Sampling Theory, Signal Processing, and Data Analysis (2024) 22:19
https://doi.org/10.1007/s43670-024-00093-x

ORIGINAL ARTICLE

Generalizing geometric nonwindowed scattering transforms on compact Riemannian manifolds

Albert Chua[1] · **Yang Yang**[2]

Received: 20 February 2024 / Accepted: 19 October 2024 / Published online: 11 November 2024

Abstract

We study nonwindowed geometric scattering transforms on compact Riemannian manifolds without boundary. These transforms are formulated as $\mathbf{L}^q$-norms, with $1 < q \leq 2$, of a cascade of geometric wavelet transforms and modulus operators. We provide weighted measures for these operators, and prove that these operators are well-defined under specific conditions on the manifold, invariant to the action of isometries, and stable to diffeomorphisms for bandlimited functions.

Keywords Wavelet scattering · Geometric deep learning · Spectral geometry · Riemannian manifolds

1 Introduction

In recent years, deep convolutional neural networks have shown strong performance on various vision-related tasks [25, 37, 39]. However, because of how complex deep convolutional architectures are, it is not entirely clear what mechanisms enable deep convolutional networks to get strong performance on these tasks. In an effort to better understand the properties of deep convolutional architectures, Mallat proposed the scattering transform [6, 28], which uses a cascade of specified filters and nonlinearities to mimic the behavior of a deep convolutional neural network. Using a specific class of wavelet filters, Mallat found that the scattering transform had many desirable properties for machine learning tasks such as translation invariance and stability to small deformations. Additionally, the scattering transform and its generalizations have been found various applications as a general feature extractor, such as in [1–4, 7, 8,

✉ Albert Chua
chuaalbe@msu.edu

Yang Yang
yangy5@msu.edu

1 Department of Mathematics, Michigan State University, Wells Hall, 619 Red Cedar Rd, East Lansing, MI 48824, USA

2 Department of Computational Mathematics, Science and Engineering, Michigan State University, 428 South Shaw Lane, East Lansing, MI 48824, USA

17, 26, 29, 30, 35, 36, 38]. Authors have also explored extensions of the scattering transform to semi-discrete frames [27, 42, 43] and more general Gabor frames [15, 16].

However, certain forms of data, such as point cloud data have a geometry than is non-Euclidean, which motivate manifold learning models [12, 40, 41] and geometric deep learning [5]. As an extension of the scattering transform, researchers have considered graph scattering transforms in [18–20, 32, 44] constructed via graph wavelets [12, 24]; additionally [33, 34] extend wavelets and the scattering transform to simplicial complexes; the overarching idea is that these extensions of the scattering transform have similar desirable stability properties, and present success on non-Euclidean datasets.

The scattering transform has also been extended to compact Riemannian manifolds, known as the geometric scattering transform, via defining the wavelet transform on compact Riemannian manifolds using eigenfunctions of the Laplace–Beltrami operator. In particular, for the windowed geometric scattering transform, it is proved that the representation was locally invariant to isometries for all square-integrable functions and stable to diffeomorphic deformations under mild restrictions [31]. The concept of nonwindowed geometric scattering transform is also proposed in [31], which is proved invariant to isometries and stable to diffeomorphic deformations.

Regarding nonwindowed geometric scattering transforms, their Euclidean counterparts, nonwindowed scattering transforms, have been effective for applications in quantum chemistry, audio synthesis, and physics have appeared in [1, 8, 17, 26, 38]; theoretical results involving stability to deformations have also been provided in [11]. The main idea behind nonwindowed scattering transforms is that they provide a small number of descriptive features for high dimensional data. As a natural extension, [10] used scattering moments on manifolds for classifications tasks involving point cloud data. However, there were limited theoretical results in [9] for scattering moments provided in [10], which motivates this paper.

The major contribution of this paper is as follows:

- We provide a well-defined weighted measure for scattering moments for $\mathbf{L}^q(\mathcal{M})$-functions with $q \in (1, 2]$. Here $\mathbf{L}^q(\mathcal{M})$ denotes the space of q-integrable functions on a compact Riemannian manifold without boundary $\mathcal{M}$. Compared to the work [31], our results provide theoretical justification for nonwindowed scattering moments with $q \in (1, 2)$. From a practical perspective, the extension allows dealing with real data sets sampled from a wider class of functions. However, our weighted measure is only defined for an aribtrary, finite number of layers, and requires restrictions on the regularity of the manifold.
- We show the measure also has a diffeomorphism stability result for bandlimited functions, similar to [31]. This covers a wide class of functions, but is not fully general like in [11, 28]. A diffeomorphism stability result for non-bandlimited functions is left to future work.

2 A review of the geometric scattering transform on manifolds

2.1 Notation

We introduce some notations for the purposes of this paper. Henceforth, $\mathcal{M}$ will denote a compact, smooth, n-dimensional Riemannian manifold without boundary contained in $\mathbb{R}^d$, where $d \geq n$, with geodesic distance between two points $x_1, x_2 \in \mathcal{M}$ given by $r(x_1, x_2)$ and Laplace–Beltrami operator denoted as Δ. The notation $\mathbf{L}^q(\mathcal{M})$ denotes the set of all functions $f : \mathcal{M} \to \mathbb{R}$ such that $\int_{\mathcal{M}} |f(x)|^q \, d\mu(x) < \infty$, where $d\mu(x)$ is integration with respect to the Riemannian volume, whose measure is given by μ. We use the notation $\mathrm{Isom}(\mathcal{M}_1, \mathcal{M}_2)$ be the set of isometries between manifolds $\mathcal{M}_1$ and $\mathcal{M}_2$. Lastly, the set of diffeomorphisms on $\mathcal{M}$ will be denoted by $\mathrm{Diff}(\mathcal{M})$, and the maximum placement of $\gamma \in \mathrm{Diff}(\mathcal{M})$ will be given by $\|\gamma\|_\infty := \sup_{x \in \mathcal{M}} r(x, \gamma(x))$.

2.2 Spectral filters and the geometric wavelet transform

The convolution of two compactly-supported functions $f, g \in \mathbf{L}^2(\mathbb{R}^n)$ is usually defined in space as

$$(f * g)(x) = \int_{\mathbb{R}^n} f(y) g(x - y) \, dy.$$

However, for a general manifold, even under the conditions we have prescribed, a notation of translation does not necessarily exist, To motivate our definition below, consider $f, g \in L^2([0, 1])$. A basis is given by $\{c_n(x)\}_{n \geq 0}$. Via a fourier series expansion, we have

$$(f * g)(x) = \sum_{n \geq 0} \langle (f * g), \overline{e}_n \rangle e_n(x).$$

Based on the intuition above, one can create a spectral definition of convolution via the spectral decomposition of $-\Delta$. Denote $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. Because our manifold is compact, it is well known that $-\Delta$ has a discrete spectrum, and we can order the eigenvalues in increasing order and denote them as $\{\lambda_n\}_{n \in \mathbb{N}_0}$. We will denote the corresponding eigenfunctions as $\{e_n(x)\}_{n \in \mathbb{N}_0}$, which form an orthonormal basis for $\mathbf{L}^2(\mathcal{M})$.

Suppose $f \in \mathbf{L}^2(\mathcal{M})$. Since the set of functions $\{e_n(x)\}_{n \in \mathbb{N}_0}$ forms a basis in $\mathbf{L}^2(\mathcal{M})$, we decompose

$$f(x) = \sum_{n \in \mathbb{N}_0} \langle f, e_n \rangle e_n(x) = \sum_{n \in \mathbb{N}_0} \left(\int_{\mathcal{M}} f(y) \overline{e_n(y)} \, d\mu(y) \right) e_n(x), \tag{1}$$

which is similar to a Fourier series. Since $e_n(y)$ is a replacement for a Fourier mode, it is natural to let

$$\hat{f}(n) = \int_{\mathcal{M}} f(y) \overline{e_n(y)} \, d\mu(y) \tag{2}$$

and define convolution on $\mathcal{M}$ between functions $f, h \in \mathbf{L}^2(\mathcal{M})$ as

$$f * h(x) = \sum_{n \in \mathbb{N}_0} \hat{f}(n)\hat{h}(n)e_n(x). \tag{3}$$

Defining the operator $T_h f(x) := f * h(x)$, it is easy to verify that

$$T_h f(x) = \int_{\mathcal{M}} \tilde{K}_h(x, y) f(y)\, d\mu(y), \quad \tilde{K}_h(x, y) := \sum_{n \in \mathbb{N}_0} \hat{h}(n) e_n(x)\overline{e_n(y)}. \tag{4}$$

Similar to how convolution commutes with translations on $\mathbb{R}^n$, it is important for convolution on $\mathcal{M}$ to be equivariant to a group action on $\mathcal{M}$. The authors of [31] construct an operator by convolving with functions that commute with isometries since the the geometry of $\mathcal{M}$ should be preserved by a representation. To accomplish this goal, we use a similar definition for spectral filters. A filter $h \in \mathbf{L}^2(\mathcal{M})$ is a *spectral filter* if $\lambda_k = \lambda_\ell$ implies $\hat{h}(k) = \hat{h}(\ell)$. One can prove that there exists $H : [0, \infty) \to \mathbb{R}$ such that

$$H(\lambda_n) = \hat{h}(n), \qquad \forall n \in \mathbb{N}_0.$$

Let $G : [0, \infty) \to \mathbb{R}$ be nonnegative and decreasing with $G(0) > 0$. A low-pass spectral filter ϕ is given in frequency as $\hat{\phi}(n) := G(\lambda_n)$ and its dilation at scale 2^j for $j \in \mathbb{Z}$ is $\hat{\phi}_j(n) := G(2^j \lambda_n)$. Using the set of low-pass filters, $\{\hat{\phi}_j\}_{j \in \mathbb{Z}}$, we define *wavelets* by

$$\hat{\psi}_j(n) := [|\hat{\phi}_{j-1}(n)|^2 - |\hat{\phi}_j(n)|^2]^{1/2} = [|G(2^{j-1}n)|^2 - |G(2^j n)|^2]^{1/2}, \tag{5}$$

Note that these are wavelets modeled after the wavelets from [13].

Fix $J \in \mathbb{Z}$. Define the operators

$$A_J f := f * \phi_J,$$
$$\Psi_j f := f * \psi_j, \qquad j \leq J.$$

The *windowed geometric wavelet transform* is given by

$$W_J f := \{A_J f, \Psi_j f : \quad j \leq J\} \tag{6}$$

and the *nonwindowed geometric wavelet transform* is given by

$$W f := \{\Psi_j f : \quad j \in \mathbb{Z}\}. \tag{7}$$

We have the following theorem, which provides a condition for when our wavelet frame is a nonexpansive frame.

