mathcal{R}_{\mathcal{F}}(n) \leq M_2(1/n)^{\alpha/(2\alpha+d)} \quad \text{for } n \geq n_0\,. \tag{4.21}$$

(Note that (4.20) resembles Lemma 2.12, which was proved for all $1 \leq p < \infty$, but for $0 < \alpha < 1$ only.)

The existence of A_2 follows from Corollary 2.10, and the existence of M_2 is a consequence of (4.2). Once the existence of M_1 has been proved, the existence of A_1 will follow from (4.3). So it remains to prove the left-hand side inequality in (4.21). In order to do this, we use Assouad's Lemma (its proof can be found for example in [6]):

THEOREM 4.22 (Assouad's Lemma). *For $r \in \mathbb{N}$ set $Z_r = \{-1, 1\}^r$; for $z \in Z_r$ and $1 \leq i \leq r$ define*

$$z_{i+} = (z_1, \ldots, z_{i-1}, 1, z_{i+1}, \ldots, z_r)\,,$$
$$z_{i-} = (z_1, \ldots, z_{i-1}, -1, z_{i+1}, \ldots, z_r)\,.$$

Let $\{f_z : z \in Z_r\}$ be a family of 2^r distinct densities on $\mathbb{R}^d$. Assume that there exists a partition of $\mathbb{R}^d$ into measurable sets $A_1, \ldots, A_r$ such that for any $z \in Z_r$ and $1 \leq i \leq r$

$$\int_{A_i} |f_{z_{i+}}(x) - f_{z_{i-}}(x)|\,dx \geq p > 0\,.$$

Moreover, let

$$\int_{\mathbb{R}^d} \sqrt{f_{z_{i+}}(x) f_{z_{i-}}(x)}\,dx \geq q > 0\,.$$

Then for any density estimator F_n

$$\sup_{z \in Z_r} E_f D(f_z, F_n) \geq \frac{rp}{2} \max\left(1 - \sqrt{2 - 2q^n}, \frac{1}{2}q^{2n}\right).$$

Now we prove the left-hand side inequality in (4.21). For a given $m \in \mathbb{N}$ take $k \in \mathbb{N}$, $k \geq m+2$; recall that $N^{(k)}(x) = k[0, \ldots, k; (\cdot - x)_+^{k-1}]$. Choose $h > 0$ and $r \in \mathbb{N}$ such that $2khr = 1$ and put

$$g_{r,k}(t_1, \ldots, t_d) = N^{(k)}(t_1/h) \ldots N^{(k)}(t_d/h)\,.$$

Then $\operatorname{supp} g_{r,k} = [0, kh]^d$; for $i_1, \ldots, i_d \in \{0, 1, \ldots, r-1\}$ define

$$\begin{aligned} A_{(i_1,\ldots,i_d)} &= [2khi_1, 2kh(i_1 + \tfrac{1}{2})] \times \ldots \times [2khi_d, 2kh(i_d + \tfrac{1}{2})] \\ g_{(i_1,\ldots,i_d)} &= g_{r,k}(t_1 - 2khi_1, \ldots, t_d - 2khi_d) \\ &\quad - g_{r,k}(t_1 - 2kh(i_1 + \tfrac{1}{2}), \ldots, t_d - 2kh(i_d + \tfrac{1}{2}))\,. \end{aligned}$$

For each $i \in \{1, 2, \ldots, r^d\}$ choose a multi-index $u_i \in \{0, 1, \ldots, r-1\}^d$ so that $u_{i_1} \neq u_{i_2}$ for $i_1 \neq i_2$.

Let $a \in \mathbb{R}$, $0 < a < 1$; for each $z \in Z_{r^d}$, $z = (z_1, \ldots, z_{r^d})$, set

$$f_z(\mathbf{t}) = \begin{cases} 1 + a\sum_{i=1}^{r^d} z_i g_{u_i}(\mathbf{t}) & \text{if } \mathbf{t} \in \mathbb{I}^d\,, \\ 0 & \text{if } \mathbf{t} \in \mathbb{R}^d \setminus \mathbb{I}^d\,. \end{cases}$$

First note that $\int_{\mathbb{I}^d} f_z(x)\,dx = \int_{\mathbb{R}^d} f_z(x)\,dx = 1$ and $f_z \geq 0$ for any $z \in Z_{r^d}$. The parameters a and r will be chosen in such a way that f_z restricted to $\mathbb{I}^d$ is an element of $\mathcal{F}$.

The properties of the functions $N_{i,h}^{(k)}$ mentioned in Section 3 imply that there exists a constant $C_m > 0$ such that for every $z \in Z_{r^d}$, a, r and h

$$\omega_{m,1}(f_z, \delta) \leq a\sum_{i=1}^{r^d} \omega_{m,1}(g_{u_i}, \delta) \leq C_m a \frac{1}{(2k)^d} \min((\delta/h)^m, 1)\,.$$

Therefore, for given $0 < \alpha < m$ and $k \geq m+2$ it is possible to find a constant $C_{m,k,\alpha}$, independent of a and r, such that $\omega_{m,1}(f_z, \delta) \leq C_{m,k,\alpha} a r^\alpha \delta^\alpha$ for any $z \in Z_{r^d}$. Note that

$$\int_{A_{u_i}} |f_{z_{i+}}(x) - f_{z_{i-}}(x)|\,dx = 4ah^d = 4a\frac{1}{(2kr)^d} = p\,,$$

$$\begin{aligned} \int_{\mathbb{R}^d} \sqrt{f_{z_{i+}}(x) f_{z_{i-}}(x)}\,dx &= \int_{\mathbb{I}^d} \sqrt{f_{z_{i+}}(x) f_{z_{i-}}(x)}\,dx \\ &\geq 1 - 2(kh)^d + 2(kh)^d\sqrt{1 - a^2} \geq 1 - 2(kh)^d a^2 = q\,. \end{aligned}$$

