Chin. Ann. Math. Ser. B
42(4), 2021, 519–528
DOI: 10.1007/s11401-021-0275-4

Chinese Annals of Mathematics, Series B
© The Editorial Office of CAM and
Springer-Verlag Berlin Heidelberg 2021

Sample Numbers and Optimal Lagrange Interpolation of Sobolev Spaces W_1^r*

Guiqiao XU[1] Zehong LIU[1] Hui WANG[1]

Abstract This paper investigates the optimal recovery of Sobolev spaces $W_1^r[-1,1], r \in \mathbb{N}$ in the space $L_1[-1,1]$. They obtain the values of the sampling numbers of $W_1^r[-1,1]$ in $L_1[-1,1]$ and show that the Lagrange interpolation algorithms based on the extreme points of Chebyshev polynomials are optimal algorithms. Meanwhile, they prove that the extreme points of Chebyshev polynomials are optimal Lagrange interpolation nodes.

Keywords Worst case setting, Sampling number, Optimal Lagrange interpolation nodes, Sobolev space
2000 MR Subject Classification 41A05, 41A25

1 Introduction and Main Results

Let F be a Banach space of functions defined on a compact set D that can be continuously embedded in $C(D)$, BF is the unit ball of F, and $G\,(\supseteq F)$ is a normed linear space with norm $\|\cdot\|_G$. We want to approximate functions f from BF by using a finite number of arbitrary function values $f(t)$ (standard information) for some $t \in D$. We consider only nonadaptive information. For $\mathbf{x} = (x_1, x_2, \cdots, x_n) \in D^n$, we use $I_{\mathbf{x}}$ to denote the nonadaptive information operator, i.e.,

$$I_{\mathbf{x}}(f) := (f(x_1), f(x_2), \cdots, f(x_n)) \in \mathbb{R}^n, \quad f \in F.$$

We say that $A_n = \varphi \circ I_{\mathbf{x}}$ is an algorithm based on the information operator $I_{\mathbf{x}}$, where φ is an arbitrary mapping from $\mathbb{R}^n$ to G. We also consider linear algorithms, i.e., algorithms of the form

$$A_n^{\rm lin}(f) = \varphi^{\rm lin} \circ I_{\mathbf{x}}(f) := \sum_{j=1}^{n} f(x_j) h_j, \quad h_j \in G,\ x_j \in D,\ j = 1, \cdots, n.$$

We use an algorithm A_n to reconstruct functions from BF. The worst case error of the algorithm A_n for BF in G is defined by

$$e(BF, A_n, G) := \sup_{f \in BF} \|f - A_n(f)\|_G. \tag{1.1}$$

Manuscript received September 26, 2020. Revised March 13, 2021.
[1]Department of Mathematics, Tianjin Normal University, Tianjin 300387, China.
E-mail: Xuguiqiao@aliyun.com llhcliuzehong@qq.com 929280544@qq.com
*This work was supported by the National Natural Science Foundation of China (Nos. 11871006, 11671271).

For a given $\mathbf{x} = (x_1, x_2, \cdots, x_n) \in D^n$, the worst case error for BF in G based on the information operator $I_{\mathbf{x}}$ is defined by

$$e(BF, I_{\mathbf{x}}, G) := \inf_{\varphi} \sup_{f \in BF} \|f - \varphi \circ I_{\mathbf{x}}(f)\|_G,$$

where the infimum is taken over all mappings φ from $\mathbb{R}^n$ to G.

We define the linear sampling numbers and the sampling numbers for BF in G by

$$g_n^{\text{lin}}(BF, G) := \inf_{A_n^{\text{lin}}} e(BF, A_n^{\text{lin}}, G),$$

and

$$g_n(BF, G) := \inf_{A_n} e(BF, A_n, G) = \inf_{\mathbf{x} \in D^n} e(BF, I_{\mathbf{x}}, G), \tag{1.2}$$

respectively. If there exists an information operator $I_{\mathbf{x}^*}$ and a mapping φ^* such that the algorithm $A_n^* = \varphi^* \circ I_{\mathbf{x}^*}$ satisfies

$$e(BF, A_n^*, G) = g_n(BF, G),$$

then we call $I_{\mathbf{x}^*}$ the nth optimal information and A_n^* the nth optimal algorithm.