Theorem 1 *Let $G : [0,\infty) \to \mathbb{R}$ be nonnegative, decreasing, and continuous with $0 < G(0) = C$, $\lim_{x\to\infty} G(x) = 0$, and $\{\psi_j\}_{j\in\mathbb{Z}}$ is a set of wavelets generated by using the low-pass filter $\hat{\phi}(k) = G(\lambda_k)$ in Eq. 5. Then we have*

$$\sum_{j\in\mathbb{Z}} \|f * \psi_j\|_2^2 = C^2 \|f\|_2^2. \tag{8}$$

Proof For fixed $I, J > 1$, we telescope to get

$$\begin{aligned}\sum_{j=-I}^{J} |\hat{\psi}_j(k)|^2 &= \sum_{j=-I}^{J} \left[|G(2^{j-1}\lambda_k)|^2 - |G(2^j\lambda_k)|^2\right] \\ &= |G(2^{-I-1}\lambda_k)|^2 - |G(2^J\lambda_k)|^2.\end{aligned}$$

Since $\lim_{I\to\infty} |G(2^{-I-1}\lambda_k)|^2 = C^2$ and $\lim_{J\to\infty} |G(2^J\lambda_k)|^2 = 0$ by the assumption on G, it follows that

$$\sum_{j\in\mathbb{Z}} |\hat{\psi}_j(k)|^2 = \lim_{I\to\infty} |G(2^{-I-1}\lambda_k)|^2 - \lim_{J\to\infty} |G(2^J\lambda_k)|^2 = C^2.$$

We can write

$$\|f * \psi_j\|_2^2 = \sum_{k\in\mathbb{N}_0} |\hat{\psi}_j(k)|^2 |\hat{f}(k)|^2.$$

Thus, it follows that

$$\begin{aligned}\sum_{j\in\mathbb{Z}} \|f * \psi_j\|_2^2 &= \sum_{j\in\mathbb{Z}} \sum_{k\in\mathbb{N}_0} |\hat{f}(k)|^2 |\hat{\psi}_j(k)|^2 \\ &= \sum_{k\in\mathbb{N}_0} |\hat{f}(k)|^2 \left(\sum_{j\in\mathbb{Z}} |\hat{\psi}_j(k)|^2\right) \\ &= C^2 \|f\|_2^2.\end{aligned}$$

□

In the case where we choose $G(\lambda) = e^{-\lambda}$, which corresponds to our kernel being the heat kernel on $\mathcal{M}$, we have $C = 1$ and the relation above is an isometry.

2.3 The geometric scattering transform

Here, we recall the geometric scattering transform as defined in [31], which is a geometric analog to the Euclidean definition of the scattering transform. This transform is useful as it provides a representation that meaningfully encodes high frequency information of a signal f. Define the propagator as

$$U[j]f := |\Psi_j f| \qquad \forall j \in \mathbb{Z}, \tag{9}$$

which is convolution of a wavelet and applying a nonlinearity; we can also define the windowed propagator as

$$U_J[j]f := |\Psi_j f| \qquad \forall j \leq J. \tag{10}$$

Similar to scattering transforms on Euclidean space, one can apply a cascade of convolutions and modulus operators repeatedly. In particular, for $m \in \mathbb{N}$, let $j_1, \ldots, j_m \in \mathbb{Z}$. The m-layer propagator is defined as

$$U[j_1, \ldots, j_m]f := U[j_m] \cdots U[j_1]f = |||f * \psi_{j_1}| * \psi_{j_2}| \cdots * \psi_{j_m}| \tag{11}$$

and the m-layer windowed propagator is defined as

$$U[j_1, \ldots, j_m]f := U[j_m] \cdots U[j_1]f = |||f*\psi_{j_1}|*\psi_{j_2}| \cdots *\psi_{j_m}|, \quad j_1, \ldots, j_m \leq J \tag{12}$$

with $U[\emptyset]f = f$ and $U_J[\emptyset]f = f$. To aggregate low frequency information and get local isometry invariance, one can apply a low-pass filter in a manner similar to pooling to each windowed propagator to get windowed scattering coefficients:

$$S_J[j_1, \ldots, j_m] = A_J U_J[j_1, \ldots, j_m]f = U_J[j_1, \ldots, j_m]f * \phi_J,$$

where we defined $S_J[\emptyset]f = f * \phi_J$. The *windowed geometric scattering transform* is given by

$$S_J f = \{S_j[j_1, \ldots, j_m]f : m \geq 0, \quad j_i \leq J, \quad \forall 1 \leq i \leq m\}. \tag{13}$$

It is proved that this windowed scattering operator was nonexpansive, invariant to isometries up to the scale of the low-pass filter, and stable to diffeomporhisms under mild assumptions [31].

On the other hand, for applications such as manifold classification, one desires full isometry invariance instead of isometry invariance up to scale 2^J. We see that

$$\lim_{J \to \infty} S[j_1, \ldots, j_m]f = \mathrm{vol}(\mathcal{M})^{-1/2} \|U[j_1, \ldots, j_m]f\|_1. \tag{14}$$

As a proxy, one can consider

$$\overline{S}f[j_1, \ldots, j_m] = \|U[j_1, \ldots, j_m]f\|_1. \tag{15}$$

This motivates defining the *nonwindowed geometric scattering transform* [31] as

$$\overline{S}f = \{\overline{S}[j_1, \ldots, j_m]f : m \geq 0, \quad j_i \in \mathbb{Z}, \quad \forall 1 \leq i \leq m\}. \tag{16}$$

In this paper, we will extend the domain of the nonwindowed geometric scattering transform: Instead of considering $\mathbf{L}^1(\mathcal{M})$ norms of m-layer propagators, we will instead consider $\mathbf{L}^q(\mathcal{M})$ norms of m-layer propagators for $q \in (1, 2]$, which we define as m-layer q-nonwindowed geometric scattering coefficients (which are also referred to as scattering moments in other works):

$$\overline{S}_q^m[j_1, \ldots, j_m]f = \|U[j_1, \ldots, j_m]f\|_{\mathbf{L}^q(\mathcal{M})} \qquad \forall (j_1, \ldots, j_m) \in \mathbb{Z}^m, \tag{17}$$

which has seen application in quantum chemistry [17, 26, 38] and for point cloud data [10]. As shorthand notation, we will use the following notation for one layer coefficients:

$$\overline{S}_q[j]f = \|U[j]f\|_{\mathbf{L}^q(\mathcal{M})} \qquad \forall j \in \mathbb{Z}. \tag{18}$$

To measure stability, invariance, and equivariance, we define the following norm for q-nonwindowed geometric scattering coefficients:

$$\|\overline{S}_q^m f\|_{\ell^2(\mathbb{Z}^m)}^q := \left(\sum_{j_m \in \mathbb{Z}} \cdots \sum_{j_1 \in \mathbb{Z}} |\overline{S}_q^m[j_1, \ldots, j_m]f|^2 \right)^{q/2}, \tag{19}$$

which follows the definition in [11]. Since many of the results in [11] rely on results of Littlewood Paley theory, we will provide extensions of these results to compact manifolds, with some extra restrictions added.

3 Some results related to littlewood paley theory

Denote by $\ell^2(\mathbb{Z})$ the space of square-summable sequences indexed by integers, that is,

$$\ell^2(\mathbb{Z}) := \{(\ldots, a_{-2}, a_{-1}, a_0, a_1, a_2, \ldots) : a_j \in \mathbb{C} \text{ for each } j \in \mathbb{Z}, \sum_{j=-\infty}^{\infty} |a_j|^2 < \infty\}.$$

We define some $\ell^2(\mathbb{Z})$-valued function spaces to be used later.

For $1 \leq p < \infty$, define the $\mathbf{L}^p_{\ell^2(\mathbb{Z})}(\mathcal{M})$-norm as

$$\|g\|^p_{\mathbf{L}^p_{\ell^2(\mathbb{Z})}(\mathcal{M})} := \int_{\mathcal{M}} \|g(x)\|^p_{\ell^2(\mathbb{Z})} \, d\mu(x) \, .$$

and the $\mathbf{L}^{p,\infty}_{\ell^2(\mathbb{Z})}(\mathcal{M})$-norm as

$$\|g\|_{\mathbf{L}^{p,\infty}_{\ell^2(\mathbb{Z})}(\mathcal{M})} := \sup_{\delta > 0} \delta \cdot \mu(\{x \in \mathcal{M} : \|g(x)\|_{\ell^2(\mathbb{Z})} > \delta\})^{1/p} \, ,$$

The spaces $\mathbf{L}^p_{\ell^2(\mathbb{Z})}(\mathcal{M})$ and $\mathbf{L}^{p,\infty}_{\ell^2(\mathbb{Z})}(\mathcal{M})$ consist of $\ell^2(\mathbb{Z})$-valued functions that have finite $\mathbf{L}^p_{\ell^2(\mathbb{Z})}(\mathcal{M})$-norm and $\mathbf{L}^{p,\infty}_{\ell^2(\mathbb{Z})}(\mathcal{M})$-norm, respectively. It is clear that for any $\delta > 0$,

$$\begin{aligned}\delta^p \cdot \mu(\{x \in \mathcal{M} : \|g(x)\|_{\ell^2(\mathbb{Z})} > \delta\}) &= \int_{\{x\in\mathcal{M}:\|g(x)\|_{\ell^2(\mathbb{Z})}>\delta\}} \delta^p \, d\mu(x) \\ &\leq \int_{\{x\in\mathcal{M}:\|g(x)\|_{\ell^2(\mathbb{Z})}>\delta\}} \|g(x)\|^p_{\ell^2(\mathbb{Z})} \, d\mu(x) \\ &\leq \int_{\mathcal{M}} \|g(x)\|^p_{\ell^2(\mathbb{Z})} \, d\mu(x).\end{aligned}$$

Taking the supremum in $\delta > 0$ concludes

$$\|g\|_{\mathbf{L}^{p,\infty}_{\ell^2(\mathbb{Z})}(\mathcal{M})} \leq \|g\|_{\mathbf{L}^p_{\ell^2(\mathbb{Z})}(\mathcal{M})}.$$

Hence $\mathbf{L}^p_{\ell^2(\mathbb{Z})}(\mathcal{M})$ embeds continuously into $\mathbf{L}^{p,\infty}_{\ell^2(\mathbb{Z})}(\mathcal{M})$.

