



**MACQUARIE**  
University  
SYDNEY · AUSTRALIA

## Macquarie University PURE Research Management System

---

**This is the accepted author manuscript version of an article published as:**

Brandolini, L., Chen, W. W. L., Colzani, L., Gigante, G., & Travaglini, G. (2019). Discrepancy and numerical integration on metric measure spaces. *Journal of Geometric Analysis*, 29(1), 328-369.

**Access to the published version:** <https://doi.org/10.1007/s12220-018-9993-6>

This is a post-peer-review, pre-copyedit version of an article published in *Journal of Geometric Analysis*. The final authenticated version is available online at:  
<https://doi.org/10.1007/s12220-018-9993-6>

## Discrepancy and numerical integration on metric measure spaces

Luca Brandolini · William W. L. Chen ·  
Leonardo Colzani · Giacomo Gigante ·  
Giancarlo Travaglino

Received: date / Accepted: date

**Abstract** We study here the error of numerical integration on metric measure spaces adapted to a decomposition of the space into disjoint subsets. We consider both the error for a single given function, and the worst case error for all functions in a given class of potentials. The main tools are the classical Marcinkiewicz–Zygmund inequality and *ad hoc* definitions of function spaces on metric measure spaces. The same techniques are used to prove the existence of point distributions in metric measure spaces with small  $L^p$  discrepancy with respect to certain classes of subsets, for example metric balls.

**Keywords** Discrepancy · numerical integration · metric measure spaces

**Mathematics Subject Classification (2010)** 65D30 · 11K38

---

L. Brandolini

Dipartimento di Ingegneria Gestionale, dell'Informazione e della Produzione, Università degli Studi di Bergamo, Viale Marconi 5, 24044 Dalmine, Bergamo, Italia.

E-mail: luca.brandolini@unibg.it

W.W.L. Chen

Department of Mathematics, Macquarie University, Sydney, NSW 2109, Australia.

E-mail: william.chen@mq.edu.au

L. Colzani

Dipartimento di Matematica e Applicazioni, Università di Milano–Bicocca, Via R.Cozzi 55, 20125 Milano, Italia.

E-mail: leonardo.colzani@unimib.it

G. Gigante

Dipartimento di Ingegneria Gestionale, dell'Informazione e della Produzione, Università degli Studi di Bergamo, Viale Marconi 5, 24044 Dalmine, Bergamo, Italia.

E-mail: giacomo.gigante@unibg.it

G. Travaglino

Dipartimento di Matematica e Applicazioni, Università di Milano–Bicocca, Via R.Cozzi 55, 20125 Milano, Italia.

E-mail: giancarlo.travaglino@unimib.it

## 1 Introduction

The starting point of this research is a simplified version of Lemma 5 in [16] which gives an upper bound of higher norms of the discrepancy of a random set of points in the unit square  $[0, 1]^2$ , treated as a torus. Let  $N = M^2$  and consider a random set of  $N$  points  $\mathcal{P}$  as follows: Split the unit square into  $N$  small squares  $\{S_j\}_{j=1}^N$  of area  $N^{-1}$  in the usual way. In each small square  $S_j$  there is a random point  $x_j$ , uniformly distributed in the small square, independently of the distribution of all the other random points in the other small squares.

Suppose that  $\mathcal{B}$  is a convex set in  $[0, 1]^2$ . Let  $\mathcal{J}$  denote the set of all values of  $j$  for which the small squares  $S_j$  intersect the boundary  $\partial\mathcal{B}$  of  $\mathcal{B}$ . Then it is easy to see that the cardinality of  $\mathcal{J}$ ,  $|\mathcal{J}|$ , is  $O(M)$ . For each  $j \in \mathcal{J}$ , write

$$\xi_j = \chi_{\mathcal{B}}(x_j) = \begin{cases} 1 & \text{if } x_j \in \mathcal{B}, \\ 0 & \text{otherwise,} \end{cases}$$

and let  $\eta_j = \xi_j - \mathbb{E}\xi_j$ . Then,  $|\eta_j| \leq 1$  and  $\mathbb{E}\eta_j = 0$ . Furthermore if we define the discrepancy as

$$D[\mathcal{P}, \mathcal{B}] := \frac{1}{N} \sum_{j=1}^N \chi_{\mathcal{B}}(x_j) - |\mathcal{B}|, \quad (1)$$

then

$$D[\mathcal{P}, \mathcal{B}] = \frac{1}{N} \sum_{j \in \mathcal{J}} \eta_j.$$

We now want to estimate  $\mathbb{E}(|D[\mathcal{P}, \mathcal{B}]|^p)$  from above, where  $p$  is an even positive integer. Note first that

$$|D[\mathcal{P}, \mathcal{B}]|^p = \frac{1}{N^p} \sum_{j_1 \in \mathcal{J}} \cdots \sum_{j_p \in \mathcal{J}} \eta_{j_1} \cdots \eta_{j_p}$$

and so

$$\mathbb{E}(|D[\mathcal{P}, \mathcal{B}]|^p) = \frac{1}{N^p} \sum_{j_1 \in \mathcal{J}} \cdots \sum_{j_p \in \mathcal{J}} \mathbb{E}(\eta_{j_1} \cdots \eta_{j_p}) \quad (2)$$

The random variables  $\eta_j$ , where  $j \in \mathcal{J}$ , are independent because the distribution of the random points are independent of each other. If one of  $j_1, \dots, j_p$ , say  $j_i$ , is different from all the others, then

$$\mathbb{E}(\eta_{j_1} \cdots \eta_{j_p}) = \mathbb{E}(\eta_{j_i}) \mathbb{E}(\eta_{j_1} \cdots \eta_{j_{i-1}} \eta_{j_{i+1}} \cdots \eta_{j_p}) = 0.$$

It follows that the only non-zero contribution to the sum (2) comes from those terms where each of  $j_1, \dots, j_p$  appears more than once. It can be shown that the major contribution comes when they appear in pairs, and there are

$$O_p \left( \binom{|\mathcal{J}|}{p/2} \right) = O_p(|\mathcal{J}|^{p/2}) = O_p(M^{p/2}) = O_p(N^{p/4})$$

such pairs. Bounding each of such terms  $\mathbb{E}(\eta_{j_1} \cdots \eta_{j_p})$  trivially by 1, we obtain the estimate

$$\mathbb{E}(|D[\mathcal{P}, \mathcal{B}]|^p) = O_p(N^{-3p/4}). \quad (3)$$

This result is the first step towards the proof of the existence of point sets with small  $L^p$  discrepancy with respect to, say, all discs in the square. Indeed, let  $B(x, r)$  be the ball centered in the point  $x$  and with radius  $r$ . Then, an application of the above estimate to the sets  $B(x, r)$  gives

$$\int_0^{1/2} \int_{[0,1]^2} \mathbb{E}(|D[\mathcal{P}, B(x, r)]|^p) dx dr = O_p(N^{-3p/4}) \quad (4)$$

and Fubini's theorem immediately implies the existence of an  $N$ -point set  $\mathcal{P}$  such that

$$\left( \int_0^{1/2} \int_{[0,1]^2} |D[\mathcal{P}, B(x, r)]|^p dx dr \right)^{1/p} = O_p(N^{-3/4}). \quad (5)$$

By the monotonicity of the  $L^p$  norms, one obtains these estimates for all  $p < +\infty$ .

This argument can be easily extended to a very general setting. In some sense, all that one needs for the argument to work is

1. a partition of the ambient space into  $N$  parts with the same measure and similar diameter (the analog of the "small squares" in the previous argument);
2. a collection of sets with uniformly regular boundary, in such a way that the cardinality of the collection of indices  $\mathcal{J}$  can be controlled uniformly in terms of the diameter of the "small squares".

We will therefore be able to replace the unit square with a compact Riemannian manifold, or more generally, with metric measure spaces  $\mathcal{M}$  having finite measure and with the property that for any integer  $N$  the space can be partitioned as required in point 1 above. By a recent result [21] this property holds under very general hypotheses.

In fact, we can replace the characteristic function of the set  $\mathcal{B}$  with more general functions, so that our results are actually results on numerical integration. Consider the integral

$$\int_{\mathcal{M}} f(x) dx$$

of a function  $f(x)$  over the metric measure space  $\mathcal{M}$  with finite measure  $dx$  and distance between two points  $x$  and  $y$  denoted with  $|x - y|$ , and the Riemann sums

$$\sum_{j=1}^N \omega_j f(x_j),$$

where  $\{x_j\}_{j=1}^N$  are nodes in  $\mathcal{M}$  and  $\{\omega_j\}_{j=1}^N$  are given weights. We are interested in the rate of decay of the error

$$\sum_{j=1}^N \omega_j f(x_j) - \int_{\mathcal{M}} f(x) dx$$

as  $N \rightarrow +\infty$ . This decay depends on the smoothness of the function  $f(x)$ , the weights  $\{\omega_j\}_{j=1}^N$  and the distribution of the nodes  $\{x_j\}_{j=1}^N$  in  $\mathcal{M}$ . For references on this problem when the metric space is a torus, a sphere, or more generally a compact Riemannian manifold, see, for example, [8], [9], [10], [11], [12], [13], [14],

[25], [26], [27], [28], [29], [33]. For some results related to metric measure spaces, see [35], [49], [50].

Here we proceed as in the situation described before for the study of the discrepancy, and partition  $\mathcal{M}$  into  $N$  disjoint measurable sets  $\mathcal{X}_1, \dots, \mathcal{X}_N$  with positive measure, set  $\omega_j = |\mathcal{X}_j|$ , where  $|\cdot|$  denotes the measure, and consider random choices of points  $x_j \in \mathcal{X}_j$ .

To fix the notation we write  $\boldsymbol{\omega} = (\omega_1, \dots, \omega_N)$ ,  $\mathbf{x} = (x_1, \dots, x_N)$ ,  $\mathbf{X} = \mathcal{X}_1 \times \dots \times \mathcal{X}_N$ ,  $d\mathbf{x} = \frac{dx_1}{\omega_1} \times \dots \times \frac{dx_N}{\omega_N}$ , and consider the probability space  $(\mathbf{X}, d\mathbf{x})$ . We also write the error as

$$\mathcal{E}_{\mathbf{x}, \boldsymbol{\omega}}(f) = \sum_{j=1}^N \omega_j f(x_j) - \int_{\mathcal{M}} f(x) dx. \quad (6)$$

Notice the analogy with the discrepancy (1). There will be however an important difference: In order to measure the smoothness of our functions we will use suitable Besov spaces or potential spaces adapted to this more general context and obtain estimates of  $\mathcal{E}_{\mathbf{x}, \boldsymbol{\omega}}(f)$  for functions in such spaces. The previous combinatorial argument, however, works only when the integrability exponent  $p$  is an even integer, whereas in this case in order to obtain sharp results, we need estimates that work for generic values of  $p$ . The main idea is to replace such combinatorial argument with a generalization of the classical Khintchine inequality for sums of random variables, due to Marcinkiewicz and Zygmund [36, 37]. It says that for every  $1 \leq p < +\infty$  and for every sequence of independent random variables  $\{f_j\}_{j=1}^N$

$$\mathbb{E} \left( \left| \sum_{j=1}^N (f_j - \mathbb{E}(f_j)) \right|^p \right) \approx_p \mathbb{E} \left( \left( \sum_{j=1}^N |f_j - \mathbb{E}(f_j)|^2 \right)^{p/2} \right).$$

For the case of discrepancy, this immediately gives

$$\mathbb{E}(|D[\mathcal{P}, \mathcal{B}]|^p) \approx_p \frac{1}{N^p} \mathbb{E} \left( \left( \sum_{j \in \mathcal{J}} |\eta_j|^2 \right)^{p/2} \right) \leq \frac{1}{N^p} \mathbb{E}(|\mathcal{J}|^{p/2}) = O(N^{-3p/4}).$$

We will see that one such argument can be also used to deduce estimates on the error in numerical integration. In particular, we will study two types of problems. In the first case we will focus on the worst case numerical integration error, which determines how bad the error of a given fixed quadrature rule can be when applied to all integrands whose norm has an upper bound. The function space that we will consider for this type of problem is a space of potentials: We will say that  $f \in \mathbb{H}_p^\Phi(\mathcal{M})$  for  $1 \leq p \leq \infty$  and a suitable kernel  $\Phi(x, y)$  defined on  $\mathcal{M} \times \mathcal{M}$  (see §2 for the precise definition) if there is a  $g \in \mathbb{L}^p(\mathcal{M})$  such that

$$f(x) = \int_{\mathcal{M}} \Phi(x, y) g(y) dy,$$

and its norm is  $\|f\|_{\mathbb{H}_p^\Phi(\mathcal{M})} = \inf_g \|g\|_{\mathbb{L}^p(\mathcal{M})}$ , where infimum is taken over all  $g(x)$  which give the potential  $f(x)$ . Observe that when  $\mathcal{M}$  is the Euclidean space  $\mathbb{R}^d$  and  $\Phi(x, y) = |x - y|^{\alpha-d}$ ,  $0 < \alpha < d$ , is the Riesz kernel, then for  $1 < p < \infty$  the potential space  $\mathbb{H}_p^\Phi(\mathcal{M})$  coincides with the homogeneous fractional Sobolev

space  $\dot{\mathbb{H}}_p^\alpha(\mathbb{R}^d)$ . The (non-homogeneous) fractional Sobolev space  $\mathbb{H}_p^\alpha(\mathbb{R}^d)$  can be obtained similarly, via the Bessel kernel. Also, when  $\mathcal{M}$  is a compact Riemannian manifold, the Sobolev space  $\mathbb{H}_p^\alpha(\mathcal{M})$  can be defined as a potential space via the Bessel kernel, see Example 6.5 here or [8] for details on this. We will show here the following

**Theorem I** *Let  $\mathcal{M}$  be a metric measure space with the property that there exist  $d$  and  $c$  such that for every  $y \in \mathcal{M}$  and  $r > 0$ ,*

$$|\{x \in \mathcal{M} : |x - y| \leq r\}| \leq cr^d.$$

*Assume also that  $\mathcal{M}$  can be decomposed into a finite disjoint union of sets in the form  $\mathcal{M} = \mathcal{X}_1 \cup \dots \cup \mathcal{X}_N$ , with  $\omega_j = |\mathcal{X}_j| \approx N^{-1}$  and  $\delta_j = \text{diam}(\mathcal{X}_j) \approx N^{-1/d}$ . Assume that for some  $0 < \alpha < d$*

$$|\Phi(x, y)| \leq c|x - y|^{\alpha-d} \quad \text{for every } x \text{ and } y,$$

$$|\Phi(x, y) - \Phi(z, y)| \leq c|x - z||x - y|^{\alpha-d-1} \quad \text{if } |x - y| \geq 2|x - z|.$$

*Finally, assume that  $1 < p \leq +\infty$ ,  $1/p + 1/q = 1$ , and  $d/p < \alpha < d$ . Then*

$$\left\{ \int_{\mathbf{X}} \sup_{\|f\|_{\mathbb{H}_p^\alpha} \leq 1} |\mathcal{E}_{\mathbf{x}, \omega}(f)|^q d\mathbf{x} \right\}^{1/q} \leq \begin{cases} cN^{-\alpha/d} & \text{if } \alpha < d/2 + 1, \\ cN^{-1/2-1/d} (\log N)^{1/2} & \text{if } \alpha = d/2 + 1, \\ cN^{-1/2-1/d} & \text{if } \alpha > d/2 + 1. \end{cases}$$

The first observation is that the Bessel kernel on a compact Riemannian manifold satisfies the hypotheses required by this theorem. In fact, the particular case given by the case  $\mathcal{M}$  a compact Riemannian manifold,  $\Phi$  the Bessel kernel,  $\alpha < d/2 + 1$ , and  $p = 2$  had been proved in [8, Theorem 2.7] (see also [14, Theorem 24] for the case of the sphere).

We also show that under rather natural hypotheses on the space  $\mathcal{M}$  and on the kernel  $\Phi$ , the estimates from above hold from below as well.

**Theorem II** *Let  $\mathcal{M}$  be a metric measure space with the property that there exist  $H, K$ , and  $d$  such that for every  $y \in \mathcal{M}$  and  $0 < r < r_0$ ,*

$$Hr^d \leq |\{x \in \mathcal{M} : |x - y| \leq r\}| \leq Kr^d.$$

*Assume also that  $\mathcal{M}$  can be decomposed into a finite disjoint union of sets in the form  $\mathcal{M} = \mathcal{X}_1 \cup \dots \cup \mathcal{X}_N$ , with  $\omega_j = |\mathcal{X}_j| \approx N^{-1}$  and  $\delta_j = \text{diam}(\mathcal{X}_j) \approx N^{-1/d}$ . Suppose that there exists  $0 < \alpha < d$ , such that for any  $j = 1, \dots, N$  and any  $z \in \mathcal{X}_j$ , and for any  $y$  such that  $\text{dist}(y, \mathcal{X}_j) \geq 2\delta_j$ ,*

$$\int_{\mathcal{X}_j} |\Phi(x, y) - \Phi(z, y)| dx \geq cN^{-1-1/d} (\text{dist}(y, \mathcal{X}_j))^{\alpha-d-1}.$$

*Suppose also that for any  $y \in \mathcal{M}$ , the function  $x \mapsto \Phi(x, y)$  is continuous in  $x \neq y$ . Finally, assume that  $1 < p \leq +\infty$ ,  $1/p + 1/q = 1$ , and  $d/p < \alpha < d$ . Then*

$$\left\{ \int_{\mathbf{X}} \sup_{\|f\|_{\mathbb{H}_p^\alpha} \leq 1} |\mathcal{E}_{\mathbf{x}, \omega}(f)|^q d\mathbf{x} \right\}^{1/q} \geq \begin{cases} cN^{-\alpha/d} & \text{if } \alpha < d/2 + 1, \\ cN^{-1/2-1/d} (\log N)^{1/2} & \text{if } \alpha = d/2 + 1, \\ cN^{-1/2-1/d} & \text{if } \alpha > d/2 + 1. \end{cases}$$

Once again, it should be observed that the Bessel kernel on a compact Riemannian manifold satisfies the hypotheses in the above theorem, and that the particular case given by the case where  $\mathcal{M}$  is the sphere,  $\Phi$  the Bessel kernel, and  $p = 2$  is contained in [14, Theorems 24 and 25].

