Some lower bounds for optimal sampling recovery of functions with mixed smoothness

A. Gasnikov and V. Temlyakov

Abstract

Recently, there was a substantial progress in the problem of sampling recovery on function classes with mixed smoothness. Mostly, it has been done by proving new and sometimes optimal upper bounds for both linear sampling recovery and for nonlinear sampling recovery. In this paper we address the problem of lower bounds for the optimal rates of nonlinear sampling recovery. In the case of linear recovery one can use the well developed theory of estimating the Kolmogorov and linear widths for establishing some lower bounds for the optimal rates. In the case of nonlinear recovery we cannot use the above approach. It seems like the only technique, which is available now, is based on some simple observations. We demonstrate how these observations can be used.

1 Introduction

Recently, there was a substantial progress in the problem of sampling recovery on function classes with mixed smoothness. This paper is a followup of the recent papers [7], [20], [21], and [12]. In this paper we address the problem of lower bounds for the optimal rates of sampling recovery. The problem of sampling recovery on function classes with mixed smoothness has a long history with first results going back to the 1963 (see [11]). In many cases this problem is still open. We refer the reader to the books [5] and [18] for the corresponding historical discussion.

In this section we describe the problem setting and present some known upper bounds. In Sections 3 – 5 we obtain some new results. Section 6 contains a discussion. In this paper we admit the following convenient and standard notation agreement. We use C, C' and c, c' to denote various positive constants. Their arguments indicate the parameters, which they may depend on. Normally, these constants do not depend on a function f and running parameters m, v, u. We use the following symbols for brevity. For two nonnegative sequences $a = \{a_n\}_{n=1}^\infty$ and $b = \{b_n\}_{n=1}^\infty$ the relation $a_n \ll b_n$ means that there is a number $C(a,b)$ such that for all n we have $a_n \le C(a,b)b_n$. Relation $a_n \gg b_n$ means that $b_n \ll a_n$ and $a_n \asymp b_n$ means that $a_n \ll b_n$ and $a_n \gg b_n$. For a real number x denote $[x]$ the integer part of x, $\lceil x \rceil$ – the smallest integer, which is greater than or equal to x.

We study the multivariate periodic functions defined on $\mathbb{T}^d := [0,2\pi]^d$. Denote for $1 \le p < \infty$

$$\|f\|_p := \left((2\pi)^{-d}\int_{\mathbb{T}^d}|f(\mathbf{x})|^p d\mathbf{x}\right)^{1/p}, \qquad \mathbf{x} = (x_1,\dots,x_d)$$

and for $p = \infty$

$$\|f\|_\infty := \sup_{\mathbf{x}\in\mathbb{T}^d}|f(\mathbf{x})|.$$

We begin with the definition of classes $\mathbf{W}^r_q$.

Definition 1.1. *In the univariate case, for $r > 0$, let*

$$F_r(x) := 1 + 2\sum_{k=1}^\infty k^{-r}\cos(kx - r\pi/2) \tag{1.1}$$

and in the multivariate case, for $\mathbf{x} = (x_1,\dots,x_d)$, let

$$F_r(\mathbf{x}) := \prod_{j=1}^d F_r(x_j).$$

Denote

$$\mathbf{W}^r_q := \{f : f = \varphi * F_r, \quad \|\varphi\|_q \le 1\},$$

where

$$(\varphi * F_r)(\mathbf{x}) := (2\pi)^{-d}\int_{\mathbb{T}^d}\varphi(\mathbf{y})F_r(\mathbf{x}-\mathbf{y})d\mathbf{y}.$$

The classes $\mathbf{W}^r_q$ are classical classes of functions with *dominated mixed derivative* (Sobolev-type classes of functions with mixed smoothness). The reader can find results on approximation properties of these classes in the books [18] and [5].

We now proceed to the definition of the classes $\mathbf{H}^r_p$, which is based on the mixed differences. In this paper we obtain new results for these classes.

Definition 1.2. *Let* $\mathbf{t} = (t_1, \dots, t_d)$ *and* $\Delta^l_{\mathbf{t}} f(\mathbf{x})$ *be the mixed* l*-th difference with step* t_j *in the variable* x_j*, that is*

$$\Delta^l_{\mathbf{t}} f(\mathbf{x}) := \Delta^l_{t_d,d} \dots \Delta^l_{t_1,1} f(x_1, \dots, x_d).$$

Let e *be a subset of natural numbers in* $[1, d]$*. We denote*

$$\Delta^l_{\mathbf{t}}(e) := \prod_{j\in e} \Delta^l_{t_j,j}, \qquad \Delta^l_{\mathbf{t}}(\varnothing) := Id \text{ -- identity operator.}$$

We define the class $\mathbf{H}^r_{p,l} B$*,* $l > r$*, as the set of* $f \in L_p$ *such that for any* e

$$\left\|\Delta^l_{\mathbf{t}}(e) f(\mathbf{x})\right\|_p \le B \prod_{j\in e} |t_j|^r. \tag{1.2}$$

In the case $B = 1$ *we omit it. It is known (see, for instance, [18], p.137) that the classes* $\mathbf{H}^r_{p,l}$ *with different* $l > r$ *are equivalent. So, for convenience we fix one* $l = [r] + 1$ *and omit* l *from the notation.*

It is well known that in the univariate case ($d = 1$) the approximation properties of the above $\mathbf{W}$ and $\mathbf{H}$ classes are similar. It is also well known that in the multivariate case ($d \ge 2$) asymptotic characteristics (for instance, Kolmogorov widths, entropy numbers, best hyperbolic cross trigonometric approximations and others) have different rate of decay in the majority of cases. Recently, a new scale of classes has been introduced and studied. It turns out that this scale is convenient for simultaneous analysis of optimal sampling recovery of both the $\mathbf{W}$ and the $\mathbf{H}$ classes. We give a corresponding definitions now. Let $\mathbf{s} = (s_1, \dots, s_d)$ be a vector whose coordinates are nonnegative integers

$$\rho(\mathbf{s}) := \left\{\mathbf{k} \in \mathbb{Z}^d : [2^{s_j-1}] \le |k_j| < 2^{s_j}, \qquad j = 1, \dots, d\right\},$$

where $[a]$ means the integer part of a real number a. For $f \in L_1(\mathbb{T}^d)$

$$\delta_{\mathbf{s}}(f, \mathbf{x}) := \sum_{\mathbf{k}\in\rho(\mathbf{s})} \hat{f}(\mathbf{k}) e^{i(\mathbf{k},\mathbf{x})}, \quad \hat{f}(\mathbf{k}) := (2\pi)^{-d} \int_{\mathbb{T}^d} f(\mathbf{x}) e^{-i(\mathbf{k},\mathbf{x})} d\mathbf{x}.$$

Definition 1.3. *Consider functions with absolutely convergent Fourier series. For such functions define the Wiener norm (the A-norm or the ℓ_1-norm)*

$$\|f\|_A := \sum_{\mathbf{k}} |\hat{f}(\mathbf{k})|.$$

The following classes, which are convenient in studying sparse approximation with respect to the trigonometric system, were introduced and studied in [17]. Define for $f \in L_1(\mathbb{T}^d)$

$$f_j := \sum_{\|\mathbf{s}\|_1 = j} \delta_{\mathbf{s}}(f), \quad j \in \mathbb{N}_0, \quad \mathbb{N}_0 := \mathbb{N} \cup \{0\}.$$

For parameters $a \in \mathbb{R}_+$, $b \in \mathbb{R}$ define the class

$$\mathbf{W}^{a,b}_A := \{f : \|f_j\|_A \le 2^{-aj}(\bar{j})^{(d-1)b}, \quad \bar{j} := \max(j,1), \quad j \in \mathbb{N}\}.$$

The following embedding result follows easily from known results. We give a detailed proof of it in Section 2.