Recall that we have defined the operator $Wf : \mathcal{M} \to \ell^2(\mathbb{Z})$ given by

$$W(f)(x) := \{\Psi_j f(x)\}_{j\in\mathbb{Z}} = \left\{\int_{\mathcal{M}} K_{2^{-j/2}}(x, y) f(y) \, d\mu(y)\right\}_{j\in\mathbb{Z}}.$$

with kernel given by $\vec{K} = \{K_{2^{-j/2}}\}_{j\in\mathbb{Z}}$ associated to the wavelets $\{\psi_j\}_{j\in\mathbb{Z}}$ generated using a low-pass filter $G \in \mathcal{S}(\mathbb{R}^+)$ satisfying the conditions of Theorem 1 for Eq. 5. Since

$$\|W(f)(x)\|_{\ell^2(\mathbb{Z})} = \left(\sum_{j\in\mathbb{Z}} |f * \psi_j(x)|^2\right)^{1/2},$$

we conclude from Theorem 1 that

$$\|Wf\|_{\mathbf{L}^2_{\ell^2(\mathbb{Z})}(\mathcal{M})} = C\|f\|_{\mathbf{L}^2(\mathcal{M})}.$$

Our goal is to extend this operator and prove that for all $q \in (1, 2)$, there exists C_q such that

$$\|Wf\|_{\mathbf{L}^q_{\ell^2(\mathbb{Z})}(\mathcal{M})} \leq C_q\|f\|_{\mathbf{L}^q(\mathcal{M})},$$

Before providing any proofs, we will state preliminary lemmas that will be vital to our approach. The first few lemmas concern the kernel of our convolution operator.

Lemma 2 ([23, Corollary 2.2]) *Let n be the dimension of $\mathcal{M}$. Suppose that $F \in \mathcal{S}(\mathbb{R}^+)$, the space of Schwartz functions (all functions with rapidly decreasing derivatives of all orders), restricted to the nonnegative real axis and $F(0) = 0$. For the kernel*

$$K_t(x, y) = \sum_{n \in \mathbb{N}} F(t^2 \lambda_n) e_n(x) \overline{e_n(y)},$$

the following pointwise bound holds for some $C_n > 0$ and for all $t > 0$ and all $x, y \in \mathcal{M}$:

$$|K_t(x, y)| \leq \frac{C_n t^{-n}}{\left(1 + \frac{r(x,y)}{t}\right)^{n+1}}.$$

The next result is proved in [22, Proposition 3.1]. We include its proof (following the idea for [22, Proposition 3.1]) to demonstrate that the constants c_1, c_2 can be chosen to depend continuously on the metric g. Such continuous dependence is subsequently necessary for the remark preceding Example 1. Recall that given an open cover $\mathcal{U}$ on the manifold $\mathcal{M}$, a number $\gamma > 0$ is called a *Lebesgue number* for $\mathcal{U}$ if for all $x \in \mathcal{M}$, there exists $\mathcal{U}_x \in \mathcal{U}$ such that $B(x, \gamma) \subset \mathcal{U}_x$, where we define

$$B(x, r_0) := \{y \in X \, : \, r(x, y) < r_0\}. \tag{20}$$

Lemma 3 ([22, Proposition 3.1]) *Cover $\mathcal{M}$ with a finite collection of open sets P_i with $1 \leq i \leq I$ such that the following properties hold for each index i:*

1. *there exists a chart (V_i, φ_i) with $\overline{P}_i \subset V_i$ ($\overline{P}_i$ denotes the closure of P_i)*
2. *$\varphi_i(P_i)$ is a ball in $\mathbb{R}^n$.*

Choose $\delta > 0$ so that 3δ is a Lebesgue number for the covering $\{P_i\}$. Then there exist $c_1, c_2 > 0$ such that for any $x \in \mathcal{M}$ and any $B(x, 3\delta) \subset P_i$, the following statements hold in the coordinate system on P_i obtained from φ_i:

1. *For all $y, z \in P_i$, we have $r(y, z) \leq c_2 |y - z|$.*
2. *For all $y, z \in B(x, \delta)$, we have $r(y, z) \geq c_1 |y - z|$.*

Here, $|\cdot|$ denotes the Euclidean norm induced by the φ_i-coordinates.

Proof This proof follows the idea for [22, Proposition 3.1] yet with explicitly identified constants c_1, c_2 to demonstrate their continuous dependence on the metric g.

Using the coordinates given by φ_i, we identify $g(x)$ with an $n \times n$ smooth matrix for each $x \in \overline{P}_i$. We write $|\cdot|_g$ for the norm induced by g, and $|\cdot|$ for the Euclidean norm in the φ_i-coordinates. For each $x \in \overline{P}_i$ and a tangent vector $v_x \in T_x\mathcal{M}$, we have

$$\underline{\Lambda}(x) |v_x|^2 \leq |v_x|^2_{g(x)} = \sum_{\alpha,\beta=1}^{n} v_x^\alpha g_{\alpha\beta}(x) v_x^\beta \leq \overline{\Lambda}(x) |v_x|^2.$$

Here $\underline{\Lambda}(x)$, $\overline{\Lambda}(x)$ denote the minimum and maximum eigenvalues of the matrix $g(x)$ respectively. We conclude

$$\left(\min_{x\in\overline{P}_i}\underline{\Lambda}(x)\right)|v_x|^2 \le |v_x|^2_{g(x)} \le \left(\max_{x\in\overline{P}_i}\overline{\Lambda}(x)\right)|v_x|^2$$

uniformly for $(x, v) \in T\overline{P}_i$. As both $\underline{\Lambda}(x)$ and $\overline{\Lambda}(x)$ are positive continuous functions of x on the compact subset $\overline{P}_i$, their minimum and maximum are strictly positive.

For the first statement, take $y, z \in P_i$ and define $\gamma : [0, 1] \to P_i$, $\gamma(t) := tz + (1-t)y$ where the coordinates of y, z are given by φ_i. Then γ is a smooth curve connecting y, z and $|\gamma'(t)| = |y - z|$ for all t. By the definition of the Riemannian distance $r(y, z)$, we have

$$\begin{aligned} r(y, z) &\le \text{ length of } \gamma = \int_0^1 |\gamma'(t)|_{g(\gamma(t))}\, dt \\ &\le \left(\max_{x\in\overline{P}_i}\overline{\Lambda}(x)\right)^{\frac{1}{2}} \int_0^1 |\gamma'(t)|\, dt = \left(\max_{x\in\overline{P}_i}\overline{\Lambda}(x)\right)^{\frac{1}{2}} |y - z|. \end{aligned}$$

One choice for c_2 is

$$c_2 := \max_{i=1,\ldots,I}\left(\max_{x\in\overline{P}_i}\overline{\Lambda}(x)\right)^{\frac{1}{2}}.$$

For the second statement, take $y, z \in B(x, \delta)$ and let $\gamma_k : [0, 1] \to \mathcal{M}$ be a sequence of piecewise C^1 curves connecting y, z such that their lengths $\ell(\gamma_k) \to r(y, z)$ as $k \to \infty$. For large k, we have

$$r(\gamma_k(t), x) \le r(\gamma_k(t), y) + r(y, x) \le r(z, y) + \delta \le r(z, x) + r(x, y) + \delta \le 3\delta$$

for all $t \in [0, 1]$, hence $\gamma_k \subset P_i$. Therefore, we have for large k that

$$\begin{aligned} \ell(\gamma_k) &= \int_0^1 |\gamma_k'(t)|_g\, dt \ge \left(\min_{x\in\overline{P}_i}\underline{\Lambda}(x)\right)^{\frac{1}{2}} \int_0^1 |\gamma_k'(t)|\, dt \\ &\ge \left(\min_{x\in\overline{P}_i}\underline{\Lambda}(x)\right)^{\frac{1}{2}} \left|\int_0^1 \gamma_k'(t)\, dt\right| \ge \left(\min_{x\in\overline{P}_i}\underline{\Lambda}(x)\right)^{\frac{1}{2}} |y - z|. \end{aligned}$$

Letting $k \to \infty$ proves the statement, and one choice for c_1 is

$$c_1 := \min_{i=1,\ldots,I}\left(\min_{x\in\overline{P}_i}\underline{\Lambda}(x)\right)^{\frac{1}{2}}.$$

□

For the rest of this paper, we fix the collections $\{P_i\}$, $\{V_i\}$, $\{\varphi_i\}$, and constants δ, c_1, c_2 from the previous Lemma.

Lemma 4 *Suppose that $r(y,z) < \min\left\{\frac{1}{2}r(x,y), \delta\right\}$ so that y and z lie on the same ball of the covering. Assume that there exist c_1 and c_2 from Lemma 3 with $c_1c_2 < 2$. Then there exists a constant C_δ such that*

$$|K_t(x,y) - K_t(x,z)| \leq C_\delta \frac{r(y,z)t^{-n-1}}{\left(1 + \frac{r(x,y)}{t}\right)^{n+1}}.$$

Proof Using the proof of Theorem 5.5 in [22], for each $x \in \mathcal{M}$, there exists a point w_x on the segment connecting y to z such that

$$|K_t(x,y) - K_t(x,z)| \leq C_1 \frac{r(y,z)t^{-n-1}}{\left(1 + \frac{r(x,w_x)}{t}\right)^{n+1}}. \tag{21}$$

Now notice that triangle inequality implies that

$$r(x,y) \leq r(x,w_x) + r(y,w_x).$$

By Lemma 3, since w_x lies on the line segment between y and z, we see that

$$r(y,w_x) \leq c_2|y - w_x| \leq c_2|y - z| \leq c_1c_2r(y,z).$$

It follows that

$$r(x,y) - c_1c_2r(y,z) \leq r(x,w_x).$$

Since $c_1c_2 < 2$ and $r(x,y) \geq 2r(y,z)$, we see that

$$r(x,y) - \frac{c_1c_2}{2}r(x,y) \leq r(x,w_x)$$

so that $1 - \frac{c_1c_2}{2} > 0$. This leads to $r(x,y) \leq C_r r(x,w_x)$ for some constant C_r independent of x. Finally, we can make a replacement in the right side of (21) to get

$$|K_t(x,y) - K_t(x,z)| \leq C_\delta \frac{r(y,z)t^{-n-1}}{\left(1 + \frac{r(x,y)}{t}\right)^{n+1}}.$$

□

Now we provide the necessary tools from classical harmonic analysis for extension. The idea is similar to the proof in the Euclidean case; we wish to prove a weak-type (1, 1) bound and extend by interpolation. Let X be a set, β a metric, and μ a measure on X such that $0 < \mu(B(x,r)) < \infty$ for all $x \in X$ and $r > 0$. We say that a space (X, β, μ) is of homogeneous type, though often with the metric and measure omitted when implied, if for all $x \in X$ and $r > 0$ there exists a constant C_D such that

$$\mu(B(x, 2r)) \leq C_D \mu(B(x, r)), \tag{22}$$

where $B(x, r)$ is a ball of radius r centered at x for (X, β). The property above is also known as the doubling property. It is well known that a C^∞ compact Riemannian manifold using the standard Riemannian metric and volume is of homogeneous type.