The sampling numbers are closely related to many classical approximation problems such as width and information-based complexity, and they have a wide range of applications in numerical analysis. The aim of studying sampling numbers is to find optimal or nearly optimal information, construct optimal or nearly optimal algorithms according to the known standard information, and determine orders (or values) of the sampling numbers.

Let $L_1 \equiv L_1[-1, 1]$ be the space of measurable functions defined on $[-1, 1]$, for which the norm

$$\|f\|_1 := \int_{-1}^{1} |f(x)| \mathrm{d}x$$

is finite. Denote by $W_1^r \equiv W_1^r[-1, 1], r \in \mathbb{N}$ the class of all functions f such that $f^{(r-1)}$ ($f^{(0)} := f$) are absolutely continuous and $f^{(r)} \in L_1$.

In recent years, the study of sampling numbers has attracted much interest, and a great number of interesting results have been obtained (see [1–15]). This paper investigates the sampling numbers of Sobolev spaces W_1^r in L_1. We remark that, in most cases, we can achieve only weak equivalences (orders) of the sampling numbers. In this paper, we obtain the values of the sampling numbers of Sobolev spaces W_1^r in L_1. To show our results, we introduce the following Lagrange interpolation algorithms.

Let $x_1, x_2, \cdots, x_n$ be n distinct points in $[-1, 1]$. Write $\mathbf{x} = (x_1, x_2, \cdots, x_n)$. Then, the Lagrange interpolation polynomial $L_{\mathbf{x}}(f)$ of a function $f : [-1, 1] \to \mathbb{R}$ based on the knots $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ is defined by

$$L_{\mathbf{x}}(f) \in \mathcal{P}_{n-1}, \quad L_{\mathbf{x}}(f, x_k) = f(x_k), \quad k = 1, 2, \cdots, n, \tag{1.3}$$

where and in the following, $\mathcal{P}_n$ represents the space of all algebraic polynomials of degree at most n. The classical Lagrange interpolation formula gives

$$L_{\mathbf{x}}(f,x)=\sum_{k=1}^{n} f(x_k)\ell_k(x),$$

where

$$\ell_k(x)=\frac{W_{\mathbf{x}}(x)}{(x-x_k)W'_{\mathbf{x}}(x_k)},\quad W_{\mathbf{x}}(x)=\prod_{k=1}^{n}(x-x_k).$$

First, we obtain the following results.

Theorem 1.1 *For $r\in\mathbb{N}$, we have*

$$g_r(BW_1^r,L_1)=g_r^{\rm lin}(BW_1^r,L_1)=e(BW_1^r,L_{\mathbf{x}_r},L_1)=\frac{C_r}{r!},\tag{1.4}$$

where

$$\mathbf{x}_r=\Big(\cos\frac{r\pi}{r+1},\cos\frac{(r-1)\pi}{r+1},\cdots,\cos\frac{\pi}{r+1}\Big)\tag{1.5}$$

is the set of extreme points of $(r+1)$*th* Chebyshev *polynomial* $T_{r+1}(x)=\cos((r+1)\arccos x)$, *and*

$$C_r=\Big\|(1-\cdot)^r-2\sum_{i=1}^{r}(-1)^{i-1}\Big(\cos\frac{i\pi}{r+1}-\cdot\Big)_+^r\Big\|_\infty,\quad x_+^r=\begin{cases}x^r, & x\ge 0;\\ 0, & x<0.\end{cases}\tag{1.6}$$

Choosing nodes is important for interpolation algorithms. Given a sufficiently smooth function, if nodes are not suitably chosen, then the interpolation polynomials do not converge to the function as the number of nodes tends to infinity. A well-known example is the Runge's phenomenon. Hence the study of optimal interpolation nodes is a hot topic, see [16–19] and the references therein. In general, if nodes $\mathbf{c}=(c_1,c_2,\cdots,c_n)\in[-1,1]^n$ satisfies

$$e(BF,L_{\mathbf{c}},G)=\inf_{\mathbf{x}=(x_1,x_2,\cdots,x_n)\in[-1,1]^n} e(BF,L_{\mathbf{x}},G),\tag{1.7}$$

then we call $\mathbf{c}=(c_1,c_2,\cdots,c_n)$ the nth optimal Lagrange interpolation nodes and $L_{\mathbf{c}}$ the nth optimal Lagrange interpolation algorithm for BF in G. The value $e(BF,L_{\mathbf{c}},G)$ is called the nth optimal Lagrange interpolation error for BF in G and we denote it as $e(n,BF,G)$.

