SAMPLING THEOREMS FOR INVERSE PROBLEMS ON RIEMANNIAN MANIFOLDS

G. S. ALBERTI[1] ⓘ, E. DE VITO[1] ⓘ, B. GARIBOLDI[2] ⓘ, AND G. GIGANTE[2] ⓘ

ABSTRACT. We consider inverse problems consisting of the reconstruction of an unknown signal f from noisy measurements $y = Ff + \text{noise}$, where Ff is a function on a Riemannian manifold without boundary $\mathcal{M}$. We consider the case when only pointwise samples are available, namely $y_j = (Ff)(x_j) + \eta_j$, where $\{x_j\}_{j=1}^n \subseteq \mathcal{M}$ is a Marcinkiewicz-Zygmund family. We derive sampling theorems providing explicit bounds on the reconstruction error depending on n, the smoothness of f and the properties of F. We study in detail the case when F is a convolution on a compact two-point homogeneous space. As a corollary, we state a sampling theorem for convolutions on the two-dimensional sphere, and discuss four relevant examples related to terrestrial and celestial measurements.

1. INTRODUCTION

1.1. **Sampling and approximation.** The classical Shannon sampling theorem, a cornerstone of signal processing and harmonic analysis, asserts that a bandlimited function $f \in L^2(\mathbb{R})$ with spectral support contained in an interval $[-\Omega, \Omega]$ is entirely determined by its values on the discrete set $\{x_j = j/(2\Omega)\}_{j \in \mathbb{Z}}$, and can be exactly reconstructed via the Shannon-Whittaker interpolation formula [71, 54, 60]. This fundamental result highlights the intrinsic link between the bandwidth Ω — the measure of frequency content — and the density $\frac{1}{2\Omega}$ of the samples required for perfect reconstruction. In higher dimensions, analogous results hold for functions in $L^2(\mathbb{R}^d)$ whose Fourier transforms are supported in compact sets: such functions can be reconstructed from their values on suitable lattices, provided the sampling rate exceeds the Nyquist threshold determined by the volume of the spectral support [42, 9].

On compact domains such as the torus $\mathbb{T}^d$, the sphere S^d, or a compact Riemannian manifold $\mathcal{M}$ without boundary, where the frequency domain is discrete, the situation becomes even more structured: bandlimited functions f become finite-dimensional, and exact recovery from a finite number of measurements $\{f(x_j)\}_j$ is possible, provided the samples $\{x_j\}_j$ are suitably taken [48, 29, 46, 45]. This can be formalized with the concept of Marcinkiewicz-Zygmund family, whose construction is guaranteed whenever the sampling space can be divided into regions, each with a prefixed volume and small diameter [49, 47, 22, 21]. This type of partition, on the other hand, exists in every connected compact orientable Riemannian manifold without boundary [27, 19], and is explicit in simple settings, such as for the two-dimensional sphere

[56]. As with the Shannon theorem, the bandwidth of f is explicitly linked with the density of the sampling points $\{x_j\}_j$. All these results have an "approximation" counterpart in the cases where f is not bandlimited but smooth, an assumption that is quantified by the decay of the generalized Fourier coefficients of f.

1.2. **Inverse problems.** Inverse problems appear when an unknown signal f belonging to a space X has to be recovered from indirect measurements of the form Ff, where $F\colon X \to Y$ is the so-called forward map, and Y is the measurement space [10, 36, 59]. Inverse problems are typically ill-posed: whenever measurements are noisy, so that only an approximation y of Ff is available, where $\|y - Ff\|_Y$ is small, the reconstruction of an approximation of f is problematic, because the inverse of F is not continuous. In the linear setting, this is often a consequence of the compactness of F. For example, this is the case whenever F is given as an integral operator with a suitable kernel, such as a convolution operator or the Radon transform [51, 57]. The classical method to address ill-posedness is regularization [20, 38].

Even if X and Y are usually infinite-dimensional function spaces, which is due to the fact that the models often involve integral or differential operators, the available measurements are always finite dimensional [37, 43, 3, 39, 8, 4]. In many settings, assuming that Y is a space of functions defined on a manifold $\mathcal{M}$, the finite measurements can be written as noisy pointwise samples of Ff on a suitable family of points $\{x_j\}_j \subset \mathcal{M}$, namely,

$$y_j = (Ff)(x_j) + \eta_j, \qquad j = 1, \dots, N, \tag{1}$$

where $\{\eta_j\}_j$ represents measurement noise. This problem is often studied in a random design setting, in which the points x_j are sampled i.i.d. with respect to a fixed probability distribution on $\mathcal{M}$ [17]. Several results exist in this context, regarding approaches based on statistical inverse learning (mostly for linear problems) [12, 58, 16, 15, 32, 1] and approaches studying the statistical consistency for nonlinear problems [2, 28, 52, 53]. Most of these works are based on smoothness/Gaussian priors on f; for sparsity-based priors, several works have investigated the use of compressed sensing techniques to further reduce the number of measurements in inverse problems, see e.g. [33, 7, 18, 44, 5, 6]. In the context of a deterministic design, we refer the reader to [64, 65, 50], where the authors study the case of Fourier Integral Operators F.

1.3. **Our contribution: sampling for inverse problems.** In this work, we address the inverse problem consisting of the recovery of f from the measurements $\{y_j\}_j$ given in (1) by applying the techniques based on sampling and approximation discussed in §1.1. In contrast to the approaches based on random design points mentioned above, here the points $\{x_j\}_j$ are deterministic and chosen to form a Marcinkiewicz-Zygmund family, which as we mentioned before can be achieved e. g. in any orientable compact connected Riemannian manifold without boundary. We consider abstract inverse problems under suitable smoothness assumptions on the forward map F. We prove recovery estimates and explicitly quantify the error as a function of the noise level and of the number of samples.

Furthermore, we leverage the flexibility of considering general Riemannian manifolds beyond the Euclidean setting and consider convolution operators F on compact two-point homogeneous spaces, which allows us to write F as Fourier multiplier (with respect to the basis given by the

eigenfunctions of the Laplace-Beltrami operator on $\mathcal{M}$). In this setting, the inverse problem consists in the reconstruction of a function f from noisy pointwise samples of $(h * f)(x_j)$, where h is a given filter. As a key application, we consider the convolution on the two-dimensional sphere [68, 30, 41, 40, 69], in the spirit of the results obtained in [34] with pointwise samples. We obtain explicit error estimates, and quantify how the smoothness of f and the degree of ill-posedness of F (namely, the decay of the generalized Fourier coefficients of h) affect the reconstruction error and the number of required samples. These estimates are then applied to four examples of deconvolution problems on the sphere [62, 61].

1.4. **Structure of the paper.** In Section 2 we anticipate our results specialized for the convolution problem on the sphere, in order to better illustrate the abstract approach presented in the rest of the paper. These results are applied to four different examples in Section 3. In Section 4 we review some basic facts about sampling and approximation for functions on manifolds, based on Marcinkiewicz-Zygmund families. In Section 5 we apply these results to abstract inverse problems. Finally, in Section 6 we apply the abstract derivation to convolutions on compact two-point homogeneous spaces.

2. Convolution on the sphere

In this section, motivated by several applications, we anticipate our results specialized for the convolution operator on the sphere of the Euclidean space.

2.1. **The abstract result for the sphere** S^d**.** We briefly introduce the notation and setting. Let S^d be the sphere in $\mathbb{R}^{d+1}$ endowed with the Riemann measure μ normalized as $\mu(S^d) = 1$. The orthogonal group $SO(d + 1)$ acts transitively on S^d and its action is denoted by gx where $g \in SO(d + 1)$ and $x \in S^d$. Letting $o \in S^d$ denote the north pole, the stability subgroup at o is

$$\{g \in SO(d + 1) : go = o\} = SO(d),$$

so that we identify S^d with the quotient space $SO(d + 1)/SO(d)$ by means of the map

$$SO(d + 1) \ni g \mapsto go \in S^d.$$

For a fixed filter function $h \in L^1(S^d, \mu)$, the convolution operator $F \colon L^2(S^d) \to L^2(S^d)$ by h, $Ff = h * f$, is defined as

$$(Ff)(x) = \int_{SO(d+1)} h(g^{-1}x)f(go)dg,$$

where dg is the Haar measure of $SO(d+1)$ normalized to 1. It is well known that the convolution operator has a nice structure in terms of spherical harmonics. Indeed, let Δ be the Laplace-Beltrami operator on S^d, regarded as a self-adjoint positive operator on $L^2(S^d)$. Its (distinct) eigenvalues are

$$\lambda_m^2 = m(m + d - 1) \qquad m \in \mathbb{N}.$$

For each $m \in \mathbb{N}$, let $\mathcal{H}_m$ be the space of the eigenfunctions of Δ with eigenvalue λ_m^2. The space $\mathcal{H}_m$ is finite dimensional, and we denote its dimension by

$$\dim \mathcal{H}_m = \delta_m = \binom{d+m}{d} - \binom{d+m-2}{d} \asymp m^{d-1}$$

and the corresponding orthogonal projection by Π_m. Finally, let

$$\mathcal{Q}_m = \bigoplus_{\ell=0}^{m} \mathcal{H}_\ell \qquad m \in \mathbb{N}$$

be the space of diffusion polynomials of bandwith λ_m.

Assume now that the filter function $h \in L^2(S^d) \subset L^1(S^d)$. By setting

$$b_m = \frac{1}{\delta_m} (\Pi_m h)(o) \qquad m \in \mathbb{N}, \tag{2}$$

the convolution operator F with filter h reads as

$$Ff = \sum_{m=0}^{\infty} b_m \Pi_m f \qquad f \in L^2(S^d), \tag{3}$$

i.e. F is a Fourier multiplier. Note that, for any bounded sequence $\{b_m\}_{m\in\mathbb{N}}$, Eq. (3) defines a bounded operator, which in general is not a convolution operator. For example, if $b_m = 1$ for all $m \in \mathbb{N}$, then F is the identity operator. We stress that our results hold true also for this larger class of operators.

We need to recall the definition of Sobolev spaces $H^\sigma(S^d)$ on the sphere. For $\sigma \geq 0$, we have that

$$H^\sigma(S^d) = \left\{ f \in L^2(S^d) : \sum_{m=0}^{+\infty} (1 + \lambda_m^2)^\sigma \|\Pi_m f\|^2 < +\infty \right\}$$

is a Hilbert space with respect to the norm

$$\|f\|_{H^\sigma(S^d)}^2 = \sum_{m=0}^{+\infty} (1 + \lambda_m^2)^\sigma \|\Pi_m f\|^2.$$

The following proposition provides some smoothing properties of the convolution operator; it follows from the general case treated in Proposition 24.

Proposition 1. *Choose $h \in L^2(S^d)$, set b_m as in (2) and assume that*

$$|b_m| \leq c(1 + \lambda_m^2)^{-\frac{\gamma}{2}} \tag{4}$$

for some $c, \gamma \geq 0$. Fix $\omega \geq 0$ and $\sigma = \omega + \gamma$, then F is a bounded operator from $H^\omega(S^d)$ into $H^\sigma(S^d)$, with norm bounded by c.

The next result shows the existence of a nice discretization family of points. It is a particular case of Theorems 12 and 13.

Proposition 2. *Fix $0 < \varepsilon < 1$ and $m \in \mathbb{N}$. There exists a finite family $\{x_j\}_{j=1}^{L_m}$ of distinct points of S^d such that*

$$(1 - \varepsilon)\|p\|_2^2 \leq \frac{1}{L_m} \sum_{j=1}^{L_m} |p(x_j)|^2 \leq (1 + \varepsilon)\|p\|_2^2, \qquad p \in \mathcal{Q}_m \tag{5}$$

and $L_m = \lceil C_1(\lambda_m^2)^{\frac{d}{2}} \rceil \sim m^d$ for some constant $C_1 > 0$.

It is possible to provide an explicit construction of the discretization family $\{x_j\}_{j=1}^{L_m}$; we describe the case $d = 2$ in Section 2.2. The points $\{x_j\}_{j=1}^{L_m}$ are called a Marcinkiewicz-Zygmund family. The map

$$p \in \mathcal{Q}_m \longmapsto \{p(x_j)\}_{j=1}^{L_m} \in \mathbb{C}^{L_m}$$

is ε-close to an isometry, which allows for the stable recovery of $p \in \mathcal{Q}_m$ from its pointwise samples $\{p(x_j)\}_j$.

We now state our main result for the sphere, which is a particular case of Theorem 25 and Corollary 27. The focus is the reconstruction of $f^\dagger$, where the measurement data are the noisy values of the function $Ff^\dagger = h * f^\dagger$ sampled at the discretization points $\{x_j\}_{j=1}^{L_m}$, namely,

$$y_j = (h * f^\dagger)(x_j) + \eta_j,$$

where the noise η_j is bounded by β.

Theorem 3. *Let $h \in L^2(S^d)$ such that $b_0 \neq 0$ and Eq. (4) holds true for some $\gamma \geq 0$. Fix $f^\dagger \in H^\omega(S^d)$ with $\omega \geq 0$, and assume that for some $\zeta \geq 0$, $\sigma = \omega + \gamma > d/2 + \zeta$. Given $m \in \mathbb{N}$, let $\{x_j\}_{j=1}^{L_m}$ be the discretization family of points given by Proposition 2 and let $\{y_j\}_{j=1}^{L_m}$ be the corresponding family of noisy outputs such that*

$$|y_j - Ff^\dagger(x_j)| \leq \beta \quad j = 1, \ldots, L_m$$

where $\beta \geq 0$ is the noise level. Choose $p_m^\beta \in \mathcal{Q}_m$ as

$$p_m^\beta \in \underset{p \in \mathcal{Q}_m}{\operatorname{argmin}} \frac{1}{L_m} \sum_{j=1}^{L_m} |y_j - Fp(x_j)|^2, \tag{6}$$

then

$$\|Ff^\dagger - Fp_m^\beta\|_{H^\zeta(S^d)} \leq \sqrt{\frac{1 + \kappa}{\sigma - \zeta - \frac{d}{2}}} \|Ff^\dagger\|_{H^\sigma(S^d)} (1 + \lambda_m^2)^{-\frac{\sigma - \zeta}{2} + \frac{d}{4}} + \sqrt{\kappa}\beta(1 + \lambda_m^2)^{\frac{\zeta}{2}}$$

where $\kappa = \frac{1 + \varepsilon}{1 - \varepsilon}$. If furthermore $\zeta \geq \gamma$ and for some constant $c_0 > 0$,

$$|b_m| \geq c_0(1 + \lambda_m^2)^{-\zeta/2} \text{ for all } m \geq 0, \tag{7}$$

then we have

$$\|f^\dagger - p_m^\beta\|_2 \leq c_0^{-1}\|Ff^\dagger - Fp_m^\beta\|_{H^\zeta(S^2)}.$$

The function p_m^β is a solution of a least squares problem having the space of diffusion polynomials (with a given bandwith) $\mathcal{Q}_m$ as hypothesis space (as in regularization by projection [38]), and depends only on the measured noisy data y_j. Note that $\mathcal{Q}_m$ has dimension $\sum_{\ell=0}^m \delta_\ell \asymp m^d$ and the number of discretization points is $L_m \asymp m^d$. Our result shows that, for a sufficiently small noise level β, Fp_m^β is a good approximation (in the H^ζ-norm) of the true value $h * f^\dagger$, and that, under hypothesis (7) on h, p_m^β is a good approximation (in the L^2 norm) of $f^\dagger$.