The first result we will need is a Calderon-Zygmund decomposition:

Theorem 5 ([14], Corollary 2.3) *Suppose that X is a space of homogeneous type. Suppose that $f \in \mathbf{L}^1(X)$ and choose $\alpha > 0$ such that $\alpha^{-1}\|f\|_1 < \mu(X)$. Then we can decompose $f := g + b$ such that*

$$\|g\|^2_{\mathbf{L}^2(\mathcal{M})} \leq C_1 \alpha \|f\|_{\mathbf{L}^1(\mathcal{M})},$$
$$b = \sum_i b_i,$$

where $C_1 > 0$ is a constant, supp$(b_i) \subset B(x_i, r_i)$ for some countable collection of balls $\{B(x_i, r_i)\}$, and each b_i satisfies

$$\int_X b_i(x) d\mu(x) = 0,$$
$$\|b_i\|_1 \leq C\alpha\mu(B(x_i, r_i)),$$
$$\sum_i \mu(B(x_i, r_i)) \leq C\alpha^{-1}\|f\|_{\mathbf{L}^1(\mathcal{M})}.$$

Theorem 6 *Suppose that we choose wavelets $\{\psi_j\}_{j\in\mathbb{Z}}$ generated using $G \in \mathcal{S}(\mathbb{R}^+)$ in Eq. 5 that satisfy the conditions of Theorem 1 and $c_1c_2 < 2$ in Lemma 3. The following weak (1, 1) bound holds for some $A > 0$:*

$$\|Wf\|_{\mathbf{L}^{1,\infty}_{\ell^2(\mathbb{Z})}(\mathcal{M})} \leq A\|f\|_{\mathbf{L}^1(\mathcal{M})}.$$

Proof First, for any α such that $\alpha^{-1}\|f\|_{\mathbf{L}^1(\mathcal{M})} > \mu(\mathcal{M})$, we see that

$$\mu(\{x \in \mathcal{M} : \|Wf(x)\|_{\ell^2(\mathbb{Z})} > \alpha\}) \leq \mu(\mathcal{M}) \leq \alpha^{-1}\|f\|_{\mathbf{L}^1(\mathcal{M})}.$$

Now, we consider the case where $\alpha^{-1}\|f\|_{\mathbf{L}^1(\mathcal{M})} \leq \mu(\mathcal{M})$. We use our Caulderon-Zygmund decomposition and write $f = g + b$. It follows that

$$\begin{aligned}\mu(\{x\in\mathcal{M}:\|Wf(x)\|_{\ell^2(\mathbb{Z})}>\alpha\}) &\le \mu(\{x\in\mathcal{M}:\|Wg(x)\|_{\ell^2(\mathbb{Z})}>\alpha/2\})\\ &\quad+\mu(\{x\in\mathcal{M}:\|Wb(x)\|_{\ell^2(\mathbb{Z})}>\alpha/2\})\\ &:= I_1+I_2.\end{aligned}$$

For I_1, we apply Chebyshev inequality, $\mathbf{L}^2_{\ell^2(\mathbb{Z})}(\mathcal{M})$ boundedness of W, and our assumption on g to find that

$$\begin{aligned}\mu(\{x\in\mathcal{M}:\|Wg(x)\|_{\ell^2(\mathbb{Z})}>\alpha/2\}) &\le \frac{4}{\alpha^2}\|Wg\|^2_{\mathbf{L}^2_{\ell^2(\mathbb{Z})}(\mathcal{M})}\\ &\le \frac{4}{\alpha^2}\|W\|^2_{\mathbf{L}^2(\mathcal{M})\to\mathbf{L}^2_{\ell^2(\mathbb{Z})}(\mathcal{M})}\|g\|^2_{\mathbf{L}^2(\mathcal{M})}\\ &\le \frac{4C_1}{\alpha}\|W\|^2_{\mathbf{L}^2(\mathcal{M})\to\mathbf{L}^2_{\ell^2(\mathbb{Z})}(\mathcal{M})}\|f\|_{\mathbf{L}^1(\mathcal{M})}.\end{aligned}$$

For I_2, let $B=\bigcup_i B(x_i,2r_i)$. Then it follows that

$$\begin{aligned}I_2 &\le \mu(B)+\mu(\{x\in B^c:\|Wb(x)\|_{\ell^2(\mathbb{Z})}>\alpha/2\}\\ &\le \frac{C_D}{\alpha}\|f\|_{\mathbf{L}^1(\mathcal{M})}+\frac{2}{\alpha}\|Wb\|_{\mathbf{L}^1_{\ell^2(\mathbb{Z})}(B^c)}\\ &\le \frac{C_D}{\alpha}\|f\|_{\mathbf{L}^1(\mathcal{M})}+\frac{2}{\alpha}\sum_i\|Wb_i\|_{\mathbf{L}^1_{\ell^2(\mathbb{Z})}(B(x_i,2r_i)^c)},\end{aligned}$$

where the constant C_D comes from the fact that our measure has the doubling property.

To estimate $\|Wb_i\|_{\mathbf{L}^1_{\ell^2(\mathbb{Z})}(B(x_i,2r_i)^c)}$, we notice that

$$\begin{aligned}\|Wb_i\|_{\mathbf{L}^1_{\ell^2(\mathbb{Z})}(B(x_i,2r_i)^c)} &= \int_{B(x_i,2r_i)^c}\|Wb_i(x)\|_{\ell^2(\mathbb{Z})}\,d\mu(x)\\ &= \int_{B(x_i,2r_i)^c}\left\|\int_{B(x_i,r_i)}\vec{K}(x,y)b_i(y)\,d\mu(y)\right\|_{\ell^2(\mathbb{Z})}d\mu(x).\end{aligned}$$

Now, since each of the functions b_i integrate to zero over the balls $B(x_i,r_i)$,

$$\begin{aligned}&\int_{B(x_i,2r_i)^c}\left\|\int_{B(x_i,r_i)}\vec{K}(x,y)b_i(y)\,d\mu(y)\right\|_{\ell^2(\mathbb{Z})}d\mu(x)\\ &\quad=\int_{B(x_i,2r_i)^c}\left\|\int_{B(x_i,r_i)}(\vec{K}(x,y)-\vec{K}(x,x_i))b_i(y)\,d\mu(y)\right\|_{\ell^2(\mathbb{Z})}d\mu(x).\end{aligned}$$

Examining

$$\left\|\int_{B(x_i,r_i)}(\vec{K}(x,y)-\vec{K}(x,x_i))b_i(y)\,d\mu(y)\right\|_{\ell^2(\mathbb{Z})}.$$

more closely, we have

$$\begin{aligned}
&\left\| \int_{B(x_i,r_i)} (\vec{K}(x,y) - \vec{K}(x,x_i)) b_i(y)\, d\mu(y) \right\|_{\ell^2(\mathbb{Z})} \\
&= \left(\sum_{j\in\mathbb{Z}} \left| \int_{B(x_i,r_i)} (\vec{K}(x,y) - \vec{K}(x,x_i)) b_i(y)\, d\mu(y) \right|^2 \right)^{1/2} \\
&\leq \sum_{j\in\mathbb{Z}} \int_{B(x_i,r_i)} |\vec{K}(x,y) - \vec{K}(x,x_i)|\, |b_i(y)|\, d\mu(y).
\end{aligned}$$

Thus, after an application of Fubini's theorem, we have

$$\begin{aligned}
&\int_{B(x_i,2r_i)^c} \left\| \int_{B(x_i,r_i)} (\vec{K}(x,y) - \vec{K}(x,x_i)) b_i(y)\, d\mu(y) \right\|_{\ell^2(\mathbb{Z})} d\mu(x) \\
&\leq \int_{B(x_i,r_i)} |b_i(y)| \int_{B(x_i,2r_i)^c} \sum_{j\in\mathbb{Z}} \left| K_{2^{-j/2}}(x,y) - K_{2^{-j/2}}(x,x_i) \right| \, d\mu(x)\, d\mu(y) \\
&= \int_{B(x_i,r_i)} |b_i(y)| \left(\sum_{j\in\mathbb{Z}} \int_{B(x_i,2r_i)^c} \left| K_{2^{-j/2}}(x,y) - K_{2^{-j/2}}(x,x_i) \right| \, d\mu(x) \right) d\mu(y),
\end{aligned}$$

where $\vec{K}$ is the kernel defined on page 7. Now we consider the term inside the parentheses. We will break this argument into cases. First, consider if $2r_i \geq \delta$. We see that

$$\begin{aligned}
&\sum_{j\in\mathbb{Z}} \int_{B(x_i,2r_i)^c} |K_{2^{-j/2}}(x,y) - K_{2^{-j/2}}(x,x_i)|\, d\mu(x) \\
&\leq C \sum_{j\in\mathbb{Z}} \int_{B(x_i,2r_i)^c} \frac{2^{nj/2}}{\left(1 + 2^{j/2} r(x,x_i)\right)^{n+1}} + \frac{2^{nj/2}}{\left(1 + 2^{j/2} r(x,y)\right)^{n+1}}\, d\mu(x)
\end{aligned}$$

where the inequality follows from Lemma 2.

Now, since x_i is the center of $B(x_i, r_i)$, if $x \in B(x_i, 2r_i)^c$, then $r(x, x_i) \geq 2r_i \geq \delta$. Similarly, since $y \in B(x_i, r_i)$, it follows that $r(y, x_i) < r_i$ and we have $2r(y, x_i) \leq r(x, x_i)$. Apply triangle inequality to get

$$r(x,x_i) \leq r(x,y) + r(y,x_i) \leq r(x,y) + r_i \leq r(x,y) + \frac{1}{2} r(x,x_i),$$

which means that $r(x, x_i) \leq 2r(x, y)$.

Going back to the integral, there exists C_1 such that

$$\begin{aligned}
&\sum_{j\in\mathbb{Z}}\int_{B(x_i,2r_i)^c}\frac{2^{nj/2}}{\left(1+2^{j/2}r(x,y)\right)^{n+1}}+\frac{2^{nj/2}}{\left(1+2^{j/2}r(x,x_i)\right)^{n+1}}\,d\mu(x)\\
&\quad\le C_1\sum_{j\in\mathbb{Z}}\int_{r(x,x_i)\ge 2r(y,x_i)}\frac{2^{nj/2}}{\left(1+2^{j}r(x,x_i)\right)^{n+1}}\,d\mu(x)\\
&\quad= C_1\sum_{j\ge 0}\int_{r(x,x_i)\ge 2r(y,x_i)}\frac{2^{nj/2}}{\left(1+2^{j/2}r(x,x_i)\right)^{n+1}}\,d\mu(x)\\
&\qquad+C_1\sum_{j<0}\int_{r(x,x_i)\ge 2r(y,x_i)}\frac{2^{nj/2}}{\left(1+2^{j/2}r(x,x_i)\right)^{n+1}}\,d\mu(x)\\
&\quad:= J_1+J_2.
\end{aligned}$$

For J_1, since $r(x,x_i)>\delta$,

$$\begin{aligned}
&\sum_{j\ge 0}\int_{r(x,x_i)\ge 2r(y,x_i)}\frac{2^{nj/2}}{\left(1+2^{j/2}r(x,x_i)\right)^{n+1}}\,d\mu(x)\\
&\quad\le\sum_{j\ge 0}\int_{r(x,x_i)\ge 2r(y,x_i)}\frac{2^{nj/2}}{(2^{j/2}r(x,x_i))^{n+1}}\,d\mu(x)\\
&\quad=\sum_{j\ge 0}2^{-j/2}\int_{r(x,x_i)\ge 2r(y,x_i)}r(x,x_i)^{-n-1}\,d\mu(x)\\
&\quad<\infty.
\end{aligned}$$

