Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2014, Article ID 402918, 6 pages
http://dx.doi.org/10.1155/2014/402918

Research Article

ε-Coverings of Hölder-Zygmund Type Spaces on Data-Defined Manifolds

Martin Ehler[1] and Frank Filbir[2,3]

[1] *Department of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria*
[2] *Faculty of Mathematics, Technische Universität München, Boltzmannstraße 3, 85748 Garching, Germany*
[3] *Helmholtz Zentrum München, Ingolstädter Landstraße 1, 85764 Neuherberg, Germany*

Correspondence should be addressed to Martin Ehler; martin.ehler@univie.ac.at

Received 13 February 2014; Accepted 19 May 2014; Published 17 June 2014

Academic Editor: Carlo Cattani

We first determine the asymptotes of the ε-covering numbers of Hölder-Zygmund type spaces on data-defined manifolds. Secondly, a fully discrete and finite algorithmic scheme is developed providing explicit ε-coverings whose cardinality is asymptotically near the ε-covering number. Given an arbitrary Hölder-Zygmund type function, the nearby center of a ball in the ε-covering can also be computed in a discrete finite fashion.

1. Introduction

Data processing in the digital era often deals with finitely many high-dimensional data chunks stemming from measurements that obey some continuous physical model. The implementation and numerical evaluation require estimates on the accuracy of the discretization with respect to the underlying model. As an elementary tool providing accuracy guarantees, we will address ε-coverings of some function spaces related to information theory and machine learning.

As a standard concept in discrete mathematics, the *ε-covering number* $n_\varepsilon(Y)$ is the minimal number of balls of radius ε that cover a compact metric space Y. An arbitrary element in Y can be represented by a nearby center preserving precision up to ε. As such, ε-coverings are also an integral part of approximation theory, especially if Y is some function space. Covering numbers capture the complexity of Y and the approximation aspects are used in many fields such as information theory, statistics, nonparametric density estimation, and machine learning. There are estimates on the asymtotics of the ε-covering numbers of the standard function spaces (cf. [1, 2]), but some fields such as machine learning involve data lying on some manifold, so that target functions are naturally defined on this manifold. To clarify the terminology, we consider smoothness spaces on manifolds as somewhat nonstandard function spaces. It may be possible that the covering number of a function space on some compact Riemannian manifold can be assembled by covering numbers of standard function spaces on Euclidian spaces derived from the charts. However, it is also important to derive explicit ε-coverings whose cardinality is near the benchmark given by the ε-covering number. We believe that explicit coverings may be harder to construct using the charts due to interface problems, and therefore we will not pursue this direction and, instead, we will follow a more global approach.

In general, there is still demand for computing coverings of many discrete and continuous spaces [3]. As an important additional requirement, any covering of a function space needs to come with an algorithmic scheme to determine some function's nearby center in an effective manner. At first sight, the latter seems simple enough as we can take the center whose distance is minimal. However, determining the distance between two functions is eventually a continuous operation, and one is particularly interested in finite methods.

In this paper, we first determine the asymtotics of the ε-covering number for the unit ball of some Hölder-Zygmund type space $Y = \mathscr{C}^s(\mathbb{X})$ on an underlying smooth compact Riemannian manifold $\mathbb{X}$ (without boundary and with nonnegative Ricci curvature). In fact, we determine the asymtotics of the *metric entropy* $\log_2(n_\varepsilon(Y))$, which is the number of bits needed to enumerate the ε-covering (cf. [1]). Moreover, we compute an explicit ε-covering, such that

$$\frac{\log_2(\widehat{n}_\varepsilon)}{\log_2(n_\varepsilon)} \lesssim \left(1 + \log_2\left(\frac{1}{\varepsilon}\right)\right), \tag{1}$$

where $\widehat{n}_\varepsilon$ is the cardinality of the constructed covering and $\lesssim$ means that the left-hand side can be bounded by a generic constant times the right-hand side. Hence, our covering is optimal up to a logarithmic factor by means of the metric entropy. We allow the underlying manifold to be unknown in our scheme and, instead, to be represented through a finite sampling. This sampling must be chosen carefully and is the key to obtaining a finite scheme. The centers of our ε-covering can then be determined through a finite process, and we can measure any function's distance to these centers in a finite manner.

For constructions of ε-coverings on periodic smoothness spaces, for instance, we refer to [4, 5]. The concept of ε-entropy is also closely related to entropy numbers; see [6–8].

The outline of this paper is as follows. In Section 2 we introduce the setting, define the Hölder-Zygmund type space $\mathscr{C}^s(\mathbb{X})$, and determine the metric entropy for its unit ball. An explicit covering is computed in Section 3.