Proposition 1.1. *We have for $r > 1/q$*

$$\mathbf{W}^r_q \hookrightarrow \mathbf{W}^{a,b}_A \quad \text{with} \quad a = r - 1/q,\, b = 1 - 1/q, \quad 1 < q \le 2; \tag{1.3}$$

$$\mathbf{H}^r_q \hookrightarrow \mathbf{W}^{a,b}_A \quad \text{with} \quad a = r - 1/q,\, b = 1, \quad 1 \le q \le 2. \tag{1.4}$$

We give a very brief history of the recent development of the sampling recovery on these classes. We refer the reader to the books [5] and [18] for the previous results. In this paper we study the following characteristic of the optimal sampling recovery. Let Ω be a compact subset of $\mathbb{R}^d$ with the probability measure μ on it. For a function class $W \subset \mathcal{C}(\Omega)$, we define (see [23])

$$\varrho^o_m(W, L_p) := \inf_{\xi} \inf_{\mathcal{M}} \sup_{f \in W} \|f - \mathcal{M}(f(\xi^1), \dots, f(\xi^m))\|_p,$$

where $\mathcal{M}$ ranges over all mappings $\mathcal{M} : \mathbb{C}^m \to L_p(\Omega, \mu)$ and ξ ranges over all subsets $\{\xi^1, \cdots, \xi^m\}$ of m points in Ω. Here, we use the index o to mean optimality. Clearly, the above characteristic is a characteristic of nonlinear recovery. For a discussion of the sampling recovery by linear methods see Section 6.

The authors of [6] (see Corollary 4.16 in v3) proved the following bound for $1 < q < 2$, $r > 1/q$ and $m \ge c(r,d,q)v(\log(2v))^3$, $v \in \mathbb{N}$,

$$\varrho_m^o(\mathbf{W}_q^r, L_2(\mathbb{T}^d)) \le C(r,d,q)v^{-r+1/q-1/2}(\log(2v))^{(d-1)(r+1-2/q)+1/2}. \tag{1.5}$$

The authors of [3] proved the following bound

$$\varrho_m^o(\mathbf{W}_q^r, L_2(\mathbb{T}^d)) \le C'(r,d,q)v^{-r+1/q-1/2}(\log(2v))^{(d-1)(r+1-2/q)} \tag{1.6}$$

provided that

$$m \ge c'(r,d,q)v(\log(2v))^3. \tag{1.7}$$

In the above mentioned results the sampling recovery in the L_2 norm has been studied. The technique, which was used in the proofs of the bounds (1.5) and (1.6) is heavily based on the fact that we approximate in the L_2 norm. The following upper bound was proved in [7].

Theorem 1.1 ([7]). *Let $1 < q \le 2 \le p < \infty$ and $r > 1/q$. There exist two constants $c = c(r,d,p,q)$ and $C = C(r,d,p,q)$ such that we have the bound*

$$\varrho_m^o(\mathbf{W}_q^r, L_p(\mathbb{T}^d)) \le Cv^{-r+1/q-1/p}(\log(2v))^{(d-1)(r+1-2/q)} \tag{1.8}$$

for any $v \in \mathbb{N}$ and any m satisfying

$$m \ge cv(\log(2v))^3.$$

Thus, Theorem 1.1 extends the previously known upper bound for $p = 2$ to the case $p \in [2,\infty)$.

In [7] Theorem 1.1 was derived from the embedding (1.3) and the following result for the $\mathbf{W}_A^{a,b}$ classes (see [7], Theorem 5.3, Remark 5.1, and Proposition 5.1).

Theorem 1.2 ([7]). *Let $p \in [2,\infty)$. There exist two constants $c(a,p,d)$ and $C(a,b,p,d)$ such that we have the bound*

$$\varrho_m^o(\mathbf{W}_A^{a,b}, L_p(\mathbb{T}^d)) \le C(a,b,p,d)v^{-a-1/p}(\log(2v))^{(d-1)(a+b)} \tag{1.9}$$

for any $v \in \mathbb{N}$ and any m satisfying

$$m \ge c(a,d,p)v(\log(2v))^3.$$

Theorem 1.2 and embedding (1.4) imply the following analog of the bound (1.8) for the **H** classes. There exist two constants $c = c(r,d,p,q)$ and $C = C(r,d,p,q)$ such that we have the bound for $r > 1/q$

$$\varrho_m^o(\mathbf{H}_q^r, L_p(\mathbb{T}^d)) \le Cv^{-r+1/q-1/p}(\log(2v))^{(d-1)(r+1-1/q)} \tag{1.10}$$

for any $v \in \mathbb{N}$ and any m satisfying

$$m \ge cv(\log(2v))^3.$$

However, we point out that the bound (1.10) is weaker than the corresponding known bound for the linear recovery (see Section 6 for a discussion).

The following lower bound for the **H** classes is the main result of this paper.

Theorem 1.3. *For $1 \le q \le p < \infty$, $p > 1$, $r > 1/q$, we have*

$$\varrho_m^o(\mathbf{H}_q^r, L_p) \ge c(d)m^{-r+1/q-1/p}(\log m)^{(d-1)/p}.$$

Theorem 1.3 is a direct corollary of Lemma 3.2, which is proved in Section 3. The reader can find a discussion of this result in Sections 3 and 6. Note that a new nontrivial feature of Theorem 1.3 is the logarithmic factor $(\log m)^{(d-1)/p}$, which shows that some logarithmic in m factor is needed.

In Section 4 we derive the following lower bound from the known results developed in numerical integration.

Proposition 1.2. *We have for $r > 0$*

$$\varrho_m^o(\mathbf{H}_\infty^r, L_1) \gg m^{-r}(\log m)^{d-1}.$$

In Section 5 we formulate the setting of the sampling recovery in the general space $L_p(\Omega,\mu)$, $1 \le p < \infty$, and instead of the trigonometric system $\mathcal{T}^d$ we study a general uniformly bounded system $\Psi = \{\psi_\mathbf{k}\}_{\mathbf{k}\in\mathbb{Z}^d}$. We prove there a lower bound for the new classes defined with respect to the trigonometric system $\mathcal{T}^d$.

2 Preliminaries

We need some classical trigonometric polynomials. The univariate Fejér kernel of order $j-1$:

$$\mathcal{K}_j(x) := \sum_{|k|\le j}\big(1-|k|/j\big)e^{ikx} = \frac{(\sin(jx/2))^2}{j(\sin(x/2))^2}.$$

The Fejér kernel is an even nonnegative trigonometric polynomial of order $j-1$. It satisfies the obvious relations

$$\|\mathcal{K}_j\|_1 = 1, \qquad \|\mathcal{K}_j\|_\infty = j. \tag{2.1}$$

Let $\mathcal{K}_{\mathbf{j}}(\mathbf{x}) := \prod_{i=1}^d \mathcal{K}_{j_i}(x_i)$ be the d-variate Fejér kernels for $\mathbf{j} = (j_1,\dots,j_d)$ and $\mathbf{x} = (x_1,\dots,x_d)$.