For J_2, it is routine to see that

$$\sum_{j<0}\int_{r(x,x_i)\ge 2r(y,x_i)}\frac{2^{nj/2}}{\left(1+2^{j/2}r(x,x_i)\right)^{n+1}}\,d\mu(x)\le\sum_{j<0}2^{nj/2}\mu(\mathcal{M})<\infty.$$

Now we consider the case where $2r_i<\delta$. In this case, we see that $r(y,x_i)<r_i<\delta$, and we still have $2r(y,x_i)<r(x,x_i)$. Thus, the bound

$$|K_{2^{-j/2}}(x,y)-K_{2^{-j/2}}(x,x_i)|\le C\frac{2^{nj/2}}{\left(1+2^{j/2}r(x,x_i)\right)^{n+1}}$$

still applies. We can also apply Lemma 4 to get

$$|K_{2^{-j/2}}(x,y)-K_{2^{-j/2}}(x,x_i)|\le C_\delta\frac{r(y,x_i)2^{(n+1)j/2}}{\left(1+2^{j/2}r(x,x_i)\right)^{n+1}}.$$

Taking the geometric mean, for any $s \in [0, 1]$, we have

$$|K_{2^{-j/2}}(x, y) - K_{2^{-j/2}}(x, x_i)| \leq C_2 \frac{2^{nj/2}(2^{j/2}r(y, x_i))^s}{\left(1 + 2^{j/2}r(x, x_i)\right)^{n+1}}$$

for some constant C_2. It now follows that for $C_3 = \max\{C_\delta, C_2\}$, we have

$$\begin{aligned}
&\sum_{j\in\mathbb{Z}} |K_{2^{-j/2}}(x, y) - K_{2^{-j/2}}(x, x_i)| \\
&\leq \sum_{2^{j/2} < \frac{2}{r(x,x_i)}} |K_{2^{-j/2}}(x, y) - K_{2^{-j/2}}(x, x_i)| + \sum_{2^{j/2} \geq \frac{2}{r(x,x_i)}} |K_{2^{-j/2}}(x, y) - K_{2^{-j/2}}(x, x_i)| \\
&\leq C_\delta \sum_{2^{j/2} < \frac{2}{r(x,x_i)}} \frac{r(y, x_i)2^{(n+1)j/2}}{\left(1 + 2^{j/2}r(x, x_i)\right)^{n+1}} + C_2 \sum_{2^{j/2} \geq \frac{2}{r(x,x_i)}} \frac{2^{nj/2}(2^{j/2}r(y, x_i))^{1/2}}{\left(1 + 2^{j/2}r(x, x_i)\right)^{n+1}} \\
&\leq C_3 \left(r(x_i, y) \sum_{2^{j/2} < \frac{2}{r(x,x_i)}} 2^{(n+1)j/2} + r(x_i, y)^{1/2} \sum_{2^{j/2} \geq \frac{2}{r(x,x_i)}} \frac{2^{(n+1/2)j/2}}{(2^{j/2}r(x, x_i))^{n+1}} \right) \\
&\leq C_4(r(x_i, y)r(x, x_i)^{-n-1} + r(x_i, y)^{1/2}r(x, x_i)^{-n-1/2}).
\end{aligned}$$

Integrating over $2r(x_i, y) \leq r(x, x_i)$ yields a constant independent of r_i. It now follows that

$$\|Wb_i\|_{\mathbf{L}^1_{\ell^2(\mathbb{Z})}(B(x_i,r_i)^c)} \leq C_5 \|b_i\|_{\mathbf{L}^1(\mathcal{M})}$$

for some constant C_5. Using the Calderón-Zygmund decomposition,

$$I_2 \leq \left(\frac{C_D}{\alpha} + \frac{2C_5}{\alpha}\right) \|f\|_{\mathbf{L}^1(\mathcal{M})}.$$

□

Recall the following result, which is a vector-valued version of Marcinkiewicz Interpolation:

Lemma 7 ([21], Theorem 1.18) *Let $\mathcal{A}_1, \mathcal{A}_2$ be Banach spaces, $T : \mathcal{A}_1 \to \mathcal{A}_2$ be quasilinear on $\mathbf{L}^{p_0}_{\mathcal{A}_1}(X)$ and $\mathbf{L}^{p_1}_{\mathcal{A}_1}(X)$ with $0 < p_0 < p_1$. If T satisfies*

$$\|Tf\|_{\mathbf{L}^{p_i,\infty}_{\mathcal{A}_2}(X)} \leq M_i \|f\|_{\mathbf{L}^{p_i}_{\mathcal{A}_1}(X)}$$

for $i = 0, 1$, then

$$\|Tf\|_{\mathbf{L}^p_{\mathcal{A}_2(X)}} \leq N_p \|f\|_{\mathbf{L}^p_{\mathcal{A}_1}(X)} \qquad \forall p \in (p_0, p_1),$$

where N_p is dependent on p.

The following corollary is a direct result of interpolation now:

Corollary 8 *Suppose that we choose wavelets $\{\psi_j\}_{j\in\mathbb{Z}}$ generated by using $G \in \mathcal{S}(\mathbb{R}^+)$ in Eq. 5, G satisfies the conditions of Theorem 1, and $c_1c_2 < 2$. We have*

$$\|Wf\|^q_{\mathbf{L}^q_{\ell^2(\mathbb{Z})}(\mathcal{M})} \leq C_q \|f\|^q_{\mathbf{L}^q(\mathcal{M})}$$

for some constant $C_q > 0$, where $q \in (1, 2)$.

By duality, the result of Corollary 8 actually holds for $q \in (1, \infty)$. However, for generalizing the nonwindowed geometric scattering transform, we only need results for $q \in (1, 2)$ since $\mathbf{L}^2(\mathcal{M}) \subset \mathbf{L}^q(\mathcal{M})$. For $q > 2$, since our manifold is compact, we have $\mathbf{L}^q(\mathcal{M}) \subset \mathbf{L}^2(\mathcal{M})$, so previous results in [31] are applicable, and further theoretical analysis is not as significant.

Additionally, although the result of Corollary 8 seems restrictive because one needs $c_1c_2 < 2$, the result applies for a variety of different manifolds. If one finds a metric where the condition above holds, a class of metrics can be found by perturbing the metric. This is because the choice of the constants c_1, c_2 in the proof of Lemma 3 depend continuously on the metric g. Thus if $c_1c_2 < 2$ for g, the same strict inequality holds for all metrics that are sufficiently close to g. We provide a simple example below where the conditions of Lemma 3 hold. The result of the example below can also be extended to n-torii without much difficulty.

Example 1 Consider $\mathcal{M} = \mathbb{S}^1$, the unit circle that is embedded in $\mathbb{R}^2$, along with the charts:

$$\begin{aligned}
V_1 &:= \{(x_1, x_2) : x_1^2 + x_2^2 = 1,\ x_1 > 0\}, \qquad \varphi_1 : V_1 \to \mathbb{R},\ (x_1, x_2) \mapsto x_2 \\
V_2 &:= \{(x_1, x_2) : x_1^2 + x_2^2 = 1,\ x_2 > 0\}, \qquad \varphi_2 : V_2 \to \mathbb{R},\ (x_1, x_2) \mapsto x_1 \\
V_3 &:= \{(x_1, x_2) : x_1^2 + x_2^2 = 1,\ x_1 < 0\}, \qquad \varphi_3 : V_3 \to \mathbb{R},\ (x_1, x_2) \mapsto x_2 \\
V_4 &:= \{(x_1, x_2) : x_1^2 + x_2^2 = 1,\ x_2 < 0\}, \qquad \varphi_4 : V_4 \to \mathbb{R},\ (x_1, x_2) \mapsto x_1.
\end{aligned}$$

These are clearly diffeomorphisms onto their ranges. Choose

$$\begin{aligned}
P = &\{(-\tfrac{\pi}{3} + \omega, \tfrac{\pi}{3} - \omega), (\tfrac{\pi}{6} + \omega, \tfrac{5\pi}{6} - \omega), (\tfrac{2\pi}{3} + \omega, \tfrac{4\pi}{3} - \omega), (\tfrac{7\pi}{6} + \omega, \tfrac{11\pi}{6} - \omega)\} \\
&:= \{P_1, P_2, P_3, P_4\}.
\end{aligned}$$

Here $\omega \in (0, \frac{\pi}{12})$ is a small angle. The covers P_i clearly satisfy the first two conditions laid out in Lemma 3. Equip $\mathbb{S}^1$ with the standard metric induced by the inclusion $\mathbb{S}^1 \hookrightarrow \mathbb{R}^2$. We can verify the desired estimates in Lemma 3 as follows: It is clear that any arc of $\mathbb{S}^1$ with length less than $\frac{\pi}{6}$ is contained in one of $P_1, \ldots, P_4$. If we choose $\delta \in (0, \frac{\pi}{36})$, then for any $x \in \mathbb{S}^1$, $B(x, 3\delta) \subset P_i$ for some i. Suppose

$$\begin{aligned}
y &= (y_1, y_2) = (\cos\theta, \sin\theta), \\
z &= (z_1, z_2) = (\cos\tilde{\theta}, \sin\tilde{\theta}),
\end{aligned}$$

$$\theta, \tilde{\theta} \in (-\tfrac{\pi}{3} + \omega, \tfrac{\pi}{3} - \omega),$$
$$y_2, z_2 \in (\sin(-\tfrac{\pi}{3} + \omega), \sin(\tfrac{\pi}{3} - \omega)).$$

The φ_1-coordinates of y, z are y_2, z_2, respectively. Hence,

$$|y - z| = |y_2 - z_2| = |\sin\theta - \sin\tilde{\theta}|$$
$$r(y, z) = |\theta - \tilde{\theta}|.$$

By the mean value theorem:

$$|y - z| = |\sin\theta - \sin\tilde{\theta}| = |\cos\xi|_{\xi\in(-\frac{\pi}{3}+\omega, \frac{\pi}{3}-\omega)}|\theta - \tilde{\theta}| \leq |\theta - \tilde{\theta}| = r(y, z).$$

and

$$\begin{aligned} r(y, z) &= |\theta - \tilde{\theta}| \\ &= |\arcsin y_2 - \arcsin z_2| \\ &= \left|\frac{1}{\sqrt{1-\eta^2}}\right|_{\eta\in(\sin(-\frac{\pi}{3}+\omega), \sin(\frac{\pi}{3}-\omega))} |y_2 - z_2| \\ &\leq \frac{1}{\sqrt{1 - \sin^2\left(\frac{\pi}{3} - \omega\right)}}|y_2 - z_2| \\ &= \frac{1}{\sqrt{1 - \sin^2\left(\frac{\pi}{3} - \omega\right)}}|y - z|. \end{aligned}$$

This suggests the choice $c_1 = 1$ and $c_2 = \frac{1}{\sqrt{1-\sin^2(\frac{\pi}{3}-\omega)}}$. We have

$$c_1 c_2 = \frac{1}{\sqrt{1 - \sin^2\left(\frac{\pi}{3} - \omega\right)}} < \frac{1}{\sqrt{1 - \sin^2\left(\frac{\pi}{3}\right)}} = 2.$$

The analysis for $y, z \in P_3$ is similar, only with the difference that $\theta, \tilde{\theta} \in (\frac{2\pi}{3} + \omega, \frac{4\pi}{3} - \omega)$ and we have $y_2, z_2 \in (\sin(\frac{4\pi}{3} - \omega), \sin(\frac{2\pi}{3} + \omega)) = (\sin(-\frac{\pi}{3} + \omega), \sin(\frac{\pi}{3} - \omega))$.