In order to understand the two above results it could be useful to recall the following result [8, Theorem 2.16]: Let  $\mathcal{M}$  be a compact Riemannian manifold. For every  $1 \leq p \leq \infty$  and  $\alpha > d/p$  there exists  $c > 0$  such that for every distribution of points  $\mathbf{x} = \{x_j\}_{j=1}^N$  and weights  $\omega = \{\omega_j\}_{j=1}^N$ , one has

$$\sup_{\|f\|_{\mathbb{H}_p^\alpha} \leq 1} |\mathcal{E}_{\mathbf{x}, \omega}(f)| \geq cN^{-\alpha/d}.$$

In other words, the worst case error for any quadrature rule cannot have a better decay than  $N^{-\alpha/d}$ . Thus, Theorem I says that a random choice of points  $x_j \in \mathcal{X}_j$ ,  $j = 1, \dots, N$ , gives the best possible decay for the worst case error in  $\mathbb{H}_p^\alpha(\mathcal{M})$  when  $d/2 < \alpha < d/2 + 1$ , while Theorem II says that when  $\alpha \geq d/2 + 1$  the stratification strategy does not lead, on average, to quadrature rules with the desired decay  $N^{-\alpha/d}$  of the worst case error in  $\mathbb{H}_p^\alpha(\mathcal{M})$ .

By the above mentioned result in [21] on the partitioning of  $\mathcal{M}$  into  $N$  regions of equal measure and small diameter, under the hypotheses on  $\mathcal{M}$  contained in Theorem II (see §5 below), the above Theorems I and II apply with equal weights, that is  $\omega_j = |\mathcal{M}|/N$  for all  $j = 1, \dots, N$ . In [12] and [14] a sequence of point configurations on the sphere that gives the best possible decay  $N^{-\alpha/d}$  for the worst case error in the equal weight case has been called a Quasi Monte Carlo (QMC) design sequence for  $\mathbb{H}_p^\alpha(\mathcal{M})$ .

We would like to emphasize that all the cited results in [8] and [14] are based on Hilbert space techniques ( $p = 2$ ), while we were able to obtain  $\mathbb{L}^p$  integrability results thanks to the Marcinkiewicz–Zygmund inequality. Moreover, we work in the more general setting of metric measure spaces, and some of the results are new even in the particular case of compact Riemannian manifolds.

Concerning the non-Hilbert space setting, we mention here that when  $\mathcal{M}$  is the  $d$ -dimensional sphere,  $1 \leq p \leq +\infty$ , and  $\alpha > d/p$ , the existence of QMC design sequences for  $\mathbb{H}_p^\alpha(\mathcal{M})$  has been proved with different techniques in [12].

So far, we have considered the worst case error in numerical integration, that is the error for a whole class of functions. The second type of estimate that will be discussed here concerns the numerical integration error for a given fixed function. In particular, we will consider functions in the homogeneous Hajlasz–Besov space  $\dot{\mathbb{B}}_{p, \infty}^\alpha(\mathcal{M})$ , as defined in [32]. For details on these spaces, see §3 below. Here we should mention that when  $0 < \alpha < 1$ , the spaces  $\mathbb{H}_p^\alpha(\mathcal{M})$  as in Theorem I are embedded into  $\dot{\mathbb{B}}_{p, \infty}^\alpha(\mathcal{M})$ , and that when  $\mathcal{M}$  is the Euclidean space  $\mathbb{R}^d$  and  $0 < \alpha < 1$  then the spaces  $\dot{\mathbb{B}}_{p, \infty}^\alpha(\mathcal{M})$  coincide with the classical homogeneous Besov spaces defined via Littlewood–Paley decomposition. The main result in this context is the following

**Theorem III** *Assume that a metric measure space  $\mathcal{M}$  can be decomposed into a finite disjoint union of sets in the form  $\mathcal{M} = \mathcal{X}_1 \cup \dots \cup \mathcal{X}_N$ , with measure  $0 < |\mathcal{X}_j| = \omega_j \approx N^{-1}$  and  $0 < \text{diam}(\mathcal{X}_j) \approx N^{-1/d}$ . Then for every  $1 \leq p \leq 2$  there is a constant  $c$  such that*

$$\left\{ \int_{\mathbf{x}} |\mathcal{E}_{\mathbf{x}, \omega}(f)|^p d\mathbf{x} \right\}^{1/p} \leq cN^{1/p-1-\alpha/d} \|f\|_{\dot{\mathbb{B}}_{p, \infty}^\alpha(\mathcal{M})}$$

and for every  $2 \leq p < +\infty$  there is a constant  $c$  such that

$$\left\{ \int_{\mathbf{X}} |\mathcal{E}_{\mathbf{x},\omega}(f)|^p d\mathbf{x} \right\}^{1/p} \leq c N^{-1/2-\alpha/d} \|f\|_{\mathbb{B}_{p,\infty}^\alpha(\mathcal{M})}.$$

Of course, here a random choice of points  $x_j \in \mathcal{X}_j$ ,  $j = 1, \dots, N$  gives better estimates than those obtained in Theorem I (and better than  $N^{-\alpha/d}$ ). This is natural, since in this case we are looking for point distributions which give a small error for a given integrand  $f$ , whereas in the situation described by Theorem I we were looking for point distributions which give a small error for all integrands in our space at the same time.

Theorem III and its sharpness will be discussed in §7 below.

We believe that our effort in the search for the minimal properties that guarantee the validity of certain results in discrepancy theory and numerical integration can be of some help towards a deeper understanding of these results, even in the classical cases. In fact, to the best of our knowledge, the above Theorems I, II and III are new even when applied to a general compact Riemannian manifold. There has recently been some interest in this type of problems in spaces as general as those considered here, or even more. See for example [35], [41], [45], [46], [47] for discrepancy and numerical integration on metric spaces and [49], [50] for analysis on fractals.

The plan of the paper is the following. In §2 and §3 we introduce the appropriate Sobolev-type spaces, and we recall a few results on how these spaces relate to each other. These matters are not completely new, but can be of some help for the unfamiliar reader. In §4 we introduce in some detail the Marcinkiewicz–Zygmund inequality. In §5 we recall the above mentioned result in [21] concerning the partitioning of a metric measure space into regions of equal measure and small diameter. In §6 and §7 we give all the details on our results on numerical integration, with examples. Finally, §8 contains our results on the  $L^p$  (and  $L^\infty$ ) discrepancy that generalize (3) and (5).

## 2 Sobolev spaces and potentials on measure spaces

Our estimates on the worst case error described in Theorems I and II require a definition of Sobolev spaces via potentials. For a classical approach to potentials, see, for example, [48].

**Definition 2.1** Let  $\mathcal{M}$  be a measure space, let  $1 \leq p, q \leq +\infty$  with  $1/p + 1/q = 1$ , and let  $\Phi(x, y)$  be a measurable kernel on  $\mathcal{M} \times \mathcal{M}$ . Assume that for every  $x$ ,

$$\begin{aligned} \int_{\mathcal{M}} |\Phi(x, y)|^q dy &< +\infty \text{ if } q < +\infty, \\ \text{ess sup}_{y \in \mathcal{M}} \{|\Phi(x, y)|\} &< +\infty \text{ if } q = +\infty. \end{aligned}$$

Then every function  $g(x)$  in  $\mathbb{L}^p(\mathcal{M})$  has a pointwise well defined potential

$$f(x) = \int_{\mathcal{M}} \Phi(x, y) g(y) dy.$$



The space  $\mathbb{H}_p^\Phi(\mathcal{M})$  is the space of all potentials of functions in  $L^p(\mathcal{M})$ , with norm

$$\|f\|_{\mathbb{H}_p^\Phi(\mathcal{M})} = \inf_g \left\{ \int_{\mathcal{M}} |g(x)|^p dx \right\}^{1/p}.$$

The infimum is taken over all  $g(x)$  which give the potential  $f(x)$ .

Observe that this definition does not even require a metric. Potentials can also be defined under weaker assumptions on the kernel, but the above assumptions guarantee that these potentials are defined pointwise everywhere, and this will be necessary in the sequel. In particular, when  $\mathcal{M}$  is the Euclidean space  $\mathbb{R}^d$  and  $\Phi(x, y) = |x - y|^{\alpha-d}$  with  $0 < \alpha < d$  is the Riesz kernel, then  $\mathbb{H}_p^\Phi(\mathcal{M})$  is the homogeneous fractional Sobolev space  $\dot{\mathbb{H}}_p^\alpha(\mathbb{R}^d)$ . However, the cases  $p = 1$  and  $p = +\infty$  require some extra care. For interesting examples of generalized potential spaces, see, for example, [31].

### 3 Besov and Triebel–Lizorkin spaces on metric measure spaces

Our estimates in Theorem III require a definition of Sobolev spaces, more appropriately Besov or Triebel–Lizorkin spaces, via upper gradients. Let  $\mathcal{M}$  be a metric measure space, that is, a metric space equipped with a positive Borel measure. With a small abuse of notation we denote by  $|\mathcal{X}|$  the measure of a measurable set  $\mathcal{X}$  and by  $|x - y|$  the distance between two points  $x$  and  $y$ . We will often denote with  $B(x, r)$  the open balls  $\{y \in \mathcal{M} : |x - y| < r\}$  with center  $x$  and radius  $r$ . Simple examples are Riemannian manifolds, or not necessarily smooth surfaces in a Euclidean space with the inherited measure and distance. In [24] Hajlasz has given a purely metric definition of a Sobolev space: A measurable function  $f(x)$  is in the Sobolev space  $\mathbb{W}_p^1(\mathcal{M})$ ,  $1 \leq p \leq +\infty$ , if there exists a non-negative function  $g(x)$  in  $L^p(\mathcal{M})$  such that for almost every  $x, y \in \mathcal{M}$ ,

$$|f(x) - f(y)| \leq |x - y| (g(x) + g(y)).$$

For example, it is proved in [24] that in Euclidean spaces one can choose as an upper gradient  $g(x)$  a suitable multiple of the Hardy–Littlewood maximal function of the gradient  $\nabla f(x)$ .

The following is a natural generalization of upper gradient and associated Sobolev space.

**Definition 3.1** Let  $\mathcal{M}$  be a metric measure space and  $\varphi(t)$  a non-negative increasing function in  $t \geq 0$ . A measurable non-negative function  $g(x)$  is a  $\varphi$ -gradient of a measurable function  $f(x)$  if there exists a set  $\mathcal{N}$  with measure zero such that for all  $x$  and  $y$  in  $\mathcal{M} \setminus \mathcal{N}$ ,

$$|f(x) - f(y)| \leq \varphi(|x - y|) (g(x) + g(y)).$$

**Definition 3.2** A measurable function  $f(x)$  is in the Hajlasz–Sobolev space  $\dot{\mathbb{M}}_p^\varphi(\mathcal{M})$ ,  $0 < p \leq +\infty$ , if it has a  $\varphi$ -gradient in  $L^p(\mathcal{M})$ . We set

$$\|f\|_{\dot{\mathbb{M}}_p^\varphi(\mathcal{M})} = \inf \|g\|_{L^p(\mathcal{M})}.$$

The infimum is taken over all  $\varphi$ -gradients  $g(x)$  of  $f(x)$ .

In [24] Hajlasz proved that when  $\varphi(t) = t$ ,  $1 < p \leq +\infty$  and  $\mathcal{M} = \mathbb{R}^d$ , then this space coincides with the classical homogeneous Sobolev space  $\dot{\mathbb{H}}_p^1(\mathbb{R}^d)$ . The above definition has been extended by Koskela, Yang and Zhou [32] who have defined Besov and Triebel–Lizorkin spaces on a general metric measure space. In particular they proved that, when  $\varphi(t) = t^\alpha$  with  $0 < \alpha < 1$  and  $\mathcal{M} = \mathbb{R}^d$ , then the space  $\dot{\mathbb{M}}_p^\alpha(\mathcal{M})$  is larger than the classical fractional Sobolev space, and it coincides with the homogeneous Triebel–Lizorkin space  $\dot{\mathbb{F}}_{p,\infty}^\alpha(\mathbb{R}^d)$ .

What follows is in the spirit of the definitions of Besov and Triebel–Lizorkin spaces in [32]. In order to define these spaces one needs to introduce families of gradients localized at different scales.

**Definition 3.3** Let  $\mathcal{M}$  be a metric measure space and  $\varphi(t)$  a non-negative increasing function in  $t \geq 0$ . Let  $n_0 = \log_2(\text{diam}(\mathcal{M}))$ , possibly infinity. A sequence of non-negative measurable functions  $\{g_n(x)\}_{n=-n_0}^{+\infty}$  is a  $\varphi$ -gradient for a measurable function  $f(x)$  if there exists a set  $\mathcal{N}$  with measure zero such that for all  $x$  and  $y$  in  $\mathcal{M} \setminus \mathcal{N}$  with  $|x - y| \leq 2^{-n}$ ,

$$|f(x) - f(y)| \leq \varphi(2^{-n})(g_n(x) + g_n(y)).$$

**Definition 3.4** A measurable function  $f(x)$  is in the Hajlasz–Triebel–Lizorkin space  $\dot{\mathbb{F}}_{p,q}^\varphi(\mathcal{M})$ ,  $0 < p < +\infty$  and  $0 < q \leq +\infty$ , if  $f(x)$  has a  $\varphi$ -gradient  $\{g_n(x)\}$  with

$$\left\| \left\{ \sum_{n=-n_0}^{+\infty} |g_n(x)|^q \right\}^{1/q} \right\|_{\mathbb{L}^p(\mathcal{M})} < +\infty \text{ if } 0 < q < +\infty,$$

$$\left\| \sup_{n \geq -n_0} |g_n(x)| \right\|_{\mathbb{L}^p(\mathcal{M})} < +\infty \text{ if } q = +\infty.$$

The infimum of the above expression taken over all  $\varphi$ -gradients defines the semi-norm  $\|f\|_{\dot{\mathbb{F}}_{p,q}^\varphi(\mathcal{M})}$ .

**Definition 3.5** A measurable function  $f(x)$  is in the Hajlasz–Besov space  $\dot{\mathbb{B}}_{p,q}^\varphi(\mathcal{M})$ ,  $0 < p \leq +\infty$  and  $0 < q \leq +\infty$ , if  $f(x)$  has a  $\varphi$ -gradient  $\{g_n(x)\}$  with

$$\left\{ \sum_{n=-n_0}^{+\infty} \|g_n(x)\|_{\mathbb{L}^p(\mathcal{M})}^q \right\}^{1/q} < +\infty \text{ if } 0 < q < +\infty,$$

$$\sup_{n \geq -n_0} \|g_n(x)\|_{\mathbb{L}^p(\mathcal{M})} < +\infty \text{ if } q = +\infty.$$

The infimum of the above expression taken over all  $\varphi$ -gradients defines the semi-norm  $\|f\|_{\dot{\mathbb{B}}_{p,q}^\varphi(\mathcal{M})}$ .

Observe that the above spaces are homogeneous, and constant functions have semi-norms equal to zero. Also observe that when  $q = +\infty$  and  $\varphi(t)$  is doubling, that is  $\varphi(2t) \leq c\varphi(t)$ , then the space  $\dot{\mathbb{F}}_{p,\infty}^\varphi(\mathcal{M})$  coincides with the previously defined Hajlasz–Sobolev space  $\dot{\mathbb{M}}_p^\varphi(\mathcal{M})$ . It suffices to define  $g(x) = \sup \{g_n(x)\}$ . In particular, the straightforward generalization of a Hajlasz–Sobolev space is a Hajlasz–Triebel–Lizorkin space.

When  $\varphi(t) = t^\alpha$  the above definition is nothing but the definition of Hajlasz–Besov and Hajlasz–Triebel–Lizorkin given in [32]. To be precise, the definition of  $\varphi$ -gradient in [32] requires

$$|f(x) - f(y)| \leq 2^{-\alpha n}(g_n(x) + g_n(y))$$

only for  $x$  and  $y$  with  $2^{-n-1} \leq |x - y| \leq 2^{-n}$ . On the other hand, if one defines

$$G_n(x) = \sum_{k=0}^{+\infty} 2^{-\alpha k} g_{n+k}(x),$$

then

$$|f(x) - f(y)| \leq 2^{-\alpha n} (G_n(x) + G_n(y))$$

for  $x$  and  $y$  with  $|x - y| \leq 2^{-n}$ , and the seminorms defined via  $\{g_n(x)\}$  and  $\{G_n(x)\}$  are equivalent. In the same paper, it is proved that when  $\mathcal{M}$  is the Euclidean space  $\mathbb{R}^d$  and  $\varphi(t) = t^\alpha$  with  $0 < \alpha < 1$ , then the spaces  $\mathbb{B}_{p,q}^\varphi(\mathcal{M})$  and  $\mathring{\mathbb{F}}_{p,q}^\varphi(\mathcal{M})$  coincide with the classical Besov and Triebel–Lizorkin spaces defined via a Littlewood–Paley decomposition. See, for example, [3] for the relevant definitions.

The lemma below gives an example of a function in the spaces  $\mathbb{B}_{p,q}^\varphi(\mathcal{M})$  and  $\mathring{\mathbb{F}}_{p,q}^\varphi(\mathcal{M})$ .

**Definition 3.6** For every subset  $\mathcal{B}$  in  $\mathcal{M}$ , define

$$\psi_{\mathcal{B}}(t) = |\{x \in \mathcal{B} : \text{dist}\{x, \mathcal{M} \setminus \mathcal{B}\} \leq t\}| + |\{x \in \mathcal{M} \setminus \mathcal{B} : \text{dist}\{x, \mathcal{B}\} \leq t\}|.$$

For example, if  $\mathcal{M}$  is a  $d$ -dimensional Riemannian manifold and  $\mathcal{B}$  is a bounded open set with regular boundary, then  $\psi_{\mathcal{B}}(t) \approx t$ , while if  $\psi_{\mathcal{B}}(t) \approx t^\beta$ , then the boundary has Minkowski fractal dimension  $d - \beta$ .