The univariate de la Vallée Poussin kernels are defined as follows

$$\mathcal{V}_m := 2\mathcal{K}_{2m} - \mathcal{K}_m.$$

We also need the following special trigonometric polynomials. Let s be a nonnegative integer. We define

$$\mathcal{A}_0(x) := 1, \quad \mathcal{A}_1(x) := \mathcal{V}_1(x) - 1, \quad \mathcal{A}_s(x) := \mathcal{V}_{2^{s-1}}(x) - \mathcal{V}_{2^{s-2}}(x), \quad s \ge 2,$$

where $\mathcal{V}_m$ are the de la Vallée Poussin kernels defined above. For $\mathbf{s} = (s_1,\dots,s_d) \in \mathbb{N}_0^d$ define

$$\mathcal{A}_{\mathbf{s}}(\mathbf{x}) := \prod_{j=1}^d \mathcal{A}_{s_j}(x_j), \qquad \mathbf{x} = (x_1,\dots,x_d).$$

We now prove Proposition 1.1.

Proof of Proposition 1.1. First, we prove (1.3). It is well known (see, for instance, [14], Ch.2, Theorem 2.1) that for $f \in \mathbf{W}^r_q$ one has for $1 < q < \infty$

$$\|f_j\|_q \le C(d,q,r)2^{-jr}, \quad j \in \mathbb{N}. \tag{2.2}$$

The known results (see Theorem 2.3 below) imply for $1 < q \le 2$

$$\|f_j\|_A \le C(d,q)2^{j/q}j^{(d-1)(1-1/q)}\|f_j\|_q \tag{2.3}$$

$$\le C(d,q,r)2^{-(r-1/q)j}j^{(d-1)(1-1/q)}. \tag{2.4}$$

Therefore, class $\mathbf{W}^r_q$ is embedded into the class $\mathbf{W}^{a,b}_A$ with $a = r-1/q$ and $b = 1-1/q$.

Second, we prove (1.4). Here we need a well known result on the representation of the $\mathbf{H}$ classes (see, for instance, [18], p.137).

Theorem 2.1. *Let $f \in \mathbf{H}^r_{q,l}$. Then for* $\mathbf{s} \geq \mathbf{0}$

$$\big\|A_{\mathbf{s}}(f)\big\|_q \leq C(r,d,l)2^{-r\|\mathbf{s}\|_1}, \qquad 1 \leq q \leq \infty, \tag{2.5}$$

$$\big\|\delta_{\mathbf{s}}(f)\big\|_q \leq C(r,d,q,l)2^{-r\|\mathbf{s}\|_1}, \qquad 1 < q < \infty. \tag{2.6}$$

Conversely, from (2.5) or (2.6) it follows that there exists a $B > 0$, *which does not depend on* f, *such that* $f \in \mathbf{H}^r_{q,l}B$.

By Theorem 2.1 we obtain that for $f \in \mathbf{H}^r_q$ one has for $1 \leq q \leq \infty$

$$\|A_{\mathbf{s}}(f)\|_q \leq C(d,r)2^{-r\|\mathbf{s}\|_1}, \quad \mathbf{s} \in \mathbb{N}^d_0. \tag{2.7}$$

It is known and easy to see that for $q \in [1,2]$

$$\|A_{\mathbf{s}}(f)\|_A \leq C(d)2^{\|\mathbf{s}\|_1/q}\|A_{\mathbf{s}}(f)\|_q \leq C'(d,r)2^{-(r-1/q)\|\mathbf{s}\|_1}. \tag{2.8}$$

Therefore,

$$\|f_j\|_A \leq C''(d,r)2^{-(r-1/q)j}j^{(d-1)}, \tag{2.9}$$

which completes the proof of (1.4).

We formulate some known results from harmonic analysis and from the hyperbolic cross approximation theory, which will be used in our analysis.

We begin with the problem of estimating $\|f\|_u$ in terms of the array $\big\{\|\delta_{\mathbf{s}}(f)\|_v\big\}$. Here and below in this section u and v are scalars such that $1 \leq u, v \leq \infty$. Let an array $\varepsilon = \{\varepsilon_{\mathbf{s}}\}$ be given, where $\varepsilon_{\mathbf{s}} \geq 0$, $\mathbf{s} = (s_1, \dots, s_d)$, and s_j are nonnegative integers, $j = 1, \dots, d$. We denote by $G(\varepsilon, v)$ and $F(\varepsilon, v)$ the following sets of functions ($1 \leq v \leq \infty$):

$$G(\varepsilon,v) := \big\{f \in L_v : \big\|\delta_{\mathbf{s}}(f)\big\|_v \leq \varepsilon_{\mathbf{s}} \qquad \text{for all } \mathbf{s}\big\},$$

$$F(\varepsilon,v) := \big\{f \in L_v : \big\|\delta_{\mathbf{s}}(f)\big\|_v \geq \varepsilon_{\mathbf{s}} \qquad \text{for all } \mathbf{s}\big\}.$$

The following theorem is from [14], p.29 (see also [18], p.94). For the special case $v = 2$ see [13] and [14], p.86.

Theorem 2.2. *The following relations hold:*

$$\sup_{f \in G(\varepsilon,v)} \|f\|_u \asymp \left(\sum_{\mathbf{s}} \varepsilon^u_{\mathbf{s}} 2^{\|\mathbf{s}\|_1(u/v-1)}\right)^{1/u}, \qquad 1 \leq v < u < \infty; \tag{2.10}$$

$$\inf_{f \in F(\varepsilon,v)} \|f\|_u \asymp \left(\sum_{\mathbf{s}} \varepsilon^u_{\mathbf{s}} 2^{\|\mathbf{s}\|_1(u/v-1)}\right)^{1/u}, \qquad 1 < u < v \leq \infty, \tag{2.11}$$

with constants independent of ε.

We will need a corollary of Theorem 2.2 (see [14], Ch.1, Theorem 2.2), which we formulate as a theorem. Let Q be a finite set of points in $\mathbb{Z}^d$, we denote

$$\mathcal{T}(Q) := \left\{ t\colon\, t(\mathbf{x}) = \sum_{\mathbf{k}\in Q} a_{\mathbf{k}} e^{i(\mathbf{k},\mathbf{x})} \right\}$$

and

$$Q_n := \bigcup_{\mathbf{s}:\|\mathbf{s}\|_1 \le n} \rho(\mathbf{s}).$$

Theorem 2.3. *Let $1 < q \le 2$. For any $t \in \mathcal{T}(Q_n)$ we have*

$$\|t\|_A := \sum_{\mathbf{k}} |\hat{t}(\mathbf{k})| \le C(q,d) 2^{n/q} n^{(d-1)(1-1/q)} \|t\|_q.$$

3 The case $1 \le q \le p \le \infty$

Let us discuss lower bounds for the nonlinear characteristic $\varrho^o_m(W, L_p)$. Denote for $\mathbf{N} = (N_1, \dots, N_d)$, $N_j \in \mathbb{N}_0$, $j = 1, \dots, d$,

$$\Pi(\mathbf{N}, d) := \{\mathbf{k} \in \mathbb{Z}^d \,:\, |k_j| \le N_j, j = 1, \dots, d\}$$

and

$$\mathcal{T}(\mathbf{N}, d) := \left\{ f = \sum_{\mathbf{k}\in\Pi(\mathbf{N},d)} c_{\mathbf{k}} e^{i(\mathbf{k},\mathbf{x})} \right\}, \quad \vartheta(\mathbf{N}) := \prod_{j=1}^{d} (2N_j + 1).$$

In this section $\Omega = \mathbb{T}^d$ and μ is the normalized Lebesgue measure on $\mathbb{T}^d$. The following Lemma 3.1 was proved in [21].