Next, consider $y, z \in P_2$. Suppose

$$y = (y_1, y_2) = (\cos\theta, \sin\theta),$$
$$z = (z_1, z_2) = (\cos\tilde{\theta}, \sin\tilde{\theta}),$$
$$\theta, \tilde{\theta} \in (\tfrac{\pi}{6} + \omega, \tfrac{5\pi}{6} - \omega),$$
$$y_1, z_1 \in (\cos(\tfrac{5\pi}{6} - \omega), \cos(\tfrac{\pi}{6} + \omega)).$$

The φ_2-coordinates of y, z are y_1, z_1, respectively. Hence,

$$|y - z| = |y_1 - z_1| = |\cos\theta - \cos\tilde{\theta}|$$
$$r(y, z) = |\theta - \tilde{\theta}|.$$

By the mean value theorem:

$$|y - z| = |\cos\theta - \cos\tilde{\theta}| = |\sin\xi|_{\xi\in(\frac{\pi}{6}+\omega, \frac{5\pi}{6}-\omega)}|\theta - \tilde{\theta}| \le |\theta - \tilde{\theta}| = r(y, z).$$

and

$$\begin{aligned} r(y, z) &= |\theta - \tilde{\theta}| \\ &= |\arccos y_1 - \arccos z_1| \\ &= \left|\frac{1}{\sqrt{1-\eta^2}}\right|_{\eta\in(\cos(\frac{5\pi}{6}-\omega),\cos(\frac{\pi}{6}+\omega))} |y_1 - z_1| \\ &\le \frac{1}{\sqrt{1-\cos^2(\frac{\pi}{6}+\omega)}}|y_1 - z_1| \\ &= \frac{1}{\sqrt{1-\cos^2(\frac{\pi}{6}+\omega)}}|y - z|. \end{aligned}$$

This suggests the choice $c_1 = 1$ and $c_2 = \frac{1}{\sqrt{1-\cos^2(\frac{\pi}{6}+\omega)}}$. We have

$$c_1c_2 = \frac{1}{\sqrt{1-\cos^2(\frac{\pi}{6}+\omega)}} < \frac{1}{\sqrt{1-\cos^2\left(\frac{\pi}{6}\right)}} = 2.$$

Note that the choice agrees with the case $y, z \in P_1$. The analysis for $y, z \in P_4$ is similar as well, which proves the desired claim.

4 Generalizing nonwindowed geometric scattering

Now that we have developed the machinery necessary for the rest of the paper, we prove the q-nonwindowed scattering transforms are bounded operators with respect to (19) and outline basic properties of the representation.

4.1 The 2-nonwindowed geometric scattering norm

We start by proving that 2-nonwindowed scattering transforms are bounded operators with respect to (19).

Theorem 9 *Suppose that G satisfies the conditions of Theorem 1 and $\{\psi_j\}_{j\in\mathbb{Z}}$ be a set of spectral filters generated by using G in Eq. 5. Then we have*

$$\|\overline{S}_2^m f - \overline{S}_2^m g\|_{\ell^2(\mathbb{Z}^m)}^2 \le C^{2m}\|f-g\|_{\mathbf{L}^2(\mathcal{M})}^2$$

for all $f, g \in \mathbf{L}^2(\mathcal{M})$.

Proof In the case of $m = 1$, we see that

$$\begin{aligned}\sum_{j\in\mathbb{Z}} |\overline{S}_2 f[j] - \overline{S}_2 g[j]|^2 &= \sum_{j\in\mathbb{Z}} |\|f*\psi_j\|_2 - |g*\psi_j\|_2|^2 \\ &\le \sum_{j\in\mathbb{Z}} \|f*\psi_j - g*\psi_j\|_2^2 \\ &= \sum_{j\in\mathbb{Z}} \|(f-g)*\psi_j\|_2^2 \\ &\le C^2\|f-g\|_{\mathbf{L}^2(\mathcal{M})}^2.\end{aligned}$$

Now assume that we have

$$\|\overline{S}_2^n f - \overline{S}_2^n g\|_{\ell^2(\mathbb{Z}^n)}^2 \le C^{2n}\|f-g\|_{\mathbf{L}^2(\mathcal{M})}^2$$

for some $k \ge 1$. For the $m = k+1$ case, we see that

$$\begin{aligned}&\sum_{(j_1,\dots,j_{k+1})\in\mathbb{Z}^{k+1}} \left|\overline{S}_2^{k+1} f[j_1,\dots,j_{k+1}] - \overline{S}_2^{k+1} g[j_1,\dots,j_{k+1}]\right|^2 \\ &= \sum_{(j_1,\dots,j_{k+1})\in\mathbb{Z}^{k+1}} \left|\|U[j_1,\dots,j_k]f*\psi_{j+1}\|_2 - \|U[j_1,\dots,j_n]g*\psi_{j+1}\|_2\right|^2 \\ &\le \sum_{(j_1,\dots,j_{k+1})\in\mathbb{Z}^{k+1}} \|(U[j_1,\dots,j_k]f - U[j_1,\dots,j_k]g)*\psi_{j+1}\|_2^2 \\ &\le C^2 \sum_{(j_1,\dots,j_k)\in\mathbb{Z}^k} \|U[j_1,\dots,j_k]f - U[j_1,\dots,j_k]g\|_2^2 \\ &= C^2\|\overline{S}_2^k f - \overline{S}_2^k g\|_{\ell^2(\mathbb{Z}^k)}^2\end{aligned}$$

Now apply the induction hypothesis to get

$$C^2\|\overline{S}_2^k f - \overline{S}_2^k g\|_{\ell^2(\mathbb{Z}^k)}^2 \le C^{2(k+1)}\|f-g\|_{\mathbf{L}^2(\mathcal{M})}^2.$$

Thus, the claim is proven. □

Corollary 10 *Suppose that G satisfies the conditions of Theorem 1 and let $\{\psi_j\}_{j\in\mathbb{Z}}$ be a set of wavelets generated by using G in Eq. 5. Then we have*

$$\|\overline{S}_2^m f\|_{\ell^2(\mathbb{Z}^m)}^2 = C^{2m}\|f\|_{\mathbf{L}^2(\mathcal{M})}^2$$

for all $f \in \mathbf{L}^2(\mathcal{M})$ *and all* $m \geq 1$.

For proper invariance, we provide a theorem that demonstrates that the 2-nonwindowed geometric scattering transform is invariant to isometries.

Theorem 11 *Let* $\xi \in Isom(\mathcal{M}, \mathcal{M}')$*, and let* $f \in \mathbf{L}^2(\mathcal{M})$*. Define* $f' = V_\xi f$ *and let* $\left(\overline{S}_2^m\right)'$ *be the corresponding* 2*-nonwindowed geometric scattering transform on* $\mathcal{M}'$ *produced by a littlewood paley wavelet satisfying the conditions described in Theorem 1. We have* $\left(\overline{S}_2^m\right)' f' = \overline{S}_2^m f$.

Proof We see that $\overline{S}_2[\emptyset]f = \|f\|_2 = \|V_\xi f\|_2$ since V_ξ is an isometry. Now suppose that we consider $p = (j_1, \ldots, j_m)$. Then since convolution using a spectral filter commutes with isometries and modulus operators (see Theorem 2.1 in [31]),

$$\begin{aligned}\overline{S}_2^m[j_1, \ldots, j_m]f &= \|U[p]f\|_{\mathbf{L}^2(\mathcal{M})}\\ &= \|V_\xi U[p]f\|_{\mathbf{L}^2(\mathcal{M})}\\ &= \|U[p]V_\xi f\|_{\mathbf{L}^2(\mathcal{M})}\\ &= \|U[p]f'\|_{\mathbf{L}^2(\mathcal{M})}\\ &= \left(\overline{S}_2^m\right)'[j_1, \ldots, j_m]f'.\end{aligned}$$

Thus, we can see that each layer is isometry invariant. □

4.2 The *q*-nonwindowed geometric scattering norm

Now we prove the q-nonwindowed Geometric Scattering Transforms, for $q \in (1, 2)$, are bounded operators with respect to (19) under mild assumptions.

Theorem 12 *Suppose that we choose wavelets* $\{\psi_j\}_{j\in\mathbb{Z}}$ *generated by using* $G \in \mathcal{S}(\mathbb{R}^+)$ *in Eq. 5, G satisfies the conditions of Theorem 1, and* $c_1c_2 < 2$*. Then*

$$\|\overline{S}_q^m f - \overline{S}_q^m g\|_{\ell^2(\mathbb{Z}^m)}^q \leq C_q^m \|f - g\|_{\mathbf{L}^q(\mathcal{M})}^q$$

for all $f, g \in \mathbf{L}^q(\mathcal{M})$*, for all* $m \geq 1$*, and some constant* C_q *dependent on* q.