**Proposition 3.7** Let  $\mathcal{B}$  be an arbitrary subset of  $\mathcal{M}$ . Then

$$\begin{aligned} \|\chi_{\mathcal{B}}\|_{\mathring{\mathbb{F}}_{p,\infty}^\varphi(\mathcal{M})} &\leq \left\{ \sum_{n=-n_0}^{+\infty} \varphi(2^{-n})^{-p} \psi_{\mathcal{B}}(2^{-n}) \right\}^{1/p}, \\ \|\chi_{\mathcal{B}}\|_{\mathbb{B}_{p,\infty}^\varphi(\mathcal{M})} &\leq \sup_{n \geq -n_0} \left\{ \varphi(2^{-n})^{-1} \psi_{\mathcal{B}}(2^{-n})^{1/p} \right\}. \end{aligned}$$

In particular, if  $\varphi(t) = t^\alpha$  and  $\psi_{\mathcal{B}}(t) \leq ct^\beta$ , then  $\chi_{\mathcal{B}} \in \mathring{\mathbb{F}}_{p,\infty}^\varphi(\mathcal{M})$  for  $p\alpha < \beta$  and  $\chi_{\mathcal{B}} \in \mathbb{B}_{p,\infty}^\varphi(\mathcal{M})$  for  $p\alpha \leq \beta$ .

*Proof* It suffices to observe that a  $\varphi$ -gradient for the characteristic function  $\chi_{\mathcal{B}}(x)$  is given by

$$g_n(x) = \begin{cases} \varphi(2^{-n})^{-1} & \text{if } x \in \mathcal{B} \text{ and } \text{dist}\{x, \mathcal{M} \setminus \mathcal{B}\} \leq 2^{-n}, \\ 0 & \text{otherwise.} \end{cases}$$

Of course, there are other possible choices for the  $\varphi$ -gradient of  $\chi_{\mathcal{B}}(x)$ , for example,

$$g_n(x) = \begin{cases} \varphi(2^{-n})^{-1} & \text{if } x \in \mathcal{M} \setminus \mathcal{B} \text{ and } \text{dist}\{x, \mathcal{B}\} \leq 2^{-n}, \\ 0 & \text{otherwise.} \end{cases}$$

□

The following is an immediate consequence of the definitions.

**Proposition 3.8** (i) If  $q_1 \leq q_2$  and  $\varphi_1(t) \leq \varphi_2(t)$ , then

$$\dot{\mathbb{B}}_{p,q_1}^{\varphi_1}(\mathcal{M}) \subseteq \dot{\mathbb{B}}_{p,q_2}^{\varphi_2}(\mathcal{M}) \quad \text{and} \quad \dot{\mathbb{F}}_{p,q_1}^{\varphi_1}(\mathcal{M}) \subseteq \dot{\mathbb{F}}_{p,q_2}^{\varphi_2}(\mathcal{M}).$$

(ii) For every  $\varphi(t)$  and  $0 < p \leq +\infty$ ,

$$\dot{\mathbb{B}}_{p,p}^{\varphi}(\mathcal{M}) = \dot{\mathbb{F}}_{p,p}^{\varphi}(\mathcal{M}) \quad \text{and} \quad \dot{\mathbb{F}}_{p,\infty}^{\varphi}(\mathcal{M}) \subseteq \dot{\mathbb{B}}_{p,\infty}^{\varphi}(\mathcal{M}).$$

In particular, for fixed  $p$  and  $\varphi(t)$ , the largest space in the scale of Hajlasz–Besov and Hajlasz–Triebel–Lizorkin spaces is  $\dot{\mathbb{B}}_{p,\infty}^{\varphi}(\mathcal{M})$ .

In the Euclidean spaces it is well known that the homogeneous potential spaces  $\dot{\mathbb{H}}_p^{\alpha}(\mathbb{R}^d)$  defined via the Riesz potentials coincide with the homogeneous Triebel–Lizorkin spaces  $\dot{\mathbb{F}}_{p,2}^{\alpha}(\mathbb{R}^d)$  defined via the Littlewood–Paley decomposition [3, 20]. We do not know under which assumptions on  $\Phi(x, y)$  and  $\varphi(t)$  and  $\mathcal{M}$  the equality  $\dot{\mathbb{H}}_p^{\Phi}(\mathcal{M}) = \dot{\mathbb{F}}_{p,2}^{\varphi}(\mathcal{M})$  holds. Anyhow, the following proposition guarantees a weaker embedding.

**Proposition 3.9** Assume that  $\psi(t)$  is an increasing function on  $0 \leq t < +\infty$  with  $\psi(0) = 0$ , and define for  $\varepsilon > 0$

$$\varphi(t) = \sum_{k=0}^{+\infty} \psi\left(2^{2-k}t\right) + \sum_{k=0}^{+\infty} 2^{-(k+1)\varepsilon} \psi\left(2^{k+2}t\right).$$

Also assume that  $\Phi(x, y)$  is a kernel on  $\mathcal{M} \times \mathcal{M}$  with the property that for some  $C > 0$

$$|\Phi(x, y)| \leq C \frac{\psi(|x-y|)}{|B(x, 6|x-y|)|} \quad \text{for every } x, y \in \mathcal{M},$$

$$|\Phi(x, y) - \Phi(z, y)| \leq C \left(\frac{|x-z|}{|x-y|}\right)^{\varepsilon} \frac{\psi(|x-y|)}{|B(x, 6|x-y|)|} \quad \text{when } |x-y| \geq 2|x-z|.$$

Define the potential

$$f(x) = \int_{\mathcal{M}} \Phi(x, y) g(y) dy.$$

Finally, define the Hardy–Littlewood maximal operator

$$Mg(x) = \sup_{r>0} \left\{ |B(x, 3r)|^{-1} \int_{\{|x-y| \leq r\}} |g(y)| dy \right\}.$$

Then,

$$|f(x) - f(z)| \leq 2C\varphi(|x-z|)(Mg(x) + Mg(z)).$$

*Proof* By the hypotheses on the kernel,

$$\begin{aligned} |f(x) - f(z)| &\leq \int_{\mathcal{M}} |\Phi(x, y) - \Phi(z, y)| |g(y)| dy \\ &\leq C \int_{\{|x-y| \leq 4|x-z|\}} \psi(|x-y|) |B(x, 6|x-y|)|^{-1} |g(y)| dy \\ &\quad + C \int_{\{|z-y| \leq 4|x-z|\}} \psi(|z-y|) |B(z, 6|z-y|)|^{-1} |g(y)| dy \\ &\quad + C \int_{\{|x-y| \geq 2|x-z|\}} |x-z|^{\varepsilon} |x-y|^{-\varepsilon} \psi(|x-y|) |B(x, 6|x-y|)|^{-1} |g(y)| dy. \end{aligned}$$

The dyadic decomposition

$$\{|x - y| \leq 4|x - z|\} = \bigcup_{k=0}^{+\infty} \{2^{1-k}|x - z| < |x - y| \leq 2^{2-k}|x - z|\}$$

gives

$$\begin{aligned} & \int_{\{|x-y| \leq 4|x-z|\}} \psi(|x-y|) |B(x, 6|x-y|)^{-1} |g(y)| dy \\ & \leq \sum_{k=0}^{+\infty} \psi(2^{2-k}|x-z|) |B(x, 6 \cdot 2^{1-k}|x-z|)^{-1} \int_{\{|x-y| \leq 2^{2-k}|x-z|\}} |g(y)| dy \\ & \leq \sum_{k=0}^{+\infty} \psi(2^{2-k}|x-z|) \sup_{r>0} \left\{ |B(x, 3r)|^{-1} \int_{\{|x-y| \leq r\}} |g(y)| dy \right\} \\ & \leq \varphi(|x-z|) Mg(x). \end{aligned}$$

Similarly,

$$\int_{\{|z-y| \leq 4|x-z|\}} \psi(|z-y|) |B(z, 6|z-y|)^{-1} |g(y)| dy \leq \varphi(|x-z|) Mg(z).$$

Finally, the dyadic decomposition

$$\{|x - y| \geq 2|x - z|\} = \bigcup_{k=0}^{+\infty} \{2^{k+1}|x - z| \leq |x - y| < 2^{k+2}|x - z|\}$$

gives

$$\begin{aligned} & \int_{\{|x-y| \geq 2|x-z|\}} |x-z|^\varepsilon |x-y|^{-\varepsilon} \psi(|x-y|) |B(x, 6|x-y|)^{-1} |g(y)| dy \\ & \leq \sum_{k=0}^{+\infty} 2^{-(k+1)\varepsilon} \psi(2^{k+2}|x-z|) |B(x, 6 \cdot 2^{k+1}|x-z|)^{-1} \int_{\{|x-y| \leq 2^{k+2}|x-z|\}} |g(y)| dy \\ & \leq \varphi(|x-z|) Mg(x). \end{aligned}$$

□

**Corollary 3.10** *With the notation of the above proposition, if  $1 < p \leq +\infty$ , then the potential space  $\mathbb{H}_p^\Phi(\mathcal{M})$  can be continuously embedded into  $\dot{\mathbb{F}}_{p,\infty}^\varphi(\mathcal{M})$ .*

*Proof* It suffices to recall that, due to the extra 3 in the definition of  $Mg(x)$ , this maximal operator  $Mg(x)$  is bounded on  $\mathbb{L}^p(\mathcal{M})$  for all  $1 < p \leq +\infty$ , even when the measure on the metric space is non doubling. See [39]. □

*Example 3.11* If  $\varphi(t) = t^\alpha$  with  $0 < \alpha < \varepsilon$ , then

$$\begin{aligned} \psi(t) &= \sum_{k=0}^{+\infty} \varphi(2^{2-k}t) + \sum_{k=0}^{+\infty} 2^{(-k-1)\varepsilon} \varphi(2^{k+2}t) \\ &= \left( 2^{2\alpha} \sum_{k=0}^{+\infty} 2^{-k\alpha} + 2^{2\alpha-\varepsilon} \sum_{k=0}^{+\infty} 2^{-k(\varepsilon-\alpha)} \right) t^\alpha = Ct^\alpha. \end{aligned}$$

Thus, if

$$|\Phi(x, y)| \leq C \frac{|x - y|^\alpha}{|B(x, 6|x - y|)|} \quad \text{for every } x, y \in \mathcal{M},$$

$$|\Phi(x, y) - \Phi(z, y)| \leq C \left( \frac{|x - z|}{|x - y|} \right)^\varepsilon \frac{|x - y|^\alpha}{|B(x, 6|x - y|)|} \quad \text{when } |x - y| \geq 2|x - z|.$$

and if  $1 < p \leq +\infty$  then the potential space  $\mathbb{H}_p^\Phi(\mathcal{M})$  can be continuously embedded into  $\mathbb{F}_{p, \infty}^\varphi(\mathcal{M})$  with  $\varphi(t) = t^\alpha$ .

#### 4 The Marcinkiewicz–Zygmund inequality

The main ingredient in what follows is the Marcinkiewicz–Zygmund inequality for sums of independent random variables.

As is well known, the variance of the sum of independent random variables is the sum of the variances. For every sequence of independent random variables  $f_j$ ,

$$\mathbb{E} \left( \left| \sum_j (f_j - \mathbb{E}(f_j)) \right|^2 \right) = \sum_j \mathbb{E} (|f_j - \mathbb{E}(f_j)|^2).$$

In fact, there is a similar result with the second moment replaced by other moments and with the equality replaced by two inequalities.

**Theorem 4.1 (Marcinkiewicz–Zygmund)** *For every  $1 \leq p < +\infty$ , there exist positive constants  $A(p)$  and  $B(p)$  such that for every sequence  $\{f_j\}$  of independent random variables,*

$$\begin{aligned} A(p) \left\{ \mathbb{E} \left( \sum_{j=1}^N |f_j - \mathbb{E}(f_j)|^2 \right)^{p/2} \right\}^{1/p} \\ \leq \left\{ \mathbb{E} \left( \left| \sum_{j=1}^N (f_j - \mathbb{E}(f_j)) \right|^p \right) \right\}^{1/p} \\ \leq B(p) \left\{ \mathbb{E} \left( \sum_{j=1}^N |f_j - \mathbb{E}(f_j)|^2 \right)^{p/2} \right\}^{1/p}. \end{aligned}$$

The Marcinkiewicz–Zygmund inequality is a generalization of the classical inequality of Khintchine for sums of random variables with Rademacher distribution that take values  $\pm 1$  with probability  $1/2$ . For a proof, see [36] and [37], or [18].

In what follows, special attention will be paid to the constants, and  $A(p)$  and  $B(p)$  will denote the best constants in the Marcinkiewicz–Zygmund inequality. If  $\bar{A}(p)$  and  $\bar{B}(p)$  are the corresponding best constants for the Khintchine inequality, then it can be proved that

$$\frac{1}{2}\bar{A}(p) \leq A(p) \leq \bar{A}(p) \quad \text{and} \quad \bar{B}(p) \leq B(p) \leq 2\bar{B}(p);$$

see, for example, [18]. In particular there is a positive constant  $c$  such that  $c \leq \bar{A}(p) \leq 1$ , while  $\bar{B}(p) = 1$  for  $1 \leq p \leq 2$  and  $\bar{B}(p) = \sqrt{2} (\Gamma((p+1)/2) / \sqrt{\pi})^{1/p}$  for  $2 \leq p < +\infty$ ; see [15] and [23].

We remark that the Marcinkiewicz–Zygmund inequality can also be extended to infinite sums of independent random variables.

From now on we will assume that  $\mathcal{M}$  is a measure space of finite measure which can be expressed as a finite union  $\mathcal{X}_1 \cup \dots \cup \mathcal{X}_N$  of disjoint sets  $\mathcal{X}_1, \dots, \mathcal{X}_N$  with measures  $0 < |\mathcal{X}_j| = \omega_j < +\infty$ .

As indicated earlier, we write  $\boldsymbol{\omega} = (\omega_1, \dots, \omega_N)$ ,  $\mathbf{x} = (x_1, \dots, x_N)$ ,  $\mathbf{X} = \mathcal{X}_1 \times \dots \times \mathcal{X}_N$  and

$$d\mathbf{x} = \frac{dx_1}{\omega_1} \times \dots \times \frac{dx_N}{\omega_N}.$$

The error incurred in a quadrature rule with sampling points  $\mathbf{x} = (x_1, \dots, x_N)$  and weights  $\boldsymbol{\omega} = (\omega_1, \dots, \omega_N)$  is the functional

$$\mathcal{E}_{\mathbf{x}, \boldsymbol{\omega}}(f) = \sum_{j=1}^N \omega_j f(x_j) - \int_{\mathcal{M}} f(x) dx.$$

The above Marcinkiewicz–Zygmund inequality has an immediate corollary that allows us to control the norm of this error.

**Corollary 4.2** *Let  $1 \leq p < +\infty$ . For every measurable function  $f(x)$  on  $\mathcal{M}$ ,*

$$\begin{aligned} A(p) \left\{ \int_{\mathbf{X}} \left( \sum_{j=1}^N \left| \omega_j f(x_j) - \int_{\mathcal{X}_j} f(y_j) dy_j \right|^2 \right)^{p/2} d\mathbf{x} \right\}^{1/p} \\ \leq \left\{ \int_{\mathbf{X}} |\mathcal{E}_{\mathbf{x}, \boldsymbol{\omega}}(f)|^p d\mathbf{x} \right\}^{1/p} \\ \leq B(p) \left\{ \int_{\mathbf{X}} \left( \sum_{j=1}^N \left| \omega_j f(x_j) - \int_{\mathcal{X}_j} f(y_j) dy_j \right|^2 \right)^{p/2} d\mathbf{x} \right\}^{1/p}. \end{aligned}$$

*Proof* It suffices to apply the Marcinkiewicz–Zygmund inequality to the independent random variables

$$f_j(\mathbf{x}) = f_j(x_1, x_2, \dots, x_N) = \omega_j f(x_j),$$

and observe that

$$\mathbb{E}(f_j) = \int_{\mathcal{X}_j} f(x_j) dx_j$$

and

$$\sum_{j=1}^N (f_j - \mathbb{E}(f_j)) = \sum_{j=1}^N \left( \omega_j f(x_j) - \int_{\mathcal{X}_j} f(x_j) dx_j \right) = \sum_{j=1}^N \omega_j f(x_j) - \int_{\mathcal{M}} f(x) dx.$$

□

## 5 Diameter bounded equal measure partition of metric measure spaces

In some of the results that follow we shall assume that a metric measure space  $\mathcal{M}$  can be decomposed into a finite disjoint union of sets in the form  $\mathcal{M} = \mathcal{X}_1 \cup \dots \cup \mathcal{X}_N$ , with  $\omega_j = |\mathcal{X}_j| \approx N^{-1}$  and  $\delta_j = \text{diam}(\mathcal{X}_j) \approx N^{-1/d}$ . In fact, under appropriate assumptions, Gigante and Leopardi in [21] proved the following more precise result.

**Theorem 5.1** *Let  $\mathcal{M}$  be a connected metric measure space with finite measure with the property that there exist positive constants  $d$ ,  $H$  and  $K$  such that for every  $y \in \mathcal{M}$  and  $0 < r < \text{diam}(\mathcal{M})$ ,*

$$Hr^d \leq |\{x \in \mathcal{M} : |x - y| < r\}| \leq Kr^d.$$

*Then, there exist two constants  $c_1$  and  $c_2$ , such that for every sufficiently large  $N$  there exists a partition  $\mathcal{M} = \mathcal{X}_1 \cup \dots \cup \mathcal{X}_N$  and points  $y_j \in \mathcal{X}_j$  with  $|\mathcal{X}_j| = |\mathcal{M}|/N$  and*

$$\left\{x \in \mathcal{M} : |x - y_j| < c_1 N^{-1/d}\right\} \subset \mathcal{X}_j \subset \left\{x \in \mathcal{M} : |x - y_j| < c_2 N^{-1/d}\right\}.$$

For example, the theorem applies to all compact Riemannian manifolds. An algorithmic construction in the particular case of the 2-dimensional sphere that pays attention to the size of the constant  $c_2$  is contained in [40] (see also [34] for the extension to higher dimensions).