In order to highlight the dependence on the discretization parameter m and the noise level β, let us write the recovery estimate without the explicit constants, by using that $\lambda_m^2 \asymp m^2$ for $m \geq 1$:
$$\|f^\dagger - p_m^\beta\|_2 \lesssim m^{\zeta-\sigma+\frac{d}{2}} + \beta m^\zeta.$$
For a fixed noise level β, this leads us to choose m in such a way that the two terms have the same order, namely $m^{\zeta-\sigma+\frac{d}{2}} \asymp \beta m^\zeta$, so that
$$m = \left\lceil \beta^{-\frac{1}{\omega+\gamma-\frac{d}{2}}} \right\rceil, \tag{8}$$
where we used that $\sigma = \omega + \gamma$. Plugging this into the bound above yields
$$\|f^\dagger - p_m^\beta\|_2 \lesssim \beta^{1-\frac{\zeta}{\omega+\gamma-\frac{d}{2}}}. \tag{9}$$
As expected, since $\omega+\gamma-\frac{d}{2} > \zeta$ by assumption, as $\beta \to 0$, we have $p_m^\beta \to f^\dagger$, provided that the number of measurements increases accordingly to (8). Furthermore, it is possible to understand the role of the parameters involved in these estimates:

- the parameter ω such that $f^\dagger \in H^\omega(S^2)$ quantifies the smoothness assumption on $f^\dagger$;
- the parameters γ and ζ quantify the smoothness of h through the decay of its generalized Fourier coefficients in (4) and (7), and will usually be of the same order.

Therefore, according to (8), fewer samples are needed as γ and ω increase (in other words, as the problem becomes more ill-posed and as $f^\dagger$ becomes smoother). Similarly, in light of (9), a smoother $f^\dagger$ (larger ω) yields a smaller reconstruction error, but a more ill-posed problem (larger γ and ζ) yields a larger error.

In [34], results similar to ours were obtained for the two-dimensional sphere, but in a setting with statistical noise and a different choice of approximating polynomial. In particular, exact knowledge of the multiplier associated with the operator is required, and the sampling points must form a spherical design—a construction that is more challenging than that of the Marcinkiewicz-Zygmund family provided by Proposition 2.

2.2. **The particular case of the two-dimensional sphere** S^2. We provide an example of a discretization family for the two-dimensional sphere. It can be shown (see [22, Theorem 5.1]) that if one can divide the manifold (in this case the sphere S^2) into N regions with the same measure and bounded diameter, then every choice of points, one in each region, gives a discretization family, as in Proposition 2. For the 2-dimensional sphere, we describe an explicit partition [56]. Fix $m \in \mathbb{N}$ and let $N = L_m$ be the size of the discretization family. Let

- $\theta_0 = \arccos(1 - 50/N)$;
- s be the largest odd integer smaller than or equal to $\sqrt{\pi N}/2$;

- $\Delta\theta = (\pi - 2\theta_0)/s$;
- $\theta'_k = \theta_0 + k\Delta\theta$ for $k = 1, 2, \ldots, s$, $\theta'_{s+1} = \pi$;
- and $y_k = N(\cos\theta'_{k-1} - \cos\theta'_k)/2$, for $k = 1, 2, \ldots, s$.

It can be proved that there exists a symmetric sequence of integers $\{\ell_i\}_{i=1}^s$ such that

(1) $\sum_{i=1}^s \ell_i = \sum_{i=1}^s y_i = N - 50$.
(2) $|y_1 - \ell_1| = |y_s - \ell_s| \leq 1/2$ and $|y_i - \ell_i| \leq 1$ for $i = 2, \ldots s - 1$.
(3) $\left|\sum_{i=1}^k (y_i - \ell_i)\right| \leq 1/2$, for $k = 1, \ldots, s$.

Finally set $\ell_0 = \ell_{s+1} = 25$. It is shown in [56] that for $N \geq 50$, the regions of the sphere described in spherical coordinates as $D_{k,j}^N = [\theta_{k-1}, \theta_k] \times [2\pi(j-1)/\ell_k, 2\pi j/\ell_k]$, for $k = 0, \ldots, s+1$, $j = 1, \ldots, \ell_k$, where $\theta_k = \arccos(1 - (2/N)\sum_{i=0}^k \ell_i)$ for $k = 0, \ldots, s + 1$, and $\theta_{-1} = 0$, form a partition of the sphere of regions with equal area $1/N$, each contained in a spherical cap of radius bounded above by $c_4 N^{-1/2}$, for some $c_4 > 0$. It can be easily shown that all regions contain a spherical cap of radius $c_3 N^{-1/2}$, for some $c_3 > 0$.

Therefore we can give the following constructive version of Proposition 2 (see also Theorem 12).

Proposition 4. *Fix $0 < \varepsilon < 1$ and $m \in \mathbb{N}$. For every possible choice of $x_{k,j} \in D_{k,j}^{L_m}$, the family $\{x_{k,j}\}_{k=0,j=1}^{s+1,\ell_k}$ is such that*

$$(1 - \varepsilon)\|p\|_2^2 \leq \frac{1}{L_m} \sum_{k=0}^{s+1} \sum_{j=1}^{\ell_k} |p(x_{k,j})|^2 \leq (1 + \varepsilon)\|p\|_2^2, \qquad p \in \mathcal{Q}_m \tag{10}$$

whenever $L_m = \lceil C_1 m(m+1) \rceil$ for some constant $C_1 > 0$.

3. Four examples

In this section, we discuss four examples related to Theorem 3 for the two-dimensional sphere S^2, which are motivated by real-world applications. For further details, the reader is referred to [62, 61].

3.1. **Sea surface temperature anomalies.** This experiment [62, 61] simulates the global map of temperature anomalies of the sea surface produced by NASA's Aqua satellite in January 2011, by sampling it in $N = 6745$ locations. Since temperatures typically have smooth variations, one may assume the anomaly function $f^\dagger$ to be an element of $H^\omega(S^2)$ for some $\omega > 1$. The inversion problem here reduces to F being the identity operator, that is $b_m = 1$ for every m (as already noted, our result holds true for any Fourier multiplier of the form (3)).

We may therefore set $\gamma = 0$, and apply Theorem 3 with $\zeta = 0$ to obtain

$$\|p_m^\beta - f^\dagger\|_2 \leq \sqrt{1 + \kappa}\|f^\dagger\|_{H^\omega(S^2)} \frac{1}{(\omega - 1)^{\frac{1}{2}}} \frac{1}{(1 + m(m+1))^{\frac{\omega-1}{2}}} + \sqrt{\kappa}\beta.$$

3.2. **Wildfires and Deforestation.** This example deals with global density maps of trees and wildfires across the globe for the year 2016. In both experiments, the data used in [62, 61] consists of tree and fire counts recorded by NASA's Aqua and Terra satellites. In [62, 61], it is assumed that the spatial density $f^\dagger$ of trees (respectively fires) is a function in $H^\omega(S^2)$ with $\omega > 1$, where we model the Earth surface by S^2. In both cases, the inverse problem consists in recovering $f^\dagger$ from a finite number of pointwise evaluations of $Ff^\dagger = h * f^\dagger$, where h is the characteristic function of a small region around the North Pole. Here we propose to replace this choice of h with the characteristic function of a spherical cap of angle $\vartheta_0 \leq \pi/2$, so that $h(x) = \chi_{[0,\vartheta_0]}(\rho(o,x))$ where $\rho(o,x)$ is the angular distance of x from the North Pole. With this choice, by Eq. (25) below and Rodrigues' formula (see [67, (4.3.1)]), if $m \geq 1$,

$$b_m = \frac{1}{m} P_{m-1}^{1,1}(\cos\vartheta_0) \left(\sin\frac{\vartheta_0}{2}\right)^2 \left(\cos\frac{\vartheta_0}{2}\right)^2.$$

Let us now prove the upper bound (4) with $\gamma = 3/2$. Kogbetliantz inequality [24, eq. (1.3)] applied to our setting reads as

$$\left|P_{m-1}^{1,1}(\cos\vartheta_0)(\sin\vartheta_0)^{\frac{3}{2}}\right| \leq \frac{4\Gamma\left(m+\frac{1}{2}\right)}{\Gamma\left(\frac{3}{2}\right)(m+1)\Gamma(m)}$$

for all $\vartheta_0 \leq \pi/2$ and for all $m \geq 1$. It follows that for $m \geq 1$

$$|b_m| \leq \left|\frac{(\sin\vartheta_0)^{\frac{1}{2}}}{4m} P_{m-1}^{1,1}(\cos\vartheta_0)(\sin\vartheta_0)^{\frac{3}{2}}\right|$$

$$\leq \left|\frac{(\sin\vartheta_0)^{\frac{1}{2}}}{m} \frac{\Gamma\left(m+\frac{1}{2}\right)}{\Gamma\left(\frac{3}{2}\right)(m+1)\Gamma(m)}\right| \leq \frac{3^{\frac{3}{4}}}{2} \frac{(\sin\vartheta_0)^{\frac{1}{2}}}{(1+m(m+1))^{\frac{3}{4}}}.$$

Since $b_0 = (1 - \cos\vartheta_0)/2 \leq (\sin\vartheta_0)^2/2 \leq 3^{3/4}(\sin\vartheta_0)^{1/2}/2$, we can deduce that indeed for all $m \geq 0$,

$$|b_m| \leq \frac{3^{\frac{3}{4}}}{2} \frac{(\sin\vartheta_0)^{\frac{1}{2}}}{(1+m(m+1))^{\frac{3}{4}}}.$$

In Figure 1a we illustrate this bound numerically.

We now move to studying the lower bound (7) with $\zeta = 3/2$. It can be proved (see [11] and also [14]) that $x = \cos(2\pi/41)$ is a 3/2-gegenbadly approximable number, which in our notation means that there exists a constant c such that for every $m \geq 1$

$$|P_{m-1}^{1,1}(x)| \geq cm^{-\frac{1}{2}}.$$

In other words, if we set $\vartheta_0 = 2\pi/41$, then for every $m \geq 1$,

$$|b_m| \geq c_0 \frac{1}{(1+m(m+1))^{\frac{3}{4}}}.$$

Numerical simulations (Figure 1b) give an estimate $c_0 \geq 0.4 * 10^{-3}$. (Similar remarks hold for other values of ϑ_0).

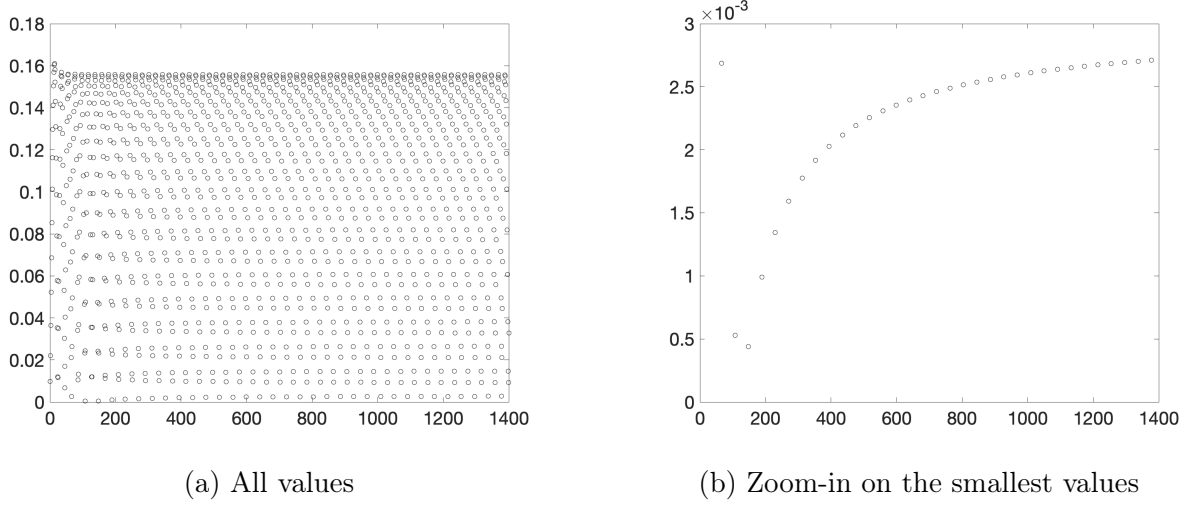(a) All values   (b) Zoom-in on the smallest values

FIGURE 1. Plot of the sequence $(1 + m(m + 1))^{3/4}|b_m|$ for $\vartheta_0 = 2\pi/41$ and $m = 1, \ldots, 1400$.