Lemma 3.1 ([21])**.** *Let $1 \le q \le p \le \infty$ and let $\mathcal{T}(\mathbf{N}, d)_q$ denote the unit L_q-ball of the subspace $\mathcal{T}(\mathbf{N}, d)$. Then we have for $m \le \vartheta(\mathbf{N})/2$ that*

$$\varrho^o_m(\mathcal{T}(2\mathbf{N}, d)_q, L_p) \ge c(d) \vartheta(\mathbf{N})^{1/q - 1/p}.$$

Let n be a natural number. Denote

$$\mathbf{H}(Q_n)_q := \{f \,:\, f \in \mathcal{T}(Q_n), \quad \|A_{\mathbf{s}}(f)\|_q \le 1\}.$$

Theorem 2.1 implies that $\mathbf{H}(Q_n)_q$ is embedded in $\mathbf{H}^r_q C 2^{rn}$ with some constant C independent of n. Moreover, $\mathbf{H}(Q_{n+b})_q$ is embedded in $\mathbf{H}^r_q C(b) 2^{rn}$ with some constant $C(b)$ independent of n. We now prove the following analog of Lemma 3.1.

Lemma 3.2. *Let $1 \leq q \leq p < \infty$, $p > 1$ and n be a natural number divisible by 3. Denote $S_n := \min_{\|\mathbf{s}\|_1 = n} |\rho(\mathbf{s})|$. Clearly, $S_n \asymp 2^n$. Then there exists a constant b independent of n such that we have for $m \leq S_n/2$*

$$\varrho_m^o(\mathbf{H}(Q_{n+b})_q, L_p) \geq c(d) 2^{n(1/q-1/p)} n^{(d-1)/p}.$$

Proof. Let a set $\xi \subset \mathbb{T}^d := [0, 2\pi]^d$ of points $\xi^1, \dots, \xi^m$ be given. Let n be a natural number divisible by 3 and let $Y_{n,3}$ denote the set of all $\mathbf{s} \in \mathbb{N}^d$ such that all the coordinates of $\mathbf{s}$ are natural numbers divisible by 3 and $\|\mathbf{s}\|_1 = n$. Clearly, $|Y_{n,3}| \asymp n^{d-1}$. Consider the subspaces

$$T(\xi, \mathbf{s}) := \{f \in \mathcal{T}(\rho(\mathbf{s})) : \ f(\xi^\nu) = 0, \quad \nu = 1, \dots, m\}, \quad \mathbf{s} \in Y_{n,3}.$$

Let $g_{\xi,\mathbf{s}} \in T(\xi, \mathbf{s})$ and a point $\mathbf{x}_\mathbf{s}^*$ be such that $|g_{\xi,\mathbf{s}}(\mathbf{x}^*)| = \|g_{\xi,\mathbf{s}}\|_\infty = 1$. We set $2^{\mathbf{s}-2} := (2^{s_1-2}, \dots, 2^{s_d-2})$ and

$$t_\mathbf{s}(\mathbf{x}) := g_{\xi,\mathbf{s}}(\mathbf{x}) \mathcal{K}_{2^{\mathbf{s}-2}}(\mathbf{x} - \mathbf{x}_\mathbf{s}^*), \quad f := \sum_{\mathbf{s} \in Y_{n,3}} t_\mathbf{s}. \tag{3.1}$$

Then $f \in \mathcal{T}(Q_{n+d})$, $f(\xi^\nu) = 0$, $\nu = 1, \dots, m$, and the bound

$$\|g_{\xi,\mathbf{s}} \mathcal{K}_{2^{\mathbf{s}-2}}\|_q \leq \|g_{\xi,\mathbf{s}}\|_\infty \|\mathcal{K}_{2^{\mathbf{s}-2}}\|_q \leq C_1(d) 2^{n(1-1/q)}, \quad \mathbf{s} \in Y_{n,3} \tag{3.2}$$

implies that for all $\mathbf{s}$

$$\|A_\mathbf{s}(f)\|_q \leq C_1(q, d) 2^{n(1-1/q)}. \tag{3.3}$$

In (3.2) we used the known bound for the L_q norm of the Fejér kernel (see [18], p.83, (3.2.7)). Moreover, our assumption that all the coordinates of $\mathbf{s}$ are natural numbers divisible by 3 implies that

$$\delta_\mathbf{u}(f) = \delta_\mathbf{u}(t_\mathbf{s})$$

with only one appropriate $\mathbf{s}$. It is easy to derive from here that for each $\mathbf{s}$ such that $\|\mathbf{s}\|_1 = n$ there exists $\mathbf{u}(\mathbf{s})$ such that $n - 2d \leq \|\mathbf{u}(\mathbf{s})\|_1 \leq n + d$ with the property

$$\|\delta_{\mathbf{u}(\mathbf{s})}(t_\mathbf{s})\|_\infty \geq c(d) \|t_\mathbf{s}\|_\infty, \qquad c(d) > 0. \tag{3.4}$$

By (2.1) we get

$$|t_\mathbf{s}(\mathbf{x}^*)| \geq C_2(d) 2^n. \tag{3.5}$$

We now apply the inequality, which directly follows from (2.11) of Theorem 2.2 with $u = p$ and $v = \infty$, and obtain

$$\|f\|_p \geq C_3(d,p)2^{n(1-1/p)}n^{(d-1)/p}. \tag{3.6}$$

Let $\mathcal{M}$ be a mapping from $\mathbb{C}^m$ to L_p. Denote $g_0 := \mathcal{M}(\mathbf{0})$. Then for $h := f(\max_{\mathbf{s}} \|A_{\mathbf{s}}(f)\|_q)^{-1}$ we have

$$\|h - g_0\|_p + \| - h - g_0\|_p \geq 2\|h\|_p. \tag{3.7}$$

Inequality (3.3) and the fact that $f \in \mathcal{T}(Q_{n+d})$ imply

$$\max_{\mathbf{s}} \|A_{\mathbf{s}}(f)\|_q \leq C_1'(d)2^{n(1-1/q)} \quad \text{and} \quad A_{\mathbf{s}}(f) = 0,\ \|s\|_1 > n + 3d. \tag{3.8}$$

Relations (3.7), (3.8), (3.6), and the fact that both h and $-h$ belong to $\mathbf{H}(Q_{n+d})_q$ complete the proof of Lemma 3.2.

$\square$

As a direct corollary of Lemma 3.2 we obtain Theorem 1.3 from the Introduction.

Remark 3.1. *By the Bernstein inequalities (see, for instance, [18], p.89) one finds out that there exists a constant $C(r,d) > 0$ such that*

$$C(r,d)\vartheta(\mathbf{N})^{-r}\mathcal{T}(2\mathbf{N},d)_q \subset \mathbf{W}^r_q.$$

Then by Lemma 3.1 we obtain

$$\varrho^o_m(\mathbf{W}^r_q, L_p) \geq c(d)m^{-r+1/q-1/p}. \tag{3.9}$$

It is well known (see, for instance, [5], p.42) that classes $\mathbf{W}^r_q$ are embedded in the classes $\mathbf{H}^r_q$. Therefore, (3.9) implies the same lower bound for the classes $\mathbf{H}^r_q$. However, it is weaker than the bound in Theorem 1.3.