Using $C^r\equiv C^r[-1,1]$, $r=0,1,2,\cdots$ represents the spaces of functions with rth order continuous derivative on $[-1,1]$, respectively. The most important optimal Lagrange interpolation nodes problem is for C^0 in L_∞. For $n=3$ and $n=4$, the results can be found in [20] and [21], respectively. For $n\ge 5$, it is still an open problem. For $r\ge 1$, it is well known that the rth optimal Lagrange interpolation nodes are the zeros of the rth Chebyshev polynomial $T_r(x)=\cos(r\arccos x)$ for C^r in L_∞. In this paper, we give the rth optimal Lagrange interpolation nodes for BW_1^r in L_1. The result is as follows.

Theorem 1.2 *Let $r \in \mathbb{N}$. Then we have*

$$e(r, BW_1^r, L_1) = e(BW_1^r, L_{\mathbf{x}_r}, L_1) = \frac{C_r}{r!},$$

where $\mathbf{x}_r$ and C_r are given by (1.5) *and* (1.6)*, respectively.*

The remainder of this paper is organized as follows. In Section 2, we give some lemmas related to the proof of our main results. The proofs of Theorems 1.1 and 1.2 are given in Section 3 respectively.

2 Background Information

First we introduce a remainder theorem about Lagrange interpolation (see [22]). Let $x_0, x_1, x_2, \cdots, x_n$ be $n+1$ distinct points in $[-1,1]$. For $0 \le i \le n$, let

$$\Delta_i = (-1)^i \begin{vmatrix} \frac{x_0^{n-1}}{(n-1)!} & \frac{x_0^{n-2}}{(n-2)!} & \cdots & \frac{x_0^{n-j+1}}{(n-j+1)!} & \frac{x_0^{n-j}}{(n-j)!} & \cdots & 1 \\ \frac{x_1^{n-1}}{(n-1)!} & \frac{x_1^{n-2}}{(n-2)!} & \cdots & \frac{x_1^{n-j+1}}{(n-j+1)!} & \frac{x_1^{n-j}}{(n-j)!} & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{x_{i-1}^{n-1}}{(n-1)!} & \frac{x_{i-1}^{n-2}}{(n-2)!} & \cdots & \frac{x_{i-1}^{n-j+1}}{(n-j+1)!} & \frac{x_{i-1}^{n-j}}{(n-j)!} & \cdots & 1 \\ \frac{x_{i+1}^{n-1}}{(n-1)!} & \frac{x_{i+1}^{n-2}}{(n-2)!} & \cdots & \frac{x_{i+1}^{n-j+1}}{(n-j+1)!} & \frac{x_{i+1}^{n-j}}{(n-j)!} & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{x_n^{n-1}}{(n-1)!} & \frac{x_n^{n-2}}{(n-2)!} & \cdots & \frac{x_n^{n-j+1}}{(n-j+1)!} & \frac{x_n^{n-j}}{(n-j)!} & \cdots & 1 \end{vmatrix}. \tag{2.1}$$

Then for $f \in W_1^n$, it follows from [22, (8)] that

$$\sum_{i=0}^{n} \Delta_i \Big(f(x_i) - \int_{-1}^{1} f^{(n)}(t) \cdot \frac{(x_i - t)_+^{n-1}}{(n-1)!} \mathrm{d}t \Big) = 0. \tag{2.2}$$

In particular, if $f(x_i) = 0$ for $1 \le i \le n$, $x_0 = x$, then (2.2) becomes

$$f(x) = \int_{-1}^{1} B_{\mathbf{x}}(x,t) f^{(n)}(t) \mathrm{d}t, \tag{2.3}$$

where $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ and

$$B_{\mathbf{x}}(x,t) = \sum_{i=0}^{n} \frac{(x_i - t)_+^{n-1}}{(n-1)!} \frac{\Delta_i}{\Delta_0}.$$