Proof We start by providing a proof for the case of $m = 1$:

$$\begin{aligned}\|\overline{S}_q f - \overline{S}_q g\|_{\ell^2(\mathbb{Z}^m)}^q &= \left(\sum_{j\in\mathbb{Z}} |\overline{S}_q[j]f - \overline{S}_q[j]g|^2\right)^{q/2}\\ &= \left(\sum_{j\in\mathbb{Z}} |\|U[j]f\|_q - \|U[j]g\|_q|^2\right)^{q/2}\end{aligned}$$

$$\leq \left(\sum_{j \in \mathbb{Z}} \|U[j]f - U[j]g\|_q^2 \right)^{q/2}$$
$$= \left(\sum_{j \in \mathbb{Z}} \left(\int_{\mathcal{M}} |U[j]f(x) - U[j]g(x)|^q \, d\mu(x) \right)^{2/q} \right)^{q/2}$$

Via Minkowski's Integral Inequality,

$$\left(\sum_{j \in \mathbb{Z}} \left(\int_{\mathcal{M}} |U[j]f(x) - U[j]g(x)|^q \, d\mu(x) \right)^{2/q} \right)^{q/2}$$
$$\leq \int_{\mathcal{M}} \left(\sum_{j \in \mathbb{Z}} |U[j]f(x) - U[j]g(x)|^2 \right)^{q/2} d\mu(x)$$
$$\leq \int_{\mathcal{M}} \left(\sum_{j \in \mathbb{Z}} |(f * \psi_j)(x) - (g * \psi_j)(x)|^2 \right)^{q/2} d\mu(x)$$
$$= \|W(f - g)\|^q_{\mathbf{L}^q_{\ell^2(\mathbb{Z})}(\mathcal{M})}.$$

Now apply Corollary 8 to get

$$\|W(f - g)\|^q_{\mathbf{L}^q_{\ell^2(\mathbb{Z})}(\mathcal{M})} \leq C_q \|f - g\|^q_{\mathbf{L}^q(\mathcal{M})}.$$

Now assume that for some $m \geq 1$, we have

$$\|\overline{S}_q^m f - \overline{S}_q^m g\|^q_{\ell^2(\mathbb{Z}^m)} \leq C_q^m \|f - g\|^q_{\mathbf{L}^q(\mathcal{M})}.$$

Similar to above, when we consider the case with $m + 1$, we can mimic the steps above to get

$$\|\overline{S}_q^{m+1} f - \overline{S}_q^{m+1} g\|^q_{\ell^2(\mathbb{Z}^m)}$$
$$= \left(\sum_{j_{m+1} \in \mathbb{Z}} \cdots \sum_{j_1 \in \mathbb{Z}} |\overline{S}_q^{m+1}[j_1, \ldots, j_{m+1}]f - \overline{S}_q^{m+1}[j_1, \ldots, j_{m+1}]g|^2 \right)^{q/2}$$
$$= \left(\sum_{(j_1, \ldots, j_m) \in \mathbb{Z}^m} \cdots \sum_{j_{m+1} \in \mathbb{Z}} \left(\int_{\mathcal{M}} |U[j_1, \ldots, j_{m+1}]f(x) \right. \right.$$
$$\left. \left. - U[j_1, \ldots, j_{m+1}]g(x)|^q \, d\mu(x) \right)^{2/q} \right)^{q/2}$$

$$= \left(\sum_{(j_1,\ldots,j_m)\in\mathbb{Z}^m} \left(\sum_{j_{m+1}\in\mathbb{Z}} \left(\int_{\mathcal{M}} |U[j_1,\ldots,j_{m+1}]f(x) \right.\right.\right.$$

$$\left.\left.\left. -U[j_1,\ldots,j_{m+1}]g(x)|^q \, d\mu(x) \right)^{2/q} \right)^{\frac{q}{2}\cdot\frac{2}{q}} \right)^{q/2}$$

$$\leq \left(\sum_{(j_1,\ldots,j_m)\in\mathbb{Z}^m} \left(\int_{\mathcal{M}} \left(\sum_{j_{m+1}\in\mathbb{Z}} |U[j_1,\ldots,j_{m+1}]f(x) \right.\right.\right.$$

$$\left.\left.\left. -U[j_1,\ldots,j_{m+1}]g(x)|^2 \right)^{q/2} d\mu(x) \right)^{2/q} \right)^{q/2}$$

$$\leq \left(\sum_{(j_1,\ldots,j_m)\in\mathbb{Z}^m} \|W(U[j_1,\ldots,j_m]f - U[j_1,\ldots,j_m]g)\|^2_{\mathbf{L}^q_{\ell^2(\mathbb{Z})}(\mathcal{M})} \right)^{q/2}$$

$$= C_q \left(\sum_{(j_1,\ldots,j_m)\in\mathbb{Z}^m} \|U[j_1,\ldots,j_m]f - U[j_1,\ldots,j_m]g\|^2_{\mathbf{L}^q(\mathcal{M})} \right)^{q/2}.$$

Now we see that we can apply the induction hypothesis to get

$$\left(\sum_{(j_1,\ldots,j_m)\in\mathbb{Z}^m} \|U[j_1,\ldots,j_m]f - U[j_1,\ldots,j_m]g\|^2_{\mathbf{L}^q_{\ell^2(\mathbb{Z})}(\mathcal{M})} \right)^{q/2}$$
$$= \|\overline{S}_q^m[j_1,\ldots,j_m]f - \overline{S}_q^m[j_1,\ldots,j_m]g\|^q_{\ell^2(\mathbb{Z}^m)}$$
$$\leq C_q^m \|f-g\|^q_{\mathbf{L}^q(\mathcal{M})}.$$

□

Corollary 13 *Suppose that we choose wavelets* $\{\psi_j\}_{j\in\mathbb{Z}}$ *generated by using* $G \in \mathcal{S}(\mathbb{R}^+)$ *in Eq. 5, G satisfies the conditions of Theorem 1, and* $c_1c_2 < 2$. *Then*

$$\|\overline{S}_q^m f\|^q_{\ell^2(\mathbb{Z}^m)} \leq C_q^m \|f\|^q_{\mathbf{L}^q(\mathcal{M})}$$

for all $f \in \mathbf{L}^q(\mathcal{M})$, *for all* $m \geq 1$, *and some constant* C_q *dependent on* q.

For the next theorem, we omit the proof since it is identical to the case when $q = 2$, but we state it for completeness.

Theorem 14 *Let* $\xi \in Isom(\mathcal{M}, \mathcal{M}')$, *and let* $f \in \mathbf{L}^q(\mathcal{M})$. *Define* $f' = V_\xi f$ *and let* $\left(\overline{S}_q^m\right)'$ *be the corresponding* q*-nonwindowed geometric scattering transform on* $\mathcal{M}'$

produced by wavelets $\{\psi_j\}_{j\in\mathbb{Z}}$ *using* $G \in \mathcal{S}(\mathbb{R}^+)$ *in Eq. 5, G satisfies the conditions of Theorem 1, and* $c_1c_2 < 2$ *in Lemma 3. We have* $\left(\overline{S}_q^m\right)' f' = \overline{S}_q^m f$.

5 Diffeomorphism stability

In machine learning tasks, it is often necessary for a representation to have some degree of invariance with respect to the action of a group. For tasks involving manifolds, one may like to have local isometry invariance. More formally, let $V_\xi f(x) = f(\xi^{-1}x)$ for $\xi \in \mathrm{Isom}(\mathcal{M})$ and consider a representation $\Phi : \mathcal{B}_1 \to \mathcal{B}_2$, where $\mathcal{B}_1, \mathcal{B}_2$ are Banach Spaces. It would be desirable to have a representation such that

$$\|\Phi f - \Phi V_\xi f\|_{\mathcal{B}_2} \leq 2^{-nJ} \|\xi\|_\infty \|f\|_{\mathcal{B}_1},$$

where J controls the degree of invariance. A simple choice is to use an averaging filter, but this potentially leads to the loss of information that can be crucial for the task.

For other tasks, such as manifold classification, a fully rigid representation may be required, and full isometry invariance is desirable. That is to say, we have

$$\|\Phi f\|_{\mathcal{B}_2} = \|\Phi V_\xi f\|_{\mathcal{B}_2}.$$

In addition to invariance, it is necessary for a representation to also have stability properties. Instead of considering an isometry, consider $\xi \in \mathrm{Diff}(\mathcal{M})$ and think of $V_\xi f$ as a small deformation of f. We want

$$\|\Phi f - \Phi V_\xi f\|_{\mathcal{B}_2} \leq K(\xi, J)\|f\|_{\mathcal{B}_1},$$

where $K(\xi, J) \to 0$ as $\|\xi\|_\infty \to 0$ for fixed J and $K(\xi, J) \to 0$ as $J \to \infty$ for fixed ξ; this is to ensure that small deformations of an input do not lead to large changes in the representation.

With the above discussion in mind, we provide diffeomorphism stability results for a generalization of bandlimited functions, λ-bandlimited functions, which are defined as functions which satisfy $\hat{f}(k) = \langle f, \phi_k\rangle = 0$ whenever $\lambda_k \geq \lambda$.

Lemma 15 ([31]) *Suppose* $\xi \in \mathit{Diff}(\mathcal{M})$. *If* $f \in \mathbf{L}^2(\mathcal{M})$ *is* λ*-bandlimited, then*

$$\|f - V_\xi f\|_{\mathbf{L}^2(\mathcal{M})} \leq C(\mathcal{M})\lambda^n \|\xi\|_\infty \|f\|_{\mathbf{L}^2(\mathcal{M})}$$

for some constant $C(\mathcal{M})$.

Theorem 16 *Suppose* $\xi \in \mathit{Diff}(\mathcal{M})$. *Let* $f \in \mathbf{L}^2(\mathcal{M})$, *and assume that* ψ *is a wavelet family satisfying the conditions of Theorem 1. Then*

$$\|\overline{S}_2^m f - \overline{S}_2^m V_\xi f\|_{\ell^2(\mathbb{Z}^m)} \leq C(\mathcal{M})\lambda^n \|\xi\|_\infty \|f\|_{\mathbf{L}^2(\mathcal{M})}.$$

Proof We apply Theorem 9, so Lemma 15 gives the desired result. □

5.1 Stability results for the q-nonwindowed geometric scattering norm

Lemma 17 *Suppose $\xi \in Diff(\mathcal{M})$ and $1 < q < 2$. If $f \in \mathbf{L}^q(\mathcal{M})$ is λ-bandlimited, then*

$$\|f - V_\xi f\|_{\mathbf{L}^q(\mathcal{M})} \leq C_q(\mathcal{M})\lambda^n \|\xi\|_\infty \|f\|_{\mathbf{L}^q(\mathcal{M})}$$

for some constant $C_q(\mathcal{M})$.