2. Covering Numbers for Hölder-Zygmund Type Spaces

We first fix the setting and list some technical assumptions used throughout the paper. Let $\mathbb{X} \subset \mathbb{R}^d$ be an α-dimensional compact and connected Riemannian manifold without boundary and with nonnegative Ricci curvature, geodesic distance ρ, and μ being the normalized Riemannian volume measure on $\mathbb{X}$; $\{\varphi_k\}_{k=0}^\infty$ are the eigenfunctions of the Laplacian on $\mathbb{X}$, and $\{-\lambda_k^2\}_{k=0}^\infty$ are the corresponding eigenvalues arranged in nonincreasing order, so that $0 = \lambda_0 \le \lambda_1 \le \cdots$. Readers who are not familiar with some terms from differential geometry that are used here may simply think of a "nice" manifold without boundary, such as the sphere, the real projective space, the (real) Grassmann manifold, or more generally compact homogeneous spaces. The above properties ensure certain estimates on the heat kernel on $\mathbb{X}$ (see [9, 10]), which were used in a series of papers [9, 11–13] to develop approximation schemes for smooth functions on the manifold. Here, we will make use of those approximation schemes, but we will keep the technical details at a minimum level.

Let N be a positive integer and most of the time we will restrict ourselves to $N = 2^j$, where j is some nonnegative integer. The space of *diffusion polynomials* up to degree N is

$$\Pi_N := \operatorname{span}\{\varphi_k : \lambda_k \le N\}. \tag{2}$$

Later, we will use the fact that the above conditions imply the following estimate on the Christoffel function:

$$\sum_{\lambda_k \le N} |\varphi_k(x)|^2 \asymp N^\alpha, \quad x \in \mathbb{X},\ N > 0 \tag{3}$$

(cf. [9–11]), so that integration and orthonormality yield $\dim(\Pi_N) \asymp N^\alpha$. Here, the symbol $\asymp$ indicates that each side is bounded by a generic positive constant times the other side.

In traditional scenarios, the accuracy of approximation by polynomials is closely related to the smoothness of the function. Therefore, the accuracy of approximation itself is nowadays considered to be a measurement of smoothness. This viewpoint is particularly useful in our setting because defining smoothness in a classical manner would require more technical details. Here, we define the *Hölder-Zygmund type space* of order $s > 0$ by $\mathscr{C}^s(\mathbb{X}) = \{f \in L_\infty(\mathbb{X}) : \|f\|_{\mathscr{C}^s(\mathbb{X})} < \infty\}$, where its norm is given by

$$\|f\|_{\mathscr{C}^s(\mathbb{X})} := \|f\|_{L_\infty(\mathbb{X})} + \sup_{N\ge1} N^s E(f, \Pi_N, L_\infty(\mathbb{X})), \tag{4}$$

with $E(f,\Pi_N,L_\infty(\mathbb{X})) := \inf_{g\in\Pi_N}\|f-g\|_{L_\infty(\mathbb{X})}$. Hence, $f \in L_\infty(\mathbb{X})$ is contained in the Hölder-Zygmund type space if and only if it can be approximated by Π_N at rate N^{-s}. Since the eigenfunctions $\{\varphi_k\}_{k=0}^\infty$ are known to be smooth and we consider the L_∞-norm, each function in $\mathscr{C}^s(\mathbb{X})$ has a continuous representative and point evaluation makes sense. The unit ball in $\mathscr{C}^s(\mathbb{X})$ is denoted by $\overline{\mathscr{C}^s}(\mathbb{X}) := \{f \in L_\infty(\mathbb{X}) : \|f\|_{\mathscr{C}^s(\mathbb{X})} \le 1\}$. To compute its covering number, we first establish compactness. Since $\mathscr{C}^s(\mathbb{X})$ is not finite-dimensional, $\overline{\mathscr{C}^s}(\mathbb{X})$ is not compact in the Hölder-Zygmund type space, but we consider it as a subspace of $L_\infty(\mathbb{X})$.

Lemma 1. *The set $\overline{\mathscr{C}^s}(\mathbb{X})$ is compact in $L_\infty(\mathbb{X})$.*

The compactness of this embedding can be derived from (4) by abstract arguments involving Kolmogorov numbers (cf. [6]). Here, we provide a simple elementary proof for the sake of completeness.