Noting that $x_0 = x$, for $i = 1, 2, \cdots, n$, from (2.1) it is easy to verify that

$$\frac{\Delta_i}{\Delta_0} = -\ell_i(x). \tag{2.4}$$

Hence, it follows from (2.4) that

$$B_{\mathbf{x}}(x,t) = \frac{1}{(n-1)!} \Big((x-t)_+^{n-1} - \sum_{i=1}^{n} (x_i - t)_+^{n-1} \ell_i(x) \Big)$$

$$= \frac{(x-t)_+^{n-1} - L_{\mathbf{x}}((\cdot - t)_+^{n-1}, x)}{(n-1)!}. \tag{2.5}$$

If $f \in W_1^n$, $\mathbf{x} = (x_1, x_2, \cdots, x_n)$, then it follows from (1.3) that

$$f(x_i) - L_{\mathbf{x}}(f, x_i) = 0, \quad i = 1, 2, \cdots, n. \tag{2.6}$$

Since $L_{\mathbf{x}}(f)$ is an algebraic polynomial of degree at most $n-1$, we conclude that $(f - L_{\mathbf{x}}(f))^{(n)}(t) = f^{(n)}(t)$ and this means $f - L_{\mathbf{x}}(f) \in W_1^n$. Combining these facts with (2.3) and (2.6), we obtain

$$f(x) - L_{\mathbf{x}}(f, x) = \int_{-1}^{1} B_{\mathbf{x}}(x,t)(f - L_{\mathbf{x}}(f))^{(n)}(t)\mathrm{d}t = \int_{-1}^{1} B_{\mathbf{x}}(x,t) f^{(n)}(t)\mathrm{d}t. \tag{2.7}$$

Now we introduce some information about the norms of integral operators. Let $K(x,t)$ be a piecewise continuous function on $[-1,1]^2$. We define

$$S(f,x) = \int_{-1}^{1} K(x,t) f(t)\mathrm{d}t.$$

It is known that S is a linear continuous operator from L_1 to L_1. Furthermore, let $\|S\|_{1,1}$ be the operator norm of S from L_1 to L_1. Then it is known that

$$\|S\|_{1,1} = \sup_{f \in L_1, f \neq 0} \frac{\|Sf\|_1}{\|f\|_1} = \sup_{-1 \le t \le 1} \int_{-1}^{1} |K(x,t)|\mathrm{d}x. \tag{2.8}$$

Lemma 2.1 *Let $-1 \le x_1 < x_2 < \cdots < x_r \le 1$, $\mathbf{x} = (x_1, x_2, \cdots, x_r)$. Then we have*

$$e(BW_1^r, L_{\mathbf{x}}, L_1) = \sup_{-1 \le t \le 1} \int_{-1}^{1} |B_{\mathbf{x}}(x,t)|\mathrm{d}x, \tag{2.9}$$

where $B_{\mathbf{x}}(x,t)$ is given by (2.5).

Proof If $f \in BW_1^r$, then it follows from (2.7) with $n = r$ that

$$f(x) - L_{\mathbf{x}}(f, x) = \int_{-1}^{1} B_{\mathbf{x}}(x,t) f^{(r)}(t)\mathrm{d}t. \tag{2.10}$$

Let

$$T(f,x) = \int_{-1}^{1} B_{\mathbf{x}}(x,t) f(t)\mathrm{d}t. \tag{2.11}$$

Then it follows from (2.8) and (2.10) that

$$\begin{aligned} \|f - L_{\mathbf{x}}(f)\|_1 = \|T(f^{(r)})\|_1 &\le \|f^{(r)}\|_1 \cdot \sup_{-1 \le t \le 1} \int_{-1}^{1} |B_{\mathbf{x}}(x,t)|\mathrm{d}x \\ &\le \sup_{-1 \le t \le 1} \int_{-1}^{1} |B_{\mathbf{x}}(x,t)|\mathrm{d}x. \end{aligned} \tag{2.12}$$

By (1.1) and (2.12), we conclude that

$$e(BW_1^r, L_{\mathbf{x}}, L_1) \le \sup_{-1 \le t \le 1} \int_{-1}^{1} |B_{\mathbf{x}}(x,t)|\mathrm{d}x. \tag{2.13}$$