## 6 Numerical integration in potential spaces

In this section we shall study the functional  $\mathcal{E}_{\mathbf{x},\omega}$  on the potential space  $\mathbb{H}_p^\Phi(\mathcal{M})$ . Let

$$\|\mathcal{E}_{\mathbf{x},\omega}\|_{\Phi,p} = \sup_{f \in \mathbb{H}_p^\Phi(\mathcal{M})} \left\{ \frac{|\mathcal{E}_{\mathbf{x},\omega}(f)|}{\|f\|_{\mathbb{H}_p^\Phi(\mathcal{M})}} \right\}$$

be the norm of this functional, also termed *worst case error*, and the following lemma gives an explicit formula for it.

**Lemma 6.1** *Assume that a measure space  $\mathcal{M}$  can be decomposed into a finite disjoint union of sets in the form  $\mathcal{M} = \mathcal{X}_1 \cup \dots \cup \mathcal{X}_N$ , with measure  $|\mathcal{X}_j| = \omega_j > 0$ . Assume also that  $1 \leq p \leq +\infty$ ,  $1 \leq q \leq +\infty$ ,  $1/p + 1/q = 1$ , and that for every  $x$ ,*

$$\left\{ \int_{\mathcal{M}} |\Phi(x, y)|^q dy \right\}^{1/q} < +\infty.$$

*Finally, assume that*

$$\left\{ \int_{\mathcal{M}} \left( \int_{\mathcal{M}} |\Phi(x, y)| dx \right)^q dy \right\}^{1/q} < +\infty.$$

*Then the functional  $\mathcal{E}_{\mathbf{x},\omega}$  is well defined and continuous on  $H_p^\Phi(\mathcal{M})$ , and its norm is*

$$\|\mathcal{E}_{\mathbf{x},\omega}\|_{\Phi,p} = \left\{ \int_{\mathcal{M}} \left| \sum_{j=1}^N \int_{\mathcal{X}_j} (\Phi(x_j, y) - \Phi(x, y)) dx \right|^q dy \right\}^{1/q}.$$



*Proof* For simplicity assume that  $q < +\infty$ , the case  $q = +\infty$  being similar. Let

$$f(x) = \int_{\mathcal{M}} \Phi(x, y) g(y) dy$$

be the potential of a function  $g(x)$  in  $\mathbb{L}^p(\mathcal{M})$ . Since

$$\int_{\mathcal{M}} |\Phi(x, y)|^q dy < +\infty,$$

$f(x)$  is pointwise well defined. Since

$$\int_{\mathcal{M}} \left( \int_{\mathcal{M}} |\Phi(x, y)| dx \right)^q dy < +\infty,$$

it follows from Fubini's theorem that  $f(x)$  is integrable, and

$$\int_{\mathcal{M}} f(x) dx = \int_{\mathcal{M}} \int_{\mathcal{M}} \Phi(x, y) g(y) dy dx = \int_{\mathcal{M}} g(y) \int_{\mathcal{M}} \Phi(x, y) dx dy.$$

This implies that  $\mathcal{E}_{\mathbf{x}, \omega}(f)$  is well defined. Moreover,

$$\begin{aligned} |\mathcal{E}_{\mathbf{x}, \omega}(f)| &= \left| \sum_{j=1}^N \omega_j f(x_j) - \int_{\mathcal{M}} f(x) dx \right| \\ &= \left| \sum_{j=1}^N \omega_j \int_{\mathcal{M}} \Phi(x_j, y) g(y) dy - \int_{\mathcal{M}} g(y) \int_{\mathcal{M}} \Phi(x, y) dx dy \right| \\ &= \left| \int_{\mathcal{M}} \left( \sum_{j=1}^N \int_{\mathcal{X}_j} (\Phi(x_j, y) - \Phi(x, y)) dx \right) g(y) dy \right| \\ &\leq \left\{ \int_{\mathcal{M}} |g(y)|^p dy \right\}^{1/p} \left\{ \int_{\mathcal{M}} \left| \sum_{j=1}^N \int_{\mathcal{X}_j} (\Phi(x_j, y) - \Phi(x, y)) dx \right|^q dy \right\}^{1/q}. \end{aligned}$$

Taking the infimum as  $g(y)$  varies among all possible functions in  $\mathbb{L}^p(\mathcal{M})$  with potential  $f(x)$ , one obtains

$$\|\mathcal{E}_{\mathbf{x}, \omega}\|_{\Phi, p} \leq \left\{ \int_{\mathcal{M}} \left| \sum_{j=1}^N \int_{\mathcal{X}_j} (\Phi(x_j, y) - \Phi(x, y)) dx \right|^q dy \right\}^{1/q}.$$

Conversely, using the standard argument for  $L^p - L^q$  duality, set

$$\begin{aligned} F(y) &= \sum_{j=1}^N \int_{\mathcal{X}_j} (\Phi(x_j, y) - \Phi(x, y)) dx, \\ g(y) &= \begin{cases} \overline{F(y)} |F(y)|^{q/p-1} & \text{if } F(y) \neq 0, \\ 0 & \text{if } F(y) = 0, \end{cases} \end{aligned}$$

and

$$f(x) = \int_{\mathcal{M}} \Phi(x, y) g(y) dy.$$

Then

$$\begin{aligned} |\mathcal{E}_{\mathbf{x},\omega}(f)| &= \left| \sum_{j=1}^N \omega_j f(x_j) - \int_{\mathcal{M}} f(x) dx \right| = \left| \int_{\mathcal{M}} F(y) g(y) dy \right| \\ &= \int_{\mathcal{M}} |F(y)|^{1+q/p} dy = \left\{ \int_{\mathcal{M}} |F(y)|^q dy \right\}^{1/q} \left\{ \int_{\mathcal{M}} |F(y)|^q dy \right\}^{1/p} \\ &= \left\{ \int_{\mathcal{M}} |F(y)|^q dy \right\}^{1/q} \left\{ \int_{\mathcal{M}} |g(y)|^p dy \right\}^{1/p} \geq \left\{ \int_{\mathcal{M}} |F(y)|^q dy \right\}^{1/q} \|f\|_{\mathbb{H}_p^\Phi}. \end{aligned}$$

This implies that

$$\|\mathcal{E}_{\mathbf{x},\omega}\|_{\Phi,p} \geq \left\{ \int_{\mathcal{M}} |F(y)|^q dy \right\}^{1/q}.$$

□

**Theorem 6.2** *Assume that a measure space  $\mathcal{M}$  can be decomposed into a finite disjoint union of sets in the form  $\mathcal{M} = \mathcal{X}_1 \cup \dots \cup \mathcal{X}_N$ , with measure  $|\mathcal{X}_j| = \omega_j > 0$ . Assume also that  $1 \leq p \leq +\infty$ ,  $1 \leq q \leq +\infty$ ,  $1/p + 1/q = 1$ , and that for every  $x$ ,*

$$\left\{ \int_{\mathcal{M}} |\Phi(x,y)|^q dy \right\}^{1/q} < +\infty.$$

Finally, assume that

$$\left\{ \int_{\mathcal{M}} \left( \int_{\mathcal{M}} |\Phi(x,y)| dx \right)^q dy \right\}^{1/q} < +\infty.$$

Define

$$\Gamma(\Phi) = \sum_{j=1}^N \left\{ \int_{\mathcal{X}_j} \int_{\mathcal{M}} \left| \int_{\mathcal{X}_j} (\Phi(x_j,y) - \Phi(z_j,y)) dz_j \right|^q dy \frac{dx_j}{\omega_j} \right\}^{1/q},$$

and

$$\Delta(\Phi) = \left\{ \int_{\mathbf{X}} \int_{\mathcal{M}} \left( \sum_{j=1}^N \left| \int_{\mathcal{X}_j} (\Phi(x_j,y) - \Phi(z_j,y)) dz_j \right|^2 \right)^{q/2} dy d\mathbf{x} \right\}^{1/q}.$$

Then for every  $1 \leq p \leq +\infty$ ,

$$\left\{ \int_{\mathbf{X}} \|\mathcal{E}_{\mathbf{x},\omega}\|_{\Phi,p}^q d\mathbf{x} \right\}^{1/q} \leq \Gamma(\Phi), \quad (7)$$

and for every  $1 < p \leq +\infty$ ,

$$A(q) \Delta(\Phi) \leq \left\{ \int_{\mathbf{X}} \|\mathcal{E}_{\mathbf{x},\omega}\|_{\Phi,p}^q d\mathbf{x} \right\}^{1/q} \leq B(q) \Delta(\Phi). \quad (8)$$

In particular, there exist choices of nodes  $\{x_j\}$  with the property that for every function  $f(x)$  in the potential space  $\mathbb{H}_p^\Phi(\mathcal{M})$ ,

$$\left| \sum_{j=1}^N \omega_j f(x_j) - \int_{\mathcal{M}} f(x) dx \right| \leq \begin{cases} \Gamma(\Phi) \|f\|_{\mathbb{H}_p^\Phi(\mathcal{M})} & \text{for every } 1 \leq p \leq +\infty, \\ B(q) \Delta(\Phi) \|f\|_{\mathbb{H}_p^\Phi(\mathcal{M})} & \text{for every } 1 < p \leq +\infty. \end{cases}$$

The constants  $A(q)$  and  $B(q)$  are the best constants in the Marcinkiewicz–Zygmund inequality. The constants  $\Gamma(\Phi)$  and  $\Delta(\Phi)$  are related to the smoothness of the kernel  $\Phi(x, y)$ . These last constants could be estimated in terms of Sobolev norms. However, in the applications, the estimates in terms of Sobolev norms are not always optimal, and it is more convenient to keep the above complicated expressions. Finally, since  $B(q) \rightarrow +\infty$  as  $p \rightarrow 1+$ , the interest of the estimate (7) is when  $p \rightarrow 1+$ .

*Proof* By Lemma 6.1 and the triangle inequality,

$$\begin{aligned} \left\{ \int_{\mathbf{X}} \|\mathcal{E}_{\mathbf{x}, \omega}\|_{\Phi, p}^q d\mathbf{x} \right\}^{1/q} &= \left\{ \int_{\mathbf{X}} \int_{\mathcal{M}} \left| \sum_{j=1}^N \int_{\mathcal{X}_j} (\Phi(x_j, y) - \Phi(z_j, y)) dz_j \right|^q dy d\mathbf{x} \right\}^{1/q} \\ &\leq \sum_{j=1}^N \left\{ \int_{\mathcal{X}_j} \int_{\mathcal{M}} \left| \int_{\mathcal{X}_j} (\Phi(x_j, y) - \Phi(z_j, y)) dz_j \right|^q dy \frac{dx_j}{\omega_j} \right\}^{1/q}, \end{aligned}$$

where we have used the fact that for a function depending only on  $x_j$ , integration on  $\mathbf{X}$  coincides with integration on  $\mathcal{X}_j$ . This gives the proof with  $\Gamma(\Phi)$ . The proof with  $\Delta(\Phi)$  is similar, with the crucial difference that we replace the triangle inequality with the Marcinkiewicz–Zygmund inequality. Indeed, by Corollary 4.2,

$$\begin{aligned} \left\{ \int_{\mathbf{X}} \|\mathcal{E}_{\mathbf{x}, \omega}\|_{\Phi, p}^q d\mathbf{x} \right\}^{1/q} &= \left\{ \int_{\mathcal{M}} \int_{\mathbf{X}} \left| \sum_{j=1}^N \int_{\mathcal{X}_j} (\Phi(x_j, y) - \Phi(z_j, y)) dz_j \right|^q d\mathbf{x} dy \right\}^{1/q} \\ &\leq B(q) \left\{ \int_{\mathcal{M}} \int_{\mathbf{X}} \left( \sum_{j=1}^N \left| \int_{\mathcal{X}_j} (\Phi(x_j, y) - \Phi(z_j, y)) dz_j \right|^2 \right)^{q/2} d\mathbf{x} dy \right\}^{1/q}. \end{aligned}$$

The proof for the lower bound is similar.  $\square$

The following corollary is a slightly generalized version of Theorem I in the Introduction.

**Corollary 6.3** *Let  $\mathcal{M}$  be a metric measure space with the property that there exist  $d$  and  $c$  such that for every  $y \in \mathcal{M}$  and  $r > 0$ ,*

$$|\{x \in \mathcal{M} : |x - y| \leq r\}| \leq cr^d.$$

*Assume also that  $\mathcal{M}$  can be decomposed into a finite disjoint union of sets in the form  $\mathcal{M} = \mathcal{X}_1 \cup \dots \cup \mathcal{X}_N$ , with  $\omega_j = |\mathcal{X}_j| \approx N^{-1}$  and  $\delta_j = \text{diam}(\mathcal{X}_j) \approx N^{-1/d}$ . Assume that for some  $\varepsilon > 0$  and  $0 < \alpha < d$ ,*

$$|\Phi(x, y)| \leq c|x - y|^{\alpha-d}$$

*for every  $x$  and  $y$ , and*

$$|\Phi(x, y) - \Phi(z, y)| \leq c|x - z|^\varepsilon |x - y|^{\alpha-d-\varepsilon}$$

*if  $|x - y| \geq 2|x - z|$ . Finally, assume that  $1 < p \leq +\infty$ ,  $1/p + 1/q = 1$  and  $d/p < \alpha < d$ . Then*

$$\left\{ \int_{\mathbf{X}} \|\mathcal{E}_{\mathbf{x}, \omega}\|_{\Phi, p}^q d\mathbf{x} \right\}^{1/q} \leq \begin{cases} cN^{-\alpha/d} & \text{if } \alpha < d/2 + \varepsilon, \\ cN^{-1/2-\varepsilon/d} (\log N)^{1/2} & \text{if } \alpha = d/2 + \varepsilon, \\ cN^{-1/2-\varepsilon/d} & \text{if } \alpha > d/2 + \varepsilon. \end{cases}$$

*Proof* The assumption  $\alpha > d/p$  ensures that the kernel  $\Phi(x, y)$  is  $q$  integrable and satisfies the hypotheses of Theorem 6.2. It then suffices to estimate

$$\Delta(\Phi) = \left\{ \int_{\mathcal{M}} \int_{\mathbf{X}} \left( \sum_{j=1}^N \left| \int_{\mathcal{X}_j} (\Phi(x_j, y) - \Phi(z_j, y)) dz_j \right|^2 \right)^{q/2} dx dy \right\}^{1/q}.$$

If  $\text{dist}(y, \mathcal{X}_j) \leq 2\delta_j$ , then for every  $x_j$  in  $\mathcal{X}_j$ ,

$$\begin{aligned} \left| \int_{\mathcal{X}_j} (\Phi(x_j, y) - \Phi(z_j, y)) dz_j \right| &\leq c \int_{\mathcal{X}_j} (|x_j - y|^{\alpha-d} + |z_j - y|^{\alpha-d}) dz_j \\ &\leq c\omega_j |x_j - y|^{\alpha-d} + c\delta_j^\alpha \leq cN^{-1} |x_j - y|^{\alpha-d}. \end{aligned}$$

If  $\text{dist}(y, \mathcal{X}_j) \geq 2\delta_j$ , then

$$\begin{aligned} \left| \int_{\mathcal{X}_j} (\Phi(x_j, y) - \Phi(z_j, y)) dz_j \right| &\leq c \int_{\mathcal{X}_j} |x_j - z_j|^\varepsilon |x_j - y|^{\alpha-d-\varepsilon} dz_j \\ &\leq c\delta_j^\varepsilon \omega_j |x_j - y|^{\alpha-d-\varepsilon} \leq cN^{-1-\varepsilon/d} |x_j - y|^{\alpha-d-\varepsilon}. \end{aligned}$$

Hence

$$\begin{aligned} &\left\{ \int_{\mathcal{M}} \int_{\mathbf{X}} \left( \sum_{j=1}^N \left| \int_{\mathcal{X}_j} (\Phi(x_j, y) - \Phi(z_j, y)) dz_j \right|^2 \right)^{q/2} dx dy \right\}^{1/q} \\ &\leq c \left\{ \int_{\mathcal{M}} \int_{\mathbf{X}} \left( N^{-2} \sum_{j: \text{dist}(y, \mathcal{X}_j) \leq 2\delta_j} |x_j - y|^{2\alpha-2d} \right)^{q/2} dx dy \right\}^{1/q} \\ &\quad + c \left\{ \int_{\mathcal{M}} \int_{\mathbf{X}} \left( N^{-2-2\varepsilon/d} \sum_{j: \text{dist}(y, \mathcal{X}_j) > 2\delta_j} |x_j - y|^{2\alpha-2d-2\varepsilon} \right)^{q/2} dx dy \right\}^{1/q}. \end{aligned}$$

Under the assumption that  $\text{diam}(\mathcal{X}_j) \approx N^{-1/d}$ , there is only a bounded number of  $\mathcal{X}_j$  with  $\text{dist}(y, \mathcal{X}_j) \leq 2 \text{diam}(\mathcal{X}_j)$ . Hence

$$\begin{aligned} &\left\{ \int_{\mathcal{M}} \int_{\mathbf{X}} \left( N^{-2} \sum_{j: \text{dist}(y, \mathcal{X}_j) \leq 2\delta_j} |x_j - y|^{2\alpha-2d} \right)^{q/2} dx dy \right\}^{1/q} \\ &\leq c \left\{ \int_{\mathcal{M}} \int_{\mathbf{X}} \sum_{j=1}^N \left( N^{-2} \chi_{\{\text{dist}(y, \mathcal{X}_j) \leq 2\delta_j\}}(y) |x_j - y|^{2\alpha-2d} \right)^{q/2} dx dy \right\}^{1/q} \\ &\leq c \left\{ N^{-q} \sum_{j=1}^N \int_{\mathcal{X}_j} \int_{\{|y-x_j| \leq cN^{-1/d}\}} |x_j - y|^{\alpha q - dq} dy \frac{dx_j}{\omega_j} \right\}^{1/q} \leq cN^{-\alpha/d}. \end{aligned}$$

Moreover,

$$\begin{aligned}
& N^{-2-2\varepsilon/d} \sum_{j:\text{dist}(y, \mathcal{X}_j) > 2\delta_j} |x_j - y|^{2\alpha-2d-2\varepsilon} \\
& \leq cN^{-1-2\varepsilon/d} \sum_{j:\text{dist}(y, \mathcal{X}_j) > 2\delta_j} \omega_j |x_j - y|^{2\alpha-2d-2\varepsilon} \\
& \leq cN^{-1-2\varepsilon/d} \sum_{j:\text{dist}(y, \mathcal{X}_j) > 2\delta_j} \int_{\mathcal{X}_j} |x - y|^{2\alpha-2d-2\varepsilon} dx \\
& \leq cN^{-1-2\varepsilon/d} \int_{\{|x-y| > cN^{-1/d}\}} |x - y|^{2\alpha-2d-2\varepsilon} dx \\
& \leq \begin{cases} cN^{-2\alpha/d} & \text{if } \alpha < \varepsilon + d/2, \\ cN^{-1-2\varepsilon/d} \log N & \text{if } \alpha = \varepsilon + d/2, \\ cN^{-1-2\varepsilon/d} & \text{if } \alpha > \varepsilon + d/2. \end{cases}
\end{aligned}$$

Hence

$$\begin{aligned}
& \left\{ \int_{\mathcal{M}} \int_{\mathbf{X}} \left( N^{-2-2\varepsilon/d} \sum_{j:\text{dist}(y, \mathcal{X}_j) > 2\delta_j} |x_j - y|^{2\alpha-2d-2\varepsilon} \right)^{q/2} d\mathbf{x} dy \right\}^{1/q} \\
& \leq \begin{cases} cN^{-\alpha/d} & \text{if } \alpha < \varepsilon + d/2, \\ cN^{-1/2-\varepsilon/d} (\log N)^{1/2} & \text{if } \alpha = \varepsilon + d/2, \\ cN^{-1/2-\varepsilon/d} & \text{if } \alpha > \varepsilon + d/2. \end{cases}
\end{aligned}$$

□

The following result shows that under some natural assumptions on the kernel, the mean value estimate in the above corollary is essentially sharp, and is a slightly generalized version of Theorem II in the Introduction.