Proof. We aim to verify that any sequence $(f_j)_{j=1}^\infty \subset \overline{\mathscr{C}^s}(\mathbb{X})$ must have an accumulation point in this set. Since each space Π_N is finite-dimensional, there are $g_{j,N} \in \Pi_N$, such that $\|f_j\|_{L_\infty(\mathbb{X})} + \sup_{N\ge1} N^s \|f_j - g_{j,N}\|_{L_\infty(\mathbb{X})} \le 1$. The latter implies that $\|g_{j,N}\|_{L_\infty(\mathbb{X})}$ is bounded for all j and N. Thus, there is $g_1 \in \Pi_1$ such that the subsequence $(g_{\pi_1(j),1})_{j=1}^\infty$ converges towards g_1. For any $N = 1, 2, \ldots$, we can recursively construct $g_N \in \Pi_N$ such that

$$g_{\pi_N(j),k} \longrightarrow g_k, \quad \forall k = 1, \ldots, N, \tag{5}$$

and $(g_{\pi_N(j),k})_{n=1}^\infty$ is a subsequence of $(g_{\pi_{N-1}(j),k})_{n=1}^\infty$. For $N' \ge N$, this construction yields that $(g_{\pi_{N'}(j),N})_{j=1}^\infty$ is a subsequence of $(g_{\pi_N(j),N})_{j=1}^\infty$, so that we derive

$$\begin{aligned} \|g_N - g_{N'}\|_{L_\infty(\mathbb{X})} &= \lim_{n\to\infty} \left\|g_{\pi_{N'}(j),N} - g_{\pi_{N'}(j),N'}\right\|_{L_\infty(\mathbb{X})} \\ &\le \lim_{n\to\infty} \left\|g_{\pi_{N'}(j),N} - f_{\pi_{N'}(j)}\right\|_{L_\infty(\mathbb{X})} \\ &\quad + \left\|f_{\pi_{N'}(j)} - g_{\pi_{N'}(j),N'}\right\|_{L_\infty(\mathbb{X})} \\ &\le N^{-s} + N'^{-s}. \end{aligned} \tag{6}$$

Therefore, $(g_N)_{N=1}^{\infty}$ is a Cauchy sequence and, hence, converges towards some $g \in L_\infty(\mathbb{X})$. Standard calculations reveal that g is an accumulation point of $(f_j)_{j=1}^{\infty}$ and is contained in $\overline{\mathscr{C}^s}(\mathbb{X})$, which concludes the proof. $\square$

We can now derive the asymptotes of the ε-covering number of $\overline{\mathscr{C}^s}(\mathbb{X})$ in $L_\infty(\mathbb{X})$.

Theorem 2. *If $s > 0$ is fixed and $0 < \varepsilon \le 1$, then*

$$\log_2\left(n_\varepsilon\left(\overline{\mathscr{C}^s}(\mathbb{X})\right)\right) \asymp \varepsilon^{-\alpha/s} \tag{7}$$

holds, where the generic constants do not depend on ε.

Analogous results can be derived for similar concepts such as different types of n-widths of functions spaces (cf. [14–16]). Theorem 2 and its proof are rather classical and can be derived from [17]. To guide the interested reader, we will provide the outline of the proof that is based on a general Banach space result and is also used in [18, Theorem 4.1]. Let X be a Banach space and let $\{\phi_k\}_{k=1}^{\infty} \subset X$ be a sequence of linearly independent elements whose linear span is dense in X, and define $X_k := \operatorname{span}\{\phi_1,\ldots,\phi_k\}$ with $X_0 = \{0\}$. Let $\{\delta_k\}_{k=0}^{\infty}$ be a nonincreasing sequence of positive numbers with $\lim_{k\to\infty}\delta_k = 0$. The full approximation space is

$$\begin{aligned}&\mathscr{A}\left(X;\{\delta_k\}_{k=0}^{\infty},\{\phi_k\}_{k=1}^{\infty}\right)\\&\quad:=\left\{f\in X : E\left(f,X_k,X\right)\le\delta_k,\text{for } k=0,1,\ldots\right\}.\end{aligned} \tag{8}$$

A proof similar to Lemma 1 yields that this space is compact, and we can formulate the result from Banach space theory that goes back to Lorentz in [17].

Theorem 3 (see [19, Theorem 3.3]). *Let $\{\delta_k\}_{k=0}^{\infty}$ be a nonincreasing sequence of positive numbers such that $\delta_{2k} \le c\delta_k$, for $k = 1,2,\ldots$ and some constant $c \in (0,1)$. For $\ell \ge 0$, let $M_\ell := \min\{k : \delta_k \le e^{-\ell}\}$. If n_ε denotes the ε-covering number of $\mathscr{A}(X;\{\delta_k\}_{k=0}^{\infty},\{\phi_k\}_{k=1}^{\infty})$ in X, then one has, for $0 < \varepsilon \le 1$,*

$$\log_2\left(n_\varepsilon\right) \asymp \sum_{\ell=1}^{L} M_\ell, \tag{9}$$

where $L := 2 + \lfloor\log(1/\varepsilon)\rfloor$.