On the other hand, for any $g \in L_1[-1,1]$, let

$$\overline{f}(x) = \frac{1}{(r-1)!}\int_{-1}^{x}(x-t)^{r-1}g(t)\mathrm{d}t.$$

By a direct computation, we obtain

$$\overline{f}^{(r)}(x) = g(x). \tag{2.14}$$

From (2.10) and (2.14) it follows that

$$\overline{f}(x) - L_{\mathbf{x}}(\overline{f}, x) = \int_{-1}^{1} B_{\mathbf{x}}(x,t)\overline{f}^{(r)}(t)\mathrm{d}t = \int_{-1}^{1} B_{\mathbf{x}}(x,t)g(t)\mathrm{d}t = T(g,x). \tag{2.15}$$

By (2.15), we obtain

$$\|\overline{f} - L_{\mathbf{x}}(\overline{f})\|_1 = \|T(g)\|_1. \tag{2.16}$$

From (1.1), (2.8) and (2.16) it follows that

$$e(BW_1^r, L_{\mathbf{x}}, L_1) \geq \sup_{\|g\|_1\leq 1}\|T(g)\|_1 = \sup_{-1\leq t\leq 1}\int_{-1}^{1}|B_{\mathbf{x}}(x,t)|\mathrm{d}x. \tag{2.17}$$

Combining (2.13) with (2.17), we obtain (2.9). This completes the proof of Lemma 2.1.

An n-dimensional subspace G of $C[-1,1]$ is called a weak Chebyshev subspace if every function $g \in G$ has at most $n-1$ sign changes. By [23, Theorem 6.3] we know that for every n-dimensional weak Chebyshev subspace of $C[-1,1]$, there exists a set of n-canonical points $t_1 < \cdots < t_n$ in $(-1,1)$, i.e., there exist $t_1 < \cdots < t_n$ in $(-1,1)$ such that

$$\sum_{i=0}^{n}(-1)^i\int_{t_i}^{t_{i+1}} g(t)\mathrm{d}t = 0 \tag{2.18}$$

holds for all $g \in G$, where $t_0 = -1$ and $t_{n+1} = 1$.

If G is a weak Chebyshev subspace of $C[-1,1]$, then the set

$$K(G) = \{f \in C[-1,1] : \mathrm{span}(G \cup f) \text{ is a weak Chebyshev subspace of } C[-1,1]\}$$

is called the convexity cone of G.

Lemma 2.2 (see [23, Theorem 6.6]) *Let G be an n-dimensional weak* Chebyshev *subspace of $C[-1,1]$. If the set $\{t_1,\cdots,t_n\}$ of canonical points of G is poised with respect to G, then every function $f \in K(G)$ has a unique best L_1-approximation g_f from G and g_f is uniquely determined by*

$$g_f(t_i) = f(t_i), \quad i = 1,\cdots,n. \tag{2.19}$$

Lemma 2.3 (see [24, Lemma 4.3]) *For $\mathbf{x} = (x_1, x_2, \cdots, x_r) \in [-1,1]^r$, we have*

$$e(BW_1^r, I_{\mathbf{x}}, L_1) := \inf_{\varphi}\sup_{\|f^{(r)}\|_1\leq 1}\|f - \varphi\circ I_{\mathbf{x}}(f)\|_1 = \sup_{\substack{\|f^{(r)}\|_1\leq 1,\\ f(x_1)=f(x_2)=\cdots=f(x_r)=0}}\|f\|_1. \tag{2.20}$$

3 Proofs of Theorem 1.1 and Theorem 1.2

Proof of Theorem 1.1 We consider the upper estimate first. Let $\mathbf{x}_r$ be given by (1.5). Then, the Lagrange interpolation algorithm $L_{\mathbf{x_r}}$ is a linear algorithm. Hence, it follows from (2.9) that

$$g_r(BW_1^r, L_1) \le g_r^{\mathrm{lin}}(BW_1^r, L_1) \le e(BW_1^r, L_{\mathbf{x}_r}, L_1) = \sup_{-1\le t\le 1} \int_{-1}^{1} |B_{\mathbf{x}_r}(x,t)|\mathrm{d}x. \tag{3.1}$$

We will compute the last integration in (3.1). For $t=-1$, we have $(\cdot - t)_+^{r-1} = (\cdot - t)^{r-1} \in \mathcal{P}_{r-1}$. For $t=1$, we have $(\cdot - t)_+^{r-1} = 0 \in \mathcal{P}_{r-1}$. Hence from (2.5) it follows that $B_{\mathbf{x}_r}(x,t)=0$ for $t=\pm 1$. This means that

$$\int_{-1}^{1} |B_{\mathbf{x}_r}(x,t)|\mathrm{d}x = 0 \quad \text{for } t=\pm 1. \tag{3.2}$$