Hence, by applying Theorem 3 and Proposition 1 with $\gamma = \zeta = 3/2$, we obtain

$$\|f^\dagger - p_m^\beta\|_2 \leq c_0^{-1}\|Ff^\dagger - Fp_m^\beta\|_{H^{\frac{3}{2}}(S^d)}$$

$$\leq c_0^{-1}\sqrt{\frac{1+\kappa}{\omega-1}}\,\|Ff^\dagger\|_{H^{\omega+\frac{3}{2}}(S^2)}(1 + m(m+1))^{-\frac{\omega-1}{2}} + c_0^{-1}\sqrt{\kappa}\beta(1 + m(m+1))^{\frac{3}{4}}$$

$$\leq \frac{3^{\frac{3}{4}}}{2}(\sin\vartheta_0)^{\frac{1}{2}}c_0^{-1}\sqrt{\frac{1+\kappa}{\omega-1}}\,\|f^\dagger\|_{H^\omega(S^2)}(1 + m(m+1))^{-\frac{\omega-1}{2}} + c_0^{-1}\sqrt{\kappa}\beta(1 + m(m+1))^{\frac{3}{4}}.$$

3.3. **Planck and the Cosmic Microwave Background.** Here, the aim is to estimate the full sky intensity map (displaying the intensity or temperature of every astronomical radio source across the celestial sphere) from a point source sky model [62]. The data consists of noisy directional samples of the map, convolved with a model of the point spread function of the radio telescope (the beamshape of the dish antenna). One wishes to estimate the intensity field $f^\dagger$ of the stars on the celestial sphere using samples obtained by steering the dish antenna towards N directions. Similarly as before, we assume $f^\dagger$ to be an element of the Sobolev space $H^\omega(S^2)$ for some $\omega > 1$. The proposed point spread function is

$$h(x) = g(\rho(x, o)),$$

where

$$g(\vartheta) = \left(\frac{\lambda_0 J_1\left(4\pi R\sin\left(\frac{\vartheta}{2}\right)\right)}{2R\sin\left(\frac{\vartheta}{2}\right)}\right)^2, \tag{11}$$

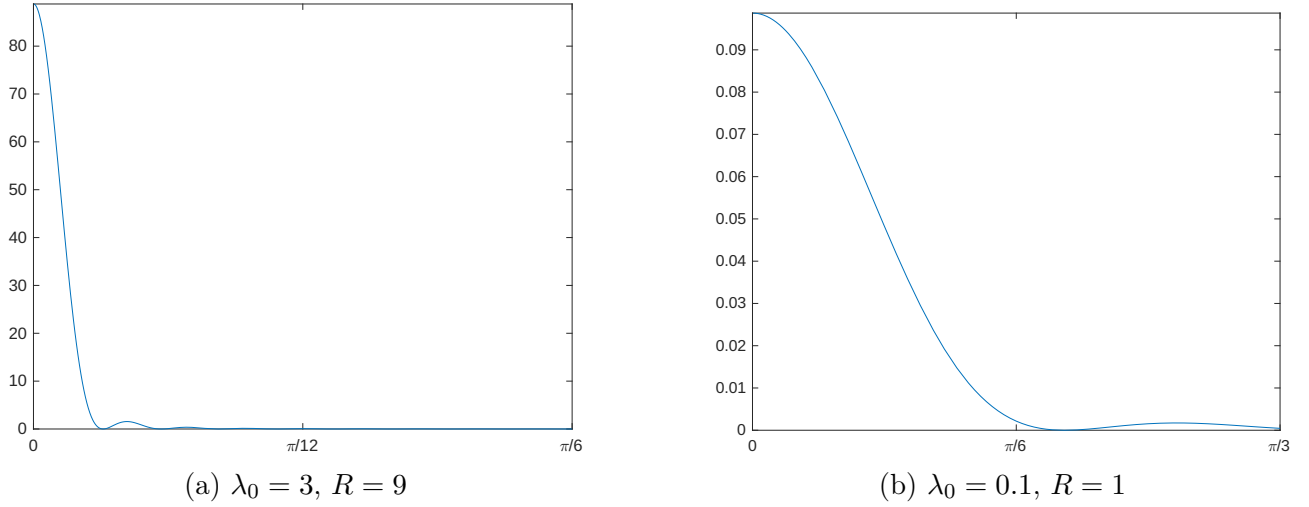(a) $\lambda_0 = 3$, $R = 9$ (b) $\lambda_0 = 0.1$, $R = 1$

FIGURE 2. The function g in (11), used to construct the convolution filter of the example from §3.3.

and J_1 is the Bessel function of the first kind and order 1, corresponding to the far-field beamshape of an ideal circular aperture with radius R m operating at a wavelength λ_0 m. As usual, $\rho(x, o)$ is the angular distance of x from the north pole o.

In [62] two possibilities are studied, one with $\lambda_0 = 3$ and $R = 9$, where $N = 768$ samples are used, and the second with $R = 1$, $\lambda_0 = 0.1$ and $N = 9248$ samples corresponding to data taken from the Planck radio telescope. See Figure 2 for an illustration of the function g in these two cases.

Since the function g can be extended to $\mathbb{R}$ as an even, C^∞, 2π-periodic function, h belongs to all Sobolev spaces $H^\sigma(S^2)$, $\sigma \geq 0$. For the sake of simplicity, we study the case $\sigma = 0$, so that for $m \geq 1$ (see Remark 29),

$$|b_m| \leq \frac{\|h\|_2}{(1 + m(m+1))^{\frac{1}{4}}}.$$

We can apply Theorem 3 with $\gamma = 1/2$ and, say, $\zeta = 0$ to obtain

$$\|F p_m^\beta - F f^\dagger\|_2 \leq \sqrt{1 + \kappa} \|h\|_2 \|f^\dagger\|_{H^\omega(\mathcal{B}, \mathcal{E})} \frac{1}{\left(\omega - \frac{1}{2}\right)^{\frac{1}{2}}} \frac{1}{(1 + m(m+1))^{\frac{\omega - 1/2}{2}}} + \sqrt{\kappa} \beta,$$

where

$$\|h\|_2 = \left(\int_0^\pi |g(\vartheta)|^2 \sin\left(\frac{\vartheta}{2}\right) \cos\left(\frac{\theta}{2}\right) d\vartheta \right)^{\frac{1}{2}} = \left(2 \int_0^1 \left| \frac{\lambda_0 J_1(4\pi Rx)}{2Rx} \right|^4 x \, dx \right)^{\frac{1}{2}}.$$

When $\lambda_0 = 3$ and $R = 9$ one gets $\|h\|_2 \approx 1.064$, while when $\lambda_0 = 0.1$ and $R = 1$, one gets $\|h\|_2 \approx 0.0106$.

Unfortunately, a lower bound of the form (7) does not hold because h is smooth and so $(b_m)_m$ has rapid decay; thus we cannot state an explicit bound for $\|p_m^\beta - f^\dagger\|_2$.

3.4. **Lunar elemental abundance maps.** In this example, the aim is to build global distribution maps of radioactive elements on the surface of the Moon using real data collected by NASA's Lunar Prospector probe [62]. These maps are used by scientists to retrace the Moon's geologic history. The emission of gamma rays at the surface of the Moon is modeled by an intensity function $f^\dagger$ belonging to some space $H^\omega(S^2)$, with $\omega > 1$. The problem here consists in recovering $f^\dagger$ from a finite number of evaluations of $Ff^\dagger = h * f^\dagger$, where h is the point spread function of the orbital gamma ray spectrometer, which is well fitted by

$$h(x) = \left(1 + \frac{R^2(\rho(x,o))^2}{2\sigma(t)^2}\right)^{-\iota(t)-1}$$

where $R = 1737.1\,\mathrm{km}$ is the radius of the Moon, $t = 30\,\mathrm{km}$ is the altitude of the spacecraft, and

$$\sigma(t) = 0.704t + 1.39\,\mathrm{km}, \qquad \iota(t) = -4.87\,\mathrm{km}^{-1} \times 10^{-4}t + 0.631.$$

As in §3.3, we use that $h \in L^2(S^2)$, so that

$$|b_m| \leq \frac{\|h\|_2}{(1 + m(m+1))^{\frac{1}{4}}}$$

and we can apply Theorem 3 with $\gamma = 1/2$ and, say, $\zeta = 0$ to obtain

$$\|Fp_m^\beta - Ff^\dagger\|_2 \leq \sqrt{1+\kappa}\|h\|_2\|f^\dagger\|_{H^\omega(S^2)}\frac{1}{\left(\omega - \frac{1}{2}\right)^{\frac{1}{2}}}\frac{1}{(1 + m(m+1))^{\frac{\omega-1/2}{2}}} + \sqrt{\kappa}\beta,$$

where

$$\|h\|_2 = \left(\int_0^\pi \left(1 + \frac{R^2\vartheta^2}{2\sigma(t)^2}\right)^{-2\iota(t)-2}\sin\left(\frac{\vartheta}{2}\right)\cos\left(\frac{\vartheta}{2}\right)d\vartheta\right)^{\frac{1}{2}} \approx 0.0061.$$

As in §3.3, we do not derive a lower bound of the form (7).

4. Sampling

In this section, we revisit some fairly recent results regarding the approximation of smooth functions from a finite number of pointwise evaluations. This collection of results begins with the work of K. Gröchenig [29], concerning the approximation of a function in a properly defined Sobolev space with a certain "polynomial" obtained by least squares from the samples of the function on the nodes of a Marcinkiewicz-Zygmund family. One of the main features of this work is that the simplicity of the arguments allows for a great generality. For example, the ambient space is just a topological space with a probability measure, and the "polynomials" are simply the finite linear combinations of the functions of any orthonormal system in L^2.

Soon after, Gröchenig's results were revisited in more explicit situations: for example, W. Lu and H. Wang [46] studied the case of the sphere, and J. Li, Y. Ling, J. Geng and H. Wang [45] the case of a compact Riemannian manifold. In these cases, the orthonormal system is given by the eigenfunctions of the Laplace-Beltrami operator and the Sobolev spaces are defined as

usual. In these situations, it is possible to improve Gröchenig's results, obtaining asymptotically optimal estimates on the decay of the approximation error.

4.1. **Sampling with Marcinkiewicz-Zygmund families.** Let $\mathcal{M}$ be a space with a topology (usually a compact space) and let μ be a probability measure on $\mathcal{M}$. Let $L^2(\mathcal{M}, \mu)$ denote the Hilbert space of (equivalence classes of) square-integrable functions, equipped with

$$\langle f, g \rangle_2 = \int_{\mathcal{M}} f\, \overline{g}\, d\mu, \qquad \|f\|_2^2 = \int_{\mathcal{M}} |f|^2\, d\mu.$$

Futhermore, we assume that

(i) $\mathcal{S} = \{\psi_k : k \in \mathbb{N}\}$ is an orthonormal system in $L^2(\mathcal{M}, \mu)$, such that each ψ_k is continuous and bounded on $\mathcal{M}$, and $\psi_0 \equiv 1$. We set $H(\mathcal{S}) = \overline{\operatorname{span}\{\psi_k : k \in \mathbb{N}\}}$.

(ii) $\mathcal{L} = \{\Lambda_k : \Lambda_k \geq 0\}$ is a non-decreasing sequence with $\Lambda_k \to +\infty$.

Let the Sobolev space associated to $\mathcal{S}$ and $\mathcal{L}$ be defined as

$$H^\sigma(\mathcal{S}, \mathcal{L}) = \left\{ f \in H(\mathcal{S}) : \|f\|_{H^\sigma(\mathcal{S},\mathcal{L})} < +\infty \right\}, \qquad \sigma \geq 0,$$

where the norm is given by

$$\|f\|_{H^\sigma(\mathcal{S},\mathcal{L})} = \left(\sum_{k=0}^{+\infty} |\langle f, \psi_k \rangle_2|^2 (1 + \Lambda_k^2)^\sigma \right)^{\frac{1}{2}}. \tag{12}$$

Notice that $H^\sigma(\mathcal{S}, \mathcal{L})$ is a Hilbert space, with the scalar product

$$\langle f, g \rangle_{H^\sigma(\mathcal{S},\mathcal{L})} = \sum_{k=0}^{+\infty} \langle f, \psi_k \rangle_2 \overline{\langle g, \psi_k \rangle_2} (1 + \Lambda_k^2)^\sigma.$$

Let $\mathcal{Q}_r(\mathcal{S}, \mathcal{L})$ be the set of "polynomials" of degree r on $\mathcal{M}$, that is

$$\mathcal{Q}_r(\mathcal{S}, \mathcal{L}) = \operatorname{span}\{\psi_k : \Lambda_k \leq r\} \subseteq L^2(\mathcal{M}).$$

Definition 5. A doubly-indexed set $\{x_{m,j} : m \in \mathbb{N}, j = 1, \ldots, L_m\} \subseteq \mathcal{M}$, a set of positive weights $\{\tau_{m,j} : m \in \mathbb{N}, j = 1, \ldots, L_m\}$, and a non-negative, increasing sequence $\{r_m\}$ of "degrees" diverging to $+\infty$, such that

$$A\|q\|_2^2 \leq \sum_{j=1}^{L_m} |q(x_{m,j})|^2 \tau_{m,j} \leq B\|q\|_2^2 \qquad \text{for all } q \in \mathcal{Q}_{r_m}(\mathcal{S}, \mathcal{L}), \tag{13}$$

with constants $A, B > 0$ independent of m, is called a Marcinkiewicz-Zygmund family $\mathcal{X}$ for $\mathcal{S}$ and $\mathcal{L}$. We set $\kappa = B/A$.

We define the series

$$C_\sigma(\mathcal{S}, \mathcal{L}) = \sup_{x \in \mathcal{M}} \left(\sum_{k=0}^{+\infty} |\psi_k(x)|^2 (1 + \Lambda_k^2)^{-\sigma} \right)^{\frac{1}{2}}$$

and the remainders

$$\phi_\sigma(r, \mathcal{S}, \mathcal{L}) = \sup_{x \in \mathcal{M}} \left(\sum_{k:\Lambda_k > r} |\psi_k(x)|^2 (1 + \Lambda_k^2)^{-\sigma} \right)^{\frac{1}{2}}.$$

When $C_\sigma(\mathcal{S}, \mathcal{L}) < +\infty$, then $\phi_\sigma(r, \mathcal{S}, \mathcal{L}) < +\infty$ too. If $\phi_\sigma(r, \mathcal{S}, \mathcal{L}) \to 0$ as $r \to +\infty$, then $C_\sigma(\mathcal{S}, \mathcal{L}) < +\infty$.