**Corollary 6.4** *Let  $\mathcal{M}$  be a metric measure space with the property that there exist  $H, K$  and  $d$  such that for every  $y \in \mathcal{M}$  and  $0 < r < r_0$ ,*

$$Hr^d \leq |\{x \in \mathcal{M} : |x - y| \leq r\}| \leq Kr^d.$$

*Assume also that  $\mathcal{M}$  can be decomposed into a finite disjoint union of sets in the form  $\mathcal{M} = \mathcal{X}_1 \cup \dots \cup \mathcal{X}_N$ , with  $\omega_j = |\mathcal{X}_j| \approx N^{-1}$  and  $\delta_j = \text{diam}(\mathcal{X}_j) \approx N^{-1/d}$ . Suppose that there exists  $0 < \alpha < d$  and  $\varepsilon > 0$ , such that for any  $j = 1, \dots, N$  and any  $z \in \mathcal{X}_j$ , and for any  $y$  such that  $\text{dist}(y, \mathcal{X}_j) \geq 2\delta_j$ ,*

$$\int_{\mathcal{X}_j} |\Phi(x, y) - \Phi(z, y)| dx \geq cN^{-1-\varepsilon/d} (\text{dist}(y, \mathcal{X}_j))^{\alpha-d-\varepsilon}.$$

*Suppose also that for any  $y \in \mathcal{M}$ , the function  $x \mapsto \Phi(x, y)$  is continuous in  $x \neq y$ . Finally, assume that  $1 < p \leq +\infty$ ,  $1/p + 1/q = 1$  and  $d/p < \alpha < d$ . Then*

$$\left\{ \int_{\mathbf{X}} \|\mathcal{E}_{\mathbf{x}, \omega}\|_{\Phi, p}^q d\mathbf{x} \right\}^{1/q} \geq \begin{cases} cN^{-\alpha/d} & \text{if } \alpha < d/2 + \varepsilon, \\ cN^{-1/2-\varepsilon/d} (\log N)^{1/2} & \text{if } \alpha = d/2 + \varepsilon, \\ cN^{-1/2-\varepsilon/d} & \text{if } \alpha > d/2 + \varepsilon. \end{cases}$$

*Proof* It follows from Lemma 6.1 that

$$\begin{aligned} \left\{ \int_{\mathbf{X}} \|\mathcal{E}_{\mathbf{x}, \omega}\|^q d\mathbf{x} \right\}^{1/q} &= \left\{ \int_{\mathbf{X}} \int_{\mathcal{M}} \left| \sum_{j=1}^N \int_{\mathcal{X}_j} (\Phi(x_j, y) - \Phi(z_j, y)) dz_j \right|^q dy d\mathbf{x} \right\}^{1/q} \\ &\geq |\mathcal{M}|^{-1/p} \int_{\mathbf{X}} \int_{\mathcal{M}} \left| \sum_{j=1}^N \int_{\mathcal{X}_j} (\Phi(x_j, y) - \Phi(z_j, y)) dz_j \right| dy d\mathbf{x}. \end{aligned}$$

By Corollary 4.2, this is bounded from below by

$$\begin{aligned} &|\mathcal{M}|^{-1/p} A(1) \int_{\mathcal{M}} \int_{\mathbf{X}} \left( \sum_j \left| \int_{\mathcal{X}_j} (\Phi(x_j, y) - \Phi(z_j, y)) dz_j \right|^2 \right)^{1/2} d\mathbf{x} dy \\ &\geq |\mathcal{M}|^{-1/p} A(1) \int_{\mathcal{M}} \int_{\mathbf{X}} \left( \sum_{j: \text{dist}(y, \mathcal{X}_j) \geq 2\delta_j} \left| \int_{\mathcal{X}_j} (\Phi(x_j, y) - \Phi(z_j, y)) dz_j \right|^2 \right)^{1/2} d\mathbf{x} dy. \end{aligned}$$

By the continuity of  $z_j \rightarrow \Phi(z_j, y)$ , there exists a point  $x_j^*$ , depending on  $y$ , such that

$$\int_{\mathcal{X}_j} \Phi(z_j, y) dz_j = \omega_j \Phi(x_j^*, y).$$

Thus

$$\begin{aligned} &\int_{\mathbf{X}} \left( \sum_{j: \text{dist}(y, \mathcal{X}_j) \geq 2\delta_j} \left| \int_{\mathcal{X}_j} (\Phi(x_j, y) - \Phi(z_j, y)) dz_j \right|^2 \right)^{1/2} d\mathbf{x} \\ &= \int_{\mathbf{X}} \left( \sum_{j: \text{dist}(y, \mathcal{X}_j) \geq 2\delta_j} \omega_j^2 |\Phi(x_j, y) - \Phi(x_j^*, y)|^2 \right)^{1/2} d\mathbf{x}. \end{aligned}$$

For any two positive sequences  $\{\alpha_j\}$  and  $\{\beta_j\}$  we clearly have

$$\sum_j \alpha_j^2 \beta_j \leq \left( \sum_j \alpha_j^2 \right)^{1/2} \left( \sum_j \alpha_j^2 \beta_j^2 \right)^{1/2}.$$

Thus if  $v_j$  is a point of the closure of  $\mathcal{X}_j$  that minimizes the distance from  $y$ ,

$$\begin{aligned}
& \int_{\mathbf{X}} \left( \sum_{j: \text{dist}(y, \mathcal{X}_j) \geq 2\delta_j} \omega_j^2 |v_j - y|^{2\alpha - 2d - 2\varepsilon} \frac{|\Phi(x_j, y) - \Phi(x_j^*, y)|^2}{|v_j - y|^{2\alpha - 2d - 2\varepsilon}} \right)^{1/2} d\mathbf{x} \\
& \geq \int_{\mathbf{X}} \left( \sum_{j: \text{dist}(y, \mathcal{X}_j) \geq 2\delta_j} \omega_j^2 |v_j - y|^{2\alpha - 2d - 2\varepsilon} \right)^{-1/2} \times \\
& \quad \times \sum_{j: \text{dist}(y, \mathcal{X}_j) \geq 2\delta_j} \omega_j^2 |v_j - y|^{2\alpha - 2d - 2\varepsilon} \frac{|\Phi(x_j, y) - \Phi(x_j^*, y)|}{|v_j - y|^{\alpha - d - \varepsilon}} d\mathbf{x} \\
& = \left( \sum_{j: \text{dist}(y, \mathcal{X}_j) \geq 2\delta_j} \omega_j^2 |v_j - y|^{2\alpha - 2d - 2\varepsilon} \right)^{-1/2} \times \\
& \quad \times \sum_{j: \text{dist}(y, \mathcal{X}_j) \geq 2\delta_j} \omega_j^2 |v_j - y|^{\alpha - d - \varepsilon} \left( \int_{\mathbf{X}} |\Phi(x_j, y) - \Phi(x_j^*, y)| d\mathbf{x} \right) \\
& \geq c \left( \sum_{j: \text{dist}(y, \mathcal{X}_j) \geq 2\delta_j} \omega_j^2 |v_j - y|^{2\alpha - 2d - 2\varepsilon} \right)^{-1/2} \sum_{j: \text{dist}(y, \mathcal{X}_j) \geq 2\delta_j} \omega_j^2 |v_j - y|^{2\alpha - 2d - 2\varepsilon} N^{-\varepsilon/d} \\
& = c \left( \sum_{j: \text{dist}(y, \mathcal{X}_j) \geq 2\delta_j} \omega_j^2 |v_j - y|^{2\alpha - 2d - 2\varepsilon} \right)^{1/2} N^{-\varepsilon/d}.
\end{aligned}$$

The desired result now follows from the estimates

$$\sum_{j: \text{dist}(y, \mathcal{X}_j) \geq 2\delta_j} \omega_j^2 |v_j - y|^{2\alpha - 2d - 2\varepsilon} \geq c \begin{cases} N^{-2\alpha/d + 2\varepsilon/d} & \text{if } \alpha < d/2 + \varepsilon, \\ N^{-1} \log N & \text{if } \alpha = d/2 + \varepsilon, \\ N^{-1} & \text{if } \alpha > d/2 + \varepsilon. \end{cases}$$

Indeed,

$$\sum_{j: \text{dist}(y, \mathcal{X}_j) \geq 2\delta_j} \omega_j^2 |v_j - y|^{2\alpha - 2d - 2\varepsilon} \geq cN^{-1} \int_{\{|x-y| > cN^{-1/d}\}} |x - y|^{2\alpha - 2d - 2\varepsilon} dx.$$

If  $\lambda \geq (2K/H)^{1/d}$ , then for every centre  $y$  and  $r < r_0/\lambda$ ,

$$|\{x \in \mathcal{M} : r < |x - y| \leq \lambda r\}| \geq Kr^d.$$

This gives

$$\begin{aligned}
& N^{-1} \int_{\{|x-y| > cN^{-1/d}\}} |x-y|^{2\alpha-2d-2\varepsilon} dx \\
& \geq N^{-1} \sum_{k=0}^{[\log_\lambda(r_0 N^{1/d})]-1} \int_{\{\lambda^k N^{-1/d} \leq |x-y| < \lambda^{k+1} N^{-1/d}\}} |x-y|^{2\alpha-2d-2\varepsilon} dx \\
& \geq N^{-1} \sum_{k=0}^{[\log_\lambda(r_0 N^{1/d})]-1} \left(\lambda^k N^{-1/d}\right)^{2\alpha-2d-2\varepsilon} K \left(\lambda^k N^{-1/d}\right)^d \\
& = KN^{-2\alpha/d+2\varepsilon/d} \sum_{k=0}^{[\log_\lambda(r_0 N^{1/d})]-1} \lambda^{k(2\alpha-d-2\varepsilon)} \\
& \geq c \begin{cases} N^{-2\alpha/d+2\varepsilon/d} & \text{if } \alpha < d/2 + \varepsilon, \\ N^{-1} \log N & \text{if } \alpha = d/2 + \varepsilon, \\ N^{-1} & \text{if } \alpha > d/2 + \varepsilon. \end{cases}
\end{aligned}$$

□

*Example 6.5* Let  $\mathcal{M}$  be a  $d$ -dimensional compact Riemannian manifold. Let  $\{\lambda^2\}$  and  $\{\varphi_\lambda(x)\}$  be the eigenvalues and a complete orthonormal system of eigenfunctions of the Laplace Beltrami operator  $\Delta$  respectively. Every tempered distribution on  $\mathcal{M}$  has Fourier transform and series

$$\mathcal{F}f(\lambda) = \int_{\mathcal{M}} f(y) \overline{\varphi_\lambda(y)} dy \quad \text{and} \quad f(x) = \sum_{\lambda} \mathcal{F}f(\lambda) \varphi_\lambda(x).$$

The Bessel kernel  $B^\alpha(x, y)$ ,  $-\infty < \alpha < +\infty$ , is a distribution defined by the expansion

$$B^\alpha(x, y) = \sum_{\lambda} \left(1 + \lambda^2\right)^{-\alpha/2} \varphi_\lambda(x) \overline{\varphi_\lambda(y)}.$$

A distribution  $f(x)$  is the Bessel potential of a distribution  $g(x)$  if

$$f(x) = \int_{\mathcal{M}} B^\alpha(x, y) g(y) dy = \sum_{\lambda} \left(1 + \lambda^2\right)^{-\alpha/2} \mathcal{F}g(\lambda) \varphi_\lambda(x).$$

Bessel potentials of functions in  $\mathbb{L}^p(\mathcal{M})$  define the fractional Sobolev space  $\mathbb{H}_p^\alpha(\mathcal{M})$ . If  $0 < \alpha < d$ , then the Bessel kernel satisfies the estimates

$$|B^\alpha(x, y)| \leq c|x-y|^{\alpha-d} \quad \text{and} \quad |\nabla B^\alpha(x, y)| \leq c|x-y|^{\alpha-d-1}.$$

See [8, Lemmas 2.5 and 2.6]. In particular, Corollary 6.3 applies with  $\Phi(x, y) = B^\alpha(x, y)$  and  $\varepsilon = 1$ . Indeed, using the Hadamard parametrix for the wave equation, see, for example, [7], one can prove a more precise result: there is a smooth positive function  $C(y)$  and positive constants  $\eta$  and  $c$  such that

$$B^\alpha(x, y) = C(y) |x-y|^{\alpha-d} + E(x, y),$$

with

$$|E(x, y)| \leq c|x-y|^{\alpha-d+\eta} \quad \text{and} \quad |\nabla E(x, y)| \leq c|x-y|^{\alpha-d-1+\eta}.$$

It then follows that Corollary 6.4 applies also. See [14, Theorems 24 and 25] for the case of the sphere.



## 7 Numerical integration in Besov spaces

The techniques of the previous section can also be used to study the error in numerical integration from a different perspective. So far we have considered the worst case error

$$\left\{ \int_{\mathbf{X}} \sup_{\|f\|_{\dot{\mathbb{B}}_p^\varphi(\mathcal{M})} \leq 1} |\mathcal{E}_{\mathbf{x},\omega}(f)|^q d\mathbf{x} \right\}^{1/q},$$

whereas now we will estimate the error

$$\left\{ \int_{\mathbf{X}} |\mathcal{E}_{\mathbf{x},\omega}(f)|^p d\mathbf{x} \right\}^{1/p}.$$

for a given  $f \in \dot{\mathbb{B}}_{p,\infty}^\varphi(\mathcal{M})$ . The following is a slightly generalized version of Theorem III in the Introduction.

**Theorem 7.1** *Assume that a metric measure space  $\mathcal{M}$  can be decomposed into a finite disjoint union of sets in the form  $\mathcal{M} = \mathcal{X}_1 \cup \dots \cup \mathcal{X}_N$ , with measure  $0 < |\mathcal{X}_j| = \omega_j < +\infty$  and  $0 < \text{diam}(\mathcal{X}_j) = \delta_j < +\infty$ . Also let  $\varphi(t)$  be a non-negative increasing function in  $t \geq 0$ , and let  $\dot{\mathbb{B}}_{p,\infty}^\varphi(\mathcal{M})$  be the associated Besov space. Then for every  $1 \leq p \leq +\infty$ ,*

$$\left\{ \int_{\mathbf{X}} |\mathcal{E}_{\mathbf{x},\omega}(f)|^p d\mathbf{x} \right\}^{1/p} \leq 2 |\mathcal{M}|^{1-1/p} \varphi(2 \sup \{\delta_j\}) \|f\|_{\dot{\mathbb{B}}_{p,\infty}^\varphi(\mathcal{M})}. \quad (9)$$

Furthermore, if  $1 \leq p \leq 2$ , then

$$\left\{ \int_{\mathbf{X}} |\mathcal{E}_{\mathbf{x},\omega}(f)|^p d\mathbf{x} \right\}^{1/p} \leq 2B(p) \sup \{\omega_j^{1-1/p}\} \varphi(2 \sup \{\delta_j\}) \|f\|_{\dot{\mathbb{B}}_{p,\infty}^\varphi(\mathcal{M})}, \quad (10)$$

and if  $2 \leq p < +\infty$ , then

$$\left\{ \int_{\mathbf{X}} |\mathcal{E}_{\mathbf{x},\omega}(f)|^p d\mathbf{x} \right\}^{1/p} \leq 2B(p) |\mathcal{M}|^{1/2-1/p} \sup \{\omega_j^{1/2}\} \varphi(2 \sup \{\delta_j\}) \|f\|_{\dot{\mathbb{B}}_{p,\infty}^\varphi(\mathcal{M})}. \quad (11)$$

Observe that the estimate (9) is of some interest only for large  $p$ . Indeed, if  $1 \leq p \leq 2$ , then (10) is better than (9), and if  $2 \leq p < +\infty$  and  $B(p) \leq |\mathcal{M}|^{1/2} \left( \sup \{\omega_j^{1/2}\} \right)^{-1}$ , then (11) is better than (9).