Next we consider $t\in(-1,1)$. It is known that $\mathcal{P}_r$ is a Chebyshev subspace of $C[-1,1]$. Furthermore, from [23, Theorem 4.10] we know that the canonical points for $\mathcal{P}_r$ on $[-1,1]$ are the extreme points of the Chebyshev polynomial T_{r+2} in $(-1,1)$, i.e.,

$$t_i = \cos\frac{(r+2-i)\pi}{r+2}, \quad i=1,\cdots,r+1. \tag{3.3}$$

From (3.3) it follows that $\mathbf{x}_r$ is the set of the canonical points for $\mathcal{P}_{r-1}$ on $[-1,1]$. Furthermore, from [23, Theorem 1.19] we know that for each $t\in(-1,1)$, $\mathrm{span}(\mathcal{P}_{r-1}\cup(\cdot - t)_+^{r-1})$ is an $(r+1)$-dimensional weak Chebyshev space of $C[-1,1]$, i.e., $(\cdot - t)_+^{r-1} \in K(\mathcal{P}_{r-1})$. Hence from Lemma 2.2 and (1.3), it follows that $L_{\mathbf{x}_r}((\cdot - t)_+^{r-1})$ is the best L_1-approximation of $(\cdot - t)_+^{r-1}$ from $\mathcal{P}_{r-1}$ on $[-1,1]$. Therefore, from [25, Theorem 10.5], (3.2) and (2.18), it follows that

$$\begin{aligned}\int_{-1}^{1} |B_{\mathbf{x}_r}(x,t)|\mathrm{d}x &= \frac{1}{(r-1)!}\int_{-1}^{1} |(x-t)_+^{r-1} - L_{\mathbf{x}_r}((\cdot - t)_+^{r-1}, x)|\mathrm{d}x \\ &= \frac{1}{(r-1)!}\Big|\sum_{i=0}^{r}(-1)^i \int_{\cos\frac{(i+1)\pi}{r+1}}^{\cos\frac{i\pi}{r+1}} (x-t)_+^{r-1}\mathrm{d}x\Big|. \end{aligned}\tag{3.4}$$

For $t\in(-1,1)$, it is obvious that there exists an N_t with $0\le N_t\le r$ such that $t\in \big[\cos\frac{(N_t+1)\pi}{r+1}, \cos\frac{N_t\pi}{r+1}\big)$. Then (3.4) becomes

$$\begin{aligned}&\int_{-1}^{1} |B_{\mathbf{x}_r}(x,t)|\mathrm{d}x \\ &= \frac{1}{(r-1)!}\Big|\sum_{i=0}^{N_t-1}(-1)^i \int_{\cos\frac{(i+1)\pi}{r+1}}^{\cos\frac{i\pi}{r+1}} (x-t)^{r-1}\mathrm{d}x + (-1)^{N_t}\int_{t}^{\cos\frac{N_t\pi}{r+1}} (x-t)^{r-1}\mathrm{d}x\Big|. \end{aligned}\tag{3.5}$$

By (3.5) and a direct computation, we obtain

$$\begin{aligned}\int_{-1}^{1} |B_{\mathbf{x}_r}(x,t)|\mathrm{d}x &= \frac{1}{r!}\Big((1-t)^r - 2\sum_{i=1}^{N_t}(-1)^{i-1}\Big(\cos\frac{i\pi}{r+1} - t\Big)^r\Big) \\ &= \frac{1}{r!}\Big((1-t)^r - 2\sum_{i=1}^{r}(-1)^{i-1}\Big(\cos\frac{i\pi}{r+1} - t\Big)_+^r\Big). \end{aligned}\tag{3.6}$$

From (3.1)–(3.2) and (3.6) we obtain the upper estimate.