Definition 6. Let $\mathcal{X}$ be a Marcinkiewicz-Zygmund family for $\mathcal{S}$ and $\mathcal{L}$, with nodes $\{x_{m,j}\}$, weights $\{\tau_{m,j}\}$ and degrees $\{r_m\}$. For any finite sequence of complex numbers $y_m = \{y_{m,j}\}_{j=1}^{L_m}$, we set

$$T^\dagger y_m = \operatorname*{argmin}_{q \in \mathcal{Q}_{r_m}(\mathcal{S}, \mathcal{L})} \sum_{j=1}^{L_m} |y_{m,j} - q(x_{m,j})|^2 \tau_{m,j}. \tag{14}$$

The following result shows that this minimization problem has indeed a unique minimizer, and provides a useful bound on its norm.

Lemma 7. *For any finite sequence of complex numbers* $y_m = \{y_{m,j}\}_{j=1}^{L_m}$, *problem* (14) *admits a unique minimizer* $T^\dagger y_m \in \mathcal{Q}_{r_m}(\mathcal{S}, \mathcal{L})$, *and*

$$\|T^\dagger y_m\|_2 \leq A^{-\frac{1}{2}} \left(\sum_{j=1}^{L_m} |y_{m,j}|^2 \tau_{m,j} \right)^{\frac{1}{2}}.$$

Proof. Let us equip $\mathbb{C}^{L_m}$ with the weighted norm

$$\|a\|_{\mathbb{C}^{L_m}}^2 := \sum_{j=1}^{L_m} |a_j|^2 \tau_{m,j}.$$

Let $T \colon \mathcal{Q}_{r_m}(\mathcal{S}, \mathcal{L}) \to \mathbb{C}^{L_m}$ be the sampling operator defined by $Tq = \{q(x_{m,j})\}_j$. By (13),

$$A^{\frac{1}{2}} \|q\|_2 \leq \|Tq\|_{\mathbb{C}^{L_m}} \leq B^{\frac{1}{2}} \|q\|_2, \tag{15}$$

and T is therefore injective. Let $\mathcal{W} = \operatorname{Im} T$ and $\Pi_{\mathcal{W}}$ be the orthogonal projection from $\mathbb{C}^{L_m}$ onto $\mathcal{W}$, then T is an invertible operator from $\mathcal{Q}_{r_m}(\mathcal{S}, \mathcal{L})$ onto $\mathcal{W}$ with inverse $S \colon \mathcal{W} \to \mathcal{Q}_{r_m}(\mathcal{S}, \mathcal{L})$. The bound (15) implies that

$$A^{\frac{1}{2}} \|Sy\|_2 \leq \|y\|_{\mathbb{C}^{L_m}} \qquad \forall y \in \mathcal{W},$$

so that $\|S\| \leq A^{-\frac{1}{2}}$. Furthermore, since

$$\operatorname*{argmin}_{q \in \mathcal{Q}_{r_m}(\mathcal{S}, \mathcal{L})} \|Tq - y_m\|_{\mathbb{C}^{L_m}}^2 = \operatorname*{argmin}_{q \in \mathcal{Q}_{r_m}(\mathcal{S}, \mathcal{L})} \|Tq - \Pi_{\mathcal{W}} y_m\|_{\mathbb{C}^{L_m}}^2 = S\Pi_{\mathcal{W}} y_m,$$

it follows that the pseudoinverse of T (see [55]) is $T^\dagger = S\Pi_{\mathcal{W}}$ and $\|T^\dagger\| \leq \|S\| \|\Pi_{\mathcal{W}}\| \leq A^{-\frac{1}{2}}$. $\square$

The following theorem holds (see [29, Theorem 3.3]). It concerns the reconstruction of a smooth function $f^\dagger \in H^\sigma(\mathcal{S}, \mathcal{L})$ from its pointwise samples $\{f^\dagger(x_{m,j})\}_j$.

Theorem 8. *Let $\mathcal{X}$ a Marcinkiewicz-Zygmund family for $\mathcal{S}$ and $\mathcal{L}$, with nodes $\{x_{m,j}\}$, weights $\{\tau_{m,j}\}$ and degrees $\{r_m\}$. Let $f^\dagger \in H^\zeta(\mathcal{S}, \mathcal{L})$ be continuous for some $\zeta \geq 0$, and define*

$$q_m = T^\dagger\{f^\dagger(x_{m,j})\} = \operatorname*{argmin}_{q \in \mathcal{Q}_{r_m}(\mathcal{S}, \mathcal{L})} \sum_{j=1}^{L_m} |f^\dagger(x_{m,j}) - q(x_{m,j})|^2 \tau_{m,j}.$$

Then for every $\sigma \geq \zeta$ and $m \in \mathbb{N}$

$$\|f^\dagger - q_m\|_{H^\zeta(\mathcal{S}, \mathcal{L})} \leq \sqrt{1 + \kappa} \|f^\dagger\|_{H^\sigma(\mathcal{S}, \mathcal{L})} \phi_{\sigma - \zeta}(r_m, \mathcal{S}, \mathcal{L}).$$

For the sake of completeness, and given that this version is slightly more general and with a somewhat refined constant than the original result in [29], we report here the elementary proof.

Proof. Let us call P_{r_m} the orthogonal projection of $H(\mathcal{S})$ onto $Q_{r_m}(\mathcal{S}, \mathcal{L})$. Given that if $f^\dagger \in H^\zeta(\mathcal{S}, \mathcal{L})$, then $f^\dagger - P_{r_m} f^\dagger$ remains orthogonal to $Q_{r_m}(\mathcal{S}, \mathcal{L})$ in $H^\zeta(\mathcal{S}, \mathcal{L})$,

$$\|f^\dagger - q_m\|^2_{H^\zeta(\mathcal{S}, \mathcal{L})} = \|f^\dagger - P_{r_m} f\|^2_{H^\zeta(\mathcal{S}, \mathcal{L})} + \|P_{r_m} f^\dagger - q_m\|^2_{H^\zeta(\mathcal{S}, \mathcal{L})}.$$

Now,

$$\|f^\dagger - P_{r_m} f^\dagger\|_{H^\zeta(\mathcal{S}, \mathcal{L})} = \left\| \sum_{\Lambda_k > r_m} \langle f^\dagger, \psi_k \rangle_2 \psi_k \right\|_{H^\zeta(\mathcal{S}, \mathcal{L})} = \left\| \sum_{\Lambda_k > r_m} \langle f^\dagger, \psi_k \rangle_2 (1 + \Lambda_k^2)^{\frac{\zeta}{2}} \psi_k \right\|_2$$

$$\leq \left\| \sum_{\Lambda_k > r_m} \langle f^\dagger, \psi_k \rangle_2 (1 + \Lambda_k^2)^{\frac{\sigma}{2}} (1 + \Lambda_k^2)^{-\frac{\sigma - \zeta}{2}} \psi_k \right\|_\infty \leq \|f^\dagger\|_{H^\sigma(\mathcal{S}, \mathcal{L})} \phi_{\sigma - \zeta}(r_m, \mathcal{S}, \mathcal{L}).$$

On the other hand, using the notation of the proof of Lemma 7, since $P_{r_m} f^\dagger \in Q_{r_m}(\mathcal{S}, \mathcal{L})$,

$$\|P_{r_m} f^\dagger - q_m\|_{H^\zeta(\mathcal{S}, \mathcal{L})} \leq (1 + r_m^2)^{\frac{\zeta}{2}} \|P_{r_m} f^\dagger - q_m\|_2$$

$$= (1 + r_m^2)^{\frac{\zeta}{2}} \|T^\dagger\{P_{r_m} f^\dagger(x_{m,j})\} - T^\dagger\{f^\dagger(x_{m,j})\}\|_2$$

$$\leq A^{-\frac{1}{2}} (1 + r_m^2)^{\frac{\zeta}{2}} \left(\sum_{j=1}^{L_m} |P_{r_m} f^\dagger(x_{m,j}) - f^\dagger(x_{m,j})|^2 \tau_{m,j} \right)^{\frac{1}{2}}$$

$$\leq B^{\frac{1}{2}} A^{-\frac{1}{2}} (1 + r_m^2)^{\frac{\zeta}{2}} \sup_x |P_{r_m} f^\dagger(x) - f^\dagger(x)|$$

$$= B^{\frac{1}{2}} A^{-\frac{1}{2}} (1 + r_m^2)^{\frac{\zeta}{2}} \sup_x \left| \sum_{\Lambda_k > r_m} \langle f^\dagger, \psi_k \rangle_2 \psi_k(x) \right|$$

$$\leq B^{\frac{1}{2}} A^{-\frac{1}{2}} (1 + r_m^2)^{\frac{\zeta}{2}} \|f^\dagger\|_{H^\sigma(\mathcal{S}, \mathcal{L})} \phi_\sigma(r_m, \mathcal{S}, \mathcal{L})$$

$$\leq B^{\frac{1}{2}} A^{-\frac{1}{2}} \|f^\dagger\|_{H^\sigma(\mathcal{S}, \mathcal{L})} \phi_{\sigma - \zeta}(r_m, \mathcal{S}, \mathcal{L}),$$

where we have used that $\sum_{j=1}^{L_m} \tau_{m,j} \leq B$ (simply set $q \equiv 1$ in (13)). □

Notice that normally $\sum_{j=1}^{L_m} \tau_{m,j} = 1 \leq B$, in which case the factor κ could be replaced by $A^{-\frac{1}{2}}$.

In the original theorem in [29], ζ is set equal to 0 and $\mathcal{S}$ is assumed to be a complete orthonormal system, so that $H(\mathcal{S}) = L^2(\mathcal{M})$. Also, in that paper, r_m is set to be equal to m.

In a real-life context, evaluations $f^\dagger(x_{m,j})$ may be subject to certain measurement errors. The following result takes this possibility into account.

Corollary 9. *Let $\mathcal{X}$ be a Marcinkiewicz-Zygmund family for $\mathcal{S}$ and $\mathcal{L}$, with nodes $\{x_{m,j}\}$, weights $\{\tau_{m,j}\}$ and degrees $\{r_m\}$. Let $f^\dagger$ be a continuous function in $H^\zeta(\mathcal{S}, \mathcal{L})$, $\zeta \geq 0$, and let $\{y_{m,j}\}$ be the corresponding family of noisy samples such that*

$$|y_{m,j} - f^\dagger(x_{m,j})| \leq \beta, \qquad j = 1, \ldots, L_m,$$

where $\beta \geq 0$ is the noise level. Set

$$q_m^\beta = T^\dagger\{y_{m,j}\} = \operatorname*{argmin}_{q \in \mathcal{Q}_{r_m}(\mathcal{S}, \mathcal{L})} \sum_{j=1}^{L_m} |y_{m,j} - q(x_{m,j})|^2 \tau_{m,j}.$$

Then for every $\sigma \geq \zeta$ and $m \in \mathbb{N}$

$$\|f^\dagger - q_m^\beta\|_{H^\zeta(\mathcal{S}, \mathcal{L})} \leq \sqrt{1 + \kappa}\|f^\dagger\|_{H^\sigma(\mathcal{S}, \mathcal{L})} \phi_{\sigma - \zeta}(r_m, \mathcal{S}, \mathcal{L}) + \sqrt{\kappa}\beta(1 + r_m^2)^{\frac{\zeta}{2}}.$$

Proof. By Lemma 7 and Theorem 8 we have

$$\|f^\dagger - T^\dagger\{y_{m,j}\}\|_{H^\zeta(\mathcal{S}, \mathcal{L})}$$

$$\leq \|f^\dagger - T^\dagger\{f^\dagger(x_{m,j})\}\|_{H^\zeta(\mathcal{S}, \mathcal{L})} + \|T^\dagger\{f^\dagger(x_{m,j})\} - T^\dagger\{y_{m,j}\}\|_{H^\zeta(\mathcal{S}, \mathcal{L})}$$

$$\leq \|f^\dagger - T^\dagger\{f^\dagger(x_{m,j})\}\|_{H^\zeta(\mathcal{S}, \mathcal{L})} + (1 + r_m^2)^{\frac{\zeta}{2}} \|T^\dagger\{f^\dagger(x_{m,j})\} - T^\dagger\{y_{m,j}\}\|_2$$

$$\leq \sqrt{1 + \kappa}\|f^\dagger\|_{H^\sigma(\mathcal{S}, \mathcal{L})} \phi_{\sigma - \zeta}(r_m, \mathcal{S}, \mathcal{L}) + A^{-\frac{1}{2}}(1 + r_m^2)^{\frac{\zeta}{2}} \left(\sum_{k=1}^{L_m} |f^\dagger(x_{m,j}) - y_{m,j}|^2 \tau_{m,j} \right)^{\frac{1}{2}}$$

$$\leq \sqrt{1 + \kappa}\|f^\dagger\|_{H^\sigma(\mathcal{S}, \mathcal{L})} \phi_\sigma(r_m, \mathcal{S}, \mathcal{L}) + B^{\frac{1}{2}} A^{-\frac{1}{2}} \beta(1 + r_m^2)^{\frac{\zeta}{2}},$$

where we have used that $\sum_{j=1}^{L_m} \tau_{m,j} \leq B$. $\qquad\square$

The following result gives a lower estimate for the function $\phi_\sigma(r_m, \mathcal{S}, \mathcal{L})$.