4 The case $1 \leq p \leq q \leq \infty$

We now proceed to the case $1 \leq p \leq q \leq \infty$ and concentrate on the special case, when $p = 1$ and $q = \infty$. It is clear that we have the following inequalities for all $1 \leq p \leq q \leq \infty$ and for all classes $\mathbf{F}^r_q$ ($\mathbf{F}$ stands for both $\mathbf{W}$ and $\mathbf{H}$)

$$\varrho^o_m(\mathbf{F}^r_\infty, L_1) \leq \varrho^o_m(\mathbf{F}^r_q, L_p).$$

The following functions were built in [15] (see also [18], pp. 264–266): For any number $n \in \mathbb{N}$ and any set of points $\{\xi^1, \dots, \xi^N\}$, $N \le 2^{n-1}$, there are functions $t_{\mathbf{s}} \in \mathcal{T}(2^{\mathbf{s}-1}, d)$ such that

$$t_{\mathbf{s}}(\xi^j) = 0, \quad j = 1, \dots, N, \qquad \|t_{\mathbf{s}}\|_\infty \le 1$$

and

$$\int_{\mathbb{T}^d} t(\mathbf{x}) d\mathbf{x} \ge c(d) n^{d-1}, \qquad t(\mathbf{x}) := \sum_{\|\mathbf{s}\|_1 = n} t_{\mathbf{s}}(\mathbf{x}). \tag{4.1}$$

Moreover, it was proved there that for $q < \infty$ one has: There exists a constant $c = c(r, q, d) > 0$ such that

$$ct2^{-rn} n^{-(d-1)/2} \in \mathbf{W}^r_q.$$

This example implies the following Proposition 4.1.

Proposition 4.1. *For any $q < \infty$ we have for $r > 1/q$*

$$\varrho^o_m(\mathbf{W}^r_q, L_1) \gg m^{-r} (\log m)^{(d-1)/2}.$$

We now show how the above example implies the lower bound for the $\mathbf{H}$ classes – Proposition 1.2 from Introduction.

Proof of Proposition 1.2. We prove that there is a positive constant $c(d, q, r)$ such that

$$c(d, q, r) t 2^{-rn} \in \mathbf{H}^r_\infty.$$

For that we estimate $\|A_{\mathbf{u}}(t)\|_\infty$ for all $\mathbf{u}$ and use Theorem 2.1. Obviously, $A_{\mathbf{u}}(t) = 0$ if for some j we have $2^{u_j - 2} > 2^{s_j - 1}$. Therefore, it is sufficient to analyze $\mathbf{u}$ such that $\|\mathbf{u}\|_1 \le n + d$. In the same way we see that

$$A_{\mathbf{u}}(t) = \sum_{\mathbf{s} \in Y_n(\mathbf{u})} A_{\mathbf{u}}(t_{\mathbf{s}}), \quad Y_n(\mathbf{u}) := \{\mathbf{s} \,:\, s_j \ge u_j - 1,\, j = 1, \dots, d,\, \|\mathbf{s}\|_1 = n\}.$$

Denote $w_j := s_j - u_j + 1$, $\mathbf{w} := (w_1, \dots, w_d)$. Then for $\mathbf{s} \in Y_n(\mathbf{u})$ we have $w_j \ge 0$, $j = 1, \dots, d$, and

$$\|\mathbf{w}\|_1 = \|\mathbf{s}\|_1 - \|\mathbf{u}\|_1 + d = n - \|\mathbf{u}\|_1 + d.$$

The total number of such $\mathbf{w}$ is $\ll (n - \|\mathbf{u}\|_1 + d)^{d-1}$. Therefore,

$$\|A_{\mathbf{u}}(t)\|_\infty \ll (n - \|\mathbf{u}\|_1 + d)^{d-1} \ll 2^{r(n - \|\mathbf{u}\|_1)}.$$

By Theorem 2.1 this implies that there exists a positive constant $c(d,q,r)$ such that $c(d,q,r)t2^{-rn} \in \mathbf{H}^r_\infty$. We now use (4.1) and complete the proof in the same way as it has been done in the proof of Lemma 3.2 above.

We now make a comment in the style of Remark 3.1, which points out that somewhat weaker than Propositions 4.1 and 1.2 results can be derived from the known results. The following Lemma 4.1 was proved in [20].

Lemma 4.1 ([20])**.** *Let $\mathcal{T}(\mathbf{N},d)_\infty$ denote the unit L_∞-ball of the subspace $\mathcal{T}(\mathbf{N},d)$. Then we have for $m \le \vartheta(\mathbf{N})/2$ that*

$$\varrho^o_m(\mathcal{T}(\mathbf{N},d)_\infty, L_1) \ge c(d) > 0.$$

Remark 4.1. *In the same way as above in Remark 3.1 by using Lemma 4.1 instead of Lemma 3.1 we obtain*

$$\varrho^o_m(\mathbf{H}^r_\infty, L_1) \gg \varrho^o_m(\mathbf{W}^r_\infty, L_1) \gg m^{-r}.$$

Comment on the Gelfand width. 4.1. It is easy to see from the construction of functions t in [15] (see also [18], pp. 264–266), mentioned above, that the sampling linear functionals can be replaced by any linear functionals. This means that Propositions 4.1 and 1.2 hold for the following asymptotic characteristics as well. The Gelfand width is defined as follows

$$c_m(\mathbf{F},X) := \inf_{\varphi_1,\dots,\varphi_m} \sup_{f\in\mathbf{F}:\, \varphi_j(f)=0,\, j=1,\dots,m} \|f\|_X$$

where $\varphi_1,\dots,\varphi_m$ are linear functionals. Thus, we have

$$c_m(\mathbf{W}^r_\infty, L_1) \gg m^{-r}(\log m)^{(d-1)/2} \tag{4.2}$$

and

$$c_m(\mathbf{H}^r_\infty, L_1) \gg m^{-r}(\log m)^{d-1}. \tag{4.3}$$

5 Sampling recovery on classes with structural condition

We formulate the setting of the sampling recovery in the general space $L_p(\Omega,\mu)$, $1 \le p < \infty$, and instead of the trigonometric system $\mathcal{T}^d$ we study a general uniformly bounded system $\Psi = \{\psi_\mathbf{k}\}_{\mathbf{k}\in\mathbb{Z}^d}$. More precisely, let Ω be a compact subset of $\mathbb{R}^d$ with the probability measure μ on it. By the L_p norm, $1 \le p < \infty$, of the complex valued function defined on Ω, we understand

$$\|f\|_p := \|f\|_{L_p(\Omega,\mu)} := \left(\int_\Omega |f|^p d\mu\right)^{1/p} \quad \text{and} \quad \|f\|_\infty := \sup_{\mathbf{x}\in\Omega} |f(\mathbf{x})|.$$

Let a uniformly bounded system $\Psi := \{\psi_{\mathbf{k}}\}_{\mathbf{k}\in\mathbb{Z}^d}$ be indexed by $\mathbf{k} \in \mathbb{Z}^d$. Consider a sequence of subsets $\mathcal{G} := \{G_j\}_{j=1}^\infty$, $G_j \subset \mathbb{Z}^d$, $j = 1, 2, \dots$, such that

$$G_1 \subset G_2 \subset \cdots \subset G_j \subset G_{j+1} \subset \cdots, \qquad \bigcup_{j=1}^\infty G_j = \mathbb{Z}^d. \tag{5.1}$$

Consider functions representable in the form of absolutely convergent series

$$f = \sum_{\mathbf{k}\in\mathbb{Z}^d} a_{\mathbf{k}}(f)\psi_{\mathbf{k}}, \qquad \sum_{\mathbf{k}\in\mathbb{Z}^d} |a_{\mathbf{k}}(f)| < \infty. \tag{5.2}$$