In order to prove Theorem 7.1, we need a Poincaré type inequality for functions in Hajlasz–Besov spaces.

**Lemma 7.2** *Let  $1 \leq p \leq +\infty$ , let  $\mathcal{M}$  be a metric measure space, and let  $\{g_n(x)\}$  be a  $\varphi$ -gradient for an integrable function  $f(x)$ . Let  $\mathcal{X}$  be a measurable subset of  $\mathcal{M}$  with  $\omega = |\mathcal{X}| > 0$  and  $\text{diam}(\mathcal{X}) \leq 2^{-n}$ , and let*

$$f_{\mathcal{X}} = \frac{1}{\omega} \int_{\mathcal{X}} f(y) dy.$$

Then

$$\left\{ \int_{\mathcal{X}} |f(x) - f_{\mathcal{X}}|^p \frac{dx}{\omega} \right\}^{1/p} \leq 2\varphi(2^{-n}) \left\{ \int_{\mathcal{X}} |g_n(x)|^p \frac{dx}{\omega} \right\}^{1/p}.$$

*Proof* For almost every  $x$  and  $y$  with  $|x - y| \leq 2^{-n}$ , we have

$$|f(x) - f(y)| \leq \varphi(2^{-n})(g_n(x) + g_n(y)).$$

Then, by Hölder's inequality, we obtain

$$\begin{aligned} \left\{ \int_{\mathcal{X}} |f(x) - f_{\mathcal{X}}|^p \frac{dx}{\omega} \right\}^{1/p} &\leq \left\{ \int_{\mathcal{X}} \int_{\mathcal{X}} |f(x) - f(y)|^p \frac{dx}{\omega} \frac{dy}{\omega} \right\}^{1/p} \\ &\leq \varphi(2^{-n}) \left\{ \int_{\mathcal{X}} \int_{\mathcal{X}} |g_n(x) + g_n(y)|^p \frac{dx}{\omega} \frac{dy}{\omega} \right\}^{1/p} \leq 2\varphi(2^{-n}) \left\{ \int_{\mathcal{X}} |g_n(x)|^p \frac{dx}{\omega} \right\}^{1/p}. \end{aligned}$$

□

*Proof (of Theorem 7.1)* Let  $f \in \mathbb{B}_{p,\infty}^{\varphi}(\mathcal{M})$ , and let  $\{g_n(x)\}$  be a  $\varphi$ -gradient for  $f(x)$ . Choose  $n$  such that  $2^{-n-1} < \sup \delta_j \leq 2^{-n}$ . Then by Lemma 7.2, we have

$$\begin{aligned} \left\{ \int_{\mathbf{X}} \left| \sum_j \omega_j f(x_j) - \int_{\mathcal{M}} f(x) dx \right|^p d\mathbf{x} \right\}^{1/p} &= \left\{ \int_{\mathbf{X}} \left| \sum_j \omega_j (f(x_j) - f_{\mathcal{X}_j}) \right|^p d\mathbf{x} \right\}^{1/p} \\ &\leq \sum_j \omega_j \left\{ \int_{\mathcal{X}_j} |f(x_j) - f_{\mathcal{X}_j}|^p \frac{dx_j}{\omega_j} \right\}^{1/p} \leq 2\varphi(2^{-n}) \sum_j \omega_j \left\{ \int_{\mathcal{X}_j} |g_n(x_j)|^p \frac{dx_j}{\omega_j} \right\}^{1/p} \\ &\leq 2\varphi(2^{-n}) \left\{ \sum_j \omega_j \right\}^{1-1/p} \left\{ \sum_j \int_{\mathcal{X}_j} |g_n(x_j)|^p dx_j \right\}^{1/p} \\ &\leq 2\varphi(2 \sup \delta_j) |\mathcal{M}|^{1-1/p} \left\{ \int_{\mathcal{M}} |g_n(x)|^p dx \right\}^{1/p}. \end{aligned}$$

The proofs of (10) and (11) are similar. Indeed, by Corollary 4.2, we have

$$\begin{aligned} \left\{ \int_{\mathbf{X}} \left| \sum_{j=1}^N \omega_j f(x_j) - \int_{\mathcal{M}} f(x) dx \right|^p d\mathbf{x} \right\}^{1/p} \\ \leq B(p) \left\{ \int_{\mathbf{X}} \left( \sum_{j=1}^N \omega_j^2 |f(x_j) - f_{\mathcal{X}_j}|^2 \right)^{p/2} d\mathbf{x} \right\}^{1/p}. \end{aligned}$$

Choose  $n$  such that  $2^{-n-1} < \sup \delta_j \leq 2^{-n}$ , and assume that  $1 \leq p \leq 2$ . By Lemma 7.2, we obtain

$$\begin{aligned} \left\{ \int_{\mathbf{X}} \left( \sum_{j=1}^N \omega_j^2 |f(x_j) - f_{\mathcal{X}_j}|^2 \right)^{p/2} d\mathbf{x} \right\}^{1/p} &\leq \left\{ \int_{\mathbf{X}} \left( \sum_{j=1}^N \omega_j^p |f(x_j) - f_{\mathcal{X}_j}|^p \right) d\mathbf{x} \right\}^{1/p} \\ &= \left\{ \sum_{j=1}^N \omega_j^p \int_{\mathcal{X}_j} |f(x_j) - f_{\mathcal{X}_j}|^p \frac{dx_j}{\omega_j} \right\}^{1/p} \leq 2\varphi(2^{-n}) \left\{ \sum_{j=1}^N \omega_j^p \int_{\mathcal{X}_j} |g_n(x_j)|^p \frac{dx_j}{\omega_j} \right\}^{1/p} \\ &\leq 2 \sup \left( \omega_j^{1-1/p} \right) \varphi(2 \sup \{\delta_j\}) \left\{ \int_{\mathcal{M}} |g_n(x)|^p dx \right\}^{1/p}. \end{aligned}$$

Similarly, if  $2 \leq p < +\infty$ , then Hölder's inequality with indices  $p/(p-2)$  and  $p/2$  yields

$$\begin{aligned}
& \left\{ \int_{\mathbf{X}} \left( \sum_{j=1}^N \omega_j^2 |f(x_j) - f_{\mathcal{X}_j}|^2 \right)^{p/2} dx \right\}^{1/p} \\
& \leq \left\{ \sum_{j=1}^N \omega_j^{(2p-2)/(p-2)} \right\}^{(p-2)/2p} \left\{ \sum_{j=1}^N \omega_j \int_{\mathcal{X}_j} |f(x_j) - f_{\mathcal{X}_j}|^p \frac{dx_j}{\omega_j} \right\}^{1/p} \\
& \leq \sup(\omega_j^{1/2}) \left\{ \sum_{j=1}^N \omega_j \right\}^{(p-2)/2p} 2\varphi(2^{-n}) \left\{ \sum_{j=1}^N \omega_j \int_{\mathcal{X}_j} |g_n(x_j)|^p \frac{dx_j}{\omega_j} \right\}^{1/p} \\
& \leq 2|\mathcal{M}|^{1/2-1/p} \sup(\omega_j^{1/2}) \varphi(2 \sup \delta_j) \left\{ \int_{\mathcal{M}} |g_n(x)|^p dx \right\}^{1/p}.
\end{aligned}$$

□

**Corollary 7.3** *Assume that a metric measure space  $\mathcal{M}$  can be decomposed into a finite disjoint union of sets in the form  $\mathcal{M} = \mathcal{X}_1 \cup \dots \cup \mathcal{X}_N$ , with  $\omega_j = |\mathcal{X}_j| \approx N^{-1}$  and  $\text{diam}(\mathcal{X}_j) \approx N^{-1/d}$  for a suitable positive constant  $d$ . Then for every  $1 \leq p < +\infty$  and every  $0 < \varepsilon < 1$ , there exists a constant  $c$  with the following property. For every function  $f(x)$  in the Besov space  $\dot{\mathbb{B}}_{p,\infty}^\varphi(\mathcal{M})$ ,  $\varphi(t) = t^\alpha$  and  $\alpha > 0$ , with probability greater than  $1 - \varepsilon$ , a random choice of points  $\{x_j\}$  in  $\{\mathcal{X}_j\}$  gives*

$$\left| \sum_{j=1}^N \omega_j f(x_j) - \int_{\mathcal{M}} f(x) dx \right| \leq \begin{cases} c \|f\|_{\dot{\mathbb{B}}_{p,\infty}^\varphi(\mathcal{M})} N^{1/p-1-\alpha/d} & \text{if } 1 \leq p \leq 2, \\ c \|f\|_{\dot{\mathbb{B}}_{p,\infty}^\varphi(\mathcal{M})} N^{-1/2-\alpha/d} & \text{if } 2 \leq p < +\infty. \end{cases}$$

*Proof* This follows from Theorem 7.1 via Chebyshev's inequality. □

The following example shows that Theorem 7.1 is essentially sharp.

*Example 7.4* As in Theorem 5.1, let  $\mathcal{M}$  be a metric measure space of finite measure with the property that there exist positive constants  $H, K, d$ , such that for every  $y \in \mathcal{M}$  and  $0 < r < \text{diam}(\mathcal{M})$ ,

$$Hr^d \leq |\{x \in \mathcal{M} : |x - y| < r\}| \leq Kr^d.$$

For every  $N$ , the space  $\mathcal{M}$  can be decomposed into a finite disjoint union of sets in the form  $\mathcal{M} = \mathcal{X}_1 \cup \dots \cup \mathcal{X}_N$ , with  $\omega_j = |\mathcal{X}_j| \approx N^{-1}$  and  $\delta_j = \text{diam}(\mathcal{X}_j) \approx N^{-1/d}$ . Moreover every  $\mathcal{X}_j$  contains a ball  $B(w_j, r_j)$  of centre  $w_j$  and radius  $r_j \approx \delta_j$ . It is possible to prove that each ball  $B(w_j, r_j)$  contains two disjoint balls  $B(y_j, \varepsilon r_j)$  and  $B(z_j, \varepsilon r_j)$ , with  $\varepsilon = 6^{-1}(2K/H)^{-1/d}$ . Fix  $1 \leq j \leq N$  and define

$$f_j(x) = \left(1 - \frac{2}{\varepsilon r_j} |x - y_j|\right)_+ - \vartheta_j \left(1 - \frac{2}{\varepsilon r_j} |x - z_j|\right)_+,$$

with  $\vartheta_j$  such that  $f$  has mean 0. Observe that  $0 < \vartheta_j < C$  independent of  $N$ . Also note that there exist constants  $A$  and  $B$  independent of  $N$  such that

$$\int |f_j(x)|^p \frac{dx}{\omega_j} \geq A,$$

and

$$|f_j(x) - f_j(y)| \leq B\delta_j^{-1} |x - y| \quad \text{for all } x, y \in \mathcal{M}.$$

If  $\varphi(t) = t^\alpha$  with  $d/p < \alpha \leq 1$ , and if  $w_j$  is the above defined point in  $\mathcal{X}_j$ , then the function

$$g_j(x) = c \min\left(\delta_j^{-\alpha}, |x - w_j|^{-\alpha}\right)$$

is a  $\varphi$ -gradient of  $f(x)$ . Indeed, if  $x \in \mathcal{X}_j$  and  $|x - y| \leq 2\delta_j$ , then

$$|f_j(x) - f_j(y)| \leq B\delta_j^{-1} |x - y| \leq c\delta_j^{-\alpha} |x - y|^\alpha,$$

while for  $x \in \mathcal{X}_j$  and  $|x - y| > 2\delta_j$ ,

$$\begin{aligned} |f_j(x) - f_j(y)| &= |f_j(x)| \leq \max\{1, \vartheta_j\} = \max\{1, \vartheta_j\} |x - y|^{-\alpha} |x - y|^\alpha \\ &\leq c |w_j - y|^{-\alpha} |x - y|^\alpha. \end{aligned}$$

In particular,

$$\|f_j\|_{\mathbb{B}_{p,\infty}^\varphi(\mathcal{M})} \leq c \left\{ \int_{\mathcal{M}} \min\left(\delta_j^{-\alpha p}, |x - w_j|^{-\alpha p}\right) dx \right\}^{1/p} \leq cN^{\alpha/d-1/p}.$$

Moreover, since  $f(x)$  has mean zero and it is supported in  $\mathcal{X}_j$ ,

$$\left\{ \int_{\mathbf{X}} \left| \sum_{k=1}^N \omega_k f_j(x_k) - \int_{\mathcal{M}} f_j(x) dx \right|^p dx \right\}^{1/p} = \omega_j \left\{ \int_{\mathcal{X}_j} |f_j(x_j)|^p \frac{dx_j}{\omega_j} \right\}^{1/p} \geq cN^{-1}.$$

Finally,

$$N^{-1} = N^{1/p-1} N^{-\alpha/d} N^{\alpha/d-1/p} \geq c \sup \left\{ \omega_j^{1-1/p} \right\} \varphi(2 \sup \{\delta_j\}) \|f_j\|_{\mathbb{B}_{p,\infty}^\varphi(\mathcal{M})}.$$

This shows that Theorem 7.1 with  $1 \leq p \leq 2$  is sharp.

In order to show that the theorem is essentially sharp also for  $2 < p \leq +\infty$ , let  $f_j(x)$  as before and define

$$f(x) = \sum_{j=1}^N f_j(x).$$

If  $\varphi(t) = t^\alpha$ , then a  $\varphi$ -gradient of  $f(x)$  is given by

$$g(x) = cN^{\alpha/d}.$$

Indeed, if  $x, y \in \mathcal{X}_j$ , then

$$|f(x) - f(y)| = |f_j(x) - f_j(y)| \leq B\delta_j^{-1} |x - y| \leq cN^{\alpha/d} |x - y|^\alpha.$$

If  $x \in \mathcal{X}_i$  and  $y \in \mathcal{X}_j$ , with  $i \neq j$  and  $|x - y| \leq N^{-1/d}$ , then

$$|f(x) - f(y)| = |f_i(x) - f_j(y)| \leq |f_i(x) - f_i(y)| + |f_j(y) - f_j(x)| \leq cN^{\alpha/d} |x - y|^\alpha.$$

If  $|x - y| \geq N^{-1/d}$ , then

$$|f(x) - f(y)| \leq c \leq cN^{\alpha/d} |x - y|^\alpha.$$

This gives

$$\|f\|_{\dot{\mathbb{B}}_{p,\infty}^{\varphi}(\mathcal{M})} \leq cN^{\alpha/d}.$$

Moreover, the Marcinkiewicz-Zygmund inequality gives for  $2 < p < +\infty$ ,

$$\begin{aligned} & \left\{ \int_{\mathbf{X}} \left| \sum_{j=1}^N \omega_j f(x_j) - \int_{\mathcal{M}} f(x) dx \right|^p d\mathbf{x} \right\}^{1/p} \geq A(p) \left\{ \int_{\mathbf{X}} \left( \sum_{j=1}^N |\omega_j f_j(x_j)|^2 \right)^{p/2} d\mathbf{x} \right\}^{1/p} \\ & \geq A(p) \left\{ \int_{\mathbf{X}} \left( \sum_{j=1}^N |\omega_j f_j(x_j)|^2 \right) d\mathbf{x} \right\}^{1/2} \geq A(p) \left\{ \sum_{j=1}^N \omega_j \int_{\mathcal{X}_j} |f_j(x_j)|^2 dx_j \right\}^{1/2} \\ & \geq A(p) \min \left\{ \omega_j^{1/2} \right\} \left\{ \int_{\mathcal{M}} |f(x)|^2 dx \right\}^{1/2} \geq cN^{-1/2}. \end{aligned}$$

Finally,

$$N^{-1/2} = N^{-1/2} N^{-\alpha/d} N^{\alpha/d} \geq c \sup \left\{ \omega_j^{1/2} \right\} \varphi(2 \sup \{ \delta_j \}) \|f\|_{\dot{\mathbb{B}}_{p,\infty}^{\varphi}(\mathbb{T}^d)}.$$

In particular, this estimate shows that Theorem 7.1 with  $2 < p < +\infty$  is sharp.

When  $p = +\infty$  let  $f(x)$  and  $\{y_j\}$  as before, so that  $f(y_j) = 1$ . Then

$$\left| \sum_{j=1}^N \omega_j f(y_j) - \int_{\mathcal{M}} f(x) dx \right| = 1 \geq c\varphi(2 \sup \{ \delta_j \}) \|f\|_{\dot{\mathbb{B}}_{p,\infty}^{\varphi}(\mathcal{M})}.$$

In particular, this estimate shows that Theorem 7.1 with  $p = +\infty$  is sharp.

## 8 Discrepancy

The following result on the expected value of the  $p$ th power of the discrepancy of a random set of points with respect to a fixed given set in a measure space extends a result in [16, Lemma 5]. The result in the latter paper concerns the case of a compact convex set in the  $d$ -dimensional unit cube, and the proof is based on the combinatorial argument described in the Introduction, which in our case is replaced by the Marcinkiewicz-Zygmund inequality.