Proposition 10. *Assume that for a doubly indexed collection of nodes $\{x_{m,j}\}_{m=1, j=1}^{+\infty, L_m}$, for corresponding weights $\{\tau_{m,j}\}$, for an increasing sequence of degrees $\{r_m\}_{m=1}^{+\infty}$ diverging to ∞, and for $\sigma \geq 0$ and $\zeta \geq 0$, there exists a function $\Psi_{\sigma, \zeta}(r_m)$ such that for every polynomial f (finite linear combination of the ψ_k's), if we define*

$$q_m = \operatorname*{argmin}_{q \in \mathcal{Q}_{r_m}(\mathcal{S}, \mathcal{L})} \sum_{j=1}^{L_m} |f(x_{m,j}) - q(x_{m,j})|^2 \tau_{m,j},$$

there holds

$$\|f - q_m\|_{H^\zeta(\mathcal{S}, \mathcal{L})} \leq \|f\|_{H^\sigma(\mathcal{S}, \mathcal{L})} \Psi_{\sigma, \zeta}(r_m).$$

Then, if we let k_m be the smallest value k for which $\Lambda_k > r_m$, we have

$$\Psi_{\sigma, \zeta}(r_m) \geq (1 + \Lambda_{k_m}^2)^{-\frac{\sigma - \zeta}{2}}.$$

Proof. Set $f = \psi_{k_m}$. Then f is orthogonal to $\mathcal{Q}_{r_m}(\mathcal{S}, \mathcal{L})$, and so

$$(1 + \Lambda_{k_m}^2)^\zeta = \|f\|_{H^\zeta(\mathcal{S},\mathcal{L})}^2 \leq \|f\|_{H^\zeta(\mathcal{S},\mathcal{L})}^2 + \| - q_m\|_{H^\zeta(\mathcal{S},\mathcal{L})}^2 = \|f - q_m\|_{H^\zeta(\mathcal{S},\mathcal{L})}^2$$
$$\leq \|f\|_{H^\sigma(\mathcal{S},\mathcal{L})}^2 (\Psi_{\sigma,\zeta}(r_m))^2 = (1 + \Lambda_{k_m}^2)^\sigma (\Psi_{\sigma,\zeta}(r_m))^2,$$

as claimed. $\qquad\square$

Assume now that $\mathcal{M}$ is a compact Riemannian manifold without boundary. Let $\mathcal{S}$ denote the orthonormal basis of eigenfunctions of the Laplace-Beltrami operator and $\mathcal{L}$ the set of the square roots of the corresponding eigenvalues. It has been proved in [45] (see also [46]) that in this case, under the hypotheses of Theorem 8 with $\zeta = 0$,

$$\|f^\dagger - q_m\|_2 \leq C(1 + \kappa^{\frac{1}{2}})r_m^{-\sigma}\|f^\dagger\|_{H^\sigma(\mathcal{S},\mathcal{L})},$$

which is asymptotically sharp in view of Proposition 10. In our context, we require the possibility for $\mathcal{S}$ to be an orthonormal system but not necessarily a basis.

It is unclear whether the above result from [45] can be extended to a proper subset of a basis of eigenfunctions of the Laplace–Beltrami operator. For this reason, we shall employ a slight refinement of Gröchenig's result (Theorem 8), which yields a better constant, though it generally does not achieve the optimal asymptotic decay.

4.2. The construction of Marcinkiewicz-Zygmund families in Riemannian manifolds.
As a first and remarkable example, let us suppose that $\mathcal{M}$ is a connected compact orientable d-dimensional Riemannian manifold without boundary and with normalized Riemannian measure $d\mu$, such that $\mu(\mathcal{M}) = 1$. Let us consider the distinct eigenvalues $\{0 = \lambda_0^2 < \lambda_1^2 < \lambda_2^2 < \ldots\}$ of the Laplace-Beltrami operator and let $\mathcal{B} = \left\{\varphi_m^\ell\right\}_{m=0,\ell=1}^{+\infty,\delta_m}$ be a corresponding orthonormal basis of eigenfunctions, so that $\Delta\varphi_m^\ell = \lambda_m^2\varphi_m^\ell$. Set $\mathcal{E} = \{\lambda_m\}_{m=0,\ell=1}^{+\infty,\delta_m}$.

In this case, the space of diffusion polynomials of bandwidth $r \geq 0$ is

$$\mathcal{Q}_r(\mathcal{B}, \mathcal{E}) = \text{span}\left\{\varphi_m^\ell : \lambda_m \leq r\right\}.$$

By Weyl's estimates on the spectrum of an elliptic operator [35, Theorem 17.5.3 and Corollary 17.5.8],

$$\dim(\mathcal{Q}_r(\mathcal{B}, \mathcal{E})) \sim r^d \quad \text{and} \quad \sum_{m:\lambda_m \leq r} \sum_{\ell=1}^{\delta_m} |\varphi_m^\ell(x)|^2 \leq W r^d \tag{16}$$

for some $W > 0$.

We begin with a definition.

Definition 11. Let $0 < a \leq 1 \leq b$ and $0 < c_1 < c_2$. We say that a collection $\mathcal{R} = \{R_j\}_{j=1}^N$ of measurable subsets of $\mathcal{M}$ is a partition of $\mathcal{M}$ with parameters a, b, c_1 and c_2 if the following conditions hold:

- $\cup_{j=1}^N R_j = \mathcal{M}$ and $\mu(R_i \cap R_j) = 0$ for every $1 \leq i < j \leq N$;
- $a/N \leq \mu(R_j) \leq b/N$ for every $j = 1, \cdots N$;
- and each R_j is contained in a geodesic ball of radius $c_2 N^{-1/d}$ and contains a geodesic ball of radius $c_1 N^{-1/d}$.

Maggioni, Mhaskar and Filbir [21, 22, 47] have proven that any choice of N points, each one in a region R_j as above, produces a Marcinkiewicz-Zygmund family with weights equal to the measures of the regions.

Theorem 12 ([22, Theorem 5.1]). *Let $0 < a \leq 1 \leq b$ and $0 < c_1 < c_2$. Then, there exist positive constants C and D such that for all $N \geq 1$ and $r > 0$ such that $N \geq Cr^d$, for all partitions $\{R_j\}_{j=1}^N$ with parameters a, b, c_1 and c_2 as in Definition 11, for all $x_j \in R_j$, for all $q \in \mathcal{Q}_r(\mathcal{B}, \mathcal{E})$ it holds*

$$\left| \int_{\mathcal{M}} |q(x)|^2 \, d\mu(x) - \sum_{j=1}^N \mu(R_j) \, |q(x_j)|^2 \right| \leq DrN^{-1/d} \int_{\mathcal{M}} |q(x)|^2 \, d\mu(x).$$

In particular, for any $\varepsilon > 0$, if $N \geq \max\{C, D^d/\varepsilon^d\}r^d$, then

$$\left| \int_{\mathcal{M}} |q(x)|^2 \, d\mu(x) - \sum_{j=1}^N \mu(R_j) \, |q(x_j)|^2 \right| \leq \varepsilon \int_{\mathcal{M}} |q(x)|^2 \, d\mu(x),$$

that is

$$(1 - \varepsilon) \int_{\mathcal{M}} |q(x)|^2 \, d\mu(x) \leq \sum_{j=1}^N \mu(R_j) \, |q(x_j)|^2 \leq (1 + \varepsilon) \int_{\mathcal{M}} |q(x)|^2 \, d\mu(x).$$

It remains to show that partitions as required by the previous theorem do exist. Indeed the following theorem holds (see [19, 26, 27]).

Theorem 13. *There exist two constants $0 < c_3 < c_4$ such that for all constants a and b with $0 < a \leq 1 \leq b$, for every $N \geq 1$ and for every choice of weights $\{\tau_j\}_{j=1}^N$ with $\sum_{j=1}^N \tau_j = 1$ and $a/N \leq \tau_j \leq b/N$, there is a partition of $\mathcal{M}$, $\mathcal{R} = \{R_j\}_{j=1}^N$, with parameters a, b, $c_3(a^2/b)^{1/d}$ and $c_4 b^{1/d}$ such that $\mu(R_j) = \tau_j$ for all $j = 1, \ldots, N$.*

5. Sampling in inverse problems

In this section, we apply the previous results to inverse problems. More precisely, we consider the problem consisting of the recovery of a signal $f^\dagger$ from sampled measurements of $Ff^\dagger$, where F is a linear map.

We make the following assumptions. The first assumption concerns the measurement space Y for the inverse problem, which will be a Sobolev space on $\mathcal{M}$.

Assumption 14. Let $\mathcal{M}$ be a topological space and μ be a probability measure in $\mathcal{M}$. Let $\mathcal{B} = \{\varphi_k : k \in \mathbb{N}\}$ be an orthonormal basis in $L^2(\mathcal{M}, \mu)$, such that each φ_k is continuous and bounded on $\mathcal{M}$, and $\varphi_0 \equiv 1$. Let $\mathcal{E} = \{\lambda_k : \lambda_k \geq 0\}$ be a non-decreasing sequence with $\lambda_k \to +\infty$. Assume that there exists a Marcinkiewicz-Zygmund family $\mathcal{X}$ for $\mathcal{B}$ and $\mathcal{E}$, with nodes $\{x_{m,j} : m \in \mathbb{N}, j = 1, \ldots, L_m\} \subseteq \mathcal{M}$, weights $\{\tau_{m,j}\}$ and degrees $\{r_m\}$, such that

$$A\|q\|_2^2 \leq \sum_{j=1}^{L_m} |q(x_{m,j})|^2 \tau_{n,j} \leq B\|q\|_2^2 \qquad \text{for all } q \in \mathcal{Q}_{r_m}(\mathcal{B}, \mathcal{E}),$$

with constants A, $B > 0$ independent of m. Let $Y = H^\sigma(\mathcal{B}, \mathcal{E})$, for some $\sigma \geq 0$.

In the second assumption, we fix a Galerkin nested sequence of subspaces of X, the space where the unknown f belongs to.

Assumption 15. Let X be a Banach space and assume that X admits a multi-resolution structure, that is a family $\{G_m\}_{m \in \mathbb{N}}$ of subspaces of X such that

  (i) each G_m is finite-dimensional,
 (ii) $G_m \subseteq G_{m+1}$,
(iii) $\overline{\cup_{m=0}^{+\infty} G_m} = X$.

The last assumption quantifies the smoothing effect of F.

Assumption 16. Let $F \colon X \to Y$ be linear and bounded, and assume that

  (i) $F(G_m) \subseteq \mathcal{Q}_{r_m}(\mathcal{B}, \mathcal{E})$,
 (ii) $1 \in F(G_0)$,
(iii) $\dim \operatorname{Im} F = +\infty$.

We now build a new orthonormal system $\mathcal{S}$ of $L^2(\mathcal{M}, \mu)$ and a new non-decreasing sequence $\mathcal{L} = \{\Lambda_k : \Lambda_k \geq 0\}$ with $\Lambda_k \to +\infty$, adapted to the operator F. Let us begin with $\mathcal{S}_0 = \{\psi_k : 0 \leq k \leq d_0\}$, an orthonormal basis for $F(G_0)$, with $\psi_0 \equiv 1$. Being $F(G_0) \subseteq \mathcal{Q}_{r_0}(\mathcal{B}, \mathcal{E})$, for each k between 0 and d_0 we set $\Lambda_k = r_0$. Let us now complete (if necessary) $\mathcal{S}_0$ to an orthonormal basis $\mathcal{S}_1 = \{\psi_k : 0 \leq k \leq d_1\}$ for $F(G_1)$. Being $F(G_1) \subseteq \mathcal{Q}_{r_1}(\mathcal{B}, \mathcal{E})$, for each k between $d_0 + 1$ and d_1 (if any) we set $\Lambda_k = r_1$. Proceeding similarly, for any $m \geq 1$, we complete (if necessary) $\mathcal{S}_m$ to an orthonormal basis $\mathcal{S}_{m+1} = \{\psi_k : 0 \leq k \leq d_{m+1}\}$ for $F(G_{m+1})$. Being $F(G_{m+1}) \subseteq \mathcal{Q}_{r_{m+1}}(\mathcal{B}, \mathcal{E})$, for each k between $d_m + 1$ and d_{m+1} (if any) we set $\Lambda_k = r_{m+1}$. We call $\mathcal{S} = \{\psi_k : k \in \mathbb{N}\}$ and $\mathcal{L} = \{\Lambda_k : k \in \mathbb{N}\}$. Note that, thanks to Assumption 16 (iii), we have $\Lambda_k \to +\infty$.

Furthermore, $\mathcal{Q}_{r_m}(\mathcal{S}, \mathcal{L}) = \operatorname{span}\{\psi_k : \Lambda_k \leq r_m\} = \operatorname{span}\{\psi_k : 0 \leq k \leq d_m\} = F(G_m) \subseteq \mathcal{Q}_{r_m}(\mathcal{B}, \mathcal{E})$.

In the following result, we investigate the recovery of $Ff^\dagger$ from a finite number of noisy sampled measurements.

Corollary 17. *Let us suppose that Assumptions 14, 15, 16 hold and let $\sigma \geq \zeta \geq 0$. Let $f^\dagger \in X$ be such that $Ff^\dagger$ is continuous and let $\{y_{m,j}\}$ be such that*

$$|y_{m,j} - Ff^\dagger(x_{m,j})| \leq \beta \quad j = 1, \dots, L_m$$

where $\beta \geq 0$ is the noise level. Let

$$p_m^\beta \in \operatorname*{argmin}_{p \in G_m} \sum_{j=1}^{L_m} |y_{m,j} - Fp(x_{m,j})|^2 \tau_{m,j}.$$

Then

$$\|Ff^\dagger - Fp_m^\beta\|_{H^\zeta(\mathcal{S}, \mathcal{L})} \leq \sqrt{1 + \kappa}\|Ff^\dagger\|_{H^\sigma(\mathcal{S}, \mathcal{L})} \phi_{\sigma - \zeta}(r_m, \mathcal{S}, \mathcal{L}) + \sqrt{\kappa}\beta(1 + r_m^2)^{\frac{\zeta}{2}}.$$

Proof. Since $F(G_m) = \mathcal{Q}_{r_m}(\mathcal{S}, \mathcal{L})$, it follows that

$$Fp_m^\beta = \operatorname*{argmin}_{q \in \mathcal{Q}_{r_m}(\mathcal{S}, \mathcal{L})} \sum_{j=1}^{L_m} |y_{m,j} - q(x_{m,j})|^2 \tau_{m,j}.$$

Since $\mathcal{Q}_{r_m}(\mathcal{S}, \mathcal{L}) \subseteq \mathcal{Q}_{r_m}(\mathcal{B}, \mathcal{E})$ by Assumption 16, it follows that $\mathcal{X}$ is a Marcinkiewicz-Zygmund family also for $\mathcal{S}$ and $\mathcal{L}$ with the same constants A and B, and by Corollary 9, the thesis follows.   □

6. Convolutions on compact two-point homogeneous spaces

Here, we apply the results of Section 5 to a particular class of forward maps F, namely, convolutions on compact two-point homogeneous spaces. Two-point homogeneous spaces are introduced in §6.1, zonal functions and convolutions are described in §6.2 and, finally, the problem of sampling convolutions is discussed in §6.3.