For $\beta \in (0, 1]$ and $r > 0$ consider the following class $\mathbf{A}^r_\beta(\Psi, \mathcal{G})$ of functions f, which have representations (5.2) satisfying conditions

$$\left(\sum_{\mathbf{k}\in G_j\setminus G_{j-1}} |a_{\mathbf{k}}(f)|^\beta\right)^{1/\beta} \le 2^{-rj}, \quad j = 1, 2, \dots, \quad G_0 := \emptyset. \tag{5.3}$$

Probably, the first realization of the idea of the classes $\mathbf{A}^r_\beta(\Psi, \mathcal{G})$ was realised in [17] in the special case, when Ψ is the trigonometric system $\mathcal{T}^d := \{e^{i(\mathbf{k},\mathbf{x})}\}_{\mathbf{k}\in\mathbb{Z}^d}$ (see [21] for a detailed historical discussion). The classes $\mathbf{A}^r_\beta(\Psi)$ studied in [21] correspond to the case of $\mathbf{A}^r_\beta(\Psi, \mathcal{G})$ with

$$G_j := \{\mathbf{k} \in \mathbb{Z}^d \,:\, \|\mathbf{k}\|_\infty < 2^j\}, \quad j = 1, 2, \dots. \tag{5.4}$$

We now define classes $\mathbf{W}^{a,b}_{A_\beta}(\Psi)$, which were introduced and studied in [12]. For

$$f = \sum_{\mathbf{k}} a_{\mathbf{k}}(f)\psi_{\mathbf{k}}, \qquad \sum_{\mathbf{k}} |a_{\mathbf{k}}(f)| < \infty, \tag{5.5}$$

denote

$$\delta_{\mathbf{s}}(f, \Psi) := \sum_{\mathbf{k}\in\rho(\mathbf{s})} a_{\mathbf{k}}(f)\psi_{\mathbf{k}}, \quad f_j := \sum_{\|\mathbf{s}\|_1=j} \delta_{\mathbf{s}}(f, \Psi), \quad j \in \mathbb{N}_0, \quad \mathbb{N}_0 := \mathbb{N} \cup \{0\}$$

and for $\beta \in (0,1]$

$$|f|_{A_\beta(\Psi)} := \left(\sum_{\mathbf{k}} |a_{\mathbf{k}}(f)|^\beta\right)^{1/\beta}.$$

Note, that if representations (5.5) are unique, then in the case $\beta = 1$ the characteristic $|f|_{A_\beta(\Psi)}$ is the norm and in the case $\beta \in (0,1)$ it is a quasi-norm. For parameters $a \in \mathbb{R}_+$, $b \in \mathbb{R}$ define the class $\mathbf{W}^{a,b}_{A_\beta}(\Psi)$ of functions f for which there exists a representation (5.5) satisfying

$$|f_j|_{A_\beta(\Psi)} \le 2^{-aj}(\bar{j})^{(d-1)b}, \quad \bar{j} := \max(j,1), \quad j \in \mathbb{N}_0. \tag{5.6}$$

In the case, when Ψ is the trigonometric system and $\beta = 1$, classes $\mathbf{W}^{a,b}_{A_\beta}(\Psi)$ were introduced in [17]. The general definition in the case $\beta = 1$ is given in [3]. We use the notation A in place of A_1. Thus, $\mathbf{W}^{a,b}_{A}(\Psi) := \mathbf{W}^{a,b}_{A_1}(\Psi)$. Note that the $\mathbf{W}^{a,b}_{A_\beta}(\Psi)$ classes can be seen as the $\mathbf{W}$ type classes with the structural condition on the coefficients in the quasi-norm A_β.

We now define an analog of the $\mathbf{W}^{a,b}_{A_\beta}(\Psi)$ classes in a style of the $\mathbf{H}$ classes. For parameters $a \in \mathbb{R}_+$, $b \in \mathbb{R}$ define the class $\mathbf{H}^{a,b}_{A_\beta}(\Psi)$ of functions f for which there exists a representation (5.5) satisfying

$$|\delta_{\mathbf{s}}(f,\Psi)|_{A_\beta(\Psi)} \le 2^{-aj}(\bar{j})^{(d-1)b}, \quad \bar{j} := \max(j,1), \quad j \in \mathbb{N}_0, \quad \|\mathbf{s}\|_1 = j. \tag{5.7}$$

Note, that the following embedding follows directly from the definitions of the classes

$$\mathbf{H}^{a,b}_{A_\beta} \hookrightarrow \mathbf{W}^{a,b'}_{A_\beta} \quad \text{with} \quad b' = b + 1/\beta. \tag{5.8}$$

We will need some simple properties of the quasi-norms $|\cdot|_{A_\beta}$.

Proposition 5.1. *Assume that Ψ is a uniformly bounded ($\|\psi_{\mathbf{k}}\|_\infty \le B$, $\mathbf{k} \in \mathbb{Z}^d$) orthonormal system. Denote for $\mathbf{N} = (N_1,\dots,N_d)$, $N_j \in \mathbb{N}_0$, $j = 1,\dots,d$,*

$$\Psi(\mathbf{N},d) := \left\{ f = \sum_{\mathbf{k} \in \Pi(\mathbf{N},d)} c_{\mathbf{k}}\psi_{\mathbf{k}} \right\}, \quad \vartheta(\mathbf{N}) := \prod_{j=1}^d (2N_j + 1).$$

Then for $q \in [1,2]$ we have for $f \in \Psi(\mathbf{N},d)$

$$|f|_A \le B^{2/q-1}\vartheta(\mathbf{N})^{1/q}\|f\|_q.$$

Proof. Denote for an array $\mathbf{v} = \{v_{\mathbf{k}}\}_{\mathbf{k}\in\Pi(\mathbf{N},d)}$, $|v_{\mathbf{k}}| = 1$, $\mathbf{k} \in \Pi(\mathbf{N}, d)$,

$$D_{\mathbf{N},\Psi}(\mathbf{v}, \mathbf{x}) := \sum_{\mathbf{k}\in\Pi(\mathbf{N},d)} v_{\mathbf{k}}\psi_{\mathbf{k}}(\mathbf{x}).$$

Let f have the representation (5.5). By the orthonormality assumption we have

$$|f|_A = \sum_{\mathbf{k}\in\Pi(\mathbf{N},d)} |a_{\mathbf{k}}(f)| = \int_\Omega D_{\mathbf{N},\Psi}(\mathbf{v}, \mathbf{x})\bar{f} d\mathbf{x}, \tag{5.9}$$

where $v_{\mathbf{k}} := \operatorname{sign} a_{\mathbf{k}}(f) := a_{\mathbf{k}}(f)/|a_{\mathbf{k}}(f)|$, if $a_{\mathbf{k}} \neq 0$ and $v_{\mathbf{k}} = 1$ if $a_{\mathbf{k}} = 0$; $\bar{f}$ is the complex conjugate to f. From (5.9) we obtain

$$|f|_A \le \|D_{\mathbf{N},\Psi}(\mathbf{v}, \cdot)\|_{q'}\|\bar{f}\|_q, \quad q' := q/(q-1). \tag{5.10}$$

Using simple relations

$$\|D_{\mathbf{N},\Psi}(\mathbf{v}, \cdot)\|_2 = \vartheta(\mathbf{N})^{1/2}, \qquad \|D_{\mathbf{N},\Psi}(\mathbf{v}, \cdot)\|_\infty \le B\vartheta(\mathbf{N})$$

and the well known inequality $\|g\|_p \le \|g\|_2^{2/p}\|g\|_\infty^{1-2/p}$, $p \in [2, \infty)$, we get

$$\|D_{\mathbf{N},\Psi}(\mathbf{v}, \cdot)\|_{q'} \le B^{2/q-1}\vartheta(\mathbf{N})^{1/q}. \tag{5.11}$$

Combining (5.10) and (5.11), we complete the proof.