At this point our preparations are complete.

Proof of Theorem 2. We aim to apply Theorem 3 with the function system $\{\varphi_k\}_{k=0}^{\infty}$ and with X being the closure of $\bigcup_{N=1}^{\infty}\Pi_N$ in $L_\infty(\mathbb{X})$. There, the index set is supposed to start with $k = 1$, so we set $\phi_k = \varphi_{k-1}$, $k = 1,2,\ldots$. To define the sequence $\{\delta_k\}_{k=0}^{\infty}$, we need some preparations. As pointed out before, integrating (3) over $\mathbb{X}$ yields $\dim(\Pi_N) \asymp N^\alpha$. By using $X_k := \operatorname{span}\{\varphi_0,\ldots,\varphi_{k-1}\}$, we derive, for $N^\alpha \le k \le (2N)^\alpha$,

$$(2N)^s E\left(f,\Pi_{2N},X\right) \lesssim k^{s/\alpha}E\left(f,X_k,X\right) \lesssim N^s E\left(f,\Pi_N,X\right). \tag{10}$$

Therefore, there are constants $C_i \ge 1$, for $i = 1,2$, such that the definitions $\delta_{1;0} = 1/2$, $\delta_{1;k} := (2C_1)^{-1}k^{-s/\alpha}$, and $\delta_{2;0} = C_2$, $\delta_{2;k} := C_2 k^{-s/\alpha}$, lead to

$$\begin{aligned}&\mathscr{A}\left(X;\{\delta_{1;k}\}_{k=0}^{\infty},\{\phi_k\}_{k=1}^{\infty}\right)\\&\quad\subset\overline{\mathscr{C}^s}(\mathbb{X})\subset\mathscr{A}\left(X;\{\delta_{2;k}\}_{k=0}^{\infty},\{\phi_k\}_{k=1}^{\infty}\right),\end{aligned} \tag{11}$$

which also yields

$$\begin{aligned}&n_\varepsilon\left(\mathscr{A}\left(X;\{\delta_{1;k}\}_{k=0}^{\infty},\{\phi_k\}_{k=1}^{\infty}\right)\right)\\&\quad\le n_\varepsilon\left(\overline{\mathscr{C}^s}(\mathbb{X})\right)\le n_\varepsilon\left(\mathscr{A}\left(X;\{\delta_{2;k}\}_{k=0}^{\infty},\{\phi_k\}_{k=1}^{\infty}\right)\right).\end{aligned} \tag{12}$$

Since $\delta_{i;2k} \le c\delta_{i;k}$, for $c := 2^{-s/\alpha} \in (0,1)$, we can apply Theorem 3. According to [18, Lemma 4.1], $\sum_{\ell=1}^{L} M_\ell \asymp e^{L\alpha/s}$, so that the choice of L in (9) implies (7). $\square$

Remark 4. The proof of Theorem 2 discovers that (7) also holds under much weaker conditions, and we have only used the fact that there is a sequence of linearly independent functions $\{\varphi_k\}_{k=0}^{\infty}$, so that the polynomial spaces in (2) satisfy $\dim(\Pi_N) \asymp N^\alpha$.

3. Near Optimal Covering

This section is dedicated to constructing our covering of the unit ball in the Hölder-Zygmund type space, which is based on localized summation kernels as developed in a series of papers [9, 11–13]. We first need some preparations. A Borel probability measure ν on $\mathbb{X}$ is called a *quadrature measure* of order N if

$$\int_{\mathbb{X}} f(x)\,g(x)\,d\mu(x) = \int_{\mathbb{X}} f(x)\,g(x)\,d\nu(x), \quad \forall f,g\in\Pi_N. \tag{13}$$

Note that our setting yields that there is a constant $a > 0$ such that $f\cdot g \in \Pi_{aN}$ for all $f,g \in \Pi_N$ and all N (cf. [11, Theorem A.1]); see also [20] for homogeneous spaces. The existence of quadrature measures with finite support is proved for fairly general smooth Riemannian manifolds in [11], where a construction procedure is outlined. In fact, the support of ν can be chosen to be contained in any sufficiently dense finite sampling $\{x_\ell\}_{\ell=1}^{m}$ of $\mathbb{X}$, so that ν can be identified with $\{x_\ell\}_{\ell=1}^{m}$ and nonnegative weights $\{\omega_\ell\}_{\ell=1}^{m}$ satisfying $\nu(\{x_\ell\}) = \omega_\ell$. Examples on the sphere, for instance, are given in [21].