6.1. **Two-point homogeneous spaces.** A d-dimensional Riemannian manifold $\mathcal{M}$ with distance ρ is said to be a two-point homogeneus space if given four points $x_1, x_2, y_1, y_2 \in \mathcal{M}$ such that $\rho(x_1, y_1) = \rho(x_2, y_2)$, then there exists an isometry g of $\mathcal{M}$ such that $gx_1 = x_2$ and $gy_1 = y_2$. Wang in [70] has completely characterized compact connected two-point homogeneous spaces. More precisely $\mathcal{M}$ is isometric to one of the following compact rank 1 symmetric spaces:

  i) the Euclidean sphere $S^d = SO(d+1)/SO(d) \times \{1\}$, $d \geqslant 1$;
  ii) the real projective space $P^n(\mathbb{R}) = O(n+1)/O(n) \times O(1)$, $n \geqslant 2$;
  iii) the complex projective space $P^n(\mathbb{C}) = U(n+1)/U(n) \times U(1)$, $n \geqslant 2$;
  iv) the quaternionic projective space $P^n(\mathbb{H}) = Sp(n+1)/Sp(n) \times Sp(1)$, $n \geqslant 2$;
  v) the octonionic projective plane $P^2(\mathbb{O})$.

In the following we will assume that $\mathcal{M}$ is one of the above symmetric spaces and that d is its real dimension. In particular the real dimension of $P^n(\mathbb{K})$ is $d = nd_0$, where $d_0 = 1, 2, 4, 8$ according to the real dimension of $\mathbb{K} = \mathbb{R}, \mathbb{C}, \mathbb{H}$ and $\mathbb{O}$ respectively. In the case of S^d it will be convenient to set $d_0 = d$. See [25, pp. 176-178], see also [31], [63] and [72]. In the following, to keep the notation simple, we will use

$$a = \frac{d-2}{2}, \qquad b = \frac{d_0 - 2}{2}. \tag{17}$$

Let μ be the Riemannian measure on $\mathcal{M}$ normalized so that $\mu(\mathcal{M}) = 1$. Let us assume that ρ is normalized so that $\operatorname{diam}(\mathcal{M}) = \pi$ and let $B_r(x) = \{y \in \mathcal{M} : \rho(x, y) < r\}$.

If o is a fixed point in $\mathcal{M}$, then $\mathcal{M}$ can be identified with the homogeneous space G/K, where G is the group of isometries of $\mathcal{M}$ and K is the stabilizer of o. We will also identify functions $f(x)$ on $\mathcal{M}$ with right K-invariant functions $\widetilde{f}(g)$ on G by setting $\widetilde{f}(g) = f(x)$ when $go = x$. The measure μ is invariant under the action of G; in other words, for every $g \in G$,

$$\int_{\mathcal{M}} f(gx) d\mu(x) = \int_{\mathcal{M}} f(x) d\mu(x).$$

Let dg and dk denote left Haar measures on G and K, respectively. These measures are normalized in such a way that $\int_G dg = 1$ and $\int_K dk = 1$. By [23, Theorem 2.51], we have

$$\int_G f(go)\, dg = \int_{\mathcal{M}} f(x) d\mu(x), \qquad f \in L^1(\mathcal{M}). \tag{18}$$

6.2. Zonal functions and convolutions.

Definition 18. A function f on $\mathcal{M}$ is a zonal function if for every $x \in \mathcal{M}$ and every $k \in K$ we have $f(kx) = f(x)$.

Lemma 19. *Let f be a zonal function. Then $f(x)$ depends only on $\rho(x, o)$. Furthermore, defining f_0 so that $f(x) = f_0(\rho(x, o))$ we have*

$$\int_{\mathcal{M}} f(x) d\mu(x) = \int_0^\pi f_0(r) A(r) dr, \tag{19}$$

where

$$A(r) = c(a,b) \left(\sin\left(\frac{r}{2}\right) \right)^{2a+1} \left(\cos\left(\frac{r}{2}\right) \right)^{2b+1}$$

and

$$c(a,b) = \left(\int_0^\pi \left(\sin\left(\frac{r}{2}\right) \right)^{2a+1} \left(\cos\left(\frac{r}{2}\right) \right)^{2b+1} dr \right)^{-1} = \frac{\Gamma(a+b+2)}{\Gamma(a+1)\Gamma(b+1)}.$$

Proof. Let $x, y \in \mathcal{M}$ be such that $\rho(x, o) = \rho(y, o)$. Since $\mathcal{M}$ is two-point homogeneous there exists $g \in G$ such that $gx = y$ and $go = o$. Thus, $g \in K$ and $f(y) = f(gx) = f(x)$. Equation (19) follows from (4.17) in [25]. $\qquad\square$

Definition 20. If f and h are two integrable functions on $\mathcal{M}$, then we define the convolution of h and f by

$$h * f(x) = \int_G h(g^{-1}x) f(go) dg.$$

The following lemma shows that this convolution with filter h coincides with the convolution with a zonal function $\widetilde{h}$.

Lemma 21. *Let $f, h \in L^1(\mathcal{M})$. Write*

$$\widetilde{h}(x) := \int_K h(kx)\, dk = \widetilde{h}_0(\rho(x, o)). \tag{20}$$

Then $\widetilde{h}$ is zonal and

$$h * f = \widetilde{h} * f.$$

Furthermore

$$\widetilde{h} * f(x) = \int_{\mathcal{M}} \widetilde{h}_0(\rho(x, y)) f(y) d\mu(y).$$

Proof. Take $x \in \mathcal{M}$ and write $x = uo$ with some $u \in G$. Recall that dg is also a right Haar measure on G. Observe that for any $k \in K$, by the change of variable $w = k^{-1}g^{-1}u$,

$$h * f(x) = \int_G h(g^{-1}uo)f(go)dg = \int_G h(kwo)f(uw^{-1}k^{-1}o)dw = \int_G h(kwo)f(uw^{-1}o)dw.$$

Integration in $k \in K$ gives

$$h * f(x) = \int_K \int_G h(kwo)f(uw^{-1}o)dwdk = \int_G f(uw^{-1}o) \int_K h(kwo)dkdw.$$

Notice that the function $\widetilde{h}$ is zonal on $\mathcal{M}$. It follows that

$$h * f(x) = \int_G f(uw^{-1}o)\widetilde{h}(wo)dw = \int_G \widetilde{h}(g^{-1}uo)f(go)dg = \widetilde{h} * f(x).$$

In other words, the convolution depends only on $\widetilde{h}$. Observe that

$$\widetilde{h} * f(x) = \int_G \widetilde{h}(g^{-1}x)f(go)dg = \int_G \widetilde{h}_0(\rho(g^{-1}x, o))f(go)dg$$

$$= \int_G \widetilde{h}_0(\rho(x, go))f(go)dg = \int_{\mathcal{M}} \widetilde{h}_0(\rho(x, y))f(y)d\mu(y),$$

where we used (18) in the last identity. $\square$

Let Δ be the Laplace-Beltrami operator on $\mathcal{M}$, let

$$0 = \lambda_0^2 < \lambda_1^2 < \lambda_2^2 < \dots,$$

be the distinct eigenvalues of Δ arranged in increasing order, let $\mathcal{H}_m$ be the eigenspace associated with the eigenvalue λ_m^2 and let δ_m be its dimension. In the case $m = 0$, we have $\mathcal{H}_0 = \text{span}\{1\}$ and $\delta_0 = 1$. It is well known that

$$L^2(\mathcal{M}) = \bigoplus_{m=0}^{+\infty} \mathcal{H}_m. \tag{21}$$

If $f(x) = f_0(\rho(x, o))$ is a zonal function on $\mathcal{M}$, then

$$\Delta f(x) = \frac{1}{A(t)} \frac{d}{dt} \left(A(t) \frac{d}{dt} f_0(t) \right) \Big|_{t = \rho(x, o)} \tag{22}$$

(see (4.16) in [25]).

Next, we study $\mathcal{H}_m$ as a reproducing kernel Hilbert space.

Definition 22. The zonal spherical function of degree $m \in \mathbb{N}$ with pole $x \in \mathcal{M}$ is the unique function $Z_x^m \in \mathcal{H}_m$, given by the Riesz representation theorem, such that for every $Y \in \mathcal{H}_m$

$$Y(x) = \int_{\mathcal{M}} Y(y)Z_x^m(y)d\mu(y).$$

The next lemma summarizes the main properties of zonal functions and is essentially taken from [66], where the case $\mathcal{M} = S^d$ is discussed in detail. For a proof of the general case, see [14, 13].

Lemma 23. *Take $m \in \mathbb{N}$ and $x, y \in \mathcal{M}$. We have that:*

(i) if $Y_m^1, \ldots, Y_m^{\delta_m}$ is an orthonormal basis of $\mathcal{H}_m \subset L^2(\mathcal{M})$, then

$$Z_x^m(y) = \sum_{\ell=1}^{\delta_m} \overline{Y_m^\ell(x)} Y_m^\ell(y);$$

(ii) the function Z_x^m is real valued and $Z_x^m(y) = Z_y^m(x)$;

(iii) if $g \in G$, then $Z_{gx}^m(gy) = Z_x^m(y)$;

(iv) for every $f \in L^2(\mathcal{M})$, $Z_o^m * f = \Pi_m f$ where $\Pi_m f$ is the orthogonal projection of f on $\mathcal{H}_m$;

(v) $Z_x^m(x) = \|Z_x^m\|_2^2 = \delta_m$;

(vi) Z_o^m is a zonal function and

$$Z_o^m(x) = \frac{\delta_m}{P_m^{a,b}(1)} P_m^{a,b}(\cos(\rho(x,o))) \tag{23}$$

where $P_m^{a,b}$ are the Jacobi polynomials;

(vii) $\{\delta_m^{-1/2} Z_o^m\}_{m=0}^{+\infty}$ is an orthonormal basis of the subspace of $L^2(\mathcal{M})$ of zonal functions;

(viii) $\lambda_m^2 = m(m+a+b+1)$;

(ix) and

$$\delta_m = (2m+a+b+1)\frac{\Gamma(b+1)}{\Gamma(a+1)\Gamma(a+b+2)}\frac{\Gamma(m+a+b+1)}{\Gamma(m+b+1)}\frac{\Gamma(m+a+1)}{\Gamma(m+1)} \sim m^{2a+1}$$

(when $m = 0$, the product $(2m+a+b+1)\Gamma(m+a+b+1)$ must be replaced by $\Gamma(a+b+2)$; this way, the case $a = b = -1/2$ can be included). In particular, for $m \geq 1$,

$$\delta_m \leq \frac{a+1}{b+1}(2m+a+b+1)m^{2a}. \tag{24}$$

Proof. Here we provide a proof only for (iv) and the last estimate of (ix), which are not given in the references above. By (i), (iii) and (18) we have that

$$Z_o^m * f(x) = \int_G Z_o^m(g^{-1}x)f(go)dg = \int_G Z_{go}^m(x)f(go)dg$$

$$= \int_G \sum_{\ell=1}^{\delta_m} \overline{Y_m^\ell(go)} Y_m^\ell(x)f(go)dg = \sum_{\ell=1}^{\delta_m} \left(\int_\mathcal{M} f(y)\overline{Y_m^\ell(y)}d\mu(y) \right) Y_m^\ell(x) = \Pi_m f(x),$$

as claimed.

Concerning (ix), inequality (24) is trivially satisfied for $a = b = -1/2$. In the remaining cases, $a \geq b \geq -1/2$ and $a \geq 0$. Furthermore, either $a + b$ or a are nonnegative integers. If $a + b$ is a nonnegative integer, then

$$\frac{\Gamma(m+a+b+1)}{\Gamma(m+1)}\frac{\Gamma(m+a+1)}{\Gamma(m+b+1)} = (m+a+b)\ldots(m+1)(m+a)\ldots(m+b+1)$$

$$\leq m^{a+b}m^{a-b}(1+a+b)\ldots(2)(1+a)\ldots(1+b+1)$$

$$= m^{2a}\frac{\Gamma(a+b+2)\Gamma(a+2)}{\Gamma(b+2)}.$$

Similarly, if a is a nonnegative integer, then

$$\frac{\Gamma(m+a+b+1)}{\Gamma(m+b+1)}\frac{\Gamma(m+a+1)}{\Gamma(m+1)} = (m+a+b)\dots(m+b+1)(m+a)\dots(m+1)$$

$$\leq m^a m^a (1+a+b)\dots(1+b+1)(1+a)\dots(2)$$

$$= m^{2a}\frac{\Gamma(a+b+2)\Gamma(a+2)}{\Gamma(b+2)}.$$

Thus, in any case,

$$\delta_m \leq (2m+a+b+1)\frac{\Gamma(b+1)}{\Gamma(a+1)\Gamma(a+b+2)}m^{2a}\frac{\Gamma(a+b+2)\Gamma(a+2)}{\Gamma(b+2)}$$

$$= \frac{a+1}{b+1}(2m+a+b+1)\,m^{2a}.$$

$\square$

Let f,h in $L^2(\mathcal{M})$. By Lemma 23 (vii) we have the following expansion of the zonal function $\widetilde{h}$ given in (20):

$$\widetilde{h} = \sum_{m=0}^{+\infty} b_m Z_o^m,$$

where

$$b_m = \frac{1}{\delta_m}\int_{\mathcal{M}}\widetilde{h}(y)Z_o^m(y)d\mu(y) = \frac{1}{\delta_m}\int_{\mathcal{M}}\int_K h(ky)Z_o^m(y)dkd\mu(y)$$

$$= \frac{1}{\delta_m}\int_K\int_{\mathcal{M}} h(x)Z_o^m(k^{-1}x)d\mu(x)dk = \frac{1}{\delta_m}\int_{\mathcal{M}} h(x)Z_o^m(x)d\mu(x).$$

Also, by (19) and (23), we have

$$b_m = \frac{1}{\delta_m}\int_{\mathcal{M}}\widetilde{h}(y)Z_o^m(y)d\mu(y) = \int_0^\pi \widetilde{h}_0(r)\frac{P_m^{a,b}(\cos r)}{P_m^{a,b}(1)}A(r)dr. \tag{25}$$

Furthermore, by Lemma 23 (iv) the convolution of h and f can be written as

$$h*f = \widetilde{h}*f = \sum_{m=0}^{+\infty} b_m Z_o^m * f = \sum_{m=0}^{+\infty} b_m \Pi_m f.$$

This means that the convolution with a function h can be seen as a multiplier operator.