Take $r > (C/C_{m,k,\alpha})^{1/\alpha}$, $a = (C/C_{m,k,\alpha})r^{-\alpha}$; then $f_z \in \mathcal{F}$ for $z \in Z_{r^d}$. It follows from Theorem 4.22 that for any density estimator F_n

$$\begin{aligned} \sup_{f \in \mathcal{F}} E_f D(f, F_n) &\geq \sup_{z \in Z_{r^d}} E_{f_z} D(f_z, F_n) \geq \tfrac{1}{2}pr^d(1 - \sqrt{2(1 - q^n)}) \\ &\geq \tfrac{1}{2}pr^d(1 - \sqrt{2n(1 - q)}) = s_1 r^{-\alpha}(1 - \sqrt{s_2 n r^{-(2\alpha+d)}})\,. \end{aligned} \tag{4.23}$$

where s_1, s_2 are some positive constants, independent of n and r. Take

$$r_n = [(4s_2 n)^{1/(2\alpha+d)}] + 1\,.$$

Then

$$\sqrt{s_2 n r_n^{-(2\alpha+d)}} \leq \frac{1}{2}\,;$$

also $r_n > (C/C_{m,k,\alpha})^{1/\alpha}$ and $r_n \le 2(4s_2n)^{1/(2\alpha+d)}$ for n large enough. This and (4.23) imply the left-hand side inequality in (4.21).

References

[1] C. de Boor, *Splines as linear combinations of B-splines*, in: Approximation Theory II, G. G. Lorentz *et al.* (eds.), Academic Press, New York 1976, 1–47.

[2] Z. Ciesielski, *Properties of the orthonormal Franklin system, II*, Studia Math. 27 (1966), 289–323.

[3] —, *Asymptotic nonparametric spline density estimation in several variables*, in: Internat. Ser. Numer. Math. 94, Birkhäuser, Basel 1990, 25–53.

[4] Z. Ciesielski and T. Figiel, *Spline approximation and Besov spaces on compact manifolds*, Studia Math. 75 (1982), 13–36.

[5] —, —, *Spline bases in classical function spaces on compact C^∞ manifolds, Part II*, ibid. 76 (1983), 95–136.

[6] L. Devroye and L. Györfi, *Nonparametric Density Estimation. The L_1 View*, Wiley, New York 1985.

[7] P. Groeneboom, *Some current developments in density estimation*, in: Mathematics and Computer Science, Proceedings of the CWI symposium, November 1983, J. W. de Bakker, M. Hazewinkel and J. K. Lenstra (eds.), North-Holland, 1986, 163–192.

[8] A. N. Kolmogorov and V. M. Tikhomirov, *ε-Entropy and ε-capacity of sets in function spaces*, Uspekhi Mat. Nauk 14 (2) (1959), 3–86 (in Russian); English transl.: Amer. Math. Soc. Transl. (2) 17 (1961), 277–364.

[9] G. G. Lorentz, *Metric entropy and approximation*, Bull. Amer. Math. Soc. 72 (1966), 903–937.

[10] I. J. Schoenberg, *Cardinal interpolation and spline functions*, J. Approx. Theory 2 (1969), 167–206.

INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES
ABRAHAMA 18
81-825 SOPOT, POLAND

Received February 26, 1992 (2905)

INFORMATION FOR AUTHORS

Manuscripts should be typed on one side only, with double or triple spacing and wide margins, and submitted in duplicate, including the original typewritten copy. Poor quality copies will not be accepted.

An **abstract** of not more than 200 words and the AMS Mathematics Subject Classification are required.

Formulas must be typewritten. A complete **list of all handwritten symbols** with indications for the printer should be enclosed. Special typefaces should be indicated according to the following code: **script letters**—by encircling the typed Roman letter in black, **German letters**—by typing the Roman equivalent and underlining in green, **boldface letters**—by straight black underlining.

Figures should be drawn accurately on separate sheets, preferably twice the size in which they are required to appear. The author should indicate in the margin of the manuscript where figures are to be inserted.

References should be arranged in alphabetical order, typed with double spacing, and styled and punctuated according to the examples given below. Abbreviations of journal names should follow Mathematical Reviews. Titles of papers in Russian should be translated into English.

Examples:

[6] D. Beck, *Introduction to Dynamical Systems*, Vol. 2, Progr. Math. 54, Birkhäuser, Basel 1978.

[7] R. Hill and A. James, *An index formula*, J. Differential Equations 15 (1982), 197–211.

[8] J. Kowalski, *Some remarks on $J(X)$*, in: Algebra and Analysis, Proc. Conf. Edmonton 1973, E. Brook (ed.), Lecture Notes in Math. 867, Springer, Berlin 1974, 115–124.

[Nov] A. S. Novikov, *An existence theorem for planar graphs*, preprint, Moscow University, 1980 (in Russian).

Authors' **affiliation** should be given at the end of the manuscript.

Authors receive only page **proofs** (one copy). If the proofs are not returned promptly, the article will be printed in a later issue.

Authors receive 50 **reprints** of their articles.

The publisher would like to encourage submission of manuscripts written in TEX. **On acceptance of the paper**, authors should send discs (preferably PC) plus relevant details to the Editorial Committee, or transmit the paper by electronic mail to: impanw@plwatu21.bitnet.