The results in [11] yield that we can even choose a sequence $(\nu_N)_{N=1}^{\infty}$ of quadrature measures of order N, respectively, such that $\#\operatorname{supp}(\nu_N) \lesssim N^\alpha$. For the remaining part of the paper, we will suppose that this estimate holds and we define, for $N = 2^j$,

$$\begin{aligned}\sigma_N(f) &:= \int_{\mathbb{X}} f(y)\,K_N(\cdot,y)\,d\nu_N(y),\\ \text{where } K_N(x,y) &= \sum_{k=0}^{\infty} h\left(\frac{\lambda_k}{N}\right)\varphi_k^*(y)\,\varphi_k(x),\end{aligned} \tag{14}$$

where $h : \mathbb{R}_{\geq 0} \to \mathbb{R}$ is an infinitely often differentiable and nonincreasing function with $h(t) = 1$ for $t \leq 1/2$ and $h(t) = 0$ for $t \geq 1$. Although we will not explicitly use it in the present paper, we want to point out that many advantageous properties of σ_N are steered by the so-called localization of the kernel K_N; that is, for fixed $S > \alpha$ and all $x \neq y$ with $N = 1, 2, \ldots$,

$$\left|K_N(x, y)\right| \lesssim \frac{N^{\alpha-S}}{\rho(x, y)^S}. \tag{15}$$

See [12, 13]. Later, we will apply

$$\sup_{x\in\mathbb{X}} \int_{\mathbb{X}} \left|K_N(x, y)\right| d\left|\nu_N\right|(y) \lesssim 1 \tag{16}$$

(cf. [11]). Those estimates are used in [12, 13] to characterize the Hölder-Zygmund type smoothness by means of σ_N.

Theorem 5. *Assume that $(\nu_N)_{N=1}^{\infty}$ is a family of quadrature measures of order N, respectively. Then, for all $f \in \mathscr{C}^s(\mathbb{X})$, one has*

$$\left\|f - \sigma_N(f)\right\|_{L_\infty(\mathbb{X})} \lesssim N^{-s}\|f\|_{\mathscr{C}^s(\mathbb{X})}, \tag{17}$$

where the generic constants do not depend on N or f. On the other hand, if, for $f \in L_p(\mathbb{X})$, there are generic constants not depending on N such that $\|f - \sigma_N(f)\|_{L_\infty(\mathbb{X})} \lesssim N^{-s}$ holds, then $f \in \mathscr{C}^s(\mathbb{X})$.

Next, by using $h(\lambda_k/N)h(\lambda_k/2N) = h(\lambda_k/N)$ and applying the quadrature property of ν_N, a straightforward calculation yields

$$\sigma_N(f, x) = \int_{\mathbb{X}} \sigma_N(f, y) K_{2N}(x, y) d\nu_N(y). \tag{18}$$

For some fixed $S > 1$, we define the actual approximation by

$$\begin{gathered}\sigma_N^{\circ}(f, x) := \int_{\mathbb{X}} I_N(f, y) K_{2N}(x, y) d\nu_N(y), \\ \text{where } I_N(f, y) = N^{-S}\left\lfloor N^S \sigma_N(f, y)\right\rfloor.\end{gathered} \tag{19}$$

In other words, we replace $\sigma_N(f, y)$ in (18) with a number on the grid $(1/N^S)\mathbb{Z}$. We define the following collection:

$$\mathscr{M}_{S,N} := \left\{\sigma_N^{\circ}(f) : f \in \overline{\mathscr{C}^s}(\mathbb{X})\right\}, \tag{20}$$

which induces a covering of $\overline{\mathscr{C}^s}(\mathbb{X})$ in $L_\infty(\mathbb{X})$.

Theorem 6. *For fixed $s > 0$ and $S > \max(1, s)$, one applies the discretization (19). Then, there is a constant $c > 0$ such that, for all $f \in \overline{\mathscr{C}^s}(\mathbb{X})$, $\|f - \sigma_N^{\circ}(f)\|_{L_\infty(\mathbb{X})} \leq cN^{-s}$ holds. Thus, for $cN^{-s} = \varepsilon \leq 1$, the collection $\mathscr{M}_{S,N}$ induces an ε-covering of $\overline{\mathscr{C}^s}(\mathbb{X})$ in $L_\infty(\mathbb{X})$. Its cardinality $\widehat{n}_\varepsilon$ satisfies*

$$\log_2(\widehat{n}_\varepsilon) \lesssim \varepsilon^{-\alpha/s}\left(1 - \log_2(\varepsilon)\right), \tag{21}$$

where the generic constant does not depend on ε.