Let us now set $\mathcal{B} = \{Y_m^\ell\}_{m=0,\ell=1}^{+\infty,\delta_m}$ and $\mathcal{E} = \{\sqrt{m(m+a+b+1)}\}_{m=0,\ell=1}^{+\infty,\delta_m}$.

Proposition 24. *Take $\omega,\gamma \geq 0$. Let $h \in L^2(\mathcal{M})$ such that $\widetilde{h} = \sum_{m=0}^{+\infty} b_m Z_o^m$, with*

$$|b_m| \leq c\big(1+\lambda_m^2\big)^{-\frac{\gamma}{2}}$$

for some $c > 0$, and $b_0 \neq 0$. Let $\sigma = \omega + \gamma$ and consider the linear bounded operator given by $Ff = h*f$. Then F maps continuously $H^\omega(\mathcal{B},\mathcal{E})$ into $H^\sigma(\mathcal{B},\mathcal{E})$, with norm bounded by c.

Proof. By the definition of the norm in H^σ (see (12)) and by Lemma 23, part (viii), we have

$$\|Ff\|_{H^\sigma}^2 = \|\widetilde{h} * f\|_{H^\sigma}^2 = \left\|\sum_{m=0}^{+\infty} b_m \Pi_m f\right\|_{H^\sigma}^2 = \sum_{m=0}^{+\infty} b_m^2 \|\Pi_m f\|_2^2 (1 + \lambda_m^2)^\sigma$$

$$\leq c^2 \sum_{m=0}^{+\infty} \|\Pi_m f\|_2^2 (1 + \lambda_m^2)^{\sigma - \gamma} = c^2 \|f\|_{H^\omega}^2. \qquad \square$$

6.3. **Sampling convolutions.** The focus of the following result is the reconstruction of $Ff^\dagger$ from some noisy sampled measurements, where F is a convolution operator.

Theorem 25. *Let $\mathcal{M}$ be a d-dimensional compact two-point homogeneous space as described above, $d \geq 2$. Assume that the nodes $\{x_{m,j}\}_{j=1}^{L_m}$, the weights $\{\tau_{m,j}\}_{j=1}^{L_m}$ and the sequence $r_m = \lambda_m$ are a Marcinkiewicz-Zygmund family for $\mathcal{B}$ and $\mathcal{E}$ with constants $1 - \varepsilon$ and $1 + \varepsilon$, with $\varepsilon > 0$.*

Let $h \in L^2(\mathcal{M})$ be such that $\widetilde{h}(x) = \sum_{m=0}^{+\infty} b_m Z_o^m(x)$, with $|b_m| \leq c(1 + \lambda_m^2)^{-\gamma/2}$ for some $c > 0$, $\gamma \geq 0$, and $b_m \neq 0$ for infinitely many values of m, including $m = 0$. Let $\omega \geq 0$, $\sigma = \omega + \gamma$ and consider the linear bounded operator $F \colon H^\omega(\mathcal{B}, \mathcal{E}) \to H^\sigma(\mathcal{B}, \mathcal{E})$ given by $Ff = h * f$. Let $\zeta \geq 0$, assume $\sigma - \zeta > d/2$ and take $f^\dagger \in H^\omega(\mathcal{B}, \mathcal{E})$. Let $\{y_{m,j}\}$ be a family of noisy outputs corresponding to $\{Ff^\dagger(x_{m,j})\}$,

$$|y_{m,j} - Ff^\dagger(x_{m,j})| \leq \beta \quad j = 1, \ldots, L_m$$

where $\beta \geq 0$ is the noise level. For every $m \in \mathbb{N}$ let

$$p_m^\beta \in \operatorname*{argmin}_{p \in \mathcal{Q}_{\lambda_m}(\mathcal{B}, \mathcal{E})} \sum_{j=1}^{L_m} |y_{m,j} - Fp(x_{m,j})|^2 \tau_{m,j}.$$

Then

$$\|Fp_m^\beta - Ff^\dagger\|_{H^\zeta(\mathcal{B}, \mathcal{E})} \leq \sqrt{1 + \kappa} \|Ff^\dagger\|_{H^\sigma(\mathcal{B}, \mathcal{E})} \phi_{\sigma - \zeta}(\lambda_m, \mathcal{B}, \mathcal{E}) + \sqrt{\kappa}\beta(1 + \lambda_m^2)^{\frac{\zeta}{2}}$$

$$\leq \sqrt{\frac{d(1 + \kappa)}{d_0(\sigma - \zeta - d/2)}} \frac{c}{(1 + \lambda_m^2)^{(\sigma - \zeta - d/2)/2}} \|f^\dagger\|_{H^\omega(\mathcal{B}, \mathcal{E})} + \sqrt{\kappa}\beta(1 + \lambda_m^2)^{\frac{\zeta}{2}}.$$

Recall that $\kappa = \frac{1 + \varepsilon}{1 - \varepsilon}$, and $d/d_0 = 1$ for the sphere, while $d/d_0 = n$ for the projective space $P^n(\mathbb{K})$.

The existence of a Marcinkiewicz-Zygmund family, as in the hypotheses of Theorem 25, is guaranteed by Theorems 12 and 13.

Proof. We want to apply Corollary 17. Assumption 14 is satisfied by hypothesis. In particular, notice that the space $\mathcal{Q}_{\lambda_m}(\mathcal{B}, \mathcal{E})$ coincides with the space $G_m := \bigoplus_{s=0}^m \mathcal{H}_s$ and Assumption 15 follows. Finally, since $F(G_m) \subseteq G_m$, $1 \in F(G_0)$ (since $G_0 = \mathcal{H}_0 = \operatorname{span}\{1\}$ and $b_0 \neq 0$), and $\dim \operatorname{Im} F = +\infty$ (since $b_m \neq 0$ for infinitely many values of m), then also Assumption 16 holds.

We may therefore apply Corollary 17, with $\mathcal{S} = \{Y_m^\ell : m = 0, \ldots, +\infty \text{ s.t. } b_m \neq 0, \ell = 1, \ldots, \delta_m\} \subseteq \mathcal{B}$ and $\mathcal{L} = \{\sqrt{m(m + a + b + 1)} : m = 0, \ldots, +\infty \text{ s.t. } b_m \neq 0, \ell = 1, \ldots, \delta_m\} \subseteq \mathcal{E}$ and we obtain

$$\|Ff^\dagger - Fp_m^\beta\|_{H^\zeta(\mathcal{S}, \mathcal{L})} \leq \sqrt{1 + \kappa} \|Ff^\dagger\|_{H^\sigma(\mathcal{S}, \mathcal{L})} \phi_{\sigma - \zeta}(\lambda_m, \mathcal{S}, \mathcal{L}) + \sqrt{\kappa}\beta(1 + \lambda_m^2)^{\frac{\zeta}{2}}.$$

Notice that $\|Ff^\dagger\|_{H^\sigma(\mathcal{S},\mathcal{L})} = \|Ff^\dagger\|_{H^\sigma(\mathcal{B},\mathcal{E})} < +\infty$ by Proposition 24 and that, since $\sigma > d/2 + \zeta \geq d/2$, then $Ff^\dagger \in H^\sigma(\mathcal{B},\mathcal{E})$ is continuous. Also,

$$\phi_{\sigma-\zeta}(\lambda_m, \mathcal{S}, \mathcal{L}) = \sup_{x \in \mathcal{M}} \left(\sum_{n>m,\ b_n \neq 0} \sum_{\ell=1}^{\delta_n} |Y_n^\ell(x)|^2 (1+\lambda_n^2)^{-(\sigma-\zeta)} \right)^{\frac{1}{2}}$$

$$\leq \sup_{x \in \mathcal{M}} \left(\sum_{n>m} \sum_{\ell=1}^{\delta_n} |Y_n^\ell(x)|^2 (1+\lambda_n^2)^{-(\sigma-\zeta)} \right)^{\frac{1}{2}} = \phi_{\sigma-\zeta}(\lambda_m, \mathcal{B}, \mathcal{E}).$$

Since $Z_x^n(x) = \delta_n$ by Lemma 23 (v), and since by (24)

$$\delta_n \leq \frac{a+1}{b+1}(2n+a+b+1)(1+\lambda_n^2)^a,$$

by Lemma 23 (i) and (17) it follows that

$$\phi_{\sigma-\zeta}(\lambda_m, \mathcal{B}, \mathcal{E})^2 = \sup_x \sum_{n>m} \sum_{\ell=1}^{\delta_n} |Y_n^\ell(x)|^2 (1+\lambda_n^2)^{-(\sigma-\zeta)}$$

$$= \sup_x \sum_{n>m} Z_x^n(x)(1+\lambda_n^2)^{-(\sigma-\zeta)}$$

$$\leq \frac{a+1}{b+1} \sum_{n>m} (2n+a+b+1)(1+\lambda_n^2)^{a+\zeta-\sigma}$$

$$\leq \frac{a+1}{b+1} \int_m^{+\infty} (2x+a+b+1)(1+x(x+a+b+1))^{a+\zeta-\sigma}\, dx$$

$$= \frac{a+1}{b+1} \int_{1+m(m+a+b+1)}^{+\infty} y^{a+\zeta-\sigma}\, dy$$

$$= \frac{a+1}{b+1} \frac{1}{\sigma-\zeta-a-1} \frac{1}{(1+m(m+a+b+1))^{\sigma-\zeta-a-1}}$$

$$= \frac{d}{d_0} \frac{1}{\sigma-\zeta-d/2} \frac{1}{(1+\lambda_m^2)^{\sigma-\zeta-\frac{d}{2}}}.$$

This concludes the proof. $\qquad\qquad\square$

Remark 26. It is clear from the above proof that the term $\phi_{\sigma-\zeta}(\lambda_m, \mathcal{B}, \mathcal{E})$ in the thesis of the theorem can be replaced by $\phi_{\sigma-\zeta}(\lambda_m, \mathcal{S}, \mathcal{L})$. Of course, when several terms of the sequence b_m vanish, then $\phi_{\sigma-\zeta}(\lambda_m, \mathcal{S}, \mathcal{L})$ may be asymptotically smaller than $\phi_{\sigma-\zeta}(\lambda_m, \mathcal{B}, \mathcal{E})$.

In the following result, we investigate the reconstruction of $f^\dagger$.

Corollary 27. *Under the hypotheses of Theorem 25, assume furthermore that $\zeta \geq \gamma$ and that, for some constant $c_0 > 0$,*

$$|b_m| \geq c_0(1+\lambda_m^2)^{-\zeta/2}, \text{ for all } m \geq 0.$$

Then

$$\|f^\dagger - p_m^\beta\|_2 \le c_0^{-1}\|Ff^\dagger - Fp_m^\beta\|_{H^\zeta(\mathcal{B},\mathcal{E})}$$

$$\le c_0^{-1}\sqrt{1+\kappa}\|Ff^\dagger\|_{H^\sigma(\mathcal{B},\mathcal{E})}\phi_{\sigma-\zeta}(\lambda_m,\mathcal{B},\mathcal{E}) + \sqrt{\kappa}\beta(1+\lambda_m^2)^{\frac{\zeta}{2}}$$

$$\le \frac{c}{c_0}\sqrt{\frac{d(1+\kappa)}{d_0(\sigma-\zeta-d/2)}}\frac{1}{(1+\lambda_m^2)^{(\sigma-\zeta-d/2)/2}}\|f^\dagger\|_{H^\omega(\mathcal{B},\mathcal{E})} + \sqrt{\kappa}\beta(1+\lambda_m^2)^{\frac{\zeta}{2}}.$$

Proof. Note that, since $b_m \ne 0$ for every m, we have that F is injective, with inverse given by

$$F^{-1}f = \sum_m b_m^{-1}\Pi_m f.$$

Thus

$$\|F^{-1}f^\dagger\|_2^2 = \sum_m |b_m|^{-2}\|\Pi_m f^\dagger\|_2^2 \le c_0^{-2}\sum_m(1+\lambda_m^2)^\zeta\|\Pi_m f^\dagger\|_2^2 = c_0^{-2}\|f^\dagger\|_{H^\zeta}^2.$$

Therefore,

$$\|f^\dagger - p_m^\beta\|_2 \le c_0^{-1}\|Ff^\dagger - Fp_m^\beta\|_{H^\zeta}.$$

Applying this to the estimates in Theorem 25, the thesis follows. $\square$

The following proposition provides an alternative description of Sobolev spaces and an explicit formula for computing the coefficients b_m.

Proposition 28. *Let K be a nonnegative integer. Then*
  (i) $H^{2K}(\mathcal{B},\mathcal{E}) = \{f \in L^2(\mathcal{M}): (-\Delta)^K f \in L^2(\mathcal{M})\}$;
  (ii) if $h = \sum_{m=0}^{+\infty} b_m Z_o^m \in H^{2K}(\mathcal{B},\mathcal{E})$, then

$$b_0 = \int_{\mathcal{M}} h(x)d\mu(x)$$

  and for $m \ge 1$

$$b_m = \frac{1}{\delta_m(m(m+a+b+1))^K}\int_{\mathcal{M}}(-\Delta)^K h(x)Z_o^m(x)d\mu(x);$$

(iii) for $m \ge 1$,

$$|b_m| \le \frac{\|(-\Delta)^K h\|_2}{\delta_m^{\frac{1}{2}}(m(m+a+b+1))^K} \le \frac{2^K\|(-\Delta)^K h\|_2}{\delta_m^{\frac{1}{2}}(1+m(m+a+b+1))^K}.$$

Proof. The equivalence in (i) is well known, and the rest follows from the definition of b_m as

$$b_m = \frac{1}{\delta_m}\int_{\mathcal{M}} h(x)Z_o^m(x)d\mu(x),$$

from the self-adjointness of Δ and the fact that $(-\Delta)^K Z_o^m = (m(m+a+b+1))^K Z_o^m$. $\square$

Remark 29. Notice that in the case of the sphere S^2, if $h \in H^{2K}(\mathcal{B}, \mathcal{E})$ for some integer K, then by Proposition 28, for $m \geq 1$

$$|b_m| \leq \frac{2^K \|(-\Delta)^K h\|_2}{(2m+1)^{\frac{1}{2}}(1+m(m+1))^K} \leq \frac{2^K \|(-\Delta)^K h\|_2}{(1+m(m+1))^{\frac{4K+1}{4}}}.$$

References

[1] Abhishake, T. Helin, and N. Mücke. Statistical inverse learning problems with random observations, pages 201–244. De Gruyter, Berlin, Boston, 2025.