**Theorem 8.1** *Assume that a metric measure space  $\mathcal{M}$  is decomposed into a finite disjoint union of sets in the form  $\mathcal{M} = \mathcal{X}_1 \cup \dots \cup \mathcal{X}_N$ , and call  $\omega_j = |\mathcal{X}_j|$  and  $\delta_j = \text{diam}(\mathcal{X}_j)$ . Let  $\mathcal{B}$  be a measurable subset of  $\mathcal{M}$ , and let*

$$\psi_{\mathcal{B}}(t) = |\{x \in \mathcal{B} : \text{dist}\{x, \mathcal{M} \setminus \mathcal{B}\} \leq t\}| + |\{x \in \mathcal{M} \setminus \mathcal{B} : \text{dist}\{x, \mathcal{B}\} \leq t\}|.$$

*If  $\mathcal{J}$  is the set of indices  $j$  such that  $\mathcal{X}_j$  intersects both  $\mathcal{B}$  and its complement, then the following hold:*

(i) *For every choice of points  $\{x_j\}$  in  $\{\mathcal{X}_j\}$ ,*

$$\left| \sum_{j=1}^N \omega_j \chi_{\mathcal{B}}(x_j) - |\mathcal{B}| \right| \leq \psi_{\mathcal{B}} \left( \sup_{j \in \mathcal{J}} \{ \delta_j \} \right).$$

(ii) For every  $1 \leq p < +\infty$ ,

$$\left\{ \int_{\mathbf{X}} \left| \sum_{j=1}^N \omega_j \chi_{\mathcal{B}}(x_j) - |\mathcal{B}| \right|^p dx \right\}^{1/p} \leq B(p) \sqrt{\sup_{j \in \mathcal{J}} \{\omega_j\} \psi_{\mathcal{B}} \left( \sup_{j \in \mathcal{J}} \{\delta_j\} \right)}.$$

Observe that the right hand side in (ii) is better than the one in (i) when

$$B(p) \leq \sqrt{\frac{\psi \left( \sup_{j \in \mathcal{J}} \{\delta_j\} \right)}{\sup_{j \in \mathcal{J}} \{\omega_j\}}}.$$

Also observe that

$$\sup_{j \in \mathcal{J}} \{\omega_j\} \leq \sum_{j \in \mathcal{J}} \omega_j \leq \psi_{\mathcal{B}} \left( \sup_{j \in \mathcal{J}} \{\delta_j\} \right).$$

For a sufficiently refined decomposition  $\{\mathcal{X}_j\}$  of the space, one should expect  $\sup_{j \in \mathcal{J}} \{\omega_j\}$  to be much smaller than  $\psi_{\mathcal{B}} \left( \sup_{j \in \mathcal{J}} \{\delta_j\} \right)$ . Hence, for a fixed value of  $p$ , estimate (ii) is in general better than (i) as  $N \rightarrow +\infty$ . On the other hand, recall that  $B(p) \rightarrow +\infty$  as  $p \rightarrow +\infty$ , hence, for a fixed decomposition of  $\mathcal{M}$ , estimate (i) wins for sufficiently large values of  $p$ .

*Proof* The proof of (i) is elementary. For every choice of  $x_j \in \mathcal{X}_j$  one has

$$\sum_{j=1}^N \omega_j \chi_{\mathcal{B}}(x_j) - |\mathcal{B}| = \sum_{j=1}^N (\omega_j \chi_{\mathcal{B} \cap \mathcal{X}_j}(x_j) - |\mathcal{B} \cap \mathcal{X}_j|).$$

If  $\mathcal{X}_j \subseteq \mathcal{B}$  or if  $\mathcal{B} \cap \mathcal{X}_j = \emptyset$  then  $\omega_j \chi_{\mathcal{B} \cap \mathcal{X}_j}(x_j) - |\mathcal{B} \cap \mathcal{X}_j| = 0$ . Moreover, for every  $j$ ,

$$|\omega_j \chi_{\mathcal{B} \cap \mathcal{X}_j}(x_j) - |\mathcal{B} \cap \mathcal{X}_j|| \leq \omega_j.$$

Then, by the triangle inequality,

$$\left| \sum_{j=1}^N (\omega_j \chi_{\mathcal{B} \cap \mathcal{X}_j}(x_j) - |\mathcal{B} \cap \mathcal{X}_j|) \right| \leq \sum_{j \in \mathcal{J}} \omega_j \leq \psi_{\mathcal{B}} \left( \sup_{j \in \mathcal{J}} \{\delta_j\} \right).$$

The proof of (ii) is similar, with the crucial difference that we replace the triangle inequality with the Marcinkiewicz-Zygmund inequality (Corollary 4.2),

$$\begin{aligned} & \left\{ \int_{\mathbf{X}} \left| \sum_{j=1}^N (\omega_j \chi_{\mathcal{B} \cap \mathcal{X}_j}(x_j) - |\mathcal{B} \cap \mathcal{X}_j|) \right|^p dx \right\}^{1/p} \\ & \leq B(p) \left\{ \sum_{j \in \mathcal{J}} \omega_j^2 \right\}^{1/2} \leq B(p) \sup_{j \in \mathcal{J}} \{\omega_j^{1/2}\} \left\{ \sum_{j \in \mathcal{J}} \omega_j \right\}^{1/2} \\ & \leq B(p) \sqrt{\sup_{j \in \mathcal{J}} \{\omega_j\} \psi_{\mathcal{B}} \left( \sup_{j \in \mathcal{J}} \{\delta_j\} \right)}. \end{aligned}$$

□

Theorems like the above are the main building block for the proof of the existence of point distributions with small  $L^p$  discrepancy with respect to given collections of subsets. A very general result of this type is the following.

**Corollary 8.2** *Assume that a metric measure space  $\mathcal{M}$  can be decomposed into a finite disjoint union of sets in the form  $\mathcal{M} = \mathcal{X}_1 \cup \dots \cup \mathcal{X}_N$ , with  $\omega_j = |\mathcal{X}_j| \approx N^{-1}$  and  $\delta_j = \text{diam}(\mathcal{X}_j) \approx N^{-1/d}$ . Let  $\mathbb{G}$  be a collection of measurable subsets of  $\mathcal{M}$  with the property that there exist positive constants  $c$  and  $\beta$  such that for all sets  $\mathcal{G} \in \mathbb{G}$*

$$\psi_{\mathcal{G}}(t) = |\{x \in \mathcal{G} : \text{dist}\{x, \mathcal{M} \setminus \mathcal{G}\} \leq t\}| + |\{x \in \mathcal{M} \setminus \mathcal{G} : \text{dist}\{x, \mathcal{G}\} \leq t\}| \leq ct^\beta.$$

*Then for any finite positive measure  $\mu$  on any sigma algebra on  $\mathbb{G}$ , and for every  $1 \leq p < +\infty$  there exists a constant  $C$  and a choice of points  $\{x_j\}$  in  $\{\mathcal{X}_j\}$  such that*

$$\left( \int_{\mathbb{G}} \left| \sum_{j=1}^N \omega_j \chi_{\mathcal{G}}(x_j) - |\mathcal{G}| \right|^p d\mu(\mathcal{G}) \right)^{1/p} \leq CN^{-1/2-\beta/2d}.$$

*Proof* By point (ii) of Theorem 8.1,

$$\begin{aligned} & \int_{\mathbb{G}} \int_{\mathbf{X}} \left| \sum_{j=1}^N \omega_j \chi_{\mathcal{G}}(x_j) - |\mathcal{G}| \right|^p dx d\mu(\mathcal{G}) \\ & \leq B(p)^p \int_{\mathbb{G}} \left( \sup_{j \in \mathcal{J}} \{\omega_j\} \psi_{\mathcal{G}} \left( \sup_{j \in \mathcal{J}} \{\delta_j\} \right) \right)^{p/2} d\mu(\mathcal{G}) \\ & \leq B(p)^p \mu(\mathbb{G}) \left( \sup_{j \in \mathcal{J}} \{\omega_j\} c \left( \sup_{j \in \mathcal{J}} \{\delta_j\} \right)^\beta \right)^{p/2}. \end{aligned}$$

This implies that there exists an  $\mathbf{x} \in \mathbf{X}$  such that the thesis of the theorem holds.  $\square$

We emphasize that under the hypotheses of Corollary 5.1 the required decomposition exists, and it is always possible to take all  $\omega_j$  equal to  $|\mathcal{M}|N^{-1}$ . The corollary has several possible applications. We now examine a few particular cases, starting with the isotropic discrepancy (the discrepancy with respect to convex sets) in the unit cube  $[0, 1]^d$ .

**Corollary 8.3** *Let  $1 \leq p < +\infty$ , and let  $\mu$  be a finite positive measure on a sigma algebra on the collection  $\mathcal{K}_u^d$  of all convex sets of the unit cube  $[0, 1]^d$ . For any integer  $N$  there exists a distribution of points  $\{x_j\}_{j=1}^N$  in  $[0, 1]^d$  such that*

$$\left( \int_{\mathcal{K}_u^d} \left| \frac{1}{N} \sum_{j=1}^N \chi_K(x_j) - |K| \right|^p d\mu(K) \right)^{1/p} \leq CN^{-1/2-1/2d}.$$

*Proof* It suffices to show that Corollary 8.2 applies with  $\beta = 1$ . First of all, the unit cube can be decomposed into a finite disjoint union of sets in the form  $\mathcal{M} = \mathcal{X}_1 \cup \dots \cup \mathcal{X}_N$ , with  $\omega_j = |\mathcal{X}_j| = N^{-1}$  and  $\delta_j = \text{diam}(\mathcal{X}_j) \approx N^{-1/d}$ . See, for example,

Theorem 5.1. It then suffices to observe that for all convex sets  $K$  the uniform estimate  $\psi_K(t) \leq 4dt$  holds. Indeed, by the coarea formula

$$\psi_K(t) = |\{x \in [0, 1]^d : \text{dist}(x, \partial K) \leq t\}| = \int_{-t}^t |\partial K_u|_{d-1} du,$$

where for  $u > 0$  we define  $K_u = \{x \in K : \text{dist}(x, \partial K) \geq u\}$  and for  $u < 0$  we define  $K_u = \{x \in [0, 1]^d : \text{dist}(x, K) \leq |u|\}$ , and  $|\cdot|_{d-1}$  is the  $(d-1)$ -dimensional Hausdorff measure. It is a well known property of convex sets that all sets  $K_u$  are convex (see e.g. [44, Chapter 3]). By the Archimedean property of monotonicity with respect to inclusion of the measure of the boundary of convex sets (see [6, Property 5, page 52]), from  $K_u \subset [0, 1]^d$  it follows  $|\partial K_u|_{d-1} \leq |\partial[0, 1]^d|_{d-1} = 2d$ .  $\square$

As an explicit example of measure, one could consider any finite measure supported on the translated, rotated and dilated copies of a fixed convex set. Hence, this result includes and extends well known results on the  $L^p$  discrepancy with respect to discs, or other collections of sets with “reasonable” shapes (see for example [16, Theorem 2D]). It is interesting to observe that if one replaces the above  $L^p$  norm with a supremum in  $K \in \mathcal{K}_u^d$ , then the above result fails. Indeed, Schmidt (see [43]) proved that for any  $N$  point distribution in the unit cube there exists a convex set with discrepancy of order  $N^{-2/(d+1)}$ .

It is perhaps worth mentioning some results about measures on the space of convex sets. It is well known that the set of nonempty convex compact subsets of  $\mathbb{R}^d$ , let us call it  $\mathcal{K}^d$ , can be made into a metric space by introducing the Hausdorff distance

$$d_H(A, B) = \max\{\sup_{a \in A} \inf_{b \in B} |a - b|, \sup_{b \in B} \inf_{a \in A} |a - b|\}.$$

A large class of sigma finite Borel measures on  $\mathcal{K}^d$ , which are positive on open sets of  $\mathcal{K}^d$  and are invariant under rigid motions, has been recently constructed by L. M. Hoffmann ([30]). Just to give a rough idea, let  $\{K_n\}_{n=1}^{+\infty}$  be an enumeration of all polytopes of  $\mathbb{R}^d$  with vertices in rational points, and let  $\sum_{n=1}^{+\infty} \alpha_n < +\infty$  be a convergent series with positive terms. Then one can define the measure

$$\mu = \sum_{n=1}^{+\infty} \alpha_n \delta_{K_n},$$

where  $\delta_{K_n}$  is the Dirac delta centered at  $K_n$ . This measure is supported on rational polytopes, but is positive on open sets since these rational polytopes are dense. A suitable clever modification can be made invariant under rigid motions. Nevertheless, it can be shown that there are more isometries of the space  $\mathcal{K}^d$  than those coming from rigid motions of  $\mathbb{R}^d$ , and in fact it has been showed by Bandt and Baraki (see [2]) that for  $d > 1$  there are no nontrivial sigma finite measures on  $\mathcal{K}^d$ , that are invariant with respect to all isometries of the whole space  $\mathcal{K}^d$ . Hence it seems that there is no “natural” measure on  $\mathcal{K}^d$ .

Similar results hold in compact Riemannian manifolds. For the sake of simplicity, we state here a result on the  $L^p$  discrepancy associated to geodesic balls.



**Corollary 8.4** *Let  $\mathcal{M}$  be a compact Riemannian manifold with injectivity radius  $r_0$ , and let  $0 < r_1 < r_0$ . Denote by  $B(x, r)$  the geodesic ball centered at the point  $x \in \mathcal{M}$  with radius  $r$ . Then for any  $1 \leq p < +\infty$ , for any finite positive measure  $\mu$  on  $\mathcal{M} \times (0, r_1)$ , and for any integer  $N$ , there exists a distribution of points  $\{x_j\}_{j=1}^N$  in  $\mathcal{M}$  such that*

$$\left( \iint_{\mathcal{M} \times (0, r_1)} \left| \frac{|\mathcal{M}|}{N} \sum_{j=1}^N \chi_{B(x, r)}(x_j) - |B(x, r)| \right|^p d\mu(x, r) \right)^{1/p} \leq CN^{-1/2-1/2d}.$$

*Proof* As before, it suffices to show that Corollary 8.2 applies with  $\beta = 1$ . Indeed,  $\mathcal{M}$  can be decomposed into a finite disjoint union of sets in the form  $\mathcal{M} = \mathcal{X}_1 \cup \dots \cup \mathcal{X}_N$ , with  $\omega_j = |\mathcal{X}_j| = |\mathcal{M}|N^{-1}$  and  $\delta_j = \text{diam}(\mathcal{X}_j) \approx N^{-1/d}$ . See, for example, Theorem 5.1. Moreover, there exists a positive constant  $c$  such that for every geodesic ball  $B(x, r)$  with  $r < r_1$  and every  $t > 0$  one has  $\psi_{B(x, r)}(t) \leq ct$ . It clearly suffices to prove this for all  $0 < t \leq (r_0 - r_1)/2$ . Indeed, by the triangle inequality,

$$\begin{aligned} \{y \in B(x, r) : \text{dist}\{y, \mathcal{M} \setminus B(x, r)\} \leq t\} &\subset B(x, r) \setminus B(x, r-t), \\ \{y \in \mathcal{M} \setminus B(x, r) : \text{dist}\{y, B(x, r)\} \leq t\} &\subset \overline{B}(x, r+t) \setminus B(x, r). \end{aligned}$$

If  $s \leq 0$  then we set  $B(x, s) = \emptyset$ . Finally, if  $r < r_1$  and  $t < (r_0 - r_1)/2$ , the exponential map diffeomorphically maps the annulus  $\overline{B}(x, r+t) \setminus B(x, r-t)$  into the tangent space in  $x$ , and by a uniform bound on the Jacobian of the exponential map, its measure is bounded above by  $c((r+t)^d - \max\{0, r-t\}^d) \leq ct$ .  $\square$

We have already mentioned that the above results fail in general when one replaces the  $L^p$  norm with a supremum. Nevertheless, an extra hypothesis concerning the complexity of the collection of sets  $\mathbb{G}$  allows to obtain the same upper bound in the supremum case too, up to a logarithmic transgression.

**Theorem 8.5** *Let  $\mathcal{M}$  be a metric measure space with finite measure with the property that there exist positive constants  $d$  and  $c_1$ , such that for every sufficiently large  $N$  there exists a partition  $\mathcal{M} = \mathcal{X}_1 \cup \dots \cup \mathcal{X}_N$  with  $|\mathcal{X}_j| = |\mathcal{M}|N^{-1}$  and  $\text{diam}(\mathcal{X}_j) \leq c_1N^{-1/d}$ . Let  $\mathbb{G}$  be a collection of measurable subsets of  $\mathcal{M}$  with the following two properties:*

(i) *There exist positive constants  $c_2$  and  $\beta$  such that for all sets  $\mathcal{G} \in \mathbb{G}$*

$$\psi_{\mathcal{G}}(t) = |\{x \in \mathcal{G} : \text{dist}\{x, \mathcal{M} \setminus \mathcal{G}\} \leq t\}| + |\{x \in \mathcal{M} \setminus \mathcal{G} : \text{dist}\{x, \mathcal{G}\} \leq t\}| \leq c_2t^\beta.$$

(ii) *There exist positive constants  $c_3$  and  $\gamma$  such that for all integers  $N$  and for all distributions  $\mathcal{P}$  of  $N$  points in  $\mathcal{M}$  there are at most  $c_3N^\gamma$  equivalence classes in  $\mathbb{G}$ , where  $\mathcal{G}, \mathcal{G}'$  in  $\mathbb{G}$  are defined to be equivalent if  $\mathcal{P} \cap \mathcal{G} = \mathcal{P} \cap \mathcal{G}'$ .*

*Then for any integer  $N$  there exists a distribution of points  $\{z_j\}_{j=1}^N$  such that*

$$\sup_{\mathcal{G} \in \mathbb{G}} \left| \frac{|\mathcal{M}|}{N} \sum_{j=1}^N \chi_{\mathcal{G}}(z_j) - |\mathcal{G}| \right| \leq CN^{-1/2-\beta/2d} \sqrt{\log N}.$$

A few words on the above condition (ii) are perhaps necessary. We say that  $\mathbb{G}$  shatters a finite subset  $\mathcal{P}$  of  $N$  points of  $\mathcal{M}$  if there are exactly  $2^N$  distinct intersections of sets of  $\mathbb{G}$  with  $\mathcal{P}$ . The Vapnik Chervonenkis dimension, or VC-dimension, of  $\mathbb{G}$  is the supremum of the sizes of all finite subsets of  $\mathcal{M}$  that are shattered by  $\mathbb{G}$ . By Sauer lemma (see [42]), condition (ii) coincides with asking that the collection  $\mathbb{G}$  has finite VC-dimension. For example, the collection of convex sets in  $\mathbb{R}^d$  has infinite VC-dimension. Indeed a set of  $N$  points on a circle can be easily shattered with convex sets. On the other hand, the collection of balls in  $\mathbb{R}^d$  has VC-dimension  $d + 1$  (see [38, Chapter 5] for an account on this subject).

The proof that we present here follows closely the lines of the classic result for discs in the unit square, as one can find in J. Matoušek's book [38].