Proof of Theorem 6. The triangle inequality yields

$$\begin{aligned}\left\|f - \sigma_N^{\circ}(f)\right\|_{L_\infty(\mathbb{X})} &\lesssim \left\|f - \sigma_N(f)\right\|_{L_\infty(\mathbb{X})} \\ &\quad + \left\|\sigma_N(f, \nu_N) - \sigma_N^{\circ}(f)\right\|_{L_\infty(\mathbb{X})}.\end{aligned} \tag{22}$$

Since Theorem 5 implies $\|f - \sigma_N(f)\|_{L_\infty(\mathbb{X})} \lesssim N^{-s}\|f\|_{\mathscr{C}^s(\mathbb{X})} = N^{-s}$, we only need to take care of the term on the farmost right. The quantization (19) immediately yields

$$\left|\sigma_N(f, y) - I_N(f, y)\right| \leq N^{-S}, \quad \forall y \in \operatorname{supp}(\nu_N), \tag{23}$$

so that (18) and (16) imply

$$\begin{aligned}&\left\|\sigma_N(f) - \sigma_N^{\circ}(f)\right\|_{L_\infty(\mathbb{X})} \\ &\quad = \left\|\int_{\mathbb{X}} (\sigma_N(f, y) - I_N(f, y)) K_N(\cdot, y) d\nu_N(y)\right\|_{L_\infty(\mathbb{X})} \\ &\quad \lesssim N^{-S} \leq N^{-s}.\end{aligned} \tag{24}$$

Hence, we have derived the estimate on $\|f - \sigma_N^{\circ}(f)\|_{L_\infty(\mathbb{X})}$.

To tackle (21), we apply (23), which yields

$$\left|I_N(f, y)\right| \lesssim \left\|\sigma_N(f)\right\|_{L_\infty(\mathbb{X})}. \tag{25}$$

According to [13, Theorem 5.1], $\|\sigma_N(f)\|_{L_\infty(\mathbb{X})} \lesssim \|f\|_{L_\infty(\mathbb{X})}$ holds. Since f is contained in the ball of radius 1, we see that

$$\left|I_N(f, y)\right| \lesssim 1. \tag{26}$$

Thus, the number of possible values of $I_N(f, y)$ for fixed y is at most $c_1 N^S$, where $c_1 \geq 1$ is a positive constant. Note that we can assume that $c_1 N^S \geq 1$ because, otherwise, $I_N(f, y)$ would be zero. Since $\#\operatorname{supp}(\nu_N) \lesssim N^\alpha$, we have $\#\{I_N(f, y) : y \in \operatorname{supp}(\nu_N)\} \lesssim N^\alpha$. Therefore, we have $\widehat{n}_\varepsilon \leq (c_1 N^S)^{c_2 N^\alpha}$, for some positive constant c_2. By using $cN^{-s} = \varepsilon \leq 1$, we obtain

$$\begin{aligned}\log_2(\widehat{n}_\varepsilon) &= c_2 c^{\alpha/s} \varepsilon^{-\alpha/s} \log_2\left(c_1\left(\frac{c}{\varepsilon}\right)^{S/s}\right) \\ &\lesssim \varepsilon^{-\alpha/s} \log_2\left(\frac{(c_1)^{s/S} c}{\varepsilon}\right) \\ &\lesssim \varepsilon^{-\alpha/s} \log_2\left(\frac{(c_1/\varepsilon)^{s/S} c}{\varepsilon}\right) \\ &\lesssim \varepsilon^{-\alpha/s}\left(1 + \log_2\left(\frac{1}{\varepsilon}\right)\right),\end{aligned} \tag{27}$$

which concludes the proof. □

According to Theorems 2 and 6, the ε-covering number n_ε of $\overline{\mathscr{C}^s}(\mathbb{X})$ and the number $\widehat{n}_\varepsilon$ of ε-balls induced by $\mathscr{M}_{S,N}$ satisfy

$$\frac{\log_2(\widehat{n}_\varepsilon)}{\log_2(n_\varepsilon)} \lesssim \left(1 + \log_2\left(\frac{1}{\varepsilon}\right)\right). \tag{28}$$

Therefore, our scheme is optimal up to a logarithmic factor by means of the metric entropy.