[2] K. Abraham and R. Nickl. On statistical Calderón problems. Math. Stat. Learn., 2(2):165–216, 2019.

[3] B. Adcock, A. C. Hansen, E. Herrholz, and G. Teschke. Generalized sampling: extension to frames and inverse and ill-posed problems. Inverse Problems, 29(1):015008, 27, 2013.

[4] G. S. Alberti, A. Arroyo, and M. Santacesaria. Inverse problems on low-dimensional manifolds. Nonlinearity, 36(1):734–808, 2023.

[5] G. S. Alberti, A. Felisi, M. Santacesaria, and S. I. Trapasso. Compressed sensing for inverse problems and the sample complexity of the sparse radon transform. Journal of the European Mathematical Society, 2025.

[6] G. S. Alberti, A. Felisi, M. Santacesaria, and S. I. Trapasso. Compressed sensing for inverse problems ii: applications to deconvolution, source recovery, and mri. arXiv preprint arXiv:2501.01929, 2025.

[7] G. S. Alberti and M. Santacesaria. Infinite dimensional compressed sensing from anisotropic measurements and applications to inverse problems in PDE. Appl. Comput. Harmon. Anal., 50:105–146, 2021.

[8] G. S. Alberti and M. Santacesaria. Infinite-dimensional inverse problems with finite measurements. Arch. Ration. Mech. Anal., 243(1):1–31, 2022.

[9] J. J. Benedetto and P. J. S. G. Ferreira, editors. Modern sampling theory. Applied and Numerical Harmonic Analysis. Birkhäuser Boston, Inc., Boston, MA, 2001. Mathematics and applications.

[10] M. Bertero, P. Boccacci, and C. De Mol. Introduction to inverse problems in imaging. CRC press, 2021.

[11] D. Bilyk, M. Mastrianni, and S. Steinerberger. Single radius spherical cap discrepancy via gegenbadly approximable numbers. Adv. Math., 452:Paper No. 109812, 18, 2024.

[12] G. Blanchard and N. Mücke. Optimal rates for regularization of statistical inverse learning problems. Found. Comput. Math., 18(4):971–1013, 2018.

[13] L. Brandolini, B. Gariboldi, and G. Gigante. Irregularities of Distribution on Two-Point Homogeneous Spaces. In The Mathematical Heritage of Guido Weiss, Appl. Numer. Harmon. Anal., pages 101–125. Birkhäuser/Springer, Cham, 2025.

[14] L. Brandolini, B. Gariboldi, G. Gigante, and A. Monguzzi. Single radius spherical cap discrepancy on compact two-point homogeneous spaces. Math. Z., 308(3):Paper No. 52, 16, 2024.

[15] T. A. Bubba, M. Burger, T. Helin, and L. Ratti. Convex regularization in statistical inverse learning problems. Inverse Probl. Imaging, 17(6):1193–1225, 2023.

[16] T. A. Bubba and L. Ratti. Shearlet-based regularization in statistical inverse learning with an application to x-ray tomography. Inverse Problems, 38(5):Paper No. 054001, 43, 2022.

[17] M. Dashti, S. Harris, and A. Stuart. Besov priors for Bayesian inverse problems. Inverse Probl. Imaging, 6(2):183–200, 2012.

[18] A. Ebner and M. Haltmeier. Convergence rates for the joint solution of inverse problems with compressed sensing data. Inverse Problems, 39(1):Paper No. 015011, 16, 2023.

[19] M. Ehler, U. Etayo, B. Gariboldi, G. Gigante, and T. Peter. Asymptotically optimal cubature formulas on manifolds for prefixed weights. J. Approx. Theory, 271:Paper No. 105632, 24, 2021.

[20] H. W. Engl, M. Hanke, and A. Neubauer. Regularization of inverse problems, volume 375 of Mathematics and its Applications. Kluwer Academic Publishers Group, Dordrecht, 1996.

[21] F. Filbir and H. N. Mhaskar. A quadrature formula for diffusion polynomials corresponding to a generalized heat kernel. J. Fourier Anal. Appl., 16(5):629–657, 2010.

[22] F. Filbir and H. N. Mhaskar. Marcinkiewicz-Zygmund measures on manifolds. J. Complexity, 27(6):568–596, 2011.

[23] G. B. Folland. A course in abstract harmonic analysis. Textbooks in Mathematics. CRC Press, Boca Raton, FL, second edition, 2016.

[24] K.-J. Förster. Inequalities for ultraspherical polynomials and application to quadrature. In Proceedings of the Seventh Spanish Symposium on Orthogonal Polynomials and Applications (VII SPOA) (Granada, 1991), volume 49, pages 59–70, 1993.

[25] R. Gangolli. Positive definite kernels on homogeneous spaces and certain stochastic processes related to Lévy's Brownian motion of several parameters. Ann. Inst. H. Poincaré Sect. B (N.S.), 3:121–226, 1967.

[26] B. Gariboldi and G. Gigante. Optimal asymptotic bounds for designs on manifolds. Anal. PDE, 14(6):1701–1724, 2021.

[27] G. Gigante and P. Leopardi. Diameter bounded equal measure partitions of Ahlfors regular metric measure spaces. Discrete Comput. Geom., 57(2):419–430, 2017.

[28] M. Giordano and R. Nickl. Consistency of Bayesian inference with Gaussian process priors in an elliptic inverse problem. Inverse Problems, 36(8):085001, 35, 2020.

[29] K. Gröchenig. Sampling, Marcinkiewicz-Zygmund inequalities, approximation, and quadrature rules. J. Approx. Theory, 257:105455, 20, 2020.

[30] D. M. Healy, Jr., H. Hendriks, and P. T. Kim. Spherical deconvolution. J. Multivariate Anal., 67(1):1–22, 1998.

[31] S. Helgason. Groups and geometric analysis, volume 83 of Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 2000. Integral geometry, invariant differential operators, and spherical functions, Corrected reprint of the 1984 original.

[32] T. Helin. Least squares approximations in linear statistical inverse learning problems. SIAM J. Numer. Anal., 62(4):2025–2047, 2024.

[33] E. Herrholz and G. Teschke. Compressive sensing principles and iterative sparse recovery for inverse and ill-posed problems. Inverse Problems, 26(12):125012, 24, 2010.

[34] R. Hielscher and M. Quellmalz. Optimal mollifiers for spherical deconvolution. Inverse Problems, 31(8):085001, 28, 2015.

[35] L. Hörmander. The analysis of linear partial differential operators. III. Classics in Mathematics. Springer, Berlin, 2007. Pseudo-differential operators, Reprint of the 1994 edition.

[36] V. Isakov. Inverse problems for partial differential equations, volume 127 of Applied Mathematical Sciences. Springer, Cham, third edition, 2017.

[37] J. Kaipio and E. Somersalo. Statistical inverse problems: discretization, model reduction and inverse crimes. J. Comput. Appl. Math., 198(2):493–504, 2007.

[38] B. Kaltenbacher, A. Neubauer, and O. Scherzer. Iterative regularization methods for nonlinear ill-posed problems, volume 6 of Radon Series on Computational and Applied Mathematics. Walter de Gruyter GmbH & Co. KG, Berlin, 2008.

[39] H. Kekkonen, M. Lassas, and S. Siltanen. Analysis of regularized inversion of data corrupted by white Gaussian noise. Inverse Problems, 30(4):045009, 18, 2014.

[40] G. Kerkyacharian, T. M. Pham Ngoc, and D. Picard. Localized spherical deconvolution. Ann. Statist., 39(2):1042–1068, 2011.

[41] P. T. Kim and J.-Y. Koo. Optimal spherical deconvolution. J. Multivariate Anal., 80(1):21–42, 2002.

[42] H. J. Landau. Necessary density conditions for sampling and interpolation of certain entire functions. Acta Mathematica, 117:37–52, 1967.

[43] M. Lassas, E. Saksman, and S. Siltanen. Discretization-invariant Bayesian inversion and Besov space priors. Inverse Probl. Imaging, 3(1):87–122, 2009.

[44] D. Lazzaro, S. Morigi, and L. Ratti. Oracle-net for nonlinear compressed sensing in electrical impedance tomography reconstruction problems. J. Sci. Comput., 101(2):Paper No. 49, 29, 2024.

[45] J. Li, Y. Ling, J. Geng, and H. Wang. Weighted least ℓ_p approximation on compact Riemannian manifolds. J. Fourier Anal. Appl., 30(5):Paper No. 57, 29, 2024.

[46] W. Lu and H. Wang. Approximation and quadrature by weighted least squares polynomials on the sphere. Pure Appl. Funct. Anal., 8(2):565–581, 2023.

[47] M. Maggioni and H. N. Mhaskar. Diffusion polynomial frames on metric measure spaces. Appl. Comput. Harmon. Anal., 24(3):329–353, 2008.

[48] J. D. McEwen and Y. Wiaux. A novel sampling theorem on the sphere. IEEE Trans. Signal Process., 59(12):5876–5887, 2011.

[49] H. N. Mhaskar, F. J. Narcowich, and J. D. Ward. Spherical Marcinkiewicz-Zygmund inequalities and positive quadrature. Math. Comp., 70(235):1113–1130, 2001.

[50] F. c. Monard and P. Stefanov. Sampling the X-ray transform on simple surfaces. SIAM J. Math. Anal., 55(3):1707–1736, 2023.

[51] F. Natterer. The mathematics of computerized tomography, volume 32 of Classics in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001. Reprint of the 1986 original.

[52] R. Nickl. Bayesian non-linear statistical inverse problems. Zurich Lectures in Advanced Mathematics. EMS Press, Berlin, [2023] ©2023.

[53] R. Nickl and G. P. Paternain. On some information-theoretic aspects of non-linear statistical inverse problems. In ICM—International Congress of Mathematicians. Vol. 7. Sections 15–20, pages 5516–5538. EMS Press, Berlin, [2023] ©2023.

[54] H. Nyquist. Certain topics in telegraph transmission theory. Transactions of the American Institute of Electrical Engineers, 47(2):617–644, 1928.

[55] R. Penrose. On best approximation solutions of linear matrix equations. Proc. Cambridge Philos. Soc., 52:17–19, 1956.

[56] E. A. Rakhmanov, E. B. Saff, and Y. M. Zhou. Minimal discrete energy on the sphere. Math. Res. Lett., 1(6):647–662, 1994.

[57] R. Ramlau and O. Scherzer, editors. The Radon transform—the first 100 years and beyond, volume 22 of Radon Series on Computational and Applied Mathematics. Walter de Gruyter GmbH & Co. KG, Berlin, [2019] ©2019.

[58] A. Rastogi, G. Blanchard, and P. Mathé. Convergence analysis of Tikhonov regularization for non-linear statistical inverse problems. Electron. J. Stat., 14(2):2798–2841, 2020.

[59] O. Scherzer, editor. Handbook of mathematical methods in imaging. Vol. 1, 2, 3. Springer, New York, second edition, 2015.

[60] C. E. Shannon. Communication in the presence of noise. Proceedings of the IRE, 37(1):10–21, 1949.

[61] M. Simeoni. Functional penalised basis pursuit on spheres. Appl. Comput. Harmon. Anal., 53:1–53, 2021.

[62] M. M. J.-A. Simeoni. Functional Inverse Problems on Spheres: Theory, Algorithms and Applications. PhD thesis, EPFL, Lausanne, 2020.

[63] M. M. Skriganov. Point distributions in two-point homogeneous spaces. Mathematika, 65(3):557–587, 2019.

[64] P. Stefanov. Semiclassical sampling and discretization of certain linear inverse problems. SIAM J. Math. Anal., 52(6):5554–5597, 2020.

[65] P. Stefanov and S. Tindel. Sampling linear inverse problems with noise. Asymptot. Anal., 132(3-4):331–382, 2023.

[66] E. M. Stein and G. Weiss. Introduction to Fourier analysis on Euclidean spaces, volume No. 32 of Princeton Mathematical Series. Princeton University Press, Princeton, NJ, 1971.

[67] G. Szegő. Orthogonal polynomials, volume Vol. XXIII of American Mathematical Society Colloquium Publications. American Mathematical Society, Providence, RI, fourth edition, 1975.

[68] A. C. M. van Rooij and F. H. Ruymgaart. Regularized deconvolution on the circle and the sphere. In Nonparametric functional estimation and related topics (Spetses, 1990), volume 335 of NATO Adv. Sci. Inst. Ser. C: Math. Phys. Sci., pages 679–690. Kluwer Acad. Publ., Dordrecht, 1991.

[69] T. Vareschi. Application of second generation wavelets to blind spherical deconvolution. J. Multivariate Anal., 124:398–417, 2014.

[70] H.-C. Wang. Two-point homogeneous spaces. Ann. of Math. (2), 55:177–191, 1952.

[71] E. T. Whittaker. On the functions which are represented by the expansions of the interpolation theory. Proceedings of the Royal Society of Edinburgh, 35:181–194, 1915.

[72] J. A. Wolf. Spaces of constant curvature. University of California, Department of Mathematics, Berkeley, CA, second edition, 1972.

[1]MaLGa Center, Department of Mathematics, University of Genoa, Via Dodecaneso 35, 16146, Genova, Italy

Email address: giovanni.alberti@unige.it

Email address: ernesto.devito@unige.it

[2]Dipartimento di Ingegneria Gestionale, dell'Informazione e della Produzione, Università degli Studi di Bergamo, Viale G. Marconi 5, 24044, Dalmine BG, Italy

Email address: biancamaria.gariboldi@unibg.it

Email address: giacomo.gigante@unibg.it