*Proof* Let  $M = N^q$ , where  $q$  is a positive integer that will be fixed later. Consider two partitions of  $\mathcal{M}$  as in the hypothesis. The first is composed by the  $N$  sets  $\mathcal{X}_1, \dots, \mathcal{X}_N$ , and the second is composed by the  $M$  sets  $\mathcal{Y}_1, \dots, \mathcal{Y}_M$ . For any  $j = 1, \dots, N$ , define  $I_j = \{i = 1, \dots, M : \mathcal{Y}_i \cap \mathcal{X}_j \neq \emptyset\}$  and, for all  $i \in I_j$  define

$$\mathcal{Y}_{j,i} = \mathcal{X}_j \cap \mathcal{Y}_i.$$

Fix a point  $y_{j,i}$  in any of the sets  $\mathcal{Y}_{j,i}$ . Clearly  $\{y_{j,i}\}_{i \in I_j}$  forms a partition of  $\mathcal{X}_j$  and  $\mathcal{Y}_{j,i} \subset \{x \in \mathcal{M} : |x - y_{j,i}| < c_1 M^{-1/d}\}$ .

For each  $j = 1, \dots, N$ , let us pick one point  $q_j$  among all the points  $y_{j,i}$  with  $i \in I_j$ . This point  $q_j$  is chosen randomly with probability  $\mathbb{P}[q_j = y_{j,i}] = N|\mathcal{Y}_{j,i}|/|\mathcal{M}|$ , the choices being independent for distinct values of  $j$ . The discrepancy of the point distribution  $\{q_j\}_{j=1}^N$  with respect to a given  $\mathcal{G} \in \mathbb{G}$  can be estimated as follows

$$\begin{aligned} & \left| \frac{|\mathcal{M}|}{N} \sum_{j=1}^N \chi_{\mathcal{G}}(q_j) - |\mathcal{G}| \right| \\ & \leq \left| \frac{|\mathcal{M}|}{N} \sum_{j=1}^N \left( \chi_{\mathcal{G}}(q_j) - \frac{N}{|\mathcal{M}|} \sum_{i \in I_j} |\mathcal{Y}_{j,i}| \chi_{\mathcal{G}}(y_{j,i}) \right) \right| + \left| \sum_{j=1}^N \sum_{i \in I_j} |\mathcal{Y}_{j,i}| \chi_{\mathcal{G}}(y_{j,i}) - |\mathcal{G}| \right| \end{aligned}$$

The second term in the above sum is deterministic and can be treated easily. By Theorem 8.1 (i) it is bounded above by

$$\psi_{\mathcal{G}} \left( c_1 M^{-1/d} \right) \leq c_2 \left( c_1 M^{-1/d} \right)^{\beta} = c_1^{\beta} c_2 N^{-\beta q/d}.$$

It is therefore sufficient to take  $q \geq d/\beta$  to obtain an estimate better than what is needed.

The other term is of a probabilistic nature. We only need to consider the values of  $j$  for which  $\mathcal{X}_j$  intersects both  $\mathcal{G}$  and  $\mathcal{M}$ . Call this set  $\mathcal{J}$  and its cardinality  $m$ . Since

$$m \frac{|\mathcal{M}|}{N} \leq \psi_{\mathcal{G}} \left( c_1 N^{-1/d} \right) \leq c_1^{\beta} c_2 N^{-\beta/d},$$

we have  $m \leq c_1^{\beta} c_2 N^{1-\beta/d} / |\mathcal{M}|$ . Let us now set

$$k_j = \frac{N}{|\mathcal{M}|} \sum_{i \in I_j} |\mathcal{Y}_{j,i}| \chi_{\mathcal{G}}(y_{j,i}),$$

and call  $F_j$  the random variable  $\chi_{\mathcal{G}}(q_j) - k_j$ . Thus we have

$$\frac{|\mathcal{M}|}{N} \sum_{j=1}^N \left( \chi_{\mathcal{G}}(q_j) - \frac{N}{|\mathcal{M}|} \sum_{i \in I_j} |\mathcal{Y}_{j,i}| \chi_{\mathcal{G}}(y_{j,i}) \right) = \frac{|\mathcal{M}|}{N} \left( \sum_{j \in \mathcal{J}} F_j \right).$$

The variables  $F_j$  are mutually independent, and  $F_j$  takes the value  $1 - k_j$  with probability  $k_j$ , and  $-k_j$  with probability  $1 - k_j$ . Therefore for every  $\Delta > 0$  we have

$$\mathbb{P} \left[ \left| \sum_{j \in \mathcal{J}} F_j \right| \geq \Delta \right] \leq 2 \exp(-2\Delta^2/m)$$

(see, for example, [1, Corollary A.1.7]). Let us fix a constant  $C > 0$ . Then we have showed that

$$\begin{aligned} & \mathbb{P} \left[ \frac{|\mathcal{M}|}{N} \left| \sum_{j \in \mathcal{J}} F_j \right| \geq CN^{-1/2-\beta/(2d)} \sqrt{\log N} \right] \\ &= \mathbb{P} \left[ \left| \sum_{j \in \mathcal{J}} F_j \right| \geq CN^{1/2-\beta/(2d)} \sqrt{\log N} / |\mathcal{M}| \right] \\ &\leq 2 \exp(-2C^2 N^{1-\beta/d} \log N / |\mathcal{M}|^2 / m) \\ &\leq 2N^{-C^2 c_1^{-\beta} c_2^{-1} |\mathcal{M}|^{-1}}. \end{aligned}$$

Finally, if  $\mathbb{F} \subset \mathbb{G}$  contains one representative for each equivalence class, then

$$\begin{aligned} & \mathbb{P} \left[ \frac{|\mathcal{M}|}{N} \left| \sum_{j \in \mathcal{J}} F_j \right| \geq CN^{-1/2-\beta/(2d)} \sqrt{\log N} \text{ for some } \mathcal{G} \in \mathbb{G} \right] \\ &= \mathbb{P} \left[ \frac{|\mathcal{M}|}{N} \left| \sum_{j \in \mathcal{J}} F_j \right| \geq CN^{-1/2-\beta/(2d)} \sqrt{\log N} \text{ for some } \mathcal{G} \in \mathbb{F} \right] \\ &\leq \sum_{\mathcal{G} \in \mathbb{F}} \mathbb{P} \left[ \frac{|\mathcal{M}|}{N} \left| \sum_{j \in \mathcal{J}} F_j \right| \geq CN^{-1/2-\beta/(2d)} \sqrt{\log N} \right] \\ &\leq 2c_3 N^{\gamma-C^2 c_1^{-\beta} c_2^{-1} |\mathcal{M}|^{-1}} < 1 \end{aligned}$$

if  $C$  is large enough. The theorem follows.  $\square$

The next Corollary shows one possible application of the above theorem.

**Corollary 8.6** *Let  $\mathcal{M}$  be a  $d$ -dimensional compact Riemannian manifold isometrically embedded in  $\mathbb{R}^D$ , and call  $B^D(x, r) = \{y \in \mathbb{R}^D : \|y - x\| < r\}$  the Euclidean  $D$ -dimensional ball of center  $x$  and radius  $r$ . Then there exist positive constants  $r_0$  and  $C$  such that for any integer  $N$  there exists a distribution of points  $\{z_j\}_{j=1}^N \in \mathcal{M}$  such that*

$$\sup_{r \leq r_0, x \in \mathcal{M}} \left| \frac{|\mathcal{M}|}{N} \sum_{j=1}^N \chi_{B^D(x,r) \cap \mathcal{M}}(z_j) - |B^D(x,r) \cap \mathcal{M}| \right| \leq CN^{-1/2-1/2d} \sqrt{\log N}.$$

Notice that, by the Nash embedding theorem, every Riemannian manifold can be isometrically embedded into some Euclidean space. In particular, if one takes as  $\mathcal{M}$  the  $d$ -dimensional unit sphere in  $\mathbb{R}^{d+1}$ , then the sets  $B^{d+1}(x, r) \cap \mathcal{M}$  of the above corollary coincide with the usual spherical caps, and one recovers Beck's estimate for the spherical cap discrepancy (see [5, Theorem 24D]).

*Proof* It is enough to show that the collection of subsets of the form  $B^D(x, r) \cap \mathcal{M}$  satisfies the two hypotheses of Theorem 8.5. By compactness, there exists a positive number  $r_0$  such that for all  $0 < r \leq r_0$  and for all  $x \in \mathcal{M}$ , the set  $\mathcal{N} = \{y \in \mathcal{M} : \|x - y\| = r\}$  is a hypersurface of  $\mathcal{M}$  with uniformly bounded  $(d - 1)$ -dimensional volume. Furthermore, the measure of the set of points of  $\mathcal{M}$  with geodesic distance from  $\mathcal{N}$  less than or equal to  $t$  is bounded above by

$$\int_{\mathcal{N}} \int_{-t}^t |f(s, n)| ds dn,$$

where  $dn$  is the  $(d - 1)$ -dimensional volume form on  $\mathcal{N}$  and  $f(s, n)$  is the (uniformly bounded) Jacobian of the exponential map of the normal bundle on  $\mathcal{N}$  in  $\mathcal{M}$  (see [22] for the details). Thus

$$\psi_{\mathcal{M} \cap B^D(x, r)}(t) \leq ct,$$

and the first hypothesis of Theorem 8.5 holds with  $\beta = 1$ . Finally, as we mention before, balls in  $\mathbb{R}^D$  satisfy the second hypotheses of the same theorem with  $\gamma = D + 1$  ([38, Chapter 5]).  $\square$

**Acknowledgements** The authors wish to thank Dmitriy Bilyk for several conversations concerning the results on discrepancy contained in this paper.

## References

1. **N. Alon, J.H. Spencer** The probabilistic method. Third edition. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley and Sons, Inc., Hoboken, NJ, 2008.
2. **C. Bandt, G. Baraki**, *Metrically invariant measures on locally homogeneous spaces and hyperspaces*, Pacific J. Math. 121 (1986), 13–28.
3. **J. Bergh, J. Löfström**, *Interpolation spaces*, Springer Verlag 1976.
4. **J. Beck, W.W.L. Chen**, *Note on irregularities of distribution II*, Proc. London Math. Soc. 61 (1990), 251–272.
5. **J. Beck, W.W.L. Chen**, *Irregularities of Distribution*. Cambridge University Press, Cambridge (1987).
6. **T. Bonnesen, W. Fenchel**, *Theory of convex bodies*. Translated from the German and edited by L. Boron, C. Christenson and B. Smith. BCS Associates (1987).
7. **L. Brandolini, L. Colzani**, *Decay of Fourier transforms and summability of eigenfunction expansions*, Ann. Sc. Norm. Super. Pisa Cl. Sci., 29 (2000), 611–638.
8. **L. Brandolini, C. Choirat, L. Colzani, G. Gigante, R. Seri, G. Travaglino**, *Quadrature rules and distribution of points on manifolds*, Ann. Sc. Norm. Super. Pisa Cl. Sci., 13 (2014), 889–923.
9. **L. Brandolini, W.W.L. Chen, G. Gigante, G. Travaglino**, *Discrepancy for randomized Riemann sums*, Proc. Amer. Math. Soc. 137 (2009), 3187–3196.
10. **L. Brandolini, L. Colzani, G. Gigante, G. Travaglino**, *On the Koksma–Hlawka inequality*, J. Complexity 29 (2013), 158–172.
11. **J.S. Brauchart, J. Dick**, *A characterization of Sobolev spaces on the sphere and an extension of Stolarsky's invariance principle to arbitrary smoothness*, Constr. Approx. 38 (2013), 397–445.

12. **J.S. Brauchart, J. Dick, E.B. Saff, I.H. Sloan, Y.G. Wang, R.S. Womersley**, *Covering of spheres by spherical caps and worst-case error for equal weight cubature in Sobolev spaces*, *J. Math. Anal. Appl.*, 431 (2015), 782–811.
13. **J.S. Brauchart, K. Hesse**, *Numerical integration over spheres of arbitrary dimension*, *Constr. Approx.* 25 (2007), 41–71.
14. **J.S. Brauchart, E.B. Saff, I.H. Sloan, R.S. Womersley**, *QMC design: optimal order quasi Monte Carlo integration schemes on the sphere*, *Math. Comp.* 83 (2014), 2821–2851.
15. **D. Burkholder**, *Sharp inequalities for martingales and stochastic integrals*, *Astérisque* 157-158 (1988), 75–94.
16. **W.W.L. Chen**, *On irregularities of distribution. III*, *J. Austral. Math. Soc.* 60 (1996), 228–244.
17. **W.W.L. Chen, G. Travaglini**, *Deterministic and probabilistic discrepancies*, *Ark. Mat.* 47 (2009), 273–293.
18. **Y.S. Chow, H. Teicher**, *Probability theory. Independence, interchangeability, martingales*, Springer Verlag, 1988.
19. **F. Filbir, H.N. Mhaskar**, *A quadrature formula for diffusion polynomials corresponding to a generalized heat kernel*, *J. Fourier Anal. Appl.* 16 (2010), 629–657.
20. **M. Frazier, B. Jawerth, G. Weiss**, *Littlewood-Paley theory and the study of function spaces*. CBMS Regional Conference Series in Mathematics, 79. Published for the Conference Board of the Mathematical Sciences, Washington, DC, by the American Mathematical Society, Providence, RI, (1991).
21. **G. Gigante, P. Leopardi**, *Diameter bounded equal measure partitions of Ahlfors regular metric measure spaces*, *Discrete Comput. Geom.* 57 (2017), 419–430.
22. **E. Heintze, H. Karcher**, *A general comparison theorem with applications to volume estimates for submanifolds*, *Ann. Sci. Ecole Norm. Sup. (4)* 11 (1978), 451–470.
23. **U. Haagerup**, *The best constant in the Kintchine inequality*, *Studia Math.* 70 (1981), 231–283.
24. **P. Hajłasz**, *Sobolev spaces on an arbitrary metric space*, *Potential Anal.* 5 (1996), 403–415.
25. **K. Hesse**, *A lower bound for the worst-case cubature error on spheres of arbitrary dimension*, *Numer. Math.* 103 (2006), 413–433.
26. **K. Hesse, H.M. Mhaskar, I.H. Sloan**, *Quadrature in Besov spaces on the Euclidean sphere*, *J. Complexity* 23 (2007), 528–552.
27. **K. Hesse, I.H. Sloan**, *Worst-case errors in a Sobolev space setting for cubature over the sphere  $S^2$* , *Bull. Austral. Math. Soc.* 71 (2005), 81–105.
28. **K. Hesse, I.H. Sloan**, *Optimal lower bounds for cubature error on the sphere  $S^2$* , *J. Complexity* 21 (2005), 790–803.
29. **K. Hesse, I.H. Sloan**, *Cubature over the sphere  $S^2$  in Sobolev spaces of arbitrary order*, *J. Approx. Theory* 141 (2006), 118–133.
30. **L.M. Hoffmann**, *Measures on the space of convex bodies*, *Adv. Geom.* 10 (2010), 477–486.
31. **J. Hu, M. Zähle**, *Generalized Bessel and Riesz Potentials on Metric Measure Spaces*, *Potential Anal.* 30 (2009), 315–340.
32. **P. Koskela, D. Yang, Y. Zhou**, *Pointwise characterization of Besov and Triebel Lizorkin spaces and quasiconformal mappings*, *Adv. Math.* 226 (2011), 3579–3621.
33. **A. Kushpel**, *Optimal cubature formulas on compact homogeneous manifolds*. *J. Funct. Anal.* 257 (2009), 1621–1629.
34. **P. Leopardi**, *Diameter bounds for equal area partitions of the unit sphere*, *Electron. Trans. Numer. Anal.* 35 (2009), 1–16.
35. **M. Maggioni, H.N. Mhaskar**, *Diffusion polynomial frames on metric measure spaces*, *Appl. Comput. Harmon. Anal.* 24 (2008), 329–353.
36. **J. Marcinkiewicz, A. Zygmund**, *Sur les fonctions indépendantes*, *Fund. Math.* 29 (1937), 60–90.
37. **J. Marcinkiewicz, A. Zygmund**, *Quelques théorèmes sur les fonctions indépendantes*, *Studia Math.* 7 (1938), 104–120.
38. **J. Matoušek** *Geometric discrepancy. An illustrated guide. Revised paperback reprint of the 1999 original. Algorithms and Combinatorics*, 18. Springer-Verlag, Berlin, 2010.
39. **F. Nazarov, S. Treil, A. Volberg**, *Weak type estimates and Cotlar inequalities for Calderón-Zygmund operators on nonhomogeneous spaces*, *Internat. Math. Res. Notices*, 9 (1998), 463–487.

40. **E.A. Rakhmanov, E.B. Saff, Y.M. Zhou**, *Minimal discrete energy on the sphere*, Math. Res. Lett. 1(6) (1994), 647–662.
41. **A. Reznikov, E.B. Saff**, *The covering radius of randomly distributed points on a manifold*, Int. Math. Res. Not. IMRN, 19 (2016), 6065–6094.
42. **N. Sauer**, *On the density of families of sets*, J. Combinatorial Theory Ser. A 13 (1972), 145–147.
43. **W.M. Schmidt**, *Irregularities of distribution IX*, Acta Arith., 27 (1975), 385–396.
44. **R. Schneider**, *Convex Bodies: the Brunn-Minkowski Theory*, Second Expanded Edition (Encyclopedia of Mathematics and its Applications). Cambridge University Press (2014).
45. **M. M. Skrikanov**, *Point distributions in compact metric spaces*, Mathematika, 63, (2017), 1152–1171.
46. **M. M. Skrikanov**, *Point distributions in compact metric spaces, II*, arXiv:1701.04007
47. **M. M. Skrikanov**, *Point distribution in compact metric spaces, III. Two-point homogeneous spaces*, arXiv:1701.04545
48. **E.M. Stein**, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press (1970).
49. **R.S. Strichartz**, *Analysis on fractals*, Not. Am. Math. Soc., 46(10), (1999) 1199–1208.
50. **R.S. Strichartz, M. Usher**, *Splines on fractals*, Math. Proc. Camb. Philos. Soc., 129(2), (2000) 331–360.