□

Let $\beta \in (0, 1)$. Then for any set of numbers $\{y_j\}_{j=1}^M$ we have by the Hölder inequality

$$\sum_{j=1}^M |y_j|^\beta \le \left(\sum_{j=1}^M |y_j|\right)^\beta M^{1-\beta}. \tag{5.12}$$

Inequality (5.12) and Proposition 5.1 imply the following Corollary 5.1.

Corollary 5.1. *Under assumptions of Proposition 5.1 for* $q \in [1, 2]$ *and* $\beta \in (0, 1]$ *we have for* $f \in \Psi(\mathbf{N}, d)$

$$|f|_{A_\beta} \le \vartheta(\mathbf{N})^{1/\beta-1}|f|_A \le B^{2/q-1}\vartheta(\mathbf{N})^{1/\beta-1+1/q}\|f\|_q.$$

We now proceed to the main result of this section – the lower bound for $\rho_m^o(\mathbf{H}^{a,b}_{A_\beta}, L_p)$.

Theorem 5.1. *Let $a > 0$ and $b \in \mathbb{R}$. Then for $\beta \in (0, 1]$ and $p \in [2, \infty)$*

$$\rho_m^o(\mathbf{H}^{a,b}_{A_\beta}(\mathcal{T}^d), L_p) \gg m^{1-1/p-1/\beta-a}(\log m)^{(d-1)(b+1/p)}.$$

Proof. Let f be the function defined by (3.1) in the proof of Lemma 3.2. Then by (3.2) and Corollary 5.1 we obtain

$$|t_{\mathbf{s}}|_{A_\beta} \ll 2^{n/\beta}, \tag{5.13}$$

which easily implies that there exists a positive constant c independent of n such that $c2^{n(-a-1/\beta)}n^{(d-1)b}f \in \mathbf{H}^{a,b}_{A_\beta}$. We now use (3.6) and complete the proof in the same way as it has been done in the proof of Lemma 3.2. $\square$

The following upper bound was proved in [12] for $p \in [2, \infty)$

$$\varrho_m^o(\mathbf{W}^{a,b}_{A_\beta}(\mathcal{T}^d), L_p(\mathbb{T}^d)) \ll \left(\frac{m}{(\log m)^3}\right)^{1-1/p-1/\beta-a} (\log(m))^{(d-1)(a+b)}. \tag{5.14}$$

Note that Theorem 5.1 does not cover the case $p \in [1, 2)$. We now present the corresponding lower bound in the case $p = 1$. In the same way as we derived Theorem 5.1 from the example built in the proof of Lemma 3.2 we can derive the following lower bound from the example presented in Section 4

$$\rho_m^o(\mathbf{H}^{a,b}_{A_\beta}(\mathcal{T}^d), L_1) \gg m^{1/2-1/\beta-a}(\log m)^{(d-1)(b+1)}. \tag{5.15}$$

This bound, the upper bound (5.14) with $p = 2$, and the embedding (5.8) show that the characteristics $\rho_m^o(\mathbf{H}^{a,b}_{A_\beta}(\mathcal{T}^d), L_p)$ have the same power decay $m^{1/2-1/\beta-a}$ for all $p \in [1, 2]$.

Comment 5.1. Similarly to Comment 4.1 we have the following analog of the bound (5.15)

$$c_m(\mathbf{H}^{a,b}_{A_\beta}(\mathcal{T}^d), L_1) \gg m^{1/2-1/\beta-a}(\log m)^{(d-1)(b+1)}. \tag{5.16}$$

6 Discussion

In this paper we focus on the study of the lower bounds for the nonlinear characteristic $\varrho_m^o(W, L_p)$. Most of the known results on optimal sampling recovery deal with the linear recovery methods. Recall the setting of the optimal linear recovery. For a fixed m and a set of points $\xi := \{\xi^j\}_{j=1}^m \subset \Omega$, let Φ be a linear operator from $\mathbb{C}^m$ into $L_p(\Omega,\mu)$. Denote for a class W (usually, centrally symmetric and compact subset of $L_p(\Omega,\mu)$)

$$\varrho_m(W,L_p) := \inf_{\text{linear } \Phi;\, \xi} \sup_{f\in W} \|f-\Phi(f(\xi^1),\dots,f(\xi^m))\|_p.$$

The characteristic $\varrho_m(W,L_p)$ was introduced and studied in [16]. The reader can find a detailed discussion of results on $\varrho_m(W,L_p)$ in the books [5] and [18]. Recently, a substantial progress in estimating $\varrho_m(W,L_p)$ and $\varrho^o_m(W,L_p)$ (mostly, the case of recovery in the L_2 norm was studied) was made in the papers [8], [9], [10], [19], [22], [6], [1], [2], [3], [4], [21], [7], [12].

We have an obvious inequality

$$\varrho^o_m(W,L_p) \le \varrho_m(W,L_p), \tag{6.1}$$

which means that the upper bounds for $\varrho_m(W,L_p)$ serve as upper bounds for $\varrho^o_m(W,L_p)$ and the lower bounds for $\varrho^o_m(W,L_p)$ serve as lower bounds for $\varrho_m(W,L_p)$. It is an interesting problem to understand for which function classes W the rates of decay of the characteristics $\varrho_m(W,L_p)$ and $\varrho^o_m(W,L_p)$ coincide. It is known that even in the classical case of classes $\mathbf{W}^r_q$, $1<q<2$, the rates of $\varrho_m(\mathbf{W}^r_q,L_2)$ and $\varrho^o_m(\mathbf{W}^r_q,L_2)$ do not coincide. It was observed in [6] that in the case $1<q<2$ for large enough d the upper bound for $\varrho^o_m(\mathbf{W}^r_q,L_2)$ and the known lower bound for $\varrho_m(\mathbf{W}^r_q,L_2)$ imply that $\varrho^o_m(\mathbf{W}^r_q,L_2) = o(\varrho_m(\mathbf{W}^r_q,L_2))$. This means that in those cases nonlinear methods give better rate of decay of errors of sampling recovery than linear methods do. We now discuss this important phenomenon in detail. We begin with the upper bounds (1.5) and (1.6) for $\varrho^o_m(\mathbf{W}^r_q,L_2)$ given in the Introduction.

For instance, (1.6) and (1.7) imply

$$\varrho^o_m(\mathbf{W}^r_q,L_2(\mathbb{T}^d)) \ll \left(\frac{m}{(\log m)^3}\right)^{-r+1/q-1/2} (\log m)^{(d-1)(r+1-2/q)}. \tag{6.2}$$

We now proceed to the lower bounds for $\varrho_m(\mathbf{W}^r_q,L_2(\mathbb{T}^d))$. It is obvious that

$$d_m(W,L_p) \le \varrho_m(W,L_p), \tag{6.3}$$

where $d_m(W,X)$ is the Kolmogorov width of W in a Banach space X:

$$d_m(W,X) := \inf_{Y\subset X, \dim Y\le m} \sup_{f\in W} \inf_{y\in Y} \|f-y\|_X.$$

Here are the known bounds on the $d_m(\mathbf{W}^r_q, L_2)$ (see, for instance, [18], p.216): For $1 < q \le 2$, $r > 1/q - 1/2$ and $2 < q < \infty$, $r > 0$

$$d_m(\mathbf{W}^r_q, L_2) \asymp \left(\frac{(\log m)^{d-1}}{m}\right)^{r-(1/q-1/2)_+}, \quad (a)_+ := \max(a, 0). \tag{6.4}$$

Taking into account that $r+1-2/q < r+1/2-1/q$ for $1 < q < 2$ we conclude from (6.2), (6.3), and (6.4) that for large enough d we have $\varrho^o_m(\mathbf{W}^r_q, L_2) = o(\varrho_m(\mathbf{W}^r_q, L_2))$.