Now we consider the lower estimate. Let $x_1, x_2, \cdots, x_r$ be r arbitrary distinct points in $[-1,1]$ and $\mathbf{x} = (x_1, x_2, \cdots, x_r)$. Combining Lemma 2.3 with (2.6) as well as $(f - L_{\mathbf{x}}(f))^{(r)}(x) = f^{(r)}(x)$, we obtain

$$\begin{aligned} e(BW_1^r, I_{\mathbf{x}}, L_1) &= \sup_{\substack{\|f^{(r)}\|_1 \le 1,\\ f(x_1)=f(x_2)=\cdots=f(x_r)=0}} \|f\|_1 \\ &\ge \sup_{\|f^{(r)}\|_1 \le 1} \|f - L_{\mathbf{x}}(f)\|_1 = e(BW_1^r, L_{\mathbf{x}}, L_1). \end{aligned} \tag{3.7}$$

From (3.7) and (2.9) it follows that

$$e(BW_1^r, I_{\mathbf{x}}, L_1) \ge \sup_{-1 \le t \le 1} \int_{-1}^{1} |B_{\mathbf{x}}(x,t)| \mathrm{d}x. \tag{3.8}$$

For any $-1 \le t \le 1$, since $L_{\mathbf{x}_r}((\cdot - t)_+^{r-1})$ is the best L_1-approximation of $(\cdot - t)_+^{r-1}$ from $\mathcal{P}_{r-1}$ on $[-1,1]$, we obtain

$$\begin{aligned} \int_{-1}^{1} |B_{\mathbf{x}}(x,t)| \mathrm{d}x &= \frac{1}{(r-1)!} \int_{-1}^{1} |(x-t)_+^{r-1} - L_{\mathbf{x}}((\cdot - t)_+^{r-1}, x)| \mathrm{d}x \\ &\ge \frac{1}{(r-1)!} \int_{-1}^{1} |(x-t)_+^{r-1} - L_{\mathbf{x}_r}((\cdot - t)_+^{r-1}, x)| \mathrm{d}x \\ &= \int_{-1}^{1} |B_{\mathbf{x}_r}(x,t)| \mathrm{d}x. \end{aligned} \tag{3.9}$$

From (1.2) and (3.8)–(3.9) we obtain the lower estimate. This completes the proof of Theorem 1.1.

Proof of Theorem 1.2 From (1.4) we obtain the upper estimate. On the other hand, from (2.9), (3.9) and (1.7) we obtain the lower estimate. The proof of Theorem 1.2 is completed.

Note 1 The values of C_r can be computed for $r = 1, 2, \cdots$, respectively. For example, $C_1 = 1, C_2 = \frac{1}{2}, C_3 = 1 - \frac{\sqrt{2}}{2}$. We guess that

$$C_r = 1 - 2 \sum_{i=1}^{[\frac{r+1}{2}]} (-1)^{i-1} \Big(\cos \frac{i\pi}{r+1} \Big)^r,$$

where $[x]$ represents the integer part of x.

Note 2 In practice, one often wants to have boundary points as interpolation nodes, i.e.,

$$\mathbf{x} = (-1, x_2, \cdots, x_{r-1}, 1).$$

Then the following question arises: For which sets of points $-1 < c_2 < c_3 < \cdots < c_{r-1} < 1$, we have

$$e(BF, L_{\mathbf{c}}, G) = \inf_{\mathbf{x}=(-1, x_2, \cdots, x_{r-1}, 1)} e(BF, L_{\mathbf{x}}, G). \tag{3.10}$$

Papers [15, 18] considered this problem recently. Obviously, from (2.9) it follows that $\mathbf{c} = (-1, c_2, \cdots, c_{r-1}, 1)$ is the solution of (3.10) for BW_1^r in L_1 if and only if

$$\sup_{-1\le t\le 1}\int_{-1}^{1}|B_{\mathbf{c}}(x,t)|\mathrm{d}x = \min_{\mathbf{x}=(-1,x_2,\cdots,x_{r-1},1)}\sup_{-1\le t\le 1}\int_{-1}^{1}|B_{\mathbf{x}}(x,t)|\mathrm{d}x. \tag{3.11}$$

To each integer $r > 2$, we can compute the solution of (3.10) for BW_1^r in L_1 by using (3.11). But the explicit solution to this problem is an open problem.

Note 3 When $n \neq r$, the values of the sampling numbers and the nth optimal Lagrange interpolation nodes of the problems given by (1.7) and (3.10) for BW_1^r in L_1 are open problems.

Acknowledgement The authors thank the referees for their valuable advices.