Proof Since f is λ-bandlimited, $f \in \mathbf{L}^2(\mathcal{M})$ as well, and the proof is nearly identical to the proof of the case when $q = 2$, but we provide the steps for completeness. We define π_λ be the operator that projects a function $f \in \mathbf{L}^2(\mathcal{M})$ onto the eigenspace E_λ and define the projection operator

$$P_\lambda := \sum_{\lambda_n \leq \lambda} \pi_{\lambda_n}$$

with kernel

$$K^{(\lambda)}(x, y) = \sum_{\lambda_n \leq \lambda} e_n(x)\overline{e_n(y)}.$$

We have $P_\lambda f = f$ μ-almost-everywhere. Thus, via Holder's inequality,

$$\begin{aligned}
|f(x) - V_\xi f(x)| &= |P_\lambda f(x) - V_\xi P_\lambda f(x)| \\
&= \left| \int_{\mathcal{M}} K^{(\lambda)}(x, y) f(y)\, dy - \int_{\mathcal{M}} K^{(\lambda)}(\xi^{-1}(x), y) f(y)\, dy \right| \\
&\leq \left| \int_{\mathcal{M}} (K^{(\lambda)}(x, y) - K^{(\lambda)}(\xi^{-1}(x), y)) f(y)\, dy \right| \\
&\leq \|f\|_{\mathbf{L}^q(\mathcal{M})} \left(\int_{\mathcal{M}} |K^{(\lambda)}(x, y) - K^{(\lambda)}(\xi^{-1}(x), y)|^p\, dy \right)^{1/p} \\
&\leq C_{q,\mathrm{Vol}(\mathcal{M})} \|f\|_{\mathbf{L}^q(\mathcal{M})} \|\xi\|_\infty \|\nabla K^{(\lambda)}\|_\infty
\end{aligned}$$

for some constant $C_{q,\mathrm{Vol}(\mathcal{M})}$ dependent on q and the volume of the manifold. Here, p is the conjugate exponent of q in the sense that $\frac{1}{p} + \frac{1}{q} = 1$. Now, by Lemma H.1 in [31], we have

$$\|\nabla K^{(\lambda)}\|_\infty \leq C(\mathcal{M})\lambda^n.$$

for some constant $C(\mathcal{M})$. Thus, the proof is complete. □

Theorem 18 *Suppose $\xi \in Diff(\mathcal{M})$. Let $f \in \mathbf{L}^q(\mathcal{M})$ be λ-bandlimited. Additionally, suppose that we choose wavelets $\{\psi_j\}_{j \in \mathbb{Z}}$ generated by using $G \in \mathcal{S}(\mathbb{R}^+)$ in Eq. 5, G satisfies the conditions of Theorem 1, and $c_1 c_2 < 2$. Then*

$$\|\overline{S}_q^m f - \overline{S}_q^m V_\xi f\|_{\ell^2(\mathbb{Z}^m)} \leq C(\mathcal{M})\lambda^n \|\xi\|_\infty \|f\|_{\mathbf{L}^q(\mathcal{M})}$$

for some constant $C(\mathcal{M})$.

Proof We apply Theorem 12 to get

$$\|\overline{S}_q^m f - \overline{S}_q^m V_\xi f\|_{\ell^2(\mathbb{Z}^m)} \le C_q \|f - V_\xi f\|_{\mathbf{L}^q(\mathcal{M})}.$$

By Lemma 17, we have

$$\|f - V_\xi f\|_{\mathbf{L}^q(\mathcal{M})} \le C(\mathcal{M})\lambda^n \|\xi\|_\infty \|f\|_{\mathbf{L}^q(\mathcal{M})},$$

which gives the desired result. □

6 Conclusions and future work

We have provided a framework for understanding nonwindowed scattering coefficients. In particular, we provide a weighted measure for distortion between nonwindowed scattering coefficients, showed our weighted measure is well-defined mapping for $\mathbf{L}^q(\mathcal{M})$ functions, and showed that nonwindowed scattering coefficients are stable to diffeomorphisms for λ-bandlimited functions. For future work, it is of interest to see if it is possible to extend our results to manifolds that are not restricted the conditions present in Sects. 4 and 5. Additionally, what are other manifolds that satisfy the conditions present in sections 4 and 5?

Acknowledgements We would like to thank Michael Perlmutter for providing feedback, which improved the quality of this initial draft. Additionally, we would like to thank the anonymous reviewer for their comments, which further improved the quality of this paper.

Funding The research of Y. Yang is partially supported by the NSF grant DMS-2006881, DMS-2237534, DMS-2220373, the NIH grant R03-EB033521 and the startup fund from Michigan State University.

Declarations

Conflict of interest The authors have no Conflict of interest.

References

1. Allys, E., Levrier, F., Zhang, S., Colling, C., Regaldo-Saint Blancard, B., Boulanger, F., Hennebelle, P., Mallat, S.: The rwst, a comprehensive statistical description of the non-gaussian structures in the ism. Astron. Astrophys. **629**, A115 (2019)
2. Andén, J., Lostanlen, V., Mallat, S.: Joint time-frequency scattering. IEEE Trans. Signal Process. **67**(14), 3704–3718 (2019)
3. Andén, J., Mallat, S.: Multiscale scattering for audio classification. In ISMIR, pages 657–662. Miami, Florida, (2011)
4. Andén, J., Mallat, S.: Deep scattering spectrum. IEEE Trans. Signal Process. **62**(16), 4114–4128 (2014)
5. Bronstein, M.M., Bruna, J., LeCun, Y., Szlam, A., Vandergheynst, P.: Geometric deep learning: going beyond euclidean data. IEEE Signal Process. Mag. **34**(4), 18–42 (2017)
6. Bruna, J., Mallat, S.: Invariant scattering convolution networks. IEEE Trans. Pattern Anal. Mach. Intell. **35**(8), 1872–1886 (2013)

7. Bruna, J., Mallat, S., Bacry, E., Muzy, J.F.R.: Intermittent process analysis with scattering moments. Ann. Stat. **43**(1), 323–351 (2015)
8. Bruna, J., Mallat, S.: Audio texture synthesis with scattering moments, (2013)
9. Chew, J., Hirn, M., Krishnaswamy, S., Needell, D., Perlmutter, M., Steach, H., Viswanath, S., Wu, H.-T.: Geometric scattering on measure spaces. arXiv preprint arXiv:2208.08561, (2022)
10. Chew, J., Steach, H., Viswanath, S., Wu, H.-T., Hirn, M., Needell, D., Vesely, M.D., Krishnaswamy, S., Perlmutter, M.: The manifold scattering transform for high-dimensional point cloud data. In Topological, Algebraic and Geometric Learning Workshops 2022, pages 67–78. PMLR, (2022)
11. Chua, A., Hirn, M., Little, A.: On generalizations of the nonwindowed scattering transform. Appl. Comput. Harmon. Anal. **68**, 101597 (2024)
12. Coifman, R.R., Lafon, S.: Diffusion maps. Appl. Comput. Harmon. Anal. **21**(1), 5–30 (2006)
13. Coifman, R.R., Maggioni, M.: Diffusion wavelets. Appl. Comput. Harmon. Anal. **21**(1), 53–94 (2006)
14. Coifman, R.R., Weiss, G.: Analyse harmonique non-commutative sur certains espaces homogènes : étude de certaines intégrales singulières. (1971)
15. Czaja, W., Li, W.: Analysis of time-frequency scattering transforms. Appl. Comput. Harmon. Anal. **47**(1), 149–171 (2019)
16. Czaja, W., Li, W.: Rotationally invariant time-frequency scattering transforms. J. Fourier Anal. Appl. **26**, 1–23 (2020)
17. Eickenberg, M., Exarchakis, G., Hirn, M., Mallat, S., Thiry, L.: Solid harmonic wavelet scattering for predictions of molecule properties. J. Chem. Phys. **148**(24), 241732 (2018)
18. Gama, F., Ribeiro, A., Bruna, J.: Diffusion scattering transforms on graphs. In International Conference on Learning Representations, (2019)
19. Gama, F., Ribeiro, A., Bruna, J.: Stability of graph scattering transforms. Adv. Neural Inf. Process. Syst. **32**, 25 (2019)
20. Gao, F., Wolf, G., Hirn, M.: Geometric scattering for graph data analysis. In International Conference on Machine Learning, pages 2122–2131. PMLR, (2019)
21. García-Cuerva, J., De Francia, J.L.R.: Weighted norm inequalities and related topics. Elsevier, Amsterdam (1985)
22. Geller, D., Mayeli, A.: Continuous wavelets on compact manifolds. Math. Z. **262**(4), 895–927 (2009)
23. Geller, D., Pesenson, I.Z.: Kolmogorov and linear widths of balls in sobolev spaces on compact manifolds. Math. Scand. **115**(1), 96–122 (2014)
24. Hammond, D.K., Vandergheynst, P., Gribonval, R.: Wavelets on graphs via spectral graph theory. Appl. Comput. Harmon. Anal. **30**(2), 129–150 (2011)
25. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 770–778, (2016)
26. Hirn, M., Mallat, S., Poilvert, N.: Wavelet scattering regression of quantum chemical energies. Multiscale Model. Simul. **15**(2), 827–863 (2017)
27. Koller, M., Großmann, J., Monich, U., Boche, H.: Deformation stability of deep convolutional neural networks on Sobolev spaces. In 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pages 6872–6876. IEEE, (2018)
28. Mallat, S.: Group invariant scattering. Commun. Pure Appl. Math. **65**(10), 1331–1398 (2012)
29. Oyallon, E., Belilovsky, E., Zagoruyko, S.: Scaling the scattering transform: Deep hybrid networks. In Proceedings of the IEEE international conference on computer vision, pages 5618–5627, (2017)
30. Oyallon, E., Mallat, S.: Deep roto-translation scattering for object classification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 2865–2873, (2015)
31. Perlmutter, M., Gao, F., Wolf, G., Hirn, M.: Geometric wavelet scattering networks on compact Riemannian manifolds. In Mathematical and Scientific Machine Learning, pages 570–604. PMLR, (2020)
32. Perlmutter, M., Tong, A., Gao, F., Wolf, G., Hirn, M.: Understanding graph neural networks with generalized geometric scattering transforms. SIAM J. Math. Data Sci. **5**(4), 873–898 (2023)
33. Saito, N., Schonsheck, S.C., Shvarts, E.: Multiscale hodge scattering networks for data analysis. arXiv preprint arXiv:2311.10270, (2023)
34. Saito, N., Schonsheck, S.C., Shvarts, E.: Multiscale transforms for signals on simplicial complexes. Sampl. Theory Signal Process. Data Anal. **22**(1), 2 (2024)
35. Sifre, L., Mallat, S.: Combined scattering for rotation invariant texture analysis. ESANN **44**, 68–81 (2012)

36. Sifre, L., Mallat, S.: Rotation, scaling and deformation invariant scattering for texture discrimination. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 1233–1240, (2013)
37. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. In International Conference on Learning Representations, (2015)
38. Sinz, P., Swift, M.W., Brumwell, X., Liu, J., Kim, K.J., Qi, Y., Hirn, M.: Wavelet scattering networks for atomistic systems with extrapolation of material properties. J. Chem. Phys. **153**(8), 084109 (2020)
39. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 1–9, (2015)
40. Tenenbaum, J.B., de Silva, V., Langford, J.C.: A global geometric framework for nonlinear dimensionality reduction. Science **290**(5500), 2319–2323 (2000)
41. Van der Maaten, L., Hinton, G.: Visualizing data using t-sne. J. Mach. Learn. Res. **9**, 11 (2008)
42. Wiatowski, T., Bölcskei, H.: A mathematical theory of deep convolutional neural networks for feature extraction. IEEE Trans. Inf. Theory **64**(3), 1845–1866 (2017)
43. Wiatowski, T., Grohs, P., Bölcskei, H.: Energy propagation in deep convolutional neural networks. IEEE Trans. Inf. Theory **64**(7), 4819–4842 (2017)
44. Zou, D., Lerman, G.: Graph convolutional neural networks via scattering. Appl. Comput. Harmon. Anal. **49**(3), 1046–1074 (2020)