It is interesting to point out that we do not know if the above effect, which holds for the $\mathbf{W}$ classes, holds for the $\mathbf{H}$ classes as well. The following upper bound is known (see, for instance, [18], p.308) for the linear recovery in the case $1 \le q < p < \infty$, $r > 1/q$,

$$\varrho_m(\mathbf{H}^r_q, L_p(\mathbb{T}^d)) \ll \left(\frac{m}{(\log m)^{d-1}}\right)^{-r+1/q-1/p} (\log m)^{(d-1)/p}. \tag{6.5}$$

In the case $p \ge 2$ the bound (6.5) is better than the bound (1.10) from Introduction. This means that in the case $1 < q \le 2 \le p < \infty$ the known upper bounds for the linear recovery are better than those in (1.10).

Let us make some comments on the technique available for proving the lower bounds for the optimal recovery. In the case of linear recovery we can use the inequality (6.3) or even a stronger one with the Kolmogorov width replaced by the linear width. This theory is well developed (see, for instance, the books [5] and [18]). In the case of nonlinear recovery we cannot use the above approach. It seems like the only technique, which is available now, is based on the following simple observation.

Proposition 6.1. *Let $W \subset X$ be a symmetric ($f \in W \Rightarrow -f \in W$) subset, consisting of continuous functions. Then*

$$\varrho^o_m(W, X) \ge \inf_{\xi^1,\dots,\xi^m} \sup_{f \in W\,:\, f(\xi^j)=0,\, j=1,\dots,m} \|f\|_X.$$

Acknowledgements. This work was supported by Ministry of Science and Higher Education of the Russian Federation (Grant No. 075-15-2024-529).

References

[1] M. Dolbeault , D. Krieg, and M. Ullrich, A sharp upper bound for sampling numbers in L_2, *Appl. Comput. Harmon. Anal.* **63** (2023), 113–134; arXiv:2204.12621v1 [math.NA] 26 Apr 2022.

[2] F. Dai and V.N. Temlyakov, Universal discretization and sparse sampling recovery, arXiv:2301.05962v1 [math.NA] 14 Jan 2023.

[3] F. Dai and V.N. Temlyakov, Random points are good for universal discretization, J. Math. Anal. Appl., **529** (2024) 127570; arXiv.2301.12536[math.FA] 5 Feb 2023.

[4] F. Dai and V.N. Temlyakov, Lebesgue-type inequalities in sparse sampling recovery, arXiv:2307.04161v1 [math.NA] 9 Jul 2023.

[5] Dinh Dũng, V.N. Temlyakov, and T. Ullrich, Hyperbolic Cross Approximation, Advanced Courses in Mathematics CRM Barcelona, Birkhäuser, 2018; arXiv:1601.03978v2 [math.NA] 2 Dec 2016.

[6] T. Jahn, T. Ullrich, and F. Voigtlaender, Sampling numbers of smoothness classes via ℓ^1-minimization, arXiv:2212.00445v3 [math.NA] 31 Jul 2023.

[7] E. Kosov and V. Temlyakov, Bounds for the sampling discretization error and their applications to universal sampling discretization, arXiv:2312.05670v2 [math.NA] 27 Jan 2024.

[8] D. Krieg and M. Ullrich, Function values are enough for L_2-approximation, *Found. Comp. Math.*, doi:10.1007/s10208-020-09481-w; arXiv:1905.02516v4 [math.NA] 19 Mar 2020.

[9] D. Krieg and M. Ullrich, Function values are enough for L_2-approximation: Part II, *J. Complexity*, doi:10.1016/j.jco.2021.101569; arXiv:2011.01779v1 [math.NA] 3 Nov 2020.

[10] N. Nagel, M. Schäfer, T. Ullrich, A new upper bound for sampling numbers, *Found. Comp. Math.*, Pub Date: 2021-04-26, DOI: 10.1007/s10208-021-09504-0; arXiv:2010.00327v1 [math.NA] 30 Sep 2020.

[11] S.A. Smolyak, Quadrature and interpolation formulas for tensor products of certain classes of functions. Dokl. Akad. Nauk SSSR, 148 (1963), 10421045; English translation in Soviet Math. Dokl., 4 (1963).

[12] A.P. Solodov and V.N. Temlyakov, Sampling recovery on function classes with a structural condition, arXiv:2404.07210v2 [math.NA] 25 Jun 2024.

[13] V.N. Temlyakov, Approximation of Periodic Functions of Several Variables by Bilinear Forms, Izvestiya AN SSSR, Ser. Mat., **50** (1986), 137–155; English transl. in Mathematics of the USSR-Izvestia, **28** (1987), 133–150.

[14] V.N. Temlyakov, Approximation of functions with bounded mixed derivative, Trudy MIAN, **178** (1986), 1–112. English transl. in Proc. Steklov Inst. Math., **1** (1989).

[15] V.N. Temlyakov, On a way of obtaining lower estimates for the errors of quadrature formulas, Matem. Sbornik, **181** (1990), 1403–1413; English transl. in Math. USSR Sbornik, **71** (1992).

[16] V.N. Temlyakov, On Approximate Recovery of Functions with Bounded Mixed Derivative, J. Complexity, **9** (1993), 41–59.

[17] V.N. Temlyakov, Constructive sparse trigonometric approximation and other problems for functions with mixed smoothness, arXiv: 1412.8647v1 [math.NA] 24 Dec 2014, 1–37; Matem. Sb., **206** (2015), 131–160.

[18] V. Temlyakov, *Multivariate Approximation*, Cambridge University Press, 2018.

[19] V.N. Temlyakov, On optimal recovery in L_2, J. Complexity **65** (2021), 101545; arXiv:2010.03103v1 [math.NA] 7 Oct 2020.

[20] V. Temlyakov, Sparse sampling recovery by greedy algorithms, arXiv:2312.13163v2 [math.NA] 30 Dec 2023.

[21] V. Temlyakov, Sparse sampling recovery in integral norms on some function classes, arXiv:2401.14670v1 [math.NA] 26 Jan 2024.

[22] V.N. Temlyakov and T. Ullrich, Bounds on Kolmogorov widths of classes with small mixed smoothness, J. Complexity, Available online 4 May 2021, 101575; arXiv:2012.09925v1 [math.NA] 17 Dec 2020.

[23] J.F. Traub, G.W. Wasilkowski, and H. Woźniakowski, Information-Based Complexity, Academic Press, Inc., 1988.

A.V. Gasnikov,
Ivannikov institute for System Programming of Russian Academy of Sciences, Moscow, Russia;
Steklov Mathematical Institute of Russian Academy of Sciences, Moscow, Russia;
Innopolis University, Tatarstan, Russia.
E-mail: gasnikov@yandex.ru

V.N. Temlyakov, University of South Carolina, USA,
Steklov Mathematical Institute of Russian Academy of Sciences, Russia;
Lomonosov Moscow State University, Russia;
Moscow Center of Fundamental and Applied Mathematics, Russia.
E-mail: temlyakovv@gmail.com