Our results are also related to the field of manifold learning, in which a function must be reconstructed from finite training data (cf. [22–25]). When actually applying our scheme, we first acquire a set of samples $\{x_\ell\}_{\ell=1}^m$ sufficiently well covering $\mathbb{X}$ and we also need the function values $\{f(x_\ell)\}_{\ell=1}^m$, which altogether build the training data. Next, we compute a quadrature measure ν_N for some maximal N such that $\operatorname{supp}(\nu_N) \subset \{x_\ell\}_{\ell=1}^m$; see [11, 21] for an algorithm. Here, we need that the sample points $\{x_\ell\}_{\ell=1}^m$ are well distributed and larger N require more samples. An element in $\mathscr{M}_{S,N}$ that is ε-close to f is simply given by $\sigma_N^\circ(f)$, whose computation only requires knowledge of f and $\{\varphi_k : \lambda_k \le N\}$ on the finite set $\operatorname{supp}(\nu_N)$; see (14) and (19). In other words, we do not need to know the entire manifold but only the finite sampling of the training data $\{x_\ell\}_{\ell=1}^m$, the sampling of the target function $\{f(x_\ell)\}_{\ell=1}^m$, and, more delicately, the sampling of the eigenfunctions $\{\varphi_k(x_\ell) : \lambda_k \le N, \ell = 1, \dots, m\}$ of the Laplacian. Those eigenfunctions, however, are not explicitly known except for few special cases, such as the sphere, projective space, the Grassmann manifold, and few more. Fortunately, approximation of those eigenfunctions is a common procedure in manifold learning. Computational schemes are based on the graph Laplacian to be built from the training data and, at least under suitable assumptions, converging towards the Laplacian on the manifold when the cardinality of the data increases (cf. [26–28] and references therein). Those schemes approximately sample the first few eigenfunctions on the training data. Thus, our proposed approach is indeed fully discrete and computationally feasible even if the eigenfunctions $\{\varphi_k\}_{k=0}^\infty$ are not explicitly known. In fact, the manifold itself can be unknown. As long as $\mathbb{X}$ satisfies the theoretical assumptions, it is simply represented by means of a finite sample.

Remark 7. The technical assumptions on the manifold $\mathbb{X}$ and the function system $\{\varphi_k\}_{k=0}^\infty$ imply certain estimates on the heat kernel on $\mathbb{X}$ (see [9, 10]), mainly used to ensure that the localization property (15) holds (cf. [12, 13]). Our assumptions also imply the existence of quadrature measures ν_N and that $f \cdot g \in \Pi_{aN}$ for all $f, g \in \Pi_N$ and some constant $a > 0$. These items lead to the characterization of the Hölder-Zygmund type space by means of σ_N in Theorem 5. Moreover, the family $(\nu_N)_{N=1}^\infty$ can be chosen with finite support, in fact with $\#\operatorname{supp}(\nu_N) \lesssim N^\alpha$ (cf. [11]). Theorem 5 and $\#\operatorname{supp}(\nu_N) \lesssim N^\alpha$ are indeed the two main ingredients of the poof of our results in Theorem 6.

Remark 8. The reader familiar with the approximation scheme developed in [9, 11–13] may expect that the presented results can be generalized to a wider class of Besov spaces on metric spaces. This is indeed true but requires more technical details and does not lead to a fully discrete scheme in the end. Here, we intended to emphasize the main ideas by keeping technical details at a minimum level and to focus on the development of a fully discrete covering algorithm. The more general approach will be described elsewhere.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

Martin Ehler has been funded by the Vienna Science and Technology Fund (WWTF) through Project VRG12-009. The research of Frank Filbir was partially funded by the Deutsche Forschungsgemeinschaft Grant FI 883/3-1. Both authors thank H. N. Mhaskar for many fruitful discussions.

References

[1] A. N. Kolmogorov and V. M. Tihomirov, "ε-entropy and ε-capacity of sets in function spaces," *Uspekhi Matematicheskikh Nauk*, vol. 14, no. 2, pp. 3–86, 1959, English translation in American Mathematical Society, vol. 17, pp. 277–364, 1961.

[2] A. G. Vituškin, *Theory of the Transmission and Processing of Information*, Pergamon, 1961.

[3] A. Schürmann and F. Vallentin, "Computational approaches to lattice packing and covering problems," *Discrete & Computational Geometry*, vol. 35, no. 1, pp. 73–116, 2006.

[4] D. Dung, "Non-linear approximations using sets of finite cardinality or finite pseudo-dimension," *Journal of Complexity*, vol. 17, no. 2, pp. 467–492, 2001.

[5] V. N. Temlyakov, "Estimates for the asymptotic characteristics of classes of functions with bounded mixed derivative or difference," *Proceedings of the Steklov Institute of Mathematics*, vol. 189, pp. 161–197, 1990.