References

[1] Byrenheid, G., Dũng, D., Sickel, W. and Ullrich, T., Sampling on energy-norm based sparse grids for the optimal recovery of Sobolev type functions in H^γ, *J. Approx. Theory*, **207**, 2016, 207–231.

[2] Byrenheid, G., Kämmerer, L., Ullrich, T. and Volkmer, T., Tight error bounds for rank-1 lattice sampling in spaces of hybrid mixed smoothness, *Numer. Math.*, **136**(4), 2017, 993–1034.

[3] Byrenheid, G. and Ullrich, T., Optimal sampling recovery of mixed order Sobolev embeddings via discrete Littlewood-Paley type characterizations, *Anal. Mathem.*, **43**(2), 2017, 133–191.

[4] Dũng, D., B-spline quasi-interpolant representations and sampling recovery of functions with mixed s-moothness, *J. Complexity*, **27**, 2011, 541–567.

[5] Dũng, D., Sampling and cubature on sparse grids based on a B-spline quasi-interpolation, *Found. Comp. Math.*, **16**, 2016, 1193–1240.

[6] Dũng, D., B-spline quasi-interpolation sampling representation and sampling recovery in Sobolev spaces of mixed smoothness, *Acta Math. Vietnamica*, **43**, 2018, 83–110.

[7] Dũng, D., Optimal adaptive sampling recovery, *Adv. Comput. Math.*, **34**, 2011, 1–41.

[8] Dũng, D., Temlyakov, V. and Ullrich, T., Hyperbolic Cross Approximation, Advanced Courses in Mathematics CRM Barcelona, Springer Nature Switzerland AG, Barecelona, 2018.

[9] Kudryavtsev, S. N., The best accuracy of reconstruction of finitely smooth functions from their values at a given number of points, *Izv. Math.*, **62**(1), 1998, 19–53.

[10] Novak, E. and Triebel, H., Function spaces in Lipschitz domains and optimal rates of convergence for sampling, *Constr. Approx.*, **23**, 2006, 325–350.

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

[12] Triebel, H., Bases in Function Spaces, Sampling, Discrepancy, Numerical Integration, EMS, Zürich, 2010.

[13] Vybíral, J., Sampling numbers and function spaces, *J. Complexity*, **23**, 2007, 773–792.

[14] Wang, H. and Wang, K., Optimal recovery of Besov classes of generalized smoothness and Sobolev classes on the sphere, *J. Complexity*, **32**, 2016, 40–52.

[15] Xu G. Q. and Wang H., Sample numbers and optimal Lagrange interpolation of Sobolev spaces, *Rocky MT. J. Math.*, **51**(1), 2021, 347–361.

[16] Babaev, S. S. and Hayotov, A. R., Optimal interpolation formulas in $W_2^{(m,m-1)}$ space, *Calcolo*, **56**, 2019, 23–45.

[17] Mastroianni, G. and Occorsio, D., Optimal systems of nodes for Lagrange interpolation on bounded intervals, A survey, *J. Comput. Appl. Math.*, **134**, 2001, 325–341.

[18] Hoang, N. S., On node distributions for interpolation and spectral methods, *Math. Comp.*, **85**, 2016, 667–692.

[19] Szabados, J. and Vèrtesi, P., Interpolation of Functions, World Scientific, Singapore, 1990.

[20] Rack, H. J. and Vajda, R., On optimal quadratic Lagrange interpolation: Extremal node systems with minimal Lebesgue constant via symbolic computation, *Serdica J. Comput.*, **8**, 2014, 71–96.

[21] Rack, H. J. and Vajda, R., Optimal cubic Lagrange interpolation: Extremal node systems with minimal Lebesgue constant, *Stud. Univ. Babes-Bolyai Math.*, **60**(2), 2015, 151–171.

[22] Birkhoff, G. D., General mean value and remainder theorems with applications to mechanical differentiation and quadrature, *Trans. Am. Math. Soc.*, **7**, 1906, 107–136.

[23] Nürnberger, G., Approximation by Spline Functions, Springer-Verlag, Beijing, 1992.

[24] Novak, E. and Woźniakowski, H., Tractability of Multivariate Problems, Volume I: Linear Information, EMS, Zürich, 2008.

[25] Devore, R. A. and Lorentz, G. G., Constructive Approximation, Springer-Verlag, New York, 1993.