[6] B. Carl and I. Stephani, *Entropy, Compactness and the Approximation of Operators*, Cambridge University Press, Cambridge, UK, 1990.

[7] D. E. Edmunds and H. Triebel, *Function Spaces, Entropy Numbers, Differential Operators*, vol. 120, Cambridge University Press, Cambridge, UK, 1996.

[8] D. D. Haroske and H. Triebel, *Distributions, Sobolev Spaces, Elliptic Equations*, European Mathematical Society, 2008.

[9] C. K. Chui and H. N. Mhaskar, "Smooth function extension based on high dimensional unstructured data," *Mathematics of Computation*, 2013.

[10] E. B. Davies, "L^p spectral theory of higher-order elliptic differential operators," *The Bulletin of the London Mathematical Society*, vol. 29, no. 5, pp. 513–546, 1997.

[11] F. Filbir and H. N. Mhaskar, "Marcinkiewicz-Zygmund measures on manifolds," *Journal of Complexity*, vol. 27, no. 6, pp. 568–596, 2011.

[12] M. Maggioni and H. N. Mhaskar, "Diffusion polynomial frames on metric measure spaces," *Applied and Computational Harmonic Analysis*, vol. 24, no. 3, pp. 329–353, 2008.

[13] H. N. Mhaskar, "Eignets for function approximation on manifolds," *Applied and Computational Harmonic Analysis*, vol. 29, no. 1, pp. 63–87, 2010.

[14] D. Dũng and T. Ullrich, "n-widths and ε-dimensions for high-dimensional approximations," *Foundations of Computational Mathematics*, vol. 13, no. 6, pp. 965–1003, 2013.

[15] E. Novak, "Optimal recovery and n-widths for convex classes of functions," *Journal of Approximation Theory*, vol. 80, no. 3, pp. 390–408, 1995.

[16] A. Pinkus, *n-Widths in Approximation Theory*, Springer, Berlin, Germany, 1985.

[17] G. G. Lorentz, "Metric entropy and approximation," *Bulletin of the American Mathematical Society*, vol. 72, pp. 903–937, 1966.

[18] H. N. Mhaskar, "On the representation of smooth functions on the sphere using finitely many bits," *Applied and Computational Harmonic Analysis*, vol. 18, no. 3, pp. 215–233, 2005.

[19] G. G. Lorentz, M. V. Golitschek, and Y. Makovoz, *Constructive Approximation, Advanced Problems*, Springer, New York, NY, USA, 1996.

[20] D. Geller and I. Z. Pesenson, "Band-limited localized Parseval frames and Besov spaces on compact homogeneous manifolds," *Journal of Geometric Analysis*, vol. 21, no. 2, pp. 334–371, 2011.

[21] Q. T. Le Gia and H. N. Mhaskar, "Localized linear polynomial operators and quadrature formulas on the sphere," *SIAM Journal on Numerical Analysis*, vol. 47, no. 1, pp. 440–466, 2009.

[22] M. Belkin and P. Niyogi, "Semi-supervised learning on riemannian manifolds," *Machine Learning*, vol. 56, no. 1–3, pp. 209–239, 2004.

[23] M. Belkin, P. Niyogi, and V. Sindhwani, "Manifold regularization: a geometric framework for learning from labeled and unlabeled examples," *Journal of Machine Learning Research*, vol. 7, pp. 2399–2434, 2006.

[24] M. Gavish, B. Nadler, and R. R. Coifman, "Multiscale wavelets on trees, graphs and high dimensional data: theory and applications to semi supervised learning," in *Proceedings of the 27th International Conference on Machine Learning (ICML '10)*, pp. 367–374, June 2010.

[25] A. D. Szlam, M. Maggioni, and R. R. Coifman, "Regularization on graphs with function-adapted diffusion processes," *Journal of Machine Learning Research*, vol. 9, pp. 1711–1739, 2008.

[26] M. Belkin and P. Niyogi, "Convergence of Laplacian eigenmaps," in *Proceedings of the NIPS*, B. Schölkopf, J. C. Platt, and T. Hoffman, Eds., pp. 129–136, MIT Press, 2006.

[27] B. Nadler, S. Lafon, R. R. Coifman, and I. G. Kevrekidis, "Diffusion maps, spectral clustering and Eigen functions of Fokker-Planck operators," in *Advances in Neural Information Processing Systems*, Y. Weiss, B. Schölkopf, and J. Platt, Eds., vol. 18, MIT Press, Cambridge, Mass, USA, 2006.

[28] U. von Luxburg, M. Belkin, and O. Bousquet, "Consistency of spectral clustering," *The Annals of Statistics*, vol. 36, no. 2, pp. 555–586, 2008.